A Professional's Guide to

DATA COMMUNICATION *in a* TCP/IP WORLD

E. Bryan Carne

A Professional's Guide to Data Communication in a TCP/IP World

For a listing of recent titles in the *Artech House Telecommunications Library*
turn to the back of this book.

A Professional's Guide to Data Communication in a TCP/IP World

E. Bryan Carne

Artech House, Inc.
Boston • London
www.artechhouse.com

Library of Congress Cataloging-in-Publication Data
Carne, E. Bryan, 1928–
A professional's guide to data communication in a TCP/IP world / E. Bryan Carne.
p. cm.
Includes bibliographical references and index.
ISBN 1-58053-909-2 (alk. paper)
1. TCP/IP (Computer network protocol). 2. Data transmissions systems. I. Title.
TK5105.585.C36 2004
004.6'2—dc22
2004053826

British Library Cataloguing in Publication Data
Carne, E. Bryan (Edward Bryan), 1928–
A professional's guide to data communication in a TCP/IP world.—(Artech House telecommunications library)
1. Computer networks 2. TCP/IP (Computer network protocol)
I. Title
004.6
ISBN 1-58053-909-2

Cover design by Gary Ragaglia

685 Canton Street
Norwood, MA 02062

International Standard Book Number: 1-58053-909-2

10 9 8 7 6 5 4 3 2 1

To Joan, Kevin, Benjamin, and Matthew
with thanks for your outstanding support

Contents

Preface

There is nothing so certain in this world as change. Throughout the ages, wise men have made this point, and for several hundred years, change, in the form of the Industrial and Electronic Revolutions, has affected us all. As technology feeds on itself, the process continues. This book is about change, about the ability of the Internet to dictate technical direction through its overwhelming presence. With more than 200 million hosts generating traffic in this network of networks, it is no wonder that TCP/IP has become the protocol suite of choice to support the exchange of messages in commercial operations and residential activities. Developed initially for point-to-point data operations, it has been adapted to local area networks, wide area networks, radio networks, and for voice services, to the detriment of all other protocol suites. Data communication is an essential part of our lives. It continues to evolve to an activity largely directed by TCP/IP.

In writing this book, I have assumed that the reader is familiar with common telecommunications terms and practices. For those who may need a refresher, Appendix A describes some of the basic concepts that are employed in the text.

My book provides a comprehensive picture of the Internet protocol stack and the role of TCP/IP in data communications. It describes the TCP/IP suite in some detail and, for handy reference, contains Appendix B, which lists the fields of frames and headers used in this activity.

The book is a guide to the protocols, networks, codes, signals, and equipment that make it possible to communicate using TCP/IP. It explains advanced LAN and WAN technologies and gives an integrated view of bridging, routing, tagging, and labeling operations. In addition, it describes local loop technologies, particularly the limitations of twisted pairs, the use of optical fibers and radio, and the potential of pervasive voice over IP. This book is a ready reference to all aspects of data communication employing TCP/IP and includes a substantial glossary to provide explanations of the special terms that are the burden of every book on communications.

Conscious of my inability to treat each topic in detail, I have not tried to write a design manual. My intention is to paint the scene, to chronicle what is involved, and to promote understanding of how the pieces fit together. Where can you get further information? I have included a list of books that I like, and use, that can be of help. However, I suggest that the way to start is to use the services of a good search engine. There are hundreds of pages available on almost every subject that can point you in the right direction. We are in a dynamic environment. Change is everywhere, and new ways of doing things are being proposed even as you read these words. Like your new computer, most printed knowledge has aged, and is becoming obsolete, even before you purchase it.

Whether you are an IT professional, a business professional with data responsibilities, or a communications engineer wanting a handbook on the application of TCP/IP in contemporary communications, I hope you will find this attempt to cover the field in one volume worthwhile. In addition, if you are an undergraduate computer science or engineering student or a continuing education student with a software or communications concentration, I hope you will explore the field of data communication with this book as your guide.

Acknowledgments

In writing my book, an anonymous reviewer suggested a reorganization that improved the presentation immensely and had helpful comments on the contents. I thank him for his insight and the time he spent with my manuscript. In addition, I want to thank Judi Stone of Artech House for showing me that her PC world and my Mac world are compatible, Mark Walsh and his staff for helping me focus my efforts, Barbara Lovenvirth for editing the final manuscript, and Jill Stoodley and Rebecca Allendorf for managing its production. Finally, I want to thank my wife Joan, my son Kevin, and my grandsons Benjamin and Matthew for keeping everything going during the writing of this book.

CHAPTER 1

A TCP/IP World?

When he received a message from Alfred Vail, Samuel Morse is said to have exclaimed, “What hath God wrought?” On May 24, 1844, the pair showed they could communicate with electricity over a wire that ran between Washington, D.C., and Baltimore. Theirs was the first practical demonstration of long-distance *digital* communication. For several years the telegraph remained a scientific curiosity. Then, as the railroads expanded, eager entrepreneurs began wiring the country. As a result, in every village and town, Civil War battles were reported within hours. Telephone soon followed. It added more wires to the layers that festooned urban areas. Now, at the beginning of the twenty-first century, we have a pervasive communication network that encompasses the globe. Over it, with the appropriate terminal, we can send data, voice, and video messages to virtually anyone. A major component of this network, the Internet, is known in every household and enterprise and is used by many. What hath God wrought, indeed!

At first, data communication meant sending a fixed format message between two points. Telegrams were sent this way. If they needed to go further than one link could carry them, they were repeated over the next link, and the next, until they arrived at the terminal closest to their destination. There, they were printed and delivered by hand. Originally converted into coded signals with a manual key and sounder, ingenious persons soon perfected ways to automate sending and receiving. Eventually, it was possible for the sender to type the message on a *teletypewriter* and for the receiver to receive a printed copy on a similar machine known as a *teleprinter*. Connections remained primarily point to point.

Not long after the development of electronic computers, inventors saw that computer uses could be enhanced if these machines would communicate with one another. They understood that creating the *information age* required collecting data from anywhere, processing them somewhere, and disseminating the information products to any points that wanted to use them. Moreover, if this was done in close to real time, many operations could be automated. Pressures such as this led to experiments and, eventually, to the OSI and Internet communication models described in Chapter 2. They add layers of software procedures that expand simple point-to-point data transfer to complex data communication tasks in ever-growing networks.

Many of the stakeholders in the OSI model were governments and international standards agencies. They worked diligently to produce an efficient protocol suite that could be adopted universally. However, while the international bodies studied the problems they were creating, ARPAnet was showing an effective protocol suite for data communication over metropolitan, continental, and intercontinental distances. Soon, it became obvious to many that what eventually became known as TCP/IP was more flexible (i.e., could accommodate any style of networking) and more scalable (i.e., could handle growing networks efficiently) than the OSI contender. These advantages remain true today.

1.1 The Internet

In 1969, the Department of Defense commissioned its *Advanced Research Projects Agency* (ARPA) to develop a data network. From a few nodes located at academic institutions, ARPAnet has grown into the Internet, the largest cooperative venture ever undertaken by mankind. Extraordinarily complex, Internet Software Consortium (http://www.isc.org) estimates that, in January 2004, 233 million hosts were advertised in the *Domain Name System* (DNS). At the beginning of 1998, they reported just 30 million hosts. Described as a network of networks, the Internet consists of local, regional, and national networks that pass traffic to each other. Three organizations contribute to the operation and evolution of the Internet; they are:

- *Internet Society:* This organization promotes cooperation and coordination. An international body, it is concerned with network architecture, the evolution of protocols, and numbering. These tasks are performed through the *Internet Activities Board* (IAB), the *Internet Engineering Task Force* (IETF), and the *Internet Research Task Force* (IRTF). The Internet Society coordinates the activities of the *Internet Assigned Numbers Authority* (IANA) with IETF.
- *Internet Registry:* This organization administers *generic Top-Level Domains* (gTLDs) in cooperation with the *Council of Registrars* (CORE).
- *World Wide Web Consortium:* This is an industry consortium that develops standards for the World Wide Web.

Committees of specialists from governments, universities, and commercial entities assist each of these organizations, and some of the work is contracted to private industry. Using documents known as *Request for Comments* (RFCs), standards, protocols, and specifications for all facets of the Internet are developed and promulgated. Under the direction of the IETF, RFCs progress through several consensus-building stages. Ultimately, they become official documents describing the Internet and are archived by the IAB. Several thousand RFCs exist. They are available electronically from a number of sites.

Network operators are divided in three tiers. *Tier 1* contains operators that provide networks with a national reach and are largely responsible for backbone operations. *Tier 2* contains operators that provide regional networks and may engage in backbone operation. *Tier 3* contains operators that provide local networks and may operate a connection to the backbone. Within their networks (called *autonomous* networks), the operators are responsible for establishing operating discipline. Furthermore, they must cooperate with their neighbors with whom they share connections and agree upon the discipline to pass traffic between their networks.

Traffic is exchanged among autonomous networks at exchange points. At the lowest level, autonomous networks exchange traffic that is generated in a metropolitan area or large local area, and provide transit to a higher-level exchange for traffic destined elsewhere. At the higher level, they exchange traffic generated by networks in a region and provide transit for traffic destined for other regions or international points. At the highest level, they exchange traffic on a national and an international level. Originally, the *National Science Foundation* (NSF) and some national carriers established four national network access points (NAPs) in San Francisco, Chicago, Washington, D.C., and New York. Since then, they have been supplemented by around 10 *metropolitan area exchanges* (MAEs) in major metropolitan areas and many more *Internet eXchange Points* (IXPs) in smaller metropolitan complexes. Internet exchanges have been established in developed (and developing) countries so that Internet traffic can flow to most regions of the world.

1.1.1 TCP/IP Suite

Communication in the Internet is facilitated by protocols identified, in short, as TCP/IP and often simply as IP. Computer protocols are procedures performed at the behest of application processes. Applications are the elements for which the entire network is established; they manipulate data and request communication to move data from place to place:

- *TCP* is an acronym for *Transmission Control Protocol*; it governs the reliable, sequenced, and unduplicated delivery of data. A related transport protocol is called UDP, an acronym for User Datagram Protocol. It provides data transport on a best-effort basis without acknowledgments or guaranteed delivery.
- *IP* is an acronym for *Internet Protocol*; its major purpose is to make origination and destination addresses available to guide data across networks. IP includes several management protocols that are essential to the operation of the Internet.

Together, TCP, UDP, IP, and associated protocols are known as the *TCP/IP suite*.

TCP/IP facilitates interconnection and internetworking. Since 1982, when the Defense Communications Agency declared it to be the protocol suite for ARPAnet, the basic technology has demonstrated both robustness and scalability. Developed initially for point-to-point operations, it has survived more than two decades of exponential growth. During that time, the suite has been adapted to local area networks, wide area networks, radio networks, and for voice services.

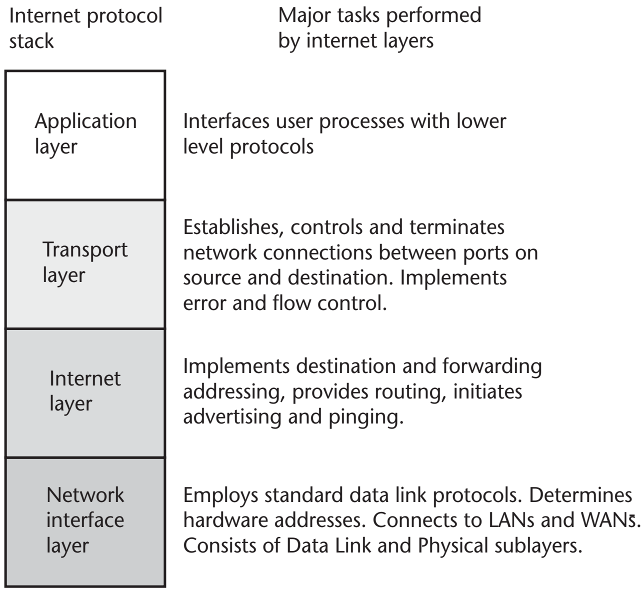
The TCP/IP suite continues to evolve as new applications develop. TCP/IP has displaced many successful alternative protocol suites to become the suite of choice for digital communication. When 200 million machines all use the same procedures, it is difficult to maintain that another set of protocols is better. Truly, the fact that TCP/IP powers this vast array of computing machines is credential enough to claim that it unites the world.

1.1.2 Internet Protocol Stack

Protocols are applied in sequence to the user's data to create a frame that can be transmitted from the sending application to the receiving application. The receiver reverses the procedure to obtain the original user's data and pass them to the receiving application. To formalize the sequential nature of employing the protocols, we construct a stack. As shown in Figure 1.1, for the Internet the stack has four layers. The top layer is the application layer. It contains the application processes that generate and manipulate data and request communication support from the lower layers. The next layer is the transport layer. It contains UDP and TCP. They initiate connectionless transport or initiate and terminate connection-oriented transport with error control and flow control. The transport layer *protocol data unit* (PDU) contains identifying numbers for the ports through which the application layer communicates with the transport layer. The next layer is the Internet layer. It contains IP and other associated protocols. They provide the frame with originating and terminating addresses to guide the PDU to its destination. The bottom layer is the network interface layer. It employs standard data link protocols and converts the data stream to a signal stream for transmission over physical facilities to the destination stack. Here, the frame is handed off from layer to layer in reverse. The bottom layer passes the PDU to the Internet layer, the Internet layer passes the PDU to the transport layer, and the transport layer passes it to the application that can use the data being delivered. In doing this, each receiving layer makes use of the information added by its corresponding sending layer. A further description of the Internet stack can be found in Chapter 2. My purpose here is to set the stage for discussion of some application layer protocols and the protocols that make up TCP/IP.

1.2 Some Application Layer Protocols

At the application layer, the user may generate information at a keyboard, or an application may generate a file. Either way, these actions make use of supporting programs to achieve certain outcomes. The more common of these programs are as follows.



Internet protocol stack	Major tasks performed by internet layers
Application layer	Interfaces user processes with lower level protocols
Transport layer	Establishes, controls and terminates network connections between ports on source and destination. Implements error and flow control.
Internet layer	Implements destination and forwarding addressing, provides routing, initiates advertising and pinging.
Network interface layer	Employs standard data link protocols. Determines hardware addresses. Connects to LANs and WANs. Consists of Data Link and Physical sublayers.

Figure 1.1 Internet Protocol stack.

1.2.1 Information Retrieval

Hypertext Transfer Protocol (HTTP) is a request/response protocol that transfers data between client computers and HTTP servers. HTTP translates digital streams into text and pictures for display on PCs.

Of the multitude of application protocols extant, HTTP finds almost universal application in support of information retrieval activities associated with pages from the World Wide Web. To retrieve information from an HTTP server, the client sends a request for a *resource* (an object or service provided by a server). The request contains a description of the action to be taken (e.g., *GET*, *PUT*, *DELETE*) and a description of the resource (uniform resource identifier) on which the action is performed. The uniform resource identifier is a standard way of describing a resource to a server. It includes two items: *uniform resource locator* (URL) and *uniform resource name* (URN). A resource is requested by location or name and may include resource-specific information. In response, the HTTP server returns the data requested.

1.2.2 File Transfer

File Transfer Protocol (FTP) is a protocol used to share and transfer files between clients and servers and to use servers for remote storage or other purposes.

Another procedure for data transfer, FTP can establish connections between server and server, as well as between client and server. FTP sessions consist of two separate connections. A *control* connection is used to negotiate communication parameters and control and monitor the status of any data connection opened between the parties. A separate duplex *data* connection is opened to transfer data between them.

File transfer is initiated by commands issued by the user *protocol interpreter* (PI) over the command channel. The user-PI initiates a control connection from a client port to the server process. The server-PI listens for user-PI connections, listens for user-PI commands, controls the server responses, and controls the server data transfer process. A user can initiate data transfer between two servers by establishing control connections with each and issuing commands that cause them to open a data connection between themselves.

1.2.3 Mail Transfer

Simple Mail Transfer Protocol (SMTP) is a procedure that facilitates the transfer of electronic mail between hosts. SMTP provides message transfer. It does not manage mailboxes or mail systems.

SMTP provides reliable, efficient processes for the transfer of electronic mail. It transfers messages between clients and servers and between servers. Communication is initiated by the user's mail system, establishing a duplex connection to an SMTP server. When the channel is established, the client informs the SMTP receiver that it wishes to send mail. The client issues one or more commands that identify the recipient(s) of the forthcoming message. The SMTP server establishes a duplex connection to the final destination. The client notifies the server of its intention to send mail and proceeds to send the message data. If the mail transfer is successful, the server issues a receipt and the client closes the channel.

1.2.4 Using Another Computer

TELNET is a remote terminal protocol that allows a user to log on to another host elsewhere on Internet. TELNET establishes a duplex connection using TCP/IP and passes the user's keystrokes directly to the target machine.

1.2.5 Resolving Names and Numbers

Domain Name System (DNS) is a process that maps host names and IP address numbers and provides one given the other (i.e., resolves names into numbers and numbers into names). It maintains a distributed database.

Keeping track of numerical addresses is easy for clients and servers, but, as the number of addresses grows, becomes more difficult for people. Accordingly, two addressing systems are employed. One, a routable number system, is used among machines. The other, a user-friendly name system, is used between people and machines. To ensure the infallible operation of DNS, both name and number must be *globally* unique. In principle, because each component of the name may be up to 63 characters long, finding unique names is not an issue. However, assigning unique numerical addresses is more difficult. Two numbering versions exist. One (IPv4) uses 32-bit addressing, and the other (IPv6) uses 128-bit addressing. IPv4 and IPv6 addresses are discussed later in this chapter.

Common *generic top-level domain* (gTLD) names are three-letter extensions that divide name addresses by establishment type. Two-letter extensions are used to divide names by geographical locations. Some of the establishment type extensions are:

- .com commercial organization;
- .edu educational institution;
- .gov agency of the U.S. government;
- .int organization established by international treaty;
- .mil U.S. military organization;
- .net network provider;
- .org nongovernment or nonprofit organization.

Some of the geographic location extensions are:

- .au Australia;
- .it Italy;
- .jp Japan;
- .uk Great Britain.

Extensions can have more than three letters, and many more extensions have been proposed to the *Internet Corporation for Assigned Names and Numbers* (ICANN). ICANN is responsible for coordinating the assignment of globally unique identifiers to Internet users.

Beneath these gTLDs the names are narrowed down until they stand for a single entity. Thus, my e-mail address used to be *bcarne@monad.net*. It has three parts. The first part is *.net*, indicating that a network provider [e.g., an *Internet Service*

Provider (ISP)] collected my e-mail. The next part was *monad*, signifying Monadnet Corporation (my ISP, based in Keene, New Hampshire, now part of Prexar Corporation, based in Bangor, Maine). The third part was my e-mail name, *bcarne*. As noted above, my e-mail name can be up to 63 characters long, leaving plenty of room for invention. The three parts together were my *universal resource name* (URN), a unique name that was easy to remember. If someone wished to send me e-mail, that person entered my URN from his or her PC. His or her SMTP program contacted a domain name server that related my URN to the address of my ISP. Then SMTP had a network address with which to route the e-mail!

1.3 User Datagram Protocol

Below the application layer is the transport layer. It contains two protocols, UDP and TCP. UDP is a simple transport layer protocol for applications that do not require reliable delivery service. When sending, UDP accepts data from the application layer, adds port numbers to guide delivery, computes a checksum to be used at the receiver to check the validity of the source and destination addresses, and sends the combination to IP. When receiving, UDP reverses these actions.

1.3.1 UDP Attributes

Commonly used for short data messages UDP provides *connectionless* service, that is, messages are sent without negotiating a connection. They carry no sequence numbers, and their receipt goes unacknowledged. UDP datagrams do not provide information on buffer storage available at the receiver or sender, are not segmented, and do not provide flow control information. Despite this list of negative attributes, the low overhead makes UDP datagrams ideal carriers for short messages, such as requests, answers, and repetitive announcements, sent to single locations using IP *unicast* addresses. In addition, UDP is used whenever data is sent to multiple locations using IP *multicast* or *broadcast* addresses. Because it has few internal controls to provide discipline, UDP is known as a laissez-faire protocol.

1.3.2 UDP Header

Figure 1.2 shows a UDP frame in which the application PDU is encapsulated by a UDP header to create a UDP PDU. The header carries the number of the source port (to identify the application creating the application PDU), the number of the destination port (to identify the application to which the PDU is sent), the length of the UDP PDU in bytes (to assist the receiver to size and process the payload data), and a checksum (to verify the integrity of the datagram at the receiver). A complete listing of the UDP header is found in Appendix B.

Port numbers 0 through 1,023 are assigned by IANA for common use and port numbers 1,024 and above by the application for specific uses. Called *well-known UDP port numbers*, some of those assigned by IANA are:

- UDP 53 Domain Name System;
- UDP 67 Dynamic Host Configuration Protocol (DHCP) Client;

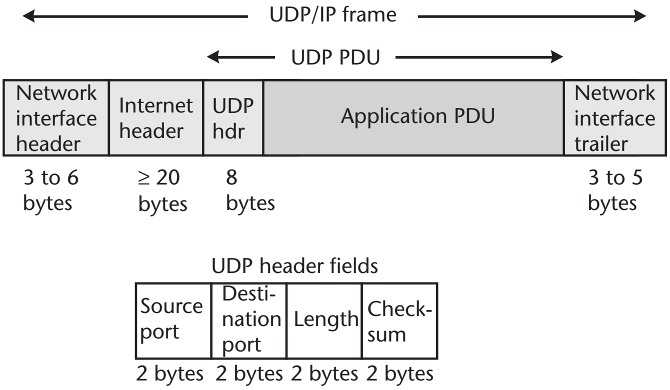


Figure 1.2 UDP header and UDP/IP frame.

- UDP 68 Dynamic Host Configuration Protocol (DHCP) Server;
- UDP 69 Trivial File Transfer Protocol (TFTP);
- UDP 137 NetBIOS Name Service;
- UDP 138 NetBIOS Datagram Service.
- UDP 161 Simple Network Management Protocol (SNMP)

By identifying the port number through which the application PDU reaches UDP in the transport layer, the application is providing an address for the return of data.

1.3.3 Checksum

The checksum is calculated by summing 16-bit words over the UDP datagram (header + payload) and a *pseudoheader*. It consists of the source IP address, the destination IP address, an unused byte, a byte that identifies the UDP protocol (0x11), and the length (in bytes) of the segment. In addition, if the number of bytes in this stream is odd, a padding byte is added. (For computation only. The padding byte is not transmitted.) Repeating the addresses (they are also contained in the Internet header) ensures that, if a routing or segmentation process modifies the values in the IP header, it is detected in the transport layer.

In more detail, the sender adds the 16-bit words in the segment and computes the *ones complement* of the sum. This is the number put in the checksum field and sent to the receiver. The receiver sums the 16-bit words *and* the ones complement. If the result is all ones, no errors have been detected. If the result contains one or more zeros, an error or errors are present. In this circumstance, the datagram is destroyed.

1.4 Transmission Control Protocol (TCP)

TCP provides *connection-oriented* services. A logical connection is set up between originating and terminating stations. Acknowledgments, error and flow controls, and other features are employed to ensure reliable data transfer. TCP is a transport layer protocol that provides reliable data transfer over point-to-point duplex channels. TCP accepts data from the application layer, adds data required to achieve reliable operation, and sends the combination to IP. TCP associates port numbers with specific applications, provides a number for every byte in the data stream, provides acknowledgments, computes timeouts to ensure the repetition of unacknowledged frames, exercises flow control, and uses special messages to establish and terminate duplex communication.

TCP is used with unicast addresses only. It cannot be used for multicast or broadcast deliveries. Before data is transferred between processes running on two hosts, a duplex connection is negotiated. At the end of the exchange, the connection is closed using a termination process. Provisions are made for recovery from untoward events.

1.4.1 Sequencing

To ensure reliable delivery service, the sender and receiver track data sent over a TCP connection. The first byte of a segment is assigned a number taken at random from 0 through 65,535, the range of numbers contained in a 2-byte field. Subsequent bytes are numbered from this number. Data streams in both directions are sequenced and positive acknowledgments are given. If an error is detected, the receiver requests retransmission from the last error-free frame. If no acknowledgment is received, the sender retransmits the segment. At the receiver, duplicate segments are discarded and out-of-sequence segments are placed in the proper order. Checksums are used to verify bit-level integrity.

1.4.2 Segmentation

To fit the application PDU within the IP datagram sent over the network interface layer link, the application PDU might be broken into segments by TCP. The sender and receiver exchange information on the maximum size segment that each can handle and adjust buffers accordingly.

1.4.3 TCP Header

Figure 1.3 shows a TCP frame in which the application PDU is encapsulated by a TCP header to create a TCP PDU. Considerably more complicated than UDP, the header contains entries necessary for the sender and receiver to establish a connection and implement reliable delivery. A complete listing of the TCP header can be found in Appendix B.

1.4.4 TCP Ports

As with UDP, the port number defines a location through which an application layer process sends a data segment to a TCP process or to which a TCP process delivers a data segment for an application layer process. Care must be taken to distinguish between UDP and TCP ports. UDP supports connectionless services. TCP supports connection-oriented services. The 1,024 numbers (0 through 1,023) are assigned by IANA. Examples are:

- TCP 20 FTP Server (data channel);
- TCP 21 FTP Server (control channel);

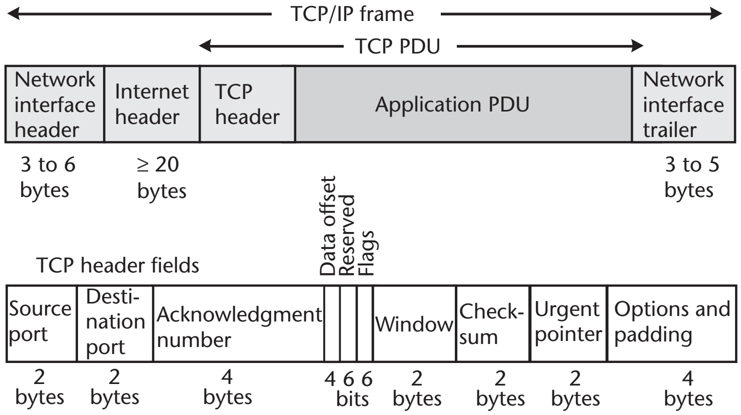


Figure 1.3 TCP header and TCP/IP frame.

- TCP 23 Telnet Server;
- TCP 25 Simple Mail Transfer Protocol (SMTP);
- TCP 80 Hypertext Transfer Protocol (HTTP);
- TCP 137 NetBIOS Session Service.

As required, numbers 1,024 and above are dynamically allocated by application processes.

1.4.5 Checksum

The checksum is calculated by summing 16-bit words over a pseudoheader, the TCP header, and the payload. The pseudoheader contains the source IP address, the destination IP address, a TCP identifier code (0x06), and the length (in bytes) of the segment. Repeating the IP addresses confirms that a routing or segmentation process has not modified these essential fields in the IP header. In addition, if the number of bytes in this stream is odd, a padding byte is added. As with UDP, the sender adds the 16-bit words in the segment and computes the ones complement of the sum. This is the number put in the checksum field and sent to the receiver. The receiver sums the 16-bit words *and* the ones complement. If the result is all ones, no errors have been detected. If the result contains one, or more, zeros, an error or errors are present. In this circumstance, the segment is destroyed.

1.4.6 Urgent Data

Under some circumstances, the data stream must be interrupted by control data. Setting the URG flag, using the urgent pointer field, and including the urgent data at the beginning of the TCP data segment accomplish this. The urgent pointer field records the number of bytes from the beginning of the TCP header to the last byte of urgent data in the payload.

1.4.7 Cumulative Acknowledgments

To achieve reliable data transfer, TCP employs *cumulative* or *selective* acknowledgments for TCP segments received. When using cumulative acknowledgments, the number in the TCP header acknowledgment field is the number of the first byte of the frame the receiver next expects to receive. Its presence explicitly acknowledges error-free receipt of all bytes up to, but not including, this byte. If a frame is received with errors, it is discarded. The receiver continues to hold the number of the first byte of the errored frame as the acknowledgment number signaling the sender to repeat the frame. When a frame is lost, it goes unacknowledged and is retransmitted after a while (see Section 1.4.10). In the cumulative acknowledgment environment, the acknowledgment number is one more than the number of the last byte of the frame that it has received without an error. It stays that way until the next frame is received perfectly.

1.4.8 Selective Acknowledgments

When using selective acknowledgments, TCP acknowledges bytes to either side of a missing or errored frame so that the sender need only repeat defective frames.

1.4.9 Flow Control

Flow control is a procedure for controlling the rate of transfer of packets between the sender and receiver so that packets are not lost due to congestion at critical points along the path or overwhelm the receiver.

Satisfactory communication requires that the receiver receives the entire message just as the sender sent it. For this to happen, the sending and receiving hosts, and the intermediate nodes, must cooperate to transport the data stream at an appropriate speed. It should not be so fast that packets can find no room in the buffers along the way and are lost to the system; it should not be too slow so that transmission takes longer than necessary. Flow control requires traffic measurements to be made, results to be fed to the receiver, controls to be invoked, and perhaps instructions sent to the sender. To do this, sequence numbers must identify the packets so that they can be tracked.

Receiver-side flow control is the process of actions taken by the receiver so that the incoming byte stream does not overload the receiver's buffer storage. As a first step in flow control, the receiver tells the sender the size of the receive buffer allocated to the exchange. In response, the sender tells the receiver the size of the message segment that it will send (segment size is less than buffer allocated). Data flow is adjusted to make maximum use of the facilities available. When possible, the receiver will increase the buffer to receive longer segments. Whenever acknowledgments are received, the sender is informed of the size of this *window*.

Sender-side flow control is the process in which, in response to guidance from the receiver, actions are taken by the sender to send the byte stream without causing congestion. At intermediate nodes packets are received, checked, and may be modified. They are held in buffer storage while tests are run, routes are found, and other traffic is processed. Should the sender send too quickly, or should there be an overwhelming amount of other traffic, the buffers fill, and there may be nowhere for the packets to wait for processing. As a result, they are lost from the system. Congestion information is passed downstream from sender to receiver. The receiver controls congestion relief. It increases the size of the receive window (buffer) and/or commands the sender to decrease the number or length of the segments it sends. In extreme cases, it may command the sender to stop sending until the congestion clears.

Changing traffic loads from other senders may affect some of the intermediate nodes. They pass congestion status information along to the receiver. In addition, the sender may send special packets to probe conditions along the path. The receiver returns these packets to the sender. On the basis of this information, the sender may reduce the transmission unit size so that the intermediate nodes can make buffer capacity available to other circuits. In other situations, the intermediate nodes may destroy packets that have been sent in excess of the rate that the network owner has guaranteed to the user. Flow control requires constant monitoring by all the nodes in the network and frequent instructions to the senders to slow down or speed up to accommodate changing conditions.

1.4.10 Retransmission Time-Out

In TCP, all segments containing data must be acknowledged. For each connection, TCP maintains a variable whose value is the amount of time within which an ACK is expected for the segment just sent. Called the *retransmission time-out* (RTO), if the sender does not receive an ACK by the time RTO expires, the segment is retransmitted. To prevent needless repetitions, RTO must be greater than the *round-trip time* (RTT) for the connection. Since the RTT is likely to vary with traffic conditions, it must be monitored continually, and the RTO adjusted accordingly.

For frames containing data, TCP uses an exponential backoff algorithm to determine the RTO of successive retransmissions. Initially, when the TCP segment is sent, the RTO is set to the value currently known for the connection (RTO1). If the retransmission timer expires without an acknowledgment, the segment is resent and the RTO timer is set to 2^n RTO1 (where $n = 0, 1, 2, \ldots$). This step is repeated until a maximum number of retransmissions are reached. At that time the connection is abandoned.

Segments that contain no data (e.g., ACKs) are not acknowledged. The sender does not set an RTO for a *data-less* segment. Thus, it does not retransmit lost data-less segments. To recover a lost ACK, the sender retransmits the segment(s) that the ACK would have acknowledged. When assembling the data stream on the basis of their sequence numbers, the receiver discards duplicate packets.

1.5 Creating a Connection

TCP employs a duplex logical circuit to implement communication between application processes running on two hosts. Each endpoint is identified by the combination of host IP address and TCP port number. The circuit is identified by the endpoints in each host (i.e., IP address 1 + TCP port 1, and IP address 2 + TCP port 2).

To create a connection, the hosts must exchange information and negotiate parameters. The three steps involved are shown in Figure 1.4. The hosts:

- Must learn the number of the first byte of data that will be sent to them. With it they can locate each field and send acknowledgments using numbers recog-

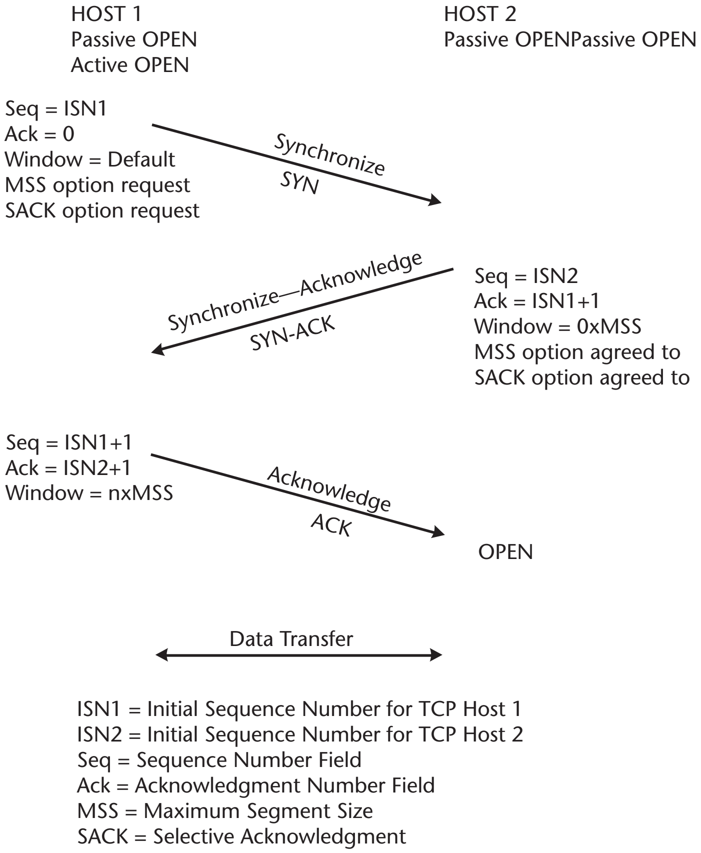


Figure 1.4 TCP connection establishment procedure.

nized by the sender. To achieve this, each must provide the other with its *initial sequence number* (ISN).

- Must determine the size of the buffer memory the other will provide for the receipt of their PDUs so that they do not send too much data at a time (and lose it).
- Must negotiate the maximum size of the segments they exchange so that communication will be as intense as possible.
- May negotiate options to satisfy specialized objectives.

1.5.1 OPEN Function Calls

To create a connection, the sending application issues an *active* OPEN function call that opens a message queue (port) from the application to the transport layer. Using the fields in the TCP header, the source and destination port numbers are entered. The *initial sequence number for Host 1* (ISN1) is placed in the sequence number field. The number 0 (because there is no exchange to acknowledge) is placed in the acknowledgment number field. As an opening move, Host 1 informs Host 2 that Host 1's receiving window is set at its default level. In addition, options may be negotiated such as varying the *maximum segment size* (MSS) depending on traffic conditions, and using a *selective acknowledgment procedure* (SACK).

Connection establishment will succeed only if the potential application in the receiver is in a *listening* mode (i.e., capable of receiving the connection request message that passes up the protocol stack to the proper port). To do this, applications issue *passive OPEN function calls* to specific port numbers or to ranges of port numbers. (This action may be part of the system start-up procedure.) If a connection is to be made, the process *must* be listening for incoming connection requests. If it is not *listening*, the connection cannot be made.

1.5.2 Flags

In the initial exchange, the sending host (Host 1) sets the *synchronize* (SYN) flag to inform the receiving host (Host 2) that Host 1 wishes to synchronize counting the forward data stream and establish other parameters. In reply, Host 2 responds with a TCP header in which both synchronize (SYN) and *acknowledge* (ACK) flags are set. The sequence number field contains the *initial sequence number for Host 2* (ISN2). The acknowledgment number field contains an acknowledgment number of ISN1 + 1, meaning Host 2 has received the frame numbered ISN1 without detecting an error and is waiting for frame ISN1 + 1. In addition, Host 2 informs Host 1 that its receive window is set to $n \times$ MSS, adjusting n is acceptable, and selective acknowledgments can be used.

Host 1 completes the connection establishment procedure with a TCP header in which the ACK flag is set. It contains a sequence number of ISN1 + 1 (the next frame in the exchange), an acknowledgment number of ISN2 + 1 (acknowledging ISN2 and waiting for ISN2 + 1), and informs Host 2 that Host 1's receive window is set to $n \times$ MSS. With this message, Hosts 1 and 2 are synchronized and ready to exchange messages.

1.5.3 Connection Denied

Should Host 2 be unable to open a connection with Host 1, Host 2 replies with the *acknowledge–reset* message shown in Figure 1.5. Both ACK and RST flags are activated. The sequence number is set to 0 since there will be no data stream to follow. The acknowledgment number is set to ISN1 + 1 to acknowledge Host 1's original frame. The receive window is closed. Upon receipt of a message carrying an RST flag, the receiving host may try again to create the connection. After three failures, the attempt is likely to be abandoned. Setting the RST flag in the middle of an

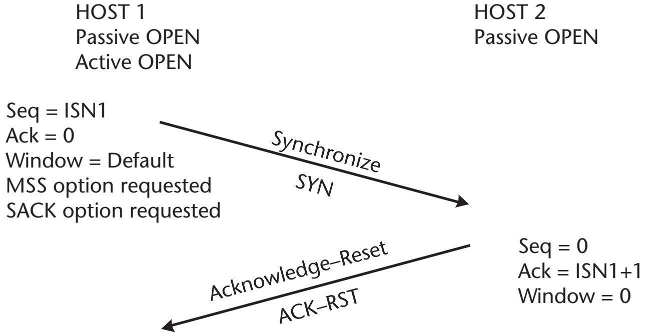
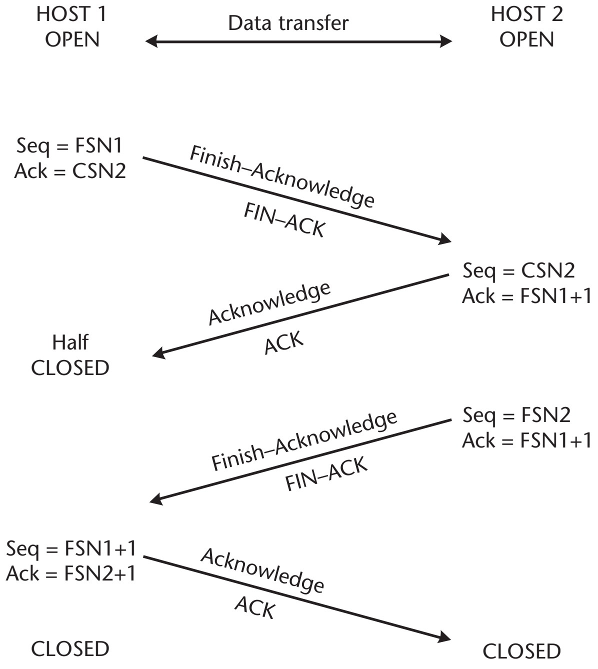


Figure 1.5 TCP connection reset procedure.



FSN1 = Final sequence number for TCP Host 1
FSN2 = Final sequence number for TCP Host 2
CSN2 = Current sequence number for Host 2

Figure 1.6 TCP Connection termination procedure.

exchange will cause the connection to be aborted. All data in transit, as well as all data in buffers waiting to be sent, is lost.

1.5.4 Connection Termination

Under normal circumstances, connection termination requires the exchange of the four messages shown in Figure 1.6. To terminate an exchange, Host 1 sends a *finish–acknowledge* message in which the ACK and FIN flags are set. The sequence number field carries the *final sequence number* (FSN1) and the acknowledgment number field carries the sequence number of the message about to be sent by Host 2 (CSN2, current sequence number). The connection is described as *half-closed*.

Assuming Host 2 has not finished its part of the data exchange and must keep its side of the connection open, it responds with a TCP header in which only the ACK flag is set. The sequence number is CSN2 and the acknowledgment number is FSN1 + 1. The header encapsulates the next segment of data from the application on Host 2. When Host 2 comes to the final data segment, it creates a *finish–acknowledge* frame. In the TCP header the FIN and ACK flags are set. The sequence number is the final sequence number (FSN2). The acknowledgment number field continues to carry FSN1 + 1. The header encapsulates the final data segment. Host 1 responds with an acknowledgment frame in which the ACK flag is set, the sequence number is FSN1 + 1, and the acknowledgment number is FSN2 + 1. The connection is closed.

1.6 Internet Protocol

The transport layer PDU (either UDP PDU or TCP PDU) is passed to the Internet layer where the *Internet Protocol* (IP) adds information necessary for routing the PDU from source to destination. IP makes a *best effort* to deliver packets to their final destination. It adds the addresses needed to route frames from source to destination and provides management and control facilities.

The combination of the transport layer PDU and the header added by the Internet layer is known as an *IP datagram*. Containing source and destination network addresses, the datagram provides connectionless, unreliable delivery service to the transport layer. When sending payloads larger than the *maximum transmission unit* (MTU) permitted by the transmission link, IP fragments the datagram. For instance, Ethernet limits the payload to approximately 1,500 bytes, and frame relay limits the payload to 8,189 bytes. When receiving, IP reassembles the fragments into a complete datagram.

1.6.1 IP Version 4

Two versions of IP are employed. The majority of users use Version 4 (IPv4). Version 6 (IPv6) was introduced in the mid-1990s to overcome a potential shortage of IPv4 addresses and update the header structure. Some government, university, and commercial organizations use it.

1.6.1.1 IPv4 Header

Figure 1.7 shows the fields of an IPv4 header. When no options are invoked, the header is 20-bytes long. When all options are invoked, it is 60 bytes long. Padding bytes are added at the end of the header to bring the total length to a multiple of 4 bytes. (The *header length* field is counted in 4-byte blocks.) Of note are:

- *Type of service (TOS) field:* This field indicates the quality of service with which the datagram is to be processed by the intermediate routers. Some rout-

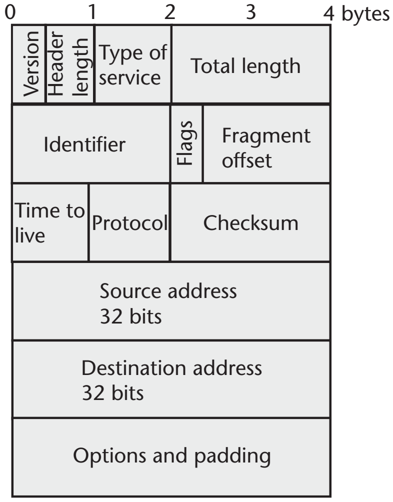


Figure 1.7 IPv4 header.

ing protocols calculate routes that optimize the values in the TOS field. Usually, the TOS byte is set to 0 × 00 by the sending host (i.e., normal precedence, delay, throughput, reliability, and cost).

- *Time to Live (TTL) field:* This field records the number of *hops* the datagram may make before being destroyed. A *hop* is the name given to the action of passing over a data link between contiguous nodes.

Each node handling the datagram reduces the TTL number by one. When TTL reaches zero, unless the node handling it is the destination, the datagram is destroyed. If the datagram is a broadcast message, TTL is set to 1 by the source. In this way, the datagram is restricted to the immediate network and is not forwarded. A complete listing of the IPv4 header is found in Appendix B.

1.6.1.2 IPv4 Addresses

In Version 4, IP addresses are 32 bits long. Divided into 4 bytes, they are written as four decimal numbers separated by dots; thus, 204.97.16.2 is an IP address. Writing the address in this fashion is known as *dotted decimal notation*. The numbers are the decimal equivalent of the binary codes in the bytes. In fact, the same address can be written in three ways; thus:

- Dotted decimal: 204.97.16.2;
- Binary: 11001100011000010001000000000010;
- Hexadecimal: 0×CC–61–10–02.

