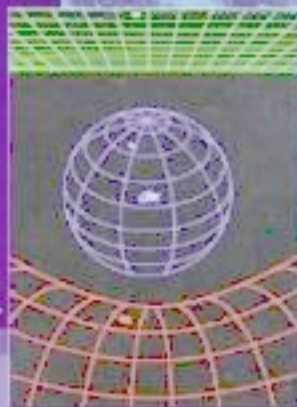


# Networks and Telecommunications

Design and Operation

Second Edition



Martin P. Clark

---

## *Networks and Telecommunications*

---



---

# ***Networks and Telecommunications***

## ***Design and Operation***

***Second Edition***

---

**M. P. Clark**

*Telecommunications Consultant  
Germany*

**JOHN WILEY & SONS**  
Chichester · New York · Weinheim · Brisbane · Singapore · Toronto



Copyright © 1991, 1997 by John Wiley & Sons Ltd  
Baffins Lane, Chichester  
West Sussex, PO19 1UD, England  
National 01243 779777  
International (+44) 1243 779777  
e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk  
Visit our Home Page on <http://www.wiley.co.uk> or <http://www.wiley.com>

Reprinted March 1998

All Rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London, W1P 9HE, UK, without the permission in writing of the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of this publication.

Neither the author nor John Wiley & Sons Ltd accept any responsibility or liability for loss or damage occasioned to any person or property through using the material, instructions, methods or ideas contained herein, or acting or refraining from acting as a result of such use. The author and Publisher expressly disclaim all implied warranties, including merchantability of fitness for any particular purpose. There will be no duty on the authors or Publisher to correct any errors or defects in the software.

Designations used by companies to distinguish their products are often claimed as trademarks. In all instances where John Wiley & Sons is aware of a claim, the product names appear in initial capital or capital letters. Readers, however, should contact the appropriate companies for more complete information regarding trademarks and registration.

#### *Other Wiley Editorial Offices*

John Wiley & Sons, Inc., 605 Third Avenue,  
New York, NY 10158-0012, USA

Wiley-VCH Verlag GmbH  
Pappelallee 3, D-69469 Weinheim, Germany

Jacaranda Wiley Ltd, 33 Park Road, Milton,  
Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #20-01,  
Jin Xing Distripark, Singapore 129809

John Wiley & Sons (Canada) Ltd, 22 Worcester Road  
Rexdale, Ontario, M9W 1L1, Canada

#### *Library of Congress Cataloging-in-Publication Data*

Clark, Martin P.

Networks and telecommunications: design and operation/Martin P. Clark. — 2nd ed.  
p. cm.

Includes bibliographical references and index.

ISBN 0-471-97346-7

1. Telecommunication systems. 2. Data transmission systems.

3. Computer networks I. Title.

TK5102.5.C53 1997

621.382 — dc21

97-9248  
CIP

#### *British Library Cataloguing in Publication Data*

A catalogue record for this book is available from the British Library

ISBN 0 47197346 7

Typeset in 10/12 pt Times by Aarontype Ltd., Easton, Bristol

Printed and bound in Great Britain by Bookcraft (Bath) Ltd

This book is printed on acid-free paper responsibly manufactured from sustainable forestry for which at least two trees are planted for each one used for paper production

---

# *Contents*

---

## **Summary**

Part 1	Fundamentals of telecommunications networks	1
Part 2	Modern telephone networks	211
Part 3	Modern data networks	339
Part 4	Multimedia networks	437
Part 5	Running a network	475
Part 6	Setting up networks	741
Part 7	Specific Businesses and networks	777

<b>Preface</b>	<b>xxi</b>
----------------	------------

<b>About the Author</b>	<b>xiii</b>
-------------------------	-------------

<b>Acknowledgements</b>	<b>xv</b>
-------------------------	-----------

## **Part 1 FUNDAMENTALS OF TELECOM NETWORKS 1**

<b>1</b>	<b>Information and its Conveyance</b>	<b>3</b>
1.1	Types of Information	6
1.2	Telecommunications Systems	6
1.3	A Basic Telecommunications System	7
1.4	Common Types of Telecommunications Systems	8
1.5	Networks	9
1.6	Connection-oriented Transport Service (COTS) and Connectionless Network Service (CLNS)	11
1.7	Circuit-, Packet- and Cell-switched Networks	12
1.8	Considerations for Network Planners	14
1.9	Technical Standards for Telecommunications Systems	15

<b>2</b>	<b>Introduction to Signal Transmission and the Basic Line Circuit</b>	<b>17</b>
2.1	Analogue and Digital Transmission	17
2.2	Telegraphy	19
2.3	Telephony	21
2.4	Received Signal Strength, Sidetone and Echo	23
2.5	Automatic Systems: Central Battery and Exchange Calling	24
2.6	Real Communications Networks	27
<b>3</b>	<b>Long-haul Communication</b>	<b>29</b>
3.1	Attenuation and Repeaters	29
3.2	Line Loading	31
3.3	Amplification	32
3.4	Two-and Four-wire Circuits	35
3.5	Equalization	36
3.6	Frequency Division Multiplexing (FDM)	37
3.7	Crosstalk and Attenuation on FDM Circuits	41
<b>4</b>	<b>Data and the Binary Code System</b>	<b>43</b>
4.1	The Binary Code	43
4.2	Electrical Representation and Storage of Binary Code Numbers	44
4.3	Using the Binary Code to Represent Textual Information	45
4.4	Morse Code	46
4.5	Baudot Code (Alphabet IA2)	46
4.6	ASCII	47
4.7	EBCDIC	49
4.8	Use of the Binary Code to Convey Graphical Images	49
4.9	Facsimile	49
4.10	Digital Transmission	52
<b>5</b>	<b>Digital Transmission and Pulse Code Modulation</b>	<b>55</b>
5.1	Digital Transmission	55
5.2	Pulse Code Modulation	57
5.3	Quantization	60
5.4	Quantization Noise	61
5.5	Time Division Multiplexing	61
5.6	Higher Bit Rates of Digital Line Systems	64
5.7	Digital Frame Formatting and 'Justification'	65
5.8	Interworking the 2 Mbit/s and 1.5 Mbit/s Hierarchies	60
5.9	Synchronous Frame Formatting	70
5.10	Line Coding	71
5.11	Other Line Codes and their Limitations	74
<b>6</b>	<b>The Principles of Switching</b>	<b>77</b>
6.1	Circuit-switched Exchanges	77
6.2	Call Blocking within the Switch Matrix	82
6.3	Full and Limited Availability	83
6.4	Fan-in–Fan-out Switch Architecture	86

6.5	Switch Hardware Types	88
6.6	Strowger Switching	88
6.7	Crossbar Switching	96
6.8	Reed Relay Switching	100
6.9	Digital Switching	101
6.10	Packet and Cell Switches	106
<b>7</b>	<b>Setting up and Clearing Connections</b>	<b>109</b>
7.1	Alerting the Called Customer	109
7.2	Automatic Networks	110
7.3	Set Up	110
7.4	Number Translation	115
7.5	Unsuccessful Calls	117
7.6	Inter-exchange and International Signalling	118
7.7	The R2 Signalling System	121
7.8	R2 Line Signalling	122
7.9	Compelled or Acknowledged Signalling	126
7.10	R2 Inter-register, Multi-frequency Code (MFC) Signalling	127
7.11	Digital Line Systems and Channel-associated Signalling	131
7.12	Signalling Interworking	132
7.13	Advanced Signalling Applications	133
7.14	Signalling Sequence Diagrams	133
7.15	Call Set-up and Information Transfer in Data Networks	136
7.16	Network Interfaces: UNI, NNI, INI, ICI, SNI	137
7.17	Information Transfer in Connectionless Networks	139
<b>8</b>	<b>Transmission Systems</b>	<b>141</b>
8.1	Audio Circuits	141
8.2	Standard Twisted Pair Cable Types for Indoor Use	143
8.3	Transverse Screen and Coaxial Cable Transmission	143
8.4	Frequency Division Multiplexing (FDM)	145
8.5	HDSL (High Bit-rate Digital Subscriber Line) and ADSL (Asymmetric Digital Subscriber Line)	148
8.6	Optical Fibres	148
8.7	Radio	153
8.8	Radio Wave Propagation	155
8.9	Radio Antennas	157
8.10	Surface-wave Radio Systems	161
8.11	High Frequency (HF) Radio	161
8.12	Very High Frequency (VHF) and Ultra High Frequency (UHF) Radio	161
8.13	Microwave Radio	163
8.14	Tropospheric Scatter	166
8.15	Satellite Systems	167
8.16	'Multiple Access' Radio and Satellite Systems	172
8.17	Electromagnetic Interference (EMI) and Electromagnetic Compatibility (EMC)	175

<b>9</b>	<b>Data Network Principles and Protocols</b>	<b>177</b>
9.1	Computer Networks	177
9.2	Basic Data Conveyance: Introducing the DTE and the DCE	178
9.3	Modulation of Digital Information over Analogue Lines Using a Modem	180
9.4	High Bit Rate Modems	181
9.5	Modem 'Constellations'	182
9.6	Computer-to-network Interfaces	186
9.7	Synchronization	189
9.8	Bit Synchronization	190
9.9	Character Synchronization: Synchronous and Asynchronous Data Transfer	191
9.10	Handshaking	192
9.11	Protocols for Transfer of Data	193
9.12	The Open Systems Interconnection Model	194
9.13	Data Message Format	199
9.14	Implementation of Layered Protocol Networks	201
9.15	The Use of Null Layers	204
9.16	Other Layered Protocols	204
9.17	Data Network Types	205
9.18	Point-to-point Data Networks	205
9.19	Circuit-switched Data Networks	206
9.20	Packet-switched Data Networks	207
9.21	Practical Computer Networks	209
<b>Part 2</b>	<b>MODERN TELEPHONE NETWORKS</b>	<b>211</b>
<b>10</b>	<b>Integrated Services Digital Network (ISDN)</b>	<b>213</b>
10.1	The Concept of ISDN	213
10.2	Bearer, Supplementary and Teleservices	214
10.3	ISDN Interfaces and End-user Applications	215
10.4	Basic Rate Interface (BRI)	216
10.5	The S/T Interface Specification	216
10.6	Use of the Basic Rate Interface	219
10.7	ISDN Terminals	221
10.8	Primary Rate Interface	222
10.9	The Public Network and ISDN	224
10.10	Deployment of ISDN	225
10.11	The Marketing of ISDN and the Early User Benefits	226
10.12	Network Interworking	227
10.13	Companies' Private ISDNs (Corporate ISDN)	227
10.14	Broadband Services over ISDN	229
<b>11</b>	<b>Intelligent Networks and Services</b>	<b>231</b>
11.1	The Concept of Intelligent Networks	231
11.2	Intelligent Network Architecture	232

11.3	The Service Control Point (SCP)	233
11.4	The Service Switching Point (SSP)	234
11.5	The Service Management System (SMS) and Service Creation Environment (SCE)	234
11.6	Benefits of Intelligent Networks	235
11.7	Intelligent Network (IN) Services	235
11.8	Calling Card	236
11.9	Freephone Service (or 800 Service)	237
11.10	900 Service	239
11.11	Centrex Service and Virtual Private Network	239
11.12	Line Information Database (LIBD)	241
11.13	Televoting	243
11.14	Cellular Radio Telephone Service	244
11.15	Network Intelligence and PBXs	245
11.16	Voicemail and Voice Response Systems	246
11.17	Considerations Before Introducing IN to a Network	248
11.18	The Future of Intelligent Networks	248
<b>12</b>	<b>Signalling System No. 7</b>	<b>249</b>
12.1	SS7 Signalling between Exchanges	249
12.2	SS7 Signalling Networks	251
12.3	The Structure of SS7 Signalling	253
12.4	The Message Transfer Part (MTP)	254
12.5	The User Parts of SS7	256
12.6	The Telephone User Part (TUP)	257
12.7	The Data User Part (DUP)	258
12.8	The Integrated Services User Parts (ISUP)	258
12.9	The Enhanced Telephone User Part (TUP+)	258
12.10	The Signalling Connection Control Part (SCCP)	258
12.11	Transaction Capabilities (TC)	261
12.12	The Mobile Application Part (MAP)	263
12.13	Operation and Maintenance Application Part (OMAP)	263
12.14	Intelligent Network Application Part (INAP)	264
12.15	The Use and Evolution of CCITT7 Signalling	264
12.16	Signalling Network Planning and Testing	265
12.17	Interconnection of SS7 Networks	266
<b>13</b>	<b>Synchronous Digital Hierarchy (SDH) and Synchronous Optical Network (SONET)</b>	<b>267</b>
13.1	History of the Synchronous Digital Hierarchy (SDH)	267
13.2	The Problems of PDH Transmission	267
13.3	The Multiplexing Structure of SDH	270
13.4	The Tributaries of SDH	273
13.5	Path Overhead	276
13.6	Section Overhead (SOH)	276
13.7	Network Topology of SDH Networks	277
13.8	Optical Interfaces for SDH	278

13.9	Management of SDH Networks	278
13.10	SONET (Synchronous Optical Network)	279
13.11	SDH and ATM (Asynchronous Transfer Mode)	280
<b>14</b>	<b>Operator Assistance and Manual Services</b>	<b>281</b>
14.1	Manual Network Operation	281
14.2	Semi-automatic Telephony	282
14.3	Calling the Operator	287
14.4	Operator Privileges	288
14.5	Typical Assistance Services	289
14.6	Cooperation between International Operators: Code 11 and Code 12 Services	291
14.7	A Modern Operator Switchroom	293
14.8	Operator Assistance on Telex Networks	294
14.9	Operator Assistance on Data Networks	294
<b>15</b>	<b>Mobile Telephone Networks</b>	<b>297</b>
15.1	Radio Telephone Service	297
15.2	Cellular Radio	299
15.3	Making Cellular Radio Calls	303
15.4	Tracing Cellular Radio Handsets	304
15.5	Early Cellular Radio Networks	305
15.6	Global System for Mobile Communications (GSM)	307
15.7	GSM Technology	308
15.8	Personal Communications Network (PCN) and DCS-1800	311
15.9	Aeronautical and Maritime Mobile Communications Services	313
15.10	Iridium, Globalstar and the Evolution Towards the Universal Mobile Telephone Service (UMTS)	314
<b>16</b>	<b>Cordless Telephony and Radio in the Local Loop (RILL)</b>	<b>319</b>
16.1	The Drive for Radio in the Local Loop	319
16.2	Fixed Networks Based on Radio Technology	320
16.3	Cordless Telephones	321
16.4	Telepoint or Cordless Telephone 2 (CT2)	322
16.5	DECT (Digital European Cordless Telephony)	323
16.6	DECT Handover	325
16.7	The Radio Relay Station Concept in DECT	325
16.8	The DECT Air Interface (D3-interface)	326
16.9	Other ISDN Wireless Local loop Systems	328
16.10	Shorthaul Point-to-multipoint (PMP) Microwave Radio	328
<b>17</b>	<b>Fibre in the Loop (FITL) and Other Access Networks</b>	<b>329</b>
17.1	Fibre Access Networks	329
17.2	Fibre to the Building (FTTB)	329
17.3	Fibre to the Curb (FTTC)	330
17.4	Fibre to the Home (FTTH)	331
17.5	Broadband Passive Optical Network	331

17.6	Access Network Interfaces	332
17.7	ETSI V5 Interfaces	333
17.8	V5.2 Interface	335
17.9	V5.1 Interface	336
17.10	Significance of the V5.x Interfaces	336
17.11	Re-use of Existing Copper Access Networks	337
17.12	HDSL (High Bitrate Digital Subscriber Line)	337
17.13	ADSL (Asymmetric Digital Subscriber Line)	337
17.14	Hybrid Fibre/Coax (HFC) Networks	338

### **Part 3 MODERN DATA NETWORKS 339**

#### **18 Packet Switching 341**

18.1	Packet Switching Basics	341
18.2	Transmission Delay in Packet-switched Networks	343
18.3	Routing in Packet-switched Networks	344
18.4	ITU-T Recommendation X.25	346
18.5	The Technical Details of X.25	348
18.6	X.25 Link Access Procedure (LAP and LAPB)	348
18.7	X.25 Packet Level Interface (Layer 3 Protocol)	350
18.8	Typical Parameter Default Settings Used in X.25 Networks	354
18.9	Packet Assembler/Disassemblers (PADs)	355
18.10	ITU-T Recommendation X.75	358
18.11	When X.25 Packet Switching May and May Not Be Used	360
18.12	Alternatives to X.25-based Packet Switching	361
18.13	IBM's 'Systems Network Architecture'	361
18.14	APPN (Advanced Peer-to-peer Networking)	366

#### **19 Local Area Networks (LANs) 367**

19.1	The Emergence of LANs	367
19.2	LAN Topologies and Standards	367
19.3	CSMA/CD (IEEE 802.3, ISO 8802.3): Ethernet	369
19.4	Token Bus (IEEE 802.4, ISO 8802.4)	371
19.5	Token Ring (IEEE 802.5)	372
19.6	Logical Link Control for LANs	374
19.7	LAN Operating Software and LAN Servers	374
19.8	Interconnection of LANs: Bridges, Routers and Gateways	375

#### **20 Frame Relay 379**

20.1	The Throughput Limitations of X.25 Packet Switching	379
20.2	The Need for Faster Response Data Networks	381
20.3	The Emergence and use of Frame Relay	383
20.4	Frame Relay UNI	383
20.5	Frame Relay SVC Service	384
20.6	Congestion Control in Frame Relay networks	384



20.7	Frame Relay NNI	386
20.8	Frame Format	386
20.9	Address Field Format	387
20.10	ITU-T Recommendations Pertinent to Frame Relay	388
20.11	FRAD (Frame Relay Access Device)	388
<b>21</b>	<b>Campus and Metropolitan Area Networks (MANs)</b>	<b>391</b>
21.1	Fibre Distributed Data Interface	391
21.2	Switched Multimegabit Digital Services (SMDS)	394
21.3	The Demise of MANs	398
<b>22</b>	<b>Electronic Mail, Internet and Electronic Message Services</b>	<b>399</b>
22.1	Videotext	399
22.2	Electronic Mail (e-mail)	400
22.3	Addressing Schemes for Electronic Mail	402
22.4	The Advantages and Disadvantages of e-mail	403
22.5	EDI: Corporate Communication with Customers and Suppliers via e-mail	403
22.6	Internet	404
22.7	TCP/IP Protocol Stack	405
22.8	Common Applications Using TCP/IP	407
22.9	The Internet Protocol	408
22.10	The Internet Control Message Protocol (ICMP)	410
22.11	Transmission Control Protocol (TCP)	410
22.12	Online Database Services	410
<b>23</b>	<b>The Message Handling System (MHS)</b>	<b>413</b>
23.1	The Need for MHS	413
23.2	The Concept of MHS	413
23.3	The MHS Model	414
23.4	Layered Representation of MHS	417
23.5	The Structure of MHS Messages and MHS Addresses	419
23.6	MHS Management Domains	420
23.7	MHS and the OSI Directory Service	421
23.8	Message Conversion and Conveyance Using MHS	421
23.9	Setting Up a Message Handling System	422
23.10	File Transfer Access and Management (FTAM)	424
23.11	Summary	424
<b>24</b>	<b>Mobile and Radio Data Networks</b>	<b>425</b>
24.1	Radiopaging	425
24.2	Mobile Data Networks	429
24.3	TETRA (Trans-European Trunked Radio System)	431
24.4	Wireless LANs	432
24.5	Radiodetermination Satellite Services (RDSS) and the Global Positioning System (GPS)	436

**Part 4 MULTIMEDIA NETWORKS 437****25 Broadband, Multimedia Networks and the B-ISDN 439**

25.1	Multimedia Applications: the Driver for Broadband Networks	439
25.2	Video Communication	441
25.3	The Emergence of the B-ISDN	441
25.4	The Services to be Offered by B-ISDN	442
25.5	The Emergence of the ATM Switching Technique as the Heart of ATM	443
25.6	Connection Types Supported by B-ISDN	446
25.7	User Device Connection to B-ISDN	446
25.8	Evolution to Broadband-ISDN	449

**26 Asynchronous Transfer Mode (ATM) 451**

26.1	A Flexible Transmission Medium	451
26.2	Statistical Multiplexing and the Evolution of Cell Relay Switching	452
26.3	The Problems to be Solved by Cell Relay	453
26.4	The Technique of Cell Relay	454
26.5	The ATM Cell Header	455
26.6	The Components of an ATM Network	456
26.7	The ATM Adaption Layer (AAL)	458
26.8	ATM Virtual Channels and Virtual Paths	458
26.9	User, Control and Management Planes	459
26.10	How is a Virtual Channel Connection (VCC) Set Up?	460
26.11	Signalling Virtual Channels and Meta-signalling Virtual Channels	461
26.12	Virtual Channel Identifiers (VCIs) and Virtual Path Identifiers (VPIs)	462
26.13	Information Content and Format of the ATM Cell Header	464
26.14	ATM Protocol Layers	465
26.15	The ATM Transport Network	465
26.16	Capability of the ATM Adaption Layer (AAL)	467
26.17	Protocol Stack when Communicating via an ATM Transport Network	468
26.18	ATM Protocol Reference Model (PRM)	469
26.19	ATM Forum Network Reference Model	471
26.20	ATM Forum Network Management Model	472

**Part 5 RUNNING A NETWORK 475****27 Telecommunications Management Network (TMN) 477**

27.1	The Problems of Managing Networks	477
27.2	Network Provisioning	479
27.3	Umbrella Network Management Systems	480
27.4	The Q3-interface, the Common Management Information Protocol (CIMP) and the Concept of Managed Objects (MO)	483
27.5	The ISO Management Model	485
27.6	TMN Management Function Model	486
27.7	The Network Management Forum (NMF), OMNIpoint and SPIRIT	487

27.8	Realization of a TMN	487
27.9	Example of Early TMN Realization	488
27.10	Simple Network Management Protocol (SNMP)	489
27.11	Summary of TMN Benefits	489
27.12	Telecommunications Intelligent Network Architecture (TINA)	490
<b>28</b>	<b>Network Routing, Interconnection and Interworking</b>	<b>491</b>
28.1	The Need for a Network Routing Plan	491
28.2	Network Routing Objectives and Constraints	494
28.3	The Administration of Routing Tables	497
28.4	Routing Protocols Used in Modern Networks	499
28.5	Network Topology State and the 'Hello State Machine'	500
28.6	Signalling Impact upon Routing and Call Set-up Delays	503
28.7	Plausibility Check During Number Analysis	504
28.8	Network Interconnection	504
28.9	Network Interconnection Services	505
28.10	Interconnect	506
28.11	Equal Access	506
28.12	Number Portability	507
28.13	Shared Use of Access Network Ducts and Cables	507
28.14	Pitfalls of Interconnection	508
28.15	The Point of Interconnection and Collocation	508
28.16	The Interconnection Contract	509
28.17	Interworking	510
<b>29</b>	<b>Network Numbering and Addressing Plans</b>	<b>513</b>
29.1	The International Telephone Numbering Plan	513
29.2	International Public Data Network Address Scheme	520
29.3	Escape Codes	521
29.4	Telex Network Numbering Plan (ITU-T F.69)	524
29.5	X.500: The Addressing Plan for the Message Handling Service (MHS)	524
29.6	Internet Addressing Scheme	525
29.7	Internet e-mail (STMP) Addresses	526
29.8	Network Addressing Schemes Used in Support of Broadband-ISDN and ATM	527
<b>30</b>	<b>Teletraffic Theory</b>	<b>529</b>
30.1	Telecommunications Traffic	529
30.2	Traffic Intensity (Circuit-switched Networks)	530
30.3	Practical Traffic Intensity (Erlang) Measurement	531
30.4	The Busy Hour	533
30.5	The Formula for Traffic Intensity	535
30.6	The Traffic-carrying Capacity of a Single Circuit	536
30.7	Dimensioning Circuit-switched Networks	539
30.8	Example Route Dimensioning	542
30.9	Call Waiting Systems	543
30.10	Dimensioning Data Networks	546

30.11	Pollaczek–Khinchine Delay Formula	550
30.12	Practical Dimensioning of Networks	551
30.13	Appendix: The Derivation of Erlang’s Formula	551
<b>31</b>	<b>Traffic Monitoring and Forecasting</b>	<b>555</b>
31.1	Measuring Network Usage	555
31.2	Usage Monitoring in Circuit-Switched Networks	556
31.3	Traffic Intensity	556
31.4	Total Usage Monitoring	558
31.5	Number of Calls Attempted	560
31.6	Number of Calls Completed	561
31.7	Monitoring Usage of Data Networks	562
31.8	Forecasting Models for Predicting Future Network Use	564
31.9	Fitting the Forecasting Model	567
31.10	Other Forecasting Models	569
<b>32</b>	<b>Network Traffic Control</b>	<b>571</b>
32.1	Networks	571
32.2	Sizing Circuit-switched Networks	572
32.3	Hierarchical Network	573
32.4	Overflow of ‘Automatic Alternative Routing’ (AAR)	577
32.5	Wilkinson–Rapp Equivalent Random Method	579
32.6	Dimensioning ‘Final Routes’	581
32.7	Trunk Reservation	581
32.8	‘Crankback’ or ‘Automatic Re-routing’ (ARR)	584
32.9	Proportionate Bidding Facility (PBF)	585
32.10	Dynamic Routing	585
32.11	Routing and Traffic Control in Data Networks	586
32.12	Network Design	588
32.13	Appendix: The Wilkinson–Rapp Route Dimensioning Method	590
<b>33</b>	<b>Practical Network Transmission Planning</b>	<b>593</b>
33.1	Network Transmission Plan	593
33.2	Send and Receive Reference Equivalents	594
33.3	Connection Reference Points and Overall Reference Equivalent	596
33.4	Measuring Network Loss	598
33.5	Correcting Signal Strength	599
33.6	The Control of Sidetone	602
33.7	The Problem of Echo	602
33.8	Echo Control and Circuit Instability	603
33.9	Signal (or ‘Propagation’) Delay	606
33.10	Noise and Crosstalk	607
33.11	Signal Distortion	608
33.12	Transmission Plan for Digital and ‘Data’ Networks	609
33.13	International Transmission Plan	612
33.14	Private Network Transmission Plan	613
33.15	Circuit and Transmission System Line-up	613

33.16	Network Resource Management	613
33.17	Circuit Provisioning Planning	615
33.18	New Cable Planning	616
33.19	Local Line Planning	618
33.20	Trunk and International Line Planning	623
33.21	Radio Transmission Systems	623
33.22	Satellite Transmission Management	628
<b>34</b>	<b>Quality of Service (QOS) and Network Performance (NP)</b>	<b>633</b>
34.1	Framework for Performance Management	633
34.2	Quality: A Marketing View	635
34.3	Quality of Service (QOS) and Network Performance (NP)	636
34.4	Quality of Service Parameters	640
34.5	Generic Network Performance Parameters	640
34.6	Performance Monitoring Functions of Modern Networks	642
34.7	Network Performance Planning and Measurement	642
34.8	A Few Practical Tips	646
34.9	Summary	646
<b>35</b>	<b>Charging and Accounting for Network Use</b>	<b>647</b>
35.1	Recompense for Network Use	647
35.2	Customer Subscription Charges	648
35.3	Customer Usage Charges	648
35.4	Pulse Metering	649
35.5	Electronic Ticketing	652
35.6	Accounting	652
35.7	Route Destination Accounting	654
35.8	Charging and Accounting on Data Networks	655
35.9	Charging and Accounting for Manual (Operator) Assistance	656
35.10	Charging and Accounting for Leased Circuits	657
35.11	Charging Payphone Calls	657
35.12	Customer Billing	657
35.13	Setting Customer Charges and Accounting Rates	658
35.14	Network Costs and How to Recharge Them	660
35.15	Future Accounting and Charging Practices	662
<b>36</b>	<b>Maintaining the Network</b>	<b>663</b>
36.1	The Objectives of General Maintenance	663
36.2	Maintenance Philosophy	663
36.3	Maintenance Organization	665
36.4	Centralized Operation and Maintenance	666
36.5	Lining Up Analogue and Mixed Analogue/Digital Circuits	667
36.6	High Grade Data Circuit Line-up	671
36.7	Lining Up Digital Circuits	673
36.8	Performance Objectives	674
36.9	Maintenance 'Access Points'	675
36.10	Localizing Network Faults	676

36.11	Hardware Faults	679
36.12	Software Faults	679
36.13	Change Control Procedure for Hardware and Software	680
<b>37</b>	<b>Containing Network Overload</b>	<b>683</b>
37.1	The Effect of Congestion	683
37.2	Network Monitoring	684
37.3	Network Management Controls	687
37.4	Expansive Control Actions	688
37.5	Restrictive Control Actions	691
37.6	Network Management Systems	694
<b>38</b>	<b>Network Economy Measures</b>	<b>695</b>
38.1	Cost Minimization	695
38.2	Frequency Division Multiplexing (FDM)	696
38.3	Time Division Multiplexing (FDM)	698
38.4	Wavelength Division Multiplexing	699
38.5	Circuit Multiplication Equipment (CME)	699
38.6	Speech Interpolation and Statistical Multiplexing	699
38.7	Analogue Bandwidth Compression and Low Rate Encoding of PCM	703
38.8	Data Multiplexors	706
38.9	Data Compression	707
38.10	Practical Uses of CME	707
38.11	Constraints on the Use of CME	709
<b>39</b>	<b>Network Security Measures</b>	<b>711</b>
39.1	The Trade-off between Confidentiality and Interconnectivity	711
39.2	Different Types of Protection	712
39.3	Encryption	713
39.4	Network Access Control	713
39.5	Path Protection	714
39.6	Destination Access Control	715
39.7	Specific Technical Risks	716
39.8	Carelessness	717
39.9	Call Records	718
39.10	Mimicked Identity	718
39.11	Radio Transmission, LANs and Other Broadcast-type Media	718
39.12	EMI (Electro-Magnetic Interference)	719
39.13	Message Switching Networks	719
39.14	Other Types of Network Abuse	720
<b>40</b>	<b>Technical Standards for Networks</b>	<b>723</b>
40.1	The Need for Standards	723
40.2	Worldwide International Standards Organizations	724
40.3	Regional and National Standards Organizations	727
40.4	Regulatory Standards Organizations	732

40.5	Other Standards-promoting 'Fora'	734
40.6	Proprietary Standards	736
40.7	The Structure and Content of ITU-T Recommendations	738

## **Part 6 SETTING UP NETWORKS 741**

### **41 Building, Extending and Replacing Networks 743**

41.1	Matching Network Capacity to Forecast Demand	743
41.2	Other Factors Affecting the Need for New Exchanges	748
41.3	Factors in Determining an Exchange Provision Programme	749
41.4	Determining a Strategy for Network Evolution	750
41.5	Comparison of Strategy Options	754
41.6	Exchange Design and Specification	754
41.7	Outline Circuit-switched Design: Circuit Numbers and Traffic Balance	756
41.8	Outline Design of Other Types of Network	759
41.9	The Effect of Low Circuit Infill on Exchange and Lineplant Planning	759
41.10	Functional Requirements of Exchanges or Line Systems	760
41.11	Methods of Network or Exchange Modernization	760

### **42 Selecting and Procuring Equipment 763**

42.1	Tendering for Equipment	000
42.2	Project Management	000
42.3	Procurement Policy	000
42.4	Planning Documentation	000
42.5	The Tender Document	000
42.6	Summary	000

## **Part 7 SPECIFIC BUSINESSES AND NETWORKS 777**

### **43 Meeting Business Needs and Creating Competitive Edge 779**

43.1	Content of an IT Strategy	779
43.2	The Study of Information Flows	780
43.3	The Tactical Development Plan	786
43.4	Business Applications of IT	787
43.5	Summary	792

### **44 Network Regulation and Deregulation 793**

44.1	Reasons for Deregulation	793
44.2	The Dilemma of Deregulation	795
44.3	Optional Methods of Regulation	797
44.4	Types of Regulatory Bodies	797
44.5	Designation of 'Customer Premises Equipment' (CPE)	798
44.6	Deregulation of Value-added Services	798
44.7	Competition in Basic Services	799
44.8	The Instruments of PTO Regulation	800

44.9	European Telecommunications Deregulation	802
44.10	Instruments of United Kingdom Regulation	806
44.11	United States Telecommunications Regulation	809
44.12	Other Countries	813
<b>45</b>	<b>Corporate Networks</b>	<b>815</b>
45.1	Telecommunications Management	815
45.2	Premises Cabling Schemes	816
45.3	Office Computer Networking	819
45.4	Private Networks	820
45.5	Architecture of Private Networks	822
45.6	Planning Private Networks	824
45.7	A Word of Warning	828
45.8	PTO Leased Circuit Offerings	828
45.9	Making Use of Mobile Radio Technology	829
<b>46</b>	<b>Public Networks and Telecommunications Service Providers</b>	<b>831</b>
46.1	Company 'Mission'	831
46.2	Identifying and Addressing the PTO's market	832
46.3	PTO Product Development	834
46.4	PTO Business Development	836
	<b>Bibliography</b>	<b>839</b>
	<b>Glossary of Terms</b>	<b>849</b>
	<b>Glossary of Abbreviations</b>	<b>865</b>
	<b>ISO two-letter country code abbreviations (ISO 3166)</b>	<b>890</b>
	<b>Index</b>	<b>895</b>





---

# Preface

---

The second edition of this text is a much larger work than the first edition, having needed to be expanded with extensive new sections on data, broadband and multimedia networks and much extended coverage of modern transmission media, radio and network management. The technology may have moved on since 1990, but the original preface is as relevant today as it was then...

The networks that I have personally had to develop and operate have bridged not only the classical subdivisions of *voice* and *data*, but also national boundaries, technological standards and regulations. As a result, I have found it necessary to become familiar with a broad range of technology, telecommunications practices and worldwide regulations. I have needed to know how things work and interwork, and the practical problems to be overcome. I have needed to understand the basic electrical engineering detail underlying telecommunications, but experience has also given me two equally valuable resources: an armoury of practical techniques and a knowledge of both technical and regulatory constraints.

In my time I have used plenty of books, each addressing one of the many individual technologies that I have had to deal with. However, to date I have not discovered a book aimed at technical managers who need a broad range of knowledge about how to develop and operate a number of different types of network in a pragmatic way. This seems perverse, since many of today's professional *telecommunications managers* are expected to know about all the various types of network making up their trade. It is also perverse given the fact that, in my experience, network problems cannot be neatly categorized by technology. I have rarely needed to know all the technical details of a particular technology and cannot remember having been able to use any network to its full capability. Seldom have I had the opportunity, because seldom have I had a problem where one technology alone will suffice, and never have I had the option anyway to scrap the established network with its constraints and start entirely afresh.

Surprisingly perhaps, my major problems have arisen not from technical incompatibility of different international networks; such problems are soluble with time and money. Rather the difficulty has been in gaining a basic understanding of how different types of network operate, and therefore how they could be made to co-exist, interwork or cooperate to serve a given user's need. It is from this background that I decided to write this book, a day-to-day encyclopaedia for practical folk like me, people faced with

*voice, data* and a whole compendium of other networks to be responsible for. Here in 46 chapters we have a picture of an evolving jungle of techniques and administrative controls, all about telecommunications to-day and how they got that way.

The book provides an insight for all telecommunications-affected parties: end users, engineers, private companies, public telecommunications operators, regulators and governments. The sort of people that I hope will read it and use it, the interested parties, are students and academics, telecommunications and computer practitioners, and technical managers. Through our greater common awareness and understanding, I hope that we in turn can help extend familiarity with telecommunications and their effective use to a much wider audience: business managers, directors, and consultants; specialist press and publicity people; chairmen and board members of companies, politicians and everyday folk.

How to read a book of this kind with advantage depends on what one hopes to get out of it: the student may plough through it from cover to cover, while the academic toys fastidiously with the bibliography, glossary and index. The chairman only reads what his advisers put in front of him. However, it will take more than that if we are to harness telecommunications for their full potential; expanding systems of linked computers, international and mobile communications are turning business and society upside down, whether we like it or not.

In all quarters more know-how is needed, more shared knowledge, more mutual awareness of what is actually going on, not only here and now under our noses, but in other minds and in other countries as well. To meet this need, this book sets out at length the raw facts and the jargon; and although the picture presented is strictly contemporary, I hope the book will give its readers the suggestion and confidence to carry things forward into the future among themselves. This, in its way, is the most hopeful and important part of the book.

I have chosen not to remove any technical coverage from the first edition. You, the reader, might be confronted by an older technology and need to deal with it. Beyond this, many old ideas find new lease of life in 'modern technologies'

For those using the book to unravel an outside world full of jargon and opportunity, the bibliography, glossaries and index will be invaluable, but the acronyms abound, beyond the possibilities of a single book. A number of excellent dictionaries will help the reader who finds himself at a loss.

MARTIN CLARK  
Eppstein, Germany  
30th November 1996

---

## *About the Author*

---

*Martin Clark* spent eight years in the international network planning department of British Telecom, where he gained knowledge and experience of the fundamentals and practicality of telecommunications. He also developed a good understanding of the technical, regulatory and other problems associated with international networks. In 1989 he became Group Telecommunications Manager with Grand Metropolitan, where his experience gave him an insight into global *corporate network* management and the business opportunities offered by telecommunications. In 1991 he moved to Germany, where he has worked on setting up a pan-European managed data network for Cable & Wireless, in managing the Deutsche Bank's *corporate network* and also helped Mannesmann in the establishment of a new public telecommunications carrier (ARCOR) to rival Deutsche Telekom. Recently he has been active with the wireless ATM company, Netro, and he also works as a freelance consultant.

Martin Clark is a Chartered European Engineer and a member of the IEE.



---

# Acknowledgements

---

When reviewing the list of sources, reference bodies and helpers who have contributed to this work, I come to think of myself more as a coordinator than a sole author, and my special thanks are due to the many who have encouraged me, contributed material or reviewed sections of text.

The material that appears in the chapters on PCM, network traffic control, network overload and ISDN is derived from articles first published in the *African Technical Review and African Review*. Two of these articles (on ISDN and PCM) were co-authored by Felix Redmill, my colleague at British Telecom, and it is to him that I owe the initial idea and the encouragement to pursue this work at all.

British Telecom supplied many of the photographic illustrations, and also an education in telecommunications spanning eight years. My particular thanks are due to John Kelly of the customer communications photographic library and David Hay, the BT Group Archivist. The pictures represent only a glimpse of the rich history of telecommunications, and I hope they will encourage others to use the BT Archives.

The American Telephone and Telegraph Corporation (AT&T) also provided a large number of photographs, as well as making available some of their world-respected Bell Laboratories staff to review portions of the text. I am sincerely grateful to Tim Granger of the National Accounts staff for arranging these contributions.

No book on telecommunications as they are today could fail to recognize the contribution to world standards made by the International Telecommunications Union and its subordinate entities ITU-R and ITU-T. You will find references littered throughout the text. Particular copyright extracts (chosen by the author, but reproduced with the prior authorization of the ITU) are labelled accordingly. The full texts of all ITU copyright material may be obtained from the ITU Sales and Marketing Service, Place des Nations, CH-1211 Geneva 20, Switzerland, Telephone: +41 22 730 6141 or Internet: Sales@itu.int.

The cartoons that bring humour and parody to the chapters on quality and technical standards are by Patrick Wright of Fernhurst, Sussex, England, but the copyright is that of the author.

In this second edition I am also indebted for photographs to the international satellite organisation INTELSAT, to Siemens AG and to Netro Corporation. Further

thanks are also due to ATM Forum for their kind permission to reproduce the ATM network reference model and network management model from their standards.

Let me thank all those personal friends who helped especially with the first edition – without it there would have been no second edition! Those few I have the space to mention are: Pettrel Adams, who toiled unfailingly with the typing; Elizabeth Bathurst, the editing assistant at John Wiley; Richard Cawdell (formerly of the European Commission); Colin Clark, my father and critic; Ann-Marie Halligan, John Wiley's editor; John Goatly, father-in-law, critic and adviser; John Green of GrandMet, for his advice on IBM networks; Martin Pigott, former CCITT rapporteur for number seven signalling; Peter Walker, formerly of British Telecom network planning department and now with Oftel; and Arthur Watson, my first telecom boss and part-time golfing friend.

Special thanks go to my close family, for enduring the months of aspiration and toil.

**Martin Clark**

PART 1

---

**FUNDAMENTALS OF  
TELECOMMUNICATIONS  
NETWORKS**

---





---

# *Information and its Conveyance*

---

The world about us brims with information. All the time our ears, eyes, fingers, mouths and noses sense the environment around us, continually increasing our 'awareness', 'intelligence' and 'instructive knowledge'. Indeed these last two phrases are at the heart of the Oxford Dictionary's definition of the word information. Communication, on the other hand, is defined as 'the imparting, conveyance or exchange of ideas, knowledge or information'. It might be done by word, image, instruction, motion, smell – or maybe just a wink! Telecommunication is communication by electrical, radio or optical (e.g. laser) means. We introduce the basic capabilities and terminology of telecommunications and networking in this chapter.

As hybrid words go, 'telecommunications' wins no prizes, but it is all we have to work with. The greek 'tele' prefix means distant, and nothing else; communication, in the sense of information passed to and fro between human beings, and animals, is an activity that goes back beyond recorded times. With a broad view, long distance communications brings to mind Armada beacons, heliographs flashing between Frontier posts, empires held together by relays of post-houses, whales singing to one another in the deep, and the family dog which conducts its social life by laying and following scent trails. For the narrower purposes of this book telecommunications is going to mean the transfer of information by electromagnetic means, (and with this will go a certain amount of accepted jargon). All systems have much in common, whatever their age. In principle each requires a transmitter, a carrying device or transmission medium, a receiver, and a supply of information which will be equally comprehensible at both ends. For lessons in technique nothing should be disregarded, however ancient: for a cheap, speedy and comprehensive message, what is there to beat a human wink?

In the science and business of telecommunications, a structured framework has been created for conveying certain types of information across long distances, with little respect for the barriers of geography. In this book we study this framework; first we understand how the forces of electricity, light and radiowaves may be tamed to provide a basis for such communication; then we focus on the pragmatic operation of networks and the quest for solutions to the business needs of information flow.

Figure 1.1 illustrates a simple but powerful model for understanding and categorizing various different means of communication. The model illustrates a number of different ways in which a business may communicate either within itself, or with its external environment of suppliers and customers. Thus we introduce the concept of an 'information environment', across which information flows in one of a number of different forms.

The simplest form of information flow (illustrated by Figure 1.1) might be directly from one person to another, by word of mouth or by a visual signal. Alternatively the information could have been conveyed on paper or electrically. The advantage of either of the latter two methods is that the information in paper or electronic form may also be readily stored for future reference.

In this book we shall use the model of Figure 1.1 twice. Here we use it to illustrate how different methods of communication may be categorized into one of the three broad types, and as a basis for explaining the prerequisite components of a telecommunications system. In Chapter 43 it is used to illustrate the analysis, simplification and planning of business information flows. This double theme runs throughout the book: understanding telecommunications technology, and explaining its exploitation in a pragmatic business-oriented manner. Well-known examples of communications methods that fall into the three categories of paper, person-to-person and electronic are given in Figure 1.2.

Some types of communication are hybrids of the three basic methods. Modern facsimile machines, for example, are capable of relaying images of paper documents over the telephone network and recreating them at a distant location. This would appear as quite a complex information path on our model of Figure 1.1, as Figure 1.3 shows.

First a person must record the relevant information on paper (shown as (a) on Figure 1.3), then he must feed it into the facsimile machine which converts it into

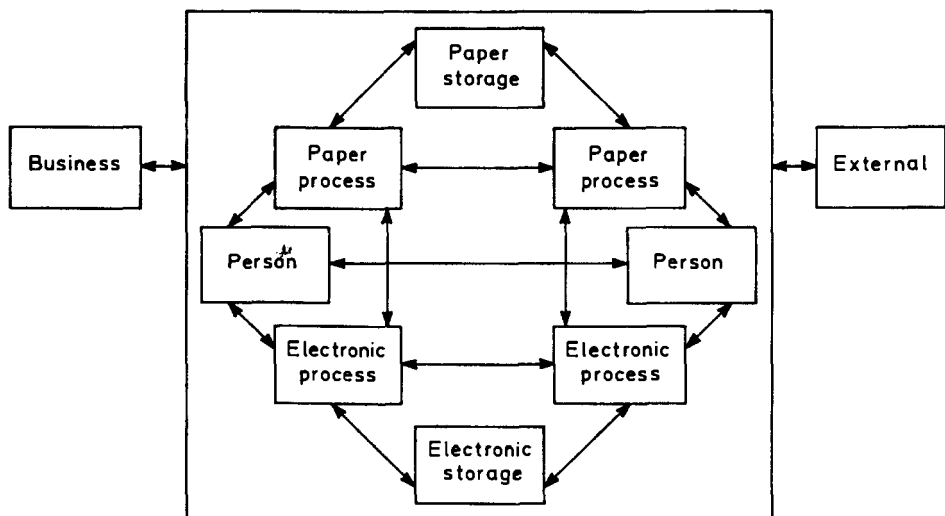


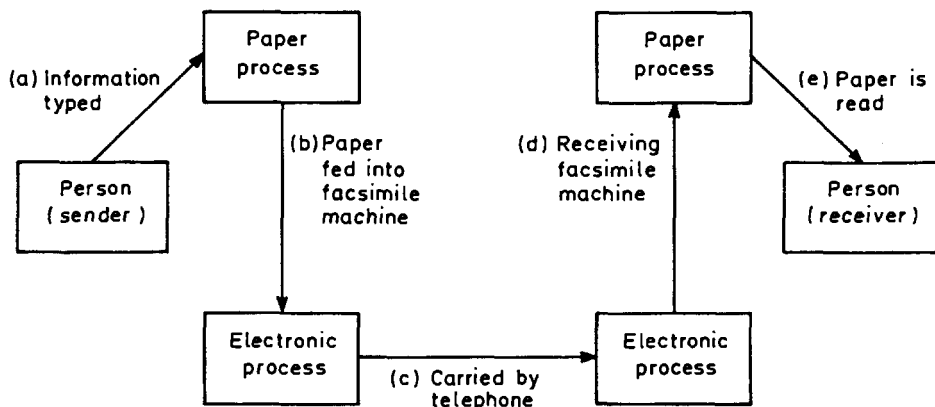
Figure 1.1 Categorizing information flows

Mode	Examples	
Paper	<u>One - to - one</u>	<u>Broadly aimed</u>
	Sending a letter Leaving a note	Advertising board
Person - to - person	Talking Winking	Acting Television Radio
Electronic	← Computer networking →	

**Figure 1.2** Categorizing simple communication methods

electronic format (b). Next the telephone network conveys the electronic information (c), before the distant facsimile machine reconverts the information to paper (d) and the receiver reads it (e).

The example we have chosen is rather convoluted, requiring several successive conversions to take place, changing the format of the information from 'personal' to 'paper' to 'electronic' and back again. All these conversions make the process inefficient, and as we shall find out later, companies who have recognized this fact have already set about converting all their key business information into electronic (computer) format, not only for conveyance, but also for storage and processing purposes.



**Figure 1.3** Information conveyance by facsimile

## 1.1 TYPES OF INFORMATION

Put specifically in the context of telecommunications, information might be a page of written text, a conversation or a television picture. The *information* usually requires conversion into an electrical signal in order to be conveyed by *telecommunication* means. However, there are many different types of information, so can they all be treated identically? The answer to this is 'no', because each type of information makes slightly different demands on the telecommunication system.

Information conveyed over telecommunications systems is usually classed as either an *analogue signal* information or as *data* (digital information). An *analogue signal* is an electrical waveform which has a shape directly analogous to the information it represents (e.g. speech or a television picture). *Data*, on the other hand, is the word given to describe information in the form of text, numbers or coded computer or video information.

Different forms of data and analogue signal information require different treatment. For example, when conversing with someone we expect their reply to follow shortly after our own speech, but when we send a letter we do not expect a reply for some days. The analogy runs directly into telecommunications. Thus, an electrical representation of conversation must allow the listener to respond instantly. However, in the case of data communication, slightly more leeway exists, as a computer is prepared to accept response times of several seconds. A human would find this length of delay intolerable in everyday speech. Another difference between electrical representations designed for different applications will be the speed with which information can be transferred. This is normally referred to as the *information rate*, the *bandwidth* or the *bitrate*. A speech circuit requires more bandwidth to carry the different voice tones than a telegraph wire needs simply to carry the same information as text. Later chapters in this book discuss the various methods of electrical representation and the technical standards used.

## 1.2 TELECOMMUNICATIONS SYSTEMS

There are four essentials for effective information transfer between two points, all of which are provided in well-designed telecommunications systems:

- a transmitting device
- a transport mechanism
- a receiving device
- the fourth requirement is that the conveyed information is coded in such a way as to be compatible with, and comprehensible to, the receiver.

All four components together form a telecommunications system.

In the example of a communication system consisting of two people talking to one another, the transmitting device is the mouth, the transport mechanism is the sound through the air, and the receiving device is the other person's ear. Provided that both people talk the same language, then the fourth requirement has also been met, and

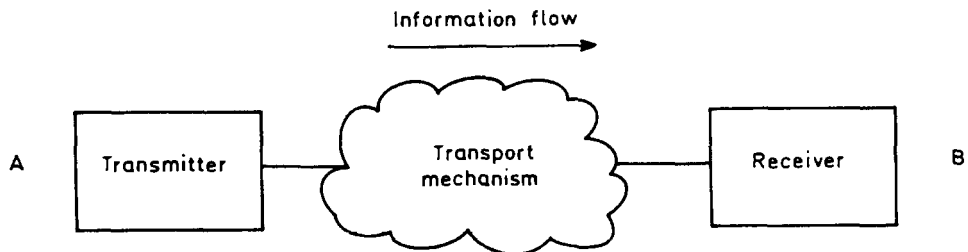
conversation can continue. However, if the talker speaks English, and the listener only understands French, then, despite the availability of the 'physical components' of the system (i.e. mouth, ear and air), communication is ineffective due to the incompatibility of the information.

The coding and method of transfer of the information over the transport mechanism is to said to be the *protocol*. In our example the *protocol* would be either the English or the French language: the fact that the talker is English and the listener French is an example of *protocol incompatibility*. *Protocol* also defines the procedure to be used. An example of the procedural part of protocol is the use of the word 'over' to signify the end of radio messages (for example 'Come in Foxtrot, Over'). The protocol in this case prompts a reply and prevents both parties speaking at once. The hardest part of telecommunications system design is often the need to ensure the compatibility of the protocol. In some cases, this necessitates the provision of *interworking devices*. In our example, the interworking device might be a human English/French interpreter.

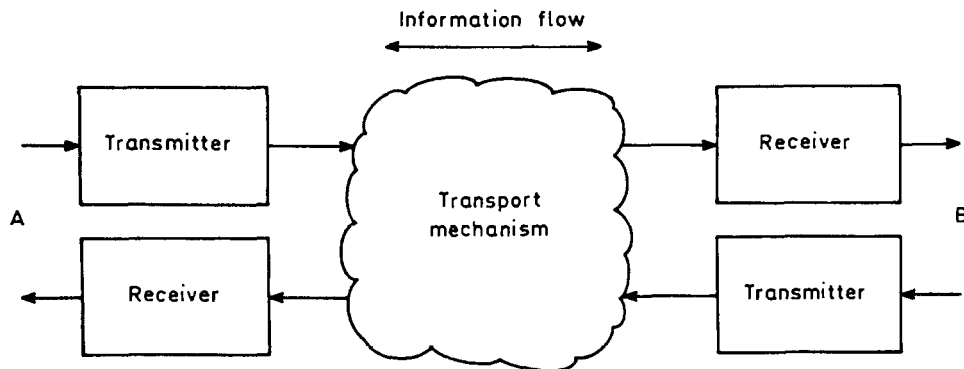
### 1.3 A BASIC TELECOMMUNICATIONS SYSTEM

Figure 1.4 illustrates the physical elements of a simple telecommunications system including the transmitter, the receiver and the transport mechanism.

As already discussed, the physical element shown in Figure 1.1 must be complemented by the use of a compatible *protocol* between transmitter and receiver. Together with such a protocol, we have all the means for communication from point A to point B in Figure 1.4. We do not, however, have the wherewithal for communication in reverse (i.e. from B to A). Such single direction communication, or *simplex operation* as it is called, may suffice for some purposes. For many more examples of communication, two way, or *duplex operation* is normally required. For duplex operation, a transmitter and a receiver must be provided at both ends of the connection, as shown in Figure 1.5. A telephone handset, for example, contains both a microphone and an earphone. Duplex operation allows both parties to talk at once and both to be able to (and have to) listen. This allows the human listener to interrupt, or two computers to send information to one another in both directions at the same time. Not all devices are capable of talking and listening at the same time as required for duplex operation.



**Figure 1.4** Basic physical elements of a telecommunications system (*simplex operation*)



**Figure 1.5** A basic duplex telecommunications system

There is also *half-duplex* operation, in which communication is possible in both directions, but not at the same time, as only one communications path is available. First the talker must stop speaking, then the listener can reply.

The transport mechanism can be one of a range of different media, ranging from sound waves passing through air to laser light pulses passing down the latest technology, optical fibre. Furthermore the transport mechanism may or may not comprise an element of switching, as we describe later in the chapter.

Most transport mechanisms demand an encoding of the information or data into a signal form suitable for conveyance over electrical transmission media. Chapters 2 to 5 describe how many of the common forms of information (e.g. speech, TV, telex, computer data, facsimile, etc.) are converted into a transmittable signal carried in either 'analogue' or 'digital' form. In Chapters 6 and 9 we discuss various methods of switching and in Chapter 8 we discuss a range of different transmission media (cables, radio systems, etc.), describing how different ones provide the optimum balance of low cost and good transmission performance for individual cases of application.

## 1.4 COMMON TYPES OF TELECOMMUNICATIONS SYSTEMS

In order to meet differing communications needs, a number of different types of telecommunications equipment have been developed over time. These include, in chronological order:

- telegraph
- telephone
- telex
- data networks using either *circuit-*, *packet-*, *frame-* or *cell-switched* conveyance
- computer *local area networks (LANs)*, *metropolitan area networks (MANs)* and *wide area networks (WANs)*

- *integrated* voice and data networks
- *multimedia* networks

This book covers the principles involved in each of the above telecommunications types; it also aims to give adequate background to enable the reader to tackle basic network planning of any of these types. The book covers networking from the simple interconnection of two telephones right up to complex, globally spread, telecommunications networks.

## 1.5 NETWORKS

Let us now consider the ideal properties of the various components of the telecommunications system illustrated in Figure 1.5. If both *stations* A and B are provided with telephones, then the transport mechanism need be no more than a single transmission line, as illustrated in Figure 1.6.

The system can also be extended to include further parties. For example, if a third station C wishes to be interconnected for private interconnection with either or both of the other two stations (A and B) then this can be achieved by duplication of the simple layout. In this way a triangular network between A, B and C is created, as shown in Figure 1.7.

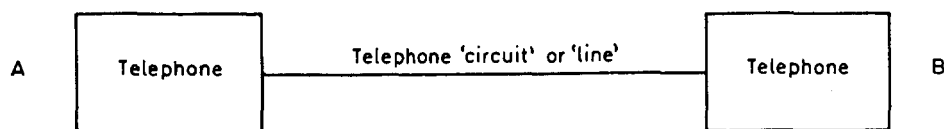


Figure 1.6 A simple two station telephone system

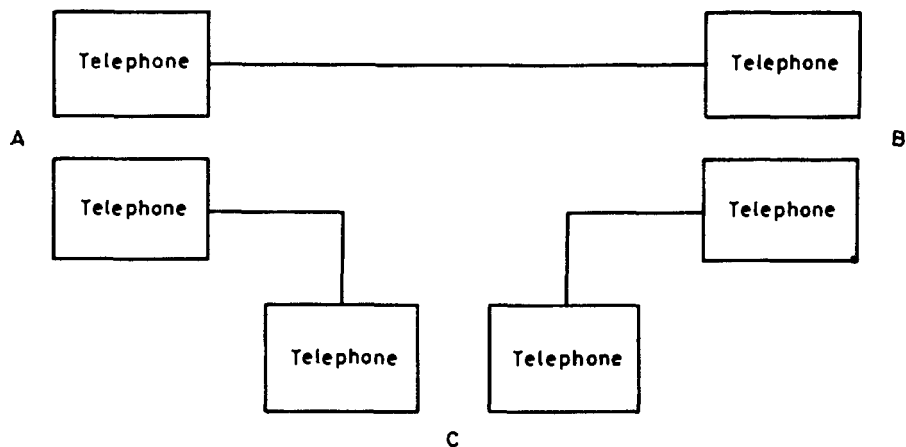


Figure 1.7 Three stations interconnected by independent telephone lines



The configuration of Figure 1.7 is used today by some companies in their private networks, where a dedicated 'private line telephone' may operate over a special telephone line, leased from a telecommunications administration, to connect different premises. However, for a network interconnecting a large number of *stations*, the configuration is uneconomical in equipment. In the three-station (A, B, C) case illustrated, six telephones and three lines are needed to interconnect the stations, but only two telephones and one line can ever be used at any one time (unless one of the people is superhuman and can talk and listen on more than one telephone at a time).

As even more stations are introduced to the configuration, the relative inefficiency grows. In a system of  $N$  stations in which each has a direct link to each other, a total of  $\frac{1}{2}N(N-1)$  telephone lines will be needed, together with  $N(N-1)$  telephone sets. If configured in the manner shown in Figure 1.7, the linking of 100 stations would need 5000 links (and 10 000 telephones) and a 10 000 station system would need 50 million lines and 100 million telephones. We need to find a more efficient configuration!

Let us limit each station in Figure 1.7 to one telephone only. To make this possible we install a switching device at each station to enable appropriate line selection, so that connection to the desired destination may be achieved on demand. This is now a simple switched network, as Figure 1.8 shows. Now the transport mechanism (stylised in Figure 1.5) is no longer just a single 'line', but is a more complex 'switch' and line arrangement.

Let us develop Figure 1.8 further, by permitting more stations (telephones in this case) to be connected to each of the three switches. Three more stations, A', B' and C' are shown in Figure 1.9. The new configuration allows the idle lines of Figure 1.8 (A-C and B-C) to be put to use.

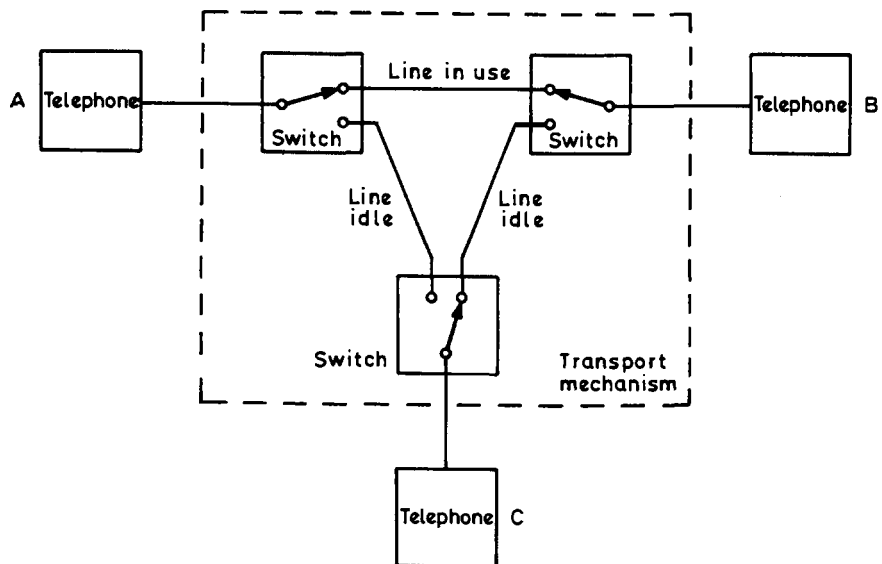
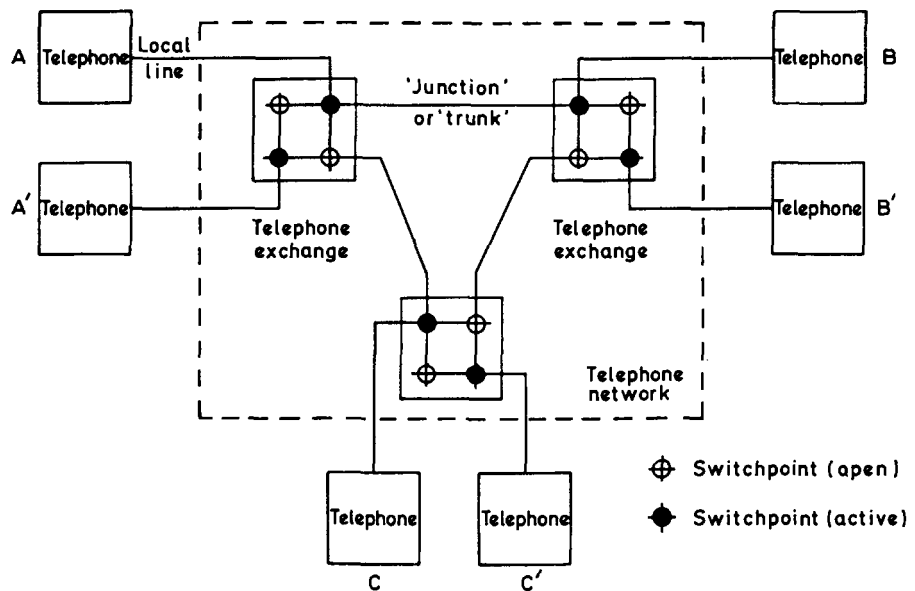


Figure 1.8 A simple three station switched network



**Figure 1.9** A simple telephone network

Figure 1.9 illustrates simultaneous calls involving A and B, B' and C', A' and C'. In this example each of the switches (which are now labelled as *exchanges*) is shared by a number of stations, each of which is switched and connected to the exchange by a *local line* or *local loop*. Our example now resembles a *public switched telephone network (PSTN)*. Because the lines between exchanges are correctly referred to as *junctions* or *trunks*, they have been labelled accordingly.

In a real telephone network the numbers of exchanges and their locations are governed by the overall number and geographical density of stations (telephone users) requiring interconnection. Similarly, the number of *junctions* or *trunks* provided between the various exchanges will be made sufficient to cater for the normal telephone call demand. In this way, far fewer junctions than stations need to be provided. This affords a significant cost saving over the configuration of Figure 1.6.

Before leaving Figure 1.9, note how each of the exchanges has been drawn as an array of individual switch points. This allows either of the telephones connected to the exchange to access either of the junctions, and this is a so-called *full availability* switch as any one of the incoming lines may be connected to any one of the available junctions. We shall come back to circuit theory of switching and availability in Chapter 6.

## 1.6 CONNECTION-ORIENTED TRANSPORT SERVICE (COTS) AND CONNECTIONLESS NETWORK SERVICE (CLNS)

In the example of the last section we justified on economic grounds alone the use of *switched* as opposed to ‘transmission line only’ networks. The particular case that we

have developed is an example of a *circuit-switched network*. Other important switched network types, especially used for data transmission, are those of *packet switching*, *message switching* and *cell switching*.

*Circuit-switched* and most *packet-switched* and *cell-switched* networks are examples of *connection-oriented switching* or *connection-oriented transport service (COTS)*. In a *connection-oriented* switching technique a *circuit*, *virtual circuit*, *connection* or *virtual connection (VC)* is established between sender and receiver before information is conveyed. Thus a telephone connection is first established by dialling, before the conversation takes place. This ensures the readiness of the receiver to receive information before it is sent. (There is no point in talking if nobody is listening). In contrast, *connectionless* switching techniques or *connectionless-network service (CONS)* allow messages to be despatched, maybe even without checking the validity of the address. Thus, for example, the postal service is analogous to a *connectionless* service. The sender posts the entire message (envelope and contents) into the post box and forgets about it. Sometime later, the receiver receives the message, either delivered directly through his letter box or by picking it up from his local post office. *Electronic mail*, today's computerized version of the postal service, is also a *connectionless network service* for sending letters directly between computers.

The advantage of *connectionless* service is that the sender need not wait for the receiver to be ready and the network need not be encumbered with the extra effort of setting-up a connection. Thus neither sender nor network need bother to keep redialling when either the receiver is already busy on another call, asleep on the other side of the world, disconnected, switched-off or otherwise unable to answer the call. Instead the message is lodged in a temporary store or 'post office'-like device. The disadvantage is that the sender has no clear guarantee or confirmation of message delivery. He is left in doubt: did the receiver not get the message or was he simply too lazy to reply?

*Message switching* networks are networks which deliver the message (e.g. letter or document) in one go. Most *message switching* networks (including perhaps the best known, *Internet*) are based on *connectionless network service*.

As an aside, sometimes the end-to-end communication is *connectionless* even though each of the individual links in the physical communication chain is a *connection-oriented* connection. Thus both sender and receiver might telephone an electronic mail *post office* to send and receive their mail. The connection between the two post offices may be a permanent connection and both telephone calls (to deliver and pick up the message) are connection-oriented, but because the sender and receiver do not both need to connect to their respective post offices at the same time, the end-to-end communication is *connectionless*.

## 1.7 CIRCUIT-, PACKET- AND CELL-SWITCHED NETWORKS

The distinguishing property of a *circuit-switched connection* is the existence throughout the communication phase of the *call*, of an unbroken physical and electrical path between origin and destination points. The path is established at *call set-up* and *cleared* after the call. The path may offer either one direction (simplex) or two direction (duplex) use. Telephone networks are circuit-switched networks.

Conversely, although *packet-switched networks* are also *connection-oriented*, an entire physical path from origin to destination will not generally be established at any time during communication. Instead, the total information to be transmitted is broken down into a number of elemental 'packets', each of which is sent in turn. An analogy might be sending the text of a large book through the post in a large number of envelopes. In each individual envelope might be just a single page. The envelopes can either be sent sequentially (say one each day), or numbered so that the receiver can reassemble the pages in order. Figure 1.10 illustrates this simple packet-switched communication system. Should the receiver in Figure 1.10 not receive any given numbered envelope, he may write back over the reverse connection and re-request it. In this way, very accurate and reliable communication may be established.

Packet-switched networks are usually *connection-oriented*. A connection *set-up* phase confirms the readiness of the receiver to receive information and it determines the route through the network which will be used to carry the packets. The *connection* which results is actually termed a *virtual connection*, because though it appears to the two end-users as though a dedicated path exists, the physical connection is actually shared with other users. By breaking the information into *packets*, *statistical multiplexing* may be used to increase the network throughput. Statistical multiplexing is the technique of sending packets from different users' *virtual connections* over the same physical connection (Figure 10). This is made possible by labelling each packet (pages of Figure 1.10) with the identity of the virtual connection to which it belongs (separate books could be sent simultaneously in Figure 1.10). The labelling allows packets to be sent from any of the virtual connections, provided that the line is at that moment idle. If the line is already busy, it may be that the new packets must wait a fraction of a second before transmission is possible. This possibility of slight delay leads to another description of packet-switching as a *store-and-forward technique*. Packet-switching is the technique behind X.25, frame relay and many other computer network techniques.

*Circuit-switched* networks are generally necessary when very rapid or instantaneous interaction is required (as is the case with speech). Conversely, packet-switched

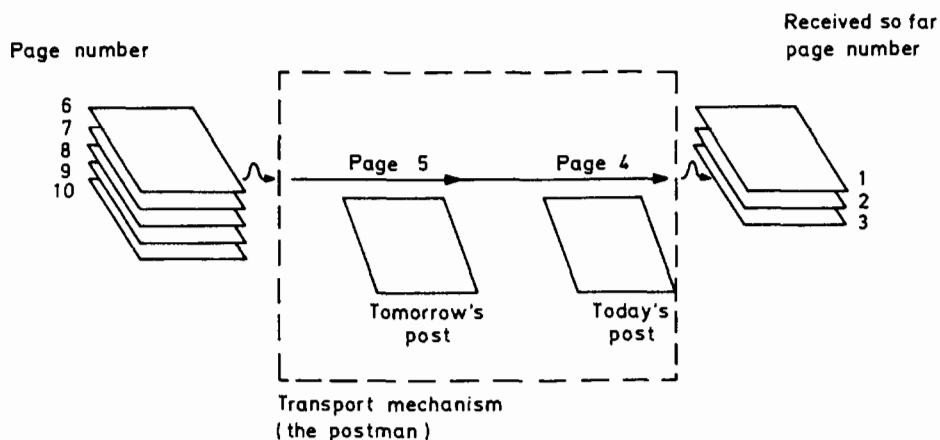


Figure 1.10 A simple 'message switched' communication system

networks are more efficient when instantaneous reaction is not required, but when very low 'corruption' (either through distortion or loss of information) is paramount.

*Cell-switching* is a specialized form of packet-switching in which the packet lengths are standardized at a fixed length. As we shall see in Chapters 25 and 26, *cell switching* or *cell relay switching* is the basis of *multimedia* networks, the *broadband integrated services digital network (B-ISDN)* and *ATM (asynchronous transfer mode)*.

## 1.8 CONSIDERATIONS FOR NETWORK PLANNERS

We have established the economic value of switched networks (as opposed to direct wire systems), and briefly addressed the needs for basic communication (the physical, switching and protocol needs). What other factors need to be provided in order to enable networks to function in a manner fit for purpose? The following list is a brief summary of some of the factors which require consideration.

### *Transmission and configuration plan*

A plan laying out guidelines according to which appropriate transmission media may be arranged so that adequate end-to-end conveyance of information is achieved. The plan will include safeguards to ensure that the received signal is loud enough, clear enough and free of noise and interference.

### *Numbering and routing plans*

The numbering plan is crucial to the ability of the network to deliver communications to the appropriate destination. Much like reading an address on an envelope, it is the inspection of the destination number that permits the determination of destination and appropriate route within the network.

### *Usage monitoring plan*

A usage monitoring plan is needed to ensure continuing and future suitability of the network. Adequate measurements of network performance are required so that early steps may be taken, as necessary, to adjust the network design, or expand its overall throughput capacity to meet demand.

### *Charging and accounting plan*

Public telecommunication operators (PTOs) need reimbursement for their services. Specific equipment may be required to monitor individuals' use of the network, so that bills can be generated.

### *Maintenance plan*

A maintenance plan is needed to guarantee that the service level meets agreed targets, that routine maintenance is carried out, and that problems are quickly seen to. Any of these items may be appropriate to any type of telecommunications network, regardless of what type of information it is carrying (for example, telephone, data or multimedia network). Each is discussed more fully in later chapters.

## 1.9 TECHNICAL STANDARDS FOR TELECOMMUNICATIONS SYSTEMS

As a way of promoting greater compatibility between various telecommunications systems in different parts of the world, a number of bodies that define technical standards are active in various parts of the world. These are discussed in more detail in Chapter 40. However, two of the most significant and influential bodies are worth mentioning now. These are:

- the International Organization for Standardization (ISO)
- the International Telecommunications Union (ITU), and specifically its sub-entities, *the standardization sector (ITU-T, formerly CCITT, consultative committee for international telephones and telegraphs)* and *the radiocommunication sector (ITU-R, formerly CCIR, consultative committee for international radiocommunication)*.

The *standards* and *recommendations* of these two bodies are commonly used throughout telecommunications, and are often referred to by this book.



---

# *Introduction to Signal Transmission and the Basic Line Circuit*

---

To make it suitable for carriage over most telecommunications networks, information must first be encoded in an electrical manner, as an electrical signal. Only such signals can be conveyed over the wires and exchanges that comprise the ‘transport mechanism’ of telecommunications networks. Over the years, a variety of different methods have been developed for encoding different types of information. This chapter categorizes all methods of information transmission into one of two basic types, analogue transmission and digital transmission, describing the principles of both types but concentrating on the technique of analogue transmission. Chapter 3 continues the subject of analogue transmission, explaining its use for information conveyance over long distances, and Chapters 4 and 5 expand the details of digital transmission.

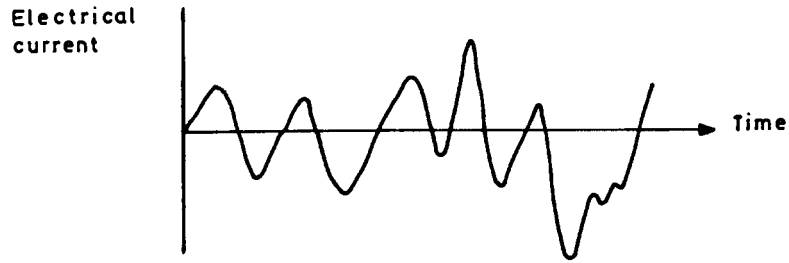
## 2.1 ANALOGUE AND DIGITAL TRANSMISSION

The simplest type of *physical medium* (transport mechanism) used in telecommunications networks is a pair of electrical wires. These allow the conveyance of an electrical current or *signal* from a transmitter at one end of the wires to the receiver at the other end.

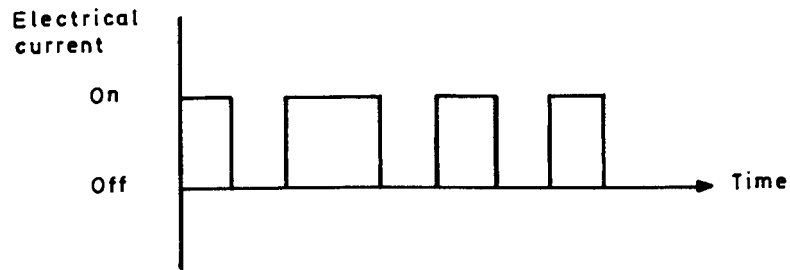
There are two basic methods for information encoding which may be applied by the transmitter and decoded by the receiver. These are analogue signal encoding and transmission, and digital encoding and transmission.

As the name suggests, analogue signal encoding involves the creation of an electrical signal waveform which is analogous to the waveform of the original information (e.g. a speech wave pattern). In this way the electrical signal is similar in shape to the speech or other information waveform that it represents. Analogue transmission lines are used to convey analogue encoded signals and, until the early 1980s, predominated as the links of the world’s telecommunications networks. *Analogue exchanges* have provided the nodal interconnections between these links. Digital encoding and transmission of information (a technique in which information is conveyed as a string of *digits*, or more





(a) Typical analogue signal



(b) Typical digital signal

**Figure 2.1** Typical analogue and digital systems

precisely, pulses of 'on' and 'off' electrical current) is rapidly taking over as the main method of telecommunications transmission. This is because of the significant benefits that digital transmission offers, in terms of performance and cost. Figure 2.1 illustrates typical analogue and digital signal waveforms.

As for the transmission line systems themselves, Table 2.1 presents a short summary of some of the *physical media* which may be used, and compares their relative merits. In general any *physical medium* may be designed to work in either *analogue* or *digital* mode, and each of the different transmission media listed in Table 2.1 has been used as the basis of both analogue and digital transmission systems, though digital systems tend today to be the system of choice. Choosing exactly which medium and encoding system should be used is a decision for the network planner, who needs to consider the relative costs and the ease with which any new type of line system could be incorporated into the existing network. We shall discuss the relative benefits of different transmission media in more detail in Chapter 8.

Switching equipment and exchanges are also classified into *analogue* and *digital* types. A switch, is after all, only a specialized and very flexible form of transmission equipment.

To develop our understanding of information encoding, we now discuss the various techniques in chronological order.

**Table 2.1** Comparison of common transmission media

Transmission medium used as basis of line system	Comments
Copper wire (or aluminium wire)	Ordinary wire. The gauge (or thickness) will be carefully selected for the particular use. The pairs of wires may be either 'parallel' or 'twisted' within the 'cable', which usually comprises many pairs. Copper and aluminium wire is the predominant system used for the 'local loop network'.
Coaxial cable	Centre conductor plus cylindrical conducting sheath. Used on long-haul or 'trunk' systems. Particularly used to carry many tens of hundreds of communications channels simultaneously; in 'multiplexed mode' (described later).
Microwave radio	Small dish, 'line-of-sight' radio system. Used for long-haul 'trunks', particularly where terrain is difficult for cable laying.
Tropospheric scatter	Radio system using high power to achieve 'over the horizon' communication. 100–600 km but susceptible to fading.
High-frequency radio (HF, VHF and UHF)	Used historically for ship-to-shore communications. Now increasingly used by the flourishing mobile cellular radio networks of portable telephone handsets.
Satellite	Satellites are generally intended to be 'geostationary' (permanently above a particular point on the equator). They give a single hop access to approximately one third of the globe. The disadvantage is the comparatively long half-second transmission time.
Optical fibre	Glass fibre, of a thickness comparable with a human hair. Used in conjunction with lasers, optical fibres are capable of the carriage of thousands of simultaneous calls.

## 2.2 TELEGRAPHY

Electric telegraphy was the earliest common method of digital transmission. It began in the early part of the 19th century, and became the forerunner of modern telecommunications. Samuel Morse, inventor of the Morse code, developed a working telegraph system in 1835. By 1837 a practical system was established from Euston railway station in London to Camden, a mile away. The system came about through the desire of railway signalmen to ensure the safety of trains, the number of which was rapidly growing as a result of the industrial revolution. In 1865 Morse code was accepted as a world standard during the first International Telegraph convention, and this gave birth to the *International Telecommunications Union (ITU)*. Public interest in telegraphy grew rapidly until about 1880, when telephones became more widely available.

Telegraphy formed the basis of the early telegram service. It provided the means of sending short textual messages between post offices so that immediate delivery could then be arranged by the 'telegram boy' carrying it by cycle to the recipient's premises.

A simple telegraph circuit is illustrated in Figure 2.2. The system works by transmitting simple on/off electrical pulses from the transmitter (the key) to the receiver

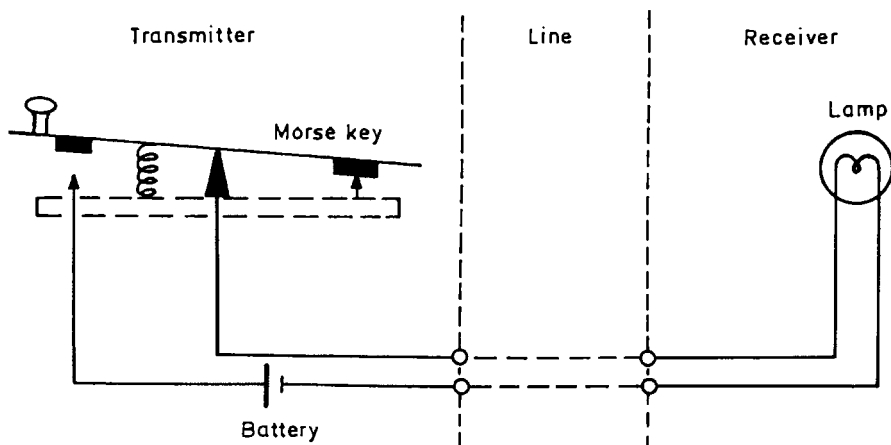


Figure 2.2 A simple telegraph circuit

(the lamp). A single pair of wires provides the transmission medium for the electrical current between origin and destination.

In the early days, telegraphy was carried out by manual *keying* (switching the circuit 'on' and 'off') according to the Morse code. The code split the message up into its component alphabetic characters, each of which is represented by a unique sequence of short or long electrical pulses (so-called *dots* and *dashes*). For those readers not familiar with the code it is described more fully in Chapter 4.

*Dots* and *dashes* are sent by pressing the *key* of Figure 2.2 for an appropriate length of time. Consecutive dots or dashes are separated by a *space* (a period of no current – key not pressed) equal in length to one *dot*. Three successive spaces are used to separate the sequences of dots and dashes which together represent a single character. Words are separated by five spaces.

As telegraph networks were gradually replaced by telex networks, which essentially are no more than *automatic telegraph* networks, so the Morse code was superseded by the Murray code, another digital code, but one which was especially designed for automatic working. It differs from the Morse code in that every alphabetic and numeric character is represented by a fixed number (five) of electrical pulses, so-called *marks*, or *spaces*. Unlike the dots and dashes of the Morse code, the *marks* and *spaces* of the Murray code are all of equal length. The code is more suitable for automatic working as it is easier to design mechanical equipment to respond to pulses (either marks or spaces) which have a predetermined duration. For example, a telex receiver or *teleprinter* can more easily select and print characters at a fixed speed than it can at a variable one. The evolution from the Morse to the Murray code took place in a number of stages, first to a manual and then an automatic teleprinter system, finally to the *Telex* system. The output of the early teleprinter devices was a narrow strip of paper which could be stuck onto a receiving Telegram form. The Telex system by contrast prints direct onto a roll of ordinary paper. Only in the late 1980s did facsimile machines start to supersede telex as a widespread means of text communication.

In both telegraph and telex networks, two-way sequential transmission over a single pair of wires is possible simply by providing a device capable of both the transmitting and the receiving functions at both ends of the line.

## 2.3 TELEPHONY

Telephony, invented by Alexander Graham Bell and patented in the USA in 1875–7 is perhaps today's most important means of long distance communication. Certainly in terms of overall volume of information carried, it is far and away the most important of the long distance telecommunications methods, though lately starting to be overtaken by the explosion in information carried by computer networks.

Telephony is the transmission of sound, in particular, speech, to a distant place. The literal translation of 'tele-phone' from the Greek means 'long distance-voice'.

Telephony is achieved by conversion of the sound waves of a speaker's voice into an equivalent electrical signal wave. Let us consider the characteristics of sound, so that we may understand how this conversion is performed. Sound is really only vibrations in the air around us, caused by localized pockets of high and low air pressure which have been generated by some form of mechanical vibration, for example an object hitting another, or the vibration of the human vocal chords during speech.

Sound waves in the air around us cause objects in the vicinity to vibrate in sympathy. The human ear detects sound by the use of a very sensitive diaphragm, which vibrates in synchronism with the sound hitting it.

The pitch of a sound (how 'high' or 'low' it sounds) depends on the *frequency* of the sound vibration. A typical range of frequencies which are audible to the human ear are 20–20 000 Hz (or vibrations per second). Low frequencies are heard as low notes, high frequencies as high notes, and most sound is a very complex mixture of frequencies as Figure 2.3 illustrates. Figure 2.3(a) is a 'pure' tone of a single frequency. To a listener this

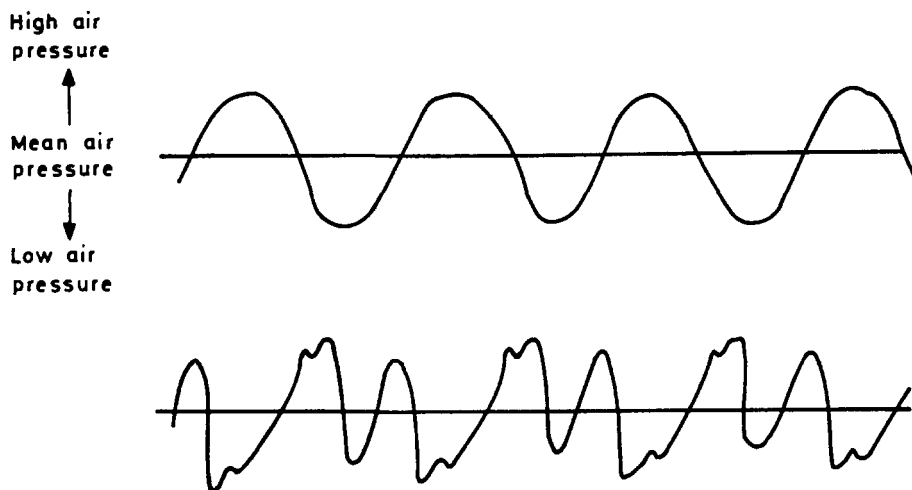


Figure 2.3 (Top) Waveform of 'pure' (single frequency) tone. (Bottom) Typical speech waveform

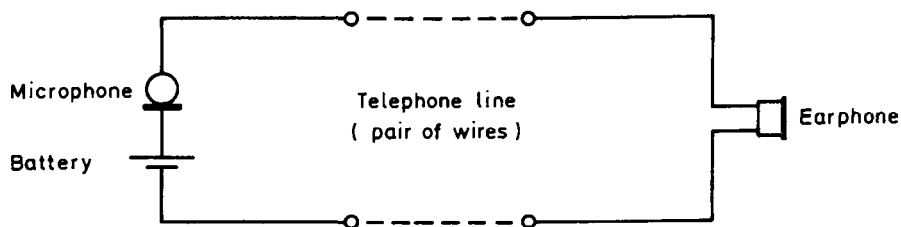


Figure 2.4 A simple telephone circuit

is the rather dull and insipid note produced by a tuning fork. By contrast, Figure 2.3(b) shows the waveform of the vowel 'o', as in 'hope'. Note how the waveform is not as smooth as before. In reality the waveform will be even more complicated, preceded and followed by other complicated patterns making up the sounds of the rest of the word. Real sounds have an irregular and spiky waveform.

Concert grade musical performances use a wide range (or *bandwidth*) of frequencies within the human aural range. However, for the purpose of intelligible speech, only the frequencies from 300–3400 Hz are necessary. If only this bandwidth of frequencies is received the quality is slightly downgraded but is still entirely comprehensible and acceptable to most listeners.

So how can the sound waveform be converted into an electrical signal? Figure 2.4 illustrates a simple telephone circuit capable of the conversion. It consists of a microphone, a battery, a telephone line and an earphone.

When no-one speaks into the microphone, then both the microphone and earphone have a steady electrical resistance, and so a constant electrical current flows around the circuit according to Ohm's law (*current = voltage divided by resistance*). If someone speaks into the microphone, then the sound waves hitting the microphone diaphragm cause the electrical resistance of the device to alter slightly. The change in resistance results in a corresponding change in the circuit current. To illustrate this, Figure 2.5 shows the circuit components in more detail. On the left of the diagram is a *carbon microphone*, common in many countries. The diaphragm of the microphone compacts or rarefies the carbon granules, and this results in the change of resistance. The change in resistance causes a change in the current around the circuit. The current varies about

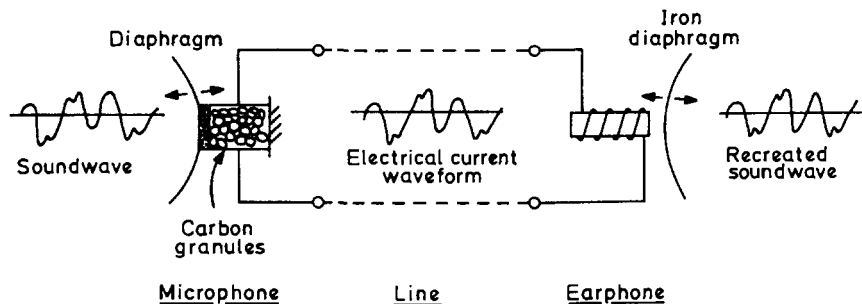


Figure 2.5 The carriage of a soundwave over a simple telephone circuit

the steady-state value in the same way that the air pressure in the sound wave fluctuates about its average pressure. The effect is to create an electrical signal fluctuating in almost the same way as the original sound wave.

Conversely, in the earphone the electrical signal is reconverted to sound. The simple earphone shown in Figure 2.5 is constructed from an iron diaphragm and an electromagnet. An electromagnet is an electrical device which acts like a magnet when current passes through its winding and tends to attract the iron diaphragm accordingly. The strength of the attraction depends on the magnitude of the current flowing in the coil of the electromagnet (i.e. on the telephone circuit current). The varying electrical signal created by the microphone results in a varying force of attraction on the diaphragm. This vibrates the surrounding air and recreates the original soundwave.

The conversion process described above is an *analogue* one, as the electrical waveform carried through the *analogue network* is analogous to the original sound waveform.

## 2.4 RECEIVED SIGNAL STRENGTH, SIDETONE AND ECHO

In an analogue network, it is important that the received electrical signal strength is sufficient to reproduce soundwaves loud enough for the listener to hear. Herein lies a problem with the circuit of Figure 2.5. The circuit works adequately at short line lengths, but if too long a line were to be used, then because the electrical resistance of the microphone would be substantially smaller than the resistance of the line itself, the microphone would not be able to stimulate much fluctuation in the circuit current. The result is faint sound reproduction. To get around this problem, the circuit is improved by the use of transformers, as shown in Figure 2.6. The greater number of turns on the secondary winding of each transformer (as compared with the primary winding) reduces the effect of the line resistance.

You will also notice in Figure 2.6, that the circuit connects the equivalent of two complete telephone handsets (i.e. it contains two microphones and two earphones). The circuit is therefore capable of two-way communication.

A real telephone circuit normally uses a more complicated circuit. The added complication is needed to control the effect of *sidetone* and reduce the chance of *echo*. Consider the following example in explanation of these effects. As party A speaks, the voice will be heard in earphone B, but also in earphone A. The speaker will therefore hear *sidetone* of his or her own voice. If the sidetone is too loud, the speaker may be distracted. Conversely, if there is no sidetone, the speaker sometimes gets the impression that the phone is 'dead'. An *anti-sidetone circuit* is normally incorporated into the telephone circuitry to control the level of sidetone. It does so by restricting the volume of speech and other noises picked up by the microphone from being heard in the speaker's own earphone. The anti-sidetone circuit also eases the difficulty of hearing weak incoming signals in noisy places (e.g. a public telephone in a cafe). The circuit has the further benefit that it controls the undesirable effect of *echo*. Consider again speaker A's position. The voice is emitted by earphone B, and if too loud is picked up by microphone B and returned to earphone A. The slight delay introduced by the signal transit time (along the line and back) means that 'A' hears a slightly delayed echo of his

or her own voice. The anti-sidetone circuit in instrument B helps to control this effect, by reducing the signal *feedback* from the earphone to the microphone of B's handset. The *echo path* from microphone A (via earphone B, microphone B and earphone A) is therefore largely eliminated. We will learn in Chapter 33, however, that this is not the only possible echo path in practical networks, and that even more precautions may need to be taken.

## 2.5 AUTOMATIC SYSTEMS: CENTRAL BATTERY AND EXCHANGE CALLING

A practical problem encountered by the circuit in Figure 2.6, when it was employed in early telephone networks was that it required an individual battery (*local battery*) to be provided in each telephone customers line. Thus a *common* (or *central*) battery (CB) system was developed. One large battery, or power system, is positioned in each exchange and it provides the necessary power for all customers' lines terminated at that exchange. The use of a CB system (in preference to individual line *local batteries*) has significant benefit in simplifying the job of maintenance.

Figure 2.7 illustrates a central battery configuration. Note that a high impedance coil has been added to the circuit. This forms part of another normal circuit feature called *the transmission bridge*. Without the *transmission bridge* coil the speech currents from one telephone would be shorted out by the battery and not heard in the receiving telephone. An inductive *coil* prevents this *shorting*. A high resistance can be used but this has the disadvantage of causing conversation *overhearing* in all the other telephones connected to the same central battery.

Returning now to the circuit of Figure 2.6, there are three more deficiencies that make it unsuitable for direct use in a real telephone network. These are that

- the circuit is always drawing electrical power
- the customer has no means of calling the exchange, to set-up calls to other customers
- the exchange has no means of *alerting* the called customer (for example, ringing the telephone)

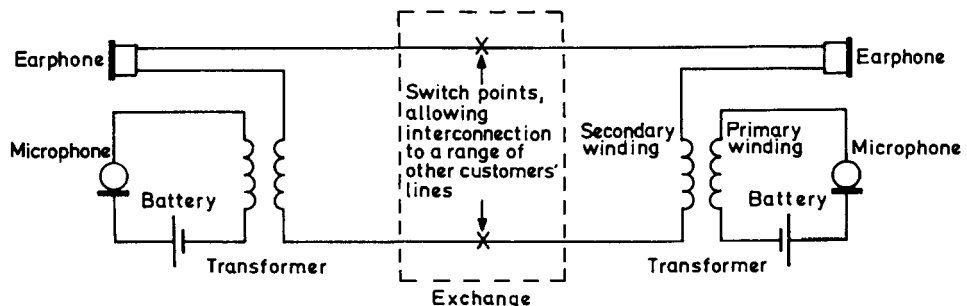


Figure 2.6 Basic two-way telephone circuit

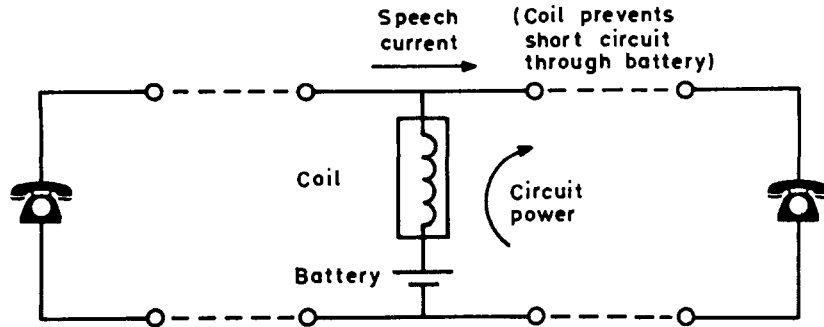


Figure 2.7 A simple transmission bridge

The first two problems above can be solved by providing a switch in the telephone handset which is activated when the handset is lifted off the *cradle*. In this way the telephone circuit is 'open circuit' when the handset is *on-hook* and does not draw power when not in use. When required for use, lifting the handset completes the circuit, an action which is sometimes called *looping* the circuit. On detection of the *loop* the exchange readies itself to receive the telephone number of the *called party*, and indicates this readiness by returning *dialling tone*. The *calling customer* may then dial the number on the telephone handset and the number is relayed to the exchange by one of the following two signalling methods:

- *loop and disconnect pulsing*
- *multi-frequency (MF) tone signalling*

*Loop and disconnect pulsing* or *LD signalling* is the older of the two methods and the most prevalent. The line is *disconnected* and re-looped in a very rapid sequence. The effect is to cause a short period, or *pulse* of disconnection. The pulses are used to indicate the digits of the destination number, one pulse for a digit of value 'one', two pulses for value 'two', etc., up to 10 pulses for digit 'zero'. The pulses are repeated at a frequency of ten per second. Thus the digit '1' will take one-tenth of a second to send, and digit '0' will take a whole second. Between each digit, a longer gap of half a second is left, so that the exchange may differentiate between the sequences of pulses representing consecutive digits of the overall numbers. The gap is called the *inter-digit pause (IDP)*. A telephone number is typically ten digits in length and takes 6–15 seconds to dial from the telephone, depending on the values of the actual digits.

Multi-frequency (MF), or *tone signalling* is a more recent invention. Its use is growing because the digits may be relayed much faster (in 1–2 seconds) and this reduces the call set-up time or *post-dialling delay (PDD)* if the exchanges are fast in their switching action. The exchange still needs to be alerted initially by looping the line, but following the dial tone the digits are dialled by the customer as a number of short burst of tones. Rather than detecting the disconnect pulses, the exchange must instead detect the frequency of the tones to determine the digit values of the destination number. Each digit is represented by a combination of two pure frequency tones. The use of two tones

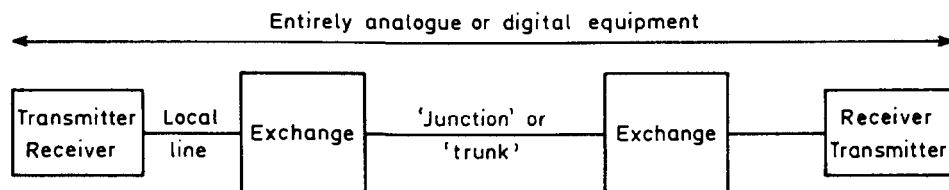


**Table 2.2** Multi-frequency (MF) tone signalling (customer to exchange)

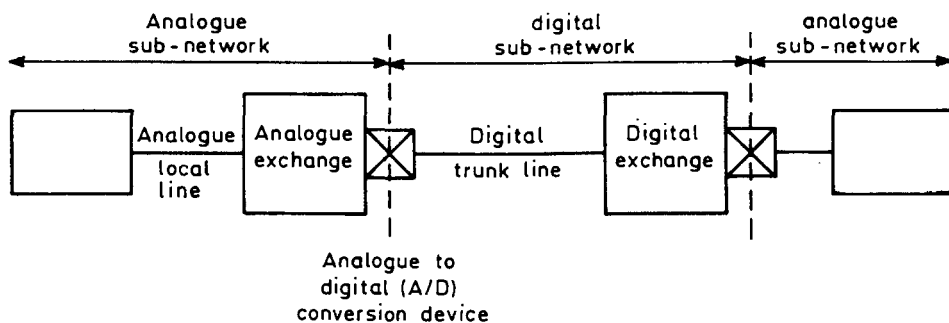
Low frequency group	High frequency group			
	1208 Hz	1336 Hz	1477 Hz	1633 Hz (not used)
697 Hz	1	2	3	spare
770 Hz	4	5	6	spare
852 Hz	7	8	9	spare
941 Hz	*	0	#	spare

in combination reduces the risk of mis-operation if other interfering noises are present on the line. Table 2.2 lists the frequencies used.

Following receipt of the number, the exchange network sets up a connection to the desired destination and *alerts* the *called customer*. *Ring*ing the telephone is the normal means of *alerting*, or attracting the called customer's attention. To achieve ringing, a large alternating current is applied to the called customer's line to activate the bell.



(a) A simple network of entirely analogue or digital components



(b) A network comprising both analogue and digital components

**Figure 2.8** Interworking dissimilar networks

Simultaneously, a *ring tone* is returned to the caller. The bell does not ring at other times because the normal line current is insufficient to cause the bellhammer vibration.

In response to the ringing telephone, the called customer picks up the handset, and therefore *loops* the line. On detection of the *loop*, the exchange *trips* (ceases) the ringing, and conversation may then take place. At the end of the conversation the caller replaces the handset thereby removing the loop, so that the line is disconnected. The exchange responds by *clearing* the rest of the connection.

## 2.6 REAL COMMUNICATIONS NETWORKS

We have now developed all the basic principles of telephone communication. Similar principles apply in nearly all communication network types, and the remaining chapters of this book develop these principles, discussing in detail the solutions to the complications that exist in practice. Before ending the chapter, however, it is valuable to point out another reality of practical networks. Although similar principles might form the foundation of nearly all telecommunications networks, it is not true to say that all networks are fully compatible, allowing direct interconnection. Indeed direct interconnection of dissimilar networks (e.g. telephone and telex) is almost impossible, with little exception. Where interconnection of dissimilar networks is possible, it is usually by the use of special *interworking* or conversion equipment at the interface of the two dissimilar networks.

Even two telephone networks, one operating using the analogue encoding technique, and the other using the digital one, cannot be directly interconnected without the use of conversion equipment. The equipment is needed to convert the information from its digitally encoded form into the equivalent analogue signal, and vice versa. The final diagram in this chapter, Figure 2.8, shows an example of interworking dissimilar networks. It illustrates the positioning of analogue to digital (and vice versa) conversion equipment at the interface of interconnected analogue and digital sub-networks.



---

# Long-haul Communication

---

None of the circuits that we have so far discussed are suitable as they stand for long haul communication. To get us anywhere with long haul we need to address ourselves to the following inescapably pertinent topics:

- attenuation (loss of signal strength over distance)
- line loading (a way of reducing attenuation on medium length links)
- amplification (how to boost signals on long haul links)
- equalization (how to correct tonal distortion)
- multiplexing (how to increase the number of 'circuits' that may be obtained from one physical cable)

In this chapter we discuss predominantly how these line issues affect analogue transmission systems and how they may be countered. The effects on digital transmission are discussed in later chapters.

## 3.1 ATTENUATION AND REPEATERS

Sound waves diminish the further they travel and electrical signals become weaker as they pass along electromagnetic transmission lines. With electrical signals the attenuation (as this type of loss in signal strength is called) is caused by the various electrical properties of the line itself. These properties are known as the *resistance*, the *capacitance*, the *leakance* and the *inductance*. The attenuation becomes more severe as the line gets longer. On very long haul links the received signals become so weak as to be imperceptible, and something needs to be done about it. Usually the attenuation in analogue transmission lines is countered by devices called *repeaters*, which are located at intervals along the line, their function being to restore the signal to its original wave shape and strength.

Signal attenuation occurs in simple wireline systems, in radio and in optical fibre systems. The effect of attenuation on a line follows the function shown in Figure 3.1, where the signal *amplitude* (the technical term for signal strength) can be seen to fade with distance travelled according to a negative exponential function, the rate of decay of the exponential function along the length of the line being governed by the attenuation constant, alpha ( $\alpha$ ).

Complex mathematics, which we will not go into here, reveal that the value of the attenuation constant for any particular signal frequency on an electrical wire line transmission system is given by the following formula:

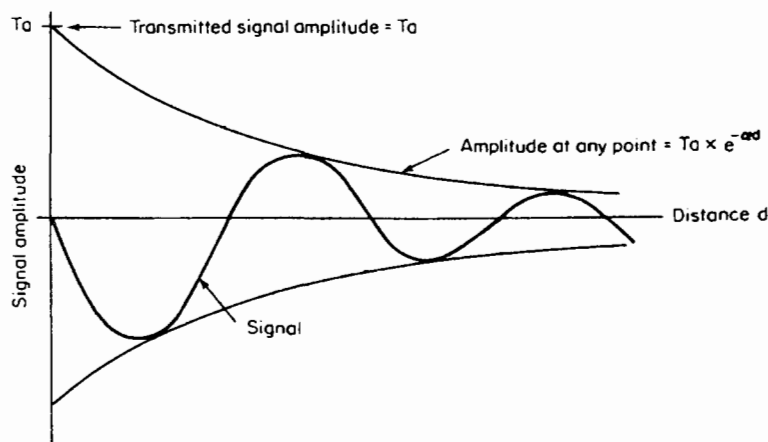
$$\alpha = \sqrt{\left(\frac{1}{2} \left\{ \sqrt{[(R^2 + 4\pi^2 f^2 L^2)(G^2 + 4\pi^2 f^2 C^2)]} + (RG - 4\pi^2 f^2 LC) \right\}\right)}$$

The larger the value of  $\alpha$ , the greater the attenuation, the exact value depending on the following line characteristics:

- $R$ : the electrical *resistance* per kilometre of the line, in ohms
- $G$ : the electrical *leakance* per kilometre of the line, in mhos
- $L$ : the *inductance* per kilometre of the line, in henries
- $C$ : the *capacitance* per kilometre of the line, in farads
- $f$ : the *frequency* of the particular component of the signal.

The *resistance* of the line causes direct power loss by impeding the onward passage of the signal. The *leakance* is the power lost by conduction through the insulation of the line. The *inductance* and *capacitance* are more complex and current-impeding phenomena caused by the magnetic effects of alternating electric currents.

It is important to realize that because the amount of attenuation depends on the frequency of the signal, then the attenuation may differ for different frequency components of the signal. For example, it is common for high frequencies (treble tones) to be disproportionately attenuated, leaving the low frequency (bass tones) to dominate. This leads to *distortion*, also called *frequency attenuation distortion* or simply *attenuation distortion*.



**Figure 3.1** The effect of distance on attenuation.  $\alpha$  = attenuation constant

### 3.2 LINE LOADING

As the *inductance* and *capacitance* work against one another, a simple means of minimizing the effect of unwanted attenuation and distortion is to increase the inductance of the line to counteract some or all of the line capacitance.

Taking a closer look at the equation given in the last section, we find that attenuation is zero when both resistance and leakance are zero, but will have a real value when either resistance or leakance is non-zero. The lesson is that both the resistance and the leakance of the line should be designed to be as low as is practically and economically possible. This is done by using large *gauge* (or diameter) wire and good quality insulating sheath.

A second conclusion from the formula for the attenuation constant is that its value is minimized when the inductance has a value given by the expression

$$L = CR/G \text{ henries/km}$$

This, in effect, reduces the electrical properties of the line to a simple resistance, minimizing both the attenuation and the distortion simultaneously. In practice, the *attenuation* and *distortion* of a line can be reduced artificially by increasing its inductance  $L$  ideally in a continuous manner along the line's length. The technique is called *line loading*. It can be achieved by winding iron tape or some other magnetic material directly around the conductor, but it is cheaper and easier to provide a *lumped loading coil* at intervals (say 1–2 km) along the line. The attenuation characteristics of unloaded and loaded lines are shown in Figure 3.2.

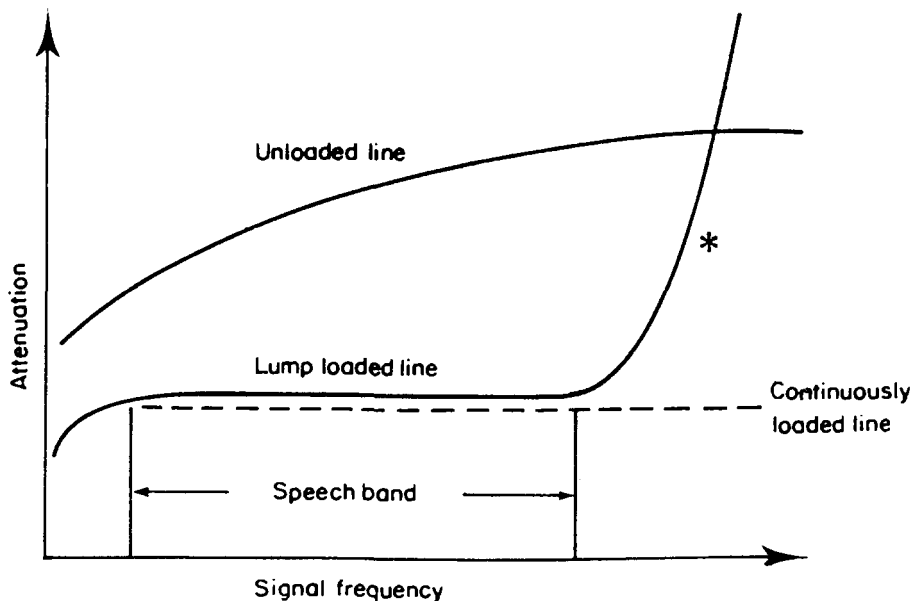


Figure 3.2 Attenuation characteristics of loaded line

The use of *line loading* has the disadvantage that it acts as a high frequency filter, tending to suppress the high frequencies. It is therefore important when *lump loading* is used, to make sure that the wanted speech band frequencies suffer only minimal attenuation. If the steep part of lump loaded line curve (marked by an asterisk in Figure 3.2) were to occur in the middle of the speech band, then the higher frequencies in the conversation would be disproportionately attenuated, resulting in heavy and unacceptable distortion of the signal at the receiving end.

### 3.3 AMPLIFICATION

Although *line loading* reduces speech-band *attenuation*, there is still a loss of signal which accumulates with distance, and at some stage it becomes necessary to boost the signal strength. This is done by the use of an electrical amplifier. The reader may well ask what precise distance line loading is good for. There is, alas, no simple answer as it depends on the gauge of the wire and on the transmission bandwidth required by the user. In general, the higher the bandwidth, the shorter the length limit of loaded lines (10–15 km is a practical limit).

Devices called *repeaters* are spaced equally along the length of a long transmission line, radio system or other transmission medium. *Repeaters* consist of amplifiers and other equipment, the purpose of which is to boost the basic signal strength.

Normally a *repeater* comprises two amplifiers, one for each direction of signal transmission. A splitting device is also required to separate *transmit* and *receive* signals. This is so that each signal can be fed to a relevant *transmit* or *receive* amplifier, as Figure 3.3 illustrates. The splitting device is called a *hybrid* or *hybrid transformer*. Essentially it converts a two-way communication over two wires into two one-way, two-wire connections, and it is then usually referred to as a *four-wire communication* (one direction of signal transmission on each pair of a two-pair set). So, while a single two-wire line is adequate for two-way telephone communication over a short distance, as soon as the distance is great enough to require amplification then a conversion to

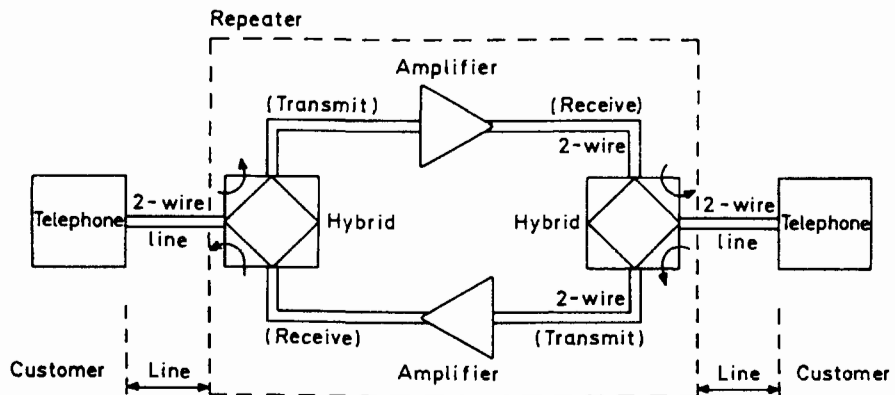


Figure 3.3 A simple telephone repeater system

four-wire communication is called for. Figure 3.3 shows a line between two telephones with a simple telephone repeater in the line. Each repeater consists of two hybrid transformers and two amplifiers (one for each signal direction).

The *amplification* introduced at each repeater has to be carefully controlled to overcome the effects of attenuation, without adversely affecting what is called the *stability* of the circuit, and without interfering with other circuits in the same cable. Repeaters too far apart or with too little amplification would allow the signal current to fade to such an extent as to be subsumed in the electrical *noise* present on the line. Conversely, repeaters that are too close together or have too much amplification, can lead to circuit instability, and to yet another problem known as *crosstalk*.

A circuit is said to be unstable when the signal that it is carrying is over-amplified, causing feedback and even more amplification. This in turn leads to even greater feedback, and so on and on, until the signal is so strong that it reaches the maximum power that the circuit can carry. The signal is now distorted beyond cure and all the listener hears is a very loud singing noise. For the causes of this distressing situation let us look at the simple circuit of Figure 3.4.

The diagram of Figure 3.4 shows a poorly engineered circuit which is electrically unstable. At first sight, the diagram is identical to Figure 3.3. The only difference is that various signal attenuation values (indicated as negative) and amplification values (indicated as positive) have been marked using the standard unit of measurement, the decibel (dB). The problem is that the net gain around the loop is greater than the net loss. Let us look more closely. The hybrid transformer H1 receives the incoming signal from telephone Q and transmits it to telephone P, separating this signal from the one that will be transmitted on the outgoing pair of wires towards Q. Both signals suffer a 3 dB attenuation during this 'line-splitting' process. Adding the attenuation of 1 dB which is suffered on the local access line by the outgoing signal coming from telephone P, the total attenuation of the signal by the time it reaches the output of hybrid H1 is therefore 4 dB. The signal is further attenuated by 5 dB as a result of line loss. Thus the input to amplifier A1 is 9 dB below the strength of the original signal. Amplifier A1 is set to more than make up for this attenuation by boosting the signal by 13 dB, so that at

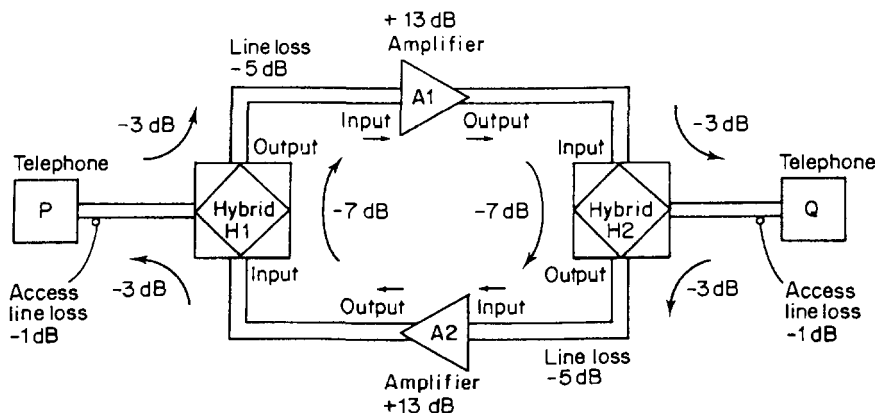


Figure 3.4 An unstable circuit



its output the signal is actually 4 dB louder than it was at the outset. However, by the time the second hybrid loss (in H2) and Q's access line attenuation have been taken into account, the signal is back to its original volume. This might suggest a happy ending, but unfortunately the circuit is unstable. Its instability arises from the fact that neither of the hybrid transformers can actually carry out their line-splitting function to perfection; a certain amount of the signal received by hybrid H2 from the output of amplifier A1 is finding its way back onto the return circuit (Q-to-P).

A well-designed and installed hybrid would give at least 30 dB separation of receive and transmit channels. However, in our example, the hybrid has either been poorly installed or become faulty, and the unwanted retransmitted signal (originated by P but returned by hybrid H2) is only 7 dB weaker at hybrid Hs's point of output than it was at the output from A1, and is then attenuated by 5 dB and amplified 13 dB before finding its way back to hybrid H1, where it goes through another undesired retransmission, albeit at a cost of 7 dB in signal strength. The strength of this signal, which has now entirely 'lapped' the four-wire section of the circuit, is 2 dB less than that of the original signal emanating from telephone P. However, on its first 'lap' it had a strength 4 dB lower than the original. In other words, the re-circulated signal is actually louder than it was on the first lap! What is more, if it goes around again it will gain 2 dB in strength for each lap, quickly getting louder and louder and out of control. This phenomenon is called *instability*. The primary cause is the feedback path available across both hybrids, which is allowing incoming signals to be retransmitted (or 'fed back') on their output. The path results from the non-ideal performance of the hybrid. In practice it is impossible to exactly balance the hybrid's resistance with that of the end telephone handset.

One way to correct circuit instability is to change the hybrids for more efficient (and probably more expensive) ones. This solution requires careful balancing of each telephone handset and corresponding hybrid, and is not possible if the customer lines and the hybrids are on opposite sides of the switch matrix (as would be the case for a two-wire customer local line connected via a local exchange to a four-wire trunk or junction). A cheaper and quicker alternative to the instability problem is simply to reduce the amplification in the feedback loop. This is done simply by adding an attenuating device. Indeed most variable amplifiers are in fact fixed gain amplifiers (around 30 dB) followed immediately by variable attenuators or *pads*. For example, in the case illustrated in Figure 3.4 a reduction in the gain of both amplifiers A1 and A2 to 12 dB will mean that the feedback signal is exactly equal in amplitude to the original. In this state the circuit may just be stable, but it is normal to design circuits with a much greater margin of stability, typically at least 10 dB. For the Figure 3.4 example this would restrict amplifier gain to no more than 7 dB. Under these conditions the volume of the signal heard in telephone Q will be 6 dB quieter than that transmitted by telephone P, but this is unlikely to trouble the listener.

*Crosstalk* is the name given to an overheard signal on an adjacent circuit. It is brought about by electromagnetic induction of an over-amplified signal from one circuit onto its neighbour, and Figure 3.5 gives a simplified diagram of how it happens.

Because the transit amplifier on the circuit from telephone A to telephone B in Figure 3.5 is over-amplifying the signal, it is creating a strong electromagnetic field around the circuit and the same signal is induced into the circuit from P to Q. As a result the user of telephone Q annoyingly overhears the user of telephone A as well as

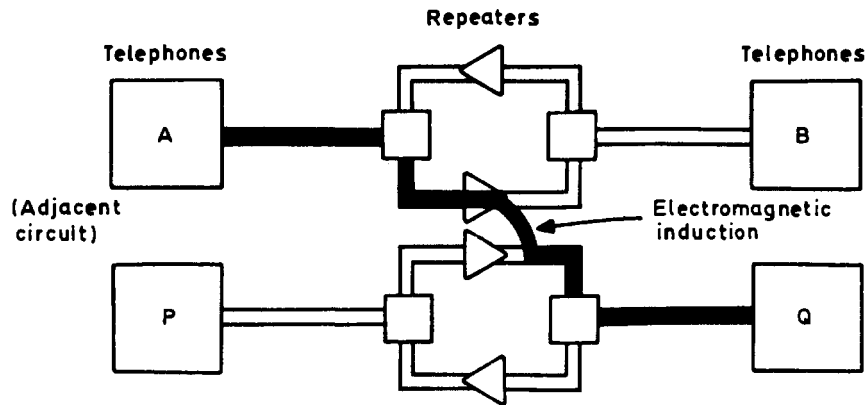


Figure 3.5 Crosstalk. Q hears A

the conversation from P. It is resolved either by turning down the amplifier or by increasing the separation of the circuits. If neither of these solutions is possible then a third, more expensive, option is available, involving the use of specially *screened* or *transverse-screened* cable. In such cable a foil *screen* wrapped around the individual pairs of wires makes it relatively immune to electromagnetic interference.

### 3.4 TWO- AND FOUR-WIRE CIRCUITS

The diagram in Figure 3.3 illustrates a single repeater, used for boosting signals on a two-wire line system. If the wire is a long one, a number of individual repeaters may be required. Figure 3.6 shows an example of a long line in which three amplifiers have been deployed.

Such a system may work quite well but it has a number of drawbacks, the most important of which is the difficulty in maintaining circuit stability and acceptable received signal volume simultaneously; this difficulty arises from the interaction of the various repeaters. An economic consideration is the high cost of the many hybrid transformers that need to be provided. All but the first and last hybrids could be dispensed with if the circuit were wired instead as a four-wire system along the total length of its repeated section, as shown in Figure 3.7. This arrangement reduces the problem of achieving circuit stability when a large number of repeaters are needed, and

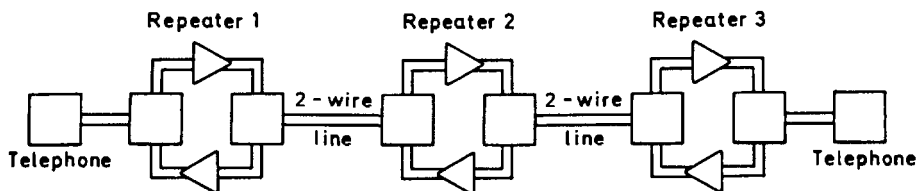


Figure 3.6 A repeated 2-wire circuit

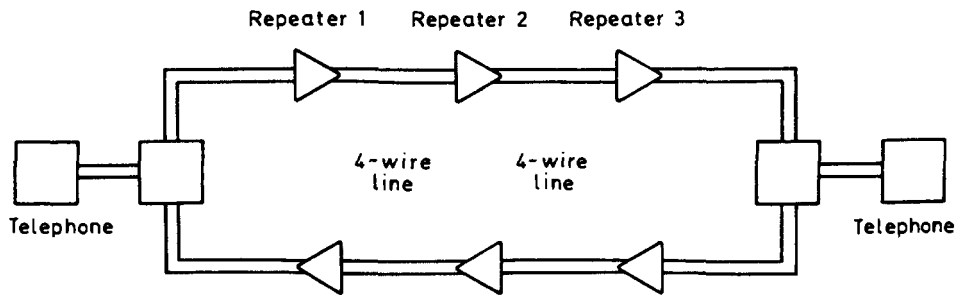


Figure 3.7 A repeatered 4-wire circuit

it eases the maintenance burden. This is why most amplified long haul (i.e. *trunk*) circuits are set up on four-wire transmission lines. Shorter circuits, typically requiring only 1–2 repeaters (i.e. junction circuits), can however make do with two-wire systems.

### 3.5 EQUALIZATION

We have mentioned the need for equalization on long haul circuits, to minimize the signal distortion. Speech and data signals comprise a complex mixture of pure single frequency components, each of which is affected differently by transmission lines. The result of different attenuation of the various frequencies is tonal degradation of the received signal; at worst, the entire high or low-frequency range could be lost. Figure 3.8 shows the relative amplitudes of individual signal frequencies of a distorted and an undistorted signal.

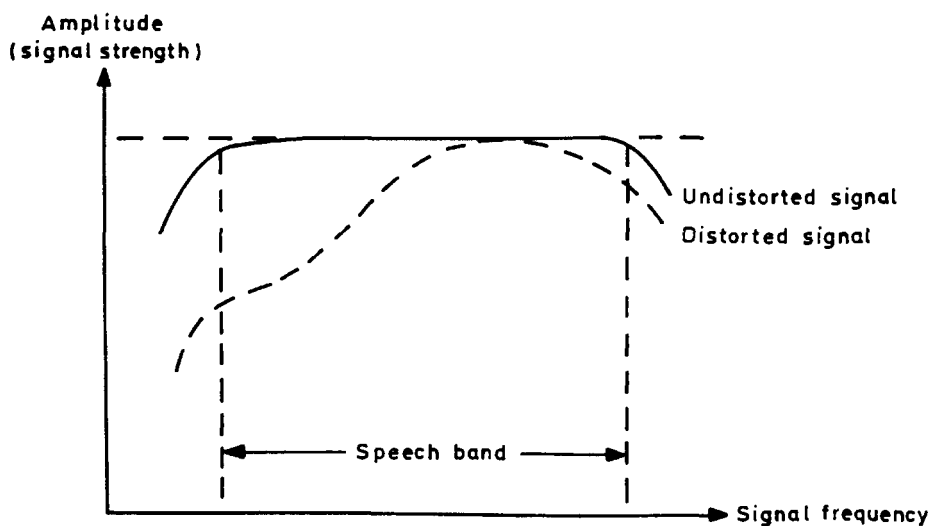


Figure 3.8 Amplitude spectrum of distorted and undistorted signals (attenuation distortion)

In Figure 3.8 the amplitude of the frequencies in the undistorted (received) signal is the same across the entire speech bandwidth, but the high and low frequency signals (high and low notes) have been disproportionately attenuated (lost) in the distorted signal. The effect is known as *attenuation distortion* or *frequency attenuation distortion*.

To counteract this effect, *equalizers* are used, which are circuits designed to amplify or attenuate different frequencies by different amounts. The aim is to 'flatten' the frequency response diagram to bring it in line with the undistorted frequency response diagram. In our example above, the equalizer would need to amplify the low and high frequencies more than it would amplify the intermediate frequencies.

Equalizers are normally included in repeaters, so that the effects of distortion can be corrected all the way along the line, in the same way that amplifiers counteract the effect of attenuation.

### 3.6 FREQUENCY DIVISION MULTIPLEXING (FDM)

When a large number of individual communication channels are required between two points a long distance apart, providing a large number of individual physical wire circuits, one for each channel, can be a very expensive business. For this reason, what is known as multiplexing was developed as a way of making better use of lineplant. Multiplexing allows many transmission channels to share the same physical pair of wires or other transmission medium. It requires sophisticated and expensive *transmission line-terminating equipment (LTE)*, but has the potential for overall saving in cost because the number of wire pairs required between the end points can be reduced.

**Table 3.1** Frequency division multiplex (FDM) hierarchy

Bandwidth name	Consists of	Bandwidth	Usual bandwidth	Number of channels
Channel (1 telephone channel)	24 telegraph subchannels 120 Hz spacing	4 kHz	0–4 kHz	1
Group	12 channels	48 kHz	60–108 kHz	12
Supergroup	5 groups	240 kHz	312–552 kHz	60
Basic hypergroup (also called a 'super mastergroup')	15 supergroups (3 mastergroups)	3.7 MHz (3.6 MHz used) (240 kHz per supergroup with 8 kHz spacing normally between each)	312–4082 kHz (4 MHz line)	900
Basic hypergroup (alternative)	16 supergroups	4 MHz	60–4028 kHz	960
Mastergroup	5 supergroups	1.2 MHz	312–1548 kHz	300
Hypergroup (12 MHz)	9 mastergroups	12 MHz	312–12 336 kHz	2 700
Hypergroup (60 MHz)	36 mastergroups	60 MHz	4404–59 580 kHz	10 800

The method of multiplexing used in analogue networks is called *frequency division multiplex (FDM)*. FDM calls for a single, high grade, four-wire transmission line (or equivalent), and both pairs must be capable of supporting a very large bandwidth. Some FDM cables have a bandwidth as high as 12 MHz (million cycles per second), or even 60 MHz. This compares with the modest 3.1 kHz (thousand cycles per second) required for a single telephone channel. The large bandwidth is the key to the technique, as it subdivides readily into a much larger number of individual small bandwidth channels.

The lowest constituent bandwidth that makes up an FDM system is a single channel bandwidth of 4 kHz. This comprises the 3.1 kHz needed for a normal speech channel, together with some spare bandwidth to create separation between channels on the system as a whole. Various other standard bandwidths are then integral multiples of a single channel. Table 3.1 illustrates this and gives the names of these standard bandwidths.

The overall bandwidth of the FDM transmission line is equal to one of the standard bandwidths (e.g. *supergroup*, *group*) named in Table 3.1, and is then broken down into a number of sub-bandwidths, called *tributaries* in the manner shown in Figure 3.9. The equipment which performs this segregation of bandwidth is called *translating equipment*. Thus *supergroup translating equipment (STE)* subdivides a supergroup into its five component *groups*, and a *channel translating equipment (CTE)* subdivides a group into twelve individual channels.

Not all the available bandwidth needs be broken down into individual 4 kHz channels. If required, some of it can be used directly for large bandwidth applications such as concert grade music or television transmission. In Figure 3.9, two 48 kHz circuits are derived from the supergroup, together with 36 individual telephone

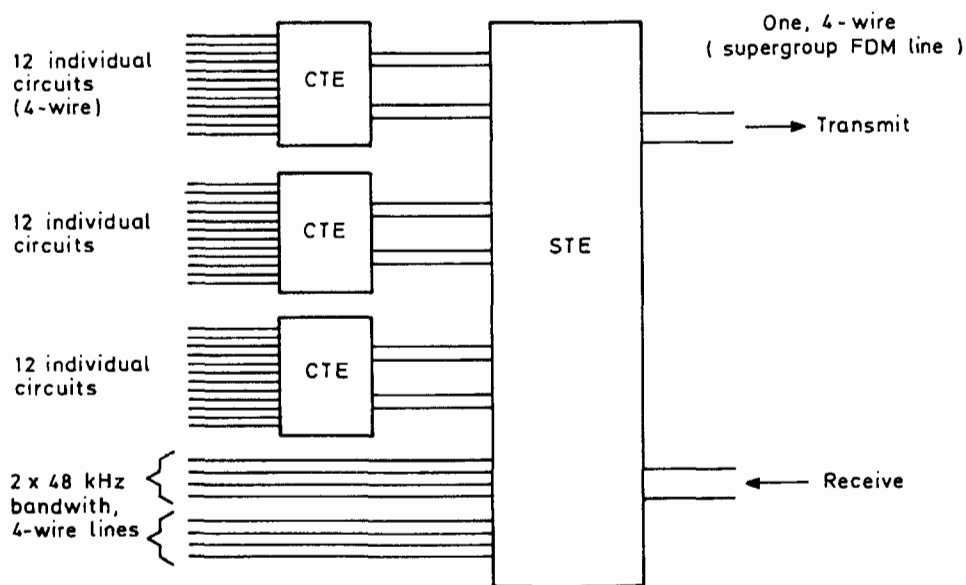


Figure 3.9 Breaking-up bandwidth in FDM

channels. The same principle can be applied at other levels in the hierarchy. Thus, for example, all the supergroups of a *hypergroup* could be broken down into their component groups using HTE and STE. Alternatively, some of the supergroups could be used directly for bandwidth applications of 240 kHz.

Before multiplexing, the audio signals which are to be multiplexed-up are first converted to four-wire transmission, if they are not so already. The signals on each transmit pair are then accurately filtered so that stray signals outside the allocated bandwidth are suppressed. In fact, a telephone channel is filtered to be only 3.1 kHz in bandwidth (of the available 4 kHz). The remaining 0.9 kHz *separation* prevents speech interference between adjacent channels. Groups are filtered to 48 kHz, supergroups to 240 kHz, etc. Each filtered signal is then modulated by a carrier frequency, which has the effect of *frequency shifting* the original signal into another part of the frequency spectrum. For example, a telephone channel starting out in the bandwidth range 300–3400 Hz might end up in the range 4600–7700 Hz. Another could be shifted to the range 8300–11 400 Hz, and so on.

Each component channel of an FDM group is frequency shifted by the CTE to a different bandwidth *slot* within the 48 kHz available; so that in total 12 individual channels may be carried. Likewise, in a supergroup, five already made-up groups of 48 kHz bandwidth are slotted in to the 240 kHz bandwidth by the STE.

The frequency shift is achieved by *modulation* of the component bandwidths with different *carrier signal* frequencies. The frequency of the *carrier signal* which is used to modulate the original signal (or *baseband*) will be equal to the value of the frequency shift required. Each carrier signal must be produced by the translating equipment.

The modulation of a signal in the frequency band 300–3400 Hz using a carrier frequency of 8000 Hz produces a signal of bandwidth from 4600 Hz (8000–3400) to 11 400 Hz (8000 + 3400). The original frequency spectrum of 300–3400 Hz is reproduced in two mirror image forms, called *sidebands*. One sideband is in the range 4600–7700 Hz and the other is in the range 8300–11 400 Hz. Both sidebands are shown in Figure 3.10.

As all the information is duplicated in both sidebands, only one of the sidebands needs to be transmitted. For economy in the electrical power needed to be transmitted to line, it is normal for FDM systems to operate in a *single sideband (SSB)* and

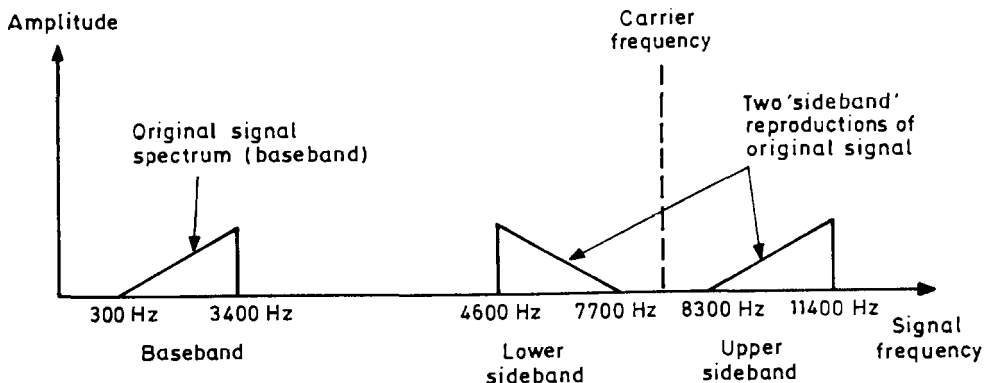


Figure 3.10 Frequency shifting by carrier modulation

*suppressed carrier* mode. The original signal is reconstructed at the receiving end by modulating (mixing) a locally generated frequency. (Note: *single sideband* operation may also be used in radio systems, but the carrier is not suppressed in this case because it is often inconvenient to make a carrier signal generator available at the receiver. When the *carrier* is not suppressed a much simpler and cheaper detector can be used.)

Each *baseband* signal to be included in an FDM system is modulated with a different carrier frequency, the *lower sideband* is extracted for conveyance. It may seem inefficient not to double up the use of carrier frequencies, adopting alternating *upper* and *lower sidebands* of different channels, but by always using the lower sideband we obtain a better overall structure, allowing easier extraction of single channels from higher order FDM systems. The carrier frequencies needed to produce a standard group are thus 64 kHz, 68 kHz, 72 kHz, ..., 108 kHz and the overall structure is as shown in Figure 3.11.

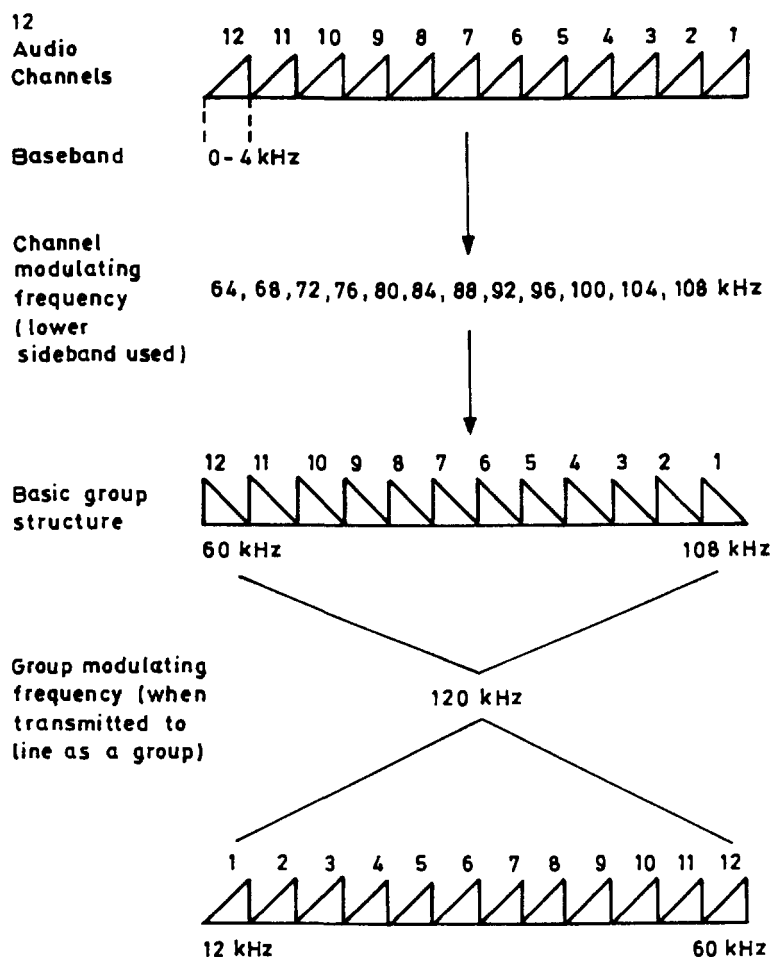


Figure 3.11 The structure of an FDM group

Supergroups and hypergroups may be modulated in a similar fashion, using appropriate carrier frequencies and single sideband operation.

### 3.7 CROSSTALK AND ATTENUATION ON FDM SYSTEMS

There can be as much crosstalk and attenuation in FDM systems as in the single channel or *audio* circuits which we discussed in the first part of the chapter. They require just as much if not more planning, as FDM systems are generally more sensitive and complex.





---

# *Data and the Binary Code System*

---

'Data', a plural noun, is the term used to describe information which is stored in and processed by computers. It is essential to know how such data are represented electronically before we can begin to understand how it can be communicated between computers, communication devices (e.g. facsimile machines) or other data storage devices. As a necessary introduction to the concept of 'digital' transmission, this chapter is devoted to a description of the method of representing textual and numeric information which is called the 'binary code'.

## **4.1 THE BINARY CODE**

*Binary code* is a means of representing numbers. Normally, numbers are quoted in *decimal* (or *ten-state*) code. A single digit in decimal code may represent any of ten different *unit* values, from nought to nine, and is written as one of the figures 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Numbers greater than nine are represented by two or more digits: twenty for example is represented by two digits, 20, the first '2' indicating the number of 'tens', so that 'twice ten' must be added to '0' units, making twenty in all. In a three digit decimal number, such as 235, the first digit indicates the number of 'hundreds' (or 'ten tens'), the second digit the number of 'tens' and the third digit the number of 'units'. The principle extends to numbers of greater value, comprising four or indeed many more digits.

Consider now another means of representing numbers using only a *two-state* or *binary* code system. In such a system a single digit is restricted to one of two values, either zero or one. How then are values of two or more to be represented? The answer, as in the decimal case, is to use more digits. 'Two' itself is represented as the two digits, one-zero or 10. In the binary code scheme, therefore, 10 does not mean 'ten' but 'two'. The rationale for this is similar to the rationale of the decimal number system with which we are all familiar.

In decimal the number one-thousand three-hundred and forty-five is written '1345'. The rationale is

$$(1 \times 10^3) + (3 \times 10^2) + (4 \times 10) + 5$$

The same number in binary requires many more digits, as follows.

$$1345(\text{decimal}) = 10101000001(\text{binary})$$

(binary)	(decimal)
= $1 \times 2^{10}$	1024
+ $0 \times 2^9$	+ 0
+ $1 \times 2^8$	+ 256
+ $0 \times 2^7$	+ 0
+ $1 \times 2^6$	+ 64
+ $0 \times 2^5$	+ 0
+ $0 \times 2^4$	+ 0
+ $0 \times 2^3$	+ 0
+ $0 \times 2^2$	+ 0
+ $0 \times 2$	+ 0
+ 1	+ 1
	= 1345

Any number may be represented in the binary code system, just as any number can be represented in decimal.

All numbers when expressed in binary consist only of 0s and 1s, arranged as a series of binary digits a term which is usually shortened to the jargon *bits*. The *string* of *bits* of a binary number are usually suffixed with a 'B', to denote a binary number. This prevents any confusion that the number might be a decimal one. Thus 41 is written '101001B'.

## 4.2 ELECTRICAL REPRESENTATION AND STORAGE OF BINARY CODE NUMBERS

The advantage of the binary code system is the ease with which binary numbers can be represented electrically. As each digit, or *bit*, of a binary number may only be either 0 or 1, the entire number can easily be transmitted as a series of 'off' or 'on' (sometimes also called *space* and *mark*) pulses of electricity. Thus forty one (101001B) could be represented as on-off-on-off-off-on, or mark-space-mark-space-space-mark. The number could be conveyed between two people on opposite sides of a valley, by flashing a torch, either on or off, say every half second. Figure 4.1 illustrates this simple binary

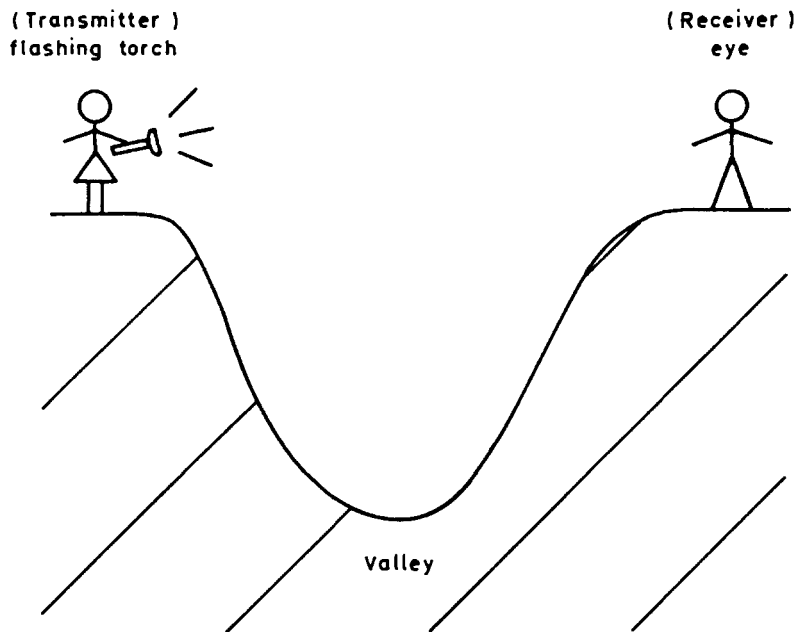


Figure 4.1 A simple binary communication system

communication system in which two binary digits (or bits) are conveyed every second. The speed at which the binary code number, or other information can be conveyed is called the *information conveyance rate* (or more briefly the *information rate*). In this example the rate is two bits per second, which can be expressed also as 2 bit/s.

Figure 4.1 illustrates a means of transmitting numbers, or other binary coded data by a series of 'on' or 'off' electrical states. Transmission of data, however, is not in itself sufficient to permit proper exchange of information between the computers or other equipment located at either end of the line; some method of data storage is needed as well. At the sending end the data have to be stored prior to transmission, and at the receiving end a storage medium is needed not only for the incoming data, but also for the computer programmes required to interpret it.

### 4.3 USING THE BINARY CODE TO REPRESENT TEXTUAL INFORMATION

The letters of the alphabet can be stored and transmitted over binary coded communication systems in the same way as numbers, provided that they have first been *binary-encoded*. There are four notable binary coding systems for alphabetic text. In chronological order these are the *Morse code*, the *Baudot code* (used in Telex, and also known as international alphabet number 2 IA2), *EBCDIC* (*extended binary coded*

A	•	N	— •	0	— — — — —
B	— • • •	O	— — — — —	1	• — — — —
C	— • — •	P	• — — — •	2	• • — — —
D	— • •	Q	— — — • —	3	• • • — —
E	•	R	• — — •	4	• • • • —
F	• • — — •	S	• • •	5	• • • • •
G	— — — •	T	—	6	— • • • •
H	• • • •	U	• • — —	7	— — — • • •
I	• •	V	• • • — —	8	— — — — • •
J	• — — — —	W	• — — — —	9	— — — — •
K	— • — — —	X	— • • — —	'	— — — • • — —
L	• — — • •	Y	— • — — —	.	• — — • — — • — —
M	— — —	Z	— — — • •	?	• • — — — • •

Figure 4.2 The Morse code

*decimal interchange code*), and *ASCII* (*American (national) standard code for information interchange*, also known as international alphabet IA5). These four coding schemes are now described briefly.

## 4.4 MORSE CODE

The Morse code system of dots and dashes was for use over key and lamp telegraph systems. It was also used for signalling by heliograph and by flag. Its two binary elements are *dit* and *da* (*dot* and *dash*). Thirty-nine characters were coded, as shown in Figure 4.2. When transmitting, a short pause is inserted to mark the beginning and end of each character; and between words there is a longer pause.

As an example of morse code, we see from the figure that the word *Morse* is transmitted as 'da da' (pause) 'da da da' (pause) 'dit da dit' (pause) 'dit dit dit' (pause) 'dit' (which would be written as --/---/.-./.../.).

## 4.5 BAUDOT CODE (ALPHABET IA2)

When the telex system was introduced, the *Baudot Code* (now called *the international alphabet IA2*) was developed, with significant advantages over the Morse code for *automatic use*. Each character is represented by five binary elements (usually called *mark* and *space*), but seven elements are transmitted in total, because start (space) and stop (mark) bits are also used. Fixing the number of elements cuts out the need for gaps or pauses between alphabetic characters, and separate words are delimited without a break by introducing the space (SP) character (00100). The regular flow of these signals suits automatic transmitting and receiving devices, and makes them easier to design. Figure 4.3 illustrates the Baudot code. Thus the sequence of seven bits sent to represent the letter A are 'space(start)-mark-mark-space-space-space-mark(stop)'.

Case	Character		Pattern					Case	Character		Pattern				
	(letters)	(figures)	5	4	3	2	1		(letters)	(figures)	5	4	3	2	1
	A	-	0	0	0	1	1		Q	1	1	0	1	1	1
	B	?	1	1	0	0	1		R	4	0	1	0	1	0
	C	:	0	1	1	1	0		S	,	0	0	1	0	1
	D	£	0	1	0	0	1		T	5	1	0	0	0	0
	E	3	0	0	0	0	1		U	7	0	0	1	1	1
	F	!	0	1	1	0	1		V	;	1	1	1	1	0
	G	&	1	1	0	1	0		W	2	1	0	0	1	1
	H		1	0	1	0	0		X	/	1	1	1	0	1
	I	8	0	0	1	1	0		Y	6	1	0	1	0	1
	J	(Bell)	0	1	0	1	1		Z	"	1	0	0	0	1
	K	(	0	1	1	1	1		Shift (figures to letters)		1	1	1	1	1
	L	)	1	0	0	1	0		Shift (letters to figures)		1	1	0	1	1
	M	.	1	1	1	0	0		Space (SP)		0	0	1	0	0
	N	,	0	1	1	0	0		Carriage Return<		0	1	0	0	0
	O	9	1	1	0	0	0		Line Feed		0	0	0	1	0
	P	0	1	0	1	1	0		Blank		0	0	0	0	0

1 = Mark (Punch hole on paper tape)  
0 = Space (No hole)

Figure 4.3 Baudot code (International Alphabet IA2)

The word *Baudot* would thus be transmitted in Baudot code as:

order of	B	A	U	D	O	T
transmit	10011	11000	11100	10010	00011	00001

In passing it is also worth mentioning that the term *Baud* is commonly used in data communications as the unit of rate of signal change on the line transmission medium (the so-called *Baud rate*). Telex networks usually operate at a rate of 50 Baud (50 signal changes per second) and they use the Baudot code. As 5 line state changes (from mark-to-space, space-to-mark, space-to-space or mark-to-mark) are required to convey each character, this produces an *information rate* of 50 divided by 5, that is to say 10 alphabetic characters per second, which incidentally corresponds roughly to ordinary human speech, when we are speaking or reading deliberately.

## 4.6 ASCII

With the advent of semi-conductors and the first computers, 1963 saw the development of a new seven-bit binary code for computer characters. This code encompassed a wider character range, including not only the alphabetic and numeric characters but also a range of new *control characters* which are needed to govern the flow of data in and around the computers. The code, named ASCII (pronounced 'Askey') is now common in computer systems. The letters stand for American (National) Standard Code for Information Interchange. It is also known as International Alphabet number 5 (IA5) and is defined by ITU-T recommendation T.50. Figure 4.4 illustrates it.

Note that the bit numbers 1–7 (top left-hand corner of the table) represent the *least to the most significant bits*, respectively. Each letter, however, is usually written most significant bit (i.e. bit number 7) first. Thus the letter C is written '1000011'. However, to confuse matters further, the least significant bit is transmitted first. Thus the order of transmission for the word 'ASCII' is

(I)	(I)	(C)	(S)	(A)	order
1001001	1001001	1000011	1010011	1000001	of
last				first	transmit

Bits				7	0	0	0	0	1	1	1	1
				6	0	0	1	1	0	0	1	1
4	3	2	1	5	0	1	0	1	0	1	0	1
0	0	0	0	0	NUL	DLE	SP	0	@	P	\	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	0	STX	DC2	"	2	B	R	b	r
0	0	1	1	1	EXT	DC3	#	3	C	S	c	s
0	1	0	0	0	EOT	DC4	\$	4	D	T	d	t
0	1	0	1	1	ENQ	DAK	%	5	E	U	e	u
0	1	1	0	0	ACK	SYN	&	6	F	V	f	v
0	1	1	1	1	BEL	ETB	'	7	G	W	g	w
1	0	0	0	0	BS	CAN	L	8	H	X	h	x
1	0	0	1	1	HT	EM	)	9	I	Y	i	y
1	0	1	0	0	LF	SUB	*	:	J	Z	j	z
1	0	1	1	1	VT	ESC	+	;	K	[	k	{
1	1	0	0	0	FF	FS	,	<	L	\	l	:
1	1	0	1	1	CR	GS	-	=	M	]	m	}
1	1	1	0	0	SO	RS	.	>	N	^	n	~
1	1	1	1	1	SI	US	/	?	O	_	o	DEL

Figure 4.4 The ASCII code (International Alphabet IA5)

The characters, may need not be transmitted directly in the form of the 35 bits shown, but are usually separated by other *control characters*. In particular delimiting bits, so-called *start* and *stop bits* may be used to separate the strings representing individual characters. We shall return to this subject in Chapter 9 when discussing *asynchronous* and *synchronous* transmission methods. Other control characters are also used by modern computer software to control the formatting of text (e.g. in Microsoft's *Word* format).

## 4.7 EBCDIC

EBCDIC or *extended binary coded decimal interchange code* is an extension of ASCII, giving more control characters. It uses an 8-bit representation for each character, as shown in Figure 4.5, and is widely used in IBM computers and compatible machines.

## 4.8 USE OF THE BINARY CODE TO CONVEY GRAPHICAL IMAGES

Besides representing numerical and alphabetical (or textual) characters, the binary code can also be used to transmit pictorial and graphical images as well as complex computer information and formatting.

Pictures are sent as binary information by sending (typically) 8-bit numbers (representing a value between 1 and 256) to represent the particular colour and shade of a miniscule dot, making up a part of the picture. Put all the coloured dots together again in the right pattern (like an impressionist painting) and the picture reappears. This is the principle by which computer images are communicated.

Send a series of pictures, one after the other at a rate of 25 Hz (25 picture frames per second) and you have a television or video signal. Alternatively, if you are willing to trade some of the dynamic quality of the picture (and thus cost), then there is *videotelephony* and *videoconferencing*, a television-like signal sent over telephone-type connections. ITU-T recommendation H.261 lays down a standard for conversion of a video signal to binary code. To illustrate the principles of graphic image transfer using the binary code, we next take the example of facsimile.

## 4.9 FACSIMILE

Facsimile machines work in pairs, separated by some form of transmission link. At the *transmitting* end of the link, one facsimile machine *scans* a piece of paper, and converts the black-and-white image which it sees into a binary-coded stream of data. This data is then transmitted to the receiving facsimile machine, where it is used to produce a black-and-white facsimile reproduction of the original paper image. The working principle of these machines is simple enough, as we may now see.

The image on the original is assumed to be composed of a very large number of tiny dots, arranged in a grid pattern on the paper. Figure 4.6, for example, shows how one word on the paper may be broken down into a grid of dots.



				4	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
				3	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
				2	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
				1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
8	7	6	5																	
0	0	0	0	NUL	SOH	STX	ETX	PF	HT	LC	DEL			SMM	VT	FF	CR	SO	SI	
0	0	0	1	DLE	DC1	DC2	DC3	RES	NL	BS	IL	CAN	EM	CC		IFS	IGS	IRS	IUS	
0	0	1	0	DS	SOS	FS		BYP	LF	EDB	PRE			SM			ENQ	ACK	BEL	
0	0	1	1			SYN		PN	RS	UC	EOT					DC4	NAK		SUB	
0	1	0	0	SP										c	.	<	(	+	1	
0	1	0	1	&										!	\$	*	)	;	7	
0	1	1	0	-	/										,	%	-	>	?	
0	1	1	1											:	#	@	,	=	"	
1	0	0	0		a	b	c	d	e	f	g	h	i							
1	0	0	1		j	k	l	m	n	o	p	q	r							
1	0	1	0			s	t	u	v	w	x	y	z							
1	0	1	1																	
1	1	0	0		A	B	C	D	E	F	G	H	I							
1	1	0	1		J	K	L	M	N	O	P	Q	R							
1	1	1	0			S	T	U	V	W	X	Y	Z							
1	1	1	1	0	1	2	3	4	5	6	7	8	9						□	

Figure 4.5 EBCDIC code





**Figure 4.8** Facsimile terminal. A group 3 facsimile terminal receiving an incoming document. Typically around 25–60 seconds is required to transmit one page, though darker documents may take longer (*Courtesy of British Telecom*)

the original image. Returning to Figure 4.6, we see how the image of the word ‘paper’ has been coded for transmission and subsequent reproduction with the aid of the scanning grid.

Not surprisingly, facsimile machines actually use slightly more sophisticated techniques than those described, but the principles are the same. The purpose of these enhancements to the basic technique is to improve the accuracy and overall speed of transmission and so reduce the time required for conveying each paper sheet.

Any type of image can be conveyed using facsimile machines: typed text, manuscript, pictures and diagrams. The scanning and image reproducing machinery works in the same way for all of them.

Since 1968, when recommendations for CCITT’s first *Group 1* standard apparatus were published, various generations of facsimile machines have been developed. The latest *Group 4* facsimile machines produce extremely high quality pictures, and can transmit a page of A4 in a few seconds, as compared with the six minutes that group 1 apparatus took over the same job.

## 4.10 DIGITAL TRANSMISSION

Nowadays most data and much other information are communicated in one or other of the binary coded forms and the ability to send all sorts of information simultaneously

over a single network has led to *multimedia* communication and computing. This is the term applied to sound, video and data signals transferred simultaneously. As binary coded data are transmitted as a sequence of 'on' or 'off' states, with each 'on' or 'off' representing the value '1' or '0' of consecutive binary digits or *bits*, all information is conveyed essentially as a string of digits, and so the process has acquired the name *digital transmission*. We go on now to assess its considerable advantages over the analogue technique, and how it can be extended to speech and other analogue signals.



---

# *Digital Transmission and Pulse Code Modulation*

---

Following trials in the 1960s, digital telecommunications systems were first widely deployed in the 1970s. Since then, the miniaturization and large scale integration of electronic components and the rapid advance in computer technology have made digital technology the obvious choice for all new telecommunications transmission and switching systems. Now, in the networks of most countries and with wider world international satellite and submarine networks, digital transmission has no rivals. So what exactly is it, and what can we gain by it?

## 5.1 DIGITAL TRANSMISSION

In investigating analogue transmission we found a relationship between bandwidth and overall information carrying capacity, and we described *frequency division multiplexing* (*FDM*). This was a method of reducing the number of physical wires needed to carry a multitude of individual channels between two points, and it worked by sharing out the overall bandwidth of a single set of four-wires (transmit and receive pairs) between all the channels to be carried. We now discuss digital transmission in detail, how it works, and the equivalents of analogue *bandwidth* and channel multiplexing. In contrast with analogue networks, digital networks are ideal for the direct carriage of data, because as the name suggests a digital transmission medium carries information in the form of individual *digits*. Not just any type of digits, but *binary digits* (*bits*) in particular.

The medium used in digital transmission systems is usually designed so that it is only electrically stable in one of two states, equivalent to 'on' (binary value '1') or 'off' (binary value '0'). Thus a simple form of digital line system might use an electrical current as the conveying medium, and control the current to fluctuate between two values, 'current on' and 'current off'. A more recent digital transmission medium using optical fibre (which consists of a hair-thin (50 microns in diameter) strand of glass) conveys the digital signal in the form of 'on' and 'off' light signals, usually generated by some sort of semiconductor electronic device such as a *laser* or a *light emitting diode* (*LED*). Any other medium capable of displaying distinct 'on/off' states could also be used.

For simultaneous two-way (or *duplex*) digital transmission a four-wire (or equivalent) transmission medium is always required. Just as with four-wire analogue transmission (as discussed in Chapter 3) one *pair* of wires (or its equivalent, for example an optical fibre) is used for the *transmit* (*Tx*) direction whereas the other pair is used for the *receive* (*Rx*) direction. These allow the digital pulses to pass in both directions simultaneously without interference. Thus they differ from simple local analogue telephone systems which can be made to work adequately in a two-way mode over only two wires (recall Figure 5.6 of Chapter 2).

The principal advantage of digital over analogue transmission is the improved quality of connection. With only two 'allowed' states on the line ('off' and 'on') it is not all that easy to confuse them even when the signal is distorted slightly along the line by electrical *noise* or *interference* or some other cause. Digital line systems are thus relatively immune to interference. As Figure 5.1 demonstrates, the receiving end only needs to detect whether the received signal is above or below a given threshold value. If the pulse shape is not a 'clean' square shape, it does not matter. Allow the same electrical disturbance to interfere with an analogue signal, and the result would be a low volume crackling noise at the receiving end, which could well make the signal incomprehensible.

To make digital transmission still more immune to noise, it is normal practice to *regenerate* the signals at intervals along the line. A regenerator reduces the risk of misinterpreting the received bit stream at the distant end of a long-haul line, by counteracting the effects of attenuation and distortion, which show up in digital signals as pulse shape distortions. In this corrective function a digital regenerator may be regarded as the equivalent of an analogue repeater.

The process of *regeneration* involves detecting the received signal and recreating a new, clean square wave for onward transmission. The principle is shown in Figure 5.2.

The *regeneration* of digital signals is all that is needed to restore the signal to its original form; there is no need to amplify, equalise or process it in any other way. The fact that the signal can be regenerated exactly is the reason why digital transmission produces signals of such high quality.

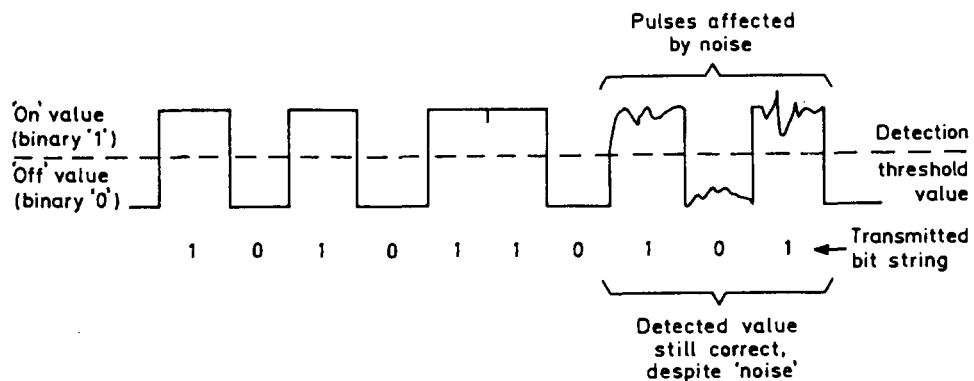


Figure 5.1 Digital signal and immunity to noise

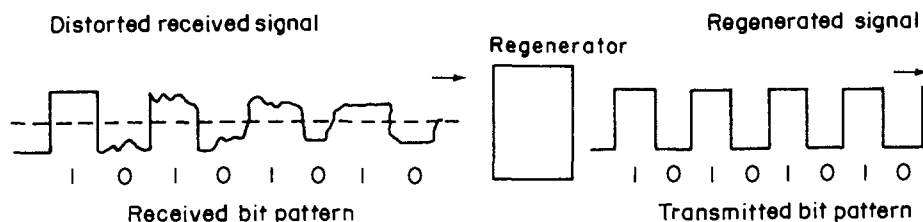


Figure 5.2 The principle of regeneration

Errors at the detection stage can be caused by *noise*, giving the impression of a pulse when there is none. Their likelihood can, however, be reduced by stepping up the electrical power (which effectively increases the overall pulse size or height), and a probability equivalent to one error in several hours or even days of transmission can be obtained (a so-called *bit error rate* or more correctly *bit error ratio (BER)* of 1 errored bit in 1 million bits is noted as  $BER = 1 \times 10^{-6}$ ). This is good enough for speech, but if the circuit is to be used for data transmission it will not be adequate; the error rate may need to be reduced still further, and that requires a special technique using an error checking code. Typical line systems nowadays have BER of  $10^{-9}$ , but with error checking techniques this can be improved to  $10^{-13}$ .

A digital line system may be designed to run at almost any bit speed, but on a single digital circuit it is usually 64 kbit/s. This is equivalent to a 4 kHz analogue telephone channel, as we shall see shortly. The bit speed of a digital line system is roughly equivalent to the bandwidth of an analogue line system; the more information there is to be carried, the greater the required bit speed. Later in the chapter we also discuss how individual 64 kbit/s digital channels can be multiplexed together on a single physical circuit, by a method known as *time division multiplex* (or *TDM*). TDM has the same multiplying effect on the circuit carrying capacity of digital line systems as FDM has for analogue systems.

## 5.2 PULSE CODE MODULATION

You may well ask, how is a speech signal, a TV signal or any other analogue signal to be converted into a form that can be conveyed digitally? The answer lies in a method of analogue to digital signal conversion known as *pulse code modulation*.

*Pulse code modulation (PCM)* functions by converting analogue signals into a format compatible with digital transmission, and it consists of four stages. First, there is the translation of analogue electrical signals into digital pulses. Second, these pulses are coded into a sequence suitable for transmission. Third, they are transmitted over the digital medium. Fourth, they are translated back into the analogue signal (or an approximation of it) at the receiving end. PCM was invented as early as 1939, but it was only in the 1960s that it began to be widely applied. This was mainly because before the day of solid state electronics we did not have the technology to apply the known principles of PCM effectively.



Speech or any other analogue signals are converted into a sequence of binary digits by sampling the signal waveform at regular intervals. At each sampling instant the waveform amplitude is determined and, according to its magnitude, is assigned a numerical value, which is then coded into its binary form and transmitted over the transport medium. At the receiving end, the original electrical signal is reconstructed by translating it back from the incoming digitalized signal. The technique is illustrated in Figure 5.3, which shows a typical speech signal, with amplitude plotted against time. Sampling is pre-determined to occur at intervals of time  $t$  (usually measured in microseconds). The numerical values of the sampled amplitudes, and their 8-bit binary translations, are shown in Table 5.1.

Because the use of decimal points would make the business more complex and increase the bandwidth required for transmission, amplitude is represented by integer values only. When the waveform amplitude does not correspond to an exact integer value, as occurs at time  $4t$  in Figure 5.3, an approximation is made. Hence at  $4t$ , value  $-2$  is used instead of the exact value of  $-2.4$ . This reduces the total number of digits that need to be sent. The signal is reconstituted at the receiving end by generating a stepped waveform, each step of duration  $t$ , with amplitude according to the digit value received. The signal of Figure 5.3 is therefore reconstituted as shown in Figure 5.4.

In the example, the reconstituted signal has a square waveform rather than the smooth continuous form of the original signal. This approximation affects the listener's comprehension to an extent which depends on the amount of inaccuracy involved. The similarity of the reconstituted signal to the original may be improved by

- increasing the sampling rate (i.e. reducing the time separation of samples) to increase the number of points on the horizontal axis of Figure 5.3 at which samples are taken, and/or

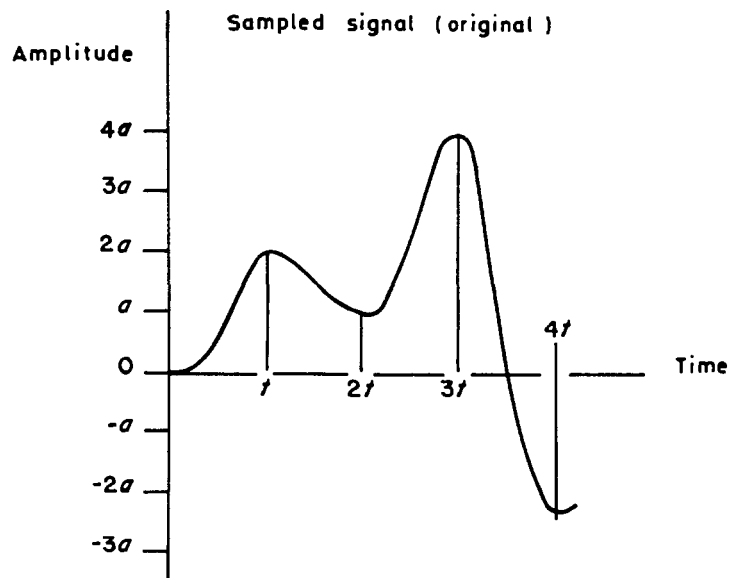


Figure 5.3 Sampling a waveform

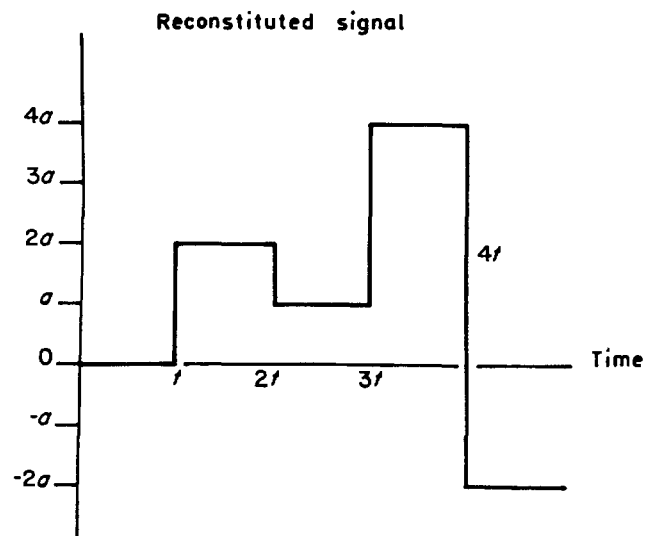
**Table 5.1** Waveform samples from Figure 4.1

Time	Amplitude	Decimal numeric value	8-bit binary translation
0	0	0	00000000
$t$	$2a$	2	00000010
$2t$	$a$	1	00000001
$3t$	$4a$	4	00000100

- increasing the number of *quantization levels* (i.e. wave amplitude levels). The quantization levels are the points on the vertical scale of Figure 5.3.

However, without an infinite sampling rate and an infinite range of quantum values, it is impossible to match an original analogue signal precisely. Consequently an irrecoverable element of *quantization noise* is introduced in the course of translating the original analogue signal into its digital equivalent. The sampling rate and the number of quantization levels need to be carefully chosen to keep this *noise* down to levels at which the received signal is comprehensible to the listener. The snag is that the greater the sampling rate and the greater the number of quantization levels, the greater is the digital bit rate required to carry the signal. Here again a parallel can be found with the bandwidth of an analogue transmission medium, where the greater is the required *fidelity* of an analogue signal, the greater is the bandwidth required.

The minimum acceptable sampling rate for carrying a given analogue signal using digital transmission is calculated according to a scientific principle known as the *Nyquist criterion*, (after the man who discovered it). The criterion states that the sample rate must be at least double the frequency of the analogue signal being sampled. For a

**Figure 5.4** Reconstruction of the waveform of Figure 4.1 from transmitted samples

standard speech channel this equals  $2 \times 4 \text{ kHz} = 8000$  samples per second, the normal bandwidth of a speech channel being 4 kHz.

The number of quantization levels found (by subjective tests) to be appropriate for good speech comprehension is 256. In binary digit (bit) terms this equates to an eight bit number, so that the quantum value of each sample is represented by eight *bits*. The required transmission rate of a digital speech channel is therefore 8000 samples per second, times 8 bits, or 64 kbit/s. In other words a digital channel of 64 kbit/s capacity is equivalent to an analogue telephone channel bandwidth of 4 kHz. This is the reason why the basic digital channel is designed to run at 64 kbit/s.

### 5.3 QUANTIZATION

When the amplitude level at a sample point, does not exactly match one of the quantization levels, an approximation is made which introduces what is called *quantization noise* (also *quantizing noise*). Now, if the 256 quantization levels were equally spaced over the amplitude range of the analogue signal, then the low amplitude signals would incur far greater percentage quantization errors (and thus distortion) than higher amplitude signals. For this reason, the quantization levels are not linearly spaced, but instead are more densely packed around the zero amplitude level, as shown in Figure 5.5. This gives better signal quality in the low amplitude range and a more

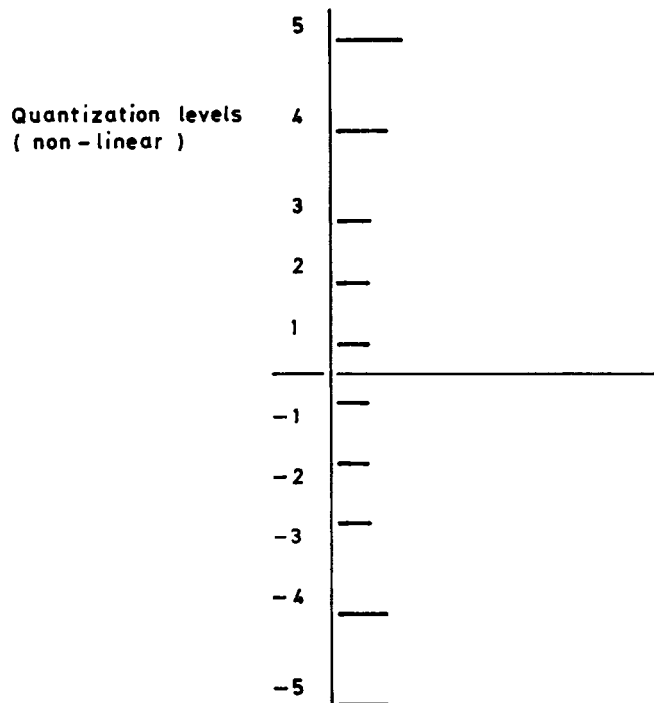


Figure 5.5 Non-linear quantization levels

consistently clear signal across the whole amplitude range. Two particular sets of quantization levels are in common use for speech signal quantization. They are called the *A-Law code* and the  $\mu$ -*Law (Mu-Law) code*. Both have a higher density of quantization levels around the zero amplitude level, and both use an eight bit (256 level) coding technique. They only differ in the actual amplitude values chosen as their respective quantum levels. The A-law code is the European standard for speech quantization and the Mu-law code is used in North America. Unfortunately, because the codes have non-corresponding quantization levels, conversion equipment is required for interworking and this adds to the quantization noise of a connection comprising both A-law and Mu-law digital transmission plant.

Conversion from A-law to  $\mu$ -law (Mu-law) code (or vice versa) amounts to a compromise between the different quantization levels. An 8-bit binary number in one of the codes corresponds to a particular quantization value at a particular sample instant. This 8-bit number is converted into the 8-bit number corresponding to the nearest value in the other quantization code. The conversion is therefore a relatively simple matter of mapping (i.e. converting) between one eight bit value and another.

## 5.4 QUANTIZATION NOISE

Most *noise* heard by the listener on a digital speech circuit is the noise introduced during quantization rather than the result of interfering electromagnetic noise added along the line, and it is minimized by applying the special A-law and  $\mu$ -law codes as already discussed. The total amount of quantization noise (quantizing noise) on a received signal is usually quoted in terms of the number of quantization levels by which the signal differs from the original.

This value is quoted as a number of *quantization distortion units (or qdus)*. Typically the acceptable maximum number of qdus allowed on a complete end-to-end connection is less than 10 (taking into account any A-to- $\mu$  law code conversion or other signal processing undertaken on the connection). Another possible type of speech processing is the technique of speech compression, and we shall see in Chapter 38 that the overall bit rate can be reduced by speech compression, at the cost of some increase in quantization noise.

Quantization noise only occurs in the presence of a signal. Thus, the quiet periods during a conversation are indeed quiet. This 'quietness' gives an improved subjective view of the quality of digital transmission.

## 5.5 TIME-DIVISION MULTIPLEXING

As digital transmission is by discrete pulses and not continuous signals, it is possible for the information of more than one 64 kbit/s channel to be transmitted on the same path, so long as the transmission rate (i.e. *bit rate*) is high enough to carry the bits from a number of channels. In practice this is done by *interleaving* the pulses from the various channels in such a way that a sequence of eight pulses (called a *byte* or an *octet*) from the first channel is followed by a sequence of eight from the second channel, and so on. The principle is illustrated in Figure 5.6, in which the TDM equipment could be imagined to

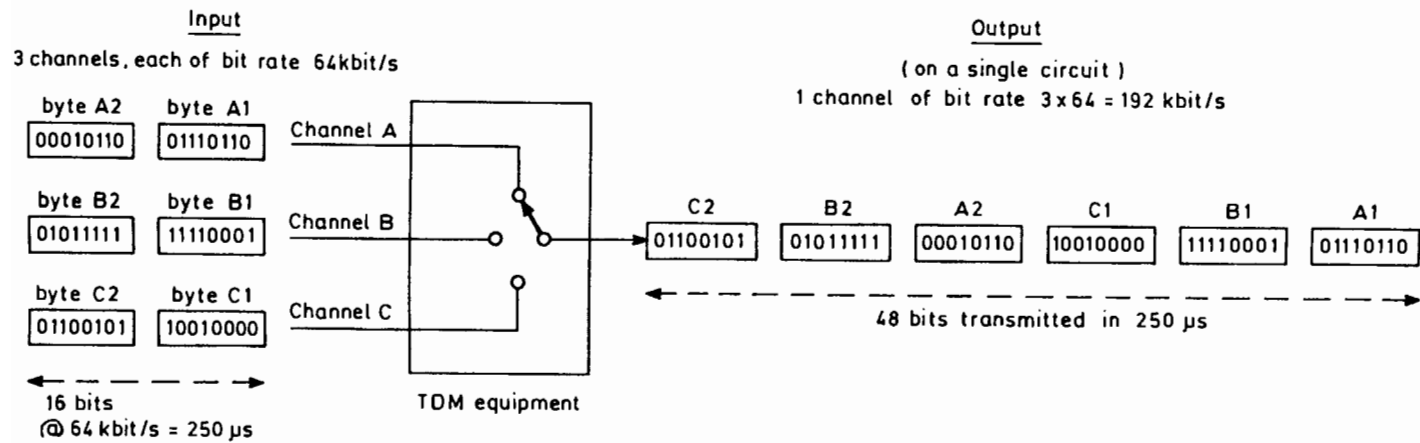


Figure 5.6 The principle of time division multiplexing (TDM)

be a rotating switch, picking up in turn 8 bits (or 1 *byte*) from each of the input channels A, B and C in turn. Thus the output bit stream of the TDM equipment is seen to comprise, in turn, byte A1 (from channel A), byte B1 (from channel B), byte C1 (from channel C), then, cycling again, byte A2, byte B2, byte C2 and so on. Note that a higher bit rate is required on the output channel, to ensure that all the incoming data from all three channels can be transmitted onward. As  $3 \times 2 = 6$  bytes of data are received on the incoming side during a time period of 250 microseconds (1 byte on each channel every 125 microseconds), all of them have to be transmitted on the outgoing circuit in an equal amount of time. As only a single channel is used for output, this implies a rate of  $6 \times 8 = 48$  bytes in 250 microseconds, i.e. 192 kbit/s. (Unsurprisingly, the result is equal to  $3 \times 64$  kbit/s). Thus, in effect the various channels 'time-share' the outgoing transmission path. The technique is known as *time-division multiplexing (TDM)*.

TDM can either be carried out by interleaving a byte (i.e. 8 bits) from each *tributary* channel in turn, or it can be done by single bit interleaving. Figure 5.6 shows the more common method of byte interleaving. The use of the TDM technique is so common on digital line systems that physical circuits carrying only 64 kbit/s are extremely rare, so that digital line terminating equipment usually includes a multiplexing function. Figure 5.7 shows a typical digital line terminating equipment, used to convert between a number of individual analogue channels (carried on a number of individual physical circuits) and a single digital bit stream carried on a single physical circuit. The equipment shown is called a primary multiplexor. A *primary multiplexor (PMUX)* contains an analogue to digital conversion facility for individual telephone channel conversion to 64 kbit/s, and additionally a time division multiplex facility. In Figure 5.7 a PMUX of European origin is illustrated, converting 30 analogue channels into A-law encoded 64 kbit/s digital format, and then time division multiplexing all of these 64 kbit/s channels into a single 2.048 Mbit/s (E1) digital line system. (2.048 Mbit/s

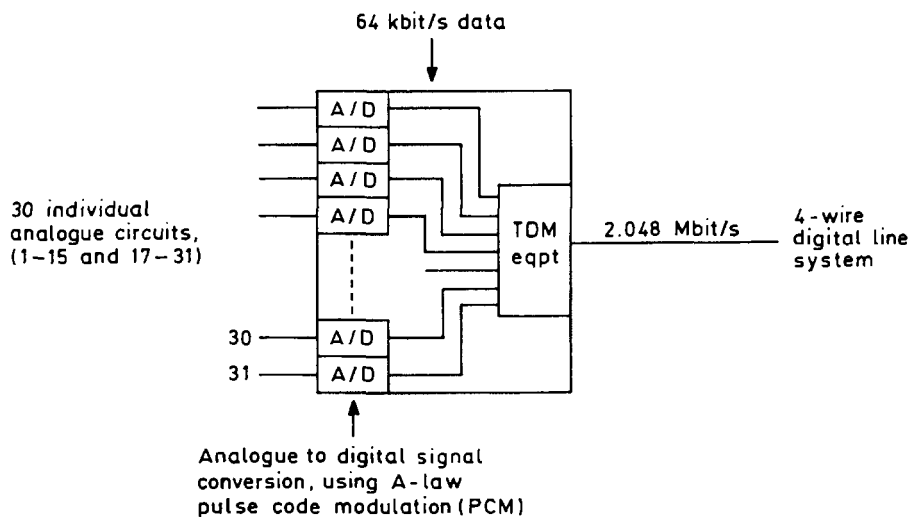


Figure 5.7 European primary multiplexor

is actually equal to  $32 \times 64$  kbit/s, but channels '0' and '16' of the European system are generally used for purposes other than carriage of information.)

We could equally well have illustrated a North American version PMUX. The difference would have been the use of  $\mu$ -law encoding and the multiplexing of 24 channels into a 1.544 Mbit/s transmission format also called a *T-span*, a *T1* line or a *DS1*. (T = Transmission, DS = digital line system) ( $1.544 \text{ Mbit/s} = 24 \times 64 \text{ kbit/s} + 8 \text{ kbit/s}$ ).

The transmitting equipment of a digital line system has the job of multiplexing the bytes from all the constituent channels. Conversely, the receiving equipment must disassemble these bytes in precisely the correct order. This requires synchronous operation of transmitter and receiver, and to this end particular patterns of pulses are transmitted at set intervals, so that alignment and synchronism can be maintained. These extra pulses are sent in channel 0 of the European 2 Mbit/s digital system, and in the extra 8 kbit/s of the North American 1.5 Mbit/s system.

## 5.6 HIGHER BIT RATES OF DIGITAL LINE SYSTEMS

The number of channels multiplexed on a carrier depends on the overall rate of bit transmission on the line. Given that each channel must be transmitted at 64 kbit/s, the overall bit speed is usually related to an integer multiple of 64 kbit/s. There are three basic hierarchies of transmission rates which have been standardized for international use, but these extend to higher bit rates than the 2.048 Mbit/s and 1.544 Mbit/s versions so far discussed.

The ITU-T (formerly CCITT, Consultative Committee for International Tele-phones and Telegraphs), CEPT (European Committee for Posts and Telecommunications) and ETSI (European Telecommunications Standards Institute) have standardized 2.048 Mbit/s as the primary digital bandwidth (E1 line system) and A-Law as the speech encoding algorithm. This has 32 channels, 30 for speech and two for alignment synchronization and signalling, more of which we shall discuss later in the chapter.

Higher transmission rates in the European digital hierarchy are attained by interleaving a number of 2 Mbit/s systems as illustrated in Figure 5.8. The standardized rates are:

2.048 Mbit/s,	referred to as E1 or	2 Mbit/s
8.448 Mbit/s,	referred to as E2 or	8 Mbit/s ( $4 \times 2 \text{ Mbit/s}$ )
34.368 Mbit/s,	referred to as E3 or	34 Mbit/s ( $4 \times 8 \text{ Mbit/s}$ )
139.264 Mbit/s,	referred to as E4 or	140 Mbit/s ( $4 \times 34 \text{ Mbit/s}$ )
564.992 Mbit/s,	referred to as E5 or	565 Mbit/s ( $4 \times 140 \text{ Mbit/s}$ )

Multiplexing equipment is available for any of the rate conversions, as Figure 5.8 shows.

In the second ITU-T standard (which currently predominates in North America), a different multiplex hierarchy is recommended and is shown below. This is based on a basic block of  $24 \times 64$  kbit/s channels plus 8 kbit/s for *framing*, giving a bitrate of 1.544 Mbit/s (T1 line system).  $\mu$ -Law encoding is used for pulse code modulation of speech signals. The principles of multiplexing, however, are largely the same, and diagrams similar to Figure 5.8 could have been drawn.

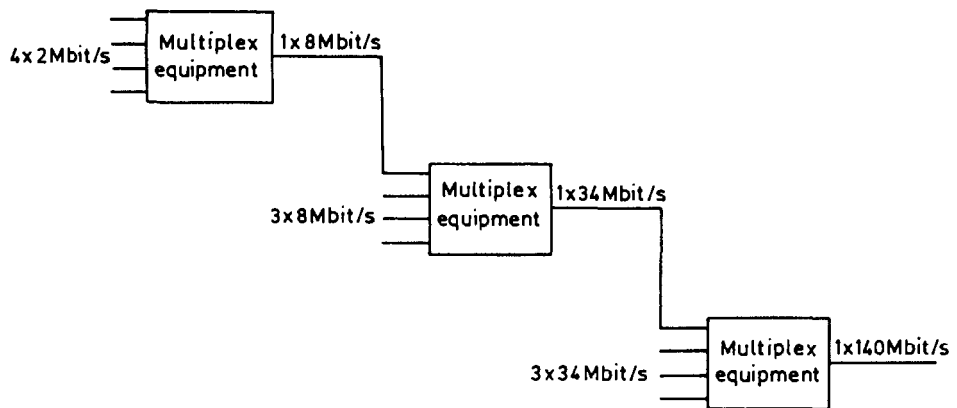


Figure 5.8 European digital multiplex hierarchy

DS0 = 64 kbit/s, the basic channel

T1 or DS1 = 1.544 Mbit/s (this is called the T-span, T1 or DS1 system)

T2 or DS2 = 6.312 Mbit/s ( $4 \times 1.5$  Mbit/s)

T3 or DS3 = 44.736 Mbit/s ( $7 \times 6$  Mbit/s), sometimes referred to as 45 Mbit/s

DS4 = 139.264 Mbit/s ( $3 \times 45$  Mbit/s)

278.176 Mbit/s ( $6 \times 45$  Mbit/s)

In the third system, predominant in Japan, yet another hierarchy is used, though there is some overlap with the North American system.  $\mu$ -Law encoding is applied to speech pulse code modulation.

DS0 = 64 kbit/s, the basic channel

J1 = 1.544 Mbit/s (this is called the T-span, T1 or DS1 system)

J2 = 6.312 Mbit/s ( $4 \times 1.5$  Mbit/s)

J3 = 32.064 Mbit/s ( $5 \times 6$  Mbit/s)

J4 = 97.728 Mbit/s ( $3 \times 32$  Mbit/s)

The various hierarchies are incompatible at all levels (including the basic speech channel level, on account of the different quantization code used by the A and  $\mu$ -law PCM algorithms). Interworking equipment is therefore required for international links between administrations employing the different hierarchies. In general, this interworking is undertaken in the country which uses the 1.544 Mbit/s standard.

Before we leave the subject of nomenclature for digital stream bitrates, we should also mention the terminology used particularly for high speed video channels. These are H0 (384 kbit/s), H11 (1536 kbit/s) and H12 (1920 kbit/s). These correspond, respectively, to  $6 \times 64$  kbit/s,  $24 \times 64$  kbit/s and  $30 \times 64$  kbit/s. All three bitrates may be supported by an E1 line system, only the first two from a T1 or J1 system.

## 5.7 DIGITAL FRAME FORMATTING AND 'JUSTIFICATION'

As we noted earlier in the chapter, it is common in a 2.048 Mbit/s system to use only thirty 64 kbit/s channels (representing only 1.920 Mbit/s) for actual carriage of



information. This leaves an additional 128 kbit/s bit rate available. Similarly, in the 1.544 Mbit/s system, the bit rate required to carry twenty-four 64 kbit/s channels is only 1.536 Mbit/s, and 8 kbit/s are left over. The burning question: what becomes of this spare capacity? The answer: it is used for synchronization and signalling functions.

Consider a 2.048 Mbit/s bit stream, and in particular the bits carried during a single time interval of 125 microseconds. During a period of 125 microseconds a single sample of 8 bits will have been taken from each of the 30 constituent or *tributary* channels making up the 2.048 Mbit/s bit stream. These are structured into an imaginary *frame*, each frame consisting of 32 consecutive *timeslots*, one *timeslot* of 8 bits for each tributary channel. Overall the *frame* represents a snapshot image, one sample of 8 bits taken from each of the 30 channels, at a frequency of one frame every 125  $\mu$ s. Each frame is structured in the same way, so that the first timeslot of eight bits holds the eight-bit sample from tributary channel 1, the second timeslot the sample from channel 2, and so on. The principle is shown in Figure 5.6. It is very like a single frame of a movie film; the only thing missing is the equivalent of the film perforations which allow a movie projector to move each 'freeze-frame' precisely. This 'film perforations' function is in fact performed by the first timeslot in the frame. It is given the name timeslot '0'. It carries so-called *framing* and *synchronization* information, providing a clear mark to indicate the start of each *frame* and an indication of the exact bit transmission speed. The principle is shown in Figure 5.9, which illustrates a single frame of 32 *timeslots*.

Timeslot 0 then provides a mark for framing. Timeslots 1-15 and 17-31 are used to carry the tributary channels. That leaves timeslot 16 which, as we shall see, is used for *signalling*.

We cannot leave timeslot zero, without briefly discussing its *synchronization* function which serves to keep the line system bit rate running at precisely the right speed. Consider a wholly digital network consisting of three digital exchanges A, B and C interconnected by 2.048 Mbit/s digital transmission links, as shown in Figure 5.10, with end users connected to exchanges A and C.

Each of the exchanges A, B and C in Figure 5.10 will be designed to input and receive data from the digital transmission links A-B and B-C at 2.048 Mbit/s. What happens if link A-B actually runs at 2 048 000 bit/s, while link B-C runs at 2 048 001 bit/s? This, or something even worse, could quite easily happen in practice if we did not take *synchronization* steps to prevent it. In the circumstances shown, the bit stream received by exchange B from exchange A is not fast enough to fill the outgoing timeslots on the link from B to C correctly, and a *slip* of 1 'wasted' bit will occur once per second. Conversely, in the direction from C to A via B, unsent bits will gradually be stored up by exchange B at a rate of 1 extra unsent bit per second, because the exchange is unable to transmit the bits to A as fast as it is receiving them from exchange C. Ultimately bits are lost when the store in exchange B overflows. Neither *slip* nor *overflow* of bits is desirable, so networks are normally designed to be synchronous at the 2.048 Mbit/s level, in other words are controlled to run at exactly the same speed. Actually, they run plesiochronously. Some of the bits in timeslot zero of a 2.048 Mbit/s line system are used to try to maintain the synchronization, but systems are not firmly locked in-step. Each system instead runs from its own *clock*. The synchronization bits adjust the speed of the clock (faster or slower) to keep it in step with other systems, but as there is more than one clock in the network, there is still a discrepancy in the synchronization of the

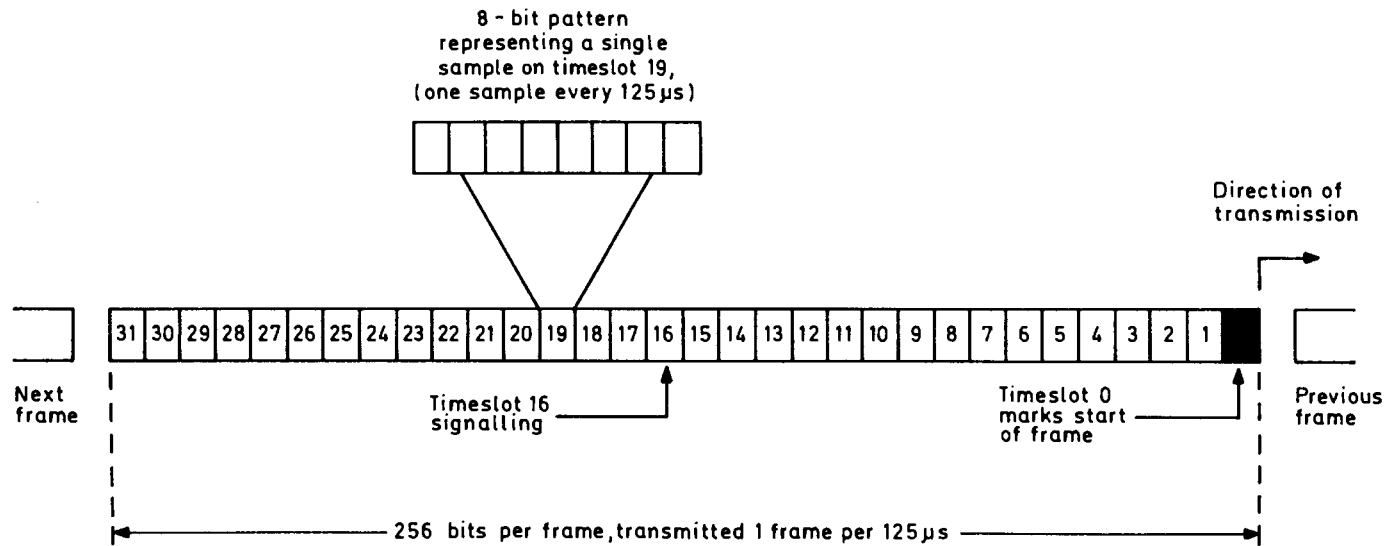


Figure 5.9 2.048 Mbit/s digital frame format

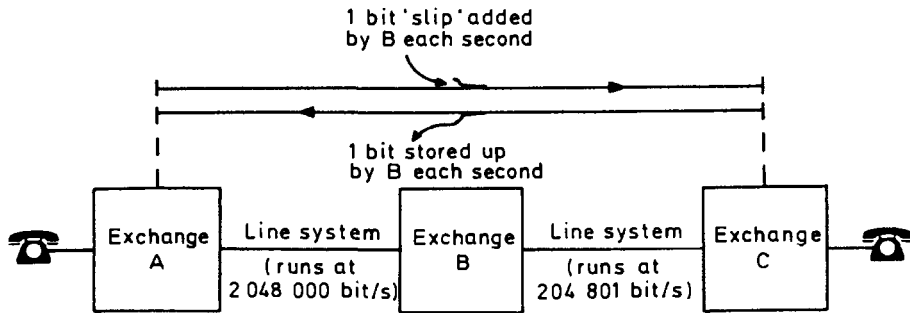


Figure 5.10 The need for synchronization

various systems, hence the term *plesio*-chronous. The same *plesiochronous* functions of *framing* and *synchronization* are carried out by the surplus 8 kbit/s capacity of the 1.544 Mbit/s digital line system.

Because the speed of the system is actually slightly greater than the sum of the *tributary* input channels, extra *dummy* bits (also called *justification bits* or *stuffing*, leading to the term *bit stuffing*) need to be added to the stream. These can be removed at the receiving multiplexor. Should one of the tributaries be supplying data (bits) slightly faster than its nominated rate, this can be accommodated by the multiplexor by substituting some of the *justification bits* with user data. Similarly, if the rate of the input channel is slightly too slow, more justification bits (J) can be added (Figure 5.11).

Now let us consider the function of timeslot 16 in a 2.048 Mbit/s digital line system. This *timeslot* is usually reserved for carrying the signalling information needed to set up the calls on the 30 *user channels*. The function of signalling information is to convey the intended destination of a call on a particular channel between one exchange and the next.

From the above, we see that the maximum usable bit rate of a 2.048 Mbit/s system is  $30 \times 64$  kbit/s or 1.920 Mbit/s. In occasional circumstances, however, this can be increased to 1.984 Mbit/s when the signalling channel (timeslot 16) is not needed.

When required on a 1.544 Mbit/s line system, a signalling channel can be made available by *stealing* a small number of bits (equivalent to 4 kbit/s) from one of the tributary

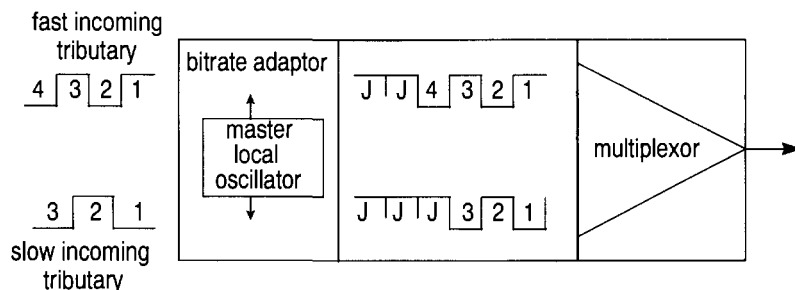


Figure 5.11 The process of justification (plesiochronous digital hierarchy)

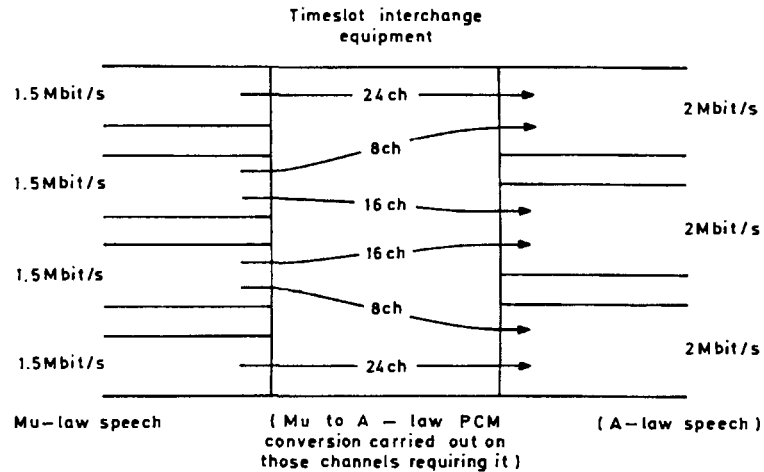
channels (thereby reducing the capacity of that particular channel to 60 kbit/s). Alternatively, and nowadays more usually, one whole 64 kbit/s channel may be dedicated for signalling use. The method of *stealing* bits to create a 4 kbit/s signalling channel is known in North America as *robbed bit signalling*. It is only permissible to *rob* the bits from a voice channel and not from a data channel. Robbing a small number of bits from a voice channel is permissible because the quality lost thereby is almost imperceptible to a human telephone listener. Robbing bits from a channel which is carrying data, however, will result in quite unacceptable data corruption. As users of telephone networks have become accustomed to transmitting data signals (e.g. fax), *robbed bit signalling* has become less acceptable. Where a signalling channel is required on a 1.544 Mbit/s digital line system carrying only data circuits, a whole channel should be dedicated to signalling. Such a dedicated signalling channel is necessary to create SS7 signalling links between computer-controlled telephone exchanges.

To return to the two different bit rate hierarchies, observant readers may have noticed that the higher bit rates of both hierarchies are not exact integer multiples of the basic 2.048 Mbit/s and 1.544 Mbit/s tributaries. Instead, some extra *framing* bits have been added once again at each hierarchical level. These are provided for the same framing reasons as have already been described in connection with the 2.048 Mbit/s line system, and illustrated in Figure 5.9. However, unlike their 2 Mbit/s or 1.5 Mbit/s tributaries, synchronization of higher bit-rate line systems in PDH (plesiochronous digital hierarchy) is not usually undertaken. Instead, higher order systems are generally allowed to *free-run*. The extra bits allow free running, as a slightly higher bit rate is available than the tributaries can feed. The higher bit rate ensures that there is no possibility of bits building up between the tributaries and the higher bit rate line system itself. Instead there will always be a few bits to spare. The benefit is that the need for synchronization at the higher bit rate is avoided, but the 'penalty' is the complicated frame structure needed at the higher rates of the hierarchy. The same problem faces users of the 1.544 Mbit/s hierarchy.

## 5.8 INTERWORKING THE 2MBIT/S AND 1.5MBIT/S HIERARCHIES

Interworking of digital line systems running in the 2 Mbit/s hierarchy and 1.5 Mbit/s hierarchy is relatively straightforward, given the availability of proprietary equipment for the conversion. At its simplest, the 24 channels of a 1.5 Mbit/s system can be carried within a 2 Mbit/s system, effectively wasting the remaining capacity of the 2 Mbit/s system. Alternatively, a 2 Mbit/s system can be entirely carried on two 1.5 Mbit/s systems, wasting 16 channels of the second 1.5 Mbit/s system. More efficiently, however, four 1.5 Mbit/s systems fit almost exactly into three 2 Mbit/s systems or vice versa. (They appear to fit exactly, but usually some bits are taken for separating the different frames so that the efficiency is reduced slightly).

The interworking of one digital hierarchy into the other needs only involve mapping the individual 8-bit timeslots from one hierarchy into corresponding timeslots in the other. The technique is called *timeslot interchange*. The only complication is when the 8-bit patterns in the timeslots are not simple data patterns (data patterns should be mapped across unchanged) but when they are sample patterns corresponding to A or



**Figure 5.12** Timeslot interchange between 1.5 Mbit/s and 2 Mbit/s

$\mu$ -law pulse code modulated speech. In this instance, an A- to  $\mu$ -law (Mu-law) speech conversion is also required at the 2 Mbit/s to 1.5 Mbit/s interworking point.

Figure 5.12 illustrates a typical timeslot interchange between four 1.5 Mbit/s and three 2 Mbit/s digital line systems.

Note that the timeslot interchange equipment in Figure 5.12 is also capable of  $\mu$ - to A-law conversion (and vice versa). This has to be available on each of the channels, but is only employed when the channel is carrying a speech call. When there are consecutive speech and data calls on the same channel, the  $\mu$ - to A-law conversion equipment will have to be switched on for the first call and off for the second. Some means is therefore needed of indicating to the timeslot interchange equipment whether at any particular time it is carrying a speech or a data call. Alternatively, particular channels could be pre-assigned either to speech or data use. In this case the  $\mu$ - to A-law conversion equipment will be permanently on and permanently off, respectively.

## 5.9 SYNCHRONOUS FRAME FORMATTING

Modern line systems, specifically *SDH* (*synchronous digital hierarchy*) and *SONET* (*synchronous optical network*) demand synchronous operation of all the line systems within a network (i.e. all must operate using the same clock). In return, it offers a simpler and more regular frame structure of 2 Mbit/s and 1.5 Mbit/s tributaries within the higher bitrates (multiples of 155 Mbit/s). As we will discuss in Chapter 13 this gives much greater flexibility in management and administration of the system. A further significant benefit is their support of both 2 Mbit/s and 1.5 Mbit/s based hierarchies, creating an easy migration path for worldwide standardization. Table 5.2 lists the basic bitrate hierarchies of both SDH and SONET. A more detailed analysis follows in Chapter 13.

**Table 5.2** SDH (synchronous digital hierarchy) and SONET (synchronous optical network)

North American SONET	Carried Bitrate Mbit/s	SDH
VT 1.5	1.544	VC-11
VT 2.0	2.048	VC-12
VT 3.0	3.152	—
VT 6.0	6.312	VC-21
—	8.448	VC-22
—	34.368	VC-31
—	44.736	VC-32
—	149.76	VC-4
STS-1 (OC-1)	51.84	—
STS-3 (OC-3)	155.52	STM-1
STS-6 (OC-6)	311.04	—
STS-9 (OC-9)	466.56	—
STS-12 (OC-12)	622.08	STM-4
STS-18 (OC-18)	933.12	—
STS-24 (OC-24)	1244.16	—
STS-36 (OC-36)	1866.24	—
STS-48 (OC-48)	2488.32	STM-16
STS-96 (OC-96)	4976.64	—
STS-192 (OC-192)	9953.28	STM-64

## 5.10 LINE CODING

The basic information to be transported over any digital line system, irrespective of its hierarchical level, is a sequence of ones and zeros, also referred to as *marks* and *spaces*. The sequence is not usually sent directly to line, but is first arranged according to a *line code*. This aids intermediate regenerator timing and distant end receiver timing, maximizing the possible regenerator separation and generally optimizing the operation of the line system. The potential problem is that if either a long string of 0s or 1s were sent to line consecutively then the line would appear to be either permanently 'on' or permanently 'off'. Effectively a direct current condition is transmitted to line. This is not advisable for two reasons. First the power requirement is increased and the attenuation is greater for direct as opposed to alternating current. Second, any subsequent devices in the line cannot distinguish the beginning and end of each individual bit. They cannot tell if the line is actually still 'alive'. The problem gets worse as the number of consecutive 0s or 1s increases. Line codes therefore seek to ensure that a minimum frequency of line state changes is maintained.

Figure 5.13 illustrates the most commonly used line codes. Generally they all seek to eliminate long sequences of 1s or 0s, and try to be *balanced* codes, i.e. producing a net zero direct current voltage (thus the three state codes CMI and HDB3 try to negate positive pulses with negative ones). This reduces the problems of transmitting power across the line. The more sophisticated modern techniques simultaneously seek to

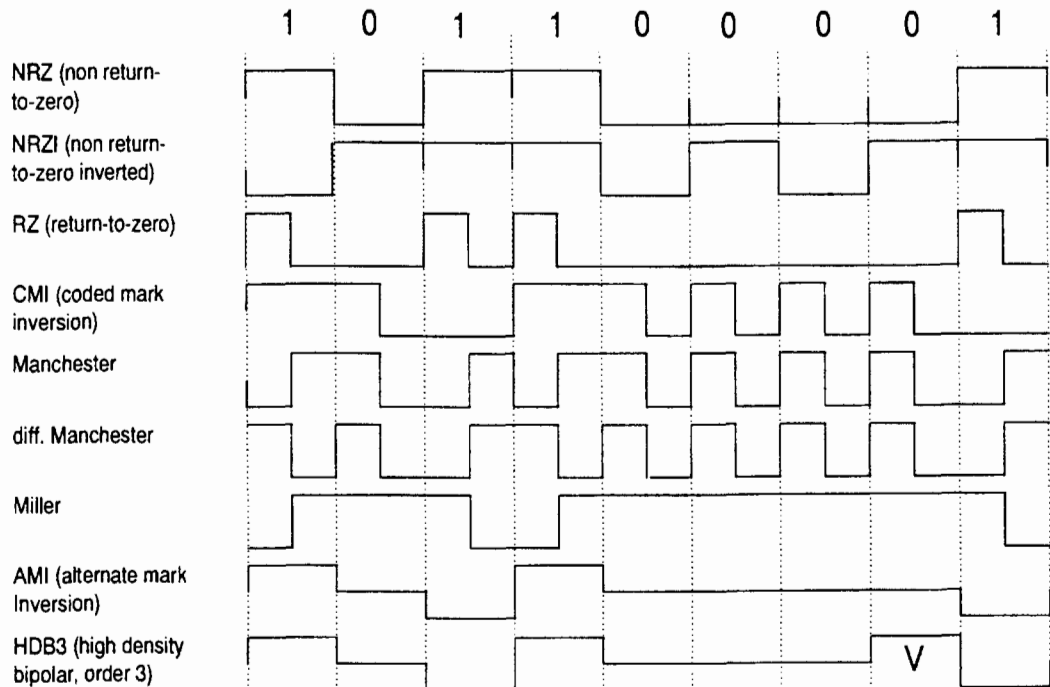


Figure 5.13 Commonly used line codes for digital line systems

reduce the frequency of line state changes (the *baud rate*) so that higher user bitrates can be carried.

The simplest line code illustrated in Figure 5.13 is a *non-return to zero (NRZ)* code in which 1 = on and 0 = off. This is perhaps the easiest to understand.

In *NRZI (non-return-to-zero inverted)* it is the presence or absence of a *transition* which represents a 0 or a 1. This retains the relative simplicity of the code but may be advantageous where the line spends much of its time in an 'idle' mode in which a string of 1s or 0s may be sent. Such is the case, for example, between an *asynchronous terminal* and a *host computer* or *cluster controller*. NRZI is used widely by the IBM company for such connections.

A *return-to-zero (RZ)* code is like NRZ except that *marks* return to zero midway through the bit period, and not at the end of the bit. Such coding has the advantage of lower required power and constant *mark* pulse length in comparison with basic NRZ. The length of the pulse relative to the total bit period is known as the *duty cycle*. Synchronization and timing adjustment can thus be achieved without affecting the *mark* pulse duration.

A variation of the NRZ and RZ codes is the *CMI (coded mark inversion)* code recommended by ITU-T. In *CMI*, a 0 is represented by the two signal amplitudes A1 and A2 which are transmitted consecutively, each for half the bit duration. 1s are sent as full bit duration pulses of one of the two line signal amplitudes, the amplitude alternating between A1 and A2 between consecutive *marks*.

In the *Manchester code*, a higher pulse density helps to maintain synchronization between the two communicating devices. Here the transition from high-to-low represents a 1 and the reverse transition (from low-to-high) a 0. The *Manchester code* is used in *ethernet LANs* (Chapter 19).

In the *differential Manchester code* a voltage transition at the bit start point is generated whenever a binary 0 is transmitted but remains the same for binary 1. The IEEE 802.5 specification of the *token ring LAN* (Chapter 19) demands differential Manchester coding. In both the *Manchester* and *differential Manchester* coding schemes, two extra *coding violation* symbols exist, J and K. These allow for *bit stuffing* as previously discussed.

In the *Miller code*, a transition either low-to-high or high-to-low represents a 1. No transition means a 0.

The *AMI* (*alternate mark inversion*) and *HDB3* (*high density bipolar*) codes defined by ITU-T (recommendation G.703) are both three-state, rather than simple two-state (on/off) codes. In these codes, as can be seen in Figure 5.13, the two extreme states are used to represent *marks*, and the mid state is used to represent *spaces*. The three states could be positive and negative values, with a mid value of 0. In the case of optical fibres, where light is used, the three states could be 'off', 'low intensity' and 'high intensity'. In both AMI and HDB3 line codes, alternative *marks* are sent as positive and negative pulses. Alternating the polarity of the pulses helps to prevent direct current being

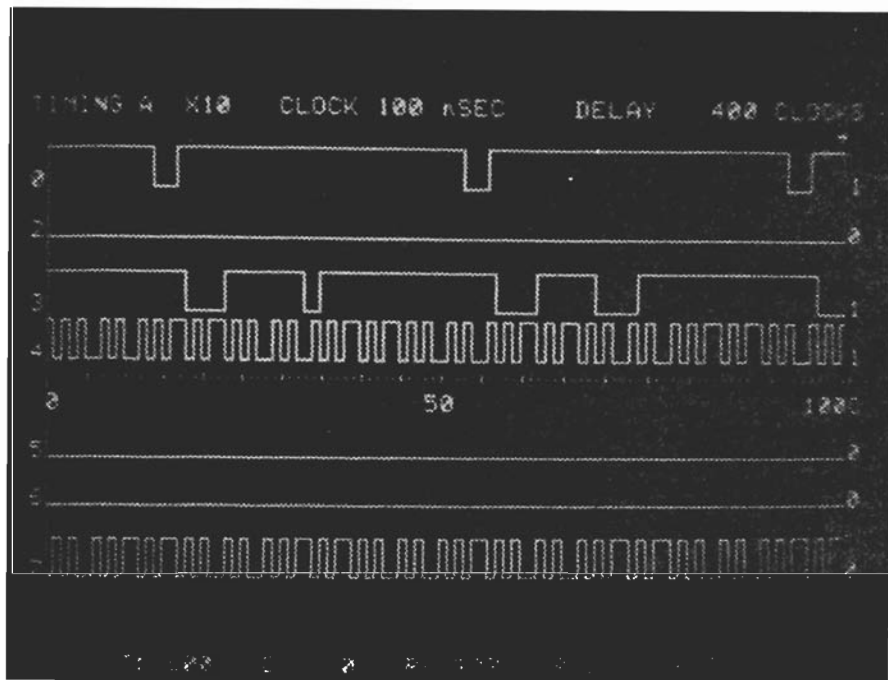


Figure 5.14 Digital signal pattern. These oscilloscope patterns result from testing of circuits using a standard line format for the Bell Systems digital network. (Courtesy of AT&T)



transmitted to line. In a two-state code, a string of marks would have the effect of sending a steady 'on' value to line.

The HDB3 code (used widely in Europe and on international transmission systems) is an extended form of AMI in which the number of consecutive zeros that may be sent to line is limited to three. Limiting the number of consecutive zeros brings two benefits: first a null signal is avoided, and second a minimum *mark density* can be maintained (even during idle conditions such as pauses in speech). A high mark density aids the regenerator timing and synchronization.

In HDB3, the fourth zero in a string of four is *marked* (i.e. forcibly set to 1) but this is done in such a way that the 'zero' value of the original signal may be recovered at the receiving end. The recovery is achieved by marking fourth zeros in *violation*, that is to say, in the same polarity as the previous 'mark', rather than in opposite polarity mark (opposite polarity of consecutive marks being the normal procedure).

## 5.11 OTHER LINE CODES AND THEIR LIMITATIONS

One of the line codes used in the past in North America in association with the 1.5 Mbit/s line system is called *zero code suppression*. The technique seeks to eliminate patterns of 8 or more consecutive zeros, but it does so in an irreversible manner by forcibly changing the value of the eighth consecutive bit of value 0, so that instead of transmitting 00000000, 00000001 is transmitted. Unfortunately, as it is only a two-state code, the receiving end device, unlike an HDB3 receiver, is unable to tell that the eighth bit value has been altered. An error results. The error is not perceptible to speech users, but would cause unacceptable corruption of data carried on a 64 kbit/s channel.

Once 'eighth bit encoding' using the zero code suppression technique had begun, it became acceptable to rob the eighth bit for other internal network uses. A *robbed bit signalling* channel, equivalent to the Europeans' *timeslot 16* signalling channel, was created, as already discussed.

Both of the above uses of 'eighth bit encoding' reduced the usable portion of the 64 kbit/s channel. For this reason, it is common for data terminals in North America to commonly use only seven of the eight available bits in each byte. This has the effect of reducing the usable bit rate to 56 kbit/s (8000 samples of 7 bits per second) even though 64 kbit/s is carried on the line. North American readers may be familiar with the 56 kbit/s user rate.

In connections from Europe to North America where the 56 kbit/s user data rate is employed, it is necessary to employ a rate adaptor at the European end to accommodate the lower rate. In essence the rate adaptor is programmed to waste the eighth bit of each byte, giving a 56 kbit/s user rate even at the European end. Alternatively a rate adaptor may be used at both ends to employ an even lower bit rate, such as the ITU-T standard bit rate of 48 kbit/s. In this case two bits of each byte are ignored. Stimulated by worldwide customer pressure for 64 kbit/s services (including ISDN, see Chapter 10), the restriction to 56 kbit/s channel capacity in North America looks set to disappear with the adoption of various new line codes. These include *B8ZS* (*bipolar 8-zero substitution*) and *ZBTSI* (*zero byte time slot interchange*). Like HDB3,

both eliminate long strings of zeros, but in a recoverable way. The B8ZS code, for example changes a string of eight zeros to the pattern '000VB0VB'.

Some US readers may additionally have come across the *64 kbit/s restricted* bandwidth. As its name suggests, this provides a bit rate close to 64 kbit/s. It derives from the use zero code suppressed channel prior to the availability of either B8ZS or ZBTSI. Any bit pattern can be sent by the user at the normal 64 kbit/s rate so long as the '00000000' pattern is never used.



---

# *The Principles of Switching*

---

In this chapter we deal with the mechanics of the switching process, describing the sequence of functions necessary to establish connections across a telecommunications network. We shall cover the principles of circuit switching (as would be used in voice or circuit data networks) as well as the statistical multiplexing switching techniques of packet and cell switching. In addition, we describe in outline some of the better known types of switch technology.

## 6.1 CIRCUIT-SWITCHED EXCHANGES

In circuit-switched networks a physical path or *circuit* must exist for the duration of a call between its point of origin and its destination, and three particular attributes are needed in all circuit-switched exchanges.

- First, the ability not only to establish and maintain (or *hold*) a physical connection between the *caller* and the *called party* for the duration of the call but also to disconnect (clear) it afterwards.
- Second, the ability to connect any circuit carrying an incoming call (*incoming circuit*) to one of a multitude of other (*outgoing*) circuits. Particularly important is the ability to select different outgoing circuits when subsequent calls are made from the same incoming circuit. During the set-up period of each call the exchange must determine which outgoing circuit is required usually by extracting it from the dialled number. This makes it possible to put through calls to a number of other network users.
- Third, the ability to prevent new calls intruding into circuits which are already in use. To avoid this the new call must either be diverted to an alternative circuit, or it must temporarily be denied access, in which case the caller will hear a *busy* or *engaged* tone or the data user will receive an equivalent message or signal.

Exchanges are usually designed as an *array* or *matrix* of *switched crosspoints* as illustrated in Figure 6.1.

The switch matrix illustrated in Figure 6.1 has five incoming circuits, five outgoing circuits and 25 switch crosspoints which may either be *made* or *idle* at any one time. Any of the incoming circuits, A to E, may therefore be interconnected to any of the outgoing circuits, 1 to 5, but at any particular instant no incoming circuit should be connected to more than one outgoing circuit, because each caller can only speak with one party at a time. In Figure 6.1, incoming circuit A is shown as connected to outgoing circuit 2, and simultaneously C is connected to 1, D to 4 and E to 3. Meanwhile, circuits B and 5 are idle. Therefore, at the moment illustrated, four calls are in progress. Any number up to five calls may be in progress, depending on demand at that time, and on whether the called customers are *free* or not. Let us assume, for example, that only a few moments before the moment illustrated, customer B had attempted to make a call to customer 3 (by dialling the appropriate number) only to find 3's line engaged. Any telephone user will recognise B's circumstance. However, only a moment later, customer E might cease conversation with customer 3, and instantly make a call to customer 5. If B then chances to pick up the phone again and redial customer 3's telephone number, the call will complete because the line to C is no longer busy. Figure 6.2 shows the new switchpoint

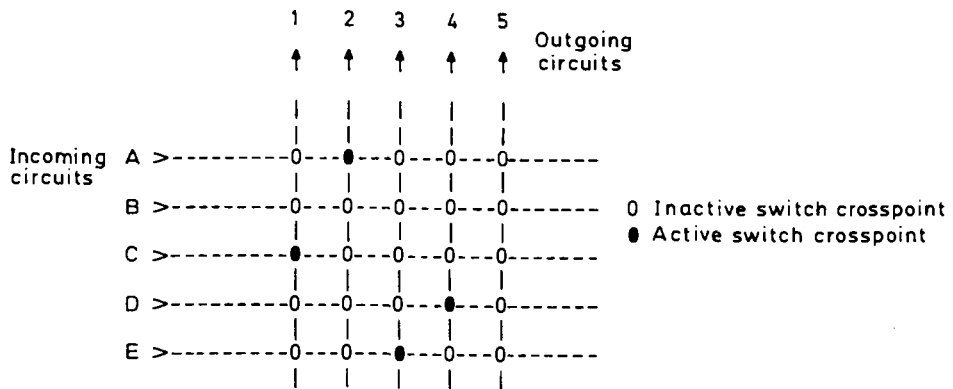


Figure 6.1 A basic switch matrix at a typical instant in time

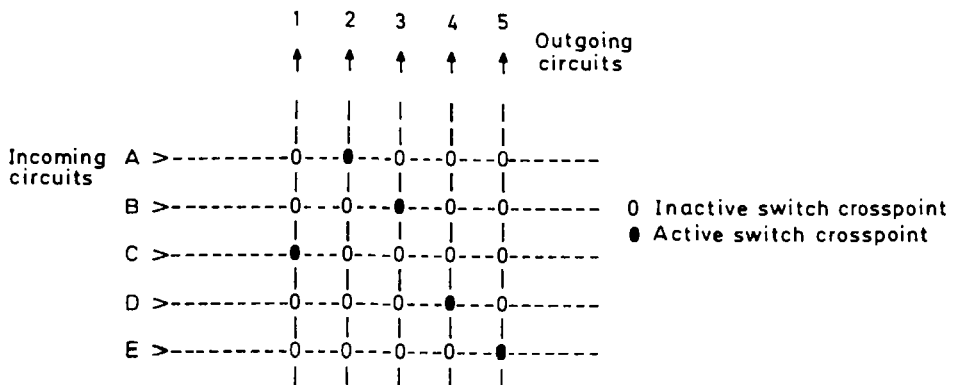


Figure 6.2 The same switch matrix a few moments later

configuration at this subsequent point in time, when five calls (the maximum for this switch) are simultaneously in progress.

At any one time, between nought and five of the twenty-five crosspoints may be in use, but never can an incoming circuit be connected to more than one outgoing circuit, nor any outgoing circuit be connected to more than one incoming circuit.

What exactly do we mean when we say a *connection* is made? In previous chapters on line transmission methods we concluded that a basic circuit needs at least two wires, and that a long distance one is best configured with four wires (a *transmit* and a *receive* pair). How are all these wires connected by the switch, and how exactly is intrusion prevented?

The answer to the first question is that each of the two (or four) wires of the connection is switched separately, but in an array similar to that in Figure 6.2. Thus a number of switch array 'layers' could be conceived, all switching in unison as shown in Figure 6.3, which illustrates the general form of a four-wire switch.

Each layer of the switch shown in Figure 6.3 is switching one wire of the four. Each of the four layers switches at the same time, using corresponding crosspoints so that all of the four wires comprising any given incoming circuit are connected to all the corresponding wires of the selected outgoing circuit (note that in Figure 6.3, circuit A is connected to circuit 5).

Additionally, in Figure 6.3 you will see that the fifth or so-called *P-wire* has also been switched through. The use of such an 'extra' wire is one way of designing switches to prevent call intrusion. Usually this method is used in an electro-mechanical exchange and it works as follows.

When any of circuits A to E are idle, there is an electrical voltage on their corresponding P-wires. When any of the callers A to E initiates a new call, the voltage on the P-wire is dropped to earth (zero volts). When the call is switched through the matrix to any of the outgoing circuits 1 to 5, the P-wire of that circuit will also be earthed as a result of being connected to the P-wire of the incoming circuit. When the call is over, the caller replaces the handset. This causes a voltage to be re-applied to

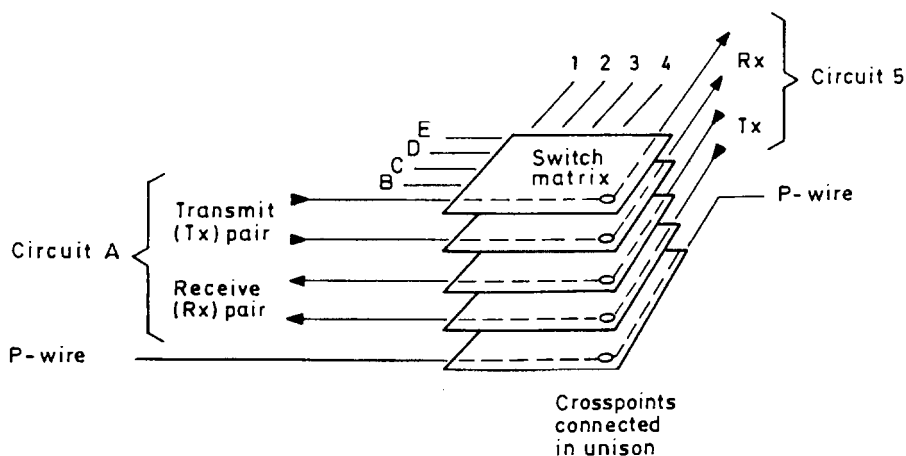


Figure 6.3 Switching a four-wire connection

the P-wire, to which the switch matrix responds by clearing the connection. Intrusion is prevented by prohibiting connection of circuits to others for which the P-wire is already in an earthed (i.e. busy) condition. In this way, the *earth* on the P-wire is used as a marker to distinguish lines in use.

The P-wire also provides a useful method of circuit holding, and for initiation of circuit *clearing*. To this end the switch is designed to maintain (or *hold*) the connection so long as the P-wire is earthed. As soon as the caller replaces the handset, the P-wire is reset to a non-zero electrical voltage, and the switch responds by clearing the connection (i.e. releasing the switch point).

In modern computer-controlled (*stored program control (SPC)*) exchanges call intrusion is prevented, and the job of *holding* the circuit is carried out by means of the exchange processor's electronic 'knowledge' of the circuits in use. *P-wires* are thus becoming obsolete.

To return to the example in Figures 6.1 and 6.2, what if party A wishes to call party E? Our diagrams show A and E as only able to make outgoing calls through the switch, so how can they be connected together? The answer lies in providing one or both circuits with access to both *incoming* and *outgoing* sides of the exchange, and it is done by *commoning* (i.e. wiring together) circuits 1 and A, 2 and B, 3 and C, 4 and D, 5 and E, as shown in Figure 6.4. (Incidentally, in the example shown in Figure 6.4 as well as in other diagrams in the remainder of the chapter, it is necessary to duplicate the layers for each of the two or four wires of the connection, as already explained in Figure 6.3. For simplicity, however, the diagrams only illustrate one of the layers).

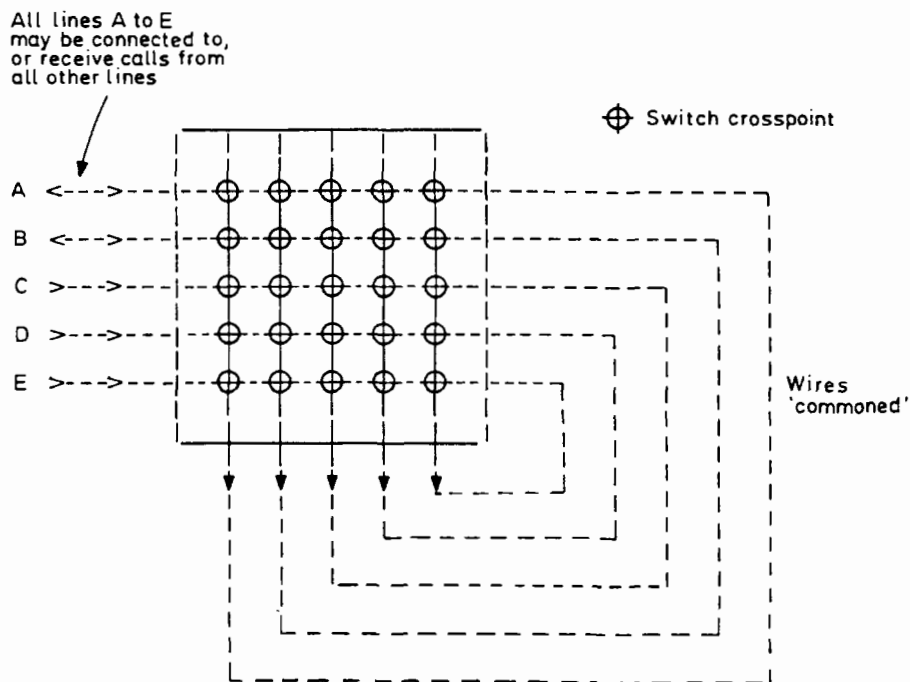


Figure 6.4 A simple small local exchange

In Figure 6.4, any of the customers A to E may either make calls to, or receive calls from, any of the other four customers. The maximum number of simultaneous calls now possible across the matrix is only two, as compared with the five possible in Figure 6.2. This is because two calls are sufficient to engage four of the five available lines. One line must therefore always be idle.

A switch matrix designed in the manner illustrated in Figure 6.4 would serve well in a small exchange with only a few customers, such as an office *private branch exchange* (PBX) or a small *local exchange* (termed *central office* or *end office* in North America) in a public telephone network. The exchange illustrated in Figure 6.4 is actually a *full availability and non-blocking system*. Fully available means that any line may be connected to any other; non-blocking indicates that so long as the destination line is free a connection path can be established across the switch matrix regardless of what other connections are already established. These terms will be more fully explained later in the chapter and we shall also discover some of the economies that may be made in larger exchanges, by introducing *limited availability*.

First, let us consider how the isolated system of Figure 6.4 can be connected to other similar systems in order to give more widespread access to other local exchanges, say, or to trunk and international exchanges. Figure 6.5 illustrates how this is done. It shows how some circuits, which are designated as incoming or outgoing *junctions*, connect the exchange to other exchanges in the network. Such inter-exchange circuits allow connections to be made to customers on other exchanges.

Networks are built up by ensuring that each exchange has at least some *junction*, *tandem* or *trunk* circuits to other exchanges. The exchanges need not be fully interconnected, however; connections can also be made between remote exchanges by the use of *transit* (also called *tandem*) routes via third exchanges, as shown in Figure 6.6. *Junction* and *trunk* circuits are always provided in multiple numbers. This means that if one particular circuit is already in use between two exchanges, a number of equally suitable alternative circuits (interconnecting the same exchanges) could be used instead. Figure 6.6 shows four circuits interconnecting exchanges P and Q; two each for incoming and outgoing directions of traffic. The consequence is that when circuit 1 from exchange P to exchange Q is already busy, circuit 2 may be used to establish another call. Only when both circuits are busy need calls be failed, and callers given the *busy tone*.

Each of the exchanges shown in the networks of Figures 6.5 and 6.6 must have its switch matrix and circuit-to-exchange connections configured as illustrated in Figure 6.7.

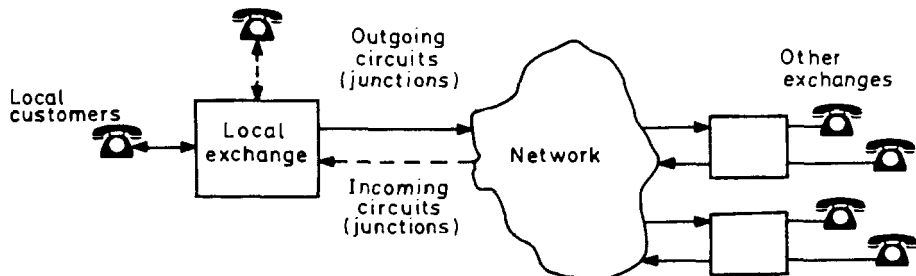


Figure 6.5 Junction connection to other exchanges



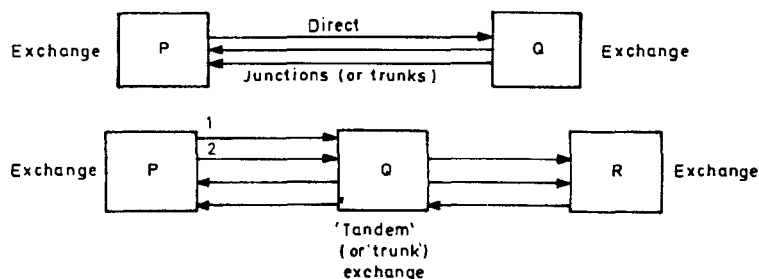


Figure 6.6 Typical networking arrangements

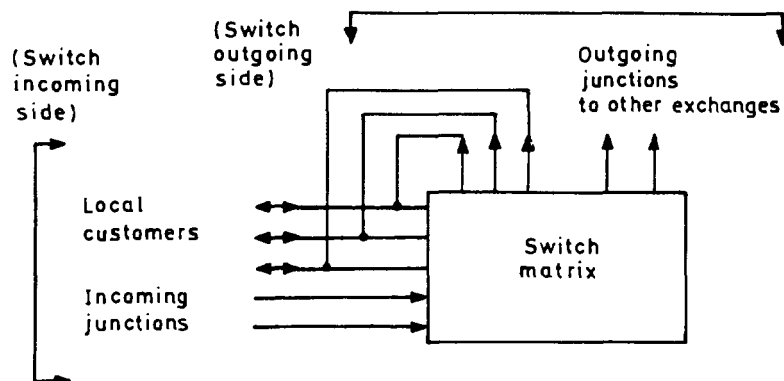


Figure 6.7 Local exchange configuration

Figure 6.7 shows how the local customers' lines are connected to both incoming and outgoing sides of the switch matrix, and how in addition a number of *uni-directional* (i.e. single direction of traffic) incoming and outgoing junctions are connected.

The junctions of Figure 6.7 could have been designed to be *bothway* junctions. In this case, like the customers' lines illustrated in Figure 6.4, they would need access to both incoming and outgoing sides of the switch matrix. In some circumstances this can be an inefficient use of the available switch ports, because it may reduce the number of calls that the switch can carry at any given time. (Remember that the matrix in Figure 6.4 may only carry a maximum of two calls at any time, while the same size matrix in Figure 6.1 could carry five calls).

## 6.2 CALL BLOCKING WITHIN THE SWITCH MATRIX

Telecommunications networks which are required to have very low call blocking probabilities need to be designed with excess equipment capacity over and above that needed to carry the average call load. Indeed, to achieve zero call blocking, we would need to provide a network of an infinite size. This would guarantee enough capacity

even in the unlikely event of everyone wanting to use the network at once. However, because an infinitely sized network is impractical, telecommunications systems are normally designed to be incapable of carrying the last very small fraction of traffic. Switching matrices are similarly designed to lose a small fraction of calls as the result of internal switchpoint congestion.

In the case of switch matrices we refer to the designed lost fraction of calls as the *switch blocking coefficient*. This coefficient exactly equates to the *grade-of-service* that we shall define in Chapter 30, and the dimensioning method is exactly the same. Thus a switch matrix with a blocking coefficient of 0.001 is designed to be incapable of completing 1 call in 1000. That one call will be lost as a direct consequence of *switch matrix congestion*. By comparison, a *non-blocking switch* is designed in such a way that no calls fail due to internal congestion.

How does switch blocking come about anyway, and how can costs be cut by designing switches with relatively large switch blocking coefficients? There are two methods of economizing on hardware, both of which inflict some degree of call blocking due to switch matrix congestion. They rely either on

- limiting circuit *availability*, or on
- employing *fan-in, fan-out* switch architecture.

and they are described separately below.

### 6.3 FULL AND LIMITED AVAILABILITY

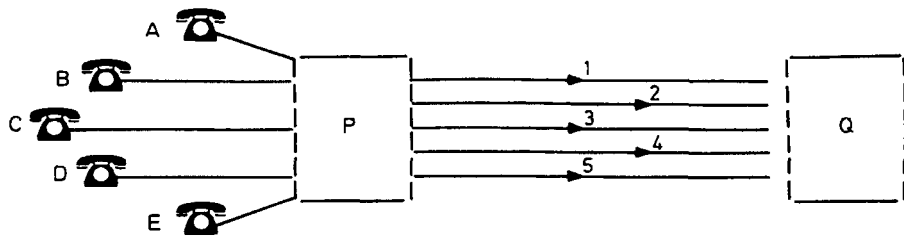
All switches fall into one of two classes:

- *full availability* switches, or
- *limited (or partial) availability* switches

The difference between the two lies in the internal architecture of the switch. The term *availability*, in this context, is used to describe the number of the circuits in a given outgoing *route* which are *available* to any individual incoming circuit. As an example, Figure 6.8 illustrates a simple network in which five customers, A to E, are connected to an exchange P, which, in turn, has five junction circuits to exchange Q. However, Figure 6.8 is not drawn in sufficient detail to show the *availability* of circuits within the group of junction circuits joining P and Q, because the architecture of the switch matrix itself is not shown.

Figures 6.9 and 6.10 illustrate two of many possible switch matrix architectures for the exchange P which was shown topologically in Figure 6.8. In Figure 6.9, all the outgoing trunk circuits from P (numbered 1 to 5) may be accessed by any of the customers lines, A to E. This is the *fully available* configuration, as all outgoing circuits are *available* to all the incoming circuits.

By contrast, in Figure 6.10 each of the customers may only access four of the five outgoing circuits. Not all of them get through to the same four, though. Customer A

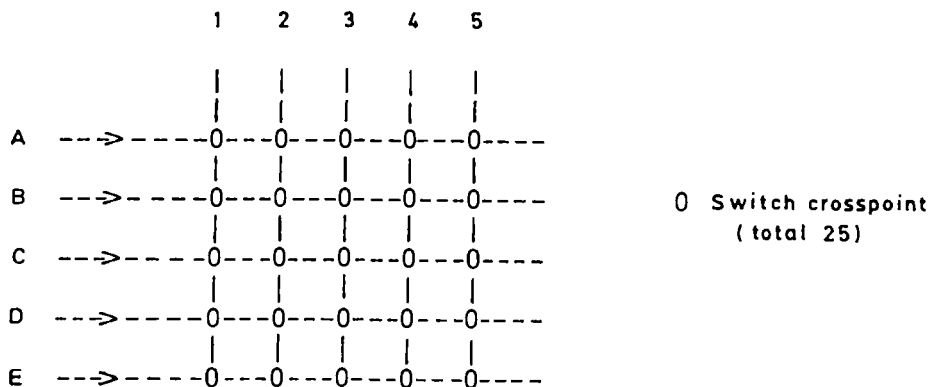


**Figure 6.8** Customers A to E on exchange P, which has five circuits to exchange Q

may access circuits 1, 2, 3, 4, customer B circuits 1, 2, 3, 5, customer C circuits 1, 2, 4, 5, customer D circuits 1, 3, 4, 5 and customer E circuits 2, 3, 4, 5. Figure 6.10 shows only one of a number of possible permutations (called *gradings*) in which the outgoing circuits could be made *available* to the incoming ones. Figure 6.10 therefore illustrates one particular *limited availability* grading. The *availability* of the grading shown is 4, as only a maximum of four outgoing circuits (within the outgoing route PQ) are *available* to any individual incoming circuit. This despite the fact that more than four circuits exist within the route as a whole. In this example the total route size PQ is five circuits.

Note in Figure 6.10 how the total number of switch crosspoints is only 20 compared with the 25 that were required in Figure 6.9. This may give the advantage of reducing the cost of the exchange, particularly if the switch matrix hardware is expensive. On the other hand it may be an unfortunate limitation of the hardware design that only four outgoing ports are possible per incoming *circuit*. As we will find later in this chapter, electromechanical switches are often not configurable as full availability switches because of the way they are made.

The disadvantage of limited availability switches is that more calls are likely to fail through internal congestion than with an equivalent full availability switch. The difference between the two is plain in Figures 6.9 and 6.10. In Figure 6.10, when circuits 1 to 4 are busy but circuit 5 is not, call attempts made on line A are *failed*, whereas the same



**Figure 6.9** Switch matrix of exchange P configured as 'fully available'

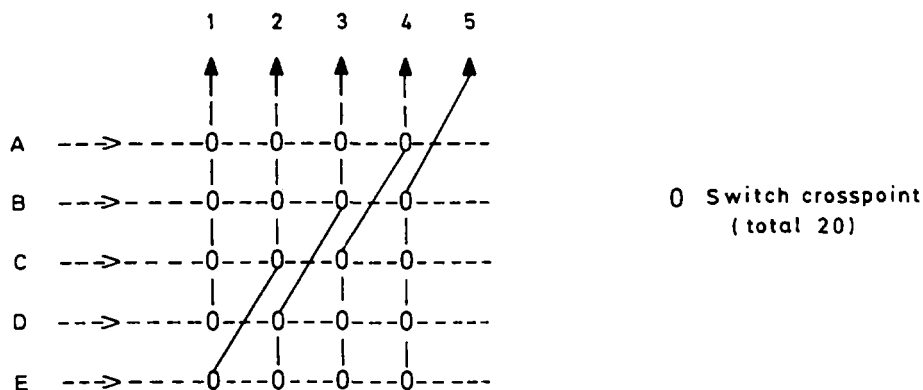


Figure 6.10 Switch matrix of exchange P configured as 'limited availability' (4)

attempt made on the configuration in Figure 6.9 will succeed, hence the lower switch blocking of the latter. Similarly, B cannot reach circuits 4, nor C circuit 3, D circuit 2 or E circuit 1.

In the past a whole statistical science grew up in order to minimize the grade of service impairments encountered in limited availability systems. It was based on the study of *grading* which involves determining the *slip-pattern* of wiring (for example see Figure 6.10) in which optimum use of the limited available circuits is achieved. The resultant *grading chart* is often diagrammatically represented in a form similar to that shown in Figure 6.11.

Figure 6.11 illustrates the grading chart of a number of selectors (switching mechanisms) sharing a common group of outgoing circuits to the same destination (e.g. a distant exchange). In total, 65 outgoing circuits are available in the grading, but of these, each individual incoming circuit (and its corresponding selector) can only be

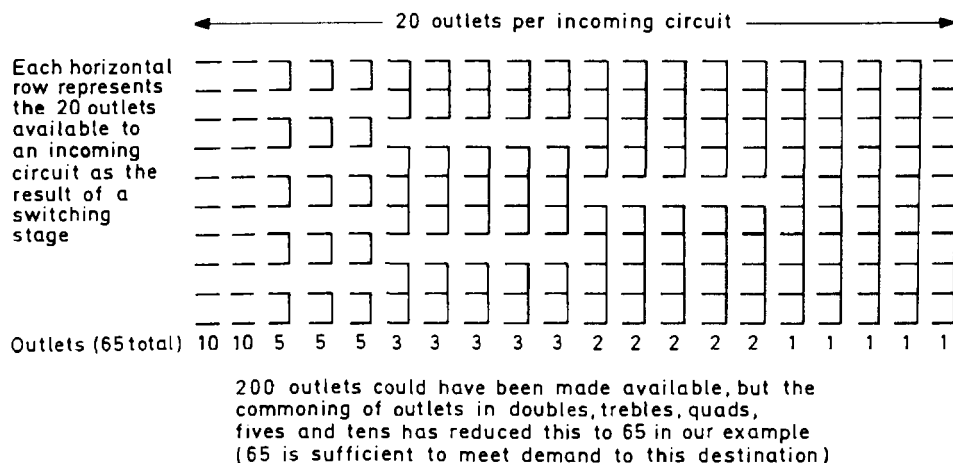


Figure 6.11 A 20-availability grading

connected to 20 of the 65. In other words, each selector (and therefore its corresponding incoming circuit) has a limited availability of 20. Thus the top horizontal row of 20 circuits shown on the grading chart of Figure 6.11 are the outlets available to a particular incoming circuit. The second row, meanwhile, represents the 20 outlets available to a different incoming circuit.

The action of a particular incoming circuit's selector mechanism is to scan across its own part of the grading (i.e. its horizontal row) from left to right, and to select the first available free circuit nearest the left-hand side. Other selectors similarly scan their rows of the grading to select free outgoing circuits. This means that outgoing circuits towards the left-hand side of the grading chart are generally more heavily used than those on the right. To counteract this effect, the right hand outlets of the grading (i.e. the later choices) are combined as *doubles*, *trebles*, *quads*, *fives* and *tens*, etc. This means that the same outgoing circuits are accessible from more than one selector, and thus incoming circuit. This helps to boost the average traffic carried by outgoing circuits in the 'later part' of the grading and so create even loading. The science of *grading* was particularly prevalent during the days of electromechanical exchange predominance, and different types of grading emerged, named after their inventors (e.g. the O'Dell grading).

In the diagram of Figure 6.11 you will see that apparently 10 incoming circuits (the number of horizontal rows) have access to a far greater number (65) of outgoing trunks circuits, all to the same destination exchange. Absurd you might think, and you would be right. There is no point in having more outlets to the same destination than the incoming demand could ever need. The explanation is that in practice the grading horizontal reflects identical wiring of ten or twenty selectors making up a whole shelf. In our case, then, 100 ( $10 \times 10$ ) or 200 ( $20 \times 10$ ) incoming circuits are vying for 65 outgoing trunks. You will agree that this is much more plausible. The reason that the grading is simplified in this way is that it is much easier to design and wire a  $10 \times 20$  grading and duplicate it than it is to create a  $100 \times 20$  or  $200 \times 20$  grading.

In our example, if 20 or fewer circuits would have sufficed to meet the traffic demand to the destination then the grading work is much easier. In this case, all the selector outlets of Figure 6.11 may be *commoned* and full availability of outgoing circuits is possible, each incoming circuit capable of accessing each outlet.

Limited availability switches are nowadays becoming less common, as technology increasingly enhances the sophistication and reduces the cost of modern exchanges, thereby removing many of the hardware constraints and extra costs associated with limited availability switch design. Among older exchange technologies, limited availability was a common hardware constraint. Typically, *Strowger* type exchanges (described later in this chapter) were either 10-, 20- or 24- availability. In other words each incoming circuit had access only to a maximum of either 10, 20 or 24 circuits.

## 6.4 FAN-IN-FAN-OUT SWITCH ARCHITECTURE

In Figure 6.4 we developed a simple exchange, suitable on a small scale to be a *fully-available* and *non-blocking* switching mechanism for full interconnectivity between five customers. It was achieved with a switch matrix of 25 crosspoints. Earlier in the chapter

we suggested that economies of scale could be made within larger switches. The main technique for achieving these economies is the adoption of a *fan-in-fan-out* switch architecture.

Consider a much larger equivalent of the exchange illustrated in Figure 6.4. A typical local exchange, for example, might have 10 000 customer lines plus a number of junction circuits, so that in using a configuration like Figure 6.4, a matrix of around  $10\,000 \times 10\,000$  switch crosspoints would be required. Bearing in mind that a typical residential customer might only contribute on average 0.5 calls to the busy hour traffic, and that each call has an average duration of 0.1 hours (6 minutes), then the likely maximum number of these switchpoints that will be in use at a given time is only around  $10\,000 \times 0.05$ , or 500. The conclusion is that a similar arrangement to Figure 6.4 is rather inefficient on this much larger scale.

Let us instead set a target maximum switch blocking of 0.0005. In other words, we intend that a small fraction (0.05%) of calls be lost as the result of internal switch congestion. This is a typical design value and such target switch blocking values can be met by using the *fan-in-fan-out* architecture illustrated in Figure 6.12.

Note how the total number of switchpoints needed has been reduced by breaking the switch matrix into *fan-in* and *fan-out* parts. 561 connections join the two parts, the significance of the value 561 being that this is the number of circuits theoretically predicted to carry 500 simultaneous calls, with a blocking probability of 0.05% (see Chapter 30). The switch is now limited to a maximum carrying capacity of 561 simultaneous calls, but the benefit is that the total number of switchpoints required is only  $2 \times 561 \times 10\,000$  or around 10 millions (a tenth of the number required by the configuration like that of Figure 6.4). This provides the potential for savings in the cost of switch hardware.

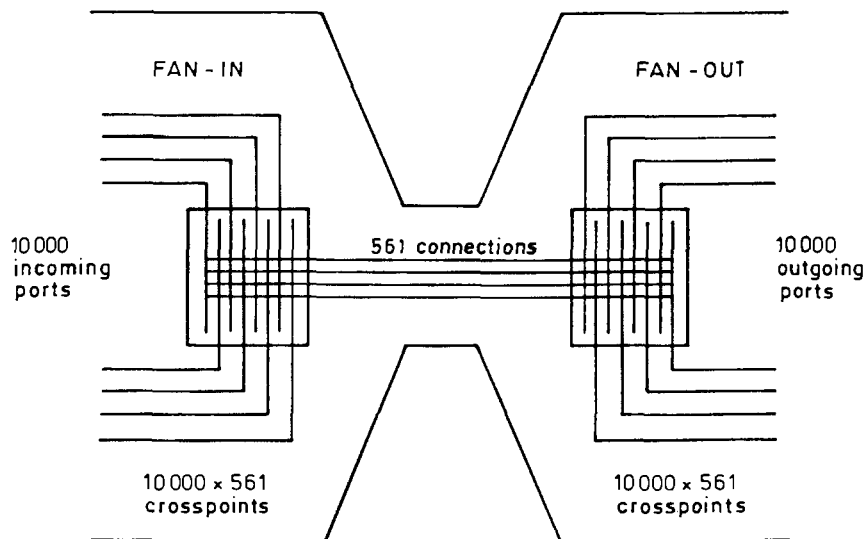


Figure 6.12 The principle of fan-in-fan-out

In our example it is intended that the internal blocking should never exceed 0.05% of calls failed due to internal switch congestion. In practice the actual switch blocking depends on the actual offered traffic, and it may be slightly higher or lower than this nominal value.

## 6.5 SWITCH HARDWARE TYPES

We have dealt with the general principles of exchange switching and the need for a number of incoming lines to be able to be connected to a range of outgoing lines, using a matrix of switched crosspoints. We now go on to discuss the different ways in which the matrix can be achieved in practice, and describe four individual switch types. In chronological order these are:

- Strowger (or *step-by-step*) switching
- crossbar switching
- reed relay switching
- digital switching

A number of other types, *Rotary*, *500-point*, *panel* and *X-Y* switching systems have been developed over the years. They are not discussed in detail here.

## 6.6 STROWGER SWITCHING

Strowger switches were the first widely used type of automatic exchange systems. They were developed by and named after an American undertaker who was keen to prevent human operators transferring calls to his competitors. His patent was filed on 12 March 1889.

Strowger exchanges are a marvel of engineering ingenuity, using precisely controlled mechanical motion to make electrical connections and, though now largely obsolete, they provide a valuable insight into the switching functions necessary in a telephone network, and an understanding of some of the historical reasons for modern telephone network functions and structures. The combination of electrical and mechanical components used in Strowger switching leads to the much-used expression *electro-mechanical switching*.

The switching components of Strowger exchanges are usually referred to as *selectors*, and they work in a manner which is marvellously easy to understand. In its simplest form, a selector consists of a moving set of contacting arms (known as a *wiper assembly*) which moves over another fixed set of switch contacts known as the *contact bank*. The act of switching consists of moving (or *stepping*) the contactor arm over each contact in turn until the desired contact is reached. Two main types of Strowger (or *step-by-step*) selectors are used in most exchanges of this type. They are called *uniselectors* and *two-motion selectors*.

A uniselector is a type of selector in which the wiper assembly rotates in one plane only, about a central axis. The contactors move along the arc of a circle, on which the fixed contact bank is arranged. Figure 6.13 illustrates the principle. A single incoming circuit is connected to the uniselector's contactor on the wiper assembly, and 25 possibly outgoing circuits are connected, one to each of the individual contacts making up the contact bank. The first contact in the bank is not connected to any outgoing circuit, but serves as the *rest* position for the wiper arm during the idle period between calls, thus in practice only 24 outlets are available.

When a call comes in on the incoming circuit, indicated by a loop (say, because the customer has picked his phone up), the uniselector automatically selects a free outgoing connection to a *first selector*. The *first selector* is a Strowger 2-motion selector which initially returns the dial tone to the caller and subsequently responds to the first dialled digit. We shall discuss the mechanism of 2-motion selectors shortly.

*Uniselectors* consist of a number of rows (or planes) of bank contacts, as the photograph in Figure 6.14 illustrates. The use of a number of planes allows all the contacts necessary for two, three or four wire and P-wire switching to be carried out simultaneously.

How do we cope with more than one incoming circuit? One answer is to provide a uniselector corresponding to each individual callers line. The outlets of these uniselectors can then be *graded* as we have seen to provide access to a suitable number of first selectors sufficient to meet traffic demand. A simple arrangement of this type is shown in Figure 6.15(a). However, unless the traffic on the incoming circuit is quite heavy then this arrangement is relatively inefficient and uneconomic, requiring a large number of uniselectors which see little use. For this reason it is normal to provide also a *hunter* or *linefinder*. This is a second uniselector, wired back-to-back with the first as shown in Figure 6.15(b). The linefinder (or hunter) is used to enable the first uniselector to be shared between a number of incoming lines. A number of linefinders (sufficient to meet customers traffic demand) are graded together, giving each individual line a number to choose from (Figure 6.15(c)).

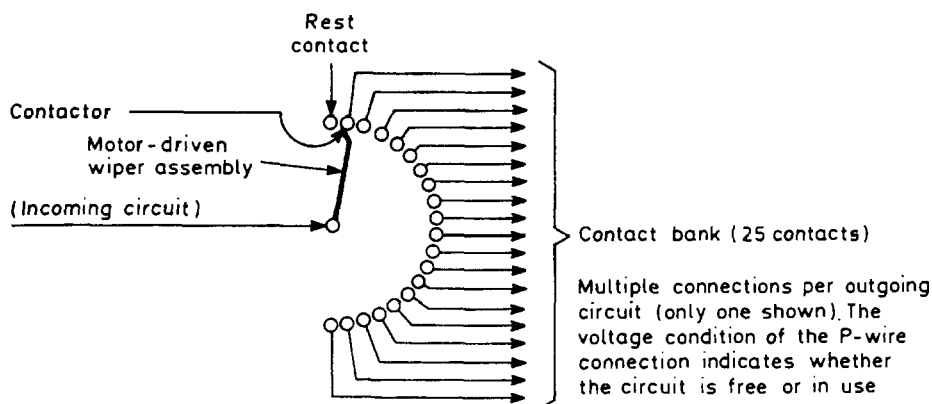
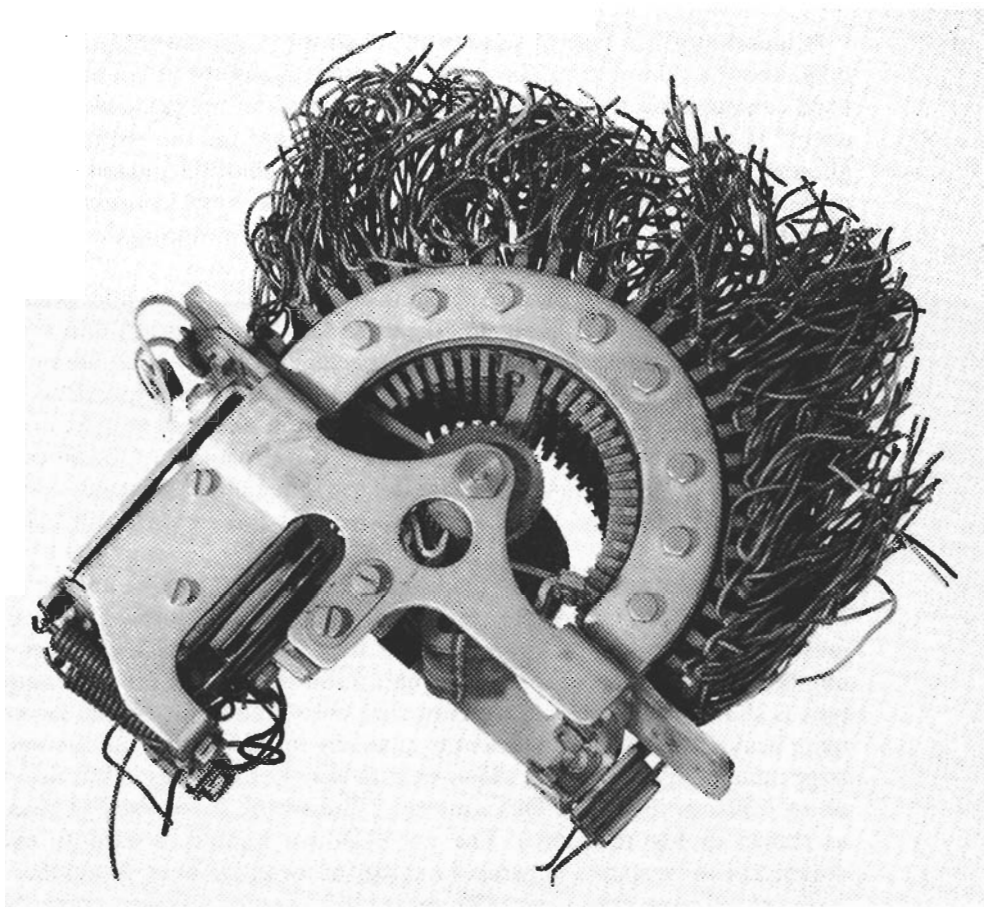


Figure 6.13 A simple uniselector

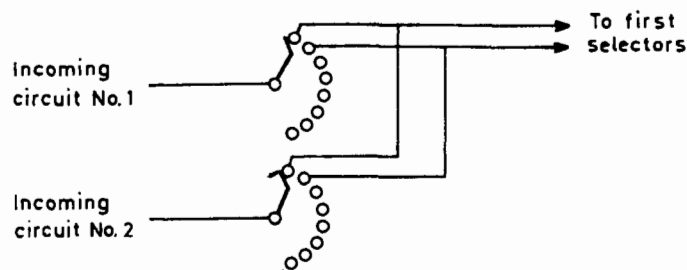




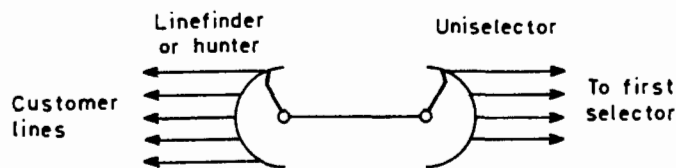
**Figure 6.14** Strowger unselector. The wipers move in the arc of a circle around 25 outlet contacts, stopping at a free outlet. Uniselectors are typically used on the customer side of an exchange, helping to find free exchange equipment to handle the call.

Different selectors in the same grading are prevented from simultaneously choosing the same outgoing circuit by the action of the P-wire, as we saw earlier in the chapter. It is also the P-wire that invokes the *release* of the selectors at the end of a call, whereupon a spring or other mechanical action returns the wiper assembly to the rest or home position.

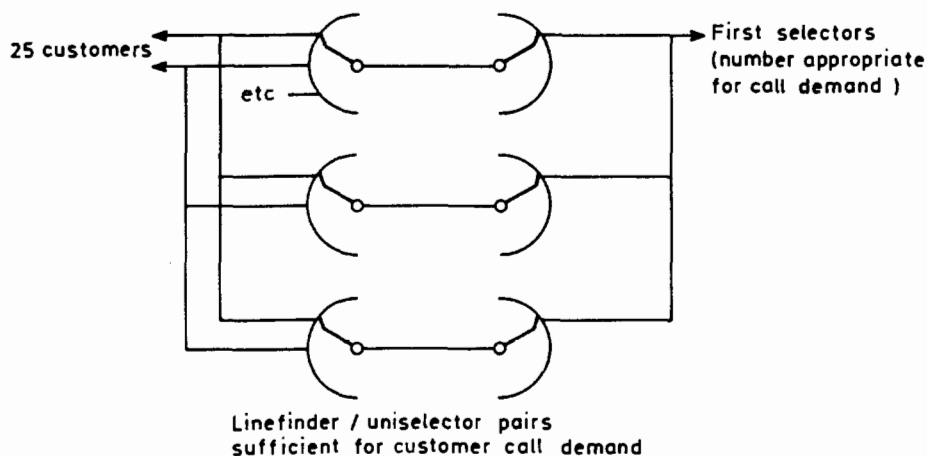
The most common type of selector found in Strowger exchanges is the *two-motion selector*. These are the type capable of responding to dialled digits. The wipers of a two-motion selector can, as the name implies, be moved in two planes. The first motion is linear, up-and-down between the ten planes (or *levels*) of bank contacts under dialled digit control. This is followed by a circular rotation into the bank itself. The second motion can be an automatic motion, scanning across the *grading* to find a free connection to a subsequent two-motion selector to analyse the next digit. Alternatively, if the two-motion selector is a *final selector*, then the selector analyses both the final two digits of the called customers number. In this case the rotary motion of the selector is controlled by a dialled digit.



(a) Uniselectors graded to share 'first' selectors



(b) Linefinder used to achieve unselector economy

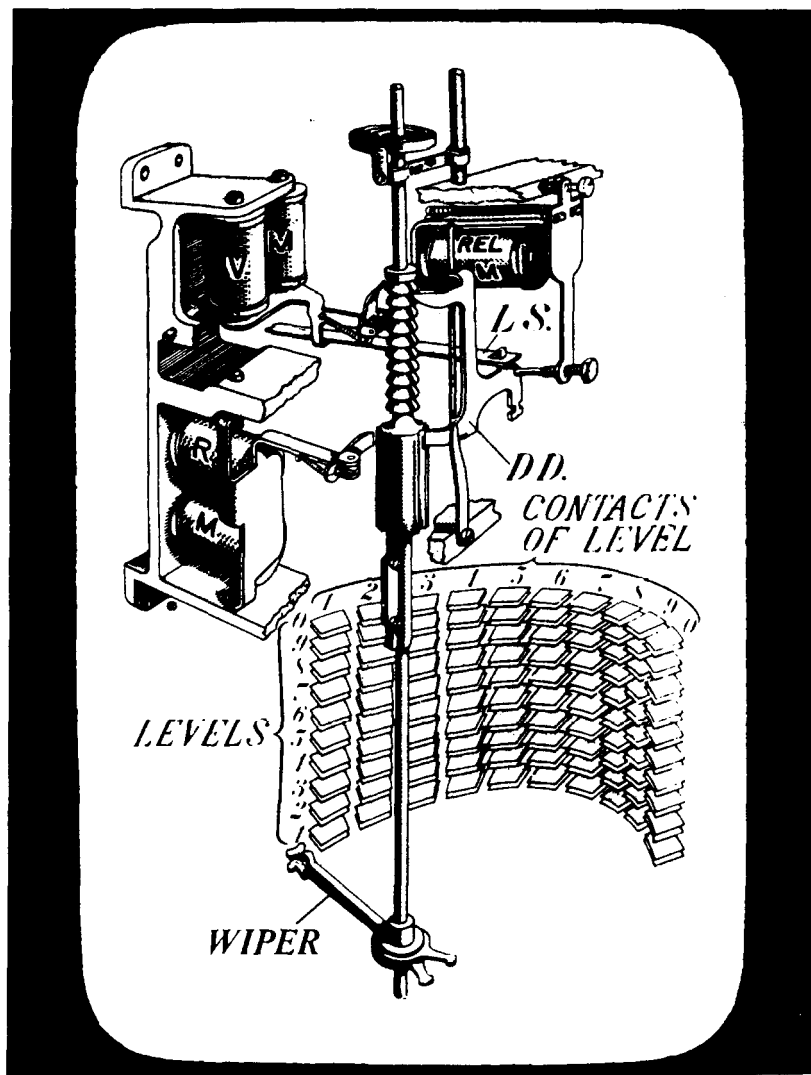


(c) Grading of linefinders and uniselectors

Figure 6.15 Using uniselectors

Each contact bank plane in a two-motion selector appears as an arch shaped layer or *level* with a number of sets of equally spaced contacts in each bank. The appearance of the selector contact bank is thus somewhat akin to a portion of a cylinder, as Figure 6.16 shows.

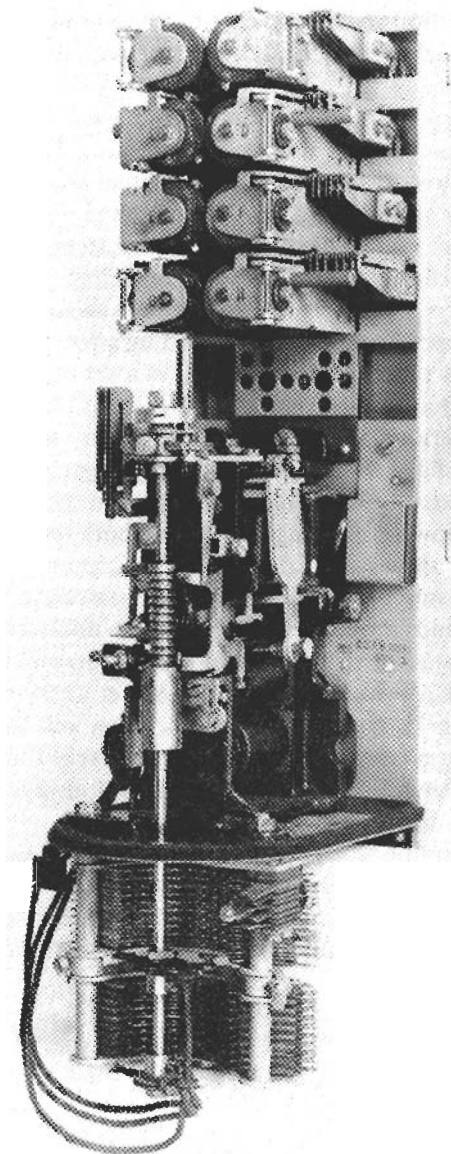
In practice the cylindrical bank of a two-motion selector needs to be duplicated to allow each of the wires of a connection (two-wire plus P-wire, etc.) to be connected



**Figure 6.16** The principle of the Strowger two-motion selector. An illustration from an old manuscript, showing the basic action of the two-motion selector. Pulsing the VM relays initially steps the wiper in the vertical plane. RM relay pulses then move the wiper an appropriate number of contacts in the horizontal plane. (Courtesy of BT Archives)

simultaneously, so that an actual two motion selector looks a little more complicated, as the photograph Figure 6.17 reveals.

Having heard the dial tone returned from the first selector (when it is ready), the caller dials the *called number*. This is indicated to the exchange by a train of electrical loop-disconnect ('off-on') pulses on the incoming line itself. These are the pulses which activate *first, second* or subsequent *two-motion selectors* accordingly. The number of 'off-on' pulses used to represent a particular digit corresponds to the value of the digit



**Figure 6.17** Strowger two-motion selector. This one is actually a final selector, used in the British Post Office network. (Courtesy of BT Archives)

dialled. Thus one pulse is the digit value 1, two pulses equals value 2 etc. The digit value 0 is represented by ten pulses. This simple form of signalling is called *loop disconnect* or *LD signalling*; it was described in Chapter 2. Each pulse steps the selector upwards in the vertical plane by one *level*. The gap between dialled trains of pulses (the *inter-digit pause*) indicates the end of the train and so marks the end of the *stepping* sequence. When a two-motion selector detects the gap between dialled digits then the rotary action of the selector commences automatically, moving the wiper into the bank to find

a free connection. Alternatively, in the case of a final selector, the *inter-digit pause* is used by the selector to ready itself for receiving a second dialled digit to control its rotary motion.

Digressing for a moment, it is worth noting that a Strowger two-motion selector has a limited availability of 10 (or sometimes 20). The constraint arises because the selector may only automatically scan the horizontal levels of the bank and on each a maximum of 10 (or 20) outlets may be accommodated.

Apart from being ideal for the direct stepping of two-motion Strowger selectors, another benefit of *loop disconnect* signalling is the ease with which pulses may be generated by a dial telephone. Figure 6.18 shows a dial telephone and how the pulses of loop disconnect signalling are created by a rotating cam attached to the telephone dial. As the dial is turned, the cam operates a set of contacts which connect and disconnect the circuit. Depending on how far the dial is turned, a number of pulses are generated. A longer period of electrical current 'on' separates the bursts corresponding to consecutive digits of the number. This longer on-period is called the *inter-digit pause* and is generated by an initial wide tooth on the rotating cam, as shown in Figure 6.18.

Actual Strowger exchanges comprise both *uniselector* and *two-motion selector* types. Figure 6.19 shows an example of a permutation of uniselectors and two-motion selectors to support up to 1000 customers on a three-digit numbering scheme. The uniselector finds a free *first selector*, which analyses the first dialled digit and then finds a free *final selector* in the correct range (corresponding to the first dialled digit). The final selector provides the final connection to the customer.

For networks consisting of more than one exchange it is clear that the same destination customer will not always be reached using the same route from the calling customer, because the starting place is not always the same. All routes, however, will comprise a number of switching stages, located in the various exchanges. These need to be fed with trains of digit impulses to operate. As the path is not common, then each caller will either have to dial a different string of digits, or else each exchange will have to be capable of *translating* a common dialled string into the string necessary to activate the selectors on the route needed from that particular exchange. In the latter case, a

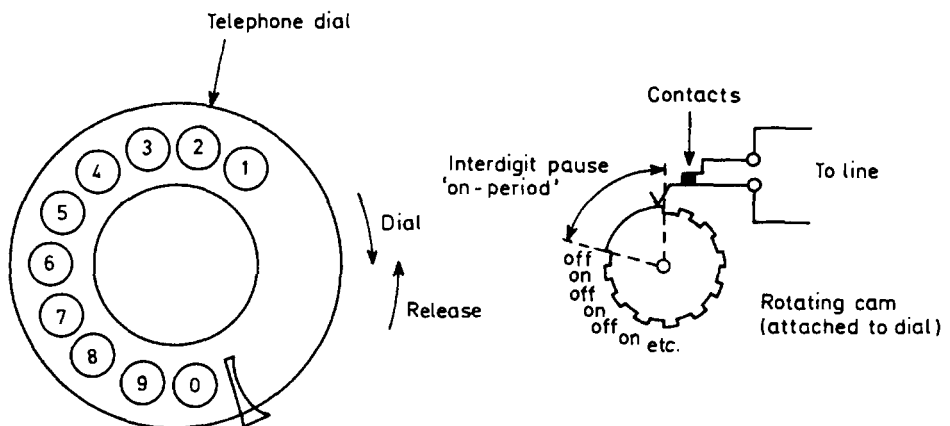


Figure 6.18 Generating 'loop-disconnect' signals

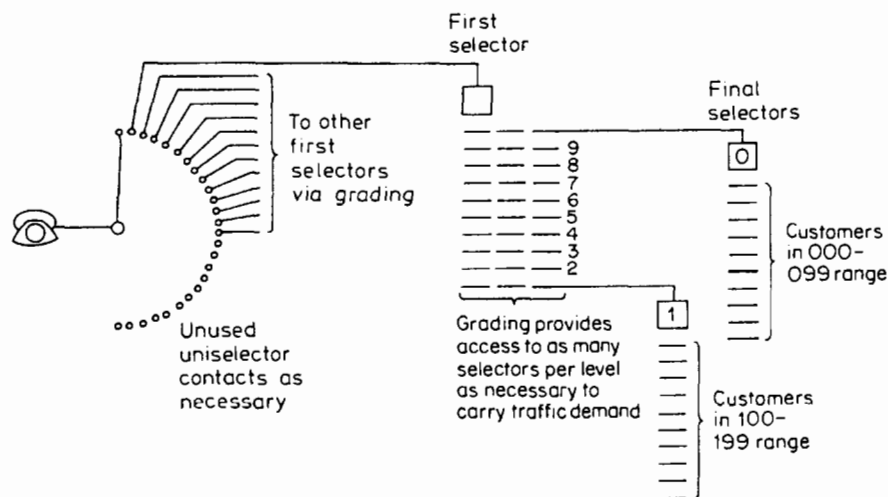


Figure 6.19 Selectors for a 1000-customer Strowger exchange

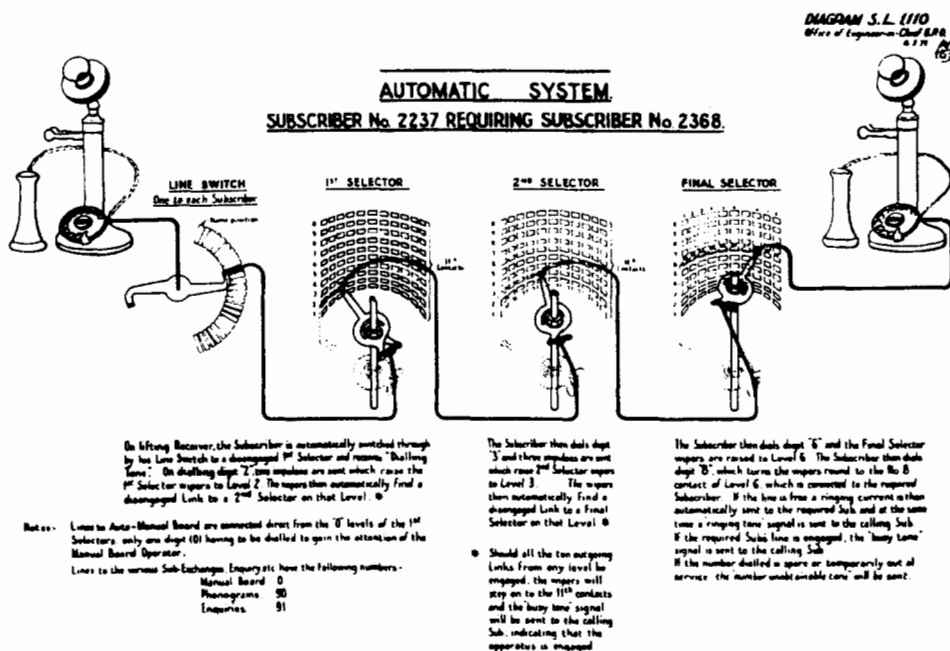


Figure 6.20 The principle of Strowger automatic switching. An extract from an early British general Post Office (GPO) article, explaining the principle of Strowger automatic switching. (Courtesy of BT Archives)

so-called *common or linked numbering scheme*, a *register/translator* is needed to store the dialled digits and *translate* (or convert) them into the string of *routing digits* which are needed to step the selectors as previously described. The function of the register and the technique of number translation are both described more fully in the next chapter.

Finally, let us close our discussion of Strowger switching by explaining briefly why the phenomenon of limited availability arises. It comes about because each selector is limited in the number of bank contacts. On a *uniselector* this is typically 24. On a *two-motion selector* it is usually only 10 or 20. Thus no matter how many circuits make up the route in total, each incoming circuit may have access only to a limited number of them (24, 10 or 20 in the examples given above).

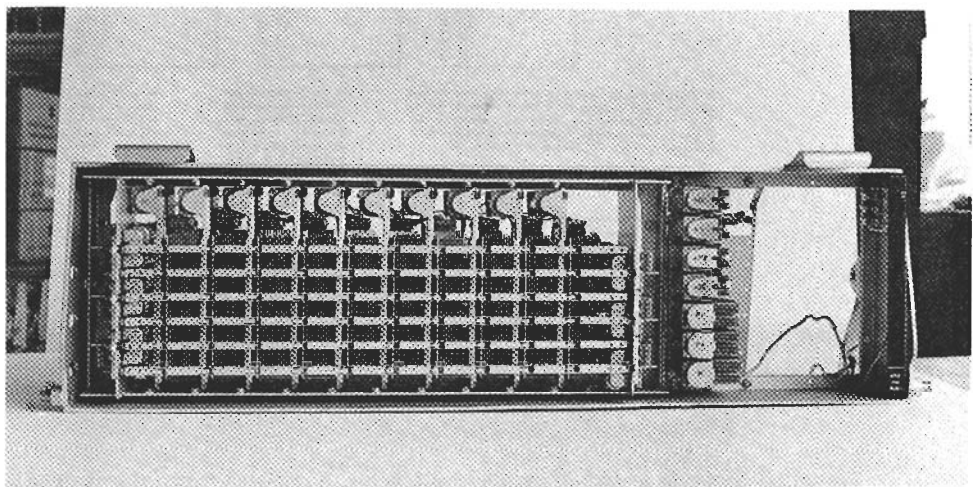
The major drawback of Strowger switching was the relatively large amount of space that it takes up, the relatively high electrical power needed for busy-hour operation, and the labour-intensive demands of maintaining it. The mechanical parts are highly prone to wear, and the electrical contacts are very sensitive to damage and dirt. As a result Strowger switching is nowadays largely obsolete, though some Strowger exchanges still operate in remote areas and may well remain in those lacking the capital to replace them.

## 6.7 CROSSBAR SWITCHING

Crossbar technology emerged in the 1940s and was partly, though not wholly, responsible for a change in equipment usage away from Strowger and other similar step-by-step exchange types. Crossbar switching offered the advantage of reduced maintenance and accommodation requirements, but was sometimes more expensive than Strowger for large and complex exchange applications.

As early as 1926 the first public crossbar exchange opened in Sweden. It worked in a step-by-step, digit-by-digit, manner which was too unwieldy for exchanges exceeding 2000 lines. It was not until the 1940s that crossbar systems became more common, following development in America of a system employing *marker control*. This became by far the most common type of crossbar exchange and it is the type described here.

The technique of crossbar switching is relatively easy to understand, because the switch matrix is immediately apparent from the physical structure of the components.



**Figure 6.21** Crossbar switching matrix. A typical basic crossbar matrix. This one comprises ten armatures and bridge magnets, six select bars and twelve select magnets. (Courtesy of British Telecom)

As the photograph in Figure 6.21 shows, the switch looks like a matrix, consisting of ordered vertical and horizontal members.

The operation of the crossbar switch is most readily understood by considering the action of two adjacent crosspoints. Figure 6.22 shows a single crosspoint of a crossbar assembly. The adjacent crosspoint of interest is a mirror image of the one shown immediately below it.

The crosspoint shown in Figure 6.22 is capable of simultaneous switching of three inlet wires to three outlet wires. The three inlet wires (a *pair* plus a P-wire) are connected to three fixed *bridge common* contacts and the three outlet wires are connected to three moveable outlet spring contacts. The three outlet spring contacts are joined together by a piece of insulating material, which also connects the outlet springs to the *lifting spring*. Thus when the lifting spring is moved to the right of the diagram the outlet springs are all simultaneously pushed into contact with the bridge common, so completing the connection. (As an aside, more wires can also be switched simultaneously, by adding more outlet springs and bridge commons).

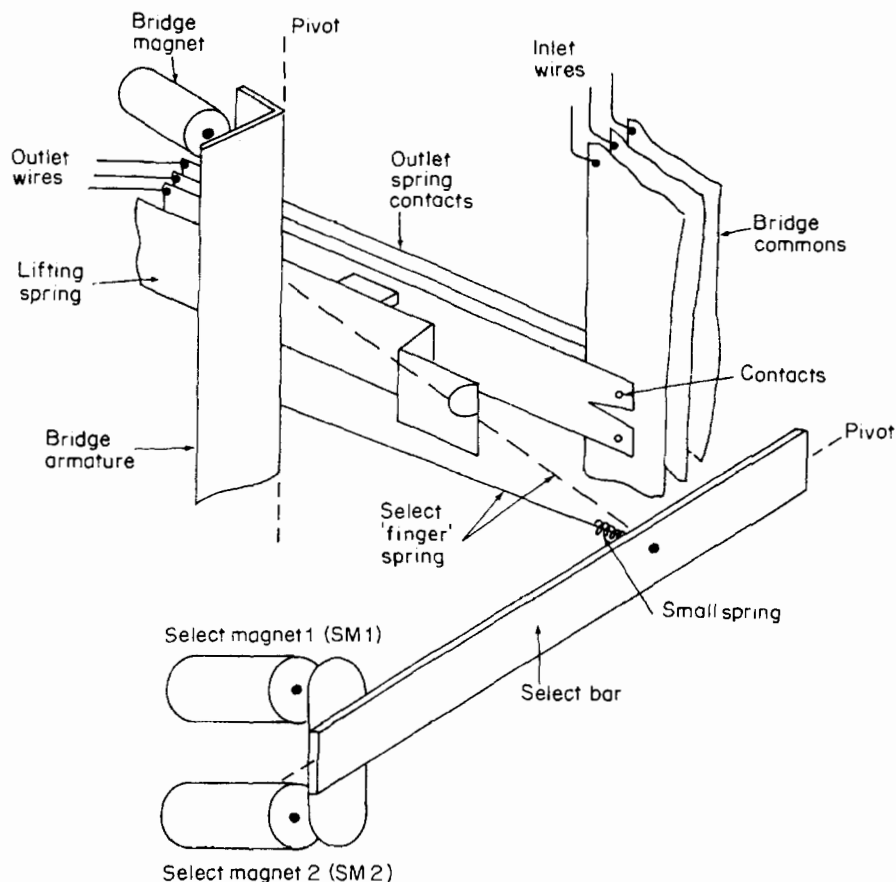


Figure 6.22 Operation of a crossbar switch



The lifting spring is actuated by the action of the *select finger* and the *bridge armature*. The *select finger* is really a thin piece of wire, connected to the *select bar* by a small spring, and this gives the *select finger* a small amount of flexibility. The select magnets are electromagnets, activated by passing an electric current through their coils. Not more than one of the select magnets is used at any one time. To activate the crosspoint shown in Figure 6.22, select magnet 2 must be operated. This has the effect of tilting the select bar, and so moving the select finger into the position marked by the dashed line on our diagram. At this point the bridge magnet is activated. This in turn pivots the *bridge armature*, with the effect of *trapping* the select finger onto the lifting spring, so moving the lifting spring towards the right of the diagram, and making contact between the outlet springs and the bridge commons. The connection is now complete. The bridge magnet must remain held as long as the connection is required, so trapping the select finger for the entire call. The select magnet, however, may be released once the connection has been made. The slight flexibility in the select finger permits it to remain trapped under the *bridge armature*, even when the select magnet is released.

At the end of the call, the bridge magnet is released, and the select finger springs back to its normal position.

By a similar mechanism, another set of outlet spring contacts (beneath the set illustrated) can be connected to the same inlet bridge commons (inlet circuit). In fact, this is the function of select magnet 1. The contacts are arranged in a mirror image beneath those of our illustration, and work in precisely the same way except that they use select magnet 1 instead of select magnet 2. Thus select magnet 1 selects one outlet circuit while select magnet 2 selects the other. It does not matter that there is only one select finger available, as we would not wish to connect the inlet to both outlets at the same time anyway.

Figure 6.23 illustrates a larger portion of a crossbar switch, showing three bridge magnets and armatures and two select bars. The form of the overall switch matrix is now apparent. To activate any particular crosspoint of the matrix one of the select magnets must be activated (according to the desired outlet circuit required) and then the appropriate bridge magnet is also activated, to connect the desired inlet circuit.

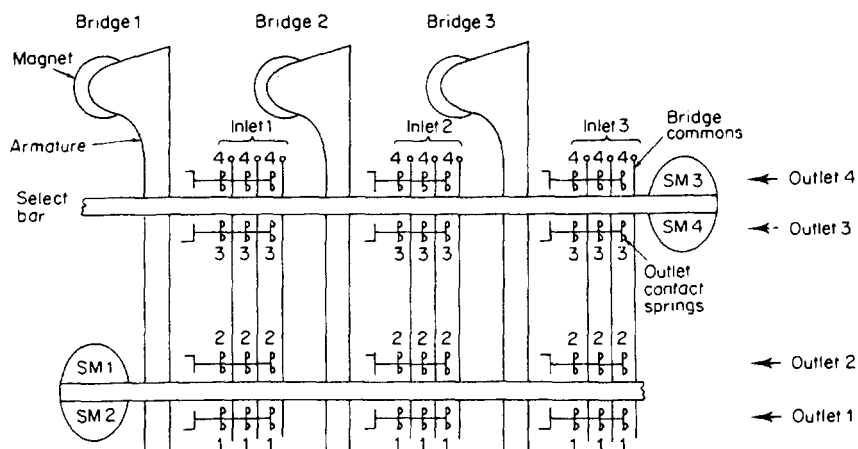


Figure 6.23 A crossbar matrix

Crossbar switch assemblies usually consist of 10 bridge armatures and 6 select bars, requiring 10 bridge magnets and 12 select magnets. (Note: some manufacturers of crossbar switches used more armatures and bridge magnets than discussed here. The principle, however, is the same.) Used in its most straightforward manner, therefore, a single crossbar switch assembly can be used for 10 inlets and 12 outlets. By combining a number of these assemblies a much bigger matrix can be built up.

The *limited availability* of 12 outlets from the basic assembly can raise problems, as we saw earlier in the chapter. So let us consider how we could improve things to allow 20 outlets. A single bridge of a 20 outlet switch is shown in Figure 6.24. Note that there are now six bridge commons and corresponding sets of outlet springs instead of only three. Notice also how the top two magnets have been labelled as *auxiliary magnets* instead of *select magnets*.

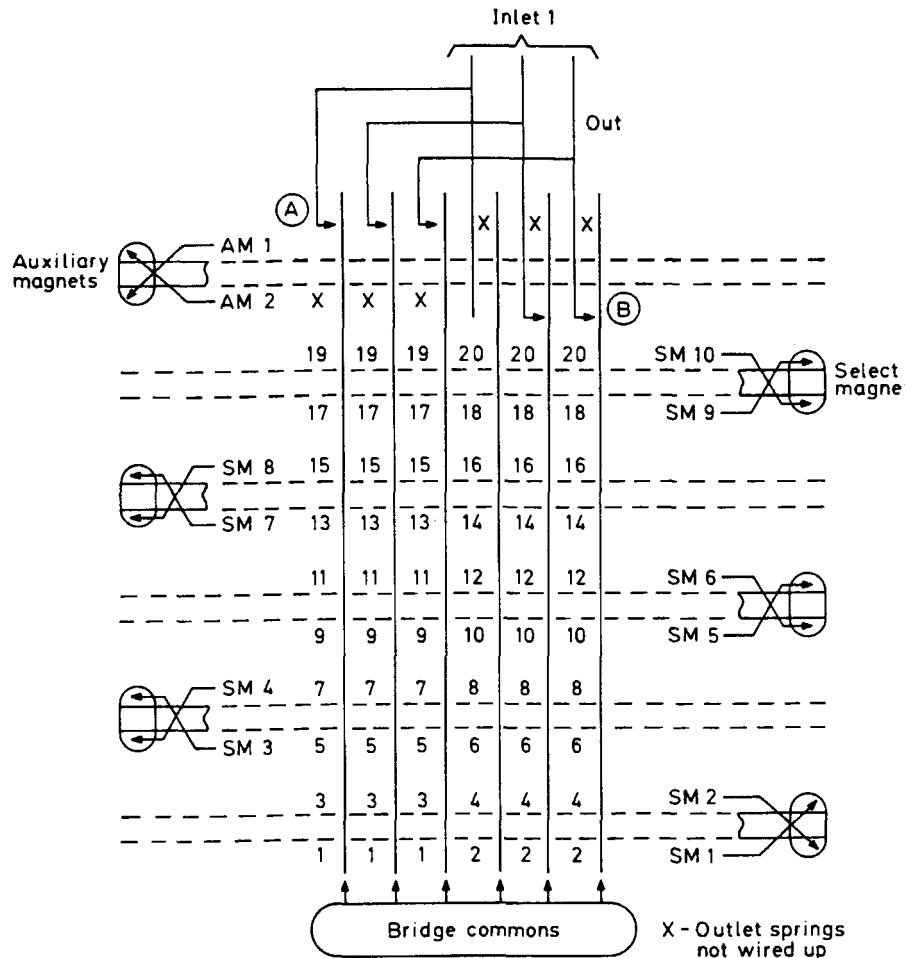


Figure 6.24 Twenty-outlet crossbar switch

In the assembly shown in Figure 6.24 each crosspoint is made by activating three magnets: one of the auxiliary magnets, one of the select magnets, and the appropriate bridge magnet. The auxiliary magnet essentially has the effect of connecting the inlet circuit either to the bridge commons corresponding to the even numbered outlet circuits, or to the bridge commons corresponding to the odd numbered outlet circuits. (The three right-hand bridge commons correspond to even numbered outlet circuits and to auxiliary magnet 2. The three left-hand bridge commons correspond to the odd numbered outlet circuits and auxiliary magnet 1). The overall assembly of 10 bridge armatures and 6 select bars is thus increased to a capacity of 10 inlets and 20 outlets. With even more sophisticated wiring and the use of 9 bridge commons, up to 28 outlet circuits can be made available to the 10 inlets.

The individual assemblies of crossbar switches are usually arranged in a *fan-in-fan-out* manner as described earlier in the chapter. In order to pick an appropriate path between inlet and final outlet circuit, the exchange must first examine the dialled digit train. This is done by the register, the action of which is described more fully in the next chapter. Having chosen the outlet circuit required, a path is usually selected by one of two methods. The easiest to explain is the method used in *stored program control (SPC)* crossbar exchanges. In SPC exchanges (i.e. the most recent ones), the exact path is determined by a special control computer from its knowledge of the instantaneous state of the exchange. The necessary cascade of crosspoints across the exchange is then activated. As an alternative we have the original method of *marker control*, which was electro-mechanical. In marker control, the final destination outlet is *marked*, and a large number of trial paths are then set up in a backwards direction across the switch in a cascade fashion. However, only the path which happens to find its way right across the chain of individual switches to reach the desired inlet is actually connected, while all the other trial paths are released.

## 6.8 REED RELAY SWITCHING

In reed relay switched exchanges the individual crosspoints consist of two nickel-iron *reeds* sealed in a glass tube, approximately 3 cm long, containing dry nitrogen. The free ends of the reeds are plated with gold to give a low resistance contact area. When the contact is not made, the reeds rest in their idle position with their contacts apart. Figure 6.25 shows this assembly.

The whole glass tube is located along the axis of an electromagnetic coil, and when the current is switched on in this coil it induces a magnetic field in the tube. The two reeds take on opposite polarity in the overlapping free ends, so that they are attracted together and a connection is made. When the current in the coil is switched off, the reeds return to their idle rest position, and the contact is again broken. This action can be used as the basis for activating the crosspoints of a switch matrix, as discussed in the early part of the chapter and re-illustrated in Figure 6.27.

Multiple wire connections are achieved by packing the electromagnetic coil or *bobbin* (as it is called), with more than one reed. Figure 6.28 illustrates a bobbin containing four reeds and which is hence capable of simultaneous four-wire connection.

Reed relay exchanges invariably have stored program control. The exchange processor analyses the dialled number and activates the appropriate switchpoints of the switch matrix by applying current to the coil bobbins.

