

COMPUTER NETWORKS



COMPUTER NETWORKS

COMPUTER NETWORKS

For

*B.E./B.Tech/AMIE/AMIETE examinations in Computer
Science and Electronics discipline*

By

Er. Vikrant Vij

B.tech (Hons.), M.E (ECE)

*Ex- Faculty in Electronics and Communication Engineering
Jaypee University of Information Technology, Himachal Pradesh*



UNIVERSITY SCIENCE PRESS

(An Imprint of Laxmi Publications Pvt. Ltd.)

An ISO 9001:2008 Company

BENGALURU • CHENNAI • GUWAHATI • HYDERABAD • JALANDHAR
KOCHI • KOLKATA • LUCKNOW • MUMBAI • RANCHI • NEW DELHI
BOSTON (USA) • NAIROBI (KENYA)

COMPUTER NETWORKS

© by Laxmi Publications (P) Ltd.

All rights reserved including those of translation into other languages. In accordance with the Copyright (Amendment) Act, 2012, 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 or otherwise. Any such act or scanning, uploading, and or electronic sharing of any part of this book without the permission of the publisher constitutes unlawful piracy and theft of the copyright holder's intellectual property. If you would like to use material from the book (other than for review purposes), prior written permission must be obtained from the publishers.

Printed and bound in India

Typeset at ABRO Enterprises, Delhi

First Edition : 2018

ISBN 978-93-5274-080-2

Limits of Liability/Disclaimer of Warranty: The publisher and the author make no representation or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties. The advice, strategies, and activities contained herein may not be suitable for every situation. In performing activities adult supervision must be sought. Likewise, common sense and care are essential to the conduct of any and all activities, whether described in this book or otherwise. Neither the publisher nor the author shall be liable or assumes any responsibility for any injuries or damages arising here from. The fact that an organization or Website if referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Website may provide or recommendations it may make. Further, readers must be aware that the Internet Websites listed in this work may have changed or disappeared between when this work was written and when it is read.

All trademarks, logos or any other mark such as Vibgyor, USP, Amanda, Golden Bells, Firewall Media, Mercury, Trinity, Laxmi appearing in this work are trademarks and intellectual property owned by or licensed to Laxmi Publications, its subsidiaries or affiliates. Notwithstanding this disclaimer, all other names and marks mentioned in this work are the trade names, trademarks or service marks of their respective owners.

Branches		
(C)	Bengaluru	080-26 75 69 30
(C)	Chennai	044-24 34 47 26, 24 35 95 07
(C)	Guwahati	0361-254 36 69, 251 38 81
(C)	Hyderabad	040-27 55 53 83, 27 55 53 93
(C)	Jalandhar	0181-222 12 72
(C)	Kochi	0484-237 70 04, 405 13 03
(C)	Kolkata	033-22 27 43 84
(C)	Lucknow	0522-220 99 16
(C)	Mumbai	022-24 93 12 61
(C)	Ranchi	0651-220 44 64

PUBLISHED IN INDIA BY



UNIVERSITY SCIENCE PRESS

(An Imprint of Laxmi Publications Pvt. Ltd.)

An ISO 9001:2008 Company
113, GOLDEN HOUSE, DARYAGANJ,
NEW DELHI - 110002, INDIA
Telephone : 91-11-4353 2500, 4353 2501
Fax : 91-11-2325 2572, 4353 2528
www.laxmipublications.com info@laxmipublications.com

C—
Printed at:

Dedicated to ...

My Loving Parents, My Living Angels, My Guides

Mrs. Muktesh Vij & Mr. Harsh Vij

For showing me light on this earth...

CONTENTS

<i>Preface</i>	(xi)
<i>Acknowledgements</i>	(xiii)

CHAPTER 1: COMPUTER NETWORK FUNDAMENTALS 1–35

1.1 Introduction	1
1.2 Computer Network Types	3
1.3 Network Topology	5
1.4 Network Connectivity and Protocols	9
1.5 Network Models	12
1.6 Network Services	29
1.7 Examples of Network Architectures	32
<i>Review Questions</i>	35

CHAPTER 2: THE PHYSICAL LAYER 36–63

2.1 Introduction	36
2.2 The Theoretical Basis for Data Communication	36
2.3 Data Communication Transmission Technology	41
2.4 Data Communication Media Technology	43
2.5 Mobile and Cellular Networks and Communication	49
2.6 Optical Switching Networks	59
<i>Review Questions</i>	63

CHAPTER 3: DATA LINK LAYER 64–99

3.1 Introduction	64
3.2 Framing	66
3.3 Error Detection and Correction	68
3.4 Feedback Error Control	82
3.5 Flow Control	84

3.6 Introduction of Link Management	88
3.7 High Level Data Link Control	89
3.8 Link Management	93
3.9 Point to Point Protocol	97
<i>Review Questions</i>	99

CHAPTER 4: THE MEDIUM ACCESS CONTROL SUBLAYER **100–162**

4.1 Introduction	100
4.2 Multiple Access Protocols	102
4.3 ETHERNET	111
4.4 Token-Ring	127
4.5 Token Bus	140
4.6 Wireless LAN (IEEE 802.11)	141
4.7 IEEE 802.16 BROADBAND WIRELESS	150
4.8 Bluetooth	153
4.9 Connecting Devices	158
<i>Review Questions</i>	160

CHAPTER 5: THE NETWORK LAYER **163–194**

5.1 Introduction	163
5.2 Services Provided to Transport Layer	163
5.3 Network Layer Issues: Delivery and Forwarding	164
5.4 The Network Layer in Internet	166
5.5 Network Address Translation (NAT)	174
5.6 IP Version 6	176
5.7 Address Translation (ARP)	178
5.8 Host Configuration (DHCP)	180
5.9 The Internet Control Message Protocol (ICMP)	182
5.10 Routing Protocols	186
5.11 Internet Protocol Security (IPSec)	189
5.12 Virtual Private Networks (VPN)	192
<i>Review Questions</i>	193

CHAPTER 6: THE TRANSPORT LAYER **195–224**

6.1 Introduction	195
6.2 Transport Services	196

6.3 Elements of Transport Layer Protocols	199
6.4 Three Protocols	206
6.5 User Datagram Protocol (UDP)	207
6.6 TCP	211
6.7 Network Security at Transport Layer	220
<i>Review Questions</i>	224

CHAPTER 7: THE APPLICATION LAYER **225–251**

7.1 Introduction	225
7.2 Domain Name System	225
7.3 Electronic Mail (SMTP, MIME, IMAP)	230
7.4 TELNET	239
7.5 World Wide Web	240
7.6 Network Management	243
<i>Review Questions</i>	251

CHAPTER 8: COMPUTER NETWORK SECURITY **252–281**

8.1 Introduction	252
8.2 Securing the Computer Network	254
8.3 Forms of Protection	255
8.4 Security Standards	258
8.5 Sources of Security Threats	261
8.6 Security Threat Management	265
8.7 Cyber Crimes and Hackers	266
8.8 Cryptography	269
8.9 Firewalls	276
<i>Review Questions</i>	280
Glossary	282–341
Index	342–345

PREFACE

In The Name of GOD and My Angel Guides.....

There are two sides of computer revolution: one is represented by the PC on your desktop and the second one by the device that remote-controls your TV, monitors and operates your car engine, and allows you to set up your answering machine and your microwave oven. Man's constant quest to communicate has resulted in a quantum leap in technology related to data communications. For the past quarter century the maximum obtainable transmission rate on many types of communications facilities has doubled every three to five years. During the past few years this growth rate has accelerated, with emerging technologies providing a transmission capability, an order of magnitude or more above what were considered high operating rates just a year or two ago. Accompanying this growth and, in many cases, providing the impetus for the technological developments that made such growth possible are communications-dependent applications.

This book examines a wide range of techniques, technologies and systems used in data communications and computer networks. In particular it addresses a variety of data transmission methods used to convey data between physically distant locations. A number of types of network are considered which may interconnect a few or many thousands of users on either a permanent or temporary, switched basis. In order to support successful communication, a set of rules and procedures has been developed, many of which are embodied in internationally agreed standards. We shall look at the standard bodies and their standards and recommendations and also how they are used in practice.

The organized material that resulted in this book is an attempt to provide a reasonable degree of balance between rigor, clarity of presentation and at the same time keeping the length of book at manageable level. The principal objective of the book is to provide an introduction to various computer network fundamentals and to develop a foundation that can be used as basis for research and further study in this field.

Chapter I is Introduction to Computer Networks highlighting Network hardware and software issues including Network models and services.

Chapter II is about Physical Layer and networks

Chapter III is about Data link layer concepts, Flow control, error control techniques.

Chapter IV is about MAC Layer including Ethernet, WLAN, Bluetooth concepts.

Chapter V is about Network layer issues, IP Address, Congestion control, Routing, etc.

Chapter VI is about Transport Layer and its relevant issues, TCP, UDP Protocols.

Chapter VII is about Application layer including DNS, WWW, e-mail, etc.

Chapter VIII is about Network security and cryptography.

Readers of this book

The book has been written to position itself within the marketplace midway between a number of excellent texts in this subject area which may be regarded as comprehensive, and almost reference works, and a number of other texts which are rather brief and tend to be merely primers. Reference texts can be too detailed and large for a newcomer to the topic and the primer type of text can lack information and be rather bland for many readers. With this in mind the book is written to appeal to two broad ranges of readers:

1. Students of computer science/electronics engineering undergraduate and postgraduate courses who are studying data communications and/or computer networks.
2. Professional engineers, managers and users who work with computer networks in their everyday work. Here the book may be used either as an introductory text for personnel moving into the field or for updating in cases where previous experience has predominantly been in the area of traditional telecommunications...

While the information contained in this book has been carefully checked for accuracy, the author assumes no responsibility or liability for its use, or any infringement of patents or other rights of third parties which would result.

—Author

ACKNOWLEDGEMENTS

Credit where credit is due

I wish to express my gratitude to *My loving Grandparents* for their support and blessings throughout my life.

I would like to thank *my dear brother Er. Abhey Vij* for his wonderful ideas; that helped me in preparation of manuscript of this book.

I am grateful to my friend *Er. Tarun Mittar*; for inspiring me to work on this project.

I would like to express my sincerest thanks to *Dr. Rajneesh Arora*, Vice Chancellor Punjab Technical University, *Dr. Parijat De*, Director Technical Education, West Bengal, *Dr. A.K. Bhandari*, Registrar Punjab University, Chandigarh for their blessings and help in my academic pursuits.

I am thankful to *Sh. Manoj Gaur Ji*, Chairman, Jaypee Group, *Dr. Yajulu Medury*, COO, Jaypee Education System, *Brig (Retd.) Balbir Singh*, JUIT Solan, *Dr. Vivek Sehgal*, JUIT Solan for providing me conducive environment during my stay at Jaypee Institutes.

Lastly, I would like to extend my sincerest thanks to Management of IET Bhaddal and *Mrs. Kulwinder Gurcharan Singh*, Chairperson KFET for helping me in initial start of my teaching career.

—Author

CHAPTER 1

COMPUTER NETWORK FUNDAMENTALS

*I must Create a System, or be enslav'd by another Man's; I will not
Reason and Compare: my business is to Create.*

—William Blake

1.1 INTRODUCTION

The basic ideas in all types of communication are that there must be three ingredients for the communication to be effective. First, there must be two entities, dubbed a sender and a receiver. These two must have something they need to share. Second, there must be a medium through which the sharable item is channeled. This is the transmission medium. Finally, there must be an agreed-on set of communication rules or protocols. These three apply to every category or structure of communication. The effectiveness of a data communications system depends on four fundamental characteristics: delivery, accuracy, timeliness, and jitter.

1. **Delivery.** The system must deliver data to the correct destination. Data must be received by the intended device or user and only by that device or user.
2. **Accuracy.** The system must deliver the data accurately. Data that have been altered in transmission and left uncorrected are unusable.
3. **Timeliness.** The system must deliver data in a timely manner. Data delivered late are useless. In the case of video and audio, timely delivery means delivering data as they are produced, in the same order that they are produced, and without significant delay. This kind of delivery is called *real-time* transmission.

4. **Jitter.** Jitter refers to the variation in the packet arrival time. It is the uneven delay in the delivery of audio or video packets. For example, let us assume that video packets are sent every 3D ms. If some of the packets arrive with 3D-ms delay and others with 4D-ms delay, an uneven quality in the video is the result.

Figure 1.1 shows a simple communication model consisting of following key elements:

1. **Source.** This device generates the data to be transmitted; examples are telephones and personal computers.
2. **Transmitter.** Usually, the data generated by a source system are not transmitted directly in the form in which they were generated. Rather, a transmitter transforms and encodes the information in such a way as to produce electromagnetic signals that can be transmitted across some sort of transmission system. For example, a modem takes a digital bit stream from an attached device such as a personal computer and transforms that bit stream into an analog signal that can be handled by the telephone network.
3. **Transmission System.** This can be a single transmission line or a complex network connecting source and destination.
4. **Receiver.** The receiver accepts the signal from the transmission system and converts it into a form that can be handled by the destination device. For example, a modem will accept an analog signal coming from a network or transmission line and convert it into a digital bit stream.
5. **Destination.** Takes the incoming data from the receiver.

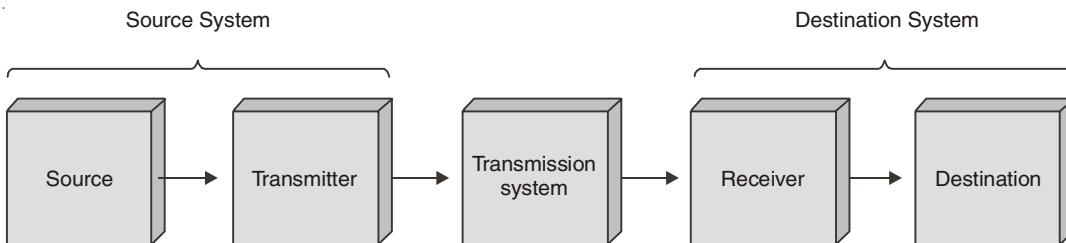


Fig. 1.1. A simple communication model.

In this chapter, we will focus on these three components in a computer network. But what is a computer network? A computer network is a distributed system consisting of loosely coupled computers and other devices. Any two of these devices, which we will from now on refer to as *network elements* or *transmitting elements* without loss of generality, can communicate with each other through a communication medium. In order for these connected devices to be considered a communicating network, there must be a set of communicating rules or protocols each device in the network must follow to communicate with another device in the network. The resulting combination consisting of hardware and software is a computer communication network or computer network in short. Figure 1.2 shows a computer network.

The hardware component is made of network elements consisting of a collection of nodes that include the end systems commonly called hosts and intermediate switching

elements that include hubs, bridges, routers, and gateways that, without loss of generality, we will call network elements.

Network elements may own resources individually, that is locally or globally. Network software consists of all application programs and network protocols that are used to synchronize, coordinate, and bring about the sharing and exchange of data among the network elements. Network software also makes the sharing of expensive resources in the network possible. Network elements, network software, and users all work together so that individual users can exchange messages and share resources on other systems that are not readily available.

Internetworking technology enables multiple, diverse underlying hardware technologies and different software regimes to interconnect heterogeneous networks and bring them to communicate smoothly. The smooth working of any computer communication network is achieved through the low-level mechanisms provided by the network elements and high-level communication facilities provided by the software running on the communicating elements.

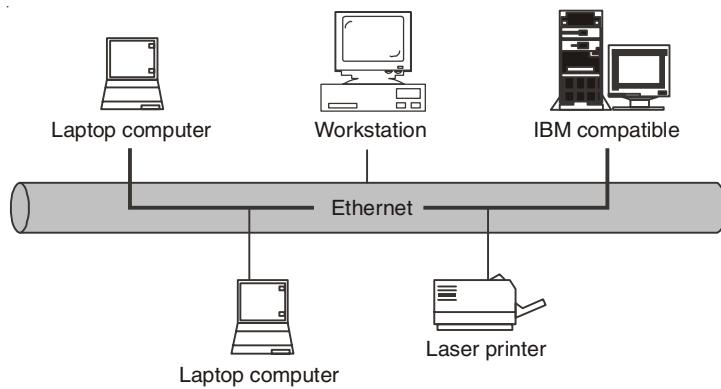


Fig. 1.2. A computer network.

1.2 COMPUTER NETWORK TYPES

Computer networks come in different sizes. Each network is a cluster of network elements and their resources. The size of the cluster determines the network type. There are, in general, two main network types: the local area network (LAN) and wide area network (WAN).

1.2.1 Local Area Networks (LANs)

A computer network with two or more computers or clusters of network and their resources connected by a communication medium sharing communication protocols and confined in a small geographical area, such as a building floor, a building, or a few adjacent buildings, is called a local area network (LAN). The advantage of a LAN is that all network elements are close together so the communication links maintain a higher speed of data movement.

Also, because of the proximity of the communicating elements, high-cost and high quality communicating elements can be used to deliver better service and high reliability. Figure 1.3 shows a LAN network.

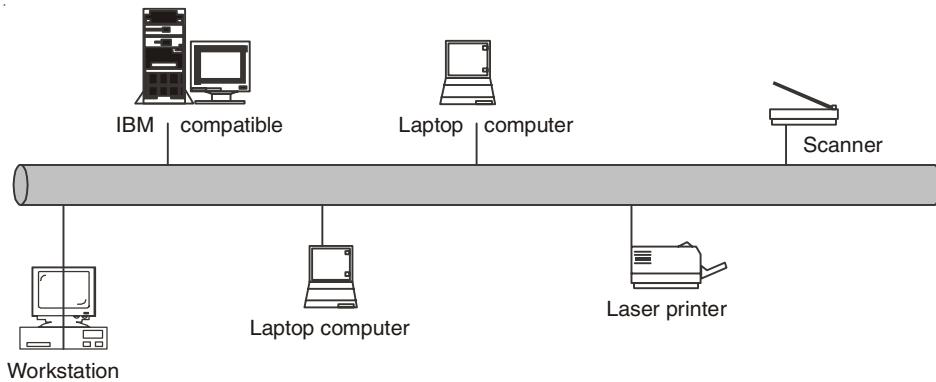


Fig. 1.3. A LAN network.

1.2.2 Wide Area Networks (WANs)

A wide area network (WAN), on the other hand, is a network made up of one or more clusters of network elements and their resources but instead of being confined to a small area, the elements of the clusters or the clusters themselves are scattered over a wide geographical area as in a region of a country or across the whole country, several countries, or the entire globe like the Internet for example.

Some advantages of a WAN include distributing services to a wider community and availability of a wide array of both hardware and software resources that may not be available in a LAN. However, because of the large geographical areas covered by WANs, communication media are slow and often unreliable. Figure 1.4 shows a WAN network.

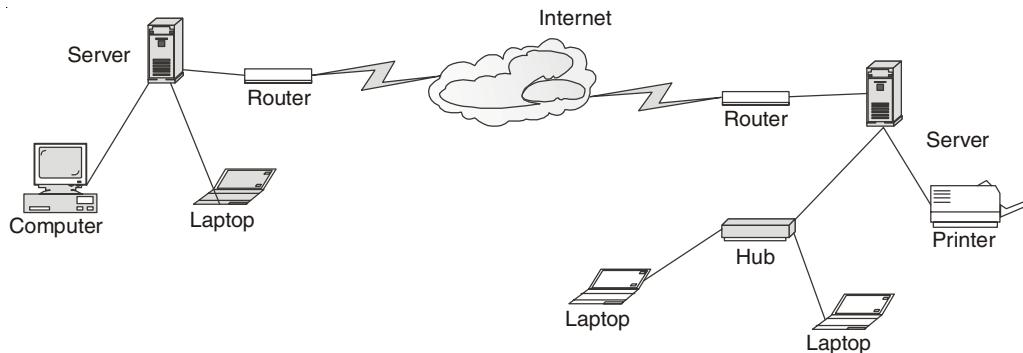


Fig. 1.4. A WAN network.

1.2.3 Metropolitan Area Networks (MANs)

Between the LAN and WAN, there is also a middle network called the metropolitan area network (MAN) because it covers a slightly wider area than the LAN but not so wide as to be considered a WAN. Civic networks that cover a city or part of a city are a good

example of a MAN. MANs are rarely talked about because they are quiet often overshadowed by cousin LAN to the left and cousin WAN to the right.

1.3 NETWORK TOPOLOGY

Computer networks, whether LANs, MANs, or WANs, are constructed based on a topology. There are several topologies including the following popular ones.

1.3.1 Mesh

A mesh topology allows multiple access links between network elements, unlike other types of topologies. The multiplicity of access links between the network elements offers an advantage in network reliability because whenever one network element fails, the network does not cease operations; it simply finds a bypass to the failed element and the network continues to function. Mesh topology is most often applied in MAN networks. Figure 1.5 shows a mesh network.

A mesh offers several advantages over other network topologies. First, the use of dedicated links guarantees that each connection can carry its own data load, thus eliminating the traffic problems that can occur when links must be shared by multiple devices. Second, there is the advantage of privacy or security. When every message travels along a dedicated line, only the intended recipient sees it. Physical boundaries prevent other users from gaining access to messages. Finally, point-to-point links make fault identification and fault isolation easy. Traffic can be routed to avoid links with suspected problems. This facility enables the network manager to discover the precise location of the fault and aids in finding its cause and solution.

The main disadvantages of a mesh are related to the amount of cabling and the number of I/O ports required. First, because every device must be connected to every other device, installation and reconnection are difficult. Second, the sheer bulk of the wiring can be greater than the available space (in walls, ceilings, or floors) can accommodate. Finally, the hardware required to connect each link (I/O ports and cable) can be prohibitively expensive. For these reasons a mesh topology is usually implemented in a limited fashion, for example, as a backbone connecting the main computers of a hybrid network that can include several other topologies.

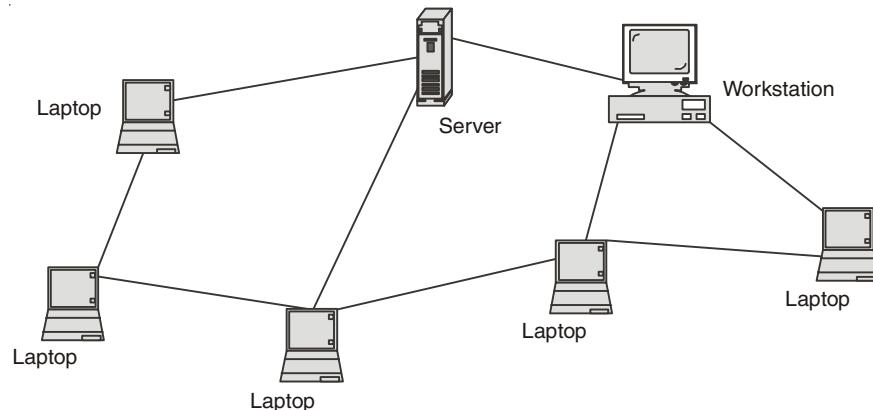


Fig. 1.5. A Mesh network.

1.3.2 Tree

A more common type of network topology is the tree topology. In the tree topology, network elements are put in a hierarchical structure in which the most predominant element is called the *root* of the tree and all other elements in the network share a child-parent relationship. As in ordinary, though inverted trees, there are no closed loops. So dealing with failures of network elements presents complications depending on the position of the failed element in the structure. For example, in a deeply rooted tree, if the root element fails, the network automatically ruptures and splits into two parts. The two parts cannot communicate with each other. The functioning of the network as a unit is, therefore, fatally curtailed. Figure 1.6 shows a network using a tree topology.

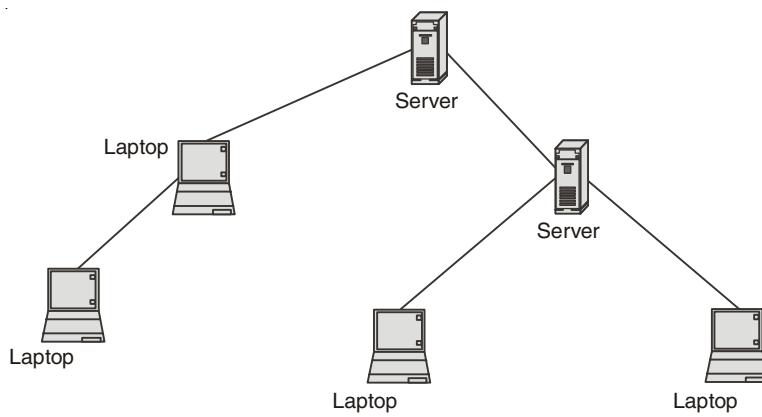


Fig. 1.6. A Tree network.

1.3.3 Bus

A more popular topology, especially for LANs, is the bus topology. Elements in a network using a bus topology always share a bus and, therefore, have equal access to all LAN resources. Every network element has full-duplex connections to the transmitting medium which allows every element on the bus to send and receive data. Because each computing element is directly attached to the transmitting medium, a transmission from any one element propagates through the entire length of the medium in either direction and therefore can be received by all elements in the network. Because of this, precautions need to be taken to make sure that transmissions intended for one element can be received by that element and no other element. The network must also use a mechanism that handles disputes in case two or more elements try to transmit at the same time. The mechanism deals with the likely collision of signals and brings a quick recovery from such a collision. It is also necessary to create fairness in the network so that all other elements can transmit when they need to do so. See Fig. 1.7.

A collision control mechanism must also improve efficiency in the network using a bus topology by allowing only one element in the network to have control of the bus at any one time. This network element is then called the bus master and other elements are considered to be its slaves. This requirement prevents collision from occurring in the network as elements

in the network try to seize the bus at the same time. A bus topology is commonly used by LANs.

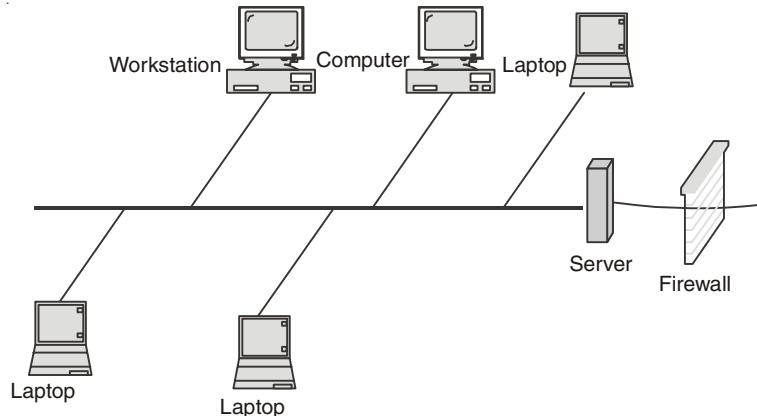


Fig. 1.7. A Bus network.

1.3.4 Star

Another very popular topology, especially in LAN network technologies, is a star topology. A star topology is characterized by a central prominent node that connects to every other element in the network. So, all the elements in the network are connected to a central element. Every network element in a star topology is connected pair wise in a point-to-point manner through the central element, and communication between any pair of elements must go through this central element. The central element or node can either operate in a broadcast fashion, in which case information from one element is broadcast to all connected elements, or transmit as a switching device in which the incoming data is transmitted only to one element, the nearest element enroute to the destination. The biggest disadvantage to the star topology in networks is that the failure of the central element results in the failure of the entire network. Figure 1.8 shows a star topology.

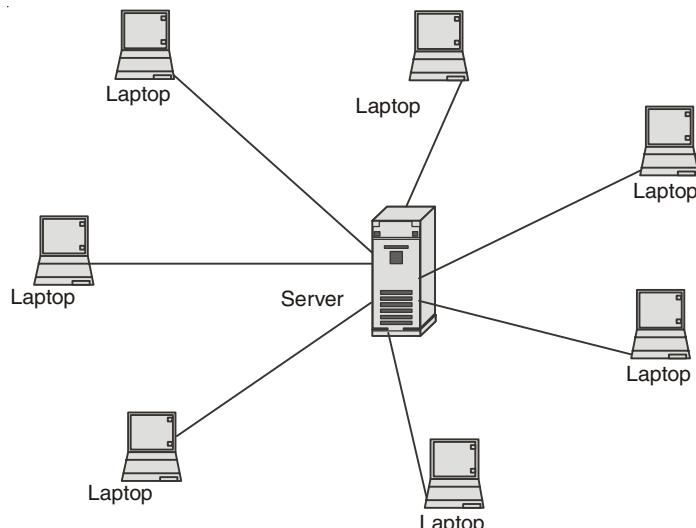


Fig. 1.8. A Star network.

1.3.5 Ring

Finally another popular network topology is the ring topology. In this topology, each computing element in a network using a ring topology is directly connected to the transmitting medium via a unidirectional connection so that information put on the transmission medium can reach all computing elements in the network through a mechanism of taking turns in sending information around the ring. Figure 1.9 shows a ring topology network. The taking of turns in passing information is managed through a *token* system. A token is a system-wide piece of information that guarantees the current owner to be the bus master. As long as it owns the token, no other network element is allowed to transmit on the bus. When an element currently sending information and holding the token has finished, it passes the token downstream to its nearest neighbor. The token system is a good management system of collision and fairness.

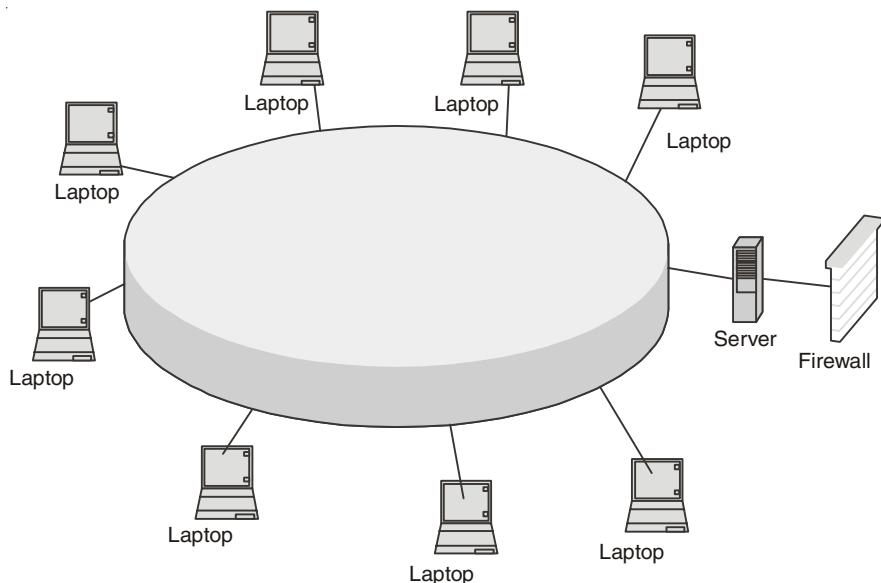


Fig. 1.9. A Ring network.

There are variants of a ring topology collectively called *hub* hybrids combining either a star with a bus or a stretched star as shown in Fig. 1.10

Although network topologies are important in LANs, the choice of a topology depends on a number of other factors, including the type of transmission medium, reliability of the network, the size of the network, and its anticipated future growth. Recently the most popular LAN topologies have been the bus, star, and ring topologies. The most popular bus- and star-based LAN topology is the Ethernet, and the most popular ring-based LAN topology is the token ring.

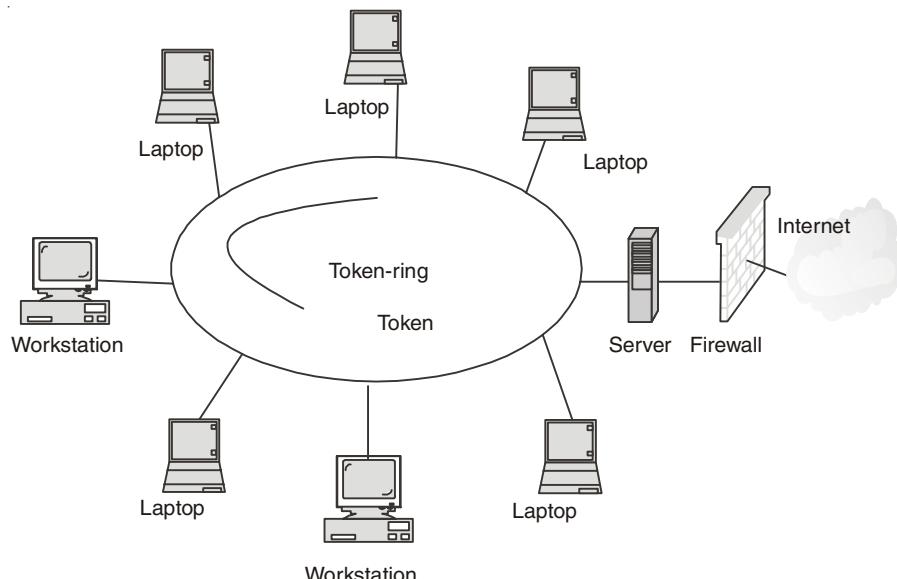


Fig. 1.10. Token Ring Hub.

1.4 NETWORK CONNECTIVITY AND PROTOCOLS

In the early days of computing, computers were used as stand-alone machines, and all work that needed cross-computing was done manually. Files were moved on disks from computer to computer. When computers, terminals, and/or other data processing devices exchange data, the scope of concern is much broader than the concerns we have discussed in Sections 1.2 and 1.3. Consider, for example, the transfer of a file between two computers. There must be a data path between the two computers, either directly or via a communication network. But more is needed. Typical tasks to be performed are:

1. The source system must either activate the direct data communication path or inform the communication network of the identity of the desired destination system.
2. The source system must ascertain that the destination system is prepared to receive data.
3. The file transfer application on the source system must ascertain that the file management program on the destination system is prepared to accept and store the file for this particular user.
4. If the file formats used on the two systems are incompatible, one or the other system must perform a format translation function.

It is clear that there must be a high degree of cooperation between the two computer systems. The exchange of information between computers for the purpose of cooperative action is generally referred to as *computer communications*.

Similarly, when two or more computers are interconnected via a communication network, the set of computer stations is referred to as a *computer network*. Because a similar level of cooperation is required between a user at a terminal and one at a computer, these terms are often used when some of the communicating entities are terminals.

In discussing computer communications and computer networks, two concepts are paramount:

Protocols Computer-communications architecture, or protocol architecture.

"A protocol is a set of rules that govern data communications". A protocol defines what is communicated, how it is communicated, and when it is communicated. The key elements of a protocol are syntax, semantics, and timing.

- **Syntax.** The term *syntax* refers to the structure or format of the data, meaning the order in which they are presented. For example, a simple protocol might expect the first 8 bits of data to be the address of the sender, the second 8 bits to be the address of the receiver, and the rest of the stream to be the message itself.
- **Semantics.** The word *semantics* refers to the meaning of each section of bits. How is a particular pattern to be interpreted, and what action is to be taken based on that interpretation? For example, does an address identify the route to be taken or the final destination of the message?
- **Timing.** The term *timing* refers to two characteristics: when data should be sent and how fast they can be sent. For example, if a sender produces data at 100 Mbps but the receiver can process data at only 1 Mbps, the transmission will overload the receiver and some data will be lost.

1.4.1 Standards

Standards are essential in creating and maintaining an open and competitive market for equipment manufacturers and in guaranteeing national and international interoperability of data and telecommunications technology and processes.

Development of standards: There are a number of advantages and disadvantages to the standards-making process. We list here the most striking ones. The principal advantages of standards are the following:

1. A standard assures that there will be a large market for a particular piece of equipment or software. This encourages mass production and, in some cases, the use of large-scale-integration (LSI) or very-large-scale-integration (VLSI) techniques, resulting in lower costs.
2. A standard allows products from multiple vendors to communicate, giving the purchaser more flexibility in equipment selection and use.

The principal disadvantages are these:

1. A standard tends to freeze the technology. By the time a standard is developed, subjected to review and compromise, and promulgated, more efficient techniques are possible.
2. There are multiple standards for the same thing. This is not a disadvantage of standards *per se*, but of the current way things are done. Fortunately, in recent

years the various standards-making organizations have begun to cooperate more closely. Nevertheless, there are still areas where multiple conflicting standards exist.

Data communication standards fall into two categories: *de facto* (meaning “by fact” or “by convention”) and *de jure* (meaning “by law” or “by regulation”).

- **De facto.** Standards that have not been approved by an organized body but have been adopted as standards through widespread use are *de facto* standards. *De facto* standards are often established originally by manufacturers who seek to define the functionality of a new product or technology.
- **De jure.** Those standards that have been legislated by an officially recognized body are *de jure* standards.

1.4.2 Standards Organizations

Standards are developed through the cooperation of standards creation committees, forums, and government regulatory agencies.

1.4.2.1 Standards Creation Committees

While many organizations are dedicated to the establishment of standards, data telecommunications in North America rely primarily on those published by the following:

- **International Organization for Standardization (ISO).** The ISO is a multinational body whose membership is drawn mainly from the standards creation committees of various governments throughout the world. The ISO is active in developing cooperation in the realms of scientific, technological, and economic activity.
- **International Telecommunication Union-Telecommunication Standards Sector (ITU-T).** By the early 1970s, a number of countries were defining national standards for telecommunications, but there was still little international compatibility. The United Nations responded by forming, as part of its International Telecommunication Union (ITU), a committee, the Consultative Committee for International Telegraphy and Telephony (CCITT). This committee was devoted to the research and establishment of standards for telecommunications in general and for phone and data systems in particular. On March 1, 1993, the name of this committee was changed to the International Telecommunication Union Telecommunication Standards Sector (ITU-T).
- **American National Standards Institute (ANSI).** Despite its name, the American National Standards Institute is a completely private, nonprofit corporation not affiliated with the U.S. federal government. However, all ANSI activities are undertaken with the welfare of the United States and its citizens occupying primary importance.
- **Institute of Electrical and Electronics Engineers (IEEE).** The Institute of Electrical and Electronics Engineers is the largest professional engineering society in the world. International in scope, it aims to advance theory, creativity, and product quality in the fields of electrical engineering, electronics, and radio as well as in all related branches of engineering. As one of its goals, the IEEE oversees the development and adoption of international standards for computing and communications.

- **Electronic Industries Association (EIA).** Aligned with ANSI, the Electronic Industries Association is a nonprofit organization devoted to the promotion of electronics manufacturing concerns. Its activities include public awareness education and lobbying efforts in addition to standards development. In the field of information technology, the EIA has made significant contributions by defining physical connection interfaces and electronic signaling specifications for data communication.

1.4.2.2 Forums

Telecommunications technology development is moving faster than the ability of standards committees to ratify standards. Standards committees are procedural bodies and by nature slow-moving. To accommodate the need for working models and agreements and to facilitate the standardization process, many special-interest groups have developed **forums** made up of representatives from interested corporations. The forums work with universities and users to test, evaluate, and standardize new technologies. By concentrating their efforts on a particular technology, the forums are able to speed acceptance and use of those technologies in the telecommunications community. The forums present their conclusions to the standards bodies.

1.4.2.3 Regulatory Agencies

All communications technology is subject to regulation by government agencies such as the **Federal Communications Commission** (FCC) in the United States. The purpose of these agencies is to protect the public interest by regulating radio, television, and wire/cable communications. The FCC has authority over interstate and international commerce as it relates to communications.

1.5 NETWORK MODELS

As computer networks have proliferated, so the need to communicate between users located on different networks has emerged. Such intercommunicating computer systems may be termed distributed computer systems and are required to process information and pass it between each other.

A new movement was, therefore, born. It was called the *open system movement*, which called for computer hardware and software manufacturers to come up with a way for this to happen. But to make this possible, standardization of equipment and software was needed. To help in this effort and streamline computer communication, the International Standards Organization (ISO) developed the Open System Interconnection (OSI) model. The OSI is an open architecture model that functions as the network communication protocol standard, although it is not the most widely used one. The Transport Control Protocol/Internet Protocol (TCP/IP) model, a rival model to OSI, is the most widely used. Both OSI and TCP/IP models use two protocol stacks, one at the source element and the other at the destination element.

In this section, we give a general idea of the layers of a network and discuss the functions of each. Detailed descriptions of these layers follow in later chapters.

1.5.1 The OSI Model

Established in 1947, the International Standards Organization (ISO) is a multinational body dedicated to worldwide agreement on international standards. An ISO standard that covers all aspects of network communications is the Open Systems Interconnection model. It was first introduced in the late 1970s. *An open system is a set of protocols that allows any two different systems to communicate regardless of their underlying architecture.*

The purpose of the OSI model is to show how to facilitate communication between different systems without requiring changes to the logic of the underlying hardware and software. The OSI model is not a protocol; it is a model for understanding and designing a network architecture that is flexible, robust, and interoperable.

A widely accepted structuring technique, and the one chosen by ISO, is layering. The communications functions are partitioned into a hierarchical set of layers. Each layer performs a related subset of the functions required to communicate with another system, relying on the next-lower layer to perform more primitive functions, and to conceal the details of those functions, as it provides services to the next-higher layer. Ideally, the layers should be defined so that changes in one layer do not require changes in the other layers. Thus, we have decomposed one problem into a number of more manageable sub problems.

The task of ISO was to define a set of layers and to delineate the services performed by each layer. The partitioning should group functions logically, and should have enough layers to make each one manageably small, but should not have so many layers that the processing overhead imposed by their collection is burdensome. The principles (ISO 7498) that guided the design effort are summarized below. The resulting reference model has seven layers.

1. Do not create so many layers as to make the system engineering task of describing and integrating the layers more difficult than necessary.
2. Create a boundary at a point where the description of services can be small and the number of interactions across the boundary are minimized.
3. Create separate layers to handle functions that are manifestly different in the process performed or the technology involved.
4. Collect similar functions into the same layer.
5. Select boundaries at a point which past experience has demonstrated to be successful.
6. Create a layer of easily localized functions so that the layer could be totally redesigned and its protocols changed in a major way to take advantage of new advances in architecture, hardware or software technology without changing the services expected from and provided to the adjacent layers.
7. Create a boundary where it may be useful at some point in time to have the corresponding interface standardized.
8. Create a layer where there is a need for a different level of abstraction in the handling of data, for example morphology, syntax, semantic.
9. Allow changes of functions or protocols to be made within a layer without affecting other layers.

10. Create for each layer boundaries with its upper and lower layer only.
11. Similar principles have been applied to sub layering; Create further sub grouping and organization of functions to form sub layers within a layer in cases where distinct communication services need it.
12. Create, where needed, two or more sub layers with a common, and therefore minimal functionality to allow interface operation with adjacent layers.
13. Allow by-passing of sub layers.

The ISO's OSI seven-layer reference model is shown in Fig. 1.11. The reference model has been developed based upon some of the principles discussed in general terms in the previous section. The seven layers and their boundaries have attempted to build upon other models which have proved successful and in such a way as to optimize the transfer of information between layers.

Layer 1, the physical layer, defines the electrical, mechanical and functional interface between a DCE and the transmission medium to enable bits to be transmitted successfully. The layer is always implemented in hardware. A common example used extensively in modems is the ITU-T's V.24 serial interface. No error control exists at layer 1 but line coding may be incorporated in order to match data signals to certain properties of the communication channel.

Layer 2 is the data link layer, the function of which is to perform error-free, reliable transmission of data over a link. Link management procedures allow for the setting up and disconnection of links as required for communication. Having established a connection, error detection, and optionally error correction, is implemented to ensure that the data transfer is reliable. Flow control is also performed to provide for the orderly flow of data (normally in the form of packets) and to ensure that it is not lost or duplicated during transmission.

Layer 3 is the network layer, whose principal task is to establish, maintain and terminate connections to support the transfer of information between end systems via one, or more, intermediate communication networks. It is the only layer concerned with routing, offering addressing schemes which allow users to refer unambiguously to each other. Apart from the control of connections and routing, the layer, by engaging in a dialogue with the network, offers other services such as a user requesting a certain quality of service or reset and synchronization procedures.

Layer 4 is the transport layer and separates the function of the higher layers, layers 5, 6 and 7, from the lower layers already discussed. It hides the complexities of data communications from the higher layers which are predominantly concerned with supporting applications. The layer provides a reliable end-to-end service for the transfer of messages irrespective of the underlying network. To fulfil this role, the transport layer selects a suitable communications network which provides the required quality of service. Some of the factors which the layer would consider in such selection are throughput, error rate and delay. Furthermore, the layer is responsible for splitting up messages into a series of packets of suitable size for onward transmission through the selected communications network.

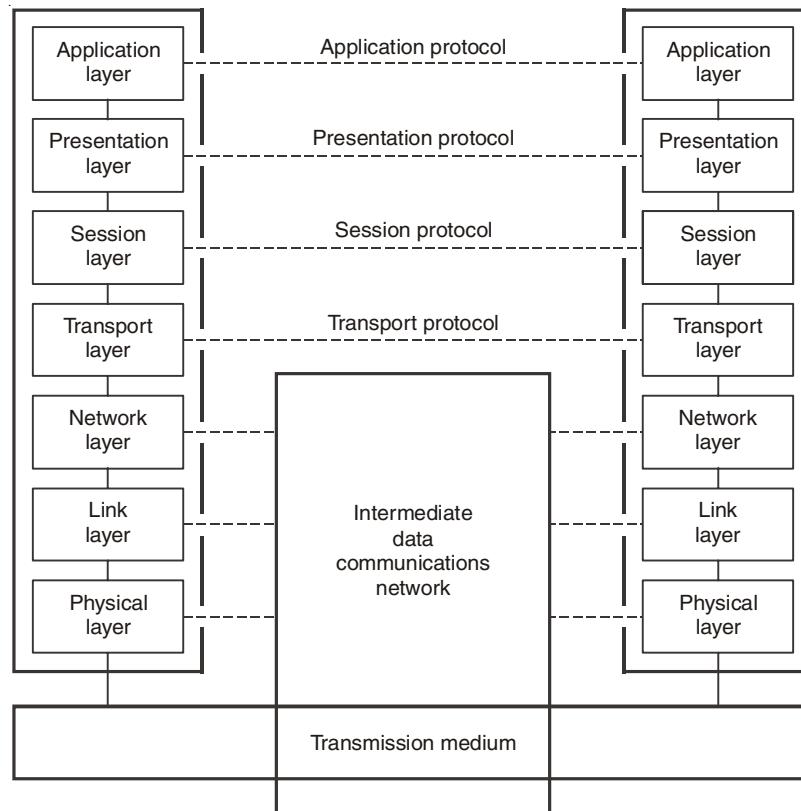


Fig. 1.11. OSI Reference Model.

Layer 5, the session layer, is responsible for establishing and maintaining a logical connection. This may include access controls such as log-on and password protection. Secondly, the session layer performs a function known as **dialogue management**. This is merely a protocol used to order communication between each party during a session. It may be best explained by way of an example. Consider an enquiry/response application such as is used for airline ticket booking systems. Although two-way communication is necessary for such an interactive application it need not be simultaneous. Suppose that the connection only provides communication in one direction at a time. The protocol must therefore regulate the direction of communication at any one instant. If, however, full simultaneous two-way communication is available then little dialogue management is required save some negotiation at set-up time. The third, and most important, function of the session layer is that of recovery (or synchronization). Synchronizing points are marked periodically throughout the period of dialogue. In the event of a failure, dialogue can return to a synchronizing point, restart and continue from that point (using back-up facilities) as though no failure had occurred.

Layer 6 is the presentation layer and presents data to the application layer in a form which it is able to understand. To that end, it performs any necessary code and/or data format conversion. In this way, there is no necessity for the application layer to be aware

of the code used in the peer-to-peer communication at the presentation layer. This means that in practice, users may operate with entirely different codes at each end and which may in turn be different again from the code used across the network for intercommunication. Encryption may also be added at layer 6 for security of messages. Encryption converts the original data into a form which ideally should be unintelligible to any unauthorized third party. Such messages may usually only be decrypted by knowledge of a **key** which of course must be kept secure.

Layer 7, the application layer, gives the end-user access to the OSI environment. This means that the layer provides the necessary software to offer the user's application programs a set of network services, for example an e-mail service. It is effectively the junction between the user's operating system and the OSI network software. In addition, layer 7 may include network management, diagnostics and statistics gathering, and other monitoring facilities.

Most standards activity has centred on the lower layers to support communications networks and their interfaces, for example ITU-T's X.25 recommendation for packet switched network operation addresses layers 1, 2 and 3, only. ISO standards have more recently addressed this imbalance with standards for some applications being available at all seven layers to support a truly open system interconnection.

1.5.2 Standardization Within the OSI Framework

The principal motivation for the development of the OSI model was to provide a framework for standardization. Within the model, one or more protocol standards can be developed at each layer. The model defines, in general terms, the functions to be performed at that layer and facilitates the standards-making process in two ways:

1. Because the functions of each layer are well-defined, standards can be developed independently and simultaneously for each layer, thereby speeding up the standards-making process.
2. Because the boundaries between layers are well-defined, changes in standards in one layer need not affect already existing software in another layer; this makes it easier to introduce new standards.

Figure 1.12 illustrates the use of the OSI model as such a framework. The overall communications function is decomposed into seven distinct layers, using the principles outlined in previous section.

These principles essentially amount to the use of modular design. That is, the overall function is broken up into a number of modules, making the interfaces between modules as simple as possible. In addition, the design principle of information-hiding is used: Lower layers are concerned with greater levels of detail; upper layers are independent of these details. Within each layer, there exist both the service provided to the next higher layer and the protocol to the peer layer in other systems.

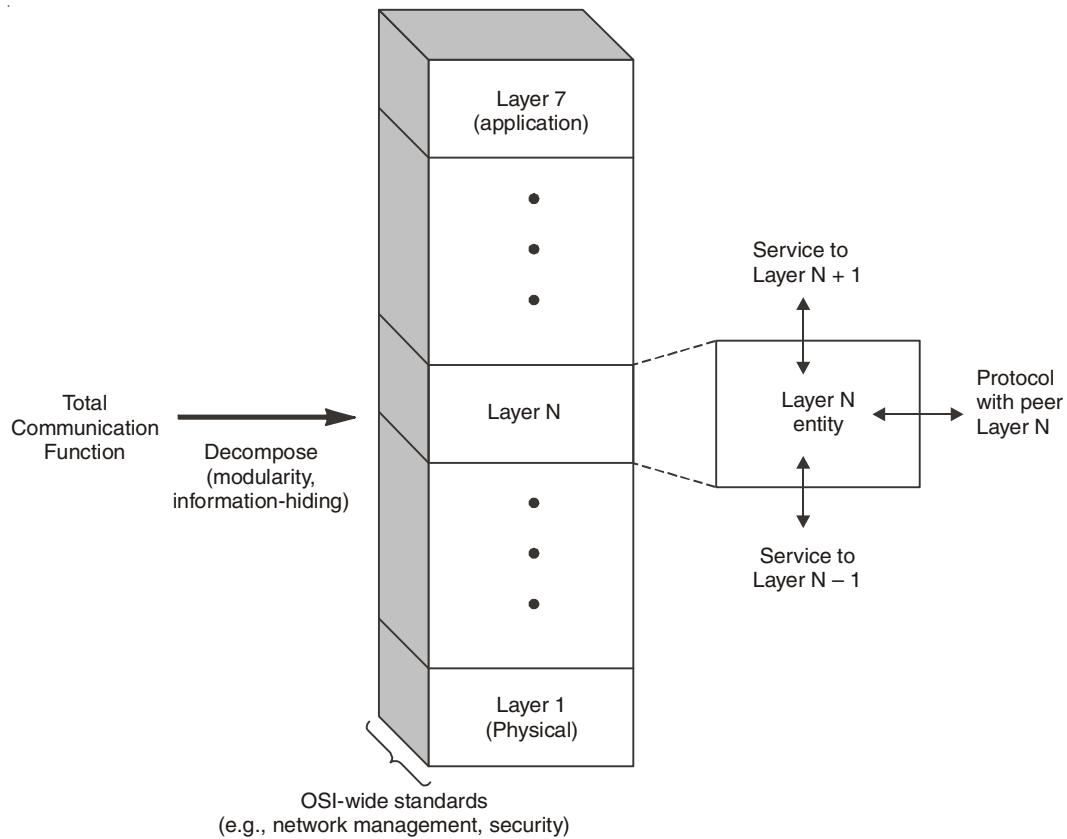


Fig. 1.12. The OSI architecture as a framework for standardization.

Figure 1.13 shows more specifically the nature of the standardization required at each layer. Three elements are key:

Protocol specification. Two entities at the same layer in different systems cooperate and interact by means of a protocol. Because two different open systems are involved, the protocol must be specified precisely; this includes the format of the protocol data units exchanged, the semantics of all fields, and the allowable sequence of PDUs.

Service definition. In addition to the protocol or protocols that operate at a given layer, standards are needed for the services that each layer provides to the next-higher layer. Typically, the definition of services is equivalent to a functional description that defines *what* services are provided, but not *how* the services are to be provided.

Addressing. Each layer provides services to entities at the next-higher layer. These entities are referenced by means of a service access point (SAP). Thus, a network service access point (NSAP) indicates a transport entity that is a user of the network service.

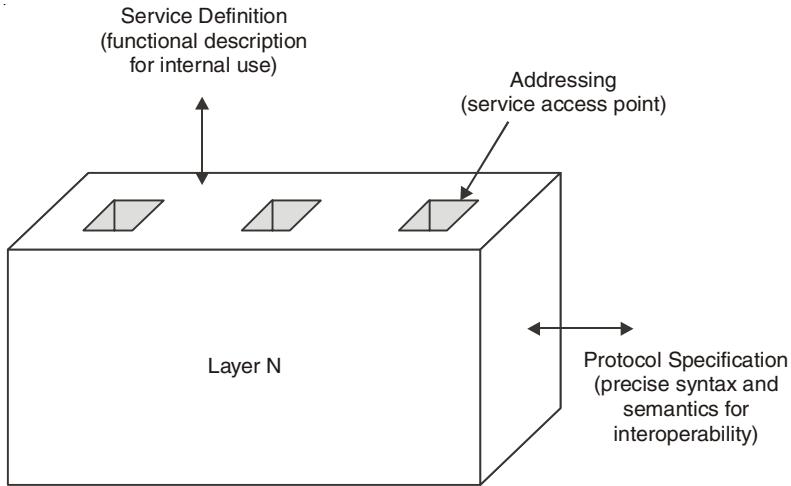


Fig. 1.13. Layer-specific standards.

The need to provide a precise protocol specification for open systems is self evident. The other two items listed above warrant further comment. With respect to service definitions, the motivation for providing only a functional definition is as follows. First, the interaction between two adjacent layers takes places within the confines of a single open system and is not the concern of any other open system. Thus, as long as peer layers in different systems provide the same services to their next-higher layers the details of how the services are provided may differ from one system to another without loss of interoperability. Second, it will usually be the case that adjacent layers are implemented on the same processor. In that case, we would like to leave the system programmer free to exploit the hardware and operating system to provide an interface that is as efficient as possible.

1.5.3 Service Primitives and Parameters

The services between adjacent layers in the OSI architecture are expressed in terms of *primitives* and *parameters*. A primitive specifies the function to be performed, and a parameter is used to pass data and control information. The actual form of a primitive is implementation-dependent; an example is a procedure call. Four types of primitives are used in standards to define the interaction between adjacent layers in the architecture.

REQUEST

A primitive issued by a service user to invoke some service and to pass the parameters needed to fully specify the requested service.

INDICATION

A primitive issued by a service provider to either

1. indicate that a procedure has been invoked by the peer service user on the connection and to provide the associated parameters, or
2. notify the service user of a provider-initiated action.

RESPONSE

A primitive issued by a service user to acknowledge or complete some procedure previously invoked by an indication to that user.

CONFIRM

A primitive issued by a service provider to acknowledge or complete some procedure previously invoked by a request by the service user.

1.5.4 Layers in the OSI Model

In this section we briefly describe the functions of each layer in the OSI model.

1.5.4.1 Physical Layer

The physical layer coordinates the functions required to carry a bit stream over a physical medium. It deals with the mechanical and electrical specifications of the interface and transmission medium. It also defines the procedures and functions that physical devices and interfaces have to perform for transmission to occur. Figure 1.14 shows the position of the physical layer with respect to the transmission medium and the data link layer.

The physical layer is also concerned with the following:

- **Physical characteristics of interfaces and medium.** The physical layer defines the characteristics of the interface between the devices and the transmission medium. It also defines the type of transmission medium.
- **Representation of bits.** The physical layer data consists of a stream of bits (sequence of 0s or 1s) with no interpretation. To be transmitted, bits must be coded into signals—electrical or optical. The physical layer defines the type of encoding (how 0s and 1s are changed to signals).
- **Data rate.** The transmission rate—the number of bits sent each second—is also defined by the physical layer. In other words, the physical layer defines the duration of a bit, which is how long it lasts.
- **Synchronization of bits.** The sender and receiver not only must use the same bit rate but also must be synchronized at the bit level. In other words, the sender and the receiver clocks must be synchronized.
- **Line configuration.** The physical layer is concerned with the connection of devices to the media. In a point-to-point configuration, two devices are connected through a dedicated link. In a multipoint configuration, a link is shared among several devices.
- **Physical topology.** The physical topology defines how devices are connected to make a network. Devices can be connected by using a mesh topology (every device is connected to every other device), a star topology (devices are connected through a central device), a ring topology (each device is connected to the next, forming a ring), a bus topology (every device is on a common link), or a hybrid topology (this is a combination of two or more topologies).
- **Transmission mode.** The physical layer also defines the direction of transmission between two devices: simplex, half-duplex, or full-duplex. In simplex mode, only

one device can send; the other can only receive. The simplex mode is a one-way communication. In the half-duplex mode, two devices can send and receive, but not at the same time. In a full-duplex (or simply duplex) mode, two devices can send and receive at the same time.

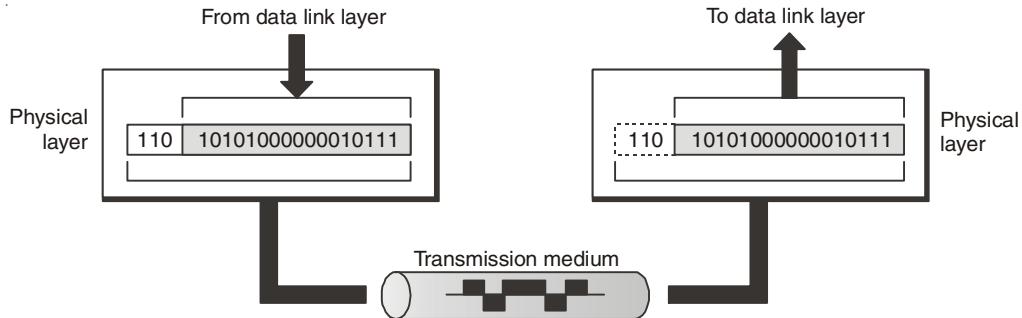


Fig. 1.14. Physical Layer.

1.5.4.2 Data Link Layer

The data link layer transforms the physical layer, a raw transmission facility, to a reliable link. It makes the physical layer appear error-free to the upper layer (network layer). Figure 1.15 shows the relationship of the data link layer to the network and physical layers.

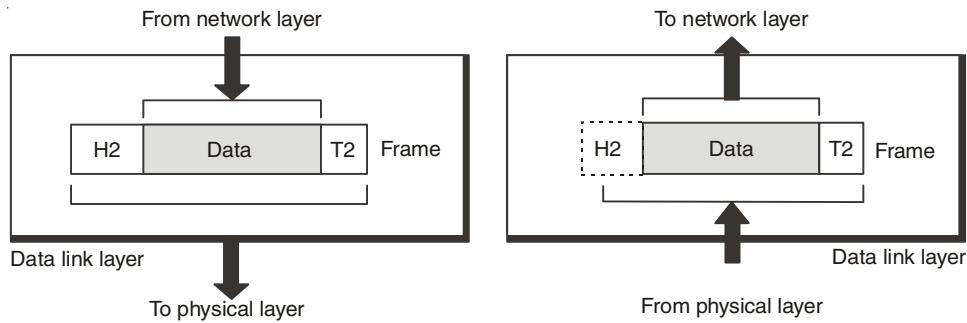


Fig. 1.15. Data Link layer.

Other responsibilities of the data link layer include the following:

- **Framing.** The data link layer divides the stream of bits received from the network layer into manageable data units called frames.
- **Physical addressing.** If frames are to be distributed to different systems on the network, the data link layer adds a header to the frame to define the sender and/or receiver of the frame. If the frame is intended for a system outside the sender's network, the receiver address is the address of the device that connects the network to the next one.
- **Flow control.** If the rate at which the data are absorbed by the receiver is less than the rate at which data are produced in the sender, the data link layer imposes a flow control mechanism to avoid overwhelming the receiver.

- **Error control.** The data link layer adds reliability to the physical layer by adding mechanisms to detect and retransmit damaged or lost frames. It also uses a mechanism to recognize duplicate frames. Error control is normally achieved through a trailer added to the end of the frame.
- **Access control.** When two or more devices are connected to the same link, data link layer protocols are necessary to determine which device has control over the link at any given time.

1.5.4.3 Network Layer

The network layer is responsible for the source-to-destination delivery of a packet, possibly across multiple networks (links). Whereas the data link layer oversees the delivery of the packet between two systems on the same network (links), the network layer ensures that each packet gets from its point of origin to its final destination.

If two systems are connected to the same link, there is usually no need for a network layer. However, if the two systems are attached to different networks (links) with connecting devices between the networks (links), there is often a need for the network layer to accomplish source-to-destination delivery. Figure 1.16 shows the relationship of the network layer to the data link and transport layers.

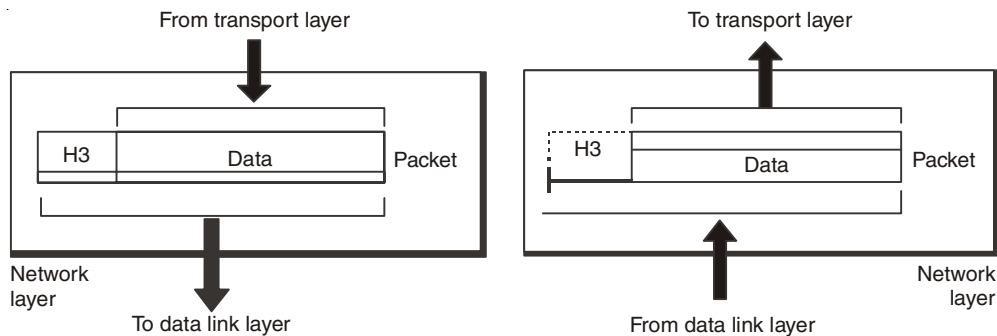


Fig. 1.16. Network layer.

Other responsibilities of the network layer include the following:

- **Logical addressing.** The physical addressing implemented by the data link layer handles the addressing problem locally. If a packet passes the network boundary, we need another addressing system to help distinguish the source and destination systems. The network layer adds a header to the packet coming from the upper layer that, among other things, includes the logical addresses of the sender and receiver.
- **Routing.** When independent networks or links are connected to create *internetworks* (network of networks) or a large network, the connecting devices (called *routers* or *switches*) *route or switch the packets to their final destination*. One of the functions of the network layer is to provide this mechanism

1.5.4.4 Transport Layer

The transport layer is responsible for process-to-process delivery of the entire message. A process is an application program running on a host. Whereas the network layer oversees

source-to-destination delivery of individual packets, it does not recognize any relationship between those packets. It treats each one independently, as though each piece belonged to a separate message, whether or not it does. The transport layer, on the other hand, ensures that the whole message arrives intact and in order, overseeing both error control and flow control at the source-to-destination level. Figure 1.17 shows the relationship of the transport layer to the network and session layers.

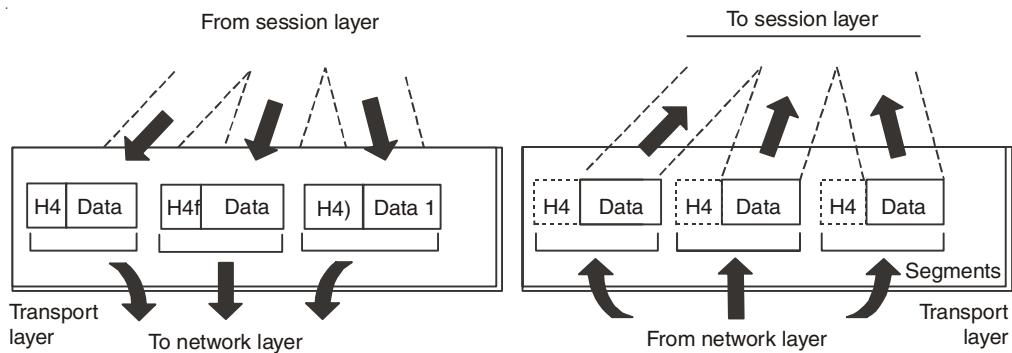


Fig. 1.17. Transport Layer.

Other responsibilities of the transport layer include the following:

- **Service-point addressing.** Computers often run several programs at the same time. For this reason, source-to-destination delivery means delivery not only from one computer to the next but also from a specific process (running program) on one computer to a specific process (running program) on the other. The transport layer header must therefore include a type of address called a *service-point address* (or port address). The network layer gets each packet to the correct computer; the transport layer gets the entire message to the correct process on that computer.
- **Segmentation and reassembly.** A message is divided into transmittable segments, with each segment containing a sequence number. These numbers enable the transport layer to reassemble the message correctly upon arriving at the destination and to identify and replace packets that were lost in transmission.
- **Connection control.** The transport layer can be either connectionless or connection oriented. A connectionless transport layer treats each segment as an independent packet and delivers it to the transport layer at the destination machine. A connection oriented transport layer makes a connection with the transport layer at the destination machine first before delivering the packets. After all the data are transferred, the connection is terminated.
- **Flow control.** Like the data link layer, the transport layer is responsible for flow control. However, flow control at this layer is performed end to end rather than across a single link.
- **Error control.** Like the data link layer, the transport layer is responsible for error control. However, error control at this layer is performed process-to process rather than across a single link. The sending transport layer makes sure that the entire message arrives at the receiving transport layer without error (damage, loss, or duplication). Error correction is usually achieved through retransmission.

1.5.4.5 Session Layer

The services provided by the first three layers (physical, data link, and network) are not sufficient for some processes. The session layer is the network *dialog controller*. It establishes, maintains, and synchronizes the interaction among communicating systems. Specific responsibilities of the session layer include the following:

- **Dialog control.** The session layer allows two systems to enter into a dialog. It allows the communication between two processes to take place in either halfduplex (one way at a time) or full-duplex (two ways at a time) mode.
- **Synchronization.** The session layer allows a process to add checkpoints, or synchronization points, to a stream of data.

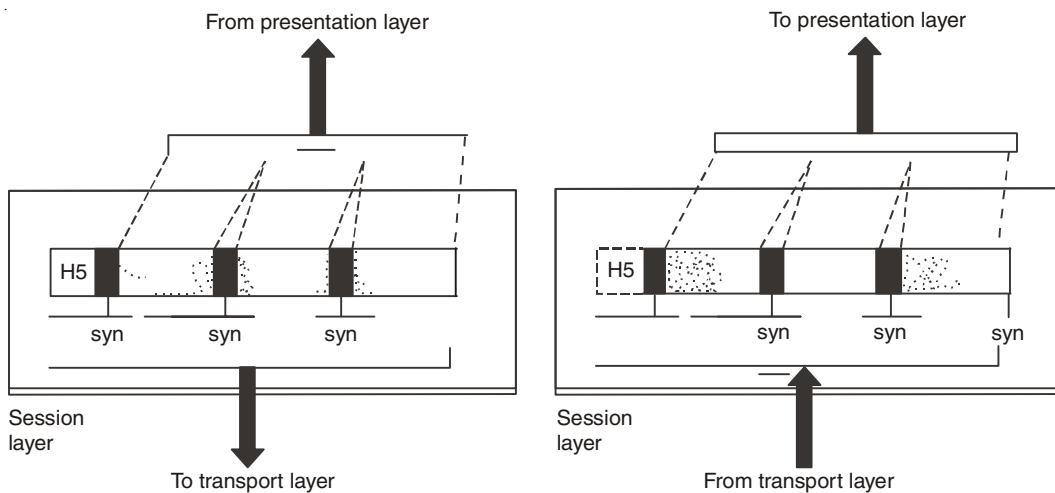


Fig. 1.18. Session Layer.

1.5.4.6. Presentation Layer

The presentation layer is concerned with the syntax and semantics of the information exchanged between two systems. Figure 1.19 shows the relationship between the presentation layer and the application and session layers.

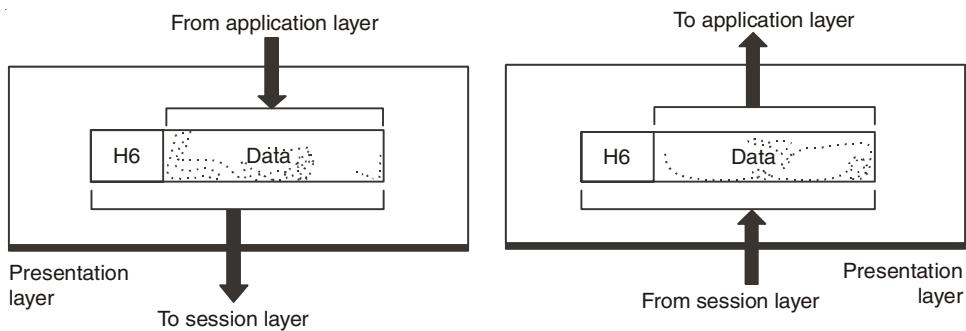


Fig. 1.19. Presentation Layer.

Specific responsibilities of the presentation layer include the following:

- **Translation.** The processes (running programs) in two systems are usually exchanging information in the form of character strings, numbers, and so on. The information must be changed to bit streams before being transmitted. Because different computers use different encoding systems, the presentation layer is responsible for interoperability between these different encoding methods. The presentation layer at the sender changes the information from its sender-dependent format into a common format. The presentation layer at the receiving machine changes the common format into its receiver-dependent format.
- **Encryption.** To carry sensitive information, a system must be able to ensure privacy. Encryption means that the sender transforms the original information to another form and sends the resulting message out over the network. Decryption reverses the original process to transform the message back to its original form.
- **Compression.** Data compression reduces the number of bits contained in the information. Data compression becomes particularly important in the transmission of multimedia such as text, audio, and video.

1.5.4.7 Application Layer

The application layer enables the user, whether human or software, to access the network. It provides user interfaces and support for services such as electronic mail, remote file access and transfer, shared database management, and other types of distributed information services.

1.5.5 The TCP/IP Reference Model

Let us now turn from the OSI reference model to the reference model used in the grandparent of all wide area computer networks, the ARPANET, and its successor, the worldwide Internet. Although we will give a brief history of the ARPANET later, it is useful to mention a few key aspects of it now.

The ARPANET was a research network sponsored by the DoD (U.S. Department of Defense). It eventually connected hundreds of universities and government installations, using leased telephone lines. When satellite and radio networks were added later, the existing protocols had trouble internetworking with them, so a new reference architecture was needed. Thus, the ability to connect multiple networks in a seamless way was one of the major design goals from the very beginning. This architecture later became known as the TCP/IP Reference Model, after its two primary protocols. It was first defined in (Cerf and Kahn, 1974). A later perspective is given in (Leiner et al., 1985). The design philosophy behind the model is discussed in (Clark, 1988).

Given the DoD's worry that some of its precious hosts, routers, and internetwork gateways might get blown to pieces at a moment's notice, another major goal was that the network be able to survive loss of subnet hardware, with existing conversations not being broken off. In other words, DoD wanted connections to remain intact as long as the source and destination machines were functioning, even if some of the machines or transmission lines in between were suddenly put out of operation. Furthermore, a flexible architecture

was needed since applications with divergent requirements were envisioned, ranging from transferring files to real-time speech transmission.

1.5.5.1 The Internet Layer

All these requirements led to the choice of a packet-switching network based on a connectionless inter network layer. This layer, called the internet layer, is the linchpin that holds the whole architecture together. Its job is to permit hosts to inject packets into any network and have them travel independently to the destination (potentially on a different network). They may even arrive in a different order than they were sent, in which case it is the job of higher layers to rearrange them, if in-order delivery is desired. Note that "internet" is used here in a generic sense, even though this layer is present in the Internet.

The analogy here is with the (snail) mail system. A person can drop a sequence of international letters into a mail box in one country, and with a little luck, most of them will be delivered to the correct address in the destination country. Probably the letters will travel through one or more international mail gateways along the way, but this is transparent to the users. Furthermore, that each country (*i.e.*, each network) has its own stamps, preferred envelope sizes, and delivery rules is hidden from the users.

The internet layer defines an official packet format and protocol called IP (Internet Protocol). The job of the internet layer is to deliver IP packets where they are supposed to go. Packet routing is clearly the major issue here, as is avoiding congestion. For these reasons, it is reasonable to say that the TCP/IP internet layer is similar in functionality to the OSI network layer. Figure 1.20 shows this correspondence.

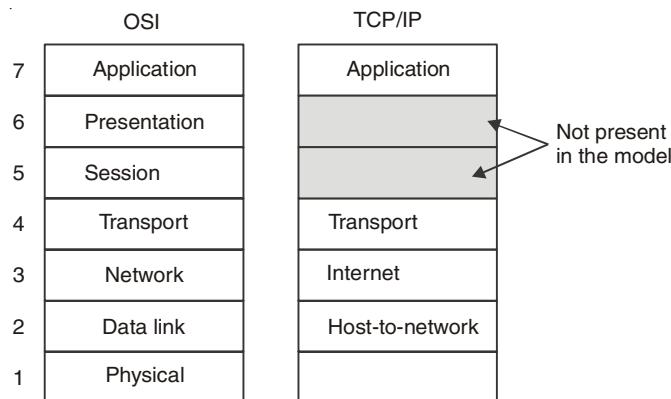


Fig. 1.20. TCP/IP Reference Model.

1.5.5.2 The Transport Layer

The layer above the internet layer in the TCP/IP model is now usually called the transport layer. It is designed to allow peer entities on the source and destination hosts to carry on a conversation, just as in the OSI transport layer. Two end-to-end transport protocols have been defined here. The first one, TCP (Transmission Control Protocol), is a reliable connection-oriented protocol that allows a byte stream originating on one machine to be delivered without error on any other machine in the internet. It fragments the incoming byte stream into discrete messages and passes each one on to the internet layer. At the destination, the

receiving TCP process reassembles the received messages into the output stream. TCP also handles flow control to make sure a fast sender cannot swamp a slow receiver with more messages than it can handle.

The second protocol in this layer, UDP (User Datagram Protocol), is an unreliable, connectionless protocol for applications that do not want TCP's sequencing or flow control and wish to provide their own. It is also widely used for one-shot, client-server-type request-reply queries and applications in which prompt delivery is more important than accurate delivery, such as transmitting speech or video.

The relation of IP, TCP, and UDP is shown in Fig. 1.21. Since the model was developed, IP has been implemented on many other networks.

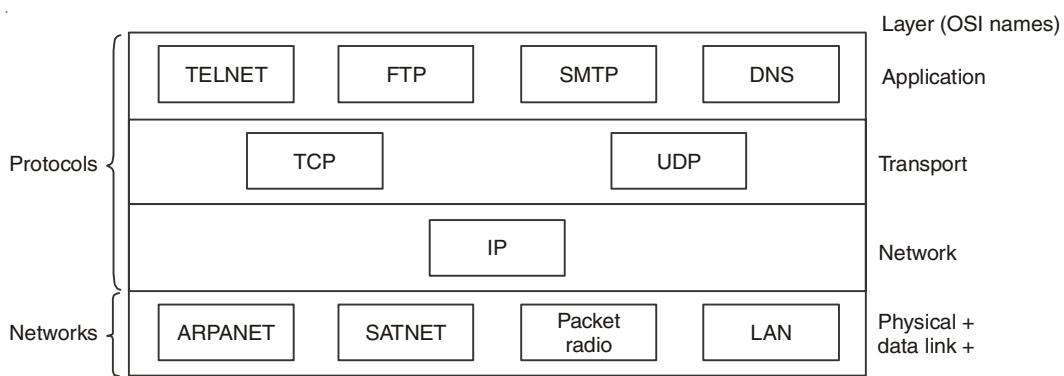


Fig. 1.21. Protocols in TCP/IP Model initially.

1.5.5.3 The Application Layer

The TCP/IP model does not have session or presentation layers. No need for them was perceived, so they were not included. Experience with the OSI model has proven this view correct: they are of little use to most applications.

On top of the transport layer is the application layer. It contains all the higher-level protocols. The early ones included virtual terminal (TELNET), file transfer (FTP), and electronic mail (SMTP), as shown in Fig. 1.21. The virtual terminal protocol allows a user on one machine to log onto a distant machine and work there. The file transfer protocol provides a way to move data efficiently from one machine to another. Electronic mail was originally just a kind of file transfer, but later a specialized protocol (SMTP) was developed for it. Many other protocols have been added to these over the years: the Domain Name System (DNS) for mapping host names onto their network addresses, NNTP, the protocol for moving USENET news articles around, and HTTP, the protocol for fetching pages on the World Wide Web, and many others.

1.5.5.4 The Host-to-Network Layer

Below the internet layer is a great void. The TCP/IP reference model does not really say much about what happens here, except to point out that the host has to connect to the network using some protocol so it can send IP packets to it. This protocol is not defined and varies from host to host and network to network.

1.5.5.5 Addressing

Four levels of addresses are used in an internet employing the TCP/IP protocols: physical (link) addresses, logical (IP) addresses, port addresses, and specific addresses . Each address is related to a specific layer in the TCP/IP architecture.

Physical Addresses

The physical address, also known as the link address, is the address of a node as defined by its LAN or WAN. It is included in the frame used by the data link layer. It is the lowest-level address. The physical addresses have authority over the network (LAN or WAN). The size and format of these addresses vary depending on the network. For example, Ethernet uses a 6-byte (48-bit) physical address that is imprinted on the network interface card (NIC). LocalTalk (Apple), however, has a 1-byte dynamic address that changes each time the station comes up.

Logical Addresses

Logical addresses are necessary for universal communications that are independent of underlying physical networks. Physical addresses are not adequate in an inter network environment where different networks can have different address formats. A universal addressing system is needed in which each host can be identified uniquely, regardless of the underlying physical network.

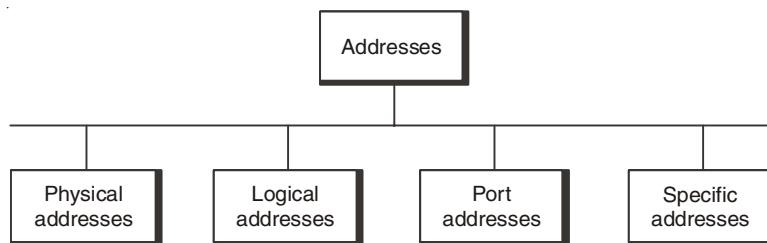


Fig. 1.22. Addresses in TCP/IP.

Port Addresses

The IP address and the physical address are necessary for a quantity of data to travel from a source to the destination host. However, arrival at the destination host is not the final objective of data communications on the Internet.

A system that sends nothing but data from one computer to another is not complete. Today, computers are devices that can run multiple processes at the same time. The end objective of Internet communication is a process communicating with another process. For example, computer A can communicate with computer C by using TELNET. At the same time, computer A communicates with computer B by using the File Transfer Protocol (FTP). For these processes to receive data simultaneously, we need a method to label the different processes.

In other words, they need addresses. In the TCPIIP architecture, the label assigned to a process is called a port address. A port address in TCP/IP is 16 bits in length.

Specific Addresses

Some applications have user-friendly addresses that are designed for that specific address. Examples include the e-mail address (for example, president@rb.nic.in) and the Universal Resource Locator (URL) (for example, www.pmoindia.nic.in). The first defines the recipient of an e-mail; the second is used to find a document on the World Wide Web. These addresses, however, get changed to the corresponding port and logical addresses by the sending computer.

1.5.6 A Comparison of the OSI and TCP/IP Reference Models

The OSI and TCP/IP reference models have much in common. Both are based on the concept of a stack of independent protocols. Also, the functionality of the layers is roughly similar. For example, in both models the layers up through and including the transport layer are there to provide an end-to-end, network-independent transport service to processes wishing to communicate. These layers form the transport provider. Again in both models, the layers above transport are application-oriented users of the transport service.

Despite these fundamental similarities, the two models also have many differences. In this section, we will focus on the key differences between the two reference models. It is important to note that we are comparing the reference models here, not the corresponding protocol stacks. The protocols themselves will be discussed later.

Three concepts are central to the OSI model:

1. Services.
2. Interfaces.
3. Protocols.

Probably the biggest contribution of the OSI model is to make the distinction between these three concepts explicit. Each layer performs some services for the layer above it. The service definition tells what the layer does, not how entities above it access it or how the layer works. It defines the layer's semantics.

A layer's interface tells the processes above it how to access it. It specifies what the parameters are and what results to expect. It, too, says nothing about how the layer works inside.

Finally, the peer protocols used in a layer are the layer's own business. It can use any protocols it wants to, as long as it gets the job done (*i.e.*, provides the offered services). It can also change them at will without affecting software in higher layers.

The TCP/IP model did not originally clearly distinguish between service, interface, and protocol, although people have tried to retrofit it after the fact to make it more OSI-like. For example, the only real services offered by the internet layer are SEND IP PACKET and RECEIVE IP PACKET.

As a consequence, the protocols in the OSI model are better hidden than in the TCP/IP model and can be replaced relatively easily as the technology changes. Being able to make such changes is one of the main purposes of having layered protocols in the first place.

The OSI reference model was devised before the corresponding protocols were invented. This ordering means that the model was not biased toward one particular set of protocols, a fact that made it quite general. The downside of this ordering is that the designers did not have much experience with the subject and did not have a good idea of which functionality to put in which layer.

With TCP/IP the reverse was true: the protocols came first, and the model was really just a description of the existing protocols. There was no problem with the protocols fitting the model. They fit perfectly. The only trouble was that the model did not fit any other protocol stacks. Consequently, it was not especially useful for describing other, non-TCP/IP networks.

Turning from philosophical matters to more specific ones, an obvious difference between the two models is the number of layers: the OSI model has seven layers and the TCP/IP has four layers. Both have (inter)network, transport, and application layers, but the other layers are different.

Another difference is in the area of connectionless versus connection-oriented communication. The OSI model supports both connectionless and connection-oriented communication in the network layer, but only connection-oriented communication in the transport layer, where it counts (because the transport service is visible to the users). The TCP/IP model has only one mode in the network layer (connectionless) but supports both modes in the transport layer, giving the users a choice. This choice is especially important for simple request-response protocols.

1.6 NETWORK SERVICES

For a communication network to work effectively, data in the network must be able to move from one network element to another. This can only happen if the network services to move such data work. For data networks, these services fall into two categories:

- Connection services to facilitate the exchange of data between the two network communicating end-systems with as little data loss as possible and in as little time as possible.
- Switching services to facilitate the movement of data from host to host across the length and width of the network mesh of hosts, hubs, bridges, routers, and gateways.

1.6.1 Connection Services

How do we get the network transmitting elements to exchange data over the network? Two types of connection services are used: the *connected-oriented* and *connectionless* services.

1.6.1.1 Connected-Oriented Services

With a connection-oriented service, before a client can send packets with real data to the server, there must be a *three-way handshake*. We will define this three-way handshake in later chapters. But the purpose of a three-way handshake is to establish a session before the actual communication can begin. Establishing a session before data is moved creates a path

of virtual links between the end systems through a network and therefore, guarantees the reservation and establishment of fixed communication channels and other resources needed for the exchange of data before any data is exchanged and as long as the channels are needed. For example, this happens whenever we place telephone calls; before we exchange words, the channels are reserved and established for the duration. Because this technique guarantees that data will arrive in the same order it was sent in, it is considered to be reliable. In short the service offers the following:

- Acknowledgments of all data exchanges between the end-systems,
- Flow control in the network during the exchange, and
- Congestion control in the network during the exchange.

Depending on the type of physical connections in place and the services required by the systems that are communicating, connection-oriented methods may be implemented in the data link layers or in the transport layers of the protocol stack, although the trend now is to implement it more at the transport layer. For example, TCP is a connection-oriented transport protocol in the transport layer. Other network technologies that are connection-oriented include the frame relay and ATMs.

1.6.1.2 Connectionless Service

In a connectionless service, there is no handshaking to establish a session between the communicating end-systems, no flow control, and no congestion control in the network. This means that a client can start communicating with a server without warning or inquiry for readiness; it simply sends streams of packets, called datagrams, from its sending port to the server's connection port in single point-to-point transmissions with no relationship established between the packets and between the end-systems. There are advantages and of course disadvantages to this type of connection service. In brief, the connection is faster because there is no handshaking which can sometimes be time consuming, and it offers periodic burst transfers with large quantities of data and, in addition, it has simple protocol. However, this service offers minimum services, no safeguards and guarantees to the sender since there is no prior control information and no acknowledgment. In addition, the service does not have the reliability of the connection-oriented method, and offers no error handling and no packets ordering; in addition, each packet self-identifies that leads to long headers, and finally, there is no predefined order in the arrival of packets. Like the connection-oriented method, this service can operate both at the data link and transport layers. For example, UDP, a connectionless service, operates at the transport layer.

1.6.2 Network Switching Services

Let us take a detour and briefly discuss data transfer by a switching element. This is a technique by which data is moved from host to host across the length and width of the network mesh of hosts, hubs, bridges, routers, and gateways. This technique is referred to as *data switching*. The type of data switching technique used by a network determines how messages are transmitted between the two communicating elements and across that network. There are two types of data switching techniques: *circuit switching* and *packet switching*.

1.6.2.1 Circuit Switching

In circuit switching networks, one must reserve all the resources before setting up a physical communication channel needed for communication. The physical connection, once established, is then used exclusively by the two end-systems, usually subscribers, for the duration of the communication. The main feature of such a connection is that it provides a fixed data rate channel, and both subscribers must operate at this rate. For example, in a telephone communication network, a connected line is reserved between the two points before the users can start using the service. One issue of debate on circuit switching is the perceived waste of resources during the so-called silent periods when the connection is fully in force but not being used by the parties. This situation occurs when, for example, during a telephone network session, a telephone receiver is not hung up after use, leaving the connection still established. During this period, while no one is utilizing the session, the session line is still open.

1.6.2.2 Packet Switching

Packet switching networks, on the other hand, do not require any resources to be reserved before a communication session begins. These networks, however, require the sending host to assemble all data streams to be transmitted into packets. If a message is large, it is broken into several packets. Packet headers contain the source and the destination network addresses of the two communicating end-systems. Then, each of the packets is sent on the communication links and across packet switches (routers). On receipt of each packet, the router inspects the destination address contained in the packet. Using its own routing table, each router then forwards the packet on the appropriate link at the maximum available bit rate. As each packet is received at each intermediate router, it is forwarded on the appropriate link interspersed with other packets being forwarded on that link. Each router checks the destination address, if it is the owner of the packet; it then reassembles the packets into the final message. Figure 1.23 shows the role of routers in packet switching networks.

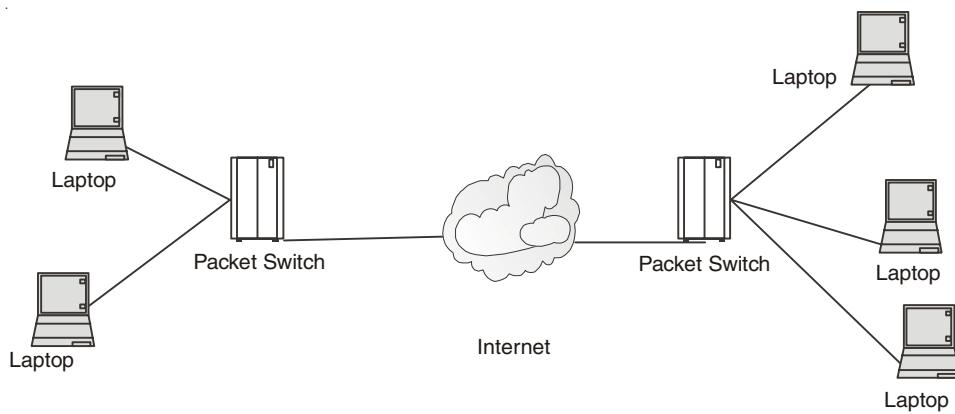


Fig. 1.23. Packet switching Network.

Packet switches are considered to be store-and-forward transmitters, meaning that they must receive the entire packet before the packet is retransmitted or switched on to the next switch. Because there is no predefined route for these packets, there can be unpredictably

long delays before the full message can be re-assembled. In addition, the network may not dependably deliver all the packets to the intended destination. To ensure that the network has a reliably fast transit time, a fixed maximum length of time is allowed for each packet.

Packet switching networks suffer from a few problems, including the following:

- The rate of transmission of a packet between two switching elements depends on the maximum rate of transmission of the link joining them and on the switches themselves.
- Momentary delays are always introduced whenever the switch is waiting for a full packet. The longer the packet, the longer the delay.
- Each switching element has a finite buffer for the packets. It is thus possible for a packet to arrive only to find the buffer full with other packets. Whenever this happens, the newly arrived packet is not stored but gets lost, a process called *packet dropping*. In peak times, servers may drop a large number of packets. Congestion control techniques use the rate of packet drop as one measure of traffic congestion in a network.

Packet switching networks are commonly referred to as *packet networks* for obvious reasons. They are also called *asynchronous* networks and in such networks, packets are ideal because there is a sharing of the bandwidth, and of course, this avoids the hassle of making reservations for any anticipated transmission. There are two types of packet switching networks:

- *virtual circuit network* in which a packet route is planned, and it becomes a logical connection before a packet is released and
- *datagram network*, which will be discussed in later part of this book.

1.7. EXAMPLES OF NETWORK ARCHITECTURES

1.7.1. IBM's System Network Architecture (SNA)

IBM's System network architecture is a method for unifying network operations. SNA describes the division of network functions into discrete layers and defines protocols and formats for communication between equivalent layers. SNA describes a network in terms of a physical network and a logical network. The physical network consists of a collection of *nodes*: host node, front end communication node, concentration node and terminal node. The host node is the central processor; the front end communication node is concerned with data transmission functions; the concentration node supervises the behaviour of terminals and other peripherals; the terminal node is concerned with the input and output of information through terminal devices. Figure 1.24 depicts a simple SNA network.

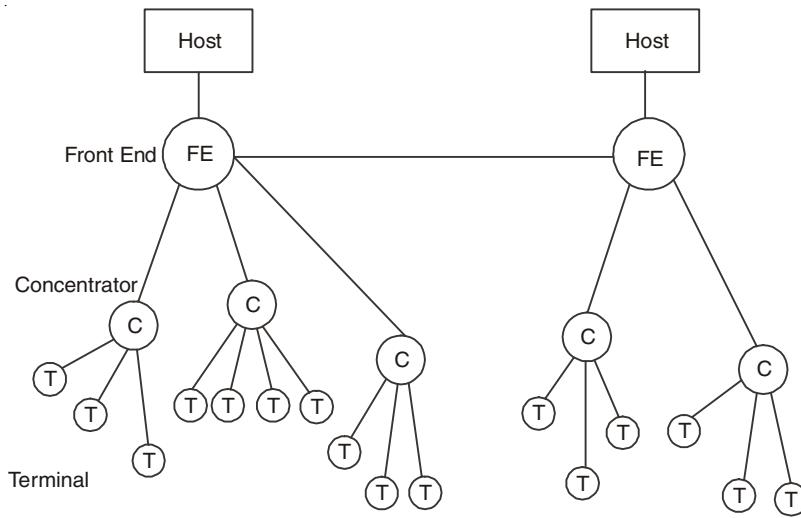


Fig. 1.24. A simple SNA Network.

The SNA logical network consists of three layers:

- (i) transmission management;
- (ii) function management; and
- (iii) application.

Each node in the SNA physical network may contain any or all of these three layers. Communication between layers is as shown in Fig. 1.25 below. The application layer consists of the user's application programs and is concerned only with the processing of information.

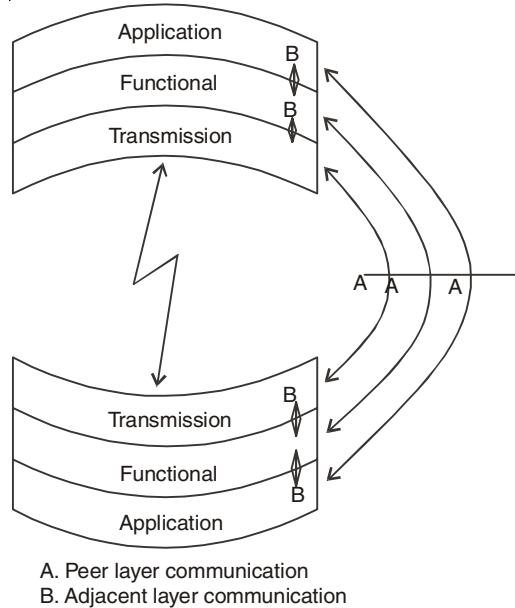


Fig. 1.25. Communication between layers in SNA Network.

The functional management layers controls the presentation format of information sent from and received by the application layer, *i.e.*, it converts the data into a form convenient to the user. The transmission management layer controls movement of user data through the network. It involves routing, scheduling and transmission functions. This layer exists in every intermediate node through which the data units flow and may utilise a variety of physical connections and protocols between the nodes of an SNA network.

Each node contains one or more *Network Addressable Units* (NAU). There are three types of NAU. A *Logical Unit* (LU) is a NAU which users use to address their process. A *Physical Unit* (PU) is a NAU which the network uses to address a physical device, without reference to which processes are using it. The third type of NAU is the *System Services Control Point* (SSCP) which has control over all front ends, remote concentrators and terminals attached to the host. The three types of NAU communicate with each other by invoking the services of the transmission management layer. The *physical link control* protocol takes care of electrical transmission of data bits from one node to another. The *data link control* protocol constructs frames from the data stream, detecting and recovering from transmission errors. This level 2 protocol is called SDLC (Synchronous Data Link Control) which we will look at later. The *path control* protocol performs the path-selection and congestion control functions within the subnet. The *transmission control* protocol initiates, recovers and terminates transport connections (called sessions) in SNA. It also controls the flow of data to and from other NAUs.

The SNA and ISO models do not correspond closely. However, a rough comparison of the architecture control levels is shown in Fig. 1.26.

Layer	OSI	SNA	DECNET
7	Application	End User	Application
6	Presentation	NAU services	
5	Session	Data Flow Control	(none)
4		Transmission Control	
3	Network	Path Control	Network Services
2	Data Link	Data Link	Transport
1	Physical	Physical	Physical

Fig. 1.26. Comparison of OSI/SNA/DECNET.

1.7.2. DECNET's DNA (Digital Network Architecture)

A DECNET is just a collection of computers (nodes) whose functions include running user programs, performing packet switching or both. The architecture of DECNET is called *Digital Network Architecture* (DNA).

DNA has five layers. The physical layer, data link control layer, transport layer and network services layer correspond to the lowest four OSI layers. DNA does not have a session layer (layer 5) and its application layer corresponds to a mixture of the presentation and application layers in OSI; see Fig. 1.26 notice that DNA's level 3 is called the transport layer, not level 4 as in the OSI Model.

Review Questions

1. Identify the five components of a data communications system.
2. What are the advantages of distributed processing?
3. What are the three criteria necessary for an effective and efficient network?
4. What are the advantages of a multipoint connection over a point-to-point connection?
5. What are the two types of line configuration?
6. What is an internet?
7. Why are protocols needed?
8. Why are standards needed?
9. List the layers of the Internet model.
10. Which layers in the Internet model are the network support layers?
11. Which layer in the Internet model is the user support layer?
12. What is the difference between network layer delivery and transport layer delivery?
13. Suppose a computer sends a frame to another computer on a bus topology LAN. The physical destination address of the frame is corrupted during the transmission. What happens to the frame? How can the sender be informed about the situation?
14. Suppose a computer sends a packet at the network layer to another computer somewhere in the Internet. The logical destination address of the packet is corrupted. What happens to the packet? How can the source computer be informed of the situation?



CHAPTER 2

THE PHYSICAL LAYER

The hand that hath made you fair hath made you good.

—William Shakespeare

2.1 INTRODUCTION

In this chapter we will look at the lowest layer depicted in the hierarchy of OSI Model. It defines the mechanical, electrical, and timing interfaces to the network. It is the layer that actually interacts with the transmission media, the physical part of the network that connects network components together. This layer is involved in physically carrying information from one node in the network to the next.

The physical layer has complex tasks to perform. One major task is to provide services for the data link layer. The data in the data link layer consists of 0's and 1's organized into frames that are ready to be sent across the transmission medium. This stream of 0's and 1's must first be converted into another entity: signals. One of the services provided by the physical layer is to create a signal that represents this stream of bits. The physical layer must also take care of the physical network, the transmission medium. The transmission medium is a passive entity; it has no internal program or logic for control like other layers.

The material covered in this chapter will provide background information on the key transmission technologies used in modern networks.

2.2 THE THEORETICAL BASIS FOR DATA COMMUNICATION

Information can be transmitted on wires by varying some physical property such as voltage or current. By representing the value of this voltage or current as a single-valued function of time, $f(t)$, we can model the behavior of the signal and analyze it mathematically. This analysis is the subject of the following sections.

2.2.1 Fourier Analysis

In the early 19th century, the French mathematician Jean-Baptiste Fourier proved that any reasonably behaved periodic function, $g(t)$ with period T can be constructed as the sum of a (possibly infinite) number of sines and cosines:

$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \sin(2\pi nft) + \sum_{n=t}^{\infty} b_n \cos(2\pi nft) \quad \dots(1)$$

where $f = 1/T$ is the fundamental frequency, a_n and b_n are the sine and cosine amplitudes of the n th harmonics (terms), and c is a constant. Such a decomposition is called a Fourier series. From the Fourier series, the function can be reconstructed; that is, if the period, T , is known and the amplitudes are given, the original function of time can be found by performing the sums of equation.

A data signal that has a finite duration (which all of them do) can be handled by just imagining that it repeats the entire pattern over and over forever (*i.e.*, the interval from T to $2T$ is the same as from 0 to T , etc.). The a_n amplitudes can be computed for any given $g(t)$ by multiplying both sides of Eq. 1 by $\sin(2\pi kft)$ and then integrating from 0 to T .

$$\int_0^T \sin(2\pi kft) \sin(2\pi nft) dt = \begin{cases} 0 & \text{for } k \neq n \\ T/2 & \text{for } k = n \end{cases} \quad \dots(2)$$

Since only one term of the summation survives: a_n . The b_n summation vanishes completely. Similarly, by multiplying Eq. 1 by $\cos(2\pi kft)$ and integrating between 0 and T , we can derive b_n . By just integrating both sides of the equation as it stands, we can find c . The results of performing these operations are as follows:

$$a_n = \frac{2}{T} \int_0^T g(t) \sin(2\pi nft) dt \quad b_n = \frac{2}{T} \int_0^T g(t) \cos(2\pi nft) dt \quad c = \frac{2}{T} \int_0^T g(t) dt$$

2.2.2 Bandwidth-Limited Signals

To see what all this has to do with data communication, let us consider a specific example: the transmission of the ASCII character “*b*” encoded in an 8-bit byte. The bit pattern that is to be transmitted is 01100010. The left-hand part of Fig. 2.1(a) shows the voltage output by the transmitting computer. The Fourier analysis of this signal yields the coefficients

$$a_n = \frac{1}{\pi n} [\cos(\pi n/4) - \cos(3\pi n/4) + \cos(6\pi n/4) - \cos(7\pi n/4)]$$

$$b_n = \frac{1}{\pi n} [\sin(3\pi n/4) - \sin(\pi n/4) + \sin(7\pi n/4) - \sin(6\pi n/4)]$$

$$c = 3/4$$

The root-mean-square amplitudes, for the first few terms are shown on the right-hand side of Fig. 2.1 (a). These values are of interest because their squares are proportional to the energy transmitted at the corresponding frequency.

No transmission facility can transmit signals without losing some power in the process. If all the Fourier components were equally diminished, the resulting signal would be reduced in amplitude but not distorted [*i.e.*, it would have the same nice squared-off shape as Fig. 2.1(a)]. Unfortunately, all transmission facilities diminish different Fourier components by different amounts, thus introducing distortion. Usually, the amplitudes are transmitted undiminished from 0 up to some frequency f_c [measured in cycles/sec or Hertz (Hz)] with all frequencies above this cutoff frequency attenuated. The range of frequencies transmitted without being strongly attenuated is called the bandwidth. In practice, the cutoff is not really sharp, so often the quoted bandwidth is from 0 to the frequency at which half the power gets through.

The bandwidth is a physical property of the transmission medium and usually depends on the construction, thickness, and length of the medium. In some cases a filter is introduced into the circuit to limit the amount of bandwidth available to each customer. For example, a telephone wire may have a bandwidth of 1 MHz for short distances, but telephone companies add a filter restricting each customer to about 3100 Hz. This bandwidth is adequate for intelligible speech and improves system-wide efficiency by limiting resource usage by customers.

Now let us consider how the signal of Fig. 2.1(a) would look if the bandwidth were so low that only the lowest frequencies were transmitted [*i.e.*, if the function were being approximated by the first few terms of Eq. 1. Figure 2.1 (b) shows the signal that results from a channel that allows only the first harmonic (the fundamental, f) to pass through. Similarly, Fig. 2.1 (c)-(e) show the spectra and reconstructed functions for higher-bandwidth channels.

Given a bit rate of b bits/sec, the time required to send 8 bits (for example) 1 bit at a time is $8/b$ sec, so the frequency of the first harmonic is $b/8$ Hz. An ordinary telephone line, often called a voice-grade line, has an artificially-introduced cutoff frequency just above 3000 Hz. This restriction means that the number of the highest harmonic passed through is roughly $3000/(b/8)$ or $24,000/b$, (the cutoff is not sharp).

For some data rates, the numbers work out as shown in Fig. 2.2. From these numbers, it is clear that trying to send at 9600 bps over a voice-grade telephone line will transform Fig. 2.1 (a) into something looking like Fig. 2.1 (c), making accurate reception of the original binary bit stream tricky. It should be obvious that at data rates much higher than 38.4 kbps, there is no hope at all for binary signals, even if the transmission facility is completely noiseless. In other words, limiting the bandwidth limits the data rate, even for perfect channels. However, sophisticated coding schemes that make use of several voltage levels do exist and can achieve higher data rates. We will discuss these later in this chapter.