A *unicast* IP address is divided in two parts—network ID and host ID. The format is shown in Figure 1.8. All nodes on the same network share the same network ID. It employs bits at the left-end of the 4-byte address field. The host ID identifies a node on the network. It employs bits at the right-end of the 4-byte address field. Two addresses are reserved for special situations. *All 1s* is the address used by broadcast messages on the local network. *All 0s* is the address used by hosts on the

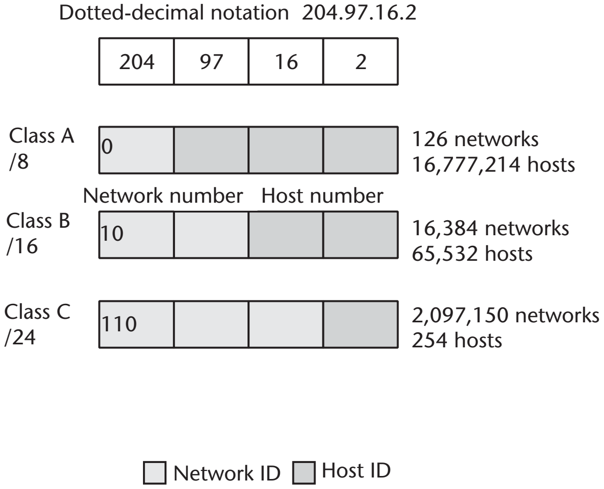


Figure 1.8 Classful addressing.

local network before they are assigned a unique ID. In addition, 127.x.y.z addresses are reserved for testing purposes.

1.6.1.3 Classful Addressing

In IPv4, the original approach to unicast addressing defined three classes for public use. Called *classful* addresses, they are:

- *Class A address:* An 8-bit network ID beginning with 0 and a 24-bit host ID.
- *Class B address:* A 16-bit network ID beginning with 10 and a 16-bit host ID.
- *Class C address:* A 24-bit network ID beginning with 110 and an 8-bit host ID.

The parameters of these address classes are given in Table 1.1.

As the network grew, the fixed address spaces of Classes A, B, and C, created difficulties in providing unique addresses. A solution that made the numbers more manageable is called *subnetting*. In it some of the bits that are reserved for host IDs are *robbed* to become parts of the network IDs. For instance, in a Class A address space, I can differentiate $2^7 - 2 = 126$ networks. If I take the four most significant bits from the first byte of the host ID field, I obtain an address space that differentiates $2^{11} - 2 = 2,046$ networks. Moving the boundary between the network ID and the host IDs has created 16 subnets for each Class A address and the original 7-bit identifier in the network ID byte can still address these subnets.

1.6.1.4 Subnet Mask

There is just one drawback. No longer is the boundary between the segments of the address fixed. How then is the processor to know how many bits in the 32-bit address space represent the network ID, and how many bits represent the host ID? A bit mask is used for this purpose. Called a *subnet* mask or an *address* mask, it contains 32 bits that are configured as follows:

- If the bit position in the mask corresponds to a bit in the network ID, it is set to 1.
- If the bit position in the mask corresponds to a bit in the host ID, it is set to 0.

By comparing the address and the subnet mask, the division between the network ID and the host ID can be found.

Table 1.1 Classful Address Parameters

	Class A or /8	*Class B or /16*	*Class C or /24*
Prefix	0	10	110
Number of addresses available	2^{31}	2^{30}	2^{29}
Number of bits in network ID	7	14	21
Number of network IDs	$2^7 - 2 = 126$	$2^{14} - 2 = 16,382$	$2^{21} - 2 = 2,097,150$
Range of network IDs	1.0.0.0–126.0.0.0	128.0.0.0–191.255.0.0	192.0.0.0–223.255.255.0
Number of bits in host ID	24	16	8
Number of host IDs	$2^{24} - 2 = 16,777,214$	$2^{16} - 2 = 65,534$	$2^8 - 2 = 254$
Range of host IDs	0.0.1–255.255.254	0.1–255.254	1–254

While subnetting made address distributions more efficient, for many applications the number of hosts required in each subnetwork can vary widely. The technique described earlier only produces equal size subnetworks. To establish networks with a varying complement of host IDs, subnetting was applied two or three times to subnetworks that already existed. To obtain sub-subnetworks with smaller numbers of host IDs, the technique of robbing right-hand bits from the host ID space was applied recursively. Each subnetwork, sub-subnetwork, and, perhaps, sub-sub-subnetwork, needed its own network mask. Because the intermediate network nodes must store routing information (IP addresses and subnet masks) for every subnetwork, subnetting began to overload the routing tables, particularly those in the backbone routers.

1.6.1.5 Supernetting

A solution to the overload problem has been found in *supernetting*. Supernetting starts with a group of Class C networks and builds upwards into the higher classes. The number of network IDs in the group must be a power of 2, and the group must have contiguous addresses. As the number of Class C address spaces bundled together increases through a power of two, the length of the subnet mask shortens by 1 bit. Hence, the requirement to bundle address spaces in powers of 2.

1.6.1.6 Classless Interdomain Routing

Using this technique, addressing is no longer associated with class structure. Class*less* addresses have replaced class*ful* addresses. Called *classless interdomain routing* (CIDR), the technique expresses a group of contiguous addresses as a single routing address by entering the lowest address of the group in the routing tables and noting the number of contiguous addresses in the group. As a result, the group of networks is addressed by a single entry. As long as the appropriate mask accompanies the CIDR block, the network ID for the CIDR block can be any number of bits. In addition, within the CIDR block, subnetting can be used to create subnetworks of convenient sizes. CIDR provides more flexibility in assigning addresses and improves the efficiency with which blocks of IDs can be addressed. It is the technique of choice for most networks.

1.6.1.7 Multicast Addresses

In addition to Class A, Class B, and Class C spaces for unicast addresses, Class D is defined for multicast addresses. The *Class D address* begins with 1110. The remaining 28 bits are used for individual IP multicast addresses ranging from 224.0.0.0 to 239.255.255.255.

An IP multicast address is a destination address associated with a group of hosts that receive the same frame(s) from a single source (one-to-many). Because routers forward IP multicast frames, the hosts can be located anywhere, and may join or leave the group at will. Managing multicast groups is the purpose of *Internet Group Management Protocol* (IGMP), described in Section 1.6.3.4. Addresses 224.0.0.0 through 224.0.0.255 are reserved for local use (same subnet traffic).

1.6.1.8 Private Addresses

Within an organization, the following *private* address spaces may be used:

- *10.0.0.0.* An address space with 24 host ID bits. Contains a single network. Host IDs range from 0.0.0 to 255.255.255.
- *172.16.0.0.* An address space with 20 host ID bits. Contains 16 network addresses that range from 172.16.0.0 through 172.31.0.0. Host IDs range from 0.0.0 through 15.255.255.
- *192.168.0.0.* An address space with 16 host ID bits. Contains 256 network addresses that range from 192.168.0.0 through 192.168.255.0.

Hosts with these private addresses are not reachable from the Internet, nor can they be connected directly to the Internet. Connections outside the organization's domain are made through a:

- *Network address translator:* This is a router that translates between private and public (Internet) addresses. In doing so, NAT must recalculate checksums. The *Source* and *Destination* addresses in the header are the network addresses of the source and destination hosts when inside the private network, or of the network address translators (NATs) serving them when in the public Internet.
- *Proxy server:* This is an application layer gateway that mediates between the private intranet and the public Internet.

These are discussed further in Chapter 6 (Section 6.2).

1.6.2 IP Version 6

The basic features of IPv6 have been available for about 10 years. Even though IPv6 can lead to improvements in operations, few users have adopted it. For one thing, the projected shortage of IPv4 addresses has not occurred in most of the Internet because of the introduction of CIDR. Also, full exploitation will require extensive changes to the backbone and existing equipment. Thus, while technology *push* is evident, market *pull* is not. Indeed, there is consumer resistance. Several strategies are being attempted to bring IPv6 into the Internet mainstream. Three of them are: create a separate IPv6 backbone; send IPv6 datagrams in IPv4 tunnels; and send IPv6 on dedicated data links. Each of them has had some success, but the *killer* application that will make IPv6 essential has yet to be discovered.

1.6.2.1 IPv6 Header

Figure 1.9 shows the fields in an IPv6 header. The most obvious change from IPv4 is the increase in size of the address space from 4 bytes (32 bits) to 16 bytes (128 bits). In addition, IPv6 eliminates some IPv4 fields that are little used and introduces eight extension headers that can be attached to provide significant flexibility. Among other things, the extensions provide routing information, fragmentation information, and path information. A complete description of the IPv6 header is found in Appendix B.

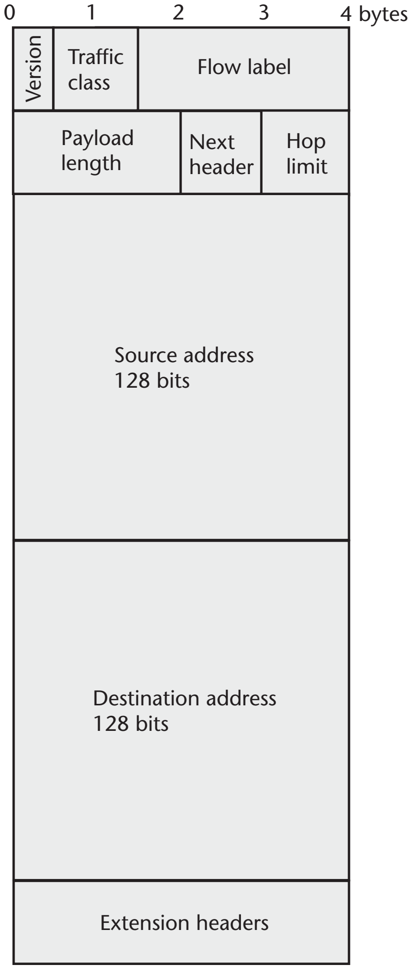


Figure 1.9 IPv6 header.

1.6.2.2 IPv6 Addresses

IPv6 addresses are 128 bits long. In the *preferred* text representation, they are written as eight 16-bit hexadecimal sections separated by colons. Thus, an IPv6 address for an *interface* might be 1234:0000:0000:CDEF:1234:0008:90AB:CDEF.

In this address block, fields containing leading zeros can be shortened. Thus, 1234:0:0:CDEF:1234:8:90AB:CDEF.

Further compression can be obtained by substituting :: for a string of zeros. However, this may be done only once in any address. Thus, 1234::CDEF:1234:8:90AB:CDEF.

In a mixed IPv4 and IPv6 environment, the six leftmost 16-bit sections are displayed in hexadecimal, and the remaining 32 bits are displayed in dotted decimal notation. Thus, 1234::CDEF:1234:8:144.171.205.239.

Portions of the address field may be used to identify special situations:

- *Format prefix*. A variable length field of leading bits that identifies the type of address. Some of them are:

    - Multicast address 11111111;
    - Aggregatable global unicast address 001;
    - Local-use unicast address 1111111010;
    - Site-local unicast address 1111111011.
- *Unspecified address*. 0:0:0:0:0:0:0:0 or :: cannot be used as a source address. Nodes in the initializing process use it before they learn their own addresses.
- *Loopback address*. 0:0:0:0:0:0:0:1 or ::1 is used by a node to send a packet to itself.
- *Aggregatable global unicast addresses*. Addresses organized into a three-tiered structure:
    - *Public topology*. Consists of 48 most significant bits that contain the format prefix (001) and the portion of address space managed by entities that provide public Internet services (45 bits).
    - *Site topology*. A second portion of the address space (16 bits) identifies an organization's internal routing paths.
    - The third portion of address space (64 bits) identifies individual interfaces on the organization's physical links.
- *Local-use unicast addresses*. Addresses used for communication over a single link. Examples are address autoconfiguration and neighbor discovery.
- *Multicast addresses*. A multicast address is assigned to a group of nodes. All nodes configured with the multicast address will receive frames sent to that address.

In principle, the increased information in the address blocks will make navigating the Internet easier and more reliable. However, the convenience comes at the expense of reworking and expanding routing tables throughout the networks, and requires a greater level of understanding of network opportunities.

1.6.3 Other Internet Layer Protocols

In addition to the transport layer protocols described earlier (i.e., UDP and TCP), IPv4 may carry other protocols (one at a time). Of major importance are *Internet Control Message Protocol* (ICMP), *Internet Group Management Protocol* (IGMP), *Address Resolution Protocol* (ARP), and *Inverse ARP* (InvARP).

1.6.3.1 Internet Control Message Protocol (ICMP)

ICMP reports errors and abnormal control conditions encountered by the first fragment of an IP datagram. There are no facilities within ICMP to provide sequencing or to request retransmission of IP datagrams. It is up to the transport layer to interpret the error and adjust operations accordingly. ICMP messages are not sent for problems encountered by ICMP error messages or for problems encountered by multicast and broadcast datagrams. An ICMP frame consists of a network interface header (whose format varies with the transmission facilities employed), an IP header, the ICMP header, a payload of ICMP message data, and a network interface trailer (variable format). A complete listing of an ICMP frame can be found in Appendix B.

1.6.3.2 Echo Request and Echo Reply Messages

Common uses for ICMP messages are determining the status and reachability of a specific node (known as *pinging*), and recording the path taken to reach it. The message sent to the node is called an *echo request* and the message returned is an *echo reply*. When the sender receives the echo reply message, the identifier, sequence number, and optional data fields are verified. If the fields are not correctly echoed, the echo reply is ignored. A listing of echo request and echo reply frames is found in Appendix B.

1.6.3.3 Destination Unreachable Messages

When a routing or delivery error occurs, a router, or the destination host, will discard the IP datagram and report the error by sending a destination unreachable message to the source IP address. To give the sender enough information to identify the datagram, the message includes the IP header and the first 8 bytes of the datagram payload. A listing of a destination unreachable frame is found in Appendix B.

1.6.3.4 Internet Group Management Protocol (IGMP)

A need for simultaneous data transfer to a number of nodes has created a demand for IP multicast traffic. Among many applications, the capability is required for audio and videoconferencing, distance learning, and television distribution. To achieve one-to-many delivery, IGMP sends a single datagram to local nodes and forwards it across routers to the distant nodes interested in receiving it. To implement this activity, IGMP provides a mechanism for hosts to register their interest in receiving IP multicast traffic sent to a specific group (multicast) address and to indicate they no longer want to receive IP multicast traffic sent to a specific group address, and for routers to query the membership of a single host group or all host groups.

1.6.3.5 Address Resolution Protocol

The IP address of a node must be converted to a hardware address before the transmission system can dispatch a message over the proper connections. This is the purpose of the *Address Resolution Protocol* (ARP) and its partner, the *Inverse Address Resolution Protocol* (InvARP).

1.6.3.6 ARP Request and Reply Messages

ARP is used to resolve the IP address of a node and its *medium access control* (MAC) address in a local area network (such as Ethernet, Token Ring, or FDDI). The resolved MAC address becomes the destination MAC address to which an IP datagram is delivered. Two messages are used:

- *ARP request message:* The forwarding node requests the MAC address corresponding to a specific forwarding IP address. The ARP request is a MAC-level broadcast frame that goes to all nodes on the physical subnetwork to which the interface requesting the address is attached.

- *ARP reply message:* The node whose IP address matches the IP address in the request message sends a reply that contains its hardware address. The reply message is a unicast frame sent to the hardware address of the requester.

A listing of ARP request and reply frames is found in Appendix B.

1.6.3.7 Gratuitous ARP and Duplicate IP Address Detection

A *gratuitous* ARP frame is an ARP request frame in which the *source protocol address* (SPA) and *target protocol address* (TPA) are set to the source's IP address. If no ARP reply frames are received, the node can assume its IP address is unique within its subnetwork. If an ARP reply is received, some other node on the subnetwork is also using the IP address and the node must obtain another address.

1.6.3.8 Inverse ARP (InvARP)

For *nonbroadcast multiple access* (NBMA)-based WAN technologies (X.25, frame relay, ATM), the network interface layer address is a virtual circuit identifier (not a MAC address). To determine the IP address of the interface at the other end, we use inverse ARP. For example, for *frame relay* (FR) connections, once the *data link connection identifiers* (DLCIs) are determined for the physical connection to an FR service provider, InvARP is used to build a table of DLCIs and corresponding IP addresses. InvARP request and InvARP reply frames have the same structure as ARP request and ARP reply frames. The operation field is set to 0×00–08 for InvARP request, and 0×00–09 for InvARP reply.

In both InvARP request and InvARP reply frames, the *sender hardware address* (SHA) is set to zero and the *target hardware address* (THA) is set to the DLCI value. The InvARP responder uses the InvARP request SHA to add an entry to its table consisting of the local DLCI and the SPA of the InvARP request. The InvARP requester uses the InvARP reply SPA to add an entry to its table consisting of the local DLCI and the SPA of the InvARP reply.

1.6.3.9 Proxy ARP

Proxy ARP facilitates answering ARP requests by a node other than the node whose IP address is carried in the request. In some circumstances, a subnetwork may be subdivided in two with the segments connected by a proxy ARP device. For each segment the proxy maintains a table of IP addresses and MAC addresses. Upon receiving an ARP request frame from a node on segment 1 for a node on segment 2, the proxy consults the table and replies with the appropriate MAC address. In addition, the proxy forwards unicast IP packets to the corresponding MAC address. This action saves time in filling routine requests.

1.6.3.10 Obtaining Configuration Information

Dynamic Host Configuration Protocol (DHCP) is a client-server protocol that manages client IP configurations and the assignment of IP configuration data.

Ensuring that networks are correctly configured at all times is an exacting task that is best left to an automatic process. For successful operation, all TCP/IP hosts must have a valid and unique IP address, a subnet mask, and the IP address of a default router/gateway. The IP addresses consist of network numbers and host numbers. Network numbers must be globally unique, that is, within the scope of the internetwork, individual networks must have unique identifiers. Host numbers must be unique within the group of hosts attached to a specific network. DHCP provides a service that dynamically allocates addresses and other information to clients as they require them.

1.7 Network Interface Layer

In order to be carried over a transmission link, network interface layer headers and trailers encapsulate the IP datagram to form an IP frame. They perform the following services:

- Indicate the start and end of the frames and distinguish the payloads from the headers and trailers.
- Identify the Internet layer protocol in use.
- Identify the hardware addresses of the source and destination nodes.
- Detect bit-level errors by use of checksums or frame check sequences.

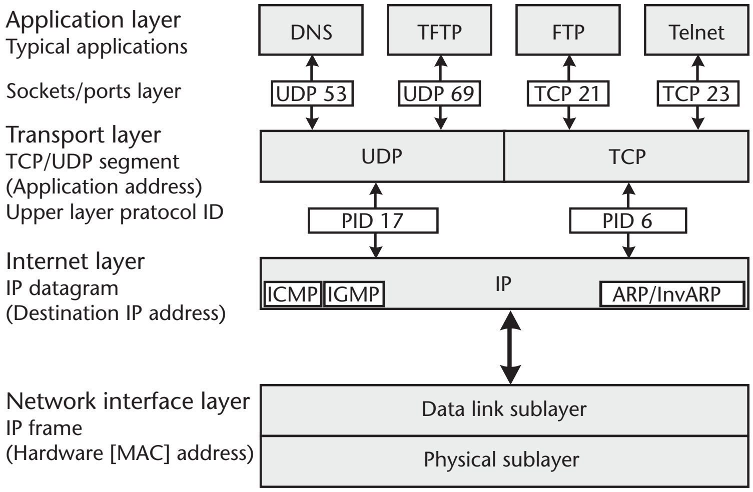
The formats of the network interface layer header and trailer depend on the type of network and the transmission equipment employed. They are addressed later in this book.

1.8 TCP/IP Protocol Stack

In this chapter, I have described the major features of the transport and Internet layers of the TCP/IP stack. The entire protocol stack is shown in Figure 1.10. Starting with some typical application layer protocols, it consists of a layer of sockets whose identification numbers (UDP ID or TCP ID) define the application for communication purposes and serve as access for any reply. They connect to UDP or TCP in the transport layer depending on whether connectionless or connection-oriented communication is to occur. At the Internet layer, the UDP or TCP segments are differentiated by separate *protocol identification numbers* (PIDs) and become IP datagrams. The Internet layer is the location for related messaging and administrative protocols (ICMP, IGMP, ARP, InvARP). From the Internet layer, the IP datagrams are passed to the network interface layer where they become IP frames.

Addresses are discovered and included at the network interface, Internet, and transport layers. The hardware or MAC address (defined and discussed in Chapters 3 and 4) is included in the frame at the network interface layer. The network or destination address is included in the IP datagram at the Internet layer. The socket number (or application address) is included in the segment at the transport layer. The diagram illustrates the basic functions needed to support data communication in a TCP/IP environment.

Finally, to avoid confusion, it is as well to repeat that IP forms datagrams. If UDP is employed as the transport layer protocol, the frame is forwarded through



DNS Domain name system
TFTP Trivial file transfer protocol
FTP File transfer protocol
Telnet terminal emulation
UDP User datagram protocol
TCP Transmission control protocol

IP Internet protocol
ICMP Internet control message protocol
IGMPInternet group management protocol
ARP Address resolution protocol
InvARP Inverse address resolution protocol

Figure 1.10 TCP/IP protocol stack.

the network on a best-effort basis without path control, no connection is established, acknowledgments are not given, and error and flow control are not used. If TCP is employed as the transport layer protocol, a duplex virtual circuit is established between sender and receiver before data transfer is initiated. With TCP able to communicate in both directions over an assigned connection, data streams can be synchronized, and acknowledgments, error control, and flow control can be employed. IP datagrams containing TCP PDUs are forwarded over the assigned channels.

CHAPTER 2

Data Communication

Data communication relies on functions performed in the terminals and equipment between originating and terminating locations. Many of these functions are implemented in software. However, with continuing improvements in the capabilities of integrated circuit chips, an increasing number of tasks at the bottom of the protocol stack are being implemented in hardware. Because they operate at *wire speeds*, processing is speeded up and response times are reduced. Nevertheless, whether realized in hardware or software, the TCP/IP suite governs the procedures involved, and the preferred format is an IP datagram.

2.1 Communication Equipment

Machines that implement data communication can be divided in three categories.

1. Those that provide an interface for users' instructions and graphical or textual outputs. Examples are:

   Terminal: A device used to input and display data. It may have native computing and data processing capabilities. A terminal relies on a host for support to accomplish the more intensive data processing tasks.

   Client: A terminal with significant computing and processing capability. A client acquires data from a server and accomplishes its tasks without outside support.

   Printer: Generally a device that provides *hard* copies of text or graphics with whatever processing power is required to produce fonts.

2. Those that process and store data. Examples are:

   Host: A host provides processing services and data support to terminals and may support clients when required. Early data processing systems were based on a mainframe computer (host) that supported many terminals (often characterized as *dumb* terminals).

   Server: A data processing device that stores data, organizes and maintains databases, and delivers copies of data files to clients, on demand. With the development of workstations and PCs, the client/server combination came into being to support central databases and make them available to *intelligent* terminals.

3. Those that facilitate the transport of frames across the network. Examples are:

Multiplexer: A device that causes several similar signals to be carried on a single physical bearer.

Repeater: A device that connects two circuits so as to extend the distance over which a signal is carried. Usually, the repeater regenerates, retimes, and reshapes the signal.

Bridge: A device that connects networks. It forwards messages between them based on a hardware address and a table of corresponding port numbers.

Router: A device that interconnects networks. It forwards messages between them based on the destination network address and a table of possible routes. Contemporary routers automatically update their knowledge of the paths available by periodically *advertising* their routing tables to one another. The path between sender and receiver is likely to contain numerous routers.

Switch: A device that selects paths or circuits so as to make real or virtual connections between sender and receiver.

Gateway: A device that interconnects networks that differ widely in performance, particularly above the network layer.

Many of these devices perform two functions. One is the processing function described earlier; the other makes the signals compatible with the transmission system in use. Conceptually, they can be divided into two parts.

- *Data terminal equipment (DTE):* The part that creates, sends, receives, and interprets data messages.
- *Data circuit-terminating equipment (DCE):* The part that assists the DTE to send or receive data messages over data circuits. DCEs *condition* (i.e., prepare) signals received from DTEs for transmission over communication connections and restore signals received from the network so as to be compatible with receiving DTEs.

These days, DTE and DCE are likely to be contained on the same network card.

Whether analog or digital signals are to be transported determines the type of DCE. If the signal is to be sent in analog form, the DCE is called a *modem*. When sending, a modem converts the binary signals received from the DTE to analog signals that match the passband of the line. When receiving, a modem converts the analog signals to binary signals and passes them on to the DTE.

If the signal is to be sent in digital form, the DCE has two components, a *data service unit* (DSU) and a *channel service unit* (CSU). The DSU/CSU performs the following functions.

When sending, the DSU/CSU:

- Converts the DTE signals to line code (namely, NRZI, 2B1Q, or other; see Appendix A).
- Inserts zeros suppression codes, idle channel codes, unassigned channel codes, and alarm codes. Zero suppression coding eliminates the possibility of too many consecutive zeros.

- When operating over T1 links, provides clear channel capability (64 kbit/s) on in-service channels by performing *binary eight zeros substitution* (B8ZS) coding or executing *zero-byte time slot interchange* (ZBTSI) (see Section 7.1.1).
- Supports superframe and extended superframe operations (see Section 7.1.1).

When receiving, the DSU/CSU:

- Converts NRZI, 2B1Q, or other signals, to a signal format compatible with the DTE.
- Removes the special codes inserted by the sending unit and notes the alarm information (if appropriate).
- Removes B8ZS coding or reconstructs ZBTSI frames.
- Supports superframe and extended superframe operations.

Most CSUs contain additional facilities that are used to detect and isolate line and equipment problems.

2.2 Making a Data Call

Consider a host (Host A) in a multilocation company that needs a data file to complete a task. The sequence of events could be as follows:

1. The application running on Host A generates the request: *Get xxxx*.
2. After polling the appropriate storage areas, the operating system (OS-A) finds no file of that name and sends a message to the operator: *File xxxx missing*. (For the sake of the story I have made the messages between the machinery and the operator understandable to the reader.)
3. After researching the matter, the operator determines the missing file is on Host B in another location. Moreover, on Host B, the file is called *yyyy*.
4. Guarding against the possibility that *yyyy* may be on Host A, the operator performs a search of Host A for *File yyyy*. It is not successful.
5. The operator makes the request: *Connect to Host B*.
6. With the help of a directory (or by other means), OS-A determines the network address of Host B is *A.b.C.d*.
7. OS-A instructs the communications processor (CP-A): *Connect to A.b.C.d*.
8. With the help of a table, CP-A determines that a private line connects directly to *A.b.C.d*.
9. CP-A opens a management file to supervise the communication session (exchange of messages) and allocates buffer memory to effect speed changing between the faster internal host circuits and the slower external communication circuits.
10. CP-A sends a *Request to Send* message to *A.b.C.d*. The request to send message includes the identity of Host A and a password.

11. CP-B consults the list of hosts from which it is permitted to accept messages. Host A and the password match an entry.
12. CP-B opens a management file to supervise the communication session and allocates buffer memory to effect speed changing in Host B.
13. CP-B sends a *Ready to Receive* message to CP-A.
14. CP-A notifies the operator that the connection is ready.
15. The operator *logs on* to Host B with a password and sends the request: *Get yyyy*. The request may include the size of the buffer allocated to receive *yyyy* and the maximum speed at which it can be received.
16. CP-B consults a list of valid users, or by other means determines that it may respond to the request.
17. CP-B requests *File yyyy* from its operating system (OS-B).
18. OS-B transfers a copy of the file to the control of CP-B.
19. CP-B conditions the file and segments it to be compatible with the communication facilities.
20. CP-B begins to send packets containing file segments to CP-A.
21. CP-A receives the packets, strips off header and trailer material, checks for errors, and begins to reassemble the file.
22. CP-A requests CP-B to re-send corrupted packets.
23. In their management files, CP-A and CP-B keep track of requests for resend to know which have been resent successfully.
24. CP-B sends the final packet and makes sure all resend requests have been honored.
25. CP-A reassembles the complete file and acknowledges error-free receipt to CP-B.
26. CP-A and CP-B terminate the connection.
27. The operator renames the file *xxxx*, formats it to suit Host A, and transfers it to the application.
28. The application completes its task.

By no means do these steps represent more than a skeleton of the communication procedure. For one thing, the scenario assumes a direct connection between the two hosts. When communication must take place across several networks, the task is significantly more complicated. However, the steps are enough to show that establishing, maintaining, and terminating data communications relies on logical routines executed in several units.

Communication procedures must promote conditions that support reliable communication, and, no matter how remote the possibility, guard against circumstances that could inhibit or degrade communication.

Satisfactory communication requires that the procedures cope with many situations. Examples are:

- *For the sender:* How is communication started? Does the sender establish a simplex channel or a duplex circuit to the receiver? Does the sender send when ready without regard to others on the network? Does the sender wait for a turn to send? How does the sender obtain permission to send? Is there a *handshake* between sender and receiver? How are data organized, and in what sequence are they sent? Does the sender repeat unacknowledged packets? How does the sender know how much data the receiver can handle? How does the sender make sure no user's data is interpreted as control data, and *vice versa*? How is communication terminated?

- *For the receiver:* Does the receiver acknowledge receipt of packets? Does the receiver report errors? How does the receiver determine the presence of errors? How does the receiver determine and keep track of the frame format? How does the receiver distinguish between control data and message data? How does the receiver notify the sender of congestion?

2.3 Open Systems Interconnection Model

The general problem of communication between cooperating dissimilar hosts situated on interconnected, but diverse, networks was studied by committees under the sponsorship of the *International Organization for Standardization* (ISO). Their work resulted in the *Open Systems Interconnection Reference Model* (OSI model, or OSIRM, for short). A model is a theoretical description of some aspect of the physical universe that identifies essential components and is amenable to analysis. Depending on the assumptions and approximations made, the subsequent results are more or less applicable to the real environment and may be extrapolated to similar situations.

2.3.1 OSI Model

As the name implies, the OSI model is designed to guide the development of *open* systems so that they can communicate with each other. Open systems are defined by the parameters of the interfaces between their functional blocks. Ideally, equipment from one vendor that implements a function will work with equipment from another vendor that implements the next function. To do this, the model does not define the equipment, only the states that must exist at their interfaces. It is the designers' problem to create equipment that satisfies these requirements. The model divides the actions of each host into seven independent activities that are performed in sequence. Figure 2.1 shows the activities arrayed in two stacks that represent the cooperating hosts. The seven layers contain *protocols* that implement the functions needed to ensure the satisfactory transfer of blocks of user's data between them. When sending, each layer accepts formatted data from the layer above, performs appropriate functions on it, adds information to the format, and passes it to the layer below. When receiving, each layer accepts formatted data from the layer below, performs some function on it, subtracts information from the format, and passes it to the layer above. Each layer shields the layer above from the details of the services performed by the layers below. Of the seven layers in the model, the top three (5, 6, and 7) focus on conditioning or restoring the user's data, and layers 1, 2, 3, and 4 implement data communication.

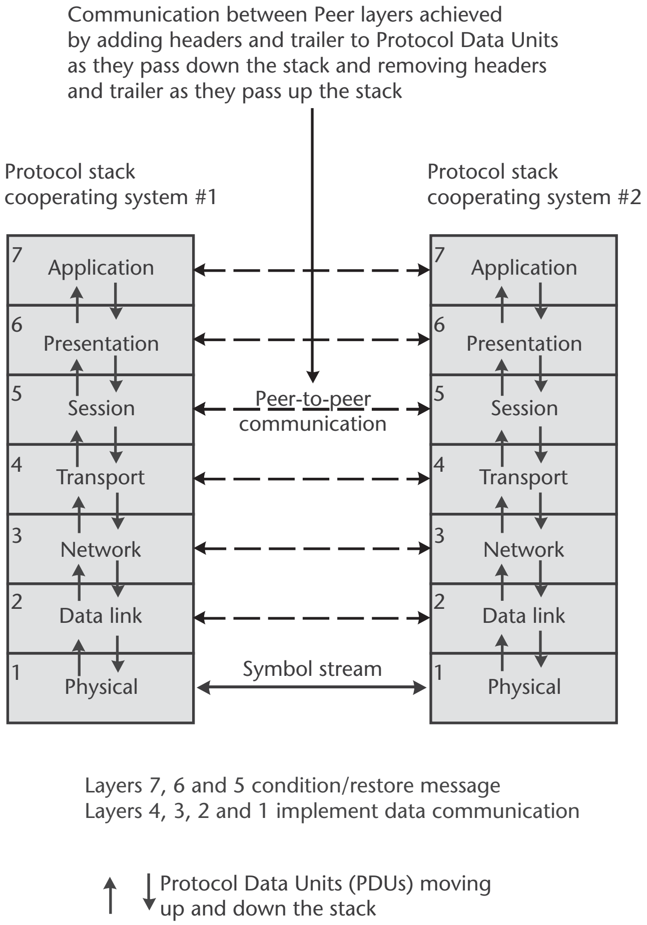

Communication between Peer layers achieved by adding headers and trailer to Protocol Data Units as they pass down the stack and removing headers and trailer as they pass up the stack
Protocol stack cooperating system #1
Protocol stack cooperating system #2
7 Application
6 Presentation
5 Session
4 Transport
3 Network
2 Data link
1 Physical
7 Application
6 Presentation
5 Session
4 Transport
3 Network
2 Data link
1 Physical
Peer-to-peer communication
Symbol stream
Layers 7, 6 and 5 condition/restore message
Layers 4, 3, 2 and 1 implement data communication
Protocol Data Units (PDUs) moving up and down the stack

Figure 2.1 OSI model of data communication between cooperating systems.

2.3.1.1 Input and Output

Users' data blocks enter the model at the application layer. In descending the protocol stack, each layer adds overhead data that manage the communication process. The extended data stream is converted to a sequence of signals that exits from the physical layer of one stack and crosses to the physical layer of the other stack on transmission facilities. There, the signals are converted back to a logical data stream that ascends the protocol stack towards the application layer of the receiving host. At each layer, the data sent by the peer layer in Stack 1 are removed and acted upon. Finally, the block of users' data emerges at the application layer of Stack 2.

2.3.1.2 Encapsulation and Decapsulation

In descending the protocol stack, the overhead data added at each layer is placed in a header, or, in the case of the data link layer, a header and trailer. This procedure is known as *encapsulation*, and the headers and trailer are said to *encapsulate* the user data. In ascending the protocol stack of the receiving system, the reverse procedure occurs; it is known as *decapsulation*, and the user data are said to be *decapsulated*. At each layer, the combination of data passed to the layer and the header (or header and trailer) added or subtracted in the layer is known as a *protocol data unit* (PDU). Figure 2.2 shows their development.

2.3.2 Layer Tasks

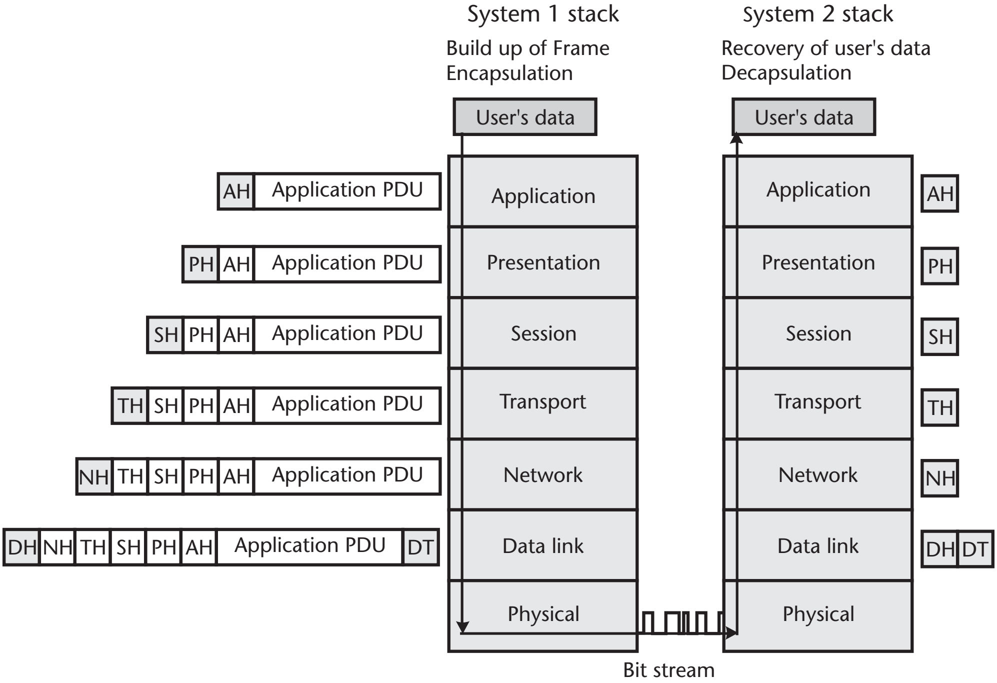
What do the protocols resident in the layers of these stacks do? Divided into those performed when sending, and those performed when receiving, the major tasks are listed in the following sections.

2.3.2.1 Application Layer

The application layer invokes generic applications (e.g., mail, file transfer, terminal emulation) in support of data generated by specific user applications. When *sending*, the application layer:

- Combines data received from the user's application with the appropriate generic function to create a user's data block.
- Encapsulates the user's data block with a header (application header, AH) that identifies this communication between specific user applications.
- Passes the *application protocol data unit* (APDU) to the presentation layer.

When *receiving*, the application layer:



AH Application Layer Header
PH Presentation Layer Header
SH Session Layer Header
TH Transport Layer Header
NH Network Layer Header
DH Data Link Layer Header
DT Data Link Layer Trailer
PDU Protocol Data Unit

Figure 2.2 Operation of the OSI model.

- Decapsulates the APDU (i.e., removes the application header from the APDU to leave the user's data block).
- Passes the user's data to the application identified by the header.

Peer-to-peer communication is required to agree upon the unique identifier for the communication. Usually it includes a port number and may include a sequence number. They are included in the application header.

2.3.2.2 Presentation Layer

The presentation layer conditions the APDU to compensate for differences in local data formats in the sender and receiver. When *sending*, the presentation layer:

- Performs translation services (e.g., code changing) and may perform data compression and encryption on the APDU.
- Encapsulates the APDU by adding a header (presentation header, PH) that identifies the specific coding, compression, and encryption employed.
- Passes the *presentation PDU* (PPDU) to the session layer.

When *receiving*, the presentation layer:

- Decapsulates the PPDU by removing the presentation header to leave the APDU;
- Performs any decoding, decompressing, and decrypting required.
- Passes the APDU to the application layer.

Peer-to-peer communication is required to agree upon coding, compression, and encryption algorithms. They are included in the presentation header.

2.3.2.3 Session Layer

The session layer directs the establishment, maintenance, and termination of the connection. It manages data transfer, including registration and password formalities, and may insert synchronization points into the information flow to facilitate restarting should a catastrophic failure occur. When *sending*, the session layer:

- Supervises the use of passwords and other checks.
- Tracks requests for retransmission and responses.
- Identifies the beginning and certifies the ending of the exchange.
- Encapsulates the PPDU by adding a header (session header, SH) that identifies any specific markers employed.
- Passes the *session PDU* (SPDU) to the transport layer.

When *receiving*, the session layer:

- Decapsulates the SPDU by removing the session header to leave the PPDU.
- Notes any specific markers.

- Passes the PPDU to the presentation layer.

Peer-to-peer communication is required to check authorizations and agree upon line discipline and the use of markers. They are functions included in the session header.

2.3.2.4 Transport Layer

The transport layer is the highest layer in the stack to be concerned with communication protocols. It ensures the integrity of end-to-end communication independent of the number of networks involved, and their performance. It is responsible for the sequenced delivery of the entire message, including error control, flow control, and quality of service requirements (if they are invoked). When *sending*, the transport layer:

- Establishes a connection-oriented duplex, or connectionless simplex, connection.
- Calculates a frame check sequence (FCS), or uses another technique, to facilitate checking the integrity of the SPDU at the receiver.
- Encapsulates the SPDU with a header (transport header, TH) to form the *transport PDU* (TPDU).
- Copies the TPDU for retransmission (if necessary).
- Passes the TPDU to the network layer.

When *receiving*, the transport layer:

- Decapsulates the TPDU by removing the transport header to form the SPDU.
- Verifies the FCS to confirm error-free reception.
- Acknowledges an error-free SPDU or discards it and may request a resend.
- May instruct the sender to modify the flow rate, if necessary.
- Passes the SPDU to the session layer.

Peer-to-peer communication is required to agree on the network(s) used for this communication, to replace corrupted frames, and to adjust data rates. This information is included in the transport header.

2.3.2.5 Network Layer

The network layer provides communications services to the transport layer. If necessary, it fragments the TPDU into packets to match the maximum frame limits of the network(s), and reassembles the packets to create the transport PDU. When *sending*, the network layer:

- Encapsulates the TPDU with a header (network header, NH) to form the *network PDU* (NPDU). The network header provides a destination address.
- May break the TPDU into packets to match the capabilities of the network(s).

- If the TPDU is segmented, encapsulates each segment with a network header to form an NPDU. The network header provides a destination address and a sequence number.
- Passes the network PDU(s) to the data link layer.

When *receiving*, the network layer:

- Removes the network header from the NPDU to form the TPDU.
- Verifies destination address and sequence number.
- Reassembles the TPDU, if necessary.
- Passes it to the transport layer.

Peer-to-peer communication is required to initiate, maintain and terminate the network level connection. These functions are performed by the network header.

2.3.2.6 Data Link Layer

The data link layer transfers data frames over a single communication link without intermediate nodes. When *sending*, the data link layer:

- Adds a header (DH) and a trailer (DT) to form the *data link PDU* (DPDU). The header includes a flag, class of frame identifier, sequence number, and hardware address of destination on the link. The trailer includes an FCS and a flag.
- Copies the frame in case retransmission is requested.
- Passes the frame to the physical layer.

When *receiving*, the data link layer:

- Reconstructs the DPDU from the bit stream received from the physical layer.
- Removes both header and trailer from the DPDU.
- Verifies FCS and other layer information.
- Discards the frame if the checks are not conclusive.
- Passes a correct NPDU on to the network layer.
- Requests resend, if necessary.

Peer-to-peer communication is required to agree on data link protocol parameters, error detection information, and error correction procedures. These are the functions of the data link header and trailer.

2.3.2.7 Physical Layer

The physical layer converts the logical symbol stream into the actual signal stream and completes the connection over which signals flow between the users. When *sending*, the physical layer:

- Converts the logical data stream to a suitable electrical signal, including signal conditioning (i.e., pulse shaping, zero stuffing, scrambling).
- Transmits a sequence of electrical symbols that represents the frame received from the data link layer.

When *receiving*, the physical layer:

- Receives a sequence of electrical signals.
- Interprets the signals as 1s and 0s.
- Deconditions the bit stream (i.e., unstuffs zeros, unscrambles).
- Passes a clean logical symbol stream to the data link layer.

Peer-to-peer communication consists of the signals that represent the total frame passed between Systems 1 and 2.

2.4 Internet Model

Contemporaneously with the development of the OSI model, the Advanced Research Projects Agency (now called DARPA, Defense Advanced Research Projects Agency) of the U.S. Department of Defense (DoD) was developing a data communication network. The objective was to enable the different networks and different computer systems deployed by organizations receiving ARPA funding to communicate. With time, ARPAnet became a four-layer model called the Internet, and the Internet has been adopted universally.

Figure 2.3 shows the approximate relationship between OSI and Internet models, and identifies the major tasks assigned to the four layers of the Internet model. Note that:

- The data link and the physical layers of the OSI model become the data link sublayer and the physical sublayer of the network interface layer of the Internet model.

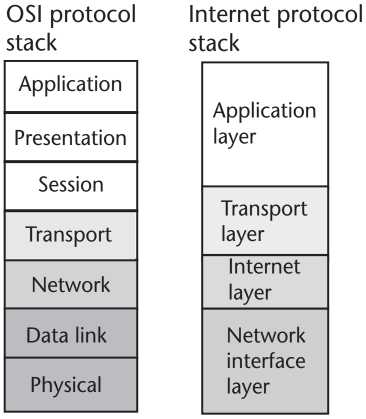


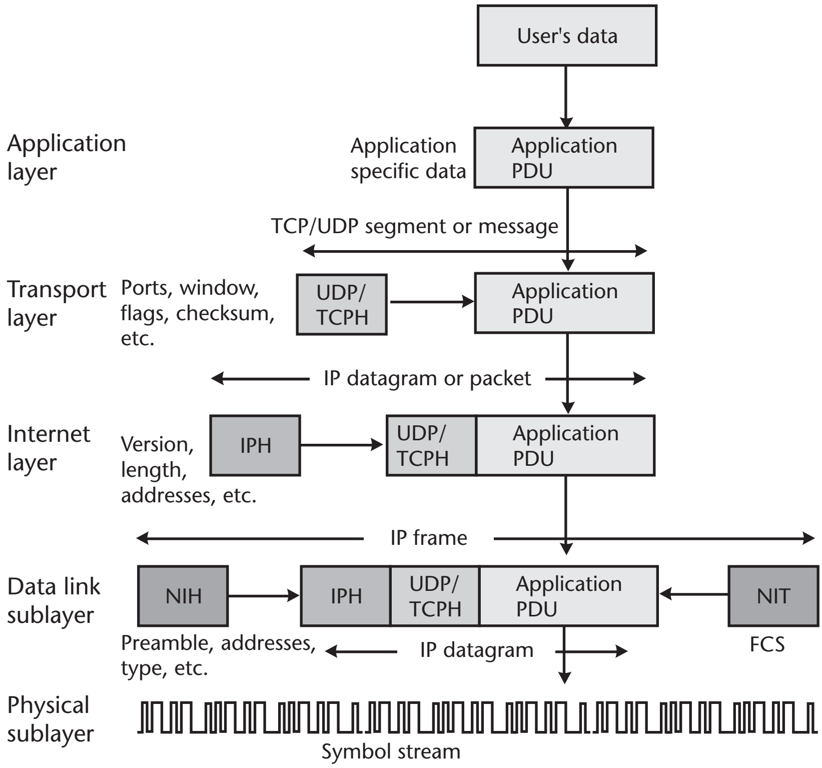
Figure 2.3 Comparison of OSI and Internet Protocol stacks.