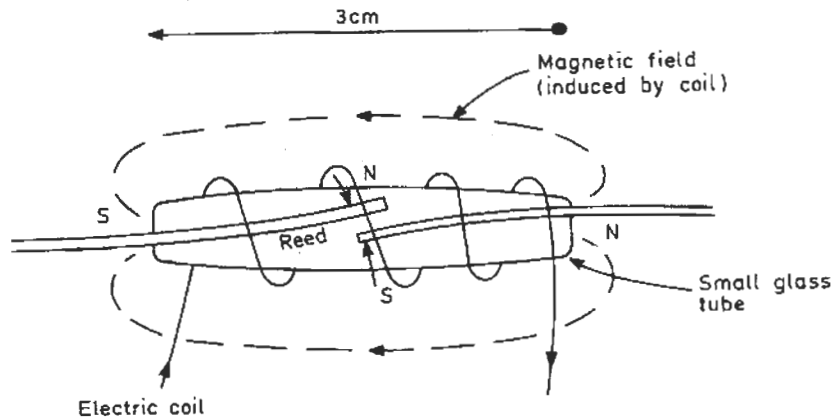


Figure 6.25 A reed relay crosspoint

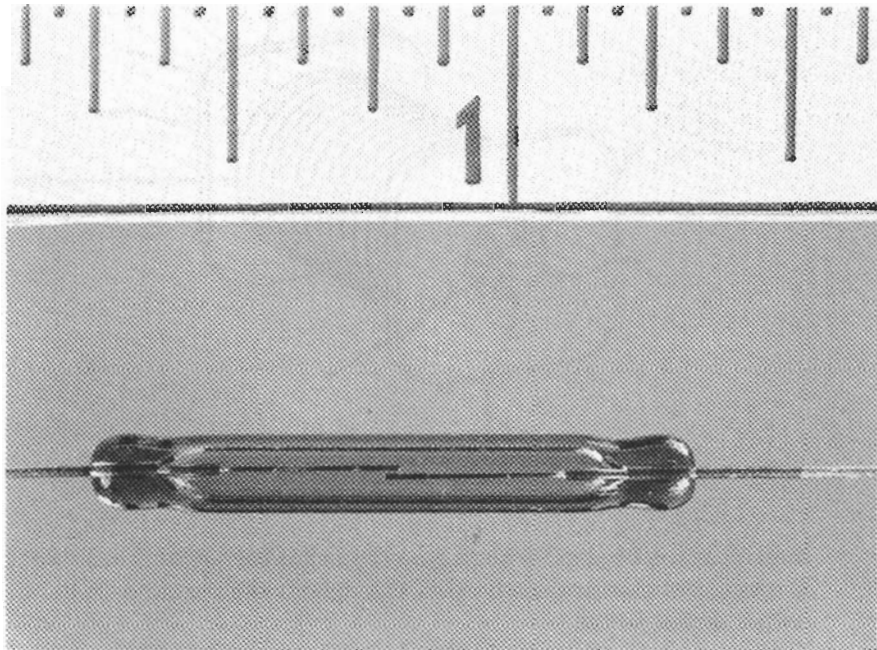


Figure 6.26 A reed relay, showing the glass encasement and contacts. This reed would be inserted into a bobbin magnet, which when activated will induce a magnetic field in the contacts, making them close. The size is around 1 inch (3 cm). (Courtesy of British Telecom)

## 6.9 DIGITAL SWITCHING

Digital switching differs considerably from the other three types of switching previously discussed in this chapter. All of the three types discussed earlier are what is called *space switching* techniques. A space switch connects and disconnects physical contacts using a

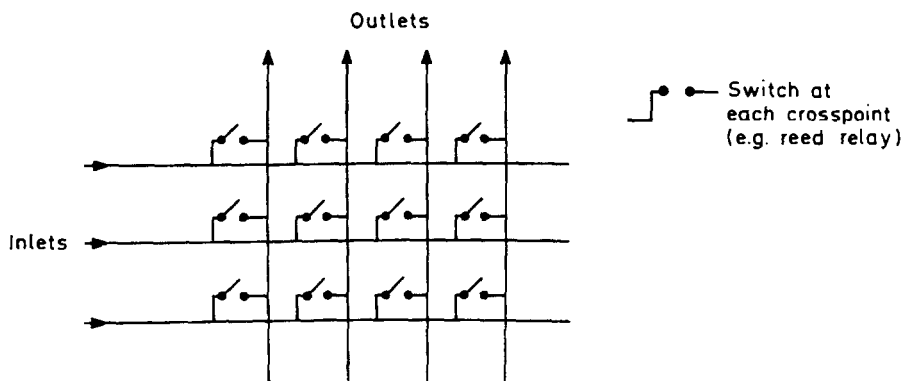


Figure 6.27 Crosspoint matrix of reed relays

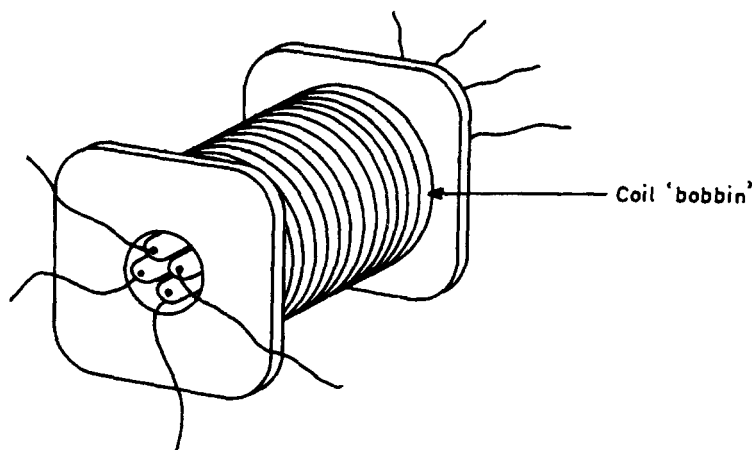


Figure 6.28 Four contact cross-point reed relay

matrix of switchpoints. When a connection has been established through a space switch, a permanent electrical path exists throughout the duration of the call. This is not so with a digital switch.

Most digital switches use instead a *time-space-time (TST)* switch architecture; they are not simple space switches. The need for *time switching* as well as *space switching* arises from the fact that the line systems connected at the periphery of the switch are not individual circuits; they are usually either 2 Mbit/s (32 channels) or 1.5 Mbit/s (24 channel) digital line systems.

As we saw in Chapter 5, a digital line system carries either 24 or 32 channels in a multiplexed form, which means that the total bit-stream carried on the system has been created by interleaving 8-bit samples of each of the constituent channels. Thus the incoming bit stream to a digital exchange appears as shown in Figure 6.29. The first eight bits are channel 0, or timeslot 0; the next eight bits are channel 1, or timeslot 1; right up to channel 23 or channel 31 as appropriate. Figure 6.29(a) shows the 2 Mbit/s line system frame structure, and Figure 6.29(b) shows the frame structure of 1.5 Mbit/s line systems.

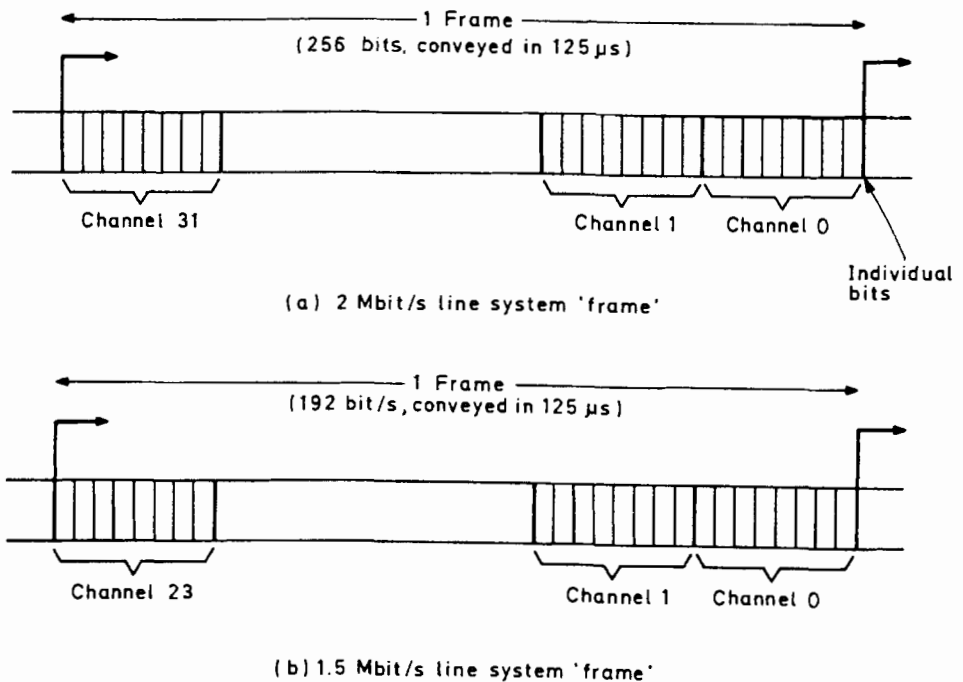


Figure 6.29 Frame structure of digital line systems

It is no good merely to space switch the digital line systems, as this would have the effect of switching all the individual 24 or 32 channels through onto the same outgoing line system. We need instead to be able to connect any channel (or *timeslot*) of one incoming digital line system onto any channel (or *timeslot*) of any of the other digital line systems which are connected to the same exchange. To do this we need both a *time-switch* capability to shift channels between timeslots, and a *space-switching* capability to enable different physical outgoing line systems to be selected.

Figure 6.30 shows a digital switch using a time-space-time architecture to switch the timeslots of three 2 Mbit/s digital line systems. On the left of the diagram the *incoming* or *receive* timeslots are shown. On the right-hand side, the corresponding *outgoing* or *transmit* timeslots of the same 2 Mbit/s line systems are shown. (Recall that any four-wire transmission system, including all digital line systems, comprises the equivalent of two pairs of wires, one pair each for the receive and the transmit directions of transmission).

The incoming (receive) line pairs are fed directly into a timeswitch, the output of which feeds the space switch (shown in the middle). The output of the space switch feeds another timeswitch to which the outgoing (transmit) line pairs are connected.

Imagine now that we wish to switch timeslot 2 in line system A to timeslot 31 in line system B. The space switch allows us to connect line system A to line system B. Point (2) of the space switch allows the receive pair from line system A to be connected to the transmit pair of line system B, while point (6) caters for the other direction of transmission (receive from line system B, and transmit on line system A). However, this by

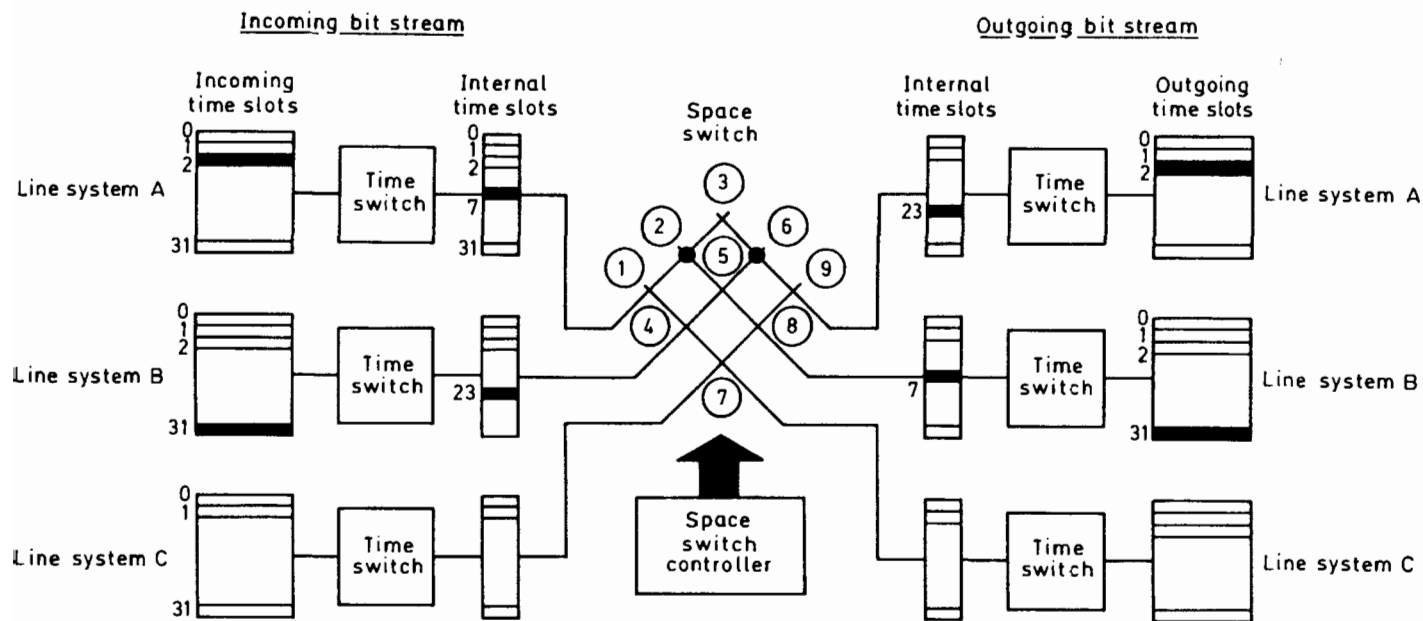


Figure 6.30 A simple time-space-time digital switch

itself is not enough because it would mean that timeslot 2 in line system A ended up in timeslot 2 of line system B, and not in the desired timeslot 31 of line system B. The two time switches allow this conveyance of bits between timeslots. A timeswitch works by storing the received pattern of 8-bits from an incoming *timeslot*, and by waiting for anything up to a whole frame (i.e. anything up to 125 microseconds) in order to feed the pattern out into any desired outgoing timeslot. For example, delaying an 8-bit pattern by  $125/32$  microseconds will move the bit pattern from one timeslot of a 2 Mbit/s system to the next, for example from timeslot 2 into timeslot 3. Similarly, a delay of  $2 \times 125/32$  microseconds will move timeslot 2 to timeslot 4 and so on. Finally, a delay of  $31 \times 125/32$  microseconds will move timeslot 2 into timeslot 1 of the next frame (Figure 29).

Why do we need 2 stages of timeswitching in Figure 6.30? For an answer, let us suppose we want not only to switch timeslot 2 of line system A into timeslot 31 of line system B, but also to switch timeslot 3 of line system A into timeslot 31 of line system C. Then, if we only had one time switch (say the one connected to the incoming line pairs), we would end up trying to superimpose both timeslot 2 and timeslot 3 on the receive side of the timeswitch onto timeslot 31 before switching the space switch across to its output side. To get around this problem, the space switch uses its own internal timeslots.

When asked for a connection between incoming and outgoing time slots, the switch control system carries out a search for free internal time slots on both the receive and transmit sides of the space switch matrix, and when an idle time slot (number 7 in the case illustrated) has been found, switching can begin. The incoming timeslot, number 2 of line system A, is timeswitched into internal timeslot 7. Thereafter whenever internal timeslot 7 comes up, contact (2) of the space switch is activated momentarily and switches the timeslot content towards the transmit pair of line system B. Finally the contents of internal timeslot 7 are time-switched again, this time into timeslot 31 on the transmit pair of line system B. Meanwhile the configuration of the space switch is being changed by the switch control system every 125 microseconds. Necessarily so, as not all incoming timeslots of line system A are to be connected to the same outgoing line system (B in this case). The space switch must be ready to send each incoming timeslot of a single line system into various timeslots of up to 32 outgoing line systems.

At the same time, internal timeslot 23 (the antiphase timeslot of timeslot 7, i.e. the timeslot + 16 timeslots) is being used in conjunction with space switch contact (6) to convey the incoming bit pattern on timeslot 31 of line system B to timeslot 2 on the transmit pair of line system A.

The antiphase timeslot is the timeslot half of the frame after the first allocated internal timeslot. By always using the antiphase timeslot, we guarantee that the return path will never suffer internal timeslot congestion, and furthermore we save the switch control system the effort of searching for a second free internal timeslot. The same method of time-space-time switching can be used in conjunction with both 1.5 Mbit/s and 2 Mbit/s line systems. In practice the time-switches are rarely designed to operate at such low rates, but instead work simultaneously on a number of such line systems which have been previously multiplexed together. Thus rather than running at 1.5 Mbit/s or 2 Mbit/s (24 or 32 channel rates), timeswitches are often designed to run at 512 channel or an even higher rate.

It is worth observing that the whole switch operation can in fact be carried out with only a time switch, and no space switch at all. This is done by multiplexing all the incoming line systems together into a very high bit stream of interleaved 8-bit channel patterns; input to the switch is then from one source only and there is no need for a



space switch. For small exchanges this single timeswitch configuration (without space switch) is practicable, and indeed some digital PBXs do function on a time-switch alone. Large public exchanges must be designed as *time-space-time (TST)* switches, as the high bit rates required to multiplex all the incoming channels together are unattainable with today's electronic technology.

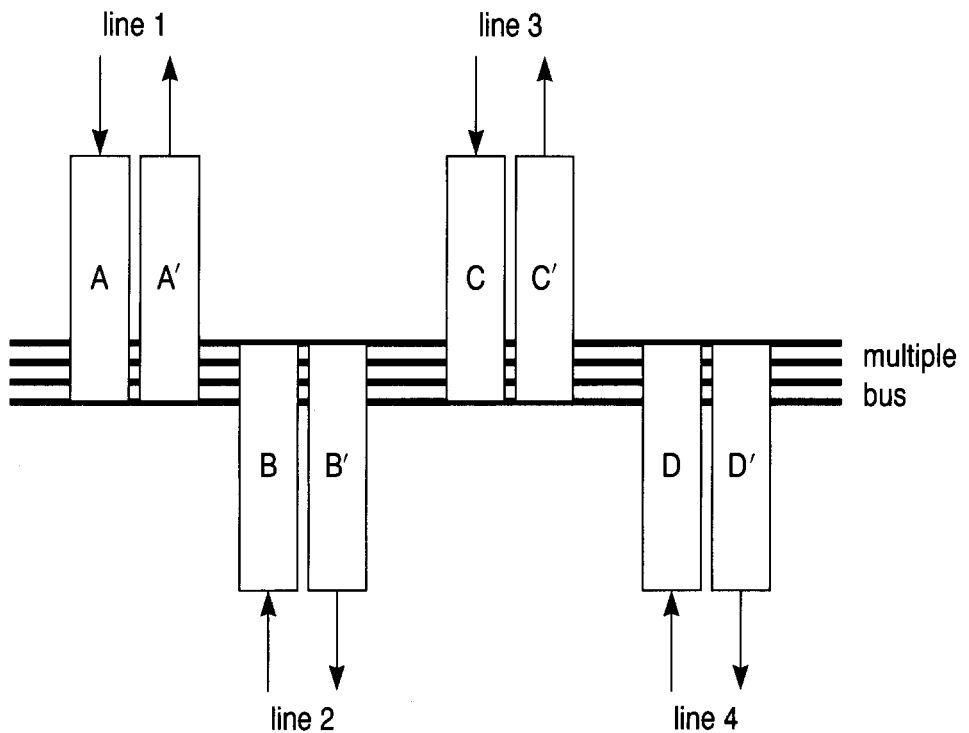
## 6.10 PACKET AND CELL SWITCHES

Packet and cell switching, as we discovered in Chapter 1, differs slightly from circuit switching. Although all three types of switching are *connection-oriented*, only in *circuit-switching* is the path permanently switched between the two endpoints for the entire duration of the connection. During connection *set-up* of a *packet-switched* or *cell-switched* connection, a path is chosen for the connection via all the relevant intermediate switches and a number of identifiers are allocated (e.g. *logical channel numbers*) by which the individual *packets* or *cells* pertaining to a certain connection may be identified during the course of the call. Each packet or cell of user data is sent with a *header* containing this identifier.

The *switching* process stacks incoming packets or cells into buffers, from which they are read out one at a time, in turn. The switch examines the header of each packet or cell and places the whole packet or cell into the appropriate output buffer (according to the *logical channel number* or *connection identifier*). When the packet reaches the front of the output buffer, it is despatched to line, so making its way link by link to the final destination. The process is somewhat akin to a postal sorting process, where at each point the postal bags are reopened, and more finely sorted. When the connection is cleared, the identifier (*logical channel number* or *connection identifier*) are made free for potential use by other new connections. Critical to the good performance of *packet-* or *cell-switching* is the ability for quick processing of packet or cell headers.

Figure 6.31 shows the internal structure of a typical data switching device (for *packet*, *frame* or *cell-switching*). It consists of a *bus*, or multiple busses, connecting together each of the incoming and outgoing ports. Packets received from any of the line inputs are first stored in input buffers. While in the buffer, the switch analyses the address in the packet, frame or cell header (i.e. the LCN, DLCI or VPI/VCI; we learn about these in Chapters 9, 18, 20 and 26), and based on this information decides for which outgoing port or trunk the packet is intended. A hardware port identification known only internally within the switch is usually appended to the front of the packet, and the timeslot on one of the internal busses is allocated to carry the packet to the intended outgoing port.

Each outgoing port permanently monitors the bus for packets bearing its hardware port identification, and relevant packets are removed from the bus and stored temporarily in an output buffer until they are next in turn for outgoing transmission to line. Thus in Figure 6.31 an incoming packet on the first line and intended for the fourth line would be read to the incoming buffer A, where it would be tagged with the hardware identification of D'. Once the allocated slot on the bus is allocated, the packet would be transferred to the outgoing buffer D', where it might need to wait briefly for its turn to be transmitted to line.



**Figure 6.31** Typical internal bus-structure of a packet, frame or cell-switch

The use of multiple busses running between the buffers of the various ports increases the paths available, so minimizing the delay caused by the 'switching' process. In many cases one of the busses acts as a *control bus*. Over the *control bus* information is passed to the various port hardware and associated buffers to regulate their use of the other (traffic-carrying) busses.



---

# Setting Up and Clearing Connections

---

The establishment of a physical connection across a connection-oriented network relies not only on the availability of an appropriate topology of exchanges and transmission links between the two end-points but also on the correct functioning of a logical call set-up and ‘cleardown’ procedure. This is the logical sequence of events for establishing calls. It includes the means by which the caller may indicate the desired destination, the means for establishment of the path, and the procedure for subsequent ‘cleardown’. In this chapter we discuss these ‘call control’ capabilities of connection-oriented networks, and we shall describe the related principles of inter-exchange signalling, and review various standard signalling systems. We commence by considering telephone networks. We move to data networks, and finally to connectionless networks.

## 7.1 ALERTING THE CALLED CUSTOMER

Figure 7.1 shows the very simple kind of communication system which we have considered earlier in this book: two telephones are directly connected by a single pair of wires, without any intervening exchange.

The users of the system in Figure 7.1, A and B, are able to talk at will to one another without fear of interruption. The problem with the equipment illustrated is the difficulty of *alerting* the other party in the first place, to bring him or her to the phone. One easy solution would be to connect a bell at both ends in parallel with each telephone set. If the bell is designed to respond to a relatively high alternating current, whenever such a current is applied from the calling end the bell at the called end rings. This was the earliest form of signalling used on telephone networks. The alternating current (properly called *ringing current*) was applied at the calling end by a manually cranked magneto-electric generator and the technique is known as *generator, bothway generator* or *ringdown* signalling. The term *ringdown* originates from the fact that call clearing in manual exchanges was made by means of the operators ringing one another a second time at the end of the call, hence *ringdown*. Figure 7.2 illustrates a possible though crude adaptation of our network to include generator signalling.

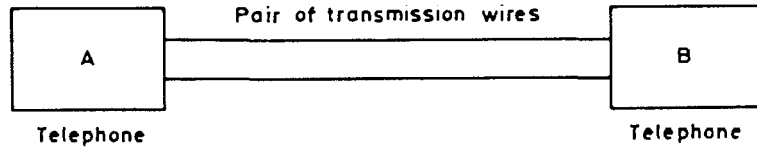


Figure 7.1 A simple communication system

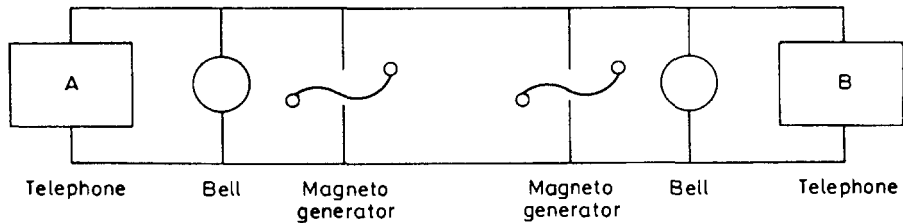


Figure 7.2 A crude 'generator signalling' circuit

Actually, the circuit would normally include an inbuilt contact to disconnect the 'own' telephone from the circuit when the handle was turned, so that it would not also ring.

Magneto-generator signalling was the only form of signalling used in early manual networks. The operator was alerted by use of the generator and was then told by the caller who it was he or she wished to call. The operator then alerted the destination customer (again using the generator) or alternatively referred the call to another operator (if the called customer was on another exchange). Finally the connection was made.

## 7.2 AUTOMATIC NETWORKS

Automatic networks are required to undertake quite a complicated logical sequence of events, first to set up calls, then to make sure they are maintained during conversation (or data transfer), and finally to clear down the connections after use. In support of this sequence of events, *call control* functions are carried out by the exchanges of automatic networks. These functions monitor the state of the call and initiate whatever actions are indicated, such as switching of the connection, applying *dial tone*, performing *number analysis* (to determine the destination), and *signalling* the desired number to a subsequent exchange. Information about the state of the call is communicated from one exchange to another by *signalling systems*, so that automatic connections can be established across a whole string of exchanges. We shall now discuss the sequence of call control functions which makes these things possible.

## 7.3 SET UP

Our first step when making a call is to tell the exchange that we want to do so, by lifting the handset from the telephone cradle or *hook*. (Actually the word hook dates from the

earliest days when the handset was hung on a hook at the side of the telephone. In those days, a horizontal cradle was no good because the early carbon microphones needed gravity to remain compacted.) This sends an *off-hook* signal to the exchange. The signal itself is usually generated by *looping* the telephone-to-exchange access line, thereby completing the circuit as shown in Figure 7.3.

On receipt of the *off-hook* signal the exchange has to establish what is called the *calling line identity (CLI)*, i.e. which particular telephone of the many connected to the exchange has generated the signal. The exchange's control system needs to know this to identify which access line *termination* requires onward cross-connection. This information also serves to monitor customers' network usage, and shows how much to charge them.

One way to identify the calling line is to use its so-called *directory number*, sometimes abbreviated as *DN*. This is the number which is dialled by a customer when calling the line. In early exchanges, particularly the Strowger type, line terminations were arranged in consecutive directory-number order; the functioning of the exchange did not allow otherwise. With the advent of computer-controlled exchanges, it is no longer necessary to use physically-adjacent line terminations for consecutive directory numbers; directory numbers can now be allocated to line terminations almost at random.

For instance it may be convenient that directory numbers 25796 and 36924 should be connected to adjacent line terminations in the exchange. (Such a situation might arise when a customer moves house within the same exchange area, and wishes to retain the same telephone number.) For convenience the line terminations in the exchange can still be numbered consecutively by using an internal numbering scheme of so-called *exchange numbers (ENs or exchange terminations (ETs))*. The extra flexibility that is required for random directory number allocation is achieved by having some form of 'mapping' mechanism of DNs to ENs, as shown in Figure 7.4.

Customers' line terminations are not alone in being given *exchange numbers*; trunks to other exchanges, and even digit sending and receiving equipment can also be allocated an exchange number, depending on the design of the exchange. Allocation of

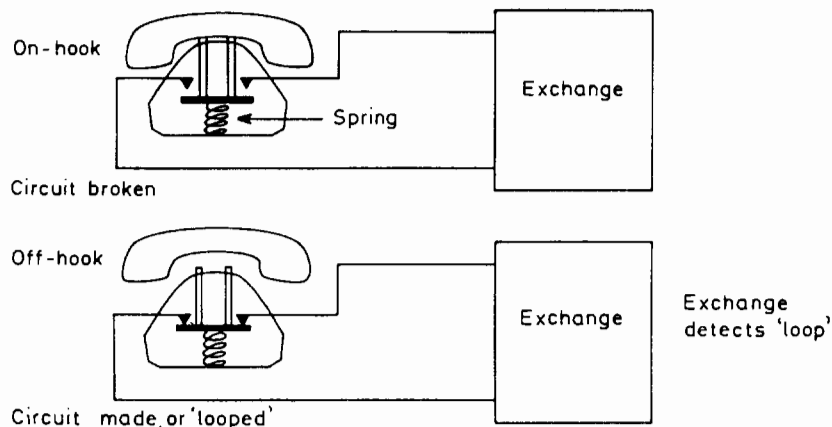


Figure 7.3 The 'off-hook' signal

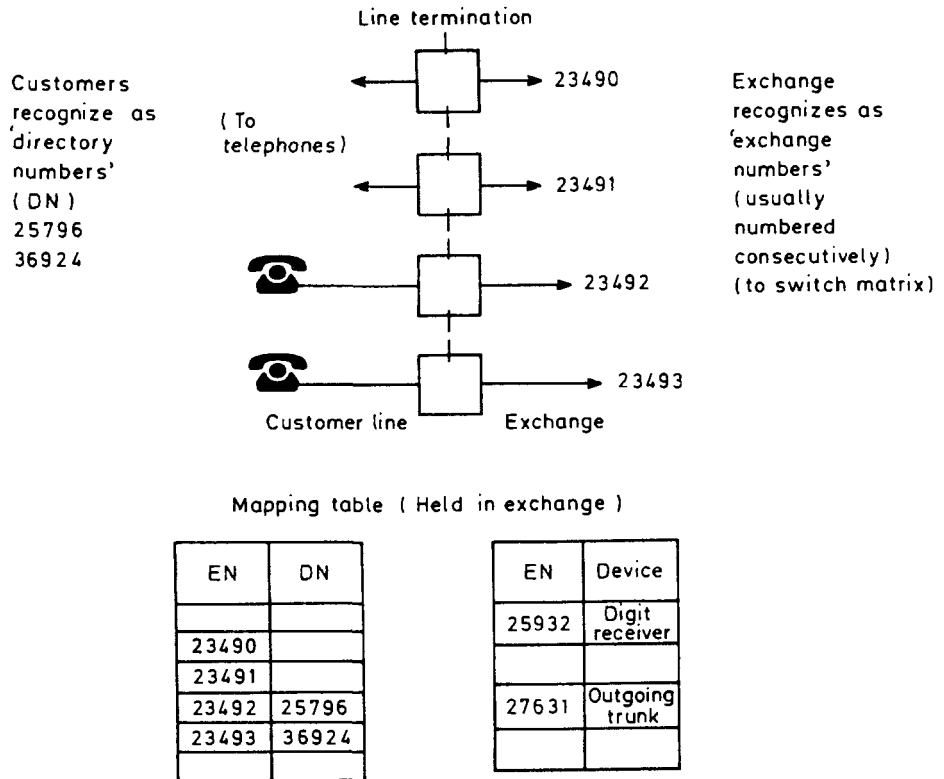


Figure 7.4 'Directory numbers' (DNs) and 'exchange numbers' (ENs)

exchange numbers to all the equipment allows the exchange to 'recognize' all the items which may need to be connected together by the switching matrix. Commands to the switching matrix may thus take the form 'connect EN23492 to EN23493'.

In the example of Figure 7.4 the command 'connect EN23492 to EN23493' would connect the directory numbers 25796 and 36924 together. Another command might connect a line to a digit receiving device. This command would be issued just prior to receiving dialled digits from the customer, at the same time that the dialtone is applied.

Having identified the calling line, the exchange's next job is to allocate and connect equipment, ready for the receipt of *dialled digits* from the customer. This equipment normally consists of two main parts, the *code receiver* which recognizes the values of the digits dialled, and the *register* which stores the received digit values, ready for analysis.

Once the exchange has prepared *code receivers* and a *register*, it announces its readiness to receive digits, and prompts the customer to dial the directory number of the desired destination. This it does by applying *dial tone*, which is the familiar noise heard by customers on lifting the handset to their ear. Because the whole sequence of events (the 'off-hook' signal, the preparation of code-receiver and register, and the return of dial tone) is normally almost instantaneous, dial tone is usually heard by the customer

before the earphone reaches the ear. Noticeable delays occur in exchanges where there are insufficient code receivers and registers to meet the call demand. The remedy lies in providing more of them.

On hearing dial tone, the customer dials the directory number of the desired destination. There are two prevalent *signalling systems* by which the digit values of the number may be indicated to the exchange: *loop disconnect signalling*, and *multi-frequency (MF) signalling*. (Multi-frequency signalling is also sometimes called *dual tone multi-frequency (DTMF) signalling*.)

In *loop disconnect* (or *LD*) signalling, as described in earlier chapters, the digits are indicated by connecting and disconnecting the local exchange access line, or *loop*.

LD signalling was first mentioned in Chapter 2. In Chapter 6 we saw how well it worked with step-by-step electromechanical exchange systems such as Strowger, and we discovered that the pulses themselves could easily be generated by telephones with rotary dials. Both these characteristics have contributed to the widespread and continuing use of LD signalling in customer's telephones.

Modern telephone exchanges also permit the use of an alternative access signalling system, (often the customer may even change the signalling type of his telephone without informing the telephone company). The alternative to LD uses multi-frequency tones (i.e. DTMF). This system has the potential for much faster dialling and call set-up (if the exchange can respond fast enough). It too was discussed briefly in Chapter 2.

DTMF telephones, almost invariably have 12 push buttons, labelled 1–9, 0, \* and #. The two extra buttons \* and # (called *star* and *hash*) are not, however, always used, and their function may vary between one network and another. Where they are used, they often indicate a request for some sort of special service; for example, \*9-58765 might be given the meaning 'divert incoming calls to another number, 58765'.

When any of the buttons of a DTMF (MF4) telephone are pressed, two audible tones are simultaneously transmitted onto the line (hence the name, *dual tone multi-frequency*). The frequencies of the two tones depend on the actual digit value *dialled*. The relationship was shown in Table 2.2 of Chapter 2.

While *dialling* in DTMF the customer hears the tones in the earpiece. Only a very short period of tone (much less than a half-second) is needed to indicate each digit value of the dialled number.

Once the exchange has started to receive digits, it can set to work on *digit analysis*. This is the process by which the exchange is able to determine the appropriate onward

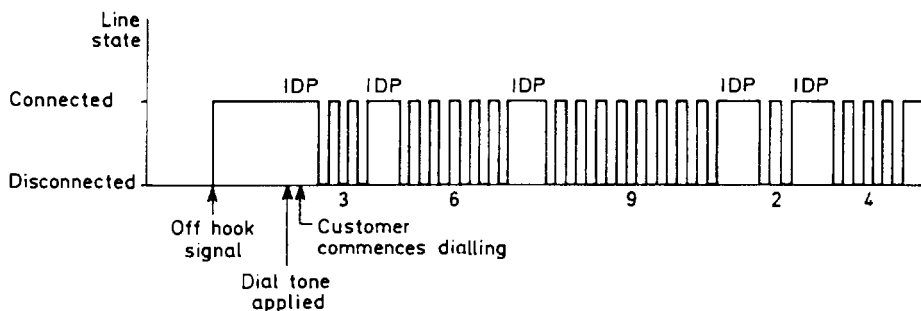


Figure 7.5 Loop disconnect signalling train



routing for the call, and the charge per minute to be levied for the call. During digit analysis, the exchange compares the *dialled number* held by the register with its own list of permitted numbers. The permitted numbers are held permanently in *routing tables* within the exchange. The routing tables give the *exchange number* identity of the outgoing route required to reach the ultimate destination.

Routing tables are normally constructed in a tree-like structure allowing a cascade analysis of the digit string. This is shown in Figure 7.6 where customer 'a' on exchange A wishes to call customer 'b' on exchange B. The number that customer 'a' must dial is 222 6129. The first three digits are the *area code*, identifying exchange B, and the last four digits, 6129, identify customer 'b' in particular. The routing table tree held by exchange A is also illustrated. As each digit dialled by customer 'a' is received by exchange A, a further stage of the analysis is made possible until, when '222' has been analysed, the command 'route via exchange B' is encountered. At this stage the switch path through exchange A may be completed to exchange B, and subsequent dialled digits may be passed on directly to exchange B for digit analysis. The *register*, *code receivers*, and other *common equipment* (i.e. control equipment which may be allocated to aid set-up or *supervise* any part of a connection) in use at exchange A is released at this point, and is made available for setting up calls on behalf of other customers. Exchange B is made aware of the incoming call by a *seizure signal* (equivalent to the off-hook signal on a calling customer's local line). The signal is sent by means of an inter-exchange signalling system which will be discussed in more detail later in the chapter.

Exchange B prepares itself to receive digits by allocating *common equipment*, including *code receivers* and *register*. Then, having received the digits, digit analysis is undertaken in exchange B, again using a routing table. This time, however, the analysis is concluded only after analysing the number '6129', at which point the command 'connect to customer's line' is issued. In principle the method of analysis of the area code '222' and the customer's number '6129' is the same. The difference is that one or more different routes may be available to get from one exchange to another, but for a

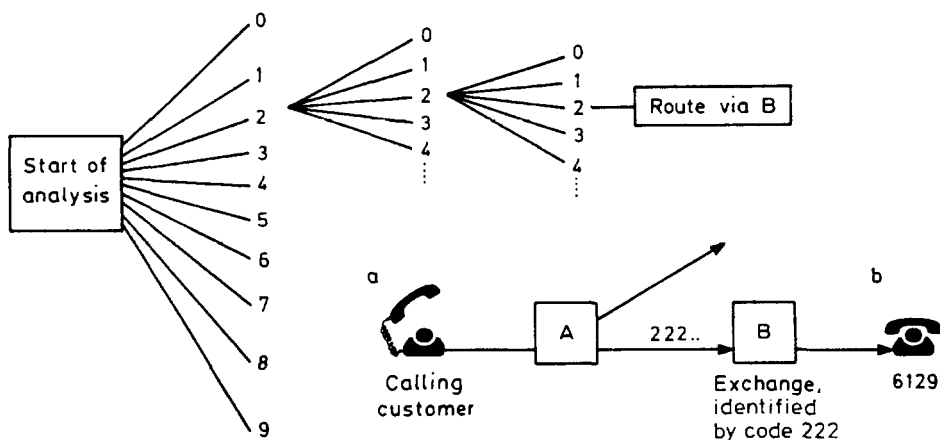


Figure 7.6 Routing table tree

customer's line, of course, only one choice is possible. The former analysis may thus require production or translation of routing digits whereas the latter only involves a line selection.

Once the connection from calling to called customer has been completed, *ringing current* will be sent to ring the destination telephone, and simultaneously ringing tone will be sent back to the calling customer a.

As soon as customer b (the called customer) *answers* the telephone (by lifting the handset), then an *answer signal* is transmitted back along the connection. This has the effect of *tripping* the *ringing current* and the *ringing tone* (i.e. turning it off), and commencing the process of charging the customer for his call.

*Conversation* (or the equivalent phase of communication) may continue for as long as required until the calling customer replaces the telephone handset to signal the end of the call. This is called the *clear signal*, and it acts in the reverse manner to the *off-hook signal*, breaking the access line *loop*. The signal is passed to each exchange along the connection, releasing all the equipment and terminating the call-charging.

Depending on the network and the type of switching equipment, it may be possible for both the calling and the called customers to initiate the *clear-down* sequence. The ability of called customers to initiate clear-down was not prevalent in all early automatic exchanges (for example, UK Strowger exchanges). However in many modern switch types either party may clear the call.

## 7.4 NUMBER TRANSLATION

Modern exchanges use stored *program control* (SPC, actually a computer processor) for the purpose of digit analysis and route determination, often using a routing data tree, as Figure 7.6 showed. The administration needed to support such exchanges and their routing tables is fairly straightforward. In the past, however, particularly in the days of electromechanical exchanges, digit analysis and call routing mechanisms were often very complex. Because they need to be formed out of hard-wired and mechanical components, their efficient operation often demanded slightly different call routing techniques. An important tool in effective call routing was, and still is, the process known as number *translation*.

Number translation is a means of reducing the number of times that digit analysis needs to be undertaken during a call connection. Digit analysis still takes place at the first exchange in the connection, and may have to be repeated at another exchange later in the connection (typically a trunk exchange), but any further digit analysis (at other exchanges) can be minimized by the use of number translation, thereby enabling subsequent exchanges to respond to the received digit string without analysing more than one digit at a time. As we learned in Chapter 6, the direct response of selectors to each digit in turn is crucial for the correct operation of some types of switching equipment (for example Strowger). Number translation also remains in use even in digital networks, for reason of flexibility to change route, signalling system or number length (e.g. abbreviated dialling).

Number translation involves detecting the actual number dialled by the customer and replacing it with any convenient string of digits which will make the operation of the

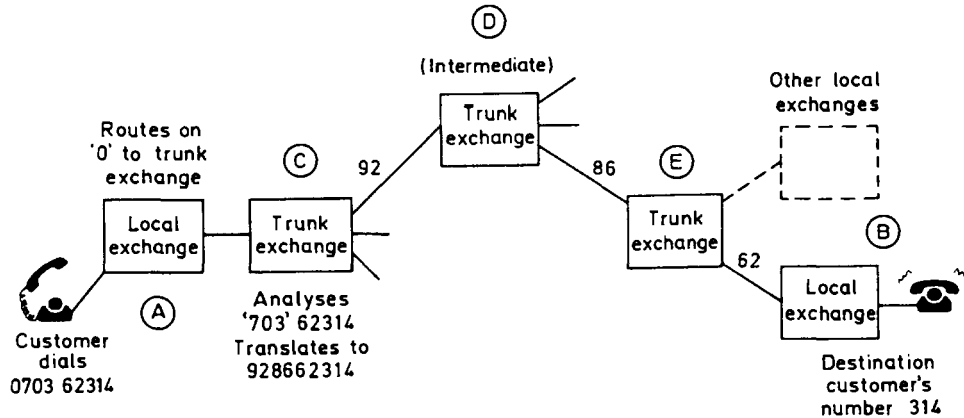


Figure 7.7 Number translation in Strowger networks

network and the connection to the called customer easier. The translated number may thus be entirely unrelated to the dialled number. Figure 7.7 gives a typical example of digit translation. The example illustrates the use of number translation in the United Kingdom *public switched telephone network (PSTN)* in the 1950s, when automatic long distance calling was introduced to the existing Strowger network. In the example, the number dialled by the customer is composed of three parts:

Trunk code '0' + Area code '703' + Customer Number '62314'.

For the same destination area, the same area code is dialled by any *calling customer* in the network, but the route taken by the call will obviously need to take account of the different starting points. Different intermediate exchanges will need to be crossed, depending on whether a particular call has originated from the north or the south of the UK. Different number translations are therefore used in each case. The sequence of events is as explained below.

On receiving the digit string from the calling customer, exchange A recognizes the first digit '0' as signifying a trunk call, and so routes the call to the nearest trunk exchange, passing on all other digits '703 62314', but deleting the '0', which has served its purpose. Trunk exchange C then analyses the next three digits (the area code) '703'. This is sufficient to establish that exchange E is the destination trunk exchange and that the route to be taken is via exchange D. Two options are now available for onward routing, either:

- (a) the call may be routed to exchange D, and the original digit string (70362314) transmitted with it, in which case exchange D will have to re-analyse the area code 703;

or:

- (b) the call may be routed to exchange D, together with a translated number string, thereby easing the digit analysis at D.

Method (a) is more commonly used with stored program control (SPC) exchanges, as it affords greater flexibility of network administration. This is because routing changes

can be made at individual exchanges without altering the *translated number* which the previous exchange must send. The disadvantage of this method is that it requires more digit analysis.

In different cases, when exchange D is of Strowger type say, or when digit analysis resources at exchange D are short, it may be important to use translated number method (b) to minimize digit analysis after exchange C. This method is described in detail below. Either method may be used, but whichever is chosen, one method usually prevails throughout the network; it is rare to find both methods simultaneously in use in the same network.

Method (b) above is the translation method. In the Figure 7.7 example the digits '92' are required by the Strowger selectors in exchange C, to select the route to exchange D. Similarly the digits '86' select the route to exchange E from exchange D. Because exchange C translates the area code '703' into the *routing digits* (9286) required by exchanges C and D, no further digit analysis will be required at either the intermediate trunk exchange D or the destination trunk exchange E. Instead the Strowger selectors absorb and respond directly to the digits. Thus even exchange C will *step* its Strowger selectors by using and absorbing the digits '92', and exchange D will do the same with the digits '86'. By the time the call reaches exchange E, only digits 62314 remain. Exchange E uses digits 62 to select the appropriate local exchange and routes the call to exchange B, the destination local exchange, where digits 314 identify the actual customer's line which will then be rung.

Translation is also used in more modern networks as a way of providing new and special services. Chapter 11, on *intelligent networks*, will describe the *freephone* or *800-service* where the calling customer dials a specially allocated '800' number rather than the actual directory number of the destination. Calls made to an 800 number (e.g. 0800 800800) are charged to the account of the destination number (i.e. to the account of the person who rents the '800' number) and are therefore free to calling customers. Although each '800' number is unique to a particular destination customer, it is not recognized by the network itself for the purpose of routing and must therefore be translated into the actual directory number of the destination. Let us imagine in Figure 7.7 that the number 0800 80800 has also been allocated to the destination customer shown, so that customers can dial either '0800 80800' or '0703 62314'. Calls to the first of these numbers are free to the caller; calls to the second will be charged. (In the former case, it is the renter of the '800' number who will be charged). In both cases, however, the call needs to be routed to the real directory number, 0703 62314. In the former case this is done by number translation, setting up a bill for the '800' account on the way.

## 7.5 UNSUCCESSFUL CALLS

We have run through the sequence of events leading up to *successful call*, when the caller gets through and a conversation follows. However, as we all find out, calls do not always succeed, and when a call fails, perhaps because of network congestion, or because the called party is busy or fails to answer, the network has to tell the caller what has happened, and then it has to clear the connection to free the network for more fruitful use.

When it is a case of network congestion or called customer busy, the caller usually hears either a standard *advisory tone*, or a *recorded announcement*. Telephone users will be miserably familiar with *busy tone* and recorded announcements of the form ‘all lines to the town you have dialled are busy; please try later’, to say nothing of the number *unobtainable tone* which tells us that we have dialled an invalid number.

A caller who hears one of the call unsuccessful advisory announcements, or a prolonged ringing tone, usually gives up and clears the connection by replacing the handset, to try again later. When, however, the caller fails to do this, the network has to force the release of the connection. *Forced release*, if needed, is put in hand between 1 and 3 minutes after the call has been dialled, and it is initiated if there has been no reply from the called party during this period, whatever the reason. Forced release once initiated, normally by the originating exchange even though the handset of the calling telephone is left off-hook, forces the calling telephone into a *number unobtainable* or *park* condition. To normalize the condition of a telephone, the handset must be returned to the cradle.

Release is also forced in a number of other ‘abnormal’ circumstances, as when a calling party fails to respond to dial tone (by dialling digits), or when too few digits are dialled to make up a valid number. In such cases the A-party (i.e. the calling) telephone may be entirely disconnected from the local exchange, silencing even its dial tone. This has the advantage of freeing code receivers, registers, and other common equipment for more worthwhile use on other customers’ calls. The calling customer who is a victim of forced release may hear either silence or *number unobtainable tone*. To restore dial tone, a new *off-hook* signal must be generated by replacing the handset and then lifting it off again.

It is clearly important for exchanges to be capable of *forced release* so that their common equipment is not unnecessarily ‘locked up’ by a backlog of unsuccessful calls.

## 7.6 INTER-EXCHANGE AND INTERNATIONAL SIGNALLING

Inter-exchange signalling is the process by which the destination number and other call control information is passed between exchanges with the object of establishing a telephone connection. Inter-exchange signalling systems have come a long way since the early days of automatic telephone switching, when ten pulse-per-second, *loop disconnect (LD)* and similar, relatively simple, *signalling systems* were the fashion. Today, many different types of inter-exchange signalling are available, and which type a particular network will use depends on the nature of the services it provides, the equipment it uses, its historical circumstances, and the length and type of the transmission medium. For example, small local networks may use relatively slow and cheap signalling systems. By contrast, in extensive public international networks, or in private networks connected to public international networks, consideration is needed for the greater demands, for faster signalling, longer numbers and greater sophistication. We can now review the signalling systems defined in ITU-T recommendations (see Table 7.1), with particular attention to the system called R2. This is a *multifrequency code (MFC)* inter-exchange signalling system, in some ways similar to the *DTMF* signalling system used for customer dialling, but far more sophisticated.

**Table 7.1** CCITT signalling systems

Signalling type	Description
CCITT 1	Now obsolete signalling system intended for manual use on international circuits. A 500 Hz tone is interrupted at 20 Hz for 2 seconds.
CCITT 2	Never implemented. CCITT 2 was a 2 Voice Frequency (VF) tone system, using 600 Hz and 750 Hz tones for 'line signalling'* and dialling pulses respectively. It was the first system for international automatic working.
CCITT 3	Now obsolete, system designed for manual and automatic operation using a 1 VF tone at 2280 Hz for both line signalling* and inter-register‡ signalling. Inter-register signalling using a binary code at 20 baud.
CCITT 4	Intra-European signalling system still using for automatic and semi-automatic use. Line signalling* using 2040/2400 Hz (2 VF) code. Inter-register‡ signalling using the same 2 VF tones; each digit comprising four elements, transmitted at 28 baud. (2040 Hz = binary 0; 2400 Hz = binary 1).
CCITT 5	System designed and still used for intercontinental operation via satellite and using circuit multiplication equipment (Chapter 38 refers). Line signalling* using 2400/2600 Hz (2 VF) code. Multifrequency (MF) inter-register‡ signalling, each digit represented by a permutation of two of six available tones.
CCITT 5 (bis)	Never used; a compelled version of CCITT 5.
CCITT 6	Common channel signalling system intended for international use between analogue SPC (stored program control) exchanges. Signalling link speed typically 2.4 kbit/s.
CCITT 7	Common channel signalling system intended for widespread use between digital SPC exchanges. Multifunctional with various different 'user parts' for different applications (see Chapter 12). Signalling link speed 64 kbit/s.
CCITT R1	Regional signalling system somewhat akin to CCITT 5 and formerly used particularly for trunk network signalling in North America.
CCITT R2	Regional signalling system used widely within Europe. Described fully later in this chapter.
CCITT R2D	Digital version of R2. Adapted particularly for use after the European Communications Satellite (ECS or 'Eutelsat').

\* Line signalling and ‡ inter-register signalling are described more fully later in this chapter.

All the systems listed in the table are for inter-exchange use. In fact they are all ITU-T standard systems designed for international use. Each of them can be used by one exchange to establish calls to another exchange, but they vary considerably in sophistication. Signalling system number 1, a simple system enabling operators in different manual exchanges to call one another up, has been described previously. In the course of time signalling system number 2 to signalling system number 7 (SS7) plus the R1 and R2 signalling systems were developed. Each tends to be slightly more sophisticated than its predecessor and therefore better attuned to the developing technologies of switching

and transmission. The most advanced of the ITU-T signalling systems developed to date is SS7. *A common channel signalling system* (this term is explained later), SS7 has a number of powerful network control features and is capable of supporting a wide range of advanced services.

The ITU-T standard signalling systems are only a small subset of the total range available, but they are the most suitable for international inter-connection of public telephone networks because they are widely available. Other systems have evolved either as national standards or have been developed specially for particular applications. Signalling systems in general can be classified into one of four different classes according to how the signalling information is conveyed over the transmission medium, as follows.

#### *Direct current (DC) signalling systems*

These use an on/off current pulse or vary the magnitude and polarity of the circuit current to represent the different signals; *loop disconnect (LD)* is an example. It has the disadvantage that it will only work when there is a distinct set of wires for each channel (i.e. on *audio* or *baseband* lineplant), and it is therefore only suitable for short ranges. DC signalling is not possible on either FDM or TDM lineplant, though the on/off states can be mimicked using speechband or voice frequency (VF) tones over FDM (or TDM) (example pulse on/off = tone on/off) or alternatively over TDM by crudely converting the pulse on/off into strings of binary 1s and 0s. The advantage of DC signalling (when possible) is its cheapness.

#### *Voice frequency (VF) signalling systems*

VF signalling is the name given to single or two-tone signalling systems (otherwise *1VF* and *2VF*). As stated above, these are similar to DC signalling systems, merely mimicking pulse 'on' and 'off' (and varying lengths and combinations of same (with tone on/tone off conditions)). The principal engineering problems associated with VF signalling systems arise from the difficulty in keeping speech frequencies and signalling tones logically separate from one another while sharing the same circuit. Signalling system number 4 is an example of a 2VF system.

#### *Multifrequency code (MFC) signalling systems*

These use tones within or close to the frequencies heard in normal speech to represent the signalling information. The advantage of MFC signalling is that it can easily be carried over FDM or TDM lineplant, the tones being processed through the multiplexor in exactly the same way as the speech frequencies. Thanks to their compatibility with FDM and TDM lineplant this type of signalling system has become very common in trunk and international networks. Signalling system number 5 (CCITT5) and R2 are examples of MFC signalling systems.

#### *Digital signalling systems*

These code their signalling information in an efficient binary code format, each byte of information having a particular meaning. This type of signalling is therefore ideal for carriage over TDM lineplant. Typically, timeslot 16 in the European 2 Mbit/s digital transmission system is reserved for digital signalling, whereas in the 1.5 Mbit/s system either an entire 64 kbit/s channel or a *robbed bit* channel is used. Alternatively, the

signals can be encoded using a *modem* (as described in Chapter 9) to make them suitable for carriage over FDM lineplant. Examples of digital signalling systems are SS6, SS7 and (in part) R2D.

R2 is typical of signalling systems in analogue network usage today, and as it provides a useful introduction to the principles of inter-exchange call control and multi-frequency signalling we shall discuss it next, returning to SS7 in Chapter 12.

## 7.7 THE R2 SIGNALLING SYSTEM

The R2 signalling is typical of the many *multi-frequency code (MFC)* systems used in the networks of the world. R2 is one of ITU-T's two 'regional' systems, and is used extensively within and between the countries of Europe. It comprises two functional parts, an *outband line signalling* system, together with an *inband* and *compelled sequence MFC inter-register signalling* system (the new terms are explained later in this section).

R2 may be used on international as well as national connections, but as there are significant differences between these two applications, it is normal to refer to the variants as if they were two separate systems, *International R2* and *National R2*.

R2 is a *channel-associated* signalling system. By this we mean that all the signals pertinent to a particular channel (or circuit) are passed down the circuit itself (in other words are *associated* with it). By contrast, the more modern *common channel signalling* systems use a dedicated *signalling link* to carry the signalling information for a large number of traffic carrying circuits. The traffic carrying (i.e. speech carrying) circuits take a separate route, so that speech and signalling do not travel together. Figure 7.8 illustrates the difference between *channel-associated* and *common channel* signalling systems.

In *channel associated signalling* systems (exchanges A and B of Figure 7.8), a large number of code senders and receivers are required, one for each circuit. By contrast, in common channel systems (exchanges C and D of Figure 7.8), a smaller number of

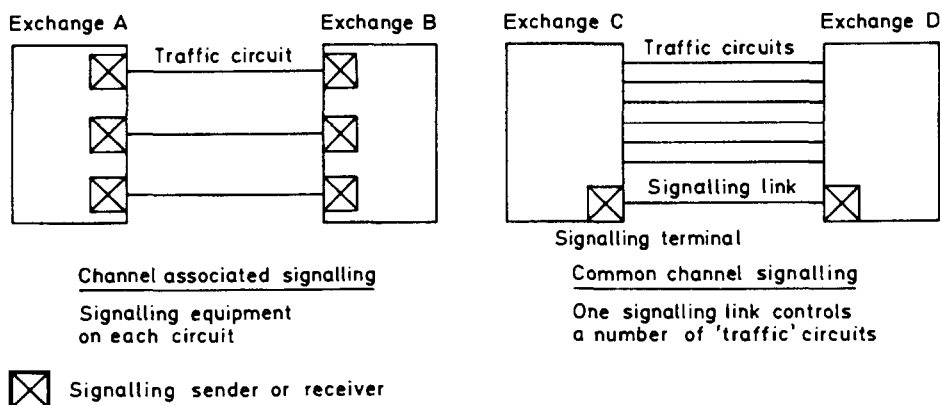


Figure 7.8 'Channel-associated' and 'common channel' signalling method



so-called *signalling terminals (ST)* are required. Examples of *channel associated signalling* systems are *loop disconnect (LD)* and *R2*. Common channel signalling systems include *SS6* and *SS7*.

## 7.8 R2 LINE SIGNALLING

Multi-frequency, channel-associated signalling systems nearly always have two parts, the *line signalling* part and the *inter-register signalling* part, each with its own distinct function. The *line signalling* part controls the line and the common equipment; it also sends *line seizures* (described earlier), and other *supervisory signals* such as the *clear-down* signal. The inter-register signalling part carries the *information*, such as *number dialled*, between exchange registers. Splitting channel-associated signalling systems into two parts in this way helps to minimize the overall number of signalling code senders and receivers required in an exchange, as we shall see.

The line signalling part is usually only a single or two-frequency system. In R2, it is a single tone (1VF), out-of-band system. *Out-of-band* means that the frequency used is *outside* the (3.1 kHz) *bandwidth* which is made available for conversation; the frequency, nonetheless, lies within the overall 4 kHz bandwidth of the circuit. Figure 7.9 illustrates the *inband* and *out-of-band (outband)* ranges of a normal 4 kHz telephone channel.

We learned about telephone circuit bandwidth in Chapter 3, and how 4 kHz of bandwidth is allocated for each individual channel on an FDM system, but that only the central 3.1 kHz bandwidth is used for conversation. The unused bandwidth provides

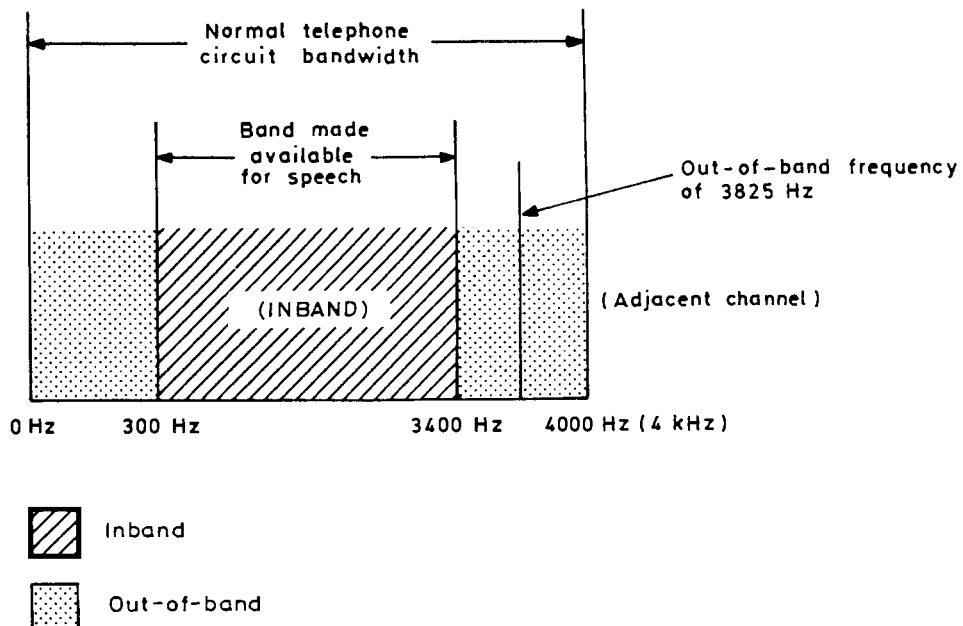


Figure 7.9 'Inband' and 'out-of-band' signals

for separation of channels as a means of reducing the likelihood of adjacent channel interference. Figure 7.9 illustrates the relationship, showing the total 4 kHz bandwidth actually allocated for the telephone circuit, and the 3.1 kHz range from 300 Hz to 3400 Hz which is available for speech. The ranges 0–300 Hz and 3400–4000 Hz are normally filtered out from the original conversation (with little customer-perceived disadvantage) and give the bandwidth separation between channels. In short, *out-of-band* means a frequency in the range 0–300 Hz or 3400–4000 Hz, and *in-band* frequencies are those in the range 300–3400 Hz.

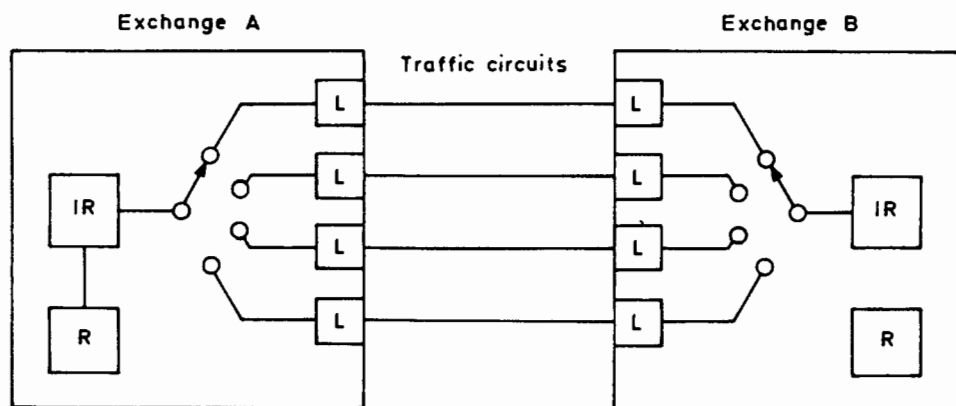
The use of an out-of-band signal (rather than an in-band one) for R2 line signalling has two advantages: first, it does not disturb the conversation (without affecting the channel separation); second, line signals cannot be sent fraudulently by a telephone customer because the *out-of-band* region is not accessible to the end user. Despite this advantage, some other signalling systems do use in-band line signalling. The advantage of using inband tones lies in simplifying the circuit configuration through FDM multiplexors.

The line signalling part of the signalling system is the only part of the signalling which is always active. It is the line signalling that actually controls the circuit. From the circuit idle state, it is the line signalling part that *seizes* the circuit (alerting the distant exchange for action), and in so doing will activate the inter-register signalling part. The *seizure* involves making a register ready at the distant (incoming) exchange, as well as activating appropriate inter-register signalling *code senders* and *receivers*. *Seizure* will be followed by a phase of inter-register signalling, to convey dialled number and other call set-up information between the exchanges; at the end of this phase the inter-register signalling equipment and register will be released for use on other circuits, but the line signalling will remain active. The line signalling has more work to do in detecting the answer condition (to meter the customer), and at the end of the call it must carry the signals necessary to clear the connection and stop the metering. Even when the call has been cleared, both exchanges must continue to monitor the line signalling to detect any subsequent call seizures.

It is because the line signalling part is always active that it is normally designed as a single or two-tone system. This reduces the complexity and cost of signalling equipment that has to be permanently active on each and every circuit. The inter-register signalling part is only used for a comparatively short period on each call, during call set up. It is necessarily more complicated because of the range of information that it must convey, but a small number of *common equipment* (code senders, code receivers and registers) may be shared between several circuits. This is cheaper than employing an inter-register sending and receiver with every circuit. The equipment is switched to an active circuit in response to a line signalling seizure, as already outlined. Figure 7.10 illustrates the inter-relationship of line and inter-register signalling parts.

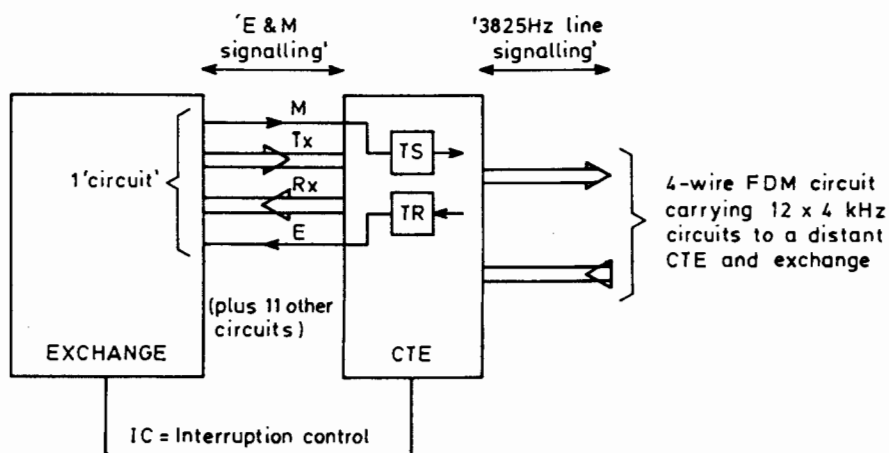
The line signalling part of R2 operates by changing the state of the single frequency (3825 Hz) tone, from on to off and vice versa. When the circuit is not in use, a tone of 3825 Hz can be detected in both *transmit* and *receive* channels of the circuit.

When R2 signalling is used directly on audio circuits (i.e. unmultiplexed analogue circuits) the line signalling tone sender and receiver is located in the exchange termination. When FDM (frequency division multiplex) is used on the circuits (more common), the tones themselves are usually generated in the FDM *channel translating equipment* (CTE), otherwise the tones would be filtered out together with other signals outside the normal speech.



**Figure 7.10** Channel-associated signalling: line and inter-register signalling parts. L, line signalling equipment (always active on each circuit); IR, inter-register signalling code senders and receivers (shared); R, register (for storing and analysing call set-up information) (shared)

Two further wires connect the exchange to the CTE, and enable the exchange to control and monitor the state of tones on incoming (receive) and outgoing (transmit) channels, whether on or off. These extra leads are called the E&M leads (hence the common term *E&M signalling*). In total therefore, each circuit between the exchange and the CTE comprises six wires: a transmit pair, a receive pair, plus E-wire and M-wire. The M-wire controls the tone state on the transmit channel, activating the CTE to send a tone when the M-wire is at high voltage, and not sending a tone when the M-wire is earthed. Similarly, the CTE conveys information concerning the state of the tone on the receive channel by using the E-wire; high voltage means there is a tone on the receive channel, earth (zero voltage) means there is not. A useful way of remembering the roles of the E and M wires is



**Figure 7.11** Typical configuration for R2 signalling. Tx, Transmit pair; Rx, Receive pair; IC, Interruption control wire; TS, 3825 Hz tone sender; TR, 3825 Hz tone receiver; CTE, channel translating equipment

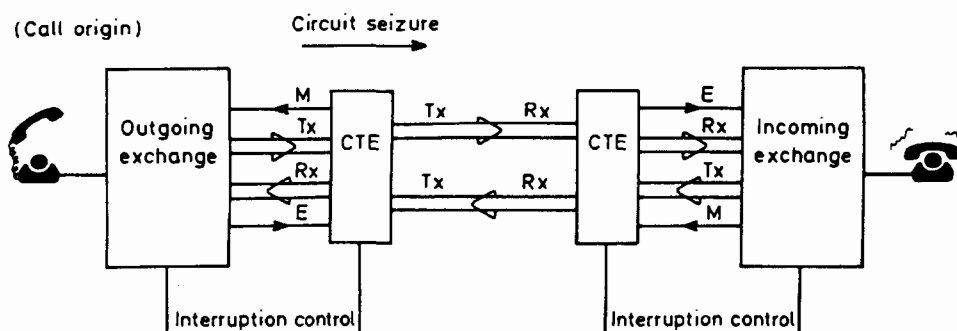


Figure 7.12 'Incoming' and 'outgoing' exchanges of an R2 controlled circuit

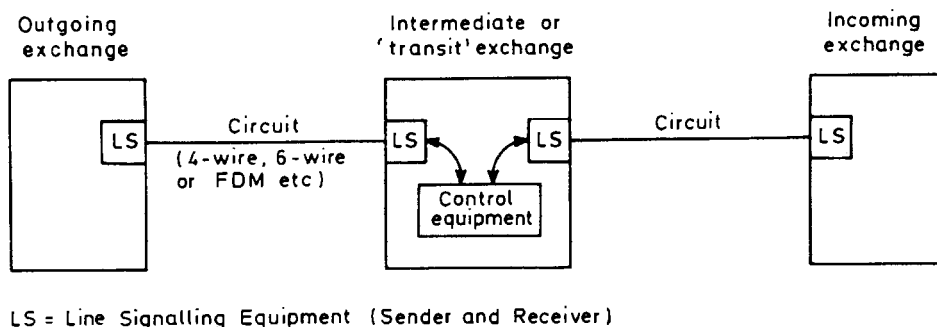
to make the association E = Ear; M = Mouth. In addition to the 72 wires needed for the 12 channels, an extra wire is provided for *interruption control* (IC), so that if the FDM carrier fails the CTE does not immediately seize all 12 circuits. Figure 7.11 shows a typical arrangement of exchange and CTE where R2 signalling is being used in conjunction with FDM lineplant.

From the tone-on-idle state (i.e. circuit not in use, with 3825 Hz tones passing in both forward and backward directions), the tones may be turned 'off' and 'on' in sequence to indicate the different stages of a call: set up, conversation and clear-down. We next discuss the signalling sequence and describe the terms *incoming* and *outgoing* exchange. The *outgoing* exchange is that originating the call. It effects the *seizure* of an idle circuit by turning 'off' the forward signal tone as shown in Figure 7.12. The *incoming* exchange (the one which did not generate the seizure) responds by allocating *common equipment*, including a register and inter-register signalling equipment.

In signalling systems other than R2 the readiness to receive digits is indicated at this stage by returning a *proceed-to-send* (PTS) signal. On receipt of the PTS the outgoing exchange sends the dialled number digits by changing over to *inter-register signalling*. As R2 has *no PTS signal* the change over to *inter-register signalling* is unprompted and the digit is sent anyway. It continues to be sent until it is *acknowledged* by the incoming end exchange, a technique known as *compelled signalling*. The call set up continues with similar inter-register and line signal interchanges. Table 7.2 shows the entire line-signalling sequence.

Table 7.2 R2 line signalling sequence. (Courtesy of ITU – derived from Table 1/Q411)

State no:	Forward signal tone	Backward signal tone	Line signal meaning	Moves to state
1	Tone-on	Tone-on	Circuit idle	2 or 6
2	Tone-off	Tone-on	Seized (by outgoing end)	3 or 5
3	Tone-off	Tone-off	Answered-conversation	4 or 5
4	Tone-off	Tone-on	Clear back	3 or 5
5	Tone-on	Tone-on or Tone-off	Release forward	1
6	Tone-on	Tone-off	Blocked (at incoming end)	1 (when unblocked)



**Figure 7.13** 'Link-by-link' operation of R2 line signalling

When the *inter-register signalling* has conveyed the necessary number of dialled digits to allow the incoming exchange to decide its action, then the connection can be assumed to be made through the incoming exchange. If a further onward link is necessary (e.g. to another trunk or international transit exchange) then either the *incoming exchange* changes its role to that of an *outgoing exchange* (for signalling purposes) or else the *outgoing exchange* retains its role and signals transparently through the previous *incoming exchange* (now switched through) to the new incoming exchange (next link). The former method is called *link-by-link signalling*, the latter *end-to-end signalling*.

The line signalling part of R2 works in a *link-by-link* mode. In other words, on a connection comprising a number of links in tandem the line signalling on each link of the connection works in an independent and sequential mode. The *clear* (or *clear-down*) signal for instance does not pass straight through from one end of the connection to the other; it is passed one link at a time (*link-by-link*), and is interpreted by the control equipment of each exchange and passed on as necessary. Independent line signalling equipment is therefore required on every link of a tandem connection, as shown in Figure 7.13.

The *link-by-link* operation of the line signalling part is in contrast to the *end-to-end* operation of the inter-register signalling part, as we shall shortly see.

## 7.9 COMPELLED OR ACKNOWLEDGED SIGNALLING

R2 line signalling is said to be a *compelled* system, meaning that each signal is sent and continues to be sent, whether forward or backward, until a signal is received from the opposite end, the return signal acting as an acknowledgement and a prompt for the next action. This ensures receipt of the signal and readiness for the next. Thus the acknowledgement of the first digit sent in the forward direction prompts the outgoing exchange to send the next. Until the signal is acknowledged the outgoing exchange merely continues to it. *Compelled* signalling is more reliable than *non-compelled*. However, the advantage that non-compelled signalling systems have is the ability to send a string of signals all in one go, so potentially reducing the time needed for call set up. This can be particularly valuable if the circuit propagation time is quite long. For if each signal has

to be acknowledged over a satellite circuit, then the *loop-delay* (there-and-back time) for signal transmission and acknowledgement over the circuit means that a maximum rate of around one signal (or digit) per second is all that can be achieved.

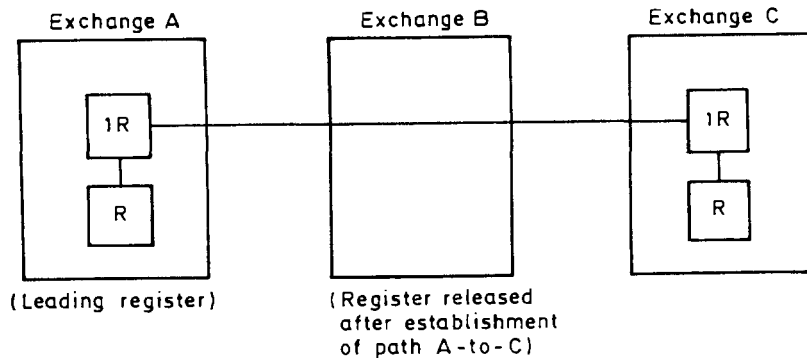
*Acknowledged* signalling is slightly different from *compelled signalling*: it is slightly less burdensome. Although compelled signalling systems are always also acknowledged systems, the reverse is not necessarily so. In an acknowledged (but not compelled) signalling system, short sequences of signals may be sent, and may be repeated if not acknowledged, but are not sent continuously. They may be acknowledged by a single signal. CCITT4 is a good example of *acknowledged* signalling; each forward digit is pulsed and acknowledged by a pulsed acknowledgement.

## 7.10 R2 INTER-REGISTER, MULTI-FREQUENCY CODE (MFC) SIGNALLING

In the previous section we saw how the line signalling part of R2 controls the circuit itself, conveying circuit seizure, answer, clearing, and other *circuit supervision* signals. It cannot convey the crucial information for actual call set up, including the dialled number, etc. This is function of the *inter-register signalling* part, which is activated following the *seizure signal* of the line signalling as we saw above.

The inter-register signalling part of R2 (R2-MFC, or *multi-frequency code*) works between R2 registers in an *end-to-end* fashion. At the outgoing exchange, an access loop signalling system (such as LD or MF4) will have stored the information required for call set up, in a register. Analysis of the digits, as we saw earlier in the chapter, then allows selection of an outgoing route to the destination. If this route is via another exchange, then inter-register signalling will be needed to relay the information to the register located in the subsequent (*incoming* or *transit*) exchange. Figure 7.10 showed how inter-register signalling equipment is configured to convey call information from the register in exchange A to that in exchange B, during call set up. If a further link, via exchange C, were added to the connection (as now shown in Figure 7.14) then the information required by the register in exchange C could be derived direct from the register in the outgoing exchange A. That being so, the register in exchange B can be released after the link B–C has been seized and the switch path through exchange B has been established. This method of signalling is described as *end-to-end* signalling. The originating register of an *end-to-end* signalling configuration (in our example, that in exchange A) is often termed the *leading register*.

End-to-end operation of inter-register signalling is common within national or international networks, but at the boundary between such networks (at the international gateway exchange for example) it is normal to undertake *signal regeneration*. By this we mean that a new *leading register* function is assumed by the gateway exchange. Thus a trunk exchange in one country is not expected to work on an end-to-end basis as an *outgoing exchange* with an *incoming* trunk exchange in some other country. Instead of that, *signal regeneration* is carried out at the outgoing international gateways and maybe the incoming one as well. In fact R2 is designed to work end-to-end from an outgoing R2 international register through to the distant local exchange. Incoming calls to Denmark, Switzerland and the Netherlands using R2 work in this way. Regeneration



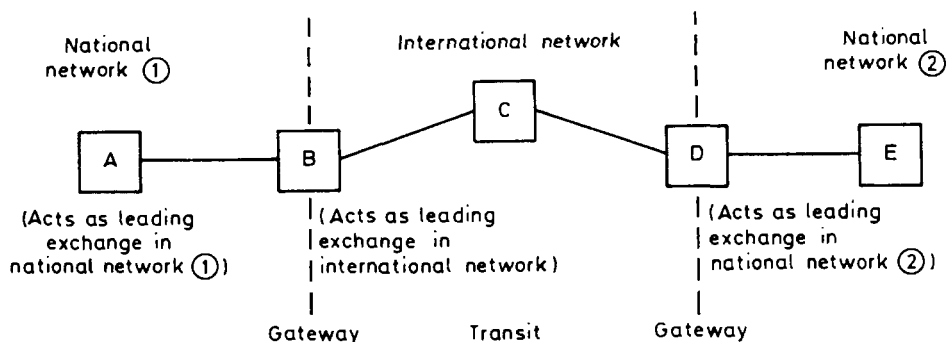
**Figure 7.14** End-to-end inter-register signalling. IR, inter-register signalling code senders and receivers; R, register

at the incoming international gateway is necessary when the national R2 is incompatible, or in cases where some other inland signalling system is used. The principle of regeneration is illustrated in Figure 7.15.

The network designer has at least two reasons for choosing to regenerate R2 signalling:

- either he will want to reduce the likelihood of excessively long holding times of *leading registers*, which can cause congestion
- or he may be anxious to accommodate the differences between national and international variants of R2.

As international as compared with national connections usually require a greater number of transmission links and exchanges to be connected in tandem, it follows that international connections generally take longer to set up. The *holding time* per call of *leading registers* which control the set up of international connections will therefore be longer than that of their solely national counterparts, and an accordingly greater



**Figure 7.15** Regeneration of R2 MFC

number of registers will be required. Therefore to prevent the congestion of national network registers which might result from a large throughput of international calls, it is normal to provide a separate set of international R2 registers at international gateways (such as exchange B of Figure 7.15). The number of international R2 registers provided at exchange B can account for the longer holding time. Furthermore, these registers can accommodate and *interwork* the slight differences between the different national and international versions of R2. Similar interworking needs call for further regeneration of national version R2 at exchange D.

Like the line signalling, the inter-register signalling in R2 MFC is carried out by means of *compelled* forward and backward signals. Subsequent signals waiting to be sent in the same direction (forward or backward) are only sent in the relevant direction as the result of requests and acknowledgements from the other end, and the previous signal removed at this time. The inter-register MFC signals themselves consist of 'two-out-of-six' *in-band* frequencies, sent together. The use of two frequencies makes the signalling more reliable and less subject to line interference, as before reacting to a signal a receiver must detect two, and only two, frequencies. In this way single stray frequencies do not cause mis-operation or the connection of 'wrong numbers'.

Fifteen *forward* and fifteen *backward* signals are available, constituted in the manner shown in Table 7.3. When R2 is used over four-wire or equivalent circuits, the use of separate frequency groups in the two directions gives no particular benefit, but over two-wire circuits it allows simultaneous signalling in both directions without interference.

Tables 7.4 and 7.5 illustrate the signals used in international R2 MFC. The first inter-register signal sent is a forward signal, usually the first dialled digit. This is acknowledged by signal A-1 requesting the next digit. Signals are sent in turn in the forward and backward directions, the incoming register controlling the sending of digits either by using the *send next digit* signal, A-1, or by requesting the re-sending of certain digits using any of signals A2, A7 or A8.

Once sufficient information has been received by the incoming register, digit analysis, route selection and the connection path switch-through may be carried out in the normal manner. International R2 MFC is more complicated than national R2 MFC,

**Table 7.3** Frequencies used in R2 MFC. (Courtesy of ITU – derived from Table 5/Q.441)

Signal frequencies		Signal value	Signal number														
Forward	Backward		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1380	1140	0	*	*		*			*				*				
1500	1020	1	*		*		*			*				*			
1620	900	2		*	*			*			*				*		
1740	780	4				*	*	*				*				*	
1860	660	7							*	*	*	*					*
1980	540	11											*	*	*	*	*



**Table 7.4** International R2 MFC forward signals. (*Courtesy of ITU – derived from Tables 6 and 7/Q.441*)

Group I: Address information etc.			
Signal number	Ia: First signal on an international circuit	Ib: Subsequent signals on an international circuit	Group II: Category of the calling party
1	Language digit: French	Digit: 1	Signals assigned for national use. If received by an international gateway they are converted to one of signals II-7 to II-10
2	English	2	
3	German	3	
4	Russian	4	
5	Spanish	5	
6	Spare (language digit)	6	
7	Spare (language digit)	7	
8	Spare (language digit)	8	Subscriber (or operator without forward transfer facility)
9	Spare (discriminating digit)	9	Data transmission call
10	Discriminating digit	0	Subscriber with priority
			Operator with forward transfer facility
11	Country code indicator: outgoing half-echo	Code 11: call to incoming operator	Spare signals for national use
12	Country code indicator: no echo suppressor required	(a) Code 12: call to operator for delay service (b) Request not accepted	
13	Code 13: call by automatic test equipment	Code 13: call to automatic test equipment	
14	Country code indicator: incoming half-echo suppressor required	Incoming half-echo suppressor required	
15	Spare	Code 15: end of pulsing	

in that more signals are generally required to carry the information pertaining to international networks, including signals which:

- indicate the country of destination
- control the selection of the type of transmission system (e.g. satellite or cable)
- request overseas operator assistance (if required), and the preferred language to be used for such assistance
- control specific items of equipment, such as echo suppressors.

Later chapters, on routing, numbering and transmission plans, as well as on operator services, describe the reason for and use of these signals.

The signals used in international signalling systems such as R2 are standardized and detailed in ITU-T recommendations. However, when used within a single operator's network some scope exists for using either a more sophisticated range of signals or a simpler one.

**Table 7.5** International R2 MFC backward signals. (*Courtesy of ITU – derived from Tables 8 and 9/Q.441*)

Signal number	Group A: Control signals	Group B: Condition of the called subscriber's line
1	Send next digit ( $n + 1$ )	Spare signal for national use
2	Send last but one digit ( $n - 1$ )	Subscriber transferred
3	Send category of calling party and changeover to reception of B-signals	Subscriber's line busy
4	Congestion in the national network	Congestion
5	Send category of calling party	Number not in use
6	Set up speech conditions	Subscriber's line free, charge
7	Send last but two digits ( $n - 2$ )	Subscriber's line free, no charge
8	Send last but three digits ( $n - 3$ )	Subscriber's line out of order
9		
	Spare signals for national use	Spare signals for national use
10		
11	Send country code indicator	
12	Send language or discriminating digit	
13	Send identity code (location) of outgoing international R2 register	Spare signals for international use
14	Is an incoming half-echo suppressor required?	
15	Congestion in an international exchange or at its output	

A more sophisticated range of signals might be invaluable in enabling more advanced call routing mechanisms to be adopted, as will be described in Chapter 32. Another reason for adopting a sophisticated signalling repertoire might be the network operator's desire to introduce some of the *network management* techniques to be described in Chapter 27. Not all these techniques, however, can be supported by the limited facilities of R2 signalling, and some more advanced signalling system such as SS7 may be needed to give network operators what they want. On the other hand a national network operator can usually reduce his overall costs by simplifying the signalling systems; in particular, R2 MFC can successfully be operated in national networks without either the forward or backward signals numbered 11 to 15. This means that some of the tone sending and receiving equipment is not needed. Specifically, both the forward tone of 1980 Hz and the backward tone of 540 Hz can be dispensed with.

Signalling System R2 is described more fully in ITU-T Recommendations Q.400-Q.490.

## 7.11 DIGITAL LINE SYSTEMS AND CHANNEL-ASSOCIATED SIGNALLING

In the earlier part of this chapter we learned how, in channel-associated signalling systems (such as CCITT1, CCITT2, CCITT3, CCITT4, CCITT5, R1 and R2), the call

set up and control information is carried over the channel itself. This is the case where the circuits are carried over analogue transmission equipment (i.e. using either audio *pairs* or FDM lineplant), but in cases where the circuits are carried by digital line systems it is usual to encode even *channel-associated* line signalling into a single channel. In 2 Mbit/s (European) digital transmission practice, the channel normally used for this purpose is timeslot 16 (as we discovered in Chapter 5), and all the line signals for the 30 usable circuits (channels 1–15 and 17–31) are encoded on this single channel. Similarly, in 1.544 Mbit/s digital transmission practice, either an entire 64 kbit/s channel can be used for the signalling, or more commonly the robbed-bit-signalling channel 9 (also described in Chapter 5) can be used.

Even though the line signalling may be concentrated on a single digital channel, it is correct to refer to CCITT1, CCITT2, CCITT3, CCITT4, CCITT5, R1 and R2 signalling systems as *channel-associated*. In all of them the multi-frequency inter-register signalling is still conveyed within the channel itself. Concentrating the line signalling on a single channel, however, enables a small economy to be made in the number of line signalling senders and receivers.

A special type of R2 line signalling was developed for use over digital line systems. The system is called R2-D (standing for R2 Digital). The line signalling of R2D is defined by ITU-T recommendation Q.421. The MFC part of R2D is identical to the analogue version. Similarly, special digital variants of the line signalling parts are to be found in other signalling systems.

## 7.12 SIGNALLING INTERWORKING

The choice of signalling system for a network is determined by a number of factors, including the network topology, the length and propagation delay of the transmission links, the services to be provided, the transmission media to be used, and of course the nature of the established network. Sunk investments in obsolescent signalling systems often mean that ancient and modern equipment practices and signalling systems have to co-exist. The job of designing and operating cheap and effective networks against such an evolving background is challenging and will remain so.

Figure 7.16 illustrates a typical situation where R2 signalling is at work on the link between exchanges A and B, with CCITT5 signalling on link B–C of the same connection. The configuration demands that similar call set up and clear-down information can be carried by both signalling systems, and even transferred from one to the other through some form of signalling interworking. Thus exchange B must comprehend both the R2 and SS5 (CCITT5) signalling repertoires and be capable of interworking between the two.

The provision of signalling interworking demands careful appreciation by exchange designers of the logical sequences and information flows of the two signalling systems. The exchange itself, or the signalling conversion device, needs to be able to interpret correctly the signals of both of the systems and convey the same meaning when converting between the two. The complexities of interworking reach their peak when the two signalling systems that are to be interworked have different signal vocabularies (e.g. one system may only have a ‘network busy’ signal whereas the other may have a

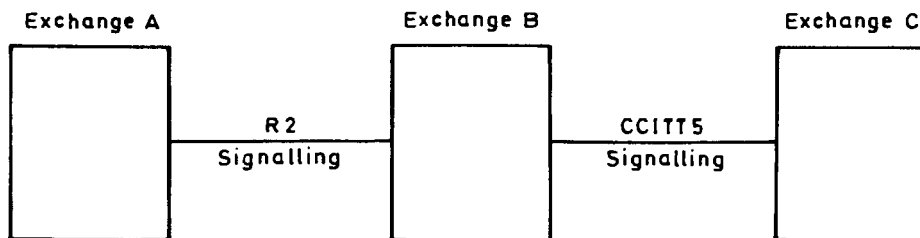


Figure 7.16 Signalling interworking

range of signals to determine whether the exchange or the called customer is busy). In these circumstances some information may well be lost.

The loss of information may, or may not, cause the network to operate in an unpredictable manner. For example, in CCITT5 there is only one backward signal to indicate that the call cannot be completed for any reason, including far-end customer congestion, network congestion, invalid number, etc. The signal is called the *busy flash* signal. However, in R2 signalling a number of different backward signals enable the cause of the unsuccessful call completion also to be returned: e.g. A4 national network congestion; A15 international network congestion; B3 subscriber line busy; and B4 congestion. So in Figure 7.16, if exchange B receives a *busy flash* signal in SS5 signalling from exchange C, which unsuccessful message signal should be conveyed in R2 signalling back to exchange A? There is no perfect answer: there is, however, a 'standard' answer, as defined by the set of ITU-T signalling interworking recommendations in the Q.600 series, to the effect that either the A4 (national network congestion) signal should be returned or the speech channel should be connected to *busy tone*.

## 7.13 ADVANCED SIGNALLING APPLICATIONS

Having described the basic principles of signalling for setting up connections on circuit switched networks, we can now approach some of the sophisticated network control methods made possible by more advanced signalling systems. These include controlling network traffic, call routing and overall transmission quality. Later chapters cover the advanced signalling requirements of what is called network management, and the most advanced of today's signalling systems SS7 (signalling system number 7).

## 7.14 SIGNALLING SEQUENCE DIAGRAMS

It is helpful, in the design, operation and maintenance of signalling systems, to understand the correct sequences of signals. This simplifies the task of understanding and tracking exchange responses to signals and gives an insight into the interaction of different signalling systems, so making easier the diagnosis of faults. The simplest means of illustrating signalling sequences, and the one commonly employed in specification documents is the signalling sequence diagram, an example of which is Figure 7.17.



Figure 7.17 illustrates a hypothetical connection between two telephones, A and B, interconnected by means of exchanges C, D and E. Signalling from customer A to exchange C is by means of *off-hook* and DTMF signalling. From exchange C to exchange D R2 signalling is in use, and from D to E CCITT5 signalling.

In reality, it is highly unlikely that customers would be directly connected to international exchanges in the manner illustrated in Figure 7.17, nonetheless it is theoretically possible. I have chosen to illustrate international signalling sequences because they are slightly more complex. International signalling systems generally need to convey more information, and here we show the conveyance of information about *country code* of destination, *language digit* (discussed in more detail in Chapter 14), and echo suppressor needs (Chapter 33). National interexchange signalling systems do not always need these complications and are generally simpler, so if you can understand the signalling sequence described here, you should be well set up for assimilating other national signalling sequence diagrams.

Ringling current and *loop* signalling provide the final means of alerting customer B and him signalling his answer.

The sequence diagram shows the signalling interactions which must occur to establish the call and clear it down afterwards. When drawn out like this, we see how many interactions the exchanges need to cope with!

Sidelined on the diagram are annotated the various stages of call set-up as described earlier in the chapter, thus *off-hook*, *dial tone*, *dialled digits*, etc.

The reason for illustrating the use of different signalling systems on links CD and DE is to contrast their different methods of operation, and the *interworking* between them that must be conducted by exchange D. You will note two major differences. First, some of the signals do not have identical meanings in the two signalling systems, so a 'best fit' approach occasionally needs to be used. Second, the sequencing of the signals is quite different in the two systems. R2 is operating in an *overlap* signalling mode. By this we mean that exchange C is retransmitting the early dialled digits on to exchange D even before the final digit has been received from the caller, A. Thus signalling on links AC and CD is in *overlap mode*. In contrast, exchange D does not commence signalling on link DE until all the digits have been received from C. This is because CCITT5 signalling is required to operate in an *en bloc signalling mode*, in which all the digits are sent together.

Historically, *overlap mode* signalling tended to be employed with early electromechanical exchanges, as it allowed exchanges late in the connection to commence their relatively slow switching actions as early in time as possible, so minimizing the overall time required for call set-up. The much more rapid digit analysis and switching action of SPC and digital exchanges favours the use of *en bloc signalling*, because the exchange processor's time is not wasted simply waiting to receive digits from a previous exchange.

The end user will not be concerned about whether the connection is established using either *overlap* or *en bloc mode* signalling, but there is one word of warning for the network operator. As the two modes do not interact efficiently (witness Figure 7.17), a connection made up of links alternating in signalling types employing both *overlap* and *en bloc* modes will tend to be relatively slow in establishing calls; 5 to 10 seconds can be added to the call set-up time. If callers are used to being connected in around 15–20 seconds and it takes 25–30 there is a risk that callers will clear calls prematurely, thinking that the set-up phase has failed.

## 7.15 CALL SET-UP AND INFORMATION TRANSFER IN DATA NETWORKS

So far, we have only considered setting up connections in telephone networks. Here, the *caller* (or *calling party*) has a need to *control* the network for the purpose of requesting a connection set-up (and for *clearing* afterwards). Once the connection is set up, there are no constraints on how he uses it, except of course that he will have to speak a language which the *called party* understands, if he wishes to communicate successfully.

In data communications, one speaks of *protocols* as opposed to *signalling*, though the principles are similar. To set up a *connection* in a *connection-oriented* data network (e.g. an X.25 packet network (Chapter 18) or an ATM network (Chapter 26)), a *control plane protocol* is used. A *control plane protocol* has the same function as the *signalling system* in a telephone network: it *establishes* and *clears* connections. Once the *connection* is established, the two end communicating devices convey information between themselves by means of the *user plane protocol*. This is the equivalent of human language in the telephone call case. The *user plane protocol* sets out how to code information and interpret it on receipt. Finally, *management plane protocols* allow for the administration, management, operation and monitoring of individual switches, ports and other devices within the network. Figure 7.18 illustrates the various types of communication which may take place.

All three types of communication (*control plane*, *user plane*, *management plane*) may occur simultaneously on the same physical connection, as we shall see in Parts 3 and 4.

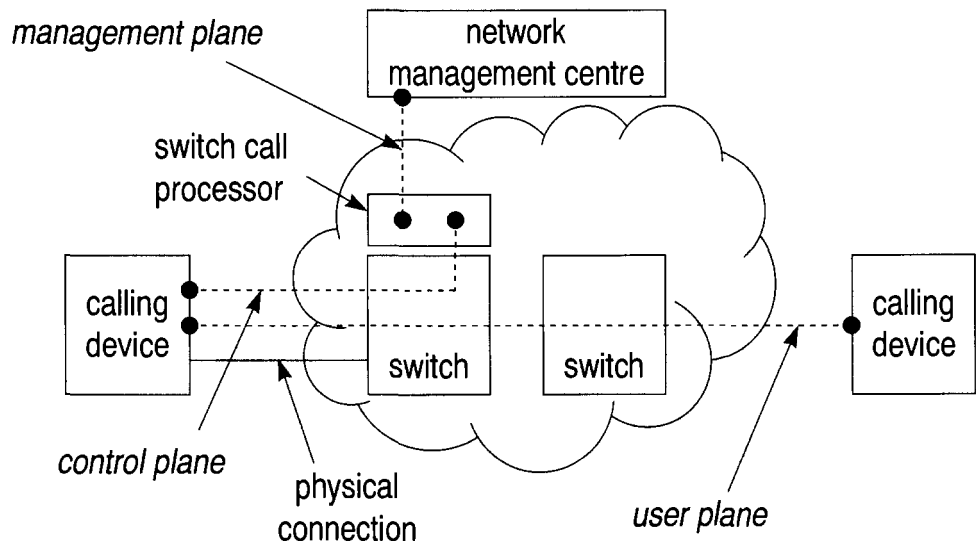


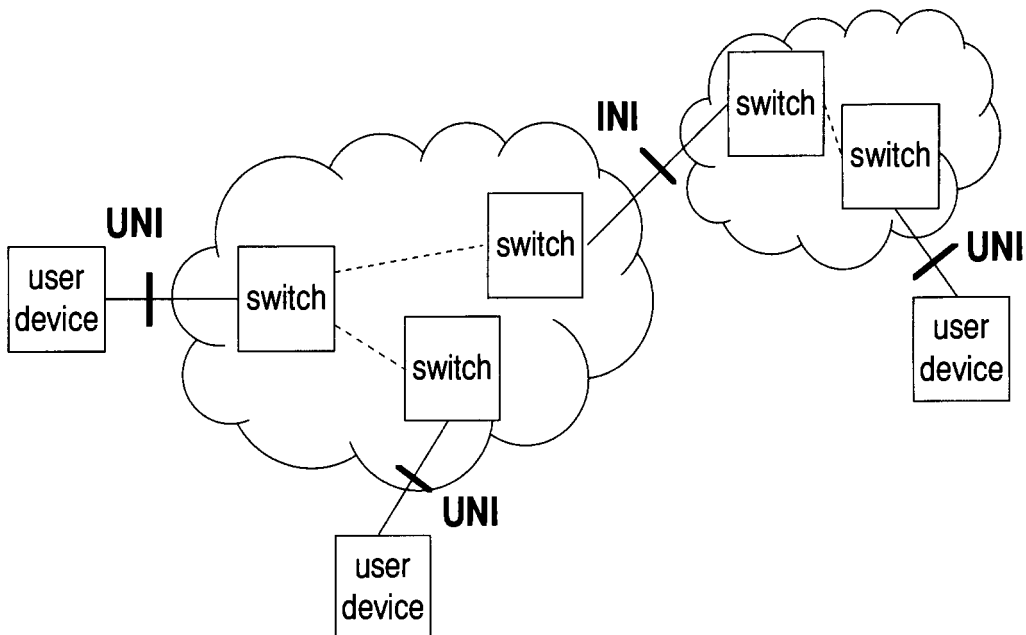
Figure 7.18 User, control and management plane communications

## 7.16 NETWORK INTERFACES: UNI, NNI, INI, ICI, SNI

In order to complete our review of the fundamentals of signalling systems and network protocols, it is valuable to introduce the various standard terms applied to describe different types of network interfaces, in particular. We shall discuss

- UNI (user–network interface)
- NNI (network–node or network–network interface)
- INI (inter-network interface)
- ICI (inter-carrier interface)
- SNI (subscriber–network interface)

The terms originate from the data networking world, where historically networks were built from switches and other equipment provided entirely by a single manufacturer. In such a world (Figure 7.19) it was not important whether a standard published (i.e. open) interface was used between switches within the network. Indeed, the manufacturers claimed added operational or management benefits from the use of their own ‘proprietary’ techniques and protocols. However, the interface used to connect users’ end devices to the network (the *user–network interface* or *UNI*) and the interface



**Figure 7.19** UNI (user–network interface) and INI (inter-network interface)





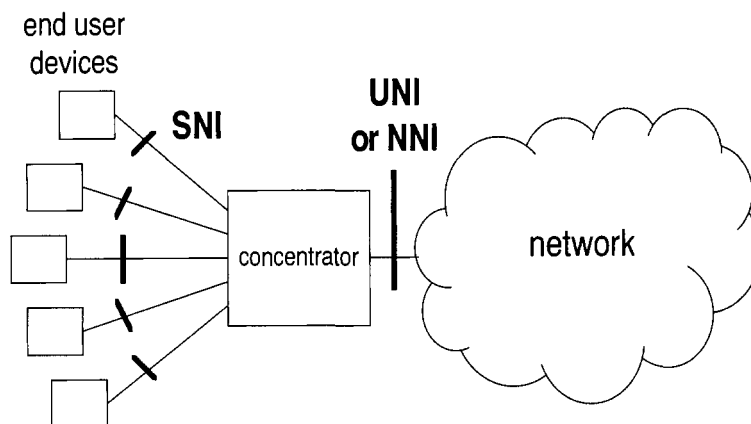


Figure 7.21 SNI (subscriber-network interface)

The *network-node interface (NNI)* allows the network of a single network operating administration to be built either of individual switches from many different manufacturers (in which case *NNI* stands for *network-node interface*) or of several sub-networks provided by different manufacturers (in which case *NNI* stands for *network-network interface*). The *network-node interface* and *network-network interface* are identical. Figure 7.20 illustrates the two cases of NNI.

At the INI call set up and clear-down must be supported but *alerting* (e.g. ringing current) is not necessary. The interface may embrace a bundle of parallel physical circuits which may be used as alternative paths for the carriage of a single connection.

The *inter-carrier interface (ICI)* is a special case of the *inter-network interface (INI)* for which the functionality has been adapted to accommodate ownership of the two networks by distinct, maybe competing, network operating organizations. Thus extra facilities have been added for quality of service monitoring and accounting as well as for the prevention of network control (i.e. management plane access) across the ICI.

Finally, *SNI (subscriber-network interface)* is the term applied to the physical connection of the end user device to the network, particularly in the case where some form of local concentrating device is used (as, for example, in Figure 7.21). A standard UNI (user-network interface) may or may not be defined for the connection of the concentrator of Figure 7.21 to the network. Alternatively, an NNI may be defined. SNIs are defined for *SMDS (switched multimegabit digital service)* and for transmission access networks (Chapter 17).

Incidentally, you may have found the terms somewhat duplicitous. This is the natural consequence of the definition of technical standards, on different timescales, by different bodies.

## 7.17 INFORMATION TRANSFER IN CONNECTIONLESS NETWORKS

*Connectionless* networks, as we discovered in Chapter 1, do not require the establishment of a *connection* prior to information transfer. Nonetheless, equivalent information must

be added to indicate the delivery *address* of each *message* or *frame*, and the content must be coded according to a suitable protocol (akin to *user plane protocol*) to enable correct interpretation on receipt.

The *Internet* protocol is an example of a *connectionless* protocol. The information is packed up in a single package and posted with a single address, coded according to the rules of the protocol. The message slowly percolates its way from switch to switch through the network. The set-up phase is far less onerous (it is not needed), and the recipient's device (e.g. personal computer) need not be switched on at the time of 'posting'. This is the major benefit of the *connectionless* transfer mode. On the other hand, unless the recipient chooses to respond, there is no acknowledgement or even guarantee of receipt.

---

# Transmission Systems

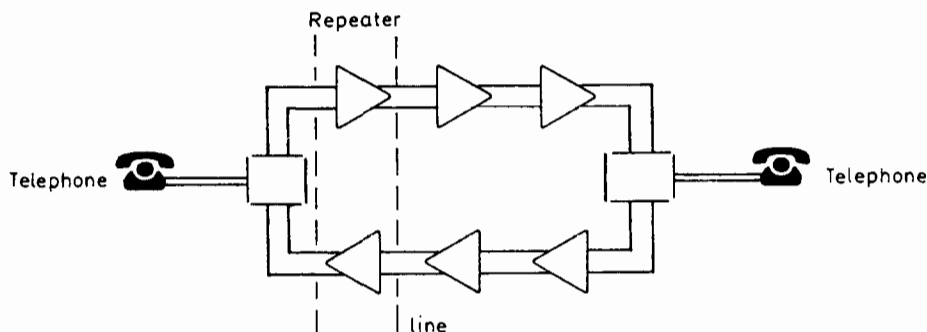
---

The basic line transmission theory that we have considered so far, together with the principles of analogue and digital signal transmission over pairs of electrical wires, is quite suitable for short range conveyance of a small number of circuits. However, it is not always practical or economic to use many multiple numbers of physical ‘pairs’ between exchanges, so we have recourse to frequency division and time division multiplexing to reduce the number of physical pairs of wires needed to convey a large number of long haul circuits between common end-points. There are, of course, different transmission media to be considered, some of them much better suited to particular applications than plain electric wire. They include coaxial cable, radio, satellites and optical fibre. In this chapter we discuss the main features of each of these transmission types, to ease the decision about which one is best suited to any particular application.

## 8.1 AUDIO CIRCUITS

The two-wire and four-wire transmission systems described in the first chapters of this book, comprising, respectively, one or two physical *pairs* of copper (or aluminium) wires together with amplifiers and equalizers as appropriate, are correctly called audio circuits. (In actual use the wires themselves are often wound or *twisted* around each other inside the cable, hence the terms *twisted pairs* and *twisted pair cables*). The title *audio* or *baseband* circuit is given because the frequencies of the electrical signals carried by the circuit are a direct match with those of the audio signal they represent. In other words, the electrical signal frequencies are in the *audio* range; the signal has not been processed in any way, neither frequency-shifted nor multiplexed. As a reminder of a typical audio circuit configuration, Figure 8.1 shows a four-wire repeated circuit as already discussed in Chapter 3 on long haul communication.

Audio circuits may be used to carry any of a wide range of bandwidths, although in general the greater the bandwidth to be carried, the shorter the maximum range of the circuit, because the electrical properties of the line (its *impedance*, *reactance* and *capacitance*) tend to interfere. It is known that high frequency signals are carried only



**Figure 8.1** A four-wire repeatered 'audio circuit'

by the outer surface of wire conductors (the *skin effect*), and this makes high frequency signals prone to interference from electrical and magnetic fields emanating from adjacent wires.

Unamplified audio circuits are commonly used to provide relatively short range telephone *junction* circuits of 4 kHz bandwidth between exchanges over a range up to about 20 km, and amplification can be used to extend this range. A typical use of an audio circuit is to connect an analogue *local* (or *end office* or *central office*) exchange to the nearest *trunk* (or *toll*) exchange. Beyond a distance of about 15 km, even if amplifiers are used, economics tend to favour other transmission means. As an example, an unamplified pair of wires using 0.63 mm diameter conductors has an insertion loss (i.e. a line loss which must be added to other circuit losses) of 3 dB at a range of 7 km. The transmission range for speech is increased to 33 km by the use of an amplifier, but this range is usually precluded by the line quality needs of the signalling system.

As also explained in Chapter 3, the rate of signal degradation (i.e. *attenuation* and *distortion*) in *twisted pair* transmission media can be reduced by *loading* the cable (see Chapter 3) and more simply, by increasing the diameter (*gauge*) of the wires themselves. Bigger conductors cause less signal attenuation because they have less resistance. However, the relative cheapness of alternative transmission media makes it uneconomic to employ very heavy gauge (larger diameter) twisted pairs over more than about 20 km. At this distance the conductors of an unamplified circuit need to be around 1 mm thick, and because the amount of conducting material required (copper or aluminium) is proportional to the square of the diameter the cost escalates rapidly as the cable gauge is increased. On a more cheerful view, the converse also applies, so that at shorter distances narrower cables are economically favourable. In one way and another a network operator has extensive information available to help him choose the right conductor (copper, aluminium, etc.), and the appropriate gauge of wire, thereby cutting costs and still giving the transmission performance required.

Because they are susceptible to external electro-magnetic interference, audio circuits are not normally suitable for use on very wide-bandwidth systems, including FDM (frequency division multiplex) systems, although early systems used screened copper wires. Audio circuits are most commonly found in small scale networks, and within the *local area* of a telephone exchange. A number of twisted pairs are normally packed together into large cables. A *cable* may vary in size from about 3 mm diameter to about

80 mm, carrying anything between one and several thousand individual pairs. Thus, running out from a 10 000 line local exchange, a small number (say 10–20) of main backbone cables of several hundred or a thousand pairs may run out along the street conduits to streetside cabinets where a simply cross-connect frame is used to extend the pairs out in a star fashion on smaller cables (say 25 or 50 pairs). Finally, at the *distribution point* individual pairs of wires may be crimped within a cable *joint* (at the top of a telegraph pole) and run out from here to individual customer's premises.

## 8.2 STANDARD TWISTED PAIR CABLE TYPES FOR INDOOR USE

For indoor use in conjunction with *structured cabling systems* (Chapter 45), there are a number of standardized *unshielded twisted pair (UTP)* and *shielded twisted pair (STP)* cabling types. Table 8.1 provides a brief overview of them.

## 8.3 TRANSVERSE SCREEN AND COAXIAL CABLE TRANSMISSION

A considerable problem with *twisted pairs* is signal distortion caused by the *skin effect*. The interference may be manifested as noise or 'hum' on the line, or in extreme cases a stronger voice or other signal may be induced from adjacent wire pairs, causing an effect known as *crosstalk*. The likelihood of such interference is increased when higher bandwidth signals (including FDM) are carried. Two options are available to reduce the interference. The first is to screen the cable with a light metal foil, usually wound into the cable sheath. This is called *screened* or *shielded cable*. In addition, transverse screen cable also includes screens between wires within the cable, isolating *transmit* from *receive* pairs. A second type of cable not prone to electromagnetic interference of this kind is coaxial cable. This is constructed so that most of the signal is carried by an electromagnetic field near the centre of the cable. Coaxial cable was the workhorse of

**Table 8.1** Standardized UTP and STP (horizontal) cabling types

UTP or STP cable type	Intended application
Category 1 (Cat. 1) and Category 2 (Cat. 2)	voice and low speed data
Category 3 (Cat. 3)	specification allows use up to 16 MHz, typically allowing use for combined voice and ethernet or 4 Mbit/s token ring LANs
Category 4 (Cat. 4)	specification allows use up to 20 MHz, allowing for 16 Mbit/s token ring LAN use as well as voice
Category 5 (Cat. 5)	specification allows use up to 100 MHz, allowing use for voice or data up to 100 Mbit/s

frequency division multiplex (FDM), high bandwidth analogue and early digital transmission systems. Now optical fibre is supplanting it for new digital lines.

Instead of a *twisted pair* of wires, a coaxial cable consists of two concentric conductors, made usually of aluminium or copper. The central wire, or pole, is separated from the cylindrical outer conductor by a cylindrical spacing layer of insulation, as shown in Figure 8.2.

The outer conductor is a metal foil or mesh, wound spirally around the insulation layer, and outside this is a layer of sheathing to provide external insulation and physical protection for the cable.

A single coaxial cable is equivalent to a single pair of twisted wires. To achieve the equivalent of four-wire transmission, two coaxial cables must be used, one to carry the *transmit* signal and one to carry the *receive* signal.

Coaxial cables function in much the same way as twisted pair circuits, but there are some important differences in performance. Unlike *twisted pair* audio circuits, coaxial cable circuits are said to be *unbalanced*. By this we mean that the two conductors in a coaxial cable do not act equally in conveying the signals. In a *balanced* circuit such as a twisted pair, where the signal is carried by the electrical currents in the two wires passing in opposite directions, the currents generate equal but opposite electromagnetic fields around them which tend to cancel each other out. By contrast, the signal in a coaxial cable is carried largely by the electromagnetic field surrounding the inner and outer conductors, which is largely induced by the central conductor. The outer conductor, sometimes called the *shield*, is usually operated at *earth* or *ground* voltage, and prevents the electromagnetic field from radiating outside the cable sheath. The signal is thereby protected from interference to some extent.

Unfortunately the overall rate of signal attenuation is greater on unbalanced circuits like coaxial cables than on balanced ones such as twisted pairs, and this needs to be counteracted by the use of amplifiers. On the other hand, unlike twisted pair circuits which are liable to lose balance as the result of damp or damage, coaxial cables are more stable and suffer less from the variable attenuation and extraneous signal reflections which sometimes affect twisted pairs. Coaxial cables are more reliable in service, are easier to install, and they require less maintenance. Further, because of their insensitivity to electromagnetic interference they perform better than twisted pairs when the cable route passes near metal objects or other electrical cables.

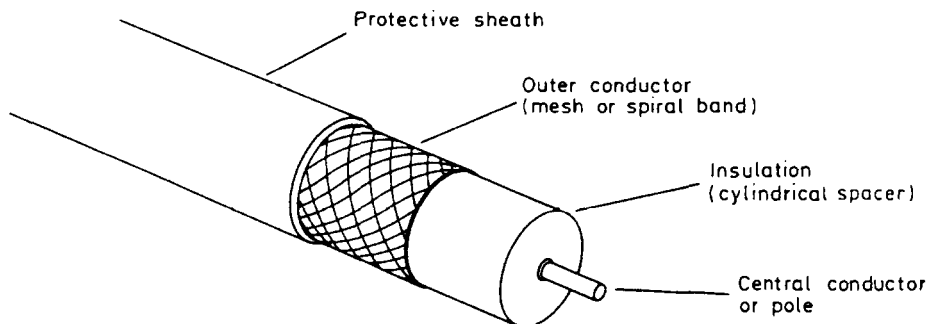


Figure 8.2 Coaxial cable

## 8.4 FREQUENCY DIVISION MULTIPLEXING (FDM)

FDM, or *frequency division multiplexing* is a means of concentrating multiple analogue circuits over a common physical four-wire circuit. It was used predominantly in the analogue era of telecommunication before glass fibres and digital techniques were available. Sophisticated equipment at the two ends of a long distance four-wire transmission line are used to break up the available signal bandwidth into many different and independent speech or other analogue signal circuits. In effect, the multiplexing equipment multiplies many times the number of circuits which may be carried, thus saving the expense of multiple long distance transmission lines.

FDM was used predominantly in conjunction with analogue trunk radio systems and long distance coaxial transmission lines. FDM equipment remains in widespread use, but the advent of cheaper and higher quality digital transmission alternatives (in particular, TDM or *time division multiplexing*) is leading to its obsolescence.

FDM allows many analogue transmission circuits (or *channels*) to share the same physical pair of wires or other transmission medium. It requires sophisticated and expensive transmission line terminating equipment and a high quality line, but it has the potential for overall cost saving because the number of wire pairs between the endpoints can be reduced. The longer the length of the line, the greater the overall saving.

Typical transmission systems used in conjunction with FDM are coaxial cables and analogue trunk radio systems. With analogue satellite systems, a slight variant of the technique is used, *FDMA* or *frequency division multiple access*. This uses a similar technique, but in addition it allows multiple earth stations to communicate using a single satellite.

The lowest constituent bandwidth that makes up an FDM system designed for the carriage of multiple telephone channels has a channel bandwidth of 4 kHz. This comprises the 3.1 kHz needed for the speech itself, together with some spare bandwidth on either side which buffers the channel from interference by its neighbouring channels. The speech normally resides in the bandwidth between 300 Hz and 3400 Hz.

In order to prepare the *baseband* signal which forms the basic building block of an FDM system, each of the individual input speech circuits are filtered to remove all the signals outside of the 300–3400 Hz band. The resulting signal is represented by the *baseband signal* (schematically a triangle) shown in Figure 8.3.

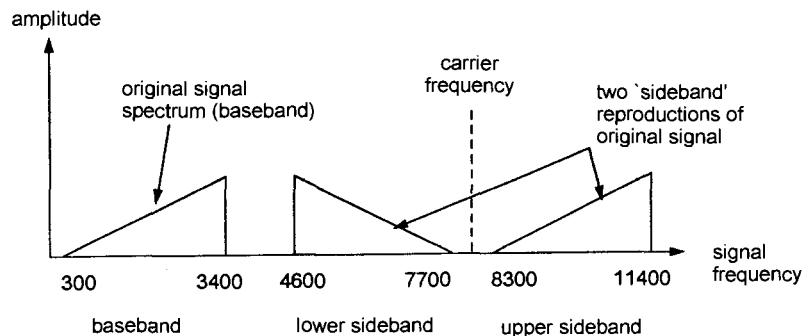
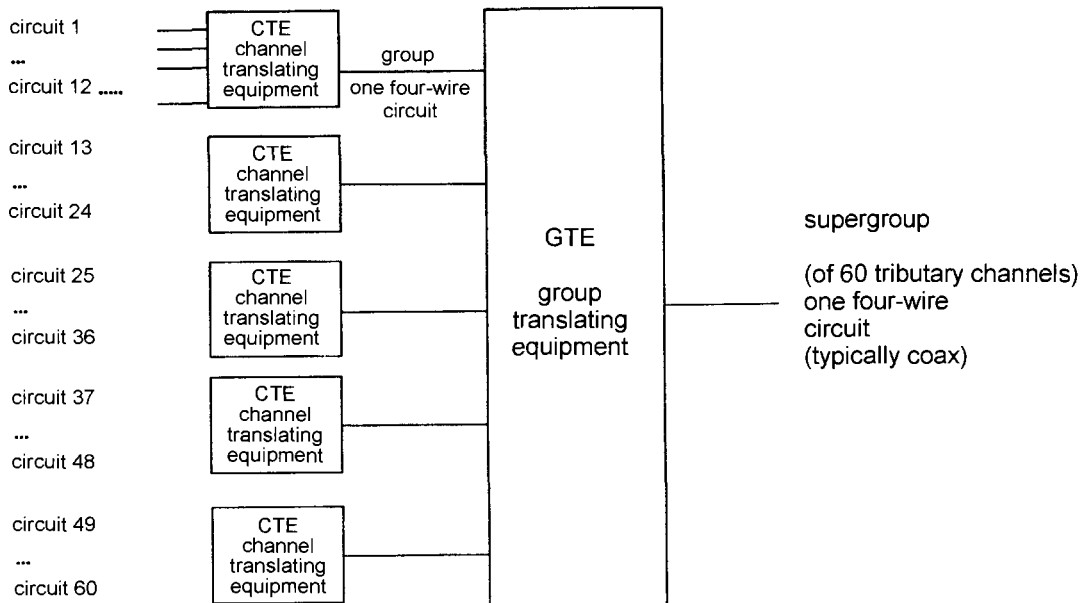


Figure 8.3 The principle of FDM, frequency shifting by carrier modulation





**Figure 8.4** Circuit multiplexing using FDM

The next step is that the signal is frequency shifted by *modulating* the baseband signal with one of a number of standard *carrier signals*. The frequency of the *carrier signal* used will be equal to the value of the frequency shift required. In our diagram, a carrier signal frequency of 8000 Hz is being used. This creates two new sideband images of the signal, the *lower sideband* and the *upper sideband*. All the original information from the baseband signal is held by each of the sideband images, so it is normal during the transmission to save electrical energy and bandwidth by transmitting only a *single sideband* having first suppressed the carrier (*suppressed carrier mode*; note, however, that suppressed carrier mode, while common for line systems is not common in radio systems as it is often inconvenient to make a carrier signal generator available at the receiver). *Demodulation* of the signal into its *tributary channels* is a similar process to that of modulation.

FDM is usually conducted in a hierarchical manner as shown in Figure 8.4. The hierarchy allows progressively more circuits to be multiplexed onto the same line. Of course, for each further stage of multiplexing a further line quality and bandwidth threshold must be exceeded for the system to operate reliably. The various different stages of the FDM hierarchy are achieved by using different *translating equipments* (the modulating and demodulating device), as we discuss next.

The carrier frequencies used to in a CTE to create a basic group are 64 kHz, 68 kHz, 72 kHz, ..., 108 kHz, and each of the lower sidebands is used. The resulting group frequency pattern is as shown in Figure 8.5.

This is the basic building block when further multiplexing into *supergroups*, but when sending to line it is normal to frequency shift the group into the lowest frequency range

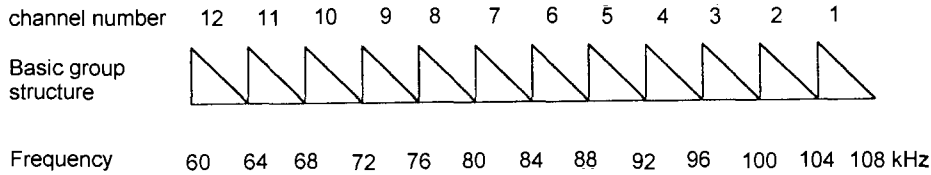


Figure 8.5 Basic group structure

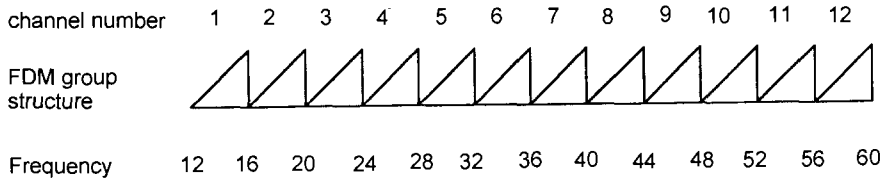


Figure 8.6 FDM group structure

by the use of a *group modulating frequency* of 120 kHz. This brings the signal into the frequency range 12–60 kHz (Figure 8.6).

The further multiplexing steps are listed briefly in Table 8.2.

FDM systems formed the backbone of public telephone networks until the mid-1970s when digital transmission and pulse code modulation took over as the cheaper and higher quality alternative. FDM was commonly used on the long distance and international links between *trunk (toll)* telephone exchanges. The transmission lines were typically either coaxial cable or analogue radio. Although the technique is now falling into disuse for optical fibre and copper line systems, it is likely to remain important in conjunction with radio systems, where the associated technique FDMA (frequency division multiple access) is a very effective way of sharing radio bandwidth between multiple communicating endpoints. We return to the subject of radio later in the chapter.

Table 8.2 FDM hierarchy

Translating equipment name	Tributary channel capacity	Signal output name
CTE channel translating equipment	12 voice channels	Group
GTE group translating equipment	5 groups (60 voice channels)	Supergroup
STE supergroup translating equipment		Hypergroup or Mastergroup

## 8.5 HDSL (HIGH BIT-RATE DIGITAL SUBSCRIBER LINE) AND ADSL (ASYMMETRIC DIGITAL SUBSCRIBER LINE)

Although modern optical fibre cable (discussed later in this chapter) has meant that copper pair and coaxial cabling has to some extent fallen out of fashion, the huge existing base of such cabling has caused recent development effort to be concentrated on seeking new methods of using it for high speed digital connection of customers. Two noteworthy methods have emerged. These are *HDSL* (*high bitrate digital subscriber line*) and *ADSL* (*asymmetric digital subscriber line*).

*HDSL* uses two or three copper pairs to provide a full duplex 2 Mbit/s access line, enabling high bitrate business applications to operate over existing copper lineplant. *ADSL*, meanwhile, provides a high bitrate of 4 Mbit/s or 6 Mbit/s in the *downstream* channel (e.g. for broadcasting a video signal to residential customers) with, say, a 64 kbit/s *upstream* channel for control of the received signal. Likely uses of *ADSL* are for *video-on-demand* (*VOD*), remote shopping (*teleshopping*), distance learning (*tele-education*) and other interactive *multimedia* services. We return to HDSL and ADSL in Chapter 17.

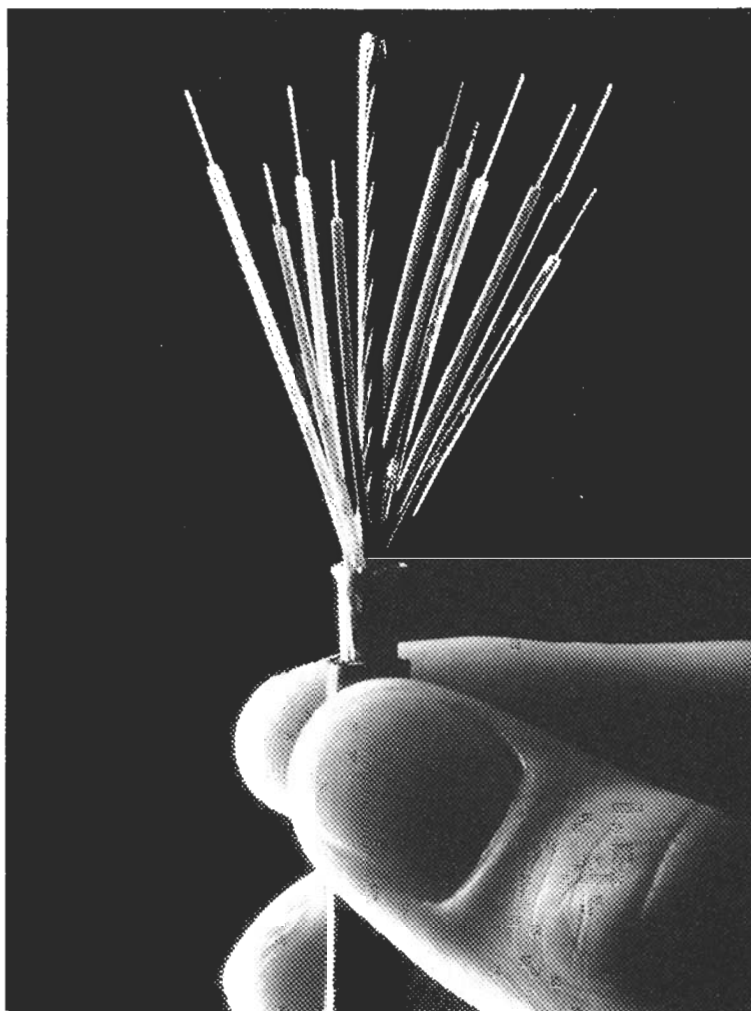
## 8.6 OPTICAL FIBRES

Most transmission types can be used to support either analogue or digital signal transmission: twisted pair circuits, coaxial cables and radio systems can all be used as the basis of either analogue or digital line transmission systems. However, the fact that digital transmission requires only two distinct line states (on/off; mark/space) has opened up a new wider range of digital propagation methods; and the most important new digital transmission medium is optical fibre.

The importance of optical fibres and their extensive use stems from their extremely high bit-rate capacity and their low cost. Optical fibres are a hair's breadth in diameter and are made from a cheap raw material: glass. They are easy to install because they are small, and because they allow the repeaters to be relatively far apart (well over 100 km spacing is possible) they are easy to maintain.

An optical fibre conveys the bits of a digital bit pattern as either an 'on' or an 'off' state of light. The light (of wavelength 1.3 or 1.5  $\mu\text{m}$ ) is generated at the transmitting end of the fibre, either by a laser, or by a cheaper device called a *light emitting diode* (*LED* for short). A diode is used for detection. The light stays within the fibre (in other words is guided by the fibre) due to the reflective and refractive properties of the outer skin of the fibre, which is fabricated in a 'tunnel fashion'.

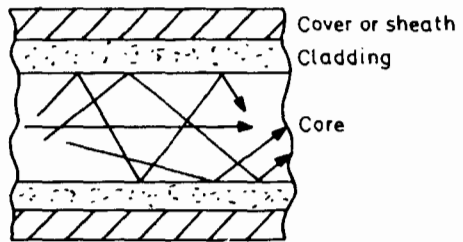
Fibre-optic technology is already into its third generation. In the first generation, fibres had two cylindrical layers of glass, called the core and the cladding. In addition a *cover* or *sheath* of a different material (say plastic) provided protection. The core and the cladding were both glass, but of different refractive index. A *step* in the refractive index existed at the boundary of core and cladding, causing reflection of the light rays at this interface, so guiding the rays along the fibre. Unfortunately however, because of the relatively large diameter of the *core*, a number of different light ray paths could be produced, all reflecting off the cladding at different angles, an effect known as



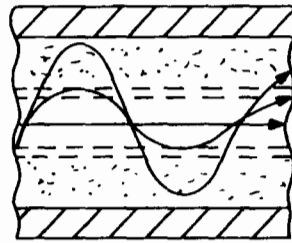
**Figure 8.7** An optical fibre cable, showing clearly the overall construction. At the centre, a steel twisted wire provides strength, enabling the cable to be pulled through ducts during installation. Around it are ten plastic tubes, each one containing and protecting a hair's breadth fibre, each capable of carrying many megabits of information per second. (Courtesy of British Telecom)

*dispersion*. The different overall ray paths would be of different lengths and therefore take different amounts of time to reach the receiver. The received signal is therefore not as sharp; it is dispersed. It is most problematic when the user intends very high bit rate operation. The different wave paths (or modes) explain the name *step index multimode* fibre which has been given to these fibres. *Step index multimode* fibre is illustrated in Figure 8.8(a).

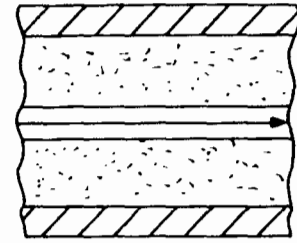
A refinement of step index multimode fibre is achieved by using a fibre with a more gradual grading of the refractive index from core to cladding. A *graded index multimode* fibre is shown in Figure 8.8(b). This has a slightly improved high bit rate performance



(i) Step index multimode



(ii) Graded index multimode



(iii) Monomode

Figure 8.8 Types of optical fibre

over step-index multimode fibres. Finally, *monomode fibres* (illustrated in Figure 8.8(c)) are the third generation of optical fibre development, and have the best performance. In *monomode fibre*, more advanced fibre production techniques have produced a very narrow core area in the fibre, surrounded by a cladding area; there is a step change in the refractive index of the glass at the boundary of the core and the cladding. The narrow core of a *monomode fibre* allows only one of the ray paths (or *modes*) to exist. As a result there is very little light pulse dispersion in *monomode fibre*, and much higher bit rates can be carried.

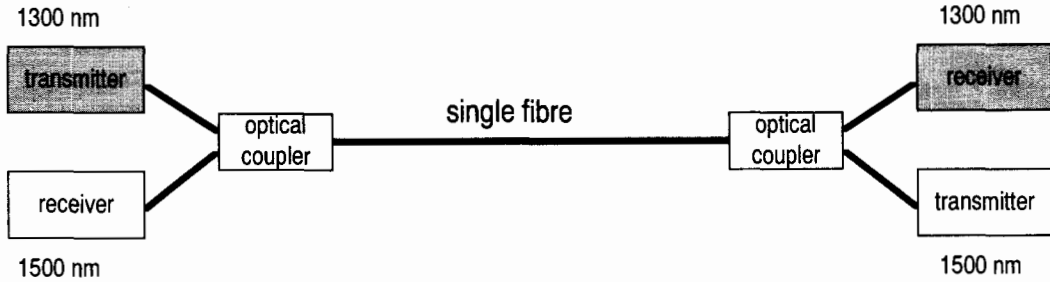
The qualities of different types of optical fibre are laid out in ITU-T recommendations G.652, G.653 and G.654. Recommendation G.652 specifies a monomode fibre optimized for operation at 1310 nm. Recommendations G.653 and G.654 specify monomode cables optimized for operation at 1550 nm. All the types of cable may be used for both 1310 nm and 1550 nm, but may not operate at maximum efficiency for both wavelengths.

For in-building and campus cabling of office, industrial or university sites it is now commonplace to use optical fibre at least in the building *risers* (the conduits between floors or buildings). The cable generally used contains *multimode* fibre, usually suitable for transmission between relatively cheap, LED transmitting devices over distances up to 3 km or 15 km. Although the *multimode fibre* itself is nowadays more expensive than monomode fibre, this is usually compensated for by the much cheaper end device costs.

In contrast, *monomode fibre* is the preferred technology of network operators, and this is the cable mainly deployed for metropolitan, national and undersea usage. The high volumes of production have brought the cost below that of multimode fibre, despite the more complex manufacture. The main benefit is the very high bitrates achievable and the lengthy inter-repeater distances. The disadvantage is the need for expensive laser devices as transmitters to produce the high quality *single mode* (single wavelength) light required.

An important factor in the design and installation of optical fibre networks is the careful jointing or interconnection of fibre sections. The reflections caused at the joint, and the losses caused by slight misalignment of the two fibre cores can lead to major losses of signal strength. The latest generation of cabling splicing devices and cable connectors have improved real network performance dramatically in recent years. The ST-connector (a bayonet-type connector) now common for optical fibre patch frames and equipment connections, for example, has been designed to ensure correct alignment of the fibres. The SC-connector specially developed for ATM (asynchronous transfer mode, see Chapter 26) also shares the ability for exact alignment.

Recent development in optical fibre technology has concentrated on achieving longer operational life of installed cable, without the need to replace 'active' components. One of the problems with the original systems was the use of optical signal regenerators specific to a given light wavelength and transmission bitrate. This means that the early optical fibres cannot easily be upgraded in bitrate without major re-investment in regenerators and other active components. For this reason, considerable effort has been put into optical amplifiers and regenerators which work without reconverting the signal to electrical form and then re-creating a new light signal for onward transmission. Such amplifiers can be built to give greater range of optical wavelength and bitrate. Bitrate, wavelength or multiplexing upgrades can thus to some extent be limited to the end devices, giving scope for new developments like *wavelength division multiplexing* (*WDM*) and *passive optical networks*, which we discuss next.



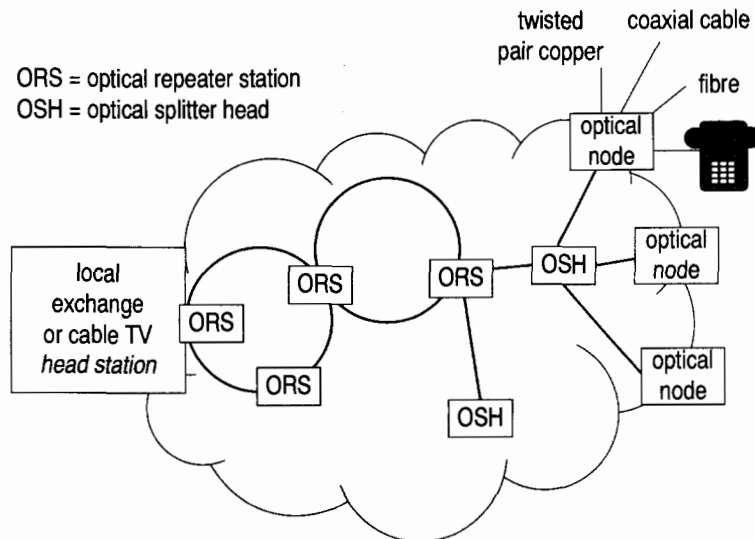
**Figure 8.9** The principle of wave division multiplex (WDM)

### 8.6.1 Wavelength Division Multiplexing (WDM)

*Wave division multiplexing (WDM)* is the use of a single fibre to carry multiple optical signals. Thus, for example, two signals at 1300 nm and 1500 nm light wavelength could carry independent 155 Mbit/s digital signals over the same fibre. In this way, *transmit* and *receive* signals could be consolidated to a single fibre, rather than requiring separate fibres (Figure 8.9). Alternatively, the capacity of two existing fibres can be doubled.

### 8.6.2 Broadband Passive Optical Network (BPON)

*Broadband passive optical networks (BPONs)* are intended for deployment in customer connection networks between network operators *local exchange* sites (*central offices*) and customer premises. Here the challenge is to achieve maximum efficiency in the use of local lineplant, creating the effect of a star-topology (direct connection of each



**Figure 8.10** An optical fibre access network

customer to the exchange) without needing thousands of fibres and also giving the potential of signal broadcasting (for cable television for example) to each customer over a common cabling infrastructure. Figure 8.10 illustrates a typical optical fibre access network. *Optical repeater stations* are coupling units for boosting signal strength or for connecting various fibre rings, meanwhile *splitters* are *passive* components which use an optical *cavity* of the appropriate length to divide the incoming signal between two fibres. The optical nodes of Figure 8.10 are the receiving devices.

## 8.7 RADIO

In 1887 Heinrich Hertz produced the first man-made radio waves. Radio waves, like electricity and light, are forms of electromagnetic radiation; the energy is conveyed by *waves* of magnetic and electrical fields. In a wire, these waves are induced and guided by an electrical current passing along an electrical conductor, but that is not the only way of propagating *electromagnetic (EM)* waves. By using a very strong electrical signal as a transmitting source, an electromagnetic wave can be made to spread far and wide through thin air. That is the principle of radio. The radio waves are produced by transmitters, which consist of a radio wave source connected to some form of antenna, examples of which are

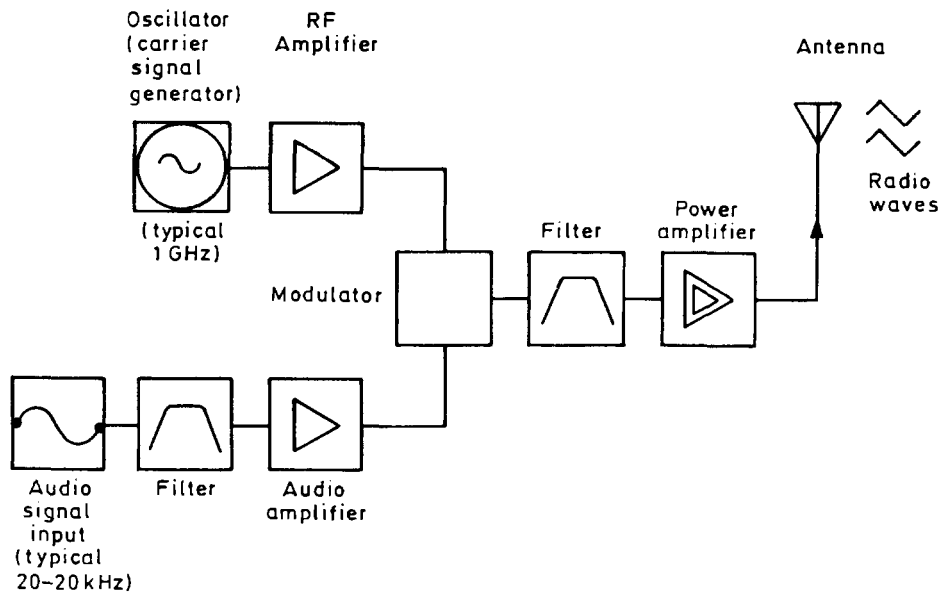
- low frequency antenna (aerial) or radio mast
- HF (high frequency), VHF (very HF), or UHF (ultra-HF) antenna or mast
- microwave dish antenna
- tropospheric scatter antenna
- satellite antenna

Radio is a particularly effective means of communication between remote locations and over difficult country, where cable laying and maintenance is not possible or is prohibitively expensive. It is also effective as a means of broadcasting the same information to multiple receivers and cost-effective for low volume use-sharing resources between many users.

One way of communicating information by radio waves is by encoding (or more correctly *modulating*) a high frequency *carrier* signal prior to transmission. The technique of modulation in radio is very similar to that used in FDM, and *single sideband (SSB)* operation, as we discussed earlier in the chapter is common, though the carrier is rarely suppressed. A distinctive feature of a radio carrier signal is its high frequency relative to the frequency bandwidth of the information signal. The frequency of the carrier needs to be high for it to propagate as radio waves.

The modulation of the radio carrier signal may follow either an analogue or a digital regime. Analogue modulation is carried out in a similar way to FDM. Digital modulation can be by ‘on/off’ carrier signal modulation (i.e. by switching the carrier on and off), or by other methods such as *frequency shift keying (FSK)*, *phase shift keying (PSK)* or *quadrature amplitude modulation (QAM)*, details of which are given in the next chapter.



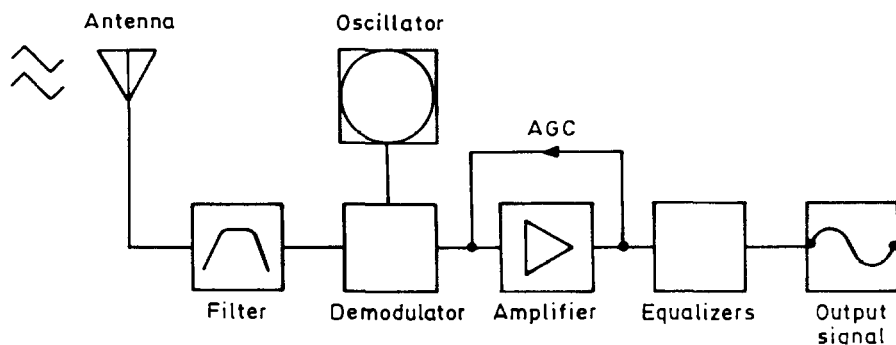


**Figure 8.11** A simplified radio transmitter

After modulation of the carrier frequency, the combined signal is amplified and applied to a radio antenna. Amplification boosts the signal strength sufficiently for the antenna to convert the electrical current energy into a radio wave.

Figure 8.11 illustrates a simple radio transmitter: an audio signal is being filtered (to cut out extraneous signals outside the desired bandwidth) and amplified. Next it is modulated onto a *radio-frequency (RF)* carrier signal which is produced by a high quality oscillator; then the modulated signal is filtered again to prevent possible interference with other radio waves of adjacent frequency. Finally the signal is amplified by a high power amplifier and sent to the antenna where it is converted into radio waves.

Figure 8.12 illustrates a corresponding radio receiver, the components of which are similar to those of the transmitter shown in Figure 8.11. The radiowaves are received by the antenna and are reconverted into electrical signals. A filter removes extraneous and



**Figure 8.12** A simplified radio receiver

interfering signals before demodulation. As an alternative to a filter, a *tuned circuit* could have been used with the antenna. A tuned circuit allows the antenna to select which radio wave frequencies will be transmitted or received.

Then comes demodulation. In the receiver shown in Figure 8.12 (typical of a microwave receiver) demodulation is carried out by *subtracting* a signal equivalent to the original carrier frequency, leaving only the original audio or *information* signal. A cruder and cheaper method of demodulation, not requiring an oscillator, employs a *rectifying circuit*. This serves to have the same subtracting effect on the carrier signal. After demodulation, the information signal is processed to recreate the original audio signal as closely as possible. The signal is next amplified using an amplifier with *automatic gain control (AGC)* to ensure that the output signal volume is constant, even if the received radio wave signal has been subject to intermittent *fading*. Finally, the output is adjusted to remove various signal distortions called *group delay* and *frequency distortion* by devices called equalizers. We will consider the cause and cure of these distortions in Chapter 33.

The simplified transmitter and receiver shown in Figures 8.11 and 8.12 could be used together for *simplex* or one-way radio transmission. Such components in use are thus equivalent to one *pair* of a four-wire transmission line discussed in Chapter 3. For *full duplex* or two-way transmission, equivalent to a four-wire system, the equipment must be duplicated, with two radio transmitters and two receivers, one of each at each end of the radio link. Two slightly different carrier frequencies are normally used for the two directions of transmission, to prevent the two channels interfering with one another. Alternatively, by duplicating senders and receivers at each end of the link, but using only one radio channel, *half-duplex* operation could be used (communication in both directions, but in only one direction at a time).

## 8.8 RADIO WAVE PROPAGATION

When radio waves are transmitted from a point, they spread and propagate as spherical wavefronts. The wavefronts travel in a direction perpendicular to the wavefront, as shown in Figure 8.13.

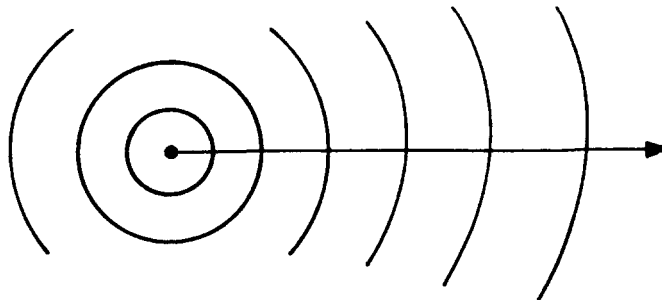
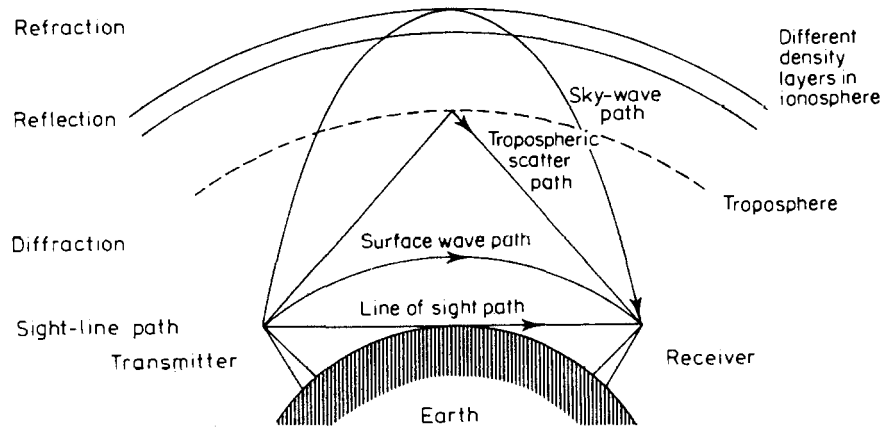


Figure 8.13 Radio wave propagation



**Figure 8.14** Different modes of radio wave propagation

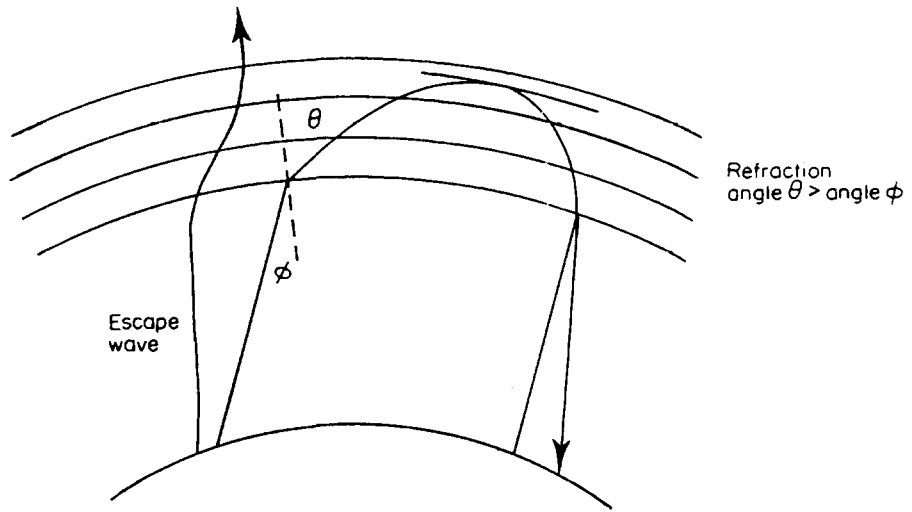
Radiowaves and lightwaves are both forms of electromagnetic radiation, and they display similar properties. Just as a beam of light may be reflected, refracted (i.e. slightly bent), and diffracted (slightly swayed around obstacles), so may a radio wave, and Figure 8.14 gives examples of various different wave paths between transmitter and receiver, resulting from these different wave phenomena.

Four particular modes of radio wave propagation are shown in Figure 8.14. A radio transmission system is normally designed to take advantage of one of these modes. The four modes are

- *line-of-sight* propagation
- *surface wave* (diffracted) propagation
- *tropospheric scatter* (reflected and refracted) propagation
- *skywave* (refracted) propagation

A *line-of-sight* (LOS) radio system (microwave radio above about 2 GHz) relies on the fact that waves normally travel in a straight line. The range of a line-of-site system is limited by the effect of the earth's curvature, as Figure 8.14 shows. Line-of-sight systems therefore can reach beyond the horizon only when they have tall masts. Radio systems can also be used beyond the horizon by utilizing one of the other three radio propagation effects also shown in Figure 8.14.

Surface waves have a good range, depending on their frequency. They propagate by diffraction using the ground as a *waveguide*. Low frequency radio signals are the best suited to surface wave propagation, because the amount of bending (the effect properly called *diffraction*) is related to the radio wavelength. The longer the wavelength, the greater the effect of diffraction. Therefore the lower the frequency, the greater the bending. A second means of over-the-horizon radio transmission is by *tropospheric scatter*. This is a form of radio wave reflection. It occurs in a layer of the earth's atmosphere called the *troposphere* and works best on ultra high frequency (UHF) radio waves. The final last example of over-the-horizon propagation given in Figure 8.14 is



**Figure 8.15** Refraction causing skywaves

known as skywave propagation. This comes about through refraction (deflection) of radiowaves by the earth's atmosphere, and it occurs because the different layers of the earth's upper atmosphere (the *ionosphere*) have different densities, with the result that radiowaves propagate more quickly in some of the layers than in others, giving rise to wave deflection between the layers, as Figure 8.15 shows in more detail. However, as Figure 8.15 also demonstrates, not all waves are refracted back to the ground; some escape the atmosphere entirely, particularly if their initial direction of propagation relative to the vertical is too low (angle  $\phi$ ).

## 8.9 RADIO ANTENNAS

Each individual radio system is required to perform slightly different tasks. The distinguishing qualities of a radio system are

- its range
- its transmit and receive signal power
- its directionality (i.e. whether the transmitted signal is concentrated in a particular direction or not)
- its mode of use (e.g. *point-to-point* or *point-to-multipoint*)

These qualities are determined by the mode of propagation for which the system is designed (e.g. surface wave, tropospheric scatter, line-of-sight, skywave, etc.), by the frequency of the radio carrier frequency, by the power of the system and most particularly by the suitability of the transmitting and receiving antennas. Antenna design has a significant effect on radio systems' range, directionality and signal power.

The simplest type of radio antenna is called a *dipole*. A dipole antenna consists of a straight metal conductor. Examples can be found fitted on many household domestic radios, to receive *VHF* (or *FM*) radio signals. Another, larger, example of a dipole aerial is a radio mast. Radio masts may be up to several hundred feet in height, consisting of one enormous dipole.

The qualities of dipole antennas make them suitable for broadcast or *omnidirectional* transmission and receipt of radio signals. A dipole antenna sends out a signal of equal strength in all directions, and a receiving antenna can receive signals equally well from most directions. Dipole aerials are therefore much used for *broadcast* transmission from one transmitting station to many receiving stations. Dipole antennas are also used for point-to-point transmission when the location of the receiver or transmitter is unknown, as in the case of mobile radio or ships at sea.

The disadvantage of omnidirectional receiving antennas is their susceptibility to radio interference from undesired sources. Omnidirectional transmitting antennas also have a disadvantage: they expend a lot of power by transmitting their radio wave in all conceivable directions, but only a small number of points within the overall coverage area pick up a very small fraction of the signal power; the remainder is wasted. For this reason, some antennas are especially designed to work directionally. One example of a directional antenna is the dish-shaped type used for microwave, satellite, and tropospheric scatter systems; see Figure 8.16(a). Dish-shaped antennas work by focusing the transmitted radio waves in a particular direction. Another type called an *array antenna* is illustrated in Figure 8.16(b). It is similar to the type used for domestic receivers.

By using a directional rather than an *omnidirectional* antenna, the range of the radio system and the signal power transmitted in or received from a given direction can be deliberately controlled.

Directional antennas are ideal for use on point-to-point radio links; the overall power needed for transmission is reduced, and the received signal suffers less from interference.

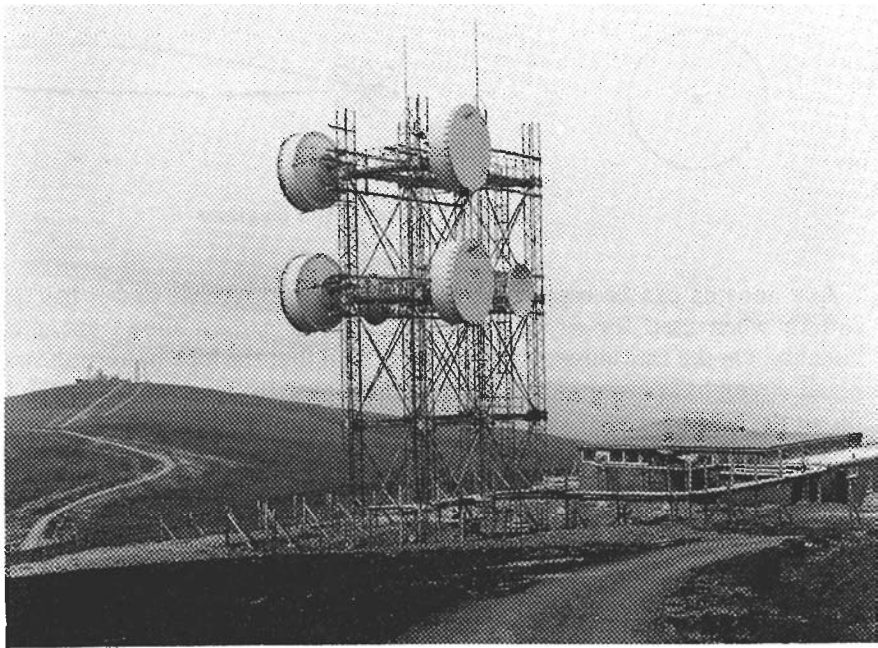
The directionality of an antenna is best illustrated by its *lobe diagram*, and two examples are given in Figure 8.17.

The lobes of an antenna lobe diagram show the transmitting and receiving efficiency of the antenna in the various directional orientations. The length of a lobe on a lobe diagram is proportional to the relative signal strength of the radiowave transmitted by the antenna in that particular direction. Hence, the circle in the omnidirectional antenna lobe diagram represents an equal strength of signal transmitted in all directions (*isotropic* radiation). Compare with this the directional aerial which has one very strong lobe in one particular direction, and a number of much smaller ones. This sort of lobe pattern is typical of many directional aerial types.

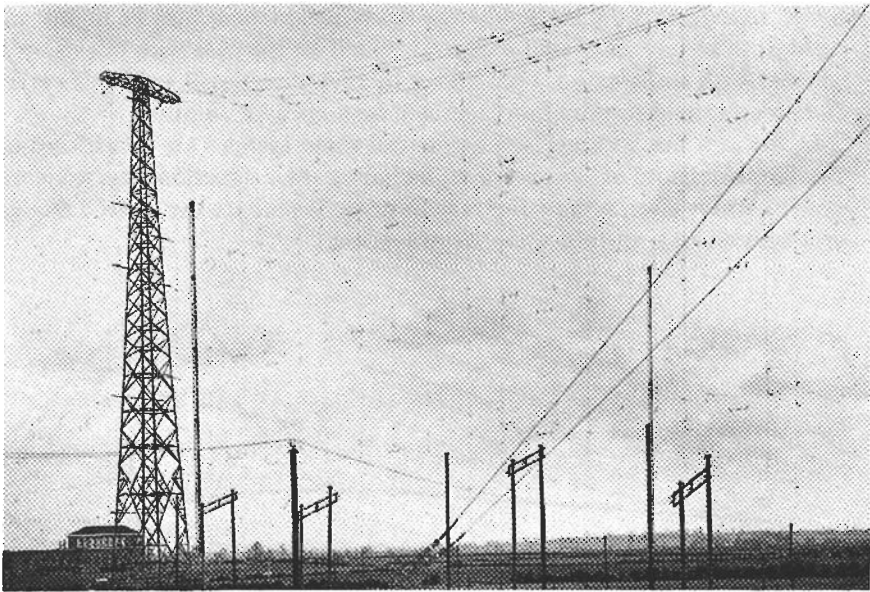
---

**Figure 8.16** (a) Microwave radio repeater station. A 68-foot tower and associated microwave radio repeater station, at the brow of a hill. The antenna dishes on opposite sides of the tower are slightly offset in direction to prevent the possibility of signal overshoot. (*Courtesy of BT Archives*) (b) Array antenna. Sixteen metre sterba array antenna. The elements of the antenna are the barely visible wires, strung from the main pylon, and arranged in a regular and rectangular formation. The pattern of repeated elements makes the antenna highly directional – cutting out all signals except that from the desired direction. (*Courtesy of BT Archives*)

(a)



(b)



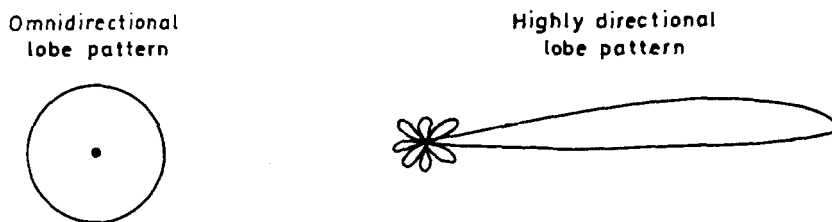


Figure 8.17 Aerial lobe diagrams

Any antenna can be used either to transmit or to receive, and it has the same lobe pattern when used for either purpose. The lobe pattern shows an antenna's directionality. On the transmission side we have seen that the lobe pattern shows the relative signal strengths transmitted in various directions. In 'receive mode', the lobe pattern shows the antenna's relative effectiveness in picking up radio waves from sources in the various directions. As we have said, a directional receive antenna also helps to reduce the likelihood of interference from other radio sources; the lobe pattern shown in Figure 8.18 shows a directional antenna being used to eliminate interference from a second radio source.

The receiving antenna R in Figure 8.18 is oriented so that its principle lobe is pointing at the transmitting source T1. Meanwhile, source T2 aligns with a *null* in the lobe pattern. The receiving antenna will thus be much more effective in reproducing the signal from T1 than in reproducing the signal from T2, so that the signal from T2 is suppressed relative to that from T1.

An antenna is required at each end of any radio link, but the two antennas need not be similar. For example, the transmitting antenna may be a high power, omnidirectional, broadcast radio mast, whereas receiving antennas may be highly directional, and perhaps relatively small and cheap, as would be the case with television broadcasting; whereas with ships' radio it is difficult to keep directional antennas correctly oriented, and so an omnidirectional antenna will be a cheaper solution.

In the next few sections, some practical radio systems are considered in more detail. The characteristics of the antennas, including their directionality, are described and an explanation is given of how the various antennas induce the desired mode of radiowave propagation (e.g. surface wave, skywave, etc.).

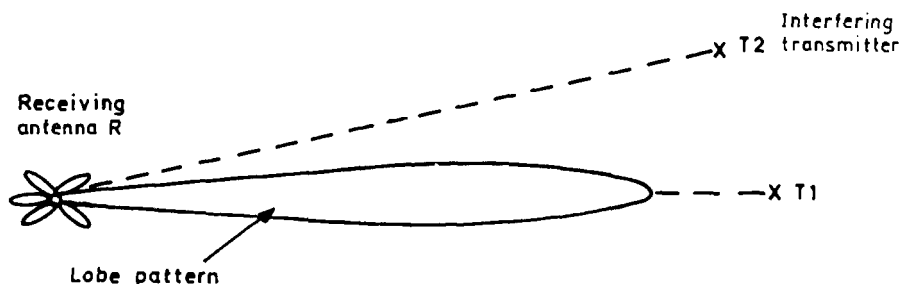


Figure 8.18 Excluding interfering radio transmitters

## 8.10 SURFACE-WAVE RADIO SYSTEMS

Surface-wave radio systems have a good range when using relatively low frequency radio waves. (In this context low frequency is the range 50 kHz–2 MHz). Surface-wave radio is typically used for broadcast radio transmissions, particularly for public radio stations, maritime radio, and navigation systems. Broadcast, surface-wave radio transmitters are usually very tall radio masts, several hundred feet high, as illustrated in Figure 8.19.

Broadcast radio transmitting antennas usually combine high power transmitters with an omnidirectional lobe pattern. The high power enables domestic radio receivers to be relatively unsophisticated, and therefore cheap. The omnidirectional pattern allows them to transmit over a wide area.

## 8.11 HIGH FREQUENCY (HF) RADIO

Radio waves in the frequency band 3–30 MHz propagate over greater distances in a *skywave* form than they do as a surface wave. For skywave propagation, a directional antenna is set up to transmit a radio wave at a specific angle to the horizon. The angle must not be so close to the vertical as to allow an escape wave such as we saw in Figure 8.15. The signal is transmitted via the earth's ionosphere (a region about 200–350 km above the earth's surface). Layers of the ionosphere which have different physical properties cause the signal to be refracted and eventually reflected back to the earth's surface, an effect shown in Figure 8.15. Unfortunately, high frequency skywave radio is very sensitive to weather conditions in the ionosphere. Time-of-day and sunspot cycles can break up reliable service, giving rise to fading and interference. Nonetheless, high frequency radio using skywaves was one of the earliest transmission methods employed for international telephone service. As early as 1927, the first commercial radio telephone service was in operation between Britain and the United States, many years before first transatlantic telephone cable was laid in 1956.

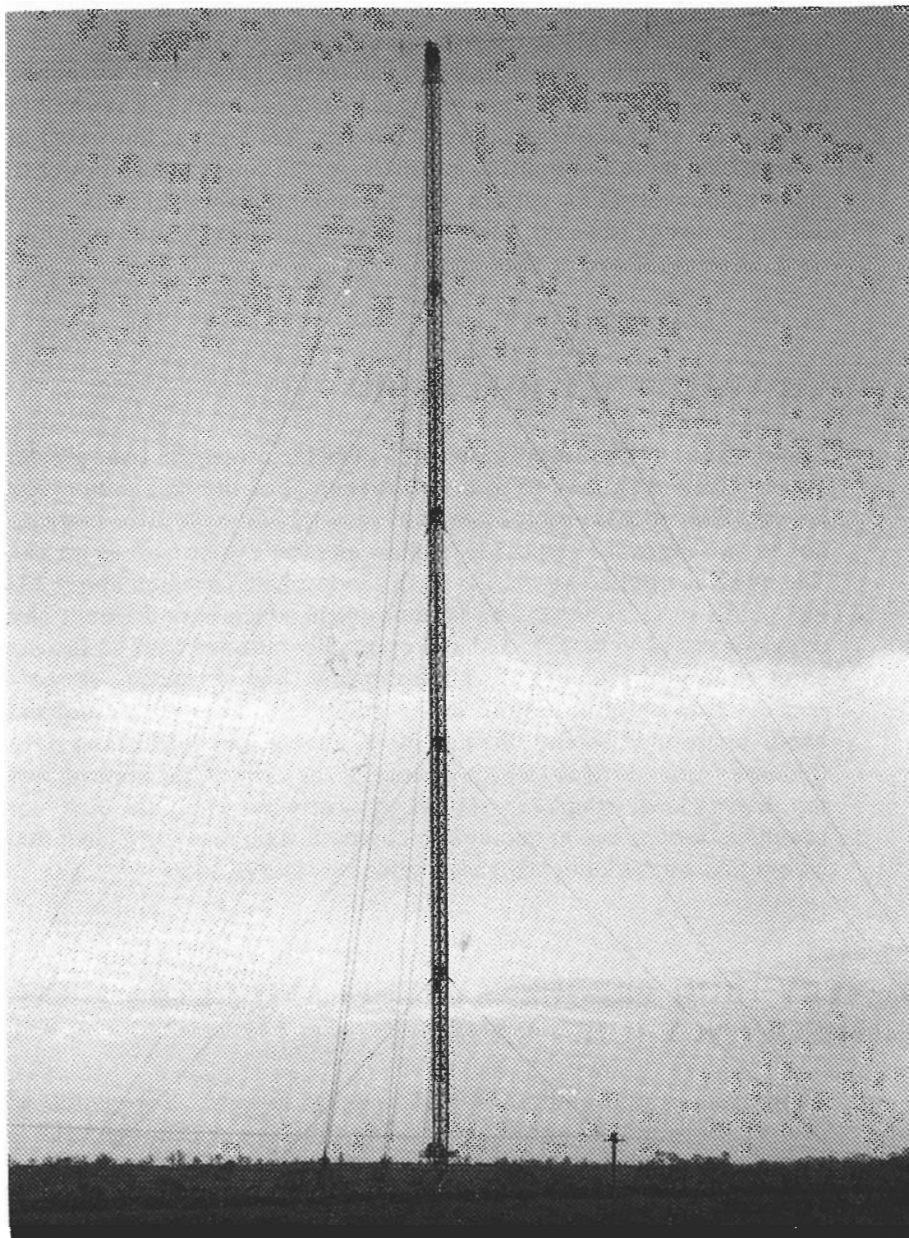
## 8.12 VERY HIGH FREQUENCY (VHF) AND ULTRA HIGH FREQUENCY (UHF) RADIO

At frequencies above about 30 MHz, neither skywave propagation nor surface-wave over-the-horizon propagation is possible. For this reason the VHF and UHF radio spectrum is used mainly for *line of sight* radio transmission systems. *Just-over-the-horizon* transmission is also possible, but only if the antennas are elevated clear of the *electrical ground effects* of the earth's surface.

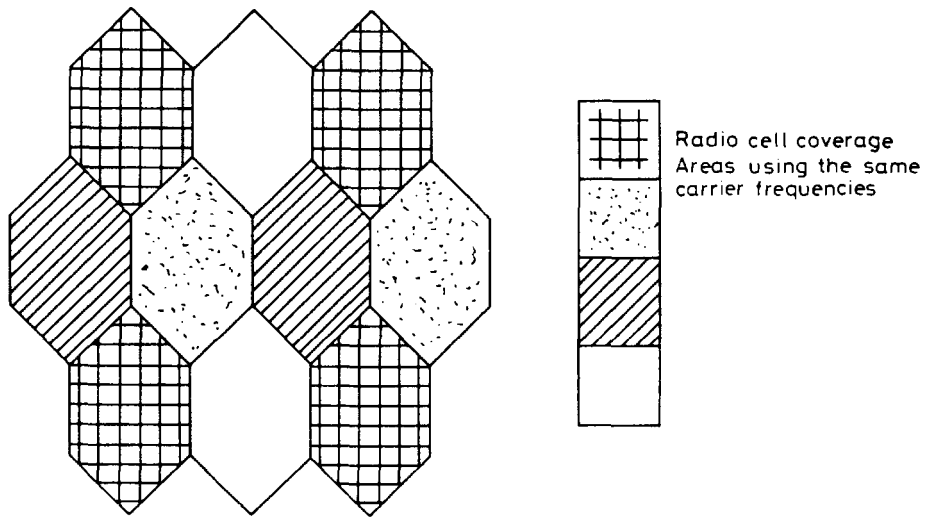
VHF and UHF radio systems are becoming increasingly common as the basis of a large number of applications:

- local radio stations
- citizens band (CB) radio





**Figure 8.19** Radio mast. 250 metre high radio mast, weighing 200 tonnes. The visible wires are probably steel, provided merely to support the mast. The active element is mounted within the main pylon. (Courtesy of BT Archives)



**Figure 8.20** Radio spectrum reuse in cellular radio networks

- cellular radio (mobile telephones)
- radiopagers

The advantage of VHF and UHF radio is that much smaller antennas can be used. It has made possible the wide range mobile communications handsets, some of which have antennas only a few inches long. A drawback is the restricted range of VHF and UHF systems, although this has been turned to advantage in cellular radio telephone systems, where a number of transmitters are used, each covering only a small *cell* area. Each cell has a given bandwidth of the radio spectrum to establish telephone calls to mobile stations within the cell. Adjacent cells use different radio bandwidth, thereby preventing radio interference between the signals at the cell boundaries. The fact that the radio range is short also means that a non-adjacent cell can *re-use* the same radio spectrum without chance of interference, so that very high radio spectrum utilization can be achieved. Figure 8.20 illustrates the *re-use* of radio spectrum, as achieved in cellular radio networks for portable telephone sets. Cells marked with the same shading are using identical radio bandwidth.

We shall return to the subject of cellular radio in Chapter 15.

### 8.13 MICROWAVE RADIO

Microwave is the name given to radio waves in the frequency above 1000 MHz (1 GHz). The prefix *micro* is in recognition of the very short wavelength (only a few centimetres). Microwave systems are commonly used as high capacity, point-to-point transmission systems in telecommunication networks, such as high-capacity trunk telephone network connections between major cities, or on a smaller scale (using smaller antennas)

between company offices. The high frequency and short wavelength of microwave radio allow high capacity radio systems to be built using relatively small but highly directional antennas. The small scale yields benefits in terms of cost, installation and maintenance.

Microwave antennas are operated in a line-of-sight mode, up to 40–50 kilometres apart as governed by the radio frequency in use, by the radio propagation conditions (including the terrain and the weather; see Chapter 33) and practically by the availability of suitable locations for radio towers.

Over the last few years, microwave radio has gained significantly in importance as major corporations and new network operators alike have sought to use radio techniques to bypass the transmission resources of the old monopoly network providers. Table 8.3 illustrates the range capabilities of the new frequency bands which have come into use, and for reference the uses allocated to each of the bands by ITU-R.

Distances beyond single hop links of 40–50 km are achieved using a multilink path, passing through a number of intermediate radio repeater stations. The path itself is

**Table 8.3** Frequency bands allocated by ITU-R for public network and corporate microwave radio links, either point-to-point or point-to-multipoint

Frequency band	Countries of application	Application
1 GHz	worldwide	cellular mobile radio
2 GHz	worldwide	cellular mobile radio DECT some point to point microwave
3 GHz	worldwide	3.5 GHz rural radio
4 GHz	worldwide	high capacity point to point (e.g. STM-1) satellite communications
6 GHz	worldwide	satellite communications
7–8 GHz	worldwide	longhaul point to point microwave (up to 50 km)
10 GHz		US: cellular base station backhaul UK: point to multipoint wireless local loop
11 GHz	worldwide	medium capacity (E3, T3)
12, 14 GHz	satellite communication	direct broadcast by satellite
13, 15 GHz	worldwide	short to medium haul point to point (25 km)
18, 23, 26 GHz	worldwide	short haul point to point (10–15 km)
28 GHz	worldwide	point to multipoint, multimedia applications
38 GHz	worldwide	short haul (5–7 km)
above 40 GHz	worldwide	extremely short haul (1–3 km)

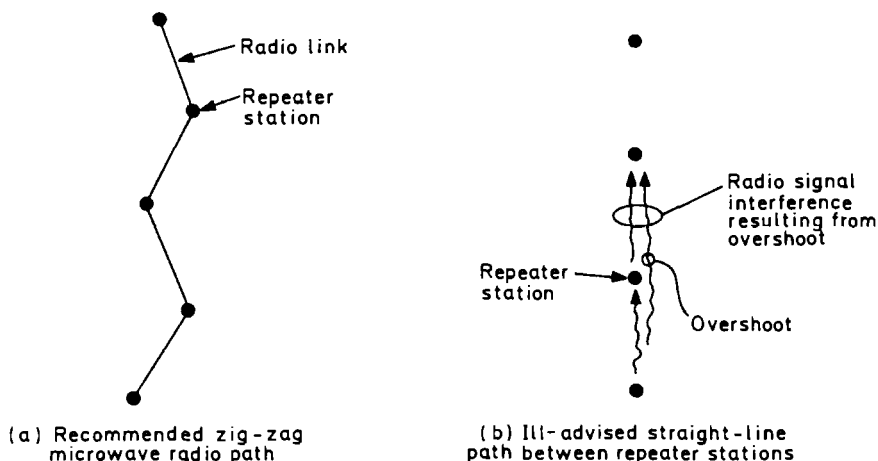


Figure 8.21 Microwave radio route planning

usually arranged in a slight zig-zag formation of *repeater stations*, as shown in Figure 8.21(a). At each repeater station there is a tower or radio mast, covered in a large number of microwave antennas. For each transmission path passing through the repeater stations there will be at least one dish antenna facing one *hop* of the path and one facing the next. (Two dishes may face each link, depending on whether the same antenna reflector is used for both transmit and receive purposes.)

The radio signal is received by one antenna and retransmitted in a slightly different direction on the next hop or link of the path by another antenna. As Figure 8.21(a) shows, the use of a zigzag path allows the use of the same radio carrier frequency on all the links of the overall path, without any risk of radio interference which might result from *overshoot* on a straight line path between repeater stations (Figure 8.21(b)).

Slight refraction caused by the earth's atmosphere causes radio waves to propagate along a curved path as shown in Figure 8.22, which has to be taken into account when setting up the antennas. Figure 8.22 also shows some of the path interference problems associated with trees and other obstacles which can be experienced with microwave radio links and need to be taken into account during installation.

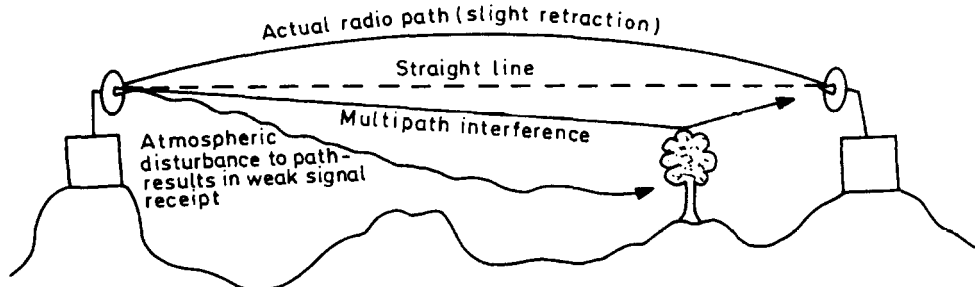


Figure 8.22 Microwave radio path and disturbances

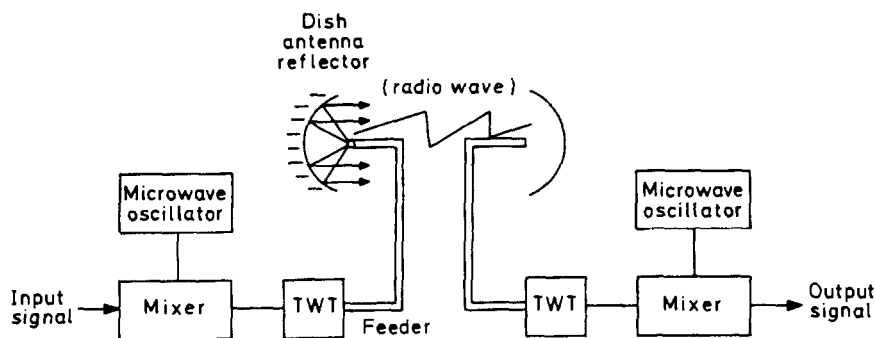


Figure 8.23 A microwave radio system

Like other forms of radio transmission, microwave radio is an effective way of spanning difficult terrain, although providing power for the intermediate *repeater stations* on long haul, multiple *hop* links can be a problem. Like other forms of radio propagation, the main problem with microwave systems is the susceptibility to fading and atmospheric signal absorption or interference due to the prevailing weather conditions, particularly heavy rain. Disturbed weather can cause radio path interference as shown in Figure 8.22. Owing to the short wavelength of microwaves, reflections off buildings and other obstacles near to the radio path, cause greater interference than experienced by longer wavelength signals. This is called *multipath interference*, and it too is shown in Figure 8.22. Careful path planning and choice of site for the antenna towers helps to minimize this problem.

The microwave carrier is produced by a microwave oscillator and is modulated in a similar way to high frequency radio signals, using modulating equipment called a mixer. The radiowave is then amplified immediately prior to transmission with the aid of a sophisticated and very high power gain amplifier, called a *travelling wave tube (TWT)*. The radio wave is then fed into the antenna itself using a hollow metal tube called a *radio waveguide* or *feeder*. The signal is usually emitted from the *focus* of a parabolic reflector antenna to produce a highly directional single beam. At the receiving end a similar aerial and set of electronics works in reverse order to reproduce the original signal. Figure 8.23 illustrates a simple microwave radio system.

## 8.14 TROPOSPHERIC SCATTER

It takes *tropospheric scatter* to make over-the-horizon communication possible. Radio waves appear to be reflected by the earth's atmosphere. It is not a well understood phenomenon and various explanations have been offered. Perhaps the easiest to grasp is the notion of radiowave reflection (commonly called *scatter*), caused by irregularities in the tropospheric region of the earth's atmosphere. The scatter occurs in a *common volume* of the earth's atmosphere, corresponding to the region 'visible' to both the receiving and transmitting tropospheric scatter antennas (Figure 8.24). The scatter angle is the angle of path deviation (or the *reflection angle*) caused by the scatter effect.

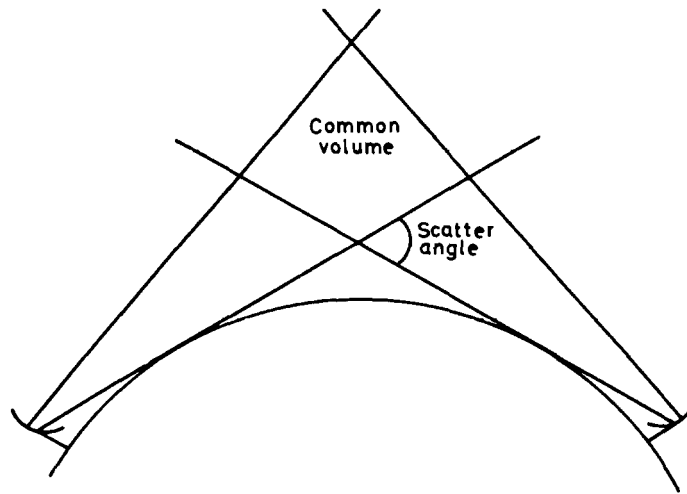


Figure 8.24 Tropospheric scatter

Large dish-shaped antennas are used for tropospheric scatter systems, together with microwave radio frequencies in the range of 800 MHz–5 GHz. Communication distances achievable with tropospheric scatter radio systems are 100–300 km. Their main drawback is the bad and continually varying signal fading that is experienced.

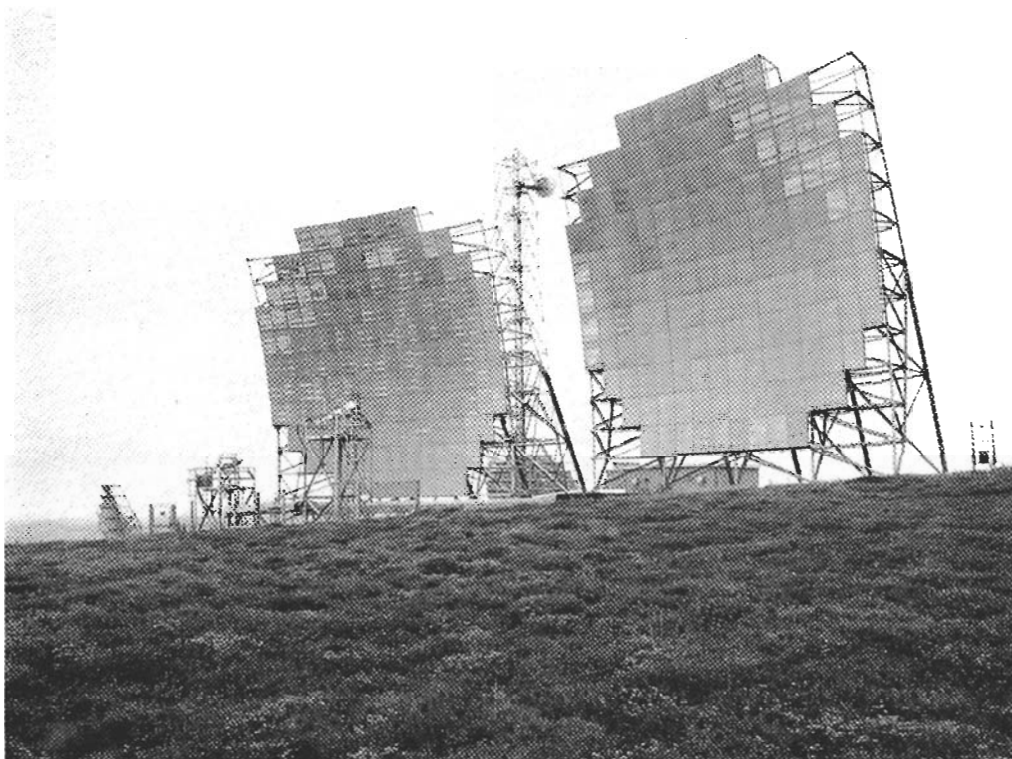
In the United Kingdom the tropospheric scatter method of *trans-horizon radio* transmission has become a standard method of communicating with offshore oil drilling platforms in the North sea.

## 8.15 SATELLITE SYSTEMS

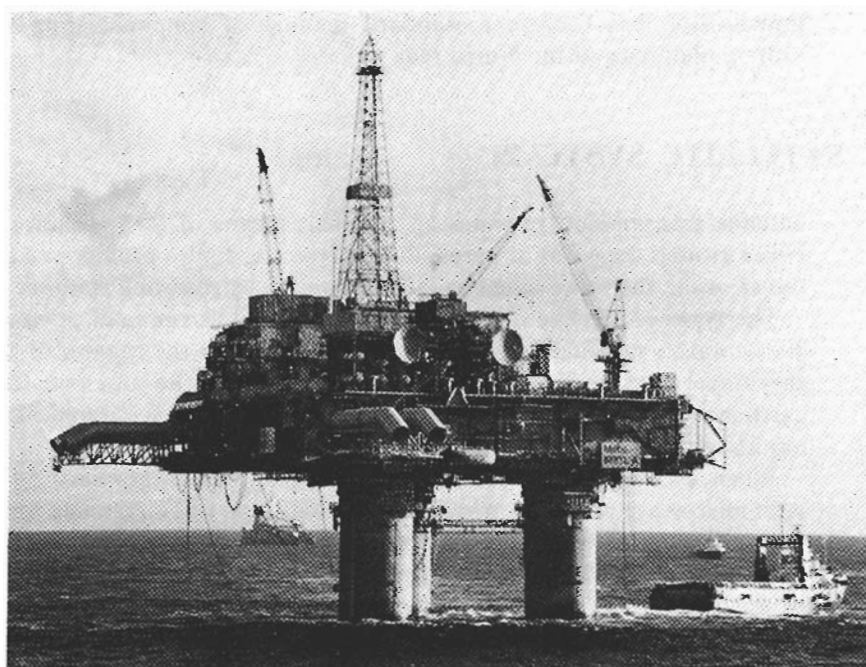
Satellite transmission provides an excellent means of long distance communication, either around the globe or across difficult terrain. It also provides an effective means of *broadcasting* the same signal to a large number of receiving stations.

The types of satellite most commonly used in telecommunications networks, called *geostationary* satellites, orbit the earth directly above the equator at such a height that they travel once around the earth's axis every 24 hours. Because both the satellite and the earth move at the same speed, the satellite appears to be geographically (or geo)stationary above a particular location on the equator.

When used for telecommunications purposes, a geostationary satellite is equipped with microwave antennas, which allow *line-of-sight* radio contacts between the satellite and other microwave antennas at *earth stations* on the ground. Communication between two earth stations can then be established by a tandem connection consisting of an *uplink* from the transmitting station to the satellite, and a *downlink* from the satellite to the receiving station. Onboard the satellite the uplink is *connected* to the downlink by a *responder* (receiver) for each uplink and a *transmitter* for each downlink, and because the two normally work in pairs they are often designed as a single piece of equipment usually referred to as a *transponder*.



**Figure 8.25** Land radio station for tropospheric scatter radio system. This one is located on the north-east coast of England, providing telecommunications services to oil rigs in the North Sea, about 100–300 km offshore. (*Courtesy of British Telecom.*)

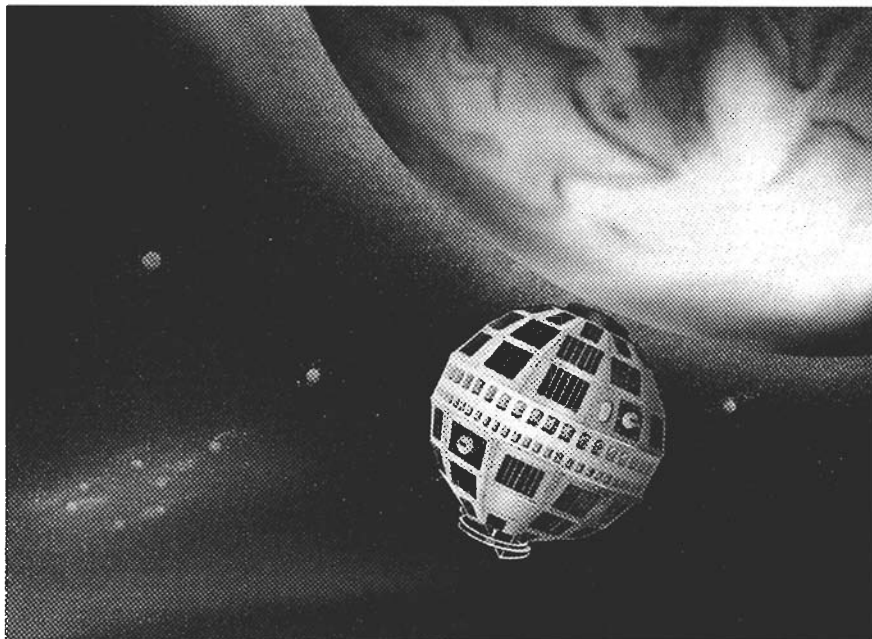


**Figure 8.26** Tropospheric scatter radio station. The oil rig receiving antennas for a tropospheric scatter radio system. (*Courtesy of BT Archives*)

The first international telecommunications satellite, called *Telstar* (Figure 8.27), was launched in 1962. This was a low orbiting satellite rather than a geostationary one, and it orbited the earth's axis faster than the earth's own spin rate. Consequently the satellite was mutually visible to American and European earth stations for three or four periods a day, but each period was of 30–40 minutes duration. Sophisticated *tracking* mechanisms were therefore required on the early earth stations to keep the reflector of the antenna pointing directly at the satellite as it made its path from one horizon to the other in about 30 minutes.

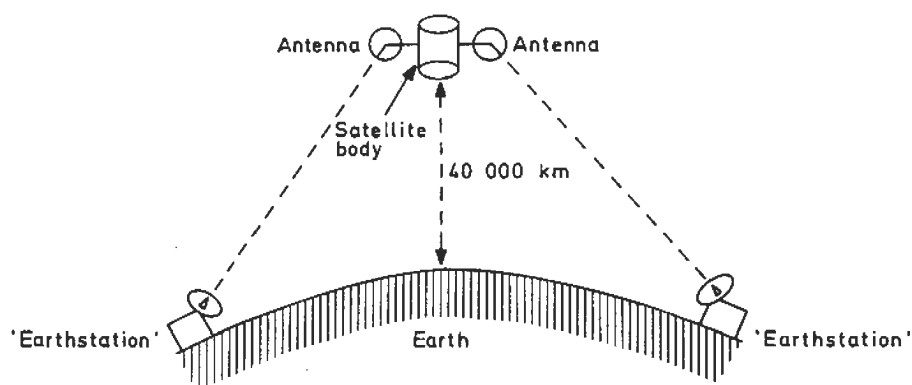
A major breakthrough in international satellite telecommunications came in 1964 with the establishment of the International Telecommunications Satellite consortium, or INTELSAT for short. The first generation of INTELSAT's series of Satellites, INTELSAT 1 (*Early Bird* to give it its popular name) was launched in 1965. *Early Bird* was a geostationary satellite, and the consortium of 11 countries (Australia, Canada, Denmark, France, Italy, Japan, Netherlands, Spain, UK, USA and Vatican City) which set up INTELSAT were able to take advantage of its services.

The INTELSAT satellites brought on a boom in international telecommunications throughout the late 1960s, 1970s and 1980s, owing to the high capacity of satellite transmission and the relatively low cost of the extremely long connections which it provided. (Satellite transmission costs are not determined by the distance between earth stations). The satellite's one drawback is the significant signal delay resulting from the



**Figure 8.27** Telstar satellite. The first international telecommunications satellite, Telstar, launched in 1962. It was a low orbiting satellite, about 0.9 metre in diameter, and supported 12 telephonic channels. (Courtesy of British Telecom.)

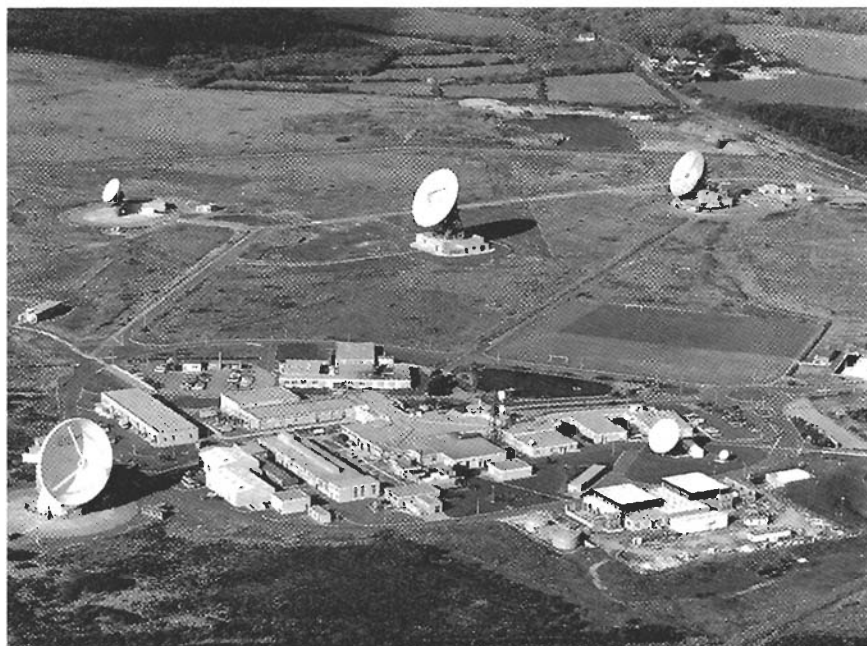




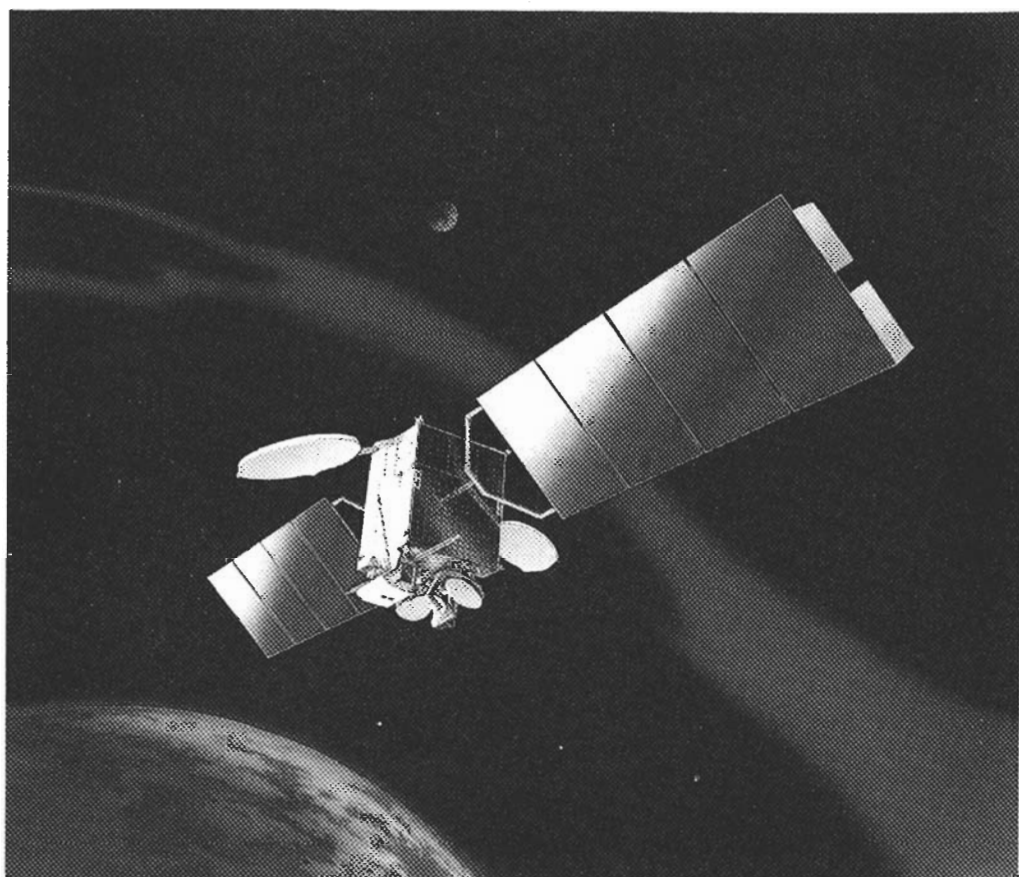
**Figure 8.28** A satellite transmission system

time it takes a microwave signal to travel up to and back from the satellite, 40 000 km above the earth's surface. Figure 8.28 illustrates a simple satellite transmission system.

Today's INTELSAT satellites are a far cry from their predecessors. Compared with the Telstar satellite, which was 0.9 m in diameter and 77 kg in mass, INTELSAT's eighth generation satellites, INTELSAT VIII which will serve the world's telecommunication needs in the early 21st century are truly 'spacecraft'. Towering edifices, 24 m in



**Figure 8.29** Satellite earth station. Goonhilly satellite earth station on the Lizard peninsula, Cornwall, England. One of the world's first established telecommunications satellite earth stations. Five main dishes and a number of smaller ones are visible. Note also the microwave tower, providing onward connection into the UK terrestrial network. (Courtesy of British Telecom.)



**Figure 8.30** INTELSAT VIII satellite. Artist's impression of the enormous INTELSAT VIII satellite, the workhorse for starting the 21st century. Twenty-four metres in height, 2 tonnes and 110 000 telephone channels (*Courtesy of INTELSAT*)

height and 2 tonnes in mass, each satellite supports up to 110 000 telephone channels. This compares with the meagre 12 circuits supported by the first Telstar.

Although global optical fibre cables are to some extent taking over from satellites as the preferred means of telecommunications transmission for telephone calls between developed countries, satellite systems are being rapidly deployed for improving telecommunications access to underdeveloped and remotely lying areas. Furthermore, their effectiveness in broadcasting information has brought a boom in their use for satellite television broadcasting. No doubt this trend will continue into the video broadcasting and *multimedia* age. Finally, another new application of satellite technology will come into use at around 1998 when a number of competing organizations intend to launch global mobile telephone services based on networks of multiple low orbiting satellites. In Chapter 15 we describe the planned systems of *Iridium*, *Globalstar*, *ICO* and *Teledesic* corporations.

## 8.16 'MULTIPLE ACCESS' RADIO AND SATELLITE SYSTEMS

We have said that one of the advantageous features of satellite and radio systems is their ability for broadcasting and point-to-multipoint usage. Thus a single radio channel and single signal can either be received by multiple parties simultaneously, or used in turn by each of the parties on an *on-demand* basis. The ability to broadcast is employed in modern small-dish satellite TV networks. The point-to-multipoint capability can be used in telephone-type networks. Thus a handful of radio frequencies may be sufficient to provide telephone network connections for thousands of end customers (provided they do not all need the telephone at once). The question is, how does one share the available radio bandwidth between the multiple users who want to use it simultaneously? The answer is, using a *multiple access* technique. Four main methods are worthy of consideration. These are *FDMA* (*frequency division multiple access*), *TDMA* (*time division multiple access*), *frequency hopping* (a second form of *TDMA*) and *CDMA* (*code division multiple access*).

### 8.16.1 FDMA (Frequency Division Multiple Access)

*FDMA* is based on *FDM* (*frequency division multiplex*, already discussed in Chapter 3 and earlier in this chapter). The radio base station (typically at the network local exchange) splits up the available radio bandwidth into different smaller frequency *bands* or *channels*, and allocates these to the individual end stations according to need. Thus a total available radio bandwidth of 100 kHz could be split into  $4 \times 25$  kHz individual channels as shown in Figure 8.31.

The frequencies could either be allocated for all-time (*PAMA*, *pre-assigned multiple access*), or *on-demand* (*DAMA*, *demand assigned multiple access*). By allocating the frequencies *on-demand*, they can be shared between many end users, the exact number depending on the number of maximum simultaneous connections needed by the users as a whole.

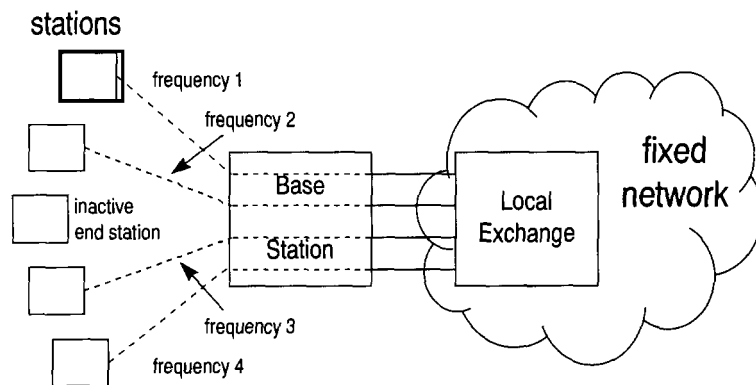


Figure 8.31 FDMA (frequency division multiple access)

### 8.16.2 TDMA (Time Division Multiple Access)

An alternative means of subdividing the available bandwidth is by means of *TDMA* (*time division multiple access*). *TDMA* is based on *TDM* (*time division multiplex*, see Chapter 5). In *TDMA* the entire radio bandwidth (in our example, of 100 kHz) is used by each of the end stations in turn, each using the signal for  $\frac{1}{4}$  of the available time. A burst of information is sent by the end station in a predetermined *timeslot* (say of 125  $\mu$ s duration) allocated by the base station. The base station transmits continuously on its transmit frequency, each station selecting the information only from the *timeslot* which is relevant to itself (Figure 8.32).

Both *FDMA* and *TDMA* can be used as the basis of digital radio systems. It is erroneous to consider *FDMA* to be only an *analogue* technique, because each of the individual channels can be coded digitally. The prime advantage of *TDMA* when connecting the radio system to a digital fixed network, however, is that the output of the base station towards the fixed network (for example, telephone exchange) is usually a single, high bitrate multiplexed digital bitstream (e.g. 2 Mbit/s). In the *FDMA* case multiple 64 kbit/s telephone channels would be presented. Another advantage of *TDMA* over *FDMA* is that only a single, high bitrate radio modem is required in the base station (as opposed to one modem per channel in the case of *FDMA*). The disadvantage is that the control system required for *TDMA* is more complex than for *FDMA*.

The prime advantage of *FDMA* is the relative simplicity of the system. The disadvantages are the multiple modems needed at the base station, the multiple connections needed between the base station and the fixed network and the relatively fixed subdivisions of the available capacity (i.e. discrete, fixed bitrate channels). In

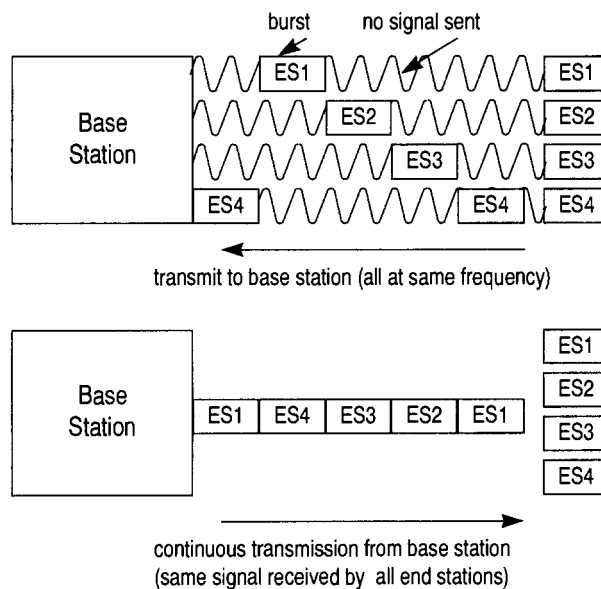


Figure 8.32 TDMA (time division multiple access)

comparison, TDMA can be relatively easily adapted to provide almost any subdivision of the capacity (channels of almost any bitrate), simply by adjusting the duration of the *timeslot* allocated to each end station.

### 8.16.3 Frequency Hopping

*Frequency hopping* could be thought of as a mix of the FDMA and TDMA techniques. A number of bearer frequencies are used by each of the end stations in turn, but each only transmits on one of the frequencies for a short burst or *timeslot*. Immediately after the timeslot, the station moves to another bearer frequency. Each station thus hops in a given pattern between a set of pre-defined frequencies (*frequency hopping*). Figure 8.33 illustrates the technique. Thus end station 1 (ES1) transmits cyclically on frequencies f1, f3, f4, f2 and then f1 again, etc.

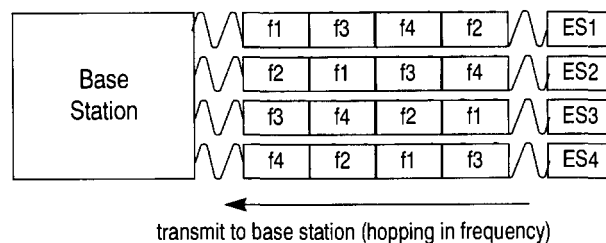
The advantage of *frequency hopping*, one of a number of *spread spectrum techniques*, is the relative immunity to radio *fading*. As it is unlikely that each radio channel will be simultaneously degraded by *fading*, the effect of a single channel fading can be virtually eliminated by 'diluting' the degradation over a large number of active user connections. Another benefit is the much greater complexity of equipment needed by criminals wishing to 'tap' the radio channel.

### 8.16.4 CDMA (Code Division Multiple Access)

In the most recently developed *spread spectrum technique*, *CDMA* (*code division multiple access*), the carrier frequency is first modulated by the digital signal to be carried, and is then multiplied by a digital signal of a much higher bit rate (called the *chip rate*). The *chip rate* signal is actually a specific binary number (or code). It has the effect of *spreading* the modulated carrier over a very large frequency band, equal to the spreading factor (the ratio of the *chip rate* to the carried *bit rate*).

At the receiver, the signal is decoded using the same binary *code*, by division of the received signal. This regenerates the original modulated carrier.

Many signals may be modulated onto the same carrier frequency by using other *codes*. These are similarly recovered from the 'mixture' of modulated signals by division by their respective codes. During the division process, all the signals are of course divided by the *code*, but only one of the modulated signals is reduced to a signal closely



**Figure 8.33** Frequency hopping

grouped around the original carrier frequency. The other coded signals can thus be removed by *filtering* before the carrier is demodulated to recover the original digital bit stream.

The advantages of *CDMA*, like frequency hopping, is the resistance to multipath fading and interference resulting from the *spread spectrum* technique. In addition, CDMA can be realized relatively cheaply using digital signal processing, requires little coordination between the stations, and requires a lower transmit power for low bit rate signals. The disadvantage is the relatively high bandwidth required for efficient operation and the close coordination required to ensure good isolation of the individual channels from one another.

## 8.17 ELECTROMAGNETIC INTERFERENCE (EMI) AND ELECTROMAGNETIC COMPATIBILITY (EMC)

We have seen how the interference of electrical signals is one of the prime factors limiting the performance of telecommunication systems. The subject has become one of major concern, and extensive technical standards have been developed to define the measurement of stray emissions from devices and the sensitivity of devices to interference from ambient electromagnetic radiation (e.g. signal ‘spikes’ on the power supply, radio signal interference or lightning). Within many countries nowadays, devices are legally required to conform to new *EMI* (*electromagnetic interference*) and *EMC* (*electromagnetic compatibility*) standards. These ensure the device’s insensitivity to interference (EMC) and guarantee they themselves do not produce stray emissions (EMI). The commonly known standards are those enforced by the United States *FCC* (*Federal Communications Commission*) and *ETSI* (*European Telecommunications Standards Institute*).



---

# *Data Network Principles and Protocols*

---

We have described the binary form in which data are held by computer systems, and how such data are conveyed over digital line transmission systems, but we shall need to know more than this before we can design the sort of devices which can communicate sensibly with one another in something equivalent to human conversation. In this chapter we shall discuss in detail the conveyance of data between computer systems, the networks required, and the so-called 'protocols' they will need to ensure that they are communicating properly. In the second half of the chapter some practical data and computer network topologies and the equipment needed to support them are described.

## 9.1 COMPUTER NETWORKS

Between 1950 and 1970 computers were large and unwieldy, severely limited in their power and capabilities, and rather unreliable. Only the larger companies could afford them, and they were used for *batch-processing* scientific, business or financial data on a large scale. Data storage in those days was laborious, limited in capacity, long in preparation, not at all easy to manage. Many storage mechanisms (e.g. paper tape and punched cards) were very labour-intensive; they were difficult to store, and prone to damage. Even magnetic tape, when it appeared, had its drawbacks; digging out some trivial information 'buried' in the middle of a long tape was a tedious business, and the tapes themselves had to be painstakingly protected against data corruption and loss caused by mechanical damage or nearby electrical and magnetic fields.

Computing was for specialists. Computer centre staff looked after the *hardware*, while *software* experts spent long hours improving their computer programmes, squeezing every last drop of 'power' out of the computers' relatively restricted capacity.

By the mid-1970s all this began to change, and very rapidly. Cheap semiconductors heralded the appearance of the microcomputer, which when packaged with the newly developed floppy diskette systems, opened a new era of cheap and widespread computing activity. *Personal computers (PCs)* began to appear on almost every manager's



desk, and many even invaded peoples' homes. Suddenly, computing was within reach of the masses, and the creation and storage of computer data was easy, cheap and fast.

All sorts of individuals began to prepare their own isolated databases and to write computer programmes for small scale applications, but these individuals soon recognized the need to share information and to pass data between different computers. This could be done by transferring floppy diskettes from one machine to another, but as time went on that method on its own proved inadequate. There was a growing demand for more geographically widespread, rapid, voluminous data transfer. More recently, the demands of *distributed processing* computer networks comprising *clients* and *servers* have created a boom in demand for data networks.

A very simple computer or *data network* consists of a computer linked to a piece of peripheral equipment, such as a printer. The link is necessary so that the data in the computer's memory can be reproduced on paper. The problem is that a 'wires-only' direct connection of this nature is only suitable for very short connections, typically up to about 20 metres. Beyond this range, some sort of *line driver* telecommunications technique must be used. A number of techniques are discussed here. A long distance point-to-point connection may be made using *modems*. A slightly more complex computer network might connect a number of computer terminals in outlying buildings back to a *host* (mainframe computer) in a specialized data centre. Another network might be a *Local Area Network (or LAN)*, used in an office to interconnect a number of desktop computing devices, laser printers, data storage devices (e.g. file servers), etc. More complex computer networks might interconnect a number of large mainframe computers in the major financial centres of the world, and provide dealers with 'up-to-the-minute' market information.

The basic principles of transmission, as set out in the early chapters of this book, apply equally to data which are communicated around computer networks. So circuit-switched networks or simple point-to-point lines may also be used for data communication. Data communication, however, makes more demands on its underlying network than a voice or analogue signal service, and additional measures are needed for coding the data in preparation for transmission, and in controlling the flow of data during transmission. Computers do not have the same inherent 'discipline' to prevent them talking two at a time. For this reason special *protocols* are used in data communication to make quite sure that information passing between computers is correct, complete and properly understood.

## 9.2 BASIC DATA CONVEYANCE: INTRODUCING THE DTE AND THE DCE

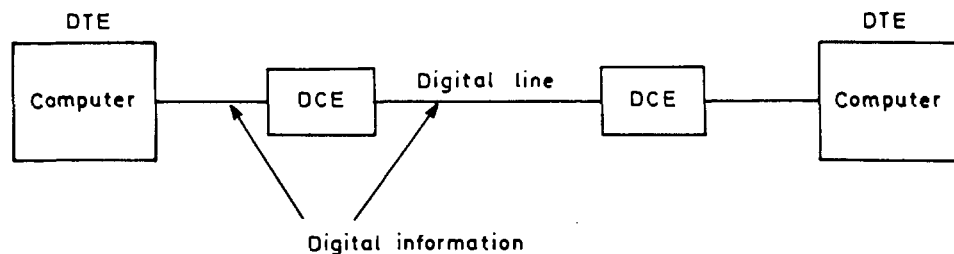
As we learned in Chapter 4, data are normally held in a computer or computer storage medium in a binary code format, as a string of digits with either value '0' or value '1'. A series of such binary digits can be used to represent alphanumeric characters (e.g. ASCII code), or graphical images, such as those transmitted by facsimile machines, video or *multimedia* signals.

In Chapter 5 we went on to discuss the principles of digital transmission, and found that it was ideal for the conveyance of binary data. Digital transmission has become the

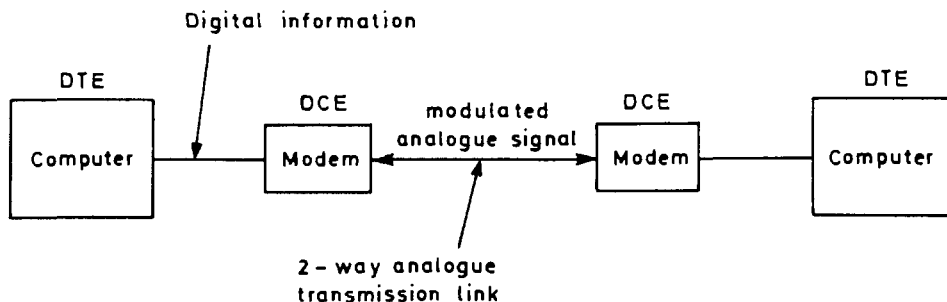
backbone of both private and public networks. However, despite the increasing availability and ideal suitability of digital transmission for data communication, it is unfortunately not always available. In circumstances where it is not, digitally-oriented computer information must instead be translated into a form suitable for transmission across an analogue network. This translation is carried out by a piece of equipment called a *modulator/demodulator*, or *modem* for short. Modems transmit data by imposing the binary (or *digital*) data stream onto an audio frequency carrier signal. The process is very similar to that used in the frequency division multiplexing of voice channels described in Chapter 3.

Figure 9.1 illustrates two possible configurations for data communication between two computers using either a digital or an analogue transmission link. The configurations look very similar, comprising the computers themselves (these are specific examples of *data terminal equipment (DTE)*); sandwiched between them in each case is a line and a pair of *data circuit terminating equipments (DCE)*.

A digital DCE (Figure 9.1(a)) connects the customer's digital DTE to a digital transmission line, perhaps provided by the public telecommunications operator (PTO). The DCE provides several network functions. In the transmit direction, it regenerates the digital signal provided by the DTE and converts it into a standardized format, level and line code suitable for transmission on the digital line. More complex DCEs may also interpret the signals sent by the DTE to the network to indicate the address desired



(a) Digital line connection



**Figure 9.1** Point-to-point connection of computers. DTE, Data Terminal Equipment; DCE, Data Circuit Terminating Equipment

when establishing a connection, and help in setting up the call. In the receive direction, the DCE establishes a reference voltage for use of the DTE and reconverts the received line signal into a form suitable for passing to the DTE. In addition, it uses the clocking signal (i.e. the exact bit rate of the received signal) as the basis for its transmitting bit rate. The received clocking signal is used because this is derived from the highly accurate *master clock* in the PTO's network. Two interfaces need to be standardized. These are the DTE/DCE interface and the DCE/DCE interface.

Digital DCEs can be used to provide various digital bit speeds, from as low as 2.4 kbit/s, through the standard channel of 64 kbit/s, right up to higher order systems such as 1.544 Mbit/s (called T1 or DS1), 2.048 Mbit/s (called E1), 45 Mbit/s (called T3 or DS3) or higher. When bit speeds below the basic channel bit rate of 64 kbit/s are required by the DTE, then the DCE has an additional function to carry out, breaking down the line bandwidth of 64 kbit/s into smaller units in a process called *sub-rate multiplexing*. The process derives a number of lower bit rate channels, such as 2.4 kbit/s, 4.8 kbit/s, 9.6 kbit/s, etc., some or all of which may be used by a number of different DTEs.

In their various guises, digital DCEs go by different names. In the United Kingdom, British Telecom's *Kilostream* digital private line service uses a DCE called a *NTU* (*network terminating unit*). Meanwhile, in the United States, AT&T's *digital data system (DDS)* and similar digital line services are provided by means of *channel service units (CSU)* or *data service units (DSU)*. Sometimes the term *LTU* (*line terminating unit*) is used.

In the case of analogue transmission lines, the DCE (Figure 9.1(b)) cannot rely on the network to provide accurate clocking information, because the bit rate of the received signal is not accurately set by the PTO's analogue network. For this reason, an internal clock is needed within the DCE in order to maintain an accurate transmitting bit rate. The other major difference between analogue and digital DCEs is that analogue DCEs (modems) need to convert and reconvert digital signals received from the DTE into an analogue signal suitable for line or radio transmission.

### 9.3 MODULATION OF DIGITAL INFORMATION OVER ANALOGUE LINES USING A MODEM

Three basic datamodulation techniques are used in modems for conveying digital end user information across DCE/DCE interfaces comprising analogue lines or radio systems, but there are more sophisticated versions of each type and even hybrid versions, combining the various techniques. Each modem is in reality a specialized device for a specific line or radio system type. The three techniques are illustrated in Figure 9.2 and described below.

#### *Amplitude modulation*

Modems employing *amplitude modulation* alter the amplitude of the carrier signal between a set value and zero (effectively 'on' and 'off') according to the respective value '1' or '0' of the modulating bit stream. Alternatively two different, non-zero values of amplitude may be used to represent '1' and '0'.

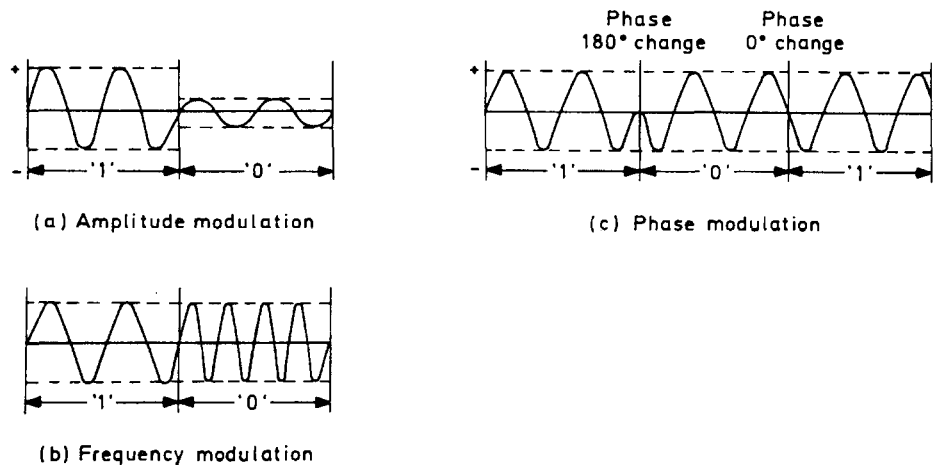


Figure 9.2 Datamodulation techniques

### Frequency modulation

In *frequency modulation* (Figure 9.2(b)), it is the frequency of the carrier signal that is altered to reflect the value '1' or '0' of the modulating bit stream. The amplitude and phase of the carrier signal is otherwise unaffected by the modulation process. If the number of bits transmitted per second is low, then the signal emitted by a frequency modulated *modem* is heard as a 'warbling' sound, alternating between two frequencies of tone. Modems using frequency modulation are more commonly called *FSK*, or *frequency shift key* modems. A common form of *FSK* modem uses four different frequencies (or tones), two for the transmit direction and two for the receive. This allows simultaneous sending and receiving (i.e. *duplex* transmission) of data by the modem, using only a two-wire circuit.

### Phase modulation

In *phase modulation* (Figure 9.2(c)), the carrier signal is advanced or retarded in its *phase* cycle by the modulating bit stream. At the beginning of each new bit, the signal will either retain its phase or change its phase. Thus in the example shown the initial signal phase represents the value '1'. The change of phase by 180° represents next bit '0'. In the third bit period the phase does not change, so the value transmitted is '1'. *Phase modulation* (often called *phase shift keying* or *PSK*) is conducted by comparing the signal phase in one time period to that in the previous period; thus it is not the absolute value of the signal phase that is important, rather the phase change that occurs at the beginning of each time period.

## 9.4 HIGH BIT RATE MODEMS

The transmission of high bit rates can be achieved by modems in one of two ways. One is to modulate the carrier signal at a rate equal to the high bit rate of change of the

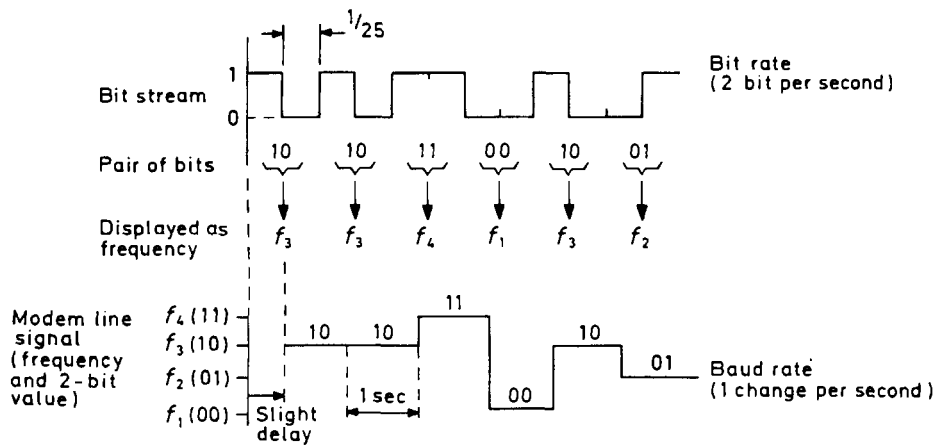


Figure 9.3 Multilevel transmission and lower baud rate

modulating signal. Now the rate (or frequency) at which we fluctuate the carrier signal is called the *baud rate*, and the disadvantage of this first method of high bit rate carriage is the equally high *baud rate* that it requires. The difficulty lies in designing a modem capable of responding to the line signal changes fast enough. Fortunately an alternative method is available in which the baud rate is lower than the bit rate of the modulating bit stream. The lower baud rate is achieved by encoding a number of consecutive bits from the modulating stream to be represented by a single line signal state. The method is called *multilevel transmission*, and is most easily explained using a diagram. Figure 9.3 illustrates a bit stream of 2 bits per second (2 bit/s being carried by a modem which uses four different line signal states. The modem is able to carry the bit stream at a *baud rate* of only 1 per second (1 Baud)).

The modem used in Figure 9.3 achieves a lower baud rate than the bit rate of the data transmitted by using each of the line signal frequencies  $f_1, f_2, f_3$  and  $f_4$  to represent two consecutive bits rather than just one. This means that the line signal always has to be slightly in delay over the actual signal (by at least one bit as shown), but the benefit is that the receiving modem will have twice as much time to detect and interpret the datastream represented by the received frequencies. *Multi-level transmission* is invariably used in the design of very high bit rate modems.

## 9.5 MODEM 'CONSTELLATIONS'

At this point we introduce the concept of modem constellation diagrams, because these assist in the explanation of more complex *amplitude* and *phase-shift-keyed (PSK)* modems. Figure 9.4 illustrates a *modem constellation* diagram composed of four dots. Each dot on the diagram represents the relative phase and amplitude of one of the possible line signal states generated by the modem. The distance of the dot from the origin

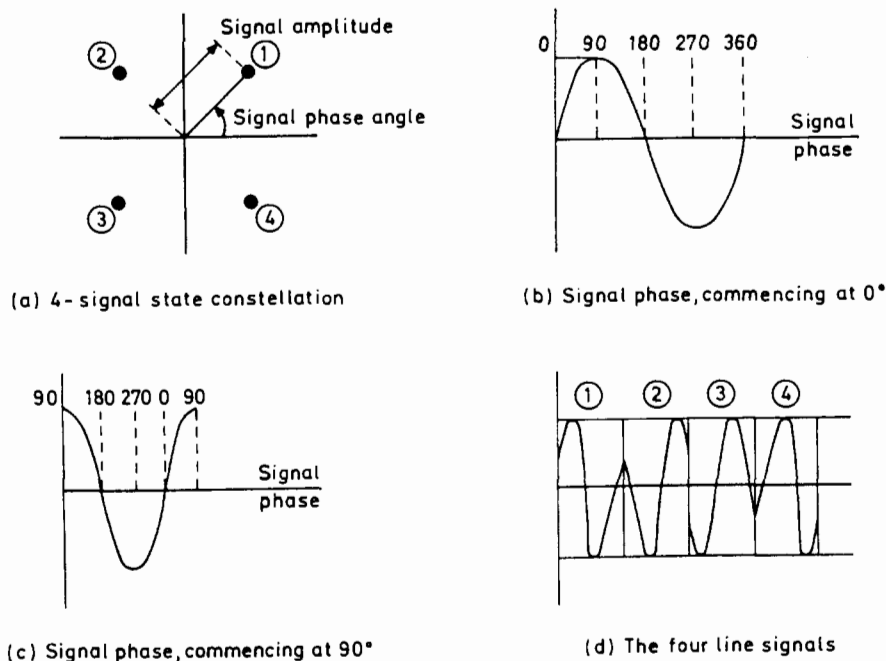


Figure 9.4 Modem constellation diagram

of the diagram axes represents the amplitude of the signal, and the angle subtended between the X-axis and a line from the point of origin represents the signal phase relative to the signal state in the preceding instant of time.

Figure 9.4(b) and (c) together illustrate what we mean by *signal phase*. Figure 9.4(b) shows a signal of  $0^\circ$  phase, in which the time period starts with the signal at zero amplitude and increases to maximum amplitude. Figure 9.4(c), by contrast, shows a signal of  $90^\circ$  phase, which commences further on in the cycle (in fact, at the  $90^\circ$  phase angle of the cycle). The signal starts at maximum amplitude but otherwise follows a similar pattern. Signal phases, for any phase angle between  $0$  and  $360^\circ$  could similarly be drawn. Returning to the signals represented by the constellation of Figure 9.4(a) we can now draw each of them, as shown in Figure 9.4(d) (assuming that each of them was preceded in the previous time instant by a signal of  $0$  phase). The phase angles in this case are  $45^\circ$ ,  $135^\circ$ ,  $225^\circ$ ,  $315^\circ$ .

We are now ready to discuss a complicated but common modem modulation technique known as *quadrature amplitude modulation*, or *QAM*. *QAM* is a technique using a complex hybrid of phase (or *quadrature*) as well as *amplitude modulation*, hence the name. Figure 9.5 shows a simple eight-state form of *QAM* in which each line signal state represents a three-bit signal (values nought to seven in binary can be represented with only three bits). The eight signal states are a combination of four different relative phases and two different amplitude levels. The table of Figure 9.5(a) relates the individual three-bit patterns to the particular phases and amplitudes of the signals that represent them. Note: Figure 9.5(b) illustrates the actual line signal pattern that would result if we sent the signals in the table consecutively as shown. Each signal

Bit combination	Line signal	
	Amplitude	Phase shift
000	low	0°
001	high	0°
010	low	90°
011	high	90°
100	low	180°
101	high	180°
110	low	270°
111	high	270°

(a) Bit combinations and line signal attributes

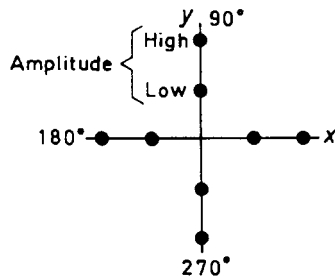
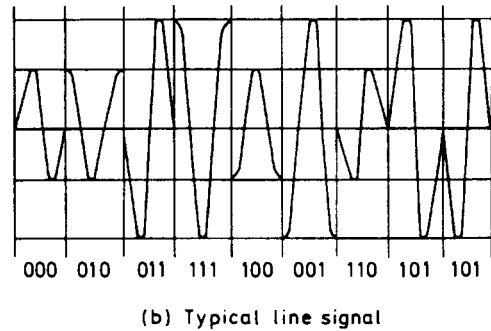


Figure 9.5 Quadrature amplitude modulation

is shown in the correct phase state relative to the signal in the previous time interval (unlike Figure 9.4 where we assumed that each signal individually had been preceded by one of 0 phase). Thus in Figure 9.5(b) the same signal state is not used consistently to convey the same three-bit pattern, because the phase difference with the previous time period is what counts, not the absolute signal phase. Hence the eighth and ninth time periods in Figure 9.5(b) both represent the pattern '101', but a different line is used, 180° phase shifted.

Finally, Figure 9.5(c) shows the constellation of this particular modem.

To finish off the subject of modem constellations, Figure 9.6 presents, without discussion, the constellation patterns of a couple of very sophisticated modems, specified by ITU-T recommendations V.22 bis and V.32 as a DCE/DCE interface. As in Figure 9.5, the constellation pattern would allow the interested reader to work out the respective 16 and 32 line signal states. Finally, Table 9.1 lists some of the common modem types and their uses. When reading the table, bear in mind that *synchronous* and *asynchronous operation* is to be discussed later in the chapter, and that *half-duplex* means that two-way transmission is possible but in only one direction at a time. This differs from simplex operation (as discussed in Chapter 1) where one-way transmission only is possible.

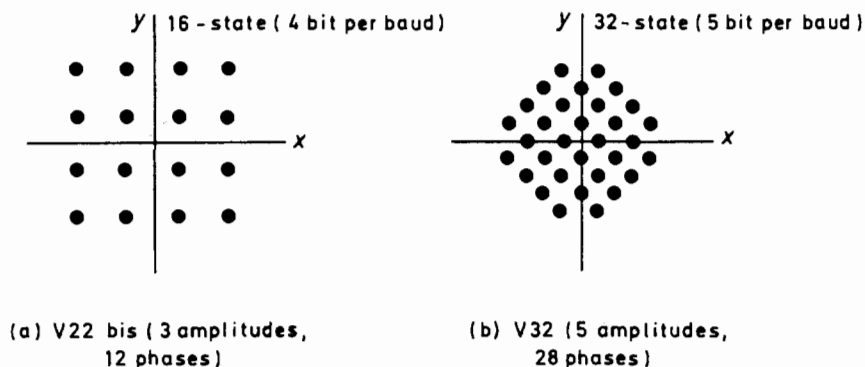


Figure 9.6 More modem constellations

Table 9.1 Common modem types

Modem type (ITU-T recommendation)	Modulation type	Bit speed (bit/s)	Synchronous (S) or asynchronous (A) operation	Full or half duplex	Circuit type required
V.21	FSK	up to 300	A	Full	2w telephone line
V.22	PSK	1 200	S or A	Full	2w telephone line
V.22 bis	QAM	2 400	S or A	Full	2w telephone line
V.23	FSK	600/1 200	A	Full	4w leaseline
				Half	2w telephone line
V.26	PSK	2 400	S	Full	4w leaseline
V.26 bis	PSK	2 400/1 200	S	Half	2w telephone line
V.27	PSK	4 800	S	Half	2w leaseline
V.27 ter	PSK	4 800	S	Half	2w leaseline
V.29	QAM	9 600	S	Full	4w leaseline
V.32	QAM	up to 9 600	S or A	Full	2w telephone line
V.33	QAM	14 400	S or A	Full	2w telephone line
V.34	QAM	up to 33 600	S or A	Full	2w telephone line
V.35	AM	48 000	wideband	Full	groupband leaseline
V.36	AM	48 000	wideband	Full	groupband leaseline
V.37	AM	72 000	wideband	Full	groupband leaseline
V.42	—	—	convert synchronous to asynchronous format	—	Error correcting protocol
V.42 bis	—	up to 30 kbit/s in association with V.32 modem	—	—	Data compression technique



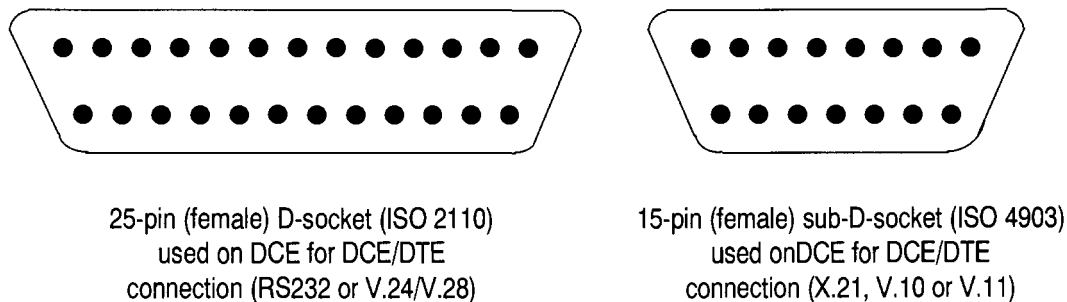
## 9.6 COMPUTER-TO-NETWORK INTERFACES

Returning now to Figure 9.1, we can see that we have discussed in some detail the method of data conveyance over the telecommunications line between one DCE and the other (or between one modem and the other) but what about the interface that connects a computer to a DCE or modem? This interface is of a generic type (DTE/DCE, i.e. between DTE and DCE) and can conform to any one of a number of standards, as specified by various organizations. The interface controls the flow of data between DTE and DCE, making sure that the DCE has sufficient instructions to deliver the data correctly. It also allows the DCE to prepare the distant DTE and confirm receipt of data if necessary. Physically, the interface usually takes the form of a multiple pin connection (plug and socket) typically with 9, 15, 25 or 37 pins arranged in a 'D-formation' (so-called *D-socket* or *sub-D-socket*). Computer users will be familiar with the type of sockets shown in Figure 9.7.

The interface itself is designed either for *parallel data transmission* or *serial data transmission*, the latter being the more common. The two different methods of transmission differ in the way each eight-bit data pattern is conveyed. Internally, computers operate using the parallel transmission method, employing eight *parallel* circuits to carry one bit of information each. Thus during one time interval all eight bits of the data pattern are conveyed. The advantage of this method is the increased computer processing speed which is made possible. The disadvantage is that it requires eight circuits instead of one. Parallel data transmission is illustrated in Figure 9.8(a), where the pattern '10101110' is being conveyed over the eight parallel wires of a computer's *data bus*. Note: nowadays 16-bit, and even 32-bit *data buses* are used in the most advanced computers. Examples of parallel interfaces are those specified by ITU-T recommendations V.19 and V.20.

*Serial transmission* requires only one transmission circuit, and so is far more cost-effective for data transmission on long links between computers. The parallel data on the computer's bus is converted into a serial format simply by 'reading' each line of the bus in turn. The same pattern '10101110' is being transmitted in a serial manner in Figure 9.8(b). Note that the baud rate needed for serial transmission is much higher than for the equivalent *parallel transmission* interface.

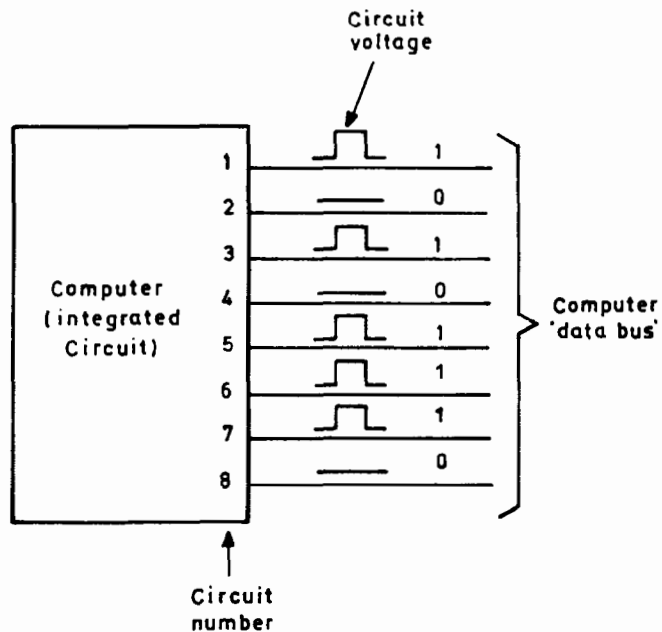
Many DTE-to-DCE (or even short direct DTE-to-DTE connections) use a serial transmission interface conforming to one of ITU-T's suites of recommendations, either X.21 or X.21bis. The basic functions of all types of DTE/DCE interface are similar:



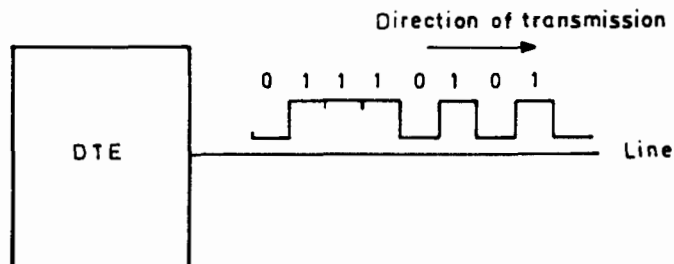
**Figure 9.7** Plug and socket arrangement for common DTE/DCE cables

- *clocking* and *synchronizing* the data transfer
- regulating the bit rate, so that the receiving device does not become swamped with data.

In Figure 9.9 we summarize the complete suite of ITU-T recommendations detailing the most important *physical layer* DTE/DCE interfaces. These are those defined by ITU-T recommendations X.21 and X.21 bis as shown.



(a) Parallel transmission



(b) Serial transmission

Figure 9.8 'Parallel' and 'serial' transmission

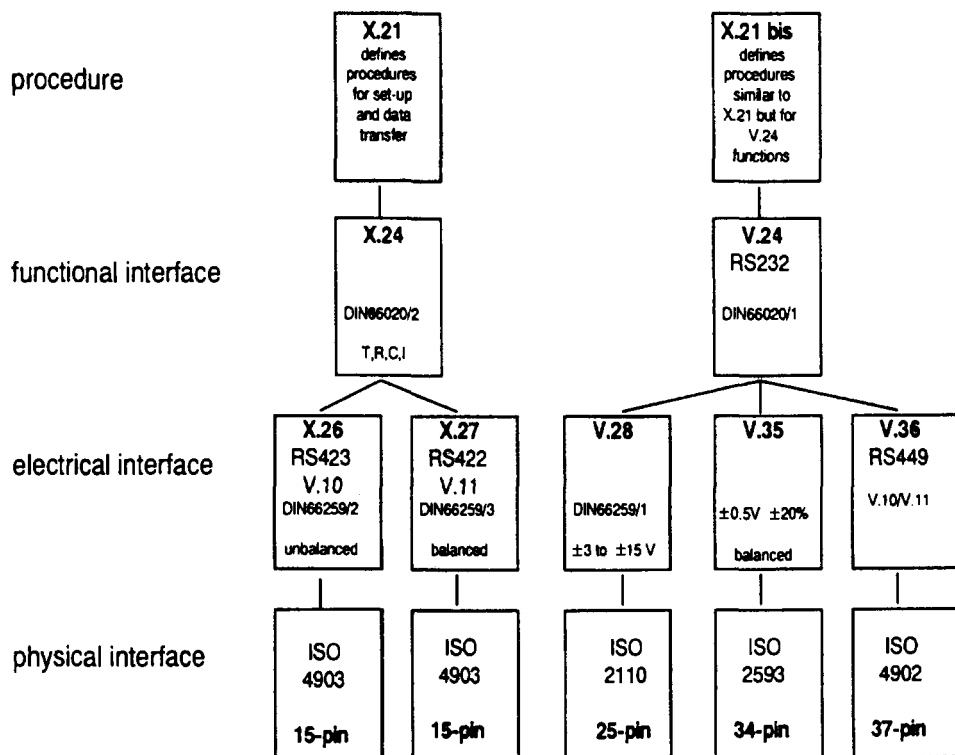


Figure 9.9 Common interfaces for DTE/DCE attachment

The ITU-T recommendations listed in Figure 9.9 define for each of the interfaces: the mechanical nature of the plugs and sockets (i.e. their physical structure), the electrical currents and voltages to be used, the functional purpose or meaning of the various circuits which appear on the various pins of the plug/socket connection, and the valid pin combinations and order (or procedure) in which the various functions may be used.

There are two broad groups of *physical layer* interfaces. The first broad group are the group sometimes loosely referred to as V.24 (RS-232) or X.21 bis which were developed in conjunction with modems and are relatively low speed devices (up to about 28 kbit/s and about 15 m cable length) allowing computer data to be transferred to modems for carriage across analogue telephone lines. The second group is the X.21 suite (sometimes only the subsets X.24 or V.11 are used). This second group was developed later for higher speed lines, which became possible in the advent of digital telephone leaselines (64 kbit/s and above and up to 100 m cable length).

In Europe, the standard interfaces offered by modems and digital leaseline DCEs (variously called *line terminating unit (LTU)*, *network terminating unit (NTU)*, *channel service unit (CSU)*, *DSU (data service unit)*) are V.24 for analogue leaselines and X.21 (or the subsets X.24 or V.11) for digital lines. In North America RS-232 (V.24) is widely used for analogue lines, but V.35 and V.36 predominate over X.21 for digital lines.

You might wonder how you connect a V.35 DTE in the USA to an X.21 DTE in Europe. The answer is quite straightforward. You order an international digital leaseline with V.35 DCE interface in the USA and X.21 interface in Europe. The fact that the two DTE-DCE connections have different forms is immaterial.

## 9.7 SYNCHRONIZATION

An important component of any data communications system is the clocking device. The data consists of a square, ‘tooth-like’ signal, continuously changing between state ‘0’ and state ‘1’, as we recall in Figure 9.10.

The successful transmission of data depends not only on the accurate coding of the transmitted signal (Chapter 4), but also on the ability of the receiving device to decode the signal correctly. This calls for accurate knowledge (or *synchronization*) of where each bit and each message begins and ends.

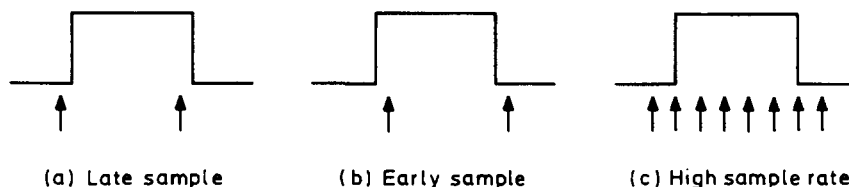
The receiver usually samples the communication line at a rate much faster than that of the incoming data, thus ensuring a rapid response to any change in line signal state, as Figure 9.11 shows. Theoretically it is only necessary to sample the incoming data at a rate equal to the nominal bit rate, but this runs a risk of data corruption. If we choose to sample near the beginning or end of each bit we might lose or duplicate data as Figures 9.11(a) and (b) show, simply as the result of a slight fluctuation in the time duration of individual bits. Much faster sampling ensures rapid detection of the start of each '0' to '1' or '1' to '0' transition, as Figure 9.11(c) shows. The signal transitions are interpreted as bits.

Variations in the time duration of individual bits come about because signals are liable to encounter time shifts during transmission, which may or may not be the same for all bits within the total message. These variations are usually random and they combine to create an effect known as *jitter*, which can lead to incorrect decoding of incoming signals when the receiver sampling rate is too low, as Figure 9.12 shows.

*Jitter*, together with a slight difference in the timing of the incoming signal and the receiver's low sampling rate have created an error in the received signal, inserting an



**Figure 9.10** Typical data signal



**Figure 9.11** Effect of sample rate

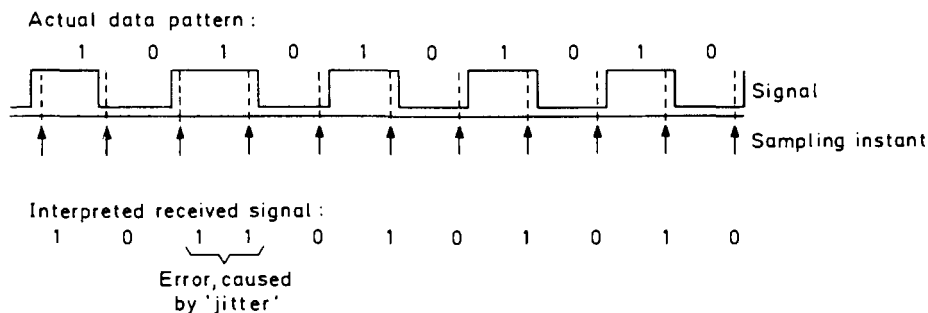


Figure 9.12 The effect of jitter

extra '1'. Thus the signal at the top is intended to represent only a ten bit pattern, but because the bit durations are not exactly correct, (due to jitter), the pattern has been misinterpreted. An extra bit has been inserted by the receiving device.

To prevent an accumulation of errors over a period of time, we use a sampling rate higher than the nominal data transfer rate as already discussed; we also need to carry out a periodic *synchronization* of the transmitting and receiving end equipments.

The purpose of *synchronization* is to remove all short, medium and long term time effects. In the very short term, synchronization between transmitter and receiver takes place at a bit level, by *bit synchronization*, which keeps the transmitting and receiving clocks in step, so that bits start and stop at the expected moments. In the medium term it is also necessary to ensure *character* or *word synchronization*, which prevents any confusion between the last few bits of one character and the first few bits of the next. If we interpret the bits wrongly, we end up with the wrong characters. Finally there is *frame synchronization*, which ensures data reliability (or *integrity*) over even longer time periods.

## 9.8 BIT SYNCHRONIZATION

Figure 9.13 shows two different pulse transmission schemes used for bit synchronization. The first, Figure 9.13(a), is a *non-return to zero (NRZ)* code which looks like a

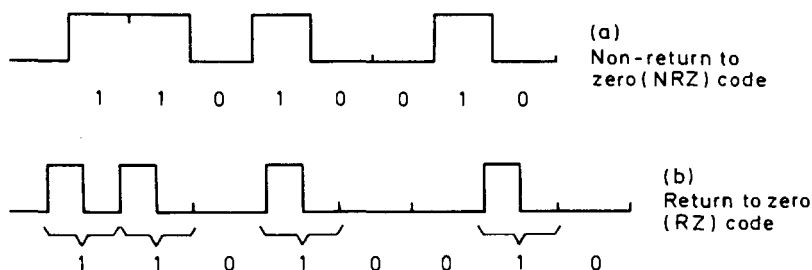


Figure 9.13 Methods of bit synchronization

string of '1' and '0' pulses, sent in the manner with which we are already familiar. In contrast Figure 9.13(b) is a *return-to-zero (RZ)* code, in which '1's are represented by a short pulse which returns to zero at the midpoint. RZ code therefore provides extra 0 to 1 and 1 to 0 transitions, most noticeably within a string of consecutive '1's. By so doing, it better maintains synchronization of clock speeds (or bit synchronization) between the transmitter and the receiver. A number of alternative techniques also exist. The technique used in a given instance will depend on the design of the equipment and the accuracy of bit synchronization required. More often than not it is the interface standard that dictates which type will be used. The synchronization code used widely on IBM *SDLC* networks is called *NRZI (non-return to zero inverted)*. These and a number of other line codes are illustrated in Chapter 5.

## 9.9 CHARACTER SYNCHRONIZATION: SYNCHRONOUS AND ASYNCHRONOUS DATA TRANSFER

Data conveyance over a transmission link can be either *synchronous* (in which individual data characters (or at least a predetermined number of bits) are transmitted at a regular periodic rate) or *asynchronous* mode (in which the spacing between the characters or parts of a message need not be regular). In *asynchronous data transfer* each data character (represented say by an eight-bit 'byte') is preceded by a few additional bits, which are sent to mark (or *delineate*) the start of the eight-bit string to the receiving end. This assures that *character synchronization* of the transmitting and receiving devices is maintained.

Between characters on an *asynchronous* transmission system the line is left in a *quiescent state*, and the system is programmed to send a series of '1's during this period to 'exercise' the line and not to generate spurious start bits (value '0'). When a character (consisting of eight bits) is ready to be sent, the transmitter precedes the eight bit pattern with an extra *start bit* (value '0'), then it sends the eight bits, and finally it suffixes the pattern with two 'stop bits', both set to '1'. The total pattern appears as in Figure 9.14, where the user's eight bit pattern 11010010 is being sent. (Note: usually nowadays, only one stop bit is used. This reduces the overall number of bits which need to be sent to line to convey the same information by 9%.)

In *asynchronous* transmission, the line is not usually in constant use. The idle period between character patterns (the quiescent period) is filled by a string of 1s. The receiver can recognize the start of each character when sent by the presence of the start bit transition (from state '1' to state '0'). The following eight bits then represent the character pattern.

The advantage of *asynchronous transmission* lies in its simplicity. The start and stop bits sent between characters help to maintain synchronization without requiring very accurate clocking, so that devices can be quite simple and cheap. Asynchronous transmission is quite widely used between computer terminals and the computers themselves because of the simplicity of terminal design and its consequent cheapness. Given that human operators type at indeterminate speeds and sometimes leave long pauses between characters, it is ideally suited to this use. The disadvantage of asynchronous transmission lies in its relatively inefficient use of the available bit speed.

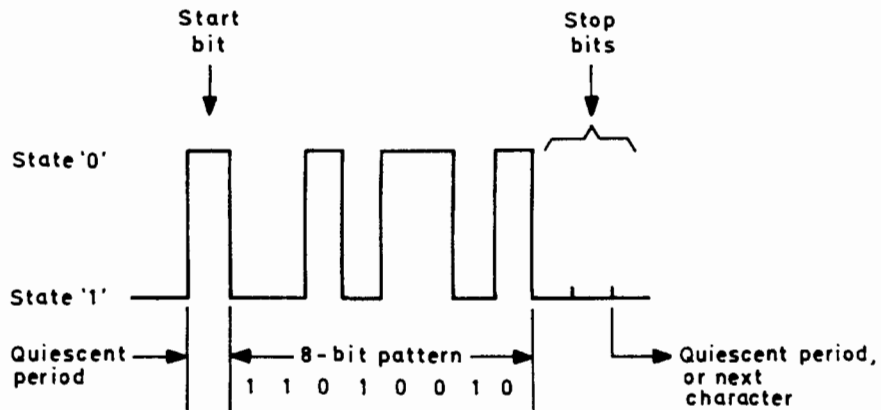


Figure 9.14 Asynchronous data transfer

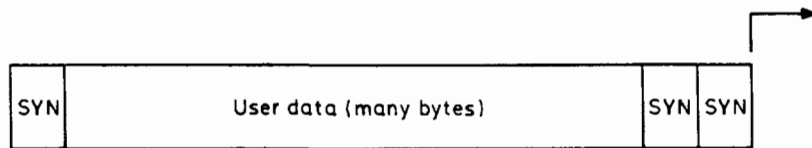


Figure 9.15 Synchronous data transfer

As we can see from Figure 9.14, out of eleven bits sent along the line, only eight represent useful information.

In *synchronous data transfer*, the data are *clocked* at a steady rate. A highly accurate clock is used at both ends, and a separate circuit may be used to transmit the timing between the two. Provided that all the data bit patterns are of an equal length, the start of each is known to follow immediately the previous character. The advantage of synchronous transmission is that much greater line efficiency is achieved (because no start bits need be sent), but its more complex arrangements do increase the cost as compared with asynchronous transmission equipment. Byte synchronization is established at the very beginning of the transmission or after a disturbance or line break using a special synchronization (SYN) pattern, and only minor adjustments are needed thereafter. Usually an entire user's data field is sent between the synchronization (SYN) patterns, as Figure 9.15 shows.

The SYN byte shown in Figure 9.15 is a particular bit pattern, used to distinguish it from other user data.

## 9.10 HANDSHAKING

We cannot leave the subject of modems without at least a brief word about the *Hayes command set* (nowadays also called the *AT command set*), for many readers will at some time find themselves faced with the question of whether a modem is 'Hayes compatible'. By this, we ask whether it uses the *Hayes command set*, the procedure by

which the link is set up and the data transfer is controlled; if you like, the *handshake* and etiquette of conversation. It allows, for example, a PC to instruct a modem to dial a given telephone number and confirm when the connection is ready. The Hayes command set has become the *de facto* standard for personal computer communication over telephone lines. An alternative scheme is today offered by ITU-T recommendation V.25 bis.

## 9.11 PROTOCOLS FOR TRANSFER OF DATA

Unfortunately the X.21, V.24/V.28 (RS-232) and RS449 standards and their inter-DCE equivalents (i.e. the line transmission standards used by modems or digital links such as V.22, V.32 or V.34) are still not enough in themselves to ensure the controlled flow of data across a network. A number of additional layered mechanisms are necessary, to indicate the coding of the data, to control the message flow between the end devices and to provide for *error correction* of incomprehensible information. Unlike human beings engaged in conversation, machines can make no sense at all of corrupted messages, and they need hard and fast rules of procedure to be able to cope with any eventuality. When they are given a clear procedure they are able, unlike most human beings, to carry on several ‘conversations’ at once. These rules of procedure are laid out in *protocols*.

Now, sit down and make yourself comfortable . . .

Once upon a time protocols were relatively unsophisticated like the simple computer-to-terminal networks which they supported and they were contained within other computer *application* programmes. Thus the computer, besides its main processing function, would be controlling the line transmission between itself and its associated terminals and other peripheral equipment. However, as organizations grew in size and data networks became more sophisticated and far-flung, the supporting communications software and hardware developed to such an extent as to be unwieldy and unmaintainable. Many computer items (particularly different manufacturers’ equipments) were incompatible. Against this background the concept of layered protocols developed with the objective of separating out the overall telecommunications functions into a layered set of sub-functions, each *layer* performing a distinct and self-contained task but being dependent on sub-layers. Thus complex tasks would comprise several layers, while simple ones would need only a few. Each layer’s simple function would comprise simple hardware and software realization and be independent of other layer functions. In this way we could change either the functions or the realization of one functional layer with only minimal impact on the software and hardware implementations of other layers. For example, a change in the routing of a message (i.e. the topology of a network) could be carried out without affecting the functions used for correcting any corrupted data (or errors) introduced on the line between the end terminals.

Most data transfer protocols in common use today use a *stack* of layered protocols. By studying such a protocol *stack* we have a good idea of the whole range of functions that are needed for successful data transfer. We need to consider the functions of each protocol layer as laid out in the international standards organization’s (ISO’s) *open*



*systems interconnection (OSI)* model. The *OSI model* is not a set of protocols in itself, but it does carefully define the division of functional layers to which all modern protocols are expected to conform. The protocols of each of the layers are defined in individual ISO standards and ITU-T recommendations.

## 9.12 THE OPEN SYSTEMS INTERCONNECTION MODEL

The open systems interconnection model, first standardized by ISO in 1983, classifies data transfer protocols in a series of layers. It sets worldwide standards of design for all data telecommunication protocols, ensuring interworking capability of equipments provided by different manufacturers.

To understand the need for the model, let us start with an analogy, drawn from a simple exchange of ideas in the form of a dialogue between two people. The speaker has to convert his ideas into words; a translation may then be necessary into a foreign language which can be understood by the listener; the words are then converted into sound by nerve signals and appropriate muscular responses in the mouth and throat. The listener meanwhile is busy converting the sound back into the original idea. While this is going on, the speaker needs to make sure in one way or another that the listener has received the information, and has understood it. If there is a breakdown in any of these activities, there can be no certainty that the original idea has been correctly conveyed between the two parties.

Note that each function in our example is independent of every other function. It is not necessary to repeat the language translation if the receiver did not hear the message, a request (prompt) to replay a tape of the correctly translated message would be sufficient. The specialist translator could be getting on with the next job so long as the less-skilled tape operator was on hand. We thus have a layered series of functions. The idea starts at the top of the talker's *stack* of functions, and is converted by each function in the stack, until at the bottom it turns up in a soundwave form. A reverse conversion *stack*, used by the listener, re-converts the soundwaves back into the idea. Figure 9.16 shows our example.

Each function in the protocol *stack* of the speaker has an exactly corresponding, or so-called *peer* function in the protocol stack of the listener. The functions at the same layer in the two stacks correspond to such an extent that if we could conduct a direct *peer-to-peer* interaction then we would actually be unaware of how the functions of the lower layers protocols had been undertaken. Let us, for example, replace layers 1 and 2 by using a telex machine instead. The speaker still needs to think up the idea, correct the grammar and see to the language translation, but now instead of being aimed at mouth muscles and soundwaves, finger muscles and telex equipment do the rest, provided that the listener also has a telex machine. We cannot, however, simply replace only the speaker's layer 1 function (the mouth), if we do not carry out simultaneous *peer protocol* changes on the listener's side because an ear cannot pick up a telex message. The principle of layered protocols is that so long as the layers interact in a *peer-to-peer* manner, and so long as the interface between the function of one layer and its immediate higher and lower layers is unaffected, then it is unimportant how the function of that individual layer is carried out. This is the principle of the *open systems interconnection (OSI)*

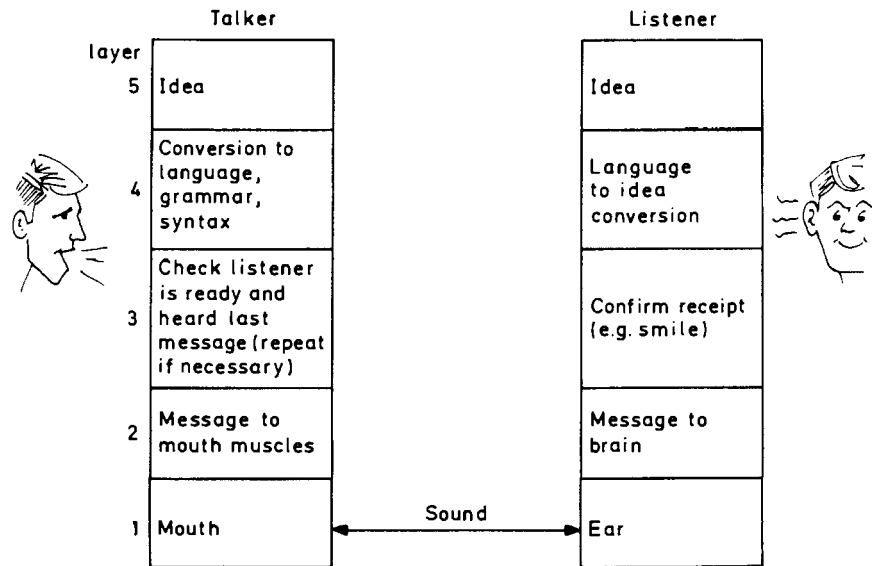


Figure 9.16 A layered protocol model for simple conversation

model. The OSI model sub-divides the function of data communication into a number of layered and peer-to-peer sub-functions, as shown in Figure 9.17.

In all, seven layers are defined. Respectively, from layer seven to layer one these are called: the *application layer*, the *presentation layer*, the *session layer*, the *transport layer*, the *network layer*, the *data link layer* and the *physical layer*.

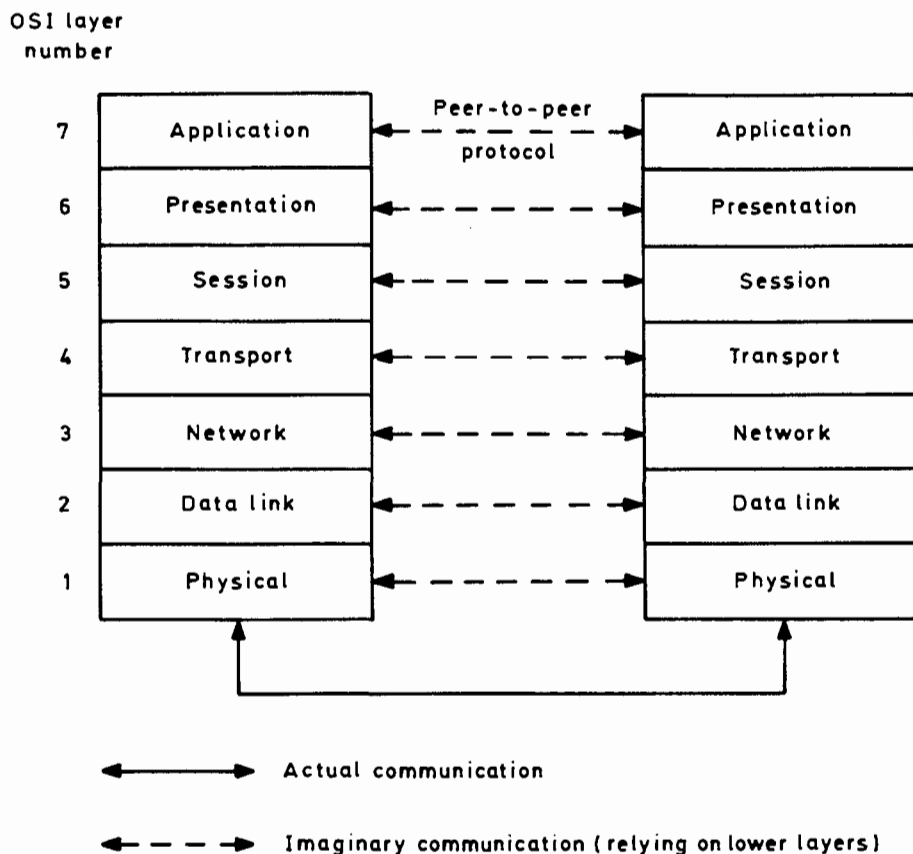
Each layer of the OSI model relies on the *service* of the layer beneath it. Thus the transport layer (layer 4) relies on the *network service* which is provided by the *stack* of layers 1–3 beneath it. Similarly the transport layer provides a *transport service* to the session layer, and so on.

The functions of the individual layers of the OSI model are defined more fully in ISO standards (ISO 7498), and in ITU-T's X.200 series of recommendations; in a nutshell they are as follows.

### 9.12.1 Application Layer (Layer 7)

This is the layer that provides communication services to suit all types of data transfer between *cooperating* computers. It comprises a number of service elements (SEs), each suited for a particular purpose or application. They may be combined in various permutations to meet the needs of more complex applications.

A wide range of *application layer protocols* will be defined over time, to accommodate all sorts of different computer equipment types, activities, controls and other applications. These will be defined in a modular fashion, the simplest common functions being termed *application service elements (ASEs)*, which are sometimes grouped in specific functional combinations to form *application entities (AEs)*. These are the functions coded as protocols.



**Figure 9.17** The Open Systems Interconnection (OSI) Model. (Courtesy of CCITT – derived from Figures 12 and 13/X200)

Although the *application layer protocol* will differ according to the particular situation, AE or ASE, the ISO standards at least provide a standardized notation (shorthand language) in which the *application layer protocols* may be defined and written to enable fairly rapid interpretation by experienced technicians and equipment designers. This notation is the *abstract syntax notation 1 (ASN.1)*. It is laid out in ITU-T recommendations X.208 (old version) and X.680 (new version). Examples of *application layer protocols* covered by this book are the *transaction capabilities of signalling system 7* (see Chapter 12) and the *message handling system* (Chapter 23).

### 9.12.2 Presentation Layer (Layer 6)

As we found in Chapter 4, data may be coded in various forms such as binary, ASCII, ITU-T IA5, EBCDIC, fax encoding, *multimedia* signal format, etc. The task of the presentation layer is to negotiate a mutually agreeable technique for data encoding and

punctuation (called the *data syntax*) and to provide for any conversion that may be necessary between different code or data layout formats, so that the application layer receives the type it recognizes.

### 9.12.3 Session Layer (Layer 5)

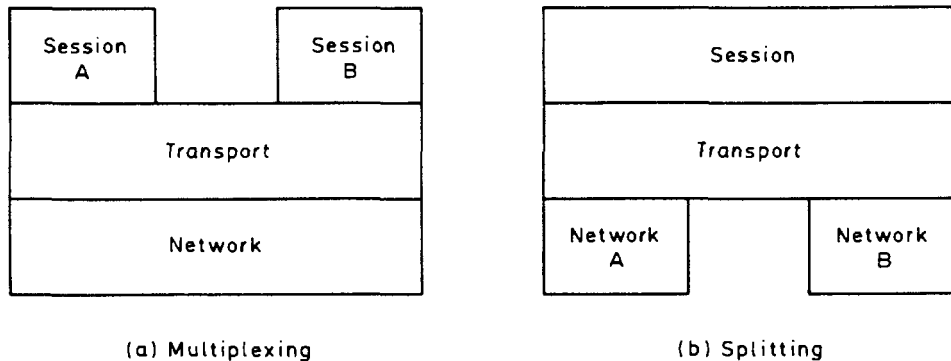
When established for a *session* of communication, the two devices at each end of the communication medium must conduct their 'conversation' in an orderly manner. They must listen when spoken to, repeat as necessary, and answer questions properly. The session protocol thus includes commands such as *start*, *suspend*, *resume* and *finish*. The session protocol is rather like a tennis umpire. He cannot always tell how hard the ball has been hit, or whether there is any spin on it, but he knows who needs to hit the ball next and whose turn it is to serve, and he can advise on the rules when there is an error, so that the game can continue. The session protocol negotiates for an appropriate type of session (e.g. full duplex or half duplex data block lengths) to meet the communication need, and then it manages the session.

### 9.12.4 Transport Layer (Layer 4)

The transport service provides for the end-to-end data relaying service needed for a communication session. The transport layer itself establishes a transport connection between the two end-user devices by making the network connection that best matches the session requirements in terms of quality of service, data unit size, flow control, and error correction needs. If more than one network is available (e.g. packet network, circuit switched data, telephone network, etc.), it chooses between them. In addition, where a number of networks must be traversed in succession (e.g. a connection travels from *LAN (local area network)* via a *wide area network* to a second *LAN*), it can arrange this too.

The transport layer provides the network addresses needed by the network layer for correct delivery of the message. The *mapping* function provided by the transport layer, in converting *transport addresses* (provided by the session layer to identify the destination) into network-recognizable addresses (e.g. telephone numbers) shows how independent the separate layers can be: the conveyance medium could be changed and the *session*, *presentation* and *application* protocols could be quite unaware of it.

The transport protocol is also capable of some important *multiplexing* and *splitting* functions (Figure 9.18). In its multiplexing mode the *transport protocol* is capable of supporting a number of different sessions over the same connection, rather like playing two games of tennis on the same court. Humans would become confused about which ball to play, but the transport protocol makes sure that computers do nothing of the kind. Conversely, the *splitting* capability of the transport protocol allows (in theory) one *session* to be conducted over a number of parallel network communication paths, like a grand master playing umpteen simultaneous chess games. Thus two *sessions* from a mainframe computer to a PC (personal computer) in a remote branch site of a large shopping chain might be used simultaneously to control the building security system



**Figure 9.18** Protocol multiplexing and splitting

and (separately) to communicate customer sales. Different software programmes (for security and for sales) in both the mainframe computer and in the PC could thus share the same telecommunications line without confusion.

The transport protocol also caters for the end-to-end conveyance, segmenting or concatenating (stringing together) the data as the network requires.

### 9.12.5 Network Layer (Layer 3)

This layer sets up and manages an end-to-end connection across a single real network, determining which permutation of individual links to be used and ensuring the correct transfer of information across the single network (e.g. *LAN* or *wide area network*). A layer 3 protocol commonly used for data communication is the X.25 packet interface standard, discussed in Chapter 18. The 1984 and later versions of X.25 conform with the *OSI model*, providing a packet-mode version of the *OSI network service*.

### 9.12.6 Datalink Layer (Layer 2)

The datalink layer operates on an individual link of a connection, managing the transmission of the data across a particular physical connection so that the individual bits are conveyed over that link without error. ISO's standard datalink protocol, specified in ISO 3309, is called *high level data link control (HDLC)*. Its functions are to

- synchronize the transmitter and receiver (i.e. the link end devices)
- control the flow of data bits
- detect and correct errors of data caused in transmission
- enable multiplexing of several *logical channels* over the same physical connection.

Typical commands used in datalink control protocols are thus *ACK* (*acknowledge*), *EOT* (*end of transmission*), etc.

### 9.12.7 Physical Layer (Layer 1)

The physical layer protocol conveys the data over the medium; it is a combination of the hardware and software which converts the data bits needed by the datalink layer into the electrical pulses, modem tones, optical signals or other means which are going to transmit the data. The physical layer ensures that the data are conveyed over the link and presented at both ends to the *datalink layer* in a standard form. Earlier in this chapter we looked at V.24/V.28 and RS-232C interfaces which form part of ITU-T's X21-bis recommendation, the layer 1 standard interface. As we also saw, alternative layer-1 interfaces for DTE/DCE connection are X.21, V.35 and V.36, and for DCE/DCE connection V.22, V.32 and V.34, among a large list of others.

## 9.13 DATA MESSAGE FORMAT

The key to any layered protocol lies in its use of headers. Each protocol layer adds a header containing information for its own use. The overall message is thus longer than that received from the higher layer protocol. The headers carry the information that the protocol needs (the *protocol control information (PCI)*) to do its work. They are stripped from the received message by the corresponding protocol handler in the distant end device, and the remainder of the message is passed up to the next higher protocol layer. Messages are passed normally in a synchronous manner, and Figure 9.19 illustrates a typical data message format showing some of the fields corresponding to the network service functions of layers 1–3. The exact format depends on the particular protocols in use.

The flag is a layer 2 function. It indicates the start of each message and includes the synchronization byte (SYN) described earlier in the chapter.

The control bits control the flow of data, passing *receipt acknowledgement* and other information for data link control (layer 2).

The header contains the *address* of the destination. The *address* varies in nature according to the various layers. The term can be confusing at first, as it does not always directly identify the end *network address* (e.g. telephone number) of the destination connection. Rather, the *address* should be regarded as being only an identifying tag for the next higher layer. Thus at level 2 (*logical link layer*), the *layer 2 address* is termed the *logical channel number* (in X.25), *data link channel identifier* (in frame relay) or *virtual path identifier/virtual channel identifier* (in ATM). This *address* indicates to which of

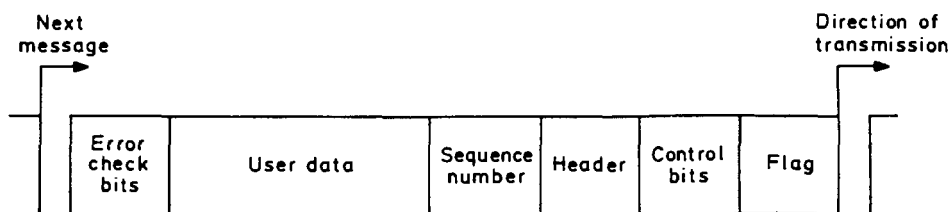
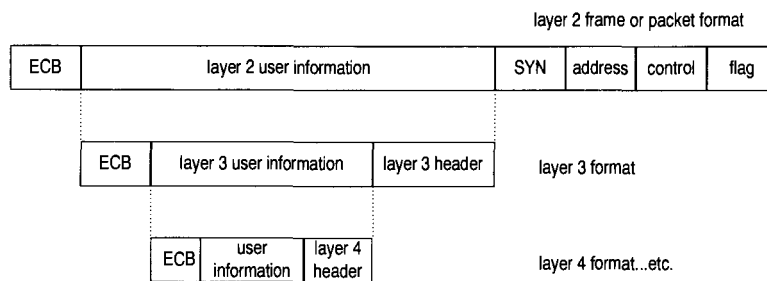


Figure 9.19 Typical data message format





**Figure 9.21** The frame-, packet- or cell-format of a structured protocol

**Table 9.2** Cyclic redundancy check codes

CRC-n type	Field multiplier	Generator polynomial
CRC-1		(simple parity check)
CRC-3	$x^3$	$x^3 + x + 1$
CRC-4	$x^4$	$x^4 + x + 1$
CRC-5	$x^5$	$x^5 + x^4 + x^2 + 1$
CRC-6	$x^6$	$x^6 + x + 1$
CRC-7	$x^7$	$x^7 + x^3 + 1$
CRC-8		
CRC-10	$x^{10}$	$x^{10} + x^9 + x^5 + x^4 + x + 1$
CRC-16	$x^{16} + x^{15} + x^{14} + x^{13} + x^{12} + x^{11} + x^{10} + x^9 + x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$	$x^{16} + x^{12} + x^5 + 1$

request the retransmission of the message, identifying it by the appropriate *sequence number*. Other more powerful means of error detection and corection are provided by *cyclic redundancy checks (CRCs)*, *Hamming Codes*, *Reed–Solomon* and *Viterbi codes*. These methods use extra bits and complex mathematical algorithms to dramatically reduce the probability of errors surviving detection and correction. They therefore greatly increase the dependability of transmission systems.

Cyclic redundancy check (CRC) code: a CRC-n is an n-bit cyclic redundancy check code. The value of the code is set by first multiplying the field by a given *multiplier*, then dividing (in binary, or *modulo 2*) the result of the multiplication by a *generator polynomial*. The *remainder* is used to set the value of the CRC field. Common CRC codes used are shown in Table 9.2.

## 9.14 IMPLEMENTATION OF LAYERED PROTOCOL NETWORKS

In reality, most of the protocol layers of OSI exist only in software (as part of a computer programme) and cannot be identified as physical items. Nonetheless, the different protocol layers need not all be implemented within the same computer programme or indeed be carried out by the same piece of hardware.



We have already seen in the first part of the chapter how it is sometimes convenient that data terminating equipment (DTE) does not have to conduct line communication functions. Instead these activities are usually delegated to a piece of specialized equipment called a data circuit terminating equipment (DCE). In essence this equipment converts between two different physical layer protocols, on the DTE side RS-232 or an equivalent interface is used, whereas on the line side, digital line codes, clocking functions, etc. have to be carried out. Figure 9.22 now illustrates the application of the OSI model to this example. Note in the figure how the various protocols are all communicating peer-to-peer. At layer 1 the DCE is converting between different protocols on DTE/DCE and DCE/DCE interfaces, but at all the other layers the peer-to-peer interaction is end-to-end.

A common approach used by many computer manufacturers, for example, is to split the layer 1–3 functions away from the *host* or *application processor* and to implement them instead on a specialized *front end processor (FEP)* or *communications controller*. In this way the host machine is relieved of the burden of communications link supervision, link synchronization, error correction and conversion of data from the internal computer bus format into the standard serial transmission line format. A typical configuration is shown in Figure 9.23. The *host* or *mainframe* computer of Figure 9.23 is that on which the main computer software programme runs. This programme communicates as necessary with other computer devices (e.g. terminals, printers, databases, other computers) for the input or output of information. The communication relies on an application layer protocol, resident in the host, which is invoked every time a command such as *print*, *read*, *input* or *output* appears in the computer program. The application protocol cooperates in a peer-to-peer fashion with a corresponding function in a distant terminal. Layer 4–6 protocols also exist in both the host computer and the distant terminal controller, but layer 1–3 functions (the real spade work of running the telecommunication network) are carried out at the host's end by a collocated *front end*

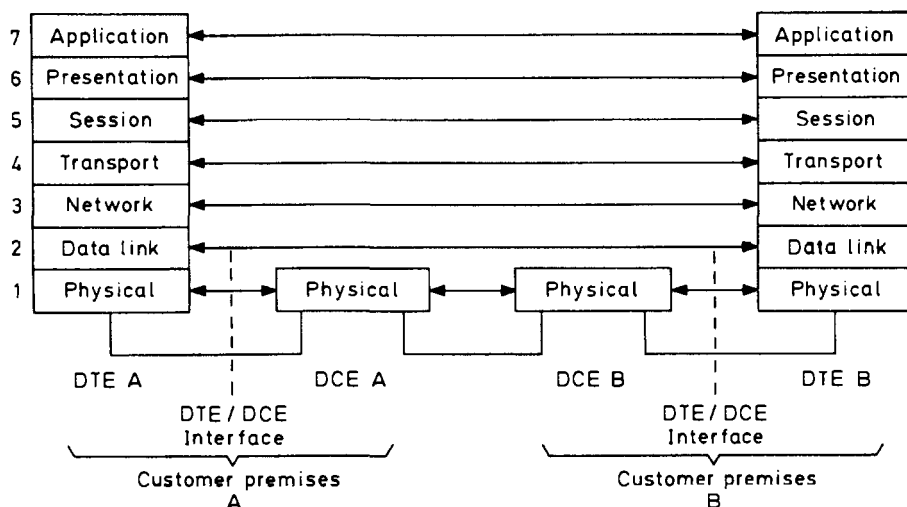


Figure 9.22 OSI model applied to DTE/DCE

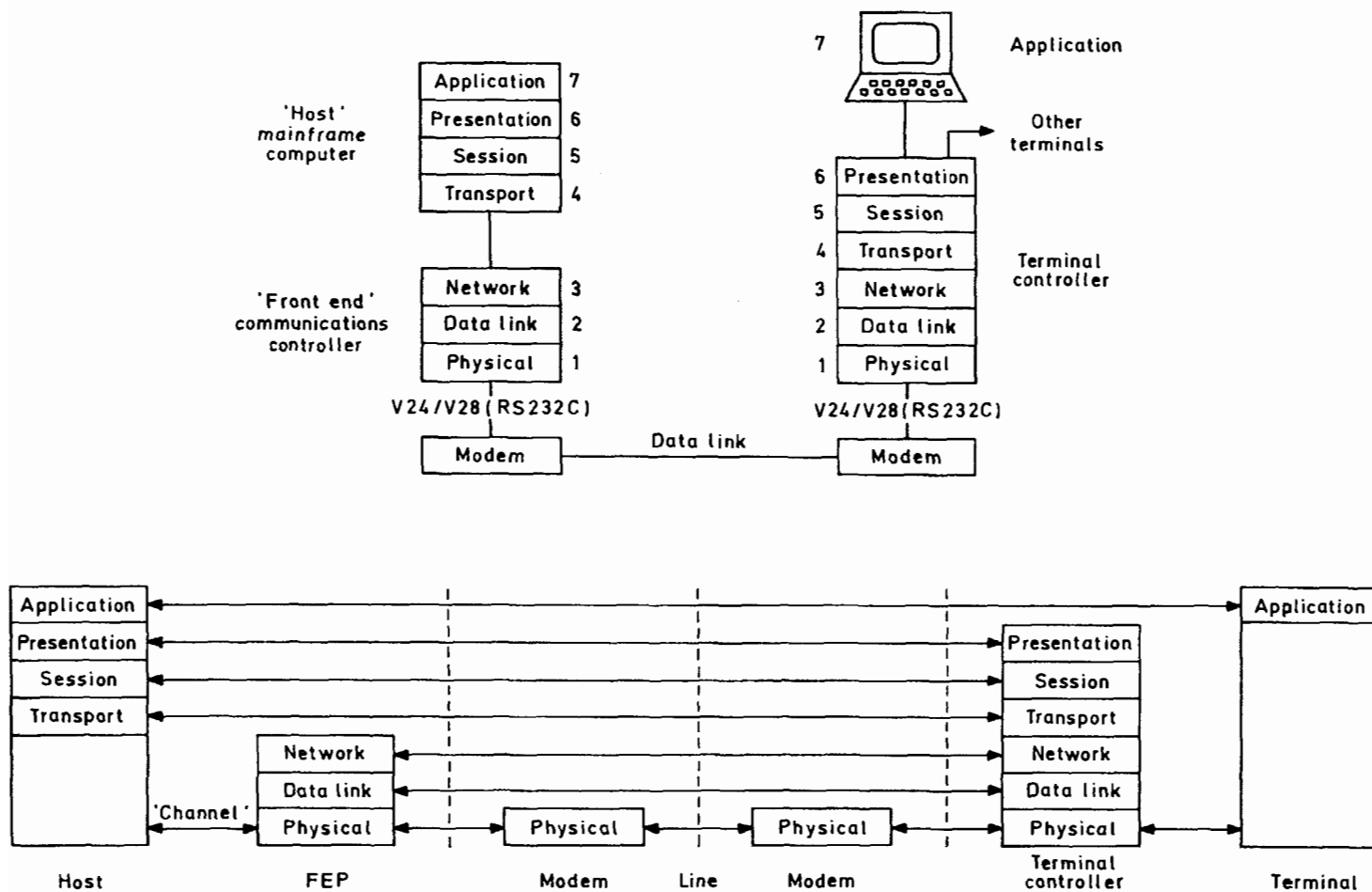


Figure 9.23 OSI model applied to host/terminal connection

communications controller. The two inter-communicate through the normal internal computer wiring (in IBM company terminology, the *channel*), essentially providing the functions of OSI layers 1–3 for this part of the connection.

As not all the protocols used between a *host* and its associated terminals are OSI standard, the various equipments are not fully interchangeable between one manufacturer's equipment and another. In our diagram, however, the protocols are shown to be fully OSI conformant at all protocol layers used on the line, between FEP and modem, and between modem and terminal controller. Thus the host and FEP together could be used with another manufacturer's OSI-conformant terminal controller/terminal combination. Likewise, the terminal controller/terminal combination shown could be used with other host/FEPs.

The actual protocols used in any given case are not exclusively defined by the OSI model, and a number of standardized sets of protocols are now appearing which conform with it. Examples include *MAP/TOP* (*manufacturing automation protocol/technical and office protocol*) which defines a set of protocols at all seven layers, which together are suitable for manufacturing and office interworking applications; other examples include protocols for the individual layers. ITU-T X.25 (1984 version or later) supports the *OSI network service*, ISO 8073 is a *transport protocol*, ISO 8327 is a *session protocol*, ISO 8823 is a *presentation protocol* and ISO 8571 *FTAM* (*file transfer, access and management*) is an example of an application layer protocol, intended for the purpose of direct data transfer between mainframe computers.

Because a large number of different application protocols are likely to be conceived, addressing the needs of the many and varied applications possible, ISO has standardized a formal language for the application layer. It is called *ASN.1* or *abstract syntax notation 1*. It defines protocol data structures and encoding rules for application layer protocols and is set out in standards ISO 8824 and ISO 8825 and in ITU-T recommendations X.208 and X.680 (new version).

## 9.15 THE USE OF NULL LAYERS

An important point to note about the OSI layered model is that it gives immense scope, allowing very sophisticated networks to be developed. Not all of the very sophisticated functions are required in every case, and their implementation may increase the cost and burden of administration. The model is designed to permit null protocols to be used at some of the layers. For example, on a network using similar end devices the syntax conversion capabilities of the presentation layer are unnecessary and can be dispensed with. Likewise, on a point-to-point data connection layer 1 and 2 protocols may be sufficient without needing much effort at layers 3 and 4 (the *network* and *transport* layers).

## 9.16 OTHER LAYERED PROTOCOLS

The OSI model defines standard peer protocol functions that ultimately will stimulate the development of wide ranging protocols allowing any two computer devices running any types of applications to intercommunicate without difficulty. However, because

some of today's layered protocol sets predate the OSI model, not all conform exactly to it. For example, the *systems network architecture (SNA)*, announced by the IBM computer company as the future direction for its product range in 1973, defines a stack of protocols that allow computers to communicate. The principles and layers of SNA are similar to those of the OSI model, although some significant differences exist that make the two incompatible. The equivalent of *HDLC* (OSI layer-2), for example, is called *SDLC (synchronous data link control)*; we shall discuss SNA in Chapter 18. Another proprietary layered protocol is the *DECNET* of the Digital Equipment Corporation.

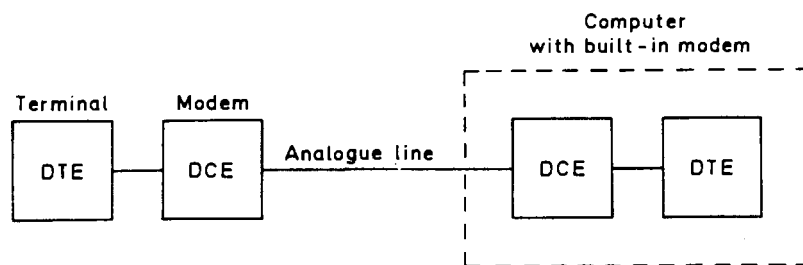
## 9.17 DATA NETWORK TYPES

Next we review the different types of real networks that are used for conveyance of data, starting with simple point-to-point technology, and moving on to more complex networks and *local area networks (LANs)*. We return to detailed technical discussion of the various network types in Part 3.

## 9.18 POINT-TO-POINT DATA NETWORKS

Point-to-point networks involving the simple interconnection of two pieces of equipment are relatively simple to establish. As we have seen, they may use either digital lines and DCEs, or analogue lines and modems. The DCE itself (either digital version or modem) may take the form of a distinct 'box' of kit, or it may be provided as part of a *line controller* or *linecard* built into the computer itself. Figure 9.24 shows a typical set-up.

Provided that the protocols at both ends of the link match, the data terminal equipments (DTEs) converse easily. In the point-to-point case most of the protocol layers (e.g. layers 3–7) described in the OSI model are redundant, and relatively simple communication software is necessary in the two DTEs. At its simplest, a point-to-point network can be run in an asynchronous, character-by-character mode. This is a common method of connecting remote terminals to a computer. The use of such a simple data



**Figure 9.24** Point-to-point data network. DTE, data terminal equipment; DCE, data circuit terminating equipment (e.g. a modem for an analogue line)

transmission technique considerably reduces the complexity and cost of hardware and software required in the (remote) computer terminals. This type of connection does not conform to the OSI ideal, as is plain from the fact that few manufacturers' computer terminals of this type can be used with other manufacturers' computers. One of the drawbacks of the OSI dream is the burden of software and hardware required on each transmitting and receiving device.

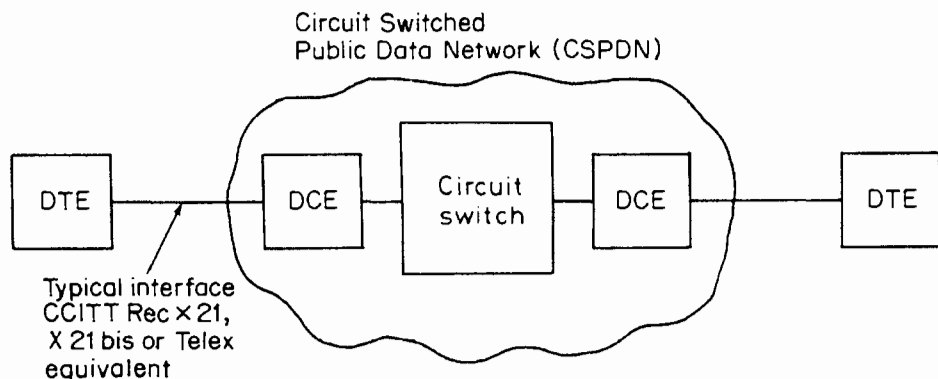
## 9.19 CIRCUIT-SWITCHED DATA NETWORKS

In our discussion of circuit-switched networks (Chapter 6) we were chiefly interested in making a connection on demand between any two customers of a telephone network. Each connection would be established for the length of their call, and would provide a clear transmission path for their conversation. At the end of the call, the connection would be cleared. Subsequent connections could then be made to other destinations as required.

The benefit of circuit-switched networks lies in the great range of destinations available on different calls, and this means that circuit-switched networks are just as useful for data conveyance, an example being the telex network. Once a telex connection is established, characters are transmitted between the two telex terminals at a 50 baud line rate, using the IA2 alphabet (see Chapter 4).

Other examples of circuit-switched data networks include AT&T's *Accunet switched-56 kbit/s network* and the German Bundespost's *Datex-L network*. Both of these are *public circuit switched digital data networks (PSDNs)* but are now practically extinct. *ISDN (integrated services digital network)*, discussed in Chapter 10, is another example of a network offering circuit-switched data capability.

Figure 9.25 shows the usual configuration of a circuit-switched network. Note how the network 'starts' at the DCE, which consequently is usually provided by the public telecommunications operator (PTO), though it may be located on customers' premises.



**Figure 9.25** Circuit switched data network. DCE, Data circuit terminating equipment; DTE, data terminal equipment

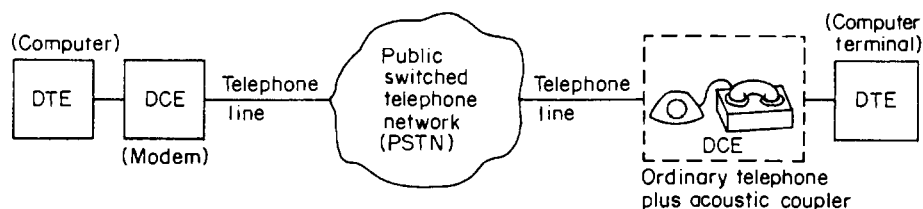


Figure 9.26 Conveying data over the telephone network

Connections in a circuit-switched data network are established in the same way as telephone connections. Once the connection is established, the data protocols in the two end terminals control the flow of data across what is then a point-to-point connection.

An alternative means of providing circuit-switched data communication to a range of destinations is by the use of ordinary telephone network, as Figure 9.26 shows.

On the left hand of Figure 9.24 a modem sits between the DTE and the telephone line. On the right-hand side the DCE is a combination of an ordinary telephone and a device called an *acoustic coupler*. The combination (which is rather outdated nowadays) performs the same function as the modem. The acoustic coupler converts the data from the DTE into the series of soundwave tones equivalent to a standard ITU-T V-series DCE/DCE interface (e.g. V.22, V.32 or V.34), and the telephone converts these tones into their electrical analogue signal equivalents. The modem simply converts the original data directly into the electrical equivalent of the acoustic tones, and is slightly more efficient and reliable. The acoustic coupler may be more suited to some (mobile) customers who may use a portable terminal with dial-up computer access from a number of different locations.

## 9.20 PACKET-SWITCHED DATA NETWORKS

A limitation of circuit-switched networks when used for data is their inability to provide variable bandwidth connections. This leads to inefficient use of resources when only a narrow bandwidth (or low bit-rate) is required. Conversely, when short bursts of high bandwidth are required, there may be data transmission delays. A more efficient means of data conveyance, packet switching, emerged in the 1970s. This has become the basis of X.25 packet networks, *frame relay* and *local area networks (LANs)*.

Packet switching is so-called because the user's overall message is broken up into a number of smaller *packets*, each of which is sent separately. We illustrated the concept in Figure 1.9 of Chapter 1. Each packet of data is labelled with the address of its intended destination, and a number of control fields are added (see Figure 9.18) before it is sent. The receiving end re-assembles the packets in their proper order, with the aid of the sequence numbers.

With packet switching, variable bandwidth becomes possible because the exchanges are allowed to operate in a *store-and-forward* fashion, each packet being routed across the network in the most efficient way available at the time. *Packet switched exchanges (PSEs)* are computers with a large data storage capacity and powerful communication software.

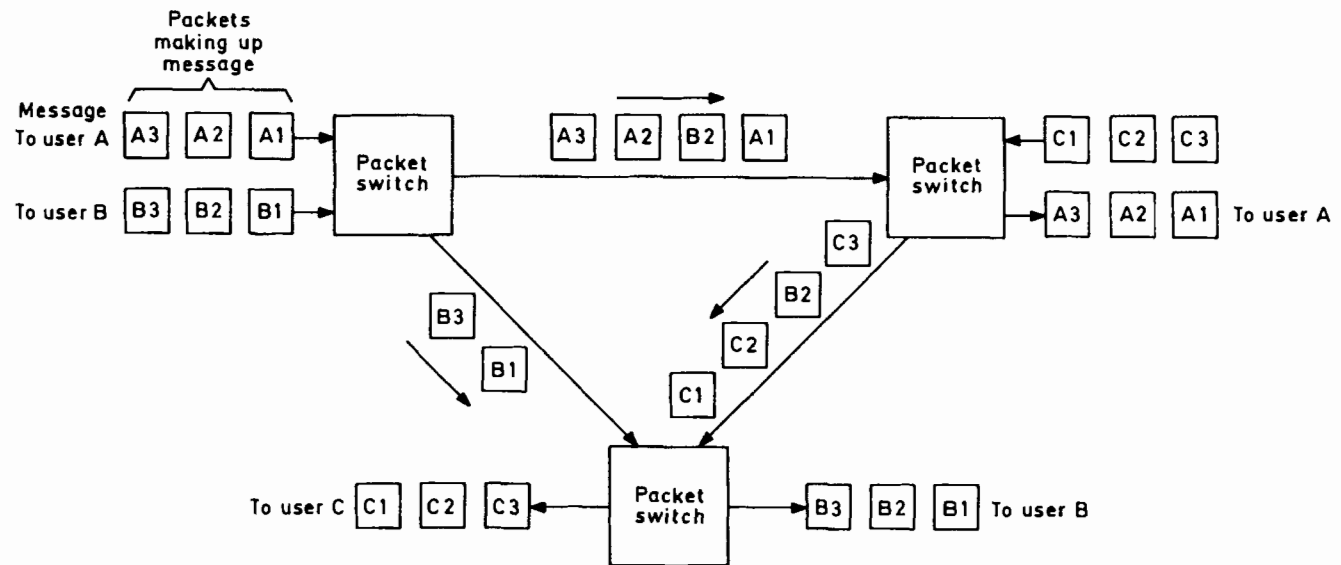


Figure 9.27 Packet switching

The principle of *packet switching* is illustrated in Figure 9.27. Note how each of the individual packets, even those with the same overall message, may take different paths through the network (*datagram* relaying). Thanks to this flexible routing, variable bandwidth can be accommodated, and transmission link utilization is optimized. More normal, however, is *path-oriented* routing, in which all the packets belonging to a given connection take the same route. This greatly reduces the problem of sorting out *mis-sequenced* packets (i.e. those out of order) on receipt, and provides for more consistent propagation delays. On failure or loss of a particular link in the path, an alternative path is usually selected without breaking the connection or losing packets.

None of the links is dedicated to the carriage of any one message. On the contrary, the individual packets of a message may be jumbled up with packets from other messages. The effect is as if a permanent channel existed between the two ends, although in reality this is not the case. The result is known as a *logical channel*, *virtual channel* or *virtual circuit*. These are the *logical channels* referred to in the layer 2 (HDLC) address field.