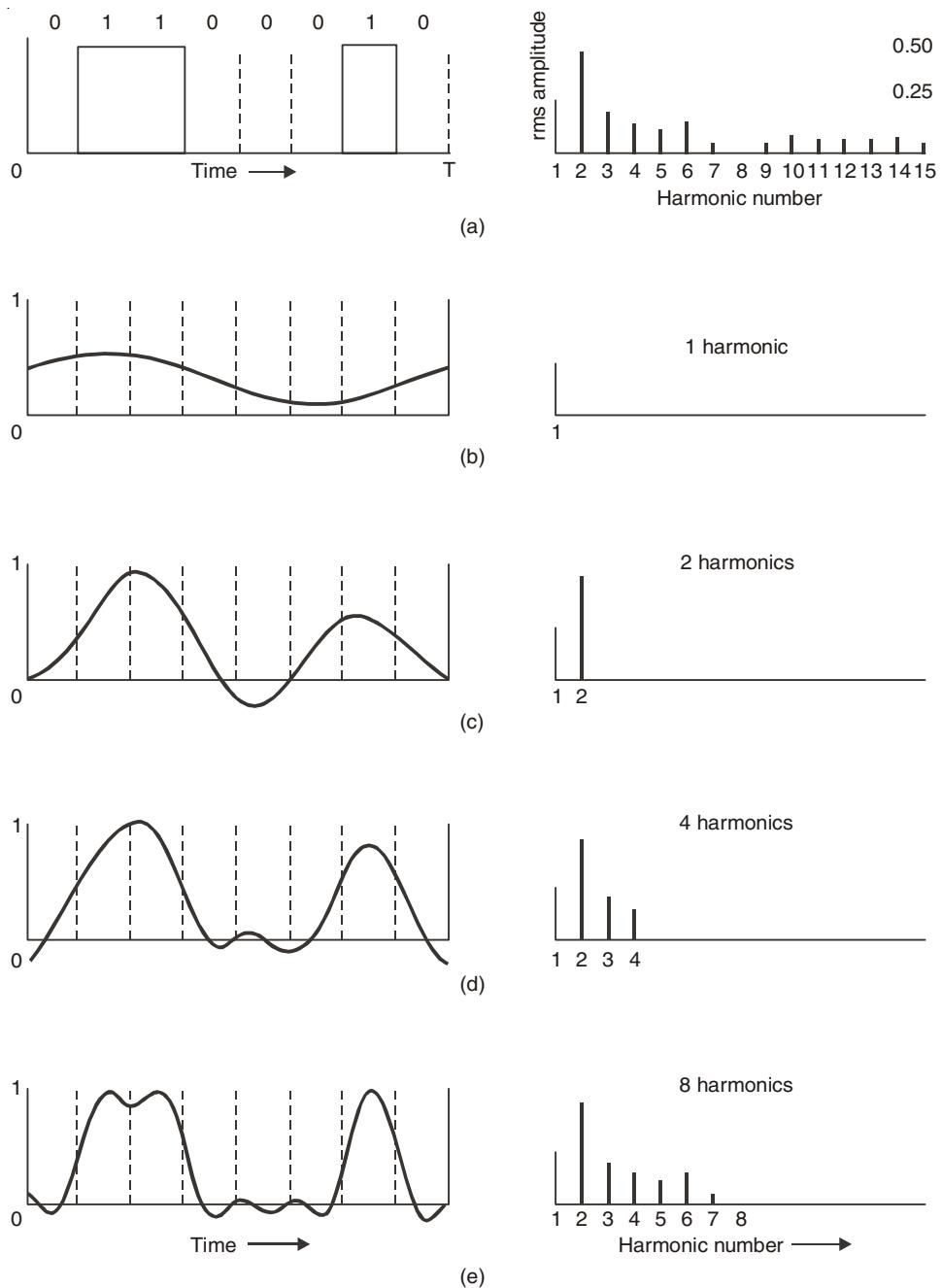


Fig. 2.1. (a) A binary signal and its root-mean-square Fourier amplitudes.
 (b)-(e) Successive approximations to the original signal.

Bps	T (msec)	First harmonic (Hz)	# Harmonics sent
300	26.67	37.5	80
600	13.33	75	40
1200	6.67	150	20
2400	3.33	300	10
4800	1.67	600	5
9600	0.83	1200	2
19200	0.42	2400	1
38400	0.21	4800	0

Fig. 2.2. Relation between data rate and harmonics.

2.2.3 The Maximum Data Rate of a Channel

As early as 1924, an AT&T engineer, Henry Nyquist, realized that even a perfect channel has a finite transmission capacity. He derived an equation expressing the maximum data rate for a finite bandwidth noiseless channel. In 1948, Claude Shannon carried Nyquist's work further and extended it to the case of a channel subject to random (that is, thermodynamic) noise (Shannon, 1948). We will just briefly summarize their now classical results here.

Nyquist proved that if an arbitrary signal has been run through a low-pass filter of bandwidth H , the filtered signal can be completely reconstructed by making only $2H$ (exact) samples per second. Sampling the line faster than $2H$ times per second is pointless because the higher frequency components that such sampling could recover have already been filtered out. If the signal consists of V discrete levels, Nyquist's theorem states:

$$\text{maximum data rate} = 2H \log_2 V \text{ bits/sec}$$

For example, a noiseless 3-kHz channel cannot transmit binary (*i.e.*, two-level) signals at a rate exceeding 6000 bps. So far we have considered only noiseless channels. If random noise is present, the situation deteriorates rapidly. And there is always random (thermal) noise present due to the motion of the molecules in the system. The amount of thermal noise present is measured by the ratio of the signal power to the noise power, called the signal-to-noise ratio. If we denote the signal power by S and the noise power by N , the signal-to-noise ratio is S/N . Usually, the ratio itself is not quoted; instead, the quantity $10 \log_{10} S/N$ is given. These units are called decibels (dB). An S/N ratio of 10 is 10 dB, a ratio of 100 is 20 dB, a ratio of 1000 is 30 dB, and so on. The manufacturers of stereo amplifiers often characterize the bandwidth (frequency range) over which their product is linear by giving the 3-dB frequency on each end. These are the points at which the amplification factor has been approximately halved (because $\log_{10} 3 \approx 0.5$).

Shannon's major result is that the maximum data rate of a noisy channel whose bandwidth is H Hz, and whose signal-to-noise ratio is S/N , is given by

$$\text{maximum number of bits/sec} = H \log_2 (1 + S/N)$$

For example, a channel of 3000-Hz bandwidth with a signal to thermal noise ratio of 30 dB (typical parameters of the analog part of the telephone system) can never transmit much more than 30,000 bps, no matter how many or how few signal levels are used and no matter how often or how infrequently samples are taken. Shannon's result was derived from information-theory arguments and applies to any channel subject to thermal noise. Counter examples should be treated in the same category as perpetual motion machines. It should be noted that this is only an upper bound and real systems rarely achieve it.

2.3 DATA COMMUNICATION TRANSMISSION TECHNOLOGY

The media through which information has to be transmitted determine the signal to be used. Some media permit only analog signals. Some allow both analog and digital. Therefore, depending on the media type involved and other considerations, the input data can be represented as either *digital* or *analog* signal. In an analog format, data is sent as continuous electromagnetic waves on an interval representing things such as voice and video and propagated over a variety of media that may include copper wires, twisted coaxial pair or cable, fiber optics, or wireless. We will discuss these media soon. In a digital format, on the other hand, data is sent as a digital signal, a sequence of voltage pulses that can be represented as a stream of binary bits. Both analog and digital data can be propagated and many times represented as either analog or digital.

Transmission itself is the propagation and processing of data signals between network elements. The concept of representation of data for transmission, either as analog or digital signal, is called an *encoding scheme*. Encoded data is then transmitted over a suitable transmission medium that connects all network elements.

There are two encoding schemes, *analog* and *digital*. Analog encoding propagates analog signals representing analog data such as sound waves and voice data. Digital encoding, on the other hand, propagates digital signals representing either an analog or a digital signal representing digital data of binary streams by two voltage levels.

2.3.1 Analog Encoding of Digital Data

Recall that digital information is in the form of 1s or 0s. To send this information over some analog medium such as the telephone line, for example, which has limited bandwidth, digital data needs to be encoded using modulation and demodulation to produce analog signals. The encoding uses a continuous oscillating wave, usually a sine wave, with a constant frequency signal called a *carrier* signal. The carrier has three modulation characteristics: *amplitude*, *frequency*, and *phase shift*.

The scheme then uses a *modem*, a modulation–demodulation pair, to modulate and demodulate the data signal based on any one of the three carrier characteristics or a

combination. The resulting wave is between a range of frequencies on both sides of the carrier as shown below:

- *Amplitude* modulation represents each binary value by a different amplitude of the carrier frequency. The absence of or low carrier frequency may represent a 0 and any other frequency then represents a 1. But this is a rather inefficient modulation technique and is therefore used only at low frequencies up to 1200 bps in voice grade lines.
- *Frequency* modulation also represents the two binary values by two different frequencies close to the frequency of the underlying carrier. Higher frequencies represent a 1 and low frequencies represent a 0. The scheme is less susceptible to errors.
- *Phase shift* modulation changes the timing of the carrier wave, shifting the carrier phase to encode the data. A 1 is encoded as a change in phase by 180 degrees and a 0 may be encoded as a 0 change in phase of a carrier signal. This is the most efficient scheme of the three and it can reach a transmission rate of up to 9600 bps.

2.3.2 Digital Encoding of Digital Data

In this encoding scheme, which offers the most common and easiest way to transmit digital signals, two binary digits are used to represent two different voltages. Within a computer, these voltages are commonly 0 volt and 5 volts. Another procedure uses two representation codes: *nonreturn to zero level* (NRZ-L), in which negative voltage represents binary one and positive voltage represents binary zero, and *nonreturn to zero, invert on ones* (NRZ-I). See Figs. 2.3 and 2.4 for an example of these two codes.

In NRZ-L, whenever a 1 occurs, a transition from one voltage level to another is used to signal the information. One problem with NRZ signaling techniques is the requirement of a perfect synchronization between the receiver and transmitter clocks. This is, however, reduced by sending a separate clock signal. There are yet other representations such as the Manchester and differential Manchester, which encode clock information along with the data.

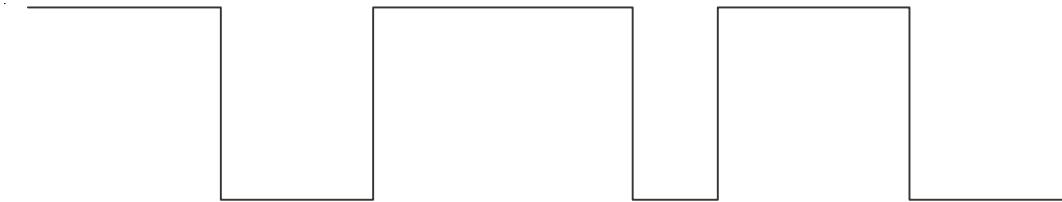


Fig. 2.3. NRZ-L N Nonreturn to zero level representation code.

One may wonder why go through the hassle of digital encoding and transmission. There are several advantages over its cousin, analog encoding. These include the following:

- Plummeting costs of digital circuitry.
- More efficient integration of voice, video, text, and image.

- Reduction of noise and other signal impairment because of use of repeaters.
 - Capacity of channels is utilized best with digital techniques.
 - Better encryption and hence better security than in analog transmission.

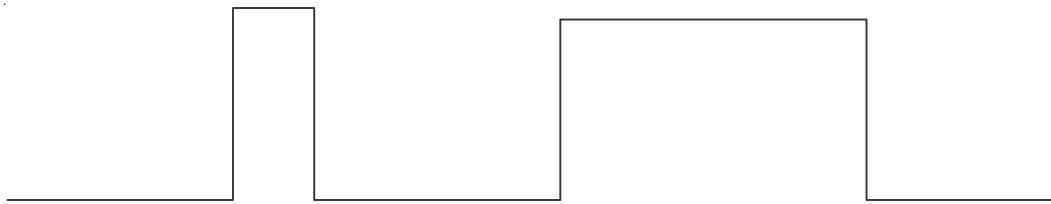


Fig. 2.4. NRZI Nonreturn to zero Invert on ones representation code.

2.3.3 Multiplexing of Transmission Signals

Quite often during the transmission of data over a network medium, the volume of transmitted data may far exceed the capacity of the medium. Whenever this happens, it may be possible to make multiple signal carriers share a transmission medium. This is referred to as *multiplexing*. There are two ways in which multiplexing can be achieved: time-division multiplexing (TMD) and frequency-division multiplexing (FDM).

In FDM, all data channels are first converted to analog form. Since a number of signals can be carried on a carrier, each analog signal is then modulated by a separate and different carrier frequency, and this makes it possible to recover during the demultiplexing process. The frequencies are then bundled on the carrier. At the receiving end, the demultiplexer can select the desired carrier signal and use it to extract the data signal for that channel in such a way that the bandwidths do not overlap. FDM has an advantage of supporting full-duplex communication.

TDM, on the other hand, works by dividing the channel into time slots that are allocated to the data streams before they are transmitted. At both ends of the transmission, if the sender and receiver agree on the time-slot assignments, then the receiver can easily recover and reconstruct the original data streams. So multiple digital signals can be carried on one carrier by interleaving portions of each signal in time.

2.4 DATA COMMUNICATION MEDIA TECHNOLOGY

2.4.1 Twisted-pair Cable

The common twisted-pair cable, which many readers are familiar with from wiring a telephone extension in the home, forms the basis for many measurements used in the communications field. Although we may have only one telephone in our home, the twisted-pair cable also forms the basis for large diameter cables that connect up to 25 instruments to switchboards and automatic switching equipment in an office environment.

The twisted-pair cable consists of two insulated conductors that are twisted together to form a transmission line. The pairs are combined in a cable that typically contains 2, 4 or 25-pairs, with the cable wrapped with a protective jacket of plastic or similar material. Two- and four-pair cables are used in both the home and office. From the home, the cable pair will be routed to an end office, whereas in many offices the cable pair will be routed to a switchboard or to a wiring closet.

In an office environment the wiring closet functions as a wire distribution center, permitting telephones on a floor or within a particular geographic area to be cabled to a common point. From the wiring closet a 25-pair cable is typically used to connect up to 25 telephones to a switchboard or electronic switching device known as a private branch exchange (PBX). From the PBX, calls to other telephone numbers connected to the switch are routed through the switch and never leave the customers' premises. Calls destined to telephone numbers not serviced by the PBX are first routed via trunks that connect the organization's PBX to a telephone company office. From there, the call is routed over the public switched telephone network to its destination.

Signal degradation

Regardless of the method of call routing, there are numerous electrical properties associated with the twisted-pair cable that will affect the quality of an electrical signal moving through the cable. Four of the more prominent electrical properties that affect an electrical signal transmitted over a twisted pair are attenuation, capacitance, crosstalk and delay distortion.

➤ Attenuation

As an electrical signal travels through a cable, it becomes weaker due to the resistance offered by the cable to flow. This weakness is called attenuation and it refers to the reduction in the amplitude or height of a transmitted signal. In voice communications, attenuation reduces the loudness of a conversation. To compensate for the effect of attenuation, telephone companies install amplifiers in their facilities at selective locations to boost signal levels. Figure 2.5 illustrates the effect of attenuation on an analog signal and the use of an amplifier to rebuild the signal level. When we speak of an analog signal, we are referring to a continuous signal as opposed to a digital signal which is discrete.

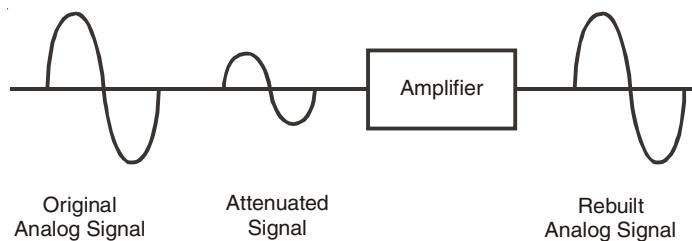


Fig. 2.5. Attenuation of an analog signal. Amplifiers are used by the telephone company to boost the signal strength of an analog signal.

A second cause of attenuation is the result of the creation of the telephone channel passband by the use of electrical filters. Telephone companies use low and high pass filters to pass only a small subset of the 20000 Hz bandwidth audible to the human ear. As a

result of the use of electrical filters at approximately 300 and 3300 Hz and the fact that high frequencies attenuate more rapidly than low frequencies, the ideal passband of a telephone channel becomes skewed. The result is an increase in attenuation as frequencies increase as well as at both filter cut-off frequencies. Figure 2.6 illustrates the typical amplitude-frequency response across a voice channel.

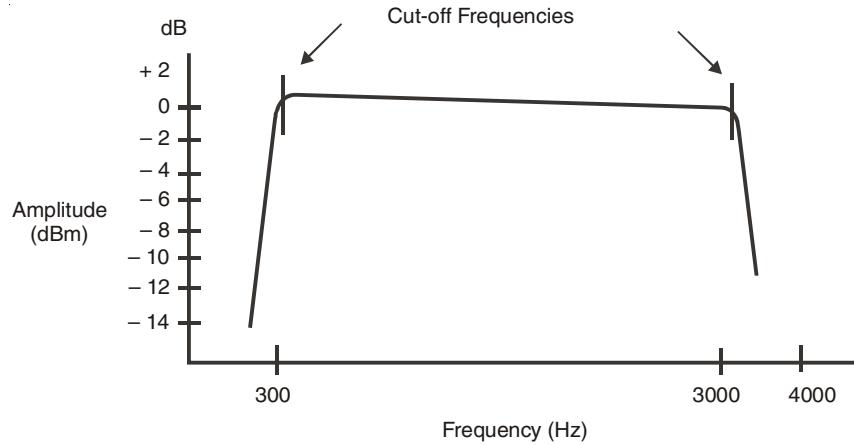


Fig. 2.6. Typical amplitude vs frequency response across a voice channel.

Due to the use of low pass and high pass filters, a large degree of attenuation occurs both below 300 and above 3300 Hz. In addition, because high frequencies attenuate more rapidly than low frequencies, the amplitude frequency response is non-linear from 300 to 3300 Hz, with the amount of signal attenuation increasing as the frequency increases.

The effect of attenuation upon a digital signal is similar to its effect upon an analog signal. That is, attenuation of a digital signal reduces the height of the square waves that form the signal.

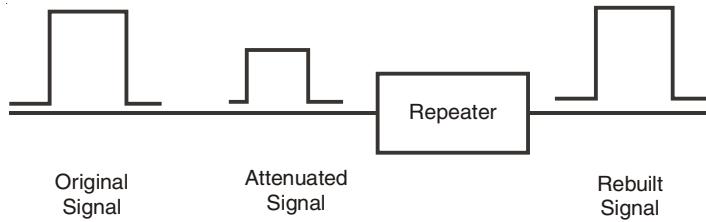


Fig. 2.7. Effect of attenuation on digital signals. A digital repeater samples the line for the occurrence of a pulse and regenerates the pulse at its original height and width.

In place of amplifiers used on analog circuits, digital repeaters, also called data regenerators, are used on a digital circuit to compensate for the loss in signal strength. The digital repeater as its name implies accepts a pulse and regenerates it at its original height and width, eliminating any previous distortion to the pulse. Figure 2.7 illustrates the attenuation of a digital signal and the use of a digital repeater to 'rebuild' the signals on a digital transmission medium.

➤ **Crosstalk**

The presence on a cable of a signal which originated on a different cable is called crosstalk. Although telephone cables always exhibit a degree of crosstalk, since a signal on an 'excited' pair always induces a signal on a 'quiet' pair, in most instances the effect of crosstalk is negligible. When the effect of crosstalk becomes relatively large, its effect upon both voice and data can become considerable.

In general, crosstalk is proportional to the dielectric constant of a cable. A cable with a lower dielectric constant will have less capacitance than a cable with a higher dielectric constant. Since the amount of capacitance on a cable is proportional to its level of crosstalk, the level of crosstalk is proportional to the dielectric constant of the cable.

2.4.2 Coaxial Cable

A coaxial cable consists of an inner conductor and an outer conductor that are insulated from one another. The insulation is called the dielectric. Figure 2.8 illustrates the composition of a coaxial cable.

Within the protective jacket the cable contains two conductors in concentric circles to one another. A solid wire forms the inner conductor, while the outer conductor functions as a shield and is usually grounded.

Coaxial cables can transmit signals ranging in frequency from 1 kHz to 1 GHz per second with little loss, distortion or interference to or from outside signals.

There are many types of coaxial cable that include cable with multiple conductors. However, coax, a term used to denote this type of cable, can normally be classified into those used for baseband and those used for broadband transmission. Baseband refers to the transmission of one signal at a time, and baseband coax normally has a characteristic impedance of 50Ω . By comparison, broadband refers to the ability to simultaneously transmit two or more signals on a cable by transmitting each signal at a different frequency. Broadband coax is the 75Ω cable used with CATV systems.

Due to the larger bandwidth of coaxial cable than of the twisted pair, they have a larger data transmission capacity. Typical usage of coaxial cable includes Community Antenna Television (CATV) and Local Area Networks (LANs).

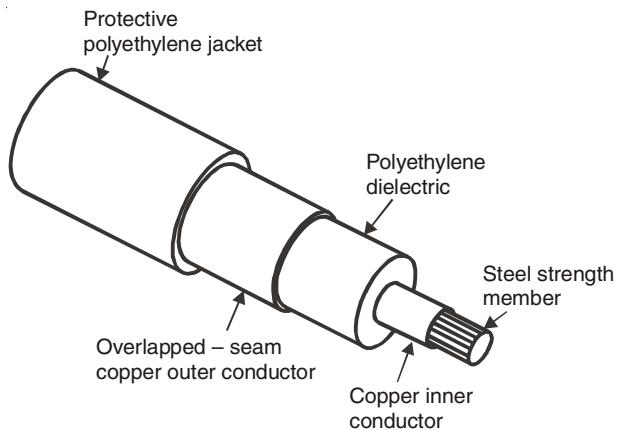


Fig. 2.8. Composition of coaxial cable.

2.4.3 Microwave

Microwave is a line-of-sight transmission medium that provides the bandwidth and capacity of coaxial cable without requiring the laying of a physical cable. Since microwave is a line-of-sight transmission, its use requires the construction of towers which, due to the curvature of the Earth, are limited to distances of 30 miles or less from one another. Figure 2.9 illustrates the use of microwave towers.

The large bandwidth of microwave communications permits communications carriers to simultaneously transmit thousands of calls between offices connected by microwave towers. As you travel around the globe the microwave tower is a common sight, especially in small towns and cities where its height makes it stand out as the communications link of the town or city to the rest of the world.

The transmission of a microwave signal to a point above the Earth forms the basis for satellite communications. A communications satellite contains a number of transponders (transmitters–receivers) which function as a microwave relay in the sky. The satellite receives the microwave signal transmitted from an Earth station and rebroadcasts it back to Earth.

Unlike transmission via a cable, microwave transmission can be affected by sunspots and weather: heavy rain, thunderstorms and dust storms all have an adverse effect upon microwave transmission.

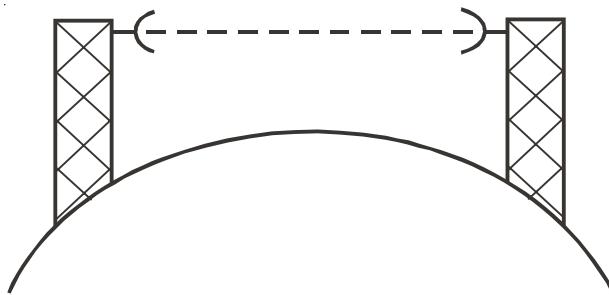


Fig. 2.9. Microwave Transmission.

During the 1950s and 1960s microwave towers literally dotted the landscape of many countries, providing the most common method for supporting inter-city communications. During the 1970s communications carriers turned to the use of fiber optic cable as a replacement for microwave. Because light transmitted over fiber is immune to electrical disturbances, by the mid—to late 1980s the use of fiber optics resulted in a substantial reduction in the use of microwave towers.

2.4.4 Fiber-optic Transmission

Once a laboratory and consumer product curiosity, optical transmission via fiber is now used for low cost, high data rate transmission. The major components of an optical system are similar to a conventional transmission system, requiring a transmitter, transmission medium and receiver. The transmitter, an electrical-to-optical (E/O) converter, receives electronic signals and converts them to a series of light pulses. The transmission medium is an optical fiber cable of plastic or glass. The receiver, an optical-to-electric (O/E) converter, changes the received light signals into corresponding electrical signals.

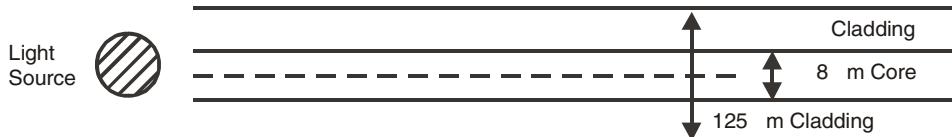
The bandwidth of a typical optical fiber can range up to 10 GHz. Normally a series of optical fibers, each shielded and bundled together, are routed between locations. To provide an indication of the transmission capability of an optical fiber, consider the potential use of two strands of fiber in a bundle. Each pair of strands can carry up to 24000 telephone calls, with calls routed in different directions on each fiber strand in a pair as until recently they were used for unidirectional transmission. An equivalent capacity established through the use of metallic twisted pair would require 48000 copper wires to establish the same number of telephone calls.

Although optical fiber has a relatively high bandwidth in comparison to copper, the growth in the use of the Internet and other applications began to tax that bandwidth. Beginning in the late 1990s a technique similar to FDM was used on many optical circuits. Based upon the subdivision of an optical fiber by wavelength, a technique referred to as wavelength division multiplexing (WDM) enables one optical fiber to transport up to 100 or possibly more simultaneous light sources, significantly increasing the transmission capacity of each optical fiber.

There are two general types of optical fiber cable used to transmit information: single-mode and multimode. A single-mode fiber has a core diameter of approximately 8 microns which is approximately the wavelength of light. This results in light flowing on one route through the fiber. In comparison, multimode fiber has a core diameter of approximately 62.5 microns, although 50, 85 and 100 micron cables are available. This diameter is large enough to permit light to travel on different paths, with each path considered a mode of propagation, hence the term multimode. Single-mode fiber is designed to transmit laser-generated light signals at distances up to 20 to 30 miles. This type of optical fiber is primarily used by communications carriers. In comparison, multimode was designed to carry relatively weak light emitting diode (LED) generated signals over relatively short distances, typically under a mile. Multimode fiber is primarily used within buildings and forms the infrastructure for optical LANs.

Optical fiber cable is typically specified using two numeric identifiers of the form x/y . Here x represents the diameter of the core, and y represents the cable diameter. Figure 2.10 compares the transmission of light in single and multimode fiber.

Typical 8/125 Single-mode Fiber



Typical 62.5/125 Multimode Fiber

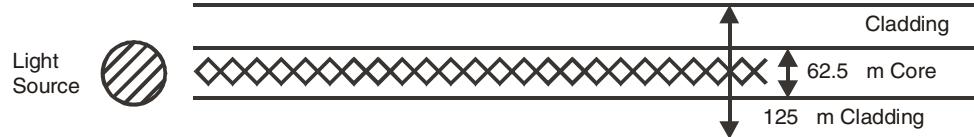


Fig. 2.10. Comparison of Single-mode and multimode fibers operation.

2.5 MOBILE AND CELLULAR NETWORKS AND COMMUNICATION

Cellular networks operate by dividing the service coverage area into zones or cells, each of which has its own set of resources or channels, which can be accessed by users of the network. Usually cellular coverage is represented by a hexagonal cell structure to demonstrate the concept, however, in practice the shape of cells is determined by the local topography. Sophisticated planning tools are used extensively by terrestrial cellular operators to assist with the planning of their cellular networks.

The shape and boundary of a cell is determined by its base station (BS), which provides the radio coverage. A BS communicates with mobile users through signalling and traffic channels (TCH). Signals transmitted in the direction from the BS to the mobile are termed the forward link or downlink, and conversely, the reverse link or uplink is in the direction of mobile to BS. Signalling channels are used to perform administrative and management functions such as setting up a call, while TCHs are used to convey the information content of a call. The allocation of channels to a cell is therefore divided between the TCHs, which form the majority, and signalling channels. These are allocated for both forward and reverse directions.

In order to increase the capacity of a network, there are three possibilities, either:

1. a greater number of channels are made available;
2. more spectrally efficient modulation and multiple access techniques are employed; or
3. the same channels are reused, separated by a distance which would not cause an unacceptable level of co-channel interference.

Cellular networks, which are limited in terms of available bandwidth, operate using the principle of frequency reuse. This implies that the same pool of frequencies is reused in cells that are sufficiently separated so as not to cause harmful co-channel interference. For a hexagonal cell structure, it is possible to cluster cells so that no two adjacent cells are using the same frequency. This is only achievable for certain cell-cluster sizes.

A seven-cell frequency reuse pattern is shown in Fig. 2.11. The total bandwidth available to the network is divided between cells in a cluster, which can then be used to determine the number of calls that can be supported in each cell. By reducing the number of cells per cluster, the system capacity can be increased, since more channels can be available per cell. However, a reduction in the cluster size will also result in a reduction in the frequency reuse distance, hence the system may become more prone to co-channel interference.

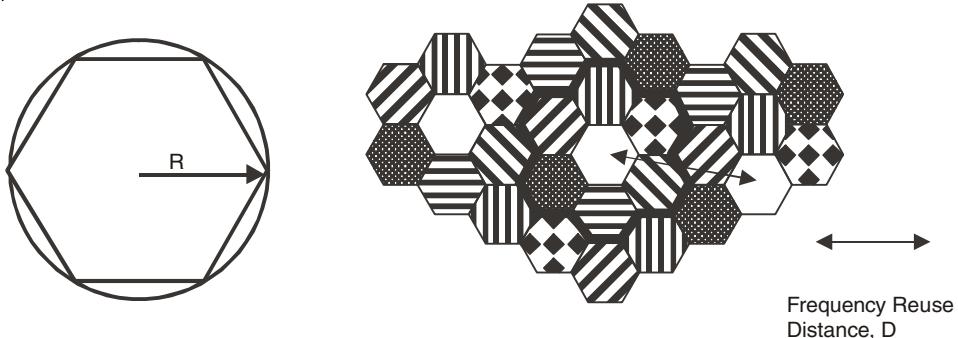


Fig. 2.11. A seven cell frequency reuse pattern.

How the mobile user gains access to the available channels within a cell is governed by the multiple access technique used by the network. Analogue cellular networks employ frequency division multiple access (FDMA), whereas digital networks employ either time division multiple access (TDMA) or code division multiple access (CDMA). For FDMA, a seven cell reuse pattern is generally employed, whereas for CDMA a single-cell frequency reuse pattern is achievable.

In a terrestrial mobile environment, reception cannot rely on line-of-sight communications and is largely dependent upon the reception of signal reflections from the surrounding environment. The resultant scattering and multipath components arrive at the receiver with random phase. The propagation channel can be characterised by a combination of a slow-fading, long-term component and a fast-fading, short-term component. As a consequence of the local terrain, the change in a mobile's position relative to that of a transmitting BS will result in periodic nulls in the received signal strength. This is due to the fact that the vector summation of the multipath and scattering components at the receiver results in a signal envelope of the form of a standing wave pattern, which has signal nulls at half-wave intervals. For a signal transmitting at 900 MHz, which is typical for cellular applications, a half-wavelength distance corresponds to approximately 17 cm. This phenomenon is known as slow-fading and is characterised by a log-normal probability density function.

As the mobile's velocity, n , increases, the variation in the received signal envelope becomes much more pronounced and the effect of the Döppler shift on the received multipath signal components also has an influence on the received signal.

This phenomenon is termed fast-fading and is characterised by a Rayleigh probability density function. Such variations in received signal strength can be as much as 30 dB below or 10 dB above the root mean square signal level, although such extremes occur infrequently.

In rural areas, where the density of users is relatively low, large cells of about 25 km radius can be employed to provide service coverage. This was indeed the scenario when mobile communications were first introduced into service. In order to sustain the mobile to BS link over such a distance requires the use of a vehicular-type mobile terminal, where available transmit power is not so constrained in comparison with hand-held devices. With an increase in user-density, the cell size needs to reduce in order to enable a greater

frequency reuse and hence to increase the capacity of the network. Urban cells are typically of 1 km radius. This reduction in cell size will also correspond to a reduction in BS and mobile terminal transmit power requirements. This is particularly important in the latter case, since it paves the way for the introduction of hand-held terminals.

When a mobile moves from one cell to another during the course of an on-going call, a handover (also termed handoff) of the call between BSs must be performed in order to ensure that the call continues without interruption. Otherwise the call will be dropped and the mobile user would need to re-initiate the call set-up sequence. Handover between BSs involves monitoring of the signal strength between the mobile to BS link. Once the signal strength reduces below a given threshold, the network initiates a procedure to reserve a channel through another BS, which can provide a channel of sufficient signal strength (Figure 2.12).

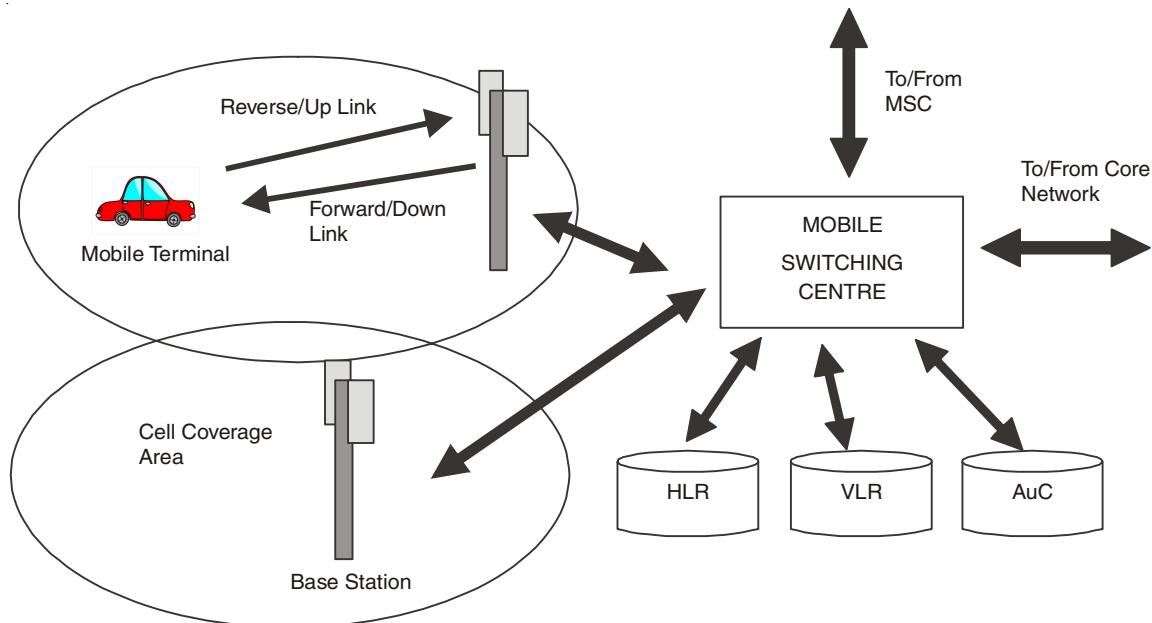


Fig. 2.12. Basic Cellular Network Architecture.

A number of BSs are clustered together via a fixed-network connection to a mobile switching centre (MSC), which provides the switching functionality between BSs during handover and can also provide connection to the fixed or core network (CN) to allow the routing of calls. The clustering of BSs around a MSC is used to define a Location Area, which can be used to determine the latest known location of a mobile user. This is achieved by associating Home and Visitor Location Areas to a mobile. Each mobile is registered with a single home location register (HLR) upon joining the network. Once a mobile roams outside of its Home Location Area into a new designated Location Area, it temporarily registers with the network as a visitor, where its details are stored in a visitor location register (VLR) associated with the MSC. Each MSC in the network has an associated VLR and HLR. The mobile's location is relayed back to its HLR, a database containing various

information on the mobile terminal, some of which is then forwarded to the VLR. The network also comprises of other databases that can be used to verify that the mobile has access to the network, such as the Authentication Centre (AuC), for example. These procedures are described later in the chapter for the GSM system.

2.5.1 First Generation Analog Cellular Systems

In the future mobile information society, where mobile-multimedia delivery will be the major technological driving force, analogue cellular technology has little, if any significance. Indeed, in many countries across Europe, mobile operators are now switching off their analogue services in favour of digital technology. However, analogue technologies still play an important role in many countries around the world, by being able to provide established and reliable mobile voice telephony at a competitive price. This section considers three of the major analogue systems that can still be found with significant customer databases throughout the world.

➤ Nordic Mobile Telephone (NMT) System

On 1 October, 1981, the Nordic NMT450 became the first European cellular mobile communication system to be introduced into service. This system was initially developed to provide mobile communication facilities to the rural and less-populated regions of the Scandinavian countries Denmark, Norway, Finland and Sweden. NMT450 was essentially developed for in-car and portable telephones. By adopting common standards and operating frequencies, roaming between Scandinavian countries was possible. Importantly, the introduction of this new technology provided network operators and suppliers with an early market lead, one that has been sustained right up to the present day.

As is synonymous of 1G systems, NMT450 is an analogue system. It operates in the 450 MHz band, specifically 453–457.5 MHz (mobile to BS) and 463–467.5 MHz (BS to mobile). FDMA/FM is employed as the multiple access scheme/modulation method for audio signals, with a maximum frequency deviation of +/- 5 kHz. Frequency shift keying (FSK) is used to modulate control signals with a frequency deviation of +/- 3.5 kHz. NMT450 operates using a channel spacing of 25 kHz, enabling the support of 180 channels. Since its introduction, the NMT450 system has continued to evolve with the development of the NMT450*i* (where *i* stands for improvement) and NMT900 systems.

NMT900 was introduced into service in 1986, around about the same time as other Western/European countries were starting to introduce their own city based mobile cellular-based solutions. NMT900 is designed for city use, catering for hand-held and portable terminals. It operates in the 900 MHz band with the ability to accommodate higher data rates and more channels.

The NMT system continues to hold a significant market share throughout the world and, significantly, the system continues to evolve, through a series of planned upgrades. In Europe, the NMT family has a particularly large market share in Eastern European countries, where mobile telephony is only now starting to become prevalent.

Since 1981, Nordic countries have continued to lead the way with now over 60% of the population in Finland and Norway having a mobile phone. The Scandinavian-based companies Nokia and Ericsson are world leaders in mobile phone technology and both are driving the phone's evolution forward.

➤ Advanced Mobile Phone Service (AMPS)

Bell Labs in the US developed the AMPS communications system in the late 1970s [BEL-79]. The AMPS system was introduced into commercial service in 1983 by AT&T with a 3-month trial in Chicago. The system operates in the US in the 800 MHz band, specifically 824–849 MHz (mobile to BS) and 869–894 MHz (BS to mobile). These bands offer 832 channels, which are divided equally between two operators in each geographical area. Of these 832 channels, 42 channels carry only system information. The AMPS system provides a channel spacing of 30 kHz using FM modulation with a 12 kHz peak frequency deviation for voice signals.

Signalling between mobile and BS is at 10 kbit/s employing Manchester coding. The signals are modulated using FSK, with a frequency deviation of ± 8 kHz. The AMPS system specifies six one-way logical channels for transmission of user and signalling information. The Reverse TCH and Forward TCH are dedicated to the transmission of user data on a one-to-one basis. Signalling information is carried to the BS on the channels reverse control channel (RECC) and reverse voice channel (RVC); and to the mobile using the channels forward control channel (FOCC) and forward voice channel (FVC).

The forward and reverse control channels are used exclusively for network control information and can be referred to as Common Control Channels. To safeguard control channels from the effect of the mobile channel, information is protected using concatenated pairs of block codes. To further protect information, an inner code employs multiple repetition of each BCH (Bose–Chadhuri–Hocquenghem) code word at least five times, and 11 times for the FVC.

In order to identify the BS assigned to a call, AMPS employs a supervisory audio tone (SAT), which can be one of three frequencies (5970, 6000 and 6030 Hz). At call set-up, a mobile terminal is informed of the SAT at the BS to which it communicates. During a call, the mobile terminal continuously monitors the SAT injected by the BS. The BS also monitors the same SAT injected by the mobile terminal. Should the received SAT be incorrect at either the mobile terminal or the BS, the signal is muted, since this would imply reception of a source of interference.

Like NMT450, the AMPS standard has continued to evolve and remains one of the most widely used systems in the world. Although market penetration did not reach Europe, at least in its unmodified form, it remains a dominant standard in the Americas and Asia.

➤ Total Access Communications System (TACS)

By the mid-1980s, most of Western Europe had mobile cellular capability, although each country tended to adopt its own system. For example, the C-NETZ system was introduced in Germany and Austria, and RADIOCOM 2000 and NMT-F, the French version of NMT900 could be found in France. This variety of technology made it impossible for international commuters to use their phones on international networks, since every national operator had its own standard. In the UK, Racal Vodafone and Cellnet, competing operators providing technically compatible systems, introduced the TACS into service in January 1985. TACS was based on the American AMPS standard with modifications to the operating frequencies and channel spacing. TACS offers a capacity of 600 channels in the bands 890–905 MHz (mobile to BS) and 935–950 MHz (BS to mobile), the available bandwidth

being divided equally between the two operators. Twenty-one of these channels are dedicated for control channels per operator. The system was developed with the aim of serving highly populated urban areas as well as rural areas. This necessitated the use of a small cell size in urban areas of 1 km. In TACS, the cell size ranges from 1 to 10 km. TACS provides a channel spacing of 25 kHz using FM modulation with a 9.5 kHz peak deviation for voice signals. In highly densely populated regions, the number of available channels is increased to up to 640 (320 channels per operator) by extending the available spectrum to below the conference of European Posts and Telegraphs (CEPT) cellular band. This is known as extended TACS (ETACS). Here, the operating frequency bands are 917–933 MHz in the mobile to BS direction and 872–888 MHz in the BS to mobile.

Fifteen years after TACS was first introduced into the UK, the combined Vodafone and Cellnet customer base amounted to just under half a million subscribers out of a total of 31 million. The future of analogue technology in developed markets is clearly limited, particularly with the re-farming of the spectrum for the 3G services. Nevertheless, analogue systems such as TACS have been responsible for developing the mobile culture and in this respect, their contribution to the evolution of the mobile society remains significant. Within Europe, TACS networks can also be found in Austria, Azerbaijan, Ireland, Italy, Malta and Spain. A variant of TACS, known as J-TACS, operates in Japan.

2.5.2 Second-Generation (2G) Systems

➤ Global System for Mobile Communications (GSM)

Following a proposal by Nordic Telecom and Netherlands PTT, the Group Special Mobile (GSM) study group was formed in 1982 by the CEPT. The aim of this study group was to define a pan-European public land mobile system.

In 1987, 13 operators and administrators signed the GSM memorandum of understanding (MoU) agreement and the original French name was changed to the more descriptive Global System for Mobile communications (GSM), although the acronym remained the same. By 1999, 296 operators and administrators from 110 countries had signed the GSM MoU. Significantly, in 1987, following the evaluation of several candidate technologies through laboratory and field trial experiments, agreement was reached on the use of a regular pulse excitation-linear predictive coder (RPE-LPC) for speech coding and TDMA was selected as the multiple access method.

In 1989 responsibility for the GSM specification was transferred to the ETSI and a year later Phase 1 GSM specifications were published. Commercial GSM services began in Europe two years later in mid-1991. In addition to voice services, the SMS was created as part of the GSM Phase 1 standard. This provides the facility to send and receive text messages from mobile phones. Messages can be up to 160 characters in length and can be used to alert the user of an incoming e-mail message, for example. It is a store-and-forward service, with all messages passing through an SMS centre. The SMS has proved to be hugely popular in Europe, with the transmission of in excess of 1 billion messages per month as of April 1999. In 1997, Phase 2 specifications came on-line, allowing the transmission of fax and data services.

At the end of 1998, ETSI completed its standardisation of GSM Phase 21 services high speed circuit switched data (HSCSD) and general packet radio service (GPRS). These two new services are aimed very much at exploiting the potential markets in the mobile data sector, recognising the influence of the Internet on mobile technologies. Responsibility for the maintenance and future development of the GSM standards is now under the control of the 3G partnership project (3GPP).

A simplified form of the GSM network architecture is shown in Fig. 2.13.

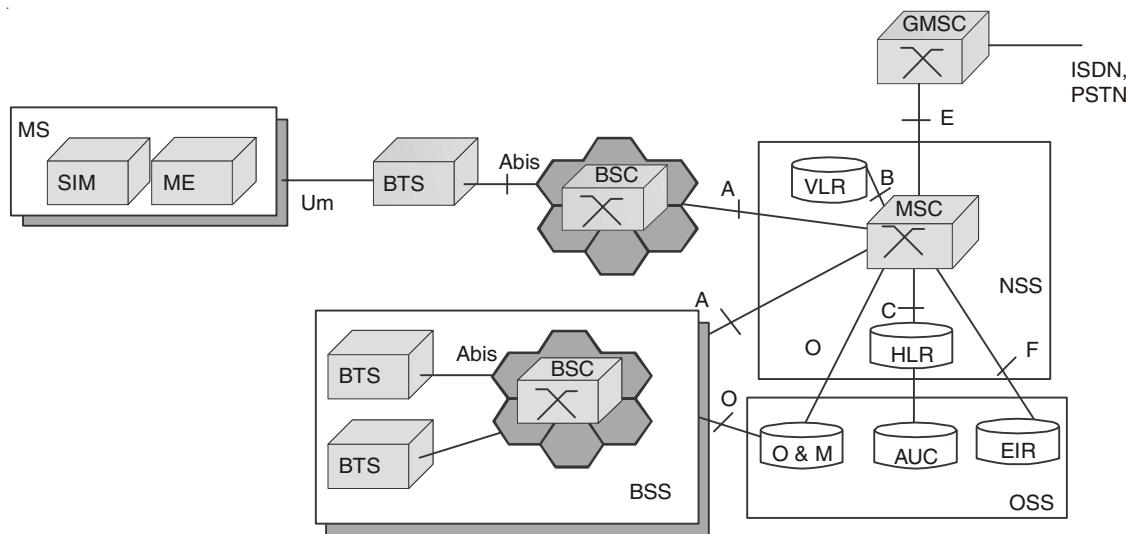


Fig. 2.13. A simplified GSM Architecture.

Mobile Station (MS) A subscriber uses an MS to access services provided by the network. The MS consists of two entities, the mobile equipment (ME) and subscriber identity module (SIM). The ME performs the functions required to support the radio channel between the MS and a BTS. These functions include modulation, coding, and so on. It also provides the application interface of the MS to enable the user to access services. A SIM card provides the ability to personalise a mobile phone. This is a smart card that needs to be inserted into the mobile phone before it can become operational. The SIM card contains the user's international mobile subscriber identity (IMSI), as well as other user specific data including an authentication key. Similarly, a terminal is identified by its international ME identity (IMEI). The IMSI and IMEI provide the capability for personal and terminal mobility, respectively. The radio interface between the MT and the BTS is termed the Um-interface and is one of the two mandatory interfaces in the GSM network.

Base Station System (BSS) The BTS forms part of the base station system (BSS), along with the base station controller (BSC). The BTS provides the radio coverage per cell, while the BSC performs the necessary control functions, which include channel allocation and local switching to achieve handover when a mobile moves from one BTS to another under the control of the same BSC. A BTS is connected to a BSC via an Abis-interface.

Network Management and Switching Subsystem (NMSS) The NMSS provides the connection between the mobile user and other users. Central to the operation of the NMSS is the MSC. A BSS is connected to an MSC via an A-interface, the other mandatory GSM interface. The coverage area of an MSC is determined by the cellular coverage provided by the BTSs that are connected to it. The functions of an MSC include the routing of calls to the appropriate BSS, performing handover between BSSs and inter-working with other fixed networks. A special type of MSC is the gateway MSC (GMSC), which provides connection to fixed telephone networks and vice versa. A GMSC is connected to an MSC via an E-interface. Central to the operation of the MSC are the two databases: HLR, connected via the C-interface; and VLR, connected via the B-interface. The HLR database contains management information for an MS. Each MS has an associated HLR. The information contained within an HLR includes the present location of an MS, and its IMSI, which is used by the authentication centre (AuC) to authorise a subscriber's access to the network.

Each MSC has an associated VLR. Whenever an MS is switched on in a new location area or roams into a new location area covered by an MSC, it must register with its VLR. At this stage, the visiting network assigns a MS roaming number (MSRN) and a temporary mobile subscriber identity (TMSI) to the MS. The location of the MS, usually in terms of the VLR's signalling address, is then conveyed to the HLR. The VLR, by using subscriber information provided by the HLR, can perform the necessary routing, verification and authentication procedures for an MS that would normally be performed by the HLR. An MSC also provides connection to the SMS centre (SMSC), which is responsible for storing and forwarding messages.

Operation Subsystem (OSS): The OSS provides the functions for the operation and management of the network. The Network Operation and Maintenance Centre performs all the necessary functionalities necessary to monitor and manage the network. It is connected to all of the major network elements (BTS, MSC, HLR, VLR) via an O-interface using an X.25 connection. The equipment interface register (EIR) is used by the network to identify any equipment that may be using the network illegally. The MSC-EIR connection is specified by the F-interface. The AuC also forms part of the OSS.

➤ Digital Cellular System 1800 (DCS1800)

The first evolution of GSM came about with the introduction of DCS1800, which is aimed primarily at the mass-market pedestrian user located in urban, densely populated regions. DCS1800 was introduced under the personal communications network (PCN) concept, also known as personal communication services (PCS) in the US. In 1989, the UK Government's Department of Trade and Industry outlined its intention to issue licenses for personal communication networks in the 1700–2300 MHz band. It was recognised that the new service, which would be aimed primarily at the pedestrian user, would be an adaptation of the GSM standard. Subsequently, ETSI produced the Phase 1 DCS1800 specification in January 1991, which detailed the generic differences between DCS1800 and GSM. This was followed by a Phase 2 specification detailing a common framework for PCN and GSM. DCS1800 operates using largely the same specification as GSM, making use of the same

network architecture but, as its name implies, it operates in the 1800 MHz band. Here, parallels can be drawn with the evolution of the NMT450 to the NMT900 system from earlier discussions.

The bands that have been allocated for DCS1800 operation are 1710–1785 MHz for the mobile to BS link and 1805–1880 MHz for the BS to mobile link. Taking into account the 200-kHz guard-band, 374 carriers can be supported. Apart from the difference in the operating frequency, the only other major difference is in the transmit power specification of the mobile station. Two power classes were defined at 250 mW and 1 W peak power, respectively. As DCS1800 is intended primarily for urban environments, the cells are much smaller than those of GSM; hence the transmit power requirements are reduced.

The UK was the first country to introduce PCN into operation, through two network operators: Mercury's One2One, which was introduced into service in September 1993; and Hutchinson's Orange, which was introduced into service in April 1994. Dual-mode 900/1800 MHz terminals are now available on the market, as are triple-mode 900/1800/1900 MHz terminals, allowing roaming into North America (in the US, the 1900 MHz band is used for PCS). Towards the end of 1999, Orange and One2One accounted for a third of the UK digital cellular market at just under 5 million subscribers, with virtually an equal market share between the two of them.

2.5.3 Third Generation Cellular Systems

ITU has called the future cellular networks 3G networks or IMT-2000; the previous term was Future Public Land Mobile Telephone System (FPLMTS). The performance for IMT-2000 air interference can be summarized as:

- Wideband CDMA systems
- Spectrum bandwidth 5 MHz
- Full coverage and mobility for a data rate of 144 kbps to 384 kbps
- Limited coverage and mobility or no mobility for 2 Mbps
- High spectrum efficiency compared to 2G system
- High flexibility to introduce new and multimedia services

The 3G features:

- Provision of multirate services
- Packet data
- A user-dedicated pilot for a coherent uplink
- An additional DL pilot channel for beam forming
- Intercarrier handover
- Fast power control
- Multiuser detection

IMT-2000 has published a minimum performance requirement of a 3G wireless system, which is for both circuit-switched (CS) and packet-switched (PS) data:

- Data rate of 144 kbps in the vehicular environment.
- Data rate of 384 kbps in the pedestrian environment.
- Data rate of 2 Mbps in the fixed indoor and pico cell environment.

The 3G systems concentrating on the three ITU-adopted systems using CDMA technology are WCDMA-UTRA (Europe), WCDMA-ARIB (Japan), and cdma2000 (North America). 3G consists of the three systems shown in Fig. 2.14, also called three modes because in the future, the three systems can have the features of hooks and extension to make intersystem connections among them.

	FDD Direct Spread	FDD Multicarrier	TDD
Bandwidth	5 MHz	5 MHz/ 1.25 MHz	5 MHz/ 1.6 MHz
Chip Rate	3.84 Mcps	3.6864 Mcps/ 1.228 Mcps	3.84 Mcps/ 1.28 Mcps
Common Pilot	CDM	CDM	TDM
Dedicated Pilot	TDM	CDM	TDM
Synchronization	Asynchronous/ Synchronous	Synchronous as cdma2000	Synchronous
Core Network	GSM-MAP/ IP	ANSI-41/ IP	GSM MAP/ IP

Fig. 2.14. 3G Standard systems.

The wireless loop area network (WLAN)¹ can offer a data speed much higher than 3G due to the different technologies. WLAN (including Wi-Fi and WiMAX) were later called the B3G (Beyond 3G) system by the industry. The comparison of the data rate with different standards in WLAN is shown in Fig. 2.15.

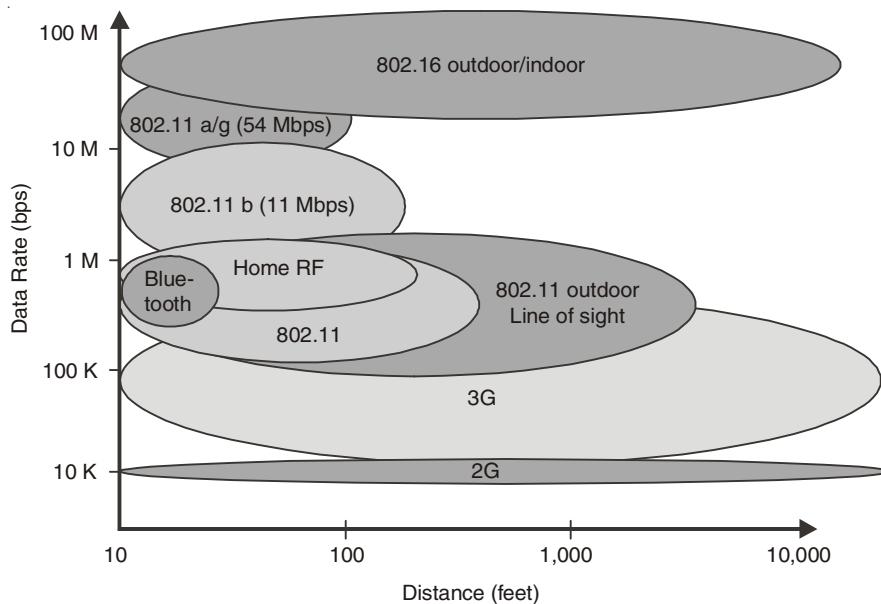


Fig. 2.15. Comparison of various Wireless systems.

2.6 OPTICAL SWITCHING NETWORKS

Optical fiber provides an unprecedented bandwidth potential that is far in excess of any other known transmission medium. A single strand of fiber offers a total bandwidth of 25000 GHz. To put this potential into perspective, it is worthwhile to note that the total bandwidth of radio on Earth is not more than 25 GHz. Apart from its enormous bandwidth, optical fiber provides additional advantages such as low attenuation loss. Optical networks aim at exploiting the unique properties of fiber in an efficient and cost-effective manner.

2.6.1 Optical Point-to-Point Links

The huge bandwidth potential of optical fiber has been long recognized. Optical fiber has been widely deployed to build high-speed optical networks using fiber links to interconnect geographically distributed network nodes. Optical networks have come a long way. In the early 1980s, optical fiber was primarily used to build and study point-to-point transmission systems. As shown in Fig. 2.16 (a), an optical point-to-point link provides an optical single-hop connection between two nodes without any (electrical) intermediate node in between. Optical point-to-point links may be viewed as the beginning of optical networks. Optical point-to-point links may be used to interconnect two different sites for data transmission and reception. At the transmitting side, the electrical data is converted into an optical signal (EO conversion) and subsequently sent on the optical fiber. At the receiving side, the arriving optical signal is converted back into the electrical domain (OE conversion) for electronic processing and storage. To interconnect more than two network nodes, multiple optical single-hop point-to-point links may be used to form various network topologies

(e.g., star and ring networks). Figure 2.16 (b) shows how multiple optical point-to-point links can be combined by means of a star coupler to build optical single-hop star networks. The star coupler is basically an optical device that combines all incoming optical signals and equally distributes them among all its output ports. In other words, the star coupler is an optical broadcast device where an optical signal arriving at any input port is forwarded to all output ports without undergoing any EO or OE conversion at the star coupler. Similar to optical point-to-point links, optical single-hop star networks make use of EO conversion at the transmitting side and OE conversion at the receiving side. Besides optical stars, optical ring networks can be realized by interconnecting each pair of adjacent ring nodes with a separate optical single-hop point-to-point fiber link, as depicted in Fig. 2.16 (c). In the resultant optical ring network, each node performs OE conversion for incoming signals and EO conversion for outgoing signals. The combined OE and EO conversion is usually referred to as OEO conversion. A good example of an optical ring network with OEO conversion at each node is the fiber distributed data interface (FDDI) standard, which can be found in today's existing optical network infrastructure.

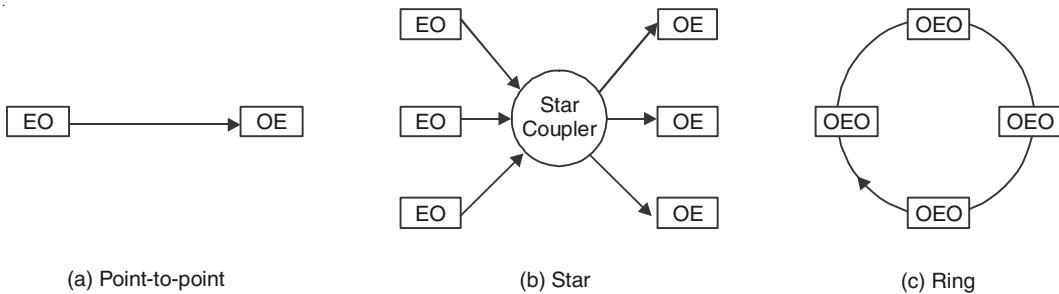


Fig. 2.16. Optical single-hop connections: (a) point-to-point, (b) star, and (c) ring configurations.

2.6.2 SONET/SDH

One of the most important standards for optical point-to-point links is the Synchronous Optical Network (SONET) standard and its closely related synchronous digital hierarchy (SDH) standard. The SONET standardization began during 1985 and the first standard was completed in June 1988. The goals of the SONET standard were to specify optical point-to-point transmission signal interfaces that allow interconnection of fiber optics transmission systems of different carriers and manufacturers, easy access to tributary signals, direct optical interfaces on terminals, and to provide new network features. SONET defines standard optical signals, a synchronous frame structure for time division multiplexed (TDM) digital traffic, and network operation procedures. SONET is based on a digital TDM signal hierarchy where a periodically recurring time frame of $125 \mu\text{s}$ can carry payload traffic of various rates. Besides payload traffic, the SONET frame structure contains several overhead bytes to perform a wide range of important network operations such as error monitoring, network maintenance, and channel provisioning.

SONET is now globally deployed by a large number of major network operators. Typically, SONET point-to-point links are used in ring configurations to form optical ring networks with OEO conversion at each node, similar to the one depicted in Fig. 2.16(c). In

SONET rings there are two main types of OEO nodes: the add-drop multiplexer (ADM) and the digital cross-connect system (DCS). The ADM usually connects to several SONET end devices and aggregates or splits SONET traffic at various speeds. The DCS is a SONET device that adds and drops individual SONET channels at any location. One major difference between an ADM and a DCS is that the DCS can be used to interconnect a larger number of links. The DCS is often used to interconnect SONET rings.

2.6.3 Multiplexing: TDM, SDM, and WDM

Given the huge bandwidth of optical fiber, it is unlikely that a single client or application will require the entire bandwidth. Instead, traffic of multiple different sources may share the fiber bandwidth by means of multiplexing. Multiplexing is a technique that allows multiple traffic sources to share a common transmission medium. In the context of optical networks, three main multiplexing approaches have been deployed to share the bandwidth of optical fiber: (1) Time Division Multiplexing (TDM), (2) Space Division Multiplexing (SDM), and (3) Wavelength Division Multiplexing (WDM).

- **Time Division Multiplexing:** We have already seen that SONET is an important example for optical networks that deploy TDM on the underlying point-to-point fiber links. Traditional TDM is a well-understood technique and has been used in many electronic network architectures throughout the more than 50-year history of digital communications. In the context of high-speed optical networks, however, TDM is under pressure from the so-called “electro-optical” bottleneck. This is due to the fact that the optical TDM signal carries the aggregate traffic of multiple different clients and each TDM network node must be able to operate at the aggregate line rate rather than the substrate that corresponds to the traffic originating from or destined for a given individual node. Clearly, the aggregate line rate cannot scale to arbitrarily high values but is limited by the fastest available electronic transmitting, receiving, and processing technology. As a result, TDM faces severe problems to fully exploit the enormous bandwidth of optical fiber.
- **Space Division Multiplexing:** One straightforward approach to avoid the electrooptical bottleneck is SDM, where multiple fibers are used in parallel instead of a single fiber. Each of these parallel fibers may operate at any arbitrary line rate (*e.g.*, electronic peak rate). SDM is well-suited for short-distance transmissions but becomes less practical and more costly for increasing distances due to the fact that multiple fibers need to be installed and operated.
- **Wavelength Division Multiplexing:** WDM appears to be the most promising approach to tap into the vast amount of fiber bandwidth while avoiding the aforementioned shortcomings of TDM and SDM. WDM can be thought of as optical frequency division multiplexing (FDM), where traffic from each client is sent on a different carrier frequency. In optical WDM networks the term wavelength is usually used instead of frequency, but the principle remains the same. As shown in Fig. 2.17, in optical WDM networks each transmitter i sends on a separate wavelength λ_i , where $1 \leq i \leq N$. At the transmitting side, a wavelength multiplexer collects all wavelengths and feeds them onto a common outgoing fiber. At the receiving side,

a wavelength demultiplexer separates the wavelengths and forwards each wavelength λ_i to a different receiver i . Unlike for TDM, each wavelength channel may operate at any arbitrary line rate well below the aggregate TDM line rate. By using multiple wavelengths the huge bandwidth potential of optical fiber can be exploited. As opposed to SDM, WDM takes full advantage of the bandwidth potential of a single fiber and does not require multiple fibers to be installed and operated in parallel, resulting in significant cost savings. Optical WDM networks have been attracting a great deal of attention by network operators, manufacturers, and research groups around the world.

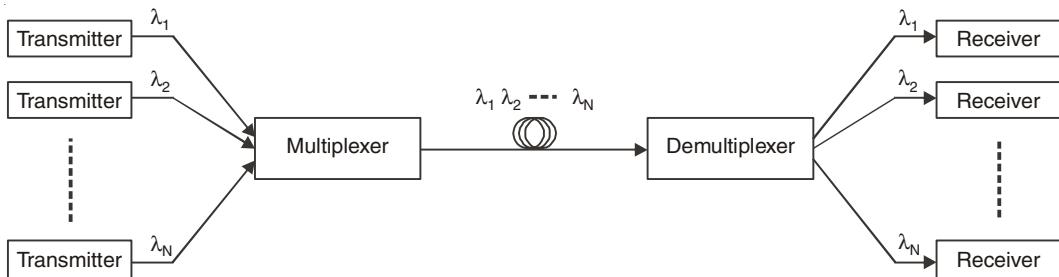


Fig. 2.17. Wavelength division multiplexing.

2.6.4 Applications of Optical Networks

Many of today's applications can be roughly broken down into the following two categories:

Latency-Critical Applications: This type of application comprises small-to medium size file transfers which require low latency. Examples of latency-critical applications are broadcast television, interactive video, video conferencing, security video monitoring, interactive games, telemedicine, and telecommuting.

Throughput-Critical Applications: This type of application includes large-size file transfers whose latency is not so important but which require much bandwidth. Examples of throughput-critical applications are video on demand (VoD), video and still-image email attachments, backup of files, program and file sharing, and file downloading (e.g., books).

Applications have a significant impact on the traffic load and throughput-delay performance requirements of (optical) networks. Depending on the application, traffic loads and throughput-delay performance requirements may change over time. For instance, web browsing has led to rather asymmetric network traffic loads with significantly less upstream traffic than downstream traffic. Web browsing is based on the client-server paradigm, where each client sends a short request message to a server for downloading data files. In future (optical) networks, the traffic load is expected to become less asymmetric due to the fact that so-called peer-to-peer (P2P) applications become increasingly popular. Napster and its successors are good examples of P2P applications, where each client also acts as a server from which other clients may download files (e.g., photos, videos, and programs). P2P applications traffic keeps growing and may already represent the major traffic load in existing access networks. The steadily growing P2P application traffic will eventually render the network traffic loads more symmetric.

The demand for more bandwidth is expected to keep increasing in order to support new emerging and future applications.

Review Questions

1. Compute the Fourier coefficients for the function $f(t) = t$ ($0 \leq t \leq 1$).
2. A noiseless 4-kHz channel is sampled every 1 msec. What is the maximum data rate?
3. Television channels are 6 MHz wide. How many bits/sec can be sent if four-level digital signals are used? Assume a noiseless channel.
4. If a binary signal is sent over a 3-kHz channel whose signal-to-noise ratio is 20 dB, what is the maximum achievable data rate?
5. What signal-to-noise ratio is needed to put a T1 carrier on a 50-kHz line?
6. What is the difference between a passive star and an active repeater in a fiber network?
7. How much bandwidth is there in 0.1 micron of spectrum at a wavelength of 1 micron?
8. It is desired to send a sequence of computer screen images over an optical fiber. The screen is 480×640 pixels, each pixel being 24 bits. There are 60 screen images per second. How much bandwidth is needed, and how many microns of wavelength are needed for this band at 1.30 microns?
9. Is the Nyquist theorem true for optical fiber or only for copper wire?
10. What is the essential difference between message switching and packet switching?
11. What is the available user bandwidth in an OC-12c connection?
12. Enlist applications of Optical switching Networks.



CHAPTER 3 *DATA LINK LAYER*

It is a mistake to look too far ahead. Only one link in the chain of destiny can be handled at a time.

—Winston Churchill

3.1 INTRODUCTION

The simplest network possible is one in which all the hosts are directly connected by some physical medium. This may be a wire or a fiber, and it may cover a small area (e.g., an office building) or a wide area (e.g., transcontinental). Connecting two or more nodes with a suitable medium is only the first step, however. There are some additional problems that must be addressed before the nodes can successfully exchange packets.

In this section we will study the design of layer 2, the data link layer (also known as the physical link control layer). The purpose of the data link layer is to transfer blocks of data without error between two adjacent devices. Adjacent devices are physically connected by a communication channel such as telephone lines, coaxial cables, optical fibres, or satellites. The implication of such a physical link is that the data bits are delivered in exactly the same order in which they are sent. The physical link has no inherent storage capacity, therefore the delay involved is the propagation delay over the link. Transmission of data over the link would be very simple indeed if no error ever occurred. Unfortunately, this is not so in a real physical link for a number of reasons:

- Natural phenomena such as noises and interference are introduced into the link causing errors in detecting the data.
- There is a propagation delay in the link.
- There is a finite data processing time required by the transmitting and receiving stations.

A data link protocol thus has to be designed to ensure an error-free transmission and also to achieve an efficiency of the data transfer as high as possible. To see the need for data link control, we list some of the requirements and objectives for effective data communication between two directly connected transmitting-receiving stations:

- **Frame synchronization.** Data are sent in blocks called frames. The beginning and end of each frame must be recognizable.
- **Flow control.** The sending station must not send frames at a rate faster than the receiving station can absorb them.
- **Error control.** Any bit errors introduced by the transmission system must be corrected.
- **Addressing.** On a multipoint line, such as a local area network (LAN), the identity of the two stations involved in a transmission must be specified.
- **Control and data on same link.** It is usually not desirable to have a physically separate communications path for control information. Accordingly, the receiver must be able to distinguish control information from the data being transmitted.
- **Link management.** The initiation, maintenance, and termination of a sustained data exchange requires a fair amount of coordination and cooperation among stations. Procedures for the management of this exchange are required.

To accomplish these goals, the data link layer takes the packets it gets from the network layer and encapsulates them into frames for transmission. Each frame contains a frame header, a payload field for holding the packet, and a frame trailer, as illustrated in Fig. 3.1. Frame management forms the heart of what the data link layer does. In the following sections we will examine all the above-mentioned issues in detail.

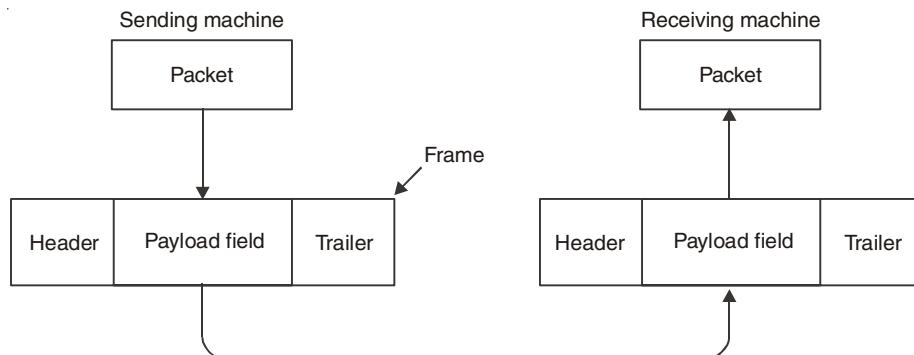


Fig. 3.1. Relationship between packets and frame.

Although this chapter is explicitly about the data link layer and the data link protocols, many of the principles we will study here, such as error control and flow control, are found in transport and other protocols as well. In fact, in many networks, these functions are found only in the upper layers and not in the data link layer. However, no matter where they are found, the principles are pretty much the same, so it does not really matter where we study them. In the data link layer they often show up in their simplest and purest forms, making this a good place to examine them in detail. It is important to mention here that data link layer provides services to network layer and in turn uses services from physical layer.

3.2 FRAMING

Data transmission in the physical layer means moving bits in the form of a signal from the source to the destination. The physical layer provides bit synchronization to ensure that the sender and receiver use the same bit durations and timing.

The data link layer, on the other hand, needs to pack bits into frames, so that each frame is distinguishable from another. Our postal system practices a type of framing. The simple act of inserting a letter into an envelope separates one piece of information from another; the envelope serves as the delimiter. In addition, each envelope defines the sender and receiver addresses since the postal system is a many-to-many carrier facility.

Framing in the data link layer separates a message from one source to a destination, or from other messages to other destinations, by adding a sender address and a destination address. The destination address defines where the packet is to go; the sender address helps the recipient acknowledge the receipt. Although the whole message could be packed in one frame, that is not normally done. One reason is that a frame can be very large, making flow and error control very inefficient. When a message is carried in one very large frame, even a single-bit error would require the retransmission of the whole message. When a message is divided into smaller frames, a single-bit error affects only that small frame.

Frames can be of fixed or variable size. In fixed-size framing, there is no need for defining the boundaries of the frames; the size itself can be used as a delimiter. An example of this type of framing is the ATM wide-area network, which uses frames of fixed size called cells. In variable-size framing, we need a way to define the end of the frame and the beginning of the next.

In this section we will look at four commonly used methods of framing:

1. Character count
2. Flag bytes with byte stuffing
3. Starting and ending flags, with bit stuffing
4. Physical layer coding violations.

1. Character count

The first framing method uses a field in the header to specify the number of characters in the frame. When the data link layer at the destination sees the character count, it knows how many characters follow and hence where the end of the frame is. This technique is shown in Fig. 3.2 (a) for four frames of sizes 5, 5, 8, and 8 characters, respectively.

The trouble with this algorithm is that the count can be garbled by a transmission error. For example, if the character count of 5 in the second frame of Fig. 3.2 (b) becomes a 7, the destination will get out of synchronization and will be unable to locate the start of the next frame. Even if the checksum is incorrect so the destination knows that the frame is bad, it still has no way of telling where the next frame starts. Sending a frame back to the source asking for a retransmission does not help either, since the destination does not know how many characters to skip over to get to the start of the retransmission. For this reason, the character count method is rarely used anymore.

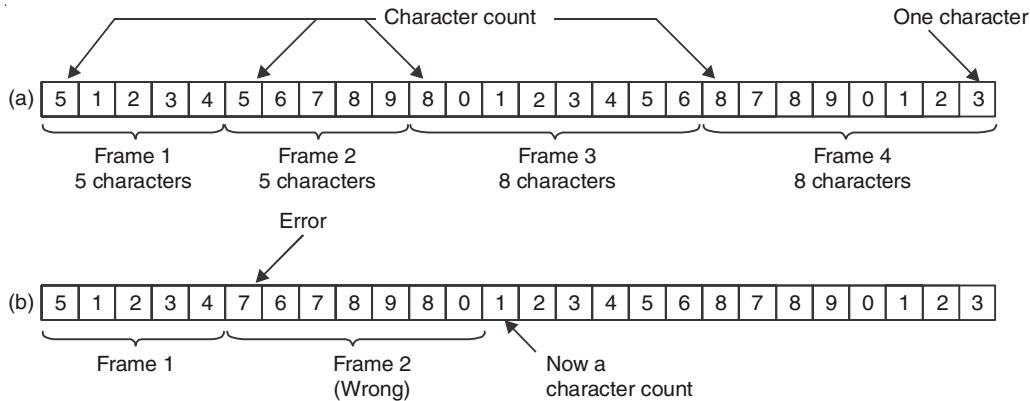


Fig. 3.2. A character stream. (a) Without errors. (b) With one error.

2. Flag bytes with byte stuffing

Character-oriented framing was popular when only text was exchanged by the data link layers. The flag could be selected to be any character not used for text communication. Now, however, we send other types of information such as graphs, audio, and video. Any pattern used for the flag could also be part of the information. If this happens, the receiver, when it encounters this pattern in the middle of the data, thinks it has reached the end of the frame. To fix this problem, a byte-stuffing strategy was added to character-oriented framing. In byte stuffing (or character stuffing), a special byte is added to the data section of the frame when there is a character with the same pattern as the flag. The data section is stuffed with an extra byte. This byte is usually called the escape character (ESC), which has a predefined bit pattern. Whenever the receiver encounters the ESC character, it removes it from the data section and treats the next character as data, not a delimiting flag.

Byte stuffing by the escape character allows the presence of the flag in the data section of the frame, but it creates another problem. What happens if the text contains one or more escape characters followed by a flag? The receiver removes the escape character, but keeps the flag, which is incorrectly interpreted as the end of the frame. To solve this problem, the escape characters that are part of the text must also be marked by another escape character. In other words, if the escape character is part of the text, an extra one is added to show that the second one is part of the text. Figure 3.3 shows the situation.

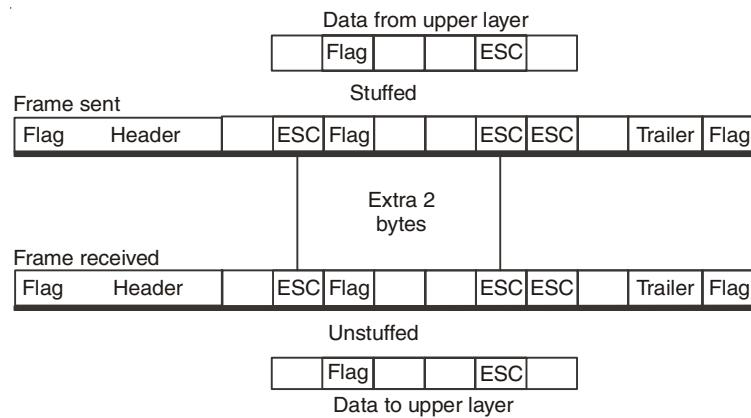


Fig. 3.3. Concept of Byte stuffing.

3. Starting and ending flags, with bit stuffing

This technique allows data frames to contain an arbitrary number of bits and allows character codes with an arbitrary number of bits per character. It works like this. Each frame begins and ends with a special bit pattern, 01111110 (in fact, a flag byte). Whenever the sender's data link layer encounters five consecutive 1s in the data, it automatically stuffs a 0 bit into the outgoing bit stream. This bit stuffing is analogous to byte stuffing, in which an escape byte is stuffed into the outgoing character stream before a flag byte in the data.

When the receiver sees five consecutive incoming 1 bits, followed by a 0 bit, it automatically destuffs (*i.e.*, deletes) the 0 bit. Just as byte stuffing is completely transparent to the network layer in both computers, so is bit stuffing. If the user data contain the flag pattern, 01111110, this flag is transmitted as 011111010 but stored in the receiver's memory as 01111110. Figure 3.4 gives an example of bit stuffing.

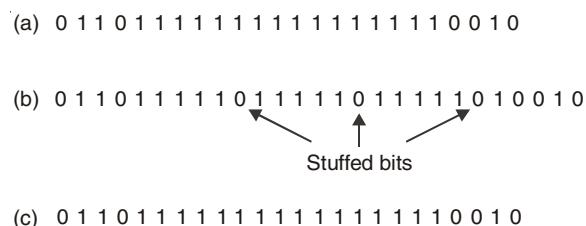


Fig. 3.4. Concept of Bit stuffing.

With bit stuffing, the boundary between two frames can be unambiguously recognized by the flag pattern. Thus, if the receiver loses track of where it is, all it has to do is scan the input for flag sequences, since they can only occur at frame boundaries and never within the data.

4. Physical layer coding violations

The last method of framing is only applicable to networks in which the encoding on the physical medium contains some redundancy. For example, some LANs encode 1 bit of data by using 2 physical bits. Normally, a 1 bit is a high-low pair and a 0 bit is a low-high pair. The scheme means that every data bit has a transition in the middle, making it easy for the receiver to locate the bit boundaries. The combinations high-high and low-low are not used for data but are used for delimiting frames in some protocols.

As a final note on framing, many data link protocols use a combination of a character count with one of the other methods for extra safety. When a frame arrives, the count field is used to locate the end of the frame. Only if the appropriate delimiter is present at that position and the checksum is correct is the frame accepted as valid. Otherwise, the input stream is scanned for the next delimiter.

3.3 ERROR DETECTION AND CORRECTION

Once the data is dispatched over the transmission medium the characteristics of the medium normally conspire to alter the transmitted data in various ways so that the signals received at the remote end of a link differ from the transmitted signals. The effects of these adverse

characteristics of a medium are known as **transmission impairments** and they often reduce transmission efficiency. In the case of binary data they may lead to errors, in that some binary zeros are transformed into binary ones and vice versa. To overcome the effects of such impairments it is necessary to introduce some form of error control. The first step in any form of error control is to detect whether any errors are present in the received data. Having detected the presence of errors there are two strategies commonly used to correct them: either further computations are carried out at the receiver to correct the errors, a process known as **forward error control**; or a message is returned to the transmitter indicating that errors have occurred and requesting a retransmission of the data, which is known as **feedback error control**. Error control is a function of the data link layer of the OSI reference model.

3.3.1 Types of Errors

Whenever bits flow from one point to another, they are subject to unpredictable changes because of interference. This interference can change the shape of the signal. In a single-bit error, a 0 is changed to a 1 or a 1 to a 0. In a burst error, multiple bits are changed. For example, a 11100 s burst of impulse noise on a transmission with a data rate of 1200 bps might change all or some of the 12 bits of information.

Single-bit error

The term *single-bit error* means that only 1 bit of a given data unit (such as a byte, character, or packet) is changed from 1 to 0 or from 0 to 1.

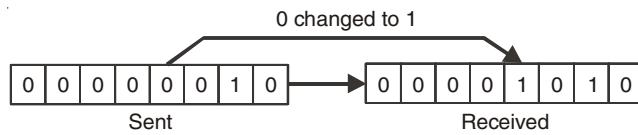


Fig. 3.5. Single-bit error.

Single-bit errors are the least likely type of error in serial data transmission. To understand why, imagine data sent at 1 Mbps. This means that each bit lasts only 1/1,000,000 s. For a single-bit error to occur, the noise must have a duration of only 1 micro sec., which is very rare; noise normally lasts much longer than this.

Burst error

The term *burst error* means that 2 or more bits in the data unit have changed from 1 to 0 or from 0 to 1. It is important to mention here that a burst error does not necessarily mean that the errors occur in consecutive bits. The length of the burst is measured from the first corrupted bit to the last corrupted bit. Some bits in between may not have been corrupted.

A burst error is more likely to occur than a single-bit error. The duration of noise is normally longer than the duration of 1 bit, which means that when noise affects data, it affects a set of bits. The number of bits affected depends on the data rate and duration of noise. For example, if we are sending data at 1 kbps, a noise of 11100 s can affect 10 bits; if we are sending data at 1 Mbps, the same noise can affect 10,000 bits.

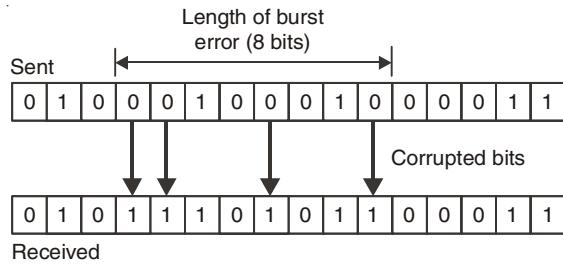


Fig. 3.6. Burst error.

3.3.2 Redundancy

The central concept in detecting or correcting errors is redundancy. To be able to detect or correct errors, we need to send some extra bits with our data. These redundant bits are added by the sender and removed by the receiver. Their presence allows the receiver to detect or correct corrupted bits.

3.3.3 Forward Error Control

The need to detect and correct errors was mentioned in Section 3.3.1 but not the means by which the detection process can be carried out. Error detection (and correction) is also known as **channel coding**. Channel coding is the process of coding data prior to transmission over a communications channel so that if errors do occur during transmission it is possible to detect and possibly even to correct those errors once the data has been received. In order to achieve this error detection/correction some bit patterns need to be identified as error free at the receiver, whereas other bit patterns will be identified as erroneous. To increase the number of identifiable bit patterns at the receiver above the bare minimum required to represent the data, additional bits, known as **redundant bits**, are added to the data or information bits prior to transmission. Various different types of code are available for use in channel coding but the most commonly used are called **linear block codes**.

3.3.4 Linear Block Codes

These constitute the simplest and most commonly used type of channel code. Data is transmitted as a fixed-length block. Prior to transmission the data is treated as a binary number and some form of linear mathematical process is carried out on a group of information bits so as to generate additional redundant bits which are known as **check bits**.

The check bits are transmitted along with the information bits, normally at the end of the block. At the receiver, a similar mathematical process is used to determine whether there are errors or not. Typical mathematical processes used are addition and division. A study of these codes inevitably requires some knowledge of mathematics but although the theory underlying the codes is complex, the following treatment has been kept fairly straightforward without being too simplistic.

3.3.4.1 Hamming codes

This is an important group of early error-correcting codes pioneered by R.W. Hamming in the 1950s. They involve the production of check bits by adding together different groups

of information bits. The type of addition used is known as modulo-2 addition and is equivalent to normal binary addition without any carries. The best way to see how the check bits are obtained is to consider a particular code as an example.

We shall consider a Hamming (7, 4) code, in which three check bits (c_1 , c_2 and c_3) are combined with four information bits (k_1 , k_2 , k_3 and k_4) to produce a block of data of length $n = 7$. This block of data is known as a **codeword**. A block of data of length 7 is too short to be appropriate for a practical data communications system, but the mathematics involved in longer blocks would become tedious. Three check equations are used to obtain the three check bits of this Hamming (7, 4) code as follows:

$$\begin{aligned}c_1 &= k_1 \oplus k_2 \oplus k_4 \\c_2 &= k_1 \oplus k_3 \oplus k_4 \\c_3 &= k_2 \oplus k_3 \oplus k_4\end{aligned}$$

where \oplus represents modulo-2 addition. The rules of modulo-2 addition are:

$$\begin{aligned}0 \oplus 0 &= 0 \\0 \oplus 1 &= 1 \\1 \oplus 1 &= 0 \text{ (no carry)}\end{aligned}$$

If we choose the information bits 1010 as an example then $k_1 = 1$, $k_2 = 0$, $k_3 = 1$ and $k_4 = 0$ and the check bits obtained from the three check equations above are as follows:

$$\begin{aligned}c_1 &= k_1 \oplus k_2 \oplus k_4 = 1 \oplus 0 \oplus 0 = 1 \\c_2 &= k_1 \oplus k_3 \oplus k_4 = 1 \oplus 1 \oplus 0 = 0 \\c_3 &= k_2 \oplus k_3 \oplus k_4 = 0 \oplus 1 \oplus 0 = 1\end{aligned}$$

The codeword is obtained by adding the check bits to the end of the information bits and therefore the data 1010101 will be transmitted (information bits first). A complete set of code words can be obtained in a similar way:

Codeword								Codeword							
no.	k_1	k_2	k_3	k_4	c_1	c_2	c_3	no.	k_1	k_2	k_3	k_4	c_1	c_2	c_3
0	0	0	0	0	0	0	0	8	1	0	0	0	1	1	0
1	0	0	0	1	1	1	1	9	1	0	0	1	0	0	1
2	0	0	1	0	0	1	1	10	1	0	1	0	1	0	1
3	0	0	1	1	1	0	0	11	1	0	1	1	0	1	0
4	0	1	0	0	1	0	1	12	1	1	0	0	0	1	1
5	0	1	0	1	0	1	0	13	1	1	0	1	1	0	0
6	0	1	1	0	1	1	0	14	1	1	1	0	0	0	0
7	0	1	1	1	0	0	1	15	1	1	1	1	1	1	1

An error that occurs in a transmitted codeword can be detected only if the error changes the codeword into some other bit pattern that does not appear in the code. This means that the code words transmitted over a channel must differ from each other in at least two bit positions. If two code words differ in only one position and an error occurs

in that position then one codeword will be changed into another codeword and there will be no way of knowing that an error has occurred. Inspection of the set of code words of the Hamming (7, 4) code reveals that they all differ from each other in at least three places. Taking code words 3 and 8 as an example, we have:

Codeword 3 0 0 1 1 1 0 0

Codeword 8 1 0 0 0 1 1 0

These two code words differ in positions 1, 3, 4 and 4 (counting from the left). The number of positions by which any two code words in a code differ is known as the **Hamming distance** or just the distance, so that the distance between these two words is four. Since all linear block codes contain the all-zeros codeword, then an easy way to find the **minimum distance** of a code is to compare a non-zero codeword which has the minimum number of ones with the all-zeros codeword. Thus, the **minimum distance** of a code is equal to the smallest number of ones in any non-zero codeword, which in the case of this Hamming (7, 4) code is three. If the code words of a code differ in three or more positions then error correction is possible since an erroneous bit pattern will be 'closer' to one codeword than another (this assumes that one error is more likely than two, two more likely than three, and so on). If we take codewords 8 and 10 as an example, we have:

Codeword 8 1 0 0 0 1 1 0

Codeword 10 1 0 1 0 1 0 1

The distance between these two codewords is three. If codeword 8 is transmitted and an error occurs in bit 3 then the received data will be:

1 0 1 0 1 1 0

This is not one of the other 15 Hamming (7, 4) codewords since an error has occurred. Furthermore, the most likely codeword to have been transmitted is codeword 8 since this is the nearest to the received bit pattern. Thus, it should also be possible to correct the received data by making the assumption that the transmitted codeword was number 8. If, however, a second error occurs in bit 7 then the received bit pattern will be:

1 0 1 0 1 1 1

It should still be possible to detect that an error has occurred since this is not one of the 16 codewords. However, it is no longer possible to correct the errors since the received bit pattern has changed in two places and is no longer closer to codeword 8 than any other (it is, in fact, now closer to codeword 10). Thus, this Hamming (7, 4) code is able to detect two errors but correct only one error. In general, if the minimum distance of a code is d , then $d-1$ errors can normally be detected using a linear block code and $\text{mod}(d-1)/2$ can be corrected.

A feature of all linear block codes which arises out of the mathematical rules used to determine the check bits is that all the codewords are related by these rules. Received data which contains errors no longer conforms to the mathematical rules, and it is this fact that is used to carry out the detection and correction processes at the receiver. If we take the

example of the Hamming (7,4) code then the encoding process will restrict the number of different codewords that can be transmitted to the 16 listed above. As a result of errors that may occur in the transmission process, the data arriving at a receiver in 7-bit blocks can have any one of $2^7 = 128$ different 7-bit patterns. This allows the receiver to detect whether errors have occurred since it is aware of the rules used in the encoding process and can apply them to see whether the received data is one of the 14 ‘legal’ codewords or not.

Furthermore, it is the case that all Hamming codes (indeed, all linear block codes) possess the mathematical property that if we add any two codewords together (modulo-2 addition) then the resulting sum is also a codeword. For example, if we add codewords 1 and 2 from the earlier list we obtain:

$$\begin{array}{r} 0\ 0\ 0\ 1\ 1\ 1 \\ \oplus\ 0\ 0\ 1\ 0\ 0\ 1 \\ \hline 0\ 0\ 1\ 1\ 1\ 0 \end{array}$$

which is codeword 3

This allows us to represent a whole code by means of a small ‘subset’ of codewords, since further codewords can simply be obtained by modulo-2 addition. In the case of the Hamming (7,4) code this is not important, since there are only 14 codewords. However, with longer block lengths the number of codewords becomes unmanageable. For example, a short block of 32 bits involves $2^{32} = 4\ 294\ 967\ 296$ different codewords.

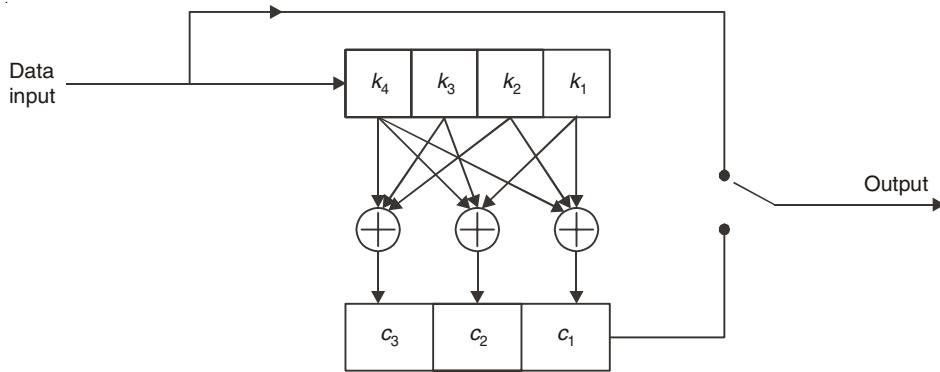
The subset of codewords is often expressed as a matrix known as a **generator matrix**, G . The codewords chosen are normally powers of 2, that is codewords 1, 2, 4, 8, A suitable generator matrix for the Hamming (7,4) code consists of the following four codewords:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

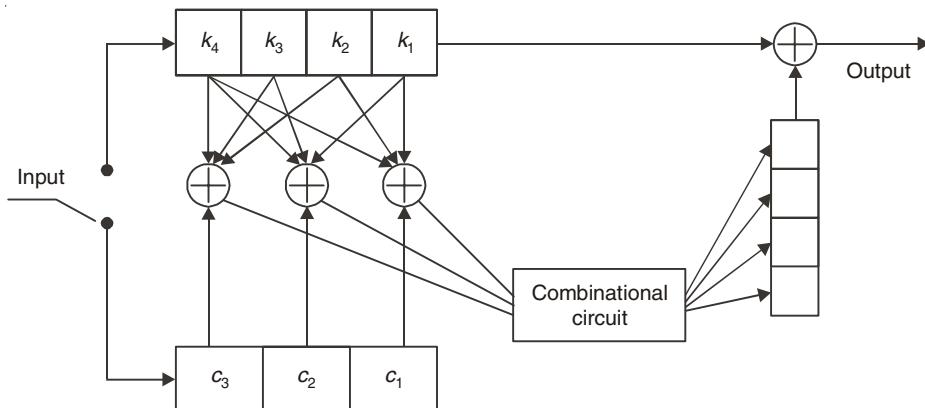
The matrix has four rows and seven columns, that is it has dimensions 4×7 ($k \times n$). The whole code can be generated from this matrix just by adding together rows, and it is for this reason that it is called a generator matrix. A further reason for the generator matrix being so named is that it can be used to generate codewords directly from the information bits without using the check equations. This is achieved by multiplying the information bits by the generator matrix using matrix multiplication.

Encoding and decoding circuits

An attractive feature of Hamming codes in the early days of error correction was that they could be easily implemented in hardware circuits, particularly if the block length was fairly short. Since modulo-2 addition is identical to an exclusive-or (EX-OR) function, a number of multiple input EX-OR gates can be used to determine the check bits, as in the circuit for a Hamming (7,4) encoder shown in Fig. 3.7.

**Fig. 3.7. Hamming Encoder.**

The information bits are fed into a 4-bit shift register and the check bits are calculated by the EX-OR circuits and are held in a 3-bit shift register. The switch is in the up position to transmit the information bits and down for the check bits. Figure 3.8 shows a corresponding decoder.

**Fig. 3.8. Hamming Decoder.**

With the receive switch up the information bits are received into a 4-bit shift register and with it down the check bits flow into a 3-bit shift register. The left-hand EX-OR gate works out the modulo-2 sum $k_2 \oplus k_3 \oplus k_4 \oplus c_3$. This equals zero if no errors have occurred during transmission. The output from the three EX-OR gates thus represents the syndrome which is fed into a combinational logic circuit to determine the position of any error. The error is corrected by means of a k -stage shift register at the output of the combinational logic circuit. This will contain all zeros apart from any erroneous position which will contain a one. The output from the shift register is added (modulo-2) serially to the received information bits and any bit position within the received data which gets added to a one will change from 0 to 1 (or 1 to 0), thus correcting the error. Unfortunately, circuits such as these, although simple in the case of the Hamming (7,4) code, become excessively complicated for the longer block lengths used in data communications networks. Hamming codes do, however, find uses in situations which do not require large block lengths, such as remote control of robotic systems.

3.3.4.2 Cyclic codes

As mentioned above, simple linear codes such as the Hamming code have a limitation in that if large block lengths are used, for example in data communications, then the encoding and decoding circuitry becomes very complex. Paradoxically, the circuitry can be made simpler if the mathematical structure of the code is made more complex.

A cyclic code is one in which all the codewords are related by the fact that if a codeword is rotated, it becomes another codeword. The following code is obtained from the single check equation $c_1 = k_1 \oplus k_2$ and, although trivial, is cyclic:

k_1	k_2	c_1
0	0	0
0	1	1
1	0	1
1	1	0

All four of these codewords can be rotated in either direction and will result in another codeword. Consequently, to define this code, it is only necessary to have one non-zero codeword, since all the other codewords can be obtained from it (the all zeros codeword is obtained by adding any codeword to itself). Cyclic codes are usually defined by a single codeword expressed as a polynomial, known as a **generator polynomial**. For example, the cyclic code used in the High-Level Data Link Control (HDLC) protocol has a generator polynomial $x^{16} + x^{12} + x^5 + 1$, where $x = 2$ since the code is binary. This is expressed as a polynomial rather than the binary number 10001000000100001 because the latter is rather unmanageable. The highest power of a generator polynomial is called its **degree** and is always equal to the number of check bits in the code. Since cyclic codes are invariably linearblock codes, they can also be described by a generator matrix, which can be readily obtained from the generator polynomial.

The fact that the codewords of a cyclic code are obtained by shifting and adding the generator polynomial leads to the important characteristic that all codewords are a multiple of the generator polynomial. To encode incoming information bits, a cyclic encoder must therefore generate check bits which, when added to the information bits, will produce a codeword which is a multiple of $G(x)$, the generator polynomial. This is achieved as follows: firstly, the information bits, normally a lengthy bit pattern, are represented as a polynomial $K(x)$, where x is, in practice, 2. Secondly, let the information bits $K(x)$, followed by c zeros (*i.e.*, a codeword with all the check bits set to zero), be represented by the polynomial $F(x)$ which is in fact $K(x)$ shifted by c places, that is $x^c K(x)$. If $F(x)$ is now divided by the generator polynomial $G(x)$ then:

$$\frac{F(x)}{G(x)} = Q(x) + \frac{R(x)}{G(x)}$$

where $Q(x)$ is the quotient and $R(x)$ the remainder, that is: $F(x) = Q(x)G(x) + R(x)$

If the remainder $R(x)$ is now added (modulo-2) to $F(x)$, we obtain:

$$F(x) + R(x) = Q(x)G(x)$$

since addition and subtraction will give the same result in modulo-2. It is this bit sequence, $F(x) + R(x)$, which is transmitted, since it is always a multiple of the generator polynomial $G(x)$ and is therefore always a legal codeword. Thus, encoding for a cyclic code consists of adding c zeros to the end of the information bits and dividing by the generator polynomial to find the remainder. (Note that modulo-2 division is used.) The remainder is added to the information bits in place of the c zeros and the resulting codeword transmitted. At the receiver, a decoder tests to see whether the received bit sequence is error free by dividing again by the generator polynomial. An error-free transmission results in a zero remainder. Such a process is also known as a **cyclic redundancy check (CRC)**.

Modulo-2 division circuits

Modulo-2 division is achieved in an electronic circuit by repeated shifting and subtraction. This can be very easily implemented using shift registers and, bearing in mind that modulo-2 subtraction is identical to addition, EX-OR gates. A circuit that will divide by $1 + x + x^3$ is given in Fig. 3.9.

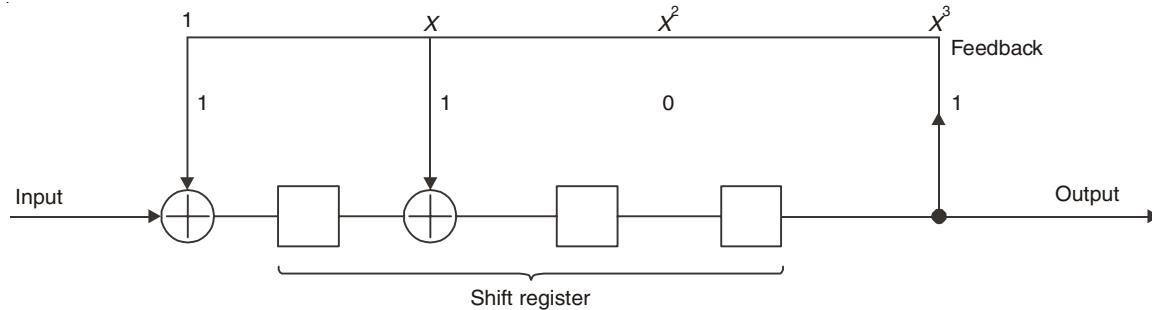


Fig. 3.9. Modulo-2 Division Circuit.

Table 3.1. Modulo-2 Division

Step J	Input on Jth shift	Feedback on Jth shift	Shift register after Jth shift	Output after Jth shift
0	–	–	0 0 0	0
1	1	0	1 0 0	0
2	1	0	1 1 0	0
3	0	0	0 1 1	1
4	0	1	1 1 1	1
5	0	1	1 0 1	1
6	1	1	0 0 0	0
7	0	0	0 0 0	0

Cyclic encoding and decoding circuits

Encoding for a cyclic code consists of dividing the polynomial $F(x)$ (incoming information bits with zeros in the check bit positions) by the generator polynomial $G(x)$ to find a remainder $R(x)$ which is then added to $F(x)$ to give the codeword $F(x) + R(x)$. To obtain the

zero check bits, the information bits $K(x)$ are shifted c places prior to the division process. An encoding circuit, known as a Meggitt encoder, that achieves this for $G(x) = 1 + x + x^3$ is shown in Fig. 3.10.

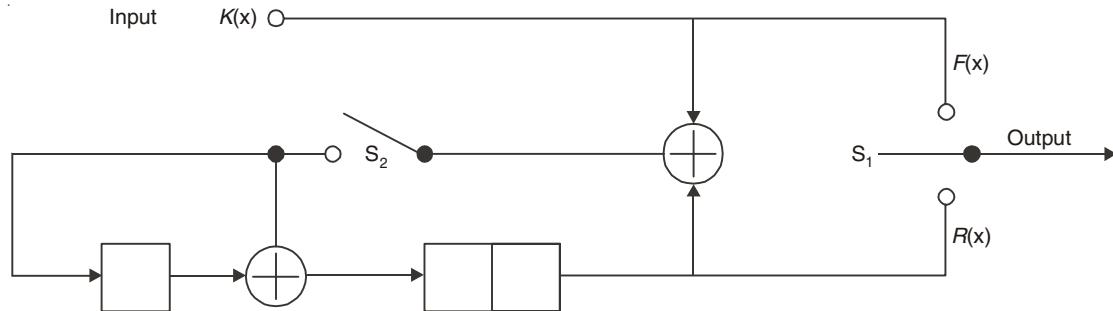


Fig. 3.10. Cyclic Encoding circuit.

Bringing the input information bits $K(x)$ into the circuit at the point shown rather than to the left of the circuit as in the division circuit of Fig. 3.9 is equivalent to shifting c places ($c = 3$ in this case). Initially S_1 is in the up position and S_2 is closed, while $F(x)$ is simultaneously transmitted and processed in the division circuit. After k shifts, the shift register will contain the remainder $R(x)$. S_1 then moves to the down position and S_2 opens so that the remainder is then transmitted.

A similar arrangement is shown in the decoder circuit of Fig. 3.11. Initially the information bits are received into a k -stage shift register with S_1 and S_2 closed. At the same time, the check bits are recalculated by the division circuit and then fed into a combinational logic circuit. The check bits are received with S_1 and S_2 open and are fed straight into the combinational logic circuit where they are compared with the recalculated check bits to determine a syndrome. S_1 and S_2 are then closed again and the syndrome is used by the combinational circuit to correct the data at the output of the decoding circuit.

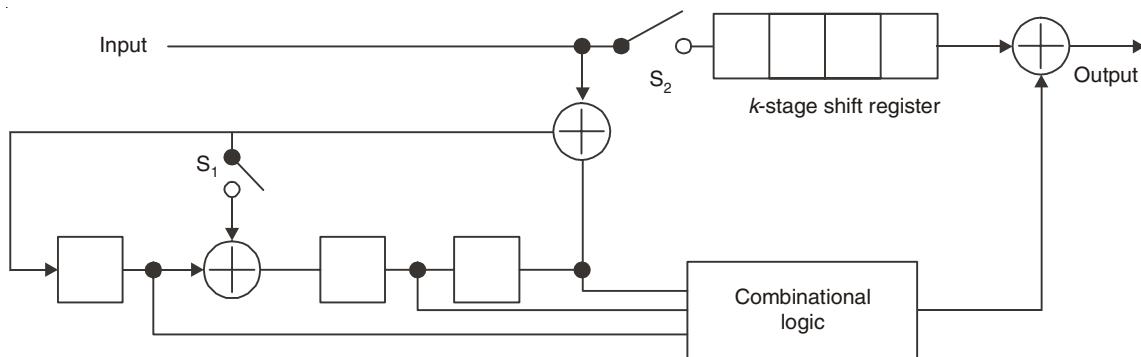


Fig. 3.11. Cyclic Decoding circuit.

Practical cyclic redundancy checks

Although the material in this section appears quite theoretical, it is of great practical importance to most data communications networks. This is because most networks use a

cyclic redundancy check as part of the level 2 protocol error checking. Table 3.2 shows some commonly used generator polynomials.

Table 3.2. Commonly used CRC

<i>Protocol(s)</i>	<i>Generator polynomial</i>
ATM (Header error check)	$x^8 + x^2 + x + 1$
ATM (OAM cells)	$x^{10} + x^9 + x^5 + x^4 + x + 1$
ATM adaptation layer 1	$x^3 + x + 1$
ATM adaptation layer 2	$x^5 + x^2 + x + 1$
Ethernet (and FDDI)	$x^{32} + x^{24} + x^{23} + x^{22} + x^{14} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$
X.25 (HDLC)	$x^{16} + x^{12} + x^5 + 1$

Advantages of Cyclic Codes

Cyclic codes have a very good performance in detecting single-bit errors, double errors, an odd number of errors, and burst errors. They can easily be implemented in hardware and software. They are especially fast when implemented in hardware. This has made cyclic codes a good candidate for many networks.

3.3.4.3 Convolutional codes

In order to carry out a cyclic redundancy check on a block of data, it is necessary for the complete block of data to be received into an input buffer at a network node within a data communications network. This is convenient for most data communications protocols, including IP, which transmit data in the form of blocks known as packets or frames. Convolutional codes work in a fundamentally different manner in that they operate on data continuously as it is received (or transmitted) by a network node. Consequently, convolutional encoders and decoders can be constructed with a minimum of circuitry. This has meant that they have proved particularly popular in mobile communications systems where a minimum of circuitry and consequent light weight are a great advantage. It should be noted that these considerations have become less significant with advances in electronic circuitry. As well as carrying out encoding and decoding continuously on a transmitted or received bit stream, the way a convolutional code treats a particular group of bits depends on what has happened to previous groups of bits. A simple convolutional encoder is illustrated in Fig. 3.12.