- The network layer of the OSI model becomes the Internet layer of the Internet model.
- A portion of the session layer and the transport layer of the OSI model are combined in the transport layer of the Internet model.
- The application and presentation layers, and most of the session layer, of the OSI model, are combined in the application layer of the Internet model.

By no means is the mapping exact, nor can it be, because many common functions are implemented in different ways. Figure 2.4 shows the process of encapsulation from the application PDU to the signals of the physical sublayer for the Internet model. The major tasks performed by the protocols in the application, transport and Internet layers are listed in the following sections.

2.4.1 Application Layer

The application layer accepts user's data and combines it with software to achieve generic tasks such as information retrieval, file transfer, and mail transfer. When *sending*, the application layer:



PDU Protocol data unit
H Header
T Trailer
UDP User datagram protocol
TCP Transmission control protocol
IP Internet protocol
NI Network interface

Figure 2.4 Formation of IP datagram and IP frame.

- Combines user's data with generic function software to create a user's data block identified as information retrieval, file transfer, and mail.
- Encapsulates the user's data block with a header (application header, AH) and identifies the source port from which it is sent, and to which any reply must be addressed.
- Passes the *application protocol data unit* (APDU) to the transport layer.

When *receiving*, the application layer:

- Removes the application header from the APDU to leave the user's data block.
- Provides any processing required to complete the transaction.
- Passes the user's data to the user's application.
- Confirms that the process is completed.

2.4.2 Transport Layer

Two modes of operation are possible in the transport layer. The header may support a simple, connectionless procedure called *User Datagram Protocol* (UDP), or may support a connection-oriented procedure called *Transmission Control Protocol* (TCP). The transport layer PDU is called a *segment* or *message*. When *sending* in the *connectionless* mode, the transport layer:

- Accepts the APDU from the application layer.
- Records both source and destination ports.
- Calculates a checksum and transmits the ones complement.
- Encapsulates the APDU with a header (TH) containing this information.
- Passes the TPDU to the Internet layer.

When *receiving* in the *connectionless* mode, the transport layer:

- Accepts the TPDU from the network interface layer.
- Checks the length and confirms it matches the value contained in TH. If it does not agree, it discards the TPDU.
- Calculates a checksum and confirms it is all ones when added to the ones complement transmitted in the checksum field. If it is not, it discards the frame.
- Passes the APDU to the receiving port identified in the TPDU.

When *sending* in the *connection-oriented* mode, the transport layer:

- Establishes a duplex connection (real or virtual).
- Accepts the APDU from the application layer.
- Records source and destination ports.
- Provides the number of the first byte to be sent.
- Acknowledges receipt of previous frame (if any).

- Identifies size of storage allocated to this segment.
- Calculates a checksum and transmits the ones complement.
- Requests options such as selective acknowledgement, larger window size, and so forth from the destination.
- Encapsulates APDU with a header (TH) containing this information to form TPDU.

When *receiving* in the *connection-oriented* mode, the transport layer:

- Accepts the TPDU from the Internet layer.
- Identifies the receiving application on the basis of both sending and receiving ports.
- Synchronizes bytes with the sender on the basis of the sequence number received.
- Using the acknowledgement field, determines whether destination has received all bytes satisfactorily.
- Implements error and flow controls.
- Responds to flags to establish duplex connection.
- Notes window size of destination and any options requested by destination.
- Calculates a checksum and confirms it is all ones when added to the ones complement transmitted in the checksum field. If it is not, it discards the frame.
- Notes requests for options.
- Passes APDU to port designated for this application.

2.4.3 Internet Layer

The Internet layer supports a connectionless procedure called *Internet Protocol* (IP). The output of the layer is a *packet* called an IP datagram. When *sending*, the Internet layer:

- Accepts the TPDU from the network interface layer.
- Provides information on the version of IP in use and the lengths of the Internet header (IH) and IP datagram.
- Adds a quality of service level, if required.
- Fragments the datagram, if necessary.
- Adds time to live.
- Identifies the protocol in the TH of the TPDU.
- Calculates a checksum and transmits the ones complement.
- Adds source and destination IP addresses.
- Requests options such as record route, source routing, and time stamp.
- Encapsulates the TPDU with the Internet header to form the IPDU.

When *receiving*, the Internet layer:

- Accepts the IPDU from the network interface layer.
- Notes the version of IP in use.
- Uses header and datagram lengths to determine the start and the length of the data segment.
- Notes fragmentation (if any) and reassembles the TPDU.
- Decrements the time to live and discards the datagram if the value is zero.
- Calculates a checksum and confirms it is all ones when added to the ones complement transmitted in the checksum field and if it is not, discards the frame.
- Notes any requests for options.
- Passes the TPDU to the Internet layer.

2.4.4 Network Interface Layer

The network interface layer consists of two sublayers:

- In the *data link sublayer*, hardware addresses are discovered, conditions for access to the transport medium are accommodated, and a header and trailer are constructed. Added to the IP datagram, they form the IP *frame*.
- In the *physical sublayer* the logical data stream is converted to a signal stream to match the transmission facilities in use.

Local area networks, such as Ethernet, Token Ring, and Fiber Ring (FDDI), and wide area networks, such as packet, frame relay and *asynchronous transfer mode* (ATM), are served by extensions of the network interface layer. They are described in Chapters 3 and 4.

CHAPTER 3

Local Area Networks

Local area networks (LANs) interconnect data processing devices that serve communities of users. Operating within the network interface layer, they receive IP datagrams from the Internet layer and return them to it. Originally restricted to a limited geographical area, their reach has been extended to metropolitan areas by the availability of optical fibers. Furthermore, terminals have been freed to roam in airports and similar locations by the availability of radio (see Section 7.5).

Two styles of local area network are in use. One is known as *Ethernet* and the other as *Token Ring*. In their common form, both employ wire pairs. In addition, there is an optical fiber ring known as Fiber Distributed Data Interface (FDDI). Beginning with speeds in the lower megabit range, advanced LANs now operate in the lower gigabit range.

3.1 Ethernet

Conceived by Xerox Corporation as a shared medium data communication device that served a local community of users, *Ethernet* was developed by a team consisting of Xerox, Digital Equipment Corporation, and Intel Corporation. Later, the IEEE 802 committees added new features. I have chosen to call the original version *Classic Ethernet* to distinguish it from the IEEE 802.3 LAN that is universally called Ethernet. It is the most popular LAN in use today. Along the way, it has shed many of the original features to boost speed and throughput and make administration and reconfiguration easier.

3.1.1 Classic Ethernet

Figure 3.1 shows the concept of Classic Ethernet. It consists of a common coaxial cable bus to which all stations are connected. Operation is half-duplex. Only one station can transmit data at a time, and, when transmitting, it cannot receive. Each station monitors the activity on the bus to determine when to send frames.

3.1.1.1 Carrier Sense Multiple Access with Collision Detection

To provide access to the common channel, Classic Ethernet employed a procedure known as *carrier sense multiple access with collision detection* (CSMA/CD). When activity on the common channel ceases, in case the frame just sent is one of a series, the station with a frame to send waits for a time equal to the *Ethernet interframe gap*. The end of an Ethernet frame is not marked explicitly. Instead, a gap is left between frames that is equivalent to 96 bit times. The station then waits a further

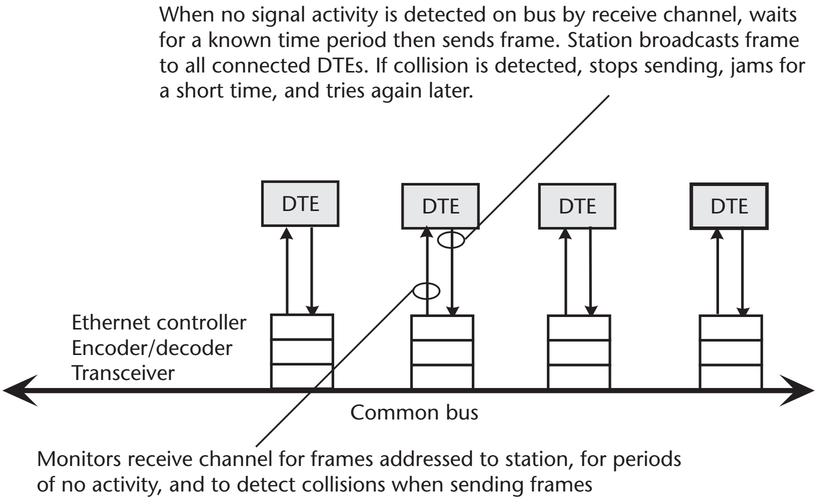

When no signal activity is detected on bus by receive channel, waits for a known time period then sends frame. Station broadcasts frame to all connected DTEs. If collision is detected, stops sending, jams for a short time, and tries again later.
DTE
DTE
DTE
DTE
Ethernet controller
Encoder/decoder
Transceiver
Common bus
Monitors receive channel for frames addressed to station, for periods of no activity, and to detect collisions when sending frames

Figure 3.1 Principle of Classic Ethernet LAN.

time period that is a random multiple of the slot time. [Slot time is the round-trip transmission time between a node at one end of the network and a node at the other end of the network. Usually, a slot time is assumed to be 512 bit times (i.e., 51.2 μsecs for a 10-Mbps LAN).] If there is still no activity, the station may send the frame. Once any station has begun transmission, other stations should detect the activity and withhold their own frames. If two, or more, stations begin to transmit at the same time, a collision will occur. They will detect they are interfering with each other, and will *jam* one another for a short time, so that all stations can hear that a collision has occurred. Then they cease transmitting. The jamming signal is 4-bytes long (usually 0×AA-AA-AA-AA). More precisely, a collision will occur if two stations begin transmissions within the time it takes signals to propagate from one to the other. For this reason, limits are placed on the distances separating terminals. On ceasing to send, the stations *back off* for a random number of slot times and try again. If the network is encountering heavy traffic, a collision may occur (with a different station) on the second attempt. The station will jam and back off again. After a number of unsuccessful attempts, the station will abandon the effort to send its message. Figure 3.2 provides a basic flowchart summary of CSMA/CD. Each terminal constantly monitors the state of activity on the LAN and follows the decision sequences on the chart.

3.1.1.2 Ethernet Frame Encapsulation

Internet Protocol (IP) datagrams and Address Resolution Protocol (ARP) messages sent over a Classic Ethernet network link are encapsulated as shown in Figure 3.3. Appendix B includes a listing of the fields in a Classic Ethernet frame.

In an *Ethernet header* the preamble serves to synchronize the receiver with the frame. The destination address follows. It may be unicast, multicast, or broadcast. The source address is a unicast address. These 6-byte addresses are assigned to the source and destination hardware at the time of manufacture. To complete the header, the EtherType field contains code that identifies the upper layer protocol in the payload.

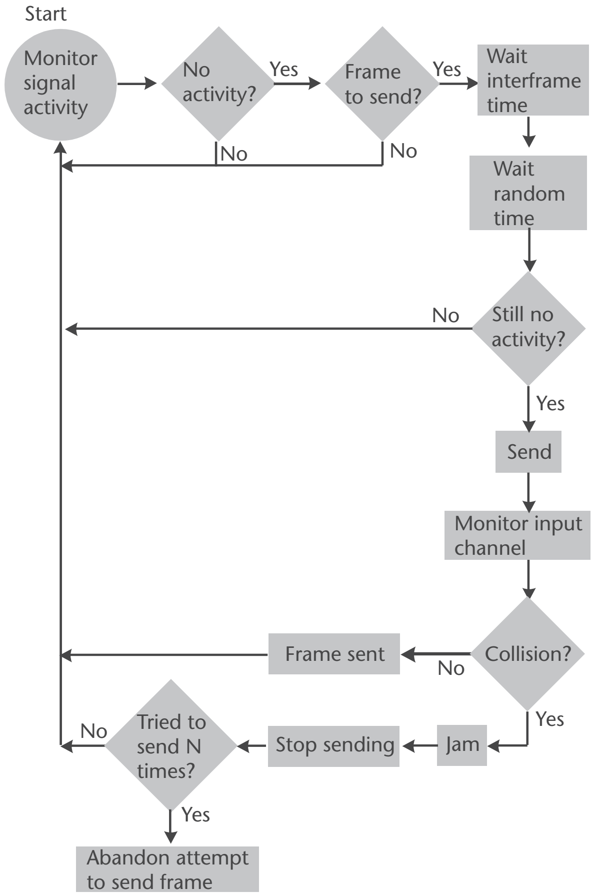


Figure 3.2 Principle of carrier sense multiple access with collision detection.

An *Ethernet trailer* consists of a 4-byte frame check sequence (FCS) generated by the source. Independently, the receiver calculates a FCS. If it agrees with the source FCS, it is highly likely that the frame has been received without error. If it does not agree, the receiver discards the frame.

3.1.2 IEEE 802.3 (Ethernet) LAN

The IEEE extended the performance of Classic Ethernet with respect to message handling. To do this, they added additional fields to the header.

3.1.2.1 LLC and MAC Sublayers

In the IEEE LAN model, layer #2 of the OSI model is divided into the logical link control (LLC) sublayer and the medium access control (MAC) sublayer. Figure 3.4 compares them with the data link and physical layers of the OSI model, and the network interface layer of the Internet layer. The functions of these sublayers are:

- *Logical link control (LLC) sublayer:* Defines the format and functions of the *protocol data unit* (PDU) passed between *service access points* (SAPs) in the source and destination stations. SAPs are ports within the sending or receiving

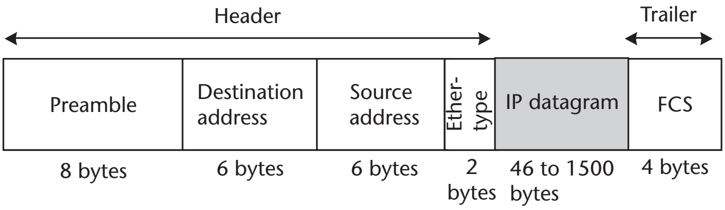
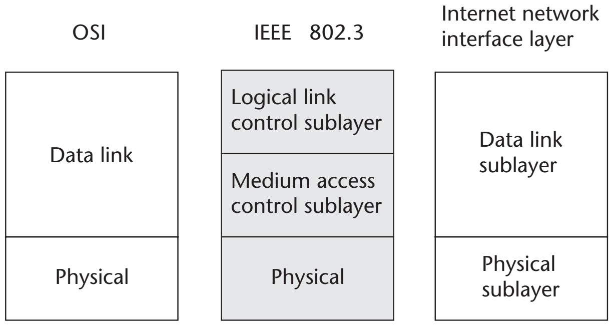


Figure 3.3 Classic Ethernet frame.


OSI
IEEE 802.3
Internet network interface layer
Data link
Physical
Logical link control sublayer
Medium access control sublayer
Physical
Data link sublayer
Physical sublayer

Logical Link Control Sublayer: defines format and functions of PDUs passed between SAPs (service access points) in source and destination

Medium Access Control Sublayer: defines format and functions of Headers and Trailers that are added to PDUs

Figure 3.4 Comparison of layers in OSI, IEEE 802.3, and Internet models.

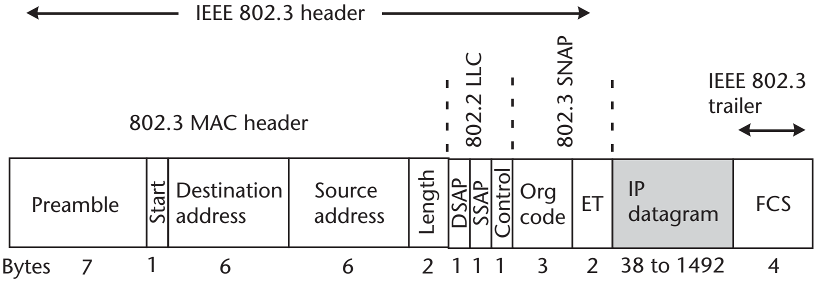
device that permit PDUs to flow to/from the upper level protocol agent identified by the EtherType entry. SAPs are associated with specific applications so that messages created by executing the applications can be identified and correlated. The LLC sublayer is standardized in IEEE 802.2.

- *Medium access control (MAC) sublayer:* Defines the format and functions of headers and trailers that encapsulate the PDUs. The MAC sublayer contains the hardware addresses of source and destination. The MAC sublayer is standardized in IEEE 802.3.

3.1.2.2 IEEE 802.3 Ethernet Frame

An IEEE 802.3 frame is shown in Figure 3.5 and listed in Appendix B. A comparison of Figures 3.3 and 3.5 shows that the simplicity of the Classic Ethernet header stands in strong contrast to the header of the IEEE 802.3 Ethernet LAN. The header consists of three sections.

- *IEEE 802.3 MAC header:* The combination of the preamble field and start delimiter is the same as the 8-byte preamble at the beginning of the Classic Ethernet frame. In the address fields, the two addresses must be the same length; they can be 2 or 6 bytes long. The former accommodates private network addresses generated locally. (Two-byte addresses are hardly ever used.) The latter accommodates the 6-byte hardware addresses assigned to equipment at



DSAP = Destination Service Access Point
SSAP = Source Service Access Point
ET = Ether Type
FCS = Frame Check Sequence

Figure 3.5 IEEE 802.3 Ethernet frame.

the time of manufacture. The length field indicates how many bytes are contained in the remaining two headers and the payload so that the receiver can detect the frame check sequence. The length will be less than 1,500 bytes (i.e., ≤0×05-DC). A value of ≤ 0×05-DC identifies the frame as an IEEE 802.3 Ethernet frame. A value ≥ 0×05-DC identifies the frame as a Classic Ethernet frame in which this field is EtherType. The lowest EtherType value is 0×06-00.

- *IEEE 802.2 LLC header:* The *destination and source SAP* (DSAP and SSAP) fields identify the points to which the payload is to be delivered in order to reach the proper upper-layer protocol. DSAP and SSAP act as upper-layer protocol identifiers. For IP, the value of both source and destination SAPs is 0×06. When used in conjunction with a SNAP header, DSAP and SSAP are set to 0×AA. This passes responsibility for identifying the upper-layer protocol to the SNAP header. The control field is 1 or 2 bytes long, depending on whether the LLC-encapsulated data is part of a connectionless communication (identified as Type 1) or a connection-oriented communication (identified as Type 2). IP datagrams and ARP messages are sent as Type 1.
- *IEEE 802.3 SNAP header:* The organization code field identifies the organization that maintains the meaning of the EtherType field that follows. For IP datagrams and ARP messages, the organization code is set to 0×00-00-00. The EtherType field is set to 0×08-00 for IP datagrams, and to 0×08-06 for ARP messages.

3.1.2.3 Subnetwork Access Protocol

IEEE 802.3 Subnetwork Access Protocol (SNAP) was created to permit protocols designed to operate with a Classic Ethernet header to be used in IEEE 802.3 applications. Messages sent over an IEEE 802.3 LAN use SNAP headers to identify the upper level protocols in use. The header contains a 3-byte organization code that identifies the organization responsible for defining the EtherType field that follows. For an IP datagram, or an ARP message, the organization code is set to 0×00-00-00. A 2-byte EtherType field that identifies the upper-layer protocol in use in the payload follows the Organization code. For an IP datagram, it is set to 0×08-00, and for an ARP message, it is set to 0×08-06. To keep the length ≤ 1,500 bytes, and accommodate the length of the extra headers (3 bytes for LLC and 5 bytes for SNAP), the payload is reduced by 8-bytes.

3.1.2.4 Additional Services

The additional information contained in the header permits three classes of services to be provided by IEEE 802.3 Ethernet. They are:

- *Connection-oriented service:* A logical connection is set up between originating and terminating stations. Acknowledgments, error and flow controls, and other features are employed to ensure reliable data transfer. For this reason, the IEEE 802.3 header contains internal logical connection points (SAPs) for both source and destination. They are used to ensure the source's frame(s) and the receiver's response(s) are delivered to the proper upper-layer protocols.
- *Acknowledged connectionless service:* The receiver acknowledges messages, but a logical connection is not established. This technique is used when the overhead (error control, flow control) associated with connection-oriented service would make the operation too slow, yet it is important to know that the message was received.
- *Unacknowledged connectionless service:* The receiver does not acknowledge messages. Error control and flow control are not employed. The service is used in applications where the occasional loss or corruption of a PDU can be corrected by procedures invoked by the upper layer communicating software entities.

In the source address and destination address fields of Classic Ethernet and IEEE 802.3 Ethernet frames, special bits are defined:

- The *Individual/Group* (I/G) bit (bit 1 in byte 0 of destination address) indicates whether the address is unicast (0) or multicast (1). For a broadcast address (which is a special case of multicast), the I/G bit is set to 1.
- The *universal (global)/local* (U/I) bit (bit 2 in byte 0 of destination and source addresses) indicates whether the address is globally unique (0) or locally administered (1). Globally unique addresses are controlled by IEEE and assigned to manufacturers to imprint during the manufacturing process.
- The *routing information indicator* bit (bit 1 in byte 0 of the source address) indicates whether Token Ring source routing information is present (1). Source routing allows a Token Ring sending node to discover and specify a route to the destination in a Token Ring segment.

3.1.3 New Configurations

Obviously, the throughput an Ethernet station achieves depends on the number of active stations and the speed of the bus. As the number of users increases, their average speed falls off, and the throughput of individual stations may become unacceptable. In addition, as the number of users grows, it is likely that the number of rearrangements that must be made to accommodate them increases. With a shared cable medium, this means constant splicing and rerouting as the cable is moved to include new, and/or eliminate old, stations.

In the early 1990s, technical improvements made it possible to connect the stations in a star configuration with twisted pairs. Pairs leading to a hub in a wiring closet replaced the shared cable. Now, changing connections on a wiring strip could add or delete stations. Later, a switch replaced the hub. The operation moved to 100 Mbps and 1,000 Mbps, and some connections use optical fibers.

Fast Ethernet products (i.e., those that operate at 100 and 1,000 Mbps) employ block coding. At 100 Mbps, the code is designated 4B/5B. Five bits substitute 4 bits in the data frame. Code patterns are selected so that the number of 1s and the number of 0s differ by no more than one. The signaling rate for 100 Mbps products is 125 Mbps. At 1,000 Mbps, the code is 8B/10B. Ten bits substitute 8 bits in the data frame. Code patterns are selected so that the number of 1s and the number of 0s differ by no more than two. The signaling rate for 1,000 Mbps products is 1,250 Mbps. More information can be found in Appendix A.

3.1.3.1 Ethernet Hub

The implementation of a common hub to which each station is attached by separate twisted pair cables, drastically modified the shared bearer approach to Ethernet. The *hub* is a combiner and a repeater. It may perform amplification, retiming, and reshaping in order to prepare the signal for retransmission. It provides a separate port for each attached station and creates the equivalent of a shared environment. It uses the same CSMA/CD algorithm to allocate the channel capacity to individual stations. Single repeaters provide from 8 to 24 ports. The combination of hub/repeater and attached stations is referred to as a *collision domain*. The repeater performs the following functions:

- Receives data from a transmitting station, restores the amplitude, timing, and shape of the received signal, and retransmits it on all ports except the port on which it was received.
- Detects simultaneous activity on two or more input ports and broadcasts a collision alert (jamming signal).
- *May* detect and disconnect stations that have failed in a continuous transmit mode (jabbering mode).

Figure 3.6 shows the principle of a *repeater* hub. Two pairs are used to connect each port to a single station. All stations must operate at the same data speed.

3.1.3.2 Switched Ethernet

The hub configuration suggests that the network might be modified to substitute a nonblocking, high-speed switch for the connection plane of the repeater hub. Then the two stations involved in a message transfer can be connected directly over a high-speed channel. Collisions are eliminated. CSMA/CD is no longer needed. Stations do not have to wait for the bus to be quiet, and they can operate at the full bit

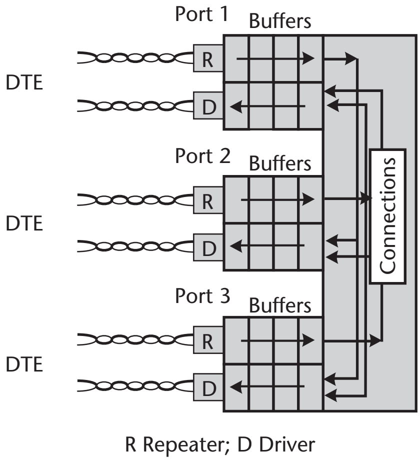


Figure 3.6 Principle of repeatered Ethernet hub.

rate of the switching fabric. Figure 3.7 shows the principle of a *switched* hub. Two methods of operation are employed:

- *Store-and-forward:* The entire frame is received and stored in the input buffer before being forwarded over a switch path to the buffer serving the port connected to the destination. In the process of storing the frame, the buffer logic may check for errors and perform other frame management functions.
- *Cut-through:* As soon as the destination address is received in the input buffer, the number of the output port is obtained from a table of ports and addresses. If a path through the switch to the designated port is available, the frame is fed to it. Should the port be busy with other traffic, the frame is stored in the input buffer to wait for the interfering traffic to clear.

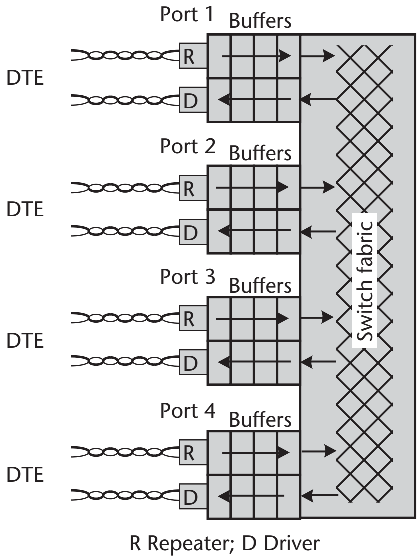


Figure 3.7 Principle of switched Ethernet hub.

For slower-speed operation (10 Mbps), the switch can be a crossbar. Crossbar switches have a plurality of horizontal and vertical paths and a means for interconnecting any one of the vertical paths with any of the horizontal paths. For higher-speed operation (100 Mbps or 1 Gbps) the switch can be a self-directing, high-speed switching fabric such as that used in *asynchronous transfer mode* (ATM) switches. The switches can be *blocking* (i.e., setting up an arbitrary switching path may not be possible because of an existing switching path) or *nonblocking* (i.e., an existing switching path cannot prevent the setting up of another switching path). Most switched Ethernets employ a nonblocking architecture.

Because the switch makes a direct connection from sender to receiver, it is possible to host 10 Mbps, 100 Mbps, and 1,000 Mbps stations on the same LAN. Of course, connections can only be made between stations operating at the same speed. This behavior is in direct contrast to a shared repeater hub on which all stations must operate at the same speed.

Switched hubs permit the linking of several shared LANs into a common data space without expanding their individual *collision domains*. Figure 3.8 shows the principle. Three repeater hub Ethernets are connected by a switched hub. Within each LAN, the stations employ CSMA/CD and are governed by the carrier sense, collision detect, backoff, and try-again rules. Between the LANs, frames are passed across the switch without hindrance. However, the switch ports must obey the CSMA/CD rules when moving frames back into a collision domain.

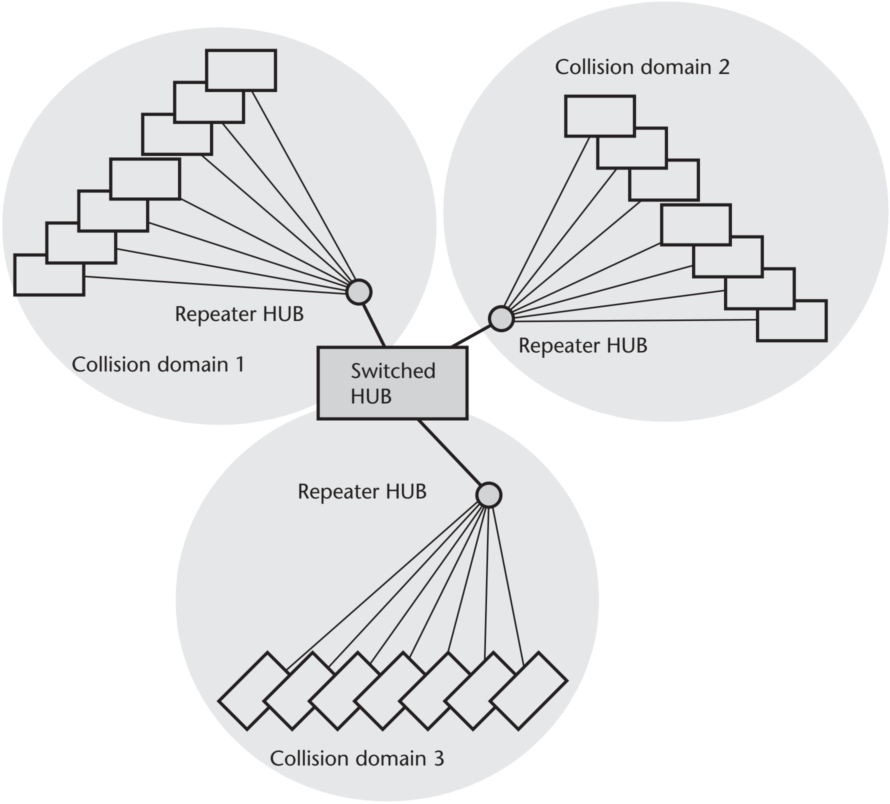


Figure 3.8 Use of switched hub to link Ethernets and separate collision domains.

3.1.3.3 Ethernet Designations

Different styles of Ethernet are identified as follows:

- *Bus connected:* In the designator, 10 = 10 Mbps speed; BASE = baseband signal; 5 = 500m; 2 = approximately 200m.
  - 10BASE5. 0.40-inch diameter coaxial cable bearer limited to segments of 500m and 100 nodes per segment when operating at 10 Mbps with Manchester signaling.
  - 10BASE2. 0.25-inch diameter coaxial cable bearer limited to segments of 185m and 30 nodes per segment when operating at 10 Mbps with Manchester signaling.
- *Hub connected:* In the designator, 10 = 10 Mbps speed; 100 = 100 Mbps speed; 1,000 = 1 Gbps speed; BASE = baseband signal; T = unshielded twisted pair; F = optical fiber. Some examples are:
  - *10BASE-T*. Operates at 10 Mbps. Employs two *unshielded twisted pairs* (UTPs) connected in a star. Each pair of UTPs supports a single station that is no more than 100m from the hub. Capable of full-duplex operation.
  - *10BASE-F*. Operates at 10 Mbps. Employs two multimode optical fibers to connect hubs separated by up to 2 kms. Fibers are run between the hubs. Each hub is connected to its community of users by UTPs. Capable of full-duplex operation.
  - *100BASE-TX*. Operates at 100 Mbps. Employs two Category 5 UTPs, or *shielded twisted pairs* (STPs) and two multimode optical fibers to interconnect hubs. Uses 4B/5B coding. Stations are limited to less than 100m from a hub. Capable of full-duplex operation.
  - *100BASE-FX*. Operates at 100 Mbps. Employs two multimode optical fibers to connect stations to hub. Uses 4B/5B coding. Fibers are limited to 2 kms. Capable of full-duplex operation.
  - *1000BASE-CX*. Operates at 1,000 Mbps. Employs two balanced copper cables. Uses 8B/10B coding. Stations are limited to 25m from hub. Capable of full-duplex operation.
  - *1000BASE-TX*. Operates at 1,000 Mbps. Employs four pairs of Category 5 UTP and multimode optical fibers to interconnect hubs. Uses 8B/10B coding. Stations are limited to 100m from hub. Capable of full-duplex operation.

3.2 IEEE 802.5 Token-Ring LAN

In a Token Ring LAN each station is connected to two others to form a single loop that connects all stations. Each station:

- Receives the data stream from the station preceding it on the ring;
- Regenerates it;
- May add to or change it;

- Sends the data stream to the next station on the physical ring.

The cabling system uses twisted-pairs with Manchester signaling. Data speeds of 4 Mbps, 16 Mbps, and 100 Mbps are in use. A *multistation access unit* (MAU) provides the ability to connect stations by UTP wiring to a central device in which the token ring is implemented. Figure 3.9 shows the concept. Furthermore, MAUs can be connected together in a ring so as to connect communities of stations. If the ring consists of dual cables (or fibers), it can be made *self-healing* by arranging for one of the cables/fibers to reverse itself to provide loopback in the event of a failure.

3.2.1 What Is a Token?

A token is an access control byte with start and end delimiters. The byte contains:

- Three priority bits (PPP), which identify the level of priority a station must have to seize the token.

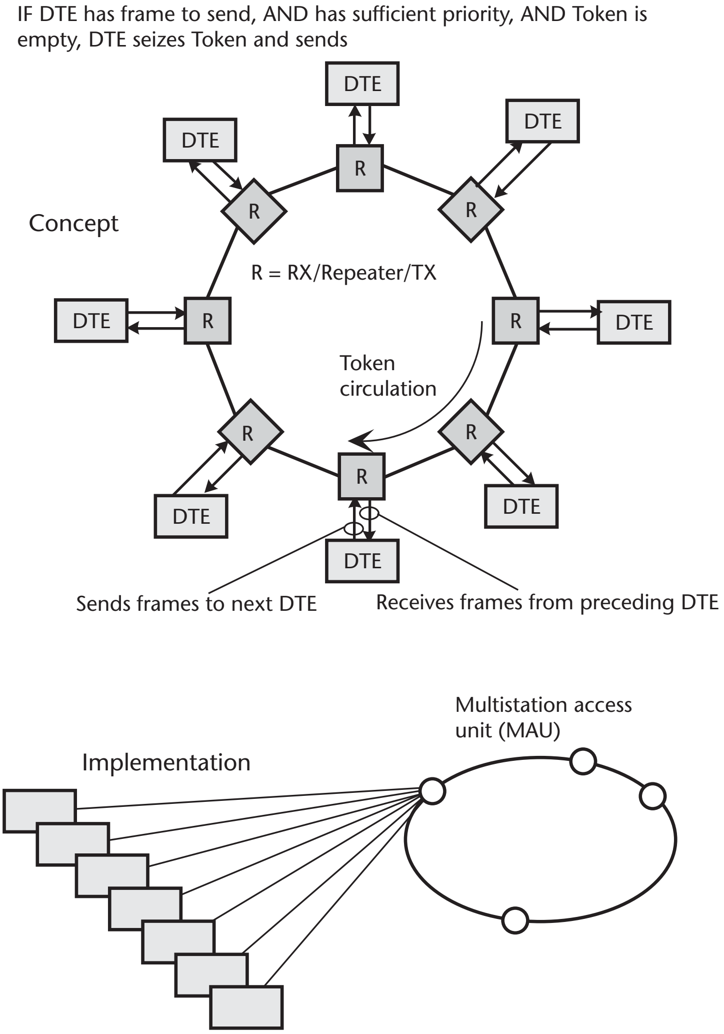

IF DTE has frame to send, AND has sufficient priority, AND Token is empty, DTE seizes Token and sends
Concept
DTE
R
R = RX/Repeater/TX
Token circulation
Sends frames to next DTE
Receives frames from preceding DTE
Implementation
Multistation access unit (MAU)

Figure 3.9 Principle of Token Ring LAN.

- A token bit (T), which gives the token status. If it is 0, the token has not been taken and a station that has sufficient priority may seize it. If it is 1, the token has been seized by another station and the frame is in use.
- A monitor bit (M), which is used to detect unclaimed frames.
- Three reservation bits (RRR), which provide a mechanism for lower priority devices to request the opportunity to transmit.

Figure 3.10 shows the sequence of activities associated with receiving a frame, determining whether the token is available, and influencing the availability of the token at some future time.

3.2.2 Token Ring Frame

Figure 3.11 shows a token and the fields in a frame containing an IP datagram. The frame consists of an IEEE 802.5 header, an IEEE 802.2 LLC header, an IEEE 802.3 SNAP header, the payload (IP datagram), and an IEEE 802.5 trailer. Appendix B includes a listing of the fields of an IEEE 802.5 Token Ring frame. They are summarized here:

- *IEEE 802.5 header:* The start delimiter field alerts the receiver to the incoming frame and provides a synchronizing signal. It contains two nondata symbols

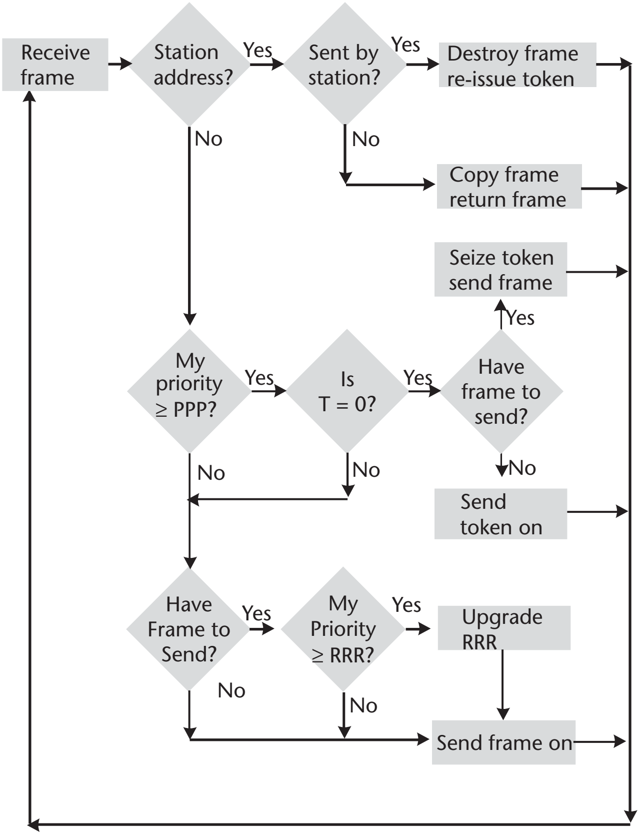


Figure 3.10 Major procedures in Token Ring LAN.

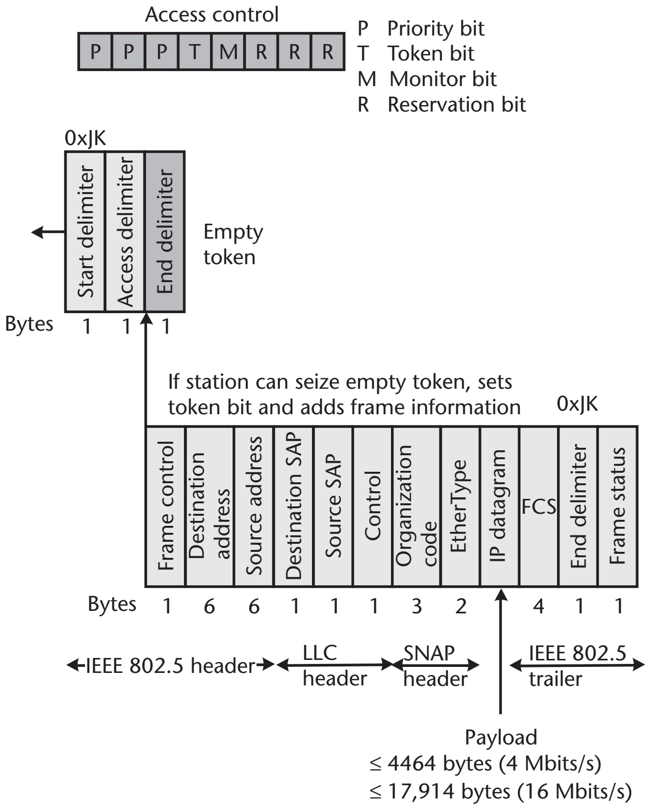


Figure 3.11 Token Ring frame.

(called J and K) that are violations of the signaling scheme. The J symbol is an encoding violation of a 1 and the K symbol is an encoding violation of a 0. The access control field is the key to token management and has been discussed above. The frame control field contains 2 bits reserved for future use and 6 active bits. They identify the frame that follows as a Token Ring MAC management frame or a Token Ring data frame. The address fields contain the unicast hardware addresses of the destination and source or multicast or broadcast addresses.

- *IEEE 802.2 LLC header:* For IP datagrams and ARP messages, the SNAP header preempts the LLC header. Accordingly, DSAP and SSAP are set to 0×AA, and the control field is set to 0×03. For other upper-layer protocols, the SNAP header may not be used. In this case, values that identify the points of origination and delivery of data to upper-layer protocols are present.
- *IEEE 802.3 SNAP header:* The organization code is set to 0×00-00-00 for IP datagrams and ARP messages. The EtherType code is set to 0×08-00 for IP datagrams and 0×08-06 for ARP messages.
- *IEEE 802.5 trailer:* The FCS is calculated over the data stream between the access control byte and the end of the payload. This allows the access control and frame status fields to be changed as needed to reflect operations without recalculating the FCS. The FCS is checked at each node. The end delimiter

contains J and K nondata symbols. In addition, it contains an intermediate frame indicator bit that identifies whether this frame is the last in a sequence (0), or there are more frames to follow (1). The end delimiter byte also contains an error detected indicator bit. Should the FCS fail, the node performing the check sets this bit and the destination node does not copy the frame. The frame status field contains duplicate address recognized indicator and frame copied indicator bits. They are used by the destination to inform the sender that the node recognized its address and successfully copied the frame. The bits are duplicated because the field is not included in the FCS.

3.3 Fiber Distributed Data Interface

Fiber distributed data interface (FDDI) employs a ring topology and uses a shared multimode fiber medium. Figure 3.12 shows the concept of FDDI. It can include a dual-fiber ring so that the system can recover from a single catastrophic fault. FDDI uses block coding (4B/5B). The signaling rate is 125 Mbps. A version of FDDI that works over wire pairs is available. It is limited to a maximum length of 100m. FDDI provides a relatively expensive solution to obtaining a local or metropolitan area network operating at 100 Mbps. It is being displaced by 100BaseTX and 1000BaseTX Ethernets.

Providing connectionless delivery using 48-bit addressing and token passing similar to IEEE 802.5 Token Ring, FDDI can be bridged to Ethernet. Standard protocol stacks communicate over FDDI in the same way they communicate over the Ethernet. Figure 3.13 shows an FDDI frame that encapsulates an IP datagram. Intentionally, it is very similar to frames for IEEE 802.3 and IEEE 802.5. Like them, when transporting IP datagrams and ARP messages, FDDI uses a SNAP header to identify the upper-layer protocol carried in the frame. The contents of the fields of an FDDI frame are listed in Appendix B.

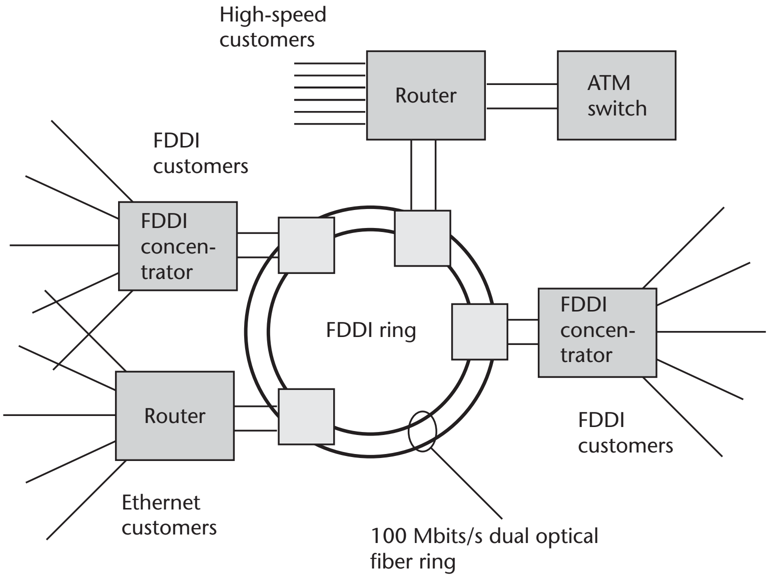


Figure 3.12 Principle of FDDI.