Packets are routed across the individual paths within the network according to the prevailing traffic conditions, the link error reliability, and the shortest path to the destination. The routes chosen are controlled by the (layer 3) software of the packet switch, together with *routing information* pre-set in the network switches by the network operator.

Packet switching gives good end-to-end reliability; with well designed switches and networks it is possible to bypass network failures (even during the progress of a 'call'). Packet switching is also efficient in its use of network links and resources, sharing them between a number of calls thereby increasing their utilization.

Packet switching forms the basis of ITU-T's X.25 and *frame relay* protocols. It has also provided the spur to more recently developed *LAN* (*local area network*) and *ATM* (*asynchronous transfer mode*) networks, as we discuss in more detail in Part 3.

## 9.21 PRACTICAL COMPUTER NETWORKS

We complete the chapter with a short discussion of the components that go to make up a real computer network, to familiarize readers with the relevant terminology used and to describe where the various *OSI layers* we discussed earlier in this chapter reside.

Five components are illustrated in Figure 9.28. The mainframe computer (sometimes also called the *host*) is on the left-hand side. This is interconnected via three other components (*front end processor* (*FEP*), *data network*, and *terminal controller*) to the user's computer terminals, which are shown on the right-hand side of the diagram.

The *front end processor* contains all the necessary communication software for setting up the connections and conversing with the terminals (i.e. protocols equivalent to OSI layers 1 to 3). A particular piece of equipment which is integrated with the *front end processor* is called the line controller. This provides the physical (layer 1) interface to the data network. It is possible to integrate the front end functions into the host itself, but the advantage of using a dedicated front end processor is that it removes a considerable burden from the *host*, allowing it to get on with the main processing job in hand.



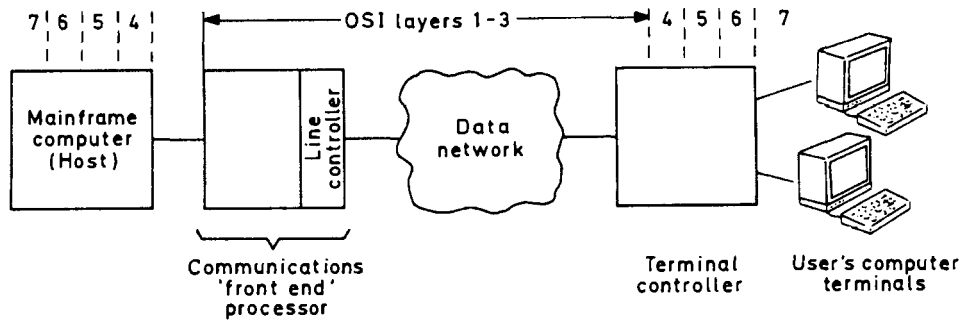


Figure 9.28 Typical computer data network

The data network might be one of a variety of types. It might be an analogue line with modems or a digital circuit, a circuit-switched, packet switched network (Chapter 18) or a LAN (Chapter 19).

Finally, the *terminal controller* interfaces a number of remote users terminals to the data network. Like the *communications front end processor*, it incorporates a line controller and some communications software to perform the functions of the data protocols. Most terminal controllers connect with terminals using the simple asynchronous, and character-oriented technique with which we are already familiar. The application (layer 7 function) thus exists to convey information to the computer about which terminal keys have been pressed, and relay back to the terminal the information to be displayed on the screen.

PART 2

---

# **MODERN TELEPHONE NETWORKS**

---



---

# *Integrated Services Digital Network (ISDN)*

---

Most public telecommunication operators throughout the world have now at least commenced the modernization of their networks by introducing digital transmission and digital switching into the public switched telephone network (PSTN). Simultaneously many operators are converting their networks to integrated services digital networks (ISDN) which will allow customers access to a variety of services while reducing the cost of provision of these services both to the administration and to the customer. So what is an ISDN? What is its value? This chapter answers these questions, and goes on to explain the technical detail of how ISDN operates, and the user benefits that can be gained.

## 10.1 THE CONCEPT OF ISDN

The internationally agreed definition of ISDN is a 'network evolved from the telephony *integrated digital network (IDN)* that provides end-to-end digital connectivity to support a wide range of services, including voice and non-voice services, to which the users have access by a limited set of standard multi-purpose customer interfaces'.

The above definition of ITU-T makes a number of points. First, that ISDN requires a digital network. Second, that this digital network is not one between exchanges, but it extends to the customers. Third, it provides not only telephony, but a variety of services. Fourth, a customer does not require a separate interface for each service provided by ISDN, but can use all services via one access point or, at worst, via 'a limited set of standard multi-purpose customer interfaces'. Figure 10.1 illustrates these principles.

*ISDN (integrated services digital network)* will remove the need for telecommunication customers to have separate physical links to telephone and low speed networks. Under ISDN, a single physical connection (or rather one of two types) is provided to a customer's premises and a range of services can be made available from it. Furthermore, if required, the services may be used simultaneously, because access is not restricted to each individual service in turn. However, as well as greatly enhancing the established services, and reducing their cost of provision, ISDN will introduce a powerful range of new ones.

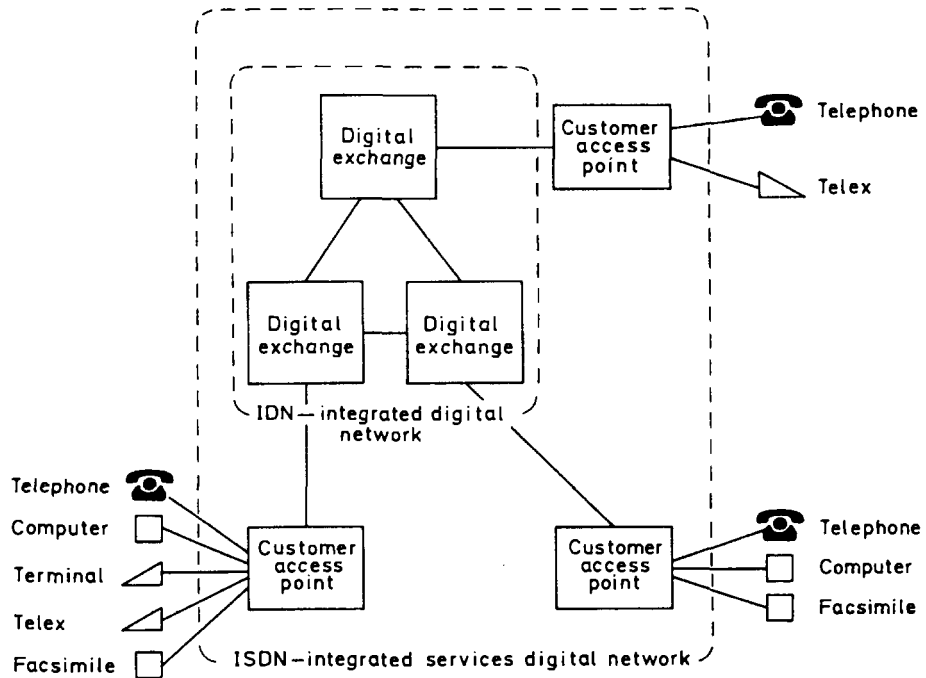


Figure 10.1 ISDN

The services of ISDN are classified into three types, *bearer services*, *supplementary services* and *teleservices*. An introduction to each of these precedes our technical and business benefits discussion of ISDN, which takes up the remainder of this chapter.

## 10.2 BEARER, SUPPLEMENTARY AND TELESERVICES

The new ISDN interfaces enable the integration of telecommunications services or *teleservices* by breaking them down into smaller component parts, called *bearer* and *supplementary* services. A *teleservice* comprises all elements necessary to support an end user's particular purpose or application. Thus examples of teleservices are *telephony*, *facsimile*, *videotelephone*, etc. The specification of a teleservice usually calls upon *bearer services* and *supplementary services* but may include additional information specific to the particular application. Thus the *Group 4* facsimile teleservice specification calls up a 64 kbit/s *bearer* capability and the ITU-T T.90 recommendation on image-encoding.

A *bearer service* is a simple information carriage service, at one of a number of available bandwidths (or rather bit-rates). At its simplest, this might be a normal switched speech (public switched telephone network) service. At its most advanced, it could be a high-speed, 64 kbit/s data service, capable of carrying a switched 'picture phone' or

slow-speed video service. The range of bandwidths available also gives scope for carriage of today's voice-band data (e.g. facsimile), telex and *packet mode* (Chapter 18) or *frame mode* (Chapter 20) services.

*Supplementary services* are so-called because, on their own, they have no purpose or value. They are always provided in conjunction with one of the bearer services, to which they add a *supplementary* value. Examples of supplementary services are *call diversion* (redirecting incoming calls to another number) and *calling-line identity* (the identification of the caller to the called party before or during answer). Another example is the *ring back when free* (correctly called *call connection to busy subscriber, CCBS*) service, which saves callers from futile repeat attempts when the called party is busy, by leaving the network to establish the call as soon as both parties are free. In all of these examples, the associated bearer service is *speech*. In some networks, some supplementary services are available even in the PSTN, but for many more ISDNs supplementary services will add a powerful new dimension.

Summarizing, examples of the three types of services included as component parts of ISDN are:

- bearer service:                      speech,                      at an appropriate bit-rate  
   3.1 kHz                      (the bearer service required for  
                        voiceband data, i.e. dial-up modems  
                        and analogue facsimile machines)
- 64 kbit/s                      high-speed data
- supplementary services:      call diversion  
   calling line identity  
   ring back when free
- teleservices:                      facsimile  
   telephony  
   videotelephone

### 10.3 ISDN INTERFACES AND END-USER APPLICATIONS

To use the ISDN, customers either have to re-equip themselves with PBXs and user terminals designed with one of the new ISDN interfaces (of which there are currently two as explained below), or use *terminal adaptors (TA)* in conjunction with existing terminals. A simple terminal might be an ISDN telephone, or an ISDN telex terminal. A more advanced device might be a personal computer with sophisticated file-transfer capabilities, or a *group 4* facsimile machine offering high-speed facsimile service of almost photocopier quality.

Worldwide standards for the new ISDN network interfaces are set by the I (ISDN), G and Q series of ITU-T recommendations. Two types of customer-to-network interface are specified, these being the *basic rate interface (BRI or BRA, basic rate access)* and *primary rate interface (PRI or PRA, primary rate access)*.

Both basic and primary rate interfaces comprise a number of B (*bearer* or *user information*) channels plus a D (*data*, but best thought of as a *signalling*) channel. The

*B* channels carry the bearer services while the *D* channel enables the customer to signal to the network how each bearer channel is to be assigned and used at time (for example, 'connect bearer (B) channel number 1 to Mr A for telephone service').

## 10.4 BASIC RATE INTERFACE (BRI)

The *BRI* (or *BRA*, *basic rate access*) comprises three channels, so-called  $2B + D$ , and is the interface that is used to connect most end-user terminal equipments, either directly to the public network or via a company *private branch exchange* (*PBX*). The bandwidth itself can be made available to the terminal in a number of different ways, using any of the *S*, *T* or *U* variants of the interface, as the diagram of Figure 10.2 shows. Each of the variants *S*, *T*, and *U* are *basic rate interfaces* but they differ in their physical realization.

The *T* interface provides for a connection point between new *ISDN terminal equipment* (*TE*) and the *ISDN network terminating equipment* (*NT1*). The *NT1*, generally provided as part of the network operator responsibility in Europe, is merely an exchange line terminating device.

The *U* interface, often known as the *wires only* interface, is a two-wire line interface, up to about 7 km in length, incoming from the public exchange, and designed to work as far as possible on existing two-wire copper pairs. Its definition came about because of regulatory requirements in the United States which debar PTOs from offering the *S/T* interface which the European PTTs intend to offer, because in the United States the *NT1* is deemed to be *customer premises equipment* (*CPE*). The *U*-interface will also be important for cellular radio based ISDNs.

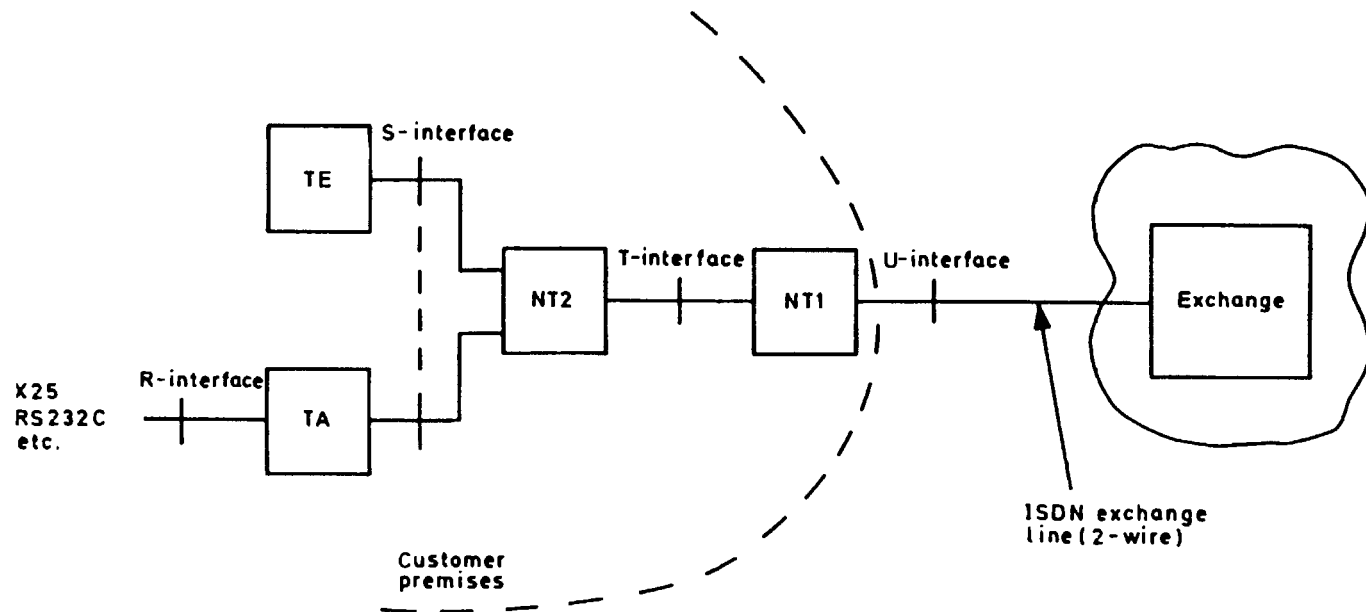
The *S* interface is physically and electrically identical to the *T* interface, but for the academics among us it is the name we give to the interface after it has passed through a call routing equipment (*NT2*). An example of an *NT2* is an *ISDN PBX*. Because the *S* and *T* interfaces are the same, some people refer to the *S/T* interface.

The final interface shown in Figure 10.2 is the *R* interface. This is the general nomenclature used to represent any type of existing telecommunication interface (such as *V.24*, *RS-232* or *X.21*, as we discussed in Chapter 9). By using a *terminal adaptor* (*TA*) we make possible the use of the *ISDN* to carry information between existing data, and telephone terminals thus easing the user's problem of deciding when to changeover from older style networks (e.g. the telephone or public packet data network).

## 10.5 THE S/T INTERFACE SPECIFICATION

The *S/T* interface is a four-wire or equivalent interface, allowing 192 kbit/s data carriage in both directions. The 192 kbit/s comprises 144 kbit/s of *user information* (data supplied by the terminal), together with an overhead of 48 kbit/s which is needed to administer and control the interface, to control the *passive bus*, for example, as described below.

The 144 kbit/s is used by the terminal as two *B* channels (both 64 kbit/s) and one *D* channel (16 kbit/s). Either or both of the *B* channels may be in use at any one time, and when used simultaneously are not constrained to be connected to the same destination.



**Figure 10.2** The basic rate interface (BRI). NT1=Network Terminating Equipment 1; NT2=Network Terminating Equipment 2; TE=Terminal Equipment (ISDN terminal); TA=Terminal Adaptor.



The passive bus is a feature of the S/T interface that enables up to eight terminal devices to be connected simultaneously to the same basic-rate interface. It is shown in Figure 10.3.

The capability for multiple-terminal connection to the BRI in this manner is crucial to the support of incoming calls because, prior to the receipt of an incoming call, it is not possible to know which device (e.g. facsimile machine, telex machine, telephone) should be connected to the line, because the type of the sending device is not known. A procedure called *terminal compatibility checking* which is part of the *D channel signalling* ensures that the call is answered only by the correct terminal device (i.e. the one of up to eight connected to the bus which is compatible with the sending device).

The *S/T interface* and the *passive bus* are intended for within-building use on customer premises. They are defined by ITU-T Recommendation I.420. They are provided directly from the NT1. On the *terminal side* of the NT1, there may be one of three wiring configurations:

- a short and full facility passive bus up to 150 m in length (the *S<sub>0</sub>-bus*)
- an extended passive bus up to 500 m in length
- a single point access line up to 1000 m in length

Protocols defined for use over the basic rate interface are defined by ITU-T Recommendation I.420 (and the European version of it is called NET3 or *Euro ISDN*). The B channels are left as clear channels for users, while a three-layer protocol stack, conforming to the first three layers of the OSI model, is tightly defined for the D channel, as Figure 10.4 shows.

The physical layer protocol (ITU-T Rec I.430) provides a *contention resolution scheme* to ensure that messages from different terminals on the passive bus do not collide (compare with the *CSMA/CD* technique for *LANs*, discussed in Chapter 19). At the second layer (I.441 or Q.921) the protocol is called *LAPD* (*link access procedure in the D-channel*). Like *HDLC* (see Chapter 9), the procedure controls the data flow over

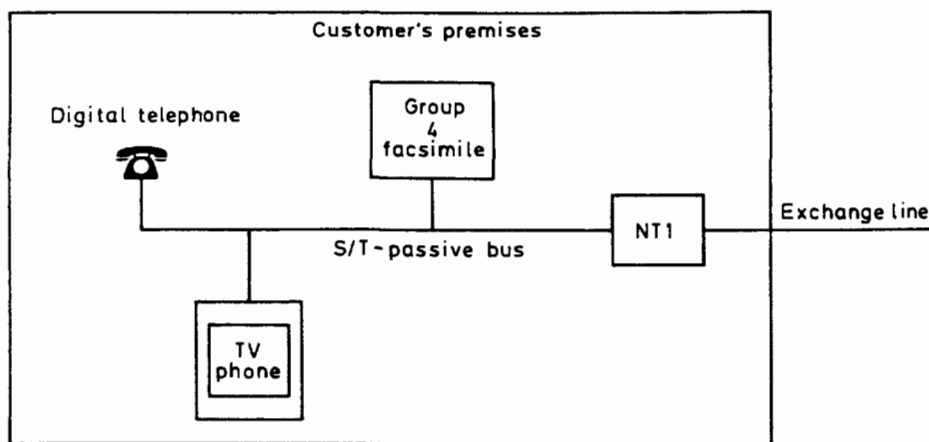


Figure 10.3 Passive bus at a customer's premises

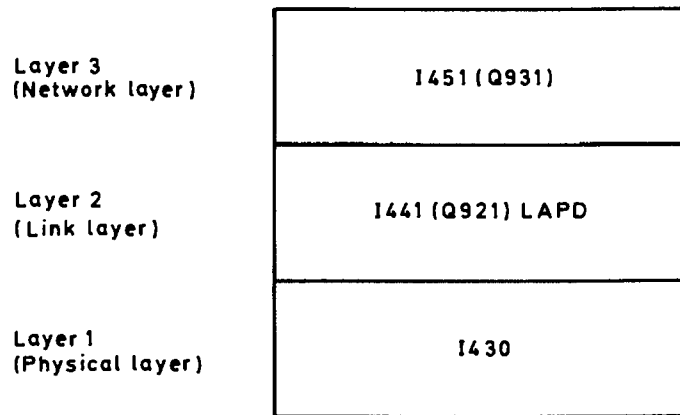


Figure 10.4 Protocol stack in the ISDN D channel

the D channel, ensuring that the data buffers are not overfilled and that the synchronization of data is maintained. It includes some degree of data error checking. In particular, LAPD provides for a number of connections to be simultaneously maintained with the multiple terminals on the passive bus. Finally, the network layer protocol (layer 3 protocol called *digital subscriber signalling 1* or *DSS1*) is defined by ITU-T Recommendation I.451 (and duplicated in recommendation Q.931). This protocol allows the transfer of dial-up signalling information and for the set-up, control and clearing of B channel connections. It allows the user's terminal to negotiate with the network set up of an appropriate terminal device at the destination end. It is this protocol that includes the *terminal compatibility* checking procedure.

## 10.6 USE OF THE BASIC RATE INTERFACE

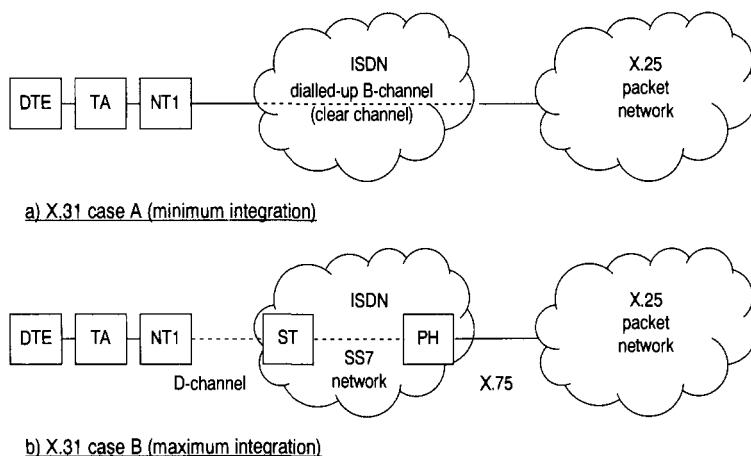
Given the small number of channels, the BRI is well-suited for digital telephone customers, computer networks, small-business users and sophisticated residential-customers (e.g. users of the *Internet*, or a businessman accessing his company computers from home). *Circuit-switched calls* at any of the three bit rates corresponding to speech, voice-band data (e.g. facsimile) or 64 kbit/s data (e.g. inter-computer links) are set up over one of the B channels, using the D channel to signal the request and managing the connection. The ISDN exchange responds to such requests by extending the relevant B channel using a circuit-switched inter-exchange connection through to the destination, as indicated by the dialled ISDN number. The ISDN number is similar to the PSTN number; see Chapter 28). However, although the customer-to-customer interface will use a B channel with 64 kbit/s capability, not all of this bit rate will necessarily be extended. Instead, within the inter-exchange part of the ISDN, a connection with a bit rate appropriate to the requested bearer service is provided. Thus in the case of a 64 kbit/s bearer service, a *clear* digital 64 kbit/s path will be chosen. However, when only *speech* bearer service is requested, the exchange may select either a standard digital telephone channel of 64 kbit/s or may choose instead a low bit rate digital connection,

routed via *circuit multiplication equipment (CME)* (see Chapter 38) or perhaps even a link carried over analogue transmission plant. Similarly, for a 3.1 kHz bearer service, a connection of at least 3.1 kHz bandwidth will be selected. Calls connected over ISDN in a circuit-switched manner include telephone calls, telex calls and 64 kbit/s circuit-switched data calls.

*Packet-switched* connections (Chapters 9 and 18) may also be established over ISDN, using either the B channel or the D channel. In the B channel mode (ITU-T recommendations X.31 case A and X.32), a circuit switched (B channel) connection is first established to a nearby *PAD* (*packet assembler/disassembler*) on the packet network, and then X.25 protocols are used during the 'conversation' phase of the call. However, a more efficient method of packet-switched connection is to use the D channel (ITU-T recommendation X.31 case B). In this instance the user's data packets are interleaved with other signalling messages on the D channel, and are transferred between ISDN exchanges either using the SS7 inter-exchange signalling network (Chapter 12) or via the established X.25 packet network. Figure 10.5 illustrates both cases. Figure 10.5(a) illustrates access to an existing packet network using dial-in via a B-channel of ISDN (*minimum integration scenario*). X.32 defines this procedure and the user identification measures which should be conducted when either dialling-in or dialling-out of the packet network in this way.

Figure 10.5(b), meanwhile, illustrates a *maximum integration scenario*, in which data information is sent from the user device (either DTE or TA) via the D channel. The data information is removed by the *signalling terminal (ST)* at the end of the D signalling channel and concentrated via the SS7 network to a *packet handler (PH)* which provides for an X.75 gateway connection to the packet network.

In addition to the support of *packet relaying services* (or *packet mode service*), ISDNs have lately been further developed also to support *frame relaying service* (*frame mode*, or



DTE = data terminal equipment    NT1 = network termination 1    PH = packet handler  
ST = signalling terminal    TA = terminal adaptor

**Figure 10.5** Carriage of packet data via ISDNs (ITU-T recommendation X.31)

*frame relay*) as defined ITU-T recommendation Q.92 (Chapter 20). The optional configurations for support of *frame relaying services* closely align with those of ISDN packet mode services. Thus a *frame handler* replaces the *packet handler* of Figure 10.5(b).

## 10.7 ISDN TERMINALS

Terminals used in conjunction with the BRI must conform with the I.420 S/T interface. This interface alone is not sufficient to ensure the correct operation of terminals for particular applications, because it only ensures correct establishment of the connection between like terminals. In addition, higher layer protocols (corresponding to the higher application layers of the OSI model) must be defined so that the two end terminals can interpret the data they are sending to one another. For example, two TV-phones operating over the ISDN must use the same picture coding technique (e.g. ITU-T's H.261 code) if the picture is to be transferred successfully. Other potential application

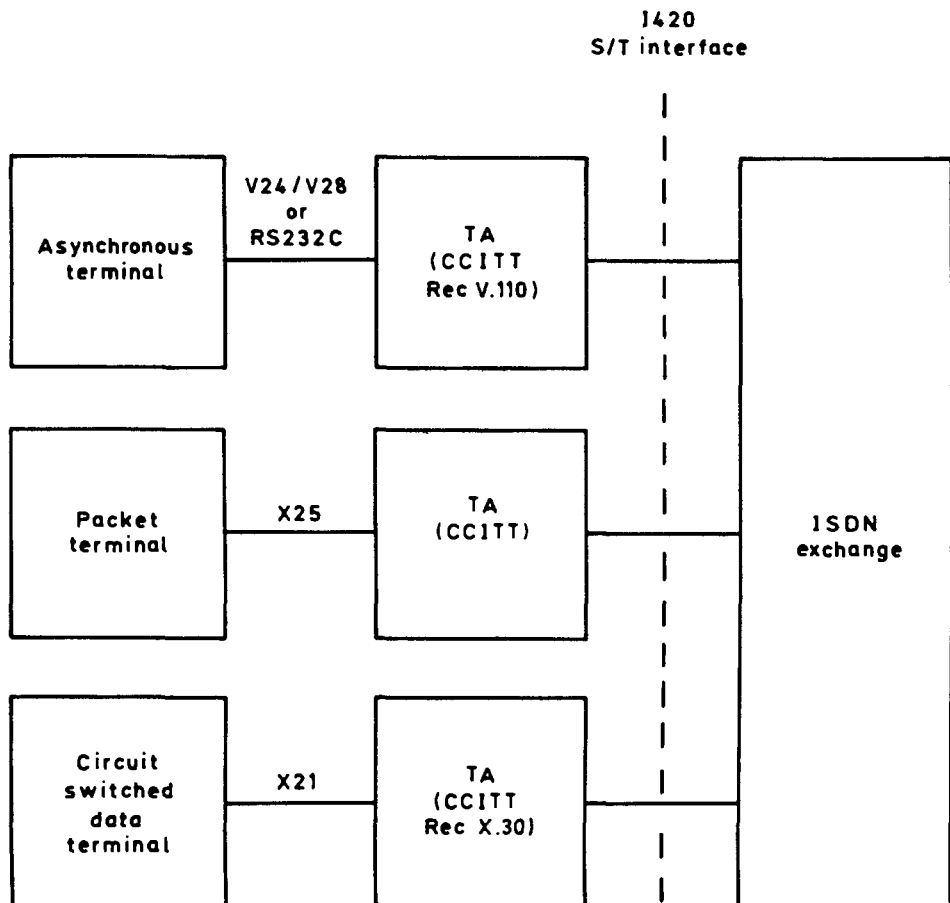
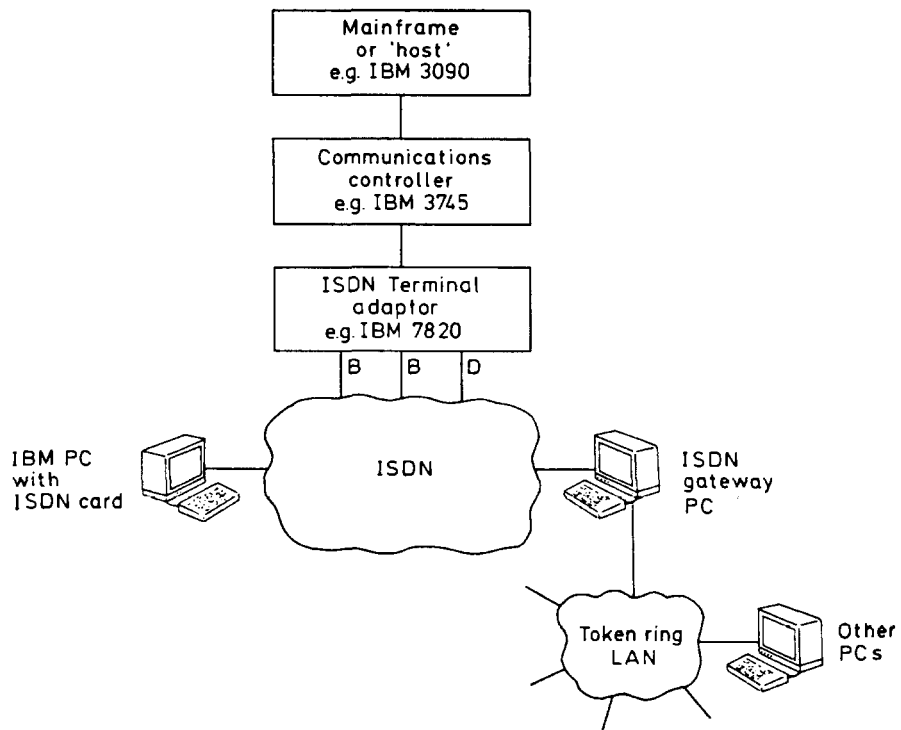


Figure 10.6 Common ISDN terminal adapters



**Figure 10.7** ISDN in an IBM computer network

layers are ISDN telephony (using A-law or  $\mu$ -law PCM), *group 4 facsimile* (using ITU-T Recommendation T.90 code) and 64 kbit/s personal computer file transfer protocols.

Developers of the ISDN recognized early on that the development of, and investment in, new terminals would be a constraining influence on ISDN take up, and so a range of terminal adaptors (TAs) have been defined in ITU-T recommendations. These convert established interfaces into a suitable format for connection to the ISDN S/T interface.

The diagram in Figure 10.6 shows some of the possible terminal adaptor configurations, giving the ITU-T recommendation numbers pertaining to them. Figure 10.7 illustrates a typical mainframe and personal computer network of the future employing IBM equipment, the ISDN and terminal adapters.

## 10.8 PRIMARY RATE INTERFACE

The *PRI* (also called the *PRA*, or *primary rate access*) is a four-wire interface optimized for connection between the public network and customer PBXs, and is best suited to the needs of medium to large customers. The format of the European PRI is  $30B + D$  (and is specified in *NET5*); i.e. 30 separate 64 kbit/s B channels, plus one 64 kbit/s D channel. (Note that the D channel is given greater capacity than in the BRI,



**Figure 10.8** ISDN terminal in use. ISDN telephones on executive secretaries' desks at Bell Laboratories in Indian Hill, Illinois, USA. These 40-button digital telephones provide such features as calling number identification, displayed on the liquid crystal display. (*Courtesy of AT&T*)

commensurate with the larger number of channels to be controlled. However PRI and BRI use a similar protocol stack.) The PRI format has been carefully designed to conform with the standard 32-channel, 2048 kbit/s pulse-code-modulation transmission format specified in ITU-T recommendation G.703. To conform with the ITU-T 24-channel PCM standard (1544 kbit/s), an alternative primary rate interface of  $23B + D$  is also defined. This is the one used in North America and Japan.

The primary rate interface is covered in ITU-T Recommendation I.421. The PRI is illustrated in Figure 10.9, where the PBX is providing a concentration function between a number of end users' individual BRIs and the network PRI.

For economic reasons, the exact traffic value at which a customer should subscribe to PRI service, as opposed to multiple BRI service, depends on the pricing policy of the local administration. Typically, PRI is worthwhile at values above about five erlangs (five simultaneous calls, around 12 exchange lines).

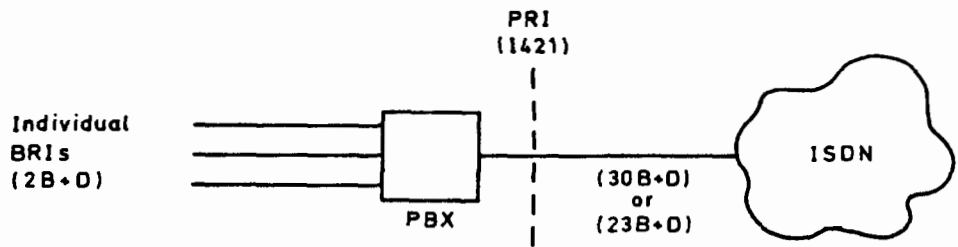


Figure 10.9 Primary rate interface (PRI)

## 10.9 THE PUBLIC NETWORK AND ISDN

A radical new style of customer/network interface as implied by ISDN must be accompanied by equally radical changes to the public network. The exchanges must be fully digital, switching individual channels at 64 kbit/s, and must support a common-channel-signalling protocol between exchanges. The SS7 *integrated services user part* (ISUP) described in chapter 12 (SS7 ISUP: ITU-T recommendations Q.761–Q.764), is intended for the support of ISDN. This must be used with the SS7 *message transfer part* (MTP) and a network of digital transmission links illustrated in Figure 10.10.

An alternative signalling system to SS7 ISUP which may be used as an interim measure is the TUP+ (*telephony user part enhanced*) signalling system, specified by CEPT/ETSI particularly for use in Europe. TUP+ provides for restricted ISDN capabilities that will eventually be superseded by the ISUP.

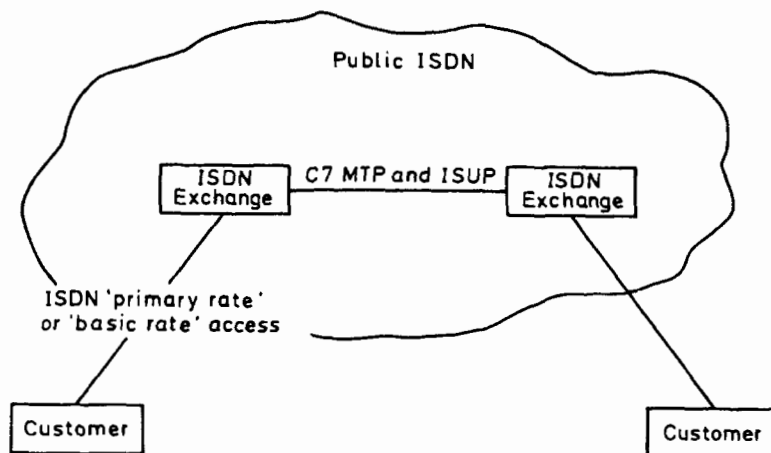


Figure 10.10 The ISDN

## 10.10 DEPLOYMENT OF ISDN

In any modern network, there needs to be a staged plan for the changeover from analogue to digital communications and a further plan for the introduction of ISDN. As ISDN requires a digital network, and because it will be paid for mainly by the business community, it may be possible for a network operator to arrange for the first digital exchanges, with digital transmission links interconnecting them, to be introduced in business areas. Then ISDN can be provided at marginal cost. Thus, although the changeover takes place in the rest of the country, the business community is already enjoying most of the benefits of ISDN, and the national economy is enjoying the investment which it generates.

It must be realized that the largest business investors are the multinational companies. Although they require good national telecommunications, technology is now such that they are coming to expect the same facilities internationally. The exchange of data between computers in different countries is already commonplace, and facilities such as international videoconferencing and fast facsimile will rapidly become requirements of many large companies.

A national ISDN is not enough. A PTO must be confident that it can interface with ISDNs in other countries so that its own ISDN achieves its full value. Meanwhile, for the multinational corporation, the availability of ISDNs in all relevant countries may be a necessary precursor to the switch of certain corporate telecommunication services to ISDN. How can compatibility be guaranteed when telecommunication system manufacturers in various parts of the world are independently developing their products? The answer is that all systems should work to the international standards recommended by the International Telecommunications Union Standardization sector (ITU-T). The first ISDN standards were documented in CCITT's *Red Book* in October 1984. They were updated in the *Blue Book* of 1988, and intense effort is continuing through the issue of *White Book* ITU-T recommendations (those issued individually subsequent to 1988) to ensure that different countries do not install incompatible systems. Even so, there are still significant differences between the various early national ISDNs. This may mean that terminal equipment compatible with one network is incompatible with that provided in another.

Given the fact that ITU-T standards for ISDN took a long time to be finalized, there are a number of reasons why a PTO may have elected to deploy its own pre-standard national ISDN.

First, it may have given the PTO an opportunity to gain experience with the new technology and simultaneously start to recoup some of the huge development costs already incurred in ISDN development and the installed digital network.

Second, it may have provided a much-demanded interim solution to the high speed digital connection services required by large corporations, and helped to stimulate end-user terminal manufacturers to invent and develop new applications that may have not previously been possible.

Third, such early ISDNs have helped to break the deadlock between CPE manufacturers and network operators, both waiting in a chicken-and-egg fashion for the other to make the first ISDN move. After all, ISDN terminals are of no value without an ISDN, and an ISDN is no use without terminals.



The appearance of CCITT's 1988 recommendations, including stable protocols for basic call control in both the Q.931 (D-channel) and the ISUP (inter-exchange) signalling systems, enabled rapid subsequent deployment of widescale ISDNs with international interconnections.

## 10.11 THE MARKETING OF ISDN AND THE EARLY USER BENEFITS

The early years of ISDN are likely to see marketing stress and user benefit in three main areas:

- the quality, cost, and 'frills' benefits made possible for telephone users of the ISDN
- the new high speed data (64 kbit/s) applications (e.g. high speed and high quality facsimile, fast computer file transfer, 'on line' services, videotelephones, etc.)
- the fast connection set-up times possible with ISDN

The quality improvements to ordinary telephone calls result from the introduction of digital transmission to the local access line, historically a major source of signal *noise* (e.g. 'fried eggs' noise). Cost benefits are likely to accrue to PBX users of the primary rate interface (PRI), as the PTO can pass on the cost advantage of getting 30 telephone circuits (Europe) (23 circuits in North America and Japan) from a single four-wire access line. Thus a cost saving should be possible by turning 30 (or 23) analogue exchange lines over to ISDN primary rate interface. A problem facing some PTOs, however, is that 1.5 Mbit/s or 2 Mbit/s digital access lines may already be available to customer PBXs, and the extra cost of ISDN signalling upgrade may not be perceived by the customer to be justified by the additional service benefits of ISDN.

ISDN service benefits form the central theme of many PTOs marketing campaigns for ISDN. These will accrue from the new supplementary services as they become available, e.g. *ring back when free*, *calling line identification (CLI)*, *call redirection*. The latter may be particularly useful to businesses receiving calls from customers, because it could be used for automatic presentation of the customers details on a computer screen, ready for quick reference. A new interface, designed to use ISDN-like signalling and called *ASAI (adjunct/switch application interface)*, will enable adjunct devices (e.g. computers) attached to ISDN exchanges to control or respond to supplementary service and other ISDN signals. Thus an ISDN PBX receiving a call and an associated *calling line identification signal*, could trigger action from an adjunct computer connected to the PBX. The call receiver might then be provided with a range of information on his computer screen (e.g. customer name and address) simultaneously being presented with the call. Alternatively, the adjunct computer might take charge of the call, instructing the PBX to divert the call or take some other action.

New high speed ISDN data applications can be targeted and promoted in an entirely different fashion, aimed at customers' needs that had not previously been met. The explosion of interest in the Internet, for example, has created a large demand for high

speed dial-up data connections. Meanwhile high speed facsimile, videoconferencing and picture phone devices are experiencing a boom in demand and the use of ISDN as a means of *dial back-up* for 64 kbit/s leaselines is also popular.

The fast connection set-up times possible with ISDN make 'online' services easier for computer users to access. No longer need the dial-in and log-on phase last for more than a minute. Instead the connection can be made and remade extremely quickly, reacting to the various requests of the user and the response of the 'online' server.

Perhaps the simplest, and single most beneficial effect of ISDN will be the common specification for worldwide network access. Ultimately this will mean that any I.420 or I.421 standard device can be connected to any ISDN anywhere in the world. In turn, the cost benefits of large scale production will mean that the future cost of ISDN devices and network ports will undercut the costs of predecesing technologies. In Europe, the EDSS1 (European DSS1 signalling as defined by ETSI based on Q.931) signalling protocol and physical interface will be standard throughout the European Union, because European mandate obliges all the dominant network operators to support it. On the other hand, regional variants are likely to continue to exist and thus thwart the goal of a uniform global network.

A company's ISDN portfolio will depend on its assessment of its important markets and applications and on its networking constraints. Thus some network operators have decided to push the 'improved telephone' aspect of the primary rate interface first, and other operators and corporate network users and planners have concentrated on the new high speed data application potential of ISDN.

Given the difficulty in matching up and interconnecting early ISDNs, a typical evolution path is likely to be: first, an unsophisticated 64 kbit/s switched service, followed later by the supplementary service benefits of fully-developed ISDN signalling. Finally, consistent international services will be possible.

## 10.12 NETWORK INTERWORKING

Network operators need not switch instantly from their current networks to an entirely new ISDN. The provision of network interworking avoids this need. Many administrations are providing early ISDN services simply by deploying a small overlay network, serving only a restricted number of ISDN customers, but allowing cross-connection to customers of existing telephone and data networks. This has the benefit of constraining the amount of early investment, but has the disadvantage of having to interwork with a number of separate existing networks, each with its own protocols (as in Figure 10.11). Such an overlay network allows an ISDN customer to communicate with customers on all the other networks, but it may have severe service constraints.

## 10.13 COMPANIES' PRIVATE ISDNS (CORPORATE ISDN)

ISDN in the private and corporate network arena will be as important as public ISDN. After all, it is the needs of the major telecommunication users that it aims to serve. Companies thus face a major strategic decision on when to start their (expensive)

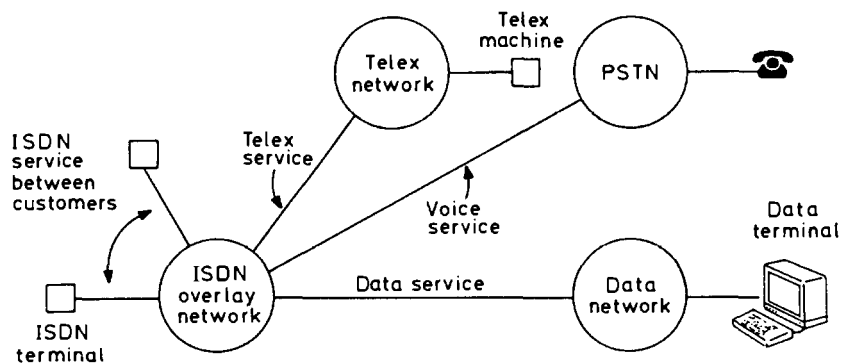


Figure 10.11 ISDN interworking

network evolution to ISDN and what technical standards to adopt. The decision needs to be tempered by the geographical locations in which they operate and the demands of the telecommunication services they use.

World standards for private ISDN have lagged behind those already developed for the public network, and a range of interim and proprietary standards still face the private network planner in 1996. The standards bodies have developed adapted (symmetrical) version of the PRI for use on private network connections between ISDN PBXs (ISPBXs). The interface is known generically as the *Q interface*, and is illustrated by Figure 10.12. It uses *Q-SIG* signalling protocol. Although it offers the potential to interconnect ISPBXs supplied by different manufacturers to provide private ISDN network services, it is still not as 'potent' as the 'proprietary' inter-PBX signalling systems that it is meant to replace.

In the absence of *Q-SIG*, the established *de facto* standard for inter-PBX ISDN signalling in UK, Canada and Scandinavian countries is the *digital private network signalling system (DPNSS)*, which was originally defined by PBX manufacturers and

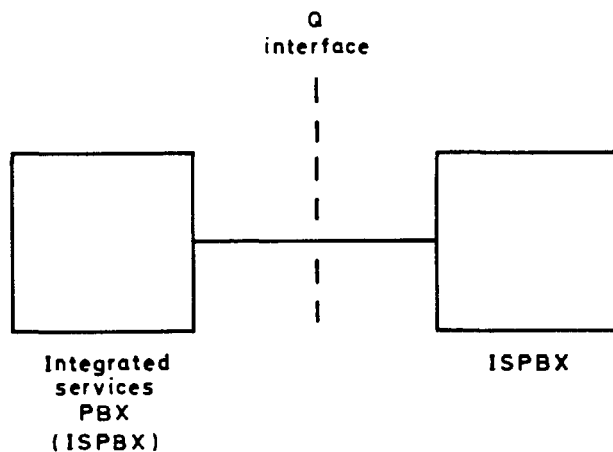
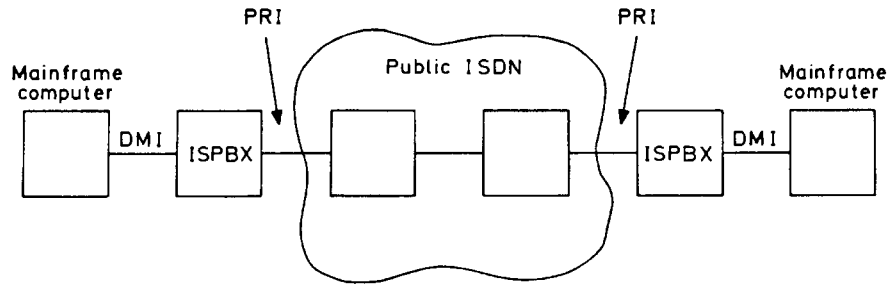


Figure 10.12 The Q interface between ISPBXs



**Figure 10.13** AT&T's digital multiplex interface (DMI)

network operators led by British Telecom in the United Kingdom. In other countries, the *de facto* standard has generally been set by the dominant PBX supplier (e.g. Siemens *Cornet* signalling is predominant in Germany).

Convergence on *Q-SIG* signalling is now likely, as it is the only method by which different devices made by a wide range of manufacturers will be able to interwork to provide private or corporate ISDN services.

Another interface of importance to private ISDN operators will be that enabling a mainframe computer to be connected to an ISPBX for high speed data transfer over the ISDN. This arena also is the realm only of proprietary standards at the moment, and AT&T's *digital multiplex interface (DMI)* provides one solution, as shown in Figure 10.13.

Finally, we should not leave the subject of private companies ISDNs without at least mentioning the scope for direct connection of company telephone exchanges with the public network using signalling system 7. Such access is likely to be available in the most developed countries. The prime benefit is the increased network control capabilities of SS7 (common channel signalling system 7) when compared with the PRI. However, security measures may need to be built into SS7 to make it more robust as a customer-connection protocol, before network operators are likely to open the interface for widespread use.

## 10.14 BROADBAND SERVICES OVER ISDN

Despite the fact that 64 kbit/s is a high bit rate compared with those possible over many established data networks, it is nonetheless inadequate for many applications of very fast inter-computer communication or video image transfer. In this context, the type of ISDN that we have described in this chapter is sometimes called *narrowband ISDN (N-ISDN)*. The emergence of new *broadband ISDNs (B-ISDN)* promises much greater bandwidth, as Chapter 25 reveals. However, in the shorter term, an adaptation of *narrowband ISDN* gives scope for greater bit rates than 64 kbit/s simply by combining an integral number  $n$  of individual 64 kbit/s channels as an  $n \times 64 \text{ kbit/s}$  bit rate path (i.e. 128 kbit/s, 192 kbit/s, 256 kbit/s, etc.). This is not quite as easy as it sounds. In concept it may seem simple to take consecutive bytes of a 128 kbit/s bit stream and feed

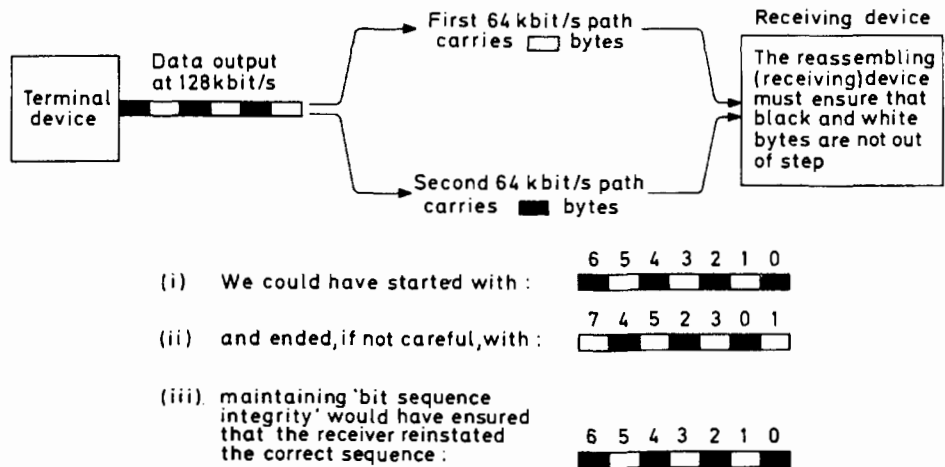


Figure 10.14 Maintaining bit sequence integrity

them alternately onto two independent 64 kbit/s paths (a *reverse multiplexing* process). In practice, having done so, it is much more difficult at the receiving end to ensure that the bytes carried on the two separate paths go back into a single string in the correct sequence. We have to maintain *bit sequence integrity*. Figure 10.14 attempts to illustrate the problem.

In addition to the problem of maintaining bit sequence integrity there is the simple problem of being able to set up the required number of 64 kbit/s connections. In theory, it may be possible to use a 32 channel *reverse multiplexor* to back-up a 2 Mbit/s lease-line. In practice, however, not all the necessary channels may be free within the network at the time required, so that it may take some minutes to re-establish the connection and for the necessary synchronization process to take place.

---

# *Intelligent Networks and Services*

---

By storing a massive ‘memory’ of customer and service information in a network, and referring to it while setting up calls, and as a historical record of network use, a phenomenal new range of services becomes possible. The effect is almost as if the network had some degree of ‘intelligent’ power of thought. This chapter commences by describing the ‘intelligent networks’ as a concept, and then goes on to give examples of the new services that we can expect from it.

## 11.1 THE CONCEPT OF INTELLIGENT NETWORKS

The concept and development of *intelligent networks (INs)* originated in North America. The forerunner was AT&T’s database *800 service*, and AT&T continue to be a key driver of the technology. Subsequently much work has also originated from the *RBOCs* (the American *Regional Bell Operating Companies*, or local telephone companies), in conjunction with their jointly funded research arm, Bellcore. More recently, ETSI (the European Telecommunications Standards Institute) has been very active.

The concept is based on the premise that all services can be broken down into elemental capabilities called *functional components* or *service-independent building blocks (SIBs or SIBBs)*. For example, a simple service may include providing dial tone, collecting digits, performing number translation, switching the connection, and charging at an appropriate rate. If we were now to examine a second service, then we would find that some of the functional components used in that service would overlap those already identified in the first. If a comprehensive set of these functional components (SIBs) could be implemented at every exchange (so-called *service switching point (SSP)*) and if a suitable means of controlling the exchanges, from new powerful and remote computers called *service control points (or SCPs)*, could be found, then new and much more powerful services could be implemented simply by writing software (a *service script*) for the SCP, enabling it to manipulate the SSP(s).

## 11.2 INTELLIGENT NETWORK ARCHITECTURE

In the past, each exchange had at least a small amount of 'intelligence', comprising software programs and related data for call routing and service control of 'basic' telephone services. However, when we speak of an *intelligent network* we mean a network equipped with a much larger information reference store and with software capable of controlling more powerful services.

The *intelligence* can be added to the network on either a *distributed* or a *centralized* basis, according to the circumstances of the established network, the equipment to be used and the service to be provided. Here we compare and contrast the two architectures as a means of illustrating the scope of possibilities.

In a network employing *distributed intelligence* the information required for advanced call routing and service control is spread over a large number of sites or exchanges. Each exchange stores a large store of information necessary for the set up and control of the wide range of services it is expected to offer. This will include a store of customer-specific data (the information pertinent to a given customer's network), as well as some *service logic* to tell the exchange exactly how each sophisticated service works, and the procedure for setting up calls. This sort of intelligent network could be created by continual enhancement of today's exchanges, progressively adding software and hardware to cope with new service needs. The advantage of such an approach (storing information at a large number of exchanges) is that the service becomes available at all existing exchanges and the call handling capacity is large. The disadvantages are that the exchange software becomes very complex and the job of keeping all the exchanges' software up to date is unmanageable. Not only that, but the software and data duplication increases the risk of inconsistencies and may affect the

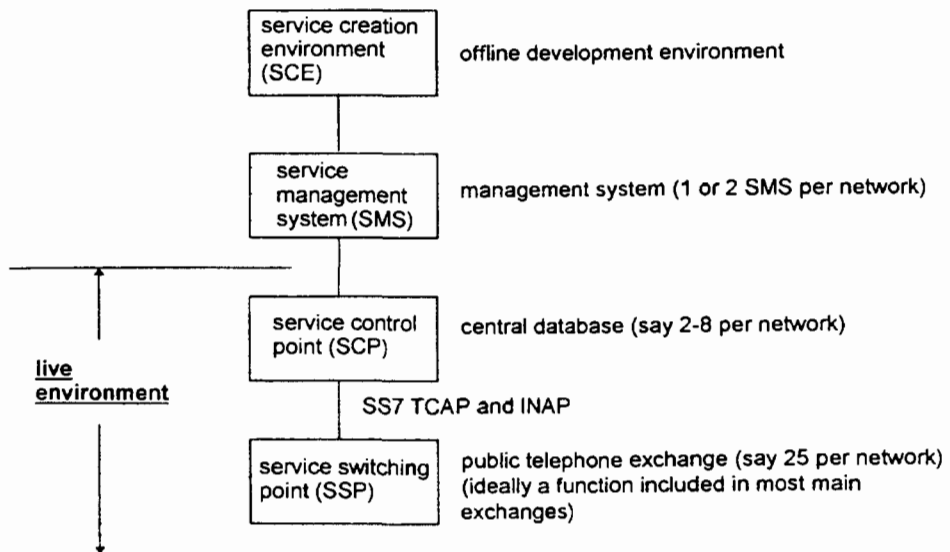


Figure 11.1 Intelligent network architecture

smooth running of both the service and the network as a whole. (For example, two exchanges may hold conflicting data because they were updated by different people at different times.)

Figure 11.1 illustrates the standard 'centralized' *intelligent network (IN)* architecture. In the lowest tier of a centralized intelligent network are a number of *service switching points* (or *SSPs*). These contain a *service switching function (SSF)*. These are enhanced telephone exchanges which have been developed to include a new *intelligent network* interface. The interface allows the exchange to refer the call control of advanced service calls to the *service control point (SCP)*, allowing the SCP (a centralized control point) to manipulate subsequent actions of the exchange.

### 11.3 THE SERVICE CONTROL POINT (SCP)

The service control point (SCP) is a specialized computer, distinct and often remote from the exchange, connected by a *signalling system number 7 (SS7)*, see Chapter 12) signalling link. The SCP comprises the *service control function (SCF)*, the 'knowledge' of how a service works and the customer-specific data required to perform it.

In response to an exchange request to deal with an 'advanced service' call attempt, the SCP sends a sequence of *primitive* commands to the exchange, using SS7 signalling. The commands direct the exchange to perform the necessary sequence of simple switching actions which combine to appear as a more complex service offering (e.g. 'collect digits', 'connect switch path'). Thus the call control is carried out by the SCP rather than the exchange (SSP). However, the connection itself never passes through the SCP. Connections are always only switched by exchanges (the SSPs).

During call set-up, call processing is suspended so that the SSP may refer to the SCP. The reference may reveal, for example, whether a given caller is permitted to be connected to a given number. Alternatively a credit card number might be validated before accepting call charges, a dialled *freephone* number might be interpreted into another 'real' telephone number, or some other 'intelligent' action may be undertaken.

The SSP sends an SS7 message to the SCP containing the dialled number and any other known information about the called or calling party. The SCP interprets the call request using the received information and its own store of data, and then returns the sequence of commands back to the SSP. The specially developed *user parts* of SS7 signalling which enable this interaction are called the *signalling connection and control part (SCCP)* and the *transaction capability application part (TCAP)*, ITU-T Q.771-Q.775) and the newly defined *INAP (intelligent network application part)*. These are described in more detail in Chapter 12.

The number of SCPs deployed in any given *intelligent network* depends on a number of factors, including the complexity of the service logic required to support the advanced services and the traffic demand for them. One option is to allocate one SCP for each individual advanced service, but for a large number of services we would need a large number of SCPs. Some experts therefore favour SCPs which are capable of handling a number of different services, so that the number of SCPs in a network can be kept down to a handful. In this case each will cater for a number of services, but each service will be duplicated over more than one SCP to prevent total loss of the service in the event of an SCP computer failure.



## 11.4 THE SERVICE SWITCHING POINT (SSP)

The *service switching point (SSP)* or *service switching function (SSF)* is a modified telephone exchange. Over and above the normal functions of a telephone exchange it contains an 'intelligent network' functionality comprising

- *trigger tables*
- *transaction capabilities*
- *intelligent peripherals (IPs)*

Every telephone exchange has a number look-up table of some form to enable it to switch calls through to their correct destinations. In the case of an SSP *trigger table*, the information needed to complete the call set-up is not contained in the table itself; instead there is a *trigger* (typically activated by the dialled number) to commence a query *transaction* with the SCP. The SSP next collects all necessary information about the call (caller's number, class of service, dialled number, etc.) and forwards this information to the SCP to request further control information.

The information and short dialogue which then follows between SSP and SCP is called a *transaction* and is conducted using SS7 TCAP signalling (*transaction capabilities application part*). During the dialogue the SCP returns a number of control commands to the SSP to control its switching and charging functions, and also to activate any necessary *intelligent peripherals (IPs)*.

*Intelligent peripherals* could be any number of different types of device affording different types of advanced service. At the simplest an IP might be a recorded announcement machine (say, thanking a televoting caller for his interest). At a more complex level it might be a voice interaction unit.

## 11.5 THE SERVICE MANAGEMENT SYSTEM (SMS) AND SERVICE CREATION ENVIRONMENT (SCE)

The *service management system (SMS)* or *service management point (SMP)* appears above the SCP and is used to control the SCPs in a network. SMSs are *offline* computer systems used to prepare database and configuration tables of network and customer-specific data before downloading them to the live SCPs. The *SMS* ensures that the data held in all SCPs is comprehensive and consistent. The fact that only one SMS exists, or a small number of SMSs, makes the manual task of administering data held within the network a great deal easier. The service risk of running only a single SMS is not material because it is only an 'updating machine'. The service availability is affected mainly by the reliability of the SCPs and SSPs.

The *SCE (service creation environment)* is the platform for SCP software development and testing (i.e. it is the tool with which new 'intelligent' services can be developed). Usually it comprises sophisticated software *debugging* tools, capable of 'stepping' through the programmed commands within a new software *service script*. These help the service designer ensure that the service is realized in the manner intended.

## 11.6 BENEFITS OF INTELLIGENT NETWORKS

The main advantages of an intelligent network over predecesing public telephone networks is the ease with which complex, and particularly network-wide, services may be managed. Instead of having to configure routing tables and other network control elements which are distributed across many exchanges, the network operator only needs to maintain the data in the central SCP. This guarantees a higher level of data consistency within the network and thus of service reliability. In addition, the network operator is able to react faster in the introduction of new services. This leads to

- minimal impact on existing network and switching equipment during the rapid introduction of new services (the introduction requires only the downloading of new service script software and configuration data to the SCP (from SCE and SMS))
- reduced cost of introducing and enhancing services
- higher quality of service
- the ability for rapid re-configuration of services, allowing continual retuning to meet changing market needs (the use of a single SMS means that the job of coordinating network upgrades is largely eliminated)
- the ability to give the limited customer control and management facilities if required (by providing special customer terminals connected to the SMS, customers could be authorized to make some changes specific to their own networks)

## 11.7 INTELLIGENT NETWORK (IN) SERVICES

Certain types of telephone services are best realized using an intelligent network. This applies to those types of service where either the charging requirements of the service are complicated (e.g. there is a need to charge the person called and not the caller), or where the handling is complicated (e.g. caller authorization is necessary or complicated *translation* of the dialled number is necessary, as for example by freephone numbers where a dialled 800 telephone number must be converted to the standard telephone number of the called party).

Examples of intelligent network services include the following.

### *Virtual private network (VPN)*

A service in which a company-specific network (a telephone closed user group with a specific telephone numbering plan) may be created for individual corporate customers of the public network. The public network thus appears to the corporate customer much as a private network would, with a 'tailored' company numbering plan.

### *Freephone*

The intelligent network converts an 800 dialled *freephone number* into a standard telephone number allowing the SSP to complete the call set-up, while simultaneously

creating a call charge record for the call receiver's account (rather than for the caller's account as is normal).

#### *Premium rate service*

The ability to charge a premium rate for calls over and above normal telephone charges so that the caller shall be charged for information services (e.g. weather forecast, traffic information, etc.). The extra charges collected are furthered to the information *service provider*.

#### *Calling card service*

A service which allows telephone company calling card holders (e.g. AT&T calling card) to make calls from any telephone in the public network, invoicing their call charges to their personal calling card account. At call set-up, IN verifies the card account number and requests caller authorization by means of his personal identification number (PIN).

#### *Televoting*

A service conceived to complement television game shows in which viewers are invited to call different telephone numbers to register their vote for the best participant in, for example, a television game show. IN counts the total number of calls to each dialled number and connects the caller to a recorded announcement which thanks him for his call.

#### *Universal number service*

This service enables customers of the public telephone network to move around the country while remaining available under the same telephone number. As the user moves to a new location, he must register with the SCP where he now is, so that future calls to his number may be furthered to him. This may become the basis of *number portability* service, the ability for a customer to change his telephone network provider without being forced to change his telephone number. The difficulty caused by changing number, requiring one to print new letterheads and advise business partners, might otherwise dissuade a change of telephone network provider, so that *number portability* is increasingly viewed as an essential enabler of competition between public telephone service providers.

#### *Universal personal telephone service*

This service is an extension of the universal number service, allowing the customer not only to roam within the 'fixed' telephone network but also to connections of mobile telephone and other types of telecommunication networks.

## 11.8 CALLING CARD

Using network intelligence to validate *calling card* or *credit card* account numbers and as a historical record of transactions, an alternative means of paying for calls is possible. Calls made from any telephone can be charged to a special 'calling card'

account. The bill can then be sent to the card holders address. Maybe such a service will make obsolete the familiar public payphones, or at least the ones for which you need handfuls of coins to operate.

There are three ways of initiating a call. In one the caller tells the operator the card number and the *personal identification number (PIN)*. The operator types this information into a computer which interrogates the SCP to check that the card is valid, and subsequently charges the cost of the call to the appropriate account. An alternative is an automatic version that relies on the customer being prompted to dial in his card account number and PIN using a DTMF telephone. Finally, a specially designed telephone with a 'card-wipe' system might also be available. In this instance, a magnetic strip on the reverse of the card is 'wiped' through a narrow channel on the telephone. The telephone 'reads' the magnetic strip to derive the *calling card* (or standard credit card) type and number, and automatically validates the card, notes its expiry date and other details by using the network intelligence in the same way as above. If the card is not valid, or if the caller dials in the wrong PIN then the call is not permitted.

The beauty of using central intelligence of the SCP to validate cards is the scope that it gives for tailoring calling capabilities of the card to its owner's needs. A student's parents can give their son a calling card with which he can only 'call home' (as in MCI's 'friends and family' service). Other calls are at the student's expense. Similarly, a company representative can be given the means to call his office.

Calling card service is growing in popularity in countries where it is already available, and most major PTOs are planning to introduce it. Figure 11.2 illustrates an example of a telephone designed especially for automatic card validation.

## 11.9 FREEPHONE SERVICE (OR 800 SERVICE)

*Freephone, toll-free, nulltarif* or *800 service* is available in a number of countries. In the UK callers who dial a number in the 0800 range, and in the US callers who dial a 1-800 range number have those calls completed entirely free of charge. The call charge is paid by the recipient of the call.

*Freephone* service gives companies a way of persuading people to call them. A company may wish to promote calls to follow-up an advertisement campaign, or to allow customers to call the service department, or maybe to allow their travelling representatives to call the office.

Network intelligence plays two key roles in support of the freephone service. First, the 0800 number dialled (say, 0800 12345) must be converted into the receiving company's actual number, say 071-246 8021, otherwise the normal telephone network will be incapable of completing the call. The second role is to record the total number and duration of calls made, so that the call recipient can be charged in due course.

Figure 11.3 shows a diagram of automatic freephone service. The caller has dialled the number 0800 12345 into the network. The SSP sends the number to the SCP, which returns the normal telephone number to the network (to allow routing), and records the time of day, call origin and call duration, so that the recipient (071-246 8021) may be charged for the call by normal quarterly account.



**Figure 11.2** Credit card telephone. Telephone specially designed to allow payment for public telephone calls by credit or calling card. (Courtesy of British Telecom)

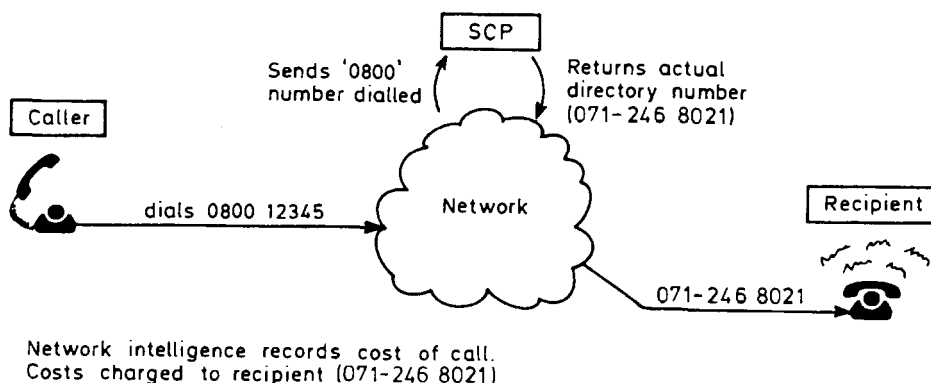


Figure 11.3 Automatic freephone service

## 11.10 900 SERVICE

The United States *900 service* uses a similar intelligent network to that of the 800 service, but rather than calls being free to callers they are charged at a premium. The premium charge covers not only the cost of the call itself, but also the cost of *value-added* information provided during the call. Thus a typical 900 service might be 'dial up weather forecast' or 'dial up sport news'.

The value-added information is provided by a service independent of the PTO who pays for the provision of 900 service facilities but receives revenue from the PTO for each call made.

The role of the SCP in the 900 service is to ensure translation of dialled 900 numbers and to record call attempts for later settlement of account between PTO and service provider.

In other countries the service may be known under different names; the UK equivalent, for example, is the *0898 service*.

## 11.11 CENTREX SERVICE AND VIRTUAL PRIVATE NETWORK

Many companies run their own automatic telephone and data networks on their own premises, using automatic *private branch exchanges (PBXs)* and private packet switches, etc. Some of these companies also lease transmission capacity from public telecommunication operators (PTOs) to connect together a number of geographically widespread sites into a single, company-wide, network. These private networks are always tailored to the company's particular needs, often supporting service facilities which are not available from the public network. For example, on a company's own telephone network, the allocation of extension numbers may be set according to departmental or company whim. In addition, other special features may be made available, such as *ring back when free*, *conference calls* and special call barring facilities (to prevent some extensions from dialling trunk or international calls).

The decreasing cost of private networks equipment, coupled with the restricted service facilities of some public networks, has recently stimulated a rapid growth of private networks. If allowed to continue by the public telecommunication operators (PTOs), this could pose a threat to revenue income, as less income is available from leased circuits than from the equivalent public network service. From their point of view it will be worse still in countries whose governments allow the resale of private network services. Faced with this, a number of public telecommunications operators and main exchange manufacturers have been developing new services to protect their market shares. *Centrex and virtual private network (VPN)* services are both products of the counter-reaction.

The *centrex* service provides facilities similar to that of a PBX, but from the public network's *local exchange* (or *central office*). This gives the customer benefits equivalent to owning an on-site PBX but without the 'up-front' capital investment, and without the ongoing need for expertise and accommodation to maintain it. All of the customer's 'on-site' telephones are connected directly to the public network's local exchange, which acts as if it were a PBX. For example, the customer may determine the extension numbering plan. In addition, features like *call interrupt*, or *ring-back when free*, etc., may be made available between extension numbers. Furthermore, just as in a PBX, only the *extension number* need be dialed to call other on-site company *extensions*. Figure 11.4 compares a *centrex* service with the comparable service provided using a PBX.

To the user of extension number 2435, on either the *centrex* (Figure 11.4(b)) or the PBX network (Figure 11.4(a)), it is not apparent which type of network is being used as the network and special service capabilities are identical. This makes it feasible for a small company to consider first subscribing to *centrex* service from the public telecommunication operator, and later installing an on-site PBX, when its cost is justified.

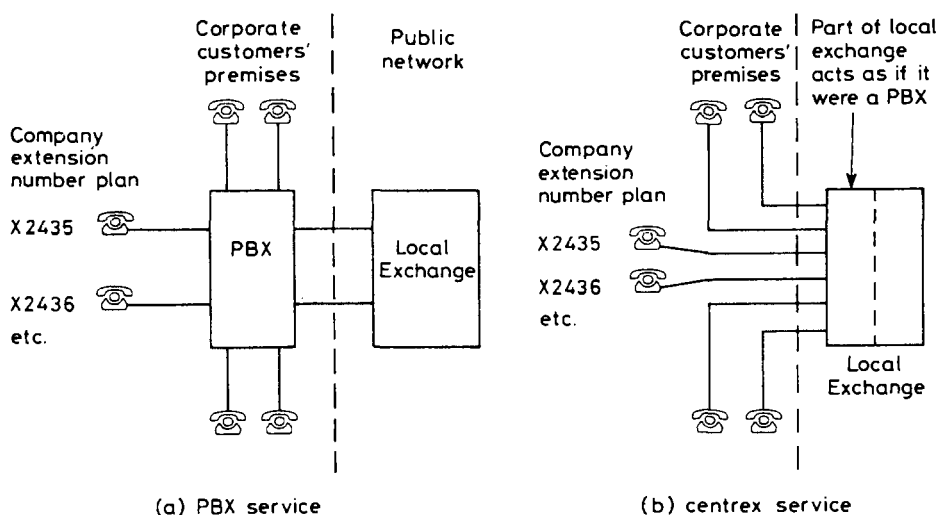


Figure 11.4 Company extension number plan using PBX centrex service

A major advantage of the *centrex* service is that it allows companies which are spread thinly over a number of different buildings within a locality to have their own PBX. If any of the locations is altered, there is no need for PBX upheaval; the PTO can adjust the *centrex* service to match.

A limiting factor of some PTO's *centrex* services is that the company's 'on-site' network must all lie within a single local exchange area. Where the company network spans a large area, spread over a number of local exchange catchment areas, then *centrex* service is inadequate, and an *enhanced centrex* or *networked centrex*, or a *virtual private network (VPN)* service is needed instead. The distinction between the two is not clear-cut; both use public network resources to provide for networking needs of geographically-diverse companies spanning several local exchange services. In this chapter we shall not labour the fine distinctions other than to say that while *enhanced centrex* assumes that the customer has no PBXs at all, VPN assumes that PBXs are in use at some of the sites.

As its name suggests, a *virtual private network* provides features and services similar to a multiple-site *private network*, while making use of the public network's resources to do so. The economies of scale available from a public network, both in switching and transmission, allow virtual private networks to be cost-competitive when compared with *private networks*.

In Figure 11.5(a) a company's *private network*, comprising two PBXs and a *leased line* interconnection between them, has allowed different extensions to receive private network style services, using the company's four-digit extension numbering plan. Extension numbers 2434-7 are illustrated: the user of extension 2436 need only dial the digits '2435' to call the user of that extension. By contrast, in Figure 11.5(b) the same services have been provided using a virtual private network. The leased line is not required, the connections between sites being made by the *virtual private network (VPN)* embedded in the public network. Nevertheless the user regards the network as a *private network*, and the user on extension 2436 need only dial '2435' to call that extension.

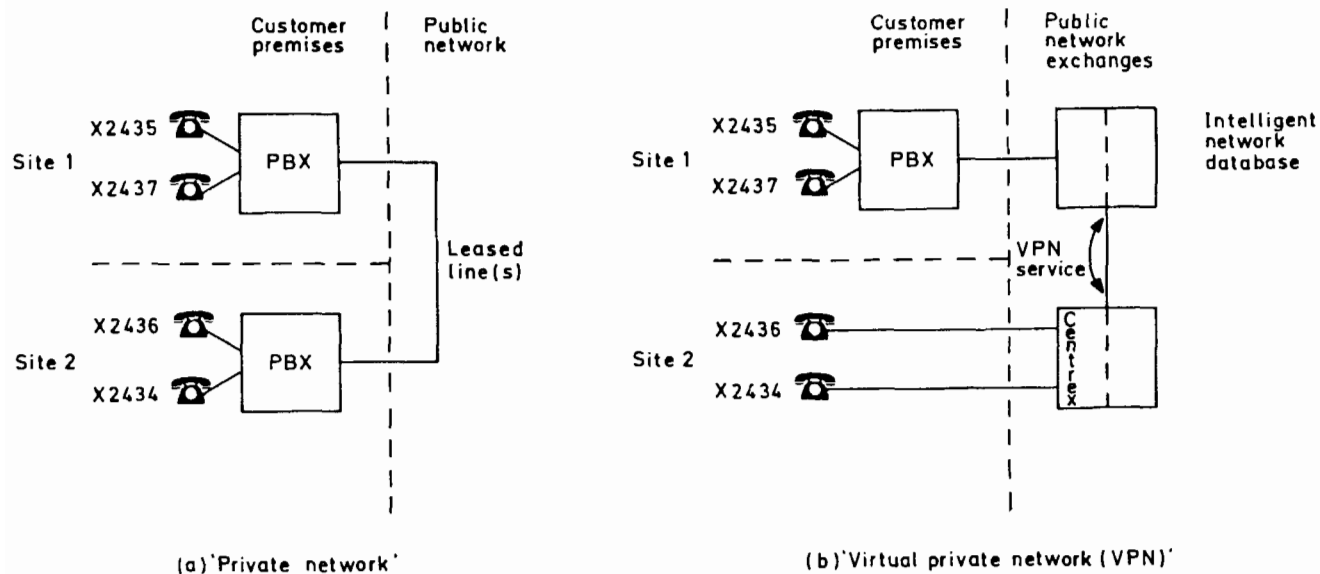
A feature of VPN service illustrated in Figure 11.5(b) is that a PBX or a *centrex* service may be used as the interface between the extension lines and the local exchanges of the public network. Thus extensions 2435 and 2437 are connected via a PBX (on site 1) while at the customer's second site, extensions 2434 and 2436 are *centrex* subscribers.

Both *centrex* and *virtual private network (VPN)* services rely heavily on network intelligence. As in freephone service, the number dialled by the calling customer requires number translation before it can be routed through the network. A charge for use may have to be recorded, in line with the PTO's overall tariff structure.

## 11.12 LINE INFORMATION DATABASE (LIDB)

An early application of network intelligence in North America was the *line information database (LIDB)*. In this application, the SCP contains a wealthy store of information about each line, and reveals what services are subscribed to by that customer. This enables the network to determine, at call set-up, whether a call attempt





**Figure 11.5** Company extension number plan extended to several sites using virtual private network

to a particular service should be allowed or barred. An example of the use of LIDB in North America, is its use by local telephone companies to record customers' preferred toll (or long distance) telephone company. It has been a requirement, since deregulation of telephone networks in the United States, for customers to inform their local telephone company which toll carriers they wish to pre-subscribe to. The requirement is laid out in the regulations of so-called *equal access* (1986). Few telephone exchanges at the time had the ability to record and react to the pre-subscription, hence the development of LIDB.

Another use of LIDB is to hold the current status of a particular piece of information, for instance whether a customer is currently 'at home', or instead wishes incoming calls to be re-directed to an alternative number. If re-direction of calls is required, the customer may programme the LIDB with the alternative number using a special dialling procedure. On receipt of each incoming call, interrogation of the LIDB by the customer's local exchange secures the re-direction.

### 11.13 TELEVOTING

An increasing practice on television programmes is to take an instant poll of viewers' opinions. Viewers are asked to dial one of a set of different directory numbers according to their answer to, or opinion on, a question issue. The following might be an example, on a TV sports programme:

#### *Question to TV viewers*

Which footballer, who has played in the World Cup, do you believe most warrants the accolade 'Best Footballer Ever'?

#### *Alternative answers*

For Pele	ring	0898 222001
For Diego Maradona	ring	0898 222002
For Bobby Charlton	ring	0898 222003
For Johan Cruyff	ring	0898 222004
For Jürgen Klinsmann	ring	0898 222005

Viewers ring one of the numbers to cast their vote, and a poll can be taken. This can be done manually by answering each telephone call in person, or automatically by using an intelligent network to record and count up all the votes, even going so far as to give 'votes-so-far' before the voting period ends. The service is called *televoting*. It is implemented by instantaneous or 'real-time' call counting at the SCP, while the customer is connected to a recorded announcement at the SSP, thanking him for his call and confirming that his vote has been registered.

## 11.14 CELLULAR RADIO TELEPHONE SERVICE

A *cellular radio* telephone allows its user to roam around a wide geographic area, making and receiving calls anywhere within that area. Outgoing calls from the handset are made via a nearby radio base station in response to the user dialling the public telephone number he wants. Incoming calls may be received by the handset, but herein lies a problem: how do we know where the mobile subscriber is, so that the call can be sent to the radio base station nearest to him? Well, in the early days of mobile telephones, callers were expected to know the geographic location of the mobile user before making the call, and to dial an appropriate area code.

In modern *cellular radio* networks, intelligence embedded within the network 'keeps an eye' on the location of the mobile user by a continual polling (or *registration*) process. This makes it possible to use the same area code and directory number for calls made to the mobile user, wherever he may be.

Cellular radio networks are split up into a number of cell areas, each served by its own transmitting *base station* which provides for radio contact between the *fixed telephone network* and mobile telephone users within the cell. When a user migrates to another cell, radio contact must be transferred to the base station in the new cell (hand-off). The transfer is initiated by a *mobile switching centre (MSC)*, an exchange which controls a number of base stations. While the mobile user stays within any of the cells of a mobile switching centre area, the mobile may be *polled* to receive its incoming call. Should it enter the area corresponding to a different MSC then the user must be re-registered in the new area. In essence, *re-registration* is an act of updating an *intelligent network* database, called the *home location register (HLR)*, which records each user's up-to-date location. Any MSC wishing to know the location of a user (in order to complete a call) asks the HLR for the information and then resumes call set-up. (Note: actually the HLR concept is defined by CEPT's GSM digital cellular radio scheme. TACS and AMPS work in a similar way, although the HLR and VLR are not specifically named as such.)

Figure 11.6 shows the registration procedure in outline; the subject is discussed further in Chapter 15.

The mobile user in the car shown on Figure 11.6 has migrated from the cell of a base station in MSC1's catchment area into the catchment area of MSC2. All incoming and outgoing calls must be set up via MSC2, and this is recorded by the database associated with MSC2 in the mobile's *home location register* (at MSC3). Note how, in this example, the SCP function has actually been *distributed* around the various MSCs. Although only one copy of each customer's data exists, not all the data reside in one SCP location. In some instances there is a clear benefit in distributing the SCP function in this way, as it allows for more traffic to be handled. As traffic growth outstrips the capacity of a single SCP to handle it, the *service logic* and customer data can be shared out between a number of exchanges so that many more call attempts can be handled simultaneously. (In essence some or all of the SCP's functions are re-located in the exchanges). The SSP part of an exchange still needs to make the referral to the SCP part to take over advanced call control, but this function is no longer located in a specialized and distant SCP computer. A special form of SS7 signalling, the *mobile application part (MAP)*, was developed especially for this purpose. The SMS updating systems works in exactly the same way – but now direct to the SSP/SCP combined unit.

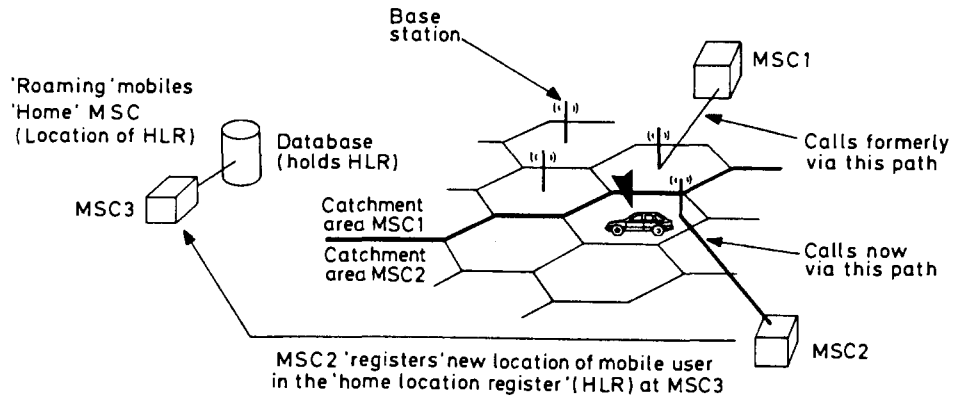


Figure 11.6 Cellular radio registration procedure

## 11.15 NETWORK INTELLIGENCE AND PBXS

By adding 'intelligence' to a PBX, a wealth of new services became possible. Three examples are as follows.

- Automatic 'wake-up' call for hotel customers. The request for the wake up call is made by the hotel customer by dialling a service selection code, plus the time. The 'intelligence' stores this information and initiates the PBX to make the wake-up call at the right time.
- Selection of the cheapest public network carrier. In a country where the public telephone service has been deregulated, it sometimes happens that a PBX is connected to the networks of two or more competing public network carriers. At a particular time of day, one of the carriers may be the 'preferred carrier' on grounds of cost or network congestion but at a different time circumstances may favour the other carrier. By giving the PBX some appropriate 'intelligence', the 'preferred' public network carrier can be adjusted according to the time of day.
- Computer prompting on incoming calls. In company telephone bureaux receiving large numbers of incoming calls, and where the operators key information into a computer during those calls, it is valuable for the computer and telephone equipment to be linked. The effect can generate a fresh computer 'form' automatically for each new caller. In addition, the computer can signal to the telephone equipment when the last form has been completed – making it ready to receive the next caller.

Figure 11.7 shows an intelligent network based upon a small computer and a PBX. In the example shown, switch and signalling developments have been necessary on both the PBX and the computer. However, unlike the public network case, SS7 signalling has not been used. Instead the intelligent network signalling is of a special type, called HCI or *host computer interface*, but there is no over-riding standard as yet. All the 'intelligent' interfaces in the PBX market are *proprietary* to particular manufacturers and several different versions exist.

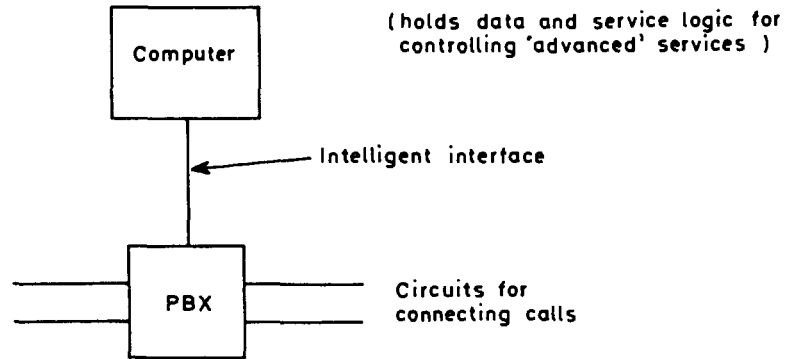


Figure 11.7 An intelligent PBX architecture

## 11.16 VOICEMAIL AND VOICE RESPONSE SYSTEMS

Voice response systems are becoming widely used in association with intelligent network services, and are important also in their own right as *voicemail*. We discuss these services next.

Voicemail works as if it were a central bank of answerphones. A caller wishing to leave a message calls the central 'voicebank' and leaves a verbal message for his addressee. This is usually quite short and informal in style. Subsequently, the addressee, who should be in the habit of calling the voicebank several times during his day, may find a number of short messages from various business colleagues when he does so. The system allows him to listen to each in turn, and immediately following each one gives him a number of options:

- to reply to the caller, perhaps copying the message to some other interested individuals
- to forward the message, perhaps to a subordinate for some action
- to store the message for later consideration
- to note mentally the message and then delete it

To execute any of these actions requires only a few simple key actions on either a tone-type telephone or on a special pocket-sized tone keypad, and then the human voice is needed to record the reply. The user is saved the inconvenience of ringing back any of his callers and manages to handle all of their enquiries at a time suitable to himself, rather than being disturbed by the telephone in the middle of an important meeting.

Voicemail suits users within a business community of interest, who might otherwise find it difficult to communicate, either because of:

- being permanently on the move (for example, field service or sales staff), or because of
- international timezone separation (there is only a very short and inconvenient time window for somebody in the UK to telephone someone in Australia), or simply because of
- being 'away from my desk' or 'already talking on the phone'.

Keen users of voicemail point to the ease with which broadcast messages may be generated, perhaps the weekly report to the troops. In addition, they praise the direct and informal nature of the messages which users soon get in the habit of leaving. Unlike the telephone, there is no need for chit-chat and social niceties, just straight down to the business.

Recently, it has become common for large companies to tie a voice mailbox to each employee's telephone line extension. Now the mailbox owner can have messages diverted to the mailbox either when he is away from his desk, or when he is busy. In replying to messages, he can either call the person directly or may send a message to their voice mailbox. When out of the office, he may log in to his mailbox remotely. Provided the user checks his mailbox frequently (several times per day), even the most-travelled employee may appear to be only a few yards from his desk.

An example of the effective use of voicemail might be its application in a sales force. The regional sales manager is easily able to broadcast targets and his weeks news. The individual salesmen can report back orders, or perhaps direct customer questions.

A further potential of voicemail technology is in providing voice-prompted control of systems. On calling a given telephone number, for example, a *voice response system* could answer and ask the caller the nature of his call: 'do you want our customer service department (dial 1), our sales department (dial 2), an account enquiry (dial 3) or some other matter (dial 0 for human assistance). According to the caller's dialled response, the system could either:

- provide answers itself to specific caller enquiries
- ask further questions
- direct the call to an appropriate department, or specific individual
- record details (e.g. of a customer order), either 'dialled' by the customer and stored directly in a computer, or stored in a verbal form for handling by a human. This might allow 24 hour and weekend ordering, for example.

In North America, *voice response* systems are already very common. You can, for example, call Montreal airport, punch in the flight number and get up-to-date news on flight arrival or departure times. Meanwhile, Air Canada's frequent flyer program allows you in your next call to check your latest mileage credit, find out about award claims and this month's special offers or merely request further literature. Service information bureaux allow you to listen to the latest weather forecast, sports reports, snow depths and even the recommended wax to be used on cross-country ski trails.

Applications in Europe are currently only in their early stages of development, but are evolving fast. Initial European applications have included the delivery confirmation

and fleet management of trucking distribution companies, simple order process systems, and 'noticeboard' systems for police 'community neighbourhood watch' schemes. Voice response systems in Europe continue to be held back by the relatively low penetration of tone-signalling phones, though hand-held signalling units have provided an adequate alternative for some of the applications already described.

### 11.17 CONSIDERATIONS BEFORE INTRODUCING IN TO A NETWORK

Before the invention of *IN*, *freephone*, *VPN* and some of today's other *intelligent network services* were realized using telephone switch-based (i.e. distributed) control software. By most telephone companies such realization is no longer considered manageable.

To date, *all intelligent networks* are to some extent based on manufacturer-proprietary interfaces, protocols and *service script* software. There are no large scale networks in which an SSP from one manufacturer works in conjunction with an SCP of another manufacturer, developed in isolation. However, due to the huge investment in existing exchanges, most public telephone network operators have an interest to create SSPs by software/hardware upgrades to some or all of their most modern exchanges. This has created pressure for standardization of the SSP/SCP interface.

There are not yet stable technical standards for the other IN interfaces (SMS/SCP, inter-network SCP/SCP, etc.). This will be a limiting factor in the speed of development of intelligent network services spanning network and sub-network boundaries. Effort will continue on developing these interfaces, fuelled by the desire particularly of European and United States governments to ensure competition in *intelligent* and *value added* public telecommunications services.

Although IN represents a significant step forward in the management of voice telephone networks, particularly making feasible some of the more advanced services, it is not yet realistic to expect all telephone calls within the network as intelligent network calls. Although this might have some operational management advantages, the difficulties of processing all the call set-ups at a single central point (the SCP) are with today's technology insurmountable.

### 11.18 THE FUTURE OF INTELLIGENT NETWORKS

Intelligent networks are a sophisticated but complicated service-enabling architecture, and so their achievement on a network-wide basis will take an extensive programme of signalling and switch enhancements over many years. Definition of appropriate standards will be essential to the interworking of intelligent networks around the world and even between different vendors' equipment within the same network. Only in this way will the network providers be able to offer effective and cost-efficient services meeting the bulk of customer needs. The enhancements, however, will mean many more powerful and flexible services for the network users of tomorrow.

---

# Signalling System

## No. 7

---

The ITU-T signalling system number 7, SS number 7, SS7, CCITT7, C7 or number seven signalling system is the most recently developed of telephone network signalling systems. It is already widely deployed in digital telephone networks and ISDNs across the world, and will also be a 'cornerstone' of 'intelligent networks' and broadband ISDNs (B-ISDN). It is a complex, common channel signalling system, which enables the controlling processors of two digital exchanges or databases to communicate directly and interact with one another in a manner optimized for digital transmission media. SS7 has also formed the basis of a number of further-developed regional signalling systems. In the United States, for example, 'ANSI SS7' is a derivative, while the UK national version is 'C7/BT'. This chapter describes the overall structure and capabilities of SS7.

### 12.1 SS7 SIGNALLING BETWEEN EXCHANGES

The SS7 signalling system is described in the Q.700 series of ITU-T recommendations. A *common channel signalling system*, optimized for digital networks, it allows direct transfer of call information transfer between exchange processors. Comprising a number of layered and modular parts, each with a different function, it is a powerful general-purpose signalling system capable of supporting a range of applications and administrative functions, including

- *ISDN (integrated services digital network)*
- *intelligent networks (INs)*
- *mobile services (e.g. cellular radio)*
- *network administration, operation and management*



In addition, its modular nature lends itself to the development of new *user parts* which may be designed to support almost any new service that can be conceived. The *user parts* of the system that have been developed so far are

- MTP      message transfer part
- SCCP    signalling connection and control part
- TUP      telephone user part
- DUP      data user part
- ISUP    ISDN services user part
- TC       transaction capabilities (used by *intelligent networks*)
- TCAP    transaction capabilities application part
- OMAP    operation and maintenance application part
- INAP    intelligent network application part
- MAP      mobile application part

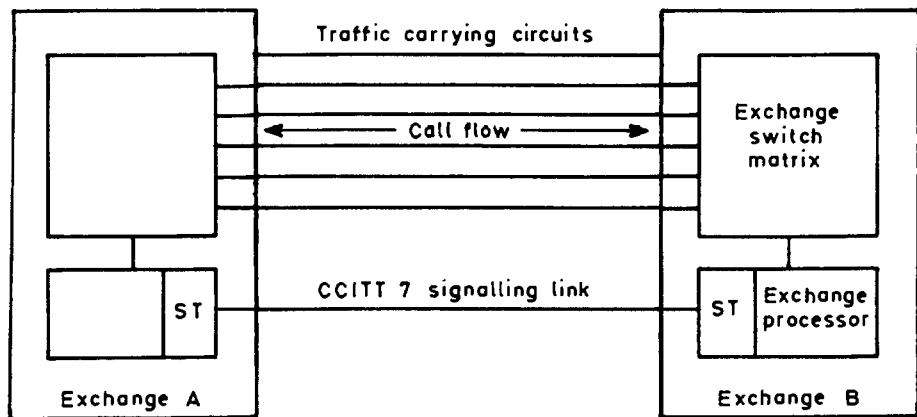
The MTP and SCCP form the ‘foundations’ of the system, providing for carriage of messages. The TUP, DUP and ISUP use the MTP and/or SCCP to convey messages relating to call control, for telephone, data, and ISDN networks, respectively. The OMAP, MAP and INAP are other *application parts* for operation and maintenance interaction, mobile network control and intelligent network services, respectively.

Initially the SS7 system was designed so that the MTP could be used in association with any or all of the telephone, data and ISDN *user parts*. However, following the emergence of the OSI model, the SCCP was developed as an adjunct to the MTP; the two in combination provide the functions of the OSI network service (layers 1–3).

SS7 signalling can be installed between two exchanges, provided that the necessary signalling functions are available in both exchanges. The functions reside in a unit termed a *signalling point*. This may be a separate piece of hardware to the exchange, but usually it is a software function in the exchange central processor. SS7 *signalling points* (*SPs*), basically exchanges, intercommunicate via *signalling links* and are said to share a *signalling relation*.

A single SS7 *signalling link* enables information to be passed directly between two exchange processors, allowing the set-up, control, and release of not just one, but a large number of traffic-carrying circuits between the exchanges. Messages over the unit take the form ‘connect circuit number 37 to the called customer number 01-234 5678’. The term *common channel signalling* aptly describes this method of operation, distinguishing it from the *channel-associated* signalling method, wherein call set-up signals pertinent to a particular circuit are sent down that circuit. SS7 is not the first common channel signalling system to be developed; CCITT 6 (SS6) was also a common channel system, but CCITT 6 was less flexible than SS7 and not so suitable for digital network use.

Having a *common channel* for conveyance of signalling messages saves equipment at both exchanges, because only one ‘sender’ and one ‘receiver’ is required at each end of the link, as against the one per circuit required with channel-associated systems. The



ST = signalling terminal

Figure 12.1 Linking two exchanges using SS7 signalling

combination of a SS7 sender and receiver is normally referred to as a *signalling terminal*. In practice signalling terminals are a combination of a software function in the exchange central processor and some hardware to terminate the line and undertake the basic bit transfer function (OSI layer 2, datalink).

A label attached to each message as it passes over the signalling link enables the receiving signalling point to know which of the many circuits it relates to. Figure 12.1 illustrates the network configuration of a simple SS7 signalling link. It shows calls flowing over a large number of traffic-carrying circuits which are connected to the switch matrix part of the exchange. Meanwhile all these circuits are controlled according to the information passed directly between the exchange processors. The *signalling terminal* (ST) function is shown residing within the exchange processor.

## 12.2 SS7 SIGNALLING NETWORKS

Networks employing SS7 signalling comprise two separate subnetworks. One subnetwork is the network of traffic-carrying circuits interconnecting the exchanges. The second subnetwork is that of the *signalling links*. In Figure 12.1 we saw this separation of traffic-carrying circuits from signalling link as it would apply on a single connection between two exchanges. Figure 12.2 now shows a more complicated example to illustrate another powerful feature of SS7: the fact that signalling networks and traffic-carrying networks may be designed and implemented almost in isolation from one another. Just because there are direct traffic-carrying circuits between two exchanges (they have a direct *traffic-carrying relation*) it does not follow that the signalling information (or *signalling traffic*) has to travel over direct signalling links, though clearly a *signalling relation* of some sort is needed.

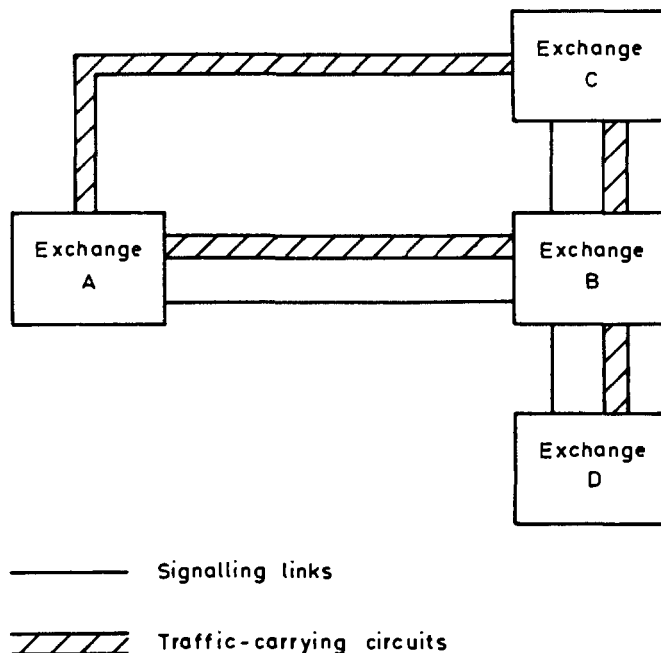


Figure 12.2 Traffic-carrying and signalling networks in SS7

Figure 12.2 shows the traffic-carrying networks and signalling networks interconnecting four exchanges, A, B, C and D. The traffic circuits directly connect A–C, A–B, B–C and B–D. All traffic to or from exchange D passes via exchange B and all traffic to or from exchange A passes either via B or C, and so on. The *signalling network*, however, is different. Signalling links only exist between A–B, B–C and B–D, so that *signalling traffic* has to be routed differently from the actual traffic. In the case of the actual traffic from A to B, there exist both direct traffic circuits and a direct signalling link. In effect, this is the same as Figure 12.1, so that both signalling messages and traffic can be passed directly between the two. Similarly exchange B may pass signalling messages and traffic directly either to exchange C or exchange D, and may also act as a normal *transit exchange* for two-link routing of traffic from exchange A to either of exchanges C or D. These are all examples of *associated mode signalling*, in which signalling links and traffic circuits have a similar configuration, and signalling messages and traffic both route in the same manner. In short, there is a signalling link associated with each link of direct traffic-carrying circuits.

By contrast, although exchange A is directly connected to exchange C by traffic-carrying circuits, there is no direct signalling link. Signalling information for these circuits must be passed on another route via exchange B. This is known as the *quasi-associated mode* of signalling, and the signalling point (SP) in exchange B is said in this instance to perform the function of a *signal transfer point (STP)*, as illustrated in Figure 12.3.

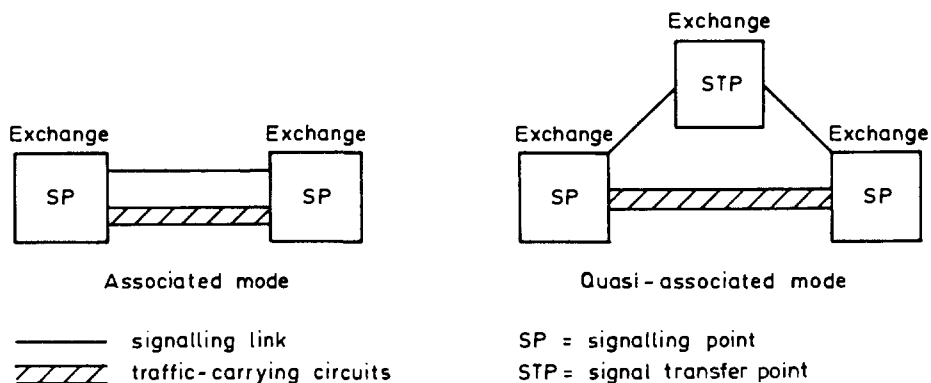


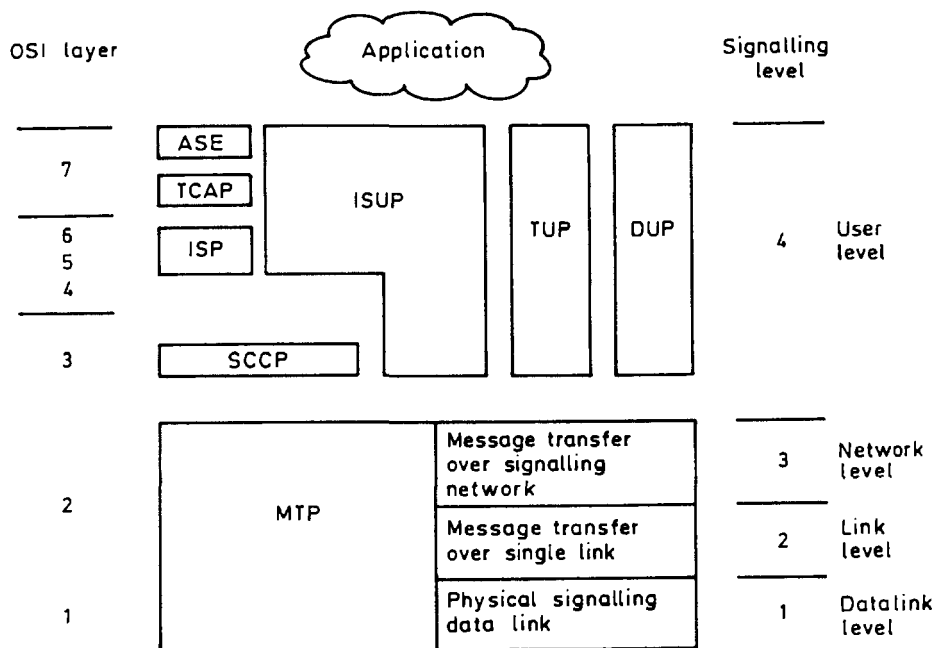
Figure 12.3 Modes of SS7 signalling

Signalling information is passed over SS7 *signalling links* in short bursts; indeed a SS7 *signalling network* is like a powerful packet-switched data network. To identify each of the signalling points for the purpose of signalling message delivery around the network, each is assigned a numerical identifier, called a *signalling point code (SPC)*. This code enables an SP to determine whether received messages are intended for it, or whether they are to be transferred (in *STP mode*) to another SP. The codes are allocated on a network by network basis. Thus the code is only unique within, say, national network A, national network B or the international network.

## 12.3 THE STRUCTURE OF SS7 SIGNALLING

Thanks to the modular manner in which the SS7 system has been designed, it encourages the development of new modules in support of future telecommunications services and functions. Figure 12.4 illustrates the functional structure of the SS7 system, relative to the layers of the Open Systems Interconnection (OSI) model (see Chapter 9).

In the same way as the OSI model has a number of functional layers, each an important foundation for the layers above it, so SS7 signalling is designed in a number of functional *levels*. Note in Figure 12.4, that the component *levels* and *parts* of SS7 do not align with the OSI layered model. The lack of alignment of *signalling levels* with *OSI layers* is unfortunate and it arises from the fact that the two models were developed concurrently but for different purposes. The lack of alignment of *levels* with *layers* means that not all higher layer OSI protocols are currently suitable for use in conjunction with the lower *levels* of SS7 signalling. The various standards development bodies are trying to rationalize the component parts of SS7 to conform with the OSI model. The *signalling connection and control part (SCCP)*, for example, delivers the OSI network service (OSI layer 3 service), so that a communication system can use the SCCP (and MTP below it) to support layers 4–7 OSI-based protocols. The levels in SS7 signalling provide a convenient separation of signalling functions, and in the remainder of the chapter the signalling level model is used in explanation.



**Figure 12.4** The structure of SS7 signalling. ASE, Application service element; TCAP, Transaction capability; ISP, Intermediate service part; ISUP, ISDN services user part; TUP, Telephone user part; DUP, Data user part; SCCP, Signalling connection and control part; MTP, Message transfer part

## 12.4 THE MESSAGE TRANSFER PART (MTP)

The foundation level of the SS7 signalling system is the message transfer part defined by ITU-T Recommendations Q.701–Q.707. The message transfer part takes care of the conveyance of messages, fulfilling signalling level functions 1 to 3 (sometimes labelled MTP1, MTP2, MTP3) as follows.

### *Level 1 (datalink functions)*

The first level defines the physical, electrical and functional requirements of the signalling data link. The level one function is attuned to the particular transmission medium as laid down by ITU-T G series recommendations. The level 1 function allows for an unstructured bit stream to be passed between SPs over an isolated signalling data link.

### *Level 2 (signalling link functions)*

This level defines the functions and procedures relating to the structure and transfer of a signal. Message flow control, and error detection and correction are included. (Flow control prevents the over-spill and consequent loss of messages that result if a message is sent when the receiving end was not ready to receive it; error detection and correction procedures eliminate message errors introduced on the link.)

### Level 3 (signalling network functions)

This level defines the functions and procedures for conveying signalling messages around an entire signalling network. It provides for the routing of messages around the signalling network. In this role it has a number of 'signalling network management' capabilities including 'load sharing' of signalling traffic between different signalling links and routes (illustrated in Figure 12.5) and re-routing around signalling link failures. Link sharing on the same route between signalling points (SPs) guards against lineplant failure (Figure 12.5(a)). Route sharing may additionally provide protection against failure of STPs. Thus in Figure 12.5 the signalling traffic from SP A to SPs B and C is shared over the two STPs, D and E. In the event of a failure of any of the routes shown, signalling messages could be re-routed.

MTP is useless on its own for setting up telephone or other connections. To perform these functions MTP needs to be used in association with one of the SS7 *user parts* which are *level 4* or *user functions*. Examples are the telephone user part (TUP) and the *integrated services digital network user part* (ISDN-UP or ISUP). These define the content and interpretation of the message, and they provide for connection control.

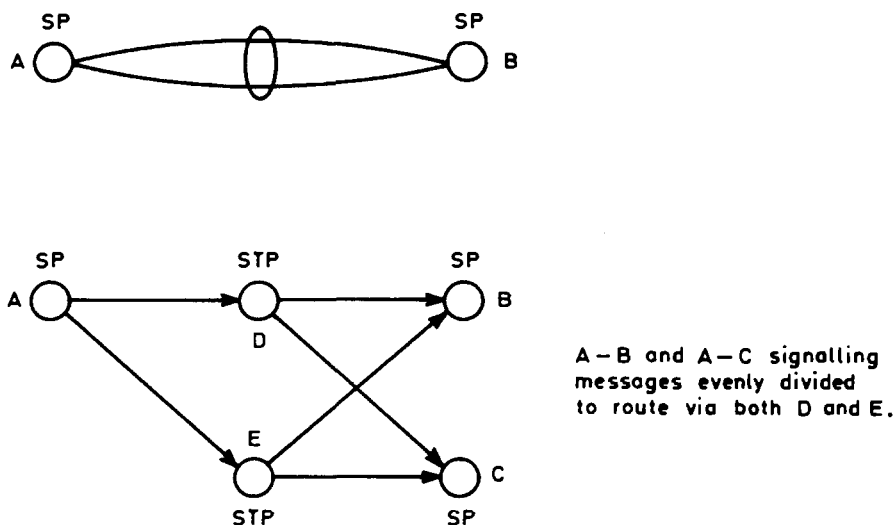
The structure of an MTP message is shown in Figure 12.6. It comprises four parts, transmitted in the following order.

### Flag

The *flag* is the first pattern of bits sent. This is an unmistakable pattern to distinguish the beginning of each message, and delimit it from the previous message. It is comparable to the synchronization (SYN) byte in data communications (Chapter 9).

### MTP information

The flag is followed by a number of *fields* of information, which together ensure the correct message transfer. These fields include: the *message sequence numbers* that keep



**Figure 12.5** Load sharing over signalling. A-B and A-C signalling messages evenly divided to route via both D and E

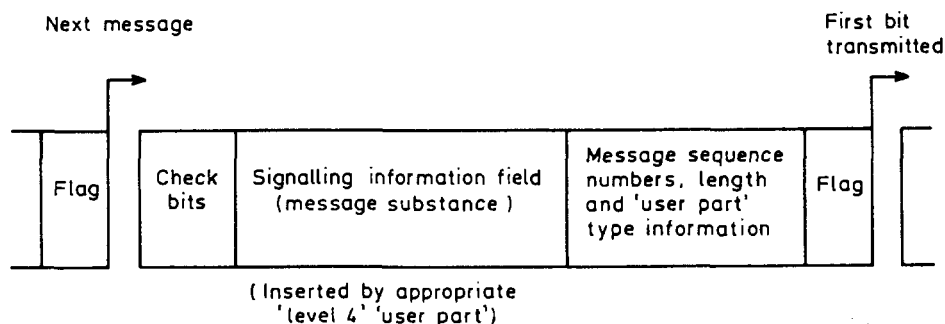


Figure 12.6 CCITT 7 MTP message structure

the messages in the correct order on receipt, and allow lost messages to be resent; and information about the type and length of the information held in the main 'signalling information field'; it might say which *user part* is in operation and record the length of the message.

### Signalling information field

This is the main information field or the 'substance' of the message. The information is inserted by one of the user parts, as appropriate for the particular application (e.g. telephone user part (TUP), or integrated services user part (ISUP)). The structure of this *field* depends on which *user part* is in use.

### Check bits

Finally, each MTP message is concluded with a *check bit* field. This is the data (*cyclic redundancy check code* or *CRC*) needed to perform the error detection and correction mechanism of the MTP level 2. The *check bits* are followed by the flag at the start of the next message.

## 12.5 THE USER PARTS OF SS7

The various user parts of SS7 are alternative functions meeting the requirements of level 4 of the signalling level model. The user parts may be used in isolation, or sometimes may be used together. Thus the *telephone user part* (TUP) and the MTP together are sufficient to provide telephone signalling between exchanges. The data user part (DUP), ISDN user part (ISUP) and other user parts need not be built into a pure telephone exchange. An example where more than one user part is employed is the combination of SCCP (signalling connection and control part), ISP (intermediate service part) and TCAP (transaction capability application part). These are all necessary for the support of the *intelligent networks* described in Chapter 11). The remainder of the chapter describes the capabilities of each of the level 4 *user parts* of SS7.

## 12.6 THE TELEPHONE USER PART (TUP)

The telephone user part comprises all the signalling messages needed in a telephone network to set up telephone calls (we described the sequence of call set-up in Chapter 7). Thus an exchange using the SS7 signalling system carries out the normal process of digit analysis and route selection, *seizes* the outgoing circuit and sends the dialled digit train onto the next exchange in the connection by using the SS7 signalling link, conveying TUP encoded messages using the MTP. Crudely put, an example of a TUP message might be 'connect the call on circuit number 56 to the destination directory number 071-234 5678'. Backward messages such as 'destination busy' are also included in the telephone user part.

The structure of TUP messages is shown in Figure 12.7. TUP messages occupy the *signalling information field* of the underlying MTP message. The messages comprise a TUP signalling information field which is used to convey 'dialled digits', 'line busy', 'answer' signals, and other circuit-related information, together with four administrative fields as follows.

### *Destination point code (DPC)*

This code identifies the signalling point to which the signalling message is to be delivered by the MTP. (The destination of a signalling message is not necessarily the same as the final destination of the call.) The signalling point is in the exchange that forms the next link of the connection (for *forward messages*) or in the previous exchange (for *backward messages*).

### *Originating point code (OPC)*

This code identifies the signalling point which originated the message (again not necessarily the origination point of the call).

### *Circuit identification code (CIC)*

This is a number that indicates to the exchange at the receiving end of the signalling link which traffic circuit each message relates to.

The telephony user part is defined in ITU-T Recs. Q.721–Q.725.

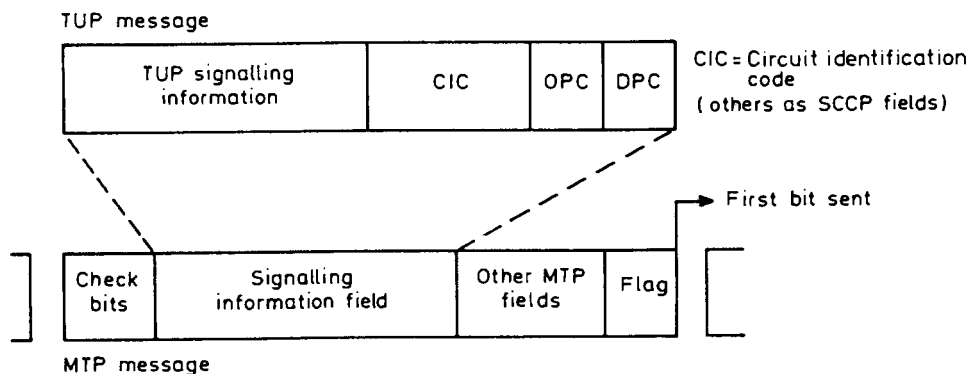


Figure 12.7 TUP message structure and relation to MTP



## 12.7 THE DATA USER PART (DUP)

The Data user part is similar to the telephone user part, but it is optimized for use on circuit-switched data networks. The message structure of the DUP is very similar to that of the TUP, illustrated in Figure 12.7. The DUP is defined by ITU-T recommendations Q.741 but was hardly ever used. It has been largely superseded by the ISUP.

## 12.8 THE INTEGRATED SERVICES USER PART (ISUP)

Used in conjunction with the MTP, the SS7 *integrated services digital network user part*, *ISDN-UP* or *ISUP*, is the signalling system designed for use in ISDNs. In effect it is a combination of capabilities similar to TUP and DUP, which allow voice and data switched services to be integrated within a single network. The message structure is similar to that of TUP and DUP, but the messages used are incompatible with both of the other systems. ISUP is defined by CCITT Rec Q.761–Q.764.

The ISDN user part (ISUP) interacts as necessary with the ISDN D-channel, signalling (*DSS1*, *digital subscriber signalling 1*, as defined by recommendation Q.931) to convey end-to-end information between ISDN user terminals. Such information includes the *terminal compatibility* checking procedure which ensures that a compatible receiving terminal is available at the location dialled by the caller. As we learned in Chapter 10, the procedure prevents, for example, the connection of a group 4 facsimile machine to a videoconference at the receiving end.

## 12.9 THE ENHANCED TELEPHONE USER PART (TUP+)

The TUP+ is an enhanced version of the TUP, though incompatible with it. It was developed by CEPT as recommendation T/SPS 43-02 for use as an interim ISDN-like signalling system supporting an early pan-European ISDN. It is used in Europe by France Telecom for international ISDN signalling, but is likely to be superseded by ISUP.

## 12.10 THE SIGNALLING CONNECTION CONTROL PART (SCCP)

The SCCP is used to convey *non-circuit-related* information between exchanges or databases, between an exchange and a database or between two exchanges (for certain types of ISDN supplementary services). By *non-circuit-related* we mean that although a signalling relation is established between an exchange and a database, no traffic circuit is intended to be set up. In essence the SCCP (in conjunction with the TC and relevant *application part*) provides a means for *querying* a reference store of information, as is necessary during call set-up on *intelligent networks*. It is an ideal data transfer mechanism for

- interrogation of a central database (Chapter 11)
- updating cellular radio ‘location registers’ (Chapters 11 and 15)
- remote activation and control of services or exchanges
- data transfer between network management or network administration and control centres

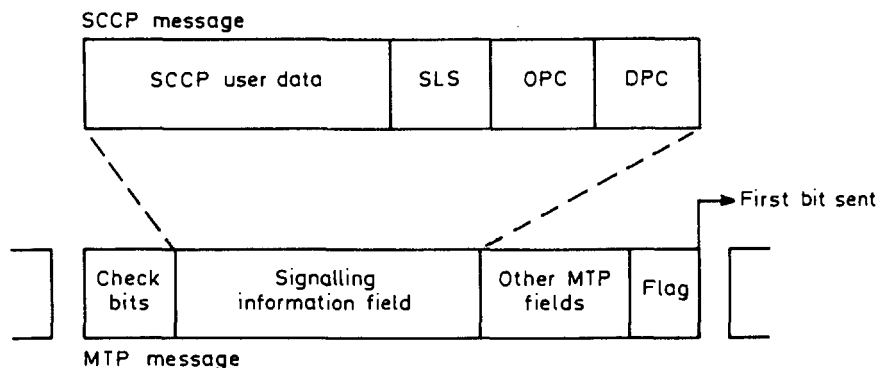
The SCCP is a *user part* which in conjunction with the MTP allows a SS7 signalling network to conform to the OSI network service (OSI layers 1–3), and to support protocols designed according to layers 4–7 of the OSI model. Most importantly, this allows new user parts to conform with the OSI model.

SCCP controls the type of connection made available between the two signalling points in the exchange and the database. In effect it establishes the *signalling relation* in preparation for one of the higher level *application parts*. Four classes of *transfer service* can be made available as defined by the OSI model. These can be grouped into *connectionless* and *connection-oriented* types, as we learned in Chapters 1 and 7. These are shown in Table 12.1.

In the context of the SCCP classes shown in Table 12.1, *connection-oriented* data transfer (classes 2 and 3) is that in which an association is established between sender and receiver before the data are sent. In reality this means the establishment of a *virtual* or *packet-switched* connection between the ends (as we discussed in Chapter 9), and not a *circuit-switched* connection. Thus before the application part signalling commences operation over the SS7 signalling link, a *virtual connection* is created by the SCCP. In the alternative *connectionless mode* of data transfer (SCCP classes 0 and 1), messages are despatched onto the SS7 signalling network without first ensuring that the recipient is ready to receive them. *Connection-oriented* procedures are useful when a large amount of data needs to be transferred. The *connectionless mode* is better suited to small and short messages, because it avoids the burden of extra messages to establish the *connection*. By their nature, connectionless messages always include address information.

**Table 12.1** The classes of SCCP

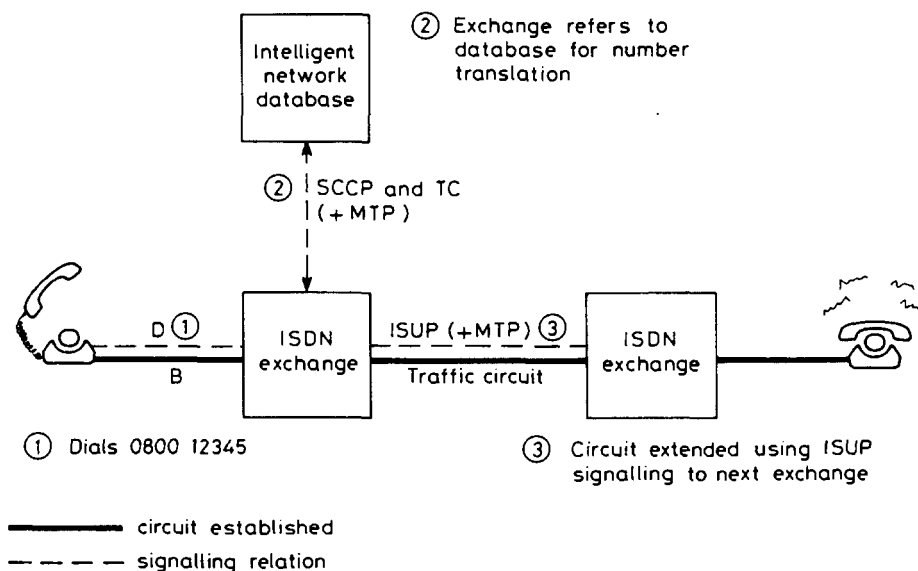
Class	Type
Class 0	Connectionless Message sequence not guaranteed
Class 1	Connectionless Message sequence guaranteed
Class 2	Basic connection-oriented Message segmentation and reassembly
Class 3	Connection oriented Message segmentation and reassembly Flow control Detection of message loss and mis-sequence



**Figure 12.8** SCCP message structure and relation to MTP. SLS, signalling link selection; OPC, originating point code; DPC, destination point code

The structure of SCCP messages is similar to that of the TUP, as we can see from Figure 12.8, except that the SCCP includes no *circuit identification code (CIC)*. The CIC is superfluous because no circuit will be established.

Figure 12.9 shows an example of the use of SCCP and TC for a database query during the call set-up phase of a Freephone (800) call. The caller (who happens to be an ISDN subscriber) dials the 0800 number, which is conveyed to the ISDN exchange by the D-channel signalling protocol, DSS1 (Q.931). Following analysis of the number, the ISDN exchange realizes that it must refer to the intelligent network databases for a number translation. It does so using the SCCP and TC. Meanwhile the circuit set-up is



**Figure 12.9** A database query using SCCP and TC

suspended. When the database interaction is over and the ISDN exchange has the appropriate information, the circuit can be connected using the standard ISUP signalling.

The SCCP is defined by ITU-T Recs Q.711–Q.714.

## 12.11 TRANSACTION CAPABILITIES (TC)

Building on the foundation of MTP and SCCP the *transaction capabilities (TC)* are that part of the SS7 signalling system which conveys *non-circuit-related information*. Its development has been intertwined with the development of *intelligent networks* (Chapter 11). The *transaction capability* is ideally suited to supervising short ‘ping-pong’ style dialogue between signalling points, typically between an exchange and an *intelligent network* database. *Transaction capabilities (TC)* break down into three component parts, and undertake the functions of OSI layers 4–7. The underlying foundation (OSI layers 1–3) is the SCCP and MTP. The *intermediate service part (ISP)*

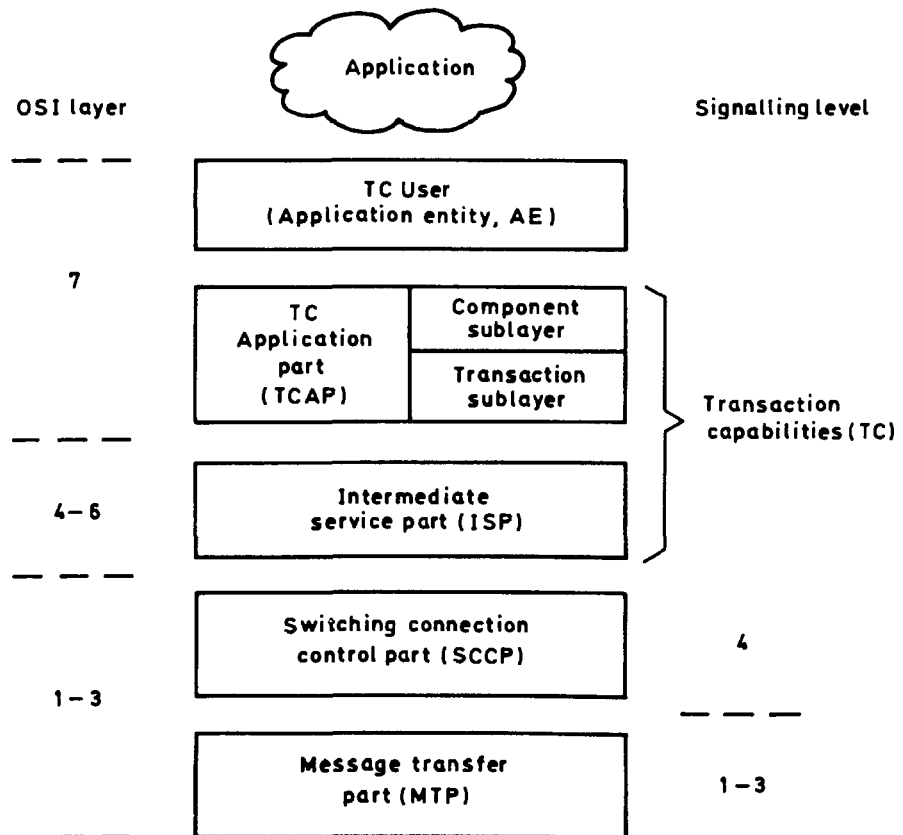


Figure 12.10 The transaction capabilities

carries out the functions of OSI layers 4–6, while the *component sublayer* and *transaction sublayer* exist within OSI layer 7 and together form the *transaction capabilities application part (TCAP)*.

*Transaction capabilities* exist to serve a *TC-user*, normally called an *application entity (AE)*. An AE contains the necessary functions to serve a particular application. In addition, every application entity also contains the transaction capabilities application part (TCAP). TCAP is, in essence, a copy of the ‘rules’ which enable the messages to be interpreted. (For example an AE may support VPN service. Another might support *freephone*.) Figure 12.10 illustrates the architecture of TC.

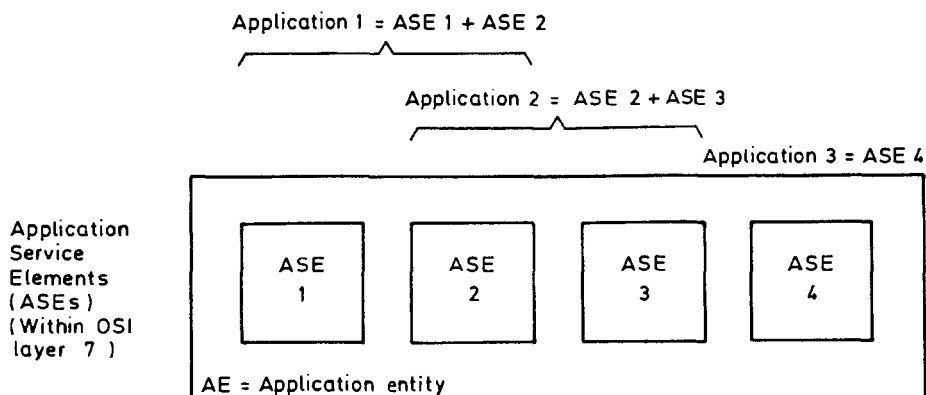
The ISP is required only when large amounts of data are to be transferred, using one of the SCCP *connection-oriented* classes.

The TCAP controls the dialogue between the exchange and the database, overseeing requests (questions or instructions) and making sure that corresponding responses are generated. The content of the request (the question itself) is prepared by the TC user (i.e. the *application entity (AE)*).

An *application entity* is the logical set of questions, responses and instructions which constitute the dialogue necessary to support an application. An AE comprises one or a number of simple functions, called *application service elements (ASEs)*. The idea is that a small set of multi-purpose ASEs can be combined together in different permutations to serve different applications. The intelligent network (IN) architecture, for example, defines a set of primitive network actions which it calls *functional components* or *service independent building blocks*. These are examples of ASEs. Figure 12.11 illustrates the concept of *application service elements*.

Three well known *application entities* defined by ITU-T are the *mobile application part (MAP)*, used to support *roaming* mobile telephone networks, the *operation and maintenance application part (OMAP)*, used for control and maintenance of remote equipment and exchanges and the *intelligent network application part (INAP)*, used in intelligent networks. MAP comprises one complex ASE, OMAP comprises two: MRVT and SRVT (the MTP and SCCP Routing Verification Tests).

Returning to the TCAP itself, let us briefly describe the functions of the *transaction* and *component* sublayers. The *transaction sublayer* is responsible for initiating,



**Figure 12.11** Permutation of ASEs to serve different applications

maintaining and closing the dialogue between the signalling points. Classification of messages into one of the four types listed below helps the two end signalling point devices to relate each message back to the previous dialogue and to check that the communication is occurring in an orderly fashion. Thus each message also has a transaction identity code.

- *Begin (dialogue or transaction)*
- *Continue*
- *End*
- *Abort*

The *component sublayer* provides machine discipline to the dialogue, controlling the *invocation* of requests and making sure that they receive proper responses. The ASE information within the requests and responses is thus classified into one of five types

- invoke an action
- return the final response (to a sequence)
- return an intermediate (but not final) response
- return a message to signal an error
- reject a message (if a request is not understood, or is out of sequence)

Although the information content of the requests and responses is not known by the component layer (it is understood only by the ASE or ASEs), the component sublayer is able to make sure that commands are undertaken and responses are given.

The transaction capabilities are defined in ITU-T Recs Q.771–Q.775.

## 12.12 THE MOBILE APPLICATION PART (MAP)

The *mobile application part* is an example of an *application entity* of SS7 signalling, developed to serve a particular application. It is used between a mobile telephone network exchange and an *intelligent network* database, called a *home (HLR)* or *visitor location register (VLR)*. The database is kept informed of the current location of the mobile telephone handset. Thus the mobile telephone customer's incoming and outgoing calls can be handled at any time. Chapter 15 on cellular telephone networks describes this application more fully.

## 12.13 OPERATION AND MAINTENANCE APPLICATION PART (OMAP)

OMAP is another *application entity* of SS7 signalling. It provides for network maintenance as well as other network operations and management functions of remote exchanges and equipment. OMAP contains 2 ASEs: the *MTP routing verification test*

(*MVRT*) and the *SCCP routing verification test (SRVT)*. These are procedures designed to enable the network operator to test the integrity of the signalling networks and identify faults.

## 12.14 INTELLIGENT NETWORK APPLICATION PART (INAP)

The *intelligent network application part (INAP)* comprises a set of application functions or *service building blocks* from which complex intelligent network services can be built. The initial functionality defined by the best standardized INAP (that of ETSI) comprises functions defined in its *capability set 1 (CSI)*. These are a range of relatively simple *intelligent network services*, but will be very important because they will be standardized across many different manufacturers' switch and service control point equipment.

## 12.15 THE USE AND EVOLUTION OF CCITT7 SIGNALLING

SS7 is an adaptable and continuously evolving signalling system that has been designed to meet the challenging and ever-changing service needs of public and major private network exchanges making up the ISDN, the B-ISDN and the *intelligent network*. Network operators may choose to implement the subset of user parts which most matches their needs, adopting new user parts as they become available.

Depending on their particular circumstances, some network operators may choose to implement an adapted version of some of the SS7 standards, taking up some of the permitted signalling *options*. The options allow the operator to 'tailor' the system to particular national requirements (e.g. C7/BT is the UK national version of TUP/ISUP and T1-ISUP is the version of ISUP used between public networks in the United States). Because of the considerable capital investment already committed to SS7, there is a pressure for use of a common system, and there is pressure also from established SS7 users to ensure that new developments are *backward compatible* with previous versions of the system.

One of the ways in which backward compatibility is ensured is by building into all SS7 implementations a mechanism for handling unrecognized information. Such information is bound to be sent occasionally from a more advanced exchange to an exchange with an older version of SS7. The common methods for dealing with this information are

- to discard it
- to ignore it
- to assume that some other *expected* response (or *default*) was actually received
- to reply with a message of 'confusion'
- to terminate the call and reset
- to raise an alarm to a human

Such a *backward compatibility* mechanism obviates any need to stop the development of SS7, or the extension of its services. If backward compatibility is not built into any new signalling standard then joint discussions between the operators interconnected networks using different versions will be necessary to agree amendments allowing compatible operation.

## 12.16 SIGNALLING NETWORK PLANNING AND TESTING

The signalling links of a SS7 *signalling network* need careful planning and implementation just like any other data network. The links need to be sufficient in number to handle the overall signalling traffic demand, and to be oriented in a topology that gives good resilience to network failures.

Two possible topologies are illustrated in Figure 12.12. Figure 12.12(a) shows a meshed network in which exchanges are capable of both SP and STP functions and rely on one another for the resilience of their signalling relations. In contrast, Figure 12.12(b) shows a topology commonly used in North America, where dedicated and duplicated computers perform the STP function alone; the exchanges are not capable of the STP function.

Because of the very complex nature of SS7 signalling, and because of the heavy network reliance on it, it is normal to undertake a comprehensive validation testing programme prior to the introduction of each new link and exchange. Exchanges built

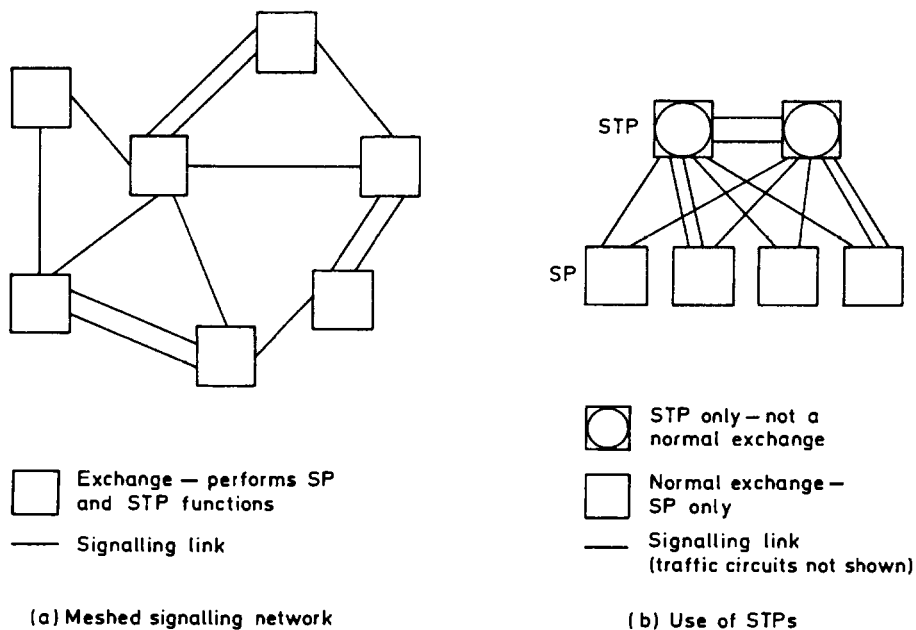


Figure 12.12 Typical SS7 signalling networks



by different manufacturers can sometimes be incompatible at first, and a testing programme is invaluable in ensuring that their problems are ironed-out before being brought into service.

## 12.17 INTERCONNECTION OF SS7 NETWORKS

Within a network owned and operated by a single network operator, the SS7 signalling system (or a variant of it) will make for close control and monitoring of the network and highly efficient call routing. However, when networks belonging to different network operators are connected together, neither operator is likely to want the other to have full control of his network.

In an international network, an operator in one country is unlikely to let the operator in another control his network management and routing rearrangements.

Competing network operators in a single country may have their networks interconnected but each guards the monitoring and control of his own network fiercely.

Public network operators may allow direct SS7 signalling from company private exchanges but they are unlikely to give up control of the network.

For these reasons, a number of options are permitted within the signalling system specifications. These allow operators by mutual agreement to restrict the capability of the signalling system when it is used for network interconnection. Thus a period of negotiation is necessary prior to the interconnection of networks. A mutual testing period confirms that the options have been selected correctly.

---

## ***Synchronous Digital Hierarchy (SDH) and Synchronous Optical Network (SONET)***

---

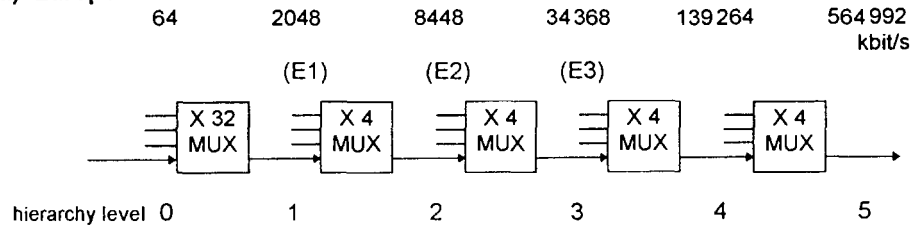
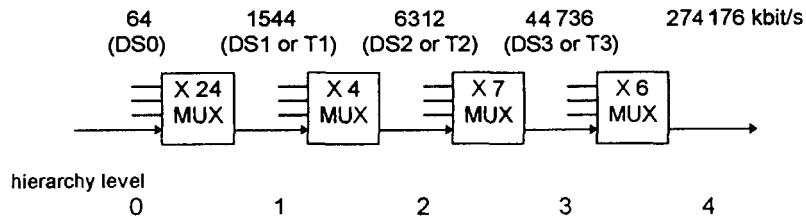
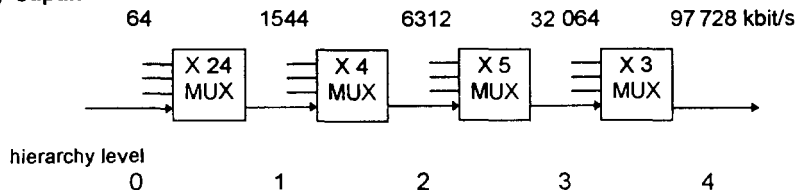
The synchronous digital hierarchy (SDH) is emerging as the universal technology for transmission in telecommunications networks. Since the first publication of international standards by ITU-T in 1989, SDH equipment has been rapidly developed and deployed across the world, and is rapidly taking over from its predecessor, the Plesiochronous Digital Hierarchy (PDH). This chapter describes SDH and the North American equivalent of SDH, SONET (Synchronous Optical Network), from which it grew. In particular, the chapter describes the features of SDH which characterize its advantages over PDH.

### **13.1 HISTORY OF THE SYNCHRONOUS DIGITAL HIERARCHY (SDH)**

The *synchronous digital hierarchy (SDH)* was developed from its North American forerunner *SONET (synchronous optical network)*. SDH is the most modern type of transmission technology, and as its name suggests it is based on a synchronous multiplexing technology. The fact that SDH is synchronous adds greatly to the efficiency of the transmission network, and makes the network much easier to manage.

### **13.2 THE PROBLEMS OF PDH TRANSMISSION**

Historically, digital telephone networks, modern data networks and the transmission infrastructures serving them have been based on a technology called *PDH (Plesiochronous Digital Hierarchy)* as we discussed in Chapter 5. As we also discussed, three distinct PDH hierarchies evolved, as we summarize in Figure 13.1. They share three common attributes.

**(a) Europe****(b) North America****(c) Japan****Figure 13.1** The various plesiochronous multiplexing hierarchies (ITU-T/G.571)

They are all based on the needs of telephone networks, i.e. offering integral multiples of 64 kbit/s channels, synchronized to some extent at the first multiplexing level (1.5 Mbit/s or 2 Mbit/s).

They require multiple multiplexing stages to reach the higher bitrates, and are therefore difficult to manage and to measure and monitor performance, and relatively expensive to operate.

They are basically incompatible with one another.

Each individual transmission line within a PDH network runs *plesiochronously*. This means that it runs on a clock speed which is nominally identical to all the other line systems in the same operator's network but is not locked *synchronously* in step (it is *free-running* as we discussed in Chapter 5). This results in certain practical problems. Over a relatively long period of time (say one day) one line system may deliver two or three bits more or less than another. If the system running slightly faster is delivering bits for the second (slightly slower) system then a problem arises with the accumulating extra bits.

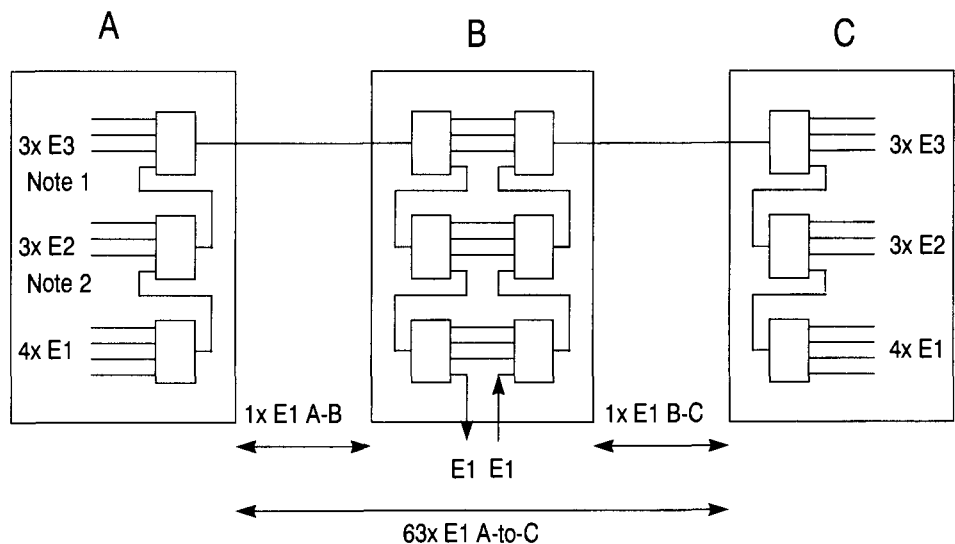
Eventually, the number of accumulated bits becomes too great for the storage available for them, and some must be thrown away. The occurrence is termed *slip*. To keep this problem in hand, *framing* and *stuffing* (or *justification*) bits are added within the

normal multiplexing process, and are used to compensate. These bits help the two end systems to communicate with one another, speeding up or slowing down as necessary to keep better in step with one another. The extra framing bits account for the difference, for example, between  $4 \times 2048(\text{E1 bitrate}) = 8192 \text{ kbit/s}$  and the actual E2 bitrate (8448 kbit/s, see Figure 13.1).

Extra framing bits are added at each stage of the PDH multiplexing process. Unfortunately this means that the efficiency of the higher order line systems (e.g. 139 264 kbit/s, usually termed 140 Mbit/s systems) are relatively low (91%). More critically still, the framing bits added at each stage make it very difficult to break out a single 2 Mbit/s *tributary* from a 140 Mbit/s line system without complete demultiplexing (Figure 13.2). This makes PDH networks expensive, rather inflexible and difficult to manage.

SDH, in contrast with PDH, requires the synchronization of all the links within a network. It uses a multiplexing technique which has been specifically designed to allow for the *drop and insert* of the individual *tributaries* within a high speed bit rate. Thus, for example, a single *drop and insert multiplexor* is required to break out a single 2 Mbit/s *tributary* from an *STM-1* (*synchronous transport module*) of 155 520 kbit/s (Figure 13.3).

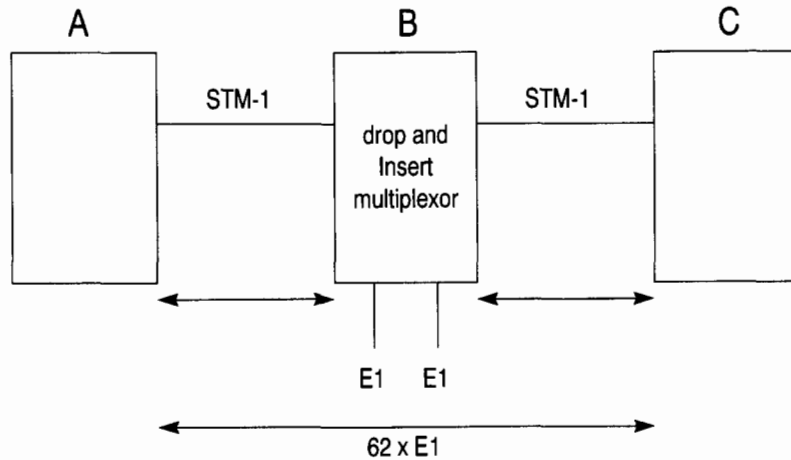
Other major problems of the PDH are the lack of tools for network performance management and measurement now expected by most public and corporate network managers, the relatively poor availability and range of high speed bit rates and the inflexibility of options for line system back-up (Figure 13.4).



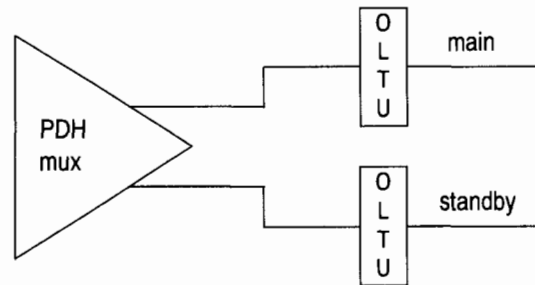
Note 1:  $3 \times \text{E3} = 12 \times \text{E2}$  or  $48 \times \text{E1}$  after demultiplexing

Note 2:  $3 \times \text{E2} = 12 \times \text{E1}$  after demultiplexing

**Figure 13.2** Breaking-out 2 Mbit/s (E1) from a 140 Mbit/s line system at an intermediate exchange



**Figure 13.3** Drop and insert multiplexor used to break-out 2 Mbit/s (E1) from a 155 Mbit/s (STM-1) line at an intermediate exchange



**Figure 13.4** Optical fibre back-up using PDH system components

Before SDH, networks had to be built up from separate multiplex and *line terminating equipment (LTE)*, as the optical equipment interfaces in particular were manufacturer-specific (i.e. proprietary). Back-up tended to be on a *1 main + 1 standby* protection basis, making back-up schemes costly (Figure 13.4) and difficult to manage. These problems have been eliminated in the design of SDH through in-built flexibility of the bitrate hierarchy, integration of the optical units into the multiplexors, ring structure topologies and in-built performance management and diagnostic functions.

### 13.3 THE MULTIPLEXING STRUCTURE OF SDH

As is shown in Figure 13.5, the *containers* (i.e. available bitrates) of the *synchronous digital hierarchy* have been designed to correspond to the bit rates of the various PDH hierarchies. These *containers* are multiplexed together by means of *virtual containers* (abbreviated to *VCs* but are not to be confused with *virtual channels* which are also

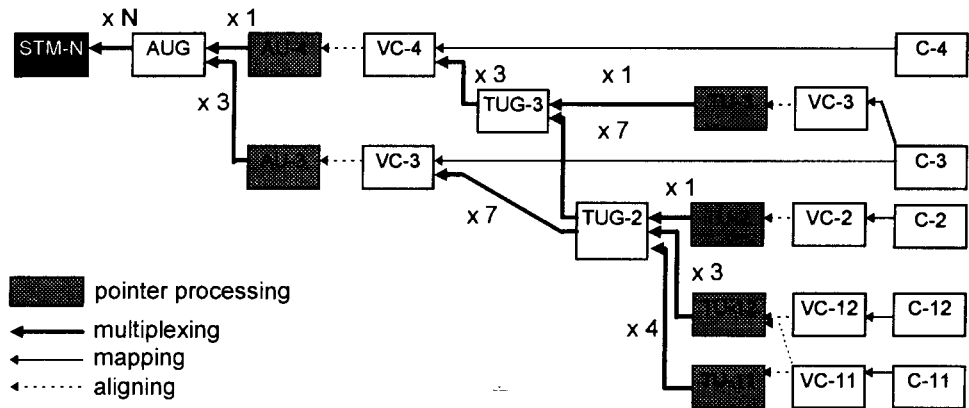


Figure 13.5 Synchronous digital hierarchy (SDH) multiplexing structure (ITU-T/G.709)

so abbreviated), *tributary units (TU)*, *tributary unit groups (TUG)*, *administrative units (AU)* and finally *administrative unit groups (AUG)* into *synchronous transport modules (STM)*.

The basic building block of the SDH hierarchy is the *administrative unit group (AUG)*. An AUG comprises one AU-4 or three AU-3s. The AU-4 is the simplest form of AUG, and for this reason we use it to explain the various terminology of SDH (*containers, virtual containers, mapping, aligning, tributary units, multiplexing, tributary unit groups*).

The *container* comprises sufficient bits to carry a full *frame* (i.e. one cycle) of user information of a given bitrate. In the case of *container 4 (C-4)* this is a field of  $260 \times 9$  bytes (i.e. 18 720 bits). In common with PDH, the *frame repetition rate* (i.e. number of cycles per second) is 8000 Hz. Thus a *C4-container* can carry a maximum user throughput rate (*information payload*) of 149.76 Mbit/s ( $18\,720 \times 8000$ ). This can either be used as a raw bandwidth or, say, could be used to transport a PDH link of 139.264 Mbit/s.

To the container is added a *path overhead (POH)* of 9 bytes (72 bits). This makes a *virtual container (VC)*. The process of adding the POH is called *mapping*. The POH information is communicated between the point of *assembly* (i.e. entry to the SDH network) and the point of *disassembly*. It enables the management of the SDH system and the monitoring of its performance.

The *virtual container* is *aligned* within an *administrative unit (AU)* (this is the key to synchronization). Any spare bits within the AU are filled with a defined filler pattern called *fixed stuff*. In addition, a *pointer* field of 9 bytes (72 bits) is added. The *pointers* (3 bytes for each VC, up to three VCs in total (9 bytes maximum)) indicate the exact position of the virtual container(s) within the AU frame. Thus in our example case, the AU-4 contains one 3 byte pointer indicating the position of the VC-4. The remaining 6 bytes of pointers are filled with an idle pattern. One AU-4 (or three AU-3s containing three pointers for the three VC-3s) are *multiplexed* to form an AUG.

To a single AUG is added  $9 \times 8$  bytes (576 bits) of *section overhead (SOH)*. This makes a single *STM-1* frame (of 19 440 bits). The SOH is added to provide for *block framing* and for the maintenance and performance information carried on a transmission line *section*

basis. (A *section* is an administratively defined point-to-point connection in the network, typically an SDH-system between two major exchange sites, between two intermediate multiplexors or simply between two regenerators). The SOH is split into 3 bytes of *RSOH* (*regenerator section overhead*) and 5 bytes of *MSOH* (*multiplex section overhead*). The RSOH is carried between, and interpreted by, SDH line system *regenerators* (devices appearing in the line to *regenerate* laser light or other signal, thereby avoiding signal degeneration). The MSOH is carried between, and interpreted by the devices assembling and disassembling the AUGs. The MOH ensures integrity of the AUG.

As the *frame repetition rate* of an STM-1 frame is 8000 Hz, the total line speed is 155.52 Mbit/s ( $19\,440 \times 8000$ ). Alternatively, power of four (1, 4, 16, etc.) multiples of AUGs may be multiplexed together with a proportionately increased section overhead,

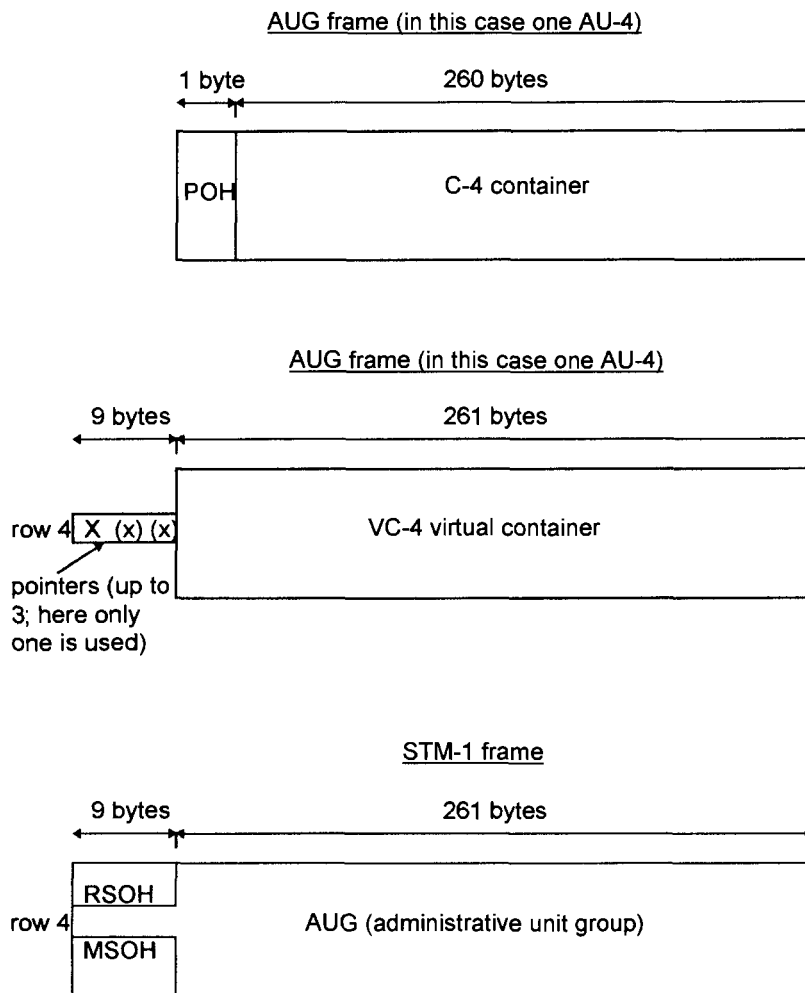


Figure 13.6 Basic structure of an STM-1 frame

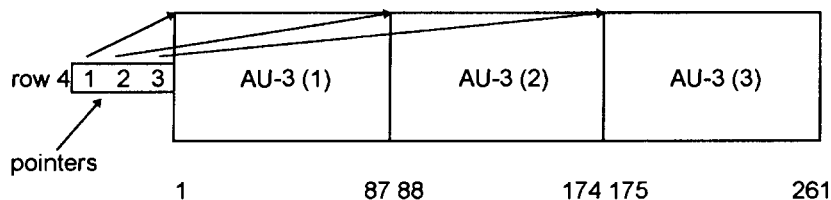
to make larger STM frames. Thus an STM-4 frame (4AUGs) has a frame size of 77 760 bits, and a line rate of 622.08 Mbit/s. An STM-16 frame (16 AUGs) has a frame size of 311 040 bits, and a line rate of 2488.32 Mbit/s.

*Tributary unit groups (TUGs)* and *tributary units (TUs)* provide for further break-down of the VC-4 or VC-3 payload into lower speed tributaries, suitable for carriage of today's T1, T3, E1 or E3 line rates (1.544 Mbit/s, 44.736 Mbit/s, 2.048 Mbit/s or 34.368 Mbit/s).

Figure 13.6 shows the gradual build up of a C-4 container into an STM-1 frame. The diagram conforms with the conventional diagrammatic representation of the STM-1 frame as a matrix of 270 columns by 9 rows of bytes. The transmission of bytes, as defined by ITU-T standards is starting at the top left hand corner, working along each row from left to right in turn, from top to bottom row. The structure is defined in ITU-T recommendations G.707, G.708 and G.709.

### 13.4 THE TRIBUTARIES OF SDH

The structure of an AUG comprising 3 AU-3s is similar to that for an AUG of one AU-4, except that the area used in Figure 13.6 for VC-4 is instead broken into 3 separate areas of 87 columns, each area carrying one VC-3 (Figure 13.7). In this case all three pointers are required to indicate the start positions within the frame of the three separate VCs. The various other TU and VC formats follow similar patterns to the AUs and VCs presented (TUs also include pointers like AUs). Table 13.1 presents the various



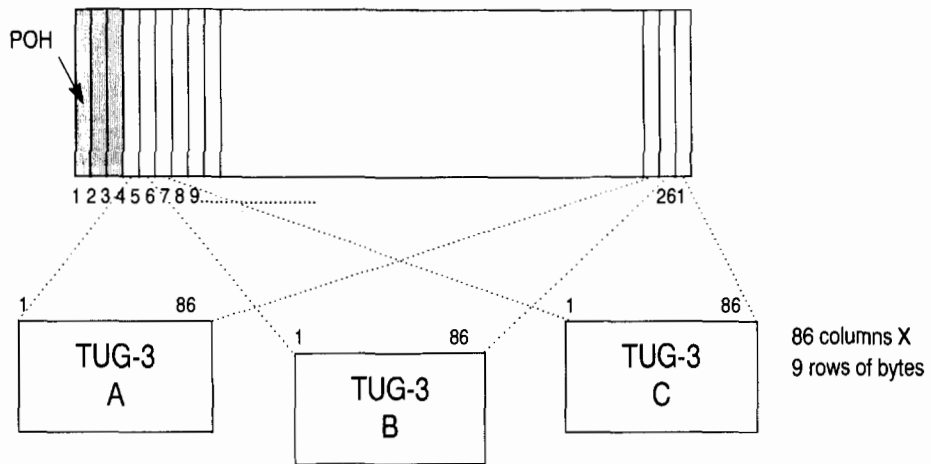
AU-3 = 87 columns x 9 rows of bytes

**Figure 13.7** AUG frame arranged as 3 x AU-3

**Table 13.1** Payload rates of SDH containers

Container type	Container frame size	Frame repetition rate	Capable of carrying PDH line type
C-11	193 bits	8000	T1 (1544 kbit/s)
C-12	256 bits	8000	E1 (2048 kbit/s)
C-21	789 bits	8000	T2 (6312 kbit/s)
C-22	1056 bits	8000	E2 (8448 kbit/s)
C-31	4296 bits	8000	E3 (34 368 kbit/s)
C-32	5592 bits	8000	T3 (44 736 kbit/s)
C-4	260 x 9 bytes	8000	139 264 kbit/s

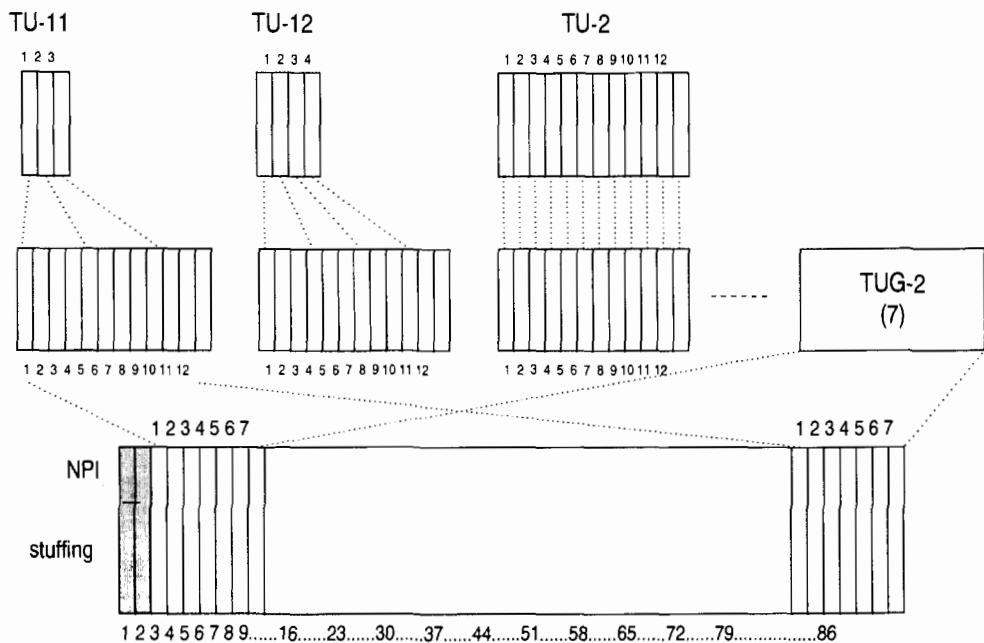




**Figure 13.8** VC-4 submultiplexing scheme as  $3 \times \text{TUG-3}$  using byte interleaving

container rates available within SDH. Note that the terminology C-12 is intended to signify the hierarchical structure and should not therefore be called C-twelve, but instead C-one-two. The relevant VC is VC-one-two, etc.

Figure 13.8 shows an alternative demultiplexing scheme, based upon the submultiplexing of a VC-4 container into three *tributary unit groups-3* (*TUG-3s*). In this case, the first three columns are used as path overhead, and each TUG occupies a total



**Figure 13.9** TUG-3 submultiplexing into  $7 \times \text{TUG-2}$ ; TUG-2 submultiplexing

of 86 columns, but the individual TUGs are *byte interleaved*. This sub-multiplexing scheme lends itself better to the carriage of PDH signals.

The sub-multiplexing of the TUG-3s themselves may be continued as shown in Figure 13.9, where each TUG-3 is sub-multiplexed into  $7 \times$  TUG-2, also using *byte interleaving*. Finally, as Figure 13.9 also shows, the TUG-2s may be subdivided into *byte interleaved* TU-11 tributaries (for T1 rate of 1.544 Mbit/s) or TU-12 tributaries (for E1 rate of 2.048 Mbit/s).

The individual *containers* (C-11 or C-12) may be packed into the TU-11 (synonymous with VC-11) or TU-12 (synonymous with VC-12) in one of three manners, using either

- no framing (i.e. *asynchronously*)
- *bit synchronous framing*
- *byte synchronous framing*

BIP-2		FEBE	unused	signal label			remote alarm
				L1	L2	L3	
1	2	3	4	5	6	7	8

a) VC-1 / VC-2 POH byte

1	2	3	4	5	6	7	8
J1							
B3							
C2							
G1							
F2							
H4							
Z3							
Z4							
Z5							

b) VC-3 / VC-4 POH (9 rows  $\times$  1 byte[8 bits])

Figure 13.10 Path overhead (POH) formats for VC-1, VC-2, VC-3 and VC-4

The *asynchronous* and *bit synchronous* framing methods allow a certain number of bits for *justification*. This enables 1.5 Mbit/s or 2 Mbit/s tributaries of an SDH transmission network to operate in conjunction with PDH or other networks running on separate clocks (i.e. not running synchronously with the SDH network; we covered the subject of *justification* in Chapter 5). *Byte synchronous framing*, in contrast, demands common clocking. The advantage is the ability to directly access 64 kbit/s subchannels within the 1.5 Mbit/s or 2 Mbit/s tributary using *drop and insert* methods (Figure 13.3). In addition, *byte synchronous* streams are simpler for the equipment to process.

### 13.5 PATH OVERHEAD

Figure 13.10 illustrates the *path overhead (POH)* formats used for creating VC-1, VC-2, VC-3 and VC-4 containers. This information is added to the corresponding *container*.

The meanings and functions of the various bits and fields are given in Table 13.2.

### 13.6 SECTION OVERHEAD (SOH)

The diagram and table of Figure 13.11 illustrate the constitution of the section overhead.

**Table 13.2** Meaning and function of the fields in the SDH path overhead (POH)

Field	Name	Function
BIP-2	bit inserted parity	error check function
FEBE	far end block error	indication of received BIP error
L1, L2, L3	signal label	indication of VC payload type
remote alarm	remote alarm	indication of receiving failure to transmitting end
J1	path trace	verification of VC-n connection
B3	BIP-8 parity code	error check function
C2	signal label	indication of VC payload type and composition
G1	path status	indication of received signal status to transmitting end
F2	path user channel	provides communication channel for network operating staff
H4	multiframe indicator	multiframe indication
Z3, Z4, Z5	bytes reserved for national network operator use	reserved

Columns (bytes)									
	1	2	3	4	5	6	7	8	9
Rows 1	A1	A1	A1	A2	A2	A2	C1		
2	B1			E1			F1		
3	D1			D2			D3		
* 4									
5	B2	B2	B2	K1			K2		
6	D4			D5			D6		
7	D7			D8			D9		
8	D10			D11			D12		
9	Z1	Z1	Z1	Z2	Z2	Z2	E2		

\* Row 4 is used for the AUG frame pointers

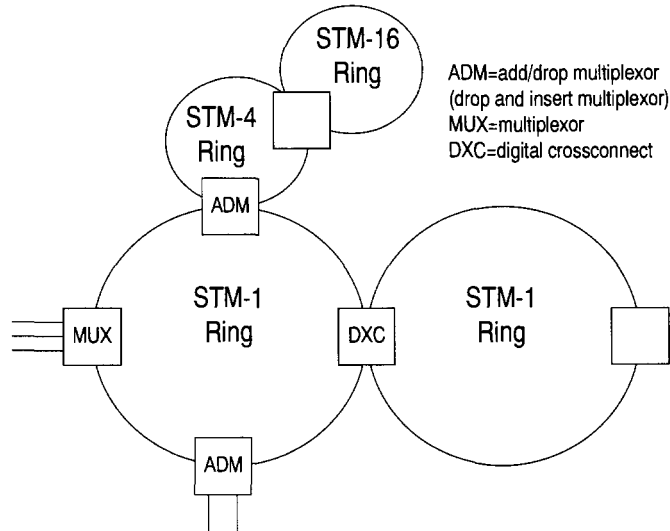
Field	Function
A1, A2	framing
B1, B2	parity check for error detection
C1	identifies STM-1 in STM-n frame
D1–D12	data communications channel (DCC – for network management use)
E1, E2	orderwire channels (voice channels for technicians)
F1	user channel
K1, K2	automatic protection switching (APS) channel
Z1, Z2	reserved

**Figure 13.11** The SDH section overhead (SOH)

## 13.7 NETWORK TOPOLOGY OF SDH NETWORKS

SDH equipment is designed to be used in the construction of synchronous (in particular, optical fibre) transmission networks in redundant ring topologies. A number of specific equipment types are foreseen by the standards as the building blocks of such networks. These are illustrated in Figure 12.12.

SDH multiplexors allow 2 Mbit/s and other sub STM-1 rate tributaries to be multiplexed for carriage by an SDH network. *Drop and insert multiplexors* (also called



**Figure 13.12** Ring topology and generic equipment types used in SDH networks

*add/drop multiplexors (ADM)* allow tributaries to be removed from the line at an intermediate station without complete demultiplexing (as we discussed in Figures 13.2 and 13.3). *Crossconnectors* (or *DXC*, *digital crossconnectors*) allow for the flexible interconnection and reconfiguration of tributaries between separate sub-networks or rings. STM-4 and STM-16 multiplexors allow concentration of STM-1 signals onto high speed 622 Mbit/s (STM-4) or 2.5 Gbit/s (STM-16) backbone networks.

Structuring of the network in interconnected rings allows for easy back-up (*restoration*) of failed connections in the network. In a highly meshed network  $n:1$  (as opposed to  $1:1$ ) restoration is possible by choosing any alternative route to the destination. In simpler networks and single rings a  $1:1$  restoration may be possible, but leaves at least 50% of the capacity unused for most of the time.  $N:1$  restoration is useful in reducing the amount of normally unused plant (and thus costs) in cases where failures are rare (see Chapter 37).

## 13.8 OPTICAL INTERFACES FOR SDH

ITU-T recommendation G.957 covers the optical interfaces defined for use with SDH, according to the light wavelength to be used and the application. A number of SDH system types are defined (Table 13.3).

## 13.9 MANAGEMENT OF SDH NETWORKS

Compared with PDH networks, SDH networks are more efficient and easier to administrate (due to the availability of *drop and insert (add/drop)* multiplexors). Using

**Table 13.3** Classification of optical fibre interfaces for SDH equipment

		Application					
		Intra-office	Inter-office				
			Shorthaul		Longhaul		
Wavelength	nominal/nm	1310	1310	1550	1310	1550	
Fibre type		G.652	G.652	G.652	G.652	G.652 G.654	G.653
Distance	kilometres	<2	15	40	40	60	
STM level	STM-1	I-1	S-1.1	S-1.2	L-1.1	L-1.2	L-1.3
	STM-4	I-4	S-4.1	S-4.2	L-4.1	L-4.2	L-4.3
	STM-16	I-16	S-16.1	S-16.2	L-16.1	L-16.2	L-16.3

a C-4 container at its full capacity (i.e. 149.76 Mbit/s) we achieve a system efficiency using SDH of 96% (c.f. 91% with PDH). However, apart from these benefits there is one other significant advantage: SDH networks are much easier to manage in operation. Partly this is due to the fact that SDH was conceived as a technology for a whole network (rather than a set of individual links); partly this is due to the fact that SDH is simply more modern, and therefore the available network management tools are more advanced.

The SDH standards, in contrast to PDH standards, set out a set of functions for monitoring and reconfiguration of remote equipment. This is achieved by the dedication of a defined management channel within the section overhead. This is referred to as the *data communication channel (DCC)* or sometimes the *embedded communication channel (ECC)*. A number of ITU-T recommendations are in course of preparation to define functions compatible with the *telecommunications management network (TMN)*, Chapter 27) which can be supported by the DCC. These will include facilities for performance monitoring (by sending a continuous given bit pattern and measuring the received signal), remote loopback and testing facilities, as well as remote configuration capability.

### 13.10 SONET (SYNCHRONOUS OPTICAL NETWORK)

SONET is the name of the North American variant of SDH. It is the forerunning technology which led to the ITU's development of SDH. The principles of SONET are very similar to those of SDH, but the terminology differs. The SONET equivalent of an SDH synchronous transfer module (STM) has one of two names, either *optical carrier (OC)* or *synchronous transport system (STS)*. The SONET equivalent of an SDH virtual container (VC) is called a *virtual tributary (VT)*. Some SDH STMs and VCs correspond exactly with SONET STS and VT equivalents. Some do not. Table 13.4 presents a comparison of the two hierarchies.

**Table 13.4** Comparison of SDH and SONET hierarchies

North American SONET	Carried Bitrate/Mbit/s	SDH
VT 1.5	1.544	VC-11
VT 2.0	2.048	VC-12
VT 3.0	3.152	—
VT 6.0	6.312	VC-21
—	8.448	VC-22
—	34.368	VC-31
—	44.736	VC-32
—	149.76	VC-4
STS-1 (OC-1)	51.84	—
STS-3 (OC-3)	155.52	STM-1
STS-6 (OC-6)	311.04	—
STS-9 (OC-9)	466.56	—
STS-12 (OC-12)	622.08	STM-4
STS-18 (OC-18)	933.12	—
STS-24 (OC-24)	1244.16	—
STS-36 (OC-36)	1866.24	—
STS-48 (OC-48)	2488.32	STM-16
STS-96 (OC-96)	4976.64	—
STS-192 (OC-192)	9953.28	STM-64

### 13.11 SDH AND ATM (ASYNCHRONOUS TRANSFER MODE)

Finally, it is worth mentioning here the integration of SDH into the specifications for *broadband-ISDN (B-ISDN)* and *ATM (asynchronous transfer mode)*. These techniques will form the basis of future *broadband networks*. The C-4 container may be used directly for carriage of ATM, and will be one of the standard speeds at which ATM will be used. ATM cells (of 53 bytes or *octets*) do not fit an integral number of times into the C-4 frame (2340 bytes), but this is not important. The SDH standards require only that the ATM *octets* are aligned with the bytes of the SDH container. Individual ATM cells can be split between container frames when necessary. We return to B-ISDN and ATM in Chapters 25 and 26.

---

## *Operator Assistance and Manual Services*

---

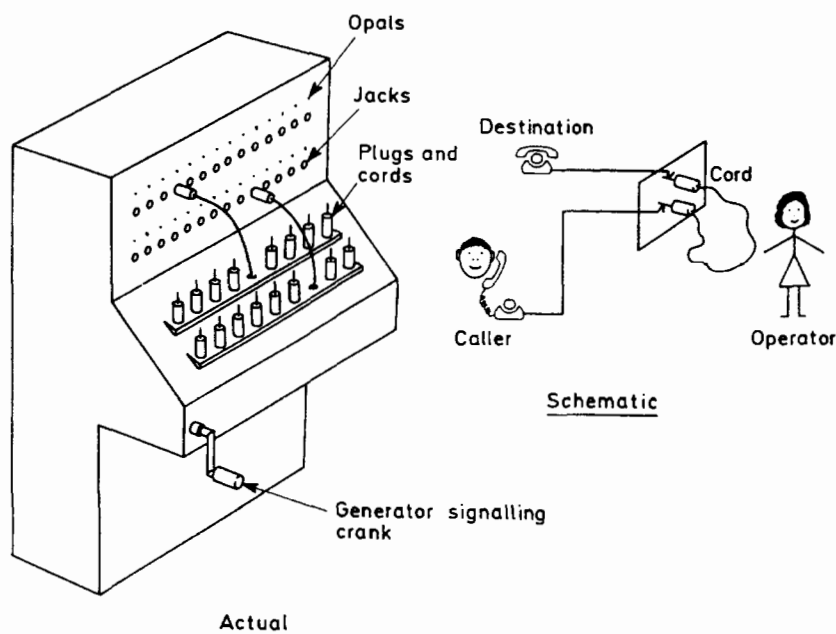
Early telephone networks were all manually operated. In the 1950s automatic networks began to take over, but even today they have failed to supplant all manual ‘assistance services’. In the public network human operators provide a ‘safety net’ of assistance and advice for customers, and in some private networks human PBX operators are still employed to answer incoming calls from the public network and to connect them to the required extension. In this chapter we discuss the operator assistance services which form a critical supplement to automatic switched services in meeting the high expectations of today’s telephone customer.

### 14.1 MANUAL NETWORK OPERATION

In a manual network, the connection of caller to destination is carried out by human operator. This is done by *plugging* cords into individual line sockets or *jacks*, one *jack* corresponding to each possible destination user. Figure 14.1 illustrates an early manual switchboard, and Figure 14.2 a typical telephone used on such a manual network. Instead of a numbered *dial* there is just a cradle for the handset and a magneto generator to call the operator.

The routine for making a call on a manual network is as follows. The caller lifts the handset, and rings the magneto generator by turning the handle. This has the effect of alerting the operator and lighting an *opal* (a light) on the operator’s switchboard (Figure 14.1). In some cases, the operator was alerted merely by rattling the cradle. This had the effect of flashing the opal. There is an opal above each incoming line jack, indicating precisely which caller wishes to make a call. To answer the request, the operator uses one of the cords mounted on the console part of the switchboard, which is pulled out and plugged into the relevant *jack* socket immediately below the *opal*. The operator is now able to speak to the caller and ask for the name of the person he wishes to call. The operator records the caller’s name, the destination number and the time of day, on a *ticket* for later *billing* of the caller. The destination party is then alerted by the operator, who rings his telephone with another hand-cranked generator. The connection





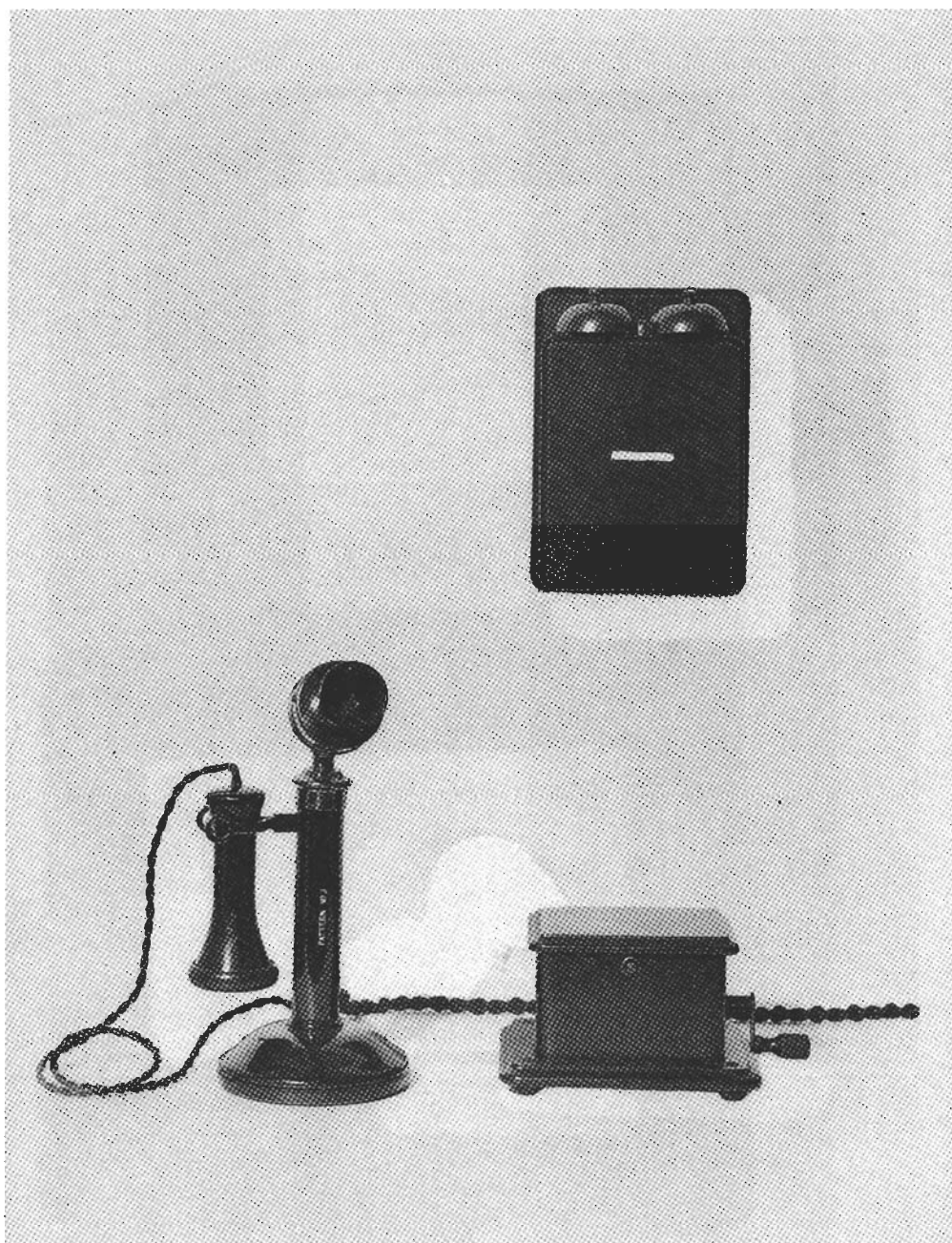
**Figure 14.1** Early manual, or 'sleeve control' switchboard

is completed by plugging the other end of the cord into the *jack* of the destination party. In this way the pair of plugs and the cord connect the two corresponding line *jacks* to caller and destination. At the end of the call the caller replaces the handset, extinguishing the *opal*. On noticing this, the operator removes the plugs and cord from the jacks, ready for use on another call.

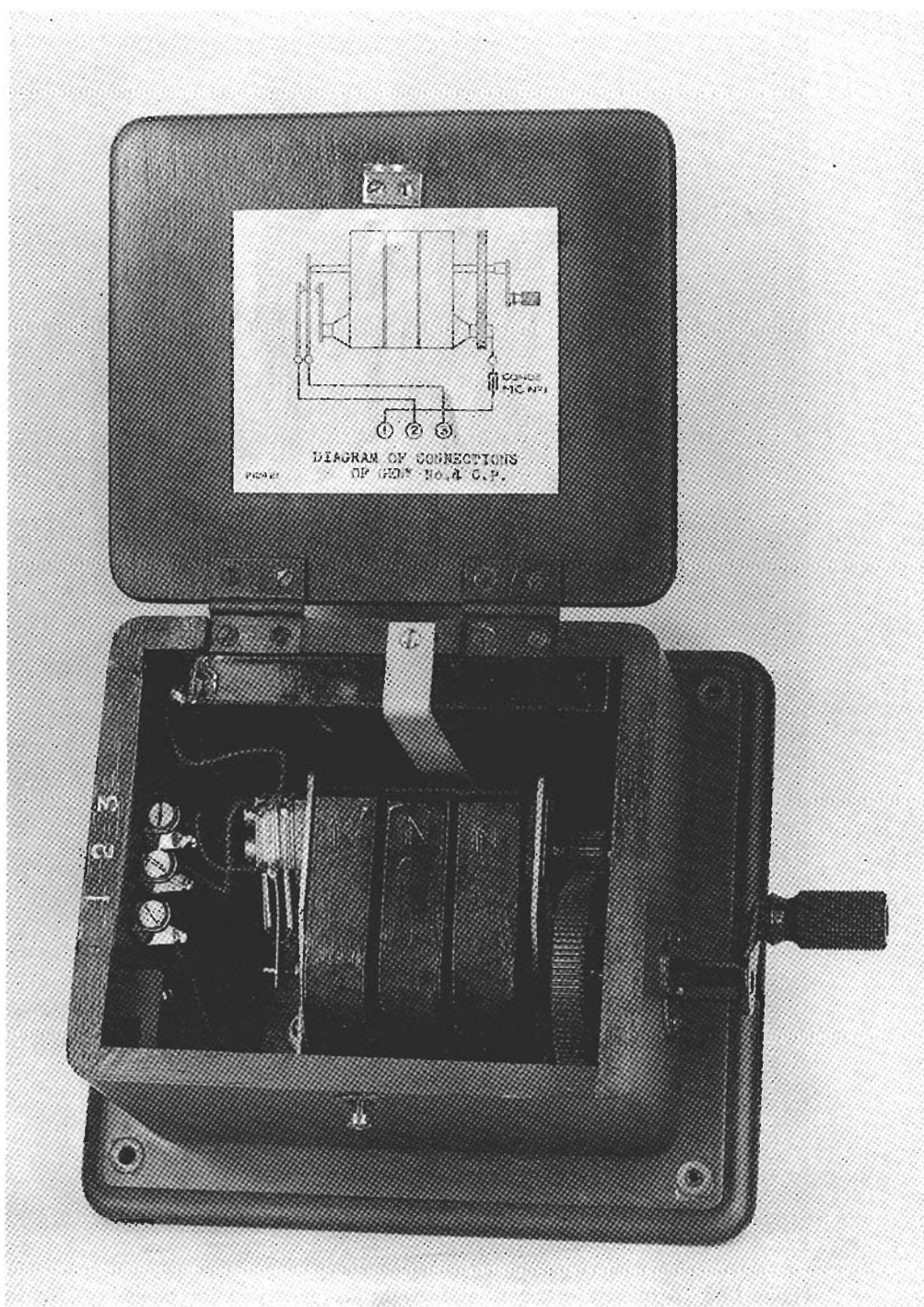
To make calls to customers on other exchanges the operator has a number of *trunk line jacks*. To use them, the operator must relay the call details to the operator on the second exchange, and forward the connection. The second operator either completes the call or forwards it to another operator, as necessary. In manual networks, setting up telephone calls is highly labour-intensive, and in the early days the majority of the workforce in public telephone companies were *telephone operators*.

## 14.2 SEMI-AUTOMATIC TELEPHONY

*Semi-automatic telephony* is the term used to describe connections which are set up by an operator across an *automatic network*. Semi-automatic telephony was common when telephone networks were first being automated, especially when some exchanges had been automated while others remained manual. Callers on the manual exchange, who wished to call others already connected to an automatic exchange, would have their calls connected *semi-automatically* by the operator. The caller would first contact the operator, and the operator would then connect the call onward using special equipment to control the automatic network rather than routing through further operators.



**Figure 14.2** A manual telephone system, the customer premises equipment. To the left is the telephone. On the wall is the bellset and to the right is the hand generator signalling set. (Courtesy of BT Archives)



**Figure 14.3** Inside of a hand generator signalling set, showing the magneto coils. (Courtesy of BT Archives)



**Figure 14.4** Early operator switchboard. (*Courtesy of British Telecom*)

## THE EXCHANGE AT WORK.

An explanation which will prevent many common misunderstandings.

The telephone lines in a large city are divided into groups, called *Exchanges*.



An Exchange has two sides, the A-side and the B-side. The A-side of an Exchange deals directly only with *Ringers up* in that Exchange; the B-side with those *rung-up* in that Exchange: e.g., Museum 605 wants Hop 3000.

4

A Museum A-side operator answers and gets into touch with a Hop B-side operator, who finds out if Hop 3000 is free and helps the Museum operator to connect the lines. If conversely Hop 3000 wants Museum 605, it is a Hop A-side operator who answers and gets into touch with a Museum B-side Operator.



If Central 6000 wants Central 3000, a Central A-side operator gets into touch with a Central B-side operator.

In every call, therefore, two Post Office operators\*

\* In reality, of course, there are four operators: two Post Office servants and two amateurs—that is, the two subscribers connected. We, in fact, are members of the unpaid Post Office telephone staff. Hence these notes.

5

**Figure 14.5** The exchange at work. An extract from an early British Post Office publication providing an explanation which will prevent many common misunderstandings. (Courtesy of BT Archives)

Ironically, in the reverse direction, any caller who was connected to an automatic exchange would have to dial a code to get hold of an operator to obtain a manual connection to any destination customer still connected to a manual exchange.

Manual exchanges have progressively given way to today's predominantly automatic networks, but even today callers resort to dialling for *assistance* from the operator in a number of instances

- to call a user on a residual manual exchange (particularly in remote overseas locations)
- to receive assistance following difficulty on an automatic connection
- to receive the answer to a general enquiry
- to enquire for the directory number of another user
- to make a call to the emergency services (fire, police or ambulance)
- to make a special service call, such as a *reverse-charge call* (also known as a *collect call*), or a *personal call*, etc.

In addition, many company switchboards and hotels still retain operators for the connection of incoming calls to individual *extensions* or hotel room numbers, despite the fact that *direct dial-in* (*DDI*, also *direct inward dialling*, *DID*) nowadays makes direct dialling possible.

### 14.3 CALLING THE OPERATOR

If all we want is to get through to the operator and tell him we need assistance, the effect of cranking the magneto is very much the same as dialling the right number on an automatic network. On a *sleeve-controlled* switchboard (one using plugs, cords and jacks, as illustrated in Figures 14.1 and 14.4), incoming calls are indicated to the



**Figure 14.6** Switchboard operators at work. A picture giving an idea of the tangle of hands and leads the frenetic operation of manual exchanges. (Courtesy of British Telecom)



operator by lighting the *opals*. In modern 'cordless' operator switchrooms the call is administered by *call queueing equipment* or *automatic call distribution (ACD)* equipment. This equipment stacks up calls in the order in which they are received, and allocates them to telephone operators in the switchroom as they become free from dealing with previous calls. Operators simply press a button to indicate that they are ready to handle another call.

While waiting in the queue for an operator to become free, the caller may hear *ringing tone*, or may instead be given a recorded message, something like 'this is the assistance service: an operator will deal with your enquiry shortly'. Recorded messages have the benefit of confirming that callers have 'got through', giving reassurance that they are not waiting in vain. The recorded message also allows callers who have accidentally dialled the number for the operator service (when meaning to dial some other number), to hang up their calls up and try again.

Because connections made via the operator are multi-link rather than single link connections, special measures are required, to ensure that the end-to-end quality of the connection is acceptable, and to charge the customer correctly. Callers are normally connected to the nearest switchroom. This ensures that the end section connection is of the best available quality. This part of the connection (i.e. from caller to operator) is not normally automatically metered for charging purposes; the call charges are derived either from paper *tickets* written by the operator, or from *electronic tickets* produced on the operator's computer consoles.

Operator switchrooms are designed to be efficient workplaces, and staff numbers and rosters are planned to meet customers' call demand. Just as an automatic telephone network must be provided with sufficient circuits to meet the traffic, so must the number of *positions* manned by operators at any given time of day match the traffic demand at that time. A useful quality target for staff providing an operator service is to aim to answer all calls within a given time (say 25 seconds), or perhaps to aim to answer 90% of calls within say 15 seconds. The latter statistic is often written in shorthand as  $PCA15 = 90\%$ , i.e. the *percentage of calls answered* in 15 seconds = 90%.  $PCA25$  can also be used as a performance statistic (measuring the percentage of calls answered in 25 seconds), but the average caller on a public network who wants operator assistance, is not satisfied with such a long wait. There is a relationship between the measured value of PCA and the staffing level of the switchroom, the use of more staff generally increasing the PCA value. Going to one extreme, to employ a very large number of operators queueing up to answer calls would ensure almost instantaneous answer. At the other extreme, with too few operators it is the caller who does the queueing and waiting. A modification of the *Erlang formula* presented in Chapter 30 is called the *Erlang waiting-time formula*. It can be used to calculate the number of operators required in a switchroom to keep the waiting time down to a target figure (i.e. to calculate PCA values)

## 14.4 OPERATOR PRIVILEGES

In the onward connection of calls, operators may be given a number of special networking privileges. They may have exclusive use of particular routes or of certain

circuits within a route, to give their callers a better chance of getting through than normal customers (especially when the network is busy). This enables the operator to be of real assistance to the caller in cases of difficulty, and furthermore reduces the likelihood of wasted operator time spent in futile repeat attempts. Other privileges explained below include *manual hold*, *circuit monitoring* and *interruption*, and *forward transfer*. However, with the increasing development and automation of networks, and the small number of human operators available for network *policing*, these features are becoming obsolete.

Manual hold allows the operator to *hold* the connection even after the calling subscriber has replaced the handset. This makes it possible to trace the origin of a malicious call in a case when a caller has given a false identity. It also prevents the caller making any further calls. This use of the facility is now largely superseded by *calling line identity (CLI)* information in automatic networks and many networks today no longer have the facility. An alternative use of the *manual hold* facility permits tracing the cause of faulty connections. The ability to trace emergency service calls (e.g. fire, police, ambulance) is of special importance.

Circuit monitoring and interruption: sometimes the operator is given the facility to monitor or interrupt customers calls while in progress. This can be useful in investigating customer complaints, including account discrepancies. It can also be used to break in on conversation already in progress, when an important incoming call is received, and this would have been done historically if a trunk call was received while only a local was in progress (hence the term for this facility, *trunk offer*).

Forward transfer: another facility becoming largely obsolete, the forward transfer facility allows the operator who has previously established a semi-automatic connection for the call, to request assistance from the operator at the destination exchange; international operators used it to provide *language assistance* (i.e. translation). The operators might speak an intermediate language (typically French or English) between themselves and their mother tongue to their own customers to resolve any difficulties during call set-up. (Example for person-to-person calls). The desired language of assistance is indicated by a special digit called the *language digit*, which is inserted by the originating operator's exchange into the called customer's dialled digit string during the signalling at call set-up. This is discussed in more detail later in the chapter.

## 14.5 TYPICAL ASSISTANCE SERVICES

As most calls are made automatically nowadays, operators need to provide only a range of *assistance* services to complement the automatic service; here are some of the more common ones.

### *'Station call' service*

A *station call* is the name given to an 'ordinary' call between two telephone stations, when it is made via the operator. A station call may be made (via the operator as opposed to automatically) either because the call cannot be dialled directly, or because customers prefer it, or perhaps because customers have had difficulty in getting



through. Another reason may be that the customer wishes to be rung back immediately after the call has finished to be *advised of the duration and charge (ADC)* (also called *time and charges*).

*'Reverse charge' or 'collect call' service*

For any of a variety of reasons, callers when travelling may not wish to pay for calls themselves, preferring to transfer the charges to the call recipients. The service which does this for them is the *reverse charge* or *collect call* service. The reason for transferring the charge may be shortage of change when using a payphone, or to avoid leaving ones host with a large telephone bill when staying away from home. Whatever the cause, collect call service must always be made via the operator. Before receiving a *collect call*, recipients are asked by the operator whether they are willing to accept the call charges. If so, the call is connected and an operator *ticket* records the call details in the normal way, except that the bill is sent to the recipient rather than to the caller. A similar service, self-explanatory, is also sometimes available: 'Bill call to third party' (this might allow a payphone caller to charge the call to his home account).

*'Personal call' service*

(Also called a *person-to-person call*.) Sometimes callers may wish to contact particular people who share their telephone with a number of others. A caller may not want to make an automatic call and pay for a connection, but to find out that the desired individual is not available. In these circumstances it is appropriate to make a *personal call*, via the operator. The caller gives the operator the name and telephone number of the individual required, and the operator then makes the call and checks that the right recipient is available to come to the telephone. If so, the caller is charged from the moment when the operator allows conversation to commence. Usually either a surcharge or a higher charge per minute of conversation is levied on personal calls. If the person wanted is not available, the connection is cleared without conversation and the caller is not charged.

*'Directory enquiry' service*

(Also called *directory assistance*.) To make an automatic call a caller must have the number of the destination telephone station. Without that number the network has no indication of what connection the caller wants. If the caller does not know it, perhaps because he has not called that particular person or company before, then one way of 'looking-up' numbers is to use the paper directory, issued to telephone customers. This gives an alphabetic list of the names of all customers with their numbers. However, the sheer number of customers nowadays has tempted many telephone companies to issue only a telephone directory covering the immediately surrounding geographical area. The operator *directory enquiry (DQ) service* provides a more comprehensive nationwide service. Access to the service in the normal way is by dialling an access code and waiting in a queue at the nearest directory enquiry switchroom. The operator looks up the number in the appropriate paper directory or by *querying* a computerized directory system. When the number is found it is given verbally to the caller, who then presumably follows up the enquiry by placing a call over the automatic network. Alternatively the operator

can immediately put a call through. Some network companies charge customers for each directory enquiry; others give a free service, which they believe stimulates more automatic calls anyway.

### *General enquiry service*

Whether publicized directly or not, the operator is also often called upon by customers to answer general enquiries about the telephone or another service. A typical enquirer might ask for the area code to be used for a particular town, or an international operator may be asked for the time-of-day difference in hours between originating and destination countries. The answers to these enquiries, and further reference information to help operators in the undertaking of all their services is usually provided in the form of a *visual index file (VIF)*. This is a handy reference book, kept on each operator's console. Operator *tickets* may or may not be made out to record general enquiries, and charges may or may not be levied.

## 14.6 COOPERATION BETWEEN INTERNATIONAL OPERATORS: CODE 11 AND CODE 12 SERVICE

When networks belonging to more than one company are interconnected, there are times when the operator in one network requires assistance from the operator in the other. This usually comes about because neither of the operators has quite the same privileged control over the other network, as they have over their own. Alternatively, in international networks operators may require no more than language assistance (i.e. translation of the distant end mother tongue into a comprehensible intermediate language).

To give operators of different networks a means of calling one another, ITU-T's telephone signalling systems are provided with three signals specifically for the use of operators; these are

- *code 11* signal
- *code 12* signal
- *language digit* signal

The *code 11* signal is used by the operator of the originating network to obtain assistance from the operator of the terminating network, in cases where the call cannot be connected semi-automatically because the terminating network is a manual one. The signal gives access to appropriately equipped operators in the terminating country. In countries where the network is already fully automatic, *code 11* service will have been withdrawn. The *code 11* signal itself is single digit signal, just as the values 1, 2, 3, 4, 5, 6, 7, 8, 9, 0 are. The difference is that code 11 cannot be dialled by ordinary telephone customers. This precaution prevents ordinary telephone subscribers from masquerading as telephone operators, and from using this privileged role to defraud overseas network companies.

The *code 12* signal, like the *code 11* signal, is also used by the operator of an originating network to get assistance from an operator in the terminating network. Also

like the *code 11* signal, the *code 12* signal is a single digit signal which cannot be dialled by ordinary telephone customers. *Code 12* is to be used when difficulty has previously been experienced in establishing an automatic or semi-automatic connection, and it gives access to an appropriately equipped operator in the terminating country. Unlike the *code 11* signal, *code 12* is often used with extra digits.

e.g. (Code 12) + ABCD

The extra digits ABCD may be used to identify an individual or particular group of operators within a switchroom, so helping overseas operators to return to the same assisting operator.

*Code 11* and *code 12* signals are always sent immediately following the country code and language digit of the international number.

The *language digit* is an extra digit, inserted into the digit train of all semi-automatic operator controlled telephone calls to indicate the language that the calling operator would prefer distant operator to speak in, should language assistance be required. Not all countries' telecommunications companies offer language assistance appropriate to all the language digits, but the following values are allocated.

1 = French

2 = English

3 = German

4 = Russian

5 = Spanish

6 to 9 (spare)

(0 = dimensioning digit for automatic working)

The *language digit (LD)* is always inserted immediately after the country code of the number dialled, or in the alternative position demanded by the international signalling system. Thus on a call to the number

44 171 234 5678 (a number in Central London, UK)

the train

44 (1) 171 234 5678

might be used by an operator overseas requiring language assistance from the UK operator in French. The digit is not inserted by the operator manually, but rather is systematically added by the exchange on a route by route basis. Thus LD=2 for English might set on the route to Japan and LD=1 for French might be appropriate on a route to a French African colony. When the language digit is set at value 0, it is called the *discriminating digit*, and it is used to identify automatic (i.e. non-operator-controlled) calls. The discriminating digit (value 0 of the language digit) is inserted systematically by automatic telephone exchanges. Thus the receiving network can

distinguish between automatic origin calls and operator controlled (i.e. semi-automatic) calls. If necessary, and in response to either a *code 11*, *code 12*, or *forward transfer* signal, *language assistance* can be given. Having examined the language digit, and in so doing performed appropriate routing and call accounting, the international exchange in the destination country (but not an international transit exchange) will remove the language digit before sending out the destination national number into the network of the terminating country. The language digit therefore only exists on international telephone links.

## 14.7 A MODERN OPERATOR SWITCHROOM

Since the early days of operator-controlled manual telephone networks, operator switchroom equipment has moved through several phases of technology, first to cordless boards and latterly to fully computerized systems. It may seem ironic that a computer system should be designed to work in a mode requiring human intervention rather than setting out to eliminate the human element; but the fact is, telephone customers still find comfort in spoken assistance.

Today, computerized operator exchanges give optimized ergonomic conditions for the human operators without compromising the network efficiency. Some of these exchanges have a central switching equipment, where the switching and control of circuits takes place; but in addition, more luxurious remote switchrooms where the operators sit are connected via computers and computer data links. From their remote switchroom (or maybe even from a workplace at home) the operators can issue commands, using the keyboard of a computer terminal, to instruct the central switch what call routing, or other control action should be performed. Meanwhile the computer can automatically generate its own electronic tickets, storing the A number (callers number), B number (destination number), call duration, time of day, etc. Some of the information can be typed in by the operator when talking to the caller (A party) (e.g. the B number), and some of the information can be derived automatically (e.g. time of day and call duration). Figure 14.7 illustrates schematically the network arrangement of this type of computer-controlled operator exchange.

As well as the ergonomics and comfort benefits of this type of switchroom (more like an office, and less like the factory-like switchrooms of the past), there are a number of other benefits.

- The operator time required to handle individual calls is reduced.
- Call re-attempts can be made more easily (by pressing a 'last number re-dial' button).
- The accuracy of operator tickets is improved, and the information is directly available in a computer format for subsequent computerized bill calculation.
- The central switching equipment can be located with other automatic exchanges, and maintained by a common technical staff, while the switchrooms may be located near an available workforce.

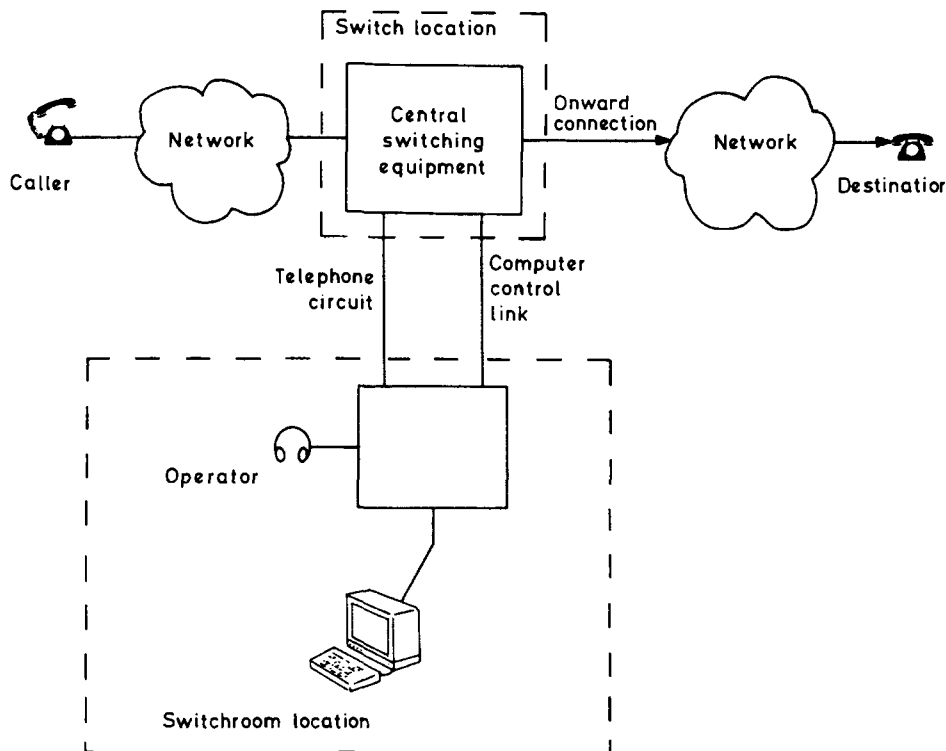


Figure 14.7 Network schematic of modern computer operator exchange

- If more than one remote switchroom or *homeworking* operators are connected to the same central switching equipment better staff rostering can be achieved, say by closing some switchrooms at night, and handling all the overnight traffic at one switchroom. This was not generally possible in the past.

## 14.8 OPERATOR ASSISTANCE ON TELEX NETWORKS

The principles of telex operator assistance networks are similar to those of telephone operator assistance networks. The difference is that the operator has to 'speak' to the caller through a telex machine and not by voice. Other slight differences exist in the actual menu of services offered; for example, the telex operator might offer a multiple destination service to telex callers.

## 14.9 OPERATOR ASSISTANCE ON DATA NETWORKS

Operator assistance is not normally available on packet-switched and other data networks. However, for the purpose of point-to-point data connection over the telephone

network, some companies (declining in number) allow their telephone operators, when asked by callers, to select particularly high grade circuits designed specifically for the transmission of voice band data. The customer uses such circuits in conjunction with data modems located on the customer premises at either end of the connection.



---

# Mobile Telephone Networks

---

Being away from a telephone, a telex or a facsimile machine has become unacceptable for many individuals, not only because they cannot be contacted, but also because they may be deprived of the opportunity to refer to others for advice or information. For these individuals, the advent of mobile communications promises a new era, one in which there will never be an excuse for being 'out-of-touch'. This chapter discusses modern mobile radio communication technologies, covering the principles of 'radio telephone service', 'trunk mobile radio (TMR)' cordless telephones, the 'global system for mobilecommunication (GSM)', as well as describing telephone communication with ships, aircraft, and trains and the emerging satellite mobile networks (e.g. Iridium and Globalstar). Mobile datacommunication networks are covered in Chapter 24.

## 15.1 RADIO TELEPHONE SERVICE

The first radio telephone services were manually operated in the high frequency radio band. These supported international telephone services, as well as communication with ships and aircraft. Immediately after World War 2, a new breed of VHF (very high frequency) radio transmitters and receivers (*transceivers*) were developed. First used for applications such as the police, the fire service and in taxis, later development lead to their use as full *radio telephones*, connected to the public switched telephone network (PSTN) for the receipt and generation of ordinary telephone calls.

The technology used a radio mast located on a hill and equipped with a powerful multi-channel radio transceiver. The mobile stations were weighty but were nonetheless popular for commercial 'car telephone' service, which grew rapidly in popularity in the mid-1970s.

For calls made to or from the radiotelephone user, the public telephone network is connected via *mobile switching centres (MSCs)* to a number of transmitter *base stations* which emit and receive radio signals from the mobile telephones. Two radio channels (using different frequencies) are needed to connect each telephone during conversation, one radio channel for each direction of conversation. Figure 15.1 illustrates a typical automatic radio telephone system of the late 1970s.



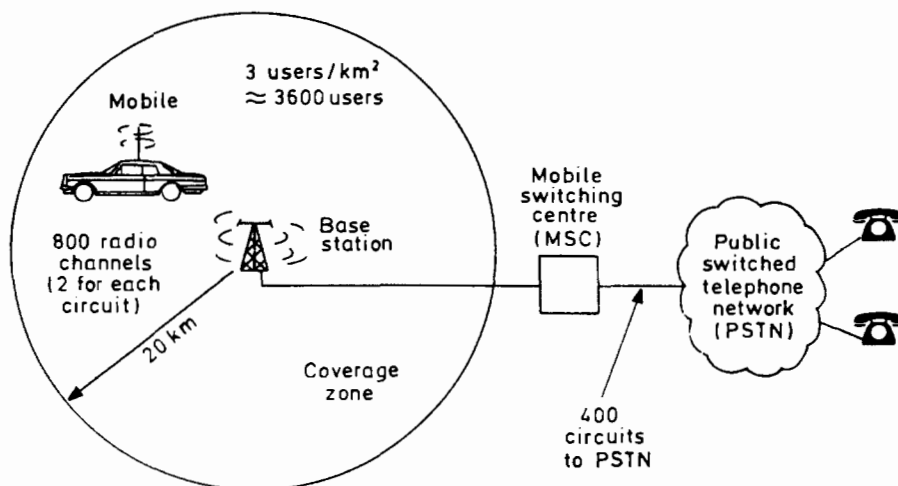


Figure 15.1 A simple radio telephone system

Figure 15.1 illustrates a single mobile switching centre and a single transmitter base station, serving 3600 mobile customers within an area of  $1200 \text{ km}^2$ , within a radius of about 20 km from the base station. Calls are made or received by the mobile telephone station by monitoring a particular radio channel, called the *signalling* or *call up* channel. When making a call, the mobile station signals a message to the base station in much the same way as an ordinary telephone sends an *off-hook* signal and a dialled digit train (Chapter 7). The base station next allocates a free pair of radio channels to be used between the base station and the mobile for the subsequent period of conversation, and the two switch over to the allocated channels simultaneously.

At the end of the call the radio channels are released for use on another call. Incoming calls to the radio telephone connect in the same way. An ordinary telephone number is dialled by an ordinary customer connected to the public switched telephone network (PSTN). A particular *area code* within the normal telephone numbering plan is usually allocated to identify the mobile network, and all calls with this dialled area code are routed to the mobile switching centre and via the appropriate base station transmitter to reach the desired mobile receiver.

A hurdle to be overcome in the design of radio telephone systems is the need for incoming calls to be routed only to the nearest base station. Failure to do so causes the failure of incoming calls, because the mobile cannot guarantee to be within the coverage area of a particular transmitting base station. Early radio telephone systems overcame the problem by requiring the caller to know the whereabouts of the mobile that he wished to call. The same customer number would be used on all incoming calls, but a different area code would have to be used, depending on which base station the caller wished the call to be routed to (i.e. according to which zone the caller thought the mobile was in). Figure 15.2 illustrates this principle.

We see here that if the mobile customer's number is '12345', then when a caller believes that the mobile is in zone 1, the number 0331 12345 should be dialled. Similarly for zones 2 and 3, the numbers 0332 12345 and 0333 12345 are appropriate. Some more advanced

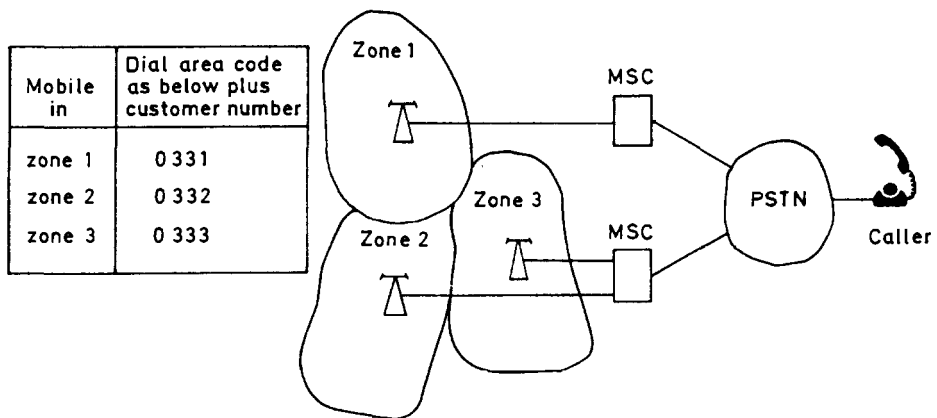


Figure 15.2 Area code routing for incoming radio telephone calls

radio telephone systems were developed with the ability to 'remember' the last location from which a call was made and route the call there. Others used 'trial-and-error' methods, but these systems were all superseded by the advent of *cellular radio*, which has eliminated the market for old-style radio telephones used by roaming users.

Another drawback of the early radio telephone systems was their low user capacity, and their 'unfriendly' usage characteristics. Not only did the automatic systems rely on the caller to know where the mobile was, but some demanded the 'press-to-speak' action of the mobile user. The systems allowed the mobile user to continue to move about during the call, but if he or she moved out of the coverage area of the base station, the call would be terminated. There were no facilities for transferring to other base stations during the course of a call. Together, these drawbacks lead to the demise of radiotelephony in its original guise and to its replacement with *cellular radio*, discussed next. The technology has, however, survived for localized radio network coverage, such as required by taxi companies or regional haulage companies. The terms *private mobile radio (PMR)* or *trunk mobile radio (TMR)* (in Germany *Bündelfunk*) are more commonly used nowadays. Some interest has been shown to extend the usage and life of the technology, as demonstrated in the recent development by ETSI of new standards for *TETRA (trans-European trunk radio)*, which aims to provide relatively low speed integrated voice and data networking capability to and from mobile stations located within relatively large geographical radio coverage areas. *TETRA* is discussed more fully in Chapter 24.

## 15.2 CELLULAR RADIO

A difficulty with early radio telephone systems was the small capacity in number of users. This arose because the early systems were designed to have very wide area coverage, but had a limited number of radio channels available within each zone. Furthermore, the *re-use* of radio channels in other zones was precluded by the risk of

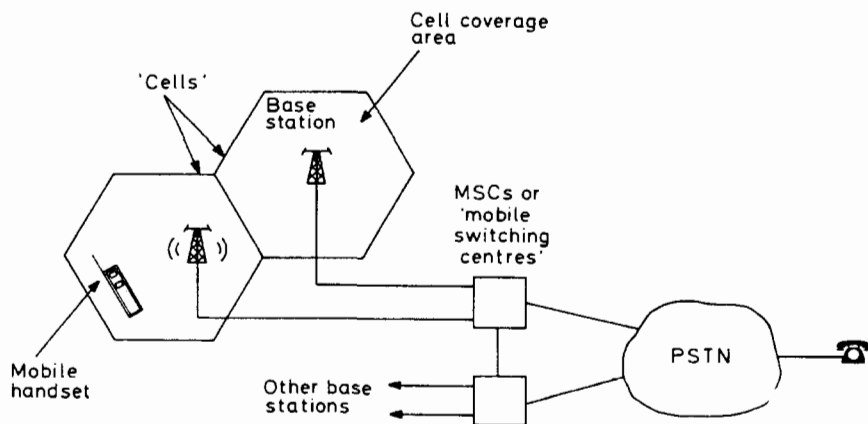
interference except where base stations were separated by large distances. Because the 1970s saw a boom in radio telephone demand, and because radio channel availability was (and still is) a limited resource, there was pressure for development of revised methods, more efficiently using the radio bandwidth, thereby greatly increasing the available user capacity. The system that evolved has become known as *cellular radio*. The basic components of cellular radio networks are *mobile switching centres (MSCs)*, *base stations*, and *mobile units*, as Figure 15.3 shows.

Cellular radio networks make efficient use of the radio spectrum, *re-using* the same radio channel frequencies in a large number of base stations or *cells*. Oriented in a 'honeycomb' fashion, each *cell* is kept small, so that the radio transmitting power required at the transmitting *base station* can be kept low. This limits the area over which the radio signal is effective, and so reduces the area over which radio signal interference can occur. Outside the *interfering zone* of the transmitter, i.e. in a non-adjacent cell at a sufficient distance away from the first, the same radio channel frequencies may be *re-used*. Figure 15.5 illustrates the interfering zone of a given cell base station, and shows another cell using the same radio channel frequencies.

A whole honeycomb of cells is established, re-using radio channels between the various cells according to a pre-determined plans. A seven cell *re-use pattern* is shown in Figure 15.6. Seven different radio channel frequency schemes are repeated over each cluster of seven hexagonal cells, each cell using a different set of frequencies. By such planning the same radio frequency can be used for different conversations two or three cells away.

Figure 15.6 shows a 7-cell *re-use pattern*. Other patterns, some involving as few as three cells, and some more than thirty cells, can also be used. Large repeat patterns are necessary to cater for heavy traffic demand in built-up areas where small non-adjacent cells may still interfere with one another.

Each cell is served initially by a single base station at its centre and is complemented as traffic grows with directional antennas and more radio channels. Using directional antennas helps to overcome radio wave *shadows*. For example, to locate three



**Figure 15.3** The basic components of a cellular radio network



Figure 15.4 Cellular radio carphone: mounted and in use in a car

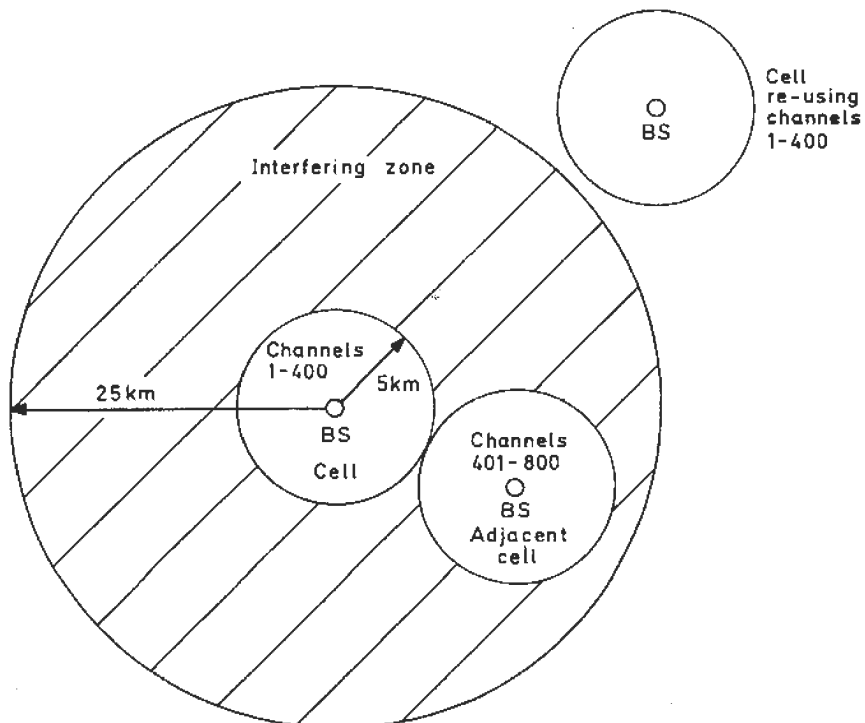


Figure 15.5 Cellular radio channel interference and re-use. BS = Base Station

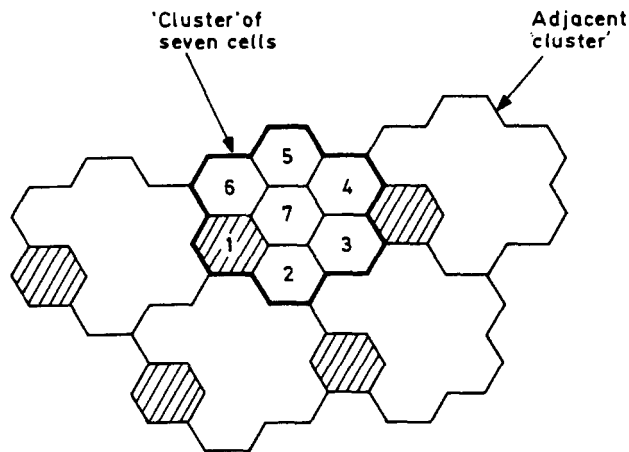


Figure 15.6 Cellular radio in re-use pattern

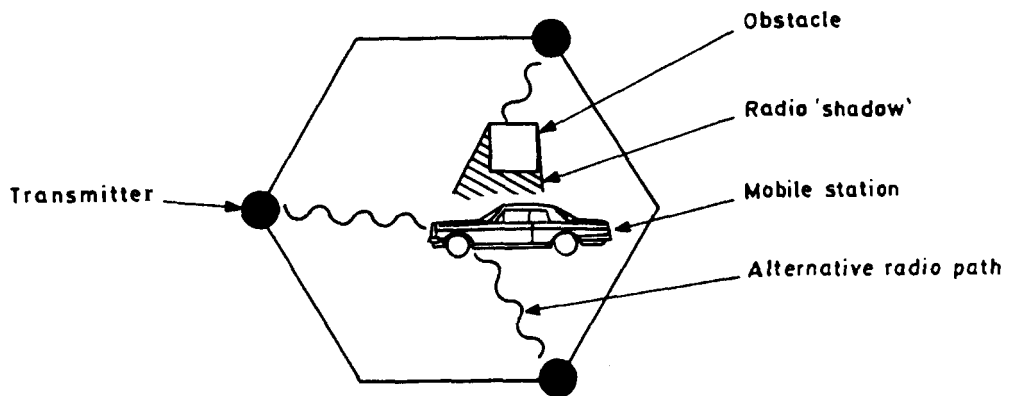
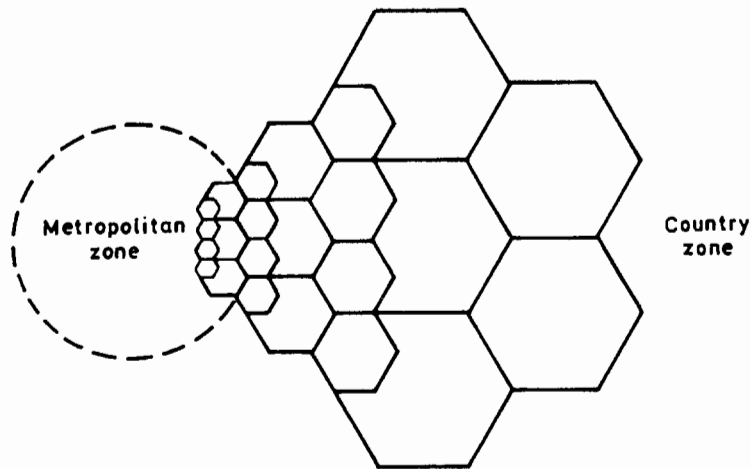


Figure 15.7 Location of multiple base stations

directional antennas, one at each alternate corner of the cell, helps to overcome the *shadow* effects that might otherwise occur near tall buildings by giving an alternative transmission path, as Figure 15.7 shows.

A feature of cellular radio networks is their ability to cope with an increasing level of demand first by using more radio channels and more antennas in the cell, and then by reducing the size of cells, *splitting* the old cells to form a multiplicity of new ones. Only a limited number of radio channels can be made available in a cell at the same time, and this limits the number of simultaneous telephone conversations. By increasing the number of cells, the overall call capacity can be increased. The number of channels needed in a given cell is determined by the normal *Erlang formula* (Chapter 30).

The total call demand during the busy hour of the day depends on the number of callers within the cell at the time. Reducing the size of the cell has the effect of reducing the number of mobile stations that is likely to be in it at any time, and so relieves



**Figure 15.8** Cell splitting to increase cell capacity

congestion. Figure 15.8 shows a simple splitting of cells and a gradual reduction in cell size in the transitional region between a low traffic (country) area and the high traffic region surrounding a major metropolitan zone. When splitting cells in this manner, due care needs to be taken when allocating radio frequencies to the new cells, and a new frequency re-use plan may be necessary to prevent inter-cell interference.

### 15.3 MAKING CELLULAR RADIO CALLS

One or more *control* or *paging* radio channels are used to make or receive calls between the mobile station. If a mobile user wants to make a call, the mobile handset scans the pre-determined channels to determine the strongest control channel, and monitors it to receive network status and availability information. When the telephone number of the destination has been dialled by the customer and the *send key* has been pressed, the mobile handset finds a free control channel and broadcasts a request for a *user* (i.e. radio telephone) channel. All the base stations use the same control channels, and monitor them for call requests. On the receipt of such a request by any base station, a message is sent to the nearest *mobile switching centre (MSC)*, indicating both the desire of the mobile station to place a call and the strength of the radio signal received from the mobile. The MSC determines which base station has received the greatest signal strength, and, based on this, decides which cell the mobile is in. It then requests the mobile handset to identify itself with an authorization number that can be used for call charging. The authorization procedure eliminates any scope for fraud.

Following authorization of an outgoing call, a free radio channel is allocated in the appropriate cell for the carriage of the call itself, and the call is extended to its destination on the public switched telephone network. Incidentally, the appropriate cell need not necessarily imply the nearest base station; more sophisticated cellular networks might also choose to use adjacent base stations if this will help to alleviate

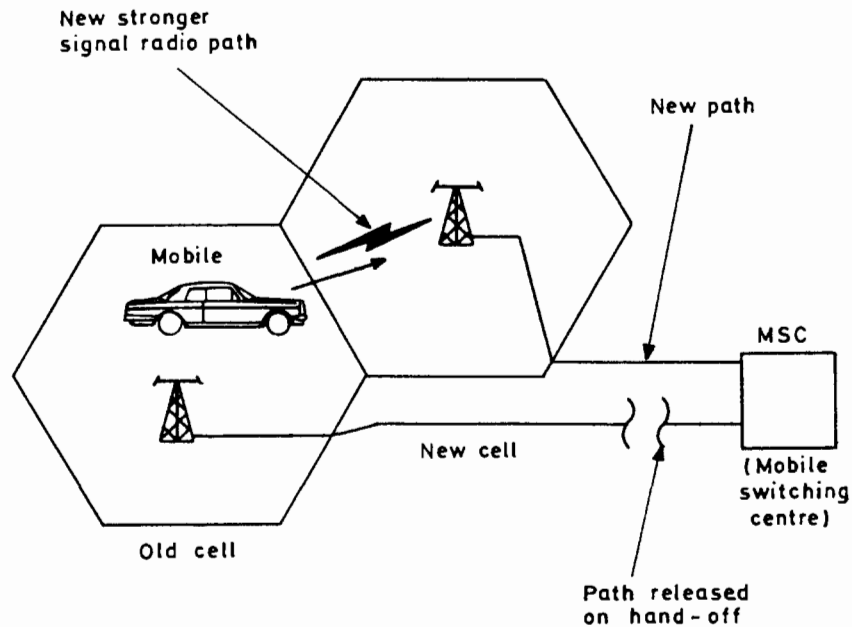


Figure 15.9 'Hand-off' during a call

channel congestion. At the end of the call, the mobile station generates an *on-hook* or *end of call* signal which causes release of the radio channel, and reverts the handset back to monitoring the control and paging channel.

Each mobile switching centre controls a number of radio *base stations*. If, during the course of a call, the mobile station moves from one cell to another (as is highly likely, because the cells are small), then the MSC is able to transfer the call to route via a different base station, appropriate to the new position. The process of changeover to the new cell occurs without disturbance to the call, and is known as *hand-off* or *handover*. This is one of the most important capabilities of a mobile telephone network. *Hand-off* is initiated either by the active base station, or by the mobile station, depending upon the network and system type. The relative strengths of signals received at all the nearest-base-stations are compared with one another continuously, and when the current station signal strength falls below a pre-determined threshold, or is surpassed by the signal strength available via an adjacent base station, then *hand-off* is initiated. The mobile switching centre establishes a duplicate radio and telephone channel in the new cell, and once established, the call is transferred to the new radio path by a control message to the mobile handset. When confirmed on the new channel and base station, the original connection is cleared, as Figure 15.9 illustrates.

## 15.4 TRACING CELLULAR RADIO HANDSETS

Whenever the mobile handset is switched on, and at regular intervals thereafter, it uses the control channel to register its presence to the nearest mobile switching centre. This

enables the local mobile switching centre at least to have some idea of the location of the mobile user. If outside the geographical area covered by the base stations controlled by its *home MSC*, the local MSC (i.e. the nearest to the mobile user) undertakes a *registration procedure*, in which it interrogates the *home MSC* (or the intelligent network database associated with it) for details of the mobile, including the authorization number and other information. The information is held by the home MSC in a database called the *home location register*, or *HLR*. It contains the mapping information necessary for completing calls to the mobile user from the PSTN (its network identity, authorization and billing information). The local MSC duplicates some of this information in a temporary *visitor location register* or *VLR*, until the caller leaves the MSC area. Once the visiting location register has been established in the local MSC, outgoing calls may be made by the mobile user.

The *registration procedure* is a crucial part of the mechanism used for tracing the whereabouts of mobile users, so that incoming calls can be delivered. Incoming calls are first routed to the nearest mobile switching centre (MSC) to the point of origin (i.e. the caller). This MSC interrogates the *home location register* for the last known location of the mobile user (this is known as a result of the most recent mobile registration). The call can then be forwarded to the mobile switching centre where the mobile was 'last heard', whereupon a paging mechanism, using the base station control channels, can locate the exact cell in which the mobile is currently located. A suitable free radio channel may then be selected for completion of the call.

Both the handsets and the network infrastructure needed to support cellular radio are complex and expensive, although the increase in user demand is reducing the cost. The mobile handsets come in a number of different forms, from the traditional car-mounted telephone, to the pocket versions. The latter are expensive not only because of the feats of electronics miniaturization that has been necessary but also because the trends towards pocket telephones has necessitated advanced battery technology and the use of sophisticated battery conservation methods (which, in effect, turn much of the unit off for as much of the time as possible, to save power). The *mobile switching centres* (MSCs) rely on advanced computers capable of storing and updating large volumes of customer information, and also of rapidly interrogating other MSCs for the location of out-of-area, or *roaming* mobiles. The interrogation relies on the use of the *mobile application part* (MAP) and the *transaction capability* (TCAP) user parts of SS7 signalling (which we discussed in Chapter 12).

## 15.5 EARLY CELLULAR RADIO NETWORKS

A number of different cellular radio standards have evolved, with the result that hand portables purchased for use on one system are unsuitable for use on another. The most common of the systems currently in use worldwide are as follows.

### *AMPS (Advanced Mobile Telephone System)*

This was first introduced in Chicago in 1977 by Illinois Bell, at that time one of the operating companies of AT&T in the United States. Developed by AT&T's Bell



Laboratories, this became the most commonly used system in North America up to the early 1990s, when digital systems began to take over. It operates in the 800 MHz and 900 MHz radio bands, at a channel spacing of 30 kHz.

### *NMT (Nordic Mobile Telephone Service)*

This commenced service in the Nordic countries in 1981 and is used in other countries in Europe (Austria, Belgium, Czechoslovakia, France, Hungary, Netherlands, Spain, Switzerland). The original system was 450 MHz based but has been gradually extended and to some extent replaced after 1986 with the later NMT-900 version.

### *C (Network C)*

Network C was a development of Network B, a digitally controlled mobile radio system introduced in the Federal Republic of Germany in 1971. Network B was 450 MHz based. Its replacement, Network C can be either 450 MHz or 900 MHz based (hence C-450 and C-900); it is still in use in Germany and Portugal, but being replaced by GSM systems.

### *TACS (Total Access Communication System)*

This is a derivative of the American AMPS standard, modified to operate in the 900 MHz band, with a more efficient 25 kHz radio channel spacing. TACs was the system introduced into the UK and Ireland during 1985. *ETACS* or *extended TACS* is a compatible derivative of TACs which gives greater channel availability, particularly in very congested areas like the metropolitan London area. The system is also used in Austria, Italy and Spain.

Other mobile telephone standards include *NAMTS* (*Nippon Automatic Mobile Telephone System* used in Japan), *Radiocom 2000* (used in France), *RTMS* (second generation Mobile Telephone used in Italy), *UNITAX* (used in China and Hong Kong) and *Comvik* (used in Sweden).

**Table 15.1** Comparison of analogue cellular radio network types

	AMPS (USA)	C 450 (Germany)	NMT 450 (Scandinavia)	NMT 900 (Scandinavia)	TACS (UK)
Uplink band	824–849 MHz	450–455 MHz	453–458 MHz	890–915 MHz	890–915 MHz
Downlink band	869–894 MHz	461–466 MHz	463–468 MHz	935–960 MHz	935–960 MHz
Channel spacing	30 kHz	20 kHz	25 kHz	25 kHz	25 kHz
Multiplexing	FDMA	FDMA	FDMA	FDMA	FDMA
Modulation	PSK	FSK	FSK	FSK	PSK
Number of channels	833	222	180	1000	1000

The trend of the above systems to migrate to the 900 MHz follows the allocation of frequencies in the band by the *World Administrative Radio Council (WARC)* in 1979. Subsequently there has been pressure to digitalize the radio speech channels, as well as the control channel, to improve radio spectrum usage. Digital channels allow closer channel spacing, due to reduced interference between channels. The adoption of digital technology enabled, first, greater efficiency in the use of available radio bandwidth and therefore higher traffic volumes; second, dramatic price reductions resulting from miniaturization and large scale production of components. Further, the introduction of competing cellular radio telephone network operators as the first stage of deregulation and competition for the traditional monopoly telephone companies simultaneously resulted in much reduced prices for mobile telephone services. The result has been a worldwide boom in mobile telephony. The predominant technical standards are those of *GSM (global system for mobile communication)* and *PCN (personal communications network)*, also known as *DCS-1800: digital cellular system/1800 MHz*.

## 15.6 GLOBAL SYSTEM FOR MOBILE COMMUNICATIONS (GSM)

In 1982, *CEPT (the European Conference of Posts and Telecommunications)* decided to start work on specifying new technical standards for the support of a pan-European mobile telephone service, based entirely on digital radio transmission. The task group set up to perform the specification work was the *CEPT/GSM* group, *GSM* being initially an acronym of *Groupe Speciale-Mobiles*. It was an ambitious programme which targetted a network spanning the whole of Europe by 1991, which would allow mobile handset users to roam anywhere on the European land mass and still be able to make and receive calls.

The GSM programme was adopted by the European Commission of the European Union, who recognized its potential for spurring standardization and deregulation of the telecommunications market in Europe. In particular, GSM became the potential for standardization of common networks across Europe, thus opening a *common market* for end user telephone equipment. In parallel, it presented an opportunity for governments of the European Union member countries to commence *deregulation* of their monopoly telephone markets without undue threat to the vast established *terrestrial* (or *fixed*) telephone networks. Finally, the programme was further reinforced by European Commission pressure on member countries to participate in a *memorandum of understanding (MOU)* initially signed in mid-1988, committing the network operators in European countries to ensure operation of the networks by 1991, allowing in addition, full roaming of mobile users between the networks and countries.

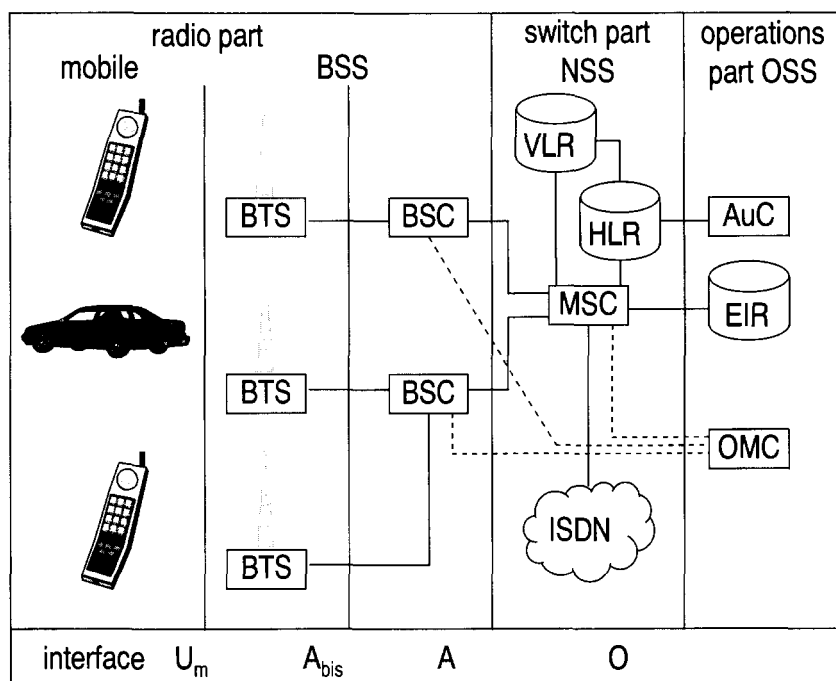
Although the 1991 target was not met by all countries, it led to a flurry of second operator licensing across Europe, and to the rapid deployment of GSM networks. In most European countries, at least two GSM networks have been established, one owned by the monopoly operator, one operated by a new competing carrier. Most of the networks have been fully operational since 1993, with extensive coverage equalling or bettering the previous analogue mobile networks. With call prices under pressure from competition, the GSM mobile telephone boom took hold in late 1993.

A major attraction of the GSM system has been the ability of its users to make and receive telephone calls anywhere in the GSM world (anywhere in Europe, Australia, New Zealand, South Africa, Hong Kong, etc.) without having to advise callers of any change in telephone number (this is called *international roaming*).

## 15.7 GSM TECHNOLOGY

Figure 15.10 presents the structure and main elements of a GSM network.

In general, the system works exactly as described earlier in the chapter. The *mobile switching centres (MSCs)* are the main controlling elements of the networks (a national network typically comprises 10–25 MSCs). Each controls a given geographic area over which a number of *base transmitter stations (BTS)* are spread (typically 5–8 km radius cells). The *BSC (base station switching centre)* is the control element for the *base transmitter stations*, but need not be collocated with the BTS. Thus in a dense metropolitan area, several antenna sites may be used, but they require only one small BSC switching



AuC	authorization centre	HLR	Home location register
BSC	base station switching centre	MSC	mobile switching centre
BSS	base station sub-system	NSS	network and switching sub-system
BTS	base transmitter station	OMC	operations and maintenance centre
EIR	equipment identity register	OSS	operations and support sub-system

Figure 15.10 Main components of a GSM network

site. *Home location registers (HLR)* and *visitor location registers (VLR)* have the functions described earlier in this chapter. Typically one HLR and one VLR is associated with each MSC. As explained in Chapter 12, the *mobile application part (MAP)* is used in communications between MSC, HLR and VLR.

The MSC provides for call control and switching, and for gateway functions to other mobile or fixed networks. The GSM system has a number of additional security features compared with its predecessors. These aim to reduce the problem of stolen handsets. A *SIM card*, containing a small user identification chip and information concerning the users configuration, must be inserted into the user's handset before it will operate. This chip enables an authorization procedure to be carried out at each call set up between mobile station, MSC and the *equipment identification register (EIR)*. Should a handset or a card be lost or stolen, then the EIR may be updated to record this information. The EIR contains a *blacklist* of barred equipment and a *grey list* of equipment not functioning correctly or for which no services are registered. A *white list* contains all registered users and relevant subscription services.

The radio part of the GSM system uses a 25 MHz radio band. The *uplink* channels (mobile to base station) occupy the band 890–915 MHz. The *downlink* (base station to mobile) channels are between 935–960 MHz, whereby *uplink* and *downlink* channels of a particular channel pair are separated by 45 MHz.

The radio band is subdivided into 124 carrier frequency pairs, each of 200 kHz *uplink* and 200 kHz *downlink* bandwidth. Each carrier is coded digitally using *TDMA (time division multiple access)* as explained in Chapter 8). The normal TDMA frame used in GSM is illustrated in Figure 15.11.

The usable individual circuit bitrate is around 24 kbit/s (114 bits per circuit, every 4.615 ms). Frequency hopping (Chapter 8) provides for protection against radio fading caused by *multipath* effects, and also gives some protection against criminal overhearing of the radio signal. A further GSM measure against overhearing is a key-coded data *encryption*, whereby the key is held on the *SIM card* in the mobile station and the *authorization centre (AuC)* provides the encryption algorithm.

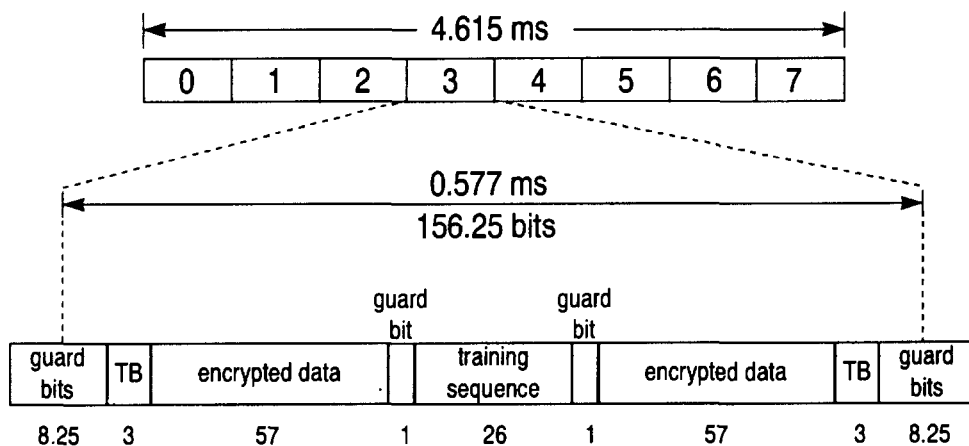
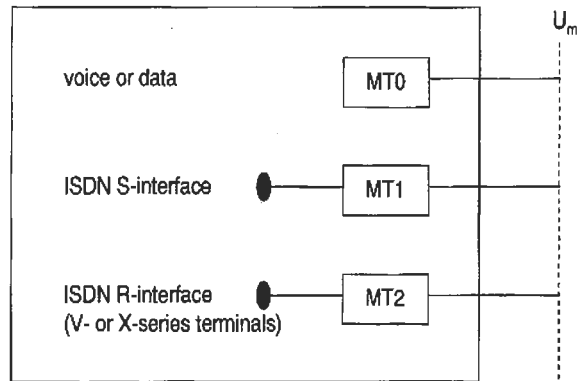


Figure 15.11 TDMA frame used in GSM



**Figure 15.12** Mobile station termination types for GSM

The radio modulation is *GMSK* (*Gaussian minimum shift keying*), a form of phase shift keying (PSK, *binary phase modulation*). The *PCM* (*pulse code modulation*) coding of the speech signal uses an *ADPCM* (*adaptive differential PCM*) algorithm (see Chapter 38) with a bitrate around 10 kbit/s. The remaining bitrate is needed for signalling, error correction, encryption and other protocol functions (e.g. *hand-off*, etc.).



**Figure 15.13** Modern GSM handheld telephone (handy). (Courtesy of Siemens AG)

The  $U_m$ -interface is the mobile telephone network equivalent of the ISDN U-interface (Chapter 10). It is the standard radio interface and protocol that allows mobile handsets from many different manufacturers access to the mobile network. Figure 15.12 illustrates the classification of the various ISDN-like end user termination tapes.

The boom in number of GSM subscribers and traffic volumes has prompted fast further development of the GSM system. The initial offering of carphones, requiring 20 Watt battery power have been rapidly supplanted by 8 Watt and 5 Watt portable equipments, and finally by 2 Watt *handies* (*handheld terminals*). These are the palm-sized mobile telephones now available in most electrical stores. The lower output power of the mobile stations equates to lower range and therefore smaller effective cell size. However, as the boom in traffic has demanded smaller cells anyway (the number of calls per cell being limited), this is no longer a practical or economic problem.

New developments of the GSM system will increase the intelligence and functionality of the system. Already, sophisticated mailbox services, short message services (displayed on the display of the mobile station as a short text), fax and data services are becoming common network offerings.

## 15.8 PERSONAL COMMUNICATIONS NETWORK (PCN) AND DCS-1800

The PCN system, now also known as *DCS-1800* (*digital communications system at 1800 MHz*), was originally intended to herald a new era of 'personal communications'. In the late 1980s a mass-market was foreseen for handheld mobile telephones (like GSM-*handies*), but the 900 MHz GSM standards were not thought at the time to be capable of satisfying the demand, because it was not thought that the mobile sets could be created small enough and with sufficient power and battery conservation to work effectively. An initiative for PCN commenced, and a new radio band at 1800 MHz was allocated.

In the event, the PCN initiative was largely overtaken by GSM, and the DCS-1800 system is little more than a modified version of GSM for the 1800 MHz radio frequency band. The network licensing programme of European governments then sealed PCN technology as a direct competitor to standard GSM by issuing further mobile network licences to new network operators. The interest of these network operators has largely been to 'cash-in' on the mobile telephone boom of the GSM operators, so that DCS-1800 network operators such as *E-plus* in Germany and *One-to-One* and *Orange* in UK cannot be distinguished by most customers from the classical cellular and GSM operators *D1* and *D2* in Germany and *Cellnet* and *Vodafone* in UK. A further reduction in market prices has resulted.

Table 15.2 compares the DCS-1800 system with that of GSM and the US and Japanese digital cellular systems.

One of the major lessons learned from both DCS-1800 and GSM has been how quickly viable alternative networks to the traditional, terrestrial-based public telephone can be established. GSM operators have become very adept at finding and buying or leasing small *footprint* sites where they may erect radio towers which serve

**Table 15.2** Comparison of DCS-1800 (PCN) with GSM, USDC and PDC

Parameter	GSM	DCS-1800	USDC (ADC) (US or American digital cellular system)	PDC (JDC) (Japanese personal digital cellular system)
Uplink band	890–915 MHz	1710–1785 MHz	824–849 MHz	940–960 MHz
Downlink band	935–960 MHz	1805–1855 MHz	869–894 MHz	810–830 MHz
Channel spacing	200 kHz	200 kHz	30 kHz	25 kHz
Duplex spacing	45 MHz	95 MHz	45 MHz	130 MHz
Multiplexing	TDMA/FDD	TDMA/FDD	TDMA/FDD	TDMA/FDD
Modulation	GMSK	GMSK	DQPSK	DQPSK
Speech data rate	13 kbit/s (6.5 kbit/s)	13 kbit/s	<13 kbit/s	<11 kbit/s
Frequencies	124	374	832	800
Time slots per radio channel	8 (16)	8	3	3
Data service	9.6 kbit/s	9.6 kbit/s	4.8 kbit/s	4.8 kbit/s
Maximum speed of mobile station	250 km/h	250 km/h	100 km/h	100 km/h
Output of handheld unit	2, 5, 8 or 20 Watt	0.25 or 1 Watt		

simultaneously as base stations for the radio links to mobile stations and masts for the installation of point-to-point microwave radio links as backhaul links to *mobile switching centres*.

The boom in demand for mobile telephones has caused operational problems for the cellular radio network operators, who in some cases have not been able to expand their networks in step with the growing demand. The result has been congestion. At first, the relief of congestion could be brought about by reducing cell sizes and so introducing more cells. However, as time has progressed, the radio bandwidth available has come to be a problem. Initially, DCS was seen as a solution. Nowadays, *DECT* (*digital European cordless telephony* described in Chapter 16) technology is increasingly being seen as a complement to GSM. The idea is that a *dual-mode telephone* handset could log into whichever network was available in a particular area. The claimed advantage of DECT is that much higher subscriber density can be supported, so overcoming the *hotspot* problem currently existing in some pure GSM networks.

## 15.9 AERONAUTICAL AND MARITIME MOBILE COMMUNICATIONS SERVICES

High frequency radio communications have long been used in aeronautical and maritime applications for navigational purposes and distress calls. Furthermore, many public telephone network operators have for many years offered ordinary telephone customers the opportunity to place or receive HF radio telephone calls to and from ships in coastal waters. Unfortunately high frequency radio systems, when used for communication over long distances at sea, suffer from poor signal propagation, atmospheric disturbance, interference, and signal fading. As a result, the ships' telephone radio service is of poor quality, and unreliable.

As a means of getting over this problem, a number of organizations during the late 1960s and early 1970s were studying the use of satellites for aeronautical and maritime communications. These studies led in 1976 to the launch of *MARISAT*, the world's first commercial maritime satellite system. It was conceived by a consortium of United States companies, and it comprised three *MARISAT* satellites in geostationary orbit, providing communications services for military and civilian use. At the outset heavy use of the system was made by the US Navy.

Further worldwide interest, followed up an initiative of the Inter-Governmental Maritime Consultative Organisation (IMCO) in 1973, led to the signing of an agreement by 26 countries in 1976, and the establishment in 1979 of the *International Maritime Satellite Organization (INMARSAT)*.

*INMARSAT's* satellites have revolutionized the world of communications with ships. Geostationary satellites now placed over the Indian, Atlantic and Pacific Oceans make possible communication to any suitably equipped ship, no matter where it is located around the globe. Automatic telephone calls, dialled directly by customers of some countries' public switched networks can be made to people aboard ships; the only requirement is that the caller knows which ocean the ship is in, so that the appropriate ocean country code number (part of the international telephone number) can be dialled, directing the call via the appropriate ocean's satellite. Similarly, persons aboard ships may make direct-dialled calls to other customers connected to the international public telephone network.

Figure 15.14 illustrates the main components of the maritime satellite network, comprising earth station equipments both on shore and aboard ship, together with a sophisticated *Access Control and Signalling Equipment*. The *ACSE* performs similar functions to those of cellular radio network *base stations* and *mobile switching centres (MSCs)*. It allocates radio channels for individual calls and carries out billing, accounting and other administrative functions.

A cheap form of satellite communication, allowing 'telex-like' communication to and from terminals on board smaller craft, vehicles and lorries is the *INMARSAT Standard C* service. This allows two-way communication using an omnidirectional antenna about the size of a saucepan.

By the end of 1987 the number of *INMARSAT* signatory countries had increased to 50. *INMARSAT's* original aim was to provide maritime communications, thereby improving the safety and management of ships. In 1985 the charter was expanded to include aeronautical satellite communications.



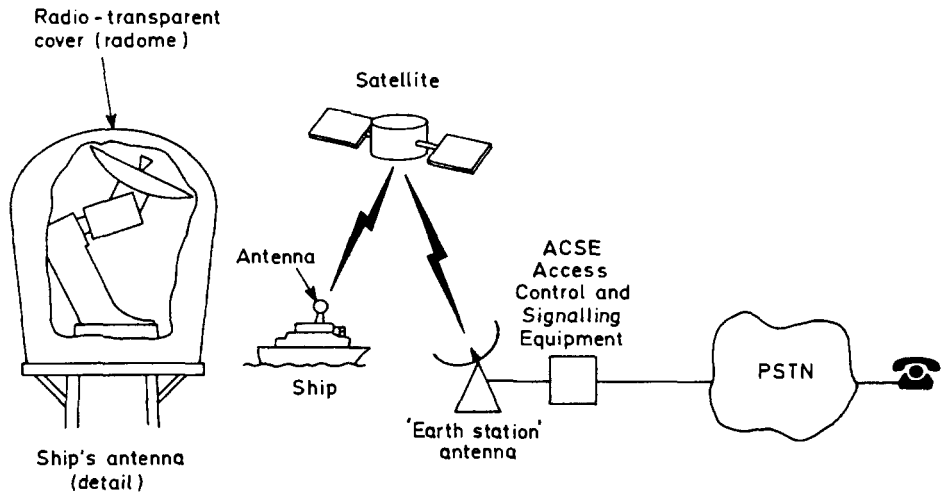


Figure 15.14 Maritime satellite communications

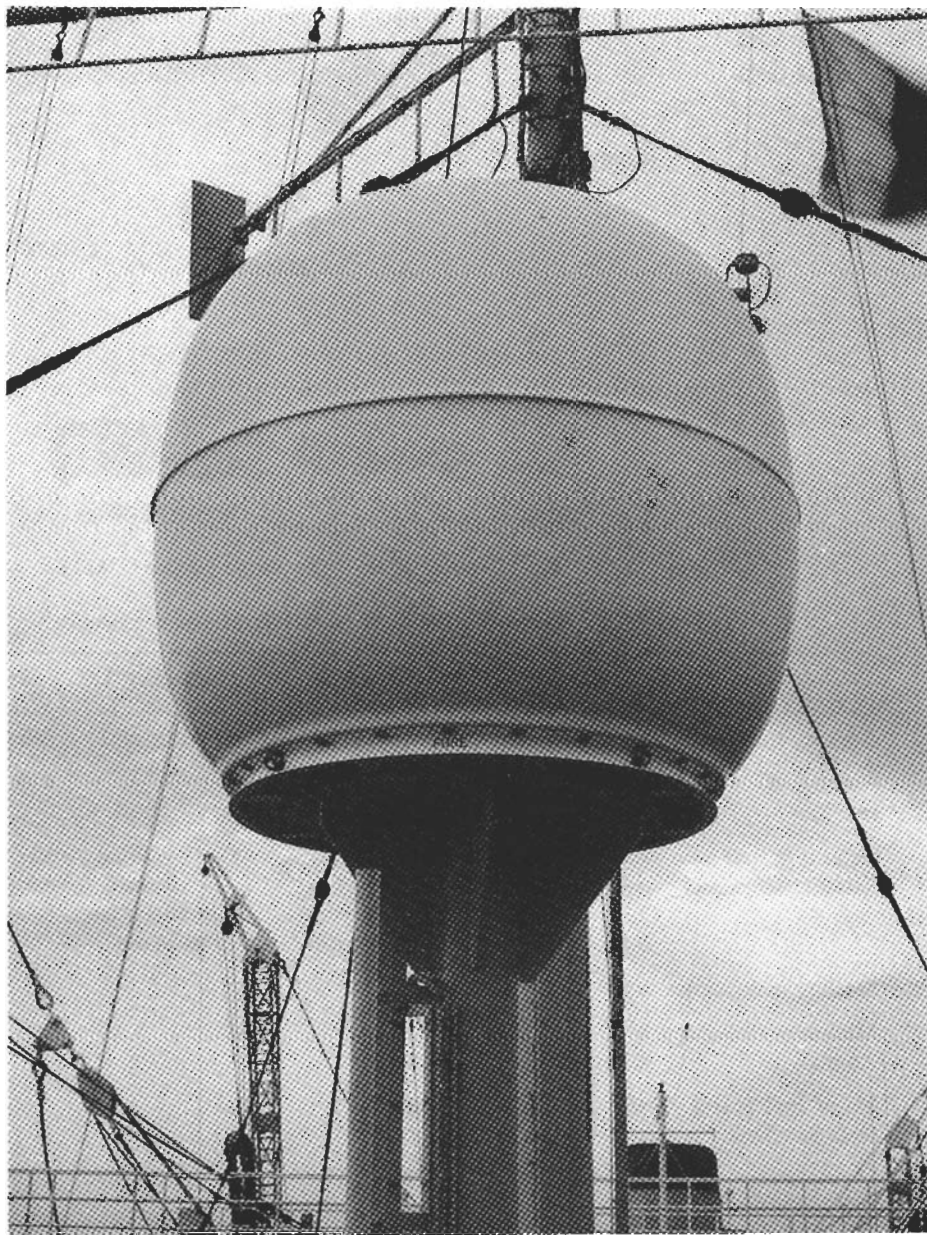
A number of initiatives have provided aeronautical satellite services. These bring the public telephone network to passengers and crew aboard commercial airliners.

### 15.10 IRIDIUM, GLOBALSTAR AND THE EVOLUTION TOWARDS THE UNIVERSAL MOBILE TELEPHONE SERVICE (UMTS)

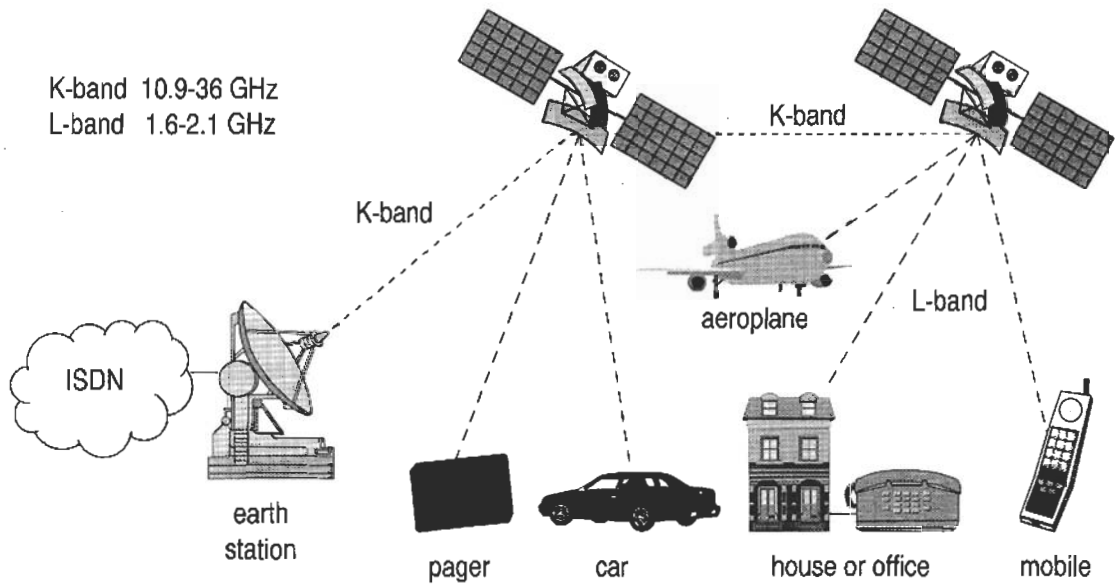
With the mobile telephone boom of the early 1990s in full swing in the developed countries, attention swung towards extending the possibilities of mobile roaming to worldwide coverage and to the increased penetration of mobile telephone services in less developed parts of the world. The result has been a number of new satellite technology initiatives. In these, the *base transmitter station (BTS)* and *base switching centre (BSC)* functions of GSM are combined into one unit which is now carried by a low orbiting satellite. The low orbit is necessary to reduce the transmit power needed in the mobile stations and also to reduce the undesirable long propagation time of signals transmitted to and from *geostationary* orbits at 37 000 km altitude (Chapter 8).

The low orbit creates a significant technological challenge, because each individual satellite now moves at considerable velocity relative to the earth's surface, being at best, 'visible' from a particular point on the ground for a maximum of 15 minutes per orbit.

Like GSM, each of the proposed satellite systems, divides the coverage area (the earth's surface) into a number of *cells*, within each of which a number of radio channels permit individual users to make telephone calls. The difference is that the *cells* are now typically the size of a European country, so that the capacity in terms of traffic density is much less than for a terrestrial-based GSM system. The advantage, of course, of the satellite systems, is the global coverage and roaming capability. Figure 15.16 illustrates the concept in simple terms and Table 15.3 compares four of the systems proposed to go into service during the 1998–2000 timeframe.



**Figure 15.15** Inmarsat ship's antenna. The radome cover protecting a ship's INMARSAT telecommunications satellite antenna. (Courtesy of British Telecom)



**Figure 15.16** The Universal Mobile Telephone Service (UMTS)

**Table 15.3** Comparison of satellite mobile telephone systems

	Iridium	Globalstar	Odyssey	Teledscic
Consortium Leader	Motorola	Loran, Qualcomm	TRW-Matra	Bill Gates
Number of satellites	66	48	10	840
Orbit type and altitude	LEO (low earth, circular orbit), 780 km, 6 orbit paths	LEO (low earth, circular orbit), 1414 km, 14 orbit paths	MEO (medium earth, circular orbit) 10 355 km, 3 orbit paths	LEO (low earth orbit) 840 km
Satellite inclination	86.4°	47/65°	55°	
Minimum elevation	8°	10°	18°	40°
2.6 ms	4.7 ms	34.5 ms		
No of spot beams (cells/satellite)	48	6	163	
Multiplexing	FDMA/TDMA	FDMA/CDMA	FDMA/TDMA	TDMA/SDMA
Modulation	QPSK	QPSK	DPSK	
Bitrate	2.4/4.8 kbit/s	4.8/9.6 kbit/s	4.8 kbit/s	
Satellite switching	on-board	transparent	transparent	
Satellite weight	689 kg	262 kg	1130 kg	
Date of service	1998	1998	1999	2001

Resulting from these initiatives, new interest has been generated in creating a *universal mobile telephone service (UMTS)* by enabling *roaming* between cells of a land-based GSM-type mobile network and a satellite-based cell. The idea is that a single, but perhaps *dual-mode*, handset could register for terrestrial GSM network service in those countries where it was available, and where not, it could be reached by a satellite service. In this way, the coverage of the satellite-based system could be complemented by the higher traffic capacity of land-based systems in more telephone traffic-dense areas.



---

## *Cordless Telephony and Radio in the Local Loop (RILL)*

---

The rapid deregulation of telephone network services taking place during the 1990s has brought a large number of new public network operators to the market, each of which has an interest in optimizing the cost of customer connection to his network. Much interest, in particular, has been channelled into radio technologies (so-called 'radio-in-the-local loop' or 'wireless local loop', WLL), as these are seen as a quick and economic way to create new access infrastructure, bypassing the dependence on the established monopoly operators for 'last-mile' connections. In this chapter we discuss some of the most important technologies in this sector. We also discuss cordless telephone technology as a means for providing 'limited mobility' access to fixed networks.

### **16.1 THE DRIVE FOR RADIO IN THE LOCAL LOOP**

It was historically the case that a monopoly existed on both the public telephone service and the construction and operation of telecommunications transmission networks. The state-owned monopoly carrier had the sole right to lay cables in the street or construct radio transmission links. Although competition in public telephone network services may have been introduced in many countries, there has not necessarily been a relaxation of the transmission network monopoly. In consequence, the new telephone carriers (network operators) may be dependent on their strongest competitors for the supply of all transmission links. Thankfully for the new operators, if a little slowly, the national transmission monopolies are also being removed. Unfortunately, however, this does not immediately remove the dependence of the new operators on the ex-monopoly carrier, because the large base of established lineplant and investment is difficult for the new carriers to duplicate quickly. The best hope for them lies in the rapid construction of an overlay, radio-based infrastructure.

## 16.2 FIXED NETWORKS BASED ON RADIO TECHNOLOGY

Figure 16.1 illustrates the typical configuration of a new telephone network based largely on radio transmission links.

Traditional *point-to-point* (PTP) microwave radio technology is ideally suited for the long point-to-point links between switching centres (i.e. between local exchanges and regional switching centres and for the trunks between regional switching centres). There may be some regulatory and administrative matters to be resolved with respect to licensing of the required radio frequencies, but as the frequencies are required on a strict point-to-point basis and not over wider areas, this should be achievable both from a regulatory and planning point of view, because the number of links of this type is relatively small.

By contrast, there has been relatively little attention paid to radio connection of end customers by the monopoly network players, so that new effort needs to be applied to develop economic technology and to finding suitable radio frequency bands for this new application. Here, the nature of the connection is *point-to-multipoint* (PMP), requiring potential connection of many thousands of endstations with dynamic allocation of radio bandwidth. The radio path length need only extend to about 5 km and the endpoints are fixed, so this removes the need for much of the complexity of the GSM system (e.g. for *roaming* and *hand-off*) but the radio design is complicated by the physical properties of available radio bands (most having relatively short range and maybe requiring *line of sight* (LOS)). Further problems are posed by the difficult radio operating conditions of urban environments (*radio shadows*, *multipath*, *interference*).

For heavily used lines with bitrates above 2 Mbit/s, *point-to-point* (PTP) microwave remains the predominant method because of the strong signal strength needed to support high bitrates reliably over appreciable distances. This is best achieved with highly directional antennas, focussing the radio signal along a single path. Frequency

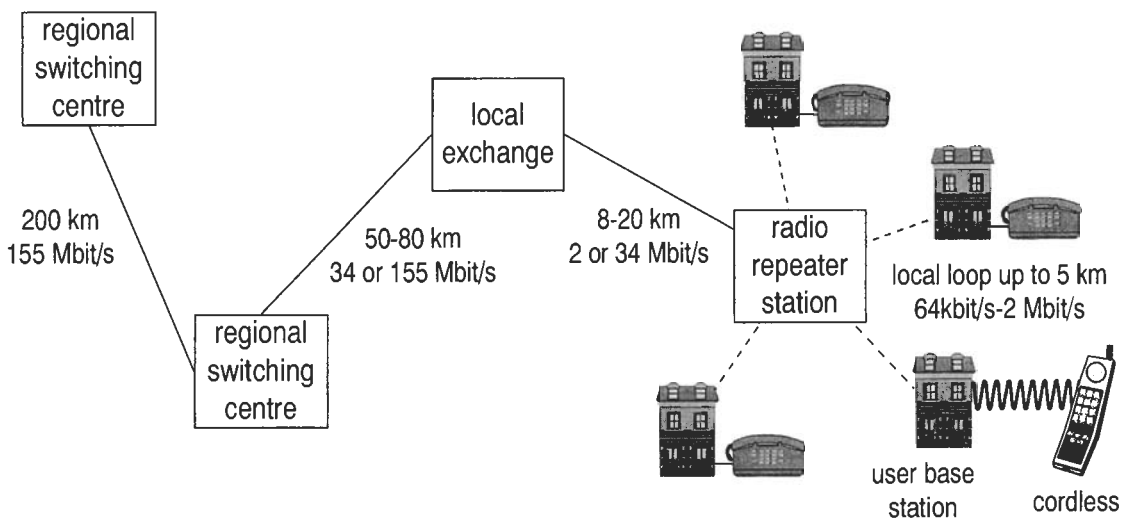


Figure 16.1 New telephone network structure based on radio technology

bands at 18 GHz, 23 GHz and 38 GHz are now allocated for so-called *shorthaul microwave* radio systems. The range of systems drops dramatically with higher frequency, so that while 15 km range is realistic within much of Europe for 18 GHz systems, 5–7 km is the reckoned range at 38 GHz.

There are many unallocated radio frequency ranges above 40 GHz, but the relatively short range of radio signals at these frequencies and the need for unimpeded *line of sight* between the antennae (because the radio waves, unlike at lower frequencies, are less capable of even slight *diffraction* around corners and past obstacles). Much attention is thus focussed on the radio range between 400 MHz and about 40 GHz. There have been three distinct technological approaches, but the different approaches are likely to converge. The three approaches are

- cordless telephony
- wireless ISDN
- shorthaul *point-to-point* (PTP) and *point-to-multipoint* (PMP) microwave radio

We discuss each in turn.

## 16.3 CORDLESS TELEPHONES

Cordless telephony is the term used to describe telephone sets connected to the ordinary (*fixed*) telephone network, but in which the handset communicates with the network by a radio transmission link instead of wires. The cradle part of a cordless telephone terminal acts as a radio transceiver or *base station* to connect a radio path to the handset, which also acts as a radio transceiver. The base station is connected to the public telephone network in the normal way. Figure 16.2 illustrates a typical cordless telephone configuration. The maximum range of these systems is typically 50 metres.

Cordless telephones were popular for some time in North America and Japan before they took off in Europe. The problem was that the European (CEPT) design specifications were more complex, making the products comparatively expensive. The exception was West Germany, where cordless phones were rented out by the Bundespost at little more than the rental cost of ordinary telephones.

Cordless telephones are very simple in comparison with cellular radio telephones, comprising a (duplex) two-way conversational radio channel, with a relatively simple signalling system. A major hurdle in the design of cordless telephones is ensuring that telephones in adjacent customers' premises do not interfere with one another and cannot be maliciously overheard. A customer is not prepared to pay for the next-door-neighbour's calls, made on the wrong base station. This may happen if a handset *interferes* with the base station next door, and was the main reason for the very strict CEPT specifications.

The advantage of cordless telephones is the freedom to carry them about the house, down the garden, around the workshop, so saving users from being away from the phone and not hearing the phone ring. Simple cordless telephones can be used only within range of their own base station. They are thus useless away from *home*, but make the customer more mobile about his own premises.



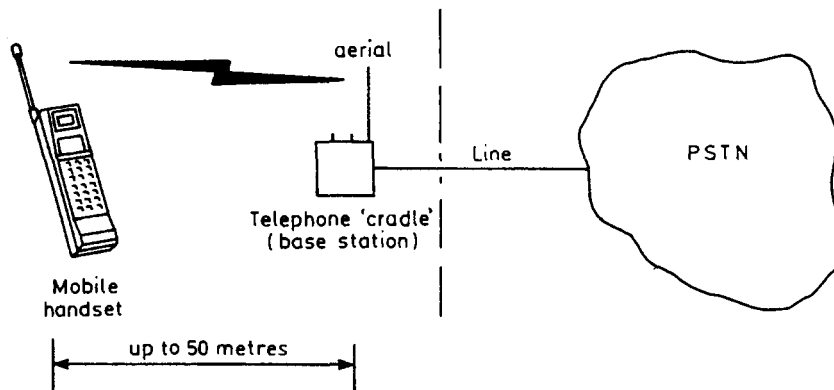


Figure 16.2 A cordless telephone

From basic cordless telephony (radio path within the end customer's premises) have evolved second and third generation technologies in which the base station is shifted to the public network operator's site. First came *telepoint* or *CT2* (2nd generation cordless telephony). *DECT* (digital European cordless telephony) followed.

## 16.4 TELEPOINT OR CORDLESS TELEPHONE 2 (CT2)

An extension of the idea of cordless telephones is the concept of *telepoint*, French *pointel* or *wide area cordless telephone*. In *telepoint* a new type of digital cordless telephone is used with a number of base stations. Besides the base station in his house, the customer has access to public *telepoint* base stations situated in well populated locations, such as airports, stations, and street corners (much as public payphones are located today). A *common air interface (CAI)* ensures compatibility of mobile handsets from various manufacturers with the *base* stations of the public network. Standing within 50–200 m of a telepoint, a caller with a telepoint handset is able to make outgoing calls into the public switched telephone network in a similar way to a cellular radio customer making an outgoing call, except that he may not move from one base station to another during the call. Incoming calls, however, are not possible other than at the home base station (i.e. the subscriber's home). Telepoint hardware includes a mobile handset and a number of base stations, each connected directly to the public switched telephone network, as Figure 16.3 illustrates.

To make a call, the handset sends a signal, including a special handset identity code, over a control channel to the base station, which confirms the identity and authorization of the user, and then allocates a radio channel in a way similar to that used in cellular radio. Onward connection of the call is made directly via the PSTN, applying dial tone, collecting dialled digits, etc., while the base station records call details for later billing of the customer. (There is one exception to this, and that is when the customer has installed a private base station in his own premises. In this case the customer pays for public network calls in the normal way as recorded by the PSTN operator.)

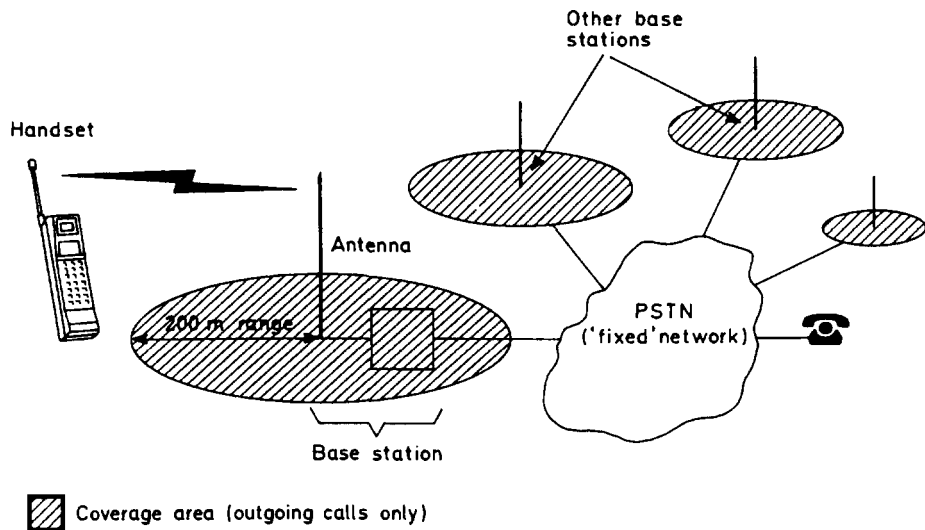


Figure 16.3 Telepoint service

As we have seen, *telepoint* or *second generation cordless telephone (CT2)* as it is also known is capable only of making outgoing calls. The technological problems of tracking the mobile handset location were not solved by CT2, so that incoming calls to users in roaming locations were not possible. There was talk of building radiopaging receivers into CT2 handsets, so that the users could be *paged* with a displayed telephone number to dial when next he was near a telepoint base station, but this would have made the cost of the handsets and their ongoing operation more than that of cellular telephone handsets.

A further problem was that CT2 did not provide for a *hand-off* procedure for moving between base station zones, so users were forced to stay in range of a single base station for the duration of each call. Thus a car needed to be parked in a *telepoint* car park, it could not be on the move.

Despite attempts at commercial service of CT2 in several countries, the CT2 standard failed, but the basic ideas and the technology survived in a third generation version, *DECT (digital European cordless telephony)*.

## 16.5 DECT (DIGITAL EUROPEAN CORDLESS TELEPHONY)

The DECT standards, developed by ETSI, have grown to be a sophisticated set, starting to rival GSM in terms of the degree of complexity. The initiative for their development grew from the desire to develop a *common air interface (CAI)* for digital *wide area cordless telephones*. Along the way, a number of other features have been built in

- security measures against unauthorized use of the handset and overhearing of conversation

- mobile station tracking so that incoming calls can be forwarded to the DECT telephone user, no matter where he is
- full handover of mobile stations from one cell to the next
- 64 kbit/s traffic carrying capability, to provide for correct functioning of ISDN data terminals over DECT
- OSI compatibility of the DECT protocols

Figure 16.4 illustrates the reference model of the DECT system.

The most important part of DECT is the radio common air interface, D3. This allows for the connection of *portable radio terminals (PT)* (the *portable part, PP*, of the system) to *fixed radio terminations (FT)* (being the *fixed part, FP*, of the system using a *cordless* (i.e. radiolink) connection. This interface can be used on its own in a similar manner to the *CAI (common air interface)* of CT2. Thus straightforward cordless telephones for home or office use are already being marketed for use by a single customer in his own premises. Meanwhile, for those with a public DECT network service subscription, use of the handsets in the *wide area* may also be possible.

The advantage of the DECT interface over predecesing cordless telephone technologies is the high speech quality afforded by a digital radio connection and the extra security measures added to guard against overhearing and unauthorized use, be it

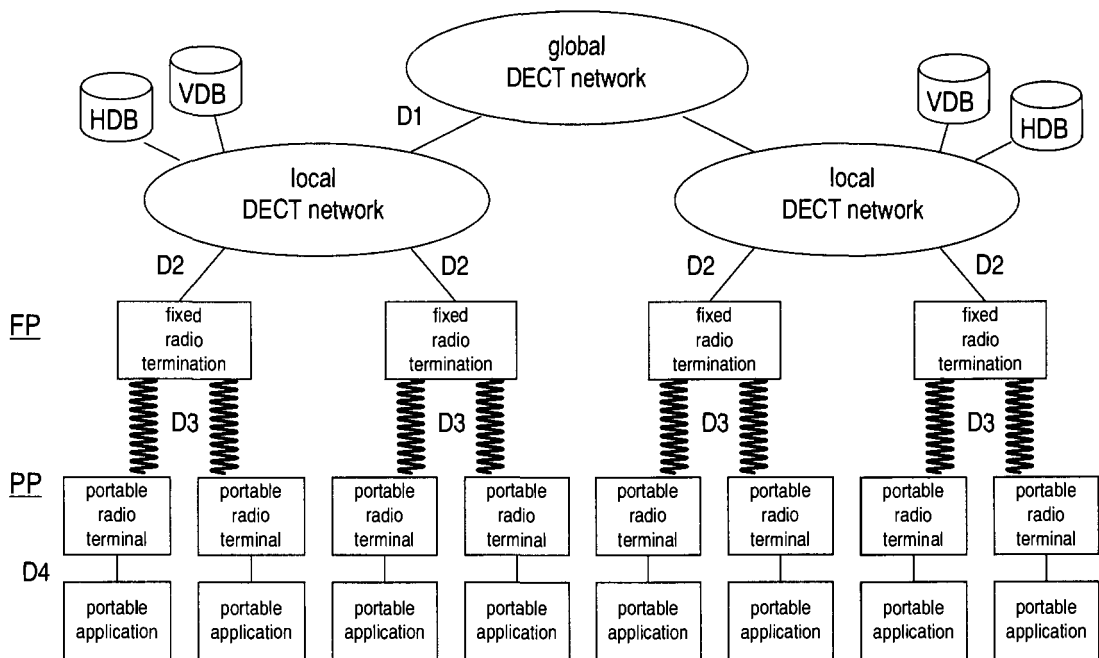


Figure 16.4 DECT reference model

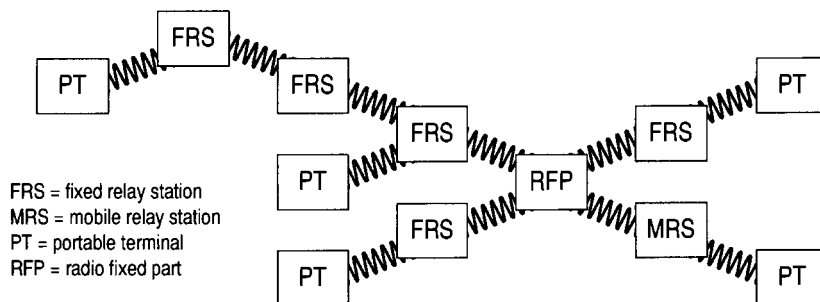
malicious or unintended. The extra security is afforded by means of data *encryption* and by *smart card* (so-called *DECT authorization module, DAM*) user identification in a manner similar to that employed by the GSM system (Chapter 15).

## 16.6 DECT HANDOVER

The interfaces D1 and D2 and the functions *HDB* (*home data base*) and *VDB* (*visitor data base*) are additional to those available in CT2. These support the ability to receive incoming calls in a wide area DECT network, also support *roaming* between cells. Each *fixed radio termination (FT)* controls a cell within a DECT radio network. Roaming between cells is controlled by a local network function comprising *home data bases (HDB)* and *visitor data bases (VDB)*, which perform similar functions to the *home location register (HLR)* and *visitor location register (VLR)* of the GSM system (Chapter 15). Unlike GSM, however, the *handover* in DECT is by means of *mobile controlled handover (MCHO)*, in which the mobile station alone decides when to handover and controls the process. This is claimed to lead to faster and more reliable handover. This method compares with the *mobile assisted handover (MAHO)* of GSM in which the *mobile switching centre* and *base stations* control the handover based on information provided by the mobile. The decision to initiate handover in the DECT system is based upon the mobile unit's measurement of the *RSSI* (*received signal strength indicator*), *C/I* (*carrier to interference*) and *BER* (*bit error rates*) of alternative signals.

## 16.7 THE RADIO RELAY STATION CONCEPT IN DECT

As the range of a single hop within the DECT system is relatively limited (typically 200 metres, although under ideal conditions with specific technical configurations several kilometres have been achieved), there has been a need to find a means of extending the range. The radio relay station concept allows for relaying of connections (i.e. concatenation of several radio links) to allow the *portable radio termination (PT)* to stray



**Figure 16.5** DECT fixed and mobile relay stations

somewhat further away from the *fixed radio termination* (or *radio fixed part*, *RFP*). Relay stations may be either *fixed relay stations*, *FRS*, or *mobile relay stations*, *MRS*, as Figure 16.5 illustrates. Up to three relay stations may be traversed, but the topology must be a star centred on the RFP.

The drawback of DECT *relaying* is that multiple radio channels are used to connect a single connection or call, making it impracticable for high traffic volume networks. In addition, the connection quality is likely to be degraded.

## 16.8 THE DECT AIR INTERFACE (D3-INTERFACE)

The DECT air interface is designed to be OSI-compliant (see Chapter 9). It therefore comprises layered protocols for *physical layer*, *medium access control* and *data link control* for both the *control-plane* (*c-plane*) and *user plane* (*u-plane*) as Figure 16.6 illustrates.

The *c-plane* protocol stack, as we discussed in Chapter 7, is used to set up and controlling connections (like telephone signalling). The *u-plane* protocol stack is that used during the *conversation phase* of a call or connection, to convey the user's speech or data. The lower layer management entity is the set of network management functions provided to monitor and reconfigure the protocols as necessary for network operation.

The characteristics of the *physical layer* of the radio interface are listed in Table 16.1. The multiple access scheme is based on TDMA, as illustrated in Figure 16.7.

A single slot may comprise either a *basic physical packet P32* (a full slot), a *short physical packet P00* (for a short signalling burst) or two half slots (*low capacity physical packet P08*).

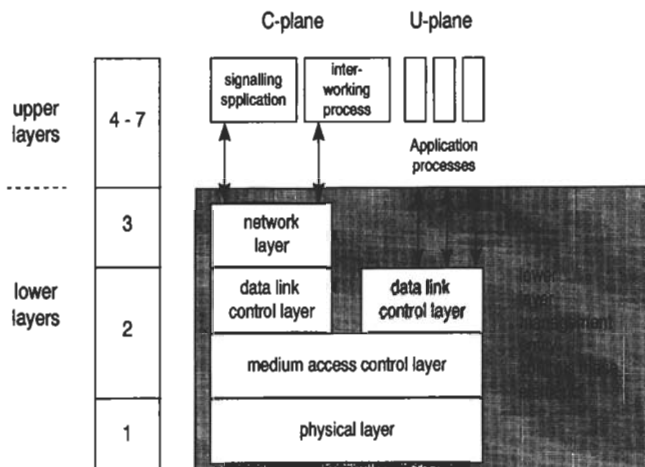
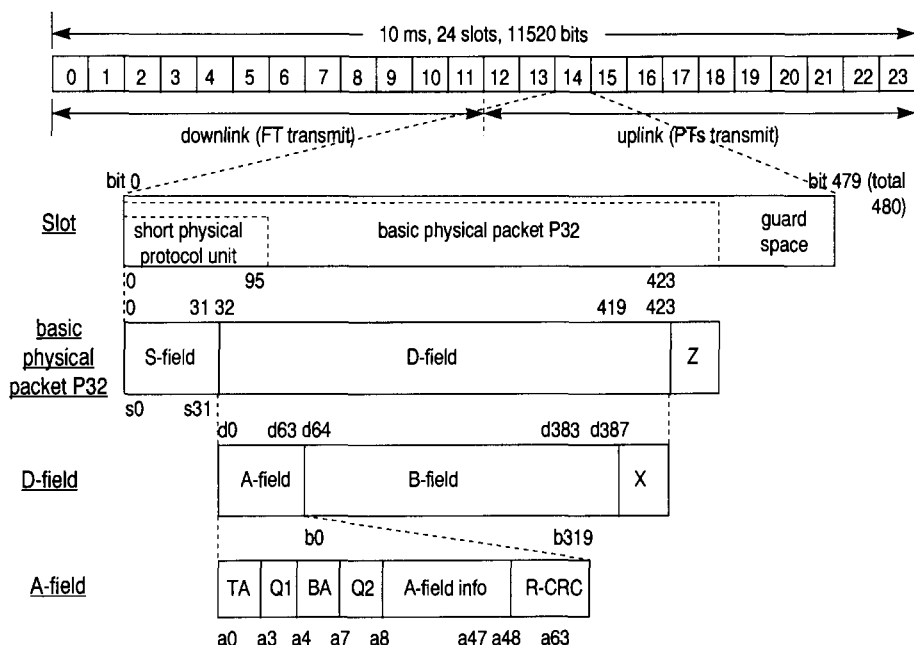


Figure 16.6 DECT protocol reference model

**Table 16.1** DECT air interface, physical layer

Radio band	1880–1900 MHz
Number of radio channels	10
Radio channel separation	1.728 MHz
Transmitter power (max)	250 mW
Channel multiplexing	TDMA (time division multiple access)
Duplex modulation	TDD (time division duplexing)
TDMA frame duration	10 ms
Timeslots per TDMA frame	24
Modulation	GFSK (Gaussian frequency shift keying)
Total bit rate	1152 kbit/s per cell
User channels	B channel: 32 kbit/s (user) A channel: 6.4 kbit/s (signalling)

**Figure 16.7** TDMA frame structure in DECT

The basic physical packet P32 is of 424 bytes length, subdivided into the *S-field* (for synchronisation) and the *D-field* (for carriage of data). The *D-field* is further subdivided into *A-* and *B-fields*, whereby the *A-field* is a permanent signalling channel (for *c-plane* protocol) and the *B-field* is the user data information filed (*u-plane* protocol). The various parameters within the *A-field* have the functions listed in Table 16.2.

**Table 16.2** DECT signalling parameters (A-field)

Parameter	Purpose
TA	A-field information type
Q1, Q2	Quality control bits used as handover criteria
BA	B-field information type
A-field information	The field used for carriage of MAC and higher layer c-plane protocol information
R-CRC	Cyclic redundancy check (for error detection)

Full 64 kbit/s ISDN *bearer channels* (i.e. user information channels for ISDN 64 kbit/s data) may be transmitted over DECT networks by the occupation of two B channels.

## 16.9 OTHER ISDN WIRELESS LOCAL LOOP SYSTEMS

Partly due to the scepticism about the suitability of DECT as a means for large scale mass market connection of fixed network customers to a telephone network, and partly due to the fact that DECT is presently (1997) only a European standard, other systems have also been developed for ISDN wireless local loop, aiming to provide for telephone and full 64 kbit/s connection service. These systems use a variety of different and as yet unstandardized techniques. Example technologies include those of *Ionica* (a British company aiming to offer a full scale telephone network across the UK – based on technology called *Proximity i* developed in conjunction with *Northern Telecom*, *NORTEL*), *Airspan* (a system developed by *DSC* in cooperation with British Telecom) and *Airloop* (a *Lucent Technologies* equipment developed by *Bell Laboratories* for use in the deregulating US and Dutch markets). Which of these systems or DECT survives in the long term remains to be seen. Critical will be the cost per user, as well as the technical system performance.

## 16.10 SHORThAUL POINT-TO-MULTIPOINT (PMP) MICROWAVE RADIO

Meanwhile, the manufacturers of traditional point-to-point microwave radio systems have not been idle. Several manufacturers have started developments of *point-to-multipoint* (PMP) systems of shorthaul microwave radio for use in the microwave band above 10 GHz. These developments foresee dynamic allocation of radio bandwidth within a cell, allowing a fixed or maybe even mobile end user to request various different bitrates (64 kbit/s or multiples thereof) on an on-demand (i.e. call-by-call) basis. These systems may not tap the initial market for ISDN radio in the local loop, but in a later generation they may be the obvious choice for broadband services, including radio in the local loop for *ATM* (*asynchronous transfer mode*).

---

## *Fibre in the Loop (FITL) and Other Access Networks*

---

The advent of optical fibre communication has coincided with a worldwide trend towards deregulation of public telecommunication network services. This has caused rapid heavy investment in optical fibre networks, including access networks for the connection of customers. This in turn has brought not only focus onto the development of new modulation and multiplexing technologies for use in conjunction with optical fibres themselves, but also the development of new techniques to enable the better usage of existing copper and coaxial cable access network lineplant, as incumbent operators attempt to make the best of the installed lineplant. In this chapter we review some of the most important of these new technologies.

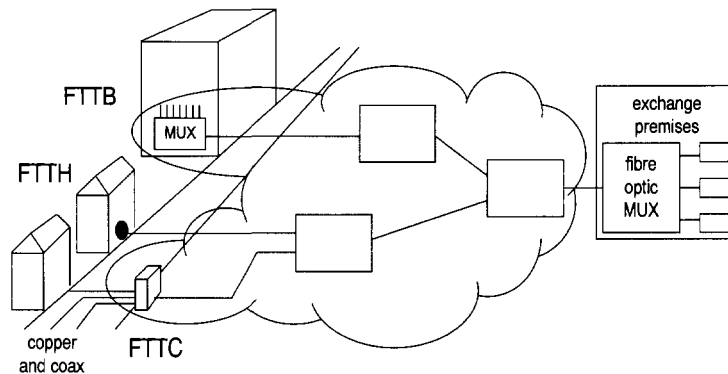
### 17.1 FIBRE ACCESS NETWORKS

A number of new terms have appeared to describe different initiatives developing solutions for the deployment of fibre cables in business and residential customer access networks. These can all be classified as various forms of *fibre in the loop (FITL)*. Subcategories of *FITL* are *fibre to the building (FTTB)*, providing direct fibre connection of business customers, office buildings of campus sites; *fibre to the home (FITH)*, providing *video on demand (VoD)*, cable television and telephone services to residential premises and *fibre to the curb (FTTC)*, whereby the fibre extends only as far as the streetside cabinet, from which existing copper or coaxial lineplant can be used to connect customer premises. Figure 17.1 illustrates these various concepts.

### 17.2 FIBRE TO THE BUILDING (FTTB)

The main driver for *FTTB* (*fibre to the building*) has been the boom in demand from business telecommunications users for line capacity. It is nowadays usually most economic for network operators to lay fibre optic cable directly into large business





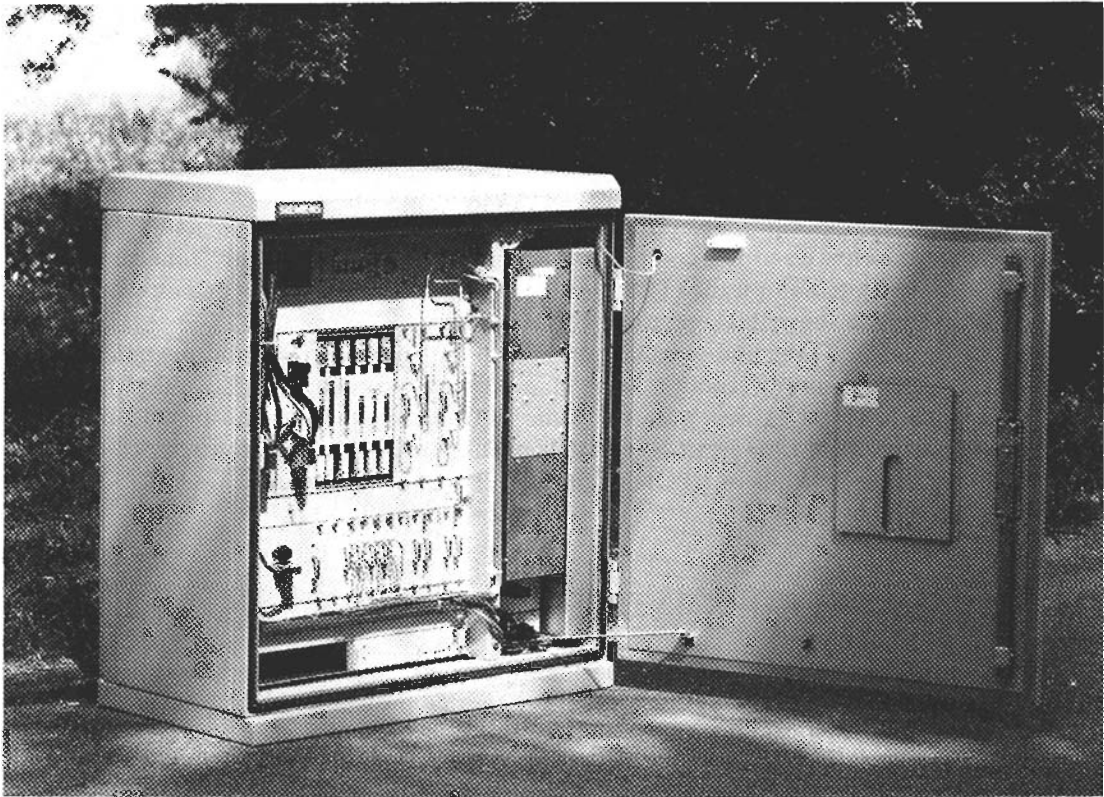
**Figure 17.1** Fibre in the loop (FITL)

premises rather than multiple pair copper cables. First, the fibre optic cable occupies far less valuable duct space; in addition, it does away with the need for amplifiers and other regenerative devices within the access network; finally, optical fibre provides plentiful capacity to meet future customer orders for bandwidth.