The data to be encoded is entered into a 3-bit shift register one bit at a time. For each bit that is fed in at the input two bits are transmitted from the output, one each from terminals X and Y. Table 3.3 shows the outputs that will appear as a result of different combinations of existing shift register contents and new input values.

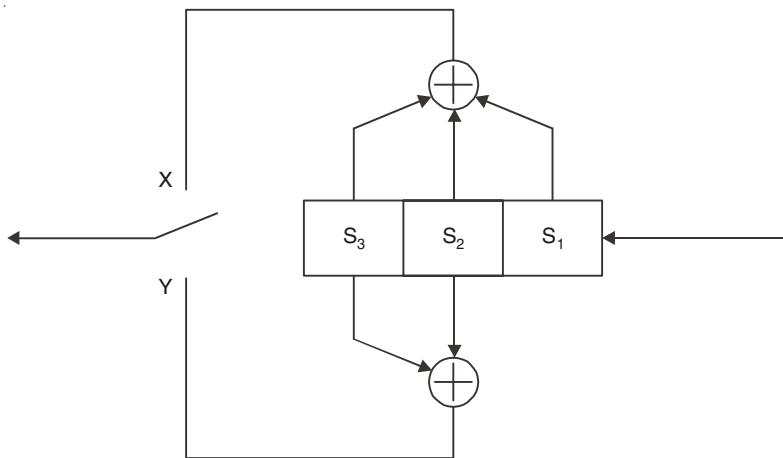


Fig. 3.12. A simple convolution encoder.

Table 3.3. Convolution Encoder Output Values

<i>Shift register</i> <i>S3 S2 S1</i>	<i>New input</i>	<i>Output</i> <i>X Y</i>	<i>Shift register</i> <i>S3 S2 S1</i>	<i>New input</i>	<i>Output</i> <i>X Y</i>
0 0 0	0	0 0	0 0 0	1	1 0
0 0 1	0	1 1	0 0 1	1	0 1
0 1 0	0	1 1	0 1 0	1	0 1
0 1 1	0	0 0	0 1 1	1	1 0
1 0 0	0	0 0	1 0 0	1	1 0
1 0 1	0	1 1	1 0 1	1	0 1
1 1 0	0	1 1	1 1 0	1	0 1
1 1 1	0	0 0	1 1 1	1	1 0

This table is of limited usefulness since it does not give any indication of the output resulting from a prolonged stream of bits at the input. In order to observe this a slightly more complicated representation, called a **Trellis diagram**, is used. There are two branches from every point in a Trellis diagram. The upper branch represents 0 input, and the lower branch represents 1 input, and the resulting output is written alongside each branch. This is illustrated in Fig. 3.13 that represents the first row of Table 3.3. At point A, it is assumed that the contents of the shift register are 000. Thus, the branch AB that represents a 0 input has an output of 00 alongside it since that is the output indicated in Table 3.3 for these values. Similarly, the lower branch AC that represents a 1 input has a value 01 alongside it. A full Trellis diagram, so called because it resembles a garden trellis, for the above encoder is illustrated in Fig. 3.14.

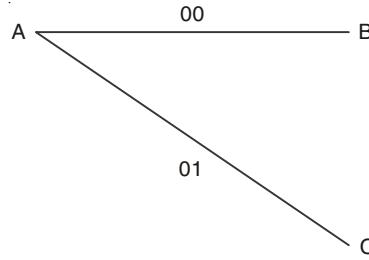


Fig. 3.13. Initial Part of Trellis diagram.

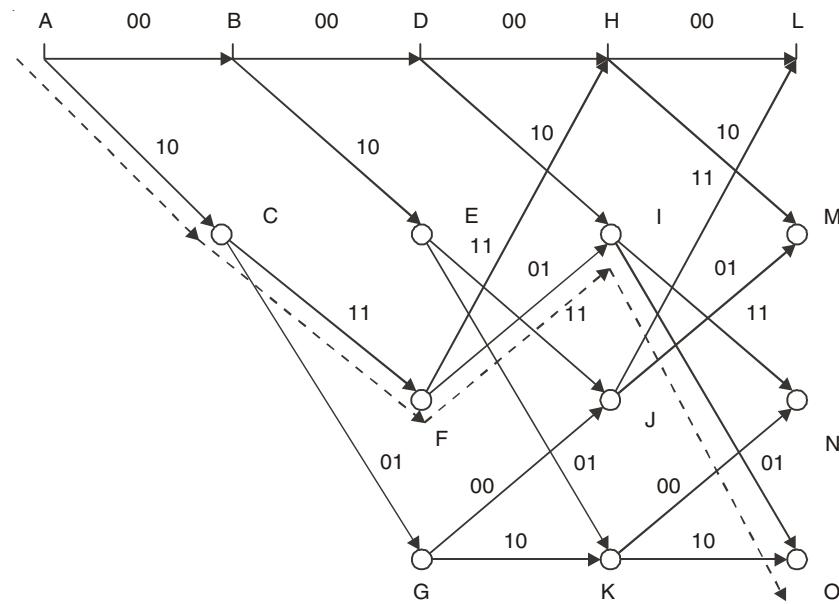


Fig. 3.14. Trellis diagram.

Note that the pattern of branches AB and AC is repeated all along the top of the trellis diagram with branches BE and BD and so on. This is because the pair of output bits will always be 00 as long as the shift register contains 000 and a zero arrives at the input. Branches CF and CG represent the second row of Table 3.3 where the contents of the shift register are 001; branches FH and FI represent the third row where the contents of the shift register are 010. The Trellis diagram can be used to indicate the output produced by a particular input bit stream by drawing a line through the trellis to represent the input bit stream. This is illustrated by the line in Fig. 3.14 that follows the path ACFIO representing an input of 1011. To obtain the resulting output, we note the pairs of digits along path ACFIO, namely 10110101.

Decoding and error correction

The principle involved in decoding convolutional codes and subsequently correcting errors is the same as that involved in linear block codes; that is, error-free transmission will produce certain received bit patterns (codewords in the terminology of linear block codes)

and errors are likely to produce different bit patterns that cannot possibly be received in error-free transmissions. An inspection of Fig. 3.14 immediately reveals that certain bit patterns are not possible. For example, consider the bit pattern of 11111110 that cannot be obtained from any of the paths through the trellis. If such a pattern is received in a system using the decoder of Fig. 3.12 then it is assumed that an error (or errors) has occurred in transmission.

If no errors occur during transmission, the path taken through the trellis diagram will be the same at the receiver as at the transmitter. Errors that occur during transmission will cause the path taken at the receiver to be different, and will eventually result in an impossible path. At this point, the receiver can notify the transmitter of an error and it may be able to correct the error by deducing the path that would have been the most likely in the absence of errors. In order to do this we need to have a measure of the different paths through the trellis. In the linear block codes, we used the concept of Hamming distance, which was defined as the number of positions in which two codewords differ and the minimum distance, that is the number of positions that a codeword differs from the all-zero codeword. In convolutional codes, we retain the term **distance** to indicate the number of positions in which two paths differ and we define the **weight** as the number of bit positions in which a path differs from the all-zero path, that is the number of logical ones that appear in a particular path. Thus, in Fig. 3.14, path ACFIN is defined by the output sequence 10110111 and therefore possesses a weight of 6 since it contains six logical ones. We can also define the **free distance**, d_{free} , as the smallest weight of any path that diverges from the all-zero path and subsequently returns to it. This will be comparable with the minimum Hamming distance of a linear block code and is used to define the error detecting/correcting capabilities of the code as follows:

$d_{\text{free}} - 1$ errors can be detected.

$\text{mod}(d_{\text{free}} - 1)/2$ errors can be corrected.

An algorithm for decoding convolutional codes was developed by Viterbi (1967). In order to understand the algorithm, let us consider the received bit pattern 11111110 mentioned earlier. It is immediately obvious that this is an erroneously received bit pattern as the first two bits are not compatible with the trellis diagram. The Viterbi algorithm takes the received bit stream, compares it with the trellis diagram and calculates the distance between each line in the trellis diagram and the received bit pattern. This is illustrated in Fig. 3.15.

If we look at the first stage of the trellis we note that AB is marked 2 and AC is marked 1. This is because AB represents the output 00 which is a distance of 2 from the first two received bits (11) and AC represents 10 which is a distance of 1 from the first two received bits. The other stages of the trellis are marked in the same way. The Viterbi algorithm now computes the minimal path through this trellis. It does this in a step-by-step fashion, the details of which are beyond the scope of this text.

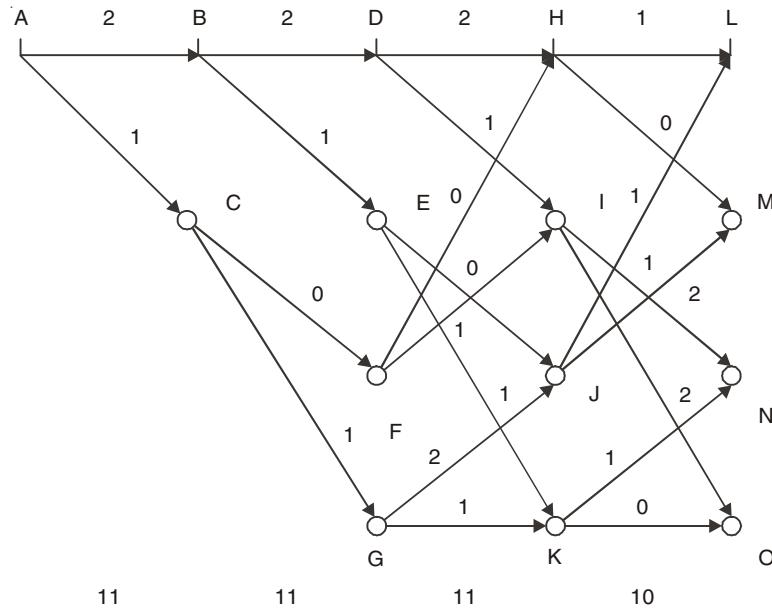


Fig. 3.15. Trellis showing distance from received bits.

In this particular case, the minimal path is ACFHM which should have a distance of 1 from the received bit sequence of 11111110. It can be seen from Fig. 3.15 that path ACFHM represents an output of 10111110, which is indeed a distance of 1 from the received bit sequence. A convolutional code receiver using the Viterbi algorithm would therefore assume that the error-free received data was 10111110 and decode it accordingly as 1001. It is worth noting that, like most other error-correcting systems, the algorithm relies on the fact that one error is more likely than two, two errors are more likely than three, and so on.

3.4 FEEDBACK ERROR CONTROL

Even very powerful error-correcting codes may not be able to correct all errors that arise in a communications channel. Consequently, many data communications links provide a further error control mechanism, in which errors in data are detected and the data is retransmitted. This procedure is known as **feedback error control** and it involves using a channel code to detect errors at the receive end of a link and then returning a message to the transmitter requesting the retransmission of a block (or blocks) of data. Alternatively, errors in received data can be detected and corrected up to a certain number of errors and a retransmission requested only if more than this number of errors occurs. The process of retransmitting the data has traditionally been known as **Automatic Repeat Request (ARQ)**. There are three types of ARQ that have been used: namely, **stop-and-wait**, **go-back-n** and **selective-repeat**.

Stop-and-wait ARQ

This technique is the simplest method of operating ARQ and ensures that each transmitted block of data or **frame** is correctly received before sending the next. At the

receiver, the data is checked for errors and if it is error free an acknowledgement (ACK) is sent back to the transmitter. If errors are detected at the receiver a negative acknowledgement (NAK) is returned. Since errors could equally occur in the ACK or NAK signals, they should also be checked for errors. Thus, only if each frame is received error free and an ACK is returned error free can the next frame be transmitted.

Case I: If, however, errors are detected either in the transmitted frame or in the returned acknowledgement, then the frame is retransmitted.

Case II: A further (and hopefully remote) possibility is that a frame or acknowledgement becomes lost for some reason. To take account of this eventuality, the transmitter should retransmit if it does not receive an acknowledgement within a certain time period known as the **timeout interval**.

Case III: Finally, a frame may be received correctly but the resulting ACK may be lost, resulting in the same frame being transmitted a second time. This problem can be overcome by numbering frames and discarding any correctly received duplicates at the receiver.

Although stop-and-wait offers a simple implementation of an ARQ system it can be inefficient in its utilization of the transmission link since time is spent in acknowledging each frame.

Go-back-n ARQ

If the error rate on a link is relatively low, then the link efficiency can be increased by transmitting a number of frames continuously without waiting for an immediate acknowledgement. This strategy is used in go-back-n ARQ which is used in a number of standard protocols including HDLC (High-level Data Link Control) and, in modified form, TCP (Transmission Control Protocol), both of which are dealt with later in the text. The go-back number n determines how many frames can be transmitted without an acknowledgement having been received and is often referred to as a transmission **window**.

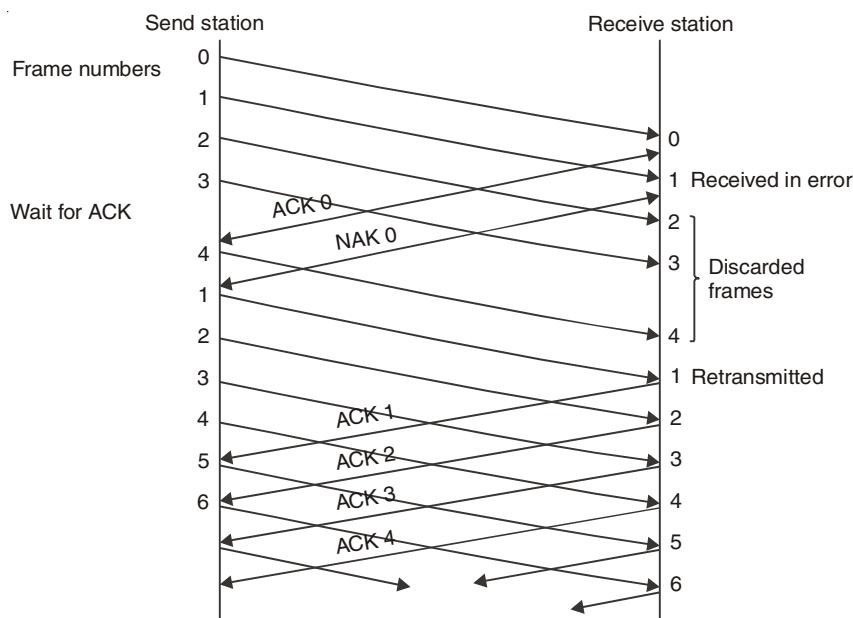


Fig. 3.16. Frame transfer go-back-n.

Frame $i + n$ cannot be transmitted until frame i has been acknowledged. Figure 3.16 shows a typical transfer of frames for a go-back-4 ARQ system over a full-duplex link.

Note that the send node transmits four frames (numbered 0, 1, 2, 3) without receiving an acknowledgement. There then follows a short wait before receipt of the first ACK which contains the number of the last correctly received frame (0 in this case). Although this acknowledgement is returned immediately after the correct receipt of frame 0, transmission delays mean that it does not reach the send node until after frame 3 has been transmitted (the send node having carried on transmitting as a result of the go-back-4 strategy). If we now assume that an error occurs in frame 1, the receive node will reply with NAK 0, indicating that an error has been detected and that the last correctly received frame was frame 0. The send node will now *go back* and retransmit the frame after the last correctly received frame, which in this case is frame 1. Meanwhile the receive node will have received frames 2, 3 and 4 which, since they have not been acknowledged, need to be discarded. In the absence of further errors, the transmit node continues transmitting frames 2, 3, 4, 5, ... for as long as ACK signals continue to be returned. A go-back- n strategy allows for the efficient transfer of frames without the need for any substantial buffer storage by receive equipment as long as error rates are low. However, if errors occur during transmission then frames that have already been received correctly need to be retransmitted along with erroneous frames. If buffer storage is available at the receive equipment then it seems reasonable that some form of selective retransmission strategy would be more efficient than go-back- n in the presence of errors.

Selective-repeat ARQ

In selective-repeat ARQ only those frames that generate a NAK are retransmitted. Although this appears more efficient than go-back- n , it requires sufficient storage at the transmitter to save all frames that have been transmitted but not acknowledged in case a frame proves to be erroneous. In this system a go-back number is still used to determine how many frames can be transmitted without receiving an acknowledgement. In the past, selective-repeat ARQ was not a particularly popular ARQ strategy because of the increased memory requirements mentioned, but, as memory capabilities have become more readily available, it offers potential efficiency gains in the presence of high error rates.

3.5 FLOW CONTROL

Flow control is a technique for assuring that a transmitting entity does not overwhelm a receiving entity with data. The receiving entity typically allocates a data buffer of some maximum length for a transfer. When data are received, the receiver must do a certain amount of processing before passing the data to the higher-level software. In the absence of flow control, the receiver's buffer may fill up and overflow while it is processing old data. Flow control coordinates the amount of data that can be sent before receiving an acknowledgement and is one of the most important duties of the data link layer. In most protocols, flow control is a set of procedures that tells the sender how much data it can transmit before it must wait for an acknowledgement from the receiver.

3.5.1 Simplest Protocol for Noiseless Channel

Our first protocol, which we call the Simplest Protocol for lack of any other name, is one that has no flow or error control. It is a unidirectional protocol in which data frames are traveling in only one direction—from the sender to receiver. We assume that the receiver can immediately handle any frame it receives with a processing time that is small enough to be negligible. The data link layer of the receiver immediately removes the header from the frame and hands the data packet to its network layer, which can also accept the packet immediately. In other words, the receiver can never be overwhelmed with incoming frames.

The data link layers of the sender and receiver provide transmission services for their network layers. The data link layers use the services provided by their physical layers (such as signaling, multiplexing, and so on) for the physical transmission of bits. Fig. 3.17 shows a design.

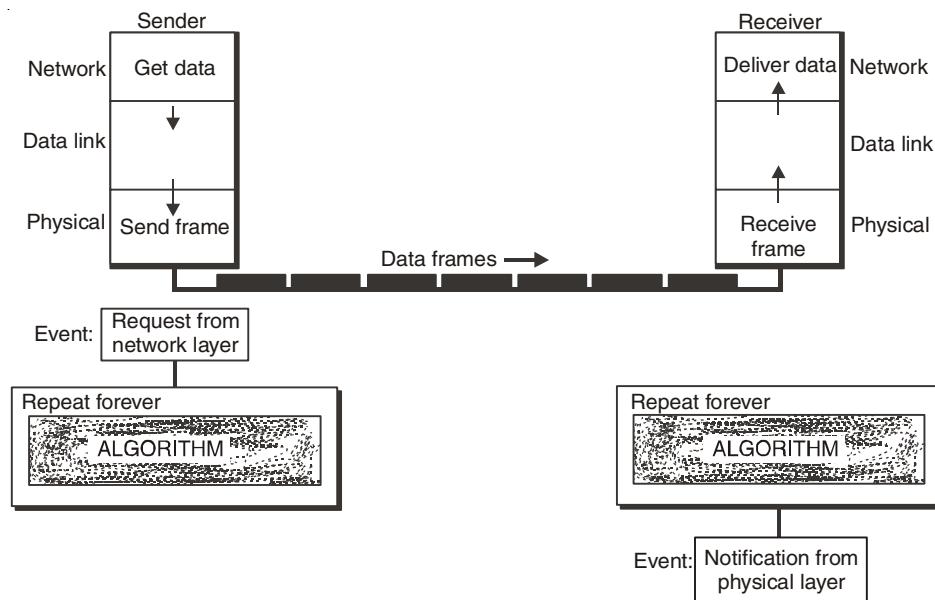


Fig. 3.17. A Simplex protocol for noiseless channel.

3.5.2 Stop-and-Wait Protocol

If data frames arrive at the receiver site faster than they can be processed, the frames must be stored until their use. Normally, the receiver does not have enough storage space, especially if it is receiving data from many sources. This may result in either the discarding of frames or denial of service. To prevent the receiver from becoming overwhelmed with frames, we somehow need to tell the sender to slow down. There must be feedback from the receiver to the sender. The protocol we discuss now is called the Stop-and-Wait Protocol because the sender sends one frame, stops until it receives confirmation from the receiver (okay to go ahead), and then sends the next frame. We still have unidirectional communication for data frames, but auxiliary ACK frames (simple tokens of acknowledgement) travel from the other direction. We add flow control to our previous protocol.

Figure 3.18 illustrates the mechanism. Comparing this figure with Fig. 3.17, we can see the traffic on the forward channel (from sender to receiver) and the reverse channel. At any time, there is either one data frame on the forward channel or one ACK frame on the reverse channel. We therefore need a half-duplex link.

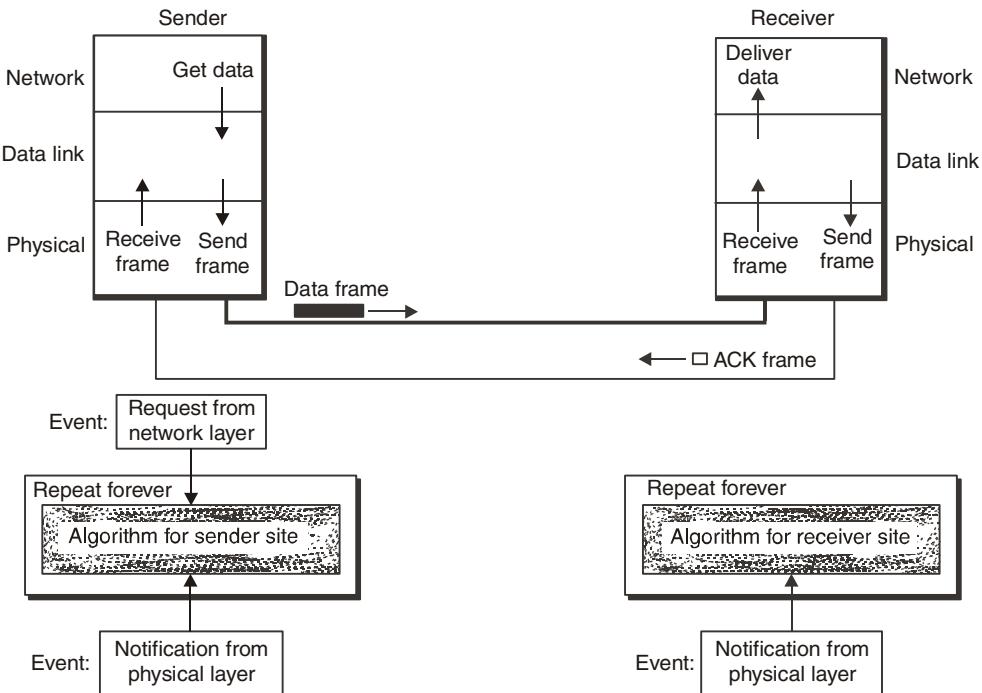


Fig. 3.18. A Stop and wait Protocol.

This procedure works fine and, indeed, can hardly be improved upon when a message is sent in a few large frames. However, it is often the case that a source will break up a large block of data into smaller blocks and transmit the data in many frames. This is done for the following reasons:

- The buffer size of the receiver may be limited.
- The longer the transmission, the more likely that there will be an error, necessitating retransmission of the entire frame. With smaller frames, errors are detected sooner, and a smaller amount of data needs to be retransmitted.
- On a shared medium, such as a LAN, it is usually desirable not to permit one station to occupy the medium for an extended period, as this causes long delays at the other sending stations.

With the use of multiple frames for a single message, the stop-and-wait procedure may be inadequate. The essence of the problem is that only one frame at a time can be in transit. In situations where the bit length of the link is greater than the frame length, serious inefficiencies results.

3.5.3 Window Mechanisms

If a frame-oriented system uses a continuous form of ARQ such as go-back- n or selective-repeat then, as long as information is available for transmission, a send node can continue sending information frames before receiving an acknowledgement. The send node will be provided with a predetermined amount of buffer storage and it is important that this storage does not become overloaded. It is common, therefore, for an additional flow control mechanism to be used with these systems that limits the number of information frames that can be transmitted before receiving an acknowledgement. The send node keeps a copy of those frames transmitted but not acknowledged so that it can retransmit if necessary. A maximum limit is set on the number of copies that are being held at the send node which is known as the **send window**.

If the send node reaches its maximum window size it stops transmitting and, in the absence of any acknowledgements, it does not transmit any more frames. When the send node finally receives an acknowledgement it can start transmitting again. The window size is chosen so that it does not impede the flow of frames. As well as the amount of send buffer storage available, the frame size and transmission rate are also taken into account in determining the window size. The operation of a send window is illustrated in Fig. 3.19.

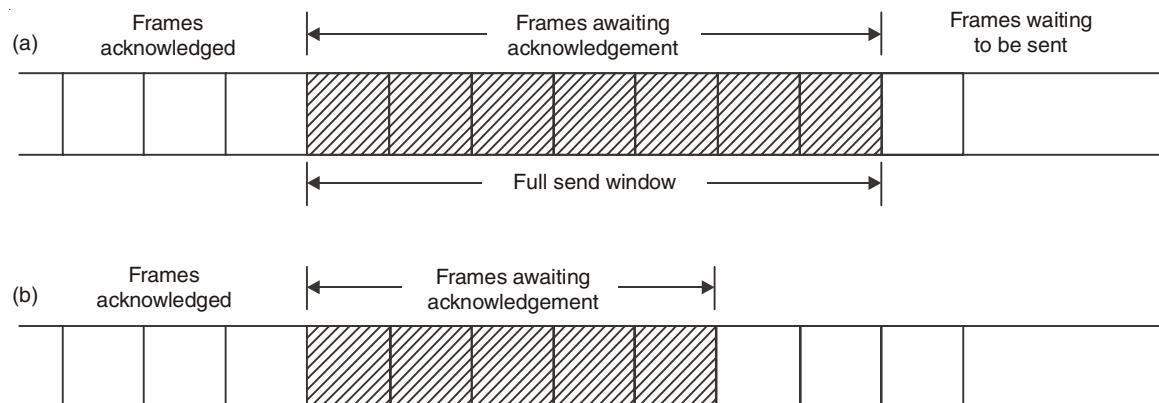


Fig. 3.19. Operation of send window: (a) window full; (b) continuous flow possible.

A list of information frames transmitted but, as yet, unacknowledged is kept at the send node. Each time an information frame is transmitted the list is increased by one and each time an acknowledgement is received it is reduced by one.

If two-way data transmission is used then a window is required at each end of the link. In this case a technique known as **piggybacking** is often used, in which an acknowledgement signal is returned inside an information frame. In links in which propagation delay is high, piggybacking improves the link throughput considerably since separate acknowledgements can be dispensed with. When a data frame arrives, instead of immediately sending a separate control frame, the receiver restrains itself and waits until the network layer passes it the next packet. The acknowledgement is attached to the outgoing data frame (using the ack field in the frame header). In effect, the acknowledgement gets a free ride on the next outgoing data frame. The technique of temporarily delaying outgoing acknowledgements so that they can be hooked onto the next outgoing data frame is known as piggybacking.

The principal advantage of using piggybacking over having distinct acknowledgement frames is a better use of the available channel bandwidth. The ack field in the frame header costs only a few bits, whereas a separate frame would need a header, the acknowledgement, and a checksum. In addition, fewer frames sent means fewer “frame arrival” interrupts, and perhaps fewer buffers in the receiver, depending on how the receiver’s software is organized. In the next protocol to be examined, the piggyback field costs only 1 bit in the frame header. It rarely costs more than a few bits.

However, piggybacking introduces a complication not present with separate acknowledgements. How long should the data link layer wait for a packet onto which to piggyback the acknowledgement? If the data link layer waits longer than the sender’s timeout period, the frame will be retransmitted, defeating the whole purpose of having acknowledgements. If the data link layer were an oracle and could foretell the future, it would know when the next network layer packet was going to come in and could decide either to wait for it or send a separate acknowledgement immediately, depending on how long the projected wait was going to be. Of course, the data link layer cannot foretell the future, so it must resort to some *ad hoc* scheme, such as waiting a fixed number of milliseconds. If a new packet arrives quickly, the acknowledgement is piggybacked onto it; otherwise, if no new packet has arrived by the end of this time period, the data link layer just sends a separate acknowledgement frame.

3.6 INTRODUCTION OF LINK MANAGEMENT

The flow control and the error control techniques are both concerned with the transfer of data across a link. Flow control ensures that the data is in the correct sequence and error control that the data is received correctly without errors. However, before either of these functions can take place the link needs to be set up and after the data transfer has taken place it needs to be disconnected. These two functions are known as **link management**. In the case of a physically short link these functions can be carried out by using separate control lines over which handshaking signals can be exchanged. An example of such a procedure is the ITU-T V.24 protocol.

In frame-oriented links, which are normally established over longer distances, it is common to exchange separate **supervisory** frames over the same channel as the information frames. Clearly, these frames, as in the case of acknowledgements, need only be of a short length compared with the information frames. Figure 3.20 shows the signal flow diagram of a typical link setup and disconnection procedure. As can be seen, two supervisory frames are used, namely a SETUP-frame and a DISC-frame. On the transmission of the SETUP-frame, frame numbers are set to zero and send and receive windows are initialized. Note that the supervisory frames need to be acknowledged since they may be corrupted by errors. Once the link is established, information frames and acknowledgements can be exchanged. Once the data transfer has ended a DISC-frame is used to terminate the logical connection between the two nodes. This process seems almost trivial at first sight but the situation becomes more complex if a failure occurs on a link or at a node. A problem arises when frames have been accepted for transmission over a link but have not reached a

receive node before a failure occurs. Link management procedures need to be able to cope with such failures.

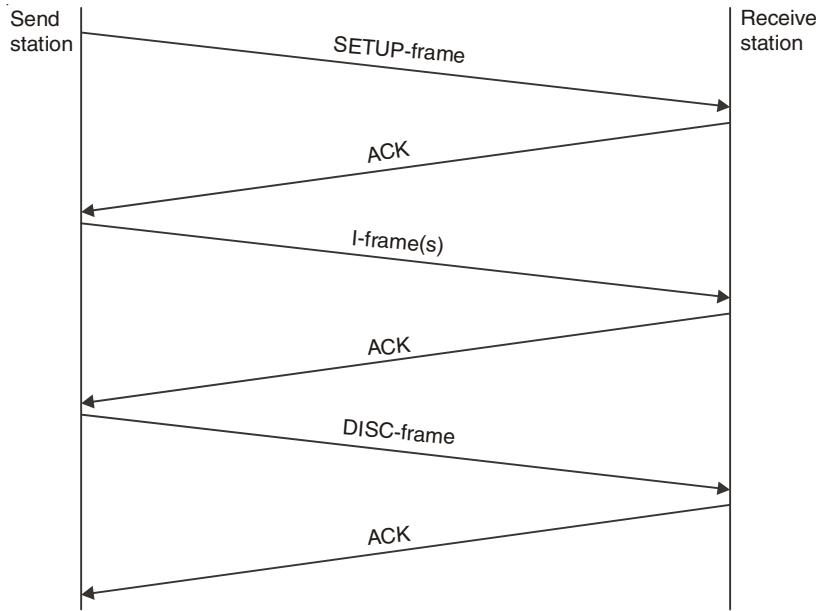


Fig. 3.20. Link set up and disconnect.

3.7 HIGH LEVEL DATA LINK CONTROL

The most important data link control protocol is HDLC (ISO 33009, ISO 4335). Not only is HDLC widely used, but it is the basis for many other important data link control protocols, which use the same or similar formats and the same mechanisms as employed in HDLC. Accordingly, in this section we provide a detailed discussion of HDLC.

3.7.1 Basic Characteristics

To satisfy a variety of applications, HDLC defines three types of stations, two link configurations, and three data-transfer modes of operation. The three station types are:

1. **Primary station.** Has the responsibility for controlling the operation of the link. Frames issued by the primary are called commands.
2. **Secondary station.** Operates under the control of the primary station. Frames issued by a secondary are called *responses*. The primary maintains a separate logical link with each secondary station on the line.
3. **Combined station.** Combines the features of primary and secondary. A combined station may issue both commands and responses.

3.7.2 The two link configurations are:

Unbalanced configuration. Consists of one primary and one or more secondary stations and supports both full-duplex and half-duplex transmission.

Balanced configuration. Consists of two combined stations and supports both full-duplex and half-duplex transmission.

3.7.3 The three data transfer modes are:

Normal response mode (NRM). Used with an unbalanced configuration. The primary may initiate data transfer to a secondary, but a secondary may only transmit data in response to a command from the primary.

Asynchronous balanced mode (ABM). Used with a balanced configuration. Either combined station may initiate transmission without receiving permission from the other combined station.

Asynchronous response mode (ARM). Used with an unbalanced configuration. The secondary may initiate transmission without explicit permission of the primary. The primary still retains responsibility for the line, including initialization, error recovery, and logical disconnection.

NRM is used on multidrop lines, in which a number of terminals are connected to a host computer. The computer polls each terminal for input. NRM is also sometimes used on point-to-point links, particularly if the link connects a terminal or other peripheral to a computer. ABM is the most widely used of the three modes; it makes more efficient use of a full-duplex point-to-point link as there is no polling overhead. ARM is rarely used; it is applicable to some special situations in which a secondary may need to initiate transmission.

3.7.4 HDLC Frame Structure

HDLC uses synchronous transmission with data being transmitted in frames. All frames have the common format shown in Fig. 3.21. The address and control fields are known collectively as a **header** and the error-checking bits are called the frame check sequence (FCS) or **trailer**.

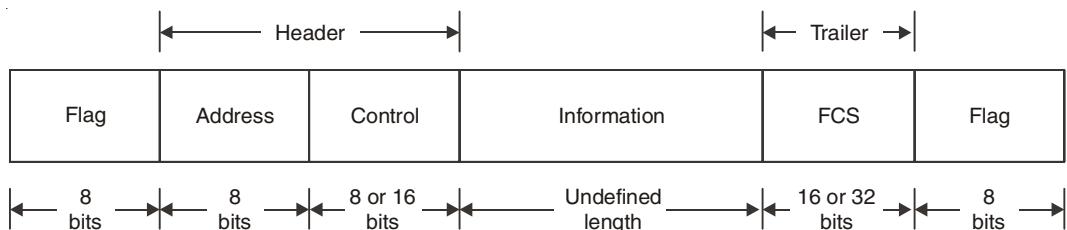


Fig. 3.21. HDLC Frame structure.

Flag fields

These two fields mark the start and finish of the frame with the bit sequence 01111110. A single flag field may be used to mark the end of one frame and the start of another if one frame follows immediately after another. Since receive nodes are continuously

looking for the flag sequence of six consecutive binary 1 bits to indicate the start or finish of a frame, it is essential that this sequence is prevented from occurring elsewhere in the frame. This is achieved by a process known as **bit stuffing** which takes place at the send node. When the frame is being assembled, if a sequence of five consecutive ones appears within a frame then an extra zero is inserted (stuffed) immediately after the sequence. At the receive node, when a sequence of five ones is received, the next bit is examined. If this bit is a zero it is removed. If the sequence is followed by the two bits 10 then the combination is accepted as a flag. If the sequence is followed by two more ones it is assumed that some form of error has occurred.

Address field

The contents of the address field depend on the mode of operation being used. In an unbalanced configuration it contains an 8-bit address which always identifies the secondary station, whether it is the primary or secondary that is transmitting. Alternatively, a group of secondary stations may have the same address, known as a **group address**, in which case a frame is transmitted from the primary station to all secondary stations in the group. The unique address containing all ones is used to allow a primary to broadcast a frame to all the secondary stations connected to it. The protocol also allows for the length of the address field to be extended in the event of the number of secondaries being too large to be addressed by an 8-bit address. In a balanced configuration the address field always contains the address of the destination. Since a balanced configuration involves a point-to-point link, the destination address is not strictly required but is included for consistency. Note that the address is not used for any routing purposes since routing is a function of the network level of the ISO model and HDLC is primarily a link-level protocol.

Control field

The control field distinguishes between the three different types of frame used in HDLC, namely information, control and unnumbered frames. The first one or two bits of the field determine the type of frame. The field also contains control information which is used for flow control and link management.

Information field

The information field does not have a length specified by HDLC. In practice, it normally has a maximum length determined by a particular implementation. Information frames (also known as I-frames) are the only frames that carry information bits which are normally in the form of a fixed-length block of data of several kilobytes in length. All other types of frame normally have an empty information field.

Frame check sequence

The FCS field contains error-checking bits, normally 16 but with a provision for increasing this to 32 in the event of systems operating in an unreliable environment or with particularly long I-frames.

3.7.5 HDLC Frame Types

The different types of frame are distinguished by the contents of the control field. The structure of all three types of control field is shown in Fig. 3.22.

Frame type	1	2	3	4	5	6	7	8	Bits
Information	0		N(S)		P		N(R)		
Supervisory	1	0	F		P		N(R)		
Unnumbered	1	1	F		P		F		

N(S) = send sequence number, N(R) = receive sequence number, F = function bits,
P = poll/final bit used for polling in normal response mode.

Fig. 3.22. Control field structure.

Information Frames

An I-frame is distinguished by the first bit of the control field being a binary 0. Note also that the control field of an I-frame contains both a send sequence number, N(S), and a receive sequence number, N(R), which are used to facilitate flow control. N(S) is the sequence number of frames sent and N(R) the sequence number of frames successfully received by the sending node prior to the present frame being sent. Thus, the first frame transmitted in a data transfer has send and receive sequence numbers 0, 0. Since 3 bits are available for each of the sequence numbers N(S) and N(R), they can have values only between 0 and 7, that is they use modulo-8 numbering. This imposes a limit on the size of the windows used for flow control. I-frames also contain a poll/final (P/F) bit (as do other frames). This acts as a poll bit when used by a primary station and a final bit by a secondary. A poll bit is set when a primary is transmitting to a secondary and requires a frame or frames to be returned in response, and the final bit is set in the final frame of a response. Since there are no primaries or secondaries in asynchronous balanced mode, the P/F bit is used differently in this mode.

Supervisory frames

Supervisory frames are distinguished by the first 2 bits of the control field being 10. These frames are used as acknowledgements for flow and error control. HDLC allows for both go-back- n and selective-repeat ARQ. Note that the supervisory frames contain only a receive sequence number since they relate to the acknowledgement of I-frames and not to their transmission. They also contain two function bits which allow for four functions as shown in Table 3.4 which lists the supervisory commands/responses.

Table 3.4. Supervisory Command and response

Name	Function
Receive Ready(RR)	Positive acknowledgement (ACK), ready to receive I-frame
Receive Not Ready (RNR)	Positive acknowledgement, not ready to receive I-frame
Reject (REJ)	Negative acknowledgement (NAK), go-back- n
Selective Reject (SREJ)	Negative acknowledgement, selective-repeat

Unnumbered frames

Unnumbered frames do not contain any sequence numbers (hence their name) and are used for various control functions. They have five function bits which allow for the fairly large number of commands and responses listed in Table 3.5.

The first five commands are used to initialize, change or terminate modes and are known as **mode-setting commands**. A node receiving such a command normally acknowledges its receipt with the UA response. A change of mode causes the I-frame sequence numbers to be reset to zero. Other responses that may result from a mode setting command are RIM, RD and DM. The UI and UP frames are used to exchange control information between nodes. The response FRMR is returned when an error occurs in a received frame. It is normally followed by a RSET command, resulting in send and receive sequence numbers being reset.

Table 3.5. Unnumbered Command/Response

Name	Function
Set Normal Response Mode (SNRM)	
Set Asynchronous Response Mode (SARM)	
Set Asynchronous Balanced Mode (SABM)	
Set Initialization Mode (SIM)	Used to initialize or change modees
Disconnect (DISC)	Used to initialize a link
Unnumbered Acknowledgement (UA)	Causes the link connection to be terminated
Request Initialization Mode (RIM)	Acknowledges the above mode-setting commands
Request Disconnect (RD)	Requests SIM command when initialization required
Disconnected Mode (DM)	Requests a disconnection of a link
Unnumbered Poll (UP)	Indicates responding station disconnected
Unnumbered Information (UI)	Used to request control information
Frame Reject (FRMR)	Used to exchange control information
Reset (SET)	Reports that unacceptable frame has been received
Test (TEST)	Resets sequence numbers
Exchange Identification (XID)	Used to exchange test signals
	Used to exchange identity and status

3.8 LINK MANAGEMENT

For data to be exchanged over an HDLC link, a connection must first be set up. Normally, this is achieved by the transfer of either an SNRM or an SABM command, depending on whether normal response or asynchronous balanced mode is being established. The receive node will respond with a UA frame if it is in a position to set up the link. A typical transfer of frames illustrating the link management aspects of an ABM link is shown in the signal flow diagram of Fig. 3.23.

In this mode, the setting up or clearing of a link may be initiated by either node. The link is set up by an SABM command and cleared by a DISC command. The UA response is used in both cases to acknowledge the successful acceptance of the commands.

Error control

This is achieved in HDLC by the use of ARQ and a cyclic error-detecting code. Prior to transmission the block of data consisting of address, control and information fields is treated as a single binary number and modulo-2 divided by a generator polynomial which is specified by the ITU-T as

$x^{16} + x^{12} + x^5 + 1$ (10001000000100001 in binary). The remainder of this division constitutes the FCS. The flags are then added and the frame transmitted. At the receive node the received frame is stripped of the flags and divided by the same generator polynomial. The remainder of this division provides a syndrome which is zero if no errors have occurred. In the event of a non-zero syndrome, an REJ frame is returned if go-back- n ARQ is being used and an SREJ frame is returned if selective-repeat ARQ is being used.

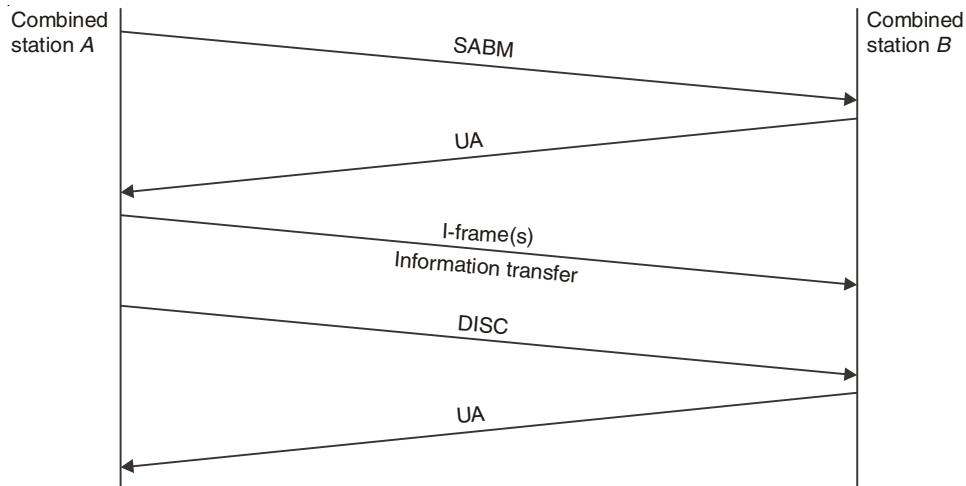


Fig. 3.23. Frame transfer in an ABM Link.

Flow control

The flow control aspects of HDLC vary slightly depending on whether NRM or ABM is being used. In NRM, data flow is carried out under the control of the primary station. A typical NRM data transfer between primary and secondary stations using a go-back- n ARQ strategy is shown in Fig. 3.24.

In this figure, data flow is from the primary to the secondary only, so that the I-frames are acknowledged by supervisory frames. Each node keeps a count of the send sequence number, $F(S)$, of the next I-frame to be sent and a count of the receive sequence number, $F(R)$, of the next I-frame to be received and it is these counts, along with the sequence numbers inside frames, that allow the flow control to function. When a node receives a frame it compares its own receive sequence number, $F(R)$, with the frame's send sequence number, $N(S)$. If these two numbers are equal then the frame is accepted; if they are not

equal the frame is rejected. The receive node then uses the value of its receive sequence number in the resulting acknowledgement. Thus I-frame (0,0) in this example is accepted because $F(R) = N(S) = 0$ and the frame is positively acknowledged by the supervisory frame RR(1) which has a receive sequence number of 1, indicating that one frame has been received correctly. Remember, as above, that supervisory frames contain only a receive sequence number whereas I-frames contain both send and receive sequence numbers. The orderly flow of I-frames from the primary is disrupted in this case by the detection at the secondary of an error in I-frame (2,0) which is immediately discarded. When I-frame (3,0) arrives at the secondary there is now a sequence mismatch between the frame, which has a send sequence number of 3, and the secondary, which has a receive count of 2 and is therefore expecting to receive a frame with a send sequence number of 2. This causes I-frame (3,0) to be rejected and a negative acknowledgement of REJ(2) to be returned indicating that the primary should retransmit I-frame (2,0). The primary will now go back and transmit this frame and the subsequent I-frame (3,0) again.

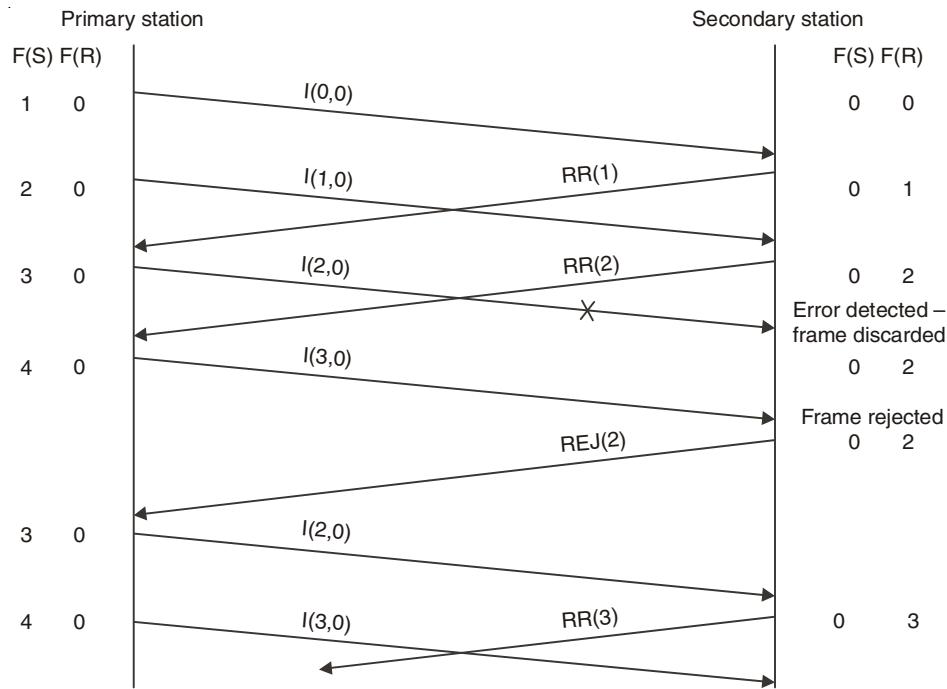


Fig. 3.24. NRM data transfer.

In the case of the ABM, both nodes can transmit I-frames independently, that is there is full-duplex transmission. Once again the best way to understand the flow control procedure is by means of a flow diagram. A typical ABM data transfer between two nodes using a go-back-n ARQ strategy is shown in Fig. 3.25. The link employs a window mechanism which, since 3 bits are allocated for a frame's send and receive sequence numbers, operates with modulo-8 numbers. This means that the window size is restricted to 7; that is, the maximum number of frames that can be transmitted without receiving an acknowledgement is 7. Since I-frames are flowing in each direction and each frame

contains both send and receive sequence numbers, these sequence numbers can be used to acknowledge the correct receipt of a frame rather than using separate acknowledgement frames as in the case of NRM operation. This type of acknowledgement process is known as piggybacking.

The flow control procedure operates in a similar way to NRM. As each frame is received, the frame send sequence number, $N(S)$, is compared with the receive count of the receive node, $F(R)$. If these two are the same then the frame is accepted; if they are not the same the frame is rejected. The frame receive sequence number, $N(R)$, is then used as an acknowledgement that the frame has been successfully received at the remote end as shown. Thus, the first frame to be received by node B , node A 's frame $I(0,0)$, is acknowledged by the next I-frame to be sent from node B which carries the sequence numbers (2,1) indicating that, at this point, node B has sent two previous frames and has successfully received one frame. Likewise, the receipt by node A of frame $I(3,2)$ from node B is an acknowledgement that frame $I(1,0)$ from node A has been correctly received at node B . This is because frame $I(3,2)$ has a receive count equal to 2 and node B has correctly received two I-frames $I(0,0)$ and $I(1,0)$). The procedure for dealing with erroneously received frames is the same as in NRM.

Frame $I(2,1)$ from node A in Fig. 3.25 is received erroneously at node B and is discarded.

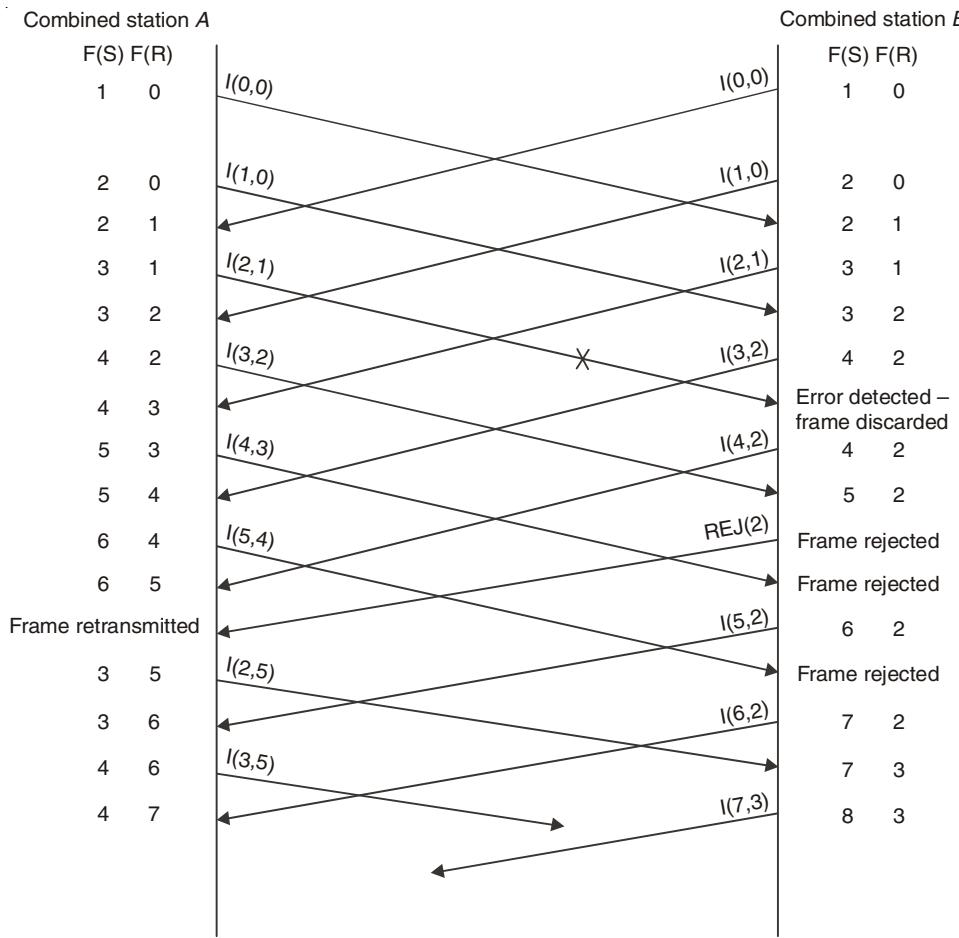


Fig. 3.25. ABM data transfer using Go-back- n ARQ.

The next frame transmitted by node *B* is frame I(4,2) indicating that only two frames have been received successfully. The next frame to arrive at node *B* (frame I(3,2) from node *A*) is rejected because its send sequence number of 3 does not match node *B*'s receive sequence number which has remained at 2, and the negative acknowledgement REJ(2) is sent by node *B*. The effect of full-duplex working in this example is such that, by the time the REJ(2) frame reaches node *A*, it has transmitted two further frames (I(4,3) and I(4,4)), both of which are rejected at node *B* because their frame send sequence numbers do not match the receive count at node *B*. Once node *A* receives the frame REJ(2) it retransmits its original frame I(2,1) which is now renumbered I(2,5) as a result of further frames having been successfully received by node *A*. It can be seen from this example that, for a go-back-*n* strategy to work successfully, there needs to be a manageable level of errors or there will be a loss of throughput due to the retransmission of frames. ABM allows for efficient full-duplex transmission on a point-to-point link and this mode of operation is incorporated in the link level of the ITU-T X.2 protocol for packet-switched networks.

3.9 POINT TO POINT PROTOCOL

Although HDLC is a general protocol that can be used for both point-to-point and multipoint configurations, one of the most common protocols for point-to-point access is the Point-to-Point Protocol (PPP). Today, millions of Internet users who need to connect their home computers to the server of an Internet service provider use PPP. The majority of these users have a traditional modem; they are connected to the Internet through a telephone line, which provides the services of the physical layer. But to control and manage the transfer of data, there is a need for a point-to-point protocol at the data link layer. PPP is by far the most common.

3.9.1 PPP Services

PPP provides several services:

1. PPP defines the format of the frame to be exchanged between devices.
2. PPP defines how two devices can negotiate the establishment of the link and the exchange of data.
3. PPP defines how network layer data are encapsulated in the data link frame.
4. PPP defines how two devices can authenticate each other.
5. PPP provides multiple network layer services supporting a variety of network layer protocols.
6. PPP provides connections over multiple links.
7. PPP provides network address configuration. This is particularly useful when a home user needs a temporary network address to connect to the Internet.

Point-to-point Protocol is defined in RFC 1661 , encapsulates IP packets inside the data field of a modified HDLC frame, as illustrated in Fig. 3.26. Note that, as with HDLC, the

frame header contains 1-byte address and control fields along with an additional 2-byte protocol field. PPP does not assign individual station addresses and the address field contains the binary sequence 11111111.

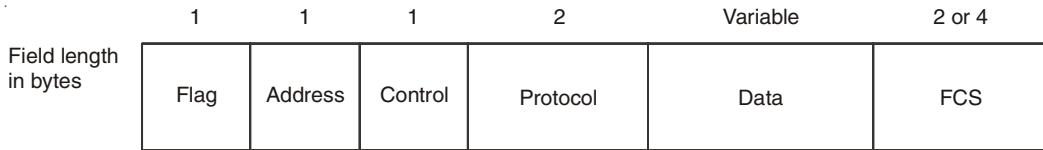


Fig. 3.26. Frame structure of point-to-point protocol.

1. **Flag.** A PPP frame starts and ends with a 1-byte flag with the bit pattern 01111110. Although this pattern is the same as that used in HDLC, there is a big difference. PPP is a byte-oriented protocol; HDLC is a bit-oriented protocol.
2. **Address.** The address field in this protocol is a constant value and set to 11111111 (broadcast address). During negotiation, the two parties may agree to omit this byte.
3. **Control.** This field is set to the constant value 11000000 (imitating unnumbered frames in HDLC). As we will discuss later, PPP does not provide any flow control. Error control is also limited to error detection. This means that this field is not needed at all, and again, the two parties can agree, during negotiation, to omit this byte.
4. **Protocol.** The protocol field defines what is being carried in the data field: either user data or other information. This field is by default 2 bytes long, but the two parties can agree to use only 1 byte.
5. **Data/Payload field.** This field carries either the user data or other information. The data field is a sequence of bytes with the default of a maximum of 1500 bytes; but this can be changed during negotiation. The data field is byte stuffed if the flag byte pattern appears in this field. Because there is no field defining the size of the data field, padding is needed if the size is less than the maximum default value or the maximum negotiated value.
6. **FCS.** The frame check sequence (FCS) is simply a 2-byte or 4-byte standard CRC.

In addition to IP, PPP supports other protocols, including Novell's Inter network Packet Exchange (IPX) and IBM's Synchronous Network Architecture (SNA). The FCS performs an identical function to that in an HDLC frame.

PPP uses a **Link Control Protocol** (LCP) to establish, configure and test the data link connection that goes through four distinct phases: Firstly, link establishment and configuration negotiation occur. Before any network layer packets (*e.g.*, IP) can be exchanged, LCP first must open the connection and negotiate configuration parameters. This phase is complete when a configuration-acknowledgement frame has been both sent and received. This is followed by an optional link-quality determination phase. In this phase, the link is tested to determine whether the link quality is sufficient to support the network layer protocols. Transmission of network layer protocol information is delayed until this phase is complete. At this point, a network layer protocol configuration negotiation occurs. PPP

is designed to allow the simultaneous use of multiple network layer protocols and network layer protocols can be configured separately and can be brought into use and taken down at any time. If the LCP closes the link, it informs the network layer protocols so that they can take appropriate action. Finally, link termination can occur. This is usually carried out at the request of a user but can happen because of a physical event, such as the loss of line signals or the expiration of an idle-period timer.

Three classes of LCP frames exist. Link-establishment frames are used to establish and configure a link; link-termination frames are used to terminate a link; and link maintenance frames are used to manage and debug a link.

Review Questions

1. Briefly describe the services provided by the data link layer.
2. Define framing and the reason for its need.
3. Compare and contrast byte-oriented and bit-oriented protocols. Which category has been popular in the past (explain the reason)? Which category is popular now (explain the reason)?
4. Compare and contrast byte-stuffing and bit-stuffing. Which technique is used in byte-oriented protocols? Which technique is used in bit-oriented protocols?
5. Compare and contrast flow control and error control.
6. What are the two protocols we discussed for noiseless channels in this chapter?
7. What are the three protocols we discussed for noisy channels in this chapter?
8. Explain the reason for moving from the Stop-and-Wait ARQ Protocol to the Go Back-NARQ Protocol.
9. Compare and contrast the Go-Back-NARQ Protocol with Selective-Repeat ARQ.
10. Compare and contrast HDLC with PPP. Which one is byte-oriented; which one is bit-oriented?
11. Define piggybacking and its usefulness.
12. A long-distance half-duplex RF link uses go-back-3 ARQ and has the following characteristics:
 - Data transmission rate: 128 kbps
 - Frame size: 256 bytes
 - Information bytes per frame: 251
 - Propagation delay: 20 ms
 - Processing delay: 4 msDetermine:
 - (a) the efficiency and throughput in the absence of errors,
 - (b) the throughput in the presence of a bit error rate of 0.00004.



CHAPTER **4**

THE MEDIUM ACCESS CONTROL SUBLAYER

Nature seems ... to reach many of her ends by long circuitous routes.

—Rudolph Lotze

4.1 INTRODUCTION

As we pointed out in Chapter 1, networks can be divided into two categories. Those using point-to-point connections and those using broadcast channels. This chapter deals with broadcast networks and their protocols.

In any broadcast network, the key issue is how to determine who gets to use the channel when there is competition for it. Many protocols for solving the problem are known and form the contents of this chapter. In the literature, broadcast channels are sometimes referred to as **multiaccess channels** or **random access channels**.

The protocols used to determine who goes next on a multiaccess channel belong to a sublayer of the data link layer called the **MAC (Medium Access Control)** sublayer. The MAC sublayer is especially important in LANs, many of which use a multiaccess channel as the basis for communication. WANs, in contrast, use point-to-point links, except for satellite networks. Because multiaccess channels and LANs are so closely related, in this chapter we will discuss LANs in general, including a few issues that are not strictly part of the MAC sublayer.

Technically, the MAC sublayer is the bottom part of the data link layer, so logically we should have studied it before examining all the point-to-point protocols in Chapter 3. Nevertheless, for most people, understanding protocols involving multiple parties is easier after two-party protocols are well understood. For that reason we have deviated slightly from a strict bottom-up order of presentation.

Protocols defined specifically for LAN and MAN transmission address issues relating to the transmission of blocks of data over the network. In OSI terms, higher layer protocols (layer 3 or 4 and above) are independent of network architecture and are applicable to LANs, MANs, and WANs. Thus, a discussion of LAN protocols is concerned principally with lower layers of the OSI model.

Figure 4.1 relates the LAN protocols to the OSI architecture. This architecture was developed by the IEEE 802 committee and has been adopted by all organizations working on the specification of LAN standards. It is generally referred to as the IEEE 802 reference model. Working from the bottom up, the lowest layer of the IEEE 802 reference model corresponds to the *physical layer* of the OSI model, and includes such functions as encoding and decoding of signals, Preamble generation and removal (for synchronization), Bit transmission and reception.

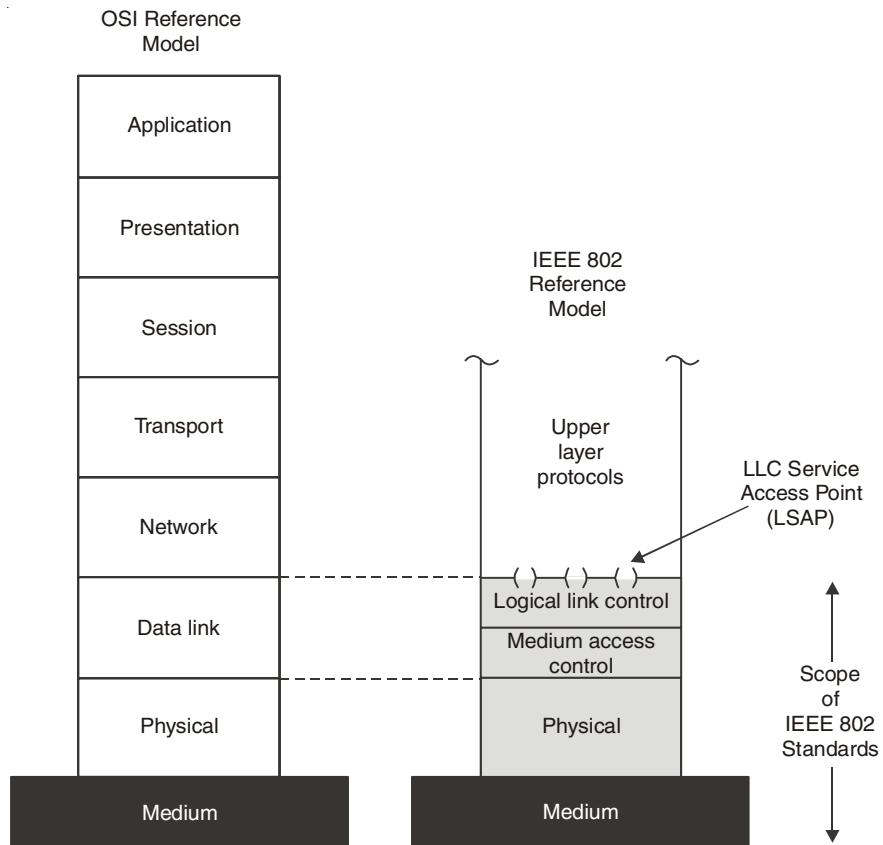


Fig. 4.1. Comparison of IEEE 802 and OSI Model.

In addition, the physical layer of the 802 model includes a specification of the transmission medium and the topology. Generally, this is considered *below* the lowest layer of the OSI model. However, the choice of transmission medium and topology is critical in LAN design, and so a specification of the medium is included. Above the physical layer are the functions associated with providing service to LAN users. These include:

1. On transmission, assemble data into a frame with address and error detection fields.
2. On reception, disassemble frame, perform address recognition and error detection.
3. Govern access to the LAN transmission medium.
4. Provide an interface to higher layers and perform flow and error control.

These are functions typically associated with OSI layer 2. The set of functions in the last bulleted item are grouped into a *logical link control* (LLC) layer. These functions are treated as a separate layer, called *medium access control* (MAC). The separation is done for the following reasons:

1. The logic required to manage access to a shared-access medium is not found in traditional layer-2 data link control.
2. For the same LLC, several MAC options may be provided.

4.2 MULTIPLE ACCESS PROTOCOLS

When nodes or stations are connected and use a common link, called a multipoint or broadcast link, we need a multiple access protocol to coordinate access to the link. The problem of controlling the access to the medium is similar to the rules of speaking in an assembly. The procedures guarantee that the right to speak is upheld and ensure that two people do not speak at the same time, do not interrupt each other, do not monopolize the discussion, and so on.

The situation is similar for multipoint networks. Many algorithms for allocating a multiple access channel are known. We categorize them into three groups.

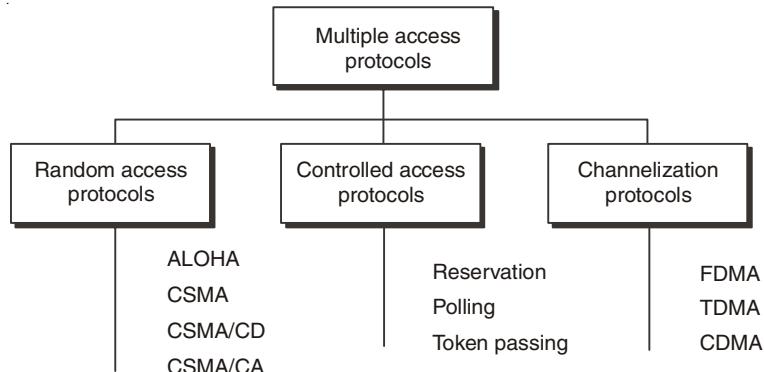


Fig. 4.2. Multiple Access Protocols

4.2.1 Random Access

In random access or contention methods, no station is superior to another station and none is assigned the control over another. No station permits, or does not permit, another station to send. At each instance, a station that has data to send uses a procedure defined by the protocol to make a decision on whether or not to send. This decision depends on the state

of the medium (idle or busy). In other words, each station can transmit when it desires on the condition that it follows the predefined procedure, including the testing of the state of the medium.

Two features give this method its name. First, there is no scheduled time for a station to transmit. Transmission is random among the stations. That is why these methods are called *random access*. Second, no rules specify which station should send next. Stations compete with one another to access the medium. That is why these methods are also called *contention* methods.

In a random access method, each station has the right to the medium without being controlled by any other station. However, if more than one station tries to send, there is an access conflict-collision and the frames will be either destroyed or modified. To avoid access conflict or to resolve it when it happens, each station follows a procedure that answers the following questions:

- When can the station access the medium?
- What can the station do if the medium is busy?
- How can the station determine the success or failure of the transmission?
- What can the station do if there is an access conflict?

The random access methods we study in this section have evolved from a very interesting protocol known as ALOHA, which used a very simple procedure called multiple access (MA). The method was improved with the addition of a procedure that forces the station to sense the medium before transmitting. This was called carrier sense multiple access. This method later evolved into two parallel methods: carrier sense multiple access with collision detection (CSMA/CD) and carrier sense multiple access with collision avoidance (CSMA/CA). CSMA/CD tells the station what to do when a collision is detected. CSMA/CA tries to avoid the collision.

ALOHA

In the 1970s, Norman Abramson and his colleagues at the University of Hawaii devised a new and elegant method to solve the channel allocation problem. Their work has been extended by many researchers since then (Abramson, 1985). Although Abramson's work, called the ALOHA system, used ground-based radio broadcasting, the basic idea is applicable to any system in which uncoordinated users are competing for the use of a single shared channel.

Whenever a station has a frame to send, it does so. The station then listens for an amount of time equal to the maximum possible round-trip propagation delay on the network (twice the time it takes to send a frame between the two most widely separated stations) plus a small fixed time increment. If the station hears an acknowledgement during that time, fine; otherwise, it resends the frame. If the station fails to receive an acknowledgement after repeated transmissions, it gives up. A receiving station determines the correctness of an incoming frame by examining a framecheck—sequence field, as in HDLC. If the frame is valid and if the destination address in the frame header matches the receiver's address, the station immediately sends an acknowledgement. The frame may be invalid due to noise on

the channel or because another station transmitted a frame at about the same time. In the latter case, the two frames may interfere with each other at the receiver so that neither gets through; this is known as a *collision*. If a received frame is determined to be invalid, the receiving station simply ignores the frame.

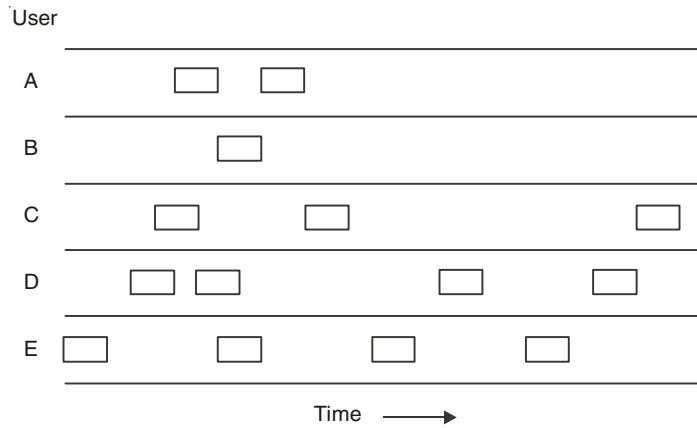


Fig. 4.3. In pure ALOHA, frames are transmitted at completely arbitrary times.

A sketch of frame generation in an ALOHA system is given in Fig. 4.3. We have made the frames all the same length because the throughput of ALOHA systems is maximized by having a uniform frame size rather than by allowing variable length frames. Whenever two frames try to occupy the channel at the same time, there will be a collision and both will be garbled. If the first bit of a new frame overlaps with just the last bit of a frame almost finished, both frames will be totally destroyed and both will have to be retransmitted later. The checksum cannot (and should not) distinguish between a total loss and a near miss. Bad is bad.

ALOHA is as simple as can be, and pays a penalty for it. Because the number of collisions rise rapidly with increased load, the maximum utilization of the channel is only about 18%.

To improve efficiency, a modification of ALOHA, known as slotted ALOHA, was developed. In this scheme, time on the channel is organized into uniform slots whose size equals the frame transmission time. Some central clock or other technique is needed to synchronize all stations. Transmission is permitted to begin only at a slot boundary. Thus, frames that do overlap will do so totally. This increases the maximum utilization of the system to about 37%.

Both ALOHA and slotted ALOHA exhibit poor utilization. Both fail to take advantage of one of the key properties of both packet radio and LANs, which is that propagation delay between stations is usually very small compared to frame transmission time. Consider the following observations. If the station-to-station propagation time is large compared to the frame transmission time, then, after a station launches a frame, it will be a long time before other stations know about it. During that time, one of the other stations may transmit a frame; the two frames may interfere with each other and neither gets through. Indeed, if the distances are great enough, many stations may begin transmitting, one after the other, and

none of their frames get through unscathed. Suppose, however, that the propagation time is small compared to frame transmission time. In that case, when a station launches a frame, all the other stations know it almost immediately. So, if they had any sense, they would not try transmitting until the first station was done. Collisions would be rare because they would occur only when two stations began to transmit almost simultaneously. Another way to look at it is that a short delay time provides the stations with better feedback about the state of the network; this information can be used to improve efficiency. The foregoing observations led to the development of carrier sense multiple access (CSMA).

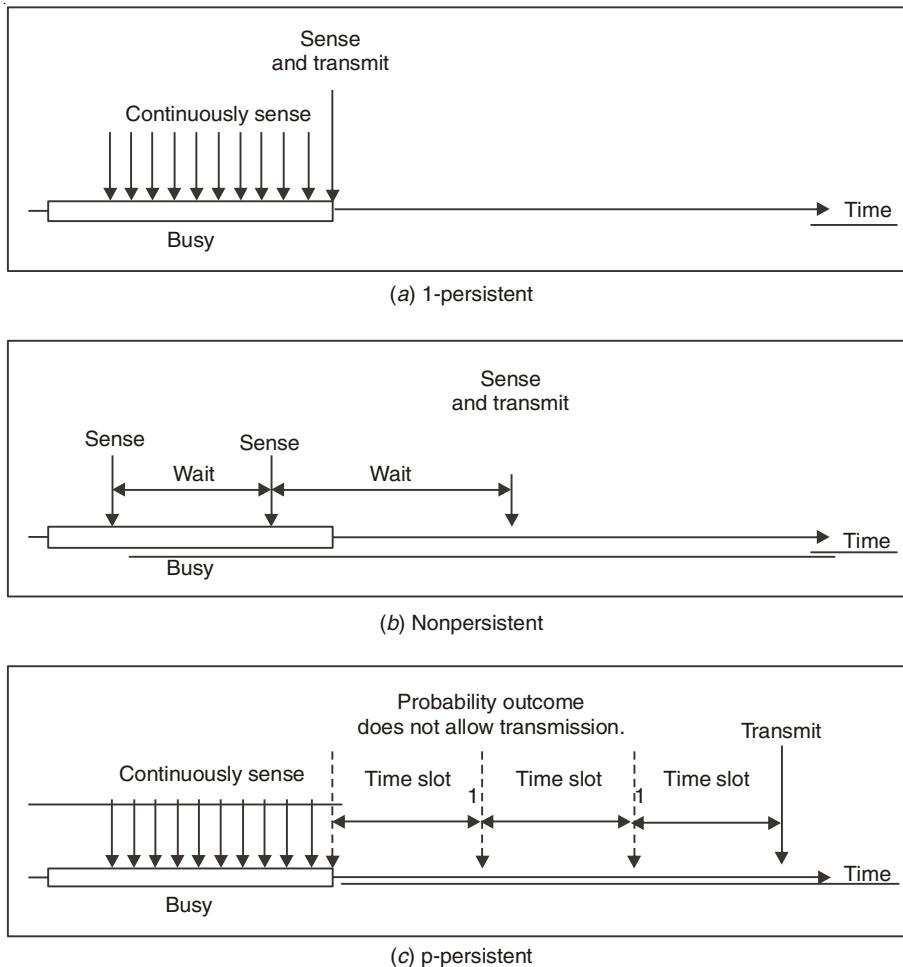
Carrier Sense Multiple Access Protocols

Protocols in which stations listen for a carrier (*i.e.*, a transmission) and act accordingly are called **carrier sense protocols**. A number of them have been proposed. Kleinrock and Tobagi (1975) have analyzed several such protocols in detail. Below we will mention several versions of the carrier sense protocols.

- (a) **1-persistent CSMA** (Carrier Sense Multiple Access): When a station has data to send, it first listens to the channel to see if anyone else is transmitting at that moment. If the channel is busy, the station waits until it becomes idle. When the station detects an idle channel, it transmits a frame. If a collision occurs, the station waits a random amount of time and starts all over again. The protocol is called 1-persistent because the station transmits with a probability of 1 when it finds the channel idle.

The propagation delay has an important effect on the performance of the protocol. There is a small chance that just after a station begins sending, another station will become ready to send and sense the channel. If the first station's signal has not yet reached the second one, the latter will sense an idle channel and will also begin sending, resulting in a collision. The longer the propagation delay, the more important this effect becomes, and the worse the performance of the protocol. Even if the propagation delay is zero, there will still be collisions. If two stations become ready in the middle of a third station's transmission, both will wait politely until the transmission ends and then both will begin transmitting exactly simultaneously, resulting in a collision. If they were not so impatient, there would be fewer collisions. Even so, this protocol is far better than pure ALOHA because both stations have the decency to desist from interfering with the third station's frame. Intuitively, this approach will lead to a higher performance than pure ALOHA. Exactly the same holds for slotted ALOHA.

- (b) **Nonpersistent CSMA**: In this protocol, a conscious attempt is made to be less greedy than in the previous one. Before sending, a station senses the channel. If no one else is sending, the station begins doing so itself. However, if the channel is already in use, the station does not continually sense it for the purpose of seizing it immediately upon detecting the end of the previous transmission. Instead, it waits a random period of time and then repeats the algorithm. Consequently, this algorithm leads to better channel utilization but longer delays than 1-persistent CSMA.

**Fig. 4.4. Carrier Sense Multiple Access.**

- (c) **p-persistent CSMA:** It applies to slotted channels and works as follows. When a station becomes ready to send, it senses the channel. If it is idle, it transmits with a probability p . With a probability $q = 1 - p$, it defers until the next slot. If that slot is also idle, it either transmits or defers again, with probabilities p and q . This process is repeated until either the frame has been transmitted or another station has begun transmitting. In the latter case, the unlucky station acts as if there had been a collision (*i.e.*, it waits a random time and starts again). If the station initially senses the channel busy, it waits until the next slot and applies the above algorithm. Figure 4.5 shows the computed throughput versus offered traffic for all three protocols, as well as for pure and slotted ALOHA. Throughput is maximum number of successful transmissions over a given period of time.

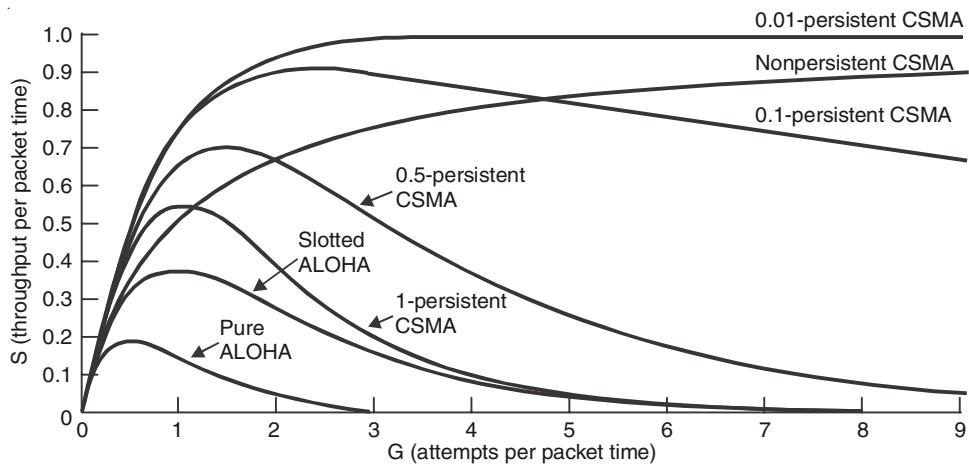


Fig. 4.5. Comparison of channel utilization vs Load for Random access methods.

Carrier-Sense Multiple Access with Collision Detection (CSMA/CD)

Carrier-Sense Multiple Access with Collision Detection (CSMA/CD) can be categorized as a listen then send access method. CSMA/CD is one of the earliest developed access techniques, and it is the technique used in Ethernet. Under the CSMA/CD concept, when a station has data to send, it first listens to determine if any other station on the network is talking. The fact that the channel is idle is determined in one of two ways, based on whether the network is broadband or baseband.

In a CSMA/CD network, if the channel is busy, the station will wait until it becomes idle before transmitting data. Since it is possible for two stations to listen at the same time and discover an idle channel, it is also possible that the two stations could then transmit at the same time. When this situation arises, a collision will occur. On sensing that a collision has occurred, a delay scheme will be employed to prevent a repetition of the collision. Typically, each station will use either a randomly generated or predefined time-out period before attempting to retransmit the message that collided. Since this access method requires hardware capable of detecting the occurrence of a collision, additional circuitry required to perform collision detection adds to the cost of such hardware.

Figure 4.6 illustrates a CSMA/CD bus-based local area network. Each workstation is attached to the transmission medium, such as coaxial cable, by a device known as a bus interface unit (BIU). To obtain an overview of the operation of a CSMA/CD network, assume that station A is currently using the channel and that stations C and D wish to transmit. The BIUs connecting stations C and D to the network would listen to the channel and note that it is busy. Once station A has completed its transmission, stations C and D attempt to gain access to the channel. Because station A's signal takes longer to propagate down the cable to station D than to station C, C's BIU notices that the channel is free slightly before station D's BIU. However, as station C gets ready to transmit, station D now assumes that the channel is free. Within an infinitesimal period of time, C starts transmission, followed by D, resulting in a collision. Here, the collision is a function of the propagation delay of the signal

and the distance between two competing stations. CSMA/CD networks therefore work better as the main cable length decreases.

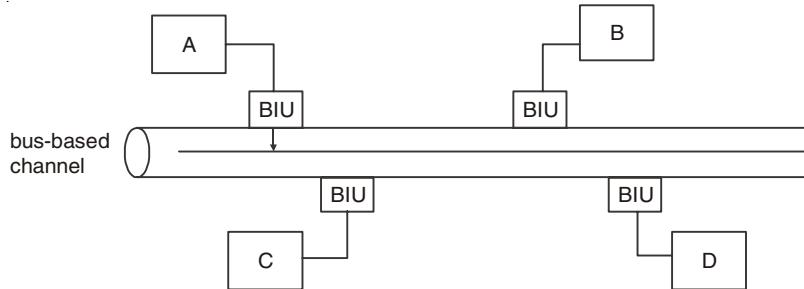


Fig. 4.6. CSMA/CD network operation. In a CSMA/CD network, as the distance between workstations increases, the resulting increase in propagation delay time increases the probability of collisions.

The CSMA/CD access technique is best suited for networks with intermittent transmission, since an increase in traffic volume causes a corresponding increase in the probability of the cable being occupied when a station wishes to talk. In addition, as traffic volume builds under CSMA/CD, throughput may decline, because there will be longer waits to gain access to the network, as well as additional time-outs required to resolve collisions that occur. In spite of those deficiencies it is relatively easy to implement the CSMA/CD access protocol, resulting in Ethernet networks currently accounting for well over 85% of all LANs.

4.2.2 Controlled Access

In controlled access, the stations consult one another to find which station has the right to send. A station cannot send unless it has been authorized by other stations. We discuss two popular controlled access methods.

Reservation

In the reservation method, a station needs to make a reservation before sending data. Time is divided into intervals. In each interval, a reservation frame precedes the data frames sent in that interval. If there are N stations in the system, there are exactly N reservation minislots in the reservation frame.

Each minislot belongs to a station. When a station needs to send a data frame, it makes a reservation in its own minislot. The stations that have made reservations can send their data frames after the reservation frame.

Figure 4.7 shows a situation with five stations and a five-minislot reservation frame. In the first interval, only stations 1, 3 and 4 have made reservations. In the second interval, only station 1 has made a reservation.

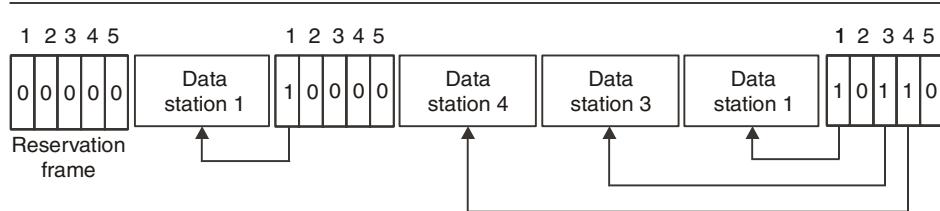


Fig. 4.7. Reservation Access methods.

Token passing

In a token passing access method, each time the network is turned on, a token is generated. The token, consisting of a unique bit pattern, travels the length of the network, either around a ring or along the length of a bus. When a station on the network has data to transmit, it must first seize a free token. On a Token-Ring network, the token is then transformed to indicate that it is in use. Information is added to produce a frame, which represents data being transmitted from one station to another. During the time that the token is in use, other stations on the network remain idle, eliminating the possibility of collisions. Once the transmission is completed, the token is converted back into its original form by the station that transmitted the frame, and becomes available for use by the next station on the network.

Figure 4.8 illustrates the general operation of a token passing in a Token-Ring network using a ring topology. Since a station on the network can only transmit when it has a free token, token passing eliminates the requirement for collision detection hardware. Due to the dependence of the network on the token, the loss of a station can bring the entire network down. To avoid this, the design characteristics of Token-Ring networks include circuitry that automatically removes a failed or failing station from the network as well as other self-healing features. This additional capability is costly; a Token-Ring adapter card in 1999 was typically priced at two to three times the cost of an Ethernet adapter card.

4.2.3 Channelization

Channelization is a multiple-access method in which the available bandwidth of a link is shared in time, frequency, or through code, between different stations.

In frequency division multiple access (FDMA), the available bandwidth is divided into frequency bands. Each station is allocated a band to send its data. In other words, each band is reserved for a specific station, and it belongs to the station all the time. Each station also uses a bandpass filter to confine the transmitter frequencies. To prevent station interferences, the allocated bands are separated from one another by small *guard bands*. FDMA specifies a predetermined frequency band for the entire period of communication. This means that stream data (a continuous flow of data that may not be packetized) can easily be used with FDMA.

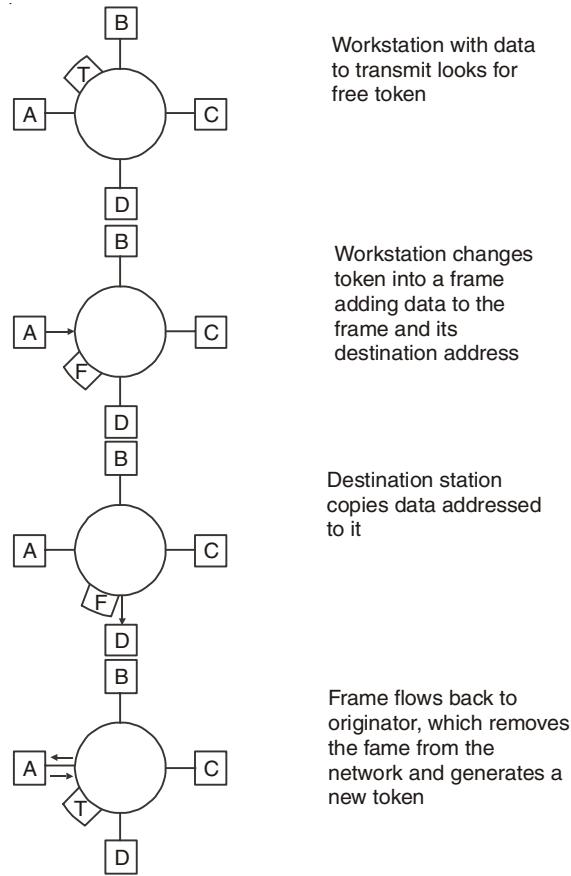


Fig. 4.8. Token-Ring operation.

In time division multiple access (TDMA), the stations share the bandwidth of the channel in time. Each station is allocated a time slot during which it can send data. Each station transmits its data in its assigned time slot. The main problem with TDMA lies in achieving synchronization between the different stations. Each station needs to know the beginning of its slot and the location of its slot. This may be difficult because of propagation delays introduced in the system if the stations are spread over a large area. To compensate for the delays, we can insert *guard time*.

In Code Division Multiple Access, channel carries all transmissions simultaneously. CDMA differs from FDMA because only one channel occupies the entire bandwidth of the link. It differs from TDMA because all stations can send data simultaneously; there is no timesharing. In nutshell, CDMA is based on coding theory, where each station is assigned a unique code, known only to intended receiver. CDMA, thus in a way adds to system security.

4.3 ETHERNET

The Ethernet is easily the most successful local area networking technology of the last 20 years. Developed in the mid-1970s by researchers at the Xerox Palo Alto Research Center (PARC), the Ethernet is a working example of the more general Carrier Sense Multiple Access with Collision Detect (CSMA/CD) local area network technology.

As indicated by the CSMA name, the Ethernet is a multiple-access network, meaning that a set of nodes send and receive frames over a shared link. You can, therefore, think of an Ethernet as being like a bus that has multiple stations plugged into it. The “carrier sense” in CSMA/CD means that all the nodes can distinguish between an idle and a busy link, and “collision detect” means that a node listens as it transmits and can therefore detect when a frame it is transmitting has interfered (collided) with a frame transmitted by another node.

Digital Equipment Corporation and Intel Corporation joined Xerox to define a 10 Mbps Ethernet standard in 1978. This standard then formed the basis for IEEE standard 802.3. With one exception that we will see in Section 2.6.2, it is fair to view the 1978 Ethernet standard as a proper subset of the 802.3 standard; 802.3 additionally defines a much wider collection of physical media over which Ethernet can operate, and more recently, it has been extended to include a 100 Mbps version called Fast Ethernet and a 1000 Mbps version called Gigabit Ethernet. The 10 Mbps Ethernet is typically used in multiple-access mode and we are interested in how multiple hosts share a single link. Both 100 Mbps and 1000 Mbps Ethernets are designed to be used in full-duplex, point-to-point configurations, which means that they are typically used in switched networks. This section will focus on the different types of Ethernet networks by closely examining the components and operating characteristics of Ethernet and then comparing its major features to the different networks defined by the IEEE 802.3 standard. Once this has been accomplished, we will focus our attention on the wiring, topology, and hardware components associated with each type of IEEE 802.3 Ethernet network, as well as the Ethernet frame used to transport data.

4.3.1 Original Network Components

The 10 Mbps bus-based Ethernet network standard originally developed by Xerox, Digital Equipment Corporation and Intel was based on the use of five hardware components. Those components include a coaxial cable, a cable tap, a transceiver, a transceiver cable, and an interface board (also known as an Ethernet controller). Figure 4.9 illustrates the relationship among the original 10 Mbps bus-based Ethernet network components.

Coaxial cable

One of the problems faced by the designers of Ethernet was the selection of an appropriate medium. Although twisted-pair wire is relatively inexpensive and easy to use, the short distances between twists serve as an antennae for receiving electromagnetic and radio frequency interference in the form of noise. Thus, the use of twisted-pair cable restricts the network to relatively short distances. Coaxial cable, however, has a dielectric shielding the conductor. As long as the ends of the cable are terminated, coaxial cable can transmit over greater distances than twisted-pair cable. Thus, the initial selection for Ethernet transmission medium was coaxial cable.

There are two types of coaxial cable that can be used to form the main Ethernet bus. The first type of coaxial cable specified for Ethernet was a relatively thick $50\ \Omega$ cable, which is normally colored yellow and is commonly referred to as thick Ethernet. This cable has a marking every 2.5 m to indicate where a tap should occur, if one is required to connect a station to the main cable at a particular location. These markings represent the minimum distance by which one tap must be separated from another on an Ethernet network. A second type of coaxial cable used with Ethernet is smaller and more flexible; however, it is capable of providing a transmission distance only one third of that obtainable on thick cable. This lighter and more flexible cable is referred to as thin Ethernet and also has an impedance of $50\ \Omega$.

Two of the major advantages of thin Ethernet over thick cable are its cost and its use of BNC connectors. Thin Ethernet is significantly less expensive than thick Ethernet. Thick Ethernet requires connections via taps, whereas the use of thin Ethernet permits connections to the bus via industry standard BNC connectors that form T-junctions.

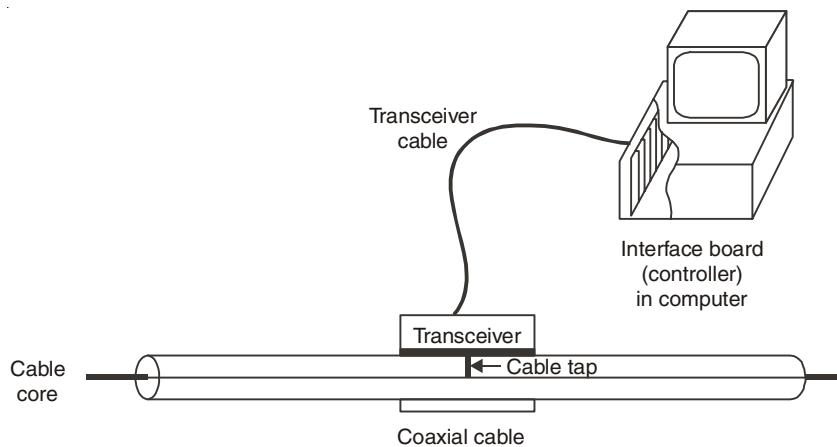


Fig. 4.9. Initial Bus based Ethernet Hardware components.

Transceiver and transceiver cable

Transceiver is a shortened form of transmitter receiver. This device contains electronics to both transmit onto and receive signals carried by the coaxial cable. The transceiver contains a tap that, when pushed against the coaxial cable, penetrates the cable and makes contact with the core of the cable. In books and technical literature the transceiver, its tap, and its housing are often referred to as the medium attachment unit (MAU).

The transceiver is responsible for carrier detection and collision detection. When a collision is detected during a transmission, the transceiver places a special signal, known as a jam on the cable. This signal is of sufficient duration to propagate down the network bus and inform all the other transceivers attached to the bus that a collision occurred. The cable that connects the interface board to the transceiver is known as the transceiver cable. This cable can be up to 50 m (165 feet) in length, and it contains five individually shielded twisted pairs. Two pairs are used for data in and data out, and two pairs are used for control signals in and out. The remaining pair, which is not always used, permits the power from the

computer in which the interface board is inserted to power the transceiver. Since collision detection is a critical part of the CSMA/CD access protocol, the original version of Ethernet was modified to inform the interface board that the transceiver collision circuitry is operational. This modification resulted in each transceiver's sending a signal to the attached interface board after every transmission, informing the board that the transceiver's collision circuitry is operational. This signal is sent by the transceiver over the collision pair of the transceiver cable, and must start within 0.6 ms after each frame is transmitted. The duration of the signal can vary between 0.5 and 1.5 ms. Known as the Signal Quality Error and also referred to as the SQE or heartbeat, this signal is supported by Ethernet Version 2.0, published as a standard in 1982, and by the IEEE 802.3 standard.

Interface board

The interface board, or network interface card (NIC), is inserted into an expansion slot within a computer, and is responsible for transmitting frames to and receiving frames from the transceiver. This board contains several special chips, including a controller chip that assembles data into an Ethernet frame and computes the cyclic redundancy check used for error detection. Thus, this board is also referred to as an Ethernet controller.

Repeaters

A repeater is a device that receives, amplifies and retransmits signals. Since a repeater operates at the physical layer, it is transparent to data and simply regenerates signals. Figure 4.10 illustrates the use of a repeater to connect two Ethernet cable segments. As indicated, a transceiver is taped to each cable segment to be connected, and the repeater is cabled to the transceiver. When used to connect cable segments, a repeater counts as one station on each connected segment. Thus, a segment capable of supporting up to 100 stations can support only 99 additional stations when a repeater is used to connect cable segments.

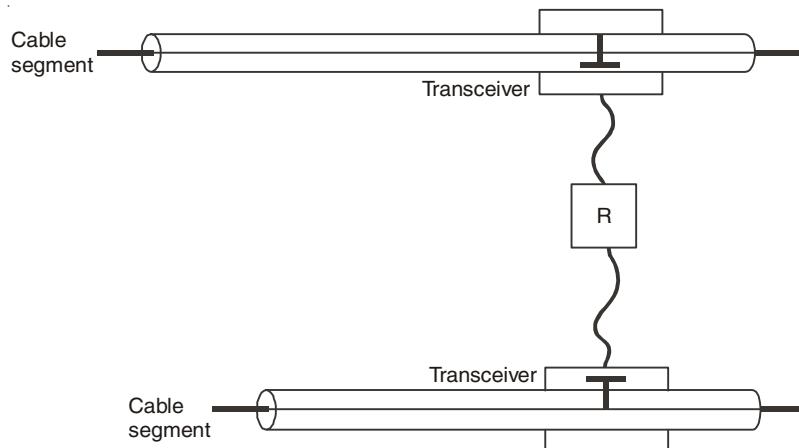


Fig. 4.10. Use of Repeaters for Network expansion.

4.3.2 IEEE 802.3 Networks

The IEEE 802.3 standard is based on Ethernet. However, it has several significant differences, particularly its support of multiple Physical layer options, which include 50 and 75 Ω coaxial

cable, unshielded twisted-pair wire, an operating rate of 100-Mbps for two standardized versions of Ethernet, and support for Gigabit Ethernet over several types of fiber optic and copper media. Other differences between various types of IEEE 802.3 networks and Ethernet include the data rates supported by some 802.3 networks, their method of signaling, the maximum cable segment lengths permitted prior to the use of repeaters, and their network topologies.

Network names

The standards that define IEEE 802.3 networks have been given names that generally follow the form ‘s-type 1’. Here, s refers to the speed of the network in Mbps, type is BASE for base band and BROAD for broadband, and 1 refers to the maximum segment length in 100 m multiples. Thus, 10BASE-5 refers to an IEEE 802.3 baseband network that operates at 10-Mbps and has a maximum segment length of 500 m. One exception to this general form is 10BASE-T, which is the name for an IEEE 802.3 network that operates at 10-Mbps using unshielded twisted pair (UTP) wire.

10BASE-5

An examination of the operating characteristics of Ethernet and 10BASE-5 indicates that these networks are the same. Figure 4.11 illustrates the major terminology changes between Ethernet and the IEEE 802.3 10BASE-5 network. These changes are in the media interface: the transceiver cable is referred to as the Attachment Unit Interface (AUI), and the transceiver, including its tap and housing, is referred to as the Medium Attachment Unit (MAU). The Ethernet controller, also known as an interface board, is now known as the Network Interface Card (NIC).

Both Ethernet and the IEEE 802.3 10BASE-5 standards support a data rate of 10-Mbps and a maximum cable segment length of 500 m. 10BASE-5, like Ethernet, requires a minimum spacing of 2.5 m between MAUs and supports a maximum of five segments in any end-to-end path through the traversal of up to four repeaters in any path. Within any path, no more than three cable segments can be populated (have stations attached to the cable) and the maximum number of attachments per segment is limited to 100.

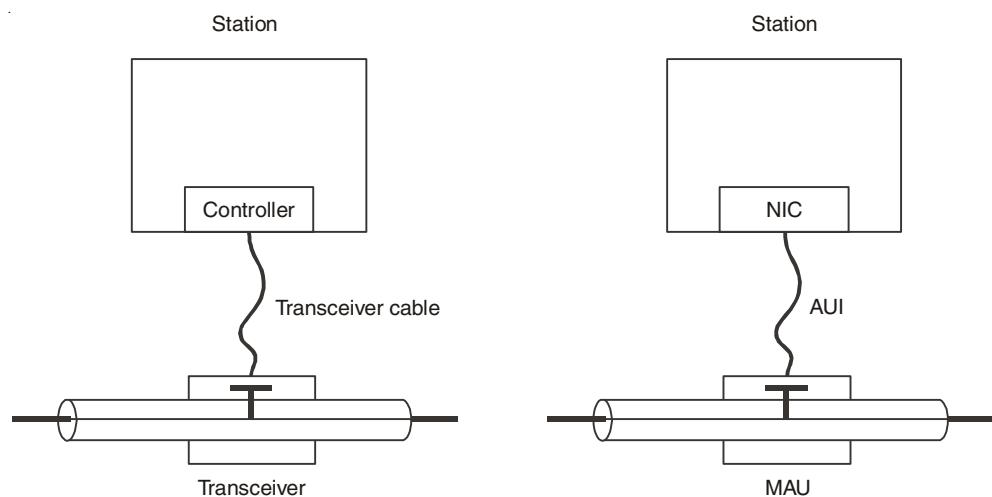


Fig. 4.11. Ethernet and 10Base-5 Terminology differences.

10BASE-2

10BASE-2 is a smaller and less expensive version of 10BASE-5. This standard uses a thinner RG-58 coaxial cable, thus earning the names 'cheapnet' and 'thinnet', as well as 'thin Ethernet'. Although 10BASE-2 cable is both less expensive and easier to use than 10BASE-5 cable, it cannot carry signals as far as 10BASE-5 cable.

Under the 10BASE-2 standard, the maximum cable segment length is reduced to 185 m (607 feet), with a maximum of 30 stations per segment. Another difference between 10BASE-5 and 10BASE-2 concerns the integration of transceiver electronics into the network interface card under the 10BASE-2 standard. This permits the NIC to be directly cabled to the main trunk cable. In fact, under 10BASE-2 the thin Ethernet cable is routed directly to each workstation location and routed through a BNC T-connector, one end of which is pressed into the BNC connector built into the rear of the network interface card.

Figure 4.12 illustrates the cabling of a one-segment 10BASE-2 network, which can support a maximum of 30 nodes or stations. BNC barrel connectors can be used to join two lengths of thin 10BASE-2 cable to form a cable segment, as long as the joined cable does not exceed 185 m in length. A BNC terminator must be attached to each end of each 10BASE-2 cable segment. One of the two terminators on each segment contains a ground wire that should be connected to a ground source, such as the screw on an electrical outlet.

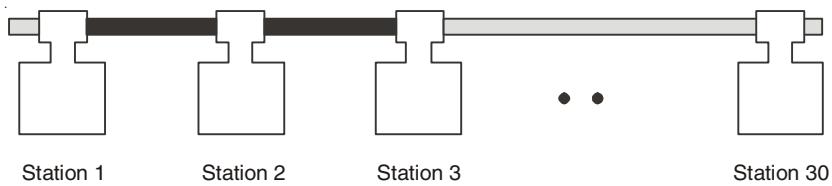


Fig. 4.12. Cabling a 10BASE-2 network.

10BROAD-36

10BROAD-36 is the only broadband network based on the CSMA/CD access protocol standardized by the IEEE. Unlike a baseband network, in which Manchester encoded signals are placed directly onto the cable, the 10BROAD-36 standard requires the use of radio frequency (RF) modems. Those modems modulate non-return to zero (NRZ) encoded signals for transmission on one channel at a specified frequency, and demodulate received signals by listening for tones on another channel at a different frequency.

A 10BROAD-36 network is constructed with a 75Ω coaxial cable, similar to the cable used in modern cable television (CATV) systems. Under the IEEE 802.3 broadband standard, either single or dual cables can be used to construct a network. If a single cable is used, the end of the cable (referred to as the headend) must be terminated with a frequency translator. That translator converts the signals received on one channel to the frequency assigned to the other channel, retransmitting the signal at the new frequency. Since the frequency band for a transmitted signal is below the frequency band 10BROAD-36 receivers scan, we say the frequency translator up converts a transmitted signal and retransmits it for reception by other stations on the network. If two cables are used, the headend simply functions as a relay point, transferring the signal received on one cable onto the second cable.

A broadband transmission system has several advantages over a baseband system. Two of the primary advantages of broadband are its ability to support multiple transmissions occurring on independent frequency bands simultaneously, and its ability to support a tree structure topology carrying multiple simultaneous transmissions. Using independent frequency bands, you can establish several independent networks. In fact, each network can be used to carry voice, data and video over a common cable. As for topology, broadband permits the use of a tree structure network, such as the structure shown in Fig. 4.13. In this example, the top of the tree would be the headend; it would contain a frequency translator, which would regenerate signals received at one frequency back onto the cable at another predefined frequency.

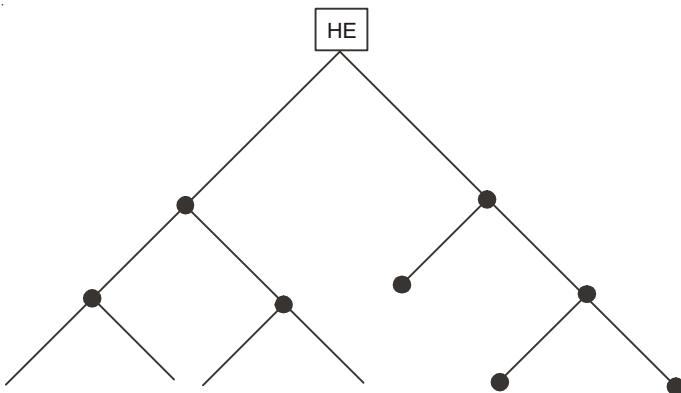


Fig. 4.13. *Broadband Tree Topology support.*

The higher noise immunity of 75Ω coaxial cable permits a 10BROAD-36 network to span 3600 m, making this medium ideal for linking buildings on a campus. In addition the ability of 10BROAD-36 to share channel space on a 75Ω coaxial cable permits organizations that have an existing CATV system, such as one used for video security, to use that cable for part or all of their broadband CSMA/CD network. Although these advantages can be significant, the cost associated with RF modems has limited the use of broadband primarily to campus environments that require an expanded network span.

In addition, the rapid development and acceptance of 10BASE-T and the decline in the cost of fiber cable to extend the span of 10BASE-T networks have severely limited what once many persons anticipated as a promising future for 10BROAD-36 networks.

1BASE-5

The 1BASE-5 standard was based on AT&T's low-cost CSMA/CD network known as StarLan. Thus, 1BASE-5 is commonly referred to as StarLan, although AT&T uses that term to refer to CSMA/CD networks operating at both 1 and 10 Mbps using unshielded twisted-pair cable. The latter is considered the predecessor to 10BASE-T. The 1BASE-5 standard differs significantly from Ethernet and 10BASE-5 standards in its use of media and topology, and in its operating rate. The 1BASE-5 standard operates at 1 Mbps and uses unshielded twisted pair (UTP) wiring in a star topology; all stations are wired to a hub, which is known as a Multiple Access Unit (MAU). To avoid confusion with the term media access unit, we will refer to this wiring concentrator as a hub.