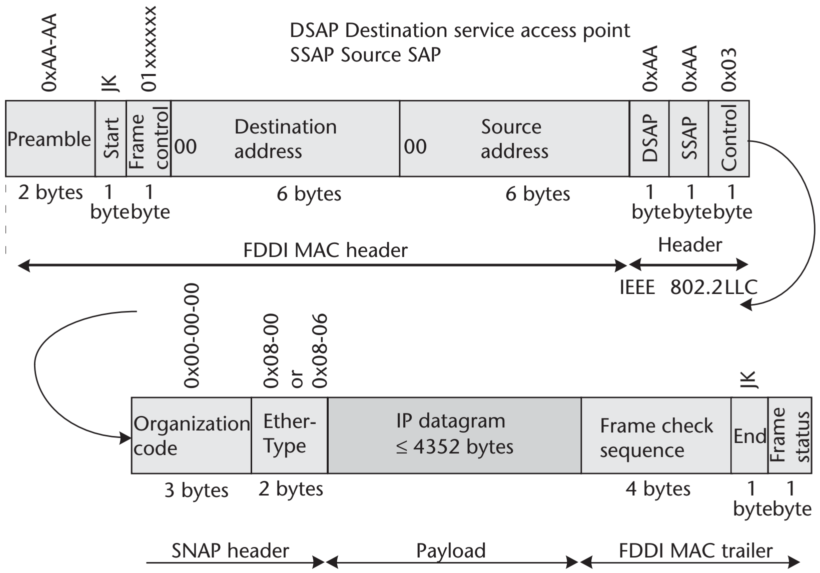


Figure 3.13 FDDI frame.

3.4 Bit Ordering

Ethernet uses little endian bit order and Token Ring/FDDI use big endian order. To make MAC address transmissions consistent between the two styles of LANs, Token Ring/FDDI systems store multibyte addresses in bit-reversed order compared to Ethernet. Figure 3.14 gives an example of the same 6-byte address stored in the Ethernet and the Token Ring/FDDI:

- In the Ethernet, the least significant bit in each byte occupies the rightmost bit position. Data streams are formed up beginning with the LSB. Bytes are taken in order from left to right.
- In the Token Ring/FDDI, the least significant address bit in each byte is stored in the rightmost bit position. Addresses are read out to data streams beginning with the rightmost bit in each byte. Bytes are taken in order from left to right.

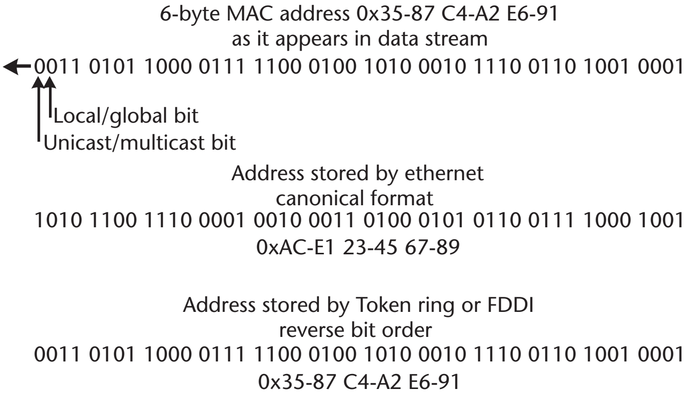


Figure 3.14 Difference in Ethernet and Token Ring/FDDI storage conventions.

In the data stream, a MAC address might read 0×35-87-C4-A2-E6-91. When stored in an Ethernet LAN it will be 0×AC-E1-23-45-67-89. When stored in Token Ring or FDDI LANs it will be 0×35-87-C4-A2-E6-91. (The 0×AC-E1-23-45-67-89 and 0×35-87-C4-A2-E6-91 are different representations of the same address.)

CHAPTER 4

Wide Area Networks

Wide area networks (WANs) consist of long-distance links joined together at various points by nodes that perform switching or routing functions. The nodes move frames from one link to another to guide them between the sending local network and the receiving local network. Because long-distance transport is expensive, all links will carry several channels multiplexed together. The links employ a variety of transmission techniques. Optical fibers and microwave radios probably carry the bulk of WAN traffic. They are supported by twisted pairs and other telephone cables and, in some cases, by wireless networks and communication satellite circuits.

Operations in the physical sublayer are synchronous or asynchronous:

- *Synchronous operation:* Actions occur at specific times in synchrony with other units in the network. A hierarchy of clocks synchronizes the entire network. They provide timing to all major facilities. The receiver uses one of these clocks to identify the boundaries between codes in the frames it receives. Synchronous operation is used in digital telephone networks. The frames require rudimentary headers and/or trailers. Examples are T-carrier networks, ISDNs, and SONETs. In addition, synchronous facilities are used to carry data traffic.
- *Asynchronous operation:* Nodes operate with similar internal clocks, but their actions are not synchronized or coordinated. To identify the boundaries between codes, the receiver recovers timing from bit transitions in the flag, or other synchronizing characters in the bit stream. Primarily, asynchronous operation is used in data networks. Examples are modem-mediated data connections over telephone lines, Ethernet LANs, and X.25 packet networks.

Before transfer to the physical medium, IP datagrams are encapsulated by network interface layer headers and trailers. They perform the same services as their LAN counterparts (i.e., delimitation, protocol identification, addressing, and bit-level integrity checking). WAN connections are divided into:

- *Point-to-point links:* They form a network segment with two terminal nodes. These links include telephone lines, ISDN circuits, digital subscriber lines, and T-carrier links.
- *Nonbroadcast multiple access (NBMA) links:* They connect more than two nodes but do not provide multicast or broadcast services. The physical link supports multiple virtual circuits that may connect to different nodes and dif-

ferent *service access points* (SAPs). NBMA links include those that operate with X.25, frame relay, and *asynchronous transfer mode* (ATM). In an IP environment, *inverse ARP* (InvARP) is used to discover the IP addresses of the nodes on the other ends of the virtual circuits.

4.1 Point-to-Point Links

Normally, private data circuits are enabled (turned up) at system generation. Absent users' traffic, they exchange short frames continuously. These frames serve to synchronize receivers to data streams and confirm that stations are ready to send or receive traffic. Frames are moved over point-to-point links by simple protocols such as PPP and SLIP. PPP employs the basic data link protocol, HDLC.

4.1.1 High-Level Data Link Control Protocol

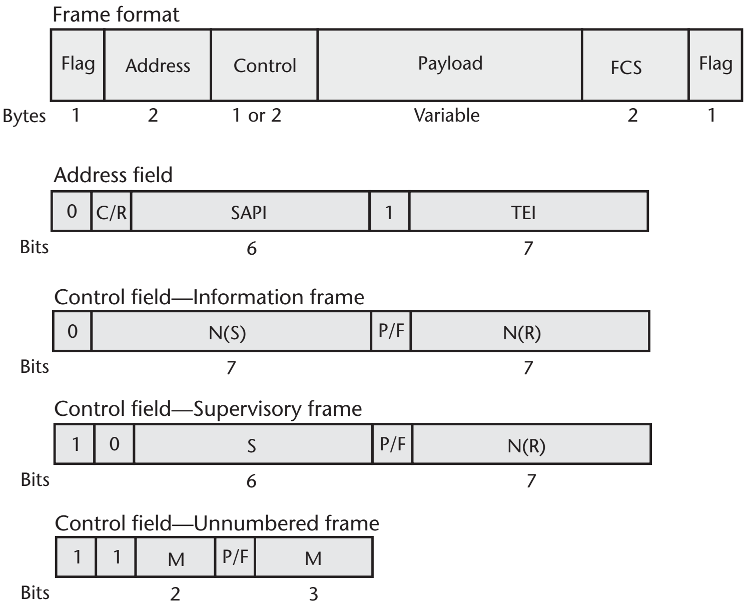
High-Level Data Link Control Protocol (HDLC) was first designed to work with packet networks. Standardized by ISO, HDLC makes use of a special character, the flag character (01111110 or 0×7E), to mark the beginning and ending of the frame. Between these markers, the header and the trailer fields are of predetermined lengths. The data that lie between the header and trailer are the payload. Over time, several variations of HDLC have appeared:

- *LAP-B:* Link Access Protocol—Balanced, first applied to the *user-network interface* (UNI) of X.25 packet-switched networks. Works in *asynchronous balanced mode* (ABM). The stations have equal status and each station may initialize, supervise, recover from errors, and send frames at any time. LAP-B served as the model for LAP-D and LAP-F.
- *LAP-D:* Link Access Protocol—Channel D, first applied to the data channel (D-channel) in ISDN. Works in ABM.
- *LAP-F:* Link Access Procedure—Frame Mode, first applied to frame mode services over the ISDN UNI on B-, D-, or H-channels.
- *PPP:* Point-to-Point Protocol, provides full-duplex data link services between peers (discussed later in this chapter).

Since LAP-D is included in PPP, I will describe its features in more detail. Figure 4.1 shows the format of a LAP-D frame and details the structure of the address and control fields.

4.1.1.1 LAP-D Address Field

The 2-byte address field marks the beginning of the first byte with 0 (bit 1) and the beginning of the second byte with 1 (bit 9). In byte 1, bit 2 identifies the frame as a command or response. A *command* frame requires an answer from the receiver. A *response* frame is the reply. The remaining bits of the 2-byte address field are divided between the *terminal endpoint identifier* (TEI, bits 3 through 8) and the *service access point identifier* (SAPI, bits 10 through 16):



C/R Command/Response
SAPI Service access point identifier
TEI Terminal endpoint identifier
N(S) Transmitter send sequence number
N(R) Transmitter receive sequence number
S Supervisory function codes
M Modifier function codes
P/F Poll/Final bit

Figure 4.1 HDLC Link Access Protocol—Channel D.

- *Terminal endpoint identifier (TEI):* Each physical node is assigned an address identifier. Assignment may be manual or automatic. The values are:
  - 0 through 63, manual assignment;
  - 64 through 126, automatic assignment;
  - 127 for temporary use during automatic TEI assignment.
- *Service access point identifier (SAPI):* Each node may support several Internet layer protocols. SAPI values are assigned that identify the buffer/queue (SAP, service access point) serving the specific protocol in the destination machine.

Called a *data link connection identifier* (DLCI), the combination of TEI and SAPI identifies a unique logical connection to an Internet layer protocol in a specific receiving device. The sending terminal may support several DLCIs simultaneously. They can be logical connections to different Internet layer protocols (control, network, or management protocols, for instance) in the same terminal or connections to different terminals (and Internet layer protocols). A given SAP is connected by a single DLCI to the sending/receiving machine.

4.1.1.2 LAP-D Control Field

LAP-D employs three types of frames. They are distinguished by the format of the control field. It occupies 1 or 2 bytes. The three types of frames are:

Information (I) frame. In the 2-byte control field:

- To identify an I-frame, the first bit of the first byte of the control field is set to 0.
- Bits 2 through 8 contain the number [N(S), 0 through 127] of this frame in the sending sequence.
- The first bit (bit 9) of the second byte is the P/F bit. In command frames, it is known as the *poll* (P) bit. When set to 1, it identifies this frame as requiring a response from the receiver. When set to 0, a response is not required. In response frames, the P/F bit is known as the *final* (F) bit. When set to 0, it identifies this frame as one of a continuing sequence. When set to 1, it is the final frame in the sequence.
- Bits 10 through 16 contain the number N(R) of the frame the sender expects to receive (0 through 127). It serves to acknowledge all frames up to N(R).

The information field must be an integral number of bytes. When user's data (payload) is sent, an information frame executes acknowledged operation. The N(S) and N(R) values provide the basis for error control (go-back-*n*) and flow control. In addition, I-frames carry control and management information.

Supervisory (S) frame. In the 2-byte control field:

- To identify an S-frame, the first 2 bits of the first byte of the control field are set to 01.
- Bits 3 through 8 contain codes for error and flow control: Receiver Ready (RR, 000000), Receiver Not Ready (RNR, 100000), and Reject (REJ, 010000). A supervisory frame is used when the receiver has no data ready to send in reply. RR signifies a positive acknowledgement and indicates ready to receive the next I-frame [N(R)]. RNR signifies a positive acknowledgment and indicates the receiver is not ready to receive next I-frame [N(R)]. REJ signifies a negative acknowledgment and indicates the sender must resend from N(R)].
- Bit 9 is the P/F bit.
- Bits 10 through 16 contain the number [N(R), 0 through 127] of the frame the sender expects to receive. It serves to acknowledge all frames up to N(R).

Unnumbered (U) frame. This frame provides unacknowledged service without flow control. Error detection is implemented, but not error correction. Upon detecting an error, the frame is discarded. In the 1-byte control field:

- To identify a U-frame, the first 2 bits of the first byte of the control field are set to 11.
- Bits 3 and 4, and bits 6 through 8 are codes that initiate communication, configure stations, test capabilities, and so forth.
- Bit 5 is the P/F bit.

4.1.2 PPP and SLIP

Point-to-Point Protocol (PPP) and *Serial Line Internet Protocol* (SLIP) are used to transport IP datagrams over point-to-point connections.

4.1.2.1 PPP

PPP encapsulates an IP datagram with an HDLC header and trailer. The frame is listed in Appendix B. Because it is a point-to-point connection, the three fields of the HDLC header—address, control, and protocol—can be omitted, or set as 0×FF (address), 0×30 (control), meaning an unnumbered information (UI) frame with poll/final bit set to 0, and 0×00–21 (protocol). The default value of the maximum size PPP frame [the *maximum receive unit* (MRU)] is 1,500 bytes (to be compatible with Ethernet). Other values (higher or lower) can be negotiated. PPP is used with SONET and SDH (see Section 7.4) and other transport systems.

4.1.2.2 Transparent Operation

On asynchronous links (such as modem mediated analog telephone lines), so that a flag character or an escape character within the IP datagram payload shall not interrupt transmission, PPP employs *character* stuffing to change the meaning of the offending character:

- In the IP datagram, a character that mimics the flag character (0×7E) is replaced by the sequence 0×7D–5E. 0×7D is the ESC character. At the receiving node, 0×7D–5E is replaced by 0×7E.
- An escape character within the IP datagram is replaced by 0×7D–5D. At the receiving node, 0×7D–5D is replaced by 0×7D.
- If the IP datagram contains the sequence 0×7D–5E, it is replaced by 0×7D–5D–5E.

In addition, a combination of character stuffing and bit stuffing is used to prevent characters in an IP datagram with values less than decimal 32 (i.e., less than 0×20) being interpreted as control characters. The ESC character is placed ahead of the character and the 6th bit is set to 1. [For instance, character 00010001 (0×11) becomes 0×7D–31 (i.e., 01111101 – 00110001)].

On synchronous links (such as T-carrier, ISDN, and SONET), *bit* stuffing is used between the framing flags to break up strings of 1s into segments of five 1s. Without regard to byte boundaries, 0 is stuffed after a sequence of five 1s. In this way, only the beginning and ending flags contain six consecutive 1s. As an example, consider the following data stream which has been divided into bytes for easier reading:

⇐01111110/01011111/11111101/11111011/01111110

The first 8 bits and the final 8 bits are underlined—they are the beginning and ending flags (07E, 01111110). In between, there is a section of the data stream (also underlined) that mimics the flag and extends over 2 bytes. Before transmission, between the beginning and ending flags, the transmitter inserts a 0 (denoted $\underline{0}$ for clarity) after sequences of five 1s. This makes the transmitted data stream

$$\Leftarrow 0111111001011111\underline{0}11111\underline{0}1011111\underline{0}101101111110$$

At the receiver, the zeros after five ones are removed to leave the original data stream.

4.1.2.3 Serial Line Internet Protocol

Another encapsulation that can be used to transmit IP datagrams over a point-to-point link is *Serial Line Internet Protocol* (SLIP). It is a very simple packet-framing protocol that only provides frame delimitation services. SLIP uses a special character called an END character (0×C0, 11000000). It is placed at the beginning and ending of each IP datagram. Two or more frames are sent in sequence with no space between them. The two END characters distinguish successive frames. In the IP datagram, to prevent the occurrence of the END character providing a false reading at the receiver, SLIP employs *character stuffing*.

- END characters within the IP datagram are replaced by the sequence 0DB–DC. At the receiving node, 0×DB–DC is changed back to 0×C0.
- ESC characters (0×DB) within the IP datagram are replaced by 0×DB–DD, and the sequence 0×DB–DD in the IP datagram is changed to 0×DB–DD–DC.

When SLIP links are used in conjunction with Ethernet networks, a maximum packet size of 1,500 bytes is used to prevent the fragmentation of IP datagrams.

4.2 Nonbroadcast Multiple Access Links

In packet-based systems, several logical circuits are established on the same physical conductor by assigning different identifiers to the traffic carried over each channel. Described as *virtual* circuits, they connect Internet layer entities in the sending terminal with Internet layer entities in one or more receiving terminals. X.25 packet switching, frame relay, and ATM are three examples of modern networks that employ NBMA links. Both X.25 packet switching and frame relay were designed for the bursty environment of data communication. ATM has been designed for simultaneous low-delay voice and video, as well as bursty data.

4.2.1 Packet-Switched Networks

In the 1970s, network developers focused on ways to transport bursty data traffic over long distances. The result was an innovative architecture called *packet switching*. Since then, the technology has evolved significantly, but the basic operations have remained the same. In part, this is due to a series of ITU Recommendations (X.25 et al.) that define the architecture and performance of the network. Known by some as *softswitches*, to distinguish them from circuit or *hardswitches*, packet switches are being used in some telephone carriers' central offices where they support asynchronous operations related to multimedia broadband applications and relieve the digital circuit switches of an uncertain load.

4.2.1.1 Architecture

ITU Recommendation X.25 describes the user-network interface. Figure 4.2 shows how a user's data file is segmented into fixed-length packets and formed into frames. Between the originating terminal and the node that serves as entry to the packet network, X.25 defines a three-layer protocol stack. Figure 4.3 shows the formal structure of the protocol stacks between the user and the network.

- In the *packet layer*, or *X.25-3 layer*, the user's data is divided into fixed length segments by the *packet layer protocol* (PLP), and a 3-byte *packet layer* header is added. In addition, PLP:
  - Multiplexes packets over the links on virtual circuits using *logical channel numbers* (LCNs) to identify the channels.
  - Performs flow control.
  - On the *receive* side, acknowledges receipt of frames and requests retransmission to correct errored frames (go-back-*n* or selective repeat ARQ).
  - On the *transmit* side, repeats unacknowledged frames.

  Packets of the same type are the same length. Originally, data packets were 128 bytes; later 512 bytes were used. X.25 allows payloads from 16 bytes (for *control* packets) to 4,094 bytes (for data packets).
- In the *data link layer*, or *X.25-2 layer*, the packet is encapsulated in an HDLC frame that implements Link Access Protocol–Balanced (LAP-B). LAP-B is similar to LAP-D. It uses 3 or 7 bits for packet numbering so that the receive window is 7 or 127 packets. Employs go-back-*n* ARQ, a 17-bit prime number as divisor for FCS , and an 8-bit address field. In addition, LAP-B:
  - Recognizes flags (to define frame limits).

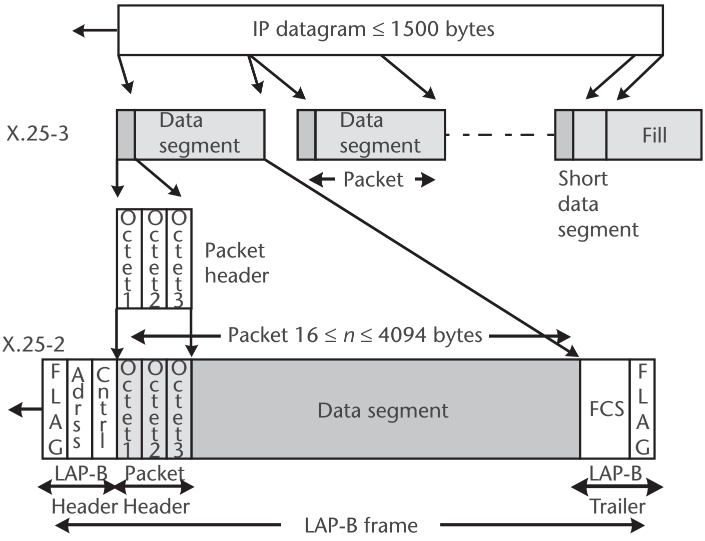


Figure 4.2 Illustrating the formation of a packet and its encapsulation in a LAP–B frame.

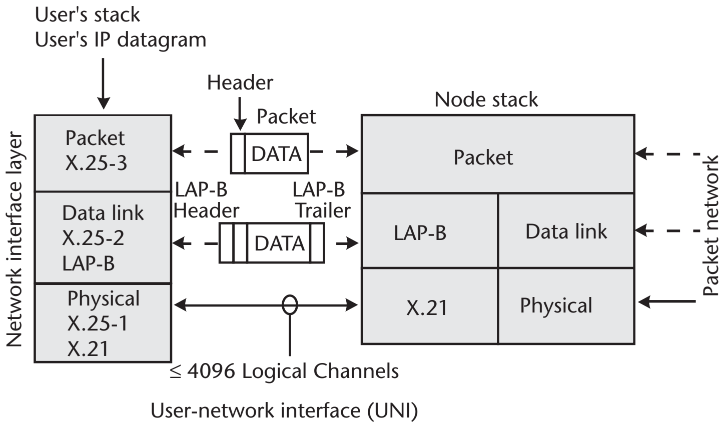


Figure 4.3 X.25 architecture.

- Executes bit stuffing (to achieve bit-transparency).
- On the *transmit* side, generates *frame check sequences* (FCSs).
- On the *receive* side, confirms FCSs.

- In the *physical layer*, or *X.25-1 layer*, the frame is transmitted over a logical channel (virtual channel) to the network node.

Figure 4.4 shows packet header formats for two data packets and a control packet. All include a 4-bit group number and an 8-bit channel number that, taken together, define 4,094 possible virtual circuits. The data packets differ in the number of bits assigned to the number of this packet [P(S)], and the number of the packet the sender expects to receive [P(R)]. With 3 bits, P(S) and P(R) ≤ 7; with 7 bits, P(S) and

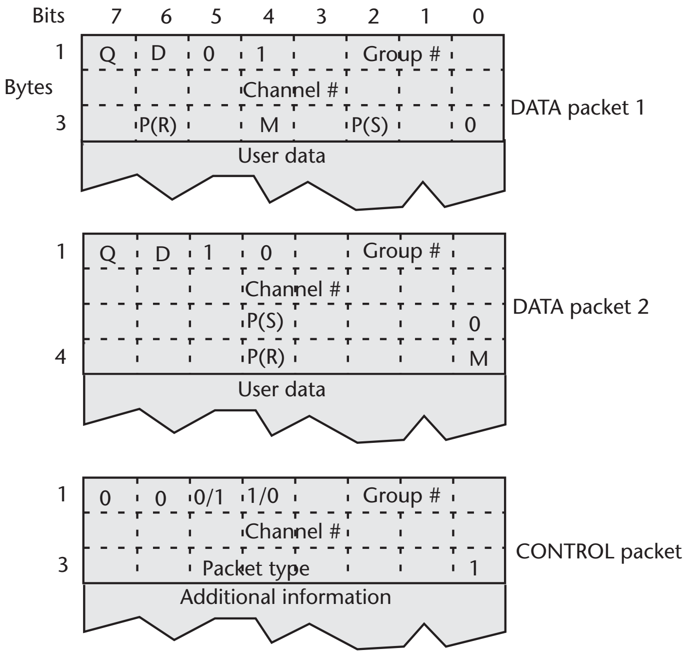


Figure 4.4 Packet formats.

P(R) ≤ 127. Using 3 bits, the sender must wait for an acknowledgment after sending seven frames. Only after all seven have been acknowledged as good can the sender begin the next packet number cycle. Using 7 bits, the sender can send up to 127 frames before waiting for an acknowledgment. Bits M, D, and Q support special functions.

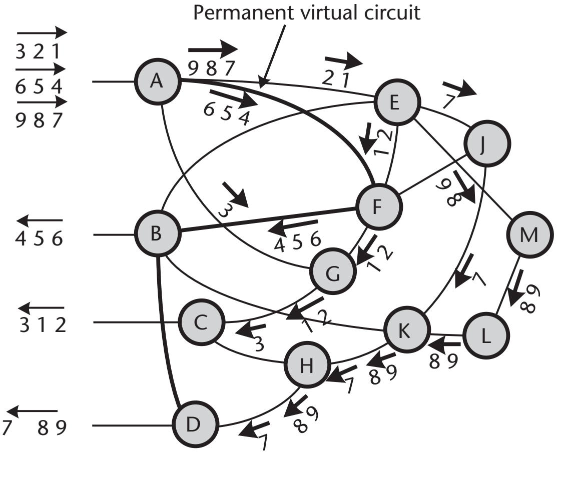
4.2.1.2 Routing

How frames are routed over a packet-switched network depends on the instructions given by the users. Three basic styles, similar to the routing techniques employed in router driven networks, can be distinguished:

- *Distributed routing:* On the basis of information about traffic conditions and equipment status (network map, port status), each node decides which link the frame shall take to its destination.
- *Centralized routing:* A primary (and perhaps an alternate) path is dedicated to a pair of stations at the time of need.
- *Permanent virtual circuit routing:* A virtual connection is permanently assigned between two stations.

Examples of each of these techniques are given in Figure 4.5:

- *Frames 1, 2, and 3* are sent from A to C using distributed routing. On the basis of the traffic distribution (links AF and AG are assumed to be congested), frames 1 and 2 are launched on link AE. Although it is not the shortest, this is a link that will connect to C. When frame 3 is presented to A, the link AG is less congested than AE. A sends frame 3 over link AG. Because frame 3 takes the path AGC, and frames 1 and 2 take the path AEFGC, frame 3 arrives at C ahead of frames 1 and 2.



Frames 1, 2, and 3 are sent from A to C with distributed routing
Frames 4, 5, and 6 are sent from A to B over a permanent virtual circuit
Frames 7, 8, and 9 are sent from A to D using centralized routing

Figure 4.5 Packet-switched network routing techniques.

- *Frames 4, 5, and 6* are sent from A to B over a permanent virtual circuit. They trace the route AFB in sequence.
- *Frames 7, 8, and 9* are sent from A to D using centralized routing. AEJKHD is defined as the primary route and AEMLKHD is an alternative. After frame 7 is sent over link EJ, a fault occurs that takes the link out of service. Frames 8 and 9 take the alternate route EMLK. The frames arrive in sequence at D but there is a delay between 7 and 8 because of the greater number of hops in the alternate route.

In the same way that the telephone numbers of the calling and called parties identify a telephone circuit, the originating and terminating logical channel numbers identify a virtual circuit.

A 128-byte packet can contain approximately 20 average words—and that may be less than two lines of text. Strings of frames, then, are common, and flow control procedures are needed to ensure that they are not sent so rapidly as to block the network links, or the receiving node.

4.2.1.3 Improving the Speed of Operations

When packet-switched networks were developed, the quality of the available transmission links was poor. As a result, every node spends time checking for errors. Consequently, packet-switched networks are slow. With the upgrading of transmission facilities to permit the introduction of digital services and the appearance of optical fibers, it has been possible to relax some of these requirements. In one approach, known as *cell relay*:

- Checking functions are dropped from intermediate nodes.
- Checking and control are moved to the edges of the network.
- 53-byte cells replace the standard packet.

In a second approach, known as *frame relay*:

- The user’s data are kept in variable length frames.
- LAP-D is applied in two steps. The data link layer protocol is changed to a limited set of capabilities known as LAP–D core and the other activities in LAP–D (known as LAP–D remainder) are completed end to end.

Figure 4.6 compares the network interface protocol stacks for packet switching, frame relay, and cell relay (ATM). Note that, in packet switching, full error control occurs with each link. Error detection results in discarding the packet and requesting retransmission. In frame relay and cell relay, error detection may occur, but error correction is left to upper level protocols.

4.2.2 Cell Relay

Cell relay service (CRS) transports voice, video, and data messages in streams of short, fixed-length cells. By dividing the payload in short segments, cell relay achieves short processing delays. Such performance is ideal for transporting voice

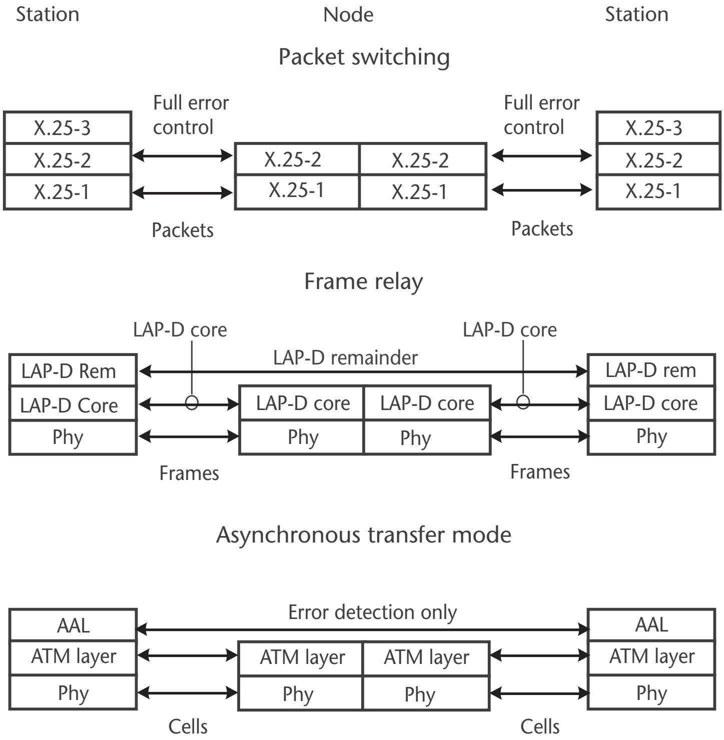


Figure 4.6 Protocol stacks for packet switching, frame relay, and ATM.

and video streams that are sensitive to delay and is not detrimental to data communication. Voice is carried as a *constant bit rate* (CBR) stream with low delay and low cell loss. Video is carried as a CBR stream or a real-time *variable bit rate* (VBR) stream. The bit rate cannot exceed the *peak cell rate* (PCR) negotiated with the network. Data is carried as a VBR stream, as a stream that uses the *available bit rate* (ABR), or as a stream for which the bit rate is unspecified (UBR). With UBR, the sender transmits as fast as it can (up to its PCR). Cell relay is implemented as ATM.

ATM is a packet switching technology that uses 53-byte, fixed-length cells to implement cell relay service. ATM employs virtual circuits (duplex) that are assigned by a signaling network prior to message transmission. ATM supports the transport of:

- Isochronous streams (a synchronizing process in which the timing information is embedded in the signal; a voice or video data stream);
- Connectionless data packets;
- Connection-oriented data packets.

ATM switches are deployed in data, voice, and video applications. In the Internet backbone they carry point-to-point traffic at speeds of 622 Mbps.

4.2.2.1 ATM Call Setup

Signaling is achieved over a separate, permanently assigned network. Each station is connected to one controller. Call setup (and termination) information is sent over a signaling connection to the network controller serving the originating node. The controllers communicate with one another over dedicated high-speed connections. Because the channel is set up before cells are transmitted, there is no need for source and destination addressing with a call. Thus, in Figure 4.9, the IEEE 802.3 header in the IP datagram frame is omitted.

4.2.2.2 Virtual Paths and Virtual Circuits

Over an ATM network, stations communicate using virtual circuits. To divide them into manageable groups, *virtual channels* (VCs) are grouped in *virtual paths* (VPs). When a request for a new connection is received, the traffic controller attempts to place it on an existing VP where resources are available, and the call will have no effect on in-use circuits. If this cannot be done, the controller may elect to place the call on the path and accept service degradation on the calls in progress, add resources to the path, seek another existing path, establish a new path, or refuse the call.

4.2.2.3 ATM Architecture

The architecture of ATM consists of the cell, the *user-node interface* (UNI), the *node-network interface* (NNI), and ATM protocol layers.

- *Cell*. This consists of 48 bytes of payload and 5 bytes of header information. If necessary, the first 4 bytes of the payload are used to identify and sequence the remaining 44-byte segments. Figure 4.7 shows the structure of an ATM cell. The fields are listed in Appendix B. In addition, Figure 4.7 shows a resource management cell. Its use will be explained in Section 4.2.2.5.
- *ATM UNI header*. This consists of:
  - 4-bit *generic flow control* (GFC) field intended to assist in controlling the flow of local traffic at the UNI;
  - 24-bit connection identifier [16-bit *virtual channel identifier* (VCI) and an 8-bit *virtual path identifier* (VPI)];
  - 3-bit *payload type identifier* (PTI) that indicates whether the cell contains upper-layer header information or user data;
  - 1-bit *cell loss priority* (CLP) field used to identify lower priority cells that, in the event of congestion, should be discarded first;
  - 8-bit *header error control* (HEC) that is used for error detection in the header.
- *ATM NNI header*. This is similar to UNI except that the GFC field is replaced by four additional VPI bits to make the VPI field 12 bits.

4.2.2.4 ATM Protocol Stack

Figure 4.8 shows the ATM protocol stack. It consists of three layers that occupy the network interface layer of the Internet model:

- *ATM adaptation layer (AAL):* When sending, AAL converts IP datagrams into sequences of cells for use by the ATM layer. When receiving, AAL converts

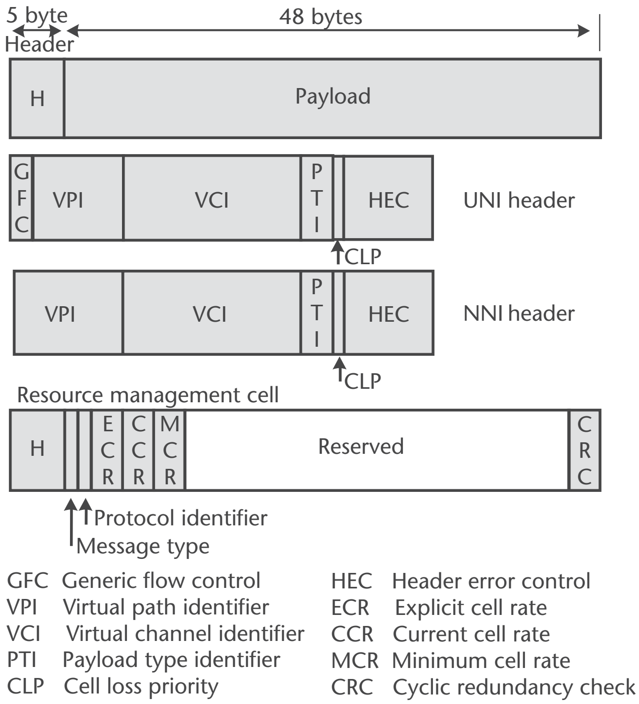


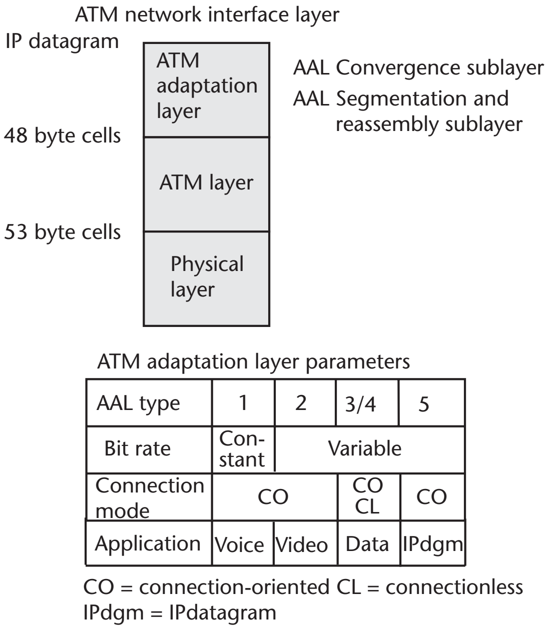
Figure 4.7 ATM cells.

sequences of cells to IP datagrams for use by upper layers. AAL is divided in two sublayers.

- *Convergence sublayer (CS):* When sending (i.e., receiving a PDU from the Internet layer), the CS constructs a CS PDU that consists of the payload, a pad to maintain a 48-byte alignment, and a trailer. When receiving, accepts CS PDU from SAR, strips off trailer, reconstructs PDU received from Internet layer, confirms error-free reception, and delivers PDU to the Internet layer. If the reception is not error-free, the CS discards the CS PDU and notifies the Internet layer.
- *Segmentation and reassembly sublayer (SAR):* When sending, SAR divides CS PDU into 48-byte SAR PDUs and delivers them to the ATM layer. When receiving, receives 48-byte SAR PDUs from ATM layer, reconstructs CS PDUs, and sends them to CS.

- *ATM layer (ATM):* When sending, adds 5-byte header (UNI or NNI, as appropriate) to 48-byte SAR PDUs, multiplexes 53-byte cells to message streams identified by VCIs and VPIs, and delivers them to the physical layer. When receiving, demultiplexes cells, deletes 5-byte header from 53-byte cells, checks error-free reception of header, and delivers SAR PDUs to SAR.
- *Physical layer:* Transports digital signals over multiplexed connections in a synchronous digital network.

Each type of AAL has been designed to handle a specific class of traffic. Figure 4.8 includes a table that summarizes their traffic handling ability.



AAL type	1	2	3/4	5
Bit rate	Constant	Variable		
Connection mode	CO		CO CL	CO
Application	Voice	Video	Data	IPdgm

CO = connection-oriented CL = connectionless
IPdgm = IPdatagram

Figure 4.8 ATM protocol layers.

- *AAL 1* provides a connection-oriented, constant bit rate voice service. AAL1 performs segmentation and reassembly, may detect lost or errored information, and recovers from simple errors.
- *AAL 2* is a connection-oriented variable bit rate video service. AAL2 performs segmentation and reassembly and detection and recovery from cell loss or wrong delivery.
- *AAL 3/4* is a combination of two services designed for connection-oriented and connectionless data services. AAL3/4 is an all-purpose layer that supports connection-oriented and connectionless variable bit-rate data services. Two operating modes are defined.
  - *Message mode:* Each *service data unit* (SDU) is transported in one *interface data unit* (IDU). Employs cyclic redundancy checking and sequence numbers.
  - *Streaming mode:* Variable-length SDUs are transported in several IDUs that may be separated in time.
- AAL5 was created by an industry forum to send frame relay and IP traffic over an ATM network. AAL5 supports connection-oriented, variable-bit-rate, and bursty data services on a best-effort basis. It performs error detection but does not pursue error recovery. AAL5 is essentially a connection-oriented-only AAL3/4 layer. AAL5 is also known as the *simple and efficient layer* (SEAL).

As an example, suppose an IEEE 802.3 Ethernet frame is sent using AAL5. Before division into cells, the IEEE 802.3 header is removed. Four bytes are inserted in the IEEE 802.3 trailer to create the AAL 5 trailer. In this trailer the length of the payload is recorded so that the receiver can discard any pad. As usual, the FCS is used to check the integrity of the frame before it is delivered to the Internet layer at its ATM destination. Figure 4.9 shows the division of an IP/UDP datagram with a 256-byte application PDU into seven ATM cells. The last cell includes a pad of 8 bytes. The fields are listed in Appendix B.

4.2.2.5 Available Bit Rate Service

To transfer cells as quickly as possible, a sender may try to use the bit rate (bandwidth) that is not allocated to other traffic. To do so without loss of data, the source must adjust its sending bit rate to match conditions as they fluctuate within the network. To control the source bit rate when using ABR service, *resource management* (RM) cells (see Figure 4.7) are introduced periodically into the sender's stream. RM cells are sent from sender to receiver (*forward* RM cells), and then turned around to return to the sender (*backward* RM cells). Along the way, they provide rate information to the nodal processors and may pick up congestion notifications. When an RM cell reaches the receiver, it (the receiver) changes the direction bit ready to return the cell to the source. If the destination is congested, it sets the *congestion indication* (CI) bit and reduces the *explicit cell rate* (ECR) value to a rate it can support. On the return of the RM cell to the source, the sending rate is adjusted accordingly. If the RM cell returns to the source without the CI bit set, the sender can increase the sending rate and set a higher ECR.

4.2.3 Frame Relay

Frame relay is a connection-oriented, network interface layer, packet-switching technology that transfers variable length frames (262 to 8,189 bytes). Originally, this was done at DS–1/E–1 speeds (1.544/2.048 Mbps). More recently, speeds up to 140 Mbps have been reported. Frame relay is well suited to data transport. By handling long datagrams without segmentation, it eliminates most of the delay in processing strings of packets. Of course, the longer the individual frames, the longer the time required to assemble them by the sender and the longer the time required to evaluate them at the receiver. Generally, delays of this sort are not serious issues in data communication; however, they pose problems for voice and video streams.

The frame relay user network interface employs a set of core functions derived from LAP–D. It uses 7 bits for packet numbering so that the receive window is 127 packets, employs go-back-*n* ARQ, and a 17-bit prime number as divisor for FCS (10001000000010001). The LAP–D core: supports limited error detection (but not

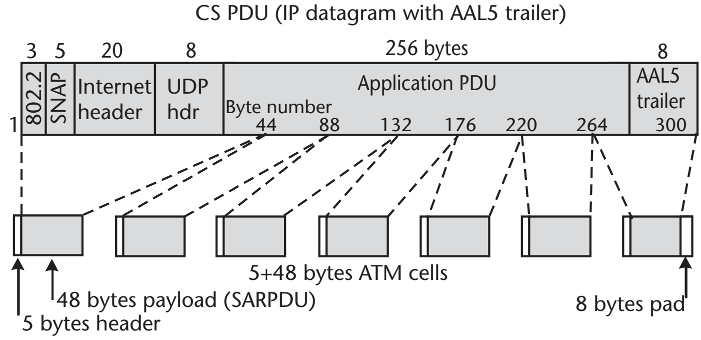


Figure 4.9 Division of CS PDU (IP datagram with AAL 5 trailer) into ATM cells.

correction) on a link-by-link basis. It recognizes flags (to define frame limits), executes bit stuffing (to achieve bit-transparency), generates or confirms frame check sequences, destroys errored frames, and, using logical channel numbers, multiplexes frames over the links.

The remaining LAP–D functions are performed end-to-end. The LAP–D remainder acknowledges receipt of frames, requests retransmission of destroyed frames, repeats unacknowledged frames, and performs flow control.

4.2.3.1 Limits to Frame Relay Operation

Frame relay does not guarantee faultless delivery of data:

- It detects, but does not correct, transmission, format, and operational errors.
- It may discard frames to clear congestion or because they contain errors. When an invalid frame is detected (for any reason), the node discards the frame.
- It is left to the receiving end-user system to acknowledge frames or request retransmission of frames.

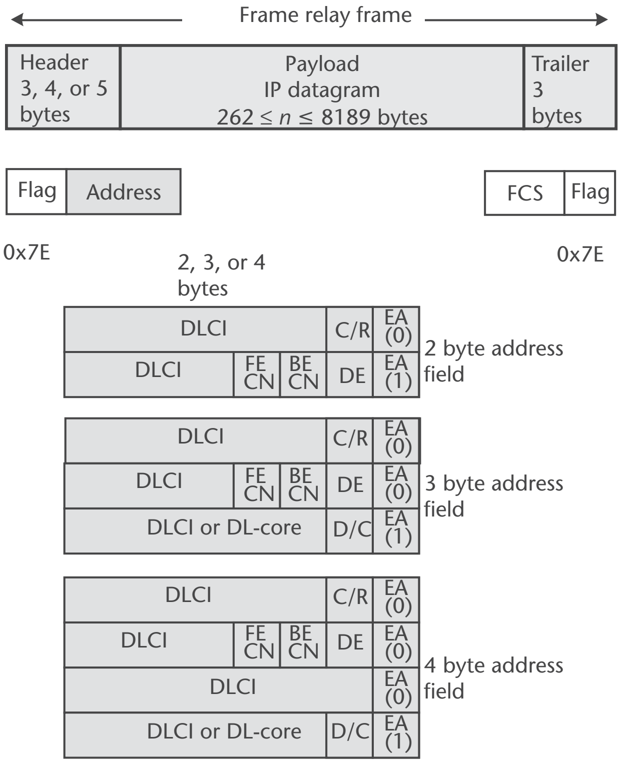
Despite these caveats, frame relay is a technique of choice for data networks that interconnect LANs separated by substantial distances over reliable transmission facilities.