In the simplest realization of FTTB only a standard multiplexor and an *optical line terminating unit (OLTU)* are required in the customer's premises and at the exchange to provide a range of different line connections with different bitrates and line interfaces. A number of new metropolitan network operators have emerged which are building exactly such networks. *Metropolitan Fibre Systems (MFS)*, *City of London Telecommunications (COLT)* and *Teleport* are examples. Their networks consist of fibre optic access networks, built in redundant multiple ring topologies, using *SDH (synchronous digital hierarchy)* transmission technology (Chapter 13). They, and similar network operators have existing networks in many large cities across the globe. The in-building multiplexor at the customer site may either be dedicated to a single customer and installed in his offices, or in many cases be shared by a number of different tenants of a large office complex; in this case being installed in a small equipment room rented by the network operator within the building as a common attachment point.

### 17.3 FIBRE TO THE CURB (FTTC)

*Fibre to the curb (FTTC)* is a natural extension of the second case of *fibre to the building*. In *fibre to the curb* a shared multiplexor is installed in a streetside cabinet rather than in an equipment room on a customer's premises. From this point, existing copper lineplant is used to connect to individual customers. The main benefit of FTTC is the ability to rationalize the copper junction cabling (i.e. that between the streetside *distribution cabinets* and the exchange) as a first step in access network modernization, without requiring the upheaval or investment that would result from wholesale replacement of all copper lineplant.



**Figure 17.2** A streetside cabinet providing for fibre to the curb (FTTC) (*Courtesy of Siemens AG*)

The *OPAL* (optical access line) system of the Deutsche Telekom is an example of an FTTC system. The *OPAL* system was used extensively in the penetration and modernization of the former East German telephone network following the reunification of Germany in 1990.

## 17.4 FIBRE TO THE HOME (FTTH)

The cabling of fibre directly into residential customers' homes is usually carried out with the main objective of providing cable television or other video entertainment services like *video on demand* (*VoD*). Here, the emphasis of new *passive optical networks* (*PON*) has been the establishment of low cost, low maintenance access networks that do not require active electronic components installed in the street environment.

## 17.5 BROADBAND PASSIVE OPTICAL NETWORK

*Broadband Passive Optical Network* (*BPON*) is a simple approach to broadband networking with a very clearly focussed commercial application. It is a technology

proposed and developed by British Telecom that is intended to bring fibre to the home. Basically, it is a network composed of monomode fibres (either in a ring or star topology) connecting telephone exchanges and peoples homes, allowing not only basic telephony but also the 'broadcasting' of cable television and video programmes.

The foundation of BPON is a technique known as *TPON* (*telephony passive optical network*). This is the method by which telephone services are provided in residential homes by means of fibre connected back to the exchange (again in either a ring or star topology). Optical couplers enable the various fibre distribution joints to be made without active electronic components (Figure 15.3). Individual calls from customers are time division multiplexed (TDM) at the exchange and selectively demultiplexed by the appropriate subscriber's receiver.

By using TDM and a technique known as *wavelength division multiplex* (*WDM*, basically the use of another laser of a different wavelength light), other broadband signals may be carried over the same fibre network. Thus the broadcast of cable television and video services is possible simultaneously with the telephone operation. This is the principle of BPON (Figure 17.3).

## 17.6 ACCESS NETWORK INTERFACES

The emergence of new active technology in the access network between customer premises and the exchange site has naturally brought with it new problems and opportunities. The problems arise from the need to devote effort to standardization of new interfaces, the opportunity is the new service functionality thereby made possible, together with the scope for network restructuring and cost optimization.

Two types of interface are now being addressed by standardization work on transmission technology for the access network. These are *local exchange (LE)* to *access network (AN)* interfaces (designated *V5 interfaces* by ETSI) and the *subscriber-network interface (SNI)*. Figure 17.4 illustrates these interfaces.

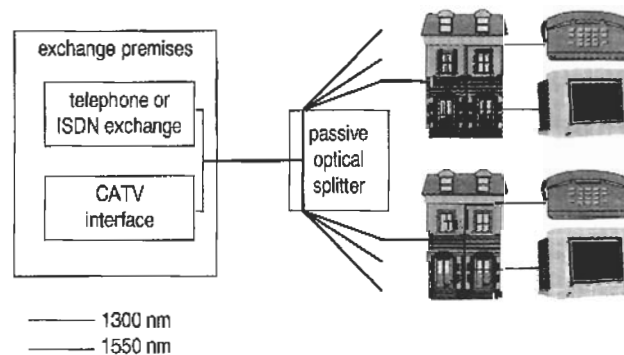


Figure 17.3 Broadband passive optical network (BPON)

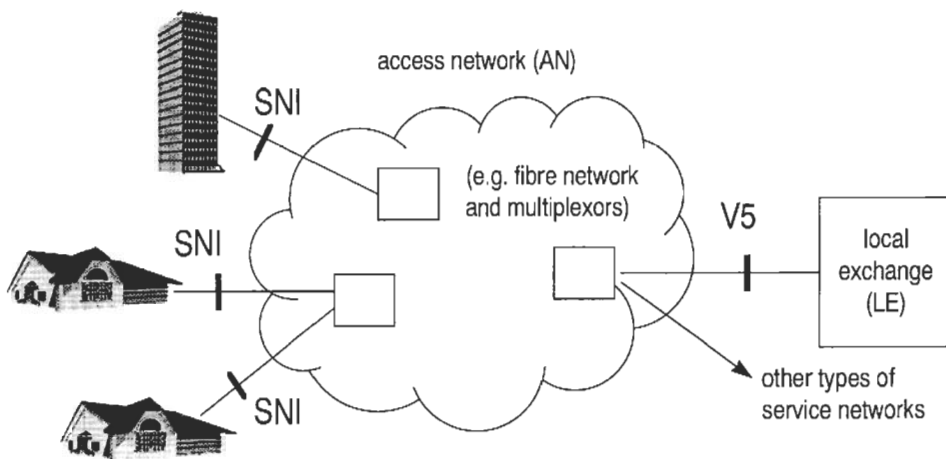


Figure 17.4 Access network interfaces

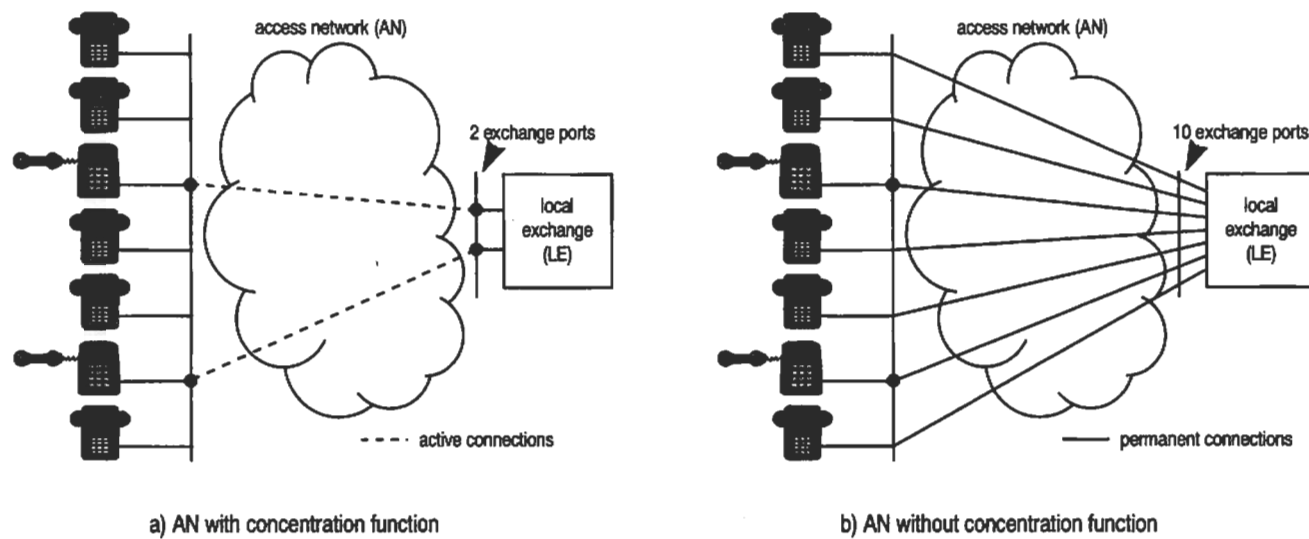
## 17.7 ETSI V5 INTERFACES

In conjunction with the modernization of the East German telephone network and the introduction of its *OPAL* (*optical access line*) technology, Deutsche Telekom recognized the potential for savings in access network lineplant, in the number of customer ports needed on telephone exchanges and in the number of telephone exchanges needed to supply a given region. This could be done by inclusion of *concentration* functions within the *OPAL* network. This led to the development of the ETSI V5 interfaces.

As Figure 17.5(a) illustrates, the access network need only support sufficient connections across itself for the actual number of telephone calls in progress. Historically, copper access networks had provided a permanent connection line for each end user (Figure 17.5(b)). This configuration requires many more connections within the *access network (AN)* and many more *local exchange (LE)* ports.

In the example of Figure 17.6, ten user terminals are connected to the local exchange. It is assumed that only a maximum of two of these terminals are in use at any one time. In the case of Figure 17.5(a), a concentrating function (i.e. simple switching function) within the access network ensures that only two through connections are required to be carried and only two ports are required at the exchange. In Figure 17.5(b), no *concentration* is undertaken by the access network, so that ten connections and ten exchange ports are necessary.

Before the access network can undertake the concentration function, a new signalling procedure must first be defined, because it would otherwise no longer be possible for the exchange to know (merely by port of origin) which customer was wishing to make a call. The local exchange requires this information so that the correct customer is billed for the call. Similarly, for incoming calls, the local exchange must be able to signal to the access network which destination customer is to be connected. This signalling is defined in the ETSI specifications for its V5.1 and V5.2 interfaces.



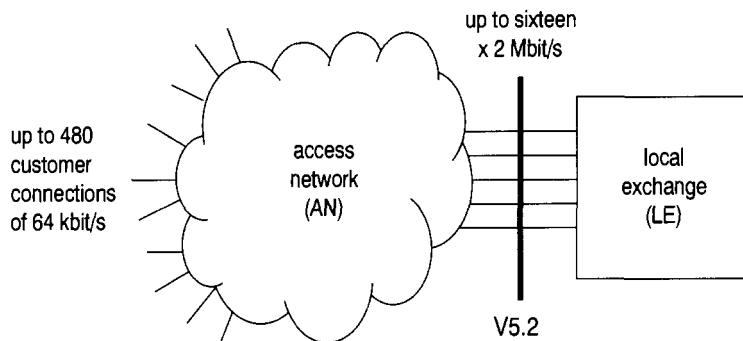
**Figure 17.5** The effect of a concentration function in the access network

## 17.8 V5.2 INTERFACE

The V5.2 interface is defined in ETSI standard ETS 300 347. It defines a method for connecting up to 480 customer lines of 64 kbit/s capacity (480 simple telephone lines, 240 ISDN basic rate access lines or 16 ISDN primary rate access lines or an appropriate mix thereof) via an *access network (AN)* to a telephone or ISDN *local exchange (LE)*. The *access network* may be connected using up to sixteen 2 Mbit/s lines to the *local exchange*. Figure 17.6 illustrates the V5.2 interface.

As the V5.2 interface provides for a concentration function (like Figure 17.5(a)) to be undertaken by the access network, the number of traffic-carrying channels at the V5.2 interface (between AN and LE) may be less than the number of customer connections required from the AN to customer premises. The protocol of V5.2 is complex and not covered in detail here. It bears some resemblance to ISDN D-channel signalling (ITU-T Q.931) and is OSI model compliant. Main elements and terminology of the interface are as follows.

- **Bearer channel:** this is a channel with a bitrate of 64 kbit/s (or an integral multiple thereof) which is used to carry customer telephone signals or ISDN data services.
- **Bearer channel connection (BCC) protocol:** this is a protocol which allows the LE to control the AN in the allocation of *bearer channels*. It is one of the types of information which may be carried by an *information path*.
- **Communication path (c-path):** this is the path needed to carry signalling or data-type information across the V5.2 interface. Apart from the *BCC protocol*, a *c-path* is also used for carriage of the ISDN D-channel signalling and packet or frame data originated by the various customer ISDN connections.
- **Communication channel (c-channel):** this is a 64 kbit/s allocation at the V5.2 interface configured to carry a *communication path*.
- **Logical communication channel (logical c-channel):** this is a group of one or more *c-paths*.



**Figure 17.6** V5.2 interface between local exchange and access network

- Physical communication channel (physical c-channel): this is an actual 64 kbit/s timeslot allocated at the V5.2 interface for carrying logical *c-channels*. A *physical c-channel* may be configured for communication and signalling and may not be used to carry *bearer channels*.
- Active c-channel: this is a *physical c-channel* which is currently carrying a *logical c-channel*. When not carrying a *logical channel*, the same *physical c-channel* becomes a *standby c-channel*.
- Standby c-channel: this is a *physical c-channel* which is not currently carrying a logical c-channel.

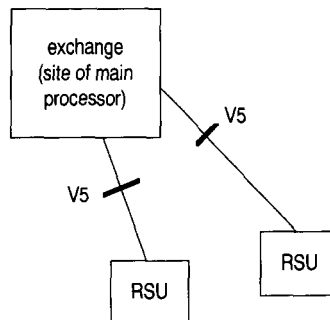
Thus the 64 kbit/s *timeslots* traversing the V5.2 interface are subdivided into *bearer channels* and *c-channels* by assignment (i.e. when the network is configured). The *bearer channels* serve to carry user telephone and ISDN or data connections. The *c-channels* serve (on an as-needed basis) to carry the *BCC protocol* for allocation of *bearer channels* to individual calls and to carry the ISDN D-channel signalling and data information between the end user terminal and the local telephone or ISDN exchange.

## 17.9 V5.1 INTERFACE

The V5.1 interface is a simpler version of the V5.2 interface in which the concentration feature (Figure 17.5(a)) is not included. V5.1 should be seen as the first step to V5.2. It allowed the adoption of new generation access network technology while the full specification and development of the concentration function (V5.2) took place.

## 17.10 SIGNIFICANCE OF THE V5.x INTERFACES

The V5.1 and V5.2 interfaces (or V5.x interfaces, as they are collectively known) provide for a standard means of connecting *remote switching units (RSUs)* of ISDN or telephone exchanges back to a central main exchange site (Figure 17.7).



**Figure 17.7** Use of a V5 type interface to support remote switching units

A network topology comprising central main exchange sites and remote switching units interlinked by standardized interfaces (V5) has significant benefits for public telephone network and ISDN operators. First, the number of exchange processor sites may be dramatically reduced (typically to a number of tens in a given country). This has significant investment and operational cost benefits. Second, the *remote switching units* (*RSUs*) may be purchased from multiple vendors, thus giving the network operator more leverage on price for these devices, which are needed in relatively high volume. This may lead to reluctance on behalf of the switch equipment manufacturers to developing it.

## 17.11 RE-USE OF EXISTING COPPER ACCESS NETWORKS

The boom in demand for 2 Mbit/s and higher bit rate connection services created by optical fibre technology has also had a spin-off in stimulating the development of technologies which attempt to re-use the existing copper and coaxial cable infrastructure. Three new technologies of particular interest are

- *HDSL (high bitrate digital subscriber line)*
- *ADSL (asymmetric digital subscriber line)*
- *HFC (hybrid fibre/coax networks)*

We discuss them in turn.

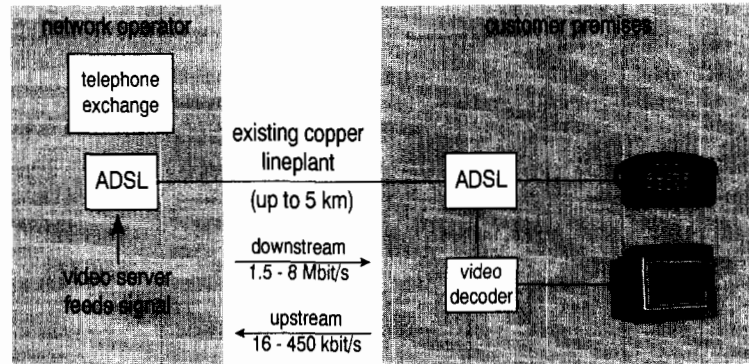
## 17.12 HDSL (HIGH BITRATE DIGITAL SUBSCRIBER LINE)

HDSL is a technique providing for full duplex 2 Mbit/s access lines using two or three copper pairs. The technique is particularly designed to serve high speed business user needs over distances up to 3 km without having to replace the copper access network or lay new lineplant. As well as being used to provide the complete access line from customer site to exchange building, HDSL is also likely to play an important role as a complement to *FTTC (fibre to the curb)* networks. In such usage HDSL provides for the final few metres from the FTTC street cabinet into the customer premises, thus potentially saving the need for a new cable or cableduct into the customer premises.

## 17.13 ADSL (ASYMMETRIC DIGITAL SUBSCRIBER LINE)

ADSL uses technology similar to HDSL, but instead of providing 2 Mbit/s bitrates in both directions, an asymmetric pair of bitrates are provided. *Downstream* (i.e. from the exchange to the customer) a high bitrate of between 1.5 Mbit/s and 8 Mbit/s is intended to be provided for broadcast and *video-on-demand (VoD)* services, as well as telephony.





**Figure 17.8** ADSL (asymmetric digital subscriber line)

*Upstream* (i.e. from customer to exchange) the bitrate provided is much lower (between 16 and 450 kbit/s). This bitrate is only intended to be sufficient for telephony and for control of the network services (e.g. to say which video should be delivered). Figure 17.8 illustrates ADSL.

## 17.14 HYBRID FIBRE/COAX (HFC) NETWORKS

There is considerable interest amongst coaxial cable TV companies to upgrade their networks for the needs of coming *interactive video* and *multimedia* services, and in the short term simply to offer public telephone service in addition to television broadcast service to their customers. This has led to a number of developments for telephony over coaxial cable TV networks and integration of coaxial cable networks (for attachment of customer premises) into fibre networks. These are sometimes referred to as *fibre to the curb* (FTTC) technologies, sometimes more specifically as *HFC* (hybrid fibre/coax).

PART 3

---

# **MODERN DATA NETWORKS**

---



---

# Packet Switching

---

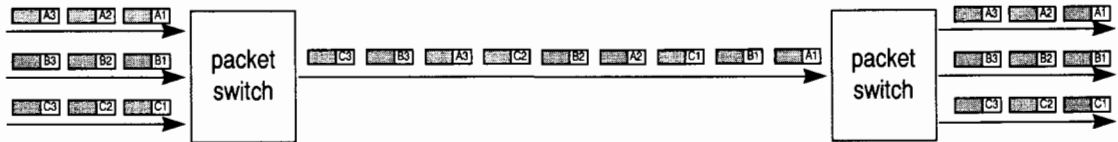
Packet switching emerged in the 1970s as an efficient means of data conveyance. It overcame the inability of circuit-switched (telephone) networks to provide efficiently for variable bandwidth connections for bursty-type usage as required between computers, terminals and storage devices. In this chapter we discuss the basics of packet switching and ITU-T's X.25 recommendation, nowadays the worldwide technical standard interface to packet-switched networks. We then also go on to discuss the IBM company's SNA (systems network architecture), a proprietary form of packet switching, important because of its dominant role in IBM computer networks.

## 18.1 PACKET SWITCHING BASICS

*Packet switching* is so-called because the user's overall message is broken up into a number of smaller packets, each of which is sent separately. We illustrated the concept in Figure 1.10 of Chapter 1. Each packet of data is labelled to identify its intended destination, and *protocol control information (PCI)* is added, as we saw in Chapter 9, before it is sent. The receiving end re-assembles the packets in their proper order, with the aid of *sequence numbers* and the other PCI fields. Each packet is carried across the network in a *store-and-forward* fashion, taking the most efficient route available at the time.

Packet switching is a form of *statistical* multiplexing, as we discovered in Chapter 9. Figure 18.1 illustrates how a link within a packet switching network is used to carry the jumbled-up packets of various different messages and the use of the information carried in the packet *header* to sort arriving packets at the destination end into the separate *logical channels*, *virtual circuits (VCs)* or *virtual calls (VCs)*.

Transmission capacity between pairs of nodes in a packet-switched network is generally not split up into rigidly separate physical channels, each of a fixed bandwidth. Instead, the entire available bandwidth between two nodal points (switches) in the network is bundled together as a single high bitrate pipe, and all packets to be sent between the two endpoints of the link share the same pipe (Figure 18.1). In this way, the entire bandwidth (i.e. full bitspeed) can be used momentarily by any of the *logical channels* sharing the connection. This means that individual packets are transported more quickly and bursts of transmission can be accommodated.

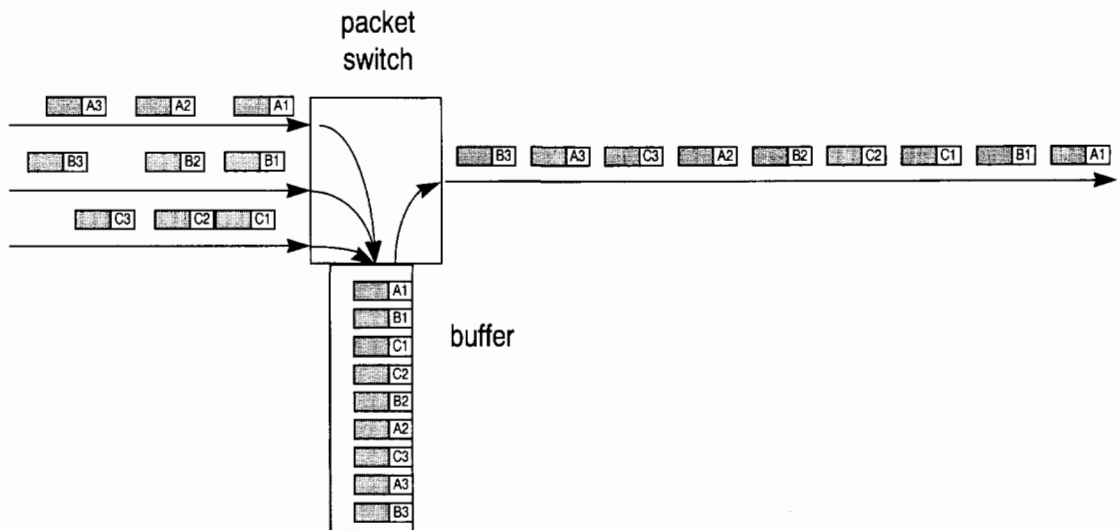


**Figure 18.1** The statistical multiplexing principle of packet switching

A problem arises when more than one or all *logical channels* try to send packets at once. This is accommodated by *buffers* at sending and receiving ends of the connection as shown in Figure 18.2. These delay some of the simultaneous packets for an instant until the line becomes free.

By use of buffers as shown in Figure 18.2, it is possible to run the transmission link at very close to 100% utilization. This is achieved by sharing the capacity between a number of end devices (each with a *logical channel*). The statistical average of the total bitrate of all the *logical channels* must be slightly lower than the line bitrate so that all packets may be carried, but at any individual point in time the buffers may be accumulating packets or emptying their contents to the line.

Packet switching is able to carry *logical channels* of almost any average bitrate. Thus a 128 kbit/s trunk between two packet switches might carry 6 logical channels of mixed and varying bitrates 5.6 kbit/s, 11.4 kbit/s, 12.3 kbit/s, 22.1 kbit/s, 28.7 kbit/s, 43.0 kbit/s and still have capacity to spare. This compares with the two channels which a telephone network would be able to carry using the same trunk capacity. (The excess capacity of the telephone channels simply has to be wasted, and the other four channels cannot be carried.)



**Figure 18.2** The use of a buffer to accommodate simultaneous sending of packets by different logical channels

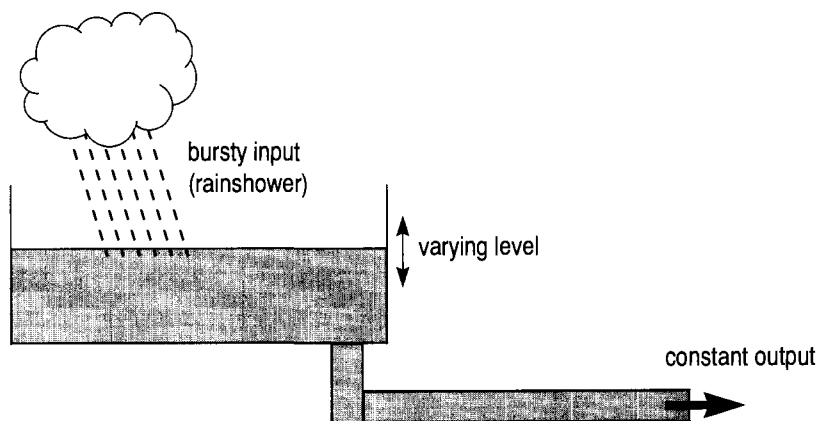
## 18.2 TRANSMISSION DELAY IN PACKET-SWITCHED NETWORKS

When using the trunks in a packet-switched network at very close to full utilization, very large buffers are required for each of the *logical channels*, to smooth out the bursts from individual channels into a smooth output for carriage by the line. (This is rather like having a very large water reservoir, collecting water during showers of rain, and varying in water depth, but always capable of outputting a constant volume of water for municipal use (Figure 18.3). The water reservoir is analogous to the data buffers, the showers of rain to the bursts of data information, and the constant output to the information carried by the line.)

We can make sure that the packets accumulated in the buffer are despatched on a *first-in-first out (FIFO)* basis to fairly share out the *queueing delays* which result, but it is critical to ensure that the queueing delay does not become unacceptably long. The chance of a very long delay is much greater when close to 100% utilization of the line is expected. (Imagine waiting in line for a bus, all of the seats of which had to be full before it pulled away; either the bus doesn't come very often, or there is a very long queue to ensure that all the seats can be filled).

A certain amount of queueing delay caused by buffering is not noticeable to computer users (a  $\frac{1}{4}$  second is a very long queueing delay in packet switching network terms). Even if a typed character did not appear on the computer screen until a  $\frac{1}{4}$  second after hitting the keyboard, the user is unlikely to notice. A *variation* in the delay (sometimes a  $\frac{1}{4}$  second, and sometimes no delay) is also unimportant. (The fact that some characters appear on the screen more quickly than the  $\frac{1}{4}$  second maximum delay will not be noticed.) On the other hand, once the average delay becomes much longer, then computer work may become frustrating, so that much longer queueing delays are unacceptable.

There is an entire statistical science used to estimate queueing delays. The most important formula is the *Erlang call-waiting formula*, which we will discuss in Chapter 30. In simple terms, however, the unacceptability of long queueing delays means that the



**Figure 18.3** The delay caused by buffering (analogy of a water reservoir)

trunks in packet-switching networks may not be utilized at 100%. Typical acceptable maximum utilization is around 50%. Despite this fact, they are still more efficient than circuit-switched networks to carry *data traffic* (i.e. information between computers).

### 18.3 ROUTING IN PACKET-SWITCHED NETWORKS

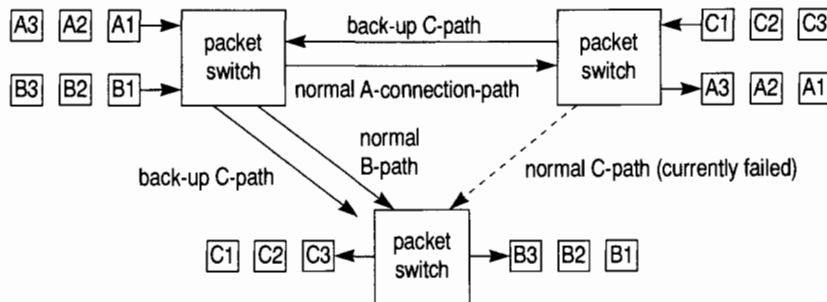
Packets are routed across the individual paths within the network according to the prevailing traffic conditions, the link error reliability, and the shortest path to the destination, according to one of two main techniques, so-called *path-oriented routing* and *datagram routing*. The routes chosen are controlled by the (layer 3) software of the packet switch, together with routing information pre-set by the network operator.

*Path-oriented routing* is nowadays the most common technique. In *path-oriented routing*, a fixed path is chosen for a given *logical channel* (i.e. *virtual circuit, VC*) at the time of call set up. The path itself is chosen based on the current loading of the network and the available topology. In the case of the *virtual call* (i.e. *switched virtual connection*) service, the required destination of the path is indicated using an X.121 address carried by a layer 3 packet, called the *call set-up packet*. This packet has the equivalent function to the dialled digits of a telephone number in setting up a telephone call.

Should any link in the path become unavailable during the course of the call (say because of a transmission failure), then an alternative path is sought, without breaking the connection. The packets are stored for a short while, while the new path is found and then sent over this path. (Figure 18.4).

The advantage of *path-oriented routing* is that the packets pertaining to a given *logical connection* all take the same path, all suffer about the same amount of queuing delay in the buffers and arrive pretty much in the same order as they were sent (allowing for any lost on the way). This makes the job of *re-sequencing* the packets at the receiving end much easier, as well as the job of directing the packets through the network. It also leads to more predictable delay performance for the end user or *computer application*. The packet switching network components can therefore be relatively simple and cheap.

The disadvantage of *path-oriented routing* is the inability of the network on a more immediate basis to employ alternative routing to better utilize the network as a whole



**Figure 18.4** Circumventing a transmission link failure using path-oriented routing

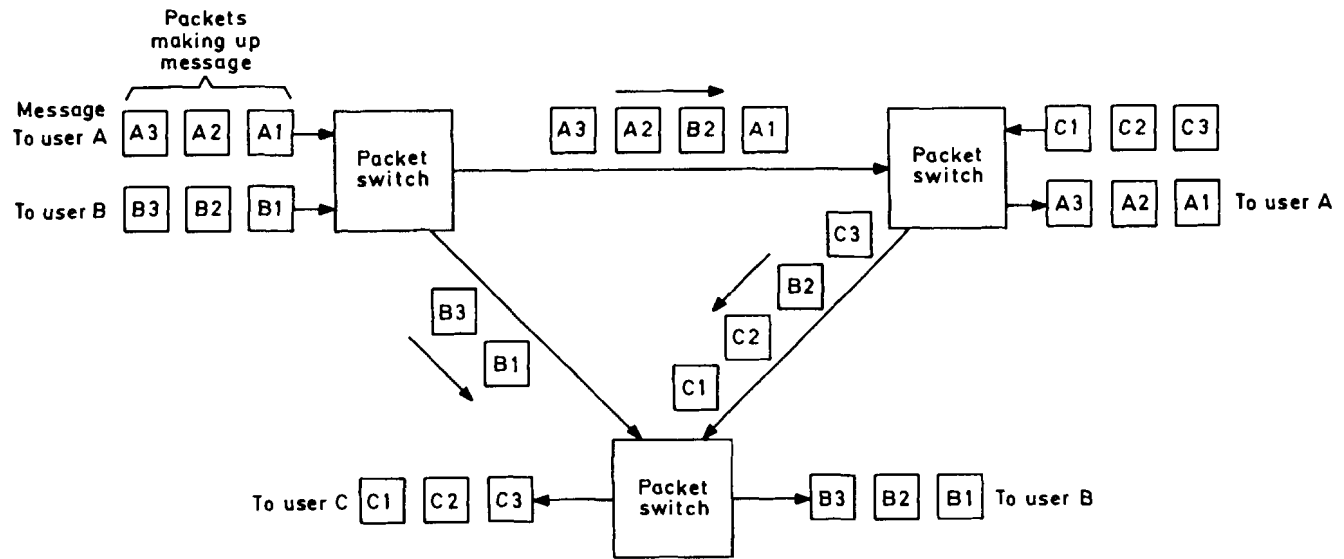


Figure 18.5 Datagram-mode packet switching



during periods of sudden surge in demand resulting from simultaneous packet bursts by many logical channels sharing the same path. The second type of routing, *datagram routing*, allows for more dynamic routing of individual packets (Figure 18.5), and thus has the potential for better overall network efficiency. The technique, however, requires more sophisticated equipment, and powerful switch processors capable of determining routes for individual packets.

Packet switching gives good end-to-end reliability, with well-designed switches and networks it is possible to bypass network failures (even during the progress of a *call*). Packet switching is also efficient in its use of network links and resources, sharing them between a number of calls, thereby increasing their utilization.

## 18.4 ITU-T RECOMMENDATION X.25

Most packet-switched networks use the protocol standards set by ITU-T's recommendation X.25. This sets out the manner in which a *data terminal equipment* (DTE) should interact with a *data circuit terminating equipment* (DCE), forming the interface to a packet-switched network. The relationship is shown in Figure 18.6.

The X.25 recommendation defines the protocols between DTE (e.g. personal computer or computer *terminal controller* (e.g. IBM 3174)) and DCE (i.e. the connection point to a *wide area network*, *WAN*) corresponding to OSI layers 1, 2 and 3 (Figure 18.7) which we learned about in Chapter 9).

The physical connection may either be X.21 (digital leaseline) or X.21 bis (V.24/V.28 modem in conjunction with an analogue leaseline: Chapter 9). Alternatively, the X.31 recommendation (Chapter 10) specifies how the physical connection (DTE/DCE) may be achieved via an ISDN (integrated digital services network). Finally, recommendation X.32 specifies the use of a dial-up connection for a *packet mode connection* via the telephone or ISDN network to an X.25 packet exchange.

The X.25 recommendation itself defines the OSI layer 2 and layer 3 protocols. These are called the *link access procedures* (*LAPB* and *LAP*) and the *packet level interface*. The *link access procedure* assures the correct carriage of data across the link connecting DTE

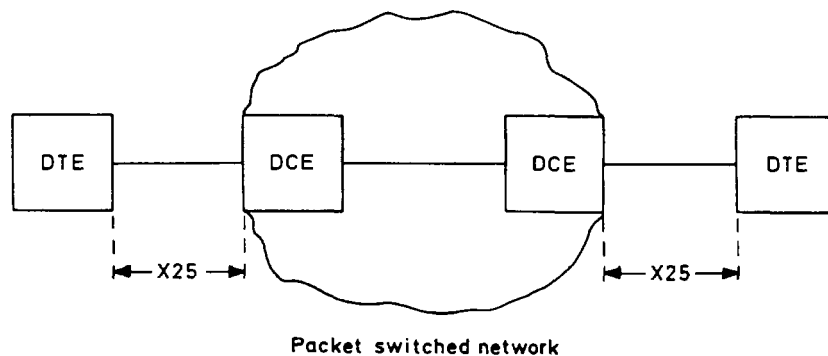


Figure 18.6 The X.25 interface to packet switched networks

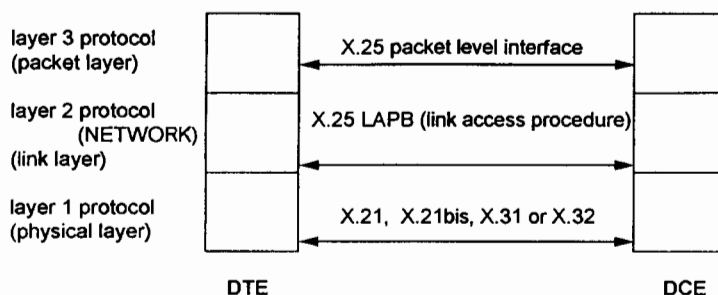


Figure 18.7 OSI layered model representation of ITU-T recommendation X.25

to DCE and for multiplexing of *logical channels*; the *packet level interface* meanwhile guarantees the end-to-end carriage of information across the network as a whole.

The *LAPB* (*link access procedure balanced*) protocol provides for the ISO *balanced class of procedure* and also allows for use of multiple physical circuits making up a single logical link. *LAP* is an older and simpler procedure only suitable for single physical circuits without balanced operation. The link access procedures use the principles and terminology of *high-level data link control (HDLC)* as defined by ISO. This procedure ensures the correct and error-free transmission of data information across the link from DTE to DCE. It does not, however, enable the DCE (in the form of a *data switching exchange, DSE*) to determine where the information should be forwarded to within the network or ensure its correct and error-free arrival at the distant side of the packet network. This is the job of the OSI layer 3 protocol, the X.25 packet level interface.

During the set-up of a *switched virtual circuit (SVC)*, also called a *virtual call (VC)*) it is a level 3 *call set up packet* which delivers the DSE the data network address of the remote DTE. Level 3 *packets* confirm the set-up of the connection to the initiating DTE and then pass end to end through the network, allowing user data to be carried between the DTEs.

A packet of data carried by the X.25 protocol may be anything between three and about 4100 octets (bytes: 8 bits). Up to 4096 alphanumeric characters of *user information* may be carried in a single packet.

In slang usage, many people refer to 'X.25 networks'. In general they mean packet switching networks to which X.25 compliant DTEs may be connected, for recommendation X.25 describes only the *UNI* (*user-network interface*, see Chapter 7). The X.25 protocol allows DTEs made by any manufacturer to communicate across the network. You should not be tempted into believing that the protocol used between the various packet *data switching exchanges (DSEs)* within the network is also X.25. Generally, packet-switched data networks are built from a number of individual exchanges, but all of them provided by the same manufacturer. The protocol used for the carriage of the data between the exchanges is normally an X.25-like, but enhanced 'proprietary' protocol. Examples include those used by *Northern Telecom (Nortel)*, *Telenet*, *BBS*, *Tymnet* and France's *Transpac*.

Proprietary trunk protocols typically allow the carriage of sophisticated network management and charging information back to the network control centre. In addition,

they may allow for dynamic adjustment of the traffic paths taken through the network, so giving better overall network performance during heavy traffic loading.

Where separate packet-switched sub-networks (provided by different manufacturers) need to be interconnected, the X.75 (NNI) protocol is used. We discuss this later in the chapter.

## 18.5 THE TECHNICAL DETAILS OF X.25

X.25 was one of the first data protocols to be well defined and standardized. As such it has formed the basis on which later data transport protocols have been developed. Understanding the principles in detail will give the reader a very good understanding of all other data switching protocols, all of which use similar principles. There thus follows a very detailed description.

## 18.6 X.25 LINK ACCESS PROCEDURE (LAP AND LAPB)

The *link access procedure* can be performed either in the basic mode ( $B = 0$ , called LAP) or in the more advanced *balanced mode* ( $B = 1$ , called LAPB). Nowadays the LAPB mode is more common.

There are two forms of LAPB; the basic form is called LAPB *modulo 8*, the extended form is called LAPB *modulo 128*. Only the modulo 8 form is universally available. The difference between the two forms is only the maximum value of the sequence number given to consecutive packets before resetting to value '0'. LAPB allows for data *frames* to be carried across a physical layer connection between a DTE and a DCE. The frame is structured in the manner shown in Figure 18.8.

The *flag* is a *delimiter* between frames. The *address* (perhaps confusingly named, as we discovered in Chapter 9) is a means of indicating whether the frame is a *command* or a *response frame*, and whether control is with the DTE or DCE. It is coded as shown in Table 18.1.

Flag 01111110	Address (8 bits)	Control (8 bits)	Information (N bits)	Frame Check Sequence (16 bits)	Flag (of next frame)
------------------	---------------------	---------------------	-------------------------	--------------------------------------	-------------------------

Figure 18.8 X.25 LAPB modulo 8 frame

Table 18.1 X.25 LAPB address field coding

	Single link procedure	Multiple link procedure
command from DCE to DTE	address A – 11000000	address C – 11110000
response of DTE to DCE	address A – 11000000	address C – 11110000
command from DTE to DCE	address B – 10000000	address D – 11100000
response of DCE to DTE	address B – 10000000	address D – 11100000

**Table 18.2** X.25 LAPB modulo 8 control field command and response coding

Format	Command	Response	bit 1	bit 2	bit 3	bit 4	bit 5	bit 6	bit 7	bit 8
Information transfer	I (information)		0		N(S)		P		N(R)	
Supervisory	RR (receive ready)		1	0	0	0	P/F		N(R)	
	RNR (receive not ready)		1	0	1	0	P/F		N(R)	
	REJ (reject)		1	0	0	1	P/F		N(R)	
Unnumbered	SABM (set asynchronous balanced mode)		1	1	1	1	P	1	0	0
	DISC (disconnect)		1	1	0	0	P	0	1	0
		DM (disconnect mode)	1	1	1	1	F	1	1	0
		UA (unnumbered acknowledgement)	1	1	0	0	F	1	1	0
		FRMR (frame reject)	1	1	1	0	F	0	0	1

The *control field* contains either a *command* or a *response*, and sequence numbers where applicable as a reference when acknowledging receipt of a previous frame. There are three types of control field format, corresponding to the numbered information transfer of *I-frames* (I format), numbered supervisory functions (S format) and unnumbered control functions (U format).

The control field is coded as detailed in Table 18.2.

The *frame check sequence (FCS)* field is a string of bits which help to determine at the receiving end whether the data in the frame has in any way been corrupted. It is a 16 bit field, created using the properties of a cyclic code, hence the term *cyclic redundancy check*. The exact 16 bit sequence sent is the ones complement of the sum (in binary) of the following two parts.

- (1) The remainder of 
$$x^k(x^{15} + x^{14} + x^{13} + x^{12} + x^{11} + x^{10} + x^9 + x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x^1 + x + 1)$$
 divided (in binary) by 
$$x^{16} + x^{12} + x^5 + 1$$
 where  $k$  is the number of bits in the frame excluding flag and FCS.
- (2) The remainder of the product of the frame content (excluding flag and FCS) and  $x^{16}$ , when divided by 
$$x^{16} + x^{12} + x^5 + 1$$

There are six configurable parameters when setting up LAPB connections. These, their meaning and typical value settings are given in Table 18.3.

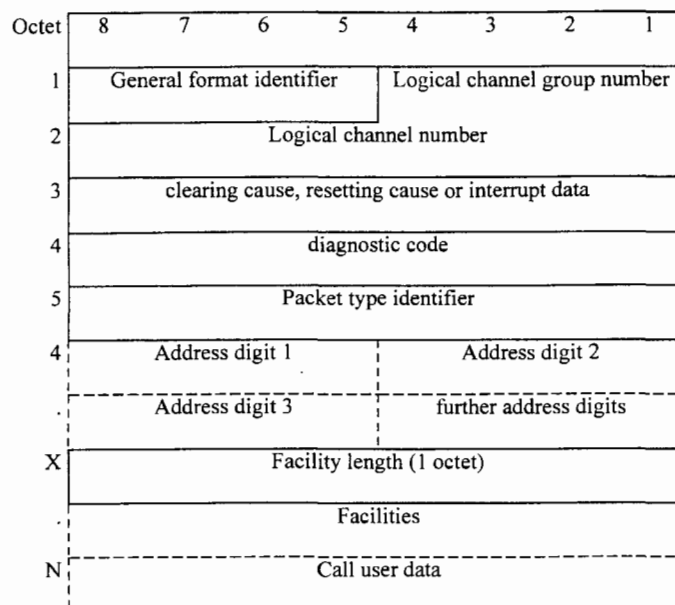
**Table 18.3** Optional parameter settings for X.25 LAPB

Parameter	Meaning	Typical value
Frame window size	This is the maximum permitted number of outstanding frames which may be sent without receiving an acknowledgement	7
N1 parameter	This is the maximum number of bits in an I-frame	2096
N2 parameter	This is the maximum number of attempts which may be used to complete a transmission	10 attempts
Timer T1	The timer T1 is the timeout period after which a retransmission may be initiated	3 seconds
Parameter T2	This is the maximum time allowed before an acknowledging frame must be initiated	0.2 seconds
Timer T3	Should the channel remain idle longer than this time, then the link shall be assumed to be non-active	0 (not used)

## 18.7 X.25 PACKET LEVEL INTERFACE (LAYER 3 PROTOCOL)

The information carried within a LAPB I-frame will be a *packet* of user data information structured in the format as defined by the X.25 packet level interface.

The general format of a *packet* is as shown in Figure 18.9.

**Figure 18.9** X.25 packet format (layer 3)

**Table 18.4** X.25 packet-coding of the general format identifier

General format identifier		bit 8	bit 7	bit 6	bit 5
Call set-up packets	sequence number modulo 8	×	×	0	1
	sequence number modulo 128	×	×	1	0
Clearing packets	sequence number modulo 8	×	0	0	1
	sequence number modulo 128	×	0	1	0
Flow control, interrupt, reset, restart, registration and diagnostic packets	sequence number modulo 8	0	0	0	1
	sequence number modulo 128	0	0	1	0
Data packets	sequence number modulo 8	×	×	0	1
	sequence number modulo 128	×	×	1	0
general format identifier extension		0	0	1	1
reserved for other applications		*	*	0	0

The minimum number of octets in a packet is three, the general format identifier, the logical channel identifier and the packet type identifier. The other packet types are added as required. As is normal, the least significant bits of each octet (i.e. bit 1) is transmitted first.

The *general format identifier* field provides information about the nature of the rest of the packet header (the first three octets of the packet). It is coded according to the values set out in Table 18.4.

The *logical channel number* is a reference number allowing the DTE and DCE to distinguish to which logical connection (or statistically multiplexed *virtual channel*) the packet belongs. Thus, in theory, up to 4096 logical channels may be supported by a single DTE /DCE connection simultaneously.

The *packet type identifier* is coded according to Table 18.5. As can be seen from the table, certain packet types are needed for *virtual calls* (VCs, also called *switched virtual channels*, or SVCs), and a lesser number of packet types are required to support *permanent virtual channels* (or PVCs). The difference between an SVC and a PVC is essentially the difference between an ordinary telephone line and a point-to-point leaseline. With an ordinary telephone line and an SVC each connection is set up on demand by dialling the number of the desired destination. With a telephone leaseline or a PVC the connection is permanently connected between the same two endpoints.

In clear request and clear indication packets, the clearing cause is included at octet 4. In reset request packets the reset cause appears in octet 4, while in interrupt data packets. *Interrupt data*, when sent, also appears in octet 4. This is data which is not subject to normal flow control, in effect allowing the DTE to override previous commands to the distant DTE (like a 'break' or 'escape' key). *Call request*, *call accepted* and *call connected* packets contain from octet 4 onwards the *address block* field, and optional *facility length* and *facility field* and the *called user data*. The

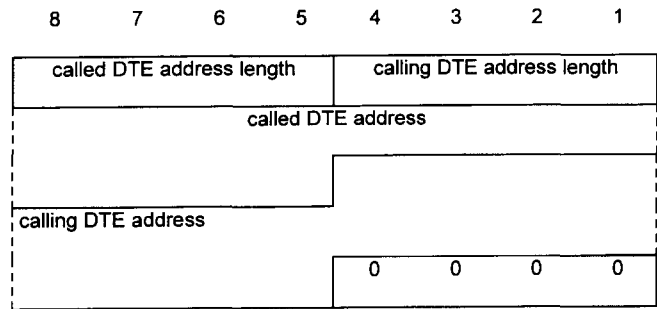
Table 18.5 X.25 packet identifier coding

Packet type											
From DCE to DTE	From DTE to DCE	VC	PVC	bit 8	bit 7	bit 6	bit 5	bit 4	bit 3	bit 2	bit 1
Call set-up and clearing											
incoming call	call request	✓		0	0	0	0	1	0	1	1
call connected	call accepted	✓		0	0	0	0	1	1	1	1
clear indication	clear request	✓		0	0	0	1	0	0	1	1
DCE clear confirmation	DTE clear confirmation	✓		0	0	0	1	0	1	1	1
Data and interrupt											
DCE data	DTE data	✓	✓	×	×	×	×	×	×	×	0
DCE interrupt	DTE interrupt	✓	✓	0	0	1	0	0	0	1	1
DCE interrupt confirmation	DTE interrupt confirmation	✓	✓	0	0	1	0	0	1	1	1
Flow control and reset											
DCE RR (modulo 8)	DTE RR (modulo 8)	✓	✓	×	×	×	0	0	0	0	1
DCE RR (modulo 128)	DTE RR (modulo 128)	✓	✓	0	0	0	0	0	0	0	1
DCE RNR (modulo 8)	DTE RNR (modulo 8)	✓	✓	×	×	×	0	0	1	0	1
DCE RNR (modulo 128)	DTE RNR (modulo 128)	✓	✓	0	0	0	0	0	1	0	1
	DTE REJ (modulo 8)	✓	✓	×	×	×	0	1	0	0	1
	DTE REJ (modulo 128)	✓	✓	0	0	0	0	1	0	0	1
reset indication	reset request	✓	✓	0	0	0	1	1	0	1	1
DCE reset confirmation	DTE reset confirmation	✓	✓	0	0	0	1	1	1	1	1
Restart											
restart indication	restart request	✓	✓	1	1	1	1	1	0	1	1
DCE restart confirmation	DTE restart confirmation	✓	✓	1	1	1	1	1	1	1	1
Diagnostic		✓	✓	1	1	1	1	0	0	0	1
Registration											
	registration request	✓	✓	1	1	1	1	0	0	1	1
registration confirmation		✓	✓	1	1	1	1	0	1	1	1

*diagnostic code* is included in diagnostic type packets and also optionally in clear request and reset request packets.

The *address block* field is coded as shown in Figure 18.10. The calling DTE and called DTE address length fields are coded in binary. The address digits themselves, however, are coded in four digit blocks, binary coded decimal, according to ITU-T recommendation X.121. If necessary, bits 1–4 of the last octet are filled with 0s to maintain octet alignment.

The *facility length* field merely indicates as a binary number the length of the facility field which follows. The *facilities* field allows the calling DTE to request a number of optional services. These (where made available) are as listed in Table 18.6.



**Figure 18.10** Format of the X.25 packet address block

**Table 18.6** X.25 network facilities

X.25 Optional facilities	Description
on-line facility registration	if subscribed to, allows user to request facility registration
extended packet sequence numbering	packets numbered modulo 128 rather than normal modulo 8
D bit modification	support of D-bit procedure
packet retransmission	allows DTE to request DCE to retransmit packets
incoming calls barred	prevents incoming calls being presented to DTE
outgoing calls barred	prevents DCE from accepting outgoing calls
one-way logical channel outgoing	restricts logical channel use to originating outgoing virtual calls only
one-way logical channel incoming	restricts logical channel use to incoming virtual calls only
non-standard default packet sizes	provides for the selection of non-default packet sizes
non-standard default window sizes	provides for the selection of non-default window sizes
default throughput classes assignment	provides for the selection of default throughput classes
flow control parameter negotiation	allows the DTE to alter window and packet sizes for each virtual call
throughput class negotiation	allows throughput class (i.e. bitspeed) negotiation for each call
closed user group	restricts incoming calls to the DTE to be from other specific DTEs which are members of a closed user group
bilateral closed user group	a closed user group of only two DTEs
fast select	allows the call request packet to contain up to 128 octets of user information, thereby speeding the sending of short user messages
fast select acceptance	authorizes DCE to forward fast select packets to the DTE



**Table 18.6** *(continued)*

X.25 Optional facilities	Description
reverse charging	allows calls to be requested for charging to recipients rather than callers
reverse charging acceptance	DTE authorization to the DCE that it is willing to accept reverse charge calls
local charging prevention	prevents the DCE from allowing virtual calls to be set up for which the local DTE will be charged
network user identification (NUI)	an authorisation procedure allowing for the dial-in to an X.25 network port across a public telephone network
charging information	DTE may request charging information
RPOA related facilities	allows the calling DTE to specify transit networks through which the call should be routed
hunt group	allows several separate DTE/DCE network connections to share the same network address. Incoming calls are routed to any free logical channel within at any of the DTEs
call redirection and call deflection	allows deflection on busy or redirection on no answer
called line address modified notification	notifies the calling DTE that the called address has been modified (diverting to a new destination)
transit delay selection and indication	allows the calling DTE to specify a desired transit delay
TOA/NPi address subscription	DTE/DCE to use TOA/NPi address format

## 18.8 TYPICAL PARAMETER DEFAULT SETTINGS USED IN X.25 NETWORKS

A typical public X.25 network might be set up with the following default settings for the packet level interface

- window size 2
- packet size 128
- window negotiation allowed
- packet negotiation allowed
- throughput negotiation allowed
- lowest logical channel number 1
- highest logical channel number 16
- customer DTE to select logical channel number, from highest number descending

## 18.9 PACKET ASSEMBLER/DISASSEMBLERS (PADS)

The X.25 protocol is suitable for the connection of *synchronous* communication devices (i.e. computers) to a packet switching network. It is widely used, and many devices are available worldwide (both DTEs and packet DSEs (data switching exchanges) which support it. However, directly X.25 protocol is unsuitable for connecting slow and relatively unintelligent *asynchronous* computer terminals to an X.25 packet network. Instead a procedure is defined by ITU-T recommendation X.28.

Necessary interworking is achieved by means of a piece of equipment provided in place of the DCE at the packet-switched exchange. This equipment is called a *packet assembler/disassembler* or *PAD* for short. As its name suggests, it assembles packets from the asynchronous terminal data for onward transmission, and it disassembles packets received from the far end DTE, conveying them on to the asynchronous terminal as individual characters.

Three ITU-T recommendations define the operation of PADs.

- X.3 defines the functions of the PAD (i.e. the conversion of asynchronous *characters* into synchronous *packets*).
- X.28 defines the control interface between PAD and asynchronous terminal. The X.28 parameters (Figure 18.18) define how the PAD is to react to the characters sent to it, how often to despatch groups of characters as a *packet*, and which characters typed by the end terminal are to be interpreted directly by the PAD itself as *control characters*.
- X.29 defines the interface between PAD and remote X25 DTE (usually some form of *host computer*).

Figure 18.11 illustrates the inter-relationship of the three recommendations and Table 18.7 lists the PAD control parameters of recommendation X.28. In the final column of

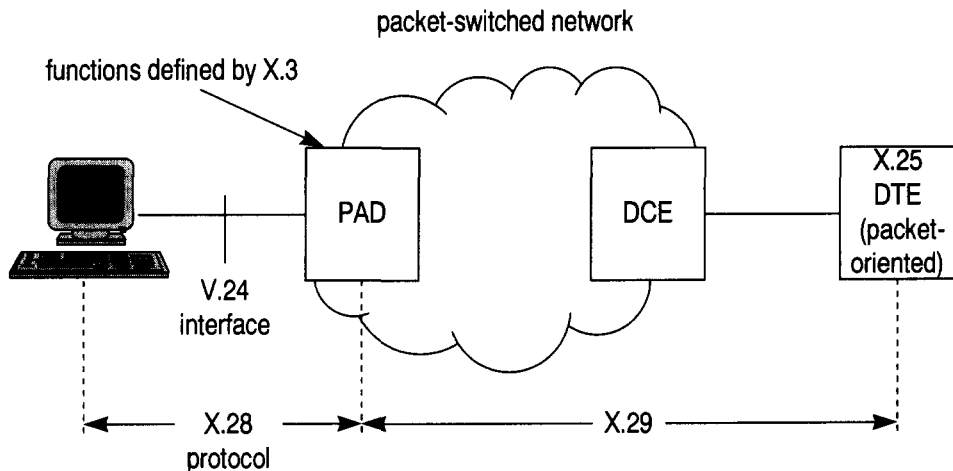


Figure 18.11 The packet assembler/disassembler (PAD)

**Table 18.7** PAD control parameters defined by ITU-T recommendation X.28

Parameter reference	Purpose	Permitted values	Simple standard profile (value)
parameter 1	PAD recall using character	0 = no escape 1 = escape 32–126 = ASCII code of 'escape' character	1
parameter 2	Echo (i.e. character is returned to be displayed on screen)	0 = no echo 1 = echo	1
parameter 3	Selection of data forwarding characters (e.g. 'carriage return' key)	0 = forwarded only on 'poll' from network 2 = forward on 'carriage return' 8 = forward on <CR>, <ESC>, <DEL>, <ENQ>, <ACK> 18 = forward on <CR>, <EDT>, <ETX> 126 = forward on all control characters	126
parameter 4	Selection of idle timer delay	0 = no delay 1–255 (delay, multiple of 10 ms)	0
parameter 5	Ancillary device control	0 = X-ON and X-OFF not in use 1 = X-ON and X-OFF in use 2 = X-ON and X-OFF in use during data transfer	1
parameter 6	Control of PAD service and command signals	0 = no signals sent to DTE 1 = signals sent to DTE other values define which signals may be used	1
parameter 7	Selection of reaction of PAD to the 'break' key	0 = no reaction to 'break' 1 = assembled characters and 'interrupt' sent to host 2 = reset on 'break' other values define other actions	2
parameter 8	Discard output	0 = deliver data 1 = discard data	0
parameter 9	Padding after carriage return	0–255 (number of characters to be inserted after 'carriage return')	0
parameter 10	Line folding	0–255 (defines length of a line in characters)	0
parameter 11	Binary speed	12 = 2 400 bit/s 13 = 4 800 bit/s 14 = 9 600 bit/s 16 = 19 200 bit/s 19 = 14 400 bit/s 20 = 28 800 bit/s 21 = 38 400 bit/s other values give other speeds	—

Table 18.7 (continued)

Parameter reference	Purpose	Permitted values	Simple standard profile (value)
parameter 12	Flow control of the PAD by the DTE	0 = no use of X-ON and X-OFF for flow control 1 = use of X-ON and X-OFF for flow control	1
parameter 13	'Linefeed' insertion after 'carriage return'	0 = no linefeed insertion 1 = linefeed after 'carriage return' other values define other actions	0
parameter 14	Linefeed padding	0-255 (number of characters inserted after linefeed)	0
parameter 15	Editing	0 = editing only in command mode 1 = editing in command and data transfer modes 2 = extended editing	0
parameter 16	Character delete	0-127 (ASCII code of character used as character delete)	127
parameter 17	Line delete	0-127 (ASCII code of character used as line delete)	24
parameter 18	Line display	0-127 (ASCII code of character used to request line display)	18
parameter 19	Editing PAD service signals	0 = no editing 1 = editing for print terminals 2 = editing for display terminals 8, 32-126 this character replaces deleted characters	1
parameter 20	Echo mask	0 = no echo mask 1 = no echo of 'carriage return' 2 = no echo of 'linefeed' other values define other control characters not echoed	0
parameter 21	Parity treatment	0 = parity not generated or checked 1 = parity checked of input from DTE 2 = parity generated for output to DTE 3 = parity generated and checked (combination of 1 and 2)	0
parameter 22	Page wait	0-255 (number of line feeds which may be generated before an 'abort')	0
parameter 23	Size of input field	0-255 (number of characters)	0
parameter 24	End of frame signal	0 = not used 1-127 ASCII code of character to signify end of frame	0

**Table 18.7** (continued)

Parameter reference	Purpose	Permitted values	Simple standard profile (value)
parameter 25	Additional data forwarding character	0 = no extra data forwarding 1–127 = ASCII code of extra forwarding character	0
parameter 26	Display 'interrupt' character	0 = no abort character 1–127 = ASCII code of 'interrupt' character	0
parameter 27	Display interrupt confirmation ('prompt' character)	0 = no confirmation 1–127 = ASCII code of 'interrupt' confirmation character	0
parameter 28	Diacritic character coding scheme	0 = not used	0
parameter 29	Extended echo mask	0 = not used	0

Table 18.7, the parameter settings are given for the *simple standard profile*. This is a standard X.28 parameter setting set for emulating a low speed 'transparent' leaseline between an asynchronous computer terminal and its terminal controller.

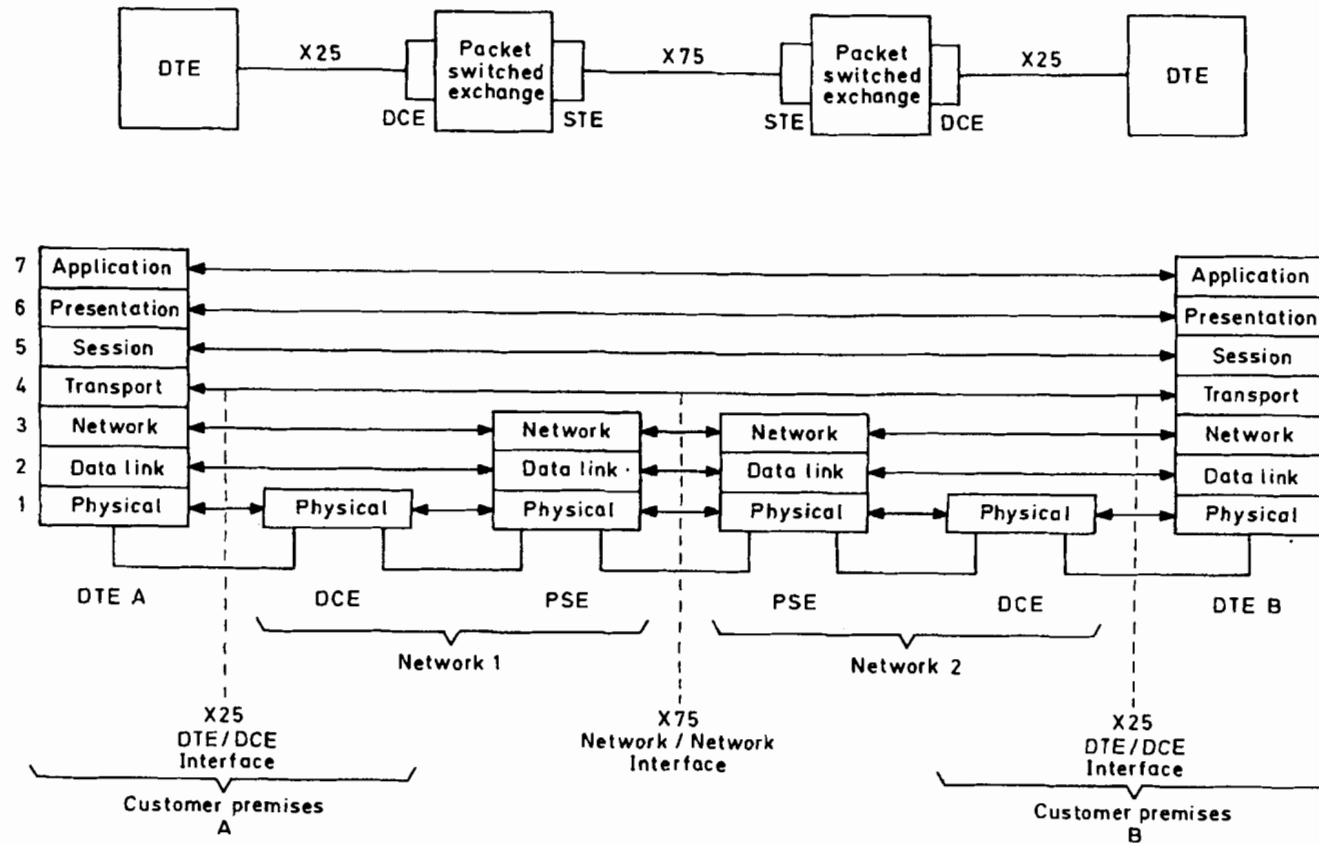
In case you are left wondering, having studied Table 18.7, what exactly 'X-ON' and 'X-OFF' are, these are the alternative states of a given control lead within a standard DTE-to-DCE interface (such as V.24 as defined in Chapter 9). When the lead is set 'X-ON' then the DTE may send transmit data. When instead the lead is set 'X-OFF' then no data may be transmitted. This ensures that the DCE is ready to receive data before the DTE may send. Before call set up the lead is set X-OFF, then X-ON after call set-up. During the call, the lead may be reset to X-OFF to slow up the receipt of data (i.e. for *flow control*) if the DCE or network becomes overloaded.

A call can be set up from an *asynchronous DTE* (so-called *start/stop terminal*) simply by typing in the X.121 data network address (Chapter 29). This procedure for call set-up from a start-stop terminal is also defined by X.28.

## 18.10 ITU-T RECOMMENDATION X.75

To connect different packet switched networks together, a *network-network interface (NNI)* or *gateway protocol* is defined by ITU-T in its recommendation X.75. X.75 can be thought of as a super set of the X.25 protocol (Figure 18.12). It was originally developed for international interconnection of packet networks).

Like X.25, recommendation X.75 defines all the protocols for layers 1–3. At layer 1, a 64 kbit/s bearer conforming to Recommendation G.703 is stipulated. At layer 2, the



**Figure 18.12** CCITT Recommendation X.75 and its interrelationship with X.25. DTE, Data terminal equipment; DCE, Data circuit terminating equipment; PSE, Packet-switched exchange; STE, Signalling terminal

*single link procedure (SLP)* and the *multilink procedure (MLP)* are defined. SLP is based on LAP (the *link access procedure* used in X.25) and should be used where *packet-switched exchanges (PSEs)* are interconnected by a single datalink. Where more than one parallel link is available between exchanges the more advanced MLP must be used. The protocol of X.75 is similar to X.25 at layer 3, with a few extra messages added to cater for inter-exchange signalling requirements. Communication using X.75 occurs between logical functions called *signalling terminals (STE)* which are software programs within the *packet-switched exchanges (PSEs)*.

## 18.11 WHEN X.25 PACKET SWITCHING MAY AND MAY NOT BE USED

The use of either a private corporate or public X.25 packet switching network is ideally suited for the connection of a large number of data terminal devices back to a central computer centre. In particular, the use of X.25 is usually economic where the volume of data transferred from each individual device is relatively low, so that by *statistically multiplexing* the data onto a common network of lines line costs may be saved. When, however, the data volume between two specific points exceeds a certain economic threshold, it will be cheaper to use a direct circuit.

Common applications are the connection of point-of-sale and credit card authorization devices in shops and stores back to the headquarters computer centre or bank authorizing centre. Other applications might include connection of remote small offices equipped with IBM 3174 terminal controllers back to a central computer centre site offering *NPSI (network packet switching interface)*, IBM's mainframe software product for enabling X.25).

X.25-based packet networks are not economically viable where relatively high volumes of data are to be carried between only a small number of sites. They are also technically unsuitable where prolonged point-to-point carriage rates much above 256 kbit/s are required.

The strength of X.25-based packet networks is their high reliability in transporting data accurately across even quite poor lines. Their weakness is their relative slowness compared with more modern techniques (e.g. frame relay) and the relatively high delays which can be incurred when transporting individual packets of information. This delay can cause particular technical difficulties or user annoyance for certain types of computer applications. Where the computer application has been designed without specific thought for X.25 communication, it may run into technical difficulties.

A common problem incurred with X.25 networks is that which occurs when applications conduct an unnecessarily ping-pong nature of dialogue, e.g. 'please send me next information... (information)... received... can I send next information?... please send next information (information, part 2), etc., etc.'. Although the delay for sending each message across the X.25 network is quite low (say 100 ms), the fact that four messages are actually sent by the application for each one which is really needed multiplies the delay to 400 ms). The best applications run with minimum requests for, and acknowledgements of, information transfer.

## 18.12 ALTERNATIVES TO X.25-BASED PACKET SWITCHING

Where all the equipment within a computer data network is provided by a single supplier (e.g. IBM, DEC, Siemens Nixdorf, Unisys) it may be possible and advisable to use the supplier's proprietary data communication protocol (e.g. SNA, DECNET, etc.). This probably helps to optimize the management of the computer environment and keep down the computer equipment costs. In addition, software programs may run better and faster. The disadvantage of 'proprietary' protocols are that they generally rule out the use of public packet switching networks as the communication transport medium. The effect may be to demand a higher number of relatively expensive point-to-point private lines (leaselines) to be rented from the telephone company, for dedicated use. Conversion of the network (usually by encapsulating the same information within X.25 packets) enables the packet switch network to be used but it is usually linked to further software and hardware investment costs.

At higher data volumes and higher speeds, frame relay protocol is becoming the accepted replacement for X.25. It shares the *statistical multiplexing* benefits which X.25-based packet networks offer, but it is prone to lower individual *frame* (rather than *packet*) delays. At very high speeds ATM (asynchronous transfer mode) is the emerging standard.

For extremely high speeds and for applications requiring very low jitter (variability of delay) such as high quality speech, music or video, there is no alternative to direct leaselines. Also, for very high point-to-point data volumes, direct leaselines and a proprietary protocol may be cheapest.

## 18.13 IBM'S 'SYSTEMS NETWORK ARCHITECTURE'

IBM's *systems network architecture* (SNA) defines the functions of a layered protocol set specifically intended for use in 'pure' IBM computer networks. It is a 'proprietary' protocol, specified only by the IBM company, but is widely supported by other manufacturers' equipment, due to the very large numbers of IBM mainframe computers and their associated computer networks.

SNA was first introduced by IBM in 1974 and it has evolved through several generations to its current form. It was developed as an all-purpose data networks architecture, a replacement for the growing number of incompatible communication products which IBM had at the time. The latest SNA networks allow a large number of *host* machines, *communication controllers*, and *cluster controllers* (IBM's name for terminal controllers) to share a sophisticated communication network. By so doing, SNA creates

- an ability for terminal users to switch dynamically between different application programmes or even different *host* computers
- an ability to attach dissimilar devices to the same network
- an ability to concentrate (or *multiplex*) a number of data connections to share a common circuit or network



**Table 18.8** SNA physical unit (PU) classes

Class type	Class purpose and example
5	Host computer or system services control point (SSCP) (network configuration software residing on a mainframe computer)
4	Communications controller (e.g. IBM 3725 or 3745) or network control program (NCP) (a software residing on a host)
3	never existed
2.1	'independent' cluster controller capable of peer-to-peer operation across an SNA network (this class was conceived for the IBM series of AS400 midrange computers to be able to communicate directly with one another)
2	'dependent' cluster controller (e.g. IBM 3174 or 3274), a device allowing a number of different end user terminals (dumb terminals) to share the same statistically multiplexed line to the mainframe host computer
1	remote devices, including terminal nodes such as remote printing devices now virtually obsolete

- an ability for direct interconnection between *host* computers, PCs and/or minicomputers, so enabling *distributed processing* and storage of data

SNA caters for communication between five different classes of *physical units (PUs)* such as terminals and computers, classified as shown in Table 18.8.

The communication itself takes place between logical units (LUs) which are software programs resident within the PUs. A *session* of communication ultimately takes place

**Table 18.9** Systems network architecture (SNA)

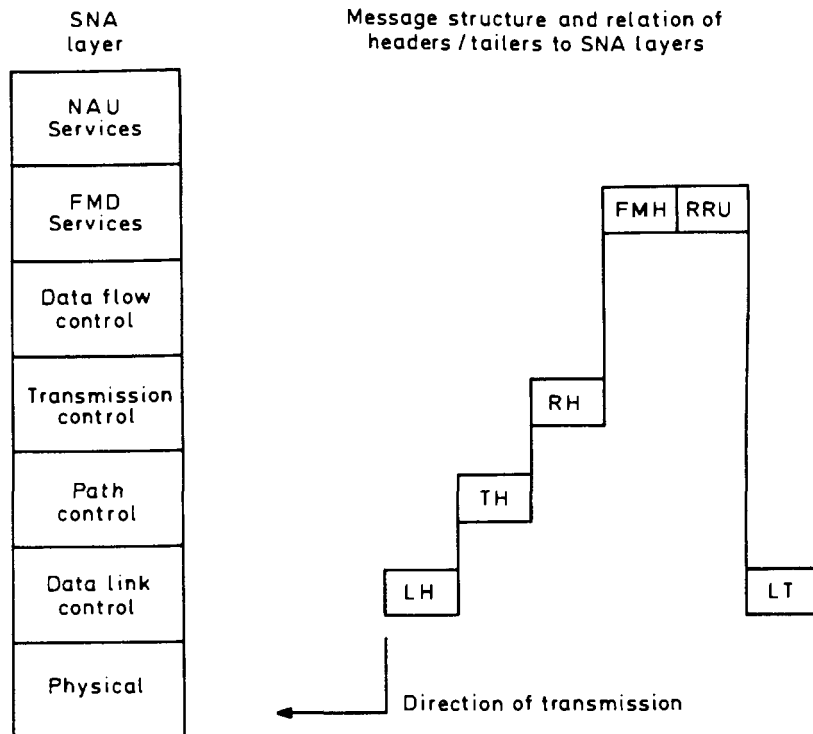
SNA layer name	Equivalent	Function
NAU services (NAU = Network Addressable Unit)	7	Data exchange between logical units (LUs)
FMD services (FMD = File Management Data)	6	Syntax. Data compression and compaction type. ASCII, EBCDIC code type
Data flow control	5	Dialogue control. (Session control)
Transmission control	4	Activating and deactivating data flow within a session
Path control	2/3	Routing and flow control. Message Packetization
Datalink control	2	Managing bit oriented data flow. SDLC
Physical	1	X21. RS232-C

to serve the communication needs of these programs. Thus an LU is the equivalent of the OSI application layer.

Serving a *logical unit (LU)* at either end of an SNA network, there are up to seven layered protocols, as shown in Table 18.9. The protocols are similar in purpose to those of the OSI model but there are some significant differences in the functions of individual layers.

The layered nature of SNA is immediately apparent in the message structure, as we show in Figure 18.13. Just as in the OSI case, each progressively lower layer adds an information field, called a *header* or *trailer*, to the original data which is held in the *request/response unit*. At the receiving end each progressively higher layer removes its appropriate header and passes to the next higher layer all that is left.

The layered nature of SNA is intended to provide for modular design of computer network software, giving greater ease of maintenance and more easily modifiable implementations. Between them the layers are intended to meet all the data communication requirements of an entire distributed data processing network. The layers themselves may thus be realized in a number of different forms. Usually the layers are either all realized in the same physical unit (PU), or, as is done with many host computers, the protocols are implemented in two parts; the lowest three layers (*physical*,



**Figure 18.13** SNA message structure. LH, Link Header; TH, Transmission Header; RH, Response Header; FMH, Function Management Header; RRU, Request Response Unit; LT, Link Trailer

*datalink* and *path control*) making up the *path control network* are conducted by the communication controller (front end processor), and the higher layers together form a *network addressable unit (NAU)* which resides in the host processor itself. The NAU is held in software in the host, and it relies on the *network control program (NCP)* software in the *front end processor* or *communication controller*. A number of alternative IBM software products are available for installation as NAUs on IBM or compatible equipment. The most commonly used communication product on IBM mainframe computers is *VTAM* (*virtual telecommunication access method*). An older, less frequently used product is *TCAM* (*advanced communication function/telecommunication access method*). Both provide the necessary host software for telecommunications in support of a logical unit (LU) in the application program of a host computer.

Both *TCAM* and *VTAM* include a function called the *SSCP* (*system services control point*). This is a software function, and each SNA network must have at least one of them to control the overall network configuration, activating and deactivating the network and establishing communication sessions. This it does by interacting with appropriate software in each of the other *physical units (PUs)* of the network.

The *communication controller*, if provided as a separate physical entity from the host, can be any of a number of IBM 37X5 series machines. These are specialized telecommunication hardware devices which conform to the SNA model when loaded with IBM *network control program (NCP)* software and act to relieve the host of communication functions. A *cluster controller* (e.g. IBM 3174 or 3274) allows the connection of dumb terminals to the network. A typical configuration is shown in Figure 18.14.

As already discussed, one of the most common SNA communication regimes is often referred to as *3270 communication*. This is the method used for communication between a mainframe computer and a 'dumb', 327X-type (e.g. 3270) terminal. However, the emergence of the personal computer (PC) in recent years has strained the capabilities of this method. Nowadays it is common to use PCs as part-time mainframe computer terminals, but to do so requires that they *emulate* the dumb terminals they substitute. So we have *3270 emulation*. Either by hardware means (a communication

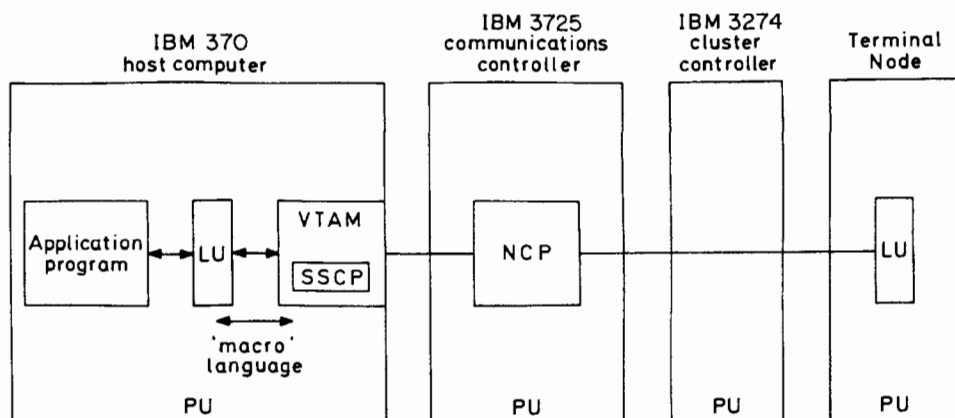


Figure 18.14 A typical SNA network

board added to the back of the PC) or under software control, the PC pretends to be a dumb terminal. Interaction may progress but not without its constraints; the session can only be established from the PC end of the connection. New OSI protocols should alleviate this constraint.

Where instead the PC (or LAN gateway) emulates the cluster controller rather than just the terminal, and by so doing can be connected directly to the front end processor, then we refer to it as an *SDLC adapter* or *SDLC gateway*.

An alternative to SDLC, predating SNA but still used today, is *BSC* (*binary synchronous communication* or *BISYNC*). Invented in 1964, this is a protocol in which the bit strings representing individual characters are delineated by control character sequences, rather than being synchronous as SDLC. For completeness, why not also mention 2780 and 3780 protocols? These were the pre-SNA protocols originally used in the days of batch computing when *jobs* were submitted by users to remote data centres. At these centres, computer staff would schedule and run the jobs in a priority order. 2780 (or the enhanced version, 3780) was the protocol used for 'remote job entry' by these staff. 2780/3780 protocol remains in common use between IBM mainframes and operation/maintenance devices.

Finally, back to SNA and a mention of the classification for the various different types of *logical user (LU) sessions* which is made possible by SNA. There are six classes of *LU session*, characterized in terms of the *session partners*. Each session type has slightly different needs and therefore puts slightly different demands on the protocol layers. These demands are met by using different options within the protocols. The overall set of options used is called the *session profile*. The six different session types are listed in Table 18.10.

Of the LU session types shown in Table 18.10 we mention only *LU6* specifically, and this only to record the widespread interest among computer systems designers for the

**Table 18.10** SNA logical user (LU) session types

LU session class	End users
0	Session conducted between end users of unknown types
1	Session conducted between an application program and an input/output device in an interactive, 'batch' or 'distributed data processing environment'
2	Session conducted between an application program and a single interactive operator display terminal (3270 type)
3	Session conducted between an application program and a single printer terminal in an interactive environment (3270 type printer)
4	Session conducted between an application program and multiple interactive operator display terminals, or between two operator display terminals
6	Session conducted for file or data transfer between application programs in a 'cooperative' distributed data-processing environment

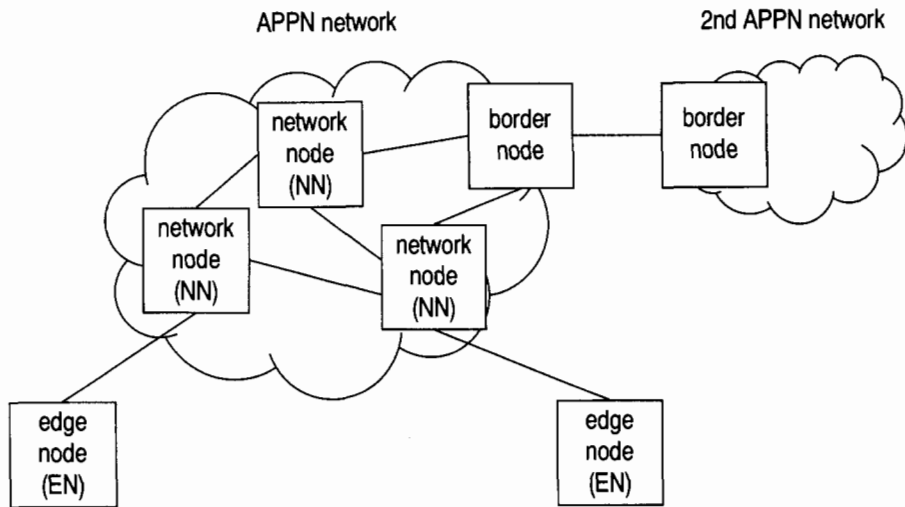


Figure 18.15 Advanced peer-to-peer networking (APPN)

second generic version LU6.2 which has set the *de facto* standard for program-to-program communications.

## 18.14 APPN (ADVANCED PEER-TO-PEER NETWORKING)

SNA was developed in the days when computer networks served purely to connect the 'dumb' user terminals to the *mainframe* or *host* computer. No other types of computer existed at the time. In recent years, however, the computing world has become dominated by PCs, *workstations* and powerful *server* computers, which share resources and software, and have a great need for intercommunication. Instead of a network dominated by a single large computer and a star-network out to the terminals, the network is required to serve many thousands of computers which share the network on an equal basis and are capable of transmitting files or commands in both directions of communication. These needs stand in stark contrast to the old *host*-dominated networks, where the commands came from the terminals, but most of the data flow was from the computer to the terminal, in the form of 'display' information. Nowadays the computers need to intercommunicate on a *peer-to-peer* basis. IBM's proprietary communication system and protocol for this purpose is called *APPN* (*advanced peer-to-peer networking*). It is in effect an extension of the PU2.1 and LU6.2 *physical* and *logical units* defined within SNA.

As illustrated by Figure 18.15, APPN is conducted between *end nodes* (EN) via *network nodes* (NN) and *border nodes*. Typical *end nodes* are AS400 *midrange* computers or *UNIX-workstations* or *servers* (such as the IBM RS6000). Meanwhile a new software has been developed for the *communications controllers* (e.g. IBM 3745) to enable them to conduct the *network node* (i.e. switching) function. A *border node* acts a little like a *network node*, but with the extra responsibility to manage the interconnection of two distinct APPN networks.

---

## *Local Area Networks (LANs)*

---

We have seen how packet switching has contributed greatly to the efficiency and flexibility of 'wide area' data networks, involving a large number of devices spread at geographically diverse locations. Packet switching, however, is not so efficient for smaller scale networks, those limited to linking personal computers within an office building; that is the realm of an alternative type of packet-switched-like network called a local area network or LAN for short. In this chapter we discuss the concept of a LAN and the various technical realizations which are available.

### **19.1 THE EMERGENCE OF LANs**

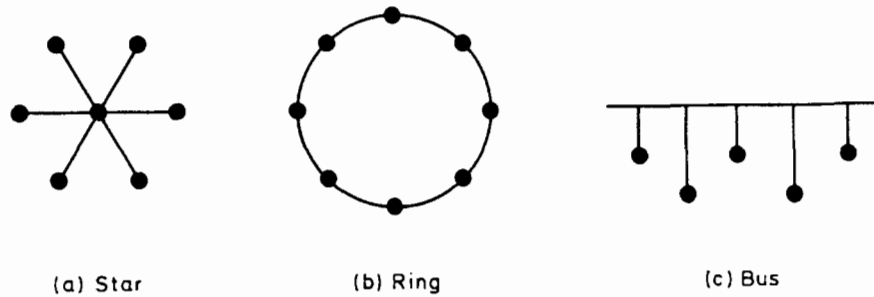
LANs emerged in the late 1980s as the most important means of conveying data between different computers and computer peripheral devices (printer, file server, electronic mail server, fax gateway, host gateway, computer printer, scanner, etc.) within a single office, office building or small campus. LANs are constrained by their mode of operation to a geographically limited area, but are ideally suited for short distance data transport.

A high bit speed LAN can carry high volumes of data with rapid response times. Such performance is crucial for most office applications, and has made them the ideal foundation for the new generation of 'electronic offices' comprising electronic workstations, word processors, shared printers, electronic filing cabinets, electronic mail systems and so on.

Most LANs conform to one of the different types specified in the Institution of Electrical and Electronic Engineers' IEEE 802 series of standards. All the types have been developed from proprietary LANs, developed earlier by individual companies or organizations, but have now achieved American and worldwide recognition, as *ISO 8802* standards.

### **19.2 LAN TOPOLOGIES AND STANDARDS**

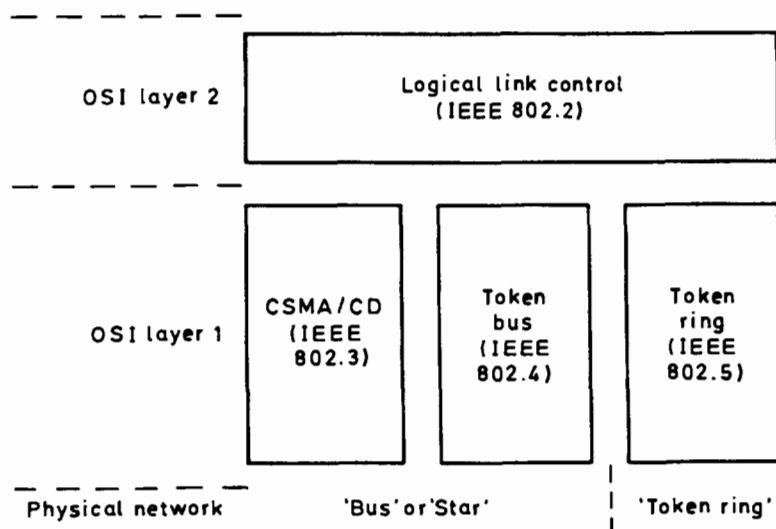
The different types of LAN are characterized by their distinctive topologies. They all comprise a single transmission path interconnecting all the data terminal devices, with a



**Figure 19.1** Alternative LAN topologies

bit speed typically between 1 and 30 Mbit/s, together with appropriate *protocols* (called the *logical link control* and the *medium access control (MAC)*) to enable data transfer. The three most common topologies are illustrated in Figure 19.1, and are called the *star*, *ring* and *bus* topologies.

Slightly different protocol standards apply to the different topologies. For example, *IEEE 802.3* defines a physical layer protocol called *CSMA/CD* (*carrier sense multiple access with collision detection*) which may be used with a bus or star topology. Used with a *bus* form medium, such LANs are normally referred to as *ethernets*. *IEEE 802.4* (ISO 8802.4) defines an alternative layer-1 protocol for a *token bus*, again suitable for either a bus or star topology. *IEEE 802.5* defines a layer 1 protocol suitable for use on a *token ring* topology. Finally, *IEEE 802.2* (ISO 8802.2) defines a *logical link control* protocol (equivalent to the OSI layer 2) that can be used with any of the above. This provides for the transfer of information between any two devices connected to the LAN. The information to be transported (i.e. information frame or packet) is submitted



**Figure 19.2** The IEEE 802 LAN standards

to the *logical link control* (LLC) layer together with the address of the device to which it is to be transmitted. Much like HDLC in X.25 (Chapter 18), the LLC assures successful transfer, error detection, retransmit, etc. Figure 19.2 shows the relationship of the various standards.

Which physical layer protocol and which topology of LAN to use depend largely on individual preference and the compatibility of the existing computer kit needing to be connected to the LAN. To a lesser degree, the geographic circumstances and the network's performance requirements are also factors. All the possible protocols transfer data between the nodes, using a packet mode of transmission; they differ in how they prevent more than one terminal using the bus or ring at the same time. The various protocols and their relative merits are now considered in turn.

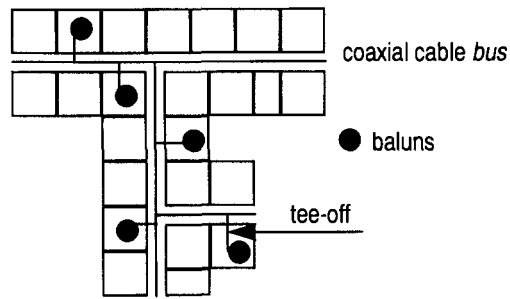
### 19.3 CSMA/CD (IEEE 802.3, ISO 8802.3): ETHERNET

*CSMA/CD* stands for *carrier sense multiple access with collision detection*. It is a *contention protocol*. On a CSMA/CD LAN the terminals do not request permission from a central controller before transmitting data on the transmission channel; they *contend* for its use. Before transmitting a packet of data, a sending terminal 'listens' to check whether the path is already in use, and if so it waits before transmitting its data. Even when it starts to send data, it needs to continue checking the path to make sure that no other stations have started sending data at the same time. If the sending terminal's output does not match that which it is simultaneously monitoring on the transmission path, it knows there has been a *collision*. To receive data, the *medium access control* (MAC) or layer 1 software in each terminal monitors the transmission path, decoding the destination address of each packet passing through to find out whether it is the intended destination. If it is, the data is read and decoded; if not, the data are ignored.

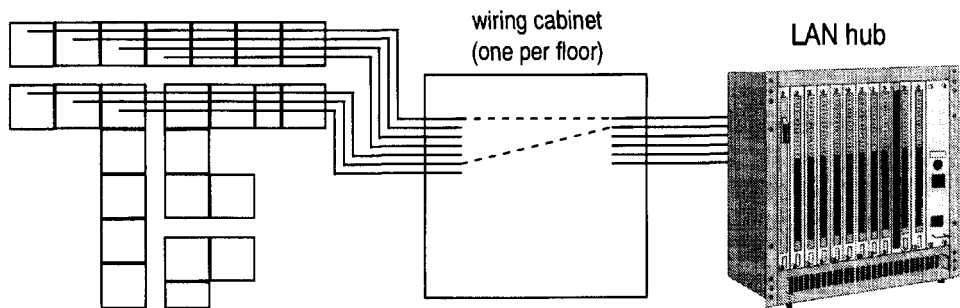
The most important type of network that employs the CSMA/CD is called *ethernet*. Ethernet was originally a proprietary LAN standard (predating the IEEE 802.3 standard) developed by the Xerox corporation of USA. The original design was based on a length of coaxial cable, with 'tee-offs' to individual work stations, with a maximum of around 500 stations. The idea was to simplify the cabling needs of offices in which many personal computers were in use. Simply by laying a single coaxial cable along each of the corridors and connecting all the cables together, a *bus* could be created over which all the office computer devices could intercommunicate. Each time a new device was installed, a new tee-off could be installed from the corridor into the particular office where the device was situated (Figure 19.3(a)). Meanwhile, no new cabling needed to be installed along the corridor, so saving space in the conduits and averting the constant removal and replacement of the ceiling tiles.

The technology for basic *ethernet* (10base5) developed rapidly. First, clever devices for the tee-off points were developed, which enabled new devices to be connected very quickly without first severing the main coaxial cable bus. The devices pressed directly into the cable. This reduced the time needed for new installations and reduced the disturbance to existing users. *Thin-ethernet* (*cheapernet* or 10base2) appeared. This





a) ethernet as coaxial cable bus



b) ethernet as structured twisted pair cabling

**Figure 19.3** Typical coaxial cable and twisted pair wiring configurations for Ethernet

allowed the use of narrower gauge coaxial cable as the main bus in smaller networks, and helped to reduce the installation costs. Meanwhile, the numbers of computer devices in offices were multiplying rapidly. Multiple ethernet networks became necessary, and increased flexibility was demanded to enable users to move offices without major cabling disturbances. This caused the development of LANs on *structured cabling*, using LAN *hubs* and *twisted pair* telephone cabling, in a star configuration. The ethernet *10baseT* standard was born (Figure 19.3(b)).

As Figure 3(a) illustrates, in the coaxial cable realisation, a single cable bus, usually installed in the cable conduit in the office corridor provides the main network element. Tee-offs into individual offices are installed as needed, either by tee-ing directly into the main bus, or by using pre-installed sockets and connectors. A baluns, usually built into the coaxial cable socket in the end location, provides for correct impedance matching (50), whether or not the device is connected into the socket.

When installed as part of a *structured cabling scheme* (nowadays the most common realisation of ethernet), twisted pair cabling provides for the transmission medium. Multiple twisted pair cables are usually installed in each individual office and near each

individual desk during office renovation, and wired back to a wiring cabinet, of which there is usually one per floor, installed in an equipment room. Usually next to the wiring cabinet, or even in the same rack, a LAN *hub* is installed. The *hub* replaces the coaxial cable backbone, so that the arrangement is sometimes referred to as a *collapsed backbone* topology. The *bus* topology still exists, but now only within the *hub* itself, which provides for the interconnection of all the devices forming the LAN, ensuring physical connection and appropriate electrical impedance matching.

Should new devices need to be added, a spare cable can be *patched* through at the wiring cabinet and a new port card can be slid into the *hub*. Should any of the devices need to be moved from one office to another this can be achieved by re-*patching* at the wiring cabinet. The *adds* and *changes* are thus far less disruptive both to other LAN users and the office furnishings. In the *structured cabling* scheme, the *baluns* is no longer needed, since the hub provides for this function.

Ethernet LAN components are relatively cheap. The bus topology is easy to realize and manage and is resilient to transmission line failures. As a result, ethernet has become the predominant type of LAN. The fact that any station may use the transmission path, so long as it was previously idle, means that fairly good use can be made of the LAN even when some destinations are unavailable because of a transmission path break, a capability which is not enjoyed by LANs employing more sophisticated data transmission, as we shall see later.

Theory suggests that the random collisions of a large number of competing devices all trying to communicate over the same CDMA/CDLAN lead to rapid network performance degradation under heavy load. In practice, however, the traffic is rarely random, because most users communicate with the various main central *server* devices within the network which regulate the communication. However, should poor performance under heavy load be a problem, it can usually be overcome by subdivision into smaller, interconnected LANs.

## 19.4 TOKEN BUS (IEEE 802.4, ISO 8802.4)

A *token bus* LAN controls the transmission of data onto the transmission path by the use of a single *token*. Only the terminal with the *token* may transmit packets onto the bus. The token can be made available to any terminal wishing to transmit data. When a terminal has the token it sends any data frames it has ready, and then passes the token on to the next terminal. To check that its successor has received the token correctly the terminal makes sure that the successor is transmitting data. If not, the successor is assumed to be on a failed part of the network, and to prevent 'lock-up' of the LAN, the original terminal creates a new successor by generating a new token. Transmission faults in the LAN bus can therefore be circumvented to some extent. However those parts of the LAN that are isolated from the token remain cut-off.

Token bus networks are not commonly used in office environments where ethernet and token ring networks predominate. Token bus networks are most common in manufacturing premises, often operating as broadband (high speed) networks for the tooling and control of complex robotic machines.

## 19.5 TOKEN RING (IEEE 802.5)

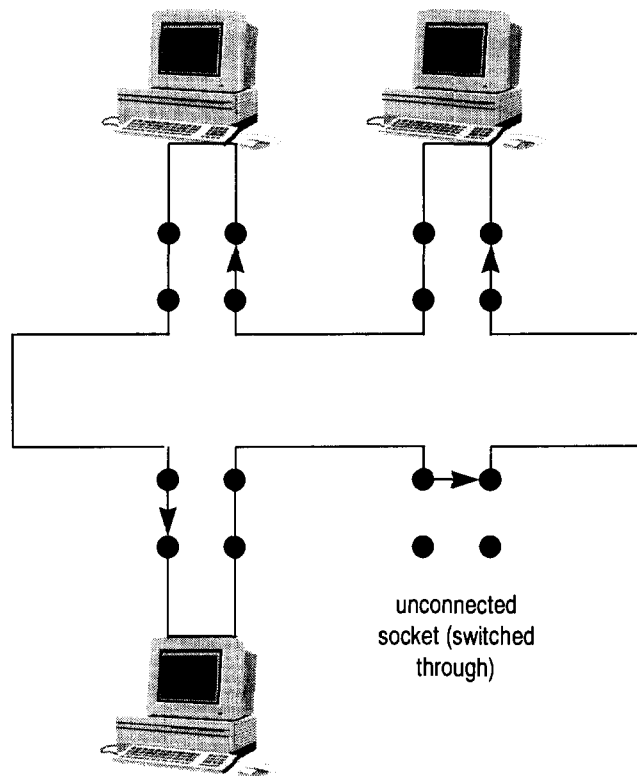
The *token ring* standard is similar in operation to the token bus, using the token to pass the 'right to transmit data' around each terminal on the ring in turn. The sequence of token passing is different: the token itself is used to carry the packet of data. The transmitting terminal sets the token's *flag*, putting the destination address in the *header* to indicate that the token is full. The token is then passed around the ring from one terminal to the next. Each terminal checks whether the data is intended for it, and passes it on; sooner or later it reaches the destination terminal where the data is read. Receipt of the data is confirmed to the transmitter by changing a bit value in the token's flag. When the token gets back to the transmitting terminal, the terminal is obliged to empty the token and pass it to the next terminal in the ring.

The feature of IEEE 802.5 MAC protocol is its ability to establish priorities among the ring terminals. This it does through a set of priority indicators in the token. As the token is passed around the ring, any terminal may request its use on the next pass by putting a *request* of a given priority in the reservation field. Provided no other station makes a higher priority request, then access to the token is given next time around. The reservation field therefore gives a means of determining demand on the LAN at any moment by counting the number of requests in the flag, and in addition the system of prioritization ensures that terminals with the highest pre-assigned authority have the first turn. High speed operation of certain pre-determined, time-critical devices is likely to be crucial to the operation of the network as a whole, but they are unlikely to need the token on every pass, so that lower priority terminals have a chance to use the ring when the higher priority stations are not active.

*Token ring* was developed by the IBM company, and is most common in office installations where large IBM mainframe and midrange computers (particularly AS400) are in use, in addition to large numbers of IBM PCs. The original form required specialized cabling (IBM type 1) and operated at 4 Mbit/s form. The initial idea was that a single cable loop could be laid through all the offices on a floor or in a building and devices added on demand. To avoid the disturbances and complications which might arise when connecting new devices to the ring (any break in the ring renders the LAN inoperative), IBM developed a sophisticated cabling system, including the various IBM special cables. The cable loop was pre-fitted with a number of sockets at all possible user device locations. The sockets ensured that when no device was connected, the ring was through-connected. However, on plugging in a new device, the ring is diverted through that device (Figure 19.4). The special socket for early token ring networks thus catered not only for correct impedance matching, but also for the ring continuity.

Token ring cards in the individual end user computer devices connected to token ring LANs also need to be designed to ensure ring continuity in the case that the device is switched off. Thus the card reverts to a 'switched-through' state when no power is applied, so that even though the end device itself plays no active part in token passing while switched off, the tokens nonetheless still have a complete ring available.

The further development of the token ring technology (mainly by IBM) has brought about the ability to use *twisted pair* cabling, and the emergence of a 16 Mbit/s as well as the original 4 Mbit/s version. In the 16 Mbit/s version, higher quality cabling (e.g. category 5 cable, as discussed in Chapter 8) may be required.



**Figure 19.4** Socket design in token ring LANs to ensure ring continuity

Token ring LAN *hubs* have also developed alongside ethernet hubs, and allow for similar *collapsed backbone* topologies in conjunction with *structured cabling systems*. Thus a token ring LAN today is difficult to distinguish from an ethernet LAN (Figure 19.3(b)). The ring topology is collapsed into the hub itself, and two sets of wires to each individual user station allow for the extension of the ring to each user device. The switch-through function previously performed by the socket is also undertaken at the hub, so reducing the complexity and cost of individual sockets, so that standard telephone sockets may be used.

The *token ring LAN* may differ from the *ethernet LAN* only in the port cards used within the hub and the LAN cards used in the individual PCs. Otherwise cabling, wiring cabinet and LAN hub unit may be identical. Indeed, in some companies, ethernet and token ring LANs exist alongside one another, without the user being aware to which type of LAN he is connected.

Token rings, like ethernet, are common in office environments, linking personal computers for the purpose of data file transfer, electronic messaging, mainframe computer interaction or file sharing. Some LAN administrators are emotional about whether *ethernet* or *token ring* offers the best solution, but in reality for most office users there is little to choose between them. Token ring LANs perform better than ethernet at near full capacity or during overload but can be more difficult and costly to install, especially when only a small number of users are involved.

In most cases, the choice between ethernet and token ring comes down to the recommendation of a user's computer supplier, as hardware and software of a particular computer type may have been developed with one or other type of LAN in mind. Thus *token ring* remains the recommendation of the IBM company, whereas in all other environments *ethernet* has gained the upper hand.

*Slotted ring* and other types of LAN also exist, but are not covered in detail in this book because they are rare. ISO 8802.7, for example, describes a slotted ring LAN used primarily by the UK academic community.

## 19.6 LOGICAL LINK CONTROL FOR LANs

A *local area network (LAN)* provides for the establishment of direct (OSI layer 2) connection between any two end devices directly connected to the LAN. However, although various different physical forms and topologies are possible (e.g. *ethernet*, *token ring*, etc.), it was quickly realized that all LANs are expected to be capable of the same basic function: carriage of data between software or *applications* running on two different computers. It therefore made sense to define a standard interface between the LAN and the computer software intended to communicate across it. This standard interface is called the *logical link control (LLC)* protocol. It is defined by IEEE standard 802.2 (ISO 8802.2).

The *logical link control (LLC)* provides a standard communication interface equivalent to that provided by OSI layer 2 to OSI layer 3 (datalink service, see Chapter 9). LLC in combination with the *medium access control (MAC)* protocol specific to the particular LAN (e.g. *ethernet*, *token bus*, *token ring*) is equivalent to an OSI layer 2 protocol.

The information carried by LLC consists of four fields, which together are termed the *LLC protocol data unit (PDU)*. The four fields are

- the *destination service access point (DSAP)*, an address which identifies the application or software session to be activated in the destination computer to receive the packet
- the *source service access point (SSAP)*, an address that identifies the application which sent the packet
- *control* information, which includes details of the type of connection (e.g. *connection-mode*, *connectionless*, *acknowledged connectionless*), the protocol in use at the next higher layer (e.g. TCP/IP, IPX, Appletalk, etc.)
- the *user data* (i.e. the raw data being transported)

## 19.7 LAN OPERATING SOFTWARE AND SERVERS

So far we have talked about the physical structure of LANs, and the logical procedures used to convey the information packets across them. This alone, however, is not a sufficient basis for creation of an office LAN. In addition, a *LAN operating system*

(software) is required. At the start, a number of different manufacturers offered alternative proprietary systems. Over time, the systems in use have reduced to five: *Novell Netware*, *IBM LAN Manager*, *Appletalk*, *Windows for workgroups* and *WindowsNT*.

LAN operating systems provide for the software *sockets* (i.e. interface) between normal computer operating software (e.g. *Microsoft DOS*, *Windows*, *Windows95*, *Apple Macintosh*, etc.) and the new functions made possible by LAN networks (e.g. file server, host gateway, fax server, common printer, etc.). LAN operating systems are closely linked to network protocols, many of which have a proprietary nature. Thus, for example, *Novell Netware* (network operating system software) in conjunction with the Novell IPX network protocol allows the personal computer user to use various types of 'network' services. For example, a PC on the LAN may be able to choose between any of the printers connected to the LAN rather than being limited to the one directly connected to his computer. Sometimes he might choose the fast black and white laser printer, whereas on other occasions the colour printer is more appropriate. Alternatively, a common *file server* might allow the LAN users to share a common data filing system. In this way each individual user has a wider choice of facilities, and overall less equipment is needed because the printers and other devices can be shared between large groups of users. Equivalent functionality can be provided using the *UNIX* operating system software in conjunction with TCP/IP or the *Apple Macintosh* system in conjunction with *Appletalk*.

*Servers* are typically powerful and expensive computers, capable of faster processing and additional functions useful to the workgroup as a whole. *Servers* are connected to the LAN, and usually remain in operation for 24 hours per day.

A *file server* is usually a computer with a large amount of storage capability which may be rapidly accessed and easily *backed up* by specialist computer staff on a once per day or once per week basis. It provides for secure storage of information and easy sharing of information basis on a workgroup or defined *closed user group* basis.

A *mail server* provides for the transmission of *electronic mail* letters between individual PC users connected to the LAN without the need for the users both to be connected to the network and have their PCs switched on at the time of sending or receiving the letter.

A *facsimile server* allows individual PC users to send printed documents directly to a remote facsimile machine without the need first to print the document to paper. The facsimile server itself is a device like a PC which is connected simultaneously to the LAN and to a facsimile/telephone connection. To the LAN, the facsimile server appears like a printer with LAN operating software, but instead of printing directly, the document is converted to facsimile (fax) format and transmitted over the telephone line to the remote fax device.

## 19.8 INTERCONNECTION OF LANS: BRIDGES, ROUTERS AND GATEWAYS

The interconnection of numerous LANS, perhaps of different types, or the connection of a LAN to a mainframe computer or other external network or device requires the use of *bridges*, *routers* or *gateways*. We discuss these in turn.

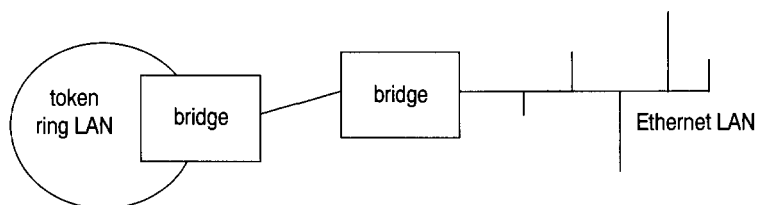
A *bridge* is used to link two separate LANs together as if they were a single LAN, typically enabling the maximum capacity of a single LAN to be surpassed, or two separate LANs in locations remote from one another to be connected as if they were a single LAN (a so-called *remote bridge*). The bridge is an intelligent hardware connected to the LAN, which examines the address in the *LLC (logical link control)* header of each packet or frame. For relevant frames, the packet is removed, passed across the bridge connection to the second bridge, where it is injected into the second LAN (Figure 19.5), which may have a different physical form (e.g. ethernet/token ring). Either a table of the relevant addresses must be kept up to date in each of the bridges, to determine which packets must be transferred into the second LAN, or simply all packets are bridged.

A *remote bridge* differs from a *local bridge* only in that a *wide area* type connection (e.g. X.25 connection or leased line connection) is used to connect the bridges. Usually only the packets destined for the remote LAN are bridged by a remote bridge, so that the lower bitrate of the bridge connection (typically 9600 bit/s) does not become a bottleneck.

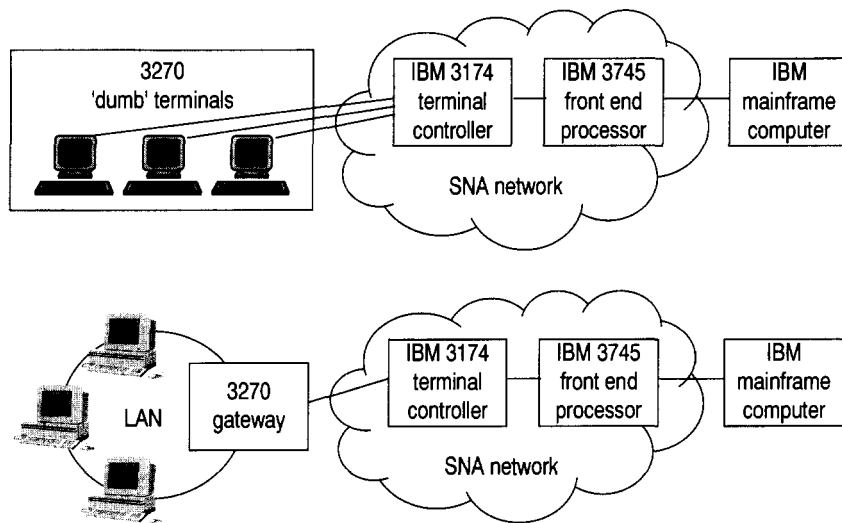
Although bridges provide for a relatively cheap means of interconnecting LANs, they are not to be recommended in large, complex networks, because they result in very complicated topologies which are extremely difficult to manage. Thus, for example, a bridge network of three LANs could have three bridge connections, connecting the LANs in a triangle fashion. The problem now is to ensure that the appropriate bridge connection (i.e. the direct one) is used when transporting frames between any pair of the LANs. In very large networks the chance of optimal path-finding is very low, so that there is a great risk of endless circular paths. To overcome this problem, the *router* appeared.

*Routers* are much more intelligent devices than *bridges*. They are designed to 'learn' the topology of complicated networks (even ones which are constantly growing or changing) and accordingly route frames or packets across them to the destination indicated in the header. Routers learn about network changes through experience. Crudely put, if they receive a packet or message for a destination which they do not recognize, they choose a route at random and see if it is successful. On following occasions, the previously successful route is selected. In this way, communication is possible even across very complicated and cumbersome networks which have been built by different parties and simply connected together.

Many LAN protocols (e.g. *Novell's IPX*, *Appletalk*, etc.) may be routed in their native (i.e. raw) form, but it is nowadays increasingly common instead to use *the transmission control protocol/internet protocol (TCP/IP)* as the main protocol to interconnect complicated LAN networks.



**Figure 19.5.** LAN bridge



**Figure 19.6** Replacement of dumb terminals using a LAN and 3270 gateway

TCP/IP uses a worldwide standardized addressing scheme (*Internet addressing*) to identify end user stations uniquely. This provides the ability to connect all LANs worldwide into a single common network, the *Internet*, thus extending the information sharing and electronic mail capabilities of single LANs to the world computer community as a whole. TCP/IP was originally developed as a transport protocol before the OSI layer 4 was fully defined. Its development was sponsored by the US government and military, particularly to provide for widescale interconnection of *UNIX* computers and workstations.

The problem with networks of multiple routers (including the *Internet* itself) is that the individual routes through the network are difficult to monitor and manage. It is difficult to know which networks are being transitted along the way, so that optimal network loading and the security of information cannot be guaranteed, though emerging standards will address these shortcomings. We discuss the *Internet* and *TCP/IP* protocol in Chapter 22.

A *LAN gateway* provides access for a LAN user to an external service, such as a mainframe computer. Typically a *gateway* consists of a Personal Computer (PC) on the LAN used entirely to run gateway software. An example of a widely used *LAN gateway* is a 3270- or *SNA-gateway*. *IBM 3270* is the communication protocol used between the host computer and the terminal of an IBM mainframe. The protocol (part of SNA, Chapter 18) allows the computer to interpret keyboard interaction at the terminal and control the exact image appearing on the terminal screen, without there needing to be 'intelligence' in the terminal. Thus 3270 type terminals are sometimes described as *dumb terminals*. The conversion necessary to make an intelligent terminal (such as a personal computer) appear to talk to a host computer like a dumb terminal is carried out by *terminal emulation* or *3270 emulation software*. Where this software resides in a *LAN gateway*, then it is called a *3270 gateway* (Figure 19.6).





---

## Frame Relay

---

For data transfer, X.25-based packet switching has established itself worldwide as a standard and very reliable means. However, X.25 is not a technique suited to the higher quality and speeds of modern data communications networks, and so it is beginning to be supplanted by new techniques, among them 'frame relay'. In this chapter we start by discussing the shortcomings of X.25-based packet switching in carrying highspeed bitrates and explain how frame relay was designed to overcome these problems. We conclude with a more detailed review of the frame relay protocols themselves.

### 20.1 THE THROUGHPUT LIMITATIONS OF X.25 PACKET SWITCHING

The reliability of X.25 packet switching has resulted from worldwide accepted standards and the huge availability of compatible hardware and software products enabling computer devices made by different manufacturers and strewn around the world to intercommunicate without difficulty. X.25 was the first universal data communication protocol and it stimulated rapid growth in data communication traffic volumes, because of its reliability and its robustness. Paradoxically, its robustness is now leading to the demise of X.25, because one of the main limitations of packet switching based on X.25 is its unsuitability for carriage of high speed information channels and its relative inefficiency when used in conjunction with high quality transmission networks.

When X.25 was developed in the late 1970s, the relative speed of the communicating devices was very low (in comparison with today's devices, typically under 9600 bit/s) and the quality of wide area digital lines was comparatively poor. As a result (and to their credit) X.25 packet networks are highly robust against poor line quality. X.25 networks are able to survive and even recover from even extensive bit errors on digital lines. The problem is that the cost of this robustness is the very limited linespeeds which are possible, and the relative inefficiency of line utilization in the case of higher quality lines.

The problems which arise when attempting to operate X.25 protocol at high speeds are due to the *windowing* technique employed by X.25 to help avoid errors. To

illustrate the problem we consider trying to use a 2 Mbit/s line to carry X.25 data over a distance of 1000 km.

As Figure 20.1 illustrates, on a high speed data transmission line there are always a large number of bits in transit on the line at any point in time (because of its length), in our example around 20 000 bits or 2500 bytes (line length  $\times$  bitrate/speed of light). (That should blow any preconception you might have had that electricity travels so fast that we can consider sender and transmitter to be in synchronism with one another!) These bits in transit on the line must be considered when designing high speed data networks, if the network is to operate efficiently.

X.25 lays a very high priority on the safe arrival of bits, in the correct order and without errors. One of the methods used to ensure safe arrival is the use of an acknowledgement *window*. Only so many packets (as defined by the *window size*, typically 7) may be transmitted by the sending device before an acknowledgement is received confirming safe arrival. As the typical maximum packet size is defined as 256 bytes, this means that a maximum of 1792 bytes ( $7 \times 256$ ) may be transmitted by the sender before an acknowledgement is returned by the receiver to confirm safe arrival. This compares with the 2500 bytes actually on the line, so that even before considering the inefficiencies caused by packet *overheads* (the *protocol control information* in the X.25 header) the X.25 *window* will constrain the efficiency of the line of Figure 20.1 to a maximum of 1792/2500 (maximum bits allowed in transit/available bits in transit) or around 70%!

Simple, you might say: increase the maximum window size! Unfortunately this only generates new problems. First, the end devices need to provide much greater storage buffers for retaining copies of the sent but unacknowledged information. Second, because the window size is greater, so is the likelihood of errors within a window. The probability of the need for a retransmission of the information to eliminate the errors is thus also greater. Also, because of the increased window size, the time required for retransmission is longer. So increasing the window size may actually reduce throughput!

Today's digital transmission is several orders of magnitude better quality than that of the 1970s, so that the heavy duty error detection and correction techniques used by X.25

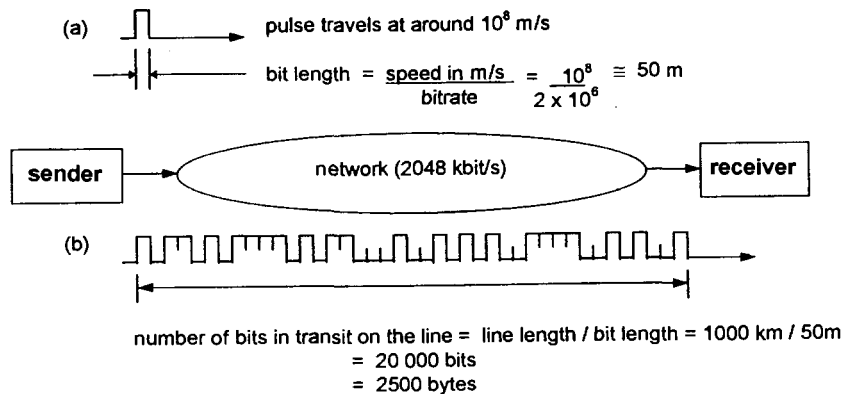


Figure 20.1 Bits in transit in an X.25 packet-switched data network

have become redundant. Windowing and acknowledgement are now largely superfluous. *Frame relay* (or *frame relaying*) was one of the first techniques designed to dispense with heavy duty error detection and correction techniques. Instead of it being undertaken by the network, the job of error detection and correction or recovery is left to *higher layer protocols* (i.e. the end user's device, computer or software) to sort out. The frame relay network, meanwhile, may concentrate on the raw information carriage and is thus more efficient and also more capable of higher information throughput. Thus for *wide area* computer and LAN-to-LAN connection needs of 64 kbit/s or greater, *frame relay* is today's preferred method. However, for bitrates above 2 Mbit/s, native ATM (see Chapter 26) should be considered.

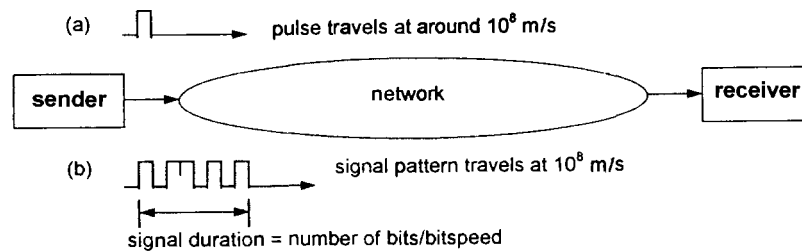
## 20.2 THE NEED FOR FASTER RESPONSE DATA NETWORKS

Although the basic need to carry high bandwidth signals drove the need for the development of the *frame relay* protocols, so did the need for faster responding networks. It may not be immediately obvious, but the time required to propagate even low bandwidth information across a digital network is dependent on the bitspeed employed: the higher the bitspeed, the lower the propagation delay. Even data applications with limited information transport needs appear to run faster when carried by a high speed network, even though sufficient bandwidth may have previously already been available. The following discussion gives two reasons why.

Let us imagine two rainwater conduits, one of small bore and one of large bore. Let us assume that the first has throughput capacity of 5 litres/s and the second of 10 litres/s. Now let us assume that the rainfall rate is 4 litres/s. Why should I bother with the large bore conduit? The answer is that the rainfall rate is not constant. Over the course of time the rate may vary between, say 2 and 6 litres per second, so that during moments in time when the rate of rainfall exceeds 5 litres/s, water will be accumulating in the roof gutter rather than flowing down the conduit. The accumulation clears when the rainfall rate drops momentarily below 5 litres/s. As a result of the momentary accumulation, some of the rainwater is delayed slightly, thereby increasing the *propagation delay*. The analogy is relevant to data transmission, where the rate of generation of typed characters or other data to be transferred is not constant, but can fluctuate wildly. The first reason why high speed lines give better performance is that they cope with short high speed bursts better. The second reason is that frames can be conveyed more quickly, as we explain next.

Figure 20.2 illustrates a more detailed example of data transmission across a telecommunication transmission line. In this case, the carriage of the electrical signals (i.e. the waveform pulses representing individual *bits* of a digital signal pattern) is at a speed close to the speed of light. Thus the leading edge of an individual pulse traverses the network at around  $10^8$  metres/s (Figure 20.2(a)). The time required, however, to transmit an entire frame of 1 byte (8 bits) is sensitive to the transmission bitspeed. The propagation time in this case is equal to the sum of the raw propagation time and the signal duration (Figure 20.2(b)).

Taking the example of a 9600 bit/s dataline of 100 km length (as might be employed in a corporate packet switching network today), we can calculate propagation times for both cases (a) and (b) of our example:



**Figure 20.2** Signal propagation time across a data network

- single bit (pulse) propagation time =  $10^5 \text{ m} / 10^8 \text{ ms}^{-1} = 10^{-3} \text{ s} = 1 \text{ ms}$
- byte propagation time = pulse propagation time + signal duration  
 $= 1 \text{ ms} + 8 \text{ bits} / 9600 \text{ bit s}^{-1}$   
 $= 1.8 \text{ ms}$

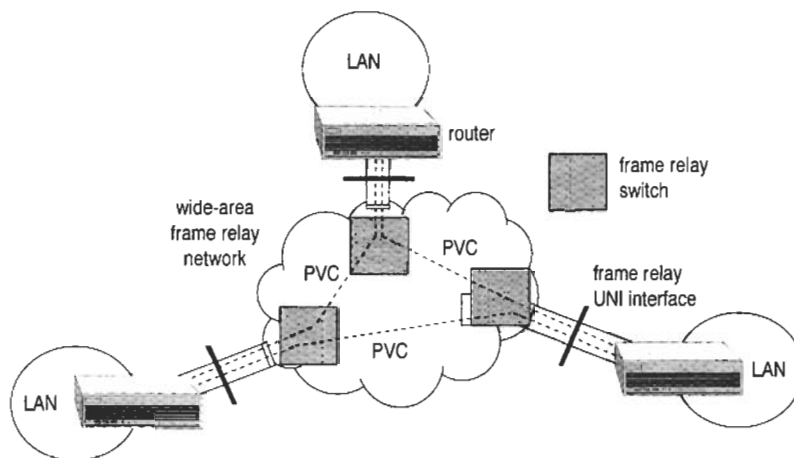
In other words, before enough bits (8) have been received to interpret the frame (in our case a data or ASCII character), 1.8 ms have elapsed. This compares with the 1 ms needed for conveyance of a pulse across the line. Despite the fact that the average throughput required from the line may be far less than the 9600 bit/s available (say 2400 bit/s or 300 characters/s), the effective propagation time of characters across the line is much longer than the 1 ms that you might expect.

No human being will notice the extra 0.8 ms, you might say? Indeed they will not where a simple one-way transmission is involved with a human end user. However, where an interactive dialogue is taking place between two computers (question-answer-question-answer), then this will take around 80% longer to conduct. A human waiting for the computer's response may see a response in around 4 seconds, where previously it was around 2 seconds. Such intercomputer dialogues are the main cause of delays for modern computer software. (Typical dialogues run 'please send first character' – 'first character' – 'received first character OK, please send next character' – 'second character' – 'received second character OK...' etc.)

The perhaps surprising reality is that it may indeed make sense to use a network with 64 kbit/s transmission links rather than 9600 bit/s links even though the average throughput is only 4000 bit/s ! (Reduction in byte propagation time from 1.8 ms to 1.1 ms). The following points are thus critical in the response time performance of data networks and associated computer applications software

- transmission line bitspeed (e.g. 9600 bit/s, 64 kbit/s, 2 Mbit/s, etc.)
- message, packet or frame length in number of bits
- the number of inter-computer interactions (request and response dialogue) necessary to complete an action before responding to the human user.

Though our example is of a lowspeed data application, similar principles apply for all types of data applications. Thus the higher the bitspeeds employed in the network, the faster the application response time.



**Figure 20.3** Typical use of frame relay to improve the wide area efficiency of LAN/router networks

## 20.3 THE EMERGENCE AND USE OF FRAME RELAY

The recent explosion in the number of *LANs* (*local area networks*, networks connecting personal computers within office buildings) and the need to interconnect LANs, as well as the growing number of *client/server* computing (UNIX) environments, have created the demand for high speed networking in data networks. Frame relay provides for relatively cheap *wide area* data communication at rates between 9600 bit/s and 2 Mbit/s, and has proved a viable alternative to leaselines, particularly in *router networks*.

Initially, *frame relay* networks only supported *PVC* (*permanent virtual channel*) service between pairs of fixed end-points. The service to the user was therefore somewhat akin to a 64 kbit/s leaseline, but without the full costs of a leaseline, because *statistical multiplexing* in the wide area part of the network allowed resources to be shared across a number of users and therefore costs to be saved by each of them (Figure 20.3). The pricing strategy of public network operators has also encouraged the use of frame relay service as a leaseline replacement service where high speed but relatively low volume usage is required, because flat rate charging based on the *committed information rate* (*CIR*) has become the industry standard.

## 20.4 FRAME RELAY UNI

In the arrangement of Figure 20.3, which is typical for a frame relay network, each of the routers is connected to the frame relay network using a single connection, typically of 64 kbit/s, employing the *Frame relay UNI* (*user-network interface*). Over this single physical connection, up to 1024 ( $2^{10}$ ) logical channels (*PVCs*, in the correct frame relay terminology, *data links*) may be connected, each to a separate end destination. These

channels are available on a permanent basis, but the capacity of the wide area part of the network is only used when there are actually *frames* requiring to be *relayed* from one end of the *data link* to the other.

In the *frame header* (like the HDLC header of X.25) is a numbered value identifying the logical connection to which the frame belongs. This is called the *data link channel identifier (DLCI)*. Unlike X.25, there are no real end-to-end network functions supported at OSI layer 3. These are the functions which provide within the network for reliable end-to-end transfer. Saving these functions simplifies the frame relay protocol (in comparison with X.25), making it more efficient and faster running.

As is also shown in Figure 20.3, it is common in frame relay networks to build fully meshed networks of logical connections (PVCs) between individual routers (a triangle in our case). This circumvents the need for the routers themselves to act as transit nodes for inter-router traffic, and so improves the overall performance perceived by the LAN users without having much effect on the overall cost of either the network hardware or the transmission lines needed in the *wide area network (WAN)*.

## 20.5 FRAME RELAY SVC SERVICE

The main drawback of the arrangement shown in Figure 20.3 is the management effort needed to establish and maintain the large number of PVC connections within the network. Potentially, each time a link in the network fails or a new link is added, administrative work may be necessary to reconfigure some or all of the PVCs to new, more efficient paths. To get around this problem, the Europeans in particular have driven the development of an enhancement of the frame relay UNI to include the capability for *on-demand* establishment of *data links* (i.e. *switched virtual circuits, SVCs*) using a dial-up procedure. Although this adds layer 3 functions to the frame relay protocol stack, these are only connection and clearing functions, which do not affect the subsequent end-to-end carriage characteristics (high speed, low delay as discussed).

The main benefit of an SVC network is that individual data links need only be established when needed and may be cleared afterwards. This simplifies the network and its management, and in addition has the effect of automatically optimizing the routing of connections each time they are newly established.

## 20.6 CONGESTION CONTROL IN FRAME RELAY NETWORKS

The high speed of the computer devices using frame relay networks leads to the need for special measures to control network congestion, because the layer 3 protocol is almost non-existent. Figure 20.4 illustrates a case in which one of the intermediate links within the *wide area* or *backbone* part of the network is in congestion. As a result, frames are accumulating rapidly (and unabated) in the buffer immediately preceding the congested link, as they wait to be transmitted over the link. Ultimately, the buffer will overflow and frames will be lost, so affecting all the *data links* sharing the congested link. Worse still, once the end user devices detect the loss of information, retransmission will commence, and the load on the network only further increases.

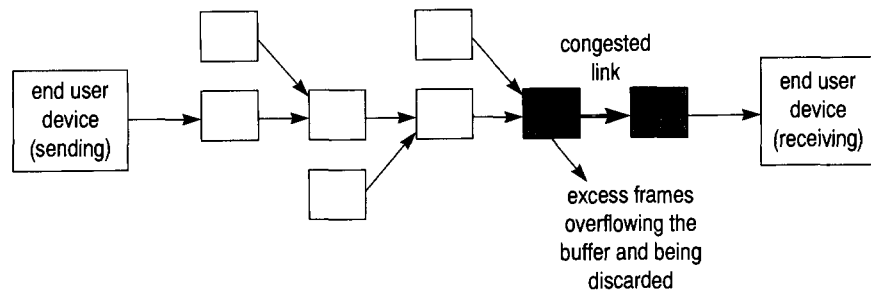


Figure 20.4 Congestion in a frame relay network

As the connection from the sending device of Figure 20.4 to the network is not congested, the sending of frames would continue unabated, were it not for the *congestion notification* procedures within the frame relay protocols.

Any time a frame underway within a frame relay network encounters congestion (the exceeding of a given waiting time or a given buffer queue length) then the frame header is tagged with a notification message, called the *forward explicit congestion notification (FECN)*. This notifies the receiving device and the switch to which it is connected of the congestion, which may then be communicated back to the source end by means of a *backward explicit congestion notification (BECN)* message. The sending device may voluntarily respond to receipt of the BECN by reducing its transmitted output, or may be forced by the network to do so. By reducing the frame transmission rate from relevant causal sources, the congestion will ease, and the transmission rate restriction may be removed. Meanwhile, further service degradation within the network as a whole is avoided.

A further refinement of the congestion control procedure of frame relay is provided by the *committed information rate (CIR)* and *excess information rate (EIR)* parameters. The *committed information rate (CIR)* is the agreed minimum bitrate to be provided by the network between the two ends of the frame relay *data link*. The CIR is agreed at the time of setting up the connection. Provided the frame transmission rate of the sending device is at or below the CIR, then the network is not permitted to force a reduction in the frame sending rate of the sending device, and is not permitted wilfully to discard frames. However, where the sending device is exceeding the CIR at a time when the BECN message is received, then the network may first request reduction in the rate of frame transmission to the CIR. Should the reduction not be undertaken (for example, because the sending device cannot respond to the request), then the network is permitted to discard the *excess* frames.

At times of no congestion, sending devices are permitted for defined short periods of time (called the *excess burst*, or *excess burst duration*,  $B_e$ ) to transmit at bitrates higher than the CIR. The maximum bitrate at which the device may send is termed the *EIR (excess information rate)*. The EIR is always greater or equal to the CIR. It is a management decision for the network operator how high EIR and CIR may be set for a given connection, and usually these values are included in the contract or order for the user's connection (for a PVC) or negotiated at connection set-up time for an SVC. The ability



to handle short bursts of high speed information (above the CIR) is what makes frame relay networks attractive to data applications requiring fast response times, as we discussed earlier in the chapter.

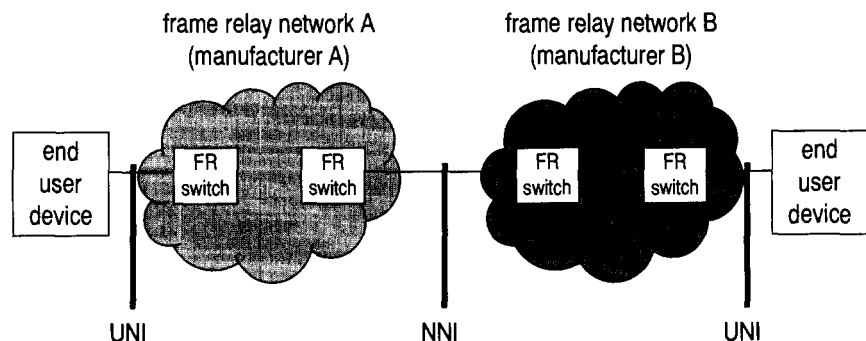
## 20.7 FRAME RELAY NNI

As in packet-switched data networks, it is common for all the switches within a given operator's network to be purchased from and supplied by a single manufacturer. The leading manufacturers of frame relay network components are the *Stratacom* and *Cascade* companies and *Northern Telecom (Nortel)*. As the switches are supplied by a single manufacturer, it is not necessary to use a standardized interface between the *nodes* within the network. As a result, the individual manufacturers have tended to develop extra congestion controls, network management and service features over and above those required by the frame relay standards, to try to improve the market value of their products. All very well; except, of course when a given frame relay connection needs to be switched across two different networks or sub-networks, supplied by different manufacturers (as in Figure 20.5). For this a standardized interface, the *NNI* (*network-network interface*) is required.

Although the frame relay NNI allows for interconnection of sub-networks of switches supplied by different manufacturers, and although reliable data transfer is possible, it is true that the congestion control and management capabilities of the combined network are much more restricted than the capabilities available within each of the sub-networks independently. This reflects the relative youth of the frame relay NNI standard.

## 20.8 FRAME FORMAT

Figure 20.6 illustrates the format of a single frame. It consists of five basic information fields, much like the data link layer format of X.25 (i.e. HDLC). The flag marks the beginning of the frame, delineating it from the previous frame. The address field carries



**Figure 20.5** Frame relay NNI (network-network interface)

Flag	Address Field	Control	Information field	Frame check sequence (FCS)

Figure 20.6 Frame format for frame relay

the DLCI (data link connection identifier), the equivalent of the *logical channel number (LCN)* of HDLC (i.e. is an OSI layer 2 address). In addition, the address field also contains control information (*command/response*), the *forward* and *backward explicit congestion notification (FECN and BECN)* discussed previously in this chapter, the *discard eligibility (DE)* indication and some extra fields used for *extended addressing*.

The control field contains supervision indication for the connection like *receiver ready (RR)*, *receiver not ready (RNR)*, etc. For user information frames, this field indicates the length of the frame. Such controls were discussed more fully in chapter 18 on X.25. These enable the two end devices to coordinate one another for the communication.

The *information field* contains the user information. This may be up to 65 536 bytes in length. Finally, the *frame check sequence (FCS)* is a *cyclic redundancy check (CRC)* code providing for error detection. We discussed CRC codes in Chapter 9.

## 20.9 ADDRESS FIELD FORMAT

Figure 20.7 illustrates the two octet (i.e. two byte) address field used in frame relay. The main function of the address field is to carry the *data link connection identifier (DLCI)*, which identifies the end device to which the frame is to be sent (OSI layer 2 address). The DLCI is a 10 bit field, allowing up to 1024 separate *virtual connections* to share the same physical connection.

1	2	3	4	5	6	7	8
(transmitted							
first							
DLCI (msb)				C/R		EA	
DLCI (lsb)		FECN		BECN		DE	EA

BECN = backward explicit congestion notification

C/R = command/response bit

DE = discard eligibility

DLCI = data link connection identifier

EA = extended addressing

lsb = least significant bits

msb = most significant bits

Figure 20.7 Address field format for frame relay

network layer link layer	Higher layer information	Higher layer information
	Q.933 (signalling)	Q.933
	Q.922 (core aspects)	Q.922 (Core aspects)
	physical layer	physical layer

Note: Q.933 protocol is the network layer protocol used for establishing switched connections (SVCs). It is not used in the PVC (permanent virtual connection) service.

**Figure 20.8** Protocol stack for frame relay

Should congestion be encountered by a given frame during its transit through the network, then the affected intermediate switch will toggle the *FECN* (*forward explicit congestion notification*) as we discussed earlier in the chapter. This alerts the receiving device of the congestion. In response, returned frames are marked using the *BECN* (*backward explicit congestion notification*). This allows the flow control procedures to be undertaken. Should congestion become so serious that frames need to be discarded, then frames marked with the *discard eligibility* (*DE*) set to '1' will be discarded first. The DE bit is set to '1' by the first frame relay switch (near the origin) on *excess* frames (i.e. those causing the information rate to exceed the *committed information rate* (*CIR*)).

## 20.10 ITU-T RECOMMENDATIONS PERTINENT TO FRAME RELAY

The following ITU-T recommendations define frame relay.

- Recommendation I.233 describes the frame relay service.
- Recommendation I.122 defines the framework of recommendations which specify frame relay, referring to the complete list of relevant recommendations.
- Recommendation Q.922 is perhaps the most important. It defines the *core aspects* of frame relay, specifically the *data link* procedure (i.e. frame format, address field etc.).
- Recommendation I.370 defines the congestion management procedures.
- Recommendation Q.933 defines the *signalling* procedures to set up *switched virtual connections*. This recommendation is not relevant for *permanent virtual circuit* (*PVC*) service.

Figure 20.8 shows the layered protocol structure of frame relay.

## 20.11 FRAD (FRAME RELAY ACCESS DEVICE)

Frame relay has historically been offered by public telecommunication operators as a cheap alternative to a leaseline in cases where high bitrates were desirable but the number of hours of usage per day was relatively low. To access a frame relay network, a

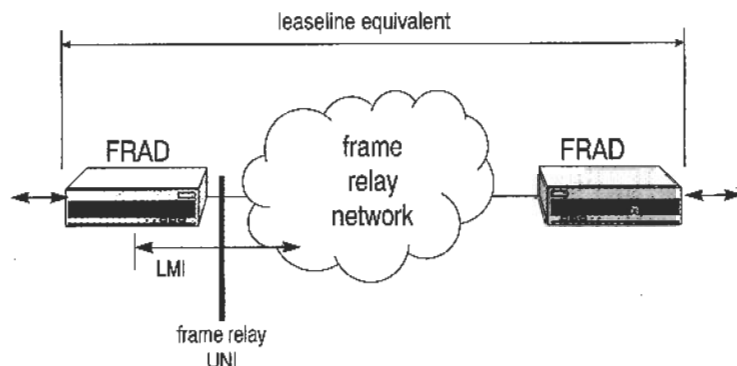


Figure 20.9 A frame relay access device (FRAD)

DTE (such as a router) must either incorporate a frame relay interface card, or alternatively use some other standard interface and an external conversion device. The most important of the available external conversion devices is the FRAD. *Frame relay access devices (FRADs)* provide for the conversion of *continuous bit stream oriented (CBO)* data signals into a *frame format*, in much the same way that a *PAD (packet assembler/disassembler)* converts the data stream from a 'dumb' computer terminal into the packet format required by an X.25 network. A FRAD may thus be used to convert a continuous data stream from an end user device normally expecting to use a 'transparent' lease line-like connection into a format suitable for carriage by a frame relay network. Figure 20.9 illustrates the principle.

Figure 20.9 illustrates the equivalent of a data lease line, created using two *frame relay access devices (FRADs)* and a frame relay network. This is usually more economical than the equivalent lease line where the CIR (committed information rate) is lower than the maximum speed of the lease line. Usually the EIR is set to the maximum speed of the equivalent lease line, and CIR is set somewhat above the average actual user information throughput rate. This virtually guarantees throughput of the user data (at the CIR). Bursts up to the maximum lease line speed (the EIR) are still possible for short periods, but at more 'quiet' times other users may share the trunk resources within the network on a 'statistical' basis. This is reflected in the lower costs of frame relay service compared to 'transparent' lease line service.

Also shown in Figure 20.9 is the *local management interface (LMI)*. This allows *status* and other *network management* information to be shared between the end terminal and the network in an SNMP-like format (see Chapter 27), so enabling overall monitoring and management of the network.



---

## *Campus and Metropolitan Area Networks (MANs)*

---

Metropolitan area networks (MANs) are network technologies similar in nature to local area networks (LANs), but with the capability to extend the reach of the LAN across whole cities or metropolitan areas, rather than being limited to, say, 100–200 metres of cabling. MANs have evolved because of the desire of companies to extend LANs throughout company office buildings spread across a campus or a number of different locations in a particular city. They provide for high speed data transport (at over 100 Mbit/s) and are ideal for the interconnection of LANs. There was some effort to extend MAN capabilities to include the carriage of telephone and video signals as an ‘integrated’ network, but this work has largely been overtaken by ATM (asynchronous transfer mode), so that the MAN technologies themselves are already obsolescent. We review here, but only briefly, the most important MAN techniques, FDDI (fibre distributed data interface), and SMDS (switched multimegabit digital service) which is based on the DQDB (distributed queue dual bus) technique.

### **21.1 FIBRE DISTRIBUTED DATA INTERFACE**

The *fibre distributed data interface (FDDI)* is a 100 Mbit/s token ring network. It is defined in IEEE 802.8 and ISO 8802.8. FDDI can be used to interconnect LANs over an area spanning up to 100 km, allowing high speed data transfer. Originally conceived as a high speed link for the needs of broadband terminal devices, FDDI is now perceived as the optimum backbone transmission system for campus-wide wiring schemes, especially where network management and fault recovery are required. In particular, FDDI became popular in association with the very first optical fibre building cabling schemes, because it provided one of the first means to connect LANs on different floors of a building or in different buildings on a campus via optical fibre. Unfortunately, due to its expensive nature and the rapid development of ATM (asynchronous transfer mode, see Chapter 26) as well as alternative building cabling schemes, FDDI has fallen into decline, no longer being recommended or further developed by most LAN and computer manufacturers.

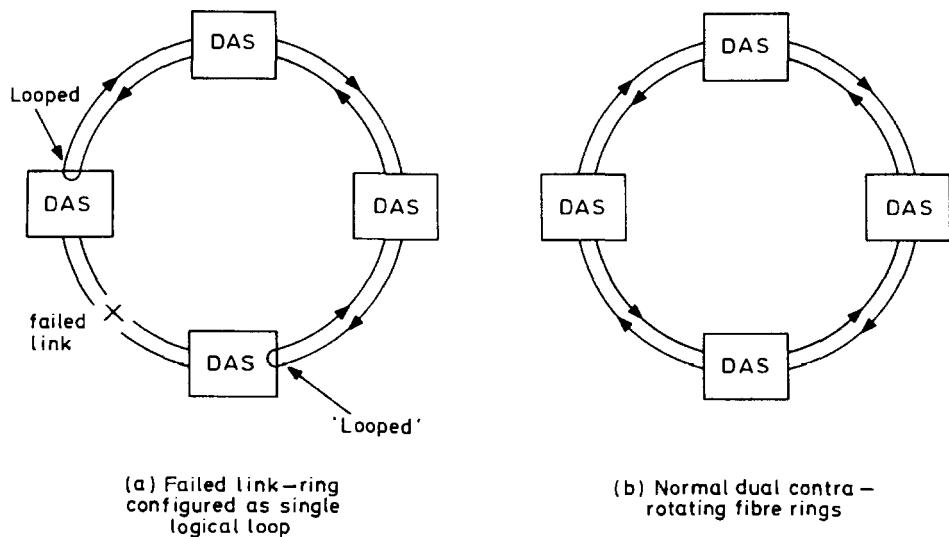
A second generation version of FDDI, FDDI-2, was developed to include a capability similar to circuit-switching to allow voice and video to be carried reliably in addition to packet data, but these capabilities were never widely used.

The FDDI standard is defined in four parts

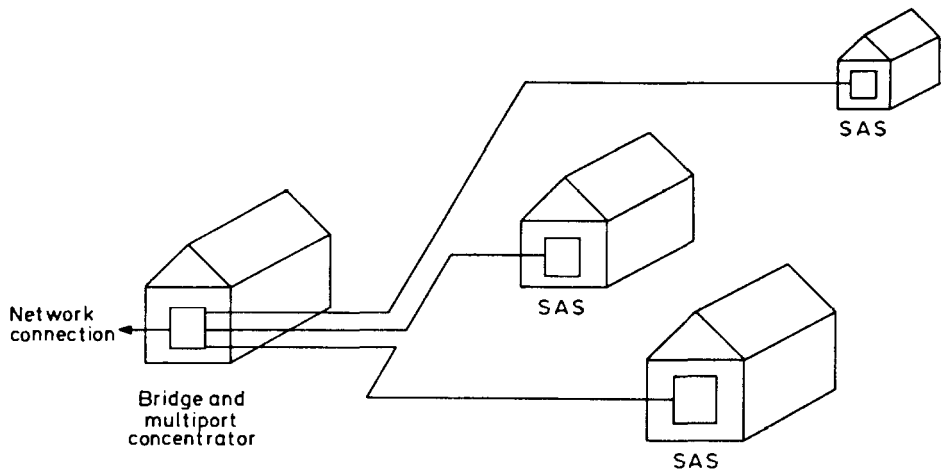
- *media access control (MAC)*, like IEEE 802.3 and 802.5 (see Chapter 19) defines the rules for token passing and packet framing
- *physical layer protocol (PHY)* defines the data encoding and decoding
- *physical media dependent (PMD)* defines drivers for the fibre optic components
- *station management (SMT)* defines a multi-layered network management scheme which controls MAC, PHY and PMD

The ring of an FDDI is composed of dual optical fibres interconnecting all stations. The dual ring allows for fault recovery even if a link is broken by reversion to a single ring, as Figure 21.1(a) shows. The fault need only be recognized by the *CMTs* (*connection management mechanisms*) of the station immediately on either side of the break. To all other stations the ring will appear still to be in its normal contra-rotating state (Figure 21.1(b)).

When configured as a ring, each of the stations is said to be in dual-attached connection. Alternatively, a fibre star connection can be formed using *single-attached stations* with a *multiport concentrator* at the hub (Figure 21.2). *Single-attached stations (SASs)* do not share the same capability for fault recovery as *double-attached stations (DASs)* on a dual ring.



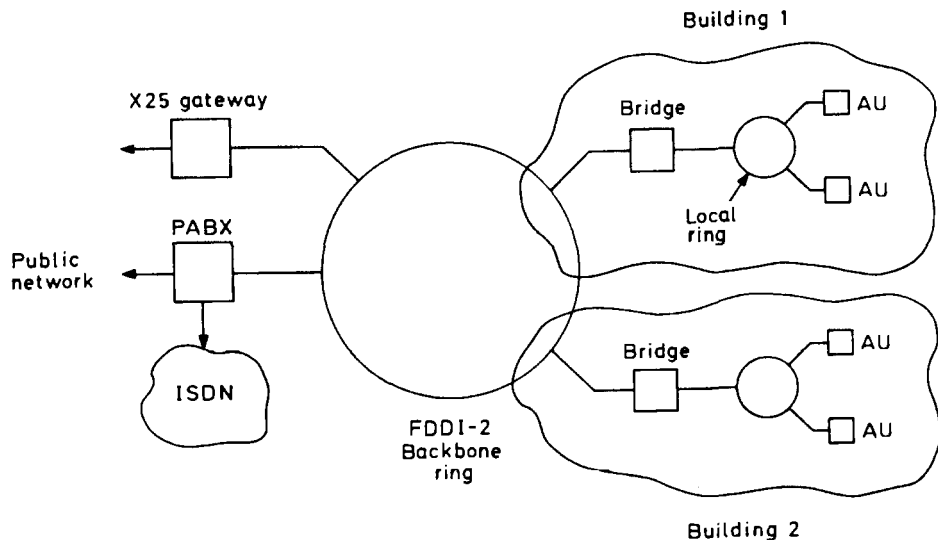
**Figure 21.1** The fibre distributed data interface (FDDI) fault recovery mechanism for double attached stations. DAS, double attached station



**Figure 21.2** Star configuration of FDDI. SAS, single attached station

Like *token ring LANs* (IEEE 802.5) and *ethernet LANs* (IEEE 802.3), FDDI is essentially only a physical layer (OSI layer 1) and data-link layer (OSI layer 2) standard. At layers 3 and above, protocols such as X.25, TCP/IP may be used.

FDDI-2, the second generation of FDDI (Figure 21.3) has a maximum ring length of 100 km and a capability to support around 500 stations including telephone and packet data terminals. Because of this, it was intended to support entire company telecommunications requirements. ATM, however, has proved a more popular prospect for



**Figure 21.3** The fibre distributed data interface-2 (FDDI-2). AU, access unit



providing these capabilities, and is now widely available from network and computer equipment manufacturers.

The FDDI-2 ring is controlled by one of the stations, called the cycle master. The cycle master maintains a rigid structure of cycles (which are like packets or data *slots*) on the ring. Within each cycle a certain bandwidth is reserved for circuit-switched traffic (e.g. voice and data). This guarantees bandwidth for established connections and ensures adequate delay performance. Remaining bandwidth within the cycle is available for packet data use.

The voice and video carriage capability of FDDI-2 is possible because of its interworking with the *integrated voice data (IVD)* LAN standard defined in IEEE 802.9.

## 21.2 SWITCHED MULTIMEGABIT DIGITAL SERVICE (SMDS)

*SMDS* (*switched multimegabit digital service*) networks conform to IEEE 802.6 and use a protocol called *distributed queue dual bus (DQDB)*. DQDB was co-developed by Telecom Australia, the University of Western Australia and their joint company, QPSX Communications Limited. It was designed to provide a basis for initial broadband *metropolitan area* interconnection of networks, but also give a possible migration path to B-ISDN (Chapter 25), for which it is now an optional access protocol. As a public data communication service, the *switched multimegabit digital service (SMDS)* became available in the United States in 1991.

The *DQDB* protocol uses two *slotted buses* of bitrates up to 155 Mbit/s to transport *segments* of information between communicating broadband devices. *Segments* are 48 byte frames of user data information.

Figure 21.4 illustrates the structure of a network using the DQDB protocol. Two unidirectional high speed buses run out from *master* and *slave frame generators* at opposite ends of the ribbon topology. Each of the devices (*nodes*) connected to the network are connected to both buses to send and receive data.

The role of the frame generators is to structure the bit stream carried along the buses into 53 byte *slots*. These *slots* are filled by nodes wishing to send user information and

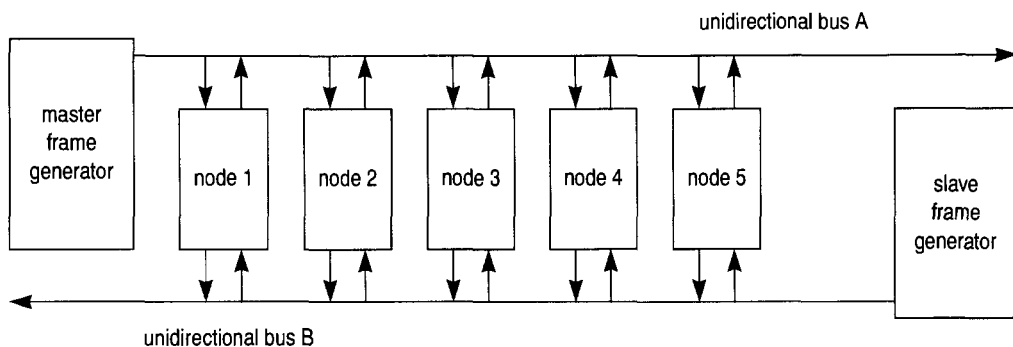


Figure 21.4 Bus structure of DQDB

are then carried downstream along the bus. The relevant receiving node reads information out of the slot being sent to it, but does not delete the slot contents. The slot thus remains on the bus, travelling further downstream until it falls off the end.

When a node wishes to send information it may do so in the first available empty slot, but in doing so must follow the procedure set out in the *medium access control (MAC)* protocol. The MAC protocol is intended to ensure a fair use of the available bandwidth of the buses between all the devices wishing to send information.

Before sending information, a sending node must know the relative position of the receiving node on the bus. It then sends a request in the opposite direction of the receiving node on the relevant bus. For example, say node 2 of Figure 21.4 wished to transmit to node 5, then it would send a request on bus B. This advises the *upstream* nodes of bus A (i.e. node 1 in our case) that it requires capacity on bus A. Node 2 must then wait until all other *previously pending* requests from other *downstream* nodes on bus A have been cleared. Once these are cleared, it may send in any free slot, and may continue to fill slots until a further slot request appears from a downstream node.

It is a simple and yet very effective *medium access control*. Requests for use of bus A are sent on bus B. Meanwhile the use of bus B is governed by the requests on bus A. The control of the use of the network is decentralized, so that each node may independently determine when it may transmit information, but must be capable of keeping track of the pending requests.

When a node is not communicating on one of the buses (say bus A), it monitors the requests for use of the bus, keeping a running total of the outstanding requests using its *request counter*. Each time a request passes on bus B, the *request counter* is incremented, and when a free slot goes by on bus A the counter is decremented. In this way it can keep track of whether a free slot on bus A is *available* to it or not. The *request counter* is never decremented to a value less than zero.

Each time a node has a segment it wishes to send on bus A, it generates a *waiting counter*. The initial value copied into the *waiting counter* is that currently held in the *request counter*. The *waiting counter* is decremented each time a free slot passes on bus A until the value reaches '0', when the segment may be sent in the next free slot.

When transmitted onto one of the buses the 48 byte *segment* of user information is supplemented with a 4 byte *segment header*, a 1 byte *access control field* and a 4 byte *slot header* as shown in Figure 21.5, so that the total length of a *slot* is 57 bytes.

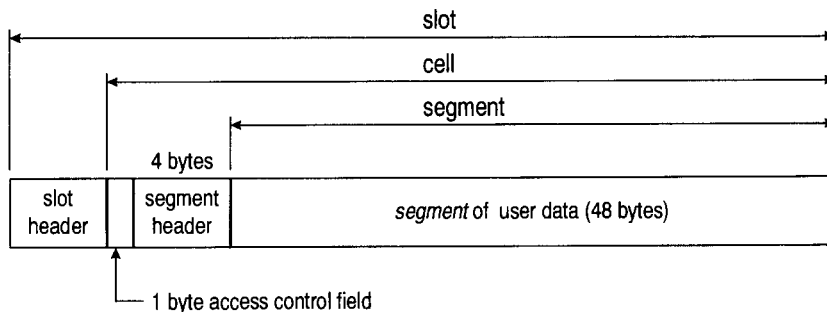


Figure 21.5 Slot and segment structure of DQDB

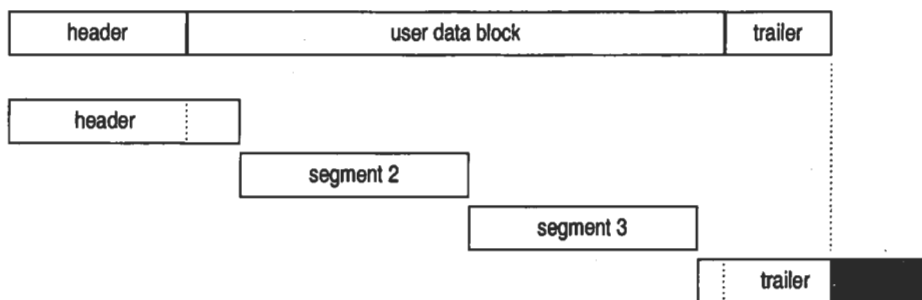


Figure 21.6 Segmenting a data block for transmission using DQDB

The *slot header* carries a 2 byte *delimiter field* and 2 bytes of control information used by the physical layer for the layer management protocol. The *access control field* may be written to by any of the nodes on the bus. This is the field in which the slot requests are transmitted. The *segment header* carries a 20-bit *virtual channel identifier*, like the *logical channel number* (OSI layer 2 address) of HDLC. This identifies the cells to the appropriate receiving node.

Data blocks to be carried by DQDB are formatted in the standard manner of frame *header*, the *user data block* and the frame *trailer*. The *frame header* contains the address of the originating and destination nodes. The *user data block* is the data frame to be carried which may be up to 9188 bytes in length (192 segments), and the *trailer* includes the *frame check sequence*. Data blocks must be broken down into individual segments and then formatted as slots for transmission. If necessary, the last segment is filled with *padding* (Figure 21.6).

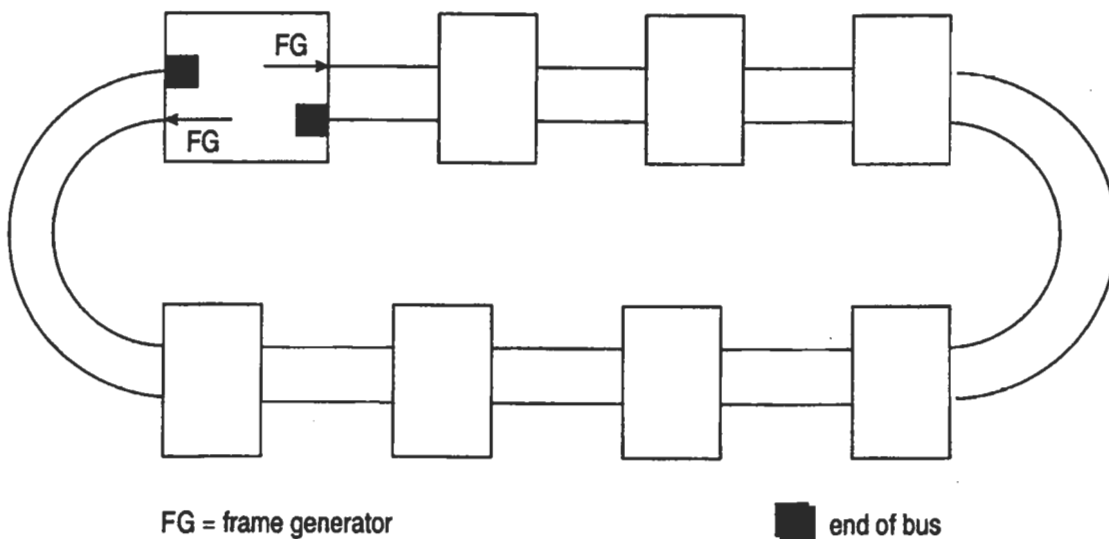


Figure 21.7 DQDB or SMDS configured in a looped bus topology

Networks using the DQDB protocol may also be configured in a *looped-bus* topology. In this case the bus is looped so that the two frame generators (Figure 21.4) are contained in the same node. This node also contains two *ends of bus* devices (Figure 21.7). In real terms, the network is still two independent buses, but there may be a practical advantage in not needing two separate frame generator nodes.

When offered as a public network service, SMDS is usually configured as shown in Figure 21.8, the public network node acting as the master frame generator and access point for a wide area broadband network which may use a protocol other than DQDB for wide area transport of information. In this way SMDS may provide an access network protocol for a broadband network based upon ATM (Chapter 26). As you may note from comparing the two technologies, they have a number of features in common (cell size of 53 bytes, *virtual channel identification* of individual channels, etc.)

Although the DQDB protocol has the charm of being a very simple and purportedly 'fair' protocol, one of the debates that dogs its wider acceptance is the doubt which exists over its 'fairness'. The slot request procedure used in the MAC does indeed help to share out the bandwidth resources between all the competing nodes, but it does not work well when many of the nodes wish to send at a bitrate close to that of the line. Let us return to Figure 21.4 and assume that the network has been idle, but that now both nodes 1 and 4 wish to transmit to node 5, both at the maximum bitrate. Node 1 starts sending immediately on bus A in every slot. Node 4, meanwhile, must first lodge a slot request on bus B. The request takes a little time to propagate along bus B until reaching node 1, whence node 1 must leave a free slot on bus A. It then goes on to use all subsequent free slots. As node 4 is only allowed to have one outstanding *slot request*, it must wait until this request is used up before generating the next one. Meanwhile node 1 is hogging all the slots.

The 'fairness' problem is particularly acute when a very long bus is used, because an entire slot is only about 900 metres long at a bitrate of 155 Mbit/s ( $57 \times 8$  [bits per slot]  $\times 3 \times 10^8$  (speed of propagation in m/s/ $155 \times 10^6$  bits/s)). Thus for a 10 km bus there will always be 11 slots between nodes 1 and 4, always with one reserved for use of node 4 and the other ten in use by node 1.

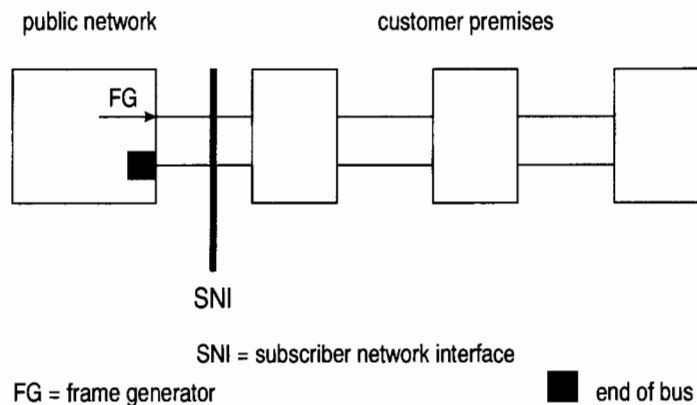


Figure 21.8 SMDS subscriber network interface

### 21.3 THE DEMISE OF MANs

Because of the emergence of *ATM (asynchronous transfer mode)* as a universal network technology for the carriage of all types of voice, video and data information in *local, metropolitan* and *wide area networks*, the MAN technologies are already obsolescent. This is strong evidence of the rapid pace of development of modern technology, but a chilling reminder of the costs and risks involved in investing in the development of or purchase of new equipment.

---

## *Electronic Mail, Internet and Electronic Message Services*

---

The ability to connect two computers together and, with relative ease, to send information from one to the other, is bringing a revolution in the way in which business and life as a whole is conducted. Today it is possible to run your bank account from home, book your holiday, send electronic messages to your work colleagues or friends, and look up to see what is on at the theatre in London or New York City. Alternatively companies may make their orders to and pay their bills from their suppliers by computer program and 'electronic data interchange'. A number of technologies have enabled this revolution: videotext, electronic mail, electronic data interchange (EDI) and Internet. We review the telecommunication aspects of these technologies in this chapter.

### 22.1 VIDEOTEXT

Videotext was the first type of device specially designed to allow telephone network customers to use a cheap device to access information from a large public database. The original technology and standards, including special terminals and modem techniques allowing asymmetric transmission (a high bitrate channel for information download to the customer with a low speed control channel for his ordering of different *pages* of information) have now been overtaken by modern personal computer based technology, but the appearance of the worldwide *Internet* has stimulated recent rapid growth for the long established videotext service providers (British Telecom's Prestel, Deutsche Telekom's *Bildschirmtext*, *BtX* or *Datex-J*, and France Telecom's *Minitel*).

The idea of videotext is that, using a low cost terminal in the form of a small television, a customer can make a phonecall to a public central database, where he could access all sorts of pages of information which he could then have displayed on his screen. Thus, for example, he might access tomorrow's weather forecast, current flight arrival information, information about financial markets, about holiday offers, about dating services or, via an extended connection to the railway company or his bank, might even order a train ticket or pay his bills.

Videotext in its original form in the United Kingdom (as *Prestel*) was quickly accepted by the travel industry, which used it widely as a means for holiday companies to advise travel agents of available itineraries and to book tickets on behalf of clients. It came at a time of revolutionary computerization of the airline industry in particular, when simultaneously the airlines were cooperating to establish and operate common and inter-linked computer ticket and booking systems.

In France, videotext became the standard method for telephone directory enquiries. The *Minitel* terminals were given away free by France Telecom. They justified doing this initially by the cost savings brought about by fewer telephone enquiries to human operators, but nowadays they also make money from the telephone calls which customers are making to the other information services which have subsequently sprung up. Some services, including the *pink pages* of dating services, are world renowned.

In Germany, the Deutsche Telekom's (formerly Deutsche Bundespost's) *T-Online* service (the new name for the modern version of *Bildschirmtext*, *Btx* or *Datex-J*) had been a commercial non-runner for years. Suddenly, however, during 1995 demand for connections started to increase rapidly. The reason was the cheap availability of access to the *Internet* using standard personal computers. It is the *Internet* which has become the world's electronic noticeboard. We discuss it later in this chapter.

## 22.2 ELECTRONIC MAIL (E-MAIL)

Electronic mail (*e-mail*) is a reliable means of message communication between human users equipped with computer terminals or personal computers. Large tracts of text, and diagrams too, can be quickly delivered across great geographical distances, and if necessary printed to a high quality paper format using local computer printing resources.

To send an *e-mail*, a user simply calls up the electronic mail software on his terminal. This prompts him for the name of the user that he wishes to send his message to, and for the names of any individuals to whom the message is to be copied. Having filled in this information, it prompts him to type the main text of the message. If he wishes, the sender can add further electronic attachments (say a document previously created using his *word processor*, his *spreadsheet* program or *presentation* software).

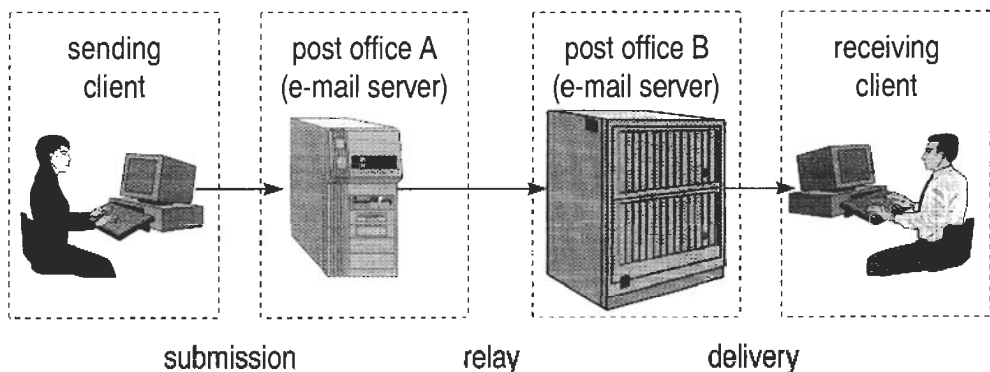
Once the *e-mail* is ready for sending, the sender simply clicks his mouse on the 'send' button. At this point, he may be asked to set a priority rating for the message, to decide whether confirmation of receipt is required, and (optionally) to set an exact time and date for delivery (if not to be immediate; this feature can be used as a reminder function, for example by sending time-delayed messages to himself). The delivery into the electronic mailboxes of the intended recipients is almost immediate.

Should any of the recipients be at their computer terminals at the time of the message receipt in their mailbox, then they may be advised of its arrival (say by a beeping noise and a short message or *icon* displayed at the bottom of their computer screen). They have the option to read the message immediately or later, depending on how they regard its priority. Those recipients who are not concurrently using their terminals will be advised of the message next time they log on. As messages are read, confirmation is returned to the sender (if required). Each recipient has the choice to reply to the message, forward it, file it, print it out, amend it or delete it.

Electronic mail systems are usually constructed in a *client/server* configuration. The *client* is the human user and his personal computer, workstation or other computer terminal. The *client* prepares and reads electronic messages. The client software prepares the *e-mail* for *submission* to the *server* or *post office*. The *post office* has a distribution function for *electronic mail* similar to the function of the postal service for letter mail (Figure 22.1).

The client software for electronic mail usually runs on a personal computer, and normally has software interfaces to other PC software such as word processing software, spreadsheet software and presentation software, so enabling easy attachment of such software to e-mail messages. Thus, for example the Microsoft Office package of software includes Microsoft Word, Excel, Powerpoint and Microsoft Mail software. Together these allow a personal computer user to prepare various types of different text, spreadsheet and presentation documents and submit them to an electronic mail post office as e-mail messages.

The post office function which forms the core of electronic mail usually resides on a server, often nowadays connected to the individual PCs by means of a LAN (local area network), though other types of X.25 packet network or dial-up connection are also possible. The post office is a combination of software and hardware, capable of storing e-mail messages, sorting them and transmitting them. The post office function works both on a store-and-forward and store-and-retrieve fashion. At the sending end, the user submits his message to the post office (post office A of Figure 22.1). This post office temporarily stores the entire message while it examines the address of the intended destination and establishes a communication means to the post office (post office B) serving the destination user. The communication means may take any number of different forms (X.25, telephone connection, internet connection etc.). Once the lower layer connection is established between origin and destination post office, or between origin post office and intermediate (transit) post office, the higher layer protocol (i.e. OSI layer 7 protocol for electronic mail (e.g. ITU-T X.400, SMTP or equivalent)) takes over, to ensure the appropriate relay of the entire message to the next post office. In each post office along the way, the message is stored and forwarded in this way.



**Figure 22.1** The basic components of an electronic mail system



Finally, at the destination post office, the message is stored in the electronic mailbox of the destination user until that user retrieves it. This is done when he logs into his own electronic mailbox account to read or send messages.

ITU-T, in its X.400-series of recommendations, specifies a message handling system (MHS) for use as an OSI layer 7 protocol in support of electronic mail. This is described in detail in Chapter 23. MHS defines P1 and P3-interfaces for relay and submission/delivery respectively. In reality, however, the P1-interface has proved to be the more important of these two, since this is the interface which allows electronic mail servers (post offices) to inter-communicate with one another. Thus the *electronic mail systems* of different companies, and maybe supplied by different software manufacturers are able to transfer *e-mails* between themselves. This is the X.400-interface talked of by *electronic mail* software suppliers. The P3-interface is only needed when a remote electronic mail user (*client*) wishes to submit or receive electronic mail messages directly from a post office operated by a public X.400-based electronic mail service provider. As, however, most electronic mail systems and post offices are based on corporately operated systems, it is usual for *client* and *post office (server)* software to be provided by the same supplier (e.g. *Microsoft mail* or *Ccmail*). In this case, it is not necessary for the software manufacturer to use the P3-interface, and instead a *proprietary* interface is employed.

An alternative to the use of the X.400 interfaces for inter-post-office interfaces is the use of the Internet. The *Internet* interface looks likely to become the predominant interface. E-mail based on the *Internet* is already firmly established in North America, and is fast supplanting X.400 where this already exists (mainly in Europe).

In addition to the X.400 and *Internet*-based electronic mail systems, a number of other *proprietary* systems have emerged. One of the first was introduced by *Compu-serve*, one of the world market leaders in *on-line services*. *Compu-serve* created the capability for companies and private individuals, from their personal computers, to access database information and send messages between one another. Its success has been largely based on being the first company to establish a worldwide user group for electronic communication. Currently it appears to be re-positioning to encompass *Internet* technology. *Microsoft*, on the other hand, has chosen to build a *Microsoft mail* capability into its *Windows95* operating system software for personal computers. This *client* capability can be used either in conjunction with a corporate *post office*, or alternatively, may use the *Microsoft Network (MSN)*, a public *post office* network operated worldwide by Microsoft Corporation. The initiative lays down a tough challenge for other X.400- and *Internet*-based electronic mail service providers.

## 22.3 ADDRESSING SCHEMES FOR ELECTRONIC MAIL

Three main types of addressing scheme are used nowadays for *electronic mail*. These are

- proprietary addressing schemes
- *Internet*-based electronic mail addresses
- X.500 addresses (for the *X.400 message handling system*)

Proprietary addressing schemes are normally employed within a corporation's electronic mail environment. The form of these addresses is often determined by the corporation itself, within the bounds set by the software product on which the system is based.

Internet-based electronic mail addresses (correctly SMTP addresses) usually take a numerical form, punctuated with dashes (–), dots (.) and @ signs (e.g. 12345.67890@compuserve.com). More commonly nowadays, an *alias* address appears on businessmen's visiting cards, e.g. Martin.Clark@serviceprovider.corporation. This form allows for easy recognition and memory of the address.

X.500 addresses are less common, but form the basis of the ITU and ISO *directory service* and are used in X.400 electronic mail networks. We describe the X.500 addressing scheme in Chapter 28.

## 22.4 THE ADVANTAGES AND DISADVANTAGES OF E-MAIL

*E-mail* messages are delivered quickly. Broadcasting of messages is also quickly and easily achieved. Editing of text and returning or forwarding the amended version can be achieved with minimal re-typing. Messages can be filed and quickly retrieved later.

Messages can be posted for exactly timed delivery, and can be prioritized according to the urgency with which they need to be dealt with. Furthermore, users are able to check their electronic mailboxes even when they are away from their normal offices, either by using somebody else's terminal or perhaps by dialling in to the post office using a portable *laptop* computer from a hotel room.

The *Internet*, in particular, has done a lot to release the full power of electronic mail, by enabling a large worldwide community of computer users to inter-communicate electronically, thus destroying the previous communication boundaries between companies and overcoming all the barriers of geography and time zones.

Companies who have successfully introduced electronic mail have observed a beneficial change in the whole culture of how they do business. Questions and responses have been much quicker and more direct, messages have been much shorter and less formal; they have been typed by the managers rather than by their secretaries. Workgroups composed of members in widespread locations have evolved, and it is possible to draw together new teams for previously impossible tasks.

## 22.5 EDI: CORPORATE COMMUNICATION WITH CUSTOMERS AND SUPPLIERS VIA E-MAIL

Where electronic mail is used as the basis of formal communications between companies for orders, payments and confirmations, it is termed instead *electronic data interchange (EDI)*. EDI between companies has become an important way of doing business. Many manufacturers and retailers demand that their suppliers accept orders for goods electronically. This gives the scope for more frequent ordering (in the case of supermarkets, for example, daily based on yesterday's sales) and so maintain their shelf stocks (and value of stocks) at a minimum level. This is *just-in-time (JIT)* provision.

Not only the buyer benefits, but also the supplier can ensure that products on the shelves with his name on are fresher and closer to their best. Even more important for the supplier, EDI allows him to bill for the goods more quickly.

The most important of standards defining the format and content of messages for *electronic data interchange (EDI)* are those published by the United Nations as *EDIFACT (electronic data interchange for administration, commerce and transport)*, but there also are a number of regional, national and industry-specific standards (e.g. *Odette* used by the automobile industry).

EDI products are available in ready-packaged software solutions for integration into a company's existing computer and electronic mail set-up. Specialized service providers have also appeared in several countries concentrating on the communication needs of companies within a specific industry segment. In this way, a *community of interest* may be established for communication between customers and suppliers (for example, within the retailing industry or the car manufacturing industry).

Setting up a network and software for EDI is actually quite straightforward. What is much harder, and what the company therefore needs to be prepared for, is the adjustment of its culture and practices to an entirely new way of business, one requiring much quicker reaction to customer demands.

## 22.6 INTERNET

The *Internet* emerged from a United States government and military initiative to enable the interconnection of different, mainly UNIX-based computer systems for intercommunication. As UNIX was proclaimed to be the first 'portable' operating system for computers, enabling software developed on a particular manufacturer's computer hardware to be easily *ported* (i.e. transferred for operation) to another manufacturer's hardware, it was natural also to develop means for easy transfer of data between systems. This led to *TCP/IP (transport control protocol/internet protocol)*.

TCP/IP was quickly adopted by the academic community in the United States, and soon afterwards by academics worldwide, because it allowed for rapid sharing of scientific information and the electronic mail communication necessary for its rapid discussion and analysis, both within and between university campuses. A worldwide community of inter-linked computers rapidly emerged. Finally, as businesses recognized the potential of the Internet as a large interconnected group of *electronic mail users* and *noticeboard readers*, they began to exploit it for communication and marketing purposes. Though the *Internet* does not have the rigid structure, network management and security controls of other public telecommunication network services, it has a very persuasive appeal; there are already plenty of people to communicate with. This has driven the explosive growth in numbers of registered *Internet* users and *pages* of information.

In its original form, the *Internet* and the *internet protocol (IP)* provided a means for interconnecting computer *servers* (typically *UNIX* computers) together. The internet addressing scheme allowed individual workstations, personal computers or software applications running on either the server or on any of the workstations to direct and send information to other applications on other servers or distant *LANs*. Being a unique

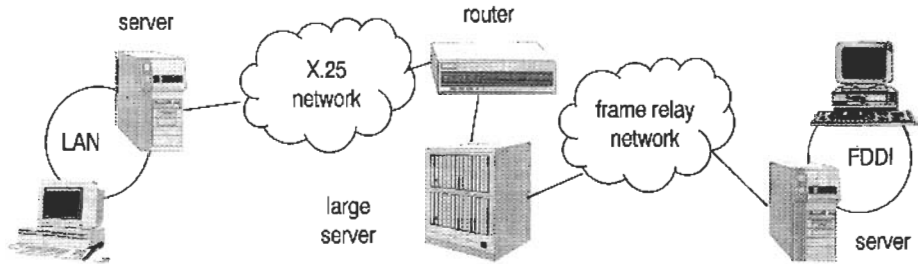


Figure 22.2 Typical Internet communication path

address, the *Internet* address allowed an end user to be identified, no matter how many transit servers, routers or networks would have to be traversed along the way (Figure 22.2). The single *network address* of the destination port in the destination network does not suffice, because the intermediate networks are unable to recognize this address.

A suite of new protocols arrived with *Internet*. Among others, these included *SMTP* (*simple mail transport protocol*), *TFTP* (*trivial file transfer protocol*) and *FTP* (*file transfer protocol*), which together form the basis of the *Internet electronic mail service*.

As the popularity of *Internet* has grown, so have the number of *servers* and *routers* making up the network. New *Internet* service providers have provided for dial-up access to the servers from private individuals using their PCs at home, and new information providers have provided more *Internet pages* of information. The *Netscape browser*, a computer software allowing users to 'surf' the *Internet* and *World Wide Web (WWW)*, by seeking information from any of the connected *servers* by means of a menu-driven screen software which drives a *hypertext* search and *browse* capability.

However, the reason for the rapid growth of the *Internet* also provides a major challenge for the next stage of its development. The factor enabling rapid growth was the possibility to add further servers and routers to the network at almost any point in the network without considering a master plan. The number of connected devices could therefore rapidly increase, as the routers (Chapter 19) are always able to find the desired destination somehow. The problem is the lack of control (and even lack of knowledge) of the path taken. The traffic flows in the network as a whole are therefore largely unmanaged and unmanageable. The only solution to slow response or congestion can be to add more capacity. Whether this capacity is added at the most appropriate point is a matter of chance.

Messages within the *Internet* may go undelivered without trace for any number of reasons, and there is no record of whether electronic copies of information have been retained by any of the intermediate parties. Much attention is currently being focussed by the computer industry on solving the security and network management difficulties of the *Internet*, to stimulate a further surge in demand.

## 22.7 TCP/IP PROTOCOL STACK

Figure 22.3 illustrates the various protocols going to make up the *TCP/IP stack*. At the heart are the *Internet protocol (IP)*, which is equivalent to an OSI layer 3 *network*

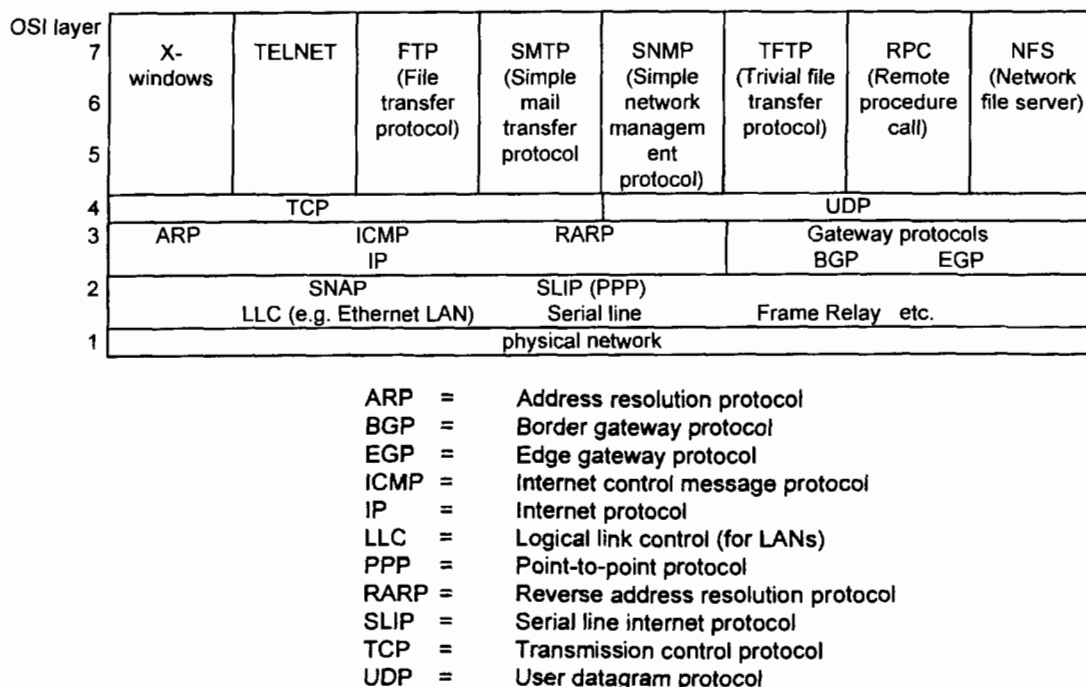


Figure 22.3 TCP/IP protocol stack and associated applications

protocol, and the *transport control protocol (TCP)* an approximate equivalent of OSI layer 4. Also very important for some common multivendor computer-networked applications is the *UDP (user datagram protocol)*.

The *Internet protocol* is a network layer-like protocol, typically running across an ethernet LAN or between LANs via a *router* network (say comprising inter-router connections running on frame relay). The benefit, however, of IP over OSI layer 3 protocols is that *Internet addresses* are widely used in, and are unique across, all computer networks worldwide. This gives the potential for a software application or a user connected anywhere in the world to the *Internet* to access any other computer or software application.

The *transport control protocol (TCP)* is usually used in conjunction with IP and *ICMP (internet control message protocol)* to guarantee *reliable* transmission. On an end-to-end basis, TCP ensures correct sequencing of arriving frames of data, and requests retransmissions when necessary. ICMP is an addition to the basic IP allowing network problems to be reported back to a communicating device. Thus ICMP is able to report the inability to deliver a message and the cause, or the need for *fragmentation*. An alternative to TCP is *UDP (user datagram protocol)*, a simpler protocol which does not perform retransmissions. Instead this job is left to the *application* if necessary.

The *address resolution protocol (ARP)* is used in association with the IP to translate IP addresses into physical hardware addresses. Thus, for example, ARP is capable of determining the appropriate ethernet LAN address (LLC). This address corresponds to

a given IP address. RARP performs the reverse function. When, for example, a diskless computer workstation boots, it obtains its hardware address from the network interface card to which it is attached. It does not, however, know its IP address, this is *resolved* by RARP.

The *gateway protocols* (*BGP*, *border gateway protocol* and *EGP*, *exterior gateway protocol*) are used to connect together different sub-networks of the Internet. They provide for inter-network routing information and communication exchange.

*SNAP* (*sub network access protocol*) is an extended version of the LAN *logical link control* (Chapter 19) which enables the *internet protocol* (*IP*) to be carried over LANs. *SLIP* (*serial line internet protocol*) or its successor, *PPP* (*point-to-point protocol*) are equivalent to OSI layer 2 protocols. They enable transport of IP packets across a simple serial communications line (such as a telephone connection or leaseline). The protocol does not include an address field. The main function is simply to ensure the delineation of packets.

Other well known *routing protocols* of the TCP/IP suite (e.g. *RIP*, *OSPF*) are discussed in Chapter 28.

## 22.8 COMMON APPLICATIONS USING TCP/IP

A number of the most common applications used in association with TCP/IP are also shown in Figure 22.3. These are

- X-windows
- TELNET
- FTP (file transfer protocol)
- SMTP (simple mail transfer protocol)
- SNMP (simple network management protocol)
- TFTP (trivial file transfer protocol)
- RPC (remote procedure call)
- NFS (network file server)

We discuss each of these briefly in turn.

### *X-windows*

This is a windows-based computer operating system intended to allow computer workstation users anywhere within a network of computers supplied by *multiple vendors* to access software applications running on remote hardware platforms.

### *TELNET*

This is a simple TCP/IP-related application that enables a remote log-on from one computer to an application running on another; if you like, a simple version of X-windows. Thus a user of of personal computer using *Windows95* software has the option to

commence a *TELNET* log-on, enabling him to conduct a *session* on a remote *UNIX* computer server. It is the simplest form of TCP/IP connection. As such, it is worth mentioning at this point the widely used *PING* (Packet Internet Groper) technique for confirming the availability of a IP connection. A *PING* is an IP-message that contains the address of its source and the destination. On receipt by the destination terminal a confirmatory *PING* is returned. The procedure confirms not only that the network connection is available but also that the destination device is 'alive'.

### *FTP (file transfer protocol)*

This is a simple but efficient protocol to transfer data files from one computer to another. Although nowadays it is also possible to include data files (e.g. Microsoft word data or Microsoft Powerpoint presentation) as an attachment to an e-mail message, this can be a very slow and cumbersome means of transferring large files. Drawing files from a remote computer can be effected much more quickly using a remote log-on and then the FTP.

### *SMTP (simple mail transfer protocol)*

SMTP has become the basis of the now world-renowned *Internet electronic mail* system (with the addresses that run *president@whitehouse.gov*). It provides for transfer of mail messages between mail servers running on the *Internet*.

### *SNMP (simple network management protocol)*

SNMP has become the *de facto* standard method of transporting network management information around computer and router networks. We discuss it in a little more depth in Chapter 27.

### *TFTP (trivial file transfer protocol)*

This is intended to be a simple protocol, and it is particularly well suited for the downloading of software to and initiation of a remote device.

### *RPC (remote procedure call)*

This provides for a software routine or other application to be called and executed on a remote computer or server.

### *NFS (network file server)*

This application makes files on remote servers and other computers appear to reside on the local user's workstation. It enables him to read and process the data as if it were residing in his own machine.

## **22.9 THE INTERNET PROTOCOL (IP)**

Figure 22.4 illustrates the fields comprised in an IP frame (or *datagram*). The *version* indicates the format of the IP *header*, specifically which version of the protocol is in use. (For example, the 1993 version was version 4). The *internet header length (IHL)*

Version	Internet header length (IHL)	Type of service		Total length
Identification		Flags	Fragment Offset	
Time to Live		Protocol		Header Checksum
Source Address				
Destination Address				
Options				Padding
Data				

**Figure 22.4** Format of an Internet Protocol Datagram

indicates the length of the header in 4 byte (i.e. 32 bit) *words*. The type of service indicates how the *datagram* (as defined in Chapter 18) will be switched and otherwise treated during its transit through the network. The *total length* indicates the total length of the IP datagram, including both header and data.

The *flags* are three bits indicating whether *fragmentation* of the *datagram* is allowed or not. The *fragment offset* indicates the position of the associated datagram *fragment* in the overall message. The *time to live* is the remaining time that the *datagram* is allowed to exist within the *Internet*. When time is up, the datagram is destroyed. This ensures that messages lost in the *Internet* (e.g. datagrams with invalid addresses) are not left around to clog up the network. The *protocol* indication determines whether TCP or UDP protocol is used at the next higher layer, and therefore determines to which protocol agent the datagram should be delivered.

The *header checksum* is akin to a *cyclic redundancy check (CRC)* or *frame check sequence (FCS)* as already discussed in earlier chapters (e.g. Chapter 9). It is only applied to the *IP header* and is recomputed each time the header is amended during passage through the network.

The source and destination *addresses* are both of 32 bits. As we shall see later in Chapter 29, it is normal to write these addresses as four decimal numbers separated by dots, e.g. 123.231.312.123. Each of the four decimal values may only have a value between 0 and 255 (i.e. one of 28 combinations).

A number of options may also be included in the header. Examples are

- security (a mechanism intended to provide for secure communication)
- loose source and record route (the source provides information to aid the routing across the network)
- strict source and record route (the source is able to give information to steer the route taken by the datagram across the network)
- record route (a record of the route taken through the network)
- timestamp (the time and date when the datagram was handled by a router)

The *padding* merely fills up the header to make it an exact number of 32 bit *words*.



Message Type
Code
Checksum
Parameters
Original IP header and data

**Figure 22.5** Internet control message protocol (ICMP) datagram format

## 22.10 THE INTERNET CONTROL MESSAGE PROTOCOL (ICMP)

ICMP can be used in conjunction with IP to provide a more *reliable* means of information transfer. This allows more network status information to be transferred to the end computers to react in the case of non-delivery or corruption of messages. The format of the ICMP amended IP datagram is shown in Figure 5, and the various message types and codes are shown in Tables 22.1 and 22.2.

## 22.11 TRANSMISSION CONTROL PROTOCOL (TCP)

The *transmission control protocol (TCP)* governs the format of the *user data* carried within an *IP datagram*. A *TCP segment* is structured as shown in Figure 22.6.

## 22.12 ONLINE DATABASE SERVICES

Stimulated by the huge population of personal computers now in use in businesses and even in private homes, a new *Online* industry has emerged to sell information on-demand in the form of raw data, news and current affairs, business and trade

**Table 22.1** ICMP message types

Message type	Message
0	Echo reply
3	Destination unreachable
4	Source quench
5	Redirect
8	Echo request
11	Time exceeded for a datagram
12	Parameter problem on a datagram
13	Timestamp request
14	Timestamp reply
17	Address mask request
18	Address mask reply

**Table 22.2** ICMP codes and meanings

Code	Meaning
0	Network unreachable
1	Host unreachable
2	Protocol unreachable
3	Port unreachable
4	Fragmentation needed
5	Source route failed
6	Destination network unknown
7	Destination host unknown
8	Source host isolated
9	Communication prohibited with destination network
10	Communication prohibited with destination host (computer)
11	Network unreachable for service type
12	Host unreachable for service type

	Function
Source port	identifies upper layer application
Destination port	identifies upper layer application
Sequence number	sequence number of this frame within complete message
Acknowledgement number	signals position of data previously acknowledged
Data offset	signals where data field begins
Reserved	(set to zero)
Urgent	indicates whether the urgent pointer is in use
Acknowledgement	indicates whether the acknowledgement number field is in use
Push	
Reset	indicates whether the connection needs to be reset
Synchronizer	used to synchronize sequence numbers
Finished	indicates sender has no more data to send
Window	indicates how much data the receiver can accept before acknowledgement
Checksum	error detection function
Urgent pointer	if used, indicates that urgent data follows
Options	
Padding	ensures header length is integral multiple of 32 bits
Data	user data field - information transferring between applications

**Figure 22.6** Structure of a TCP segment

information, software programmes, graphic images, videos or library information. Huge new global businesses have been created to provide for information and *content* trade. Perhaps the best-known of these companies are *Compuserve*, *Microsoft Network (MSN)*, *America Online* and *Europe Online*.



---

## *The Message Handling System (MHS)*

---

The message handling system (MHS) is a concept developed by ITU-T that is intended to lead to the interconnectivity of all different types of message conveying systems, e.g. telephone, telex, facsimile, electronic mail, etc. MHS sets out a simple model of basic interconnection between systems. As it does so it defines a new dictionary of standard terms and jargon to describe the various phases of a communication. In places it may even seem trivial, but the careful definition is worthwhile if it will solve today's incompatibilities. This chapter seeks to explain the model, unravel the jargon and describe the initiatives that will result.

### **23.1 THE NEED FOR MHS**

It is commonplace today for a wide variety of telecommunications technologies to be provided in the same office. Often a number of technologies provide similar but separate means for conveying identical information, but many users still employ more than one of the methods. Telex, facsimile and electronic mail all provide means of transferring text or documents. One or other of the different devices may be favoured to convey information to a given destination, depending on the facilities available at the far end, for despite the commonality of each of the different methods of document transfer, they are mutually incompatible and will not interwork. MHS defines a simple model of the procedure needed for documents to be passed electronically in a *store-and-retrieve* fashion between a wide range of different types and makes of office business machines (e.g. word processor, facsimile machine, teletype, maybe even the telephone as well). In time, the model will lead on to the development of communication standards, interfaces and equipment needed so that machines of dissimilar types may intercommunicate.

### **23.2 THE CONCEPT OF MHS**

The underlying assumption on which MHS is founded is that during any period of communication only the information content of a message is important. The corollary

is that the means of message conveyance is immaterial, and the format of the information (e.g. facsimile, telex, electronic mail, etc.) may be changed between the originating and destination stations. Thus the originator need not worry about whether a compatible facsimile or telex machine is available at the destination, because MHS will provide conversion to the required format.

ITU-T's recommendation X.400 defines a model of a *message handling (MH)* service, provided by means of a *message handling system (MHS)*. The *MH service* allows the conveyance of messages across a network on a *store-and-forward* or *store-and-retrieve* basis. Thus information between sender and recipient can be sent without interrupting the recipient from his current activities: the message is dealt with when it is convenient. This is important in computing activities if the receiving machine is already busy. It is also important for human recipients who might be away from their desks. Much like a postal system, messages may be *posted* into the system at any time of day and will be delivered to the recipient's *mailbox*. The recipient is free to sort through the incoming *mail* at any time of day, and as soon as an item is accepted by the designated recipient, automatic confirmation of receipt can be despatched back to the sender. This eliminates any possible concern about whether the recipient has seen a given item.

### 23.3 THE MHS MODEL

The MHS model provides a design aid to the development of networks providing message handling services. It is based upon the principles of the International Organization for Standardization (ISO) *open systems interconnection (OSI)* model (as described in Chapter 9). Like OSI, MHS is a layered model with general-purpose application. Unlike the OSI model, only *higher layer* functions are described in the MHS model. In fact, the MHS model is functionally equivalent to the OSI layer 7, or *application layer*.

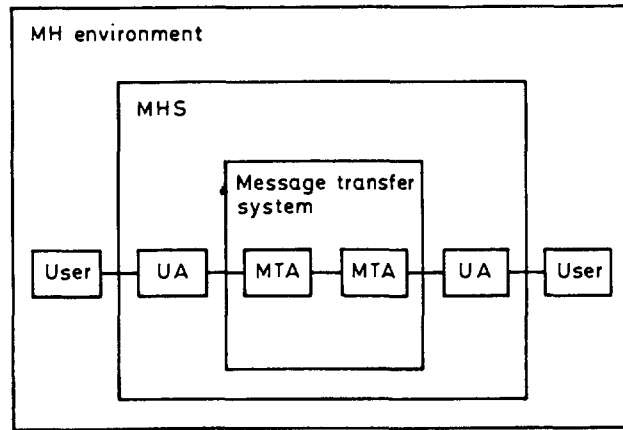
Any realization of a higher layer protocol of the OSI Model relies on implementation also of the underlying layers. The *message handling system* is no exception. The MHS defines the format of messages that may be sent between end points and the procedure for doing so. It does not state any particular physical or electrical means of conveyance. Thus any information which may be represented in binary code may be conveyed on a MHS. Equally, any network capable of carrying the information and conforming to OSI layers 1–6 will do. The network must

- establish the transport service (OSI layers 1–4). Suitable protocols in conjunction with any of the telephone, packet-switched data, LAN or leased line networks will do
- establish a *session* for *reliable transfer* of messages (OSI layer 5)

but MHS uses a null *presentation* (data syntax) layer (OSI layer 6)

Figure 23.1 shows how MHS operates to provide a service to users within a *message handling environment*. All the functions within this environment need to be standardized. For this purpose the model defines a number of terms, as follows.

The two basic functions of MHS are the *user agent (UA)* and the *message transfer agent (MTA)*. These are not usually distinct pieces of equipment in their own right, but a number of these functions may be contained within a single physical piece



**Figure 23.1** Functional model of MHS. MH, message handling; UA, user agent; MTA, message transfer agent. (Courtesy ITU – derived from Figure 1/X.400)

of equipment. The *user agent (UA)* function is undertaken by a personal computer (or similar piece of office equipment) which converts the user's message into a standard form, suitable for transmission. Having done the conversion, the user agent *submits* the message to the local *message transfer agent (MTA)*. The message is conveyed via a number of MTAs. Collectively called the *message transfer system* to the recipient's MTA. This is called the *message transfer function*. Each MTA is usually a mainframe computer or other device capable of storing and forwarding information. Interconnection of the UA to MTA and between the MTAs may use any type of network which provides the functions of OSI layers 1–6.

The two services provided by message handling systems are called the *inter-personal message service (IPMS)* and the *message transfer service (MTS)*. These correspond to the services provided by the *user agent* and the *message transfer agent*, respectively, as Figure 23.2 demonstrates.

Figure 23.2 shows three different implementations of *user agent (UA)*. This underlines the point that the user agent is only a conceptual function, which can be implemented in a number of ways provided that it interacts with the MTA function in the standard fashion. The three different implementations of user agents shown in Figure 23.2 are

- a user agent within a computer or word processor; the IPMS provided by this machine is utilized by a dumb (asynchronous) computer terminal
- an intelligent terminal, within which the user agent function is contained
- an interworking device, interfacing an existing telematic (text based) network (telex or facsimile, etc.) to one of the MTAs

The end-users of a MHS may be either persons or computer applications. Users may be either *originators* or *recipients* of messages. Messages are carried between originator

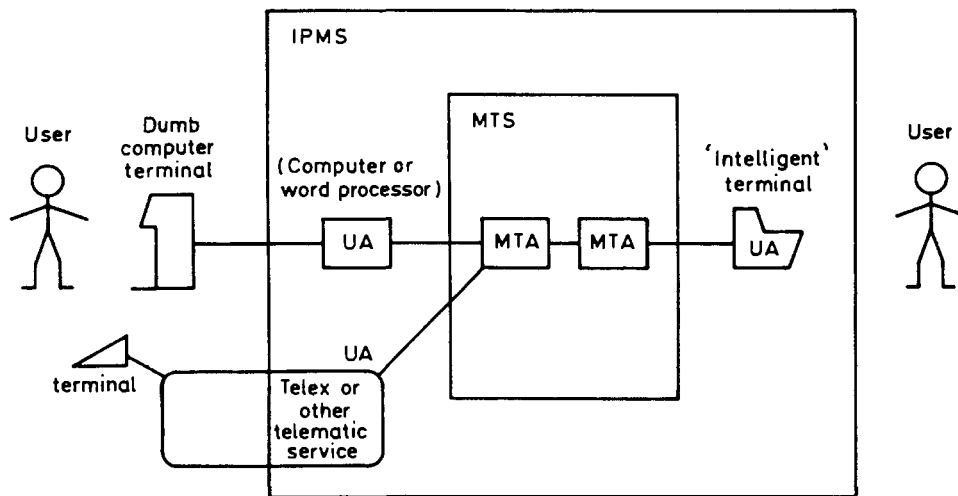


Figure 23.2 Message transfer (MTS) and interpersonal message service (IPMS)

and recipient in a sealed electronic *envelope*. Just as with the postal equivalent, the envelope provides a means of indicating the name of the recipient, and of making sure it reaches the right address with the contents intact. In addition further information may be added to the envelope for special delivery needs. For example, a confidential marking could be added to ensure that only the named addressee may read the information.

The basic structure of all messages in X.400 is thus an *envelope*, with a *content* of information. The *content* is the information sent by an *originating user agent (UA)* to one or more *recipient user agents (UAs)*. This simple message structure is shown in Figure 23.3.

MHS defines three different types of envelope (although all three are very similar in structure). These are called the *submission envelope*, the *relaying envelope* and the

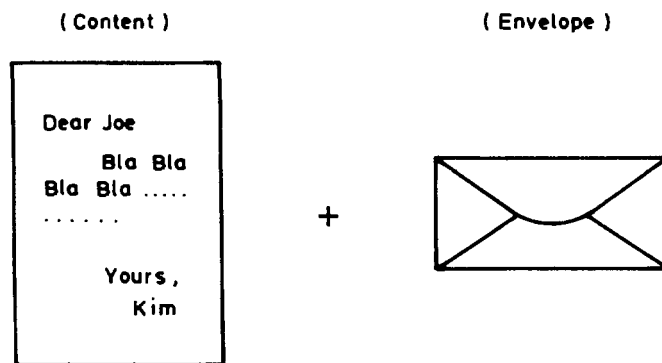


Figure 23.3 Basic message structure in MHS. (Courtesy of ITU – derived from Figure 2/X.400)

*delivery envelope*. The *submission envelope* contains the information prepared by the UA, and is labelled with the information that the MTS will need to convey the message, including the address and any markings such as *confirmed delivery* and *priority*. This envelope is *submitted* by the user agent (UA) to the message transfer agent (MTA). On reaching the MTA, the content of the message is transferred to a *relaying envelope* for relay through the message transfer system to the destination MTA. Finally the message is delivered to the recipient's user agent in a delivery envelope.

In reality, the *envelopes* are simply some extra data, sent in a standard format, as a sort of package around the users real message. All three types of envelope are very similar in structure, though slightly different information is pertinent in each case.

## 23.4 LAYERED REPRESENTATION OF MHS

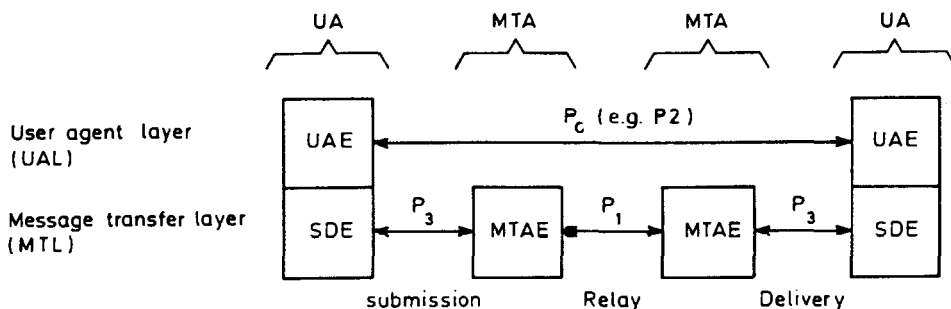
As with the OSI reference model, it is easy to conceive MHS as consisting of functional layers for which

- similar functions are grouped in the same layers
- interactions between layers are minimal
- interactions within the layers are supported by one or more distinct *peer protocols*

The principles of layered protocols were covered in Chapter 9, where the OSI model and *peer protocols* are described. In particular, MHS comprises two functional layers, as shown in Figure 23.4, and as MHS itself resides at layer 7 of the OSI model so these two layers are sub-functional layers of OSI layer 7.

A number of new terminologies are introduced by Figure 23.4 and need explanation. The two layers of MHS are the *user agent layer (UAL)* and the *message transfer layer (MTL)*. These are the layers in which the *user agents* and *message transfer agents* operate.

The user agent layer comprises *user agent entities (UAEs)* and an interacting protocol, called  $P_c$ . The protocol  $P_c$  governs the content, format, syntax and semantics



**Figure 23.4** Layered representation of MHS. (Courtesy of ITU – derived from Figures 12 and 13/X.400)



to be used when transferring information between the two user agents. ( $P_c$  stands for *content protocol*). A particular type of content protocol is called  $P_2$ , and is also known as the *interpersonal messaging protocol*. It is covered by ITU-T's recommendation X.420. The *user agent entity* is the capability, within the user agent which is required to convert the base information into the form demanded by  $P_c$  (or  $P_2$ ).

The *message transfer layer (MTL)* comprises *submission and delivery entities (SDEs)*, *message transfer agent entities (MTAEs)* and interacting protocols called *message transfer protocol* (or  $P_1$ ), and the *submission and delivery protocol* (or  $P_3$ ).

The *submission and delivery entity (SDE)* is the part of the user agent that places the information (or *content*) provided by the *user agent entity (UAE)* into the *submission envelope*. The envelope is then addressed using the protocol  $P_3$ . The *message transfer agent entity (MTAE)* is the functional part of the *message transfer agent (MTA)* which relays the message in the *relaying envelope*. The *message transfer protocol*,  $P_1$ , is used to perform the message relay. Finally the message is delivered to the destination SDE in the *delivery envelope*, using protocol  $P_3$ . The full messaging sequence is thus as follows.

1. The *user agent entity (UAE)*, which is a functional part of the *user agent (UA)*, assists the user to compose the content of the message in  $P_2$  (or some other  $P_c$ ) format.
2. The *submission and delivery entity (SDE)*, also a functional part of the *user agent*, envelopes and addresses the message, and then *submits* it to the *message transfer agent entity (MTAE)* using  $P_3$  protocol.
3. The *message transfer agent entity (MTAE)*, which is a functional part of the *message transfer agent*, re-envelopes the message in a  $P_1$  *relaying envelope*, and relays it in a store and forward fashion via the MTAEs of other MTAs, until reaching the destination MTA. For the relay, protocol  $P_1$  is used.
4. The destination MTAE returns the message to its  $P_3$  envelope and delivers it to the SDE of the destination UA, which removes the envelope.
5. The UAE part of the destination UA reformats the information to the original style and conveys it to the user.

As an aside, it is interesting to note the difference between the 1984 and 1988 versions of the  $P_3$  protocol, even though this is now largely history, because it exemplifies how initial versions of standards may not be completely reliable, even though published. Whereas CCITT's 1984 version contained a version of  $P_3$  protocol which was a *store-and-forward* protocol, this had the weakness that it had no way of preventing an MTA from attempting to forward a message to a UA, such as a personal computer, which might have been switched off at the time. In that case the message would be lost. By 1988, the CCITT recommendation for the  $P_3$  protocol had been advanced to  $P_{3+}$  as a *store-and-retrieve* protocol. In the later version, the message is retained by the destination MTA until the recipient user agent chooses to retrieve it.

## 23.5 THE STRUCTURE OF MHS MESSAGES AND MHS ADDRESSES

Let us now consider the structure of MHS messages and the structure of messages which are used to support the inter-personal message service (IPMS). Figure 23.5 shows the imaginary nested structure of a message, sent in reality as a long stream of binary coded data.

This structure is apparent also in Figure 23.5, although some other parts of the message are also shown. The content may be either user information, such as text, facsimile diagrams, etc., formatted in the P<sub>2</sub>-style as *heading* and *body*, or it can take the form of a *status report*. A *status report* is a special type of message required for the correct functioning of UAEs. A *status report* may be used, for instance, to convey *receipt confirmation messages*. The *heading* of the message characterizes the user information with elements such as 'subject', 'to', and 'copy'. The body is the text and diagrams. The content of the message must be formatted using a content protocol, P<sub>C</sub>. One type of content protocol already discussed is the P<sub>2</sub> *inter personal messaging protocol*. If required, however, a different type of content protocol can be specially developed for a given application.

The envelope of a message carries the address of the recipient. ITU-T recommendation X.400 suggests that the address should be one of the following:

- a unique identifier of entirely numeric characters; this is called a *primitive name*, and indicates the destination user agent in terms of its country, its *management domain* (explained below), and a unique customer number
- a unique name-style identifier, correctly called a *descriptive name*, comprising a personal name, an organization name and the name of an organizational unit

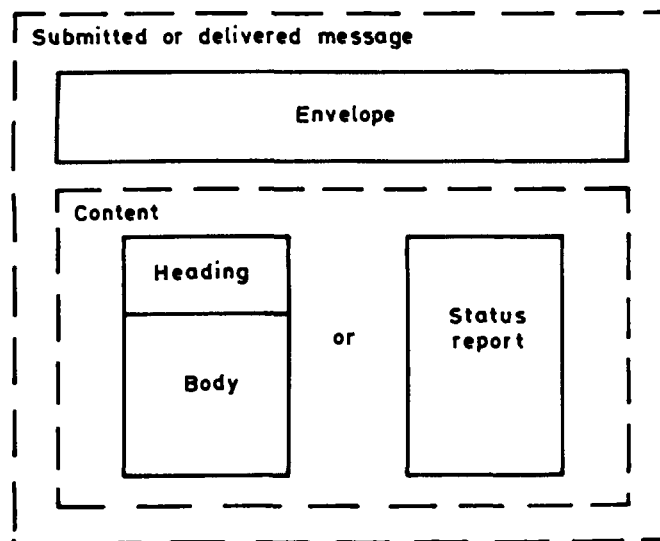


Figure 23.5 Message structure for IPMS. (Courtesy of ITU – derived from Figure 2/X.420)

All addresses in MHS must be unique, and denote exactly one user. An example of a primitive name might be a number such as 12345926, etc. By contrast an example of a descriptive name is

‘The chairman of the XYZ company’

The scope for using descriptive names is helpful to human users, who no longer have to remember arbitrary numbers for calling their friends or close associates. However, the name must be sufficient for the message handling system to recognize the user uniquely. A name such as

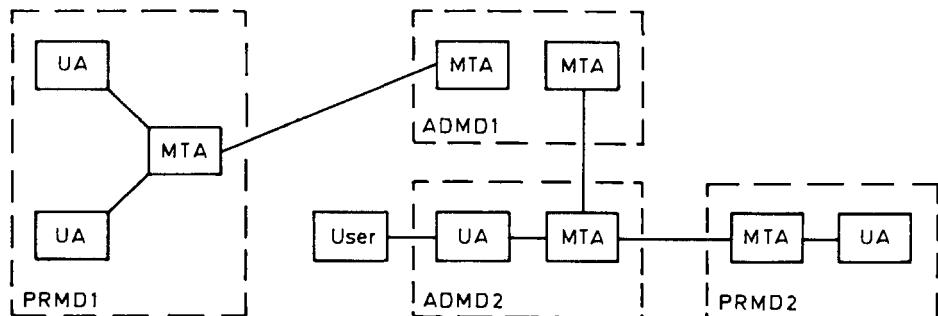
‘the floor cleaner of the XYZ company’

will not do as a descriptive name, unless the XYZ company only has one floor cleaner.

## 23.6 MHS MANAGEMENT DOMAINS

A *management domain* might be a complete MHS network, or it may comprise a small number of MTAs owned and run by a single organization. A *management domain* might be a large company’s private MHS network, or a network run by the public telephone company. The two types of domain are called, respectively, a *private management domain (PRMD)* and an *administration management domain (ADMD)*. Figure 23.6 illustrates the management domains concept, and shows the interconnection of a number of different domains. The use of management domains allows the addresses in each domain to be allocated independently of addresses in other domains. Thus the same address ‘The chairman’ may be used to apply to two different people, being the individual users of separate companies’ MHS networks.

Figure 23.6 reflects the different examples encountered in real life, where some public telecommunications operators may provide only message transfer (i.e. network) facilities while others provide customer premises equipment as well. Interconnection between all the management domains is by standard  $P_1$  (message transfer) protocol.



**Figure 23.6** Administration and private management domains

## 23.7 MHS AND THE OSI DIRECTORY SERVICE

So that the MHS may deliver messages to the correct recipients, it needs to

- validate the address names of originators and recipients (*originator/recipient, or O/R name*)
- determine the necessary call routing, and finally
- determine whether data format, or code conversion is required

A large quantity of reference information, giving details and addresses of all the users, is required so that, first, the sender can address the message correctly to the intended recipient; and, second, so that the UA and MTA can determine where this person is and so deal with the message accordingly. The information needed is drawn from the *OSI directory service* which is prescribed by ITU-T recommendations in the X.500 series. These recommendations lay out a standard structure in which information may be held in a database, and they provide for a standard method of enquiry (see Chapter 29). Besides determining the correct routing paths for MHS messages, the service can also be used to answer users' directory enquiries. A widely held view is that large scale corporate and international interconnection of electronic mail and other messaging devices cannot take place using the X.400 standard until the X.500 directory service has been fully defined and widely realized.

## 23.8 MESSAGE CONVERSION AND CONVEYANCE USING MHS

MHS is intended to recognize all of the standard binary codes (e.g. telex, facsimile, etc.) and provide for translation between them by interaction with the OSI presentation layer (layer 6). In the case of a telex machine using the message conversion capabilities of a message handling system to send a message to a facsimile machine, MHS would convert from telex characters into a facsimile message format. In time, some people hope that even speech to text conversion may be possible. So far, however, not all translation functions are understood and further technical development will be required before we have a system which provides for all types of translation.

The standard codes understood by MHS include

- text in IA5 computer code (international alphabet no. 5)
- telex (international alphabet no 2, or IA2)
- facsimile
- digitalized voice
- word processor formats

There is no restriction on the number of different types of code which may be used in a single message. Thus messages of mixed text, facsimile diagrams and even voice can

be sent, although there may of course be a limitation on what the receiving device will comprehend.

Early application of MHS by companies and other users is to convey documents in electronic form. A prime user of the standard X.400 protocols will be to interconnect today's incompatible electronic mail systems. In addition MHS has provided the foundation for a new era of paperless trading between companies, called *EDI* (*electronic data interchange*). In this environment it is invaluable for large scale distribution of documents, for rapid confirmation of receipt, and for delivery of messages in electronic form, allowing subsequent processing if necessary.

### 23.9 SETTING UP A MESSAGE HANDLING SYSTEM

Figure 23.7 illustrates a practical implementation of a message handling system, showing the items of equipment that the user may be familiar with.

Figure 23.7 shows a network comprising a personal computer, a mainframe computer and another mainframe computer. The *message handling system* is only a small function, but one which each of the computers must be capable of performing. The MHS itself is a small amount of software in each computer, and so invisible to the user.

The personal computer on the left of Figure 23.7 has within it communication software which performs a user agent function, as well as the functions of OSI layers 1–6. The physical network connecting it to the message transfer agent (MTA), which resides in the mainframe computer in the centre of the figure, is a packet-switched network. In the jargon of MHS the personal computer is said to be an *S1-system*, because it comprises only a *user agent entity* (UAE) and a *submission and delivery entity* (SDE) but no *message transfer agent entity*.

In contrast, the mainframe computer shown in the centre of Figure 23.7 is the main 'databank' of the MHS system. It comprises only a *message transfer agent entity* (MTAE) together with communication software supporting the functions of OSI layers 1–6. The MTAE performs the function of relaying message between user agents (UAs), and when necessary it stores messages for later retrieval by recipient UAs. As this mainframe computer contains a MTAE, it is called, in MHS jargon, an *S2-system*. The computer

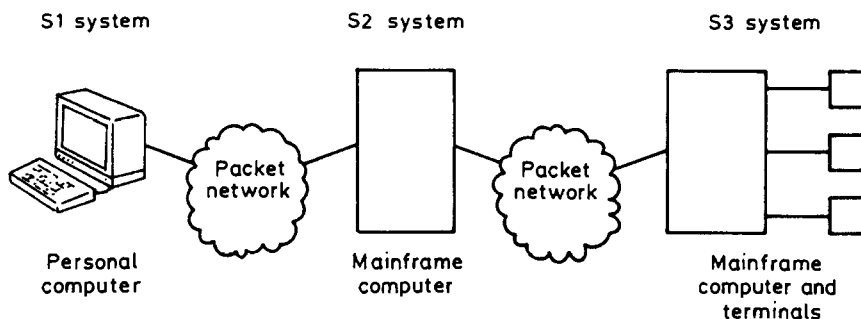


Figure 23.7 Physical example of MHS realization

may perform other normal data processing functions in addition to its MHS communication functions, although many companies may choose to devote a mainframe computer for use as a message handling system (S2-system). Finally, the computer shown on the right-hand side of Figure 23.7, comprises both a user agent entity (UAE) and a message transfer agent entity (MTAE). This is called an *S3-system*. Mainframe computers dedicated to use as MHS S3-systems are also likely to become common.

To sum up, we have distinguished three types of systems as realizations of MHS

- systems containing only *user agent (UA)* functions (i.e. UAE and SDE); these are termed *S1-systems*
- systems containing only *message transfer agent (MTA)* functions (i.e. MTAE), termed *S2-systems*
- systems containing both UA and MTA functions (i.e. UAE and MTAE), termed *S3-systems*

The diagram in Figure 23.8 shows the functional structure of the equipment of Figure 23.7. It illustrates the *stack* of OSI protocols and functions required within the actual equipment, to realize MHS in practice. The diagram shows how the *protocols* at each of the OSI layers are communicating between S1, S2 and S3 systems on a *peer* level. A packet network based MHS is shown, using the ITU-T X.25 packet protocol at layers 1–3. Meanwhile layer 4 and 5 protocols being used are those defined in ITU-T recommendations X.224 and X.225. The *presentation* layer has a null protocol, and then  $P_1$ ,  $P_3$  and  $P_c$  are used at layer 7. Layers 1–3 might equally have been based on a LAN (local area network), or a telephone network, and in these cases the X.25 packet protocols would have been unsuitable.

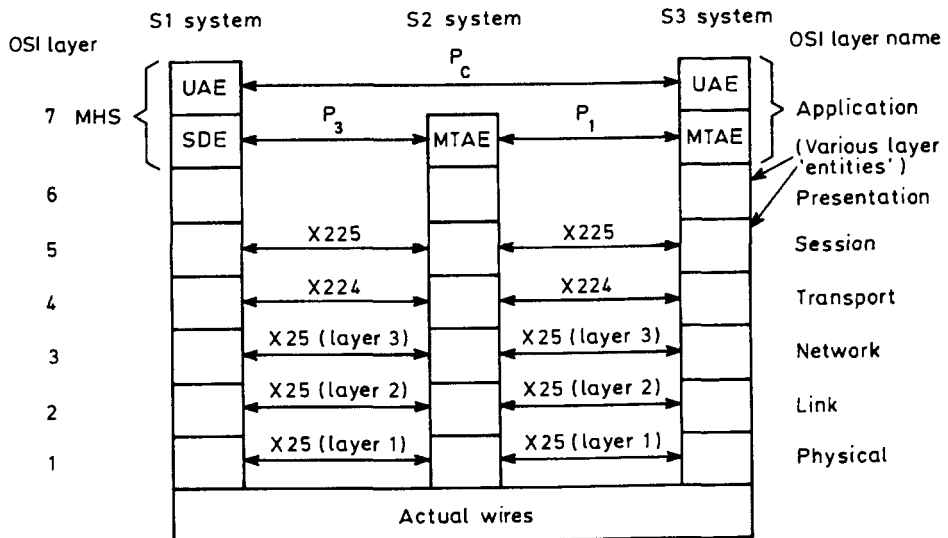


Figure 23.8 Packet based realization of MHS

## 23.10 FILE TRANSFER ACCESS AND MANAGEMENT (FTAM)

We learned in Chapter 22 how it can be useful to be able to manipulate and access files on remote computers, and be able to transfer them quickly and efficiently across computer networks. These functions are carried out in the *Internet* by protocols such as *NFS* (*network file server*) and *FTP* (*file transfer protocol*). The equivalent and more powerful (though perhaps also more cumbersome) protocol in the open systems world is called *FTAM* (*file transfer, access and management* or *file transfer and access method*).

## 23.11 SUMMARY

The Message Handling System X400 will provide for users an extremely flexible way of interconnecting different types of office equipment. It operates in a store-and-retrieve fashion and will be an important tool in computer communication, particularly for

- passing electronic mail and documents between proprietary computer systems
- data file transfer between computers separated by time zones and geography – using the public packet network

The usefulness of the X.400 service will depend on the availability of directory information so that messages can be delivered to the correct recipient. This will govern the rate at which today's different electronic mail networks may be linked together. In essence this means that many potential users must wait for full development of the new X.500 directory service.

Ultimately, X.400 will provide a most powerful means of interconnection, allowing paperless interaction not only between different equipment but also between different companies. It heralds the first real opportunity for paperless trade based on *electronic data interchange (EDI)*, and it may eventually lead to a genuinely 'paperless office'.

---

## *Mobile and Radio Data Networks*

---

Just as computers and datacommunication are revolutionizing office life, so mobile and radio data networks are enabling corporate computer networks to be extended to every part of the company's business, including the mobile sales force, the haulage fleet and the travelling executive. Data network techniques can now also be used to trace and pinpoint trucks on the road or ships at sea. This chapter discusses some of the most recent radio data network technologies, covering the principles of 'radiopaging', mobile data networking, 'wireless LANS', as well as describing radiodetermination services.

### 24.1 RADIOPAGING

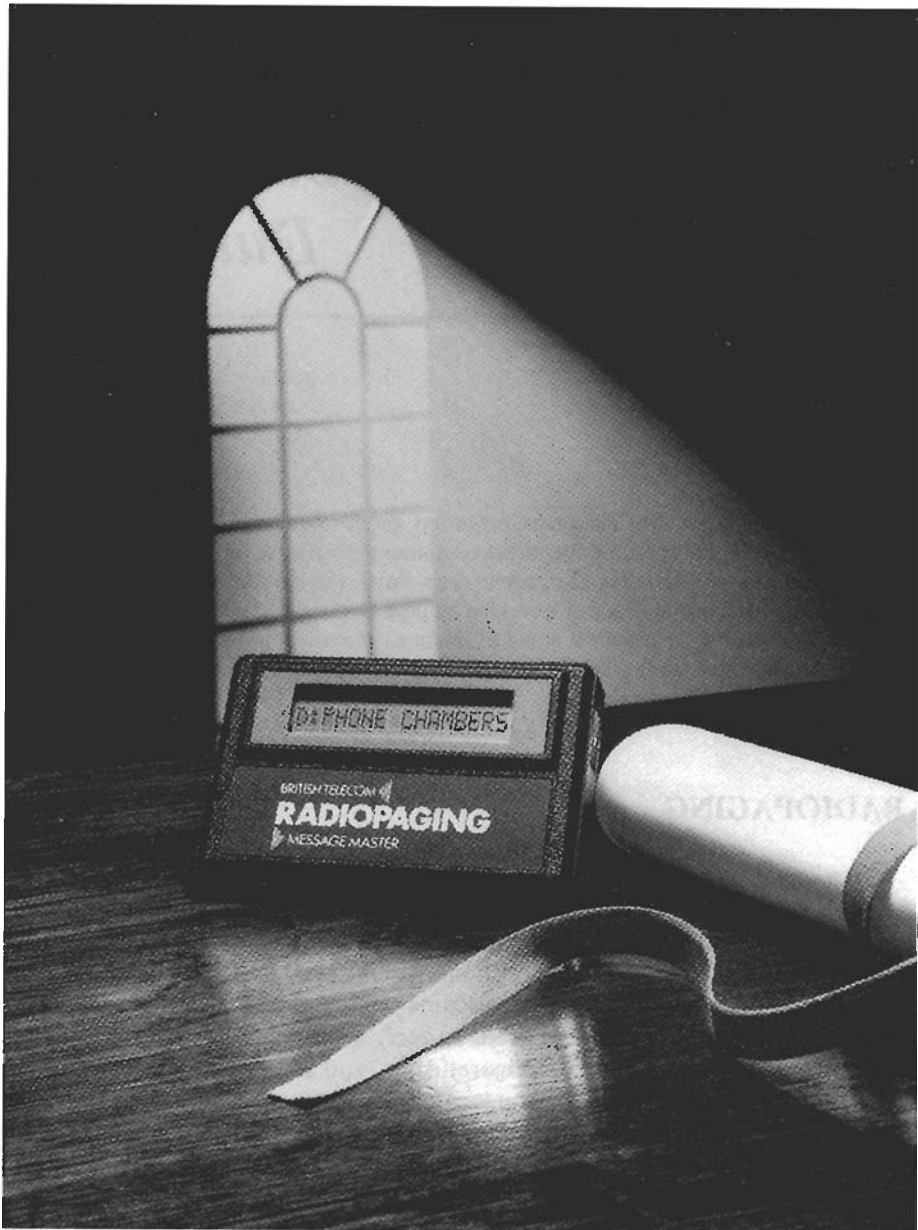
*Radiopaging* was the first major type of network that enables transfer of short data messages to mobile recipients. Initially it was a method of alerting an individual in a remote or unknown location (typically by 'bleeping' him) to the fact that someone wishes to converse with him by phone. Subsequently, the possibility to send a short text message to the mobile recipient became commonplace.

To be paged an individual needs to carry a special radio receiver, called a *radiopager*. The unit is about the size of a cigarette box, and is designed to be worn on a belt or clipped inside a pocket. The person carrying the pager may roam freely and can be *paged* provided they are within the radiopaging *service area*. The service may provide a full nationwide coverage. Figure 24.1 illustrates a typical radiopaging receiver.

The initial radiopagers were allocated a normal telephone number as if they were standard telephones. Paging was achieved by dialling this number, as if making a normal telephone call. Instead of being connected through to the radiopager the caller either speaks to a radiopaging service operator, or hears a recorded message confirming that the radiopager has been *paged*. *Paging* is done by sending a radio signal to the radiopager, causing it to emit an audible 'bleeping signal' to alert its wearer.

The simplest types of radiopager, even today, provide no further information to the wearer than the bleep. Having been alerted, the wearer must find a nearby telephone and





**Figure 24.1** Message display radiopager. A relatively sophisticated radiopager, allowing not only bleeping facilities, but also the conveyance of a short textual message. (Courtesy of British Telecom)

ring a pre-arranged telephone number (say the radiopaging operator, or the wearer's own secretary) to be given the message or the telephone number of the caller who wishes to speak with him. Thus for a caller to *page* an individual they must first inform the intermediary office (the radiopaging operator or the 'roaming individual's' secretary). The caller leaves either a message for the paged person, or a telephone number to be called. Figure 24.2 gives a general schematic view of a radiopaging network.

The key elements of a radiopaging system are the *paging access control equipment* (PACE), the paging *transmitter* and the paging *receiver*. The PACE, contains the electronics necessary for the overall control of the radiopaging network. It is the PACE which codes up the necessary signal to alert only the appropriate receiver. This signal is distributed to all the radio transmitters serving the whole of the geographic area covered by the radio-paging service. On receiving its individual alerting signal the receiver bleeps.

A special code is used between the PACE and all the paging receivers. It enables each receiver to be distinguished and alerted. The earliest codes used a discrete signal tone, modulated onto a radio frequency to identify each receiver. Such systems, available in the late 1950s, could address a small number of receivers. Two-tone systems rapidly followed in the 1960s, as the popularity of radiopaging grew. In the two-tone system, up to around 70 tones are used, any two of which are sent in consecutive short bursts, allowing up to  $70 \times 70 = 4900$  receivers to be alerted individually. Two-tone systems were used for on-site paging applications, such as summoning medical staff in a large hospital.

Two-tone systems were too small in capacity to be considered for use over a wide area covering large cities or a nation. Hence followed the development of systems using more bursts of tone. A number of proprietary five-tone systems were developed typically

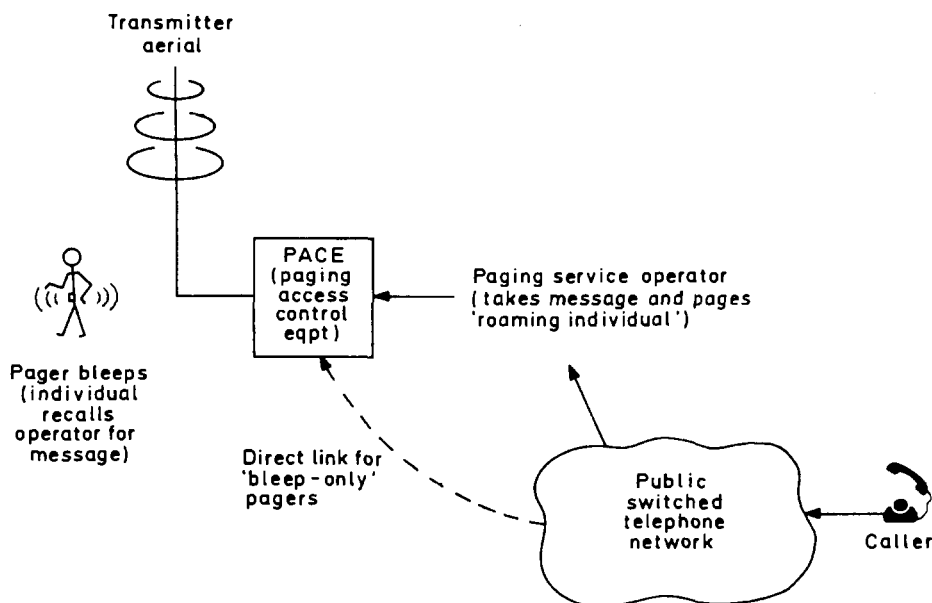


Figure 24.2 Radio paging a 'roaming individual'

using a repertoire of ten different tone frequencies, and allowing up to 25 bursts of tone per second. This amounts to a system capacity of  $10 \times 10 \times 10 \times 10 \times 10 = 100\,000$  receivers, and a *calling or paging rate* of 25/5 calls per second. However, even five-tone systems were unable to cope with the explosion in demand that many of the radiopaging operators saw in the late 1970s, and new digital coding systems for paging became necessary.

The digital codes were not as sensitive as their predecessors, but they had enhanced performance capabilities in terms of overall calling rate and capacity. Furthermore, they offered the scope for short alphanumeric messages to be paged to the receiver and promised lower overall unit costs, both of the PACE and of the individual receivers. A number of digital codes were developed in the late 1970s and early 1980s, among them the Swedish MBS code (1978), the American GSC code (1973), and the Japanese NTT code (1978). The most important code, now common throughout the world, is that stimulated by the British Post Office. Known as the POCSAG code, after the advisory group that developed it (the *Post Office code standardisation advisory group*), it was developed over the period 1975–1981 and was accepted by the CCIR (*Consultative Committee for International Radio*, the forerunner to ITU-R) as the first international radiopaging standard. It has a capacity of 2 million pagers (per zone) and a paging rate of up to 15 calls per second. Furthermore, it has the capability for transmitting short alphanumeric messages to the paging receiver. It works by transmitting a constant digital bit pattern of 512 bits per second. The bit pattern is segregated into *batches*, with each batch sub-divided into eight *frames*. A particular pager will be identified by a 21-bit *radio identity code*, transmitted within one (and always the same one) of the eight *frames*. It is this code, when recognized by the paging receiver, that results in the alerting bleep.

An extra feature of the POCSAG code is that an extra two bit code can be used to provide four different bleeping cadences in each pager. These can be assigned with different telephone numbers for paging, and they correspond to four different recall telephone numbers. This may be useful for a user who for most of the time is contacted by a small number of different people, because it potentially removes the need for the intermediary. Figure 24.3 shows how each caller uses a different telephone number to page the roaming individual, and produce a distinctive bleeping cadence.

Caller no.	Caller calls telephone number	Roaming individual hears bleep cadence	Recall telephone number
1	A	. . . . .	6 2 1 1 1 (corresponds to caller 1)
2	B	. — . — . — .	7 2 3 7 2
3	C	— — — —	8 4 9 2 3
4	D	. . — . . —	5 3 2 2 4

Figure 24.3 Different paging cadence identifies appropriate recall number

Text messages (consisting of alphanumeric characters) were initially conveyed by the radiopaging operator. It is nowadays sometimes also possible to input the message using videotext or a similar data network service. The most advanced receivers, when used in a suitably equipped radiopaging network, are capable of messages up to 80 characters long.

The pager itself is a small, cheap and reliable device. Most are battery-operated, but if the pager were to be on all of the time, the battery life would be very short, so a technique of battery conservation has become standard. We have already described how the *radio identity code* is always transmitted in the same frame of an eight frame batch to a particular receiver. This means that receivers need 'look' only for their own identity code in one particular frame, and can be 'switched off' for seven-eighths of the time. This prolongs battery life.

Paging receivers include a small wire loop aerial, and because of the low battery power can only detect strong radio signals. This fact needs to be taken into account by the radiopaging system operator when establishing transmitter locations and determining transmitted power requirements, and by the user when expecting important calls. The *radio fade* near large buildings can be a major contributor to the low probability of paging success.

The *paging access control equipment (PACE)* stores the database of information to determine which zones the customer has paid for, and to convert the telephone numbers dialled by callers into the code necessary to alert the pagers, and in addition it performs the coding of textual alphanumeric messages. The PACE also has the ability to queue up calls if the incoming calling rate is greater than that possible for alerting receivers over the radio link. Furthermore, the PACE prepares records of customer usage, for later billing and overall network monitoring.

The most advanced modern paging radiopaging systems are *satellite paging systems*. These work in exactly the same way as terrestrial radiopaging systems, except that the transmitted signal is relayed via a satellite to achieve a global coverage area. This enables the roaming individual to receive his messages wherever he is in the world.

## 24.2 MOBILE DATA NETWORKS

Mobile radio is an awkward medium for carrying data. Interference, fading, screening by obstacles, and the hand-off procedure between cells all conspire to increase errors, so although the digital fixed telephone network may expect to achieve error rates no greater than 1 in  $10^5$  bits, the error rate over mobile radio can be as high as 1 in 50 bits.

Very basic systems with slow transmission speeds (say 300 bit/s) have been used. At these rates few data are lost and connections that are lost can be re-established manually. However, for more ambitious applications error-correcting procedures must be used, normally a technique employing *forward error correction (FEC)* and automatic *re-request retransmission*. In this technique sufficient redundant information is sent for data errors to be detected and the original data reconstructed even if individual bits are corrupted during transmission. Typical speeds achieved are 2.4–4.8 kbit/s.

The appearance of mobile data networks was largely stimulated by the taxi industry. *Press to speak* private mobile radio systems first appeared in taxis as a means for controlling taxi fleet movements. A taxi customer calls a telephone number, where a

number of operators act to accept orders and *despatch* available taxis to pick clients up. The *despatching* process occurs by radio. After each 'drop-off' a taxi driver registers his position and receives instructions about where he can 'pick-up' his next client.

By the mid-1980s the *press to speak* despatch systems had become unable to cope with the size of some of the large metropolitan taxi despatch consortia. It was becoming difficult to be able reliably to contact all the drivers, and wearing on the drivers always to have to listen out for calls. Computer despatch systems were being introduced for the automation of taxi route planning, and the natural extension was direct computer readout to the individual drivers of their planned activities. By computer automation it became possible to ensure despatch of a client order to a particular taxi driver, who could be automatically prompted to acknowledge its receipt and his acceptance of the order. Simple confirmation by the driver ensures precise computer tracking of pick-up time and a successfully completed fare. Subsequent computer analysis of journey time statistics could further help future journey planning.

Now there was a need for data networking via radio. Most of the systems developed to answer this need evolved from the previous press-to-speak *private trunk mobile radio (PTMR)* systems used in the taxi and regional haulage business beforehand. As a result they tend to use a similar radio frequency range for operation, and a similar 12.5 kHz or 25 kHz channel spacing. The derived user data bitrates achievable are typically around 7200 bit/s per connection, but once the overheads necessary to ensure the reliable and bit error free transport of the user data are removed, the effective data rate of some systems does not exceed 2400 bit/s. Miserable, you might think, when compared to fixed network data applications running at 64 kbit/s or even higher rates, but quite adequate for the short packet (i.e. around 2000 byte packet messages (approximately 2000 characters)) for which the systems were developed.

The three best known manufacturers of low speed mobile data networks are Motorola (its *Modacom* system), Ericsson (Eritel's *Mobitex* system) and *ARDIS*. The systems find their main application in private network applications within metropolitan or regional operations (for haulage or taxi companies) or on campus sites, essentially providing radio-based X.25 packet networks, as Figure 24.4 shows. There have been a

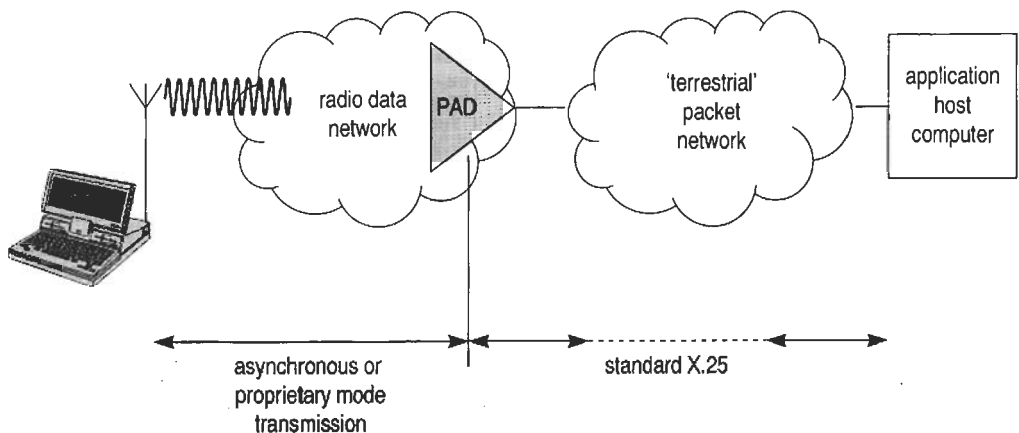


Figure 24.4 Typical arrangement of a mobile data network

number of attempts at providing commercial nationwide and even international public service networks, but these have not been a great success.

### 24.3 TETRA (TRANS-EUROPEAN TRUNKED RADIO SYSTEM)

Despite the relatively low interest in low speed mobile data networks, and the emergence of the GSM and DECT systems (Chapters 15 and 16) as overpowering competitors both for voice service via *trunk mobile radio* and data carriage via *Modacom*-like low speed mobile data networks, there has been continued effort applied by ETSI to agree the *TETRA (trans-European trunk radio)* series of standards. These are intended to provide for harmonization of trunk mobile radio networks across Europe, opening the way for pan-European services and the use of identical equipment.

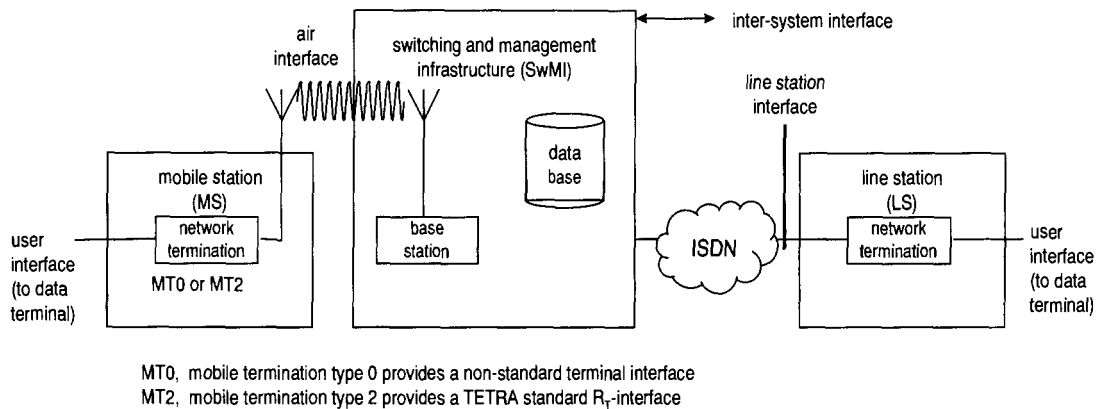
Work on the TETRA standards started in ETSI in 1988, when a system to be called *mobile digital trunk radio system (MDTRS)* was foreseen. This was renamed TETRA in 1991. A series of standards have now been published, which can be classified into two different broad system categories

- TETRA V+D is a system for integrated *voice and data*
- TETRA PDO is a system for *packet data only*

The first of these systems is intended as an ISDN-like replacement for analogue trunk mobile radio systems (Chapter 15). The second system is a standardized version of the *Modacom*-like systems, but with higher data throughput capabilities. Table 24.1 lists the bearer and teleservices planned to be made available.

**Table 24.1** Bearer and teleservices supported by the various TETRA standards

	TETRA V + D (voice and data)	TETRA PDO (packet data only)
Bearer Services	7.2–28.8 kbit/s circuit-switched voice or data (without error control) 4.8–19.2 kbit/s circuit-switched voice or data (some error control) 2.4–9.6 kbit/s circuit-switched voice or data (strong error control) connection-oriented (CONS) point-to-point packet data (X.25) connectionless (CLNS) point-to-point packet data (X.25) connectionless (CLNS) point-to-point or broadcast packet data in non-X.25-standard format	connection-oriented (CONS) point-to-point packet data (X.25) connectionless (CLNS) point-to-point packet data (X.25) connectionless (CLNS) point-to-point or broadcast packet data in non-X.25-standard format
Teleservices	4.8 kbit/s speech encrypted speech	



**Figure 24.5** Basic architecture of the TETRA system

Figure 24.5 illustrates the basic architecture of the TETRA system. The concept foresees a normal connection between a *line station (LS)* and a *mobile station (MS)* via a *base station* and *switching and management infrastructure (SwMI)*. Thus a typical example would be a taxi computer despatch centre as a line station connected to the fixed ISDN network, accessing one or more (typically many) mobile stations. Similar to the DECT system, the data base is conceived to take over *home data base* and *visitor data base* functions, to allow roaming of mobile stations between different base stations and even between different TETRA networks. The *inter-system interface (ISI)* allows for interconnection of TETRA networks operated by separate entities. The various *c-plane* and *u-plane* *air interfaces (AI)* are designed to conform with OSI.

Table 24.2 presents a brief technical overview of the TETRA system.

## 24.4 WIRELESS LANS

The idea of *wireless LANs (WLANs)* has been around for as long as LANs themselves. Indeed the first LAN, developed by the Xerox company based on the ALOHA-protocol, which became the basis of ethernet, was based on a radio medium.

There are two main benefits wireless LANs when compared with cable-based LANs

- ability to support mobile data terminals (for example, employees using laptop computers at various different desk locations within a given office building)
- ability to connect new devices without the need to lay more cabling

Two standards for wireless LANs have been developed. These are the IEEE 802.11 standard and the ETSI *HIPERLAN (high performance LAN)* standard. We describe here the ETSI HIPERLAN system.

**Table 24.2** Technical overview of the TETRA system

Radio Bands	Uplink: 380–390 MHz 410–420 MHz 450–460 MHz 870–888 MHz	Downlink: 390–400 MHz 420–430 MHz 460–470 MHz 915–933 MHz
Channel separation	25 kHz	
Channel multiplexing	V + D: TDMA (time division multiple access), with S-ALOHA on the random access channel PDO: S-ALOHA with data sense multiple access (DSMA)	
Duplex modulation	FDD (frequency division duplex), 10 MHz spacing	
Frame structure	V + D: 14.17 ms/slot, 510 bits per slot, 4 slots per frame PDO: 124 bit block length with forward error correction (FEC). Continuous downlink transmission, burst uplink ALOHA	
Modulation	$\pi/4$ DQPSK (differential quaternary phase shift keying)	
Connection set-up time	circuit switched connection, less than 300 ms connection-oriented data, less than 2 s	
Propagation delay	V + D: less than 500 ms for connection-oriented services 3–10 seconds for connectionless services PDO: less than 100 ms for 128 byte packet	

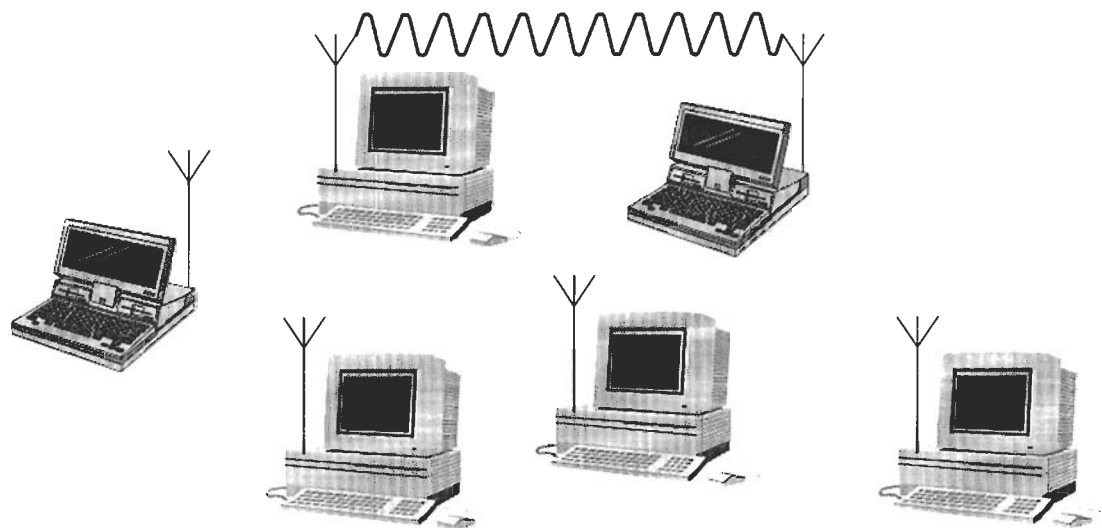
In a wireless LAN each of the devices to be connected to the LAN is equipped with a radio transmitter and receiver suited to operate at one of the defined system radio channel frequencies. For the *HIPERLAN* system, five different channels are available, either in the band 5.15–5.30 GHz or in the band 17.1–17.3 GHz, but only one of the channels is used in a single LAN at a time. The radio channel has a total bitrate close to 24 Mbit/s but the maximum user data throughput rate is around 10–20 Mbit/s, i.e. of similar capacity to a cable-based ethernet or token ring LAN.

When a device wishes to send information, this is transmitted in a manner similar to that used in an ethernet LAN. In other words, the information is simply transmitted to all other terminals in the LAN, as soon as the radio channel is available. All devices participating in the LAN ‘listen’ to the radio channel at all times, but only ‘pick up’ and decode data relevant to themselves. The structure of the LAN is therefore very simple, as Figure 24.6 illustrates, but all devices must lie within about 50 metres of one another, because of the 1 Watt maximum radio transmit power allowed.

The multiple radio frequencies (five per band) defined in the *HIPERLAN* standard allow multiple LANs to exist beside one another and even overlapping one another. Without multiple frequencies different LANs in adjacent offices might not be possible, and multiple LANs in the same office certainly not.

The 50 metre maximum diameter of the LAN could also be a major constraint in some circumstances. For this reason, the radio MAC (medium access control) provides a forwarding (or relay) function. When the forwarding function is configured into the

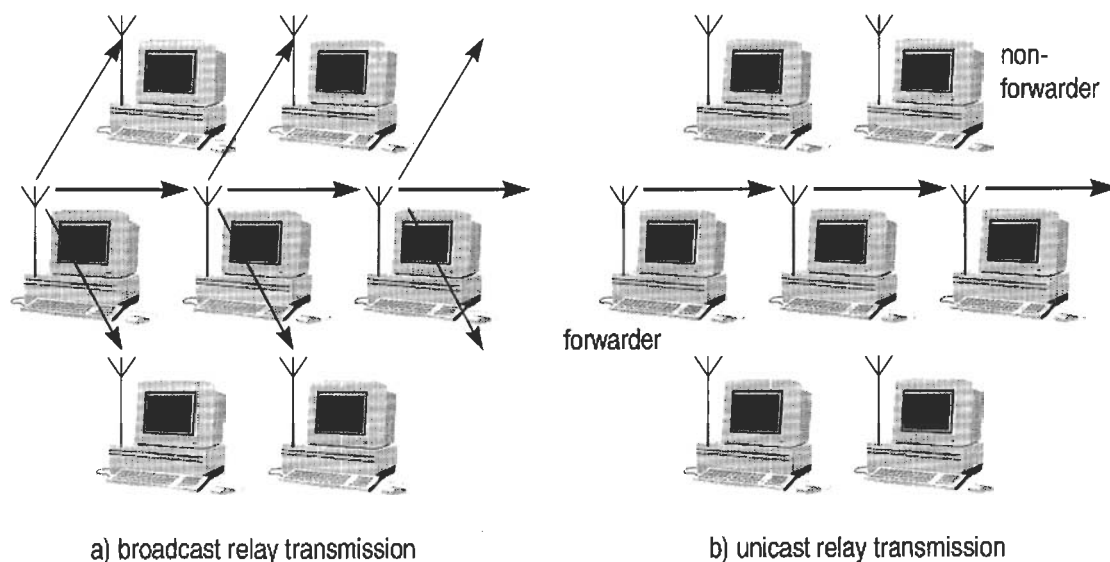




**Figure 24.6** The structure of a HIPERLAN wireless LAN

wireless LAN, a number of the stations are defined to be forwarders. The other stations are non-forwarders. The forwarding stations relay messages as necessary between themselves, broadcasting the message to any relevant non-forwarding stations within their sector as shown in Figure 24.7.

We shall not discuss the detailed radio channel access methodology and medium access control techniques here, but as an overview Figure 24.8 presents the protocol



**Figure 24.7** The forwarding function of HIPERLAN

OSI reference model	HIPERLAN protocol layers
higher layers	higher layers
data link layer (layer 2)	logical link control (LLC) IEEE802.2
	medium access control (MAC)
	channel access control (CAC)
physical layer (layer 1)	radio medium

Figure 24.8 HIPERLAN protocol reference model

reference model. Note that the use of the standard *IEEE 802.2 (ISO 8802.2) logical link control (LLC)* enables *HIPERLAN* to be used as a one-for-one replacement of an existing LAN. Unlike a normal LAN, however, two further sublayers are used beneath the LLC layer. In addition to a *medium access control (MAC)* sublayer, a *channel access control (CAC)* sublayer is also used. This is necessary to accommodate the control mechanisms necessary for the radio channel.

The *logical link control (LLC)* sublayer of OSI layer 2 provides for correct and secure delivery of information between two terminals connected to the LAN (error detection, correction, etc.). The *medium access control (MAC)* sublayer provides for the delivery of the information to the correct endpoint, providing for LAN addressing, data encryption and relaying as necessary. The *channel access control (CAC)* sublayer codes the MAC information into a format suitable for transmission across a radio medium. The technique used is called *non-pre-emptive priority multiple access (NPMA)*. NPMA breaks up the radio channel into a number of *channel access cycles*, each of which is further sub-divided into three cycle sub-phases

- *a priority resolution phase*
- *a contention resolution phase*
- *a transmission phase*

In the priority phase, any stations which do not currently claim the highest priority transmission status, are refused permission to transmit. The remaining stations compete for use of the radio channel during the next phase and any *contention* is resolved. The remaining transmission phase is then allocated to successful stations surviving both the *priority resolution* and *contention resolution* phases. This is the phase when user data are transmitted. Priorities will change from one cycle to the next to ensure that all stations have an equal ability to send data.

So much for the strengths of wireless LANs. The greatest difficulty is in achieving complete radio signal coverage throughout an office. *Multipath* effects, interference and propagation difficulties can lead to *blackspots* suffering very deep radio *fade* (i.e. poor transmission). For static devices, the problem of a fade caused by multipath of interference can be solved by moving the device only a small distance. For mobile terminals continuous good quality transmission may not be possible.

## 24.5 RADIODETERMINATION SATELLITE SERVICES (RDSS) AND THE GLOBAL POSITIONING SYSTEM (GPS)

In July 1985 the *Federal Communications Commission (FCC)* of the USA authorized the use of a new satellite radio service to be called the *radiodetermination satellite service (RDSS)*. It has widespread benefit in the field of navigation and in personal communication technology. The technical and operational standards adopted by the FCC became the basis for worldwide standards agreement under the auspices of ITU (the *global positioning system, GPS*).

GPS is a set of techniques combining radio and computer capabilities which is capable of determining precise geographical locations of points on the ground. It is used for applications such as tracking ships at sea or truck fleets on land. Its potential includes the scope for keeping track of individuals in support of global cellular radio services. Indeed some claim that GPS offers a more efficient means of tracking cellular handsets than the current method of continual updating.

The GPS system consists of a set of geostationary satellites, a control centre and a number of user terminals, called *transceivers*. The transceivers are typically quite small nowadays, even available in 'handheld' form, as many yachtsmen will be familiar with.

The control centre repeatedly sends an interrogation signal via the satellites to all the user terminals. Signals from the satellites are interpreted by each terminal and, if relevant, a response message is generated to

- alert the control centre of current position (if requested)
- reply to or request some other information from the control centre

The relative position of the user terminal from one satellite is computed by the control centre from the round-trip alert-and-response signal time scaled by the velocity of light. The relative range from three different geostationary satellites enables the control centre to compute exactly the position of the terminal in three-dimensional coordinates of latitude, longitude and altitude. This information is then associated with the terminal *identity (ID)* code and can either be passed to a third party or transmitted back to the user terminal. Thus a fleet operator could trace a lorry on the road or a ship owner could determine a ship's position at sea. Shipping companies will be familiar with the *Navstar* system.

PART 4

---

# MULTIMEDIA NETWORKS

---



---

## *Broadband, Multimedia Networks and the B-ISDN*

---

The emergence of 'multimedia' computers and software which use all sorts of different audio, data, image and video signals simultaneously has heralded a new generation of computers and computer 'applications' and spurred the need to develop and deploy a new universal technology for telecommunications. The broadband-ISDN (B-ISDN) will be this new universal technology. It will be a powerful network, capable of carrying many different types of services and communication bitrates.

### **25.1 MULTIMEDIA APPLICATIONS: THE DRIVER FOR BROADBAND NETWORKS**

A *multimedia* application is one in which different types of communication signal are transmitted simultaneously. Thus, for example, *multimedia* computers are able simultaneously to present spreadsheets, work processing and presentation foil 'windows' on the screen while simultaneously also showing a video, projecting a 'video telephone call' and playing a soundtrack (Figure 25.1). Multimedia applications greatly increase the possibilities of the modern computing and entertainment industries, and they are finding their way into all walks of modern life:

- *Teleworking*
- *Telemedicine*
- *Tele-education*
- *Teleshopping*



**Figure 25.1** A modern multimedia computer workstation (Courtesy of Siemens AG)

The simplest type of *teleworking* is that allowing a company executive to log his laptop PC into the company computer network and electronic mail facility no matter what his location. Thus working at home and during time away on a business trip becomes possible. Other forms of *teleworking* include applications designed to bring together groups of remotely located individuals into a *workgroup* (*workgroup applications*).

A *telemedicine* application might allow a group of specialist surgeons (in different locations) to carry out an emergency operation together, while sharing sight of a video signal broadcast from the operating theatre, as well as full reference to a patient's medical records and historic X-rays. A distant surgeon could indicate directly on the video exactly where an incision should be made, or maybe even conduct some of the surgery himself.

*Tele-education* offers the prospect of sharing the best possible educational facilities, information resources and equipment between a large number of dispersed locations. Thus books, video lectures and demonstrations can be available to students, even in their own homes.

Multimedia applications are also flourishing in the home environment. Here, cable television companies already widely offer 'cable' connections bringing simultaneous live television, public telephone service and *video-on-demand* services, as well as *teleshopping* (catalogue shopping from the television or computer screen).

## 25.2 VIDEO COMMUNICATION

The advance in video coding schemes and videocommunications has been one of the main contributing factors in the rapid emergence of multimedia applications. A number of standards have laid the basis for standardized coding of video information, providing for different qualities and bitrates. At the low bitrate end of the scale, ITU-T recommendation H.261 provides a standard algorithm for a videoconference *codec* (*coder/decoder*), suitable for coding relatively static images for transmission across low bitrate digital lines. Transmission at rates even as low as 64 kbit/s are possible, giving acceptable quality for a 'picture telephone' conversation or a *videoconference* between seated (and therefore relatively slow moving) participants in two or more different meeting room locations.

The H.261 algorithm works in a similar manner to ADPCM (Chapter 38). In effect the whole area of a video screen is assumed to be made up of a matrix of dots, each of a given colour (typically one of 256 shades). Thus each dot (or *pixel*, *picture element*) may be represented by an 8-bit digital value. The entire picture can be represented by sending all the digital values corresponding to the complete picture dot matrix (or bit map like the fax image discussed in Chapter 4). As a large number of dots make up the picture, it takes a little while at the start to establish the initial video *frame*. However, once the first frame has been established, it is sufficient to send the information corresponding to the difference between 'freeze frames' to recreate all subsequent frames. Provided that the movement of the image is relatively restricted, sending only the differences in the picture enables a much lower transmission bitrate to be used.

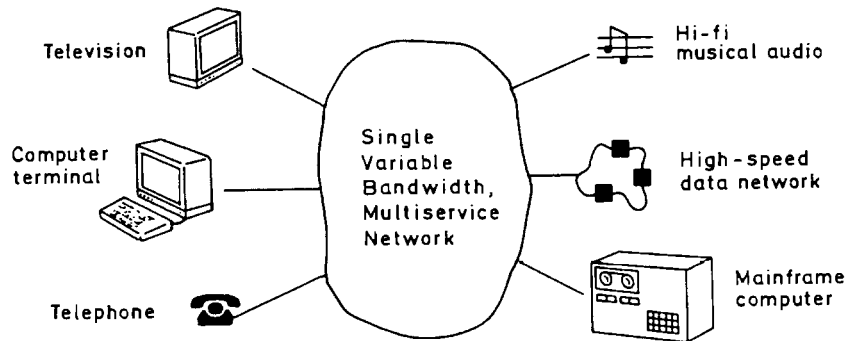
At 384 kbit/s, fairly high quality video images can be reproduced using the H.261 algorithm, and even public television signals are easily recognizable. The quality, however, does not match that of the original signal. For this reason, a number of other coding techniques are also used, where very high quality video images are to be stored and transmitted in conjunction with modern computer storage devices and telecommunications networks. These are the standards of the *Motion Pictures Experts Group* (*MPEG*), an industry common-interest group of major companies, cooperating to promote standards in this area.

## 25.3 THE EMERGENCE OF THE B-ISDN

The demands of new graphical computer software, the emergence of videotelephony and cable television, and the political pressure for the development of the *Information Highway* have all combined in their various ways to stimulate the development of switched *broadband* networks. The result has been the development of *broadband ISDN* (*broadband integrated services digital network*, or *B-ISDN*).

We saw in Chapter 10 how the *integrated services digital network* (*ISDN*) was one step in the development chain of a universal *integrated network*, and how it may replace the conventional telephone and data networks in the next few years. However, *narrow-band ISDN* (*N-ISDN*) in its current form falls a long way short of the ultimate needs of a multi-service network, because of the relatively restricted bit rate available on individual channels and because of the lack of service flexibility arising from the fact that it evolved





**Figure 25.2** The concept of the Broadband Integrated Services Digital Network (B-ISDN)

out of telephone network principles. The maximum rate currently possible on ISDN is  $n \times 64 \text{ kbit/s}$ , up to  $2 \text{ Mbit/s}$ . Such rates are not enough for very high speed file transfer between mainframe computers, or for interactive switching of broad bandwidth signals such as high definition television (HDTV) and video.

Increasing the unit bit rate ( $64 \text{ kbit/s}$ ) of ISDN would increase bit rates and make *broadband* and *multimedia* services possible, but it would be at the expense of gross inefficiency in network resources, particularly when carrying bursty data signals. So B-ISDN is not merely an extension of N-ISDN, but instead also capitalizes on the latest data switching techniques.

Current world technical standards for B-ISDN describe the general functions which are to be performed by a B-ISDN, and define the network interfaces to be used, but the technological details are not yet fully defined.

## 25.4 THE SERVICES TO BE OFFERED BY B-ISDN

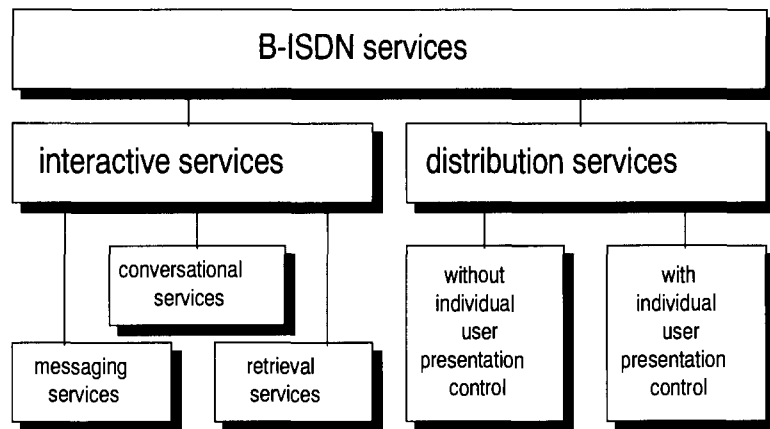
The services offered by B-ISDN networks split into two categories: *interactive services* and *distribution services* (Figure 25.3 shows).

*Interactive services* are normal communications between just two parties. There are three subtypes: messaging, conversational and retrieval services.

An example of a *conversational service* is a telephone conversation or a point-to-point data connection, where the two end-points of the communication are in *real-time* connected with one another and thus *converse*.

A *message service* is a telecommunication service similar to the postal service. A message is submitted to the network (as a letter is posted). Sometime later, the message is delivered to the given address. Message services are usually not *guaranteed*. The network is not able to check whether the recipient address is valid, and may return no confirmation to the sender of receipt. Thus 'no reply' may result either because the recipient never got the message or because he chose not to respond.

A *retrieval service* is one in which a caller accesses a central server, database or storage archive, requesting the delivery of certain specified information. He might, for example, receive data about holiday or request news video clips, etc.



**Figure 25.3** The service types offered by B-ISDN

*Distribution services* are services in which the information from a single source is distributed to many recipients at the same time. Distribution services are subdivided into those with or without individual *user presentation control*. An example of a distribution service without individual *user presentation control* is the broadcasting of national television. All the recipients receive the same signal at the same time. An example of a distribution service with individual user presentation control is *video-on-demand*. Here only those users who wish to pay and receive a given film do so.

## 25.5 THE EMERGENCE OF THE ATM SWITCHING TECHNIQUE AS THE HEART OF ATM

Much of the initial development work which led to B-ISDN started in the data networking field, as engineers tried to develop *fast packet switching* techniques suitable for the carriage of emerging video and 'image' *applications*. The switching technique

**Table 25.1** Potential applications of the various B-ISDN service categories

Category	Service type	Potential applications
Interactive services	Conversational services	Voice telephony
	Messaging services	Internet electronic mail
	Retrieval services	On-line database service
Distribution services	Without user control	Broadcast TV
	With user control	Pay-on-view TV or video-on-demand

developed as the basis for B-ISDN is an adaption of the packet switching technique used in datacommunications networks. It is called *ATM* (*asynchronous transfer mode*).

A common failing of telecommunication switching techniques previous to ATM was their restriction to low bit rate data conveyance or fixed bandwidth circuit switching. Thus the major difficulty which had to be overcome in developing ATM was the need to optimize B-ISDN to carry simultaneously both low and very high bit rate services, and to meet simultaneously the exacting demands and rapid signal conveyance necessary for speech and equivalent signals.

In classical data packet switching networks the emphasis is placed on assured and error-free delivery of data. This rules out such networks for the carriage of live speech and video signals due to the prolonged and uneven delays experienced during propagation and delivery of the user's information.

Figure 25.4 illustrates the degradation of a packetized speech channel, sharing a packet network with other speech and data users. The network is subject to congestion, so that occasional packets are held up in the buffer. The result is a rather disjointed signal, with 'staccato', 'snowy' and 'clipping' effects.

Another limitation of pre-ATM data networks, when used to carry the bitrates typically used for voice and video signals, was their relative inefficiency. When sending single packets of a small number of bits (eight for a single speech *sample*), the network is kept heavily loaded just carrying and processing packet headers! Furthermore the error detection and correction procedures which ensure that each packet value is received correctly are largely irrelevant for speech and video signals, because the ear or eye tends to 'fill-in' any slight inconsistencies in the signal anyway.

Despite these problems, statistical multiplexing (as used in data networks) is considered to be the ideal basis for broadband networks, as large numbers of connections can be accommodated of different and varying bandwidths. A large number of packets per second give a high bandwidth for one user, whereas a smaller number of packets per second yield lower bit rates for others.

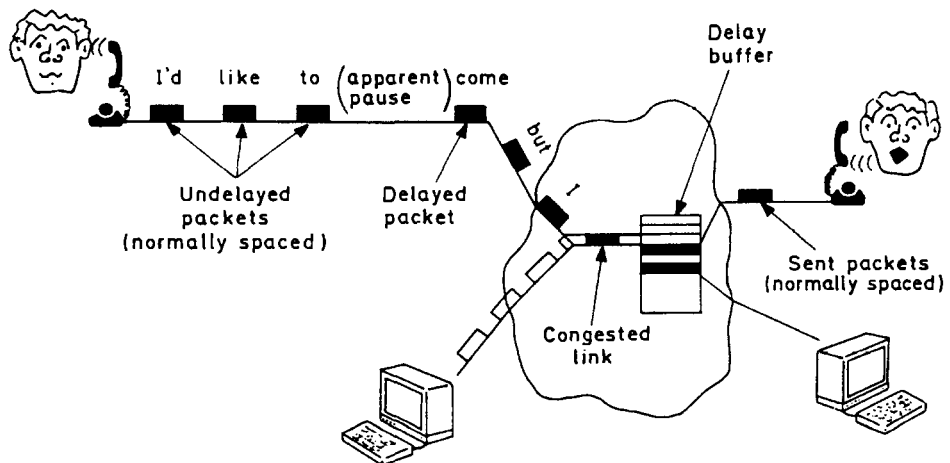


Figure 25.4 Degradation of speech on normal packet networks

In 1988, a statistical multiplexing technique called ATM was provisionally accepted by CCITT as the switching basis for future B-ISDNs, and a number of detailed CCITT (now ITU-T) *recommendations* were issued to define the basic technique. These provide for

- simultaneous carriage of mixed telecommunication service
- a connection establishment *contract*, in which the end-user device may specify during connection set-up the required connection bitrate and connection quality parameters, including maximum permissible delay and delay variation; this is then to be *guaranteed* by the network for the duration of the call
- network overload control, preventing new connections being established, and allocating available capacity to higher priority or delay sensitive (e.g. voice and video) connections first
- a limit to the delay experienced by video, voice and other delay-sensitive signals
- relatively high efficiency in the use of bandwidth for all types of applications, achieved by removing error correction methods in cases (e.g. voice and video applications) where these are superfluous

The ‘packets’ conveyed by ATM are called *cells*. Like the *packets* of X.25, each *cell* of ATM consists of an information field (in which the user data is held and conveyed) and a *header* which ‘tells’ the network where to deliver the packet, and provides a sequence number so that cells can be re-assembled in the correct order at the receiving end). But, unlike the packets of X.25, the cells in ATM are all of a fixed length (48 byte *payload* plus a 5 byte header). The fixed length gives the scope for overcoming the limitations of earlier packet networks.

The relatively short length of the cell (as compared with a normal data *packet* or *frame*) means that user data messages must be broken up into a large number of individual cells for transmission. This is inefficient, but on the positive side the smaller cells allow a transmission priority scheme to be established. High priority cells (e.g. deriving from a voice or video connection) can now be inserted between the cells which make up a single data frame, rather than having to wait until the complete frame has been transmitted. They thus experience significantly lower delay than they would have done in the case of a classical packet-switched data network.

Unfortunately, from a voice and video perspective the 48 byte information content of an ATM cell is very large compared to the single byte samples normally transmitted. The inefficiency associated with processing cell headers militates against sending very small (e.g. one byte) cells, corresponding to individual speech samples. Instead we must make do with conveying a number of consecutive speech samples in a single cell. A packet size of 48 bytes allows 48 separate samples to be sent simultaneously. This corresponds to 6 milliseconds worth of conversation (at 64 kbit/s). Thus a 6 ms propagation delay is inflicted on the signal as we wait for the cell to fill up before we can send it. Nonetheless, this is a small (and maybe unnoticed) price to pay for high network efficiency and the scope of B-ISDN.

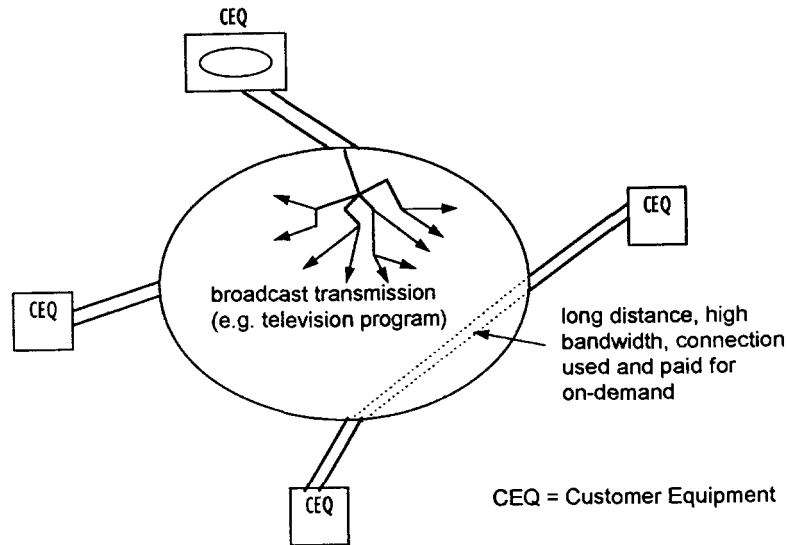


Figure 25.5 The basic switching capabilities of B-ISDN

## 25.6 CONNECTION TYPES SUPPORTED BY B-ISDN

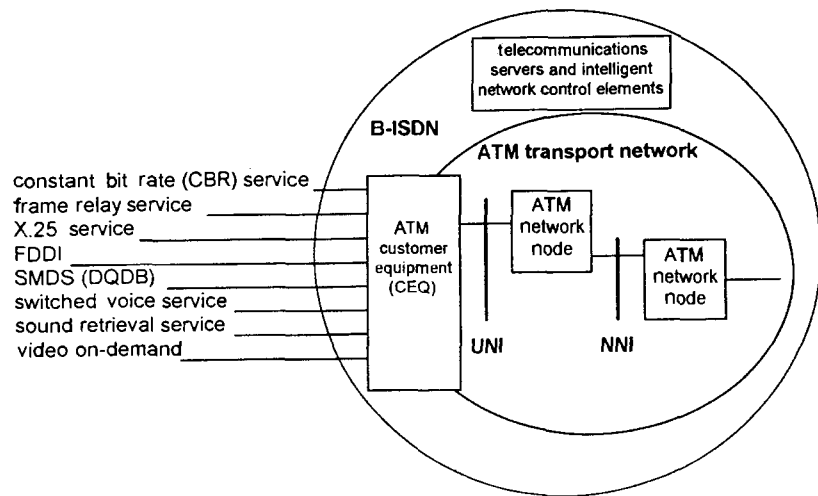
Over and above the switching of straightforward point-to-point connections, B-ISDN (through ATM) distinguishes itself from its predecessor networks not only in its ability to switch high bitrate connections, but also in its ability to set up broadcast (i.e. *multicast*) distributions of high bandwidth signals (e.g. video or television programmes). The broadcast capability, meanwhile, will enable cable TV companies to develop economically efficient metropolitan fibre cable networks for the distribution of pay-TV or *video-on-demand* (Figure 25.5).

Figure 25.6 illustrates the main types of connections and services which B-ISDN and ATM networks will offer. The diagram also shows how ATM relates to B-ISDN.

## 25.7 USER DEVICE CONNECTION TO B-ISDN

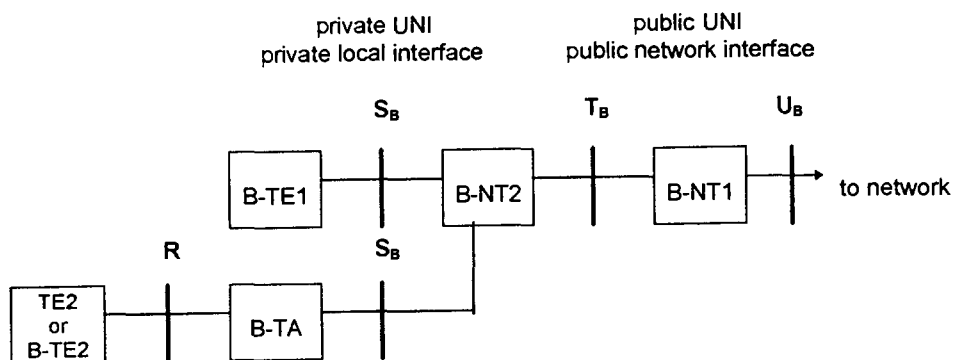
The *user-network interface (UNI)* illustrated in Figure 25.6 is the standardized interface for connecting users end devices to a B-ISDN network for the purpose of communication. Various *reference points* (defined by ITU-T recommendation I.413) are slightly different variants of the UNI interface, according to how the device is connected to the network. These are similar to the narrowband ISDN reference points, as Figure 25.7 shows.

The  $T_B$  interface is the basic UNI interface by which user equipments are connected to a B-ISDN at the B-NT1. The *B-NT1 (broadband ISDN network termination type 1)* provides line terminating functions for the public network operator (power feeding to line, management test functions, etc.). Where the user device is connected directly to the



<i>Term</i>	<i>Meaning</i>
CBR	constant bit rate (a point-to-point connection across an ATM network providing service similar to a private wire or leaseline)
DQDB	dual queue dual bus (the technology behind SMDS)
FDDI	fibre distributed data interface (a campus technology for interconnecting LAN routers)
NNI	network-node interface (an ATM network interface)
SMDS	switched multimegabit data service (an existing broadband network type)
UNI	user-network interface (an ATM network interface)

**Figure 25.6** ATM and B-ISDN: the connection types available



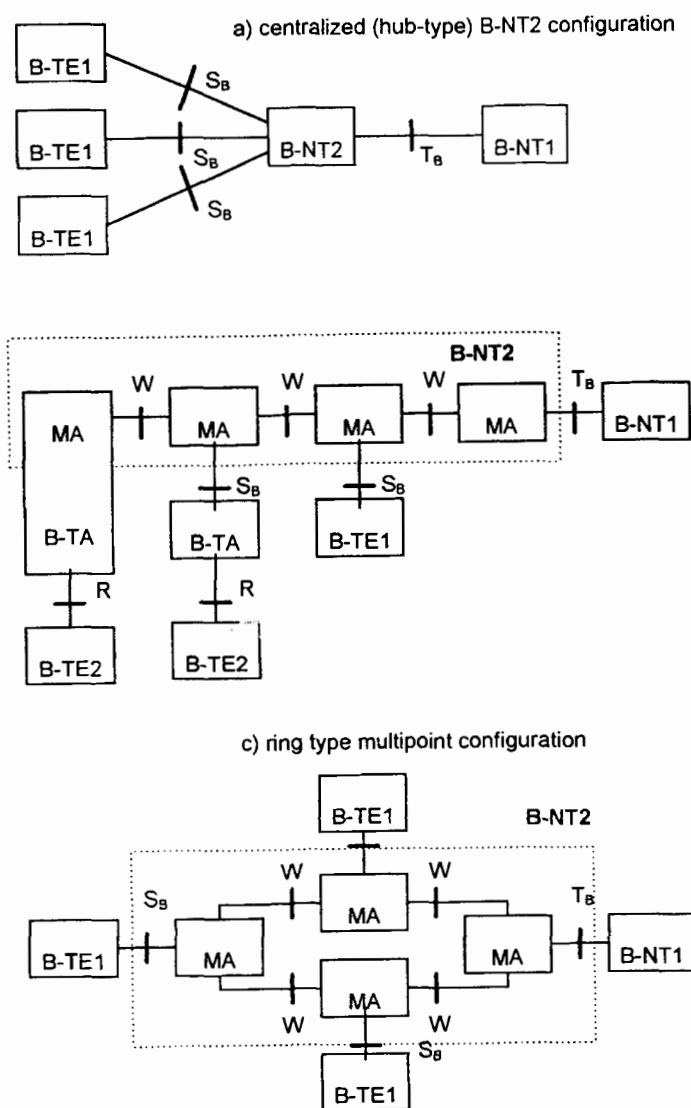
**Figure 25.7** B-ISDN reference configurations at the UNI

line the  $U_B$  interface is used (e.g. the case where the optical version of the UNI is employed).

Where an electrical interface is used at the UNI, a B-NT1 may be required to perform electrical to optical conversion of the signal for connection to a fibre network connection. Alternatively, the B-ISDN network operator may choose to install a device to

provide for network monitoring or management functions. If so, this device is also classified as a B-NT1, and the user side interface to it is then the UNI, reference  $T_B$ . The  $T_B$  interface is equivalent to the T interface of narrowband ISDN. The B suffix is merely to denote its connection with B-ISDN.

*B-NT2 (broadband ISDN network termination 2)* is the notation used to denote a user-provided switching device which permits multiple end user devices to share the same network connection. Thus a B-NT2 has one  $T_B$  connection, but may have multiple  $S_B$  ports. The equivalent in a narrowband ISDN is an NT2. An example of an NT2



**Figure 25.8** Multipoint configurations of the B-ISDN user-network interface (UNI)

device is an office PBX, an office-installed telephone exchange which switches a large number of individual extensions and a small number of main exchange lines. By using an NT2, better use can be made of the main exchange lines. The  $S_B$  interface will be very similar (if not identical) to the  $T_B$  interface.

*B-TE1* (*broadband ISDN terminal equipment type 1*) is the notation used to denote customer equipment (CEQ) which has been developed to operate over B-ISDN (i.e. ATM) networks. A B-TE1 is a B-ISDN-compatible end device.

A *B-TA* is a *broadband ISDN terminal adaptor*. This is a device provided to convert signals from non-B-ISDN-compatible end devices (*B-TE2*) into a format suitable for carriage over a B-ISDN (ATM network). Thus the R reference point is merely a notation used to refer to other types of end user device interfaces (i.e. any interface not covered by B-ISDN or ATM specifications).

ITU-T Recommendation I.413 also defines a W reference point in conjunction with a *medium adaptor* (MA). A number of medium adaptors perform the function of a B-NT2, allowing multiple end devices to be connected to a single UNI network connection. This is a *multipoint application* (Figure 25.8). The W reference point is to be found between MAs. The technical realization of the W reference point may be a non-standard (i.e. manufacturer proprietary) interface.

The various configurations that appear in Figure 25.8 reflect considerable scope and complexity for connection to B-ISDN. However, do not let yourself be put off by the mass of jargon before you notice that the possible configurations of the UNI mirror the classical office wiring configurations of LANs as we discussed in Chapter 19. The bus topology (Figure 25.8b) is like that of a classical coaxial cable *ethernet* LAN. The ring topology (Figure 25.8c) is like that of a *token ring* LAN or an FDDI network.

## 25.8 EVOLUTION TO BROADBAND-ISDN

The evolution of existing networks and applications to B-ISDN will be complex and frustrating, as are all complicated development projects. It will be many years before we see a fully integrated variable bit rate, broadband and multi-service network. Leading up to this time we will see the emergence of specialized broadband networks, optimized for particular applications. However, as technology develops and demand grows, a gradual migration of services from existing networks is inevitable. Cynics might say 'what demand?', and 'how can we possibly wish to convey so much information?' But cynics have been confounded before. It is only a matter of time!





---

## *Asynchronous Transfer Mode (ATM)*

---

ATM looks set to become the first universal telecommunication technology, capable of switching and transporting all types of telecommunication connection (e.g. voice, data video, multimedia). It will form the basis of the future broadband integrated services digital network (B-ISDN). Because of the anticipated importance of ATM, we discuss here the technical principles and terminology in depth, defining the main jargon and explaining what marks out ATM from its predecessors. In particular, we discuss the principles of statistical multiplexing and the specifics of cell switching.

### **26.1 A FLEXIBLE TRANSMISSION MEDIUM**

An ATM-equipped transmission line or telecommunication network is able to support

- usage by multiple users simultaneously
- different telecommunication needs (e.g. telephone, data transmission, LAN inter-connection, videotransmission, etc.)
- each application running at different transmission speeds (i.e. with differing bandwidth needs)

However, these capabilities are also offered by predecessor technologies, so why bother with ATM, you might ask? What distinguishes ATM from its predecessors is that it performs these functions more efficiently. ATM is capable of an instant-by-instant adjustment in the allocation of the available network capacity between the various users competing for its use. Rather than allocating fixed capacity between the two communicating parties for the duration of a *call* or *session*, ATM ensures that the line capacity is optimally used on a moment-by-moment basis, by carrying only the needed, or ‘useful’, information.

The dynamic allocation of bandwidth is achieved by ATM using a newly developed technique called *cell relay switching*. Understanding its principles is the key to understanding ATM, its strengths and limitations. In our discussion, we shall refer frequently to the work of the *ATM Forum*, an industry-wide common interest group, comprising telecommunications equipment manufacturers, network operators and users who have combined to speed the process of developing and agreeing technical standards for ATM.

## 26.2 STATISTICAL MULTIPLEXING AND THE EVOLUTION OF CELL RELAY SWITCHING

ATM is based upon a *statistical multiplexing* technique called *cell relay switching*. *Statistical multiplexing*, as we discussed in Chapters 9 and 18, is widely used to improve the efficiency of data networks. As we shall see, it can also be used effectively to carry speech connections, provided extra provisions are made to control the propagation delay.

The major benefit of statistical multiplexing is that the useful carrying capacity of the line is maximized by avoiding the unnecessary transmission of redundant information (i.e. pauses in speech or idle periods on data lines). In addition, as the full capacity of the line (i.e. its full speed) is made available for each individual connection for carriage of information, the transmission time (propagation time) may be reduced.

In the example of Figure 26.1, we recap the technique of statistical multiplexing. Three separate users (represented by sources A, B and C) are to communicate over the same transmission line, sharing the line resources by means of statistical multiplexing. The three separate source circuits are fed into a statistical multiplexor, which is connected by a single line to the demultiplexor at the receiving end. (A similar arrangement, using a second line for the receive channel, but with multiplexor on the right and demultiplexor on the left, will also be necessary but this is not shown).

The statistical multiplexor sends whatever it receives from any of the source channels directly onto the transmission line. Why is it called *statistical multiplexing*? Simply because it relies on the statistical unlikelihood of all three channels wanting to send information simultaneously. Actually, for a short period of time, the multiplexor is designed to be able to cope with simultaneous transmission from all of the sources. This

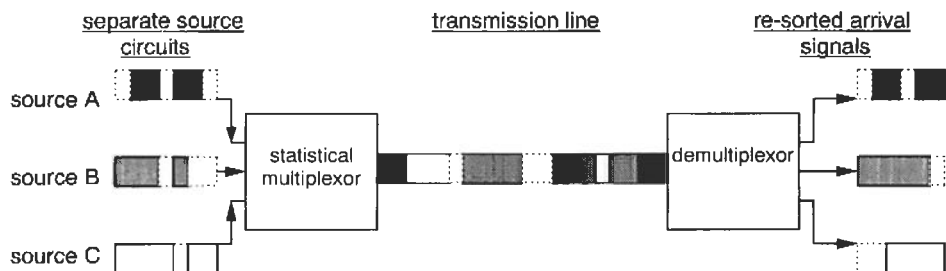


Figure 26.1 The principle of statistical multiplexing

is done by sending the most important signal directly to line and storing the lesser important signal in a *buffer* for an instant until the first signal has been transmitted. This causes a slight *variation* in the time taken by signal *packets* to traverse the network.

Variation in the time packets take to traverse a data network is not important. So long as the average delay is not great, computer users do not notice whether some typed characters appear imperceptibly faster or slower than others. As a result, *packet-switched* networking techniques have dominated the world of data transmission. Voice networks meanwhile have remained circuit switched networks, because they cannot tolerate variable delays, because they tend to 'chop up' the signal.

The strength of *circuit switching* is the guaranteed throughput and fast and constant signal propagation time over the resulting connection. This is critical so that acceptable voice quality can be achieved (any variation in the signal propagation time manifests itself to the listener as a rather broken up form of the original signal, an effect known as *jitter*). A telephone call connected in a circuit-switched manner is much like an empty pipe between two telephone users. Whatever one speaker talks into the pipe comes out at the other end in an identical format, but only one pair of callers can use the pipe during any particular call.

And so it came to pass . . . that there were networks for data . . . and separate networks for telephones. The *cell relay* technique of statistical multiplexing used in ATM is designed to contain the variation in propagation delay experienced by delay-sensitive signals such as voice and video. It is heralded as the first technique capable of efficient voice and data integration within a single network.

## 26.3 THE PROBLEMS TO BE SOLVED BY CELL RELAY

Returning to the statistical multiplexing example of Figure 26.1, a typical data packet contains between 1 and 256 characters (i.e. between 8 and 2048 bits), and the linespeed is typically 9600 bit/s. Therefore the delay at a time when two sources try to send simultaneously and one of the packets has to be temporarily stored in a buffer will be up to 200 ms (2048/9600 s) longer than when there is no simultaneous transmission. In other words, there may be up to 200 ms of *jitter*. This is unacceptable for speech transmission, but this is not the only difficulty to be overcome. Another problem is in minimizing the loss of line bandwidth to the management *overhead* of statistical multiplexing, as we discuss next.

To allow a statistical demultiplexor (Figure 26.1) to sort out the various packets belonging to the different *logical connections*, and forward them to the correct destinations (A to A, B to B, C to C, etc.) there needs to be a *header* attached to each packet to say which *logical connection* (i.e. telephone conversation or data communications session) it belongs to. The *header* (Figure 26.2) is crucial, but has the disadvantage that it adds to the information which must be carried between multiplexor and demultiplexor. At the demultiplexor, the header is removed so it does not disturb the receiver, but meanwhile it has generated an *overhead* load for the transmission line. It is thus impossible using statistical multiplexing techniques to achieve 100% loading with raw user information. Some of the capacity needs to be given up to carry the *overhead*.

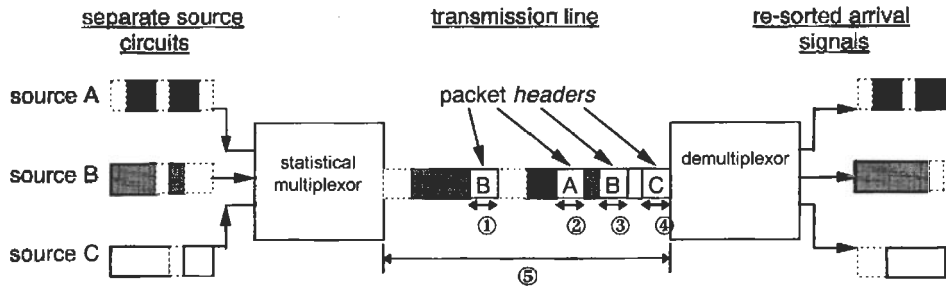


Figure 26.2 Statistical multiplexing headers and overhead

The major challenges for ATM developers are to minimize the jitter experienced by speech, video and other delay-sensitive applications, while simultaneously optimizing line efficiency by minimizing network *overhead*. As we shall see, these demands contend with one another.

## 26.4 THE TECHNIQUE OF CELL RELAY

*Cell relay* is a form of statistical multiplexing that is similar in many ways to *packet switching*, except that the *packets* are instead called *cells*. Each of the *cells* is of a fixed rather than a variable size. The fixed cell size defined by ATM standards is 48 octets (bytes) plus a 5 octet header (i.e. 53 octets in all, see Figure 26.3). The transmission line speeds currently foreseen to be used are either 2 Mbit/s, 12 Mbit/s, 25 Mbit/s, 34 Mbit/s, 45 Mbit/s, 52 Mbit/s, 155 Mbit/s or 622 Mbit/s. We may thus conclude that

- the overhead is at least 5 bytes in 53 bytes, i.e.  $>9\%$
- the duration of a cell is at most (i.e. at 2 Mbit/s line rate)  $53 \times 8 \text{ bits} / 2 \text{ Mbits}^{-1} = 0.2 \text{ ms}$  ( $12 \mu\text{s}$  at  $34 \text{ Mbits}^{-1}$ )

As the cell duration is relatively short, provided that a priority scheme is applied to allow cells from delay-sensitive signal sources (e.g. speech, video, etc.) to have access to the next cell *slot*, then the jitter (variation in signal propagation delay) can be kept very low (i.e. of the order of 0.2 ms); not zero as is possible with circuit switching, but at least low enough to give a subjectively acceptable quality for telephone listeners or video watchers. Jitter-insensitive traffic sources (e.g. datacommunication channels) can be made to wait for the allocation of the next *low priority slot*.