Each station in a 1BASE-5 network contains a Network Interface Card (NIC), cabled via UTP on a point-to-point basis to a hub port. The hub is responsible for repeating signals and detecting collisions. The maximum cabling distance from a station to a hub is 250 m; up to five hubs can be cascaded together to produce a maximum network span of 2500 m. The highest level hub is known as the header hub, and it is responsible for broadcasting news of collisions to all other hubs in the network. These hubs, which are known as intermediate hubs, are responsible for reporting all collisions to the header hub.

AT&T's 1 Mbps StarLan network, along with other 1BASE-5 systems, initially received a degree of acceptance for use in small organizations. However, the introduction of 10BASE-T, which provided an operating rate 10 times that obtainable under 1BASE-5, severely limited the further acceptance of 1BASE-5 networks.

10BASE-T

In the late 1980s a committee of the IEEE recognized the requirement of organizations for transmitting Ethernet at a 10 Mbps operating rate over low cost and readily available unshielded twisted-pair cable. Although several vendors had already introduced equipment that permitted Ethernet signaling via UTP cabling, such equipment was based on proprietary designs and was not interoperable. Thus, a new task of the IEEE was to develop a standard for 802.3 networks operating at 10 Mbps using UTP cable. The resulting standard was approved by the IEEE as 802.3i in September 1990, and is more commonly known as 10BASE-T.

The 10BASE-T standard supports an operating rate of 10 Mbps at a distance of up to 100 m (328 feet) over UTP Category 3 cable without the use of a repeater. The UTP cable requires two pairs of twisted wire. One pair is used for transmitting, and the other pair is used for receiving. Each pair of wires is twisted together, and each twist is 180°. Any electromagnetic interference (EMI) or radio frequency interference (RFI) is therefore received 180° out of phase; this theoretically cancels out EMI and RFI noise while leaving the network signal. In reality, the wire between twists acts as an antenna and receives noise. This noise reception resulted in a 100 m cable limit, until repeaters were used to regenerate the signal.

➤ **10 BASE-T Network components**

A 10BASE-T network can be constructed with network interface cards, UTP cable, and one or more hubs. Each NIC is installed in the expansion slot of a computer and wired on a point-to-point basis to a hub port. When all of the ports on a hub are used, one hub can be connected to another to expand the network, resulting in a physical star, logical bus network structure. Most 10BASE-T network interface cards contain multiple connectors, which enable the card to be used with different types of 802.3 networks. For example, most modern NICs include a RJ-45 jack as well as BNC and DB-15 connectors. The RJ-45 jack supports the direct attachment of the NIC to a 10BASE-T network, and the BNC connector permits the NIC to be mated to a 10BASE-2 T-connector. The DB-15 connector enables the NIC to be cabled to a transceiver, and is more commonly referred to as the NIC's attachment unit interface (AUI) port.

The wiring hub in a 10BASE-T network functions as a multiport repeater: it receives, retimes and regenerates signals received from any attached station. The hub also functions as a filter: it discards severely distorted frames. A 10BASE-T hub tests the integrity of the

link from each hub port to a connected station by transmitting a special signal to the station. If the device does not respond, the hub will automatically shut down the port, and may illuminate a status light-emitting diode (LED) to indicate the status of each port. Hubs monitor, record and count consecutive collisions that occur on each individual station link. Since an excessive number of consecutive collisions will prevent data transfer on all of the attached links, hubs are required to cut off or partition any link on which too many collisions occurred. This partitioning enables the remainder of the network to operate in situations where a faulty NIC transmits continuously. Although the IEEE 802.3 standard does not specify a maximum number of consecutive collisions, the standard does specify that partitioning can be initiated after 30 or more consecutive collisions occur. Thus, some hub vendors initiate partitioning when 31 consecutive collisions have occurred, whereas other manufacturers use a higher value.

Although a wiring hub is commonly referred to as a concentrator, this term is not technically correct. A 10BASE-T wiring hub is a self-contained unit that typically includes 8, 10 or 12 RJ-45 ports for direct connection to stations, and a BNC and/or DB-15 AUI port to expand the hub to other network equipment. The BNC and AUI ports enable the 10BASE-T hub to be connected to 10BASE-2 and 10BASE-5 networks, respectively. For the latter, the AUI port is cabled to a 10BASE-5 MAU (transceiver), which is tapped into thick 10BASE-5 coaxial cable. One 10BASE-T hub can be connected to another with a UTP link between RJ-45 ports on each hub.

Figure 4.14 illustrates the connectors on a typical 10BASE-T hub. On some hubs, one RJ-45 jack is labeled uplink/downlink for use in cascading hubs, whereas other vendors permits any RJ-45 port to be used for connecting hubs.

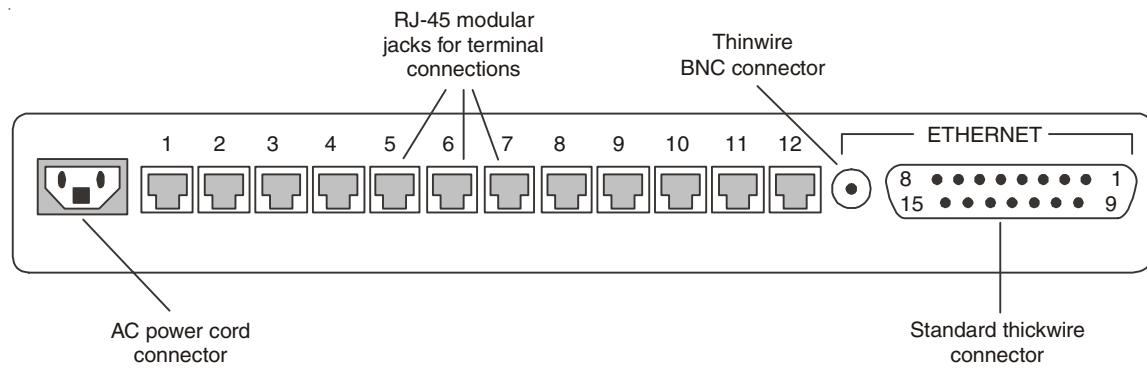


Fig. 4.14. 10BASE-T Hub Connector.

Unlike a hub, a concentrator consists of a main housing into which modular cards are inserted. Although some modular cards may appear to represent hubs, and do indeed function as 10BASE-T hubs, the addition of other modules permit the network to be easily expanded from one location and this allows additional features to be supported. For example, the insertion of a fiber optic inter-repeater module permits concentrators to be interconnected over relatively long distances of approximately 3 km.

➤ Expanding a 10BASE-T network

A 10BASE-T network can be expanded with additional hubs once the number of stations serviced has used up the hub's available terminal ports. In expanding a 10BASE-T network,

the wiring that joins each hub together is considered to represent a cable segment, and each hub is considered as a repeater. Under the 802.3 specification, no two stations can be separated by more than four hubs connected together by five cable segments.

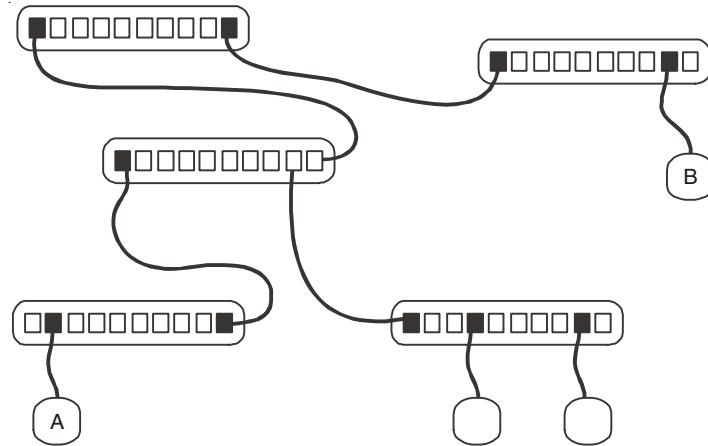


Fig. 4.15. Expanding 10BASE-T Network.

100BASE-T

The standardization of 100BASE-T, commonly known as Fast Ethernet, required an extension of previously developed IEEE 802.3 standards. In the definition process of standardization development, both the Ethernet Media Access Control (MAC) and physical layer required adjustments to permit 100 Mbps operational support. For the MAC layer, scaling its speed to 100 Mbps from the 10BASE-T 10 Mbps operational rate required a minimal adjustment, since in theory the 10BASE-T MAC layer was developed independently of the data rate. For the physical layer, more than a minor adjustment was required since Fast Ethernet was designed to support three types of media, resulting in three new names which fall under the 100BASE-T umbrella.

100BASE-T4 uses three wire pairs for data transmission and a fourth for collision detection resulting in T4 being appended to the 100BASE mnemonic. 100BASE-TX uses two pairs of category 5 UTP with one pair employed for transmission, and the second is used for both collision detection and reception of data. The third 100BASE-T standard is 100BASE-FX which represents the use of fiber optic media.

Using work developed in the standardization process of FDDI in defining 125 Mbps full-duplex signaling to accommodate optical fiber, UTP and STP through Physical Media Dependent (PMD) sublayers, Fast Ethernet borrowed this strategy. Because a mechanism was required to map the PMD's continuous signaling system to the start-stop 'half-duplex' system used at the Ethernet MAC layer, the physical layer was subdivided. This subdivision is illustrated in Fig. 4.16. The PMD sublayer supports the appropriate media to be used, whereas the convergence sublayer (CS), which was later renamed the physical coding sublayer, performs the mapping between the PMD and the Ethernet MAC layer. Although Fast Ethernet represents a tenfold increase in the LAN operating rate from 10BASE-T, to ensure proper collision detections the 100BASE-T network span was reduced to 250 m, with a maximum

of 100 m permitted between a network node and a hub. The smaller network diameter reduces potential propagation delay. When coupled with a ten-fold operating rate increase and no change in network frame size, the ratio of frame duration to network propagation delay for a 100BASE-T network is the same as for a 10BASE-T network.

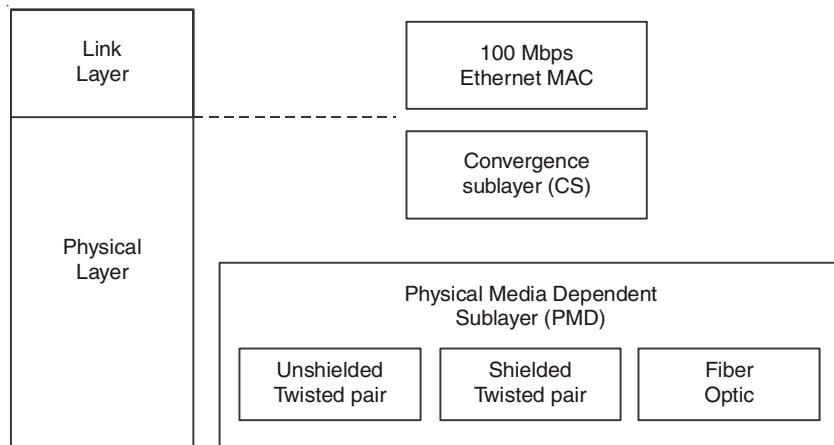


Fig. 4.16. Overview of the Fast Ethernet physical layer subdivision.

➤ Physical layer

The physical layer subdivision previously illustrated in Fig. 4.16 as indicated in the title of the Figure, presents an overview of the true layer subdivision. In actuality, a number of changes were required at the physical layer to obtain a 10 Mbps operating rate. Those changes include the use of three wire pairs for data (the fourth is used for collision detection), 8B6T ternary coding (for 100BASE-T4) instead of Manchester coding, and an increase in the clock signaling speed from 20 MHz to 25 MHz.

When the specifications for Fast Ethernet were being developed it was recognized that the physical signaling layer would incorporate medium dependent functions if support was extended to two pair cable (100BASE-TX) operations. To separate medium-dependent interfaces to accommodate multiple physical layers, a common interface referred to as the Medium Independent Interface (MII) was inserted between the MAC layer and the physical encoding sublayer. The MII represents a common point of interoperability between the medium and the MAC layer. The MII can support two specific data rates, 10 Mbps and 100 Mbps, permitting older 10BASE-T nodes to be supported at Fast Ethernet hubs. To reconcile the MII signal with the MAC signal, a reconciliation sublayer was added under the MAC layer, resulting in the subdivision of the link layer into three parts: a logical link control layer, a media access control layer and a reconciliation layer. The top portion of Fig. 4.16 illustrates this subdivision.

That portion of Fast Ethernet below the MII, which is the new physical layer, is now subdivided into three sublayers. The lower portion of Fig. 4.16 illustrates the physical sublayers for 100BASE-T4 and 100BASE-TX. The physical coding sublayer performs the data encoding, transmit, receive and carrier sense functions. Since the data coding method differs between 100BASE-T4 and 100BASE-TX, this difference requires distinct physical coding sublayers for

each version of Fast Ethernet. The Physical Medium Attachment (PMA) sublayer maps messages from the physical coding sublayer (PCS) onto the twisted-pair transmission media and vice versa.

The Medium Dependent Interface (MDI) sublayer specifies the use of a standard RJ-45 connector. Although the same connector is used for 100BASE-TX, the use of two pairs of cable instead of four results in different pin assignments.

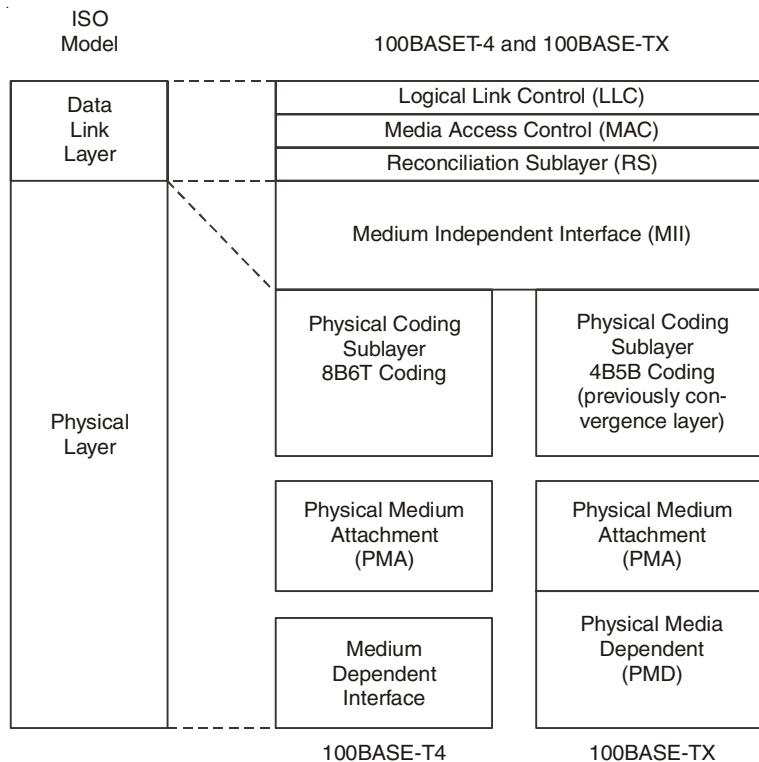


Fig. 4.17. 100BASE-T4 versus 100BASE-TX physical and data link layers.

Gigabit Ethernet

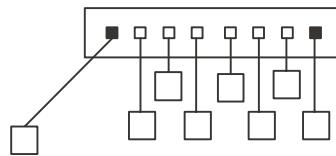
Gigabit Ethernet represents an extension to the 10 Mbps and 100 Mbps IEEE 802.3 Ethernet standards. Providing a data transmission capability of 1000-Mbps, Gigabit Ethernet supports the CMSA/CD access protocol, which makes various types of Ethernet networks scalable from 10 Mbps to 1 Gbps.

➤ Components

Similar to 10BASE-T and Fast Ethernet, Gigabit Ethernet can be used as a shared network through the attachment of network devices to a 1 Gbps repeater hub providing shared use of the 1 Gbps operating rate or as a switch, the latter providing 1 Gbps ports to accommodate high speed access to servers while lower operating rate ports provide access to 10 Mbps and 100 Mbps workstations and hubs. Although very few organizations can be expected to require the use of a 1 Gbps shared media network as illustrated in Fig. 4.18(a), the use of Gigabit switches can be expected to play an important role in providing a high-

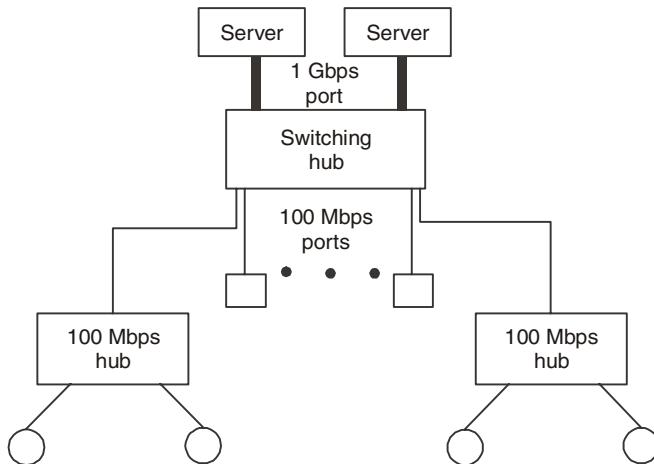
speed backbone linking 100 Mbps network users to large databases, mainframes, and other types of resources that can tax lower speed networks. In addition to hubs and switches, Gigabit Ethernet operations require workstations, bridges and routers to use a network interface card to connect to a 1 Gbps network.

(a) Shared media hub use



In a shared media environment the Gbps bandwidth provided by Gigabit Ethernet is shared among all users

(b) Switching hub use



In a switch environment each 1 Gbps port can provide a full-duplex 2 Gbps data transfer capability

Fig. 4.18. Using Gigabit Ethernet.

➤ Media support

Similar to the recognition that Fast Ethernet would be required to operate over different types of media, the IEEE 802.3z committee recognized that Gigabit Ethernet would also be required to operate over multiple types of media. This recognition resulted in the development of a series of specifications, each designed to accommodate different types of media. Thus, any discussion of Gigabit Ethernet involves an examination of the types of media the technology supports and how it provides this support. There are five types of media supported by Gigabit Ethernet—single-mode fiber, multi-mode fiber, short runs of coaxial cable or shielded twisted pair, and longer runs of unshielded twisted pair. Table 4.1 summarizes the

'flavors' of Gigabit Ethernet, indicating the IEEE designator used to reference Gigabit operations on a specific type of medium and the maximum transmission distance associated with the use of each type of medium. The actual relationship of the Gigabit 802.3z reference model to the ISO Reference Model is very similar to Fast Ethernet. Instead of a Medium Independent Interface (MII), Gigabit Ethernet uses a Gigabit Media Independent Interface (GMII). The GMII provides the interconnection between the MAC sublayer and the physical layer to include the use of an 8-bit data bus that operates at 125 MHZ plus such control signals as transmit and receiver clocks, carrier indicators and error conditions.

Table 4.1 Flavors of Gigabit Ethernet

<i>Media designator</i>	<i>Media type</i>	<i>Transmission distance</i>
1000BASE-LX	SMF	3 km
1000BASE-LX	MMF, 50 μ	550 m
1000BASE-LX	MMF, 62.5 μ	440 m
1000BASE-SX	MMF, 50 μ	550 m
1000BASE-SX	MMF, 62.5 μ	260 m
1000BASE-CX	shielded balanced copper (coax or STP)	25 m
1000BASE-T	UTP, Category 5	100 m
Key: SMF single-mode fiber		
MMF multi-mode fiber		
UTP unshielded twisted pair		
STP shielded twisted pair		

4.3.3 Frame Composition

Figure 4.19 illustrates the general frame composition of Ethernet and IEEE 802.3 frames. You will note that they differ slightly. An Ethernet frame contains an eight-byte preamble, whereas the IEEE 802.3 frame contains a seven-byte preamble followed by a one-byte start-of-frame delimiter field. A second difference between the composition of Ethernet and IEEE 802.3 frames concerns the two-byte Ethernet type field. That field is used by Ethernet to specify the protocol carried in the frame, enabling several protocols to be carried independently of one another. Under the IEEE 802.3 frame format, the type field was replaced by a two-byte length field, which specifies the number of bytes that follow that field as data. Not shown in Fig. 4.19 is the format of the Gigabit Ethernet frame.

The Gigabit Ethernet frame format is the same as the IEEE 802.3 frame format; however, if the frame length is less than 512 bytes carrier extension symbols are added to ensure that the minimum length of the frame is 512 bytes in an adapter or 520 bytes when the preamble and start of frame delimiter fields are added when a frame is transmitted.

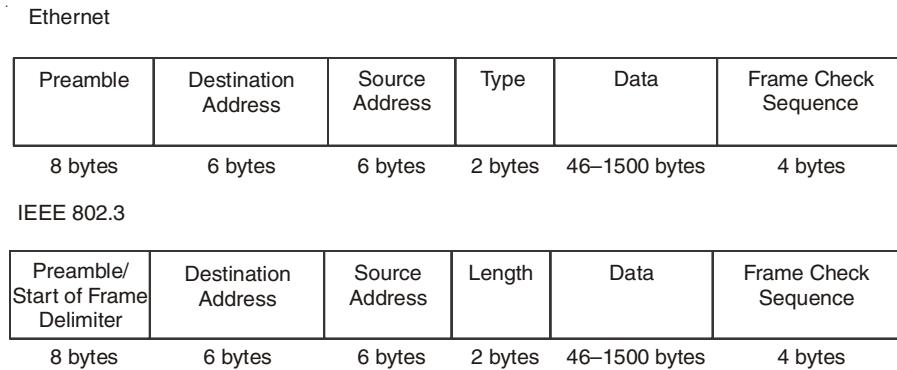


Fig. 4.19. Ethernet and IEEE 802.3 Frame Format.

Now that we have an overview of the structure of Ethernet and 802.3 frames, let us probe deeper and examine the composition of each frame field. We will take advantage of the similarity between Ethernet and IEEE 802.3 frames to examine the fields of each frame on a composite basis, noting the differences between the two when appropriate.

➤ Preamble field

The preamble field consists of eight (Ethernet) or seven (IEEE 802.3) bytes of alternating 1 and 0 bits. The purpose of this field is to announce the frame and to enable all receivers on the network to synchronize themselves to the incoming frame. In addition, this field by itself (under Ethernet) or in conjunction with the start-of-frame delimiter field (under the IEEE 802.3 standard) ensures there is a minimum spacing period of 9.6 ms between frames for error detection and recovery operations.

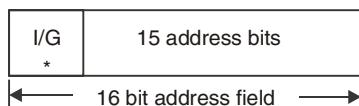
➤ Start of frame delimiter field

This field is applicable only to the IEEE 802.3 standard, and can be viewed as a continuation of the preamble. In fact, the composition of this field continues in the same manner as the format of the preamble, with alternating 1 and 0 bits used for the first six bit positions of this one-byte field. The last two bit positions of this field are 11; this breaks the synchronization pattern and alerts the receiver that frame data follows. It is important to note that the preamble and start of frame delimiter fields are only applicable for frames flowing on a network and are not included when a frame is formed in a computer, bridge or router. Thus, the addition of those fields occurs automatically by the network adapter when a frame is transmitted. This explains an area of confusion concerning minimum and maximum Ethernet frame lengths in many trade publications that deserves an elaboration. The minimum length Ethernet frame is 64 bytes when in an adapter card and 72 bytes when placed on a LAN. Similarly, the maximum length Ethernet frame is 1518 bytes when in an adapter and 1526 bytes when placed on a LAN. The preceding is applicable for both Ethernet and IEEE 802.3 versions of Ethernet, with the exception of Gigabit. As previously mentioned in this section, a Gigabit Ethernet frame must be a minimum of 512 bytes in length when in an adapter and 520 bytes when placed on a LAN.

➤ Destination address field

The destination address identifies the recipient of the frame. Although this may appear to be a simple field, in reality its length can vary between IEEE 802.3 and Ethernet frames. In addition, each field can consist of two or more subfields, whose settings govern such network operations as the type of addressing used on the LAN, and whether or not the frame is addressed to a specific station or more than one station. To obtain an appreciation for the use of this field, let us examine how this field is used under the IEEE 802.3 standard as one of the two field formats applicable to Ethernet. Figure 4.20 illustrates the composition of the source and destination address fields. As indicated, the two-byte source and destination address fields are applicable only to IEEE 802.3 networks, whereas the six-byte source and destination address fields are applicable to both Ethernet and IEEE 802.3 networks. A user can select either a two or six-byte destination address field; however, with IEEE 802.3 equipment, all stations on the LAN must use the same addressing structure. Today, almost all 802.3 networks use six-byte addressing.

A. 2 byte field (IEEE 802.3)



B. 6 byte field (Ethernet and IEEE 802.3)



I/G bit subfield '0' = individual address '1' = group address
 U/L bit subfield '0' = universally administrated addressing
 '1' = locally administrated addressing

* Set 1 '0' in source address field

Fig. 4.20. Source and destination address field formats.

➤ I/G subfield

The one-bit I/G subfield is set to a 0 to indicate that the frame is destined to an individual station, or 1 to indicate that the frame is addressed to more than one station: a group address. One special example of a group address is the assignment of all 1s to the address field. Hex FFFFFFFFFFFF is recognized as a broadcast address, and each station on the network will receive and accept frames with that destination address.

When a destination address specifies a single station, the address is referred to as a unicast address. A group address that defines multiple stations is known as a multicast address, whereas a group address that specifies all stations on the network is, as previously mentioned, referred to as a broadcast address. As an example of addressing, assume a frame has a destination address of all 1s, or hex FFFFFF. This denotes a broadcast address, resulting

in each adapter card in a station on a LAN copying the frame from the network. In comparison, if the address was a unicast address only the adapter in the station that has the destination address of the frame would copy the frame off the network.

➤ **U/L subfield**

The U/L subfield is applicable only to the six-byte destination address field. The setting of this field's bit position indicates whether the destination address is an address that was assigned by the IEEE (universally administered) or is assigned by the organization via software (locally administered).

➤ **Universal versus locally administered addressing**

Each Ethernet Network Interface Card (NIC) contains a unique address burned into its read-only memory (ROM) at the time of manufacture. To ensure that this universally administered address is not duplicated, the IEEE assigns blocks of addresses to each manufacturer. These addresses normally include a three-byte prefix, which identifies the manufacturer and is assigned by the IEEE, and a three-byte suffix, which is assigned by the adapter manufacturer to its NIC. For example, the prefix 02608C identifies an NIC manufactured by 3Com, and a prefix of hex 08002 identifies an NIC manufactured by Digital Equipment Company, the latter now owned by Compaq.

Although the use of universally administered addressing eliminates the potential for duplicate network addresses, it does not provide the flexibility obtainable from locally administered addressing. For example, under locally administered addressing, you can configure mainframe software to work with a predefined group of addresses via a gateway PC. Then, as you add new stations to your LAN, you simply use your installation program to assign a locally administered address to the NIC instead of using its universally administered address. As long as your mainframe computer has a pool of locally administered addresses that includes your recent assignment, you do not have to modify your mainframe communications software configuration. Since the modification of mainframe communications software typically requires recompiling and reloading, the attached network must become inoperative for a short period of time. Because a large mainframe may service hundreds or thousands of users, such changes are normally performed late in the evening or on a weekend. Thus, the changes required for locally administered addressing are more responsive to users than those required for universally administered addressing.

➤ **Source address field**

The source address field identifies the station that transmitted the frame. Like the destination address field, the source address can be either two or six bytes in length.

The two-byte source address is supported only under the IEEE 802.3 standard and requires the use of a two-byte destination address; all stations on the network must use two-byte addressing fields. The six-byte source address field is supported by both Ethernet and the IEEE 802.3 standard. When a six-byte address is used, the first three bytes represent the address assigned by the IEEE to the manufacturer for incorporation into each NIC's ROM. The vendor then normally assigns the last three bytes for each of its NICs.

➤ **Type field**

The two-byte type field is applicable only to the Ethernet frame. This field identifies the higher-level protocol contained in the data field. Thus, this field tells the receiving device

how to interpret the data field. Under Ethernet, multiple protocols can exist on the LAN at the same time. Xerox served as the custodian of Ethernet address ranges licensed to NIC manufacturers and defined the protocols supported by the assignment of type field values. Under the IEEE 802.3 standard, the type field was replaced by a length field, which precludes compatibility between pure Ethernet and 802.3 frames. While a pure IEEE 802.3 frame is limited to transmitting only one protocol, the IEEE recognized the necessity for its version of Ethernet to support multiple protocols. This was accomplished by the IEEE subdividing the data field into a series of fields to form what is referred to as an Ethernet-SNAP frame whose operation and utilization are described later in this section.

➤ **Length field**

The two-byte length field, applicable to the IEEE 802.3 standard, defines the number of bytes contained in the data field. Under both Ethernet and IEEE 802.3 standards, the minimum size frame must be 64 bytes in length from preamble through FCS fields. This minimum size frame ensures that there was sufficient transmission time to enable Ethernet NICs to detect collisions accurately based on the maximum Ethernet cable length specified for a network and the time required for a frame to propagate the length of the cable. Based on the minimum frame length of 64 bytes and the possibility of using two-byte addressing fields, this means that each data field must be a minimum of 46 bytes in length.

➤ **Data field**

As previously discussed, the data field must be a minimum of 46 bytes in length to ensure that the frame is at least 64 bytes in length. This means that the transmission of 1 byte of information must be carried within a 46 byte data field; if the information to be placed in the field is less than 46 bytes, the remainder of the field must be padded. Although some publications subdivide the data field to include a PAD subfield, the latter actually represents optional fill characters that are added to the information in the data field to ensure a length of 46 bytes. The maximum length of the data field is 1500 bytes.

➤ **Frame check sequence field**

The frame check sequence field, applicable to both Ethernet and the IEEE 802.3 standard, provides a mechanism for error detection. Each transmitter computes a cyclic redundancy check (CRC) that covers both address fields, the type/length field and the data field. The transmitter then places the computed CRC in the four-byte FCS field.

4.4 TOKEN-RING

There are five major types of Token-Rings. The first three types of Token-Ring networks are the focus of this section: 4, 16 and 100 Mbps Token-Ring networks that operate according to the IEEE 802.5 standard. Two additional Token-Ring networks are the Fiber Distributed Data Interface (FDDI) that operates at 100 Mbps and FDDI transmission over copper wiring, an evolving standard commonly referred to as Copper Distributed Data Interface (CDDI).

4.4.1 Topology

Although the term Token-Ring implies a ring structure, in actuality this type of LAN is either a star or star-ring structure, with the actual topology based upon the number of stations to be connected. The term star is derived from the fact that a grouping of stations and other devices, including printers, plotters, repeaters, bridges, routers and gateways, are connected in groups to a common device called a Multistation Access Unit (MAU).

Figure 4.21 illustrates a single ring formed through the use of one MAU in which up to eight devices are interconnected. Thus, for a very small Token-Ring LAN consisting of a mixture of eight or less devices, the structure actually resembles a star.

When IBM introduced its 4 Mbps Token-Ring network, its first MAU, known as the 8228, was a 10-port device, of which two ports were used for Ring-In (RI) and Ring-Out (RO) connectors which enable multiple MAUs to be connected to one another to expand the network. The remaining eight ports on the 8228 are designed to connect devices to the ring. Since IBM's eight-port 8228 reached the market, other vendors have introduced similar products with different device support capacities. You can now commonly obtain MAUs that support 4, 8, and 16 devices.

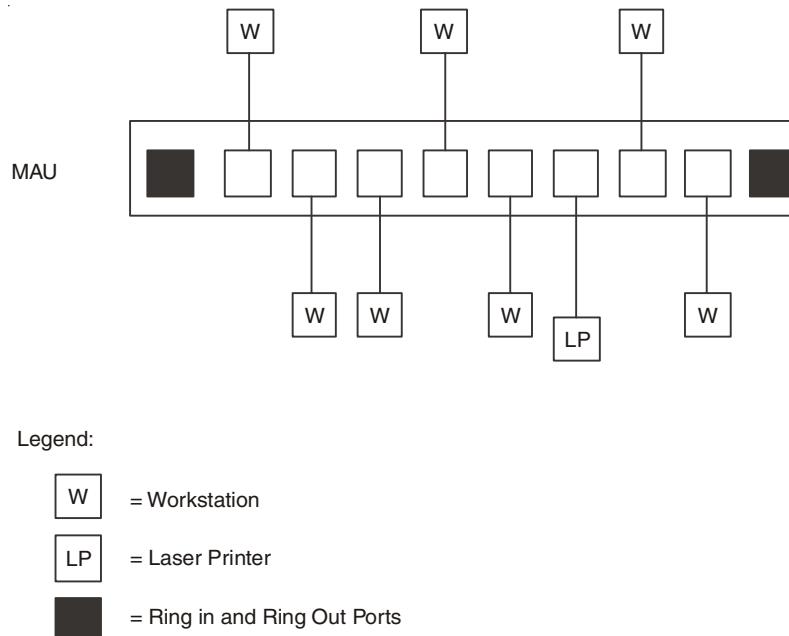


Fig. 4.21. Single-ring LAN. A single-ring LAN can support up to eight devices through their attachment to a common MAU.

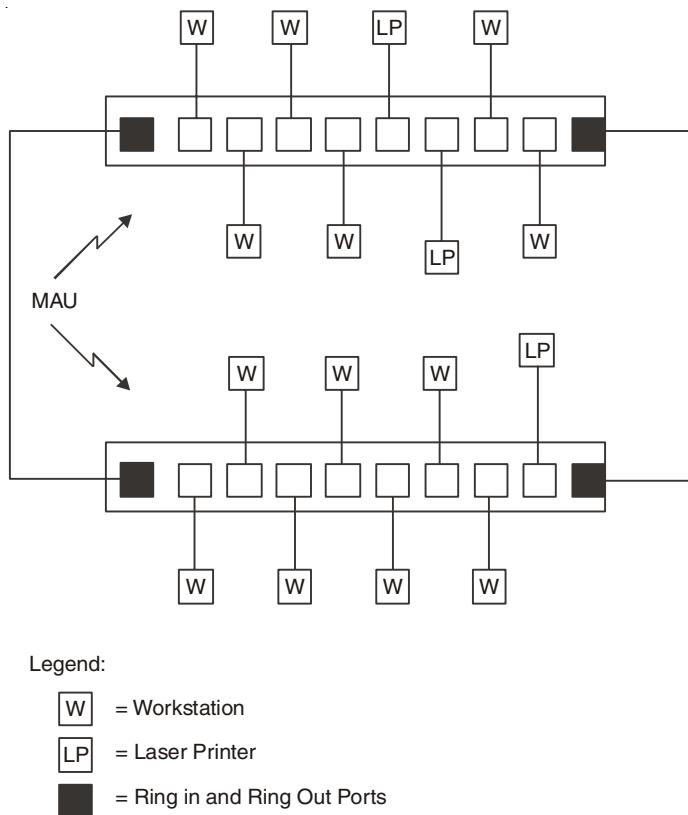


Fig. 4.22. Developing a star-ring topology.

In examining Fig. 4.21, note that an eight-port MAU is illustrated, which enables up to eight devices to be interconnected to a Token-Ring LAN. The MAU can be considered the main ring path, as data will flow from one port to another via each device connected to the port. If you have more than eight devices, you can add additional MAUs, interconnecting the MAUs via the Ring-In and Ring-Out ports located at each side of each MAU. When this interconnection occurs, by linking two or more MAUs together you form a star-ring topology, as illustrated in Fig. 4.22. In this illustration, the stations and other devices form a star structure, while the interconnection of MAUs forms the ring; hence you obtain a star-ring topology.

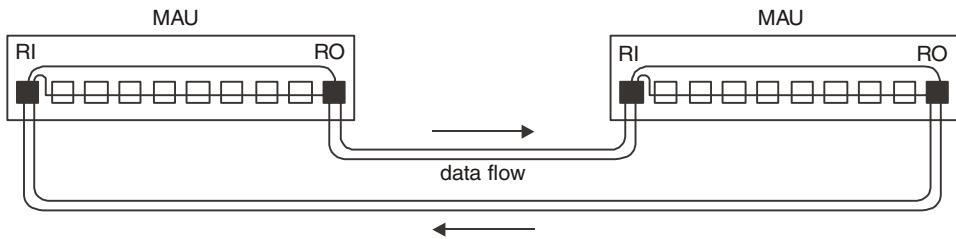
4.4.2 Redundant versus Non-redundant Main Ring Paths

When two or more MAUs are interconnected, the serial path formed by those interconnections is known as the main ring path. Connections between MAUs can be accomplished through the use of one or two pairs of wiring. One pair will be used as the primary data path, and the other pair functions as a backup data path.

The top of Fig. 4.23 illustrates the formation of a ring consisting of two MAUs in which both primary and backup paths are established to provide a redundant main ring path. If one of the cables linking the MAUs becomes disconnected, cut or crimped, the network can

continue to operate since the remaining wiring pair provides a non-redundant main ring path capability as shown in the lower portion of Fig. 4.23.

Redundant Main Ring Path



Non-Redundant Main Ring Path

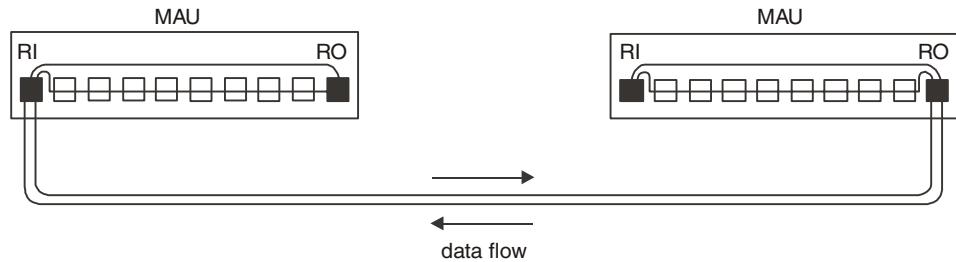


Fig. 4.23. Redundant and Non-Redundant Ring Path.

The backup capability provided by redundant main ring paths is established through the use of loop back plugs or a built-in MAU self-shorting feature.

4.4.3 Cabling and Device Restrictions

The type of cable or wiring used to connect devices to MAUs and interconnect MAUs is a major constraint that governs the size of a Token-Ring network. Since the IBM cabling system provides a large number of common types of wiring, let us first examine the type of cable defined by that cabling system.

➤ IBM cabling system

The IBM cabling system was introduced in 1984 as a mechanism to support the networking requirements of office environments. By defining standards for cables, connectors, faceplates, distribution panels, and other facilities, IBM's cabling system is designed to support the interconnection of personal computers, conventional terminals, mainframe computers, and office systems. In addition, this system permits devices to be moved from one location to another, or added to a network through a simple connection to the cabling system's wall plates or surface mounts.

The IBM cabling system specifies seven different cabling categories. Depending on the type of cable selected, you can install the selected wiring indoors, outdoors, under a carpet, or in ducts and other air spaces. The IBM cabling system uses wire which conforms to the

American Wire Gauge or AWG. The IBM cabling system uses wire between 22 AWG (0.644 mm) and 26 AWG (0.405 mm). Since a large-diameter wire has less resistance to current flow than a small one, a smaller AWG permits cabling distances to be extended in comparison with a higher AWG cable. It is important to note that the IBM cabling system was introduced prior to the EIA/TIA-568 specification that defined five categories of unshielded twisted-pair cabling as well as fiber for use in buildings. Although the IBM cabling system is no longer marketed by IBM, many independent vendors continue to offer cables according to the various categories supported by that system. More modern Token-Ring products are designed for UTP connectors, with Category 3 UTP able to support 4 and 16 Mbps Token-Ring communications, while Category 5 cable is required to support communications to devices that operate according to the recently standardized 100 Mbps High Speed Token-Ring specification. In most organizations it is quite common for all copper cabling to involve the use of Category 5 cable, as this enables an organization to retain its cabling infrastructure even if it replaces one type of network with another.

➤ **Type 1**

The IBM cabling system Type 1 cable contains two twisted pairs of 22 AWG conductors. Each pair is shielded with a foil wrapping, and both pairs are surrounded by an outer braided shield or with a corrugated metallic shield. One pair of wires uses shield colors of red and green, and the second pair of wires uses shield colors of orange and black. The braided shield is used for indoor wiring, whereas the corrugated metallic shield is used for outdoor wiring. Type 1 cable is available in two different designs: plenum and non-plenum. Plenum cable can be installed without the use of a conduit, and non-plenum cable requires a conduit. Type 1 cable is typically used to connect a distribution panel or multistation access unit and the faceplate or surface mount at a workstation.

➤ **Type 2**

Type 2 cable is actually a Type 1 indoor cable with the addition of four pairs of 22 AWG conductors for telephone usage. Due to this, Type 1 cable is also referred to as data-grade twisted-pair cable, while Type 2 cable is known as two data-grade and four-grade twisted pair. Due to its voice capability, Type 2 cable can support PBX interconnections. Like Type 1 cable, Type 2 cable supports plenum and non-plenum designs. Type 2 cable is not available in an outdoor version.

➤ **Type 3**

Type 3 cable is conventional twisted pair telephone wire, with a minimum of two twists per foot. Both 22 AWG and 24 AWG conductors are supported by this cable type. One common use of Type 3 cable is to connect PCs to MAUs in a Token-Ring network.

➤ **Type 5**

Type 5 cable is fiber optic cable. Two 100/140 nm optical fibers are contained in a Type 5 cable. This cable is suitable for indoor non-plenum installation or outdoor aerial installation. Due to the extended transmission distance obtainable with fiber-optic cable, Type 5 cable is used in conjunction with the IBM 8219 Token-Ring Network Optical Fiber Repeater to interconnect two MAUs up to 6600 feet (2 km) from one another.

➤ **Type 6**

Type 6 cable contains two twisted pairs of 26 AWG conductors for data communications. It is available for non-plenum applications only and its smaller diameter than Type 1 cable makes it slightly more flexible. The primary use of Type 6 cable is for short runs as a flexible path cord. This type of cable is often used to connect an adapter card in a personal computer to a faceplate which, in turn, is connected to a Type 1 or Type 2 cable which forms the backbone of a network.

➤ **Type 8**

Type 8 cable is designed for installation under a carpet. This cable contains two individually shielded, parallel pairs of 26 AWG conductors with a plastic ramp designed to make under-carpet installation as unobtrusive as possible. Although Type 8 cable can be used in a manner similar to Type 1, it only provides half of the maximum transmission distance obtainable through the use of Type 1 cable.

➤ **Type 9**

Type 9 cable is essentially a low-cost version of Type 1 cable. Like Type 1, Type 9 cable consists of two twisted pairs of data cable; however, 26 AWG conductors are used in place of the 22 AWG wire used in Type 1 cable. As a result of the use of a smaller diameter cable, transmission distances on Type 9 cable are approximately two-thirds of those obtainable through the use of Type 1 cable. The color coding on the shield of Type 9 cable is the same as that used for Type 1 cable. All seven types of cables defined by the IBM cabling system can be used to construct Token-Ring networks. However, the use of each type of cable has a different effect on the ability to connect devices to the network, the number of devices that can be connected to a common network, the number of wiring closets in which MAUs can be installed to form a ring, and the ability of the cable to carry separate voice conversations. The latter capability enables a common cable to be routed to a user's desk where a portion of the cable is connected to their telephone, and another portion of the cable is connected to their computer's Token-Ring adapter card.

➤ **Connectors**

The IBM cabling system includes connectors to terminate both data and voice conductors. The data connector has a unique design based on the development of a latching mechanism which permits it to mate with another identical connector.

Figure 4.24 illustrates the IBM cabling system data connector. Its design makes it self-shorting when disconnected from another connector. This provides a Token-Ring network with electrical continuity when a station is disconnected. Unfortunately, the data connector is very expensive in comparison to RS-232 and RJ telephone connectors with the typical retail price of the data connector between \$4 and \$5, whereas RS-232 connectors cost approximately \$1 and an RJ telephone connector can be purchased for a dime or so.

Due to the high cost of data connectors and cable, the acceptance of the IBM cabling system by end-users never reached its potential. Instead, Category 3 and Category 5 structured wiring with RJ45 connectors is commonly used to construct Token-Ring networks.

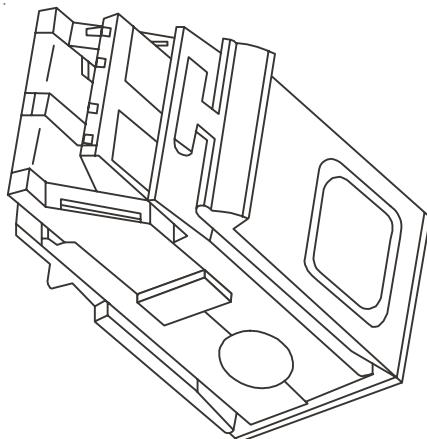


Fig. 4.24. IBM Cabling system Connector.

4.4.4 Design Constraints

There are several cabling and device constraints you must consider when designing a Token-Ring network to ensure that the network will work correctly. First, you must consider the maximum cabling distance between each device and an MAU that will service the device. The cable between the MAU and the device is referred to as a lobe, with the maximum lobe distance being 100 m (330 feet) at both 4 and 16 Mbps. This means that you must consider the lobe distance in conjunction with the cabling distance restrictions between MAUs if you have more than eight devices to be connected to a Token-Ring LAN. In addition, for larger networks, you must also consider restrictions on the number of MAUs in the network and their placement in wiring closets, as well as a parameter known as the adjusted ring length, because they collectively govern the maximum number of devices that can be supported.

➤ **Intra-MAU cabling distances**

Table 4.2 lists the maximum intra-MAU cabling distances permitted on a Token-Ring network for the two most commonly used types of IBM cables. Those distances can be extended through the use of repeaters; however, their use adds to both the complexity and cost of the network. Because 100 Mbps Token-Ring does not presently support shared media operations, it is not used to interconnect MAUs and was omitted from the table.

As indicated in Table 4.2, the cabling distance between MAUs depends on both the operating rate of the LAN (4 Mbps or 16 Mbps) and the type of cable used. Type 1 is a double-shielded pair cable, and Type 3 is non-shielded twisted-pair telephone wire.

In examining the entries in Table 4.2, it may appear odd that the maximum intra-MAU cable distance is the same for both 4 and 16 Mbps networks when Type 1 cable is used. This situation occurred because, at the time IBM set a 100 m recommended limit for a 4 Mbps network, the company took into consideration the need to reuse the same cabling when customers upgraded to a 16 Mbps operating rate. When using Type 3 cable, distances are shorter because signal line noise increases in proportion to the square root of frequency. Thus, upgrading the operating rate from 4 to 16 Mbps with Type 3 cable decreases the

maximum permissible distance. As you plan to extend your network to interconnect additional devices, you must also consider the maximum number of MAUs and devices supported by a Token-Ring network. If you use Type 1 cable, you are limited to a maximum of 33 MAUs and 260 devices. If you use Type 3 cable, you are limited to a maximum of 9 MAUs and 72 devices. These limitations are applicable to both 4 and 16 Mbps networks; however, they represent a maximum number of MAUs and devices and do not indicate reality in which a lesser number of MAUs may be required due to the use of multiple wiring closets or a long adjusted ring length. Due to the role played by the adjusted ring length in governing the number of MAUs, let us examine what an adjusted ring length is and how it functions as a constraint.

➤ Adjusted ring length

To fully understand the reason why we must consider the adjusted ring length (ARL) of a Token-Ring network requires a discussion on the network's ability to operate with a faulty cabling section. To illustrate this capability, consider the three-MAU network illustrated in Fig. 4.25 Under normal network operation, data are transmitted from RO to RI between MAUs and over the main ring path which connects the last MAU's RO port to the first MAU's RI port. If an attached device or a lobe cable fails, the lack of voltage on the MAU's port causes the port to be bypassed, permitting information from other stations to flow on the main ring path. If an MAU or a cable interconnecting two MAUs fails, the previously described built-in backup capability of MAUs permits the network to continue operating. This backup capability permits the Token-Ring to be re-established in one of two ways, dependent on the capability of the MAUs used in the network. Some MAUs have a self-shorting capability, which means that RI and RO connectors are joined together without requiring the use of a cable or plug to complete a ring. Other MAUs require the use of a cable or plug between the RI and RO connectors. By using the self-shorting capability, or using a cable or plug in the input and output connectors of the MAUs located at both ends of a failed cable, the ring can be reconfigured for operation. Figure 4.25 illustrates the reconfigured ring.

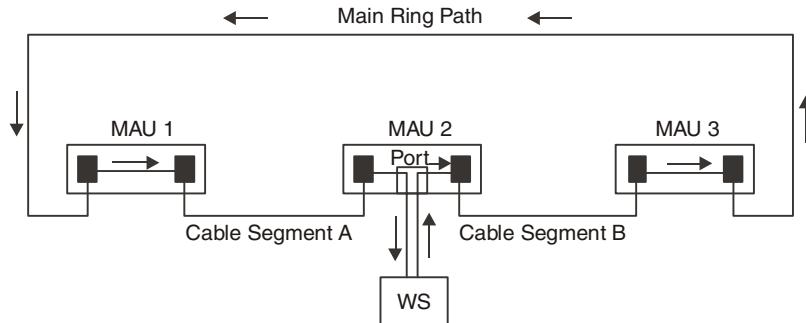
Table 4.2. Intra MAU cable distance (In feet)

<i>Operating rate</i>	<i>Type of cable</i>	
	<i>Type 1</i>	<i>Type 3</i>
4-Mbps	330	1000
16-Mbps	330	250

Note that the total cable length of the main ring path and available cable segments represents the adjusted ring length. In actuality, the adjusted ring length is the total ring length (main ring path plus all cable segments) less the shortest cable segment between MAUs. The reason why we subtract the shortest cable segment is that doing so provides the longest total cable distance for a reconfigured ring. It is that distance that a signal must be capable of flowing around a ring without excessive distortion adversely affecting the signal. For example, suppose that the main ring path is 300 feet and cable segments A and B are

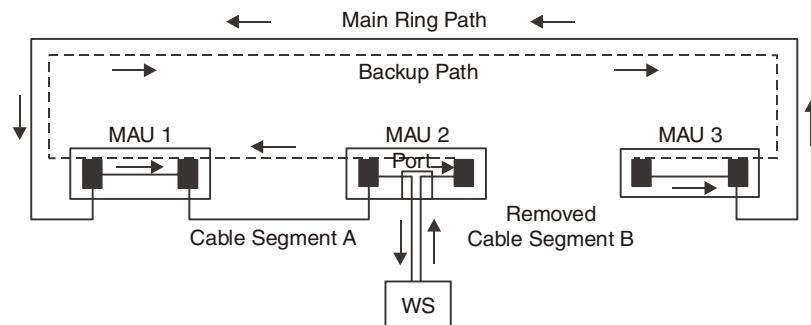
200 and 150 feet, respectively. Then, the total drive distance is $300+200+150$, or 650 feet, and the adjusted ring length is $650-150$, or 500 feet.

a. Normal network operation



$$\text{Total cable length} = \text{Main ring path} + \text{Cable segment A} + \text{Cable segment B}$$

b. Reconfigured ring due to faulty cable segment



$$\text{Adjusted ring length} = \text{Main ring path} + \text{Cable segment A}$$

Fig. 4.25. Computing the adjusted ring length.

4.4.5 Transmission Formats

Three types of transmission formats are supported on a Token-Ring network: token, abort and frame.

➤ Token

The token format, as illustrated in the top of Fig. 4.26, is the mechanism by which access to the ring is passed from one computer attached to the network to another device connected to the network. Here the token format consists of three bytes, of which the starting and ending delimiters are used to indicate the beginning and end of a token frame. The middle byte of a token frame is an access control byte. Three bits are used as a priority indicator, three bits are used as a reservation indicator, and one bit is used for the token bit, and another bit position functions as the monitor bit. When the token bit is set to a binary

0, it indicates that the transmission is a token. When it is set to a binary 1, it indicates that data in the form of a frame is being transmitted.

➤ Abort

The second Token-Ring frame format signifies an abort token. In actuality, there is no token, since this format is indicated by a starting delimiter followed by an ending delimiter. The transmission of an abort token is used to abort a previous transmission. The format of an abort token is illustrated in Fig. 4.26 (b).

➤ Frame

The third type of Token-Ring frame format occurs when a station seizes a free token. At that time the token format is converted into a frame which includes the addition of frame control, addressing data, an error detection field and a frame status field. The format of a Token-Ring frame is illustrated in Fig. 4.26 (c). By examining each of the fields in the frame, we will also examine the token and token abort frames, due to the commonality of fields between each frame.

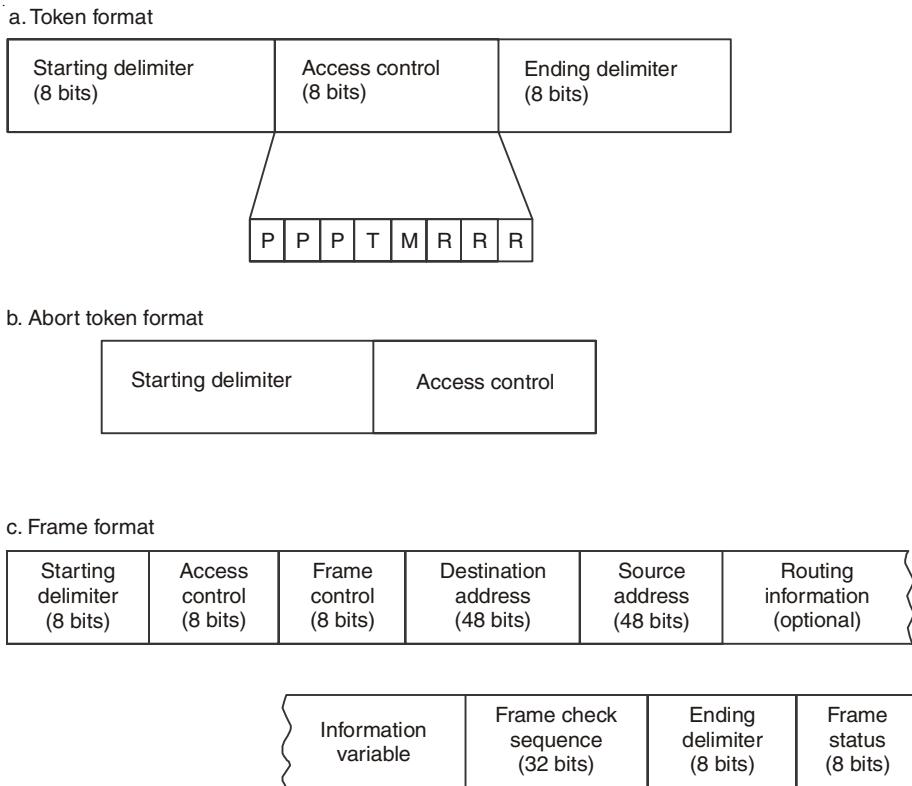


Fig. 4.26. Token, Abort and Frame format.

Starting/ending delimiters

The starting and ending delimiters mark the beginning and ending of a token or frame. Each delimiter consists of a unique code pattern which identifies it to the network.

Non-data symbols

Under Manchester and Differential Manchester encoding there are two possible code violations. Each code violation produces what is known as a non-data symbol, and it is used in the Token-Ring frame to denote starting and ending delimiters similar to the use of the flag in an HDLC frame. However, unlike the flag whose bit composition 0111110 is uniquely maintained by inserting a 0 bit after every sequence of five set bits and removing a 0 following every sequence of five set bits, Differential Manchester encoding maintains the uniqueness of frames by the use of non-data J and non-data K symbols. The two non-data symbols each consist of two half-bit times without a voltage change. The J symbol occurs when the voltage is the same as that of the last signal, and the K symbol occurs when the voltage becomes opposite of that of the last signal.

Access control

The second field in both token and frame formats is the access control byte. As illustrated at the top of Fig. 4.26, this byte consists of four subfields, and it serves as the controlling mechanism for gaining access to the network. When a free token circulates the network, the access control field represents one-third of the length of the frame since it is prefixed by the start delimiter and suffixed by the end delimiter. The lowest priority that can be specified by the priority bits in the access control byte is 0 (000), whereas the highest is seven (111), providing eight levels of priority.

Table 4.3 lists the normal use of the priority bits in the access control field. Workstations have a default priority of three, whereas bridges have a default priority of four. To reserve a token, a workstation inserts its priority level in the priority reservation subfield. Unless another station with a higher priority bumps the requesting station, the reservation will be honored and the requesting station will obtain the token. If the token bit is set to 1, this serves as an indication that a frame follows instead of the ending delimiter.

Table 4.3. Priority Bit Settings

Priority bits	Priority
000	Normal user priority, MAC frames that do not require a token and response type MAC frames
001	Normal user priority
010	Normal user priority
011	Normal user priority and MAC frames that require tokens
100	Bridge
101	Reserved
110	Reserved
111	Specialized station management

Frame control

The frame control field informs a receiving device on the network of the type of frame that was transmitted and how it should be interpreted. Frames can be either Logical Link

Control (LLC) or reference physical link functions according to the IEEE 802.5 media access control (MAC) standard. A media access control frame carries network control information and responses, and a logical link control frame carries data. The eight-bit frame control field has the format FFZZZZZZ, where FF are frame definition bits. The top part of Table 4.4 indicates the possible settings of the frame bits and the assignment of those settings. The ZZZZZZ bits convey media access control (MAC) buffering information when the FF bits are set to 00. When the FF bits are set to 01 to indicate an LLC frame, the ZZZZZZ bits are split into two fields, designated rrrYYY. Currently, the rrr bits are reserved for future use and are set to 000. The YYY bits indicate the priority of the logical link control (LLC) data. The lower portion of Table 4.4 indicates the value of the Z bits when used in MAC frames to notify a Token-Ring adapter that the frame is to be expressed buffered.

Table 4.4 Frame Control Subfields

F bit settings	Assignment
00	MAC frame
01	LLC frame
10	Undefined (reserved for future use)
11	Undefined (reserved for future use)
Z bit settings	Assignment
000	Normal buffering
001	Remove ring station
010	Beacon
011	Claim token
100	Ring purge
101	Active monitor present
110	Standby monitor present

Destination address

Although the IEEE 802.5 standard is similar to the 802.3 standard in its support of 16- and 48-bit address fields, almost all implementations of Token-Ring use 48-bit addresses. The destination address field is made up of five subfields as illustrated in Fig. 4.27. The first bit in the destination address identifies the destination as an individual station (bit set to 0) or as a group (bit set to 1) of one or more stations. The latter provides the capability for a message to be broadcast to a group of stations.

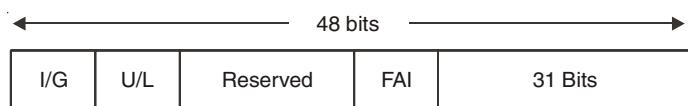


Fig. 4.27. Destination Address subfields.

Universally and locally administered addresses

Similar to the IEEE 802.3 standard, universally administered addresses are assigned in blocks of numbers by the IEEE to each manufacturer of Token-Ring equipment, with the manufacturer encoding a unique address into each adapter card. Locally administered addressing permits users to temporarily override universally administered addressing and can be used to obtain addressing flexibility.

Functional address indicator

The functional address indicator subfield in the destination address identifies the function associated with the destination address, such as a bridge, active monitor, or configuration report server. The functional address indicator indicates a functional address when set to 0 and the I/G bit position is set to a 1, the latter indicating a group address. This condition can only occur when the U/L bit position is also set to a 1, and it results in the ability to generate locally administered group addresses that are called functional addresses.

Source address

The source address field always represents an individual address which specifies the adapter card responsible for the transmission. The source address field consists of three major subfields as illustrated in Fig. 4.28. When locally administered addressing occurs, only 24 bits in the address field are used, because the 22 manufacturer identification bit positions are not used. The routing information bit identifier identifies the fact that routing information is contained in an optional routing information field. This bit is set when a frame is routed across a bridge using IBM's source routing technique.

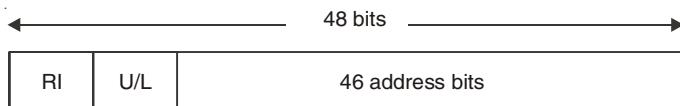


Fig. 4.28. Source Address Field.

Frame check sequence

The frame check sequence field contains four bytes which provide the mechanism for checking the accuracy of frames flowing on the network. The cyclic redundancy check data included in the frame check sequence field cover the frame control, destination address, source address, routing information and information fields. If an adapter computes a cyclic redundancy check that does not match the data contained in the frame check sequence field of a frame, the destination adapter discards the frame information and sets an error bit (E bit) indicator. This error bit indicator actually represents a ninth bit position of the ending delimiter, and it serves to inform the transmitting station that the data were received in error.



A = Address-Recognized Bits

B = Frame-Copied Bits

r = Reserved Bits

Fig. 4.29. Frame Status Field.

Frame status

The frame status field serves as a mechanism to indicate the results of a frame's circulation around a ring to the station that initiated the frame, Fig. 4.29 indicates the format of the frame status field. The frame status field contains three subfields that are duplicated for accuracy purposes, since they reside outside of CRC checking. One field (A) is used to denote whether an address was recognized, and a second field (C) indicates whether the frame was copied at its destination. Each of these fields is one bit in length. The third field, which is two bit positions in length (rr), is currently reserved for future use.

4.5 TOKEN BUS

This type of LAN has mainly been used in automated factory applications. It is now obsolescent and only included for the sake of completeness, but will be dealt with briefly. Physical layer specifications include a number of coaxial-based media derived from CATV technology with the idea of utilizing readily available, low-cost components. Table 4.5 shows some token bus variants.

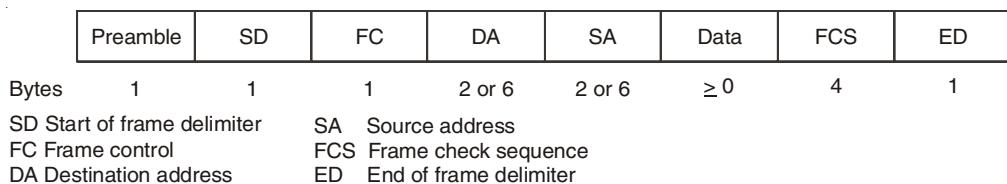
Although a bus topology is used, a logical ring is established. Each transmitted frame appears at all stations. Stations are so arranged in a ring by declaring, for a given station, the address of the next station as a **successor** and that of the preceding station a **predecessor**. Tokens and frames then circulate around the ring passing from each station to its successor in turn. In order to effect ring-like transmission it is the responsibility of a successor station only to retransmit a frame that is not destined for itself. It would be impracticable, if not chaotic, if each station, all of which receive each frame, were also to retransmit each frame that was not destined for itself. Provision exists within the protocol for stations to be added to, or removed from, the ring and to recover from abnormal or fault conditions.

The frame structure is shown in Fig. 4.30. As in most bus-based topologies, there is no continuous frame transmission on the medium. Therefore all frames commence with a preamble to enable clock resynchronization. This is followed by a start of frame delimiter (SD). Although the frame is similar to that of IEEE 802.3 it differs in that rather than specifying length, an end of frame delimiter (ED) is used. Additionally, there is a frame control (FC) field. Two FC field bits indicate:

- 00 MAC control frame
- 01 LLC data frame
- 10 Station management data frame
- 11 Special-purpose data frame

Table 4.5 IEEE 802.4 Token Bus Physical Medium Variants

	<i>Transmission medium</i>	<i>Signalling technique</i>	<i>Signalling rate (Mbps)</i>	<i>Maximum segment length (m)</i>
Broadband	Coaxial cable (75 Ω)	Broadband (AM/PSK)	1, 5, 10	Not specified
Carrier band	Coaxial cable (75 Ω)	Broadband (FSK)	1, 5, 10	7600
Optical fibre	Optical fibre	ASK-Manchester	5, 10, 20	Not specified

**Fig. 4.30. IEEE 802.4 Token Bus Frame.**

4.6 WIRELESS LAN (IEEE 802.11)

Wireless LANs do not use any wires or cables to interconnect stations physically. They operate using either infrared (IR) or radio propagation, although IR operation is rare. Current technology may only support operation at best of the order of several tens of Mbps. The dominant transmission technique by radio is in what can only be described as a hostile environment. Transmission is subject to a variety of fading, noise and interference difficulties and physical obstacles may give rise to dead spots.

LANs are now available which operate at 100 Mbps or more. The question arises: why install a wireless LAN (WLAN) with relatively low speed of operation? There are a number of applications of WLANs which are attractive. In large, open areas, such as stock exchange trading floors and historic buildings, it would be impracticable or visually unacceptable to cable stations together. An area of local area networking now proving popular is that of home LANs where two, or more, computers, printers and so forth may be interconnected as a WLAN and so avoid similar difficulties of cabling to those above.

Although WLANs may be free standing, they may alternatively be connected to either a cable-based LAN serving other physical areas within the site, or a backbone. In such circumstances a WLAN offers a **LAN extension** facility for connection of more awkward areas which are then linked to a cable LAN. A single **access point** (AP) is required within the WLAN. The access point may be implemented as a station and is the 'conduit' through which the WLAN interconnects to the wired LAN or backbone.

Another developing area for the use of WLANs is *ad hoc* networks where a peer-to-peer network may be quickly set up and dismantled. An example might be to network some

computers quickly at a stand in a trade fair. No temporary cabling is necessary and the physical siting and mobility of computers is very flexible.

Nomadic access provides a wireless link between a LAN and a mobile computer such as a laptop or personal digital assistant (PDA). This enables easy transfer of data between a fixed network and a computer used by people who for instance travel about and need to update data upon return to a central office. Other applications enable guest users to access networks at different premises, or various stock control applications in retail stores, warehouses and so on.

Another application of WLANs is for interconnection between buildings, or **wireless interconnect**. This is suitable for buildings that are relatively close together and where conventional cabling may be impossible for a variety of reasons, for example a river may separate them. Strictly speaking, wireless interconnect is not a LAN as such. Rather it is a network in its own right but without end-users. Nevertheless this application is typically considered within the scope of WLANs. Such an arrangement would usually include bridging or routing to inhibit intra-frame traffic within the WLAN being unnecessarily forwarded over the wireless interconnect.

4.6.1 IEEE 802.11 Recommendation

Operation may be at IR, with a range of a few tens of metres, or be radio-based and have a range about one order more. RF operation is by means of one, or other, of the unlicensed industrial, scientific and medical (ISM) bands at either 2.4 or 5 GHz. The use of 2.4 GHz is possible virtually anywhere in the world. However, although 5 GHz is available for use in North America, it is not universally available throughout Europe or Asia.

The IEEE introduced its first WLAN recommendations, IEEE 802.11 (also known as WiFi), in 1997 operating at 2.4 GHz. Since then further development has occurred. IEEE 802.11b was introduced to overcome the speed limitations of, and be compatible with, IEEE 802.11. IEEE 802.11a (also known as WiFi5 as it operates at 5 GHz) was introduced to improve further upon speed performance but, since a different transmission technique and frequency were employed, this meant that it was no longer compatible with the other two standards. The choice of 5 GHz operation, to provide the additional bandwidth for an improvement in speed, has restricted its use primarily to North America. There is, as always, a trade-off between speed of operation and distance. IEEE 802.11a, although faster than IEEE 802.11b, has a range which is approximately one order less. Nevertheless, for a given distance, IEEE 802.11a is anticipated to outperform IEEE 802.11b on speed.

In a wired LAN there must be a high degree of security since it is relatively easy for unauthorized parties physically to access the network, or its signals. WLANs are far less secure since radio signals especially are not neatly confined to the geographical area of a network. It is relatively straightforward both to monitor radio signals and even attempt to interconnect a WLAN station, whilst outside of the building or site containing a LAN. With this in mind IEEE 802.11 includes an authentication procedure to ensure unauthorized parties do not gain access to a network. In addition the **Wired Equivalent Privacy** (WEP), a security protocol, is specified to provide encryption and privacy. The WEP algorithm is specified to provide privacy and authentication. Symmetrical key encryption is provided in

which communicating stations share a secret key, based upon an encryption algorithm known as RC4. In addition two levels of authentication are provided: open system authentication, whereby both parties exchange their network identities; and shared key authentication, where each party shares an agreed secret key and enables full authentication of both parties to each other.

The **protocol stack** for the various recommendations is shown in Fig. 4.31. The LLC sublayer is broadly similar to that described in IEEE 802.2 earlier in this chapter and common to all derivatives of IEEE 802.11. However, there are significant differences regarding the MAC layer protocol, each of which will be described shortly. At the physical layer we see that transmission is by means of either IR or radio propagation using either **spread spectrum (SS)** techniques or **orthogonal frequency division multiplexing (OFDM)**.

At the physical layer IR may be used, operating at a wavelength of around 900 nm. Line of sight propagation is not essential since transmission may succeed by means of diffusion. Alternatively, RF propagation centres upon a carrier frequency of 2.4 GHz. In either case data may be transmitted at 1 or 2 Mbps. Two forms of SS operation are recommended. One form is frequency hopping SS (FH-SS) which uses up to 79.1 MHz channels. The precise number of channels is dependent upon the territory in which the WLAN is operated, for example 23 channels may be used in Japan whereas 70 are used in the USA. Different territories have different numbers of channels dictated by their national frequency licensing authorities to minimize the risk of interference with other services. Data is only transmitted upon a single 1 MHz bandwidth channel at any one time at the full 1 or 2 Mbps. The carrier frequency of the channel then ‘hops’ from channel to channel at a certain rate. The hopping sequence across the available channels is pseudorandom both to improve transmission performance and also to help prevent eavesdropping.

Logical Link Control					Data link layer
MAC layer					
FH-SS 2.4 GHz 1 or 2 Mbps	DS-SS 2.4 GHz 1 or 2 Mbps	IR 1 or 2 Mbps	OFDM 5 GHz ≤ 54 Mbps	DS-SS 2.4 GHz 5.5 or 11 Mbps	Physical layer
IEEE 802.11		IEEE 802.11a		IEEE 802.11b	

Fig. 4.31. IEEE 802.11 Protocol stack.

The attractions of SS operation in WLANs are:

1. The radio environment often has a large amount of interference and spurious noise from lighting, machinery, electrical equipment, etc. By rapidly changing channels some immunity to narrowband and frequency-dependent interference may be achieved.
2. At the frequency of operation employed **multipath propagation** is common whereby signals propagate from transmitter to receiver over a number of different ray paths (due to reflections). In consequence the received signal is enhanced if the rays are

constructive or reduced, or even cancelled out, if the rays are destructive. If, on a particular channel, destructive interference is experienced then, when a hop to a new frequency occurs, it is relatively unlikely that interference will continue to be destructive. Therefore, a satisfactory signal is guaranteed some of the time.

The other form of SS specified, which also operates at 1 or 2 Mbps, is that of direct sequence SS (DS-SS). A number of channels may be employed, each operating at the full 1 or 2 Mbps. As with FH-SS the actual number of channels that may be used is governed by national regulation. Japan only permits one channel, in the USA seven channels may be used and most European countries use 13 channels. The data is multiplied by a higher rate pseudorandom binary sequence prior to modulation, and hence the spectrum is spread. This then modulates a 2.4 GHz carrier frequency and is then transmitted, as illustrated in Fig. 4.32. DS-SS has an equivalent bandwidth to that of FH-SS and enjoys the same advantages in regard to combating noise and interference discussed earlier.

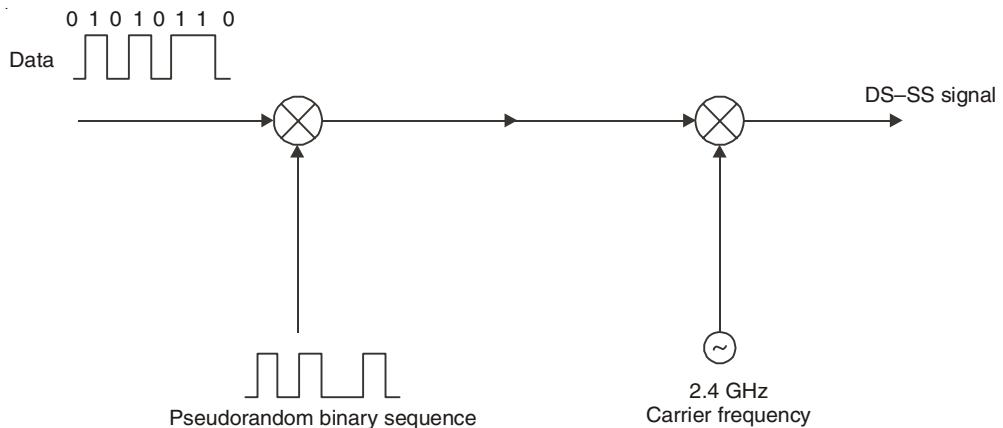


Fig. 4.32. Direct sequence spread spectrum.

4.6.2 IEEE 802.11a Recommendations

In 1997 a derivative of IEEE 802.11, 802.11a, was introduced. At the physical layer orthogonal FDM (OFDM) is used in the 5 GHz unlicensed ISM band. Up to 52 individual carrier frequencies, and hence subcarriers, may exist within the band. The subcarriers are orthogonal in order to minimize co channel interference, or inter channel interference (ICI). The use of multiple, contiguous, carriers is similar to frequency division multiplexing (FDM). (This signal has similarities to that of DMT used in ADSL.)

Although similar in principle to that of FH-SS inasmuch that multiple subcarriers, or channels, exist, hopping is not employed. Rather the serial data stream is converted into parallel form using an n -bit register, n depending upon the precise number of subcarriers. Each parallel bit is then used individually to modulate a single subcarrier. Finally the subcarriers are summed and then transmitted as a composite signal. Modulation is either PSK or QAM and various sizes of signal constellations are used. As a result a variety of data rates are possible and provides a much higher data rate, up to 54 Mbps, than IEEE 802.11 recommendations. The precise rate used is dynamically adapted and depends upon the

quality of RF transmissions and the network loading. In addition, convolutional encoding is employed to provide FEC in order to combat errors introduced in transmission.

The advantages of OFDM are:

1. As with FH-SS, the use of multiple channels means that frequency-selective interference, whereby some channels may experience significant noise and interference and others may not, is mitigated since some channels will operate satisfactorily. Therefore, some data will succeed and FEC may be successful in correcting erroneous data.
2. However, the main advantage stems from the fact that for n channels the signaling rate per channel is reduced by a factor n . Therefore, the symbol length is extended by the same factor. The effect of multipath propagation gives rise to multiple rays, all arriving at the receiver at different intervals in time. As a result the edges of symbols suffer significant distortion and so cause ISI. Now, since symbol length is increased with OFDM, the impact of ISI can be made far less significant. This is because the time spread between the arrival of the first and the last multiple ray becomes relatively small in comparison with the length of a symbol.

4.6.3 IEEE 802.11b Recommendation

This variant, introduced in 1999, is similar to that of IEEE 802.11 in that it too uses FH-SS, but at a higher data rate of either 5.5 or 11 Mbps, and yet retains transmissions with the same bandwidth. It achieves increased speed for the same bandwidth using a complex technique known as **complementary code keying** (CCK). CCK is beyond the scope of this text but, briefly, data is taken 8 bits at a time, encoded, re-serialized and finally applied to a QPSK modulator.

4.6.4 IEEE 802.11g Recommendation

IEEE 802.11a and b standards differ in regard to speed and frequency of operation. In addition they are incompatible with each other. The IEEE 802.11 Committee is developing a new standard, IEEE 802.11g. It is aiming to produce higher data rates at 2.4 GHz and be backward compatible with IEEE 802.11a and b. The attraction of 2.4 GHz operation is to enable widespread use in Asia and Europe and so become a universal standard available throughout the world.

If IEEE 802.11g offers comparable speed with that of IEEE 802.11a, the latter may become redundant – not least since if 2.4 GHz operation could satisfy every application, manufacture of radio components around a single frequency band would lead to larger scale manufacture and an attendant reduction in cost. Backward compatibility from IEEE 802.11g means that existing systems may still be used during migration to the newer standard. However, transmission would be limited to the maximum data rate of IEEE 802.11b, which is 11 Mbps.

4.6.5 IEEE 802.11 Medium Access Control Protocol

The IEEE 802.11 Committee recommended a new protocol, the Distributed Foundation Wireless MAC (DFWMAC) protocol, for the control of station access to the physical layer,

or medium, for use with WLANs. This protocol is used with all of the variants described above. The protocol has two elements: **Distributed Coordination Function (DCF)** and **Point Coordination Function (PCF)**, both illustrated in Fig. 4.33. DCF distributes the decision to transmit to each station and makes use of carrier sensing, is contention based and suitable for transmission of asynchronous frames. Optionally PCF, which uses centralized control where access is controlled or coordinated from a single point, may be overlaid upon DCF to provide contention free access.

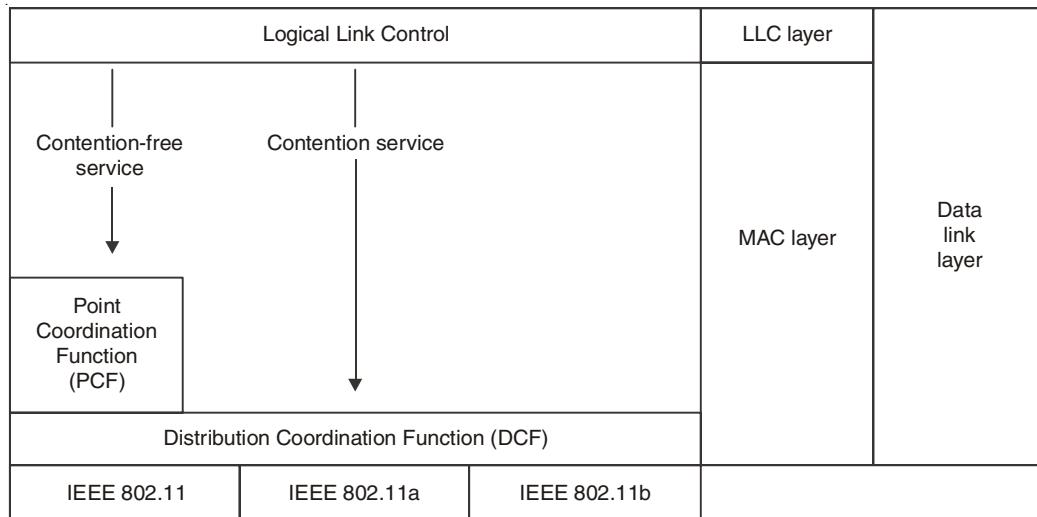


Fig. 4.33. IEEE 802.11 MAC Protocol.

➤ Distributed Coordination Function

Radio systems which transmit and receive on a single frequency cannot do so simultaneously. Therefore, WLANs must operate in half-duplex mode and while a station is transmitting, it is unable to monitor reception as the relatively strong transmitted signal would mask out any weak received signal. As a result, the Collision Detection aspect of the Ethernet CSMA/CD protocol is not possible with WLANs. Instead, the DCF sublayer uses CSMA with Collision Avoidance (CSMA/CA). Collision Avoidance seeks to arrange that transmitters only transmit at optimal times when it is unlikely that other stations may transmit and hence avoid, or at least minimize, the risk of a collision.

When a station wishes to transmit, the CSMA aspect of the protocol is as follows:

1. Station performs CSMA. If free, it waits out a delay known as the **Interframe Space (IFS)** delay and, if the medium is still free, transmits.
2. If the medium is busy (initially when CSMA is performed, or because another station transmits during the IFS delay) station defers using 1-persistence; that is, monitor until medium becomes free. When the medium becomes free, station then again waits out IFS delay and if the medium is still free backs off using the binary exponential back-off. The station re-examines the medium to see if it is still free and, if so, transmits.

3. If in 2 above, another station should gain access to the medium while a station is waiting out the IFS delay, the station again defers using 1-persistence and waits for the medium to go free.

The reason that the IFS delay is invoked when the medium is, or becomes, free before a frame may be transmitted is that it prevents a station sending frames continuously and so ensures fair-share operation. That is, once the medium is, or if necessary forced to be, idle by means of an IFS, any station may then contend for transmission of the next frame.

Two or more stations waiting for the medium to go free are unlikely to collide. This is because, when the medium does go free, each waiting station only transmits after waiting out a random back-off interval. Therefore one station, almost inevitably, will transmit before any other. Any other waiting stations, when their back-offs have elapsed, then find that the (first) station's transmissions are under way and so do not transmit. It is this feature that provides the collision avoidance feature of the protocol. IFS delay has in fact one of four possible values enabling support of priority based access. The delay values, shortest delay first, are termed Short IFS, Point Coordination Function IFS, Distributed Coordination IFS and Extended IFS, abbreviated SIFS, PIFS, DIFS and EIFS, respectively. Normal transmission of asynchronous MAC data frames as described above use a relatively large delay, DIFS, and therefore have a low priority. SIFS, the smallest delay, offers the greatest priority and is used for immediate response actions. EIFS is invoked by a station indicating it has received an erroneous, or unknown, frame. Under such circumstances priority is in effect set low. The intermediate length delay, PIFS, is used by the centralized controller in PCF operation and described shortly.

An example of the use of SIFs is in the transmission of an LLC PDU. Commonly a number of MAC frames will form a single PDU. Each PDU MAC frame is acknowledged with an ACK frame, which uses an SIFS delay. Upon receipt of the ACK the sending station of the LLC MAC frame responds immediately, after waiting out the SIFS delay, with the next MAC frame in the PDU. The use of SIFS in this way assigns higher priority to an LLC PDU enabling the LLC layer to maintain control of the medium until its transmission is completed.

If a sending station in a wired LAN appears to have succeeded in transmitting a frame onto the medium the station may reasonably assume that the frame has almost certainly reached the destination. In a wireless LAN such an assumption is less sound. A WLAN station may be confident of successfully transmitting a frame but whether or not a receiving station receives the frame depends on whether or not the two stations are within radio (or IR) contact of each other. In consequence, lost and corrupted frames due to the poorer quality transmission medium are far more frequent in WLANs.

In general such frames are either dealt with by the LLC layer at layer 2 or at layer 4, the transport layer. In Ethernet the LLC layer generally operates in unacknowledged connectionless mode, which does not afford any reliability. Therefore, reliability must be implemented at the transport layer. However, appreciable time may elapse before lost or corrupted frames are recovered from this layer. IEEE 802.11 has included a more efficient reliable transfer mechanism at the MAC layer. A handshake is used which ensures that stations are logically connected prior to transmission of data. A sending station with

data to send transmits a Request to Send (RTS) frame. Communication may then take place provided a Clear to Send (CTS) frame is returned by the receiving station. Each MAC data frame transmitted must be individually acknowledged.

If an ACK becomes corrupted, or lost, a retransmission occurs. The use of RTS/CTS and ACKs improves efficiency, medium usage and throughput. The handshake arrangement is shown in Fig. 4.34.

The above mechanism may optionally be extended further to enhance the collision avoidance feature of the MAC protocol. Other stations may monitor for RTS/CTS frames. When they have a frame to send they can allow sufficient time to elapse to ensure that the current RTS/CTS and data/ACK sequence completes before attempting transmission.

➤ Point Coordination Function

PCF operation can override DCF's contention-based transmission of asynchronous traffic by means of a centralized master, or **point coordinator**, to poll a station. This approach contrasts with that of DCF where access control is fully distributed. Polling and responding stations use an SIF delay. Since this delay is shorter than that used in DCF, PCF polling effectively locks out normal DCF frame transmission. The use of a round-robin approach with PIF means that operation is effectively contention free.

An example of the use of PCF to effect priority is where a number of stations have time-sensitive frames to transmit. Polling in round-robin fashion ensures equal access to the medium. To guard against the point controller hogging the medium by locking out asynchronous, or DCF, traffic indefinitely, a period of time called a **superframe** is defined, Fig. 4.35. Super frames are contiguous. During the contention-free period of the superframe the point-controller polls certain stations, each of which make their response. The length of the superframe is so chosen that after one cycle of poll/response activity time remains before the next superframe commences during which contention-based DCF operation may take place. Should asynchronous traffic be using the medium at the end of, and into the next, superframe, the point-controller must defer until it is able to seize control of the medium upon cessation of DCF activity.

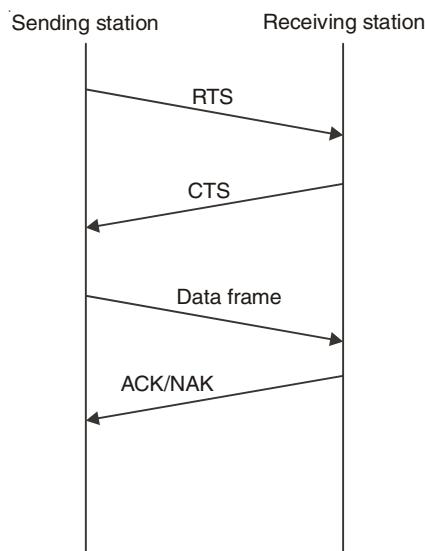


Fig. 4.34. Hand shake.

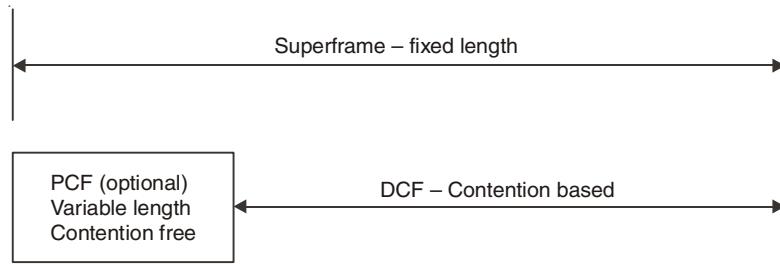


Fig. 4.35. Super Frame Structure.

PCF operation then commences but, since superframes are of fixed duration, the remaining time available for PCF activity is correspondingly reduced. During PCF operation, contention-based DCF operation is suspended in favour of round-robin polling. DCF operation is suitable for networks.

Stations are having peer-to-peer relationship (*e.g., ad hoc* networks) and no particular priority is required. If network traffic is predominantly bursty, PCF is less suitable. Bursty traffic may arise because only some stations are active, some stations only transmit irregularly, or a combination of the two. Hence PCF could waste much time, and therefore network capacity, by needlessly polling stations that have no traffic to send. PCF is most suited to time-critical and real-time frame transmission, and where frames may be reliably transmitted within a fairly narrow time window.

➤ MAC frame

The MAC frame structure is shown in Fig. 4.36. The FC field indicates whether it is a control, management or data frame and also has a number of other functions. Address fields include indication of source, destination, transmitting station and receiving station. The sequence control (SC) field is used to provide sequence numbers and caters for segmentation of frames. The frame body field carries the payload and contains either an LLC-PDU or MAC control information.

FC (2)	D/I (2)	Address (6)	Address (6)	SC (2)	Address (6)	Frame body (0 to 2312)	CRC (4)
-----------	------------	----------------	----------------	-----------	----------------	---------------------------	------------

(#) bytes
 FC Frame control
 D/I Duration/connection ID
 SC Sequence control
 CRC Cyclic redundancy check

Fig. 4.36. MAC Frame structure.

4.6.6 IEEE 802.11 Services

IEEE 802.11 offers nine services. These are provided to ensure that a WLAN offers an equivalent degree of functionality to that of a wired LAN. A WLAN may be a single entity and may, or may not, be connected to a wired LAN or backbone via an access point. In such circumstances it is analogous to that of a single, isolated, cell in a mobile radio communications

system such as GSM and has a single point coordinator. Alternatively two, or more, separate WLANs may be interconnected via an intermediate network called a **distribution system**.

There are five services which relate to distribution systems. They are called **distribution services** and are concerned with managing station membership within a cell and also interaction between stations in different cells interconnected by a distribution system. Four other services, **station services**, are concerned with a single cell only. These services relate to station mobility within a cell; that is, stations joining and leaving the cell. Table 4.6 illustrates the nine services, where they are provided and a brief outline of their purpose.

Table 4.6 IEEE 802.11 Services

	<i>Service</i>	<i>Provider</i>	<i>Purpose</i>
Distribution service	Association	DS	For a new station to associate itself to a cell
	Disassociation	DS	For a station to disassociate itself from a cell before leaving, or for a base station to alert stations it is going out of service
	Reassociation	DS	Used by a station moving from one cell to another
	Distribution	DS	Used to invoke routing (if required): frames destined for a station in another cell are transmitted to an AP station rather than intra-cell frames which are simply 'broadcast'
	Integration	DS	Used to transmit frames through an intermediate non-802.11 network. Deals with address translation, frame formatting, etc.
Station service	Authentication	Station	To enhance security, stations must be authenticated to ensure bogus stations do not attempt to join a cell
	Deauthentication	Station	Used when a station leaves a network and ensures it may not rejoin without reactivating authentication
	Privacy	Station	Manages encryption procedures for secure transmission
	Data delivery	Station	Service which provides data transmission

4.7 IEEE 802.16 BROADBAND WIRELESS

Many people in the industry realized that having a broadband wireless standard was the key element missing, so IEEE was asked to form a committee composed of people from key companies and academia to draw up the standard. The next number available in the 802 numbering space was **802.16**, so the standard got this number. Work was started in July 1999, and the final standard was approved in April 2002. Officially the standard is called "Air Interface for Fixed Broadband Wireless Access Systems." However, some people prefer

to call it a **wireless MAN (Metropolitan Area Network)** or a **wireless local loop**. We regard all these terms as interchangeable.

Like some of the other 802 standards, 802.16 was heavily influenced by the OSI model, including the (sub)layers, terminology, service primitives, and more. Unfortunately, also like OSI, it is fairly complicated. In the following sections we will give a brief description of some of the highlights of 802.16, but this treatment is far from complete and leaves out many details.

Why a new standard? There are some very good reasons for not using 802.11, primarily because 802.11 and 802.16 solve different problems. Before getting into the technology of 802.16, it is probably worthwhile saying a few words about why a new standard is needed at all.

The environments in which 802.11 and 802.16 operate are similar in some ways, primarily in that they were designed to provide high-bandwidth wireless communications. But they also differ in some major ways. To start with, 802.16 provides service to buildings, and buildings are not mobile. They do not migrate from cell to cell often. Much of 802.11 deals with mobility, and none of that is relevant here. Next, buildings can have more than one computer in them, a complication that does not occur when the end station is a single notebook computer. Because building owners are generally willing to spend much more money for communication gear than are notebook owners, better radios are available. This difference means that 802.16 can use full-duplex communication, something 802.11 avoids to keep the cost of the radios low.

Because 802.16 runs over part of a city, the distances involved can be several kilometers, which means that the perceived power at the base station can vary widely from station to station. This variation affects the signal-to-noise ratio, which, in, turn, dictates multiple modulation schemes. Also, open communication over a city means that security and privacy are essential and mandatory.

Furthermore, each cell is likely to have many more users than will a typical 802.11 cell, and these users are expected to use more bandwidth than will a typical 802.11 user. After all it is rare for a company to invite 50 employees to show up in a room with their laptops to see if they can saturate the 802.11 wireless network by watching 50 separate movies at once. For this reason, more spectrum is needed than the ISM bands can provide, forcing 802.16 to operate in the much higher 10-to-66 GHz frequency range, the only place unused spectrum is still available.

But these millimeter waves have different physical properties than the longer waves in the ISM bands, which in turn requires a completely different physical layer. One property that millimeter waves have is that they are strongly absorbed by water (especially rain, but to some extent also by snow, hail, and with a bit of bad luck, heavy fog). Consequently, error handling is more important than in an indoor environment. Millimeter waves can be focused into directional beams (802.11 is omnidirectional), so choices made in 802.11 relating to multipath propagation are moot here.

Another issue is quality of service. While 802.11 provides some support for real-time traffic (using PCF mode), it was not really designed for telephony and heavy-duty multimedia

usage. In contrast, 802.16 is expected to support these applications completely because it is intended for residential as well as business use.

In short, 802.11 was designed to be mobile Ethernet, whereas 802.16 was designed to be wireless, but stationary, cable television. These differences are so big that the resulting standards are very different as they try to optimize different things.

4.7.1 IEEE 802.16 Protocol Stack

The 802.16 protocol stack is illustrated in Fig. 4.37. The general structure is similar to that of the other 802 networks, but with more sublayers. The bottom sublayer deals with transmission. Traditional narrow-band radio is used with conventional modulation schemes. Above the physical transmission layer comes a convergence sublayer to hide the different technologies from the data link layer. Actually, 802.11 has something like this too, only the committee chose not to formalize it with an OSI-type name.

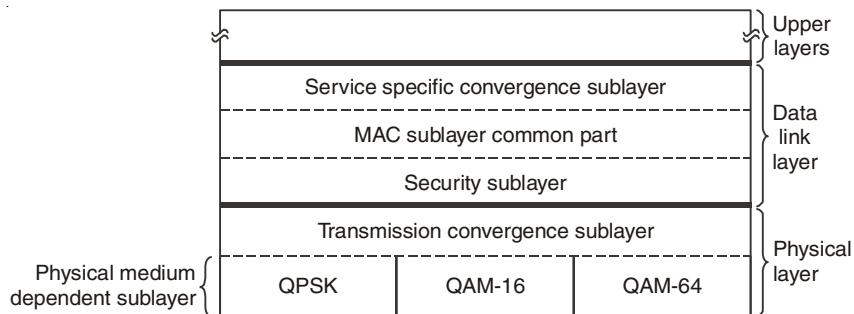


Fig. 4.37. IEEE 802.16 Protocol Stack.

Although we have not shown them in the figure, work is already underway to add two new physical layer protocols. The 802.16a standard will support OFDM in the 2-to-11 GHz frequency range. The 802.16b standard will operate in the 5-GHz ISM band. Both of these are attempts to move closer to 802.11.

The data link layer consists of three sublayers. The bottom one deals with privacy and security, which is far more crucial for public outdoor networks than for private indoor networks. It manages encryption, decryption, and key management.

Next comes the MAC sublayer common part. This is where the main protocols, such as channel management, are located. The model is that the base station controls the system. It can schedule the downstream (*i.e.*, base to subscriber) channels very efficiently and plays a major role in managing the upstream (*i.e.*, subscriber to base) channels as well. An unusual feature of the MAC sublayer is that, unlike those of the other 802 networks, it is completely connection oriented, in order to provide quality-of-service guarantees for telephony and multimedia communication. The service-specific convergence sublayer takes the place of the logical link sublayer in the other 802 protocols. Its function is to interface to the network layer. A complication here is that 802.16 was designed to integrate seamlessly with both datagram protocols (*e.g.*, PPP, IP, and Ethernet) and ATM. The problem is that packet protocols are connectionless and ATM is connection oriented. This means that every ATM

connection has to map onto an 802.16 connection, in principle a straightforward matter. But onto which 802.16 connection should an incoming IP packet be mapped? That problem is dealt with in this sublayer.

4.7.2 The 802.16 Frame Structure

All MAC frames begin with a generic header. The header is followed by an optional payload and an optional checksum (CRC), as illustrated in Fig. 4.38. The payload is not needed in control frames, for example, those requesting channel slots. The checksum is (surprisingly) also optional due to the error correction in the physical layer and the fact that no attempt is ever made to retransmit real-time frames. If no retransmissions will be attempted, why even bother with a checksum?

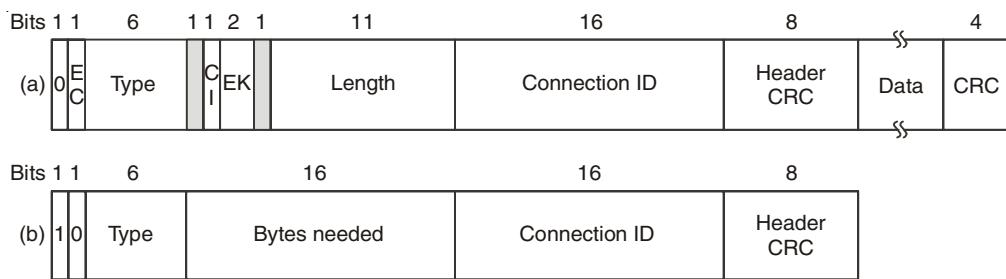


Fig. 4.38. (a) A generic frame (b) Bandwidth request frame.

The *EC* bit tells whether the payload is encrypted. The *Type* field identifies the frame type, mostly telling whether packing and fragmentation are present. The *CI* field indicates the presence or absence of the final checksum. The *EK* field tells which of the encryption keys is being used (if any). The *Length* field gives the complete length of the frame, including the header. The *Connection identifier* tells which connection this frame belongs to. Finally, the *Header CRC* field is a checksum over the header only, using the polynomial $x^8 + x^2 + x + 1$.

A second header type, for frames that request bandwidth, is shown in Fig. 4.38 (b). It starts with a 1 bit instead of a 0 bit and is similar to the generic header except that the second and third bytes form a 16-bit number telling how much bandwidth is needed to carry the specified number of bytes. Bandwidth request frames do not carry a payload or full-frame CRC.

4.8 BLUETOOTH

In 1994, the L. M. Ericsson company became interested in connecting its mobile phones to other devices (*e.g.*, PDAs) without cables. Together with four other companies (IBM, Intel, Nokia, and Toshiba), it formed a SIG (Special Interest Group, *i.e.*, consortium) to develop a wireless standard for interconnecting computing and communication devices and accessories using short-range, low-power, inexpensive wireless radios. The project was named **Bluetooth**, after Harald Blaatand (Bluetooth) II (940-981), a Viking king who unified (*i.e.*, conquered) Denmark and Norway, also without cables.

Bluetooth is a wireless LAN technology designed to connect devices of different functions such as telephones, notebooks, computers (desktop and laptop), cameras, printers, coffee makers, and so on. A Bluetooth LAN is an *ad hoc* network, which means that the network is formed spontaneously; the devices, sometimes called gadgets, find each other and make a network called a piconet. A Bluetooth LAN can even be connected to the Internet if one of the gadgets has this capability. A Bluetooth LAN, by nature, cannot be large. If there are many gadgets that try to connect, there is chaos.

Even though IEEE approved the first PAN standard, 802.15.1, in 2002, the Bluetooth SIG is still active busy with improvements. Although the Bluetooth SIG and IEEE versions are not identical, it is hoped that they will soon converge to a single standard.

4.8.1 Bluetooth Architecture

Bluetooth defines two types of networks: piconet and scatternet.

➤ Piconets

A Bluetooth network is called a piconet, or a small net. A piconet can have up to eight stations, one of which is called the primary station; the rest are called secondaries. All the secondary stations synchronize their clocks and hopping sequence with the primary. Note that a piconet can have only one primary station. The communication between the primary and the secondary can be one-to-one or one-to-many. Figure 4.39 shows a piconet.

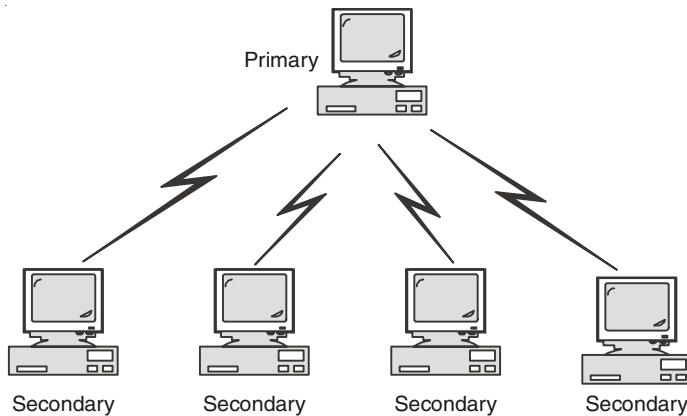


Fig. 4.39. Bluetooth Piconet.

Although a piconet can have a maximum of seven secondaries, an additional eight secondaries can be in the *parked state*. A secondary in a parked state is synchronized with the primary, but cannot take part in communication until it is moved from the parked state. Because only eight stations can be active in a piconet, activating a station from the parked state means that an active station must go to the parked state. **Primary/secondary terminology may be represented as master/slave terminology.**

➤ Scatternet

Piconets can be combined to form what is called a scatternet. A secondary station in one piconet can be the primary in another piconet. This station can receive messages from

the primary in the first piconet (as a secondary) and, acting as a primary, deliver them to secondaries in the second piconet. A station can be a member of two piconets.

4.8.2. Physical Links

Two types of links can be created between a primary and a secondary: SCQ links and ACL links.

➤ SCQ

A synchronous connection-oriented (SCQ) link is used when avoiding latency (delay in data delivery) is more important than integrity (error-free delivery). In an SCQ link, a physical link is created between the primary and a secondary by reserving specific slots at regular intervals. The basic unit of connection is two slots, one for each direction. If a packet is damaged, it is never retransmitted. SCQ is used for real-time audio where avoiding delay is all-important. A secondary can create up to three SCQ links with the primary, sending digitized audio (PCM) at 64 kbps in each link.

➤ ACL

An asynchronous connectionless link (ACL) is used when data integrity is more important than avoiding latency. In this type of link, if a payload encapsulated in the frame is corrupted, it is retransmitted. A secondary returns an ACL frame in the available odd-numbered slot if and only if the previous slot has been addressed to it. ACL can use one, three, or more slots and can achieve a maximum data rate of 721 kbps.

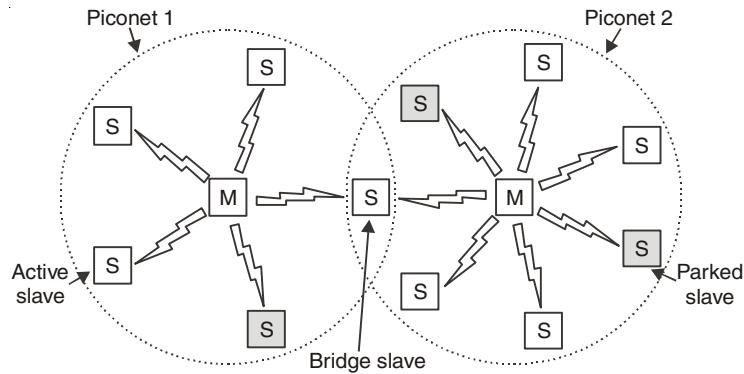


Fig. 4.40. Illustration of a scatternet.

4.8.3 The Bluetooth Protocol Stack

The Bluetooth standard has many protocols grouped loosely into layers. The layer structure does not follow the OSI model, the TCP/IP model, the 802 model, or any other known model. However, IEEE is working on modifying Bluetooth to shoehorn it into the 802 model better. The basic Bluetooth protocol architecture as modified by the 802 committee is shown in Fig. 4.41.

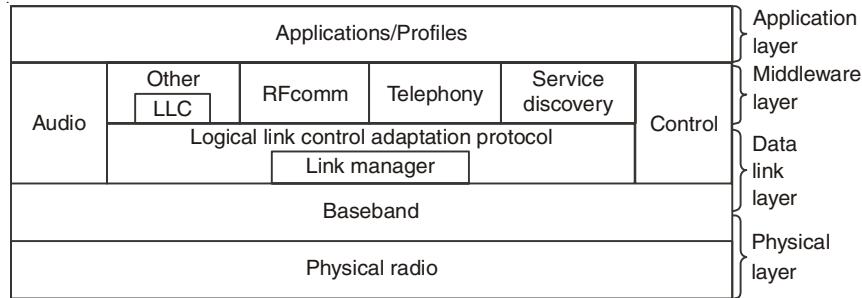


Fig. 4.41. IEEE 802.15 Protocol Architecture.

The bottom layer is the physical radio layer, which corresponds fairly well to the physical layer in the OSI and 802 models. It deals with radio transmission and modulation. Many of the concerns here have to do with the goal of making the system inexpensive so that it can become a mass market item.

The baseband layer is somewhat analogous to the MAC sublayer but also includes elements of the physical layer. It deals with how the master controls time slots and how these slots are grouped into frames.