4.2.3.2 Frame Relay UNI

Just as X.25 is directed to the *user and network interface* (UNI), so frame relay is a network access technique. Within the network [i.e., over the *network node interface* (NNI)], the procedures employed may be frame relay, cell relay, X.25 or ISDN. Often, a *frame relay access device* (FRAD) connects the user to an FR network. As shown in Figure 4.10, a header and a trailer encapsulate the payload (e.g., IEEE 802.3 Ethernet frame). In the header, the address field is 2, 3, or 4 bytes long. In these addresses, the major entry is the *data link connection identifier* (DLCI). With 10, 16, or 24 bits, it identifies the virtual circuit over which the frame is sent. The last bit of each byte tells whether this is the last byte of the address (1), or the address continues for at least one more byte (0). Frames are divided into *commands or responses* (C/R bit). The former requires a response; the latter is the response to a command or a frame that does not require a reply. Control bits are included for flow control (FECN and BECN) and *discard eligibility* (DE). A frame relay frame with 2-byte addressing is listed in Appendix B.

4.3 Quality of Service

Long-distance communication is characterized by multiplexing—the placing of more than one signal on the same bearer—in order to reduce transmission costs. Under normal circumstances, this sharing of resources is not detrimental to performance. However, when the number of signals exceeds the normal capacity of the system, the service that each frame receives will be degraded, some frames may be delayed, and others may be denied transport.



DLCI Data Link Connection Identifier
BECN Backward Explicit Congestion Notifier
C/R Command/Response Indication
EA Address Field Extension Bits
DE Discard Eligibility
FECN Forward Explicit Congestion Notification
FCS Frame Check Sequence
D/C DCLI or DL-core Control Indicator

Figure 4.10 Frame relay frames.

In the IP header (described in Section 1.3 and listed in Appendix B), there is a one-byte field entitled *type of service*. Its purpose is to indicate the level of service that the *sender* expects intermediate routers to give to the frame. For most frames, the byte is set to 0×00 by the sending host, i.e., normal precedence, delay, throughput, reliability, and cost. However:

- If there is some urgency about the contents of the frame, the sender can set the three-bit *precedence* to a value between 0 and 7. For routers able to respond, frames with precedence of 6 or 7 will be moved to the head of any queues they may encounter. When several frames are marked for preferential treatment, the one with highest precedence will be served first.
- If timeliness is important to the sender, low delay can be requested by setting the *delay* bit to 1.
- If the rate at which bits are delivered is important to the sender, high throughput (i.e., high bandwidth) can be requested by setting the *throughput* bit to 1.

- If it is important to the sender to send the frame over reliable circuits, high reliability links are requested by setting the *reliability* bit to 1.
- Finally, if none of the above is necessary, the sender may request low cost by setting the *cost* bit to 1.
- The eighth bit is reserved for future use.

Of course, merely setting the bits is no guarantee that the requests will be honored. The terms must be negotiated with each intermediate node before transmission begins. This can be done using *Resource Reservation Protocol* (RSVP). RSVP requests a path from a sender to a receiver (or multiple receivers) with given performance (i.e., bandwidth, delay, reliability). RSVP sends a *path* message specifying the requirements to all intermediate routers in the general direction of the receiver(s). If they can, the routers will respond affirmatively and agree to supply the requested performance. If they cannot, they refuse the request. Under this circumstance, the sender may seek an alternate path, modify the requirement, or postpone the activity. In addition, when made aware of the sender's request, the receiver(s) will send *reserve* messages confirming the requirement back through the intermediate routers to the sender. When the session ends, the reservation is made void with another series of messages, and the resources are freed ready for re-allocation by their respective routers.

4.3.1 Differentiated Services

The 7 active bits in the type of service field of the IP header provide an opportunity for the sender to request 128 different sets of conditions. Is it reasonable to expect routers to discriminate among so many classes of frames and respond in 128 distinct ways? Absolutely not! Accordingly, the IETF has modified the meaning of the type of service field seeking relatively simple and coarse solutions to providing *differentiated services* (DS). Their approach uses the first six bits (0 through 5) to form a *differentiated services codepoint* (DSCP) and leaves bits 6 and 7 undefined. The 64 codepoints are mapped to a few service definitions that can be provided by the router. The first 3 bits of the codepoint provide a precedence value. Intermediate routers provide differentiated levels of services to IP packets and forward them in accordance with *per hop behaviors* (PHBs). Each PHB is a service definition that is applied to a group of codepoints. Frames that receive the same PHB treatment are said to belong to a *per domain behavior* (PDB).

4.3.2 T-1 Performance Measures

In Section 7.2.1, I describe the error-detecting format employed in T-1 systems that use *extended superframe* (ESF). With a fixed number of channels and synchronous transmission, performance is defined by the number of errored frames received. Error performance is measured by loss of synchronization evidenced by incorrect framing bits, and a 6-bit *frame check sequence* (FCS). (The bit stream is divided by a 7-bit polynomial [1000011] to give a 6-bit FCS.) The six frame check (C) bits provide a cyclic redundancy check that monitors the error performance of the 4,632-bit superframe. Some of the conditions used to describe link performance are:

- *ESF error*. An OOF event, or a CRC-6 error event, or both, has (have) occurred. The meanings of these events are:
  - *Out of frame (OOF):* Condition when 2 out of 4 consecutive framing bits are incorrect (i.e., do not match the 101010 pattern).
  - *CRC-6 error:* Condition when the FCS calculated by the receiver does not equal the FCS delivered with the frame.
- *Errored second (ES)*. A second in which one, or more, ESF error condition(s) is (are) present:
  - *Bursty second (BS):* A second in which from 2 to 319 ESF error events are present.
  - *Severely errored second (SES):* A second in which from 320 to 333 ESF error events are present.
- *Failed seconds state (FS)*. Ten consecutive SESs have occurred. This state remains active until the facility transmits 10 consecutive seconds without an SES.

Error event data are analyzed and stored in the CSUs (channel service units) that terminate the link. An ESF controller (see Figure 7.6 in Chapter 7) maintains surveillance on a group of links and interrogates the CSUs on a routine basis. Depending on circumstances, the controller will report emergencies and prepare operating reports that detail performance. Collecting these measures has made it possible to describe performance and establish standards for T-1 links.

4.3.3 ATM Performance Measures

Among many other parameters, an agreement for ATM services may specify:

- *Peak cell rate (PCR):* The maximum rate at which cells are presented to the network.
- *Sustainable cell rate (SCR):* The rate at which cells can be presented to the network and assured of delivery.
- *Maximum burst size (MBS):* The greatest number of cells that are presented in a sequence.
- *Minimum cell rate (MCR):* The minimum rate at which cells are presented to the network.
- *Cell loss rate (CLR):* The difference between the number of cells sent and the number of cells received *divided* by the number of cells sent.
- *Cell misinsertion rate (CMR):* The number of cells received not intended for the receiver *divided* by the number of cells sent.

The values agreed for these parameters bind both parties. Should the corporate user exceed the agreed values, the provider is not obliged to transport the signals, nor subject to penalties for noncompliance. Should the corporate user run within these limits, the provider is subject to penalties for nonperformance.

The rate at which traffic enters the network is critical to maintaining service levels. At call setup time the host signals its requirements to the network. Each ATM switch in the path determines if sufficient resources are available to set up the connection as requested. If a switch cannot support the level, the setup message is rerouted to another switch along an alternate path to the destination. If the network is unable to support the request for call setup, it is rejected. The potential sender has the option to accept a lesser requirement, or wait until resources are available.

The ATM Forum defines five service levels, which, because ATM is a multimedia switch, include levels for data, voice, and video applications:

- *Class 1:* Supports constant bit rate video. The performance is comparable to a digital private line.
- *Class 2:* Supports variable bit rate audio and video. It is intended for packetized video and audio in teleconferencing and multimedia applications.
- *Class 3:* Supports connection-oriented data transfer. It is intended for interoperation of connection-oriented protocols such as TCP.
- *Class 4:* Supports connectionless data transfer. It is intended for interoperation of connectionless data transfer protocols such as UDP.
- *Class 5:* No objective is specified for the performance parameters. It is intended to support users who can regulate the traffic flow into the network and adapt to time-variable available resources.

4.3.4 Frame Relay Performance Measures

Frame relay may be implemented directly over T-1 links or with a core network of ATM switches. In the former case, performance is related to the discussion of T-1. In the latter case, performance is related to the discussion of ATM. Among many other parameters, an agreement for frame relay services may specify:

- *Committed information rate (CIR):* The rate at which the network agrees to transfer data.
- *Excess information rate (EIR):* The rate at which bits are sent *minus* the CIR.
- *Error rate:* In a given time, the number of errored frames received *divided* by the number of frames sent.
- *Residual error rate (RER):* The total number of frames sent *minus* the number of good frames received *divided* by the total number of frames sent.

4.3.5 QoS

The potential for service at a level different from that which the sender requests has given rise to concerns for the *quality of service* (QoS). This is particularly true for corporate users who seek to contract for specific capacity and performance levels. For them, best effort is no longer acceptable. Driven by competition for long-distance customers, providers have responded by specifying the anticipated performance of their facilities.

In a strict sense, quality is not measurable. It falls in the *I-know-it-when-I-see-it* category of human experiences. The measures and statistics listed earlier provide quantitative descriptions of performance that can be related in some way to the wishes of customers. Furthermore, they can be the basis for contracts and agreements between buyers and sellers. Fortunately, data communication is a robust art and the primary ingredient of success is accurate delivery. When all else fails, it is obtained by repetition.

CHAPTER 5

Connecting Networks Together

LANs can be connected to other LANs to make a common work environment and create larger, transparent networks called *catenets*. A catenet is an aggregate of networks that behaves as a single logical network. To create them, bridges and routers are used. The choice depends on the degree of difficulty of the communication process.

5.1 More Than One Network

Figure 5.1 shows an arrangement in which the communicating client and server is separated by several networks. More than likely, they are connected to their immediate neighbors over local area networks. These LANs are connected to other LANs by local facilities that link them in regional networks, and a long-distance network interconnects the regional networks. The regional and long-distance facilities are *wide area networks* (WANs). In order for Client A to communicate with Server B, moving frames over Client A's LAN to a regional WAN is required. Then, the frames are moved to a long-distance network (another WAN) that connects to another regional network and to Server B's LAN. Subject to different traffic patterns and operating conditions, these networks employ different technologies. Linking them together requires the use of specialized equipment.

5.1.1 Repeaters, Bridges, Routers, and Gateways

Key to the operations in Figure 5.1 are the interface matching devices. Their capabilities depend on the highest layer of the Internet model in which differences exist between the two networks they are connecting.

If differences only exist in the physical sublayers of the network interface layers, the interface-matching device is called a *repeater*. It accommodates differences in implementation of the transmission facilities. Repeaters handle electrical-to-optical conversions, signal and level changing, and other tasks.

If differences exist in the physical sublayers and/or the data link sublayers of the network interface layers, the interface-matching device is called a *bridge*. It accommodates differences in implementation in data stream formats and in transmission facilities. Thus, bridges handle changes in data formats (control bits, sequence numbers, hardware addresses, error control procedures, and flow control), as well as changes associated with transmission facilities.

If differences exist in the network interface layer and/or Internet layers, the interface-matching device is called a *router*. It accommodates differences in imple-

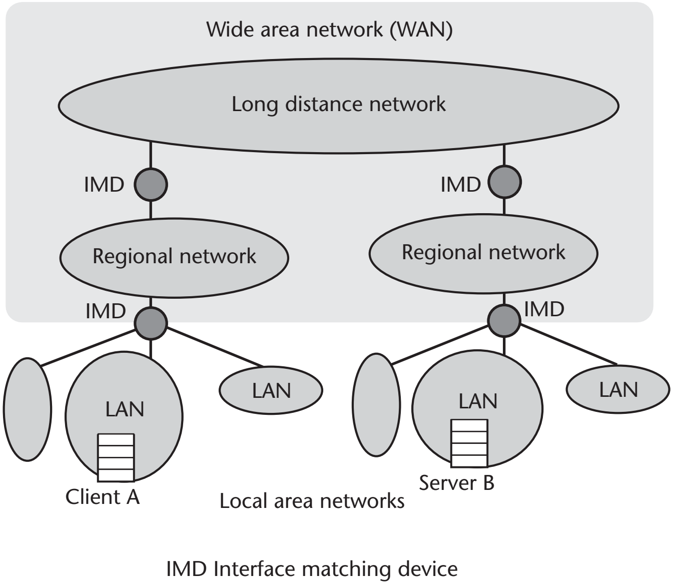


Figure 5.1 Connecting Client A to Server B.

mentation in forwarding and addressing, in data formats, and in transmission facilities. Thus, changes in routes, forwarding addresses, and segment sizes, as well as changes associated with the data stream and transmission facilities, are handled by routers.

If differences exist above the Internet layer, the interface-matching device is called a *gateway*. It accommodates differences in implementation at the higher layers of the protocol stacks. Thus, a gateway is required to interface different spreadsheets or different drafting systems, for instance.

Figure 5.2 shows the protocol stacks for a repeater, a bridge, a router, and a gateway, and illustrates the use of bridges and routers to connect clients and servers. In the layers of the protocol stacks intermediate between Client A and Server B, headers and trailers are removed, modified to reflect network differences, and replaced so that the frames can continue on their journey. Much of the discipline of data communication is devoted to ensuring that proper values are included in these headers and trailers, and they are altered appropriately at each intermediate handling point.

By way of illustration, Figure 5.3 shows the frame makeup when transferring an IP frame between two hosts connected by a router. Headers and trailers (TH1, IH1, NH1, NT1, ...) are added and subtracted along the way as user's data is passed from System 1 to System 2. Below the stacks are the PDUs that are passed from host to router, and router to host, over the two transmission systems. The combinations *IH1 + TH1 + Application PDU* and *IH2 + TH1 + Application PDU* are IP datagrams. A network interface header and trailer encapsulate each of them. Above the router stack is the transport layer PDU that was created originally in the transport layer of System 1. It has been recovered by decapsulating the frame as it passes up the router stack. Above the protocol stacks of System 1 and System 2 is the block of user's data that is transferred from one to the other.

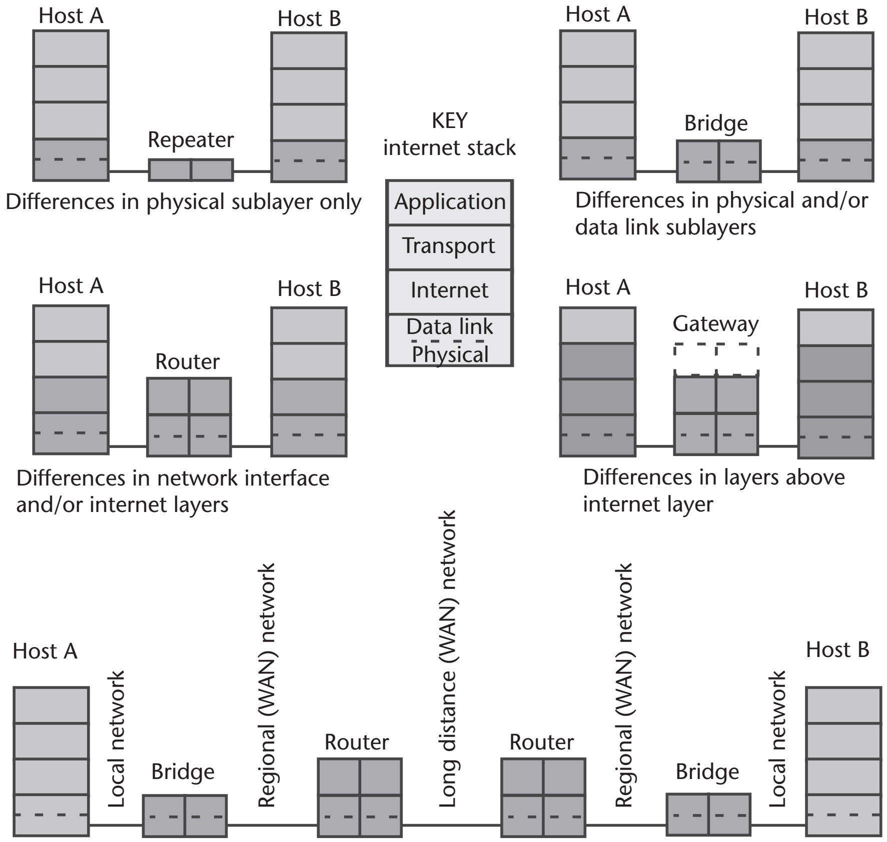


Figure 5.2 Protocol stacks for repeaters, bridges, routers, gateways, and multinode wide area network.

Note that the process employs only one transport layer header. No matter how many intermediate routers are encountered between the sending and receiving hosts, this header does not change. In addition, the process employs two Internet layer headers, two data link sublayer headers, and two data link sublayer trailers. They will change at each router as addresses and times to live change and checksums and FCSs must be recalculated.

5.1.2 Layer 2 and Layer 3 Switches

Bridges, routers, and gateways were based on special-purpose, software-driven platforms that required programs of varying complexity. Because of the cycles required, execution was relatively slow, and, as network speeds increased, they became bottlenecks. Steadily, as advances were made in the density and complexity of integrated circuit chips, more of the logic was committed to hardware. Operating at *wire speeds*, these hardware implementations have reduced response times. In addition, miniaturization has concentrated more powerful performance in smaller spaces. The result is that today's bridges and routers look different and perform significantly better than yesterday's models. Seeking to emphasize this point and differentiate the new from the old, some vendors have named these products *Layer 2 and Layer 3 switches*. The terms Layers 2 and 3 imply an OSI model. In an Internet world, the naming is understandable, if not precise. Notwithstanding the name

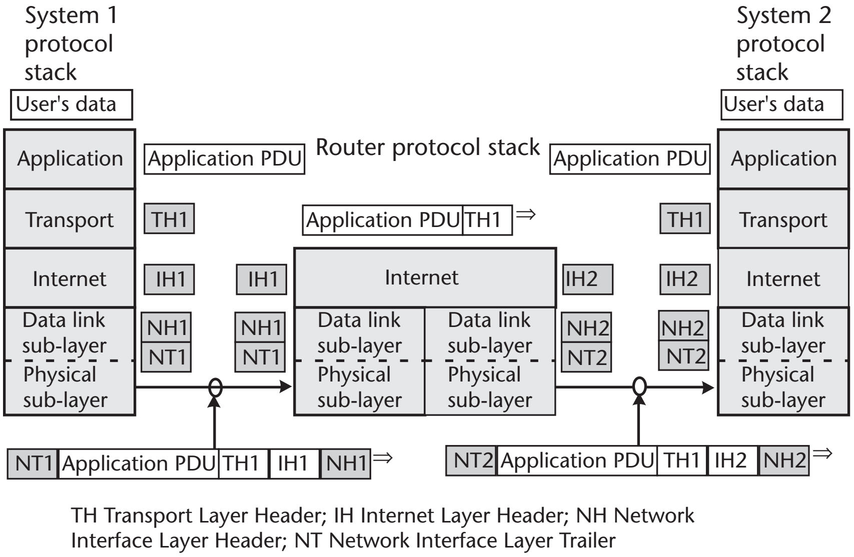


Figure 5.3 Headers/trailers employed in host–router–host path.

change, a Layer 2 switch performs the functions of a bridge, and a Layer 3 switch performs the functions of a router. They just do them faster.

5.2 Bridging

Joining several LANs together at the data link sublayer requires the capabilities of a bridge. The complexity of its task depends on the number and kind of LANs involved.

5.2.1 Bridging Identical LANs

Figure 5.4 shows an arrangement in which a bridge is used to connect five Ethernets to create a catenet. I could have chosen a catenet of Token Ring or FDDI LANs. The important requirement is that they be identical so that the bridge is solely a director of traffic. It does not have to engage in technology mediation as well. The bridge receives copies of all frames sent on each Ethernet. Because it overhears everything, the bridge is said to be operating in *promiscuous* mode. Further, it maintains a table that lists the 6-byte MAC addresses of all stations on all Ethernets, and the number of the port to which each station is connected. Stations communicate as if they were on the same LAN. Figure 5.5 shows the basic functions performed by the bridge.

When a station on Ethernet 1 sends a frame, all stations on Ethernet 1 plus Port 1 of the bridge receive it. The bridge examines the target destination address in the frame and searches the table for an entry that identifies the port on the bridge to which the destination station is attached.

If the target destination is attached to Port 1 (i.e., it is on Ethernet 1, the LAN from which the frame originated), the bridge assumes the frame has been processed in the normal way. It discards its copy of the frame. The bridge is said to *filter* all frames whose target addresses reside on the same port as that on which the frame arrived.

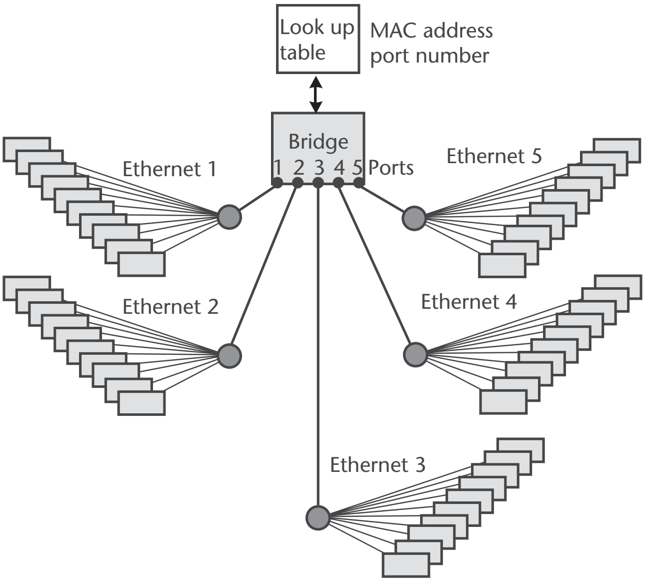


Figure 5.4 Bridging Ethernets.

If the target destination is not on Ethernet 1, and the table contains an entry, the bridge transfers the frame to the port identified by the entry. When the target Ethernet is quiet, the port launches the frame. If there is no collision, the frame will be delivered to its destination. If there is a collision, the port backs off and sends again, as required by the CSMA/CD routine.

If the target destination is not on Ethernet 1, and there is no entry in the table, Port 1 may destroy its copy of the frame. More likely, if traffic conditions permit, it will provide duplicate copies of the frame to Ports 2 through 5. As soon as they can seize the network, these ports will *flood* their Ethernets with the frame. If the target address exists on any network, the frame will be delivered.

To build a table, the bridge examines all frames received for the addresses of the *sending* stations. The addresses and the number of the ports on which they were received are used to build the look-up table. In this way, the bridge can keep an up-to-date record of all active stations, and stations that have not been active for some time can be removed from the list.

5.2.1.1 Table Search Algorithms

Conceptually, the idea of a table of station addresses and corresponding port numbers has merit. However, if all addresses are unicast and global, the number of *variable* address bits is 46; 2^{46} is approximately 7×10^{13}. To search such a space entry-by-entry in a reasonable time is impossible. A straightforward strategy is *binary searching*. With the address table sorted in numerical order, the input address is compared to the address at the center of the table. If it is larger than the center value, the address must be in the bottom half of the table. If it is less than the center value, the address must be in the upper half of the table. The search proceeds to the center of the half in which the address is located. If the address is less than the new center value, it must be in the upper half of that half of the table. If the address

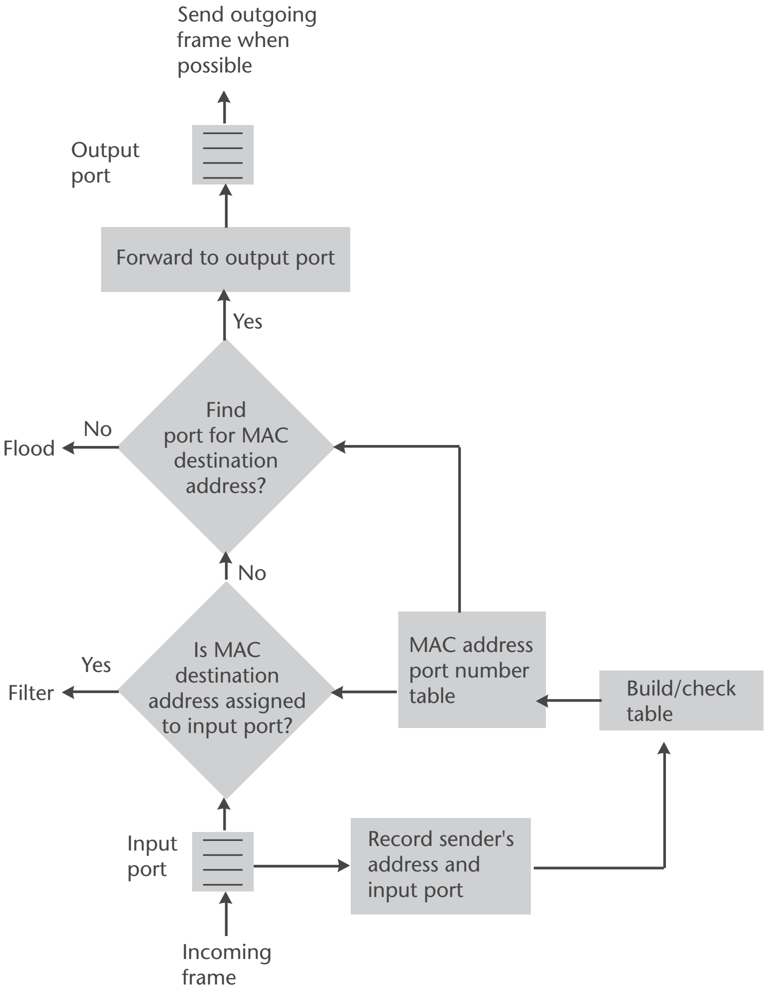

Send outgoing frame when possible
Output port
Forward to output port
Yes
Find port for MAC destination address?
No
Flood
No
MAC address port number table
Is MAC destination address assigned to input port?
Yes
Filter
Build/check table
Input port
Record sender's address and input port
Incoming frame

Figure 5.5 Bridge functions.

is more than the new center value, it must be in the lower half of that half of the table. The search then divides the quarter in which the address is located into halves and repeats the procedure. The maximum number of divisions to perform a complete search is $\log_2 N + 1$, where N is the number of entries in the table.

Binary searching is efficient and can be implemented in special-purpose silicon chips called *application-specific integrated circuits* (ASICs). It relies on having a numerically ordered table. Since the table cannot be used for searching while being updated and reordered, two copies are maintained that can be interchanged as convenient—one for updating and reordering, and the other for searching. A second technique uses *hashing*, which is a procedure that maps address space into a smaller *pointer* space so that an address search is started by searching the smaller pointer field. The *hashing* function must produce a consistent hash value for the same address and, for any arbitrary set of addresses, produce an approximately uniform distribution of pointers.

A way of providing a hash function is to use the *cyclic redundancy checking* (CRC) process. Normally, the entire frame is divided by a prime number to produce the *frame check sequence* (FCS). During the procedure, the first 48 bits to be divided are the destination address. At the end of this interval, the result will be a pseudorandom function related to the destination address. By using one or two bytes from this number to represent it, the first stage search can be reduced to searching for an 8-bit or 16-bit number in 256 or 65,536 locations. The hash numbers are said to identify *hash buckets*; each contains approximately *M*/256 or *M*/65,536 destination addresses (where *M* is the number of destination addresses in the table). Another technique for accessing the table of addresses and ports makes use of *content addressable memory* (CAM), which is a silicon-intensive solution that employs the content (hardware address of destination) as the key for retrieving associated data (e.g., port to which destination is attached).

Content-addressable memory is *hard-wired* and responds instantly to a request (identified by the destination address) with information concerning the port to which the destination device is attached. Such memory chips are expensive and have a limited storage capacity.

5.2.2 Bridging Dissimilar LANs

Figure 5.6 shows an arrangement in which a bridge is used to create a catenet of one FDDI, two Token Rings, and two Ethernet LANs. As mentioned before (Figure 5.3), the bridge receives copies of all frames sent on each network. The table lists the 6-byte MAC addresses of all stations and the number of the port to which each station is connected. The ports are equipped so that they are legitimate stations on the LANs to which they are attached. The question is: Can stations using different LAN technologies communicate transparently, that is, as if they were on the same LAN? The answer is: with some difficulty.

A comparison of Figures 3.3, 3.5, 3.11, and 3.13 in Chapter 3 and the tables in Appendix B shows that LAN types:

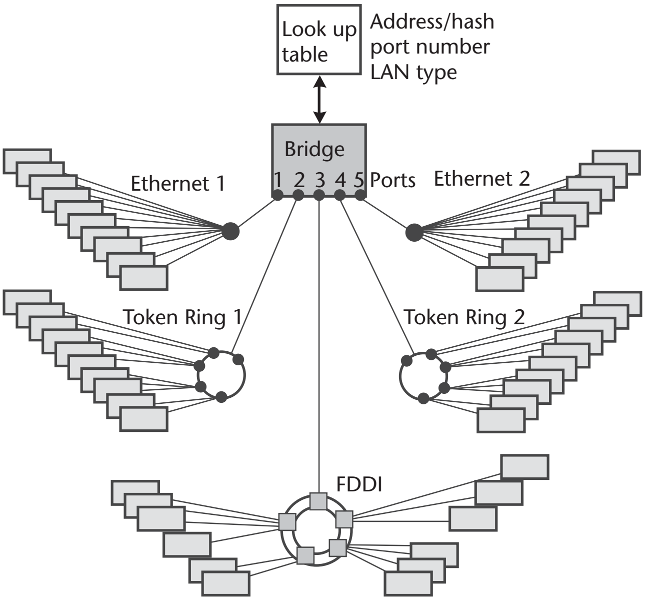


Figure 5.6 Bridging dissimilar LANs.

- Differ with respect to medium access controls, frame formats, frame semantics (i.e., the meaning of the fields within the frame), and frame lengths.
- Use the same 6-byte globally unique addresses administered by a single authority (IEEE).
- Use the same 4-byte frame check sequence procedure.
- May use fields whose equivalents do not exist in other LANs.

Furthermore, the differences and similarities may depend on the upper-layer protocol that is in use.

5.2.2.1 Translating Bridge

To allow a bridge to connect dissimilar LANs, solutions must be worked out for translating between the six dissimilar pairs of LANs formed from Classic Ethernet, IEEE 802.3 Ethernet, Token Ring, and FDDI. Table 5.1 shows the differences between frames carrying IP datagrams or address resolution (ARP) messages. A *translating* bridge will resolve them as follows.

- *Preamble and starting delimiter* can be discarded or added by the bridge, as required.
- *Access control* is peculiar to Token Ring. As required, the bridge can generate it. This information is not passed to other LANs.
- *Frame control* is peculiar to Token Ring and FDDI. It distinguishes between management and data frames. Management frames remain on the ring; only data frames are bridged. In addition, 2-byte addresses occur in FDDI, but not in other LANs. Thus, the bridge can to generate a frame control byte when needed.
- *Destination* and *source addresses* are 6-byte unique identifiers. All LANs use the same format, although storing them requires adherence to big Endian or little Endian rules.
- *Type/length* fields occur in Ethernets. For Ethernet, the type field is ≥0×05-DC and is the same as EtherType in IEEE 802.3, Token Ring, and FDDI LANs. For IEEE 802.3, the length field is <1,500 bytes. The bridge can calculate it readily.
- *Destination* and *source SAPs* are the same for IEEE 802.3, Token Ring, and FDDI LANs. They are not used in Ethernet.
- *Control* is not used in Ethernet. It is the same for IEEE 802.3, Token Ring, and FDDI LANs.
- *Organization code* is not used in Ethernet. It is the same for IEEE 802.3, Token Ring, and FDDI LANs.
- *EtherType* is the same for IEEE 802.3, Token Ring, and FDDI LANs. In Ethernet, it is entered in the type field.
- *Payload* has a maximum length that is different for each LAN. Forwarding a frame that is longer than the destination LAN, or intermediate LANs, can process will result in one of the bridges discarding it. Segmenting a large frame

Table 5.1 Comparison of Frames on Different LANs

Field	*Size*	*Ethernet*	*IEEE 802.3*	*Token Ring*	*FDDI*
Preamble	Variable	0×AA-AA-AA-AA-AA-AA-AA-AB	0×AA-AA-AA-AA-AA-AA-AA-AA	No	0×AA-AA
MAC Header					
Starting delimiter	1 byte	No	0×AB	JK	JK
Access control	1 byte	No	No	Yes	No
Frame control	1 byte	No	No	Yes	01xxxxxx
Destination address	6 bytes	Yes	Yes	Yes	Yes
Source address	6 bytes	Yes	Yes	Yes	Yes
Type/length	2 bytes	Type: 0×08-00 or 0×08-06	Length: n<1,500 (*i.e.*, n≤0×05-*DC*)	No	No
LLC Header					
Destination SAP	1 byte	No	0×AA	0×AA	0×AA
Source SAP	1 byte	No	0×AA	0×AA	0×AA
Control	1 byte	No	0×03	0×03	0×03
SNAP Header					
Organization code	3 bytes	No	0×00-00-00	0×00-00-00	0×00-00-00
EtherType	2 bytes	No	0×08-00 or 0×08-06	0×08-00 or 0×08-06	0×08-00 or 0×08-06
Payload					
IP datagram	Variable	46≤n≤1,500	38≤n≤1,492	0≤n≤4,464 or 0≤n≤17,914	0≤n≤4,352
MAC Trailer					
Frame check sequence (FCS)	4 bytes	33-bit generating function	33-bit generating function	33-bit generating function	33-bit generating function
Ending delimiter	1 byte	No	No	JK	JK
Frame status	1 byte	No	No	Yes	Yes

Type or EtherType: 0×08-00 designates Internet Protocol (IP); 0×08-06 designates Address Resolution Protocol (ARP).

to several smaller frames will be ineffective since the destination station is unlikely to be able to reassemble the segments. However, segmentation and reassembly of IP packets are possible using the Internet layer.

- *Frame check sequence* is calculated the same for all LANs. To reflect changes made in the translation, the bridge must recalculate it.
- *Ending delimiter* can be discarded or added by the bridge, as required.
- *Frame status* is used by Token Ring and FDDI. When transferring frames from Token Ring or FDDI, the bridge can stand as proxy for the destination and set the address recognized (1) and frame copied (1) bits. (Some object to this strategy because it means only that the frame reached the bridge. It does not signify delivery to the destination. Nor does it indicate that the destination is in service.) When transferring Ethernet frames to Token Ring or FDDI, the bridge can create a frame status byte with 0s for the address recognized and frame copied bits.

With care, then, when TCP/IP is used, a *translating* bridge can connect dissimilar LANs and implement virtually transparent transfers between them.

5.2.2.2 Encapsulating Bridge

Under some conditions, rather than translate frames to pass them across a foreign LAN, they can be encapsulated in a frame that is compatible with the foreign LAN. Thus, Figure 5.7 shows LANs connected to bridges that are connected to an FDDI LAN. It serves as the backbone for this network. To send a frame from Ethernet 1 to Ethernet 2, the bridge places it in the payload section of an FDDI frame that carries the addresses of the appropriate ports on the FDDI ring. When the frame arrives at the FDDI destination port, it is stripped of FDDI information and forwarded to the destination bridge. To accomplish this routing, a mechanism must be in place that permits sharing of connection data for the FDDI ports. Information concerning the entrance and exit ports on the FDDI LAN is needed by the bridge to be able to enter sending and destination addresses in the FDDI frame. To send a frame from Token Ring 1 to Ethernet 2 in Figure 5.3, the sending bridge will translate from Token Ring to Ethernet format, and then encapsulate the Ethernet frame in an FDDI frame.

Simple encapsulation (not translation and encapsulation) allows the original frame to be carried through the network from end-to-end. This includes the original FCS. It will detect errors introduced during processing within the network. When translation and encapsulation are required, the bridge recalculates the FCS. Under this circumstance, any error introduced at the bridge will not be found.

5.2.2.3 Loops and Spanning Trees

As more and more networks are bridged together to create a common work environment, chances increase that there will be more than one path between any two stations. Multiple paths raise the possibility that some traffic will be duplicated and some traffic may end up in loops. Left on their own, the loops and duplications will degrade network performance and may create deadlock in localized areas of the

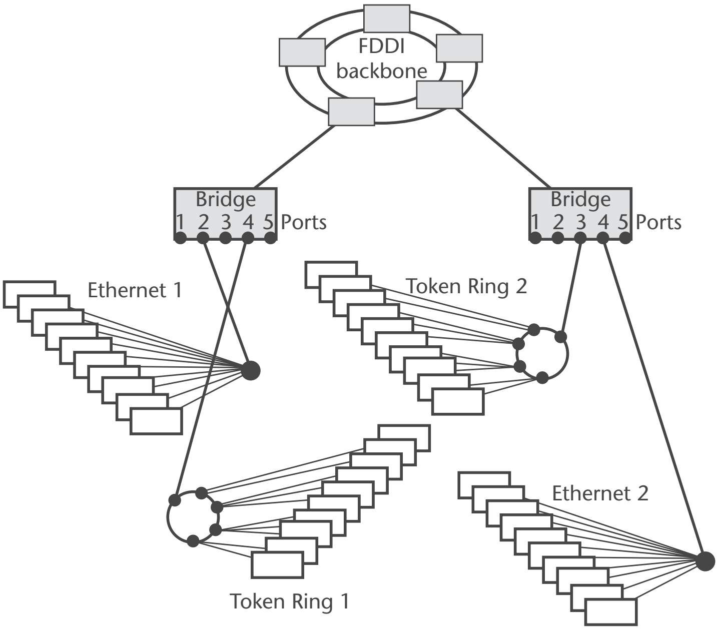


Figure 5.7 Encapsulating bridges.

catenet. To prevent this from happening, IEEE 802.1d specifies a *Spanning Tree Protocol* (STP) that can be invoked to ensure frames sent between one station and another use the single, most efficient (least cost) path. If that path fails, STP configures a new least cost path. By doing this, STP eliminates active loops in a bridged catenet.

What measure shall we use to determine efficiency? IEEE has said the cost of a given link is inversely proportional to the data rate. The faster the final path, the more efficient and more costly will be the transfer.

5.2.2.4 Source Routing

In Token Ring and FDDI catenets, a technique known as *source routing* is available. Before a communication session, the source station *discovers* the routes to each station with which it is likely to communicate. During the session the source station selects the least cost route and inserts this routing information immediately following the source address. In addition, the source sets the first bit in the first byte of the source address to 1 to indicate the frame carries source routing information. Nodes along the route read the information and route the frame accordingly. Up to 14 segments can be specified in the *route descriptors* field. Each segment terminates on a node attached to a particular ring. (See Appendix B for specific field information.) For routes that are not source-routed, a Spanning Tree Protocol can be invoked.

As its name implies, source routing is a source-directed function. Route discovery, route selection (if more than one route is available), and inserting in the frame the rings and bridges of the chosen route are all done by the source. Individual bridges are unaware of the route; they do as instructed by the frame information. In contrast, when implementing spanning tree, the bridges do the discovery and selection; the source is unaware of the route.

5.3 Routing

Routing is the process of forwarding unicast or multicast packets from a sending host to a destination host or hosts. It employs a node that furnishes the physical and logical connections between two networks so that packets are forwarded along a path that connects the sending host to the destination host. At each router, forwarding is accomplished in the Internet layer and may require different network interface layers to match the characteristics of the input and output networks. Each router *advertises* its status and capabilities and *discovers* the status and capabilities of its neighbors. Routers make forwarding decisions based on the contents of their local routing tables.

When WANs and LANs form an internetwork, network addresses, segmentation and reassembly, and other capabilities will be required to route frames. These are functions contained in the Internet layer. Figure 5.8 shows the principle of routing. If the destination host (Node 2) is on the same network as the sender (Node 1), the sending node resolves the MAC address of the destination and delivers the packet(s) *directly*. If the destination node (Node 3) is not on a directly attached network segment, the sending node makes an *indirect* delivery. It forwards the

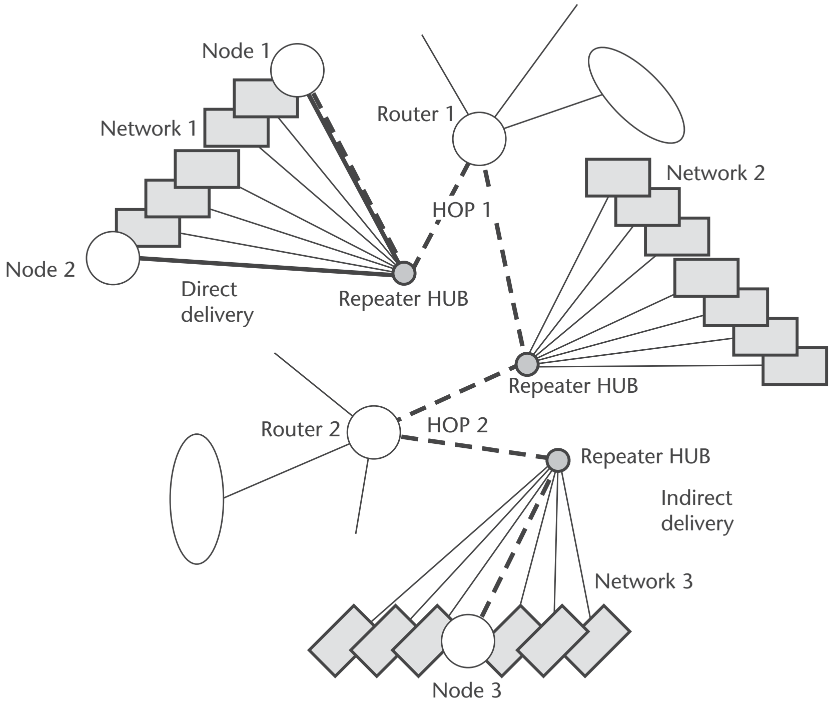


Figure 5.8 Direct and indirect delivery.

packet(s) through a series of routers (Router 1 and Router 2) to the destination node.

5.3.1 Routing over Broadcast Links

A *broadcast* link has more than two nodes on the same network segment. Ethernet links, for example, are broadcast links. Unicast, multicast, and broadcast packets sent by any node are received by all nodes on the segment. For a given forwarding IP address, ARP is used to resolve the intermediate (or final) destination MAC address. For broadcast frames, the address all-1s is used. It needs no resolution, and is not forwarded by routers (because time to live is set to 1). If they were, we would quickly fill up the LANs with broadcast messages.

5.3.2 Routing over Point-to-Point Links

A *point-to-point* link has only two nodes. Leased-line and circuit-switched WAN links, such as analog telephone lines, T-carrier, and ISDN, are examples of point-to-point links. Because there are only two nodes, and if one is the final destination, the IP address is irrelevant and ARP is not needed to resolve the destination MAC address. If the receiver is not the final destination, the IP destination address will be required to obtain further MAC addresses.

5.3.3 Routing over Nonbroadcast Multiple Access Links

Nonbroadcast multiple access (NBMA) links are characterized by virtual circuits that support more than two nodes over the same bearer. They provide point-to-point unicast services. Packet-switched WAN links such as X.25, frame relay, and ATM are examples of NBMA links. The forwarding network address for the route in the routing table is mapped to the virtual circuit identifier using a table maintained by the sending node. Inverse ARP is used to discover the network addresses of nodes on the other ends of the virtual circuits.