The jitter could be reduced still further by reducing the cell size, but this would increase the proportion of the line capacity needed to carry the cell headers, and thus reduce the line efficiency.



Figure 26.3 ATM 53 byte cell format



Figure 26.4 The cells and slots of cell relay

## 26.5 THE ATM CELL HEADER

The *cell header* carries information sufficient to allow the ATM network to determine to which connection (and thus to which destination port and end user) each cell should be delivered. We could draw a comparison with a postal service and imagine each of the cells to be a letter of 48 alphanumeric characters contained in an envelope on which a five-digit postcode appears. You simply drop your letters (cells) in in the right order and they come out in the same order at the other end, though maybe slightly jittered in time. Just like a postal service has numerous vans, lorries and personnel to carry different letters over different stretches and sorting offices to direct the letters along their individual paths, so an ATM network can comprise a mesh of transmission links and switches to direct individual cells by inspecting the *address* contained in the *header* (Figure 26.5).

The ATM 'switch' acts in much the same way as a postal sorter. On its incoming side is a *FIFO* (*first in-first out*) buffer, like a pile of letters. At the front of the buffer (like the top letter in the pile) is the cell which has been waiting longest to be switched. New cells arriving are added to the back of the buffer. The switching process involves looking at each cell in turn, and determining from the address held in the header which outgoing line should be taken. The cell is then added to the FIFO output buffer which is queueing cells waiting to be transmitted on this line. The cell then proceeds to the next exchange.

You may think that a five-digit postcode is rather inadequate as a means of addressing all the likely users of an ATM network, and it might be, were it not for several provisions of the ATM specifications. First, the 'digits' are whole octets (base 256) rather than decimal digits (base 10). This means that the header has the range for over  $10^{12}$  combinations (40 bits), though only a maximum of 28 bits ( $2.7 \times 10^8$  combinations) are ever used for addressing. Second, the addresses (correctly called *identifiers*) are only allocated to active connections.

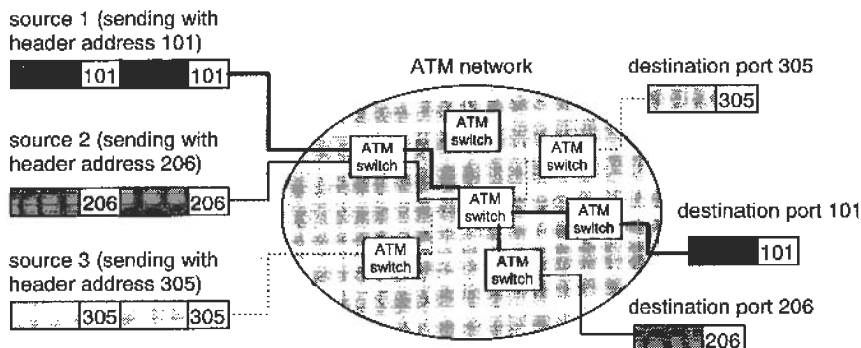


Figure 26.5 Switching in an ATM network

ATM connections are allocated an *identifier* during call set-up, and this is re-allocated to another connection when the connection is cleared. In this way the number of different identifiers need not directly reflect the number of users connected to the network (which may be many millions), only the number of simultaneously active connections. In addition, various subregions of the network may use different identifier schemes, thus multiplying the available capacity, but then demanding the ability of network nodes to *translate* (i.e. amend) identifiers in the five-octet header.

By highly efficient usage of the information carried in the header, the length of the header can be kept to a minimum. As a result, the network *overhead* is minimized.

## 26.6 THE COMPONENTS OF AN ATM NETWORK

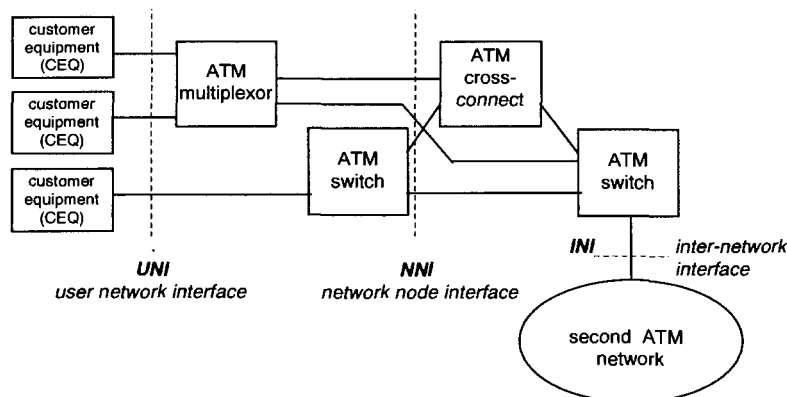
There are four basic types of equipment which go to make up an ATM network. These are

- *customer equipment (CEQ)*, also called *B-TE (broadband-ISDN terminal equipment)*
- *ATM switches*
- *ATM crossconnects*
- *ATM multiplexors*

These elements combine together to make a network as shown in Figure 26.6. A number of standard interfaces are also defined by the ATM specifications as the basis for the connections between the various components. The most important interfaces are

- the *User–Network Interface (UNI)*
- the *Network–Node Interface (NNI)*
- the *Inter-Network Interface (INI)*

These are also shown in Figure 26.6.



**Figure 26.6** The components of an ATM network (ITU-T network reference model for ATM)

ATM *customer equipment (CEQ)* is any item of equipment capable of communicating across an ATM network. One of the most popular of today's visions is the concept of *multimedia* applications, devices capable of enabling their users simultaneously to transmit synchronized video, electronic mail, data applications and telephone messages over the same line at the same time.

The ATM *user network interface (UNI)* is the standard technical specification allowing ATM customer equipment (CEQ) from various different manufacturers to communicate over a network provided by yet another manufacturer. It is the interface employed between ATM customer equipment and either ATM multiplexor, ATM crossconnect or ATM switch. It consists of a set of layered protocols as we shall discuss later.

Customer equipments communicate with one another across an ATM network by means of a *virtual channel (VC)*. The VC may either be set-up and cleared down on a call-by-call basis similar to a telephone network, in which case the connection is a *switched virtual circuit (SVC)* or it may be a permanently dedicated connection (like a lease line or private wire), in which case it is a *permanent virtual circuit (PVC)*.

An ATM *multiplexor* allows different *virtual channels* from different ATM UNI ports to be bundled for carriage over the same physical transmission line. Thus two or three customers outlying from the main exchange (Figure 26.6) could share a common line. Returning to our analogy with the postal system, the ATM multiplexor performs a similar function to a postal sack; it makes easier the task of carrying a number of different messages to the sorting station (ATM switch) by bundling a number of virtual channels into a single container, a *virtual path (VP)*. More about *virtual channels* and *virtual paths* later in the chapter (Figures 26.10–26.13).

An ATM *crossconnect* is a slightly more complicated device than the ATM multiplexor. It is analogous to a postal depot, where the various vanloads of mail are unloaded, the various sacks sorted and adjusted into different van loads. At the postal depot, the individual sacks remain unopened, and at the ATM crossconnect, the *virtual path* contents, the individual *virtual channels* remain undisturbed. The ATM crossconnect appears again in Figure 26.12.

A full ATM *switch* is the most complex and powerful of the elements making up an ATM network. It is capable not only of cross connecting *virtual paths*, but also of sorting and switching their contents, the *virtual channels* (Figure 26.13). It is the equivalent of a full postal sorting office, where sacks can either pass through unopened, or can be emptied and each letter individually re-sorted. It is the only type of ATM node device capable of interpreting and reacting upon user or network signalling for the establishment of new connections or the clearing of existing connections.

The ATM *network node interface (NNI)* is the interface used between nodes within the network or between different sub-networks. A standardized NNI gives the scope to build an ATM network from individual nodes or sub-networks supplied by different manufacturers.

The *inter-network interface (INI)* allows not only for intercommunication, but also for clean operational and administrative boundaries between interconnected networks. It is based on the NNI but includes more features for ensuring security, control and proper administration of inter-carrier connections (i.e. where networks of two different operators are interconnected). ATM Forum calls this interface *B-ICI (broadband inter-carrier interface)*.



## 26.7 THE ATM ADAPTATION LAYER (AAL)

An extra functionality is added to a basic ATM network (correctly called the *ATM Layer*) to accommodate the carriage of various different types of connection-oriented and connectionless network services (Chapter 25, Figure 26.6). This functionality is contained in the *ATM adaptation layer*. The *ATM adaptation layer (AAL)* lays out a set of rules on how the 48 byte cell payload can be used, and how it should be coded. These special codings enable the end devices which are communicating across the *ATM Layer* to communicate with one another using any of the possible *connection-oriented* or *connectionless* service as desired.

The services offered by the *ATM adaptation layer (AAL)* are classified into four *classes* or *types* (the standards use both terminologies). The distinguishing parameters of the various classes are as illustrated in Table 26.1.

An example of a Class A service is *circuit emulation* (i.e. a connection service providing for 'clear channel' connections like hard-wired digital circuits). In the ATM specifications such services are referred to as *constant bit rate (CBR)* or *circuit emulation services (CES)*. Thus a constant bit rate video or speech signal would be an AAL Class A service and would use AAL1.

*Variable bit rate (VBR)* video and audio is an example of a class B service. Thus an audio speech signal which sends no information during silent periods is an example of a Class B VBR service and would use AAL2.

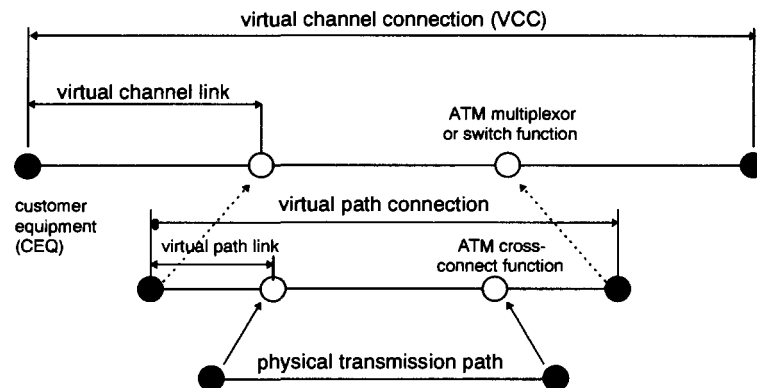
Class C and Class D cover the *connection-oriented* and *connectionless* data transfer services. Thus an X.25 packet switching service would be supported by a Class C service, and connectionless data services like electronic mail and certain types of LAN router service would be Class D. Both classes C and D use AAL types AAL3/4 or AAL5.

## 26.8 ATM VIRTUAL CHANNELS AND VIRTUAL PATHS

A *virtual channel* extended all the way across an ATM network (*ATM Layer*) is actually a *virtual channel connection (VCC)*. This connection is composed of a number of shorter length *virtual channel links*, which when laid end-to-end make up the VCC.

**Table 26.1** Service classification of the ATM adaption layer (AAL)

Transmission characteristic	Class A	Class B	Class C	Class D
AAL Type	AAL Type 1 (AAL1)	AAL Type 2 (AAL2)	AAL Type 3/4 (AAL3/4), AAL Type 5 (AAL5)	AAL Type 3/4 (AAL3/4), AAL Type 5 (AAL5)
Timing relation between source and destination	Required		Not required	
Bit rate	Constant	Variable		
Connection mode	Connection-oriented			Connectionless



**Figure 26.7** The relationship between virtual channels, virtual paths and physical transmission paths

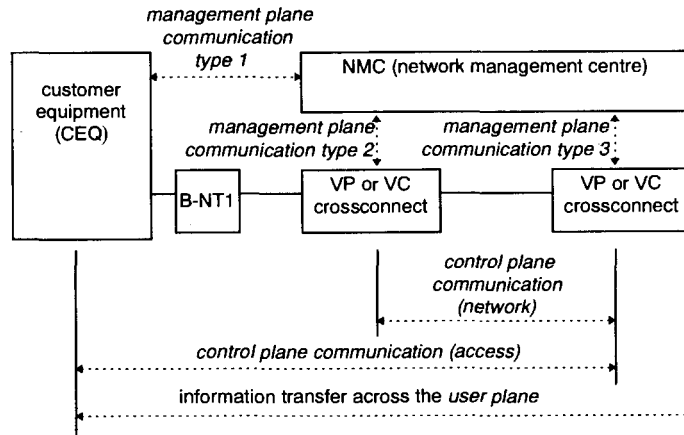
A *virtual channel link* is a part of the overall VCC, and shares the same endpoints as a *virtual path connection (VPC)* (Figure 26.7). The idea of a *virtual path (VP)* is valuable in the overall design and operation of ATM networks. As we have already discovered in the earlier part of a chapter, a *virtual path* has a function rather like a postal sack. In the same way that a postal sack helps to ease the handling of letters which all share a similar destination, so a virtual path helps to ease the workload of the ATM network nodes by enabling them to handle bundled groups of *virtual channels*. Thus a *virtual path (VP)* carries a number of different *virtual channel links*, which in their own separate ways may be concatenated with other virtual channel links to make VCCs.

Just like *virtual channels*, *virtual paths* can be classified into *virtual path connections (VPCs)* and *virtual path links*, where a VPC is made up by the concatenation of one or more virtual path links. A *virtual path link* is derived directly from a physical transmission path.

## 26.9 USER, CONTROL AND MANAGEMENT PLANES

Before two *customer equipments (CEQ)* may communicate with one another (i.e. transfer information) across the *user plane (U-plane)* of an ATM network, a connection must first be established. The connection is established by means of a *control* or a *management* communication between the CEQ and the network. This communication may take one of five forms (Figure 26.8)

- *control plane communication (access)*
- *control plane communication (network)*
- *management plane communication type 1*
- *management plane communication type 2*
- *management plane communication type 3*



**Figure 26.8** User, control and management planes of an ATM network (ITU-T recommendation I.311)

A *control plane communication (access)* is a one conducted between *CEQ* (customer equipment) and an ATM switch. During such a communication, which uses *UNI signalling*, a connection is established or released (in the case of *SVCs*, *switched virtual circuits*) much like dialling a telephone number in a telephone network. *Control plane communications (network)* will follow, as the ATM switch communicates (using network signalling) with other nodes in the network to establish the complete network connection. Once the connection is established, the user *transfers information* (i.e. communicates) across the *user plane*.

The connection could also have been established manually by the service technicians at the network management centre (a *PVC*, *permanent virtual circuit*). In this case, the user uses a *management plane communication type 1* from his *CEQ* to the *NMC* to request the establishment of a permanent connection. This could be carried by *UNI signalling* or could simply be a telephone call. The various switches and other network elements are then configured from the *NMC* by means of messages sent by *management plane communication type 2*.

*Management plane communication type 3* is initiated by ATM switches which require to refer to the *NMC* for information, authority or other assistance in the process of connection set-up. (It may be, for example, that certain high bandwidth connections require authorization from the *NMC* to prevent network congestion at peak times).

## 26.10 HOW IS A VIRTUAL CHANNEL CONNECTION (VCC) SET UP?

A *UNI signalling virtual channel (SVC)*, but not to be confused with *SVC*, *switched virtual connection*) is a virtual channel or virtual path connection at a *UNI* dedicated specifically to *UNI signalling*. Signalling virtual channels may also exist at an *NNI* interface.

A signalling message sent over the *SVC (signalling VC)* might be 'set up virtual connection number one between user A and user B'. Another message might be 'clear the connection between A and B'. ATM uses a dedicated channel for signalling information (*common channel signalling* as we discussed in Chapter 7). Drawing a parallel between ATM signalling and *narrowband ISDN* signalling, the UNI signalling interface is equivalent to the narrowband ISDN signalling defined by ITU-T recommendation Q.931 (and indeed is based on it and specified in Q.2931), and ATM NNI signalling is based on *signalling system 7 (SS7)* as used in narrowband ISDN.

At the time when a CEQ requests to set up a new *SVC (switched virtual circuit)* connection across an ATM network, it must first negotiate with the network over the *UNI signalling VC*, declaring the required *peak cell rate*, *quality of service (QOS) class* and other parameters needed. The *connection admission control (CAC)* function at the ATM switch then decides whether sufficient resources are available to allow immediate connection. If so, the connection is set up. If not, the connection request is rejected to protect the quality of existing connections. (The analogy is the telephone user's receipt of busy tone when no more lines are available). During the negotiation, virtual paths and connections between the various nodes and other equipments are allocated, and the reference numbers of these connections, the combination of *virtual path identifiers (VPIs)* and *virtual channel identifiers (VCIs)*, are confirmed over the signalling channel. These values (VPI and VCI) then appear in the *header* of any cells sent, to identify all those cells which relate to this connection.

## 26.11 SIGNALLING VIRTUAL CHANNELS AND META-SIGNALLING VIRTUAL CHANNELS

Both management and control communication in an ATM network take place via *signalling virtual channels (SVCs)*. At NNI interfaces, SVCs are usually permanently configured between the various servers (i.e. control processors) controlling a particular B-ISDN service (e.g. *video on demand*, picture telephone service, etc.). However, unlike narrowband ISDN, *signalling virtual channels (SVC)* are not normally permanently available at UNI. Instead, they are established on demand by means of a *meta-signalling virtual channel (MSVC)*. This is a permanently allocated UNI signalling channel of a fixed bandwidth. It is found in the virtual path VPI=0 and has a VCI value standard to the particular network.

By means of the *meta-signalling virtual channel*, the end device (CEQ) can establish an SVC (signalling VC) to the ATM switch (c-plane) or to the network management centre (m-plane) for signalling communication. A *service profile identifier (SPID)* carried in the meta-signalling determines which service the user requires, and enables a signalling virtual channel to the appropriate *signalling point* server to be established.

The functionality or device which exists at the end of a signalling virtual channel and conducts the act of signalling is called a *signalling point (SP)*. Such functionality exists in customer equipment (CEQ) and in ATM switches. A *signalling transfer point (STP)* is a switching point for the information carried in signalling virtual channels. Using a single *signalling VC* via an STP, an SP may communicate signalling messages to

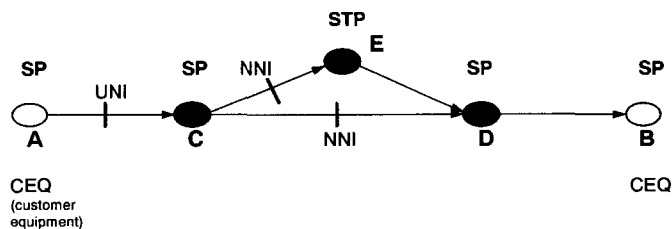


Figure 26.9 Signalling points (SPs) and signalling transfer points (STPs)

numerous other SPs using either *associated-mode* or *quasi-associated mode* signalling, as we discussed in Chapter 12. STPs improve the efficiency and reliability of the signalling network (Figure 26.9).

## 26.12 VIRTUAL CHANNEL IDENTIFIERS (VCIs) AND VIRTUAL PATH IDENTIFIERS (VPIs)

As we saw in Figure 26.7, *virtual channel connections* comprise concatenated *virtual path connections*. Each is *identified* by reference numbers carried by the cell headers of active connections called *virtual channel identifiers (VCIs)* and *virtual path identifiers (VPIs)*.

Figure 26.10 illustrates how a physical connection may be subdivided into a number of different virtual paths, each with a unique VPI. Each VPI in turn may be subdivided into several virtual channels, each with a separate VCI. The combination of VPI and VCI values is unique to each UNI or NNI and is sufficient to identify any active connection at the interface (i.e. on the same physical connection).

An ATM multiplexor, as we discussed earlier, allows a number of *virtual channels* from separate virtual paths to be combined over a single *virtual path*. Thus the virtual channels carried by physical links 1, 2 and 3 of Figure 26.11 are combined together into a single virtual path carried by physical link 4. In this way three separate end user devices use separate virtual channels to share a single physical connection line from ATM multiplexor to the ATM network.

An ATM crossconnect allows rearrangement of *virtual paths* without disturbance of the virtual channels which they contain. Thus in Figure 26.12, the contents of incoming

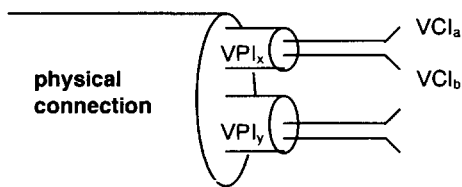


Figure 26.10 Virtual path and virtual channel identifiers

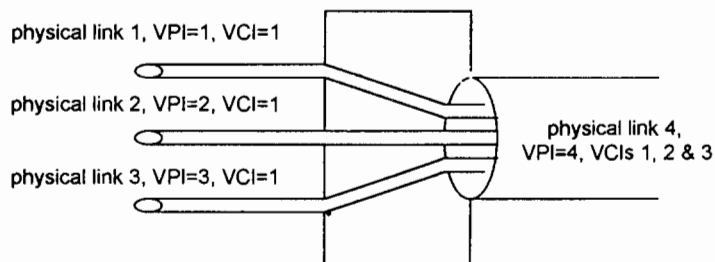


Figure 26.11 ATM multiplexor

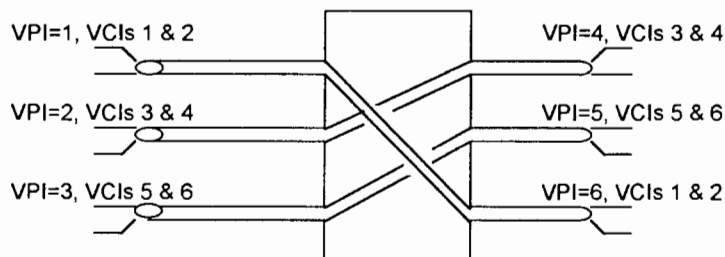


Figure 26.12 ATM crossconnect

virtual path VPI = 1 are crossconnected to outgoing virtual path VPI = 6, the VCIs remaining unchanged. An ATM crossconnect is thus a simple form of ATM switch, but one which need only process (and *translate* (i.e. amend)) VPI values.

A full ATM switch (Figure 26.13) must have the capability not only to crossconnect virtual paths, but also to switch virtual channels between different virtual paths. This requires the additional ability to process and translate VCIs held in cell headers. It is thus a relatively complex device. Good performance depends on very fast processing of both VPIs and VCIs in the cell headers. A full ATM switch is consequently a costly device.

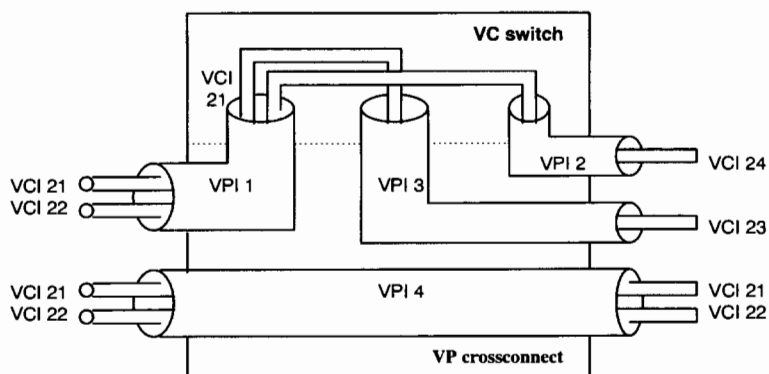


Figure 26.13 Full ATM switch

## 26.13 INFORMATION CONTENT AND FORMAT OF THE ATM CELL HEADER

The main function of the ATM cell header is to carry the VPI and VCI information which allows the active network elements to switch the cells of active connections through the network. The exact format of the cell header is as shown in Figure 26.14.

The cell header comprises 40 bits, of which 24 (UNI) or 28 (NNI) are used for the virtual path and virtual channel identifiers. Together, the VPI/VCI fields are called the *routing field*. There are four other *fields* which occupy the remainder of the header. The *PT* (*payload type*) field is occupied by the *payload type identifier (PTI)*. This identifies the contents of the cell (the *information field* or *payload*) as either a user data cell, a cell containing network management information, or a resource management cell. The *cell loss priority (CLP)* bit (when set to value 1 is used to identify less important cells which may be discarded first at a time of network or link congestion. The *generic flow control (GFC)* field is used to control the cell transmission between the customer equipment (CEQ) and the network (Figure 26.15). Finally, the *header error control (HEC)* field serves to detect errors in the cell header caused during cell transmission.

When there is no trunk congestion (i.e. there is no appreciable accumulation of cells waiting in the multiplexor buffer to be transmitted over the trunk) then the GFC field is set to the *uncontrolled transmission mode*. However, if there is a sudden surge of cells from all of the CEQ devices and the multiplexor experiences congestion (i.e. the filling of its cell buffers to a critical threshold level) then the GFC field is used to subject the

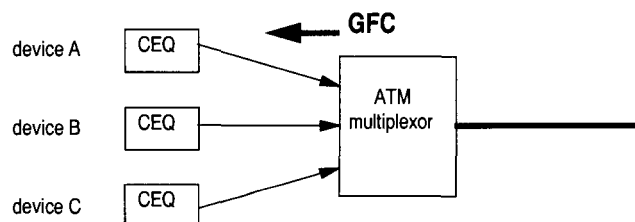
8	7	6	5	4	3	2	1	byte octet
GFC (at UNI), VPI (at NNI)				VPI				1
VPI				VCI				2
				VCI				3
VCI				PT		CLP		4
				HEC				5

**GFC** = Generic Flow Control  
**VPI** = Virtual Path Identifier  
**VCI** = Virtual Channel Identifier

**PT** = Payload Type  
**CLP** = Cell Loss Priority  
**HEC** = Header Error Control

**Note:** the GFC is used only at the UNI, at the NNI bits 5-8 of octet 1 are used as VPI.

**Figure 26.14** Structure of the ATM cell header



**Figure 26.15** Generic flow control regulates trunk congestion at a multiplexor

**Table 26.2** The functional layers of ATM

Layer name	Sublayer name	Further sublayer
Higher layers		
ATM adaptation layer (AAL)	● Convergence sublayer (CS)	● Service specific (SS) ● Common part (CP)
	● Segmentation and reassembly (SAR) sublayer	
ATM layer	● VC level	
	● VP level	
Physical layer	● Transmission convergence (TC) sublayer	
	● Physical medium (PM)	

- Service access point (SAP) – an imaginary point between functional layers.

cell flow from the various CEQ devices to *controlled transmission*. This limits the rate at which the CEQ devices may continue to send cells of one or more different types to the network.

An ATM cell is transmitted in the order of octets (i.e. octet 1 first, followed by octet 2, then octet 3, etc.). The *most significant bit (MSB)* of each octet (i.e. bit 8) is transmitted first. Thus first the *header* and then the *payload* is transmitted, MSB first.

## 26.14 ATM PROTOCOL LAYERS

Table 26.2 illustrates the protocol layers of ATM. We use this in the discussion which follows to define the terminology of ATM and to explain the relationships of the various layers to one another.

## 26.15 THE ATM TRANSPORT NETWORK

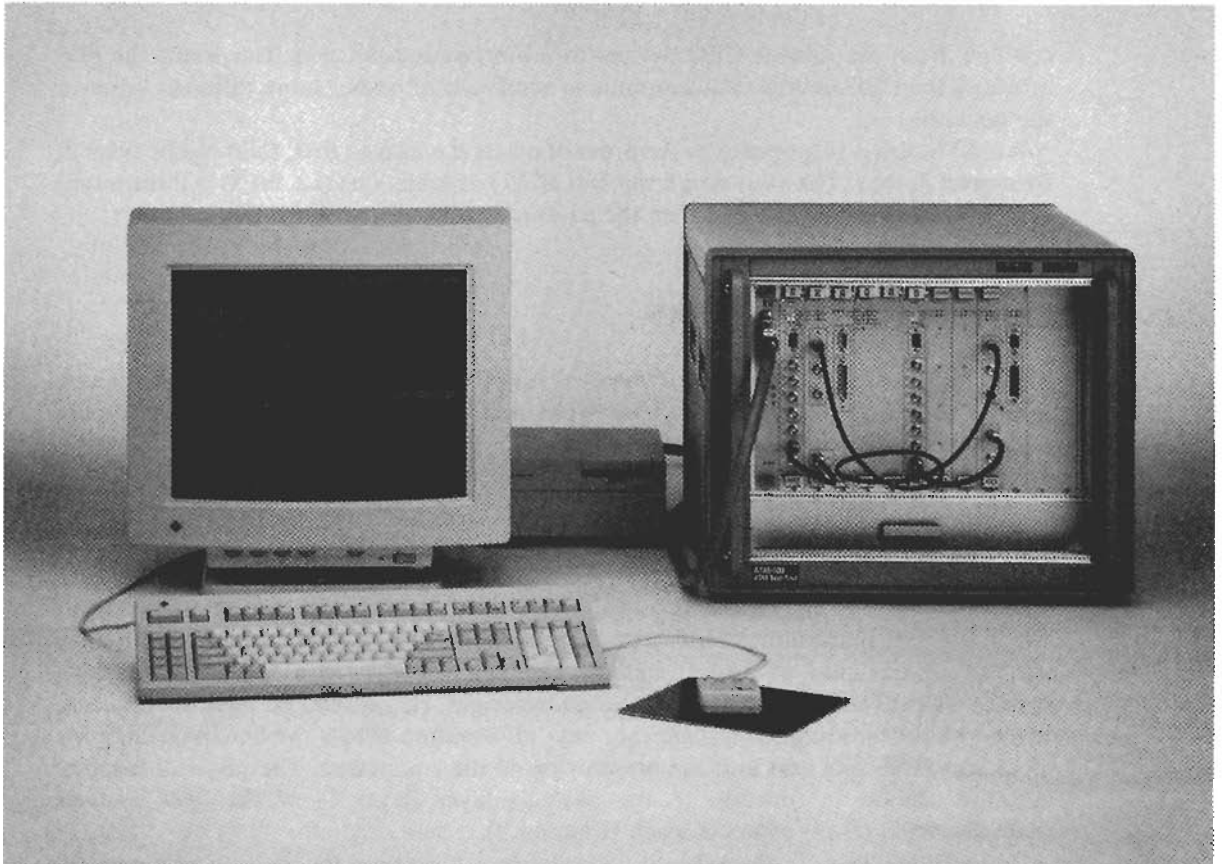
The foundation of the various protocol layers (the *protocol stack*, the set of functions which together make information transfer possible) is the physical medium used for the carriage of electrical or optical signals. The *physical layer* is a specification which defines what electrical or optical signals and voltages, etc., should be used. In addition, it sets out a procedure for transferring data information across the line, providing for clocking of the bits sent and the monitoring of the equipment. The physical layer of ATM is similar in function to the physical layer (layer 1) of the *open systems interconnection (OSI)* protocol stack (Chapter 9).

The physical layer is divided into two sublayers. These are the *physical medium* sublayer and the *transmission convergence (TC)* sublayer. The physical medium sublayer defines the exact electrical and optical interface, the line code and the bit timing. The TC sublayer provides for framing of cells, for cell delineation, for cell rate adaption to



the information carriage capacity of the line, and for operational monitoring of the various line components (*regenerator section (RS)*, *digital section (DS)* or *transmission path (TP)*). Preferred physical media defined for use with ATM include optical fibre and coaxial cable. There is also some scope to use twisted pair cable.

It is the ATM layer which controls the transport of cells across an ATM network, setting up *virtual channel connections* and controlling the submission rate (generic flow control) of cells from user equipment. The *service* provided to the ATM layer by the physical layer is the physical transport of a valid flow of cells. This is 'delivered' at a conceptual 'point' called the physical layer *service access point (PL-SAP)*. The flow of cells is correctly called a *service data unit (SDU)*, in fact the *PL-SDU (physical layer service data unit)*.



**Figure 26.16** Maintenance test tool for ATM (*Courtesy of Siemens AG*). The operation and maintenance (OAM) cells of ATM provide for advanced measurement of network performance without affecting live connections.

The ATM layer controls the service provided to it by the physical layer by means of *service primitive* commands. These are standardized requests and commands exchanged between the control function within the ATM layer (in the jargon called the *ATM layer entity* (*ATM-LE*)) and the *physical layer entity* (*PL-LE*). They allow, for example, a particular ATM-LE to request transfer of a flow of cells (service data unit). Conversely, the physical layer may wish temporarily to halt the transfer of cells to it by the ATM layer because of a problem with the physical medium.

The *transmission convergence sublayer* receives information in the form of cells provided to it by the ATM layer. This is the PL-SDU, or more specifically, the TC-SDU. These cells are supplemented by further information, including PL-cells (physical layer cells) and OAM cells (operations and maintenance cells). The extra information, an example of *protocol control information* (*PCI*) turns the TC-SDU into a *TC-PDU* (*protocol data unit*). It ensures the correct transmission of information across the physical medium. The *TC-PDU* is passed to the physical medium sublayer, where it is called the *PM-SDU* (*physical medium service data unit*).

Finally, the *PM-SDU* is converted to the *PM-PDU* by addition of further *PCI* and is passed to the medium itself. The form of the PM-PDU (and thus the conversion performed by the physical medium sublayer) is dependent upon the type of medium used (e.g. electrical, optical, etc.). To accommodate a change of the *physical medium*, only the *physical medium sublayer* need be swapped. Other hardware and software components (e.g. corresponding to the ATM layer) can be re-used.

Together the ATM layer, the physical layer and the physical medium itself are called an *ATM transport network*. An *ATM transport network* is capable of conveying information in the form of cells between network end-points. However, so that the information content carried by an ATM transport network can be correctly interpreted by the receiver, there are further higher layer protocols defined. The most important of these is the *ATM adaptation layer* (*AAL*).

## 26.16 CAPABILITY OF THE ATM ADAPTATION LAYER (AAL)

As the name suggests, the *ATM adaptation layer* (*AAL*) provides for the conversion of the *higher layer information* into a format suitable for transport by an ATM transport network. The *higher layers* are information, devices or functions of unspecific type which require to communicate across the ATM network. *Higher layer information* carried by the ATM network may be either

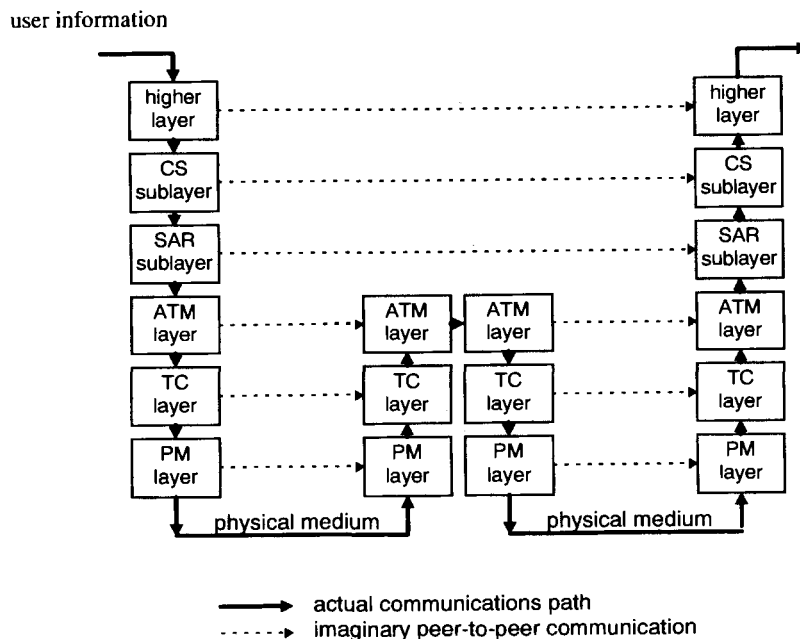
- user information (*user plane*) of one of a number of different forms (e.g. voice, data, video, etc.) as categorized by the AAL service classes (Table 26.1)
- control information (*control plane*) for setting up or clearing connections
- network management information (*management plane*) for monitoring and configuring network elements or for sending requests between network management staff.

Like the other layers, the AAL accepts *AAL-SDUs* from the higher layers and passes an *AAL-PDU* to the layer below it (the ATM layer), where it is known as an ATM-SDU. However, unlike the ATM and physical layers a number of different alternative services

can be made available to the *higher layers* above the AAL, thus allowing different types of information to be *adapted* for carriage across a common ATM transport network. It is the ATM adaptation layer which gives ATM networks their capability to transfer all sorts of different information types. It is split into two sublayers: the *convergence sublayer*, *CS* (where the alignment of the various information types into a common format takes place and division into cells occurs); and the *segmentation and reassembly sublayer*, *SAR* (where the cells are numbered sequentially to allow reconstruction in the right order at the receiving end).

## 26.17 PROTOCOL STACK WHEN COMMUNICATING VIA AN ATM TRANSPORT NETWORK

Figure 26.17 illustrates the peer-to-peer communications which take place when two user end devices communicate with one another by means of an ATM transport switch. The ATM switch supports only the lowest three protocol layers, and ‘speaks’ peer-to-peer with each of the ends, translating protocol information such as VPIs and VCIs as necessary and relaying user information. Meanwhile, at the *ATM adaptation layer (AAL)* and the higher layers, the two end devices communicate peer-to-peer directly over the connection established by the lower three layers. This information remains uninterpreted and passes ‘transparently’ through the network.



**Figure 26.17** Protocol layer representation of two end devices communicating via ATM layer switch

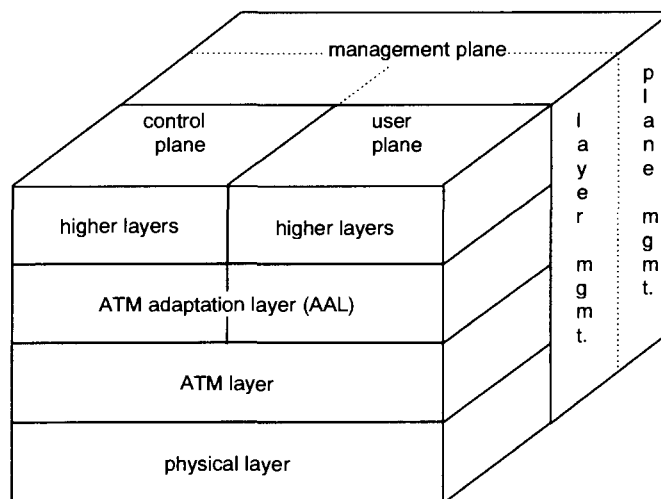
If we were to monitor the wire between either of the end devices and the ATM switch of Figure 26.17 then we would observe communication at each of the layers. What we actually observe are cells, but structured a little like a Russian doll. The smallest doll (right inside) is the information that we want to carry between the users (the *higher layer information*). All the other dolls are the protocol information (PCI), one doll for each of the lower layers, each providing a function critical to the reliable carriage and correct interpretation of the message.

## 26.18 ATM PROTOCOL REFERENCE MODEL (PRM)

Strictly, Figure 26.17 illustrates only the protocol stack for the *user plane* (i.e. for the transfer of information between end devices once the connection has been established). The AAL protocols used on the *control* and *management* planes will usually differ from the AAL protocols used on the *user plane* of the same connection, though identical protocols will be used at the ATM transport layers. This is illustrated schematically in the *ATM protocol reference model (PRM)* as shown in Figure 26.18.

In the case of the *control* and *management* planes, the network itself must interpret the higher layer information and react to it. The *control plane* AALs (for the user-network signalling in setting up a switched virtual circuit, SVC) will typically need to be suited for data information transfer. Similar protocols will also be necessary for the *management plane*. In contrast, in the case of communication across the *user plane*, the higher layer information may take any number of different forms (speech, data, etc.), but the individual network elements themselves (e.g. switches) may be incapable of recognizing and interpreting these various forms.

In real switches and ATM end user devices, common or duplicate hardware and software may thus be used for management, control and user planes at ATM and physical

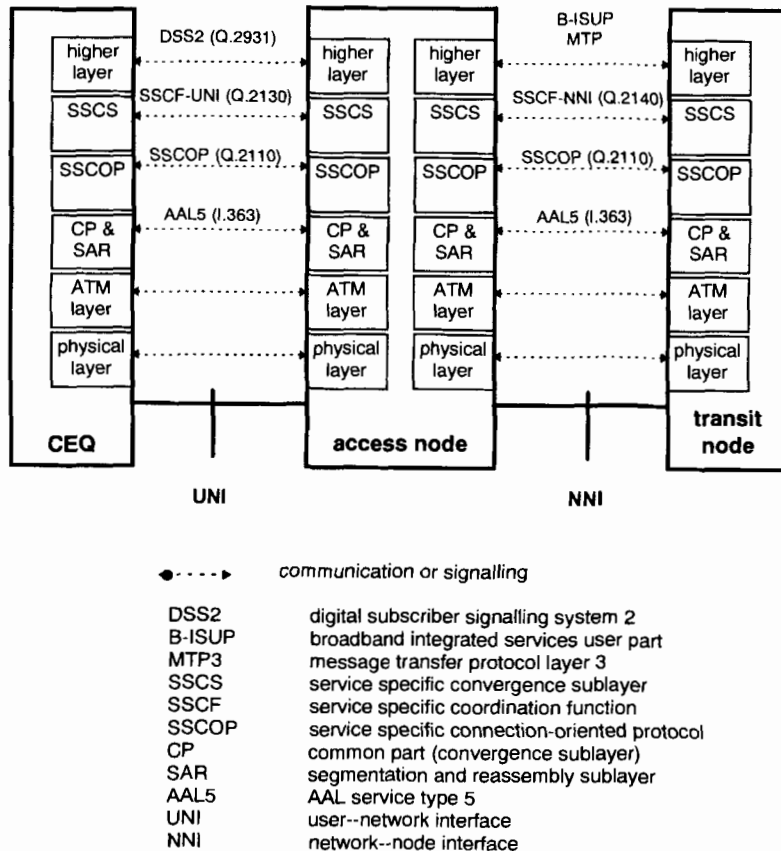


**Figure 26.18** The B-ISDN protocol reference model (*Courtesy of ITU*)

layers, but distinct hardware and software will be necessary for signalling and user information transfer at the AAL and higher layers of these *planes*.

The different types of *user*, *control* and *management* services are carried by the AAL by means of *service-specific convergence services*. Examples of specific *user plane* services offered by the *ATM adaptation layer (AAL)* are

- frame relay *SSCS (service specific convergence sublayer)* service
- SMDS (*switched multimegabit data service*) *SSCS* service
- reliable data delivery *SSCS* service (a packet-network like data network service)
- LAN emulation *SSCS* service
- desktop quality video *SSCS* service
- entertainment quality video *SSCS* service
- further services still in the stage of development by ATM forum



**Figure 26.19** UNI and NNI protocols used on the control plane

There are a set of service specific protocols used on the *control plane*. These are known as the *signalling AAL (SAAL)* and are defined by ITU-T recommendations Q.2110, Q.2130 and Q.2140. They operate in association with AAL5. The SAAL in turn carries either *UNI signalling* or *NNI signalling*. *UNI signalling* is a combination of *MTP3 (message transfer protocol layer 3)* and *DSS2 (digital subscriber signalling system version 2)*. *NNI signalling* instead uses *B-ISUP (broadband integrated services user part)* and *MTP3*. Figure 26.19 illustrates them.

The higher layer protocols of the control plane (DSS2 and B-ISUP/MTP) are equivalent to the DSS and ISUP/MTP signalling protocols of narrowband ISDN. These are the signalling systems used, respectively, between an ISDN telephone and the network exchange (UNI) and between nodes in the network (NNI). As their names suggest DSS2 and B-ISUP are based heavily upon their narrowband ISDN counterparts, DSS1 and ISUP (Chapters 10 and 12) and as we shall see in Chapter 28, the numbering plan used in B-ISDN is based on that used in modern digital telephone and ISDN networks (ITU-T E.164).

## 26.19 ATM FORUM NETWORK REFERENCE MODEL

The network reference model of ATM forum differs slightly from that of ITU-T (Figure 26.6), distinguishing between the *private* and *public* parts of an ATM network. Thus the model caters for the connection of *private* corporate and campus ATM networks to a *public* ATM network.

The ATM forum network defines the *public UNI* (also called the *public network interface*) and the *private UNI* (the *private local interface*). In addition, there are two types of the NNI, the *PNNI (private network–node interface or private network–network interface)* and *B-ICI (broadband inter-carrier interface)*. At a basic level, these both use the NNI. They differ in the types of network management and administration possible over them.

As we see in Figure 26.20, ATM forum depicts the interfaces UNI and NNI between sub-networks and not between individual nodes. This reflects the philosophy particularly of data network equipment manufacturers who have traditionally used ‘proprietary’

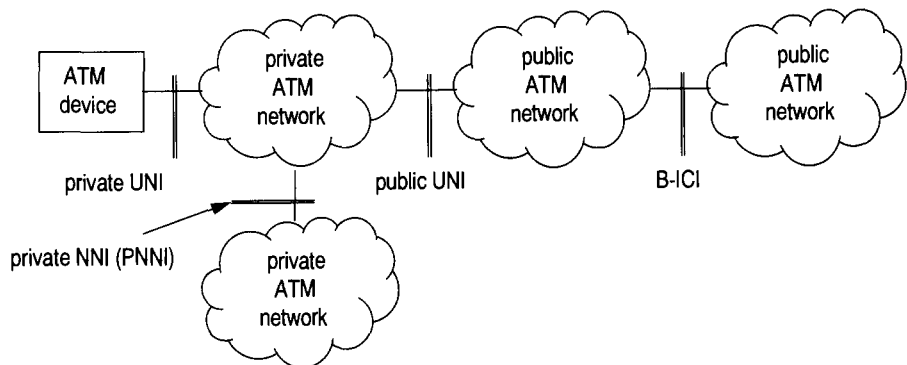


Figure 26.20 ATM forum network reference model. (Copyright The ATM Forum 1995)

interfaces between the nodes within a sub-network to enable them to differentiate their products from those of their competitors by offering improved network management and service features.

The ATM forum UNI is the best specified of all interfaces available. Three versions are currently available in most ATM switches. These are the versions 3.0 (UNI v3.0), 3.1 (UNI v3.1) and 4.0 (UNI v4.0). Version 3.1 incorporates the ITU-T recommendation Q.2931 (DSS2) for call set-up of *switched virtual circuits (SVCs)*. Version 3.0 is an earlier version, and in practical terms only supports *permanent virtual connections (PVCs)*. Version 4.0 meanwhile is the latest version, in which AAL types 2 and 3/4 have been eliminated by concentration on types AAL1 and AAL5. The specifications are published by Prentice Hall.

PNNI includes a number of functions for *discovery* of the network topology and for routing through it. This is well suited, for example, to a university or large campus private ATM network in which individual departments may be responsible for adding new switches and end-user devices to the network on their own initiative. The PNNI *topology state* functions enable other switch nodes within the network to keep abreast of the topology and connected devices.

B-ICI defines a more secure interface than that of PNNI, reinforcing the organizational boundary between different public network operators. The B-ICI interface adds specific functions to the standard NNI to allow for *contracting traffic*, and for monitoring and management of an *interconnect* between different operators networks (e.g. in the USA the *interconnect* between *LECs (local exchange carriers)* and *IECs (inter-exchange carriers)*). The B-ICI interface supports the capabilities necessary for transiting networks (like IEC networks), and allows for the support of *cell relay service (CRS)*, *frame relay service (FRS)* and *SMDS (switched multimegabit data service)*. The B-ICI interface is thus an important precursor to the regulated interconnection of public ATM networks, at least in the USA.

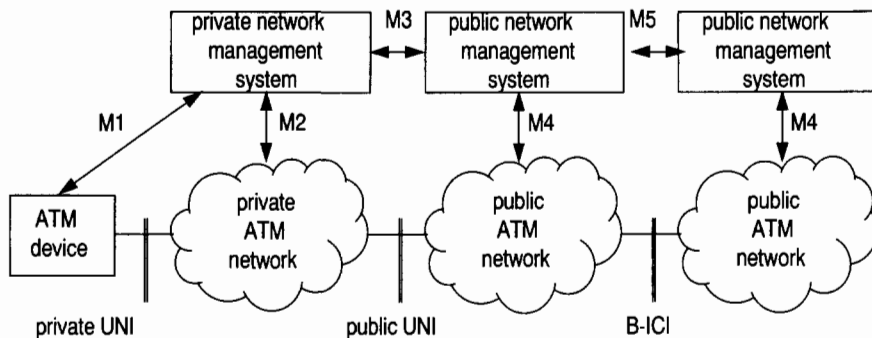
ATM forum has also defined three *interim* interfaces for use at the PNNI, B-ICI and for management control of the network. These are, respectively, the

- *IISP (interim inter-switch signalling protocol, a forerunner of the PNNI signalling protocol based on ATM forum UNI v3.1)*
- *BISSI (broadband inter-switching system interface, a forerunner of the NNI signalling protocol for use in and between public ATM networks)*
- *ILMI (interim local management interface, a protocol for management of ATM network devices based on the Internet simple network management protocol, SNMP)*

ITU-T has not recognized these interim standards.

## 26.20 ATM FORUM NETWORK MANAGEMENT MODEL

The ATM forum network management model (Figure 26.21) also differs from that of ITU-T (Figure 26.8) by the strict separation of public and private networks. Confusingly, ATM forum refers to M-interfaces rather than to *management plane* types, and



**Figure 26.21** ATM forum network management reference model (Copyright The ATM Forum 1995)

uses different numbering (ATM forum interface M1 is equivalent to ITU-T management plane type 1, but ATM forum interface M3 is not equivalent to ITU-T's management plane 3).

ATM forum's M3- and M4-interfaces are already extensively defined by ATM forum, respectively, for *customer network management (CNM) of public ATM network service* and for network management of public ATM networks. The other interfaces are less well defined.





## PART 5

---

# **RUNNING A NETWORK**

---



---

# *Telecommunications Management Network (TMN)*

---

The goal of the ‘telecommunications management network’ or ‘TMN’ is to provide for consistent and efficient management of complex telecommunications networks. The TMN model describes the basic operating and management functions which a network operator has to conduct and the standard interfaces to be used between network components and network management systems. Key elements of the TMN concept are the management model, the Q3-interface and the CMIP (common management information protocol). We discuss them all in this chapter, which in parallel gives an insight into the problems of managing complex networks. At the end, we also briefly discuss the SNMP (simple network management protocol), which is used as an alternative to CMIP in many corporate and Internet/router networks.

## **27.1 THE PROBLEMS OF MANAGING NETWORKS**

Figure 27.1 illustrates a typical telecommunications network of today. It provides a valuable insight into the complexity of managing even relatively simple telecommunications networks. The figure shows a simple data network composed of four different component types, but each provided by different manufacturers. They each have independent *network management systems (NMS)*. Each NMS is capable of monitoring and controlling every aspect its manufacturer’s network equipment, but is incapable of monitoring or controlling other manufacturers’ devices. The component types shown are

- data switches, managed by a ‘proprietary’ network management system, NMS1
- SDH (synchronous digital hierarchy) transmission technology (links some of the switches and is managed by NMS2)
- TDM (time division multiplex) transmission technology (links some of the switches and is managed by NMS3)

- direct leaselines provide for remaining switch links. These are supplied by another, public telecommunications network operator, for which there is no management system available to our operator

A simple line failure is illustrated in Figure 27.1. An SDH connection has failed. The result is a plethora of alarms generated by both NMS1 and NMS2. The network status screens of both NMS are likely to light up like Christmas trees in all sorts of colours.

The SDH network management system, NMS2, will pinpoint the root cause of the problem, probably with the appropriate link alarm indicated in red. Meanwhile, NMS1 is also showing red alarms, because one of the adjoining switches has a malfunctioning port and the other adjoining switch is isolated. Other switches may report 'yellow alarms' as connections to the isolated switch have been lost. The exact cause of the problem, however, will not be clear from the information that appears on NMS1. Unfortunately, this is likely to lead to a certain amount of confusion and wasted effort...

...The human operator managing the SDH network naturally starts about his repair... meanwhile...

...His colleague, the switch network manager watching the screen of NMS1 is unaware of the root cause. Instead he addresses what appears to him to be the urgent problem of the isolated switch. Perhaps there has been a power failure? Maybe the switch needs to be restarted? ... He calls his colleague at the isolated switch site. Having done so, he is able to rule out a local cause. So next he calls his SDH colleague and, of course, discovers the most likely cause of his own alarms. He hands over responsibility and starts to get on with something else.

But, you might ask, why did the switch manager not call the SDH manager to start with? The answer is that he is unable to determine easily the single cause of a lot of alarms. He must therefore address each alarm in turn, diagnosing the most critical alarms first.

Considerable effort may go into checking each alarm and concluding a single root cause. Only afterwards can the human network manager truly rest in peace as his SDH

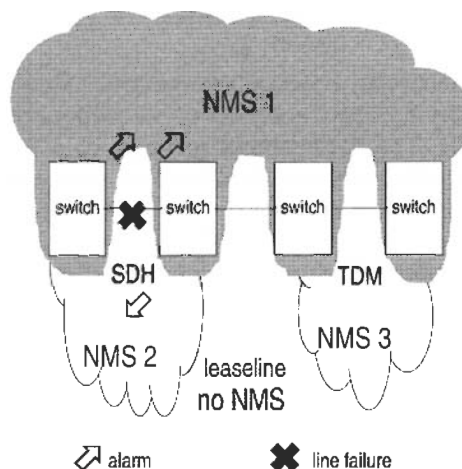
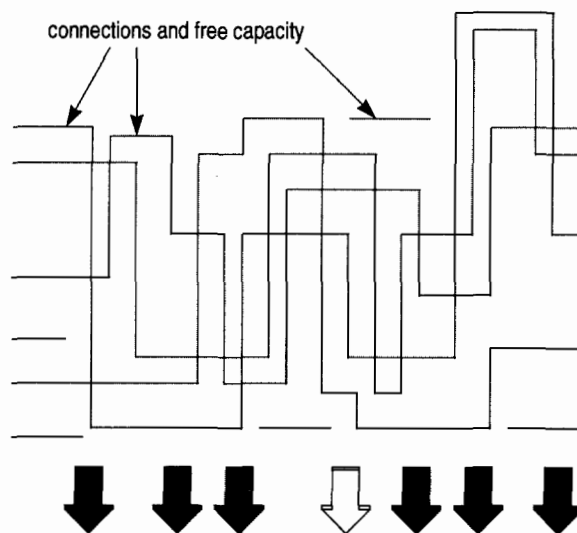


Figure 27.1 A typical telecommunications network



**Figure 27.2** Recognizing that a network is like spaghetti and the alarms are like Christmas tree lights

colleague gets on with the job of network *restoration*. Afterwards, despite his clear conscience, the alarms on his own screen do not go away... So when another fault arises before the first is cleared, the diagnosis becomes more confused and difficult as the alarms inter-mingle.

In real networks, many simultaneous faults are likely to be present at any point in time, even if they are of a non-critical nature. Working out which alarm belongs to which bit of network equipment is a complex task, like finding the knot in spaghetti! (Figure 27.2).

## 27.2 NETWORK PROVISIONING

Provisioning of new lines in the network is just as complex and labour-intensive as fault finding. Imagine trying to set up a single permanent connection (say a frame relay PVC) across the networks of either Figure 27.1 or Figure 27.2. First someone must sit down to 'design' the connection (work out the path), allocate ports and network addresses. Next, each of the TDM and telephone switch ports need to be programmed. The SDH links must be configured. The leaselines must be ordered. Finally comes the hope that all the installation staff in the chain do their jobs on-time and accurately, otherwise the connection fails an end-to-end acceptance check. It would be preferable to define the desired end-points of the connection and let a computer coordinate and execute the rest. Enormous saving in effort is possible, together with higher quality and faster installation work.

... And by being able quickly to trace the complete path or trail of a single connection through the network (e.g. Figure 27.2) using sophisticated computer network management tools, operations staff can more quickly diagnose faults and resolve

problems if they arise later. Furthermore, by computer 'filtering' of the alarms, it may be possible to reduce the 'Christmas tree lights' to the one or two alarms which are most likely to be the direct cause of a problem. This would greatly help our fault-correcting friends in Figure 27.1.

### 27.3 UMBRELLA NETWORK MANAGEMENT SYSTEMS

A number of network operators, computer manufacturers and software developers started during the 1980s to develop the idea of *umbrella network management* systems (Figure 27.3). These were to be powerful network management systems capable of controlling the network as a whole. They would be able to correlate and present the various pieces of information from individual *network element managers* (NEM) thus removing this arduous coordination task from human operators. (Note: NMS1, NMS2 and NMS3 of Figure 27.1 are *network element managers* because they are network management systems capable only of managing a particular type of *network element*.)

A single command issued to the umbrella system by a human operator would be all that is needed to configure a new trunk between any two of the switches of Figure 27.1. The umbrella system managers see to it that the plethora of commands needed to be issued to the individual network elements are generated (one command to configure a new port at one switch, a second command for the second switch, a third command to the SDH NMS to establish the transmission path, etc.).

For the umbrella network management system to work properly, a defined, standard interface between the umbrella network management system and each of the *network element managers* is needed. Over this interface the umbrella network management

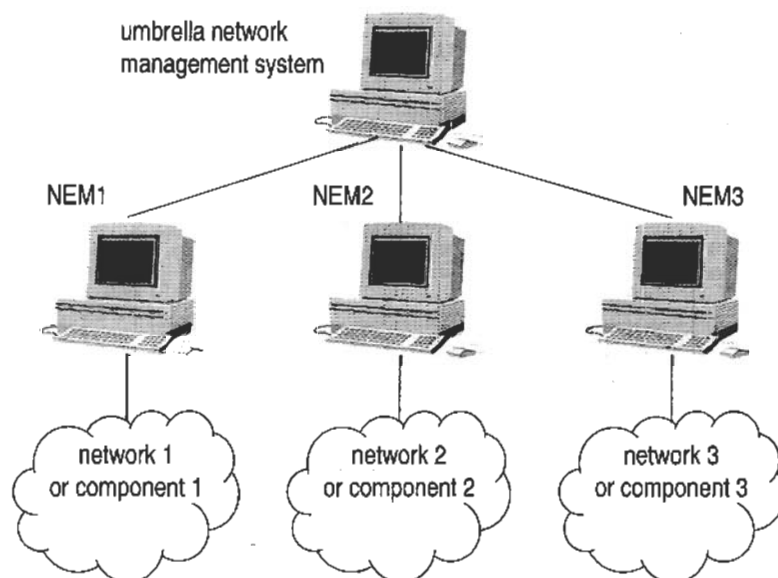


Figure 27.3 The concept of an umbrella network management system

system needs to be capable of receiving information on request about the status of specific network elements and to be able to issue commands for reconfiguration or control of the various sub-networks.

One of the first steps taken by CCITT (now ITU-T) in defining its standards for *telecommunications management network (TMN)* was to define a model of the various functions and interfaces going to make up a complete network management system. This is laid out in ITU-T recommendation M.3010 and is illustrated in Figure 27.4.

The key components of the TMN model (Figure 27.4) are the *operating system (OS)*, the *network element* (actually the *network element manager*) and the *workstation (WS)*. The key interfaces are the *Q3-interface*, the *X-interface* and the *F-interface*. The *data communications network (DCN)* merely provides a means for connecting together network management devices and network components in different locations.

The *operating system (OS)* is analogous to what we previously called an *umbrella network management system*, except that there is no longer any presumption that all the various network operations, administration and management functions (e.g. provisioning, troubleshooting, etc.) can be handled by a single computer system. Instead, a number of *operating systems* combined together will handle the full functionality of our previous *umbrella* system. A standard interface between operating systems (the *X-interface*) allows this subdivision.

Individual *network elements (NE)* probably each have their own proprietary network management systems (*network element managers*), which cater for the specialized functions and operating methods of a particular manufacturers hardware. The *Q3-interface* provides for a standard means of communication between the *operating systems* (i.e. 'umbrella' managers) and individual NEMs. The network element manager acts as an

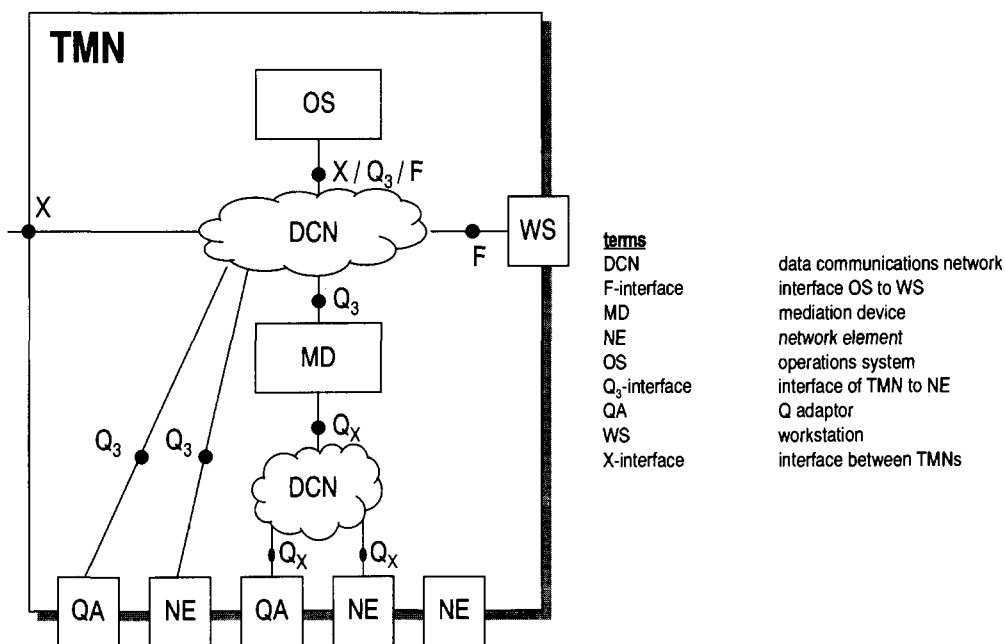
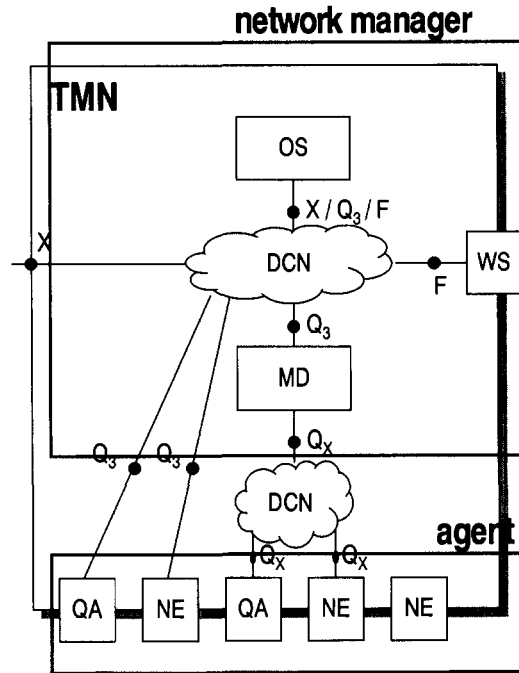


Figure 27.4 TMN model: defined functions and interfaces (courtesy of ITU-T)





**Figure 27.5** Q3- and X-interfaces are the important interfaces between network managers and agents

*agent*, interpreting the commands issued across the *Q3-interface* and reacting upon them in conjunction with the individual *network elements*. Thus status information can be sent to the manager and control commands can be executed in the network by the *agent*.

When necessary, a translation function can be used to convert *Q3-interface* commands into the proprietary format needed by a specific network element. This translation function is termed the *Q-adaptor function (QA)*. Where the individual network elements can respond directly to *Q3* commands, it is not necessary.

The term *mediation device* is applied to any device carrying out a conversion of protocol or data format necessary for the interconnection of devices. The diagram should not be taken to imply that a single *mediation device* will cater for all necessary mediation needs.

The *work station (WS)* is equivalent to a technicians laptop computer. A standard interface (the *F-interface*) allows him to log into the *manager (operating system)* wherever he may be (in the exchange building or in the field).

Figure 27.5 duplicates the TMN model as already shown in Figure 27.4, but now we see more clearly the typical system boundaries of real network management systems. The *manager* is a so-called *umbrella network management system*, capable of consolidating information fed to it by separate proprietary *network element managers (NEM)*. The manager communicates with the various individual network element managers by means of the *Q3-interface*, interpreting the standardized Q3 enquiries and commands and translating them as necessary into the proprietary format of the particular network component.

## 27.4 THE Q3-INTERFACE, THE COMMON MANAGEMENT INFORMATION PROTOCOL (CMIP) AND THE CONCEPT OF MANAGED OBJECTS (MO)

Crucial for the successful communication between manager and agent over the Q3-interface are

- the definition of a standard protocol (set of rules) for communication (this is the *common management information protocol, CMIP*)
- the definition of standardized network status information and control messages for particular standardized types of network components (so-called *managed objects*)

*CMIP* (*common management information protocol*) delivers the *common management information service (CMIS)* to the *operating system* of Figure 27.4. CMIS is the service allowing a *CMISE* (*common management information service entity*, i.e. a software function) in the manager to communicate status information and network control commands with a CMISE in each of the various agents. CMIP itself is an OSI layer 7 protocol, which sets out the rules by which the information and commands (CMISE services) may be conveyed from *manager* to *agent* or vice-versa. These basic CMISE services are restricted in number. They are listed in Table 27.1.

However, alone, CMIS and CMIP do not convey any useful information from manager to agent. They are simply the ‘rules of orderly conversation’. In the same way that the project manager of a major building project can make sure that instructions are

**Table 27.1** The basic CMISE (common management information service entity) services

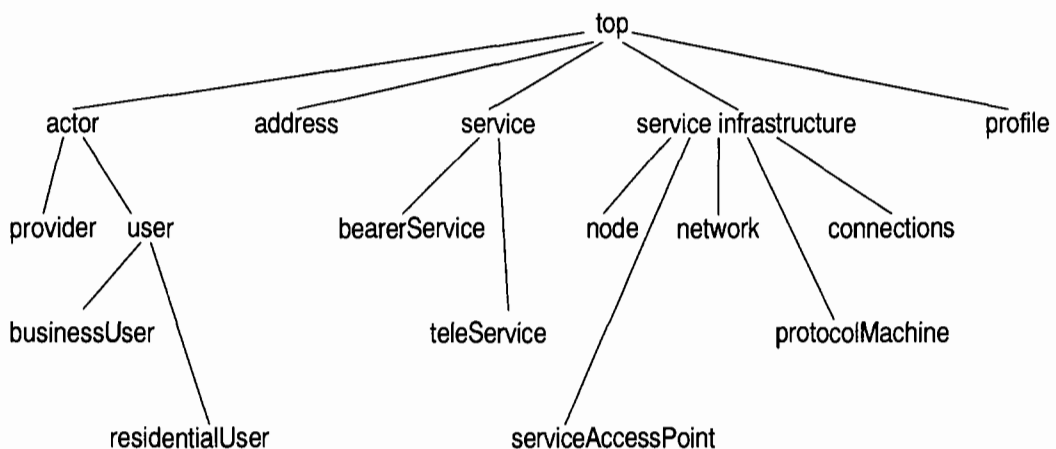
Service name	Function
M-ACTION	This service requests an action to be performed on a managed object, defining any conditional statements which must be fulfilled before committing to action
M-CANCEL-GET	This service is used to request cancellation of an outstanding previously requested M-GET command
M-CREATE	This service is used to create a new instance of a managed object (e.g. a new port now being installed)
M-DELETE	This service is used to delete an instance of a managed object (e.g. a port being taken out of service)
M-EVENT-REPORT	This service reports an event, including managed object type, event type, event time, event information, event reply and any errors
M-GET	This service requests status information (attribute values) of a given managed object. Certain filters may be set to limit the scope of the reply
M-SET	This service requests the modification of a managed object attribute value (i.e. is a command for parameter reconfiguration)

issued to the electrician, the brick layer and the plumber and that each of the craftsmen understands his instructions and completes them on time, so CMIS is able to 'project manage' the task of managing a network. The actual content of the network management 'work instructions' depends on the particular network component being managed, and is coded according to the appropriate *management information base (MIB)* of *managed objects*.

A *managed object* is a standardized definition of a particular type of network component, and the normalized states in which it can exist. Imagining a water bottle to be described as a standard *managed object* it might be 'a vessel for storing one litre of liquid' (exactly what shape it is, is unimportant). Its top might be a screw top, a cap-top or a cork, but as a *managed object* all three are simply 'watertight stoppers' which may either be 'secured' or 'opened'. Standardized states may thus be 'full', 'empty', 'half full', etc., while control commands might be 'fill up', 'pour out', 'secure stopper', 'undo stopper', etc. Provided that all manufacturers of water bottles produced products able to respond to these commands you can buy whichever device is most aesthetically pleasing without the concern of not being able to get the lid off!

Managed objects are nowadays defined for most new computer and telecommunications products. They may be defined either on an industry standard basis or by the individual manufacturer, where industry standards do not exist. They are usually grouped into sets corresponding to individual device types or network components. These are called *management information bases (MIBs)*. Thus there is an *MIB* for *routers* used on the *Internet*. There is also an *MIB* for the ISDN basic rate interface (Chapter 10). Of course there are many other MIBs, but many more will need to be defined.

Having well-defined MIBs for each of the network components is key the realization of the dream of a complete *umbrella manager*, capable of coordinating all network components. It is thus the MIBs which provide the critical information content of the messages. It is, after all, of no use to tell someone to 'act on' or 'carry out' something unless you also tell him you are talking about the 'water bottle' and he is to 'open it').



**Figure 27.6** A simple managed object inheritance hierarchy

The definition of *managed objects (MO)* and *managed information bases (MIBs)*, like other OSI layer 7 protocol objects, is carried out in the *abstract syntax notation 1 (ASN.1)* about which we learned in Chapter 9. The procedure of definition, including advice on how to group and subdivide MOs is defined in ITU-T recommendation X.722. These are the *guidelines for the definition of managed objects (GDMO)*. They include strict rules about the naming and spelling conventions and the hierarchical structure.

There are a number of commercial software products available to help with the task (so-called *GDMO browsers*). Figure 27.6 illustrates a relatively simple *managed object inheritance hierarchy*, showing the various states and attributes of a given managed object (in this case a network service).

## 27.5 THE ISO MANAGEMENT MODEL

*ISO (International Organization for Standardization)* has developed a standard model classifying the functional processes which must be undertaken in the management of computer and telecommunications networks. All management tasks are classified into one of the following five (FCAPS) categories

- Fault management
- Configuration management
- Accounting management
- Performance management, and
- Security management

This model is referred to widely by ITU-T's recommendations on a *telecommunications management network*, and is reproduced in the ITU-T X.700-series recommendations.

*Fault management* is the logging of reported problems, the diagnosis of both short and long term problems, 'blackspot' analysis and the correction of faults.

*Configuration management* is the maintenance of network topology information, routing tables, numbering, addressing and other network documentation and the coordination of network configuration changes.

*Accounting management* is the collection and processing of network usage information as necessary for the correct billing and invoicing of customers and the settlement with operators of interconnected networks.

*Performance management* is the job of managing the traffic flows in a live network, monitoring the traffic flows on individual links against critical threshold values and extending the capacity of links or changing the topology by adding new links. It is the task of ensuring performance meets design objective (e.g. *service level agreement*).

*Security management* functions include

- identification, validation and authorization of network users and TMN system users
- security of data

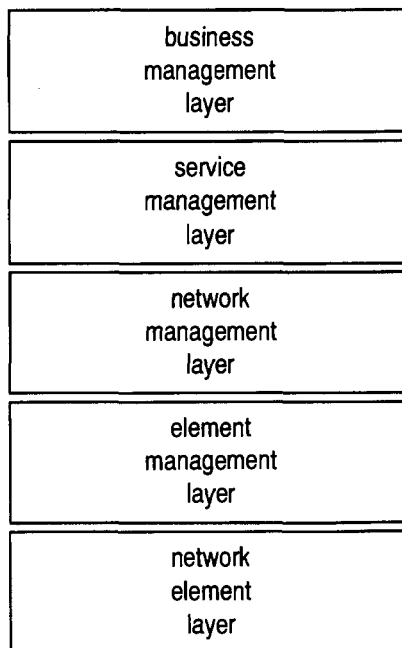
- confirmation of requested network management actions, particularly commands creating or likely to create major network or service upheaval
- logging of network management actions (for recovery purposes when necessary, and also for fault auditing)
- maintenance of data consistency using strict *change management*

## 27.6 TMN MANAGEMENT FUNCTION MODEL

Figure 27.7 illustrates the five layers of functionality defined by the OSI management model for a TMN. The layers are intended to help in the clear and rational definition of TMN *operating system* boundaries, so simplifying the definition, design and realisation of systems, by simplifying their data and communication relationships to one another.

At the lowest functional layer of Figure 27.7 (*network element layer, NEL*) are the *network elements* themselves. These are the active components making up the networks. Above them, in the second layer of the hierarchy, is the *element management layer (EML)* containing *element managers*. *Element managers* are devices which control single network components or sub-networks (e.g. a local management terminal or a proprietary *network management system*).

At the third layer, the *network management layer (NML)* are *managers* which control the main aspects of a given network type (e.g. ISDN or ATM).



**Figure 27.7** The functional layers of TMN

The *service management layer (SML)* contains *service managers* which monitor and control all aspects of a given service, on a network independent basis. Thus, for example, it is possible in theory to conceive a *frame relay service* provided over three different network types: *packet switched*-type network, ISDN and ATM. The service manager ensures installation of a given customers connection line needs on the appropriate network(s) and monitors the service delivered against a contracted frame relay *service level agreement*.

The *business management layer (BML)* contains functionality necessary for the management of a network operator's company or organization as a whole. Thus purchasing, bookkeeping and executive reporting and information systems are all resident in this layer.

## 27.7 THE NETWORK MANAGEMENT FORUM (NMF), OMNIPOINT AND SPIRIT

In addition to the standardization activities of ISO and ITU-T, a number of industry initiatives have been commenced over the last six years to speed progress in the development of umbrella network management systems.

The first major initiative, started in 1989, was the *Network Management Forum (NMF)*. This is composed of major network operators (e.g. British Telecom) and computer manufacturers (e.g. IBM, DEC, SUN, Hewlett Packard), who have together formed a common interest group to speed the realization and acceptance of common network management standards. Valuable outputs of this group have been the *OMNIpoint* and *SPIRIT* specifications.

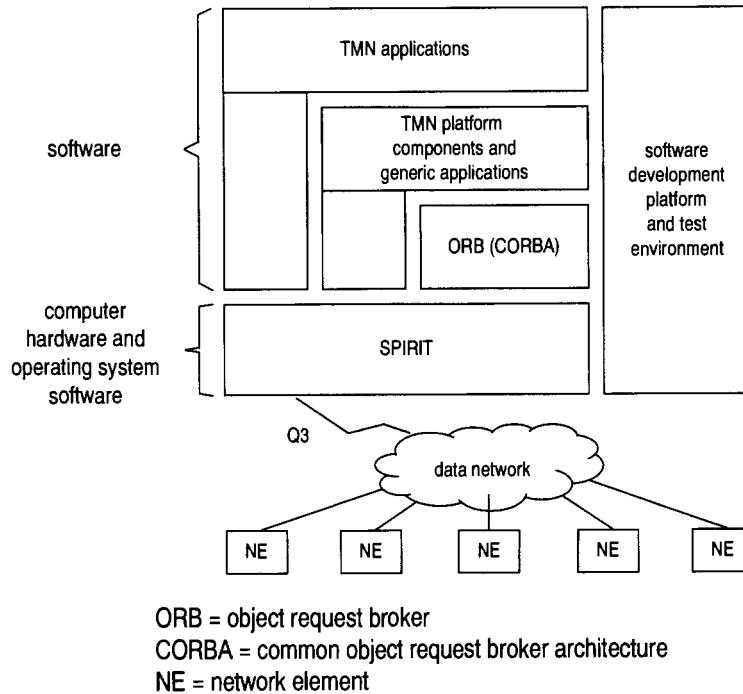
The *OMNIpoint* specifications are intended as a set of specifications and guidelines for implementors of network management systems. They attempt to review standards issued by main standards bodies and build inputs from various sources into a consolidated (and realisable) whole. Where standards are vague, NMF in its *OMNIpoint* specifications aims to provide further clarifying specifications.

A number of hardware and software manufacturers claim already to have *OMNIpoint*-compliant systems. These, in the main, are systems compliant with *OMNIpoint1*. Unfortunately, compliance to OMNIpoint1 does not mean much. Compatibility of compliant systems is not guaranteed. More interesting is *OMNIpoint2*, which has recently (November 1995) been published by NMF. This aims to be a fuller set of specifications for TMN-compliant systems.

A second valuable output of NMF has been the *SPIRIT (service providers integrated requirements for information technology)* specifications. These attempt to lay out a standard hardware and computer operating system software platform for the realization of development and run-time platforms for TMN-compliant network management systems. SPIRIT 2.0 was issued in November 1994.

## 27.8 REALIZATION OF A TMN

How does all this fit together? Figure 27.8 shows a possible layered software configuration of a TMN system, based on a client/server computer hardware platform



**Figure 27.8** Possible computer hardware and software realization of TMN

according to SPIRIT, with the latest object-oriented computing software (CORBA) loaded upon it. Above this, there are some generic tools and processes, thereby limiting the amount of software which has to be written for specific services or network operators. The layered approach leads to more rapid development, with greater sharing of development costs and benefits.

## 27.9 EXAMPLE OF EARLY TMN REALIZATION

The operations and maintenance functions defined for ATM networks are designed according to the principles of the *telecommunications management network (TMN)*. Specifically, the ATM standards define the following management functions

- performance monitoring
- defect and failure detection (by continuous monitoring and generation of alarms)
- network and system reconfiguration (including automatic changeover to backup facilities)
- fault localization

In addition to TMN interfaces, some early ATM equipment is likely to offer the *SNMP* (*simple network management protocol*), because this is the basis of the ATM *interim local management interface (ILMI)*. ILMI enables the management of *UNI managed objects* (switches, devices, parameter settings, etc.), thus making management control of devices possible across the ATM UNI. Together the *managed objects* make up the *UNI management information base (MIB)*.

The short term attraction of SNMP as the basis for ILMI lies in its widespread deployment in existing corporate data networks, routers, LANs (local area networks) and *Internet* components.

Future standardization work on ATM network management is likely to concentrate on migration to the *CMIP* protocol (probably by integration of certain SNMP procedures into it). In addition, a huge amount of work must be applied to the definition of the standard *MIB (management information base)* of *managed objects* relevant to ATM networks and devices. Without these definitions, it will be impossible to develop the necessary tools for management of the network as a whole. At the moment these are poorly defined.

## 27.10 SIMPLE NETWORK MANAGEMENT PROTOCOL (SNMP)

The *simple network management protocol (SNMP)* is a protocol in many ways similar to TMN's *common management information protocol (CMIP)*. It also allows for creation of umbrella network management systems by allowing for communication of status information and commands (i.e. *management information base, MIB* information) by means of an industry-wide accepted protocol. The main technical difference between SNMP and CMIP is its simplicity. SNMP was developed by the *Internet Engineering Task Force (IETF)* and is defined in its specification RFC 1157.

Although ITU-T was developing CMIP as a robust protocol, conforming with all the requirements of the OSI model, the engineers involved in developing and installing routers and other networks for the *Internet* and other corporate data networks were keen to have a short term solution. SNMP was the answer. It is widely deployed in corporate data networks, where the term *SNMP manager* is applied to several types of available devices capable of acting as basic *umbrella* network managers.

A new version of SNMP, SNMP2 has recently been agreed. This is significantly more complex than SNMP. Meanwhile, SNMP and CMIP protocols are beginning to converge as the work on defining common *management information bases* proceeds.

## 27.11 SUMMARY OF TMN BENEFITS

Assimilation of the various functions and interfaces of the TMN architecture into the network management plans of network operating organizations has a number of benefits

- the opportunity to develop high quality network management methods and processes supported by industry-standard products



- the scope to force suppliers of network switches and other components to develop new functionality crucial to effective network management and TMN as a whole
- in particular, the Q3-interface and CMIP (common management information protocol) are already widely adopted standards

## 27.12 TELECOMMUNICATIONS INTELLIGENT NETWORK ARCHITECTURE (TINA)

Finally, the *telecommunications intelligent network architecture (TINA)* should be mentioned before leaving the subject of TMN, because there are some areas of overlap. *TINA* is a relatively new initiative, based upon the observation that both *intelligent networks* (Chapter 11) and *TMN* are addressing the problems of management of complex networks. *TINA* is an attempt to combine TMN and intelligent network ideas and to re-define the ideal network architecture of future networks.

---

## *Network Routing, Interconnection and Interworking*

---

The control of the routing of calls and connections (so-called ‘traffic’) across telecommunications networks is the most difficult but most important responsibility of a network operator. Only by careful planning and management of appropriate call and traffic routing plans can the network operator ensure successful connection of calls and the efficient use of network resources. In this chapter we discuss the techniques used by network operators in establishing efficient call routing patterns and the special problems caused by network interconnection and interworking, when calls or connections originated in one network have to be passed to another operator’s network for completion.

### **27.1 THE NEED FOR A NETWORK ROUTING PLAN**

We might choose, laudable as it may seem, to attempt to run our telecommunications network by completing as many calls as possible or delivering the greatest proportion of data messages. The problem is that if we attempt to do so, we are bound adversely to affect the intelligibility of messages and the time it takes to deliver them.

To attempt the ‘complete if you can’ routing philosophy, we simply programme the exchanges to route all messages ‘any way possible’, rather than ever fail anything when a circuit is free. Unfortunately, the network will perish of congestion and suffer appalling signal quality. Studying the scenario, however, is highly instructive and we shall look at an example or two shortly.

The rational and rewarding alternative to the ‘any way possible’ regime is to have *network routing plan*, together with a supporting *numbering plan*. The appropriate routing algorithms laid out by the routing and numbering plans are selected to control network traffic and to comply with the overall constraints which the transmission plan and engineering guidelines impose on end-to-end connections made across the network. To work within these various constraints, and still to achieve a network that is reasonably cheap as well as highly efficient is an arduous test of planning and administration;

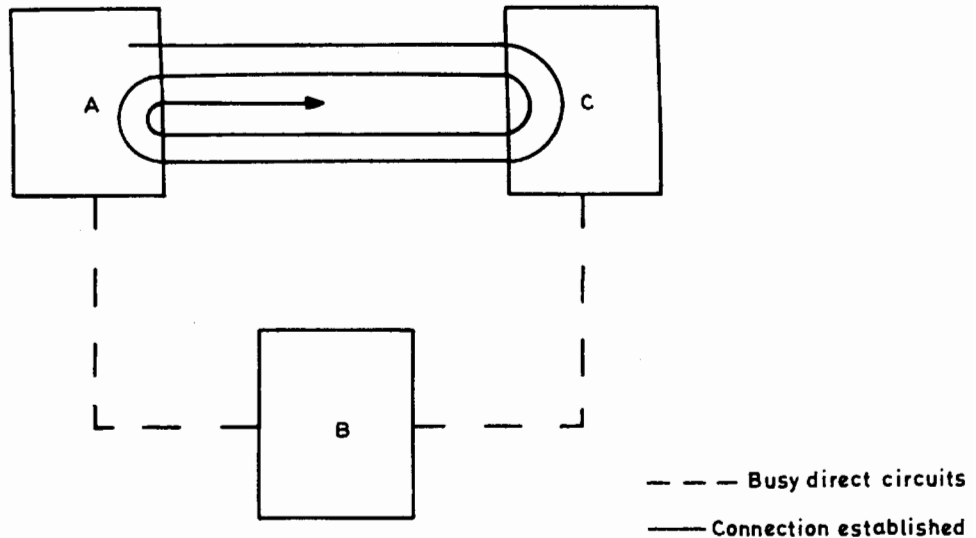


Figure 28.1 Uncontrolled circular routing

but it is well worth while when we consider the alternative of uncontrolled network routing and its disastrous effect on network congestion and the quality of connections, which the examples in Figure 28.1 and 28.2 will illustrate.

In the example shown in Figure 28.1, a *circuit-switched* connection (such as a telephone call) is desired from exchange A to exchange B. Exchange A has direct circuits to B, but these are currently busy, so exchange A has made a connection to exchange C and passed the call on, intending to transit this exchange and connect via the direct circuits from C to B. Unfortunately these circuits are also busy, and by taking no cognisance of the call's previous history, exchange C extends the connection back to exchange A using a similar logic. ('My direct circuits are busy, but I know there is also a route via A') The process continues in a circular fashion until either all the circuits between A and C also become busy, so that the call eventually fails, or finally a circuit becomes available on either of the routes A-B or C-B, in which case the call eventually completes. In either eventuality the circular routing between A and C ties up network resources, restricting communication between customers on exchanges A and C. Furthermore, even if the call does eventually complete, the transmission quality is likely to be appalling or the delay in packet data delivery may be intolerably long, due to the large number of links in the connection.

A second effect of uncontrolled routing is shown in Figure 28.2, where a communication path has finally been completed over 9 individual links, transmitting 8 intermediate exchanges. As in the circular routing example of Figure 28.1, even though the call has completed, an undue quantity of network resources have been tied up, causing possible congestion for other traffic. In addition, the large number of links again lead to poor transmission quality or intolerable delay. The quality will be particularly poor if one or more of the nine links passes over a satellite connection. In this case the

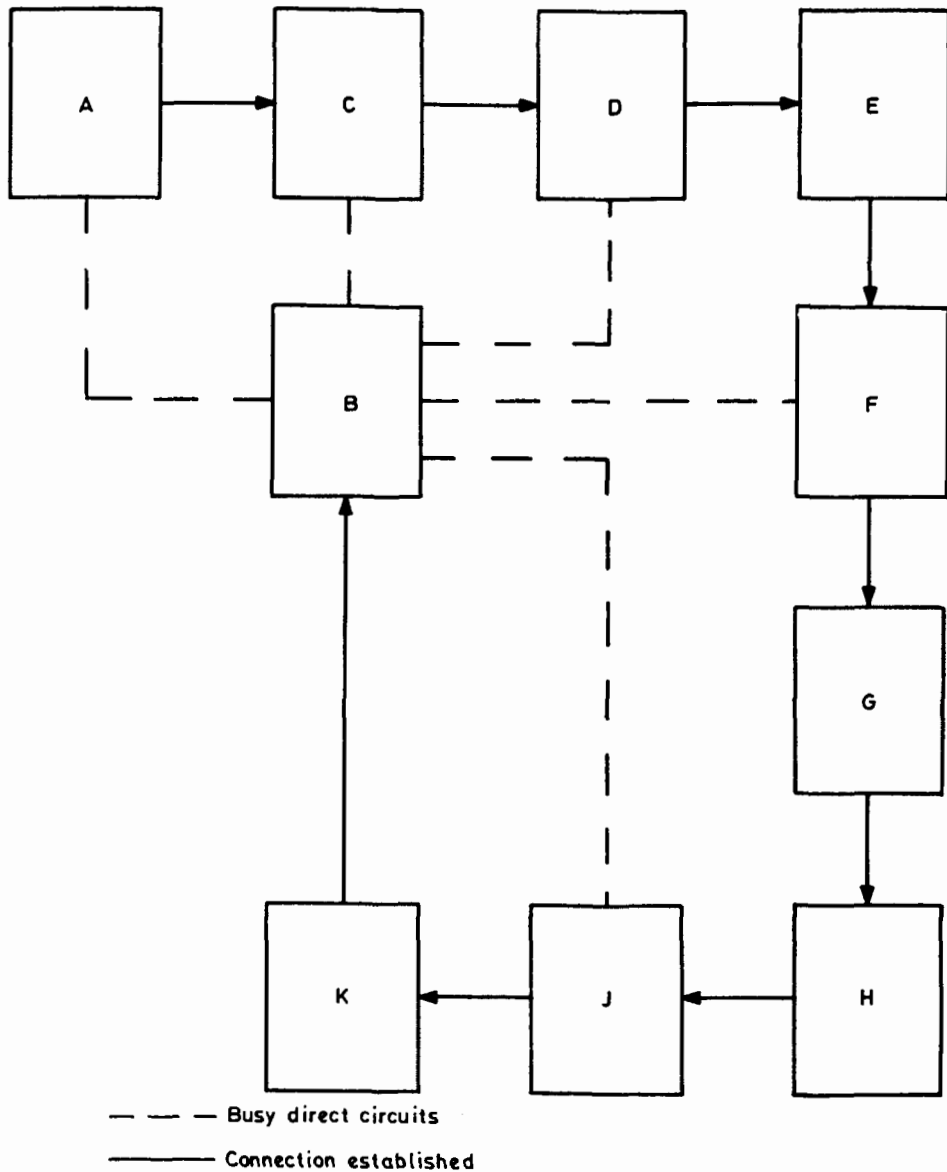


Figure 28.2 Uncontrolled transit routing

end-to-end propagation time may be several seconds. On a packet mode data connection, the data throughput rate is likely to be severely limited by such long propagation delays, particularly if the protocol requires acknowledgement of individual packets.

In practice, cases as extreme as those illustrated in Figures 28.1 and 28.2 are unlikely to arise. Nonetheless, the problems of *circular routing* and the need to set a maximum *number of hops* are real. Circular routing does not typically take place between only two

exchanges as shown in Figure 28.1, more likely in a ring of three, four or five transit exchanges. Engineering guidelines governing connection quality typically set maximum number of hops to values like 3, 5, 7 or 9 depending upon the type of network and the use to which it is being put.

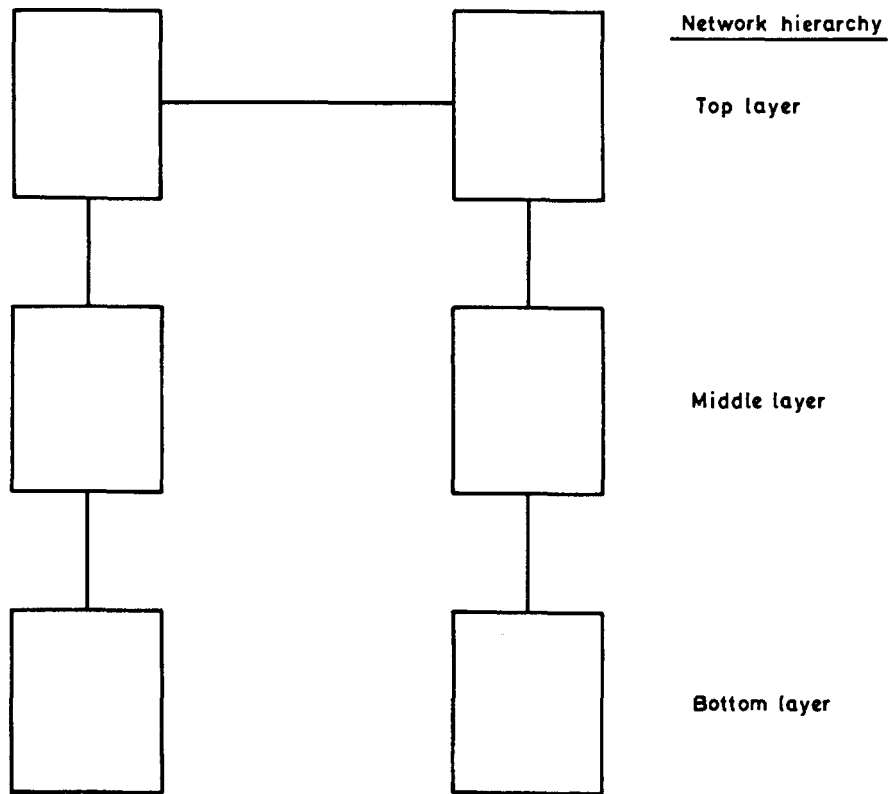
## 28.2 NETWORK ROUTING OBJECTIVES AND CONSTRAINTS

To maintain high transmission quality and to ensure the minimization of time delays both on call set up and on message or speech propagation, it is desirable to minimize the overall number of links and exchanges making up a connection. In addition it is desirable to limit the number of concatenated links of certain transmission types (e.g. satellite links, or links using low rate speech), since the tandem connection of such devices can lead to unacceptable transmission impairment, as we shall see in Chapter 33.

Historically networks were designed in a hierarchical fashion. This enabled the number of links required on a *maximally adverse connection* between any two endpoints to be limited. The maximum hop limit is set by the number of layers in the network hierarchy. Network topology rules ensure full interconnectivity of all exchanges in the top layer of the hierarchy and connection of each lower level exchange with at least one exchange in the next higher layer. Thus a hierarchical network structure consisting of  $n$  layers needs, at most, only  $(2n - 1)$  links to interconnect any two exchanges. For example, in a network comprising a three layer hierarchy, any exchange may be connected to any other without the need to use more than  $(3 \times 2 - 1) = 5$  links, as Figure 28.3 shows.

In more modern networks, the strict hierarchical method of network design is becoming less common, in favour of simpler and more flexible network topologies and routing schemes. Less rigidly structured networks prevail in which each exchange recognizes

- the need to select an economical routing conforming with engineering guidelines and the requirements of the transmission plan (Chapter 33)
- the need to contain the likelihood of rapidly escalated network congestion
- the need to charge for calls in line with the incurred costs
- the need for flexibility of routing, in order that network operators may take advantage of the non-coincidence of busy periods (so-called *route busy hours*) via transit exchanges
- the need to minimize the overall number of links in a connection, and in particular to limit the number of satellite links or bandwidth compression equipments which may be used in tandem
- the policy of preferred transmission media, say when alternative satellite and cable links are available to the same destination
- the need to avoid circular routings



**Figure 28.3** Five-link connection in a three-tier network hierarchy

Careful network design and a shrewd call routing programme at each exchange will ensure conformance to the routing plan, but this may require considerable administrative effort in establishing appropriate routing tables at each switch or exchange within the network as we discuss later.

Signals which accompany the call or connection setup message are intended to help to convey the previous history of the call or packet (for example, the existence of a previous satellite link) and make appropriate routing choice easier.

In the example of Figure 28.4, caller 1, connected to exchange A and wishing to call B, has reached exchange C by means of a satellite link. Although both cable and satellite links are available from exchange C to exchange B, the call is only allowed to mature if a circuit is available on the cable link. If instead the call were to be permitted to overflow to the satellite link, then the connection would not meet the required transmission quality standard. (If, however, the connection was only possible by the use of a double satellite link, then the call could have been permitted to mature).

By contrast, caller 2, (on exchange C, may be connected either over the satellite or the cable link. Two alternative routing policies are available to the owner and operator of exchange C to ensure optimum routing of both caller's calls. In the one shown, the operator has chosen to make the satellite first choice for caller 2's calls. This inflicts

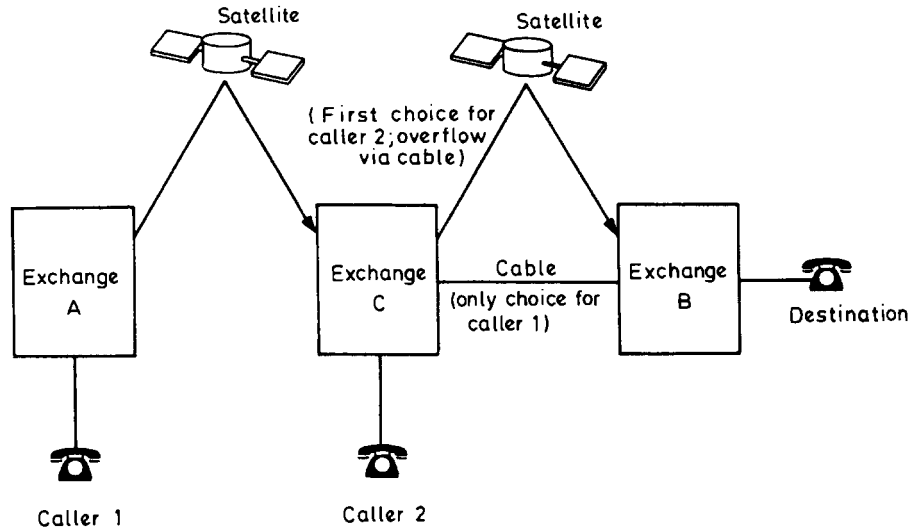


Figure 28.4 Routing based on call history

the propagation delays associated with satellite links on a large proportion of caller 2's calls to exchange B, but has the advantageous effect of maximizing the availability of cable circuits for connection of caller 1's calls to exchange B, so preventing the failure of calls in the instance when otherwise only an unacceptable 'double satellite path' were available.

In the alternative scheme the operator of exchange C could have chosen to make the cable link to exchange B first choice, even for caller 2's calls. This would have the effect of minimizing the propagation time of caller 2's calls, and may be desirable when caller 1 is a customer of a different network operator.

A common feature of all good routing schemes is their simplicity. Complicated routing schemes can lead to administrative difficulties and oversights. Apart from network congestion and poor transmission quality, slow call set-up and a burden of exchange data maintenance can also result.

All routing schemes rely upon the exchanges to analyse the *dialled number* or *network address* (i.e. OSI layer 3 address) to determine the destination of the call. Additionally, signalling information about required supplementary services (e.g. *closed user group* or *intelligent network* services) and the call's previous history (e.g. 'previous satellite link') help to determine the selection of an appropriate route to the destination and an appropriate charge.

*Closed user group (CUG)* information carried at connection setup time can be used to ensure that only certain customer lines or ports may be connected together. This might help, for example, to prevent unauthorized dial-in to a computer centre. Only members of the *CUG* may be connected to the centre. Intelligent network services include, among others, freephone, in which the charges for the call are invoiced to the receiver rather than the call originator. Finally, the connection history (e.g. 'previous satellite link') or required quality attributes of the connection (e.g. for frame relay the *committed information rate (CIR)*) may also affect call setup or connection routing.

The switches in all types of networks therefore need to analyse the *network address* to determine the intended destination of a connection and other service parameters and quality information to assess any constraints on the path to the destination. Ideally, only the minimum amount of information is analysed at any particular switch or exchange, to minimize time and effort required to determine the next step in the path. Thus, for example, at an outgoing international telephone exchange at least the *country code* indicator digits of the dialled number need to be inspected to select the appropriate route to the country concerned. A trunk telephone exchange must inspect only the *area code* to determine the onward route selection. Finally, a destination local exchange needs to examine all the digits of the destination customer's local number to select the exact line required.

### 28.3 THE ADMINISTRATION OF ROUTING TABLES

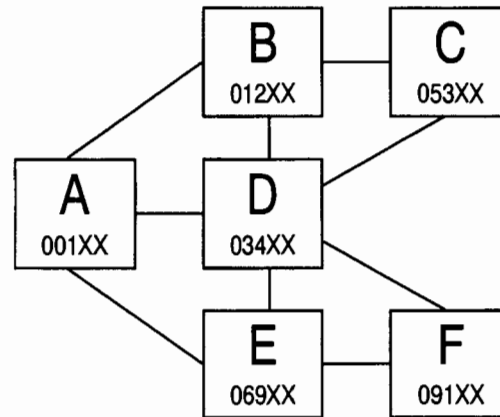
Historically, network routing plans were administered by means of *routing tables* in each of the individual switches or exchanges. Each exchange thus had a 'look-up' table of permissible address code (e.g. telephone area codes), and alongside each code, a list of the alternative routes available for completion of relevant calls. Thus in the example of Figure 28.5, we illustrate the network topology of six interconnected nodes, and the routing tables resident in exchanges A and B to reach the various telephone number blocks, 001XX (at A), 012XX (at B), 034XX (at D), 053XX (at C), 069XX (at E) and 091XX (at F).

The example of Figure 28.5 illustrates the complexity of setting up and administering the routing plan, as well as the problems of circular routing and maximum hop count already discussed. The first observation is that each of the exchanges requires a separate routing table. There is little or no commonality between the routing tables (the other four for exchanges C, D, E and F are not shown), so that considerable manual effort is required first to work out the tables and second to type them in to the *configuration data* of each of the individual exchanges. There is a very high probability in complex networks of errors in the routing plan design and further potential for errors during the typing-in stage.

If we now examine closely the routing commands given to exchanges A and B in Figure 28.5 for the handling of codes 053XX (to exchange C) we can see the potential for a circular route being set up, for if exchange D is told to use 'via A' as a third choice route to exchange C (code 053XX) then at times when the links B-C and D-C are overloaded or out-of-service due to network failure calls to code 053XX may be passed in endless loop A-B-D-A, etc.

We also see the problem of minimizing the maximum hop count. The intention of the designer of the network in Figure 28.5 is that the maximum hop count shall be three. Thus, for example, the third choice route from A-to-C is A-E-D-C. However, the third choice route from E-to-C might also be via three hops (E-A-B-C), so how do we prevent the circuitous routing (exceeding the maximum hop count) A-E-A-B-C? The answer is that the routing tables need also to take account of the *origin* of the call as well as the intended destination. Thus calls arising at exchange E but *originated* by exchanges other than exchange E should not be allowed the third option to





routing table at A

dialed code	route via:		
	first choice	second choice	third choice
001XX	direct	none	none
012XX	B	via D	via E
034XX	D	via B	via E
053XX	via B	via D	via E
069XX	E	via D	via B
091XX	via E	via D	via B

routing table at B

dialed code	route via:		
	first choice	second choice	third choice
001XX	via A	via D	none
012XX	direct	none	none
034XX	via D	via A	via C
053XX	via C	via D	none
069XX	via D	via A	none
091XX	via D	via A	none

Figure 28.5 Routing tables in a typical telephone or ISDN network

exchange C. Similarly, calls appearing at exchange E directly from exchange A should not be passed directly back again.

Further increasing the problem, the dialled digit train may need to be altered. Historically, this was necessary because the switching action of electromagnetic exchanges was triggered by the pulsed digit train, so that the digits were literally 'used' to activate the switching. As a result each subsequent exchange sent fewer digits to the next along the chain of the connection. So that, for example, an electromagnetic exchange at point A in Figure 28.5, might expect only to receive the digits XX when accepting calls to the digit range 001XX, the '001' having been used or deleted by previous exchanges in the connection. Modern computer controlled exchanges generally relay the entire dialled number, but when signalling to older electromechanical exchanges they may have to adapt the train to the *routing digits* required to activate the switching (Chapter 6).

The problems of call origin and call history dependent routing described above make for complicated signalling between the exchanges and complicated routing tables (based on the *route origin*) within the exchanges. Worse still, every time further capacity or new trunks are added to the network topology, all the routing tables in each of the exchanges may need to be amended. Routing table administration remains one of the major operational burdens of telephone and ISDN network operators.

## 28.4 ROUTING PROTOCOLS USED IN MODERN NETWORKS

In contrast to telephone networks, where typically the individual switches (exchanges) are supplied by different equipment manufacturers, data networks have often been built from switches all supplied by a single manufacturer, with a common *network management* system. The common manufacturer and network management system shared by all the switches enables the use of proprietary signalling and control mechanisms to be applied to traffic routing within the network. Thus most data network management systems require only the association of groups of destination network addresses to particular switches. The routing tables for all other switches are then generated according to the network management system's knowledge of the current network topology, using a set of automated routing design rules and routing *algorithms* (e.g. preference for high capacity routes over low capacity routes, preference for low hop count path, etc.). The human task of administering routing tables in modern data networks is thus far more straightforward than telephone network routing table administration.

Once the route is set-up for a particular connection (i.e. in a connection-oriented network such as X.25 packet switching, frame relay or ATM), it is not usually altered during the duration of the call (i.e. the period of communication). Leaving the routing of the connection unaltered (path oriented routing) means that the transmission propagation time across the network between the two devices is not subject to any unnecessary jitter (variability of delay). In addition, there is much reduced risk of cells which might otherwise have taken different paths from getting out of order. It is also much easier to determine and manage a network loading scheme, because nominal bandwidth allocations may be made to each of the connections which must statistically share a given physical transmission path.

In the most modern of networks (e.g. router and ATM networks), the entire routing administration is automated, so that switches within the network are programmed to 'learn' about the topology of the network the ideal route to a given destination (network address). A *routing protocol* is employed by such networks so that the individual nodes can discover the network topology automatically and keep themselves abreast of changes. Examples of *routing protocols* are used in the *Internet* are

- routing information protocol (RIP)
- open shortest path first (OSPF)
- border gateway protocol (BGP)
- exterior gateway protocol (EGP)

Routing protocols are used widely in the *Internet* to pass information between routers about the various sub-networks making up the network. One of the first protocols developed was the *exterior gateway protocol (EGP)* defined by RFCs 827, 888 and 904. This was a protocol intended to be used between router on a sub-network (say university campus) and an inter-site network (internet) so that *internal* UNIX computers on the sub-network could locate and establish connections to *exterior* ones in bordering networks. EGP has subsequently been largely replaced by the *border gateway protocol (BGP)* defined by RFC 1267.

Within most router networks (e.g. Cisco, Wellfleet, 3Com, etc.) it is common to use proprietary routing protocols (*interior routing protocols, IRP*), but the *RIP (routing information protocol)* defined by RFC 1058 set the initial standard for transfer of routing topology information, so that a routing table could be maintained by a *source router*. The table enables the router (near the source of a message) to determine the best path across the *Internet*. RIP complements the *hello protocol* of RFC 891 which is used to register and synchronise new connections in the network.

The *OSPF (open shortest path first)* protocol is a newer, more complex and more sophisticated protocol than RIP but intended to bring about a simplification of the topology of the *Internet* by introducing a structured hierarchy of routing nodes. It has become the accepted 'standard' routing protocol in router networks, *Intranets* (corporate router networks) and the *Internet*.

As an example of the way in which switches within a modern network may be programmed automatically to discover the network topology and keep abreast of all changes made to it, thus enabling optimal routing of calls, connection and traffic at all times, we discuss next how the *hello state machine* defined in the ATM network standards (see also Chapter 26) enables constant updating of the *ATM network topology state*.

## 28.5 NETWORK TOPOLOGY STATE AND THE 'HELLO STATE MACHINE'

The ATM forum is developing, as part of its *PNNI (private network-node interface, based on the ATM UNI v3.1)* specification, a sophisticated *source routing* control mechanism, in many ways similar to the techniques used in the *Internet*.

By keeping a record in its *topology database* of all information supplied to it about the topology of the network as a whole, an ATM network node always has a view of the entire private ATM network routing domain. The node is thus able to determine the route from itself to any *reachable* address.

The information about the topology and any changes made to it are conveyed as *topology state elements*, including *topology state parameters*. These are conveyed between the nodes in the network by means of *topology state packets*. *Topology state parameters* are classified into two types

- attributes (these influence routing decisions; a security *attribute* of a particular node may cause the set-up of a particular connection to be refused)
- metrics (these are values which accumulate over the path of the connection as a whole to determine whether it is acceptable, e.g. the propagation delays of individual links in the connection are added as *metrics*)

When a new link or node is added to the network, then the directly affected nodes communicate with one another over the new link using the *hello* procedure. This is a standardized protocol enabling the two nodes to identify themselves to one another and work out the change in topology of the network as a whole. The new topology information is then flooded (i.e. broadcast) to the other nodes in their *peer group* (i.e. sub-network) or *advertised* to neighbouring border nodes of neighbouring peer groups by means of *topology state packets*. These inform the other nodes of any new addresses which are now *reachable* and also which exit route to take from the *peer group (PG)*.

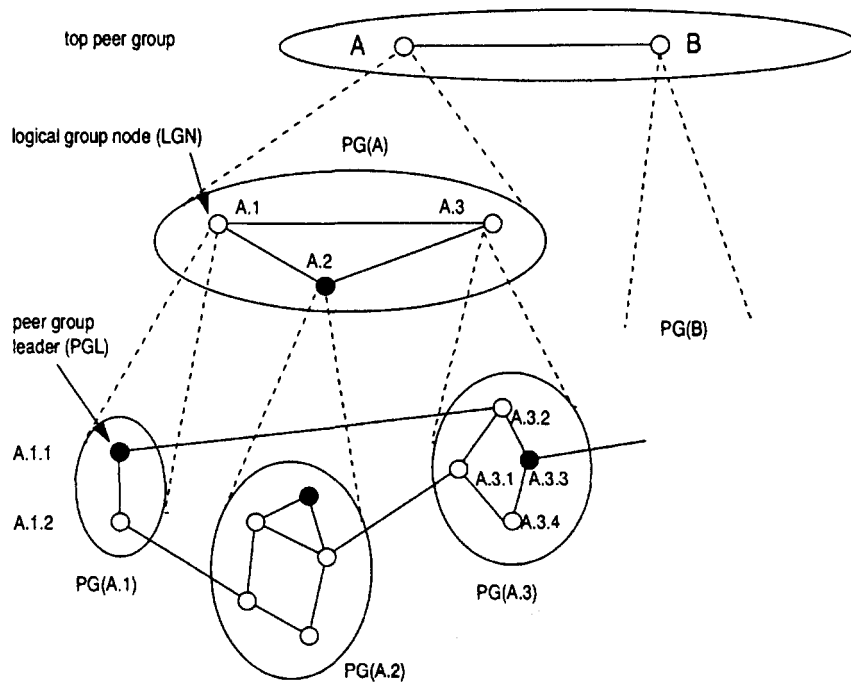
The routing information is stored in the topology database as 'address A *reachable* through entity B', where B is a known node within the peer group. Routing to outside of the *peer group*, as we shall see, is catered for by the *peer group leader (PGL)*.

A *peer group* as defined by ATM forum's PNNI specification, is a collection of nodes sharing the same *peer group ID (identifier)*. As the peer group ID is defined during the initial configuration of the node by the human installer, a *peer group* is in effect a human operator-defined sub-network.

Within a *peer group* an *election* determines the *peer group leader (PGL)*. The *election* is an ongoing process which results in the node with the highest *leadership priority* taking over certain of the more important network routing (inter-sub-network) tasks on behalf of the peer group as a whole. The peer group leader is automatically promoted to the next layer in the routing hierarchy. However, should another node achieve a higher leadership priority (as a result of some topology or capacity change within the peer group) then it will take over the PGL function.

At the next layer of the hierarchy each *peer group* appears only as a single node, a *logical group node (LGN)*. It is represented in the topology management process at this level by the peer group leader. At the highest level in the hierarchy is the top peer group (Figure 28.6).

When neighbouring nodes running the *hello protocol* conclude that they belong to the same *peer group* then they synchronize their databases (recording the sub-network or *peer group* structure). They then *flood* (i.e. broadcast) this information to all other members of the peer group. In the case where the nodes do not belong to the same peer group, then they are *border nodes* in adjacent peer groups, and an *uplink* is said to exist



**Figure 28.6** Peer groups (PG), logical group nodes (LGN) and the hierarchy of PNNI routing domains

from the border node to the *peer group leader* of the neighbouring *logical group node*. The communication between the border node and its partner's *peer group leader* is equivalent to the hello procedure but this time between *logical group nodes* concluding that they are interconnected. The new topology information in this case is said to be *advertised* to other *peer group leaders* at the same hierarchical level. The *exterior reachable addresses* (ERA, i.e. those outside the peer group) are defined in a special routing table called a *designated transit list* (DTL) held by the peer group leader.

In contrast to *uplinks*, *horizontal links* are logical links between nodes in the same peer group.

Once the route to a given destination has been determined by a source node (either from its own database or from information provided by the peer group leader), normal UNI signalling procedures can be used to set up the connection. If necessary (e.g. due to current traffic conditions in the network causing a particular route to be unsuitable or overloaded) *crankback* and *alternative routing* may be invoked (in other words the first route choice is abandoned and a second path is attempted).

The node names (or addresses) used to identify nodes in a PNNI network are similar in style to *Internet* addresses: lots of numbers, dots and letters (Chapter 29). Thus the four nodes in peer group PG(A.3) of Figure 28.6 are called A.3.1, A.3.2, A.3.3, A.3.4. This is the style of information held by the *topology database* and *designated transit list* (Address A reachable via B3.2, C4.4, D5.7, ..., etc.).

## 28.6 SIGNALLING IMPACT UPON ROUTING AND CALL SET-UP DELAYS

We next consider the way in which the network signalling can impact upon the time taken to analyse destination network addresses and setup connections across a network. Our example is pitched in a telephone or ISDN network but similar effects could equally impact the propagation of packets or frames across a data network.

Most telephone network and ISDN signalling systems allow the number analysis and route selection to be carried out in one of two ways, either in an *en bloc* manner, or in the alternative *overlap* manner. In the *en bloc* manner, the first local exchange waits for the customer to dial all the digits of the destination number before the number analysis is completed and the outgoing route is selected. All necessary digits of the dialed number and other information are then sent together (or *en bloc*) to the subsequent exchange. The subsequent exchanges are thus not bothered with setting up calls until all the information about the call and its destination is available. The exchange processor load on subsequent exchanges is thereby minimized. Figure 28.7(a) illustrates *en bloc* call set-up.

In the alternative *overlap* manner of call set up, each exchange in the connection selects the outgoing route as soon as it has sufficient information to do so (even if not all the information about the destination has been received) and passes on subsequent information as it receives it. Thus in the diagram of Figure 28.7(b), the connection may already have been made right through to exchange 'C' even before the customer has finished all the digits of the destination number. The same would not be true in the *en bloc* case. It is this feature that gives the *overlap* signalling method its edge over *en bloc*

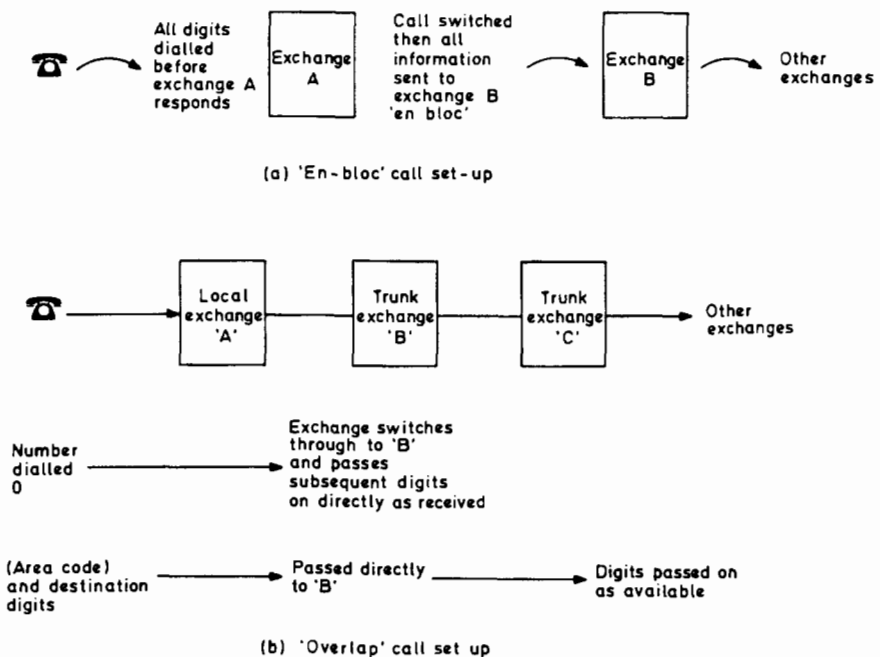


Figure 28.7 'En bloc' and 'overlap' signalling at call set-up

signalling, in being faster at setting up calls. The disadvantage of the method is the greater processing time wasted at all exchanges waiting to receive and relay digits of the dialled number.

The same problems arise in data networks.

In a frame relay network, frames may be many thousand bytes in length, each preceded by a *header* which identifies the intended destination. The time to propagate the frame from end-to-end across the network thus depends heavily upon whether it is relayed in its entirety from each node to the next and wholly received before it is relayed on (akin *en bloc mode*), or whether onward transmission may start as soon as the destination has been identified (i.e. immediately after analysis of the header, akin to *overlap mode*). In many cases in datacommunication, where a *cyclic redundancy check (CRC)* code is applied to detect and correct errors in the contents of the frame, the *en bloc* mode has to be used, because the CRC must be checked before the frame is relayed onwards. The CRC is usually transmitted at the end of the frame. Although this improves the accuracy of delivered frames, it increases the time needed for propagation through the network.

## 28.7 PLAUSIBILITY CHECK DURING NUMBER ANALYSIS

A common failing of network operators, and one which may seem attractive (particularly to those operators using the *en bloc* method of signalling call or connection set-up information), is to carry out undue *plausibility checks* on the destination *network address* or dialled number. Thus, for example, telephone exchanges could be made to look up and check whether a valid number of digits have been dialled, or whether a particular area code within a destination country is valid, etc. Such *plausibility checks* can have the benefit of removing the burden of spurious traffic from the network. Unfortunately, however, the updating of these *plausibility checks* is often overlooked when new area codes are made available in the destination country, or when number length changes are made. The result is that the exchanges may fail calls to newly valid numbers, with understandable customer annoyance. Had the *plausibility check* never been used, the problem would not have arisen. Furthermore where plausibility checks are instigated in a network using overlap signalling, the call set up may be unnecessarily delayed. Special administrative care is therefore required in the use of such checks.

## 28.8 NETWORK INTERCONNECTION

Until the early 1980s most telecommunications networks were owned and operated by state monopoly telephone companies. Thus within a given country or territory the telephone network or public data network was operated by a single entity, so that technical interface standards and quality levels were uniform across the network and routing plans could be determined unilaterally. For international interconnection of national monopoly the ITU (International Telecommunications Union) existed to agree and regulate common international standards for *gateway* connections and routing between different national networks. These set the standards for technical interworking and operational cooperation.

However, by the late 1980s, the traditional way of operating networks was no longer coping with the increasing demands and expectations of customers. As a result of the pressure for deregulation and competition between operators, the old ITU scheme for network interconnection and management began to fall apart and be replaced by a more commercially, 'free-trade'-oriented regulation scheme (Chapter 44). A flood of new operators and new networks have appeared and new rules have had to be found to regulate the *interconnection* of networks.

*Interconnection* of modern networks is a far more controversial subject than that of *gateways* and *interworking* agreements formerly regulated by the ITU, since *interconnection* is crucial to the existence of the new competing networks. The commercial terms for *interconnection* often determine the profitability or failure of the new network operators. In the next section we consider the most important services which need to be considered by companies when *interconnecting* networks.

## 28.9 NETWORK INTERCONNECTION SERVICES

For the basic interconnection of networks managed by network operators who do not have competing interests, rather only a common interest to ensure greater interconnectivity of their respective customers, the technical standards for gateway connections (e.g. *network-network interfaces (NNI)*, frame relay NNI, ITU-T X.75 for packet switched networks, etc.) and for interworking of networks as laid out by ITU are likely to suffice. However, for the interconnection of networks operated by competing organizations, it has been necessary at a legal or regulatory level to set out rules for *interconnection*, including technical standards, service levels, operational practices and

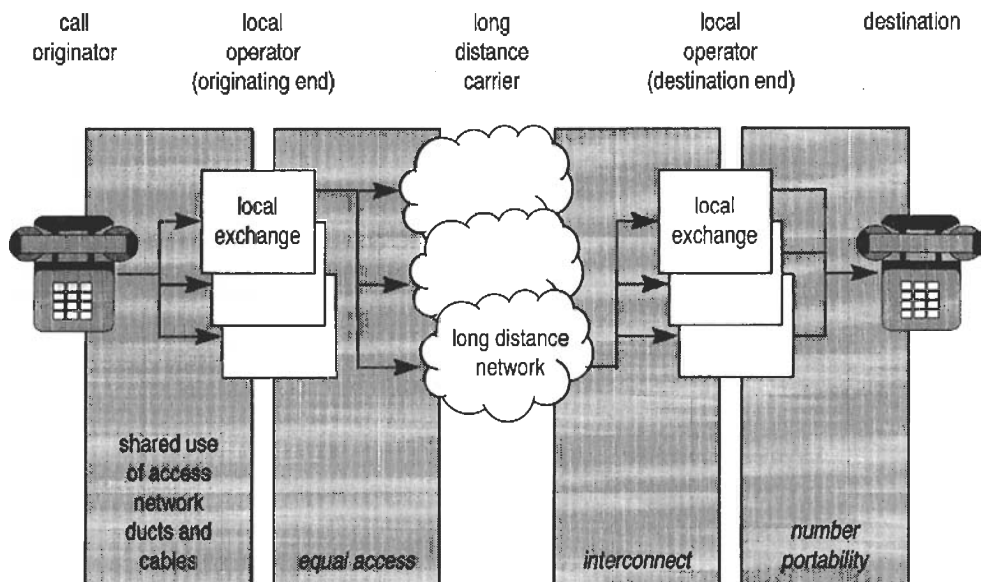


Figure 28.8 Important interconnection services



interconnection tariffs. Regional and national regulators play an important role in setting and overseeing these rules.

Figure 28.8 illustrates four of the main types of *interconnection* which are now usually stipulated by regulators to be made available by ex-monopoly and market-dominant network operators. They are

- shared use of access network ducts and cables
- *equal access*
- *interconnect*
- *number portability*

## 28.10 INTERCONNECT

*Interconnect* is the most important of the services. This is the basis by which customers connected to one network may call customers of a network operated by a second network operator. Technical standards for *interconnect* (the delivery of calls by one operator on behalf of another) are typically based on ITU-T recommendations, amended as necessary to meet local operating conditions or constraints. These standards have tended to be set by bilateral agreement between the parties or trilaterally with the involvement of the industry regulator. More recently, standards bodies are beginning to define *inter-network (INI)* and *inter-carrier interfaces (ICI)* which accommodate not only technical interfaces suitable for use between networks of competing carriers, but also setting out appropriate operational, administrative and management practices. An example of such a standard is the *broadband inter-carrier interface (B-ICI)* which has been developed by ATM Forum for interconnection of broadband ISDN or ATM networks. A second example (for interconnection of private ATM networks) is *PNNI (private network-network interface)*.

Where the parties are unable to agree on the tariffs to be paid by the network operator of the call originator to the network operator delivering the call, the *regulator* may be required to make a *determination*. Thus, for example, the *FCC (Federal Communications Commission)* needed to *determine* charges for *interconnect* between US network operators in the 1980s, as did *Oftel (Office of Telecommunications)* for the UK operators British Telecom and Mercury.

## 28.11 EQUAL ACCESS

*Equal access* was first invented in 1986 by the FCC (Federal Communications Commission) in the United States. It is a form of network interconnection between a *local exchange carrier (LEC)* and an *inter-exchange carrier (IEC)*. The idea was to ensure fair competition between the IECs (e.g. AT&T, MCI and Sprint) for the carriage of long distance telephone calls. Long distance calls are defined in the USA as calls crossing *LATA (local access transport area)* boundaries. *LATAs* are the small regional

areas within which LECs operate (maybe as a monopoly). For an *inter-exchange carrier (IEC)* to be able to offer his long distance telephone service to customers within a LATA, he needed to demand equal access service from the LEC at the *point-of-presence (POP)*. In this way it was made possible for LEC customers to choose between various IECs for carriage of long distance calls.

An LEC customer could either *subscribe* to any of the *IECs* for all his long distance calls (*pre-selection*), or could on a *call-by-call* basis elect to dial an *equal access code* to choose the carrier he wished to use for a particular call. His long distance calls were invoiced to him directly by the chosen IEC.

Where a customer subscribes to *pre-selection*, only the normal trunk or toll telephone number must be dialled when calling long distance or international. Where the customer elects for *call-by-call equal access* a carrier code (e.g. 10XXX, where XXX is a three digit combination identifying a specific IEC) must be dialled prior to the destination customer telephone number.

Equal access is not available in all countries, not even all countries where competition has been established. It is one of Mercury's grudges, that it has not been introduced in the UK, where Mercury feels it is necessary to enable stronger competition against British Telecom. It will be introduced in Germany and other European countries.

## 28.12 NUMBER PORTABILITY

The demand of new operators for *number portability* has arisen only in the last couple of years, as competition has been introduced at the *local exchange carrier (LEC)* level as well as at the long distance (*IEC, inter-exchange carrier*) level. New LECs have realized with dismay, that customers are reluctant to change operator if this will result in a change of telephone number or network address, since this is associated with considerable cost, upheaval and effort. To change a telephone number, letterheads, product packaging and documentation all need to be changed and customers and suppliers must all be informed. To change data network addresses, large numbers of computers and network devices may need to be re-programmed.

*Number portability* is a service enabling the customer to retain his telephone number or other network address despite being connected to a different *local exchange carrier's (LEC)* network. In effect, calls arriving in the network where the customer was previously connected are forwarded to his new connection in another network.

Number portability is beginning to appear in the most advanced networks (USA, UK, Germany).

## 28.13 SHARED USE OF ACCESS NETWORK DUCTS AND CABLES

A significant economic problem faced by new LECs or cable companies who commence competition against ex-monopoly carriers, is the enormous investment required to establish a cable and duct infrastructure equivalent to the existing access network of their main competitor. The size of the investment required severely limits the scope for highly attractive pricing in the short term. As a result, pressure from new carriers has

grown in recent years to force by regulatory means the ex-monopoly operators to offer the shared use of their ducts and cables. The idea is that a new operator could connect customers directly to his own local telephone exchange (central office) or other switch using an existing pair of copper wires leased on favourable terms from the ex-monopoly carrier. The wires would be diverted to the new operator's site via a *distribution frame*, cabling *patch panel* or equivalent.

Shared use of cables and ducts, however, is fraught with operational management difficulties. It is likely to be resisted strongly by incumbent (ex-monopoly) operators, and is unlikely to be available in most countries.

## 28.14 PITFALLS OF INTERCONNECTION

Ex-monopoly operators, understandably, have not generally been willing partners in agreeing terms for *interconnection* with new carriers competing against them. This factor, combined with the relative inexperience of many of the new carriers has tended to lead to one-sided interconnection agreements.

Examples of subjects sometimes inadequately covered by initial interconnect agreements include

- very slow inter-network signalling, leading to much longer call set-up times for customers of the new network
- unavailability of intelligent network and other special services (e.g. freephone, information services)
- unavailability of *calling line identity (CLI)* and *ISDN supplementary services* (e.g. *ring back when free*)
- inability to support *reverse charge* calls
- unavailability of directory information
- no access to emergency services (e.g. police, fire, ambulance service)
- no operator assistance service

To date, much of the focus of *interconnection* has been on public telephone services, because most of the new operators have viewed this market as the most financially attractive due to the large base of existing demand. As the data network service market explodes, and Internet, broadband and multimedia services appear, the focus of attention is bound to shift. This may reveal new subjects for the regulators of *interconnection* to consider.

## 28.15 THE POINT OF INTERCONNECTION AND COLLOCATION

In the early days of public telephone network interconnection, the ex-monopoly players were keen to offer *points of interconnection (POI)* only on their own premises. The

rationale was their belief that they could then control interconnection on their own terms. In addition, it would make interconnection for the new operator harder, because new lines would have to be laid by the new operators right into the POI, adding to initial cost and effort required for interconnection.

In the meantime, the ex-monopolists have realized that POIs in their own premises are not necessarily a good thing. The large number of new operators demanding interconnection from the old monopoly operators has increased the demand for space at the POI, and required ever increasing numbers of cable duct entry points to the building and greater personnel access by technicians of the new operators installing and maintaining their equipment. As a result there is now a trend towards *in-span interconnection (ISI)* in which the interconnecting operators meet at an agreed manhole or streetside cabinet.

Despite the general desire to move away from POI in the premises of the ex-monopoly operator, there remains demand from new operators for *collocation* of new network equipment next to existing exchange equipment (in the exchange buildings of the ex-monopoly operator). This is motivated partly by economic and partly by technical considerations. The access network of the ex-monopoly operator is likely to have been optimized to have one end in the premises of the local exchange. Extension of such connections to a remote site only adds cost and may not be possible due to technical constraints (such as a limit of the maximum allowed line length).

## 28.16 THE INTERCONNECTION CONTRACT

The first negotiations between ex-monopoly operators and their new rivals were drawn out affairs, coordinated by large teams of lawyers and closely umpired by regulators. There was simply too much at stake, and no experience to indicate the likely outcome. The resulting interconnection contracts were typically cumbersome affairs, each tailor-made for and individually negotiated between a single pair of interconnecting parties. As experience has grown, the form of the contracts has become more standardized and the relationship between the *interconnect requester* and *interconnect provider* has become more balanced. Nowadays the ex-monopoly players are also able to see the opportunity as well as the threat caused by interconnection and the new operators recognize that heavily biased contractual conditions may work against them when the incoming and outgoing traffic streams begin to balance out after a couple of years of operation. Generally therefore a more balanced view of interconnection is appearing, with opportunity and responsibility for all involved.

The effort required particularly by the ex-monopoly operators to administer a gamut of different interconnection agreements, with different services and tariffs has become unmanageable, so that they are as keen as the regulators and the new operators to see the establishment of standard interconnection contract terms, though prices may differ on a bilateral basis. The European Commission is playing a leading role in helping to define which topics should be covered by such a contract. These are listed in Table 28.1.

There are as yet no standard contracts, though significant progress has been made in the United Kingdom, spearheaded by *Oftel (Office of Telecommunications)*, British Telecom and the *other licensed operators (OLO)* committee.

**Table 28.1** Topics to be covered by interconnection agreements or contracts

	Topics to be covered
Ex-ante conditions for interconnection to be set by the national regulatory authority	<ul style="list-style-type: none"> <li>● dispute resolution procedure</li> <li>● requirement for publication of terms</li> <li>● requirement for equal access and number portability</li> <li>● requirements for facility sharing, including collocation and resource sharing</li> <li>● maintenance practices and standards</li> <li>● requirements for allocation of numbering resources and access to directory and emergency services</li> <li>● expected end-to-end quality standards</li> <li>● determined interconnection charges</li> </ul>
Subjects to be covered by interconnection agreements	<ul style="list-style-type: none"> <li>● description of interconnect services</li> <li>● terms of payment and billing procedure</li> <li>● locations of points of interconnection</li> <li>● technical standards for interconnection</li> <li>● measures ensuring compatibility and conformance</li> <li>● intellectual property rights</li> <li>● definition and limitation of liability and indemnity</li> <li>● definition of interconnection charges</li> <li>● dispute resolution procedure prior to escalation to national regulator</li> <li>● duration and renegotiation of agreement;</li> <li>● procedure in the event of alterations to the network or services proposed by one of the parties</li> </ul>
Optional issues to be covered by interconnection agreements	<ul style="list-style-type: none"> <li>● achievement of equal access</li> <li>● resource sharing</li> <li>● access to advanced services beyond those legally required to be offered</li> <li>● traffic and network management</li> <li>● confidentiality of non-public parts of the agreement</li> <li>● training of staff</li> </ul>

## 28.17 INTERWORKING

*Interworking* is the term applied by to the interconnection of dissimilar networks (e.g. an interconnection between an X.25-based packet switched data network and the ISDN (integrated digital services network)). It requires an *interworking function (IWF)* or *interworking unit (IWU)*.

It is rarely the case that the *interworking* carried out is able to support the full range of services available within each of the individual networks. Usually only a small subset of the services (the *common denominator subset*) can be interworked. Other services may be ruled out either due to the incompatibility of the different networks, or due to the fact that no equivalent service exists in the other network.

When *interworking* the modern ISDN with the old telephone network, for example, it is not possible to carry ISDN supplementary services across the point of interconnection. Thus neither customers connected to the telephone network nor those connected to the ISDN may receive an indication of the *calling line identity* on incoming calls though a restricted set of *supplementary services* may be available to the ISDN customer (e.g. call forwarding of incoming calls).

The problems and potential pitfalls of interworking should not be underestimated.



---

## *Network Numbering and Addressing Plans*

---

The network numbering or addressing plan is an important part of the network routing plan, because the network address is the identification used by a caller to identify the customer or network port to which he wishes to be connected. Based on the network address, switched connections within all types of network are established. In this chapter we discuss the five basic numbering and addressing schemes, and the principles that go along with them. These schemes are

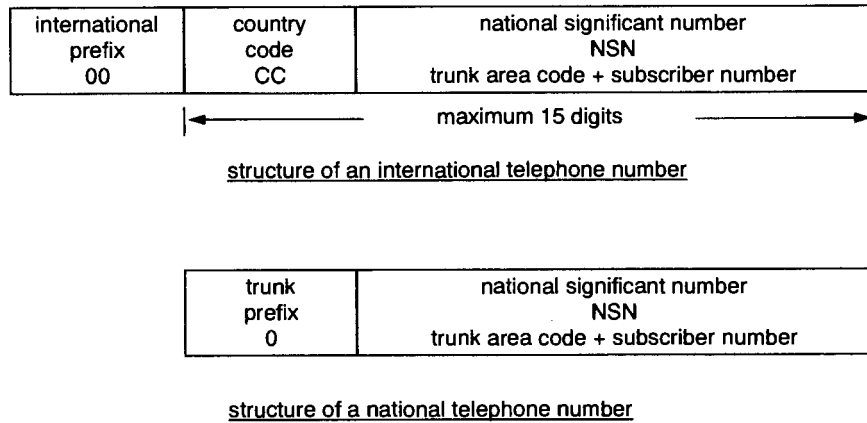
- international telephone service/ISDN numbering plan according to ITU-T recommendations E.163 and E.164 (recently combined into one recommendation, E.164)
- international public data network numbering scheme according to ITU-T recommendation X.121
- international telex service numbering according to ITU-T recommendation F.69
- addressing scheme for the message handling service (ITU-T recommendations X.500)
- Internet addressing scheme

### **29.1 THE INTERNATIONAL TELEPHONE NUMBERING PLAN**

The international telephone numbering plan was originally defined in 1964 by ITU-T's recommendation E.163, which laid down the principles of numbering pertinent to public telephone networks. It remains substantially the same, but has recently been combined with the subsequent recommendation E.164 'numbering for the ISDN era' into a single recommendation E.164.

Recommendation E.164 defines the recommended prefixes for international calls (00) and trunk calls (0), the allocation of *country codes*, and the maximum number length excluding international prefix (15 digits). Figure 29.1 illustrates an E.164 number or *network address*.





**Figure 29.1** E.164 format of international telephone numbers

The E.164 numbering plan ensures the allocation of a unique number (string of digits) to identify each individual telephone line connected to the worldwide telephone network. These numbers are analysed by telephone exchanges to determine the appropriate call routing and the appropriate call charging rate. They are designed to allow exchanges to select an economical and satisfactory onward connection by analysing only a minimum number of digits. The recommendation does not control each individual country's numbering plan, but allocates instead a large series of numbers for use in each individual national network. This flexibility allows each national network operator to prepare a *national numbering plan*, optimized for their own particular purposes.

Most network operators choose to adopt a three-tier numbering plan. The three tiers correspond to overseas (i.e. international) calls, long distance (i.e. *toll* or *trunk*) calls and local calls, and the procedures for each of these types will be discussed in order.

To place an international call over the automatic telephone network, a customer must dial first the *international prefix code* (to indicate that the digits immediately following indicate a destination overseas). The next digits will be a 1, 2 or 3-digit *country code (CC)* to indicate the particular country required, then the area code and destination customer number. Figure 29.1 illustrates the component parts of an example international number.

The example of Figure 29.2 shows a London, UK, telephone number, when dialled from Switzerland. The ITU-T recommended international prefix '00' is followed by country code digits '44' to signify the United Kingdom, digits '171' to identify Central London, digits '234' to specify the destination exchange and digits '5678' to earmark the customer. Successive exchanges within the connection gradually discard the early digits and analyse progressively more of the later ones to determine the required routing.

The standard '00' prefix is used in many countries, including Switzerland, but some countries for historical reasons currently use other prefixes. For example, the international prefix for calls made from USA is currently '011', and from France it is '19'. ITU-T recommends that all countries eventually migrate to the common '00' prefix, to ease travelling customers' dialling difficulties.

## General international number – structure

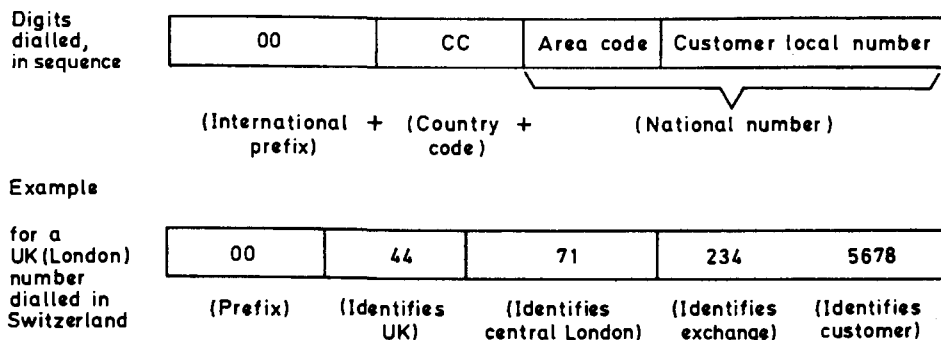


Figure 29.2 International telephone number

The country codes of E.164 numbers, identifying each individual country, are allocated by ITU-T, and the same country code is used to identify a particular country no matter where in the world the telephone caller is situated. Figure 29.3 illustrates the demographic location of world numbering zones.

The entire list of country code allocations is given in Table 29.1.

The North American country code is only one digit. This is because there is an integrated area code scheme, called the *North American dial plan (NADP)* or *number plan of America (NPA)*. This plan covers the whole of the United States, Canada and the Caribbean islands. Telephone calls placed within this area need only be dialled as *toll* numbers. The table of Figure 28.4 shows the allocation of the three-digit area, or correctly *number plan of America (NPA)*, codes which identify the various regions. A typical New York, USA number is thus +1 212 345 6789.

Recommendation E.164 sets the maximum number of digits allowed in an international number at 15 digits plus the international prefix. This limits each national numbering plan (area code plus customer number) to a maximum length of  $15 - n$  digits, where  $n$  is the length in digits of the corresponding country code. Limiting the number length in this way ensures that each network in the world can be designed to cope with the maximum number length. The recommendation also states that it should not be necessary for an outgoing international exchange to analyse more than four digits after the international prefix to determine the routing and charging information for any call. (Four digits correspond to the country code and part or all of the *area code*).

Returning to the example shown in Figure 29.2, Figure 29.5 illustrates the different digit strings for dialling the same London customer from either Switzerland (i.e. overseas calling), Birmingham, UK (i.e. trunk calling), or from a *local* customer in London. Note how the trunk number comprises only the latter part of the full international number, the international prefix and country code having been replaced with a simpler *trunk prefix*. The standard ITU-T trunk prefix '0' is used in the UK, but as is the case with the international prefix, some countries have not yet moved on to the use of the standard trunk prefix, for various historical reasons. (Example: the trunk prefix used in North America is '1'.) Finally, for local calling the customer's number is normally dialled without prefix, as Figure 29.5 shows.

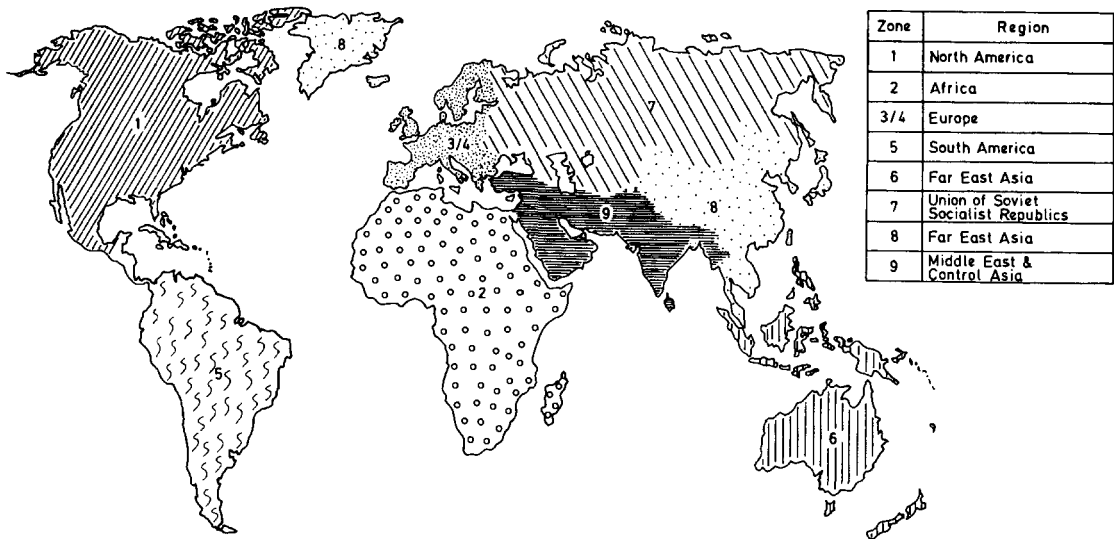


Figure 29.3 World telephone numbering zones (CCITT Recommendations E163/E164)

Four key features of the E.164 'numbering plan for the ISDN era' mark it out from its predecessor, E.163, the telephone numbering plan

- Extended international numbers (up to 15 digits rather than only 12 following the international prefix). This has expanded 1000-fold the quantity of numbers available in each country and so caters for the needs of the foreseeable future.
- The concept of *direct dialling-in (DDI)*. In DDI the last few digits at the end of the ISDN subscriber number are transferred to the customer's PBX or other equipment, so enabling the call to be completed direct to the recipient's desk telephone without the assistance of a human PBX 'switchboard' operator.
- The concept of *sub-addressing* (also called *network address extension, NAE*). A sub-address comprises up to 40 additional decimal digits on top of the ISDN number, allowing routing with established local area networks on customer premises at the distant end of a public ISDN.
- The concept of two-stage call set-up for the support of interworking. The very nature of integrated services digital networks (ISDNs) demands that different services can interwork over a common network. Sometimes, however, this is not possible without the use of a specialised interworking unit between the ISDN and the dedicated network. As Figure 29.6 shows, two-stage dialling can be invaluable in this case.

**Table 29.1** Country codes according to ITU-T recommendation E.164

Country code	Country	Country code	Country	Country code	Country
1	Canada	231	Liberia	27	South Africa
1	United States of America	232	Sierra Leone	290	Saint Helena
1809	Anguilla	233	Ghana	295	San Marino
1809	Antigua and Barbuda	234	Nigeria	296	Trinidad and Tobago
1809	Bahamas	235	Chad	297	Aruba
1809	Barbados	236	Central African Republic	298	Faroe Islands
1809	Bermuda	237	Cameroon	299	Greenland
1809	British Virgin Islands	238	Cape Verde	30	Greece
1809	Cayman Islands	239	Sao Tome and Principe	31	Netherlands
1809	Dominican Republic	240	Equatorial Guinea	32	Belgium
1809	Grenada	241	Gabonese Republic	33	France and Monaco
1809	Jamaica	242	Congo	34	Spain
1809	Montserrat	243	Zaire	350	Gibraltar
1809	Saint Kitts and Nevis	244	Angola	351	Portugal
1809	Saint Lucia	245	Guinea-Bissau	352	Luxembourg
1809	Saint Vincent and the Grenadines	246	Diego Garcia	353	Ireland
1809	Turks and Caicos Islands	247	Ascension	354	Iceland
20	Egypt	248	Seychelles	355	Albania
210	Morocco	249	Sudan	356	Malta
211	Morocco	250	Rwanda	357	Cyprus
212	Morocco	251	Ethiopia	358	Finland
213	Algeria	252	Somali Democratic Republic	359	Bulgaria
214	Algeria	253	Djibouti	36	Hungary
215	Algeria	254	Kenya	38	Yugoslavia
216	Tunisia	255	Tanzania	39	Italy
217	Tunisia	256	Uganda	40	Romania
218	Libya	257	Burundi	41	Switzerland and Liechtenstein
219	Libya	258	Mozambique	42	Czech and Slovak Republics
220	Gambia	259	Zanzibar	43	Austria
221	Senegal	260	Zambia	44	United Kingdom
222	Mauritania	261	Madagascar	45	Denmark
223	Mali	262	Reunion	46	Sweden
224	Guinea	263	Zimbabwe	47	Norway
225	Côte d'Ivoire	264	Namibia	48	Poland
226	Burkina Faso	265	Malawi	49	Germany
227	Niger	266	Lesotho	500	Falkland Islands
228	Togolese Republic	267	Botswana	501	Belize
229	Benin	268	Swaziland	502	Guatemala
230	Mauritius	269	Comoros		

Table 29.1 (continued)

Country code	Country	Country code	Country	Country code	Country
503	El Salvador	670	Mariana Islands	855	Kampuchea
504	Honduras	671	Guam	856	Lao People's Democratic Republic
505	Nicaragua	672	Australian External Territories	86	China
506	Costa Rica	673	Brunei Darussalam	87	<i>Maritime Mobile Service</i>
507	Panama	674	Nauru	880	Bangladesh
508	St Pierre and Miquelon	675	Papua New Guinea	90	Turkey
509	Haiti	676	Tonga	91	India
51	Peru	677	Solomon Islands	92	Pakistan
52	Mexico	678	Vanuatu	93	Afghanistan
53	Cuba	679	Fiji	94	Sri Lanka
54	Argentina	680	Palau	95	Myanmar
55	Brazil	681	Wallis and Futuna Islands	960	Maldives
56	Chile	682	Cook Islands	961	Lebanon
57	Colombia	683	Niue Island	962	Jordan
58	Venezuela	684	American Samoa	963	Syria
590	Guadeloupe	685	Western Samoa	964	Iraq
591	Bolivia	686	Kiribati	965	Kuwait
592	Guyana	687	New Caledonia and Dependencies	966	Saudi Arabia
593	Ecuador	688	Tuvalu	967	Yemen Arab Republic
594	Guiana	689	French Polynesia	968	Oman
595	Paraguay	690	Tokelan	969	Yemen, People's Democratic Republic
596	Martinique	691	Freestate of Micronesia	971	United Arab Emirates
597	Suriname	692	Marshall Islands	972	Israel
598	Uruguay	7	USSR	973	Bahrain
599	Netherlands Antilles	81	Japan	974	Qatar
60	Malaysia	82	South Korea	975	Kingdom of Bhutan
61	Australia	84	Vietnam	976	Mongolian People's Republic
62	Indonesia	850	North Korea	977	Nepal
63	Philippines	852	Hong Kong	98	Iran
64	New Zealand	853	Macao		
65	Singapore				
66	Thailand				

201 New Jersey	202 Washington DC	203 Connecticut
204 Manitoba (Canada)	205 Alabama	206 Washington
207 Maine	208 Idaho	209 California
212 New York City	213 Los Angeles	214 Dallas
215 Pennsylvania	216 Ohio	217 Illinois
218 Minnesota	219 Indiana	301 Maryland
302 Delaware	303 Colorado	304 West Virginia
305 Florida	306 Saskatchewan (Canada)	307 Wyoming
308 Nebraska	309 Illinois (311 Calling Card Service)	312 Chicago
313 Detroit	314 Missouri	315 New York
316 Kansas	317 Indianapolis	318 Louisiana
319 Iowa	401 Rhode Island	402 Nebraska
403 Alberta (Canada)	404 Georgia	405 Oklahoma City
406 Montana	408 San Jose	412 Pittsburgh
413 Springfield, Mass.	414 Milwaukee	415 San Francisco
416 Toronto (Canada)	417 Missouri	418 Quebec (Canada)
419 Ohio	501 Arkansas	502 Kentucky
503 Oregon	504 Louisiana	505 New Mexico
506 New Brunswick	507 Minnesota	509 Washington
510 4 Row TWX (USA) *	512 Texas	513 Ohio
514 Montreal (Canada)	515 Iowa	516 New York
517 Michigan	518 New York	519 Ontario (Canada)
601 Mississippi	602 Arizona	603 New Hampshire
604 British Columbia (Canada)	605 South Dakota	606 Kentucky
607 New York	608 Wisconsin	609 New Jersey
610 4 Row TWX (Canada) *	612 Minneapolis	613 Ottawa (Canada)
614 Columbus, Ohio	615 Tennessee	616 Michigan
617 Massachusetts	618 Illinois (700 Value Added Services)	701 North Dakota
702 Nevada	703 Virginia	704 North Carolina
705 Ontario (Canada)	707 California	709 Newfoundland
710 4 Row TWX (USA) *	712 Iowa	713 Houston, Texas
714 San Diego	715 Wisconsin	716 New York
717 Pennsylvania	718 New York City	800 800 Service ‡
801 Utah	802 Vermont	803 South Carolina
804 Virginia	805 California	806 Texas
807 Ontario (Canada)	808 Hawaii	809 Caribbean
810 4 Row TWX (USA) *	812 Indiana	813 Florida
814 Pennsylvania	815 Illinois	816 Kansas City
817 Fort Worth, Texas	819 Quebec (Canada)	900 900-Service ‡
901 Memphis	902 Nova Scotia (Canada)	903 Mexico
904 Florida	905 Mexico City	906 Michigan
907 Alaska	910 4 Row TWX (USA) *	912 Georgia
913 Kansas	914 New York	915 Texas
916 Sacramento	918 Oklahoma	919 North Carolina

\* TWX is the American equivalent of the Telex service, introduced by the Bell company in 1931. Telex was not available until Western Union introduced it in the 1950s.

‡ 800 and 900 service are described in Chapter 26.

Figure 29.4 The number plan of America

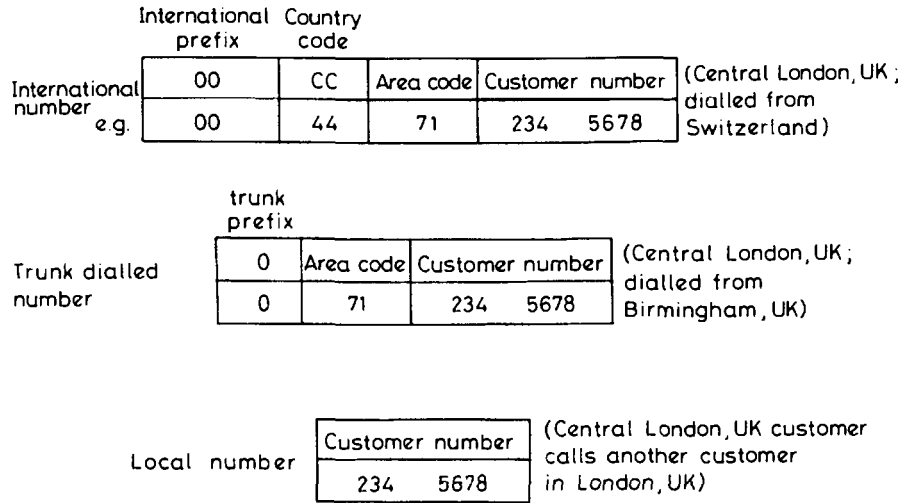


Figure 29.5 International, trunk and local numbers

Two-stage dialling may require the return of a second dial tone and the sending of extra digits after the completion of the first stage, and it can often be necessary on calls passing between two normal telephone network, either from or into some other specialized network.

## 29.2 INTERNATIONAL PUBLIC DATA NETWORK ADDRESS SCHEME

Network addresses in public data networks conform to the format set out in ITU-T recommendation X.121. These are used by the layer 3 protocol (OSI layer 3, e.g. X.25 packet level protocol) to identify customer or device connections to the network as we saw in Chapter 9. The X.121 address format is as shown in Figure 29.7.

An X.121 number (data network address) has a maximum length of 14 decimal digits, composed of a *data network identification code (DNIC)* and a *network terminal*

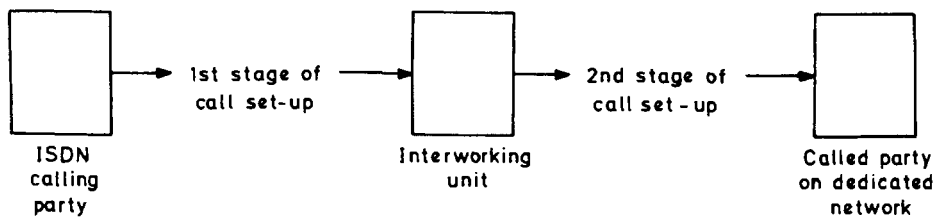


Figure 29.6 Two-stage call set-up

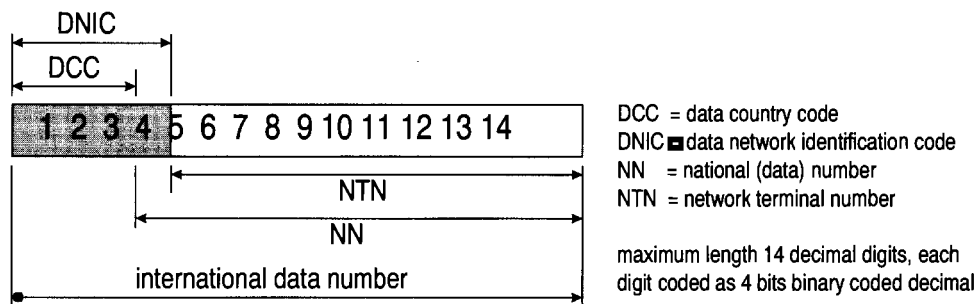


Figure 29.7 X.121 format for public data network addresses

number (NTN). The *DNIC* identifies a particular network and operator within a given country. This is composed of a three-digit *data country code* (DCC) and a one-digit *network identification code* (NIC). Following the *DNIC* is the *network terminal number* (NTN). *Data country codes* (DCC) are listed Table 29.2.

The *national data number* (NN) comprises the network identification digit (fourth digit) and the *network terminal number* (NTN). In some countries and networks it is only necessary to dial the *national number* for calls made within the network or country. However, for calls made to other data networks, the full *international data number* (DNIC plus NTN) must be dialled from the customer DTE.

The X.121 address of the destination port is carried in the *call setup packet* of the *X.25 packet level interface* (OSI layer 3, or *network layer protocol*) as we discussed in Chapter 18. It uniquely identifies a port connected to a public data network. The digit values of the address are always decimal, to ease their carriage across ISDN and other network types, but are coded in the *X.25 packet level interface* as four-bit, *binary coded decimal* (BCD) characters (Figure 18.9 of Chapter 18).

## 29.3 ESCAPE CODES

In some cases it is necessary when setting up calls across a network to identify the called address as belonging to a different numbering scheme than that usually used. Thus, for example, we learned in Chapter 10 how ITU-T recommendation X.31 defines the methodology for accessing data terminals connected to ISDNs from a port on a packet-switched public data network. In such cases, it is necessary to identify the ISDN port address as being an E.164-format (i.e. ISDN network) address. This is not the usual address format used within a public data network. The normal format is the X.121 format. Recommendation X.121, however, makes provision for the conversion of an E.164-format address to an X.121-format address. This is made simply by the addition of an *escape code* prefix. Thus the prefix '0' (when an X.121-format address is normally expected) signifies an E.164 number and a call requiring *escape* from the public data network to the ISDN using a digital connection. The prefix '9' signifies an E.164 number and a call requiring analogue *escape* from the public data network to the ISDN or telephone network. A further escape code '8' is also available within X.121 for an as yet undefined use.



**Table 29.2** Data Country Code (DCC) values defined by ITU-T recommendation X.121

Data country code (DCC)	Country	Data country code (DCC)	Country	Data country code (DCC)	Country
202	Greece	282	Georgia	401	Kazakhstan
204, 205	Netherlands	283	Armenia	404	India
206	Belgium	284	Bulgaria	410	Pakistan
208, 209	France	286	Turkey	412	Afghanistan
212	Monaco	288	Faroe Islands	413	Sri Lanka
214	Spain	290	Greenland	414	Myanmar
216	Hungary	292	San Marino	415	Lebanon
220	Yugoslavia	302	Canada	416	Jordan
222	Italy	303	Canada	417	Syria
225	Vatican City	308	St Pierre and Miquelon	418	Iraq
226	Romania	310–316	USA	419	Kuwait
228	Switzerland	330	Puerto Rico	420	Saudi Arabia
230	Czech and Slovak Republics	332	Virgin Islands	421	Yemen (Republic of)
232	Austria	334	Mexico	422	Oman
234–237	United Kingdom	338	Jamaica	423	Yemen (Republic of)
238	Denmark	340	French Antilles	424	United Arab Emirates
240	Sweden	342	Barbados	425	Israel
242	Norway	344	Antigua and Barbuda	426	Bahrain
244	Finland	346	Cayman Islands	427	Qatar
246	Lithuania	348	British Virgin Islands	428	Mongolia
247	Latvia	350	Bermuda	429	Nepal
248	Estonia	352	Grenada	430, 431	United Arab Emirates
250, 251	Russian Federation	354	Montserrat	432	Iran
255	Ukraine	356	St. Kitts	434	Uzbekistan
257	Belarus	358	St. Lucia	436	Tajikistan
259	Moldova	360	St. Vincent & the Grenadines	437	Kyrgyzstan
260	Poland	362	Netherlands Antilles	438	Turmenistan
262–265	Germany	364	Bahamas	440–443	Japan
266	Gibraltar	366	Dominica	450	South Korea
268	Portugal	368	Cuba	452	Vietnam
270	Luxembourg	370	Dominican Republic	453, 454	Hong Kong
272	Ireland	372	Haiti	455	Macao
274	Iceland	374	Trinidad and Tobago	456	Cambodia
276	Albania	376	Turks and Caicos Islands	457	Lao People's Democratic Republic
278	Malta	400	Azerbaijan	460	China
280	Cyprus			466	Taiwan

Table 29.2 (continued)

Data country code (DCC)	Country	Data country code (DCC)	Country	Data country code (DCC)	Country
467	North Korea	608	Senegal	641	Uganda
470	Bangladesh	609	Mauritania	642	Burundi
472	Maldives	610	Mali	643	Mozambique
480, 481	South Korea	611	Guinea	645	Zambia
502	Malaysia	612	Côte d'Ivoire	646	Madagascar
505	Australia	613	Burkina Faso	647	Reunion
510	Indonesia	614	Niger	648	Zimbabwe
515	Philippines	615	Togolese Republic	649	Namibia
520	Thailand	616	Benin	650	Malawi
525	Singapore	617	Mauritius	651	Lesotho
528	Brunei Darussalam	618	Liberia	652	Botswana
530	New Zealand	619	Sierra Leone	653	Swaziland
534	Northern Marianas	620	Ghana	654	Comoros
535	Guam	621	Nigeria	655	South Africa
536	Nauru	622	Chad	702	Belize
537	Papua New Guinea	623	Central African	704	Guatemala
539	Tonga		Republic	706	El Salvador
540	Solomon Islands	624	Cameroon	708	Honduras
541	Vanuatu	625	Cape Verde	710	Nicaragua
542	Fiji	626	Sao Tome and Principe	712	Costa Rica
543	Wallis and Futuna	627	Equatorial Guinea	714	Panama
	Islands	628	Gabon	716	Peru
544	American Samoa	629	Congo	722	Argentina
545	Kiribati	630	Zaire	724	Brazil
546	New Caledonia	631	Angola	730	Chile
547	French Polynesia	632	Guinea-Bissau	732	Colombia
548	Cook Islands	633	Seychelles	734	Venezuela
549	Western Samoa	634	Sudan	736	Bolivia
550	Micronesia	635	Rwanda	738	Guyana
602	Egypt	636	Ethiopia	740	Ecuador
603	Algeria	637	Somali Democratic	742	Guiana
604	Morocco		Republic	744	Paraguay
605	Tunisia	638	Djibouti	746	Suriname
606	Libya	639	Kenya	748	Uruguay
607	Gambia	640	Tanzania		

## 29.4 TELEX NETWORK NUMBERING PLAN (ITU-T F.69)

ITU-T recommendation F.69 sets out the numbering plan for public telex networks. A telex number is used within the telex network in a similar manner to the use of telephone numbers within a telephone network, to signal the desired telex port destination of calls at call setup time. A *international telex number* conforming to ITU-T F.69 has a maximum length of 12 decimal digits, of which the first two or three comprise the *telex destination code*, the telex equivalent of the *data network identification code* (DNIC) of recommendation X.121. The allocation of the codes is carried out by ITU according to the rough schedule of Table 29.3.

As well as having a unique *telex destination code*, each telex network is allocated a unique one or two letter (alphabetic character) *telex network identification code* (TNIC). This code is used as part of the *answerback* procedure. It appears on at the top of a telex being sent, to confirm the network to which the *connected line* (as opposed to the dialled line) is attached.

## 29.5 X.500: THE ADDRESSING PLAN FOR THE MESSAGE HANDLING SERVICE (MHS)

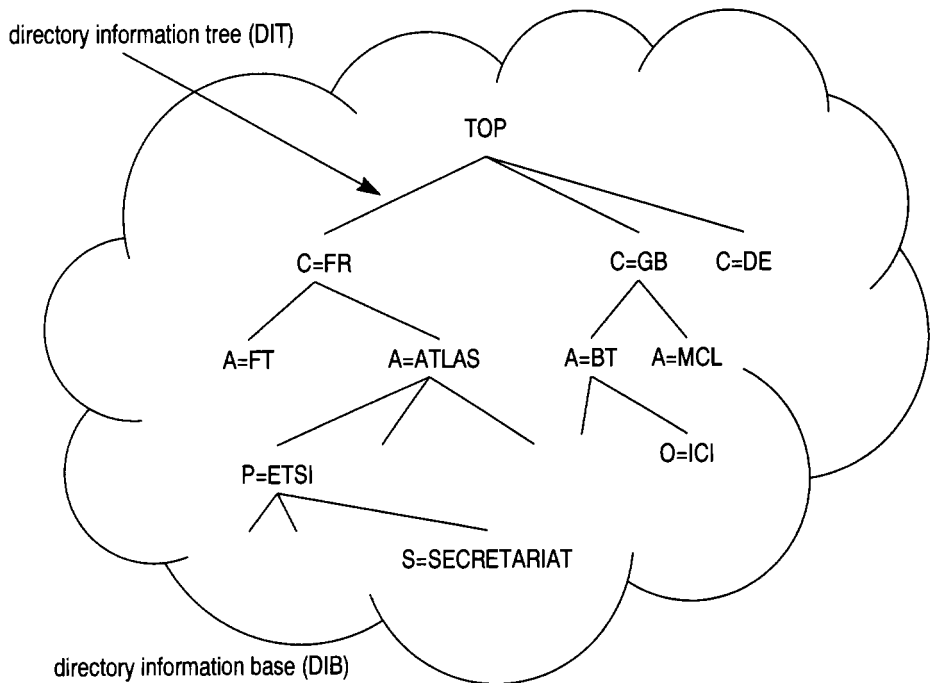
The addresses used in the *message handling service* (X.400, Chapter 23) conform to the OSI and ISO directory service, laid out in the ITU-T X.500 series of recommendations. Recommendation X.500 lays out the general form of standard addresses, conforming to a standard directory scheme. The addresses themselves are held in a *directory information base* (DIB) and they consist of a set of individual *directory information trees* (DIT). At the first layer of the tree is always the ISO-assigned two-digit *country identification* ('C='). At the next layer is the *administrative domain* ('A='). At the lower layers of individual trees, separate structures (correctly called *directory schemas*) may be used, as befitting the situation of the individual *administrations*.

Figure 29.8 illustrates the actual X.500 address of the ETSI secretariat (located in Sophia Antipolis in France). Written out in full, the address is

C = FR;	A = ATLAS;	P = ETSI;	S = SECRETARIAT
(country)	(administration)	(private domain)	(section)

**Table 29.3** Allocation of telex destination codes (ITU-T F.69)

First digit of code	Region
2	North America
3	South America
4, 5, 6	Europe and maritime services
7	Pacific
8	Middle East and Far East
9	Africa



**Figure 29.8** ITU-T recommendation X.500: A directory information base (DIB) and directory information tree (DIT)

Also shown in Figure 29.8 are other hypothetical addresses in Great Britain (GB) and Germany (DE). The imaginary British Telecom *domain* has been shown as if segregated first according to *organization* (*O*) rather than *private domain* (*P*).

## 29.6 INTERNET ADDRESSING SCHEME

Internet addresses are used to identify hardware ports and software applications within the worldwide *Internet*<sup>7</sup> computer network. They are used and interpreted by the *Internet protocol* (*IP*). The addresses comprise 32 bits when coded in binary form, but when written are usually quoted in a decimal form of four numbers, each of values between 0 and 255 (corresponding to four 8 bit numbers), separated by dots. Thus a typical *Internet address* is written 252.158.32.195.

Internet addresses are classified into five different types as shown in Figure 29.9. The different types of class address are assigned by agencies worldwide who administer the address plan. *Class A addresses* comprise a 1 byte *network address* assigned by the Internet numbering authority. The three remaining bytes are left for host addresses (i.e. computer addresses), which can be assigned 'locally' by the network operator (i.e. the *class A* address owner). Close to 17 million computers may be connected to a class A network. Class A addresses are only available to major Internet network operators.

*Class B addresses* comprise a two byte *network address* assigned by the numbering authority and a two byte *host address* space administered by the network operator. Up to 65 536 computers may be connected to a *class B* network. Class B addresses are generally only available to major corporations, who must be able to justify the extensive addressing allocation.

*Class C addresses* comprise a three byte *network address* assigned by the numbering authority and a 1 byte *host address* space. Up to 256 devices may be connected to a *class C* network. These are the types of addresses allocated to most operators of small scale networks connected to the *Internet*.

*Class D addresses* are used for broadcasting messages. Such messages will be received by all connected stations. *Class E addresses* are used only for experimental purposes.

In addition to the 32 bit *Internet address*, some networks additionally use a further 32 bit address field known as the *sub-network address*. This gives a huge scope for massively increasing the number of applications accessible within the connected end devices.

*Alias* addressing is also possible. Thus a *UNIX* server might have the *Class B internet address* 167.23.1.146 and the *alias* RS6000. Where *alias addressing* is used a file must exist on a *UNIX* server somewhere within the local network to convert the *alias address* to the 32 bit binary address. This file is called the '/etc/hosts' file and is part of the set up configuration of the server.

## 29.7 INTERNET E-MAIL (SMTP) ADDRESSES

Internet *e-mail addresses* should not be confused with ordinary *Internet addresses*. The difference is akin to the difference between an MHS (X.500) address and an X.25

	Byte 1	Byte 2	Byte 3	Byte 4
Class A	Network Address	.Host Address	.Host Address	.Host Address
Class B	Network Address	.Network Address	.Host Address	.Host Address
Class C	Network Address	.Network Address	.Network Address	.Host Address
Class D	Broadcast Address	.Broadcast Address	.Broadcast Address	.Broadcast Address
Class E	Experimental	.Experimental	.Experimental	.Experimental

Address type	Value of byte 1
Class A	0–127
Class B	128–191
Class C	192–223
Class D	224–239
Class E	240–255

**Figure 29.9** Internet address format and classification

network address. e-mail addresses are those recognized by the *SMTP* (*simple mail transfer protocol*). They typically have the form

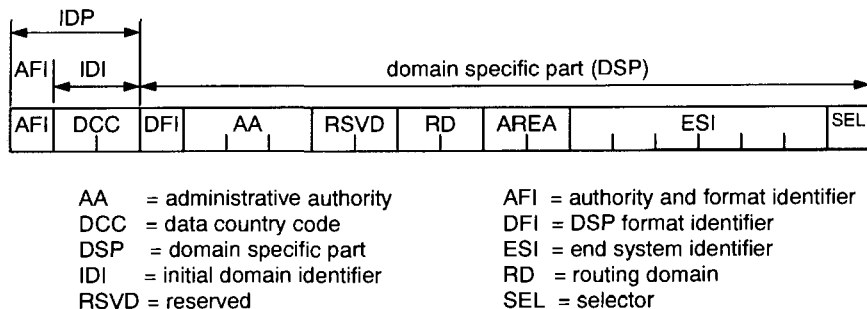
- 1789748.34543@compuserve.com
- martin.clark@onlineservice.de

The first of these forms is the actual *numerical address*. The part of the address after the @ symbol identifies the *mail server* or *post office* (the equivalent of a given mail operators 'domain'). The second form given above (i.e. the name form) is actually an *alias address*. On reaching the *mail server* the part of the address preceding the @ symbol has to be converted to the actual numerical address in order that the SMTP or local mail protocol can effect the final delivery. To perform the conversion a look-up table must exist at the destination mail server.

## 29.8 NETWORK ADDRESSING SCHEMES USED IN SUPPORT OF BROADBAND-ISDN AND ATM

Broadband-ISDN or ATM network addresses (correctly *ATM logical network addresses*) consist of three general parts, comprising nine separate fields as illustrated in Figure 29.10. The three main parts are the *AFI* (*authority and format identifier*), the *IDI* (*initial domain identifier*) and the *DSP* (*domain specific part*), where the AFI and IDI are known together as the *IDP* (*initial domain part*). The AFI allows for a number of different coding schemes to be used for the IDI, and determines which of these codings is in use.

The *NSAP* (*network service access point*) format defined by ISO 8348 and ITU-T recommendation X.213 allows the AFI to be set to indicate one of three forms of the IDI. In the first, *DCC* form (AFI=39) the IDI is the *data country code* (*DCC*) as defined by ISO 3166 (ITU-T recommendation X.121). In the second form (AFI=47), the *ICD* or *international code designator* form, the IDI identifies international organiza-



**Figure 29.10** ATM addressing plans (conformant to OSI NSAP format)

tions as defined by British Standards. In the third form ( $AFI = 45$ ), the IDI takes the form of a full E.164 international ISDN or telephone number, and the separate fields DFI and AA do not appear. Maybe a fourth form will appear using Internet addresses.

The *DFI* (*DSP format identifier*) defines the format (and therefore meaning) of the DSP (domain specific part). The *administrative authority* (*AA*) identifies the organization which administrates the addresses in this field (usually the public or private network operator). The *routing domain* (*RD*) together with the *area* field indicate the destination sub-network and area. The *ESI* (*end system identifier*) indicates the end device. The *selector* field is an extension of the address not used for routing within the network, but may be required to indicate to the end system which local device (e.g. telephone extension) is to be connected or which mode of communication should be expected (e.g. telephone or fax, etc.).

---

# Teletraffic Theory

---

Telecommunication networks, like roads, are said to carry ‘traffic’, consisting not of vehicles but of telephone calls or data messages. The more traffic there is, the more circuits and exchanges must be provided. On a road network the more cars and lorries, the more roads and roundabouts are needed. In any kind of network, if traffic exceeds the design capacity then there will be pockets of congestion. On the road this means traffic jams; on the telephone the frustrated caller receives frequent ‘busy tones’; in a data network unacceptably long ‘response times’ are experienced.

Short of providing an infinite number of lines, it is impossible to know in advance precisely how much equipment to build into a telecommunications network to meet demand without congestion. However, there is a tool for ‘dimensioning’ network links and exchanges. It is the rather complex statistical science of ‘teletraffic theory’ (sometimes called ‘teletraffic engineering’). This is the subject of this chapter.

We begin with the teletraffic dimensioning method used for circuit-switched networks, first published in 1917 by a Danish scientist, A. K. Erlang. Erlang defined a number of parameters and developed a set of formulae, which together give a framework of rules for planners to design and monitor the performance of telephone, telex and circuit-switched data networks. The latter part of the chapter deals with the dimensioning of data networks, reviewing the techniques in such a way as to offer the reader a practical method of network design.

## 30.1 TELECOMMUNICATIONS TRAFFIC

*Traffic* is the term to describe the amount of telephone calls or data messages conveyed over a telecommunications network, but this could cover any number of different scientific definitions. Possible definitions of *traffic* include

- the total number of calls or messages
- the total *conversation time* (i.e. the number of calls multiplied by the *conversation time* on each)
- the total circuit *holding time*. This is the number of calls multiplied by the *holding time* on each; the *holding time* includes the time period during which the parties are



in conversation and also the time prior to conversation when the call is being *set-up*; holding time is the total time for which the network is in use

- the total number of data characters conveyed

All the above statistics impact on the performance of telecommunication networks, but the *holding time* is particularly influential to the carrying capacity and congestion of telephone and other *circuit-switched networks*. As far as the science of teletraffic is concerned, the traffic volume is normally defined by the third definition above. In other words, the total *traffic volume* is equal to the total network holding time. The traffic volume, however, cannot be directly used in the determination of exactly how many circuits on a trunk or what size of exchanges will need to be provided. What we need is some idea of the maximum usage of the network at any one time. For this reason, it is normal to measure the traffic intensity of circuit-switched networks.

## 30.2 TRAFFIC INTENSITY (CIRCUIT-SWITCHED NETWORKS)

The *traffic intensity* of a *circuit-switched network* is defined to be the average number of calls simultaneously in progress during a particular period of time. It is measured in units of *Erlangs*. Thus an average of one call in progress during a particular period would represent a *traffic intensity* of one Erlang. The traffic intensity on any route between two exchanges can also be quoted in Erlangs. It is measured by first summing the total holding time of all the circuits within the route and then dividing this by the time period  $T$ , over which the measurement was made.

In some countries, including the United States, *traffic intensity* is measured not in *Erlangs* but in units called *CCS* (*hundred call seconds*). *CCS* is a measure of the total call *holding time* during the network or route busy hour. The two units, *CCS* and Erlang are very simply related because

$$1 \text{ Erlang} = 3600 \text{ call seconds} = 36 \text{ CCS}$$

The definition of *traffic intensity* is not restricted to traffic between exchanges. Cross-exchange traffic (that passing across an exchange from incoming ports to outgoing ports) can also be measured and quoted in Erlangs.

As an example, the route between exchanges A and B in Figure 30.1 consists of five circuits. The chart in Figure 30.1 also shows the periods of usage of each of these circuits during a 10 minute period.

It will be seen from the chart in Figure 30.1 that the total number of minutes of circuit usage during the 10 minute period was 35 minutes. This is the sum of the individual circuit holding times, respectively: 6 min, 7.5 min, 7.5 min, 8 min, 6 min. Dividing the total of 35 minutes by the length of the monitoring period (10 minutes) we may deduce that the *traffic intensity* on the route was 3.5 Erlangs. In other words, an average of 3.5 circuits were in use throughout the whole of the period.

However, usually the challenge facing a telecommunications network planner is how many circuits to provide to carry a given volume of traffic. The general principle is that more circuits will be needed on a route than the numerical value of the route traffic intensity measured in *Erlangs*. It would be easy to succumb to the mistaken belief that a

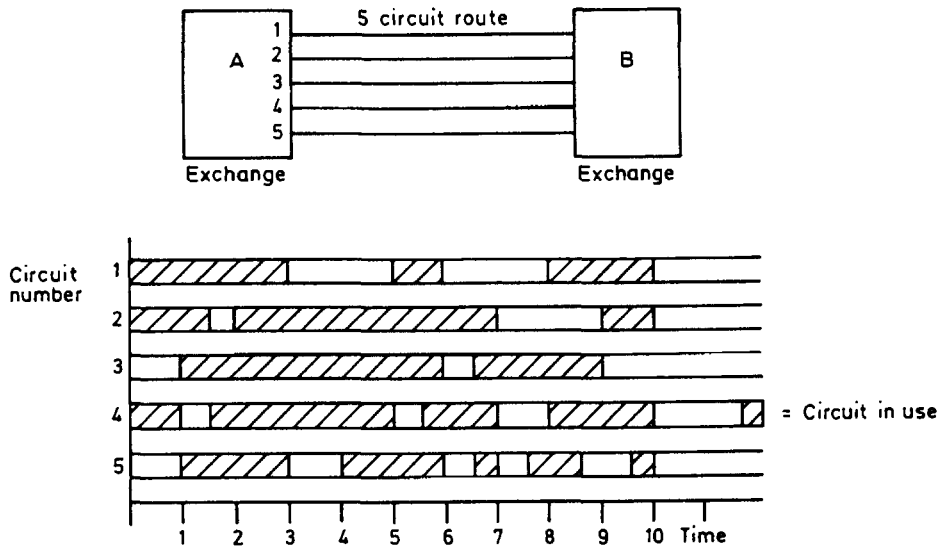


Figure 30.1 Circuit usage on a route between exchanges

traffic intensity of 3.5 Erlangs (3.5 simultaneous calls) could be carried on four circuits. Figure 30.1 clearly illustrates an example in which this is not the case. All five circuits were simultaneously in use twice, between times 2 and 3, and between times 5.5 and 6.

Here we must digress briefly to examine the concepts of *offered* and *carried* traffic, though their names may give them away. *Offered traffic* is a theoretical concept. It is a measure of the unsuppressed *traffic intensity* that would be transported on a particular route if all the customers' calls were connected without congestion. *Carried traffic* is that resultant from the carried calls, and it is the value of traffic intensity actually measured. For a network without congestion the *carried traffic* is equal to the *offered traffic*. However, if there is congestion in the network, then the offered traffic will be higher than that carried, the difference being the calls which cannot be connected.

Teletraffic theory leads to a set of tables and graphs which enable the required network circuit numbers to be related to the *offered traffic* in Erlangs (i.e. the demand).

### 30.3 PRACTICAL TRAFFIC INTENSITY (ERLANG) MEASUREMENT

The *traffic intensity* on a given route may be measured using one of two main methods, either by sampling or by an absolute measurement. The latter method was used in the example of Figure 30.1. Historically, however, electro-mechanical exchanges did not lend themselves to easy calculation of the absolute value of the total holding time. Instead, it was common practice to use a sampling technique. We can obtain an estimate of the total circuit holding time by looking at the instantaneous state of the circuits at a number of sample points in time. If we take ten samples, after  $\frac{1}{2}$  min,  $1\frac{1}{2}$  min,

**Table 30.1** Samples of number of circuits in use

Elapsed time	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$4\frac{1}{2}$	$5\frac{1}{2}$	$6\frac{1}{2}$	$7\frac{1}{2}$	$8\frac{1}{2}$	$9\frac{1}{2}$
Circuits in use	3	5	5	3	4	5	4	2	4	4

$2\frac{1}{2}$  min, ...,  $9\frac{1}{2}$  min, we could assume that this was representative of the respective time periods, 0–1 min, 1–2 min, ..., etc. Thus, because after  $\frac{1}{2}$  minute has elapsed, we find that three circuits are in use; we assume that three circuits will be in use for the whole period of the first minute. By this method we obtain 10 ‘snapshots’ of the number of circuits in use, corresponding to the 10 individual one minute periods which we sampled. The results are as shown in Table 30.1.

Averaging the sample values will give us an estimate of the traffic intensity over the whole period. This value comes out at an estimated 3.9 Erlangs. This is calculated as the sum of the ten sample values (39), divided by the overall duration (10). Compare this with the actual value of 3.5 Erlangs. The error arises from the fact that we are only sampling the circuit usage rather than using exact measurement. The error can be reduced by increasing the frequency of samples. Table 30.2 gives the values calculated at a  $\frac{1}{2}$  minute (as opposed to a 1-minute) *sampling rate* (sometimes also called *scan rate*).

The increased sampling rate of Table 30.2 estimates the traffic correctly as 3.5 Erlangs ( $70 \times 0.5/10$ ). The exactness of the estimate on this occasion is fortuitous. Nonetheless, the principle is well illustrated that too slow a sampling rate will produce unreliable traffic intensity estimates and that the estimate is improved in accuracy by a higher sampling rate. A good guide is that the sample period length should be no more than about one-third of the *average call holding time*. In the example of Figure 30.1, the average call holding time is 35 min/17 calls = 2.06 min, so a sample should be taken at least every 0.7 min to derive a reliable estimate of the traffic intensity. The minimum monitoring period should be about three times the average call holding time. This helps us to avoid unrepresentative peaks or troughs in call intensity.

Nowadays, *stored program controlled* (SPC) (i.e. computer-controlled) exchanges have made the absolute measurement of call holding times relatively easy. So today many exchanges produce the measurements of traffic intensity which are exact. They do so by calculating the value using data records storing the start and end time of each individual call or connection.

**Table 30.2**  $\frac{1}{2}$  minute scan rate

Elapsed time	$\frac{1}{4}$	$\frac{3}{4}$	$1\frac{1}{4}$	$1\frac{3}{4}$	$2\frac{1}{4}$	$2\frac{3}{4}$	$3\frac{1}{4}$	$3\frac{3}{4}$	$4\frac{1}{4}$	$4\frac{3}{4}$
Circuits in use	3	3	4	4	5	5	3	3	4	4
Elapsed time	$5\frac{1}{4}$	$5\frac{3}{4}$	$6\frac{1}{4}$	$6\frac{3}{4}$	$7\frac{1}{4}$	$7\frac{3}{4}$	$8\frac{1}{4}$	$8\frac{3}{4}$	$9\frac{1}{4}$	$9\frac{3}{4}$
Circuits in use	4	5	2	4	1	2	4	3	3	4

## 30.4 THE BUSY HOUR

In practice, telecommunications networks are found to have a discernible *busy hour*. This is the given period during the day when the *traffic intensity* is at its greatest. Traditionally measurements of traffic intensity have been made over a full 60-minute period at the *busy hour* of day, to calculate the *busy hour traffic* (a shortened term for the *busy hour traffic intensity*). By taking samples over a few representative days, future *busy hour traffic* can usually be predicted well enough to work out what the equipment quantities and route sizes of the network should be. The *dimensioning* is made by using the predicted busy hour traffic as an input to the *Erlang formula* (presented later in this chapter). When that formula has pronounced the scale of circuits and equipment that are going to be needed at the busy hour, we can feel secure at less hectic times of day.

In more recent times it has been found that exchanges are developing more than one busy hour; maybe as many as three, including morning, afternoon and evening *busy hours*. The morning and afternoon busy hours are usually the result of business traffic. The evening one results from residential, international traffic and nowadays also from *Internet* access traffic (dial-up connections from home PC-users to *Internet servers* and bureaux. Some examples of daily traffic distribution are given in Figure 30.2.

The first distribution in Figure 30.2 is a typical business-serving exchange, with morning and afternoon busy hours. The second example shows the traffic in a residential exchange, where morning and afternoon *busy hours* are less than the evening busy

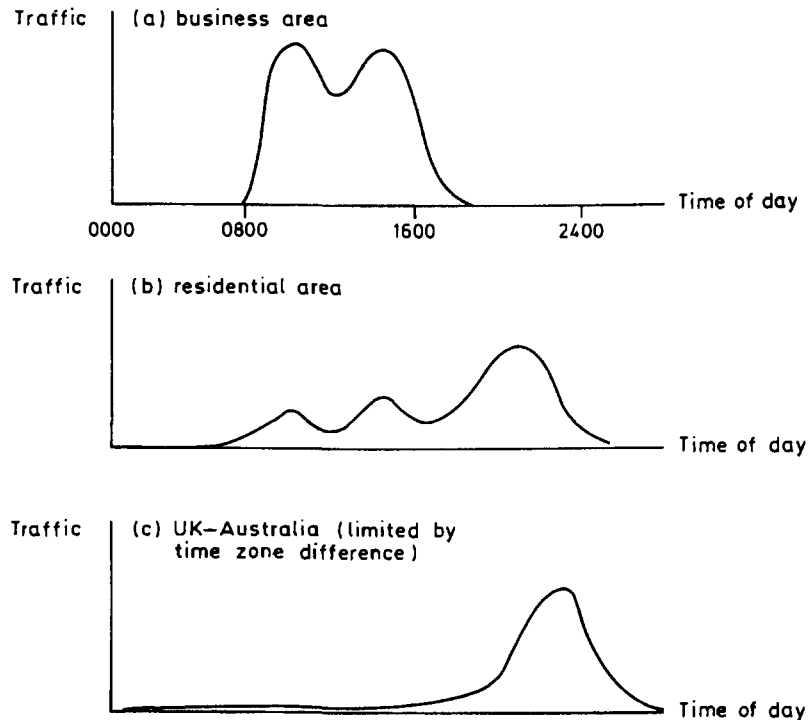


Figure 30.2 Some typical daily traffic distributions

hour. The third example is of an international route, where there is often a single peak, constrained in duration by the time zone difference between the countries.

Multiple busy hours necessitate the calculation of separate busy-hour traffic values for each individual busy hour, and design of the network to meet each one. It is not enough to use only one of the sets of busy-hour traffic values. In Figure 30.2 for example, the overall effect of distributions (a) and (b), when combined, is a two-humped pattern with overall busy hours in morning and afternoon of approximately equal magnitude. However, if we were to design the whole network on the basis of its morning and afternoon busy hour traffic values alone, we would not provide enough circuits at the residential exchange to meet its evening busy hour. Normal practice is therefore to dimension each exchange matrix according to its own *exchange busy hour*, and each route according to its individual *route busy hour*.

Another point to consider when dimensioning networks is that a route or exchange is unlikely to be equally busy throughout the entire 60 minute busy hour period. To meet the peak demand, which might be significantly higher than the hour's average traffic, it is sometimes convenient to redefine the busy hour as being of less than 60 minute duration. It may seem paradoxical to have a busy hour of less than 60 minutes, but Figure 30.3 illustrates a case of traffic which has a short duration peak of between 15 and 30 minutes.

In the example of Figure 30.3, if we were to use a 60 minute busy hour measurement period, we would estimate a busy-hour traffic of value  $Q$  as shown in the diagram. This would be a gross underestimate of the actual traffic peak, and would guarantee that the busiest period would be heavily congested. By redefining the *busy hour* to be of only 30 min duration we obtain an estimate  $P$  which is much nearer the actual traffic peak. Surprisingly, however, the use of much shorter busy hour periods and attempts to pinpoint peaks of traffic which last only a few minutes are not really important; customers who suffer congestion at this time are more than likely to get through on a repeat call attempt within a few minutes anyway!

To sum up, it is usual to monitor *exchange busy hour* and *route busy hour* traffic values as a gauge of current customer usage and network capacity needs. Future network design and dimensioning can then be based on our forecasts of what the values of these parameters will be in the future (forecasting methods are covered in Chapter 31).

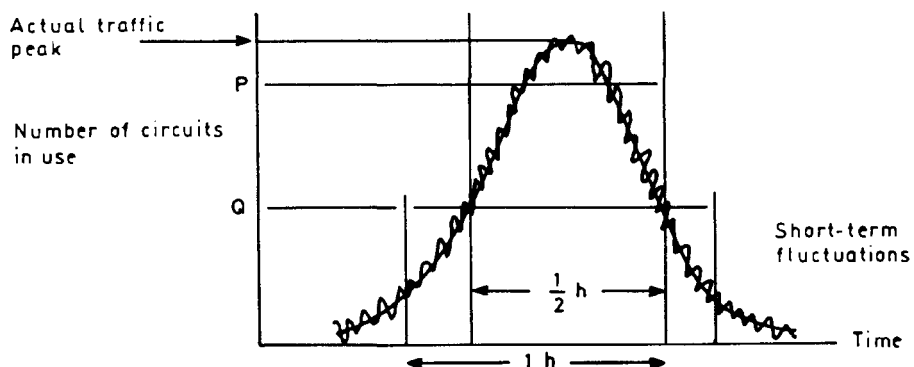


Figure 30.3 A short duration busy hour

### 30.5 THE FORMULA FOR TRAFFIC INTENSITY

Our next step is to develop the formula for traffic intensity, as a basis for subsequent discussion of the Erlang method of network dimensioning. Recapping in mathematical terms, the traffic intensity is given by the expression

$$\text{Traffic intensity} = \frac{\text{the sum of circuit holding times}}{\text{(carried traffic) the duration of the monitoring period}}$$

Now let

$A$  = the traffic intensity in Erlangs

$T$  = the duration of the monitoring period

$h_i$  = the holding time of the  $i$ th individual call

$c$  = the total number of calls in the period of mathematical summation

Then, from above

$$A = \frac{\sum_{i=1}^c h_i}{T}$$

Now, because the sum of the holding times is equal to the number of calls multiplied by the average holding time, then

$$\sum_{i=1}^c h_i = ch$$

where  $h$  = average call holding time, and therefore

$$A = \frac{ch}{T}$$

It is interesting to calculate the call arrival rate, in particular the number of calls expected to arrive during the average holding time. Let  $N$  be this number of calls, then

$$\begin{aligned} N &= \text{no of call arrivals during a period equal to the average holding time} \\ &= h \times \text{call arrival rate per unit of time} \\ &= h \times c/T \\ &= ch/T = A \end{aligned}$$

In other words, the number of calls expected to be generated during the average holding time of a call is equal to the traffic intensity  $A$ . This is perhaps a surprising result, but one which sometimes proves extremely valuable.

### 30.6 THE TRAFFIC-CARRYING CAPACITY OF A SINGLE CIRCUIT

In this section we discuss the traffic-carrying capacity of a single circuit. This leads on to a mathematical derivation of the Erlang formula, which is the formal method to calculate the traffic-carrying capacity of a circuit group of any size.

For our explanation let us assume we have provided an infinite number of circuits, laid out in a line or *grading*, as shown in Figure 30.4. The infinite number provides enough circuits to carry any value of traffic intensity. Now let us further assume that each new call scans across the circuits from the left-hand end until it finds a free circuit. Then let us try to determine how much traffic each of the individual circuits carries.

First, let us consider circuit number one. Figure 30.5 shows a timeplot of the typical activity we might expect on this circuit, either *busy* carrying a call, or *idle* awaiting for another call to *arrive*. The timeplot of Figure 30.5 starts with the arrival of the first call. This causes the circuit to become busy for the duration of the call. While the circuit is busy a number of other calls will arrive, which circuit number one will be incapable of carrying. These other calls will scan across towards a higher-numbered circuit (circuit number two, then three, and so on) until the first free circuit is found. Finally, at the end of the call on circuit number one, the circuit will be returned to the idle state. This state will prevail until the next new call arrives.

Let us try to determine the proportion of time for which circuit number one is busy. For this purpose, let us assume each call is of a duration equal to the average call holding time  $h$ . This is not mathematically rigorous but it makes for simpler explanation. Let us also invent an imaginary cycle of activity on the circuit.

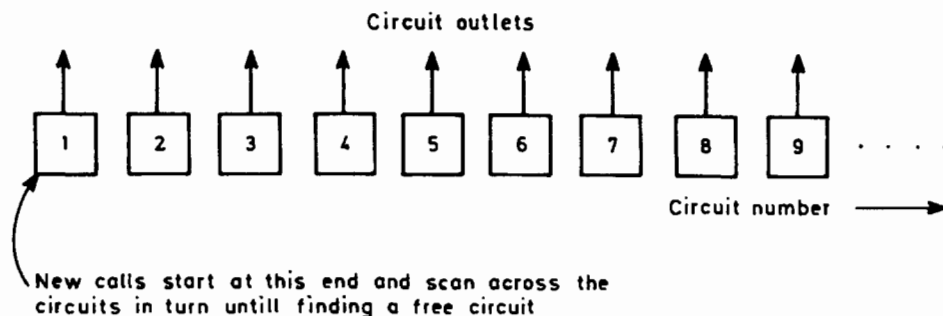


Figure 30.4 Scanning for a free circuit

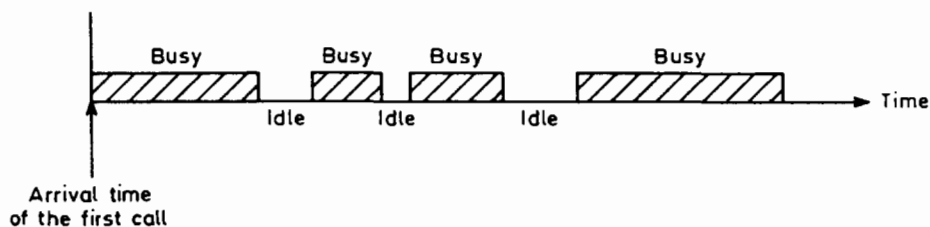


Figure 30.5 Activity pattern of circuit number 1

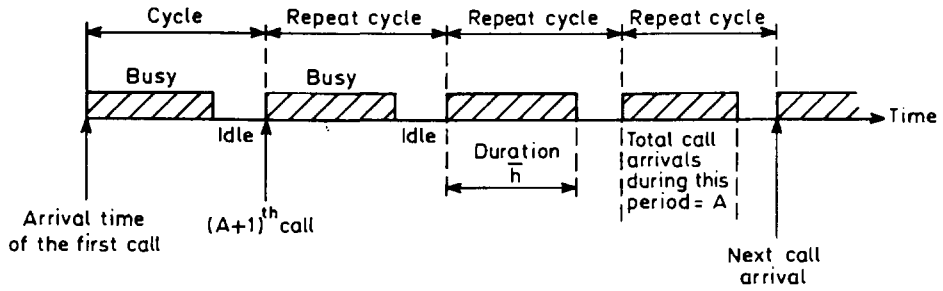


Figure 30.6 Average activity cycle on circuit number 1

Our imaginary cycle is as follows.

After the arrival of the first call, we expect circuit number one to be busy for a period of time equal to  $h$ . As we learned in the last section we can expect a total of  $A$  calls to arrive during the average holding time, where  $A$  is the offered traffic.  $(A - 1)$  of these calls (i.e. all but the first) will scan over circuit number one to find a free circuit among the higher numbered circuits. At the end of the first call circuit number one will be released, and an *idle* period will follow until the next call arrives (i.e. the  $(A + 1)$ th). We can imagine this cycle repeating itself over and over again.

As we can see from our imaginary 'average' cycle, shown in Figure 30.6, the total number of calls arriving during the cycle is  $A + 1$ . The total duration of the cycle is therefore

$$h \times \frac{A + 1}{A} = \frac{h(A + 1)}{A}$$

We also know that circuit number one is busy during each cycle for a period of duration  $h$ . Therefore the average proportion of the time for which circuit number one is busy is given by

$$\text{occupancy of circuit number 1} = h \times \frac{A}{h(1 + A)} = \frac{A}{1 + A}$$

This value is the so-called *circuit occupancy* of circuit number one. It is numerically equal to the average number of calls in progress on circuit number one, and is therefore the intensity of the traffic *carried* on circuit number one, and is measured in Erlangs accordingly. Thus if one Erlang were offered to the grading ( $A = 1$ ), then the first circuit would carry half an Erlang. The remaining half Erlang is carried by other circuits.

Taking one last step, if we assume that new calls *arrive* at random instants of time, then the proportion of calls rejected by circuit number one is equal to the proportion of time during which the circuit is busy, i.e.  $A/(1 + A)$ . In our simple case, if only one circuit were available, then  $A/(1 + A)$  proportion of calls cannot be carried. This is called the blocking ratio  $B$ , and is usually written

$$B(\text{blocking ratio - for one circuit}) = \frac{A}{1 + A}$$

where  $A$  = offered traffic.



Though not proven above in a mathematically rigorous fashion, the above result is the foundation of the Erlang method of circuit group dimensioning. Before going on, however, it is worth studying some of the implications of the formula a little more deeply for two cases.

First, take the case of one Erlang of offered traffic to a single circuit. Substituting in our formula  $A = 1$ , we conclude that the blocking value is  $1/(1 + 1) = 1/2$ . In other words, half of the calls fail (meeting congestion), and only half are carried. This confirms our earlier conclusion that the circuit numbers needed to carry a given intensity of traffic are greater than the numerical value of that traffic.

With traffic intensity of  $A = 0.01$ , then the proportion of blocked calls would have been only  $0.01/1.01 (=0.01)$ , or about one call in one-hundred blocked. This is the proportion of lost calls targetted by many network operating companies. Put in practical terms, the carrying capacity of a single circuit in isolation is around only 0.01 Erlangs.

Next let us consider a very large traffic intensity offered to our single circuit. In this case most of the traffic is blocked (if  $A = 99$ , then the formula states that 99% blocking is incurred, i.e. is not carried by our particular circuit). However, the corollary is that the traffic carried by the circuit (equal to the proportion of the time for which the circuit is busy) is 0.99 Erlangs. In other words the circuit is in use almost without let-up. This is what we expect, because as soon as the circuit is released by one caller, a new call is offered almost immediately.

In the appendix the full Erlang *lost call formula* is derived using a more rigorous mathematical derivation to gain an insight into the traffic-carrying characteristics of all the other circuits in Figure 30.4. For the time being, however, Figure 30.7 simply states the formula.

To confirm the result from our previous analysis, let us substitute  $N = 1$  into the formula of Figure 30.5. As before, we obtain a proportion of lost calls for a single circuit (offered traffic  $A$ ) of

$$B(1, A) = \frac{A/(1 + A)}{1} = \frac{A}{1 + A}$$

$$\begin{array}{l} E(N, A) \\ \text{or} \\ B(N, A) \end{array} = \frac{A^N}{N!} \left/ \left( 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^N}{N!} \right) \right.$$

where

$E(N, A)$  = proportion of lost calls, and probability of blocking

$A$  = offered traffic intensity

$N$  = available number or circuits

$N!$  = factorial  $N$

Figure 30.7 The Erlang lost-call formula

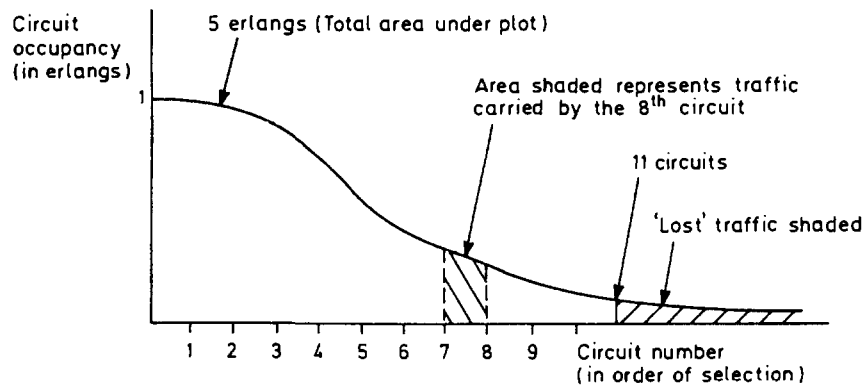


Figure 30.8 Circuit occupancies

This of course is also equal to the *circuit occupancy* (the traffic carried by it). The advantage of our new formula is that we may now calculate the occupancy of all the other circuits of Figure 30.4. By calculating the lost traffic from two circuits we can derive the carried traffic. Subtracting the traffic carried on circuit number one we end up with that carried by circuit number two. In a similar manner, the traffic carrying contributions of the other circuits can be calculated. Eventually we are able to plot the graph of Figure 30.8, which shows the individual circuit occupancies when 5 Erlangs of traffic is offered to an infinite circuit *grading*.

As expected, the low-numbered circuits carry nearly 1 Erlang and are in near-constant use, whereas higher numbered circuits carry progressively less traffic. The traffic carried by the first eleven circuits is also shown. From the formula of Figure 30.7, this is the number of circuits needed to guarantee less than 1% proportion of calls lost. Thus the right-hand shaded area in Figure 30.10 represents the small proportion of lost calls if 11 circuits are provided. The ability to calculate individual circuit occupancies is crucial to grading design (see Chapter 6).

## 30.7 DIMENSIONING CIRCUIT-SWITCHED NETWORKS

The future circuit requirements for each route of a circuit-switched network (i.e. telephone, telex, circuit switched data) may be determined from the Erlang lost call formula. We do so by substituting the predicted *offered* traffic intensity  $A$ , and using trial-and-error values of  $N$  to determine the value which gives a slightly better performance than the target blocking or *grade of service B*. A commonly used grade of service for interchange traffic routes is 0.01 or 1% blocking.

It is not an easy task by direct calculation to determine the value of  $N$  (circuits required), and for this reason it is usual to use either a suitably programmed computer or a set of *traffic tables*.

In recent years, numerous authors and organizations have produced modified versions of the Erlang method, more advanced and complicated techniques intended

to predict accurately the traffic-carrying capacity of various sized circuit groups for different *grades of service*. All have their place but in practice it comes down to finding the most appropriate method by trying several for the best fit for given circumstances. In my own experience, the extra effort required by the more refined and complicated methods of dimensioning is unwarranted. In practice the traffic demand may vary greatly from one day or month to the next and the practicality is such that circuits have to be provided in whole numbers, often indeed in multiples of say 12 or 30. The decision then is whether 1 or 2, 12 or 24, 30 or 60 circuits should be provided. It is rather academic to decide whether 23 or 24 circuits are actually necessary when at least 30 will be provided.

Table 30.3 illustrates a typical traffic table. The one shown has been calculated from the Erlang lost-call formula. Down the left hand column of the table the number of circuits on a particular route are listed. Across the top of the table various different grades of service are shown. In the middle of the table, the values represent the maximum offered Erlang capacity corresponding to the route size and grade of service chosen. Thus a route of four circuits, working to a design grade of service of 0.01, has a maximum offered traffic capacity of 0.9 Erlangs.

We can also use Table 30.3 to determine how many circuits are required to provide a 0.01 grade of service, given an offered traffic of 1 Erlang. In this case the answer is five circuits. The maximum carrying capacity of five circuits at 1% grade of service is 1.4 Erlangs, slightly greater than needed, but the capacity of four circuits is only 0.9 Erlangs.

The problem with traffic routes of only a few circuits is that only a small increase in traffic is needed to cause congestion. It is good practice therefore to ensure that a minimum number of circuits (say five) are provided on every route.

**Table 30.3** A simple Erlang traffic table

Number of circuits	Grade of service ( $B(N, A)$ )			
	1 lost call in			
	50 (0.02)	100 (0.01)	200 (0.005)	1000 (0.001)
	Erlangs	Erlangs	Erlangs	Erlangs
1	0.020	0.010	0.005	0.001
2	0.22	0.15	0.105	0.046
3	0.60	0.45	0.35	0.19
4	1.1	0.9	0.7	0.44
5	1.7	1.4	1.1	0.8
6	2.3	1.9	1.6	1.1
7	2.9	2.5	2.2	1.6
8	3.6	3.2	2.7	2.1
9	4.3	3.8	3.3	2.6
10	5.1	4.5	4.0	3.1
11	5.8	5.2	4.6	3.6
12	6.6	5.9	5.3	4.2

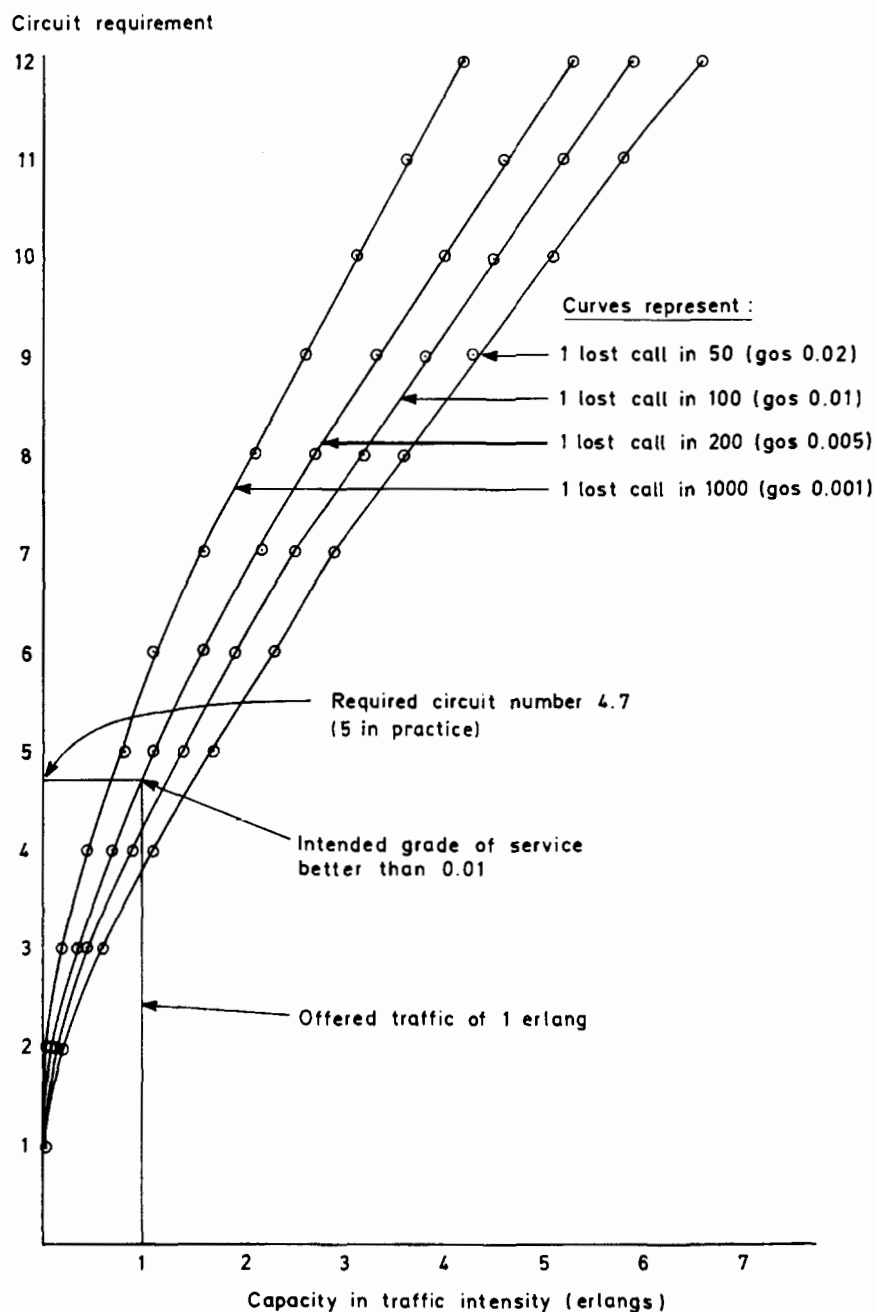


Figure 30.9 Graphical representation of the Erlang formula

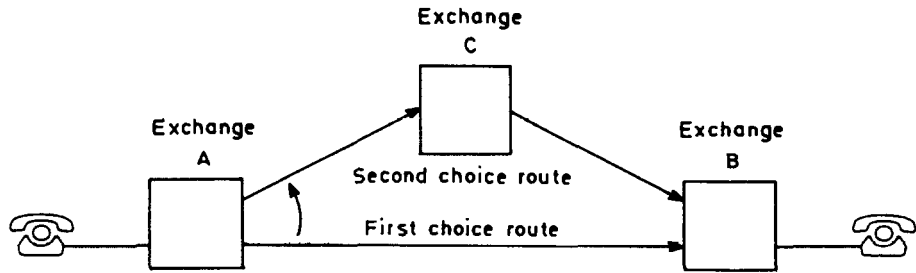


Figure 30.10 Overflow routing

The information held in Table 30.3 is sometimes presented graphically, and Figure 30.9 illustrates this. The traffic offered in Erlangs is usually plotted along the horizontal axis, and the circuit numbers up the vertical axis. A number of different curves then correspond to different grades of service. To determine how many circuits are required for a given offered traffic, the offered Erlang value is read along the horizontal axis, then a vertical line is drawn upwards to the curve corresponding to the required grade of service, and a horizontal line is drawn from this point to the vertical axis, where the circuit requirement can be read off. Our earlier example is repeated on the figure, confirming that five circuits are needed to carry 1 Erlang at 1% grade of service.

The traffic-to-circuit relationship shown in Table 30.3 and Figure 30.9, and the Erlang lost call formula on which they are based are only suitable for dimensioning *full-availability* networks (as defined in Chapter 6). For *limited availability* networks, slightly different formulae and traffic tables must be used. Furthermore, another slightly different set of traffic tables is called for when dimensioning networks which use overflow routing. Such tables are available from various publishers. Figure 30.10 shows an example of overflow routing in which traffic from exchange A to exchange B first tries to route directly, but is allowed to *overflow* via exchange C if all direct circuits are busy. Only the link A–B in this network can be dimensioned by the simple Erlang lost-call formula (provided we specify an overflow grade of service, e.g. 10% of calls), but as we shall find out in Chapter 32, only a slight modification to the method is necessary to permit us to dimension all the other links.

### 30.8 EXAMPLE ROUTE DIMENSIONING

As an example of the use of the Erlang lost-call formula in route dimensioning, let us calculate the number of circuits required to carry offered traffic  $A = 55$  Erlangs at a grade of service  $B = 0.01$ .

We shall show in Chapter 32 that an estimate of the number of circuits  $B$  required for 1% grade of service is given by the formulae

$$N = 6 + (A/0.85) \quad \text{where } A < 75$$

$$N = 14 + (A/0.97) \quad \text{where } 75 < A < 400$$

$$N = 29 + A \quad \text{where } 400 < A$$

so that for our case,  $A = 55$

$$N(\text{approx}) = 70 \text{ circuits}$$

So far so good, but from here on we rely on trial and error and the formula in Figure 30.7.

$$\text{substituting } A = 55, N = 70 \quad \text{we get } B = 0.007$$

This is slightly better than we need, so let us try

$$A = 55, N = 69. \quad \text{This time } B = 0.009$$

Again

$$A = 55, N = 68, \quad \text{then } B = 0.012$$

68 circuits give a slightly worse grade of service than we need, so 69 circuits must be provided!

## 30.9 CALL WAITING SYSTEMS

Whereas in telephone and other circuit switched networks the dimensioning is carried out to a given *grade-of-service* (or call loss), this technique is not suitable for dimensioning all types of network. Data networks and *server*-based services (such as *Internet* bureaux and operator switchrooms or *telemarketing* bureaux) are examples of queuing (i.e. *buffer*) or *call-waiting systems*. A *queue* (or *buffer*) minimizes, and may even eliminate, any loss of offered traffic. As we saw in Chapter 9, arriving packets from the separate *virtual connections* can be queued (*buffered*) in order until the line is free. Similarly, a caller accessing an Internet service must wait a little longer at times of heavy traffic demand on the network, but will be answered eventually.

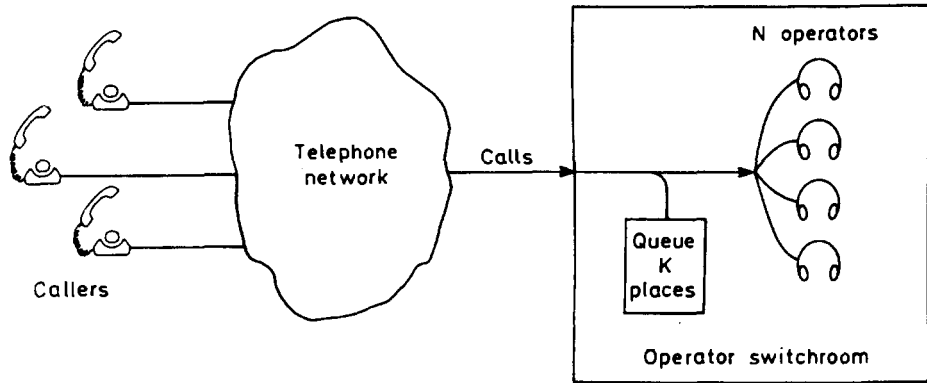
In the case of operator services or call-in telemarketing bureaux, incoming callers wait in a queue listening to music or a recorded message until a human operator becomes free.

In such networks the dimensioning techniques must be oriented towards ensuring that the capacity of a datalink, the number of servers or the number of human operators is sufficient to keep the queue waiting time within acceptable bounds. If there are insufficient operators the queue and the consequent waiting time will get longer and longer. We must also ensure that the queue capacity is long enough for all the people waiting or the data packets being temporarily stored in the *buffer*.

Consider the operator switchroom or telemarketing bureaux first, because what happens there follows on from the Erlang lost-call formula already discussed. Figure 30.11 illustrates a network in which an operator switchroom is serving customers of the telephone network.

In all,  $N$  operators (called *servers* in teletraffic jargon) are working and  $K$  places (confusingly also called *servers*) are available in the queue.

The switchroom manager wants to make sure that the number of operator positions staffed ( $N$ ) at any point of time during the day is great enough to cope with the traffic and so keep down the queue waiting time. Furthermore, the switchroom designer wants to ensure that the queue capacity ( $K$ ) is great enough to minimize the number of lost calls.



**Figure 30.11** Call queueing in an operator switchroom

Both  $N$  and  $K$  can be determined from an adapted form of the Erlang *lost-call* method, using various forms of the *Erlang Waiting Call* formula, as shown in Table 30.4.

The formulae of Table 30.4 are unfortunately complex and this is not the place to describe their derivation or any of the more complex versions of them that have also been developed. We can, however, discuss their practical use, as follows. A tolerable delay  $t$  is set for the time which telephone callers in Figure 30.11 will have to wait for an operator (e.g. 15 seconds), and we then decide what percentage of calls will be answered within this delay time (a typical target would be 95%). It is then a trial-and-error exercise, using formulae (i) and (iv) of Table 30.4, to give the number of servers ( $N$ ) which will meet the target constraints.

The value  $N$  will always be greater than the offered traffic  $A$ . If it were not so, the operators would be reducing the queue at a rate slower than that at which the queue was forming. The queue and the delay could only become longer. (Substitute  $N = A$  in formula (ii) and you will see that the average delay is infinite).

Let us use an example to illustrate the method. We shall assume that the switchroom handles an average of 3000 calls per hour, and that they take an average of  $d = 60$  seconds of the operator's time (i.e. *service time*) to deal with.

... Then the *offered traffic*,  $A$  (average number of calls in progress)

$$= 3000 \times 60/3600 = 50$$

Guessing a value of  $N = 55$  to meet our target that 95% calls will be answered within 15 seconds, we calculate in order ...

$A = 50$ Erlangs	( <i>offered traffic</i> )
$N = 55$	( <i>active servers (operators)</i> )
$B = 0.054$	from the <i>lost-call formula</i>
$C = 0.388$	from formula (i)
$t = 15$ seconds	( <i>target answer time</i> )
$d = 60$ seconds	( <i>average service time</i> )

Probability that delay exceeds 15 seconds from formula (iv) = 0.11

**Table 30.4** The Erlang waiting call formula

- 
- (i) Probability of delay (i.e. that a customer will have to wait to be served)

$$= C = \frac{NB}{N - A(1 - B)}$$

This is called the Erlang call-waiting formula.

- (ii) Average delay (held in queue)

$$= D = \frac{C}{N - A} \times d$$

- (iii) Average number of waiting calls

$$= \frac{AC}{N - A}$$

- (iv) Probability of delay exceeding  $t$  seconds

$$= C e^{-(N-A)t/d}$$

- (v) Probability of  $j$  or more waiting calls

$$= C \left( \frac{A}{N} \right)^j$$

- (vi) Probability of  $x$  servers being busy

- (a) No queue,  $X$  = number of operators busy

$$p(x) = \frac{A^x}{X!} p(0) \quad 0 \leq x \leq N$$

- (b) Queue, all operators busy,  $j$  queue places taken

$$x = N + j$$

$$p(N + j) = c \left( 1 - \frac{A}{N} \right) \left( \frac{A}{N} \right)^j$$

#### Notation

$N$  = number of servers (those processing the calls, e.g. operators)

$j$  = number of calls in queue

$B$  = lost call probability if there were no queues – derived from the Erlang lost call formula  
 $= E(N, A)$

$A$  = offered traffic in Erlangs

$d$  = average service time required by an active server to process a call

---

As this proportion is higher than the target of 0.05, we will need more servers, but probably not too many more because we are not too far adrift.

Repeating for  $N = 56$ , we find the probability of answer in 15 seconds is 0.07. Almost, but not quite! Repeat again for  $N = 57$ , and the probability of answer taking longer than 15 seconds is only 0.04. Thus 57 operators will be needed in our switchroom.



Having determined the number of operators ( $N$ ), the number of queue places needed is found by using formula (v) of Table 30.4. The planner first decides what small proportion of calls it is acceptable to lose because there are no free places in the queue. Then he finds the value  $j$  from the formula which satisfies the condition that the probability of lost calls (more waiting calls than queue positions) is met.

Going back to our example, let us aim to lose only 1% of calls because of inadequate queue capacity. Then the probability of  $j$  or more waiting calls must be less than 0.01, where  $j$  is the designed capacity of the queue. The calculation is as follows...

$A = 50$ Erlangs	offered traffic
$N = 57$	active operators
$B = 0.039$	Erlang lost call formula
$C = 0.246$	Erlang waiting call formula

So, from formula (v)

$$0.01 = C(A/N)^j$$

rearranging

$$j = (\log 0.01 - \log C) / (\log A - \log N)$$

$$j = 24.4$$

rounding up, 25 queue places will be needed.

Out of curiosity, why not calculate the average waiting time and the average number of callers in the queue? The average waiting time is 2.1 seconds, and the average number of waiting calls is 1.76.

Calculating the number of operators needed to answer calls on a company's office switchboard is carried out in an identical way to that used above.

Theoretically, different formulae must be used on occasions where the offered traffic distribution does not follow Erlang's detailed mathematical assumptions (given in the appendix). Inappropriate use of the formulae may lead to errors in the available network capacity, but on the other hand most practitioners prefer to use simple dimensioning methods and rely on their practical experience rather than worry too much about complex statistical theory. If interested, a number of more complex methods and their detailed derivations are given in numerous specialist texts on queueing theory.

## 30.10 DIMENSIONING DATA NETWORKS

Terminals and other devices sending information on data networks generally do so in a packet-, frame- or cell-oriented format over a packet-, frame- or cell-switched network.

In the case of Figure 30.12(b), the data network system has been designed to flatten the profile of the traffic by using buffers and statistical multiplexing. Such a network demands the use of a *call-waiting* dimensioning technique to decide the required data link bit rate and buffer capacity.

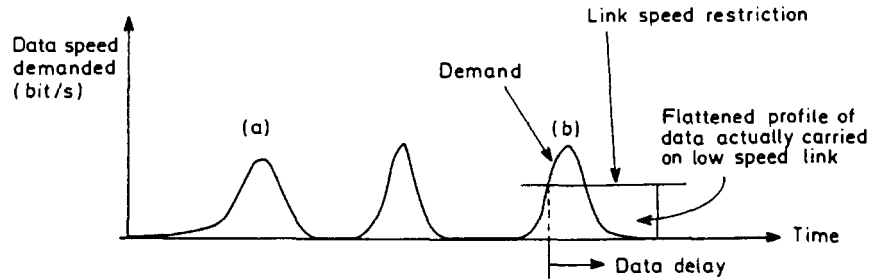


Figure 30.12 Peaked bursts of data demand

The time for which an individual packet, frame or cell must wait in the buffer for the line to become free may be a significant proportion of the total time needed for it to traverse the network as a whole. Thus the *waiting time* in the buffer contributes significantly to the *propagation time*. The propagation time in turn may affect the apparent speed of response of a computer reacting to the typed or other commands of the computer user. Most computer applications can withstand some propagation delay, though above a given threshold value further delay may be unacceptable. Thus, for example, many packet-switched data networks are designed to keep delays lower than 200 ms.

The switch and network designers of a data network must ensure that the bitspeed of the line is greater than the average offered bitrate of all the packets, frames or cells. If the line were not fast enough, a build-up of data would occur in the buffer at the transmitting end, causing considerable delay. In addition, the buffer must be large enough to temporarily store all waiting data. If the buffer were not big enough, then arriving packets will be lost at times when it is full and 'overflows'.

The formulae of Table 30.4 may be adapted for data network dimensioning by substituting  $N = 1$ . This corresponds to a single connection between points. It is equivalent to a single data link between terminals or packet-switched data exchanges. The link will have a given bit speed capacity, and the data flowing over the link could be thought of as being measured in Erlangs. For example, an average bit rate of 7500 bit/s being carried on a 14 400 bit/s link is equivalent to 7500/14 400 or 0.52 Erlangs.

Table 30.5 The Erlang call-waiting formula applied to data networks

- |       |   |
|-------|---|
| (i)   | probability of delay = $C = A$ = average line loading (typically monitored in %)                        |
| (ii)  | average delay = $D = D \frac{p}{(1/A - 1)L}$  |
| (iii) | average number of waiting packets or frames = $\frac{A^2}{1 - A}$                                       |
| (iv)  | probability of packet delay (due to output buffering) exceeding $t$ seconds<br>= $A \exp(-(1 - A)tL/p)$ |
| (v)   | probability of $j$ or more waiting packets = $A^{j+1}$  |

where  $p$  = data packet, frame or cell size in bits

$L$  = datalink line speed in bit/s

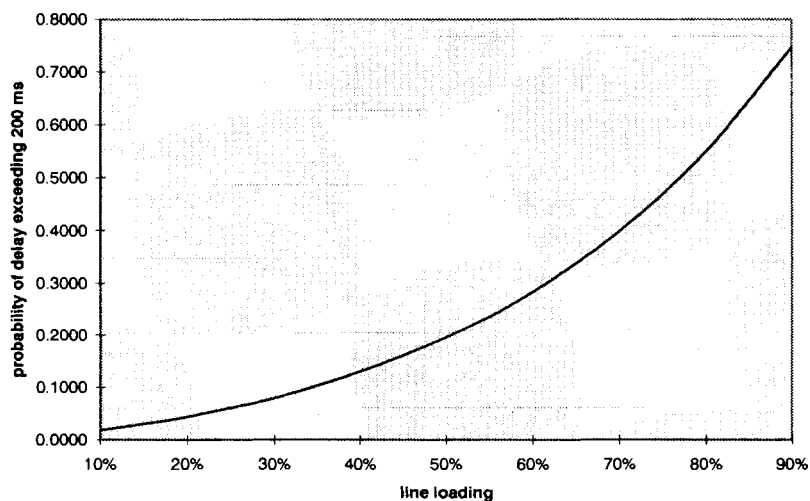
Setting  $N = 1$  into the Erlang lost-call formula to calculate  $B$ , and then substituting this value into the formulae of Table 30.4, and by renaming some of the parameters to some more appropriate to data communication, we obtain the formulae of Table 30.5.

It is interesting to substitute a few values into the formulae of Table 30.5 to give practical insight into the operation of data networks. First, let us consider the maximum acceptable line loading for a typical X.25-based packet switching network based upon 9600 bit/s trunk lines. To do so, we consider formula (iv), substituting  $L = 9600$ , a packet size  $p$  of 128 bytes = 1024 bits and an 'acceptable' buffering delay of 200 ms, together with a 'maximum acceptable' buffering delay of 500 ms. Substituting different values of  $A$ , the line loading, we obtain the probabilities of the packet delays exceeding the 200 ms and 500 ms thresholds, as shown in the graphs of Figures 30.13 and 30.14.

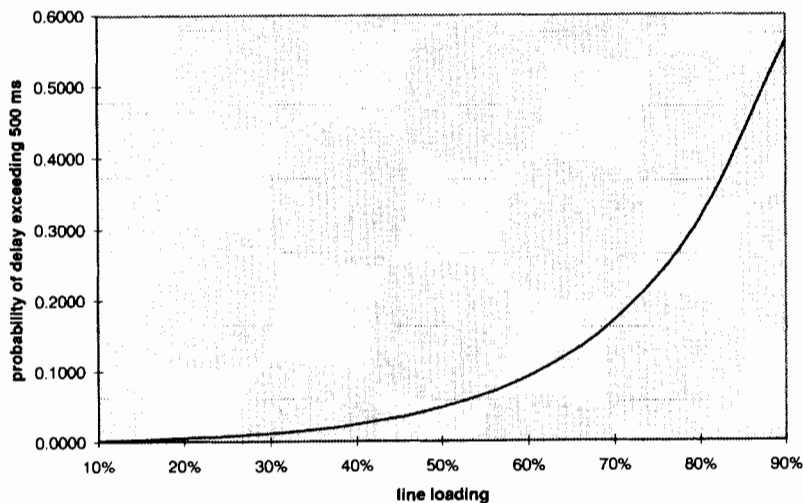
Note how the probability of the buffering delay exceeding the 500 ms 'maximum acceptable' increases rapidly above line loadings of about 40%. At 40% line loading, this probability is only around 2% (Figure 30.14). The conclusion is that we should not expect to run a 9600 bit/s packet-switched data network much above about 40–50% line loading, if we do not want to experience 'unacceptable' delays.

What about the effect of increasing the linespeed while keeping the packet size constant you might ask? What if we increase the linespeed fourfold to 28 800 bit/s? This is a speed widely available nowadays using analogue modems. The result is shown in Figure 30.15, where we again plot the probability of exceeding the 500 ms 'maximum acceptable' buffering delay limit. Now we are able to operate the line at around 70% loading without exceeding the 500 ms delay limit more than 1% of the time.

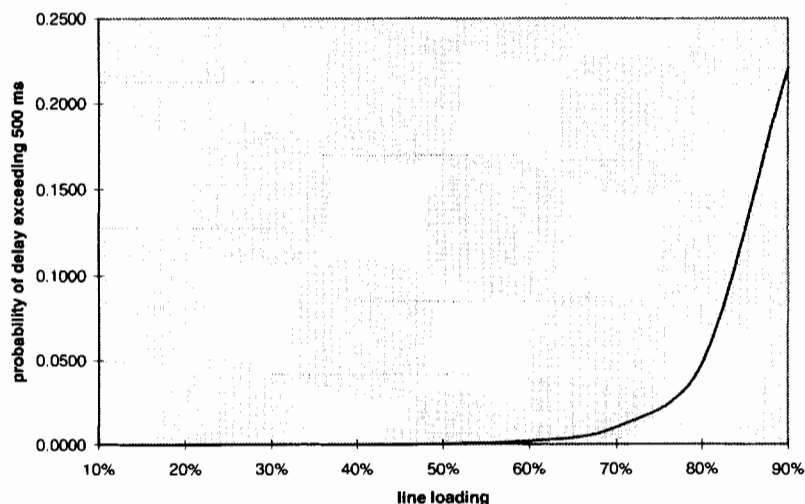
Finally, let us also consider the practical problem of how large the waiting packet buffer must be for our two examples. Here we decide on a 'maximum acceptable' proportion of packets which may be lost or corrupted (due to buffer overflow) of say 0.01%. We use formula (v) to calculate the required buffer size and derive the graph shown in Figure 30.16.



**Figure 30.13** Probability of buffering delay exceeding 200 ms (for linespeed = 9600 bit/s, packet size 128 bytes). Total packet propagation delay around 300 ms

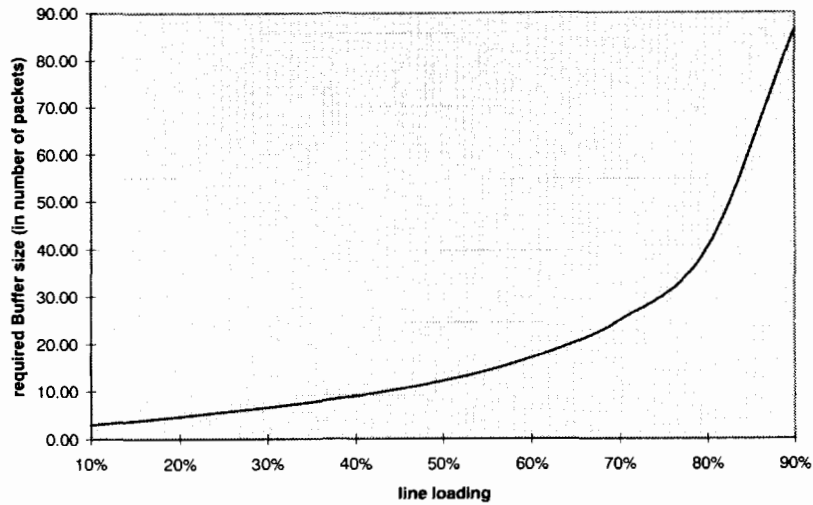


**Figure 30.14** Probability of buffering delay exceeding 500 ms (for linespeed = 9600 bit/s, packet size 128 bytes). Total packet propagation delay around 600 ms. Note: the total propagation delay of a packet through the network (first bit sent to last bit received) is the sum of the delay caused by buffering (formula 5(iv)) plus the time required to transmit the packet to line ( $p/L$ ) plus the electrical signal propagation delay (distance/speed of propagation). In the examples above the packet transmission time adds 107 ms and the electrical propagation around 3 ms per 1000 km.



**Figure 30.15** Probability of delay exceeding 500 ms (for linespeed = 28 800 bit/s, packet size 128 bytes)

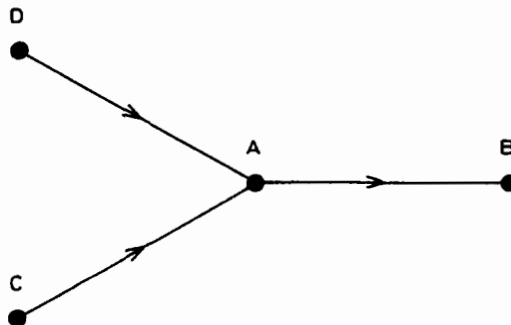
In more complex data networks, as exemplified by Figure 30.17, the same delay dimensioning method may be used for the individual links of the network. Thus link D–A of Figure 30.17 can be dimensioned according to the needs resulting from the sum of D–A and D–B traffic. Link A–B is dimensioned according to the aggregate needs of D–B, C–B and A–B traffic.



**Figure 30.16** Required packet buffer size for loss or corruption of not more than 0.01% of packets

### 30.11 POLLACZEK-KHINCHINE DELAY FORMULA

The formulae presented so far in this chapter all derive from Erlang's work of 1917 and are therefore subject to the assumption that the offered traffic follows a particular *Poisson distribution* and has a *negative exponential distribution* of holding times. Although in practice these assumptions are found to give reasonable modelling of the performance of many networks, they are not always applicable, especially in some types of data network. We therefore conclude the chapter by presenting without further discussion the more general *Pollaczek-Khinchine* formula for the average data delay. The formula given in Figure 30.18 is always applicable.



**Figure 30.17** Dimensioning meshed data networks

$$\text{Average delay, } D = \frac{A(1 + V^2) \bar{h}}{2(1 - A)}$$

$A$  = probability of delay (=traffic in erlangs)  
 $V = s/\bar{h}$  = coefficient of variation of the holding time  
 $s$  = standard deviation of the holding time.  
 $\bar{h}$  = average holding time

Figure 30.18 Pollaczek–Khinchine delay formula

### 30.12 PRACTICAL DIMENSIONING OF NETWORKS

Many network operators do not apply sophisticated traffic modelling techniques in support of their network dimensioning. Instead the more pragmatic and empirical approach prevails . . . ‘if too many calls are being lost add some more circuits’ or ‘if the computer network response time is too slow, upgrade the speed of the data transmission lines’. Although such methodology may appear crude, there is much to recommend it, as only very large networks are likely to carry sufficient traffic to be accurately characterized by statistical methods. In addition, many computer applications running on data networks may be designed to run at certain fixed times (e.g. hourly update of all the sales made in the outlets of a particular shopping chain). Here again, statistics may be of lesser value than experience.

The empirical method should, however, be used with caution. The data network is not the only source of slow responding computer applications, and merely ‘throwing more capacity at the data network’ may not provide a solution at all. Also to be reviewed is the design of the computer program itself, and the protocols being used (see Chapter 20 on network response time effects).

Adding more capacity at one point in a network sometimes has the effect of stimulating more traffic or of unstabilizing the routing of calls through the network. This in turn can lead to greater rather than lesser problems. Some self-routing networks (for example *Internet* and other *router*-based networks) may decide the best route to a destination based upon the speed or capacity of the intermediate links. Thus adding capacity to a single link in isolation may merely have the effect of attracting more traffic to it.

Ultimately, there is no better tool for dimensioning a network than a thorough knowledge of the traffic routes within it and the uses to which it is being put.

### 30.13 APPENDIX: THE DERIVATION OF ERLANG’S FORMULA

The dimensioning of telephone and other circuit-switched networks is based on a technique developed by a Danish mathematician A. K. Erlang in 1917. The *Erlang formula* or *Erlang lost-call formula* and its derivatives enable network designers to determine how many circuits are required on a given route to control the proportion of blocked calls (a typical target is for losses of one call in one-hundred or less).

In arriving at his statistical formula, Erlang made a number of assumptions, based on his observations of telephone callers' behaviour, as follows.

- (i) Calls are generated at random and only one-at-a-time (this kind of behaviour has been described as the *Poisson process*, after the mathematician who invented it).
- (ii) Statistical equilibrium exists. By this we mean that the probability of a given number of circuits being in use does not change over the relevant period.
- (iii) Calls *arriving* in the system when all the circuits are busy are lost, their holding time is zero, and they do not generate *repeat attempts*. This is the *lost-calls cleared* assumption that lends its name to the *lost-call formula*. Other more advanced formulae model the behaviour when calls which cannot be immediately completed are either queued up, or are considered to stimulate customer repeat attempts.
- (iv) Call holding times follows a negative exponential distribution.

Let us imagine that at a particular time,  $x$  circuits are in use. Now let us consider a time only a tiny fraction of a second later (which we denote as  $dt$ ). Given assumption (i) above, only one of three things can have happened

- (a) one more circuit has become busy (raising the number to  $x + 1$ )
- (b) one circuit has become idle (reducing the number busy to  $x - 1$ )
- (c) the number of busy circuits has remained constant

Now the probability that event (a) has happened is equal to the probability of a single call arriving during the period  $dt$ . From the earlier part of the chapter we know this to be  $A dt/h$ . Conversely, the probability of any particular one of the established calls terminating during the period  $dt$  is independent of the new traffic being offered, and is equal to  $dt/h$ . Multiplying by the number in progress we obtain the probability that any one of the calls clears as  $x dt/h$ .

Next we apply assumption (ii). As the system is in statistical equilibrium, then the probability that there were  $x$  busy circuits to start with and that one more circuit became busy must equal the probability that there were  $x + 1$  busy circuits to start with, and that one circuit was released. If this were not the case we would find that the circuits either became generally busier over a period of time, or generally less busy, depending on the state of the statistical unequilibrium. Assuming  $p(x)$  to be the probability of  $x$  circuits being busy, then the probability of  $(x + 1)$  busy circuits will be  $p(x + 1)$ .

From our results above, the probability that there were  $x$  busy circuits to start with, and that there has been a transition from  $x$  to  $x + 1$  busy circuits is

$$p(x)A dt/h$$

and the probability that there were  $(x + 1)$  busy circuits to start with followed by a transition from  $x + 1$  to  $x$  busy circuits is

$$p(x + 1)(x + 1) dt/h$$

As we said, we may equate the two values. Hence

$$p(x) \frac{A}{h} dt = p(x+1) \frac{(x+1) dt}{h}$$

In other words

$$p(x+1) = \frac{A}{x+1} p(x)$$

By substituting different values for  $x$ , we determine that

$$p(1) = \frac{A}{1} p(0)$$

$$p(2) = \frac{A}{2} \frac{A}{1} p(0) = \frac{A^2}{2!} p(0)$$

$$p(3) = \frac{A}{3} \frac{A}{2} \frac{A}{1} p(0) = \frac{A^3}{3!} p(0)$$

$$p(x) = \frac{A^x}{x!} p(0)$$

The value  $(x)$  is the probability that  $x$  circuits are in use. From it we derive the *probability density curve* shown in Figure 30.19, which gives the relative probabilities of different circuit numbers being in use. The curve is also called the *Erlang distribution*. As can be seen from the curve, and as we would expect, the most likely number of circuits in use is nearly equal to the offered traffic intensity  $A$ .

Finally, we use assumption (iii): that calls arriving when all circuits are busy are lost. They may lead to delayed customer re-attempts but we assume that these do not distort the statistical distribution of call arrivals. By doing so, we can thus determine the value

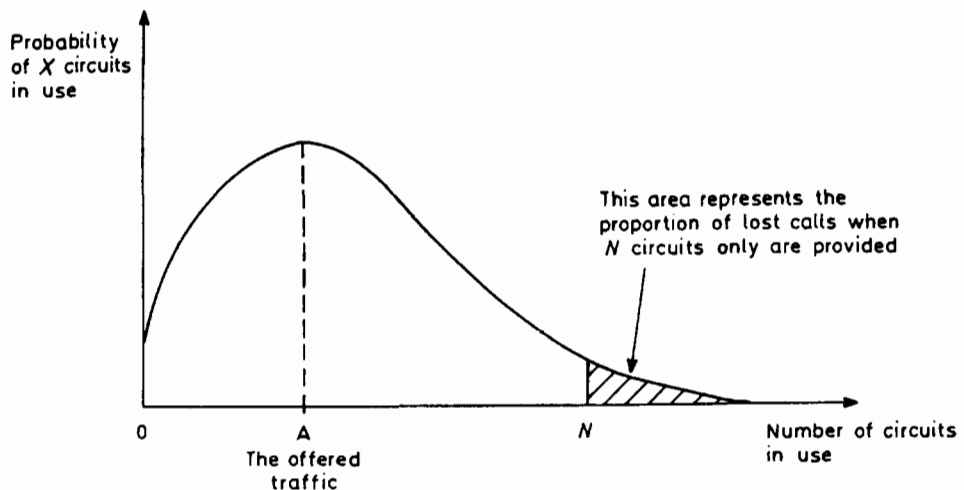


Figure 30.19 The Erlang distribution



$$E(N, A) \text{ or } B(N, A) = \frac{A^N}{N!} \left/ \left\{ 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^N}{N!} \right\} \right.$$

where  $E(N, A)$  = proportion of lost calls, and probability of blocking  
 $A$  = offered traffic intensity  
 $N$  = available number of circuits

Figure 30.20 The Erlang lost-call formula

of  $p(0)$ . We add all the probabilities and adjust them so that their cumulative probabilities are equal to 1. Thus

$$p(0) + p(1) + p(2) + p(3) + \cdots + p(N) = 1$$

From this, and our earlier result we conclude that

$$p(0) = \frac{1}{(1 + A + (A^2/2!) + \cdots + (A^N/N!))}$$

$$p(x) = \frac{A^x/X!}{(1 + A + (A^2/2!) + \cdots + (A^N/N!))}$$

This is the formula of the Erlang distribution, already illustrated in Figure 30.7. Its most useful application is in derivation of a further formula for calculating the proportion of lost calls offered to an  $N$  circuit group. This is the *Erlang lost-call formula*, usually denoted  $E(N, A)$ , or sometimes  $B(N, A)$ . It calculates the probability of call blocking, given  $N$  available circuits and an offered traffic intensity of  $A$  Erlangs. Numerically, the value is equal to the probability of finding all  $N$  circuits busy, and it is derived from the Erlang distribution by substituting  $x = N$ . Figure 30.20 presents the formula.

---

## *Traffic Monitoring and Forecasting*

---

Having explained the underlying principles of electrical communication and the statistical ‘laws’ of telecommunications traffic, we can now consider the practical design and operation of networks. A prime concern is to ensure that there are adequate resources to meet the traffic demand, or to prioritize the use of resources when shortfalls are unavoidable. Two things have to be done to keep abreast of demand. These activities are the monitoring and future forecasting of network use. In this chapter we shall review the parameters to be applied in measuring traffic activity and go on to provide an overview of some forecasting models for the prediction of future demand. Used as an input to the teletraffic engineering formulae of Chapter 30, the forecast values of future traffic can be used to predict the future network equipment requirements on which the overall planning process will be based.

### 31.1 MEASURING NETWORK USAGE

We discussed different methods of defining the volume of network usage in Chapter 30. These methods in various ways served to measure, over a predetermined period of time, any of the following parameters

- the total number of calls or messages
- the total conversation time
- the total *holding time* (holding time includes both conversation time and call set-up time)
- the total number of data characters (e.g. *packets*, *frames* or *cells* conveyed)

We have covered the concept of *traffic intensity*, which is a measure of the average call demand. We have shown how traffic intensity measured in *Erlangs* determines the number of circuits needed on a route in a circuit-switched network to maintain a given *grade of service*. For this reason it is usual practice to monitor the magnitude of the

traffic intensity, and to predict its future values by forward forecasting. This provides the basis for planning a future circuit provision programme.

Forecasting, however, need not be confined to predicting future traffic intensity. The measuring and forecasting of other parameters can also be valuable predictors of network traffic demand. For example, in telephone networks where customers pay for calls based on the total time used in conversation, the measurement of conversation time not only allows the preparation of customer bills, but also enables the network operator to calculate the revenue due and the resulting profits, immediately after the end of each financial period. Meanwhile, packet data network customers usually pay for the total volume of data carried (measured in *segments*). The actual amount of use and predicted future revenue may well be vital to the network operator on the business and financial side.

Forecasting also allows investments to be planned to meet future demand, commensurate with predicted levels of profits and financial resources. Forecasting also helps to determine when established will become unprofitable, and therefore when it may be appropriate to consider run-down or discontinuation of the service. No commercially-minded company can put up with loss-making products or services for long.

## 31.2 USAGE MONITORING IN CIRCUIT-SWITCHED NETWORKS

The parameters commonly used to monitor the usage of circuit-switched or point-to-point (leased) networks are listed below

- the traffic intensity
- the total number of minutes of usage
- the total number of calls completed
- the total number of calls attempted

The reason for monitoring each of these parameters, and the methods of measurement, are described in the following sections.

## 31.3 TRAFFIC INTENSITY

*Traffic intensity*, as we learned in Chapter 30, is the measure of the average number of simultaneous calls in progress on a route between two exchanges, or across an exchange, during a given period of time. It is normally measured during the hour of greatest traffic activity (or so-called *busy hour*) and is quoted *Erlangs*. There are two principal methods of measuring traffic intensity: either by frequent sampling of the number of circuits actually in use and calculation of the statistical average, or by using a call logging system which records the start and finish time of each individual call, so allowing the total circuit usage time to be measured accurately. Total usage during the period is divided by the duration of the period to give the average number of circuits in use.

A check at least once per month of the busy-hour traffic intensity on each route in a network is essential to make sure there are enough circuits on each individual route. Circuit requirements for each route are calculated from these values according to the Erlang formula, using the forecast traffic and the desired grade of service as inputs to the formula.

On their sampling days, many network operators are not content to monitor the traffic intensity during the anticipated busy hour (though this is an accepted practice); they also monitor the traffic activity profile on each route throughout the whole day. The profile is invaluable as an indicator of overall calling patterns, revealing short or long periods during the day when the traffic on a particular route is either very heavily congested or is under-utilizing the available capacity. In the latter case, the capacity may be useful as a means of relieving a congested route (by transit routing in the manner discussed in the next chapter). Alternatively, an advertising campaign could be used to stimulate extra network usage.

Traffic profiles on different routes vary greatly according to the nature of the route (e.g. local, trunk or international) and of its users, and they may change slowly over a period of months. Some profiles slowly distort towards a highly peaked, busy-hour oriented pattern, and others reflect a steadier traffic load throughout the day.

Highly peaked and busy-hour oriented profiles, of which Figure 31.1 is an example, can be problematic for network operators for two reasons. First congestion is highly likely during the period of peak traffic demand (even if circuits are relatively well provided), and second over the day as a whole there will be relatively little network use, with correspondingly low revenue.

Highly peaked traffic profiles can sometimes be flattened by charging a heavy price premium for calls made at the peak time. The idea is to persuade customers to make their calls either a little earlier or later during the day. If it can be achieved, a very flat profile such as that shown in Figure 31.2, makes highly efficient use of the network. Charges which depend on the time of day tend to have an effect on domestic call demand, but the effect is less marked on business calls, because few of the callers worry about the cost.

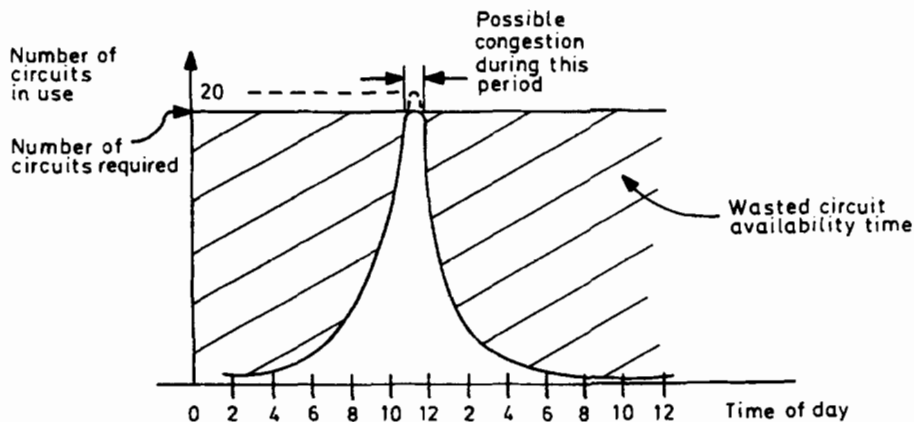


Figure 31.1 Busy hour oriented traffic profile

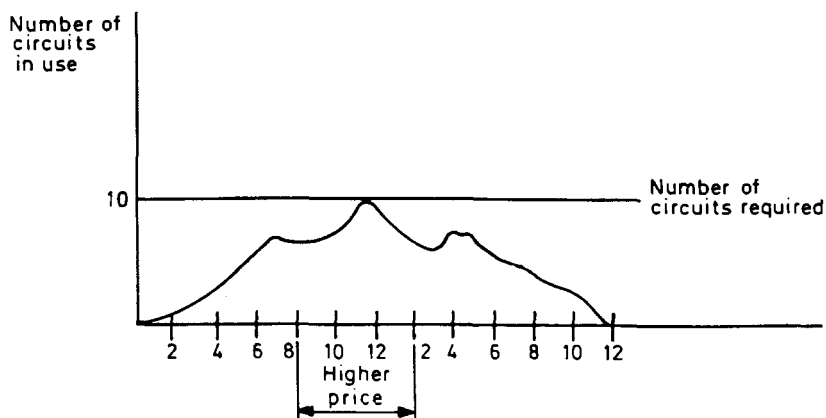


Figure 31.2 Traffic redistribution

Figure 31.2 illustrates an extreme redistribution of the traffic previously shown in Figure 31.1. In the course of the day the same number of calls have been completed in Figure 31.2 as in Figure 31.1 and thus a similar revenue has been earned, but the flatter traffic profile has meant a lower busy-hour traffic value and a reduced circuit requirement. The profile has been flattened by stimulating 'time-shift' of some of the peak traffic, by charging a higher price for the peak period 8 a.m.–2 p.m.

There are times when significantly higher busy-hour prices fail to dissuade customers from calling at the busiest time, and none of the traffic is timeshifted. When this happens the only way to improve overall network utilization and revenue is by stimulating new off-peak traffic.

### 31.4 TOTAL USAGE MONITORING

Besides coping with its route busy-hour demand, a network must be designed to carry the total volume of demand. Exchanges may be limited in their capability to record the details of individual calls, so that above a certain limit, billing or other call detail information may be lost. In addition, the amount of exchange equipment (switching points, registers, etc.) which is required will depend on the total volume of demand. There are two related measures of the total usage of a circuit-switched network or a leased point-to-point connection. The two measures are the *paid time* and the *holding time*.

The *paid time* equals the total number of usage minutes for which customers have been charged. In telephone network terms this is the number of *conversation minutes*. In a telex, or a circuit-switched data network, the paid time is the *connected time*, the duration of which has a direct relationship with the total amount of textual information that could be carried.

By contrast, the total *holding time* is the time for which the network itself has been in use, and it is always slightly greater than the *paid time*, as Figure 31.3 shows. It includes not only the 'connected' or 'conversation' time, but also the time needed for call set-up

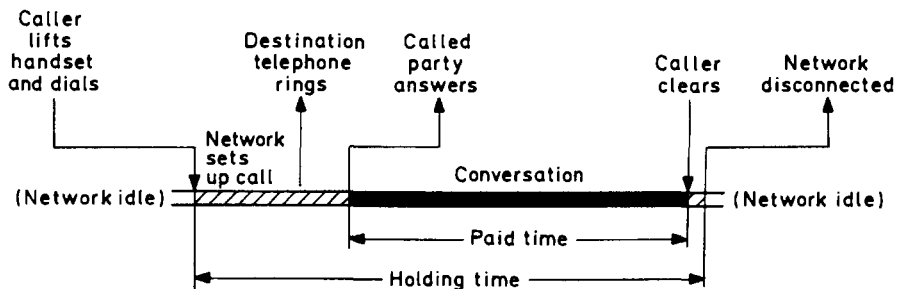


Figure 31.3 'Holding' and 'paid' time

and clear-down and the ringing time of the destination telephone. The holding time is the better indicator of network resource requirements, but network operators may choose to monitor either *paid* or *holding time*, or both.

### 31.4.1 Paid Time

The paid time is a direct measure of customers' actual usage. This is the time during which the customer actually communicates, and could be considered to be the best measure of 'real' demand. The more paid-time use of a network is made, the greater is the volume of communication. This contrasts with the traffic intensity, which is only a measure of how many people choose to communicate at the same time. It is paid time that attracts revenue and is therefore the most important financial parameter. The paid time accumulated by particular customers or on particular routes between individual exchanges gives a strong indication to network operators of where they should concentrate their resources and efforts, and where they are most at risk from competitors. Operators can ill-afford to lose valuable customers or the traffic on profitable routes.

There are three methods of either measuring or estimating paid time. The first and most obvious method is to sum the usage made by individual customers. However, although summation gives valuable route-by-route paid minutes statistics when derived from itemized customer bills, in those networks which only record their customers' usage as a bulk total number of units on simple cyclic meters (as Chapter 35 discusses), summation only reveals total network use, without any route-specific paid minute information. However, all is not lost, because the paid time on each route may be measured directly by sampling the traffic intensity on the route itself at a number of regular time intervals, and applying one of the following conversion formulae for estimating the paid minutes.

$$\text{Busy hour paid minutes} = \text{Busy hour traffic in Erlangs} \times 60 \times \text{efficiency factor}$$

$$\text{where efficiency factor} = \frac{\text{paid time}}{\text{holding time}}$$

(the value is estimated from historical information)

daily paid minutes = 5 or 6 times busy hour paid minutes

(again the value is obtained from historical records)

total monthly paid minutes = approx.  $22 \times$  daily paid minutes

(It is a good enough assumption to estimate an entire month's traffic as equivalent to that of 22 working days).

### 31.4.2 Holding Time

In contrast to the paid time, the longer period of *holding time* represents the total time when network resources are in use, and the difference between holding and paid time volumes is the time when the network *common equipment* (used for call set-up and cleardown) is in use. This quantity is called the common equipment holding time. The total holding time, the common equipment holding time, and the traffic intensity values all help to determine how many of each of the various equipments must be provided in the network.

The amount of a particular type of common equipment (e.g. signalling receivers) needed in an exchange is usually calculated by the normal Erlang method, with the common equipment traffic load calculated by multiplying the expected number of calls by the average common equipment holding time per call. Typical grades of service demanded of common equipment might be 0.5% lost calls, 0.1% or even 0.05%.

## 31.5 NUMBER OF CALLS ATTEMPTED

The total number of calls attempted (also called the number of bids) is the best measure of unconstrained customer demand, because unlike the paid minute volume or the traffic intensity actually carried by a network, it is not limited by network congestion. At a time of network congestion, the *busy hour call attempt (BHCA)* count may continue to increase (as unsatisfied customer demand continues to grow) whereas paid minute volumes and traffic intensities may saturate at the maximum capacity of the network. A very large number of unsatisfied call attempts is an almost certain sign of congestion, either through underprovision of equipment or resulting from short term network failure, and is a good means of spotting suppressed traffic within a network. Unfortunately the exact amount of suppressed traffic is difficult to estimate, because the *persistence* of customers in repeating call attempts many times over affects the overall bid count. Figure 31.4 illustrates an example of the failure of some call attempts.

The number of call attempts can only be measured accurately by monitoring each individual customer's line. A value measured at any point deeper in the network will not be an accurate measure, as it will have been reduced by the effects of any congestion at the network fringe.

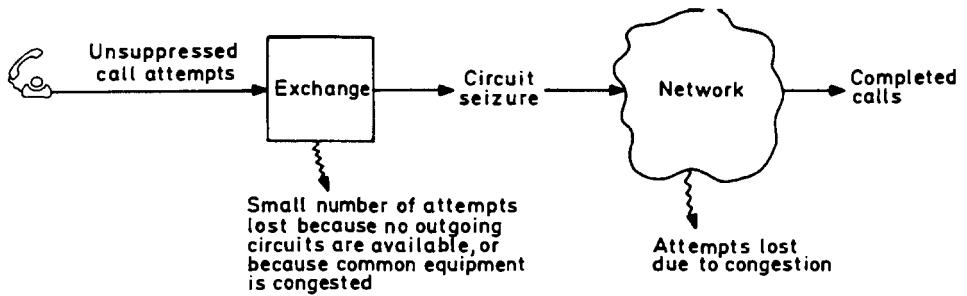


Figure 31.4 Call attempts and congestion

In cases where the call attempt count suggests that traffic is being suppressed, then the unexpurgated busy-hour traffic intensity and paid minute demand can be estimated from one of the following conversion formulae

$$\begin{aligned} \text{busy hour traffic, in Erlangs} &= \text{No. of call attempts in the busy hour} \\ &\quad \times (\text{average call holding time})/60 \end{aligned}$$

$$\begin{aligned} \text{busy hour paid time} &= \text{No. of call attempts in the busy hour} \\ &\quad \times \text{average paid time per call} \end{aligned}$$

Sometimes exchange monitoring limitations make it impossible for network operators to measure accurately the number of call attempts from a given source to a given destination. If so, the the number of call attempts may be estimated as the number of outgoing circuit *seizures*. As shown in Figure 31.4, however, this estimate will be lower than the actual because a small number of *call attempts (bids)* fail in the first exchange due to congestion, and so do not mature into an outgoing circuit *seizure*.

## 31.6 NUMBER OF CALLS COMPLETED

The number of calls completed in a network sense (i.e. reaching ringing tone or *answer*), when compared with the number of calls attempted, gives another measure of the state of network congestion. The proportion of busy hour calls completed, when expressed as a percentage of the number of calls attempted, should equal the design grade of service.

Hence

$$\text{grade of service} = \frac{(\text{number of busy hour call attempts}) - (\text{number of busy hour call completions})}{(\text{number of busy hour call attempts})} \times 100\%$$

The *grade of service* is a measure of the frustration that a customer will experience when trying to complete a call during the busiest hour of the day. We could calculate the average daily percentage of lost call attempts, but this is not so commonly done, because psychological analysis of customer behaviour suggests that it is of no relevance.



The number of calls completed by the network is a difficult quantity to measure, because not all signalling systems indicate the 'network-completed' state. In the American network a *peg-count overflow* was used to monitor this value but its absence in other networks means that it is also common to measure only the number and proportion of *answered* calls. This number is clearly lower than the number completed by the network as some calls are bound to encounter either a *subscriber busy* state or a *ring tone-no reply* condition. The proportion of answered to attempted calls, and that of answered calls to *seizures*, are known as the *answer bid ratio* (*ABR*) and the *answer seizure ratio* (*ASR*), respectively; they are defined mathematically below. Measured over relatively short periods of time (5–15 minutes), both are good indicators of instantaneous network congestion. The higher the network congestion, the lower the ABR or ASR. The converse, that the higher the ABR or ASR the greater the congestion, is not necessarily true because calls may remain unanswered for a range of other reasons (e.g. people are simply not answering their phones). These same uncertainties mean that no conclusion can be drawn from the actual value of the ABR or ASR. A conclusion can only be drawn from the value relative to its 'normal'.

$$\text{answer bid ratio (ABR)} = \frac{\text{no. of answers}}{\text{no. of call attempts}}$$

$$\text{answer seizure ratio (ASR)} = \frac{\text{no. of answers}}{\text{no. of seizures}}$$

In Chapter 37 the use of ABR and ASR statistics as tools for short term *network management* surveillance of the network will be described further.

## 31.7 MONITORING USAGE OF DATA NETWORKS

### 31.7.1 Packet-, Frame- and Cell-switched Networks, LANs and MANs

There are two prime factors of importance to data network users. These are the overall throughput capacity and the *response time* (*network transaction time* or *propagation time*) of the network.

The throughput capacity is the amount of data that can be sent over the network during a given period. Usually a network is designed to cope with the maximum demand of the peak hour, although network cost savings can be realized by queueing up less important data, for transmission outside the peak hour. Such queueing leads to more effective 24-hour use of the network.

Throughput capacity can be measured in bits per second (bit/s), messages per second, packets (frames or cells) per second, or segments per hour (1 segment = 512 bits). In designing and upgrading network capacity, it is important to consider the practical network throughput rather than just the transmission line speed capacity. The practical network throughput will always be less than the line speed capacity, first because not all messages are received without errors and some have to be re-transmitted, and second

because some of the available line capacity is effectively 'wasted' in gaps between messages. The two effects are summarized in the formulae below.

$$\text{successful message throughput (TP)} = \frac{\text{correct messages received}}{\text{time duration taken}}$$

writing it another way

$$TP = \frac{\text{total messages} \times \text{proportion with no errors}}{\text{time taken for message transmission} + \text{wasted time between messages}}$$

$$TP = \frac{M(1 - P)}{(M/R) + T}$$

where  $M$  = average message length in bits;  $P$  = probability of errors in the received message;  $R$  = line speed in bits per second;  $T$  = average line time 'wasted' between messages;  $M/R$  = time required to transmit average message.

Rearranging the formulae, so that we can relate the required network line speed to the required user data throughput

$$R = TP \frac{(1 + TR/M)}{1 - P}$$

This formula ensures an adequate average throughput capacity of the network, but another equally important characteristic of a data network is the response time. The computer systems linked by data networks will have been designed to carry out a given set of functions within a given period of time, to maintain for example a database of share prices updated at least once every hour, or (more onerously in terms of response time) to control the trains on a metropolitan underground railway.

The overall response time of a computer and data network includes the time taken to transmit an error-free message along the line and back, plus the computer processing time at the far end of the communication link, and the line turn-around time (if the link is *half-duplex*). The line transmission time includes not only the period during which the line is conveying data, but also any time spent waiting for any previous messages to be transmitted or *acknowledged*. If the response time of the network is too long, then the network designer must increase the network throughput capacity, which he can do either by upgrading line speeds (as might be appropriate on a LAN or on a point-to-point leased circuit network, using say 9600 bit/s modems rather than 4800 bit/s ones) or by providing a larger number of circuit connections (as might be appropriate in a circuit- or packet-switched data network).

The dimensioning method of calculating what overall bit throughput is required to meet a given response time constraint is based upon the Erlang waiting time formula described in Chapter 30. As might be expected, the faster the required response time, the greater is the transmission linespeed required (see also Chapter 20).

In conclusion, no matter which dimensioning method is used, if the network is to have adequate capacity to meet the user's throughput demand, the transmission line speed must be chosen with a relatively higher capacity, sufficient to counteract the effect of errors and the 'wasted' time between messages. The 'wasted' time, incidentally,

includes not only periods of instantaneous non-use, but also includes some of the data overheads which accompany each message (e.g. the header and destination codes of the network *protocol* about which we learned in Chapter 9).

Because they have a marked impact on the performance of data networks, it is worth digressing for a moment to discuss the effect of data *packet*, *block* or *frame lengths*. As we learned in Chapter 9, the protocol used to convey data messages breaks these messages down into a number of *frames*, *blocks* or *packets*, and transmits each with some other *overhead* information which is needed to ensure delivery to the correct destination and to control the frequency of errors. During periods of line noise disturbance, longer packets of data are more likely to be affected by the noise than shorter ones. The result is a lower effective data throughput and a slower response time because of the extra workload imposed by the need for large scale data re-transmission. On the other hand, shorter blocks have the disadvantage of decreasing the throughput at all times (whether the network is busy or not). This is because of the overhead of protocol *headers*, one each for the larger number of packets or frames.

### 31.8 FORECASTING MODELS FOR PREDICTING FUTURE NETWORK USE

The operation of networks, their efficient utilization, carriage of traffic, and their successful evolution all depend critically on accurate estimation of future needs. It is vital for network operators to have an efficient method of forecasting the traffic the network must carry so that a network of sufficient size can be maintained to meet users' needs.

Forecasts need to cover both short and long-term periods so that the whole business of providing for the future, the planning and the capital investment, can be put in hand in good time. For straight circuit provision a twelve month forecast may be adequate, but for normal extensions of major switching and transmission equipment it is necessary to look 1–5 years into the future, in line with the ordering lead time. Even longer forecasts may be needed when planning new cables, laying new ducts, and setting up major new sites. The longer term forecasts are inevitably less accurate than the shorter, but because there is time for adjustment, this is acceptable.

Forecasting uses historic measurements of network usage, and particularly of growth in usage, to predict likely future customer demand. A number of different forecasting models have been developed, but none of the methods are 100% reliable. This is hardly surprising as they all involve attempts at predicting future human behaviour. The reliability of any forecast relies on the accuracy of the historical information and the period over which it has been collected. The shorter the period of historical knowledge and the lower the number of observations on which it is based, the less reliable is the forecast. In general, the selection of an appropriate forecasting method ensures that the forecast is accurate for a period into the future approximately equal to one-third of the period of the historical information. Thus a one year forward forecast should be based on at least three years of historical information. The best way to select a particular forecasting method for use in any given circumstance is by experience. If the model appears to give reliable results in practice and therefore is a 'good predictor', then use it, if not; try another method.

In this chapter we shall not attempt to cover all the available forecasting methods, but instead review a few simple methods, particularly those described in ITU-T's recommendation E.506 on 'forecasting of traffic'.

As we have seen in the earlier part of this chapter the most important parameters measuring the overall network usage are the maximum instantaneous traffic demand (i.e. the *traffic intensity*, or its equivalent) and the maximum volume demand (i.e. the *paid minutes* or their equivalent). These parameters more than any others govern how much capacity is needed in the network and the revenues that result. As we have also seen, the two values are related by the daily traffic profile, so that paid minutes may be estimated from the busy-hour traffic intensity or vice-versa. Because of this fact, either or both of the parameters can be forecast, and a forecast for the other parameter can be calculated using the conversion factors presented earlier. Forecasting the values directly is called *direct forecasting*, whereas deriving them from conversion formulae is called *composite forecasting*. By using both methods, the forecast can be double-checked.

In preparation for a traffic forecast, the forecaster should identify any regular or irregular disturbances which have affected past traffic and may affect the future forecast. These may include discontinuities in traffic growth or decline, resulting from disturbances such as changes in tariffs, modernization of the network, or removal of congestion. The use of composite forecasting as a double check comes into its own when a large number of discontinuities or irregular jumps in volume are present. Figure 31.5 shows an example of a leap in traffic growth caused by the stimulation of traffic resulting from the introduction of automatic service. Such leaps in telephone traffic growth, even growth of many times over, are not uncommon.

The process of developing a forecast comprises five steps. First, one of a number of different types of forecasting models must be chosen. Examples include simple models which might assume linear, quadratic or exponential growth in traffic as shown in Figure 31.6. Alternatively, more complex forecasting models such as econometric models may be used. Econometric models attempt to predict future demand by assuming that the traffic depends upon a number of socio-economic variables such as

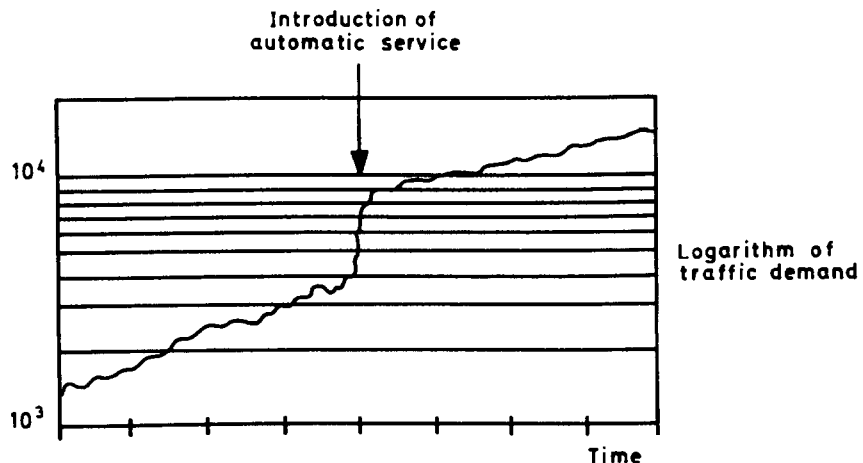
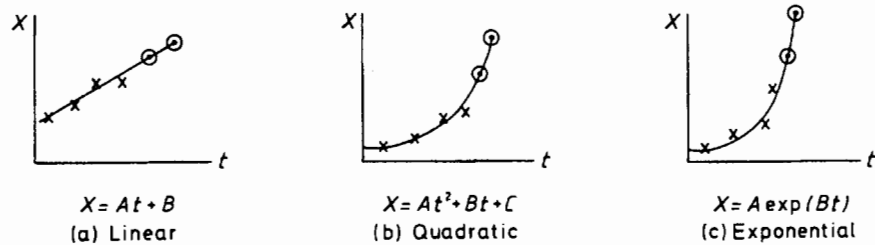


Figure 31.5 Discontinuity in traffic growth

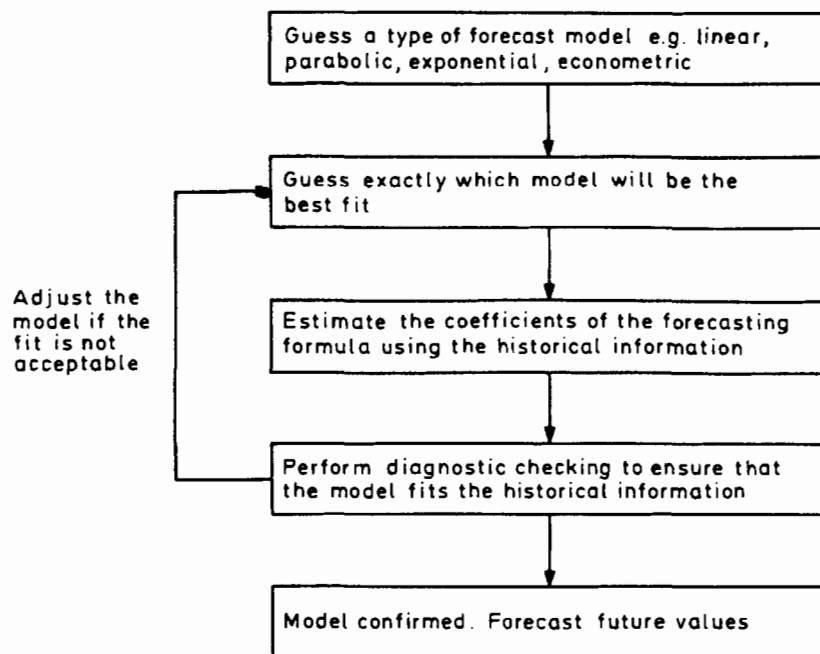


**Figure 31.6** Simple forecasting models. X, historic values; ○, forecast future values

the retail price index, the population count, etc. The equations of an econometric model reflect a relationship between these variables and the traffic parameters.

The second step is to guess exactly which model within the class fits the historical values most accurately, and the third step is to evaluate the unknown coefficients in the appropriate formula (i.e. values  $A$ ,  $B$  and  $C$  in Figure 31.6).

The accuracy of the model should be confirmed in the fourth step by a method known as *diagnostic checking*. This ensures that there are no major discrepancies between the model and the historical information. If there are discrepancies, it attempts to remedy them by adding correction factors or by adjusting the formula coefficients. Finally, the future forecast can be made. The work process is shown in the flow diagram of Figure 31.7.



**Figure 31.7** Forecasting procedure

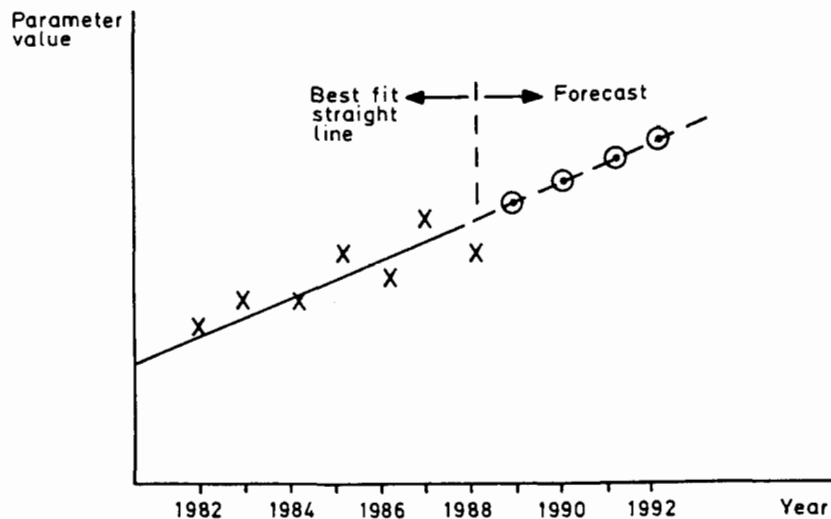
### 31.9 FITTING THE FORECASTING MODEL

Identification of the appropriate model and of the explanatory variables is the most difficult aspect of forecasting. Even when a suitable form of equation for estimating the behaviour has been determined, the evaluation of the coefficients is complex and a number of methods are available. Perhaps the simplest method of estimating the coefficient values is called *regression*. This is the only method that we discuss here, and the example that follows shows how it can be applied to a linear forecasting model. In such an instance the method is known as *linear regression*.

In linear regression we assume that the growth in traffic demand (or the change in value of some other parameter) will change over time in a linear or straight line manner. Linear regression therefore attempts to 'fit' the best straight line possible onto the historically recorded values of the parameter. When this has been done, the straight line is extended into the future to give a forecast of subsequent values, as Figure 31.8 illustrates.

In Figure 31.8 the value of a parameter has been measured every year from 1982 to 1988; based on this information, a forecast of future values corresponding to the years 1989 to 1992 is required. (The forecast is likely to be valid for around two years (one third of the period of historical information.) The forecast values are those predicted by the straight line shown in Figure 31.8. This is the *best-fit* straight line through the historical values.

The best fit straight line can be determined by minimizing the total of the vertical distances of each of the historic parameter values from the *best-fit line* as illustrated by Figure 31.9. In this example the best-fit line is taken to be that line for which the sum of the errors ( $a + b + c + d$ ) is minimized. This is the so-called *least-difference* regression line.



**Figure 31.8** Forecasting by linear regression. X, measured historic values; ⊙, forecast future values

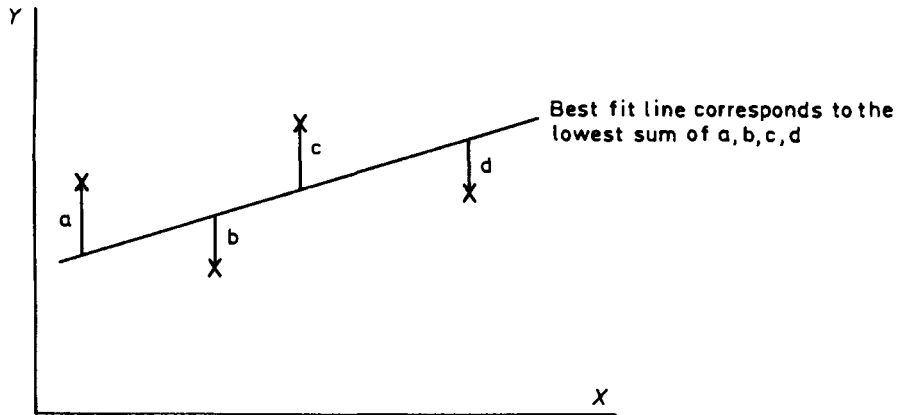


Figure 31.9 Least-difference regression

It is more normal, however, when using the regression method of line fitting, to work on *least squares regression* rather than *least difference*. In least squares regression the *best-fit* line is that corresponding to the minimum value of the squared errors, in other words the minimum value of  $(a^2 + b^2 + c^2 + d^2)$ .

The method produces an equation for the *best-fit* or *regression* line as follows

$$Y = a + bX$$

In this book we shall not discuss in detail how to derive the values of  $a$  and  $b$ , and if interested the reader should refer to an advanced statistics or mathematics book. However, for the benefit of those of a mathematical mind the formulae are as shown in the table of Figure 31.10. Future values of the parameter  $Y$  are predicted one at a time by substituting relevant values of  $x$  (the future dates),  $a$  and  $b$  into the equation shown in Figure 31.10.

Forecasting and regression methods need not be confined to straight line predictions of future events, and Figure 31.11 demonstrates the application of either the *least difference* or *least squares* regression technique to determine the coefficients of a curved forecasting model.

Best fit line equation $Y = a + bX$	
$a = M_y - bM_x$	
$b = \frac{\sum x_i y_i - M_x M_y}{S_x^2}$	
where	$M_x$ = mean of the $x$ -values $M_y$ = mean of the $y$ -values $S_x$ = standard deviation of the $x$ -values

Figure 31.10 Least square regression equation

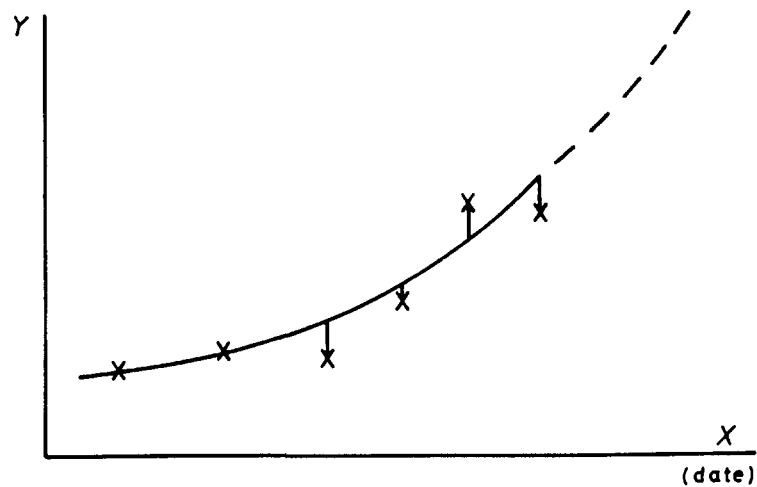


Figure 31.11 Using a curve for forecasting

### 31.10 OTHER FORECASTING MODELS

Many parameters exhibit a curved growth pattern similar to that already shown in Figure 31.11, and this is typical of a graph of total traffic demand on many different types of network. Almost any shape of curved forecasting model may be used, and some examples are given in Figure 31.12. The method remains the same: the forecaster needs to make a *hypothesis* as to the shape of curve that best predict future events, and then to conduct regression analysis to fix the coefficients of the corresponding formula which represents the forecast curve.

As we said earlier, the type of curve that is used is entirely up to the forecaster, and previous success with a given curve and forecasting method is usually the best guide. There is no 'correct' method, and only time will tell whether the prediction is correct. The only recourse is to try something different next time. The forecaster must check that any forecast he makes is sensible. Simple questions give a good guide. For example a 10 000 line local exchange is unlikely to generate 3000 Erlangs of traffic; that would require 0.3 Erlang per exchange line (18 minutes of use each day at the peak time).

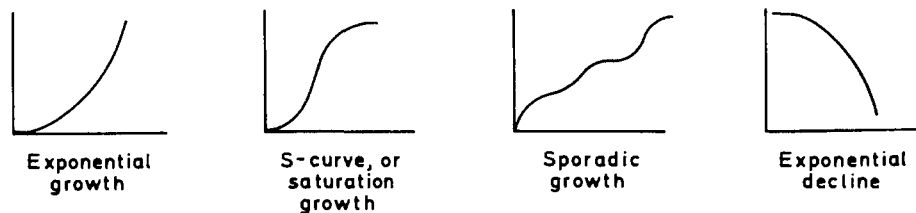


Figure 31.12 Possible prediction curves



Among the large public telecommunications companies, econometric forecasting models are becoming the most popular. By using such models, the forecasters assume that the total volume of public network traffic is related to the overall economics, social conditions and trade of the company or nation. There is no doubt about their value because the companies continue to use them. However, to end on a gloomy note, it is important to recognize that forecasts are seldom accurate, and it is no disgrace to be under or over forecast. What is inexcusable is for plans to be drawn up without sufficient contingency and flexibility to be adapted as time goes on and the real level of demand makes itself known.

---

## *Network Traffic Control*

---

Traffic control is as much art as it is science, and the profits are gratifying. The right control mechanisms and routing algorithms in the right places will determine the overall performance and efficiency of our network. What is more, they will give us simpler administration, better management and lower costs. This chapter describes a number of these admirable devices: first the simple methods commonly used for optimizing network routing under typical 'normal-day loading' conditions; then we look at recent more powerful and complex techniques, together with some of the practical complications facing telecommunications today. Finally we note how several network operators have found ways of interconnecting with the multiple carriers that have emerged lately as a result of market deregulation.

### 32.1 NETWORKS

Any metropolitan, international, trunk, transit or corporate network, by virtue of its nature and sheer size has to include a considerable number of switches (exchanges). If there are many inter-connections between these switches, any individual call crossing the network will have plenty of alternative paths open to it. It is in fact the number of path permutations which determines the robustness of the network to individual link failure and sets the level or *grade of service* provided (i.e. the probability of successful call connection). It offers the customer what he most wants, an acceptable chance of making a successful call or information transfer each time. Particular attention must be paid to choosing the call-control mechanisms which are going to determine the routing of our customers' individual calls, connections or messages, bearing in mind that in an ideal network we aim overall to achieve a controlled flow of traffic throughout the network at reasonable cost and without the penalty of complicated administration. Various methods and preliminary calculations can now be considered in turn.

## 32.2 SIZING CIRCUIT-SWITCHED NETWORKS

First of all, to find how many circuits a telecommunications network will need to meet traffic demand, a mathematical model is required to predict network performance, and it will comprise at least two parts

- a statistical distribution to represent the number of calls in progress at any given time
- a forecast of the overall volume of traffic

Erlang's model provides a statistical method for approximating telephone traffic, which is based on his measurements of practical telephone networks. A short reminder follows.

Erlang devised a method of calculating the probability of any given number of calls being in progress at any instant in time. The *probability density function* used for the calculation is termed the *Erlang distribution*. The *Erlang formula*, a related and complex iterative mathematical formula, can be manipulated into a number of different forms. The most common form is used to calculate the required number of circuits (or circuit group size) to carry a given traffic volume between two exchanges, under the constraint of having to meet a given grade of service. The traffic value input into this formula is the measurement of *traffic intensity*. The *intensity* of traffic on a route between any two exchanges in a network is equal to the average number of calls in progress, and is measured in *Erlangs*.

In Chapter 30 we discussed how, for planning purposes, it was normal to measure the route traffic intensity during the busiest hour of activity, or so-called *route busy hour*. Using this value and the *Erlang formula*, the number of circuits required for the route can be calculated according to a given target grade of service.

The *grade of service (GOS)* of a telephone route between any two exchanges in a circuit-switched network is the fractional quantity of calls which cannot be completed due to network congestion. The lower the numerical value of GOS, the better the performance. A typical target value used in many trunk and international networks is 1%, or 0.01 (in other words 1% of calls cannot be completed due to network congestion, the other 99% can). The calling customer, however, probably only perceives an end-to-end grade of service, on a call-by-call basis, of around 5% (i.e. 5% lost calls). This is because most connections comprise a number of links and exchanges, each of which is likely to be designed to inflict a 1% loss.

The common method is to dimension circuit groups within a network using a target grade of service and the Erlang formula; so that

$$\text{Circuits Required} = E(A, \text{GOS})$$

where  $E$  represents the appropriate form of the Erlang formula or function, with inputs  $A$  = route busy hour traffic and GOS = target grade of service.

Alternatively, for packet or message-switched networks, we discussed an alternative formula, which sought to model the waiting time distribution of packets, rather than the proportion of lost calls; instead of dimensioning in accordance with grade of service, we can dimension according to waiting time. The method was again covered by Chapter 30.

### 32.3 HIERARCHICAL NETWORK

The Erlang model is a good predictor of the traffic behaviour of most telecommunications networks. A feature of telephone routes dimensioned using the Erlang method is that routes comprising a large number of circuits are proportionately more efficient (i.e. require proportionately fewer circuits per carried call) than smaller ones, when designed to the same grade of service, and are therefore more economic. This is apparent from the general shape of the graph of circuits against traffic-carrying capacity. Figure 32.1 illustrates the graph of 1% loss, or 1% grade of service. Note that the graph is almost linear for circuit group with circuit numbers above 20 circuits, each extra circuit adding approximately 0.85 Erlangs of traffic capacity. Thus by adding ten circuits to a group of

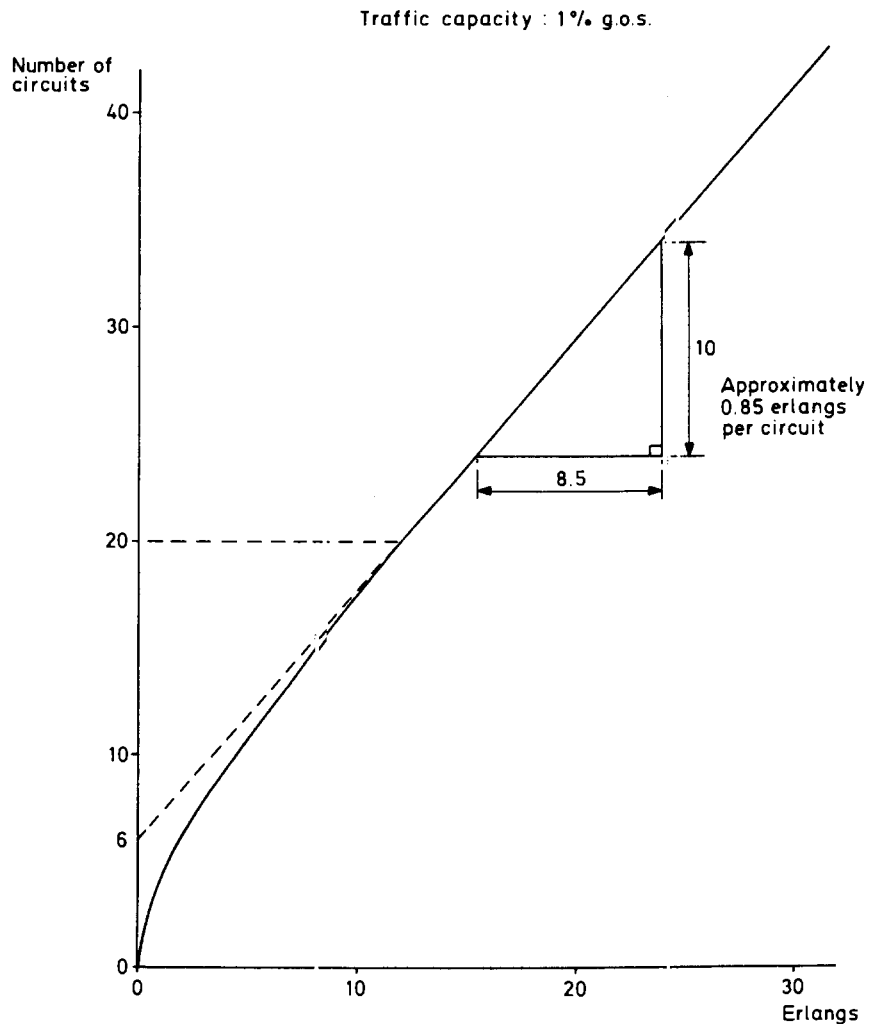


Figure 32.1 The inefficiency of small routes

20 circuits we make it suitable for carrying 20.5 Erlangs rather than 12 (an increase in capacity of 8.5 Erlangs). Notice that an isolated group of 10 circuits is only suitable for 4.5 Erlangs of offered traffic. Smaller circuit groups are thus far less efficient, as Figure 32.1 shows.