Next comes a layer with a group of somewhat related protocols. The link manager handles the establishment of logical channels between devices, including power management, authentication, and quality of service. The logical link control adaptation protocol (often called L2CAP) shields the upper layers from the details of transmission. It is analogous to the standard 802 LLC sublayer, but technically different from it. As the names suggest, the audio and control protocols deal with audio and control, respectively. The applications can get at them directly, without having to go through the L2CAP protocol.

The next layer up is the middleware layer, which contains a mix of different protocols. The 802 LLC was inserted here by IEEE for compatibility with its other 802 networks. The RFcomm, telephony, and service discovery protocols are native. RFcomm (Radio Frequency Communication) is the protocol that emulates the standard serial port found on PCs for connecting the keyboard, mouse, and modem, among other devices. It has been designed to allow legacy devices to use it easily. The telephony protocol is a real-time protocol used for the three speech-oriented profiles. It also manages call setup and termination. Finally, the service discovery protocol is used to locate services within the network.

The top layer is where the applications and profiles are located. They make use of the protocols in lower layers to get their work done. Each application has its own dedicated subset of the protocols. Specific devices, such as a headset, usually contain only those protocols needed by that application and no others.

4.8.4 The Bluetooth Frame Structure

There are several frame formats, the most important of which is shown in Fig. 4.42. It begins with an access code that usually identifies the master so that slaves within radio range of two masters can tell which traffic is for them. Next comes a 54 bit header containing typical MAC sublayer fields. Then comes the data field, of up to 2744 bits (for a five-slot transmission). For a single time slot, the format is the same except that the data field is 240 bits.

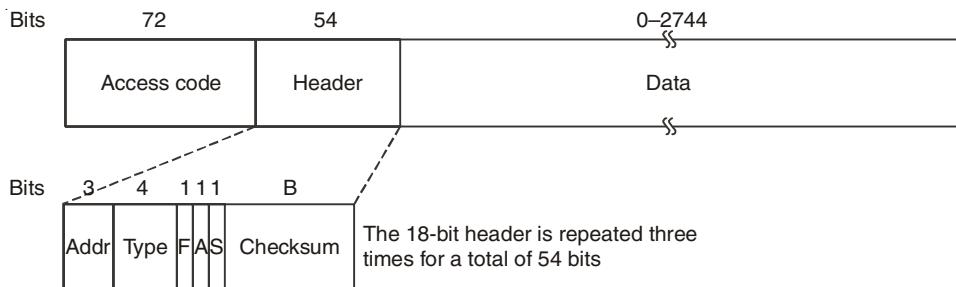


Fig. 4.42. Bluetooth Frame structure.

Let us take a quick look at the header. The *Address* field identifies which of the eight active devices the frame is intended for. The *Type* field identifies the frame type (ACL, SCO, poll, or null), the type of error correction used in the data field, and how many slots long the frame is. The *Flow* bit is asserted by a slave when its buffer is full and cannot receive any more data. This is a primitive form of flow control. The *Acknowledgement* bit is used to piggyback an ACK onto a frame. The *Sequence* bit is used to number the frames to detect retransmissions. The protocol is stop-and-wait, so 1 bit is enough. Then comes the 8-bit header *Checksum*. The entire 18 bit header is repeated three times to form the 54 bit header shown in Fig. 4.42. On the receiving side, a simple circuit examines all three copies of each bit. If all three are the same, the bit is accepted. If not, the majority opinion wins. Thus, 54 bits of transmission capacity are used to send 10 bits of header. The reason is that to reliably send data in a noisy environment using cheap, low-powered (2.5 mW) devices with little computing capacity, a great deal of redundancy is needed.

Various formats are used for the data field for ACL frames. The SCO frames are simpler though: the data field is always 240 bits. Three variants are defined, permitting 80, 160, or 240 bits of actual payload, with the rest being used for error correction. In the most reliable version (80-bit payload), the contents are just repeated three times, the same as the header.

Since the slave may use only the odd slots, it gets 800 slots/sec, just as the master does. With an 80 bit payload, the channel capacity from the slave is 64,000 bps and the channel capacity from the master is also 64,000 bps, exactly enough for a single full-duplex PCM voice channel (which is why a hop rate of 1600 hops/sec was chosen). These numbers mean that a full-duplex voice channel with 64,000 bps in each direction using the most reliable format completely saturates the piconet despite a raw bandwidth of 1 Mbps. For the least reliable variant (240 bits/slot with no redundancy at this level), three full-duplex voice channels can be supported at once, which is why a maximum of three SCO links is permitted per slave. A Bluetooth device has a built-in short-range radio transmitter. The current data rate is 1 Mbps with a 2.4-GHz bandwidth. This means that there is a possibility of interference between the IEEE 802.11b wireless LANs and Bluetooth LANs.

4.9 CONNECTING DEVICES

LANs do not normally operate in isolation. They are connected to one another or to the Internet. To connect LANs, or segments of LANs, we use connecting devices. Connecting devices can operate in different layers of the Internet model.

The five categories contain devices which can be defined as:

1. Those which operate below the physical layer such as a passive hub.
2. Those which operate at the physical layer (a repeater or an active hub).
3. Those which operate at the physical and data link layers (a bridge or a two-layer switch).
4. Those which operate at the physical, data link, and network layers (a router or a three-layer switch).
5. Those which can operate at all five layers (a gateway).

4.9.1 Passive Hubs

A passive hub is just a connector. It connects the wires coming from different branches.

In a star-topology Ethernet LAN, a passive hub is just a point where the signals coming from different stations collide; the hub is the collision point. This type of a hub is part of the media; its location in the Internet model is below the physical layer.

4.9.2 Repeaters

A repeater is a device that operates only in the physical layer. Signals that carry information within a network can travel a fixed distance before attenuation endangers the integrity of the data. A repeater receives a signal and, before it becomes too weak or corrupted, regenerates the original bit pattern. The repeater then sends the refreshed signal. A repeater can extend the physical length of a LAN, as shown in Fig. 4.43.

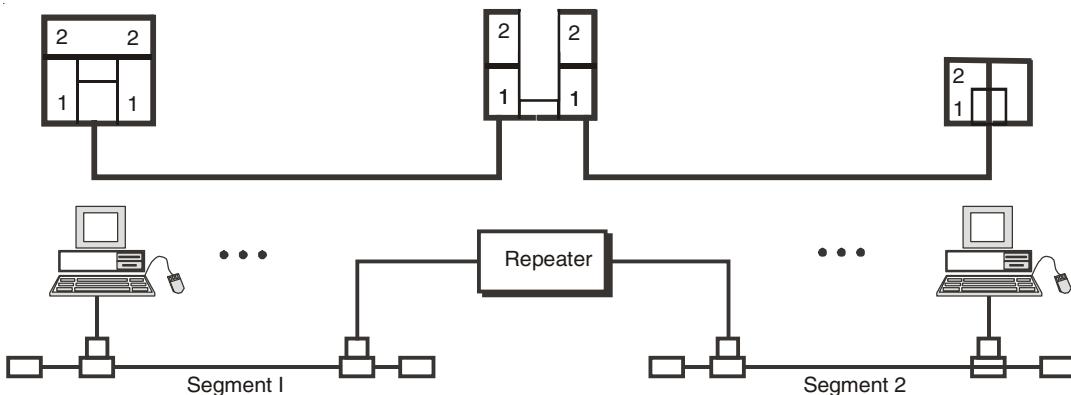


Fig. 4.43. Repeater connecting two LAN's.

It is tempting to compare a repeater to an amplifier, but the comparison is inaccurate. An amplifier cannot discriminate between the intended signal and noise; it amplifies equally

everything fed into it. A repeater does not amplify the signal; it regenerates the signal. When it receives a weakened or corrupted signal, it creates a copy, bit for bit, at the original strength. The location of a repeater on a link is vital. A repeater must be placed so that a signal reaches it before any noise changes the meaning of any of its bits. A little noise can alter the precision of a bit's voltage without destroying its identity. If the corrupted bit travels much farther, however, accumulated noise can change its meaning completely. At that point, the original voltage is not recoverable, and the error needs to be corrected. A repeater placed on the line before the legibility of the signal becomes lost can still read the signal well enough to determine the intended voltages and replicate them in their original form.

4.9.3 Active Hubs

An active hub is actually a multipart repeater. It is normally used to create connections between stations in a physical star topology. We have seen examples of hubs in some Ethernet implementations (10Base-T, for example). However, hubs can also be used to create multiple levels of hierarchy, as shown in Fig. 4.44. The hierarchical use of hubs removes the length limitation of 10Base-T (100 m).

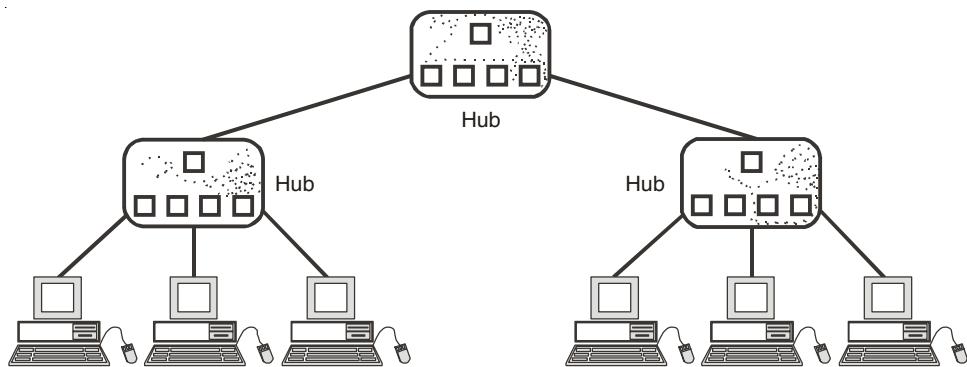


Fig. 4.44. Hierarchy of Hubs.

4.9.4 Bridges

A bridge operates in both the physical and the data link layer. As a physical layer device, it regenerates the signal it receives. As a data link layer device, the bridge can check the physical (MAC) addresses (source and destination) contained in the frame. A bridge has filtering capability. It can check the destination address of a frame and decide if the frame should be forwarded or dropped. If the frame is to be forwarded, the decision must specify the port. A bridge has a table that maps addresses to ports.

➤ Transparent Bridges

A transparent bridge is a bridge in which the stations are completely unaware of the bridge's existence. If a bridge is added or deleted from the system, reconfiguration of the stations is unnecessary. According to the IEEE 802.1 d specification, a system equipped with transparent bridges must meet three criteria:

1. Frames must be forwarded from one station to another.

2. The forwarding table is automatically made by learning frame movements in the network.
3. Loops in the system must be prevented.

➤ **Source routing Bridges**

In source routing, a sending station defines the bridges that the frame must visit. The addresses of these bridges are included in the frame. In other words, the frame contains not only the source and destination addresses, but also the addresses of all bridges to be visited.

The source gets these bridge addresses through the exchange of special frames with the destination prior to sending the data frame. Source routing bridges were designed by IEEE to be used with Token Ring LANs. These LANs are not very common today.

4.9.5 Two-Layer Switches

When we use the term *switch*, we must be careful because a switch can mean two different things. We must clarify the term by adding the level at which the device operates. We can have a two-layer switch or a three-layer switch. A **three-layer switch** is used at the network layer; it is a kind of router. The **two-layer switch** performs at the physical and data link layers. A two-layer switch is a bridge, a bridge with many ports and a design that allows better (faster) performance. A bridge with a few ports can connect a few LANs together. A bridge with many ports may be able to allocate a unique port to each station, with each station on its own independent entity. This means no competing traffic (no collision, as we saw in Ethernet).

A two-layer switch, as a bridge does, makes a filtering decision based on the MAC address of the frame it received. However, a two-layer switch can be more sophisticated. It can have a buffer to hold the frames for processing. It can have a switching factor that forwards the frames faster. Some new two-layer switches, called *cut-through* switches, have been designed to forward the frame as soon as they check the MAC addresses in the header of the frame.

Higher layer connecting devices will be discussed in subsequent chapters.

Review Questions

1. Discuss the differences between LANs and WANs with respect to their geographical area of coverage, data transmission and error rates, ownership, government regulation and data routing.
2. Describe and discuss five possible LAN applications.
3. Describe the advantages and disadvantages associated with five LAN topologies.
4. What is the difference between broadband and baseband signaling?
5. Discuss the rationale for using Manchester or Differential Manchester coding for LAN signaling.
6. Compare and contrast CSMA/CD and token passing access methods.

7. How can a collision occur on a CSMA/CD network?
8. What are the two types of coaxial cable used to construct Ethernet networks? What are the advantages and disadvantages associated with the use of each type of cable?
9. What is the purpose of a jam signal?
10. What is a 10BASE-5 network?
11. What is a 10BASE-2 network?
12. What is the purpose of a headend in a broadband network?
13. Discuss the three key components required to construct the physical infrastructure for a 10BASE-T network.
14. Describe the functions performed by a 10BASE-T hub.
15. Describe Ethernet signal limitations associated with the so-called '5-4-3' rule.
16. What is the difference between 100BASE-T4 and 100BASE-TX with respect to cable use and signaling method employed?
17. Describe two 100BASE-T applications.
18. What type of media does Gigabit Ethernet operate over?
19. What is a buffered distributor?
20. What is the difference between universally administered addressing and locally administered addressing?
21. Why must an Ethernet frame have a minimum length of 64 bytes in an adapter and 72 bytes when flowing on a LAN?
22. What is the purpose of an Ethernet-SNAP frame?
23. What is meant by the term connectionless transmission?
24. How can software determine a particular type of Ethernet frame as the frame flows on a network?
25. Why are carrier extensions required on certain types of Gigabit Ethernet frames when half-duplex shared media transmission is supported?
26. What is the purpose of a Jumbo frame?
27. Draw a two MAU Token-Ring network illustrating ring-in and ring-out connections.
28. How does a Token-Ring MAU provide a redundant ring path?
29. Why is the use of RJ-45 connectors more popular than the use of IBM cabling system data connectors for cabling Token-Ring network components?
30. What type of application is 100 Mbps Token-Ring designed to be used for?
31. What is the purpose of an abort token?
32. What are J and K symbols?
33. What is the purpose of the monitor bit? Who is responsible for monitoring the monitor bit?
34. How is a MAC frame distinguished from an LLC frame on a Token-Ring network?
35. How is OFDM different from FDM?
36. What is the access method used by wireless LANs?

37. What is the purpose of the NAV?
38. Compare a piconet and a scatternet.
39. Match the layers in Bluetooth and the Internet model.
40. What are the two types of links between a Bluetooth primary and a Bluetooth secondary?
41. In multiple-secondary communication, who uses the even-numbered slots and who uses the odd-numbered slots?
42. How much time in a Bluetooth one-slot frame is used for the hopping mechanism? What about a three-slot frame and a five-slot frame?



CHAPTER 5 *THE NETWORK LAYER*

Every seeming equality conceals a hierarchy.

—Mason Cooley

5.1 INTRODUCTION

The network layer is responsible for the source-to-destination delivery of a packet, possibly across multiple networks (links). Whereas the data link layer oversees the delivery of the packet between two systems on the same network (links), the network layer ensures that each packet gets from its point of origin to its final destination. The network layer adds a header that includes the logical addresses of the sender and receiver to the packet coming from the upper layer. If a packet travels through the Internet, we need this addressing system to help distinguish the source and destination. When independent networks or links are connected together to create an internetwork, routers or switches route packets to their final destination. One of the functions of the network layer is to provide a routing mechanism. How can you find an efficient path through a network with millions, or perhaps billions, of nodes? Closely related to this is the problem of *addressing*, the task of providing suitable identifiers for all those nodes. In this chapter we will study all these issues and illustrate them, primarily using the Internet and its network layer protocol, IP, although wireless networks will also be addressed.

5.2 SERVICES PROVIDED TO TRANSPORT LAYER

The network layer provides services to the transport layer at the network layer/transport layer interface. An important question is what kind of services the network layer provides to the transport layer. The network layer services have been designed with the following goals in mind.

1. The services should be independent of the router technology.
2. The transport layer should be shielded from the number, type, and topology of the routers present.
3. The network addresses made available to the transport layer should use a uniform numbering plan, even across LANs and WANs.

Given these goals, the designers of the network layer have a lot of freedom in writing detailed specifications of the services to be offered to the transport layer. This freedom often degenerates into a raging battle between two warring factions. The discussion centers on whether the network layer should provide connection-oriented service or connectionless service.

One camp (represented by the Internet community) argues that the routers' job is moving packets around and nothing else. In their view (based on 30 years of actual experience with a real, working computer network), the subnet is inherently unreliable, no matter how it is designed. Therefore, the hosts should accept the fact that the network is unreliable and do error control (*i.e.*, error detection and correction) and flow control themselves.

This viewpoint leads quickly to the conclusion that the network service should be connectionless, with primitives SEND PACKET and RECEIVE PACKET and little else. In particular, no packet ordering and flow control should be done, because the hosts are going to do that anyway, and there is usually little to be gained by doing it twice. Furthermore, each packet must carry the full destination address, because each packet sent is carried independently of its predecessors, if any.

The other camp (represented by the telephone companies) argues that the subnet should provide a reliable, connection-oriented service. They claim that 100 years of successful experience with the worldwide telephone system is an excellent guide. In this view, quality of service is the dominant factor, and without connections in the subnet, quality of service is very difficult to achieve, especially for real-time traffic such as voice and video.

These two camps are best exemplified by the Internet and ATM. The Internet offers connectionless network-layer service; ATM networks offer connection-oriented network-layer service. However, it is interesting to note that as quality-of-service guarantees are becoming more and more important, the Internet is evolving. In particular, it is starting to acquire properties normally associated with connection-oriented service, as we will see later.

5.3 NETWORK LAYER ISSUES: DELIVERY AND FORWARDING

Delivery refers to the way a packet is handled by the underlying networks under the control of the network layer. Forwarding refers to the way a packet is delivered to the next station.

The delivery of a packet to its final destination is accomplished by using two different methods of delivery, direct and indirect. In a direct delivery, the final destination of the packet is a host connected to the same physical network as the deliverer. Direct delivery occurs when the source and destination of the packet are located on the same physical network or when the delivery is between the last router and the destination host. If the

destination host is not on the same network as the deliverer, the packet is delivered indirectly. In an indirect delivery, the packet goes from router to router until it reaches the one connected to the same physical network as its final destination.

Forwarding means to place the packet in its route to its destination. Forwarding requires a host or a router to have a routing table. When a host has a packet to send or when a router has received a packet to be forwarded, it looks at this table to find the route to the final destination. However, this simple solution is impossible today in an internetwork such as the Internet because the number of entries needed in the routing table would make table lookups inefficient.

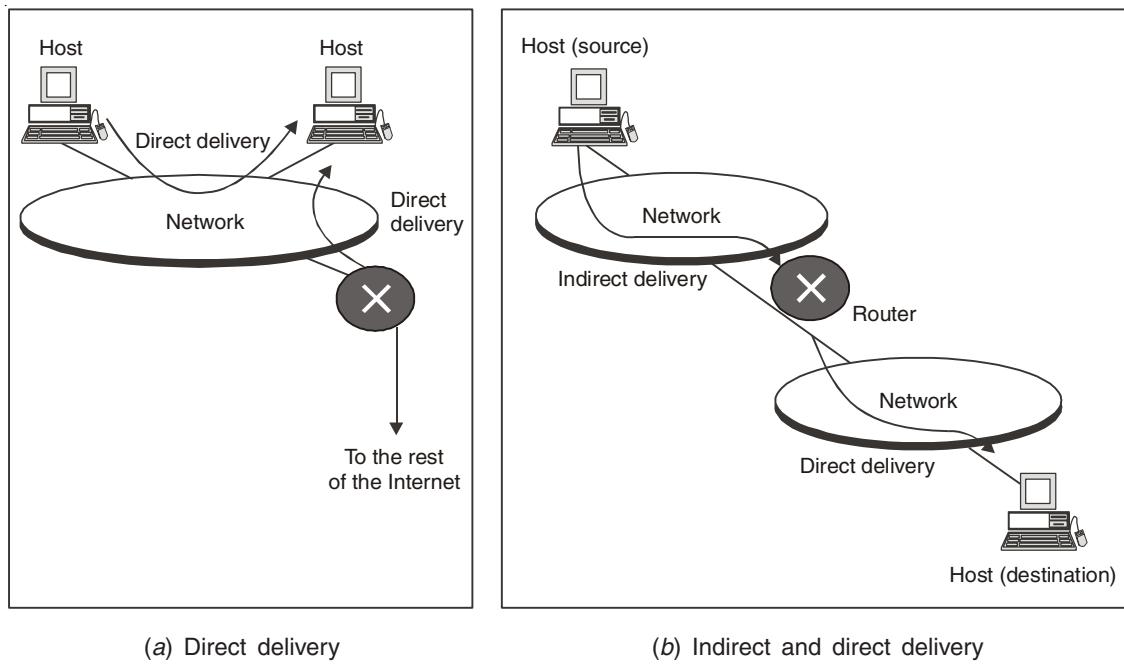
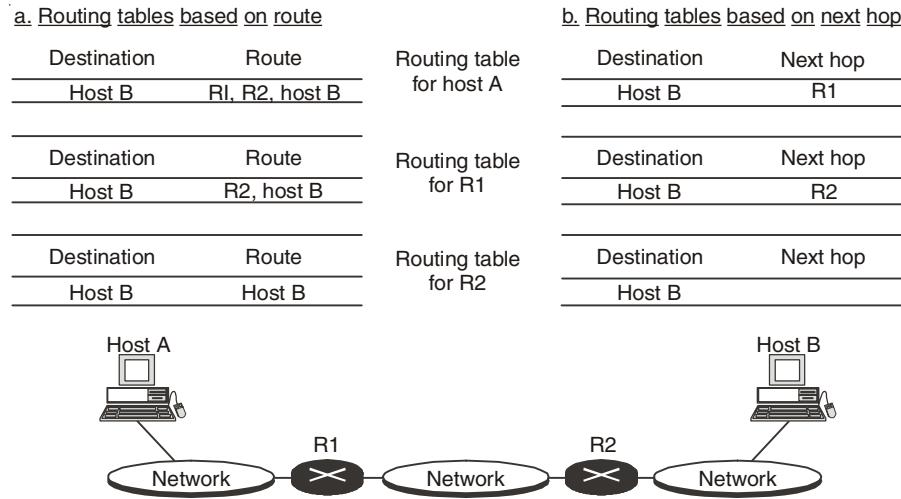


Fig. 5.1. Direct and Indirect delivery.

Several techniques can make the size of the routing table manageable and also handle issues such as security. We briefly discuss these methods here.

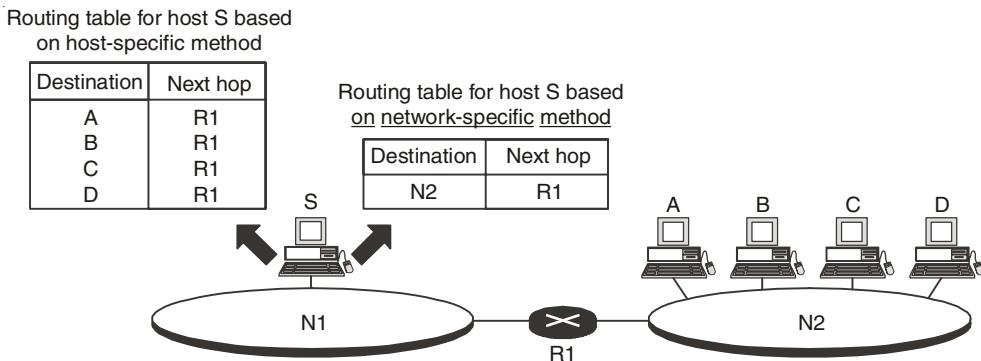
➤ Next-Hop Method Versus Route Method

One technique to reduce the contents of a routing table is called the next-hop method. In this technique, the routing table holds only the address of the next hop instead of information about the complete route (route method). The entries of a routing table must be consistent with one another. Figure 5.2 shows how routing tables can be simplified by using this technique.

**Fig. 5.2. Next-Hop Method Versus Route Method.**

➤ **Network-Specific Method Versus Host-Specific Method**

A second technique to reduce the routing table and simplify the searching process is called the network-specific method. Here, instead of having an entry for every destination host connected to the same physical network (host-specific method), we have only one entry that defines the address of the destination network itself. In other words, we treat all hosts connected to the same network as one single entity.

**Fig. 5.3. Host specific versus Network specific method.**

5.4 THE NETWORK LAYER IN INTERNET

Before getting into the specifics of the network layer in the Internet, it is worth taking at look at the principles that drove its design in the past and made it the success that it is today. All too often, nowadays, people seem to have forgotten them. These principles are enumerated and discussed in RFC 1958. Some of the principles are:

1. **Make sure it works.** Do not finalize the design or standard until multiple prototypes have successfully communicated with each other. All too often designers first write a 1000 page standard, get it approved, then discover it is deeply flawed and does not work. Then they write version 1.1 of the standard. This is not the way to go.
2. **Keep it simple.** When in doubt, use the simplest solution. William of Occam stated this principle (Occam's razor) in the 14th century. Put in modern terms: fight features. If a feature is not absolutely essential, leave it out, especially if the same effect can be achieved by combining other features.
3. **Make clear choices.** If there are several ways of doing the same thing, choose one. Having two or more ways to do the same thing is looking for trouble. Standards often have multiple options or modes or parameters because several powerful parties insist that their way is best. Designers should strongly resist this tendency. Just say no.
4. **Exploit modularity.** This principle leads directly to the idea of having protocol stacks, each of whose layers is independent of all the other ones. In this way, if circumstances that require one module or layer to be changed, the other ones will not be affected.
5. **Expect heterogeneity.** Different types of hardware, transmission facilities, and applications will occur on any large network. To handle them, the network design must be simple, general, and flexible.
6. **Avoid static options and parameters.** If parameters are unavoidable (*e.g.*, maximum packet size), it is best to have the sender and receiver negotiate a value than defining fixed choices.
7. **Look for a good design; it need not be perfect.** Often the designers have a good design but it cannot handle some weird special case. Rather than messing up the design, the designers should go with the good design and put the burden of working around it on the people with the strange requirements.
8. **Be strict when sending and tolerant when receiving.** In other words, only send packets that rigorously comply with the standards, but expect incoming packets that may not be fully conformant and try to deal with them.
9. **Think about scalability.** If the system is to handle millions of hosts and billions of users effectively, no centralized databases of any kind are tolerable and load must be spread as evenly as possible over the available resources.
10. **Consider performance and cost.** If a network has poor performance or outrageous costs, nobody will use it.

At the network layer, the Internet can be viewed as a collection of subnetworks or **Autonomous Systems (ASes)** that are interconnected. There is no real structure, but several major backbones exist. These are constructed from high-bandwidth lines and fast routers. A sketch of this quasi-hierarchical organization is given in Fig. 5.4.

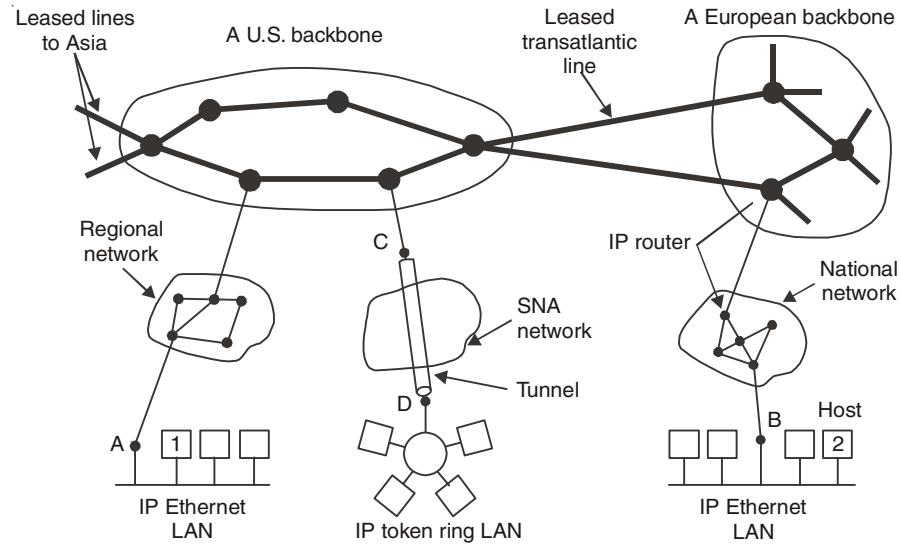


Fig. 5.4. Internet as collection of many networks.

The Internet Protocol is the key tool used today to build scalable, heterogeneous internetworks. It was originally known as the Kahn-Cerf protocol after its inventors. One way to think of IP is that it runs on all the nodes (both hosts and routers) in a collection of networks and defines the infrastructure that allows these nodes and networks to function as a single logical internetwork. For example, Fig. 5.6 shows how hosts H1 and H8 are logically connected by the internet in Fig. 5.5 including the protocol graph running on each node. Note that higher-level protocols, such as TCP and UDP, typically run on top of IP on the hosts.

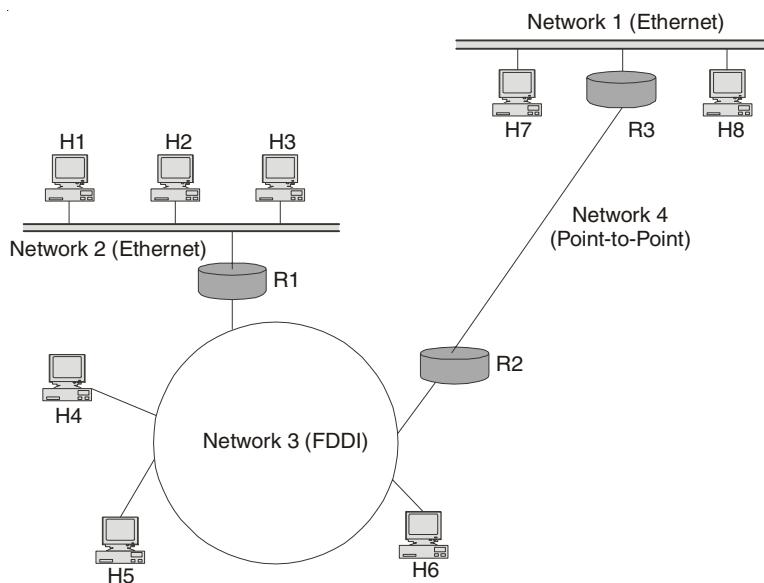


Fig. 5.5. A simple internetwork.

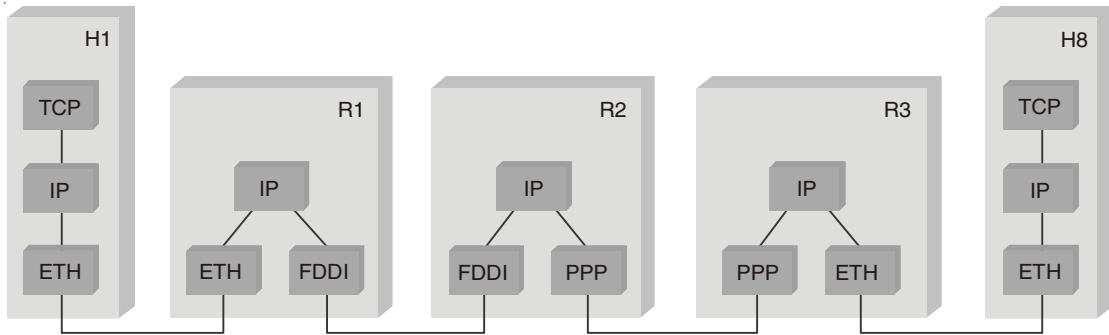


Fig. 5.6. A simple internetwork, showing the protocol layers used to connect H1 to H8 in Fig. 5.5

5.4.1 IP Protocol

The Internet Protocol (IP) is part of the TCP/IP protocol suite, and is the most widely-used internetworking protocol. It is functionally similar to the ISO standard connectionless network protocol (CLNP). As with any protocol standard, IP is specified in two parts:

- The interface with a higher layer (e.g., TCP), specifying the services that IP provides
- The actual protocol format and mechanisms

In this section, we first examine IP services and then the IP protocol. This is followed by a discussion of IP address formats.

➤ IP Services

IP provides two service primitives at the interface to the next-higher layer (Fig. 5.7). The Send primitive is used to request transmission of a data unit. The Deliver primitive is used by IP to notify a user of the arrival of a data unit. The parameters associated with the two primitives are:

Source address. Internetwork address of sending IP entity.

Destination address. Internetwork address of destination IP entity.

Protocol. Recipient protocol entity (an IP user).

Type of service indicators. Used to specify the treatment of the data unit in its transmission through component networks.

Identifier. Used in combination with the source and destination addresses and user protocol to identify the data unit uniquely. This parameter is needed for reassembly and error reporting.

Don't-fragment identifier. Indicates whether IP can segment (called fragment in the standard) data to accomplish delivery.

Time to live. Measured in network hops.

Data length. Length of data being transmitted.

Option data. Options requested by the IP user.

Data. User data to be transmitted.

Send (Deliver (
Source address	Source address
Destination address	Destination address
Protocol	Protocol
Type of service indicators	Type of service indicators
Identifier	
Don't-fragment identifier	
Time to live	
Data length	Data length
Option data	Option data
Data	Data
))

Fig. 5.7. IP Service Primitives and Parameters.

Note that the *identifier*, *don't-fragment identifier*, and *time-to-live* parameters are present in the Send primitive but not in the Deliver primitive. These three parameters provide instructions to IP that are not of concern to the recipient IP user. The sending IP user includes the *type-of-service* parameter to request a particular quality of service. The user may specify one or more of the services. This parameter can be used to guide routing decisions. For example, if a router has several alternative choices for the next hop in routing a datagram, it may choose a network of a higher data rate if the high throughput option has been selected. This parameter, if possible, is also passed down to the network access protocol for use over individual networks. For example, if a precedence level is selected, and if the subnetwork supports precedence or priority levels, the precedence level will be mapped onto the network level for this hop.

The options parameter allows for future extensibility and for inclusion of parameters that are usually not invoked. The currently defined options are:

1. **Security.** Allows a security label to be attached to a datagram.
2. **Source routing.** A sequenced list of router addresses that specifies the route to be followed. Routing may be strict (only identified routers may be visited) or loose (other intermediate routers may be visited).
3. **Route recording.** A field is allocated to record the sequence of routers visited by the datagram.
4. **Stream identification.** Names reserved resources used for stream service. This service provides special handling for volatile periodic traffic (*e.g.*, voice).
5. **Time stamping.** The source IP entity and some or all intermediate routers add a timestamp (precision to milliseconds) to the data unit as it goes by.

➤ IP Protocol

The protocol between IP entities is best described with reference to the IP datagram format, shown in Fig. 5.8. The fields are **Version (4 bits):** Indicates the version number, to allow evolution of the protocol.

Internet header length (IHL) (4 bits): Length of header in 32 bit words. The minimum value is five, for a minimum header length of 20 octets.

Type of service (8 bits). Specifies reliability, precedence, delay, and throughput parameters.

Total length (16 bits). Total datagram length, in octets.

Identifier (16 bits). A sequence number that, together with the source address, destination address, and user protocol, is intended to uniquely identify a datagram. Thus, the identifier should be unique for the datagram's source address, destination address, and user protocol for the time during which the datagram will remain in the internet.

Flags (3 bits). Only two of the bits are currently defined. The More bit is used for segmentation (fragmentation) and reassembly, as previously explained. The Don't-Fragment bit prohibits fragmentation when set. This bit may be useful if it is known that the destination does not have the capability to reassemble fragments. However, if this bit is set, the datagram will be discarded if it exceeds the maximum size of an en route subnet. Therefore, if the bit is set, it may be advisable to use source routing to avoid subnetworks with small maximum packet size.

Fragment offset (13 bits). Indicates where in the original datagram this fragment belongs, measured in 64 bit units, implying that fragments other than the last fragment must contain a data field that is a multiple of 64 bits.

Time to live (8 bits). Measured in router hops.

Protocol (8 bits). Indicates the next higher level protocol that is to receive the data field at the destination.

Header checksum (16 bits). An error-detecting code applied to the header only. Because some header fields may change during transit (e.g., time to live, segmentation-related fields), this is reverified and recomputed at each router. The checksum field is the 16 bit one's complement addition of all 16 bit words in the header. For purposes of computation, the checksum field is itself initialized to a value of zero.

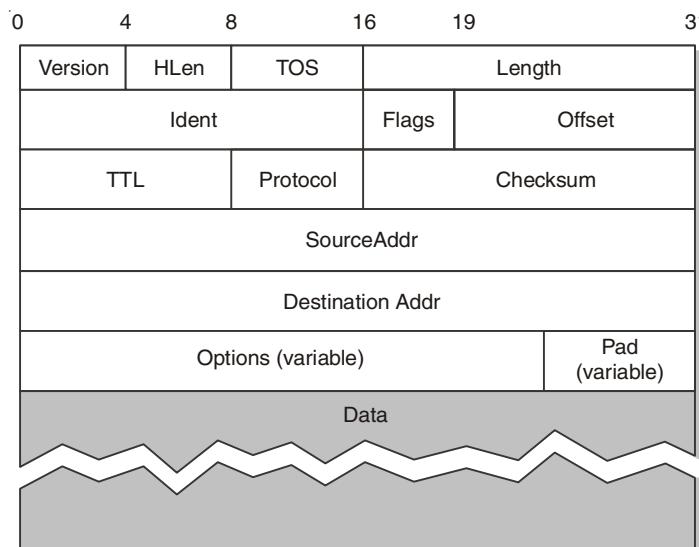


Fig. 5.8. IPv4 Packet format.

Source address (32 bits). Coded to allow a variable allocation of bits to specify the network and the end system attached to the specified network (7 and 24 bits, 14 and 16 bits, or 21 and 8 bits).

Destination address (32 bits). As above.

Options (variable). Encodes the options requested by the sending user.

Padding (variable). Used to ensure that the datagram header is a multiple of 32 bits.

Data (variable). The data field must be an integer multiple of 8 bits. The maximum length of the datagram (data field plus header) is 65,535 octets.

It should be clear how the IP services specified in the Send and Deliver primitives map into the fields of the IP datagram.

➤ IP Addresses

An IPv4 address is a 32 bit address that *uniquely* and *universally* defines the connection of a device (for example, a computer or a router) to the Internet. The source and destination address fields in the IP header each contain a 32 bit global internet address, generally consisting of a network identifier and a host identifier.

The address is coded to allow a variable allocation of bits to specify network and host, as depicted in Fig. 5.9. This encoding provides flexibility in assigning addresses to hosts and allows a mix of network sizes on an internet. In particular, the three network classes are best suited to the following conditions:

Class A. Few networks, each with many hosts.

Class B. Medium number of networks, each with a medium number of hosts.

Class C. Many networks, each with a few hosts.

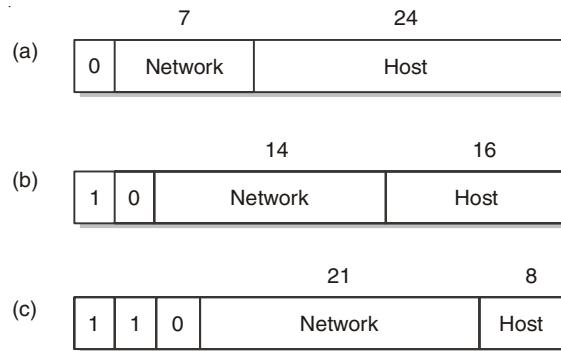


Fig. 5.9. IP Addresses: Class A, Class B, Class C.

The class of an IP address is identified in the most significant few bits. If the first bit is 0, it is a class A address. If the first bit is 1 and the second is 0, it is a class B address. If the first two bits are 1 and the third is 0, it is a class C address. Thus, of the approximately 4 billion possible IP addresses, half are class A, one quarter are class B, and one-eighth are class C. Each class allocates a certain number of bits for the network part of the address and the rest for the host part. Class A networks have 7 bits for the network part and 24 bits for the host part, meaning that there can be only 126 class A networks (the values 0 and 127

reserved), but each of them can accommodate up to $2^{24} - 2$ (about 16 million) hosts (again, there are two reserved values). Class B addresses allocate 14 bits for the network and 16 bits for the host, meaning that each class B network has room for 65,534 hosts. Finally, class C addresses have only 8 bits for the host and 21 for the network part. Therefore, a class C network can have only 256 unique host identifiers, which means only 254 attached hosts (one host identifier, 255, is reserved for broadcast, and 0 is not a valid host number). However, the addressing scheme supports 2^{21} class C networks.

In a particular environment, it may be best to use addresses all from one class. For example, a corporate inter network that consists of a large number of departmental local area networks may need to use class C addresses exclusively. However, the format of the addresses is such that it is possible to mix all three classes of addresses on the same inter network; this is what is done in the case of the Internet itself. A mixture of classes is appropriate for an inter network consisting of a few large networks, many small networks, plus some medium-sized networks.

➤ Subnetting

Ethernet-based LANs may use repeaters or hubs to form a larger network and connect more hosts. Any frame generated by a host on one segment is effectively broadcast to each host on every other segment. Since a destination host is only on one segment this may lead to a substantial increase in collision activity and degrade network performance. (In UTP networks this may be resolved by installing switches instead of hubs.) However, a popular solution is to use IP for networking and a router to split the network into a number of subnets, each subnet being connected to one port of the router. In this way datagrams destined for another host on the same subnet are not forwarded by the router to other subnets. Datagrams from a host on one subnet to a host on another subnet are only forwarded to the port to which the destination host is connected.

A LAN which does not use subnets, or LAN switches, may be thought of as comprising a single **collision domain** where each originating frame is passed to every segment and may collide with a frame from any other host in the network. In subnetting each subnet is regarded as a *single* collision domain and overall collision activity is greatly reduced. For maximum advantage in subnetting, users should be grouped into suitable subnets so that most datagrams are destined for the same subnet as the source. For instance, accounts personnel could be assigned to one subnet, management to another, and so on.

Subnetting is achieved by 'stealing' 2 bits or more (the leftmost significant) from the host address portion of an IP address. These bits are known as **subnet bits**. Suppose a network has a Class C address. It has eight host bits. There must be at least two subnet bits. This is because the all zero and all one pattern within the subnet bits is prohibited. (All ones are used for broadcast addresses and all zeros are used as a network address.) In consequence we may use 2 bits, 3 bits and so on. There is a tradeoff. The more subnet bits, and hence subnets, there are, the fewer bits are available for assigning to host addresses.

➤ Supernetting

The time came when most of the class A and class B addresses were depleted; however, there was still a huge demand for midsize blocks. The size of a class C block with a maximum number of 256 addresses did not satisfy the needs of most organizations. Even

organization needed more addresses. One solution was supernetting. In supernetting, an organization can combine several class C blocks to create a larger range of addresses. In other words, several networks are combined to create a supernet or a supemet. An organization can apply for a set of class C blocks instead of just one. For example, an organization that needs 1000 addresses can be granted four contiguous class C blocks. The organization can then use these addresses to create one supernet. Supernetting decreases the number of Is in the mask.

5.5. NETWORK ADDRESS TRANSLATION (NAT)

IP addresses are scarce. An ISP might have a/16 (formerly class B) address, giving it 65,534 host numbers. If it has more customers than that, it has a problem. For home customers with dial-up connections, one way around the problem is to dynamically assign an IP address to a computer when it calls up and logs in and take the IP address back when the session ends. In this way, a single/16 address can handle up to 65,534 active users, which is probably good enough for an ISP with several hundred thousand customers. When the session is terminated, the IP address is reassigned to another caller. While this strategy works well for an ISP with a moderate number of home users, it fails for ISPs that primarily serve business customers.

To make matters worse, more and more home users are subscribing to ADSL or Internet over cable. Two of the features of these services are (1) the user gets a permanent IP address and (2) there is no connect charge (just a monthly flat rate charge), so many ADSL and cable users just stay logged in permanently. This development just adds to the shortage of IP addresses. Assigning IP addresses on-the-fly as is done with dial-up users is of no use because the number of IP addresses in use at any one instant may be many times the number the ISP owns.

And just to make it a bit more complicated, many ADSL and cable users have two or more computers at home, often one for each family member, and they all want to be on-line all the time using the single IP address their ISP has given them. The solution here is to connect all the PCs via a LAN and put a router on it. From the ISP's point of view, the family is now the same as a small business with a handful of computers. Welcome to Jones, Inc.

The problem of running out of IP addresses is not a theoretical problem that might occur at some point in the distant future. It is happening right here and right now. The long-term solution is for the whole Internet to migrate to IPv6, which has 128 bit addresses. This transition is slowly occurring, but it will be years before the process is complete. As a consequence, some people felt that a quick fix was needed for the short term. This quick fix came in the form of **NAT (Network Address Translation)**, which is described in RFC 3022 and which we will summarize below.

The basic idea behind NAT is to assign each company a single IP address (or at most, a small number of them) for Internet traffic. *Within* the company, every computer gets a unique IP address, which is used for routing intramural traffic. However, when a packet exits the company and goes to the ISP, an address translation takes place. To make this

scheme possible, three ranges of IP addresses have been declared as private. Companies may use them internally as they wish. The only rule is that no packets containing these addresses may appear on the Internet itself. The first range provides for 16,777,216 addresses (except for 0 and -1, as usual) and is the usual choice of most companies, even if they do not need so many addresses.

The operation of NAT is shown in Fig. 5.10. Within the company premises, every machine has a unique address of the form 10.x.y.z. However, when a packet leaves the company premises, it passes through a **NAT box** that converts the internal IP source address, 10.0.0.1 in the figure, to the company's true IP address, 198.60.42.12 in this example. The NAT box is often combined in a single device with a firewall, which provides security by carefully controlling what goes into the company and what comes out. We will study firewalls in Chapter 8. It is also possible to integrate the NAT box into the company's router.

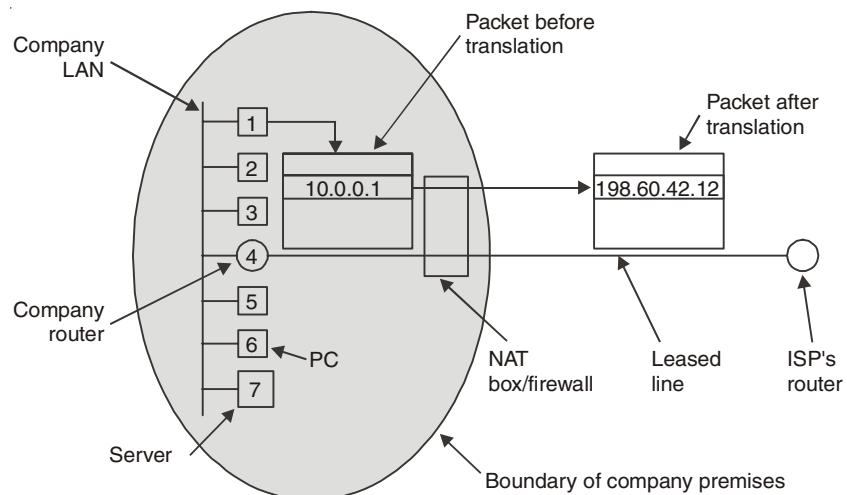


Fig. 5.10. A NAT Box in operation.

When the reply comes back (e.g., from a Web server), it is naturally addressed to 198.60.42.12, so how does the NAT box know which address to replace it with? Here in lies the problem with NAT. If there were a spare field in the IP header, that field could be used to keep track of who the real sender was, but only 1 bit is still unused. In principle, a new option could be created to hold the true source address, but doing so would require changing the IP code on all the machines on the entire Internet to handle the new option. This is not a promising alternative for a quick fix.

What actually happened is as follows. The NAT designers observed that most IP packets carry either TCP or UDP payloads. When we study TCP and UDP in Chapter 6, we will see that both of these have headers containing a source port and a destination port. Below we will just discuss TCP ports, but exactly the same story holds for UDP ports. The ports are 16 bit integers that indicate where the TCP connection begins and ends. These ports provide the field needed to make NAT work.

When a process wants to establish a TCP connection with a remote process, it attaches itself to an unused TCP port on its own machine. This is called the **source port** and tells the TCP code where to send incoming packets belonging to this connection. The process also

supplies a **destination port** to tell who to give the packets to on the remote side. Ports 0–1023 are reserved for well-known services. For example, port 80 is the port used by Web servers, so remote clients can locate them. Each outgoing TCP message contains both a source port and a destination port. Together, these ports serve to identify the processes using the connection on both ends.

When a packet arrives at the NAT box from the ISP, the *Source port* in the TCP header is extracted and used as an index into the NAT box's mapping table. From the entry located, the internal IP address and original TCP *Source port* are extracted and inserted into the packet. Then both the IP and TCP checksums are recomputed and inserted into the packet. The packet is then passed to the company router for normal delivery using the 10 *x.y.z* address.

The chief drawback of NAT is that it breaks a key assumption of the IP service model—that all nodes have globally unique addresses. It turns out that lots of applications and protocols rely on this assumption. In particular, many protocols that run over IP (e.g., application protocols) carry IP addresses in their messages. These addresses also need to be translated by a NAT box if the higher-layer protocol is to work properly, and thus NAT boxes become much more complex than simple IP header translators. They potentially need to understand an ever-growing number of higher-layer protocols. This in turn presents an obstacle to deployment of new applications. It is probably safe to say that networks would be better off without NAT, but its disappearance seems unlikely. Widespread deployment of IPv6 would almost certainly help.

5.6 IP VERSION 6

The version of IP described above is known as IP version 4 (IPv4). IP has been extraordinarily successful in becoming the universal choice of layer 3 protocol for internetworking. However, its very success has produced a problem. This has not, however, deterred the inexorable growth of IP as a *de facto* standard protocol. The problem is that IP is running out of addresses. Although a 32 bit address might, at first sight, seem adequate, the way that the addresses are organized is very wasteful.

For most organizations, a Class A address, which allows for 16 million addresses, is too big and a Class C network, with 256 addresses, is too small. Consequently, Class B addresses, which allow 65 536 addresses, are in great demand but they only allow up to 16 384 networks. It was mainly to solve this lack of addresses that a new version of IP was first outlined in 1995 as RFC 1752. Known as IPv6 (an IP version 5 had already been proposed as an early real-time protocol), it is more recently defined in RFC 1883. The new features of IPv6 can be grouped into four main categories as follows:

1. Address size: Instead of 32 bits, IPv6 uses 128 bit addresses.
2. Header format: An IPv6 header is much simpler than the IPv4 header.
3. Extension headers: IPv6 allows for several headers, in that the basic header can be followed by a number of further extension headers, followed by data.
4. Support for multimedia traffic: Users can establish a high-quality path through an

underlying network and associate voice and video traffic with that path. IPv6 terms this path a **flow** and associates a flow **label** with the datagrams in the flow. This is a dramatic departure from the traditional connectionless mode of IP working.

Figure 5.10 shows the base header with its eight fields. These fields are as follows:

- o **Version.** This 4 bit field defines the version number of the IP. For IPv6, the value is 6.
- o **Priority.** The 4 bit priority field defines the priority of the packet with respect to traffic congestion.
- o **Flow label.** The flow label is a 3 byte (24 bit) field that is designed to provide special handling for a particular flow of data.
- o **Payload length.** The 2 byte payload length field defines the length of the IP datagram excluding the base header.
- o **Next header.** The next header is an 8 bit field defining the header that follows the base header in the datagram. The next header is either one of the optional extension headers used by IP or the header of an encapsulated packet such as UDP or TCP. Each extension header also contains this field.
- o **Hop limit.** This 8 bit hop limit field serves the same purpose as the TIL field in IPv4.
- o **Source address.** The source address field is a 16 byte (128 bit) Internet address that identifies the original source of the datagram.
- o **Destination address.** The destination address field is a 16 byte (128 bit) Internet address that usually identifies the final destination of the datagram. However, if source routing is used, this field contains the address of the next router.

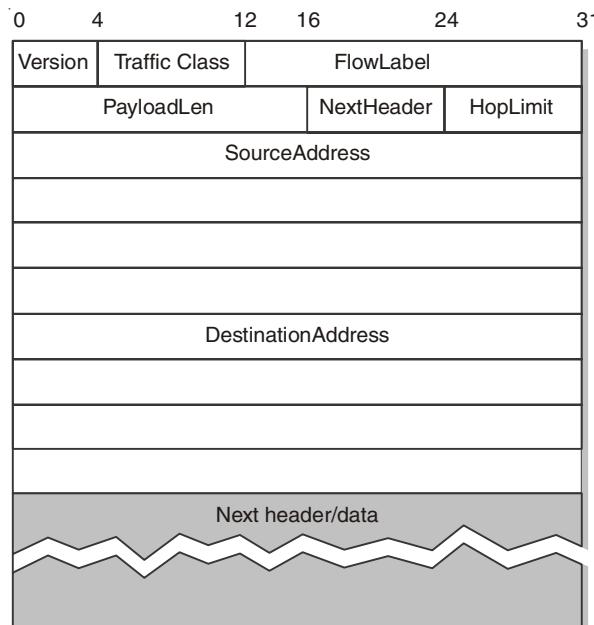


Fig. 5.11. IPv6 Header Format.

Despite offering some interesting solutions to the problems faced by IPv4, IPv6 has not, at the time of writing, been widely adopted. This is partly due to the overwhelming success of IPv4 and the resulting massive investment in it such that the cost of converting to IPv6 is formidable.

5.7 ADDRESS TRANSLATION (ARP)

IP datagrams contain IP addresses, but the physical interface hardware on the host or router to which you want to send the datagram only understands the addressing scheme of that particular network. Thus, we need to translate the IP address to a link-level address that makes sense on this network (*e.g.*, a 48 bit Ethernet address). We can then encapsulate the IP datagram inside a frame that contains that link-level address and send it either to the ultimate destination or to a router that promises to forward the datagram toward the ultimate destination.

One simple way to map an IP address into a physical network address is to encode a host's physical address in the host part of its IP address. For example, a host with physical address 00100001 01001001 (which has the decimal value 33 in the upper byte and 81 in the lower byte) might be given the IP address 128.96.33.81. While this solution has been used on some networks, it is limited in that the network's physical addresses can be no more than 16 bits long in this example; they can be only 8 bits long on a class C network. This clearly will not work for 48 bit Ethernet addresses.

A more general solution would be for each host to maintain a table of address pairs; that is, the table would map IP addresses into physical addresses. While this table could be centrally managed by a system administrator and then copied to each host on the network, a better approach would be for each host to dynamically learn the contents of the table using the network. This can be accomplished using the Address Resolution Protocol (ARP). The goal of ARP is to enable each host on a network to build up a table of mappings between IP addresses and link-level addresses. Since these mappings may change over time (*e.g.*, because an Ethernet card in a host breaks and is replaced by a new one with a new address), the entries are timed out periodically and removed. This happens on the order of every 15 minutes. The set of mappings currently stored in a host is known as the ARP cache or ARP table.

ARP takes advantage of the fact that many link-level network technologies, such as Ethernet and token ring, support broadcast. If a host wants to send an IP datagram to a host (or router) that it knows to be on the same network (*i.e.*, the sending and receiving node have the same IP network number), it first checks for a mapping in the cache. If no mapping is found, it needs to invoke the Address Resolution Protocol over the network. It does this by broadcasting an ARP query onto the network. This query contains the IP address in question (the "target IP address"). Each host receives the query and checks to see if it matches its IP address. If it does match, the host sends a response message that contains its link-layer address back to the originator of the query. The originator adds the information contained in this response to its ARP table.

The query message also includes the IP address and link-layer address of the sending host. Thus, when a host broadcasts a query message, each host on the network can learn the sender's link-level and IP addresses and place that information in its ARP table. However, not every host adds this information to its ARP table. If the host already has an entry for that host in its table, it "refreshes" this entry; that is, it resets the length of time until it discards the entry. If that host is the target of the query, then it adds the information about the sender to its table, even if it did not already have an entry for that host. This is because there is a good chance that the source host is about to send it an application-level message, and it may eventually have to send a response or ACK back to the source; it will need the source's physical address to do this. If a host is not the target and does not already have an entry for the source in its ARP table, then it does not add an entry for the source. This is because there is no reason to believe that this host will ever need the source's link-level address; there is no need to clutter its ARP table with this information.

Figure 5.12 shows the ARP packet format for IP-to-Ethernet address mappings.

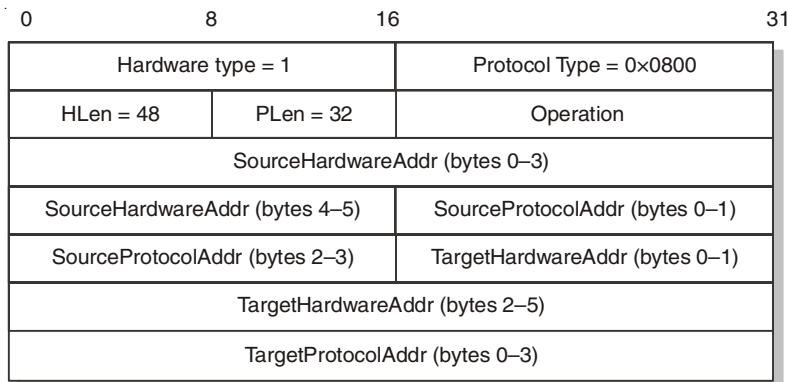


Fig. 5.12. ARP Packet format.

The fields are as follows:

- **Hardware type.** This is a 16 bit field defining the type of the network on which ARP is running. Each LAN has been assigned an integer based on its type. For example, Ethernet is given type 1. ARP can be used on any physical network.
- **Protocol type.** This is a 16 bit field defining the protocol. For example, the value of this field for the IPv4 protocol is 080016, ARP can be used with any higher-level protocol.
- **Hardware length.** This is an 8 bit field defining the length of the physical address in bytes. For example, for Ethernet the value is 6.
- **Protocol length.** This is an 8 bit field defining the length of the logical address in bytes. For example, for the IPv4 protocol the value is 4.
- **Operation.** This is a 16 bit field defining the type of packet. Two packet types are defined: ARP request (1) and ARP reply (2).
- **Sender hardware address.** This is a variable-length field defining the physical address of the sender. For example, for Ethernet this field is 6 bytes long.

- **Sender protocol address.** This is a variable-length field defining the logical (for example, IP) address of the sender. For the IP protocol, this field is 4 bytes long.
- **Target hardware address.** This is a variable-length field defining the physical address of the target. For example, for Ethernet this field is 6 bytes long. For an ARP request message, this field is all 0s because the sender does not know the physical address of the target.
- **Target protocol address.** This is a variable-length field defining the logical (for example, IP) address of the target. For the IPv4 protocol, this field is 4 bytes long.

5.8 HOST CONFIGURATION (DHCP)

IP addresses are not only must be unique on a given internetwork, but also must reflect the structure of the internetwork. As noted above, they contain a network part and a host part, and the network part must be the same for all hosts on the same network. Thus, it is not possible for the IP address to be configured once into a host when it is manufactured, since that would imply that the manufacturer knew which hosts were going to end up on which networks, and it would mean that a host, once connected to one network, could never move to another. For this reason, IP addresses need to be reconfigurable.

In addition to an IP address, there are some other pieces of information a host needs to have before it can start sending packets. The most notable of these is the address of a default router—the place to which it can send packets whose destination address is not on the same network as the sending host.

Most host operating systems provide a way for a system administrator, or even a user, to manually configure the IP information needed by a host. However, there are some obvious drawbacks to such manual configuration. One is that it is simply a lot of work to configure all the hosts in a large network directly, especially when you consider that such hosts are not reachable over a network until they are configured. Even more importantly, the configuration process is very error-prone, since it is necessary to ensure that every host gets the correct network number and that no two hosts receive the same IP address. For these reasons, automated configuration methods are required. The primary method uses a protocol known as the Dynamic Host Configuration Protocol (DHCP).

DHCP relies on the existence of a DHCP server that is responsible for providing configuration information to hosts. There is at least one DHCP server for an administrative domain. At the simplest level, the DHCP server can function just as a centralized repository for host configuration information. Consider, for example, the problem of administering addresses in the internetwork of a large company. DHCP saves the network administrators from having to walk around to every host in the company with a list of addresses and network map in hand and configuring each host manually. Instead, the configuration information for each host could be stored in the DHCP server and automatically retrieved by each host when it is booted or connected to the network. However, the administrator would still pick the address that each host is to receive; he would just store that in the server. In this model, the configuration information for each host is stored in a table that is

indexed by some form of unique client identifier, typically the “hardware address” (e.g., the Ethernet address of its network adaptor).

A more sophisticated use of DHCP saves the network administrator from even having to assign addresses to individual hosts. In this model, the DHCP server maintains a pool of available addresses that it hands out to hosts on demand. This considerably reduces the amount of configuration an administrator must do, since now it is only necessary to allocate a range of IP addresses (all with the same network number) to each network.

Since the goal of DHCP is to minimize the amount of manual configuration required for a host to function, it would rather defeat the purpose if each host had to be configured with the address of a DHCP server. Thus, the first problem faced by DHCP is that of server discovery. To contact a DHCP server, a newly booted or attached host sends a DHCPDISCOVER message to a special IP address (255.255.255.255) that is an IP broadcast address. This means it will be received by all hosts and routers on that network. (Routers do not forward such packets onto other networks, preventing broadcast to the entire Internet.) In the simplest case, one of these nodes is the DHCP server for the network. The server would then reply to the host that generated the discovery message (all the other nodes would ignore it). However, it is not really desirable to require one DHCP server on every network because this still creates a potentially large number of servers that need to be correctly and consistently configured. Thus, DHCP uses the concept of a *relay agent*. There is at least one relay agent on each network, and it is configured with just one piece of information: the IP address of the DHCP server. When a relay agent receives a DHCPDISCOVER message, it unicasts it to the DHCP server and awaits the response, which it will then send back to the requesting client. The process of relaying a message from a host to a remote DHCP server is shown in Fig. 5.13.

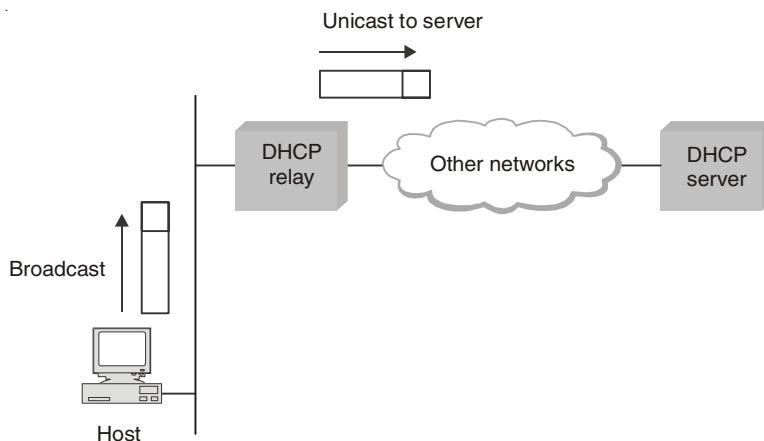


Fig. 5.13. A DHCP relay agent receives a broadcast DHCPDISCOVER message from a host and sends a unicast DHCPDISCOVER message to the DHCP server.

Figure 5.14 shows the format of a DHCP message. The message is actually sent using a protocol called UDP (the User Datagram Protocol) that runs over IP. UDP is discussed in detail in the next chapter, but the only interesting thing it does in this context is to provide a demultiplexing key that says, “This is a DHCP packet.”

DHCP is derived from an earlier protocol called BOOTP, and some of the packet fields are thus not strictly relevant to host configuration. When trying to obtain configuration information, the client puts its hardware address (e.g., its Ethernet address) in the chaddr field. The DHCP server replies by filling in the yiaddr (“your” IP address) field and sending it to the client. Other information such as the default router to be used by this client can be included in the options field. In the case where DHCP dynamically assigns IP addresses to hosts, it is clear that hosts cannot keep addresses indefinitely, as this would eventually cause the server to exhaust its address pool. At the same time, a host cannot be depended upon to give back its address, since it might have crashed, been unplugged from the network, or been turned off. Thus, DHCP allows addresses to be “leased” for some period of time. Once the lease expires, the server is free to return that address to its pool. A host with a leased address clearly needs to renew the lease periodically if in fact it is still connected to the network and functioning correctly.

Operation	HType	HLen	Hops
xid			
Secs		Flags	
ciaddr			
yiaddr			
siaddr			
giaddr			
chaddr (16 bytes)			
sname (64 bytes)			
file (128 bytes)			
options			

Fig. 5.14. DHCP packet format.

5.9 THE INTERNET CONTROL MESSAGE PROTOCOL (ICMP)

The IP standard specifies that a compliant implementation must also implement ICMP (RFC 792). ICMP provides a means for transferring messages from routers and other hosts to a host. In essence, ICMP provides feedback about problems in the communication environment. Examples of its use are: When a datagram cannot reach its destination, when the router does not have the buffering capacity to forward a datagram, and when the router can direct the station to send traffic on a shorter route. In most cases, an ICMP message is sent in response to a datagram, either by a router along the datagram’s path, or by the intended destination host.

Although ICMP is, in effect, at the same level as IP in the TCPIIP architecture, it is a user of IP. An ICMP message is constructed and then passed down to IP, which encapsulates the message with an IP header and then transmits the resulting datagram in the usual fashion. Because ICMP messages are transmitted in IP datagrams, their delivery is not guaranteed and their use cannot be considered reliable.

Figure 5.15 shows the format of the various ICMP message types. All ICMP message start with a 64-bit header consisting of the following:

Type (8 bits). Specifies the type of ICMP message.

Code (8 bits). Used to specify parameters of the message that can be encoded in one or a few bits.

Checksum (16 bits). Checksum of the entire ICMP message. This is the same checksum algorithm used for IP.

Parameters (32 bits). Used to specify more lengthy parameters.

These fields are generally followed by additional information fields that further specify the content of the message. In those cases in which the ICMP message refers to a prior datagram, the information field includes the entire IP header plus the first 64 bits of the data field of the original datagram. This enables the source host to match the incoming ICMP message with the prior datagram. The reason for including the first 64 bits of the data field is that this will enable the IP module in the host to determine which upper-level protocol or protocols were involved. In particular, the first 64 bits would include a portion of the TCP header or other transport-level header.

ICMP messages include the following:

- Destination unreachable
- Time exceeded
- Parameter problem
- Source quench
- Redirect
- Echo
- Echo reply
- Timestamp
- Timestamp reply
- Address mask request
- Address mask reply

Type	Code	Checksum	
Unused			
IP Header + 64 bits of original datagram			

(a) Destination unreachable;
time exceeded; source quench

Type	Code	Checksum	
Pointer	Unused		
IP Header + 64 bits of original datagram			

(b) Parameter problem

Type	Code	Checksum	
Gateway internet address			
IP Header + 64 bits of original datagram			

(c) Redirect

Type	Code	Checksum	
Identifier	Sequence number		
IP Header + 64 bits of original datagram			

(d) Echo, echo reply

Type	Code	Checksum	
Identifier	Sequence number		
Address mask			

(e) Timestamp

Type	Code	Checksum	
Identifier	Sequence number		
Origin timestamp			

(f) Timestamp reply

Type	Code	Checksum	
Identifier	Sequence number		
Receive timestamp			
Transmit timestamp			

(g) Address mask request

Type	Code	Checksum	
Identifier	Sequence Number		
Address mask			

(h) Address mask reply

Fig. 5.15. ICMP Formats.

The **destination-unreachable** message covers a number of contingencies. A router may return this message if it does not know how to reach the destination network. In some networks, an attached router may be able to determine if a particular host is unreachable, and then return the message. The destination host itself may return this message if the user protocol or some higher-level service access point is unreachable. This could happen if the corresponding field in the IP header was set incorrectly. If the datagram specifies a source route that is unusable, a message is returned. Finally, if a router must fragment a datagram but the Don't-Fragment flag is set, a message is returned.

A router will return a **time-exceeded** message if the lifetime of the datagram expires. A host will send this message if it cannot complete reassembly within a time limit.

A syntactic or semantic error in an IP header will cause a *parameter-problem* message to be returned by a router or host. For example, an incorrect argument may be provided with an option. The parameter field contains a pointer to the octet in the original header where the error was detected.

The *source-quench* message provides a rudimentary form of flow control. Either a router or a destination host may send this message to a source host, requesting that it reduce the rate at which it is sending traffic to the internet destination. On receipt of a source-quench message, the source host should cut back the rate at which it is sending traffic to the specified destination until it no longer receives source-quench messages; this message can

be used by a router or host that must discard datagrams because of a full buffer. In this case, the router or host will issue a source-quench message for every datagram that it discards. In addition, a system may anticipate congestion and issue such messages when its buffers approach capacity. In that case, the datagram referred to in the source-quench message may well be delivered. Thus, receipt of the message does not imply delivery or nondelivery of the corresponding datagram.

The *echo* and *echo-reply* messages provide a mechanism for testing that communication is possible between entities. The recipient of an echo-message is obligated to return the message in an echo-reply message. An identifier and sequence number are associated with the echo message to be matched in the echo-reply message. The identifier might be used like a service access point to identify a particular session, and the sequence number might be incremented on each echo request sent. The *timestamp* and *timestamp-reply* messages provide a mechanism for sampling the delay characteristics of the internet. The sender of a timestamp message may include an identifier and sequence number in the parameters field and include the time that the message is sent (originate timestamp). The receiver records the time it received the message and the time that it transmits the reply message in the timestamp-reply message. If the timestamp message is sent using strict source routing, then the delay characteristics of a particular route can be measured.

The *address-mask-request* and *address-mask-reply* messages are useful in an environment that includes what are referred to as subnets. The concept of the subnet was introduced to address the following requirement. Consider an internet that includes one or more WANs and a number of sites, each of which has a number of LANs. We would like to allow arbitrary complexity of interconnected LAN structures within an organization, while insulating the overall internet against explosive growth in network numbers and routing complexity. One approach to this problem is to assign a single network number to all of the LANs at a site. From the point of view of the rest of the internet, there is a single network at that site, which simplifies addressing and routing. To allow the routers within the site to function properly, each LAN is assigned a subnet number. The host portion of the internet address is partitioned into a subnet number and a host number to accommodate this new level of addressing.

Within the subnetted network, the local routers must route on the basis of an extended network number consisting of the network portion of the IP address and the subnet number. The bit positions containing this extended network number are indicated by the address mask. The address mask request and reply messages allow a host to learn the address mask for the LAN to which it connects. The host broadcasts an address mask request message on the LAN. The router on the LAN responds with an address mask reply message that contains the address mask. The use of the address mask allows the host to determine whether an outgoing datagram is destined for a host on the same LAN (send directly) or another LAN (send datagram to router). It is assumed that some other means (*e.g.*, manual configuration) is used to create address masks and to make them known to the local routers.

5.10 ROUTING PROTOCOLS

A major function of the Internet layer, and network layers in general, is to route traffic from a source to a destination. Routing decisions are not simply based upon destination addresses alone. They must also take account of the network topology and prevailing traffic conditions. For instance, many internet topologies contain loops and routing tables can, if correctly configured, ensure that a datagram does not follow a loop and thus reappear at a node already traversed.

Routers (and switches) operate by discovering the topology of their network and then use this topology to build routing tables containing the best routes to destinations. In order to build topology and routing tables a router cannot operate alone, it needs to exchange information with other routers. This exchange of information is achieved by routers 'talking' to each other by means of **routing protocols** which are also a part of the TCP/IP suite and reside at the Internet layer. Three routing protocols are considered here, namely Routing Information Protocol (RIP), Open Shortest Path First (OSPF) and Border Gateway Protocol (BGP).

➤ Routing Information Protocol

Routing Information Protocol (RIP) is a simple routing protocol, originally defined in 1988 as RFC 1058 and more recently as RFC 1723, based upon the original ARPANET routing algorithm. RIP involves a router calculating the best route to all other routers in a network using a **shortest path** algorithm attributable to Bellman (1957) and Ford and Fulkerson (1962). The shortest path in this case is the one that passes through the least number of routers. Each router traversed is known as a **hop**. Therefore, the shortest path is described by a **hop count**, or **distance vector**. This is a crude measure of distance or cost to reach a destination. It takes no account of other factors such as propagation delay or available bandwidth. RIP then builds a routing database that contains tables of the best routes to all the other routers. Each router then advertises its own routing tables to all other routers. Although RIP is simple to implement it is only efficient in small networks since, as the size of a network grows, RIP datagrams can become very long, thus consuming substantial amounts of bandwidth.

➤ Open Shortest Path First

A more powerful routing protocol developed subsequent to RIP, defined originally as RFC 1131 and more recently as RFC 2178, is called **Open Shortest Path First** (OSPF). It is the preferred routing protocol for medium or large networks which, in OSPF, are referred to as **autonomous systems** (ASs). OSPF endeavours to establish a least-cost shortest route within an autonomous system. Cost does not necessarily involve monetary considerations, but means that parameters are used that are of particular importance to the network operator. They may be financial or could be based on delay or transmission rate. Such parameters are known as **metrics**. Whereas RIP is a distance-vector-based protocol, OSPF is described as a **link state** routing protocol. This is because it only advertises the changes in the state of its routing tables to other routers using **link state advertisements** rather than the full tables. Link state advertisements that are exchanged between routers produce much less traffic than is generated by RIP datagrams. Each router holds a database, each containing the same

information, as a result of the exchange of link state update messages. It is worth noting that, unlike RIP, OSPF only exchanges changes in a network rather than complete topologies. This is a major advantage over RIP and results in much less information being exchanged. Cost metrics are indicated at the output ports of each router and may be deduced by router software or configured by a network administrator.

Since autonomous systems can be large, OSPF allows for them to be divided into numbered areas such that topology information is largely contained within a single area. Area 0 is a special case, termed the backbone area, and is arranged so that all other areas can be interconnected through it. Routers operating in an OSPF environment can be categorized by their connectivity with other routers and the type of traffic that they carry as illustrated in Fig. 5.16. A **stub router** has only one entry/exit point to the router and all traffic passes through this one point, whereas **multihomed routers** have more than one connection to other routers.

OSPF, in common with certain other routing protocols, can also use equal-cost multipath routing to avoid some parts of the network becoming congested while other parts are not fully utilized. Such procedures are not part of the OSPF protocol and an equal-cost multipath algorithm is analysed in RFC 2992. Equal-cost multipath routing, as the name implies, is a technique for routing datagrams along multiple paths of equal cost. The forwarding algorithm identifies paths by next-hop and the router must then decide which next-hop (path) to use when forwarding a datagram. For example, a round-robin technique might be used whereby each eligible path is used in turn. However, such an approach is not suitable for TCP sessions, which perform better if the path they flow along does not change while the stream is connected. A more useful method for determining which next-hop to use is known as a **hash-threshold**.

The router first selects a **key** by performing a cyclic redundancy check (known as a **hash**) over the datagram header fields that identify a flow (typically the source and destination IP addresses). With the very simplest implementation of the algorithm, the combined source and destination addresses are divided by the number of equal-cost routes and the remainder of this division is the key. The router then uses the key to determine which next-hop to use. This should result in a more balanced use of the available paths and is known as **load balancing**.

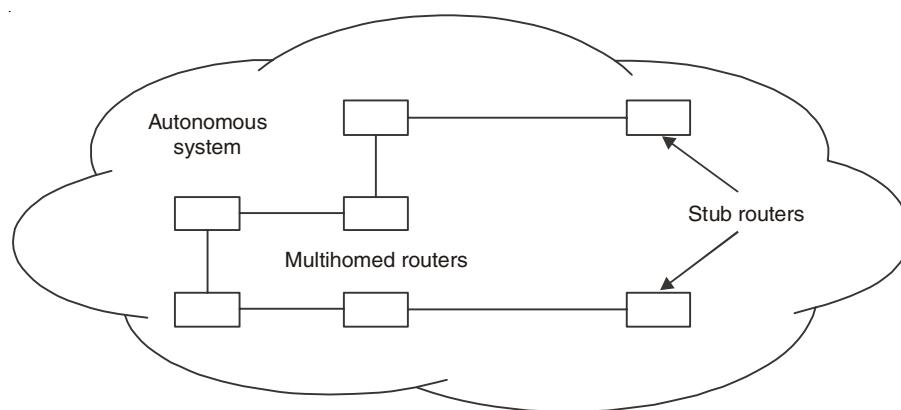


Fig. 5.16. Stub and Multihomed Routers.

➤ Border Gateway Protocol

The Border Gateway Protocol (BGP) is a routing protocol used to exchange network reachability information between autonomous systems. It is an example of an exterior gateway protocol rather than an internal gateway protocol such as OSPF or RIP. BGP was defined in RFC 1771 and subsequently RFC 1997 and in its version 4 is currently the routing protocol of choice for the Internet. BGP provides a set of mechanisms that allow support for advertising an abbreviated IP address, known as a prefix, and thus eliminate the concept of network ‘class’ from an IP address. As such, it is known as a **classless** routing protocol.

Routers, known as BGP **speakers**, exchange BGP routing information over a TCP connection. The routing information contains a list of reachable autonomous systems with **path attributes** that convey routing information, such as preference for a particular hop or route. If the two BGP speakers are in the same AS, the connection between them uses the interior BGP (IBGP). As long as an interior gateway protocol such as OSPF is running in the AS, two speakers are reachable using IP and an IBGP connection will be established. If the two BGP speakers are in different ASs then the connection between them uses the exterior BGP. It is not necessary, and is in fact undesirable, for every router in an AS to be running BGP: as long as at least one router in an AS uses BGP, the AS will remain in contact with other ASs. Of particular importance to BGP are what are known as **import and export policies**. These define which types of traffic can be received or transmitted by a particular BGP speaker. For example, a carrier will be willing to carry its own customers’ traffic but may be unwilling to allow other carriers’ customer traffic to transit through its IP network. Policies are manually configured on each router that supports BGP. Thus, a BGP speaker may have policies configured that permit it to distribute addresses from one external AS within its own AS but not the addresses from another AS. These policies are not only of importance for political reasons. They also have the effect of reducing the amount of routing information that is distributed between ASs. This is important in large groups of networks such as the Internet, as the amount of routing information could otherwise become unmanageable. Import and export policies can also be used with OSPF, but they do not assume the same importance as when used by BGP, as the distribution of routing information by OSPF is contained within a single AS. BGP also shares the nomenclature used by OSPF to describe different routers as multi homed, stub, etc. An AS that has just one single connection to another AS, described as a **stub AS**, is illustrated in Fig. 5.17. A further type of AS, known as a **transit AS**, in which multiple connections to other ASs exist and through which traffic can transit, is also shown. Clearly all transit ASs must be multihomed since, by definition, they connect to more than one other AS.

It is possible for an AS to have more than one connection to other ASs but for traffic to be unable to transit through the AS, in which case it is termed a multihomed non-transit AS.

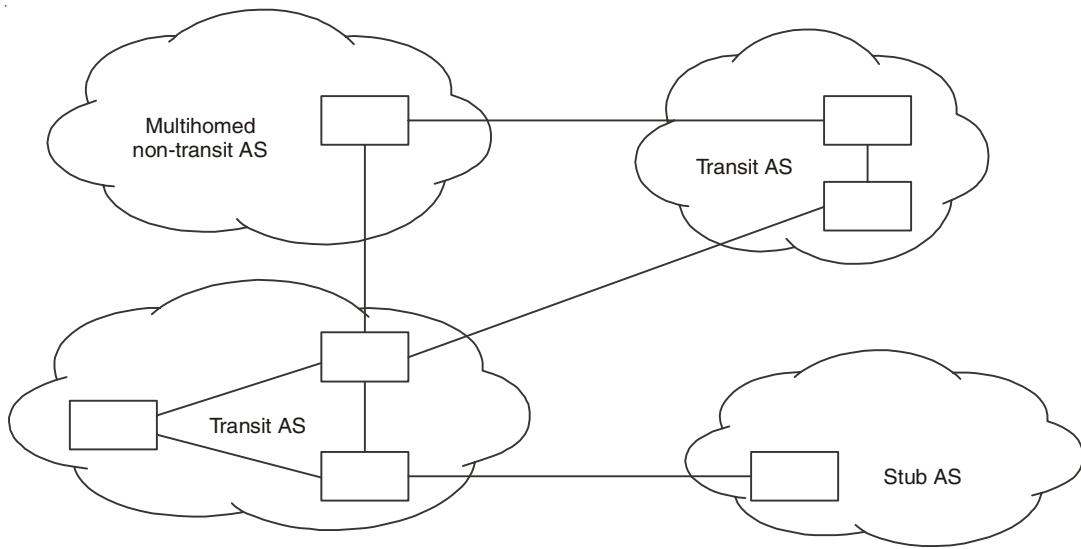


Fig. 5.17. Stub and Transit Autonomous systems.

5.11 INTERNET PROTOCOL SECURITY (IPSec)

IPSec is a suite of authentication and encryption protocols developed by the Internet Engineering Task Force (IETF) and designed to address the inherent lack of security for IP-based networks. IPSec, unlike other protocols we have discussed so far, is a very complex set of protocols described in a number of RFCs including RFC 2401 and 2411. It runs transparently to transport layer and application layer protocols which do not see it. Although it was designed to run in the new version of the Internet Protocol, IP Version 6 (IPv6), it has also successfully run in the older IPv4 as well. IPSec sets out to offer protection by providing the following services at the network layer:

- **Access control** – to prevent an unauthorized access to the resource.
- **Connectionless integrity** – to give an assurance that the traffic received has not been modified in any way.
- **Confidentiality** – to ensure that Internet traffic is not examined by nonauthorized parties. This requires all IP datagrams to have their data field, TCP, UDP, ICMP, or any other datagram data field segment, encrypted.
- **Authentication** – particularly source authentication so that when a destination host receives an IP datagram, with a particular IP source address, it is possible to be sure that the IP datagram was indeed generated by the host with the source IP address. This prevents spoofed IP addresses.
- **Replay protection** – to guarantee that each packet exchanged between two parties is different.

IPSec protocol achieves these two objectives by dividing the protocol suite into two main protocols: Authentication Header (AH) protocol and the Encapsulation Security Payload (ESP) protocol [8]. The AH protocol provides source authentication and data integrity but no confidentiality. The ESP protocol provides authentication, data integrity, and confidentiality. Any datagram from a source must be secured with either AH or ESP. Figures 5.17 and 5.18 show both IPSec's ESP and AH protections.

➤ **Authentication Header (AH)**

AH protocol provides source authentication and data integrity but not confidentiality. This is done by a source that wants to send a datagram first establishing an SA, through which the source can send the datagram. A source datagram includes an AH inserted between the original IP datagram data and the IP header to shield the data field which is now encapsulated as a standard IP datagram.

Upon receipt of the IP datagram, the destination host notices the AH and processes it using the AH protocol. Intermediate hosts such as routers, however, do their usual job of examining every datagram for the destination IP address and then forwarding it on.

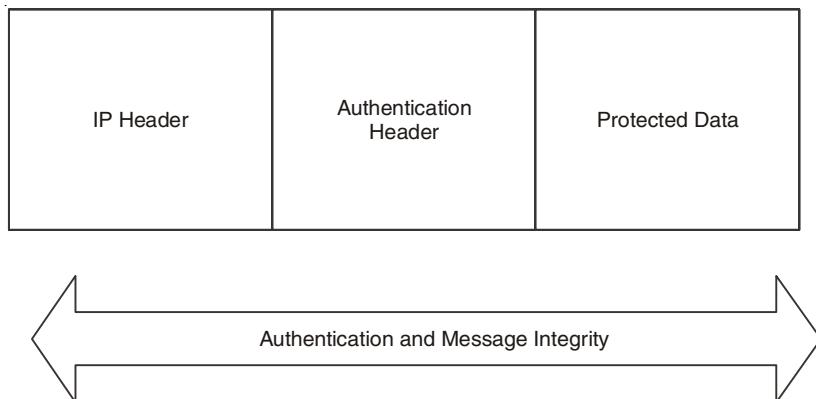


Fig. 5.18. IPsec's AH Protocol Protection.

➤ **Encapsulating Security Payload (ESP)**

Unlike the AH protocol, ESP protocol provides source authentication, data integrity, and confidentiality. This has made ESP the most commonly used IPSec header. Similar to AH, ESP begins with the source host establishing an AS which it uses to send secure datagrams to the destination. Datagrams are secured by ESP by surrounding their original IP datagrams with a new header and trailer fields all encapsulated into a new IP datagram. Confidentiality is provided by DES_CBC encryption. Next to the ESP trailer field on the datagram is the ESP Authentication Data field.

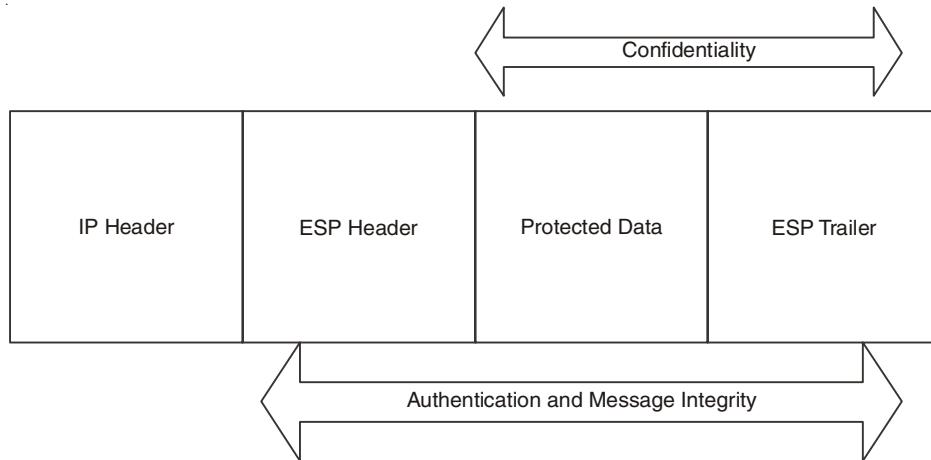


Fig. 5.19. IPSec's ESP Protocol Protection.

➤ Security Associations

In order to perform the security services that IPSec provides, IPSec must first get as much information as possible on the security arrangement of the two communicating hosts. Such security arrangements are called *security associations* (SAs). A security association is a unidirectional security arrangement defining a set of items and procedures that must be shared between the two communicating entities in order to protect the communication process. In the usual network IP connections, the network layer IP is connectionless. However, with security associations, IPSec creates logical connection-oriented channels at the network layer. This logical connection-oriented channel is created by a security agreement established between the two hosts stating specific algorithms to be used by the sending party to ensure confidentiality (with ESP), authentication, message integrity, and anti-replay protection.

Since each AS establishes a unidirectional channel, for a full duplex communication between two parties, two SAs must be established. An SA is defined by three parameters:

- **Security Parameter Index (SPI)** – a 32 bit connection identifier of the SA. For each association between a source and destination host, there is one SPI that is used by all datagrams in the connection to provide information to the receiving device on how to process the incoming traffic.
- **IP Destination Address** – address of a destination host.
- A Security Protocol (AH or ESP) to be used and specifying if traffic is to be provided with integrity and secrecy. The protocol also defines the key size, key lifetime, and the cryptographic algorithms.
- **Secret key** – which defines the keys to be used.
- **Encapsulation mode** – defining how encapsulation headers are created and which parts of the header and user traffic are protected during the communication process.

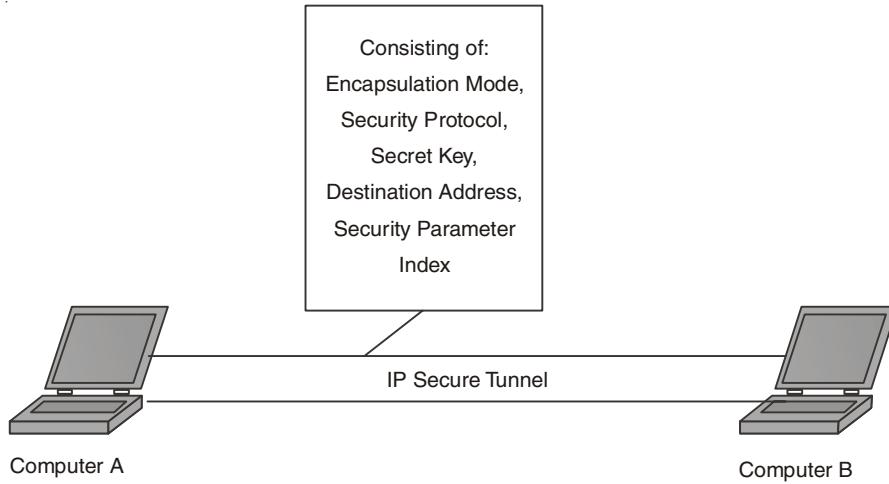


Fig. 5.20. A General Concept of IPSec's Security Association.

The security associations discussed above are implemented in two modes: transport and tunnel. This means that IPSec is operating in two modes. Transport mode provides host-to-host protection to higher layer protocols in the communication between two hosts in both IPv4 and IPv6. Tunnel mode offers protection to the entire IP datagram both in AH and ESP between two IPSec gateways.

5.12 VIRTUAL PRIVATE NETWORKS (VPN)

A VPN is a private data network that makes use of the public telecommunication infrastructure, such as the Internet, by adding security procedures over the unsecure communication channels. The security procedures that involve encryption are achieved through the use of a tunneling protocol. There are two types of VPNs: remote access which lets single users connect to the protected company network and site-to-site which supports connections between two protected company networks.

In either mode, VPN technology gives a company the facilities of expensive private leased lines at much lower cost by using the shared public infrastructure like the Internet. See Fig. 5.21.

Figure 5.21 shows two components of a VPN:

- Two terminators which are either software or hardware. These perform encryption, decryption and authentication services. They also encapsulate the information.
- A **tunnel** – connecting the end-points. The tunnel is a secure communication link between the end-points and networks such as the Internet. In fact this tunnel is virtually created by the end-points.

VPN technology must do the following activities:

- **IP encapsulation** – this involves enclosing TCP/IP data packets within another packet with an IP-address of either a firewall or a server that acts as a VPN endpoint. This encapsulation of host IP-address helps in hiding the host.

- **Encryption** – is done on the data part of the packet. Just like in SSL, the encryption can be done either in transport mode which encrypts its data at the time of generation or tunnel mode which encrypts and decrypts data during transmission encrypting both data and header.
- **Authentication** – involves creating an encryption domain which includes authenticating computers and data packets by use for public encryption.

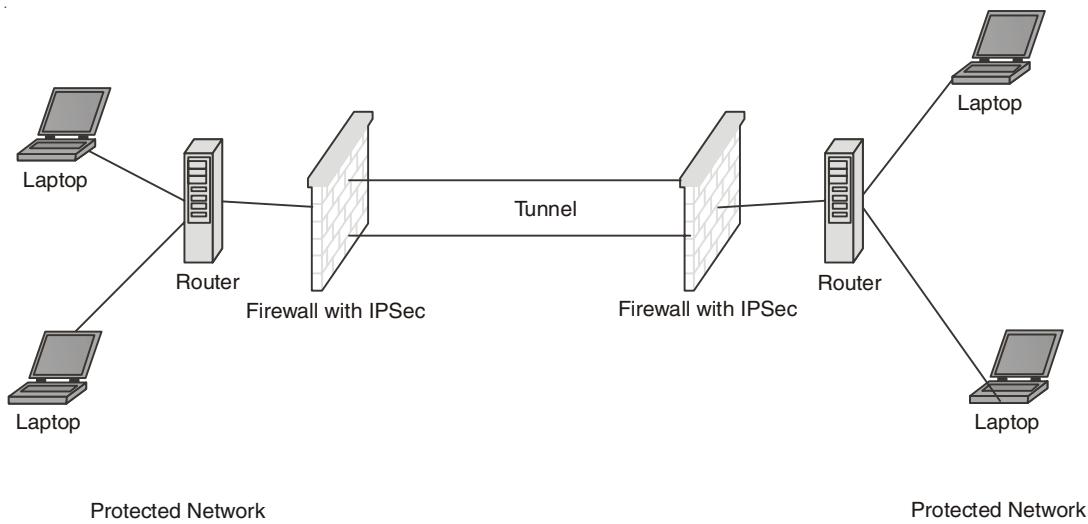


Fig. 5.21. Virtual Private network Model.

VPN technology is not new; phone companies have provided private shared resources for voice messages for over a decade. However, its extension to making it possible to have the same protected sharing of public resources for data is new. Today, VPNs are being used for both extranets and wide-area intranets. Probably owing to cost savings, the popularity of VPNs by companies has been phenomenal.

Review Questions

1. What is the number of bits in an IPv4 address? What is the number of bits in an IPv6 address?
2. What is dotted decimal notation in IPv4 addressing? What is the number of bytes in an IPv4 address represented in dotted decimal notation? What is hexadecimal notation in IPv6 addressing? What is the number of digits in an IPv6 address represented in hexadecimal notation?
3. What are the differences between classful addressing and classless addressing in IPv4?
4. Explain why most of the addresses in class A are wasted. Explain why a medium-size or large-size corporation does not want a block of class C addresses.
5. Briefly define subnetting and supernetting. How do the subnet mask and supernet mask differ from a default mask in classful addressing?

6. How can we distinguish a multicast address in IPv4 addressing? How can we do so in IPv6 addressing?
7. What is NAT? How can NAT help in address depletion?
8. Explain the need for options in IPv4 and list the options mentioned in this chapter with a brief description of each.
9. Compare and contrast the fields in the main headers of IPv4 and IPv6. Make a table that shows the presence or absence of each field.
10. Is the size of the ARP packet fixed? Explain.
11. What is the size of an ARP packet when the protocol is IPv4 and the hardware is Ethernet?
12. What is the broadcast address for Ethernet?
13. Why is there a restriction on the generation of an ICMPv4 message in response to a failed ICMPv4 error message?



CHAPTER 6 *THE TRANSPORT LAYER*

Victory is the beautiful, bright colored flower. Transport is the stem without which it could never have blossomed.

—Winston Churchill

6.1 INTRODUCTION

The transport layer is responsible for process-to-process delivery of the entire message. A process is an application program running on a host. Whereas the network layer oversees source-to-destination delivery of individual packets, it does not recognize any relationship between those packets. It treats each one independently, as though each piece belonged to a separate message, whether or not it does. The transport layer, on the other hand, ensures that the whole message arrives intact and in order, overseeing both error control and flow control at the source-to-destination level.

Computers often run several programs at the same time. For this reason, source to destination delivery means delivery not only from one computer to the next but also from a specific process on one computer to a specific process on the other. The transport layer header must therefore include a type of address called a *service-point address* in the OSI model and port number or port addresses in the Internet and TCP/IP protocol suite.

The transport protocol is the keystone of the whole concept of a computer communications architecture. Lower-layer protocols are needed, to be sure, but they are less important for (1) pedagogical purposes, and (2) designing purposes. For one thing, lower-level protocols are better understood and, on the whole, less complex than transport protocols. Also, standards have settled out quite well for most kinds of layer 1 to 3 transmission facilities, and there is a large body of experience behind their use.

Viewed from the other side, upper-level protocols depend heavily on the transport protocol. The transport protocol provides the basic end-to-end service of transferring data between users and relieves applications and other upper-layer protocols from the need to deal with the characteristics of intervening communications networks and services.

This chapter basically deals with Transport layer services, TCP, UDP Protocols.

6.2 TRANSPORT SERVICES

We begin by looking at the kinds of services that a transport protocol can or should provide to higher-level protocols. Figure 6.1 places the concept of transport services in context. In a system, there is a transport entity that provides services to TS users, which might be an application process or a session-protocol entity. This local transport entity communicates with some remote-transport entity, using the services of some lower layer, such as the network layer.

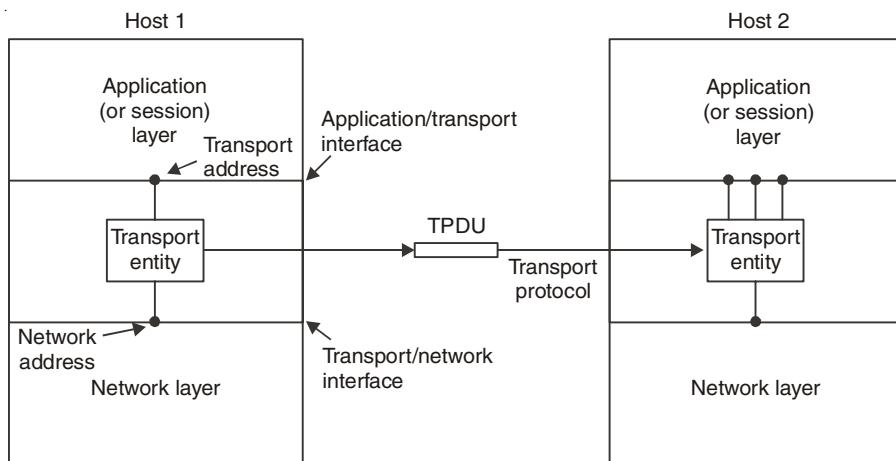


Fig. 6.1. The Network, Transport and Application Layer.

We have already mentioned that the general service provided by a transport protocol is the end-to-end transport of data in a way that shields the TS user from the details of the underlying communications systems. To be more specific, we must consider the specific services that a transport protocol can provide. The following categories of service are useful for describing the transport service:

- Type of service
- Quality of service
- Data transfer
- User interface
- Connection management
- Expedited delivery
- Status reporting
- Security

➤ Type of Service

Two basic types of service are possible: connection-oriented and connectionless, or datagram service. A connection-oriented service provides for the establishment, maintenance, and termination of a logical connection between TS users. This has, so far, been the most common type of protocol service available and has a wide variety of applications. The connection-oriented service generally implies that the service is reliable.

The strengths of the connection-oriented approach are clear. It allows for connection related features such as flow control, error control, and sequenced delivery. Connectionless service, however, is more appropriate in some contexts. At lower layers (internet, network), connectionless service is more robust. In addition, it represents a “least common denominator” of service to be expected at higher layers. Further, even at transport and above, there is justification for a connectionless service. There are instances in which the overhead of connection establishment and maintenance is unjustified or even counterproductive. Some examples follow:

Inward data collection. Involves the periodic active or passive sampling of data sources, such as sensors, and automatic self-test reports from security equipment or network components. In a real-time monitoring situation, the loss of an occasional data unit would not cause distress, as the next report should arrive shortly.

Outward data dissemination. Includes broadcast messages to network users, the announcement of a new node or the change of address of a service, and the distribution of real-time clock values.

➤ **Quality of Service**

The transport protocol entity should allow the TS user to specify the quality of transmission service to be provided. The transport entity will attempt to optimize the use of the underlying link, network, and internet resources to the best of its ability, so as to provide the collective requested services.

Examples of services that might be requested are:

- Acceptable error and loss levels.
- Desired average and maximum delay.
- Desired average and minimum throughput.
- a Priority level.

The TS user of the quality-of-service feature needs to recognize that depending on the nature of the transmission facility, the transport entity will have varying degrees of success in providing a requested grade of service. There is bound to be a trade-off among reliability, delay, throughput, and cost of services. Nevertheless, certain applications would benefit from, or even require, certain qualities of service and, in a hierarchical or layered architecture, the easiest way for an application to extract this quality of service from a transmission facility is to pass the request down to the transport protocol.

Examples of applications that might request particular qualities of service are as follows:

1. A file transfer protocol might require high throughput. It may also require high reliability to avoid retransmissions at the file transfer level.
2. A transaction protocol (*e.g.*, web browser-web server) may require low delay.
3. An electronic mail protocol may require multiple priority levels.

One approach to providing a variety of qualities of service is to include a quality-of-service facility within the protocol; we have seen this with IP and will see that transport protocols typically follow the same approach. An alternative is to provide a different transport

protocol for different classes of traffic; this is to some extent the approach taken by the ISO-standard family of transport protocols.

➤ Data Transfer

The whole purpose, of course, of a transport protocol is to transfer data between two transport entities. Both user data and control data must be transferred, either on the same channel or separate channels. Full-duplex service must be provided. Half-duplex and simplex modes may also be offered to support peculiarities of particular TS users.

➤ User Interface

It is not clear that the exact mechanism of the user interface to the transport protocol should be standardized. Rather, it should be optimized to the station environment. As examples, a transport entity's services could be invoked by

1. Procedure calls.
2. Passing of data and parameters to a process through a mailbox.
3. Use of direct memory access (DMA) between a host user and a front-end processor containing the transport entity.

A few characteristics of the interface may be specified, however. For example, a mechanism is needed to prevent the TS user from swamping the transport entity with data. A similar mechanism is needed to prevent the transport entity from swamping a TS user with data. Another aspect of the interface has to do with the timing and significance of confirmations. Consider the following: A TS user passes data to a transport entity to be delivered to a remote TS user. The local transport entity can acknowledge receipt of the data immediately, or it can wait until the remote transport entity reports that the data have made it through to the other end. Perhaps the most useful interface is one that allows immediate acceptance or rejection of requests, with later confirmation of the end-to-end significance.

➤ Connection Management

When connection-oriented service is provided, the transport entity is responsible for establishing and terminating connections. A symmetric connection-establishment procedure should be provided, which allows either TS user to initiate connection establishment. An asymmetric procedure may also be provided to support simplex connections.

Connection termination can be either *abrupt* or *graceful*. With an abrupt termination, data in transit may be lost. A graceful termination prevents either side from shutting down until all data have been delivered.

➤ Expedited Delivery

A service similar to that provided by priority classes is the expedited delivery of data. Some data submitted to the transport service may supersede data submitted previously. The transport entity will endeavor to have the transmission facility transfer the data as rapidly as possible. At the receiving end, the transport entity will interrupt the TS user to notify it of the receipt of urgent data. Thus, the expedited data service is in the nature of an interrupt mechanism, and is used to transfer occasional urgent data, such as a break character from a terminal or an alarm condition. In contrast, a priority service might dedicate resources and adjust parameters such that, on average, higher priority data are delivered more quickly.

➤ Status Reporting

A status reporting service allows the TS user to obtain or be notified of information concerning the condition or attributes of the transport entity or a transport connection.

Examples of status information are

Performance characteristics of a connection (*e.g.*, throughput, mean delay)

1. Addresses (network, transport)
2. Class of protocol in use
3. Current timer values
4. State of protocol “machine” supporting a connection
5. Degradation in requested quality of service.

➤ Security

The transport entity may provide a variety of security services. Access control may be provided in the form of local verification of sender and remote verification of receiver. The transport service may also include encryption decryption of data on demand. Finally, the transport entity may be capable of routing through secure links or nodes if such a service is available from the transmission facility.

6.3 ELEMENTS OF TRANSPORT LAYER PROTOCOLS

The transport service is implemented by a **transport protocol** used between the two transport entities. In some ways, transport protocols resemble the data link protocols we studied in detail in Chapter 3. Both have to deal with error control, sequencing, and flow control, among other issues.

However, significant differences between the two also exist. These differences are due to major dissimilarities between the environments in which the two protocols operate. These difference has many important implications for the protocols, as we shall see in this chapter.

6.3.1 Addressing

Whenever we need to deliver something to one specific destination among many, we need an address. At the data link layer, we need a MAC address to choose one node among several nodes if the connection is not point-to-point. A frame in the data link layer needs a destination MAC address for delivery and a source address for the next node’s reply.

At the network layer, we need an IP address to choose one host among millions. A datagram in the network layer needs a destination IP address for delivery and a source IP address for the destination’s reply.

At the transport layer, we need a transport layer address, called a port number, to choose among multiple processes running on the destination host. The destination port number is needed for delivery; the source port number is needed for the reply.

In the Internet model, the port numbers are 16-bit integers between 0 and 65,535. The client program defines itself with a port number, chosen randomly by the transport layer software running on the client host. This is the ephemeral port number.

The server process must also define itself with a port number. This port number, however, cannot be chosen randomly. If the computer at the server site runs a server process and assigns a random number as the port number, the process at the client site that wants to access that server and use its services will not know the port number. Of course, one solution would be to send a special packet and request the port number of a specific server, but this requires more overhead. The Internet has decided to use universal port numbers for servers; these are called well-known port numbers. There are some exceptions to this rule; for example, there are clients that are assigned well-known port numbers. Every client process knows the well-known port number of the corresponding server process.