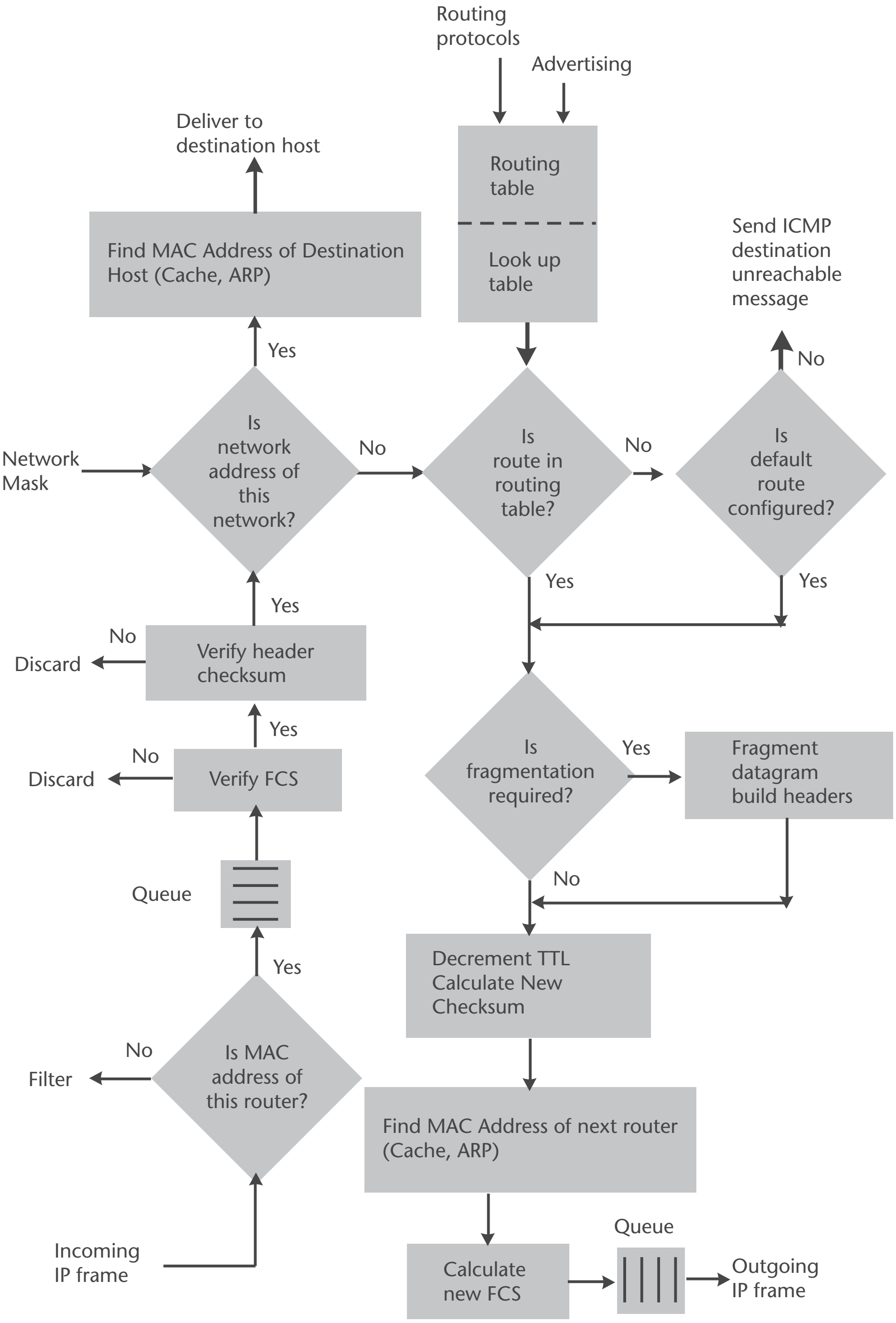

Routing protocols
Advertising
Deliver to destination host
Routing table
Look up table
Find MAC Address of Destination Host (Cache, ARP)
Send ICMP destination unreachable message
Yes
No
Network Mask
Is network address of this network?
No
Is route in routing table?
No
Is default route configured?
Yes
Yes
Yes
No
Discard
Verify header checksum
Yes
No
Discard
Verify FCS
Is fragmentation required?
Yes
Fragment datagram build headers
Queue
No
Yes
Decrement TTL Calculate New Checksum
No
Filter
Is MAC address of this router?
Find MAC Address of next router (Cache, ARP)
Queue
Incoming IP frame
Calculate new FCS
Outgoing IP frame

Figure 5.9 Router functions.

5.3.4 Router

Figure 5.9 is a functional diagram of a router. A database of routes is stored and maintained by all routers. Called a *routing table*, it contains information concerning routes between the node owning the table and the potential destination nodes. At a minimum it includes the destination ID, intermediate interface ID(s) and forwarding address(es), and information to distinguish the best route to use when multiple routes are possible. It is significantly more complex than the table maintained by bridging devices. However, its extent is limited to the immediately reachable nodes that surround it, so that it is significantly smaller. Searching a routing table is a relatively simple task. For each route, a typical routing table will include the following fields:

- *Destination address:* The IP address of the node to which the source directs the packet to be delivered. For *direct* deliveries, the destination IP address carries the same network ID as the router. For *indirect* deliveries, the destination address does not carry the same network ID as the router, and the datagram is sent to the forwarding address contained in the table entry.
- *Network mask:* A bit mask is used to determine the network ID of the destination IP address. An IP datagram with a destination IP address that contains the specific network ID for this route will be forwarded over it.
- *Forwarding IP address:* For *indirect* deliveries, the IP address of a directly reachable router to which the IP datagram is forwarded for eventual delivery to the destination IP address. The IP address to which the IP datagram is to be forwarded on its next hop.

While the routing table contains information on all routes within the router's purview, the router maintains a separate *look-up* table in which all recently used routes are recorded. If they are not used again within a specified time, they are purged. Because it does not have to search the larger routing table for directions, the router can provide rapid service if the routes are called for again before time runs out. Priority routes can be stored permanently in the look-up table.

5.3.5 Static Routing

Static routing employs manually configured routes. Because of the work involved, static routing is limited to relatively small networks. Static routing does not *scale* well. Often, static routes are used to connect to an ISP router. To make the destination unambiguous, a network mask or masks accompanies each route. By definition, a static router cannot adjust its routing table. That can only be done by manual intervention. Therefore, a static router is unable to react to the state of contiguous routers, and neighboring routers cannot update the static router's table.

5.3.6 Dynamic Routing

Dynamic routers employ routing protocols to dynamically update their routing tables. When a route becomes unreachable, it is removed from the routing table. When a router becomes unreachable, alternate routes are worked out and shared between routers. In a dynamic routing environment, routers are in regular touch with each other concerning the state and capabilities of the network. Two common routing protocols used in autonomous networks are *Routing Information Protocol* (RIP) and *Open Shortest Path First* (OSPF).

5.3.6.1 Routing Information Protocol (RIP)

RIP is a simple routing protocol with a periodic route-advertising routine that can be used in small- to medium-size networks. RIP is described as a *distance vector* routing protocol. The distance is the number of hops between the router and a specific network ID. RIP recognizes a maximum distance of 15 hops. Destinations with 16 or more hops are described as *unreachable*.

When an RIP router is initialized, it announces the routes in its table to all interfaces. In RIPv2, to support *classless* addressing, the announcement includes a network ID and a network mask. The router continues with an RIP general request to all interfaces. All routers on the same network segment as the router sending the request respond with the contents of their routing tables. With these, the requesting router builds its initial routing table. Learned routes persist for 3 minutes (default value) before being removed by RIP from the routing table. After initialization, the RIP router announces the routes in its routing table every 30 seconds (default value).

5.3.6.2 Open Shortest Path First (OSPF)

OSPF is described as a *link state* routing protocol and a classless routing protocol. Routing information is disseminated as *link state advertisements* (LSAs) that contain the IDs of connected networks, network masks, and the cost. The cost of each router interface is a dimensionless number assigned by the network administrator. It can include delay, bandwidth, and monetary cost.

The LSA of each OSPF router is distributed throughout the network through logical relationships between neighboring routers known as *adjacencies*. When all current LSAs have been disseminated, the network is described as *converged*. Based on the link state database, OSPF calculates the lowest-cost path for each route. They become OSPF routes in the IP routing table.

To control the size of the link state database, OSPF allows contiguous networks to be grouped into *areas*. A router at the border of an OSPF area can be designated an *area border* router. Reached by a single route from outside routers, it aggregates routing information for the area. The formation of areas and the use of route aggregation permit OSPF networks to scale gracefully to large IP networks.

5.3.7 Border Gateway Routing

The foregoing discussion of routing has assumed it takes place in contiguous networks administered by a single entity (such as an enterprise or an ISP). In these *autonomous* networks, the operator stipulates the internal procedures and formats. The *internal* routers share common routing policies and can communicate with each other without difficulty. What if an autonomous network needs to communicate outside itself with autonomous networks operated by other administrators? This is accomplished by *border* routers running *Border Gateway Protocol* (BGP).

BGP is a dynamic routing protocol. When running between autonomous networks, BGP is called *external BGP*. It learns routes from internal routers (using static routing, RIP, or OSPF) and announces them to border gateway peers. BGP neighbors exchange full routing information when a TCP connection is first established between them. Thereafter, changes are advertised as they occur. If BGP receives multiple advertisements for the same route, using a set of criteria based on local circumstances, it selects the *best* path, puts it in its routing table, and advertises it to its peers. In addition, BGP is used within an autonomous network to distribute information used by internal routers to direct traffic to the *best* border router. In this application it is called *internal BGP*.

5.3.8 Intermediate System-to-Intermediate System

An intermediate system is OSI terminology for a router. *Intermediate System-to-Intermediate System* (IS-IS) was developed by OSI as part of the OSI protocol stack. Because it is scalable to very large networks, IS-IS is used by large ISPs to route traffic to backbones and other Internet service providers. Like OSPF, IS-IS recognizes adjacencies, regularly advertises link-state information, and supports point-to-point and broadcast applications.

5.4 Virtual LANs

Significant changes in operation and topology have been achieved in Ethernet networks by substituting repeatered hubs in place of a shared bus, substituting switched hubs to provide individual station-to-station connections, adding duplex capability to allow each station to send and receive simultaneously, and increasing speeds from 10 Mbps to 1,000 Mbps. Of the shared cable network with access governed by CSMA/CD that is described at the beginning of Chapter 3, only the frame format remains. However, once installed and configured, changes in the number and distribution of stations or subnetworks still require changing the physical connections that define the catenet. *Virtual LAN* technology takes the next step. Irrespective of their position in the catenet, a given set of stations is able to communicate as if they are connected in a dedicated LAN. At the expense of having to logically define the associations between new and existing stations, or redefine the associations between existing stations, additions and moves can be made without changing physical connections.

5.4.1 Tags

One way to form a *virtual LAN* (VLAN) is to add an identifying *tag* to each frame and provide routers and switches with the ability to forward frames to VLANs based on these tags.

5.4.1.1 What Is a Tag?

For an IEEE 802.3 format frame encapsulating an IP datagram, it is a 2-byte field inserted between the EtherType field of the SNAP header and the payload. Shown in Appendix B, the EtherType field contains the VLAN protocol identifier—0×81-00. It indicates the frame is VLAN-tagged, and the next 2 bytes contain tag control information. In the *tag control information field* (TCIF):

- The first 4 bits in the first byte of TCIF, and the entire second byte, are used to identify the VLAN. Reserving the all 0s and all 1s values for special purposes, a total of 4,094 separate VLANs can be distinguished.
- Bit 5 of the first byte of TCIF is the *Canonical Format Indicator*. Set to 0, it shows that the bit ordering is *little Endian*; set to 1, it shows that the bit ordering is *big Endian*.
- Bits 6, 7, and 8 of the first byte of TCIF are a *priority* field. With values from 0 through 7, it indicates the user's priority for the frame. (See Appendix B for more information.)

5.4.1.2 Tagging

If the stations are *VLAN-aware*, the tag can be placed in the frame when the frame is first generated. In addition, source routing instructions can be attached to ensure that the frame is forwarded by a specific route through the intervening catenet. With the same format as Token Ring source routing, up to 14 *route descriptors* are entered in the frame. (See Appendix B for more information.) A 2-byte routing control field that contains data to assist the nodes to route the frame properly precedes the route descriptors. Tags are used with Ethernet, Token Ring, and FDDI formatted frames. Because Ethernet reads bits little Endian and Token Ring and FDDI read bits big Endian, great attention must be paid to the nature of the data stream, and its history. All three styles of LANs read bytes left to right (or top to bottom, if written in stacks).

The sending station is the obvious location at which to introduce a tag. Where else is more information readily available? True enough, but to do this will require modifying all terminals currently in use—even though many of them may not operate routinely in a VLAN environment. Only in new terminals is adding tags at the sending station a practical proposition.

Where, then, to introduce tags? Figure 5.10 shows a popular solution. A catenet of several LANs is tied together in an *enterprise network* by a multiswitch backbone. The backbone switches form two subsystems. Frames are fed from the LANs to the backbone through *edge switches*. In turn, the edge switches pass them on to *core switches* that move the frames over the backbone to other edge switches. Using the parlance of the VLAN environment, the edge and core switches are said to be *VLAN-aware*. The edge switches do the tagging, and the core switches direct the tagged frames over the backbone to the destination edge switches. The receiving edge switches *untag* the frames and send them to the LANs on which the target stations reside. The majority of stations remain *VLAN-unaware*. Only the backbone, which is responsible for moving frames between LANs, has to deal with tags.

Figure 5.11 shows how the catenet of Figure 5.10 can be divided into four virtual LANs by tags applied by edge switches. While the stations retain their physical connections, by means of tag identifiers they can be associated in new ways. In Figures 5.10 and 5.11, the perimeter LANs may be bridged catenets.

To successfully tag the frames, edge switches must:

- Read specific fields in the frame.
- Analyze the data by employing the classification rules provided by the network administrator.

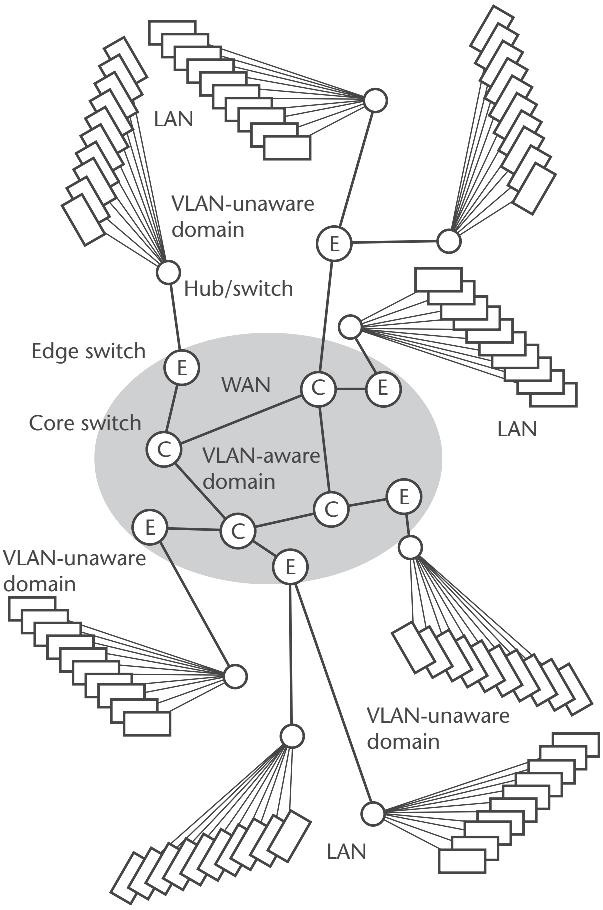


Figure 5.10 VLAN domains.

- Use the results to associate the frame with a particular VLAN.
- Insert the appropriate tag information in the frame.

Quantities such as the port number, source address, protocol type, application identifier, and other data will be the basis for assigning a VLAN identifier. Once the tag is in place, the edge switch calculates a new FCS and sends the frame over the backbone to the edge switch serving the LAN on which the VLAN station or stations exist(s). If the stations are VLAN-unaware, the terminating edge switch will remove the tag, recalculate the FCS, and send the frame to the hub. If it is a switched hub, the frame will be directed to the destination station(s) only. If it is a repeatered hub, the frame will be directed to all stations attached to the hub.

In addition, the edge switch collects information with which to extend and check its database. To make sensible decisions, the switch needs to know the topological and membership status of all nodes with which it is likely to have contact. How better to obtain this than recording the origins and destinations of traffic in the network? Tagging can add 32 bytes to the length of the frame. This does not seem to cause a problem with most equipment. As a matter of good engineering practice, the designs have more than minimum-size buffers.

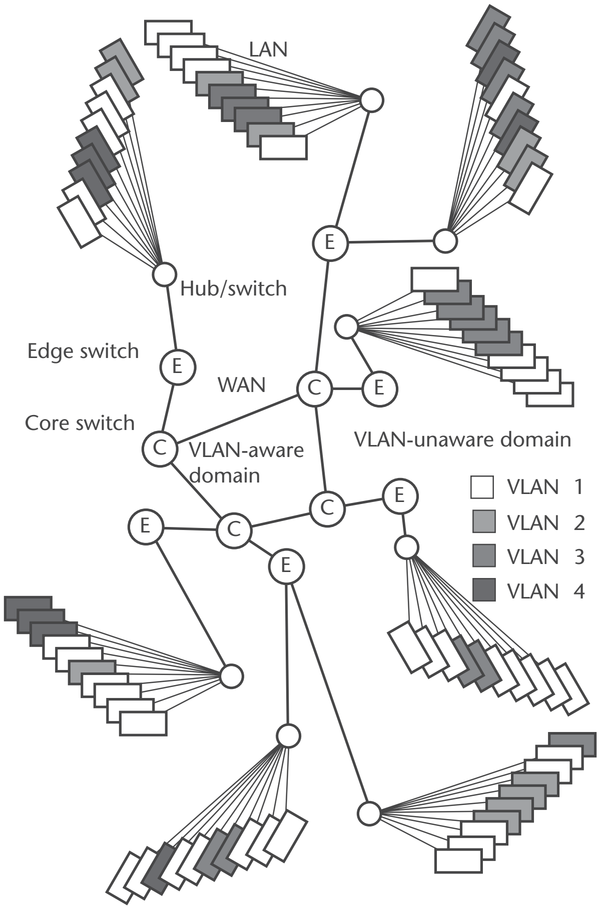


Figure 5.11 Four VLANs.

5.4.1.3 Implicit and Explicit Tags

It is customary to distinguish between implicit and explicit tags.

- *Implicit tag:* A tag implied by the contents of an untagged frame generated by a VLAN-unaware station or switch. An implicit tag resides anonymously in a normal frame emitted by a conventional station, or forwarded by a VLAN-unaware device. The frame has the potential of being tagged when a VLAN-aware device processes it. Hence, the frame is *implicitly* tagged.
- *Explicit tag:* A tag created by applying VLAN association rules to frame data. Explicit tags are created by VLAN-aware stations or by the first VLAN-aware switch. They must be removed before passing the frame to a tag-unaware device. Adding or removing a tag requires the tag-aware device to calculate a new FCS value.

5.4.2 Edge and Core Switches

The switches that connect devices in VLAN-unaware domains to devices in VLAN-aware domains are known as *edge* switches. The devices in the VLAN-unaware zone(s) are likely to be LAN's or bridged catenets. The devices in the VLAN-aware zone are known as core switches.

5.4.2.1 Switch Operation

To forward an untagged frame, the switch converts the implicit tag it carries to an explicit tag using the rules it has been given, and forwards it on the basis of this tag. If there is no basis for explicit tagging, the switch is likely to assign the frame to a default port. If it is available, the switch will use *explicit routing information* (ERI) to forward the frame along a tested route. To forward a tagged frame to the members of the frame's VLAN, the switch must know which of its ports connect to the LANs that host members of the VLAN identified by the tag. To prevent misunderstandings, if the receiving entity is tag-unaware, the terminating edge switch must strip the tag from the frame before forwarding it.

5.4.2.2 Ingress, Progress, and Egress

The actions of edge and core switches can be described in three phases. Known as *ingress*, *progress*, and *egress* processes, on each incoming port, they perform the following functions:

- The *ingress process* uses the following to tag frames and discard those assigned to VLANs not recognized by the incoming port:
  - *Acceptable frame filter:* A logical filter with two states. It allows all received frames to proceed to the rules module, or restricts passage to only those frames that are tagged. In this case, frames without tags are discarded.
  - *Rules module:* VLAN association rules are also known as ingress rules. They are applied to incoming frames and are designed and configured by network administrators. They are distributed automatically to VLAN-aware switches. Simple rules are based on port ID, MAC address, protocol type, application, and so forth. More complex rules require the use of a microprocessor or finite-state machine to parse the relevant information fields. If the received frame is already tagged it is simply necessary to assign it to the VLAN indicated on the tag. If the incoming frame is untagged, one or more of the association rules are used to assign it to a single VLAN. If a VLAN cannot be assigned using these rules, the frame is tagged with a default identifier.
  - *Ingress filter:* A filter configured to discard frames assigned to VLANs not recognized by the incoming port.
- The *progress process* forwards the tagged frame to the egress port and maintains the switching database. Frames are transported through a switching fabric and queued for transmission. The egress port is determined by the VLAN identifier and the MAC address of the destination. By observing traffic flow, the switch maps VLANs to ports to ensure an up-to-date database.
- The *egress process* uses the following to determine whether, and in what format (tagged or untagged), to transmit the frames:

- *Egress rules:* Determine if every station that is a member of the VLAN to which the frame is sent is tag-aware. If not, strips the tag from the frame.
- *Egress filter:* Discards frames because the VLAN identified in the frame is not connected to the output port. In addition, may discard or correct frames because bit ordering is not correct for the destination LAN.

5.5 Multiprotocol Label Switching

Multiprotocol label switching (MPLS) is a project of IETF designed to address problems of scalability, speed, and quality of service in today and tomorrow's networks. Intended to extend to various packet-based technologies, the work has concentrated on speeding up the passage of IP frames across a network consisting of edge routers and core switches on *label switched paths* (LSPs). LSPs are defined by labels located at each intermediate node between the source and destination. Created by the edge router first receiving the data, or by the passage of data through the network, LSPs are said to be *control* driven when they are established before data transport, and *data* driven when predicated on data flow. Sequences of packets between the same sender and receiver follow the same LSP. They are known as a *forwarding equivalence class* (FEC). All receive the treatment afforded the first packet. An LSP is one directional; for duplex working, a second path must be created in the opposite direction.

5.5.1 Label Distribution

Labels are distributed using *Label Distribution Protocol* (LDP), RSVP, OSPF, or BGP. Completion of this action creates a switched path through the network (an LSP) for a class of packets (an FEC) sent to the same destination. Three basic methods are:

- *Topology-based:* A control-driven action. Uses OSPF and BGP routing protocols that have been enhanced to incorporate label creation.
- *Request-based:* A control-driven action. Uses RSVP enhanced to incorporate label creation.
- *Traffic-based:* A data-driven action. Uses the reception of a frame to create and distribute labels with LDP.

LDP is designed to manage label functions. It includes the ability to support routing based on QoS requirements.

5.5.2 Label Location

For MPLS core networks comprised of ATM or frame relay switches, their labels are contained within the network interface headers. For ATM, the label is the combination of virtual path and virtual circuit identifiers (VPI/VCI). For frame relay, it is the *data link connection identifier* (DLCI). For other networks, labels are contained in a 32-bit field known as an *MPLS Shim* situated between the network interface header and the rest of the frame. Figure 5.12 shows labels in the lead position in

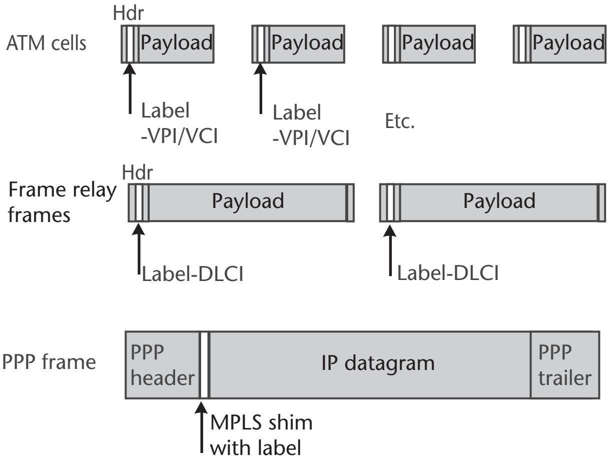


Figure 5.12 MPLS labels.

ATM cells, immediately following the flag in frame relay, and following the network interface header when PPP is used. Labels are placed at the beginning of the packet so that, without having to consult switching tables, the receiving intermediate node can route the packet quickly to the next node. Labels are only locally significant and define one hop. As required, the intermediate routers change the values for the next hop.

5.5.3 MPLS Operation

The action of assigning a specific label to a particular class of packets (FEC) is known as *binding*. Before packet flow begins, decisions to bind labels and FECs are made by edge routers. The binding is stored in a *label information base* (LIB) where it is available to each network node. LDP is responsible for maintaining this database. LSPs are created backwards from destination edge routers to source edge routers. Each node (edge router or core switch) inquires of its downstream neighbor for a label. When the process is completed, an LSP exists across the core network. Negotiations for specific QoS performance are included in the creation of the path.

With a path established, the sending edge router consults the LIB for the first downstream core switch in the LSP, inserts the label for the FEC, and transmits the packet. Subsequent switches read the incoming label, replace it by the outgoing label, and send the packet on its next hop. When the packet reaches the egress side of the destination edge router, the label is removed and the packet is transported to its destination in the usual way.

Whether they are called bridges and routers, or edge and core switches, tags or labels, the subjects I have discussed in this chapter, are key to pervasive commercial operations. Bridges make a common work environment possible and routers create vast, transparent networks. Furthermore, by taking advantage of the frame structure and using tags or labels, most of the drawbacks attendant on deploying and reconfiguring networks can be lessened or eliminated, and transport can be speeded up. There remains a major concern. As the networks expand, and communication becomes simple and acceptable to all users, how can promiscuous users be discouraged, and private information be kept just that? Some remedies are described in the next chapter.

CHAPTER 6

Protecting Enterprise Catenets

There are as many unique data catenets as there are enterprises that build and operate them. Each organization has different users, different objectives, different topologies, and different equipment. Moreover, they have different numbers of users with different skill levels that work with different applications. In addition, they are likely to have mixtures of equipment that reflect their historical evolution. Some still operate with a base of 10 Mbps shared medium Ethernets. Others will have 100 Mbps repeatered and switched hubs supporting desktop operations fed by 1,000-Mbps servers. Yet others will have Ethernets, Token Rings, and FDDI networks operating at various speeds. Transport will be by twisted pairs, optical fiber, or radio at speeds from 28.8 kbit/s to 622.08 Mbps. Because of the multitude of possibilities, no two catenets are exactly alike.

6.1 Operating Environment

Consider the environment in which enterprise catenets operate. If we define a *catenet* as several individual networks linked together to facilitate the execution of distributed data operations, and we define a network as a (complex) tool that facilitates the execution of distributed data applications, we have a description that does not depend on the business purpose for which the owning enterprise exists. Furthermore, we can generalize the nature of the data traffic that flows in the network. File transfers, application sharing, e-mail, and printer sharing produce the majority of the traffic. These activities are manifest by bursts of data separated by periods of silence.

6.1.1 Enterprise Catenet

Figure 6.1 shows an enterprise catenet. It is a hierarchical network with four levels. They are designated as follows.

- *Desktop:* Several interconnected clients, servers, and printer stations, perhaps on a single floor. Consists of individual stations connected by a LAN (Ethernet or Token Ring) that employs a common bus or a repeatered or switched hub. Each port may support a single user or a small number of end users. A desktop network is the lowest level of the catenet hierarchy.
- *Workgroup:* Interconnected desktop networks (LANs) that may be situated in several areas (floors, bays, and so forth). Consists of two or more desktop

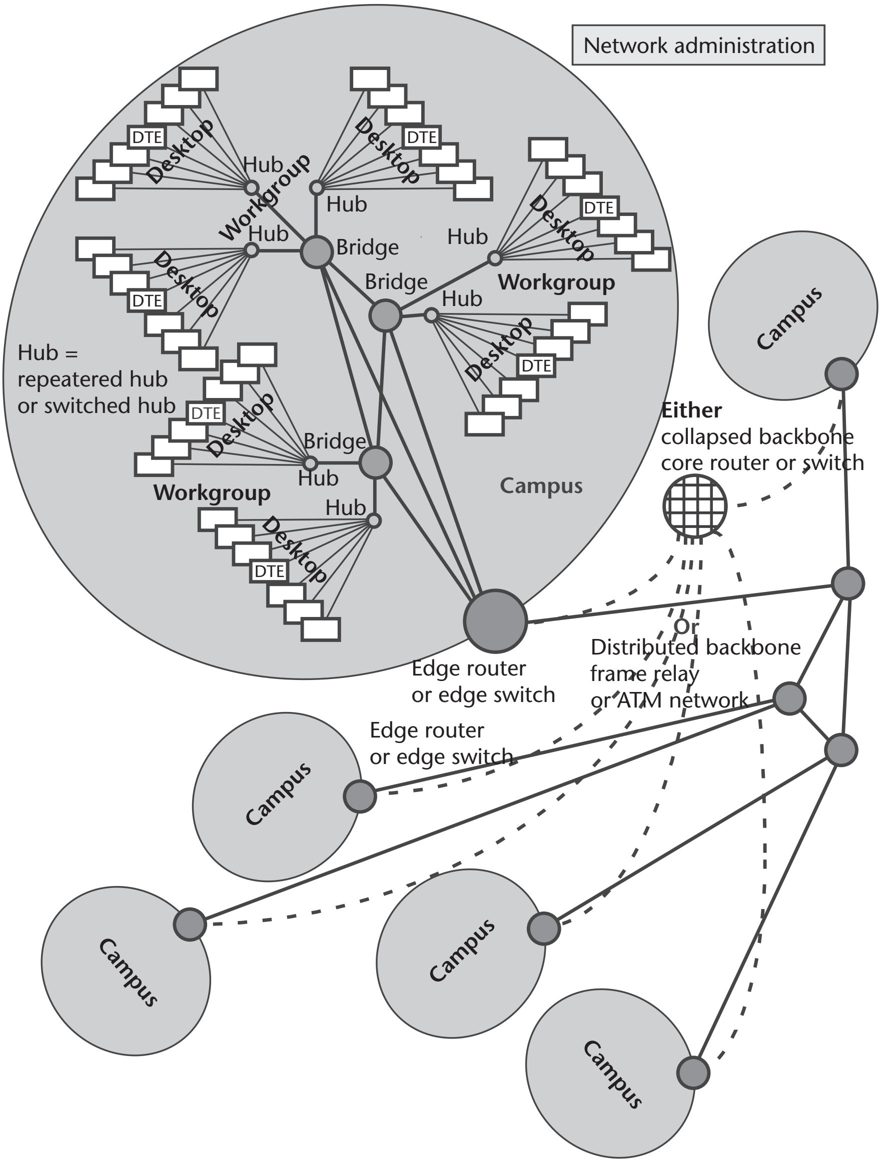

Network administration
Desktop
DTE
Hub
Workgroup
Bridge
Hub = repeatered hub or switched hub
Campus
Edge router or edge switch
Either collapsed backbone core router or switch
Or
Distributed backbone frame relay or ATM network

Figure 6.1 Enterprise catenet.

networks bridged together. Provides intercommunication among desktop networks in the workgroup.

- *Campus:* Interconnects workgroup networks within a single location. Consists of one or more workgroup networks bridged together and connected to an edge switch or edge router. Provides communication among workgroup bridges on a campus and facilitates communication to other campus networks.
- *Backbone:* Interconnects campus networks. The connection may be *distributed* or *collapsed*:
  - *Distributed backbone:* A (wide area) network (e.g., frame relay or ATM network) that interconnects campus networks to create an enterprise

catenet. It provides moderate to high bandwidth over moderate to long distances.

- *Collapsed backbone:* A single core switch or router that interconnects all campus networks in the enterprise catenet. It can provide very large aggregate bandwidth.

In Figure 6.1, both styles of backbone are shown. The distributed backbone is represented as a set of nodes in a frame relay or ATM network. It might be suited to a larger corporation with worldwide operations. The collapsed backbone is a single switch that can give faster service to a smaller network. They are shown in the same diagram for comparison purposes. It is unlikely they would be used in tandem.

6.1.2 Interconnections

In Figure 6.1, the campus networks are likely to be owned (or leased) by the enterprise. The links, bridges, hubs, and desktop stations are focused on producing the value-added services the enterprise provides. In linking the campus networks together, the enterprise owner may use:

- *Private* facilities owned or leased exclusively by the enterprise. This arrangement prevents the acquisition of company data by external operators and preserves its confidentiality for the enterprise.
- *Leased* facilities, such as permanent virtual circuits from a frame relay network provider or virtual circuits from an ATM provider. This arrangement preserves confidentiality with respect to most external operators. It is probably no impediment for a determined hacker.
- *Internet* facilities, the arrangement of which links the campus networks to the world. As soon as a public connection is added to a private network, it becomes vulnerable to unauthorized access by the curious, the mischievous, and the criminally motivated. Special techniques must be employed to restore privacy yet retain the ability to use the Internet to the advantage of the enterprise.

The combination of campus networks and collapsed backbone shown in Figure 6.1 could be an example of a catenet formed from private facilities. All the campus edge routers/switches are connected by a single core router/switch. The entire network has one purpose—to further the internal communications of the enterprise.

The combination of campus networks and distributed backbone shown in Figure 6.1 could be an example of an enterprise catenet using some leased facilities. The edge switches are connected to core switches in a frame relay or ATM network. In the frame relay network, the enterprise owner has use of specific permanent virtual circuits that interconnect the campus networks. In the ATM network, the enterprise owner has use of certain virtual circuits in defined paths that link the campus networks. As long as the connection tables limit the use of the virtual circuits to frames addressed to terminations in the catenet, the owner will have a catenet that is focused on facilitating the objectives of the enterprise.

With the maturing of the Internet, enterprise catenets need no longer be limited to accepting frames from and delivering them to stations within the enterprise. Now it is possible for communications to span the globe and connect to distant resources. Figure 6.2 shows the campus networks' end routers connected to *Internet service providers* (ISPs) that give access to the Internet. The Internet can be used for interconnecting campus network to campus network, connecting campus networks to sources of public information, and connecting between stations inside and outside the catenet. It is a distributed backbone of immense proportions.

The extension of the catenet to global distances provides the opportunity for enterprise stations to address the stations (clients or servers) in the catenet or stations anywhere within the millions of users in the Internet community. In addition, it gives the opportunity for competitors and others to read (and perhaps sabotage) the data communications of the enterprise.

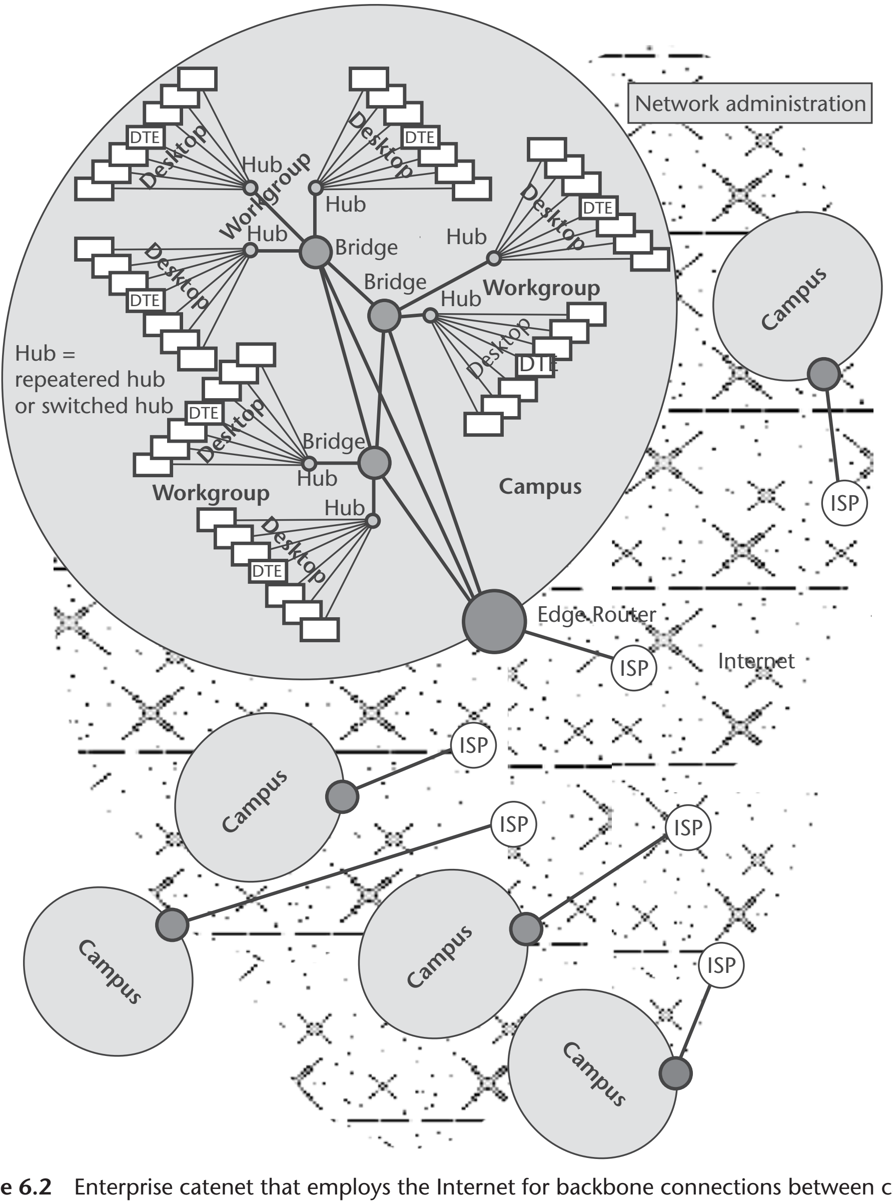


Figure 6.2 Enterprise catenet that employs the Internet for backbone connections between campus networks.

Connecting a private network to the Internet has certain advantages. Among other things, doing so facilitates the acquisition of public information, the exchange of e-mail between enterprise members and persons in other organizations, and the supply of information on enterprise products to persons in other organizations or to members of the public.

In addition, connecting a private network to the Internet has certain disadvantages. Doing so permits enterprise employees to browse the Internet for personal reasons, outsiders to access the enterprise network for illegal purposes, and virus attacks, denial of service, and other nuisances. To restore integrity to a catenet that employs the Internet (or other public network), address translation, proxies, encryption, and encapsulation techniques have been developed.

6.2 Combating Loss of Privacy

Loss of privacy can be countered by simple rules attached to internal addresses, more complex rules known as *proxies* that entail evaluating relationships between frames ,and by creating secure connections between specific stations in the Internet and stations in the private network.

6.2.1 Network Address Translation

In Section 1.6.1, I noted that private IP address spaces have been created for use by organizations. Specifically, they are:

- 10.0.0.0 to 10.255.255.255;
- 172.16.0.0 to 172.31.255.255;
- 192.168.0.0 to 192.168.255.255.

These addresses do not appear in Internet tables. When access to the Internet is required, *network address translation* (NAT) must be performed. It creates an Internet readable address that is used to return data. The principle is shown in Figure 6.3.

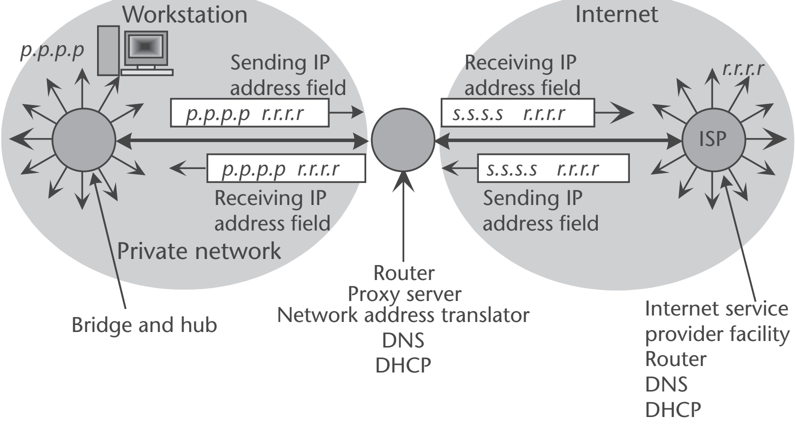


Figure 6.3 Enterprise catenet with network address translation service for connections to the Internet.

Suppose a station with an IP address *p.p.p.p* in the private network wishes to communicate with a station with an IP address *r.r.r.r* in the Internet. The IP address field in the frame sent from the sending station to the edge router will be *p.p.p.p|r.r.r.r→, where p.p.p.p* is the sending address, and *r.r.r.r* is the destination address. Because *p.p.p.p* is not recognized in the Internet, it must be changed at the edge router to a valid Internet address. Suppose this is *s.s.s.s*. On entering the Internet, the frame will have a destination address of *r.r.r.r* and a sending address of *s.s.s.s*. When information is returned, the address field will read ← *s.s.s.s|r.r.r.r* in the Internet, and ← *p.p.p.p|r.r.r.r* in the private network. Because the private addresses do not appear in the public network, they are unknown to the public stations. Thus, knowledge of the topology of the private network is denied to public stations and the task of predators becomes more difficult.

6.2.2 Proxies

In the network world, a proxy is a package of software or hardware that performs a function defined by the proxy giver. A proxy is a rule that is applied to traffic within its purview. Thus, a list and supporting logic for denied destinations of frames from users with certain privileges are a proxy. Situated between the private catenet and the edge router, a proxy server can *filter* frames using lists of sites that are specifically permitted or denied to users with different levels of privilege. Particular sites can be blocked outright, and others can be controlled based on the identity of the user, the service requested, the port, or the IP domain. A proxy server can implement the address translation function. Further, it may provide *domain name system* (DNS) *service*, *Dynamic Host Configuration Protocol* (DHCP) *service*, and other functions. A proxy server can be used at other locations in the private network to restrict or prevent traffic between sections of the catenet. In this application, address translation is not required.

The complexity of the proxies employed depends on the value the network owner places on protecting the products in the private network. In addition, the complexity of the proxies depends on the imagination of the network administrator. Three levels of proxies are:

- *Frame filtering:* After checking the address fields and contents of the frame for keywords, passage of the frame to its destination is permitted or denied. Working from lists, frame filtering is relatively easy to design and relatively fast to execute. It is also relatively crude.
- *Circuit-level filtering:* By observing the grouping of frames, a connection between client and server is detected. Using rules to determine whether the source and destination are *compatible* (i.e., are likely to have legitimate business to transact), the passage of information is permitted or denied. Circuit-level filtering requires more reference information, may not be that difficult to design, but takes longer to execute because of the number of frame evaluations that have to be made.
- *Application-level filtering:* By testing the data contained in frames that constitute a communication by the characteristics of the destination, the acceptability of the communication is determined and the passage of information is

permitted or denied. Application-level filtering can be the most complex strategy. It requires evaluation of the data being passed. Therefore, it must be custom designed for each application. Because it requires the observation of several frames, execution is likely to be slow. If the owner values the data highly enough, the simultaneous application of two or three strategies can be considered.

6.2.3 Tunnels

In Figure 6.2, the campus networks are connected into the enterprise catenet by a distributed backbone formed from Internet circuits. The data they carry is vulnerable to eavesdropping and alteration by wrongdoers. To prevent these acts, the enterprise owner can construct a tunnel between each pair of campus networks. A *tunnel* is a secure temporary connection between two points in an insecure public network.

Because users within each campus network may attempt to eavesdrop and alter messages, tunneling may be extended to the users' interfaces. Figure 6.4 shows a tunnel that connects a secure client in one campus network to a secure server in another campus network. Connections between campus networks are not the only application for this technique. No matter where they are situated, tunneling can be applied between stations that communicate over a public network to create a temporary private connection.

The techniques of encapsulation and encryption are used to create tunnels. *Tunneling* is the action of encapsulating an encrypted datagram inside another data-

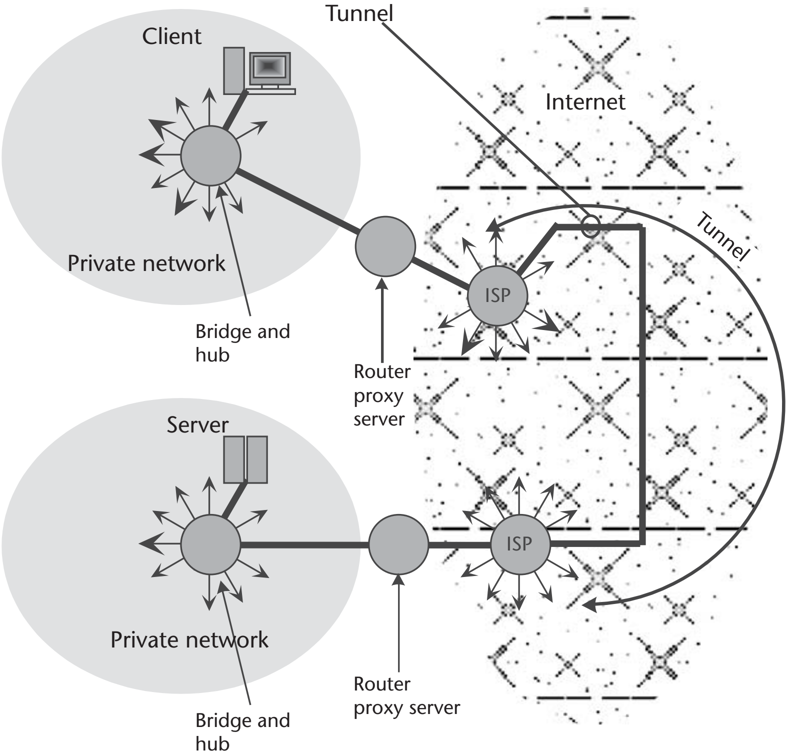


Figure 6.4 Tunnel between private networks.

gram so that it can be forwarded between two points over an insecure temporary connection without revealing its contents.