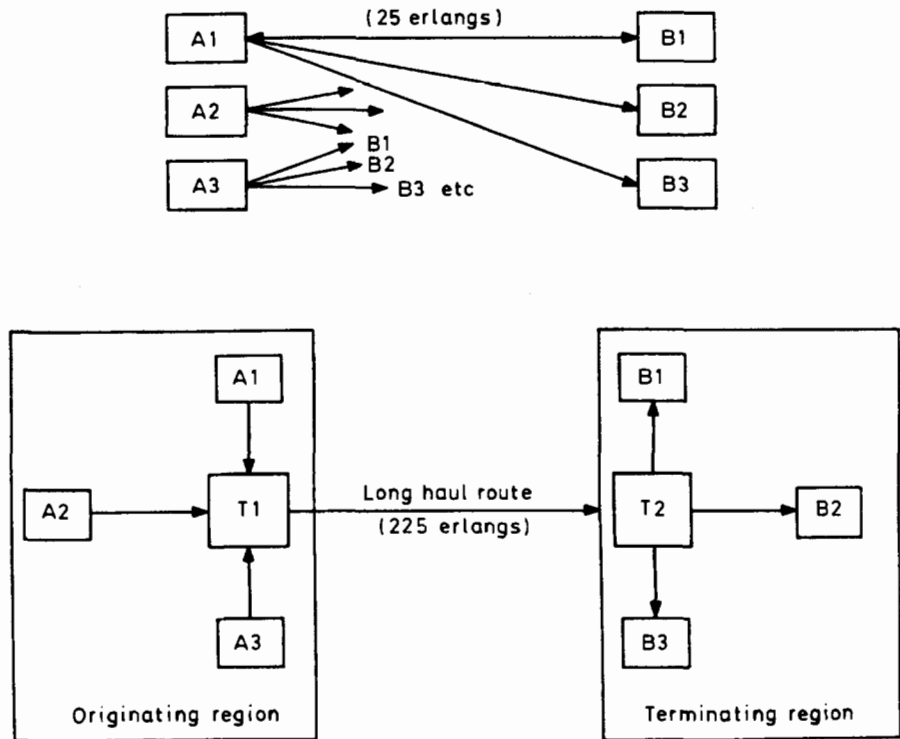
Looking at the graph it is useful to imagine that the first six circuits carry no traffic at all, and that all subsequent circuits have a capacity of 0.85 Erlangs each. It is almost as if there were a 'penalty' price of six circuits for having a route at all! The line representing this assumption is shown in Figure 32.1. The equation of this line provides a handy method of estimating the number of circuits required to carry a given offered traffic load, but it is only valid up to about  $A =$  Erlangs. Above this value, other estimating formulae can be used as follows

$$N = 6 + (A/0.85) \quad \text{for } A < 75$$

$$N = 14 + (A/0.97) \quad \text{for } 75 < A < 400$$

$$N = 29 + A \quad \text{for } 400 < A$$

The easiest way to minimize the circuit penalty which arises on small routes is to adopt a network which has a hierarchical structure, similar to the networks shown in



**Figure 32.2** The benefit of combining small routes via transit switches

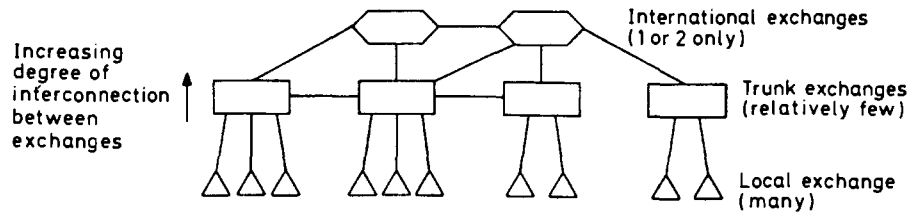


Figure 32.3 Hierarchical network structure

Figures 32.2 and 32.3. In such a structure, a small number of main exchanges are fully interconnected with one another, and various tiers of less important exchanges have progressively less direct connections to other exchanges. If no direct route is available from one given exchange to another, then the call is referred to an exchange in the next higher tier. By using such a structure we can reduce the number of circuits on long-haul routes. Under such a hierarchical scheme, routes are combined so that groups may be dimensioned (or sized) to carry multiple-traffic streams, and so reap the economic benefits of the larger scale. The hierarchical network technique works as shown in Figure 32.2.

Let us assume that the callers on each originating exchange (A1, A2, A3) in Figure 32.2 are generating calls equivalent to a traffic intensity of 25 Erlangs (i.e. an average of 25 simultaneous calls in progress) to each terminating exchange (B1, B2, B3). Let us also assume that exchanges A1, A2, A3 are quite close together, as are exchanges B1, B2, B3, but that region A and region B are a long way apart.

Without a hierarchical network structure, each A exchange would need a set of long-haul circuits to connect with each B exchange, as shown in Figure 32.2(a). This would give a total of nine routes (A1–B1, A1–B2, etc.). Using Erlang's formula for 1% GOS, each of these routes or *circuit groups* would require 36 circuits, amassing a total requirement of  $9 \times 36 = 324$  long haul circuits.

Alternatively, if the traffic is concentrated through collecting exchanges T1 and T2 (called *transit* or *tandem* exchanges) within the originating and terminating regions, only a single route is required. That route has to carry all 225 Erlangs, but only 247 long-haul circuits (again using Erlang's formula) are needed. Figure 32.2(b) illustrates this alternative network structure. Of course, when comparing the cost of the the two structures we must not overlook the additional switches T1 and T2, as Comparison 1 below indicates.

### 32.3.1 Comparison 1: Direct versus Hierarchical Structure

(Traffic between each A–B pair: 25 Erlangs)

All direct routes (as Figure 32.2(a))	324 long haul circuits ( $9 \times 36$ )
Hierarchical structure (as Figure 32.2(b))	247 long haul circuits 2 $\times$ transit switches (T1, T2); 225 Erlangs each 324 short haul circuits (A–T1 and T2–B)

So then, if 77 long-haul circuits (324–247) costs more than two 225 Erlang switches (T1, T2) plus the short haul access circuits (A–T1 and T2–B), hierarchical structure (Figure 32.2(b)) is the more cost effective.

In individual cases the particular circumstances will decide whether a direct or a hierarchical network structure is cheaper, though the general rule applies that in very-long-haul situations (international networks for example) the hierarchical structure usually wins. Hierarchical structure can also be cost-effective when point-to-point route traffic is very small (as in rural networks). To illustrate this point, in Comparison 2 the example of Figure 32.2 has been repeated with much lower traffic values (and with circuit numbers re-calculated, again using the Erlang method).

### 32.3.2 Comparison 2: Direct versus Hierarchical Structure: Low Traffic Only

(Traffic between each A–B pair: 3 Erlangs only)

All direct routes	72 long haul circuits ( $9 \times 8$ ) (8 circuits required for each 3 Erlang route)
Hierarchical structure	38 long haul circuits (for 27 Erlangs) 2 transit switches (T1, T2); 27 Erlangs each 72 short haul circuits

Note how in Comparison 2 the proportion of long-haul circuits saved as the result of a network hierarchy is much greater than it was in Comparison 1 (47% saving as opposed to 24%). This is because the inefficiency of very small routes is very marked, as we saw in Figure 32.1. There are therefore greater proportional savings to be made from combining very small routes.

Hierarchical structure is common in many of the world's telephone and ISDN networks, in which a large number of local exchanges (or end offices) route their trunk traffic via a smaller number of trunk (or toll) exchanges. At the highest tier in the hierarchy there are probably only one or two international exchanges, each lower tier has a greater number of exchanges, but each with only a restricted degree of long-haul inter-connection. Figure 32.3 illustrates a typical national hierarchy.

An added advantage enjoyed by hierarchical networks lies in their capacity for getting the most out of their circuits at times when the busy hours of various transmission routes do not coincide. For an example look back again at Figure 32.2(b), and we can see that if A1–B1 is busy in the morning and A3–B3 in the afternoon, then thanks to hierarchical structure the same circuits can be used for both traffic streams. On the other hand, separate direct circuit groups (as in Figure 32.2(a)) would be inefficient, as one or other of the groups would always be idle.

A disadvantage of hierarchical networks when compared with direct-circuited networks, is their greater susceptibility to congestion under network overload. There are two causes.

- Fewer overall circuits are available in hierarchical than in equivalent direct-circuited networks.

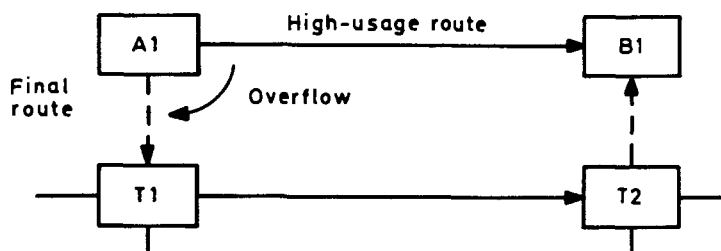
- Congestion between only one pair of exchanges (e.g. A1-B1 of Figure 32.2(b)) will result in congestion on all other routes (e.g. A2-B3), because all calls have to compete for the same circuits (T1-T2). This can rapidly lead to further congestion, as customers dial merrily on regardless.

## 32.4 OVERFLOW OR 'AUTOMATIC ALTERNATIVE ROUTING' (AAR)

A simple means of improving purely-hierarchical networks is to apply the technique of *overflow*, and for this we require the *automatic alternative routing* (AAR) mechanism.

AAR is also known as *overflow* or *alternative routing*, and most exchanges are capable of it. It can be understood as a priority listing of the route choices leading to any given destination. Turning back to our example of Figure 32.2(b), let us assume that we introduce an additional link between exchanges A1 and B1. This might be a first choice, or *high-usage* (HU) route (as described below). In this case, an AAR table for traffic from A1 to B1 might be as shown in Figure 32.4.

Cost benefits can be gained by judicious application of AAR (as in Figure 32.4) rather than using the simple hierarchical structure (as in Figure 32.2(b)), particularly if the route traffic between points A1 and B1 is large. The saving is achieved by making quite sure that the first choice route between A1 and B1 does not have enough circuits



For traffic from A1 to B1 AAR table is :

AAR table	
First choice	Use of the direct route to B1
Second choice (if all first choice circuits are busy)	Overflow the traffic via T1

Figure 32.4 Automatic alternative routing (AAR)



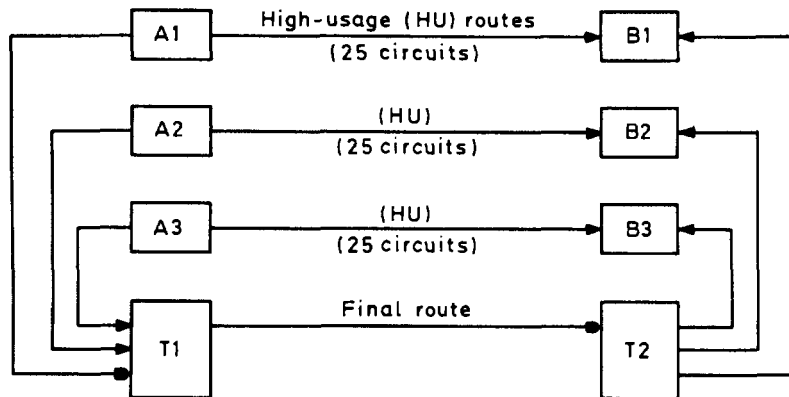


Figure 32.5 The benefit of high-usage working

to carry the whole traffic. A remainder is then left which is *forced* to overflow via T1. Because of the way they are used, the two routes A1–B1 and A1–T1–T2–B1 are called *high-usage (HU)* and *final* routes, respectively. In fact A1–B1 is a *primary high-usage route*, as it is first choice; however, secondary, tertiary, and so on, high-usage routes (i.e. extra route choices between *primary HU* and *final*) may also be used, and they have their place in the AAR table.

To evaluate the savings of high-usage working, let us assume that we provide high-usage routes of 25 circuits between each pair of exchanges in Figure 32.2(b) (i.e. nine HU routes; A1–B1, A1–B2, etc). Then the network adopts the pattern shown in Figure 32.5.

To compare the overall circuit requirement with earlier examples, we must dimension the network to the same overall 1% grade of service performance. However, the dimensioning of the overflow links (A–T1 and T2–B), as well as of the *overflow* or *final route* T1–T2, presents a special problem for which the Erlang formula has to be modified. This is because overflow traffic does not have a random nature as the Erlang dimensioning method requires. Consequently, in Comparison 3, the more complex Wilkinson–Rapp *equivalent random* method (explained later in the chapter) has been employed, and the network has been dimensioned so that all customers receive a grade of service equivalent to or better than 1%. By this method, we determine that 43 *final route* circuits are required on T1–T2. We thus need  $(9 \times 25) + 43 = 268$  circuits in total. Comparison 3 below compares this value with the direct and the hierarchical network structures.

### 32.4.1 Comparison 3: Direct versus Hierarchical versus HU Structure

(Traffic between each A–B pair: 25 Erlangs)

All direct routes:	324 long haul circuits
Hierarchical structure:	247 long haul circuits
	2 × transit switches (225 Erlangs each)
	324 short haul circuits

Overflow structure:	268 long haul circuits
	$2 \times$ transit switches (32 Erlangs each)
	99 short haul circuits

Comparison 3 shows that it is possible (by choosing the optimum size of high-usage (HU) circuit groups) to reduce significantly the size of the transit switches required (between the *hierarchical* and *overflow structures*), without significantly increasing the long-haul circuit requirement. This clearly reduces the cost. However, our example is not very realistic. In practice, if any of the individual traffic streams are too small, the benefit of providing high-usage circuits for the corresponding direct route disappears. In practice then, it is often useful to provide high-usage routes only for those traffic streams exceeding a given Erlang threshold value. In other words, 'if more than  $X$  Erlangs of traffic exist between any two nodes, then a direct HU route is justified'. The value of  $X$  will depend upon the relative costs and of lineplant and exchanges equipment and on whether the exchanges are local, trunk, or international ones.

A problem facing planners using overflow in practical networks is that the number of circuits required cannot always be cut to a minimum without adding another control mechanism such as trunk reservation, as explained later in this chapter.

Furthermore, the high-usage structure is not as advantageous as the hierarchical when route busy hours of the various traffic streams do not coincide, because the circuits are 'less available' for shared use, e.g. by one stream in the morning and by another in the afternoon.

HU structures do, however, have a better overload performance, in that individual traffic streams are to some extent protected from congestion on other routes (caused for example by very high seasonal or short term demand).

## 32.5 WILKINSON-RAPP EQUIVALENT RANDOM METHOD

In the previous section when we were attempting to dimension routes carrying overflow traffic we ran into the difficulty that because the distribution of traffic which is *overflowed* from a high usage circuit group is not random, the Erlang formula may not be used directly. So, introduce the Wilkinson-Rapp *equivalent random* method instead. This is in fact the Erlang method slightly adapted.

Imagine two traffic streams  $a$  and  $b$  which have high-usage routes of the A and B circuits available, respectively. These two traffic streams overflow amounts of traffic  $a'$  and  $b'$  onto a common final route of  $N$  circuits, which is also the first choice for the *first-offered* traffic stream  $c$ . Figure 32.6 illustrates this schematically. The problem is to find the value of  $N$  which will guarantee the appropriate grade of service on all traffic streams.

Circuit group  $N$  is subjected to overflow traffic  $a' + b'$  and first-offered traffic  $c$ , and we dimension it by the Wilkinson-Rapp *equivalent random* method as follows. The method assumes that the  $N$  circuits will behave in the same way as a set of  $N$  circuits part way down a larger group of  $(N_{\text{EQ}} + N)$  circuits when subjected to an imaginary single stream of *equivalent random traffic*,  $A_{\text{EQ}}$ . Figure 32.7 illustrates the set of  $N$  circuits and the imaginary set  $N_{\text{EQ}}$ . The problem is finding the values of  $N_{\text{EQ}}$  and  $A_{\text{EQ}}$  so that the traffic overflowed from the  $N_{\text{EQ}}$  circuits exactly matches the characteristics of the real

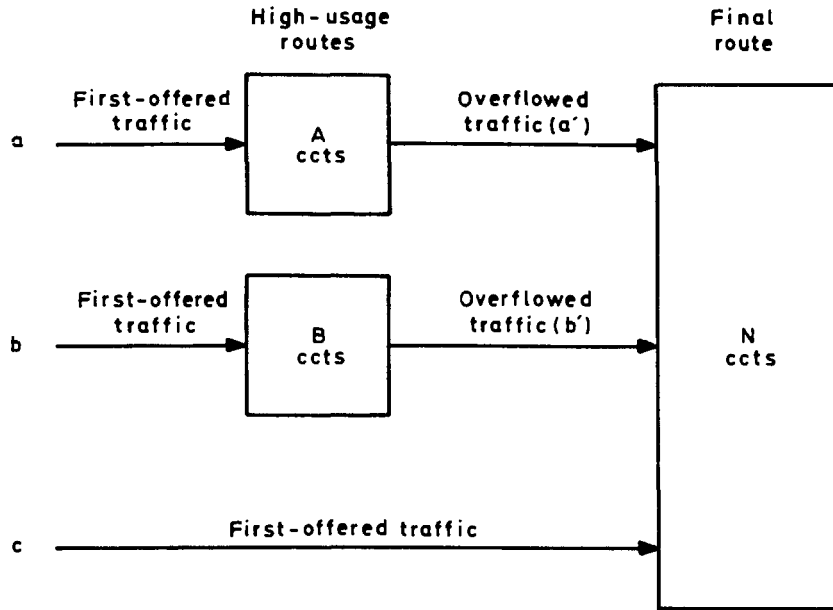


Figure 32.6 Dimensioning final routes

traffic  $a' + b' + c$ . The mean  $M$  and variance  $V$  characterize the traffic  $a' + b' + c$  which we imagine to overflow from the  $N_{EQ}$  circuits, and by choosing the values of  $A_{EQ}$  and  $N_{EQ}$  carefully we can equate the values exactly with the corresponding mean and variance values of the real traffic

By knowing the values  $A_{EQ}$  and  $N_{EQ}$ , the traffic lost by the  $N$  circuit group can easily be determined. It is done by using the normal Erlang formula, assuming that an *equivalent random* traffic value  $A_{EQ}$  is offered to a total number of circuits  $N + N_{EQ}$ . The *grade of service* determined by the formula is the overall lost traffic ( $L$  in Figure 32.7) quoted as a proportion of the imaginary original traffic value  $A_{EQ}$ .

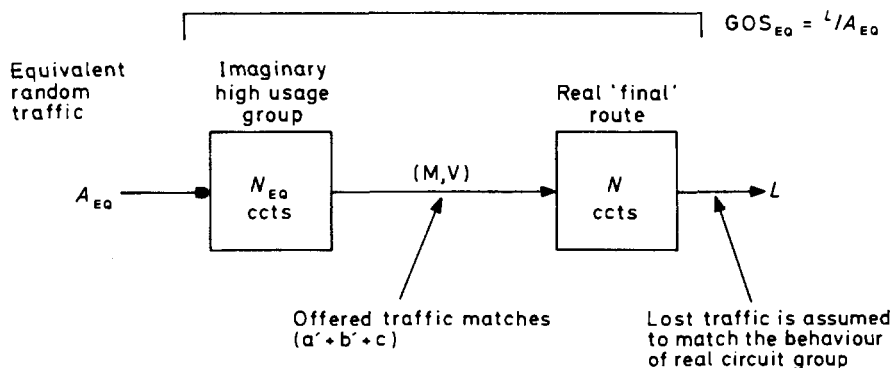


Figure 32.7 The Wilkinson-Rapp equivalent random method

It is normal to dimension the group of  $N$  circuits by first calculating the permissible traffic loss in Erlangs (i.e. the amount of the real traffic which need not be carried,  $L$  in Figure 32.7), and then calculating this as a proportion of  $A_{EQ}$ . This is the imaginary or *equivalent* grade of service required when traffic  $A_{EQ}$  is offered to  $N + N_{EQ}$  circuits, and allows us to calculate the value of  $(N + N_{EQ})$  using the Erlang method as follows

$$\text{maximum permissible lost traffic} = L$$

$$L = \text{GOS required} \times \text{traffic on } (a' + b' + c)$$

$$\text{imaginary GOS required on } N + N_{EQ} \text{ circuits} = \text{GOS}_{EQ} = L$$

$$N + N_{EQ} = E(A_{EQ}, \text{GOS}_{EQ}), \quad \text{where } E \text{ represents the Erlang formula}$$

Having determined the value of  $N + N_{EQ}$ , the real number of circuits required for the real traffic ( $N$ ) is found by subtracting value  $N_{EQ}$ .

The values  $N_{EQ}$  and  $A_{EQ}$  are relatively straightforward to calculate, but the process requires a number of mathematical steps. An appendix at the end of the chapter shows how to calculate these values and gives an example of the whole Wilkinson–Rapp overflow route dimensioning method for those readers who are interested.

The Wilkinson–Rapp method is now quite old (1956) and in consequence more complex methods have evolved which seek to improve it. As with the alternatives to the basic Erlang method it rests with the user to decide the method which suits his circumstance the best.

## 32.6 DIMENSIONING 'FINAL ROUTES'

It is normal practice to dimension final routes for only a 1% loss of the mean traffic offered to them. This ensures a 1% grade of service for any traffic which is *first offered* to the final route. This practice was adopted in the preceding section when mixed overflow and first-offered traffic shared the same circuit group.

In instances where there is no *first offered* traffic on the final route, then the circuit numbers may be reduced, commensurate with an overall 1% loss on each of the individual traffic streams. Thus if only 10% of traffic overflows from the high usage route (a typical value), only 10% of the original traffic is offered to the final route. The caller experiences a net 1% grade-of-service even if as much as 10% of the traffic offered to the final route is lost. Thus the grade of service of the final route need only be 10% in this case!

## 32.7 TRUNK RESERVATION

To some *trunk reservation* is a relatively new technique, for although it has been around for more than 30 years, only recently has the advent of stored program control (SPC) exchanges made it more widely available. It provides a method of *priority*

*ranking* traffic streams, which can be particularly useful in conjunction with overflow network schemes, when the network planner may wish to give higher priority to *first-offered traffic* than to *overflow traffic*.

As we saw in the example shown in Figure 32.6, final routes must be dimensioned large enough to ensure the chosen grade of service (usually 1%) for any *first-offered traffic* streams. The consequence of this is an unavoidably and unnecessarily good grade-of-service for the overflow traffic (far better than 1%), and the penalty regrettably is the need to provide extra circuits over and above the minimum. Trunk reservation, however, gives a way out.

Developing our example of Figure 32.2(b) a little further to illustrate the principle of *trunk reservation*, let us now recognize that in practice the route traffic between individual A and B nodes is not identical in value and varies widely. Let us assume that exchanges A1, A2, B1, B2 and B3 are in large towns, and A3 is in a much smaller town and consequently generates and receives much less telephone traffic. It might well be that while A1, A2, B1, B2 and B3 justify direct interconnection by high-usage (HU) routes, overflowing via T1 and T2, the traffic originated at A3 might be insufficient to justify such direct HU groups. In this case the traffic passing over the circuit group from T1 and T2 is a mixture of

- overflow (i.e. not all) traffic generated by A1 and A2
- all traffic generated by A3

Under such circumstance, it seems reasonable to take some precaution to ensure that the A3-originated traffic has some type of priority for use of the T1–T2 circuits so that the grade of service to A3 customers is on a par with the performance available to customers at A1 and A2. Such a mechanism is provided by one form or another of *trunk reservation*.

Numerous slight variants of *trunk reservation* exist. In the simplest, each traffic stream competing for a given route is allocated a trunk reservation value (typical values range from 0 to 15). The value corresponds to a number of idle circuits. If at any given moment only this number of idle circuits are available within the group (the others being busy), then new calls from the particular traffic stream to which the trunk reservation value applies, are rejected (i.e. are given network-busy tone and not completed). In other words, the trunk reservation value ‘disadvantages’ the traffic stream, to the benefit of other streams: the larger the value set, the larger the ‘disadvantage’.

Each traffic stream competing for the circuit group may be allocated a different trunk reservation value (or disadvantage) but at least one stream should be given value 0 (top priority), otherwise some circuits will always be idle.

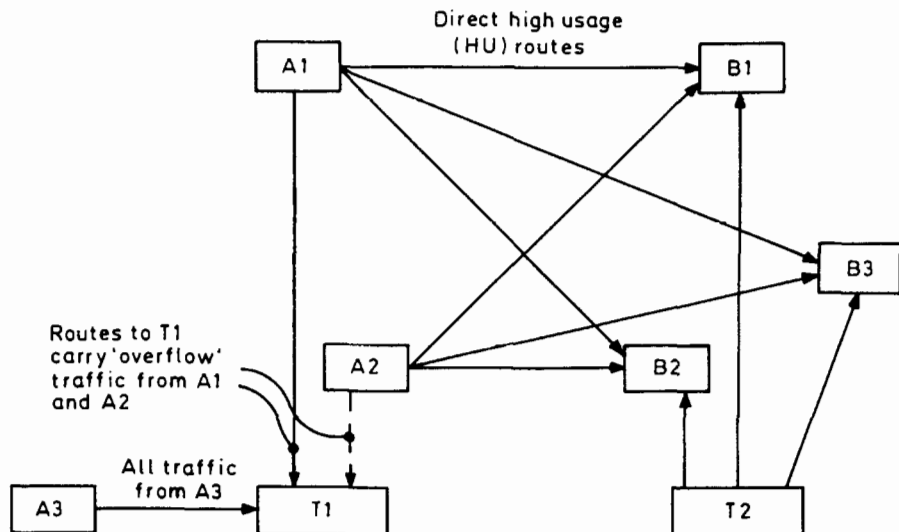
By using *trunk reservation* we can ensure that the various traffic streams are completed according to a priority order during periods when the circuit group is heavily loaded (i.e. few circuits are idle). In this manner we can nearly equalize the grades of service of the traffic streams, by disadvantaging overflow traffic streams which have had previous route options.

When further circuits become free, fewer traffic streams are disadvantaged, so that more of the traffic streams are allowed to complete calls. It is important to note that the algorithm does not take into account which particular circuits within the group are free, only the total number which are idle.

For our example illustrated in Figure 32.8 it would probably be appropriate to set values as shown in Figure 32.8.

The mathematics needed to calculate *trunk reservation* values are exceptionally complex. However, it is found in practice that values of 3 or 4 establish a significant priority and can be used in a trial-and-error fashion. More accurate calculations are possible by the use of mathematical tables or computer programs, but as any one value of trunk reservation has much the same effect on all sizes of traffic streams, whether large or small, this only strengthens the case for trial-and-error use of values 3 and 4.

Trunk reservation is highly effective in providing a priority list based on 'disadvantage', and a value as low as 15 may virtually cut off a given stream. As we shall see later in the chapter, routing techniques are moving from hierarchical (i.e. structured *overflow*) schemes towards non-hierarchical (i.e. dynamic) schemes. In consequence, new teletraffic modelling methods are necessary, capable of modelling end-to-end congestion rather than simple link-by-link analysis. Simple capacity flow models are used



Traffic source	Trunk reservation value (applied at exchange T1, according to the source of traffic)
A3	0 (top priority)
A1 or A2	3 or 4 (typical value)

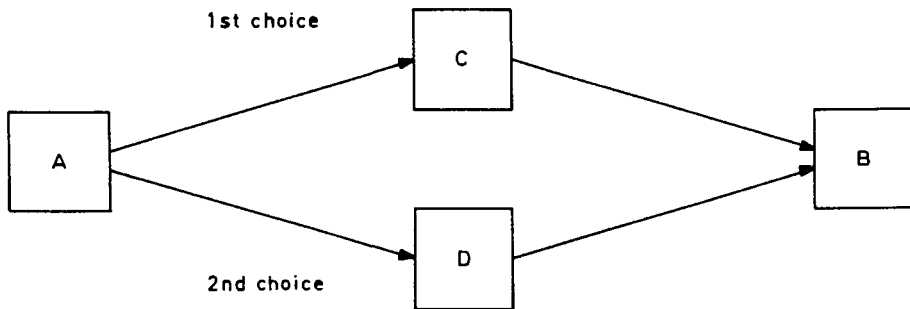
Figure 32.8 Typical trunk reservation values

for this analysis, but the technique of trunk reservation is no less relevant in protecting priority streams on particular links. In a dynamic routing environment, *trunk reservation* values need to be higher than in simple overflow schemes (i.e. AAR), typically 10.

### 32.8 'CRANKBACK' OR 'AUTOMATIC RE-ROUTING' (ARR)

A more powerful form of overflow routing is that of automatic re-routing (ARR), also called *crankback*. This mechanism, which relies on sophisticated switch and signalling interaction, enables an alternative transit route to be chosen even if it is the second link of a previous transit route choice that is busy. Figure 32.9 shows how it works.

This mechanism is particularly useful in highly interconnected networks; it gives much better grade-of-service where most of the available paths between two points are via transit exchanges, and it is made up of a multiple number of links.



Alternative routes A to B	
ARR first choice	Via : C (i) if A_C congested then ARR via A_D (ii) if A_C free and C_B free then forward call to B otherwise if C_B busy go on to ARR second choice
ARR second choice	Via : D ( if D_B busy then fail the call )

Figure 32.9 Automatic re-routing (ARR)

### 32.9 PROPORTIONATE BIDDING FACILITY (PBF)

When there is more than one route to a given destination, it is sometimes an advantage to be able to control the exact proportion of traffic offered to each route. *Proportionate bidding* provides a mechanism for this purpose. Call attempts are 'dealt' in turn around all the available routes according to a set percentage regime (one for you, three for you, two for you, etc.). The 'dealing' is made by an exchange software program, which works in almost the same way that a croupier deals a pack of cards. Both calls and cards are subdivided into random fractional samples. Figure 32.10 provides an example.

Useful applications of this algorithm might be as follows.

**Route balancing:** suppose that exchange 'A' is a trunk exchange, and B and C are international exchanges with say 60% and 40%, respectively, of circuits to a given overseas country. Traffic can be split by 'A' and routed to exchanges B and C in appropriate fractions (60/40), using PBF as a routing mechanism at exchange A.

**Multiple carrier environment:** suppose this time that exchanges A, B, C, D are international exchanges, A being in one country while B, C and D are owned by separate carriers (competing telecommunications companies) in another. Outgoing traffic from exchange A can be split, using PBF, in the same proportion as incoming traffic is received from carriers B, C and D. PBF in this case provides a fair proportionate return solution to the problems of connecting with multiple carriers. The method has already been adopted by a number of telephone network operators as a means of interconnecting with multiple telecommunications carriers operating in a destination country.

### 32.10 DYNAMIC ROUTING

Another new technique is that of dynamic routing. A number of slightly different *dynamic routing* techniques have been developed, but all of them work on the same general premise: that the traffic pattern in a large network is continually changing

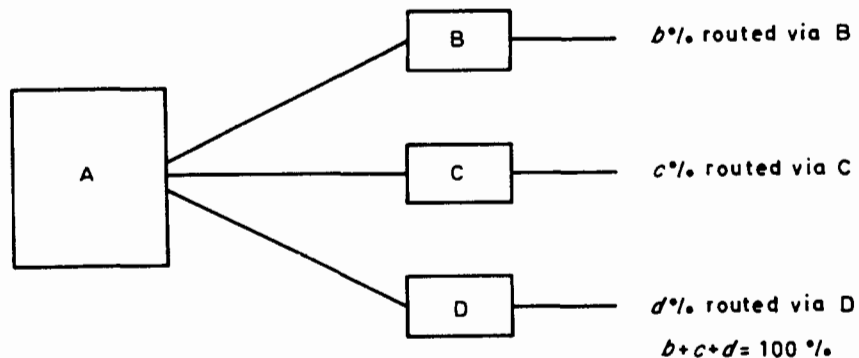


Figure 32.10 Proportionate bidding facility (PBF)



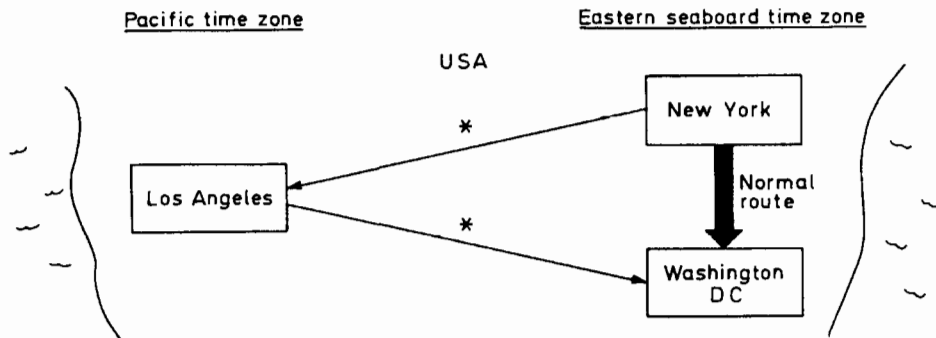


Figure 32.11 Dynamic routing

according to the time of day, the day of the week, and the season of the year. At any given time some routes will be busy while others are slack. If, then, one could employ some of the slack capacity to serve the busy routes, there would be an overall saving in network resources, even though some of the individual calls might have to be routed circuitously via exchanges in other timezones. Figure 32.11 shows an example of traffic being routed from New York to Washington via Los Angeles.

In the Figure 32.11 example, as customers in Los Angeles are waking up and are starting to ring New York, the overall routing pattern needs to be adjusted dynamically, to prevent the New York/Washington traffic from causing congestion on the trans-USA links. At this time of day European and African exchanges will be downloading as their business days end, and these switches could be used as alternative transit points for the New York/Washington traffic (instead of Los Angeles). Clever control mechanisms are needed to achieve optimum routing patterns, and to keep adapting the network so as to maintain the optimum. A number of such mechanisms have been patented. They include American Telephone and Telegraph's (AT&T's) *dynamic non hierarchical routing (DNHR)* and British Telecom's (BT's) *dynamic alternative routing (DAR)*.

## 32.11 ROUTING AND TRAFFIC CONTROL IN DATA NETWORKS

In contrast to circuit-switched networks, in particular the public telephone network, data networks (e.g. packet, frame, ATM, LAN, MAN and *Internet* networks) have tended to use more sophisticated automatic routing control techniques. Typical in telephone networks is that the routing table in each individual exchange must be manually programmed (as we saw in Figure 32.4). In data networks, more sophisticated routing protocols and algorithms (e.g. *spanning tree protocol*, a *source routing* protocol or some other protocol such as the TCP/IP method *OSPF, open shortest path first*) are generally used to select the route to the destination. The advantage is the simplicity of setting up and maintaining routing mechanisms in complicated networks, the disadvantage is the difficulty in measuring and predicting traffic flows.

In the case of *path-oriented*, connection-oriented networks, all the packets, frames or cells take the same route through the network. Path orientation helps to ensure that the

packets arrive in the same order in which they were sent, thus minimizing the job of any necessary re-sequencing. The delays encountered by each of the packets is similar, thus giving approximately even performance of the network as perceived by the user. In addition, the path traffic volumes are relatively predictable and sometimes manually programmable using techniques similar to those discussed for *circuit-switched* networks.

The alternatives to path-oriented networks are *datagram networks*. These are networks in which individual packets constituting a given communication find their own individual way through the network. The advantage of the *datagram* approach, as opposed to the *path-oriented* approach (both are illustrated in Figure 32.12) is the greater ability to spread traffic more evenly through the network according to instantaneous data packet loading and availability of individual links.

As not all the packets take the same route through a *datagram* network, it is more difficult for a third party to try to 'tap' the call, also making it more secure. Datagram networks are also more adept at coping with individual link failures within the network on a dynamic basis. Packets simply divert via another route. When a link failure occurs in a path-oriented network, the network usually takes a little longer to notice the failure, and then usually requires a little time to determine a new path. The *virtual connection* is not usually lost, but for maybe a few seconds (say around 10) it may not carry any further packets until the path is re-established over a new route.

In *connectionless* networks, a *virtual connection* is not established. Instead, each of the packets or frames making up the message is simply despatched into the network in a manner similar to *datagram* switching. There is no confirmation of receipt of messages. The main advantage of connectionless networks is the effort saved in establishing connections and the high security resulting from the use of different paths and the absence of identifiable connections. *Connectionless* routing is used widely in the *Internet*.

As is evident from the above, data networks have much more powerful automatic routing capabilities than has historically been the case in public telephone networks. This significantly reduces the manual effort required to establish and administer routing tables. In addition, it enables complex router networks (such as the *Internet* itself) to be put together simply by plugging a large number of sub-networks operated by different organizations together, without the need for an organization planning the network and routing as a whole.

So much for the advantages of automated routing algorithms and protocols; the disadvantage is the resulting chaos. The *Internet*, for example, has historically lacked a rigid structure and tight overall network planning authority, with the result that no-one

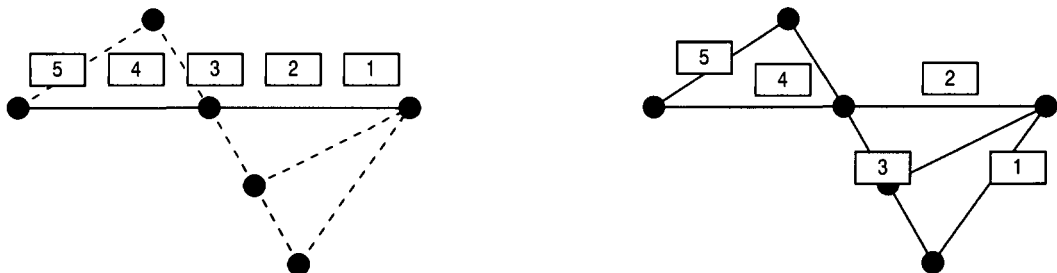


Figure 32.12 Packet, frame or cell routing in data networks

can accurately predict the path taken by an individual message traversing it. This leads to significant difficulties in designing an appropriate overall topology and ensuring adequate quality of network performance. Many customers complain of unacceptable response times when 'surfing' the *Internet*.

To optimize the topology of a data network requires a thorough understanding of the protocols and algorithms at work in determining the routing of individual connections and packets or frames. Thus, for example, many different *metrics* and *attributes* may be used in the determination of the appropriate path, as we discussed in Chapter 28.

It is a difficult job to achieve optimum traffic flow through a data network which uses complex routing protocols and algorithms. It requires very careful establishment within the network configuration database of the values of each of the metrics and attributes for each of the nodes and links, though this may be to some extent automated within the *network management system*. The easiest way to ensure success, as in telephone networks, is the use of a hierarchical network structure. Such a structure is rapidly appearing within the Internet, where only about four major nodes (all in the USA) switch most of the messages traversing the network.

## 32.12 NETWORK DESIGN

A number of traffic-controlling mechanisms have been presented in this chapter. In practice, any one or any combination of mechanisms may be used by network planners in striving towards the perfect network; and always the optimum topology makes some form of compromise between least cost, overall performance and robustness under overload. The constraints will arise from the control mechanisms actually available in any given case. None of the constraints and impositions are ever the same. All in all, network design is indeed a complex and imprecise art!

Figure 32.13 provides a flowchart for network design, outlining the steps leading to the determination of a suitable topology. It starts by assuming that a network already exists and that the current point-to-point traffic demand between exchanges can be obtained by measurement. If this is not the case, current demand will have to be estimated. Alternatively, the scope for using an entirely new network configuration might be considered.

Having measured the current traffic, the next step is to identify any problems of congestion or unreliable equipment. We must then estimate all the future requirements of the network, first by forecasting the future traffic demand (see Chapter 31), and then by reviewing any other reasons for undertaking network changes.

The network may need to be changed to accommodate routes to new exchanges, or to connect to networks in previously unserved areas. Alternatively, its topology may need to change for economic or service reasons. Instead of being one made up of a large number of small outdated exchanges, it may become one comprising a small number of large, modern technology ones. Often these factors have the most impact on the evolving network structure.

When the future demands on the network are known, we review the efficiency and suitability of the current network. For its intended future purpose we determine as many different options as possible for its evolution. At the simplest, an option may only require

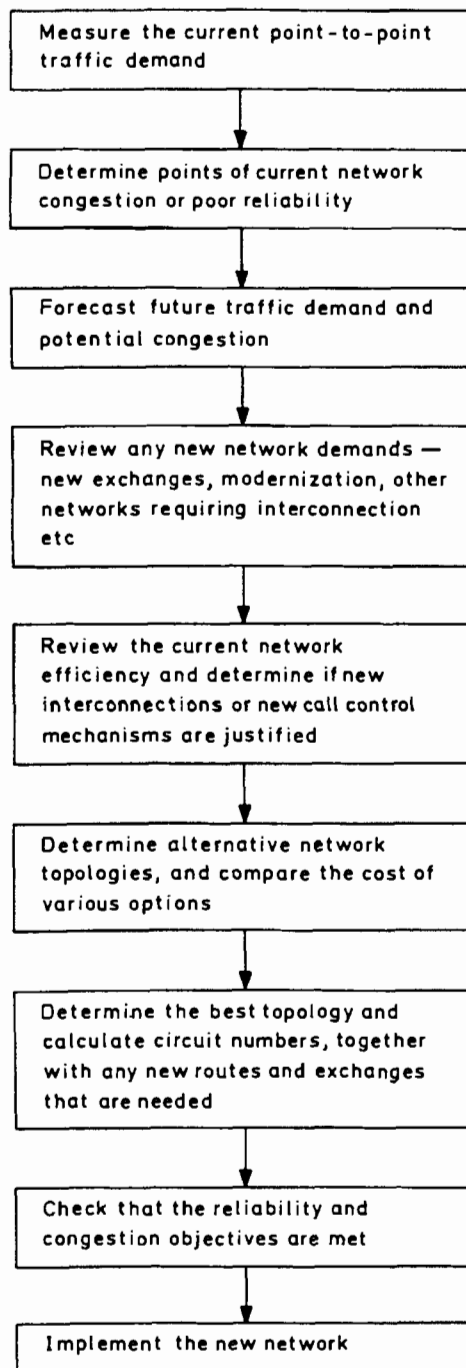


Figure 32.13 Network design

adjustment to the capacity of the exchanges or of the transmission links interconnecting them. More complex options might demand the use of new call control techniques (such as ARR, PBF or dynamic routing). However, the deployment of such techniques is likely to incur considerable cost and needs to be weighed against the benefits, both in overall network performance and in overall network cost.

Once the choice has been made, the final circuit numbers can be determined and a plan worked out for the new circuits, routes and exchanges that will be needed. Then it is up to the implementors.

Depending on the nature of the network and its customers, the plan will need periodic, if not frequent, review and readjustment. Nonetheless, the more comprehensive and forward looking is the plan, the greater is the likelihood of smooth network administration and performance. Contrariwise, insufficient forward forecasting or lack of anticipation of future needs can lead to short-term 'firefighting', with a need to adjust continually the structure of the network. Under such circumstances few networks reach their optimum of performance and economic efficiency.

### 32.13 APPENDIX: THE WILKINSON-RAPP ROUTE DIMENSIONING METHOD

In the main part of the text we briefly described the Wilkinson-Rapp method for dimensioning overflow circuit groups, but although we described the method we did not demonstrate how to calculate the values  $A_{EQ}$  and  $N_{EQ}$ . For those interested readers we now cover the necessary calculations and give an example of the use of the method. The explanation refers back to the diagrams of Figures 32.6 and 32.7 and describes the method in mathematical notation.

In Figure 32.6 the traffic offered to the  $N$  circuit group is

$a' + b'$  overflow traffic

$c$  first-offered traffic

The question is 'how many circuits ( $N$ ) are required in order that a 1% GOS is achieved?' The value can be calculated as below.

First, the separate streams within the offered traffic are characterized in mathematical terms by the statistical means and variances of their traffic intensities, which are calculated using the normal Erlang formula. Provided that the streams are statistically independent of one another, then the mean and variance of the combined stream  $a' + b' + c$  can be calculated by simple summation. These values in turn allow  $A_{EQ}$  and  $N_{EQ}$  to be worked out.

The method is then as explained in the main text of the chapter. Refer to Figure 32.7. The actual traffic which may be lost,  $L$ , is calculated as  $1\% \times (a' + b' + c)$  and from this the equivalent grade of service  $GOS_{EQ}$  can be worked out.

Finally, the total size of the equivalent circuit group  $N + N_{EQ}$  is obtained from the normal Erlang formula and the value of  $N$  deduced by subtraction of value  $N_{EQ}$ .

There now follows a short summary in mathematical notation.

First, the mean of the traffic overflowing from the group of  $A$  circuits is worked out by multiplying the traffic offered,  $a$ , by the grade of service experienced. The grade of service is obtained using the Erlang method (although the Erlang formula must be used in the manipulated form that allows the grade of service from the offered traffic and the number of circuits.

$$\text{grade of service, GOS} = E(A, a)$$

(the Erlang formula, see Chapter 30) thus mean overflow traffic is

$$M(a') = a \times E(A, a)$$

The statistical *variance* of the same traffic comes from the formula

$$V(a') = M(a') \times \{1 - M(a') + a/(A + 1 + M(a') - a)\}$$

Similarly  $M(b')$  and  $V(b')$ , the mean and variance of the  $b'$  stream are calculated as

$$M(b') = b \times E(B, b)$$

$$V(b') = M(b') \times \{1 - M(b') + b/(B + 1 + M(b') - b)\}$$

The mean and variance of the first offered traffic are both equal to  $c$ . This fact results from the assumption that the arrival rate of first-offered calls is entirely random. This contrasts with the two cases above, where the variance is greater than the mean because the overflowed calls tend to arrive only in short bursts of 'peaky' activity, the HU having removed the more predictable and steady 'base-load'. Thus

$$M(c) = V(c) = C$$

The method requires us to calculate the mean and variance of the total traffic offered to the  $N$  circuit group. This is done by addition. Let us use  $M$  and  $V$  to represent the total. These values are shown on Figure 32.7.

$$M = M(a') + M(b') + M(c)$$

$$V = M(a') + V(b') + V(c)$$

This now allows us to calculate an important characteristic of the offered traffic stream known as the *peakedness*. The peakedness measures the tendency of calls to arrive together, or in 'peaks', rather than in a more uniform rate of arrival. The peakedness is the value obtained by dividing the offered traffic variance, by the offered traffic mean. Thus

$$\text{Peakedness} = Z = V/M$$

As we saw above, normal or *first-offered traffic* has a *peakedness* of value 1 (i.e.  $M = V$ ), and overflowed traffic has a peakedness value higher than 1 (i.e.  $V > M$ ).

Finally, the equivalent random traffic  $A_{EQ}$  and the equivalent (imaginary) high usage group size  $N_{EQ}$  (as shown in Figure 32.7) can now be determined as

$$A_{EQ} = V + 3Z(Z - 1)$$

$$N_{EQ} = \frac{A_{EQ}(M + Z)}{M + Z - 1} - M - 1$$

where

$M$  = mean traffic load offered

$V$  = offered traffic variance

$Z$  = peakedness =  $V/M$

Next we calculate the imaginary grade of service required of the imaginary circuit group  $N + N_{EQ}$ . This is equal to the permissible lost traffic ( $L = 0.01 \times M$ , for 1% GOS) divided by the equivalent random traffic  $A_{EQ}$ .

$$GOS_{EQ} = L/A_{EQ}$$

Finally we determine the total number of circuits required in  $N + N_{EQ}$  to meet the grade of service  $GOS_{EQ}$  when offered traffic  $A_{EQ}$  using the normal Erlang method. We then deduce the number of real circuits required ( $N$ ) to carry the real traffic  $M$  by simple subtraction

$$GOS_{EQ} = E(A_{EQ}, \{N + N_{EQ}\})$$

(normal trial-and-error Erlang method used to determine the value  $\{N + N_{EQ}\}$ ; see Chapter 30)

$$N = (N + N_{EQ}) - N_{EQ}$$

---

## *Practical Network Transmission Planning*

---

Transmission media exist to convey information across networks. No matter what form the information takes (be it voice, video, data, or some other form), the prime requirement is that the information received at the destination should as closely as possible match that originally transmitted. The signal should be free from noise, echo, interference and distortion, and should be of sufficient strength (or volume) as to be clearly distinguishable by the receiver. The reliability of the service is also important. Meeting these objectives requires careful network transmission planning. This chapter covers two aspects of planning, first describing the electrical engineering design guidelines usually laid out in a formal network ‘transmission plan’, and then going on to discuss the general administrative and operational practicalities of ‘lining-up’ and operating a transmission network. Resource management is crucial to ensure that lineplant and circuits are available when needed, that radio bandwidth is available without interference, that cables have been laid and satellites put in orbit.

### **33.1 NETWORK TRANSMISSION PLAN**

A set of guidelines, laying down simple ‘rules’ for planning and commissioning new line systems or new circuits, is usually set out in a formal *transmission plan*. These guidelines are intended to ensure that the electrical principles of telecommunications theory are adhered to. The transmission plan should include, for example, stringent rules on the use, positioning, and strength of amplifiers or *regenerators* to correct the effects of attenuation. The plan needs to give guidelines on minimization of noise and interference, on the use of echo control devices to eliminate echo, and on the use of equalizers to rectify frequency attenuation distortion. A good transmission plan is a guideline for network design, ensuring that circuits are electrically stable and perform to a standard acceptable to the majority of their users. A typical target is ‘to ensure that 90% of people consider the performance to be “fair”, or “better”’. Acceptable values of attenuation and distortion can be established by subjective tests, and then the network can be designed to accord with them.



ITU-T's recommendations on transmission planning include design guidance on

- the overall signal volume at all points through the network
- the control of signal loss and the circuit's electrical stability
- the limits on acceptable signal propagation times (excessive propagation times manifest themselves as a delay in transmission, a potential cause of slow data throughput and response times, or of unacceptable gaps and pauses in conversation)
- the limits on acceptable noise disturbance
- the control of *sidetone* and signal *echo* (both these effects manifest themselves to speakers, who hear their own words echoed back to them)
- the minimization of signal *distortion*, *crosstalk* and *interference*
- (for digital circuits) maximum allowed bit error rate (or bit error ratio, BER), jitter and quantisation distortion
- maximum line length (e.g. approximately 15 metres for V.24 interface, 100 metres for X.21, etc.)

The recommendations lay down a standard reference system, against which national and international transmission networks may be designed or calibrated. These facilitate the correct placing of circuit conditioning equipment, and the establishment of correct signal volumes and controlled levels of signal distortion at all points along a connection.

### 33.2 SEND AND RECEIVE REFERENCE EQUIVALENTS

When designing telephone or other voice transmission networks, the strength of the electrical signal induced by the microphone at the sending end and the final sound volume produced by the receiver is all-important. It varies according to the proximity of the speaker to the microphone and the listener to the ear piece, the total loss being composed of three basic parts; the loss incurred when inducing the electrical signal in the microphone, the electrical signal loss across the transmission network itself, and finally the sound induction loss in the receiver. The relationship is straightforward to understand, as Figure 33.1 illustrates, but the problem for the network designer lies in deciding just how much signal loss is acceptable in the main part of the network itself.

With no overall loss, the listener will hear too loud a signal, equivalent to a talker at normal volume speaking directly into his ear. On the other hand too much loss will result in an inaudible signal.

To ease the design dilemma, the system of *reference equivalents* was developed by CCITT (the forerunner to ITU-T). The *send reference equivalent* (SRE) of a telephone microphone is the signal loss expressed in dB (decibels) when the electrical signal volume is compared with the original speech volume. Similarly, the *receive reference equivalent* (RRE) of the earpiece is the signal loss when the final (heard) sound volume is compared with the electrical signal volume input to the earpiece. Both SRE and RRE

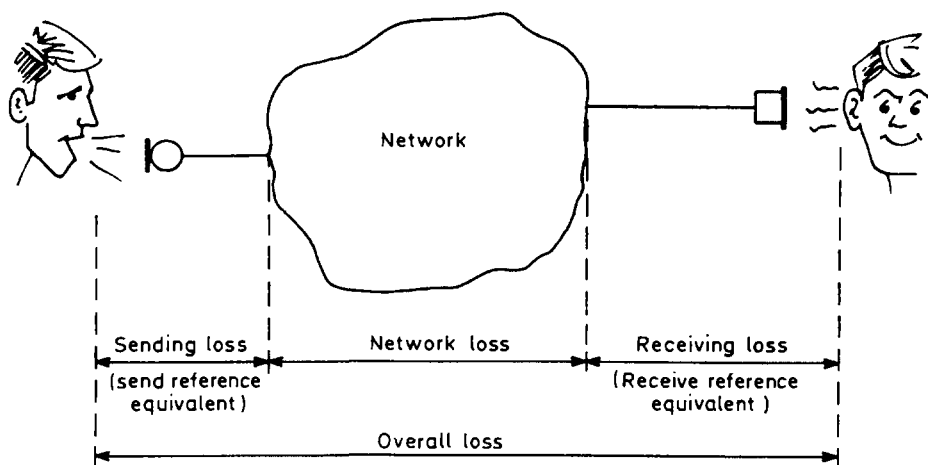


Figure 33.1 Send and receive reference equivalents

are shown on Figure 33.1. To measure the sending reference equivalent (SRE) of a handset, a special *NOSFER* equipment such as that shown in Figure 33.2 is used. This compares the equipment with a standard device, named after this system of calibration, *nouveau système fondamental pour la détermination des équivalents de référence* (new fundamental system for determining reference equivalents).

The *NOSFER* equipment in Figure 33.2 consists of a high quality microphone with a device which maintains it at a fixed distance from the speaker's mouth. This equipment is connected to the listener's earpiece via a large attenuator, of strength  $A$  dB. To perform the calibration, a person talks alternately into microphone  $a$  and microphone  $b$ , while the listener adjusts the variable attenuation  $B$  to a value at which the volumes appear to be the same. The difference between values  $B$  and  $A$  will then give a measure of the relative microphone efficiencies. The more that attenuation  $B$  must be reduced (i.e. the lower the value of  $B$ ) the less efficient the microphone and the larger the numerical value of  $A - B$ , which is termed the *send reference equivalent* (SRE). In other

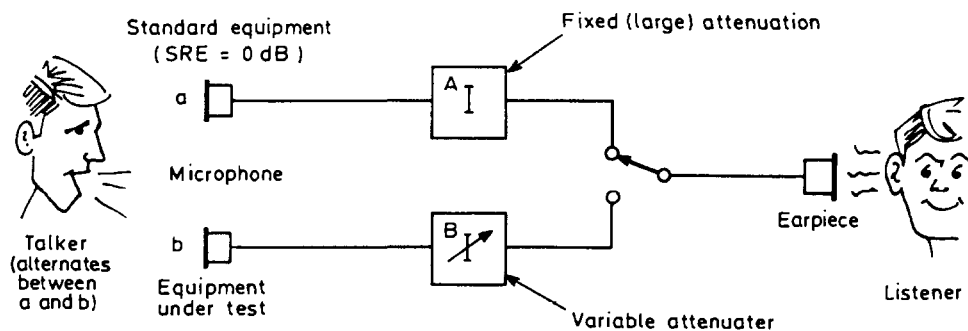
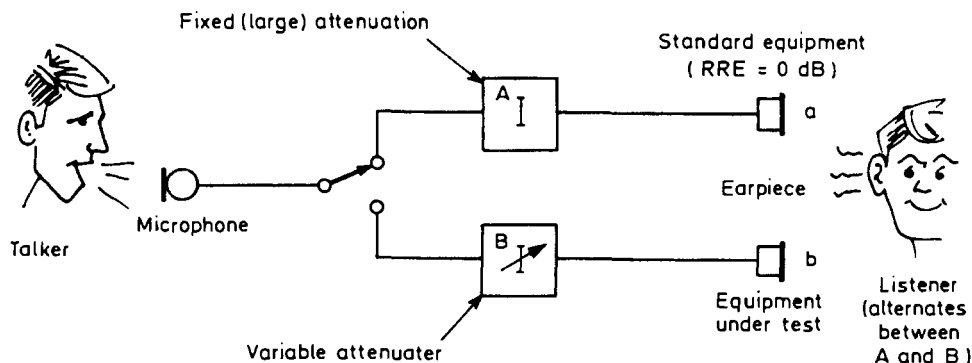


Figure 33.2 Measuring send reference equivalent (SRE)



**Figure 33.3** Measuring receive reference equivalent (RRE)

words, the higher the value of SRE, then the lower the efficiency of the microphone and the greater the loss of signal strength in it. Thus SRE is a measure of the signal loss in the sending device.

*Receive reference equivalents (RREs)* similarly are a measure of the signal loss in the receiving device, and they may be measured by an equivalent NOSFER equipment; but in this case the comparison must be made against a standard high quality receiver, as Figure 33.3 shows.

### 33.3 CONNECTION REFERENCE POINTS AND OVERALL REFERENCE EQUIVALENT

For the design of the network itself, all telephone handsets are assumed to have typical nominal SRE and RRE values. The values chosen are usually quoted for the handset loss relative to an imaginary *reference point* on the customer's line side of the telephone exchange, and they take account not only of typical handset SREs and RREs, but also of the loss encountered on the local line. Typical values are 13 dB SRE and 2–3 dB RRE. In practice, customer lines have SRE and RRE of varying values, depending on the exact handset in use and the length of the line. However, choosing a nominal design value is important as a guide for handset design and also as a benchmark for the required performance of the network itself.

In Figure 33.4 two reference points (a) and (b), have been marked, one at each end of the connection, corresponding to the points relative to which SRE and RRE values apply. These points are marked on the diagram as 0 dBr. The nomenclature dBr stands for *decibels-relative*, and it is used to indicate the received signal loudness at any point, measured relative to the signal strength at the reference point. Hence, at the reference point itself, where the relative strength is equal to the signal strength at the reference point, the value will be zero dBr.

If the network introduces a loss of 3 dB in the transmission direction from A to B, and 3.5 dB in the transmission direction from B to A, then the received signal strengths at points (b) and (a), respectively, will be –3 dBr and –3.5 dBr (i.e. signals 3 dB and

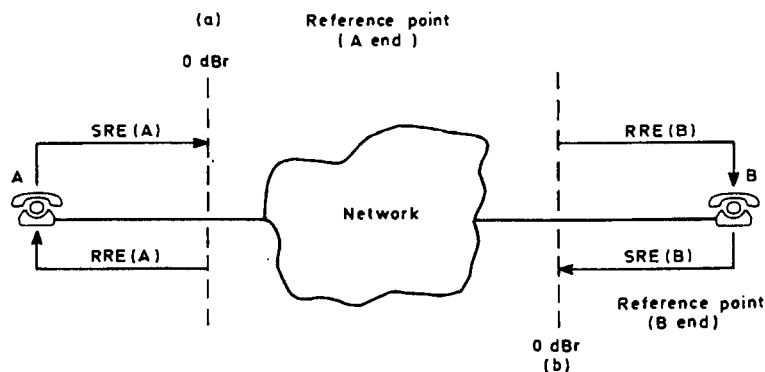


Figure 33.4 Connection 'reference point'

3.5 dB weaker than those sent from the opposite reference points). These values have been marked on the diagram in Figure 33.5.

Now we can meaningfully discuss the end-to-end loss, called the *overall reference equivalent* (or ORE). Surprisingly, the ORE may not be precisely equal to the sum of the SRE, the RRE and the cross-network loss. This is because both the SRE and RRE are only subjective measurements made in isolation on the network's general speech carrying performance, and the ORE is a more stringent measure of the actual end-to-end loss, usually measured using a single frequency tone of 800 Hz. Thus the formula

$$\text{ORE} = \text{SRE} + \text{network loss} + \text{RRE}$$

does not apply.

Initially, network planners thought the discrepancy was small enough to be ignored, thereby greatly easing the task of network design, and CCITT used to recommend that the network loss be adjusted to conform with an overall reference equivalent not exceeding

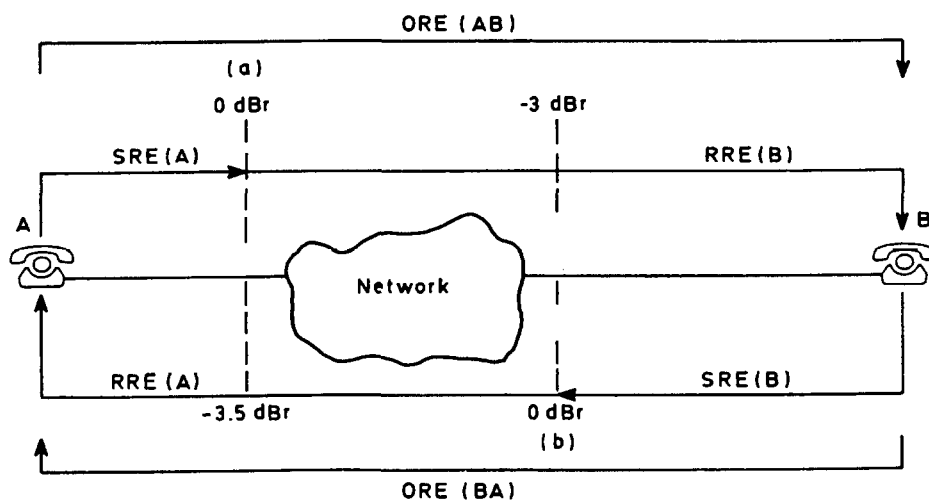


Figure 33.5 Overall reference equivalent

40 dB. However, a number of network operators found difficulty with the system, reporting discrepancies of up to 5 dB, so in 1976 CCITT developed a new measure of performance called the *loudness rating*. Still using the NOSFER equipment as a fundamental reference, and based on similar principles and reference points, the loudness rating method of network design was developed in such a way that loudness ratings (LRs) added to give a reliable and accurate algebraic sum. *Send loudness ratings (SLRs)*, *receive (RLRs)* and *overall loudness ratings (OLR)* were defined in a similar manner to the equivalent RES, but now the formula holds true

$$\text{OLR} = \text{SLR} + \text{network loss} + \text{RLR}$$

The method achieves greater reliability because it uses standard equipment, which although similar in principle to a NOSFER equipment, more closely responds to the frequencies of speech. The equipment is called an *intermediate reference system (IRS)*. First the IRS is calibrated against NOSFER, then the system under test is calibrated. The difference in the two attenuation values (needed for each system individually to give performance equal to NOSFER) is the *loudness rating (LR)*. Values of loudness rating differ from the reference equivalent by varying amounts, up to 3 dB. *Loudness ratings (LR)* are covered in recommendation P.76. The *intermediate reference system (IRS)* used in their measurement is covered in Recommendation P.48, and the methodology for determining LR is given in recommendation P.65.

Following on from the successful introduction of its *loudness rating* method, CCITT in 1980 upgraded its original method of *reference equivalents*, applying a *correction factor* to them to enable the values to be added algebraically. A simple mathematical formula converts *reference equivalents (REs)* into *corrected reference equivalents (CREs)*. Corrected reference equivalents are today's standard method for designing networks against transmission loss. The following relationship applies.

$$\text{OCRE} = \text{SCRE} + \text{network loss} + \text{RCRE}$$

*Corrected reference equivalent (CRE)* values typically differ by a variable amount (up to 3 dB) from corresponding *reference equivalents (REs)*, but are usually a fixed amount (5 dB) greater than corresponding *loudness ratings (LRs)*.

ITU-T nowadays recommends that the maximum *overall loudness rating (OLR)* should not exceed 29 dB and has set an objective optimum value of 10 dB. Such targets should be easily achievable with the excellent performance of modern digital networks. In practice the OLR or OCRE may differ slightly even from the planned value, but this will not matter if sufficient safety margin is allowed during planning, and provided that the differential losses in the two directions of transmission are not too dissimilar. In Figure 33.5, for example, the OREs in the two directions are not quite the same. ITU-T recommends that the difference be limited to 8 dB.

### 33.4 MEASURING NETWORK LOSS

That the loss in decibels incurred by a signal traversing a network will depend on the frequency of that signal is a fact that we learned in Chapter 3. With that in mind, how

can we meaningfully quote a value for network loss? The answer is that it is defined as the signal loss incurred by a standard tone of 1020 Hz frequency (formerly 800 Hz and 1000 Hz tones were also used). It is measured using a standard 1020 Hz tone source which is set to generate a known absolute signal power at a reference point known as the *transmission reference point (TRP)* or 0 dBr point. Reference (dBr) values can then be measured at any other required points in the same transmit path, and compared with the nominal values laid down by the formal network transmission plan.

### 33.5 CORRECTING SIGNAL STRENGTH

The transmission plan (ITU-T recommendations G.101, G.102, G.103, G.111 and G.121) lays out a rigid framework for the expected signal strength at all points along a connection. Any necessary adjustments in strength are achieved by the use of *amplifiers* and variable *attenuators* (also called *pads*). Weak signals are most common and are boosted by inserting amplifiers into the connection. The position of the amplifiers is critical. If insufficient amplification is used the signal strength becomes too weak, becomes inaudible, and is affected by noise interference. Subsequent amplification (if attempted) does not correct the situation, because it amplifies the noise as much as the original signal (in other words a given minimum *signal-to-noise (S/N)* ratio needs to be maintained).

If over-amplification is used it is likely to cause the electrical circuit to saturate (i.e. overload), resulting in signal distortion and electrical instability (manifested as a whistling and loud *feedback* signal). Furthermore, over-amplification can lead to interference in adjacent circuits. Unduly strong signals are easily corrected either by reducing the amplification or by the introduction of attenuators.

There are standard positions in the circuit, where amplifiers or attenuators may be used. The strength necessary is determined by comparing the actual signal strength with the pre-determined value appropriate for that particular reference point. As well as there being a reference point in the customer's local exchange, reference points can also be defined at other points in the network, at least at each exchange. This allows each individual section of an overall transmission link to be designed and *lined-up* accordingly. The sub-sections may then be joined to form a high quality, end-to-end connection. If instead the circuit is lined up once (as a single long section) there could be a counteracting effect between the performance of the various sub-sections which would mask some of the impairment that exists. We have, for example, already noted that it is not acceptable for us to allow the signal to fade to a strength equal to that of the interfering noise before we start amplifying it. Hence the need for a segmented approach to transmission planning. Figure 33.6 illustrates the transmission plan for the North American telephone network while it was analogue. (The illustrated plan has been superseded by the digital transmission plan, but it is still useful in illustrating the use of such a plan). Each link is marked with its maximum permitted loss.

Any new circuit between any two exchanges in the network must be lined up to have an overall loss within the ranges shown on the plan. Thus the maximum end-to-end network loss is assured to be 19.5 dB. The Americans call this end-to-end loss between end offices the *via net loss (VNL)*.

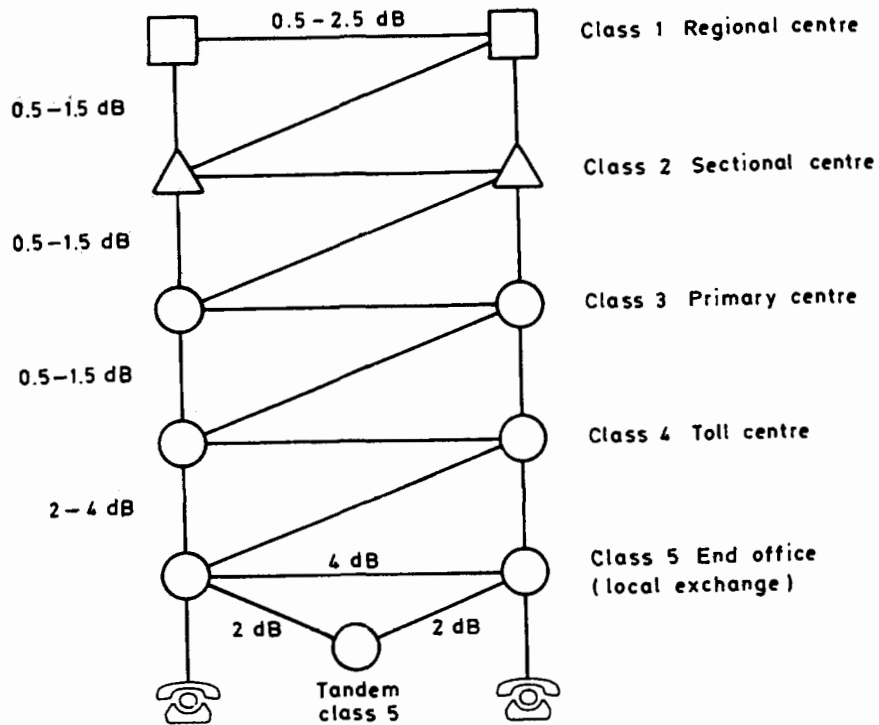
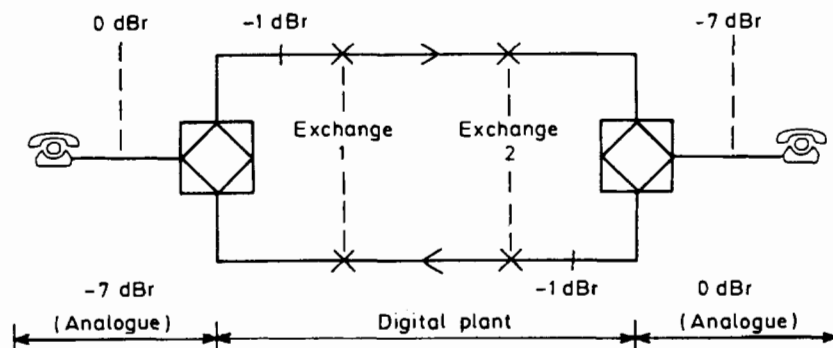


Figure 33.6 The North American telephone network transmission plan

It is best for longer links to be *lined-up* with loss values towards the upper end of the permitted range. This ensures greater electrical stability. Another important feature of the plan is that greater losses are permitted on the links at the bottom of the network hierarchy (i.e. between Class 4 and Class 5 exchanges). This is because the vast majority of lines are in this category, and considerable savings are possible if amplification can be avoided.

Analogue and digital networks both need transmission plans, but they take different forms. In digital networks, the signal strength in decibels is only an important consideration in mixed networks at the points of analogue-to-digital signal conversion, because the regeneration of digital signals by amplification is quite unnecessary in pure digital networks. The digital transmission plan concerns itself instead with such factors as maximum permissible bit error ratios (BERs) and the extent of the quantization limit, which we shall describe more fully later in the chapter. In the meantime, Figure 33.7 shows the signal loss transmission plans for the UK analogue and new UK digital networks.

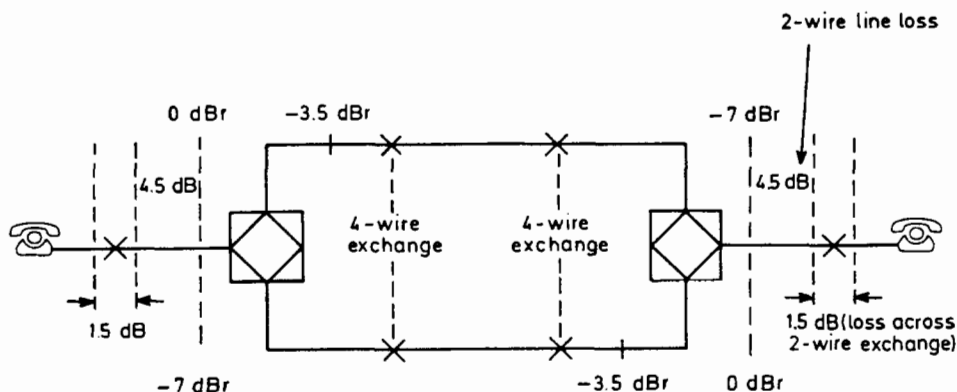
Figure 33.7(a) shows the connection of two normal telephones via two-wire analogue lines to their respective local exchanges, and then over four-wire digital transmission across both exchanges and the interconnecting link. At the originating exchange the speech signal is converted from a two-wire analogue form into a four-wire digital form by an analogue/digital conversion device, adjusted to produce a net loss of 1 dB. The signal is then switched through both exchanges in a digital form without further loss



2 to 4 wire and analogue to digital converter



Exchange switch matrix



2 to 4 wire converter



Exchange switch matrix

**Figure 33.7** (a) The UK digital network transmission plan. (b) The UK analogue network transmission plan.

until it reaches the digital-to-analogue converter at the other end. Unlike the first converter, this one adds a further 6 dB loss, giving an overall network performance of 7 dB loss between reference points. The ORE can then be calculated by adding the relevant SRE and RRE values. This part of the transmission plan is very similar to the analogue plan (Figure 33.7(b)) which it replaces, except that although the total four-wire section loss is 7 dB the recommended signal strengths at the various intermediate reference points in this part of the connection are different. Another difference is that the analogue plan allows a small number of two-wire links to be used in the connection.



### 33.6 THE CONTROL OF SIDETONE

*Sidetone* is the name given to an effect on two-wire systems (e.g. basic analogue telephones) where the speaker hears his own voice in his own earphone while he is speaking (we first introduced it in Chapter 2). Too little sidetone can make speakers think their telephone is dead, but too much leads them to lower their voices. ITU-T recommends *sidetone reference equivalents* of at least 17 dB.

### 33.7 THE PROBLEM OF ECHO

By ensuring that the ITU-T recommended loss of 6 to 7 dB at least is encountered between the two reference points at either end of the four-wire part of a connection (be it digital or analogue), we not only safeguard the circuit against the effects of instability; we also make it unlikely that any signal echo is generated by the two-to-four-wire converter (the so-called *hybrid* converter). These echoes are caused by reflection of the speaker's voice back from the distant receiving end, and Figure 33.8 shows a hybrid converter causing a signal echo of this kind. The problem of echo arises whenever a *hybrid* converter is used, and it results from the non-ideal electrical performance (i.e. the *mismatch*) of the device. No matter whether the four-wire line is digital or analogue, the result is an echo.

To understand the cause of the *echo* we have first to consider the composition of the *hybrid* itself. It consists of a *bridge* of four pairs of wires, one pair corresponds to the two-wire circuit, two more pairs correspond to the *receiver* and *transmit* pairs of the four-wire circuit, and the final pair is a *balance circuit*, the function of which is explained below. The wires are connected to the *bridge* in such a way as to create a separation of the receive and transmit signals on the two-wire circuit from or onto the corresponding receive and transmit pairs of the four-wire circuit. It is easiest to understand the type of hybrid which uses a pair of *cross-coupled* transformers as a bridge. This is shown in Figure 33.9.

Any signal generated at the telephone in Figure 33.9 appears on the transmit pair (but not on the receive pair), and any incoming signal on the receive pair appears on the telephone (but not on the transmit pair). It works as follows.

Electrical signal output from the telephone produces equal magnetic fields around windings W1 and W2. Now the resistance in the balance circuit is set up to be equal to that of the telephone to induce equal fields around windings W3 and W4, but the cross-coupling gives them opposite polarity. The fields of windings W1 and W3 tend to act together and to induce an output in winding W5. Conversely, the fields of windings W2

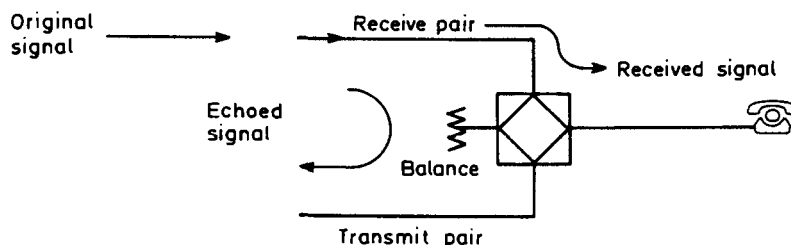


Figure 33.8 Echo caused by a hybrid converter

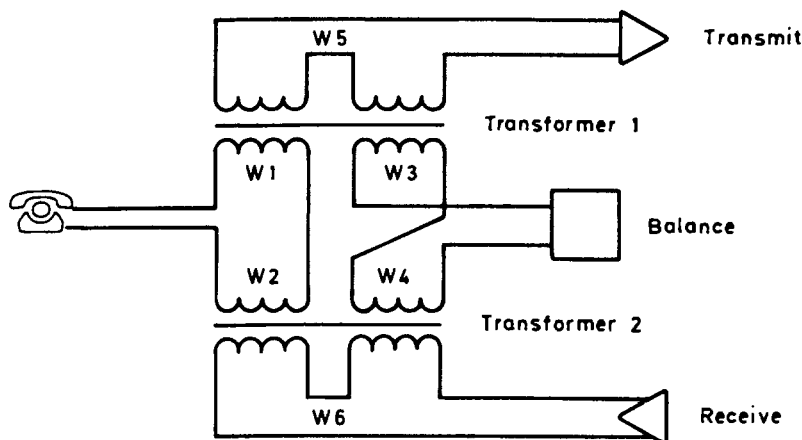


Figure 33.9 A hybrid transformer

and W4 cancel one another (due to the cross-coupling of W4), with the result that no output is induced in winding W6. This gives an output on the transmit pair as desired, but not on the receive pair.

In the receive direction, the field around winding W6 induces fields in W2 and W4. This produces cancelling fields in windings W1 and W3 because of the cross-coupling of windings W3 and W4. An output signal is induced in the two-wire telephone circuit but not in the transmit pair, winding W5.

Unfortunately, the balance resistance of practical networks is rarely matched to the resistance of the telephone. For one thing, this is because the tolerance of workmanship in real networks is much greater. In addition, the use of exchanges in the two-wire part of the circuit (if relevant) means that it is impossible to match the balance resistance to the resistance of all the individual telephones to which the hybrid may be connected. The fields in the windings therefore do not always cancel out entirely as intended. So, for example, when receiving a signal via the receive pair and winding W6, the fields produced in windings W1 and W3 may not quite cancel, and a small electric current may be induced in winding W5. This manifests itself to the speaker as an echo. The strength of the echo is usually denoted in terms of its decibel rating relative to the incoming signal. This is a value called the *balance return loss*, or sometimes, the *echo return loss*. The more efficient the hybrid, the greater the balance return loss (the isolation between receive and transmit circuits).

### 33.8 ECHO CONTROL AND CIRCUIT INSTABILITY

A variety of problems can be caused by echo, the three most important of which are

- electrical circuit *instability* (and possible *feedback*)
- talker distraction
- data corruption

If the returned echo is nearly equal in volume to that of the original signal, and if a rebounding *echo* effect is taking place at both ends of the connection, then the volume of the signal can increase with each successive echo, leading to distortion and circuit overload. This is circuit *instability*, and as we already know the chance of it occurring is considerably reduced by adjusting the four-wire circuit to include more signal attenuation. The total round-loop loss in the UK digital network shown in Figure 33.7 is at least 14 dB, probably inflicting at least 30 dB attenuation on echoes even if the hybrid has only a modest isolating performance. Talker distraction (or data corruption) is another effect of echo, but if the time delay of the echo is not too long then distraction is unlikely, because all talkers hear their own voices anyway while they are talking.

The echo delay time is equal to the time taken for propagation over the transmission link and back again, and is thus related to the length of the line itself. The longer the line, the greater the delay. Should the one-way signal propagation time exceed around 8 ms, giving an echo delay of 15 ms or more, then corrective action is necessary to eliminate the echo which most telephone users find obtrusive. A one-way propagation time of 8 ms is inevitable on all long lines over 2500 km, so that undersea cables of this length and all satellite circuits usually require echo suppression. Further propagation delay can also be caused by certain types of switching and transmission equipment. Indeed a significant problem encountered with digital transmission media is that the time required for intermediate signal regeneration (detection and waveform reshaping) means that the overall speed of propagation is actually reduced to only 0.6 of the speed of light. This means that even quite short digital lines require echo suppression. Two methods of controlling echoes on long distance transmission links are common. These are termed *echo suppression* and *echo cancellation*. Echo cancellation is nowadays most common.

An *echo suppressor* is a device inserted into the transmit path of a circuit. It acts to suppress retransmission of incoming receive path signals by inserting a very large attenuation into the transmit path whenever a signal is detected in the receive path. Figure 33.10 illustrates the principle.

Actually, the device in Figure 33.10 is called a *half echo suppressor* as it acts to suppress only the transmit path. A full echo suppressor would suppress echoes in both transmit and receive paths.

It is normal for a long connection to be equipped with two half-echo suppressors, one at each end, the actual position being specified by the formal transmission plan. Ideally half echo suppressors should be located as near to the source of the echo as possible (i.e. as near to the two-to-four-wire conversion point as possible), and works best when near the ends of the four-wire part of the connection. In practice it may not be economic to provide echo suppressors at all exchanges in the lower levels of the hierarchy, and so they are most commonly provided on the long lines which terminate at regional (class 1 of Figure 33.6) and international exchanges.

Sophisticated inter-exchange signalling is used to control the use of half echo suppressors. Such signalling ensures that on tandem connections of long-haul links intermediate echo suppressors are 'turned off' in the manner illustrated by Figure 33.11. This ensures that a maximum of two half echo suppressors (one at each end of the four-wire part of the connection) are active at any one time.

The amount of suppression (i.e. attenuation) required to reduce the subjective disturbance of echo depends on the number of echo paths available, the echo path propagation

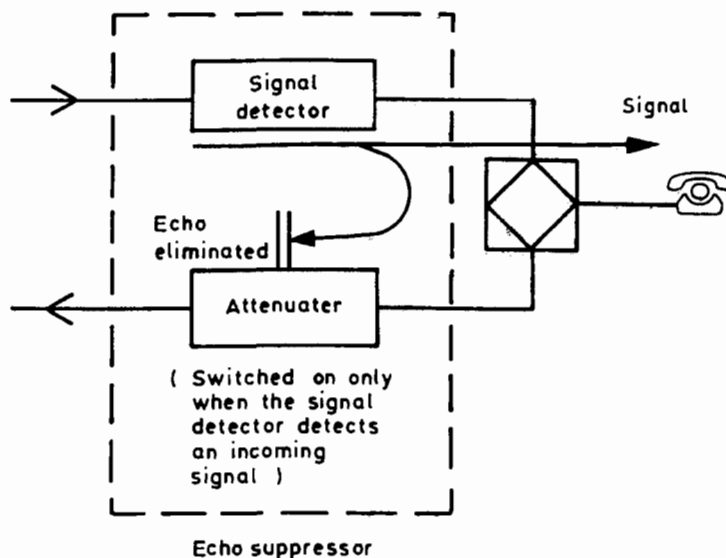


Figure 33.10 The action of an echo suppressor

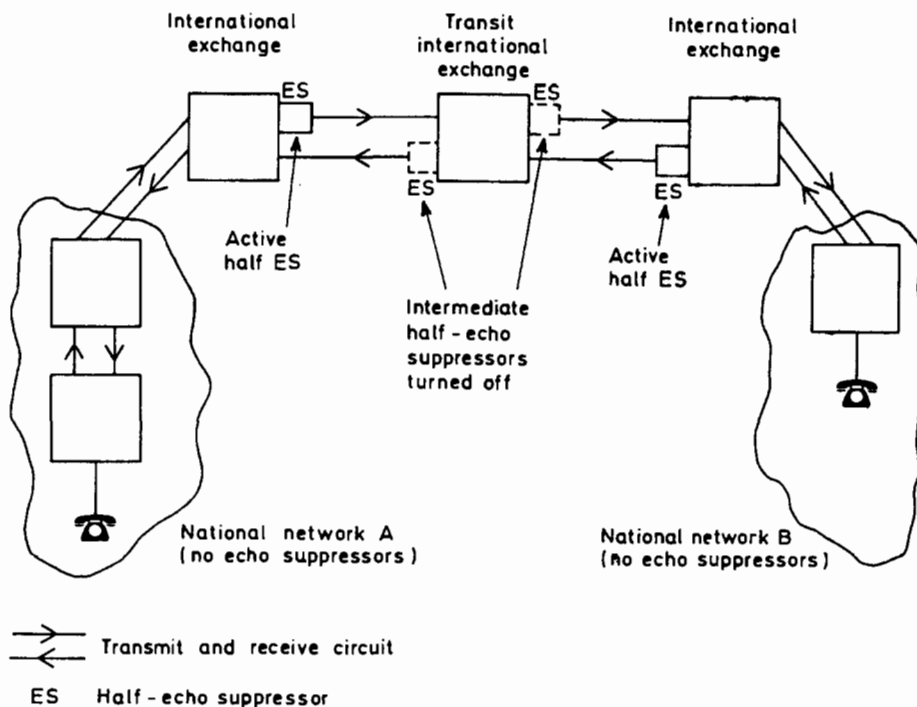


Figure 33.11 Controlling intermediate half-echo suppressors

time, and on the tolerance of the telephone users (or data terminal devices). ITU-T recommends that echo suppression should exceed  $(15 + n)$  dB where  $n$  is the number of links in the connection.

Unfortunately echo suppressors cannot be used on circuits carrying data, because the switching time between attenuation-on and attenuation-off states is too slow and can itself cause loss or corruption of data. Most data modems designed for use on telephone circuits are therefore programmed to send an initiating 2100 Hz tone over the circuit, to disable the echo suppressors.

Another form of echo control device, called an *echo canceller*, can be used on either voice or data circuits, and is the most common form of echo control device used in conjunction with digital circuits. Like an echo suppressor, an echo canceller has a signal detector unit in the receive path. However, instead of using it to switch on a large attenuator, it predicts the likely echo signal (digitally) and literally subtracts this prediction from the returning 'transmit' signal, thereby largely 'cancelling' out the real echo signal. Other signals in the transmit path should be unaffected. Listeners rate the performance of echo cancellers to be better than that of echo suppressors. This, coupled with the fact that they do not corrupt data signals, has made them standard equipment.

### 33.9 SIGNAL (OR 'PROPAGATION') DELAY

An important consideration of the network transmission plan is the overall signal delay or *propagation time*. Excessive delay brings with it not only the risk of echo, but also a number of other impairments. In conversation, for example, long propagation times between talker and listener can lead to confusion. In the course of a conversation, when we have said what we want to say, we expect a fairly prompt response. If we are met with a silent pause, caused by a propagation delay, then we may well be tempted to speak again, to check that we have been heard ('Are you still there?'). Sure as fate, as soon as we do that, the other party appears to start saying its piece, and everyone is talking at once.

On video the effect of signal delay is even more revealing. For example, on live satellite television broadcasts whoever is at the far end always gives the impression of pausing unduly before answering any question.

Nothing can be done to reduce the delay incurred on a physical cable or satellite transmission link. Thus intercontinental telephone conversations via satellite are bound to experience a one-way propagation delay of about 1/4 second, giving a minimum pause between talking and response of 1/2 second. Furthermore, the extremely rapid bit speeds and response times that computer and data circuitry are capable of, can be affected by line lengths of only a few centimetres or metres. Line lengths should therefore be minimized and circuitous routings avoided as far as possible. It is common for maximum physical line lengths to be quoted for data networks. Similarly, in telephone networks, rigid guidelines demand that double or treble satellite hops or other excessive delay paths (i.e. those of 400 ms one-way propagation time or longer) are avoided whenever possible. Excessive delays can be kept in check by appropriate network routing algorithms, as we saw in Chapter 28.

With the emergence of ATM and other ‘fast packet switching’ technology in some voice telephone networks (e.g. corporate telephone networks) the problem of speech propagation delay has been exacerbated by the time required to fill an ATM cell of 48 bytes ( $48/8000 = 0.6$  ms) and by the delay caused by the exit buffer (sometimes up to 30 ms). The exit buffer is intended to eliminate quality problems caused by *cell delay variation (CDV)* as we saw in Chapter 26, but leads to its own problems, not the least of which is the obligation to use echo cancellation devices to control echo.

The use of speech compression techniques also leads to increased signal propagation delays (primarily due to the processing time needed to compress the signal, but maybe additionally due to the increased time now needed to fill a frame or ATM cell (at 32 kbit/s an ATM cell takes 12 ms to fill)). Multiple occurrences of compression and decompression are therefore to be avoided. This requires careful design of the network topology and planning of relevant echo cancellation.

### 33.10 NOISE AND CROSSTALK

*Noise* and *crosstalk* are unwanted signals induced onto the transmission system by adjacent power lines, electrostatic interference, or other telecommunication lines. The only reliable way of controlling them is by careful initial planning and design of the transmission system and the route. One source of noise results from the induction of signals onto telecommunications cables which pass too close to high power lines. Another source of noise is poorly soldered connections or component failures. Both of these are easily avoided. However, by far the most serious source of noise in telecommunications networks and the type most difficult to contain is that caused by electromechanical exchanges themselves. This type of noise results from the electrical noise associated with the electrical pulses needed to activate the exchange, and also from the mechanical ‘chatter’.

Noise is minimized by ensuring that the signal strength is never allowed to fade to a volume level comparable with that of the surrounding noise. Thus a minimum *signal-to-noise* ( $S/N$ ) ratio of signal strengths is maintained throughout the connection. This ensures that the real signal is still perceptible amongst all the background. If the signal becomes too weak in comparison with the noise; unfortunately amplification is then of little value because it boosts signal and noise strength equally. It is difficult to remove noise without affecting the signal itself, but there is some scope for removing noise which lies outside the frequency spectrum of the signal itself by simple filtering. This can slightly improve the *signal-to-noise ratio* ( $S/N$ ).

The noise caused by atmospheric interference and magnetic or electrostatic induction has a random nature and is heard as low level hum, hiss or crackle. When this type of noise is measured using a special type of filter, weighted to reflect the human audible range, then the noise strength can be quoted in *psophometric* units. ITU-T recommends that the strength of *psophometric noise* should not exceed an *electromotive force* (*e.m.f.*) at the receiving end of 1 millivolt (1 mV). Depending on the length of the connection, the transmission network designer may decide the maximum allowable noise disturbance per kilometre of line. Typically this value is around 2–3 picowatts (a very small unit of power) per kilometre. This might be written 2 pW0p/km, where pW stands for picowatts,

0 indicates that the value is measured relative to the 0 dBr reference point, and p denotes psophometric noise.

*Crosstalk* is the interference caused by induction of telecommunications signals from adjacent transmission lines (see Chapter 3). The presence of crosstalk is usually an indication that the circuit has a fault which is causing it to perform outside its design range. Perhaps an amplifier is turned up too much or some other fault has caused an abnormally high volume signal in the adjacent transmission line, or perhaps there is a short circuit or an insulation breakdown caused by damp. The transmission plan normally seeks to minimize crosstalk by ensuring that the transmit and receive signal levels at any point along the connection never differ by more than 20 dB.

### 33.11 SIGNAL DISTORTION

In Chapter 3 we discussed the use of *equalizers* to counteract the *attenuation distortion* of signals which is caused by differential attenuation of the component frequencies. Equalizers are generally located in transmission centres, alongside amplifiers. The equalizers which correct *frequency distortion* (also called *attenuation distortion*) make sure that the attenuation of component frequencies is less than a specified maximum. ITU-T recommendations for speech channels suggest a maximum differential attenuation of 9 dB, using appropriate equalization to ensure that the frequency response of the circuit is within the *mask* illustrated in Figure 33.12.

However, *frequency distortion equalizers* are not the only type of equalizer. Another type is called a *group delay equalizer*. This acts by removing the distorting effects of *group delay*, which is an effect of signal phase distortion of the received signal caused by

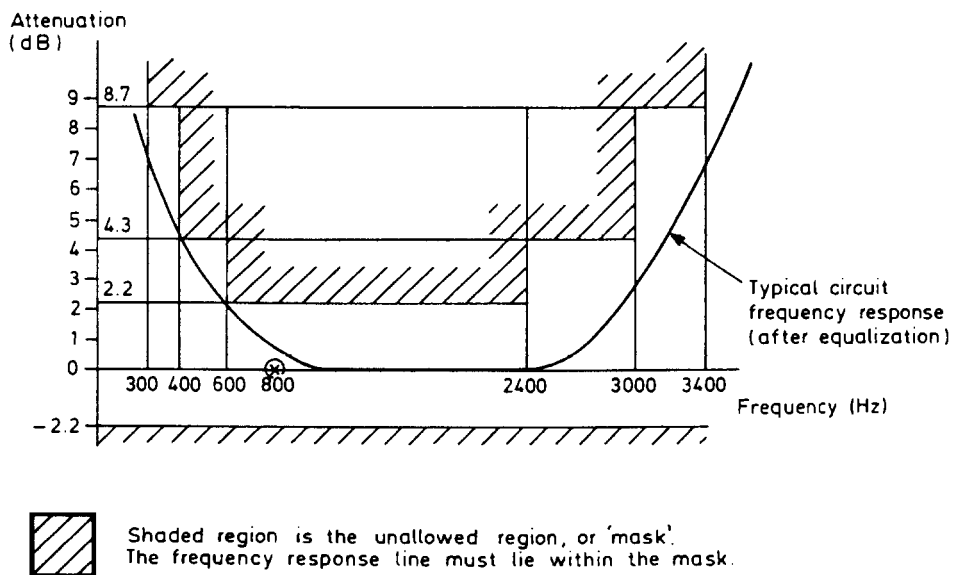


Figure 33.12 Frequency distortion: CCITT's G132 'Mask'

slightly different propagation times of the component frequencies within the signal. Group delay is particularly harmful to data signals passing over *FDM* (*frequency division multiplex*) or radio carriers, and it can be corrected by a device which adds delay to those frequency components which are received first, to even up the delay incurred by all the frequency components.

Both normal frequency attenuation equalizers and group delay equalizers are calibrated when the circuit is initially lined up. A range of different frequencies of known phase and signal strength is sent along the transmission line, and the characteristics of the received signals are carefully measured. The equalizers are then adjusted accordingly and placed in the circuit. A quick re-test should reveal perfect circuit equalization.

### 33.12 TRANSMISSION PLAN FOR DIGITAL AND 'DATA' NETWORKS

Transmission planning of a digital network, as compared with its analogue equivalent, is relatively straightforward primarily because the technique of regeneration practically eliminates the problems of *crosstalk*, *noise*, *attenuation* and *interference*. However, a number of new factors need to be taken into account in a digital transmission plan; they are

- the digital line *bit error rate* (or *bit error ratio*, *BER*)
- the synchronization of the network
- the *quantization distortion*
- (during the period of network transition) the number of analogue-to-digital and digital-to-analogue conversions

Let us take them in order.

*Regeneration* of the digital bit pattern (as described in Chapter 5) must be undertaken so frequently along the length of the line to ensure that *marks* (1s) and *spaces* (0s) are not confused by the receiver. In addition, although less prone than analogue lines, it is still prudent to protect a digital line as far as possible from noise interference and crosstalk to prevent spurious bit errors. The frequency of such error is measured in terms of the error rate (the *bit error rate* or *bit error ratio*, *BER*). Typical acceptance values of BER range from around 1 bit error in  $10^5$  bits ( $BER = 10^{-5}$ ) to one error in  $10^9$  bits, depending on the application. Some types of modern transmission system (including fibre transmission systems and modern digital radio systems) operate at very low bit error rates ( $BER = 10^{-12}$ ). The error rate is usually checked when the digital line system is first established. All channels in the same line system (e.g. each 64 kbit/s channel of a 2 Mbit/s line system) will experience the same *bit error rate* (*BER*).

The synchronization of digital networks ensures that there is no build-up or loss of information in the line system. Without synchronization, if a transmitter sent data faster than the receiver was ready to receive it, the result would be a build-up, and ultimately loss of information. Conversely, if the receiver was expecting data to arrive at a rate faster than the transmitter could send them, then imaginary 'fill-in' data would need to be created for



the missing bits. Synchronization of digital networks is usually carried out in a hierarchical manner, with one exchange designated to house the *master clock*, providing *synchronization* for other exchanges. Figure 33.13 illustrates a typical three-tier synchronization plan. At the top of the hierarchy, a single exchange has a master clock. In the second tier a number of main exchanges receive synchronization (i.e. a data transmitting and receiving rate) from the master clock exchange; additionally they are locked together by two-way synchronization links, keeping them all rigidly in step. Finally, at the bottom of the hierarchy the smaller exchanges merely receive synchronization sources from the higher level exchanges, but do not have two-way synchronization links. Any digital clocks which are located any of the exchanges of the network can be adjusted to run faster or slower to keep track with the synchronization clock. In case the main clock fails it is usual for one of the second tier exchanges to act as a standby master clock, ready to take over should the exchange with the master clock go off-air.

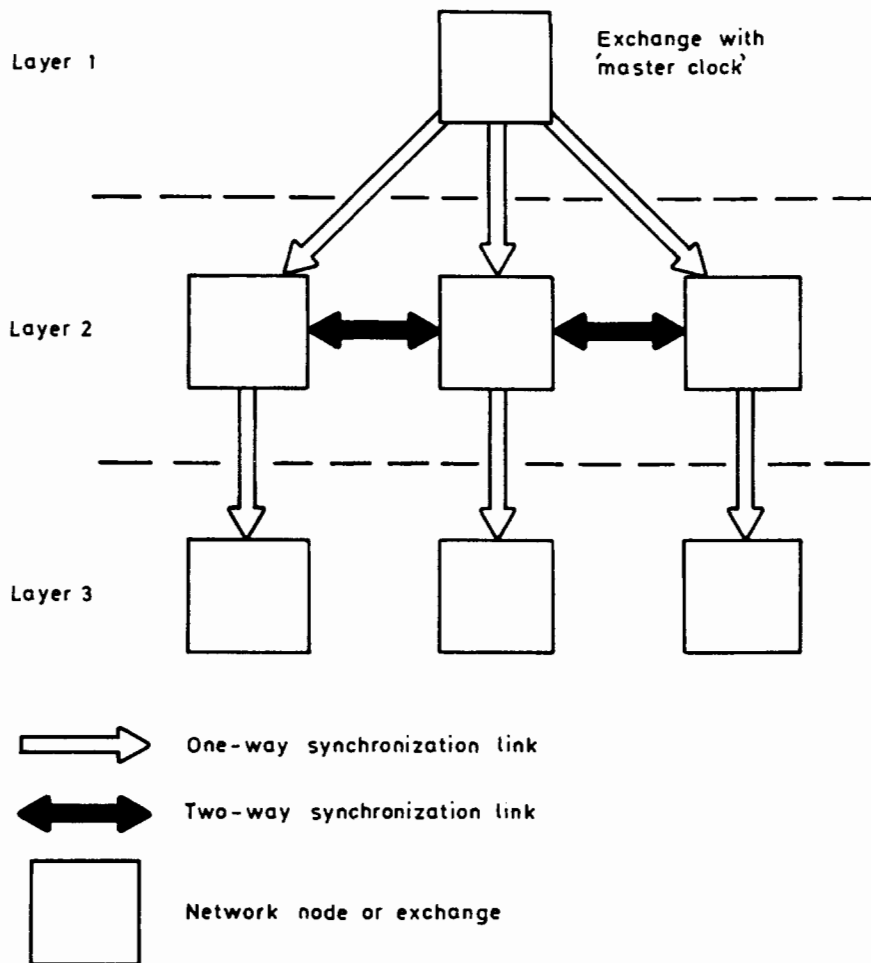


Figure 33.13 Hierarchical synchronization plan

The third important element of a digital transmission plan is the control of *quantization distortion* (also called *quantizing distortion*). As we learned in Chapter 5, quantization distortion arises because the quantum amplitude values which are available to represent the signal always differ slightly from the actual value. The relationship between the representative digital signal and the original is therefore less than perfect. In turn, there will be differences between the final signal and the original, manifested as slight (maybe even imperceptible) distortion. This type of quantization distortion occurs during the initial conversion of any analogue (e.g. speech) signal into its digital equivalent and cannot be recovered on re-conversion. Further quantization distortion can occur should certain other transmission equipments be installed on the digital transmission path. The use of such equipments may be unavoidable, but the transmission plan should clearly set out how much extra distortion is permissible. Either the equipments need to be designed to conform with these limits, or the quality of transmission will suffer. Examples are echo cancellers (discussed earlier in this chapter) and circuit multiplication devices (discussed in Chapter 38).

Quantization distortion is measured in *quantization distortion units (q.d.u.s)*. One q.d.u. is equivalent to a difference in amplitude between the digital signal and the original analogue signal equal to one digital quantum level (for explanation of quantum levels, return to Chapter 5). Thus if the final signal is reproduced with amplitudes varying from that of the original signal by a whole quantum amplitude step, then 1 q.d.u. of distortion has been encountered. A digital transmission plan needs to lay out strict limits on the maximum quantization distortion that can be allowed in any link or part of the network. CCITT 1984 recommendations (still valid today) suggested a maximum end-to-end quantisation distortion of 14 q.d.u.s, allocating 5 q.d.u.s for each national network and a 4 q.d.u. limit for the international connection in between. Typical values of quantization distortion are given in the table of Table 33.1. They reflect the *planning values* given in ITU-T recommendation G.113.

From the table of Table 33.1 it is clear that *quantization distortion* occurs as a result of any form of signal processing, and is particularly sensitive to analogue/digital conversion processes and speech compression using either 32 kbit/s ADPCM or 16 kbit/s CELP algorithms (as we discuss more in Chapter 38). For this reason, connections of interleaved analogue and digital sections, as well as multiple speech compression/decompression stages should be avoided as far as possible. ITU-T states this requirement by recommending the limitation of the number of unintegrated PCM digital sections to 3 or 4 and no more than 7. Of course, as the network evolves and digital transmission becomes more widespread this problem will disappear.

**Table 33.1** Typical quantization distortion values

Digital process	Quantization distortion units
A/D conversion by 8-bit PCM	1
A/D conversion by 7-bit PCM	3
A-law to Mu-law conversion	0.5
Digital attenuator	0.7
ADPCM (see Chapter 20)	3.5

Digital data signals should not be recoded into an analogue form using a normal analogue/digital speech conversion device. On the other hand it is permissible to pass analogue encoded data over PCM lineplant, provided that the number of analogue/digital conversions is limited as stated above.

### 33.13 INTERNATIONAL TRANSMISSION PLAN

As always, ITU-T has some advice for international network transmission planning. This advice is to be found in its G-series recommendations. The principles described therein are exactly as discussed in this chapter, although it is briefly worth explaining ITU-T's concept of a *virtual switching point (VSP)*. This is a *hypothetical reference point*, like our other reference points. In particular it is the point at which a national network is assumed to be connected to the international network. The VSP concept allows ITU-T to lay out recommendations ensuring an acceptable transmission performance of the individual parts (national and international) of a connection. By so doing, the overall acceptability of the end-to-end network is assured without unnecessary strain on the internal plans adopted any of the individual sub-sections. This is clearly important for international interconnection of networks, making it possible for dissimilar networks to work together. Figure 33.14 illustrates the concept, and the national and international network sub-components. Individual recommendations in ITU-T G-series may apply to one or more of the network sub-sections. By convention, the transmit VSP resides at the  $-3.5$  dBr point in analogue networks, and a nominal  $0.5$  dB loss is allocated for each international transmission link. The signal reference value at a receiving international gateway in a direct link connection is thus  $4.0$  dBr. In Figure 33.14, however, the international connection comprises two links, so the receive level should therefore be set up to  $-4.5$  dBr as shown.

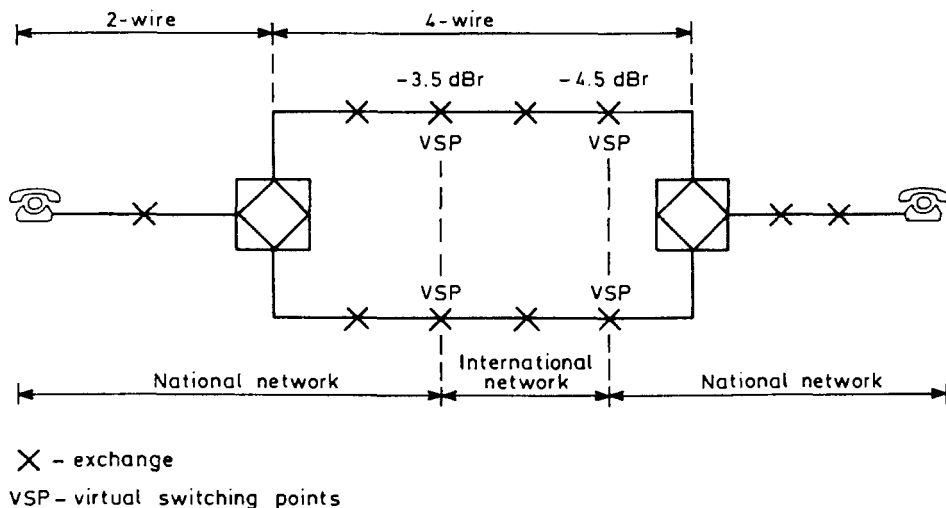


Figure 33.14 National and international networks and VSPs

In more recent ITU-T recommendations on transmission network planning, the VSP is sometimes also marked *ISC* (for *international switching centre*).

### 33.14 PRIVATE NETWORK TRANSMISSION PLANNING

In the foregoing analysis we have been preoccupied with transmission planning for the international public telephone network, but the principles apply in exactly the same way to national and international private telephone networks, and to other forms of data or telecommunications networks as well. The ideal parameter values, however, may vary from network to network.

### 33.15 CIRCUIT AND TRANSMISSION SYSTEM LINE-UP

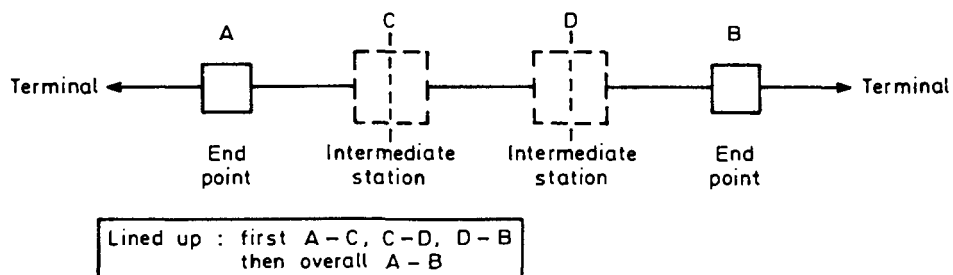
When setting up new circuits, a procedure called *lining-up* ensures that they conform with the requirements of the transmission plans. They are normally lined up first on a section-by-section basis, and then end-to-end. This ensures that the performance is as near to the design as possible. For leased circuits, the end-to-end check can literally be from the terminal in one customer's premises to the other. However, in the case of switched networks, the network operator usually has to be content with *lining-up* and checking only the individual links between adjacent exchanges. Practical constraints (the huge number of possible permutations of inter-exchange connections) prevent verification of each of the end-to-end permutations of the various links and exchanges. Figure 33.15 shows diagrammatically the difference in emphasis of the line up procedure, comparing point-to-point links with switched networks.

The usual line-up procedure is first to install and line-up any higher order transmission systems (e.g. an analogue FDM system, or a high bit rate digital system). This ensures that the overall bandwidth broadly meets requirements. Then each individual circuit is carefully lined up. This ensures correct circuit alignment even when a circuit is actually routed over a concatenation of different higher order transmission systems. Figure 33.16 shows such a complicated circuit routed from exchange A to exchange B over different higher order systems A-C and C-B, and directly connected with wire (*hard-wired*) within the building of exchange C.

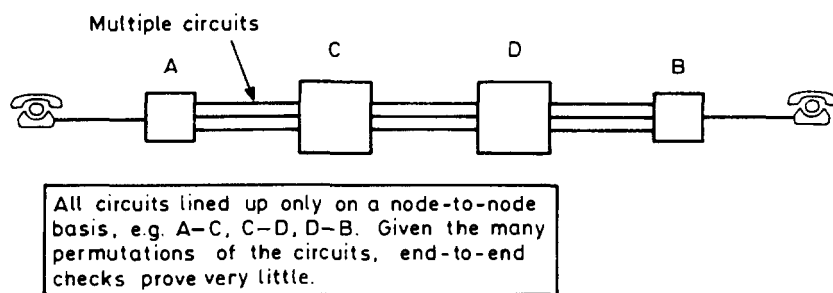
The equipment needed and the operational techniques and procedures for line-up are covered in more detail in Chapter 36 (maintenance).

### 33.16 NETWORK RESOURCE MANAGEMENT

One of the key elements of successful administration of large scale network is the proper planning and management of resources. The remainder of this chapter discusses some of the practical factors pertinent to the operation of cable, radio and satellite systems, and it starts with a few tips on planning circuit provision.



(a) Line up of point-to-point (e.g. leased) circuits



(b) Line up of switched network links

Figure 33.15 Circuit line-up procedures

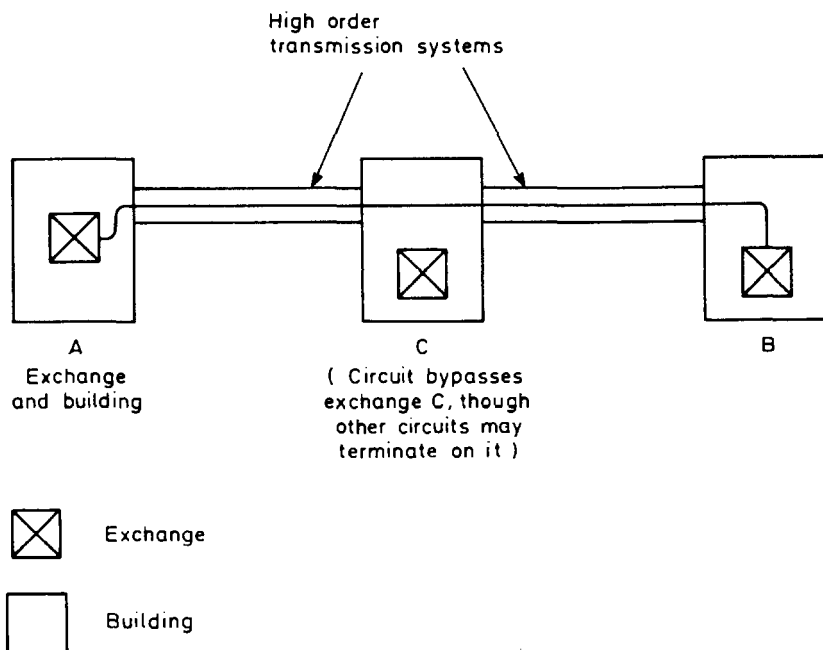


Figure 33.16 The need for end-to-end circuit line-up

### 33.17 CIRCUIT PROVISION PLANNING

*Circuit provision* is the terminology applied to the continuous process of preparing new and re-configuring old circuits in a network. In any expanding network, there is a continuous programme of circuit provision and re-arrangement. Unplanned circuit provision can lead rapidly to such a spaghetti of entangled circuits that any future evolution or re-configuration of the network becomes exceedingly complex and laborious. Pre-allocation of lineplant and of exchange capacity helps to avoid this unnecessary tangle, and allows shortfalls of equipment to be identified earlier, so that extra equipment can be ordered in good time. A computerized data base of circuit information is an extremely useful aid to this process.

In cases where a single transmission network is used to carry a range of different circuit types, a detailed circuit routing policy should be determined. For instance it should be decided whether telephone and telex circuits should be treated separately as far as possible, or mixed together, on common plant. Other circuit types which might need to be covered by the routing policy include circuits between packet-switched exchanges and point-to-point circuits between data terminal equipment.

The disadvantage of mixing all the circuits up with carefree abandon is that it is difficult to re-arrange all the circuits being used for one particular use. For example, it makes it more difficult to shut down a telephone exchange at a particular site and re-direct all the telephone circuits to the replacement exchange at a different site, while leaving a number of telex circuits to a telex exchange behind. The re-direction is much easier if all the telephone circuits are generally *groomed* together. Figure 33.17 illustrates two networks in which different circuit provision policies have been adopted. In each network there is a single site which has a telex, a telephone, a data exchange and three transmission cables to interconnect the site with another. In the first, the *groomed* network (Figure 33.17(a)), each of the transmission cables has been reserved for a particular type of circuit; one is dedicated to telex, one to telephone and one to data. If the telephone exchange is now closed down, then the cable can be extended

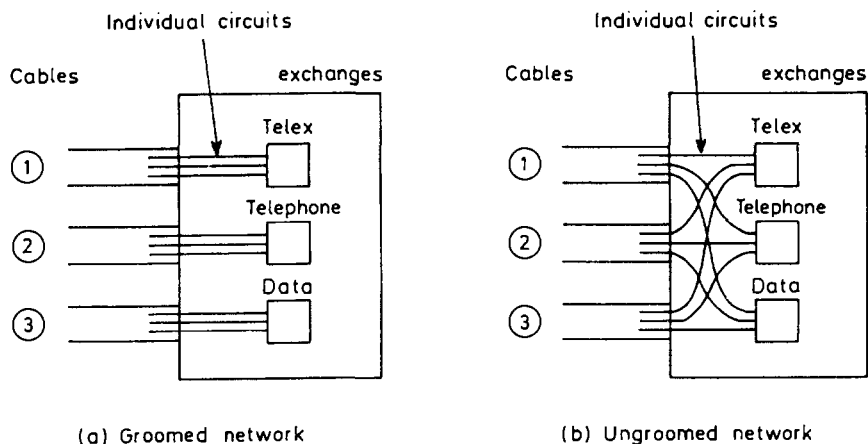


Figure 33.17 Circuit provision policies

very simply to the replacement exchange at a new site. These groomed networks are liable to breakdown; if one of the cables goes out-of-service for any reason, then one or other of the telephone, telex or data exchanges will be isolated. Another disadvantage of a groomed network is the proportionately large amount of spare capacity lying idle within it.

By contrast, Figure 33.17(b) shows a similar configuration of exchanges and cables, but in this case each cable is used to carry all types of circuits. This makes the job of reconfiguring particular circuit types more difficult, but reduces the probability of a complete cut-off of one or more of the telephone, telex and data services. Mixing circuit types also reduces the overhead of spare capacity, because a fourth cable need be purchased only when all three of the cables are full (an improvement on Figure 33.17(a) where a fourth cable is needed as soon as any one of the cables is full).

In practice, it is usual to use an amalgam of both methods of circuit routing, the groomed and the ungroomed. Each of the services passing over more than one cable is diversely routed as far as is practical and economic (to reduce service susceptibility to failure), but chunks of bandwidth within each cable are dedicated to the use of a particular circuit type (to ease reconfiguration). So, for example, one FDM group within an analogue cable could be dedicated to telephone use, whereas another group on the same cable might be dedicated for data circuits. Likewise a whole 2 Mbit/s worth of capacity on a higher bit-rate digital cable could be dedicated to a particular circuit type. Where, however, only a few circuits are required, or where there is not much spare lineplant, it is inevitable that different circuit types have to be more closely mixed together in the *ungroomed* way.

### 33.18 NEW CABLE PLANNING

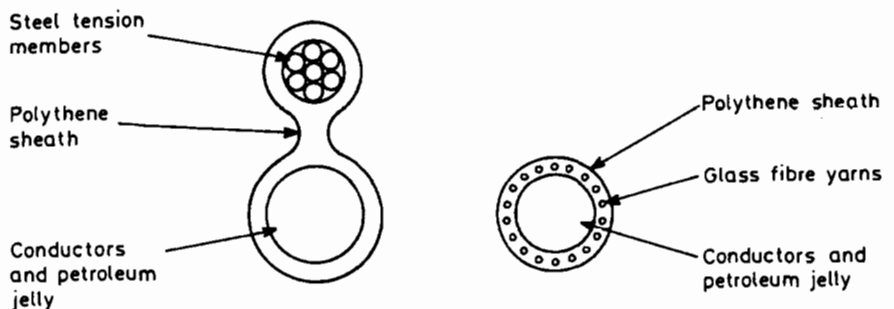
New cable planning has to take into account not only the forecast of circuit numbers required, but also the technology to be used (e.g. digital, analogue, optical fibre, etc.), the route to be taken, and the physical strength that the cable is going to need. Thus customers' individual lines and office networks require different treatment from street backbone wiring of local telephone exchanges. Inter-exchange and international cables need to be planned on an even grander scale.

Routing any type of telecommunications cables too close to high power transmission lines can lead to noise interference; routing of data network cables around an office needs similar care, and accurate measurement is frequently needed to ensure that the maximum permissible line length is not exceeded.

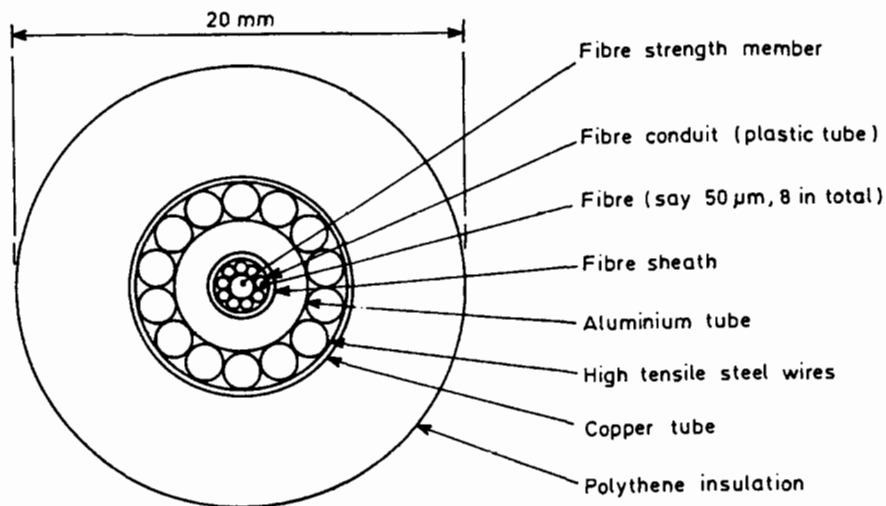
The physical strength and robustness of cables must be attuned to the working conditions. All cables are covered with a sheath to protect them from physical damage and provide electrical insulation, but some are additionally *screened* or *shielded* with metal foil to suppress electrical noise. A light plastic sheath may be adequate where the cable is supported along its entire length in an unexposed cable run (such as an office conduit), but where conditions are more harsh, as in a street conduit, extra protection may be required against mud and damp. For this purpose many cables are often packed with grease, or pressurized to keep water out (water can result in undesirable 'fried egg' circuit noise).

The tension applied to the cable, either during installation (as it is pulled through a conduit) or during use (if it is strung as an overhead wire between telegraph poles), may damage the cable if it does not have sufficient tensile (pulling) strength. Optical fibre cables, for instance, usually comprise not only the fibres and the sheath but also a carbon fibre rope, which provides tensile strength and prevents the optical fibres from being stretched during installation in a conduit.

Cables with inadequate tensile strength gradually sag, gaining resistance, and so become less reliable in performance and more prone to attenuation. An undersea cable, laid on the ocean floor, will not lie completely flat on the sea bed, but may rest on rocks or lie across caverns and ravines. It has therefore to be strong enough to carry its own weight, and it needs protection against damage from fishing trawlers and sharks, etc. As examples of robust cable designs, Figure 33.18 illustrates two designs of overhead cables designed to carry their own weight, and also a possible design for a deep sea optical fibre cable. Note how large a proportion of the structure of each of the cables is there to provide strength and protection.



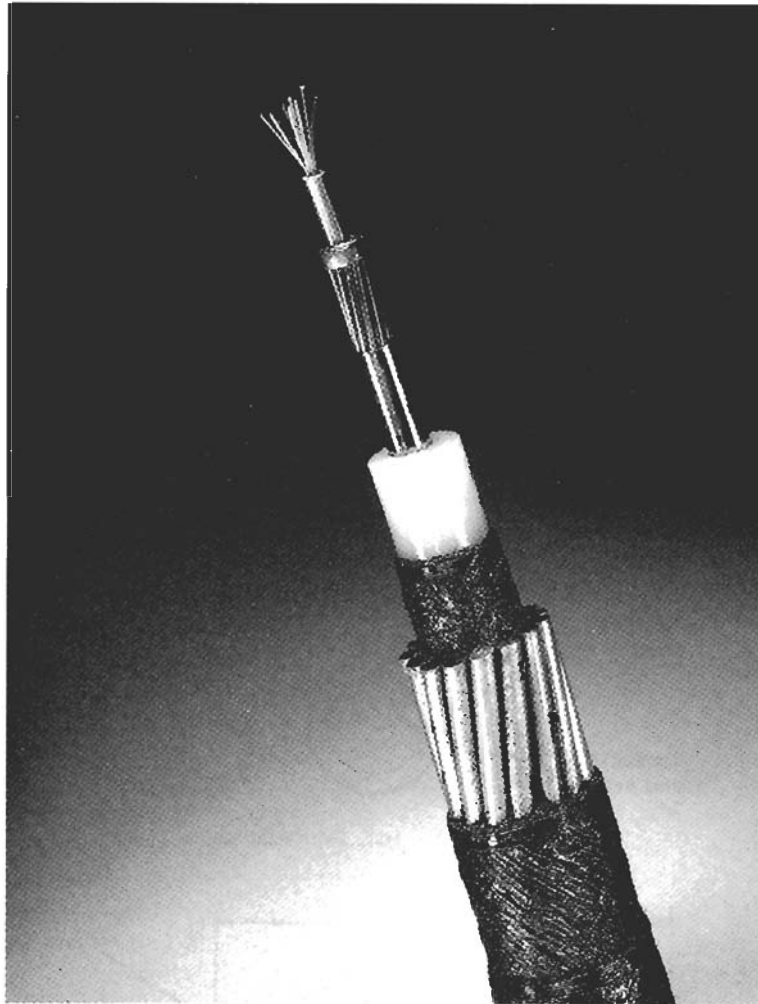
(a) Overhead cable types



(b) Deep sea optical fibre cable

Figure 33.18 Robust cable constructions





**Figure 33.19** A deep sea optical fibre cable. Only four pairs of fibres in this one, but a huge amount of protection and tensile support to stand up to the rigors of deep-sea cable laying, and a life on the very uneven sea bed. (Courtesy of British Telecom)

However, even the undersea cable in Figure 33.18(b) is not the ultimate in robust cable, because where an undersea cable comes onto the 'continental shelf' near land, a further sheathing of *rock armour* is added to give protection against fishing trawlers nets, etc. This can nearly double again the diameter of the cable.

### 33.19 LOCAL LINE PLANNING

The local line is that part of a telephone or other public network which provides connection between a customer's premises and the *local exchange* (central office or

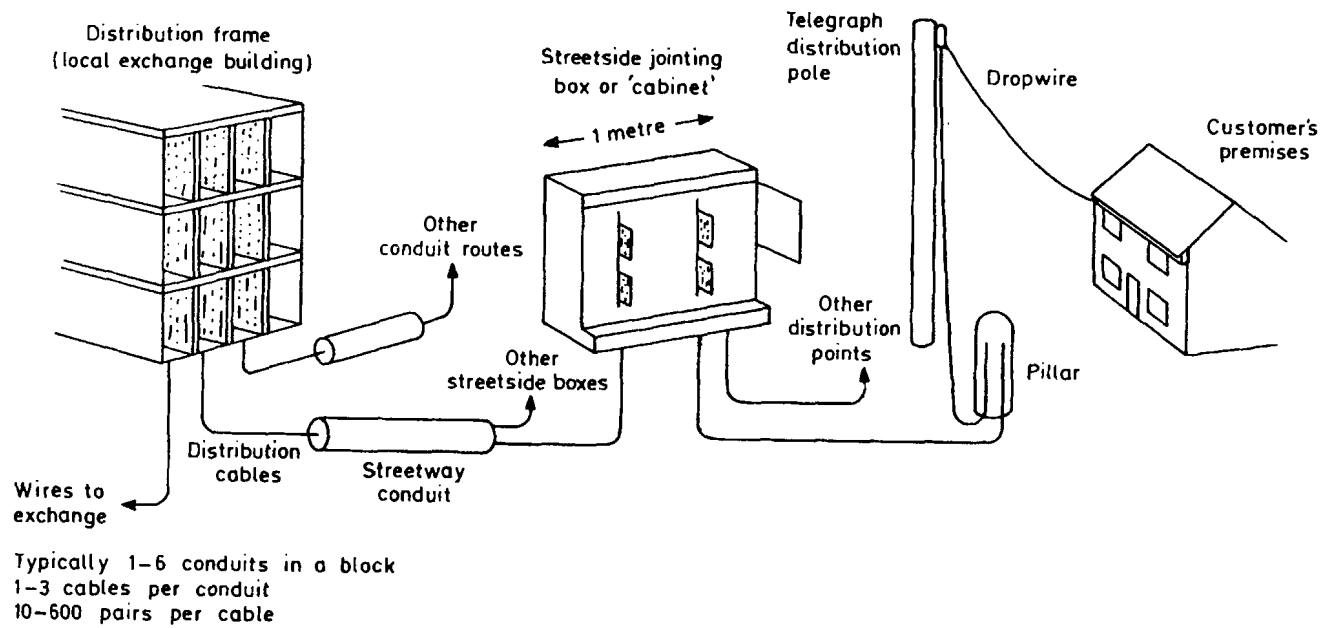


Figure 33.20 Local line network

*end-office*). Traditionally, local lines for telephone networks have been two-wire copper or aluminium *audio circuits*, with a maximum attenuation of about 10 dB and a resistance no greater than  $1000\ \Omega$ . Such limits are practical guides which enable local line planners to design circuits which conform with the transmission plan and the signalling needs, while allowing scope for long lines to be run out from the exchange to relatively remote customers, several kilometres away. Planning the local line network involves careful positioning of distribution cables, and appropriate choice of their capacity (in terms of the number of pairs of conductors). Usually, a number of high capacity *distribution cables* (10–600 pair cables) cables lead out in a star fashion from a *distribution frame* in the exchange (where they are connected to exchange equipment) to a hierarchy of cross-connection points (street-side jointing boxes). At each of these boxes a multiplicity of smaller capacity cables (or single *dropwire* pairs) can be run on to further cross-connection points (streetside *cabinets* or *pillars*) or direct into customer premises. Figure 33.20 illustrates the typical layout of a local line network.

The planner must ensure that the conductors in each of the cables are of sufficient gauge (i.e. diameter) to enable the resistance and attenuation limits to be met for each of the circuits to all the customer premises. Usually the gauge of distribution cables is the greatest, because these cover the longest part of each access line and therefore need the lowest resistance per kilometre. The shortest section, the dropwire, need only be fairly narrow gauge. The difference in resistance between narrow and wide gauge wire is not important over this short distance. For four-wire communication paths, including high grade point-to-point data circuits, two pairs of wire are needed, and where a particularly high bandwidth or digital bit speed is required, coaxial or some other special type of cable may be called for.

Like trunk networks, local line networks are being revolutionized by the introduction of optical fibre. For customers with very high digital bit speed requirements, the fibre can be run directly into digital multiplexors on their premises, but even for those with more modest requirements, optical fibres offer a cheap and reliable alternative to metal distribution cables. They provide for high capacity, low attenuation links from the exchange to a digital multiplexor in the streetway box, and in many of cities fibre-optic cable has already been deployed in this way to alleviate previous conduit congestion resulting from the large number and size of individual metal pair conductors.

The advent of optical fibres usage in local exchange networks has brought with it new topologies of local line networks (see also Chapter 17). Recently developed by British Telecom in the United Kingdom, the *telephony passive optical network (TPON)* and the *broadband passive optical network (BPON)* are examples of the new techniques. These networks use either a single fibre ring (Figure 33.24(a)) or a multiplexed star arrangement (Figure 33.24(b)) to backhaul the customer lines into the local exchange.

TPON is a network designed for basic telephony including normal leased lines, speech and data networks and ISDN. BPON goes one stage further, delivering video and cable television to the home. Who knows, maybe radio will evolve to the point where a fixed local line is unnecessary. These topics are covered in earlier chapters.

The local exchange is also evolving and miniaturizing. Already some small exchanges can actually be located within streetside cabinets and manholes.



**Figure 33.21** Streetside cabinet and jointing box. An engineer checking the wiring on the cross-connection frame inside a streetside cabinet. This provides for flexible connection of the many cables and individual pairs out to customers' premises and back to the exchange. In the foreground, the cables running away into the street conduits via a jointing box. (*Courtesy of BT Archives*)



**Figure 33.22** Streetside joint repair. Checking a joint during the installation of telephone service in 1935. Phoning the exchange. (Courtesy of BT Archives)



**Figure 33.23** Metropolitan transmission system. In a sunlit manhole in New Jersey, USA, an installer checks out the wiring in a maintenance case – part of the Bell System's metropolitan transmission system. (Courtesy of AT&T)

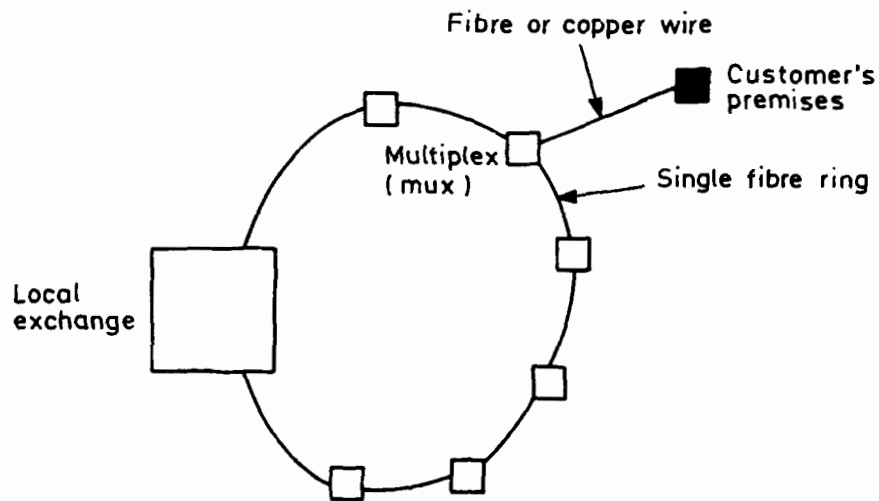
## 33.20 TRUNK AND INTERNATIONAL LINE PLANNING

Basically, the planning of trunk and international lines are no different from local ones. It is simply that the capacity needed of the individual cables and the transmission plan constraints on its performance mean that much higher quality equipment has to be used.

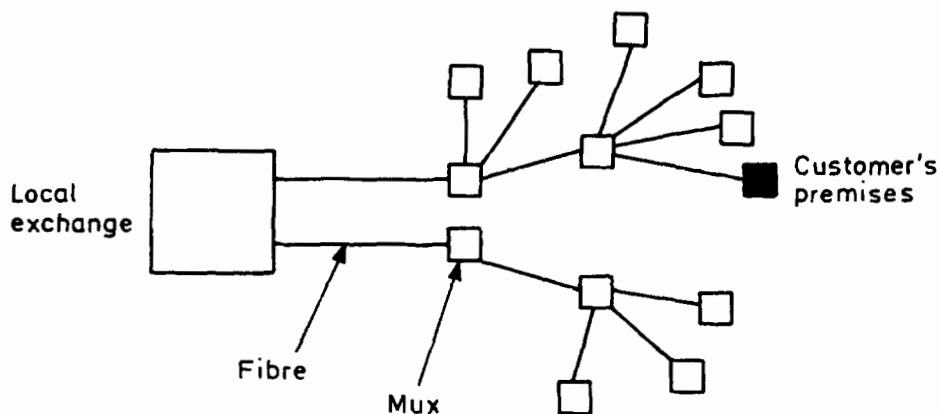
## 33.21 RADIO TRANSMISSION SYSTEMS

The most important factors in management and operation of radio systems are the allocation of the radio bandwidth and the location of radio antenna sites.

The International Telecommunications Union, Radio Communication Sector (ITU-R, formerly known as *CCIR, international radiocommunication consultative committee*; *IFRB, international frequency registration board*; and *WARC, world administrative radio council*) administers the specification of technical radio standards, the



(a) Fibre ring



(b) Multiplexed star

**Figure 33.24** Optical fibre in the local network

registration of radio users and the allocation of radio bandwidth to ensure that interference of signals is minimized. Bands are thus allocated on an international basis for such uses as

- longhaul *point-to-point* (PTP) radio (up to 100 km)
- shorthaul *point-to-point* radio (up to 20 km)
- *point-to-multipoint* (PMP) radio

- satellite *uplinks* and *downlinks*
- broadcast services (including television, public broadcast radio, etc.)
- mobile telephone networks
- military usage

Once a given radio spectrum has been earmarked for a given purpose, technical standards bodies can set about laying down radio spectrum *masks* that ensure that individual radio transmission and receiving devices use the spectrum efficiently, do not disturb neighbouring systems and, as far as necessary, are compatible with one another.

The actual use of radio spectrum is regulated on a national basis. Typically the transmission or receipt of a given radio frequency is regulated by national government bodies, to whom individual users must apply for registration. In the UK, for example, this is carried out by the *Radiocommunication agency* of the *DTI* (*Department of Trade and Industry*), in Germany the task is conducted by *BAPT* (*Bundesamt für Post und Telekommunikation*), and in the United States it is the responsibility of *FCC* (*Federal Communications Commission*).

*Corporate* or *public* telecommunications network operators who wish to employ radio systems within their network require to apply for a licence for the use of particular radio spectrum. A given radio channel or radio band allocation is usually made either

- on a regional basis for *point-to-multipoint* or mobile networks, or
- on a *point-to-point* basis for simple transmission links

In this way the radio regulator is able to ensure the most efficient usage of the spectrum and proper control over possible interference between systems.

Finally, it is down to the radio network operator to plan and install his system. The most critical part of this work is the radio network and *path planning*. Only with good path planning can an operator achieve freedom from interference and radio *fading*, and a high system *availability*. We cannot hope to be fair to this complex subject in this book, but by way of introduction to the subject we present a few ideas to illustrate the problems.

Radio path planning begins with determining the likely range of a given radio system (the length of a point-to-point link) or the radius of a given *cell* coverage area. ITU-R recommendation PN.525-2 provides a simple formula for calculating the *free space transmission loss*,  $L_{bf}$  (i.e. the attenuation of the radio signal)

#### FREE SPACE TRANSMISSION LOSS

$$L_{bf} \text{ (in dB)} = 20 \log (4\pi d / \lambda)$$

where  $\lambda$  = wavelength in km,  $d$  = path length in km

alternatively

$$L_{bf} \text{ (in dB)} = 32.4 + 20 \log f + 20 \log d$$

$f$  = frequency in MHz,  $d$  = path length in km



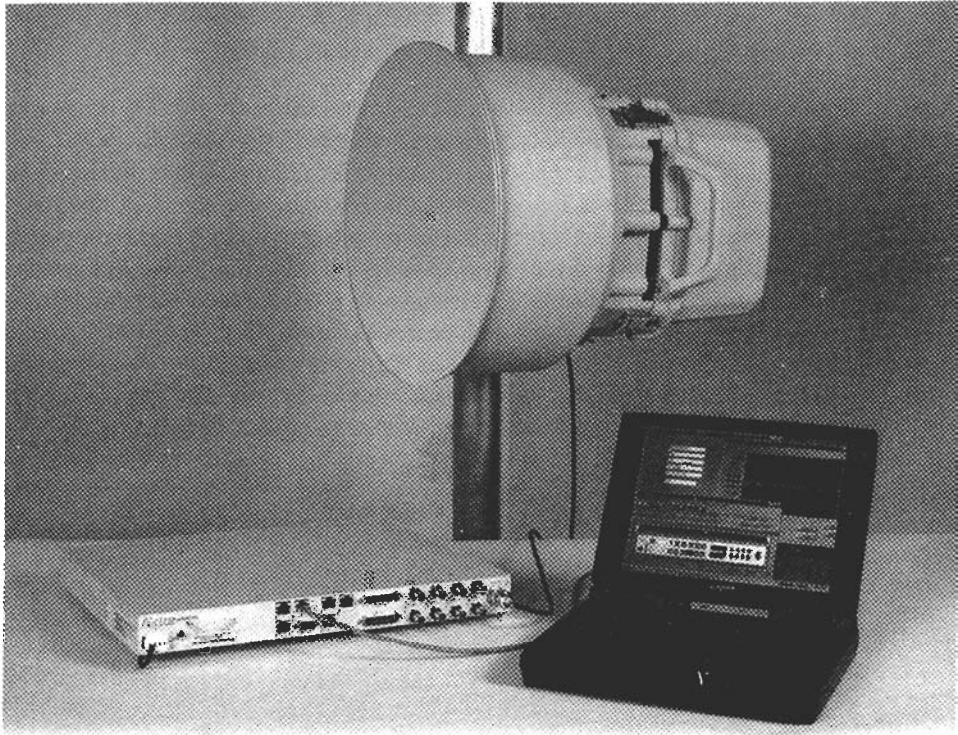
The *free space transmission loss* gives the likely signal degeneration during good weather. Thus by knowing the output transmitter power and the *gain* of both the receive and transmit antennae, the expected *receive signal level (RSL)* can easily be calculated.

#### EXPECTED RECEIVE SIGNAL LEVEL (RSL)

$$\text{RSL (in dB)} = P_{\text{TX}} + G_{\text{TA}} + G_{\text{RA}} - 32.4 - 20 \log f - 20 \log d$$

where  $P_{\text{TX}}$  = transmit power in dB,  $G_{\text{TA}}$  and  $G_{\text{RA}}$  are the transmit and receive antenna gains in dB<sub>i</sub> (i.e. gain relative to an *isotropic antenna*, the relative focussing of the signal in the desired direction),  $f$  = frequency in MHz and  $d$  = distance in km

Taking a typical *shorthaul* radio system operating in the 38 GHz band, the typical range of the system is 6.5 km. The typical transmit power is 15 dBm (32 mWatt), and with 30 cm antennae, the transmitter and receiver antennae will have gains of around 39 dB<sub>i</sub>. From our formula, the RSL in good weather will thus be around -47 dBm. Given that the *receiver sensitivity* is likely to be able reliably to detect a much weaker



**Figure 33.25** 38 GHz microwave radio terminal with 30 cm antenna. (Courtesy of Netro Corporation)

signal, of strength around only  $-80$  dBm, this gives a margin for *fading* of the radio signal. In our example the *fade margin* is 33 dB.

Attenuation, propagation loss or fading of radio signals may occur for any number of different reasons

- frequency *absorption* due to atmospheric gases
- loss due to path obstruction or partial obstruction
- fading due to *multipath*, whereby different reflections of the original signal interfere harmfully with one another
- *fading* due to rain, fog, snow or other weather effects
- attenuation due to sand and dust storms
- magnetic and electrical effects of the earth's surface and geography

The different causes of fading have varying effects on different radio frequency bands, so that, for example, rain fading is the most significant cause of attenuation for frequencies above about 1 GHz, but can be ignored at lower frequencies. This significantly affects the reliability and *availability* of *line of sight* (LOS) microwave radio systems. For lower frequencies, with much greater range, the effects of path refraction within the earth's atmosphere are significant. We end this short discussion about the considerations of radio path planning with two more simple formulae drawn from ITU-R recommendations.

*Rain fading* as affecting *line of sight* microwave radio systems operating above 1 GHz can be estimated according to ITU-R recommendations 838 and PN.837-1. These state that the attenuation,  $\gamma_R$  (in dB/km), caused by rain is as follows.

ATTENUATION DUE TO RAIN  
(ITU-R recommendation 838 for frequencies above 1 GHz)

$$\gamma_R \text{ (in dB/km)} = kR^\alpha$$

where  $k$  and  $\alpha$  are values which depend on local weather and climatic conditions and  $R$  is the rainfall rate in mm/hour. Values of these constants over the entire earth's surface are plotted in recommendation ITU-R PN.837-1

Returning to our example of the 38 GHz link of 6.5 km, the values of the constants for 38 GHz in world zone E (American midwest and northern Europe) are  $k = 0.279$  and  $\alpha = 0.943$  (ITU-R recommendation 838). Meanwhile, recommendation PN.837-1 tells us that for less than 0.01% of each average year the rainfall exceeds 22 mm/hour. We can thus expect the attenuation per kilometre to exceed  $\gamma_R = kR^\alpha = 5.1$  dB/km for (at most) 0.1% of the time. Therefore (for 99.99% of the time) we can expect the attenuation due to rain over a 6.5 km transmission link not to exceed  $5.1 \times 6.5 = 33$  dB. The 33 dB is exactly the *fade margin* we calculated earlier, you might say. This is no coincidence; the design criterion which limited the system range to 6.5 km was the requirement to achieve 99.99% availability of the link.

For radios operating at lower frequencies, the effects of multipath refraction (i.e. interference caused by the atmosphere) can be more acute. Here ITU-R recommendation PN.530-5 offers a simple formula for calculating the probability of a given radio fade, as follows.

**RADIO FADING DUE TO MULTIPATH EFFECTS  
CAUSED BY REFRACTION**

$$\text{fade probability (\% of year incidence)} \approx 10^{-\Delta G/10} K d^{3.6} f^{0.89} 10^{-A/10}$$

where

$$\Delta G = 11.05 - 2.8 \log d \text{ (for latitude } 50^\circ, \text{ Europe)}$$

$A$  = depth of fade in dB

$f$  = frequency in GHz

$d$  = distance in km

$$K \approx 10^{-6.5} P_L^{1.5}$$

$P_L$  = factor depending upon atmospheric refractivity, for Europe  
and North America the value lies between 5 and 20

### 33.22 SATELLITE TRANSMISSION MANAGEMENT

Satellite transmission is only a specialized form of microwave radio, and so like other types of radio transmission it requires proper frequency management and careful siting of both ground based antennae (called *earth stations*) and orbit vehicles (called *satellites*).

The satellites used for telecommunications are mostly *geostationary*. By this we mean that they orbit around the earth's equator at a speed equal to the spin speed of the earth itself. They therefore appear to be geographically stationary above a point on the earth's equator, which allows them to be tracked easily by earth station antennas. Any given earth station will either be able to 'see' a satellite for 24 hours a day, every day (dependent on the satellite's longitudinal position), or will never be able to see it. The world's leading organization for the development and exploitation of satellite communications is called *INTELSAT* (the International Telecommunications Satellite consortium). Intelsat is a jointly owned consortium, set up in 1964 by the world's leading public telecommunications operators to provide a means for worldwide transmission. It is not the only satellite body. Others include *INMARSAT* (the International Maritime Satellite organization), *EUTELSAT* (the European equivalent of INTELSAT) and a number of privately owned and run companies which operate satellites for a range of national telecommunications purposes (e.g. direct broadcast by satellite, telephone circuit access

to remote areas, or most recently the *universal mobile telephone service*, UMTS, a satellite-based system intended to extend mobile telephone networks to global coverage).

The geographical coverage area, or *footprint* of a satellite in orbit (i.e. the area on the ground in which earth stations are able to work to it) is affected by its geographical position above the equator and by the design and power of its antennae. The position above the equator is usually stated as the geographic longitude. The antennae may either be directed at a small area of the earth's surface in a *spot beam*, or broadcast to all the hemispherical zone on the earth's surface that the satellite can see. The latter type of antenna needs to be much larger and it demands radio signals of considerably greater power. The INTELSAT system comprises a number of satellites, positioned in clusters over the middle of the world's three great oceans, the Atlantic the Pacific and the Indian Ocean, and give the coverage areas shown in Figure 33.26. Between them the satellites allow most countries to have telecommunication with almost any other country in the world.

European countries have access to the Atlantic and the Indian Ocean INTELSAT satellites, and countries in North America can view the Pacific and Atlantic satellites (from west and east coasts, respectively).

A major benefit offered by satellite transmission over cable is the fact that a single earth station in one country and a single satellite in space may be used to provide transmission links to not just one, but a number of other countries. The signal is sent up by the transmitting earth station to the satellite (the *uplink*). The satellite responds by amplifying and re-transmitting a downpath radio signal over the whole coverage area. Any earth station within the coverage area can pick up the signal, and select from it the information intended for that particular destination. This type of operation is called *multiple access* working.

The responding and transmitting radio equipment on board the satellite is usually referred to as a *transponder*. A number of transponders usually share the same dish antennae on board the satellite. Transponders are either analogue or digital (i.e. frequency division multiplexed or time division multiplexed), and may be used either for point-to-point use or for multiple access as explained above. Each distant earth station receives all the radio carrier signals but selects only the appropriate carrier frequencies for demodulation.

The major disadvantage of satellite transmission is the one-way propagation delay of around 270 ms. This is the minimum time that it takes for the radio signal to travel the distance up to the satellite and back down to earth, and in a two-way telephone conversation it causes a half-second silence between asking a question and hearing the first part of the answer. The allocation of radio bandwidth and planning for new satellites in the Intelsat system are carried out annually at the *Global Traffic Meeting (GTM)*, held in Washington, DC, United States. At the meeting, all constituent operators and users of the Intelsat system declare and specify their forward five year bandwidth requirements. The main types of radio carriers used by Intelsat are called FDMA, TDMA, SCPC, IDR and IBS. Explanation of these terms is as follows.

FDMA (frequency division multiple access) uses analogue *frequency division modulated* carriers on up and downlinks. *Broadcast* of the *downlink* signal is the key to *multiple access* operation, allowing the same signal to be picked up at a number of distant end earth stations, but only the relevant receive channels are picked out by each destination earth station. Thus a contiguous *uplink* bandwidth must be allocated to each earth station to meet only its outgoing transmission requirements.

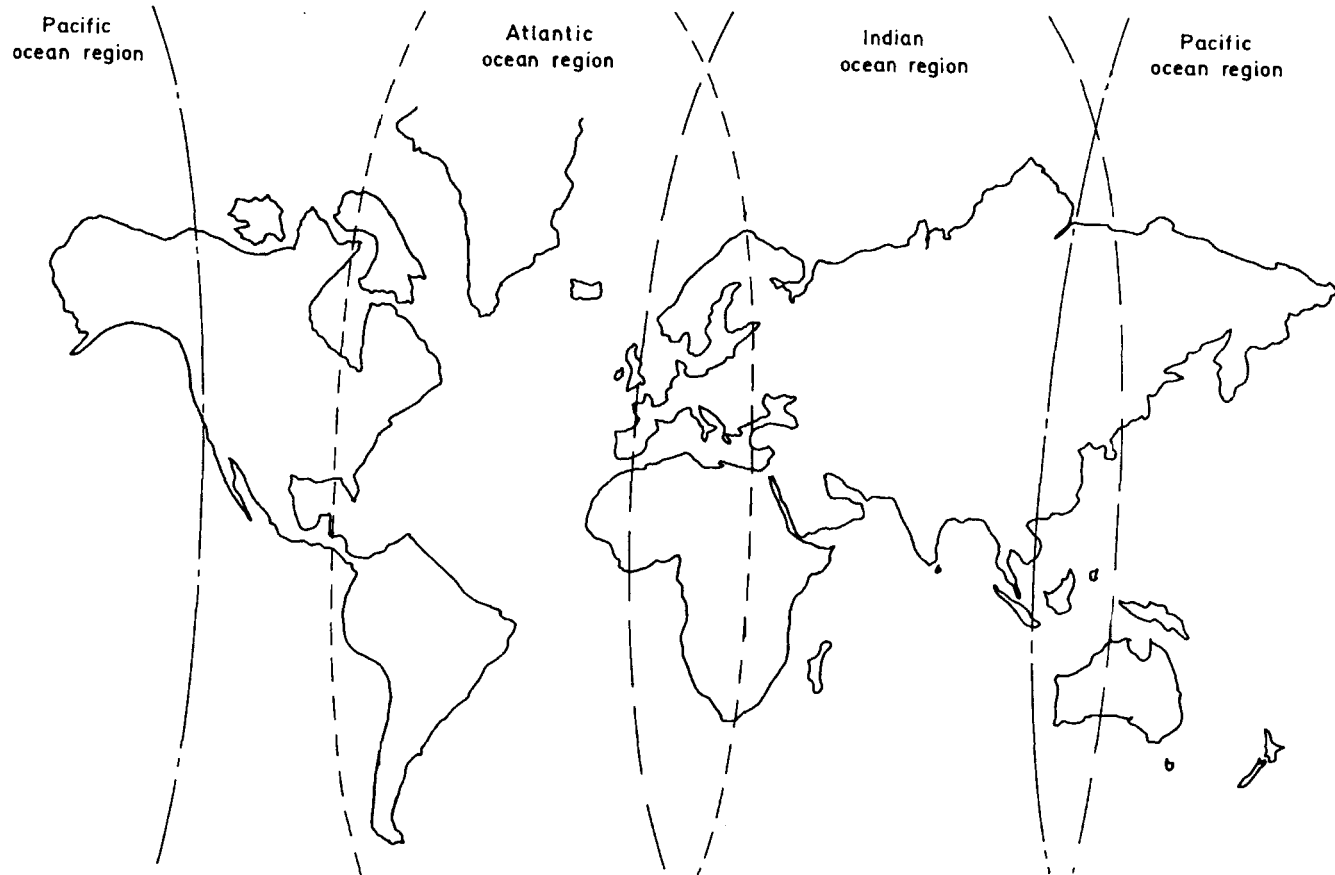


Figure 33.26 Approximate INTELSAT coverage regions

TDMA (time division multiple access) uses digitally modulated, *time division multiplexed* signals. As with FDMA, when used in the *multiple access* mode, one transmit carrier is used by each earth station, and a number of received carriers need to be scanned for the appropriate return circuits. The difference is that the transmit *carrier* is in reality a pre-allocated 'burst' of a high bit speed signal. The bursts from the different earth stations are interleaved in a time-shared fashion like ordinary TDM.

SCPC (single channel per carrier): as the name suggests, single channel per carrier systems only support one channel on each carrier. They are generally used for point-to-point circuits when only a small number of overall circuits are required between end points.

IDR (intermediate data rate) these are the latest type of carriers, of varying digital bandwidths from 1.5 Mbit/s and 2 Mbit/s upwards.

IBS (international business system) IBS transponders support point-to-point business applications between small dishes and are ideal for bandwidths up to 2 Mbit/s.



---

## *Quality of Service (QOS) and Network Performance (NP)*

---

The maintenance of good quality for any product or service (i.e. its 'fitness for purpose' and its price) is of supreme importance to the consumer and therefore requires utmost management attention. However, although it is easy enough to test a tangible product to destruction, measurement of the quality of a service is more difficult. In telecommunication the customer is left with nothing more tangible than his or her own perception of how well the communication went. On a datalink the errors that have had to be corrected automatically are barely appreciated, while in conversation unobtrusive bursts of line noise may go unnoticed. This is not to say that loss of data throughput on a datalink caused by continual error correction is of no concern to the customer, nor does it mean that noises which disturb conversation are acceptable. What we really mean is that in measuring service quality, due regard must be paid to anything of importance that the customer perceives and remembers. Concentration on setting and meeting quality targets should be paramount in planning and administration. Insufficient attention to them is the road to customer dissatisfaction and loss of business. This chapter reviews some aspects of communications quality and the practices of management, with examples of the commoner quality parameters and control measures used by the world's major operators.

### **34.1 FRAMEWORK FOR PERFORMANCE MANAGEMENT**

Good management of whatever industry demands the use of simple, structured and effective monitoring tools and control procedures to maintain the efficiency of the internal business processes and the quality of the output. When all is running smoothly a minimum of effect should be required. However, to be able quickly to correct defects or cope with abnormal circumstances, measurable means of reporting *faults* or *exceptions* and rapid procedures for identifying actionable tasks are required. The framework for doing so needs to be structured and comprehensive. In Table 34.1 a number of simple management 'dimensions' together with possible management performance tools/parameters within each dimension are presented. As you will note the dimensions cover a range of areas, some requiring more tangible monitoring measures and control procedures than others.



**Table 34.1** Framework for performance management

Dimension	Measures/controls
Formal plan and delivery cycle (organizational framework)	Formal <i>presentation</i> and <i>agreement</i> of company strategy, policy, guidance and plans. Regular review. Framework for business 'value' assessment and assurance.
User support (fit for purpose)	User comprehension of services (by questionnaire). Swift elimination of problems (full complaint log).
Supplier performance (quality of supply)	Maintaining professional attitude with suppliers. Monitoring and demanding swift lead times. Meeting targets for repair performance.
Quality performance (effectiveness)	Meeting user needs (number of complaints). Number of lost messages (e.g. percentage of data lost) Percentage of 'down-time' during normal business hours. Measure of delay (e.g. propagation delay, or backlog of orders/messages). Measure of congestion (e.g. percentage calls lost – callers frustrated). Repair performance (e.g. percentage not repaired in target time). Connection quality (e.g. percentage customers satisfied). 'Cost of poor quality' (i.e. that correcting avoidable faults).
Financial performance (efficiency)	Return on capital investment. Shareholders financial return. Cost actually spent can be compared with competitors or alternative suppliers. Internal chargeout rate for telecom services (e.g. 'pounds per bit-mile').
Technical performance (efficiency)	Compatibility of networks can be adjudged by measuring the money spent on interworking or upgrading and comparing it with total system value.
Resource management (efficiency)	Asset inventory and management (needs to be accurate and up to date). Cost per task. Adequate resourcing to meet target service lead times. Head count and man-hours should be maintained on target.
Evolution of networks (responsiveness)	Network upgrades should be properly planned and meet their budget. There should be steps to give 'benefits' assurance. The response time of new software and computer networks should be according to plan. The networks should be adaptable.

### 34.2 QUALITY: A MARKETING VIEW

To look at a network from the customer's viewpoint we need to understand his reasons for using telecommunications services. It is a mistake to generalize about all customers as 'wanting to convey information over long distances', as that takes no account of the service in use, or of each customer's individual purpose or *application*. It is, after all, a

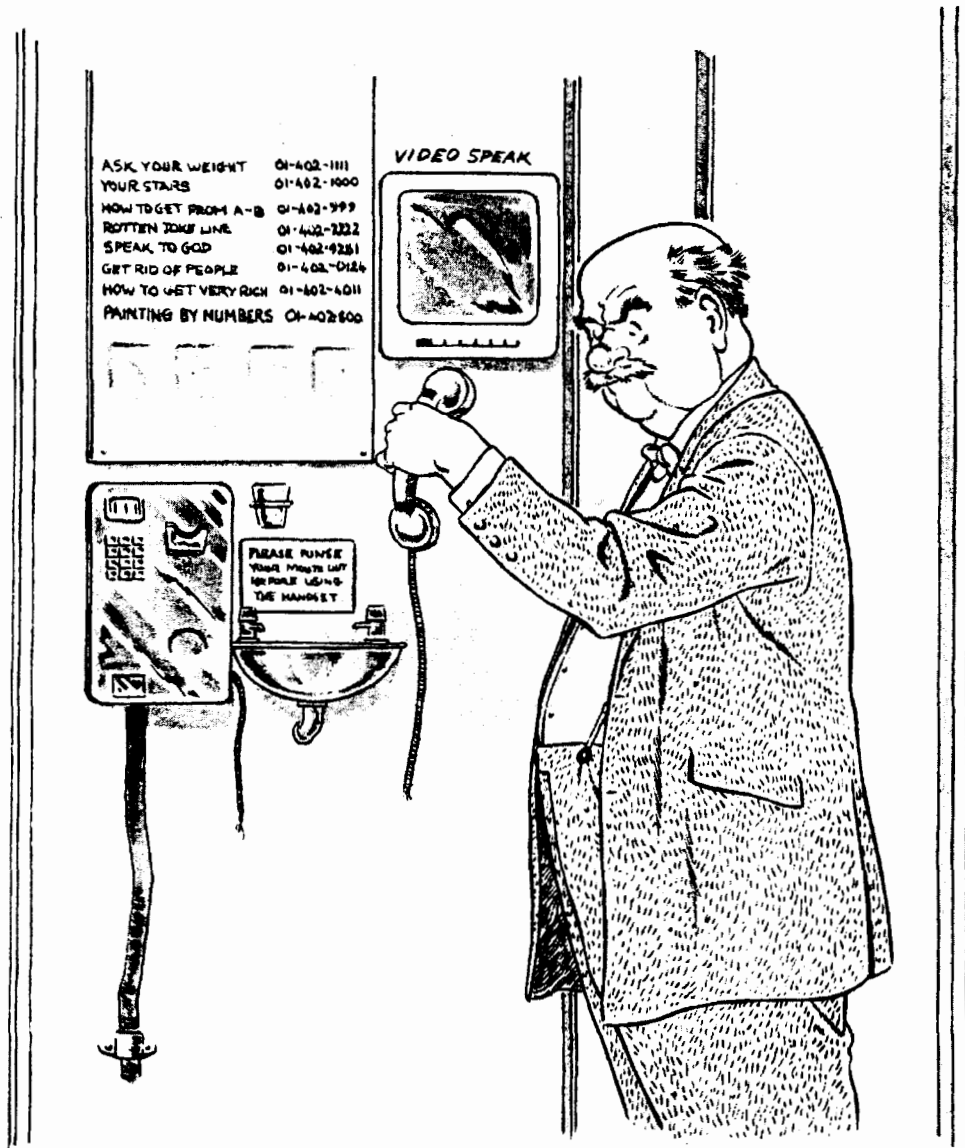


Figure 34.1 A customer's perception of quality. Get the basics right before worrying with the frills. Drawing by Patrick Wright. (Courtesy of M. P. Clark)

fundamental of marketing that products should be attuned to the markets and customers they are intended to serve. Thus the motivations and interests of someone who is out for the evening in calling home from a payphone, are quite different from the needs of a large multi-national company wishing to convey vast volumes of computer data around the world on sophisticated and dedicated networks at any hour of the day or night.

It may be that the same network infrastructure can accommodate the service needs of a number of different markets or customer 'clubs' but it must do so without compromise. One example of a network capable of supporting a number of services is the public telephone network, and it must run at optimum quality not only for human conversation, but also for facsimile machine interconnection. A more advanced example of integrated services networks is the *integrated services digital network (ISDN)* discussed in Chapter 10 and the *B-ISDN* of Chapter 25.

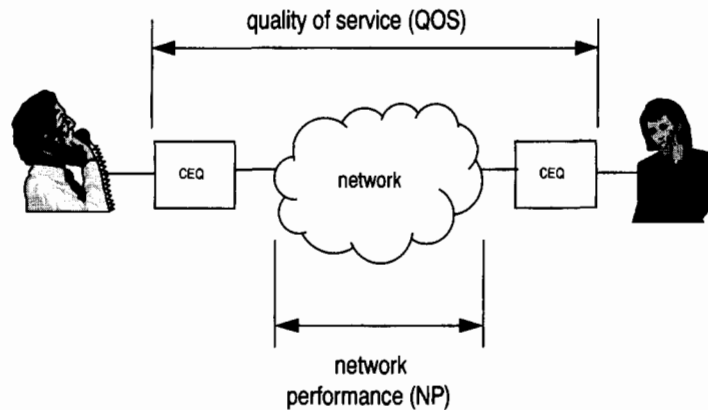
Customers' communication needs can often be met in a number of alternative fashions, ranging from travelling in person, using the postal service, the telephone, telex, packet switching, and high speed data telecommunications services. Pitching a given service to meet a given need is the job of telecommunications marketing specialists who determine its relevant qualities, expressed as a number of measurable parameters. An example of such a 'quality parameter', used commonly by telephone network operators is the *percentage of calls completed*. Not only is this a fair measure of network congestion, it is also a good predictor of the level of customer frustration.

### 34.3 QUALITY OF SERVICE (QOS) AND NETWORK PERFORMANCE (NP)

You may have realized in our example above that the percentage of answered calls is dependent not only on network congestion, but also on the availability of somebody at the destination end to answer the telephone. Thus the *quality of service* enjoyed by the customer depends on a number of factors in addition to the performance of the network. ITU-T has also drawn this distinction, and its general recommendations on the quality of telecommunications services recognize two separate categories of performance measurement

- quality of service (QOS)
- network performance (NP)

*Quality of service* measurements help a telecommunications service or network provider to gauge customers' perceptions of the service. *Network performance* parameters, on the other hand, are direct measurements of the performance of the network, in isolation from customer and terminating equipment effects. Thus *quality of service* encompasses a wider domain than *network performance*, so that it is possible to have a case of poor overall *quality of service* even though the *network performance* may be excellent. The relationship of *quality of service* to *network performance* is shown in Figure 34.2.



**Figure 34.2** The relationship between quality of service (QOS) and network performance (NP)

Similar parameters may be used to measure both *quality of service* and *network performance* (e.g. propagation delay, *bit error ratio* (*BER*), % congestion, etc.). Normally the measured *quality of service* is lower than the measured *network performance*. The difference is due to the performance degradations caused by the user's own end equipment (Figure 34.2).

The measured quality of service will differ greatly from the measured network performance values where a connection is composed of a number of connections, traversing several different networks and end user equipments. Although it is of utmost importance to the end user, the problem with *quality of service* as a performance measure is that it is difficult to measure, and would need to be measured for each individual customer separately. This is the reason for the development of the concept of *network performance*. *Network performance* can be more easily measured within the network, and provides for meaningful performance targets for the technicians and network managers operating the network.

In short, QOS parameters are user-oriented, and provide useful input to the network design process, but they are not necessarily easy to translate into meaningful technical specifications for the network. Network performance parameters, on the other hand, provide a directly usable technical basis for network designers and operational managers, but may not be meaningful to end users. Sometimes, the same parameters are appropriate for both QOS and NP, but this is not always the case.

The distinction between *quality of service* (*QOS*) and *network performance* (*NP*) is somewhat artificial, but it has become necessary as the result of recent regulatory changes in some countries, where the *public telephone operator* (*PTO*) is allowed to provide service only up to a socket in the customer's premises, and end-user equipment (i.e. *terminal*) manufacturers slug it out in the *customer premises equipment* (*CPE*) market. Government regulation should keep PTO network performance up to the mark, but overall service may be of poor quality if it is let down by faulty or badly designed customer apparatus.

Parameters should be chosen to reflect high quality end-to-end service and should be expressed quantitatively as far as possible. These QOS parameters should then be correlated with one or a number of directly related network performance (NP)

parameters, each NP parameter reflecting the performance of a component part of the network, and therefore contributing to the end-to-end quality. In this way quality of service problems can be traced quickly to their root cause.

In the example of Figure 34.3, QOS criteria have laid it down that the signal loudness received by the listener should not be less than 14 dB below the original signal transmitted. The end-to-end *QOS parameter* is therefore *signal loudness* and is measured in dB. The network between the two people in conversation comprises two telephone exchanges and three linking circuits, which are designed to individual *network performance* loudness losses of 1 dB. The overall network performance is therefore  $3 + 1 + 3 + 1 + 3 = 11$  dB, and when the losses in the telephone handset, the end-to-end (QOS) loudness is 13 dB, i.e. giving a 1 dB margin within the maximum allowed loss of 14 dB.

Considering only the *network performance (NP)* between the two telephone sockets of Figure 34.3, an acceptable end-to-end QOS is delivered even if one of the links or exchanges in the network degrades in performance by up to 1 dB (provided that all the other parts continue to work at their designed losses). In other words, a network performance of 12 dB is sufficient to deliver an acceptable end-to-end QOS of 14 dB loss. Likewise, if either of the customer telephone sets degrades in performance to the extent of adding up to 1 dB, then the overall QOS of 14 dB is still maintained, provided other components remain stable.

In normal operation it is difficult to monitor the end-to-end QOS performance of all connections, particularly when a large number of permutated switched connections are possible, and the concept of network performance parameters is invaluable. Each element of the network may be monitored in isolation and maintained within a pre-determined network performance range. Thus in the example of Figure 34.3 we might aim to keep the individual exchange losses below 1.5 dB, and the individual link losses below 3.5 dB. Although theoretically this might mean that the overall network performance could include a loss of  $3.5 + 1.5 + 3.5 + 1.5 + 3.5 = 13.5$  dB, which is 1.5 dB outside the maximum allowed network performance of 12 dB, it is highly unlikely that all the links and exchanges will simultaneously be performing in the most adverse manner. As a safeguard against such an adverse occurrence, small occasional samples of the end-to-end quality of some representative connections will need to be taken.

Sometimes the relationship between the end-to-end quality of service parameter measurements and the network performance parameters may not be as direct as in the example just discussed. In that example we were able to use similar parameters, all measured using the same scientific units (in our case dB, or decibels). However, in the QOS parameter *percentage of calls reaching answer*, there is a much more complex relationship with measurable network performance parameters, depending not only on the state of network congestion, but also on the availability of alternative routes, the incidence of faults, and other abnormalities such as *misrouting* or *misinsertion*. Some of the contributions may have negligible effect under normal conditions, so that reliable approximations may be used. For example, if the percentage of calls being answered at a given destination starts to deteriorate then the network operator's first action is to look for any network links that are congested. If any are found, then these are more than likely the cause of the congestion, and extra resources should be provided accordingly. If no individual link or exchange congestion is immediately apparent then the operator will have to look for less common causes and faults.

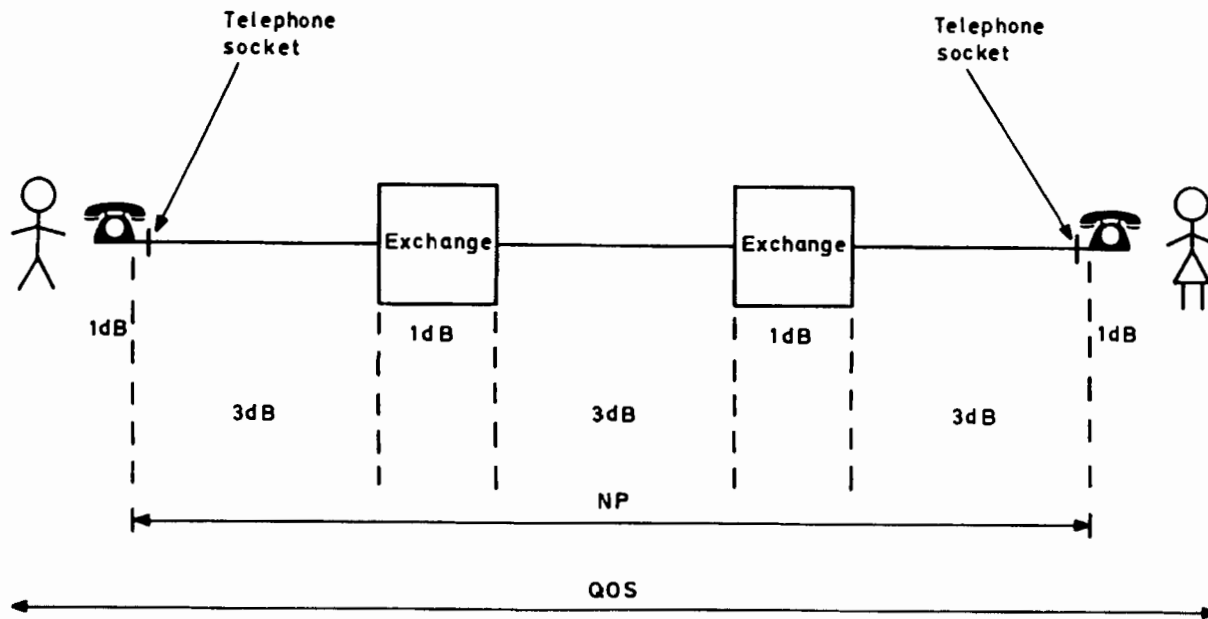


Figure 34.3 Relating QOS and NP parameters

There are no hard-and-fast rules as to which parameters should be used to measure quality of service and network performance, although ITU-T has created a generic model and a set of *generic parameters*.

### 34.4 QUALITY OF SERVICE PARAMETERS

Quality of service parameters focus on the user's likely perception of the service, rather than on the technical cause of specific degradations of service. For this reason, different quality of service parameters should be tailored for each different type of network application.

QOS parameters are not comprehensively defined by ITU-T standards. For certain types of services (e.g. leaselines, telephone service, etc.), ITU-T defines QOS parameters, and these same parameters should be used to measure end-to-end quality of such services, independent of what type of transport network is used. Thus, for example, the quality of service of a digital leaseline could be measured in terms of the *bit error ratio* (*BER*), the *jitter*, the accuracy of the *bitrate*, the propagation delay, the *availability* of the connection (the percentage of a one-month or three month period for which the circuit was not out-of-service) and the accuracy of the invoice.

In addition, the network operator has an interest to conceive network performance parameters which would help him to meet the user's expected QOS level. In this respect, he has a little help, as ITU-T has defined a number of standard network performance parameters for some types of networks. The network operator's task thus becomes relating these to the specific expectations of individual services' QOS, and setting appropriate network performance targets.

### 34.5 GENERIC NETWORK PERFORMANCE PARAMETERS

For connection-oriented networks (such as the telephone network, packet and frame-switched data networks and ATM networks), ITU-T defines a set of three different types of NP parameters, and recommends that these should be used as measures of three different functional aspects of the connection (access (connection set-up), user information transfer and disengagement). Table 34.2 shows the nine different types of

**Table 34.2** The generic primary performance parameters

Performance criterion function	Speed	Accuracy	Dependability
Access	Access speed	Access accuracy	Access dependability
User information transfer	Information transfer speed	Information transfer accuracy	Information transfer dependability
Disengagement	Disengagement speed	Disengagement accuracy	Disengagement dependability

**Table 34.3** Example primary performance parameters relevant to telephone, ISDN, data and ATM networks and services

Performance criterion function	Speed	Accuracy	Dependability
Access	Connection set-up delay	Incorrect set-up probability (misrouted connection ratio)	Probability of set-up denial (connection set-up denial ratio)
User information transfer	Successful transfer rate Propagation delay Cell transfer delay Cell delay variation (CDV) Cell transfer capacity Jitter	Bit error rate (BER) Packet, frame or cell misinsertion rate (CMR) Packet, frame or cell error ratio (CER) Severely errored frame or cell block ratio (SECBR) Errored seconds	Probability of information loss Packet, frame or cell loss ratio (CLR)
Disengagement	Delay in connection clearing	Premature release ratio Incorrect release ratio	Release failure ratio

network performance parameter which result. These are termed the *generic primary performance parameters*. Similar performance parameters could also be conceived for *connectionless* networks.

Examples of specific primary performance parameters within each of the generic performance parameter classes are given in Table 34.3. A subset (perhaps all) of these parameters should be regularly measured by network operators (as network performance measures). The target values for the chosen parameters need to be set to ensure relevant customer satisfaction with respect to quality of service. Where no direct relationship exists between quality of service and network performance parameters, experience will help to determine acceptable threshold values for NP parameters.

In addition to primary performance parameters, *ITU-T also defines the concept of derived performance parameters*. The most important derived performance parameters are those of *availability* and *acceptability*. *Availability* is a measure of the cumulative outage (i.e. non-service) time of the network as a whole (or of a given customer's part of the network) during a given measurement period (e.g. one month or three months). *Availability* is measured as the percentage of the total period for which the service was not in outage. Thus the higher the measured availability, the lower the network outage. Typical target values are 99%, 99.5% and 99.9%.

*Acceptability* has also been proposed as a derived performance parameter. This would be intended to give a qualitative measure of likely subjective customer opinion of the service level. This gives the potential for inclusion of other more general factors in measuring overall customer satisfaction with the network.



### 34.6 PERFORMANCE MONITORING FUNCTIONS OF MODERN NETWORKS

In the most modern network technologies (e.g. SDH, synchronous digital hierarchy (Chapter 13) and ATM, asynchronous transfer mode (Chapter 26)), extended performance measurement and monitoring techniques are built-in. Thus, for example, it may be possible to measure the instantaneous *bit error ratio (BER)* of a given connection, or the propagation delays being experienced in a live network. This is done by inserting extra performance management traffic into the network and monitoring the experienced performance. This is intended to be a good predictor of likely network performance as perceived by end customers.

The performance monitoring function of ATM operates by sending performance monitoring cells on an end-to-end basis after each block of  $N$  user cells. This function can be used to monitor end-to-end *errored blocks*, *misinsertion* (incorrectly sequenced or incorrectly delivered cells), *cell transfer delay* and other performance parameters on a specific connection.

As an example of the network monitoring and diagnosis tools available within ATM, it is, for example, possible to apply a loopback to *OAM (operations, administration and management)* cells connections within an ATM network without affecting other virtual connections sharing the same physical connection. This may be useful to confirm circuit continuity and to measure loop propagation times, without having to interrupt the live user connections.

### 34.7 NETWORK PERFORMANCE PLANNING AND MEASUREMENT

The remainder of the chapter proposes a pragmatic approach for establishing QOS and NP parameters for telecommunications services. The approach is designed to generate a balanced mix of parameters that will prove valuable in monitoring all aspects of service quality, including not only the quality of the connections themselves, but also the important support activities such as maintenance and provision of service.

By defining a standard set of categories relating to telecommunications service quality, we can ensure that selected parameters within each category give a balanced picture of current network performance and quality of service as against target. Without such discipline in the past, many network operators have made extensive measurements of the technical aspects of performance with too little respect for the broader *service attributes* giving them a distorted view of their quality, and they have suffered accordingly. For although they may have been returning exemplary technical performance statistics, their customers may have been far from happy because they could never get a reply from the fault reporting bureau, or could never get assistance about how to use a service. A possible set of categories from which to choose a balanced set of performance parameters is shown below. Each category is explained, and some examples of actual parameters are given in the paragraphs that follow

- customer service
- service availability

- provision and alteration of service
- service reliability
- making calls (and *clearing* them)
- quality of conversation or information transfer
- fairness, reliability and accuracy of service charges

### 34.7.1 Customer Service

Parameters in this category measure general customer perceptions of the ‘helpfulness’ of network operator staff, their courteousness, their ‘understanding’ of problems and their ‘willingness to help’. Some parameters might additionally reflect customer feeling about the availability and usefulness of literature, operating instructions and advertisements. Parameters in this category have historically been the most overlooked, perhaps because it is difficult to measure them accurately by market research, and perhaps because they are sensitive to swings in public opinion. Typical parameters might well include ‘percentage of satisfied customers’. Great care is required when selecting the precise wording of market research questions.

### 34.7.2 Service Availability

A customer’s perception of the quality of service may depend on its availability. For example, the geographic coverage may be too restricted, either prohibiting the customer from service subscription in the first place, or being too constraining in the number of egress points available. One’s view of a payphone service, for example, is much affected by the time it takes to find a kiosk at the time when wanting to make a call. It is also influenced by the number of payphones available at busy sites such as stations and airports where customers face intolerable delays. A third factor influencing the quality of payphone service might be the destinations which are available from it. For example, payphones which only accept small-denomination coins have little appeal for foreign tourists making international calls to their home country. Typical parameter measures in this category might thus be

- percentage of customers desiring service who are not in the coverage area (waiting list length)
- average queueing time (for a payphone)
- percentage of desired destinations not available (from market research)

### 34.7.3 Provision and Alteration of Service

Having ordered a service, the customer expects it to be provided promptly and conveniently. The time taken to complete customer orders, and perhaps the percentage of

appointments kept are worth measuring. Other parameters could also measure the standard of installation work. Typical measures might be

- percentage of contracts for the provision of service completed within a target number of days
- percentage of appointments for provision of service kept
- percentage of installations meeting customers' quality expectation
- number of installation revisits required

#### 34.7.4 Service Reliability

Having had a line established on the premises, a customer will wish to make and receive messages or calls according to whim. Loss of service occurs each time the user is unable to make or receive messages or calls for any reason other than customer mis-operation. Furthermore, when he has suffered a loss of service, the customer is anxious to be restored to full service within an acceptable period of time, and with minimum inconvenience. Thus *time to repair* or *time for repair* are important. Other measures in this category might be

- total period of lost service within one year
- number of occasions of service loss within one year
- average time lapse between reporting a fault and satisfactory; restoration of service
- percentage of faults cleared within target time period
- average number of faults per customer line per year
- percentage of payphone time out of order
- percentage of payphone time out of order due to full coinbox

#### 34.7.5 Making Calls (and Clearing Them)

Perhaps the most over-indulged category from the historical point of view. Having had the line and terminal equipment installed, the customer will be interested in establishing calls or sending messages over the network. The customer expects prompt service, and correct delivery of messages, or connection of calls to the right number. Network congestion is a frustration. A plethora of easily quantifiable and measurable parameters are available

- percentage of calls receiving dial tone within target time (fractions of a second)
- *percentage of calls* connected to the correct number

- *percentage of calls failed* due to network congestion (a network performance measure)
- average time taken between commencement of dialling and receipt of ringing tone
- average delivery time required (for a packet data message)
- percentage of calls not clearing correctly at the end of a call
- *answer bid ratio (ABR)*: the ratio of calls answered to the total number of call *bids*
- connection accessibility (*availability*) objective (see ITU-T Rec G.180).

### 34.7.6 Conversation or Information Transfer

The prime reason for the service to convey conversation or other information. The customer demands clarity of speech, lack of noise, sufficient loudness, absence of interruptions or interference, and security of the connection (from cut-off, or intrusion by a third party). For data, accurate conveyance of information is required (e.g. low jitter (variance in speed), low *bit error ratio (BER)* (introduced errors), etc.). Measurements of parameters in this category have historically been made by taking a small statistical sample of human service observations. A small percentage of calls are listened into by human operators, to confirm intelligibility of speech, to detect any apparent dissatisfaction of customers, and to trace the cause of call failures during set up. With the growth in telephone traffic and the increased sophistication of services there is now a demand for exchange manufacturers to develop more accurate computer-controlled sampling methods. Typical parameters in this category include

- percentage of calls with too faint conversation
- percentage of calls with excessive distortion
- percentage of data calls with excessive jitter
- percentage of calls cut off prematurely
- average time to deliver a data packet
- proportion of misrouted calls, undelivered or *misinserted* data packets, frames or cells

### 34.7.7 Service Charges

The charge invoiced to a customer reflects the accuracy and adequacy of the invoicing procedure as well as that of the exchange call metering device. Errors can be introduced by the mechanism of collecting call data, by the transfer of that data to the billing

system, by the aggregation of individual call charges, and by the transfer of that information from electronic to printed form. All these functions must work at an acceptable standard. Typical check measures might be

- percentage of (correctly) disputed bills
- percentage of individual calls incorrectly charged
- percentage bill inaccuracy on a per customer basis
- probability of an individual customer's bill being wrong

### 34.8 A FEW PRACTICAL TIPS

Best network performance is achieved when the network itself is designed from the start to specific *quality of service* and *network performance* targets. As far as possible these should be directly measurable parameters, and the threshold values should also be defined, if possible together with remedy actions to be taken. The remedy may not always be a simple one. As we learned in the various chapters on data networking, a slow propagation time of packets or frames across a network and the associated response time of a given computer application may not be significantly improved simply by increasing the bitrate of the line. Indeed, as we learned in Chapter 30, the upgrading of a line in a data network to a higher speed may lead only to greater problems.

Heterogeneous networks (i.e. those comprising many sub-networks of different technologies and topologies) are particularly hard to manage, because the monitoring of end-to-end connections is difficult to achieve without considerable effort and external measurement equipment. The task of narrowing down the exact source of a problem may be nearly impossible, being difficult to re-create in simulated conditions within the test laboratory.

Simple network design guidelines and defined network performance measures are the best formula for success. Complex networks, with complicated topologies and very sophisticated automatic control and routing mechanisms (e.g. the *Internet*) are very difficult to monitor and manage. Providing a better guarantee of successful message delivery and acceptable propagation times will be a major challenge for the many network operators and service providers involved with the *Internet*.

### 34.9 SUMMARY

This chapter has stressed the importance of good quality in telecommunications services and has outlined the two concepts *quality of service* and *network performance*, categorizing both to allow the easy establishment of a balanced set of parameters for continuous monitoring and correction of service performance.

---

## *Charging and Accounting for Network Use*

---

So far we have concerned ourselves entirely with the technical and operational side of running networks and providing telecommunications services. Network operators, if they want to stay in business, also need to think about capital and operational costs, and how to recover them from their customers. For a network to be financially sound and independent its tariff structure needs to cover the total expenditure made up of

- provision of lineplant, exchanges and buildings (i.e. capital costs)
- service maintenance and administration costs
- interest on capital
- depreciation of equipment
- research and development
- future capital investment
- profit and shareholders' earnings

Government regulation in some countries set tight financial limits on profits and targets, whereas in other countries the tariff structure is tailored to meet political ends, and financing is direct from government coffers. Nonetheless, the financial practices of charging and accounting for network use are common throughout the world, and they reflect the principles explained in this chapter.

### **35.1 RECOMPENSE FOR NETWORK USE**

Public telecommunications customers will be familiar with the arrival of their monthly or quarterly bill, comprising both subscription charges (paying for the rental of their line), and a charge for usage (the number of calls made, or packets of data sent, etc.). The practice of charging for network use is not confined to public networks; private network operators, especially the large corporations, are increasingly levying internal

‘transfer charges’ on individual departments for their use of telecommunications services. This is to keep close control of areas of company activity representing either ‘profit centres’ or ‘cost centres’. Additionally, accounting between telecommunications operators is becoming more common, as a result of the recent large increase in the interconnection of networks.

ITU-T defines two different types of financial transaction, to each of which different practices and a very particular vocabulary apply. The two types of financial transaction are described below, and a number of key items of vocabulary are highlighted.

### 35.1.1 Charging (Also Called Billing)

This is the name given to the practice of recovering financial costs from *end-users* for their use of the network. *Charges*, including subscription charges and *usage* charges, are accumulated over a standard period, typically three months, and are *billed* to the customer. Monies are then *collected* from the customer, typically a telephone line subscriber.

### 35.1.2 Accounting

The term *accounting* is normally applied only to the practice of financial *settlement* between different network operators, where they mutually cooperate to provide services. An accounting settlement is usually made by the owner of the originating network (i.e. the one collecting the customer charges) to all the other operators participating in the connection. This reimburses the other operators for the use of their networks.

## 35.2 CUSTOMER SUBSCRIPTION CHARGES

Customer charges, also known as *tariffs*, usually comprise two parts, a subscription charge and a usage charge. The subscription charge is made to offset the *standing* or *fixed* costs, for example, the rental of a telephone or other end terminal, the provision and upkeep of the local exchange and of the access line, and the operator’s general administrative overheads. The actual rate of the subscription charge depends on the type of access line, the end equipment provided, and the services to which the customer subscribes. They are easily calculable from a fixed formula and are usually charged by quarterly bill. Examples of services which attract a higher subscription charge are circuits specially designed for carriage of data, or lines with supplementary services such as *closed user group (CUG)*, *diversion of incoming calls* (on *customer busy* or *customer absent*), *three-party-call facility*, etc.

## 35.3 CUSTOMER USAGE CHARGES

Customer usage charges are made in addition to subscription charges, to offset the use-dependent (the *variable*) costs of network provision. As we have seen in the chapter on

teletraffic theory, the greater the amount of traffic between any two exchanges, the greater the number of trunks that are required between them. The cost of these extra trunks is recovered in the usage charge, which is usually calculated on a *cost-plus-mark-up* basis to reflect the network resources required to provide the service. The charge for any given call (or for the carriage of a data message) thus usually depends on the following parameters

- the type of network used
- the duration in time of the call, or
- the volume of data carried, and
- the distance of the communication

In addition, more sophisticated factors may also be taken into account.

- whether operator assistance, or some other *value-added service* was provided
- whether the call (or message) was conveyed during the network's peak usage period (the charge may be time-of-day dependent)
- how many call attempts (whether abortive or not) were made; making a charge for each call attempt dissuades users from congesting the network by incessantly recalling busy numbers

To calculate the charge due from each individual customer, the network operator must have a means of monitoring and recording each incidence of usage. In most networks charges are levied only on the call (or message) originator (so that the receiver does not pay). Usage is normally monitored by the originating exchange, which must determine the origin and destination of each call (or message), its duration, and the time-of-day when conveyed. In a circuit-switched network the origin of a call can be determined simply by detecting which incoming line is in use. The destination is then determined by analysis of the number dialled during call set-up. The chargeable duration of the call (the *call duration*) is normally taken to be the *conversation time*, which is determined by timing the period between the *answer signal* and the *cleardown* signal at the end of the call. Call duration on manual networks, or of calls established *semi-automatically*, can be measured directly by the human operator, but for automatic networks different procedures are used. There are two principal means for automatic customer charge monitoring: *pulse metering* and *electronic ticketing*, as explained next.

## 35.4 PULSE METERING

The older and simpler charging method (still used in many telephone networks) is to levy call charges on the basis of the number of pre-determined intervals of time used in the course of the call. This method is called *pulse metering*. In it a fixed price is charged for a small time period or *unit* of the conversation phase. When the call is first answered, the first *unit* (of time) commences. After a pre-determined period of time (the



length of which will depend on the location of origin and destination and the time of day), the first *unit* will expire and a second *unit* of time will commence. The charge payable is calculated by multiplying the number of units used by the fixed price per unit. The customer is usually unaware when each *unit* ends and the next begins, or indeed how much of a partially used unit remains.

On different calls, e.g. to destinations at different distances away, or on calls to the same destination made at different times of day, the time duration of a *unit* may be varied. In this way, the average per-minute-charge made for different calls can be varied. One caution, however: it is not generally understood that the upper limit of metering pulse rate may not actually be determined by the capability of the pulsing unit, but on the ability of the transmission line to carry the pulse correctly and the meter to clock properly.

Over a given period of time (say three months) the number of *units* used by the customer is accumulated on a counting device (called a *cyclic meter*), located at the exchange. Location at the exchange eases the job of reading the meters, and reduces the scope for customer fraud.

During the course of individual calls, the cyclic meter is triggered to step onward one count at the beginning of each new *unit* of time. This is done by sending a periodic electrical pulse to the meter. The customer's invoice shows the total number of *units* used, and a financial charge which is calculated by multiplying this figure by the fixed price per unit.

While calls are being set up by the exchange, the exchange control and routing system (*common equipment*) obtains the destination of the call from the number dialled, to determine the appropriate outgoing route and charge rate. For routing, as we described in Chapter 28, a fair number of digits may need to be examined to decide to which precise exchange the call should next be routed. However, although the same piece of common equipment may determine both the route and the call charge rate, it is likely that fewer digits will need to be analysed for charging purposes. This is because most network operators choose to run *chargeband* schemes, in which destinations within a fairly wide geographical *zone* (or *chargeband*) are charged at the same rate. This eases the administration of the scheme and the customer's comprehension of it. As an example of a simple chargeband scheme, a public telephone company might opt to levy charge for international calls based only on the first digit of the *country code*. This would divide the world into eight distinct zones, as we saw in Chapter 29 (North America, Africa, Europe, South America, Far East Asia, Russia and ex-Socialist Republics, Middle East, and Central Asia). A similar national chargeband scheme could equally well be developed corresponding to the area code digits of the national telephone number.

Each of the chargebands needs a *charging programme*. The *programme* divides the day into a number of time periods, for each of which a different per-minute-charge will apply. At the busiest times, we expect higher rates, whereas cheap rates at off-peak times may help to stimulate extra traffic and revenue without needing more equipment. As an example, within the United Kingdom, British Telecom sub-divides its chargebands into times of day called *peak rate* (corresponding to the busiest times of day, when the highest per-minute-charge is levied), *cheap rate* (corresponding to periods of low activity, when the lowest per-minute-charge is made) and *standard rate* (for which a medium charge is made). Over the course of a weekday, peak rate applies from 9 a.m. to 1 p.m.; standard rate from 1 p.m. to 6 p.m., and from 8 a.m. to 9 a.m.;

and cheap rate applies each evening from 6 p.m. to 8 a.m. The per-minute-charge during the peak rate period is made more expensive by reducing the time duration of the unit (remember that the *units* have a fixed price). Figure 35.1 illustrates the relationship between price-per-minute and unit duration on the British Telecom scheme, and shows the pulsing signal required to step the customer's meter.

Exchanges which employ pulse metering to charge their customers usually have a number of machines or electronic devices to generate the meter pulse signals, one for each of the pulse supplies. A selector device, controlled by the dialled number analysis and by the time-of-day, connects the customer meter to the appropriate pulse supply for any particular call. More than one pulse meter supply may exist. For example, local calls will be pulsed from the local exchange (if metered at all). However, trunk and international call charge pulses will most likely be generated at the trunk or international exchange which analyses the destination *country code* digits. A simplified diagram of a pulse metering device is shown in Figure 35.2.

A very simple form of pulse metering, in which only one meter pulse is generated for each call, independent of call duration, is used in some networks, particularly for metering local calls. The advantage of such a charging philosophy is the simplicity of the

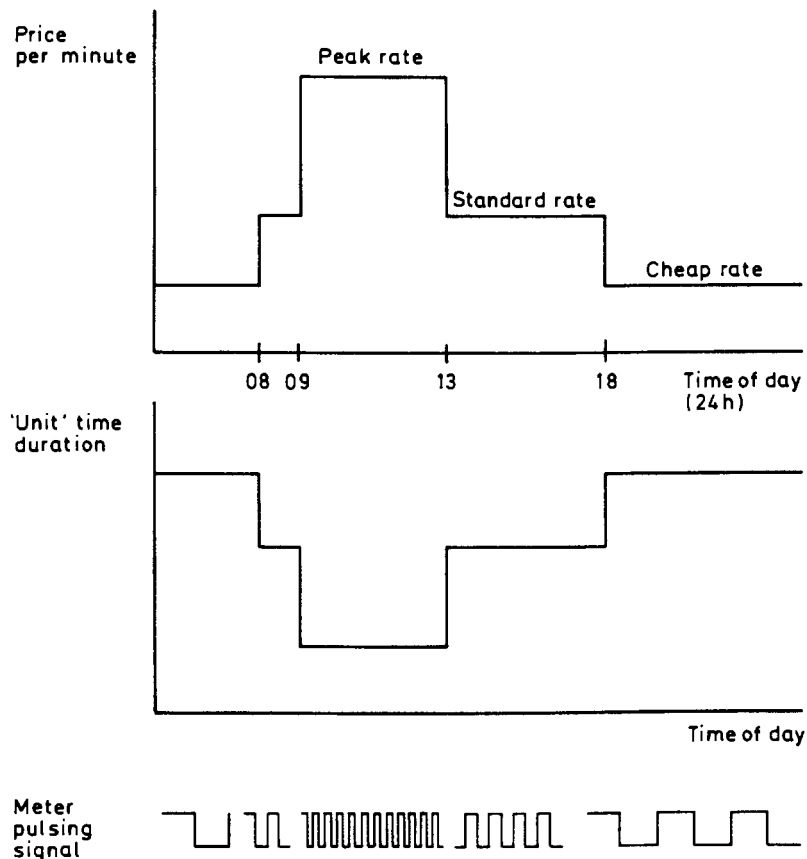
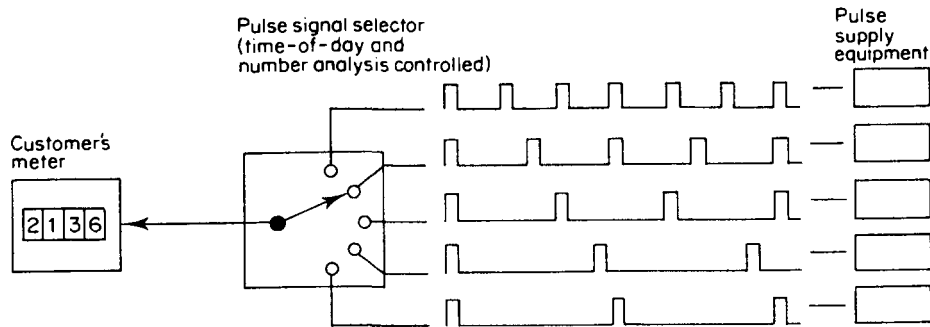


Figure 35.1 Pulse metering



**Figure 35.2** Pulse metering selector or 'tariff' equipment'

equipment required and the ease of administration. However, as the need to recover usage-dependent costs becomes more acute with the boom in demand (e.g. night-time dial-in calls to the Internet), so the method is falling into obsolescence. Another drawback is the tendency to promote network congestion and phone box queues!

## 35.5 ELECTRONIC TICKETING

*Electronic ticketing* (also called automated message accounting or *toll ticketing*) is far more accurate and reliable than pulse metering, and it is becoming common in consequence of the spread of stored program control exchanges. In this method the exchange control computer monitors and records information about each call, including the number dialled, the duration, the time of call set up, the time of call clear, etc. The information is stored in an electronic *call record*, usually on computer storage disk or magnetic tape, until the time when the customer is due to be invoiced. The invoice is calculated by adding the amounts due from each individual call, derived from the details on the individual call records. *Electronic ticketing* provides the scope for more complex chargeband structures, and the capability (much demanded by customers) for an itemized record of the details and cost of each call made.

## 35.6 ACCOUNTING

*Accounting* is the method of financial settlement between network operators, when more than one network is traversed by a call. Accounting has historically been necessary to allow reimbursement between the networks of different countries in recompense for their handling of international calls, but in addition accounting is also becoming common even between networks operating in the same country, as deregulation allows or even demands ever greater interconnection.

Most *accounting* is nowadays undertaken using computer monitoring devices to record the details of individual calls in a manner resembling *electronic ticketing*. The financial settlement is made on the basis of total usage over a given period of time, and is

paid from the call originating network directly to all other network operators. Alternatively, the accounting settlements may be *cascaded* between network operators, starting at the originating end, if more than two networks have been used (see Figure 35.3).

The direct method of accounting *settlement* (Figure 35.3(a)) is only possible when the owner of network A knows the exact path taken by calls in reaching network C, after traversing network B. For example, if there were also a route from network B to network C via a fourth network E, then the direct settlement method may not be possible, because owner A may not be able to calculate an apportionment in settlement for E. This is the main reason for the use of the *cascade* method of settlement shown in Figure 35.3(b). In this method the transit party settlements are proportioned and cascaded in sequence.

Accounting *settlements* depend on total usage measured in *paid minutes*. As we learned in Chapter 31, the number of accountable or *paid minutes* on each call is equal to the total useful time that has passed between answer and call clear-down. The settlement is calculated by multiplying the paid minute total by the *accounting rate*. The accounting rate itself is fixed by negotiation directly between the interconnected network operators, and is intended to reflect the costs associated with

- the destination of the call and the distance
- the services used

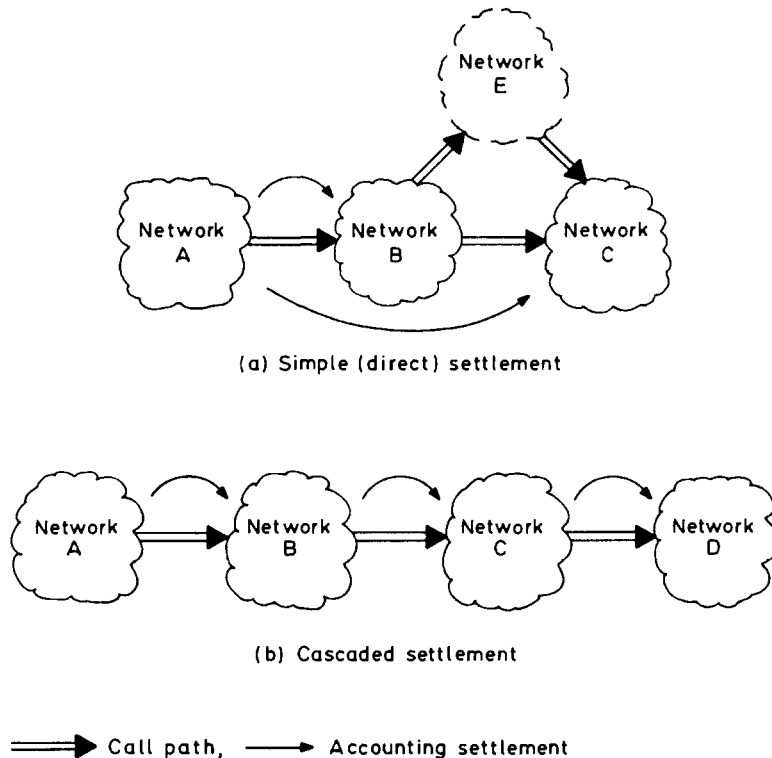


Figure 35.3 Cascaded settlement of account

- the call duration
- the time of day
- the type of switching
- the route taken, etc.

The *accounting rate* may be set in any accepted international currency (e.g. US Dollars, Sterling, *gold Francs*, *special drawing rights (SDRs)*), and may either be negotiated by the bilateral agreement of the parties, or in case of difficulty paid at standardized rates such as those recommended by the ITU-T D-series recommendations.

### 35.7 ROUTE DESTINATION ACCOUNTING

When there is more than one route to a given destination, *route destination* accounting must be undertaken; an example is shown in Figure 35.4, where calls between network A and network B may either be routed directly or may *overflow* via network C.

Correct reimbursement of both operators B and C is possible only if operator A records total paid minutes by route and destination (further sub-divided into peak, standard and cheap rates according to time of day if necessary). *Destination settlement* is due to the operator of network B for all calls, whether routed directly or not, but the payment may be slightly lower for calls transiting network C. This would reflect an apportioning of revenue to operator C and also the effective cost saving that results from the reduced number of direct (A–B) circuits that are required.

ITU-T recommendations demand that the operator of network A periodically (e.g. monthly or quarterly) *declares* two *Route–Destination* (or RD) accounts, corresponding to the total originated paid minutes sent

- to *Destination B*; by *Route direct*
- to *Destination B*; by *Route via C*

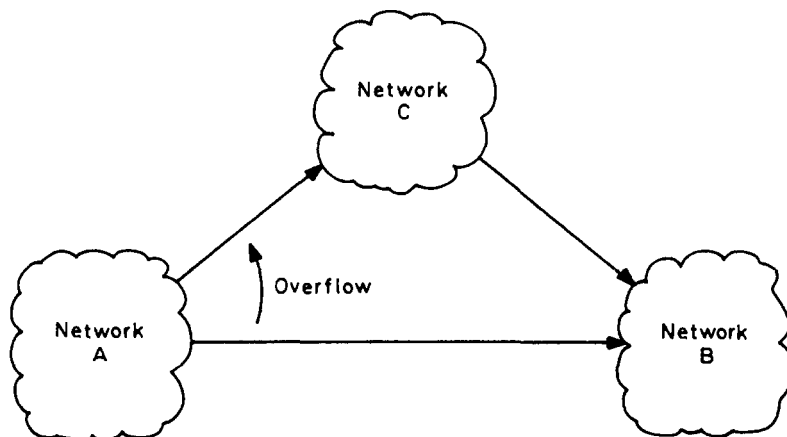


Figure 35.4 'Route destination' accounting

Declaration of accounts between network operators consists of exchanging information about each other's total originated paid minutes of traffic, categorized as Route/Destination and peak/standard/cheap time-of-day (as necessary). The due payment is calculated as the difference between outpayments due for outgoing traffic and revenue due from incoming calls; and one of the operators pays the outstanding balance accordingly. When incoming and outgoing traffic are roughly balanced, very little settlement is required, but usually there is a net financial gain for settling accounts with an incoming balance of traffic, and a net outpayment when most of the traffic is outgoing. For some developing countries the net balance can either provide a significant service of hard currency, influencing the speed of development of their networks. For others, the debt is a severe problem and embarrassment.

### 35.8 CHARGING AND ACCOUNTING ON DATA NETWORKS

When data are carried by a circuit-switched network, customer charges and inter-operator accounts are usually calculated according to the total time for which the circuit was available to the end users. The usage time, or *paid time*, is the time that elapses between the answer signal and the circuit cleardown. This period is equivalent to the conversation time of circuit-switched telephone network, and the same principles apply as have already been described. In addition a subscription charge is usually applied to the use of each customer network port. Data network types charged and accounted in this way include telex, circuit-switched data networks and ISDN.

In cases where data is carried by a *packet-switched*, *frame relay* or *ATM* (i.e. *cell-switched*) network, the accounting and charging principles are usually different. The subscription charges are typically on a per-port basis, but in addition, further charges may apply according to the number of simultaneous *logical channels* (i.e. *virtual channels*) which may be set up. Subscription charges may also apply for *supplementary services* including, for example, the *closed user group* (CUG) facility which allows only members of the CUG to communicate with one another. This may be a useful feature in a corporate data network, where the corporation does not wish outsiders to be able to access their computers.

Usage charges in data networks are usually based on one or more of the following factors

- the speed of the connection established (i.e. the bitrate, typically the *committed information rate*, CIR) is the determinant
- the duration of the connection in time
- the number of packets, frames or cells sent or some other measure of the transmitted volume of data

In packet-switched (i.e. X.25-based) networks it is usual to charge and account according only to the total volume of data carried by the network. The volume is quoted in *segments*. The standard *segment* size is 64 octets (or bytes), equivalent to  $64 \times 8 = 512$  bits. (Confusingly, however, a *segment* of SMDS data is only 48 bytes long.)

Most *frame relay* networks were built up offering only *PVC* (*permanent virtual circuit*) service. A frame relay PVC can be set up only by subscription, and is a permanent connection between two defined endpoints. So far as the end customer is concerned, a frame relay PVC appears much like a leaseline with a bitrate equal to the *committed information rate* (*CIR*), but at times of low network usage the customer may be able to send much higher speed data for short bursts (*excess bursts*). Frame relay connections have thus to date been charged in a similar manner to leaselines, based on a fixed monthly subscription charge according to the *CIR*. As *switched virtual circuit* (*SVC*) or *virtual call* (*VC*) frame relay connections become possible, these are more likely to be charged on a volume-related basis similar to the approach discussed above which is used in X.25 packet networks.

ATM networks are only just emerging, and no fixed pattern is yet established. As ATM is expected to be a 'universal' technology capable of providing many different types of voice, data and video transport services, I expect the emergence of different types of customer-charging principles, mimicking the charge structures of the equivalent services provided using 'legacy' technologies.

Methods of charging and invoicing customers, and for accounting and settling the inter-network account, are in general similar to the electronic ticketing technique used in circuit-switched networks.

### 35.9 CHARGING AND ACCOUNTING FOR MANUAL (OPERATOR) ASSISTANCE

Calls set up on behalf of the end user, and controlled by the network operator (i.e. manual or *semi-automatic calls*), are normally charged and accounted according to the details of the call, as recorded on the operator's paper ticket or electronic equivalent.

The chargeable (and accountable) time on an operator call is usually less than the total *conversation time* (the time between the called-subscriber-answer and the call-clear). This is because some time is taken up by the operator before the two end parties can speak. According to the nature of the call, the unchargeable conversation time may have been used by the operator

- to confirm that a particular person has answered the call, if a *personal call* has been requested
- to check that the receiver is willing to accept call charges if a *collect* (i.e. *reverse charge*) call has been requested
- for some other similar reason

Operator-controlled calls often carry a premium charge. This offsets the cost of the operator's time, and the extra unchargeable network time required at the commencement of each call. Operator-assisted calls are also accounted between network operators at a slightly higher rate than automatic calls, for similar reasons. The *manual* accounting rate is used when a second operator is also used in the terminating country, otherwise the *semi-automatic* accounting rate applies.

As we learned in Chapter 14, *semi-automatic* calls which route via *automatic* exchanges may be recognized by the presence of a *language digit* in the dialled number train, rather than by the automatic call *discrimination digit* (i.e. *language digit* value '0'). This fact prevents the possibility of accounting the same call twice, once on the operator's ticket and once on the automatic exchange accounting record.

### 35.10 CHARGING AND ACCOUNTING FOR LEASED CIRCUITS

Besides providing switched services, network operators are often willing to provide transmission bandwidth on a point-to-point basis. Circuits conforming to any of the standard bandwidths may be leased from the network either by private individuals as *private circuits* (otherwise called *leased circuits* or *leaselines*) or by international network operators (from other international operators) as *administration leases*. For both types charges and accounting rates are normally applied according to the straight line distance either between the end points of the leased section of the circuit or between the countries' capital cities.

For private circuits, commercial market rates are set by the network operators, but for administration leases, a lower accounting rate (nearer to cost) may be used. Within Europe, *TEUREM* accounting rates used to govern the charges made between network operators for administration leases. Nowadays these rates are tending to disappear under competitive pressure as *deregulation* leads to a more open market.

### 35.11 CHARGING PAYPHONE CALLS

Payphones present a particular challenge for network designers. Unlike normal subscription telephones, the call charges must be collected prior to or during the call, because to attempt to collect the money afterwards, or by invoice, is an invitation to fraud. Special telephones are required for the purpose; they collect the money 'up-front', calculate when the period paid for is about to run out and then they prompt the caller to insert more money. Forced clearing is applied to the call if the caller fails to insert more money when the paid time has expired. As with operator-controlled calls, charges for payphone calls are generally slightly higher than normal calls. The extra charge offsets the greater costs of the telephone, its maintenance, and upkeep of the kiosk, which clearly cannot be defrayed against a subscription charge.

### 35.12 CUSTOMER BILLING

Compared with some of the other administrative and management tasks associated with running a network, the process of billing is relatively straightforward, involving collecting the available information from several sources and calculating an 'amount due' sum relevant to each customer based on a series of product-dependent tariff formulae and



customer-specific negotiated discount schemes (as recorded in the contract archive). This does not mean to say that it is done well. Many of the world's public network operators have enormous problems with their billing systems

The most challenging part of a good billing system is achieving sufficient flexibility in the tariff models which may be applied to products, and sufficient reliability to achieve high turnout of bills at month end and to satisfy even the most stringent auditor. Simultaneously to providing all this flexibility in software routines, the same system must be capable of processing huge quantities of data accurately, reliably and auditably, meeting the billing date each month.

### 35.13 SETTING CUSTOMER CHARGES AND ACCOUNTING RATES

Setting customer charges (*tariffs*) and accounting rates is a skilled task in which network operators have not only to ensure coverage of their costs, a reasonable profit, and a return on investment, but also to give customers a 'fair deal'. The setting of *tariffs* and *accounting rates* therefore demands careful consideration, together with acute awareness of market expectations and willingness-to-pay. Anticipation of competitors' reactions is important, as tariff or accounting rate changes, particularly increases, tend to have a deterrent effect on demand, and may have adverse political or public relations consequences, factors which network operators cannot afford to ignore.

Figure 35.5 illustrates revenue breakdowns for the *originated* and *terminated* traffic of an imaginary network. The *calls* or *connections* have been connected in cooperation with another operator and relevant account settlements are shown.

In Figure 35.5(a), the overall revenue (area of the pie chart) is large, reflecting the call collections from telephone user tariffs. Unfortunately, however, over half of this revenue has needed to be paid out to the terminating network operator in settlement of accounts.

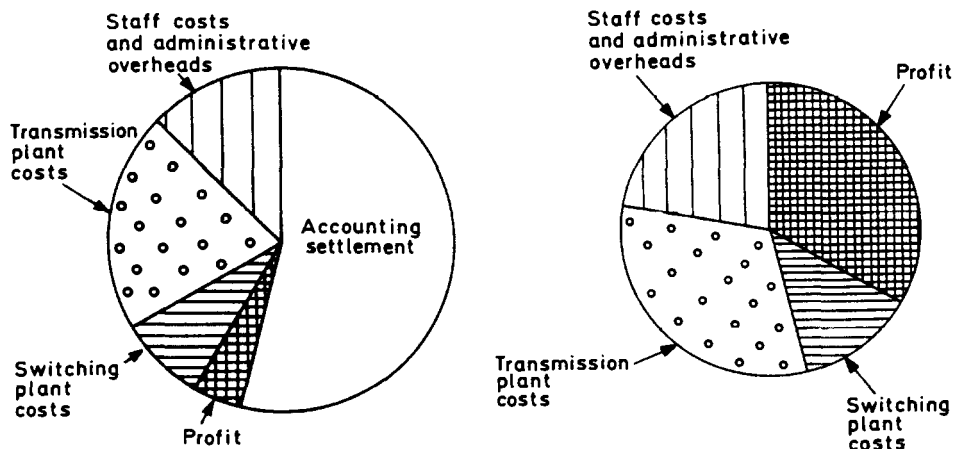


Figure 35.5 Apportionment of revenue

You might think it fairer that the settlement should be almost exactly half of the collection revenue, and historically this was so; but although the cost of equipment has fallen and the tariffs charged for calls have been forced down by market pressures, *interconnection charges* and *international accounting rates* have not always fallen appropriately quickly. Both *interconnection charges* and *international accounting rates* are typically difficult to negotiate down because the balance of traffic (and therefore the net payment) often favours the other or more dominant network operator. Thus there is a network outflow of traffic, the high *interconnection charges* and *accounting rates* which exist for historic reasons can be heavily penalizing. This is a particularly acute problem for national telephone companies in developing countries, and can be a major problem for the 'new' operators emerging as competitors to the old state monopoly telephone companies in the developed nations.

In developing countries it is often the case that an outdated and congested national network is the cause of more successful outgoing international calls than incoming ones. The net outpayment settling the account can be an extreme embarrassment and problem for the country's overall balance of trade. Surprisingly perhaps, even between more advanced nations, a PTO (public telecommunications operator) may make considerably more money from incoming calls received from the distant PTO than from outgoing ones generated from its own customers. This despite the fact that his own outgoing customer tariff may be nearly twice his compatriot PTO's price for incoming calls! In time this accounting rate injustice is likely to be redressed, but as you can imagine it is a very sensitive issue between world governments.

In the case of *interconnection charges* it has been realized by most national telecommunications regulators that these charges can create a major obstacle in the establishment of a more open and *deregulated* market. The *interconnection charges* set by dominant market players are thus nowadays usually kept under close scrutiny. Nonetheless, it is very difficult for regulators, given the difficulty in allocating the costs of the ex-state telephone monopolies, to determine what realistic fair charges for *interconnection* should be. Many regulators have *determined* interconnection charges, having observed continuing unresolved disputes between incumbent operators and new players, but their success in setting fair charges has been varied. Some new operators receive the assurance of a new business 'on a plate', enabling them easily to undercut the prices of their established competitor. In other countries the new players struggle to develop market share because of difficulties in achieving an adequate financial margin.

The next major element of expenditure shown in the pie chart of Figure 35.5 is the transmission plant costs, then the company overheads and finally the switching costs. All this leaves the margin available for profit.

For a contrast, look at the terminated (i.e. incoming) call revenue breakdown shown in Figure 35.5(b). The overall revenue (the size of the pie chart) is smaller, but it is greater than half of the outgoing call revenue (provided that the accounts settlement is more than half of the outgoing collections, and there is a rough balance of traffic). Meanwhile the costs are similar (for similar traffic demand), so much larger amount of the revenue goes to profit.

The example of Figure 35.5 is typical of the revenues and profits that an international telephone network operator may earn from a given overseas route. Note the extreme sensitivity of the profits and revenue to the call charge tariff and the accounting rate. Note also the difference in profitability between *originated* and *terminated* accounts.

### 35.14 NETWORK COSTS AND HOW TO RECHARGE THEM

Networks rely on the cooperation of individuals to share communications, means and costs. No-one would use the public telephone network if he had to pay for it all; it would be far too expensive. Because everyone else in the country subscribes to a line and pays for it, a customer has a wide choice of people that he can call for a moderate and affordable cost. Once a network has reached a certain critical size and has become stable, its financing and its cost allocation become relatively straightforward, based on the *average long term marginal cost*. We return to exactly what this is later. For smaller networks, particularly those which are either growing or shrinking, cost allocation is more problematic.

If all the cost incurred before the first customer were all charged to the first customer, there would never be a first customer. Not only would he have no-one to call, it would cost him a fortune if he could do so. The problem is that just at the time when most money is needed, the network has least to offer. The usual ways around the problem are first to seek an initial critical mass of customers, so that shared between them both the service and the costs are acceptable. The second way is to accept the need for a fairly lengthy payback period on the project (i.e. by depreciating the capital cost over a longer period, say ten years instead of seven), so lowering the annual costs which must be recharged to customers. This second approach puts the network operator into a risk venture business.

The risks are twofold: first that the equipment will not last the planned lifetime, second that customers find they no longer need the service and stop paying for it earlier than anticipated by the budget. In this case, there may be a temptation to recalculate costs, dividing by the number of remaining customers. The remaining customers in consequence find their costs higher than expected, and are likely to re-evaluate their use of the network. As more customers pull out, the problem only becomes worse.

The morals are simple; first recognize that operating a network is a business risk venture. Second, a direct division of telecommunication network costs between all the customers is not necessarily a good approach to determining appropriate charges.

Consider a value-added service provider installing a new voicemail system. Let us imagine that the minimum size of equipment which may be installed would serve 100 customers and that this can subsequently be outfitted for the needs of individual extra users up to a maximum of 400 customers. To provide for the needs of more than 400 users a second system must be purchased. The overall costs (capital investment) plotted against the number of users are then as shown in Figure 35.6.

The question is, what is the right price to set for the initial customers, assuming that the limit of 400 customers will be relatively quickly surpassed? We discuss two possibilities, shown as pricing line 1 and pricing line 2 in Figures 35.6 and 35.7.

You may consider pricing line 1 (Figure 35.6) to be the fairest method of recharging users. This represents the exact division of the costs for a fully utilized (400-user) system between the 400 users. This would be the *average long term marginal cost* that we discussed earlier. Using it at start-up leads to almost certain problems. First, if the system is only ever subscribed to by 350 users then there is a lot of money short (because the cost line on our diagram is above the price line (i.e. recouped costs) at this point. Worse still, if 401 users turn up then virtually the entire initial cost (\$200) of the second system will be in arrears.

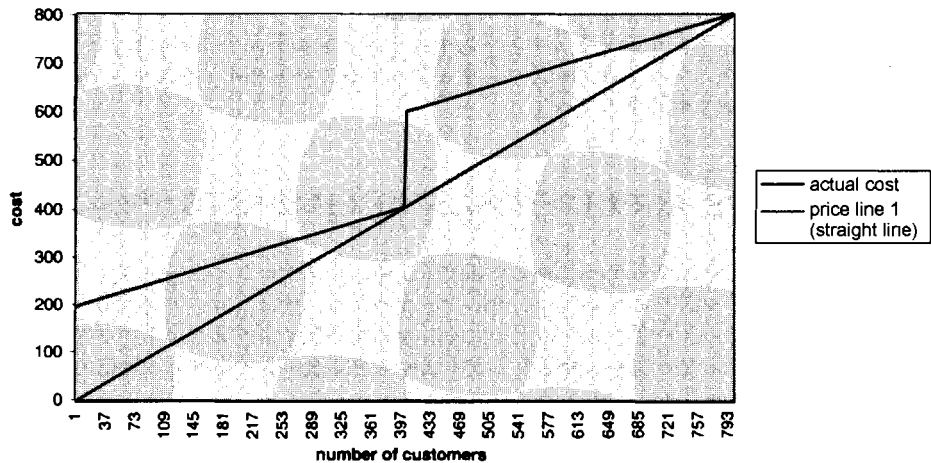


Figure 35.6 Alternative pricing strategies based on average long term marginal cost

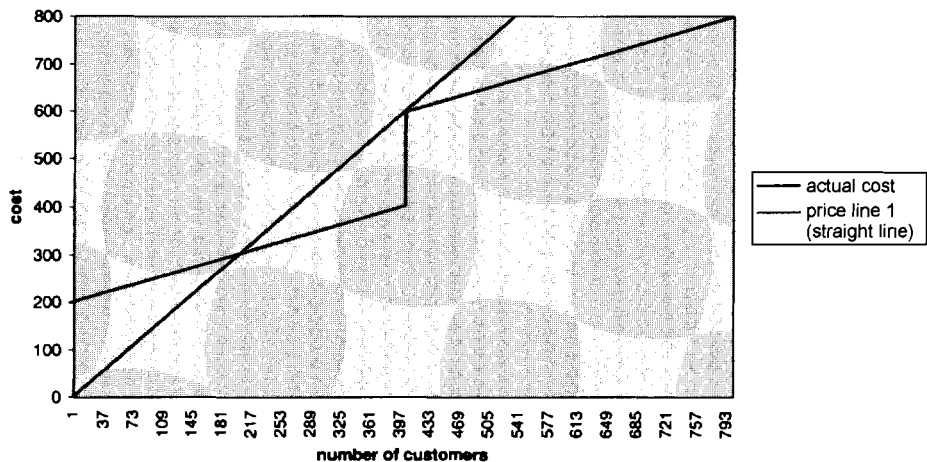


Figure 35.7 Alternative pricing strategies: initial recovery of investment

A better pricing line to use initially is pricing line 2 (Figure 35.7). This divides the full cost of the first 400 user system *and* the costs of the initial part of the second system between the first 400 users. This has two benefits; first, the venture is in profit after about 200 users are added to the first system. Second, there is enough money available to install the second system as soon as the need arises. However, as other businessmen notice the market potential they are likely to enter the market with their own competing services. In the long term, therefore, the price will be oriented to the *average long term marginal cost* (pricing line 1) because this represents the minimum cost per customer. The end-customer price in an established market is likely to be this cost plus a reasonable profit margin.

It is the average long term marginal cost which is the goal of regulators when determining interconnection charges. The problem is in trying to find out exactly what these costs are (which costs of a large ex-monopoly should be apportioned to which customers and services?).

### **35.15 FUTURE ACCOUNTING AND CHARGING PRACTICES**

More and more new telecommunications services emerge and the world is becoming littered with new and competing network operators, new charging and accounting methods and new monitoring techniques. They are bound to be more complicated than those that we have already, and they will demand skilful administrators and engineers to realize them and set the rates.

---

## *Maintaining the Network*

---

No matter how much careful planning goes into the design of a network, and no matter how reliable the individual components are, corrective action will always be required in some form or another, to prevent or make good network and component failures, and maintain overall service standards. However, attitudes towards maintenance and the organization behind it vary widely, ranging from the 'let it fail then fix it' school of thought right through to 'prevent faults at any cost'. This chapter describes a typical maintenance regime in its philosophical, organizational and procedural aspects.

### **36.1 THE OBJECTIVES OF GENERAL MAINTENANCE**

As succinctly stated by ITU-T, the objective of a general maintenance organization is to minimize the occurrence of failures, and to ensure that in case of failure

- the right personnel can be sent to
- the right place with
- the right equipment at
- the right time to perform
- the right corrective actions.

### **36.2 MAINTENANCE PHILOSOPHY**

In pursuing these objectives, the wise network operator establishes a maintenance philosophy closely linked with overall targets for network quality and for the

proportion of time that the network is intended to be fault-free (*available*). In this task he will take due account of network economics and the most likely causes of failure. Networks fail for all sorts of reasons; common examples are

- cable or connector damage or disturbance
- equipment overheating
- electronic component failure
- mechanical equipment jamming, or other failure
- mechanical wear
- dirty (high resistance) relay or switch contacts (a diminishing problem as electro-mechanical exchanges are withdrawn)
- power supply loss
- vandalism (e.g. to public payphones)
- software errors
- erroneous exchange data
- poor connections between cables or other components (e.g. *dry soldered joints*)
- interference (e.g. due to electromagnetic disturbances, recent standards *on EMC, electromagnetic compatibility*, are designed to ensure that equipment does not cause electromagnetic disturbance and is itself not unduly sensitive to such interference, i.e. is *electromagnetically protected*)

Each cause of failure has its own cure, but broadly speaking, there are three main approaches

- *corrective* maintenance
- *preventive* maintenance
- *controlled* maintenance

*Corrective maintenance* is carried out after the failure has been diagnosed, and it consists of the repair or replacement of faulty components.

*Preventive maintenance* is to eliminate the accumulation of faults. Preventive maintenance usually consists of *routine* testing and *correction* of working equipment (as opposed to failed equipment) to prevent degradation in performance before any failure actually occurs.

*Controlled maintenance* is a more systematic approach, combining both the corrective and preventive methods. The underlying philosophy of controlled maintenance is to prevent network failure. This is done by using special analysis techniques to monitor day-to-day network performance and degradation, thereby avoiding maintenance work.

The advantage of controlled maintenance is that it concentrates on areas where the customer is likely to benefit most, and it reduces the extent of preventive maintenance and the complications of corrective maintenance. When new networks are being designed or extended, or when new capabilities are being added to existing networks, consideration needs to be given to the maintenance philosophy and to the organization, the maintenance facilities, and the test equipment that will support it. The best *controlled* mix of both corrective and preventive philosophies depends on the number and nature of problems. These, in turn, depend on the overall network structure and the component equipment types. So, in the days of widespread electromechanical switches and relays when a frequent cause of failure was mechanical wear, a good deal of time was spent on preventive type maintenance, oiling the moving parts as it were. Nowadays, when hardware faults in modern electronic equipment are relatively rare and software faults often take some to present themselves (the exchange operating fault-free for extended periods between occurrences), a corrective philosophy is adopted. Faulty component boards are completely replaced without even attempting a diagnosis, and software faults are debugged as they arise.

### 36.3 MAINTENANCE ORGANIZATION

Real networks are in a constant state of change throughout their lives. To match traffic demand, new circuits are continually established between exchanges. Established circuits may need to be re-arranged as transmission systems are upgraded or taken down and faulty equipment needs to be repaired, replaced or avoided by a diversion. For optimum efficiency the organizations set up to establish these maintenance tasks, and the tools with which they are provided, should be planned in such a way that lifetime costs will be minimized. There is a choice, for instance, between paying more at the start for expensive but reliable equipment, or using cheaper equipment and incurring higher ongoing running costs. Lifetime cost analysis must include

- initial cost of equipment
- cost of spares and test equipment
- ongoing running and maintenance costs
- costs associated with periods of lost service

High wages and skills shortages in recent years have weighed the scales in favour of using more reliable equipment and a smaller maintenance workforce. Indeed, in some instances the field workforce has been pared to the minimum of two people, one worker plus a stand-in to cover annual leave and periods of sickness. Some observers question the sense of this, pointing to the fact that so complex are the devices, so computerized the routine activities and so rare the faults, that the field maintenance staff often do not have the experience to cope.

For this reason a comprehensive headquarters *maintenance support*, *technical support* or back-up organization (sometimes called *second line support* or *third line*



*support*) is needed in addition to the direct maintenance staff, to perform the following functions.

- To provide detailed equipment and maintenance documentation.
- To provide maintenance training on new equipment and document 'fail-safe' maintenance procedures (explaining the general methods to be adopted, and making sure that unintended disturbance to other customers is not caused by maintenance action).
- To develop and put in place a fault-reporting procedure.
- To repair complicated items of equipment or resolve complex software problems.
- To develop and procure the necessary test equipment.
- To maintain an appropriate store of spare parts and to call for re-design of poor equipment, taking duly into account the failure rate of each item, the number of items in operation, and the actual repair turn around time (e.g. repair time, or delivery time for a part not held in stock), and calculating the service and revenue risk if no spare part is available.
- To develop and maintain an equipment identification and inventory scheme for tracking equipment in use and spare equipment either in stock or on order (in addition, maintenance staff in different exchanges need to be able to indicate faulty lines or circuits to one another).
- To prepare a list of *contact points* and telephone numbers through which the maintenance staffs in different maintenance centres may communicate.

The direct maintenance workforce is usually collocated with the exchange, some staff being switching experts while others have transport so that they can go out and deal with exterior plant problems. The staff located within the exchange provide a 'control point' for new circuit *lineups*, and for the initial reporting and diagnosis of faults.

The number of staff located at any exchange depends on the size, complexity and reliability of the exchange. Not every exchange can justify its own on-site maintenance staff, and out-stationed staff may be posted to the exchange either on a regular preventive maintenance schedule, or simply when there is a fault.

### 36.4 CENTRALIZED OPERATION AND MAINTENANCE

A modern practice, aimed at reducing the maintenance workforce, is to leave exchanges unmanned. Computer technology and *extended alarms* allow staff in a single *centralized operation and maintenance (CO&M)* centre to monitor and control a number of different exchanges in *real time* (giving instantaneous and *live* control of each exchange). Figure 36.1 illustrates a typical centralized operation and maintenance scheme.

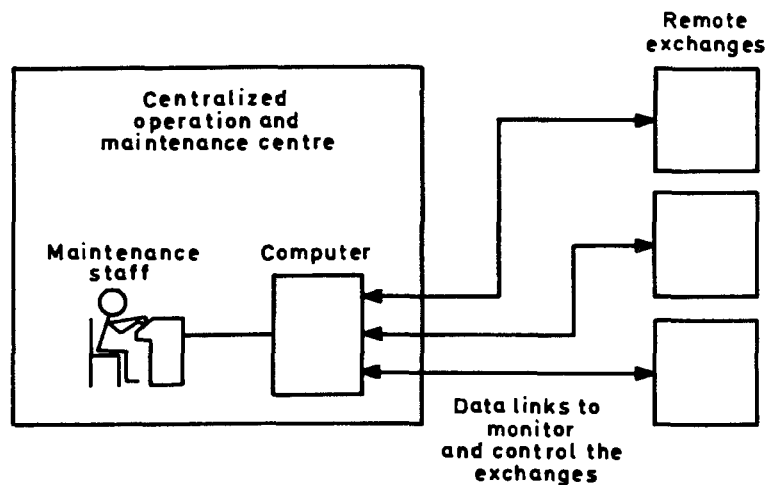


Figure 36.1 Centralized operation and maintenance

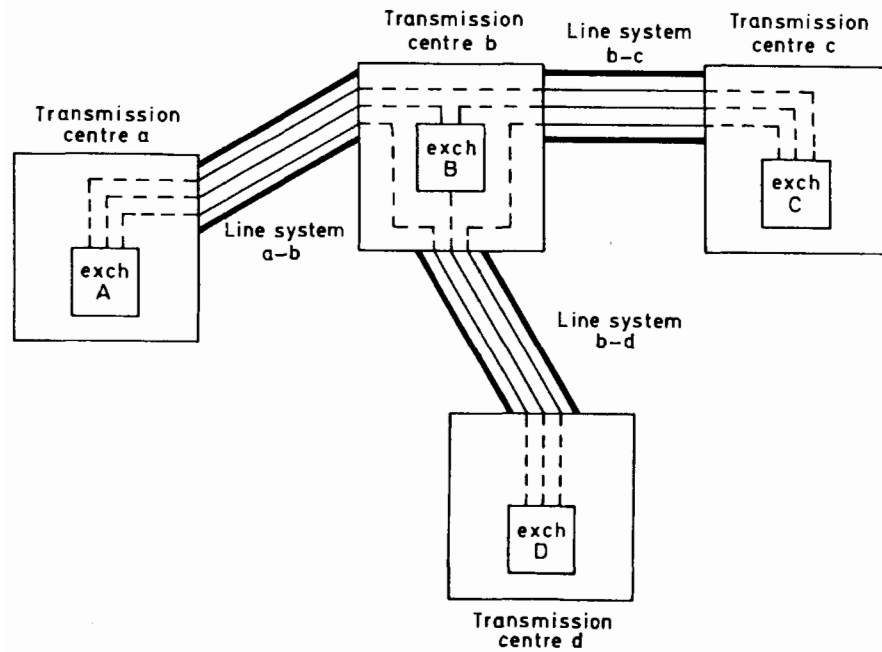
Centralized operation and maintenance (CO&M) can be introduced only when the remote exchanges are computer-controlled, and are designed to be capable of self-diagnosis of faults. A datalink back to a computer at the centralized operation and maintenance centre allows the maintenance staff to monitor the exchange performance, noting any problems and applying any necessary controls. Under a CO&M scheme, the exchanges are designed with duplicated items of equipment, which remain idle until they are activated electronically to take over the function of a failed item. The exchange can thus continue to work at full load, while a member of the maintenance team is sent out to the exchange site, to repair the faulty equipment, or to replace it completely by a *circuit-board* change.

### 36.5 LINING UP ANALOGUE AND MIXED ANALOGUE/DIGITS CIRCUITS




The transmission links between exchanges are commissioned (or *lined up*) using a two-stage method as follows. First, the lineplant itself is established in sections which are tested and calibrated in turn, and then connected together. A number of reference measurements are made along the entire length to check and calibrate the overall end-to-end performance. When the line system as a whole has been established, multiplexing and other terminal equipment is applied to its ends to obtain the individual circuits or groups which may be tested individually. The individual circuit testing is necessary because, as Figure 36.3 shows, a real circuit (or group) is likely to traverse a number of line systems, which may interact adversely. So although each section in isolation may be within limits, the combination may not be so. The calibration of each group and circuit is thus carried out on an end-to-end basis.



**Figure 36.2** Maintenance workstation. The AT&T 5ESS telephone exchange now includes a video monitor that displays colour-coded diagrams, each of which indicates the status of a certain part of the system. A central office technician is checking the status of digital trunks. (*Courtesy of AT&T*)



Key :

-  line system (e.g. coaxial cable, microwave system etc.)
-  individual circuit within line system
-  jumper wires
- exch exchange

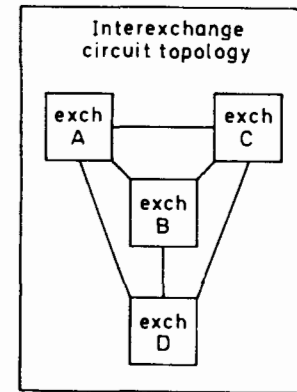


Figure 36.3 Circuit set-up

Figure 36.3 shows four exchanges A, B, C and D, located in transmission centres a, b, c and d. From the inset diagram of Figure 36.3 we see that the exchanges are configured topologically in a fully interconnected manner. Each exchange has a circuit to every other exchange but the total of six circuits has been achieved with the use of three line systems only: a–b, b–c and b–d. Where the transmission centres are not directly interconnected by a line system, a circuit has been provided by the concatenation of two line systems, with a *jumper wire* completing the connection across the intermediate transmission centre. Thus for example the circuit from exchange A to exchange C uses line systems a–b and b–c, and a jumper wire across transmission centre b.

The number and diversity of calibration measurements and adjustments necessary during circuit line-up, and the amount of deviation allowed in the ongoing values depend on the type of circuit (e.g. analogue or digital), and on the use to which the circuit is being put (e.g. voice or data). High grade data circuits, for example, have more stringent line conditioning requirements than simple voice grade circuits. The measurements and adjustments ensure that the circuit conforms with the transmission plan (see Chapter 33). Thus on analogue and mixed analogue/digital line systems and circuits *line-up*, measurements will be made of

- overall *loss* in signal strength (in dB)
- amplitude loss/frequency *attenuation distortion*
- *group delay* (particularly if the circuit is to be used for data)
- *noise, crosstalk, echo*, etc.
- inter-exchange signalling tests (if appropriate)

Various equipments, including *amplifiers, equalisers, filters*, and *echo controllers*, are then adjusted to bring the line conditions within the set limits, using measuring equipment as follows

- signal tone generators (calibrated for precise frequencies and signal strengths)
- calibrated frequency and signal strength detectors
- noise meters
- equipment for inter-exchange signalling or data protocol testing

First, the end-to-end circuit loss is determined by sending a calibrated 1020 Hz (or in purely analogue networks, 800 Hz) signal of a known strength, and measuring the received strength of this allows the circuit amplification to be adjusted accordingly. Next, a range of different calibrated signal frequencies across the whole circuit bandwidth is sent, and the received signal strengths are again measured. This allows the frequency distortion equalizers to be adjusted. Then the group delay distortion is corrected using group delay equalizers. This equalization is particularly important for high speed modem data circuits, and it is achieved by measuring the relative phase of different frequencies relative to the 1020 Hz (or 800 Hz) signal. Following this,

psophometric noise checks are conducted, together with tests of inter-exchange signalling systems or of data protocols (e.g. X.25, frame relay or IBM's SNA), and finally a test call is established.

Where a pure tone signal of calibrated strength needs to be injected into the digital part of a mixed analogue/digital connection or network, this can be done either using an analogue tone sender and an analogue/digital converter, or by the use of a *digital reference sequence (DRS)*. A *digital reference sequence* is a digital bit pattern corresponding to a particular analogue signal frequency and strength. Such a pattern is easy to store in computer-like memory and is a very reliable means of reproducing an accurately calibrated signal.

## 36.6 HIGH GRADE DATA CIRCUIT LINE-UP

In high grade data circuits which use modems over analogue or mixed analogue/digital plant, a number of extra *line-up* measurements may be necessary, as follows

- *weighted noise*
- *notched noise*
- *impulse noise*
- *phase hits and gain hits*
- *harmonic disturbance (or inter-modulation noise)*
- *frequency shift distortion*
- *jitter*

Weighted noise is a measure of the noise in the middle of the channel bandwidth. Noise frequencies in this range are most likely to cause modem errors. *Psophometric* (European) and *C-message filters* (United States) are used to measure this type of noise.

Notched noise is measured by applying a pure frequency tone at one end and removing it with a *notch filter* at the other; the remaining noise is then measured. The notched noise itself arises from the way in which signals have been digitized or otherwise processed over the course of the link. It is thus similar to *quantization distortion*, which was discussed in Chapter 5.

Impulse noise is characterized by large 'spikey' waveforms and arises from unsuppressed power surges or mechanical switching noise. It is most common on electro-mechanical switched networks.

Phase hits and gain hits are intermittent but only moderate and relatively short (less than 200 ms) disturbances in the phase or amplitude of a signal. Typically, less than 10 should be recorded in a 15 minute test period. Special test equipment is required. More serious *gain hits* are called *dropouts*. *Gain hits* are most troublesome in voice use; phase hits manifest themselves as bit errors in data signals.

Harmonic disturbance may result from the intermodulation of two different signal frequencies  $F_1$  and  $F_2$  when passed through nonlinear processing devices. New stray signals of frequencies  $F_1 + F_2$ ,  $F_1 - F_2$ ,  $F_1 + 2F_2$ , etc., are produced. This type of disturbance is measured by a spectrum analyzer.

Frequency shift is also measured by a spectrum analyzer. Frequency shift is obviously a problem for a data modem if the frequency received differs to such an extent from that sent as to be mis-interpreted. It is most likely to occur when a carrier system or other frequency modulating signal processing has been used.

Jitter or *phase jitter* arises when the timing of the pulses on incoming data signal varies slightly, so that the pulse pattern is not quite regular. The effects of *jitter* can accumulate over a number of regenerated links, and they result in received bit errors. It can be reduced by reading the incoming data into a store and then reading it back out at an accurate rate, using a highly stable clock controlled by a *phase-locked loop (PLL)* circuit.

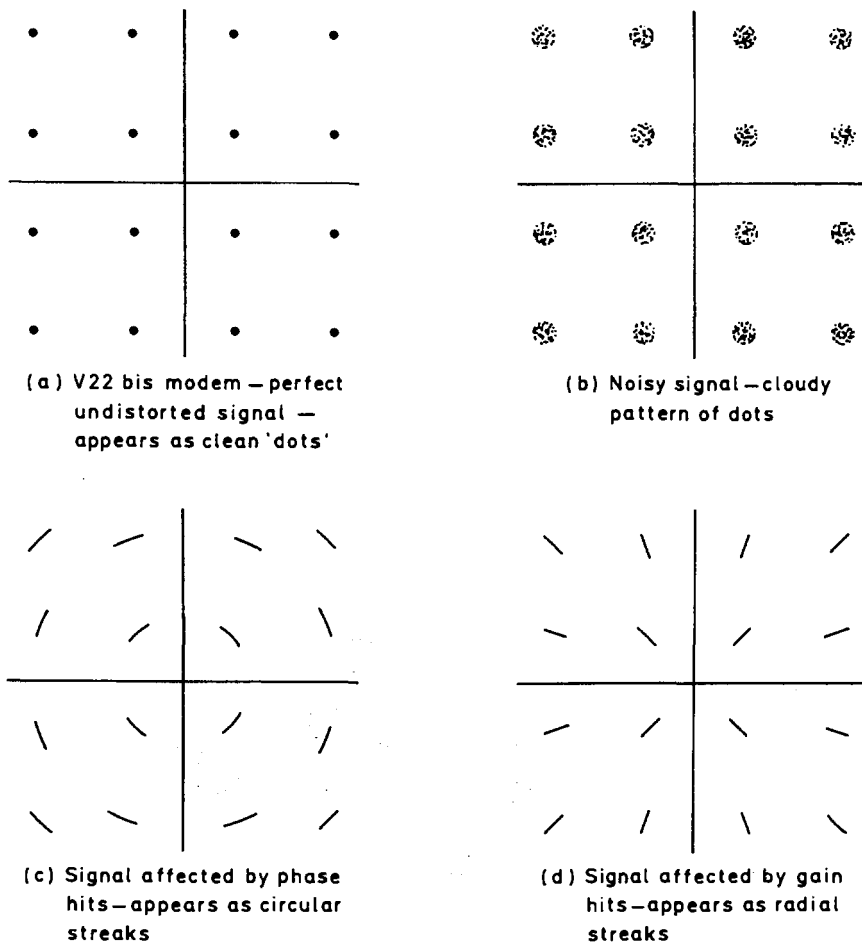


Figure 36.4 Detecting analogue line disturbances using constellation diagrams

All of the above parameters should be checked when the line system or circuit is first established. Problems are likely to reflect poorly designed or poorly installed equipment, and the best remedy is prevention: check the quality of work which is made, because correction circuits are expensive and not entirely effective.

In Chapter 9 we introduced the idea of constellation diagrams for modems. We are now in a position to illustrate their practical use for detecting noise, phase hits and gain (or amplitude) hits on analogue lines employing modems. Special test equipment may be connected at the receiving end of the line in place of the receiving modem. The test equipment displays on an oscilloscope-like screen the constellation pattern of the received data signal. Disturbances appear as shown in Figure 36.4.

As is apparent from the patterns of Figure 36.4, the problem with noise, phase and gain hits is that if they become too great, they result in the incorrect interpretation of the signal; the received bit pattern then includes errors.

## 36.7 LINING UP DIGITAL CIRCUITS

Digital line systems and their *tributary* bit streams must conform to a different transmission plan from their analogue equivalents and are therefore lined up in a different manner. The important parameters, as we saw in Chapter 33, are

- the *bit error ratio* (*BER*)
- the network *synchronization*
- the *quantization* or *quantizing distortion*

In practice, the network *synchronization* (i.e. the *jitter* and clock accuracy) and the *quantization distortion* are set by the design of a circuit and can be improved only by re-design. The lining-up process can only address the question of error rate. The accuracy of synchronization depends on the use of highly stable clocking sources and inter-exchange synchronization links, as described in Chapter 33. Quantization distortion is a form of noise, which affects analogue voice and data signals when they are carried over digital media using *pulse code modulation* (*PCM*) and other signal processing techniques (see Chapter 5). It can be reduced only by re-designing the circuit to avoid multiple signal conversions (e.g. analogue to digital conversion, *signal compression*, etc.). In the case of digital data circuits, these should never be designed to include any form of signal processing because digital data devices are intolerant of the high bit error rates that arise from *quantization distortion*.

Standard practice for digital circuit line up is to perform a test of digital error performance. This may involve a lengthy stability test over several days or simply be a *quick-check* (15 minute) test. A *pseudo-random bit pattern generator* (a digital signal generator) is used to provide a test digital signal. During the test the proportion of bit errors (the *bit error ratio*, or *BER*) is measured, along with the proportion of *error free seconds* (*EFS*). Expected values are typically no more than 1 error in  $10^5$  or  $10^9$  for *BER* and at least 99.5% *EFS*. The errors, if excessive, may have arisen as the result of any number of different impairments. Some causes can be eliminated easily (e.g. by increasing the transmitted power to reduce the effect of noise). Other causes need more radical circuit checks.



Like their analogue equivalents, digital bit streams are lined-up in two stages, first at higher order (e.g. at 140 Mbit/s, if this is the line system bit rate) and then on an end-to-end basis for each tributary stream (e.g. 2 Mbit/s, 1.5 Mbit/s). The secondary checking of lower order tributary streams is necessary for the same reasons as in the comparable analogue case exemplified by Figure 36.3.

## 36.8 PERFORMANCE OBJECTIVES

In recognition of the fact that practical networks can never match idealized performance objectives, ITU-T recommendation G.102 sets out different operating limits for the target *performance objective*, the *design objective*, the *commissioning objective* (also known as the *line-up limit*) and the *maintenance limit*.

The *performance objective* is the ideal performance level for the particular application.

The *design objective* transfers this objective into a realizable target range, within which economically designed equipment can be expected to operate, given optimum conditions of power supply, temperature, humidity, etc.

The *line-up limit* recognizes that the optimum conditions can rarely be achieved in practice, but nonetheless sets a stringent practical range within which the equipment must operate on first establishment. Any faults identified during line-up, which cause the circuit to operate outside this range, should be cleared and the circuit re-lined, before taking it into service.

The *maintenance limit* is the least stringent range of operating conditions, but still within a 'tolerable' range as far as the application is concerned. The limit is chosen so that, if exceeded, a fault is considered to exist. The fault should be cleared and the circuit re-lined-up. As Figure 36.5 shows, each of the last three performance ranges described is a slight relaxation of its predecessor, so giving targets which are achievable in practice. Thus even if a slight degradation in quality occurs after circuit *line-up*, the circuit still operates within the *maintenance limit* operating range.

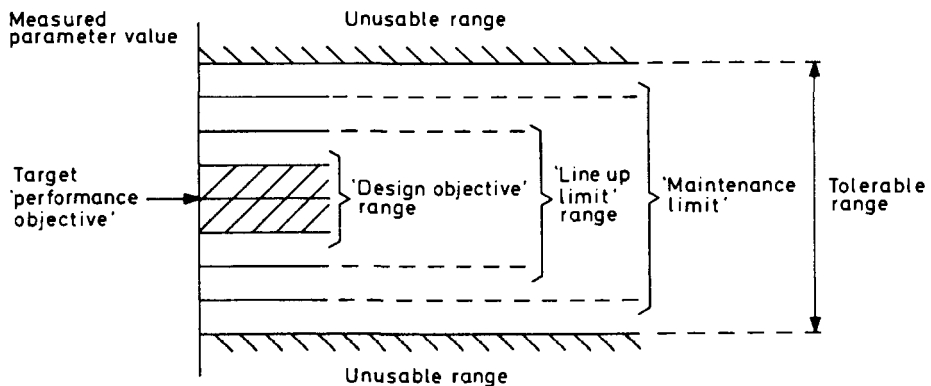


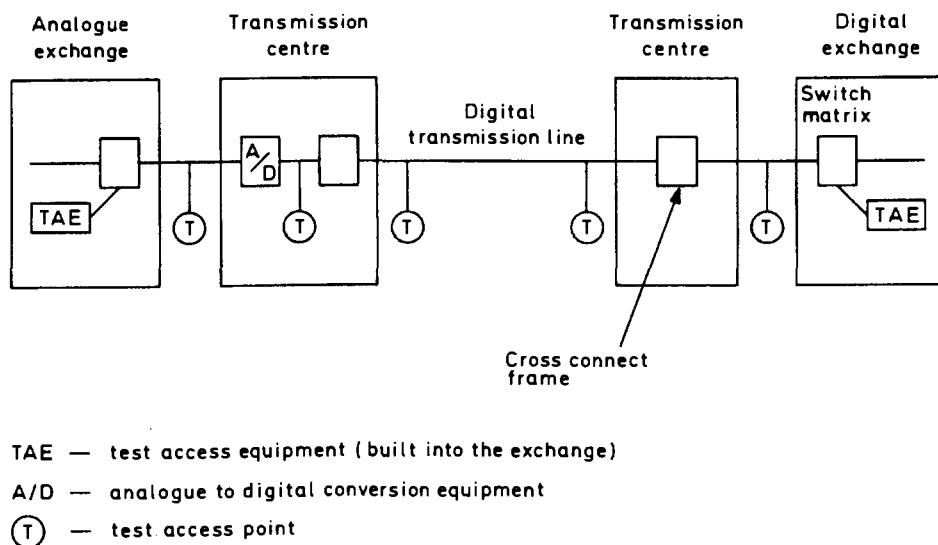
Figure 36.5 Performance objectives: 'line up' and 'maintenance'

A steady deterioration in performance of items in service can be expected as the result of operational use, occasional overload and general ageing as well as genuine faults. Often the degradation is slow enough to be imperceptible and not to warrant routine checks. For this reason, it is common practice to use automatic alarms to alert maintenance staff to the need for corrective action, so that faults can promptly be eliminated and the equipment returned to its line-up operating range. Two levels of alarm are possible, one at the *maintenance limit* level, and one at the 'unusable' performance level. Faults on the latter level clearly need more urgent attention than those on the former.

### 36.9 MAINTENANCE 'ACCESS POINTS'

For the purpose of network lining-up and subsequent maintenance, it is necessary to provide a number of *test access points* for maintenance. Ideally, *test access points* are provided at a number of different points in a network to enable easy localization of faults and initial segment-by-segment circuit alignment. A very large number of test points, increases the overall network and equipment costs, and may in itself be a source of unreliability. Figure 36.6 shows possible test access arrangements for a simple network of two exchanges (one analogue and one digital) interconnected by a digital transmission link. Note how test access points have been provided between all the major items of equipment. This allows the causes of any faults or line-up problems to be quickly narrowed down.

As is typically the case, both exchanges in Figure 36.6 have built-in test access equipment for diagnosing and correcting faults within the exchange itself. In addition,



**Figure 36.6** Test access points. TAE, test access equipment (built into the exchange); A/D, analogue to digital conversion equipment; T, test access point

five external test access points enable the exchange's external conditions to be verified and faults in other equipment along the link to be *localized*, diagnosed and corrected.

## 36.10 LOCALIZING NETWORK FAULTS

In an efficient maintenance organization, faults are discovered before the customer finds them. In this way some of the faults can be corrected even before the customer is aware of them. Means for fault detection include

- equipment *alarms*
- equipment routine testing (so-called *routing*)
- live network monitoring
- customer fault reporting

As before, a combination of techniques is usually appropriate. The four are discussed in turn. Equipment alarms are used to indicate an abnormal state (i.e. when the maintenance limit has been exceeded) in an equipment or an environment of crucial importance. (An environment *fault* might be too high a temperature: computer equipment is prone to failure under such conditions.) Alarms are usually ranked in order of importance, and are usually generated by continuous monitoring of some type of 'heartbeat' signal. When the heartbeat stops, it is time for action. For example, the *pilot* signal on analogue transmission systems is a low level and single tone frequency outside the normal bandwidth range which is continuously transmitted. Absence of the pilot signal at the receiving end is interpreted as a transmission link failure, and it is usually notified to maintenance staff as an audible and/or visible alarm).

Equipment routine testing, on the other hand, is suited to any equipment which is fault- or wear-prone (e.g. mechanical equipment). It can be carried out manually or by automatic test equipment, the choice depending on the number and type of devices, their complexity, and the periodicity of tests. For routine network testing, ITU-T has described a number of *automatic transmission, measuring and signalling test equipments (ATME)* in its O-series of recommendations.

Live network monitoring helps to pick up transient faults and then avert them; network overload can thus be foreseen, and its adverse effects corrected, as we shall see in Chapter 37. Useful means of monitoring live network performance include call completion rate and congestion statistics, and quality sampling of a small number of connections. For example, a sudden flood of calls or a rapid drop in the call completion rate may indicate the onset of congestion caused by network failure or quality degradation.

Not all faults, however, can be detected other than by the user, and so in the last resort an efficient customer or user fault reporting and handling procedure is paramount.

After detecting a fault, the next step is to localize it and diagnose its nature, all of which may be directly apparent or correctly reported, but usually more diagnosis is required. For example, when a major transmission system fails, a whole host of alarms

may go off in the maintenance centre, indicating failure not only of the main transmission system, but also of each of the derived circuits or tributaries. The alarms must be cleared in priority order. Correcting the main transmission fault first may clear the alarms on each individual tributary, but will not do so if lesser problems remain on particular tributaries. Thus, in our example, maintenance staff should first attempt to localize the main transmission failure to a given section of the link. This they can do by making use of the various test access points available to them, and by liaison with staff in other maintenance centres through which the link passes. The example is illustrated in Figure 36.7, where the failure in one of the channels in section A–B and the link A–B–C has been detected by maintenance staff in station C. The absence of the main transmission pilot signal in the receive channel at station C has raised an alarm. Telephone liaison with the staff in station B reveals that though this alarm actually notifies one of the loss of a whole analogue or digital group of circuits, it is evident that this is the likely cause of the problem on the particular A–C channel. If instead the fault had lain on the transmission link between stations B and C, the alarm in station B would not have gone off. Furthermore, if the fault had been in station C's transient channel, the alarms would have gone off at one or both of stations A and B.

Some faults in switched networks, can be extremely difficult to trace. A common problem that is encountered when trying to locate faults in switched networks is trying to trace which precise links and exchanges were traversed at the time of the fault. It is like knowing that your friend is driving between London and Birmingham but not knowing which route he took. You know he has broken down, but where do you look first? Faults of this nature can persist for many months and are finally cleared, either as the result of routine maintenance, or because routine testing of one of the individual network components revealed the fault.

Not all switched network faults are difficult to trace. For example, a phenomenon known to all experienced maintenance staff is the occurrence of a *killer trunk*. Imagine that the first choice circuit on the route between the main London-to-Birmingham telephone route has become faulty. Imagine also that the fault results in any call attempting to use the circuit being immediately failed and released. Affected callers receive nothing at all! No so bad, you might think: one circuit faulty in hundreds will not make much difference. However, the fault has an incredibly wide-ranging effect, nearly 'killing' all the traffic on the route. The reason is that, being the first choice circuit, all calls will attempt to seize it before trying other circuits. Of course, any call that attempts to do so is immediately failed and released. So nearly all calls fail! Hence

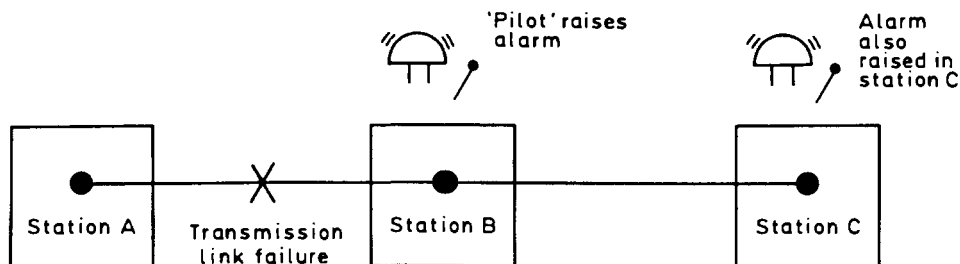


Figure 36.7 Transmission failure and alarm

the name *killer trunk*. To the experienced maintenance man, however, the killer trunk phenomenon shows up like a sore thumb, because traffic records reveal

- a huge number of short holding time calls
- a very low call success (i.e. answering) rate
- very little overall traffic in erlangs
- virtually no activity on many circuits

The condition can be cleared in the first instance simply by *busy*ing out the circuit (i.e. making it unavailable for use). A repair can then be carried out. The worst effects of a killer trunk can also be prevented by making sure that circuits within the route are chosen randomly, not always scanned in the same order.

Before leaving the subject of network test points and fault localization procedures we should also mention the use of loopback techniques for detecting faults within the *tail* part of a circuit between the network operator's site and the end user's terminal. This part of the circuit can be the most inaccessible (say in unmanned premises), or it may be subject to quite adverse conditions in congested ducts, strung between telegraph poles or even draped across desks. It is just as prone to failure as any other section, and the network operator needs to be able to localize faults here as anywhere else, ideally being able to distinguish such faults from faults in the user's terminal equipment, and preferably without even visiting the user's location. What makes this possible is a circuit *loopback* or a responding equipment.

A *loopback* simply *loops* the user's receive channel directly to the transmit channel, so enabling the network operator to send a test signal both ways along the tail and to confirm its correct return. Loopbacks may be either manually operated by the user (on request from the network operator) or invoked by the network operator using a

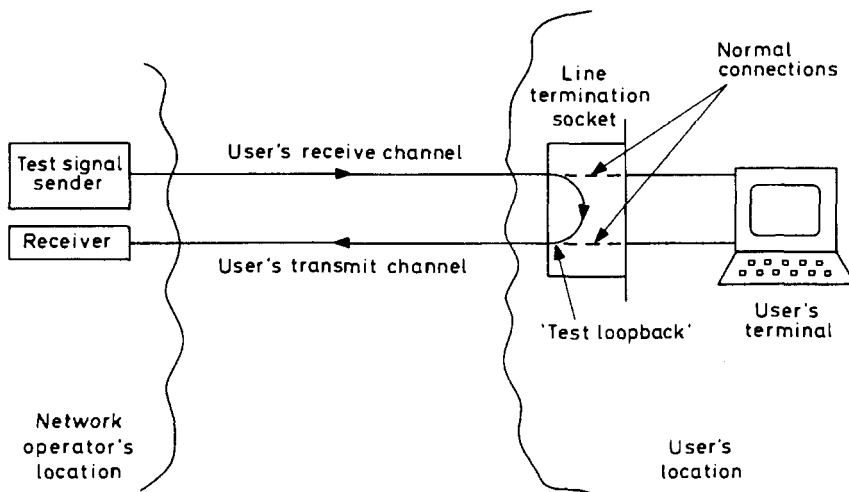


Figure 36.8 Testing a connection tail using a loopback

2713 Hz tone or equivalent signal to switch a remotely controlled equipment within the line termination socket at the user's premises. As we see from Figure 36.8, the loopback enables the faults to be localized beyond doubt as being either on the line or in the user's terminal equipment, all without a visit to the user's premises.

Loopback techniques are widely used on point-to-point data connections, where users are intolerant of any significant downtime. PTOs provide them on the tails of their private leased circuits, and other data network providers put them on their equipment circuits. Loopbacks for modems are described in ITU-T Recommendation V.54.

Another way of testing lines to remote unmanned locations is to use responding equipment which answers calls made to the location, giving standard test signals and other responses in accordance with commands.

### 36.11 HARDWARE FAULTS

Historically the most common faults have come from mechanical equipment or hardware. Some maintenance organizations are well prepared to deal with such component failures, and general wear and tear, by preventive action or by repair, but it must be recognized that the semiconductor age has brought with it much greater levels of reliability and complexity and much greater risk associated with failure. All this has forced change on the handling of software faults.

Nowadays, it is common to design electronic equipment to be tolerant of faults; computer processors and network exchanges have their crucial parts duplicated, the two halves sharing the traffic load and continually monitoring one another for faults. If a fault is detected in either half, then it is shut down and the other half takes over the full traffic load until the faulty half is repaired by maintenance staff, in response to the alarm.

Self-diagnosis software within the computer or exchange is used to localize the fault to a particular circuit board. Correction is achieved merely by sliding the board out and replacing it with a new one. The processor or exchange can then be restored to normal working, and the circuit board can be repaired at leisure, sent back to the manufacturer or thrown away.

### 36.12 SOFTWARE FAULTS

Errors in computer programs (or *software*) are becoming the most common cause of faults in telecommunications equipment. Newly developed software is often littered with errors or *bugs*, of an unpredictable nature which come to light in operation, sometimes years later, so that it is difficult to build experience. In addition, because the background of most telecommunications technicians is not in computer programming, the skills for rectification of faults are extremely rare.

Software bugs may take some time to find, and they require the expert attention of specialized *maintenance support* staff and the original designer, being outside the competence of normal direct maintenance staff.

So, you may ask, how do we keep the service on the exchange running while we sort out the problem? The answer is by initiating a manual or automatic *restart* procedure which *re-boots* (*resets*) the whole exchange to allow it to carry on working. This is done by re-loading historical data and software, overwriting any corruptions which may have crept in as a result of the software bug. The software and data resident in the exchange processor at the time of failure is downloaded for a later fault diagnosis, which attempts to trace back the processing *events* leading up to the problem. Meanwhile the exchange is likely to run quite normally for some time, until a similar sequence of events conspires against it again.

Localized *restarts* of minor functional units are likely to be automatically invoked by the exchange itself to minimize the off-air time, but more (e.g. whole exchange) restarts may require manual initiation, to prevent an automatic restart from disrupting a large number of unaffected calls at an inopportune moment.

The computerization of many maintenance tasks may have greatly reduced the human numbers required, but the extra skills and knowledge demanded of those staff that remain make them a prized commodity!

### 36.13 CHANGE CONTROL PROCEDURE FOR HARDWARE AND SOFTWARE

As telecommunications equipment has become increasingly computerized over the last few years, so has the importance of effective change control procedures for the various hardware and software releases.

It is normal nowadays for equipment manufacturers to further develop their hardware and software in one year steps, offering new hardware and software *releases* at least once a year, if not more often. Sometimes these new releases resolve previous functional or operating problems, sometimes they bring with them new service capabilities. Nearly always, even for those releases which are *backward-compatible* with previous releases, is some sort of system procedure necessary. New releases may be *backward-compatible* only with the immediately preceding release, so that any older releases of the software or hardware may need to be updated before the new release can be introduced. On other occasions, some form of network, equipment or topology configuration change may be necessary before the new release is installed and activated. The act of installing the new release must be a carefully planned and executed procedure.

The first step in introducing a new hardware or software release to the network should be the verification of its correct functioning and good quality. Here an *offline* (i.e. non-live) test network is helpful, not only in checking the correct functioning of new service capabilities but also in the *regression testing* of functions already used extensively within the network (which were already available in a previous release, but may not operate in quite the same way in the new release). The test network also serves to practice the installation of the new release, refining the most appropriate order of steps to be taken to minimize service disturbance to customers when the release is installed in the live network.

Having installed a new release of hardware or software it is very important to keep an inventory of the hardware and software *level* installed in each of the individual nodes and equipments making up a network, so that further release updates can be correctly administered and installed. Ideally, each of the nodes should be updated to run at the most recent level of hardware and software. This much reduces the problems of different software and hardware vintages having to interoperate with one another, and usually also guarantees better support from the manufacturer should a problem arise. (The specialists always tend to be most acquainted with the latest version, tending slowly to forget the idiosyncrasies of older versions). Unfortunately, practical reality does not usually allow this for various reasons

- the equipment manufacturers' pricing scheme may charge software according to the number of nodes in which it is installed; if the service benefit is only warranted in a few nodes it may be uneconomic to install it in all the nodes
- a new software release may demand hardware upgrades which are not affordable
- services may operate in a different and (from a particular user-perspective) unacceptable way in a new release

Experience with *software* demonstrates that each new release and even each new *patch* (a software correction intended to eliminate a problem or software error) should be treated with caution. You cannot treat new software as 'only for the good'; it can also bring problems.





---

## Containing Network Overload

---

Network designers like to think that their traffic models are ideally stable, and that their forecasts of future traffic will never be wrong. In real life, however, such assumptions are unwise, because any one of a number of problems may arise, resulting in network overload, and so congestion. The forecast may underestimate demand; there may be a short period of extraordinarily high pressure (for example at New Year, Christmas, or any public holiday, or following a natural disaster); or there may be a network link or switch (exchange) failure. Monitoring and controlling the network to detect and avoid network congestion and the resulting service degradation is a difficult task, but an important one. This chapter discusses these new 'network management' methods in more detail.

### 37.1 THE EFFECT OF CONGESTION

Congestion in a telecommunications network manifests itself to customers who are attempting to make calls as a 'network busy' tone, or as a delay in computer or data network response. It annoys the customer; even worse, the congestion can rapidly increase as the result of customer or equipment *repeat attempts*. (*Repeat attempts* are further call attempts, made in vain by customers hoping to make a quick connection by immediate re-dialling.) These further call attempts only exacerbate the problem, because they greatly increase the loading on exchange equipment (e.g. switch processors, number stores, senders and receivers), and they add to the overall traffic volume. The congestion can then spiral out of control. The networks most at risk from congestion are those which have been kept to minimum circuit numbers in order to keep costs down. Also very much at risk are those networks that employ a high proportion of multi-link overflow routings, because congestion on one link in this type of network rapidly affects other routes, with a consequent 'domino effect').

As a result of practical studies, teletraffic experts have produced models which show that under overload conditions the effective throughput of a network can actually fall. This makes the overload worse, further reducing throughput, as Figure 37.1 demonstrates.

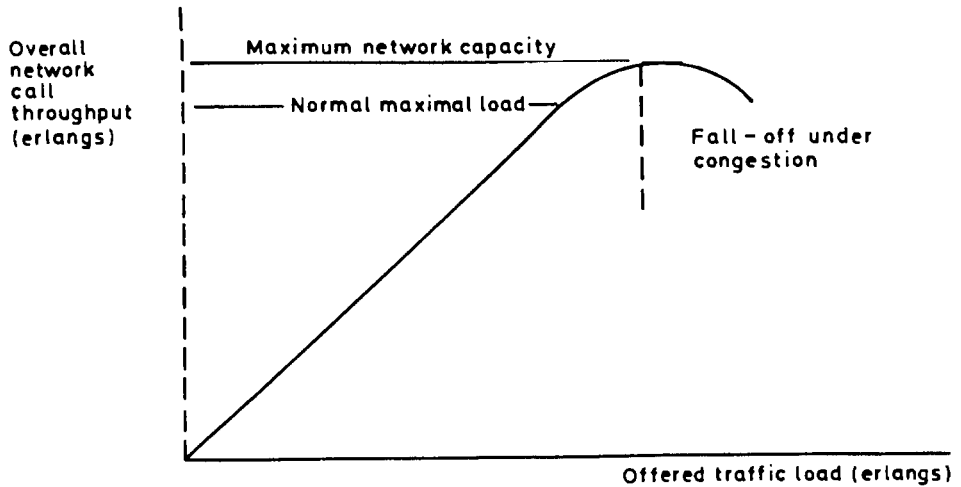


Figure 37.1 The effect of congestion

Spotting the onset of congestion and taking early appropriate action is crucial to maintaining control of the network. Only by careful control can customer annoyance and unnecessary call failure be minimized.

## 37.2 NETWORK MONITORING

The most common parameter used to measure network congestion in a circuit-switched network (such as a telex or telephony network) is the *traffic load* (usually measured in Erlangs, the number of circuits in use) over a period of time. In a packet network, the length of the delay queue (i.e. the number of packets awaiting transmission or the percentage trunk loading) can be used instead, as we saw in the formulae of Chapter 30. To serve the purposes of network management, these parameters need to be monitored frequently, so giving a human *network manager* a *real-time* perception of network performance. Other measurements, such as the *lost call count* (peg count) or machine overload information may also be available, helping quickly to diagnose congestion. However, in all cases, care must be taken when interpreting measurements taken over too short a period because short duration measurements or a very small *sample* of calls can be statistically deceptive. Traffic measurements are not reliable when taken over a time period much shorter than the average call holding time.

Traffic and network performance information nowadays usually comes from computers. Daily, weekly, and monthly usage records can be calculated by computer post-processing of information. The post-processing may only be made well after the day when the traffic itself was recorded from information stored on magnetic tape or disk. Such information is clearly of no advantage in recommending immediate alleviation action. The experience of many network operators is that network management monitoring information can be obtained only by setting up a completely distinct computer *network management system*.

The monitoring part of a network management system relies on the processing of *status* information (i.e. information about the current state of the network and its components), output in computer format by the switches or *stored program control* (SPC) exchanges. Commonly this information comprises *alarms* and other fault messages, but in addition may also include an itemized call record for each call made through the exchange. *Call records* are merely computer messages from the exchange to the network management system, giving all the details of the call (i.e. number dialled, start time, duration, whether successful or not, etc.).

From the collation of information relating to all current and very recent calls, a true picture of the current network state may be calculated by the network management system, and is displayed in some way to a human *network manager*. Most importantly, the load of traffic and the degree of congestion on each individual route into and out of the exchange can be shown.

If the network manager, in watching this information (or having it *alarmed* to his or her attention), notices that the congestion level on a particular route, link or exchange has exceeded a pre-determined threshold or that some other problem has arisen, he or she needs to evaluate the cause of the problem and take corrective action as fast as possible. A quick and correct diagnosis enables the manager to adopt the best available solution.

It is worth reflecting here that, as in road traffic jams and human diseases, the cause of a telecommunication network problem and its symptoms may not both be centred on the same place or region. Causes may lurk a long way from the effect. For illustration, consider the example shown in Figure 37.2.

As a result of failure on the link that connects the tandem exchange T with exchange C, there is growing congestion on the link between exchanges A and T. The congestion is caused by customers at exchange A persisting in vain attempts to make new calls to exchange C. The repeat attempts are resulting in unusually high traffic between exchanges A and T, thereby also affecting other customers at A wishing to connect to B. In a very sophisticated network monitoring scheme, the network manager aims to achieve an instant indication and diagnosis of major problems. This may be easy, for example, as the result of an *alarm*. In Figure 37.2, the T-C link failure may register an alarm, which will instantly tell the manager the cause of network congestion. In practice, however, real-time network monitoring facilities may only be available at a limited

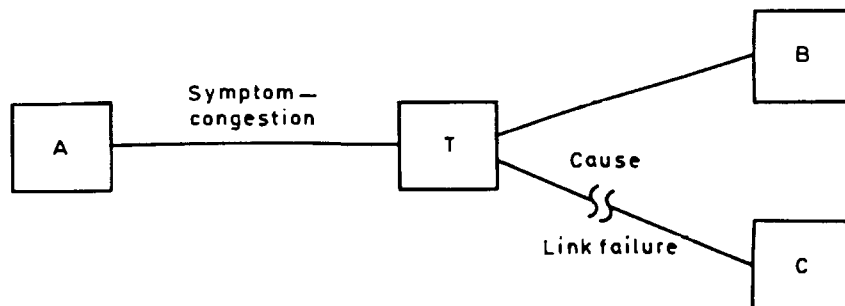


Figure 37.2 Example of symptom and cause of congestion

number of nodes or within given sub-network areas corresponding to the data network technology of a particular supplier, so that the network manager may have only an incomplete knowledge of the network status as a whole. In this case instant diagnosis and resolution of problems will require some guesswork. If in Figure 37.2, exchange A is the only one with network management capabilities (perhaps because it is the only exchange that is modern enough), then the network manager associated with exchange A could not directly know that the link T–C had failed. In this case he has to rely more heavily on ‘gut feel’ and experience. So what other tools are available to help diagnosis?

The monitoring of a second parameter comes in useful as an aid to diagnosis. The second parameter used is either the *answer/bid ratio (ABR)* or the *answer/seizure ratio (ASR)*. They are similar. ABR is the percentage ratio of answered calls to a given destination, compared with the number of call attempts (*bids*) intended for that particular destination. ASR is nearly the same, but instead is calculated using the number of outgoing seizures from the originating exchange as the denominator. The only reason for using one value rather than the other might be that it is this value that the particular exchange normally measures.

For network management purposes either ABR or ASR should be measured at each originating exchange on a destination basis. Let us assume for the remainder of this chapter that we have chosen to measure the ABR by destination as well as each route’s traffic intensity in Erlangs. Therefore, at exchange A in Figure 37.2 two ABR values are monitored, corresponding to destinations B and C, and the route traffic A–T is also measured. The ABR value is performed at A as follows

$$\text{ABR for destination C (measurement at A)} = \frac{\begin{array}{l} \text{number of answered calls to destination C} \\ \text{originated at exchange A during time } T \end{array}}{\begin{array}{l} \text{total number of calls to destination C} \\ \text{originated at exchange A during same time } T \end{array}}$$

In the example of Figure 37.2 the network manager at A, on noticing that the Erlang threshold has been exceeded on route A–T will also notice that the ABR from A to C has also significantly deteriorated. Meanwhile, calls are still completing from A to exchange B. The network manager at A will not know the root cause of the failure, but this is not necessary. The manager has two courses of action to choose from

- call personnel at exchange C to request further information on the problem, and thus receive advice on which network controls to apply
- go ahead and implement unilateral network management controls on the basis of the diagnosis to hand

Which course of action and the exact controls used by the manager will depend on the individual circumstance.

We have now seen that the monitoring of ABR values is invaluable in diagnosing congestion, but it need not be restricted to a follow-on action which is considered only after the traffic intensity threshold has been exceeded. There are additional benefits to be gained by monitoring destination ABR values continuously and comparing figures against *threshold*, or *action required* values. Figure 37.2 is again useful in illustration. Consider the case when a failure of link T–C occurs at night or some other period when

demand is so low that, because of the low calling rate on the network as a whole, there is no resulting traffic congestion on link A–T. The traffic threshold value in this instance would not be exceeded; and if the network manager were not additionally monitoring ABR on destination C (the value will have dropped to zero completion rate), the problem in completing calls to C might pass unnoticed, and no corrective action could be taken.

In data networks it is nowadays common for *traffic policing* algorithms to regulate the influx of traffic (i.e. packets, frames or cells) from a given source. Where a given source exceeds the *committed information rate (CIR)*, a frame rate contracted to be ‘guaranteed’ by the network at the time of set-up of the connection, then *excess* frames or cells may either be *marked* to be the first to be thrown away when congestion is encountered or may be directly discarded. This is achieved using special signalling techniques. *Forward explicit congestion notification (FECN)* and *backward explicit congestion notification (BECN)* messages are tagged to actual user message frames, to alert the user end equipments and the switches to which they are connected that congestion has been encountered somewhere along the path. These messages allow alleviation action to be triggered. Equivalent messages also exist in other data protocols (e.g. ATM) but may have other names.

Automatic congestion notification procedures built into data networks may avoid the need, at least to some extent, for special monitoring and congestion alleviation measures to be taken by human network managers of data networks. Further measures available to human network managers include taking individual *virtual connections* out of service or re-scheduling particular computer devices run their programs to run at a later (quieter) point in time.

The simplest method of detecting network problems is to monitor all available alarms. Alarms may simply activate bells and lamps, but a sophisticated *network management system* can have equipment *alarm* presentation built in, and it can be capable of some automatic problem diagnosis. As with other network status information, alarm information, along with ABRs, etc., is most effectively presented to the network manager as a combination of graphical and tabular displays on video terminals. These allow the manager at a glance to gauge overall network status, make a rapid diagnosis of problems, and take corrective or fault-containing action in good time.

The collation of all alarms by a central *network management system* (sometimes also called an *umbrella network management system*) has become the standard means of network monitoring and control. Specialist network management *platforms* (suites of computer hardware and software) have been developed specifically for this purpose, as we saw in Chapter 27.

### 37.3 NETWORK MANAGEMENT CONTROLS

With the problem located, what sort of network management control will be appropriate? All network management actions can be classified into one of two categories

- *expansive* control actions
- *restrictive* control actions

The correct action to be taken in any individual circumstance needs to be considered in the light of a set of guiding principles.

- Use all available equipment to complete calls or deliver data packets, frames or cells.
- Give priority to calls or packets, most likely to complete.
- Prevent exchange (nodal) congestion and its spread.
- Give priority to calls or data packets that can be completed using only a small number of links.

In an *expansive action* the network manager makes further resources or capacity available to alleviate the congestion, whereas in a *restrictive action*, having decided that there are insufficient resources within the network as a whole to cope with the demand, the manager can simply fail any new calls coming in to the *hard to reach* (i.e. temporarily congested) destination(s). It makes good sense to fail these calls close to their point of origin, because early rejection of calls frees as many network resources as possible, which can then be put to good use in completing calls between unaffected points of the network.

### 37.4 EXPANSIVE CONTROL ACTIONS

There are many examples of expansive actions. Perhaps the two most worthy of note are

- *network restoration*
- *temporary alternative re-routing (TAR)*

*Network restoration*, by no means a new technique, has been carried out by many network operators for years; but for some of them the introduction of computerized network management centres has provided an opportunity to apply centrally controlled network restoration for the first time.

*Network restoration* is made possible by providing more plant in the network than the normal traffic load requires. During times of failure this 'spare' or *restoration* plant is used to 'stand-in' for the faulty equipment, for example, a failed cable or transmission system, or even an entire failed exchange. By restoring service with spare equipment, the faulty line system or exchange can be removed from service and repaired more easily.

Network restoration techniques have historically been applied to transmission links, where it has been usual to plan spare capacity into the network in the form of analogue spare groups, supergroups, hypergroups or as digital bit-carrying capacity. The spare capacity is built into new transmission systems on a *1 for N* basis; i.e. 1 unit of restoration capacity for every *N* traffic-carrying units.

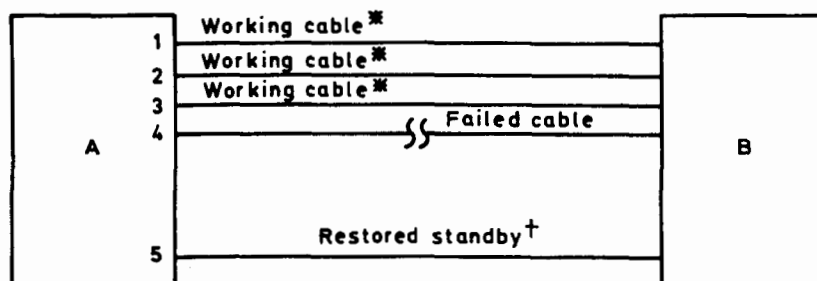
The following example shows how 1 for *N* restoration works. Between two points of a network, A and B, a number of transmission systems are required to carry the traffic. These are to be provided in accordance with a 1-in-4 restoration scheme. One example

of how this could be met is with five systems, operated as four fully loaded transmission line plus a separate spare (which should be kept warm; inactive plant tends not to work when called into action). Automatic *changeover* equipment can be used to effect instant restoration of any of the other cables, should they fail. An alternative but equally valid 1-in-4 configuration is to load each of the five cables at four-fifths 'stand-in', should any of the other four fail (Figure 37.3(a)). In the latter case, the failed circuits must be restored in four parts, each of the other cables taking a quarter of the failed portion (as shown in Figure 37.3(b)).

In practice, not all cables are of the same capacity and it is not always practicable or economic to restore cables on a one-for-one basis as shown in Figure 37.3(a). Both methods prevail. Another common practice used for restoration is that of 'triangulation' (concatenating a number of restoration links via third points to enable full restoration). Figure 37.4 illustrates the principle of triangulation. In the simple example shown, a cable exists from exchange A to exchange B, but there is no direct restoration path. Restoration is provided instead by plant which is made available in the triangle of links A-C and C-B. These restoration links are also used individually to restore simpler cable failures, i.e. on the one-link connections such as A-C or B-C.

Because of the scope for triangulation, *restoration networks* (also called *protection networks*) are often designed on a network-wide basis. This enables overall network costs to be minimized without seriously affecting their resilience to problems.

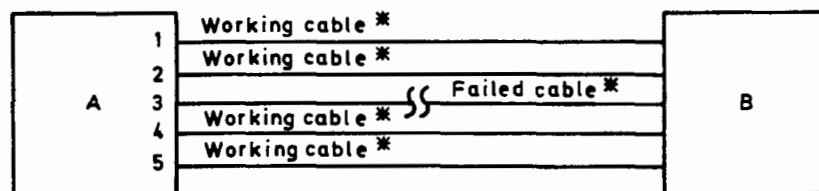
(a) Entire spare cable under normal condition



\* Each working cable carries its normal load

† Cable 5 'stands -in' to carry the entire load of cable 4  
Normally cable 5 is idle

(b) Spread load under normal condition



\* Each cable normally only 80 per cent loaded

On failure of any one cable, four 20 per cent components  
are spread over the other four cables

Figure 37.3 Two methods of 1 in 4 restoration



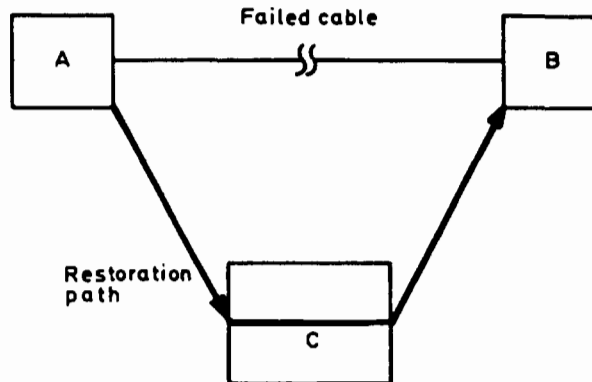
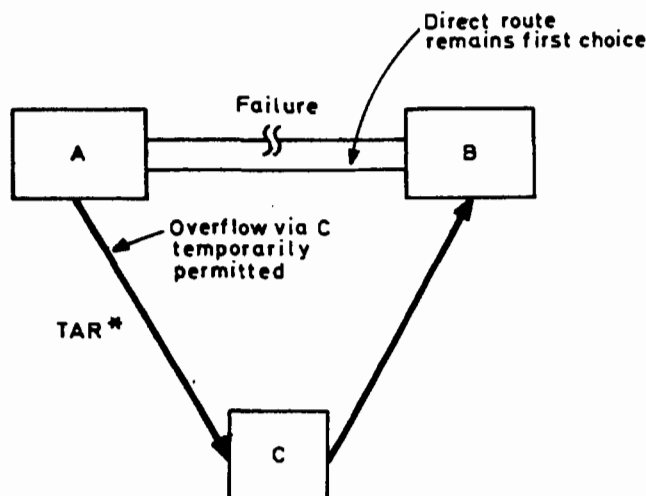


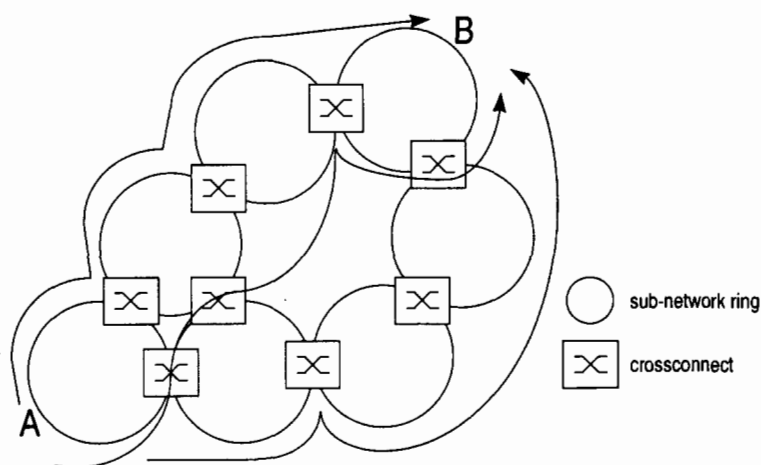
Figure 37.4 Restoration by triangulation

The use of idle capacity via third points is the basis of a second expansive action, termed *temporary alternative re-routing (TAR)*. This method is generally invoked only from computer-controlled switches where routing data changes can be made easily. It involves either routing calls to a particular destination via temporarily different routes, or allowing more route-overflow options than normal. In Figure 37.5, some direct capacity between A and B has failed, resulting in congestion. On noticing this, the network manager has also observed that the routes A–C and C–B are currently lightly loaded (again by chance this traffic has a different busy hour from that of A–B). A temporary overflow route via C would help to relieve congestion from A to B. The manager can adapt the normal routing choice list (used at exchange A) to include a temporary overflow via C. This expedient route path is termed a *temporary alternative route*, or TAR.



\*Transit traffic from A to B via C not normally allowed

Figure 37.5 Temporary alternative re-routing



**Figure 37.6** Alternative paths A–B in an SDH network made up of sub-network rings

In modern transmission technology (SDH and SONET as discussed in Chapter 13), *restoration* capabilities are built-in. Thus in both SDH and SONET it is intended that highly resilient transmission networks should be built up from inter-meshed ring sub-networks. The ring topology alone leads to the possibility of alternative routing around the surviving ring arc, should one side of the ring become broken due to a link failure. Crossconnect points between ring sub-networks further ensure a multitude of alternative paths through larger networks, as is clear from Figure 37.6. The possibilities are limited only by the capabilities of the network planner to dimension the network and topology appropriately and the ability of the *network management system* to execute the necessary path changes at times when individual links fail or come back into service.

In data networks (e.g. packet, frame and cell-switched networks) alternative routing is usually automatically inherent within the normal routing algorithms of the network, and is undertaken automatically by affected switches, sometimes in conjunction with the *network management system*.

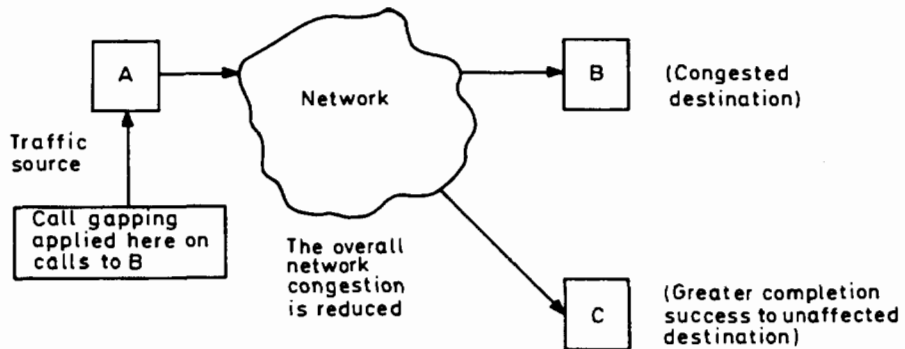
Where *temporary alternative routing* can be undertaken at the data or voice switching level, it may make sense to include all available transmission capacity into the switched network rather than keep some capacity free for transmission level *restoration* (e.g. using SDH). This saves the need for additional *restoration switching* equipment at the transmission level and in addition will afford much better network performance to customers when no links are out of service, because more bandwidth will normally be available.

## 37.5 RESTRICTIVE CONTROL ACTIONS

Unfortunately, there will always be some condition under which no further *expansive action* is possible (in Figure 37.5, routes A–C and C–B may already be busy with their

own direct traffic, or may not be large enough for the extra demand imposed by A-B traffic). In this state, congestion cannot be alleviated. Even worse, all calls made in vain to the problematic destination will be a nuisance to other callers, because these fruitless calls lead to congestion of traffic attempting to reach other destination. In this case, the best action is to refuse (or at least restrain) calls to the affected destination as near to their points of origination as possible. Perhaps the most flexible of restrictive actions is *call gapping* (in the case of a data network, the similar and equivalent action is called *pacing, flow control or ingress control*).

Under call gapping or flow control the traffic demand is 'diluted' at all originating exchanges. A restricted number of call attempts (or data frames) to the affected destination are allowed to pass from each originating exchange into the network as a whole. Within the wider network, this reduces the network overload, relieves congestion of traffic to other destinations, giving a better chance of completion. There are two principle sub-variants of *call gapping*, as shown in Figure 37.7. These are the *1-in-N* and *1-in-T* types. The *1-in-N* method allows every  $N$ th call to pass into the network. The remaining proportion of calls,  $(N - 1)/N$ , are failed immediately at their originating exchange (in this case, A). These callers hear network busy tone. The *1-in-T* method, by comparison, performs a similar call dilution by allowing only 1 call every  $T$  seconds to mature.



Call gapping type	Effect
1-in- $N$	Of $N$ calls generated at a particular exchange for the affected destination, only one is allowed to mature into a network call attempt. All others are failed at A regardless
1-in- $T$	Only 1 call attempt every $T$ seconds is permitted to mature beyond the originating exchange

Figure 37.7 Call gapping

In the first method, the value of  $N$  may be varied, to control the level of call dilution (for  $N = 2$ , 50% of calls would be allowed into the network; for  $N = 3$ , 33%, etc.). Similarly, in the  $I$ -in- $T$  method, the value  $T$  may be adjusted according to the level of congestion. The greater the value at which  $T$  or  $N$  are set at any particular instant in time, the greater the diluting effect on calls. In data flow control, any incoming frames over and above a rate allowed by a given threshold value are either marked for preferential discarding (i.e. will be thrown away before other frames should congestion be encountered) or may be immediately discarded.

If congestion continues to get worse, the value  $N$  or  $T$  can be increased accordingly. With a very large value of  $N$  or  $T$ , nearly all calls are blocked at the originating point. The action of complete blocking is quite radical, but nonetheless it is sometimes necessary. This measure may be appropriate following a public disaster (earthquake, riot, major fire, etc.). Frequently in these conditions, the public are given only one telephone number as a point of enquiry, and inevitably there is an instant flood of calls

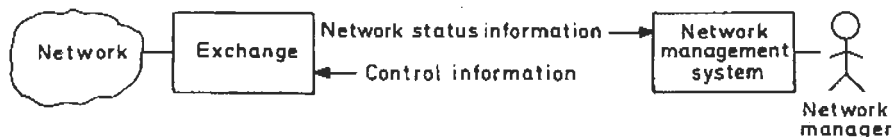


Figure 37.8 A network management system



**Figure 37.9** Network management centre. AT&T's Network Operations Centre at Bedminster in New Jersey, USA. It controls the AT&T worldwide intelligent network, the most advanced telecommunications network in the world. The centre controls data and voice calls over 2.3 billion circuit miles worldwide, handling more than 75 million calls a day. It is active 24 hours a day, 365 days a year. (Courtesy of AT&T)

to the number, few of which can be completed. In this instance, call gapping is a powerful tool for diluting calls, thereby increasing the likelihood of successful call completion for other network users.

## 37.6 NETWORK MANAGEMENT SYSTEMS

Today it is common for network management computer systems to be developed as an integral part of the modern switches (i.e. exchanges) or sub-networks which they control. Direct datalink connections between exchange processors or sub-networks and *network management systems* allow real-time network status information to be presented to the network management systems, and for traffic control signals to be returned. In this way, switch data changes or other network control methods can be made quickly. An example might be a switch routing data change to amend routing patterns by the addition of TARs. Another example of a control signal might be an indication to the exchange to perform call gapping on calls to an appropriate destination. Figure 37.8 illustrates the functional architecture of a typical *network management system*, but we discuss the subject in more depth in Chapter 27.

Network management systems can be procured either from computer manufacturers or computer software companies, but they are increasingly being offered by switch manufacturers as integral parts of new telephone or data exchanges. As the competition between suppliers becomes ever greater, perhaps we will have to wait only a few years before someone achieves a truly 'self-healing' network, one that remains congestion-free without the continuous assistance of human network managers.

---

## *Network Economy Measures*

---

People who plan telecommunications networks are always searching for equipment economy measures: reducing the cost of a network for a given information carrying capacity; or, conversely, increasing the information throughput of a fixed network resource. To achieve their aims, they can either maximize the electrical bandwidth available from a given physical transmission path, or (if it is the other kind of economy they want) they can reduce the amount of electrical bandwidth required to carry individual messages or connections. This chapter describes a few of the practical economy measures open to them.

### 38.1 COST MINIMIZATION

Reducing the cost of equipment required for a given information throughput is important for public and private network operators alike; both will be keen to reduce the quantity and the cost of lineplant and switch gear.

If a given resource, say a transmission link, is already laid on then there is not much to be gained by applying economy measures which have the sole effect of making some of the available capacity redundant. In such circumstances it may be advantageous to 'squeeze' extra capacity from the line, especially if it is nearing its limit. This can be valuable for one of three reasons; first, it enables expenditure on more capacity to be delayed; second, it may be the only practicable means; or third, the cost of duplication may be prohibitive. The first reason might postpone the need for a private network operator to lease more capacity from the PTO (public telecommunication operator). The second case might arise because of a need to make more telephone channels available from a limited radio bandwidth. The third might reflect a lack of resources to finance the prohibitive cost of a transatlantic undersea cable.