➤ IANA Ranges

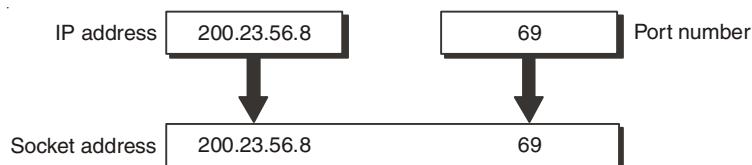
The IANA (Internet Assigned Number Authority) has divided the port numbers into three ranges: well-known, registered, and dynamic (or private).

- **Well-known ports.** The ports ranging from 0 to 1023 are assigned and controlled by IANA. These are the well-known ports.
- **Registered ports.** The ports ranging from 1024 to 49,151 are not assigned or controlled by IANA. They can only be registered with IANA to prevent duplication.
- **Dynamic ports.** The ports ranging from 49,152 to 65,535 are neither controlled nor registered. They can be used by any process. These are the ephemeral ports.

➤ Socket Addresses

Process-to-process delivery needs two identifiers, IP address and the port number, at each end to make a connection. The combination of an IP address and a port number is called a socket address. The client socket address defines the client process uniquely just as the server socket address defines the server process uniquely.

A transport layer protocol needs a pair of socket addresses: the client socket address and the server socket address. These four pieces of information are part of the IP header and the transport layer protocol header. The IP header contains the IP addresses; the UDP or TCP header contains the port numbers.



6.3.2 Connection Establishment

Establishing a connection sounds easy, but it is actually surprisingly tricky. At first glance, it would seem sufficient for one transport entity to just send a CONNECTION REQUEST TPDU to the destination and wait for a CONNECTION ACCEPTED reply. The problem occurs when the network can lose, store, and duplicate packets. This behavior causes serious

complications. Imagine a subnet that is so congested that acknowledgements hardly ever get back in time and each packet times out and is retransmitted two or three times. Suppose that the subnet uses datagrams inside and that every packet follows a different route. Some of the packets might get stuck in a traffic jam inside the subnet and take a long time to arrive, that is, they are stored in the subnet and pop out much later.

This problem can be attacked in various ways, none of them very satisfactory. One way is to use throw-away transport addresses. In this approach, each time a transport address is needed, a new one is generated. When a connection is released, the address is discarded and never used again. But this is not even a satisfactory approach.

Another possibility is to give each connection a connection identifier (*i.e.*, a sequence number incremented for each connection established) chosen by the initiating party and put in each TPDU, including the one requesting the connection. After each connection is released, each transport entity could update a table listing obsolete connections as (peer transport entity, connection identifier) pairs. Whenever a connection request comes in, it could be checked against the table, to see if it belonged to a previously-released connection.

Unfortunately, this scheme has a basic flaw: it requires each transport entity to maintain a certain amount of history information indefinitely. If a machine crashes and loses its memory, it will no longer know which connection identifiers have already been used.

Instead, we need to take a different tack. Rather than allowing packets to live forever within the subnet, we must devise a mechanism to kill off aged packets that are still hobbling about. If we can ensure that no packet lives longer than some known time, the problem becomes somewhat more manageable.

Packet lifetime can be restricted to a known maximum using one (or more) of the following techniques:

1. Restricted subnet design.
2. Putting a hop counter in each packet.
3. Time stamping each packet.

The first method includes any method that prevents packets from looping, combined with some way of bounding congestion delay over the (now known) longest possible path. The second method consists of having the hop count initialized to some appropriate value and decremented each time the packet is forwarded. The third method requires each packet to bear the time it was created, with the routers agreeing to discard any packet older than some agreed-upon time. This latter method requires the router clocks to be synchronized, which itself is a nontrivial task unless synchronization is achieved external to the network, for example by using GPS or some radio station that broadcasts the precise time periodically.

In practice, we will need to guarantee not only that a packet is dead, but also that all acknowledgements to it are also dead, so we will now introduce T , which is some small multiple of the true maximum packet lifetime. The multiple is protocol dependent and simply has the effect of making T longer. If we wait a time T after a packet has been sent, we can be sure that all traces of it are now gone and that neither it nor its acknowledgements will suddenly appear out of the blue to complicate matters.

Tomlinson (1975) introduced the **three-way handshake**. This establishment protocol does not require both sides to begin sending with the same sequence number, so it can be used with synchronization methods other than the global clock method. The normal setup procedure when host 1 initiates is shown in Fig. 6.2. Host 1 chooses a sequence number, x , and sends a CONNECTION REQUEST TPDU containing it to host 2. Host 2 replies with an ACK TPDU acknowledging x and announcing its own initial sequence number, y . Finally, host 1 acknowledges host 2's choice of an initial sequence number in the first data TPDU that it sends.

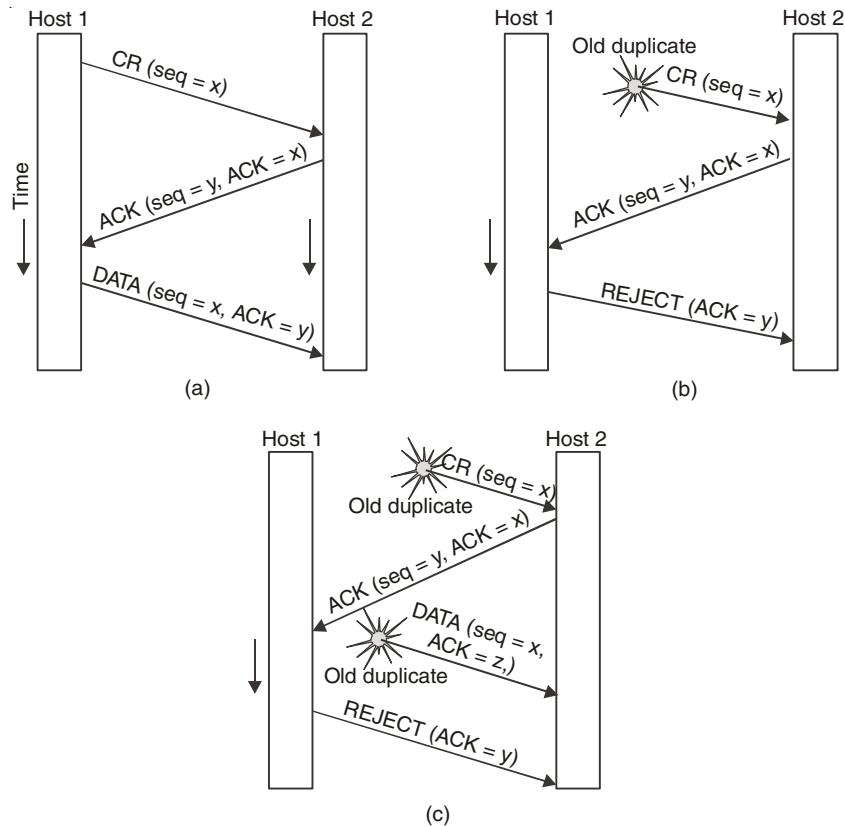


Fig. 6.2 Three protocol scenarios for establishing a connection using a three-way handshake. CR denotes CONNECTION REQUEST. (a) Normal operation. (b) Old duplicate CONNECTION REQUEST appearing out of nowhere. (c) Duplicate CONNECTION REQUEST and duplicate ACK.

Now let us see how the three-way handshake works in the presence of delayed duplicate control TPDUs. In Fig. 6.2 (b), the first TPDU is a delayed duplicate CONNECTION REQUEST from an old connection. This TPDU arrives at host 2 without host 1's knowledge. Host 2 reacts to this TPDU by sending host 1 an ACK TPDU, in effect asking for verification that host 1 was indeed trying to set up a new connection. When host 1 rejects host 2's attempt to establish a connection, host 2 realizes that it was tricked by a delayed duplicate and abandons the connection. In this way, a delayed duplicate does no damage.

The worst case is when both a delayed CONNECTION REQUEST and an ACK are floating around in the subnet. This case is shown in Fig. 6.2 (c). As in the previous example, host 2 gets a delayed CONNECTION REQUEST and replies to it. At this point it is crucial to realize that host 2 has proposed using y as the initial sequence number for host 2 to host 1 traffic, knowing full well that no TPDUs containing sequence number y or acknowledgements to y are still in existence. When the second delayed TPDU arrives at host 2, the fact that z has been acknowledged rather than y tells host 2 that this, too, is an old duplicate. The important thing to realize here is that there is no combination of old TPDUs that can cause the protocol to fail and have a connection set up by accident when no one wants it.

6.3.3 Connection Release

Releasing a connection is easier than establishing one. Nevertheless, there are more pitfalls than one might expect. As we mentioned earlier, there are two styles of terminating a connection: asymmetric release and symmetric release. Asymmetric release is the way the telephone system works: when one party hangs up, the connection is broken. Symmetric release treats the connection as two separate unidirectional connections and requires each one to be released separately.

Asymmetric release is abrupt and may result in data loss. Consider the scenario of Fig. 6.3. After the connection is established, host 1 sends a TPDU that arrives properly at host 2. Then host 1 sends another TPDU. Unfortunately, host 2 issues a DISCONNECT before the second TPDU arrives. The result is that the connection is released and data are lost.

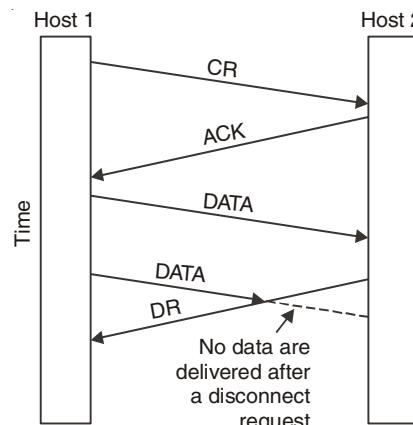


Fig. 6.3. Abrupt disconnect with loss of data.

Clearly, a more sophisticated release protocol is needed to avoid data loss. One way is to use symmetric release, in which each direction is released independently of the other one. Here, a host can continue to receive data even after it has sent a DISCONNECT TPDU. Symmetric release does the job when each process has a fixed amount of data to send and clearly knows when it has sent it. In other situations, determining that all the work has been done and the connection should be terminated is not so obvious.

6.3.4 Multiplexing and Demultiplexing

The addressing mechanism allows multiplexing and demultiplexing by the transport layer, as shown in Fig. 6.4.

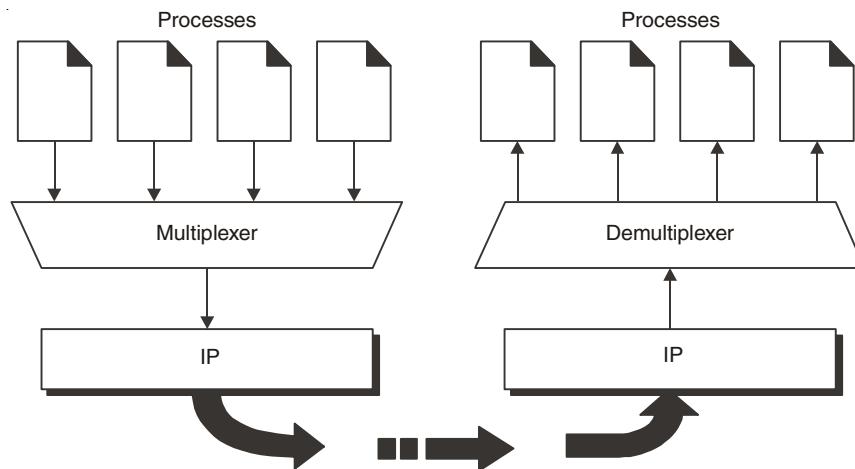


Fig. 6.4. Multiplexing and Demultiplexing.

Multiplexing

At the sender site, there may be several processes that need to send packets. However, there is only one transport layer protocol at any time. This is a many-to-one relationship and requires multiplexing. The protocol accepts messages from different processes, differentiated by their assigned port numbers. After adding the header, the transport layer passes the packet to the network layer.

Demultiplexing

At the receiver site, the relationship is one-to-many and requires demultiplexing. The transport layer receives datagrams from the network layer. After error checking and dropping of the header, the transport layer delivers each message to the appropriate process based on the port number.

6.3.5 Flow Control

Whereas flow control is a relatively simple mechanism at the link layer, it is a rather complex mechanism at the transport layer, for two main reasons:

1. Flow control at the transport level involves the interaction of TS users, transport entities, and the network service.
2. The transmission delay between transport entities is generally long compared to actual transmission time, and, what is worse, it is variable.

If the network service is unreliable, the sender must buffer all TPDUs sent, just as in the data link layer. However, with reliable network service, other trade-offs become possible. In particular, if the sender knows that the receiver always has buffer space, it need not retain

copies of the TPDUs it sends. However, if the receiver cannot guarantee that every incoming TPDU will be accepted, the sender will have to buffer anyway. In the latter case, the sender cannot trust the network layer's acknowledgement, because the acknowledgement means only that the TPDU arrived, not that it was accepted. Even if the receiver has agreed to do the buffering, there still remains the question of the buffer size. If most TPDUs are nearly the same size, it is natural to organize the buffers as a pool of identically-sized buffers, with one TPDU per buffer, as in Fig. 6.5 (a). However, if there is wide variation in TPDU size, from a few characters typed at a terminal to thousands of characters from file transfers, a pool of fixed-sized buffers presents problems. If the buffer size is chosen equal to the largest possible TPDU, space will be wasted whenever a short TPDU arrives. If the buffer size is chosen less than the maximum TPDU size, multiple buffers will be needed for long TPDUs, with the attendant complexity.

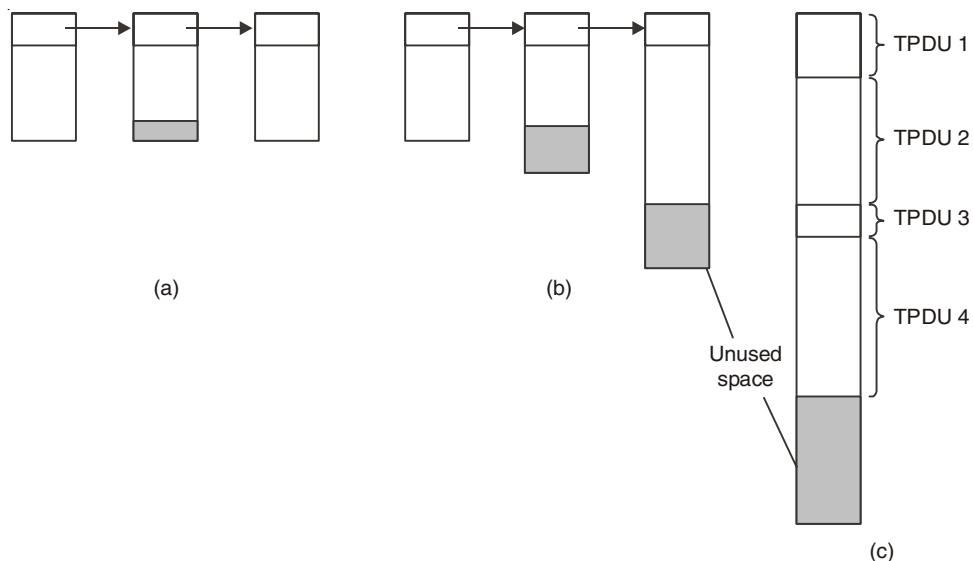


Fig. 6.5. (a) *Chained fixed-size buffers.* (b) *Chained variable-sized buffers.* (c) *One large circular buffer per connection.*

Another approach to the buffer size problem is to use variable-sized buffers, as in Fig. 6.5 (b). The advantage here is better memory utilization, at the price of more complicated buffer management. A third possibility is to dedicate a single large circular buffer per connection, as in Fig. 6.5 (c). This system also makes good use of memory, provided that all connections are heavily loaded, but is poor if some connections are lightly loaded.

The optimum trade-off between source buffering and destination buffering depends on the type of traffic carried by the connection. For low-bandwidth bursty traffic, such as that produced by an interactive terminal, it is better not to dedicate any buffers, but rather to acquire them dynamically at both ends. Since the sender cannot be sure the receiver will be able to acquire a buffer, the sender must retain a copy of the TPDU until it is acknowledged. On the other hand, for file transfer and other high-bandwidth traffic, it is better if the receiver does dedicate a full window of buffers, to allow the data to flow at maximum

speed. Thus, for low-bandwidth bursty traffic, it is better to buffer at the sender, and for high bandwidth smooth traffic, it is better to buffer at the receiver.

As connections are opened and closed and as the traffic pattern changes, the sender and receiver need to dynamically adjust their buffer allocations. Consequently, the transport protocol should allow a sending host to request buffer space at the other end. Buffers could be allocated per connection, or collectively, for all the connections running between the two hosts.

A reasonably general way to manage dynamic buffer allocation is to decouple the buffering from the acknowledgements, in contrast to the sliding window protocols of Chapter 3. Dynamic buffer management means, in effect, a variable-sized window. Initially, the sender requests a certain number of buffers, based on its perceived needs. The receiver then grants as many of these as it can afford. Every time the sender transmits a TPDU, it must decrement its allocation, stopping altogether when the allocation reaches zero. The receiver then separately piggybacks both acknowledgements and buffer allocations onto the reverse traffic.

6.4 THREE PROTOCOLS

The original TCP/IP protocol suite specifies two protocols for the transport layer: UDP and TCP. We first focus on UDP, the simpler of the two, before discussing TCP. As a layered communications protocol, TCP/IP groups functions into defined network layers. Fig. 6.6 illustrates a portion of the TCP/IP protocol suite and indicates the relationship between TCP/IP protocols and the seven-layer OSI Reference Model. The left portion of Fig. 6.6 indicates the seven layers of the ISO's Open System Interconnection Reference Model.

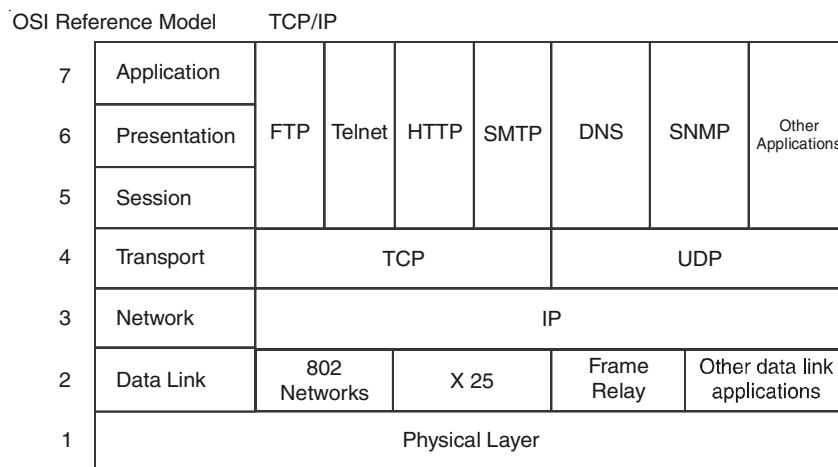


Fig. 6.6. Position of UDP, TCP, in TCP/ IP suite.

6.5 USER DATAGRAM PROTOCOL (UDP)

The simplest possible transport protocol is one that extends the host-to-host delivery service of the underlying network into a process-to-process communication service. There are likely to be many processes running on any given host, so the protocol needs to add a level of demultiplexing, thereby allowing multiple application processes on each host to share the network. Aside from this requirement, the transport protocol adds no other functionality to the best-effort service provided by the underlying network.

The Internet's User Datagram Protocol (UDP) is an example of such a transport protocol. The only interesting issue in such a protocol is the form of the address used to identify the target process. Although it is possible for processes to *directly* identify each other with an OS-assigned process id (pid), such an approach is only practical in a closed distributed system in which a single OS runs on all hosts and assigns each process a unique id. A more common approach, and the one used by UDP, is for processes to *indirectly* identify each other using an abstract locator, often called a *port* or *mailbox*. The basic idea is for a source process to send a message to a port and for the destination process to receive the message from a port.

The header for an end-to-end protocol that implements this demultiplexing function typically contains an identifier (port) for both the sender (source) and the receiver (destination) of the message. The UDP Header is given in Fig. 6.7.

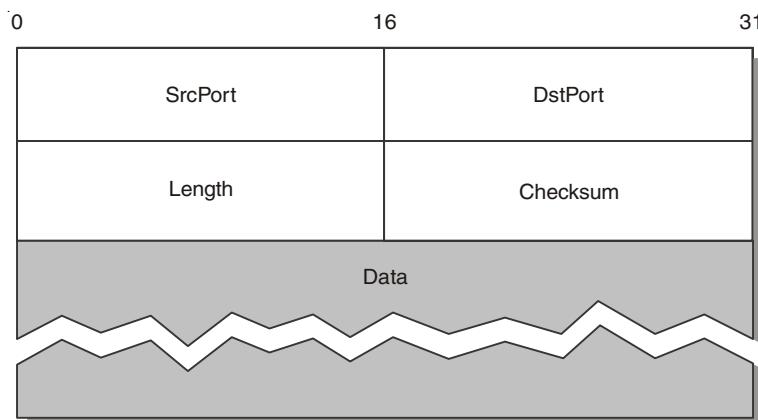


Fig. 6.7. UDP Header.

Source port number. This is the port number used by the process running on the source host. It is 16 bits long, which means that the port number can range from 0 to 65,535. If the source host is the client (a client sending a request), the port number, in most cases, is an ephemeral port number requested by the process and chosen by the UDP software running on the source host. If the source host is the server (a server sending a response), the port number, in most cases, is a well-known port number.

Destination port number. This is the port number used by the process running on the destination host. It is also 16 bits long. If the destination host is the server (a client sending

a request), the port number, in most cases, is a well-known port number. If the destination host is the client (a server sending a response), the port number, in most cases, is an ephemeral port number. In this case, the server copies the ephemeral port number it has received in the request packet.

Length. This is a 16 bit field that defines the total length of the user datagram, header plus data. The 16 bits can define a total length of 0 to 65,535 bytes. However, the total length needs to be much less because a UDP user datagram is stored in an IP datagram with a total length of 65,535 bytes. The length field in a UDP user datagram is actually not necessary. A user datagram is encapsulated in an IP datagram. There is a field in the IP datagram that defines the total length. There is another field in the IP datagram that defines the length of the header. So if we subtract the value of the second field from the first, we can deduce the length of a UDP datagram that is encapsulated in an IP datagram. However, the designers of the UDP protocol felt that it was more efficient for the destination UDP to calculate the length of the data from the information provided in the UDP user datagram rather than ask the IP software to supply this information. We should remember that when the IP software delivers the UDP user datagram to the UDP layer, it has already dropped the IP header.

Checksum. This field is used to detect errors over the entire user datagram (header plus data).

The UDP checksum calculation is different from the one for IP and ICMP. Here the checksum includes three sections: a pseudoheader, the UDP header, and the data coming from the application layer.

The pseudoheader is the part of the header of the IP packet in which the user datagram is to be encapsulated with some fields filled with Os.

32 bit source IP address		
32 bit destination IP address		
A lOs	8 bit protocol (17)	16 bit UDP total length
Source port address 16 bits		Destination port address 16 bits
UDP total length 16 bits		Checksum 16 bits

Fig. 6.8. UDP Pseudo Header.

If the checksum does not include the pseudoheader, a user datagram may arrive safe and sound. However, if the IP header is corrupted, it may be delivered to the wrong host. The protocol field is added to ensure that the packet belongs to UDP, and not to other transport-layer protocols. We will see later that if a process can use either UDP or TCP, the destination port number can be the same. The value of the protocol field for UDP is 17. If

this value is changed during transmission, the checksum calculation at the receiver will detect it and UDP drops the packet. It is not delivered to the wrong protocol. Note the similarities between the pseudoheader fields and the last 12 bytes of the IP header.

6.5.1 UDP Operation

UDP uses concepts common to the transport layer. These concepts will be discussed here briefly, and then expanded in the next section on the TCP protocol.

Connectionless Services

As mentioned previously, UDP provides a connectionless service. This means that each user datagram sent by UDP is an independent datagram. There is no relationship between the different user datagrams even if they are coming from the same source process and going to the same destination program. The user datagrams are not numbered. Also, there is no connection establishment and no connection termination, as is the case for TCP. This means that each user datagram can travel on a different path. One of the ramifications of being connectionless is that the process that uses UDP cannot send a stream of data to UDP and expect UDP to chop them into different related user datagrams. Instead each request must be small enough to fit into one user datagram. Only those processes sending short messages should use UDP.

Flow and Error Control

UDP is a very simple, unreliable transport protocol. There is no flow control and hence no window mechanism. The receiver may overflow with incoming messages. There is no error control mechanism in UDP except for the checksum. This means that the sender does not know if a message has been lost or duplicated. When the receiver detects an error through the checksum, the user datagram is silently discarded. The lack of flow control and error control means that the process using UDP should provide these mechanisms.

Encapsulation and Decapsulation

To send a message from one process to another, the UDP protocol encapsulates and decapsulates messages in an IP datagram.

➤ **Queuing**

A port is purely an abstraction. Exactly how it is implemented differs from system to system, or more precisely, from OS to OS. For example, the socket API is an implementation of ports. Typically, a port is implemented by a message queue, as illustrated in Fig. 6.9. When a message arrives, the protocol (e.g., UDP) appends the message to the end of the queue. Should the queue be full, the message is discarded. There is no flow-control mechanism that tells the sender to slow down. When an application process wants to receive a message, one is removed from the front of the queue. If the queue is empty, the process blocks until a message becomes available.

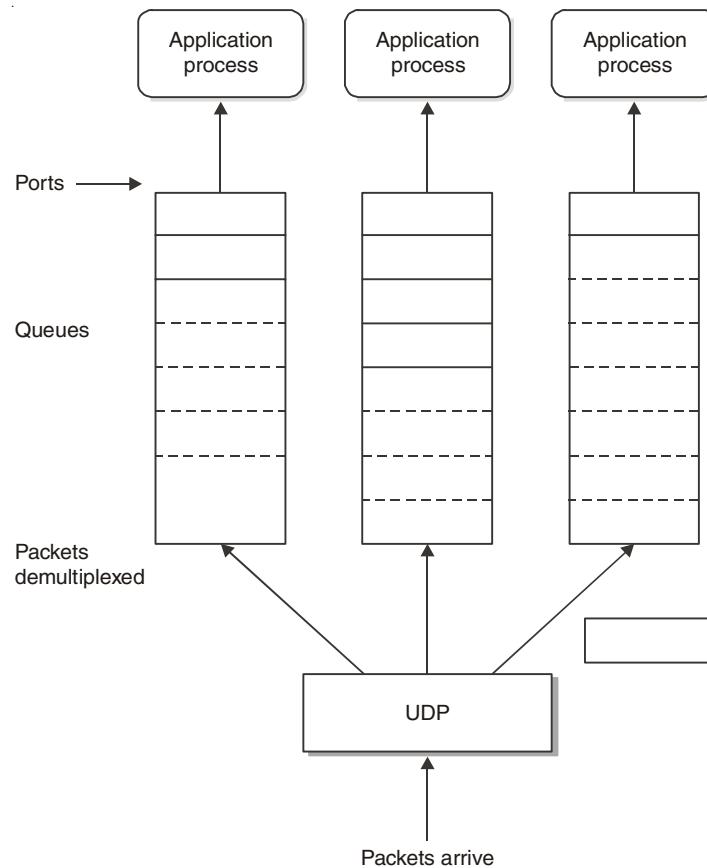


Fig. 6.9. UDP message queue.

The following lists some uses of the UDP protocol:

- UDP is suitable for a process that requires simple request-response communication with little concern for flow and error control. It is not usually used for a process such as FrP that needs to send bulk data.
- UDP is suitable for a process with internal flow and error control mechanisms. For example, the Trivial File Transfer Protocol (TFTP) process includes flow and error control. It can easily use UDP.
- UDP is a suitable transport protocol for multicasting. Multicasting capability is embedded in the UDP software but not in the TCP software.
- UDP is used for management processes such as SNMP.
- UDP is used for some route updating protocols such as Routing Information Protocol (RIP).

6.6 TCP

In contrast to a simple demultiplexing protocol like UDP, a more sophisticated transport protocol is one that offers a reliable, connection-oriented, byte-stream service. Such a service has proven useful to a wide assortment of applications because it frees the application from having to worry about missing or reordered data. The Internet's Transmission Control Protocol (TCP) is probably the most widely used protocol of this type; it is also the most carefully tuned. It is for these two reasons that this section studies TCP in detail, although we identify and discuss alternative design choices at the end of the section.

In terms of the properties of transport protocols given in the problem statement at the start of this chapter, TCP guarantees the reliable, in-order delivery of a stream of bytes. It is a full-duplex protocol, meaning that each TCP connection supports a pair of byte streams, one flowing in each direction. It also includes a flow-control mechanism for each of these byte streams that allows the receiver to limit how much data the sender can transmit at a given time. Finally, like UDP, TCP supports a demultiplexing mechanism that allows multiple application programs on any given host to simultaneously carry on a conversation with their peers. In addition to the above features, TCP also implements a highly tuned congestion-control mechanism. The idea of this mechanism is to throttle how fast TCP sends data, not for the sake of keeping the sender from overrunning the receiver, but to keep the sender from overloading the network.

6.6.1 TCP Services

TCP is designed to provide reliable communication between pairs of processes (TCP users) across a variety of reliable and unreliable networks and internets. Functionally, it is equivalent to Class 4 ISO Transport. In contrast to the ISO model, TCP is stream oriented. That is, TCP users exchange streams of data. The data are placed in allocated buffers and transmitted by TCP in segments. TCP supports security and precedence labeling. In addition, TCP provides two useful facilities for labeling data, push and urgent:

Data stream push. Ordinarily, TCP decides when sufficient data have accumulated to form a segment for transmission. The TCP user can require TCP to transmit all outstanding data up to and including those labeled with a push flag. On the receiving end, TCP will deliver these data to the user in the same manner; a user might request this if it has come to a logical break in the data.

Urgent data signaling. This provides a means of informing the destination TCP user that significant or "urgent" data is in the upcoming data stream. It is up to the destination user to determine appropriate action.

As with IP, the services provided by TCP are defined in terms of primitives and parameters. The services provided by TCP are considerably richer than those provided by IP, and, hence, the set of primitives and parameters is more complex.

Table 6.1 lists TCP service request primitives, which are issued by a TCP user to TCP, and Table 6.2 lists TCP service response primitives, which are issued by TCP to a local TCP user. Table 6.3 provides a brief definition of the parameters involved. Several comments are in order.

Table 6.1 TCP Service Request Parameters

<i>Primitive</i>	<i>Parameters</i>	<i>Description</i>
Unspecified Passive Open	source-port, [timeout], [timeout-action], [precedence], [security-range]	Listen for connection attempt at specified security and precedence from anyremote destination.
Fully Specified Passive Open	source-port, destination-port, destination-address, [timeout], [timeout-action], [precedence], [security-range]	Listen for connection attempt at specified security and precedence from specified destination.
Active Open	source-port, destination-port, destination-address, [timeout], [timeout-action], [precedence], [security]	Request connection at a particular security and precedence to a specified destination.
Active Open with Data	source-port, destination-port, destination-address, [timeout], [timeout-action], [precedence], [security], data, data-length, PUSH-flag, URGENT-flag	Request connection at a particular security and precedence to a specified destination and transmit data with the request.
Send	local-connection-name, data, data-length, PUSH-flat, URGENT-flag, [timeout], [timeout-action]	Transfer data across named connection
Allocate	local-connection-name, data-length	Issue incremental allocation for receive data to TCP
Close	local-connection-name	Close connection gracefully
Abort	local-connection-name	Close connection abruptly
Status	local-connection-name	Query connection status

Note: Square brackets indicate optional parameters.

Table 6.2 TCP Service Report Parameters

<i>Primitive</i>	<i>Parameters</i>	<i>Description</i>
Open ID	local-connection-name, source-port destination-port*, destination-address*	Informs TCP user of connection name assigned to pending connection requested in an Open primitive
Open Failure	local-connection-name	Reports failure of an Active Open request
Open Success	local-connection-name	Reports completion of pending Open request
Deliver	local-connection-name, data, data-length, URGENT-flag	Reports arrival of data

Closing	local-connection-name	Reports that remote TCP user has issued a Close and that all data sent by remote user have been delivered
Terminate	local-connection-name, description	Reports that the connection has been terminated; a description of the reason for termination is provided
Status	local-connection-name, source-port,	Reports current status of connection
Response	source-address, destination-port, destination-address, connection-state, receive-window, send-window, amount- awaiting-ACK, amount-awaiting-receipt, urgent-state, precedence, security, timeout	
Error	local-connection-name, description	Reports service-request or internal error

* = Not used for Unspecified Passive Open.

Table 6.3 TCP Service Parameters

Source Port	Local TCP user.
Timeout	Longest delay allowed for data delivery before automatic connection termination or error report; user specified.
Timeout-action	Indicates whether the connection is terminated or an error is reported to the TCP user in the event of a timeout.
Precedence	Precedence level for a connection. Takes on values zero (lowest) through seven (highest); same parameter as defined for IP
Security-range	Allowed ranges in compartment, handling restriction, transmission control codes, and security levels.
Destination Port	Remote TCP user.
Destination Address	Internet address of remove host.
Security	Security information for a connection, including security level, compartment, handling restriction, and transmission control code; same parameter as defined for IP.
Data	Block of data sent by TCP user or delivered to a TCP user.
Data Length	Length of block of data sent or delivered.
PUSH flag	If set, indicates that the associated data are to be provided with the urgent data stream push service.
URGENT flag	If set, indicates that the associated data are to be provided with the urgent data signaling service.
Local Connection Name	Identifier of a connection defined by a (local socket, remote socket) pair; provided by TCP.
Description	Supplementary information in a Terminate or Error primitive.

Source Address	Internet address of the local host.
Connection State	State of referenced connection (CLOSED, ACTIVE OPEN, PASSIVE OPEN, ESTABLISHED, CLOSING).
Receive Window	Amount of data in octets the local TCP entity is willing to receive.
Send Window	Amount of data in octets permitted to be sent to remote TCP entity.
Amount Awaiting ACK	Amount of previously transmitted data awaiting acknowledgement.
Amount Awaiting Receipt	Amount of data in octets buffered at local TCP entity pending receipt by local TCP user.
Urgent State	Indicates to the receiving TCP user whether there are urgent data available or whether all urgent data, if any, have been delivered to the user.

The two Passive Open commands signal the TCP user's willingness to accept a connection request. The Active Open with Data allows the user to begin transmitting data with the opening of the connection.

6.6.2 TCP Header Format

TCP uses only a single type of protocol data unit, called a TCP segment. The header is shown in Fig. 6.10 . Because one header must serve to perform all protocol mechanisms, it is rather large, with a minimum length of 20 octets. The fields are:

- **Source port (16 bits).** Source service access point.
- **Destination port (16 bits).** Destination service access point.
- **Sequence number (32 bits).** Sequence number of the first data octet in this segment except when SYN flag is set. If SYN is set, it is the initial sequence number (ISN), and the first data octet is ISN + 1.
- **Acknowledgement number (32 bits).** A piggybacked acknowledgement. Contains the sequence number of the next data octet that the TCP entity expects to receive.
- **Data offset (4 bits).** Number of 32-bit words in the header.
- **Reserved (6 bits).** Reserved for future use.
- **Flags (6 bits).**

URG: Urgent pointer field significant.

ACK: Acknowledgement field significant.

PSH: Push function.

RST: Reset the connection.

SYN: Synchronize the sequence numbers.

FIN: No more data from sender.

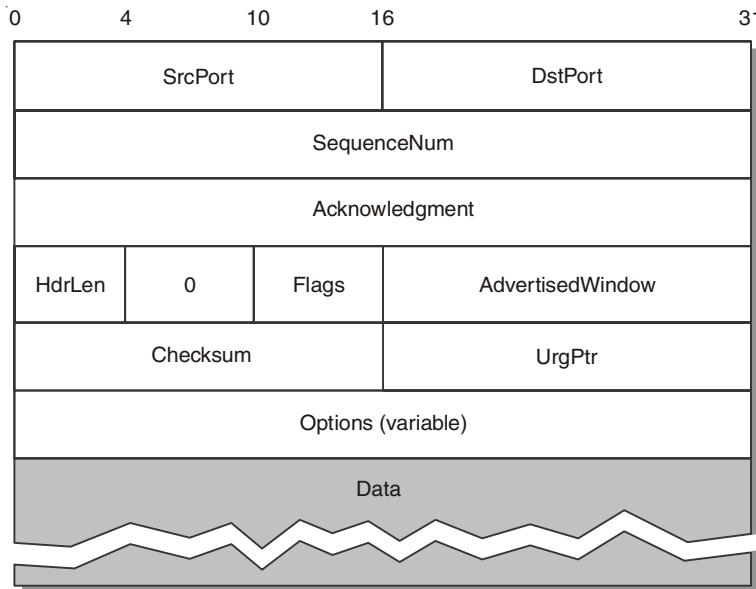


Fig. 6.10. TCP Header Format.

- **Window (16 bits).** Flow control credit allocation, in octets. Contains the number of data octets beginning with the one indicated in the acknowledgement field that the sender is willing to accept.
- **Checksum (16 bits).** The one's complement of the sum modulo 2¹⁶-1 of all the 16 bit words in the segment, plus a pseudo-header. The situation is described below.
- **Urgent Pointer (16 bits).** Points to the last octet in a sequence of urgent data; this allows the receiver to know how much urgent data is coming.
- **Options (Variable).** At present, only one option is defined, which specifies the maximum segment size that will be accepted.

Several of the fields in the TCP header warrant further elaboration. The *source port* and *destination port* specify the sending and receiving users of TCP. As with IP, there are a number of common users of TCP that have been assigned numbers; these numbers should be reserved for that purpose in any implementation. Other port numbers must be arranged by agreement between communicating parties.

The *sequence number* and *acknowledgment number* are bound to octets rather than to entire segments. For example, if a segment contains sequence number 1000 and includes 600 octets of data, the sequence number refers to the first octet in the data field; the next segment in logical order will have sequence number 1600. Thus, TCP is logically stream-oriented: It accepts a stream of octets from the user, groups them into segments as it sees fit, and numbers each octet in the stream.

The *checksum* field applies to the entire segment, plus a pseudo-header prefixed to the header at the time of calculation (at both transmission and reception). The pseudo-header includes the following fields from the IP header: source and destination internet address

and protocol, plus a segment length field. By including the pseudo-header, TCP protects itself from mis delivery by IP. That is, if IP delivers a segment to the wrong host, even if the segment contains no bit errors, the receiving TCP entity will detect the delivery error. If TCP is used over IPv6, then the pseudo-header is different.

6.6.3 TCP Mechanisms

A TCP connection begins with a client (caller) doing an active open to a server (callee). Assuming that the server had earlier done a passive open, the two sides engage in an exchange of messages to establish the connection. Only after this connection establishment phase is over do the two sides begin sending data. Likewise, as soon as a participant is done sending data, it closes one direction of the connection, which causes TCP to initiate a round of connection termination messages. Notice that while connection setup is an asymmetric activity (one side does a passive open and the other side does an active open), connection teardown is symmetric (each side has to close the connection independently). Therefore, it is possible for one side to have done a close, meaning that it can no longer send data, but for the other side to keep the other half of the bidirectional connection open and to continue sending data.

➤ Three-Way Handshake

The algorithm used by TCP to establish and terminate a connection is called a *three-way handshake*. We first describe the basic algorithm and then show how it is used by TCP. The three-way handshake involves the exchange of three messages between the client and the server, as illustrated by the timeline given in Fig. 6.11.

The idea is that two parties want to agree on a set of parameters, which, in the case of opening a TCP connection, are the starting sequence numbers the two sides plan to use for their respective byte streams. In general, the parameters might be any facts that each side wants the other to know about. First, the client (the active participant) sends a segment to the server (the passive participant) stating the initial sequence number it plans to use (Flags = SYN, Sequence Num = x). The server then responds with a single segment that both acknowledges the client's sequence number (Flags = ACK, Ack = $x + 1$) and states its own beginning sequence number (Flags = SYN, Sequence Num = y). That is, both the SYN and ACK bits are set in the Flags field of this second message. Finally, the client responds with a third segment that acknowledges the server's sequence number (Flags = ACK, Ack = $y + 1$). The reason that each side acknowledges a sequence number that is one larger than the one sent is that the Acknowledgement field actually identifies the "next sequence number expected," thereby implicitly acknowledging all earlier sequence numbers. Although not shown in this timeline, a timer is scheduled for each of the first two segments, and if the expected response is not received, the segment is retransmitted.

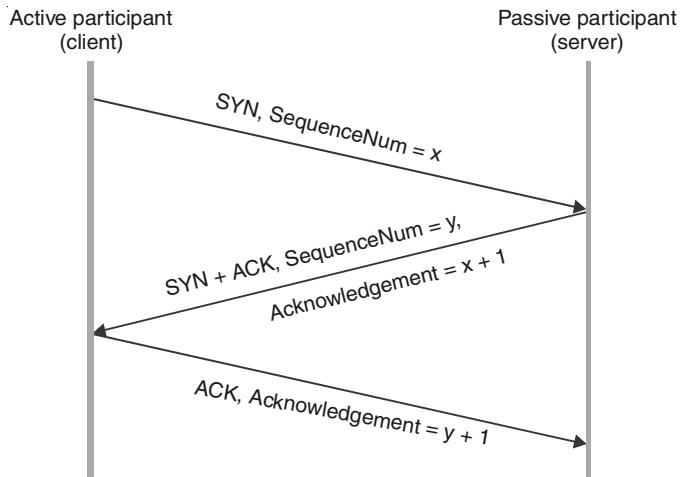


Fig. 6.11. Timeline for three way Hand Shake mechanism.

You may be asking yourself why the client and server have to exchange starting sequence numbers with each other at connection setup time. It would be simpler if each side simply started at some “well-known” sequence number, such as 0. In fact the TCP specification requires that each side of a connection select an initial starting sequence number at random. The reason for this is to protect against two incarnations of the same connection reusing the same sequence numbers too soon, that is, while there is still a chance that a segment from an earlier incarnation of a connection might interfere with a later incarnation of the connection.

6.6.4 State Transition Diagram

TCP is complex enough that its specification includes a state transition diagram. A copy of this diagram is given in Fig. 6.12. This diagram shows only the states involved in opening a connection (everything above ESTABLISHED) and in closing a connection (everything below ESTABLISHED). Everything that goes on while a connection is open—that is, the operation of the sliding window algorithm—is hidden in the ESTABLISHED state.

TCP’s state transition diagram is fairly easy to understand. Each circle denotes a state that one end of a TCP connection can find itself in. All connections start in the CLOSED state. As the connection progresses, the connection moves from state to state according to the arcs. Each arc is labelled with a tag of the form *event/action*. Thus, if a connection is in the LISTEN state and a SYN segment arrives (*i.e.*, a segment with the SYN flag set), the connection makes a transition to the SYN RCVD state and takes the action of replying with an ACK + SYN segment.

Notice that two kinds of events trigger a state transition: (1) a segment arrive from the peer (*e.g.*, the event on the arc from LISTEN to SYN RCVD), or (2) the local application process invokes an operation on TCP (*e.g.*, the *active open* event on the arc from CLOSE to SYN SENT). In other words, TCP’s state transition diagram effectively defines the *semantics* of both its peer-to-peer interface and its service interface.

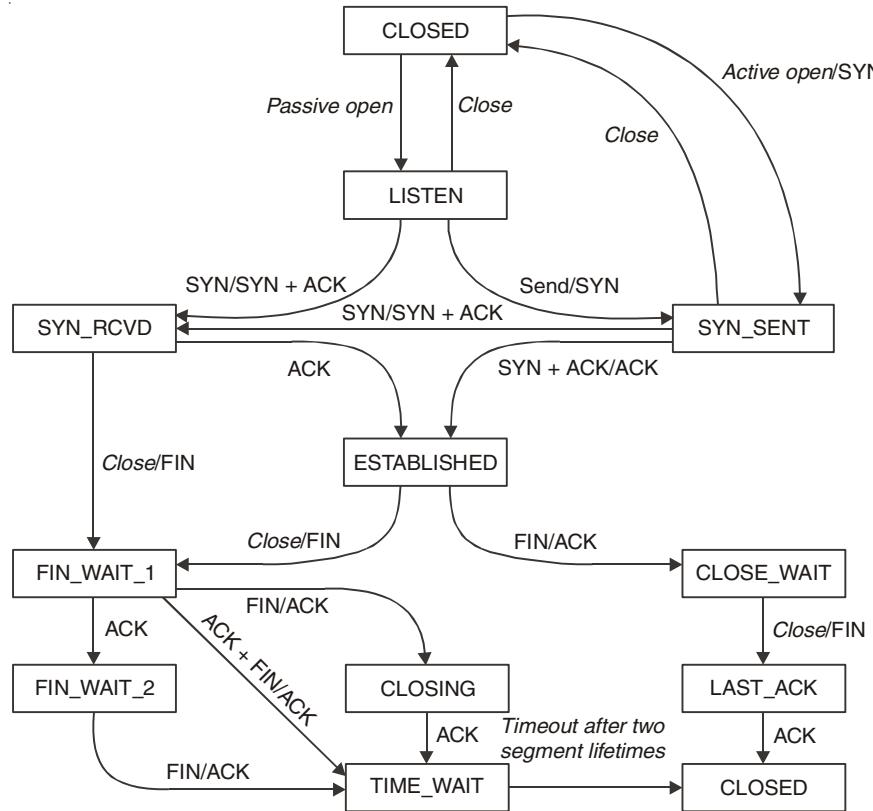


Fig. 6.12. State Transmission in TCP.

Now let's trace the typical transitions taken through the diagram in Fig. 6.12. Keep in mind that at each end of the connection, TCP makes different transitions from state to state. When opening a connection, the server first invokes a passive open operation on TCP, which causes TCP to move to the LISTEN state. At some later time, the client does an active open, which causes its end of the connection to send a SYN segment to the server and to move to the SYN SENT state. When the SYN segment arrives at the server, it moves to the SYN RCVD state and responds with a SYN+ACK segment. The arrival of this segment causes the client to move to the ESTABLISHED state and to send an ACK back to the server. When this ACK arrives, the server finally moves to the ESTABLISHED state. In other words, we have just traced the three-way handshake.

There are three things to notice about the connection establishment half of the state transition diagram. First, if the client's ACK to the server is lost, corresponding to the third leg of the three-way handshake, then the connection still functions correctly. This is because the client side is already in the ESTABLISHED state, so the local application process can start sending data to the other end. Each of these data segments will have the ACK flag set, and the correct value in the Acknowledgement field, so the server will move to the ESTABLISHED state when the first data segment arrives. This is actually an important point about TCP—every segment reports what sequence number the sender is expecting to see next, even if this repeats the same sequence number contained in one or more previous segments.

The second thing to notice about the state transition diagram is that there is a funny transition out of the LISTEN state whenever the local process invokes a *send* operation on TCP. That is, it is possible for a passive participant to identify both ends of the connection (*i.e.*, itself and the remote participant that it is willing to have connect to it), and then to change its mind about waiting for the other side and instead actively establish the connection. To the best of our knowledge, this is a feature of TCP that no application process actually takes advantage of the final thing to notice about the diagram is the arcs that are not shown. Specifically, most of the states that involve sending a segment to the other side also schedule a timeout that eventually causes the segment to be resent if the expected response does not happen. These retransmissions are not depicted in the state transition diagram. If after several tries the expected response does not arrive, TCP gives up and returns to the CLOSED state.

Turning our attention now to the process of terminating a connection, the important thing to keep in mind is that the application process on both sides of the connection must independently close its half of the connection. If only one side closes the connection, then this means it has no more data to send, but it is still available to receive data from the other side. This complicates the state transition diagram because it must account for the possibility that the two sides invoke the *close* operator at the same time, as well as the possibility that first one side invokes close and then, at some later time, the other side invokes close. Thus, on any one side there are three combinations of transitions that get a connection from the ESTABLISHED state to the CLOSED state:

- This side closes first:

ESTABLISHED → FIN WAIT 1 → FIN WAIT 2 → TIME WAIT → CLOSED.

- The other side closes first:

ESTABLISHED → CLOSE WAIT → LAST ACK → CLOSED.

- Both sides close at the same time:

ESTABLISHED → FIN WAIT 1 → CLOSING → TIME WAIT → CLOSED.

There is actually a fourth, although rare, sequence of transitions that leads to the CLOSED state; it follows the arc from FIN WAIT 1 to TIME WAIT.

The main thing to recognize about connection teardown is that a connection in the TIME WAIT state cannot move to the CLOSED state until it has waited for two times the maximum amount of time an IP datagram might live in the Internet (*i.e.*, 120 seconds). The reason for this is that while the local side of the connection has sent an ACK in response to the other side's FIN segment, it does not know that the ACK was successfully delivered. As a consequence, the other side might retransmit its FIN segment, and this second FIN segment might be delayed in the network. If the connection were allowed to move directly to the CLOSED state, then another pair of application processes might come along and open the same connection (*i.e.*, use the same pair of port numbers), and the delayed FIN segment from the earlier incarnation of the connection would immediately initiate the termination of the later incarnation of that connection.

6.7 NETWORK SECURITY AT TRANSPORT LAYER

Although several protocols are found in Transport layer for network security , we are only going to discuss two: Secure Socket Layer (SSL) and Transport Layer Security (TLS). Currently, however, these two are no longer considered as two separate protocols but one under the name SSL/TLS, after the SSL standardization was passed over to IETF, by the Netscape consortium, and Internet Engineering Task Force (IETF) renamed it TLS. Figure 6.13 shows the position of these protocols in the network protocol stack.

6.7.1 Secure Socket Layer (SSL)

SSL is a widely used general purpose cryptographic system used in the two major Internet browsers: Netscape and Explorer. It was designed to provide an encrypted end-to-end data path between a client and a server regardless of platform or OS. Secure and authenticated services are provided through data encryption, server authentication, message integrity, and client authentication for a TCP connection through HTTP, LDAP, or POP3 application layers. It was originally developed by Netscape Communications and it first appeared in a Netscape Navigator browser in 1994. The year 1994 was an interesting year for Internet security because during the same year, a rival security scheme to SSL, the S-HTTP, was launched. Both systems were designed for Web-based commerce. Both allow for the exchange of multiple messages between two processes and use similar cryptographic schemes such as digital envelopes, signed certificates, and message digest.

Although these two Web giants had much in common, there are some differences in design goals, implementation, and acceptance. First, S-HTTP was designed to work with only Web protocols. Because SSL is at a lower level in the network stack than S-HTTP, it can work in many other network protocols. Second, in terms of implementation, since SSL is again at a lower level than S-HTTP, it is implemented as a replacement for the sockets API to be used by applications requiring secure communications. On the other hand, S-HTTP has its data passed in named text fields in the HTTP header. Finally in terms of distribution and acceptance, history has not been so good to S-HTTP. While SSL was released in a free mass circulating browser, the Netscape Navigator, S-HTTP was released in a much smaller and restricted NCSA Mosaic. This unfortunate choice doomed the fortunes of S-HTTP.

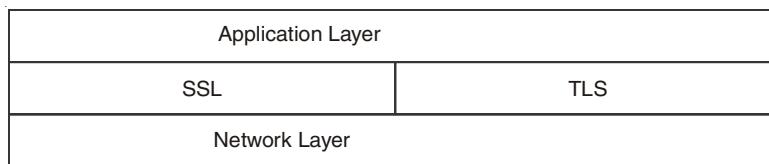


Fig. 6.13. Transport Layer Security Protocols.

The stated SSL objectives were to secure and authenticate data paths between servers and clients. These objectives were to be achieved through several services that included data encryption, server and client authentication, and message integrity:

- **Data encryption** – to protect data in transport between the client and the server from interception and could be read only by the intended recipient.

and clients. These objectives were to be achieved through several services that included data encryption, server and client authentication, and message integrity:

- **Data encryption** – to protect data in transport between the client and the server from interception and could be read only by the intended recipient.
- **Server and client authentication** – the SSL protocol uses standard public key encryption to authenticate the communicating parties to each other.
- **Message integrity** – achieved through the use of session keys so that data cannot be either intentionally or unintentionally tampered with.

These services offer reliable end-to-end secure services to Internet TCP connections and are based on an SSL architecture consisting of two layers: the top layer, just below the application layer, that consists of three protocols, namely the SSL Handshake protocol, the SS Change Cipher Specs Protocol, and the SSL Alert protocol. Below these protocols is the second SSL layer, the SSL Record Protocol layer, just above the TCP layer. See Fig. 6.14.

The SSL Handshake

Before any TCP connection between a client and a server, both running under SSL, is established, there must be almost a process similar to a three-way handshake we discussed in Section 6.6. This get-to-know-you process is similarly called the SSL handshake.

During the handshake, the client and server perform the following tasks:

- Establish a cipher suite to use between them.
- Provide mandatory server authentication through the server sending its certificate to the client to verify that the server's certificate was signed by a trusted CA.
- Provide optional client authentication, if required, through the client sending its own certificate to the server to verify that the client's certificate was signed by a trusted CA. The CA may not be the same CA who signed the client's certificate. CAs may come from a list of trusted CAs. The reason for making this step optional was a result of realization that since few customers are willing, know how, or care to get digital certificates, requiring them to do this would amount to locking a huge number of customers out of the system which would not make business sense. This, however, presents some weaknesses to the system.
- Exchange key information using public key cryptography, after mutual authentication, that leads to the client generating a session key (usually a random number) which, with the negotiated cipher, is used in all subsequent encryption or decryption. The customer encrypts the session key using the public key of the merchant server (from the merchant's certificate). The server recovers the session key by decrypting it using its private key. This symmetric key, which now both parties have, is used in all subsequent communication. For details in security Algorithms refer to Chapter 8.

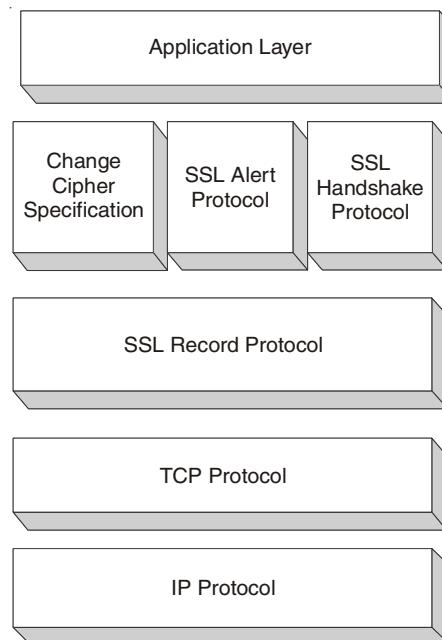


Fig. 6.14. SSL Protocol Stack.

SSL Cipher Specs Protocol

The SSL Cipher Specs protocol consists of an exchange of a single message in a byte with a value of 1 being exchanged, using the SSL record protocol, between the server and client. The bit is exchanged to establish a pending session state to be copied into the current state, thus defining a new set of protocols as the new agreed on session state.

SSL Alert Protocol

The SSL Alert protocol, which also runs over the SSL Record protocol, is used by the two parties to convey session warning messages associated with data exchange and functioning of the protocol. The warnings are used to indicate session problems ranging from unknown certificate, revoked certificate, and expired certificate to fatal error messages that can cause immediate termination of the SSL connection. Each message in the alert protocol sits within two bytes, with the first byte taking a value of (1) for a warning and (2) for a fatal error. The second byte of the message contains one of the defined error codes that may occur during an SSL communication session.

SSL Record Protocol

The SSL record protocol provides SSL connections two services: confidentiality and message integrity:

- **Confidentiality** is attained when the handshake protocol provides a shared secret key used in the conventional encryption of SSL messages.
- **Message integrity** is attained when the handshake defines a secret shared key used to form a message authentication code (MAC).

In providing these services, the SSL Record Protocol takes an application message to be transmitted and fragments the data that needs to be sent, compresses it, adds a MAC, encrypts it together with the MAC, adds an SSL header, and transmits it under the TCP protocol. The return trip undoes these steps. The received data is decrypted, verified, and decompressed before it is forwarded to higher layers. The record header that is added to each data portion contains two elementary pieces of information, namely, the length of the record and the length of the data block added to the original data. See Fig. 6.15.

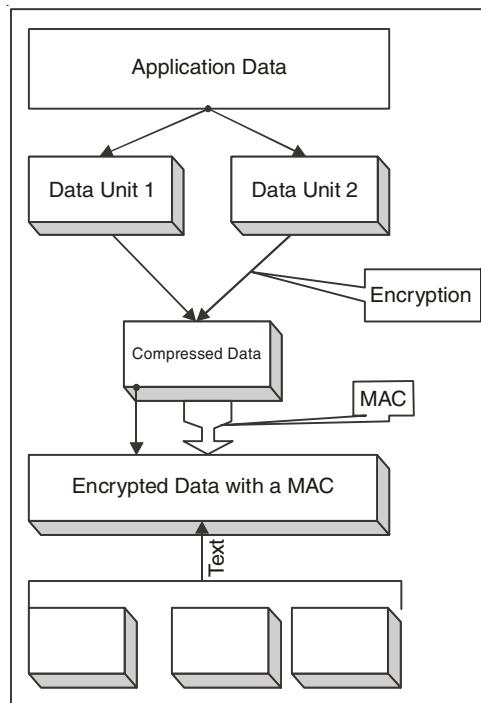


Fig. 6.15. SSL Record Protocol operation process.

SSL protocols are widely used in all Web applications and any other TCP connections. Although they are mostly used for Web applications, they are gaining ground in e-mail applications also.

6.7.2 Transport Layer Security (TLS)

Transport Layer Security (TLS) is the result of the 1996 Internet Engineering Task Force (IETF) attempt at standardization of a secure method to communicate over the Web. The 1999 outcome of that attempt was released as RFC 2246 spelling out a new protocol – the Transport Layer Security or TLS. TLS was charged with providing security and data integrity at the transport layer between two applications. TLS version 1.0 was an evolved SSL 3.0. So, as we pointed out earlier, TLS is the successor to SSL 3.0. Frequently, the new standard is referred to as SSL/TLS. Since then, however, the following additional features have been added.

- **Interoperability** – ability to exchange TLS parameters by either party, with no need for one party to know the other's TLS implementation details.
- **Expandability** – to plan for future expansions and accommodation of new protocols.

Review Questions

1. In cases where reliability is not of primary importance, UDP would make a good transport protocol. Give examples of specific cases.
2. Are both UDP and IP unreliable to the same degree? Why or why not?
3. Do port addresses need to be unique? Why or why not? Why are port addresses shorter than IP addresses?
4. What is the dictionary definition of the word *ephemeral*? How does it apply to the concept of the ephemeral port number?
5. What is the minimum size of a UDP datagram?
6. What is the maximum size of a UDP datagram?
7. What is the minimum size of the process data that can be encapsulated in a UDP datagram?
8. What is the maximum size of the process data that can be encapsulated in a UDP datagram?
9. Compare the TCP header and the UDP header. List the fields in the TCP header that are missing from UDP header. Give the reason for their absence.
10. UDP is a message-oriented protocol. TCP is a byte-oriented protocol. If an application needs to protect the boundaries of its message, which protocol should be used, UDP or TCP?
11. What can you say about the TCP segment in which the value of the control field is one of the following?
 - (a) 000000
 - (b) 000001
 - (c) 010001
12. What is the maximum size of the TCP header? What is the minimum size of the TCP header?
13. Highlight various security issues in Transport layer.



CHAPTER 7

THE APPLICATION LAYER

Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning

—Winston Churchill

7.1 INTRODUCTION

The application layer enables the user, whether human or software, to access the network. It provides user interfaces and support for services such as electronic mail, file access and transfer, access to system resources, surfing the world wide web, and network management. However, even in the application layer there is a need for support protocols, to allow the applications to function. All of the protocols and functions described so far are geared toward one objective: the support of distributed applications that involve the interaction of multiple independent systems. In the OSI model, such applications occupy the application layer and are directly supported by the presentation layer. In the TCP/IP protocol suite, such applications typically rely on TCP or UDP for support.

In this chapter, we will examine a number of quite different applications that give the reader a feel for the range and diversity of applications supported by a communications architecture.

7.2 DOMAIN NAME SYSTEM

Up to this point, we have been using addresses to identify hosts. While perfectly suited for processing by routers, addresses are not exactly user friendly. It is for this reason that a unique *name* is also typically assigned to each host in a network. This section describes how a naming service can be developed to map user-friendly names into router-friendly addresses. Such a service is often the first application program implemented in a network since it frees

other applications to identify hosts by name rather than by address. Name services are sometimes called *middleware* because they fill a gap between applications and the underlying network.

Host names differ from host addresses in two important ways. First, they are usually of variable length and mnemonic, thereby making them easier for humans to remember. (In contrast, fixed-length numeric addresses are easier for routers to process.) Second, names typically contain no information that helps the network locate (route packets toward) the host. Addresses, in contrast, sometimes have routing information embedded in them; *flat* addresses (those not divisible into component parts) are the exception.

Before getting into the details of how hosts are named in a network, we first introduce some basic terminology. First, a *name space* defines the set of possible names. A name space can be either *flat* (names are not divisible into components) or *hierarchical* (Unix file names are the obvious example).

Second, the naming system maintains a collection of *bindings* of names to values. The value can be anything we want the naming system to return when presented with a name; in many cases it is an address.

Finally, a *resolution mechanism* is a procedure that, when invoked with a name, returns the corresponding value. A *name server* is a specific implementation of a resolution mechanism that is available on a network and that can be queried by sending it a message.

Because of its large size, the Internet has a particularly well-developed naming system in place—the *domain name system* (DNS). We therefore use DNS as a framework for discussing the problem of naming hosts. Note that the Internet did not always use DNS. Early in its history, when there were only a few hundred hosts on the Internet, a central authority called the Network Information Center (NIC) maintained a flat table of name-to-address bindings; this table was called hosts.txt. Whenever a site wanted to add a new host to the Internet, the site administrator sent email to the NIC giving the new host's name/address pair. This information was manually entered into the table, the modified table was mailed out to the various sites every few days, and the system administrator at each site installed the table on every host at the site. Name resolution was then simply implemented by a procedure that looked up a host's name in the local copy of the table and returned the corresponding address.

It should come as no surprise that the hosts.txt approach to naming did not work well as the number of hosts in the Internet started to grow. Therefore, in the mid-1980s, the domain naming system was put into place. DNS employs a hierarchical name space rather than a flat name space, and the “table” of bindings that implements this name space is partitioned into disjoint pieces and distributed throughout the Internet. These sub-tables are made available in name servers that can be queried over the network.

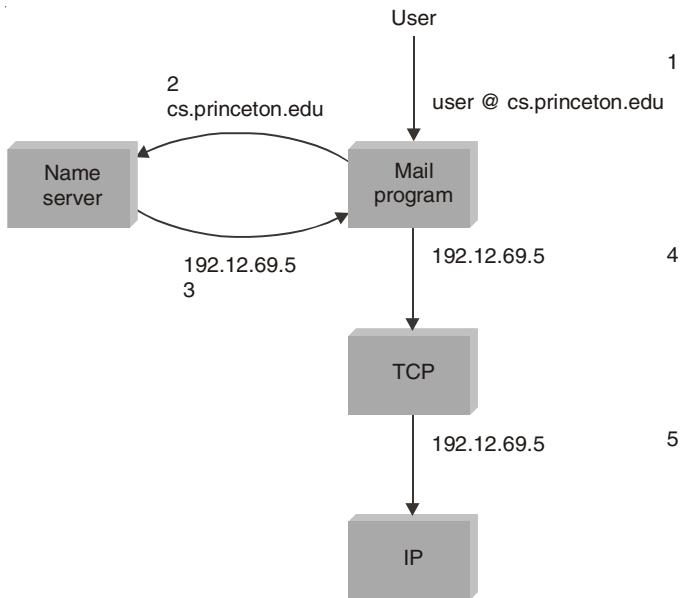


Fig. 7.1. Names translated into addresses, where the numbers 1–5 show the sequence of steps in the process.

What happens in the Internet is that a user presents a host name to an application program (possibly embedded in a compound name such as an email address or URL), and this program engages the naming system to translate this name into a host address. The application then opens a connection to this host by presenting some transport protocol (e.g., TCP) with the host’s IP address. This situation is illustrated (in the case of sending email) in Fig. 7.1.

7.2.1 Domain Hierarchy

DNS implements a hierarchical name space for Internet objects. Unlike Unix file names, which are processed from left to right with the naming components separated with slashes, DNS names are processed from right to left and use periods as the separator. (Although they are “processed” from right to left, humans still “read” domain names from left to right.) An example domain name for a host is `academicia.cs.princeton.edu`. Notice that we said domain names are used to name Internet “objects.” What we mean by this is that DNS is not strictly used to map host names into host addresses. It is more accurate to say that DNS maps domain names into values.

Like the Unix file hierarchy, the DNS hierarchy can be visualized as a tree, where each node in the tree corresponds to a domain and the leaves in the tree correspond to the hosts being named. Figure 7.2 gives an example of a domain hierarchy. Note that we should not assign any semantics to the term “domain” other than that it is simply a context in which additional names can be defined.

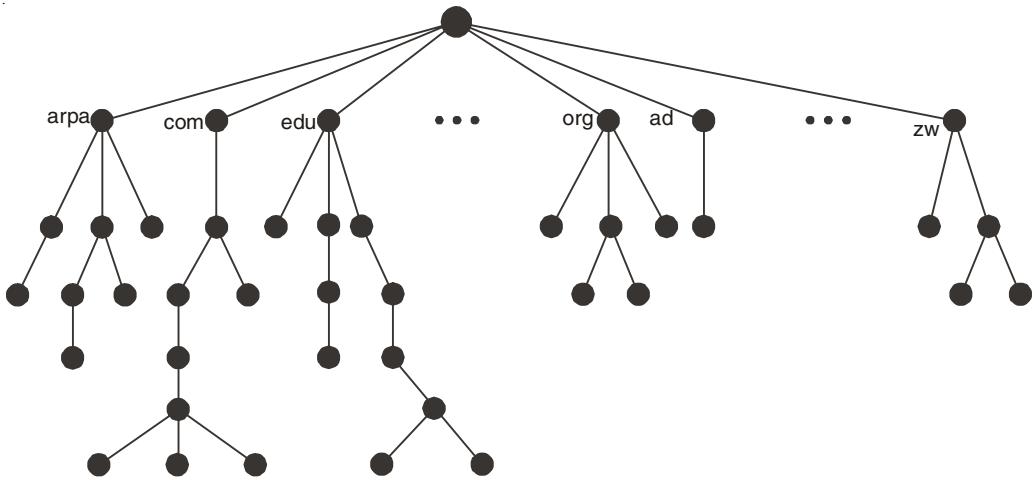


Fig. 7.2. Domain Hierarchy.

Since the complete domain name hierarchy cannot be stored on a single server, it is divided among many servers. What a server is responsible for or has authority over is called a zone. We can define a zone as a contiguous part of the entire tree. If a server accepts responsibility for a domain and does not divide the domain into smaller domains, the *domain* and the *zone* refer to the same thing. The server makes a database called a *zone file* and keeps all the information for every node under that domain. However, if a server divides its domain into sub domains and delegates part of its authority to other servers, *domain* and *zone* refer to different things. The information about the nodes in the sub domains is stored in the servers at the lower levels, with the original server keeping some sort of reference to these lower-level servers. Of course the original server does not free itself from responsibility totally: It still has a zone, but the detailed information is kept by the lower-level servers.

7.2.2 Resource Records

Every domain, whether it is a single host or a top-level domain, can have a set of *resource records* associated with it. For a single host, the most common resource record is just its IP address, but many other kinds of resource records also exist. When a resolver gives a domain name to DNS, what it gets back are the resource records associated with that name. Thus, the primary function of DNS is to map domain names onto resource records.

A resource record is a five-tuple. Although they are encoded in binary for efficiency, in most expositions, resource records are presented as ASCII text, one line per resource record. The format we will use is as follows:

Domain_name	Time_to_live	Class	Type	Value
-------------	--------------	-------	------	-------

The *Domain_name* tells the domain to which this record applies. Normally, many records exist for each domain and each copy of the database holds information about multiple domains. This field is thus the primary search key used to satisfy queries. The order of the

records in the database is not significant.

The *Time_to_live* field gives an indication of how stable the record is. Information that is highly stable is assigned a large value, such as 86400 (the number of seconds in 1 day). Information that is highly volatile is assigned a small value, such as 60 (1 minute).

The third field of every resource record is the *Class*. For Internet information, it is always *IN*. For non-Internet information, other codes can be used, but in practice, these are rarely seen.

The *Type* field tells what kind of record this is. The most important types are listed in Fig. 7.3.

Type	Meaning	Value
SOA	Start of Authority	Parameters for this zone
A	IP address of a host	32-Bit integer
MX	Mail exchange	Priority, domain willing to accept e-mail
NS	Name Server	Name of a server for this domain
CNAME	Canonical name	Domain name
PTR	Pointer	Alias for an IP address
HINFO	Host description	CUP and OS in ASCII
TXT	Text	Uninterpreted ASCII text

Fig. 7.3. The principal DNS resource record types for IPv4.

7.2.3 DNS in the Internet

DNS is a protocol that can be used in different platforms. In the Internet, the domain name space (tree) is divided into three different sections: generic domains, country domains, and the inverse domain.

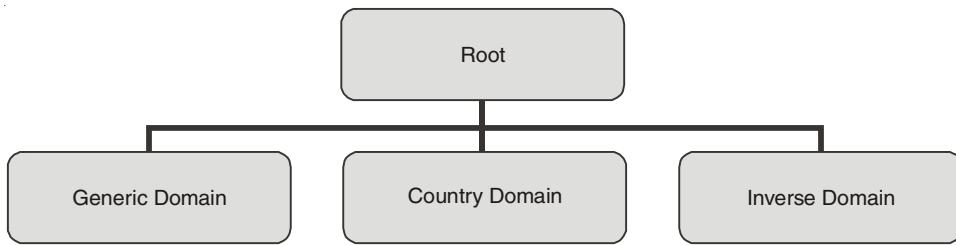


Fig. 7.4. DNS used in the Internet.

Generic Domains

The generic domains define registered hosts according to their generic behavior. Each node in the tree defines a domain, which is an index to the domain name space database.

Country Domains

The country domains section uses two-character country abbreviations (*e.g.*, in for India).

Second labels can be organizational, or they can be more specific, national designations.

Inverse Domain

The inverse domain is used to map an address to a name. This may happen, for example, when a server has received a request from a client to do a task. Although the server has a file that contains a list of authorized clients, only the IP address of the client (extracted from the received IP packet) is listed. The server asks its resolver to send a query to the DNS server to map an address to a name to determine if the client is on the authorized list.

This type of query is called an inverse or pointer (PTR) query. To handle a pointer query, the inverse domain is added to the domain name space with the first-level node called *arpa* (for historical reasons). The second level is also one single node named *in-addr* (for inverse address). The rest of the domain defines IP addresses.

The servers that handle the inverse domain are also hierarchical. This means the net id part of the address should be at a higher level than the subnetid part, and the subnetid part higher than the hostid part. In this way, a server serving the whole site is at a higher level than the servers serving each subnet. This configuration makes the domain look inverted when compared to a generic or country domain. To follow the convention of reading the domain labels from the bottom to the top, an IP address such as 132.34.45.121 (a class B address with netid 132.34) is read as 121.45.34.132.in-addr. arpa.

7.3 ELECTRONIC MAIL (SMTP, MIME, IMAP)

The most heavily used application in virtually any distributed system is electronic mail. Electronic mail, or **e-mail**, as it is known to its many fans, has been around for over two decades. Before 1990, it was mostly used in academia. During the 1990s, it became known to the public at large and grew exponentially to the point where the number of e-mails sent per day now is vastly more than the number of **snail mail** (*i.e.*, paper) letters.

E-mail, like most other forms of communication, has its own conventions and styles. In particular, it is very informal and has a low threshold of use. People who would never dream of calling up or even writing a letter to a Very Important Person do not hesitate for a second to send a sloppily-written e-mail. The first e-mail systems simply consisted of file transfer protocols, with the convention that the first line of each message (*i.e.*, file) contained the recipient's address. As time went on, the limitations of this approach became more obvious.

Sending a message to a group of people was inconvenient. Managers often need this facility to send memos to all their subordinates. Messages had no internal structure, making computer processing difficult. For example, if a forwarded message was included in the body of another message, extracting the forwarded part from the received message was difficult. The originator (sender) never knew if a message arrived or not. The user interface was poorly integrated with the transmission system requiring users first to edit a file, then leave the editor and invoke the file transfer program. It was not possible to create and send messages containing a mixture of text, drawings, facsimile, and voice. As experience was gained, more elaborate e-mail systems were proposed. In 1982, the ARPANET e-mail proposals were published as RFC 821 (transmission protocol) and RFC 822 (message format). Minor revisions, RFC 2821 and RFC 2822, have become Internet standards, but everyone still

From the start, the Simple Mail Transfer Protocol (SMTP) has been the workhorse of the TCP/IP protocol suite. However, SMTP has traditionally been limited to the delivery of simple text messages. In recent years, there has been a demand for the delivery mail to be able to contain various types of data, including voice, images, and video clips. To satisfy this requirement, a new electronic mail standard, which builds on SMTP, has been defined: the Multi-Purpose Internet Mail Extension (MIME). In this section, we first examine SMTP, then look at MIME.

7.3.1 Architecture and Services

In this section we will provide an overview of what e-mail systems can do and how they are organized. They normally consist of two subsystems: the user agents, which allow people to read and send e-mail, and the message transfer agents, which move the messages from the source to the destination.

➤ Services Provided by a User Agent

A user agent is a software package (program) that composes, reads, replies to, and forwards messages. It also handles mailboxes. Fig. 7.5 shows the services of a typical user agent.

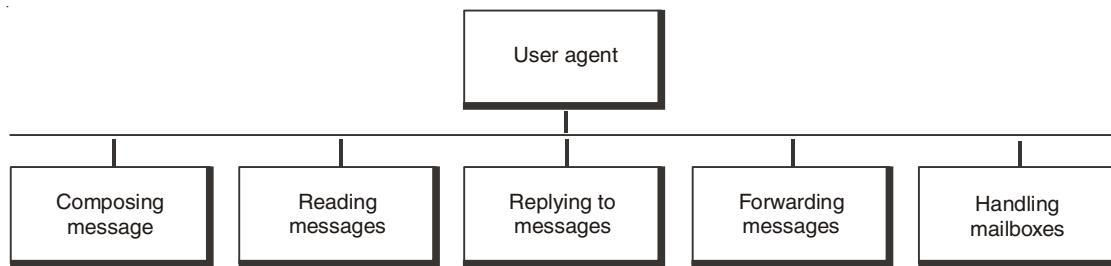


Fig. 7.5. Services provided by User Agent.

Composing Messages. A user agent helps the user compose the e-mail message to be sent out. Most user agents provide a template on the screen to be filled in by the user. Some even have a built-in editor that can do spell checking, grammar checking, and other tasks expected from a sophisticated word processor. A user, of course, could alternatively use his or her favorite text editor or word processor to create the message and import it, or cut and paste it, into the user agent template.

Reading Messages. The second duty of the user agent is to read the incoming messages. When a user invokes a user agent, it first checks the mail in the incoming mailbox. Most user agents show a one-line summary of each received mail. Each e-mail contains the following fields.

1. A number field.
2. A flag field that shows the status of the mail such as new, already read but not replied to, or read and replied to.
3. The size of the message.
4. The sender.
5. The optional subject field.

Replies to Messages. After reading a message, a user can use the user agent to reply to a message. A user agent usually allows the user to reply to the original sender or to reply to all recipients of the message. The reply message may contain the original message (for quick reference) and the new message.

Forwarding Messages. *Replying* is defined as sending a message to the sender or recipients of the copy. *Forwarding* is defined as sending the message to a third party. A user agent allows the receiver to forward the message, with or without extra comments, to a third party.

Handling Mailboxes. A user agent normally creates two mailboxes: an inbox and an outbox. Each box is a file with a special format that can be handled by the user agent. The inbox keeps all the received e-mails until they are deleted by the user. The outbox keeps all the sent e-mails until the user deletes them. Most user agents today are capable of creating customized mailboxes.

User Agent Types. There are two types of user agents: command-driven and GUI-based. Command-driven user agents belong to the early days of electronic mail. They are still present as the underlying user agents in servers. A command-driven user agent normally accepts a one-character command from the keyboard to perform its task. For example, a user can type the character r, at the command prompt, to reply to the sender of the message, or type the character R to reply to the sender and all recipients. Some examples of command-driven user agents are *mail*, *pine*, and *elm*. Modern user agents are GUI-based. They contain

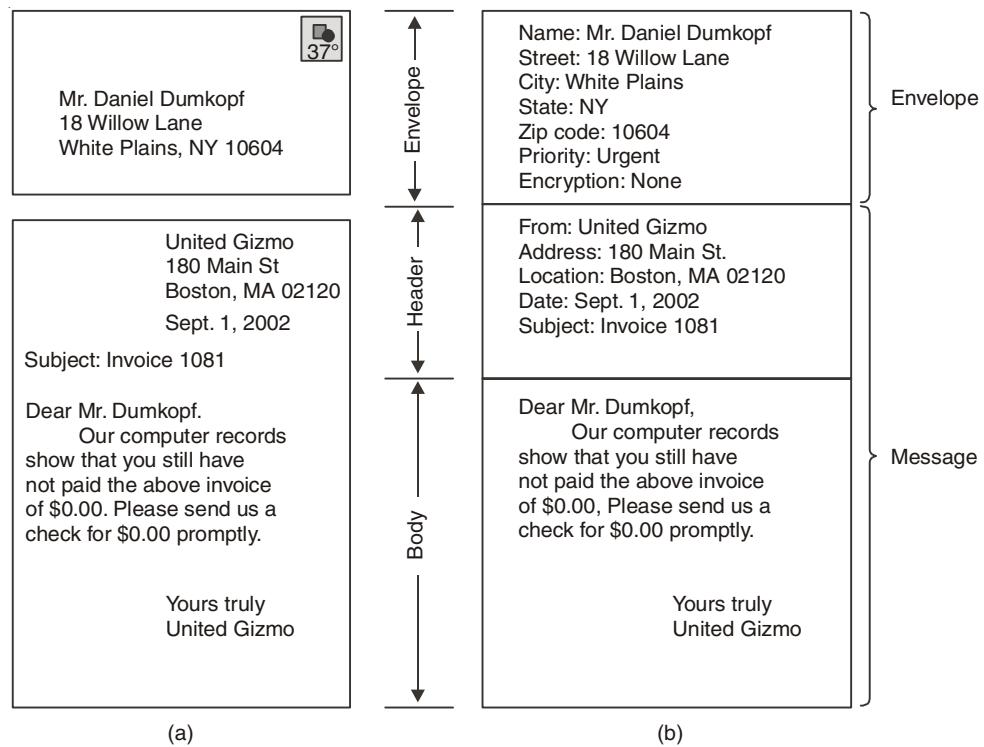


Fig. 7.6. Envelopes and messages. (a) Paper mail. (b) Electronic mail.

graphical user interface (GUI) components that allow the user to interact with the software by using both the keyboard and the mouse. They have graphical components such as icons, menu, bars etc.

Sending Mail. To send mail, the user, through the UA, creates mail that looks very similar to postal mail. It has an *envelope* and a *message*.

The envelope usually contains the sender and the receiver addresses. The message contains the header and the body. The header of the message defines the sender, the receiver, the subject of the message, and some other information (such as encoding type, as we see shortly). The body of the message contains the actual information to be read by the recipient.

Receiving Mail. The user agent is triggered by the user (or a timer). If a user has mail, the VA informs the user with a notice. If the user is ready to read the mail a list is displayed in which each line contains a summary of the information about a particular message in the mailbox. The summary usually includes the sender mail address, the subject, and the time the mail was sent or received. The user can select any of the messages and display its contents on the screen.

Addresses. To deliver mail, a mail handling system must use an addressing system with unique addresses. In the Internet, the address consists of two parts: a local part and a domain name, separated by an @ sign. The local part defines the name of a special file, called the user mailbox, where all the mail received for a user is stored for retrieval by the message access agent. The second part of the address is the domain name. An organization usually selects one or more hosts to receive and send e-mail; the hosts are sometimes called *mail servers* or *exchangers*. The domain name assigned to each mail exchanger either comes from the DNS database or is a logical name (for example, the name of the organization).

7.3.2 MIME—The Multipurpose Internet Mail Extensions

In the early days of the ARPANET, e-mail consisted exclusively of text messages written in English and expressed in ASCII. For this environment, RFC 822 did the job completely: it specified the headers but left the content entirely up to the users. Nowadays, on the worldwide Internet, this approach is no longer adequate. The problems include sending and receiving:

1. Messages in languages with accents (e.g., French and German).
2. Messages in non-Latin alphabets (e.g., Hebrew and Russian).
3. Messages in languages without alphabets (e.g., Chinese and Japanese).
4. Messages not containing text at all (e.g., audio or images).

A solution was proposed in RFC 1341 and updated in RFCs 2045–2049. This solution, called **MIME (Multipurpose Internet Mail Extensions)** is now widely used. We will now describe it. For additional information about MIME, see the RFCs.

The basic idea of MIME is to continue to use the RFC 822 format, but to add structure to the message body and define encoding rules for non-ASCII messages. By not deviating from RFC 822, MIME messages can be sent using the existing mail programs and protocols. All that has to be changed are the sending and receiving programs, which users can do for themselves.

MIME defines five new message headers, as shown in Fig. 7.8. The first of these simply tells the user agent receiving the message that it is dealing with a MIME message, and which version of MIME it uses. Any message not containing a *MIME-Version:* header is assumed to be an English plaintext message and is processed as such.

Header	Meaning
Date:	The date and time the message was sent
Reply-To:	E-mail address to which replies should be sent
Message-Id:	Unique number for referencing this message later
In-Reply-To:	Message-Id of the message to which this is a reply
References:	Other relevant Message-Ids
Keywords:	User-chosen keywords
Subject:	Short summary of the message for the one-line display

Fig. 7.7. Some fields used in the RFC 822 message header.

The Content-Description. header is an ASCII string telling what is in the message. This header is needed so the recipient will know whether it is worth decoding and reading the message. If the string says: "Photo of Barbara's hamster" and the person getting the message is not a big hamster fan, the message will probably be discarded rather than decoded into a high-resolution color photograph.

The Content-Id. header identifies the content. It uses the same format as the standard *Message-Id:* header.

The Content-Transfer-Encoding. tells how the body is wrapped for transmission through a network that may object to most characters other than letters, numbers, and punctuation marks. Five schemes (plus an escape to new schemes) are provided. The simplest scheme is just ASCII text. ASCII characters use 7 bits and can be carried directly by the e-mail protocol provided that no line exceeds 1000 characters.

Header	Meaning
MIME-Version:	Identifies the MIME version
Content-Description:	Human-readable string telling what is in the message
Content-Id:	Unique identifier
Content-Transfer-Encoding:	How the body is wrapped for transmission
Content-Type	Type and format of the content

Fig. 7.8. RFC 822 headers added by MIME.

The next simplest scheme is the same thing, but using 8 bit characters, that is, all values from 0 up to and including 255. This encoding scheme violates the (original) Internet e-mail protocol but is used by some parts of the Internet that implement some extensions to the original protocol. While declaring the encoding does not make it legal, having it explicit may at least explain things when something goes wrong. Messages using the 8 bit encoding must still adhere to the standard maximum line length.

Even worse are messages that use binary encoding. These are arbitrary binary files that not only use all 8 bits but also do not even respect the 1000-character line limit. Executable programs fall into this category. No guarantee is given that messages in binary will arrive correctly, but some people try anyway.

The correct way to encode binary messages is to use **base64 encoding**, sometimes called **ASCII armor**. In this scheme, groups of 24 bits are broken up into four 6-bit units, with each unit being sent as a legal ASCII character. The coding is "A" for 0, "B" for 1, and so on, followed by the 26 lower-case letters, the ten digits, and finally + and / for 62 and 63, respectively. The == and = sequences indicate that the last group contained only 8 or 16 bits, respectively. Carriage returns and line feeds are ignored, so they can be inserted at will to keep the lines short enough. Arbitrary binary text can be sent safely using this scheme.

In summary, binary data should be sent encoded in base64 or quoted-printable form. When there are valid reasons not to use one of these schemes, it is possible to specify a user-defined encoding in the *Content-Transfer-Encoding:* header.

The last header shown in Fig. 7.8 is really the most interesting one. It specifies the nature of the message body. Seven types are defined in RFC 2045, each of which has one or more subtypes. The type and subtype are separated by a slash, as in

Content-Type: video/mpeg

Type	Subtype	Description
Text	Plain	Unformatted text
	Enriched	Text including simple formatting commands
Image	Gif	Still picture in GIF format
	Jpeg	Still picture in JPEG format
Audio	Basic	Audible sound
Video	Mpeg	Movie in MPEG format
Application	Octet-stream	An uninterpreted byte sequence
	postscript	A printable document in PostScript
Message	Rfc822	A MIME RFC 822 message
	Partial	Message has been split for transmission
	External-body	Message itself must be fetched over the net
Multipart	Mixed	Independent parts in the specified order
	Alternative	Same message in different formats
	Parallel	Parts must be viewed simultaneously
	Digest	Each part is a complete RFC 822 message

Fig. 7.9. The MIME types and subtypes defined in RFC 2045.

7.3.3 Message Transfer Agent: SMTP

The actual mail transfer is done through message transfer agents. To send mail, a system must have the client MTA, and to receive mail, a system must have a server MTA. The formal protocol that defines the MTA client and server in the Internet is called the Simple

Mail Transfer Protocol (SMTP). As we said before, two pairs of MTA client/server programs are used in the most common situation (fourth scenario). Fig. 7.10 shows the range of the SMTP protocol in this scenario.

SMTP is used two times, between the sender and the sender's mail server and between the two mail servers. As we will see shortly, another protocol is needed between the mail server and the receiver. SMTP simply defines how commands and responses must be sent back and forth. Each network is free to choose a software package for implementation. We discuss the mechanism of mail transfer by SMTP in the remainder of the section.

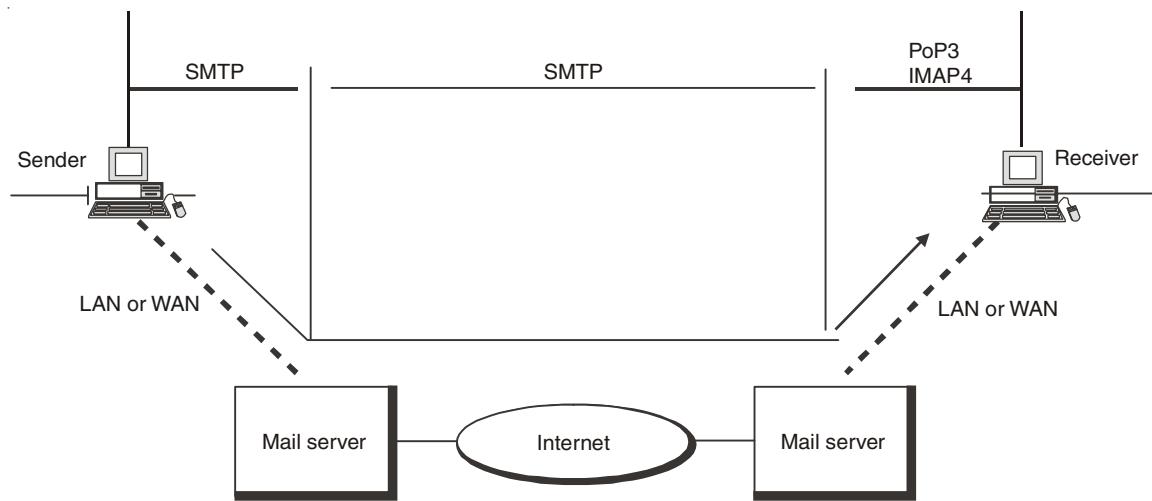


Fig. 7.10. SMTP range.

➤ *Commands and Responses*

SMTP uses commands and responses to transfer messages between an MTA client and an MTA server.



Fig. 7.11. Command and Responses.

Each command or reply is terminated by a two-character (carriage return and line feed) end-of-line token. Commands are sent from the client to the server. It consists of a keyword followed by zero or more arguments. SMTP defines 14 commands. The first five are mandatory; every implementation must support these five commands. The next three are often used and highly recommended. The last six are seldom used. Responses are sent from the server to the client. A response is a three digit code that may be followed by additional textual information.

The process of transferring a mail message occurs in three phases: connection establishment, mail transfer, and connection termination.

Keyword	Argument(s)
HELO	Sender's host name
MAIL FROM	Sender of the message
RCPT TO	Intended recipient of the message
DATA	Body of the mail
QUIT	
RESET	
VRFY	Name of recipient to be verified
NOOP	
TURN	
EXPN	Mailing list to be expanded
HELP	Command name

Fig. 7.12. Some of the Commands.

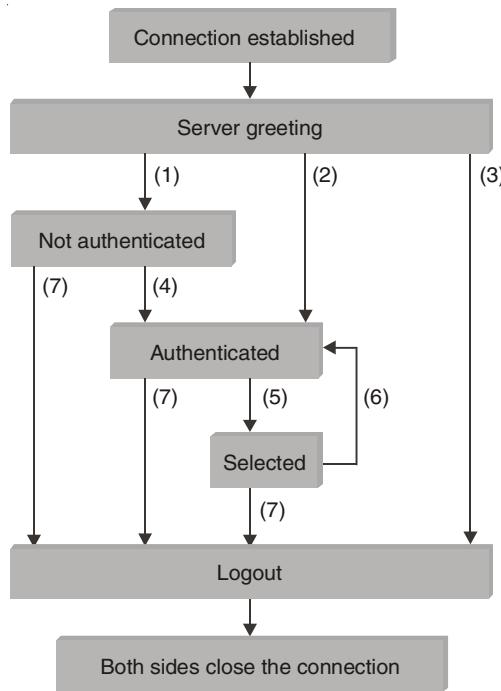
7.3.4 Message Access Agent: IMAP

The final step is for the user to actually retrieve his or her messages from the mailbox, read them, reply to them, and possibly save a copy for future reference. The user performs all these actions by interacting with a mail reader. In many cases, this reader is just a program running on the same machine as the user's mailbox resides, in which case it simply reads and writes the file that implements the mailbox. In other cases, the user accesses his or her mailbox from a remote machine using yet another protocol, such as the Post Office Protocol (POP) or the Internet Message Access Protocol (IMAP). It is beyond the scope of this book to discuss the user interface aspects of the mail reader, but it is definitely within our scope to talk about the access protocol. We consider IMAP, in particular.

IMAP is similar to SMTP in many ways. It is a client/server protocol running over TCP, where the client (running on the user's desktop machine) issues commands in the form of <CRLF> terminated ASCII text lines and the mail server (running on the machine that maintains the user's mailbox) responds in kind. The exchange begins with the client authenticating him or herself, and identifying the mailbox he or she wants to access. This can be represented by the simple state transition diagram shown in Fig. 7.13. In this diagram, LOGIN, AUTHENTICATE, SELECT, EXAMINE, CLOSE, and LOGOUT are example commands that the client can issue, while OK is one possible server response. Other common commands include FETCH, STORE, DELETE, and EXPUNGE, with the obvious meanings. Additional server responses include NO (client does not have permission to perform that operation) and BAD (command is ill formed).

When the user asks to `FETCH` a message, the server returns it in `MIME` format and the mail reader decodes it. In addition to the message itself, IMAP also defines a set of message *attributes* that are exchanged as part of other commands, independent of transferring the message itself. Message attributes include information like the size of the message, but more interestingly, various *flags* associated with the message (*e.g.*, `Seen`, `Answered`, `Deleted`, and `Recent`). These flags are used to keep the client and server synchronized; that is, when the user deletes a message in the mail reader, the client needs to report this fact to the mail server. Later, should the user decide to `expunge` all deleted messages, the client issues an `EXPUNGE` command to the server, which knows to actually remove all earlier deleted messages from the mailbox.

Finally, note that when the user replies to a message, or sends a new message, the mail reader does not forward the message from the client's desktop machine to the mail server using IMAP, but it instead uses SMTP. This means that the user's mail server is effectively the first mail gateway traversed along the path from the desktop to the recipient's mailbox.



- (1) connection without preauthentication (OK greeting)
- (2) preauthenticated connection (PREAUTH greeting)
- (3) rejected connection (BYE greeting)
- (4) successful LOGIN or AUTHENTICATE command
- (5) successful SELECT or EXAMINE command
- (6) CLOSE command, or failed SELECT or EXAMINE command.
- (7) LOGOUT command, server shutdown, or connection close

Fig. 7.13. IMAP State Transition.

7.4 TELNET

Telnet represents an interactive remote access terminal protocol developed to enable users to log into a remote computer as if their terminal was directly connected to the distant computer. Several flavors of telnet have been developed, including TN3270 which is designed to support telnet access to IBM mainframes. TN3270 primarily differs from telnet in that it recognizes the screen control codes generated by IBM mainframes. Otherwise, the use of telnet to access an IBM mainframe would more than likely result in the display of what appears to be garbage on the terminal device. Telnet and TN3270 use a common TCP connection to transmit both data and control information.

Figure 7.14 illustrates the Net Manage Chameleon Telnet application after the application's Connect menu has been selected. The resulting dialog box labeled Connection Dialog enables a user to enter the host address that they wish to establish a telnet session with. In addition, this GUI version of telnet supports eight terminal types which are selected through the Emulate pull down menu. In Fig. 7.14 port number 23 was selected, which is the default port used for telnet communications.

The Chameleon suite of TCP/IP applications to include FTP and Telnet were originally developed for Windows 3.1, which did not include a TCP/IP protocol stack. When Microsoft added support for TCP/IP in Windows 95 and succeeding versions of the Windows operating system it also included several TCP/IP applications. One such application is FTP. Another application is Telnet. Unlike FTP, Microsoft included a GUI-based Telnet application in all releases of Windows after version 3.1.

Figure 7.15 illustrates the Microsoft Telnet application after the Remote System entry from its Connect menu was selected. Similar to the Chameleon Telnet application, you

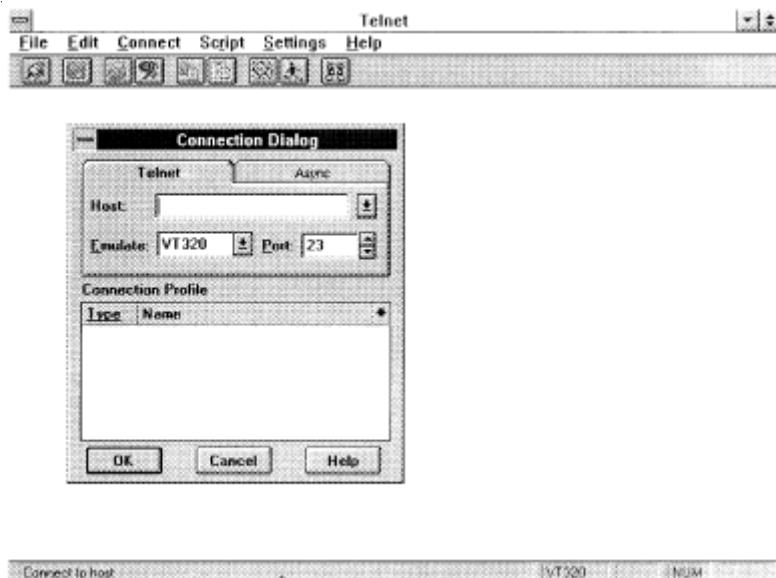


Fig. 7.14. Preparing to initiate a telnet session using the Net Manage Chameleon Telnet application.

specify a host name, port and terminal type, with a list of terminal types supported by the Microsoft version of Telnet pulled down in the illustration. Note that the Microsoft version of Telnet supports five terminals and its port configuration is limited to predefined mnemonics, such as Telnet, which results in port 23 being used. Other ports supported by the Microsoft version of Telnet include daytime, echo, gold and chargen.

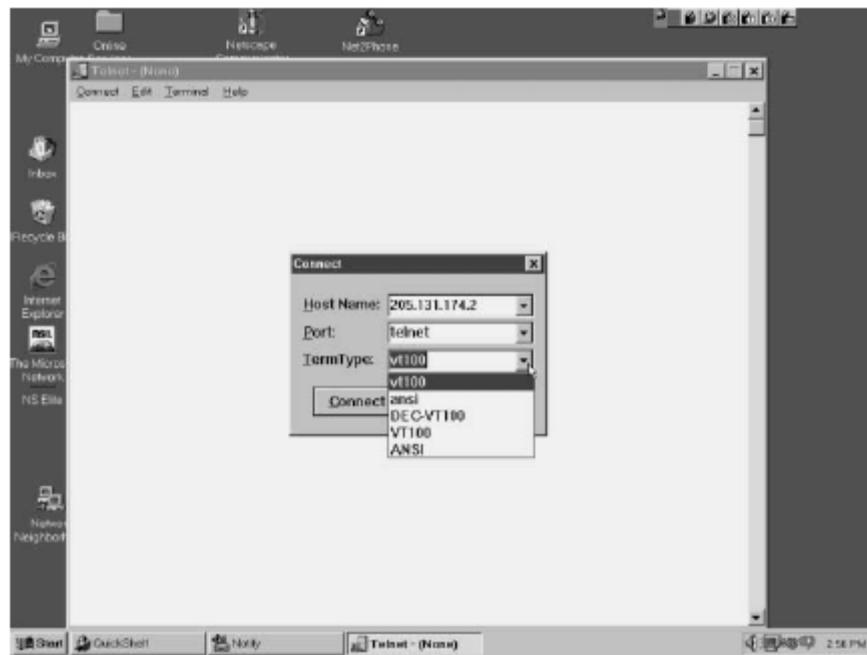


Fig. 7.15. The Microsoft Windows Telnet client supports a limited number of predefined ports, such as Telnet for port 23 use.

7.5 WORLD WIDE WEB

The **World Wide Web** (WWW) is a repository of information linked together from points all over the world. The WWW has a unique combination of flexibility, portability, and user-friendly features that distinguish it from other services provided by the Internet. The WWW project was initiated by CERN (European Laboratory for Particle Physics) to create a system to handle distributed resources necessary for scientific research. In this section we first discuss issues related to the Web. We then discuss a protocol, HTTP, that is used to retrieve information from the Web.

7.5.1 Architecture

The WWW today is a distributed client/server service, in which a client using a browser can access a service using a server. However, the service provided is distributed over many locations called *sites*, as shown in Fig. 7.16.

Each site holds one or more documents, referred to as *Web pages*. Each Web page can contain a link to other pages in the same site or at other sites. The pages can be retrieved and viewed by using browsers. Let us go through the scenario shown in Fig. 7.16. The client needs to see some information that it knows belongs to site A. It sends a request through its browser, a program that is designed to fetch Web documents. The request, among other information, includes the address of the site and the Web page, called the URL, which we will discuss shortly. The server at site A finds the document and sends it to the client. When the user views the document, she finds some references to other documents, including a Web page at site B. The reference has the URL for the new site. The user is also interested in seeing this document. The client sends another request to the new site, and the new page is retrieved.

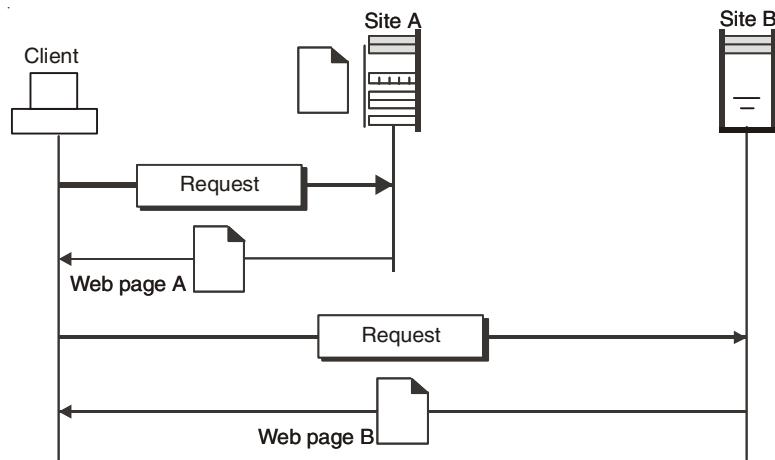


Fig. 7.16. Architecture of world wide web.

Client (Browser)

A variety of vendors offer commercial browsers that interpret and display a Web document, and all use nearly the same architecture. Each browser usually consists of three parts: a controller, client protocol, and interpreters. The controller receives input from the keyboard or the mouse and uses the client programs to access the document. After the document has been accessed, the controller uses one of the interpreters to display the document on the screen. The client protocol can be one of the protocols such as FTP or HTTP. The interpreter can be HTML, Java, or JavaScript, depending on the type of document.

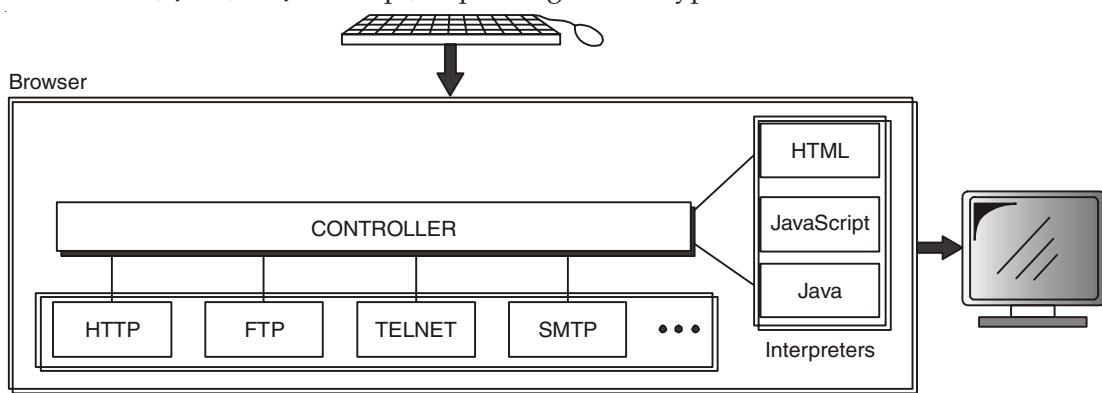


Fig. 7.17. Browser.

Server

The Web page is stored at the server. Each time a client request arrives, the corresponding document is sent to the client. To improve efficiency, servers normally store requested files in a cache in memory; memory is faster to access than disk. A server can also become more efficient through multithreading or multiprocessing. In this case, a server can answer more than one request at a time.

Uniform Resource Locator

A client that wants to access a Web page needs the address. To facilitate the access of documents distributed throughout the world, HTTP uses locators. The uniform resource locator (URL) is a standard for specifying any kind of information on the Internet. URL defines four things: protocol, host computer, port, and path.

7.5.2 HTTP

The Hypertext Transfer Protocol (HTTP) is a protocol used mainly to access data on the World Wide Web. HTTP functions as a combination of FTP and SMTP. It is similar to FTP because it transfers files and uses the services of TCP. However, it is much simpler than FTP because it uses only one TCP connection. There is no separate control connection; only data are transferred between the client and the server.

HTTP is like SMTP because the data transferred between the client and the server look like SMTP messages. In addition, the format of the messages is controlled by MIME-like headers. Unlike SMTP, the HTTP messages are not destined to be read by humans; they are read and interpreted by the HTTP server and HTTP client (browser). SMTP messages are stored and forwarded, but HTTP messages are delivered immediately. The commands from the client to the server are embedded in a request message. The contents of the requested file or other information are embedded in a response message. HTTP uses the services of TCP on well-known port 80.

➤ **HTTP Transaction**

Figure 7.18 illustrates the HTTP transaction between the client and server. Although HTTP uses the services of TCP, HTTP itself is a stateless protocol. The client initializes the transaction by sending a request message. The server replies by sending a response.