Figure 6.5 illustrates the concept of tunneling. Data to be sent in a secure way is assembled in an IP datagram by the sending station. It contains the IP network addresses of the sending station and the receiving station. I will call this datagram, D(1). D(1) is encapsulated by a network interface header and trailer, and sent to the router facing the Internet (R1). Here, the header and trailer are stripped from D(1), it is encrypted, and *wrapped* (encapsulated) in a second IP datagram. I will call this datagram D[D(1)]2 to symbolize an encrypted IP datagram [D(1)] encapsulated by a second datagram D(2). D(2) contains the IP address of the router R(2) serving the destination campus network and the IP address of the sending router R(1). At R(2), D[D(1)]2 is decrypted and *unwrapped* (decapsulated) to give D(1). D(1) is encapsulated with network interface header and trailer information and sent on to the destination address it contains.

Remote users who must use a telephone connection, can use this technique. After establishing a normal *dial-up networking* (DUN) connection to a local ISP, the remote user generates an IP datagram addressed to an enterprise destination. This datagram is encapsulated in a PPP frame and may be encrypted. It becomes the *users data* in a second IP datagram addressed to the intranet tunnel router serving the home station. The encapsulated datagram travels from tunnel server to tunnel server on the basis of the network addresses contained in the encapsulated datagram. Thus, an eavesdropper is denied the knowledge of the true origin and destination of the original datagram. At the tunnel server, the original IP datagram is *unwrapped* and forwarded to its destination. In effect, the action of tunneling has created a private connection out of public facilities.

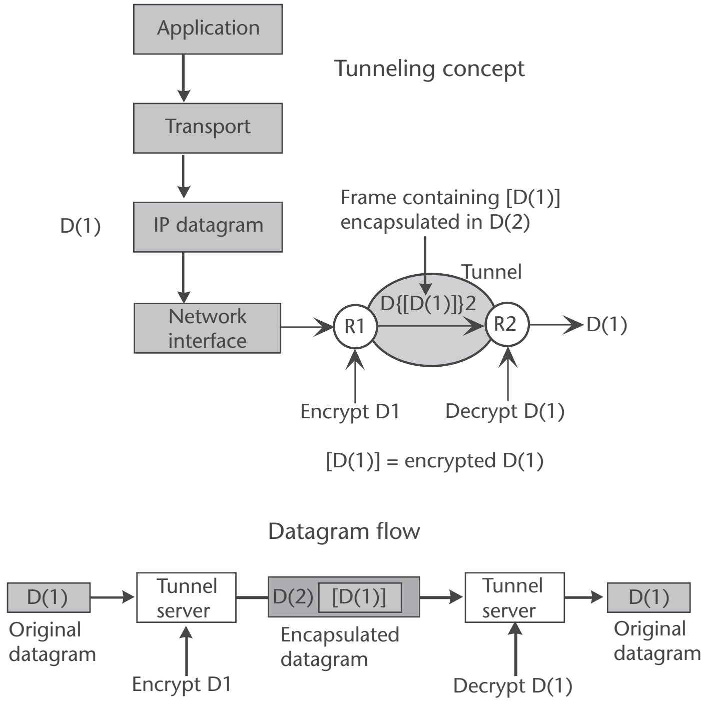


Figure 6.5 Tunneling.

If it is important that the message information be protected throughout its journey, the sender can encrypt it before forming the original frame. Decryption at the receiving station can serve to confirm (authenticate) that the message originated from the expected source (see the following).

6.2.4 Encryption, Decryption, and Authentication

Through the application of one or more rules, of encryption is the action of making readable (*clear-text*) data frames into not-readable (*cipher-text*) data frames. The rules for encryption are chosen so that the application of the same rules, or a set of rules based on them, will restore the not-readable frame to readability.

Decryption is the reverse of encryption. Through the application of one or more rules based on those employed to encrypt a packet, an encrypted frame is resotred to its original meaning.

These two rules are known as *keys*. Common encryption systems use a single key or two keys.

- *Single-key cryptography:* Also known as *secret-key cryptography*, employs the same key for encryption and decryption. Keys are bit patterns of any convenient length (40, 64, and 128 are common values). The longer the key, the harder the code is to break. To be effective, the key must be kept secret from everyone except the users.
- *Two-key cryptography:* Also known as *public-key cryptography*, employs two keys. One key is available to the public (public key); the other key is known only to its owner (private key). Either key can be used to create encrypted messages. They are decrypted by the other key.

Because of the need to keep the single key secret even though both encrypter and decrypter are using it, the management of single-key systems is more difficult than two-key systems. For this reason, most encryption systems use two-key cryptography.

Two-key systems provide other advantages. Through the use of the keys in specific order, the sender can guarantee *privacy*, provide *authentication*, and encrypt the message to achieve both privacy and authentication. Suppose there are two stations. Station 1 knows its own private (S1) and public (P1) keys, and can obtain the public key of Station 2 (P2). In similar fashion, Station 2 knows its own private (S2) and public (P2) keys, and the public key of Station 1 (P1).

If Station 1 wishes to send a private message to Station 2, it encrypts the message (M) with Station 2's public key to produce P2⊗M, where ⊗ stands for the action of encrypting or decrypting. Upon receiving P2⊗M, Station 2 uses its private key to decrypt the frame. This produces S2⊗{P2⊗M} = M. Because Station 1 used Station 2's public key to encrypt the message, only Station 2 can decrypt it using its private key. Privacy is assured, but Station 2 cannot be sure of the origin of the message.

If Station 1 wishes to send a message to Station 2 and have Station 2 know with certainty that it came from Station 1, Station 1 encrypts it with its private key. This produces S1⊗M. Station 2 decrypts S1⊗M with Station 1's public key. This produces P1⊗{S1⊗M} = M. Because Station 1 used its private key to encrypt the message, the frame can only have come from Station 1. However, any station with Station 1's public code can decrypt it. Authentication is assured, but privacy is not.

If Station 1 wishes to send a private message to Station 2 and have Station 2 know with certainty that it came from Station 1, Station 1 encrypts the message with Station 1's private key and then with Station 2's public key. This produces P2⊗S1⊗M. Station 2 decrypts P2⊗S1⊗M with its private key and then with Station 1's public key. This produces S2⊗P1⊗{P2⊗S1⊗M} = M. Privacy is obtained by encryption with P2 and decryption with S2. Authentication is obtained by encryption with S1 and decryption with P1.

Cryptography is an important ingredient in national security. For this reason, the U.S. Government is ever vigilant to ensure that commercial cryptography does not compromise national cryptography. In addition, law-enforcement agencies are anxious to limit the effectiveness of commercial cryptography so that codes used by criminals can be broken.

6.2.5 IP Security

A set of protocols known as *IPsec* (IP security) has been developed by the IETF to provide authentication and privacy services for IPv4 and IPv6. Authentication provides the receiver with the ability to check that the immutable fields in the received frame are identical to those in the frame that was sent. (Immutable fields are those that do not change during transport.) Thus, the message, the transport header, and parts of the network header are immutable. Items such as time-to-live and network checksum vary with the number of nodes the frame passes through. They are mutable and are carried as 0s when calculating the hash information.

Operating at the Internet layer, the services allow the stations to select a level of security that matches their security requirements. The parameters for each security service are collected and stored by the receiver. They are called a *security association* (SA). As a minimum, an SA includes: an identification number (security parameters index); a cryptographic algorithm; a key or keys that implement the algorithm; the lifetime of the key(s); and a list of sending stations that can use the security association. Each destination creates its own SAs. In addition, it stores a number of mandatory algorithms. To identify a specific SA requires both the security parameters index and the destination address.

In IPv4, authentication information is carried in an authentication header inserted between the Internet layer header and the transport layer header in the IP datagram. In IPv6, the IP datagram consists of a base header, extension headers, transport layer header, and message. The authentication header is one of the extension headers. Figure 6.6 shows IPv4 and IPv6 datagrams that include authentication headers. The information fields in the datagram are listed in Appendix B. The authentication header provides data integrity through the use of keyed hashing. Hash functions represent a variable-length message by a fixed-length data string. The hashing algorithm is negotiated during SA setup. It provides address and payload integrity by hashing those entries in the IP header that do not change and the entire payload. To provide additional security, IPsec can create new keys after a set amount of data has been transferred or a certain time has elapsed.

When authentication and privacy are required, IPsec employs an *encapsulating security payload* (ESP). ESP has three sections: an ESP header that is positioned

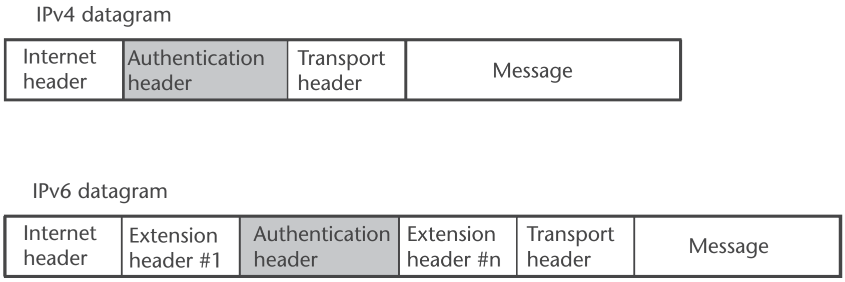


Figure 6.6 Authentication headers in IPv4 and IPv6 datagrams.

between the Internet header and the transport header, an ESP trailer that follows the message, and an ESP authentication that follows the ESP trailer. Appendix B lists the information fields in a datagram with ESP. Neither the authentication protocol, nor ESP, fits the definition of tunneling given earlier in this section. True, they provide authentication and/or encryption, but they do not wrap an encrypted datagram inside another datagram so that it can be forwarded between two points over an insecure temporary connection without making use of its contents.

IPsec defines tunneled versions of the authentication header and the encapsulating security payload. They are shown in Figure 6.7. Each contains the original IP datagram encapsulated by a second Internet header that contains the IP addresses of the tunnel ends. In addition, an authentication header or an ESP header is positioned next to the original datagram. An ESP trailer and ESP authentication field follow the original datagram in the ESP tunneling datagram.

6.2.6 Other Tunneling Protocols

Industry groups have developed other tunneling protocols. Of note are:

- *Point-to-Point Tunneling Protocol (PPTP):* A data link sublayer (Layer 2) protocol that encapsulates PPP frames in IP datagrams for transmission over an IP network. PPTP supports a single tunnel between client and server.
- *Layer 2 Tunneling Protocol (L2TP):* A data link sublayer (Layer 2) protocol that encapsulates PPP frames for transmission over IP, X.25, frame relay, or ATM. L2TP supports multiple tunnels. L2TP combines the best features of PPTP and L2F, an early product from Cisco Systems Corporation. When used in an IP network, L2TP uses UDP for tunnel creation and transmission. Both

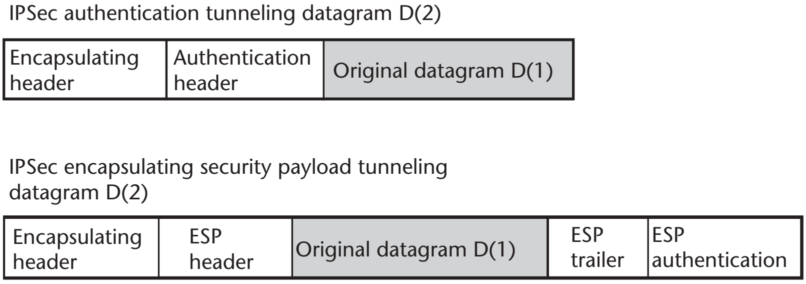


Figure 6.7 IPsec tunneling mode datagrams.

tunneled data and control frames share the same UDP stream. L2TP uses IPsec for cryptographic services. Figure 6.8 shows an L2TP datagram encapsulated by PPP and encrypted by IPsec. The original datagram is wrapped in a PPP frame. The PPP frame is then incorporated in a new IP datagram with a UDP header and an L2TP header. Adding an IPsec encapsulating security payload header and trailer and an IPsec authentication trailer provides message integrity and authentication. Finally, an IP header is attached that contains the network addresses of the beginning and ending of the tunnel.

6.2.7 Firewalls

In a catenet that has Internet connections, preventing eavesdropping, hacking, or theft of information and controlling the amount and nature of internal traffic forwarded to Internet are a formidable task. Most schemes rely on establishing and maintaining an electronic *firewall*, which is a software/hardware device that denies unauthorized callers access to a private network, and controls calls from the private network to destinations reached over the public network.

Situated between an intranet and the Internet, a firewall consists of screening routers, dedicated servers, and computer logic that implement rules to determine which connections are allowed and which are not. As noted in Section 6.2, the rules are called *proxies*. They restrict the number of services available to outside connections and prevent the manipulation of services to provide unauthorized levels of access. In addition, a firewall can be used to limit the flow of specific information to callers from within the intranet and serve as the termination of tunnels through the Internet.

Figure 6.9 generalizes the relationship between a firewall, a private network, and the Internet. Conceptually, the firewall prevents the free exchange of data frames between the private and public networks. If it compares favorably with one or more databases managed by servers and meets other tests (if applicable), a data frame will be passed around the wall. The internal router passes it on to the appropriate subnetwork. For a catenet with several campus networks connected by the Internet, a firewall is used to isolate each campus network from the Internet.

6.2.8 Functions Performed in Firewall

In Figure 6.9, a representative sampling is shown of the database and testing capabilities in the firewall servers and associated devices. For small networks, some can

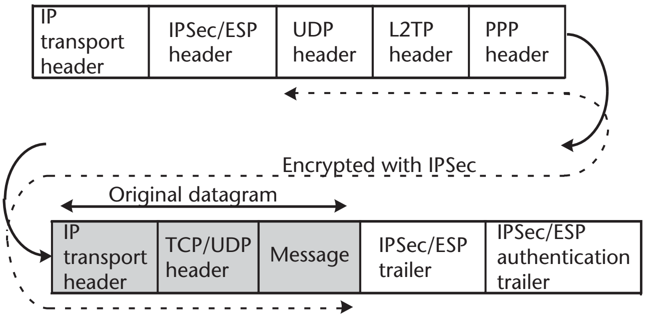


Figure 6.8 L2TP encapsulation with privacy and authentication.

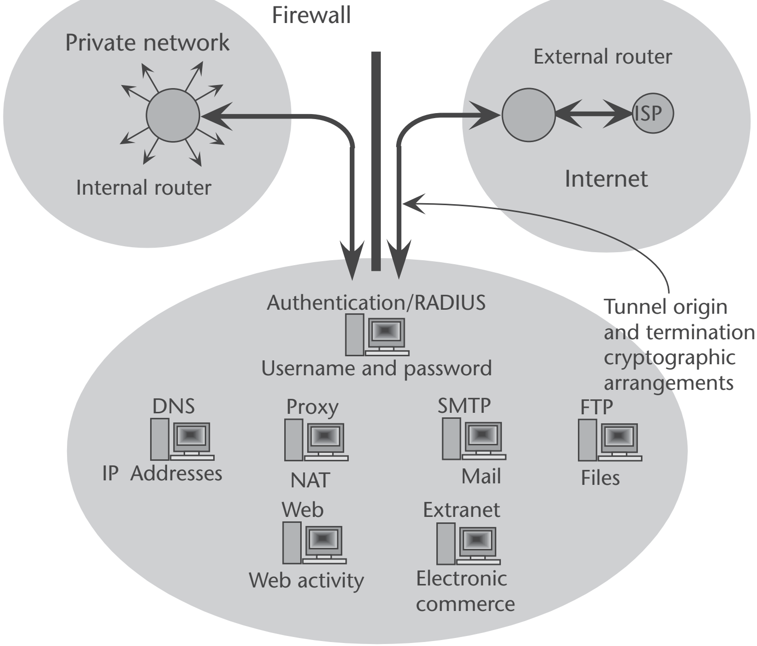


Figure 6.9 Concept of firewall and the functions it performs.

be combined, and not all of them may be necessary. In large networks, they may all be individual units, and more may be necessary to handle special situations.

When a private network is connected to the Internet, it is usual for management to be concerned about the time wasted by employees surfing the Web for personal reasons. This concern leads to a request for a policy that only authorized users may access the Internet. To implement this policy requires the manual entry of each authorized user in a database. For a large user community, this can be a lot of work, particularly if there is significant turnover. If dynamic IP addressing is in use (i.e., each station receives an address at the start of a session and is entitled to its use for a fixed time), the procedure will be complicated by changes in station addresses. If the station operator is changed frequently, the procedure may be complicated by changes in usernames and passwords. If banning all http:// traffic is impossible, perhaps the best approach is to maintain activity logs and question excessive use or the use of specific addresses.

Briefly, the functions that may be implemented at the firewall can be described as follows:

- *Authentication:* Knowing that the incoming message has not been changed on its journey through the public network and that the sender is correctly identified is important for incoming traffic. Knowing the correct identity of those that make outgoing calls to use Internet services or contact persons is equally important. Proxy and/or *Remote Authentication Dial-In Service* (RADIUS) servers make appropriate tests on the data frames. They work with username and password information and may challenge originating or terminating entities to confirm information.

- *Simple mail transport service (SMTP), domain name service (DNS), File Transfer Protocol (FTP), and World Wide Web (WWW):* Standard Internet services may require individual handling. Some users will have more privileges than others, and some may have none. All traffic should be recorded in segregated logs for review and troubleshooting.
- *Network address translation:* By using special addresses that are not recognized by Internet devices, a private network may be hidden from Internet stations. For traffic to be accepted from the Internet, the incoming addresses must be translated from Internet IP addresses to private network IP addresses.
- *Cryptography:* The firewall can serve as the origin and termination of tunnels across the Internet to other campus networks, employees on the road, and authorized customers and suppliers. The firewall must know what certificate authorities (CAs) to use, which cryptographic algorithms are authorized, and what kind of key management is expected. A certificate authority is a trusted third-party organization or company that issues digital keys (certificates) used to create digital signatures and public/private cryptographic keys. For IPsec, the encryption scheme is defined by the firewall. Other encryption schemes are determined by the destination IP address.
- *Electronic commerce:* Tunnel calls between enterprise employees and customers or suppliers are set up in accordance with agreed proxies. Both customers and suppliers are likely to be permitted only a limited group of internal contacts.

Altogether, the capability of the devices in the firewall is sufficient to create a secure network out of the combination of campus networks and Internet. They permit enterprises to have confidence in their data communication facilities, while taking advantage of the flexibility and pervasiveness of the Internet. Perhaps it is too much to hope that there will be a neat set of standardized devices in the future.

6.3 Virtual Private Networks

A *virtual private network* (VPN) is a data network composed of private and public sections that permits sending confidential data over unprotected public connections without the risk of compromise by eavesdroppers, thieves, or those who would sabotage information. To the users, a VPN appears as a private network.

The success of the Internet has inspired companies and organizations to distribute an increasing amount of information over circuits using Internet protocols. In a format made easy to read by incorporating the graphical interfaces and hypertext techniques of the Web, companies and organizations are able to provide proprietary information to employees and product information to the public. To serve them, companies and organizations use the public Internet. To serve their internal needs, companies and organizations use private internets called *intranets*.

At first, users from inside and outside the enterprise were pleased to communicate with one another and do business together. However, once the user community had suffered a few episodes of eavesdropping, hacking, or thefts of information, they sought to achieve privacy without sacrificing the flexibility acquired from using the public Internet. To do this, they created a VPN. However, it would be wrong to imagine that VPNs can be created solely from public Internet facilities. They use the full-range of communication facilities including leased telephone circuits, frame relay or ATM links, communication satellite hops, ISDN, and POTS.

6.3.1 Types of VPNs

VPNs can be divided in several ways. One set of configurations is:

- *Intranet VPN:* A VPN in which several enterprise campus networks are interconnected by tunnels over Internet connections (distributed backbone).
- *Extranet VPN:* An intranet VPN used by customers, suppliers, and vendors. Tunnels are established over Internet connections to a secure enterprise server.
- *Remote access VPN:* A VPN in which enterprise employees on the move can establish a dial-up connection to a remote ISP and create tunnels to enterprise campus networks.
- *Intracompany VPN:* A single campus network or an intranet VPN, in which encrypted communications are used to protect against security breaches within the enterprise.

Using any of these arrangements ensures the owner has a significant level of control over who can read information (i.e., read only), work with information (i.e., download), and contribute or change information (i.e., author or edit). Furthermore, they can restrict electronic mail and other traffic to within the company. In addition, the network uses a popular set of protocols that are familiar to many persons. Moreover, campus networks (intranets) can be connected over a distributed backbone supplied by the Internet.

6.3.2 Basic Connections

As pointed out at the beginning of the chapter, there are as many kinds of data networks as there are enterprises using them. It is unlikely that any fall neatly in the categories listed earlier. Privacy in the commercial world is difficult to implement and almost impossible to guarantee. It is even harder when some of the communication facilities are used by the public, and company loyalty is not what it used to be. Nevertheless, the lure of a pervasive network that is significantly cheaper than leasing private lines, is hard to refuse. For clients operating within company facilities, the keys to success are user authentication (e.g., passwords), address management (e.g., network address translation), and proxies (e.g., content filtering). For clients operating in the public domain, overriding importance must be given to encryption and tunneling. In addition, they are the keys to private connections between campus networks over the Internet.

Figure 6.10 illustrates some basic connections between the facilities that I have described. At the top of the diagram a straightforward connection to Internet is made through the campus firewall that will include many of the individual protections shown in Figure 6.9. Unauthorized communications by persons on campus and off campus can be prevented while providing access for legitimate purposes. The middle diagram shows a campus-to-campus connection. Because the informa-

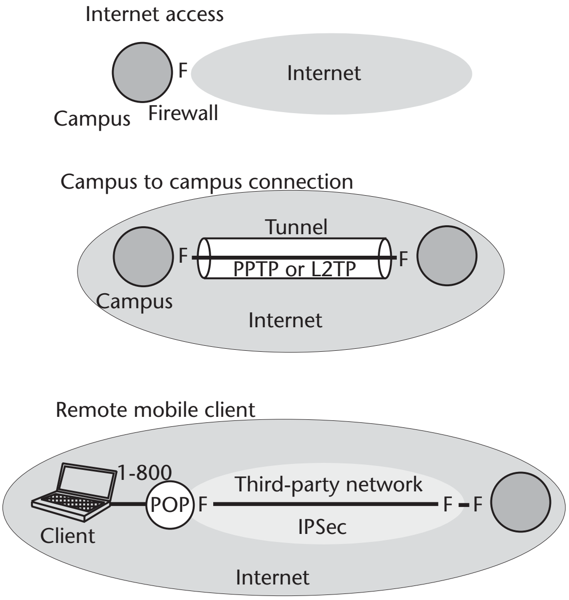


Figure 6.10 VPN basic connections.

tion exchanged is important, an encrypted tunnel is employed. At the bottom is an arrangement that a remote client can employ. The client makes use of a third party's facilities by calling an 800 number. The POP connects the call through a server and a secure connection to the campus firewall. A level of security is provided by IPsec.

Enterprises have recognized that the Internet is an affordable, worldwide medium that can be used to interconnect private networks and carry sensitive data. Their demand has created an opportunity for ISPs to offer value-added services that emphasize scalability and network management. That they can provide worldwide transport is a nonissue. Of course, they can! But can they provide worldwide security? Irrespective of their promises, security must remain the responsibility of whoever wants to preserve confidentiality. Prudent managers understand this and will institute their security measures at their firewalls.

CHAPTER 7

Transmission Facilities

Electric currents, electromagnetic waves, and optical energy carry messages on transmission facilities. The availability of ubiquitous transport is a prerequisite for the operation of the networks described in earlier chapters. It is tempting for managers to fantasize about owning all the communication facilities needed to support an enterprise. However, it soon becomes apparent that transmission equipment is expensive, sites are difficult to obtain, and maintenance by enterprise employees is virtually impossible. Consequently, most transport outside corporate buildings uses facilities owned and operated by common carriers. In this chapter, I describe some of the systems likely to be provided by the telephone companies and other entities. Because these facilities work together, all companies providing transport services operate compatible equipment.

7.1 Twisted Pairs

Twisted pairs are major components of the public telephone network. They are the dominant bearers in the local loop. In addition, twisted pairs are used extensively for on-premises wiring for enterprise installations.

A *twisted pair* is two insulated wires twisted together and contained in a cable of many pairs. Known as *tip* and *ring*, neither of the wires is connected directly to the ground. The twist keeps the conductors balanced with respect to themselves, the cable shield, and other pairs. Often, twisted pairs are called *cable pairs*. A *paired cable* is a cable whose conductors are twisted pairs.

Commonly, twisted pairs are deployed in 25- or 50-pair *bundles* wrapped in a metal sheath known as a *binder*. The sheath is grounded at the cable ends. The binders are contained in an outer sheath of plastic to create *polyolefin-insulated cable* (PIC). In common use, the number of pairs in a cable ranges from 25 or 50 to as many as 4,200. Figure 7.1 shows some of these items and identifies the signals associated with a twisted pair. They are:

- *Differential mode signals:* Signals applied between the wires of a twisted pair. Also known as *metallic* signals. Messages are always transmitted as differential signals.
- *Common mode signals:* Signals measured between the two wires and ground. Also known as *longitudinal signals*. Common mode signals are created by outside interference (noise).

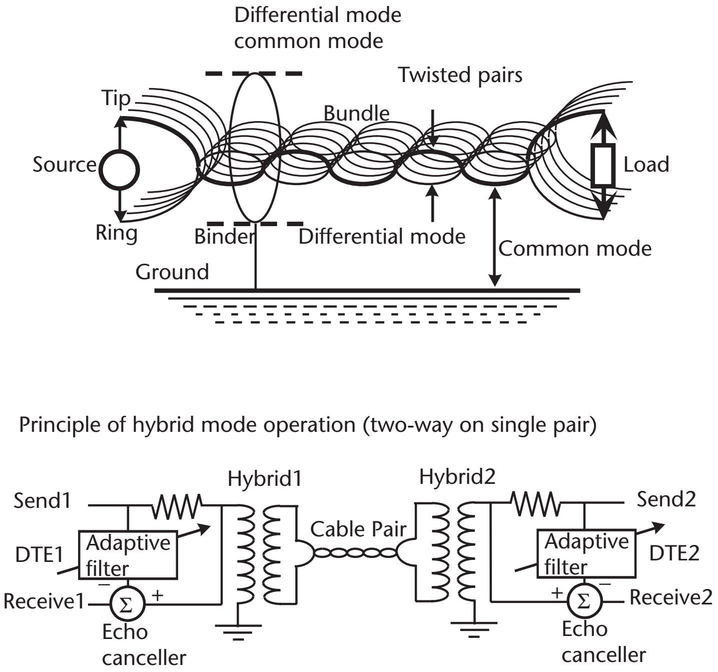


Figure 7.1 Differential, common, and hybrid modes in twisted pair operation.

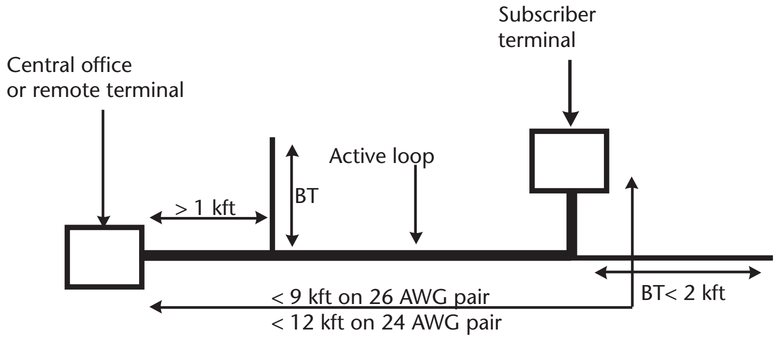
Two-way operation over a single twisted pair is achieved by the use of transformers, echo canceling devices, and adaptive filters. Called *hybrid mode operation*, the principle is shown in the lower half of Figure 7.1. When a signal is sent from terminal *Send1*, the combination of the adaptive filter and echo-canceling device prevents it from appearing at terminal *Receive1*. Simultaneously, if a signal is sent from terminal *Send2*, terminal *Receive1* receives it without interference from *Send1*. Hybrid operation eliminates the need to run a second pair to each subscriber to obtain a duplex circuit.

7.1.1 Cable Pair Impairments

Cable pairs are subject to impairments produced by installation procedures. For instance, in areas where cables have been installed in anticipation of demand, less than the full length of the cable pair may be used to serve an existing subscriber. The remainder is left attached but not terminated. It is called a *bridged tap*, which is a cable pair continued beyond the point at which the pair is connected to a subscriber or an unterminated cable pair attached to an active cable pair.

Because they load the active pair, bridged taps increase the attenuation of the signal and create impedance discontinuities. The higher attenuation lowers the signal-to-noise ratio at the receiver and the impedance discontinuities cause signal reflections that can adversely affect the data stream. Figure 7.2 shows some bridged tap arrangements. They are anathema for most data circuits, although *digital subscriber line* (DSL) equipment operates with limited tap lengths.

Another installation practice that is detrimental to digital signals is the use of *loading coils*. As the length of the cable pair increases, the attenuation increases. Because of the capacitance of the pair, the higher voice frequencies suffer more



Limitations based on carrier serving area (CSA) specifications

No more than 2 BTs
First more than 1 kft from CO
Longest BT < 2kft
Total BT length < 2.5 kft

Figure 7.2 Bridged taps.

attenuation than the lower voice frequencies. Eventually, the voice signal becomes unintelligible due to the loss of these frequencies. On long connections (over 18,000 feet), it was standard practice to add loading coils to improve voice signal performance. Loading may be present on 19-, 22-, and 24-gauge loops longer than 18,000 feet, or 26 gauge loops longer than 15,000 feet. D66 loading consists of 66-mH coils spaced 4,500 feet apart. H88 loading consists of 88-mH coils spaced 6,000 feet apart. The first load coil from the CO is located a half-section out. However, the additional inductance has an adverse effect on digital signals, and the coils must be removed before the connection can be used for data. Modern practice relies on equalizers to compensate for unequal frequency attenuation.

One further installation practice should be noted. To ensure reliable ringing (and reliable disconnects) of telephones powered from the cable pair, a current of greater than 25 milliamps is required. With a 48-volt battery in the CO, a 26-AWG (American Wire Gauge) copper wire loop can connect points up to a maximum 9,000 feet apart (carrier serving area). To serve loops longer than this, larger size wires are added. As the distance from the CO increases, the wire size is increased from 26 to 24 to 22 and (rarely) 19 AWG. If space permits in the CO cable vault, 24 AWG pairs alone can be used to 12,000 feet. At the junction points, the changes in wire diameter produce impedance changes that create reflections and may have an adverse effect on digital signals. In selecting a cable pair connection for data, the one with the least number of wire size changes is likely to provide the best performance.

4.1.2 Circuit Noise

Signals are subject to corruption by many events. Collectively, the interference is known as *noise*, which is the sum of all unwanted signals added to the message signal in the generation, transmission, and reception processes.

Figure 7.3 illustrates the transmission environment in which the major noise contributor is longitudinal current. These currents are produced in tip and ring by voltages to ground. If the loop is balanced to the ground, they are of equal magni-

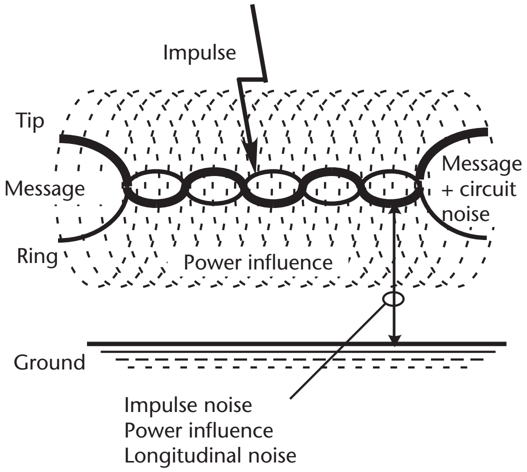


Figure 7.3 Noise components.

tude and flow in the same direction so that the voltage between tip and ring is zero. However, if the loop is unbalanced to ground, signals due to the longitudinal currents will be measured between tip and ring. On an idle circuit, this is known as *circuit noise*, which is also known as *metallic*, *background*, or *differential* noise. Using a band-limited weighting filter, it is the power measured between tip and ring when no message signal is present.

A common filter weights the noise frequencies in proportion to their perceived annoyance. The output of the filter is expressed in dBrnC, *decibels referenced to noise with C-weighting*. Circuit noise has two major components:

- *Power influence:* Noise caused by inductive interference from the public power system. Radiation from the public power system comprises fundamental (60 Hz) and harmonic (*n*×60 Hz) frequencies. Telephone equipment is susceptible to harmonics, especially those above 300 Hz. (Interference from three-phase power systems is somewhat less than from single-phase systems because even harmonics cancel out leaving only the odd harmonics to generate interference.)
- *Impulse noise:* Short, intense bursts of noise. For telephone purposes, it is defined as a voltage increase of greater than 12 dB above the root-mean-squared (rms) background noise that lasts less than 10 ms. Impulses are produced by lightning strikes, certain types of combustion engines, and sudden changes in load due to catastrophic events. A pair with circuit noise less than 20 dBrnC is rated *good*. On long rural routes, less than 26 dBrnC is acceptable. Above 40 dBrnC, the circuit is *unacceptable*.

7.1.3 Crosstalk

Other interfering signals are generated by *crosstalk* between circuits. *Crosstalk* occurs when signals between an unbalanced tip and ring (differential mode signals) generate electromagnetic fields that induces interfering signals in nearby pairs. Crosstalk is a factor in limiting the *rate* at which data can be sent, and the distance over

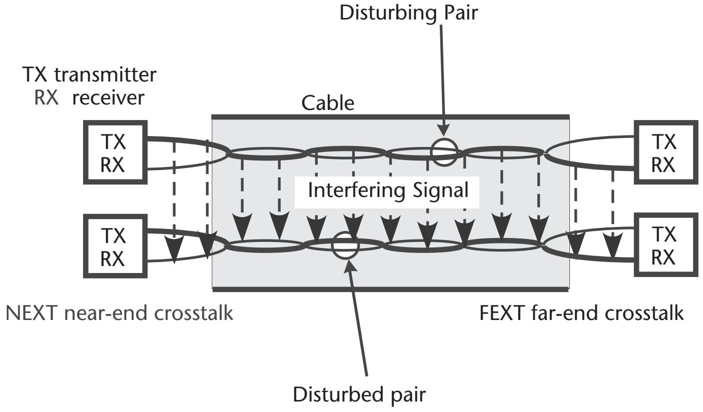


Figure 7.4 Crosstalk components.

which it may be sent (*data reach*). Figure 7.4 shows the major components of crosstalk in a paired cable. It is divided into *near-end crosstalk* and *far-end crosstalk*:

- *Near-end crosstalk (NEXT):* A condition in which a signal transmitted over a twisted pair in a paired cable creates a disturbance in other pairs at the *same* end of the cable. Near-end crosstalk is produced by interference from the transmitting wire of one pair to the receiving wire of another pair measured at the receiving point at the same end of the cable. The magnitude is independent of the length of the cable. NEXT can be a major impairment in systems that share the same frequency band for downstream and upstream transmissions. (The *downstream* direction is from the CO to the subscriber. The *upstream* direction is from the subscriber to the CO.) When different frequency bands are used, NEXT between downstream and upstream signals is avoided. NEXT can be divided into:
  - *SNEXT:* Crosstalk from the same type of signal running in the same binder (*self-crosstalk*);
  - *FNEXT:* Crosstalk from a different type of signal running in the same binder (*foreign crosstalk*).

  Near-end crosstalk is the sum of self-crosstalk and foreign crosstalk. As shown in Figure 7.4, crosstalk also affects equipment at the far end of the cable.
- *Far-end crosstalk (FEXT):* A condition in which a signal transmitted over a twisted pair in a paired cable creates a disturbance in other twisted pairs at the *far* end of the cable. Far-end crosstalk is produced by interference from the transmitting wire of one pair to the receiving wire of another pair measured at the receiving point at the far end of the cable. Its magnitude depends on the length of the cable. Like NEXT, FEXT is composed of SFEXT and FFEXT and can be avoided if different frequency bands are used for downstream and upstream signal streams.

Because larger numbers of wire pairs are bundled together in feeder cables of finer wire, crosstalk is more severe at the CO end of a connection. At the subscriber end, where there are fewer and coarser wires, the level of crosstalk is less severe. This means that the upstream signal-to-noise ratio at the central office will be less than the downstream signal-to-noise ratio at the pedestal. Accordingly, higher rate signals can be transmitted downstream to the customer than can be transmitted upstream to the central office.

7.2 Transport Based on Twisted Pairs

Twisted pairs are used to transport digital signals operating from 2.4 kbit/s to 55 Mbps and higher. Common twisted pair digital loops are:

- *Subrate digital:* 2.4–56 kbit/s; symmetrical channels (i.e., upstream and downstream channels operate at same speed); employs one pair.
- *T-1 carrier:* 1.544 Mbps; symmetrical channels; employs two pairs, one for each direction; with repeaters every 6,000 feet, operates up to 50 miles; uses AMI line code (see Appendix A).
- ISDN subscriber lines:
  - *Basic rate (BRI):* 160 kbit/s; symmetrical channels; employs one pair; operates to 18,000 feet; uses 2B1Q line code (see Appendix A).
  - *Primary rate (PRI):* 1.544 Mbps; symmetrical channels; operates over any existing DS-1 rate transmission systems (e.g., repeatered T-1 or HDSL).
- Digital subscriber lines:
  - *High bit-rate DSL (HDSL):* 1.544 Mbps; symmetrical channels; employs two pairs (dual-duplex); without repeater operates to 12,000 feet, with one repeater (doubler) operates to 24,000 feet; with two repeaters operates to 36,000 feet; uses 2B1Q line code.
  - *Single-pair high-data-rate DSL (G.shdsl):* Up to 2.32 Mbps; symmetrical channels; employs one pair; operates up to 24,000 feet without repeater.
  - *Asymmetric DSL (ADSL):* Up to 8 Mbps downstream and up to 640 kbit/s upstream, employs one pair; operates to 12,000 feet without repeater.
  - *Very high-speed DSL (VDSL):* 13 Mbps and 26 Mbps symmetrical, or 52 Mbps downstream and 6.4 Mbps upstream; employs one pair; operates over short distances between fiber access nodes and clusters of buildings.

The bit rates quoted are actual line rates. The user's data rate is something less than these rates. Some units require two twisted pairs; others use only one. The differences between the performance of DSLs reflects the year in which each was standardized and the capability of digital electronics at the time.

7.2.1 Transmission System 1 (T-1)

The first digital transmission equipment widely deployed in the Bell System was T-1 (*transmission system 1*). In its original application, it carries 24 multiplexed voice channels at a speed of 1.544 Mbps. Multiplexing is the action of interleaving several signal streams so that they can be carried on a single bearer. A multiplexer combines several digital signals into a higher speed digital stream. Each voice signal is sampled 8,000 times per second, and the sample values are companded and coded in 8-bit words. Companding (derived from the words compressing and expanding) is the action of reducing the dynamic range of a signal so an approximately equal number of samples are present at each quantizing level for digitizing. The samples are compressed so that higher-value amplitudes are reduced with respect to lower-level amplitudes. This makes more quantizing codes available to lower level signals and improves the signal-to-noise ratio. To convert compressed samples back to something close to their original levels, the amplitudes of the samples are expanded. The digital values are transmitted over two cable pairs (one for each direction) and *alternate mark inversion* (AMI) signaling is employed (see Appendix A). At least 90% of the signal energy is distributed between 0 Hz and 1.5 MHz with a peak at around 700 kHz. The signals are amplified, reshaped, and retimed by repeaters spaced 6,000 feet apart (except the first and the last which must be within 3,000 feet of the terminals). Normally, because of jitter in the timing circuits, a T-1 line is limited to no more than 50 repeaters.

T-1 established certain parameters that have permeated the modern *public switched telephone network* (PSTN). For instance, in the digitizing process, the analog voice signal is sampled at 8,000 samples per second. This limits the bandwidth of a reconstructed analog voice signal to 4 kHz (see Appendix A). With an 8-bit quantizing code, the basic digital voice rate becomes 64 kbit/s. Quantizing is the process that segregates sample values into ranges and assigns an 8-bit code to each range. Whenever a sample value falls within a range, the output is the code assigned to that range. Known as DS-0 (digital signal level 0), 64 kbit/s is the basic building block for all higher-speed services, whether voice or data. When used for data, the functions of sampling, companding, quantizing, and coding described earlier are not employed.

7.2.1.1 Data T-1

Figure 7.5 shows a T-1 configured for data-only operation. It differs from T-1 voice in that the twenty-fourth byte of each frame is used as a signaling channel. In T-1 voice, all 24 bytes are used for voice channels with per channel signaling provided by bit robbing in every sixth byte of each channel. In data operation T-1 consists of multiplexers connected to terminal repeaters that are then connected to one another over two twisted pairs punctuated by line repeaters. To emphasize the flexibility of T-1, I have included a second multiplexer that multiplexes subrate (i.e., 2.4, 4.8, 9.6, and 19.2 kbit/s) duplex data lines to 64 kbit/s. The multiplexer sends a bipolar signal to the terminal repeater and receives a similar signal from it. The terminal repeaters convert the bipolar stream to AMI format, time the outgoing signals, and regenerate the incoming signals.

Full-rate (64 kbit/s) data channels are interleaved to create a 1.544-Mbps data stream. Figure 7.6 shows the formation of a T-1 data frame. For simplicity, only one direction of transmission is shown. For duplex operation, a second frame must be created from bytes sent in the reverse direction. The frame consists of 23 bytes of payload, 1 byte of signaling data, and a *framing* bit (known as the 193rd bit). Each frame is transmitted at a speed of 1.544 Mbps in 125 μs (the voice sampling time). For the repeaters to function correctly, 12.5% (1 in 8) of the bits must be 1s, and

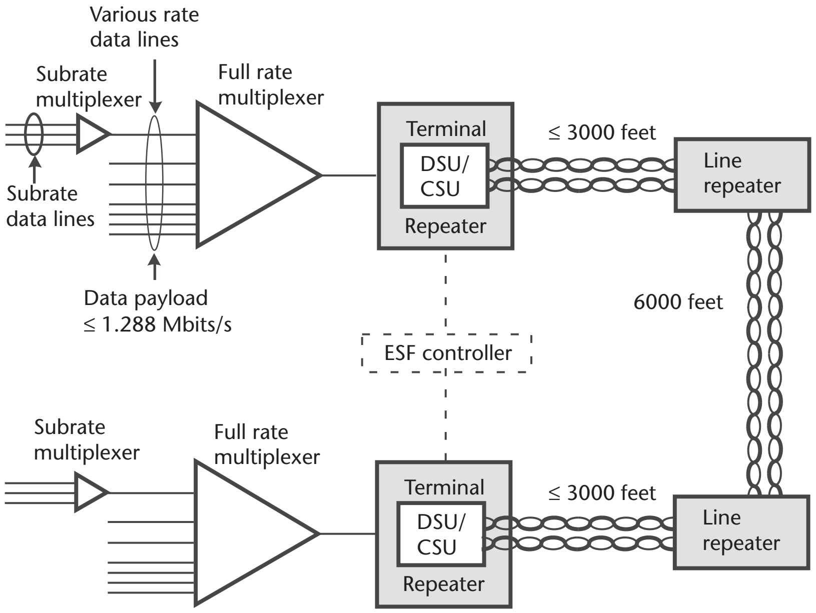


Figure 7.5 T-1 data-only configuration.

there can be no more than 15 consecutive 0s. To ensure meeting these figures the last bit of every data byte is set to 1. This action reduces the per channel data throughput to 56 kbit/s. With 23 data channels, the data throughput becomes 1.288 Mbps per T-1 line. To distinguish signaling bytes from data bytes, the eighth bit in a signaling byte is set to 0.

7.2.1.2 64-kbit/s Clear Channel

To make entire 64-kbit/s channels available to users (64-kbit/s clear channel capability), special coding that is transparent to the user is introduced into all-0s bytes. Called *bipolar with 8 zeros substitution* (B8ZS), bipolar violations are inserted in bit positions 4 and 7 of all-0s bytes. In an AMI signal, the 1s polarity alternates regularly. A bipolar violation is a 1 with the same polarity as the previous 1. Because of the violations (bits 4 and 7), the receiver can detect the pattern (bits 4, 5, 7, and 8) and remove it before processing. Each violation is followed by a normal 1 (in positions 5 and 8). Thus, 00000000 becomes 1V01V000 (Bit 8 ← Bit 1, canonical format), a pattern that more than meets the 1s requirement. The receiver reverses this substitution to produce the original data stream.