Earlier chapters in this book have covered one important means of lineplant economy, that of bandwidth multiplexing, by either the (analogue) *frequency division multiplex (FDM)* method, or the (digital) *time division multiplex (TDM)* method. This chapter briefly recapitulates these two methods, and goes on to describe some other

important techniques including *circuit multiplication equipment (CME)*, *statistical multiplexing* (used in data networks), *speech interpolation*, *low rate encoding (LRE)* and differential or adaptive differential PCM (DPCM and ADPCM).

## 38.2 FREQUENCY DIVISION MULTIPLEXING (FDM)

Frequency division multiplexing (FDM) provides a means of carrying more than one telecommunications channel over a single physical analogue bearer circuit, as Chapter 3 records. FDM relies on the carriage of a large electrical bandwidth over the circuit. Bandwidth for individual telecommunications channels is made available by subdivision of the overall bandwidth, much as some main roads are marked out into a number of lanes. Standard large bandwidths are employed over the physical circuit. These are called groups, supergroups, hypergroups, etc. They are normally exact integer multiples of a base unit of 4 kHz, which is the nominal bandwidth required for a single telephone circuit.

As we may recall from the example of Figure 38.1, a single four-wire circuit and a pair of *channel translating equipments (CTE)* enable us to derive 12 telephone channels between the end points A and B. This compares with the 12 individually wired telephone circuits which might otherwise be required.

In much the same way as 4 kHz bandwidths (individual telephone channels) can be multiplexed by *CTE* to form an *FDM group*, so FDM groups can be multiplexed by *group translating equipment (GTE)* to form *supergroups*, and *supergroups* can be formed into *hypergroups* by *STE*.

As we also recall (from Chapter 33), lineplant savings are possible by making connections out of any number of segments, each composed of a channel or circuit drawn from a number of different cables. This can save the need to lay a new direct cable. The connection from X to Y in Figure 38.2, for example, has been made using two FDM line systems, X-B and B'-Y without there being any direct wires between the two ends X and Y. Instead all actual wires converge on a single hub site at B. The use of concatenated (or *tandem*) FDM connections in this way should not be apparent to the

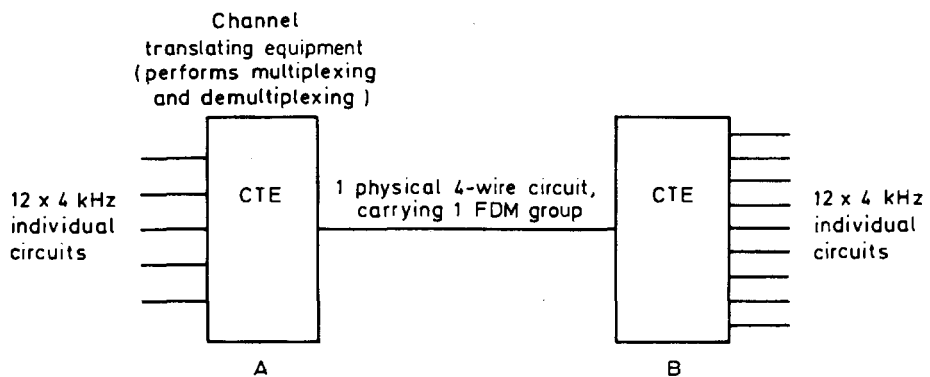


Figure 38.1 Saving lineplant using FDM

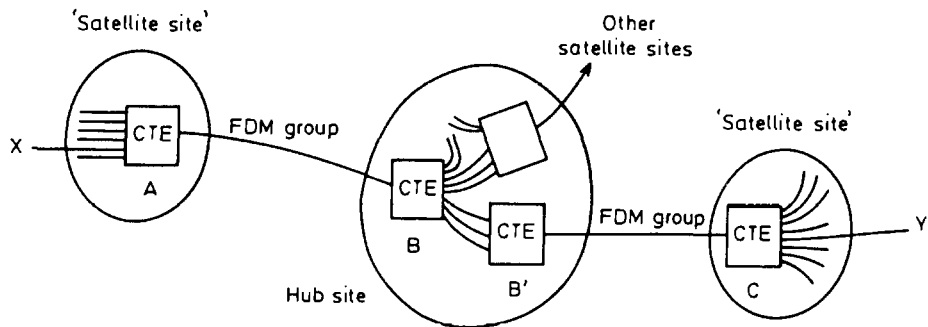


Figure 38.2 Tandem use of FDM systems

end user provided the interconnections are four-wire and not two-wire, as the full bandwidth is carried 'transparently' (i.e. without altering the signal significantly unlike the methods of compression we shall discuss later in the chapter). There is a slight degradation which results from the cumulative effect of repeated multiplexing and demultiplexing, so that the number of tandem sections should be minimized. So though a two-wire interconnection between tandem sections is possible, it is not recommended because of the nightmare combination of circuit stability and echo problems that it creates (see Chapter 33).

In the example shown in Figure 38.2, a connection between users X and Y passes through tandem FDM systems A–B and B'–C. Between A and B it is carried as part of the FDM group. At B it is demultiplexed and remultiplexed (at B') into another FDM group B'–C.

So that we may later put in context the relative economies of the techniques discussed later in the chapter, it is important to note that FDM systems work in the *four-wire* or *duplex* mode. By this we mean that simultaneous transmission in both directions is possible at all times. This means that X and Y in Figure 38.2 may talk at precisely the same time, and both messages may be conveyed simultaneously. This is possible because a permanent communication path exists in both directions at all times. The diagram of Figure 38.3 illustrates this in detail.

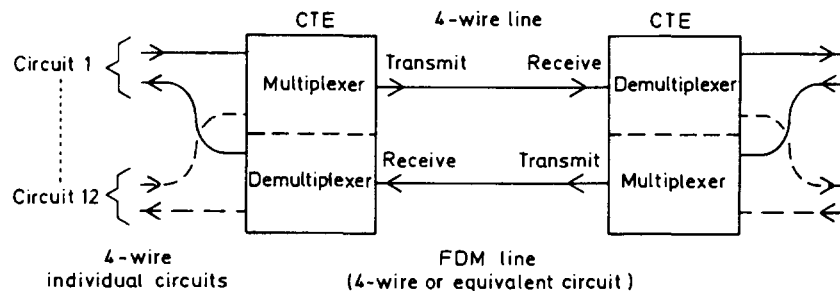


Figure 38.3 An FDM line system in detail



Another method of obtaining even greater lineplant economy using FDM is to allocate only 3 kHz bandwidth (as opposed to 4 kHz) for each individual speech channel. This was the method used on early transatlantic cables. The method has fallen out of use due to the quality impairments that results.

### 38.3 TIME DIVISION MULTIPLEXING

On digital lineplant the FDM technique is not normally applied. In rare cases, however, one of two devices, either a *codec* (coder/decoder) or a *trans-multiplexor*, is useful as a means of enabling an FDM group or supergroup to be carried on digital plant. The normal multiplexing method for digital signals on digital plant is called *time division multiplexing (TDM)*, as we learned in Chapter 5.

In TDM digital line systems the lowest bit speed (corresponding to a single telephone or data channel) is 64 kbit/s. Thus linesystems operate at bit speeds equal to, or an integer multiple of, 64 kbit/s.

Individual channels comprise a continuous series of eight-bit numbers (or octets) corresponding to the signal amplitude (or data signal value) sampled once every 125  $\mu$ s. The technique of *time division multiplexing (TDM)* combines channels together by interleaving octets taken from a number of channels in turn. As we recall, the European multiplexing hierarchy for TDM is 64 kbit/s–2 Mbit/s–8 Mbit/s–34 Mbit/s–40 Mbit/s, and the North American standard is 64 kbit/s–1.5 Mbit/s–6 Mbit/s–45 Mbit/s–140 Mbit/s. The equivalent of the CTE used in FDM is called a *primary multiplexor (PMUX)*, and is illustrated as a reminder in Figure 38.4.

The individual channels of a TDM bit stream are called *tributaries*. Where the tributary is an analogue signal, such as a speech circuit, the primary multiplexing equipment must first carry out analogue-to-digital speech encoding before multiplexing the 64 kbit/s tributaries into the 2 Mbit/s stream. The encoding method used for speech signals is called *pulse code modulation (PCM)*, and this was also described in Chapter 5.

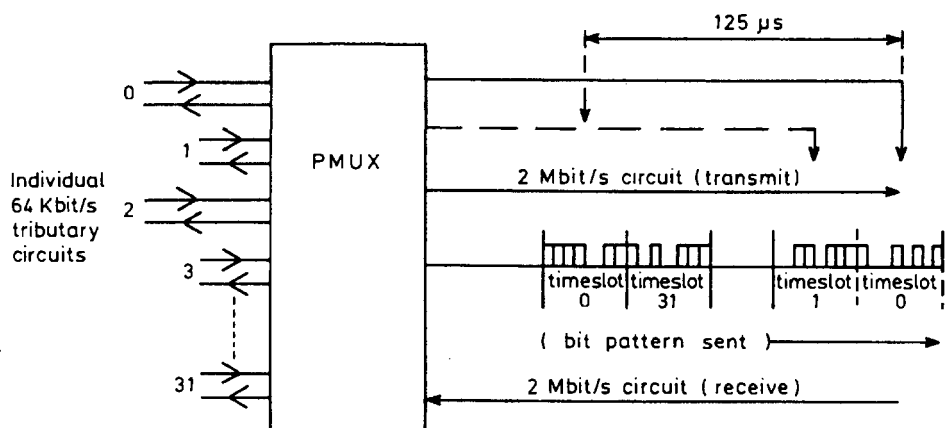


Figure 38.4 Time division multiplex (TDM)

In common with FDM signals, TDM bit streams operate in a duplex mode and require four-wire transmission. Also like FDM, individual TDM channels may be concatenated together without significant impairment of the end-to-end signal quality, because the 64 kbit/s or other bit stream is carried 'transparently'.

### 38.4 WAVELENGTH DIVISION MULTIPLEXING

*Wavelength division multiplexing (WDM)* extends still further the bit rate capacity of optical fibre. The technique relies on the sharing of a single fibre between a number of transmitting lasers or LEDs and receiving LEDs. The different transmitter/receiver pairs (e.g. at 1300 nm and 1500 nm) are able to share the fibre harmoniously merely by working at different light wavelengths. We discussed this technique in Chapter 8.

### 38.5 CIRCUIT MULTIPLICATION EQUIPMENT (CME)

*Circuit multiplication equipment (CME)* is a term used to describe various types of equipment capable of increasing the number of data or speech circuits that may be derived from a cable or a fixed bandwidth. The term, however, is not usually used to describe multiplexing equipment, such as that needed for FDM or TDM.

*Circuit multiplication equipments* tend to use 'corner cutting' methods to derive bandwidth economy, by one of a combination of the following practices

- statistical multiplexing or *interpolation* of individual channels
- bandwidth *compression* of analogue signals, or *low bit rate encoding* (LRE – of PCM encoded signals)
- data multiplexing
- data compression

Different types of circuit multiplication equipment are available, designed either for voice or data network use. Most voice network equipment is described simply as CME, but devices intended for use on data networks include *statistical multiplexors*, *data multiplexors* and *data compressors*. All these devices (data and voice) are similar in purpose, the main difference between them being the compression technique used. For example, *statistical multiplexors* employ the interpolation method of bandwidth economy, and voice CME may employ bandwidth compression as well.

### 38.6 SPEECH INTERPOLATION AND STATISTICAL MULTIPLEXING

We have noted that FDM and TDM transmission systems are designed to allow *duplex* operation (simultaneous transmission of speech or data in both directions). However, each direction of transmission is probably in use only about 40% of the time? In speech,

for example, one or other of the channels is nearly always idle, because both people seldom talk at once, and then probably because the listener wishes to interrupt the speaker. The overall utilization is certainly under 50%, but to make it worse the speaker leaves gaps between words and between sentences, so that the efficiency is unlikely to top 30–40%. Multiplying the effect, on a total of 60 speech circuits we might expect between 18 and 24 channels to be in use in each direction at any one time! The distinction drawn here between *circuits* and *channels* is made on purpose. A circuit consists of a two channels, a *receive* and a *transmit*. There are 120 channels available in total, 60 in each direction (but only about 40 are in use) between 18 and 24 transmit channels and a similar number of receive channels. Of course there will be short periods when a larger number of channels may be in use, but this is so improbable that we conclude that 80 channels are effectively being wasted. There is considerable scope for economy!

Let us err on the safe side, and allocate 30 channels in each direction (60 in total). This is equivalent to 30 circuits, and can be carried by a single 2 Mbit/s digital line system, rather than the two line systems we would have required for the full sixty circuits. The difference is that we now need a '60 derived circuits on 30 bearers' circuit multiplication device, (60/30 CME) capable of squeezing the 60 conversations into the 30 available circuits (60 channels). The first method that we discuss for doing this uses *speech interpolation*, as illustrated in Figure 38.5.

At each end of the transmission link shown in Figure 38.5 a CME terminal equipment is provided. Between the two terminals are 30 speech circuits and a control circuit, each comprising a transmit and a receive channel. The two terminal equipments are normally of identical manufacturer type, and are commissioned and brought into service simultaneously with the 30 bearer circuits. Up to sixty *derived* circuits are connected to the exterior facing side of the CMEs for carriage over the link, but the maximum number of circuits that we may *derive* will depend on local traffic characteristics, as we shall see later.

*Interpolation* relies on interleaving short bursts of conversation from a number of different *derived* telephone calls onto a single *bearer* circuit. To work well, statistics require the CME to have a large number of bearer circuits available, and to carry a larger number of derived circuits. Thus, for example, during a short burst of

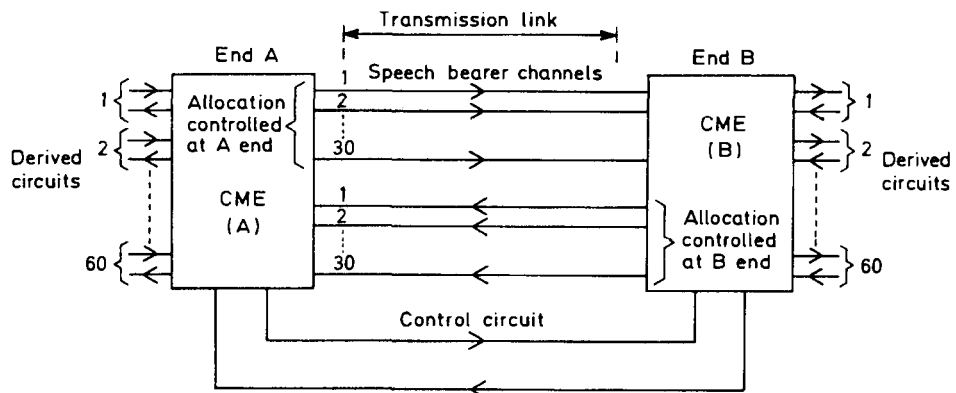


Figure 38.5 60/30 CME designed to use speech interpolation

conversation in the direction A–B, on derived circuit number 1 in Figure 38.5 the CME at the A-end can allocate the use of one of the outgoing bearer channels, say also number 1, to carry the burst. The burst might be as short as (or even shorter than) the curtailed phrase ‘I’ve got a . . .’. At the end of the burst, bearer channel 1 is made idle again. If by chance at this instant a burst of conversation commences on derived circuit number 2, then bearer channel number 1 must also be allocated to carry this subsequent burst. However, simultaneously, the A-end talker on derived circuit number 1 may start to talk again. In this instance, the A-end CME must again allocate a bearer circuit to carry the next burst, but because bearer number 1 is not now available it has to choose a different one, say number 17. This is all right provided that the B-end and CME similarly leaps around its incoming bearer channels to reconstruct the conversation.

Constantly, the A end CME will be allocating and freeing up the bearer channels, in the direction A–B, according to when the A-end person on each of the 60 derived channels is talking. Similarly the B end CME will allocate bearers in the direction B–A, for B-end speakers.

If you were to listen to any of the individual bearer channels, you would hear a train of incessant words and noises, which together would make no sense, as you would be hearing disjointed parts of up to 60 different conversations, as Figure 38.6 shows.

Thus, on bearer number one of Figure 38.6 we might hear the words: ‘I’ve got a’ . . . ‘would you please’ . . . ‘after tea’ . . . ‘thirty nine tons’ . . . ‘what about’ . . . ‘very well thanks’ . . . ‘six kilometres’ . . . ‘guarantees’ . . . ‘Agreed?’.

The conversations on the bearer channels are decoded by one of two methods. The first method, shown in Figure 38.5, uses a control channel between the two CMEs. This carries information such as ‘connect bearer number 1 to derived circuit number 1’; ‘connect bearer number 1 to derived circuit number 2’, etc. The second method is virtually the same, except that instead of using a dedicated control channel, each burst of speech is carried in a *packet*, the first end of which is coded with some extra control information that says which conversation it belongs to. The former method is common in CME designed for telephone use, whereas statistical multiplexors designed for data networks may use the second (packet switching) technique.

Both statistical multiplexors (for data networks) and CME (for voice networks) use interpolation as a method of lineplant economy. Both types of device only work best when the number of bearer channels is fairly large, and *gains* of 2 or 3 times are possible (i.e. 2–3 times as many *derived* circuits as bearers).

Gains of 2 or 3 times (i.e. 2–3 times as many *derived* circuits as *bearers*) are possible with both *statistical multiplexors* (for data networks) and *CME* (for voice application); hardware designs reflect this order of gain. However, although the hardware design of an equipment may suggest the use of a certain number of bearer and derived circuits, only the statistical characteristics of the real traffic can determine how many of each type should actually be connected. In practice few CMEs are wired to use all the bearer circuit and derived circuit ports simultaneously, and some bearer or derived circuit ports (or both) are usually unequipped. Thus the CME in Figure 38.5 could be used as a 54/30 device (1.8 gain) or a 60/24 device (2.5 gain) or a 48/24 device (2 gain). This provides a useful degree of freedom in network planning, but great care is needed in operation if the device is being run either at very high gain ratios (say above 2.5:1) or with a very limited number of bearers (say, less than 15). If, for example, you tried to run the device on a 12/6 basis, or alternatively at 5:1 gain (60/12), then severe quality

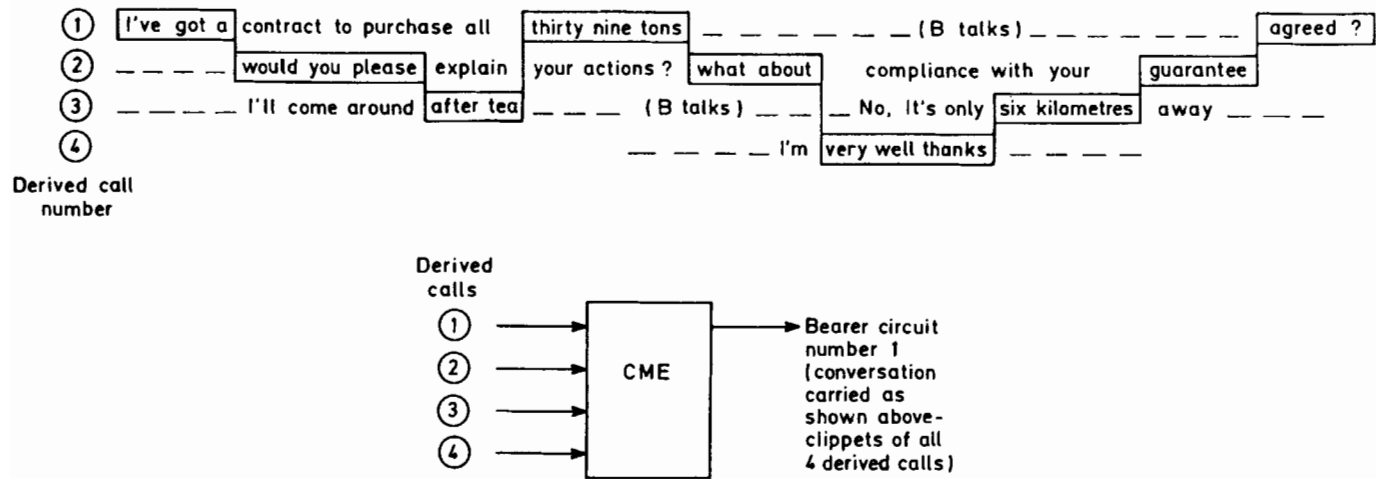


Figure 38.6 Speech interpolation

impairments would be likely on the conversations carried. These arise whenever there is no free bearer channel to carry a burst of conversation in a particular conversation. The section of conversation is lost completely, and the listener hears a rather broken-up message. This effect is known as *clipping*, or *freezeout*. Later on in the chapter methods for controlling it are described.

### 38.7 ANALOGUE BANDWIDTH COMPRESSION AND LOW RATE ENCODING OF PCM

If, prior to transmission, an analogue signal is passed through a special nonlinear electrical circuit to compress its bandwidth, then the saved bandwidth can be used to carry another signal, and an economy can be made. At the receiving end bandwidth expansion will be required to restore the original signal. This method of compression and expanding, or *companding*, although quite feasible, is not commonly used on analogue transmission systems as a means of lineplant economy because the benefits are small. The technique is, however, useful as a means of reducing noise interference on the received signal of analogue systems. It is effective as a means of reducing high frequency noise, as in the expansion stage some of the noise is shifted to a frequency above that audible to the human ear. The technique is the basis of the *Dolby* noise reduction system, well known amongst audio cassette tape users.

The real boom in the use of *bandwidth compression* has come with the development of special PCM *low rate encoding* (*LRE*) techniques to carry original analogue signals over digital lineplant. Despite the line bit speeds possible with high speed electronic components and optical fibre line systems, many research and development departments are working hard at digital signal bandwidth compression, and rates as low as 6 kbit/s are already quite feasible for carriage of speech (just imagine, 10 conversations down a single 64 kbit/s circuit). Similarly, 384 kbit/s video gives a very good quality for video-conferencing and even for low quality video (70 Mbit/s or so is required for broadcast standard television).

In the chapter on *pulse code modulation* (*PCM*), we discussed how an analogue signal could be converted into a digital bit pattern, by *sampling* the analogue signal at a high frequency to determine the amplitude of the signal. Corresponding to each sample, the amplitude is represented (*quantized*) as an integer number and transmitted as a string of binary digits, or bits. Figure 38.7 reminds us of the principle.

The 'normal' sampling frequency for speech is 8000 Hz, and the normal number of quantization levels is 256 (equivalent to 8 bits per *sample*). Thus the normal telephone bit rate, as we saw in Chapter 5, is  $8\text{ kHz} \times 8\text{ bits per sample}$ , which equals 64 kbit/s, and most digital telephone switching systems are based on this rate. However, if we can reduce either the sampling rate or the number of quantization levels without affecting too much the quality of speech, then we can make a direct saving in linespeed. Well, scientists have already found means of encoding speech at much lower rates, including 32 kbit/s, 16 kbit/s, 8 kbit/s and even 4.8 kbit/s. Indeed, provoked by the need to squeeze more channels out of limited radio bandwidth, 8 kbit/s is approximately the bit speed used to carry speech between the base stations and the mobile hand-held telephone units of modern digital cellular radio telephone systems.

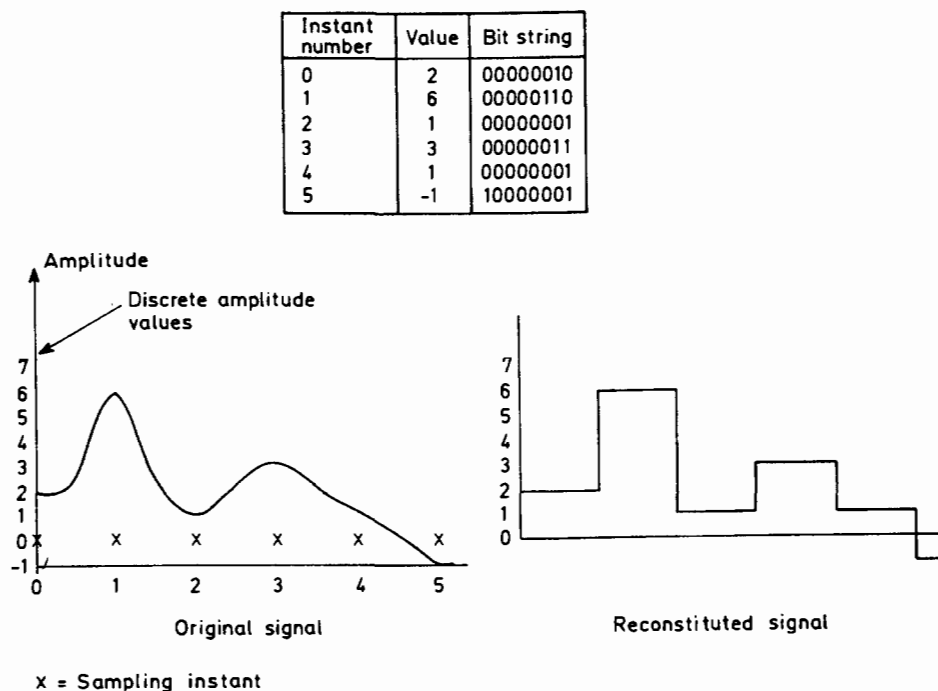


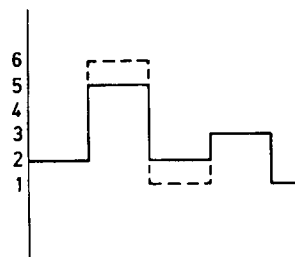
Figure 38.7 Pulse code modulation

For analogue signals (e.g. video or speech) the difference in *quantization values* between consecutive amplitude samples is usually small. So even though the amplitude value of a sample could theoretically take any one of 256 ( $2^8$ ) different values, in practice the sample amplitude rarely differs much from the previous one. *Differential PCM* (DPCM, a form of LRE) takes advantage of this fact by sending only the difference in amplitude between successive samples, rather than always sending the absolute value of each sample. The example of Figure 38.7 is repeated in Table 38.1, where the absolute sample values, and the differences between successive samples are shown.

Table 38.1 Differential PCM (DPCM)

Instant number	Simple amplitude value	Bit string (normal PCM)	Sample difference	Bit string (32 kbit DPCM)
0	2	00000010	—	—
1	6	00000110	+4	0100
2	1	00000001	-5	1101
3	3	00000011	+2	0010
4	1	00000001	-2	1010
5	-1	00000001	-2	1010

Instant number	Sample amplitude value	Actual sample difference	Bitstring (24 kbit/s D PCM)	Difference actually transmitted
0	2	—	—	—
1	6	+4	011	+3
2	1	-5	111	-3
3	3	+2	001	+1
4	1	-2	110	-2
5	-1	-2	110	-2



(a) Sample values

(b) Reconstituted signal

**Figure 38.8** Signal distortion resulting from DPCM

Notice in Table 38.1 how the difference in amplitude values of the consecutive samples can be expressed using only a four-bit string. This is equivalent to a 32 kbit linespeed, and is referred to as 32 kbit *DPCM*. As with normal PCM encoding, the first bit indicates positive or negative value (0 = positive, 1 = negative), the remaining three bits define the differential amplitude. DPCM rates other than 32 kbit/s can be used in our example or any other, but the risk is ever present that if too low a rate is used it may prove to be inadequately responsive to the fluctuation in signal amplitude. Figure 38.8, for instance, repeats our example, illustrating the difference values that would be transmitted using a 24 kbit/s DPCM device, using 3-bit difference encoding. Also illustrated in Figure 38.8 is the resulting reconstituted signal.

We clearly see that the 24 kbit/s encoding can only cope with a maximum difference between consecutive sample amplitudes of plus or minus 3. In the first part of the time period, therefore, where the signal is changing by more than this, the DPCM encoding responds as much as it can, but the lack of adequate responsiveness results in the signal distortion shown in Figure 38.8(b). However, towards the end of the time period, the reconstituted signal recovers to match the original, because the signal fluctuation has reduced.

All differential PCM devices are bound to cause short periods of signal distortion because at least some samples differ by an amount greater than the device is capable of matching, but infrequent distortion in this manner is not critical for certain types of signals. Provided that the bit rate of the DPCM device is high enough, the subjective quality of the reconstituted signal may be quite acceptable to the viewer or listener. The bit rate is too low if users begin to find the picture or sound quality either 'granular' or 'abrupt'. Normal practice is to ensure that a high proportion of users (say 99%) rate the standard of performance as either 'very good' or better.

### 38.7.1 Adaptive Differential PCM (ADPCM)

This works in a similar manner to differential PCM. However, by a further refinement on the DPCM technique it permits even lower bit speeds. In ADPCM, a *predictive* algorithm is used to estimate what the next sample value is likely to be. The actual value is then compared with the prediction, and only the bit pattern that describes the



difference of the two values is transmitted down the line. At the receiving end, the use of the same predictive algorithm together with the correction value received on the line allows the signal to be reconstituted. This technique gives very impressive economies in bit speed, provided that a suitable predictive algorithm is available for the type of signal to be sent. To date most research work has concentrated on speech and video based ADPCM algorithms, and in the field ADPCM devices have come to be more common than DPCM device because comparable signal output performance can be achieved with typically half the bit speed. ADPCM is an ITU-T defined algorithm (recommendations G.721 and G.726) and alternative bitrates are 40 kbit/s, 32 kbit/s, 24 kbit/s and 16 kbit/s.

ADPCM has the disadvantage that, although the reproduced speech quality is very good, the task of compressing and decompressing the signal is quite onerous, demanding powerful (and expensive) processing and causing a significant signal delay, which usually means that echo cancellation (Chapter 33) should also be used. A technique developed quickly after ADPCM, intended to reduce the processing load and delay, is *CELP* (*code excited linear prediction*). The general principles of *CELP* and *ADPCM* are similar in that only the difference between the actual signal amplitude and the predicted value is transmitted to line, but the algorithm used in *CELP* for the prediction is much simpler. A further development of *CELP*, *LD-CELP* (*low delay CELP*) is defined in ITU-T recommendation G.728.

## 38.8 DATA MULTIPLEXORS

A simple way of achieving lineplant economy in cases where a number of low speed data circuits are required between the same two end-points is by the use of data multiplexors. Usually these employ TDM techniques as already discussed, allowing a number of low speed applications to share the same high speed link, but to achieve lower gains than statistical multiplexors. Figure 38.9 illustrates an example network. Datamultiplexing without statistical multiplexing is becoming increasingly rare.

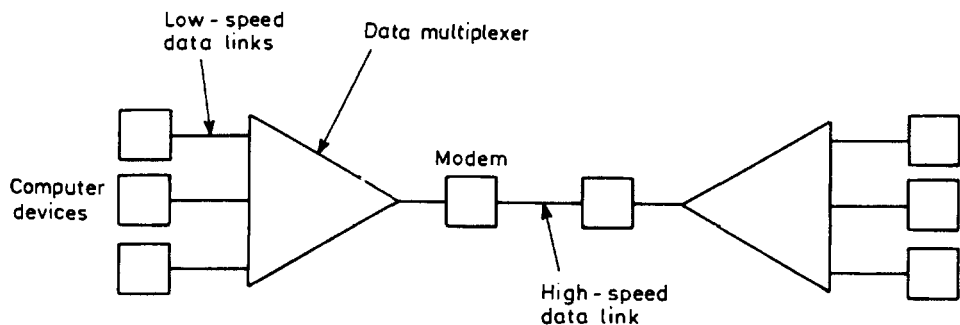


Figure 38.9 Use of data multiplexors

### 38.9 DATA COMPRESSION

A further way of economizing on data bit rate, and so lineplant, is made possible by the use of data compression techniques. These can be sub-divided into reversible and non-reversible categories. They allow a 33–75% saving in line speed.

The most important example of reversible data compression is called the *Huffman code*. This works by looking for the most commonly used data characters and using a shorthand code (say only 3 bits) to represent them, so enabling a saving of 5 bits from the standard 8-bit ASCII pattern. Unfortunately, to make the short codes for these characters possible, the less frequently used characters need to be coded with more than the normal 8 bits, but the relative frequency of use of characters means that, on average, around 6 bits per characters are needed. Irreversible data compression codes (often called *data compactors*) are less frequently used. These in essence corrupt the data signal by removing some of the detailed information content of the message which is adjudged by the transmitting device to be 'irrelevant'. Of course the information is irrecoverable at the receiving end.

### 38.10 PRACTICAL USES OF CME

Different manufacturers produce various types of CME, suitable for slightly different applications. A simple application is the use of a statistical multiplexor or data compressor within a private data network to economize on lineplant between buildings. A typical configuration might be that of a central mainframe computer located in one building with a number of users' terminals located remotely in other buildings. Figure 38.10 illustrates the use of statistical multiplexors.

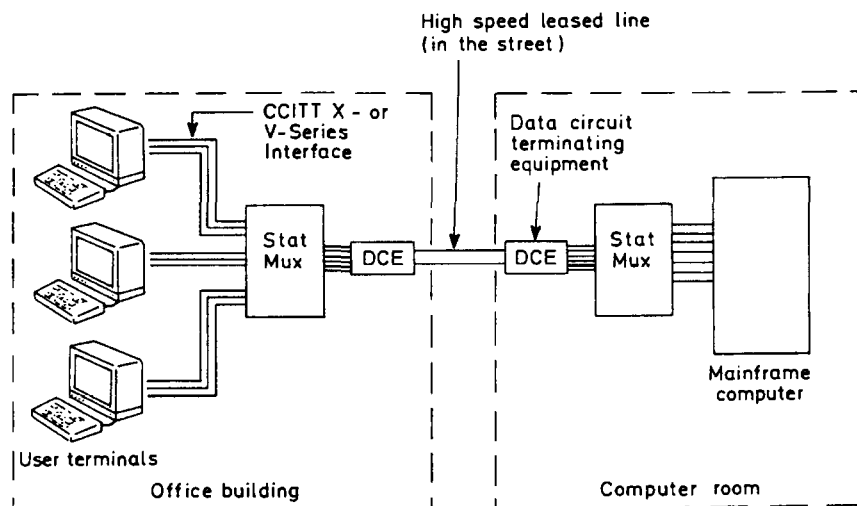


Figure 38.10 Typical use of statistical multiplexors

Within each remote users location, a number of terminals have been connected directly to a statistical multiplexor. This employs the interpolation method for allocating the available line bandwidth to the terminals which are active. When data are received from an active terminal they are formatted by the multiplexor into a standard packet format and given a label indicating which terminal they came from. They are then transmitted down the single high speed leased line, to the multiplexor at the mainframe computer site. (The *DCEs* (*data-circuit terminating equipment*) are merely the line terminating devices required to control the flow of data over the line itself). The second statistical multiplexor (which could be a built-in part of the computer) disassembles all the packets back on to individual circuits, appearing to the computer as if all the terminals were in the same room.

Messages are sent from the computer to the individual terminals in a similar way. A *buffer* (a sort of *queue* provided within the statistical multiplex) ensures that no data are lost even in the unlikely event that all the terminals try to communicate with the computer at once. If this should happen, some data are merely delayed and the user may notice a slightly reduced speed of computer response but this is rarely problematic, provided that the frequency of occurrence is low.

However, if the statistical multiplexor is overloaded with too many terminals, then the buffer may become full, so that the statistical multiplexor needs to 'tell' the terminals to stop sending data temporarily. In this case, the user notices a further reduction in the speed of response. Should the condition prevail for long periods it becomes frustrating to the user, and it should be alleviated by reducing the number of terminals connected to the statistical multiplexor or by increasing the speed of the trunk line between the multiplexors. Statistical multiplexing, as we saw in Chapter 9 and Section 3, is a 'cornerstone' of modern data network switching.

A larger scale example of CME use might be that by public telephone companies on an undersea cable. Such use has the advantage of increasing the call carrying capacity of a given cable many times over. Indeed, the *digital circuit multiplication equipment* (*DCME*) that was used as early as 1988 on TAT8 (the first transatlantic optical fibre cable) had the capability to enhance the call-carrying capacity up to five times, from around 8000 actual channels, up to about 40 000 derived ones. The DCME used a combination of 32 kbit/s ADPCM (giving 2:1 gain) combined with speech interpolation (giving a further 2.5:1 gain). Figure 38.11 shows the configuration in which the devices are used.

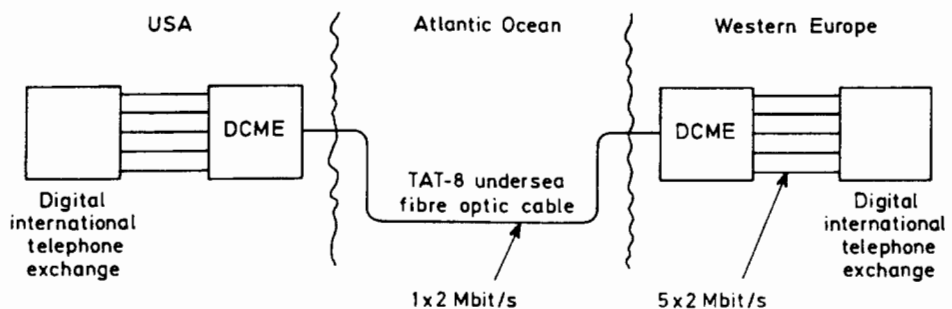


Figure 38.11 Use of DCME on TAT-8

Five 2 Mbit/s digital line systems from the international telephone exchange (or exchanges) are connected to the *derived channel* side of the DCME. Only one 2 Mbit/s line system of the transatlantic fibre system (or other transmission system) is connected to the *bearer* side. On this bearer, timeslot 0 is used in the normal way for synchronization and linesystem control, and timeslot 16 is used for the circuit allocation control, leaving 30 individual bearer circuits. Meanwhile, on each of the derived 2 Mbit/s systems timeslot 0 is used in the normal manner for synchronization and linesystem control, and timeslots 1–15 and 17–31 are used as 30 derived speech circuits, so giving up to  $5 \times 30 = 150$  derived circuits in total. Timeslot 16 of each desired 2 Mbit/s line system is used to communicate control messages between the telephone exchange itself and the DCME. These messages are necessary to prevent certain degradation conditions. For example, the acceptance of new calls can be suspended during periods of *freezeout* (as described earlier in the chapter) or during periods of bearer transmission failure. The DCME alerts these conditions to the telephone exchange, which responds by diverting new calls to alternative paths. A suitable control procedure for the purpose is specified in ITU-T recommendation Q.50.

### 38.11 CONSTRAINTS ON THE USE OF CME

The economies made possible by CME are to some extent achieved by ‘cheating’, and care is needed in their use. Three constraints need to be kept permanently under review, because the evolution of networks (both in topology and demand terms) can lead to serious impairments caused by the use of CME. The constraints are that

- CME-derived circuits should only be used for ‘approved’ applications
- CME gain ratios should be kept low enough to prevent freezeout, or suitable control measures should be implemented
- circuits comprising tandem links which are all individually circuit multiplexed should be avoided as far as possible

We explain the reasons in turn. First, particular types of CME are adapted for particular service types. Thus a statistical multiplexor is suitable for data but not voice, and a telephone network DCME is optimized for voice encoded PCM but completely corrupts a digital data signal. These are clear-cut cases, but less obvious mismatches can creep in to irritate users. For example, the DCMEs used in public telephone networks must not only perform to a high standard for speech calls, but must also support dial-up voiceband data applications, such as facsimile transmissions. Fortunately, the PTOs have recognized that such data applications are degraded by 32 kbit/s low rate encoding, and have adapted their DCMEs to employ an alternative (40 kbit/s) algorithm on the occasion that a data call is detected. Meanwhile a 24 kbit/s algorithm is used on a simultaneous voice call to make the bandwidth available. (This capability is defined in ITU-T Recommendation G.721.) Less observant network operators might have implemented the DCME and still be receiving the complaints from their facsimile users!

The second constraint is that gain ratios should be kept low enough to avoid the adverse effects of freezeout. As we have seen, freezeout (undue contention for *bearer* circuits) can cause an unacceptable delay to users of data statistical multiplexors; it manifests itself in voice CME as unacceptably ‘clipped’ and abrupt speech. The gain ratio is reduced simply by increasing the number of available bearer circuits or by trimming back on derived circuit numbers.

Finally, a warning about the ill-effects of the tandem use of CME. In data applications, tandem links of statistical multiplexors can lead to significant transmission delays. Worse still, the use of tandem CME in voice networks can cause the signal to become unintelligible. All this may seem obvious, but it can be difficult to foresee some of the permutations that can arise, particularly in switched networks! As we saw in Chapter 33, each ADPCM encoding introduces about three *quantization distortion units* (*qdu*). Within an overall budget of 14 you cannot afford to be liberal, so be careful when designing networks like Figure 38.2 using CME!

---

## *Network Security Measures*

---

Improvements in, and expansions of, communications systems and networks have left many companies open to breaches in confidentiality, industrial espionage and abuse. Sometimes such breaches go unnoticed for long periods, and can have serious business or cost implications. Equally damaging can be the impact of simple mistakes, misinterpreted, or distorted information. Increased belief in the reliability of systems and the accuracy of information has brought great gains in efficiency, but blind belief suppresses the questions which might have confirmed the need for corrections. This chapter describes the various levels of information protection which may be provided by different types of telecommunications networks, and the corresponding risks. It goes on to make practical suggestions about how a company's protection needs could be assessed, and how different types of information can best be secured in transit

### **39.1 THE TRADE-OFF BETWEEN CONFIDENTIALITY AND INTERCONNECTIVITY**

The man who sold the first telephone must have been a brilliant salesman, for there was no-one for the first customer to talk to! On the other hand, what confidence the customer could have had that there were no eavesdroppers on his conversations! The simplicity of the message should be a warning to all: the more people on your network, the greater your risk.

As the number of connections on a network increases, users are subjected to

- the risk of interception, *tapping* or 'eavesdropping'
- greater uncertainty about who they are communicating with (have you reached the right telephone or not, which caller might be masquerading as someone else?)
- the risk of time-wasting mistakes (an incorrect access to a database or a misinterpretation of data may lead to the corruption or deletion of substantial amounts of data)

- the nuisance of disturbance (wrong number calls, unsolicited calls from salesmen; worse still; forced entry by computer hackers, or abuse of the network by third parties to gain free calls at your expense)

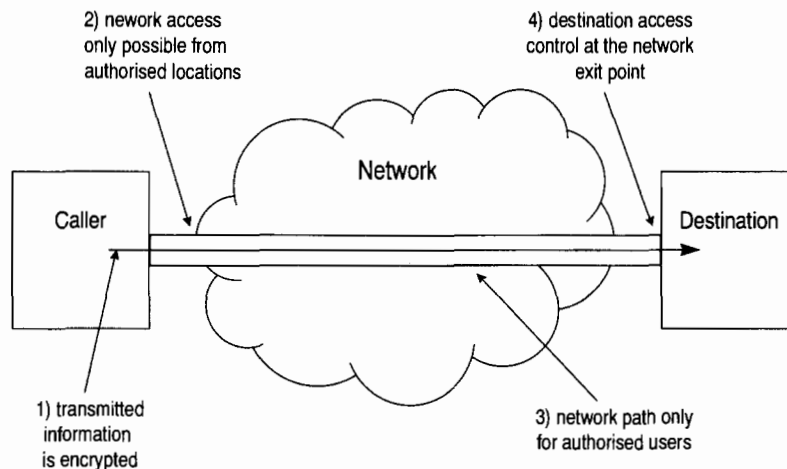
Too often, much thought goes into improving the connectivity of networks, but too little is applied to information protection. Risks creep in, often unnoticed. We discuss next the different types of protection which are available.

## 39.2 DIFFERENT TYPES OF PROTECTION

The information conveyed across communication networks may be protected from external distortion or abuse by any one of four basic means (Figure 39.1).

- encryption: coding of the information, so that only the desired sender and receiver of the information can understand it, and can tell if it has been distorted.
- network access control, allowing only authorized users to gain access to the communications network at its entry point.
- path protection, permitting only authorized users to use specific network paths.
- destination access control, allowing only authorized users to exit the network on a specific line, or to gain access to a specific user.

A combination of the four different protection methods will give the maximum overall security. Methods which are available in the individual categories set out below.



**Figure 39.1** Four aspects of communications security and protection

### 39.3 ENCRYPTION

*Encryption* (sometimes called *scrambling*) is available for the protection of both speech and data information. A *cypher* or electronic *algorithm* can be used to code the information in such a way that it appears to third parties like meaningless garbage. A combination of a known codeword (or combination of codewords) and a decoding formula are required at the receiving end to reconvert the message into something meaningful. The most sophisticated encryption devices were developed initially for military use. They continuously change the precise codewords and/or algorithms which are being used, and employ special means to detect possible disturbances and errors. One of the most secure methods was developed by the United States defence department, and it is known as *DES* (*defence encryption standard*).

To give the maximum protection, information encryption needs to be coded as near to the source and decoded as near to the destination as possible. There is nothing to compare with speaking a language which only you and your fellow communicator understand!

In a technical sense the earliest opportunity and best place for encryption is the caller's handset. Sometimes, either for technical or economic reasons, this point is not feasible and the encryption is first carried out deeper in a telecommunication network. Thus, for example, a whole site might be protected with only a few encryption devices on the outgoing lines rather than equipping each PBX extension separately. Clearly the risks are then higher.

For most commercial concerns I do not believe that the security risks arising from technical interception of signals within wide area networks are great. It is much simpler to overhear conversations on the train, read fax messages carelessly left on unattended fax machines or 'bug' someone's office than it is to intercept messages half-way across a network.

For maximum protection of data, the data themselves should always be stored in an encrypted form, and not just encrypted at times when they are to be carried across telecommunications networks. Permanent encryption of the data renders them in a meaningless or inaccessible form for even the most determined computer hacker. Thus, for example, encrypted confidential information held on an executive's laptop computer can be prevented from falling into unwanted hands, should the laptop go missing.

### 39.4 NETWORK ACCESS CONTROL

By controlling who has access to a network we minimize both intentional and unintentional disturbances to communication. In much the same way that we might reduce the road hold-ups, hazards and hijacks by limiting the number of cars, careless drivers and criminals on the road.

The simplest way of limiting network access is to restrict the number of network connections. Without a connection, a third party cannot access a network and cannot cause disturbance. The physical security of connections which do exist (i.e. lock and key) may also be important for very high security needs.



Entry to a network can be protected by password or equivalent software-based means. The simplest procedures require a user to 'log on' with a recognized username, and then further be able to provide a corresponding *authorization code* or *personal identification number (PIN)*.

The problem with simple password access control methods is that people determined to get in just keep trying different combinations until they stumble on a valid password. Aided by computers, the first *hackers* simply tried all the possible password combinations. The problem can be alleviated to some extent by limiting the number of attempts which may be made consecutively (bank cash teller machines, for example, typically retain the customer's card if he does not type in the correct authorization code within three attempts).

More secure password control systems require the user first to produce some sort of physical token (e.g. a key or a magnetic card). Without the key or card the system simply does not allow other potential intruders to start trying passwords. This method, for example, is used in modern cellular telephone networks, where a card (the *SIM card*) must be inserted into the phone to activate its potential network use. The SIM card identifies itself to a subscriber database within the network itself which holds information about authorized customers (we discussed this in Chapter 15). The SIM card itself must be activated each time the phone is switched on by the user typing in a PIN.

## 39.5 PATH PROTECTION

The communication path itself is bound to run through public places and in consequence past sources of potential eavesdropping, interception and disturbance. The best path protection depends on the right combination of physical and electrical telecommunication techniques, but from the serious eavesdropper there is no absolute protection. Encryption, as already discussed, prevents the eavesdropper from understanding what he might pick up. To reduce the risk of interception, the path should be kept as short as possible and not used if electrical disturbances are detected on it. There is nothing better than sitting in the same room!

In the early days of telephony, individual wires were used for individual calls and thus the physical paths for all callers were separate. Laying a separate cable continues to be a means of security for some. Some firms, for example, order their 'own' point-to-point leased lines from remote sites to their computer centre to ensure that only authorized callers can access their data. However, for the determined eavesdropper the physical separation may be an advantage; it is much easier to identify the right cable and *tap* into it at a manhole in the street. Alternatively, without tapping, he can surround a copper cable with a detection device to sense the electromagnetic signals passing along the cable, and interpret these for his own use.

Even glassfibre cable is not immune against eavesdropping. A glassfibre cable need not be cut at an intermediate point to insert a signal detector, it only needs to be bowed into a tight loop, whereupon some of the light signal emits through the fibre wall and can then be detected. Such procedures are now adopted in some optical fibre performance measurement and test equipment. The hacker need only put similar technology to criminal purpose.

Where radio is used as the communications path (you may not know this if you order a leased line from the telephone company), interception of the signal may be very straightforward. Overhearing of mobile telephone conversations, for example, has led to many a scandal in the press. Protection of radio (both from *radio interference* and from eavesdropping) can be achieved at least to some extent either by the use of proprietary modulation techniques or by new methods such as *frequency hopping*. In this method both transmitter and receiver jump in synchronism (every few fractions of a second) between different *carrier* frequencies. Jumping about like this reduces the possible chance of prolonged interference which may be present on a particular frequency, and makes it very difficult for eavesdroppers to catch much of a conversation.

Most modern telecommunications devices use *multiplexing* (FDM or TDM) to enable many different communications to coexist on the same physical cable at the same time. On the one hand this makes it harder to perform interception through *tapping* because the electrical signal carried by the wire has to be decomposed into its constituent parts before any sense can be made of a particular communication. On the other hand, it may mean that an electrically coded version of your information is available in the machine of someone you might like to keep it from. A message sent across a LAN, for example, may appear to go directly from one PC to another. In reality the message is broadcast to all PCs connected to the LAN and the LAN software is designed to ensure that only the intended recipient PC is activated to decode it.

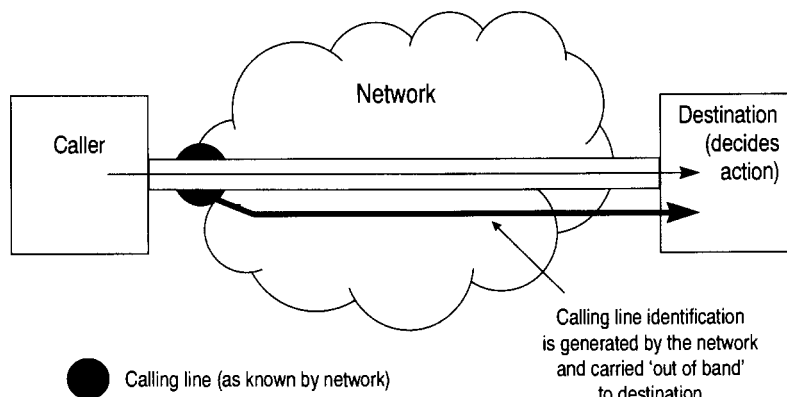
In practice, path protection across LANs and similar networks (including the *Internet*) is not possible. If such paths cannot be avoided by sensitive data transmissions then data encryption must be used. The lack of ability for such path protection has been a limiting factor in the acceptance of the Internet for transmission of sensitive commercial information. Much effort is now being focussed on improving security within the *Internet*. The techniques, however, largely rely on access control methods (e.g. *firewalls*) and key-coded encryption.

## 39.6 DESTINATION ACCESS CONTROL

Protection applied at the destination end is analogous to the keep of a medieval castle; having got past the other layers of protection, it is the last hope of preventing a raider from looting your prized possessions.

On highly interconnected access networks, destination protection may be the only feasible means available for securing data resources which must be shared and used by different groups of people. Typically, companies apply access control methods at a computer centre entry point. A much used protection method is a simple password authorization within the computer application software, but the level of security can be substantially improved by combining this with one of two types of feature which may be offered within the feeder network, either *calling line identity (CLI)* or *closed user group (CUG)*.

*Calling line identity (CLI)* is a feature available on telex networks, on X.25 *packet switched* data networks and on modern ISDN telephone networks. The network itself identifies the caller to the receiver, thus giving the receiver the opportunity to refuse the



**Figure 39.2** Calling line identity (CLI)

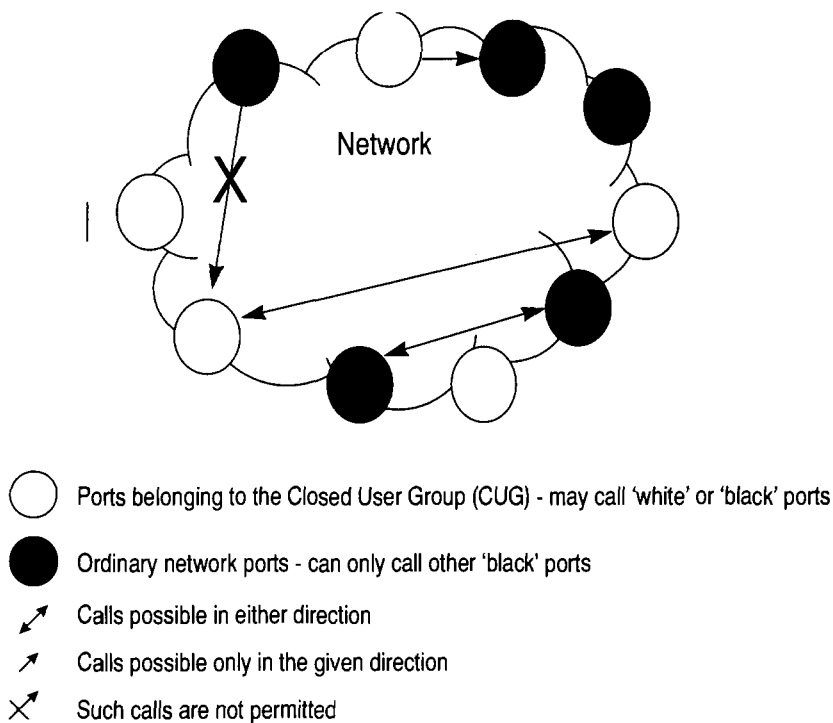
call if it is from an unauthorized calling location (see Figure 39.2). Call-in to a company's computer centre can thus be restricted to remote company locations. Password protection should additionally be applied as a safeguard against intruders in these sites.

Not all systems which might appear to offer the *calling line identity* are reliable. Fax machines, for example, often letterhead their messages with 'sent from' and 'sent to' telephone numbers. These are unreliable. They are only numbers which the machine owner has programmed in himself. It is thus very easy for the would-be criminal to masquerade under another telephone number (either as caller or as receiver) to send false information or obtain confidential papers. Even though you may have dialled a given telephone number correctly, you have no idea where you may have been automatically diverted to! The *XID* (*exchange identifier*) and *NUI* (*network user identifier*) procedures used in data networks are similarly insecure. They are in effect no more than passwords passed from the originating terminal to the network or destination terminal as a means of identification. They may be correct and adequate for most purposes but are easy to forge.

The *closed user group* (*CUG*) facility is common in data networks. To a given exit connection from the network for which a CUG has been defined, only pre-determined calling connections (as determined by the network itself) are permitted to make calls. Typically a small number of connections within a CUG are permitted to call one another. Additionally, they may be able to call users outside the CUG, but these general users will not be able to call back. In effect, communication to a member of the group is closed except for the other members of the group, hence the name. The principles of CUG are illustrated in Figure 39.3. CUG cannot be easily mimicked, as the information is generated by the network itself.

## 39.7 SPECIFIC TECHNICAL RISKS

What are the main technical risks leading to potential network abuse, breaches in confidentiality or simple corruption of information? What can be done to avoid them?



**Figure 39.3** The principle of closed user groups (CUGs)

## 39.8 CARELESSNESS

Always check addresses. I was once amazed to receive some UK government classified '*SECRET*' documents that should have been sent to one of my namesakes!

Why even think about encrypting a fax message between sending and receiving machines, if either machine is to be left unattended? Do not contemplate reading it on the train or talking about it on the bus.

Computer system passwords should be changed regularly. If possible, password software should be written so that it demands a regular change of password, does not allow users to use their own names, and does not allow any previously used passwords to be re-used.

Ex-employees should be denied access to computer systems and databanks by changing system passwords and by cancelling any personal user accounts.

Computer systems designed to restrict *write-access* to a limited number of authorized users are less liable to be corrupted by simple errors. Holding the company's entire customer records in a PC-based spreadsheet software leaves it very prone to unintentional corruption or deletion by occasional users of the data. Any changes to a database should first be confirmed by the user (e.g. 'update database with 25 new records? – Confirm or Cancel'). Subsequently, the system software should perform certain plausibility checks before the old data are replaced (e.g. can a person claiming social security really have been born in 1870?).

Ensuring proper and regular back-ups of computer data helps to guard against corruption or loss due to viruses, intruders, technical failures or simple mistakes. Daily or weekly back-ups should be archived *off-line*.

Simple precautions properly applied would dramatically reduce the risk of most commercial concerns!

### 39.9 CALL RECORDS

On very sensitive occasions, say when contemplating a company takeover, it may be important to a senior company executive that no-one should know he is even in contact with a particular company or adviser. Such company executives should be reminded of the increasing commonality of itemized call records from telephone companies, and similar call *logging* records which can be derived from in-house office telephone systems. Such devices keep a record of the telephone numbers called by each telephone line extension.

### 39.10 MIMICKED IDENTITY

Sometimes information can be gained under false pretences by claiming to be someone authorized to receive that information. Just as problematic and probably easier, false information could be fed into an organization or system to confuse or corrupt it. Virus softwares, for example, once into a computer can wreak almost unlimited damage.

Identity information which cannot be trusted should not be used (for example, the *sent from* or *sent to* telephone numbers which appear on fax messages. If the identity of a caller or destination should be *validated* using a technology which can be relied on to confirm addresses before being *authorized*.

The possibility of call diversion should not be forgotten. Modern telephone networks give householders, for example, the chance to divert calls to their holiday cottage while on vacation. They also provide an opportunity for criminals. A telex network *answer-back* confirms that the right destination has been reached, and similar *called line identity* can provide assurance on X.25, ISDN and other modern networks.

### 39.11 RADIO TRANSMISSION, LANS AND OTHER BROADCAST-TYPE MEDIA

Broadcast-type telecommunications media, although technically very reliable, are not well suited to high security applications. Diana Princess of Wales discovered to her cost just how easily analogue mobile telephones can be intercepted. However, other broadcast telecommunication media may not be so apparent to users; satellite, LANs and radio-sections of *leaselines* rented from the telephone company may also be security risk-prone.

Satellite transmission has proved to be one of the most reliable means of international telecommunication. Satellite media do not suffer the disturbances of cables by

fishing trawlers and by sharks and achieve near 100% availability over long periods of time. However, from a security standpoint, just about anyone can pick up a satellite signal. Thus satellite pay-TV channels need much more sophisticated coding equipment than do cable TV stations to prevent unauthorized viewing.

*Local area networks (LANs)* of interconnected PCs work by broadcasting information across themselves. So although LANs achieve a very high degree of connectivity (particularly those connected to the public *Internet* network), they could also present a security risk for sensitive information.

### 39.12 EMI (ELECTRO-MAGNETIC INTERFERENCE)

Electromagnetic interference has recently become a significant problem as the result of high power and high speed data communications devices (e.g. mobile telephones and office LAN systems). Although not usually of malicious origin, EMI can nonetheless lead to corruption of data information and general line degradation, particularly with intermittent and unpredictable errors.

The problem of EMI is recognized as being so acute that a range of international technical conformance standards has been developed which define the acceptable electromagnetic radiation of individual devices. In practical office communication terms, the most common problems are experienced with high speed data networks (e.g. LANs), particularly when the cabling has not been well designed. Simple precautions are

- the rigid separation of telecommunications and power cabling in office buildings
- the use of specified cable material only
- the rigid observance of specified maximum cable lengths

### 39.13 MESSAGE SWITCHING NETWORKS

Certain telecommunications networks (e.g. electronic mail networks, voicemail networks, some fax machines and fax networks and X.400 networks) carry whole messages in a *store-and-forward* fashion. The sender creates the message and posts it into the network, where it is stored in its entirety. The message subsequently progresses step-wise across the network as the availability of resources permit. Either the message will be automatically delivered to the user (e.g. fax) or it may wait for him to pick it up (e.g. electronic mail).

Message switching networks offer their users a higher level of confidence that messages will be delivered correctly and completely, and usually can give confirmation of receipt. At one level, modern message systems (e.g. electronic mail or voicemail) ensure that messages are read or heard by a manager himself rather than by his secretary. For very highly confidential information, users need to take into account the fact that a complete copy of the message is stored somewhere in the transmitting network.

‘Deletion’ of a message from your mailbox may prevent you as a user from further accessing a message, but should not be taken to imply that the information itself has

been obliterated from its storage place. A technical specialist with the right access may still be able to retrieve it.

Public telecommunication carriers in most countries are obliged by law to ensure absolute confidentiality of transmitted information and proper deletion once the transmission is completed successfully. Although this level of legal protection may be adequate for the confidentiality needs of most commercial concerns, for matters of national security it will not be.

Some modern fax machines (particularly those which offer 'broadcast' facility) also work by first storing electronically the information making up the fax. It may thus be possible for others to retrieve your message from the sending machine, even though you have removed the original paper copy.

### 39.14 OTHER TYPES OF NETWORK ABUSE

Finally, let us not forget that the most common motivation for network intrusion is the simple criminal desire to get something for nothing, perhaps telephone calls at your expense.

One of the easiest ways to create this opportunity for an outsider is to set up a network with both *dial-on* and *dial-off* capability. The scam works as follows...

Some companies provide a reverse-charge network *dial-on* capability to enable their executives to access their electronic mailboxes from home without expense. Some of these companies simultaneously offer a *dial-off* facility. Thus, for example, the London office of a company might call anywhere in the United States for domestic tariff, by first using a leased line to the company's New York office, and then 'dialling-off' into the local US telephone company.

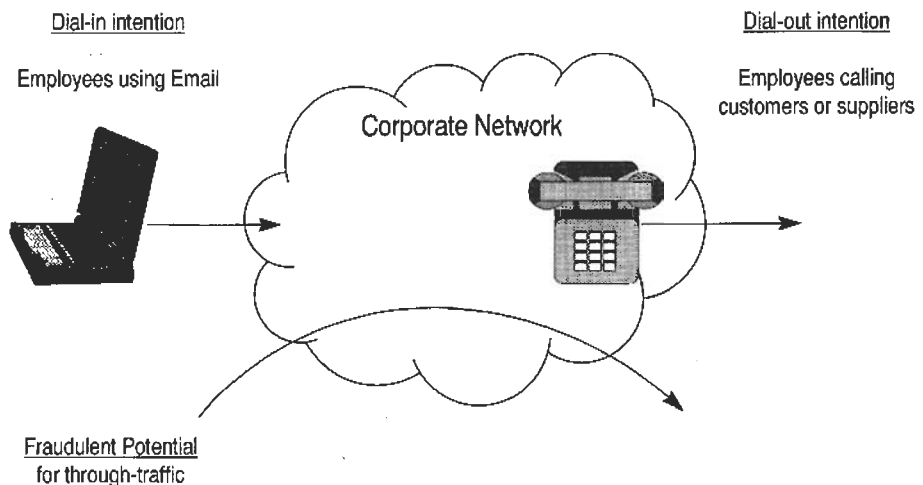


Figure 39.4 The risks of dial-on/dial-off

Now the criminal outsider can make all the calls he wants, entirely at company expense, unless the network is well enough designed to prevent simultaneous *dial-on* and *dial-off* by the same call (Figure 39.4).

Alternatively, *dial-back* can be used instead of *dial-on*. *Dial-back* similarly reverses the charges for the caller (other than the cost of the initial set-up call), but in addition enables the company to have greater confidence that only authorized callers (i.e. known telephone numbers) are originating calls.





---

## *Technical Standards for Networks*

---

In the past, telecommunications networks have been evolved in the minds of their designers to meet well defined but changing user demands. These broadly innovative influences are bound to persist and users can look forward to ever more sophisticated telecommunications services in the future. It was only in the mid-1960s that customer-dialled international telephone calls first became possible, and in those days such calls were for the rich alone. Nowadays the advance of technology has made them so much cheaper that everyone takes them for granted. In the late 1960s everyone marvelled at the first live satellite broadcasts; now there are so many satellites in space that many of them no longer have names, only longitudinal geographic location references. Every day the world's communicators move larger and larger volumes of video, text, speech, and data information around the globe, and all without a second thought. The revolution in communications that we have seen, and the ever-widening scope for information transfer, have only come about through an almost fanatical emphasis on the international compatibility of telecommunications networks and equipment, underpinned by worldwide agreement on the technical standards without which interconnection between networks in different countries and the interworking of different network types would be impossible. This chapter discusses the various bodies involved in setting these standards. It outlines in particular the recommendations of ITU (International Telecommunications Union), the world's most authoritative body on telecommunications technical standards.

### **40.1 THE NEED FOR STANDARDS**

Whenever two or more pieces of equipment, built by different people or different manufacturing companies, are called on to work harmoniously together within the same network, there is a need for technical standards. The *standard* defines comprehensively the interface, or set of interfaces to be used between various equipments. It describes the functions that each of the equipments are required to perform, and it ensures that the signals passed between the equipments are fit for their purpose, and unambiguous.

Only a network designed and built in isolation could exist without carefully defined and documented technical standards for the interfaces. Such networks are extremely

rare, for even when manufacturers develop their own special interfaces for interconnecting two pieces of their own equipment, these interfaces usually conform to their own documented proprietary standards. An example of a *proprietary standard* is the *systems network architecture (SNA)* (described in Chapter 18) which can be used as the basis for network design (called an architecture) interconnecting computer systems. SNA is similar to ITU's (or ISO's) OSI model, but was developed by the IBM corporation and was in use by their customers before the ITU standards had been agreed.

Most networks are from equipment supplied by many different manufacturers, who usually cooperate with the network operators and other interested parties to define the standards necessary to ensure correct interworking of equipment. The application or potential of any particular *standard* is usually the determining factor in how many parties need to be involved in developing and documenting it. Consequently many slightly different groupings of manufacturers, network operators, users, and regulatory bodies, both national and international, have sprung up over the years to address the development of various different sets of standards. Each grouping (often called a *standards body*) tends to specialize in the interfaces required between particular types of equipment (say computers, or facsimile machines, or *customer premises equipment (CPE)* or safety). The remainder of the chapter describes a number of the most prominent ones.

## 40.2 WORLDWIDE INTERNATIONAL STANDARDS ORGANIZATIONS

International network interconnections have always presented something of a challenge for the engineers developing them. Historically, international telecommunications between different countries have been provided by means of gateway switches, which have provided a means for interfacing and interworking a common international standard with the plethora of slightly differing national networks. International standards have thus provided the foundation of the recent boom in international communication. The boom is likely to continue so long as each national network evolves slowly towards the international standard, leading ultimately to a homogeneous worldwide network. However, this is only one of the benefits to be derived from international standards; no less important, with users of even the smallest networks insisting on conformance, terminal equipment manufacturers will be obliged to produce devices of worldwide compatibility.

The amount and quality of international standardization has improved rapidly in recent years. One of the practices which has improved standards is the definition not only of interfaces and protocols, but also of the procedures for *conformance* and *compatibility testing*. Such procedures are defined in standards called *PICS (protocol implementation conformance statements)* and *PIXITs (protocol information extra information for testing)*.

The most important worldwide standards organizations in telecommunications are the *International Organization for Standardization (ISO)* (colloquially the *International Standards Organization*) and the *International Telecommunications Union (ITU)*. A third body, but only important in the agreement of standards for satellite working, is the *International Telecommunications Satellite Organization*, called *INTELSAT*.

### *International Organization for Standardization (ISO)*

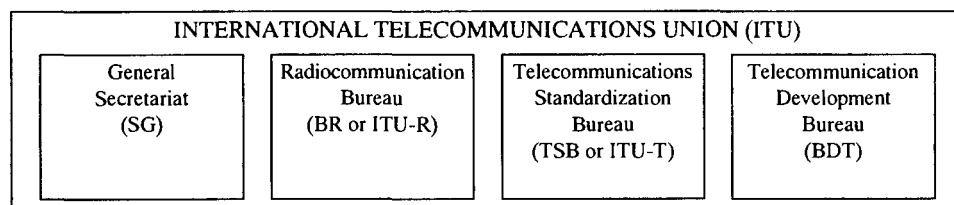
The ISO is a 'voluntary' organization, composed of and financed by the national standards organizations (e.g. *British Standards Institute (BSI)*, *American National Standards Institute (ANSI)*, etc.) of each of the member countries. Like each of its component national organizations, the ISO lays down standards for practically every conceivable item, not only telecommunications, but also the colour scheme for electrical wiring, and even the standard sizes for paper (e.g. A3, A4, etc.). The organization has a number of sub-committees, but in the main it does not produce standards from scratch. It tends merely to 'brush-up' standards submitted by its member organizations. In this way the *local area network (LAN)* standard ISO 8802 started life as IEEE802, generated initially by the *American Institution of Electrical and Electronic Engineers (IEEE)*.

#### *Address*

International Organization for Standardization  
Rue de Varembe, 1  
Case Postale 56  
CH-1211  
Geneve 20  
Switzerland  
Tel: (41) 22 749 0111

### *International Telecommunications Union (ITU)*

The ITU is an agency of the *United Nations*, responsible for overseeing all aspects of telecommunications. Organizationally, the ITU is split into four permanent parts, all of which are headquartered in Geneva. These are shown in Figure 40.1. The General Secretariat is responsible for overall administration, finance, and publication of regulations, journals and technical *recommendations* (their name for *standards*). The *Radio-communication Bureau (BR or ITU-R)* serves in two functions, first as a 'custodian of international public trust', regulating the assignment of radio frequencies throughout the world, and so preventing interference between radio *stations* (this responsibility used to be carried out by the ITU under the organization called *IFRB (International (radio) Frequency Registration Board)*). Second, the ITU-R acts as a *consultative committee* which generates technical standards and reports regarding radiocommunication (this function was formerly carried out by the predecessor body known as *CCIR (Consultative Committee for Radiocommunication)*). The *Telecommunications Standardization Bureau (TSB or ITU-T)* is the main *consultative committee* which generates



**Figure 40.1** Organization of the ITU

technical standards for telephone/telegraph networks (including data networks). The *ITU-T* (*ITU standardization sector*) replaces the former body, known as the *CCITT* (*International Telegraph and Telephone Consultative Committee*). The fourth part of the ITU, set up by the plenipotentiary meeting in June 1989 is the *BDT* (*Bureau for the Development of Telecommunications*). The BDT has the same status as the ITU-R and ITU-T in its aim is to secure 'technical cooperation' but in addition it has a prime role in raising help and finance to help the less-industrialized countries develop their telecommunications networks. More than 150 countries are members of the ITU and participate in the work of the consultative committees. Delegates to the individual *study group* meetings comprise not only the main public network operators, but also scientific and industrial organizations, private companies and equipment manufacturers. In the main, the study groups generate *recommendations* from scratch, based on the *contributions* of delegates.

ITU-T approves new *recommendations* at the *WTSC* (*World Telecommunications Standardization Conference*) which has replaced the four yearly *CCITT plenipotentiary meeting* which formally approved the coloured 'books' of technical standards.

*Address*

International Telecommunications Union

Place de Nations

CH-1211

Geneve 20

Switzerland

(Same address for ITU-R, ITU-T and BDT)

Tel: (41) 22 730 51 11

(Sales and Marketing Service)

Tel: (English): (41) 22 730 6141

Tel: (French): (41) 22 730 6142

Tel: (Spanish): (41) 22 730 6143

*International Electrotechnical Commission (IEC)*

Also known as the *Commission Electrotechnique Internationale (CEI)*, the IEC exists to promote international cooperation on electrical and electronic standards. It works in a similar manner to the ISO, gaining an international consensus of opinion and issuing standards in nine main fields, the most important to us being those affecting Telecommunications and Electronic components and those on Telecommunications equipment and information technology. The IEC has equal status and works in close cooperation with the ISO, whose prime interest is in the non-electrical fields.

*Address*

International Electrotechnical Commission

Rue de Varembe, 3

CH-1202

Geneve 20

Switzerland

Tel: (41) 22 919 02 11

### *INTELSAT*

The International Telecommunications Satellite organization is run by subscription between the world's main international public network operators. Its prime purpose is to develop, procure, and operate satellite services between countries. In support of these aims it develops its own standards specifically for the satellite service (e.g. from satellite earth station to satellite in space, or for *multiple access* operation (see Chapter 33)).

#### *Address*

INTELSAT  
3400 International Drive NW  
Washington DC 20008  
United States of America  
Tel: (1)-202-944-6800

## **40.3 REGIONAL AND NATIONAL STANDARDS ORGANIZATIONS**

A number of regional and national standards organizations are also in existence. In some cases these duplicate the international standards, and may even be at variance with them, a situation which is bound to persist, because it takes a long time to agree on an international standard. However, some users cannot wait but feel obliged to develop their own 'interim' standards, and these too tend to persist even after the emergence of the world standard, because of the investment tied up in them. Notable regional and national organizations involved in such standards are as follows.

### *Comité Européen de Normalisation (CEN)*

The CEN is the European equivalent of the ISO; constituted of member national standards bodies, it generates European Norms (ENs), but unlike ISO standards ENs are mandatory standards in the signatory countries to the CEN. *CENELEC*, a related organization, is the European equivalent of the IEC, generating European electro-technical standards.

#### *Addresses*

Comité Européen de Normalisation  
Rue de Stassart 36  
B-1050  
Brussels  
Belgium  
Tel: (32) 2 550 08 11

CENELEC  
Rue de Stassart 35  
B-1050  
Brussels  
Belgium  
Tel: (32) 2 519 68 71

### *European Conference for Posts and Telecommunications (CEPT)*

The Telecommunications Commission (T Com) of the CEPT was until 1988 prominent, in much the same way as the CCITT (the forerunner of the ITU-T), in the development of European telephone telegraph and data networks standards, and many CEPT recommendations have formed the basis of the CCITT (now ITU-T) recommendations. However, in 1988, prompted by the desires of the *European Union* (EU, then called the *European Economic Community*, EEC) to quickly develop the network interfaces needed for the establishment of a pan-European network infrastructure, the CEPT elected to hand over all its technical standards development work to a new permanently staffed organization called the *ETSI* (*European Telecommunications Standards Institute*). Nevertheless, the CEPT lives on, as an organization of European public telecommunications operators discussing mutual strategy and planning, including management of the radio spectrum in the European region.

#### *Address*

CEPT Information Desk  
European Telecommunications Office (ETO)  
Holsteinsgade 63  
DK-2100 Copenhagen  
Denmark  
Tel: (45) 35 43 25 52

### *European Telecommunications Standards Institute (ETSI)*

Funded in part by European Union money, the ETSI was founded near Nice in France in 1988 to develop the network interfaces and other technical standards necessary for a homogeneous, and pan-European telecommunications network. It has over 370 members from more than 30 countries in the CEPT and the European Free Trade Association (EFTA), as well as European Union states. The technical work is undertaken by full time staff, seconded from public network operators, manufacturers, trade associations, users, research bodies and national standards bodies. In addition, a number of special voluntary working parties have been established to deal with priority items.

The ETSI is composed of a technical assembly and a permanent secretariat, but the standards development work is carried out by working parties and project teams. Technical documents produced by ETSI may become one of three types of document, as listed below

- *EN* or *ENV* (*European Norm*) Standards; to become ENs or temporary ENs (ENVs), the standard needs to be put up for formal adoption by CEN/CENELEC
- *ETS* (*European Telecommunication Standards*) ETSs are voluntarily accepted standards and guidelines
- *NETs* (*Normes Européennes de Télécommunications*), these are the most important standards, because they are mandatory across Europe

In addition, ETSI also issues *ETRs* (*European Technical reports*) which provide useful technical information for the design of equipment and telecommunications networks, but these do not have the status of standards.

*Address*

European Telecommunications Standards Institute  
650 Route des Lucioles  
F-06921 Sophia Antipolis Cedex  
France  
Tel: (33) 4 92 94 42 00  
secretariat@etsi.fr  
Tel: (Publications office): (33) 92 94 42 41

*European Computer Manufacturers Association (ECMA)*

A consortium of European computer manufacturers, working on a voluntary basis to develop computer interconnection standards. The ECMA have often debated standards within Europe before contributing to the ITU. An example is the OSI transport protocol, ECMA-72 and ISO 8072.

*Address*

European Computer Manufacturers Association  
Rue du Rhone, 114  
CH-1204  
Geneve  
Switzerland  
Tel: (41) 22 849 6000

*American National Standards Institute (ANSI)*

ANSI is a voluntary national standards organization of the United States of America, and is a member of the ISO. It is the United States clearing house for standards, generally adopting standards proposed by smaller companies and organizations within the United States. Notable achievements of ANSI have been the computer character code *ASCII* (*American standard code for information interchange*) and the computer programming languages, *COBOL* and *FORTRAN*. ASCII was the forerunner to ITU-T's international alphabet number 5 (IA5, ITU-T recommendation T.50) which defines the alphabet in which computers 'talk' to one another.

*Address*

American National Standards Institute  
11 West 42nd Street  
New York  
NY 10036  
United States of America  
Tel: (1)-212-642-4900



*Electronic Industries Association (EIA)*

The EIA is a national trade association, developing electrical standards primarily for North America. Notable achievements of the EIA have been the RS-232 (forerunner to ITU-T V.24) and RS-449 (forerunner to ITU-T V.36) interfaces used between computer equipment.

*Address*

Electronic Industries Association  
2500 Wilson Boulevard  
Arlington  
Virginia VA22201  
United States of America  
Tel: (1)-703-907-7500

*Institute of Electrical and Electronics Engineers (IEEE)*

This is a professional institution based in the United States but for electrical engineers based anywhere in the world. It lays down professional and technical standards and codes of conduct. A notable achievement of IEEE was the IEEE *local area network* (LAN) standards in the 802 series, first published in 1983. These are now used throughout the world as the standards for *local area networks* (LANs). They have subsequently been adopted by the International Organization for Standardization as ISO 8802.

*Address*

Institution of Electrical and Electronic Engineers  
United Engineering Centre  
345 East 47th Street  
New York  
NY 10016  
United States of America  
Tel: (1)-212-705-7900

Publications from Customer Services

Tel: (1)-908-981-0060

*Exchange Carriers Standards Association (United States)*

Commonly called the T1-Committee, a leading body in agreeing standards for the deregulated networks of the United States. The *T1-committee* (formerly Exchange carriers standards association) is open to any subscribing members, but these are mainly United States exchange carriers. Standards under the auspice of T1 include most United States PTO (or *exchange carrier*) networking standards (e.g. T1-ISUP, the T1 version of the number 7 integrated services digital network user part signalling described in Chapter 12. T1-ISUP is used mainly in the United States). Another T1 standard is the *B-ICI* (*Broadband Inter-Carrier Interface*) used to interconnect ATM networks of different carriers. The standards work itself is carried out by a number of sub-committees called variously T1-D1, T1-X1, etc.

*Address*

Alliance for Telecommunications Industry Solutions  
(Standards Committee T1)  
1200 G Street NW  
Suite 500  
Washington DC 20005  
United States of America  
Tel: (1)-202-434-8845

*Society of Manufacturing Engineers (SME)*

This is an American society formed of subscribing manufacturing organizations. It is responsible for the early development work on the *manufacturing automation protocol (MAP)* and the *technical and office protocol TOP*.

*Address*

Society of Manufacturing Engineers  
PO Box 930  
1 SME Drive  
Dearborn  
Michigan 48121  
United States of America  
Tel: (1)-313-271-1500

*Telecommunication Technology Committee (TTC) (Japan)*

This body (set up in Japan at the time of liberalization of telecommunications network services) is responsible for laying down network interface standards. It is composed primarily of leading Japanese scientists and engineers from network operators and equipment manufacturers. In the same way that T1 and ETSI set the pace for telecommunications standards in North America and Europe, respectively, so the TTC sets the *de facto* standards for the South East Asian region.

*Address*

The Telecommunication Technology Committee  
Hamamatsu-cho Binato-Ku  
1-2-11 Hamamatsu-cho  
Susuki Building  
Minato-Ku  
Tokyo 105  
Tel: (81) 33 43 21 551

*British Standards Institution (BSI)*

As an example of the many other national standards organizations, the *British Standards Institution (BSI)* produces telecommunications and other standards and is a member of the ISO.

*Address*

British Standards Institution  
Linford Wood  
Milton Keynes  
MK14 6LE  
United Kingdom  
Customer service and sales,  
Tel: (44) 181 996 7000

## 40.4 REGULATORY STANDARDS ORGANIZATIONS

As many of the world's governments have recently become keen to deregulate their telecommunications markets, allowing free competition particularly between manufacturers of *customer premises equipment (CPE)* as well as between network operators, there has been an increasing need for certain network interface standards to be mandated by law, and 'policed' by a regulating body. (*Customer premises equipment (CPE)* is the general name applied to any type of equipment connected to the public network. Thus telephones, computer terminals and other similar devices are all CPE.) The use of a mandated standard for the interface between network and CPE ensures the protection of the network operator's staff and equipment, meanwhile also assuring CPE manufacturers that their equipment operates satisfactorily over the network. The following organizations are active in laying down mandatory technical standards.

### *European Telecommunications Standards Institute (ETSI)*

Already mentioned in this chapter, one of the first jobs of the ETSI was to lay out the CPE-to-network interfaces that become mandatory in each of the European Union member states and other signatories of the CEPT's memorandum of understanding. The standards are known as *NETs (Normes Européennes de Telecommunications)*, French for 'European telecommunications standards'. Once a NET has been adopted, both public network operators and CPE manufacturers need to prove that their equipment conforms to the NET interface, before legal approval is given either for the launch of a new network service or for the sale of CPE. The NET programme, however, is significant in that CPE approval and certification (the 'green dot') will be valid Europe-wide (in any one of the signatory countries). Table 40.1 lists the original mandatory NETs.

### *Federal Communications Commission (FCC)*

The FCC is the regulatory body in the United States, reporting directly to congress, which lays down the code of practice to which network operators and equipment manufacturers must conform.

*Address*

Federal Communications Commission  
1919 M Street NW  
Washington DC 20554  
United States of America  
Tel: (1) 202-418-0260

**Table 40.1** Mandatory European Telecommunications Interface Standards

Normes Européennes de Télécommunications (NETs)	
<i>Access Group</i>	
NET 1	X.21 (Chapter 9)
NET 2	X.25 (Chapter 10)
NET 3	ISDN basic rate interface (Chapter 10)
NET 4	PSTN normal telephone access
NET 5	ISDN primary rate interface (Chapter 10)
NET 6	X.32 (Chapter 10)
NET 7	ISDN terminal adaptor (Chapter 10)
<i>Mobile Group</i>	
NET 10	Cellular network access
NET 11	Telephony terminal
<i>Modems Group</i>	
NET 20	General modem
NET 21	V.21 modem
NET 22	V.22 modem
NET 23	V.22 bis modem
NET 24	V.23 modem
NET 25	V.32 modem
<i>Terminals Group</i>	
NET 30	Group 3 facsimile terminal
NET 31	Group 4 facsimile terminal
NET 32	Teletex terminal
NET 33	Telephony terminal

*Office of Telecommunications (Of tel)*

Of tel is the regulatory body for telecommunications within the United Kingdom, overseeing the activities of all public and private telecommunications operators and users. It is part of the UK government's Department of Trade and Industry. Of late, *Of tel* has focussed its regulations on controlling the conduct of operators in the UK rather than on technical standardization. Standardization of technical interfaces for *interconnection* of networks has been delegated to the *Network Interconnection Coordination Committee (NICC)* while questions of quality and network design have been left to individual user choice.

*Address*

The Office of Telecommunications  
 50 Ludgate Hill  
 London  
 EC4M 7JJ  
 United Kingdom  
 Tel: (44)-171-634 8700

*British Approvals Board for Telecommunications (BABT)*

BABT is the independent body in the United Kingdom which undertakes customer premises equipment (CPE) approvals, confirming that submitted equipment conforms to the national network interface. BABT is the UK's main registered body for approvals which will apply throughout the European Union and other NET signatory countries.

*Address*

British Approvals Board for Telecommunications  
Claremont House  
34 Molesey Road  
Hersham, Walton-on-Thames  
Surrey KT12 4RQ  
United Kingdom  
Tel: (44) 932 251200

*Bundesamt für Zulassung in der Telekommunikation (BZT)*

BZT is the German equivalent of BAPT, being the responsible agency in Germany for testing and issuing of technical approvals certificates.

*Address*

Bundesamt für Zulassung in der Telekommunikation  
Talstraße 34-42  
66119 Saarbrücken  
Germany  
Tel: (49) 681 598-0

## 40.5 OTHER STANDARDS-PROMOTING 'FORA'

Towards the end of the 1980s a number of organizations were becoming increasingly frustrated because of the bureaucratic manner in which the official standards organizations operated, because of the constraining rules of procedure, and particularly by the long maturation time then required to develop and agree new technical standards. For equipment manufacturers this was resulting in lost sales, as potential purchasers held off until the standards were agreed. For network operators this meant a hold-up in the opening dates of new services. In a number of fields the frustration was so great that a number of like-minded organizations set about forming debating and lobbying fora to promote the quicker development of standards. Also typical is a permanent member-financed company to organize the work of the forum and promote its results. Typical also is a less formal procedure of meetings and contributions. Much of the work and communication within the working groups is instead conducted via electronic mail, and members may be obliged to own an electronic mailbox as a condition of membership. The most well-known of such fora are the *Frame Relay Forum*, the *Network Management Forum* and the *ATM Forum*.

*Frame Relay Forum*

The Frame Relay Forum was established with the objective to develop, agree and promote frame relay networks

*Address*

Frame Relay Forum  
303 Vintage Park Drive  
Foster City  
California CA94404  
United States of America  
Tel: (1) 415 578 6980  
[www.frforum.com](http://www.frforum.com)

*Network Management Forum*

The Network Management Forum (NMF) was established in 1989 with the objective to speed progress in the development of standardized 'umbrella' network management systems. The members are drawn broadly from two camps, major network operators (e.g. British Telecom, which played a leading role in the initial establishment of the forum and provided the first president) and computer manufacturers (e.g. DEC, Hewlett Packard, IBM, SUN, etc.). Valuable outputs of the NMF have been the OMNIpoint and SPIRIT (service providers integrated requirements for information technology) specifications. The OMNIpoint specifications were the first attempt to consolidate the many emerging computer and network management standards and specifications into a meaningful and realizable whole.

*Address*

Network Management Forum (NMF)  
1201 Mount Kemble Avenue  
Morristown  
New Jersey NJ 07960  
United States of America  
Tel: (1) 201 425 1900

*ATM Forum*

The ATM Forum came about through the interest of four major manufacturers who wanted to speed up the standardization process of ATM. It was set up in October 1991 by Northern Telecom (now called Nortel), Sprint, Sun Microsystems and Digital Equipment Corporation (DEC). In January 1992 membership was opened to wider participation from the telecommunication industry. The purpose of the ATM forum is to promote the development of technical standards for ATM and to raise awareness in the market of its capabilities. Many ATM standards adopted by ITU-T are heavily based on (if not identical to) those previously issued by the ATM forum.

*Address*

The ATM Forum  
2570 West El Camino Real  
Suite 304

Mountain View  
California CA 94040  
United States of America  
Tel: (1) 415 949 6700  
info@atmforum.com

### *Internet Engineering Task Force (IETF)*

The *IETF* is the group which develops the technical standards (*RFCs* (*requests for comment*)) for the *Internet*. The *RFCs* themselves can be obtained from

Jon Postel  
RFC Editor  
USC Information Sciences Institute  
4676 Admiralty Way  
Marina del Rey  
CA 90292-6695  
United States of America  
Tel: (1) 213 822 1511  
postel@isi.edu

*RFCs* may also be viewed using *FTP* on the host *nisc.sri.com*

## 40.6 PROPRIETARY STANDARDS

Despite the phenomenal number of standards organizations around the world, a few of which have been mentioned, some companies nonetheless feel that they can steal a march on their competitors by developing their own *proprietary* (as opposed to *open*) technical standards. By *proprietary* we mean that the standards are not available for free use by any manufacturer; either they may not be published at all, or alternatively a 'licence fee' is demanded when the standards are used in the design of new equipment. It is usually only large and powerful organizations who can afford to retain their standards as *proprietary*. The markets that such companies can command for their proprietary technology make it worthwhile to protect the technical rights. Other companies may find it lucrative to 'buy-into' these standards, so that they in turn may sell compatible 'add-on' equipment, but the lead company retains an edge. Examples of companies who have been successful in retaining control of proprietary standards are AT&T and IBM.

### *American Telephone and Telegraph Company (AT&T)*

One of the world's largest telecommunications carriers (based in the United States), AT&T and the new equipment manufacturer, Lucent, which was created by spin-off from AT&T, draw their technical strength from their famous research arm, Bell Laboratories. An example of one of AT&T's proprietary standards is the UNIX operating system, used as the operating software in many modern computer systems.

*Address*

Lucent Technologies  
Customer Information Centre  
2855 North Franklin Road  
Indianapolis  
Indiana 46219  
United States of America  
Tel: (1)-800-432-6600  
(1)-888-582-3688

*Bell Communications Research (Bellcore)*

Bellcore is the jointly owned research arm of the seven United States Regional Bell Operating Companies (RBOCs). It was formed partly with staff from AT&T's Bell Laboratories at the time of AT&T divestiture in the mid-1980s. Bellcore initially gained considerable respect for its networking standards called *Technical References (TRs)*, in particular in the field of intelligent networks and services (Chapter 11). Preliminary versions of TRs are called TAs (Technical Advisories). In recent years, however, Bellcore has suffered from a lack of clear direction and funding from its owners, who see themselves increasingly in competition with one another.

*Address*

Customer Service, Technical Publications  
Bellcore  
444 Hoes Lane  
Piscataway  
New Jersey 08854  
United States of America  
Tel: (1)-908-758-2142

*British Telecom (BT)*

BT is the largest public telecommunications operator in the United Kingdom. It is descended from the state-owned *Post Office Telecommunications* (formerly known as *GPO (General Post Office)*). British Telecom remains a predominant force in UK network standards setting. The standards are nowadays catalogued as *British Telecom Network Requirements* (or *BTNRs*). Examples include BTNR 190 (*DPNSS, Digital Private Network Signalling System*, a widely used standard in private networking of PBXs).

*Address*

British Telecom Network Secretariat  
Postpoint 2C07  
403 Saint John Street  
London EC1V 4PL  
United Kingdom  
Tel: (44) 171 843 5188



*International Business Machines (IBM)*

Perhaps the world's most important computer manufacturer, it was IBM who gave us the standard hardware and bus configuration for many mainframe and personal computers and the *systems network architecture (SNA)* to communicate data between interconnected computers.

*UK Address*

IBM Technical Publication Centre  
Alencon House  
Alencon Link  
Basingstoke  
Hampshire RG21 7EJ  
United Kingdom  
Tel: (44) 1256 478166

*European Address*

IBM Software Manufacturing Solutions  
IBM Denmark A/S  
Sortemosevej 21  
Alleroed  
DK 3450 Denmark  
Tel: (45) 2 33 000

*United States Address*

IBM Mechanicsburg Distribution Centre  
180 Kost Road  
Mechanicsburg  
Pennsylvania  
PA 17055-0786  
United States of America  
Tel: (1)-717-796-3200  
(1)-800-879-2755tpb -1

## 40.7 THE STRUCTURE AND CONTENT OF ITU-T RECOMMENDATIONS

Finally, we take a look into the structure and content of the *recommendations* of the world's leading authority on networking standards, the ITU-T. The ITU-T *recommendations* used to be approved every four years (at CCITT plenary assemblies of all the member organizations). This used to be the culmination of four years development work, establishing new recommendations, improving existing, and deleting obsolete ones. The recommendations were then published in a series of coloured volumes, a process taking an entire year because all the transcripts had to be translated and published in all the working languages. Each four yearly issue was

given a 'colour' which was assigned to each of the bound volumes in which the recommendations appeared, and every four years with each subsequent issue the colour was changed. The whole set of volumes corresponding to a particular year thus became colloquially known as The Blue Book (1988), or The Red Book (1984). Thus

- the *yellow* book was approved in 1980
- the *red* book was approved in 1984
- the *blue* book was approved in 1988

**Table 40.2** ITU-T series of recommendations

Recommendation series	Subject coverage
A	Organization of the work of the ITU-T
B	Means of expression (symbols, terms, vocabulary, etc.)
C	General telecommunications statistics
D	Recommendations for general application, including general tariff principles and charging and accounting in international telecommunication services and recommendations for regional application
E	International telephone network and ISDN: operation, numbering routing, mobile services, quality of service, network management and traffic engineering
F	Operations and quality of service: telegraph services, mobile service, telematic service, message handling services, directory services, document communication, data transmission services, audiovisual services, ISDN services, universal personal telecommunication, human factors
G	Transmission systems and media
H	Transmission of non-telephone signals
I	Integrated services digital network (ISDN)
J	Transmission of sound-programme and television signals
K	Protection against interference
L	Construction, installation and protection of cables and other elements of outside plant
M	Maintenance and telecommunications management network (TMN)
N	Maintenance of international sound-programme and television transmission circuits
O	Specifications for measuring equipment
P	Telephone transmission quality
Q	General recommendations on telephone switching and signalling
R	Telegraphy
S	Alphabetical telegraph terminal equipment
T	Telematic services
U	Telegraph switching
V	Data communication over the telephone network
X	Public data networks
Z	Programming languages (including SDL, specification and description language; CHILL, CCITT high level language; and MML, man-machine language)

The four year cycle of issuing new recommendations was ended in 1992, when the plenary assembly restructured the ITU. One of the objectives of the re-organization of the CCITT into the ITU-T was to speed up the process of agreement and publication of standards. Thus it became possible to issue new recommendations as soon as they became available. The recommendations issued since 1992 have therefore appeared individually as separate recommendations on plain white paper. Colloquially these are known as the *white* book.

Each updated recommendation is an improvement over its predecessor, but there is no guarantee of compatibility between equipments which have been designed to different versions of the same recommendation. Thus, for example, there is no guarantee that the message handling system (MHS) of the X.400 recommendation of 1988 was compatible with the version of 1984 (and indeed it was not).

One of the problems of the new publishing regime is that the overview of the standards has to some extent been lost. Each *book* of recommendations used to be generally cohesive and compatible within itself. Nowadays, it is difficult to obtain a full picture of all the other relevant recommendations, because some of these may not yet be fully complete and thus published.

Table 40.2 provides an overview of the subject areas, covered by the various *series* of recommendations, which make up each book. Recommendations are usually known by their series letter and a number; X.25 for example is a particular type of *data communications network interface* in the X-series of recommendations, in fact the one used in packet switched networks.

PART 6

---

# SETTING UP NETWORKS

---



---

## *Building, Extending and Replacing Networks*

---

It is rare to come across a telecommunications network that is not in a state of continuous evolution. At its simplest, a network could be expanding simply to cope with increased demand. In addition the network may be expected at the same time to provide increasingly sophisticated telecommunications services. For example, since the mid-1980s, many of the world's public telecommunications operators (PTOs) have responded to customer demand by the introduction of a 'freephone service' offering the facility for call recipients to pay for the calls. As an introduction to the subject of network evolution which the operator must be able to cope with, this chapter describes how networks may be built or extended to meet capacity and service needs. Particular attention is paid to the design factors inherent in the equipment ordering process; these are crucial to the success of a network evolution plan. In addition, and because it is not always possible to achieve the evolution without the entire replacement of the network and equipment, the chapter describes various methods by which network modernization can be achieved without a major disturbance or interruption to the established service.

### **41.1 MATCHING NETWORK CAPACITY TO FORECAST DEMAND**

No matter what type of telecommunications service is provided, network costs are minimized by matching capacity to the demand. Both over-provision and under-provision increase costs. When equipment is provided before it is actually needed, higher capital expenditure is incurred at an earlier date. In addition, there are higher running costs associated with equipment maintenance, accommodation, staffing, etc. What is more, when the equipment finally comes into use it may already be obsolescent and in need of early replacement. Under-provision, which some network operators have to put up with through lack of capital, is also expensive, because of the work needed to relieve congestion and to maintain overloaded equipment. Furthermore, higher costs are incurred in rearranging the network to incorporate new exchanges, as a result of the lack of network 'manoeuvring room'.

Figure 41.1 illustrates a graph of forecast demand, and shows a typical network operator's equipment provision schedule, in which 'steps' of extra capacity are provided to meet the projected demand. The demand may represent the total (local and long distance) traffic originated in a given area, in which case each 'step' on the capacity profile could represent the provision of a local exchange extension, or of a new local exchange. Alternatively, the demand curve could correspond to a particular type of traffic. One example might be the *trunk* or international traffic generated by a number of local exchanges, covering a wide area. The steps of capacity shown in Figure 41.1 might then correspond to new trunk or international exchanges. Another example might be a case where the demand predicted in Figure 41.1 is the forecast demand for a new telecommunications service, for which some special new equipment will be required. The steps then relate to the necessary provision schedule of that equipment.

It is seldom possible in large public networks to match the fitted capacity exactly to the demand, because the practicalities of exchange provision or extension usually allow capacity to be added economically only in fairly large chunks. It is therefore normal to keep the 'steps' of the fitted capacity graph above the demand, to ensure that the new equipment is on hand before the demand seems likely to exceed capacity. This reduces the risk of under-capacity. Where a network has slipped into under-capacity, new exchanges often become congested on the day they are opened, caused by the fact that forecasters tend to underestimate the amount suppressed traffic in congested networks.

By contrast, in corporate networks and networks of smaller service providers it may be possible to match the number of ports almost exactly to the growth in customer

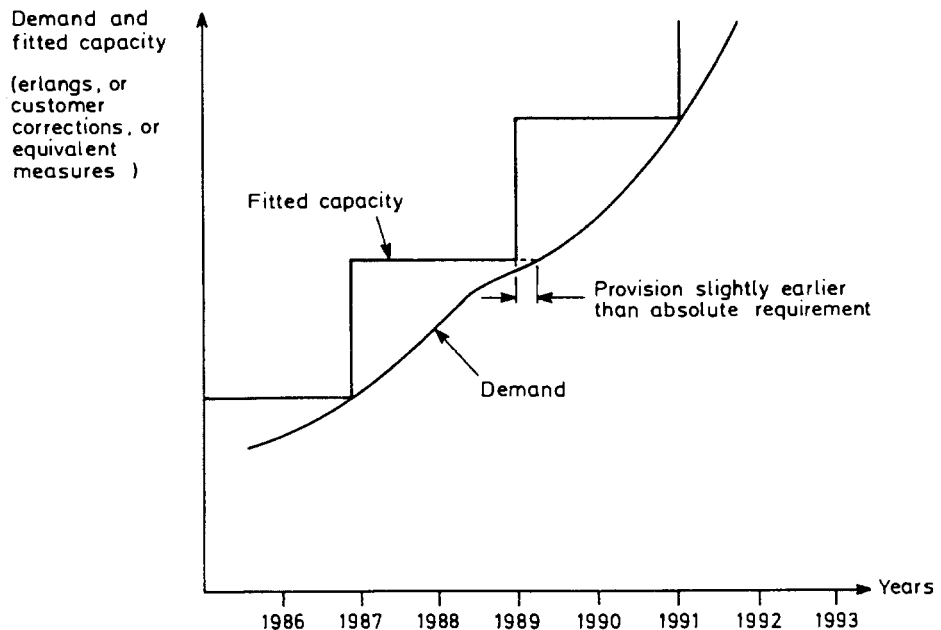


Figure 41.1 Forecast demand and fitted capacity

demand, by quoting a lead time for connection of new customers to the network which exactly matches the delivery time of a new port card from the network switch manufacturer. This is true *just-in-time (JIT)* provision of the network.

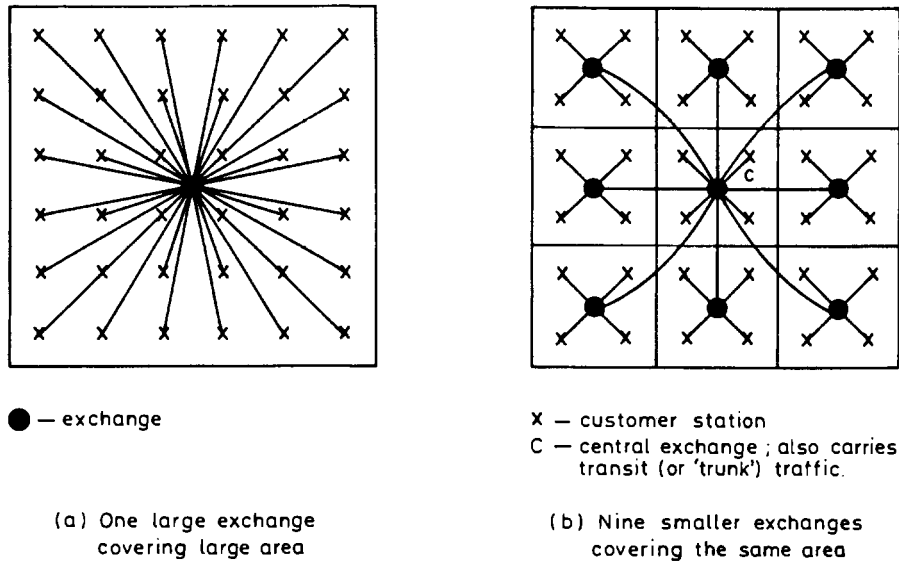
The forecast of demand is worked out as described in Chapter 31. For the purposes of exchange provision, the demand is normally quoted in terms of the highest instantaneous network throughput that is required. In circuit switched terms, the important parameters are the traffic intensity (i.e. the busy hour traffic in Erlangs) and the number of customer connections required. In data networks, the equivalent of the traffic intensity is the percentage trunk loading and/or the volume of data in segments to be carried per hour.

Forecasts form a direct tool for determining the forward network provisioning schedules, but to be useful the parameters to be forecast need first to be thought out carefully. Before making any forecast, the network planner needs to determine the geographical area which the traffic forecast is intended to cover, and the type or destination of traffic (data, voice, video, local, trunk, international) which any new switches (i.e. exchanges) will carry. In other words, he must have some idea of what type of network topology will be employed, and how particular services will be supported. This depends on a number of factors

- terrain conditions
- density of customers and/or ports or end user devices (telephones or data terminals)
- volume of traffic generated by each customer or user device
- volume and proportion of traffic to other local, trunk or international destinations
- the established network (if any), its capability for extension, its suitability for the support of any new services (if required), its state of repair, and its degree of obsolescence
- the reliability, optimum size, and service capabilities of contemporary equipment which might be used to extend or replace the established network

Designing a new network from scratch, to serve an area which has previously been ignored, has the advantage of allowing a 'clean slate' approach, but it is likely to be constrained by the amount of capital available. This may well limit the initial network design to choosing the optimum locations for such exchanges as can be purchased, and of deciding on each exchange's service or 'catchment' area. This task is best carried out by use of a map of the area to be served, marking it up to show the density of traffic likely to be originated and terminated in each spot. In areas that generate low volumes of traffic (e.g. rural areas or the branch offices of a corporate network it may prove economic to provide a medium size exchange to cover a very wide area or a number of corporate offices, but in extreme conditions it is usually efficient to employ a large number of smaller exchanges. The two diagrams in Figure 41.2 demonstrate the trade-off of exchange equipment against lineplant. In Figure 41.2(a) only one exchange is used, and each customer is connected to it by direct lines. Figure 41.2(b) shows an alternative network using smaller exchanges and it incurs greater cost in switching equipment, but the overall lineplant requirements and costs are reduced.





**Figure 41.2** The lineplant versus switch equipment trade-off. ●, exchange; x, customer station; C, central exchange; also carries transit (or 'trunk') traffic

Option (a) will generally be appropriate for urban areas, and areas where lineplant can be provided relatively easily and relatively cheaply. Option (b) might be more appropriate as a means of serving 'pockets' of remote network customers, particularly where the terrain makes the provision of lineplant difficult. An interesting feature of option (b) is the emergence of a hierarchical network structure, where the central exchange has taken on the role of transit switching between all the other exchanges. This typical occurrence in real networks was discussed more fully in Chapter 32.

For both networks shown in Figure 41.2, it is necessary for the network operator to forecast the future demand from customers, not only in terms of the traffic in Erlangs (or equivalent), but also in terms of the total number demanding an exchange connection. Shortage of either type of capacity may lead to customer dissatisfaction and loss of business. Shortage of customer connections means that new customers cannot be connected. Shortage of busy hour throughput (i.e. Erlang) capacity means that the needs of the established customers are not met. Graphs may thus be drawn, similarly to Figure 41.1, for both fitted traffic-carrying capacity in Erlangs and the number of customer connections. Such a graph enables the network planner to decide on the dates and sizes of future exchange extensions, or new exchange provisions. In this way the capacity may be matched to demand.

If the rate of growth in demand is slow, then small exchange extensions may be a good way of increasing the capacity. However, in cases when the growth is very rapid, it may be more appropriate to provide completely new exchanges, with large capacities. In addition, during rapid growth the provision of entirely new exchanges presents an ideal opportunity for adjusting the network topology, perhaps splitting an exchange area into two parts, with one area to be served solely by the new exchange while the old

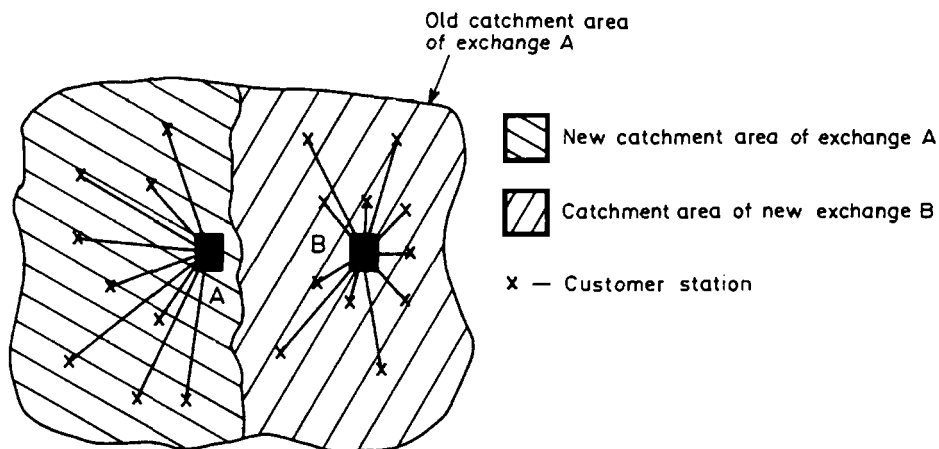
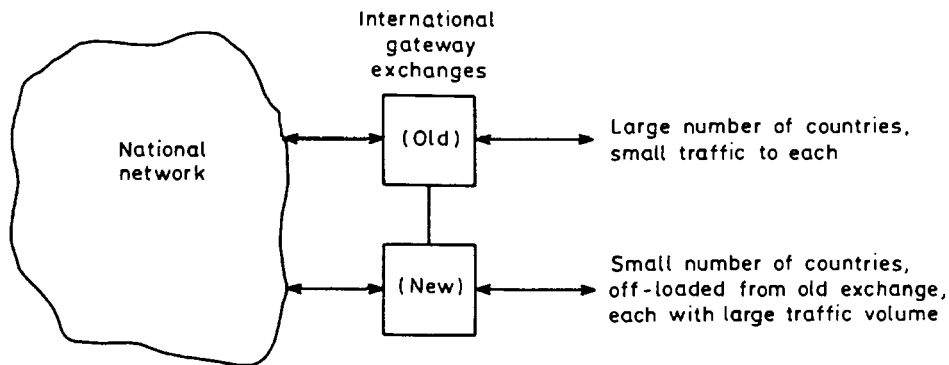


Figure 41.3 Splitting a local exchange area

exchange continues to serve the remaining area. Figure 41.3 illustrates the splitting up of an established local exchange area into two parts. Exchange A is the established exchange. Forecast growth in connections at a new town in the east of the catchment area will exhaust the capacity of exchange A, and so it has been decided to locate a new exchange in the new town (at B), and split the exchange area accordingly. This split means that the new exchange will cater for any further growth in the region around the new town, while the offload of traffic from exchange A onto the new exchange, will also allow for further growth in traffic in the western area. This traffic will be served by exchange A. The decision to split the area in this case is quite straightforward, because a large quantity of new local line wiring will be needed and can be established at the new exchange B. However, in the case in which all the wiring is already centred on exchange A, the economics would probably have forced an extension there.

Going back to Figure 41.2, option (b) can be used to demonstrate a further point. A forecast of the exchange connections and busy hour Erlangs must be drawn up for each of the exchanges shown in option (b), in the same way as already described. In addition, in the case of the central exchange C, a forecast is also required for the *transit* (or *trunk*) traffic to be carried by this exchange. This is the traffic between different outlying small exchanges which is routed via exchange C. In practice, a forecast will also be required for the traffic which exchange C will need to carry between any of the exchanges shown and to-and-from other geographic areas, not illustrated in Figure 41.2. In effect, just as local exchanges may be extended or may have their catchment areas adjusted, so may trunk and international exchanges (like exchange C).

Take as an example a single international exchange which is used initially to serve the international traffic needs of an entire country. Traffic growth may be such as to exhaust the capacity of the exchange, necessitating the provision of a new one. One way of reconfiguring the network in these circumstances would be to offload some of the overseas routes, corresponding to those overseas countries which have the largest volumes of traffic. The established exchange will thereby gain growth capacity (resulting from the offload), whereas the new exchange will serve and provide growth capacity for



**Figure 41.4** Reconfiguring international exchanges

a small number of heavy traffic routes. Figure 41.4 illustrates this example. Similar network reconfigurations to those shown in Figures 41.3 and 41.4 might also be appropriate for trunk or other transit exchanges.

## 41.2 OTHER FACTORS AFFECTING THE NEED FOR NEW EXCHANGES

Apart from a straightforward increase in demand, there are a number of other factors which a network operating company may take into account when deciding the best time to replace exchanges. These are

- the age and lifetime of the existing equipment (hardware and software)
- the obsolescence (or continuing usefulness) of the existing equipment
- the comparative running costs of existing and replacement equipment
- the new service (and retention of old services) demands of customers

Historically, electromechanical telephone exchanges were planned and operated on the basis of a 20 or 25 year lifetime. Over this period, the basic technology did not change very much, and the eventual need for replacement was governed by the increasing wear, and the consequent increase in running costs associated with maintenance, spares and labour. Modern computer-controlled equipment has a much shorter lifetime, perhaps as little as 3–7 years. The problem is not the reliability and wear of the electronic components, but the increasing rate of change of technology, and customer service expectations. Equipment lifetimes are becoming shorter, as the result of more rapid equipment obsolescence. An exchange might have plenty of copper wire termination ports but these cannot be used to terminate fibre lineplant unless converters are installed. This reduces the useful capacity of the exchange. In a similar way, older signalling system ports (for example pre-digital ones) become obsolescent and also reduce exchange capacity. Finally, equipment software needs occasional upgrading to

provide either the capability for new services or the updating of maintenance capabilities. Without upgrade, certain categories of traffic may need to be diverted to newer exchanges for processing.

With older exchanges the running costs alone may render them uneconomic. Many of the world's public telecommunications operators have been modernizing their networks prior to the originally intended life-end of the equipment. The modernization is carried out to reduce ongoing running costs, particularly in manpower terms brought about by the use of digital (as opposed to analogue) transmission and stored program control (SPC) exchanges. The new wave of modernisation following the *digitalization* which took place in the 1980s is the introduction of *intelligent networks* (Chapter 11) and sophisticated network management controls (Chapters 27 and 37).

### 41.3 FACTORS IN DETERMINING AN EXCHANGE PROVISION PROGRAMME

Figure 41.5 shows a similar demand curve to that in Figure 41.1, but in this case the diagram takes account of not only the need to match capacity to the forecast growth in demand, but also the decay in the effective capacity of the existing equipment. The decline in effective capacity comes about because of the gradual obsolescence of equipment and the eventual need for the replacement of life-expired exchanges.

Over the period up until 1990 the forecast is that the effective capacity of the existing system in Figure 41.5 will diminish due to the redundancy of equipment, in itself the result of obsolescence. Over this period the new exchanges, (1), (2) and (3), are required to handle the forecast growth in traffic, and to make good the shortfall which would otherwise arise as a result of the reducing effective capacity of existing equipment. The diagram shows the planned 'changeover' replacement of exchange (1) after a seven year lifetime. Diagrams similar to that shown in Figure 41.5 are an invaluable tool in determining future equipment provision programmes. The example shown in Figure 41.5 demands a forward provision programme as shown in Table 41.1.

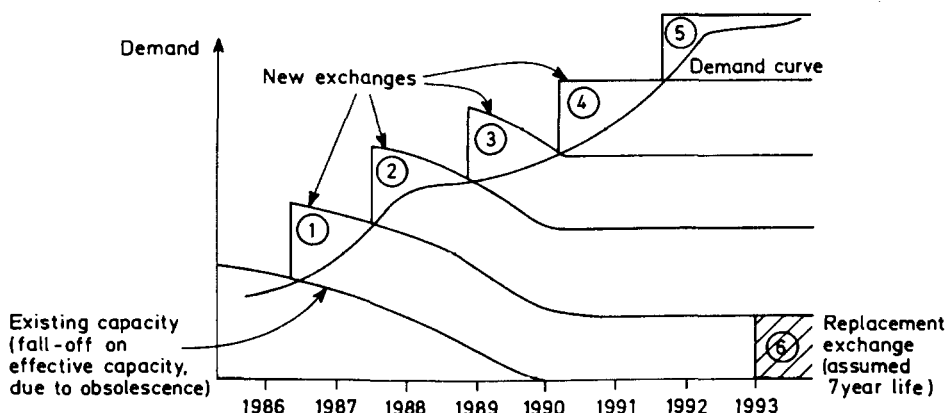


Figure 41.5 An exchange provision programme

**Table 41.1** Exchange provision programme

New exchange no.	Date of introduction
1	1986
2	1987
3	1988
4	1990
5	1991
6	1993

Any provision programme is critically dependent on the closure dates chosen for the older units, and also on the installed size of the new units.

Counteracting some of the disadvantages of over-provision (discussed earlier in the chapter), the provision of larger units in advance of demand reduces the frequency of exchange extensions and new exchange provisions, and it brings benefits in the way of reducing and simplifying the installation work programme. The use of larger exchanges also tends to simplify the job of inter-exchange network and traffic design (for example in Figure 41.2 no inter-exchange network is required for option (a), whereas a relatively complicated network is required to interconnect the nine exchanges of option (b)). On the other hand, the use of a larger number of small units may be the best answer in difficult terrain, or in cases where there is an established workforce, network or building already available. In the end the overall exchange provision is a compromise between a number of constraints.

#### 41.4 DETERMINING A STRATEGY FOR NETWORK EVOLUTION

The efficiency and reliability of a network ultimately depend on the expertise of its designer, and there is no substitute for experience in developing network evolution schemes tailored to particular circumstances. Nonetheless, it is helpful to have a system for selecting the best available evolution scheme when there is more than one to choose from. A systematic approach also guards against the accidental oversight of crucial considerations. The factors listed below set out a framework for determining and evaluating alternative network evolution schemes, and by using them the network planner can devise a forward strategy for evolution and development.

##### *Capacity exhaustion date*

This provides the 'backstop' date, by which new capacity must have been provided. Some evolution schemes may be precluded if they do not provide extra capacity early enough.

##### *Short-term network constraints*

It may be impossible to avoid the provision of further, already obsolescent, equipment to meet a short term peak in requirements (for example, a new international exchange may need to include a few lines using an archaic signalling system, to meet growth to a

particular country, prior to modernization of equipment in that country). It is no good buying such modern equipment that it will not interwork with the existing network. Some evolution schemes or particular types of equipment may therefore be precluded.

### *Objective long-term network topology*

Over the lifetime of any exchange, the network will constantly be evolving towards the goal of the long term strategy. A good network planner will optimize the procurement date and design of the exchange to maximize the value of the exchange throughout its life, and ensure its compatibility with the long term objectives.

### *Evaluation of overall costs*

The total lifetime costs of an exchange (or of a complete exchange provision programme) include not only the capital costs associated with the provision of the equipment, the network, the accommodation, the peripherals, the training, the documentation and the spares, but also the ongoing running costs associated with equipment maintenance, rates, bills, electricity, staff costs, etc. Although one type of equipment may be cheaper to buy initially, it may be significantly more expensive in running costs. A *discounted cash flow (DCF)* analysis, to determine the *net present value (NPV)* (also called the *present worth*) costs of the exchange, or of the entire provision programme is often carried out. The example in Table 41.2 compares the net present value of providing a single large exchange of one equipment type, against an alternative strategy of purchasing a different manufacturer's exchange equipment, which allows some of the capital outlay to be deferred by phasing the overall provision capacity. The equipment capital costs shown in Table 41.2 compare the provision of a large unit costing £10 million in one go, with the provision of an initially smaller unit which is extended twice, and costs £12 million in total. The running costs (maintenance and manpower) are assumed to be 5% per annum of the equipment value (a typical assumption). The accommodation costs are common to both cases, and include £1 million initial purchase and fitting out of the building, plus £0.1 million per annum rent and rates, etc. It is interesting to see from the full-life cost analysis shown in Table 41.2, how, despite the fact that the overall expenditure is cheaper for option 1 than for option 2 (£15.2 million as against £16.7 million), it is nonetheless cheaper in present value terms to adopt option 2 (£13.45 million compared with £14.13 million). The reduced present value results from the deferred outlay. In the early years, more money is therefore 'left in the bank' (we may think of it as earning interest).

From Table 41.2 the planner could equally decide that £0.68 million was cheap insurance against unplanned demand. In either case a full cost analysis would also include the cost of transmission plant provisions and rearrangements; these can be much greater than the switching costs.

In a private network, where line capacity is leased from the public telecommunications network operator, it is easy to match lineplant to demand: pay as you need.

### *The profile of installation work*

The profile of installation work may be an important factor, because it may either be impossible for the chosen manufacturer to deliver equipment at a very fast rate (for example, a rapid and large-scale network modernization may require very large

**Table 41.2** Discounted cash flow analysis

	Year							Total
	1987	1988	1989	1990	1991	1992	1993	
	(£m)	(£m)	(£m)	(£m)	(£m)	(£m)	(£m)	(£m)
<i>Option 1 (large exchange)</i>								
Equipment capital costs	10	—	—	—	—	—	—	10
Running costs	0.5	0.5	0.5	0.5	0.5	0.5	0.5	3.5
Other costs (accommodation etc.)	1.1	0.1	0.1	0.1	0.1	0.1	0.1	1.7
Total cost in year	11.6	0.6	0.6	0.6	0.6	0.6	0.6	15.2
Present value of cost (10% discount)	11.6	0.54	0.54	0.49	0.44	0.39	0.35	14.13
	Year							Total
	1987	1988	1989	1990	1991	1992	1993	
	(£m)	(£m)	(£m)	(£m)	(£m)	(£m)	(£m)	(£m)
<i>Option 2 (large exchange)</i>								
Equipment capital costs	4	—	4	—	4	—	—	12
Running costs	0.2	0.2	0.4	0.4	0.6	0.6	0.6	3
Other costs (accommodation etc.)	1.1	0.1	0.1	0.1	0.1	0.1	0.1	1.7
Total cost in year	5.3	0.3	4.5	0.5	4.7	0.7	0.7	16.7
Present value of cost (10% discount)	5.3	0.27	3.65	0.36	3.08	0.41	0.37	13.45

resources beyond the means of some suppliers). It may also prove impossible for the installation workforce to match a stop-start programme. It is much better, once an installation team has been established, to maintain as steady a workload as possible.

### *Recruitment of maintenance staff*

Consideration should be given to the maintenance of both old and new exchanges. Transferring and retraining staff from old units onto new ones may help to reduce recruitment difficulties, but the ability to do so will depend on the relative geographical locations of the two units and the skills match of the staff involved.

### *Associated network changes*

The introduction of a new exchange to support a new service, or as part of a technology modernization strategy, may have to be coordinated with other network rearrangements, with a corporate reorganization, with public announcements, or with changes in the law.

### *Ongoing network administration*

One provision strategy may lead to much easier ongoing network administration than another. Perhaps one strategy allows *centralized operations* (all maintenance staff

performing remote maintenance from a single central location). One scheme may ease the task of monitoring ongoing network congestion and quality performance. Another might ease the task of network design, reducing the planning manpower that is needed. Finally, one strategy might leave the traffic less prone to network failure, so that less traffic would be lost in the event of a single exchange or transmission link failure.

#### *Available exchange size*

Exchange costs are often ‘front-loaded’. (By this expression we refer to the fact that you need to pay for a building, a maintenance staff and a central processor for an SPC exchange, no matter what its size.) Per unit of capacity, larger exchanges therefore tend to be cheaper. The discounts available from suppliers for bigger orders are also generally more favourable.

#### *Commercial conditions for equipment provision*

Open tenders for equipment, sent to a number of potential suppliers, will often attract lower prices, because competing equipment providers are eager to gain maximum business. To allow open tendering, the requirements for the equipment should be kept as open as possible. Non-competitive tenders, for the extension of existing equipment by the original supplier, should be avoided where possible. Where an extension is anticipated, it is better to make this part of the initial tender and contract rather than negotiate it after the first installation has been made. When one supplier is ruled out, either by the stringency of the requirements or because of equipment being extended, there is little incentive for this supplier to offer competitive prices. Some network operating companies deliberately buy equipment of two compatible makes. Even though in the very short term this is clearly not the cheapest option, it protects them from becoming wholly dependent on either of the individual manufacturers. (We discuss the commercial considerations for contract placement in Chapter 42 on ‘selecting and procuring equipment’.)

#### *Equipment provision lead times*

There is a finite period of time required to develop new software for exchanges, to manufacture the hardware, and to plan, install and test the equipment. This time is called the equipment *lead-time*, and must be taken into account when determining the exchange (or other equipment) provision programme. Indeed, some manufacturers may rule themselves out of possible consideration for a contract because the lead-time that they are able to tender is not short enough. Insufficient respect for the lead-time, or over-optimistic estimation of development timescales leads to delayed opening dates, and almost inevitably to network congestion or some other problem.

#### *Suppressed demand*

A factor sometimes worthy of consideration, when injecting new capacity into congested networks, is the quantity and likely effect of any suppressed traffic demand. In particular, will an unexpected heavy load present any problems to the new exchange, and have sufficient tests been planned?



## 41.5 COMPARISON OF STRATEGY OPTIONS

When planning the evolution of networks, it is usually best to compare a number of different provision strategies, evaluating the differences (as outlined in the previous section) between perhaps

- a strategy relying on the extension of existing exchanges (i.e. switches)
- a strategy adding new exchanges
- a strategy replacing all exchanges
- a long term strategy for a small number of large exchanges
- a long term strategy for a large number of small, geographically dispersed exchanges

The options used in the evaluation should be restricted to a small number of practical options (say three or four) and compared on as many factors as possible. A complete long term provision programme should always be worked out, even though it may be necessary, in the short term, only to commit expenditure to a single new exchange or transmission line system. This ensures that consideration has been given to how the new item of equipment will operate in the prevailing network throughout its life. The analysis will finally lead to a decision about the exchange size, location, in-service date, service requirements, and possibly which manufacturer is to be asked to provide it. A similar analysis can be used to determine the required-by date and capacity for a new line system.

## 41.6 EXCHANGE DESIGN AND SPECIFICATION

Having decided on the size, location and in-service date for a new exchange (i.e. switch) or exchange extension, next comes the task of mapping out the design of the exchange itself. This is usually conducted in two phases. First an outline design sets out the functions and overall requirements, leading to an ultimate and more detailed specification. The specification must list the exact functions or services to be provided, defining the required software capabilities and performance requirements. It must also include the signalling systems to be provided, the routing and number analysis capabilities required, the billing and charging features needed, and the expected maintenance and network management features, etc. These are all features without which the exchange could not operate, either in isolation or within the network. In addition, the specification needs to list any other interfaces to be provided so that the exchange can be interconnected to, and interwork with, any peripheral equipment. Such *peripheral* equipment might include echo suppressors, circuit multiplication equipment, statistics post-processing computers (for the processing of accounting, traffic recording, and billing records), recorded announcement machines, network management systems, network configuration databases, etc.

Each interface needs to be clearly and unambiguously defined, so that the exchange correctly interworks with the peripheral equipment. The likelihood of incompatibility is increased when the two pieces of equipment are provided by different manufacturers,

and for this reason standard interfaces should be used as far as possible. The use of standard interfaces also tends to reduce the development costs incurred by the manufacturer, because these interfaces are likely to have been developed already. A lower equipment price therefore pertains (as compared with an equivalent exchange for which special development has to be undertaken). Figure 41.6 illustrates a possible exchange configuration for a modern day, digital and stored program control (SPC) exchange.

In Figure 41.6 the exchange is shown surrounded by a number of different types of equipment, each with its own particular interface. There are computer systems, one performing the *network management* functions of network status monitoring and network control (as already described in Chapter 37); the other is a database, with an up-to-date record of the network configuration surrounding the exchange. Both systems are connected electronically to the exchange to be kept automatically up to date with any changes in the network. Two X.25 datalinks are employed for the interconnection. These carry data information across a Q3-interface according to the specifications of the *telecommunications management network (TMN, Chapter 27)*.

The exchange in Figure 41.6 is shown with magnetic tape interfaces for downloaded accounting, billing, and traffic statistics, and uploading software and exchange data. The important *interface* in this instance is the format and informational content of the data on the tape. The peripheral computer needs sufficient information to carry out any necessary accounting, billing or traffic recording functions, and it needs to understand the format in which this information is passed out by the exchange, so that the data may be correctly interpreted.

Another couple of interfaces shown in Figure 41.6 are the control mechanisms for circuit multiplication equipment and for echo suppressors. These can be relatively simple, 'hard-wired' electrical leads, indicating simple on/off controls, or they may be the more sophisticated interface defined in ITU-T recommendation Q.50.

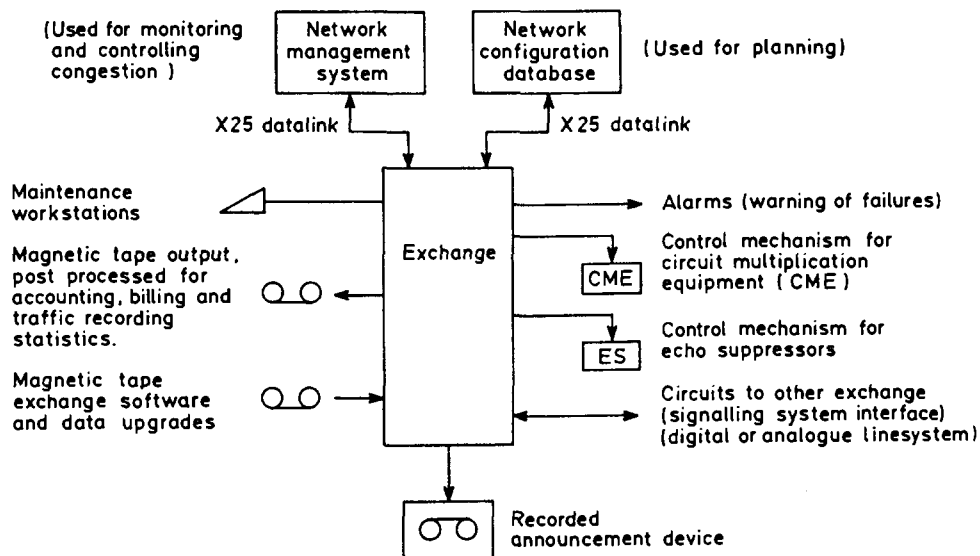


Figure 41.6 Some typical exchange interfaces and peripherals

The final interfaces shown in Figure 41.6 are those intended for maintenance workstations, for recorded announcement devices, and for reporting equipment alarms. In the case in which the peripheral devices required are provided by the exchange manufacturer as part of the overall contract for the exchange, it may not be important for the network operator to specify the interface in great technical detail. The exchange manufacturer is then free to use a *proprietary* interface, probably specially designed for the purpose. The benefit to the network operator might be simpler working practices at the maintenance workstation, or perhaps more informative alarms or recorded announcements.

Figure 41.6 is not meant to illustrate a comprehensive or exclusive set of interfaces; there are other interfaces which are appropriate, depending on the type and intended use of the exchange.

Just like exchanges, line terminating equipment, cross-connect frames and multiplexors require maintenance and alarms, and also like exchanges they are under progressive development to be capable of being remotely maintained or reconfigured using datalink interfaces. For these reasons, large portions of the specifications may be the same or similar.

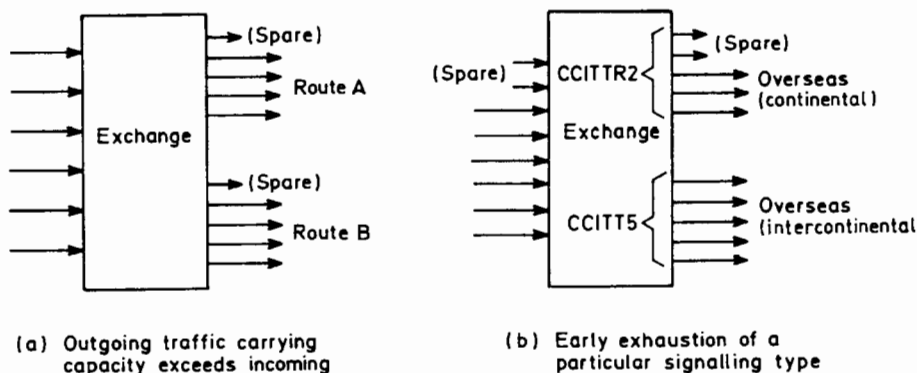
Exchange specifications are often produced in two stages. An initial outline design lays out the network topology and describes the overall *functional* requirements of the exchange, and it leads to the lengthier detailed specification. The next chapter covers in some detail the points to be considered when preparing a full specification. The remainder of this chapter concentrates on the important considerations and tactics appropriate to the outline design stage.

## 41.7 OUTLINE CIRCUIT-SWITCHED DESIGN: CIRCUIT NUMBERS AND TRAFFIC BALANCE

In designing an exchange for a circuit-switched network it is crucial to order the right number of circuits and the correct 'mix' of different signalling system termination types. If there are not enough terminations of a particular kind, or too few *outgoing* or *incoming* circuits, then the exchange cannot be fully loaded and its effective traffic-carrying capacity is reduced. Figure 41.7 shows two scenarios in which inappropriate circuit and signalling mixes have led to premature exchange exhaustion.

In Figure 41.7(a), with the circuits that have been ordered, the exchange cannot be supplied with enough incoming traffic, because the incoming circuits have been exhausted. The spare outgoing circuit terminations are wasted, and the effective exchange capacity is correspondingly reduced. A similar wastage could also have resulted from a shortage of outgoing circuits and an excess of incoming ones. In this case, incoming circuits would need to be left spare, to prevent the exchange running into congestion. The best way to alleviate either of these conditions is to carry out a small exchange extension (if possible) of incoming or outgoing circuits as appropriate.

In Figure 41.7(b) a slightly less obvious traffic imbalance has crept in. Illustrated is an international exchange, drawing incoming traffic from a national network, and connecting it onward to both continental and inter-continental overseas destinations. For continental destinations, R2 signalling is used; for inter-continental destinations,



**Figure 41.7** Poor circuit mix causing premature exchange capacity exhaustion

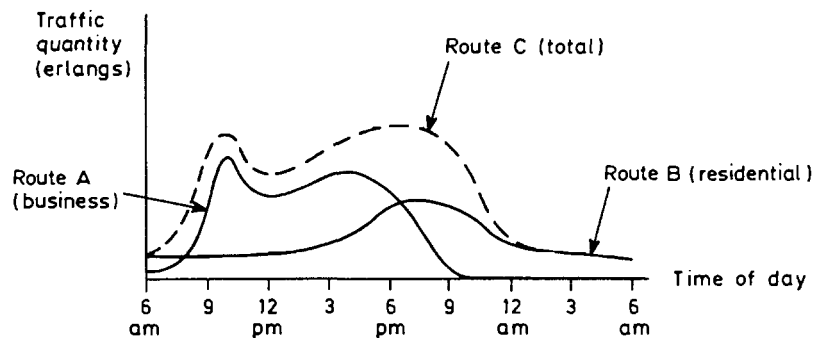
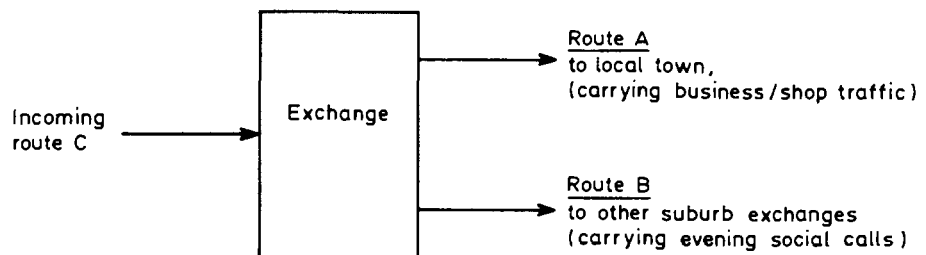
the signalling system is CCITT5 (this signalling type is rapidly becoming obsolete, but nonetheless is used here to illustrate the point). The diagram illustrates the exhaustion of CCITT5 signalling terminations. Although overall there are still incoming and outgoing circuits available (incoming from national, and outgoing in R2), any further loading of the exchange will lead to congestion on the CCITT5 routes. This premature exhaustion comes about because, when further incoming circuits are connected from the national network, they feed a mix of further traffic to both the inter-continental (CCITT5) and continental (R2) destinations. However, no further CCITT5 circuit termination can be made available. The imbalance is therefore in the relative quantities of CCITT5 and R2 signalling terminations, and the actual ratio of signalling termination types should match the balance of intercontinental to continental traffic. Two methods are possible for alleviating the type of premature exhaustion shown in Figure 41.7(b): one is to order an exchange extension to correct the CCITT5 shortfall, and the other is to pre-sort the traffic (at a prior exchange within the national network), so that a disproportionate amount of continental traffic is sent to this particular international exchange. This allows the spare incoming and outgoing R2 terminations to be used, without adding traffic to cause congestion on the inter-continental (CCITT5) routes. Of course, in this circumstance, some other international exchange will need to carry a disproportionately large share of intercontinental (CCITT5) traffic.

The following procedure will help to determine an appropriate exchange termination mix, and to avoid such traffic imbalances. It should be applied with care, and due consideration should also be given to studying the whole range of expected network configurations in which the exchange will be used throughout its life. For each configuration, do the following.

- (i) Calculate the circuit number, signalling system types, and traffic direction requirements to meet the busy hour requirements of each individual route connected to the exchange.
- (ii) Add up the total incoming and outgoing traffic in Erlangs, and compare these totals with the total incoming and outgoing circuit numbers. (Bothway circuits, if

any, should be considered as half incoming, half outgoing). Ensure that the numbers are balanced, or within about 10% of one another, but do not arbitrarily re-adjust the numbers to give a closer match. If an imbalance exists, go back and check it. One effect which can lead to apparent imbalance is that of non-coincident route busy hours. Figure 41.8 shows an example in which the optimum circuit numbers may at first glance appear misaligned, but in fact are not. Outgoing Route A carries mainly business traffic, predominantly during the day time. The busy hour itself is at 9 a.m., when the traffic is 5 Erlangs (requiring 11 circuits on 1% tables). Route B by contrast carries mainly residential traffic, with a busy hour traffic load of 4 Erlangs at 8 p.m. (requiring 10 circuits). Route C is the incoming route, serving both outgoing routes; its busy hour traffic is 7 Erlangs, at 7 p.m. (requiring 14 circuits). Summing up the route busy hour traffic values, and incoming/outgoing circuits, there appears to be an imbalance, as demonstrated by Table 41.3. However, as the analysis above shows, the imbalance of circuits is relatively extreme (being 50% extra on the outgoing side). Large real exchanges rarely show an imbalance greater than about 10%.

- (iii) Other traffic balancing studies may be appropriate. A useful one to perform on international exchanges is of national 'side' traffic to international 'side' traffic. Similar balancing of local 'side' to trunk 'side' traffic could be made for trunk (toll) exchanges.



**Figure 41.8** Traffic balance and non-coincident route busy hours

**Table 41.3** Apparent traffic and circuit imbalance

	Total incoming	Total outgoing
Route busy-hour traffic in Erlangs	7	$5 + 4 = 9$
Appropriate number of circuits	14	$11 + 10 = 21$

- (iv) Finally, it is worth calculating the peak cross-exchange traffic and comparing this with the sum of the individual route busy hours. The maximum traffic across the exchange (the cross-exchange traffic) will be less than the sum of all the route busy hours, unless all the route busy hours are coincident in time. This gives a plausibility check on the exchange traffic design. The nominal value of the cross-exchange traffic is also invaluable to the exchange manufacturer, because it is the maximum call throughput rate which the exchange common equipment (for example, the central processor) must be capable of handling.
- (v) Should any of the traffic balances highlight unexplainable imbalances, then the original design should be checked.

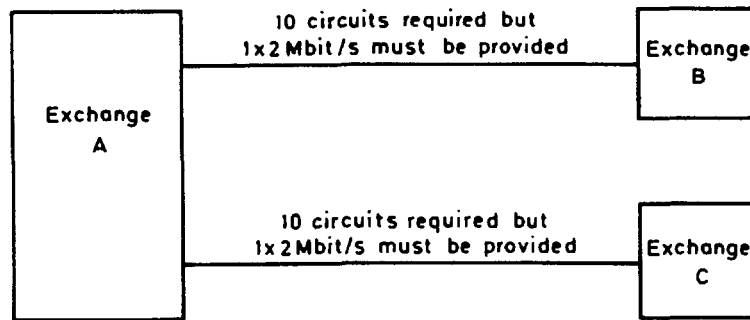
## 41.8 OUTLINE DESIGN OF OTHER TYPES OF NETWORK

Traffic balance considerations similar to those in the preceding section should be applied when designing networks other than circuit-switched ones. For example, in a traffic balance of a packet-switched exchange the overall packet volume and rate carried by each route and the exchange as a whole would be taken into account.

## 41.9 THE EFFECT OF LOW CIRCUIT INFILL ON EXCHANGE AND LINEPLANT PLANNING

One final trap for the unwary exchange or lineplant planner, is the effect of low circuit *infill*. The problem arises when circuits are multiplexed together within the exchange, either on a higher order analogue FDM system, or as a 2 Mbit/s (or 1.5 Mbits) higher bit rate digital line system. In such a case, the whole analogue or digital *group* must be provided between the two end points, even if the circuit demand is not sufficient to need all the circuits in the *group* (12 circuits for a normal analogue group, 30 circuits for a 2 Mbit/s digital line system, 24 circuits for 1.5 Mbit/s). The effect is known as low circuit *infill*. In cases where the circuit *infill* is lower than 100%, extra lineplant and exchange termination capacity may be required.

Figure 41.9 shows an example in which a total of 20 circuits is required between exchange A, and two other exchanges, B and C. The terminations available on all the



**Figure 41.9** The effect of poor circuit infill

exchanges are standard 2 Mbit/s (European) digital terminations, and the only lineplant is also 2 Mbit/s. In result, two 2 Mbit/s digital line systems must be used, together with two 2 Mbit/s exchange terminations at exchange A. In other words, unless instead the planner reverts to transit routes to access the destinations, the equivalent of 60 direct circuits must be provided, with an effective circuit *infill* of 10 circuits within each 2 Mbit/s line system, i.e. 33% *infill*.

Poor *infills* can also occur in higher order line systems. For instance, within an analogue supergroup it may be that only four FDM groups are used. (To provide the circuits using a supergroup may still be cheaper than providing separate lines for four individual analogue groups). A similar example in digital practice might be the use of only three 2 Mbit/s *tributaries* of an 8 Mbit/s digital line system, or the use of only two 34 Mbit/s *tributaries* within a 140 Mbit/s digital line system, etc. So, as with exchange terminations, extra lineplant capacity must be provided on all occasions when poor circuit infills are encountered.

## 41.10 FUNCTIONAL REQUIREMENTS OF EXCHANGES OR LINE SYSTEMS

We have now discussed in some detail the traffic flow design of exchanges, but what other functional characteristics will be required? The answer depends on the expected performance of the switch. The next chapter, on equipment procurement, at least provides a rough checklist of items which may be appropriate to the final specification.

## 41.11 METHODS OF NETWORK OR EXCHANGE MODERNIZATION

Finally, before going on in the next chapter to discuss the production of a detailed equipment specification, it is worth considering various tactical methods for network or exchange modernization. The earlier part of this chapter considered how to determine the provision programme for exchanges or new line systems, and talked about the need

for new equipment to meet growth or extension in demand, for equipment upgrade or replacement of life-expired equipment. What was not covered in any detail was the tactical methods that might be employed for network modernisation (or replacement). Basically, two methods are available for either exchange or lineplant modernization; they are the *integration* method and the *overlay* method. Such methods were used in the 1980s modernization from analogue to digital networks, and will be needed again in the upgrade from ISDN to B-ISDN.

#### 41.11.1 Integration

A new exchange or line system is provided in a manner fully integrated with the existing network. A typical example might be a one-for-one or one-for-many replacement of older equipment. Figure 41.10(a) shows a one-for-many replacement of four small analogue local exchanges with a single larger modern exchange. Figure 41.10(b) shows a one-for-one transfer of an analogue line system onto a digital replacement.

The *integration* method of replacement is generally used where the old and new technologies are expected to co-exist for a long time, as part of a gradual modernization regime. The changeover of individual exchanges or line systems may be carried out slowly, circuit by circuit, or by a special rapid wiring 'changeover' device.

#### 41.11.2 Network Overlay

When network operators are embarking on more radical modernization, perhaps undertaking a heavy capital investment programme to replace the network very quickly with modern equipment, an *overlay* network approach may be the most economic method. Network overlay is also a good method of introducing networks for new or

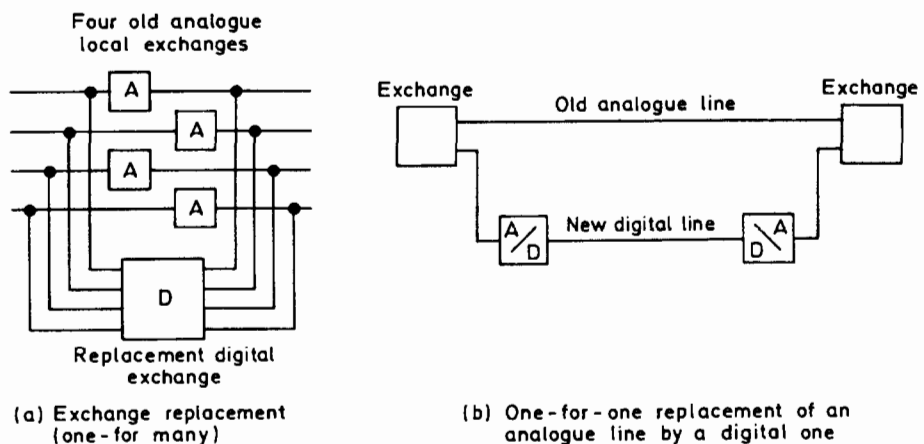
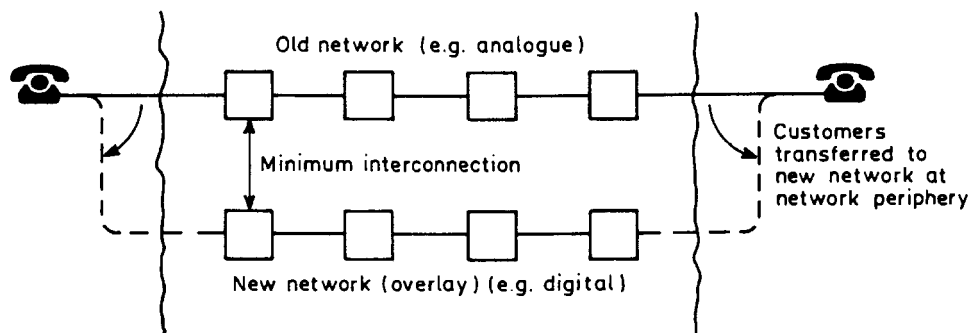


Figure 41.10 Methods of equipment replacement





**Figure 41.11** Network overlay

special services. In the *overlay* method, the new equipment is linked together as an entirely separate (or *overlaid*) network. Customers are then transferred from one network to the other at the customer-to-network interface on the periphery of the network. Figure 41.11 demonstrates the principle.

The advantage of using a network overlay is the much reduced number of interconnections required between old and new networks. Interconnection is often still needed so that customers of the 'old' and 'new' networks may inter-communicate, but the extent, and therefore the cost of large scale *interworking* between the two networks can be contained. For example, the amount of analogue-to-digital conversion equipment between an old analogue network and a new digital overlay network can be minimized. In addition, the man-effort associated with the extensive network reconfiguration of interworked networks can be largely contained.

---

## *Selecting and Procuring Equipment*

---

The preceding chapter described the factors in outline equipment design which are crucial to the success of the network evolution plan. Having a good plan is one thing; executing it is another, and in this respect the ordering and delivery of the equipment should receive considerable attention. This is an activity often referred to as 'procurement', and it can be carried out very successfully by highly mechanistic management methods. Various 'project management' techniques are available which tackle network design as if it were the input of a procurement 'process', much as 'raw material' is the input to a manufacturing 'process'. In production-line fashion, project management techniques lead the project through its design and checking stages, and on to ordering, installation, and testing. However, in the same way as a poor raw material has an effect right through to the end product of a manufacturing chain, so the defects of a poor network design cannot be made good during implementation. This chapter describes a typical methodology for selecting and procuring equipment, starting from the outline equipment design.

### **42.1 TENDERING FOR EQUIPMENT**

Except for some of the largest corporations, very few companies that operate networks also manufacture the equipment that goes to make them up; everything from personal computer links upwards needs to be bought from other manufacturers. For small pieces of equipment like modems the purchase may be straightforward, because a range of items meeting a standard specification (say, an ITU-T V-series recommendation) is available from a number of manufacturers. The purchaser then has time to concentrate on the finer differences (size, price, reliability, ease of maintenance, etc.) of the various equipments, and to choose what suits them best. More costly equipment, like major transmission systems or large exchanges, can seldom be bought 'off-the-shelf'. Instead, the differing circumstance of each network and application means that a considerable amount of 'adaptive engineering' is required in each case, to modify or upgrade the manufacturer's own 'basic' equipment design. Selecting suitable equipment is then

much more difficult, because at the time of order it may be that none has yet been developed. Under these circumstances, the normal method of equipment selection is by an *invitation to tender (ITT)*, *request for proposal (RFP)* or *request for quotation (RFQ)* and procurement follows after a contract has been placed with the manufacturer who has tendered a price and a technical conformance nearest to the specification.

As well as a technical specification, an invitation to tender includes a number of commercial and other contract conditions. The specification itself either lays out in precise detail the individual electrical components and connections to be made, or it is a *functional specification* which lays out only which general functions must be performed and which external interfaces are required. *Functional specifications* are preferred by tenderers and manufacturers alike because they place fewer constraints on the internal design of the equipment. They permit a wider range of manufacturers to consider adaptation of their products to meet the external interface requirements. For the purchaser, the larger the number of equipment manufacturers kept in competition and providing 'standard' equipment, the lower the tendered prices will be. For the manufacturer the benefit is the larger market.

Each *tender* sets out the degree to which the equipment can be adapted to conform with the specification, the degree to which the commercial conditions are considered acceptable, and the quoted price. Even if none of the tenders meet the specifications and commercial conditions in every respect, purchasers can at least gain an understanding of each piece of equipment and decide on their own trade-offs between various factors, which are

- price and conditions for payment
- amount of further technical development required by an existing product, and how much confidence there is that it can be achieved
- ease of equipment maintenance
- time required for installation, and whether the in-service date is early enough
- quality of product and installation
- modularity of kit; the ease of with which parts may be replaced or with which a subsequent equipment extension may be carried out
- warranty and after-sales service
- the area and standard of accommodation required
- robustness and reliability of equipment
- political factors (nationality and financial liquidity of manufacturer, etc.)
- software licence and conditions

The process of invitation-to-tender and response is illustrated in Figure 42.1 and is much the same, no matter what type of equipment is being purchased. The length of each of the working documents (specification, tender, etc.) depends on the complexity of the requirement and the confidence gained from previous experience, but the contract itself needs only to be a short document committing the manufacturer to

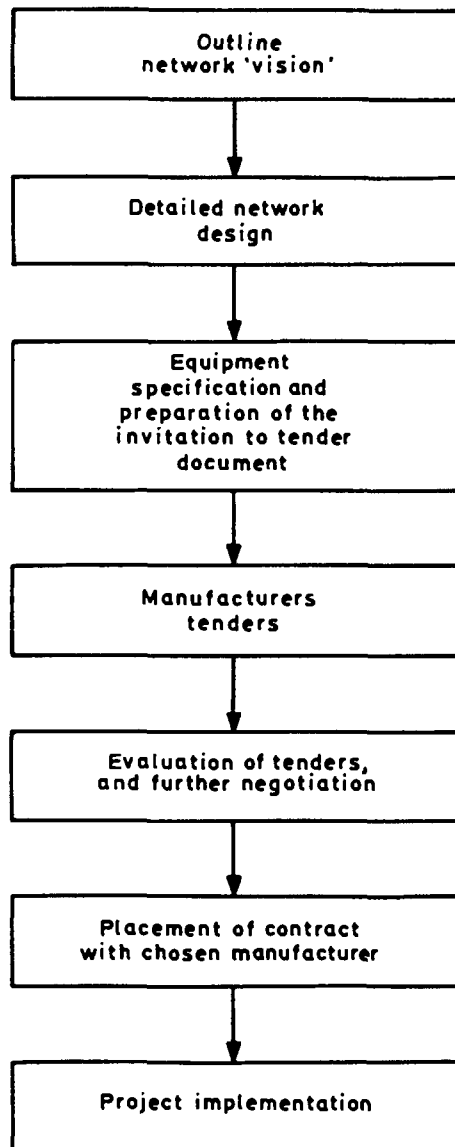


Figure 42.1 Procuring equipment

conform with the tender response. In the event of a failure to conform with the requirements of the contract, a number of penal clauses usually cover each of the possible eventualities. For example, failure on the part of the manufacturer to develop a given function may result in no payment, while failure on the part of the purchaser to pay a given instalment, or to provide some detailed information on a given date may nullify the contractual obligation placed on the supplier.

## 42.2 PROJECT MANAGEMENT

Even before the equipment goes out to tender, the job of project management must have commenced, to coordinate all the separate aspects associated with the provision. Equipment room accommodation will be needed (some to a clinical standard) – built or refurbished by a separately contracted company; power and standby generator equipment may need to be purchased and installed; and all sorts of changes in the network may need to be undertaken.

During tendering, a sound project management keeps each task on schedule, extracting prompt replies from manufacturers, insisting on rigorous evaluation and timely contract placement. At the next stage the job becomes one of overseeing the manufacturer's work and quality, and of coordinating the supporting in-house or other contractors' activities.

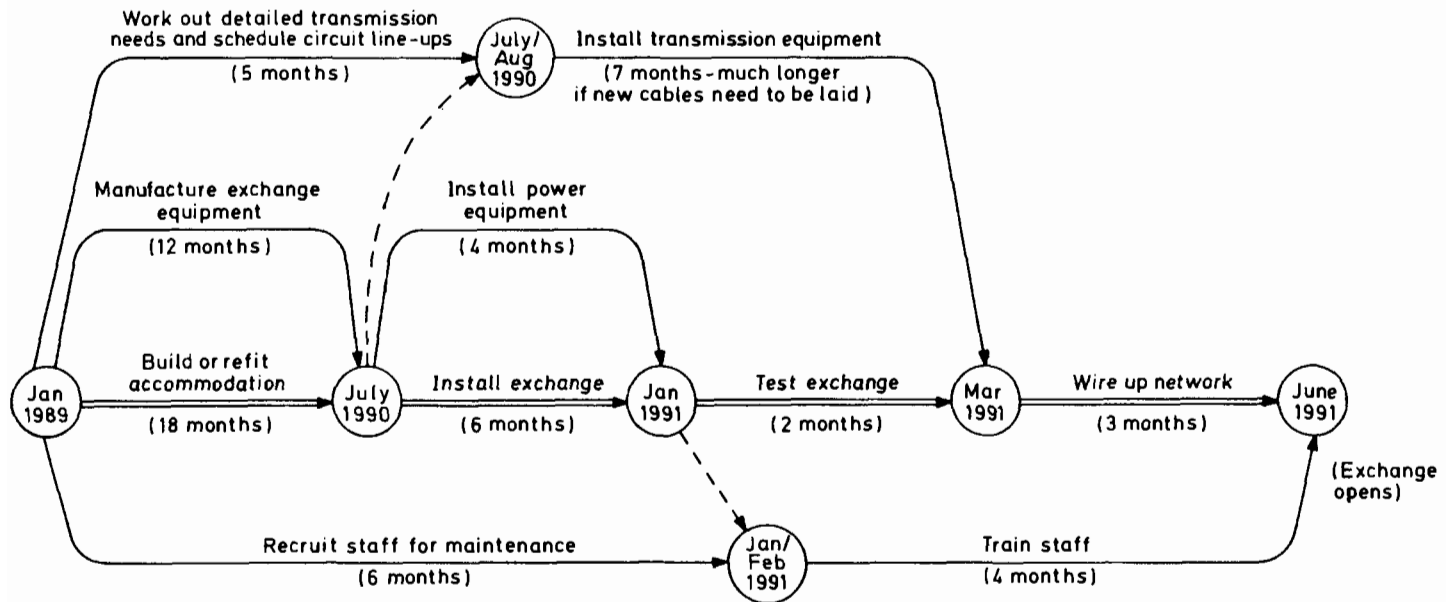
Later on in the process, prior to the *contract completion date (CCD)*, the purchaser should test the equipment, especially any new developments, to check conformance with the specification before final *acceptance* and payment.

The key to successful project management is early project planning. This must be thorough and realistic. Forgetting some activities, or programming too much or too little time for other activities, can seriously upset the smooth running of the entire project.

It is valuable during the project planning stage to undertake a *critical path analysis*. This sets out the relationship between each of the sub-projects, laying out the full sequence of *events* (or sub-projects) and the dependency of any individual *event* start date on the completion of an earlier *event*. For example, it would not be possible to fit-out a building before it had been built, but it might not matter whether the electrical wiring or the plumbing were to be carried out first. This consecutive sequence of interdependent sub-projects determines the earliest date at which the project could be completed. If any of the sub-projects were to take longer than scheduled, then the earliest possible project completion date would slip by at least the same number of days. Figure 42.2 illustrates a simple critical path analysis project schedule. A number of computer software packages are available for this purpose (e.g. *Microsoft Project*).

The exchange installation project depicted in Figure 42.2 comprises ten sub-projects or *events*. Four of these *events* make up the *critical path* of the project. These are the building and fitting out of the accommodation, followed by the exchange installation, the exchange testing, and the wiring-up of the network. None of these activities can commence before the previous one has been completed, so that the consecutive time period required for their completion determines the minimum overall project length. Slippage in any of these four sub-projects delays the project's earliest completion date (i.e. the exchange opening).

In contrast, the other six events do not lie on the critical path. This means that a degree of slack time is available. For example, a total of 29 months is available to recruit and train maintenance staff, but only 10 months are required. These tasks could therefore be delayed for 19 months after the project start date without affecting the project end date. In fact, the only constraint in this case is that the training of staff cannot commence before January 1991, after the exchange installation, otherwise there is no equipment to be trained on. This relationship is shown by the dotted line arrow on Figure 42.2, indicating a *dependent event*. The use of a dotted line arrow indicates that it is only a *one-way dependency*, rather than two-way. In other words, the exchange



## Key

- subproject
- subproject earliest and latest start dates
- == critical path
- - dependent events

Figure 42.2 Simple critical path analysis for exchange installation project

must be installed before training can commence, but in our example it is not the case that the maintenance staff need to have been recruited before exchange testing can commence.

Other events which are not time-critical are the ordering of circuits and the installation of transmission equipment. The ordering must precede the installation, but the ten months needed for the work are far less than the 26 months available. The transmission equipment installation cannot commence until July 1990 at the earliest, after the accommodation has been prepared.

Not only does a critical path analysis provide a checklist of all the tasks to be undertaken, it provides a straightforward and mechanistic method for marshalling resources and ensuring the timely completion of critical path events. Furthermore, it allows the scheduling of non-critical events in the most efficient and convenient manner possible within the timescale. For example, staff training in Figure 42.2 could start at any time between January and February 1991, according to convenience.

### **42.3 PROCUREMENT POLICY**

Before purchasing items of equipment, it is prudent for any company to determine a proper procurement policy. A procurement policy protects the interests of the purchaser. It is no good always buying the cheapest equipment if this means the eventual acquisition of a diverse range of slightly incompatible equipment, each piece requiring different spares and differently trained maintenance staff, and with insoluble compatibility problems should the equipments ever need to be integrated. On the other hand, if a company always purchases equipment from a single supplier, then dependency creeps in; the supplier's prices may escalate to give a better profit margin; worse still, the supplier may go out of business or become unreliable or uncooperative. Using even a small range of different suppliers helps to maintain competition between them. An example of a complete equipment purchasing policy might be that 'new equipment purchased should be compatible with an IBM equipment'.

At the time of tendering for equipment, due account should be taken of the company procurement policy. It makes poor sense to write a specification which precludes all but one supplier, when the company procurement policy is to promote competitive tendering, with a long term objective of maintaining at least two suppliers. Overworded and complicated specifications tend to have the effect of precluding suppliers, so that in general it is best to try to keep conditions of tender and contract as simple as possible.

### **42.4 PLANNING DOCUMENTATION**

Before commencing any major project, whether it involves procurement or not, the network operator must determine the long term network strategy or 'vision' of which the project forms a part. The strategy should set out the long term objectives for network development, and against the backdrop of short term constraints, it should

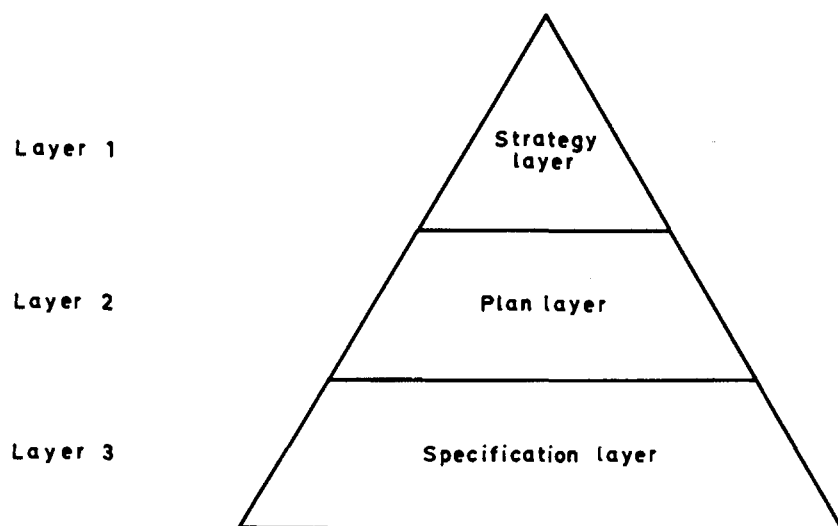
map out a series of projects to achieve the long term goal. Typically, the strategic goal may be reached in a number of different ways, and the analysis leading up to a recommendation of a particular plan should consider and reject a number of alternative options.

Having laid out the long term strategy and mapped it into a series of component projects, the planning, coordination, and documentation of each individual project must commence. Any project is best documented in a controlled and layered structure to ensure easy reference to detailed information while retaining consistency and oversight across the whole job.

Figure 42.3 shows a schematic documentation structure of three layers, oriented in a triangular fashion. At the top or 'strategy' layer the documentation is light, laying out the overall objectives, the end goal, the resources to be used, and in outline the major component projects of the strategy. At the second or 'plan' layer a number of separate planning documents exist, each comprehensively describing an individual project. The plan needs to cover everything: service and project timescales, staff responsibilities, network topology, maintenance procedures, customer service arrangements, accommodation requirements, sales, procurement, customer billing, training and all other conceivable aspects. Ideally, the document refers to the strategy that it supports, and also explains the existence and inter-relationship of any more detailed specifications which exist to support it. The third and last layer of documents, 'specifications', covers the detailed technical definitions and procedures.

Specification documents are the meat of the project. It is documents at this level of detail that are tendered to prospective manufacturers and are subsequently used for definitive reference. That is not to say that layer 1 and 2 documents are any less important, because they hold the entire project and strategy together.

Given a proper document structure, a document registration procedure should be set up to ensure proper quality inspection of documents. Documents should be checked for



**Figure 42.3** A layered documentation structure



accuracy, comprehensiveness, and clarity by persons other than their original authors. To ensure that there is no confusion over discrepancies between differently updated or amended versions of the same document, a proper re-issuing and change-control procedure should be adopted.

## 42.5 THE TENDER DOCUMENT

The tender document or *invitation to tender* describes the equipment to be supplied. It is prepared by the network-operator whenever a major item of equipment is to be procured. It describes which technical functions the equipment is expected to perform and it includes any commercial conditions which apply to the supply contract. Once prepared, the tender document is sent to one or more prospective suppliers, who consider the requirements and send back formal replies, stating what they are able to offer and the price to be charged if their offer is taken up.

The tender document and the chosen manufacturer's response to it form the documentation of contract. Any subsequent contractual dispute between the network operator and the supplier is resolved by reference to this documentation. From the network operator's point of view, it is important to ensure that the tender document makes the obligations on any contracted supplier absolutely clear. The supplier on his side aims to make it clear in the response, which of the specified items can be supplied and which are not included in the offer. The most harmonious business relationships and the most successful projects are based on accurate understanding. In the remainder of the chapter we describe items that should to be included in the tender document.

### 42.5.1 Commercial Conditions of Contract

In the opening section of the tender document a general introduction should be given, explaining briefly the type of equipment required and in what timescales. This prevents any 'no-hope' supplier from wasting further time by analysing all the detail. The commercial conditions of contract should go on to explain in detail what items of work are the obligation of the supplier, the timescales for delivery of each item, and the terms on which payment is to be made (for example, installments as each part is delivered, or 'only when the whole job is done'). The commercial conditions also need to lay out the penalties for failure to conform with the requirements of contract. These penalties sometimes include an obligation on the supplier to correct errors and may call for a sum of money to be paid to the network operator as compensation for *liquidated damage*.

### 42.5.2 Other General Conditions of Contract

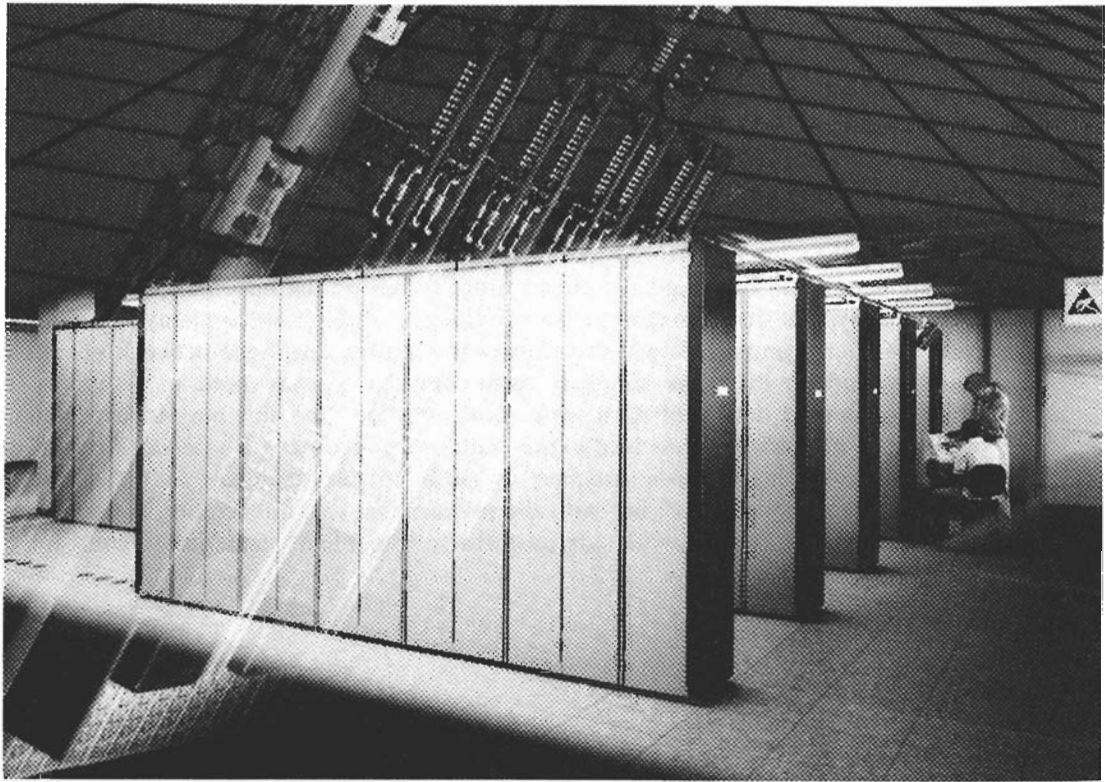
The general conditions of contract may clarify any other legal or general constraints which the network operator might wish to impose. This section might cover the ownership of any document copyright, the ownership of *patent rights* and *intellectual*

*property rights (IPRs)* of any new items that may be developed, and may cover general conditions pertaining to the disclosure of information. The section may constrain the supplier not to make press releases and should cover the procedure for resolution of contract disputes. It could cover both parties' rights to terminate the contract, and (if relevant) the privilege of one party in the event of the other's bankruptcy. Finally, the section may clarify how particular legal constraints on one of the parties affects the contract. For example, let us imagine that the network operator is obliged by law not to release information to a particular third party. This information may, however, be required by the supplier in order that the contract can be fulfilled. In such an instance the supplier should be obliged by contract not to release the information to the third party. This type of situation can easily arise. Imagine a telephone company A buying an exchange from company B, where company B is both a telephone network operating company and an exchange supplier. Now company A may need to reveal to its exchange supplier details of its network configuration, and this may include details pertinent to network business held jointly with one of company B's network operating competitors. In this instance company A needs to release details to the exchange supplier part of company B, but with the proviso that this information is not passed back to company B's network administrators, who might gain unfair competitive advantage from it.

### 42.5.3 Technical Conditions

The technical part of the specification sets out the type of equipment required and its detailed functions. Once upon a time, many of the large *public telecommunications operators (PTOs)* had a hand in designing the equipment. In those days the specification was at a level of detail of an actual equipment design, almost a detailed circuit and component design, so that contracts were principally for manufacture. Nowadays, however, it is more common for purchasers (including PTOs) to provide a *functional specification*, laying out no more than the broad functions the equipment is expected to perform, together with the details of any external interfaces enabling the new device to interwork correctly with other network components. As we have already noted, the use of a functional specification as opposed to an *equipment design specification* promotes competition among prospective suppliers, with a long term benefit of lower costs and higher quality. Items which may be pertinent in the specification are explained below

- accommodation and environment
- network configuration and topology
- operation and maintenance
- equipment size and performance
- functions
- control interfaces
- power supply



**Figure 42.4** A modern ISDN exchange (*Courtesy of Siemens AG*)

The *accommodation and environment* in which the equipment is to be installed: any constraining floor layouts, temperature considerations or other adverse conditions (e.g. siting of equipment in a manhole).

The *network configuration and topology* in which the equipment is to work, include the specification of any interfaces to other network components.

The *operation and maintenance* requirements: whether remote control or monitoring of equipment is required; which test points and test equipment is required; whether internal diagnostic programmes are required; what failure alarms are required.

The *equipment size and expected performance*; for an exchange, the total daily traffic throughput in calls, and the peak hour traffic intensity (number of simultaneous calls) should be quoted, together with performance expectations under overload. The number of *busy hour call attempts* (*BIICA* for circuit switched networks) or packets (packet switched) may be important to the dimensioning of the exchange processing capacity, and the total number of daily call attempts may govern how much storage capacity is required for any statistical information which may be required from the exchange. For a transmission system, the bandwidth of the system or the bit rate needs to be quoted, along with its jitter performance, sensitivity to noise, and other relevant information.

The *functions* of the equipment. For an exchange this includes defining the call control, call routing capabilities and services that are required, together with the method of switching, the method of changing routing data held in the exchange, and the signalling systems to be provided. In addition, recorded announcements may be required, along with particular methods or formats for output of traffic recording, call charging and accounting information. Transmission equipment may need to provide maintenance staff with some means of remote monitoring failures, and of reconfiguring equipment. An automatic switchover to an alternative line system may be required.

The *control* interfaces. It is increasingly common for network operators to expect to be able to control, monitor and reconfigure the equipment within their network from a single computerized 'control centre'. Any equipment destined for such a network requires the necessary computer interfaces to be built in. These must be covered by the specification. Conversely, the equipment supplied may be controlling other devices, and so it too must conform to a defined control interface. (An exchange, for example, may have to control external echo suppressors or circuit multiplication equipment, and a computer may be required to monitor and control any of a whole range of devices.)

The type of *power supply*. The supplier may need to provide equipment for working at a given voltage. Alternatively, the power equipment (including batteries and back-up generators) may need to be provided by the supplier. Often electronic equipment demands *uninterruptable power supply (UPS)*, because computer data can be lost as the result of instantaneous breaks in power.

#### 42.5.4 Reliability Conditions

Reliability expectations should be quoted as a series of measurable parameters, quoted with reasonable target values. Measures such as 'the number of separate instances of failure, measured month on month, shall not exceed...' are much better than more vague parameters like 'the *mean time between failures (MTBF)* shall be greater than three years', even if MTBF is an oft-quoted parameter. The problem with the latter statistic (MTBF) is that the long measurement period that is involved makes it virtually unenforceable. (If two failures happen in the first six months, does this mean that the next failure will not happen for at least five-and-a-half years?) Obligations of the supplier on failing to meet the reliability conditions should also be laid out. For systems like public telephone exchanges or high grade computer control systems, where no break in service is acceptable to customers, *availability* values as high as (or even better than) 99.9% are common, though even more stringent measures may be appropriate. (It is sometimes not hard to meet a 99.9% availability for a network as a whole, even if some individual ports are out of service for days or weeks at a time, because the overall incidence of failure is nowadays usually very low.) Other considerations of reliability are the conditions and prices for the supply of spares and for after-sales support, ideally lasting for the lifetime of the equipment. The network operator does not want to be left in the position, a short time after equipment purchase, where the supplier changes the product range, and stops producing spares for the now obsolescent equipment.

### 42.5.5 Project Management Conditions

The tender document should lay out any special procedures for project management. The purchaser, for instance may make conditions on the work practices of the supplier, constraining noise, mess, access to site or health and safety. In addition it may be necessary to clarify responsibilities: whether for instance the supplier or the purchaser is to wire-up the equipment on site. Particular installation work practices may need to be adopted. Also, for a large and lengthy project, the purchaser may wish the supplier to build check points into the development and installation, so that the rate of progress through the project can be fairly closely monitored.

### 42.5.6 Quality Assurance

The tender document should set down the particular measures that will be applied by the purchaser during *acceptance testing* to determine whether the equipment is fit-for-purpose. Usually a fair proportion of the monies are withheld until these tests have been completed. The tender document may also mandate the supplier to carry out internal inspections at various stages of development and manufacture, and may demand adherence to a quality work practice methodology such as those laid out in the ISO 9000-series standards.

### 42.5.7 Documentation and Training

In some cases, particularly when an *open tender* is undertaken between a number of competing suppliers, the purchaser may choose a type of equipment that is new to the network. In this case the purchaser requires training for the maintenance and administration staff, and full documentation for later reference. This should be stated in the tender document, so that the price for it receives full consideration. Alternatively, it could be the subject of a separate contract.

### 42.5.8 Definitions

Finally, it is usual to include a glossary of definitions within the tender document, as a ready reference for terms of jargon or unfamiliar abbreviations. This is merely good documentation practice. The following list illustrates the parts of a complete tender document, but of course the size of it and the omission of any irrelevant parts are purely at the purchaser's discretion.

- Commercial conditions
- Other general conditions
- Technical conditions
- Reliability conditions

- Project management conditions
- Quality assurance
- Documentation and training
- Definitions and glossary

## 42.6 SUMMARY

In summary, the procurement of equipment can be a relatively straightforward and mechanistic process provided that the necessary planning is undertaken at an early date, and provided that care is taken in preparing the documentation. The commonest causes of project failure stem from faults in the very early stages. The cost of rectification can be enormous.



PART 7

---

# **SPECIFIC BUSINESSES AND NETWORKS**

---





---

## *Meeting Business Needs and Creating Competitive Edge*

---

Electronic mail and messaging systems, customer databases, computer-aided design (CAD) systems, management information systems, voice recognition and response systems, image processing and video systems, together with the telecommunications networks that make them possible, are revolutionizing the way we do business, with profound effects on company success and profitability. Companies nowadays must plan and manage their information technology (IT) resources with as much concentrated effort and imagination as they apply to any other resource or activity. The potential benefits are enormous; costs and lead times can be reduced, along with the manpower and other resource needs, and quality and serviceability can be improved. It follows that a company's IT development cannot be left entirely to the data processing men and women, the computer centre staff and the telecommunications manager. IT has become the concern of the company's top people, and they must build it solidly into the company's central strategy. Too many companies have suffered from the lack of a proper and coherent IT strategy, and many now face the long and costly process either of integrating a variety of small computer or telecommunications systems that have been installed in penny numbers or clawing back the competitive ground they have lost.

### **43.1 CONTENT OF AN IT STRATEGY**

An effective IT strategy combines an understanding of company needs, mission and working practices, with practical understanding of what is technologically possible. Computer and telecommunications equipment should be planned and selected to serve the company's overall interests. The strategy will combine idealistic targets for handling and processing the company's *information*, with a recognition of the strengths and weaknesses of any existing company an IT structure, and will evolve this in a manageable way. Priorities set for development will make allowance for time pressures and short term goals, but while working towards the long term objective.

For some large companies, with the time, money and expertise it may be right to consider immediate full interconnection of all existing computer and telecommunications systems, and by so doing streamline the operation of previously diverse functions such as financial ledger, warehouse stock control, customer orders and enquiries. Such interconnection could speed up the flow of information between various company departments, with a consequent improvement in information detail and accuracy, better customer service, and faster preparation of accounts. However, the systems can only be interconnected where individual systems have been designed and built according to the plan of a previous integration strategy.

Figure 43.1 illustrates a simple network of computers and telecommunications being used to support a manufacturing operation. In total, seven compatible computer systems are in use. Each has its own function and has been designed to carry it out in an optimum way, but each is able to communicate information with the other six systems. Whereas the computer-aided design (CAD) computer relies on working hand-in-hand with the human designer for creative input, the computer controlling the manufacturing plant will instead need to run for long periods, without human interference, and at high levels of accuracy and reliability. The two systems are very different, but need to communicate the design information.

Overall, the computer and telecommunications systems of Figure 43.1 support an optimum company organizational structure and the necessary 'key information flows'. This does not mean that all the systems need to be implemented at the same time in one expensive and traumatic step. Instead the individual systems can be rolled out in an order dictated by company priorities and resources.

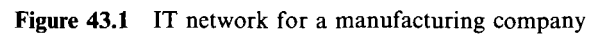
For smaller companies using computing and advanced telecommunications technology for the first time it is right to produce a long term 'information system strategy', but to invest in the first place in only one of the component parts. This allows the company to adapt more gradually to the different working practices demanded by the computer environment.

The study of information flows should not be restricted to those which can be computerized. The strategy also needs to take cognisance of business information, flowing either by word-of-mouth or on paper deliveries. Today's technology allows considerable integration and streamlining of all these modes of information flow, as we shall see.

## 43.2 THE STUDY OF INFORMATION FLOWS

Many companies have recognized the benefits of computerization and are at various stages in the automation of their main *supply chain*. Computers control the taking of orders, the design of products, the manufacture of goods and the scheduling of deliveries. Very few companies, however, have appreciated the need for complete information flow analysis. This is borne out by the huge expenditure on telephone bills, site visits and mailing of documents.

Let us return to our model of information flow methods presented in Chapter 1. The model is shown again in Figure 43.2.



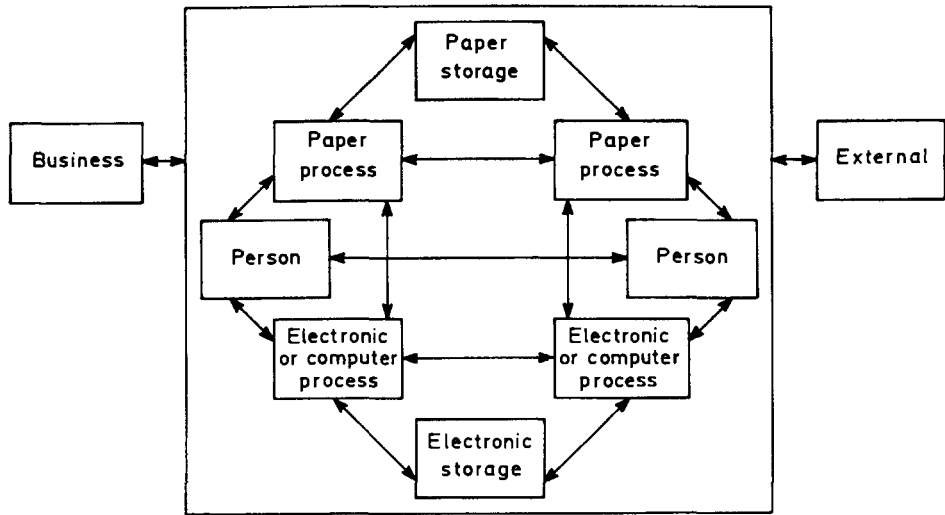


Figure 43.2 Modes of information flow

The model simplifies the issues of information flow, helping us to analyse business processes and to perform value and cost benefit analyses. The model shows three basic modes of information transfer and two methods of storage

- paper flow
- person-to-person flow
- inter-computer (electronic) flow
- paper storage
- electronic storage

Far-sighted companies have probably already set themselves a target to achieve widespread *business re-engineering*, computerization and automation of information flows, but how many have truly achieved it? A couple of examples bring our model to life.

Figure 43.3 shows an example of near achievement of the goal for entire computerization. It is drawn from the motor car industry. Note how all the basic information flows are held, processed or conveyed by computer. There is still interaction with humans for design input and tooling, but there is no human 'bottleneck' or breach in the main flow of information, from right to left in the diagram; the key information can always be retrieved from the computers.

Next we show a far less well-organized company common in many industries. Figure 43.4 illustrates a company rife with manual processes, paper and person-to-person information flows that are strangling the business, even though a considerable amount

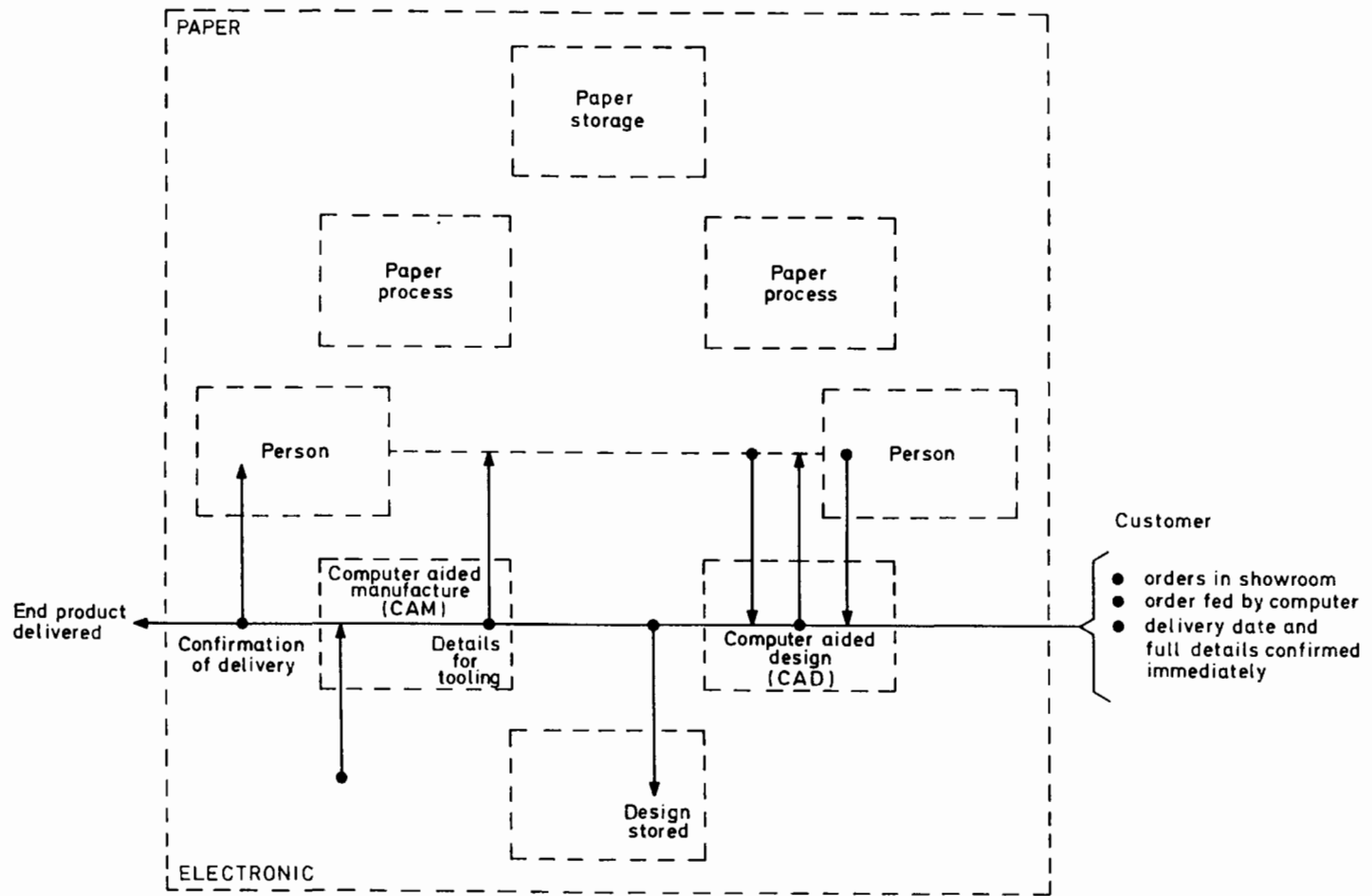


Figure 43.3 Computerization of car manufacture and supply

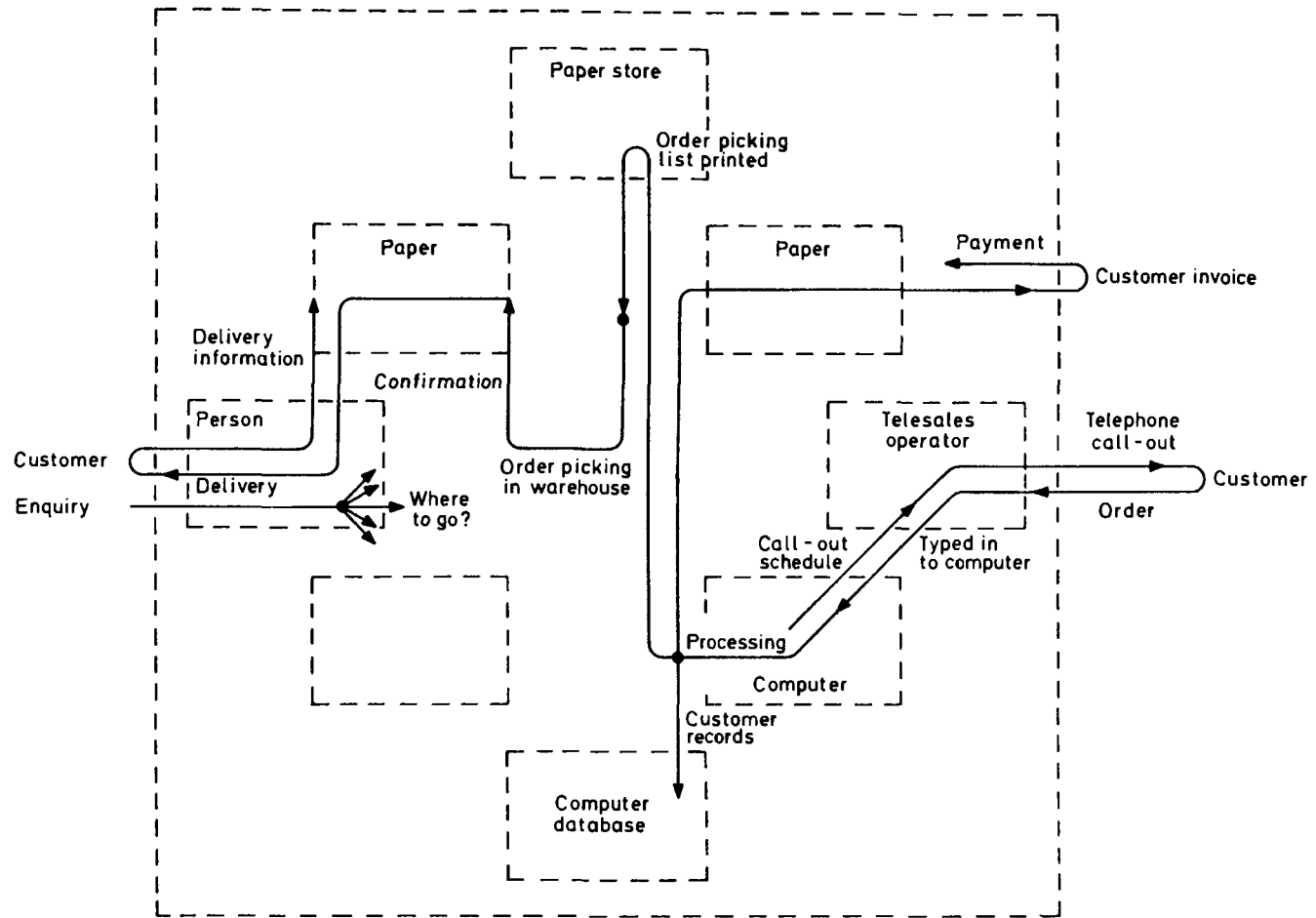


Figure 43.4 Inefficient company information flows

of computing equipment has been installed. The example is that of a simple distribution company. The company takes orders from retail and wholesale outlets and delivers goods by lorry from its warehouse. The orders are taken by *telesales* operators who make telephone calls out to customers according to a scheduled call-out list. These orders are then typed into a computer system, which records customers' invoices, and 'picks' stock from the warehouse. Next, the *order picking list* is printed out on pages in the warehouse, and a delivery schedule is produced for the lorry driver. These enable the goods to be loaded and delivered.

The problems and missed opportunities are many. For a start, by calling the customer over the telephone we miss the opportunity to allow him to order electronically from a PC or some other keyboard device. His order must be made when we ring him rather than when he wants. If our competitor were to ring him up first, we may lose the order. If instead we supply the customer with a PC and a 24 hour communication path then we may be able to eliminate our competitors. A further benefit of this approach would be greater accuracy in processing orders and greater responsiveness to changed orders.

Changed orders, traffic jams and other problems snarl up the ideal schedule. Stock may go missing, be damaged or broken. The result is that the stock available may not meet that demanded by the picking list. Thus what is loaded onto the lorry may differ from what ought to be. Some is broken on route. On an unacceptably high number of occasions, when the goods are actually delivered, they neither match the original order or the pre-prepared delivery note. So the delivery man amends the delivery note by hand, asks the customer to sign it to confirm delivery, and then takes it back to a clerk at the depot, who keys the actual delivery details back into the central computer.

If the invoice is produced at this point, then with luck the customer will not dispute it, but if, as in some companies, the invoice was produced at the time of the original order, then further correction will be necessary to sort out the inevitable customer query. The result is a spaghetti of information flows, messages and enquiries, on paper and by word-of-mouth, trying to sort out the mess. Local 'entrepreneurs' introduce their own isolated personal computer to help with their own departmental problem, and at least the job finally gets done, but not without a heavy penalty of cost, delayed payment and customer dissatisfaction. Customer enquiries are a nightmare because the information is so difficult to trace: word of mouth, computer, paper records?

It is much simpler to identify the important information flows, get them in the correct sequence, and support them with information technology at the right place. In the example of Figure 43.4, a computer terminal in the cab of the delivery truck could enable the delivery note and even the invoice to be produced at the point of delivery. This would mean that both would always be correct and delivered in a timely manner to the customer. No more queries, reduced customer frustration, and timely customer payment.

The lesson is simple and obvious: information flow studies need to address all information flows. These include not only the obvious and ideal flows but also the more 'out-of-sight' ones; a weather eye on the rising phone bill or the cost of mailed complaints gives a telling warning of things amiss.

I hope that the value of Figure 43.2 in analysing business processes is apparent already, but its benefits do not end there. The model also helps us to recognize and



compare similar methods of conveying the same information. For example, our analysis of information flows may conclude that customer enquiries and helpline support are best carried out in a person-to-person manner. We therefore next must choose the best method available; there are of course many

- site visit
- telephone call
- video or *videoconference*
- voice messaging (e.g. answerphone)
- facsimile
- telex
- electronic mail
- cellular telephone

Each method suits different needs, because each has its own characteristics of speed, cost and reliability. The challenge for users is to recognize the differences and make the appropriate choice from the portfolio. Do not be tempted, however, to use a person-to-person technique when either an inter-computer flow is more appropriate or when a paper flow is paramount (e.g. for the recording of a signature).

Many companies could learn from this analysis, perhaps the most significant benefit being the conversion of routine processes (e.g. order-taking) into an electronic form, thereby freeing people to provide personal services which today receive too little priority (e.g. customer account development, and assistance with problems).

As an example of a pertinent 'portfolio' study of inter-personal mail communications, consider the differences between mail, telex, facsimile and electronic mail. All convey a page of written text from one person to another, and have little other value. Mail is probably the cheapest, but the time for a response is several days. Telex is a bit quicker but requires an experienced telex terminal operator. A facsimile requires only a minimum of skills to operate, can transmit even hand-written text, and is often cheaper and quicker than sending a telex. Of all the electronic methods, electronic mail is cheaper on a word-for-word basis for large scale and regular information transfer, it is more reliable and has the benefit of underlying integration with other computer processes.

### 43.3 THE TACTICAL DEVELOPMENT PLAN

Having decided on a strategy and long term goal for use of *Information Technology*, a tactical development plan needs to be laid out, setting down the order in which various specific functions will be carried out, and the overall structure, technical standards, design and supply guidelines to be followed.

## 43.4 BUSINESS APPLICATIONS OF IT

The business applications of IT are wide and diverse. In the remainder of the chapter we discuss some examples of how computer and telecommunications technology has been combined harmoniously to serve business needs in a few important number market sectors.

### 43.4.1 Manufacturing

*Computer-aided design (CAD)*, *computer-aided manufacturing (CAM)* and *computer-integrated manufacturing (CIM)* have enabled faster turnaround times to be achieved, and also much closer quality control. A number of the major car manufacturers, for example, are now entirely dependent on computer-aided design and manufacture. Telecommunications networks link the various design organizations to speed the process of design, and to aid in the rapid programming of robotic production line equipment.

An important OSI initiative in the manufacturing area, undertaken by the United States Society of Manufacturing Engineers, is the development of the *MAP/TOP OSI* profile (*manufacturing application procedure/technical office protocol*). This is a particular set of OSI protocols supplemented with extra application layer protocols to meet manufacturing industry and office networking needs in particular.

### 43.4.2 Distribution and Transport

*Electronic data interchange (EDI)* allows for paperless trading, promising the elimination of huge volumes of invoices, orders and delivery confirmations in favour of much faster electronic methods.

*EDI* requires the use of common telecommunications methods between trading partners (e.g. customers/suppliers), but in addition it requires agreement on the exact format and information content of messages that will be sent. The underlying public telecommunications networks needed for *EDI* are provided by dial-up lines, leased data lines, *X.25 packet* and *frame relay* networks, and by the emerging *X.400 message handling system*; but over and above these standards, agreement is also necessary on the format, syntax and information content of messages.

In the United Kingdom the interim standard for *EDI* is the *TRADACOMS* standard developed by the UK's *Article Numbering Association (ANA)*. The *ANA* is the association set up by the retailing and distribution industry to administer *bar coding* numbers on retail goods. The *TRADACOMS* standards are based on the *United Nations Trade Data Interchange (UNITDI)* syntax and design method.

*TRADACOMS* defines formats for transaction files, orders, stock adjustment, delivery notes, invoices, delivery confirmations, credit notes, pricing lists, statements, availability reports, general text and stock snapshots, as well as the product price and customer information master files. Over 1000 companies across the complete spectrum of retail and service industries in UK use the standards, the largest *EDI* community in Europe.

Another important interim EDI standard is the *American National Standard Institute's ANSI X.12*. This standard is in widespread use in North America. In the longer term *EDIFACT (Electronic Data Interchange for Administration, Commence and Transport)* syntax and message format will be the international standard. The standards will comprise two parts, syntax and message design. The first part, EDIFACT EDI syntax, has already been designed and accepted as a full international standard, ISO 9735 but a number of project teams are still working under the direction of the EDIFACT board to develop all the *UNSMs (United Nations standard messages)* that will also be needed.

### 43.4.3 Retail Trade

*EPOS (Electronic Point of Sale)* terminals are the intelligent *cash tills* that are replacing the simple cash register in many 'chain' retail outlets. The EPOS terminal is a computer which collects detailed information about each sale. By *polling* each shop every night, the central headquarters of a major retailing chain can receive extremely up-to-date and accurate information about the market, its fast and slow selling brands. Figure 43.5 illustrates a simple EPOS arrangement.

### 43.4.4 EFTPOS (Electronic Funds Transfer at the Point of Sale)

These are sometimes also known simply as *EFT* terminals allow retailers to validate credit cards by 'swiping' them through the slot of a specially adapted telephone (Figure 43.6). Card users will be familiar with such machines in shops.

The EFT terminals look like adapted telephones, as Figure 43.6 shows. They automatically make a connection to the bank or credit card company's computer, either directly via the telephone network or via the public packet data network. Any necessary confirmation or response to the retailer is given by a liquid crystal display on the terminal.

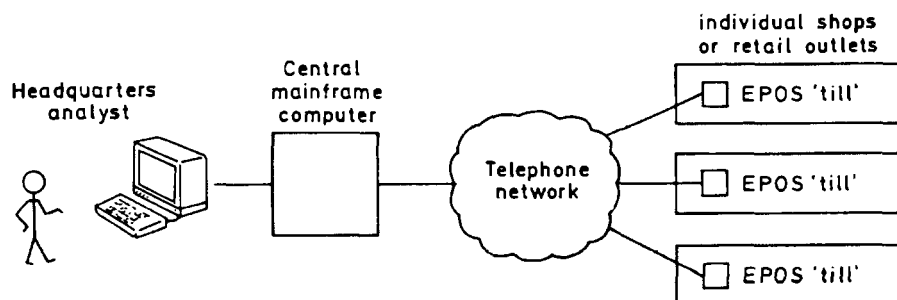
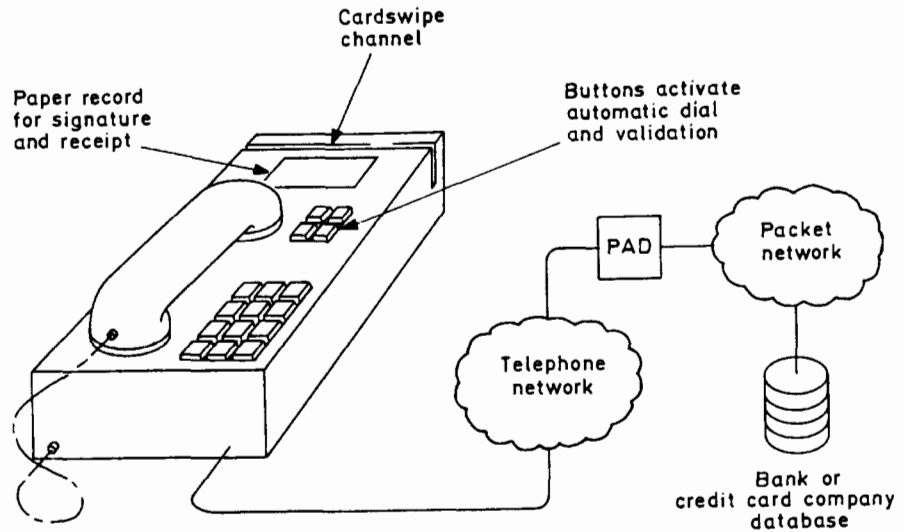


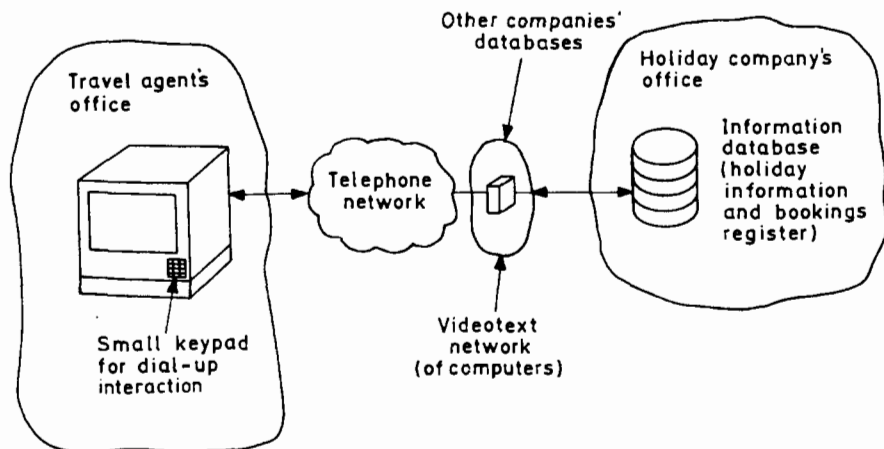
Figure 43.5 Electronic point of sale (EPOS)



**Figure 43.6** Electronic funds transfer/point of sales (EFTPOS)

#### 43.4.5 Videotext

This quickly became popular among UK travel agents. They could get up-to-the-minute information on available holidays, and interaction allows them to book holidays and reserve airline seats instantly. Figure 43.7 illustrates a general schematic of the service. In the diagram, a terminal in the travel agent's office is used to dial up the videotext computer over the telephone network. This computer holds the information and administers any necessary charging scheme for viewing pages of information.



**Figure 43.7** The videotext (or 'Prestel') service

Various proprietary names are used for Videotext Service. Examples are British Telecom's *Prestel* Service, France Telecom's *Minitel* and the Deutsche Bundespost's *Bildschirmtext*, subsequently renamed *Btx*, *Datex-J* and now *T-Online*. Nowadays, the modern version of videotext is the online database service, mostly accessed via the *Internet*.

#### **43.4.6 Satellites**

These are ideally suited to broadcasting sound, video or textual (data) information from a single origin to a large number of outstations, because they allow simultaneous reception by each of the outstations. Thus for large chains of retail stores they are an ideal medium for sending out to all stores price update information, new catalogue information, training information, new product videos and so on. Within the United Kingdom, a consortium of betting shop chains combined forces in the early 1980s to set up a satellite company called *Satellite Information Services (SIS)*. The prime business of this company was in televising horse and dog racing meetings and publishing up-to-minute betting odds and race results. The output is broadcast nationwide by satellite to more than 10 000 betting shops in the UK.

#### **43.4.7 Telesales and Telemarketing**

These are the names given to large scale telephone interaction with customers. Bureaux are often staffed by several thousands of operators equipped with computer terminals and capable of receiving or making telephone calls to customers. Either in the wake of an advertising campaign or on a schedule of 'cold calling', conversations are established for the purpose of promotion, taking customers orders, etc.

#### **43.4.8 Financial Systems**

Dealer board systems comprise a large console and advanced computer and telecommunications components. At the centre of the console, a 'touch-sensitive' screen displays the latest market prices, which are updated by a direct datalink connection to an appropriate database. The touch-sensitive nature of the screen allows the dealer to select topics and receive more information quickly about a particular item, or may record the details of a deal. Dealer boards also usually comprise a number of telephone sets so that telephone calls can be set-up quickly to other dealers (Figure 43.8).

#### **43.4.9 Commerce**

In some large organizations, an 'electronic office' environment has been created, in which each member of staff has a desktop personal computer, word-processor or other terminal which is linked to all the others using either a local area network, or by means

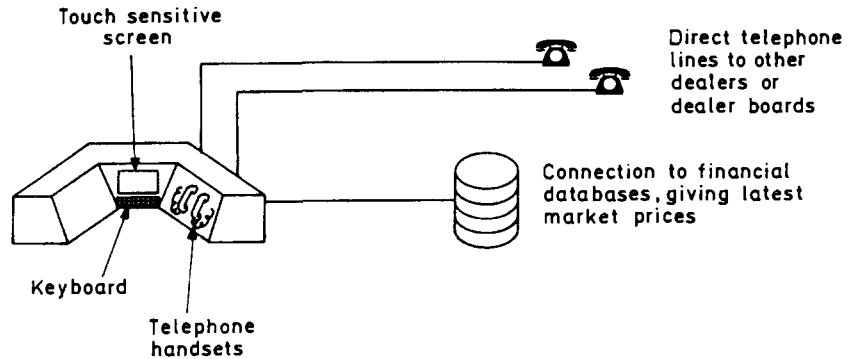


Figure 43.8 A financial 'dealer board' system

of an *electronic mail*, *store-and-forward* mailbox system. More sophisticated office networks even allow the individual users to connect their desktop personal computers to a range of other devices, including the office mainframe computer, various *servers*, the *Internet*, the fax machine, the office user printer, the public packet network, a 'library' database or just other PC users. Figure 43.9 gives a schematic of a simple electronic office network.

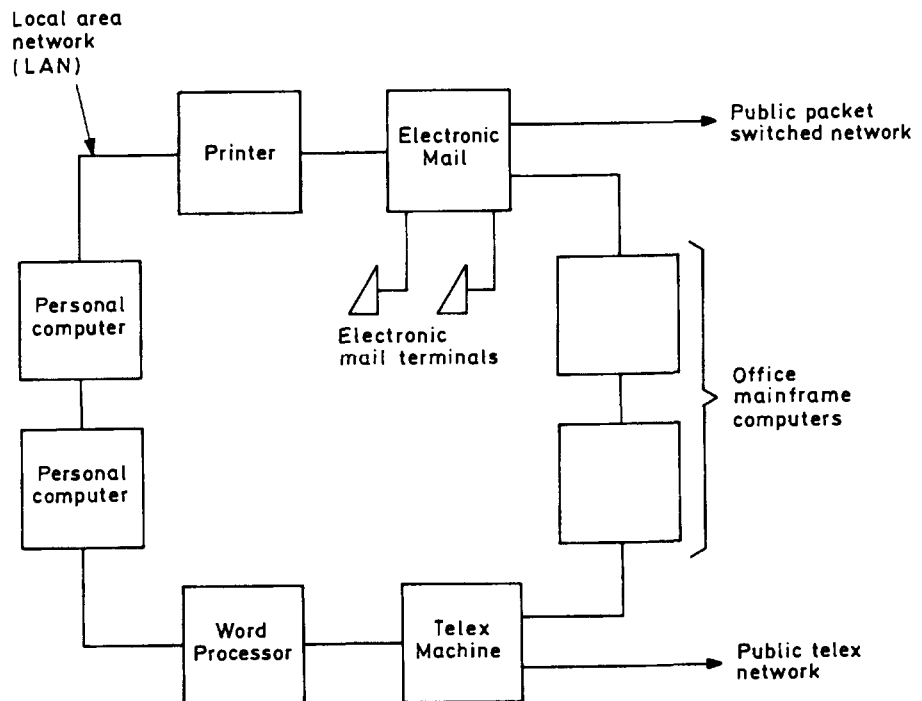


Figure 43.9 An electronic office

#### 43.4.10 Media and Advertising

The media world has demands for fast transmission of high quality pictures and graphical material. Such demands can be met by using the latest video cameras and monitors and by the high speed switched data capabilities available with ISDN and broadband networks. New video-telephones are available to transmit a moving colour picture and its sound signal down a single ISDN telephone line. The same device can send a high quality colour picture in a matter of seconds. Given that the receiver (say an advertising agency's client) has a graphics computer available to accept the colour copy, then it could be *edited* in a few minutes and returned to the advertising agency for copy-printing and publication.

Another new application that eases the transport of high quality paper documents is *group 4 facsimile*. A group 4 facsimile machine can transmit a photocopier-grade A4 page to a similar machine across an ISDN network in 8–15 seconds.

### 43.5 SUMMARY

Correctly configured, telecommunications and computer technology have already stimulated a revolution in the business practices and competitive nature of many industries. Further advances in technology mean that systems with almost human capabilities are no longer inconceivable. The equipment and techniques are complex, and require careful planning to achieve maximum business advantage.

---

## Network Regulation and Deregulation

---

Since the 1980s, governments in hot pursuit of general economic development have made dramatic changes in the regulation of industry. Monopolistic public utilities have been privatized, and the markets historically dominated by those utilities have been deregulated and opened to competition. This chapter discusses the motivations behind telecommunications deregulation, and explains some of the new regulatory measures which protect customers' interests by ensuring that quality services are available at a fair price.

### 44.1 REASONS FOR DEREGULATION

The word *deregulation* is something of a misnomer because no country has completely rescinded all the telecommunications laws and regulations. Rather, the regulations are being changed to a new framework (a *liberalized* one) to encourage competition between companies for the provision of telecommunications services. In fact the new regulations are more voluminous than the old ones, and will require more policing to ensure that companies are conforming with their obligations.

As with any change in the legal system, a government wants to convince itself of the benefit of the new regulation; not only its direct effect but also the secondary effect on other elements of the social or industrial strata. The government will wish to ensure maximum efficiency of overall resources, but it will also be interested in the fairness of a new system. In addition, from an economic sense, the impact on the country's balance of trade will be important.

The *European Economic Community*, *EEC* (nowadays called the *European Union*, *EU*) noted in 1986 that: '*the strengthening of European telecommunications has become one of the major conditions for promoting a harmonious development of economic activities and a competitive market throughout the [European] community and for achieving the completion of the community-wide market for [all] goods and services by 1992*'.



As a result, the European Commission (the 'government' of the EEC) set about establishing a European-wide network infrastructure comprising common network standards and a more open competitive market for telecommunications terminals and services.

The pressure for reform had been building in countries worldwide for many years, but different governments responded at different speeds and in different ways.

The United States led the way in the process of deregulation. As early as 1967, a small company called *Microwave Communication Inc (MCI)* lodged with the US government's *Federal Communications Commission (FCC)* an application to run a *common carrier* (transmission circuit) service between Chicago and St Louis. The proposed quality of the service was lower than that of the available leased circuit service using the established large (and licensed) carriers, a market dominated by the *American Telephone and Telegraph (AT&T)* company and the *Bell Operating Companies (BOCs)*. The 'back-up' for the MCI service was to be minimal, and no subscriber terminal equipment was to be provided (allowing the customer to connect almost anything, as desired), but the benefit was a significantly lower price, being only half the cost of the established service.

Encouraged by the significantly different nature of the service, and hopeful that competition would stimulate new activity from previously untapped markets, the FCC gave approval for the service in 1969, and went on a year later virtually to force the established *common carriers* to provide interconnection facilities. A flood of other aspirant carriers followed in MCI's footsteps, and the regulatory framework for telecommunications within the United States has been changing ever since.

A major milestone in United States deregulation was the establishment in 1986 of *equal access* (Chapter 28) for all *trunk* (or *toll*) telephone carriers. As the phrase suggests, the laws of equal access are designed to promote fair and equal competition between the long distance *inter-exchange carriers (IECs)*. They demand that local telephone companies connect in an equal manner with all licensed carriers who request interconnection. The effect is that local telephone subscribers have a 'free choice' of their *long distance* carrier.

Outside the United States, the tide of change had been slower until the 1980s, at which time a number of factors brought pressure on many of the world's governments to consider deregulation. The pressures were

- the convergence of computing technology and telecommunications (the *Information Technology (IT)* explosion), and their contribution to national wealth and economic growth
- the reduced costs and much increased service capabilities made possible by technological progress
- the high customer expectations, and in particular the demands of business customers
- growing customer dissatisfaction with the high prices charged and limited services made available by *public telecommunications operators (PTOs)*
- the willingness of private companies to invest *venture capital* into developing new telecommunications business

Together these factors brought about rapid change.

## 44.2 THE DILEMMA OF DEREGULATION

Deregulation poses a dilemma for the governments that consider it. Any government will wish to decide the optimum allocation of the country's overall resources and produce a framework of laws and regulations to encourage both public and private companies to use those resources accordingly. A substantial proportion of a country's *gross national product (GNP)* may depend on telecommunications, either directly from the sale of telecommunications services, or as a result of business conducted on the telephone. The decisions on the regulatory framework governing telecommunications are therefore crucial.

The main dilemma that faces governments is that shown pictorially in Figure 44.1. The question is whether a public service need is best served by a small number of monopolized but strongly regulated public utility companies which have an advantage in their homogeneity of service offerings and in the prices offered over a large geographic area, but tend to be inefficient and slow to respond to change; or whether the public need is better served by a more deregulated market, with a larger number of more diverse companies and services, competing for business. The smaller, competing companies will tend to respond more quickly to technological changes and fluctuations in the demands of the market. They will also tend to be cheaper to the customer, but may be so because of a compromise in quality. Furthermore, viewed from a government perspective, the companies of a competitive environment may not appear to use the country's resources efficiently, and may not appear to have an incentive to serve the needs of some crucial but unprofitable markets (e.g. remote rural telephone service).

The historical argument has been in favour of a small number of *natural monopolies*, as we know. Candidate services for a *natural monopoly* are those in which

- there is a clearly established public need for a service of a uniform nature and covering a widespread geographic area

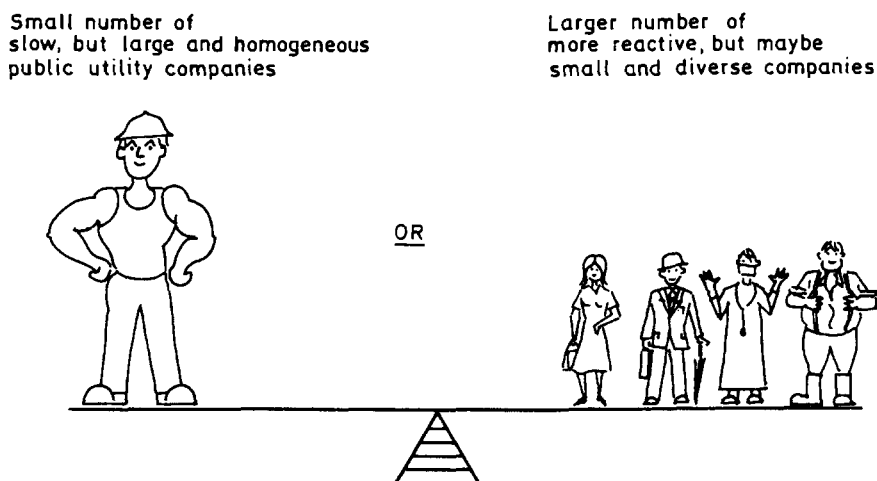


Figure 44.1 Deregulation and the monopoly dilemma

- there is economy in large scale operation
- it is cheaper for one firm to serve the market than for two or more
- the market power of a large company in competition can enable it to control prices and make entry unattractive to rivals
- the lack of information available to new market entrants and the limited resources of small companies leave them at a significant disadvantage

The services which have tended to be protected by governments as *natural monopolies*, are electricity supply, gas supply, water supply, sewerage, refuse and telecommunications. All have to some extent benefited from the large scale of the operation that has resulted from monopoly. In telecommunications networks the addition of each new customer to the early networks, sometimes at disproportionate expense, brought benefit to the new customer and to established users, for whom the available extent of connectivity had been increased.

Under a *natural monopoly* it is usual for a government to confer obligations and rights on the companies operating public utility services. These obligations and rights are summarized in Table 44.1.

The rights and obligations of public utility companies have historically been closely policed by governments keen to dictate the services which may be offered, the prices which may be charged, and the profits which may be made. Laudable though this may seem, it slows up the development of the public utility companies and their services and only adds to the inefficiencies, because more bureaucracy is involved in each decision made.

The aim of improved efficiency has been a prime motivator for *deregulation*, the aim being to 'shake up' the established monopoly carrier by introducing competition. However, most governments' actions have fallen short of removing all regulations and legal entry barriers, so that some financial protection exists for the established carriers to go on providing unprofitable, but necessary services. The challenge is to create a new regulatory regime that is simple enough to promote competition, increase efficiency and reduce prices, but without introducing the plethora of legal and bureaucratic barriers that seem necessary to protect unprofitable services. Ideally, the regulatory framework should encourage a greater variety of services, a greater responsiveness to changing customer demands, better quality, and lower costs; all made possible by more efficient use of resources and by new private investment.

**Table 44.1** The regulation of large public utility companies

---

*Obligations*

- 1 To provide service to all who apply, even if unprofitable
- 2 To provide safe, reliable and adequate service
- 3 To provide fair terms and prices for service provision

*Rights*

- 1 To make money and a fair return on investment
  - 2 The right to withdraw services under given conditions
  - 3 In a competitive environment, the freedom and scope to use initiative to create a 'market edge' for their services, and be fairly financed for unprofitable services
-

### 44.3 OPTIONAL METHODS OF REGULATION

There are two main methods by which a government may impose regulation on industry. The government can control the *structure* of the industry, dictating either what companies may produce or what services may be operated. Monopolies are often structure-controlled. Alternatively, the *de-regulated* approach is to control the *conduct* of companies laying out the actions and behaviours which are permissible.

Most regimes comprise elements of both structure and conduct regulation, and real measures are not always easily classified into one of the two types.

Structure regulation controls which companies may operate in an industry, typically restricting entry by licensing operators, and by having a strict monopolies and company merger policy.

Conduct regulation controls include those governing restrictive practices and those prohibiting 'anti-competitive' acts (such as financial cross-subsidization between two different company products).

Conduct regulation requires a 'policing' body to ensure that companies conform with their obligations. Under *structure* regulation, such a body is not necessary, due to the smaller number of active companies, and also because the public utility company is often state-owned anyway.

In countries undergoing telecommunications deregulation, the shift in the emphasis of regulation tends to be away from *structural* control, towards *conduct* control. Many countries are establishing *regulatory bodies (regulators)* to oversee the conduct of companies in the industry. In addition, it is common for the government to sell off part or all of their ownership of the established monopoly company. The main reason for this is to put the established carrier on an equal business and commercial footing with other new telecommunications market entrants. It affords the company more scope to manage itself, and allows the government to handle competing companies more equally.

### 44.4 TYPES OF REGULATORY BODIES

The regulatory body needed for policing *conduct* regulations can take one of three forms

- a government department
- an independent body, financed by the government or by the industry
- a self-regulating body, composed of a council of delegates from companies active in the industry

Central to the decision as to which form of regulatory body is appropriate, is the question of how much information is available to the regulators. A regulator cannot perform the task if unable either to collect or to comprehend information easily.

Government and independent organizations have the benefit of neutrality and their desire to see 'fair play', but have less information and expertise available to them than would be the case for a body formed of delegates from active companies. On the other

hand, self-regulating bodies (such as the UK's *Law Society*) have more information and expertise available, but only work fairly when the body has an interest in preserving a good reputation. Customers, for example, might not respect the rulings of a self-regulatory body, set up to control conduct in the second hand car market.

Countries may choose different types of regulatory body to oversee telecommunications, but in the United Kingdom and the United States the regulatory bodies are linked to the government. *Oftel* (the *Office of Telecommunications*) is linked to the UK *Department of Trade and Industry (DTI)*, while in the United States the *Federal Communications Commission (FCC)* reports directly to the US congress. This is now the model shared by most European countries, Australia and Japan. A notable exception is the regulatory structure in New Zealand, where all telecommunications-specific regulation was removed, leaving only the standard industry regulating bodies such as monopolies commission, fair trade, etc. Sweden also largely emulates the New Zealand model.

#### 44.5 DESIGNATION OF 'CUSTOMER PREMISES EQUIPMENT (CPE)'

In nearly all countries, the first step in deregulation was the opening of the customer apparatus (or *customer premises equipment (CPE)*) market to free competition. To do so, the *public telecommunication operator (PTO)* is mandated to provide a line from the telephone exchange only as far as a standard socket in the customer's premises and is deprived of the exclusive right to provide the first telephone handset. The customer is free to purchase terminal apparatus (i.e. a telephone, fax machine or other equipment) from a high street store or some other supplier. He may connect it to the network, provided that it is suitably approved (e.g. marked with a 'green dot' or labelled with an 'approved' sticker). In countries with open customer apparatus markets, fierce competition prevails, promoting a wide diversity of available equipment and low prices. Only a minimum of regulation is necessary in this market. In general, governments have found it adequate to set up *approvals bodies*, to verify that new terminal types conform to the network interface and safety standards. This is done at the design and prototype stage of terminal product development, and each manufactured unit then bears an 'approved' marking, although it is not individually conformance-tested.

#### 44.6 DEREGULATION OF VALUE-ADDED SERVICES

It is not feasible or desirable to open all parts of the market to competition, because certain parts of it can benefit from being a *natural monopoly*. On the other hand, regulation must include steps to control the market power of the large established company and its ability to prevent new companies entering the market.

In telecommunications the problem has been tackled by attempting to separate the natural monopoly segments of the market (e.g. the transmission network) from the parts which might benefit from competition (the *value-added network services (VANS)* and *customer premises equipment (CPE)* markets).

The definition of *value-added services* or *value-added network services (VANS)* varies from country to country, but it is generally worded to cover any service falling outside the definition of a 'basic transmission service'. This definition allows entrepreneurial companies to establish services at a premium price based on network and transmission equipment leased from the PTO. Examples of value-added services include

- entire *managed networks* (these are networks in which a VANS supplier contracts to provide services akin to those on a 'private' or *corporate network*, in essence taking over the role of the customer telecommunications manager); the added value is the 'management' provided
- recorded information services (e.g. 'speaking clock' or share price news)
- store and forward message service
- telephone conference service
- voicebank service (message recording service)

In nearly all deregulated countries, providers of *value added service* are not usually permitted to provide the transmission equipment between premises, though sometimes they are permitted to *resell* the 'basic services' (e.g. the simple telephone service), at a straight markup. In some countries, basic services are protected as the sole province of the public telecommunications organizations (PTOs). The situation may change in due course. In the United States many of the long distance telephone carriers established themselves in the mid-1980s by reselling telephone service based on their own purchase of freephone and leased line facilities. They used the bulk traffic discount available on the latter services as a way of creating a margin for under-cutting the normal telephone tariffs. Unfortunately, the quality of such services was sometimes seriously impaired. Other observing governments have steered their 'deregulation paths' to avoid this, but resale is bound to be permitted when the markets are more mature. In the UK, the government gave the PTOs British Telecom and Mercury Communications five years respite from the *simple resale* of telephone service when they were first licensed in 1984, but on 1 July 1989 decided that this protection was no longer necessary, because both operators had been given plenty of time to prepare for the greater competition demanded by major users.

## 44.7 COMPETITION IN BASIC SERVICES

As a pre-requisite to the licensing of a new carrier, many new legal regimes demand massive financial strength of the new entrant, since heavy capital investment is required to establish a viable network, and the government is keen to ensure protection of both investor and customer interests. For this reason, in the United Kingdom only one new major PTO was initially licensed to compete with the previously established operator, *British Telecom (BT)*. The first new PTO was *Mercury Communications Limited*. It received a similar licence to British Telecom, enabling it to provide transmission plant and 'basic' telecommunications services. It is backed by the might and telecommunications expertise of its parent, Cable and Wireless, but when originally set up it also had

the financial backing of *Barclays Bank* and *British Petroleum (BP)*. Since the early 1990s many more companies have been licensed, but in each case Oftel needed to confirm the expertise and financial credentials of each.

In the United States and Japan scores of *toll* (or *trunk*) carriers are now operating, as well as a small number (3–10) of international carriers. Other countries are following suit. In the countries of the European Union (EU), for example, a council mandate requires complete deregulation of the public telecommunications market by 1 January 1998. The refusal of public operating licences by a national regulator is only permitted where the competence or financial strength of the company is in doubt or where resources (e.g. radio bandwidth) are not available in sufficient supply to support the proposed business concept of an applicant new operator.

## 44.8 THE INSTRUMENTS OF PTO REGULATION

Historically, PTOs tended to operate in all areas of the market; not just in basic service provision, but also in the value-added service domain and in the provision of customer apparatus. Special regulatory measures may therefore be necessary so that companies newly entering the market are not disadvantaged because they are small, or because they only compete in one domain of the PTO's business. The regulations ensure that new companies can gain a fair foothold in the market. Two forms of conduct regulation can have this desired effect

- *price control*: to prevent cross-subsidy of PTO services
- *enforced separation* of the PTO's business units into 'basic network services', *value-added services*, and 'provision of CPE': to prevent unfair advantage because of the greater availability to the ex-monopolist of market information

An effective but flexible method of *price control* is based on the formula shown in Table 44.2.

By applying the formula not just to the price of a single service, but also to the prices of a 'basket' of basic telecommunications services, a degree of flexibility is permitted to

**Table 44.2** PTO price control

---

<b>Price rise limited to:</b>	$RPI - X + Y$
RPI —	Retail Price Index (Government index of inflation)
X —	expected potential for PTO cost reduction from its own improved efficiency
Y —	reflects the increase expected in uncontrollable PTO costs. (This factor is more pertinent to other public services, e.g. electricity generation, where the cost of oil/gas is outside the generating agency's control.

---

the PTO, to temper price rises between a number of similar services. Thus the formula may apply to the overall price rise of a number of services. One service may increase in price at a rate higher than the RPI provided some other service has little or no price rise in compensation.

*Enforced separation of the PTO's business units* into sub-organizations dealing with 'basic network services', 'value added services', and 'provision of customer premises equipment (CPE)' using 'chinese walls', helps to reduce the overwhelming and unfair market power that the PTOs would otherwise have over new entrants in each of the sub-markets. A regulated code of conduct prevents the PTO from *linked sales*, from *unfair cross-subsidization* of services and from *unfair use of information*. A *linked sale* would include a practice such as 'preferential network charges provided that you buy your CPE from us' or 'receive a large reduction on your CPE when you buy a new ISDN line from us'. *Cross-subsidization* is where the price of one service is kept artificially low by the profits from another service. *Unfair use of information* includes 'tip-offs' from one part of the business to another. A 'tip-off' from the part of a PTO which supplies exchange lines to the part selling customer apparatus, would enable the latter to corner the market in CPE, largely preventing potential competitors from bidding.

Figure 44.2 illustrates the enforced separation of a PTO's business units. Each is forced to operate in a manner independent of the others, with *cross-subsidization* and *unfair use of information* being prohibited. In practice such clean breaks are difficult to maintain, and we can expect some arguments, not least over the funding of standards development work (this will affect service and CPE prices) and the ownership of new technology ideas.

Critical aspects to be covered by any new *de-regulated* form of market structure are the regulations covering

- *interconnection* between networks of different operators, to ensure the full connectivity of all customers
- pricing control, to ensure in particular that ex-monopolists are not able unfairly to exploit their dominant market positions

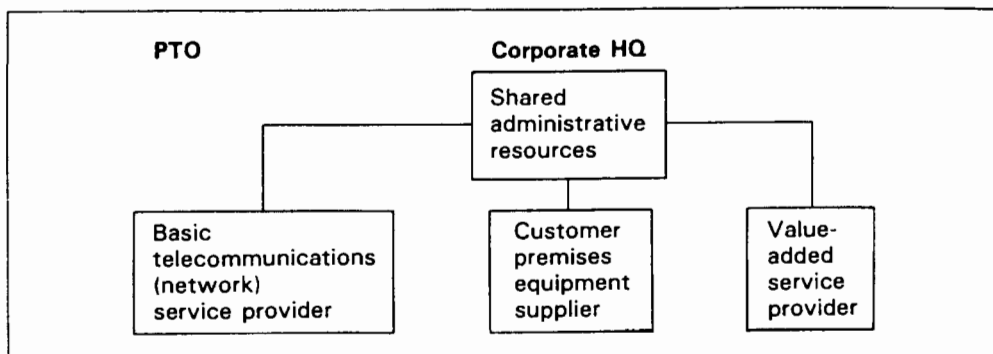


Figure 44.2 Enforced separation of PTO business units



- the securing of *universal* service ('a phone for everyone') and its funding (e.g. by obligating ex-monopoly operators to continue to provide and finance it, to take over the financing from the state, or to establish a *universal service fund* to which all active operators are required to contribute)
- fair means of 'equal access' to customers. Nowadays equal access is expected not only with respect to dialled access of long distance traffic, but also with regard to the use of shared ducts and cables with the ex-monopoly operator with regard to direct connection of customers to the network (we discussed this in Chapter 28)
- the use of radio frequencies and applicable charges for their use

## 44.9 EUROPEAN TELECOMMUNICATIONS DEREGULATION

As part of the EEC programme to generate a single European market for all products and services by 1992, removing all trade barriers between the member states, the European Commission issued a green paper in 1986 laying out plans for a European-wide network infrastructure for telecommunications services. An ambitious programme over the period up to 1992 intended to bring about a community-wide market for customer apparatus and 'portability' of equipment, by setting up a common set of network interface standards and introducing a single approvals mechanism for CPE. In addition, an ongoing period of long term R&D and EEC government investment will ensure harmonization of the constituent national networks.

The green paper of 1986 laid out a framework of recommendations to the governments of member states. The objectives were to

- encourage the development of new services and customer apparatus, so stimulating general economic activity in and between member states
- encourage market conditions favouring innovation
- stimulate the growth of value-added and information services, as these were forecast to have a major impact on the future speed and tradeability of other industrial services, as well as on the potential geographic locations for economic growth
- give wide customer choice
- provide employment and economic growth
- conform with social obligations both within and between states

The framework laid out by the EEC recommended

- the protection from competition of a small number of licensed public telecommunications operators (PTOs) to run the basic services and network infrastructure
- regulated competition in value-added and information services
- European-wide network interface standards and a common approvals procedure for network attachments

- strict guidelines concerning the conduct and policing of these guidelines
- an open market for customer apparatus (network attachments)
- separation of PTO activities by 'chinese walls' between those parts of the business acting as a basic service provider, value-added service provider, customer apparatus provider, and any other part acting on behalf of the regulatory authority (some PTOs have been assigned the job of performing network attachment approvals)
- continuous review of PTO's conduct and performance on basic service provision, to ensure that the monopoly position is not abused
- continuous conduct review of value-added providers

Following the green paper of 1986, the objective of a single market for telecommunications in Europe was endorsed by the *Council of Ministers* on 30 June 1988, and subsequent *directives* (mandates to European Union states) and green papers have been issued covering the regulation of

- terminal equipment
- telecommunications services, according to *open network provision (ONP)*
- network infrastructure

#### 44.9.1 Regulation of Terminal Equipment

For the regulation of *terminal equipment*, a number of directives were issued with the following objectives

- the creation of an entirely open market for connection and maintenance of CPE
- the loss of the PTO's exclusive right to provide CPE by the end of 1990
- the establishment of user access to public network 'terminal/network access' points
- the establishment within member states of independent bodies to oversee technical specification, licensing, type approval, etc. (*approvals bodies*)
- the right of customers to dissolve long term contracts concluded during the monopoly era

The first terminal equipment directive was 86/361/EEC of July 1986, establishing the mutual recognition by member states of equipment type approval.

#### 44.9.2 Regulation of Telecommunications Services, and Open Network Provision (ONP)

In the *telecommunications services* field, a number of directives required

- the removal of virtually all restrictions on competitive *value-added* service provision

- the review (in January 1992) of the need for continuing the exclusive rights of PTOs to provide the basic network infrastructure and the basic telephone service. This review concluded that further steps towards deregulation and further competition even in 'basic services' and network infrastructure should be commenced

A code of rules known as *ONP* (*open network provision*) was established to define the conditions under which all the basic public telecommunications in each member state should be opened up to rival private service providers. The ONP framework is the responsibility of a subgroup of the European government, called *GAP* (*Groupe d'Analyse et Prevision*, in English the Group of Analysis and Forecasting). The goals of ONP (open network provision) are

- to harmonize access to the public telecommunications networks of Europe
- to create a common framework for open use of the networks

The first draft directive (AS ART 100A) covering ONP was agreed in principle by the European council of ministers on 6 February 1990 and the full directive came into force in mid-1990. Three main areas are proposed for harmonization

- technical interfaces
- usage conditions (e.g. provision time, contract period, quality, maintenance, fault reporting, resale, operating procedures of interconnection)
- tariff principles

ONP applies to all PTOs (public telecommunications operators). Ultimately ONP will regulate all types of services and networks, but initial priority was given to leased lines, data networks and ISDN. More recently, considerable effort has also been applied to *broadband*, *mobile* and *intelligent network (IN)* services. ONP demands that PTOs offering these services do so under published conditions which are objective, equal on both parties and non-discriminatory. Limitations in the contract may relate only to essential requirements such as safety and security.

#### 44.9.3 Deregulation of Telecommunications Infrastructure

By the end of the 1980s, *European Union* (by then the *European Community*) directives had already mandated the opening of the *value-added service sector*. This step allowed new innovative service operators to start providing information-type services. More significantly, however, this mandate also deregulated and opened to competition the public data networking sector. This created scope by 1989 for the appearance of the first pan-European X.25-based packet switched data networks and *managed data networks*. The public telephone monopoly and the transmission infrastructure monopoly, however, were allowed to be retained within individual member states, according to national political will.

Some European countries, notably the United Kingdom had by 1989 already deregulated most of the telecommunication sector. Indeed in 1989 the UK regulator decided that further operators beyond the previous *duopoly* operators should be licenced (so-called *duopoly review*). Although other European states were agreeing new proposals for deregulation, these were not so far reaching as those in the UK at the time. The German proposals for the *Post Reform 1*, agreed by the *Bundestag* in April 1989, for example, laid out the steps for re-structuring the Deutsche Bundespost (the post and telecommunication monopoly) and introducing service competition for data services. The public telephone monopoly and the transmission infrastructure remained in place, effectively even blocking corporations from the ability to operate their own 'private' internal company telephone networks.

The continuing will of the *European Commission* for deregulation brought the directive for *Corporate Networks* in 1992. This forced member state governments to relax their telephone monopolies to allow the operation of private telephone networks within corporation *closed user groups* (i.e. internal within the corporation and to suppliers and customers).

Finally, in 1995 the *green paper on the liberalization of telecommunications infrastructure and cable television networks* was approved. This set out the principles, timescale and measures to be undertaken by member state governments within the European Union for 'complete' deregulation of their telecommunications markets. Key points of the green paper were

- requirement for the removal of the transmission infrastructure monopoly by 1 July 1996
- requirement for the removal of the public telephone switching monopoly by 1 July 1998
- specific rules regarding the regulation of communications *content*, specifically aimed at controlling potential dangerous monopolies being established between, for example, television programme producers and cable TV distributors; the content rules also aim to control levels of quality and public decency

Specific details of the green paper and the directives which resulted from it over the period up to early 1996 included

- the requirement for open, non-discriminatory and transparent licencing procedures for new operators
- the requirement for the number of licences not to be restricted, except as determined by 'essential requirements'
- the definition of the selection criteria permissible in the award of licences and authorizations
- guidelines on acceptable regulatory and financing measures for 'safeguarding and developing a *universal service*
- the definition of the regulatory framework for *interconnection* and *interoperability* between public network operators

- the rules of open access to infrastructure, the application of the ONP guidelines and competition rules
- the regulation of rights of way (for laying cables), radio frequencies and network numbering resources
- the maintenance of a national directory
- the requirements for data protection and the guidelines regarding *intellectual property rights (IPRs)*
- the establishment of *national regulatory authorities (NRA)* and their responsibilities

Table 44.3 gives a brief summary of the current state of telecommunications regulation within the European Union member states.

#### 44.10 INSTRUMENTS OF UNITED KINGDOM REGULATION

Within the UK the following legislation applies to telecommunications.

The *British Telecommunications Act, 1981*, and the *Telecommunications Act, 1984*, provide the main legal instruments in the UK. The 1981 act set up British Telecom (the former monopoly network operator) as a private company separate from the Post Office. The 1984 act introduced the framework for competition. Under section 7 of the 1984 act, which allows the *Secretary of State* and the *Director General of Telecommunications (DGT)* to licence PTOs, British Telecom was granted the scope to act as a public telecommunication operator for provision of basic services. Under its licence British Telecom (BT) is required to continue to provide service throughout the UK, to provide rural service, directory enquiry service, public call boxes, and emergency services (999, Police, Fire Service and Ambulance). BT must publish standard terms for the provision of service, it must wire customer premises in such a way that it does not constrain the choice of CPE, and BT is not allowed to undertake unfair cross-subsidization of services. In addition the BT licence gives the regulator a number of powers for undertaking quality checks and ensuring fair trading.

A number of other carriers have also been licensed as public telecommunications operators (PTOs) under section 7 of the 1984 act. The first of these was issued to Mercury Communications Ltd (MCL), the other initial *duopoly* operator. Mercury has an equivalent license to that of British Telecom and it provides competition for basic services, including long distance calling and international network services. Initially, other licences were relatively restricted:

- the Telephone Department of the city of Kingston-upon-Hull (who run the telephone service in that city, now called *Kingston communications*)
- a number of cable television operators
- the cellular radio telephone companies, *Vodafone* and *Cellnet*
- the public cordless telephone operators then appearing (CT2 operators)

**Table 44.3** Current European Union telecommunications regulations and regulators

Country	Legislation governing telecommunications	Regulatory agency	Approvals body
Austria	Telecoms bill due 1996		
Belgium	Two new laws expected 1996 to end infrastructure monopoly and allow operator licensing due 1996	Belgian Institute for Telecommunications	Belgian Institute for Telecommunications
Denmark	Telephone and Telegraph Act, 1987 Total liberalization, 1996	Ministry of Research National Telecom Agency	Telelaboratoriet
Finland	Telecommunications Act, 1987	Ministry of Transport and Communications; Telecommunications Administration Centre	Telecommunications Administration Centre
France	New law, 1996	Direction de la Reglementation Generale de Postes et Télécommunications (DGPT)	CNET
Germany	Telekommunikationsgesetz, (TKG), 1996	Bundesministerium für Post und Telekommunikation (BMPT) Bundesamt für Post und Telekommunikation (BAPT)	Bundesamt für Zulassung in der Telekommunikation (BZT)
Greece	Greece granted 5 year extension to 2003 to expand existing network infrastructure	Ministry of Transport and Communications	National Telecommunications Committee (NTC)
Ireland			
Italy	Telecommunications law, due 1996	Ministero delle Poste e Telecomunicazioni (MPT)	Istituto Superiore Poste e Telecomunicazioni
Luxembourg	Deregulation delayed until 2000		
Netherlands	Telecommunications Structure Reform, 1989 Telecommunications Law	Ministry of Transport, Public Works and Water Management;	Telification B.V.
Norway (not European Union)		Ministry of Transport and Communications; Statens Teleforvaltning (STF)	Statens Teleforvaltning (STF)

Table 44.3 (continued)

Country	Legislation governing telecommunications	Regulatory agency	Approvals body
Portugal	Portuguese deregulation delayed until 2000		
Spain	Law organizing Telecommunications (LOT), 1987, 1990	Ministry of Public Works, Transport and the Environment Direccion General de Telecomunicaciones (DGT)	Direccion General de Telecomunicaciones (DGT)
Sweden	Orientation of Telecommunications Policy, 1989	Post and Telecommunications Authority	Post and Telecommunications Authority
Switzerland (not European Union)		Federal Department of Transport, Communications and Energy; Bundesamt für Kommunikation (BAKOM)	Bundesamt für Kommunikation (BAKOM)
United Kingdom	British Telecommunications Act, 1981 Telecommunications Act, 1984	Department of Trade and Industry (DTI) Office of Telecommunications (OfTel)	British Approvals Board for Telecommunications (BABT)

However, since the *duopoly review* in 1989 many more operators (PTOs) have been licensed.

The terms of each licence constrain the PTO's operation of services and professional conduct.

All telecommunications network and systems other than those owned by PTOs are, by definition of the Telecommunications Act, *private telecommunications systems*. As such they need to be licensed and must conform with the provisions of the relevant licence. To obviate the need for cumbersome licensing of each individual telecommunication system, a number of *class licences* were issued. The initial class licences were the *Branch Systems General Licence (BSGL)* and the *Value-Added and Data Services (VADS)* licence. However, revision in November 1989 of the BSGL meant that the VADS licence became superfluous, and was thus withdrawn. Later, the BSGL was replaced by the *telecommunications services licence* and the *self-provision licence*. The latest versions of these licences were published on 9 September 1996.

The split into licences aimed separately at service operators and at users meeting their own needs (self-provision) reflects a general trend among regulators to focus on the conduct of the different distinct types of network operator.

The *self-provision licence* is the general class licence allowing any user to connect telephones, private branch exchanges (PBXs), or any other *branch system* equipment to the network. It is a 'catch-all' licence which allows every plain old telephone user to use the phone legally. The provisions of the *class licence* are such that most users need not register. Instead, the granting of a licence may be assumed unless there has been a special revocation. The *self-provision licence* applies to the vast majority of telecommunications systems in the UK, i.e. those run in a single building, on a campus or in a *corporate network*. The *network code of practice (NCOP)* which used to regulate the technical aspects of interconnecting private and public networks under BSGl, thereby governing the quality of connections and disturbance to the public network (e.g. caused by re-dialling) under the BSGl is no longer mandatory. Instead there is much greater scope for users to decide for themselves about the quality of connections they are prepared to live with and the reliability of the CPE they wish to connect.

The *telecommunications service licence* covers value-added service providers and network resellers, parties using the lines of the individually licenced carriers to provide public telecommunications services.

Under both class licences only approved apparatus may be connected to a public network. Unapproved equipment may be connected by means of an approved *barrier box* which provides for protection of the network against spurious mains voltages generated in error by the CPE. The technical standards for network interfaces are lodged by PTOs with the *British Approvals Board for Telecommunications (BABT)*, who must approve equipment for network attachment.

Other than the need for network attachment approval and for conformance with normal quality and safety standards, there is little constraint on the market in customer premises equipment (CPE).

Overall, the whole telecommunications market is regulated by the *Office of Telecommunications (Of tel)*, which is part of the UK government's *Department of Trade and Industry (DTI)*.

## 44.11 UNITED STATES TELECOMMUNICATIONS REGULATION

Telecommunications regulation, as all other United States law, is complicated by the two-tier federal (interstate) and state governing bodies. Thus although federal laws and regulations apply, these may be supplemented in each state with further local laws.

The *FCC (Federal Communications Commission)* is the independent United States government agency, responsible directly to *Congress*, and charged with regulating interstate and international communication by radio, television, wire, satellite and cable. The FCC Chairman is designated by the President. The jurisdiction of the FCC covers the 50 states plus the Islands. Its responsibilities and rulings may be categorized into seven broad areas

- radio regulations for commercial and amateur radio
- broadcast services and broadcast stations
- other radio services (e.g. aviation, marine, public safety)



- *common carriers* (i.e. telephone and similar service companies)
- radio licence restrictions
- call signs
- international matters

The prime instruments of federal regulation are

- the *communications act of 1934* and its amendments, notably the *telecommunications reform act of 1996*
- FCC reports and decisions
- FCC annual reports
- FCC registers (of Notices of proposed rulemaking decisions and policy statements)
- the Code of Federal Regulations (CFR)

Copies of all these documents are available from the US Government Printing Office (GPO) at

*Address*

US Government Printing Office (GPO)  
 Superintendent of Documents  
 Washington DC 20401  
 United States of America  
 Tel: (1)-202-783-3238

In particular, title 47 of the code of federal regulations (CFR) details the FCC Rules and Regulations, as shown in Table 44.4.

The *Rules and Regulations* lay down the operating constraints and procedures to be followed by companies providing inter-state and foreign communications services and by customers using them. Applications for construction of new facilities, licence of new

**Table 44.4** The FCC Rules and Regulations (Title 47 of the Code of Federal Regulations (CFR))

Vol. I	Parts 0–19	Commission organization, practice and procedure. Commercial radio service and frequency allocations
Vol. II	Parts 20–39	Common carrier services
Vol. III	Parts 40–69	Common carrier services
Vol. IV	Parts 70–79	Mass media services
Vol. V	Parts 80–end	Private radio and direct broadcast satellite services

services, discontinuance of old services, and tariff amendments for new or existing services must be made direct to the FCC, accompanied by the relevant fee. Policies and rule-making reflect the applications made. The address of the FCC is given in Chapter 40.

Prior to the reform act of 1996 the federal regulations governing telecommunications in the United States had largely come about as the result of a number of major judicial rulings, each plaintiff generally seeking a relaxation in the rules to add to that achieved by a predecessor.

Until the early 1950s the telephone companies in the United States, dominated by the *Bell Telephone System*, dictated the services that were available to customers, and laid down strict conditions about their use. In general, these were quite restrictive, for example even going so far as prohibiting subscribers from attaching subscriber-provided equipment to a telephone company circuit. However, in 1956 a US circuit court of appeals overturned an FCC ruling that had permitted the Bell Telephone system to prevent a subscriber from attaching a non-Bell device to his telephone. The device was the *Hush-a-Phone*, a mechanical device that clipped onto the handset. This commenced a period of relatively rapid regulatory change.

In June 1968 a new ruling of the FCC prohibited the Bell System from preventing subscribers using an acoustically-coupled (i.e. sound-interconnected) device known as the *Carterfone*. The *Carterfone* was designed to interconnect private mobile radio units to the fixed telephone network without special and expensive equipment. The decision opened the way for datacommunication over the telephone network, and in 1969 the FCC allowed independently manufactured (non-Bell) modems to be used on the public network.

By 1976 the FCC had introduced their *equipment registration program*. This enables devices produced by non-telephone company manufacturers to be used on the telephone network, provided that they meet certain stipulated specifications. The specifications are intended to ensure that terminal equipment does not cause physical or electrical damage to the network, does not injure telephone company employees and does not impair the service of other users.

The first real telephone network competition for the Bell System came in 1969 with the MCI private wire service. However, by the early 1980s competition was building rapidly in the *resale* of long distance telephone service, and there was growing pressure to reduce the monopoly privileges of the AT&T/Bell company. The pressure led to the drawn out AT&T *anti-trust suit*, and finally to Judge Green's *modified final judgement* which concluded it. The *modified final judgement* was the vehicle for AT&T (Bell company) divestiture, breaking the company up into AT&T (as a long distance carrier and equipment manufacturer) and seven *Regional Bell Operating Companies (RBOCs)*. The seven RBOCs (Ameritech, Bell Atlantic, Bell South, NYNEX, Pacific Bell, Southwestern Bell and US West) were of roughly equal asset size and were established to take over the local call transport services of the former AT&T/Bell company. The RBOCs, however, were prohibited from providing long distance services and from manufacturing equipment. Meanwhile AT&T was prohibited from local call transport services and from activity as a *yellow pages* operator.

As a toll provider, AT&T competes with a number of other licensed toll telephone service providers, principally MCI and US Sprint. To ensure fair competition between them, the FCC laid down *equal access* regulations governing the access arrangements

with the local telephone companies. These came into full force in 1986, but paradoxically the framework was not seen by all as entirely equal, because AT&T as the established and 'dominant' carrier was necessarily subject to greater regulation to curb its *de facto* monopoly strength, but had an established customer base and considerable assets.

The domains of *toll traffic* and *local call transport* are defined according to whether calls pass between or remain within a small geographical area known as a LATA. About 160 LATAs (*local access and transport areas*) make up the United States. *Toll* calls (inter-LATA) are those which cross LATA boundaries and were, until the reform act of 1996, the sole province of *inter-exchange carriers* (IECs). Meanwhile, local (intra-LATA) calls were (until the 1996 reform act) the domain of local telephone companies. Toll carriers come under federal jurisdiction, whereas local telephone companies are largely governed by the regulations of the relevant state public utilities commission.

Individual state *Rules and Regulations* need to be interpreted separately as they affect inter- and intra-LATA telecommunications services. Application is suggested directly to the relevant state agency.

Federal telecommunications legislation was greatly overhauled by the *telecommunications reform act of 1996*. The main provisions of this legislation were

- the first reform of United States communications law since 1934
- the ending of the monopoly in local telephone services (the RBOC monopoly over local lines and intra-LATA traffic)
- the introduction of greater customer choice in selection of cable TV companies, as both local and long distance providers can now provide these services
- the stimulation of 'advanced' digital television services
- new rules setting out criminal and civil penalties on dissemination of obscene information, particularly on the *Internet*

**Table 44.5** Review of leading deregulated telecommunications markets outside Europe and USA

Country	Legislation governing telecommunications	Regulatory agency
Australia	General policy framework, 1988	Austel
Canada	Railway Act, 1906 valid until early 1990s Competitive basis for National Common Carriers (NCCS), 1987	–
Japan	Telecommunications Business Law, 1985 NTT Law, 1985	MPT (Ministry of Posts and Telecommunications)
New Zealand	Law from 1 April 1989 removed, now subject only to normal industrial trading laws and regulations	Normal trading regulators

- the establishment of the *V-chip*, an electronic chip which would enable parents to control the amount of violence their children are able to view on television
- regulations regarding *universal service*
- controls over media ownership

#### 44.12 OTHER COUNTRIES

Other countries are also deregulating and most follow a common philosophy to that of the United States and the *European Union (EU)*, but each country learns from what it considers were the mistakes of those countries who have already deregulated. Thus, although Japan and Australia followed the European model to some degree, so the European model was previously based on the experience of United States deregulation. The regulations are far from stable. They continue to evolve at a quickening pace, but Table 44.5 provides a brief review. The table should be used only as a guide. It should not be considered to be a full, accurate or comprehensive statement of today's regulations.



---

## *Corporate Networks*

---

Although large multi-national companies can afford to employ specialist ‘telecommunications managers’, smaller companies and individuals are less privileged. For these latter groups, telecommunications are often just another subsidiary responsibility for either a computer services manager or even the general business manager. Nonetheless all telecommunications managers can affect the way in which their companies operate. In this chapter we shall discuss how. The following topics are covered

- telecommunications performance and management
- cabling for telephones and computers
- office computer networking (LANs, etc.)
- private network design and operation

### **45.1 TELECOMMUNICATIONS MANAGEMENT**

The telecommunications manager should manage all aspects of telecommunications. He needs to take ownership of the full cost and functional performance of all information flows, by whatever means, throughout his company. The job should not consist only of designing and operating the company’s private networks. This loses sight of the ‘strategic horizon’, the profound effect that effective all-round communication can have on a company’s operations.

Telecommunications management involves the day-to-day operation and administration of resources, and it should include creative and strategic questioning.

- What is our long term business goal?
- What are the most important ‘information flows’ supporting it?
- Why is ‘key’ information being carried on a particular medium?
- What is the capability of new technology and how can it affect the core business processes to advantage?

- Why should I maintain a private network rather than use public services? Which gives better 'added value'?
- Why are we spending so much money on paper mail?
- Why are we using facsimile rather than electronic mail?
- How can we gain competitive advantage?
- How can we lock-in our customers using telecommunications?
- Which are the right standards for me to use?
- How will legal regulation affect me?
- *Above all*, are our board and senior management properly clued up about telecommunications? If not, is that not my fault?

Good telecommunications management keeps a company using the right technologies, the right suppliers, the right control procedures and the right monitoring measures, maintaining optimum performance at all points in time.

From the suppliers there is a greater need for products which encompass whole business solutions, tailored packages of services, products and support from a single supplier to meet the need of a business process neatly and completely. No longer has the telecommunications manager got the time to bind together a rag-bag of bits, essentially the solution himself. Neither can he afford the time needed to resolve any subsequent problems or take on any necessary enhancements. What he has to do is to identify the need, understand it completely and find suitable suppliers. Thus he seeks a communications solution for 'order collection', 'customer payment', 'customer account management', etc.

Having found solutions to the business communication needs, a framework is needed for performance and overall cost management. We covered this in some detail in Chapters 34 and 35. Finally, the manager may turn to the technology and to the design and operation of networks. The remainder of the chapter, discusses some of the important current issues facing telecommunications managers in this particular area of responsibility

- building cabling
- computer networking
- private network design and operation

## 45.2 PREMISES CABLING SCHEMES

Until the explosion in the number of office computer workstations in the early 1980s, buildings were designed and pre-wired with the needs only of electricity and telephone cabling in mind. However, the large number of personal computers, word processors, facsimile machines, mainframe computer terminals, printers and electronic mail terminals now in use means that the needs of data terminal cabling now generally exceeds those of simple electricity and telephone.

At first, companies installed dedicated wires for each terminal as they were needed; installation work and the costs associated with it mounted up on a 'pay-as-you-go' basis. In time the wiring became a complicated spaghetti of leads strewn across false ceilings, knotted on cross-connection frames, and squeezed into congested conduits. Each time a new cable was needed, the ceiling tiles would come down again, the conduits would be opened up, and somehow another path would be found. The costs began to grow out of proportion. Each new cable was harder to install than the previous one and each new installation led to faults being inflicted on existing wiring as cables were drawn into already congested ducts.

Most companies now favour formally *structured cabling* schemes which are planned and installed at the time of initial building development or during refurbishment. Sufficient conduits, equipment cabinets and cables are provided to cater for wiring in excess of all foreseen needs.

Figure 45.1 illustrates a typical *structured cabling scheme*. It comprises cables running out from a main trunking room or patch frame. First a number of high-grade *backbone* or *riser* cables run out (in building conduits often called *risers*) to *patch panels* (passive components) or to *cabling hubs* (these are 'active' components which may convert electrical formats or interfaces as may be needed) located on a one-per-floor basis. From these secondary locations, a large number of lower grade cables run out to all possible locations within the office floor. It is common nowadays to find glass fibre cables used in the *risers* and *unshielded twisted pair (UTP)* or *shielded twisted pair (STP)* copper cables used within a given building storey.

Typically four four-pair *category 5 (Cat 5)* cables (i.e. eight wires per cable) are run from the floor patch panel to each single occupant office and terminated on eight-pin *RJ-45* sockets, and four *RJ-45* sockets per office allow for up to two telephone extensions and two data terminals per desk. Standard telephone gauge *unshielded twisted pair (UTP)* is suitable for most LANs, data and telephone needs. For higher performance needs, *shielded twisted pair (STP)*, coaxial cable or fibre optic cable may be needed (see Chapter 8).

The floor patch panel provides all the connections on that floor and for later reconfiguration simply by changing the patch wire. *Local area networks (LANs)* of logical *star*, *bus* and *ring* topologies can be created and reconfigured at will. In addition the floor patch panel is connected via *riser* cables to patch panels on other floors and back to the main patch frame. These allows for connection to public telecommunications networks and the mainframe computer installations on other floors. The riser cables are usually either *shielded twisted pair* or fibre.

Though structured cabling schemes require greater capital outlay at first, they usually reward their owners with much lower running costs, very fast and cheap terminal wiring lead times and much reduced frequency of faults. The largest single cause of wiring errors, tampering with the cables in the conduits, is eliminated by a structured scheme. Unstructured (*ad hoc*) wiring schemes should thus only be considered where very low terminal populations exists (less than about 20 terminals).

Structured wiring schemes generally require very little maintenance and operate virtually trouble-free, but good management practice demands that a proper inventory be kept of the use of each wire, and that each be properly labelled.

A number of regulatory stipulations need to be considered during the planning and installation phases of cabling. First, it is not good practice to run cables in air conditioning



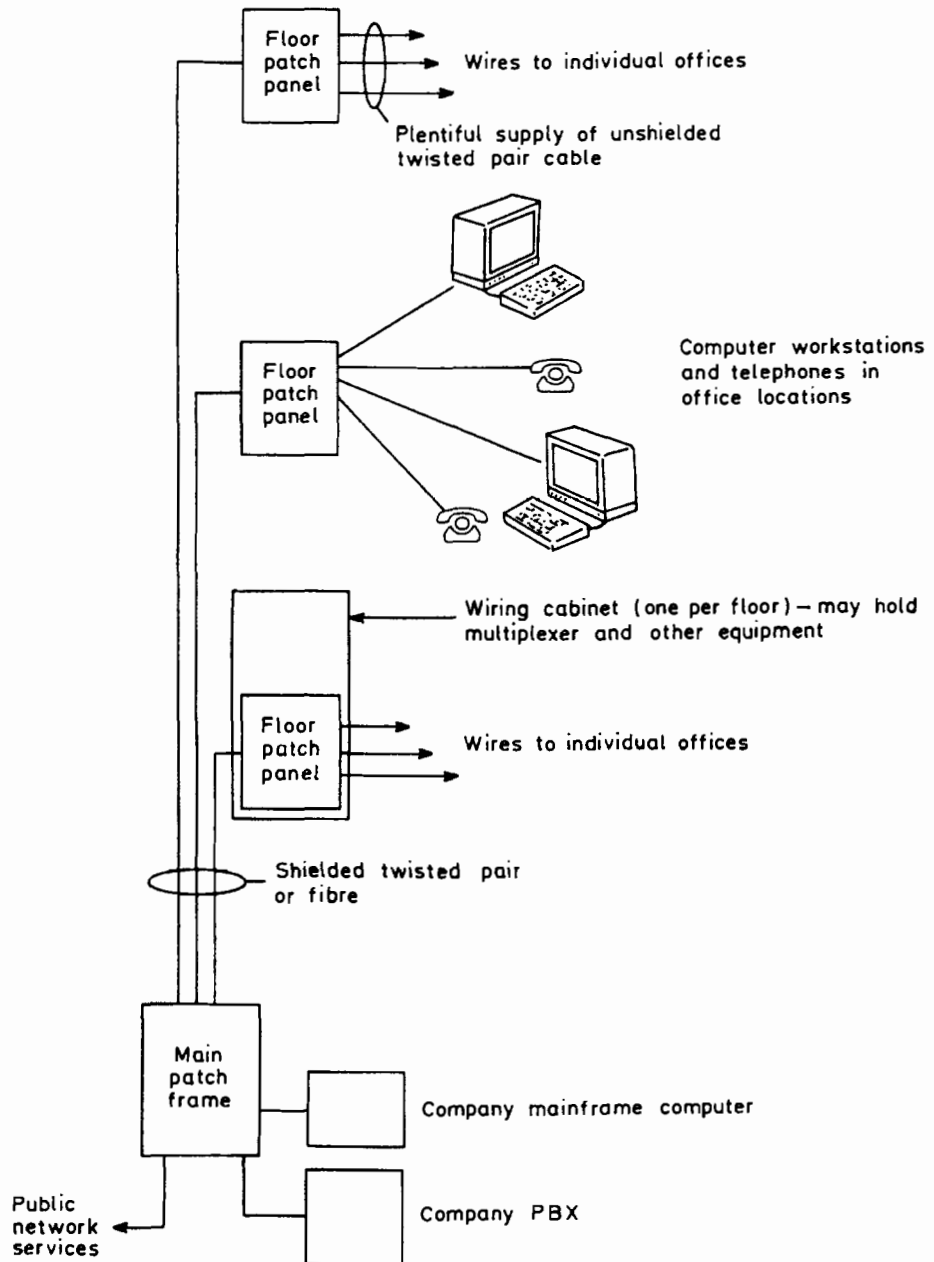


Figure 45.1 Structured cabling scheme

ducts because of the smoke and fire risk arising from the combustion of the plastic insulation. (Actually, United States law permits this but it is illegal in the UK). Separate conduits for electrical and telecommunications wiring are recommended to protect telecommunications users and equipment alike from the dangers of a mains electrical 'shock'.

In Germany, fire regulations require sealing up the inside of the riser conduits at each floor level with cement. This makes the pre-installation of many cables in a *structured cabling scheme* attractive, as the effort and cost associated with drilling through each floor, pulling a single cable and then re-sealing the cement would be very expensive and time-consuming.

However, a regulatory restriction applies in the United Kingdom which may cause companies to consider separate cabling schemes for data and telephone uses. The restriction is that all cables and patch frames carrying telephone wiring which may possibly be connected to the public telephone network (via the PBX) must be maintained by the *registered PBX maintainer*. Furthermore, any data or other equipment connected to wires in the same cable (irrespective of whether these wires are connected to the PBX) must either be *approved* for connection to the telephone network (with 'green dot'), or they must be isolated by means of a barrier box (intended to prevent high voltages leaking onto the public telephone network). These restrictions are easy to avoid by installing separate structured cabling schemes for telephone and data wiring. Thus cables and patching equipment are duplicated, but conduits and cabinets may be shared.

### 45.3 OFFICE COMPUTER NETWORKING

Issues in the design and management of office computer networks are

- meeting capacity and response time needs
- flexibility for reconfiguration
- compatibility of *application* and *network management software*

We discussed network dimensioning to meet data network capacity and user response needs in Chapter 30. Typical response times needed from a real computer network are shown in Table 45.1. Achieving such targets takes a considerable amount of skill; there are far too many detailed factors to be considered which we have not the time for here, but many excellent books, consultancy organizations and suppliers are available to

**Table 45.1** Typical response time expectations

Function	'Maximum' response time
Computer activation	3.0 s
Feedback of error	2–4 s
Respond to identify code	2.0 s
Keyboard entry	0.1 s
Respond to simple enquiry	2.0 s
Request for next page	0.5 s
Respond to 'execute problem'	15.0 s

assist with complex data networking problems. Basic guidance is easier: 'keep it as simple as possible'. Careful design of relevant software, together with the use of the formulae given in Chapter 30 will go a long way!

The flexibility to reconfigure data networks or to meet a number of needs simultaneously is derived from the use of a number of tools and techniques, e.g.

- *structured cabling*
- *local area networks (LANs)* (Chapter 19)
- *open* computer architectures (e.g. OSI and TCP/IP, see Chapters 9 and 19).

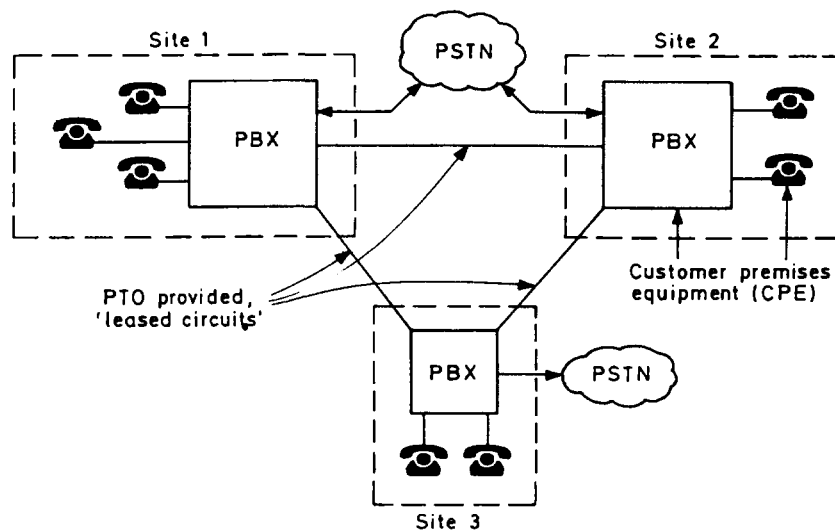
The use of *LANs* or other computer network architectures (e.g. X.25, frame relay, ATM, TCP/IP, APPN, SNA), needs to be tempered by careful preparation. None of the tools are as flexible as they may seem, and lack of appreciation of this fact can lead to equipment incompatibilities.

*Local area networks (LANs)* should be designed and run using rigorous 'mini-data centre' procedures. Out of hand, they can be very difficult to manage and problems can be hard to diagnose.

The selection of a particular LAN (e.g. *ethernet* or *token ring*) or a particular computer architecture (e.g. *TCP/IP*, *OSI*, *SNA* or *APPN*) needs to take cognizance of the applications to which it will be put (e.g. *file transfer* or *printer sharing* or *mainframe database repository*) and also of the computer hardware and software (e.g. *PC operating system software: DOS, Windows, Windows95, WindowsNT, OS/2, UNI, MVS*, etc.) which it is expected to serve. Equipment from a particular computer manufacturer tends to have been developed with one architecture in mind. For example, the IBM architecture is *SNA* or *APPN* and the preferred LAN technology is *token ring*. Meanwhile *Digital Electric Corporation's (DEC)* proprietary architecture is called *DECNET* and the preferred LAN technology is *ethernet*. In the longer term, internet protocols (e.g. *TCP/IP*) and *open system interconnection (OSI)* standards allow greater inter-changeability of computer equipment and network components.

## 45.4 PRIVATE NETWORKS

Public telecommunications networks are usually large and complex, requiring massive capital and manpower resources. As a result, network changes are usually conducted in a careful and steady manner because the cost of mistakes is magnified by the vast scale. Also, because public telecommunications operators (PTOs) are normally required by terms of licence to provide services of uniform availability over widegeographic areas, the development and rollout of innovative and technologically advanced services can be held up, and prices to main business customers can be prejudiced by the subsidisation of uneconomic rural areas. Because of this, and driven by the desire to adopt the latest techniques, some companies have developed extensive *private* or *corporate networks*. The advantage to be gained depends on the service offerings and tariffs of the public telecommunications operator(s) and also on the legal constraints placed locally on such a network.



**Figure 45.2** A private telephone network

A typical private network is shown in Figure 45.2. The example shows three *private branch exchanges (PBXs)* in different office buildings, interconnected by circuits between the buildings. The PBXs and other *customer premises equipment (CPE)* may be owned by the private company, or leased from the PTO. The wiring in each office normally belongs to the private company. However, the circuits interconnecting the different PBXs in different buildings almost certainly are leased from the PTO. Thus a private network usually comprises *customer premises equipment* (or *customer apparatus*) and *leased circuits* which interconnect the different locations.

The *private network* illustrated in Figure 45.2 is capable of switching telephone calls between any two telephones in any of the three different offices. If desired, a single numbering plan could cover all three offices, each telephone having a unique three- or four-digit *extension number*. In this way users benefit from a short dialling procedure, and the company can save money by transporting inter-office calls over the leased circuits rather than having to pay the full public telephone tariff for each call. Provided that the telephone *traffic* between different offices exceeds a given threshold value (the exact value of which depends on the relative leased circuit and PSTN call tariffs) it will almost certainly be cheaper to use direct leased circuits in this manner, as the calculation in Table 45.2 shows. The cost benefit of the *private network* (compared with using the public network) is prone to changes in PTO tariffs. As an example, in the United Kingdom the balance of public service tariffs and leased line rental charges has changed so radically since the early 1980s that many large companies are now considering the abolition of the *private networks* they established at that time. Another factor encouraging companies to give up their private networks is the desire to reduce the in-house manpower needed for the day-to-day responsibilities of network operation. So the pressure is for the public network operators to assume this greater responsibility!

**Table 45.2** Threshold for leased circuit consideration

---

Leased circuit tariff = $\pounds L$ per quarter
Public telephone tariff per minute = $\pounds p$
Telephone usage minutes per quarter = $m$
Public telephone bill per quarter = $mp$
therefore if $mp > L$
or minutes per quarter $m > L/p$
a leased circuit is worth consideration

---

The calculations shown in Table 45.2 are oversimplified. The full analysis should take account of traffic profile effects (see Chapters 30 and 31). Only the point-to-point traffic should be considered, and the number of leased circuits needed needs to take account of the grade of service (following the *Erlang formula*).

Private networks are common in medium and large companies where the economics of large scale helps them to 'pay-in', particularly for intra-company communication between main offices. For communication with suppliers or customers, or even to staff in small remote offices, private networks are often connected to the public switched telephone network (PSTN). In Figure 45.2 each of the PBXs is illustrated with a direct link to the PSTN.

Apart from cost or the capabilities of new technology, another motive for a company to establish a *private network* could be the network security afforded. Take a small private packet-switched or other data network which connects a computer holding sensitive data to a number of remote workstations. Unauthorized users with workstations which are not connected to the private network cannot gain access to the sensitive information.

When it is not possible or economic to establish a private network to ensure the security of information, *scramblers* or *encryption* devices may be employed in conjunction with the public network. *Scramblers* work in pairs, at each end of a connection, converting speech into a coded, but unintelligible signal for transmission on the public line. Equivalent devices when used to *scramble* sensitive data are usually called *encryption* devices.

## 45.5 ARCHITECTURE OF PRIVATE NETWORKS

*Public* and *private networks* are based on similar technical principles, but differ in scale and complexity. *Private networks* generally cater for much lighter traffic, but more specialized service needs. An example of the difference in technical standards is illustrated by making a comparison between the signalling systems used for conveying telephone calls between PBXs over a private network with those used between the exchanges of a public network. Inter-PBX signalling systems often only carry short digit strings (equivalent to a four-digit extension number), but in addition they need to carry information necessary for invoking special features such as *ring back when free*, etc. As a result, the speed at which individual digits of the dialled number are sent

between the PBXs may not be the critical factor, because the small number of digits constrain the call set up time anyway, but the signal vocabulary must be wide to support all the various special functions. By contrast, inter-exchange signalling systems need to carry longer digit strings. A full international number and its prefix could be up to 18 digits long, and must pass quickly between exchanges so that the set-up delay is reasonable. Exchanges in a public network may need to transfer information on charging or routing the call, as well as any customer-requested supplementary services (e.g. *calling line identity (CLI)* or *ring back when free*). Finally, the inter-exchange signalling system may be used to convey information for network administration, including *network management* information and control signals, or for remote operation and maintenance commands.

Public network standards tend to be robust, designed to withstand and control the severe and diverse demands of the wide range of uses to which public networks are exposed. They need to be resistant both to failure and fraud. *Private network* standards are nearly always derivatives of public network standards, and are often developed from the interface used to connect customer premises equipment to public network exchanges. Figure 45.3 illustrates an example pertinent in the UK, where the primary rate ISDN PBX-to-network interface, called *DASS2 (digital access signalling system No. 2)* is similar to the *digital private network signalling system (DPNSS)* which may be used on private networks between PBXs to support ISDN-like services (the ITU-T's signalling systems for the similar purposes are called *DSS-1 (digital signalling system 1)* and *Q-SIG*. These are defined in the Q.93X series of recommendations). The advantage of using similar standards for PBX-to-exchange and PBX-to-PBX interfaces is that it minimizes the complexity of PBXs which are required to operate in both modes. Similar benefits can be gained in other types of networks (e.g. packet networks, etc.) by aligning the technical standards used in private and public network variants. Minor differences are inevitable, have always existed and are bound to persist. They are the results of the contrasting network circumstances and demands on public and private networks. Thus a private packet-switched network is likely to be based on ITU-T's X.25 standard, but it may include a number of customized software features. Likewise the *message handling system*, as we saw in Chapter 23, recognizes separate public and private *management domains*.

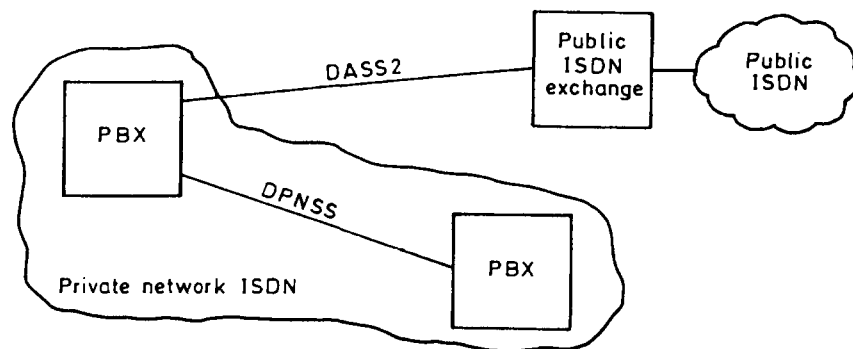


Figure 45.3 PBX-to exchange and PBX-to-PBX ISDN signalling

## 45.6 PLANNING PRIVATE NETWORKS

The smaller traffic scale of private networks, coupled with their relative freedom from licence constraints, means that their structure can differ from that of public ones. Indeed, the optimum topology is not only independent of the real cost of equipment, but can be highly distorted by the PTO's relative tariffs for *leased circuits* and public network services.

Besides the economy measures of Chapter 38, three particular network routing techniques which offer considerable scope for cost reduction to private network planners are

- establishment of telecommunications hubs
- public network overflow
- public network bypass

### 45.6.1 Telecommunications Hubs

The small traffic scale of private networks tends to encourage the use of a *star network* topology in which one, or a small number of exchanges at the *hub* of the network provide a transit switching point between all other exchanges. The configuration is illustrated in Figure 45.4, where all calls between any pair of exchanges are switched via the *hub* exchange A. By employing such a topology a handful of circuits may be sufficient to carry all traffic (let us say 3 circuits from each outlying exchange to exchange A, a total of 21 circuits), as against a more *meshed* or *fully interconnected* network of only 1 circuit between each pair of exchanges which would require 29 circuits (as shown in the inset of Figure 45.4).

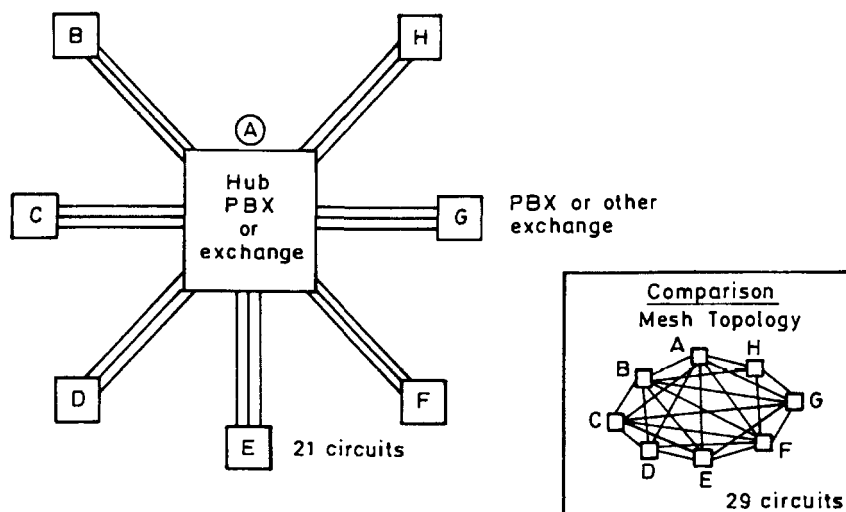


Figure 45.4 'Hub' or 'star network' topology for private networks

The star topology shown in Figure 45.4 is similar to the hierarchical network structure described for larger scale (public) networks in Chapter 32. The smaller traffic on private networks means that the use of the star topology (or *hierarchical structure*) is nearly always more cost-effective than a more meshed network of inter-exchange connections. It not only minimizes the total number of circuits required, but also tends to minimize the total number of circuit miles. Minimizing the number of circuit miles is most important, because it is this value that relates to the *leased circuit* charges levied by the *public telecommunications operator (PTO)*. Careful siting of the hub exchange is also important. In international *private networks*, the choice of the hub site may be especially important, because the price of leased circuits may not be the same in each of the countries. Let us return to Figure 45.4. It might be the case that exchange A is the geographical hub of the eight sites, that the prices of leased circuits into and out of the country in which exchange F is situated are cheaper. In such an instance, it may be more economic to use exchange F as the network hub. Similar anomalies of leased circuit prices can arise even within a single country, caused either by a banded (rather than directly distance proportioned) price structure or because of the differential price structure of competing PTOs.

Circuit numbers are calculated in the normal way for the type of network and the circuit lengths are measured as radial distances. The costs can then be worked out accordingly. Repeating the procedure for different optional networks based on the various available hub sites helps to determine the cheapest configuration.

#### 45.6.2 Public Network Overflow

The traffic between any two exchanges of a network has a 'peaked' profile. However, although public networks are mandated to carry this peak of demand, there is no similar mandate for private networks. In particular, there is no need for the private network to be able to carry the peak demand using its own resources alone. Instead it is permissible, and may be far more economic, to *overflow* the peak demand to the public network. Figure 45.5 illustrates this technique. Two PBXs (or other type of *private exchange*) called A and B, and located at different offices of the same company, are connected together as a *private network*. For most of the day the traffic demand between exchanges A and B is between 6 and 8 Erlangs, but between three and four in the afternoon it peaks at 12 Erlangs.

The private network planner has done some cost calculations. He has established a network of eight direct *leased circuits* on the *private network* between exchanges A and B, but has elected to overflow the excess 7 Erlangs of peak traffic via the public network. The total cost of the configuration is difficult to calculate, because the cost of calls overflowed via the public network must be measured or estimated from the complicated teletraffic formulae presented in Chapters 30 and 32. For each hour of the day the traffic that is carried on the eight circuits direct route, and the traffic overflowing from it, is determined using the Erlang formula, by inputting the traffic demand for that hour. Thus, until 5.00 a.m. there is no traffic and no overflow. From 9.00 a.m. to 3.00 p.m. the traffic is 7 Erlangs, and the overflow from eight circuits is 1.25 Erlangs so 75 call minutes per hour are sent via the public network. Similarly, during the hour of peak activity (3.00 p.m. to 4.00 p.m.) the offered traffic is 12 Erlangs, the overflow from the eight



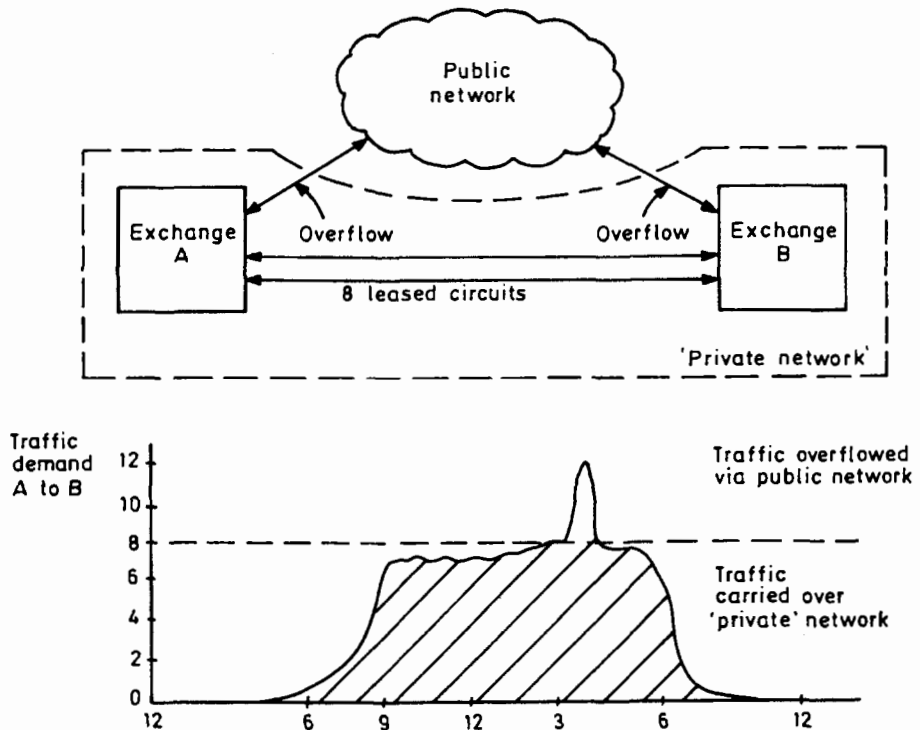


Figure 45.5 Public network overflow

circuits is 5.07 Erlangs, and 304 minutes route via the public network during this period. Over the whole day (adding also the overflowed minutes in the 4.00 p.m. to 6.00 p.m.) a total of 904 minutes overflow daily via the public network. At a cost of £ $P$  per minute and £ $L$  per quarter tariff for each leased circuit, this configuration gives an approximate overall quarterly charge of

$$\text{Estimated quarterly charge} = (L + 3 \times 22 \text{ days}^* \times 904 \times P)$$

(There are approximately 22 business days per month.) As in the last example, the optimum number of leased circuits (for minimum cost) is determined in a trial-and-error fashion by calculating two costs for various configurations.

### 45.6.3 Public Network Bypass

Where permitted by law, public network bypass may be an additional method of deriving additional telecommunications savings. A cost saving is made on long distance calls over the public network by routing the call as far as possible within the private network before handing it over to the public network for completion. By so doing, the public network is bypassed as far as possible, the shorter geographical length of the public network connection reducing the chargeable tariff. Figure 45.6 illustrates the method

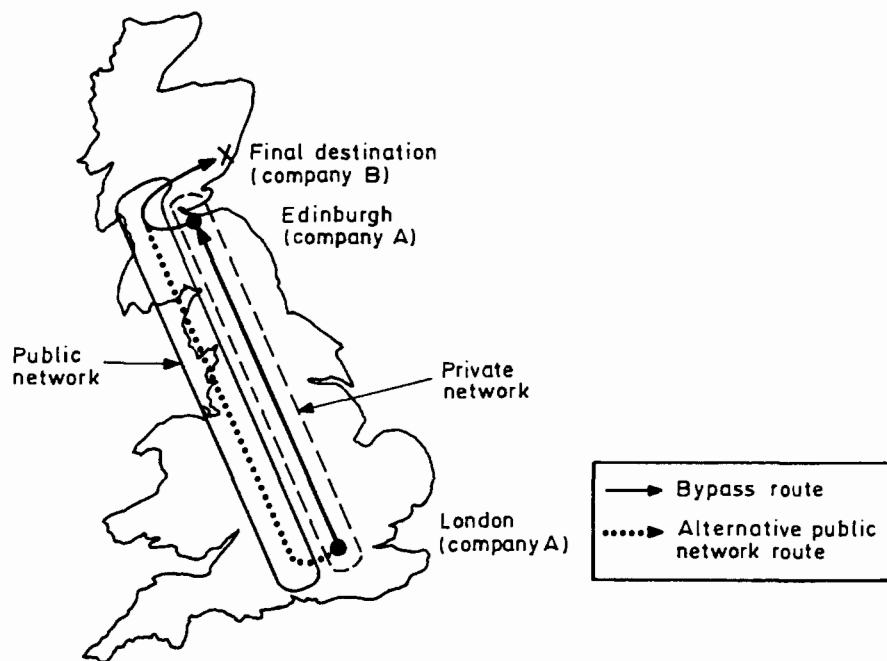


Figure 45.6 Public network bypass

of bypass. In the example shown, a company *private network* is established in the United Kingdom, linking *private exchanges* in London and Edinburgh. A caller connected to the private exchange in London wishes to make a call to an office in Edinburgh. Unfortunately the desired destination is not connected directly to the desired network. However, by routing the call over the private network from London to Edinburgh, and then into the public network for the final short hop, the public network (and its distance dependent tariff) has been bypassed for the greater part of the connection. Only the tariff of a local call is payable, rather than the much higher tariff that would have been due if the call had been routed over the public network as a *trunk call* for the entire distance between London and Edinburgh. Whether or not the overall economics favour such a bypass may change from day-to-day, depending on public network tariffs and on the existing topology of the private network. You should not assume that the example used is economic in all cases; specific evaluation is necessary on each occasion. Furthermore, the transmission quality of the end-to-end connection may preclude certain configurations. Particularly in the data networking case, the speed of data propagation is often critical.

Another means of achieving public network *bypass* is by the use of *out-of-area exchange lines*. Thus our Edinburgh office could have purchased London out-of-area exchange lines from the PTO. In effect, this makes the office appear to the public telephone user as if it were in London. It will have a London telephone number and call charges will assume that London is the endpoint. Thus calls from the Edinburgh to London office, or vice versa, count as 'local calls'. There is a rental charge, however, for the out of area line, similar to a leased circuit charge for a London/Edinburgh circuit.

## 45.7 A WORD OF WARNING

Having described a number of clever routing methodologies to reduce the costs of *private networks*, it is worth recalling, before we end the discussion, that the end-to-end connection must be properly serviceable. The transmission planning and routing 'rules', set out in Chapters 28 and 33, are applicable here. For example, it is no use creating a super-cheap end-to-end telephone connection which comprises so many component links that speech is unintelligible over it. However, the high performance of digital switching and transmission makes this much less likely.

If the *private network* is to be connected to the *public network* there may be legal restrictions on the type and technical characteristics of the signal that may be conveyed. In some countries telephone network bypass is illegal. In nearly all countries there will be signal power and technical line limitations.

## 45.8 PTO LEASED CIRCUIT OFFERINGS

Generally, PTOs offer a comprehensive range of different *leased (leaselines)* or *private circuit* types, each attuned to one or a number of particular types, such as data, telephone or packet switching. With care, the planning of private networks can be relatively straightforward.

The five most common types of leased circuit conform to ITU-T recommendations M.1020, or M.1025, M.1040, their national equivalents, or they are digital leased circuits.

Recommendation M.1020 defines the line conditioning (i.e. line-up limits, see Chapter 36) of a high grade four-wire analogue circuit, of 3.1 kHz bandwidth, suitable for high speed data or speech use. A circuit lined up according to M.1020 has a very flat frequency response (i.e. shows very little *attenuation distortion*), very good *group delay* response, and good circuit *stability*. Such a circuit is suitable for most of the analogue data modems described in Chapter 9 and can be used for point-to-point or networked (e.g. packet network) applications.

Recommendation M.1040 defines the line conditioning of a four-wire leased circuit, intended primarily for voice conversation. Its performance is not as good as that of an M.1020 circuit, but because less time and equipment is needed for line conditioning, the cost to the customer is lower. M.1025 provides a quality midway between M.1040 and M.1020.

Digital *leased circuits* are usually defined in terms of their bit rate (e.g. 64 kbit/s, 2 Mbit/s (E1), 1.5 Mbit/s (T1), 34 Mbit/s (E3), 45 Mbit/s (T3), 52 Mbit/s (STS-1) or 155 Mbit/s (STM-1)), their *bit error ratio (BER)* (typically  $10^{-5}$  or  $10^{-9}$ ) and their overall availability (typically around 99.5%). They are usually presented with one of the following interfaces: V.24(RS232), X.21 V.35, V.36(RS449), G.703 or an ISDN-like interface. Provided that these parameters suit the given application, the circuit may be used for any type of service (e.g. telephone, data).

Other types of leased circuits that are sometimes available are *wideband* analogue (e.g. FDM *group, supergroup*) or high speed digital circuits. In addition some PTOs are willing to provide just the physical medium, and allow the user to configure the bandwidth in the manner desired. Examples include the *wholesale* of optical fibre

capacity (so-called *dark fibre*) and the leasing and sale of satellite dishes for direct customer access to very high bandwidth satellite circuits (e.g. *International Business Service (IBS)*, which is a satellite service of INTELSAT allowing companies direct access for rates up to 2 Mbit/s). Finally, other specialized leased circuit services may be available from particular PTOs.

## 45.9 MAKING USE OF MOBILE RADIO TECHNOLOGY

The companies that propose to use mobile communications must give careful attention to the coordination and management of the system. Proper control must be established over

- equipment purchasing
- air-time subscription
- cost management
- user-justification rules

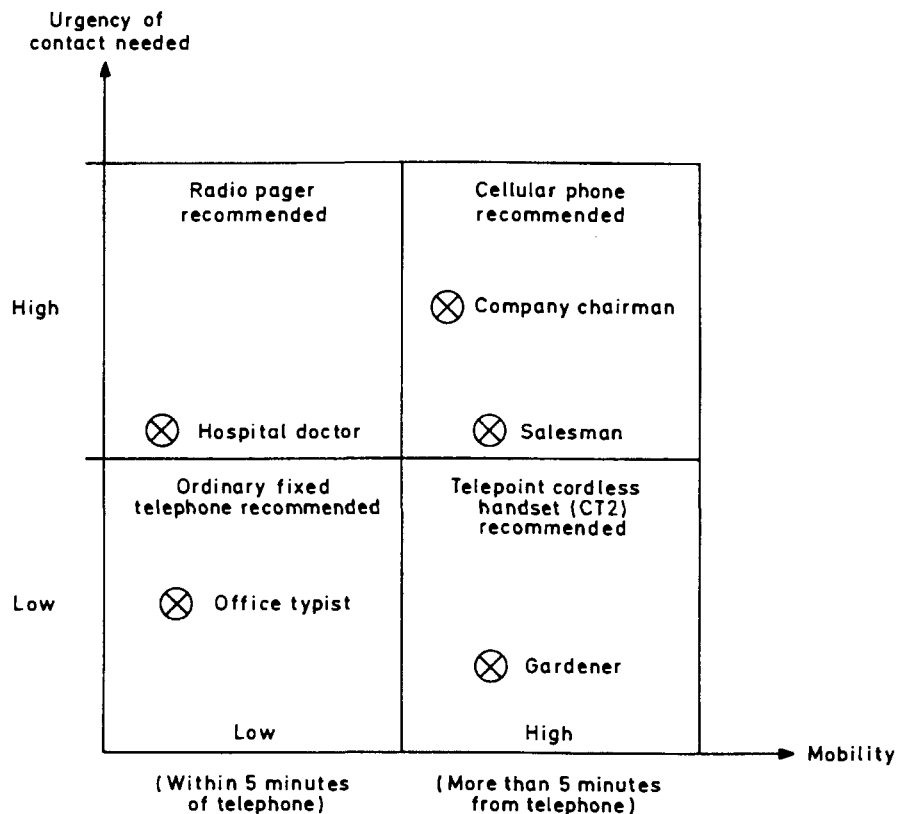


Figure 45.7 'Quantification' chart for mobile communication users

By dealing with a single supplier for carphone and hand-held units, and also for 'air-time' supply, the administration and control of costs is made considerably easier, but the 'buying-power' leverage over the supplier is also increased. The decision to purchase cellular or other mobile telephones for use by individual departments needs to be tempered by a central coordinating unit. One of the problems is that the falling cost of handsets is bringing many of them within the financial authorization of the lowest levels of management.