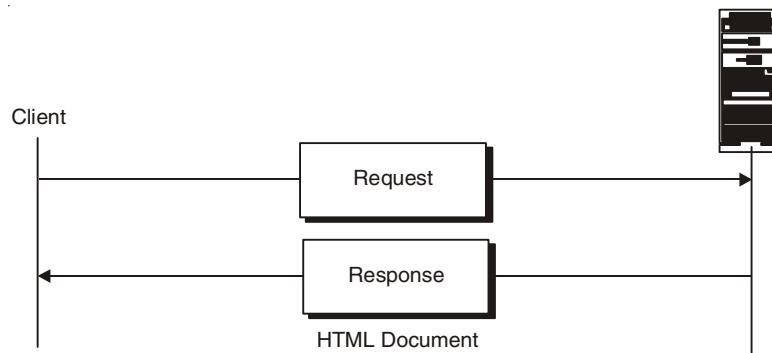


Fig. 7.18. HTTP Transaction.

➤ Messages

The formats of the request and response messages are similar; both are shown in Fig. 7.19. A request message consists of a request line, a header, and sometimes a body. A response message consists of a status line, a header, and sometimes a body.

- **Request and status lines:** The first line in a request message is called a request line; the first line in the response message is called the status line.
- **Request type:** This field is used in the request message. In version 1.1 of HTTP, several request types are defined.
- **URL:** We discussed the URL earlier in the chapter.
- **Version:** The most current version of HTTP is 1.1.
- **Status code:** This field is used in the response message. The status code field is similar to those in the FTP and the SMTP protocols. It consists of three digits. Whereas the codes in the 100 range are only informational, the codes in the 200 range indicate a successful request. The codes in the 300 range redirect the client to another URL, and the codes in the 400 range indicate an error at the client site. Finally, the codes in the 500 range indicate an error at the server site.
- **Status phrase:** This field is used in the response message. It explains the status code in text form.

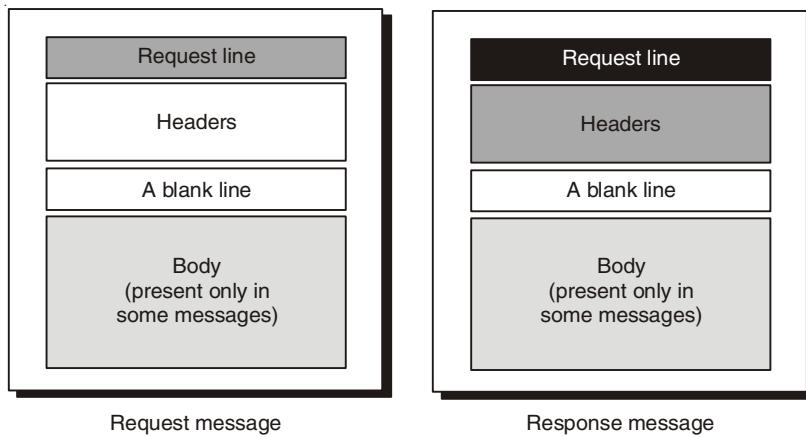


Fig. 7.19. Request and response messages.

7.6 NETWORK MANAGEMENT

A data communications network needs to operate efficiently, particularly in the event of a major failure. The cost of a network ceasing to function efficiently as a result of a failure is extremely high and most major networks incorporate some form of network management system to ensure efficient operation. The functions performed by a network management system must include the monitoring of the network's performance, handling failures when they occur and reconfiguring the network as a response to such failures.

7.6.1 Fault Management

An important feature of any network management system is the detection of faults and the subsequent repair and restoration of system performance. The first step in fault handling following the detection of a fault is normally setting an alarm indication of some kind. Traditionally, data communications equipment such as modems and multiplexers have used a crude form of alarm in which the loss of a circuit causes a lamp to extinguish. This primitive type of alarm conveys little information and can easily be overlooked. A true network management system normally provides a more positive type of alarm such as a message at a network operator's desk indicating the location and type of fault which requires some action by the operator. Some systems also set off alarms to indicate degrading situations, which may allow a serious situation to be averted.

Once an alarm has indicated a failure the next step is to restore the system to its normal operating condition. Initially, this may involve finding some short-term solution to the problem such as a fallback arrangement, but eventually the precise cause of the fault will need to be determined by using diagnostic techniques and the faulty device repaired or replaced.

➤ System restoration and reconfiguration

System restoration following a failure involves two processes. Normally, the first step is some form of **fallback switching** which involves the replacement of a failed device or circuit by an immediately available back-up. In the case of failed circuits the backup is normally an alternative circuit within the same network, although in some cases an alternative circuit may be provided by a separate back-up network. With equipment such as routers, switches and multiplexers, the back-up is usually provided by spare equipment. Fallback switching is normally automatic but may still be carried out manually, for example by using a **patch panel** to connect to alternative circuits temporarily. A patch panel is an arrangement that allows devices to be connected, or patched, together. Thus, a patch panel may be used to bypass faulty equipment and to patch in spare equipment in the event of failure. Similarly, a patch panel allows a limited amount of reconfiguration to be carried out, although the reconfiguration of a large network is normally beyond its capabilities.

➤ Test equipment

Fallback switching is only a short-term solution to network failure. The more permanent repair or replacement of equipment or circuits requires testing to locate a failure precisely. Many systems provide an integrated test facility although some may only provide access for separate test equipment. Several items of test equipment, listed below, have traditionally proved useful, particularly in smaller networks in which there is no integrated test facility.

Data analyser

This is a device which can capture the status of data on a data line and display this information either on a screen or as a printout. It displays data flow in both directions of a line as well as timing information. A typical printout is shown in Fig. 7.20. This printout has been produced using a software-based data analyser and shows a data signal above a corresponding clock signal.

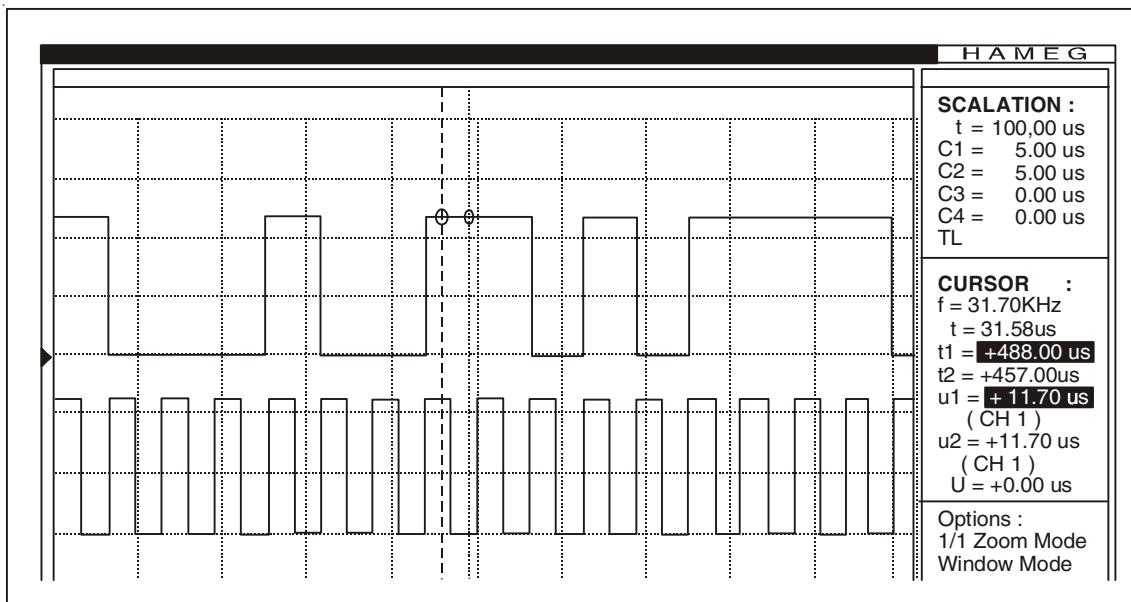


Fig. 7.20. Data Analyzer output.

Protocol analyser

A protocol analyser is an extension of the idea of a data analyser which, as well as displaying data, can also carry out simulations. It normally has two modes of operation, the more basic of which does little more than a data analyser. However, in simulation mode the protocol analyser can simulate the signals and responses expected from a particular piece of equipment or part of the network. In this way it can be used not only to determine the source of failures but also to analyse and hopefully improve network performance. Protocol analysers can be used with most common standard protocols such as Frame Relay, ATM and the IEEE 802 series of protocols.

➤ Trouble ticketing

Trouble tickets have traditionally been a major constituent of network management systems. They are used as a means of logging the problems that arise in a data network. A trouble ticket is an extended log entry containing useful information such as date, time, network device involved and the nature of the problem. It also contains information on any follow-up, such as the remedial action taken and details of any equipment or parts replaced. Manual trouble ticketing has mainly been replaced by trouble ticketing databases which can play a further role in the management and planning of a network. Thus, network managers or engineers may, for example, examine the fault history of a particular item of equipment or the action taken for particular types of faults to see whether the network is functioning efficiently.

7.6.2 Configuration Management

Configuration management involves the long-term planning and configuration of a network's topology and inventory of equipment and circuits. Most systems of configuration management

contain an inventory database with information on both active and back-up equipment and connections. All changes to a network's configuration are tracked and these databases updated. As well as aiding the reconfiguration of a network following network failures, configuration management systems allow network managers and engineers to make informed decisions on future network expansion.

7.6.3 Accounting Management

A network management system needs to keep track of the use of network resources. This is often simply for accounting reasons in networks where customers are charged for the use of the network resources. Other reasons for accounting management are:

- assisting with the planning of future network development and expansion;
- helping users to make more efficient use of the network;
- the detection of users who may be using the network inappropriately.

7.6.4 Performance Management

There are a number of measures of network performance in relation to link control and management. These included link efficiency and utilization, error rates and delay times. These measures apply equally to the network as a whole and the collection of these and other statistics forms an important part of any network management system. Statistics gathered by a network management system not only aid the efficient operation of a network but also can be used to plan the future requirements of a network. Most systems use the two further performance criteria of availability and response time, as outlined below.

➤ Availability

In most communications systems availability is an important measure of the amount of time that a system is available for use and can be accessed by the customer. Network availability is normally measured only from the originating service provider network switch node to the terminating switch node. In managed service contracts under which the customer uses the service provider's end equipment, network availability is normally measured on an end-router to end-router basis. In Internet dial-up access, availability is related to the percentage of calls that are not blocked when dialling into the Internet.

Typically availability is measured as the average amount of time a system is available per unit time. Alternatively it can be measured in terms of **mean time between failures (MTBF)** and the **mean time to repair (MTTR)** as follows:

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

➤ Response time

In many communications systems response time is an important measure of performance. It is a measure of the speed of operation of the system. It can be defined for systems in which humans carry out operations as the time between an operator pressing a transmit key and a reply appearing on a screen. This time is also known as **roundtrip delay**. A network management system will gather response time statistics for specific devices and circuits as

well as for complete networks. Once gathered, statistics are presented in a variety of ways, quite often in the form of a graphical printout. Fig. 7.21 shows a typical response time chart. This chart shows response times plotted at regular intervals, the value plotted being an average response time obtained during the time interval. Network management systems often use such statistics to produce more detailed breakdowns which allow percentiles to be obtained, as in Fig. 7.22 which has been obtained from Fig. 7.21.

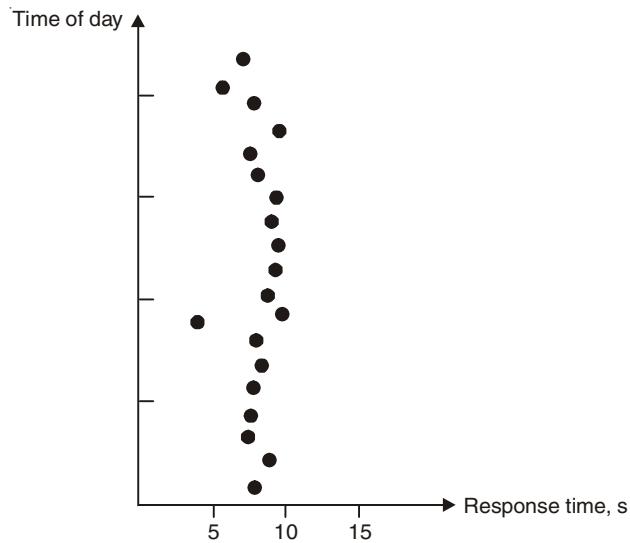


Fig. 7.21. Response Time chart.

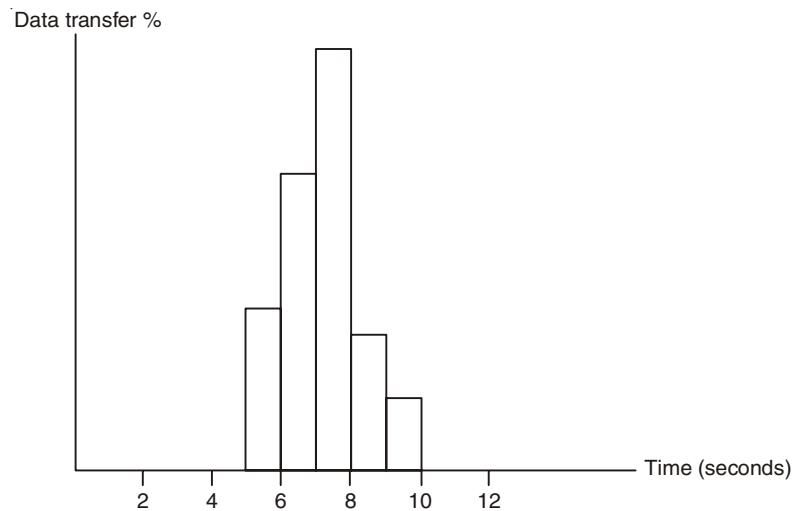


Fig. 7.22. Detailed Output charts.

7.6.5 Integrated Management Systems

Originally, the functions of performance, failure, configuration and accounting management were provided separately for a data communications network. More recently, network

management systems integrate these management functions with reporting capabilities in a centralized computer. The computer has access to a variety of local and remote monitoring devices, typically by means of SNMP, and may itself be connected to a computer network. The monitoring devices used depend on the type of system. For example, a system assembled by a company that supplies routers will use diagnostic monitoring facilities associated with its routers.

7.6.6 Network Management Standards

Standards for network management are not as well established in practical networks as in other areas of data communications. The American National Standards Institute (ANSI) Committee X3T5.4 was tasked by the ISO to develop a management standard for OSI systems. Within the ISO itself, Study Group 21 looked at network management within the OSI model in general. In its capacity as a LAN standardization body, the IEEE 802 Committee carries out work on standards for the management of LANs.

➤ OSI network management

As is the case with most areas of OSI standardization, the series of standards that have been developed by the ISO is both complicated and voluminous. The first of these standards, issued jointly by the ISO and ITU-T, was the X.700 recommendation entitled *OSI Basic Reference Model Part 4: Management Framework*, which specified the overall management framework for the OSI model and gave a general introduction to network management concepts. In this context the ISO prefers the term 'systems management' to what is more generally known as network management. A number of standards followed which constitute the ITU-T X.700 series of recommendations. Recommendation X.701 provides a general overview of the other standards.

The key elements of the series of documents are recommendations X.710 Common Management Information Services and X.720 Common Management Information Protocols. These form the basis for all the OSI network management functions. X.710 is the set of management services provided and X.720 provides a protocol for exchanging information between the points at which management functions are initiated and other points at the same level where the functions are carried out. Five functions are specified in the documentation as follows:

1. **Fault management:** The OSI fault management facility allows for the detection and identification of abnormal operation in an OSI environment. The specific facilities include the following:
 - (a) The detection of faults and the passing on of error reports.
 - (b) The carrying out of diagnostic tests on remote system resources. The ISO uses the term **managed object** to describe a resource or device that is managed by its management protocols.
 - (c) The maintenance of a log of events.
 - (d) The operation of threshold alarms which are activated if a preset threshold is crossed.

2. **Configuration management:** OSI configuration management allows a network manager to observe and modify the configuration of network components in an open system. The following facilities are included:
 - (a) The collection of data concerning the current configuration of network components.
 - (b) The alteration of network component configurations.
 - (c) The initialization and closing down of network components.
3. **Accounting management:** OSI accounting management allows network managers to identify the use of resources in an open system and, where costs are incurred in the use of resources, to calculate and allocate such costs. Two main areas of accounting are specified, namely the transmission system including the communication medium and the end-systems.
4. **Performance management:** A performance management capability is envisaged which will allow for the monitoring and collection of data concerning the current performance of network resources in an open system and also the generation of performance reports. As yet, there is no facility for the prediction of performance patterns. It is possible to make performance predictions using, among other tools. The lack of this facility is considered by some to be a disadvantage of this area of the OSI management standards.
5. **Security management:** The term 'security' is frequently associated with sensitive areas such as military systems in which data is often highly confidential or even secret. However, the term has a much wider meaning and a security management facility should ideally include the following functions:
 - (a) The control of access by users to the various resources of a system.
 - (b) The protection of both data and operations to ensure that they are carried out correctly.
 - (c) The authentication of data communications to ensure that data is from the source that it claims to be.

No systems are totally secure. Even a local, isolated system is prone to some insecurity and such problems are multiplied as the size of a network increases. An open system, therefore, is particularly vulnerable in this respect.

➤ M.3100 recommendation

Perhaps one of the more widely used applications of OSI network management can be found in the world of telecommunications management networks where much of the modelling provided is based on objects specified in the ITU-T M.3100 recommendation. M.3100 in turn bases much of its terminology on the networking nomenclature defined for the Synchronous Digital Hierarchy (SDH) by the G.803 recommendation. These recommendations provide a generic networks information model that can be used to manage **network elements** such as multiplexers, switches and routers. The managed objects could be physical components or logical components such as software. A managed object possesses **attributes** that allow a user to control and/or observe the behaviour of the object. Objects with similar attributes and behaviours may be grouped into **object classes**.

A good way to illustrate these terms is by way of a hierarchical diagram, known as an **inheritance/containment hierarchy**. An explanatory inheritance hierarchy is shown in Fig. 7.23 It should be stressed that this figure is intended to be explanatory rather than typical. The network element shown is a switch. There are two object classes shown called equipment and software.

Within each object class there are one or more managed objects which may be organized into further subclasses. It may be that there is more than one example of a particular object, in which case they are enumerated. The recommendation refers to each enumeration as an **instance**. Thus in Fig. 7.23 there are two instances (1 and 2) of an object called *LineCard*. Most objects will have a number of attributes.

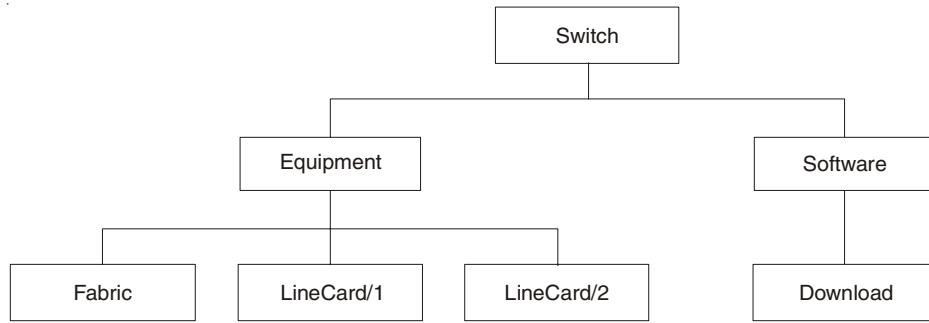


Fig. 7.23. Inheritance hierarchy of switch.

➤ Network management in the TCP/IP environment

TCP/IP evolved for many years without any formal network management capability. It was not until the late 1980s, when the Internet started to grow rapidly, that attention was given to this matter and a network management protocol called SNMP (Simple Network Management Protocol) was developed for use in a TCP/IP environment. The first SNMP products were produced in 1988 and their use has spread rapidly since to most major manufacturers. Network management within a TCP/IP environment is based on four key components:

1. **Network management station:** This is normally a standalone device which acts as the interface between a network manager and SNMP. A network management station provides typical network management functions such as data analysis, fault recovery and network monitoring facilities. It also has access to a database.
2. **Agent:** Network elements within a network such as routers and bridges can be managed from a network management station. The active software modules within a network element that communicate with a network management station are known as **agents**. Network management stations send various commands to agents which respond appropriately. The agents may also send important items of information, such as alarms, to a network management station even though they are not specifically requested.
3. **Management information base:** At each network element there is a collection of data similar to a database, which is known as a **management information base (MIB)**.

A network management station controls a network element by accessing its management information base and retrieving information from it or modifying its contents. Typical information contained in a MIB would be configuration data.

4. **Network management protocol:** This is the SNMP and it allows the management stations and the agents to communicate with each other.

Review Questions

1. How is HTTP related to WWW?
2. How is HTTP similar to SMTP?
3. How is HTTP similar to FTP?
4. What is a URL and what are its components?
5. ARP and DNS both depend on caches; ARP cache entry lifetimes are typically 10 minutes, while DNS cache is on the order of days. Justify this difference. What undesirable consequences might there be in having too long a DNS cache entry lifetime?
6. IPv6 simplifies ARP out of existence by allowing hardware addresses to be part of the IPv6 address. How does this complicate the job of DNS? How does this affect the problem of finding your local DNS server?
7. How are options negotiated in TELNET?
8. Describe the addressing system used by SMTP.
9. In electronic mail, what are the tasks of a user agent?
10. In electronic mail, what is MIME?
11. Why do we need POP3 or IMAP4 for electronic mail?



CHAPTER 8

COMPUTER NETWORK SECURITY

It is true greatness to have in one the frailty of a man and the security of a god.

—Seneca

8.1 INTRODUCTION

Before we talk about network security, we need to understand in general terms what security is. Security is a continuous process of protecting an object from unauthorized access. It is a state of being or feeling protected from harm. That object in that state may be a person, an organization such as a business, or property such as a computer system or a file. Security comes from secure which means, according to *Webster Dictionary*, a state of being free from care, anxiety, or fear.

An object can be in a *physical state* of security or a *theoretical state* of security. In a physical state, a facility is secure if it is protected by a barrier like a fence, has secure areas both inside and outside, and can resist penetration by intruders. This state of security can be guaranteed if the following four protection mechanisms are in place: deterrence, prevention, detection, and response. **Deterrence** is usually the first line of defense against intruders who may try to gain access. **Prevention** is the process of trying to stop intruders from gaining access to the resources of the system. **Detection** occurs when the intruder has succeeded or is in the process of gaining access to the system. Signals from the detection process include alerts to the existence of an intruder. **Response** is an aftereffect mechanism that tries to respond to the failure of the first three mechanisms. It works by trying to stop and/or prevent future damage or access to a facility.

A theoretical state of security, commonly known as pseudo security or security through ob-security (STO) is a false hope of security. Many believe that an object can be secure as long as nobody outside the core implementation group has knowledge about its existence. This security is often referred to as “bunk mentality” security. This is virtual security in the

sense that it is not physically implemented like building walls, issuing passwords, or putting up a firewall, but it is effectively based solely on a philosophy. The philosophy itself relies on a need to know basis, implying that a person is not dangerous as long as that person doesn't have knowledge that could affect the security of the system like a network, for example. In real systems where this security philosophy is used, security is assured through a presumption that only those with responsibility and who are trustworthy can use the system and nobody else needs to know. The belief that secrecy can make the system more secure is just that, a belief—a myth in fact. Unfortunately, the software industry still believes this myth.

Although its usefulness has declined as the computing environment has changed to large open systems, new networking programming and network protocols, and as the computing power available to the average person has increased, the philosophy is in fact still favored by many agencies, including the military, many government agencies, and private businesses.

In either security state, many objects can be thought of as being secure if such a state, a condition, or a process is afforded to them. Because there are many of these objects, we are going to focus on the security of a few of these object models. These will be a computer, a computer network, and information.

8.1.1 Computer Security

This is a study, which is a branch of Computer Science, focusing on creating a secure environment for the use of computers. It is a focus on the “behavior of users,” if you will, required and the protocols in order to create a secure environment for anyone using computers. This field, therefore, involves four areas of interest: the study of computer ethics, the development of both software and hardware protocols, and the development of best practices. It is a complex field of study involving detailed mathematical designs of cryptographic protocols. We are not focusing on this in this chapter.

8.1.2 Network Security

As we saw in chapter 1, computer networks are distributed networks of computers that are either strongly connected meaning that they share a lot of resources from one central computer or loosely connected, meaning that they share only those resources that can make the network work. When we talk about computer network security, our focus object model has now changed. It is no longer one computer but a network. So computer network security is a broader study of computer security. It is still a branch of computer science, but a lot broader than that of computer security. It involves creating an environment in which a computer network, including all its resources, which are many; all the data in it both in storage and in transit; and all its users are secure. Because it is wider than computer security, this is a more complex field of study than computer security involving more detailed mathematical designs of cryptographic, communication, transport, and exchange protocols and best practices. This chapter focuses on this field of study.

8.1.3 Information Security

Information security is even a bigger field of study including computer and computer network security. This study is found in a variety of disciplines, including computer science, business management, information studies, and engineering. It involves the creation of a state in which information and data are secure. In this model, information or data is either in motion through the communication channels or in storage in databases on server. This, therefore, involves the study of not only more detailed mathematical designs of cryptographic, communication, transport, and exchange protocols and best practices, but also the state of both data and information in motion. We are not discussing these in this chapter.

8.2 SECURING THE COMPUTER NETWORK

Creating security in the computer network model we are embarking on in this chapter means creating secure environments for a variety of resources. In this model, a resource is secure, based on the above definition, if that resource is protected from both internal and external unauthorized access. These resources, physical or not, are objects. Ensuring the security of an object means protecting the object from unauthorized access both from within the object and externally. In short, we protect objects. System objects are either tangible or non tangible. In a computer network model, the tangible objects are the hardware resources in the system, and the intangible object is the information and data in the system, both in transition and static in storage.

8.2.1 Hardware

Protecting hardware resources include protecting.

- End user objects that include the user interface hardware components such as all client system input components, including a keyboard, mouse, touch screen, light pens, and others.
- Network objects like firewalls, hubs, switches, routers and gateways which are vulnerable to hackers.
- Network communication channels to prevent eavesdroppers from intercepting network communications.

8.2.2 Software

Protecting software resources includes protecting hardware-based software, operating systems, server protocols, browsers, application software, and intellectual property stored on network storage disks and databases. It also involves protecting client software such as investment portfolios, financial data, real estate records, images or pictures, and other personal files commonly stored on home and business computers.

8.3 FORMS OF PROTECTION

Now, we know what model objects are or need to be protected. Let us briefly, keep details for later, survey ways and forms of protecting these objects. Prevention of unauthorized access to system resources is achieved through a number of services that include access control, authentication, confidentiality, integrity, and non-repudiation.

8.3.1 Access Control

This is a service the system uses, together with a user pre-provided identification information such as a password, to determine who uses what of its services. Let us look at some forms of access control based on hardware and software.

➤ **Hardware Access Control Systems**

Rapid advances in technology have resulted in efficient access control tools that are open and flexible, while at the same time ensuring reasonable precautions against risks. Access control tools falling in this category include the following:

- *Access terminal.* Terminal access points have become very sophisticated, and now they not only carry out user identification but also verify access rights, control access points, and communicate with host computers. These activities can be done in a variety of ways including fingerprint verification and real-time anti-break in sensors. Network technology has made it possible for these units to be connected to a monitoring network or remain in a stand-alone off-line mode.
- *Visual event monitoring.* This is a combination of many technologies into one very useful and rapidly growing form of access control using a variety of realtime technologies including video and audio signals, aerial photographs, and global positioning system (GPS) technology to identify locations.
- *Identification cards.* Sometimes called proximity cards, these cards have become very common these days as a means of access control in buildings, financial institutions, and other restricted areas. The cards come in a variety of forms, including magnetic, bar coded, contact chip, and a combination of these.
- *Biometric identification.* This is perhaps the fastest growing form of control access tool today. Some of the most popular forms include fingerprint, iris, and voice recognition. However, fingerprint recognition offers a higher level of security.

➤ **Software Access Control Systems**

Software access control falls into two types: point of access monitoring and remote monitoring. In *point of access* (POA), personal activities can be monitored by a PC-based application. The application can even be connected to a network or to a designated machine or machines. The application collects and stores access events and other events connected to the system operation and download access rights to access terminals.

In remote mode, the terminals can be linked in a variety of ways, including the use of modems, telephone lines, and all forms of wireless connections. Such terminals may, sometimes if needed, have an automatic calling at pre-set times if desired or have an attendant to report regularly.

8.3.2 Authentication

Authentication is a service used to identify a user. User identity, especially of remote users, is difficult because many users, especially those intending to cause harm, may masquerade as the legitimate users when they actually are not. This service provides a system with the capability to verify that a user is the very one he or she claims to be based on what the user is, knows, and has.

Physically, we can authenticate users or user surrogates based on checking one or more of the following user items:

- User name (sometimes screen name)
- Password
- *Retinal images*: The user looks into an electronic device that maps his or her eye retina image; the system then compares this map with a similar map stored on the system.
- *Fingerprints*: The user presses on or sometimes inserts a particular finger into a device that makes a copy of the user fingerprint and then compares it with a similar image on the system user file.
- *Physical location*: The physical location of the system initiating an entry request is checked to ensure that a request is actually originating from a known and authorized location. In networks, to check the authenticity of a client's location a network or Internet protocol (IP) address of the client machine is compared with the one on the system user file. This method is used mostly in addition to other security measures because it alone cannot guarantee security. If used alone, it provides access to the requested system to anybody who has access to the client machine.

8.3.3 Confidentiality

The confidentiality service protects system data and information from unauthorized disclosure. When data leave one extreme of a system such as a client's computer in a network, it ventures out into a non trusting environment. So, the recipient of that data may not fully trust that no third party like a cryptanalysis or a man-in-the middle has eavesdropped on the data. This service uses encryption algorithms to ensure that nothing of the sort happened while the data was in the wild.

Encryption protects the communications channel from sniffers. *Sniffers* are programs written for and installed on the communication channels to eavesdrop on network traffic, examining all traffic on selected network segments. Sniffers are easy to write and install and difficult to detect. The encryption process uses an encryption algorithm and key to transform data at the source, called *plaintext*; turn it into an encrypted form called *ciphertext*, usually unintelligible form; and finally recover it at the sink. The encryption algorithm can either be *symmetric* or *asymmetric*. Symmetric encryption or secret key encryption, as it is usually called, uses a common key and the same cryptographic algorithm to scramble and unscramble the message. Asymmetric encryption commonly known as public key encryption uses two different keys: a public key known by all and a private key known by only the sender and the receiver. Both the sender and the receiver each has a pair of these keys, one public and

one private. To encrypt a message, a sender uses the receiver's public key which was published. Upon receipt, the recipient of the message decrypts it with his or her private key.

8.3.4 Integrity

The integrity service protects data against active threats such as those that may alter it. Just like data confidentiality, data in transition between the sending and receiving parties is susceptible to many threats from hackers, eavesdroppers, and cryptanalysts whose goal is to intercept the data and alter it based on their motives. This service, through encryption and *hashing algorithms*, ensures that the integrity of the transient data is intact. A hash function takes an input message M and creates a code from it. The code is commonly referred to as a hash or a message digest.

A one-way hash function is used to create a signature of the message—just like a human fingerprint. The hash function is, therefore, used to provide the message's integrity and authenticity. The signature is then attached to the message before it is sent by the sender to the recipient.

8.3.5 Nonrepudiation

This is a security service that provides proof of origin and delivery of service and/or information. In real life, it is possible that the sender may deny the ownership of the exchanged digital data that originated from him or her. This service, through *digital signature* and encryption algorithms, ensures that digital data may not be repudiated by providing proof of origin that is difficult to deny. A digital signature is a cryptographic mechanism that is the electronic equivalent of a written signature to authenticate a piece of data as to the identity of the sender.

We have to be careful here because the term "nonrepudiation" has two meanings, one in the legal world and the other in the cryptotechnical world. Adrian McCullagh and Willian Caelli define "nonrepudiation" in a cryptotechnical way as follows:

- In authentication, a service that provides proof of the integrity and origin of data, both in a forgery-proof relationship, which can be verified by any third party at any time; or
- In authentication, an authentication that with high assurance can be asserted to be genuine, and that cannot subsequently be refuted.

However, in the legal world, there is always a basis for repudiation. This basis, again according to Adrian McCullagh, can be as follows:

- The signature is a forgery.
- The signature is not a forgery, but was obtained via
- Unconscionable conduct by a party to a transaction;
- Fraud instigated by a third party;
- Undue influence exerted by a third party.

8.4 SECURITY STANDARDS

The computer network model also suffers from the standardization problem. Security protocols, solutions, and best practices that can secure the computer network model come in many different types and use different technologies resulting in incompatibility of interfaces, less interoperability, and uniformity among the many system resources with differing technologies within the system and between systems. System managers, security chiefs, and experts, therefore, choose or prefer standards, if no de facto standard exists, that are based on service, industry, size, or mission. The type of service offered by an organization determines the types of security standards used. Like service, the nature of the industry an organization is in also determines the types of services offered by the system, which in turn determines the type of standards to adopt. The size of an organization also determines what type of standards to adopt. In relatively small establishments, the ease of implementation and running of the system influence the standards to be adopted. Finally, the mission of the establishment also determines the types of standards used. For example, government agencies have a mission that differs from that of a university. These two organizations, therefore, may choose different standards.

We are, therefore, going to discuss security standards along these divisions. Before we do that, however, let us look at the bodies and organizations behind the formulation, development, and maintenance of these standards. These bodies fall into the following categories:

- International organizations such as the Internet Engineering Task Force (IETF), the Institute of Electronic and Electric Engineers (IEEE), the International Standards Organization (ISO), and the International Telecommunications Union (ITU).
- Multinational organizations like the European Committee for Standardization (CEN), Commission of European Union (CEU), and European Telecommunications Standards Institute (ETSI).
- National governmental organizations like the National Institute of Standards and Technology (NIST), American National Standards Institute (ANSI), and Canadian Standards Council (CSC).
- Sector specific organizations such as the European Committee for Banking Standards (ECBS), European Computer Manufacturers Association (ECMA), and Institute of Electronic and Electric Engineers (IEEE).
- Industry standards such as RSA, the Open Group (OSF + X/Open), Object Management Group (OMG), World Wide Web Consortium (W3C)), and the Organization for the Advancement of Structured Information Standards (OASIS).
- Other sources of standards in security and cryptography. Each one of these organizations has a set of standards.

8.4.1 Security Standards Based on Type of Service/Industry

System and security managers and users may choose a security standard to use based on the type of industry they are in and what type of services that industry provides.

➤ **Public-Key Cryptography Standards (PKCS)**

In order to provide a basis and a catalyst for interoperable security based on public-key cryptographic techniques, the Public-Key Cryptography Standards (PKCS) were established. These are recent security standards, first published in 1991 following discussions of a small group of early adopters of public-key technology. Since their establishment, they have become the basis for many formal standards and are implemented widely.

In general, PKCS are security specifications produced by RSA Laboratories in cooperation with secure systems developers worldwide for the purpose of accelerating the deployment of public-key cryptography. In fact, worldwide contributions from the PKCS series have become part of many formal and de facto standards, including ANSI X9 documents, PKIX, SET, S/MIME, and SSL.

➤ **The Standards For Interoperable Secure MIME (S/MIME)**

S/MIME (*Secure Multipurpose Internet Mail Extensions*) is a specification for secure electronic messaging. It came to address a growing problem of e-mail interception and forgery at the time of increasing digital communication. So, in 1995, several software vendors got together and created the S/MIME specification with the goal of making it easy to secure messages from prying eyes.

It works by building a security layer on top of the industry standard MIME protocol based on PKCS. The use of PKCS avails the user of S/MIME with immediate privacy, data integrity, and authentication of an e-mail package. This has given the standard a wide appeal, leading to S/MIME moving beyond just e-mail. Already vendor software warehouses, including Microsoft, Lotus, Banyan, and other on-line electronic commerce services are using S/MIME.

➤ **Federal Information Processing Standards (FIPS)**

Federal Information Processing Standards (FIPS) are National Institute of Standards and Technology (NIST) approved standards for advanced encryption. These are U.S. federal government standards and guidelines in a variety of areas in data processing. They are recommended by NIST to be used by U.S. government organizations and others in the private sector to protect sensitive information. They range from FIPS 31 issued in 1974 to current FIPS 198.

➤ **Secure Sockets Layer (SSL)**

SSL is an encryption standard for most Web transactions. In fact, it is becoming the most popular type of e-commerce encryption. Most conventional intranet and extranet applications would typically require a combination of security mechanisms that include

- Encryption
- Authentication
- Access control.

SSL provides the encryption component implemented within the TCP/IP protocol. Developed by Netscape Communications, SSL provides secure web client and server communications, including encryption, authentication, and integrity checking for a TCP/IP connection.

8.4.2 Security Standards Based on Size/Implementation

If the network is small or it is a small organization such as a university, for example, security standards can be spelled out as best practices on the security of the system, including the physical security of equipment, system software, and application software.

- **Physical security.** This emphasizes the need for security of computers running the Web servers and how these machines should be kept physically secured in a locked area. Standards are also needed for backup storage media like tapes and removable disks.
- **Operating systems.** The emphasis here is on privileges and number of accounts, and security standards are set based on these. For example, the number of users with most privileged access like *root* in UNIX or *Administrator* in NT should be kept to a minimum. Set standards for privileged users. Keep to a minimum the number of user accounts on the system. State the number of services offered to clients computers by the server, keeping them to a minimum. Set a standard for authentication such as user passwords and for applying security patches.
- **System logs.** Logs always contain sensitive information such as dates and times of user access. Logs containing sensitive information should be accessible only to authorized staff and should not be publicly accessible. Set a standard on who and when logs should be viewed and analyzed.
- **Data security.** Set a standard for dealing with files that contain sensitive data. For example, files containing sensitive data should be encrypted wherever possible using strong encryption or should be transferred as soon as possible and practical to a secured system not providing public services.

8.4.3 Security Standards Based on Interests

In many cases, institutions and government agencies choose to pick a security standard based solely on the interest of the institution or the country.

➤ British Standard 799 (BS 7799)

The BS 7799 standard outlines a code of practice for information security management that further helps to determine how to secure network systems. It puts forward a common framework that enables companies to develop, implement, and measure effective security management practice and provide confidence in inter-company trading. BS 7799 was first written in 1993, but it was not officially published until 1995, and it was published as an international standard BS ISO/IEC 17799:2000 in December 2000.

➤ Orange Book

This is the U.S. Department of Defense *Trusted Computer System Evaluation Criteria* (DOD-5200.28-STD) standard known as the *Orange Book*. For a long time, it has been the *de facto* standard for computer security used in government and industry, but other standards have now been developed to either supplement it or replace it. First published in 1983, its security levels are referred to as "Rainbow Series."

➤ Homeland National Security Awareness

After the September 11, 2001, attack on the United States, the government created a new cabinet department of Homeland Security to be in charge of all national security issues. The Homeland Security department created a security advisory system made up of five levels ranging from green (for low security) to red (severe) for heightened security.

8.5 SOURCES OF SECURITY THREATS

The security threat to computer systems springs from a number of factors that include weaknesses in the network infrastructure and communication protocols that create an appetite and a challenge to the hacker mind, the rapid growth of cyberspace into a vital global communication and business network on which international commerce and business transactions are increasingly being performed and many national critical infrastructures are being connected, the growth of the hacker community whose members are usually experts at gaining unauthorized access into systems that run not only companies and governments but also critical national infrastructures, the vulnerability in operating system protocols whose services run the computers that run the communication network, the insider effect resulting from workers who steal and sell company databases and the mailing lists or even confidential business documents, social engineering, physical theft from within the organizations of things such as laptop and hand-held computers with powerful communication technology and more potentially sensitive information, and security as a moving target.

8.5.1 Design Philosophy

Although the design philosophy on which both the computer network infrastructure and communication protocols built has tremendously boosted were cyberspace development, the same design philosophy has been a constant source of the many ills plaguing cyberspace. The growth of the Internet and cyberspace in general was based on an *open architecture work in progress* philosophy. This philosophy attracted the brightest minds to get their hands dirty and contribute to the infrastructure and protocols. With many contributing their best ideas for free, the Internet grew in leaps and bounds. This philosophy also helped the spirit of individualism and adventurism, both of which have driven the growth of the computer industry and underscored the rapid and sometimes motivated growth of cyberspace. Because the philosophy was not based on clear blueprints, new developments and additions came about as reactions to the shortfalls and changing needs of a developing infrastructure. The lack of a comprehensive blueprint and the demand-driven design and development of protocols are causing the ever present weak points and loopholes in the underlying computer network infrastructure and protocols.

In addition to the philosophy, the developers of the network infrastructure and protocols also followed a policy to create an interface that is as user-friendly, efficient, and transparent as possible so that all users of all education levels can use it unaware of the working of the networks and therefore are not concerned with the details. The designers of the communication network infrastructure thought it was better this way if the system is to

serve as many people as possible. Making the interface this easy and far removed from the details, though, has its own downside in that the user never cares about and pays very little attention to the security of the system. Like a magnet, the policy has attracted all sorts of people who exploit the network's vulnerable and weak points in search of a challenge, adventurism, fun, and all forms of personal gratification.

8.5.2 Weaknesses in Network Infrastructure and Communication Protocols

Compounding the problems created by the design philosophy and policy are the weaknesses in the communication protocols. The Internet is a packet network that works by breaking the data to be transmitted into small individually addressed packets that are downloaded on the network's mesh of switching elements. Each individual packet finds its way through the network with no predetermined route and the packets are reassembled to form the original message by the receiving element.

To work successfully, packet networks need a strong trust relationship that must exist among the transmitting elements. As packets are disassembled, transmitted, and re-assembled, the security of each individual packet and the intermediary transmitting elements must be guaranteed. This is not always the case in the current protocols of cyberspace. There are areas where, through port scans, determined users have managed to intrude, penetrate, fool, and intercept the packets.

The two main communication protocols on each server in the network, UDP and TCP, use port numbers to identify higher layer services. Each higher layer service on a client uses a unique port number to request a service from the server and each server uses a port number to identify the service needed by a client. The cardinal rule of a secure communication protocol in a server is never to leave any port open in the absence of a useful service. If no such service is offered, its port should never be open. Even if the service is offered by the server, its port should never be left open unless it is legitimately in use.

In the initial communication between a client and a server, the client addresses the server via a port number in a process called a *three-way handshake*. The three way handshake, when successful, establishes a TCP virtual connection between the server and the client. This virtual connection is required before any communication between the two can begin. The process begins by a client/host sending a TCP segment with the synchronize (SYN) flag set; the server/host responds with a segment that has the acknowledge valid (ACK) and SYN flags set, and the first host responds with a segment that has only the ACK flag set. This exchange is shown in Fig. 8.1.

The three-way handshake suffers from a *half-open* socket problem when the server trusts the client that originated the handshake and leaves its port door open for further communication from the client. As long as the half-open port remains open, an intruder can enter the system because while one port remains open, the server can still entertain other three-way handshakes from other clients that want to communicate with it. Several half-open ports can lead to network security exploits including both TCP/IP and UDP Protocols: Internet Protocol spoofing (IP spoofing), in which IP addresses of the source element in the data packets are altered and replaced with bogus addresses, and SYN flooding where the server is overwhelmed by spoofed packets sent to it.

In addition to the three-way handshake, ports are used widely in network communication. There are well-known ports used by processes that offer services. For example, ports 0 through 1023 are used widely by system processes and other highly privileged programs. This means that if access to these ports is compromised, the intruder can get access to the whole system. Intruders find open ports via port scans. The two examples below from G-Lock Software illustrate how a port scan can be made.

- *TCP connect() scanning* is the most basic form of TCP scanning. An attacker's host is directed to issue a `connect()` system call to a list of selected ports on the target machine. If any of these ports is listening, `connect()` system call will succeed; otherwise, the port is unreachable and the service is unavailable.
- *UDP Internet Control Message Protocol (ICMP) port unreachable scanning* is one of the few UDP scans. Recall from Chapter 6 that UDP is a connectionless protocol; so, it is harder to scan than TCP because UDP ports are not required to respond to probes. Most implementations generate an ICMP port_unreachable error when an intruder sends a packet to a closed UDP port. When this response does not come, the intruder has found an active port.

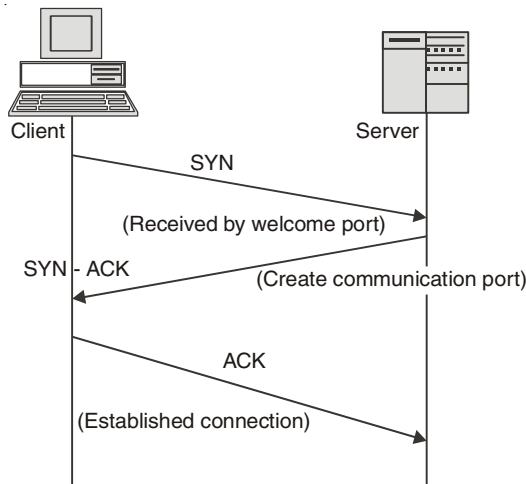


Fig. 8.1. A three way Hand Shake.

Further more , there are DDoS attacks , that are generally classified as nuisance attacks in the sense that they simply interrupt the services of the system. System interruption can be as serious as destroying a computer's hard disk or as simple as using up all the available memory of the system. DDoS attacks come in many forms, but the most common are the following: smurfing, ICMP protocol, and ping of death attacks.

8.5.3 Rapid Growth of Cyberspace

There is always a security problem in numbers. Since its beginning as ARPANET in the early 1960s, the Internet has experienced phenomenal growth, especially in the last 10 years. There was an explosion in the numbers of users, which in turn ignited an explosion in the number of connected computers.

Just less than 20 years ago in 1985, the Internet had fewer than 2000 computers connected and the corresponding number of users was in the mere tens of thousands. However, by 2001, the figure has jumped to about 109 million hosts, according to Tony Rutkowski at the Center for Next Generation Internet, an Internet Software Consortium. This number represents a significant new benchmark for the number of Internet hosts. At a reported current annual growth rate of 51% over the past 2 years, this shows continued strong exponential growth, with an estimated growth of up to 1 billion hosts if the same growth rate is sustained.

This is a tremendous growth by all accounts. As it grew, it brought in more and more users with varying ethical standards, added more services, and created more responsibilities. By the turn of the century, many countries found their national critical infrastructures firmly intertwined in the global network. An interdependence between humans and computers and between nations on the global network has been created that has led to a critical need to protect the massive amount of information stored on these network computers. The ease of use of and access to the Internet, and large quantities of personal, business, and military data stored on the Internet was slowly turning into a massive security threat not only to individuals and business interests but also to national defenses.

As more and more people enjoyed the potential of the Internet, more and more people with dubious motives were also drawn to the Internet because of its enormous wealth of everything they were looking for. Such individuals have posed a potential risk to the information content of the Internet, and such a security threat has to be dealt with.

Statistics from the security company Symantec show that Internet attack activity is currently growing by about 64% per year. The same statistics show that during the first 6 months of 2002, companies connected to the Internet were attacked, on average, 32 times per week compared to only 25 times per week in the last 6 months of 2001. Symantec reports between 400 and 500 new viruses every month and about 250 vulnerabilities in computer programs.

In fact, the rate at which the Internet is growing is becoming the greatest security threat ever. Security experts are locked in a deadly race with these malicious hackers that at the moment looks like a losing battle with the security community.

8.5.4 The Growth of the Hacker Community

Although other factors contributed significantly to the security threat, in the general public view, the number one contributor to the security threat of computer and telecommunication networks more than anything else is the growth of the hacker community. Hackers have managed to bring this threat into news headlines and people's living rooms through the ever increasing and sometimes devastating attacks on computer and telecommunication systems using viruses, worms, and DDoS.

The general public, computer users, policy makers, parents, and law makers have watched in bewilderment and awe as the threat to their individual and national security has grown to alarming levels as the size of the global networks have grown and national critical infrastructures have become more and more integrated into this global network. In some cases, the fear from these attacks reached hysterical proportions, as demonstrated in the following major attacks between 1986 and 2003 that we have rightly called the big "bungs."

8.5.5 Vulnerability in Operating System Protocol

One area that offers the greatest security threat to global computer systems is the area of software errors, especially network operating systems errors. An operating system plays a vital role not only in the smooth running of the computer system in controlling and providing vital services, but by playing a crucial role in the security of the system in providing access to vital system resources. A vulnerable operating system can allow an attacker to take over a computer system and do anything that any authorized super user can do, such as changing files, installing and running software, or reformatting the hard drive.

8.5.6 The Invisible Security Threat—The Insider Effect

Quite often, news media reports show that in cases of violent crimes such as murder, one is more likely to be attacked by someone one does not know. However, real official police and court records show otherwise. This is also the case in network security. Research data from many reputable agencies consistently show that the greatest threat to security in any enterprise is the guy down the hall.

In 1997, the accounting firm Ernst & Young interviewed 4,226 IT managers and professionals from around the world about the security of their networks. From the responses, 75 percent of the managers indicated that they believed authorized users and employees represent a threat to the security of their systems. Forty-two percent of the Ernst and Young respondents reported they had experienced external malicious attacks in the past year, while 43 percent reported malicious acts from employees .

The *Information Security Breaches Survey 2002*, a U.K. government's Department of Trade and Industry sponsored survey conducted by the consultancy firm PricewaterhouseCoopers, found that in small companies, 32 percent of the worst incidents were caused by insiders, and this number jumps to 48 percent in large companies.

Although slightly smaller, similar numbers were found in the *CBI Cybercrime Survey 2001*. In that survey, 25 percent of organizations identified employees or former employees as the main cybercrime perpetrators, compared to 75 percent who cited hackers, organized crime, and other outsiders.

Other studies have shown slightly varying percentages of insiders doing the damage to corporate security. As the data indicates, many company executives and security managers had for a long time neglected to deal with the guys down the hall selling corporate secrets to competitors.

8.6 SECURITY THREAT MANAGEMENT

Security threat management is a technique used to monitor an organization's critical security systems in real-time to review reports from the monitoring sensors such as the intrusion detection systems, firewall, and other scanning sensors. These reviews help to reduce false positives from the sensors, develop quick response techniques for threat containment and assessment, correlate and escalate false positives across multiple sensors or platforms, and develop intuitive analytical, forensic, and management reports.

As the workplace gets more electronic and critical company information finds its way out of the manila envelopes and brown folders into online electronic databases, security management has become a full-time job for system administrators. While the number of dubious users is on the rise, the number of reported criminal incidents is skyrocketing, and the reported response time between a threat and a real attack is down to 20 minutes or less. To secure company resources, security managers have to do real-time management. Realtime management requires access to real time data from all network sensors. Among the techniques used for security threat management are risk assessment and forensic analysis.

8.6.1 Risk Assessment

Even if there are several security threats all targeting the same resource, each threat will cause a different risk and each will need a different risk assessment. Some will have low risk while others will have the opposite. It is important for the response team to study the risks as sensor data come in and decide which threat to deal with first.

8.6.2 Forensic Analysis

Forensic analysis is done after a threat has been identified and contained. After containment, the response team can launch the forensic analysis tools to interact with the dynamic report displays that have come from the sensors during the duration of the threat or attack if the threat results in an attack. The data on which forensic analysis should be performed must be kept in a secure state to preserve the evidence. It must be stored and transferred, if this is needed, with the greatest care, and the analysis must be done with the utmost professionalism possible if the results of the forensic analysis are to stand in court.

8.7 CYBER CRIMES AND HACKERS

The greatest threats to the security, privacy, and reliability of computer networks and other related information systems in general are cyber crimes committed by cyber criminals, but most importantly hackers. Judging by the damage caused by past cyber criminal and hacker attacks to computer networks in businesses, governments, and individuals, resulting in inconvenience and loss of productivity and credibility, one cannot fail to see that there is a growing community demand to software and hardware companies to create more secure products that can be used to identify threats and vulnerabilities, to fix problems, and to deliver security solutions.

Industry and governments around the globe are responding to these threats through a variety of approaches and collaborations such as:

- Formation of organizations, such as the *Information Sharing and Analysis Centers* (ISACs).
- Getting together of industry portals and ISPs on how to deal with distributed denial of service attacks including the establishment of *Computer Emergency Response Teams* (CERTs).

- Increasing the use of sophisticated tools and services by companies to deal with network vulnerabilities. Such tools include the formation of Private Sector Security Organizations (PSSOs) such as SecurityFocus, Bugtraq, and the International Chamber of Commerce's Cybercrime Unit.

According to the director of the U.S. National Infrastructure Protection Center (NIPC), cyber crimes present the greatest danger to e-commerce and the general public in general. The threat of crime using the Internet is real and growing and it is likely to be the scourge of the 21st century. A *cyber crime* is a crime like any other crime, except that in this case, the illegal act must involve a connected computing system either as an object of a crime, an instrument used to commit a crime or a repository of evidence related to a crime. Alternatively, one can define a cyber crime as an act of unauthorized intervention into the working of the telecommunication networks or/and the sanctioning of an authorized access to the resources of the computing elements in a network that leads to a threat to the system's infrastructure or life or that causes significant property loss.

Because of the variations in jurisdiction boundaries, cyber acts are defined as illegal in different ways depending on the communities in those boundaries. Communities define acts to be illegal if such acts fall within the domains of that community's commission of crimes that a legislature of a state or a nation has specified and approved. Both the International Convention of Cyber Crimes and the European Convention on Cyber Crimes have outlined the list of these crimes to include the following:

- Unlawful access to information
- Illegal interception of information
- Unlawful use of telecommunication equipment.
- Forgery with use of computer measures
- Intrusions of the Public Switched and Packet Network
- Network integrity violations
- Privacy violations
- Industrial espionage
- Pirated computer software
- Fraud using a computing system
- Internet/e-mail abuse
- Using computers or computer technology to commit murder, terrorism, pornography, and hacking.

8.7.1 Cyber Criminals

Who are the cyber criminals? They are ordinary users of cyberspace with a message. As the number of users swells, the number of criminals among them also increases at almost the same rate. A number of studies have identified the following groups as the most likely sources of cyber crimes:

- *Insiders:* For a long time, system attacks were limited to in-house employee generated attacks to systems and theft of company property. In fact, disgruntled insiders are

a major source of computer crimes because they do not need a great deal of knowledge about the victim computer system. In many cases, such insiders use the system everyday. This allows them to gain unrestricted access to the computer system, thus causing damage to the system and/or data.

- *Hackers*: Hackers are actually computer enthusiasts who know a lot about computers and computer networks and use this knowledge with a criminal intent. Since the mid-1980s, computer network hacking has been on the rise mostly because of the widespread use of the Internet.
- *Criminal groups*: A number of cyber crimes are carried out by criminal groups for different motives ranging from settling scores to pure thievery. For example, such criminal groups with hacking abilities have broken into credit card companies to steal thousands of credit card numbers.
- *Disgruntled ex-employees*: Many studies have shown that disgruntled ex-employees also pose a serious threat to organizations as sources of cyber crimes targeting their former employers for a number of employee employer issues that led to the separation. In some cases, ex-employees simply use their knowledge of the system to attack the organization for purely financial gains.
- *Economic espionage spies*: The growth of cyberspace and e-commerce and the forces of globalization have created a new source of crime syndicates, the organized economic spies that plough the Internet looking for company secrets. As the price tag for original research skyrockets, and competition in the market place becomes global, companies around the globe are ready to pay any amount for stolen commercial, marketing, and industrial secrets.

8.7.2 Hackers

The word *hacker* has changed meaning over the years as technology changed. Currently, the word has two opposite meanings. One definition talks of a computer enthusiast as an individual who enjoys exploring the details of computers and how to stretch their capabilities, as opposed to most users who prefer to learn only the minimum necessary. The opposite definition talks of a malicious or inquisitive meddler who tries to discover information by poking around.

Before acquiring its current derogatory meaning, the term *hacking* used to mean expert writing and modification of computer programs. Hackers were considered people who were highly knowledgeable about computing; they were considered computer experts who could make the computer do all the wonders through programming. Today, however, hacking refers to a process of gaining unauthorized access into a computer system for a variety of purposes, including the stealing and altering of data and electronic demonstrations. For sometime now, hacking as a political or social demonstration has been used during international crises. During a crisis period, hacking attacks and other Internet security breaches usually spike in part because of sentiments over the crisis. For example, during the two Iraq wars, there were elevated levels of hacker activities.

According to the Atlanta-based Internet Security Systems, around the start of the first Iraq war, there was a sharp increase of about 37 percent from the fourth quarter of the year before, the largest quarterly spike the company has ever recorded.

There are several sub-sects of hackers based on hacking philosophies. The biggest sub-sects are crackers, hacktivists, and cyber terrorists.

➤ **Crackers**

A cracker is one who breaks security on a system. Crackers are hardcore hackers characterized more as professional security breakers and thieves. The term was recently coined only in the mid-1980s by purist hackers who wanted to differentiate themselves from individuals with criminal motives whose sole purpose is to sneak through security systems. Purist hackers were concerned journalists were misusing the term "hacker." They were worried that the mass media failed to understand the distinction between computer enthusiasts and computer criminals, calling both hackers. The distinction has, however, failed; so, the two terms *hack* and *crack* are still being often used interchangeably.

➤ **Hacktivists**

Hacktivism is a marriage between pure hacking and activism. Hacktivists are conscious hackers with a cause. They grew out of the old phreakers. Hacktivists carry out their activism in an electronic form in hope of highlighting what they consider noble causes such as institutional unethical or criminal actions and political and other causes. Hacktivism also includes acts of civil disobedience using cyberspace. The tactics used in hacktivism change with the time and the technology. Just as in the real world where activists use different approaches to get the message across, in cyberspace, hacktivists also use several approaches including automated e-mail bombs, web defacing, virtual sit-ins, and computer viruses and worms.

➤ **Cyberterrorists**

Based on motives, cyberterrorists can be divided into two categories: the terrorists and information warfare planners.

Terrorists: The World Trade Center attack in 2001 brought home the realization and the potential for a terrorist attack on not only organizations' digital infrastructure but also a potential for an attack on the national critical infrastructure. Cyber terrorists who are terrorists have many motives, ranging from political, economic, religious, to personal. Most often, the techniques of their terror are through intimidation, coercion, or actual destruction of the target.

Information Warfare Planners: This involves war planners to threaten attacking a target by disrupting the target's essential services by electronically controlling and manipulating information across computer networks or destroying the information infrastructure.

8.8 CRYPTOGRAPHY

Long ago, humans discovered the essence of secrecy. The art of keeping secrets resulted in victories in wars and in growth of mighty empires. Powerful rulers learned to keep secrets and pass information without interception; that was the beginning of cryptography. Although the basic concepts of cryptography predate the Greeks, the present word *cryptography*, used to describe the art of secret communication, comes from the Greek meaning "secret writing." From its rather simple beginnings, cryptography has grown in tandem with technology and its importance has also similarly grown. Just as in its early days, good cryptographic prowess still wins wars.

Cryptography is being increasingly used to fight off this massive invasion of individual privacy and security, to guarantee data integrity and confidentiality, and to bring trust in global e-commerce. Cryptography has become the main tool for providing the needed digital security in the modern digital communication medium that far exceeds the kind of security that was offered by any medium before it. It guarantees authorization, authentication, integrity, confidentiality, and nonrepudiation in all communications and data exchanges in the new information society.

Table 8.1 shows how cryptography guarantees these security services through five basic mechanisms that include symmetric and public key encryption, hashing, digital signatures, and certificates.

A cryptographic system consists of four essential components:

- Plaintext – the original message to be sent.
- Cryptographic system (cryptosystem) or a cipher – consisting of mathematical encryption and decryption algorithms.
- Ciphertext – the result of applying an encryption algorithm to the original message before it is sent to the recipient.
- Key – a string of bits used by the two mathematical algorithms in encrypting and decrypting processes.

A cipher or a cryptosystem is a pair of invertible functions, one for encrypting or enciphering and the other for decryption or deciphering. The word *cipher* has its origin in an Arabic word *sifr*, meaning *empty* or *zero*. The encryption process uses the cryptographic algorithm, known as the encryption algorithm, and a selected key to transform the plaintext data into an encrypted form called ciphertext, usually unintelligible form. The cipher text can then be transmitted across the communication channels to the intended destination.

A cipher can either be a stream cipher or a block cipher. Stream ciphers rely on a key derivation function to generate a key stream. The key and an algorithm are then applied to each bit, one at a time. Even though stream ciphers are faster and smaller to implement, they have an important security gap. If the same key stream is used, certain types of attacks may cause the information to be revealed. Block ciphers, on the other hand, break a message up into chunks and combine a key with each chunk, for example, 64 or 128 bits of text.

Table 8.1 Modern Cryptographic Security Services

<i>Security Services</i>	<i>Cryptographic Mechanism to Achieve the Service</i>
Confidentiality	Symmetric encryption
Authentication	Digital signatures and digital certificates
Integrity	Decryption of digital signature with a public key to obtain the message digest. The message is hashed to create a second digest. If the digests are identical, the message is authentic and the signer's identity is proven.
Nonrepudiation	Digital signatures of a hashed message then encrypting the result with the private key of the sender, thus binding the digital signature to the message being sent.
Nonreplay	Encryption, hashing and digital signature

8.8.1 Symmetric Encryption

Symmetric encryption or secret key encryption, as it is usually called, uses a common key and the same cryptographic algorithm to scramble and unscramble the message as shown in Figs. 8.2 and 8.3. The transmitted final ciphertext stream is usually a chained combination of blocks of the plaintext, the secret key, and the ciphertext.

The security of the transmitted data depends on the assumption that eavesdroppers and cryptanalysts with no knowledge of the key are unable to read the message. However, for a symmetric encryption scheme to work, the key must be shared between the sender and the receiver. The sharing is usually done through passing the key from the sender to the receiver. This presents a problem in many different ways, as we will see in further section. The question which arises is how to keep the key secure while being transported from the sender to the receiver.

Symmetric algorithms are faster than their counterparts, the public key algorithms.

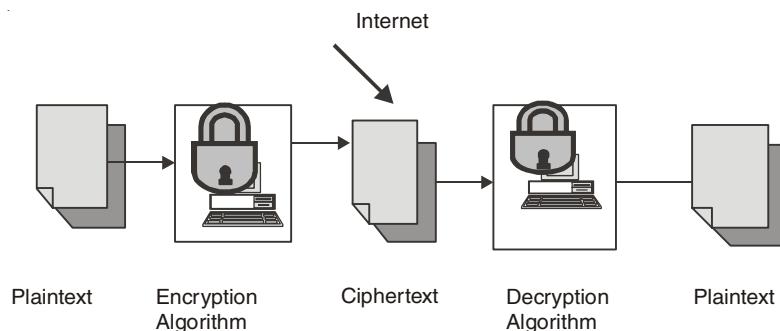


Fig. 8.2. Symmetric Encryption.

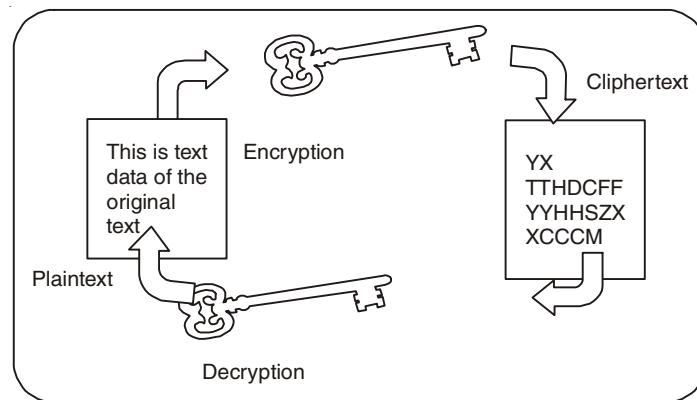


Fig. 8.3. Encryption and Decryption with Symmetric Cryptography.

➤ Symmetric Encryption Algorithms

The most widely used symmetric encryption method in the United States is the block ciphers Triple Data Encryption Standard (3DES). Triple DES developed from the original and now cracked DES uses a 64-bit key consisting of 56 effective key bits and 8 parity bits. Triple DES encrypts the data in 8-byte chunks, passing it through 16 different iterations

consisting of complex shifting, exclusive ORing, substitution, and expansion of the key along with the 64-bit data blocks. Figure 8.4 shows how Triple DES works.

Although 3DES is complicated and complex, and therefore secure, it suffers from several drawbacks including the length of its key fixed at 56 bits plus 8 bits of parity. The limited key length is making it possible for the ever-increasing speed of newer computers to render it useless as it possible to compute all possible combinations in the range $0-2^{56}-1$.

Because of this, the National Institute of Standards and Technology (NIST) has presented the Advanced Encryption Standard (AES), which is expected to replace DES. AES is Advanced Encryption Standard whose algorithm was decided to be Rijndael, developed by two Belgian researchers, Joan Daemen and Vincent Rijmen.

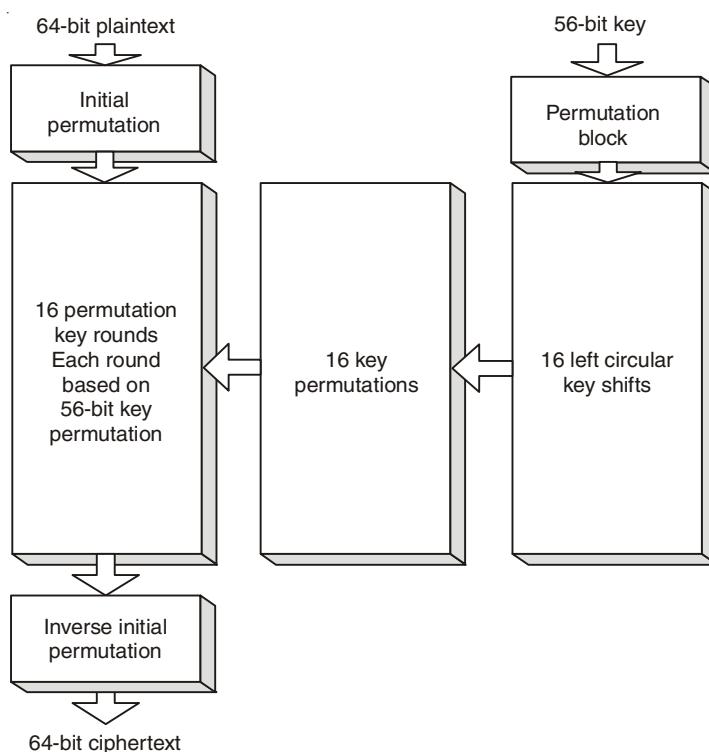


Fig. 8.4. DES Algorithm.

Several other symmetric encryption algorithms in use today include International Data Encryption Algorithm (IDEA), Blowfish, Rivest Cipher 4 (RC4), RC5, and CAST-128. See Table 8.2 for symmetric key algorithms.

➤ Problems with Symmetric Encryption

As we pointed out earlier, symmetric encryption, although fast, suffers from several problems in the modern digital communication environment. These are a direct result of the nature of symmetric encryption. Perhaps the biggest problem is that a single key must be shared in pairs of each sender and receiver. In a distributed environment with large numbers of combination pairs involved in many-to-one communication topology, it is difficult for the one recipient to keep so many keys in order to support all communication.

In addition to the key distribution problem above, the size of the communication space presents problems. Because of the massive potential number of individuals who can carry on communication in a many-to-one, one-to-many, and many-to-many topologies supported by the Internet, for example, the secret-key cryptography, if strictly used, requires billions of secret keys pairs to be created, shared, and stored. This can be a nightmare! Large numbers of potential correspondents in the many-to-one, one-to-many, and many-to-many communication topologies may cause symmetric encryption to fail because of its requirement of prior relationships with the parties to establish the communication protocols like the setting up of and acquisition of the secret key.

Besides the problems discussed above and as a result of them, the following additional problems are also observable:

- The integrity of data can be compromised because the receiver cannot verify that the message has not been altered before receipt.
- It is possible for the sender to repudiate the message because there are no mechanisms for the receiver to make sure that the message has been sent by the claimed sender.
- The method does not give a way to ensure secrecy even if the encryption process is compromised.
- The secret key may not be changed frequently enough to ensure confidentiality.

Table 8.2 Symmetric Key Algorithms

<i>Algorithm</i>	<i>Strength</i>	<i>Features (key length)</i>
3DES	Strong	64, 112, 168
AES	Strong	128, 192, 256
IDEA	Strong	64, 128
Blowfish	Weak	32–448
RC4	Weak	
RC5	Strong	32, 64, 128
BEST	Strong	
CAST-128	Strong	32, 128

8.8.2 Public Key Encryption

Since the symmetric encryption scheme suffered from all those problems we have just discussed above, there was a need for a more modern cryptographic scheme to address these flaws. The answers came from two people: Martin Hellman and Whitfield Diffie, who developed a method that seemed to solve at least the first two problems and probably all four by guaranteeing secure communication without the need for a secret key. Their scheme, consisting of mathematical algorithms, led to what is known as a *public key encryption* (PKE).

Public key encryption, commonly known asymmetric encryption, uses two different keys, a public key known to all and a private key known only to the sender and the receiver. Both the sender and the receiver own a pair of keys, one public and the other a closely guarded

private one. To encrypt a message from sender A to receiver B, as shown in Fig. 8.5 , both A and B must create their own pairs of keys.

Then A and B publicize their public keys—anybody can acquire them. When A has to send a message M to B, A uses B's public key to encrypt M. On receipt of M, B then uses his or her private key to decrypt the message M. As long as only B, the recipient, has access to the private key, then A, the sender, is assured that only B, the recipient, can decrypt the message. This ensures data confidentiality. Data integrity is also ensured because for data to be modified by an attacker, it requires the attacker to have B's, the recipient's, private key. Data confidentiality and integrity in public key encryption is also guaranteed in Fig. 8.5.

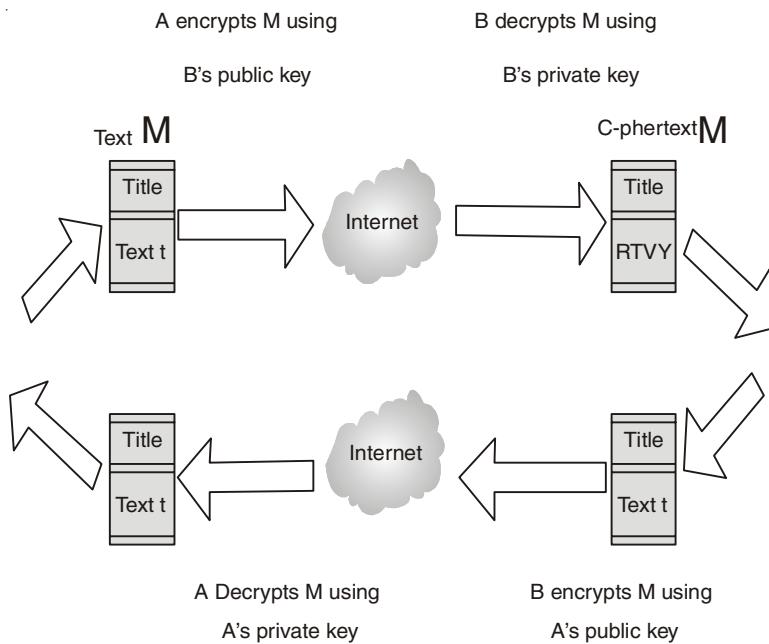
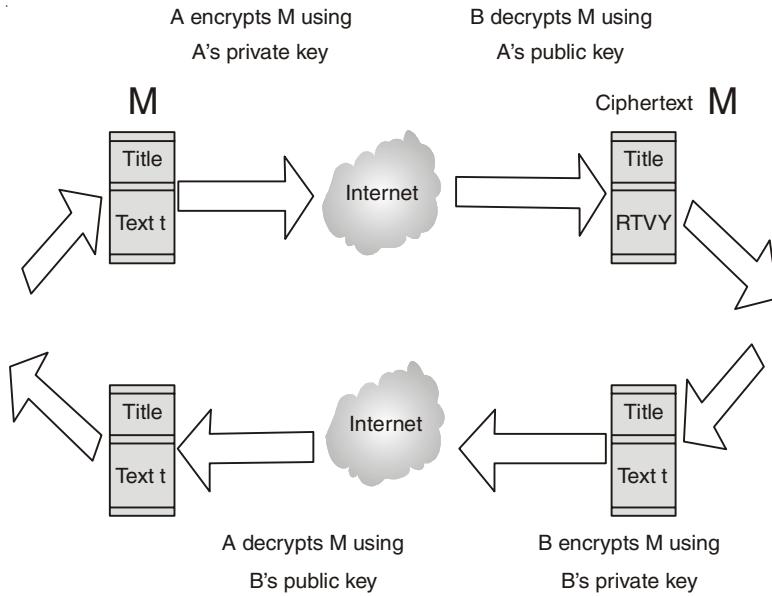
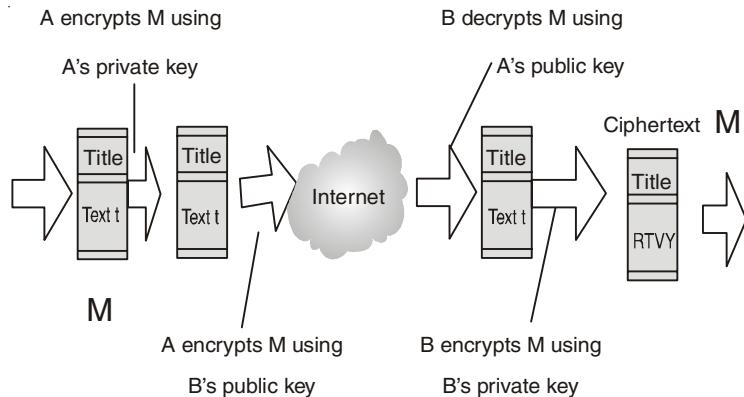


Fig. 8.5. Public Key Encryption with Data Integrity and Confidentiality.

As can be seen, ensuring data confidentiality and integrity does not prevent a third party, unknown to both communicating parties, from pretending to be A, the sender. This is possible because anyone can get A's, the sender's public key. This weakness must, therefore, be addressed, and the way to do so is through guaranteeing of sender nonrepudiation and user authentication. This is done as follows: after both A and B have created their own pairs of keys and exchanged the public key pair, A, the sender, then encrypts the message to be sent to B, the recipient, using the sender's private key. Upon receipt of the encrypted message, B, the recipient, then uses A's, the sender's public key to encrypt the message. The return route is also similar. This is illustrated in Fig 8.6 Authentication of users is ensured because only the sender and recipient have access to their private keys. And unless their keys have been compromised, both cannot deny or repudiate sending the messages.

**Fig. 8.6. Authentication and Non-repudiation**

To ensure all four aspects of security, that is data confidentiality and integrity and authentication and nonrepudiation of users, a double encryption is required as illustrated in Fig. 8.7.

**Fig. 8.7. Ensuring Data Confidentiality and Integrity and User Authentication and Non-repudiation.**

The core of public key encryption is that no secret key is passed between two communicating parties. This means that this approach can support all communication topologies including one-to-one, one-to-many, many-to-many, and many-to-one, and along with it, several to thousands of people can communicate with one party without exchange of keys. This makes it suitable for Internet communication and electronic commerce applications. Its other advantage is that it solves the chronic repudiation problem experienced by symmetric encryption. This problem is solved, especially in large groups, by the use of digital signatures and certificates. The various cryptographic algorithms used in this scheme

rely on the degree of computational difficulty encountered as an attempt is made to recover the keys. These algorithms, should be labor intensive and the amount and difficulty involved should, and actually always, increase with the key length. The longer the key, the more difficult and the longer it should take to guess the key, usually the private key.

➤ Problems with Public Key Encryption

Although public key encryption seems to have solved the major chronic encryption problems of key exchange and message repudiation, it still has its own problems. The biggest problem for public key cryptographic scheme is speed. Public key algorithms are extremely slow compared to symmetric algorithms. This is because public key calculations take longer than symmetric key calculations since they involve the use of exponentiation of very large numbers which in turn take longer to compute. For example, the fastest public key cryptographic algorithm such as RSA is still far slower than any typical symmetric algorithm. This makes these algorithms and the public key scheme less desirable for use in cases of long messages.

In addition to speed, public key encryption algorithms have a potential to suffer from the *man-in-the-middle* attack. The man-in-the-middle attack is a well known attack, especially in the network community where an attacker sniffs packets off a communication channel, modifies them, and inserts them back on to the channel. In case of an encryption channel attack, the intruder convinces one of the correspondents that the intruder is the legitimate communication partner.

8.8.3 Enhancing Security: Combining Symmetric and Public Key Encryptions

As we noted in Section 8.8.1, symmetric algorithms, although faster than public key algorithms, are beset with a number of problems. Similarly public key encryption also suffers slowness and the potential of the “man-in-the-middle” attacker. To address these concerns and to preserve both efficiency and privacy of the communication channel and increase the performance of the system, a hybrid cryptosystem that uses the best of both and at the same time mitigating the worst in each system is widely used.

8.9 FIREWALLS

The rapid growth of the Internet has led to a corresponding growth of both users and activities in cyberspace. Unfortunately, not all these users and their activities are reputable; thus, the Internet has been increasingly, at least to many individuals and businesses, turning into a “bad Internet.” Bad people are plowing the Internet with evil activities that include, among other things, intrusion into company and individual systems looking for company data and individual information that erodes privacy and security. There has, therefore, been a need to protect company systems, and now individual PCs, keeping them out of access from those “bad users” out on the “bad Internet.” So network system administrators must be able to find ways to restrict access to the company network or sections of the network from both the “bad Internet” outside and from unscrupulous inside users.

Such security mechanisms are based on a *firewall*. A firewall is a hardware, software, or a combination of both that monitors and filters traffic packets that attempt to either enter or leave the protected private network. It is a tool that separates a protected network or part of a network, and now increasingly a user PC, from an unprotected network—the “bad network” like the Internet. In many cases the “bad network” may even be part of the company network. By definition, a “firewall,” is a tool that provides a filter of both incoming and outgoing packets. Most firewalls perform two basic security functions:

- Packet filtering based on *accept* or *deny* policy that is itself based on rules of the security policy.
- Application proxy gateways that provide services to the inside users and at the same time protect each individual host from the “bad” outside users.

By denying a packet, the firewall actually drops the packet. In modern firewalls, the firewall logs are stored into log files and the most urgent or dangerous ones are reported to the system administrator. This reporting is slowly becoming real time.

In its simplest form, a firewall can be implemented by any device or tool that connects a network or an individual PC to the Internet. For example, an Ethernet bridge or a modem that connects to the “bad network” can be set as a firewall. Most firewalls products actually offer much more as they actively filter packets from and into the organization network according to certain established criteria based on the company security policy. Most organization firewalls are *bastion host*, although there are variations in the way this is set up. A bastion host is one computer on the organization network with bare essential services, designated and strongly fortified to withstand attacks. This computer is then placed in a location where it acts as a gateway or a choke point for all communication into or out of the organization network to the “bad network.” This means that every computer behind the bastion host must access the “bad network” or networks through this bastion host. Figure 8.8 shows the position of a bastion host in an organization network.

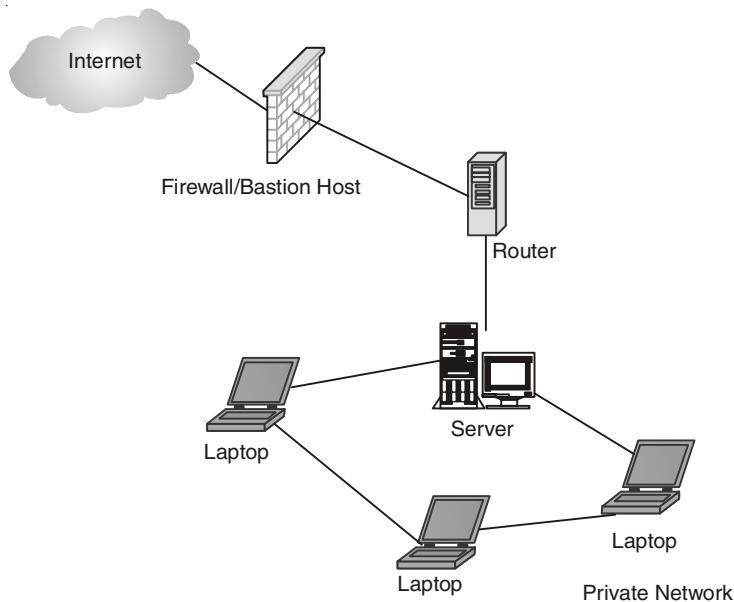


Fig. 8.8. Bastion Host between Private and Bad Network.

For most organizations, a firewall is a network perimeter security, a first line of defense of the organization's network that is expected to police both network traffic inflow and outflow. This perimeter security defense varies with the perimeter of the network. For example, if the organization has an extranet, an extended network consisting of two or more LAN clusters, or the organization has a Virtual Private Network (VPN), then the perimeter of the organization's network is difficult to define. In this case, then each component of the network should have its own firewall. See Fig. 8.9.

As we pointed out earlier, the accept/deny policy used in firewalls is based on an organization's security policy. The security policies most commonly used by organizations vary ranging from completely disallowing some traffic to allowing some of the traffic or all the traffic. These policies are consolidated into two commonly used firewall security policies.

- Deny-everything-not-specifically-allowed which sets the firewall in such a way that it denies all traffic and services except a few that are added as the organization needs develop.
- Allow-everything-not-specifically-denied which lets in all the traffic and services except those on the "forbidden" list which is developed as the organization's dislikes grow.

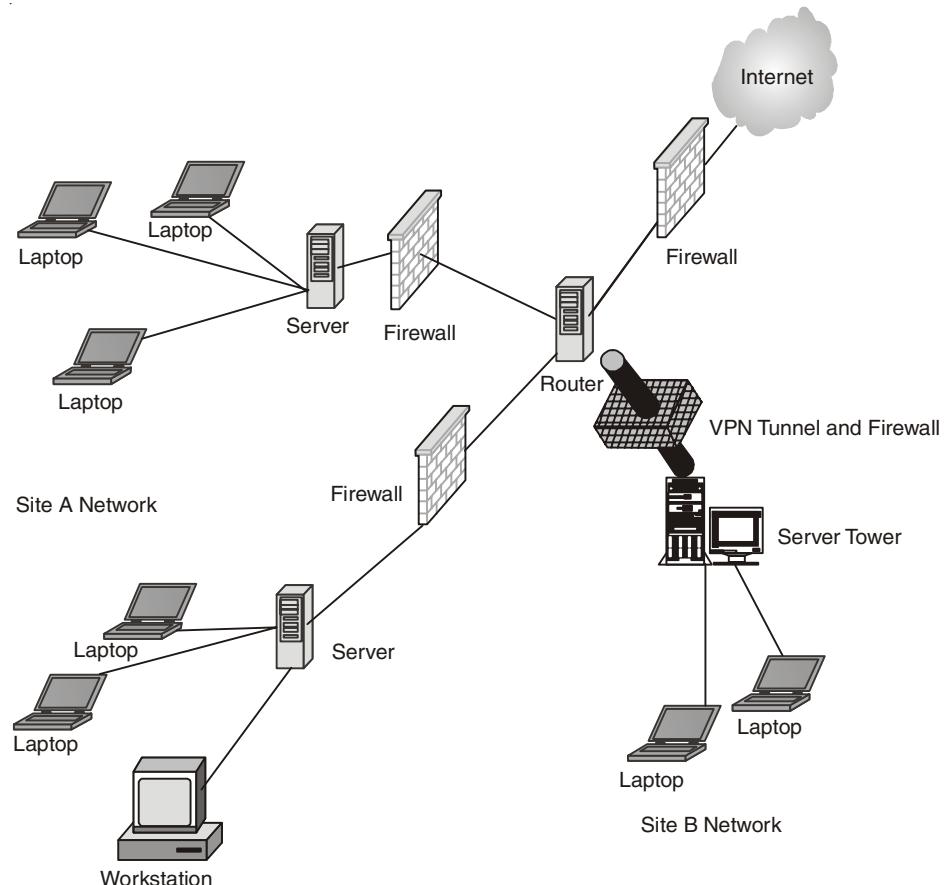


Fig. 8.9. Firewall in changing parameter security.

Based on these policies, the following design goals are derived:

- All traffic into and out of the protected network must pass through the firewall.
- Only authorized traffic, as defined by the organizational security policy, in and out of the protected network, will be allowed to pass.
- The firewall must be immune to penetration by use of a trusted system with secure operating system.

When these policies and goals are implemented in a firewall, then the firewall is supposed to

- Prevent intruders from entering and interfering with the operations of the organization's network. This is done through restricting which packets can enter the network based on IP addresses or port numbers. Prevent intruders from deleting or modifying information either stored or in motion within the organization's network.
- Prevent intruders from acquiring proprietary organization information.
- Prevent insiders from misusing the organization resources by restricting unauthorized access to system resources.
- Provide authentication, although care must be taken because additional services to the firewall may make it less efficient.
- Provide end-points to the VPN.

8.9.1 Types of Firewalls

Firewalls are used very widely to offer network security services. This has resulted in a large repertoire of firewalls. To understand the many different types of firewalls, we need only look at the kind of security services firewalls offer at different layers of the TCP/IP protocol stack.

As Table 8.3 shows, firewalls can be set up to offer security services to many TCP/IP layers. The many types of firewalls are classified based on the network layer it offers services in and the types of services offered.

Table 8.3 Types of Firewall

<i>Layer</i>	<i>Firewall services</i>
Application	Application-level gateways, encryption, SOCKS Proxy Server
Transport	Packet filtering (TCP, UDP, ICMP)
Network	NAT, IP-filtering
Data link	MAC address filtering
Physical	May not be available.

The first type is the *packet inspection or filtering router*. This type of firewall uses a set of rules to determine whether to forward or block individual packets. A packet inspection router could be a simple machine with multiple network interfaces or a sophisticated one with multiple functionalities. The second type is the *application inspection or proxy server*.

The proxy server is based on specific application daemons to provide authentication and to forward packets. The third type is the *authentication and virtual private networks* (VPN). A VPN is an encrypted link in a private network running on a public network. The fourth firewall type is the *small office or home* (SOHO) firewall, and the fifth is the network address translation (NAT).

8.9.2 Firewall Services

The broad range of services offered by the firewall are based on the following access controls:

- Service control – where the firewall may filter traffic on the basis of IP addresses, TCP, UDP, port numbers, and DNS and FTP protocols in addition to providing proxy software that receives and interprets each service request before passing it on.
- Direction control – where permission for traffic flow is determined from the direction of the requests.
- User control – where access is granted based on which user is attempting to access the internal protected network, which may also be used on incoming traffic.
- Behavior control – in which access is granted based on how particular services are used, for example, filtering e-mail to eliminate spam.

8.9.3 Limitations of Firewalls

Given all the firewall popularity, firewalls are still taken as just the first line of defense of the protected network because they do not assure total security of the network. Firewalls suffer from limitations, and these limitations and other weaknesses have led to the development of other technologies. In fact, there is talk now that the development of IPSec technology is soon going to make firewall technology obsolete. We may have to wait and see. Some of the current firewall limitations are as follows:

- Firewalls cannot protect against a threat that bypasses it, such as a dial-in using a mobile host.
- Firewalls do not provide data integrity because it is not possible, especially in large networks, to have the firewall examine each and every incoming and outgoing data packet for anything.
- Firewalls cannot ensure data confidentiality because, even though newer firewalls include encryption tools, it is not easy to use these tools. It can only work if the receiver of the packet also has the same firewall.
- Firewalls do not protect against internal threats.
- Firewalls cannot protect against transfer of virus-infected programs or files.

Review Questions

1. What is security and Information security? What is the difference between them?
2. It has been stated that security is a continuous process; what are the states in this process?

3. What are the differences between symmetric and asymmetric key systems?
4. What is PKI? Why is it so important in information security?
5. What is the difference between authentication and nonrepudiation?
6. Why is there a dispute between digital nonrepudiation and legal nonrepudiation?
7. Virtual security seems to work in some systems. Why is this so? Can you apply it in a network environment? Support your response.
8. Security best practices are security guidelines and policies aimed at enhancing system security. Can they work without known and proven security mechanisms?
9. Does information confidentiality infer information integrity? Explain your response.
10. Comment on the rapid growth of the Internet as a contributing factor to the security threat of cyberspace. What is the responsible factor in this growth? Is it people or the number of computers?
11. There seems to have been an increase in the number of reported virus and worm attacks on computer networks. Is this really a sign of an increase, more reporting, or more security awareness on the part of the individual? Comment on each of these factors.
12. Discuss the basic components of cryptography.
13. Discuss the weaknesses of symmetric encryption.
14. Discuss the weaknesses of public key encryption.
15. Why is a hybrid cryptosystem preferred over symmetric and public key encryption systems?
16. Why is PKI so vital in modern communications?
17. Discuss the role of Firewalls in Information network security.
18. Discuss the differences between a firewall and a packet filter.
19. Give reasons why firewalls do not give total security.
20. Discuss the advantages of using an application-level firewall over a network-level firewall.



GLOSSARY

Access Control List

(ACL) A list of the services available on a server, each with a list of the hosts permitted to use the service.

Anonymous FTP

An interactive service provided by many Internet hosts allowing any user to transfer documents, files, programs, and other archived data using File Transfer Protocol. The user logs in using the special user name "ftp" or "anonymous" and his e-mail address as password. He then has access to a special directory hierarchy containing the publically accessible files, typically in a subdirectory called "pub". This is usually a separate area from files used by local users.

A reference like

ftp: euagate.eua.ericsson.se /pub/eua/erlang/info

means that files are available by anonymous FTP from the host called euagate.eua.ericsson.se in the directory (or file) /pub/eua/erlang/info. Sometimes the hostname will be followed by an Internet address in parentheses. The directory will usually be given as a path relative to the anonymous FTP login directory. A reference to a file available by FTP may also be in the form of a URL starting "ftp:".

Application Layer

The top layer of the OSI seven layer model. This layer handles issues like network transparency, resource allocation and problem partitioning. The application layer is concerned with the user's view of the network (e.g. formatting electronic mail messages). The presentation layer provides the application layer with a familiar local representation of data independent of the format used on the network.

Access Point

A hardware device or a computer's software that acts as a communication hub for users of a wireless device to connect to a wired LAN.

AI

Stands for Artificial Intelligence. This is the area of computer science focusing on creating machines that can engage on behaviors that humans consider intelligent.

The ability to create intelligent machines has intrigued humans since ancient times, and today with the advent of the computer and 50 years of research into AI programming techniques, the dream of smart machines is becoming a reality. Researchers are creating systems which can mimic human thought, understand speech, beat the best human chess player, and countless other feats never before possible.

ACPI

A hardware device or a computer's software that acts as a communication hub for users of a wireless device to connect to a wired LAN.

Access Point

A hardware device or a computer's software that acts as a communication hub for users of a wireless device to connect to a wired LAN.

Actuator

Device that performs an action or outputs a signal in response to a signal from a computer.

Addressing

A method of identifying a resource (such as a program) or piece of information (such as a file) on a network. Methods of addressing vary considerably from network-to-network.

AGP

Short for Accelerated Graphics Port, a new interface specification developed by Intel Corporation. AGP is based on PCI, but is designed especially for the throughput demands of 3-D graphics. Rather than using the PCI bus for graphics data, AGP introduces a dedicated point-to-point channel so that the graphics controller can directly access main memory. The AGP channel is 32 bits wide and runs at 66 MHz. This translates into a total bandwidth of 266 Mbps, as opposed to the PCI bandwidth of 133 Mbps. AGP also supports two optional faster modes, with throughputs of 533 Mbps and 1.07 GBps. In addition, AGP allows 3-D textures to be stored in main memory rather than video memory.

AGTL Signaling

(Assisted Gunning Transistor Logic) AGTL and AGTL+ use the same signaling protocol only at different voltage levels. AGTL+ operates at 1.5V signaling levels, while AGTL operates at 1.25V.

Algorithm

A formal set of instructions that can be followed to perform a specific task, such as a mathematical formula or a set of instructions in a computer program.

AMD

(Advanced Microchip Devices) A semiconductor manufacturer and is a major competitor of Intel. They manufacture the Athlon, Duron, and K6 CPU chips.

ANSI

It stands for American National Standards Institute. The Institute's mission is to enhance both the global competitiveness of U.S. business and the U.S. quality of life by promoting and facilitating voluntary consensus standards and conformity assessment systems, and safeguarding their integrity.

AppleTalk

A protocol suite developed by Apple Computer in the early 1980s, was developed in conjunction with the Macintosh computer. AppleTalk's purpose was to allow multiple users to share resources, such as files and printers. The devices that supply these resources are called servers, while the devices that make use of these resources (such as a user's Macintosh computer) are referred to as clients. Hence, AppleTalk is one of the early implementations of a distributed client/server networking system.

AppleScript

It is an English-like language used to write script files which can automate the actions of the computer and the applications which run on it.

AppleShare

This is Apple's network system. It is to the Macintosh what FTP is to the PC.

Applet

An applet is a small program, generally written in the Java programming language that is usually translated by browsers thus allowing interactivity on Web pages.

Application

A software program designed to perform a specific task or group of tasks, such as word processing, communications, or database management.

Archie

Or ArchiePlex which is an Archie gateway for the World Wide Web. It can locate files on Anonymous FTP sites in the Internet.

ASCII

It stands for American Standard Code Information Interchange and is pronounced (ask-ee). A standard code or protocol for displaying characters and transferring data between computers and associated equipment. It was developed for the purpose of information exchange among the following:

- Associated equipment
- Data communications systems
- Data processing systems.

There are 128 standard ASCII codes each of which can be represented by a 7 digit binary number: 0000000 through 1111111.

Asynchronous

1: not synchronous; not occurring or existing at the same time or having the same period or phase. 2: of, used in, or being digital communication (as between computers)

in which there is no timing requirement for transmission and in which the start of each character is individually signaled by the transmitting device.

ASP

Stands for Active Server Pages, which is an open, compile-free application environment in which you can combine HTML, scripts, and reusable ActiveX server components to create dynamic and powerful Web-based business solutions. Active Server Pages enables server side scripting for IIS with native support for both VBScript and Jscript.

ATA

Short for Advanced Technology Attachment, a disk drive implementation that integrates the controller on the disk drive itself. There are several versions of ATA, all developed by the Small Form Factor (SFF) Committee:

- ATA: Known also as IDE, supports one or two hard drives, a 16-bit interface and PIO modes 0, 1 and 2.
- ATA-2: Supports faster PIO modes (3 and 4) and multiword DMA modes (1 and 2). Also supports logical block addressing (LBA) and block transfers. ATA-2 is marketed as Fast ATA and Enhanced IDE (EIDE).
- ATA-3: Minor revision to ATA-2.
- Ultra-ATA: Also called Ultra-DMA, ATA-33, and DMA-33, supports multiword DMA mode 3 running at 33 Mbps.
- ATA/66: A version of ATA proposed by Quantum Corporation, and supported by Intel, that doubles ATA's throughput to 66 Mbps.
- ATA/100: An updated version of ATA/66 that increases data transfer rates to 100 Mbps.

ATM

Asynchronous Transfer Mode — International standard for cell relay in which multiple service types (such as voice, video, or data) are conveyed in fixed-length (53-byte) cells. Fixed-length cells allow cell processing to occur in hardware, thereby reducing transit delays. ATM is designed to take advantage of high-speed transmission media such as E3, SONET, and T3.

ATX

The modern-day shape and layout of PC motherboards. It improves on the previous standard, the Baby AT form factor, by rotating the orientation of the board 90 degrees. This allows for a more efficient design, with disk drive cable connectors nearer to the drive bays and the CPU closer to the power supply and cooling fan.

AVI

Stands for Audio/Video Interleaved. AVI is the most common format for audio/video data on the PC.

Backbone

This term is often used to describe the main line or series of connections in a network. The backbones of the Internet are high-speed data highways serving as a major access points to which other networks connect.

Backup

To copy files to a second medium (a disk or tape) as a precaution in case the first medium fails. One of the cardinal rules in using computers is back up your files regularly. Even the most reliable computer is apt to break down eventually. Many professionals recommend that you make two, or even three, backups of all your files. To be especially safe, you should keep one backup in a different location from the others.

You can back up files using operating system commands, or you can buy a special-purpose backup utility. Backup programs often compress the data so that backups require fewer disks.

(1) The act of backing up. (2) A substitute or alternative. The term backup usually refers to a disk or a tape that contains a copy of data.

Bandwidth

How much stuff you can send through a connection. Usually measured in bits-per-second. A full page of English text is about 16,000 bits. A fast modem can move about 57,000 bits in one second. Full-motion full-screen video would require roughly 10,000,000 bits-per-second, depending on compression.

Baud

Pronounced bawd. The term is named after J.M.E. Baudot, the inventor of the Baudot telegraph code. In common usage the baud rate of a modem is how many bits it can send or receive per second. Technically, baud is the number of times per second that the carrier signal shifts value—for example a 1200 bit-per-second modem actually runs at 300 baud, but it moves 4 bits per baud ($4 \times 300 = 1200$ bits per second).

BBS

Stands for Bulletin Board Service. A computerized meeting and announcement system that allows people to carry on discussions, upload and download files, and make announcements without the people being connected to the computer at the same time.

Beta

Preliminary or testing stage of a software or hardware product; “a beta version”; “beta software”.

BIOS

Stands for Audio/Video Interleaved. AVI is the most common format for audio/video data on the PC. Stands for Basic Input/Output System. The BIOS gives the computer a little built-in starter kit to run the rest of softwares from floppy disks (FDD) and hard disks (HDD). The BIOS is responsible for booting the computer by providing a basic set of instructions.

Binary

This is a basic system of numbering using ones and zeros.

Bit

(Binary DigIT) A single digit number in base-2, in other words, either a 1 or a zero. The smallest unit of computerized data. Bandwidth is usually measured in bits-per-second.

Binary

This is a basic system of numbering using ones and zeros.

Blog

(Slang term for a Weblog) This is a publicly accessible personal journal for an individual. Similar to a personal diary, but shared over the web. The activity of updating a blog is “blogging” and someone who keeps a blog is a “blogger.” Blogs are typically updated daily using software that allows people with little or no technical background to update and maintain the blog. Postings on a blog are almost always arranged in chronological order with the most recent additions featured most prominently.

Bluetooth

Radio technology that connects electronic devices without using a cable. Data and voice can be exchanged at ranges of up to 10 meters without the need for devices to be lined up together.

BMP

(pronounced “bimp”): It’s a bitmap, an image made up of little dots.

BNC

Different sources expand BNC as Bayonet Navy Connector, British Naval Connector, Bayonet Neill Concelman or Bayonet Nut Connection. A connector widely used in the CCTV industry, usually for coaxial cable. Easy to install and reliable with little video signal loss. Pictures of BNC Connectors.

Boot Disk

A diskette from which you can boot your computer. Normally, your computer boots from a hard disk, but if the hard disk is damaged (for example, by a virus), you can boot the computer from a bootable diskette. For this reason, it’s a good idea to make sure you always have a bootable diskette on hand. In Windows 95, you can create a bootable diskette by following these steps:

1. Insert a blank, formatted diskette in the floppy drive
2. Select Start->Settings->Control Panel
3. Open Add/Remove Programs
4. Select the Startup Disk tab and press the Create Disk... button.

A bootable diskette is also called a bootable floppy, boot disk, and startup disk.

Buffer

A place, especially in RAM, for the temporary storage of data for the purpose of speeding up an operation such as printing or disk access. Data from a buffer is available more quickly than data from where the buffer got it. Typically buffers get data before it is needed so it will be ready quickly when it is needed. Similar to cache.

Browser

A browser is the software used for viewing pages on the web. Two examples are Microsoft Internet Explorer and Netscape Navigator.

BUS

A collection of wires through which data is Transmitted from one part of a computer to another. You can think of a bus as a highway on which data Travels within a computer. When used reference to Personal computers, the term bus usually refers to Internal bus.

This is a bus that connects all the internal computer components to the CPU and main memory. There's also an expansion bus that enables expansion boards to access the CPU and memory.

All buses consist of two parts—an address bus and a data bus. The data bus transfers actual data whereas the address bus transfers information about where the data should go.

The size of a bus, known as its width, is important because it determines how much data can be transmitted at one time. For example, a 16-bit bus can transmit 16 bits of data, whereas a 32-bit bus can transmit 32 bits of data.

Every bus has a clock speed measured in MHz. A fast bus allows data to be transferred faster, which makes applications run faster. On PCs, the old ISA bus is being replaced by faster buses such as PCI.

Nearly all PCs made today include a local bus for data that requires especially fast transfer speeds, such as video data. The local bus is a high-speed pathway that connects directly to the processor.

In networking, a bus is a central cable that connects all devices on a local-area network (LAN). It is also called the backbone.

Bus Mastering

A technique that allows certain advanced bus architectures to delegate control data transfers between the CPU and associated peripheral devices to an add-in board. This gives greater system bus access and higher data transfer rates than conventional systems.

Bridge

A device which forwards traffic between network segments based on data link layer information. These segments would have a common network layer address.

Every network should only have one root bridge.

Byte

Unit of memory or data needed to represent one character in binary (1s and 0s) form. One byte is usually 8 bits.

Cable modem (CM)

Client device for providing data over a cable TV network.

Cache

A special block of fast memory used for temporary storage of data for quick retrieval. Compare buffer. (Think of a buffer as a temporary holding place between two devices, and a cache as a temporary holding place for one device.)

Cache RAM

Cache (usually SRAM) stores frequently requested data and instructions. It is a small block of high-speed memory located between the CPU and the main memory. When your computer processor needs data, it will check the Cache first to see if it is there. If the data is not there, it will retrieve it from the slower main memory.

CATV

Community Antenna Television or Cable TV system. Can be all coaxial or HFC (Hybrid Fiber Coax) based.

CD-ROM

(Compact Disc Read-Only Memory) An optical storage medium that can hold about 600 MB of data and is accessed with lasers.

CGA

Stands for Color Graphics Adapter. IBM's first microcomputer color standard. CGA allowed a maximum of four colors at a resolution of 320 x 200 or two colors at 640 x 200.

CGI

The Common Gateway Interface. A standard for running external programs from a World-Wide Web HTTP server. What is returned from the CGI program is based on what was requested, and this information can be accessed and returned to the user in many different ways.

Chassis

The physical framework of the computer system that houses modules, wiring, and power supplies.

CICS

Customer Information Control System. A general purpose IBM mainframe-based transaction management system. CICS is one of IBM's most widely used database/data communications subsystems.

Clock Speed

The clock speed is the frequency which determines how fast devices that are connected to the system bus operate. The speed is measured in millions of cycles per second (MHz or megahertz) and is generated by a quartz crystal on the motherboard which acts as a kind of metronome. Devices that are synchronized with the clock may run faster or slower but their speed is determined by multiplying or dividing a factor by the clock speed.

Clustering

A technique in which two or more servers are interconnected and can access a common storage pool. Clustering prevents the failure of a single file server from denying access to data and adds computing power to the network for large numbers of users.

CMOS

Abbreviation of complementary metal oxide semiconductor. Pronounced see-moss, CMOS is a widely used type of semiconductor. CMOS semiconductors use both NMOS (negative polarity) and PMOS (positive polarity) circuits. Since only one of the circuit types is on at any given time, CMOS chips require less power than chips using just one type of transistor. This makes them particularly attractive for use in battery-powered devices, such as portable computers. Personal computers also contain a small amount of battery-powered CMOS memory to hold the date, time, and system setup parameters.

CMTS

Cable Modem Termination System. Central device for connecting the cable TV network to a data network like the internet. Normally placed in the headend of the cable TV system.

COBOL

Stands for Common Business Oriented Language. A computer programming language invented during the second generation of computers and designed to meet the needs of business. Although less often used today, it was well-suited for writing programs that process large files and generate reports.