To get around the problem of integrating new and old generation equipment in this fast-changing environment one tactic could be to rent the handsets. However, because much of the cost involved is made up of usage charges (subscription and call charges) users should not be tempted to attach too much importance to what the secondary issue of handset cost. Usage costs of cellular radio technology are typically around £25 per month per handset, plus a per minute rate for calls of up to three times the ordinary telephony charges. Cordless technology (e.g. DECT) is slightly cheaper but nonetheless can represent a substantial extra cost over normal telephones or *radiopagers*. For this reason, companies are recommended to employ some means of assessing their users' need for service. A simple technique is shown diagrammatically in Figure 45.7.

Figure 45.7 classifies users into four broad categories according to their 'mobility' (how long they are likely to be away from the telephone) and their 'need for urgent communication' (i.e. whether they are likely to need to contact or be contacted by others urgently). Examples of typical users in each of the four categories of mobile communications are shown. A hospital doctor with low mobility but high urgency of contact is recommended to have a radiopager, whereas for a company chairman or salesman with high mobility and high contact urgency, a cellular telephone is recommended. However, if this service is heavily congested a public cordless network (e.g. DECT if available) might provide some degree of alternative. For an office typist we recommend no mobile telephone at all, and our recommendation for the DECT handset is the humble gardener, assuming of course he does not intend to 'lie low'.

---

## *Public Networks and Telecommunications Service Providers*

---

Historically, telecommunications networks have been run by state-owned organizations, who have operated as legally protected monopolies. These organizations tended to run telecommunications in a manner similar to other public utility services (electricity, gas, water, etc.), providing only a few very basic services, but in a reliable and well-engineered fashion, and available at almost any location to anyone who cared to apply. Recently, the trend has been to ‘privatize’ the established ‘public telecommunications operator’ (PTO) and allow new operators to set up networks in competition. The objectives in so doing are to improve the quality and diversity of available services, while simultaneously reducing unit costs and attracting private capital investment; in other words there is an attempt to make the PTOs more commercially responsive. Where competition is now a reality, the public telecommunications operators (PTOs) are forced to minimize running costs to work at maximum efficiency and so keep their prices competitive. In addition, PTOs are forced to adopt a much more commercial attitude towards marketing and business strategy. This chapter reviews briefly how some of the basic principles of marketing may be applied to a telecommunications marketplace and in addition outlines some of the PTOs’ options for business development.

### **46.1 COMPANY ‘MISSION’**

Like any good business, a PTO needs to define its business objectives clearly, including its overall mission and its strategy for achieving it.

A typical *mission* of a PTO might be:

- to provide a sufficiently diverse range of telecommunications networks and services to meet business (and/or residential) customers needs

- to provide services of high quality at a fair price and to operate in an efficient manner, developing the business to sustain growth in company earnings
- (possibly) to accept an obligation to serve the community at large, providing telecommunications for remote areas, and other specialized services for minority groups, including the handicapped
- to establish and maintain a certain market share

The *mission* needs to be sufficiently clear to motivate company staff appropriately, but sufficiently flexible to allow for change in business circumstances.

Realizing the company mission and determining day-to-day strategy is the task of company management. Each of the various functions and organizational departments need to be kept working in concert – serving their customers. It is no good having a brilliantly well-designed product or service, if the customers for whom it is intended are so unaware of it that they do not buy it. Similarly if it is far too expensive, not well enough made, simply too difficult to purchase, or if the customer service is poor, then customers will not want to buy it.

## 46.2 IDENTIFYING AND ADDRESSING THE PTO'S MARKET

The first part of a PTO's marketing programme must be to identify a customer requirement. A good deal of insight, imagination and research is required to analyse potential ways in which a customer's business can be improved by a product or service, in return for payment. The product or service may make the customer's activity easier, more effective, cheaper, more enjoyable, or may be attractive for some other reason.

A large PTO will be wise to maintain a balanced portfolio of products and services. Dependence on too few products leaves a company prone to rapid loss of business whenever there is a sudden and significant change in market demand, or when a rival company rapidly introduces a much better product. Nonetheless, each product must be clearly defined in terms of the customers need that it serves. In classical marketing terms this means getting the four P's right: product, price, place and promotion. In PTO service terms we interpret these as follows:

### 46.2.1 The Product

Telecommunications products (or services) should not be specified only in terms of their technical characteristics; the quality and customer service aspects of the product should be defined as well. The 'quality' is not only the service or manufacturing quality, but also the target in-service reliability and back-up measures (if any). The customer service arrangements should include any necessary enquiry service, and the standard and expected speed of repair should be defined.

### 46.2.2 The Price

Price is an important consideration in the customer's final decision whether or not to purchase. The customer may be unable to afford too high a price, no matter how 'fair' the price may be for a very high quality product. Paradoxically, however, too low a price may produce customer perceptions of a low quality product.

As a first estimate of a suitable price, it is valuable to undertake a 'cost-plus' analysis. Figure 46.1 illustrates the costs and revenue associated with a given product. Costs include *fixed costs* and *variable costs*. *Fixed costs* are those costs that are not related to sales volume. For the telephone service, the *fixed costs* are the costs of the telephones and the exchanges. These items are needed by all customers, irrespective of how many calls they make.

The *variable costs* include those associated with the lines interconnecting different exchanges. The number of circuits provided (and so the cost of them) can be adjusted to

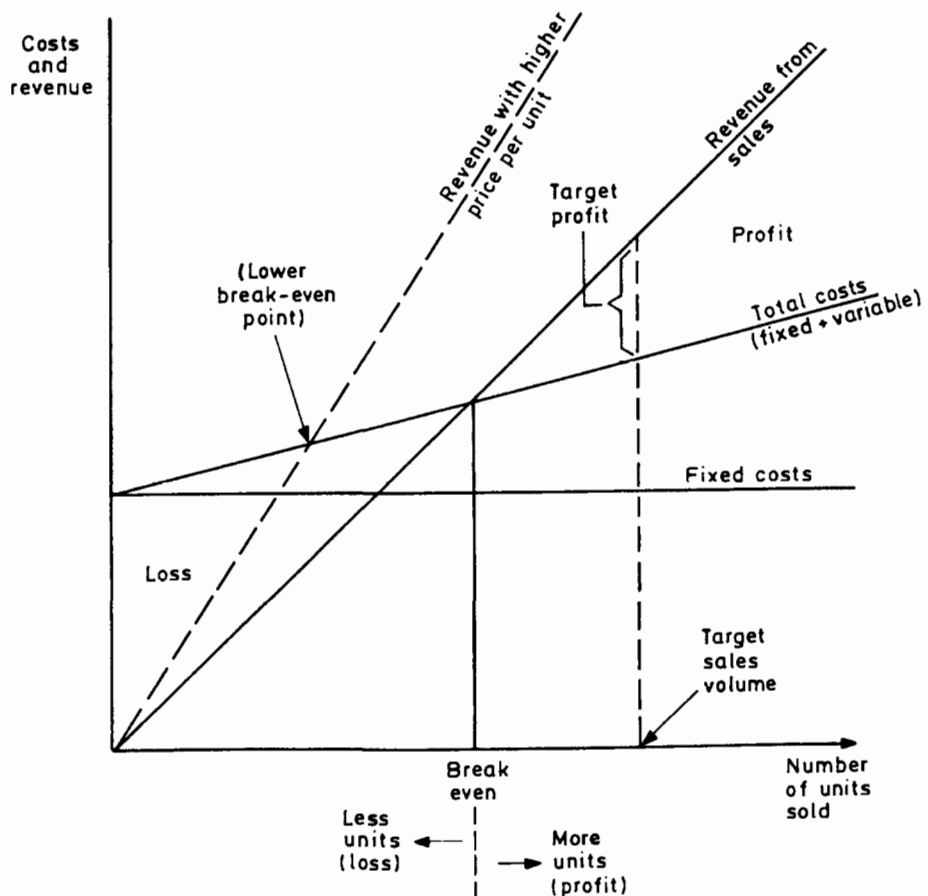


Figure 46.1 Determining product price



match the actual inter-exchange call demand. Other variable costs include those associated with staff costs (service level could be adjusted), administration, etc.

The revenue from the telephone service is the total money collected, and it is equal to the total number of call minutes times the price charged per minute. The relationship of costs and revenues is as shown in Figure 46.1.

The *break even point* shown on Figure 46.1, is the point at which the revenue just equals the costs. At this 'point', sufficient products (i.e. calls or whatever) have been made for the company to *break even*. Selling fewer units than this will result in a loss. Selling more will generate a profit.

The number of units that have to be sold to *break even* can be adjusted, by adjusting the price charged for each unit. The higher the price charged to customers for each 'unit', the smaller the number of units that need to be sold in order to break even. This is also illustrated in Figure 46.1.

The degree to which the number of sales is sensitive to price is known as the *price elasticity*. Normally, one expects a lower price to attract more sales and a higher price to deter sales; but different products have different *price elasticities*, so that with some products a higher price has little effect on the number sold, whereas for other *commodity products* the only important factor in the buyer's decision may be which competitor's product is the cheapest.

### 46.2.3 The Place

The *place* at which a product or service is available is also important. For some products, customers' loyalty will make them travel long distances to purchase a given 'brand' of product. For other products (for example, foodstuffs) the customer is likely to choose from the brands available on the nearest supplier's shelf (e.g. supermarket).

The place element of a telecommunications marketing plan needs to lay out the geographic availability of the service (for example 'lines are available nationwide' or 'lines are available in major cities'). The plan also needs to lay out the location of sales and customer support staff, and the method by which customers may 'buy' the service (e.g. by going to the local telephone office, or by calling a freephone number, etc.).

### 46.2.4 Promotion

Promotion is the means by which the intended customers of a product or service get to hear about it. Including 'advertising' and 'publicity', the reader will be familiar with many techniques used in this field.

## 46.3 PTO PRODUCT DEVELOPMENT

During product development, it is helpful to develop a set of documentation and procedures detailing the product, its launch and subsequent support. One possible means of doing this is by means of a 'product manual'. Suggested topics to be covered in such a manual are given below.

### 46.3.1 Product Manual

- A full description of the product (or service), describing its customer applications and its technical attributes, including precisely how customers' equipment is connected to the network.
- The competing products (both internal products and those of competing companies); the expected product lifetime, and the changeover strategy for products which it is planned to replace.
- The geographic availability of the product and any relevant service availability or 'upgrade' dates.
- The expected reliability of the product, and the methodology for pre-service and in-service *quality assurance* monitoring checks.
- The price of the product and any discounts; the warranty, service, maintenance and regulatory conditions.
- Safety and Technical Standards may need to be declared. *Customer premises equipment (CPE)* may have to be type-approved for connection to the network. Data protection regulations may apply to the computer storage of personal data (e.g. addresses) and further regulations are likely to govern communications secrecy, i.e. maintaining the confidentiality of information or other messages carried on behalf of customers.
- Is the product to be sold by *direct sales* or will an *agency* outlet be used? What advertising and promotional methods will be used?
- Is the product to be sold and transported to the customer's premises (e.g. an item of CPE)? Or does it need specialized installation, as in the case of a network access wire to the customer's premises (for a Network service)? Products and back-up spares should either be freely available to meet the demand of customers or be deliverable within the target lead-time.
- Ordering lead times and installation duration times need to be determined. Resources need to be allocated, and training given to company staff. An ordering procedure should to be established.
- Sales targets and incentives need to be set. An exhibition suite is needed to demonstrate complex services, and shops are needed for retail sales. A full price list must be available.
- Records of customer names, and of the applications to which products will be put, can be invaluable. For example, the knowledge that telephone line is used predominantly in conjunction with a facsimile machine may help maintenance and network design staff.
- A telephone number is needed for customer enquiries and fault reporting.
- Adequate training needs to be given to all staff, and reference information made available.

- Means and procedures are needed to measure the ongoing performance of the product, not just in technical, quality and reliability terms, but also from the marketing and sales perspective, and in financial viability terms.
- A proper procedure should be established for updating the product manual.

#### 46.4 PTO BUSINESS DEVELOPMENT

Finally, we consider the directions in which the world's major public telecommunications companies are likely to evolve and develop their business over the next few years. Historically, most PTOs were restricted to providing telecommunications network services within a single country, with international extension of network services made available as a result of *bilateral agreement* between two or more PTOs.

Figure 46.2 illustrates a typical spread of a PTOs current activities, with the predomination of business in the home country. Most business is in the 'basic interconnection market', though there is growing business resulting from international traffic interconnection, and there may be a considerable source of income from CPE sales and value-added services.

From this position, each PTO must develop its business strategy, taking into consideration a number of factors, including size, legal constraints, the degree of deregulation in its home market, the sophistication of its customers, the degree of its network development and its capital resources.

One PTO may decide that its appropriate short term goal is simply to make its established business grow in financial terms, without a significant change of emphasis in

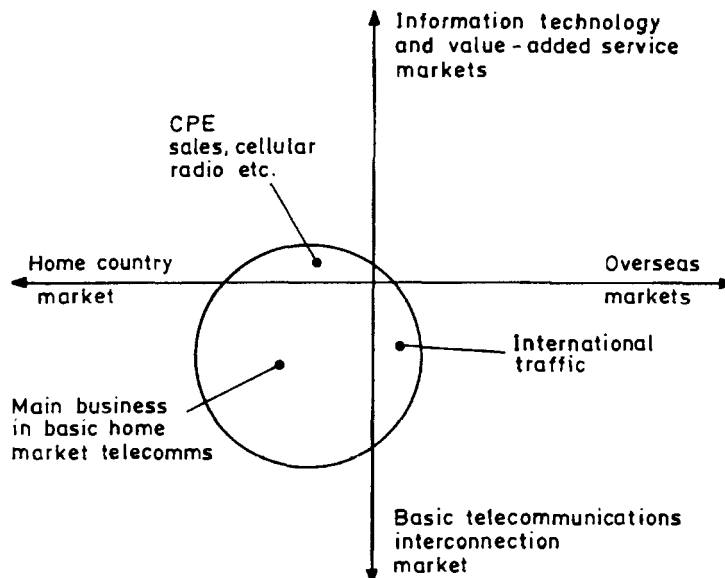


Figure 46.2 Typical PTO business (late 1980s)

its sphere of activity. In this case, the activity and resources of the PTO can be concentrated on developing the sophistication of its established network, on increasing its coverage area (perhaps by increasing the proportion of households with a telephone), and maybe by diversifying the number of basic services offered (for example, by introducing a packet switched data network). In support of this strategy, finance is needed for network developments, modernization, research and development (R&D) and product innovation in the home market, and revenue return accrues from the expanded home customer base, until the point of home market saturation is reached.

In contrast, some of the world's large PTOs, already operating in deregulated and well developed home markets, appear to wish to diversify their business into other spheres of activity. Opportunities for diversification include the specialist international service markets, and the whole field of value-added and information services. It seems possible that by the early 21st century a small number of major PTOs will have grown into large international conglomerates with spheres of business activity much wider than their previous operations. Notably British Telecom (BT), AT&T, France Telecom, Deutsche Telekom, MCI, Sprint, Unisource (a joint venture of the PTTs of the Netherlands, Sweden and Switzerland) and Cable & Wireless appear to have this goal. Figure 46.3, illustrates an example in which a PTO's business is diversified over a large range of markets throughout the world, with steady and balanced growth in each of the markets.

How will these companies achieve this level of expansion? Well, their first step will be the generation of spare cash from their established business. Then, diversification into the new markets will come about by financial venture in those countries where new deregulation permits it, by joint venture or by simple company acquisition. We can

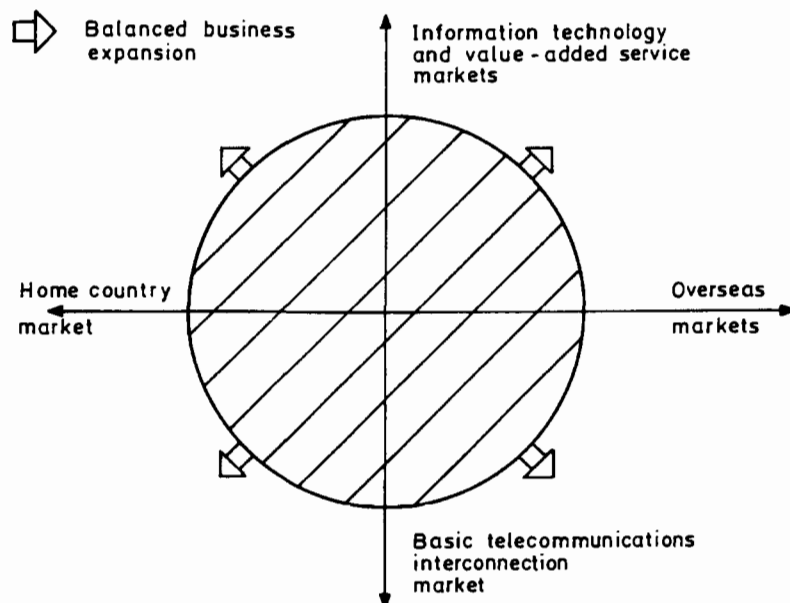


Figure 46.3 Business sphere of major international PTOs (mid to late 1990s)

therefore expect the establishment or acquisition of a number of PTO subsidiary companies. Largely autonomous at first, these companies may be kept small and dynamic to cope with the initial market entry, but will be able to call on the massive resources of the parent. Later on, fuller integration of the subsidiary organization into the main company structure may help further development of the company's 'core' business. Each PTO's prime business and its prime competitors will constantly be changing. We may see major alliances of PTOs with telecommunications manufacturers, computer hardware and software suppliers, and may even see enormous PTO conglomerates, operating in a number of regions of the world. Aspirations to become a 'global' PTO have already been demonstrated by a number of PTOs, which have sought variously to infiltrate overseas CPE markets or to run foreign networks on a franchise basis.

Some PTOs are already in close joint cooperation, specifically aiming to meet the telecommunications and value-added service needs of the world's multi-national corporations. This is a crucial market, since the top 60 corporations have annual turnover between £10 billion and £100 billion, and valuable telecommunications business. PTOs will ignore this business at their peril.

In Europe we can expect a fight for the domination of the *single European market (SEM)* in the *European Union* and a 'goldrush'-like scrambling between western PTOs as they attempt to capture market share in the under-developed parts of eastern Europe. The fight for the North American domestic market is already raging and the gloves are off for the world market battle! A major business conflict is in prospect, and we can expect some of the world's major PTOs, as well as some computer and telecommunications equipment manufacturers, to get bruised on the way! Let the fittest company survive!

---

# *Bibliography*

---

## **HISTORY**

*From Semaphore to Satellite – The Centenary of the International Telecommunication Union.* ITU Geneva, 1965.

*Beginnings of Telephony.* Fredrick Leland Rhodes. Harper & Brothers Publishers, 1929.

*Telecommunications – A technology for change.* Eryl Davies. London, Her Majesty's Stationery Office (HMSO), London Science Museum, 1973.

*A History of Engineering (1925–1975).* A. E. Joel Jr., Bell Telephone Laboratories, Inc, 1982.

*The International Telex Service.* R. N. Renton. Pitman, 1974.

*One Hundred Years of Telephone Switching (1878–1978), Part 1: Manual and Electromechanical Switching (1878–1960's) Volume 1.* Ropert Chapuis, North Holland Studies in Telecommunication, 1982.

## **GENERAL REFERENCES AND HANDBOOKS**

*Dictionary of Telecommunications.* S. J. Aries, Butterworths, 1981.

*Computer and Telecommunications Handbook.* Jeff Maynard, Granada, 1984.

*Electronic Communications Handbook.* Andrew F. Inglis, McGraw-Hill, 1988.

*NCC Handbook of Data Communications.* NCC Publications, 1982.

*Handbook of Data Communications.* John D. Lent, Prentice-Hall Inc., 1984.

*Telecommunications Transmission Handbook.* Roger L. Freeman. John Wiley & Sons, 2nd edition, 1981.

*Reference Manual for Telecommunications Engineering.* Roger L. Freeman, John Wiley & Sons, 2nd edition, 1994.

*Reference Data for Engineers: Radio, Electronics, Computer and Communications.* Edward C. Jordan, Harold W. Sams & Co Inc., 7th edition, 1985.

*Telecommunication System Engineering: Analogue and Digital Network Design.* Roger L. Freeman, John Wiley & Sons, 3rd edition, 1996.

*Telecommunication Networks.* J. E. Flood, Peter Peregrinus Ltd (for IEE), 1975.

*Telecommunications Engineer's Reference Book.* Edited by Fraidon Mazda, Butterworth-Heinemann Ltd, Oxford, 1993.

*Transmission Systems for Communications,* Bell Laboratories, 5th edition, 1982

## BASICS OF TELECOMMUNICATIONS

- Understanding Telecommunications.* Michael Overman, Lutterworth Press, 1974.  
*Telecommunications Networks: A technical introduction.* R. J. Murphy, Harold W. Sams & Co., 1987.  
*Telecommunications Systems* (TEC Level 1). P. H. Smale, Pitman. 2nd Edition, 1981.  
*Telecommunications Systems.* Pierre-Girard Fonto Uiet, Artech House Inc., 1986.  
*Telecommunications Engineering.* J. Dunlop & D. G. Smith. Van Nostrand Reinhold (UK), 1984.  
*Modern Electronic Communication.* Gary M. Miller, Prentice-Hall, 5th edition, 1996.

## TRANSMISSION LINES

- Fundamental of Transmission Lines and Electromagnetic Fields.* S. R. Seshadri, Addison-Wesley, 1971.  
*The Electrical Principles of Telecommunications.* R. Lowe and D. Nave, Macmillan, 1973.  
*Transmission Systems for Communications.* Bell Telephone Laboratories, 4th edition, 1970.  
*Digital, Analog and Data Communication.* William Sinnema, Reston Publishing Company, 1982.  
*Communication Systems.* Simon Haykin, John Wiley & Sons, 2nd edition, 1983.  
*Transmission Systems.* M. T. Hill and B. G. Evans, George Allan and Unwin Ltd., 1973.  
*Telecommunications Systems.* Piere-Girard Fonto Uiet, Artech House Inc., 1986.  
*Digital Transmission Systems.* David R. Smith, Van Nostrand Reinhold/Lifetime Learning Publications, 1985.  
*An Introduction to Digital Integrated Communications Systems.* H. Inose, Peter Peregrinus Ltd. (for IEE), 1979.  
*Digital Techniques and Systems.* D. C. Green, Pitman. 2nd edition, 1980.  
*Signal Coding and Processing: an Introduction Based on Video Systems.* J. G. Wade, Ellis Horwood Ltd., 1987.  
*ITU-T G.101: General characteristics for international telephone connections and international telephone circuits: The Transmission Plan.*  
 ITU-T G.781, G.782, G.783 recommendations defining SDH multiplexor.  
 ITU-T G.784 defines SDH management.

## SWITCHING AND SIGNALLING

- Telephony and Telegraphy: An Introduction to Instruments and Switching Systems* (Electro-mechanical and Reed-Electronic). Sydney F. Smith, Oxford University Press, 3rd edition, 1978.  
*Introduction to Telephones and Telephone Systems.* A. Michael Noll, Artech House, 1986.  
*Introduction to Telephone Switching.* Bruce E. Brieley, Addison-Wesley, 1973.  
*Telecommunications Switching.* J. Gordon Pearce, Plenum Press, 1981.  
*Basic Electronic Switching for Telephone Systems.* David Talley. Hayden Book Company Inc., 1982.  
*Electronic Switching.* Grinsec, North Holland Studies in Telecommunications, 1983.  
*Electronic Switching. Digital Central Office Systems of the World.* Amos E. Joel Jr., IEEE Press, 1982.  
*Digital Communications Switching.* John Ronayne, Pitman, 1986.  
*Signalling in Telecommunications Networks.* S. Welch, Peter Peregrinus Ltd (for IEE), 1979.

## RADIO AND SATELLITE SYSTEMS

- Radio and Line Transmission* (Vols 1 and 2). D. Roddy, Pergaman Press, 1967.
- HF Communications: A Systems Approach*. Nicholas Maslin, Pitman, 1987.
- Microwave Principles and Systems*. Nigel P. Cook, Prentice-Hall, 1986.
- Microwave Transmission-line Impedance Data*. M. A. R. Gunston, Van Nostrand Reinhold MARCONI Series, 1972.
- Microwave and Fibre Optic Communications*. Arthur F. Wickersham, Prentice-Hall, 1988.
- Communications Satellite Handbook*. W. L. Morgan and G. D. Gordon, 1989.
- Radio Receivers*. William Gosling, Peter Peregrinus Ltd. (for IEE), 1986.
- Radio Wave Propagation*. Arnel Picquenard, Macmillan, 1974.
- Radio System Design for Telecommunications* (1–100 GHz). R. L. Freeman, Wiley, 1987.
- Satellite Communications*. Stan Prentiss, Tab Books Inc., 2nd edition, 1987.
- Radio Communication Handbook*. Radio Society of Great Britain (RSGB), 5th edition, 1987.
- Manual of Satellite Communications*. Emanuel Fthenakis, McGraw-Hill, 1984.
- IFRB handbook on radio regulatory procedures*. ITU-R.
- ITU-R handbook of antenna diagrams*.
- ITU-R handbook of curves for radiowave propagation over the surface of the earth*.
- ITU-R recommendations: P1 series (1994). Propagation in ionized media*.
- ITU-R recommendations: PN series (1994). Propagation in non-ionized media*.

## MOBILE NETWORKS

- Cellular Mobile Radio Telephones*. Stephen W. Gibson, Prentice-Hall Inc., 1987.
- Digital Cellular Radio*. George Calhoun, Artech House, 1988.
- Mobile Cellular Telecommunications Systems*. William C. Y. Lee, McGraw-Hill, 1989.
- Pan-European Mobile Communications*. IBC Technical Services Ltd., Summer, 1989.
- Shipboard Antennas*. Preston E. Law Jr., Artech House, 2nd edition, 1986.
- Radiodetermination Satellite Services and Standards*. Martin A. Rothblatt, Artech House, 1988.
- Mobilfunknetze und ihre Protokolle* (Mobile networks and protocols) Prof. Bernhard Walke, Aachen University, 1995 (to be republished by Wiley in 1997).
- ETSI GSM recommendations 04.04, 1991, Layer 1, general requirements*.
- ETSI GSM recommendations 04.05, 1993, Data link layer, general aspects*.
- ETSI GSM recommendations 04.06, 1993, MS-BSS interface, Data link layer specification*.
- ETSI GSM recommendations 04.07, 1993, mobile radio interface, Layer 3, general aspects*.
- ETSI GSM recommendations 04.08, 1993, mobile radio interface, Layer 3 specification*.
- ETSI GSM recommendations 05.02, 1993, multiplexing and multiple access on the radio path*.
- ETSI GSM recommendations 05.08, 1993, radio sub-system link control*.
- ETSI digital European cordless telephone (DE/RES 3001)*.

## DATA COMMUNICATIONS

- Fundamentals of Data Communications*. Jerry Fitzgerald, Tom S. Eason, John Wiley & Sons, 1978.
- Data Communications and Teleprocessing Systems*. Trevor Housley, Prentice-Hall, 1979.
- Principles of Digital Data Transmission*. A. P. Clark, Pentech Press, 2nd edition, 1983.
- Basic Guide to Data Communications*. Edited by R. Sarch, McGraw-Hill, 1985.
- Data Transmission*. M. D. Bacon and G. M. Bull, MacDonald/London, and American Elsevier Inc., 1973.



- Practical Data Communications: Modems, Network and Protocols.* F. Jennings, Blackwell Scientific Publications, 1986.
- Data Communications Networking Devices: Characteristics Operation, Applications.* Gilbert Held, John Wiley & Sons Ltd., 2nd edition, 1989.
- The McGraw-Hill Internetworking Handbook.* D. Edgar Taylor, McGraw-Hill, 1995.
- Packet Switching and X.25 Networks.* S. Poulton, Pitman, 1989.
- Proprietary Network Architectures.* K. C. E Gee, NCC Publications, 1981.
- Systems Network Architecture: A Tutorial.* A. Meijer, Pitman, 1989.
- SNA Theory and Practice.* A. Guruga, Pergamon Infotech Limited, 1984.
- Systems Network Architecture: architecture reference*, version 2, SC30-3422-03, IBM Corporation, Research Triangle Park NC, USA, 1993.
- 3270 Information display system: introduction.* GA27-2739-22. IBM corporation, Research Triangle Park, NC, USA, 1988.
- APPN architecture and product implementations tutorial*, GG24-3816-00. IBM corporation, Raleigh, NC, USA, 1992.
- Technical Introduction to the Macintosh Family.* Addison Wesley/Apple Computer, 2nd edition, 1992.
- DECnet Digital Network Architecture (Phase V).* Digital Equipment Corporation, 1991.

## OPEN SYSTEMS

- What is OSI?* C. Pye, NCC Publications, 1988.
- Open Systems: The Basic Guide to OSI and its Implementation.* P. Judge, Computer Weekly/Systems International (Reed Business Publishing), 1988.
- OSI Explained: End to End Computer Communications Standards.* J. Henshall and S. Shaw, Ellis Horwood, 1988.
- The Open Systems Interconnection Model.* ISO 7498.
- Standards for Open Systems Interconnection.* T. Knowles, J. Larmouth and K. G. Knightson. BSP Professional Books, 1987.
- ITU-T X.200: Reference model of open systems interconnection.*
- MAP/TOP Networking.* V. C. Jones. McGraw-Hill Manufacturing and Systems Engineering Series, 1988.
- Government Open Systems Interconnection Profile (UK).* Central Computer and Telecommunications Agency, 1988.
- ITU-T X.208: Abstract syntax notation (ASN.1) old version.*
- ITU-T X.680: Abstract syntax notation (ASN.1) new version.*

## LOCAL AREA NETWORKS (LANS) AND METROPOLITAN AREA NETWORKS

- The LAN Jungle Book: A Survey of Local Area Networks.* W. S. Curne, Edinburgh University/Kinesis Computing, 1988.
- Local Area Networks (LANs) Standards IEEE 802 Series (ISO 8802 series).*
- IEEE 802.2: Logical link control (LLC).*
- IEEE 802.3: Ethernet LAN media access control (MAC).*
- IEEE 802.5: Token ring LAN media access control (MAC).*
- IEEE 802.6: Metropolitan Area Networks (MANs).*

*IEEE 802.7: Broadband LANs.*

*IEEE 802.8: FDDI/Fibre optic LANs.*

*High-speed Networking Technology, an Introductory Survey.* Dutton & Lenhard, IBM/Prentice Hall, 1992, 1995.

*Choosing a Local Area Network.* M. Gandy, NCC Publications, 1986.

*Local Area Networks: Selection Guidelines.* J. S. Fritz, C. F. Kaldenbach and L. M. Progar, Prentice-Hall, 1985.

*Broadband LAN Technology.* G. Y. Kim, Artech House, 1988.

*Broadband Data Communications and Local Area Networks.* R. G. Wilson and N. Squibb, Collins, 1986.

## ISDN

*Integrating Voice and Data.* Edited by R. Sarch. McGraw-Hill, 1987.

*Tutorial: Integrated Services Digital Networks (ISDN).* W. Stallings, IEEE/Computer Society Press, 1985.

*Integrated Services Digital Networks.* A. M. Rutkowski, Artech House, 1985.

*The Integrated Services Digital Networks: From Concept to Application.* J. Ronayne, Pitman, 1987.

*ISDN, The Integrated Services Digital Networks: Concepts, Methods, Systems.* P. Bocker, Springer-Verlag, 1988.

*ISDN Explained.* J. Griffiths, BT/John Wiley, 2nd edition, 1995.

*CCITT Basic Rate Interface (BRI).* ITU-T I.420.

*CCITT Primary Rate Interface (PRI).* ITU-T I.421.

*Network Interface Description for [USA] ISDN Customer Access [U reference point].* Bellcore TR-TSY-000776, Issue 1, Aug. 1989.

*Primary Rate Interface (PRI) Capabilities and Features.* AT&T TR41459, 1989.

*DASS2 (Digital Access Signalling System): British Telecom Network Requirement (BTNR) no. 191.*

*DPNSS (Digital Private Network Signalling System): British Telecom Network Requirement (BTNR) no. 190.*

*NET3: European Basic Rate Interface for ISDN (ETSI).*

*NET5: European Primary Rate Interface for ISDN (ETSI).*

## B-ISDN AND ATM

*ATM Networks and Principles.* Martin P. Clark, Wiley, 1996.

*ATM user-network interface (UNI) specification.* Prentice-Hall/ATM forum, version 3.1, 1995.

*ITU-T I.311: B-ISDN general network aspects,* 1993.

*ITU-T I.211: General service aspects of B-ISDN,* 1993.

*ITU-T I.321: B-ISDN protocol reference model and its application,* 1991.

*ITU-T I.327: B-ISDN functional architecture,* 1993.

*ITU-T I.361: B-ISDN ATM layer specification,* 1995.

*ITU-T I.362: B-ISDN ATM adaptation layer (AAL) functional description,* 1993.

*ITU-T I.363: B-ISDN ATM adaptation layer (AAL) specification,* 1993, addendum 1995.

*ITU-T Q.2010: B-ISDN overview – signalling capability set 1,* 1995.

*ITU-T Q.2931: Digital subscriber signalling 2 (DSS2) – UNI basic call control,* 1995.

## SIGNALLING SYSTEM NO. 7

- Message Transfer Part (MTP)*. ITU-T Q.701-709.  
*Telephony User Part (TUP)*. ITU-T Q.721-725.  
*Data User Part (DUP)*. ITU-T Q.741.  
*Integrated Services User Part (ISUP)*. ITU-T Q.760-9.  
*Signalling Connection and Control Part (SCCP)*. ITU-T Q.711-4.  
*Transaction Capabilities (TC)*. ITU-T Q.771-775  
*Operations and Maintenance Application Part (OMAP)*. ITU-T Q.795.  
*CEPT Telephony User Part (enhanced) (TUP+)*. CEPT Rec T/SPS 43-02.

## INTELLIGENT NETWORKS

- ETSI ETR 023: Intelligent networks framework*.  
*Physical plane for intelligent network capability set 1 (CS1)*. ETS 300 348 (ITU-T Q.1215), 1993.  
*Core intelligent network application protocol (INAP)*. ETS 300 374, parts 1-4.  
*Service Switching Points (SSPs) Generic Requirements*. Bellcore TR-TSY-000024, Oct. 1985.  
*Service Control Point (SCP) Node/Service Management System (SMS) Generic Interface Specification*. Bellcore TA-TSY-000365, Issue 2, July 1988.  
*Service Management Systems/Line Information Data Base (SMS/LIDB): Interface Specification*. Bellcore TA-TSY-000446, Issue 2, July 1988.  
*Service Switching Point 1 and Framework*. Bellcore TA-TSY-000921, Sept. 1988.  
*Intelligent Network Access Point (INAP) Framework*. Bellcore TA-TSY-000922, Sept. 1988.  
*Intelligent Peripheral (IP) Framework*. Bellcore TA-TSY-000923, Sept. 1988.  
*Service Logic Interpreter 1 and Framework*. Bellcore TA-TSY-000924, Sept. 1988.

## INTERNET, ONLINE SERVICES AND ELECTRONIC MAIL

- ITU-T X.400: Message handling systems: system model service elements*.  
*ITU-T X.500: OSI directory service*.  
*Electronic Message Systems and Services*. S. Roberts and T. Hay Comm. Ed. Books, 1989.  
*Internet protocol, RFC 791*, 1981.  
*Military standard internet protocol, MIL-STD-1777*, 1983.  
*Transmission control protocol (TCP), RFC 793*, 1981.  
*Military standard transmission control protocol, MTD-STD-1778*, 1983.  
*User datagram protocol (UDP), RFC 768*, 1980.  
*Internet control message protocol (ICMP), RFC 792*, 1981.  
*Ethernet address resolution protocol (ARP), RFC 826*, 1982.  
*Reverse address resolution protocol (RARP), RFC 903*, 1984.  
*Bootstrap protocol (BOOTP), RFC 951*, 1985.  
*Simple mail transfer protocol (SMTP), RFC 821*, 1982.  
*Telnet protocol specification, RFC 854*, 1983.  
*Trivial file transfer protocol, revision 2 (TFTP), RFC 783*, 1981.  
*File transfer protocol (FTP), RFC 959*, 1985.  
*Simple network management protocol (SNMP), RFC 1157*, 1990.  
*Border gateway protocol 3 (BGP), RFC 1267*, 1991.  
*Open shortest path first (OSPF), version 2, RFC 1247*, 1992.

*Point-to-point protocol for the transmission of multi-protocol datagrams over point-to-point links, RFC 1331*, 1992.

*Routing information protocol (RIP)*, RFC 1058, 1988.

*Simple internet protocol plus white paper, RFC 1710*, 1994.

## SPEECH AND DATA COMPRESSION

*Fundamentals of Speech Signal Processing*. Shyzo Saito and Kazuo Nakata, Academic Press, 1985.  
*Data Compression and Error Control Techniques with Applications*. V. Cappellini, Academic Press, 1985.

*Data Compression: Techniques and Applications, Hardware and Software Considerations*. G. Held, John Wiley & Sons, 2nd edition, 1987.

*Data Transmission: Analysis, Design, Applications*. D. Tugal and O. Tugal, McGraw-Hall, 1982.

*Digital Transmission Systems (Good review of ADPCM)*. D. R. Smith, Van Nostrand Reinhold/Lifetime Learning Publications, 1985.

*ITU-T recommendations G.721 and G.726: Adaptive Differential Pulse Code Modulation (ADPCM)*.

*ITU-T G.728: Low delay code excited linear prediction (LD-CELP)*.

*Signal Coding and Processing: An introduction based on video systems*. J. G. Wade, Ellis Horwood Ltd., 1987.

*A Method for the Construction of Minimum Redundancy Codes*. D. A. Huffman, *Proceedings IRE*, 40, 1098, 1952.

## TELETRAFFIC AND NETWORK TOPOLOGY ENGINEERING

*Principles of Telecommunication Traffic Engineering*. D. Bear, Peter Peregrinus Ltd., 3rd edition, 1988.

*Introduction to Queueing Theory*. R. B. Cooper. North Holland, 2nd edition, 1981.

*Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges*. A. K. Erlang, 1918.

*An outline of the trunking aspect of automatic telephony*. G. F. O'Dell. *Journal of the IEE*, 1927, 185–222.

*Teletraffic Engineering Manual*. Standard Elektrik Lorenz Ag., Stuttgart, 1966.

*Computer-Communication Network Design and Analysis*. M. Schwartz, Prentice-Hall Inc., 1977.

*Tariffs, Traffic and Performance: The Management of Cost Effective Telecommunications Services*. J. M. Hunter, N. Lawrie and M. Peterson, Comm. Ed. Publishing Ltd., 1988.

*Telecommunications Economics*. T. J. Morgan, Technicopy Ltd., 2nd edition, 1976.

*ITU-T E.401: Statistics for the international telephone service (number of circuits in operation and volume of traffic)*.

*ITU-T E.506: Forecasting international traffic*.

*ITU-T E.507: Models for forecasting international traffic*.

*ITU-T E.508: Forecasting new international services*.

*Theories for toll traffic engineering in the USA*. R. I. Wilkinson, *Bell System Technical Journal*, 1956, 35, 421–514.

*Planning of junction network in a multi-exchange area*. Y. Rapp, Ericsson Technics, No. 1, 1964, 77–129.

*An explicit formula for dimensioning links offered overflow traffic*. L. T. M. Berry, *Australia Telecommunications Research*, 1974, 8, 13–17.

*Economic Analysis of Telecommunications: Theory and Applications.* L. Courville, A. de Fontenay and R. Dobell, North Holland, 1983.

*Telecommunications Economics.* T. J. Morgan, Technicopy Ltd., 2nd edition, 1976.

## NETWORK MANAGEMENT AND TMN

*ITU-T recommendations, X.700 series, OSI systems management.*

*ITU-T X.720, OSI management information model.*

*ITU-T X.722, Guidelines for the definition of managed objects (GDMO).*

*ITU-T M.3000, Overview of TMN recommendations.*

*ITU-T M.3010, Principles for a TMN.*

*ITU-T M.3020, TMN interface specification methodology.*

*ITU-T M.3100, Generic network information model.*

*ITU-T M.3110, Generic network level information model.*

*ITU-T M.3180, Catalogue of TMN management information.*

*ITU-T M.3200, TMN management services overview.*

*ITU-T M.3400, TMN management functions.*

*ITU-T Q.811, 812 and 824 Q3 interface.*

*ITU-T X.710 CMIS (common management information service) definition.*

*ITU-T X.711 CMIP (common management information protocol) specification.*

*Simple network management protocol (SNMP), RFC 1157.*

## PROJECT MANAGEMENT AND TELECOMMUNICATIONS CONTRACTS

*Perspectives on Project Management.* R. N. G. Burbridge, Peter Peregrinus Ltd (for IEE), 1988.

*Critical Path Analysis and Other Project Network Techniques.* K. Lockyer, Pitman, 4th edition, 1984.

*Drafting Engineering Contracts.* H. Henkin, Elsevier Applied Science, 1988.

## TECHNICAL STANDARDS

*ISO Catalogue.*

*ITU Catalogue of Publications,* ITU, Geneva.

*Quick Reference to IEEE Standards.* IEEE, 1986.

*BSI Standards Catalogue.*

*OSI Standards and Acronyms.* A. V Stokes, Blenheim Online Publications, 2nd edition, 1988.

*Omnicom Index of Standards.* McGraw-Hill, 1988.

*McGraw-Hill's Compilation of Data Communications Standards.* Edition III, Edited by H. C. Folts, Vol. 1-3. McGraw-Hill, 1986.

## OTHER NOTABLE ITU-T RECOMMENDATIONS

*T.4: Standardisation of Group 3 facsimile apparatus for document transmission.*

*T.50: International reference alphabet, formerly International Alphabet number 5.*

- X.25: Interface between DTE and DCE for terminals operating in the packet mode and connected to public data networks by dedicated circuits.*
- X.21: Interface between data terminal equipment (DTE) and data circuit-terminating equipment (DCE) for synchronous operation on public data networks.*
- X.3: Packet assembly/disassembly facility (PAD) in a public data network.*
- X.28: DTE/DCE interface for a start-stop mode DTE accessing the PAD facility in a public data network.*
- X.75: Terminal and transit call control procedures and data transfer system on international circuits between packet-switched data networks.*
- E.170: International Routing Plan.*
- X.110: International routing principles and routing plan for public data networks.*
- E.164: Numbering plan for the ISDN era.*
- X.121: International numbering plan for public data networks.*
- F.69: Plan for telex destination codes.*
- Q.800-series recommendations on Quality of Service (QOS) and Network Performance (NP).*
- M.1020: Characteristics of special quality international leased circuits with special bandwidth conditioning.*
- M.1025: Characteristics of special quality international leased circuits with basic bandwidth conditioning.*
- M.1040: Characteristics of ordinary quality international leased circuits.*
- G.703: Physical/electrical characteristics of hierarchical digital interfaces.*
- G.704: Synchronous frame structures used at primary and secondary hierarchical levels.*
- V.24: Data communication over the telephone network: list of definitions for interchange circuits between data terminal equipment (DTE) and data circuit-terminating equipment (DCE).*
- V.32: A family of 2-wire, duplex modems operating at data signalling rates of up to 9600 bit/s for use on the general switched telephone network and on leased telephone-type circuits.*
- V.10: Electrical characteristics for unbalanced double-current interchange circuits operating at data signalling rates nominally up to 100 kbit/s.*
- V.11: Electrical characteristics for balanced double-current interchange circuits operating at data signalling rates up to 10 Mbit/s.*

## TELECOMMUNICATIONS AS A BUSINESS TOOL

- Managing to Communicate.* Martin P. Clark, Wiley, 1994.
- Competing In Time: Using Telecommunications for Competitive Advantage.* P. G. W. Keen, Ballinger Publishing, 1986.
- Telecomms for Business: A Manager's Guide.* B. Miller, Comm. Ed. Books, 1986.
- Manager's Guide to Telecommunications.* M. Gandoff, Heinemann, New Tech, 1987.
- The Computer User's Year Book 1989.* ISSN 0268-6821, VNU Business Publications, London.
- The Software User's Year Book 1989.* VNU Business Publications, London.
- The EDI Handbook: Trading in the 1990s.* Gifkins and Hitchcock, Blenheim Online Publications, 1988.
- Corporate Communications Networks.* J. E. Lane, NCC Publications, 1984.
- Private Telecommunications Networks: Design and Implementation.* R. Bell, Comm. Ed. Books, 1988.
- Communication Networks: Private Networks Within the Public Domain.* J. E. Ettinger, Pergamon Infotech Ltd., 1985.
- Worldwide Telecommunications Guide for the Business Manager.* W. M. Vignault, Wiley, 1987.
- Eosys Cabling Guide for Building Professionals.* Eosys Ltd., Clove House, The Broadway, Farnham Common, Slough, 1986.

- Reviewing Your Data Transmission Network.* P. R. D. Scott, NCC Publications, 1983.  
*The LAN Jungle Book: A Survey of Local Area Networks.* W. Scott Currie, Edinburgh University/Kinesis Computing, 1988.

## TELECOMMUNICATONS REGULATION

- Telegraph Regulations/Telephone Regulations: Final Acts of the World Administrative Telegraph and Telephone Conference.* ITU Geneva, 1973, Final Protocol/Resolutions/Recommendations/Opinions.  
*From Telecommunications to Electronic Services: A Global Spectrum of Definitions, Boundary Lines and Structures.* R. R. Bruce, J. P. Cunard and M. D. Director, Butterworths, 1986.  
*The Law and Regulation of International Space Communication.* R. L. White and H. M. White Jr., Artech House, 1988.  
*Telecommunications Policy.* Edited by C. Blackman, Butterworth Scientific, continuously updated.  
*The Future of Telecommunications.* M. Beesley and B. Laidlaw.  
*US Code of Federal Regulations Title 47.* Published annually.

## TELECOMMUNICATIONS SERVICE PROVIDERS AND STRATEGY

- Tariffs, Traffic and Performance: the management of cost effective telecommunications services.* J. M. Hunter, N. Lawrie and M. Peterson, Comm. Ed. Publishing Ltd., 1988.  
*Telecommunications Industry: Technical Change and International Competitiveness.* E. Sciberras and B. D. Payne, Longman, 1986.  
*From Telecommunications to Electronic Services – A Global Spectrum of Definitions, Boundary Lines and Structures.* R. R. Bruce, J. P. Cunard and M. D. Director, Butterworths, 1986.  
*International Communications: Blueprint for Policy.* K. W. Leeson, North Holland, 1984.  
*International Telecommunications: User Requirements and Supplier Strategies.* K. L. Lancaster, Lexington Books, 1982.

---

# *Glossary of Terms*

---

Two-wire communication	A communication path between end points using only two wires. Its difficulties and limitations arise from the shared use by transmit and receive signals of the same path.
2B + D interface	An alternative term used to describe the ISDN basic data interface.
Four-wire communication	A communication path between end points using four wires, two each for each direction of transmission, transmit and receive. Four wire communication paths have better isolation of transmit and receive signals than two-wire paths, and are thus easier to operate.
adaptive differential pulse code modulation (ADPCM)	A highly efficient, low bit speed method of digital encoding of voice signals.
advanced peer to peer networking (APPN)	A communications architecture developed by the IBM company. It is used between IBM AS400 computers and between workstations and servers.
amplification	The boosting of a faint signal, counteracting the effects of attenuation to ensure good comprehension by the receiver.
amplitude	The relative strength or volume of a signal.
amplitude modulation (AM)	A technique used to enable a low bit speed or comparatively low bandwidth information signal to be carried on a higher frequency carrier signal, by adjustment of the latter's amplitude to match the former's wave shape.
analogue transmission	A method of signal transmission in which the shape of the electrical current waveform on the telecommunications transmission line is analogous to the air pressure waves of the sound or music signal it represents.
application layer	The uppermost layer (layer 7) of the open systems interconnection (OSI) model. This layer is a software function providing a program in one computer with a communication service to a 'cooperating' computer.
ASCII	The American Standard Code for Information Interchange. A widely-used 7-bit binary code used to represent alphabetic and numeric characters in computer-understandable form.
asymmetric digital subscriber line (ADSL)	A transmission technique allowing high speed signals (e.g. 6–8 Mbit/s) to be transmitted to customer premises over existing copper lineplant. In the upstream direction typically a maximum of 384 kbit/s is available.
asynchronous data transfer	A method of data transfer in which each alphabetic or numeric character (represented by 7 or 8 bits) is preceded by start and stop bits to delineate the 7/8 bit pattern from the idle bit pattern which otherwise occupies the digital line.



asynchronous transfer mode (ATM)	A modern technology to integrate all types of telecommunications traffic (voice, data, video, multimedia) across a single network. ATM forms the technological basis of the broadband integrated services digital network.
attenuation	The effect of signal dwindling experienced with accumulating line length or distance of radio transmission.
attenuation distortion	The various frequency components of a complex sound or data signal are often attenuated by the line to different extents. The tonal balance of the original signal is thus distorted: attenuation distortion.
audio circuits	Copper or aluminium-wired circuits used to carry a single signal, typically a speech signal in the audio bandwidth.
balanced transmission	Telecommunications transmission across metallic conductors (i.e. wire pairs) in which both conductors play an equal role (e.g. twisted pair 120 Ohm cabling).
bandwidth	Measured in Hertz or cycles per second, the range of component frequencies making up a complex speech or data circuit. Thus speech comprises frequencies between 300 Hz and 3400 Hz, a total bandwidth of 3.1 kHz.
baseband	An analogue signal in its original form (prior to any form of multiplexing).
basic rate interface (BRI)	The simplest form of network access available on the ISDN (integrated services digital network). The BRI comprises 2B + D channels for carriage of signalling and user information.
bearer service	Relating in particular to ISDN, various bit rates or bandwidths (called bearer services) may be made available between the two ends of the connection. Examples include speech bearer service, 64 kbit/s data service.
binary code	A two-state (on/off) code easily interpreted by electronic circuits. 7 and 8 Binary digIT (bit) codes are usually used in computers to represent alphabetic and numeric characters.
bit error ratio (BER)	A measure of the quality of a digital transmission line, either quoted as a percentage, or more usually as a ratio, typically 1 error in 100 thousand or 1000 million bits carried. The lower the number of errors, the better quality the line.
bit rate or bit speed	The bandwidth of a digital line or signal governing the rate at which individual alphabetic or numeric characters may be carried by the line.
bridge	A device to interconnect LANs or LAN segments together.
broadband	A service or system supporting rates greater than 2 Mbit/s.
broadband integrated services digital network (B-ISDN)	A powerful network capable of carrying all types of narrow and broadband telecommunications service types (e.g. voice, data, Internet, video, multimedia, etc.)
broadband passive optical network (BPON)	An optical fibre transmission network used to connect customer premises to the public network for purpose of high bandwidth services (e.g. television distribution).
broadcast	A service providing unidirectional distribution to multiple receivers.
brouter	A device comprising both bridge and router functionality.
busy hour	The period of highest network usage.
busy hour traffic	A measure of the maximum user demand for network throughput.
byte	A binary digital signal value of 8 bits (decimal value 0–255).
carrier frequency	A high frequency signal used in FDM or radio transmission. The information is modulated onto the carrier prior to transmission.
carrier sense multiple access with collision detection (CSMA/CD)	The transmission technique employed in ethernet LANs. Data packets are sent onto a common bus when the bus is seen to be idle. Addresses in the packets allow the correct destination device to take them from the bus.
cell	A data block of fixed length (48 byte information field and 5 byte header). Cells are transmitted across a network using ATM.

cellular radio	A mobile call-in and call-out telephone service in which radio base stations are arranged to individual cellular coverage areas. Variants include TACS, AMPS, NMT, GSM, network-C, radio, Radiocom-2000.
CELP (code excited linear predication)	A fast means of low bitrate digital speech compression. Because the algorithm is linear the processing may be carried out more quickly, thus minimizing signal delay.
centrex	A telephone network service enabling PBX-type functions to be offered direct from the public exchange.
channel associated signalling (CAS)	A method of telephone signalling in which dialled digit and other call set-up and control information is carried by the associated circuit or via timeslot 16 of the associated 2 Mbit/s digital trunk circuit.
circuit-switching	A method of switching provides for a direct connection of fixed bandwidth or bitrate (e.g. 64 kbit/s) between originating and terminating points of a communication for the entire duration of a call.
circuit multiplication equipment (CME)	A device enabling line economy by processing signals (e.g. speech) in such a way that it appears that more than one circuit were available, hence circuit multiplying by statistical means.
circuit transfer mode	A telecommunication transfer technique based on circuit-switching and requiring permanent allocation of bandwidth.
coaxial cable	A cable used for high frequency and high bandwidth analogue and digital signal transmission. The cable comprises a single inner conductor plus a conducting outer sheath, in cross-section a dot in a circle.
code division multiple access (CDMA)	A 'spread spectrum' technique used in point-to-multipoint radio systems. CDMA systems spread a large number of signals across a common spectrum using a code. The main advantage of the technique over TDMA and FDMA is resistance to fading and interference.
codec (coder/decoder)	A device for interfacing a four-wire digital transmission line to a four-wire analogue. The coder codes the to-be-sent two-wire analogue signal in a form suitable for two-wire digital transmission. The decoder meanwhile decodes the received two-wire digital signal.
common channel signalling	A type of interexchange signalling in which the dialled digit, call set-up and other circuit control information for a large number of circuits (typically 200) is conveyed directly between the exchange processors on common channel signalling links.
common management information protocol (CMIP)	The standard data protocol defined by ISO and ITU-T as part of the telecommunications management network for requesting management status information from a network element or sending control commands to it. CMIP is used at the Q3-interface.
configuration management	One of the five main network management functions defined by the ISO management model. The others are accounting, security, performance and fault management.
connection-oriented	Connection-oriented communication across a telecommunications network is that in which a connection is established between source and destination before the communication contents may be transferred. This ensures that the receiver is 'ready' to receive.
connection admission control	The procedure by which a network decides whether to accept a new connection.
connectionless service	A service allowing for information transfer across a telecommunications network without prior establishment of an end-to-end connection. 'Post office'-like devices store the message prior to collection by the destination party.
constant bit rate	A service characterized by a service bit rate of constant value.

control plane	A conceptual communications connection between data switching devices for communicating control information (e.g. for establishing or clearing connections, etc.)
conversational service	An interactive service providing for bidirectional communication.
cordless telephone	A telephone handset connected to the ordinary telephone network by means of a short radio path (typically up to 500 metres).
crosstalk	The interference of telephone and other signals resulting from the induction of an unwanted signal from an adjacent telephone line (the crossed-wire).
customer premises equipment (CPE)	The telephone or data equipment on a customer's premises, that using the public telephone or data network for connection to a remote location.
cyclic redundancy check (CRC)	A binary check sum enabling the detection and (possibly also) correction of bits within a data frame which have become corrupted during passage over a long distance connection.
data	Data is the term used to describe alphabetic and numeric information stored in and processed by computers.
data circuit terminating equipment (DCE)	The equipment terminating and controlling the transmission line, and often marking the endpoint of the public data network. Data terminal equipments (DTEs) such as computers are connected directly to DCE.
data compression	Any of various techniques, of both reversible and irreversible types, designed to reduce the bit speed needed for a particular type of communication (e.g. facsimile).
data terminating equipment (DTE)	The jargon used to describe any type computer or other equipment, when connected to a data communications network.
datagram	A type of data packet-switching in which the various packets belonging to a given connection do not all traverse the same path through the network. This type of switching improves the network robustness and increases security against 'eavesdroppers'.
datalink layer	Layer 2 of the open systems interconnection (OSI) model. This software function operates on an individual link within the connection to ensure that individual bits are conveyed without error. The best known layer 2 protocol is HDLC.
decibel (dB)	A measure of signal strength, volume, amplification or signal loss.
digit analysis	The process of analysis of dialled digits carried out by an exchange to determine the destination of a call and therefore enable the selection of an appropriate outgoing route.
digital European cordless telephone (DECT)	The ETSI-standardized technology used in digital cordless telephones.
digital transmission	A modern method of signal transmission in which the speech or data information is represented as a string of binary digits.
distortion	The corruption of a signal brought about by electrical disturbance or non-uniform line properties, particularly causing different frequency components of a complex signal to be unequally affected.
distribution	A broadband service used to distribute video or audio signals to end users.
duplex	Transmission means allowing simultaneous two-way signal transmission, in both receive and transmit directions without interference of the signals.
E&M signalling	A means of telephone signalling in which the 'loop' and 'disconnect' pulses are carried on two extra separate leads corresponding to 'ear (E)' and 'mouth (M)'. This type of signalling is easily converted to tone signalling suitable for carriage via FDM.
EBCDIC	Extended binary coded decimal interchange code. An 8-bit computer code to represent alphabetic and numeric characters.

echo	The effect resulting from a delayed reflection of a signal. Echo in a telephone circuit manifests to the talker as a slightly delayed repeat of his own voice, returned from the distant end, just like a sound echo.
echo canceller	A device used to suppress telephone echo (return of the speaker's voice) which operates by simulating a negative version of the outgoing signal in the receive path, so that the returning echo is 'cancelled'.
electromagnetic compatibility (EMC)	EMC is the term applied to a series of specifications which define how telecommunications devices should be immune to electromagnetic disturbance.
electromagnetic interference (EMI)	Electromagnetic interference is the disturbance caused by a telecommunications device to other neighbouring devices.
electronic data interchange	Electronic data interchange is the transfer of computer-readable data (orders, payments, etc.) between companies trading with one another.
electronic mail	Electronic mail is the use of telecommunications networks and 'post offices' for the transfer of messages, letters and documents between humans. The best known type of e-mail is that supported by the Internet.
electronic ticketing	The process of recording call details at the time of call-set-up and conversation so that charges can be calculated and billed to the calling customer.
<i>en bloc</i> signalling	A method of interexchange signalling in which the entire dialled number is transferred from one exchange to the next as a continuous string of digits. Digits are not conveyed to any subsequent exchange until all digits have been received.
encryption	The coding ('scrambling') of data information prior to carriage across a telecommunications network, to reduce the security risk arising from potential interception of the information.
equal access	The interconnection means used between local exchange carriers and long distance telephone companies which is meant to ensure fair competition between the long distance carriers.
equalization	The process carried out on long lines to counteract the ill-effects of attenuation distortion, rebalancing the signal strengths of various component frequencies to be equal to that of the original signal.
Erlang	The measure of traffic intensity on circuit-switched networks. One Erlang of traffic represents an average demand of one call in progress continuously throughout the busy hour of a given route or exchange.
Erlang lost-call formula	Developed in 1917 by A. K. Erlang, this formula is used to calculate circuit requirements of a given interexchange route given the traffic demand in erlangs and the required grade of service.
Erlang waiting-call formula	A derivative of the lost-call formula, the Erlang waiting-call formula is used to determine circuit requirements, queue length requirements and the server numbers needed in a waiting system (e.g. a data network or an operator switchroom).
error-free seconds (EFS)	A performance measure used to specify the quality of data lines. Higher specification lines require a higher proportion of error free seconds.
ethernet LAN	The most popular type of physical infrastructure used for local area network interconnection of PCs and other computer devices within an office environment. Based on a bus structure and the use of the CSMA/CD protocols.
facsimile	A device capable of transmitting or receiving graphical (paper-based) documents or images via the telephone network.
fade margin	The reserve capacity ('gain') of a radio receiver which helps to guard against fading.
fading	The loss of radio signal caused by interference or weather attenuation effects.
fault management	One of the five main network management functions defined by the ISO management model. The others are accounting, security, performance and configuration management.

fibre distributed data interface (FDDI)	A type of metropolitan area network (MAN) capable of bitrates up to about 100 Mbit/s. FDDI is typically used to link LAN segments within a large office complex or on a campus site.
frame	A data block of variable length identified by a header label.
frame relay	A modern form of packet switching, capable of much higher bitrates than X.25-based packet-switched networks.
frame relay access device (FRAD)	A device which converts data from a data terminal equipment (DTE) for carriage by a frame relay network.
framing	The structuring framework required in digital line systems so that a number of different users or channels may share the line by time division means.
freephone	A telephone network service in which the call recipient pays call charges (e.g. 800-service).
frequency distortion	A signal impairment caused by transmission line phenomena (e.g. attenuation) which affect the various component frequencies of a complex signal unequally. The tonal balance of the original complex signal is lost, distorted.
frequency division multiple access (FDMA)	A technique used in point-to-multipoint radio systems. FDMA shares the available radio spectrum by allocating a small part of the available spectrum to each user. The main advantage of the technique over CDMA and TDMA is its simplicity.
frequency division multiplex (FDM)	A technique allowing a number of different audio signals to share the same line without impact upon one another. A high bandwidth transmission line is shared out by dividing into a number of smaller bandwidths, each used to carry a single user's signal.
frequency modulation (FM)	A technique used to enable a low bit speed or low bandwidth information signal to be carried on a much higher frequency carrier signal by adjustment of the latter's frequency by a small amount corresponding to the amplitude of the former.
frequency shift distortion	An impairment caused by slight, but different shifts in frequency of the component frequencies of a signal.
frequency shift keying (FSK)	When a carrier signal is frequency modulated with a binary data signal, the effect is merely to alternate the carrier signal between two or more fixed values. This is known as frequency shift keying.
full duplex	In contrast to half-duplex devices, full duplex ones allow permanent, simultaneous two-way transmission of information, without interaction or interference of receive and transmit signals.
gain hits	An impairment causing errors on data transmission lines. Gain hits are intermittent but only moderate and relatively short disturbances in the amplitude of a received signal.
gateway	A device used to interconnect networks of dissimilar technology, protocols or ownership.
geostationary satellite	A type of communications satellite in orbit above the earth's equator but moving at the same speed as the rotation of the earth. From the ground, the satellite therefore appears to be geographically stationary.
global system for mobilecommunication (GSM)	One of the most common types of cellular network technology, and one which allows full roaming of subscribers between networks around the world. Most common in Europe.
grade of service (GOS)	The proportion of calls in a circuit switched network which are lost due to network congestion. Because it is uneconomic to provide networks of enormous uncongested proportion, it is usual to design networks to a given grade of service, typically 1 in 100.
group delay	A signal impairment caused by different speeds of propagation of the component frequencies of a complex signal.

half duplex	A telecommunications device allowing two-way transmission of signals or other information, but only in one direction at a time. Thus a half duplex device cannot simultaneously transmit and receive, though interspersed bursts in each direction are possible.
Hamming code	An error correction/detection code used in datacommunications.
harmonic disturbance	Harmonic disturbance is an impairment resulting from the intermodulation of two different signal frequencies $F_1$ and $F_2$ when passed through nonlinear transmission devices. Stray signals, $F_1 + F_2$ , $F_1 - F_2$ , $F_1 + 2F_2$ , etc., are produced.
Hayes protocol	A command language used between a computer and a modem so that the computer can control the modem. Commands which may be issued include dial, dial number, clear, etc.
header	The bits within a block or cell which provide for correct delivery of the payload.
high bitrate digital subscriber line (HDSL)	A transmission technique allowing a high speed duplex signal of 2 Mbit/s to be transmitted to customer premises over existing copper lineplant.
high level data link control (HDLC)	A protocol conforming with the open systems interconnection (OSI) model, layer 2, allowing for the error-free conveyance of bits over a single link of a connection.
high usage route	First choice route in network topology.
holding time	The total time for which a circuit or exchange is in use on a call, commencing from the moment when the calling customer picks up the phone, or when the circuit or exchange is seized, until the time when the connection is cleared.
hub	A hardware device providing the physical interconnections and bus or ring topology necessary as the basis of a given type of local area network (e.g. ethernet, token ring, etc.)
Huffman code	A code used in datacommunications for data compression. This has the benefit of reducing the overall number of bits which need to be carried across the network, so improving speed of transmission and minimizing costs.
hybrid	A hybrid or hybrid transformer provides two-wire to four-wire transmission line conversion. Often used at each end of long transmission lines to convert the local customers line (if analogue and two-wire) to the four-wire central trunk section.
I.420 interface	An alternative expression to define the ISDN basic rate, or 2B + D interface. I.420 is the number of the CCITT recommendation in which the interface is defined.
IA2 alphabet	International alphabet number 2. A 5-bit mark/space code used between telex machines to represent the alphabetic, numeric and punctuation characters used in text.
IA5 alphabet	International alphabet number 5. A 7-bit computer code used to represent alphabetic, numeric and control characters. Developed from the ASCII code.
impulse noise	Impulse noise is characterized by large 'spikey' waveforms and it arises from unsuppressed power surges or mechanical switching noise. To the listener it manifests as a loud chatter or cracking sound.
in-band range	The term inband is usually applied to signal frequencies lying within the bandwidth of a normal speech channel, i.e. between 300 Hz and 3400 Hz.
in-band signalling	The term applied to tone signalling carried within the speech bandwidth of a telephone connection.
integrated services digital network (ISDN)	A development of the telephone and circuit-switched data network principle, in which a common, integrated and digital network is provided to cater for both voice and data applications.

intelligent network (IN)	A development of the telephone network in which the processing intelligence is centralized to service control points (SCP). This makes for more consistent management of control data and software and more rapid development of new services.
interactive service	A service allowing bidirectional communication, either conversational, messaging or retrieval.
interchange medium	The means used to interchange data between systems (either a storage or transmission means or both).
interference	A signal impairment caused by the interaction of an unwanted adjacent signal. Interference particularly affects FDM and radio systems.
intermodulation noise	Intermodulation noise is the name given to the undesirable signals which can result, of frequencies $F1 + F2$ , $F1 - F2$ , $F1 + 2F2$ , etc., when two different component signal frequencies $F1$ and $F2$ interact with one another.
international telephone and telegraph consultative committee (CCITT)	An influential subcommittee of the International Telecommunications Union (ITU), a leading body in technical and operational standards for international telephone networks. Recently renamed ITU-T.
Internet	The world's foremost data network, based on a large interconnected network of different organizations networks employing TCP/IP protocols.
Intranet	A corporate internal data network based upon Internet technology and routers.
invalid cell	A cell where the header is declared by the header error control process to contain errors.
ITU-R	International Telecommunications Union, Radiocommunications sector, responsible for the administration of radio spectrum worldwide and the development of technical standards for radiocommunication.
ITU-T	International Telecommunications Union, Standardization sector, responsible for the development of technical recommendations which are widely used as the basis for international telecommunications technology and networks.
ITU-T recommendation	A technical 'specification' of the ITU-T, however not having official 'standards' status, and not being mandatory.
jitter	Jitter or phase jitter arises when the timing of the pulses of a digital signal varies slightly, so that the pulse pattern is not quite regular. The effects of jitter can accumulate over a number of regenerated links and result in received bit errors.
junction	A circuit interconnecting two telephone or other circuit-switched exchanges. In particular, the word junction usually relates to the circuits directly interconnecting local exchanges in localized zones.
labelled channel	A temporary collection of all block payloads with a common label value
line code	The technique used when actually transmitting a digital signal onto a transmission line. The line code helps to ensure the synchronization and correct receipt of data.
line of sight (LOS)	Unobstructed line of sight between sending and receiving antennae is required for radio systems operating at frequencies much above about 1 GHz.
local area network (LAN)	A type of data network that allows the interconnection of computer devices within a relatively restricted area, typically lying along a spine or circle of length less than about 200 metres.
logical link control (LLC)	An OSI layer 2 protocol used in conjunction with local area networks (LANs).
logical signalling channel	A logical channel used to carry signalling information.
loudness rating (LR)	The loudness rating of a telephone circuit gives a measure of the received signal volume (in the listener's ear) compared with that spoken by the talker. A transmission plan allocates how much loss is acceptable overall and apportions network design limits.

managed object	A conceptual 'object' which may be monitored or controlled by means of a standard management protocol such as SNMP or CMIP. Managed objects are defined devices or functions, with a discrete number of possible states (e.g. on, off, transmit, receive, etc.)
management information base (MIB)	A collection of defined managed objects relating to given type of telecommunications device or interface (e.g. radio link, telephone exchange, subscriber ISDN line, etc.)
management plane	A conceptual communications connection between data switching devices to communicate network management information (e.g. to configure the network, setting up permanent connections or monitoring performance.)
Manchester code	The line code used in ethernet LANs.
mark	A binary digit of value 1.
media	Plural of medium.
media access control (MAC)	An OSI layer 1 protocol used in conjunction with a given LAN or MAN physical medium to provide a standard interface to the OSI layer 2 protocol.
medium	A means by which information can be perceived, expressed, stored or transmitted.
message handling system (MHS)	A conceptual model specified by ITU-T recommendation X.400, describing the way in which message devices (e.g. facsimile and telex machines, computers, word processors, etc.) intercommunicate.
message switching	A means of data communication in which complete messages are conveyed together across the transport medium.
messaging service	An interactive service based on store-and-forward and mailbox transfer.
meta-signalling	The procedure for establishing, checking and releasing signalling virtual channels.
metropolitan area network (MAN)	A type of data communications network suitable for interconnecting computer devices across a metropolitan area. Such networks typically have a bus or ring topology of up to about 100 km in total length.
microwave radio	Radio systems operating with a carrier frequency in the range 1–60 GHz. In telecommunications networks, microwave radio is typically used to span long distances and difficult terrain or to provide for rapid interconnection of two endpoints.
misinsertion	A quality problem associated with packet-switched, frame relay and ATM networks in which packets or frames of information are delivered in the wrong order or to the wrong network address. It may cause a significant information security risk.
mixed document	A document containing a mixture of text, graphics, data, image and moving pictures.
modem	A term derived from a shortening of modulator/demodulator. A modem provides for the carriage of digital data information over analogue transmission lines.
modulation	Modulation is a means of encoding information on to a carrier signal so that it may be transmitted over a radio link or transmission line.
monomode fibre	An optical fibre with a narrow central core of a different refractive index from the cladding which surrounds it. The narrowness of the core allows only very few ray paths to exist so that dispersion is minimized; a single mode of transmission prevails
motion pictures experts group (MPEG)	An industry grouping which has developed a number of coding standards for transmission of video signals (e.g. MPEG1, MPEG2 etc.)
multimedia service	A service in which interchanged information is a mixture of text, sound, graphics, video.



multimode	An optical fibre with a relatively wide central core. The relative large dimension of the core allows many different ray paths to exist, so that dispersion may be a problem. Typically multimode fibres are limited in range compared with monomode fibres.
multiplexing	Multiplexing is a technique to combine a number of full duplex channels together to share the same transmission line or radio medium. Unlike circuit multiplication, multiplexing does not affect the full duplex nature of the channels or signal quality.
multipoint	A communication or network configuration involving more than two end points.
network-node interface (NNI)	A standardized interface used between switches or sub-networks within a given (data) network.
network layer	The network layer is layer 3 of the open systems interconnection (OSI) model. This layer sets up an end-to-end connection across a real network, determining which permutation of individual links to be used.
network performance (NP)	A specific term applied to performance measurement parameters which are designed to monitor only the end-to-end network connection part of a given communication (e.g. telephone call or computer application running via a data network).
noise	Stray, unwanted and random signals interfering with the desired signal. Common forms of noise manifest as a low level crackling noise.
notched noise	Notched noise arises in data circuits, caused by the way in which the signal is processed over the course of the connection. Its name arises from the way it is measured, a pure frequency tone is applied at one end and removed with a notch filter.
NT1 (network terminating equipment 1)	The line terminating device providing customers' basic rate interface (2B + D) to the ISDN.
off-hook signal	The signal generated by a telephone on lifting the handset, indicating to the exchange that the customer wishes to make a call.
open shortest path first (OSPF)	A standardized protocol used widely in the Internet to convey routing information between routers.
open systems interconnection (OSI)	A conceptual model specified by the ITU-T X.200 series. The model describes the 7-layer process of communication between 'cooperating' computers. The model is the standard for open communication between computers made by different manufacturers.
optical fibre	A glass fibre of only a hair's breadth dimension, capable of carrying very high digital bit rates with very little signal attenuation.
OSI model	The seven layer conceptual model defining the process of communication between computer systems.
OSI network service	The service provided by a protocol conforming to the OSI network layer (layer 3). The network service provides an end-to-end connection for the use of the OSI transport layer (layer 4).
out-of-band range	The term used to describe the frequency range within an allocated 4 kHz circuit bandwidth but outside of the filtered speech range of 300–3400 Hz. Out-of-band interexchange signalling cannot easily be fraudulently emulated by end telephone users.
overflow routing	An alternative choice route in a network topology.
overlap signalling	A method of interexchange signalling in which dialled number digits are transferred from one exchange to the next exchange in the connection one-by-one. As soon as received, digits may also be onpassed to any subsequent exchange, even if the entire dialled.

packet	An information block identified by a label at layer 3 of the OSI model (e.g. X.25 packet)
packet-switching	A type of exchange or network which conveys a string of information from origin to destination by cutting it up into a number of packets and carrying each independently.
packet assembler/ disassembler (PAD)	A device enabling a character (i.e. asynchronous) terminal to be connected to a packet network.
packet transfer mode	A telecommunication transfer technique in which information is carried in packets.
parallel data transmission	Data transmission between computer devices using multiple lead wires or ribbon cables. Whole bytes of information (8 bits) are sent simultaneously by using a separate wire for each of the individual bits of the byte.
perception medium	The nature of information as perceived by the user (multimedia).
performance management	One of the five main network management functions defined by the ISO management model. The others are accounting, security, configuration and fault management.
periodic frame	A transmitted element which is repeated at intervals of equal time (e.g. 125 microseconds).
phase hits	Phase hits are impairments affecting data circuits. A phase hit is an intermittent but moderate and relatively short disturbance in the phase of the received signal. Where these accumulate over a multiple links they can affect the correct interpretation.
phase jitter	Phase jitter or simply jitter arises when the timing of the pulses on an incoming data signal varies slightly, so that the pulse pattern is not quite regular. Where these effects accumulate over a multilink connection, received bit errors can result.
phase modulation	A technique used to enable a low bit speed or comparatively low bandwidth information signal to be carried on a higher frequency carrier signal, by adjustment of the latter's phase to match the former's wave shape.
phase shift keying (PSK)	When a binary data signal is phase modulated onto a carrier the effect is to create a transmitted carrier signal of one of a number of fixed phases. Jumping between these phases according to the incoming bit stream is known as keying.
physical layer	Layer 1 of the open systems interconnection (OSI) model. The physical layer protocol is the hardware and software in the line terminating device which converts the databits needed by the datalink layer into the electrical pulses, tones or other form.
plesiochronous digital hierarchy (PDH)	Digital transmission hierarchy in which individual line systems within a network do not run in exact synchronization with one another. Instead free running or plesiochronous systems sometimes need to be corrected for slip and justification errors.
point to point protocol (PPP)	A protocol used in the Internet to connect serial end terminal devices, in particular by means of dial-up lines.
presentation layer	Layer 6 of the open systems interconnection (OSI) model. The presentation layer defines the manner in which the data is encoded, e.g. binary, ASCII, IA5, IA2, EBCDIC, facsimile, etc.
presentation medium	The means or device used to reproduce information to the user (multimedia).
primary multiplex (PMUX)	A device performing the first stage of time division multiplexing. In the UK and Europe, a primary multiplexor converts 32 analogue signals to 2.048 Mbit/s time division multiplexed signal. In the US equivalent, 24 are multiplexed to 1.544 Mbit/s.
primary rate interface (PRI)	The 23B + D or 30B + D ISDN interface typically used to connect ISDN PBXs to the public ISDN.

private branch exchange (PBX)	A small scale telephone exchange specifically intended for internal office usage.
private local interface	Synonymous with private UNI (ATM).
private network interface	Synonymous with private UNI (ATM).
proceed-to-send (PTS)	A backwards interexchange signal used to indicate the exchange's readiness, following the prompting of the forward seizure signal, to receive dialled digits for analysis, switching and onward call routing.
propagation delay	The delay encountered by a signal during the course of transmission from origin to destination.
proportional bidding	An interconnect methodology to distribute outgoing calls from one national telephone network between several competing network carriers in the destination country.
protocol	A protocol is a rule of procedure by which computer devices intercommunicate. Thus a protocol is the equivalent of a human language, with punctuation and grammatical rules.
PTO (public telecommunications operator)	A term of legal significance in the United Kingdom, applying to companies licensed to provide public network services, including telephone and data network services.
public network interface	Synonymous with public UNI.
pulse code modulation (PCM)	The method of conversion used to enable speech or some other analogue audio signal to be carried over a digital transmission path.
pulse metering	The historical method of monitoring telephone call duration for later customer invoicing. Pulses are generated at regular intervals during the call and recorded on the customers cyclic meter. A price per unit is levied on the customer's invoice.
Q-SIG	An ISDN signalling technique used on tie lines between ISDN PBXs.
Q3-interface	The standard interface between a TMN-compliant network element (network) and an umbrella network management operating system. The interface employs CMIP (common management information protocol).
quadrature amplitude modulation (QAM)	A complex method of modulation used in modern modems to allow very high data rates to be carried reliably and relatively error-free. QAM is a combination of phase and amplitude modulation.
quality of service (QOS)	Expressed as one of a number of different parameters which are designed to measure the performance of communications applications which operate by means of telecommunications networks.
quantization	A process fundamental to pulse code modulation, when converting an analogue signal into a digital equivalent. The analogue signal amplitude is coded regularly as a digital value corresponding to the nearest pre-defined quantization level.
quantization distortion unit (qdu)	When PCM encoded signals encounter nonlinear processing during their course of transmission then further quantization noise can result on final reception. The measure of overall error is made in qdus.
quantization noise	Quantization noise is heard when an analogue signal is reproduced at the receiving end of a pulse code modulated transmission link. It arises from the slight discrepancy between the reproduced <i>quantization</i> level and the original signal level.
quantizing	An alternative term for quantization.
radiopaging	A radio-based method for alerting field-based staff. Staff wear a cigarette box-sized unit which either beeps or displays a short alphanumeric message.
regenerator	A device inserted at an intermediate stage of a digital transmission line to counteract the effects of signal deterioration which occur on long lines. As far as possible the device regenerates the original bit stream.

repeater	A device inserted at an intermediate stage of an analogue transmission line to counteract the effects of signal deterioration (particularly attenuation) which occur on long lines.
representation medium	Information as described by its coded form (multimedia)
retrieval service	An interactive service providing for the accessing of data stored in data base centres.
ringing current	The current applied to a destination telephone to ring the bell and thereby alert its owner to answering it.
ringing tone	The tone, quite separate from ringing current, but applied at the same time, to inform the caller that the destination telephone is ringing.
router	A switching device used in the Internet or Intranet to determine the appropriate routing of a given data message through a data network to the desired destination.
routing table	A table held in the switch, router or exchange which relates dialled digit information to the appropriate outgoing route to be selected.
RS-232C	A standard interface, using a 25-pin D-type plug and connector by which computing devices may be physically interconnected. The standard also defines the layer 1 protocol for data transfer.
S/T interface	An alternative term used to describe the ISDN basic rate interface.
satellite	In telecommunications terms, a radio repeater station operating in the microwave frequency range, which has been placed in space to orbit the earth. Most telecommunications satellites nowadays are geostationary (described elsewhere).
SDLC (synchronous datalink control)	Synchronous data link control. IBM's equivalent of HDLC, forming the layer 2 of IBM's systems network architecture (SNA).
security management	One of the five main network management functions defined by the ISO management model. The others are accounting, configuration, performance and fault management.
segment of data	A measure of data quantity used to monitor data network (particularly packet data network) usage. A segment is equal to 512 bytes. In some technologies (e.g. DQDB) a segment is only 48 (or 64) bytes in length.
seizure signal	The interexchange signal sent forwards along the connection to alert the subsequent exchange of a desire to establish a new connection for a call currently being set up.
serial data transmission	Data transmission between computer devices using only a single circuit path. Whole bytes of information (8 bits) are sent in a sequential pattern. Serial transmission is used across long distance telecommunications networks (c.f. parallel transmission).
service attributes	Transmission capabilities and other functionality supported by a network for the carriage of a service.
service bit rate	The bit rate available to a user for information transfer.
service control elements	The primitive controls needed to set up, manage and release a multimedia service (multimedia).
session layer	Layer 5 of the open systems interconnection (OSI) model. When established for a session of communication, the two devices at each end of the overall connection must conduct their 'conversation' in an orderly manner.
shielded twisted pair (STP)	Term applied particularly to indoor cabling of a type in which twisted copper pairs are 'wrapped' immediately inside the plastic sheathing with a foil shield. The shielding serves to reduce electromagnetic disturbances from or to the cable.
sidetone	The term used to describe the talker's own voice when heard immediately in his own earpiece (if there is any delay then the effect is instead called echo). With no sidetone, telephone users think that the line is dead, but too much causes distraction.

signalling	The means by which destination address information is conveyed from the end device (e.g. telephone) to the network exchange and subsequently relayed from node to node through the network to establish a connection.
signalling virtual channel	A virtual channel for transporting signalling information.
simple network management protocol (SNMP)	A network management protocol widely used for managing network equipment, corporate data networks and the Internet.
simplex	A transmission means allowing only one direction of transmission (for example, public broadcast radio).
single sideband (SSB)	A method of bandwidth and power economy used in radio and FDM systems.
sound retrieval service	On-demand (user-initiated) retrieval of music or sound track.
space	A term used to describe a bit of binary value 0.
space switching	A technique used for switching telecommunications connections based on a matrix of separate physical switching points.
speech interpolation	A method of achieving circuit multiplication, by interspersing other clippets of other conversations in the gaps between individual words, sentences, and in the period while the distant user is talking.
spread spectrum radio	A method of modulating an information signal onto a broadband radio carrier which helps to minimize the effects of noise, radio fading and the risk of signal interception.
statistical multiplexing	A method of data lineplant economy similar in operation to speech interpolation devices. Various data communication conversations share the same line by making use of each other's idle periods.
storage medium	The physical means to store information or data (multimedia).
supplementary service	Supplementary services exist in ISDN. On their own they have no value, but serve to add a supplementary value to a bearer service such as speech or 64 kbit/s data. An example of a supplementary service is calling line identity or ring back when free.
suppressed carrier operation	Mode of radio system operation in which transmit power is saved by suppressing the carrier before transmission. The synthesised carrier must be added to the received signal prior to demodulation.
switched multimegabit data service (SMDS)	A high speed metropolitan area network (MAN) based on an extended bus topology which uses the DQDB (dual queue distributed bus) technique.
synchronization	The method by which the bit patterns appearing on digital line systems may be properly clocked and interpreted allowing the beginning of particular patterns and frame formats to be correctly identified.
synchronous data transfer	Data transfer employing a strictly regular pattern, rather than (as in asynchronous operation) using start and stop bits to distinguish character patterns from idle line operation.
synchronous digital hierarchy (SDH)	Modern digital transmission hierarchy in which individual line systems within a network are designed to run exactly synchronously with one another. This gives major management and topology administration benefits.
synchronous optical network (SONET)	The North American technology which preceded SDH (synchronous digital hierarchy) and is based on similar principles.
synchronous transfer mode	A transfer mode which offers a periodic fixed length frame to each connection.
systems network architecture (SNA)	A standard framework of communications methods used in particular in networks of IBM computer devices.
TCP/IP	The protocol suite used as the basis of the Internet.
telecommunications management network (TMN)	The technical architecture conceived by ITU to provide for overall coordinated management of telecommunications networks.

teleservice	A term used to describe a particular type of ISDN service, a teleservice comprises all elements necessary to support an end user's particular purpose or application. Examples of teleservices include telephony, facsimile, videotelephone etc.
telex	An old-style teletype device used to convey character-based messages.
terminal adaptor	A device used to convert an older style device so that it can run communicate across a modern network.
terminal equipment	The user's end equipment, connected to one end of a telecommunications connection or network port.
throughput	The number of bits in a block successfully transferred across a network per unit time.
time division multiple access (TDMA)	A technique used in point-to-multipoint radio systems. TDMA shares the available radio spectrum by allocating timeslots of a high bitrate radio bearer channel to each user. The main advantage over CDMA and FDMA is the very flexible means of bitrate sharing.
time division multiplex (TDM)	A technique allowing a number of different digital signals to share the same high bit rate digital transmission line by dividing it into a number of smaller component bit rates, each interleaved in time with the others.
time space time	A technique used for switching telecommunications connections based on a combination of time switching and a matrix of separate physical switching points. This is the most commonly used form of switching in digital switches.
token ring LAN	A popular type of physical infrastructure used for local area network interconnection of PCs and other computer devices within an office environment. It is used particularly in networks of IBM mainframes and personal computers.
traffic intensity	The measure of instantaneous demand on a network, quantified in terms of the number of users using the network simultaneously.
transfer mode	A telecommunications technique comprising transmission, multiplexing and switching.
transit delay	The time difference between the entry of the first bit and exit of the last bit from a network.
transmission bridge	The circuitry provided in telephone exchanges to power individual customers lines but retain isolation between them so that they do not overhear one another's conversations.
transmission medium	The physical means to transmit information (multimedia).
transmultiplexor (TMUX)	A device enabling an FDM group or supergroup to be carried on digital line plant.
transponder	A term derived from transmitter/responder used to describe the radio repeater device associated with telecommunications satellites.
transport layer	Layer 4 of the open systems interconnection (OSI) model. The transport layer provides for end-to-end data relaying service across any type of data network (e.g. packet network or circuit switched network).
trellis coding	An efficient form of inter-modem modulation enabling high bitrate digital signals to be transferred over analogue lines.
tropospheric scatter	A form of radio wave reflection from the earth's troposphere, allowing over the horizon communication.
trunk	A circuit interconnecting telephone exchanges. Particularly a long distance circuit.
twisted pairs	Metallic conductor wires, twisted in pairs corresponding to the legs of a two-wire circuit.
unbalanced transmission	Telecommunications transmission across metallic conductors (i.e. wire pairs) in which the two conductors do not play an equal role (e.g. 75 ohm coaxial cabling).

unshielded twisted pair (UTP)	Term applied particularly to indoor cabling of a type in which twisted copper pairs are sheathed directly in a plastic covering without a foil shield. This is the simplest and cheapest form of indoor telephone and data network cabling.
usage parameter control	The execution of appropriate action when negotiated values of information transfer are exceeded.
user-network interface (UNI)	The interface (physical interface and protocols) between a user's terminal equipment and the network.
user plane	A conceptual end-to-end communications connection between user terminal equipments for communicating user information (i.e. the useful information content of a communications session).
valid cell, packet or frame	A cell in which the header is declared by the header error control process to be free of errors.
variable bit rate service	A telecommunication service using a service bit rate based on statistical parameters.
video on-demand	A video retrieval service.
videoconferencing	A means of video communication via a telecommunications network, allowing a meeting or conference to be held between participants in diverse locations.
virtual channel connection	A concatenation of virtual channel links.
virtual channel link	ATM connection between a point where a virtual channel identifier is assigned and a point where it is removed.
virtual connection	A means of conveying data from origin to destination without there ever being a dedicated physical path between the two. An example is a packet-switched connection.
virtual path connection	A concatenation of virtual path links.
virtual path link	ATM connection between a point where a virtual path identifier is assigned and a point where it is removed.
voicemail	A means for leaving speech messages for absent or otherwise-engaged telephone network users.
weighted noise	Weighted noise is a measure of noise in the middle of a channel bandwidth (in other words the noise which is hardest to remove by simple filtering).
wide area network	A term applied to describe particularly the long distance part of a corporate data network.
worldwide web	An interconnected network of computer servers around the world, which provide for fast and easy access to online information via the Internet.

---

## *Glossary of Abbreviations*

---

1VF	One voice frequency (signalling)	AAR	Automatic alternative routing
2VF	Two voice frequency (signalling)	AARP	Appletalk address resolution protocol
2w	two-wire		
3270 protocol	Protocol used to terminal (IBM SNA)	ABM	Asynchronous balanced mode (HDLC)
4B3T	four-binary three-tertiary (ISDN line code)	ABR	Answer bid ratio
4w	four-wire	ABR	Available bit rate
A	Applicable	AC	Access control (IBM token ring LAN)
A	Agent	AC	Alternating current
A/D	Analogue-to-digital conversion	ACB	Access method control block (IBM VSAM or VTAM)
A + S	Alarm and state management	ACD	Automatic call distribution
A =	Administrative domain (X.500)	ACE	Access connection element
A-law	PCM speech encoding (European standard)	ACF	Access control field (DQDB)
AA	Administrative authority (numbering)	ACF	Advanced communication function (IBM SNA)
AAA	Autonomous administrative area	ACI	Access control information
AAI	Administration authority identifier	ACK	Acknowledgement
AAL	ATM adaptation layer	ACR	Available cell rate
AAL-IE	AAL information element	ACSE	Association control service element (OSI layer 7)
AAL-PCI	AAL protocol control information		
AAL-SDU	AAL service data unit	ADC	Advice of duration and charge
AAL1	AAL type 1 (constant bit rate service)	ADDMD	Administration directory domain
		ADM	Add/drop multiplexor (SDH)
AAL2	AAL type 2 (variable bit rate, but time critical service)	ADMD	Administration management domain (X400)
AAL3/4	AAL type 3/4 (class C frame relaying service or class D connectionless service)	ADP	Answer detection pattern
		ADPCM	Adaptive differential pulse code modulation
AAL5	AAL type 5 (class C frame relaying service or class D connectionless service), a simple alternative to AAL3/4	ADSL	Asymmetric digital subscriber line
		ADSP	Appletalk data stream protocol (Appletalk)
		AE	Application entity (OSI layer 7)



AEP	Appletalk echo protocol (Appletalk)	ARR	Automatic repeat request
AFC	Automatic frequency control (radio)	ARR	Automatic re-routing
AFI	Address format identifier or authority and format identifier (numbering)	ARS	Address resolution server
AFP	Appletalk filing protocol (Appletalk)	ARS	Automatic route selection
AGC	Automatic gain control	ASAI	Adjunct/switch application interface
AI	Air interface	ASC	Accredited standard committee
AII	Active input interface (optical fibre)	ASCII	American standard code for information interchange
AIM	ATM inverse multiplexing	ASE	Application service element (OSI layer 7)
AIP	Air interface packet	ASK	Amplitude shift keying
AIS	Alarm indication signal	ASN.1	Abstract syntax notation (old version in X.208, new version is X.680)
AL	Alignment (AAL3/4 CPCS)	ASO	Application service object
AL	Access link	ASP	Appletalk session protocol (Appletalk)
AM	Accounting management	ASR	Answer seizure ratio
AM	Amplitude modulation	AT&T	American Telephone and Telegraph Company
AMI	Alternate mark inversion (line code)	ATE	ATM terminating equipment
AMPS	Advanced mobile telephone system (American cellular radio)	ATM	Asynchronous transfer mode
AMS	Audiovisual multimedia services	ATM-SDU	ATM service data unit
AN	Access network or node	ATME	Automatic transmission, measuring and signalling text equipment
ANA	Article numbering association (UK)	ATMM	ATM management
ANI	Automatic number identification (US equivalent of CLI)	ATP	Appletalk transaction protocol (Appletalk layer 4)
ANL	Automatic noise limiter	ATPC	Automatic transmit power control (radio)
ANR	Automatic network routing	ATS	Abstract test suite
ANS	American national standard	AU	Administrative unit (SDH terminology)
ANSI	American National Standards Institute	AU- <i>n</i>	Administrative unit of order <i>n</i> (SDH)
AOI	Active output interface (optical fibre)	AuC	Authorization centre
AOM	A-profile sub-class: OSI management	AUG	Administrative unit group (SDH)
AOW	Asia-Oceania workshop (conformance testing body)	AUI	Attachment unit interface (ethernet LAN)
AP	Acknowledgement packet	AUU	ATM layer user-to-user indication
AP	Application process	AVI	Audiovisual interactive service
APDU	Application protocol data unit	AVMMS	Audiovisual multimedia service
API	Application programming interface	AW	Administrative weight
APPC	Advanced program-to-program communication (IBM LU6.2)	B	Binary
APPL	Application program	B-BC	Broadband bearer capability
APPN	Advanced peer-to-peer networking (IBM)	B-HLI	Broadband higher layer information
APS	Automatic protection switching	B-ICI	Broadband inter carrier interface
ARC	Alarm reporting capability	B-ISDN	Broadband integrated services digital network
ARE	All routes explorer	B-ISDN PRM	Protocol reference model for B-ISDN
ARL	Adjusted ring length	B-ISPBX	Private branch exchange for B-ISDN
ARP	Address resolution protocol (Internet)	B-ISUP	Broadband integrated services user part
ARPA	Advanced research project agency		
ARQ	Automatic repeat request		

B-LLI	Broadband lower layer information	BCVT	Basic class virtual terminal
B-NT	Network Termination for B-ISDN	Be	Excess burst duration (frame relay)
B-NT1	Network Termination 1 for B-ISDN	BECN	Backward explicit congestion notification (frame relay)
B-NT2	Network Termination 2 for B-ISDN (multipoint configuration)	BEDC	Block error detection code (ATM OAM performance cell)
B-SP	B-ISDN signalling point	BER	Basic encoding rules
B-STP	B-ISDN signalling transfer point	BER	Bit error ratio
B-TA	Terminal Adaptor for B-ISDN	BFO	Beat frequency oscillator
B-TE	Terminal Equipment for B-ISDN	BGP	Border gateway protocol (Internet)
B-TE1	B-ISDN Terminal Equipment type 1 (designed for ATM)	BHCA	Busy hour call attempts (telephone tariff)
B-TE2	B-ISDN Terminal Equipment type 2 (connected via B-TA)	BIND	Berkeley Internet name domain
B8ZS	Bipolar eight-zero substitution (line code)	BIP	Bit interleaved parity code (block error detection code)
BABT	British Approvals Board for Telecommunications	BIP	Bit inserted parity
BAKOM	Bundesamt für Telekommunikation (Switzerland)	BIS	Border intermediate system
balanced	A type of data transmission using twisted pair cable (usually 120 Ohm)	BIS	Bringing into service
Balun	Balance-to-unbalance transformer	BIS PDU	Border intermediate system protocol data unit
BAPT	Bundesamt für Post und Telekommunikation (Germany)	BISSI	Broadband inter-switching system interface
BAS	Bitrate allocation signal	Bisync	IBM protocol (BSC)
BASize	Buffer allocation size (AAL 3/4 CPCS)	bit/s	Bits per second
BB	Baseband	BIU	Basic information unit (IBM SNA)
BBN	Bolt, Beranek & Newman (data switch manufacturers)	BLER	Block error result (ATM OAM performance cell)
Bc	Committed burst size (frame relay)	block	A unit of information consisting of a header and an information field
BCC	Bearer channel connection protocol (V5)	BML	Business management layer (TMN management model)
BCC	Block check character	BMPT	Bundesministerium für Post und Telekommunikation (Germany)
BCD	Binary coded decimal	BN	Bridge number
BCDBS	Broadband connectionless data bearer service	BNN	Boundary network node
BCN	Backward congestion notification	BOC	Bell operating company (USA)
BCOB	Broadband connection-oriented bearer (service)	BOM	Beginning of message
BCOB-A	Broadband connection-oriented bearer A: constant bit rate ATM transport with stringent end-to-end timing needs	BOOTP	Bootstrap protocol
BCOB-C	Broadband connection-oriented bearer C: variable bit rate ATM transport with no end-to-end timing needs	BOP	Bit-oriented protocols
BCOB-X	Broadband connection-oriented bearer X: ATM-only: network will not process layers higher than ATM layer	BPON	Broadband passive optical network
		BPP	Bridge port pair (service routing descriptor)
		BPSK	Binary phase shift keying
		BRA	Basic rate access (ISDN)
		BRI	Basic rate interface (ISDN)
		BRL	Balance return loss
		Brouter	Combination bridge and router
		BS	Base station
		BSC	Base station switching centre
		BSC	Binary synchronous communication (IBM)

BSD	Berkeley system distribution	CBR	Constant bitrate
BSI	British Standards Institution	CCBS	Call connection to busy subscriber (ring back when free)
BSS	Base station sub-system	CCD	Contract completion date
BT	British Telecommunications plc	CCD	Charge-coupled device
Btag	Beginning tag (AAL3/4 CPCS)	CCDP	Co-channel dual polarized operation (radio)
BTAM	Basic telecommunications access method	CCIR	International radiocommunication consultative committee (now ITU-R, the radiocommunication sector of ITU)
BTS	Base transmitter station	CCIS	Common channel interoffice signalling
BUS	Broadcast and unknown server	CCITT	International telephone and telegraph consultative committee (now ITU-T, the telecommunication standardization sector of ITU)
BVPS	Broadband virtual path services	CCR	Commitment, concurrency and recovery services (OSI layer 7)
BWO	Backward wave oscillator	CCS	Common channel signalling
BZT	Bundesamt für Zulassung in der Telekommunikation (Germany)	CCS	Hundred call seconds (telephone traffic)
C	Characteristic	CCS	Console communication service (IBM SNA)
C	Conditional	CCU	Communications control unit
C	Capacitance	CCW	Channel command word (IBM ESCON)
C	Network C (German analogue mobile telephone network)	CD	Collision detection
C/I	Carrier to interference ratio (radio)	CD	Countdown counter (DQDB)
C/R	Command/response	CDDI	Copper-stranded distributed data interface (FDDI)
C =	Country identification (X.500)	CDLC	Cellular data link control
C-Band	Satellite transmission	CDMA	Code division multiple access (radio)
C-channel	Communication channel	CDS	Common directory service
C- <i>n</i>	Container of bitrate order <i>n</i> (SDH)	CDV	Cell delay variation (ATM)
C7/BT	British Telecom version of SS7	CDVT	Cell delay variation tolerance
CA	Customer access (TMN)	CE	Connection endpoint
CAC	Channel access control	CE	Connection element
CAC	Connection admission control	CEC	Cell error control (physical layer)
CAD	Computer-aided design	CEC	Commission of the European Community
CAD-CAM	Computer-aided design/Computer-aided manufacturing	CEI	Connection endpoint identifier
CAI	Common air interface	CELP	Code-excited linear prediction
CAI	Computer-assisted instruction	CEN	Comite Europeen de Normalisation
CAMC	Customer access maintenance centre	CEP	Connection endpoint
CAMF	Customer access management function (TMN)	CEPT	European Conference for Posts and Telecommunications
CAP	Carrierless amplitude modulation/phase modulation	CEQ	Customer equipment
CAS	Channel-associated signalling	CER	Cell error ratio (ATM)
CASE	Common application service element	CES	Connection endpoint suffix
Cat. 1	Category 1 (cabling)	CES	Circuit emulation service
Cat. 2	Category 2 (cabling)		
Cat. 3	Category 3 (cabling)		
Cat. 4	Category 4 (cabling)		
Cat. 5	Category 5 (cabling)		
CATV	Common antenna television		
CB	Central battery		
CB	Citizens band (radio)		
CBDS	Connectionless broadband data service		
CBO	Continuous bit stream oriented		

CFR	Code of federal regulations (USA)	CMISE	Common management information service entity
CFS	Common functional specifications (RACE)	CMOL	CMIP over logical link control
Channel	IBM host channel	CMOT	Common management information services over TCP/IP
CHPID	Channel path identifier (IBM ESCON)	CMR	Cell misinsertion rate (ATM)
CI	Congestion indication	CMT	Connection management mechanism (FDDI)
CI	Customer installation (TMN)	CN	Congestion notification
CIB	CRC-32 indicator bit	CN	Customer network
CIC	Circuit identification code	CNLS	Connectionless
CID	Communication identifier (IBM VTAM)	CNM	Customer network management
CIDR	Classless interdomain routing	CNR	Complex node representation
CIM	Computer-integrated manufacturing	CNT	Communications name table
CIME	Customer installation maintenance entities	CO	Central office
CIMF	Customer installation management function (TMN)	CO	Connection oriented
CIR	Committed information rate (frame relay)	CO&M	Centralized operation and maintenance
CL	Connectionless	CODEC	Coder/decoder
Clav	Cell available (flow control)	COH	Connection overhead
CLEC	Competitive local exchange carrier	COI	Connection-oriented internetworking
CLI	Calling line identity	COLT	City of London telecommunications (network operator)
CLI	Connectionless internetworking	COM	Continuation of message
CLIP	Calling line identity presentation	CON	Concentrator
CLIR	Calling line identity restriction	Conf	Confirm
Clk	Clock	CONF	Confirm
CLLM	Consolidated link layer management (FR)	CONS	Connection-oriented network service (OSI)
CLNAP	Connectionless network service access protocol	COP	Character-oriented protocols
CLNP	Connectionless network protocol	CORBA	Common object request broker architecture
CLNS	Connectionless network service (OSI)	COS	Class of service
CLP	Cell loss priority (ATM)	COSM	Class of service manager
CLR	Cell loss ratio (ATM)	COSS	Connection-oriented session service
CLSF	Connectionless service function	COTS	Connection-oriented transport service
CLTP	Connectionless transport protocol	CP	Common part
CLTS	Connectionless transport service	CP	Control point
CLU	Control logical unit (IBM)	CP	Control program
CM	Configuration management	CPCS	Common part convergence sublayer (AAL3/4 and AAL5)
CME	Circuit multiplication equipment	CPCS-CI	CPCS-congestion indication
CME	Common management environment	CPCS-LP	CPCS-loss priority
CME	Connection management entity	CPCS-UU	CPCS-user-to-user indication
CMI	Coded mark inversion (line code at the physical layer when coaxial cable is used)	CPE	Customer premises equipment
CMIP	Common management information protocol	CPI	Common part indicator (CPCS AAL3/4 or AAL5)
CMIS	Common management information service	CPN	Customer premises network
		CPU	Central processing unit

CR	Command/response	DASS2	Digital access signalling system (UK ISDN signalling for PBXs)
CR	Conditional requirement	dB	Decibel
CRC	Cyclic redundancy check	dBi	Decibel relative to isotropic antenna
CRC- <i>n</i>	Cyclic redundancy check of <i>n</i> bits length	dBm	Power in decibels relative to 1 milliWatt
CRE	Corrected reference equivalent	dBø	Power in decibels relative to 1 Watt
CRF	Connection-related function	DBPSK	Differential binary phase shift keying
CRF (VC)	Virtual channel connection-related function	DC	Direct current
CRF (VP)	Virtual path connection-related function	DCC	Data country code (X.121)
CRM	Cell rate margin	DCE	Data circuit terminating equipment
CRMA	Cyclic reservation multiple access	DCE	Distributed computer environment
CRS	Cell relay service	DCF	Data communications function
CRT	Cathode ray tube	DCF	Discounted cashflow
CS	Circuit switching	DCL	Digital command language (Digital equipment corporation)
CS	Convergence sublayer (ATM)	DCME	Digital circuit multiplication equipment
CS-PDU	Convergence sublayer protocol data unit	DCN	Data communications network
CS1	Capability set 1 (SS7 INAP)	DCS	DECT control system
CSI	Convergence sublayer indication bit (AAL1)	DCS	Digital cellular system (mobile telephone system)
CSMA	Carrier sense multiple access	DDCMP	Digital data communications message protocol (DECnet)
CSMA/CD	Carrier sense multiple access with collision detection	DDDB	Distributed directory database (IBM APPN)
CSR	Cell missequenced ratio	DDI	Direct dialling in
CSU	Channel service unit	DDP	Datagram delivery protocol (Appletalk layer 3)
CT2	second generation cordless telephony	DDS	Digital data system
Ctag	Confirmation tag	DDS	Direct digital synthesis
CTD	Cell transfer delay	DE	Directory entries (X.500)
CTE	Channel translating equipment	DE	Discard eligibility (frame relay or FR-SSCS)
CTP	Connection termination point	DEA	Directory entry attributes (X.500)
CTS	Clear to send	DEC	Digital Equipment Corporation
CTS	Common transport semantics	DECNET	Networking protocol of the Digital equipment corporation
CTS	Conformance test services	DECT	Digital European cordless telephone
CUG	Closed user group	DES	Defence encryption standard
CV	Code violation	DES	Data encryption standard
CW	Continuous wave	DEV	Device address field
D	Direction	DFA	DXI frame address
D/C	DLCI or DL-CORE control indicator (frame relay or FR-SSCS)	DFI	Domain specific part format identifier (numbering)
DA	Destination address	DFT	Discrete Fourier transform
DAL	Dedicated access line	DGP	Dissimilar gateway protocol
DAM	DECT authentication module	DGPT	Direction de la Reglementation Generale de Postes et Telecommunications (France)
DAMA	Demand assigned multiple access		
DAP	Data access protocol (DECnet)		
DAP	Directory access protocol		
DAR	Dynamic alternative routing		
DAS	Directory system agent		
DASD	Digital access storage device		

DGT	Dirección General de Telecomunicaciones (Spain)	DRRS	Digital radio relay system
DGT	Director General of Telecomunicaciones	DS	Default standby
DH	DMPDU header (DQDB)	DS	Digital section (transmission line)
DIB	Directory information base (X.500)	DS	Direct sequence
DID	Direct inward dialling	Ds-type	ISDN D-channel signalling information
DIS	Draft international standard	DS0	64 kbit/s transmission line or channel
DISC	Disconnect	DS1	Digital section level 1 (North American standard, T1, 1544 kbit/s)
DISP	Directory information shadowing protocol	DS3	Digital section level 3 (North American standard, T3, 45 Mbit/s)
DIT	Directory information tree (X.500)	DSA	Directory service agent (X.500)
DLCI	Data link connection identifier (frame relay)	DSAP	Destination service access point
DLS	Data link switching	DSE	Data switching exchange
DLU	Destination logical unit	DSG	Default slot generator (DQDB)
DM	Delta modulation	DSL	Digital subscriber line
DM	Disconnected mode	DSP	Digital signal processing
DMA	Deferred maintenance alarm	DSP	Directory system protocol
DMD	Directory management domain	DSP	Domain specific part (numbering)
DME	Distributed management environment	DSS	Digital signalling system
DMI	Digital multiplex interface	DSS	Distributed sample scrambler
DMO	Domain management organization	DSS1	Digital subscriber signalling system 1 (Q.931 for ISDN)
DMPDU	Derived MAC (media access control) PDU (DQDB)	DSS2	Digital subscriber signalling system 2 (ATM)
DMT	Discrete multitone transmission	DSSS	Direct sequence spread spectrum
DMTF	Desktop management task force	DSU	Data service unit
DN	Directory number	DSU	Digital service unit (digital line terminating device)
DN	Distinguished name	DT	DMPDU trailer (DQDB)
DNA	Digital network architecture (Digital equipment corporation)	DTE	Data terminal equipment
DNHR	Dynamic non-hierarchical routing	DTI	Department of Trade and Industry (UK)
DNIC	Data network identification code (X.121)	DTL	Designated transit list
DNS	Domain name system or server	DTMF	Dual tone multifrequency (signalling)
DOD	Department of defense (USA)	DTP	Data transfer protocol
DOP	Directory operational and binding management protocol	DTR	Data terminal ready
DOS	Disk operating system	DU	Data unit
DPC	Destination point code	DUA	Directory user agent (X.500)
DPCM	Differential pulse code modulation	DUP	Data user part (SS7)
DPL	Primary link for distribution services	DUT	Device under test
DPNSS	Digital private network signalling system (inter PBX signalling)	DXC	Digital crossconnect
DPSK	Differential phase shift keying	DXI	Data exchange interface
DQ	Directory enquiry	E&M	E & M signalling (sometimes 'ear' and 'mouth')
DQDB	Distributed queue dual bus (the technology behind SMDS)	E-DSS1	European digital subscriber signalling 1
DQPSK	Differential quaternary phase shift keying (TETRA)	E-R	Entity relationship
		E.163	ITU-T recommendation defining the telephone number scheme

E.164	ITU-T recommendation defining the ISDN numbering scheme	EL	Electrical (interface)
E1	Digital transmission section level 1 (European standard 2048 kbit/s)	EL	Element layer (TMN)
E3	Digital transmission section level 3 (European standard 34 Mbit/s)	ELAN	Emulated local area network
EA	Extended address	ELAP	EtherTalk link access protocol (Appletalk layer 2 for Ethernet LAN)
EA	Address extension bit	ELLC	Extended logical link control (IBM SNA)
EAB	Extended address bit	EM	Electromagnetic
EBCDIC	Extended binary coded decimal interchange code	EMA	Enterprise management architecture (Digital equipment corporation)
EBCN	Explicit backward congestion notification	EMC	Electromagnetic compatibility
EC	Echo canceller	EMF	Electromotive force
EC	Error correction	EMI	Electromagnetic interference
EC	European Community	EML	Element management layer (TMN management model)
ECC	Embedded control channel	EN	Edge node (IBM APPN)
ECMA	European Computer Manufacturers Association	EN	Equipment number
ECS	Encryption control signal	EN	Exchange number
ECSA	Exchange Carriers Standards Association (USA)	Enb	Enable
ED	Error detection	EOM	End of message
ED	Ending delimiter (IBM token ring LAN)	EOT	End of transmission
EDC	Error detection code (physical layer)	EP	Emulation program
EDI	Electronic data interchange	EPHOS	European procurement handbook for open systems
EDIFACT	Electronic data interchange for administration, commerce and transport	EPOS	Electronic point of sale
EEC	European economic community	EPP	Endpoint processor
EEI	External environment interface	EQ	Enquiry
EFCI	Explicit forward congestion indication	ER	Explicit route
EFCN	Explicit forward congestion notification	ERA	Exterior reachable address
EFD	Event forwarding discriminator	ERL	Echo return loss
EFI	Errored frame indicator	Erlang	Unit of traffic intensity measurement/telephone
EFS	Error-free seconds	ERM	Event report management
EFTPOS	Electronic funds transfer at point of sale	ERP	Error recovery protocol
EGP	Edge gateway protocol (Internet)	ES	Echo suppressor
EGP	Exterior gateway protocol	ES	End system
EGRP	Exterior gateway routing protocol	ES	Errored seconds
EIA	Electronic Industries Association (USA)	ESCON	Enterprise systems connection (IBM)
EIR	Equipment identity register	ESI	End system identifier (numbering)
EIR	Excess information rate (frame relay)	ESIG	European SMDS interest group
EIRP	Equivalent isotropic radiated power	ET	Exchange termination
		ETACS	Extended TACS (mobile telephone system)
		Etag	End tag (AAL3/4 CPCS)
		ETR	Early token release
		ETS	European telecommunications standard
		ETSI	European Telecommunications Standards Institute
		ETX	End of text
		EU	European Union

EWOS	European workshop for open systems (conformance testing body)	F6 state	FC3 + FC4 or FC1 + FC3 + FC4 on the user side of the UNI (physical layer)
f	Reference point between operating system function and workstation function (TMN)	F7 state	Power on state (transient state) at the user side of the UNI (physical layer)
f	Frequency	FADU	File access data unit (FTAM)
f-type data	ISDN D-channel data of frame type (i.e. SAPI = 32–62)	FAS	Frame alignment signal or sequence
F0 state	Loss of power on the user side of the UNI (physical layer)	FC	Frame control (IBM token ring LAN, FDDI)
F1 cell	Physical layer OAM cell (regenerator level)	FCAPS	Fault, configuration, accounting, performance, security (OSI management model)
F1 flow	Protocol communication at the regenerator section level (ATM physical layer)	FCC	Federal Communications Commission (USA)
F1 level	Regenerator section level (ATM physical layer)	FC <sub>n</sub>	Fault condition number <i>n</i> (physical layer); the various fault conditions are defined in ITU-T I.432
F1 state	Operational state on the user side of the UNI (physical layer)	FCS	Frame check sequence
F2 cell	Physical layer OAM cell (digital section level)	FDD	Frequency division duplex (radio)
F2 flow	Protocol communication at the digital section level (ATM physical layer)	FDDI	Fibre distributed data interface
F2 level	Digital section level (ATM physical layer)	FDM	Frequency division multiplexing
F2 state	Fault condition number 1 (FC1) on the user side of the UNI (physical layer)	FDMA	Frequency division multiple access (radio)
F3 cell	Physical layer OAM cell (transmission path level)	FDX	Full duplex
F3 flow	Protocol communication at the transmission path level (ATM physical layer)	FE	Functional elements
F3 level	Transmission path level (ATM physical layer)	FE	Functional entity
F3 states	Fault condition number 2 (FC2) on the user side of the UNI (physical layer)	FEA	Functional entity actions
F4 flow	Protocol communication at the virtual path level (ATM layer)	FEAC	Far end alarm and control
F4 level	Virtual path level (ATM layer)	FEBE	Far end block error
F4 state	Fault condition number 3 (FC3) or FC1 and FC3 on the user side of the UNI (physical layer)	FEC	Forward error correction
F5 flow	Protocol communication at the virtual channel level (ATM layer)	FECN	Forward explicit congestion notification (frame relay)
F5 level	Virtual channel level (ATM layer)	FEP	Front end processor
F5 state	Fault condition number 4 (FC4) on the user side of the UNI (physical layer)	FERF	Far end receive failure (subsequently renamed RDI, remote defect indication)
		FEXT	Far end crosstalk
		FFH	Fast frequency hopping
		FFT	Fast Fourier transform
		FG	Frame generator
		FH	Frame handler
		FH	Frequency hopping
		FI	Format identifier
		FID	Format identification
		FIFO	First-in first-out (queueing mechanism)
		FIN	Final
		FITL	Fibre in the loop
		FM	Fault management
		FM	Frequency modulation
		FMBS	Frame mode bearer service



FMH	Function management header (IBM SNA)	G10 state	FC3 + FC4 or FC2 + FC3 + FC4 on the network side of the UNI (physical layer)
FP	Fixed radio part (DECT)	G11 state	FC1 + FC2 + FC3 on the network side of the UNI (physical layer)
FPLMTS	Future public land mobile telephone system	G12 state	FC1 + FC3 + FC4 or FC1 + FC2 + FC3 + FC4 on the network side of the user interface (physical layer)
FR	Frame relay (correctly frame relaying)	G13 state	Power-on state (transient state) on the network side of the UNI (physical layer)
FR-IWP	Frame relaying interworking point	G2 state	Fault condition number 1 (FC1) on the network side of the UNI (physical layer)
FR-SSCS	Frame relaying service specific convergence sublayer	G3 state	Fault condition number 2 (FC2) on the network side of the UNI (physical layer)
FRAD	Frame relay access device	G4 state	Fault condition number 3 (FC3) on the network side of the UNI (physical layer)
frame	A block of variable length identified by a header label	G5 state	FC4 or FC4 + FC2 on the network side of the UNI (physical layer)
FRBS	Frame relaying bearer service	G6 state	FC1 + FC2 conditions on the network side of the UNI (physical layer)
FRFH	Frame relay frame handler	G7 state	FC1 + FC3 conditions on the network side of the UNI (physical layer)
FRM	Fast resource management	G8 state	FC1 + FC4 or FC1 + FC2 + FC4 conditions on the network side of the UNI (physical layer)
FRMR	Frame reject	G9 state	FC2 + FC3 conditions on the network side of the UNI (physical layer)
FRS	Frame relay service	GAP	Gateway access protocol (DECnet)
FRS	Fixed relay station (DECT)	Gbit/s	Gigabits per second (1000 million bits/s)
FRSE	Frame relay switching equipment	GBSVC	General broadcast signalling virtual channel
FRTE	Frame relay terminal equipment	GCAC	Generic connection admission control
FS	Frame status (IBM token ring LAN)	GCRA	Generic cell rate algorithm
FSBS	Frame switching bearer service	GDMO	Guidelines for the definition of managed objects (TMN)
FSK	Frequency shift keying	GECOS	General Electric computer operating system
FSM	Finite state machine	GFC	Generic flow control (ATM)
FSS	File system switch (UNIX)	GGP	Gateway to gateway protocol
FTAM	File transfer, access and management	GI	Group identifier
FTP	File transfer protocol (Internet)	GII	Global information infrastructure
FTTB	Fibre to the building		
FTTC	Fibre to the curb		
FTTH	Fibre to the home		
FTTK	Fibre to the kerb		
Full	Full availability		
FUNI	ATM Frame user network interface		
FX	Foreign exchange service		
g	Reference point between WSF and human user		
G	Conductance		
G.703	ITU-T recommendation defining a four wire high speed digital physical layer interface		
G.704	ITU-T recommendation defining the 32 timeslot framing of a 2 Mbit/s signal		
G.732	ITU-T recommendation defining the multiframe synchronization of G.703 interfaces		
G0 state	Loss of power on the network side of the UNI (physical layer)		
G1 state	Operational state on the network side of the UNI (physical layer)		

GL	Group length	HMP	Host monitoring protocol
GLC	General log control	HOB	Head of bus
GMC	General management capability	HPR	High performance routing (IBM APPN)
GME	Global management entity	HS	Half session (IBM APPN)
GMSK	Gaussian minimum shift keying (radio modulation for GSM)	HSSI	High speed serial interface (EIA/TIA 612/613)
GNIM	Generic network information model	HTE	Hypergroup translating equipment
GNP	Gross national product	HU	High usage
GOS	Grade of service	Hz	Hertz
GOSIP	Government open systems interconnection profile	I.420	ITU-T recommendation defining the ISDN basic rate interface (also So,S/T, NET3, BRI, Euro-ISDN, etc.)
GPN	Generating polynomial (network)	I/O	Input/output
GPO	General Post Office (UK)	Ia	Interface point on the user terminal side of the SB-reference point at the UNI
GPO	Government Printing Office (USA)	IA2	International alphabet number 2 (Telex)
GPS	Global positioning system	IA5	International alphabet number 5 (the ASCII code)
GPU	Generating polynomial (user)	IAB	Internet activities board
GSM	Global System for Mobilecommunications	IAB	Internet architecture board
GTF	Generalized trace facility (IBM SNA)	IASG	Internet address sub-group
GTM	Global traffic meeting (INTELSAT)	Ib	Interface point on the B-NT2 side of the SB-reference point at the UNI
GUI	Graphical user interface	IBS	International business system (INTELSAT satellite service)
H	Header	IC	Interruption control
H.261	ITU-T recommendation defining a videoconference codec standard	ICD	International code designation (address)
H0	384 kbit/s transmission rate	ICF	Information conversion function
H11	1536 kbit/s transmission rate	ICI	Inter-carrier interface
H12	1920 kbit/s transmission rate	ICIP	Inter-carrier interface protocol
HCI	Host computer interface	ICIP_CLS	ICIP connectionless services
HCS	Header check sequence	ICMP	Internet control message protocol
HDB	Home data base (DECT)	ICP	Interconnect control program
HDB3	High density bipolar three-state (line code)	ICS	Implementation conformance statement
HDLC	High level datalink control (OSI layer 2)	ICV	Integrity check value
Hdr	Header	ID	Identity, identifier
HDSL	High bitrate digital subscriber line	ID	Interface data (AAL3/4 SAR-SDU)
HDTV	High definition television	IDI	Initial domain identifier (numbering)
HDX	Half-duplex	IDL	Interface definition language
header	The bits within a block or cell which provide for correct delivery of the payload	IDMIB	Inter-domain management information base
HEC	Header error control (ATM)	IDN	Integrated digital network
HF	High frequency (radio)	IDP	Initial domain part
HFC	Hybrid fibre coax	IDP	Inter-digit pause
HIPERLAN	High performance (wireless) local area network		
HLF	Higher layer function		
HLI	Higher layer information		
HLR	Home location register (GSM)		
HMA	Human machine adaptation		

IDR	Intermediate data rate (INTELSAT satellite service)	IPI	Initial protocol identifier (ISO/IEC TR 9577)
IDU	Interface data unit	IPL	Primary link for interactive services
IE	Information element	IPMS	Inter-personal message service (X.400)
IEC	Inter-exchange carrier (USA)	IPng	Internet protocol next generation
IEC	International Electrotechnical Commission	IPv4	Internet protocol version 4
IEE	Institution of Electrical Engineers (UK)	IPX	Internetwork packet exchange (Novell Netware)
IEEE	Institute of Electrical and Electronic Engineers (USA)	IRA	Internal reachable ATM address
IEN	Internet engineering note	IRN	Intermediate routing node
IETF	Internet engineering task force	IRP	Interior routing protocols (router networks)
IF	Intermediate frequency	IRP	Internal reference point
IGM	Internet group management	IRS	Intermediate reference system
IGP	Interior gateway protocol	IRTP	Internet reliable transaction
IHL	Internet header length	IS	Intermediate system
IISP	Interim interswitch signalling protocol (ATM NNI)	IS	International standard
ILEC	Independent local exchange carrier (USA)	ISC	International switching centre
ILMI	Interim local management interface (based on SNMP)	ISDN	Integrated services digital network
IMAC	Isochronous media access control	ISDU	Isochronous service data unit (DQDB)
IMCO	Inter-governmental maritime consultative organization	ISI	Inter-system interface (TETRA)
IMP	Internet message processor	ISI	In-span interconnection
IMPDU	Initial MAC (media access control) PDU (DQDB)	ISO	International Organization for Standardization
IMSI	International mobile station identity	ISO-IP	ISO internet protocol
IMSSI	Inter-MAN (metropolitan area network) switching system interface (DQDB)	ISO 2110	Standard 25-pin plug for V.24
IN	Intelligent network	ISO DE	ISO development environment (OSI-MDE) (a software product)
INAP	Intelligent network application part (SS7)	ISP	Intermediate service part (SS7)
IND	Indicate	ISP	International standardized profile
Ind	Indication	ISP	Internet service provider
Info	CPCS payload	ISP	Interswitch signalling protocol (ATM NNI)
INI	Inter-network interface	ISPBX	ISDN private branch exchange
INN	Intermediate network node	ISR	Intermediate session routing (IBM APPN)
INTELSAT	International Telecommunications Satellite organization	ISSI	Inter-switching system interface
IOC	Input/output control	ISUP	Integrated services digital network user part (SS7)
IOCP	Input/output channel program (IBM ESCON)	IT	Information technology
IOP	Interoperability testing	IT	Information type
IP	Internet protocol	ITC	Independent telephone company
IP	Intelligent peripheral (IN)	ITT	Invitation to tender
IPCP	Internet protocol control protocol	ITU	International Telecommunications Union
		ITU-R	International Telecommunications Union Radiocommunication sector (former CCIR)

ITU-T	International Telecommunications Union Standardization sector (former CCITT)	LDDB	Local directory database
IUT	Implementation under test	LE	LAN emulation
IVD	Integrated voice data	LE	Layer entity
IWF	Interworking function	LE	Local exchange
IWP	Interworking point	LE_ARP	LAN emulation address resolution protocol
IWU	Interworking unit	LEC	Local exchange carrier (USA)
IXC	Interexchange carrier	LEC	LAN emulation client
J	Justification or justification bit	LECID	LAN emulation client identifier
J-bit	Justification bit (of line code)	LECS	LAN emulation configuration server
JIT	Just in time	LED	Light-emitting diode
K-bit	Justification bit (of line code)	LEN	Low entry networking
kbit/s	Kilobits per second	LF	Largest framesize
Kbps	Kilobits per second	LFA	Loss of frame alignment
KLM	Kernel lock manage	LFC	Local function capabilities
Ku-Band	Satellite transmission	LFSID	Local form session identifier (IBM APPN logical channel identifier)
Kx	Inter-network reference point; the point between a B-ISDN and some other sort of transit network	LGN	Logical group node
L	Inductance	LH	Link header
LA	Link acknowledgement	LI	Link identifier
LAN	Local area network	LI	Length indicator
LANE	LAN emulation	LIC	Line interface coupler
LAP	Link access procedure	LIDB	Line information database
LAPB	Link access procedure balanced (X.25)	LIFO	Last in first out (buffer)
LAPD	ISDN link access procedure (D-channel)	LIVT	Link integrity verification test
LAPM	Link access procedure modems	LLA	Logical layered architecture
LAPV5-DL	Data link sublayer of the link access protocol of V5.X interface	LLAP	LocalTalk link access protocol (Appletalk layer 2)
LAPV5-EF	Envelope function sublayer of link access protocol of V5.X interface	LLC	Lower layer compatibility (ISDN)
LASER	Light amplification by stimulated emission of radiation	LLC	Logical link control (token ring and ethernet LANs)
LATA	Local access and transport area (USA)	LLI	Lower layer information
LC-node	Last common node	LME	Layer management entity
LCGN	Logical channel group number	LMI	Local management interface
LCN	Logical channel number (statistical multiplexing)	LN	Link attention
LCP	Link control protocol	LNA	Link attention acknowledgement
LCT	Last compliance time or last conformance time (leaky bucket algorithm)	LO	Local oscillator
LD	LAN destination	LOC	Loss of cell delineation
LD	Language digit	LOF	Loss of frame
LD	Loop disconnect (signalling)	LOM	Loss of OAM cell
LD-CELP	Low delay code excited linear prediction	LOP	Loss of pointer
		LOS	Line of sight
		LPDU	Logical link control protocol data unit
		LR	Link request
		LR	Loudness rating
		LRC	Longitudinal redundancy check
		LRE	Low rate encoding
		LS	Line station (TETRA)
		LSAP	Link service access point
		LSB	Least significant bit

LSB	Lower sideband	MA	Medium Adaptor
LT	Line termination	MAC	Media access control
LT	Link transfer	MAF	management application function (software communication function within TMN functional block)
LTE	Line terminating equipment		
LTH	Length field	MAHO	Mobile-assisted handover (GSM)
LTU	Line terminating unit	MAN	Metropolitan area network
LU	Logical unit (IBM SNA)	MAP	Manufacturing automation protocol
LUNI	LAN-user-network-interface	MAP	Mobile application part (SS7)
M	Manager	MAU	Media access unit (Token ring LAN)
M	Mandatory	MaxCR	Maximum cell rate
M	More (AAL3/4 SAR-SDU)	Mb/s	Megabits per second
M-ACTION	(CMIP) request an object to perform an action	Mbit/s	Megabits per second
M-CANCEL	(CMIP) cancels previous M-GET command	Mbps	Megabits per second
-GET		MBS	Maximum burst size
M-CREATE	(CMIP) creates objects	MBS	Monitoring block size (PL-OAM cell)
M-DELETE	(CMIP) deletes objects	MCD	Maintenance cell description
M-EVENT	(CMIP) allows a network resource to announce the occurrence of an event	MCDV	Maximum cell delay variation
-REPORT		MCF	MAC convergence function (DQDB)
M-GET	(CMIP) read value of an attribute	MCF	Management communication function (software communication function within TMN functional block)
M-SET	(CMIP) add, remove or replace command	MCF	Message communication function
M reference point	Interface point between a B-ISDN and a CLSF (control function) of an external specialized service provider	MCHO	Mobile-controlled handover
M.1020	ITU-T recommendation defining a high quality analogue leaseline of 4 kHz	(DECT)	
		MCI	Microwave Communications Inc. (USA)
M.1025	ITU-T recommendation defining a medium quality analogue leaseline of 4 kHz	MCL	Mercury Communications Limited (UK)
		MCLR	Maximum cell loss ratio
M.1040	ITU-T recommendation defining a low/medium quality analogue leaseline of 4 kHz	MCP	MAC convergence protocol (DQDB)
		MCR	Maximum cell rate
M1-interface	Interface between ATM end device and private ATM network management system (ATM forum)	MCSN	Monitoring cell sequence number (ATM OAM performance management cell)
M2-interface	Interface between Private ATM network and its management system (ATM forum)	MCTD	Maximum cell transfer delay
		MD	Mediation device
M3-interface	Interface offered from public ATM network management system to enable customer network management (ATM forum)	MDE	Management development environment
		MDTRS	Mobile digital trunk radio system
M4-interface	Interface between Public ATM network and its management system (ATM forum)	MDU	Management data unit
		MF	Management function (TMN)
M5-interface	Interface between management systems of different Public ATM networks (ATM forum)	MF	Mapping function
		MF	Mediation function
		MF	More fragments (IP)
		MF	Multi-frequency (signalling)
		MF4	Alternative identification for DTMF signalling

MFA	Management functional areas	MS	Multiplex section (SDH)
MFC	Multi-frequency code (signalling)	MSAP	MAC Service access point (SMDS)
MFJ	Modified final judgement	MSB	Most significant bit
MFS	Metropolitan fibre systems	MSC	Mobile switching centre
MH	Message handling	MSDU	MAC service data unit (SMDS)
MHS	Message handling system (X.400)	MSF	Management systems framework
MIB	Management information base	MSISDN	Mobile station ISDN number
MIC	Media interface connector (jack, plug or socket)	MSL	Maximum segment lifetime (IP)
MID	Message identifier (DQDB)	MSN	Microsoft network
MID	Multiplexing identification (ATM)	MSN	Multiple systems networking
MIF	Management information file	MSOH	Multiplex section overhead (SDH)
MIL-STD	Military standard	MSP	Management service provider (TMN)
MIM	Management information model	MSRN	Mobile station roaming number
MIS	Management information system	MSS	MAN switching system (SMDS)
MIS	Mobile interface server	MST	Multiplex section termination
ML	Maximum length	MSVC	Meta signalling virtual channel
MLID	Multiple link interface driver (Novell Netware)	MTA	Message transfer agent (X.400)
MLP	Multilink procedure	MTAE	Message transfer agent entity (X.400)
MLT	Multilevel transit (line code)	MTBF	Mean time between failures
MM	Message mode (AAL3/4 or AAL5)	MTL	Message transfer layer (X.400)
MM	Multimode fibre	MTP	Message transfer part (SS7)
MMIC	Microwave monolithic integrated circuit	MTP1	Message transfer part level 1 (signalling system number 7)
MNP	Microcom networking protocol	MTP2	Message transfer part level 2 (signalling system number 7)
MNT	Management	MTP3	Message transfer part level 3 (signalling system number 7)
MO	Managed objects	MTS	Message transfer service (X.400)
MOCS	Management object conformance statement	MTTR	Mean time to repair
MOTIS	Message-oriented text interchange system (ISO/IEC 10021)	MTU	Message transfer unit
MOU	Memorandum of understanding	MTU	Maximum transfer unit (IP)
MP	Measurement point	Mu-law	PCM speech encoding (North American and Japanese standard)
MPEG	Motion pictures experts group	MUF	Maximum usable frequency
MPH	Management physical header (physical layer primitive)	MUX	Multiplexor
MPH-AI	MPH activate indication	MV	Major vector identifier (Token ring MAC)
MPH-CI <sub>n</sub>	MPH correction indication with parameter <i>n</i>	MVL	Major vector length (Token ring MAC)
MPH-DI	MPH deactivate indication	MVS	Multiple virtual storage (IBM operating system)
MPH-EI <sub>n</sub>	MPH error indication with parameter <i>n</i>	N/A	Not applicable
MPOA	Multiprotocol over ATM	N-ISDN	Narrowband ISDN
MPP	Multi-protocol package (Appletalk)	N0	(Call state at the user side of the UNI interface) null; no call exists
MPT	Ministero delle Poste e Telecomunicazioni (Italy)	N1	(Call state at the network side of the UNI interface) call initiated
MRS	Mobile relay station (DECT)	N10	(Call state at the network side of the UNI interface) active
MRVT	MTP routing verification test (SS7)		
MS	Management services		
MS	Mobile station (TETRA)		

N11	(Call state at the network side of the UNI interface) release request	NHRP	Next hop resolution protocol
N12	(Call state at the network side of the UNI interface) release indication	NIC	Network interface card (Token ring LAN)
N22	(Call state at the network side of the UNI interface) call abort	NIC	Number of included cells (PL-OAM cell)
N3	(Call state at the network side of the UNI interface) outgoing call proceeding	NIR	NEXT (near end crosstalk) loss-to-insertion ratio
N4	(Call state at the network side of the UNI interface) call delivered	NIST	National Institute for Science and Technology
N6	(Call state at the network side of the UNI interface) call present	NIU	Network interface unit
N7	(Call state at the network side of the UNI interface) call received	NLM	Network lock manager (Novel Network)
N8	(Call state at the network side of the UNI interface) connect request	NLPID	Network layer protocol identifier (ISO/IEC TR 9577)
N9	(Call state at the network side of the UNI interface) incoming call proceeding	NM	Network management
NA	Network adaptor (an interworking function between B-ISDN and narrowband ISDN)	NMB-EB	Number of monitored blocks, far end block errors (physical layer)
NADP	North American dial plan	NMB-EDC	Number of monitored blocks, for which error detection codes are included (physical layer)
NAE	Network address extension	NMC	Network management centre
NAMTS	Nippon automatic mobile telephone system	NMF	Network management forum
NAU	Network addressable unit	NML	Network management layer (TMN management model)
NBP	Name binding protocol (Appletalk)	NMS	Network management system
NCO	Numerically controlled oscillator	NMT	Nordic mobile telephone system (cellular radio)
NCOP	Network code of practice	NN	National number
NCP	Network control program (IBM SNA)	NN	Network node (IBM APPN)
NCP	Network control protocol (Novell Netware)	NNI	Network node interface
NCS	Network computing system (UNIX)	NOF	Normal operational frames
NDF	New data flag	non-P-format	A convergence sublayer format type used by AAL1
NDIS	Network driver interface specification	NOSFER	Nouveau systeme fondamental pour la determination des equivalents de reference
NE	Network element	Novell	A popular LAN operating system software
NEF	Network element function	Netware	
NEL	Network element layer (TMN management model)	NP	Network performance
NEM	Network element manager	NPA	Number plan of America
NET	Norme europeene de telecommunication	NPC	Network parameter control
NetBIOS	Network basic input/output system (IBM LAN operating system)	NPDU	Network protocol data unit
NEXT	Near end crosstalk	NPMA	Non pre-emptive priority multiple access
NFD	Net filter discrimination	NPSI	Network packet switching interface (IBM X.25)
NFS	Network file server (TCP/IP)	NPV	Net present value
		NRM	Network resource management
		NRZ	Non return-to-zero (line code)
		NRZI	Non return-to-zero inverted (line code)

NSAP	Network service access point	OMC	Operations and maintenance centre
NSDU	Network service data unit	OMG	Object management group
NSF	Network search function (IBM APPN)	ONP	Open network provision (European Union)
NSN	National significant number (E.164)	OO	Object-oriented
NSP	Network services protocol (DECnet)	OOF	Out of frame
NSR	Non-source routed	OOK	On/off keying
NSS	Network and switching sub-system	OPAL	Optical access line
NT	Network termination	OPC	Originating point code
NT1	Network terminating equipment type 1 (ISDN)	ORB	Object request broker
NT2	Network terminating equipment type 2 (ISDN)	ORE	Overall reference equivalent
NTE	Network terminating equipment	ORS	Optical repeater station
NTN	National terminal number (X.121)	OS	Operating system
NUI	Network user identification	OSF	Open software foundation
NVP	Network voice protocol (TCP/IP)	OSF	Operations systems function
NVT	Network virtual terminal (UNIX)	OSH	Optical splitter head
O	Optional	OSI	Open systems interconnection
O/R	Originator/recipient	OSIE	OSI environment
OAF	Origination address field	OSIRM	OSI reference model
OAM	Operation and maintenance	OSPF	Open shortest path first (Internet)
OAMP	Operations, administration, maintenance and provisioning	OSS	Operations and support sub-system
OAMC	Operations, administration and maintenance centre (TMN)	OUI	Organizationally unique identifier
OC-1	Optical carrier (SONET) OC-1 has bitrate of 51.84 Mbit/s	OWF	Optimum working frequency
OC- <i>n</i>	Optical carrier level- <i>n</i> (SONET)	P	Performance
OCC	Other common carrier	P/F	Poll final
OCD	Out of cell delineation (an anomaly causing LOC)	P-format	A convergence sublayer format type used by AAL1
OCS	Object conformance statement	P-NNI	Private network node interface (ATM)
octet	Data volume equal to 1 byte (8 bits)	p-type data	ISDN D-channel data of packet type (i.e. SAPI = 16)
ODA	Office document architecture	P reference point	Conceptual point within a B-ISDN; the interface between the ATM switching network and a connectionless service function
ODI	Open datalink interface (Novell Netware)	P1-interface	X.400 interface for relaying between post offices
ODP	Open distributed processing	P3-interface	X.400 interface between client and post office (submission/delivery)
ODP	Originator detection pattern	PA	Prearbitrated segment (DQDB)
off-hook signal	Signal to indicate readiness of telephone to send dialled digits	PABX	Private automatic branch exchange (nowadays synonymous with PBX)
OFN	Optical fibre node	PACE	Paging access control equipment
OfTel	Office of Telecommunications (UK)	packet	An information block identified by a label at layer 3 of the OSI model (e.g. X.25 packet)
OH	Overhead	PAD	Packet assembler/disassembler
OID	Object identifier	PAD	Padding
OLR	Overall loudness rating	PAF	Prearbitrated function (DQDB)
OLRT	Online real-time	PAM	Pulse amplitude modulation
OLTU	Optical line terminating unit	PAMA	Pre-assigned multiple access
OM	Object management		
OMAP	Operations and maintenance part (SS7)		



PANS	Pretty amazing new stuff	PL-OU	Physical layer overhead unit
PBF	Proportionate bidding facility	PLCP	Physical layer convergence protocol
PBX	Private branch exchange (an office telephone system)	PLK	Primary link
PC	Personal computer	PLL	Phased-locked loop
PC	Printed circuit	PLMTS	Public land mobile telephone system
PC	Priority control	PLP	Packet layer protocol
PCA	Percentage of calls answered	PLS	Primary link station
PCA- <i>n</i>	Percentage of calls answered in <i>n</i> seconds	PLSP	PNNI link state packet
PCI	Programming communication interface (like an API)	PLU	Primary logical unit (IBM SNA)
PCI	Protocol control information	PM	Performance management
PCM	Pulse code modulation	PM	Physical medium (sublayer)
PCN	Personal communications network (mobile telephone system)	PMD	Physical layer medium dependent
PCR	Peak cell rate	PMP	Point-to-multipoint (radio system)
PCS	Personal communications system (mobile telephone system)	PMP-node	Point-to-multipoint node
PDD	Post dialling delay	PMR	Private mobile radio
PDH	Plesiochronous digital hierarchy (transmission technique)	PMTR	Private mobile trunk radio
PDO	Packet data only	PMUX	Primary multiplexor
PDU	Protocol data unit	PN	Pseudo-noise
PF	Presentation function	PNNI	Private network-node interface, or Private network-network interface (ATM)
PG	Peer group	POCSAG	Post office code standardization advisory group (radiopaging code)
PGL	Peer group leader	POH	Path overhead (SDH)
PH	Packet handler	POI	Path overhead identifier (DQDB)
PH	Primitive used by layer 2 to control the physical layer	POI	Point of interconnection
Ph-SAP	Physical layer SAP (DQDB)	PolSK	Polarity shift keying
PHP	PNNI hello packet	PON	Passive optical network
PHTE	PNNI horizontal topology element	POP	Point of presence
PHTP	PNNI horizontal topology packet	POS	Point of sale (system)
PHY	Physical layer or physical layer protocol	POSIX	Portable operating system interface (UNIX)
PI	Parameter length	POTS	Plain old telephone service
PICS	Protocol implementation conformance statement	PP	Portable part (DECT)
PID	Protocol identifier	PPDU	Presentation protocol data unit (OSI layer 6)
PIN	Personal identification number	PPM	Pulse position modulation
PING	Packet internet groper (Internet)	PPP	Point-to-point protocol (Internet)
PIU	Path information unit (IBM SNA)	PRA	Primary rate access (ISDN)
PIXIT	Protocol implementation eXtra information for testing	PRBS	Pseudo random binary sequence
PKCS	Public key cryptosystems	PRDMD	Private directory management domain
PL	Pad length (DQDB)	PRI	Primary rate interface (ISDN)
PL	Permanent line	Private NNI	ATM NNI interface designed for use between private ATM networks
PL	Physical layer	Private UNI	ATM interface between private ATM switch (B-NT2) and ATM user (B-TE), i.e. the R or SB (UNI)
PL-OAM	Physical layer operation and maintenance cell	PRM	Protocol reference model
		PRMA	Packet reservation multiple access

PRMD	Private Management Domain (X400)	QPSK	Quarternary phase shift keying
PRR	Pulse repetition rate	QPSX	Queued packet and synchronous switch
PRS	Primary reference source (timing clock)	Qx	Proprietary interface between mediation function and network element (TMN)
PRT	Pulse repetition time	R	Requirement
PrtY	Parity	R	Reserved filed (PL-OAM cell)
PSAP	Physical service access point (OSI)	R	Resistance
PSDN	Public switched data network	Reference point	Reference point at the ATM UNI between a B-TA and a B-TE2 (i.e. an X- or V-series interface)
PSE	Packet-switched exchange	R & D	Research and development
PSH	Push	R1	Regional signalling system number 1 (ITU-T or CCITT)
PSI	Packet-switched interface	R2	Regional signalling system number 2 (ITU-T or CCITT)
PSK	Phase shift keying	R2D	Digital version of CCITT R2 signalling
PSN	PL-OAM sequence number (PL-OAM cell)	RAI	Remote alarm indication
PSTN	Public switched telephone network	RAM	Random access memory
PT	Payload type	RARP	Reverse address resolution protocol (Internet)
PT	Portable radio termination (DECT)	RBB	Residential broadband
PTE	Path terminating equipment	RBOC	Regional bell operating company (USA)
PTI	Payload type identifier	RC	Routing control
PTM	path trace management	RD	Route destination (accounting)
PTO	Public telecommunications operator	RD	Routing descriptor
PTP	Point-to-point (radio system)	RD	Routing domain (numbering)
PTR	Pointer	RDA	Remote database access
PTS	Proceed to send (signal)	RDI	Remote defect indication (formerly called FERF, far end receive failure)
PTSE	PNNI topology state element	RDM	Radio distribution module
PTSP	PNNI topology state packet	RDN	Relative distinguished name
PTT	Post, telegraph and telephone (company)	RDP	Reliable data protocol
PU	Physical unit (IBM SNA)	RDSS	Radiodetermination satellite system
Public UNI	ATM interface between a private ATM device or switch and a carrier network (UNI TB or UB)	RE	Reference equivalent
PV	Parameter value	Ref	Reference
PVC	Permanent virtual circuit	REJ	Reject
PVCC	Permanent virtual channel connection	Req	Request
PVPC	Permanent virtual path connection	REQ	Request
PWM	Pulse width modulation	RES	Reserved
Q-SIG	Q-signalling (inter-PBX ISDN signalling system)	RESP	Response
Q3	Interface at q3 reference point (TMN)	Rest 0	(Call state at the UNI) null, no transaction exists
QA	Q adaptor	Rest 1	(Call state at the UNI) restart request
QA	Quality assurance	Rest 2	(Call state at the UNI) restart request
QA	Queued arbitrated (DQDB)	REX	Remote execution service (UNIX)
QAF	Q-adaptor function (TMN)	RF	Radio frequency
QAF	Queued arbitrated function (DQDB)		
QAM	Quadrature amplitude modulation		
qdu	Quantization distortion unit		
QOS	Quality of service		

RFC	Request for comment (IETF specification document)	Rs	Sustainable cell rate
RFI	Radio frequency interference	RS-232	EIA interface RS-232 (equivalent V.24)
RFP	Radio fixed part (DECT)	RS-449	EIA interface RS-449 (V.35/V.36)
RFP	Request for proposal	RS	Repeater station
RFQ	Request for quotation	RS	Regenerator section
RFS	Ready for service	RSE	Real system environment
RFS	Remote file sharing (UNIX)	RSL	Receive signal level (radio)
RFT	Request for technology	RSOH	Regenerator section overhead (SDH)
RG	Regenerator	Rsp	Response
RI	Ring in	RSRVD	Reserved
RI	Routing information	RSS	Route selection service
RIC	Radio identity code	RSSI	Received signal strength indicator (DECT)
RII	Routing information indicator	RST	Reset
RILL	Radio in the local loop	RSU	Remote switching unit
Ringdown	Basic signalling form used only to ring a telephone at the end of a fixed link	RSVD	Reserved
RIP	Routing information protocol (Internet)	RT	Routing type
RIP	Router information protocol (Novell Netware)	RTO	Retransmission timeout
RISC	Reduced instruction set computer	RTS	Residual time stamp
RJE	Remote job entry	RTSE	Reliable transfer service element (OSI layer 7)
RL	Return loss	RTT	Round trip time
RM	Reference model	RU	Remote unit
RM	Resource management	RW	Read write
RM	Routine maintenance	Rx	Receive
RNR	Receiver not ready	RZ	Return-to-zero (line code)
RNU	Radio network unit	S & F	Store and forward
RO	Read only	S/I	Signal to interference ratio
RO	Ring out	S/N	Signal to noise ratio
ROA	Recognized operating agency	S-AIS	Section alarm indication signal (PL-OAM cell)
ROM	Read only memory	S-FEBE	Section far end block error (PL-OAM cell)
Root	Functional block which initiates a point-to-multipoint call	S-FERF	Section far end received failure (PL-OAM cell)
ROSE	Remote operations service element (OSI layer 7)	S (with suffix VCI or VPI)	Switching function for a virtual connection or virtual path
Router	Device used to interconnect LANs	SA	Source MAC address
RPC	Remote procedure call (Internet)	SAA	Services and applications
RPOA	Recognized private operating agency	SAA	Systems applications architecture (IBM architecture)
RPS	Radio protection switching	SAAL	Signalling ATM adaptation layer
RQ	Request counter (DQDB)	SAB	Subnetwork access boundary
RR	Receiver ready	SABM	Set asynchronous balanced mode
RR	Resource record	SABME	Set asynchronous balanced mode
RRE	Receiving reference equivalent	SACK	Selective acknowledgement (TCP/IP)
RRR	Radio relay regenerator	SAP	Service access point
RRT	Radio relay terminal	SAP	Service advertising protocol (Novell Netware)
RS	Reception status (AAL3/4 SAR-SDU)		
RS	Regenerator section (transmission)		

SAPI	Service access point identifier	SG	Study group
SAR	Security alarm reporting	SIB	Service independent building block (IN)
SAR	Segmentation and reassembly sublayer (ATM)	SIBB	Service independent building block (IN)
SAR-PDU	Segmentation and reassembly layer protocol data unit	SIG	SMDS interest group
SAS	Single attached station	SIM	System identification module
SAT	Security audit trail	SIP	SMDS subscriber interface protocol (SMDS)
SATF	Shared access transport facility	SIP-2	SMDS subscriber interface protocol, level 2
S <sub>B</sub>	Reference point at the user-network interface (UNI) between terminal equipment (B-TE) and B-NT2	SIP-3	SMDS subscriber interface protocol, level 3
SBS	Selective broadcast signalling	SIPP	Simple Internet protocol plus
SBSVC	Selective broadcast signalling virtual channel	SIR	Sustained information rate (SMDS)
SC	Optical fibre connector type	SLE	Sublayer entity
SC	Session connector (IBM APPN)	SLIP	Serial line interface protocol (Internet)
SCCP	Signalling connection and control part (SS7)	SLP	Single link procedure
SCE	Service creation environment (IN)	SM	Scheduling management
SCE	System control element (IBM SNA)	SM	Security management
SCF	Service control function (IN)	SM	Session manager
SCF	Synchronization and coordination (or control) function	SM	Single mode (monomode) fibre
SCP	Service control point (IN)	SM	Streaming mode (AAL3/4 or AAL5)
SCPC	Single channel per carrier (satellite system)	SMAE	System management application entity
SCR	Sustainable cell rate	SMASE	System management application service element
SD	Starting delimiter (IBM token ring LAN)	SMB	Server message block
SDDI	Shielded distributed data interface	SMDS	Switched multimegabit data service
SDE	Submission and delivery entity (X.400)	SME	Society of Manufacturing Engineering (USA)
SDF	Service data function	SMF	Single mode fibre (monomode fibre)
SDH	Synchronous digital hierarchy	SMF	System management function
SDL	Specification and description language	SMK	Shared management knowledge
SDLC	Synchronous data link control (IBM SNA)	SML	Service management layer (TMN management model)
SDN	Software defined network	SMP	Service management point
SDT	Structured data transfer (AAL1)	SMS	Service management system (IN)
SDU	Service data unit	SMT	Station management (FDDI)
SECB	Severely errored cell block	SMTP	Simple mail transfer protocol
SECBR	Severely errored cell block ratio	SN	Sequence number
SEL	Selector (address)	SN	Subarea node (IBM)
SES	Severely errored seconds	SNA	System network architecture
SFET	Synchronous frequency encoding technique	SNAP	Subnetwork address plan (IEEE 802.1)
SFH	Slow frequency hopping	SNAP	Subnetwork address protocol (Internet)
SFT	System fault tolerance (Novell Netware)	SNAP	System network architecture protocol (IBM)

SNI	SNA network interconnection (IBM SNA)	SSCF	Service specific coordination function (ATM)
SNI	Subscriber network interface	SSCOP	Service specific connection-oriented protocol (ATM)
SNMP	Simple network management protocol	SSCP	System services control point (IBM SNA)
SNP	Sequence number protection (AAL1)	SSCS	Service specific convergence sublayer (ATM)
SNPA	Sub-network point of attachment	SSF	Service switching function (IN)
SOC	Start of cell	SSI	Service specific information
SOH	Section overhead (SDH)	SSM	Single segment message
SONET	Synchronous optical network	SSP	Service switching point (IN)
SP	Service provider	ST	Optical fibre connector type
SP	Signalling point	ST	Segment type (AAL3/4 SAR)
SP	Space	ST	Signalling terminal
SP-node	Single party node	STACK	Start acknowledgement message (DECnet)
SPAG	Standards promotion and applications group	STDL	Structured transaction definition language
SPC	Signalling point code	STE	Section terminating equipment (SONET)
SPC	Stored program control	STE	Signalling terminal
SPE	Synchronous payload envelope (SONET)	STE	Spanning tree explorer
SPF	Shortest path first protocol (Internet)	STE	Supergroup translating equipment
SPID	Service profile identifier	STF	Statens Teleforvaltning (Norway)
SPIRIT	Service providers' integrated requirements for information technology	STM	Station management
SPL	Service provider link	STM	Synchronous transfer mode
SPM	FDDI-to-SONET physical layer mapping standard (FDDI)	STM-1	Synchronous transport module of SDH with bitrate of 155 Mbit/s
SPN	Subscriber premises network	STM-16	Synchronous transport module of SDH with bitrate of 2.5 Gbit/s
SPX	Sequenced packet exchange (Novell Netware)	STM-4	Synchronous transport module of SDH with bitrate of 622 Mbit/s
SR	Source routing	STM-64	Synchronous transport module of SDH with bitrate of 10 Gbit/s
SRB	Source route bridging	STM- <i>n</i>	Synchronous transport module- <i>n</i>
SRE	Sending reference equivalent	STP	Shielded twisted pair (cable)
SREJ	Select reject frame	STP	Signalling transfer point
SRF	Specifically routed frame	STP	Spanning tree protocol
SRL	Structural return loss	STS	Synchronous transport system (SONET)
SRL	Spectrum reference level	STS-1	Synchronous transport system (SONET) with bitrate of 51.84 Mbit/s
SRT	Source routing transparent	STS-3	Synchronous transport system (SONET) with bitrate of 155.52 Mbit/s
SRT	Subscriber radio terminal	STS- <i>n</i>	Synchronous transport signal level- <i>n</i> (SONET)
SRTS	Synchronous residual time stamp (AAL1)	STX	Start of text
SRTT	Smoothed round trip time	SUT	System under test
SRVT	SCCP routing verification test (SS7)	SVC	Signalling virtual channel
SS	Service specific		
SS	Start/stop transmission		
SS6	Signalling system number 6		
SS7	Signalling system number 7		
SSAP	Source service access point		
SSB	Single sideband		

SVC	Switched virtual channel	TETRA	Trans-European trunk radio
SVC	Switched virtual circuit	TFTP	Trivial file transfer protocol (Internet)
SVI	Subvector identifier (Token ring MAC)	TG	Transmission group (IBM)
SVL	Subvector length (Token ring MAC)	TH	Transmission header (IBM SNA)
SVP	Subvector parameters (Token ring MAC)	THT	Token holding time
SWG	Sub working group	TIB	Task information base
SwMI	Switching and management infrastructure (TETRA)	TIC	Token ring interface coupler
SYN	Synchronization pattern	TIG	Topology information group
T	Trailer	TINA	Telecommunications intelligent network architecture
T1	North American TDM hierarchy (1.5 Mbit/s)	TKG	Telekommunikationsgesetz (German telecom law)
T3	North American TDM hierarchy (45 Mbit/s)	TLAP	Token Talk link access protocol (Appletalk layer 2 for token ring LAN)
TA	Technical advisories	TLI	Transport level interface (Internet)
TA	Terminal adaptor	Tlr	Trailer
TACS	Total access communication system (mobile telephone)	TLV	Type/length/value
TAR	Temporary alternative re-routing	TMN	Telecommunications management network
TAT	Theoretical arrival time (leaky bucket algorithm)	TMR	Trunk mobile radio
TAT	Transatlantic telephone cable	TMSI	Temporary mobile station identity
T <sub>B</sub>	Reference point at the user-network interface (UNI) on the user side of the B-NT1	TMUX	Transmultiplexor
TB	Transparent bridging	TNIC	Telex network identification code (F.69)
T <sub>c</sub>	Committed burst duration (frame relay)	TNS	Transit network selection
TC	Transaction capabilities (SS7)	TOP	Technical and office protocol
TC	Transmission convergence sublayer (ATM)	TP	Termination point
TCAM	Telecommunication access method (IBM)	TP	Throughput
TCAP	Transaction capabilities application part (SS7)	TP	Transaction processing
TCE	Transmission connection element	TP	Transmission path
TCF	Transparent computing facility	TP	Twisted pair (cable)
TCP	Transmission control protocol (Internet)	TP	Termination point
TCP/IP	Transmission control protocol/Internet protocol	TP-AIS	Transmission path alarm indication signal (PL-OAM cell)
TCRF	Transit connection related function	TP-FEBE	Transmission path far end block error (PL-OAM cell)
TDC	Telex destination code (F.69)	TP-FERF	Transmission path far end received failure (PL-OAM cell)
TDD	Time division duplex (radio)	TPE	Transmission path endpoint
TDM	Time division multiplexing	TPI	Transport provider interface (Internet)
TDMA	Time division multiple access (radio)	TPM	Topology database manager
TE	Terminal equipment	TPON	Telephony passive optical network
TEM	Testing management	TR	Technical references
		TR	Technical report
		TR	Token ring (LAN)
		TR	Trouble resolution
		Trader	A sophisticated directory service agent

TRCC	Total received cell count (ATM OAM performance cell)	UDP	User datagram protocol (Internet)
TRF	Tuned radio frequency receiver	UDT	Unstructured data transferred
TS	Time stamp (ATM OAM performance management cell)	UHF	Ultra high frequency (radio)
Ts	Equal to 1/Rs, the inverse of the sustainable cell rate	UME	UNI management entity
$\tau_s$	Burst tolerance	UMTS	Universal mobile telephone service
TS16	Timeslot 16	unbalanced	A type of data transmission using coaxial cable (usually 75 Ohm)
TSAG	Telecommunication standardization advisory group	UNI	User-network interface
TSE	Telecommunications services environment	UNITDI	United Nations trade data interchange
TSF	Trail signal fail	UNRP	User-to-network reference points
TSI	Time slot interchange	UPC	Usage parameter control
TSO	Tine sharing option (IBM operating system)	UPS	Uninterruptible power supply
TSSI	Time slot sequence integrity	UPT	Universal personal telecommunication
TST	Time space time (digital switch matrix)	URG	Urgent
TSTP	Timestamp (ATM OAN performance cell)	USART	Universal synchronous/asynchronous receiver/transmitter
TT	Trouble ticketing	USB	Upper sideband
TTC	Telecommunication Technology Committee (Japan)	UTC	Coordinated universal time
TTL	Time to live (Internet)	UTOPIA	Universal test and operations physical interface for ATM (ATM specification for standard ATM/PHY layer interface)
TTP	Trail termination point	UTP	Unshielded twisted pair (cable)
TTR	Time to repair	UTP3	Unshielded twisted pair cable, category 3
TTRT	Target token rotation time	UTP5	Unshielded twisted pair cable, category 5
TTY	Teletype	UUS	User-to-user signalling
TU- <i>n</i>	Tributary unit of order <i>n</i> (SDH)	V	Violation
TUC	Total user cell number (ATM OAM performance management cell)	V.10	ITU-T recommendation for physical layer DTE/DCE data interface
TUG	Tributary unit group (SDH)	V.11	ITU-T recommendation for physical layer DTE/DCE data interface
TUP	Telephone user part (SS7)	V.110	ITU-T recommendation for sub-rate data framing (i.e. deriving rates below 64 kbit/s)
TWT	Travelling wave tube	V.22bis	ITU-T recommendation for inter-modem modulation (2400 bit/s)
Tx	Transmit	V.24	ITU-T recommendation V.24 (DTE/DCE interface up to about 28.8 kbit/s)
U	User option	V.25 bis	ITU-T recommendation defining modem control technique
U-interface	Two wires only ISDN BRI interface	V.32	ITU-T recommendation for inter-modem modulation up to 9600 bit/s
UA	Unnumbered acknowledgement	V.34	ITU-T recommendation for inter-modem modulation up to 28 800 bit/s
UA	User agent		
UAE	User agent entity (X.400)		
UAL	User agent layer (X.400)		
UART	Universal asynchronous receiver/transmitter		
UAS	Unavailable seconds		
U <sub>B</sub>	Reference point at the UNI on the network side of B-NT1		
UBR	Unspecified bitrate (ATM)		
UDF	User-defined		

V.35	ITU-T recommendation for physical layer DTE/DCE data interface (usually rates above 64 kbit/s in USA)	VC CPF	Virtual connection connecting point functions
V.36	ITU-T recommendation V.36 (DTE/DCE interface for digital lines used in USA)	VCC	Virtual channel connection (ATM)
V.42	ITU-T recommendation defining data compression and error correction	VCCE	Virtual channel connection endpoint
V.54	ITU-T recommendation defining control technique for applying loopbacks to remote modems	VCI	Virtual channel identifier (ATM)
V + D	Voice and data	VCL	Virtual channel link
V5.1	Interface between access node (AN) and local exchange (LE)	VCO	Voltage-controlled oscillator
V5.2	Interface between access node (AN) and telephone local exchange (LE) (up to $16 \times 2$ Mbit/s with concentrator function)	VDB	Visitor data base (DECT)
VANS	Value added network services	VF	Variance factor
VAS	Value added service	VF	Voice frequency
VASP	Value added service provider	VFS	Virtual file system
VAX	Virtual Address eXtended (Digital equipment corporation)	VHF	Very high frequency (radio)
VBR	Variable bitrate (ATM)	VIF	Visual index file
VBR-NRT	Variable bitrate (non real time)	VLR	Visitor location register (GSM)
VBR-RT	Variable bitrate (real time)	VM	Virtual machine (IBM)
VC	Virtual call	vocoding	Low bit rate voice compression
VC	Virtual channel	VoD	Video on demand
VC	Virtual circuit	VP	Virtual path
VC	Virtual connection	VP-XC	Virtual path crossconnect
VC	Virtual container	VP CEPF	Virtual path connection end point functions
VC-11	Virtual container of SDH with 1544 kbit/s transmission rate	VP CPF	Virtual path connecting point functions
VC-12	Virtual container of SDH with 2048 kbit/s transmission rate	VPC	Virtual path connection (ATM)
VC-21	Virtual container of SDH with 6312 kbit/s transmission rate	VPCE	Virtual path connection endpoint
VC-22	Virtual container of SDH with 8448 kbit/s transmission rate	VPI	Virtual path identifier (ATM)
VC-31	Virtual container of SDH with 34.368 Mbit/s transmission rate	VPL	Virtual path link
VC-32	Virtual container of SDH with 44.736 Mbit/s transmission rate	VPN	Virtual private network
VC-4	Virtual container of SDH with 149.76 Mbit/s transmission rate	VPRPC	Broadband virtual path service for reserved and permanent connections
VC- <i>n</i>	Virtual container of order <i>n</i> (SDH)	VPT	Virtual path terminator
VC CEPF	Virtual connection connection end point functions	VR	Virtual route (IBM SNA)
		VSAT	Verysmall aperture terminal (satellite)
		VSP	Virtual switching point
		VSWR	Voltage standing wave ratio
		VT	Virtual tributary (SONET)
		VT1.5	Virtual tributary of SONET with 1544 kbit/s transmission rate
		VT100	ASCII-based terminal developed by the Digital equipment corporation
		VT2.0	Virtual tributary of SONET with 2048 kbit/s transmission rate
		VT3.0	Virtual tributary of SONET with 3152 kbit/s transmission rate
		VT6.0	Virtual tributary of SONET with 6312 kbit/s transmission rate
		VTAM	Virtual telecommunications access method (IBM SNA)
		VTE	Virtual terminal environment
		VTOA	Voice over ATM
		W/U	Wanted to unwanted signal ratio



WAL2	Line code	X.500	ITU-T recommendation defining the directory service
WAN	Wide area network		
WARC	World administrative radio council	X.75	ITU-T recommendation defining inter-network interface between different packet switched networks
WDM	Wavelength division multiplexing		
WG	Working group		
wink start	A type of off-hook signalling	XDR	External data representation
WLAN	Wireless local area network	XID	Exchange identification
WLL	Wireless local loop	XIF	Cross-polarization improvement factor due to XPIC operation (radio)
WS	Workstation		
WSF	Workstation function	XMP	X manager protocol (a mapping function, common handler between CMIS/CMIP)
WTSC	World telecommunications standardization conference (ITU)		
WWW	Worldwide web (Internet)	XNS	Xerox network system
X	Interface between two management systems (TMN)	XPD	Cross-polar discrimination
x	reference point between one OSF and another (i.e. between TMNs)	XPG	X/Open portability guides
X-OFF	Transmit 'off' at DTE/DCE interface, data may not be sent	XPI	Cross-polar interference (radio)
X-ON	Transmit 'on' at DTE/DCE interface, data may be sent	XPIC	Cross-polar interference canceller (radio)
X.121	ITU-T recommendation defining public data network numbering scheme	XTI	X/Open transport interface
X.21	ITU-T recommendation X.21 (DTE/DCE interface for digital lines used in Europe)	XTP	Xpress transfer protocol
X.21	ITU-T recommendation defining a physical layer interface type (used in Europe for DTE rates above 64 kbit/s)	ZBTSI	Zero byte time slot interchange (line code)
X.21bis	ITU-T recommendation defining DCE/DTE physical layer interface where an analogue line is used between DCEs	ZIP	Zone information protocol (Appletalk)
X.25	ITU-T recommendation defining the interface between a DTE and packet network DCE		
X.28	ITU-T recommendation defining interface between an asynchronous DTE and a packet-switched network (PAD)		
X.29	ITU-T recommendation defining procedure for communication between a synchronous DTE (X.25) and an asynchronous DTE (X.28)		
X.3	ITU-T recommendation defining the functions of a packet assembler/disassembler		
X.400	ITU-T recommendation defining the message handling service (MHS)		

#### Two-letter country code abbreviations used by ISO (ISO 3166)

These codes are employed by X.400/X.500)

Afghanistan	AF
Albania	AL
Algeria	DZ
American Samoa	AS
Andorra	AD
Angola	AO
Anguilla	AI
Antarctica	AQ
Antigua and Barbuda	AG
Argentina	AR
Armenia	AM
Aruba	AW
Australia	AU
Austria	AT
Azerbaijan	A
Bahamas	BS
Bahrain	BH
Bangladesh	BD
Barbados	BB
Belarus	BY
Belgium	BE

Belize	BZ	French Guiana	GF
Benin	BJ	French Polynesia	PF
Bermuda	BM	French Southern Territories	TF
Bhutan	BT	Gabon	GA
Bolivia	BO	Gambia	GM
Bosnia and Herzegovina	BA	Georgia	GE
Botswana	BW	Germany	DE
Bouvet Island	BV	Ghana	GH
Brazil	BR	Gibraltar	GI
British Indian Ocean Territory	IO	Greece	GR
Brunei Darussalam	BN	Greenland	GL
Bulgaria	BG	Grenada	GD
Burkina Faso	BF	Guadaloupe	GP
Burundi	BI	Guam	GU
Cambodia	KH	Guatemala	GT
Cameroon	CM	Guinea	GN
Canada	CA	Guinea-Bissau	GW
Cape Verde	CV	Guyana	GY
Cayman Islands	KY	Haiti	HT
Central African Republic	CF	Heard and McDonald Islands	HM
Chad	TD	Honduras	HN
Chile	CL	Hong Kong	HK
China	CN	Hungary	HU
Christmas Island	CX	Iceland	IS
Cocos (Keeling) Islands	CC	India	IN
Colombia	CO	Indonesia	ID
Comoros	KM	Iran (Islamic Republic of)	IR
Congo	CG	Iraq	IQ
Cook Islands	CK	Ireland	IE
Costa Rica	CR	Israel	IL
Cote d'Ivoire	CI	Italy	IT
Croatia	HR	Jamaica	JM
Cuba	CU	Japan	JP
Cyprus	CY	Jordan	JO
Czech Republic	CZ	Kazakhstan	KZ
Denmark	DK	Kenya	KE
Djibouti	DJ	Kiribati	KI
Dominica	DM	Korea, Democratic People's Republic of	KP
Dominican Republic	DO	Korea, Republic of	KR
East Timor	TP	Kuwait	KW
Ecuador	EC	Kyrgyzstan	KG
Egypt	EG	Lao People's Democratic Republic	LA
El Salvador	SV	Latvia	LV
Equatorial Guinea	GQ	Lebanon	LB
Estonia	EE	Lesotho	LS
Ethiopia	ET	Liberia	LR
Falklands Islands (Malvinas)	FK	Libyan Arab Jamahiriya	LY
Faroe Islands	FO	Liechtenstein	LI
Fiji	FJ	Lithuania	LT
Finland	FI	Luxembourg	LU
France	FR	Macau	MO

Madagascar	MG	Saint Helena	SH
Malawi	MW	Saint Kitts and Nevis	KN
Malaysia	MY	Saint Lucia	LC
Maldives	MV	Saint Pierre and Miquelon	PM
Mali	ML	Saint Vincent and the Grenadines	VC
Malta	MT	Samoa	WS
Marshall Islands	MH	San Marino	SM
Martinique	MQ	Sao Tome and Principe	ST
Mauritania	MR	Saudi Arabia	SA
Mauritius	MU	Senegal	SN
Mexico	MX	Seychelles	SC
Micronesia, Federate States of	FM	Sierra Leone	SL
Moldova, Republic of	MD	Singapore	SG
Monaco	MC	Slovak Republic	CS
Mongolia	MN	Slovenia	SI
Montserrat	MS	Soloman Islands	SB
Morocco	MA	Somalia	SO
Mozambique	MZ	South Africa	ZA
Myanmar	MM	Spain	ES
Namibia	NA	Sri Lanka	LK
Nauru	NR	Sudan	SD
Nepal	NP	Suriname	SR
Netherlands	NL	Svalbard and Jen Mayen Islands	SJ
Netherlands Antilles	AN	Swaziland	SZ
Neutral Zone	NT	Sweden	SE
New Caledonia	NC	Switzerland	CH
New Zealand	NZ	Syrian Arab Republic	SY
Nicaragua	NI	Taiwan, Province of China	TW
Niger	NE	Tajikistan	TJ
Nigeria	NG	Tanzania, United Republic Of	TZ
Niue	NU	Thailand	TH
Norfolk Island	NF	The Former Yugoslav Republic of Macedonia	YM
Northern Mariana Islands	MP	Togo	TG
Norway	NO	Tokelau	TK
Oman	OM	Tonga	TO
Pakistan	PK	Trinidad and Tobago	TT
Palau	PW	Tunisia	TN
Panama	PA	Turkey	TR
Papua New Guinea	PG	Turkmenistan	TM
Paraguay	PY	Turks and Caicos Islands	TC
Peru	PE	Tuvalu	TV
Phillippines	PH	Uganda	UG
Pitcairn	PN	Ukraine	UA
Poland	PL	United Arab Emirates	AE
Portugal	PT	United Kingdom	GB
Puerto Rico	PR	United States of America	US
Qatar	QA	Uruguay	UY
Reunion	RE	Uzbekistan	UZ
Romania	RO	Vanuatu	VU
Russian Federation	RU	Vatican City State	VA
Rwanda	RW	Venezuela	VE

Vietnam	VN	Yemen, Republic of	YE
Virgin Islands (British)	VG	Yugoslavia	YU
Virgin Islands (American)	VI	Zaire	ZR
Wallis and Futuna Islands	WF	Zambia	ZM
Western Sahara	EH	Zimbabwe	ZW



---

# *Index*

---

0898 service 239  
1 for N restoration 688  
1 main + 1 standby protection 270  
1.5 Mbit/s digital line system 64  
1.920 Mbit/s bitrate 65  
1:1 restoration 278  
10-availability 86, 94  
1020 Hz tone 670  
10base2 ethernet LAN 369  
10base5 ethernet LAN 369  
10baseT ethernet LAN 370  
15-pin plug or socket 188  
155 Mbit/s signal 70  
16 Mbit/s token ring LAN 372  
1VF signalling 119  
2 Mbit/s digital line system 64  
20-availability 86, 94  
23B + D 223  
24-availability 86  
25-pin plug or socket 188  
2713 Hz tone 679  
2B + D 216  
2VF signalling 119  
3.1 kHz bearer service 215  
30B + D 222  
3270 emulation 377  
3270 gateway 377  
34-pin plug or socket 188  
37-pin plug or socket 188  
3Com 500  
4 Mbit/s token ring LAN 372  
500-point switching 88  
64 kbit/s bearer service 215, 219  
800-service 231, 237, 260

800 Hz tone 670  
838, ITU-R recommendation 627  
86/361/EEC mandate of July 1986 803  
900-service 239  
  
A-field 327  
A-law code 61, 63  
A-to-Mu law conversion 70  
AA (administrative authority) 528  
AAL-PDU (ATM adaptation layer protocol data unit) 467  
AAL-SDU (ATM adaptation layer service data unit) 467  
AAL class A 458  
AAL class B 458  
AAL class C 458  
AAL class D 458  
AAL1 (AAL type 1) 458  
AAL2 (AAL type 2) 458  
AAL3/4 (AAL type 3/4) 458  
AAL5 (AAL type 5) 458, 470  
AAR (automatic alternative routing) 577  
abbreviated dialling 115  
abort command 262  
ABR (answer bid ratio) 562, 645, 686  
absorption 166, 627  
abstract syntax notation 1 (ASN.1) 196, 204, 485  
abuse 720  
acceptability 641  
access accuracy 640  
access control 712, 713, 715  
access control field 395  
access dependability 640  
access network (AN) 329, 332, 333, 397

- access point 675
- access speed 640
- access unit (AU) 393
- accessibility 645
- accommodation 766, 771, 772
- accounting 647, 648, 652
- accounting management 485
- accounting rate 659
- accounting settlement 653
- accuracy 640
- ACD (automatic call distribution) 288
- ACK (acknowledge) 198
- acknowledged connectionless mode 374
- acknowledged signalling 126
- ACSE (access control and signalling equipment) 313
- active c-channel 336
- adaptation layer, ATM (AAL) 458, 465, 467, 468, 470
- adaptive differential pulse code modulation (ADPCM) 310, 441, 611, 696, 705
- adaptive engineering 763
- ADC (advice of duration and charge) 290
- ADC (american digital cellular system) 311
- add/drop multiplexor (ADM) 278
- address 455
- address field format 387
- address resolution protocol (ARP) 406
- address, layer 2 199
- address, layer 3 200
- addressing plan 513
- adds and changes 371
- adjacent cell 300
- ADM (add/drop multiplexor) 278
- ADMD (administration management domain) 420
- administration 752
- administration management domain (ADMD) 420
- administrative authority (AA) 528
- administrative domain 525
- administrative unit (AU) 271
- administrative unit group (AUG) 271
- ADPCM (adaptive differential pulse code modulation) 310, 441, 611, 696, 705
- ADSL (asymmetric digital subscriber line) 148, 337
- advertise 502
- advertising industry 792
- advice of duration and charge (ADC) 290
- advisory tone 118
- AE (application entity) 195, 261
- AFI (authority and format identifier) 527
- AFI = 39 527
- AFI = 45 527
- AFI = 47 527
- AGC (automatic gain control) 155
- ageing 675
- agent 482
- air-time 829
- Air Canada 247
- air interface (AI) 432
- air interface, DECT 326
- Airloop 328
- Airspan 328
- alarm 478, 676
- alert and response signal 436
- alerting 26, 109, 425
- aligning 271
- Alliance for Telecommunications Industry Solutions 731
- Aloha protocol 432
- alteration of service 643
- alternate mark inversion (AMI) code 72, 73
- alternative routing 502
- AM (amplitude modulation) 180
- America Online (AOL) 411
- Ameritech 811
- AMI (alternate mark inversion) code 72, 73
- amplification 29, 32
- amplifier 32
- amplitude 58, 183
- amplitude modulation (AM) 180
- AMPS (Advanced Mobile Telephone System) 244, 305
- ANA (Article Numbering Association) 787
- analogue-to-digital conversion 600, 609, 611, 673
- analogue modulation 153
- analogue signal 6, 17
- ANSI (American National Standards Institute) 729
- ANSI X.12 788
- answer 561
- answer bid ratio (ABR) 562, 645, 686
- answer seizure ratio (ASR) 562, 686
- answer time 544
- answerback 522, 718
- answering rate 678
- answerphone 246
- ant-trust 811
- antenna 157
- antenna lobe diagram 158
- anti-sidetone circuit 23
- antiphase timeslot 105
- Apple Macintosh 375
- Appletalk 374, 375, 376

- application entity (AE) 195, 261
- application layer (layer 7) 195, 202
- application service element (ASE) 195, 254, 262
- APPN (advanced peer-to-peer networking) 366
- approvals body 798, 803
- approvals procedure 802
- APS (automatic protection switching) 277
- ARDIS 430
- area code 114, 116
- ARP (address resolution protocol) 406
- ARR (automatic re-routing) 584
- array antenna 158
- arrival rate 535
- arriving calls 552, 553
- AS ART 100A directive on ONP 804
- AS400 366, 372
- ASAI (adjunct/switch application interface) 226
- ASCII (American standard code for information interchange) 46, 47, 48, 178, 196, 729
- ASE (application service element) 195, 254, 262
- ASN.1 (abstract syntax notation 1) 196, 204, 485
- ASR (answer seizure ratio) 562, 686
- assistance 281
- assistance service 289
- assistance, telex 294
- associated mode signalling 252
- asymmetric digital subscriber line (ADSL) 148, 337
- asynchronous data transfer 191
- asynchronous framing 276
- asynchronous terminal 355, 415
- asynchronous transfer mode (ATM) 361, 381, 397, 445, 451, 501, 527, 606, 655
- AT&T (American Telephone and Telegraph Company) 231, 305, 506, 586, 736, 794, 837
- AT&T anti-trust suit 811
- AT&T divestiture 811
- AT-command 192
- Atlantic Ocean region (AOR) 629, 630
- ATM-SDU (ATM layer service data unit) 467
- ATM (asynchronous transfer mode) 361, 381, 397, 445, 451, 501, 527, 606, 655
- ATM adaptation layer (AAL) 458, 465, 467, 468, 470
- ATM adaptation layer protocol data unit (AAL-PDU) 467
- ATM adaptation layer service data unit (AAL-SDU) 467
- ATM crossconnect 456, 457, 463
- ATM customer equipment 457
- ATM Forum 452, 457, 470, 506, 735
- ATM Forum network management model 472
- ATM Forum network reference model 471
- ATM layer 465, 468
- ATM layer entity (ATM-LE) 467
- ATM logical network address 527
- ATM multiplexor 456, 457, 463
- ATM network address 527
- ATM network node 447
- ATM network, private 473
- ATM protocol 465
- ATM switch 456, 457, 463
- ATM transport network 465, 468
- ATME (automatic transmission, measuring and signalling test equipment) 676
- atmospheric interference 607
- attempted calls 556
- attenuation 29, 31, 36, 41, 141
- attenuation constant 30
- attenuation distortion 30, 37, 608
- attenuation due to rain 627
- attenuation, radio 625, 627
- attenuator 599
- attribute 483, 501
- AU (administrative unit) 271
- audio circuit 141
- AUG (administrative unit group) 271
- aural range 22
- authorization 303, 714, 715
- authorization centre (AuC) 308, 309
- automatic alternative routing (AAR) 577
- automatic call distribution (ACD) 288
- automatic gain control (AGC) 155
- automatic network 110
- automatic protection switching (APS) 277
- automatic re-request 429
- automatic re-routing (ARR) 584
- auxiliary magnet 100
- availability 641, 645
- average call holding time 535
- average data delay 550
- average delay 545, 547
- average long term marginal cost 660
- average number of waiting calls 545
- average service time 544, 545
- B-field 327
- B-ICI (broadband inter-carrier interface) 457, 471, 472, 506, 730
- B-ISDN (broadband integrated services digital network) 394, 439, 441, 451, 527
- B-ISDN (broadband ISDN) 229



- B-ISSI (broadband inter-switching system interface) 472
- B-ISUP (broadband integrated services user part) 470, 471
- B-NT1 (broadband network termination type 1) 446
- B-TA (broadband terminal adaptor) 447
- B-TE (broadband ISDN terminal equipment) 456
- B-TE1 (broadband terminal equipment type 1) 447
- B-TE2 (broadband terminal equipment type 2) 447
- B channel 216
- B8ZS (bipolar 8-zero substitution) code 74
- BABT (British Approvals Board for Telecommunications) 734, 808
- backward compatibility 264, 680
- backward explicit congestion notification (BECN) 385, 387, 687
- BAKOM (Bundesamt für Kommunikation, Switzerland) 808
- balance circuit 34, 602
- balance of trade 659
- balanced circuit 144, 188
- balanced class of procedure 347, 348
- balanced mode 348
- balanced portfolio 832
- balancing 34
- baluns 370, 372
- bandwidth compression 703
- bankruptcy 771
- BAPT (Bundesamt für Post und Telekommunikation, Germany) 625, 734, 807
- bar coding 787
- Barclays Bank 800
- barrier box 819
- base station 244, 298, 300
- base station sub-system (BSS) 308
- base station switching centre (BSC) 308
- base transmitter station (BTS) 308
- baseband 39, 141, 145
- basic rate access (BRA) 215, 216, 217
- basic rate interface (BRI) 215, 216, 217
- basic transmission service 799
- basket of services 800
- bass tone 30
- battery conservation 429
- Baud rate 47, 182
- Baudot code 45, 46
- BBN 347
- BCD (binary coded decimal) 521
- BDT (Telecommunication Development Bureau of ITU) 725, 726
- Be (excess burst duration) 385
- bearer channel 216, 335
- bearer channel connection (BCC) 335
- bearer circuit 700
- bearer service 214, 215, 431
- BECN (backward explicit congestion notification) 385, 387, 687
- begin command 262
- Belgian Institute for Telecommunications 807
- Bell Atlantic 811
- Bell Laboratories 306, 328
- Bell South 811
- Bell telephone system 811
- Bell, Alexander Graham 21
- Bellcore (Bell Communications Research) 231, 737
- BER (bit error ratio) 57, 325, 594, 600, 609, 641, 673
- BGP (border gateway protocol) 406, 407, 500
- BHCA (busy hour call attempt) 560, 772
- bid 560, 561, 686
- Bildschirmtext 399, 400
- billing 648, 657
- binary code 43, 44
- binary coded decimal (BCD) code 521
- BISYNC 365
- bit 44
- bit error ratio (BER) 57, 325, 594, 600, 609, 641, 673
- bit inserted parity (BIP) 276
- bit rate 61
- bit sequence integrity 230
- bit stealing 68
- bit stuffing 68
- bit synchronization 190
- bit synchronous framing 275
- bit timing 465
- bit/s 45
- bits per second (bps) 45
- blackspot 435
- blocking 82, 87
- blocking probability 87
- blocking ratio 537
- blue book 739, 740
- BML (business management layer) 486, 487
- BMPT (Bundesministerium für Post und Telekommunikation, Germany) 807
- bobbin 100
- BOC (Bell Operating Companies) 794
- body of message 419
- border gateway protocol (BGP) 406, 407, 500
- border node 366

- bothway generator signalling 109
- BPON (broadband passive optical network) 152, 331, 620
- BRA (basic rate access) 215, 216, 217
- break-even 833
- BRI (basic rate interface) 215, 216, 217
- bridge 375, 376
- bridge armature 98
- bridge common 97
- British Petroleum (BP) 800
- British Standards 527
- British Telecom 328, 332, 399, 487, 506, 509, 525, 586, 620, 650, 735, 737, 799, 806, 837
- British Telecommunications Act, 1981 (UK) 806, 808
- broadband 371, 391, 439
- broadband integrated services digital network (B-ISDN) 394, 439, 441, 451, 527
- broadband inter-carrier interface (B-ICI) 457, 471, 472, 506, 730
- broadband ISDN (B-ISDN) 229
- broadband passive optical network (BPON) 152, 331, 620
- broadcast 158, 434, 625, 718
- browser 405
- BSC (base station switching centre) 308
- BSC (binary synchronous communication) 365
- BSGL (branch systems general licence, UK) 808
- BSI (British Standards Institution) 731
- BSS (base station sub-system) 308
- BTNR (British Telecom Network Requirement) 737
- BTNR 190 (DPNSS) 737
- BTS (base transmitter station) 308
- BtX (Bildschirmtext) 399, 400
- buffer 342, 384, 453, 543
- bugging 713
- building cabling 816
- Bündelfunk 299
- Bundesamt für Kommunikation (BAKOM, Switzerland) 808
- Bundesamt für Post und Telekommunikation, Germany (BAPT) 625, 734, 807
- Bundesamt für Zulassung in der Telekommunikation (BZT) 734, 807
- Bundesministerium für Post und Telekommunikation, Germany (BMPT) 807
- bursty traffic 341
- bus 394
- business development 836
- business management layer (BML) 486, 487
- busy hour 533, 556, 565
- busy hour call attempt (BHCA) 560, 772
- busy tone 77
- bypass 824, 826, 827
- byte synchronous framing 275
- BZT (Bundesamt für Zulassung in der Telekommunikation) 734, 807
- C/I (carrier to interference) 325
- C/R (command/response bit) 387
- C-450 306
- C-900
- 306 c-channel 335
- c-channel (communication channel) 335
- c-channel, active 336
- c-channel, standby 336
- c-path 335
- c-path (communications path) 335
- c-plane 326, 432, 461
- cabinet, streetside distribution 330, 619
- Cable & Wireless 799
- cable planning 616
- cable sheath 144
- cable tensile strength 617
- cabling regulations 819
- CAC (channel access control) 435
- CAD (computer-aided design) 779, 780, 787
- CAI (common air interface) 322, 324
- calibration 667
- call arrival rate 535
- call arrivals 552, 553
- call attempts 556, 560, 686
- call completions 556, 561
- call connection to busy subscriber (CCBS) 215
- call control 109
- call diversion 215, 718
- call duration 649
- call forwarding 511
- call gapping 692
- call gapping, 1 in N 692
- call gapping, 1 in T 692
- call holding time 552
- call holding time, average 535
- call logging system 556
- call re-direction 226, 243
- call set up 503, 558
- call set up time 508
- call waiting system 543
- calling card service 236
- calling line identity (CLI) 111, 215, 226, 289, 508, 715, 716, 718

- CAM (computer-aided manufacturing) 787
- campus network 391
- capability set 1 (CS1) 264
- capacitance 29, 30, 31
- capacity 540, 749
- carphone 301, 310
- carried traffic 531, 539
- carrier frequency 39, 146
- carrier signal 39, 40, 154
- Carterfone 811
- CAS (channel associated signalling) 121, 131
- cascade settlement 653
- category 1 (cat 1) cable 143
- category 2 (cat 2) cable 143
- category 3 (cat 3) cable 143
- category 4 (cat 4) cable 143
- category 5 (cat 5) cable 143, 372, 817
- CBO (continuous bit stream oriented) 389
- CBR (constant bit rate) 447, 458
- CC (country code) 514, 517, 522
- CCBS (call connection to busy subscriber) 215
- CCD (contract completion date) 766
- CCIR (Consultative Committee for Radiocommunication – now ITU-R) 725
- CCITT (International Telegraph and Telephone Consultative Committee – now ITU-T) 726
- CCITT 1 signalling 119, 132
- CCITT 2 signalling 119
- CCITT 3 signalling 119, 132
- CCITT 4 signalling 119, 132
- CCITT 5 signalling 119, 132
- CCITT 5bis signalling 119
- CCITT 6 signalling 119
- CCITT 7 signalling 119, 249
- CCITT blue book 739, 740
- CCITT R1 signalling 119, 132
- CCITT R2 signalling 119, 121, 132
- CCITT R2D signalling 119
- CCITT recommendations 738
- CCITT red book 739, 740
- CCITT yellow book 740
- Cemail 402
- CCS (common channel signalling) 121, 250
- CCS (hundred call seconds) 530
- CCS7 (common channel signalling system 7) 229
- CDMA (code division multiple access) 174
- CDV (cell delay variation) 607, 641
- CEI (Commission Electrotechnique Internationale) 726
- cell 395
- cell-switched 12
- cell delay variation (CDV) 607, 641
- cell error ratio (CER) 641
- cell header 455, 464
- cell loss priority (CLP) 464
- cell misinsertion rate (CMR) 641
- cell rate adaption 465
- cell relay 452
- cell relay service (CRS) 472
- cell size 454, 547
- cell splitting 303
- cell switching 106
- cell transfer capacity 641
- cell transfer delay 641
- Cellnet 311, 806
- cellular radio 163, 244, 299
- CELP (code excited linear prediction) 611, 706
- CEN (Comité Européen de Normalisation) 727
- CENELEC 727
- central battery (CB) 24
- central office 240
- centralized operation and maintenance (CO&M) 666, 667
- centrex 239
- CEPT (European Conference for Posts and Telecommunications) 64, 728, 307, 321
- CEQ (customer equipment) 449, 456
- CER (cell error ratio) 641
- CES (circuit emulation service) 458
- changeover 689
- channel 37
- channel access control (CAC) 435
- channel associated signalling (CAS) 121, 131
- channel interference 301
- channel service unit (CSU) 180, 188
- channel translating equipment (CTE) 38, 123, 124, 146, 696
- character synchronization 191
- chargeable duration 649
- chargeband 650
- charges 645
- charging 14, 647, 648
- cheapernet 369
- check bits 256
- checksum 409
- CHILL (CCITT high level language) 739
- chinese wall 801
- chip rate 174
- CIC (circuit identification code) 257, 260
- CIM (computer-integrated manufacturing) 787
- CIR (committed information rate) 383, 385, 389, 496, 655, 687

- circuit 12
- circuit emulation service (CES) 458
- circuit identification code (CIC) 257, 260
- circuit infill 759
- circuit monitoring 289
- circuit multiplication 700
- circuit multiplication equipment (CME) 696, 699
- circuit occupancy 537, 538
- circuit provision 615
- circuit requirement 542
- circuit stability 35
- circuit switched public data network (CSPDN) 206
- circuit switching 12, 77, 259, 453, 530, 556
- circular routing 492, 494, 497
- Cisco Systems 500
- citizens band (CB) radio 161
- class of service 234
- clear signal 115
- cleardown 109, 125, 559
- clearing 77, 80
- CLI (calling line identity) 111, 215, 226, 289, 508, 715, 716, 718
- client/server architecture 383
- climate zone 627
- clipping of speech 444, 703, 710
- closed user group (CUG) 353, 496, 655, 715, 716, 717
- CLNS (connectionless network service) 11, 431
- CLP (cell loss priority) 464
- cluster controller 361, 362, 364
- CME (circuit multiplication equipment) 696, 699
- CMI (coded mark inversion) code 72
- CMIP (common management information protocol) 477, 483, 489
- CMIS (common management information service) 483
- CMISE (common management information service entity) 483
- CMR (cell misinsertion rate) 641
- CNET (Centre nationale d'Etudes Télécommunications, France) 807
- coaxial cable 19, 143
- COBOL 729
- code 11 130, 291
- code 12 130, 291
- code 13 130
- code 15 130
- code conversion 421
- code division multiple access (CDMA) 174
- code of conduct 801
- Code of Federal Regulations (CFR) 810
- code receiver 112
- codec (coder/decoder) 698
- coded mark inversion (CMI) code 72
- codeword 713
- coding violation 73
- collapsed backbone 371, 373
- collect call 286, 290, 656
- collision 371
- collocation 508
- COLT (City of London Telecommunications) 330
- commissioning 667
- commissioning objective 674
- committed information rate (CIR) 383, 385, 389, 496, 655, 687
- common air interface (CAI) 322, 324
- common carrier 810
- common channel signalling (CCS) 121, 250
- common equipment 560
- common equipment holding time 560
- common management information protocol (CMIP) 477, 483, 489
- common management information service (CMIS) 483
- common management information service entity (CMISE) 483
- common part (CP) 465
- common volume 167
- communication channel (c-channel) 335
- communication controller 361, 362, 364
- Communications Act of 1934 (United States) 810
- communications path (c-path) 335
- compaction 707
- companding 703
- company mission 831
- compatibility 510, 724
- compelled signalling 126
- competition 802
- completed calls 556, 561
- component failure 664
- composite forecasting 565
- compression 607, 673, 699, 703, 707
- CompuServe 402, 411
- computer-aided design (CAD) 779, 780, 787
- computer-aided manufacturing (CAM) 787
- computer-integrated manufacturing (CIM) 787
- computer network 177, 210
- computer peripheral 367
- Comvik 306
- concentration function 333, 334
- conduct regulation 797
- conference call service 239, 799

- confidential information 719
- confidentiality 711
- configuration change 680
- configuration management 485
- configuration plan 14
- conformance 510, 724
- congestion 492, 530, 540, 560, 576, 634, 645, 683
- congestion control 384
- connection-oriented 11, 12, 13, 109, 259, 458, 499, 586
- connection-oriented network service (CONS) 431
- connection-oriented transport service (COTS) 11
- connection management (CMT) 392
- connection mode 374
- connection reference points 596
- connection set up delay 641
- connection set up denial ratio 641
- connectionless 11, 12, 139, 259, 374, 458, 587
- connectionless network service (CNLS) 11, 431
- CONS (connection-oriented network service) 11, 431
- constant bit rate (CBR) 447, 458
- container-4 (C-4) 271
- container (SDH) 271, 273
- container frame size 273
- content 416, 417, 419, 805
- contention resolution 435
- continental destination 756
- continue command 262
- continuous bit stream oriented (CBO) 389
- contra-rotating 392
- contract conditions 770
- contract dispute 771
- control bus 107
- control channel 304
- control plane 136, 459, 469
- control plane AAL 469
- control plane communication (access) 459
- control plane communication (network) 459
- controlled maintenance 664
- controlled transmission 465
- controller 361
- convergence of computing and telecommunications 794
- convergence sublayer (CS) 465, 468
- conversation 12, 125, 645
- conversation time 529, 555, 558
- conversational service 442
- conversion, A-to-Mu law 70
- conversion, analogue-to-digital 600, 609, 611, 673
- copper wire 19
- copyright 770
- CORBA (common object request broker architecture) 488
- cord 287
- cordless telephony 297, 319, 321
- corporate inefficiency 784
- corporate ISDN 227
- corporate networks 805, 815
- corrective maintenance 664
- corruption 603, 711
- cost minimisation 695
- cost plus (charging) 649
- COTS (connection-oriented transport service) 11
- country code (CC) 514, 517, 522
- CP (common part) 465
- CPE (customer premises equipment) 216, 724, 798
- cradle 110
- crankback 502, 584
- CRC (cyclic redundancy check) 201, 256, 349, 387, 504
- credit card telephone 238
- critical path analysis 766, 767
- cross-exchange traffic 530
- cross-subsidization 801, 806
- crossbar switching 96
- crossconnect 278, 691, 756
- crossconnect, ATM 456, 457, 463
- crosspoint 77
- crosstalk 33, 34, 41, 594, 607, 670
- CRS (cell relay service) 472
- CS (convergence sublayer) 465, 468
- CS1 (capability set 1) 264
- CSMA/CD (carrier sense multiple access with collision detection) 368, 369
- CSU (channel service unit) 180, 188
- CT2 (2nd generation cordless telephony) 322
- CTE (channel translating equipment) 38, 123, 124, 146, 696
- CUG (closed user group) 353, 496, 655, 715, 716, 717
- customer network management 473
- customer number 116
- customer premises equipment (CPE) 216, 724, 798
- customer satisfaction 635, 641
- customer service 643
- cut off 645
- cyclic redundancy check (CRC) 201, 256, 349, 387, 504
- D-field 327
- D-socket 185
- D channel 216

- D1-interface 324, 325
- D2-interface 324, 325
- D3-interface 324, 326
- DAMA (demand assigned multiple access) 172
- damage 664
- DAR (dynamic alternative routing) 586
- dash 20
- DASS2 (digital access signalling system 2) 823
- data 6, 43
- data circuit termination equipment (DCE) 179, 202, 346
- data communication 739
- data communication channel (DCC) 277, 279
- data communications network (DCN) 481
- data compactor 707
- data compression 699, 707
- data country code (DCC) 521, 522, 527
- data delay, average 550
- data flow control 362
- data link channel identifier (DLCI) 199, 384, 387
- data multiplexor 706
- data network 177, 205, 339, 739
- data network address 521
- data network identification code (DNIC) 521
- data protection 806
- data service unit (DSU) 180, 188
- data switching exchange (DSE) 347
- data terminal equipment (DTE) 179, 202, 346
- data user part (DUP) 250, 254, 258
- datagram 209, 344, 345, 408, 587
- datalink control 362
- datalink control layer, DECT 326
- datalink layer (layer 2) 198, 202
- Datex-J 399, 400
- Datex-L network 206
- DC (direct current) signalling 120
- DCC (data communication channel) 277, 279
- DCC (data country code) 521, 522, 527
- DCE (data circuit termination equipment) 179, 202, 346
- DCME (digital circuit multiplication equipment) 708
- DCN (data communications network) 481
- DCS-1800 310
- DDI (direct dialling in) 287, 519
- DE (discard eligibility) 387, 388
- dealer board 790
- DEC (Digital Equipment Corporation) 205, 361, 487, 735
- decency 805
- DECNET 205, 361
- DECT (digital European cordless telephony) 312, 322, 323, 829
- DECT air interface 326
- DECT, reference model 324
- deep sea optical fibre 618
- defect 633
- degradation 675
- delay 23
- delay-sensitive signals 445, 453, 454
- delay variation 445
- delivery envelope 417
- delivery note 785
- demand assigned multiple access (DAMA) 172
- demodulation 155
- demultiplexing 274
- dependability 640
- dependency 766
- depreciation 647
- deregulation 652, 793, 795
- derived circuit 700
- derived performance parameters 641
- DES (defence encryption standard) 713
- descriptive name 419
- design objective 674
- designated transit list (DTL) 502
- desktop quality video SSCS 470
- destination access control 712, 715
- destination point code (DPC) 257
- destination service access point (DSAP) 374
- destination settlement 654
- Deutsche Bundespost 805
- Deutsche Telekom 331, 333, 399, 400, 837
- DFI (DSP format identifier) 528
- diagnosis 675
- dial-back 721
- dial-off 720
- dial-on 720
- dial tone 25, 92, 110
- dialogue 382
- Diana, Princess of Wales 718
- DIB (directory information base) 525
- DID (direct inward dialling) 287
- differential Manchester code 72
- differential quaternary phase shift keying (DQPSK) 433
- diffraction 156
- digit analysis 114
- digital circuit multiplication equipment (DCME) 708
- digital crossconnect (DXC) 278
- digital encoding 17

- Digital Equipment Corporation (DEC) 205, 361, 487, 735
- digital information 6
- digital modulation 153
- digital multiplex interface (DMI) 229
- digital private network signalling system (DPNSS) 228, 737
- digital reference sequence (DRS) 671
- digital section (DS) 466
- digital signalling 120, 131
- digital subscriber signalling 1 (DSS1) 219, 260
- digital subscriber signalling 2 (DSS2) 470, 471, 472
- digital switching 101
- digital transmission 43, 55
- dimensioning 539, 542, 581
- diode 148
- dipole antenna 158
- Dirección General de Telecomunicaciones (DGT, Spain) 808
- direct accounting 653
- direct current (DC) signalling 120
- direct dialling in (DDI) 287, 519
- direct inward dialling (DID) 287
- Direction de la Réglementation Générale de Postes et Télécommunications (DGPT, France) 807
- directional antenna 158, 302, 320
- directionality 157
- Director General of Telecommunications (DGT) 806
- directory 806
- directory enquiry service (DQ) 290, 421, 806
- directory information 508
- directory information base (DIB) 525
- directory information tree (DIT) 525
- directory number (DN) 111
- directory schema 525
- directory scheme 525
- DISC (disconnect) 349
- discard eligibility (DE) 387, 388
- disconnect 77
- discontinuation of service 811
- discounted cash flow (DCF) 751, 752
- disengagement accuracy 640
- disengagement dependability 640
- disengagement speed 640
- disputed bill 646
- distortion 30, 36, 141, 594, 599, 604, 608, 645
- distributed data processing environment 365
- distributed intelligence 232
- distributed queue dual bus (DQDB) 394
- distribution cable 619
- distribution frame 619
- distribution function 401
- distribution industry 787
- distribution service 442
- disturbance 712, 714
- DIT (directory information tree) 525
- diversion 718
- DLCI (data link channel identifier) 199, 384, 387
- DMI (digital multiplex interface) 229
- DN (directory number) 111
- DNHR (dynamic non-hierarchical routing) 586
- DNIC (data network identification code) 521
- documentation 666, 768, 774
- Dolby noise reduction system 703
- domain specific part (DSP) 527
- dot 20
- double attached station (DAS) 392
- double satellite hop 495, 606
- downlink 306, 309, 311, 625, 629
- DPC (destination point code) 257
- DPCM (differential pulse code modulation) 696, 704
- DPNSS (digital private network signalling system) 228, 737
- DQ (directory enquiry) service 290, 421, 806
- DQDB (distributed queue dual bus) 394
- drop and insert multiplexor 269
- dropwire 619
- DRS (digital reference sequence) 671
- dry joint 664
- DS (digital section) 466
- DS0 signal 65, 268
- DS1 signal 64, 268
- DS3-signal 268
- DSAP (destination service access point) 374
- DSC (Digital Switch Corporation) 328
- DSE (data switching exchange) 347
- DSP (domain specific part) 527
- DSS1 (digital subscriber signalling 1) 219, 260
- DSS2 (digital subscriber signalling 2) 470, 471, 472
- DSU (data service unit) 180, 188
- DTE (data terminal equipment) 179, 202, 346
- DTI (Department of Trade and Industry, UK) 625, 733, 798, 808
- DTL (designated transit list) 502
- dual attached connection 392
- dumb terminal 364, 366, 377, 415
- duopoly review (UK) 805
- DUP (data user part) 250, 254, 258
- duplex operation 7, 56, 155

- DXC (digital crossconnect) 278  
dynamic alternative routing (DAR) 586  
dynamic non-hierarchical routing (DNHR) 586  
dynamic routing 585
- E&M signalling 124  
E-plus 311  
E-wire 124, 125  
E.163, ITU-T recommendation 513, 516, 518  
E.164, ITU-T recommendation 471, 513, 514, 516, 517, 518, 527  
E.506, ITU-T recommendation 565  
E1 signal 63, 180, 268, 828  
E3 signal 268, 828  
EA (extended addressing) 387  
earphone 23  
earth station 170, 628  
eavesdropping 711, 714, 715  
EBCDIC (extended binary coded decimal interchange code) 45, 49, 50, 196, 362  
ECC (embedded communication channel) 279  
echo 23, 593, 602, 603, 670  
echo canceller 606  
echo path 24  
echo suppressor 604  
ECMA-72 729  
ECMA (European Computer Manufacturers Association) 729  
econometric model 565  
ECS (European communications satellite) 119  
EDI (electronic data interchange) 399, 403, 422, 424, 787  
EDIFACT (electronic data interchange for administration, commerce and transport) 404, 788  
EDSSI (European DSSI) 227  
EEC (European Economic Community) 728, 793  
EEC framework 802  
effective capacity 749  
EFS (error free seconds) 673  
EFTA (European Free Trade Association) 728  
EFTPOS (electronic funds transfer at the point of sale) 788  
EGP (edge gateway protocol) 406, 407  
EGP (exterior gateway protocol) 500  
EIA (Electronic Industries Association) 730  
EIR (equipment identity register) 308, 309  
EIR (excess information rate) 385, 389  
election 501  
electro-mechanical exchange 79  
electromagnetic compatibility (EMC) 175  
electromagnetic interference (EMI) 143, 175, 719  
electromagnetic wave 153  
electronic data interchange (EDI) 399, 403, 422, 424, 787  
electronic mail 12, 399, 400, 719, 791  
electronic messaging 399  
electronic office 791  
electronic ticketing 652  
electrostatic induction 607  
element management layer (EML) 486  
email 400, 527  
embedded communication channel (ECC) 279  
EMC (electromagnetic compatibility) 175  
emergency service 286, 289, 508, 806  
EMI (electromagnetic interference) 143, 175, 719  
EML (element management layer) 486  
EN (end node) 366  
EN (European Norm) 727, 728  
EN (exchange number) 111  
*en bloc* signalling 503  
encryption 713  
encryption 309, 712, 713, 822  
encryption, key-coded 309, 715  
end-to-end operation 126  
end command 262  
end node (EN) 366  
end system identifier (ESI) 528  
enforced separation 800  
engaged tone 77  
engineering guidelines 494  
enhanced telephone user part (TUP+) 258  
enquiry 286, 291  
entertainment quality video SSCS 470  
entry barrier 796  
ENV (temporary European Norm) 728  
envelope 416, 419  
environment 771, 772  
EOT (end of transmission) 198  
EPOS (electronic point of sale) 788  
equal access 506, 510, 794, 802, 811  
equal access code 507  
equalization 29, 36, 608, 670  
equalizer 593  
equipment identity register (EIR) 308, 309  
equipment room 766  
equivalent random method 578, 579, 590  
equivalent random traffic (Aeq) 579  
ERA (exterior reachable address) 502  
Ericsson 430  
Eritel 430



- Erlang 530
- Erlang call-waiting formula 343, 544
- Erlang distribution 554, 572
- Erlang formula 557
- Erlang lost-call formula 539, 541, 554, 572
- Erlang waiting time formula 563
- Erlang, A.K. 529, 551
- Erlangs 556
- error correction 254, 380, 504
- error detection 254, 380, 504
- error free seconds (EFS) 673
- errored seconds 641
- escape code 521
- ESI (end system identifier) 528
- ET (exchange termination) 111
- ETACS (extended TACS) 306
- /etc/hosts' file 527
- ethernet LAN 73, 368
- ethernet LAN, 10base2 369
- ethernet LAN, 10base5 369
- ethernet LAN, 10baseT 370
- ETO (European Telecommunications Office) 728
- ETR (European Technical Report) 729
- ETS (European Telecommunication Standards) 728
- ETSI (European Telecommunications Standards Institute) 175, 231, 728, 732
- ETSI secretariat 525
- EU (European Union) 728, 793, 800, 804
- Europe Online 411
- European Commission 509, 794
- European communications satellite (ECS) 119
- European digital multiplex hierarchy 65
- European Telecommunications Standards Institute (ETSI) 175, 231, 728, 732
- EUTELSAT (European Telecommunications Satellite Organization) 628
- even parity 200
- excess burst (Be) 385
- excess burst duration (Be) 385
- excess information rate (EIR) 385, 389
- exchange busy hour 758
- exchange calling 24
- Exchange Carriers Standards Association 730
- exchange design 754
- exchange identifier (XID) 716
- exchange number (EN) 111
- exchange provision 749
- exchange termination (ET) 111
- expansive control action 687, 688
- exponential growth 569
- extended addressing (EA) 387
- extended binary coded decimal interchange code (EBCDIC) 45, 49, 50, 196, 362
- exterior reachable address (ERA) 502
- F-interface 481
- F.69, ITU-T recommendation 513, 524
- facsimile 49, 421
- facsimile server 375
- facsimile terminal 52
- fade probability 628
- fading 166, 309, 435, 627, 628
- fair trading 806
- fairness 397
- fan-in 83, 86
- fan-out 83, 86
- far end block error (FEBE) 276
- fault-reporting 666
- fault correction 675
- fault localization 676, 678
- fault management 485
- fault reporting 676
- FCC (Federal Communications Commission) 175, 436, 506, 625, 732, 794, 798, 809
- FCC registers 810
- FCC reports 810
- FCC rules and regulations 810
- FCS (frame check sequence) 348, 387
- FDD (frequency division duplex) 433
- FDDI-2 392, 393
- FDDI (fibre distributed data interface) 391, 392
- FDM (frequency division multiplexing) 37, 144, 145, 609, 696
- FDMA (frequency division multiple access) 147, 172, 629
- FEBE (far end block error) 276
- FEC (forward error correction) 429
- FECN (forward explicit congestion notification) 385, 687
- Federal Communications Commission (FCC) 175, 436, 506, 625, 732, 794, 798, 809
- feedback 34, 599
- FEP (front end processor) 202, 209
- fibre distributed data interface (FDDI) 391, 392
- fibre in the loop (FITL) 329
- fibre to the building (FTTB) 329
- fibre to the curb (FTTC) 329, 330
- fibre to the home (FTTH) 329, 331
- FIFO (first in first out) queue 343, 455
- file server 375
- file transfer protocol (FTP) 405, 406, 407, 408, 424

- file transfer protocol, trivial (TFTP) 405, 406, 407, 408
- file transfer, access and management (FTAM) 204, 424
- final route 578
- fire regulations 819
- firewall 715
- first-offered traffic 580
- first in first out (FIFO) queue 343, 455
- FITL (fibre in the loop) 329
- fitness for purpose 633
- five-tone radiopaging 427
- fixed costs 833
- fixed part (FP) 324
- fixed radio termination (FT) 324
- fixed relay station (FRS) 325, 326
- flag 255, 348, 387
- flooding 501
- flow control 254, 259, 357, 362, 464, 692, 693
- FM (frequency modulation) 181
- FMD services 362
- FMH (function management header) 363
- footprint 629
- forced release 118
- forecast traffic 557
- forecasting 555
- format 417
- FORTRAN 729
- forward error correction (FEC) 429
- forward explicit congestion notification (FECN) 385, 687
- forward transfer 289
- forwarding 433
- four-wire circuit 35, 141, 600, 602
- four-wire communication 32, 35
- four-wire switch 79
- FP (fixed part) 324
- FRAD (frame relay access device) 388
- frame 66
- frame check sequence (FCS) 348, 387
- frame discarding 693
- frame error ratio 641
- frame format 67, 70, 386
- frame generator 394, 397
- frame mode 220
- frame relay 379
- frame relay access device (FRAD) 388
- Frame Relay Forum 735
- frame relay protocol 361
- frame relay service (FRS) 472
- frame relay SSCS 470
- frame repetition rate 273
- frame size 547
- frame switching 107
- frame synchronization 190
- frame transmission rate 385
- framing 66, 268
- France Telecom 258, 399, 837
- fraud 650, 720
- free space transmission loss 625
- freephone 235, 237, 260
- freezeout 709, 710
- frequency attenuation distortion 30, 37, 593, 670
- frequency distortion 608
- frequency division duplex (FDD) 433
- frequency division multiple access (FDMA) 147, 172, 629
- frequency division multiplexing (FDM) 37, 144, 145, 609, 696
- frequency hopping 174, 309, 715
- frequency modulation (FM) 181
- frequency of errors 564
- frequency re-use 300, 301, 303
- frequency shift distortion 671, 672
- frequency shift keying (FSK) 153, 181
- frequency shifting 146
- "friends and family" 237
- front end processor (FEP) 202, 209
- FRS (fixed relay station) 325, 326
- FRS (frame relay service) 472
- FSK (frequency shift keying) 153, 181
- FT (fixed radio termination) 324
- FTAM (file transfer, access and management) 204, 424
- FTP (file transfer protocol) 405, 406, 407, 408, 424
- FTTB (fibre to the building) 329
- FTTC (fibre to the curb) 329, 330
- FTTH (fibre to the home) 329, 331
- full-life cost analysis 751
- full availability 11, 81, 83, 84
- full duplex operation 197
- function management header (FMH) 363
- functional specification 764, 771
- G.101, ITU-T recommendation 599
- G.102, ITU-T recommendation 599, 674
- G.103, ITU-T recommendation 599
- G.111, ITU-T recommendation 599
- G.113, ITU-T recommendation 611
- G.121, ITU-T recommendation 599
- G.180, ITU-T recommendation 645
- G.652, ITU-T recommendation 151

- G.653, ITU-T recommendation 151
- G.654, ITU-T recommendation 151
- G.703, ITU-T recommendation 73, 223, 358, 828
- G.721, ITU-T recommendation 706, 709
- G.726, ITU-T recommendation 706
- G.728, ITU-T recommendation 706
- G.957, ITU-T recommendation 278
- gain 701
- gain hits 671, 672
- GAP (Groupe d'Analyse et Prevision) 804
- gateway 505, 724, 375
- gateway (LAN) 377
- gateway protocol 358, 406, 407
- gauge 32, 142
- Gaussian frequency shift keying (GFSK) 327
- Gaussian minimum shift keying (GMSK) 310
- GDMO (guidelines for the definition of managed objects) 485
- GDMO browser 485
- general format identifier 351
- general maintenance 663
- general policy framework, 1988 (Australia) 812
- generator polynomial 201
- generator signalling 109, 284
- generic flow control (GFC) 464
- generic network performance parameters 640
- geostationary satellite 19, 167, 312, 436, 628
- GFC (generic flow control) 464
- global positioning system (GPS) 436
- global system for mobile communication (GSM) 244, 297, 307
- Global Traffic Meeting (GTM) 629
- Globalstar 171, 297, 314, 316
- GNP (gross national product) 795
- gold franc 654
- GOS (grade of service) 539, 540, 541, 557, 561, 571, 572
- GPO (General Post Office, UK) 737
- GPO (US government printing office) 810
- GPS (global positioning system) 436
- grade of service (GOS) 539, 540, 541, 557, 561, 571, 572
- graded index multimode fibre 149, 150
- grading 84, 91, 539
- green paper 802
- groomed network 615
- ground voltage 144
- group 37, 696, 828
- group 4 facsimile 52
- group delay 670
- group delay distortion 670
- group delay equalizer 608, 670
- group translating equipment (GTE) 146, 696
- Groupe Speciale-Mobiles (GSM) 307
- GSC code (radiopaging) 428
- GSM (global system for mobile communication) 244, 297, 307
- GTE (group translating equipment) 146, 696
- H.261, ITU-T recommendation 49, 221, 441
- H0 bitrate 65
- H11 bitrate 65
- H12 bitrate 65
- half-duplex operation 8, 155, 184, 197
- half echo suppressor 604, 605
- Hamming code 201
- hand-off 244, 304, 323
- hand generator signalling 284
- handover 304
- handshaking 192
- handy 310
- hardware faults 679
- hardware port identification 106
- hardware upgrade 681
- harmonic disturbance 671, 672
- harmonization 804
- hash 113
- Hayes command set 192
- HDB (home database) 325
- HDB3 (high density bipolar, order 3) code 72
- HDLC (high level datalink control) 198, 347
- HDSL (high bitrate digital subscriber line) 148, 337
- HDTV (high definition television) 442
- header checksum 409
- header error control (HEC) 464
- heading 419
- health and safety 774
- heartbeat 676
- HEC (header error control) 464
- hello procedure 502
- hello protocol 501
- hello state machine 500
- Hewlett Packard 487, 735
- HF (high frequency) radio 19, 153, 161
- HFC (hybrid fibre coax) 337, 338
- hierarchical network structure 494, 573
- high bitrate digital subscriber line (HDSL) 148, 337
- high density bipolar, order 3 (HDB3) code 72
- high frequency radio (HF) 19, 153, 161
- high level datalink control (HDLC) 198, 347
- high usage (HU) route 577, 578, 583
- higher layer protocol 221

- higher order transmission 613
- HIPERLAN 432, 434
- HLR (home location register) 244, 263, 305, 308
- holding 77, 80
- holding time 128, 529, 535, 552, 555, 558, 560
- holding time, short 678
- home database (HDB) 325
- home location register (HLR) 244, 263, 305, 308
- hook 110
- horizontal link 502
- host address 526
- host computer interface (HCI) 245
- host machine 361
- hus 370, 817, 824
- Huffmann code 707
- humidity 674
- hunt group 354
- hunter 89
- Hush-a-Phone 811
- hybrid 32, 35, 602
- hybrid fibre coax (HFC) 337, 338
- hybrid transformer 32, 35, 602
- hypergroup 37, 696
- hypertext 405
- I-format 349
- I.122, ITU-T recommendation 388
- I.233, ITU-T recommendation 388
- I.311, ITU-T recommendation 460
- I.363, ITU-T recommendation 470
- I.370, ITU-T recommendation 388
- I.413, ITU-T recommendation 446, 449
- I.420, ITU-T recommendation 218, 221, 227
- I.421, ITU-T recommendation 223, 227
- I.430, ITU-T recommendation 218
- I.441, ITU-T recommendation 218
- IA2 (international alphabet number 2) 45, 206, 421
- IA5 (international alphabet number 5) 46, 196, 421, 729
- IBM (International Business Machines) 341, 361, 487, 735, 738
- IBM 2780 365
- IBM 3174 346, 360, 362, 364, 377
- IBM 3270 364, 365, 377
- IBM 3274 362, 364
- IBM 327x 364
- IBM 370 364
- IBM 3725 362, 364
- IBM 3745 362, 364, 366, 377
- IBM 3780 365
- IBM 37x5 364
- IBM LAN manager 375
- IBM personal computer (PC) 372
- IBM RS6000 366
- IBM type 1 cable 372
- IBS (international business service) 631, 829
- ICI (inter-carrier interface) 137
- ICMP (internet control message protocol) 406, 410
- ICO 171
- icon 400
- identifier (ID) 455, 501
- identifier, service profile (SPID) 461
- identity code (ID) 436
- IDI (initial domain identifier) 527
- IDN (integrated digital network) 213
- IDP (initial domain part) 527
- IDP (inter-digit pause) 25, 93, 94
- IDR (intermediate data rate) 631
- IEC (inter-exchange carrier) 506, 794, 812
- IEC (International Electrotechnical Commission) 726
- IEEE (Institute of Electrical and Electronics Engineers) 730
- IEEE 802 367
- IEEE 802.11 432
- IEEE 802.2 368, 374, 391, 435
- IEEE 802.3 368, 369
- IEEE 802.4 368, 371
- IEEE 802.5 73, 368
- IEEE 802.9 394
- IETF (Internet Engineering Task Force) 489, 736
- IFRB (International Frequency Registration Board) 623, 725
- IHL (internet header length) 408
- IISP (interim inter-switch signalling protocol) 472
- ILMI (interim local management interface) 472, 489
- images, communication of 439
- IMCO (Inter-Governmental Maritime Consultative Organisation) 312
- impulse noise 671
- in-span interconnection (ISI) 509
- INAP (intelligent network application part) 232, 233, 250, 264
- inband 122
- inband signalling 122
- incoming circuit 77
- incoming traffic 757
- incorrect release ratio 641
- incorrect set up probability 641
- Indian Ocean region (IOR) 629, 630
- inductance 29, 30, 31
- induction 607

- information 3
- information conveyance rate 45
- information environment 4
- information flow 4, 780
- information loss 641
- information transfer 645
- information transfer accuracy 640
- information transfer dependability 640
- information transfer speed 640
- infrastructure 804
- ingress control 692
- INI (inter network interface) 137, 506
- INI, ATM 456, 457
- INMARSAT (International Maritime Satellite Organization) 312, 628
- INMARSAT standard C 313
- instability 33, 34, 599, 602, 603
- installation 766
- Institute of Electrical and Electronics Engineers (IEEE) 730
- Istituto Superiore Poste e Telecomunicazioni (Italy) 807
- insulation 616
- integrated digital network (IDN) 213
- integrated services digital network (ISDN) 206, 213, 249, 335, 739
- integrated services private branch exchange (ISPBX) 228
- integrated services user part (ISUP) 224, 258
- integrated voice data (IVD) LAN 394
- integration 761
- intellectual property rights (IPR) 510, 770, 806
- intelligent network 231, 235, 249, 737, 804
- intelligent network application part (INAP) 232, 233, 250, 264
- intelligent network architecture 232
- intelligent peripheral (IP) 234
- intelligent terminal 415
- INTELSAT (International Telecommunications Satellite Consortium) 169, 628, 630, 724, 727
- INTELSAT VIII satellite 171
- inter-carrier interface (ICI) 137
- inter-continental destination 756
- inter-digit pause (IDP) 25, 93, 94
- inter-exchange circuit 81
- inter-exchange signalling 114, 118, 604, 670
- inter-LATA 812
- inter-modulation noise 671
- inter-network interface (INI) 506
- inter-operability 805
- inter-PBX signalling 822
- inter-personal message service (IPMS) 415, 419
- inter-personal messaging protocol (P2) 419
- inter-register signalling 119, 121
- inter-system interface (ISI) 432
- inter network interface (INI) 137
- interactive service 442
- interception 711, 713, 715
- interconnection 491, 504, 506, 652, 805
- interconnection agreement 508
- interconnection charges 659
- interconnection contract 509
- interconnection, ex-ante conditions 510
- interconnection, in-span (ISI) 509
- interconnectivity 711
- interfaces 723, 804
- interference 56, 142, 158, 160, 166, 300, 303, 320, 435, 593, 594, 607, 616, 627, 739
- interfering zone 300
- interior routing protocol (IRP) 500
- interleaving 61
- intermediate service part (ISP) 261
- intermodulation 672
- international accounting rate 659
- international alphabet number 2 (IA2) 45, 206, 421
- international alphabet number 5 (IA5) 46, 196, 421, 729
- international code designator 527
- international exchange 81
- International Frequency Registration Board (IFRB) 623, 725
- international number 515, 520
- International Organization for Standardization (ISO) 15, 414, 724, 725
- international prefix 514
- international roaming 307
- international signalling 118
- international switching centre (ISC) 613
- International Telecommunications Union (ITU) 15, 19, 504, 723, 724, 725
- international transmission plan 612
- Internet 399, 402, 404, 405, 500, 587
- Internet access 533, 652
- Internet address classes 526
- Internet addressing 502, 525, 526
- Internet class A address 526
- Internet class B address 526
- Internet class C address 526
- Internet class D address 526
- Internet class E address 526
- Internet control message protocol (ICMP) 406, 410
- Internet Engineering Task Force (IETF) 489, 736

- internet header length (IHL) 408
- Internet numbering authority 526
- internet protocol (IP) 405, 406, 408, 526
- interpolation 696, 699, 702, 708
- interruption 289
- interworking 69, 491, 504, 510, 520
- interworking device 415
- interworking function (IWF) 510
- interworking unit (IWU) 510
- intra-LATA 812
- intrusion 80, 720
- Ionica 328
- ionosphere 157
- IP (intelligent peripheral) 234
- IP (internet protocol) 405, 406, 408, 526
- IP frame 408
- IPMS (inter-personal message service) 415, 419
- IPR (intellectual property rights) 510, 770, 806
- IPX 374, 376
- Iridium 171, 297, 314, 316
- IRP (interior routing protocol) 500
- ISC (international switching centre) 613
- ISDN (integrated services digital network) 206, 213, 249, 335, 739
- ISDN deployment 225
- ISDN in private network 227
- ISDN interworking 228
- ISDN signalling 227
- ISDN terminal 221, 225
- ISDN user part (ISUP) 250, 254
- ISI (inter-system interface) 432
- ISO (International Organization for Standardization) 15, 414, 724, 725
- ISO 2110 186, 188
- ISO 3166 527
- ISO 3309 198
- ISO 4903 186, 188
- ISO 7498 195
- ISO 8072 729
- ISO 8073 204
- ISO 8327 204
- ISO 8348 527
- ISO 8571 204
- ISO 8802 367
- ISO 8802.2 368, 374, 391, 435
- ISO 8802.3 368, 369
- ISO 8802.4 368, 371
- ISO 8802.5 368
- ISO 8802.7 374
- ISO 8823 204
- ISO 9000 774
- ISO 9735 788
- ISO management model 485
- isotropic radiation 158
- ISP (intermediate service part) 261
- ISPBX (integrated services private branch exchange) 228
- ISUP (integrated services user part) 224, 250, 254, 258
- IT (information technology) 779
- IT strategy 779
- itemized call record 685
- ITT (invitation to tender) 764
- ITU-R (ITU Radiocommunication bureau) 725
- ITU-Standardization Sector (ITU-T) 15
- ITU-T (ITU Standardization Sector) 725, 726
- ITU-T recommendations 723, 738
- ITU (International Telecommunications Union) 15, 19, 504, 723, 724, 725
- IVD (integrated voice data) LAN 394
- IWF (interworking function) 510
- IWU (interworking unit) 510
- J-bit 68
- jack 282, 287
- JDC (Japanese digital cellular system) 311
- JIT (just-in-time) 403, 745
- jitter 189, 453, 454, 499, 594, 641, 645, 671, 672
- joint 622
- Judge Green 811
- junction 11, 36, 81
- just-in-time (JIT) 403, 745
- justification 68, 268
- justification bit 68
- key 19
- key-coded encryption 715
- killer trunk 677
- Kingston communications (UK) 806
- LAN, wireless 432
- LAN-to-LAN connection 380
- LAN (local area network) 178, 367, 817
- LAN administrator 373
- LAN bridge 375, 376
- LAN emulation 470
- LAN emulation SSCS 470
- LAN gateway 377
- LAN hus 370, 373
- LAN operating system 374
- language assistance 289
- language digit 291

- LAP (link access procedure) 346, 348
- LAPB (link access procedure balanced) 347, 348
- LAPD (link access procedure in the D-channel) 218
- last mile connection 319
- LATA (local access and transport area) 506, 812
- Law organizing Telecommunications (LOT, Spain) 808
- layer 1 (physical layer) 199, 202, 465
- layer 2 (datalink layer) 198, 202
- layer 3 (network layer) 198, 202
- layer 3 protocol 350
- layer 4 (transport layer) 197, 202
- layer 5 (session layer) 197, 202
- layer 6 (presentation layer) 196, 202
- layer 7 (application layer) 195, 202, 402
- layer entity (LE) 467
- layer management 469
- layered protocol 201
- LCN (logical channel number) 199, 351
- LD-CELP (low delay CELP) 706
- LD (loop disconnect) signalling 25, 113
- LE (local exchange) 332
- lead time 564, 634, 753
- leadership priority 501
- leakance 29, 30, 31
- leased circuit (leaseline) 383, 657, 821, 828
- least difference regression 567, 568
- least significant bit (LSB) 387
- LEC (local exchange carrier) 506
- LED (light emitting diode) 55, 148
- LEO (low earth circular orbit) 316
- level 1 signalling (datalink) 254
- level 2 signalling (signalling link functions) 254
- level 3 signalling (signalling network functions) 255
- LGN (logical group node) 501
- LH (link header) 363
- licensing 799
- LIDB (line information database) 241
- lifetime 748
- lifting spring 97
- light emitting diode (LED) 55, 148
- limited availability 83, 84, 94
- line-of-sight (LOS) 19, 156, 164, 320, 627
- line code 71, 465
- line information database (LIDB) 241
- line jack 282
- line loading 29, 30, 31, 142, 548
- line signalling 119, 121, 122, 125, 126
- line station (LS) 432
- line termination equipment (LTE) 37
- line termination unit (LTU) 180, 188
- line up 599, 613, 667, 669
- linear regression 567
- linefinder 89, 91
- lineplant 746, 759
- link-by-link operation 126
- link access procedure (LAP) 346, 348
- link access procedure balanced (LAPB) 347, 348
- link access procedure in the D-channel (LAPD) 218
- link failure 587
- link header (LH) 363
- link trailer (LT) 363
- linked numbering scheme 95
- linked sale 801
- liquidated damages 770
- LLC (logical link control) 368, 374, 376, 406, 435
- LMI (local management interface) 389
- load sharing 255
- loading 142
- lobe diagram 158
- local access and transport area (LATA) 506, 812
- local area network (LAN) 178, 367, 817
- local battery 24
- local exchange (LE) 81, 240, 332
- local exchange carrier (LEC) 506
- local line network 619
- local line planning 618
- local mail protocol 527
- local management terminal 486
- local radio 161
- log on 714
- logical channel 341
- logical channel number (LCN) 199, 351
- logical communication channel (logical c-channel) 335
- logical group node (LGN) 501
- logical link control (LLC) 368, 374, 376, 406, 435
- logical network address 527
- logical unit (LU) 362
- long distance 507
- longhaul radio 624
- loop 25
- loop disconnect (LD) signalling 25, 113
- loopback 678
- looped bus topology 396
- Loran 316
- LOS (line-of-sight) 19, 156, 164, 320, 627
- loss 598
- lost calls 538, 543
- lost traffic 539
- loudness rating (LR) 598
- low orbit satellite 314

- low rate encoding (LRE) 696
- lower sideband 39, 40, 145
- LR (loudness rating) 598
- LRE (low rate encoding) 696
- LS (line station) 432
- LSB (least significant bit) 387
- LT (link trailer) 363
- LTE (line termination equipment) 37
- LTU (line termination unit) 180, 188
- LU (logical unit) 362
- LU0 365
- LU1 365
- LU2 365
- LU3 365
- LU4 365
- LU6 365
- LU6.2 366
- Lucent Technologies 328
- lumped loading 31
  
- M-ACTION command 483
- M-CANCEL-GET command 483
- M-CREATE command 483
- M-DELETE command 483
- M-EVENT-REPORT command 483
- M-GET command 483
- M-interface 472
- m-plane 461
- M-SET command 483
- M-wire 124, 125
- M.1020, ITU-T recommendation 828
- M.1025, ITU-T recommendation 828
- M.1040, ITU-T recommendation 828
- M.3010, ITU-T recommendation 481
- M1-interface 473
- M2-interface 473
- M3-interface 473
- M4-interface 473
- M5-interface 473
- MAC (medium access control) 369, 374, 392, 395, 435
- magnetic induction 607
- magnetic tape 177
- magneto-generator signalling 110, 281, 284
- MAHO (mobile assisted handover) 325
- mail server 375
- maintenance 663, 764
- maintenance access point 675
- maintenance case 623
- maintenance limit 674
- maintenance organization 665
- maintenance philosophy 663
- maintenance plan 14
- maintenance workstation 668
- maintenance, controlled 664
- maintenance, corrective 664
- maintenance, preventive 664
- MAN (metropolitan area network) 391
- managed network 799
- managed object (MO) 483
- management domain 419, 420
- management function model 486
- management information base (MIB) 484, 489
- management information system (MIS) 779
- management model 477, 485
- management plane 136, 459, 469
- management plane communication type 1 459
- management plane communication type 2 459
- management plane communication type 3 459
- manager 482
- Manchester code 72
- manual exchange 286
- manual hold 289
- manual keying 20
- manual operation 281
- manual service 281
- manual telephone 283
- manufacturing 787
- manufacturing automation protocol (MAP) 204, 731, 787
- MAP (mobile application part) 244, 250, 263, 305
- mapping 271
- marginal cost 660
- marginal cost, average long term 660
- MARISAT 312
- mark 20, 44, 148
- marker control 96
- mast antenna 158, 162
- master frame generator 394
- mastergroup 37
- Matra 316
- maximum bits in transit 380
- maximum hop count 493, 497
- maximum line length 594, 606, 719
- maximum number of digits 515
- MBS code (radiopaging) 428
- MCHO (mobile controlled handover) 325
- MCI (Microwave Communications Inc.) 506, 794, 811, 837
- MD (mediation device) 481
- mechanical wear 664
- media industry 792



- mediation device (MD) 481
- medium access control (MAC) 369, 374, 392, 395, 435
- medium access control layer, DECT 326
- MEO (medium earth circular orbit) 316
- Mercury Communications Limitef 506, 799, 806
- message format 199
- message handling system (MHS) 402, 413, 524
- message sequence number 256
- message service 442
- message switching 12, 719
- message transfer agent (MTA) 414, 420
- message transfer agent entity (MTAE) 418, 422
- message transfer layer (MTL) 417, 418
- message transfer part (MTP) 224, 250, 254, 470
- message transfer protocol (P1) 418, 420
- message transfer service (MTS) 415
- meta-signalling 461
- meta-signalling virtual channel (MSVC) 461
- metering 123
- metrid 501
- metropolitan area network (MAN) 391
- Metropolitan Fibre Systems (MFS) 330
- MF (multi-frequency) tone signalling 25, 26, 113
- MF4 signalling 113
- MFC (multi-frequency code) signalling 118, 120, 127
- MFS (Metropolitan Fibre Systems) 330
- MH environment 415
- MHS (message handling system) 402, 413, 524
- MHS model 414
- MIB (management information base) 484, 489
- microphone 22
- Microsoft Corporation 402
- Microsoft DOS 375
- Microsoft Excel 401
- Microsoft Mail 401
- Microsoft Network (MSN) 402, 411
- Microsoft Office 401
- Microsoft Powerpoint 401
- Microsoft Project 766
- Microsoft Word 49, 401
- microwave radio 19, 153, 163, 626
- microwave radio terminal 626
- minimum monitoring period 532
- Ministero delle Poste e Telecomunicazioni (MPT, Italy) 807
- Ministry of Posts and Telecommunications (Japan) 812
- Ministry of Research (Denmark) 807
- Ministry of Transport and Communications (Finland) 807
- Ministry of Transport and Communications (Greece) 807
- Ministry of Transport and Communications (Norway) 807
- Ministry of Transport, Public Works and Water Management (Netherlands) 807
- Minitel 399
- mis-interpretation 711
- mis-sequence 259
- MIS (management information system) 779
- MLP (multiple link procedure) 348, 360
- MML (man machine language) 739
- MO (managed object) 483
- mobile application part (MAP) 244, 250, 263, 305
- mobile assisted handover (MAHO) 325
- mobile controlled handover (MCHO) 325
- mobile data network 425, 429
- mobile digital trunk radio system (MDTRS) 431
- mobile relay station (MRS) 325, 326
- mobile station (MS) 432
- mobile switching centre (MSC) 244, 297, 300, 308
- mobile telephone 310
- mobile telephone network 625
- mobile telephony 297
- mobile unit 300
- Mobitex 430
- Modacom 430
- modem 180, 181, 182
- modem constellation 182, 185, 672
- modernization 751, 760
- modified final judgement (MFJ) 811
- modulation 39
- monitoring 14
- monomode fibre 150, 151
- monopolies commission 798
- monopoly 793, 795, 805
- monopoly operator 508
- Morse code 19, 45, 46
- Morse, Samuel 19
- most significant bit (MSB) 48, 387, 465
- Motorola 316, 430
- MPEG (motion pictures experts group) 441
- MRS (mobile relay station) 325, 326
- MRVT (MTP routing verification test) 262, 263
- MS (mobile station) 432
- MSB (most significant bit) 48, 387, 465
- MSC (mobile switching centre) 244, 297, 300, 308
- MSOH (multiplex section overhead) 272
- MSVC (meta-signalling virtual channel) 461
- MTA (message transfer agent) 414, 420
- MTAE (message transfer agent entity) 418, 422

- MTBF (mean time between failures) 773  
MTL (message transfer layer) 417, 418  
MTP (message transfer part) 224, 250, 254, 470  
MTP routing verification test (MRVT) 262, 263  
MTP3 (message transfer protocol level 3) 471  
MTS (message transfer service) 415  
Mu-law code 61, 64  
multi-frequency (MF) tone signalling 25, 26, 113  
multi-frequency code (MFC) signalling 118, 120, 127  
multi-level transmission 182  
multi-national company (MNC) 815  
multicast 446  
multiframe indicator 276  
multimedia 437, 451  
multimedia workstation 440  
multimode fibre 149, 150  
multipath 165, 166, 309, 320, 435, 627, 628  
multiple access 147, 172, 173, 174, 309, 435, 629, 631  
multiple link procedure (MLP) 348, 360  
multiplex section overhead (MSOH) 272  
multiplexing 29, 37, 57, 61, 62, 144, 145, 268, 269, 271, 456, 457, 463, 609, 696, 698  
multiplexor 63, 698, 701  
multiplexor, ATM 456, 457, 463  
multiplication gain 701  
multipoint 448, 449  
Murray code 20  
  
 $n \times 64$  kbit/s 229, 442  
N:1 restoration 278  
N-ISDN (narrowband ISDN) 229, 441  
N1 parameter 350  
N2 parameter 350  
NAE (network address extension) 520  
NAMTS (Nippon Automatic Mobile Telephone System) 306  
narrowband ISDN (N-ISDN) 229, 441  
national (data) number (NN) 521  
national regulatory authority (NRA) 806  
national significant number (NSN) 514  
National Telecom Agency (Denmark) 807  
National Telecommunications Committee (Greece) 807  
natural monopoly 795  
NAU (network addressable unit) 364  
NAU services 362  
Navstar 436  
NCP (network control program) 362, 364  
NE (network element) 481  
NEL (network element layer) 486  
NEM (network element manager) 480  
NET (Normes Européennes de Télécommunications) 728, 732, 733  
NET3 218  
NET5 222  
Netherlands PTT 837  
Netro Corporation 626  
network access control 712, 713  
network address 496, 500  
network address extension (NAE) 520  
network address, ATM 527  
network address, logical 527  
network addressable unit (NAU) 364  
network administration 752  
network and switching sub-system (NSS) 308  
network C 306  
network capacity 684, 743  
network configuration 771  
network control program (NCP) 362, 364  
network design 588, 589  
network element (NE) 481  
network element layer (NEL) 486  
network element manager (NEM) 480  
network evolution 750  
network identification code (NIC) 521  
network layer (layer 3) 198, 202  
network loss 598  
network management controls 687  
Network Management Forum (NMF) 487, 735  
network management layer (NML) 486  
network management model, ATM Forum 472  
network management model, ATM Forum 472  
network management system (NMS) 477, 693  
network monitoring 684  
network network interface (NNI) 137, 358, 505  
network node (NN) 366  
network node interface (NNI) 137  
network overlay 761  
network overload 576  
network performance (NP) 633, 636, 637, 639  
network performance parameters, generic 640  
network reference model, ATM Forum 471  
network restoration 688  
network routing plan 491  
network service access point (NSAP) 527  
network synchronization 673  
network terminal number (NTN) 521  
network terminating equipment 1 (NT1) 216, 217  
network terminating equipment 1, broadband (B-NT1) 446

- network terminating equipment 2 (NT2) 216, 217
- network terminating unit (NTU) 188
- network topology state 500
- network transaction time 562
- network transmission plan 593
- network transmission plan, North America 600
- network transmission plan, UK digital network 601
- network usage 555
- network user identification (NUI) 354, 716
- NFS (network file server) 406, 407, 408, 424
- NIC (network identification code) 521
- NICC (Network Interconnection Coordination Committee) 733
- NMF (Network Management Forum) 487, 735
- NML (network management layer) 486
- NMS (network management system) 477, 693
- NMT-450 306
- NMT-900 306
- NMT (Nordic Mobile Telephone Service) 306
- NN (national (data) number) 521
- NN (network node) 366
- NNI (network network interface) 137, 358, 505
- NNI (network node interface) 137
- NNI, ATM 456, 457, 464, 470
- NNI, frame relay 386
- noise 56, 593, 594, 607, 616, 670
- noise meter 670
- non-adjacent cell 300
- non-blocking 81
- non-compelled signalling 126
- non-discrimination 804, 805
- non-pre-emptive priority multiple access (NPMA) 435
- non return to zero (NRZ) code 72, 190
- non return to zero inverted (NRZI) code 72, 191
- normalized state 484
- Nortel 328, 347, 386, 735
- North America dial plan (NADP) 515, 519
- North American telephone network transmission plan 600
- Northern Telecom 328, 347, 386, 735
- NOSFER (nouveau systeme fondamental pour la determination des equivalents de reference) 595
- notched noise 671
- noticeboard 404
- Novell Netware 375
- NP (network performance) 633, 636, 637, 639
- NPA (number plan of America) 515, 519
- NPMA (non-pre-emptive priority multiple access) 435
- NPSI (network packet switching interface) 360
- NPV (net present value) 751
- NRA (national regulatory authority) 806
- NRZ (non return to zero) code 72, 190
- NRZI (non return to zero inverted) code 72, 191
- NSAP (network service access point) 527
- NSN (national significant number) 514
- NSS (network and switching sub-system) 308
- NT1 (network terminating equipment 1) 216, 217
- NT2 (network terminating equipment 2) 216, 217
- NTN (network terminal number) 521
- NTT code (radiopaging) 428
- NTT Law, 1985 (Japan) 812
- NTU (network terminating unit) 188
- NUI (network user identification) 354, 716
- null layer 204
- nulltariff service 237
- number plan of America (NPA) 515, 519
- number portability 507
- number translation 115
- number unobtainable 118
- numbering plan 14, 513
- NYNEX 811
- Nyquist criterion 59
- O/R (originator/recipient) name 421
- O-series, ITU recommendations 676
- OAM (operations, administration and management) 642
- OAM cell (operation and maintenance cell) 466, 467
- obsolescence 748
- OC-1 (SONET) 71, 280
- OC-12 (SONET) 71, 280
- OC-192 (SONET) 71, 280
- OC-3 (SONET) 71, 280
- OC-48 (SONET) 71, 280
- occupancy 537, 538
- ocean country code 313
- odd parity 200
- O'Dell grading 86
- Odette 404
- Odyssey 316
- off-hook signal 112, 115
- offered traffic 531, 537, 544, 545, 554
- office computer network 815, 819
- Ofitel (Office of Telecommunications) 506, 509, 733, 398, 808
- OLO (other licensed operator) 509
- OLR (overall loudness rating) 598
- OMAP (operation and maintenance application part) 250, 263

- OMC (operations and maintenance centre) 308
- omnidirectional 158
- OMNIpoint 487, 735
- on-hook 25
- one-for-many replacement 761
- one-for-one replacement 761
- One-to-One 311
- online database service 410, 443
- ONP (open network provision) 803, 806
- opal 288
- OPAL (optical access line) 331, 333
- OPC (originating point code) 257
- open network provision (ONP) 803, 806
- open shortest path first (OSPF) 407, 586
- open systems interconnection (OSI) model 194, 203, 253, 414, 465, 724
- operating limits 674
- operating system (OS) 481
- operation and maintenance application part (OMAP) 250, 263
- operations and maintenance centre (OMC) 308
- operations and support sub-system (OSS) 308
- operator assistance 281, 289, 508, 656
- operator privileges 288
- operator switchboard 281, 285, 293
- optical carrier (OC, SONET) 71, 280
- optical fibre 19, 148, 466
- optical line terminating unit (OLTU) 330
- Orange 311
- ORB (object request broker) 488
- order-taking 786
- order of transmission 48
- ordering lead time 564
- orderwire 277
- ORE (overall reference equivalent) 596, 601
- organization name 419
- Orientation of Telecommunications Policy, 1989 (Sweden) 808
- originating point code (OPC) 257
- originator/recipient (O/R) name 421
- OS (operating system) 481
- oscillator 154
- OSI (open systems interconnection) model 194, 203, 253, 414, 465, 724
- OSI directory service 421
- OSI protocols 423
- OSPF (open shortest path first) 407, 586
- OSS (operations and support sub-system) 308
- other licensed operator (OLO) 509
- out-of-band 122
- outband signalling 121, 122
- outgoing circuit 77
- outgoing route 114
- outgoing traffic 757
- outlet spring 97
- output buffer 455
- over-provision 743
- over-the-horizon radio 156
- overall loudness rating (OLR) 598
- overall reference equivalent (ORE) 596, 601
- overflow 66, 384, 562, 547
- overflow routing 542, 577, 584, 825
- overheads 659
- overheating 664
- overlap signalling 503
- overlay 761
- overload 576, 683
- overload control 445
- p-wire 79
- P1-interface 402
- P1-protocol (message transfer protocol) 418, 420
- P2-protocol (inter-personal messaging protocol) 419
- P3-interface 402
- P3-protocol (submission and delivery protocol) 418
- Pacific Bell 811
- Pacific Ocean region (POR) 629, 630
- pacing 692
- packet 13, 347
- packet delay 547
- packet error ratio 641
- packet handler (PH) 220
- packet identifier 352
- packet level interface 346, 350
- packet level protocol 521
- packet mode 220
- packet negotiation 354
- packet size 354, 380, 547
- packet switched exchange (PSE) 207, 360
- packet switching 12, 106, 207, 208, 259, 341, 379, 453
- packetized speech 444
- pad (attenuator) 599
- PAD (packet assembler/disassembler) 220, 355, 389, 430
- padding 396
- paging access control equipment (PACE) 427
- paging receiver 427
- paging transmitter 427
- paid minute 560, 654
- paid time 558, 559

- PAMA (pre-assigned multiple access) 172
- panel switching 88
- paper tape 177
- paper ticket 656
- paperless office 424
- parallel copper pair wire 19
- parallel transmission 187
- parity 357
- parity bit 200
- parity code 276
- partial availability 83
- passive bus 218
- passive optical networks (PONs) 151, 331
- password control system 714
- patch panel 817
- patching 371
- patent rights 770
- path-oriented routing 209, 344, 586
- path control 362
- path obstruction 627
- path overhead (POH) 271, 276
- path protection 712, 714
- path status 276
- path trace 276
- path user channel 276
- pay-on-view television 443
- payback period 660
- payload type (PT) 464
- payload type identifier (PTI) 464
- payphone calls 657
- PBF (proportionate bidding facility) 585
- PBX (private branch exchange) 81, 239
- PBX maintainer 819
- Pc-protocol 417
- PCA (percentage of calls answered) 288, 544
- PCA15 288
- PCA25 288
- PCI (protocol control information) 199, 200, 341, 380, 467
- PCM (pulse code modulation) 57, 70, 310, 673, 703, 704
- PCN (personal communications network) 310
- PDC (Japanese personal digital cellular system) 311
- PDD (post dialling-delay) 25
- PDH (plesiochronous digital hierarchy) 69, 267
- PDU (protocol data unit) 374, 467
- peer-to-peer communication 468
- peer group (PG) 501
- peer group leader (PGL) 501
- peg count 562
- percentage of calls answered (PCA) 288, 544
- perception of quality 635
- performance management 485, 633
- performance objective 674
- performance parameters 640, 641
- performance parameters, derived 641
- permanent virtual circuit (PVC) 351, 383, 457, 656
- persistence 560
- personal call 286, 290
- personal communications network (PCN) 310
- personal identification number (PIN) 236
- personal name 419
- PG (peer group) 501
- PGL (peer group leader) 501
- PH (packet handler) 220
- phase 183
- phase-locked loop (PLL) 672
- phase angle 183
- phase hits 671, 672
- phase jitter 672
- phase modulation 181
- phase shift keying (PSK) 153, 181, 310
- physical communication channel (physical c-channel) 336
- physical layer (layer 1) 199, 202, 465
- physical layer entity (PL-LE) 467
- physical layer protocol (PHY) 392
- physical media dependent (PMD) 392
- physical medium 17
- physical medium (PM) 17, 465
- physical medium layer 468
- physical packet P00 326
- physical packet P08 326
- physical unit (PU) 362
- picking list 785
- PICS (protocol implementation conformance statement) 724
- pillar 619
- pilot signal 676
- PIN (personal identification number) 236, 714
- PING (packet Internet groper) 408
- pitch 21
- pixel 51, 441
- PIXIT (protocol information extra information for testing) 724
- PL-cell (physical layer cell) 467
- PL-SAP (physical layer service access point) 466
- PL-SDU (physical layer service data unit) 466, 467
- place of sale 834
- plane management 469
- plausibility check 504, 717
- plesiochronous 68

- plesiochronous digital hierarchy (PDH) 69, 267
- PLL (phase-locked loop) 672
- plug 287
- PM (physical medium) 17, 465
- PMD (physical media dependent) 392
- PMP (point-to-multipoint) radio 157, 320, 328, 624
- PMR (private mobile radio) 299
- PMUX (primary multiplexor) 63, 698
- PN.525-2, ITU-R recommendation 625
- PN.530-5, ITU-R recommendation 628
- PN.837-1, ITU-R recommendation 627
- PNNI (private network network interface) 471
- PNNI (private network node interface) 471, 500
- POCSAG code (radiopaging) 428
- POH (path overhead) 271, 276
- point-to-multipoint (PMP) radio 157, 320, 328, 624
- point-to-point (PTP) radio 157, 320, 624
- point-to-point data service 431
- point-to-point protocol (PPP) 406, 407
- point of interconnection (POI) 508
- Pointel 322
- Poisson distribution 550
- Pollaczek–Khinchine delay formula 550
- PONs (passive optical networks) 151, 331
- POP (point of presence) 507
- portability 802
- portable part (PP) 324
- portable radio terminal (PT) 324, 325
- portfolio 832
- Post and Telecommunications Authority (Sweden) 808
- post dialling-delay (PDD) 25
- post office 401
- Post Office Telecommunications (UK) 737
- Post reform 1 (Germany) 805
- PP (portable part) 324
- PPP (point-to-point protocol) 406, 407
- PRA (primary rate access) 215, 222
- pre-assigned multiple access (PAMA) 172
- predictive algorithm 705
- premature release ratio 641
- premises cabling 816
- premium rate service 236
- Prentice Hall 472
- present worth 751
- presentation control 443
- presentation layer (layer 6) 196, 202
- presentation protocol 204
- press to speak 429
- pressurized cables 616
- Prestel 399, 789
- preventive maintenance 664
- previous satellite link 496
- PRI (primary rate interface) 215, 222
- price control 800
- price determination 833
- primary high usage route 578
- primary multiplexor (PMUX) 63, 698
- primary rate access (PRA) 215, 222
- primary rate interface (PRI) 215, 222
- primitive name 419
- prioritisation 581
- priority 688
- priority resolution 435
- private ATM network 473
- private branch exchange (PBX) 81, 239
- private ISDN 227
- private line 10
- private local interface, ATM 471
- private management domain (PRMD) 420
- private mobile radio (PMR) 299
- private network 815, 820
- private network management system 473
- private network network interface (PNNI) 471
- private network node interface (PNNI) 471, 500
- private telephone network 805
- private trunk mobile radio (PTMR) 430
- private UNI (ATM) 471, 473
- PRM (protocol reference model) 326, 469
- PRMD (private management domain) 420
- probability of blocking 538, 554
- probability of delay 545, 547
- probability of errors 563
- probability of lost calls 545, 546
- probability of packet delay 547
- proceed-to-send (PTS) 125
- procurement 763, 765
- product definition 831
- product development 834
- product manual 834, 835
- profit 659
- profit margin 661
- project management 766
- promotion 834
- propagation delay 381, 444, 445, 452, 453, 496, 501, 547, 606, 629, 634, 641
- propagation time 381, 397, 452, 453, 547, 562, 594
- proportion of lost calls 538, 554
- proportionate bidding facility (PBF) 585
- proprietary standard 723
- protection (restoration network) 270, 689
- protection (security measure) 712

- protection network 689
- protection, 1 + 1 270
- protocol 7, 177, 193, 564
- protocol control information (PCI) 199, 200, 341, 380, 467
- protocol data unit (PDU) 374, 467
- protocol multiplexing 198
- protocol reference model (PRM), ATM 469
- protocol reference model (PRM), B-ISDN 469
- protocol reference model (PRM), DECT 326
- protocol stack 194
- protocol stack, ATM 468
- protocol testing 670
- provisioning 479, 615, 643
- Proximity i 328
- PSE (packet switched exchange) 207, 360
- pseudo-random bit pattern generator 673
- PSK (phase shift keying) 153, 181, 310
- psophometric noise 607, 671
- psophometric units 607
- PSTN (public switched telephone network) 116, 213, 297
- PT (payload type) 464
- PT (portable radio terminal) 324
- PTI (payload type identifier) 464
- PTMR (private trunk mobile radio) 430
- PTO (public telecommunications operator) 794, 800, 803, 831
- PTP (point-to-point) radio 157, 320, 624
- PTS (proceed-to-send) 125
- PU (physical unit) 362
- PU1 362
- PU2 362
- PU2.1 362, 366
- PU3 362
- PU4 362
- PU5 362
- public ATM network 473
- public call box 806
- public circuit switched digital data network (PSDN) 206
- public data network address scheme 521
- public network 831
- public network interface, ATM 471
- public network management system 473
- public switched telephone network (PSTN) 116, 213, 297
- public telecommunications operator (PTO) 794, 800, 803, 831
- public telephone switching monopoly 805
- public UNI (ATM) 471, 473
- public utilities commission (United States) 812
- public utility 793, 795
- pulse code modulation (PCM) 57, 70, 310, 673, 703, 704
- pulse metering 649, 650, 651
- punched card 177
- PVC (permanent virtual circuit) 351, 383, 457, 656
- Q-adaptor (QA) 481
- Q-adaptor function 482
- Q-interface 228
- Q-SIG signalling 228, 823
- Q.50, ITU-T recommendation 709
- Q.400, ITU-T recommendation 131
- Q.421, ITU-T recommendation 132
- Q.490, ITU-T recommendation 131
- Q.701, ITU-T recommendation 254
- Q.707, ITU-T recommendation 254
- Q.711, ITU-T recommendation 261
- Q.714, ITU-T recommendation 261
- Q.721, ITU-T recommendation 257
- Q.725, ITU-T recommendation 257
- Q.761, ITU-T recommendation 224, 258
- Q.764, ITU-T recommendation 224, 258
- Q.771, ITU-T recommendation 233, 263
- Q.775, ITU-T recommendation 233, 263
- Q.921, ITU-T recommendation 218
- Q.922, ITU-T recommendation 221, 388
- Q.931, ITU-T recommendation 219, 227, 260
- Q.933, ITU-T recommendation 388
- Q.2110, ITU-T recommendation 470, 471
- Q.2130, ITU-T recommendation 470, 471
- Q.2140, ITU-T recommendation 470, 471
- Q.2931, ITU-T recommendation 461, 470, 472
- q.d.u. (quantization distortion unit) 61, 611, 710
- Q3-interface 477, 481, 483
- QA (Q-adaptor) 481
- QA (quality assurance) 774
- QAM (quadrature amplitude modulation) 153, 183
- QOS (quality of service) 633, 636, 637, 639, 739
- QPSX Communications Ltd. 394
- quadrature amplitude modulation (QAM) 153, 183
- Qualcom 316
- quality assurance (QA) 774
- quality inspection 769
- quality of service (QOS) 633, 636, 637, 639, 739
- quantization 59, 60, 673, 703
- quantization distortion 594, 609, 611
- quantization distortion unit (q.d.u.) 61, 611, 710
- quantization level 59, 61
- quantization limit 600

- quantization noise 59, 61
- quantizing distortion 611, 673
- quasi-associated mode signalling 252
- queue 543
- queue capacity 543
- queueing delay 343
- queueing system 543
- queueing time 643
- Qx-interface 481
  
- R-interface 217
- R-reference point 449
- R2 signalling 118, 119, 121
- radio 152
- radio attenuation 625
- radio fixed part (RFP) 325
- radio frequency (RF) 154
- radio frequency allocations 164
- radio in the local loop 319
- radio regulations 809
- radio relay station 325
- radio repeater 158
- radio route planning 165
- radio shadow 320
- radio telephone service 297
- radio transmission 623
- radio, longhaul 624
- radio, shorthaul 624
- Radiocom 2000 306
- Radiocommunication bureau (BR) 725
- radiodetermination satellite services (RDSS) 436
- radiopager 163, 425, 426
- radiopaging 425
- rainfall statistics 627
- ranking 582
- RARP (reverse address resolution protocol) 406
- rate adaption 465
- RBOC (Regional Bell Operating Company) 231, 737, 811
- RD (route destination) accounting 654
- RD (routing domain) 528
- RDSS (radiodetermination satellite services) 436
- re-attempts 553
- re-boot 680
- re-sequencing 344, 587
- re-use pattern 300
- reachable address 502
- reachable address, exterior (ERA) 502
- real-time network monitoring 685
- reassembly 259, 445
- receive reference equivalent (RRE) 594, 601
- receive signal level (RSL) 626
- received signal strength indicator (RSSI) 325
- receiver 3
- receiver not ready (RNR) 349, 387
- receiver ready (RR) 349, 387
- recommendations of ITU-T 723, 738
- reconstituted signal 58
- recorded announcement 118
- rectifying circuit 155
- red book 739, 740
- redundant ring topologies 277
- reed relay switching 100
- Reed Solomon code 201
- reflection 156
- refraction 156
- regeneration 56, 127, 609
- regenerator 272
- regenerator section (RS) 466
- regenerator section overhead (RSOH) 272
- register 112, 128
- register/translator 95
- registration 305
- regression testing 680
- regulation 793, 797
- regulation, European Union 802
- regulation, United Kingdom 806
- regulation, United States of America 809
- regulator 797, 806
- regulatory body 797, 806
- reject (REJ) 349
- relaying envelope 416
- release 80
- release failure ratio 641
- reliability 644, 745, 773
- reliable data delivery SSCS 470
- reliable transfer 414, 470
- remote alarm 276
- remote bridge 376
- remote job entry 365
- remote switching unit (RSU) 336
- repair 634
- repeat attempt 683
- repeater 29, 32, 35
- request response unit (RRU) 363
- resale 240
- reset 680
- resistance 29, 30, 31
- resource management 613, 634
- response header (RH) 363
- response time 382, 562, 819
- responsiveness 634



- restart 680
- restoration 479, 644, 688
- restoration network 689
- restoration plant 688
- restoration, 1 for N 688
- restoration, 1:1 278
- restoration, N:1 278
- restrictive control action 687, 691
- restoration, 1 for N 688
- retail trade 788
- retransmission 380, 384
- retrieval service 442
- return to zero (RZ) code 72, 191
- revenue 659, 833
- reverse address resolution protocol (RARP) 406
- reverse charging 286, 290, 354, 508, 656, 720
- reverse multiplexing 230
- RF (radio frequency) 154
- RFC (request for comment) 736
- RFC 1157 489
- RFC 1267 500
- RFC 827 500
- RFC 888 500
- RFC 891 500
- RFC 904 500
- RFP (radio fixed part) 325
- RFP (request for proposal) 764
- RFQ (request for quotation) 764
- RH (response header) 363
- right of way 806
- ring back when free 215, 226, 239
- ring topology 373, 691
- ringdown signalling 109
- ringing 26
- ringing current 109
- ringing tone 561
- RIP (routing information protocol) 407, 500
- riser 817
- RJ-45 socket 817
- RNR (receiver not ready) 349, 387
- roaming 305, 307, 425
- robbed bit signalling 69, 74
- rock armour 618
- rotary switching 88
- route balancing 585
- route busy hour 530, 572, 758
- route destination (RD) accounting 654
- router 375, 376
- routine testing 676
- routing 344, 421, 491, 499, 502
- routing algorithm 571, 606
- routing domain (RD) 528
- routing field 464
- routing information protocol (RIP) 407, 500
- routing plan 14, 491
- routing protocol 499
- routing table 114
- routing table administration 497, 498
- routing 676
- RPC (remote procedure call) 406, 407, 408
- RR (receiver ready) 349, 387
- RRE (receive reference equivalent) 594, 601
- RRU (request response unit) 363
- RS (regenerator section) 466
- RS232 interface 186, 188, 193, 199, 730
- RS422 188
- RS423 188
- RS449 188, 193, 730
- RSL (receive signal level) 626
- RSOH (regenerator section overhead) 272
- RSSI (received signal strength indicator) 325
- RSU (remote switching unit) 336
- running costs 748, 751
- rural service 806
- RZ (return to zero) code 72, 191
- S/N (signal to noise ratio) 599, 607
- S/T-interface 216, 221
- S-curve 569
- S-field 327
- S-format 349
- S-interface 216, 217
- S1-system 422
- S2-system 422
- S3-system 423
- SAAL (signalling AAL) 471
- SABM (set asynchronous balanced mode) 349
- safety 774, 804
- sampling 57, 531, 556
- sampling frequency 703
- sampling rate 58, 190, 532, 703
- SAP (service access point) 374, 466
- SAR (segmentation and reassembly) 259, 465, 468
- SAS (single attached station) 392
- satellite 19, 153, 167, 628
- Satellite Information Services (SIS) 790
- Sb-interface 447, 449
- SC-connector 151
- scan rate 532
- scatter angle 167
- SCCP (signalling connection and control part) 233, 250, 254, 258

- SCCP classes 259
- SCCP routing verification test (SRVT) 262, 264
- SCE (service creation environment) 232, 234
- SCF (service control function) 233
- SCP (service control point) 231, 232, 233
- SCPC (single channel per carrier) 631
- scrambler 822
- scrambling 713
- screened cable 35, 616
- SDE (submission and delivery entity) 418, 422
- SDH (synchronous digital hierarchy) 70, 267, 691
- SDL (specification and description language) 739
- SDLC (synchronous datalink control) 191, 205, 362, 365
- SDLC adapter 365
- SDLC gateway 365
- SDR (special drawing rights) 654
- SDU (service data unit) 466, 467
- SDU (service data unit) 467
- SE (service element) 195
- second dial tone 520
- second line support 665
- Secretary of State (UK) 806
- section overhead (SOH) 271, 276
- security 409, 711, 804
- security management 485
- segment 395, 556, 655
- segment header 395
- segment, TCP 410
- segmentation 259
- segmentation and reassembly (SAR) 259, 465, 468
- seizure 123, 561
- select bar 99
- select finger 98
- select magnet 98
- selecting equipment 763
- selector 85
- self-provision licence (UK) 808
- self-regulating 797
- self-routing network 551
- self healing network 694
- semantics 417
- semi-automatic accounting rate 656
- semi-automatic telephony 282
- send reference equivalent (SRE) 594, 601
- sequence number 200
- serial line internet protocol (SLIP) 406, 407
- serial transmission 187
- server 375, 401, 543
- service-independent building block (SIB or SIBB) 231
- service access point (SAP) 374, 466
- service availability 643
- service control function (SCF) 233
- service control point (SCP) 231, 232, 233
- service creation environment (SCE) 232, 234
- service data unit (SDU) 466, 467
- service element (SE) 195
- service management layer (SML) 486, 487
- service management system (SMS) 232, 234
- service profile identifier (SPID) 461
- service provider 831
- service script 234
- service specific connection-oriented protocol (SSCOP) 470
- service specific convergence sublayer (SSCS) 465
- service specific coordination function (SSCF) 470
- service switching function (SSF) 233, 234
- service switching point (SSP) 231, 232, 234
- service time 544
- service time, average 544, 545
- session 414
- session layer (layer 5) 197, 202
- session protocol 204
- set up 109, 110, 125
- settlement 653
- settlement, cascade 653
- settlement, destination 654
- settlement, simple 653
- severely errored cell block ratio (SECBR) 641
- severely errored frame ratio 641
- shared duct 507
- shield 144
- shielded cable 616
- shielded twisted pair (STP) cable 143, 817
- ship's antenna 314
- shorthaul radio 321, 624
- SIB (service-independent building block) 231
- SIBB (service-independent building block) 231
- sideband 39
- sidetone 23, 594, 602
- Siemens Cornet signalling 229
- Siemens Nixdorf 361
- signal duration 381
- signal label 276
- signal power 157
- signal strength 23
- signal to noise ratio (S/N) 599, 607
- signal transfer point (STP) 252
- signalling 739
- signalling AAL (SAAL) 471

- signalling connection and control part (SCCP) 233, 250, 254, 258
- signalling information field 256
- signalling interworking 132
- signalling link 251
- signalling network planning 265
- signalling point (SP) 252, 461
- signalling sequence diagram 133, 134
- signalling system number 7 249
- signalling terminal (ST) 220, 251
- signalling terminal (STE) 360
- signalling transfer point (STP) 461
- signalling virtual channel (SVC) 460, 461
- SIM card 714
- simple mail transfer protocol (SMTP) 401, 405, 406, 407, 408
- simple network management protocol (SNMP) 406, 407, 408, 477, 489
- simple settlement 653
- simplex operation 7, 155, 184
- single attached station (SAS) 392
- single channel per carrier (SCPC) 631
- single European market (SEM) 838
- single link procedure (SLP) 348, 360
- single sideband (SSB) 39, 153
- sky-wave path 156
- slave frame generator 394
- slip 66, 268
- slip-pattern 85
- SLIP (serial line internet protocol) 406, 407
- slot 395
- slot header 395
- SLP (single link procedure) 348, 360
- SMDS (switched multimegabit digital service) 394, 472, 655
- SMDS SSCS 470
- SME (Society of Manufacturing Engineers) 731, 787
- SML (service management layer) 486, 487
- SMS (service management system) 232, 234
- SMT (station management) 392
- SMTP (simple mail transfer protocol) 401, 405, 406, 407, 408
- SMTP address 403, 527
- SNA (systems network architecture) 205, 341, 361, 724, 738
- SNAP (subnetwork access protocol) 406, 407
- SNI (subscriber network interface) 137, 332, 397
- SNMP (simple network management protocol) 406, 407, 408, 477, 489
- So-bus 218
- Society of Manufacturing Engineers (SME) 731, 787
- software bug 679
- software faults 679
- software installation 680
- software licence 764
- software patch 681
- software problems 666
- software release 680
- software upgrade 681
- software virus 718
- SOH (section overhead) 271, 276
- SONET (synchronous optical network) 70, 267, 279, 691
- source and record route 409
- source routing 500
- source service access point (SSAP) 374
- Southwestern Bell (SBC) 811
- SP (signalling point) 252, 461
- SP (space) 20, 44, 46, 148
- space (SP) 20, 44, 46, 148
- space switching 103, 105
- spare parts 666
- SPC (stored program control) 80, 100, 532
- special drawing rights (SDR) 654
- specification 754, 764, 769, 771
- specification and description language (SDL) 739
- spectrum analyzer 672
- spectrum utilization 163
- speech compression 607
- speech interpolation 696, 699, 702, 708
- speech, packetized 444
- speed 640
- SPID (service profile identifier) 461
- SPIRIT 487
- splitting 198
- sporadic growth 569
- spot beam 629
- Sprint 506, 735, 811, 837
- SRE (send reference equivalent) 594, 601
- SRVT (SCCP routing verification test) 262, 264
- SS7 signalling 249, 461
- SSAP (source service access point) 374
- SSB (single sideband) 39, 153
- SSCF-NNI 470
- SSCF-UNI 470
- SSCF (service specific coordination function) 470
- SSCOP (service specific connection-oriented protocol) 470
- SSCP (system services control point) 362, 364
- SSCS (service specific convergence sublayer) 465

- SSF (service switching function) 233, 234
- SSP (service switching point) 231, 232, 234
- ST-connector 151
- ST (signalling terminal) 220, 251
- stability 594
- standards 723
- standby c-channel 336
- star 113
- start bit 191
- Statens Teleforvaltning (STF, Norway) 807
- station call 289
- station management (SMT) 392
- statistical equilibrium 552
- statistical multiplexing 13, 342, 383, 444, 452, 546, 696, 699
- statistical multiplexor 452, 701, 707
- status report 419
- STE (signalling terminal) 360
- STE (supergroup translating equipment) 38
- stealing bits 68
- step-by-step switching 88
- step index multimode fibre 149, 150
- stimulation of traffic 565
- STM-1 signal (SDH) 71, 269, 828
- STM-16 (SDH) 71
- STM-4 (SDH) 71
- STM-64 (SDH) 71
- STM (synchronous transport module) 71, 271
- stop bit 191
- store-and-forward 13, 401, 414
- store-and-retrieve 401, 413
- stored program control (SPC) 80, 100, 532
- STP (shielded twisted pair) cable 143, 817
- STP (signal transfer point) 252, 461
- Stratacom 386
- streetside jointing cabinet 619, 621, 622
- streetway conduit 619
- Strowger switching 86, 88
- structure regulation 797
- structured cabling 370, 373, 817, 818
- STS-1 signal (SONET) 71, 280, 828
- STS-12 (SONET) 71, 280
- STS-192 (SONET) 71, 280
- STS-3 (SONET) 71, 280
- STS-48 (SONET) 71, 280
- STS (synchronous transport system, SONET) 71, 279
- stuffing 68, 268
- sub-addressing 520
- sub-D-socket 185
- sub-multiplexing 274
- sub-network address 526
- submission 401
- submission and delivery entity (SDE) 418, 422
- submission and delivery protocol (P3) 418
- submission envelope 416
- subnetwork access protocol (SNAP) 406, 407
- subscriber network interface (SNI) 137, 332, 397
- subscriber number 514
- subscription charge 655
- successful transfer rate 641
- SUN Microsystems 487, 735
- super mastergroup 37
- supergroup 37, 146, 696, 828
- supergroup translating equipment (STE) 38
- supervisory function 349
- supplementary service 214, 215, 511, 655
- suppressed demand 753
- suppressed traffic 560
- surface wave path 156, 161
- surfing the Internet 405
- SVC (signalling virtual channel) 460, 461
- SVC (switched virtual circuit) 347, 351, 384, 457, 469, 656
- Swiss PTT 837
- switch matrix 77, 82, 96
- switchboard 281, 285
- switched multimegabit digital service (SMDS) 394, 472, 655
- switched virtual circuit (SVC) 347, 351, 384, 457, 469, 656
- switching and management infrastructure (SwMI) 432
- switching costs 659
- SYN (synchronization byte) 192, 199, 255
- synchronization 66, 189, 609, 610, 673, 709
- synchronization byte (SYN) 199, 255
- synchronous communication 355
- synchronous data transfer 191
- synchronous datalink control (SDLC) 191, 205, 362, 365
- synchronous digital hierarchy (SDH) 70, 267, 691
- synchronous frame formatting 70
- synchronous optical network (SONET) 70, 267, 279, 691
- synchronous transport module (STM) 71, 271
- synchronous transport system (STS, SONET) 71, 279
- syntax 197, 417, 787
- system services control point (SSCP) 362, 364
- systems network architecture (SNA) 205, 341, 361, 724, 738

- T-interface 216, 217
- T-Online 400
- T-span 64
- T Com (Telecommunications Commission of CEPT) 728
- T.50, ITU-T recommendation 47, 729
- T.90, ITU-T recommendation 214
- T1-committee 730
- T1-D1 730
- T1-X1 730
- T1 signal 64, 180, 268, 828
- T1 timer 350
- T2 parameter 350
- T3 signal 65, 268, 828
- T3 timer 350
- TA (terminal adaptor) 215, 217, 221, 222
- TACS (Total Access Communication System) 244, 306
- TAE (test access equipment) 675
- tandem 81
- tandem exchange 575
- TAR (temporary alternative re-routing) 688, 690
- target answer time 544
- target sales volume 833
- tariff amendment 811
- tariff model 658
- tariff principles 804
- Tb-interface 446
- TC-PDU (transmission convergence layer protocol data unit) 467
- TC-SDU (transmission convergence layer service data unit) 467
- TC (transaction capabilities) 234, 250, 261
- TC (transmission convergence) sublayer 465, 468
- TCAM (advanced communication function/telecommunications access method) 364
- TCAP (transaction capability application part) 232, 233, 234, 250, 254
- TCP (transmission control protocol) 406, 410
- TCP/IP (transmission control protocol/internet protocol) 374, 375, 376, 393, 404, 586
- TCP/IP protocol stack 405
- TDD (time division duplexing) 327
- TDM (time division multiplexing) 57, 61, 62, 698
- TDMA (time division multiple access) 173, 309, 631
- TE (terminal equipment) 216, 217
- TE2 (terminal equipment type 2) 447
- technical and office protocol (TOP) 204, 731, 787
- technical conformance 764
- technical support 665
- tee-off 369
- tele-education 439, 440
- Telecom Australia 394
- telecommunications 3
- Telecommunications Act, 1984 (UK) 806, 808
- Telecommunications Act, 1987 (Finland) 807
- Telecommunications Administration Centre (Finland) 807
- Telecommunications business law, 1985 (Japan) 812
- telecommunications management network (TMN) 279, 477, 481, 739
- Telecommunications reform Act, 1996 (United States) 810
- Telecommunications Standardization Bureau (TSB) 725
- Telecommunications Structure Reform (Netherlands) 807
- telecommunications system 9
- Teledesid 171, 316
- telegraph 8
- telegraph pole 619
- Telekommunikationsgesetz (TKG, 1996, Germany) 807
- Telelaboratoriet (Denmark) 807
- telemarketing 543, 790
- telematid 415
- telemedicine 439, 440
- Telenet 347
- telephone 8
- Telephone and Telegraph Act, 1987 (Denmark) 807
- telephone channel 37, 39, 696
- telephone network 211
- telephone user part (TUP) 250, 254, 257
- telephony 21
- telephony passive optical network (TPON) 332, 620
- telephony user part enhanced (TUP+) 224
- Telepoint 322
- Teleport 330
- teleprinter 20
- telesales 790
- teleservice 214, 215, 431
- teleshopping 439, 440
- teletraffic engineering 529
- teletraffic theory 529
- televoting 236, 243
- teleworking 439, 440
- telex 8, 20, 45
- telex destination code 524
- telex network identification code (TNIC) 522
- telex numbering plan 522
- Telia (Swedish PTT) 837

- Telification B.V. (Netherlands) 807
- Telnet 406, 407
- Telstar satellite 169
- temperature 674
- temporary alternative re-routing (TAR) 688, 690
- tender 770
- tendering 763
- terminal adaptor (TA) 215, 217, 221, 222
- terminal adaptor, broadband (B-TA) 447
- terminal equipment (TE) 216, 217
- terminal equipment type 1, broadband (B-TE1) 447
- terminal equipment type 2 (TE2) 447
- terminal equipment type 2, broadband (B-TE2) 447
- terminal equipment, broadband (B-TE) 456
- terminal equipment, regulation of 803
- terminal node 364
- terminal/network access point 803
- test access equipment (TAE) 675
- test access point 675
- test equipment 666
- testing 766
- TETRA (trans European trunk radio system) 299, 431
- TETRA PDO 431
- TETRA V + D 431
- TEUREM accounting rates 657
- TFTP (trivial file transfer protocol) 405, 406, 407, 408
- TH (transmission header) 363
- thin ethernet 369
- third line support 665
- throughput 382, 562, 683
- throughput capacity 562
- throughput class 353
- throughput negotiation 354
- ticket 281
- ticketing 652, 656
- time-of-day charging 649
- time-space-time (TST) switching 102, 104
- time division duplexing (TDD) 327
- time division multiple access (TDMA) 173, 309, 631
- time division multiplexing (TDM) 57, 61, 62, 698
- time switching 103, 105
- timeslot 0 (zero) 66
- timeslot 16 66
- timestamp 409
- TINA (telecommunicationns intelligent network architecture) 490
- tip-off 801
- TKG (Telekommunikationsgesetz, 1996, Germany) 807
- TMN (telecommunications management network) 279, 477, 481, 739
- TMN management function model 486
- TMR (trunk mobile radio) 297, 299
- TNIC (telex network identification code) 522
- token 371
- token bus LAN 368, 371
- token ring LAN 368, 372
- token ring LAN, 16 Mbit/s 372
- token ring LAN, 4 Mbit/s 372
- toll-free service 237
- tone-on-idle 125
- tone generator 670
- TOP (technical and office protocol) 204, 731, 787
- top peer group 501
- topology database 502
- topology management 501
- topology state 500
- topology state packets 501
- topology state parameters 501
- TP (transmission path) 466
- TPON (telephony passive optical network) 332, 620
- TR (Technical Reference) 737
- tracing 289
- tracing (mobile user) 305
- TRADACOMS 787
- traffic 529, 531
- traffic balance 756
- traffic carrying capacity 540
- traffic imbalance 756
- traffic intensity 530, 556, 560, 745
- traffic intensity measurement 531
- traffic mean 580
- traffic monitoring 555
- traffic peak 534
- traffic profile 557
- traffic table 540
- traffic variance 580
- training 774
- trans-multiplexor 698
- transaction capabilities (TC) 234, 250, 261
- transaction capability application part (TCAP) 232, 233, 234, 250, 254
- transaction file 787
- transceiver 297
- transfer charges 648
- transformer 23
- transit exchange 494, 574
- transit list 502
- transit node 470, 494

- translation 115, 421, 463
- transmission bridge 24
- transmission control protocol (TCP) 406, 410
- transmission convergence (TC) sublayer 465, 468
- transmission delay 343
- transmission header (TH) 363
- transmission infrastructure 805
- transmission line 9
- transmission loss, free space 625
- transmission medium 3, 8
- transmission path (TP) 466
- transmission phase 435
- transmission plan 14, 593
- transmission plan, international 612
- transmission plan, North America 600
- transmission plan, UK digital network 601
- transmission plant costs 659
- transmission priority 445
- transmission rate 61
- transmission reference point (TRP) 599
- transmit power 300
- transmitter 3
- Transpad 347
- transparency (of business activities) 805
- transponder 167
- transport industry 787
- transport layer (layer 4) 197, 202
- transport mechanism 17
- transport protocol 204
- transport service 414
- transverse screened cable 35, 143
- travelling wave tube (TWT) 166
- treble tone 30
- triangulation 689
- tributary 68, 268, 273, 674, 677
- tributary unit (TU) 271
- tributary unit group (TUG) 271
- trigger table 234
- tripping 115
- trivial file transfer protocol (TFTP) 405, 406, 407, 408
- troposphere 156
- tropospheric scatter 19, 153, 156, 166, 168
- TRP (transmission reference point) 599
- trunk 11, 81, 111
- trunk area code 514
- trunk code 116
- trunk loading 745
- trunk mobile radio (TMR) 297, 299
- trunk offer 289
- trunk reservation 579, 581
- TSB (Telecommunications Standardization Bureau) 725
- TST (time-space-time) switching 102, 104
- TTC (Telecommunication Technology Committee) 731
- TU (tributary unit) 271
- TUG (tributary unit group) 271
- tuning fork 22
- TUP (telephone user part) 250, 254, 257
- TUP+ (enhanced telephone user part) 224, 258
- twisted pair cable 19, 141, 370, 466
- two-motion selector 88, 90, 92, 94
- two-stage call set up 520
- two-to-four wire conversion 602
- two-tone radiopaging 427
- two-wire circuit 35, 600, 602
- two-wire communication 32, 35
- TWT (travelling wave tube) 166
- Tymnet 347
- U-format 349
- U-interface 216, 218
- u-plane 326, 432, 459
- UA (user agent) 414, 420
- UAE (user agent entity) 418, 422
- UAL (user agent layer) 417
- Ub-interface 447
- UDP (user datagram protocol) 406
- UHF (ultra high frequency radio) 19, 153, 156, 161
- UK digital network transmission plan 601
- ultra high frequency radio (UHF) 19, 153, 156, 161
- Um-interface 310
- umbrella network management system 479
- UMTS (universal mobile telephone service) 314, 316
- unbalanced circuit 144, 188
- uncontrolled transmission mode 464
- under-provision 743
- unfair use of information 801
- UNI (user network interface) 137, 347
- UNI signalling 502
- UNI v3.0 (ATM) 472
- UNI v3.1 (ATM) 472, 500
- UNI, ATM 456, 457, 464, 470
- UNI, B-ISDN 446, 447
- UNI, frame relay 382, 386
- unicast 434
- uninterruptible power supply (UPS) 773
- uniselector 88, 89, 91, 94
- Unisource 837
- Unisys 361

- UNITAX 306
- UNITDI (United Nations Trade Data Interchange) 787
- universal mobile telephone service (UMTS) 314, 316
- universal number service 236
- universal personal telephone service 236
- universal service 802, 805
- University of Western Australia 394
- UNIX 366, 375, 383, 404, 527, 736
- unknown destination 411
- unreachable host 410
- unreliability 675
- unshielded twisted pair (UTP) cable 143, 817
- UNSM (United Nations standard messages) 788
- unsuccessful call 117
- upgrade 681
- uplink 306, 309, 311, 625, 629
- uplink (hello procedure) 502
- upper sideband 39, 40, 145
- UPS (uninterruptible power supply) 773
- US government printing office (GPO) 810
- US West 811
- usage charges 648
- USDC (US digital cellular system) 311
- user-network signalling 469
- user agent (UA) 414, 420
- user agent entity (UAE) 418, 422
- user agent layer (UAL) 417
- user data block 396
- user datagram protocol (UDP) 406
- user network interface (UNI) 137, 347
- user part 256
- user plane 136, 459, 469
- username 714
- utilization 344, 564
- UTP (unshielded twisted pair) cable 143, 817
- V-chip 812
- V.10, ITU-T recommendation 186, 188
- V.11, ITU-T recommendation 186, 188
- V.21, ITU-T recommendation 185
- V.22, ITU-T recommendation 185, 193, 199
- V.22bis, ITU-T recommendation 184, 185
- V.23, ITU-T recommendation 185
- V.24, ITU-T recommendation 186, 188, 193, 199, 730, 828
- V.25bis, ITU-T recommendation 193
- V.26, ITU-T recommendation 185
- V.27, ITU-T recommendation 185
- V.27ter, ITU-T recommendation 185
- V.28, ITU-T recommendation 186, 188, 193, 199
- V.29, ITU-T recommendation 185
- V.32, ITU-T recommendation 184, 185, 193, 199
- V.33, ITU-T recommendation 185
- V.34, ITU-T recommendation 185, 193, 199
- V.35, ITU-T recommendation 185, 188, 199
- V.36, ITU-T recommendation 185, 188, 199, 730, 828
- V.37, ITU-T recommendation 185
- V.42, ITU-T recommendation 185
- V.42bis, ITU-T recommendation 185
- V.54, ITU-T recommendation 679
- V.110, ITU-T recommendation 221
- V5-interface 332, 333
- V5.1 interface 333, 336
- V5.2 interface 333, 335
- V5.x interfaces 336
- VADS (value-added and data services licence, UK) 808
- validation 788
- value-added service 798
- VANS (value-added network service) 798
- VBR (variable bit rate) 458
- VC-11 (virtual container of SDH) 71
- VC-12 (virtual container of SDH) 71
- VC-21 (virtual container of SDH) 71
- VC-22 (virtual container of SDH) 71
- VC-31 (virtual container of SDH) 71
- VC-32 (virtual container of SDH) 71
- VC-4 (virtual container of SDH) 71
- VC (virtual call) 341, 347, 351, 656
- VC (virtual channel) 457
- VC (virtual circuit) 12, 341
- VC (virtual container) 270, 271, 279
- VC level 465
- VCC (virtual channel connection) 459, 460
- VCI (virtual channel identifier) 199, 396, 461, 462
- VDB (visitor database) 325
- verification 680
- very high frequency radio (VHF) 19, 153, 161, 297
- VF (voice frequency) signalling 119, 120
- VHF (very high frequency radio) 19, 153, 161, 297
- via net loss (VNL) 599
- video communication 439
- video on demand (VOD) 148, 329, 331, 440, 443
- videotelephone 215
- videotext 399, 789
- VIF (visual index file) 291
- violation 73, 74
- violation symbol 73



- virtual call (VC) 341, 347, 351, 656
- virtual channel (VC) 457
- virtual channel connection (VCC) 459, 460
- virtual channel identifier (VCI) 199, 396
- virtual channel link 459
- virtual circuit (VC) 12, 341
- virtual connection 12, 13
- virtual container-3 (VC-3) 271
- virtual container-4 (VC-4) 271
- virtual container (VC) 270, 271, 279
- virtual path (VP) 457
- virtual path connection (VPC) 459
- virtual path identifier (VPI) 199, 461, 462
- virtual path link 459
- virtual private network (VPN) 235, 239
- virtual switching point (VSP) 612
- virtual tributary (VT) 279
- virus software 718
- visitor database (VDB) 325
- visitor location register (VLR) 263, 305
- visual index file (VIF) 291
- Viterbi code 201
- VLR (visitor location register) 263, 305
- VNL (via net loss) 599
- VOD (video on demand) 148, 329, 331, 440, 443
- Vodafone 311, 806
- voice frequency (VF) signalling 119, 120
- voice recognition 779
- voice response 246
- voicemail 246, 719
- VP (virtual path) 457
- VP level 465
- VPC (virtual path connection) 459
- VPI (virtual path identifier) 199, 461, 462
- VPI=0 461
- VPN (virtual private network) 235, 239
- VSP (virtual switching point) 612
- VT (virtual tributary) 279
- VT 1.5 (virtual tributary of SONET) 71, 280
- VT 2.0 (virtual tributary of SONET) 71, 280
- VT 3.0 (virtual tributary of SONET) 71, 280
- VT 6.0 (virtual tributary of SONET) 71, 280
- VTAM (virtual telecommunications access method) 364
- W-interface 448, 449
- waiting calls 546
- waiting calls, average number of 545
- WAN (wide area network) 346, 367, 384
- WARC (World Administrative Radio Council) 307, 623
- warranty 764
- wave propagation 155
- wavelength division multiplexing (WDM) 151, 152, 332, 699
- WDM (wavelength division multiplexing) 151, 152, 332, 699
- weighted noise 671
- Wellfleet 500
- wide area cordless telephony 322
- wide area network (WAN) 346, 367, 384
- Wilkinson-Rapp equivalent random method 578, 579, 590
- window 380
- window negotiation 354
- window size 350, 354, 380
- Windows 375
- Windows for workgroups 375
- Windows95 375, 402
- WindowsNT 375
- wiper 93
- wiper assembly 89
- wireless LAN 432
- wireless local loop (WLL) 319
- wiring cabinet 371, 373
- WLL (wireless local loop) 319
- word synchronization 190
- workstation (WS) 481, 482
- World Administrative Radio Council (WARC) 307, 623
- worldwide web (WWW) 405
- WS (workstation) 481, 482
- WTSC (World Telecommunications Standardization Conference) 726
- WWW (worldwide web) 405
- X-interface 481
- X-OFF 356, 358
- X-ON 356, 358
- X-windows 406, 407
- X-Y switching 88
- X.3, ITU-T recommendation 355
- X.21, ITU-T recommendation 186, 188, 193, 199, 346, 828
- X.21bis, ITU-T recommendation 186, 188, 199, 346
- X.24, ITU-T recommendation 188
- X.25 options 353
- X.25, ITU-T recommendation 138, 207, 346, 379, 423, 430, 521, 655
- X.26, ITU-T recommendation 188
- X.27, ITU-T recommendation 188
- X.28, ITU-T recommendation 355, 356

- X.29, ITU-T recommendation 355  
X.30, ITU-T recommendation 221  
X.31, ITU-T recommendation 220, 521  
X.31 case A 220  
X.31 case B 220  
X.32, ITU-T recommendation 346  
X.75, ITU-T recommendation 220, 348, 358, 359, 505  
X.121, ITU-T recommendation 344, 513, 521, 522, 527  
X.200, ITU-T recommendation 195  
X.208, ITU-T recommendation 196  
X.213, ITU-T recommendation 527  
X.224, ITU-T recommendation 423  
X.225, ITU-T recommendation 423  
X.400, ITU-T recommendation 401, 402, 403, 424, 524, 719  
X.500, ITU-T recommendation 402, 421, 424, 513, 524  
X.680, ITU-T recommendation 196  
X.700, ITU-T recommendation 485  
X.722, ITU-T recommendation 485  
Xerox 432  
XID (exchange identifier) 716  
yellow book 740  
yellow pages 811  
ZBTISI (zero byte time slot interchange) code 74  
zero byte time slot interchange (ZBTISI) code 74