Codec

1. Short for compressor/decompressor, a codec is any technology for compressing and decompressing data. Codecs can be implemented in software, hardware, or a combination of both. Some popular codecs for computer video include MPEG, Indeo and Cinepak.
2. In telecommunications, (short for coder/decoder) a device that encodes or decodes a signal. For example, telephone companies use codecs to convert binary signals transmitted on their digital networks to analog signals converted on their analog networks.
3. Translation of a binary value into a voltage that can be transmitted over a wire.

Cold Boot

A cold reboot (also known as a hard reboot) is when power to a computer is cycled (turned on and off) or a special reset signal to the processor is triggered (from a front panel switch of some sort). This restarts the computer without first performing the usual shut-down procedure. (With many operating systems, especially those with disc caches, after a hard reboot the system may well be in an “unclean” state, and require that checks and repairs to on-disc file system structures be performed before normal operation can begin.) It may be caused by power failure, be done by accident, or be done deliberately as a last resort because nothing else to retrieve the system from a “hung” state works.

Collision

When two hosts transmit on a network at once causing their packets to collide and corrupt each other.

Collision Detection

A class of methods for sharing a data transmission medium in which hosts transmit as soon as they have data to send and then check to see whether their transmission has suffered a collision with another host's.

If a collision is detected then the data must be resent. The resending algorithm should try to minimise the chance that two hosts's data will repeatedly collide. For example, the CSMA/CD protocol used on Ethernet specifies that they should then wait for a random time before re-transmitting.

COM

Stands for Component Object Module. In DOS systems, the name of a serial communications port. DOS supports four serial ports: COM1, COM2, COM3, and COM4. However, most software uses system interrupts to access the serial ports, and there are only two IRQ lines reserved. This means that the four COM ports share the same two IRQ lines. Typically, COM1 and COM3 use IRQ4, while COM2 and COM4 use IRQ3. So in general, if you have two devices, one of which is attached to COM1 and the other to COM3, you cannot use them simultaneously.

Command Line

Commands you type to run an application. You can type commands at an MS-DOS prompt or in the Run dialog box in the Program Manager of Windows. Interfaces in which you type commands rather than choose them from a menu are often called command line interfaces. MS-DOS has a command line interface while the Macintosh does not.

Compiler

This is an application that converts a programming language into a machine language program.

Config.SYS

The configuration file for DOS systems. Whenever a DOS computer boots up, it reads the CONFIG.SYS file (if it exists) and executes any commands in it. The most common commands are BUFFERS= and FILES=, which enable you to specify the buffer size and the number of files that can be open simultaneously. In addition, you can enter commands that install drivers for devices.

Contrast Ratio

TIs a method of measuring a dynamic range. The higher the contrast ratio, the more detailed the image will be. Blacks will be blacker, whites will be whiter, and particularly text on the image will be more vivid. A typical LCD monitor will have a contrast ratio of about 200:1.

Controller

A device that controls the transfer of data from a computer to a peripheral device and vice versa. For example, disk drives, display screens, keyboards, and printers all require controllers. In personal computers, the controllers are often single chips. When you purchase a computer, it comes with all the necessary controllers for standard components, such as the display screen, keyboard, and disk drives. If you attach additional devices, however, you may need to insert new controllers that come on expansion boards.

Controllers must be designed to communicate with the computer's expansion bus. There are three standard bus architectures for PCs—the AT bus, PCI (Peripheral Component Interconnect), and SCSI. When you purchase a controller, therefore, you must ensure that it conforms to the bus architecture that your computer uses.

Conventional Memory

On DOS systems, conventional memory refers to the portion of memory that is available to standard DOS programs. DOS systems have an address space of 1MB (megabyte), but the top 384K (called high memory) is reserved for system use. This leaves 640K of conventional memory. Everything above 1MB is either extended or expanded memory.

CPE

Customer Premises Equipment. Used to describe the PC and/or other equipment, that the customer may want to connect to the cable modem.

CPM: (Cost Per Thousand)

The practice of calculating a cost per 1000 ad displays. It is used by programs that pay on an impression basis—with the CPM rate being the amount you earn for every 1000 times an advertisement is displayed. For example, a \$5 CPM means you earn \$5 every time 1000 ads are displayed on your site. CPM can also be calculated for pay-per-sale, pay-per-lead and pay-per-click programs by using this formula: Amount earned/(number of impressions/1000).

CPU

Central Processing Unit. In a microcomputer, a processor on an IC chip (called a microprocessor) that serves as the heart of the computer. It interprets and carries out instructions, performs numeric computations, and controls the peripherals connected to it. Often the entire system unit is called the CPU.

CSS

Stands for Cascading Style Sheets.

Cyberspace

Author William Gibson in his novel *Neuromancer* describes a more highly developed form of the Internet and who originally coined the term Cyberspace. The word Cyberspace is currently used to describe the whole range of information resources available through computer networks.

Cyberpunk

Cyberpunk was originally a cultural sub-genre of science fiction taking place in a not-so-distant, dystopian, over-industrialized society. The term grew out of the work of William Gibson and Bruce Sterling and has evolved into a cultural label encompassing many different kinds of human, machine, and punk attitudes. It includes clothing and lifestyle choices as well.

Daisy Chain

A hardware configuration in which devices are connected one to another in a series. The SCSI interface, for example, supports a daisy chain of up to 7 devices.

Data

Anything that is recorded or used for processing. The stuff that transfers between computers needed a name—data seemed good.

Data Bus

A group of parallel conductors (circuit traces) found on the motherboard that is used by the CPU to send and receive data from all the devices in the computer. Also called the external data bus.

Data Conversion

The translation of data from one format to another. Often when data are moved from one system to another, some form of data conversion is required to convert the data to a format the receiving system can interpret. Sometimes it is necessary to have an intermediate format.

Database

Anything that accepts data is a database. A pile of newspapers is a database. A computer database has the ability to manipulate that data. It is possible to attach applications to that database to search the contents.

Data Mining

Sorting through data to identify patterns and establish relationships. Data mining parameters include:

- Association - looking for patterns where one event is connected to another event.
- Sequence or path analysis - looking for patterns where one event leads to another later event.
- Classification - looking for new patterns (May result in a change in the way the data is organized but that's ok).
- Clustering - finding and visually documenting groups of facts not previously known.
- Forecasting - discovering patterns in data that can lead to reasonable predictions about the future.

Data Rate

Speed that information moves from one item to another. This is usually in the form of bits.

Daughter Card

A printed circuit board that plugs into another circuit board (usually the motherboard). A daughter card is similar to an expansion board, but it accesses the motherboard components (memory and CPU) directly instead of sending data through the slower expansion bus.

DDR

Stands for "Double Data Rate." It is an advanced version of SDRAM, a type of computer memory. DDR-SDRAM, sometimes called "SDRAM II," can transfer data twice as fast as regular SDRAM chips. This is because DDR memory can send and receive signals twice per clock cycle. The efficient operation of DDR-SDRAM makes the memory great for notebook computers since it uses up less power.

DDS

Stands for Direct Digital Signal. A network whose infrastructure equipment is completely digital. All signals on such a network are transmitted digitally and there is no need for analog-to-digital converters.

Decoder

A circuit or device that restores a coded signal to its original form based on knowledge of the process used to code the signal.

Decryption

Decryption is the process of converting encrypted data back into its original form, so it can be understood.

Dedicated Line

This is a phone line meant specifically for one thing, like being attached to a computer.

Degauss

To remove magnetism from a device. The term is usually used in reference to color monitors and other display devices that use a Cathode Ray Tube (CRT). These devices aim electrons onto the display screen by creating magnetic fields inside the CRT. External magnetic forces—such as the earth's natural magnetism or a magnet placed close to the monitor—can magnetize the shadow mask, causing distorted images and colors. To remove this external magnetic forces, most monitors automatically degauss the CRT whenever you turn on the monitor. In addition, many monitors have a manual degauss button that performs a more thorough degaussing of the CRT. You can also use an external degausser that degausses the monitor from the outside. Since it may be impossible to remove the external magnetic force, degaussing works by re-aligning the magnetic fields inside the CRT to compensate for the external magnetism.

Delimiter

A text character that marks the beginning and/or end of a unit of data or separates different data components. For example, periods are used as delimiters in domain names, hyphens and parentheses are used in phone numbers and social security numbers, and blank spaces and commas are used in written text. In HTML the opening delimiter of an element or tag is the less than symbol, <, and the closing delimiter is greater than symbol, >.

Demodulation

This is the process of converting analog information (like over phone lines) into digital information (like in a computer). See "Modem" for more.

DHCP

Dynamic Host Configuration Protocol. This protocol provides a mechanism for allocating IP addresses dynamically so that addresses can be reused. Often used for managing the IP addresses of all the cable modems in a cable plant and the PC's connected to the cable modems.

DHTML

Stands for Dynamic HTML.

Dial-Up Line

This is a telephone line that is connected to a server. When it is called, tones are exchanged between the server and the device calling in order to attach.

Dial-Up Networking

A component in Windows 95 that enables you to connect your computer to a network via a modem. If your computer is not connected to a LAN and you want to connect to the Internet, you need to configure Dial-Up Networking (DUN) to dial a Point of Presence (POP) and log into your Internet Service Provider (ISP). Your ISP will need to provide certain information, such as the gateway address and your computer's IP address. You access DUN through the My Computer icon. You can configure a

different profile (called a connectoid) for each different online service you use. Once configured, you can copy a connectoid shortcut to your desktop so that all you need to do to make a connection is double-click the connectoid icon.

Digital

A system that defines data in a discrete, non-fluctuating (i.e., non-analogue), numerical method. Similar to a binary system.

DIMM

Short for dual in-line memory module, a small circuit board that holds memory chips. A single in-line memory module (SIMM) has a 32-bit path to the memory chips whereas a DIMM has 64-bit path. Because the Pentium processor requires a 64-bit path to memory, you need to install SIMMs two at a time. With DIMMs, you can install memory one DIMM at a time.

DIP

Acronym for dual in-line package, a type of chip housed in a rectangular casing with two rows of connecting pins on either side.

Direct X

A set of APIs developed by Microsoft that enables programmers to write programs that access hardware features of a computer without knowing exactly what hardware will be installed on the machine where the program eventually runs. DirectX achieves this by creating an intermediate layer that translates generic hardware commands into specific commands for particular pieces of hardware. In particular, DirectX lets multimedia applications take advantage of hardware acceleration features supported by graphics accelerators. DirectX 2, released in 1996, supports the Direct3D architecture. DirectX 5, released in 1998, adds new layers to the DirectX API. In addition to the low-level layer that communicates directly with multimedia hardware, DirectX 5 also includes a Media layer that enables programmers to manipulate multimedia objects and streams. DirectX 5 also supports USB and IEEE 1394 buses, AGP, and MMX.

DLL

Dynamic Link Library. A file of functions, compiled, linked, and saved separately from the processes that use them. Functions in DLLs can be used by more than one running process. The operating system maps the DLLs into the process's address space when the process is started up or while it is running. Dynamic link libraries are stored in files with the .DLL file extension.

DMA

Hardware devices attached to PCs (ranging from keyboards to sound cards) can be designed to send their instructions to and from main memory in one of two ways. The default is to ask the CPU to do the work. The more efficient way is to allocate one of the PC's DMA channels to send instructions directly to memory. This leaves the CPU free to do more important things. Like IRQs, DMA channels are limited in number, and you can't allocate one channel to more than one device (unless you

want to grind your system to a halt). Most users come in contact with DMA when they install a sound card that—if they’re lucky—picks the right channel during setup.

DNS

Domain Name Service, is the system used on the Internet for mapping names (called domain names) to the actual numerical addresses of machines on the Internet (IP addresses). Every computer on the Internet has its own number. Since humans can remember names more easily, DNS maps the numbers, such as 906.87.42.119, to names, such as www.5starsupport.com. When you type a Web page address into your browser, your computer consults a DNS server to find the actual numerical address for the machine that goes by that name.

DOCSIS

Data Over Cable Service Interface Specification. The dominating cable modem standard. Defines technical specifications for both cable modem and CMTS.

Dongle

A device that attaches to a computer to control access to a particular application. Dongles provide the most effective means of copy protection. Typically, the dongle attaches to a PC’s parallel port. On Macintoshes, the dongle sometimes attaches to the ADB port. The dongle passes through all data coming through the port so it does not prevent the port from being used for other purposes. In fact, it’s possible to attach several dongles to the same port.

DOS

Stands for Disc Operating System. It is a generic term for the many programs that accept commands to trip applications to run. The most popular is MS-DOS (MS stands for Microsoft).

Downstream

The data flowing from the Cable Modem Termination System to the cable modem.

Downstream frequency

The frequency used for transmitting data from the Cable Modem Termination System to the cable modem. Normally in the 42/65-850 MHz range depending on the actual cable plant capabilities.

Domain

PA group of computers and devices on a network that are administered as a unit with common rules and procedures. Within the Internet, domains are defined by the IP address. All devices sharing a common part of the IP address are said to be in the same domain.

Domain Name

A name that identifies one or more IP addresses. Every domain name has a suffix that indicates which top-level (TLD) domain it belongs to. There are only a limited number of such domains.

For example:

gov - Government agencies
edu - Educational institutions
org - Organizations (nonprofit)
mil - Military
com - commercial business
net - Network organizations
ca - Canada
th - Thailand

Because the Internet is based on IP addresses, not domain names, every Web server requires a Domain Name System (DNS) server to translate domain names into IP addresses.

Dot Pitch

A measurement that indicates the diagonal distance between like-colored phosphor dots on a display screen. Measured in millimeters, the dot pitch is one of the principal characteristics that determines the quality of display monitors. The lower the number, the crisper the image. The dot pitch of color monitors for personal computers ranges from about 0.15 mm to 0.30 mm. Another term for dot pitch is phosphor pitch.

DPI

Stands for Dots Per Inch.

DRAM

DRAM (dynamic random access memory) is most commonly used type of memory in computers. A bank of DRAM memory usually forms the computer's main memory. It is called Dynamic because it needs to be refreshed.

Driver

A program that controls a device. Every device, whether it be a printer, disk drive, or keyboard, must have a driver program. Many drivers, such as the keyboard driver, come with the operating system. For other devices, you may need to load a new driver when you connect the device to your computer. In DOS systems, drivers are files with a .SYS extension. In Windows environments, drivers often have a .DRV extension. A driver acts like a translator between the device and programs that use the device. Each device has its own set of specialized commands that only its driver knows. In contrast, most programs access devices by using generic commands. The driver, therefore, accepts generic commands from a program and then translates them into specialized commands for the device.

DSP: (Digital Signal Processor)

DSP chips are widely used in sound cards, fax machines, modems, cellular phones, high-capacity hard disks and digital TVs. The first DSP chip used in a commercial product was believed to be from Texas Instruments, which was used in its very

popular Speak & Spell game in the late 1970s. DSP chips are used in sound cards for recording and playback and speech synthesis. Other audio uses are amplifiers that simulate concert halls and surround-sound effects for music and home theater.

DTP

Desk Top Publisher(ing) - A PC Term that describes a program that enables you to design, create and print a variety of projects such as letterheads, birthday cards, calendars, business cards, invitations etc. that would have previously only been possible by using the services of an outside printers business.

Dumb Terminal

This a video screen that is seeing manipulation in another computer. Example: If you log in to AOL, your computer is not doing the work—AOL's computer is. You are just being offered a window into that world. That window is your screen. It's a terminal, but it's just watching—thus a dumb terminal.

DVD

Short for digital versatile disc or digital video disc, a new type of CD-ROM that holds a minimum of 4.7GB (gigabytes), enough for a full-length movie. Many experts believe that DVD disks, called DVD-ROMs, will eventually replace CD-ROMs, as well as VHS video cassettes and laser discs. The DVD specification supports disks with capacities of from 4.7GB to 17GB and access rates of 600 KBps to 1.3 MBps. One of the best features of DVD drives is that they are backward-compatible with CD-ROMs. This means that DVD players can play old CD-ROMs, CD-I disks, and video CDs, as well as new DVD-ROMs. Newer DVD players, called second-generation or DVD-2 drives, can also read CD-R and CD-RW disks. DVD uses MPEG-2 to compress video data.

Dynamic URL

A URL that results from the search of a database-driven Web site or the URL of a Web site that runs a script. In contrast to static URLs, in which the contents of the Web page do not change unless the changes are coded into the HTML, dynamic URLs are generated from specific queries to a site's database. The page is merely a template to display the results of the query. Most of the content comes from the database that is associated with the site. Instead of changing information in the HTML code, the data is changed in the database.

EBCDIC

Extended Binary Coded Decimal Interchange Code. It is also called the Extended ASCII Code, as it adds an eighth digit to the normal seven-digit code.

ECC Memory

Error Checking and Correction. A method of detecting and correcting system memory errors by adding additional bits and using a special algorithm.

EDO Memory

Short for Extended Data Output Dynamic Random Access Memory, a type of DRAM that is faster than conventional DRAM. Unlike conventional DRAM which can only access one block of data at a time, EDO RAM can start fetching the next block of memory at the same time that it sends the previous block to the CPU.

EEPROM

Acronym for electrically erasable programmable read-only memory. Pronounced double-ee-prom or e-e-prom, an EEPROM is a special type of PROM that can be erased by exposing it to an electrical charge. Like other types of PROM, EEPROM retains its contents even when the power is turned off. Also like other types of ROM, EEPROM is not as fast as RAM. EEPROM is similar to flash memory (sometimes called flash EEPROM). The principal difference is that EEPROM requires data to be written or erased one byte at a time whereas flash memory allows data to be written or erased in blocks. This makes flash memory faster.

EIDE

(Enhanced Integrated Development Environment), also called EIDE, is a term that Western Digital coined in 1994 to represent a particular set of extensions it devised to the original AT Attachment standard. At that time, the official ATA standard was rather limiting, and work was progressing towards the new ATA-2 standard. Western Digital decided that it did not want to wait for the new standard, and also that it could better position itself as a market leader by creating a new feature set for (then) future drives. The name "Enhanced IDE" was presumably selected to build upon the common name for ATA then in popular use: IDE.

E-Mail

Stands for Electronic Mail. This is a system of relaying messages across the Internet, from one Internet user to another.

Emulation

Refers to the ability of a program or device to imitate another program or device. Many printers, for example, are designed to emulate Hewlett-Packard LaserJet printers because so much software is written for HP printers. By emulating an HP printer, a printer can work with any software written for a real HP printer. Emulation tricks the software into believing that a device is really some other device. Communications software packages often include terminal emulation drivers. This enables your PC to emulate a particular type of terminal so that you can log on to a mainframe. It is also possible for a computer to emulate another type of computer. For example, there are programs that enable an Apple Macintosh to emulate a PC.

Encryption

Encryption is the process of converting data into "unreadable code" so that unauthorized people cannot understand the content. Encryption may be used to make stored data private (e.g., data that is stored on a potentially vulnerable hard disk), or to allow a nonsecure communications channel to serve as a private

communications channel. Encryption is sometimes described as the process of converting plaintext into ciphertext. To decipher the message, the receiver of the encrypted data must have the proper decryption key.

Engine: (as in “Search Engine”)

This is the working part of a database or application.

EPROM

(Acronym for erasable programmable read-only memory, and pronounced ee-prom, EPROM is a special type of memory that retains its contents until it is exposed to ultraviolet light. The ultraviolet light clears its contents, making it possible to reprogram the memory. To write to and erase an EPROM, you need a special device called a PROM programmer or PROM burner. An EPROM differs from a PROM in that a PROM can be written to only once and cannot be erased. EPROMs are used widely in personal computers because they enable the manufacturer to change the contents of the PROM before the computer is actually shipped. This means that bugs can be removed and new versions installed shortly before delivery.

Error Rate

In many cases, it may be acceptable if an input device generates a certain number of errors. This is often referred to as the error rate and the acceptable level will vary according to the input device being used and the business application. Optical character recognition, for example, is generally considered a comparatively unreliable means of entering data. At present, a typical OCR software package will have an error rate of between five and ten per cent.

ESD

Stands for Electro Static Discharge and is defined as a sudden flow of electricity between two objects at different electrical potentials. ESD is a primary cause of integrated circuit damage or failure.

ESD Testing

One kind of test that hardware usually has to pass to prove it is suitable for sale and use. The hardware must still work after it has been subjected to some level of electrostatic discharge. Some organizations have their own ESD requirements which hardware must meet before it will be considered for purchase. Different countries have different legal regulations about levels of ESD.

How to help prevent ESD

The best way to help prevent ESD is either to use a wrist strap, or a grounding mat. However most users do not have access to such items therefore you can follow the below guidelines to help prevent ESD as much as possible.

- Clothes - Insure what you are not wearing an item that conducts a lot of Electrical Charge, such as a wool sweater. Also it is generally a good idea to remove all jewelry as well.
- Weather - When working on your computer insure there is not an electrical storm outside which increases the potential of ESD.

- Cords - Insure everything is removed from the back of the computer (power cord, mouse, keyboard, Zero Potential - Insure you and the computer are at Zero Potential by continuously touching the un-painted metal chassis (computer frame) or the Power supply.
- Standing - When working inside the computer it is highly recommended that you stand at ALL times.
- Surface - It is always best to stand on a wooden surface. Avoid working on a computer in carpeted areas.

Ethernet

A networking system that enables high speed data communication over coaxial cables. The Ethernet network system supports TCP/IP, AppleTalk, Novell Netware, and other network protocols. An Ethernet (LAN) connection is 10 Mbit/s or 100 Mbit/s, and is used to connect many computers that can all "talk" directly to each other. Normally they will all talk with a few servers and printers, but the network is all-to-all. The distance is normally limited to below 1 km.

Executable File

A file in a format that the computer can directly execute. Unlike source files, executable files cannot be read by humans. To transform a source file into an executable file, you need to pass it through a compiler or assembler. In DOS systems, executable files have either a .COM or .EXE extension and are called COM files and EXE files, respectively.

Expanded Memory

Also known as EMS (Expanded Memory Specification), expanded memory is a technique for utilizing more than 1MB (megabyte) of main memory in DOS-based computers. The limit of 1MB is built into the DOS operating system. The upper 384K is reserved for special purposes, leaving just 640K of conventional memory for programs.

Expansion Bus

A group of control lines that provide a buffered interface to devices located either on the system board or on cards that are plugged into expansion connectors. Common expansion buses included on the system board are USB, PC Card, and PCI.

Extended Memory

Memory above and beyond the standard 1MB (megabyte) of main memory that DOS supports. Extended memory is only available in PCs with an Intel 80286 or later microprocessor. Two types of memory can be added to a PC to increase memory beyond 1MB: expanded memory and extended memory. Expanded memory conforms to a published standard called EMS that enables DOS programs to take advantage of it. Extended memory, on the other hand, is not configured in any special manner and is therefore unavailable to most DOS programs. However, MS-Windows and OS/2 can use extended memory.

External Modem

A modem that resides in a self-contained box outside the computer system. Contrast with an internal modem, which resides on a printed circuit board inserted into the computer. External modems tend to be slightly more expensive than internal modems. Many experts consider them superior because they contain lights that indicate how the modem is functioning. In addition, they can easily be moved from one computer to another. However, they do use up one COM port.

Extranet

An extranet is a private network that uses the Internet protocols and the public telecommunication system to share a business's information, data or operations with external suppliers, vendors or customers. An extranet can be viewed as the external part of a company's Intranet. See also: Intranet.

FAT

Stands for File Allocation Table. Basically this is a table of contents in a directory that tells the computer what all is in there. Look at your Netscape cache, you'll see a FAT. It'll be the first file.

FAT 32

A new version of the file allocation table (FAT) available in Windows 95 OSR 2 and Windows 98. FAT32 increases the number of bits used to address clusters and also reduces the size of each cluster. The result is that it can support larger disks (up to 2 terabytes) and better storage efficiency (less slack space).

Fault Tolerance

The ability of a system to respond gracefully to an unexpected hardware or software failure. There are many levels of fault tolerance, the lowest being the ability to continue operation in the event of a power failure. Many fault-tolerant computer systems mirror all operations—that is, every operation is performed on two or more duplicate systems, so if one fails the other can take over.

FAQ

Stands for Frequently Asked Questions. An FAQ is a file or document where a moderator or administrator will post commonly asked questions and their answers.

Fax Modem

A device you can attach to a personal computer that enables you to transmit and receive electronic documents as faxes. A fax modem is like a regular modem except that it is designed to transmit documents to a fax machine or to another fax modem. Some, but not all, fax modems do double duty as regular modems. As with regular modems, fax modems can be either internal or external. Internal fax modems are often called fax boards.

FDDI

Fiber Distributed Data Interface—A standard for transmitting data on optical fiber cables at a rate of around 100,000,000 bits-per-second (10 times as fast as Ethernet, about twice as fast as T-3).

FDISK

A program found in all Aptiva software loads that allows modification of the partitions and/or logical drives on the hard drive. It can Display, Delete and Create partitions and logical drives, defining them for DOS, OS/2 or Windows, depending on which version of FDISK is used and how it is used. Type FDISK and hit Enter to start the program. This is a DESTRUCTIVE command and incorrect use will result in data loss!

Fiber Optic

An alternative to copper wire for transmitting information. In fiber optics, pulses of light representing binary data are flashed along a flexible glass fiber. The advantage over copper wiring is that a single strand of optical fiber can carry thousands and thousands of different frequencies at once without data loss.

File Sharing

This is the most important feature of the Internet. This is a method of allowing one server to give the same file to many different end users.

Firmware

Software (programs or data) that has been written onto read-only memory (ROM). Firmware is a combination of software and hardware. ROMs, PROMs and EPROMs that have data or programs recorded on them are firmware.

FORTRAN

FORmula TRANslator. Developed in 1954 by IBM, it is a high-level programming language, most widely used for scientific and engineering applications because it has excellent mathematical functions. Many programmers consider it to sacrifice "elegance" for speed of numerical manipulations.

Freeware

This is a shortened version of Free Software. Programmers offer their work without wanting pay in return.

FSB

Stands for Front Side Bus, which denotes the speed at which your processor interacts with the components on the motherboard. Typically the FSB is 100 Mhz or 133 Mhz, but overclockers often manipulate this value to increase the speed at which their processor runs. *i.e.*, 100 Mhz FSB X 5.0 clock multiplier = 500 Mhz processor.

FTP

Stands for File Transfer Protocol.

Full Duplex

Refers to the transmission of data in two directions simultaneously. For example, a telephone is a full-duplex device because both parties can talk at once. In contrast, a walkie-talkie is a half-duplex device because only one party can transmit at a time. Most modems have a switch that lets you choose between full-duplex and

half-duplex modes. The choice depends on which communications program you are running. In full-duplex mode, data you transmit does not appear on your screen until it has been received and sent back by the other party. This enables you to validate that the data has been accurately transmitted. If your display screen shows two of each character, it probably means that your modem is set to half-duplex mode when it should be in full-duplex mode.

Gateway

As in Common Gateway Interface (CGI). It is a piece of software that allows two items to communicate with each other. They are used to make connections between computers and systems inside that computer.

General Protection Fault

GPF, short for General Protection Fault, is a computer condition that causes a Windows application to crash. The most common cause of a GPF is two applications trying to use the same block of memory, or more specifically, one application trying to use memory assigned to another application.

The following situations can also cause GPFs:

- Running an application with insufficient resources
- Using improper hardware device drivers
- Corrupted or missing Windows files
- Applications exchanging data that cannot be read
- GPFs are often preceded by an invalid page fault.

GIF

Pronounced "jif." Stands for Graphical Interchange Format. It is an image format created by Compuserve.

Gigabyte

2^{30} bytes (1,073,741,824) bytes. One gigabyte is equal to 1,024 megabytes. Gigabyte is often abbreviated as G or GB.

GIGO

It's an acronym that stands for Garbage In, Garbage Out.

Glyph

A graphic symbol whose appearance conveys information; for example, the vertical and horizontal arrows on cursor keys that indicate the directions in which they control cursor movement.

Gopher

A method of distributing information by computers that has waned in popularity to ftp. Most gopher files contain only text information with few images, audio, or video components. Files can be downloaded with a similar protocol like ftp.

GUI - Graphical User Interface

A program interface that takes advantage of the computer's graphics capabilities to make the program easier to use. Well-designed graphical user interfaces can free the user from learning complex command languages. On the other hand, many users find that they work more effectively with a command-driven interface, especially if they already know the command language.

Half-Duplex

Refers to the transmission of data in just one direction at a time. For example, a walkie-talkie is a half-duplex device because only one party can talk at a time. In contrast, a telephone is a full-duplex device because both parties can talk simultaneously.

Most modems contain a switch that lets you select between half-duplex and full-duplex modes. The correct choice depends on which program you are using to transmit data through the modem.

In half-duplex mode, each character transmitted is immediately displayed on your screen. (For this reason, it is sometimes called local echo — characters are echoed by the local device). In full-duplex mode, transmitted data is not displayed on your monitor until it has been received and returned (remotely echoed) by the other device. If you are running a communications program and every character appears twice, it probably means that your modem is in half-duplex mode when it should be in full-duplex mode, and every character is being both locally and remotely echoed.

Handshaking

The process by which two devices initiate communications. Handshaking begins when one device sends a message to another device indicating that it wants to establish a communications channel. The two devices then send several messages back and forth that enable them to agree on a communications protocol.

Hard Boot

A hard reboot (also known as a cold reboot) is when power to a computer is cycled (turned on and off) or a special reset signal to the processor is triggered (from a front panel switch of some sort). This restarts the computer without first performing the usual shut-down procedure. (With many operating systems, especially those with disc caches, after a hard reboot the system may well be in an "unclean" state, and require that checks and repairs to on-disc filesystem structures be performed before normal operation can begin.) It may be caused by power failure, be done by accident, or be done deliberately as a last resort because nothing else to retrieve the system from a "hung" state works.

Hardware

These are the physical items including your computer and floppy discs.

Hayes Compatible

Hayes Microcomputer Products is one of the leading manufacturers of modems and has developed a language called the AT command set for controlling modems that

has become the de facto standard. Any modem that recognizes Hayes modem commands is said to be Hayes-compatible.

This is very useful because most communications programs use Hayes modem commands. Virtually all modems manufactured today are Hayes-compatible.

Headend

Central distribution point for a CATV system. Video signals are received here from satellites and may be other sources, frequency converted to the appropriate channels, combined with locally originated signals, and rebroadcast onto the HFC plant. The headend is where the CMTS is normally located.

Heat Sink

A component designed to lower the temperature of an electronic device by dissipating heat into the surrounding air. All modern CPUs require a heat sink. Some also require a fan. A heat sink without a fan is called a passive heat sink; a heat sink with a fan is called an active heat sink. Heat sinks are generally made of a zinc alloy and often have fins.

Helper Application

This is an application your browser uses to manipulate a downloaded program.

HFC

Hybrid fiber-coaxial (cable network). Older CATV systems were provisioned using only coaxial cable. Modern systems use fiber transport from the headend to an optical node located in the neighborhood to reduce system noise. Coaxial cable runs from the node to the subscriber. The fiber plant is generally a star configuration with all optical node fibers terminating at a headend. The coaxial cable part of the system is generally a trunk-and-branch configuration.

High Memory Area

In DOS-based systems, the high memory area refers to the first 64K of extended memory.

HST

High Speed Technology- Before the invention of the CCITT V.32 modem standards for 9600 BPS modems, US Robotics invented a proprietary protocol that runs even faster at 14,400 BPS. It became popular on US bulletin board system, but never caught on outside the USA. It is gradually being replaced by V.32.

Host

A computer on a network that provides services to other computers on the network. Unless you have your own server, you need a hosting company who provides a server or computer that is connected to the internet and makes your web pages available to the rest of the internet.

Hot Fix

Novell, Inc.'s term for the feature of their network file server operating system, Novell NetWare, which handles errors in disk write operations. The OS re-reads every block it writes to disk while it holds the data to be written in memory. In the case of an error, the data block is written to a spare area on the disk. The feature lost much of its importance with the widespread use of hard disk drives with built-in error correction and bad block re-mapping.

Hotlist

List of URLs saved within the Mosaic Web browser. (Bookmark in Netscape.)

HTML

Hyper Text Markup Language. It is a collection of structuring and formatting tags used to create Web pages.

HTTP

Stands for Hyper-Text Transport Protocol. Common protocol used to communicate between World Wide Web Servers.

Hub

A common connection point for devices in a network. Hubs are commonly used to connect segments of a LAN. A hub contains multiple ports. When a packet arrives at one port, it is copied to the other ports so that all segments of the LAN can see all packets.

A passive hub serves simply as a conduit for the data, enabling it to go from one device (or segment) to another. So-called intelligent hubs include additional features that enable an administrator to monitor the traffic passing through the hub and to configure each port in the hub. Intelligent hubs are also called manageable hubs.

A third type of hub, called a switching hub, actually reads the destination address of each packet and then forwards the packet to the correct port.

Hybrid

A device or system combining two types of mechanisms, circuits, or design approaches, each of which could of itself accomplish the total function but in a different and usually less effective manner. A hybrid computer combines digital and analog computers into one functioning system.

Hypertext

This is a mark-up language that allows for non-linear transfers of data. The method allows your computer to provide the computational power rather than attaching to a mainframe and waiting for it to do the work for you.

Hyper-Threading

(HTT = Hyper Threading Technology) is Intel's trademark for their implementation of the simultaneous multithreading technology on the Pentium 4 microarchitecture. It is basically a more advanced form of Super-threading that first debuted on the Intel Xeon processors and later added to Pentium 4 processors.

IBM

Stands for International Business Machines.

ICMP

(Internet Control Message Protocol) A message control and error-reporting protocol between a host server and a gateway to the Internet. ICMP uses Internet Protocol (IP) datagrams, but the messages are processed by the IP software and are not directly apparent to the application user.

Icon

A small video display that acts as an activation link when clicked on.

IDE

(Integrated Development Environment) A programming environment integrated into an application. For example, Microsoft Office applications support various versions of the BASIC programming language. You can develop a WordBasic application while running Microsoft Word.

IMAP

PA group of computers and devices on a network that are administered as a unit with common rules and procedures. Within the Internet, domains are defined by the IP address. All devices sharing a common part of the IP address are said to be in the same domain.

Italics

A type style with slightly slanted characters, used for emphasis. Best used to set off quotes, special phrases, and foreign words, italic letters have a redesigned structure that allows them to slant to the right. The first italic type was designed by Aldus Manutius in AD 1501 and was based on the handwriting style of that time. Furthermore, lowercase letters were in italics while capital letters were Roman (or vertical stance).

Integrated Circuit

Another name for a chip, an IC is a small electronic device made out of a semiconductor material.

Interface

This is any type of point where two different things come together. Most often, the term is used to describe the programs between you and your computer like Windows, OS/2 and others. What you see on the screen is the interface between you and what your computer is doing.

Interlacing

A display technique that enables a monitor to provide more resolution inexpensively. With interlacing monitors, the electron guns draw only half the horizontal lines with each pass (for example, all odd lines on one pass and all even lines on the next pass). Because an interlacing monitor refreshes only half the lines at one time, it can display

twice as many lines per refresh cycle, giving it greater resolution. Another way of looking at it is that interlacing provides the same resolution as non-interlacing, but less expensively.

Interleaving

A recording method that reduces data errors during playback. Instead of the file being written in a contiguous data stream, the data sectors are intermixed along the recording track. If a disc should have a smudge or scratch, the entire data file is generally recoverable because a smaller amount of the file data is affected.

Internal Modem

A modem that resides on an expansion board that plugs into a computer. In contrast, an external modem is a box that attaches to a computer's COM port via cables.

Internet

The Internet is a super-network. It connects many smaller networks together and allows all the computers to exchange information with each other. To accomplish this all the computers on the Internet have to use a common set of rules for communication. Those rules are called protocols, and the Internet uses a set of protocols called TCP/IP (Transmission Control Protocol/Internet Protocol). Many people equate the World Wide Web with the Internet. In fact, the Internet is like the highway, and the World Wide Web is like a truck that uses that highway to get from place to place.

Interrupt

A signal informing a program that an event has occurred. When a program receives an interrupt signal, it takes a specified action (which can be to ignore the signal). Interrupt signals can cause a program to suspend itself temporarily to service the interrupt.

Interrupt signals can come from a variety of sources. For example, every keystroke generates an interrupt signal. Interrupts can also be generated by other devices, such as a printer, to indicate that some event has occurred. These are called hardware interrupts. Interrupt signals initiated by programs are called software interrupts. A software interrupt is also called a trap or an exception.

PCs support 256 types of software interrupts and 15 hardware interrupts. Each type of software interrupt is associated with an interrupt handler — a routine key on your keyboard, this triggers a specific interrupt handler. The complete list of interrupts and associated interrupt handlers is stored in a table called the interrupt vector table, which resides in the first 1 K of addressable memory.

Intranet

A private network for communications and sharing of information that, like the Internet, is based on TCP/IP but is accessible only to authorized users within an organization. An organization's intranet is usually protected from external access by a firewall. See also: Extranet.

IPX

Short for Internetwork Packet Exchange, a networking protocol used by the Novell NetWare operating systems. Like UDP/IP, IPX is a datagram protocol used for connectionless communications. Higher-level protocols, such as SPX and NCP, are used for additional error recovery services. The successor to IPX is the NetWare Link Services Protocol (NLSP).

IS

Stands for Information System.

ISA

The bus architecture used in the IBM PC/XT and PC/AT. It's often abbreviated as ISA (pronounced as separate letters or as eye-sa) bus. The AT version of the bus is called the AT bus and became a *de facto* industry standard. Starting in the early 90s, ISA began to be replaced by the PCI local bus architecture. Most computers made today include both an AT bus for slower devices and a PCI bus for devices that need better bus performance. In 1993, Intel and Microsoft introduced a new version of the ISA specification called Plug and Play ISA. Plug and Play ISA enables the operating system to configure expansion boards automatically so that users do not need to fiddle with DIP switches and jumpers.

ISDN

Integrated Services Digital Network. Basically a way to move more data over regular existing phone lines. ISDN is available to much of the USA and in most markets it is priced very comparably to standard analog phone circuits. It can provide speeds of roughly 128,000 bits-per-second over regular phone lines. In practice, most people will be limited to 56,000 or 64,000 bits-per-second. Unlike DSL, ISDN can be used to connect to many different locations, one at a time, just like a regular telephone call, as long the other location also has ISDN.

ISO

Stands for the International Standards Organization. Someone has to say what is the standard for transferring data. These people are it. You'll find them in Paris.

ISP

Internet Service Provider, a company that provides access to the Internet. For a monthly fee, the service provider gives you a software package, username, password and access phone number. Equipped with a modem, you can then log on to the Internet and browse the World Wide Web and USENET, and send and receive e-mail.

In addition to serving individuals, ISPs also serve large companies, providing a direct connection from the company's networks to the Internet. ISPs themselves are connected to one another through Network Access Points (NAPs). ISPs are also called IAPs (Internet Access Providers).

JAVA

A high-level programming language developed by Sun Microsystems. Java was originally called OAK, and was designed for handheld devices and set-top boxes. Oak was unsuccessful so in 1995 Sun changed the name to Java and modified the language to take advantage of the burgeoning World Wide Web.

Java is an object-oriented language similar to C++, but simplified to eliminate language features that cause common programming errors. Java source code files (files with a .java extension) are compiled into a format called bytecode (files with a .class extension), which can then be executed by a Java interpreter. Compiled Java code can run on most computers because Java interpreters and runtime environments, known as Java Virtual Machines (VMs), exist for most operating systems, including UNIX, the Macintosh OS, and Windows. Bytecode can also be converted directly into machine language instructions by a just-in-time compiler (JIT).

Java is a general purpose programming language with a number of features that make the language well suited for use on the World Wide Web. Small Java applications are called Java applets and can be downloaded from a Web server and run on your computer by a Java-compatible Web browser, such as Netscape Navigator or Microsoft Internet Explorer.

JavaScript

This is a language very close to Java that allows for more interaction with the viewer. It is much more forgiving than Java as doesn't require its own window in which to work.

JPEG

Pronounced "J-Peg." Stands for Joint Photographic Experts Group. It's an image format that allows for compression when stored.

Jumpers

A metal bridge that closes an electrical circuit. Typically, a jumper consists of a plastic plug that fits over a pair of protruding pins. Jumpers are sometimes used to configure expansion boards. By placing a jumper plug over a different set of pins, you can change a board's parameters.

K56Flex

(A technology developed by Lucent Technologies and Rockwell International for delivering data rates up to 56 Kbps over plain old telephone service (POTS). It was long believed that the maximum data transmission rate over copper telephone wires was 33.6 Kbps, but K56flex achieves higher rates by taking advantage of the fact that most phone switching stations are connected by high-speed digital lines. K56flex bypasses the normal digital-to-analog conversion and sends the digital data over the telephone wires directly to your modem where it is decoded.

Lucent and Rockwell have announced that future K56flex modems will conform to the new V.90 standard approved by the ITU. And users with older K56flex modems may upgrade their modems to support V.90.

While K56flex offers faster Internet access than normal modems, there are several caveats to using an K56flex modem:

- The high speeds are available only with downstream traffic (e.g., data sent to your computer). Upstream traffic is delivered using normal techniques, with a maximum speed of 33.6 Kbps.
- To connect to the Internet at K56flex speeds, your Internet Service Provider (ISP) must have a modem at the other end that supports V.90.
- Even if your ISP supports V.90, you might not achieve maximum transmission rates due to noisy lines.

Kbit/s

Stands for thousands of bits per second.

Kernel

The central module of an operating system. It is the part of the operating system that loads first, and it remains in main memory. Because it stays in memory, it is important for the kernel to be as small as possible while still providing all the essential services required by other parts of the operating system and applications. Typically, the kernel is responsible for memory management, process and task management, and disk management.

Kilobyte: (KB)

This is about a thousand bytes of space. In reality, it's two to the 10th power or 1,024 bytes.

KVM

Keyboard-Video-Mouse switch. A piece of hardware that connects two or more computers to a single keyboard, monitor and mouse. Imagine you have a row of 4 computers that all serve as file servers. Why waste money buying 4 monitors, 4 keyboards and 4 mice. With a KVM switch you can connect all 4 computers to one monitor, keyboard and mouse and to switch between them when needed.

LAN

A computer network that spans a relatively small area. Most LANs are confined to a single building or group of buildings. However, one LAN can be connected to other LANs over any distance via telephone lines and radio waves. A system of LANs connected in this way is called a wide area network (WAN).

Most LANs connect workstations and personal computers. Each node (individual computer) in a LAN has its own CPU with which it executes programs, but it is also able to access data and devices anywhere on the LAN. This means that many users can share expensive devices, such as laser printers, as well as data. Users can also use the LAN to communicate with each other, by sending e-mail or engaging in chat sessions.

There are many different types of LANs Ethernets being the most common for PCs. Most Apple Macintosh networks are based on Apple's AppleTalk network system, which is built into Macintosh computers.

LCD

Abbreviation of liquid crystal display, a type of display used in digital watches and many portable computers. LCD displays utilize two sheets of polarizing material with a liquid crystal solution between them. An electric current passed through the liquid causes the crystals to align so that light cannot pass through them. Each crystal, therefore, is like a shutter, either allowing light to pass through or blocking the light.

Monochrome LCD images usually appear as blue or dark gray images on top of a grayish-white background. Color LCD displays use two basic techniques for producing color: Passive matrix is the less expensive of the two technologies. The other technology, called thin film transistor (TFT) or active-matrix, produces color images that are as sharp as traditional CRT displays, but the technology is expensive. Recent passive-matrix displays using new CSTN and DSTN technologies produce sharp colors rivaling active-matrix displays.

LED: Abbreviation of light emitting diode, an electronic device that lights up when electricity is passed through it. LEDs are usually red. They are good for displaying images because they can be relatively small, and they do not burn out. However, they require more power than LCDs.

Linux

A version of UNIX that runs on a variety of hardware platforms including x86 PCs, Alpha, PowerPC and IBM's product line. Linux is open source software, which is freely available; however, the full distribution of Linux along with technical support and training are available for a fee from vendors such as Red Hat Software and Caldera. Due to its stability, Linux has gained popularity with Internet Service Providers as the Operating System of choice for hosting Web servers.

Live Script

This is the former name of Java Script. There are few updates between the two.

Login

To attach to a computer. It has also come to represent your User ID command.

Login Script

This is the small text file that is run by the server gateway to make the attachment between it and your computer.

Loopback

A diagnostic test that returns the transmitted signal back to the sending device after it has passed through a network or across a particular link. The returned signal can then be compared to the transmitted one. The discrepancy between the two help to trace the fault. When trying to locate a faulty piece of equipment, loopbacks will be repeated, eliminating satisfactory machines until the problem is found.

LVD

Low Voltage Differential. A differential logic scheme using lower voltage levels than HVD.

MAC layer

Media Access Control sub layer in the network stack.

Mainframe

Mostly a mainframe is only a mainframe when compared to a desktop computer. It's bigger and much more powerful. Sometimes it's called a server or CPU.

MBR

Short for Master Boot Record, a small program that is executed when a computer boots up. Typically, the MBR resides on the first sector of the hard disk. The program begins the boot process by looking up the partition table to determine which partition to use for booting. It then transfers program control to the boot sector of that partition, which continues the boot process. In DOS and Windows systems, you can create the MBR with the FDISK/MBR command.

An MBR virus is a common type of virus that replaces the MBR with its own code. Since the MBR executes every time a computer is started, this type of virus is extremely dangerous. MBR viruses normally enter a system through a floppy disk that is installed in the floppy drive when the computer is started up. Even if the floppy disk is not bootable, it can infect the MBR.

MCNS

Multimedia Cable Network System Partners Ltd. The consortium behind the DOCSIS standard for cable modems.

Media

1. Objects on which data can be stored. These include hard disks, floppy disks, CD-ROMs, and tapes.
2. In computer networks, media refers to the cables linking workstations together. There are many different types of transmission media, the most popular being twisted-pair wire (normal electrical wire), coaxial cable (the type of cable used for cable television), and fiber optic cable (cables made out of glass).
3. The form and technology used to communicate information. Multimedia presentations, for example, combine sound, pictures, and videos, all of which are different types of media.

Megabyte: (MB)

About a million bytes of space. Actually it's 2 raised to the 20th power or 1,048,576 bytes of space.

Memory

Internal storage areas in the computer. The term memory identifies data storage that comes in the form of chips, and the word storage is used for memory that exists on tapes or disks. Moreover, the term memory is usually used as a shorthand for physical memory, which refers to the actual chips capable of holding data. Some computers also use virtual memory, which expands physical memory onto a hard disk. Every computer comes with a certain amount of physical memory, usually referred to as

main memory or RAM. You can think of main memory as an array of boxes, each of which can hold a single byte of information. A computer that has 1 megabyte of memory, therefore, can hold about 1 million bytes (or characters) of information.

Microcomputer

A category of computer that is generally used for personal computing, for small business computing, and as a workstation attached to large computers or to other small computers on a network.

Microprocessor

A silicon chip that contains a CPU. In the world of personal computers, the terms microprocessor and CPU are used interchangeably. At the heart of all personal computers and most workstations sits a microprocessor. Microprocessors also control the logic of almost all digital devices, from clock radios to fuel-injection systems for automobiles.

Three basic characteristics differentiate microprocessors:

- Instruction set: The set of instructions that the microprocessor can execute.
- Bandwidth: The number of bits processed in a single instruction.
- Clock Speed: Given in megahertz (MHz), the clock speed determines how many instructions per second the processor can execute.
- In both cases, the higher the value, the more powerful the CPU. For example, a 32-bit microprocessor that runs at 50 MHz is more powerful than a 16-bit microprocessor that runs at 25 MHz.

In addition to bandwidth and clock speed, microprocessors are classified as being either RISC (reduced instruction set computer) or CISC (complex instruction set computer).

MIDI

Stands for Music Instrument Digital Interface. It allows a computer to store and replay a musical instrument's output.

Minislot

(Multi-Media Extensions) A set of 57 multimedia instructions built into Intel's newest microprocessors and other x86-compatible microprocessors. MMX-enabled microprocessors can handle many common multimedia operations, such as digital signal processing (DSP), that are normally handled by a separate sound or video card. However, only software especially written to call MMX instructions — so-called MMX-enabled software — can take advantage of the MMX instruction set. The first generation of computers with MMX chips hit the market in January, 1997.

Modem

This is a word created out of the beginning letters of two other words: MODulation and DEModulation. The words mean the changing of data from digital (computer language) to analog (phone line language) and then back again. It represents the purpose of your computer's modem.

Mosaic

The first Web browser to have a consistent interface for the Macintosh, Windows, and Unix environments. It was created at the National Center for Supercomputing Applications (NCSA). The success of this browser is really responsible for the expansion of the Web.

Motherboard

The main circuit board of a microcomputer. The motherboard contains the connectors for attaching additional boards. Typically, the motherboard contains the CPU, BIOS, memory, mass storage interfaces, serial and parallel ports, expansion slots, and all the controllers required to control standard peripheral devices, such as the display screen, keyboard, and disk drive. Collectively, all these chips that reside on the motherboard are known as the motherboard's chipset.

On most PCs, it is possible to add memory chips directly to the motherboard. You may also be able to upgrade to a faster CP by replacing the CPU chip. To add additional core features, you may need to replace the motherboard entirely.

MPEG

Stands for Motion Picture Experts Group. A format to make, view, and transfer both digital audio and digital video files.

MSO

Multiple Service Operator. A cable TV service provider that also provides other services such as data and/or voice telephony.

Multiplexer

This is a piece of hardware that allows one item to take the place of several. An example would be using a multiplexer to allow 10 computers to attach where only one could before.

NACS

A Stands for Netware Asynchronous Communication Services.

Nanosecond

A billionth of a second. Many computer operations, such as the speed of memory chips, are measured in nanoseconds. Nanosecond is often abbreviated as ns.

Netbeui

Netbeui is short for NetBios Enhanced User Interface. It is an enhanced version of the NetBIOS protocol used by network operating systems such as LAN Manager, LAN Server, Windows for Workgroups, Windows 95 and Windows NT.

Netbeui was originally designed by IBM for their Lan Manager server and later extended by Microsoft and Novell.

Network

This a system that sends and receives data.

Network Card

Also, Network Interface Card or NIC. This is a component of a computer that enables the computer to communicate with other computers via a direct network connection.

Network Adapter

This is a hardware unit that connects a device to a communication line. For wide area networks (WAN), these adapters connect routers to the specific type of connection (T1, BRI) that is installed. For local area networks (LAN), these adapters connect workstations to the LAN (Ethernet or TokenRing) cabling.

NLX

New Low Profile Extended (motherboard form factor). The NLX form factor features a number of improvements over the previous design LPX form factor and began heavy usage in late 1997. The popularity of the design was confirmed by massive design use in 1998. The popularity has made it Intel's flagship line and one of the profit leaders in chipsets. Its features include:

1. Support for larger memory modules and DIMMs.
2. Support for the newest microprocessors, including the Pentium II using SEC packaging.
3. Support for AGP video cards.
4. Better access to motherboard components.
5. Support for dockable designs in which the motherboard can be removed without tools.

Node

In networks, a processing location. A node can be a computer or some other device, such as a printer. Every node has a unique network address, sometimes called a Data Link Control(DLC) address or Media Access Control(MAC) address.

Noise

Interference (static) that destroys the integrity of signals on a line. Noise can come from a variety of sources, including radio waves, nearby electrical wires, lightning, and bad connections. One of the major advantages of fiber optic cables over metal cables is that they are much less susceptible to noise.

NTFS

Short for NT File System, one of the file system for the Windows NT operating system (Windows NT also supports the FAT file system). NTFS has features to improve reliability, such as transaction logs to help recover from disk failures. To control access to files, you can set permissions for directories and/or individual files. NTFS files are not accessible from other operating such as DOS.

For large applications, NTFS supports spanning volumes, which means files and directories can be spread out across several physical disks.

NTLDR

Short for NT Loader, a program loaded from the hard drive boot sector that displays the Microsoft Windows NT startup menu and helps Windows NT load.

NVRAM

(Non-Volatile Random Access Memory) Any type of memory that is made non-volatile by connecting it to a constant power source, such as a battery. Therefore, non-volatile memory does not lose its contents when the main power is turned off.

Object

Something that contains both the data and the application that operates on that data.

OEM

(Original Equipment Manufacturer) This is a designation for companies that manufacture equipment that is then marketed and sold off to other companies under their own names.

OOP

Stands for Object Oriented Program. A larger program made up of smaller objects.

Opacity

The quality that defines how much light passes through an object's pixels. If an object is 100 percent opaque, no light passes through it.

Operating System

The most important program that runs on a computer. Every general-purpose computer must have an operating system to run other programs. Operating systems perform basic tasks, such as recognizing input from the keyboard, sending output to the display screen, keeping track of files and directories on the disk, and controlling peripheral devices such as disk drives and printers.

For large systems, the operating system has even greater responsibilities and powers. It is like a traffic cop — it makes sure that different programs and users running at the same time do not interfere with each other. The operating system is also responsible for security, ensuring that unauthorized users do not access the system.

Operating systems provide a software platform on top of which other programs, called application programs, can run. The application programs must be written to run on top of a particular operating system. Your choice of operating system, therefore, determines to a great extent the applications you can run. For PCs, the most popular operating systems are DOS, OS/2, and Windows, but others are available, such as Linux.

As a user, you normally interact with the operating system through a set of commands. For example, the DOS operating system contains commands such as COPY and RENAME for copying files and changing the names of files, respectively. The commands are accepted and executed by a part of the operating system called the command processor or command line interpreter. Graphical user interfaces allow you to enter commands by pointing and clicking at objects that appear on the screen.

Overclock

To run a microprocessor faster than the speed for which it has been tested and approved. Overclocking is a popular technique for eking out a little more performance from a system. In many cases, you can force your CPU to run faster than it was intended simply by setting a jumper on the motherboard. Overclocking does come with some risks, however, such as over-heating, so you should become familiar with all the pros and cons before you attempt it.

Overclocking is sometimes called speed margining.

PPP

Stands for Point To Point Protocol. It's a software application that allows an attachment to a server.

Parallel Port

A parallel interface for connecting an external device such as a printer. Most personal computers have both a parallel port and at least one serial port. On PCs, the parallel port uses a 25-pin connector (type DB-25) and is used to connect printers, computers and other devices that need relatively high bandwidth. It is often called a Centronics interface after the company that designed the original standard for parallel communication between a computer and printer. (The modern parallel interface is based on a design by Epson.)

A newer type of parallel port, which supports the same connectors as the Centronics interface, is the EPP (Enhanced Parallel Port) or ECP (Extended Capabilities Port). Both of these parallel ports support bi-directional communication and transfer rates ten times as fast as the Centronics port. Macintoshes have a SCSI port, which is parallel, but more flexible.

Partition

A portion of a hard disk that functions as a separate unit. A single hard disk can be divided into several partitions, each of which functions as a separate drive and has its own volume name (such as D:, E:, F:, and so on). The purpose is to make the drive more efficient, as the computer can search smaller sections for a specific file rather than the entire drive. The verb to partition refers to the process of dividing the hard drive into partitions.

Path

A path can be described as a file's address on your file system, describing where the file lives: An absolute path gives the complete path, starting at the root directory, or the very top of the file system; A relative path looks for a file from the directory you are currently in down.

PCI

Acronym for Peripheral Component Interconnect, a local bus standard developed by Intel Corporation. Most modern PCs include a PCI bus in addition to a more general ISA expansion bus. Many analysts, however, believe that PCI will eventually supplant ISA entirely. PCI is also used on newer versions of the Macintosh computer.

PCI is a 64-bit bus, though it is usually implemented as a 32-bit bus. It can run at clock speeds of 33 or 66 MHz. At 32-bits and 33 MHz, it yields a throughput rate of 133 Mbps.

Although it was developed by Intel, PCI is not tied to any particular family of microprocessors.

Peer to Peer

A type of network in which each workstation has equivalent capabilities and responsibilities. This differs from client/server architectures, in which some computers are dedicated to serving the others. Peer-to-peer networks are generally simpler and less expensive, but they usually do not offer the same performance under heavy loads.

Pen drive

A small keyring-sized device that can be used to easily transfer files between USB-compatible systems. Available in a range of capacities (and in some cases, with an MP3 player built-in). Plug it in to any USB port and it will be automatically detected by the Operating System.

Peripheral

Any external device attached to a computer. Examples of peripherals include printers, disk drives, display monitors, keyboards, and mice.

PGA

Short for Pin Grid Array, a type of chip package in which the connecting pins are located on the bottom in concentric squares. PGA chips are particularly good for chips that have many pins, such as modern microprocessors. Compare with DIP and SIP.

Short for Professional Graphics Adapter, a video standard developed by IBM that supports 640x480 resolution.

Phishing

Short for Password Harvesting Fishing. It is the luring of sensitive information, such as passwords and other personal information, from a victim by masquerading as someone trustworthy with a real need for such information.

Popular targets are users of online banking services, and auction sites such as eBay. Phishers usually work by sending out spam e-mail to large numbers of potential victims. Typically the email will appear to come from a trustworthy company and contain a subject and message intended to alarm the recipient into taking action.

A common approach is to tell the recipient that their account has been de-activated due to a problem and inform them that they must take action to re-activate their account. The user is provided with a convenient link in the same email that takes the e-mail recipient to a fake web page appearing to be that of a trustworthy company. Once at that page, the user enters her personal information which is then captured by the fraudster.

PHP

(Hypertext Preprocessor) is a server-side, HTML-embedded scripting language used to create dynamic Web pages. In an HTML document, PHP script (similar syntax to that of Perl or C) is enclosed within special PHP tags. Because PHP is embedded within tags, the author can jump between HTML and PHP (similar to ASP and Cold Fusion) instead of having to rely on heavy amounts of code to output HTML. Because PHP is executed on the server, the client cannot view the PHP code. PHP can perform any task any CGI program can, but its strength lies in its compatibility with many types of databases. Also, PHP can talk across networks using IMAP, SNMP, NNTP, POP3 or HTTP.

PICT

Pronounced "Pic,t." It is another image format.

PING

Packet Internet or Inter-Network Groper; a utility used to determine whether a particular computer is currently connected to the Internet. It works by sending a packet to the specified IP address and waiting for a reply. The computer acronym "PING" was contrived to match the submariners' term for the sound of a returned sonar pulse.

Pinout

A diagram or table that describes the purpose of each pin in a chip or connector, or each wire in a cable.

PIO

(Programmed Input/Output) A method of data transfer in which the host microprocessor transfers data to and from memory via the computer's I/O ports. PIO enables very fast data transfer rates, especially in single-tasking operating systems like DOS.

Pipeline Burst Cache

A type of memory cache built into many modern DRAM controller and chipset designs. Pipeline burst caches use two techniques—a burst mode that pre-fetches memory contents before they are requested, and pipelining so that one memory value can be accessed in the cache at the same time that another memory value is accessed in DRAM. The purpose of pipeline burst caches is to minimize wait states so that memory can be accessed as fast as possible by the microprocessor.

Pixel

Short for Picture Element, a pixel is a single point in a graphic image. Graphics monitors display pictures by dividing the display screen into thousands (or millions) of pixels, arranged in rows and columns. The pixels are so close together that they appear connected.

The number of bits used to represent each pixel determines how many colors or shades of gray can be displayed. For example, in 8-bit color mode, the color monitor

uses 8 bits for each pixel, making it possible to display 2 to the 8th power (256) different colors or shades of gray.

On color monitors, each pixel is actually composed of three dots — a red, a blue, and a green one. Ideally, the three dots should all converge at the same point, but all monitors have some convergence error that can make color pixels appear fuzzy.

The quality of a display system largely depends on its resolution, how many pixels it can display, and how many bits are used to represent each pixel. VGA systems display 640 by 480, or about 300,000 pixels. In contrast, SVGA systems display 1,024 by 768, or nearly 800,000 pixels. True Color systems use 24 bits per pixel, allowing them to display more than 16 million different colors.

Platform

A combination of hardware and operating system you use, for example, the “NT platform” is a PC running the Microsoft Windows NT operating system and the “PPC platform” is a Macintosh computer with a PowerPC processor running the Mac operating system.

PLD

((Programmable Logic Device) A digital integrated circuit that can be programmed by the user to perform a wide variety of logical operations.

Plotter

A computer output device that draws images on paper using a pen. A plotter draws real lines rather than simulating them as a conventional printer would by producing a series of very close dots.

Plug-In

This is a program that your browser uses to manipulate a downloaded file. It differs from a Helper Application in that the plug-in works inside the browser window.

PNP

Short for Plug and Play, a technology developed by Microsoft and Intel that supports plug-and-play installation. PnP is built into the Windows 95 operating system, but to use it, the computer’s BIOS and expansion boards must also support PnP.

Port

This is the connecting component or hardware that allows two computers to attach to one another.

Portal

A web site that aims to be an entry point to the World-Wide Web, typically offering a search engine and/or links to useful pages, and possibly news or other services. These services are usually provided for free in the hope that users will make the site their default home page or at least visit it often. Popular examples are Yahoo and MSN. Most portals on the Internet exist to generate advertising income for their owners, others may be focused on a specific group of users and may be part of an

intranet or extranet. Some may just concentrate on one particular subject, say technology or medicine, and are known as a vertical portals.

Post

Short for power-on self test, a series of diagnostic tests that run automatically when you turn your computer on. The actual tests can differ depending on how the BIOS is configured, but usually the POST tests the RAM, the keyboard, and the disk drives. If the tests are successful, the computer boots itself. If the tests are unsuccessful, the computer reports the error by emitting a series of beeps and possibly displaying an error message and code on the display screen. The number of beeps indicates the error, but differs from one BIOS to another.

POP

Point of Presence, also Post Office Protocol Two commonly used meanings: A Point of Presence usually means a city or location where a network can be connected to, often with dial up phone lines. So if an Internet company says they will soon have a POP in Belgrade, it means that they will soon have a local phone number in Belgrade and/or a place where leased lines can connect to their network.

A second meaning, Post Office Protocol refers to a way that e-mail client software, such as Outlook, gets mail from a mail server. When you obtain an account from an Internet Service Provider (ISP) you almost always get a POP account with it, and it is this POP account that you tell your e-mail software to use to get your mail. Another protocol called IMAP is replacing POP for e-mail.

Primary Cache

Primary cache is the cache located closest to the CPU. Usually, primary cache is internal to the CPU, and secondary cache is external. Some early-model personal computers have CPU chips that don't contain internal cache. In these cases the external cache, if present, would actually be the primary (L1) cache.

Processor

A processor is a device that processes programmed instructions and performs tasks. Your processor sends and receives information from the different parts of the system (from hardware and software). The speed at which the CPU processes information internally is measured in MegaHertz (MHz) and GigaHertz (GHz). 1 GHz is equal to 1,000 MHz.

Protocol

An agreed-upon format for transmitting data between two devices. The protocol determines the following:

- The type of error checking to be used.
- Data compression method, if any.
- How the sending device will indicate that it has finished sending a message.
- How the receiving device will indicate that it has received a message. There are a variety of standard protocols from which programmers can choose. Each has

particular advantages and disadvantages; for example, some are simpler than others, some are more reliable, and some are faster. From a user's point of view, the only interesting aspect about protocols is that your computer or device must support the right ones if you want to communicate with other computers. The protocol can be implemented either in hardware or in software.

Proxy Server

A server that acts as an intermediary between a workstation user and the internet so that the enterprise can ensure security, administrative control, and caching service. A proxy server is associated with, or part of, a gateway server that separates the enterprise network from the outside network and a firewall server that protects the enterprise network from the outside intrusion.

PS/2 Port

A type of port developed by IBM for connecting a mouse or keyboard to a PC. The PS/2 port supports a mini DIN plug containing just 6 pins. Most PCs have a PS/2 port so that the serial port can be used by another device, such as a modem. The PS/2 port is often called the mouse port.

QAM

Quadrature Amplitude Modulation. A method of modulating digital signals using both amplitude and phase coding. Used for downstream and can be used for upstream.

QPSK

Quadrature Phase-Shift Keying. A method of modulating digital signals using four phase states to code two digital bits per phase shift.

Query

This is to make a computer request of a database.

RAID

Short for Redundant Array of Independent (or Inexpensive) Disks, a category of disk drives that employ two or more drives in combination for fault tolerance and performance. RAID disk drives are used frequently on servers but aren't generally necessary for personal computers.

There are number of different RAID levels. The three most common are 0, 3, and 5: Level 0: Provides data striping (spreading out blocks of each file across multiple disks) but no redundancy. This improves performance but does not deliver fault tolerance.

Level 1: Provides disk mirroring.

Level 3: Same as Level 0, but also reserves one dedicated disk for error correction data. It provides good performance and some level of fault tolerance.

Level 5: Provides data striping at the byte level and also stripe error correction information. This results in excellent performance and good fault tolerance.

RAM

(Random Access Memory) A configuration of memory cells that hold data for processing by a computer's central processing unit, or CPU; (see also memory). The term random derives from the fact that the CPU can retrieve data from any individual location, or address, within RAM.

Ranging

The process of automatically adjusting transmit levels and time offsets of individual modems, in order to make sure the bursts coming from different modems line up in the right timeslots and are received at the same power level at the CMTS.

RAS

Short for Remote Access Services, a feature built into Windows NT that enables users to log into an NT-based LAN using a modem, X.25 connection or WAN link. RAS works with several major network protocols, including TCP/IP, IPX, and Netbeui. To use RAS from a remote node, you need a RAS client program, which is built into most versions of Windows, or any PPP client software. For example, most remote control programs work with RAS.

RDRAM

Rambus DRAM technology is a system-wide, chip-to-chip interface design that allows data to pass through a simplified bus. Rambus uses a unique RSL (Rambus Signaling Logic) technology. Rambus is available in two flavors: RDRAM and Concurrent RDRAM. RDRAM is currently in production with Concurrent RDRAM production scheduled for late 1997. The third line extension, Direct RDRAM, is in development stages and scheduled for production in 1999. In late 1996, Rambus agreed to a development and license contract with Intel that will lead to Intel's PC chip sets supporting Rambus memory starting in 1999.

RealAudio

This is a method of playing sounds invented by Rob Glasser that creates a buffer between the supplying server and your computer. The file is played without downloading it completely.

Real Time

This is method of processing data the moment it is received. Batch mode is a term used for a mainframe computer dealing with data when it has the time.

Reboot

To restart a computer. In DOS, you can reboot by pressing the Alt, Control and Delete keys simultaneously. This is called a warm boot. You can also perform a cold boot by turning the computer off and then on again.

On Macs, you reboot by selecting the "Restart" option from the Special menu.

Redundant

Array of Inexpensive (or Interconnected) Disks. A performance enhancing method of storing the same data in different places on multiple hard disks. Unnecessary

redundancy can cause problems if one copy of the data is updated and another copy of the data is not. All redundant data can't be eliminated in most databases because primary keys in one table are repeated in other tables as foreign keys to create links between tables. This type of redundancy is called controlled redundancy.

Refresh

Generally, to update something with new data. For example, some Web browsers include a refresh button that updates the currently display Web pages. This feature is also called reload. To recharge a device with power or information. For example, dynamic RAM needs to be refreshed thousands of times per second or it will lose the data stored in it.

Similarly, display monitors must be refreshed many times per second. The refresh rate for a monitor is measured in hertz (Hz) and is also called the vertical frequency, vertical scan rate, frame rate or vertical refresh rate. The old standard for monitor refresh rates was 60 Hz, but a new standard developed by VESA sets the refresh rate at 75 Hz for monitors displaying resolutions of 640×480 or greater. This means that the monitor redraws the display 75 times per second. The faster the refresh rate, the less the monitor flickers.

Resolution

Refers to the sharpness and clarity of an image. The term is most often used to describe monitors, printers, and bit-mapped graphic images. In the case of dot-matrix and laser printers, the resolution indicates the number of dots per inch. For example, a 300-dpi (dots per inch) printer is one that is capable of printing 300 distinct dots in a line 1 inch long. This means it can print 90,000 dots per square inch.

For graphics monitors, the screen resolution signifies the number of dots (pixels) on the entire screen. For example, a 640-by-480 pixel screen is capable of displaying 640 distinct dots on each of 480 lines, or about 300,000 pixels. This translates into different dpi measurements depending on the size of the screen. For example, a 15-inch VGA monitor (640×480) displays about 50 dots per inch.

Printers, monitors, scanners, and other I/O devices are often classified as high resolution, medium resolution, or low resolution. The actual resolution ranges for each of these grades is constantly shifting as the technology improves.

Resource

Generally, any item that can be used. Devices such as printers and disk drives are resources, as is memory. In many operating systems, including Microsoft Windows and the Macintosh operating system, the term resource refers specifically to data or routines that are available to programs. These are also called system resources.

RFID

Radio Frequency identification (ID). Refers to the technology that uses devices attached to objects that transmit data to an RFID receiver. An alternative to bar coding. Advantages include data capacity, read/write capability, and no line-of-sight requirements.

RISC

Reduced Instruction Set Computer. A computer processing architecture that requires fewer instructions to run applications, thus increasing processing speed.

RJ-11

Short for Registered Jack-11, a four- or six-wire connector used primarily to connect telephone equipment in the United States. RJ-11 connectors are also used to connect some types of local-area networks (LANs), although RJ-45 connectors are more common.

RJ-45

Short for Registered Jack-45, an eight-wire connector used commonly to connect computers onto a local-area networks (LAN), especially Ethernets. RJ-45 connectors look similar to the ubiquitous RJ-11 connectors used for connecting telephone equipment, but they are somewhat wider.

ROM

Stands for Read-Only Memory. A semiconductor-based memory system that stores information permanently and does not lose its contents when power is switched off. ROMs are used for firmware, such as the BIOS used in the PC; and in some portable computers, application programs and even the operating system are being stored in ROM.

Router

A device that connects any number of LANs. Routers use headers and a forwarding table to determine where packets go, and they use ICMP to communicate with each other and configure the best route between any two hosts. Very little filtering of data is done through routers. Routers do not care about the type of data they handle.

Routing Switch

A switch that also performs routing operations. Usually a switch operates at layer 2 (the Data Link layer) of the OSI Reference Model while routers operate at layer 3 (the Network layer). Routing switches, however, perform many of the layer 3 functions usually reserved for routers. And because the routing is implemented in hardware rather than software, it is faster. The downside of routing switches is that they are not as powerful or as flexible as full-fledged routers.

Because they perform some layer 3 functions, routing switches are sometimes called layer-3 switches.

RPL

Request Parameter List. A VTAM (Virtual Telecommunications Access Method) control block that contains parameters necessary for processing a request (data transfer, connecting or disconnecting a terminal, etc).

Also, Relocatable Program Library. A data set used to store CICS (Customer Information Control System) application programs, which are fetched (loaded) at execution time.

RSS: (Rich Site Summary)

XML format for distributing news headlines on the Web, also known as Really Simple Syndication.

Scalable

A system or architecture is scalable when it can be changed in size or configuration to suit changing conditions. For example, a company that plans to set up a client/server network may want to have a system that not only works with the number of people who will immediately use the system, but the number who may be using it in one year, five years, or ten years.

Scalar Processing

A sequential operation in which one instruction produces one result; it starts an instruction, handles one operand or operand pair, and produces one result. Scalar processing complements vector processing by providing solutions to problems not readily adaptable to vector techniques.

Screen Flicker

The phenomenon whereby a display screen appears to flicker. Screen flicker results from a variety of factors, the most important of which is the monitor's refresh rate, the speed with which the screen is redrawn. If the refresh rate is too slow, the screen will appear to glimmer. Another factor that affects screen flicker is the persistence of the screen phosphors. Low-persistence phosphors fade more quickly than high-persistence monitors, making screen flicker more likely. Screen flicker can also be affected by lighting. Finally, screen flicker is a subjective perception that affects people differently. Some people perceive screen flicker where others do not. Most people perceive no screen flicker if the refresh rate is 72 MHz or higher.

SCSI

Abbreviation of Small Computer System Interface. Pronounced "scuzzy," SCSI is a parallel interface standard used by Apple Macintosh computers, PCs, and many UNIX systems for attaching peripheral devices to computers. Nearly all Apple Macintosh computers, excluding only the earliest Macs and the recent iMac, come with a SCSI port for attaching devices such as disk drives and printers.

SCSI interfaces provide for faster data transmission rates (up to 80 megabytes per second) than standard serial and parallel ports. In addition, you can attach many devices to a single SCSI port, so that SCSI is really an I/O bus rather than simply an interface.

The following varieties of SCSI are currently implemented:

- SCSI-1: Uses an 8-bit bus, and supports data rates of 4 Mbps
- SCSI-2: Same as SCSI-1, but uses a 50-pin connector instead of a 25-pin connector, and supports multiple devices. This is what most people mean when they refer to plain SCSI.
- Wide SCSI: Uses a wider cable (168 cable lines to 68 pins) to support 16 bit transfers.

- Fast SCSI: Uses an 8 bit bus, but doubles the clock rate to support data rates of 10 Mbps.
- Fast Wide SCSI: Uses a 16 bit bus and supports data rates of 20 Mbps.
- Ultra SCSI: Uses an 8 bit bus, and supports data rates of 20 Mbps.
- SCSI-3: Uses a 16 bit bus and supports data rates of 40 Mbps. Also called Ultra Wide SCSI.
- Ultra2 SCSI: Uses an 8 bit bus and supports data rates of 40 Mbps.
- Wide Ultra2 SCSI: Uses a 16 bit bus and supports data rates of 80 Mbps.

SDH

Synchronous Digital Hierarchy. A method used for multiplexing many circuits with a low bit rate onto fewer circuits with a higher bit rate, and vice-versa (de-multiplexing). Used primarily in the telecoms industry to carry telephony traffic. This network can also be used to carry IP traffic.

SDRAM

Short for Synchronous Dynamic Random Access Memory, a new type of DRAM that can run at much higher clock speeds than conventional memory. SDRAM actually synchronizes itself with the CPU's bus and is capable of running at 100 MHz, about three times faster than conventional FPM RAM, and about twice as fast EDO DRAM and BEDO DRAM. SDRAM is replacing EDO DRAM in many newer computers.

Today's fastest Pentium systems use CPU buses running at 100 MHz, so SDRAM can keep up with them, though barely. Future PCs, however, are expected to have CPU buses running at 200 MHz or faster. SDRAM is not expected to support these high speeds which is why new memory technologies, such as RDRAM and SDRAM, are being developed.

Secondary Cache

Short for Level 2 cache, cache memory that is external to the microprocessor. In general, L2 cache memory, also called the secondary cache, resides on a separate chip from the microprocessor chip. The Pentium Pro, however, has an L2 cache on the same chip as the microprocessor.

Semiconductor

A material that is neither a good conductor of electricity (like copper) nor a good insulator (like rubber). The most common semiconductor materials are silicon and germanium. These materials are then doped to create an excess or lack of electrons. Computer chips, both for CPU and memory, are composed of semiconductor materials. Semiconductors make it possible to miniaturize electronic components, such as transistors. Not only does miniaturization mean that the components take up less space, it also means that they are faster and require less energy.

SEO

(Search Engine Optimization) SEO is a process of arranging a web site's content to obtain high rankings in various search engines (both the site and individual pages),

and includes tailoring on-page text (such as headlines and subtitles) as well as choosing the proper keywords for a page's meta tags.

Serial Port

A port, or interface, that can be used for serial communication, in which only 1 bit is transmitted at a time. Most serial ports on personal computers conform to the RS-232C or RS-422 standards. A serial port is a general-purpose interface that can be used for almost any type of device, including modems, mice, and printers (although most printers are connected to a parallel port).

Server

This is a mainframe computer that serves the other computers attached to it.

SGRAM

Abbreviation of Synchronous Graphic Random Access Memory, a type of DRAM used increasingly on video adapters and graphics accelerators. Like SDRAM, SGRAM can synchronize itself with the CPU bus clock up to speeds of 100 MHz. In addition, SGRAM uses several other techniques, such as masked writes and block writes, to increase bandwidth for graphics-intensive functions.

Unlike VRAM and WRAM, SGRAM is single-ported. However, it can open two memory pages at once, which simulates the dual-port nature of other video RAM technologies.

Shadowing

A technique used to increase a computer's speed by using high-speed RAM memory in place of slower ROM memory (RAM is about three times as fast as ROM). On PCs, for example, all code to control hardware devices, such as keyboards, is normally executed in a special ROM chip called the BIOS ROM. However, this chip is slower than the general-purpose RAM that comprises main memory. Many PC manufacturers, therefore, configure their PCs to copy the BIOS code into RAM when the computer boots. The RAM used to hold the BIOS code is called shadow RAM.

Shareware

Software distributed on the basis of an honor system. Most shareware is delivered free of charge, but the author usually requests that you pay a small fee if you like the program and use it regularly. By sending the small fee, you become registered with the producer so that you can receive service assistance and updates. You can copy shareware and pass it along to friends and colleagues, but they too are expected to pay a fee if they use the product.

Shareware is inexpensive because it is usually produced by a single programmer and is offered directly to customers. Thus, there are practically no packaging or advertising expenses.

Shell

The outermost layer of a program. Shell is another term for user interface. Operating systems and applications sometimes provide an alternative shell to make interaction with the program easier. For example, if the application is usually command driven,

the shell might be a menu-driven system that translates the user's selections into the appropriate commands. Sometimes called command shell, a shell is the command processor interface. The command processor is the program that executes operating system commands. The shell, therefore, is the part of the command processor that accepts commands. After verifying that the commands are valid, the shell sends them to another part of the command processor to be executed.

SID (Service ID)

Used in the DOCSIS standard to defines a particular mapping between a cable modem (CM) and the CMTS. The SID is used for the purpose of upstream bandwidth allocation and class-of-service management.

SIMM

Acronym for single in-line memory module, a small circuit board that can hold a group of memory chips. Typically, SIMMs hold up 8 (on Macintoshes) or 9 (on PCs) RAM chips. On PCs, the ninth chip is often used for parity error checking. Unlike memory chips, SIMMs are measured in bytes rather than bits. SIMMs are easier to install than individual memory chips.

The bus from a SIMM to the actual memory chips is 32 bits wide. A newer technology, called dual in-line memory module (DIMM), provides a 64 bit bus. For modern Pentium microprocessors that have a 64 bit bus, you must use either DIMMs or pairs of SIMMs.

SIP

Abbreviation of single in-line package, a type of housing for electronic components in which the connecting pins protrude from one side. Compare with DIP and PGA. A SIP is also called a Single In-line Pin Package (SIPP).

Socket

In UNIX and some other operating systems, a software object that connects an application to a network protocol. In UNIX, for example, a program can send and receive TCP/IP messages by opening a socket and reading and writing data to and from the socket. This simplifies program development because the programmer need only worry about manipulating the socket and can rely on the operating system to actually transport messages across the network correctly. Note that a socket in this sense is completely soft—it's a software object, not a physical component.

A receptacle into which a plug can be inserted.

A receptacle for a microprocessor or other hardware component.

Socket 7

The form factor for fifth-generation CPU chips from Intel, Cyrix, and AMD. All Pentium chips, except Intel's Pentium Pro (Socket 8) and Pentium II (Slot 1), conform to the Socket 7 specifications. Intel has decided to phase out Socket 7 and replace it with Slot 1. But Intel's competitors, such as AMD and Cyrix, are sticking with Socket 7, and are developing an enhanced version.

Socket 8

The form factor for Intel's Pentium Pro microprocessors. The Pentium Pro was the first microprocessor not to use the venerable Socket 7 form factor. The Pentium II microprocessors use an even newer form factor called Slot 1. Socket 8 is a 387-pin ZIF socket with connections for the CPU and one or two SRAM dies for the Level 2 (L2) cache.

Software Modem

A modem implemented entirely in software. Software modems rely on the computer's processor to modulate and demodulate signals.

Source Code

Computer programs or operating systems are originally written by a human being in a programming language. This is called the source code of the software. To be actually used by a computer, the program has to be translated by the computer from the source code into the machine language that the computer understands and can execute. This translation process is referred to as compiling.

SLIP

Stands for Serial Line Interface Protocol. This is another application that allows for a connection to another computer.

SMTP

Stands for Simple Mail Transfer Protocol.

SNMP

Since it was developed in 1988, the Simple Network Management Protocol has become the *de facto* standard for Internet network management. Because it is a simple solution, requiring little code to implement, vendors can easily build SNMP agents to their products. SNMP is extensible, allowing vendors to easily add network management functions to their existing products. SNMP also separates the management architecture from the architecture of the hardware devices, which broadens the base of multi-vendor support. Perhaps most important, unlike other so-called standards, SNMP is not a mere paper specification, but an implementation that is widely available today.

Software

This is a program, the actual code the computer reads. All other stuff is hardware. A floppy disc is hardware.

Spoofing

To fool. In networking, the term is used to describe a variety of ways in which hardware and software can be fooled. Email spoofing, for example, involves trickery that makes a message appear as if it came from a legitimate business email address.

SQL

X(Structured Query Language) A specialized programming language for sending queries to databases. Most industrial-strength and many smaller database applications can be addressed using SQL. Each specific application will have its own version of

SQL implementing features unique to that application, but all SQL-capable database support a common subset of SQL.

SRAM

Short for static random access memory, and pronounced ess-ram. SRAM is a type of memory that is faster and more reliable than the more common DRAM (dynamic RAM). The term static is derived from the fact that it doesn't need to be refreshed like dynamic RAM.

While DRAM supports access times of about 60 nanoseconds, SRAM can give access times as low as 10 nanoseconds. In addition, its cycle time is much shorter than that of DRAM because it does not need to pause between accesses. Unfortunately, it is also much more expensive to produce than DRAM. Due to its high cost, SRAM is often used only as a memory cache.

SSL: (Secure Sockets Layer)

A protocol designed by Netscape Communications to enable encrypted, authenticated communications across the Internet. SSL is used mostly in communications between Web browsers and Web servers. URLs that begin with "https" indicate that an SSL connection will be used.

Subroutine

A procedure that performs a specific function; usually a process that may be needed several times or a routine that may be used in several different programs. For example, many subroutines have been created to parse URL-encoded data.

Subscriber Unit (SU)

An alternate term for cable modem.

SVGA

Stands for Super Video Graphics Adapter. It's a high level monitor.

Swap File

A swap file is an area on your hard disk used as virtual memory. It's called a swap file because virtual memory management software swaps data between it and main memory (RAM).

Swap Space

Disk space used by the kernel as "virtual" RAM. It is slower than RAM, but because disk space is cheaper, swap is usually more plentiful. Swap space is useful to the kernel for holding lesser-used data and as a fallback when physical RAM is exhausted.

Switch

In networks, a device that filters and forwards packets between LAN segments. Switches operate at the data link layer (layer 2) of the OSI Reference Model and therefore support any packet protocol. LANs that use switches to join segments are called switched LANs or, in the case of Ethernet networks, switched Ethernet LANs.

A small lever or button. The switches on the back of printers and on expansion

boards are called DIP switches. A switch that has just two positions is called a toggle switch. Another word for option or parameter—a symbol that you add to a command to modify the command's behavior.

Switching Hub

Short for port-switching hub, a special type of hub that forwards packets to the appropriate port based on the packet's address. Conventional hubs simply rebroadcast every packet to every port. Since switching hubs forward each packet only to the required port, they provide much better performance. Most switching hubs also support load balancing, so that ports are dynamically reassigned to different LAN segments based on traffic patterns. Some newer switching hubs support both traditional Ethernet (10 Mbps) and Fast Ethernet (100 Mbps) ports. This enables the administrator to establish a dedicated, Fast Ethernet channel for high-traffic devices such as servers.

Synchronous

Synchronous can refer to: (1) A communications method that transmits a group of characters as a block of data rather than as individual characters. (2) A reference to the fact that two different data streams are tied, or synchronized, to a single reference clock. (3) Data transmitted in a time-division multiplexer.

TCP/IP

Acronym for Transmission Control Protocol/Internet Protocol, the suite of communications protocols used to connect hosts on the Internet. TCP/IP uses several protocols, the two main ones being TCP and IP. TCP/IP is built into the UNIX operating system and is used by the Internet, making it the *de facto* standard for transmitting data over networks. Even network operating systems that have their own protocols, such as Netware, also support TCP/IP.

Telnet

One of the TCP/IP Protocols. It allows a connection to another computer over dedicated phone lines.

Terabyte: (TB)

2 to the 40th power (1,099,511,627,776) bytes. This is approximately 1 trillion bytes.
10 to the 12th power (1,000,000,000,000). This is exactly one trillion.

Terminal

This is what you look at when you're on the Internet. It's your computer screen.

Terminator

A device attached to the end-points of a bus network or daisy-chain. The purpose of the terminator is to absorb signals so that they do not reflect back down the line. Ethernet networks require a terminator at both ends of the bus, and SCSI chains require a single terminator at the end of the chain. A character that indicates the end of a string. In the C programming language, the null character serves as a terminator.

Terminal Emulation

This is an application that allows your terminal to act as a dumb terminal.

Thread

In online discussions, a series of messages that have been posted as replies to each other. A single forum or conference typically contains many threads covering different subjects. By reading each message in a thread, one after the other, you can see how the discussion has evolved. You can start a new thread by posting a message that is not a reply to an earlier message.

Throughput

The amount of data transferred from one place to another or processed in a specified amount of time. Data transfer rates for disk drives and networks are measured in terms of throughput. Typically, throughputs are measured in Kbps, Mbps and Gbps.

Transceiver

Short for transmitter-receiver, a device that both transmits and receives analog or digital signals. The term is used most frequently to describe the component in local-area networks (LANs) that actually applies signals onto the network wire and detects signals passing through the wire. For many LANs, the transceiver is built into the network interface card (NIC). Some types of networks, however, require an external transceiver. In Ethernet networks, a transceiver is also called a Medium Access Unit (MAU).

Transfer Rate

The speed at which a disk drive can transfer information between its platters and your CPU. The transfer rate is typically measured in megabytes per second, megabits per second, or megahertz.

Transparent

Something that occurs without being known to the user.

Transistor

A device composed of semiconductor material that amplifies a signal or opens or closes a circuit. Invented in 1947 at Bell Labs, transistors have become the key ingredient of all digital circuits, including computers. Today's microprocessors contain tens of millions of microscopic transistors.

Prior to the invention of transistors, digital circuits were composed of vacuum tubes, which had many disadvantages. They were much larger, required more energy, dissipated more heat, and were more prone to failures. It's safe to say that without the invention of transistors, computing as we know it today would not be possible.

True Color

Refers to any graphics device or software that uses at least 24 bits to represent each dot or pixel. Using 24 bits means that more than 16 million unique colors can be represented. Since humans can only distinguish a few million colors, this is more than enough to accurately represent any color image.

Turnkey System

A system that already contains all the components and programs required for operation. The vendor takes care of installation and configurations so all the user has to do is "turn the key" to begin using the system.

TWAIN

Stands for Technology Without An Interesting Name.

V.90

A standard for 56 Kpbs modems approved by the International Telecommunication Union(ITU) in February, 1998. The V.90 standard resolves the battle between the two competing 56 Kbps technologies -X2 from 3COM and K56Flex from Rockwell Semiconductor. Both manufacturers have announced that their future modems will conform to V.90. In addition, most users who already purchased 56 Kbps modems will be able to apply a software upgrade to make their modems support V.90.

VDD

Stands for Virtual Device Driver.

Veronica

Stands for Very Easy Rodent Oriented Net-wide Index to Computerized Archives. A database of menu names from a large number of Gopher servers. A quick and easy way to search Gopher resources for information by keyword.

VGA

Stands for Video Graphics Adapter. This is a lower level color monitor.

Virtual Device Driver

In Windows systems, a special type of device driver that has direct access to the operating system kernal. This allows them to interact with system and hardware resources at a very low level. In Windows 95, virtual device drivers are often called VxDs because the filenames end with the .vxd extension.

Virtual Machine

A self-contained operating enviornment that behaves as if it is a separate computer. For example, Java applets run in a Java virtual machine (VM) that has no access to the host operating system. This design has two advantages:

- System Independence: A Java application will run the same in any Java VM, regardless of the hardware and software underlying the system.
- Security: Because the VM has no contact with the operating system, there is little possibility of a Java program damaging other files or applications.

The second advantage, however, has a downside. Because programs running in a VM are separate from the operating system, they cannot take advantage of special operating system features.

Virtual Private Network (VPN)

A data network that uses the public telecommunications infrastructure, but maintains privacy through the use of a tunneling protocol and security procedures. A VPN gives a company the same capabilities as a system of owned or leased lines to which that company has exclusive access. However, costs are much lower because the VPN uses the shared public infrastructure rather than exclusive line access.

VIRUS

A program or piece of code that is loaded onto your computer without your knowledge and runs against your wishes. Most viruses can also replicate themselves. All computer viruses are manmade. A simple virus that can make a copy of itself over and over again is relatively easy to produce. Even such a simple virus is dangerous because it will quickly use all available memory and bring the system to a halt. An even more dangerous type of virus is one capable of transmitting itself across networks and bypassing security systems.

VMS

Stands for Virtual Memory System.

Voltage Regulator

A device which maintains constant voltage in an electrical line in case of brownout.

VRAM

Video Random Access Memory—A kind of high-speed memory used for the computer's display. VRAM must be fast to keep up with the speed at which the screen is scanned. The VRAM in a PC is on a display adapter card. VRAM has two ports so it can send the data for text and images to memory and to the display at the same time.

VRML

Stands for Virtual Reality Modeling Language. It's a form of application that gives a 3-D effect to pictures sometimes allowing you to "move" through them.

VTAM

Virtual Telecommunications Access Method—The SNA software that runs on IBM mainframes and implements the functions of network control, network management interface, and SNA support for host based application programs.

VxD

Virtual Device Driver—A device driver under Windows 3.x/Windows 95 running as part of the kernel and thus having access to the memory of the kernel and all running processes as well as raw access to the hardware. VxD's usually have the filename extension .386 under Windows 3.x and .vxd under Windows 95. VxD's written for Windows 3.x can be used under Windows 95 but not vice versa.

WWW

Short for World Wide Web.

WAIS

Stands for Wide Area Information Servers. Searches large indexes of information on the Internet.

Wait State

A brief delay added before a microprocessor executes an instruction, to allow time for slower memory chips or external devices to respond. A wait states may be one or more of the computer's clock cycles or may be timed differently. One wait state on each access of memory can make the processor up to 20% slower. With no wait state (called zero wait state) the processor will run faster.

WAN

Wide Area Network—A network in which computers are connected to each other over a long distance, using telephone lines and satellite communications. See local area network (LAN).

WAV

Stands for WAVEform sound format. Microsoft's format for encoding sound files.

Weblog

(Same as blog) This is a publicly accessible personal journal for an individual. Similar to a personal diary, but shared over the web. The activity of updating a blog is "blogging" and someone who keeps a blog is a "blogger." Blogs are typically updated daily using software that allows people with little or no technical background to update and maintain the blog. Postings on a blog are almost always arranged in chronological order with the most recent additions featured most prominently.

White Papers

These are documents created to help guide you in the use of a particular piece of hardware or software. Usually some kind of extra or advanced help that is provided that goes beyond the normal manual or set of instructions for that product or item. White Papers can provide special or advanced instructions on special features or setup methods.

WiFi

Wireless Fidelity otherwise known as Wireless Networking, commonly using the 802.11b protocol. Hardware that displays the WiFi logo claims 802.11b compliance should interconnect seamlessly.

WinFS

(Windows File System) WinFS is the code name for the next generation storage platform in Windows "Longhorn." Taking advantage of database technologies, Microsoft is advancing the file system into an integrated store for file data, relational data, and XML data. Windows users will have intuitive new ways to find, relate, and act on their information, regardless of what application creates the data. Also, "WinFS" will have built-in support for multi-master data synchronization across other Longhorn machines and other data sources. The platform supports rich managed Longhorn APIs as well as Win32 APIs.

WOL: (Wake-on-LAN)

This technology is used to remotely wake up a sleeping or powered off PC over a network. When the system is turned off, the managed network adapter uses an alternate power source to monitor the network and watch for a wake-up packet from the server. Once it receives a packet, it alerts the system to power up and accept any maintenance task it is given. Wake-on-LAN is a part of Intel's Wired for Management System and is a result of the Intel-IBM Advanced Manageability Alliance.

Wake-on-LAN is also called remote wake-up.

Workgroup

Persons sharing files and data between themselves.

Workstation

The computer attached to the Internet.

WPG

Stands for Word Perfect Graphics.

UDMA

A protocol developed by Quantum Corporation and Intel that supports burst mode data transfer rates of 33.3 Mbps. This is twice as fast as the previous disk drive standard for PCs, and is necessary to take advantage of new, faster Ultra ATA disk drives.

The official name for the protocol is Ultra DMA/33. It's also called UDMA, UDMA/33 and DMA mode 33.

UNIX

This is an operating system developed by AT&T. It's big push it that it allows one server to service many different end users at one time.

UPS

Uninterruptible Power Supply. A backup power unit that provides continuous power when the normal power supply is interrupted. UPS systems can be stand-by, only supplying power when the regular supply is interrupted, or fulltime, relying on regular power and/or batteries to supply it while it supplies power to the protected device. A UPS is not necessary on most computer systems, but can be important on systems that need to be up 24 hours a day, such as servers.

Upstream

The data flowing from the Cable Modem to the CMTS.

Upstream frequency

The frequency used to transmit data from the CM to the CMTS. Normally in the 5-42 MHz range for US systems and 5-65 MHz for European systems.

URL

Stands for Universal Resource Locator. It's a fancy way of saying Internet Address.

USB

Short for Universal Serial Bus, a new external bus standard that supports data transfer rates of 12 Mbps (12 million bytes per second). A single USB port can be used to connect up to 128 peripheral devices, such as mice, modems, and keyboards. USB also supports Plug-and-Play installation and hot plugging.

Starting in 1996, a few computer manufacturers started including USB support in their new machines. Since the release of Intel's 440LX chipset in 1997, USB has become more widespread. It is expected to eventually completely replace serial and parallel ports.

User

Someone attached to a server or host.

X2

A technology developed by U.S. Robotics (now 3COM) for delivering data rates up to 56 Kbps over plain old telephone service (POTS). It was long believed that the maximum data transmission rate over copper telephone wires was 33.6 Kbps, but X2 achieves higher rates by taking advantage of the fact that most phone switching stations are connected by high-speed digital lines. X2 bypasses the normal digital-to-analog conversion and sends the digital data over the telephone wires directly to your modem where it is decoded.

XML: (eXtensible Markup Language)

Like HTML, XML is a markup language, but unlike HTML, it is not limited to Web documents. XML lets Web developers and designers create customized tags that offer greater flexibility in organizing and presenting information than is possible with the older HTML document coding system.

ZIF Socket

Zero Insertion Force socket—A special socket for plugging in integrated circuits easily. The socket can be opened with a small lever or screw; the chip is dropped in, then the socket is closed.

ZIP

Stands for Zone Information Protocol. This is an application that allows for the compression of application files.



INDEX

Symbols

100BASE-T, 119
10BASE-2, 115
10BASE-5, 114
10BASE-T, 117
10BROAD-36, 115
1BASE-5, 116

A

Accounting Management, 246
Addressing, 199
Address Resolution Protocol (ARP), 178
Advanced Mobile Phone Service (AMPS), 53
American National Standards Institute, 11
Analog Encoding, 41
Architecture, 240
ARPANET, 24
Asynchronous Balanced Mode (ABM), 90
Asynchronous Response Mode (ARM), 90
Attenuation, 44
Authentication, 256
Authentication Header (AH), 190

B

Base Station System, 55
Biometric Identification, 255
Bit Stuffing, 68
Bluetooth, 153
Bluetooth Frame Structure, 156
Border Gateway Protocol, 188
Bridges, 159
Burst Error, 69
Bus, 6

C

Carrier Sense Multiple Access, 105
Channelization, 109
Character Count, 66
Circuit Switching, 31
Coaxial Cable, 46, 111
Communication, 2
Computer Network Types, 3
Computer Security, 253
Confidentiality, 256
Configuration Management, 245
Connected-Oriented Services, 29
Connection Establishment, 200
Connection Management, 198
Connection Release, 203
Convolutional codes, 78
Cryptography, 269
Cyber Criminals, 267
Cyclic Codes, 75

D

Data Field, 127
Data Link Layer, 20
Data Transfer, 198
Decapsulation, 209
Decnet's DNA, 35
Delivery and Forwarding, 164
Demultiplexing, 204
Design Constraints, 133
Design Philosophy, 261
Destination Address, 138
Destination Address Field, 125
Destination Port Number, 207

Digital Cellular System 1800 (DCS1800), 56
 Digital Encoding, 42
 Digital Network Architecture, 35
 Distributed Coordination Function, 146
 Domain Hierarchy, 227
 Domain Name System, 225

E

Electronic Industries Association, 12
 Electronic Mail, 230
 Encapsulating Security Payload (ESP), 190
 Encapsulation, 209
 Error Control, 22
 Expedited Delivery, 198

F

Fault Management, 244
 Federal Communications Commission, 12
 Federal Information Processing Standards, 259
 Feedback Error Control, 82
 Fiber-optic Transmission, 47
 Firewalls, 276
 Flag Bytes, 67
 Flow Control, 204
 Forensic Analysis, 266
 Forward Error Control, 70
 Frame Check Sequence, 139
 Frame Composition, 123
 Functional Address Indicator, 139

G

Generic Domains, 229
 Gigabit Ethernet, 121
 Global System for Mobile Communications (GSM), 54
 Group Address, 91
 Guard Bonds, 109
 Guard Time, 110

H

Hackers, 268
 Hacktivists, 269

Hamming Codes, 70
 Hamming Decoder, 74
 Hamming Encoder, 74
 Handling Mailboxes, 232
 Hardware, 254
 HDLC Frame structure., 90
 High Level Data Link Control, 89
 Hypertext Transfer Protocol (HTTP), 242

I

Identification Cards, 255
 IEEE 802, 101
 IEEE 802.11a, 144
 IEEE 802.11b, 145
 IEEE 802.11g, 145
 Information Frames, 92
 Information Security, 254
 Integrity, 257
 International Telecommunication Union, 11
 Internet Message Access Protocol, 237
 Internet Protocol Security, 189
 IP Addresses, 172
 IP Version 6, 176
 IPv4 Packet Format, 171

L

LANA Ranges, 200
 Large-Scale-Integration (LSI), 10
 Length Field, 127
 Link Control Protocol, 98
 Link Management, 88
 Load Balancing, 187
 Local Area Networks (LANs), 3
 Logical Addresses, 27
 Logical Link Control (LLC), 102
 Logical Unit, 33

M

M.3100 Recommendation, 249
 Media Support, 122

Mesh, 5

Message Access Agent, 237

Message Transfer Agent, 235

Metropolitan Area Networks (MANs), 4

Microwave, 47

Modulo-2 Division Circuits, 76

Multiaccess, 100

Multimode Fiber, 48

Multiple Access Protocols, 102

Multiplexing, 204

N

Network Connectivity, 9

Network Address Translation (NAT), 174

Network Elements, 2

Network Layer, 21

Network Management, 243

Network Security, 253

Network Services, 29

Network Topology, 5

Next-Hop Method, 165

Nonrepudiation, 257

Nordic Mobile Telephone (NMT) System, 52

Normal Response Mode (NRM), 90

Nyquist, 40

O

Open Shortest Path First, 186

Open System Interconnection, 12

Orange Book, 260

OSI Model, 13

Operation Subsystem (OSS), 56

P

Packet Switching, 32

Passive Hubs, 158

Performance Management, 246

Personal Communication Services (PCS)

Personal Digital Assistant, 142

Physical Layer, 19, 120

Piconets, 154

Point Coordination Function, 148

Port Addresses, 27

Preamble Field, 124

Presentation Layer, 23

Protocol, 9, 97

Public Key Encryption, 273

Public-Key Cryptography Standards, 259

R

Random Access, 102

Real-time Transmission, 1

Redundancy, 70

Repeaters, 158

Resource Records, 228

Ring, 8

Route Method, 165

Risk Assessment, 266

Routing Information Protocol, 186

S

Scatternet, 154

Secure Socket Layer (SSL), 220, 259

Security, 199

Security Associations, 191

Security Threat Management, 265

Selective-Repeat ARQ, 84

Service Primitives and Parameters, 18

Session Layer, 23

Simple Mail Transfer Protocol, 235

Single-Bit Error, 69

Single-Mode Fiber, 48

Socket Addresses, 200

Software, 254

Software Access Control Systems, 255

SONET/SDH, 60

Source Address Field, 126

Source Port Number, 207

SSL Protocol Stack, 222

Standards, 10

Star, 7

Status Reporting, 199

Stop-and-Wait Protocol, 85
 Subnetting, 173
 Supernetting, 173
 Supervisory Frames, 92
 Symmetric Encryption, 271
 Synchronous Digital Hierarchy, 60
 Synchronous Optical Network, 60
 System Network Architecture, 32

T

TCP Header Format, 215
 TCP Mechanisms, 216
 TCP Services, 211
 TCP/IP Reference Model, 24
 TDMA, 110
 Third Generation Cellular Systems, 57
 Three-Way Handshake, 216
 Time-division Multiple Access, 110
 Token Bus, 140
 Token-Ring, 127
 Topology, 128
 Total Access Communications System (TACS), 53
 Transceiver, 112
 Transceiver Cable, 112
 Transmission Formats, 135
 Transport Layer, 21
 Transport Layer Security (TLS), 223
 Transport Services, 196
 Tree, 6

Twisted-pair Cable, 43
 Two-Layer Switches, 160
 Type Field, 126

U

UDP Header, 207
 UDP Message Queue, 210
 UDP Pseudo Header, 208
 Unnumbered Frames, 93
 Universal Resource Locator (URL), 28
 User Agent Types, 232
 User Datagram Protocol (UDP), 25, 207
 User Interface, 198

V

Very-Large-Scale-Integration (VLSI), 10
 Virtual Circuit Network, 32
 Virtual Private Networks (VPNs), 192
 Visitor Location Register (VLR), 51
 Visual Event Monitoring, 255

W

Wavelength Division Multiplexing (WDM), 48
 Wide Area Networks (WANs), 4
 Window Mechanisms, 87
 WLAN, 142
 World Wide Web, 240
 Wireless Lan, 141
 Wired Equivalent Privacy (WEP), 142



COMPUTER NETWORKS

Er. Vikrant Vij

ABOUT THE BOOK

In this fast pace of computer world where almost everyday something new is introduced in this faculty of studies, learning computer and obtaining further knowledge in this subject has become an essential need for everyone. Considering this crucial aspect in view, an endeavor has been made to bring out such an ideal book on computer science that can benefit the students and teachers alike. And thus, the book titled, **Computer Networks** has been brought out to meet their requirements.

Examining a wide range of techniques, technologies and systems utilized in data communications and computer networks, **Computer Networks** has been prepared to address various data transmission methods used in conveying data between physically distant locations. Several types of network are brought into consideration that may interconnect thousands of users on both permanent as well as temporary, switched basis. To execute successful communication, a set of rules and procedures has been developed, many of which are embodied in internationally accepted standards.

The organized material i.e. the outcome of this book is an effort to provide a reasonable magnitude of balance between rigor, clarity of presentation while keeping the length of the book at manageable level. Consisting of eight chapters, its prime objective is to provide an introduction to different computer network fundamentals and to develop a foundation that can be utilized as the basis for research and further studies in this sphere.

ABOUT THE AUTHOR

Er. Vikrant Vij is an academician to the core. He did his B.Tech. (with distinction) in ECE from PTU, Jalandhar and M.E. in ECE from NITTTR, Chandigarh. He has contributed and presented several research and review papers in various conferences across the country. He has published a number of textbooks in communication engineering. His areas of interest include Mobile and optical communications, VLSI design and embedded applications. He is Editor-in-Chief of International Journal of VLSI and Signal Processing Applications. He is currently working as Faculty at JPTC, Samipur, Himachal Pradesh. He is Member of IAENG, ISTE, AIRCC and IACS IT. He is a dedicated social worker and Traditional Reiki master and used to conduct counseling sessions for budding engineers.



UNIVERSITY SCIENCE PRESS

(An Imprint of Laxmi Publications Pvt. Ltd.)

An ISO 9001:2015 Company

ISBN 978-93-5274-080-2



9789352740802-0395

UCN-9804-395-COMPUTER NETWORK-VIJ