Another technique requires four frames (96 bytes) to be stored in a buffer. Called *zero-byte time slot interchange* (ZBTSI), all-0s bytes are removed, and the remaining nonzero bytes consolidated at the rear of the buffer. This leaves as many spaces at the front of the buffer, as the number of all-0s bytes. Into these spaces, seven bit numbers are entered that correspond to the positions of the all-0s bytes in the stream of 96 bytes. The eighth bit in the byte is used to indicate whether more all-0s bytes follow. At the receiver, the stream is reassembled with all-0s bytes in their correct position. This processing delays the stream by approximately 1.5 ms.

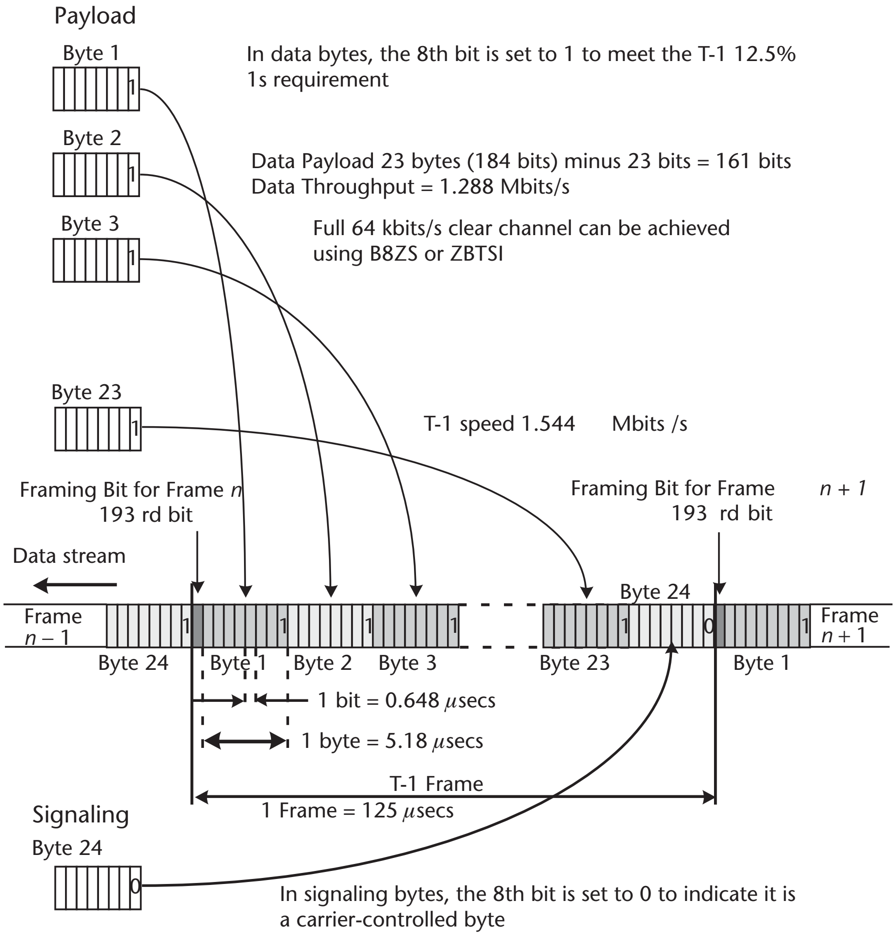

Payload
Byte 1
In data bytes, the 8th bit is set to 1 to meet the T-1 12.5% 1s requirement
Byte 2
Data Payload 23 bytes (184 bits) minus 23 bits = 161 bits
Data Throughput = 1.288 Mbits/s
Byte 3
Full 64 kbits/s clear channel can be achieved using B8ZS or ZBTSI
Byte 23
T-1 speed 1.544 Mbits /s
Framing Bit for Frame n
193 rd bit
Framing Bit for Frame n + 1
193 rd bit
Data stream
Frame n − 1
Byte 24
Byte 1
Byte 2
Byte 3
Byte 23
Byte 24
Byte 1
Frame n + 1
1
0
1 bit = 0.648 μsecs
1 byte = 5.18 μsecs
T-1 Frame
1 Frame = 125 μsecs
Signaling
Byte 24
In signaling bytes, the 8th bit is set to 0 to indicate it is a carrier-controlled byte

Figure 7.6 T-1 data frame format.

7.2.1.3 Framing Bits and Extended Superframe

The framing bit acts as a marker to synchronize the electronics and ensure the boundaries of each byte are detected correctly. Framing bits in consecutive frames are used to provide control patterns and error information. Two arrangements are a 12-frame *superframe* (SF) and a 24-frame *extended superframe* (ESF).

Figure 7.7 shows the 24-frame ESF. To make such a diagram, twenty-four 193-bit frames are stacked on top of one another. By doing this, individual channels appear as columns and the 193rd bits appear as a column at the left-hand side of the frame. They perform three functions. The six F bits in frames 4, 8, 12, 16, 20, and 24 form the pattern 101010. It is used to synchronize electronics and ensure that the receiver remains locked to the frame structure. The 12 D bits provide a 4,000-bps data link facility that forwards specific application information or historical data for maintenance use. The six C bits in frames 2, 6, 10, 14, 18, and 22 are the frame check sequence of a cyclic redundancy check that monitors the error performance of the 4,632-bit superframe. The bit stream is divided by a 7-bit polynomial (1000011) to give a 6-bit FCS. Error checking is used to measure the performance of T-1 facilities (see Section 4.3).

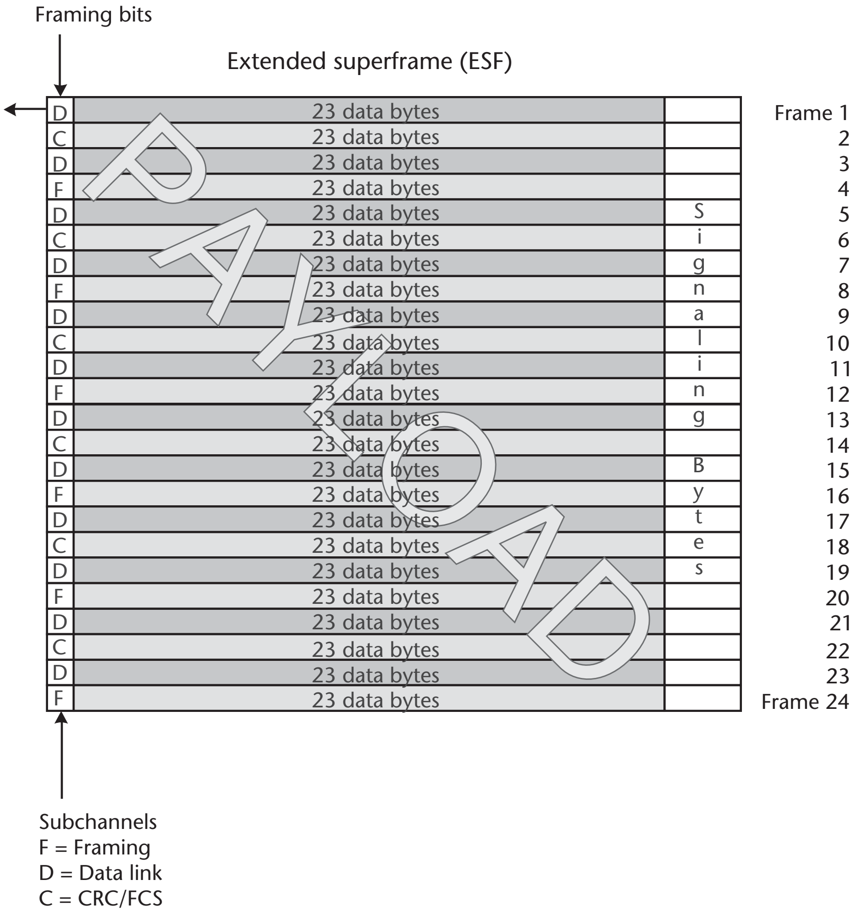


Figure 7.7 T-1 Extended superframe format.

7.2.1.4 T-Carrier Family

T-1 was the first in a hierarchy of multiplexed transmission systems developed to carry digital voice circuits in ever increasing numbers. The entire family consists of six units:

- *T-1:* Multiplexes 24 DS-0 (64 kbit/s) signals into one DS-1 (1.544 Mbps) signal (DS-1 = 24 DS-0s).
- *T-1C:* Multiplexes two DS-1 signals into one DS-1C (3.152 Mbps) signal (DS-1C = 48 DS-0s).
- *T-2:* Multiplexes four DS-1 signals into one DS-2 (6.312 Mbps) signal (DS-2 = 96 DS-0s).
- *T-3:* Multiplexes seven DS-2 signals into one DS-3 (44.736 Mbps) signal (DS-3 = 672 DS-0s). Known as T3 SYNTRAN (synchronous transmission), a special version developed for enterprise networks multiplexes 28 DS-1 signals directly to DS-3.
- *T-4NA:* Multiplexes three DS-3 signals into one DS-4NA (139.264 Mbps) signal (DS-4NA = 2076 DS-0s).

- *T-4:* Multiplexes six DS-3 signals into one DS-4 (274.176 Mbps) signal (DS-4 = 4032 DS-0s).

Only T-1 and T-1C operate on twisted pairs. Byte-level multiplexing is used in T-1 and T-3 SYNTRAN. In turn, a byte from each input line is assembled in a frame with framing and control bits, and placed on the output line. Bit-level multiplexing is used in T-1C, T-2, T-3, T-4NA, and T-4. In turn, a bit from each input line is assembled in a subframe with framing and control bits, combined with other subframes, and placed on the output line. Only T-1 and T-3 SYNTRAN have found major employment in a data environment. In many applications, digital subscriber lines are replacing T-1, and T-3 is being replaced by SONET.

7.2.2 ISDN

In the 1970s, with the development of digital computers, growing demands for data communication, and the perfection of digital voice, it became apparent to many PSTN operators that an all-digital network could carry both voice and data traffic. Called *integrated services digital network* (ISDN), it gave impetus to the development and deployment of digital switches. Later, with the invention of digital television, the concept was expanded to include video. The idea of a broadband, multimedia, digital network was born. Called *broadband ISDN* (B-ISDN), it gave impetus to the development of ATM switches, *synchronous optical network* (SONET), and *synchronous digital hierarchy* (SDH) transmission systems (see Sections 7.4.1 and 7.4.2).

Many problems had to be solved, including how to provide digital channels to individual subscribers. Presently, ISDN supports two service speeds—160 kbit/s (128- or 144-kbit/s payload) and 1.544 Mbps (1.472-Mbps payload). They provide a combination of bearer (B) channels and signaling (D, for *delta* or data) channels.

Basic Rate ISDN provides 2 × 64 kbit/s B channels, 1×16 kbit/s D channel, and 16 kbit/s overhead, for a total of 160 kbit/s. Designed to serve customers with nonloaded loops, its reach is 18,000 feet. To reduce signal attenuation over the longer loops, AMI coding was replaced by 2B1Q coding (see Appendix A). Achieving 2 bits per baud efficiency, at least 90% of the signal energy is distributed between 0 Hz and 772 kHz. Two-way operation over a single cable pair is achieved through the use of echo cancelers. Neither loading coils nor bridged taps can be present.

Primary-rate ISDN provides 23 × 64 kbit/s B channels and 1 × 64 kbit/s D channel to a customer. With a separate signaling channel, the customer has access to the full 64 kbit/s (clear-64) in the 23 B channels. B channels can be aggregated into H0 channels (384 kbit/s) and H11 channels (1.536 Mbps). For H11 channels, signaling is provided by a D channel from another primary rate interface. As in T-1, a frame consists of 24 bytes to which a framing bit (193rd bit) is added. In addition, a multiframe structure is created that consists of twenty-four 193-bit frames. Framing bits in frames 4, 8, 12, 16, 20, and 24 are used to maintain frame synchronization. However, the code is different from T-1—it is 001011. Primary rate ISDN is provided over two cable pairs using any DS-1 transmission system such as repeatered T-1 or HDSL (see Section 8.1.2).

7.3 Optical Fibers

Optical carriers used for communication are located in the infrared portion of the spectrum between 250 and 450 THz (Terahertz, 1 THz = 3×10^{14} Hz). They have wavelengths from approximately 0.85 μ to 1.6 μ (1 μ = 1 micron = 1 meter $\times 10^{-6}$). It is usual to specify them in terms of wavelength rather than frequency. Optical fibers are superior to twisted pairs in several ways:

- Because optical energy is not affected by electromagnetic radiation, it is immune from noise generated by common electromagnetic sources.
- Because the optical energy is focused in the center of the fiber and the coating (buffer) is impervious to infrared wavelengths, crosstalk is of no concern in optical fiber cables. All of the optical energy is guided along the fiber.
- Because the frequencies of optical carriers are very high compared to conceivable message bandwidths, they can be used to transport very wideband message signals.
- Because optical fiber cables can be much smaller than paired cables, in areas in which underground ducts are used, the substitution of fiber cables for paired cables frees significant space for future expansion.

Compared to copper wires, optical fibers have disadvantages:

- Optical energy propagates in only one direction along the fiber. Two fibers are needed to make a duplex circuit.
- Optical fibers are insulators; they do not conduct electricity. Therefore, they cannot carry electrical power for operating repeaters and other electrical devices. Powering equipment *through the line* is only possible if copper wires are added to the cable.
- Microbends and other mechanical insults increase fiber loss. In comparison, they have no effect on copper wires.

7.3.1 Single-Mode Fiber

The predominant design in telecommunications applications is *single-mode* fiber. It is a strand of exceptionally pure glass with a diameter about that of a human hair (125 micron = 0.005 inch). The refractive index varies from the center to the outside to focus optical energy in the center of the strand and guide it along the length. Shown in Figure 7.8, in such a fiber, the central glass core is less than 10 microns in diameter and of higher refractive index than the glass cladding. With a refractive index of 1.475, the velocity of energy in the core is approximately 200,000 km/sec (i.e., approximately two-thirds the velocity of light in free-space). A significant (and essential) fraction of the optical energy travels in the cladding glass. Because its velocity is slightly higher (around 211 km/sec) than the energy in the core, conditions are right to support single-mode propagation.

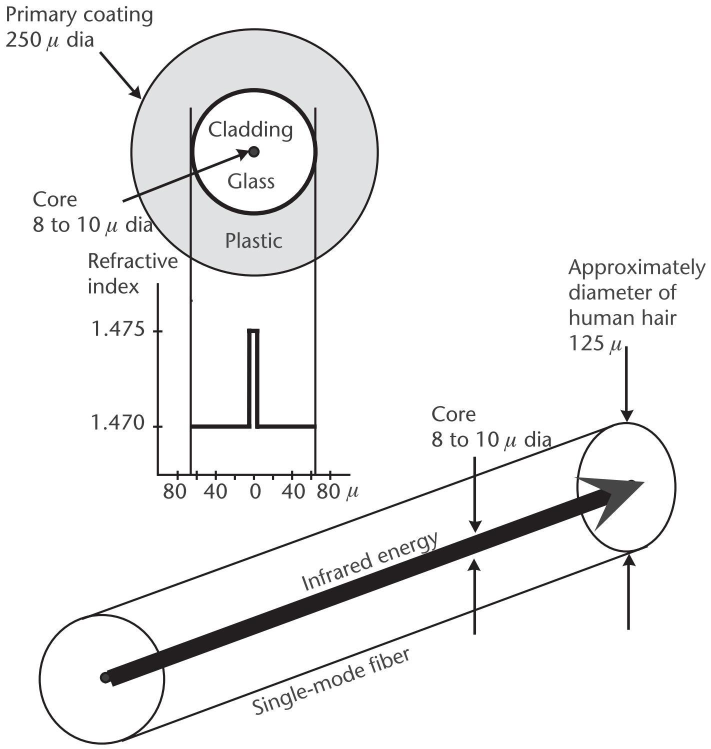


Figure 7.8 Single-mode optical fiber.

7.3.2 Optical Properties

Single-mode fibers are used with solid-state laser transmitters and photodiode detectors that operate at wavelengths around 1,550 nanometers (1 nanometer = 1 meter $\times 10^{-9}$; 1,550 nm = 1.55 micron). The lasers are switched on and off to produce pulses of infrared energy. At 1,550 nm, the fiber has an attenuation of around 0.2 dB/km (i.e., a loss of approximately 5% per kilometer, or 8% per mile). Spans of up to 60 miles can be achieved without using a repeater, and repeaterless spans of up to 130 miles have been achieved in undersea cables.

7.3.3 Wavelength Division Multiplexing

Several optical carriers can be transmitted simultaneously in the same single-mode fiber. Called *wavelength division multiplexing* (WDM), current practice employs up to 64 carriers, with the expectation that this can be upgraded to 256 carriers in the near future, and perhaps as many as 400 carriers eventually. The term *dense wavelength division multiplexing* (DWDM) is used to describe systems that employ these higher numbers of wavelengths. Crosstalk is a major concern in WDM. Interference is produced by imperfections in network components and by fiber nonlinearities that scatter the optical energy of the carriers.

7.3.4 Optical Amplifiers

Very long-distance WDM transmission is made possible by optical amplifiers. As shown in Figure 7.9, in one design a length of erbium-doped fiber is placed in the

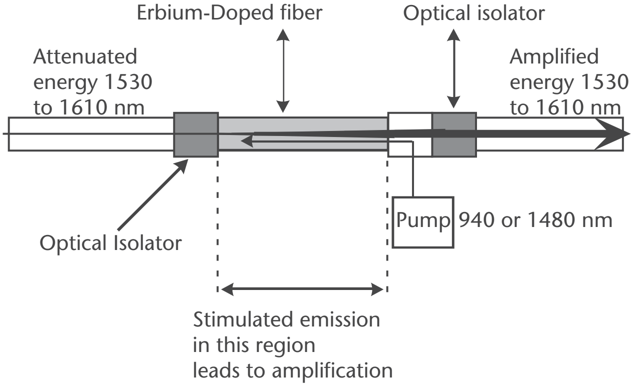

Erbium-Doped fiber
Optical isolator
Attenuated energy 1530 to 1610 nm
Amplified energy 1530 to 1610 nm
Optical Isolator
Pump
940 or 1480 nm
Stimulated emission in this region leads to amplification

Figure 7.9 Principle of Erbium-doped fiber amplifier.

optical path. Arrangements are made to pump this fiber with energy at 980 or 1,480 nm. Optical isolators are used to terminate the fiber. They restrict the pumping energy to the erbium fiber. In this fiber, the Er^{3+} ions are raised to a metastable state from which they spontaneously decay to the ground state. Because the isolators do not stop the WDM carriers, the photons of the message streams collide with (stimulate) the metastable ions. As the stimulated ion returns to the ground state, it emits a photon with the same wavelength, phase and direction as the photon it collided with (stimulated emission). Because a single photon can stimulate many ions, the result is amplified streams of coherent photons at the signal wavelengths. Ions that are not stimulated by a photon spontaneously decay to the ground state. In doing so, they emit incoherent radiation that contributes to amplifier noise. Called EDFAs, *Erbium-doped fiber amplifiers* produce gains of up to 40 dB between 1,530 and 1,610 nm (C-band, 1,530–1,565 nm; and L-band, 1,570–1,610 nm).

7.3.5 Short-Distance Facilities

For short distances, in a building or on a campus, the fiber can be made of plastic with a core of elevated refractive index or glass with a core over which the refractive index varies in a graded manner. Called *step index* and *graded index* fibers, they are shown in Figure 7.10. The energy propagates in multimode fashion along the core. Because many ray paths are possible, each with a slightly different length, the signal is dispersion-limited, and the distance-bandwidth product is significantly less than that of single-mode fiber. Nevertheless, for short distances, multimode fiber installations are reliable and relatively cheap.

7.4 Transport Based on Optical Fibers

Unlike wire, on which the signal propagates in both directions, fiber is a one-way bearer, and two are needed to complete a circuit. Pairs of optical fibers are used in point-to-point applications, and other topologies in which the need for access at intermediate points can be limited. To provide transport between major traffic junctions, telephone companies use a flexible, multipurpose, ring-like architecture. They employ two or four fiber rings to ensure fiber paths are available to recover from

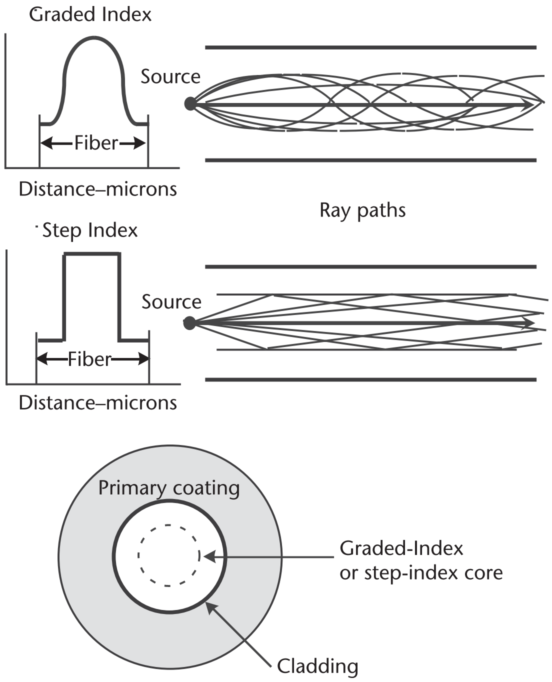


Figure 7.10 Short-distance fibers.

service interruptions. While transmission is by optical means, all signal processing is accomplished electronically.

7.4.1 Synchronous Optical Network

Synchronous optical network (SONET) is an all-digital, optical fiber transport structure that operates from 51.84 Mbps to 40 Gbps (Gbps = gigabits per second = 1,000 Mbps = 10^9 bps), and beyond. SONETs serve as very high-speed backbones in the Internet, as high-speed distribution networks in local exchange and interoffice facilities, and provide optical transport channels for private connections. Figure 7.11 shows the principle of SONET. The basic configuration is a double fiber ring in which the fibers operate in opposite directions. Should a fault occur in a link, traffic is routed back on itself to complete the journey to its destination. A SONET may contain equipment that performs the following functions:

- *Add/drop multiplexer (ADM):* Aggregates or splits SONET traffic at various speeds so as to provide access to SONET without demultiplexing the SONET signal stream. Generally, it has two equal speed network connections.
- *Terminal multiplexer (TM):* An end-point or terminating device that connects originating or terminating electrical traffic to SONET. Has only one network connection.
- *Digital cross connect (DCS):* Redistributes (and adds or drops) individual SONET channels among several STS-N links. Consolidates and segregates STS-1s, and can be used to separate high-speed traffic from low-speed traffic (to feed one to an ATM switch and the other to a TDM switch, for instance).

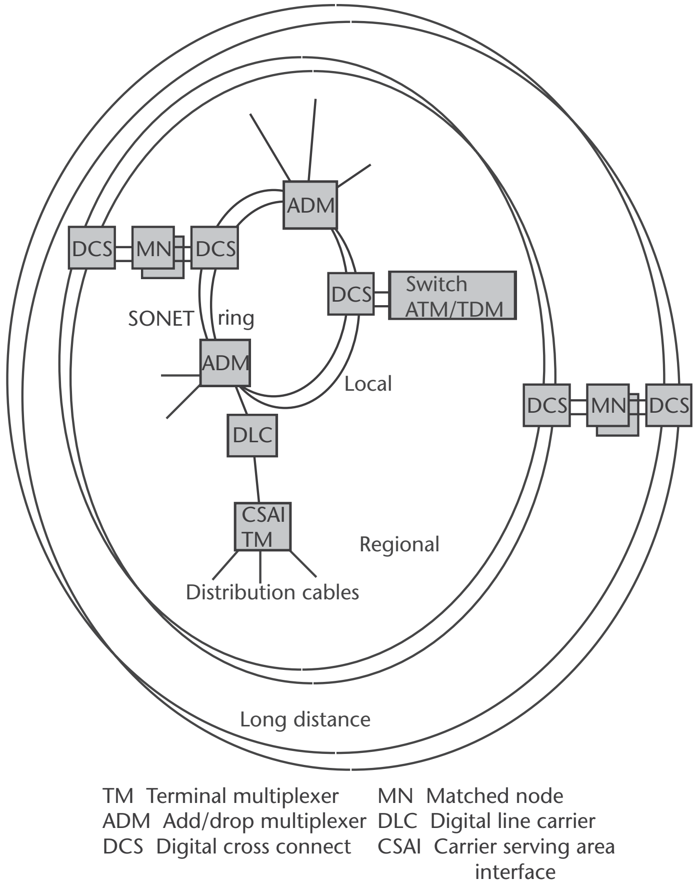


Figure 7.11 SONET rings.

- *Digital line carrier (DLC):* Used to link serving offices with *carrier serving area* (CSA) interface points. Typically, SONET DLCs concentrate DS-0 signals into OC-3 signals.
- *Matched node (MN):* Pairs of MNs are used to interconnect SONET rings and provide alternate paths for recovery in case of link failure. SONET traffic is duplicated and sent over two paths between the rings. One set of MNs provides the active path; the other set is on standby in case of failure of the active connection.
- *Drop-and-repeat node (D+R):* SONET devices configured to split SONET traffic and copy (repeat) individual channels on two or more output links. Applications include the distribution of residential video and alternate routing. (This is not shown in Figure 7.11.)

7.4.1.1 SONET Signals

While SONET is an optical transmission system, the signals at the fiber ends are converted to electrical form for processing. SONET standards define a set of optical/electronic interfaces for network transport. The electrical signal hierarchy has *N* members.

- *Synchronous transport signal level 1 (STS–1):* With a basic speed of 51.84 Mbps, STS-1 signals are designed to carry T–3 signals or a combination of T-1, T-1C, and T-2 signals that is equivalent to DS–3.
- *Synchronous transport signal level N (STS-N):* With speeds that are multiples of STS–1 (i.e., $n \times 51.84$ Mbps), STS-N signals are created by byte multiplexing N STS-1 signals. For various reasons, the values $N = 3$ (155.52 Mbps), 12 (622.08 Mbps), 24 (1244.16 Mbps), 48 (2488.32 Mbps), 96 (4,976.64 Mbps), 192 (9,953.28 Mbps), and 768 (39,813.12 Mbps) are preferred.

Corresponding to the STS signal hierarchy, the optical signals transmitted over the fiber facility are:

- *Optical carrier level 1 (OC-1):* The optical equivalent of STS-1;
- *Optical carrier level N (OC-N):* The optical equivalent of STS-N.

Similar to their electronic counterparts, optical carriers are designated OC-1, OC-3, ..., OC-768.

7.4.1.2 SONET Frames

To achieve compatibility with PSTN operations, SONET multiplexers create STS-1 frames of 125-μs duration. Figure 7.12 shows an STS-1 frame. It consists of 810 bytes, of which 774 are payload. To the payload are added 9 bytes of path overhead to form the *synchronous payload envelope* (SPE). The path overhead contains data that monitors and manages the electrical and optical connections between originating and terminating multiplexers. To the SPE are added 27 bytes of transport overhead to form a frame. The transport overhead contains data that monitors and manages the optical line between the originating and terminating SONET multiplexers.

Payloads that originate from the T-carrier family consist of a fixed number of bytes every 125 μs. Called *virtual tributaries*, they occupy 9 rows × n columns in the SPE. Thus, the virtual tributary for DS-1 consists of 27 bytes (9 rows × 3 columns). Twenty-four of them are DS-0 bytes from the T1 frame, 2 bytes are overhead related to the virtual tributary, and 1 byte is framing information. A DS-3 frame consists of 672 bytes (28 × 24). When joined with signaling bytes and stuffing bits that compensate for speed variations and fill the frame, it occupies a complete STS-1 frame.

STS-N frames are constructed by byte multiplexing lower speed frames. Of 125-μs duration, an STS-N frame is equal to $N \times$ STS-1 frames. When a signal fills more than one STS-N frame, the several frames are defined as a *concatenated* structure and designated STS-Nc. They move through the network as a single entity.

7.4.2 Synchronous Digital Hierarchy

For BISDN applications, ITU standardized a hierarchy of transport systems called *synchronous digital hierarchy* (SDH). The levels and frames [known as *synchronous transport modules* (STMs)] are exactly three times those of SONET. Thus, *synchronous transport module level 1* (STM-1) is a frame of 2,430 bytes at 155.52 Mbps (STM-1 = 3 STS-1 = STS-3); and *synchronous transport module level N* (STM-N) is

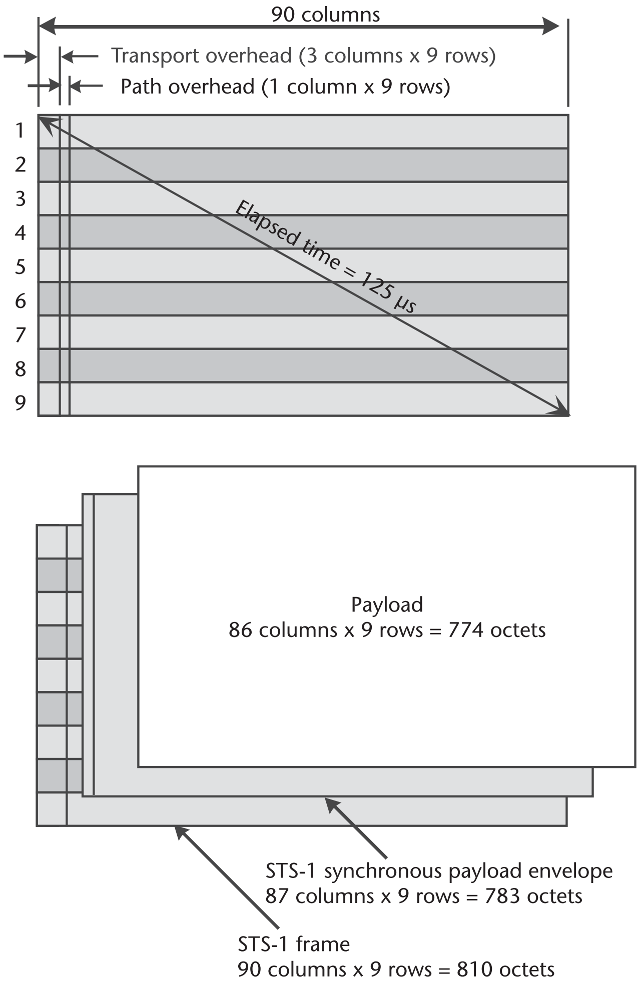

90 columns
Transport overhead (3 columns x 9 rows)
Path overhead (1 column x 9 rows)
1
2
3
4
5
6
7
8
9
Elapsed time = 125 μs
Payload
86 columns x 9 rows = 774 octets
STS-1 synchronous payload envelope
87 columns x 9 rows = 783 octets
STS-1 frame
90 columns x 9 rows = 810 octets

Figure 7.12 SONET frame.

a frame of $N \times 2430$ bytes at $N \times 155.52$ Mbps. STM-N frames are created by byte multiplexing N STM-1 frames. STM-N = N STM-1 = 3N STS-1.

In a formal sequence, STM frames are assembled from 125-μ segments of tributary signals. Figure 7.13 shows the combinations of tributaries that can form an STM-1 frame. By adding path overhead, containers (C-11, C-12, C-2, C-3, or C-4) with a 125-μ segment of a tributary signal are converted to virtual containers (VC-11, VC-12, VC-2, or VC-3). By adding pointers to indicate the start of the virtual container, VCs are converted to tributary units (TU-11, TU-12, TU-2, or TU-3). TUs are grouped together to form a tributary unit group (TUG-2 or TUG-3), and are combined with path information for the TUG to form another virtual container (VC-3 or VC-4). By adding pointers to indicate the start of these virtual containers, the VCs are converted to administrative units (AU-3 or AU-4). Finally, AU-4 or 3 AU-3s are used to create an STM-1 frame. With microwave systems and optical fibers, the STM format is employed around the world. A notable application is the undersea fiber cables that encircle the globe. Within the United States, in optical fibers, the STS format is preferred.

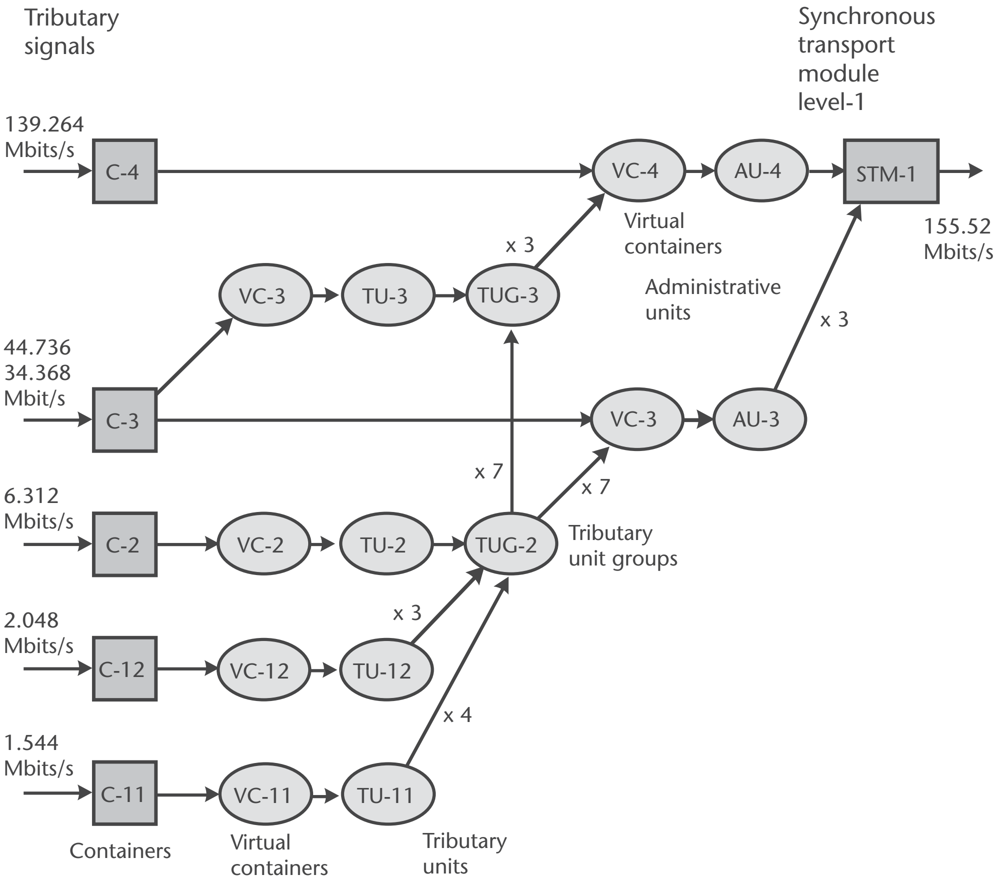


Figure 7.13 Tributary multiplexing scheme to create STM-1 frame.

7.5 Radio

Called *wireless* by Heinrich Hertz and its early developers, radio is a means of communication that employs electromagnetic waves in free space. It is this wireless property that is so important to us today. It has permitted millions of mobile users to free themselves from fixed voice networks and communicate from almost anywhere in an approximately seamless environment. Even at high speed, driving from one cell into another is accomplished without the user being aware of the change. Mobile telephones have been adopted the world over as an important adjunct to enterprise operations and as a means of keeping in touch. The next step is to provide wireless data communications as an extension of fixed data networks. However, it is not possible to provide the same transparency for data terminals. Dropping the radio connection to one access point and establishing a radio connection with another requires time during which the data stream is not transmitted. In addition, the vagaries of the electromagnetic medium make radio connections significantly less reliable than those provided by wires and fibers. Accordingly, a number of special features are included in the communication procedures that govern wireless data connections. To emphasize the difference, I use the term *movable* with data terminals in contrast to *mobile* telephone.

7.5.1 Frequencies and Modulation

Unlike wired point-to-point connections whose number could be increased until the world's copper supply is exhausted, the extent of the electromagnetic spectrum in which radio connections can be made is limited, and competition for slots is fierce. Consequently, international authorities and national governments control the use of the radio spectrum. In the United States, the FCC permits unlicensed wireless connections in three ISM (industrial, scientific, and medical) bands. They are:

- UHF ISM: 902 to 928 MHz;
- S-band ISM: 2.4 to 2.5 GHz;
- C-band ISM: 5.725 to 5.875 GHz.

In addition to wireless network connections, microwave ovens, medical imaging equipment, and other radiating devices use these bands. To accommodate these disturbing devices, the communication signal must be robust and immune to high-levels of interference. To accommodate as many users as possible in the limited bandwidths available, frequency reuse and noninterfering, low-power signals are employed. The connections use spread spectrum or orthogonal frequency division modulation techniques (see Appendix A).

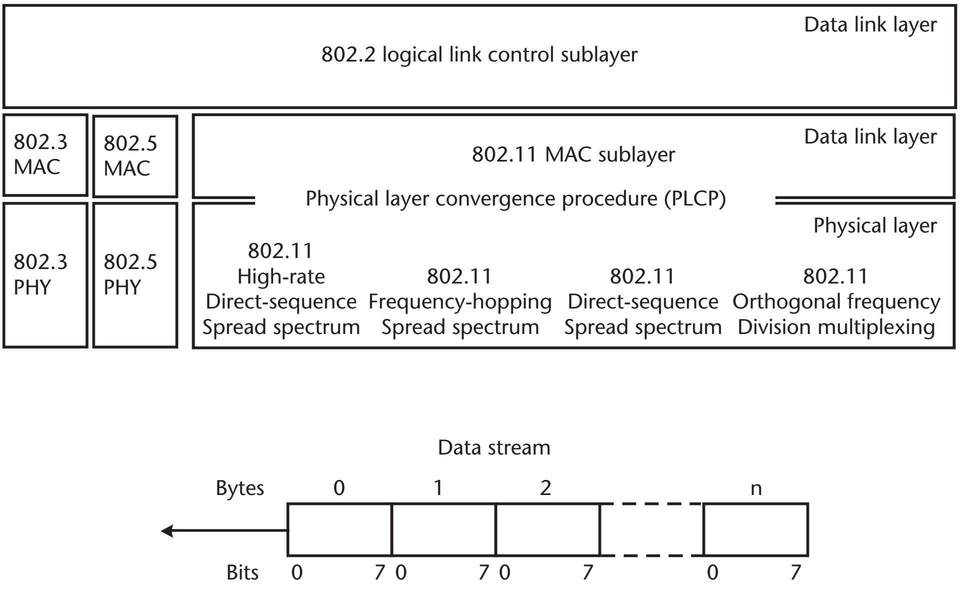
7.5.2 IEEE 802.11 Standard

Sponsored by the organization that standardized Ethernet and Token Ring LANs, IEEE 802.11 makes use of some of their features. (IEEE 802.11 has been called wireless Ethernet.) Figure 7.14 shows the relationship of IEEE 802.11 to the rest of the 802 family of specifications. It employs IEEE 802.2, the logical link sublayer of the data link layer; uses a unique MAC sublayer that includes collision avoidance; and has four physical sublayers that accommodate different implementations of the radio link. In addition, a procedure is added at the MAC/PHY interface. Called the *physical layer convergence procedure* (PLCP), it adds fields to the frame for use on the radio link. The IEEE 802.11 standard defines the infrastructure and frame formats for complete wireless networks (such as wireless LANs). In last-mile applications they are used to provide data communications between movable data terminals and fixed sites. Popular application locations are airports and other places where people gather and must wait for service.

IEEE 802.11 includes changes in the bit-ordering conventions. Bits are numbered 0 to 7 in each byte with the least significant bit on the left (bit 0), and the most significant bit on the right (bit 7). Bytes are numbered 0 to *n* and read from left to right, as usual. The change makes for easier manipulation of the bit stream. It is shown at the bottom of Figure 7.14.

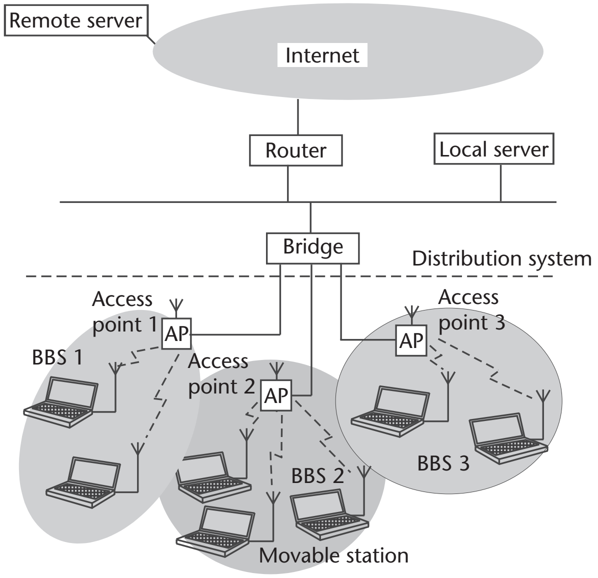
7.5.2.1 Infrastructure

Figure 7.15 shows movable stations, fixed *access points* (APs), and supporting equipment. The distribution system above the dashed line in Figure 7.15 can be configured in many ways. What the diagram suggests is one arrangement. The APs are tied to a bridge that links them together and, through a router, links them to the Internet. Servers can be positioned locally or remotely. A number of movable sta-


Data link layer
802.2 logical link control sublayer
802.3 MAC
802.5 MAC
Data link layer
802.11 MAC sublayer
Physical layer convergence procedure (PLCP)
Physical layer
802.3 PHY
802.5 PHY
802.11 High-rate Direct-sequence Spread spectrum
802.11 Frequency-hopping Spread spectrum
802.11 Direct-sequence Spread spectrum
802.11 Orthogonal frequency Division multiplexing
Data stream
Bytes
0
1
2
n
Bits
0
7
0
7
0
7
0
7

IEEE 802.11 bit and byte order

Figure 7.14 IEEE 802.11 in relation to other members of IEEE 802 family.



BBS basic service set

Figure 7.15 IEEE 802.11 basic service set and fixed facilities.

tions are associated with each AP. They form a *basic service set* (BSS). With the bridge connecting the three APs, users in different BSSs can communicate among themselves as well as access network services. When a movable station moves out of range of its associated AP, it must join another BSS by *associating* with the AP whose BSS it joins. A certain amount of downtime is required while arrangements are made to host the station and inform the routing tables of the change.

7.5.2.2 Frame Format

The format of an IEEE 802.11 frame is shown in Figure 7.16. A description of each field is given in Appendix B. The frame includes fields from an IEEE 802.3 frame that contains an IP packet. They are rearranged and augmented to take account of the radio link. The header includes four addresses. Addresses 1 and 2 are the destination and source addresses as they appear in the 802.3 header. Address 3 is required to identify the AP/BSS hosting the movable terminal. Address 4 is reserved for future use.

Because the radio link is established and synchronized in the physical connection, the preamble and start fields of the 802.3 header are discarded. In their places are a frame control field and a duration/ID field. The purpose of the frame control field is to provide the 802.11 version number and identify the type of frame that follows. They are divided into management, control, and data frames. The other bits in this 2-byte field perform specific alerting functions. The duration/ID field gives the time in microseconds the originator expects to occupy the radio channel to complete this transmission. If fragmentation is involved, it is the time to complete the entire transmission. The time is known as the *network availability vector* (NAV). It is noted by all stations in the BSS. They may not transmit during this interval.

Between Addresses 3 and 4, the sequence control field provides information that allows reconstruction of fragmented frames and detection of retransmitted and duplicate frames. The frame check sequence field checks the entire 802.11 frame.

7.5.2.3 Collision Avoidance

By reducing collisions and retransmissions, the total time required to transmit frames over the noisy environment of the ISM bands can be minimized. To do this, IEEE 802.11 specifies a MAC technique that extends the CSMA/CD routine of Ethernet to *carrier sense multiple access with collision avoidance* (CSMA/CA). A simplified diagram of a data exchange between two stations with collision avoidance is shown in Figure 7.17. Frames employ stop-and-wait ARQ. Before transmitting data, the sender sends a *request-to-send* (RTS) control frame to the receiver and

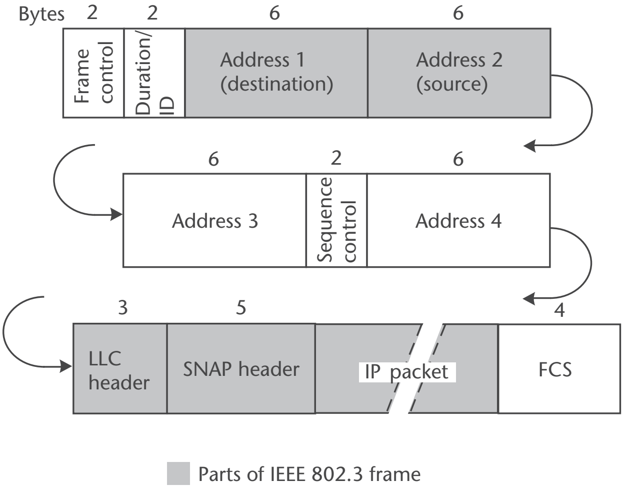


Figure 7.16 IEEE 802.11 frame incorporating IEEE 802.3 frame.

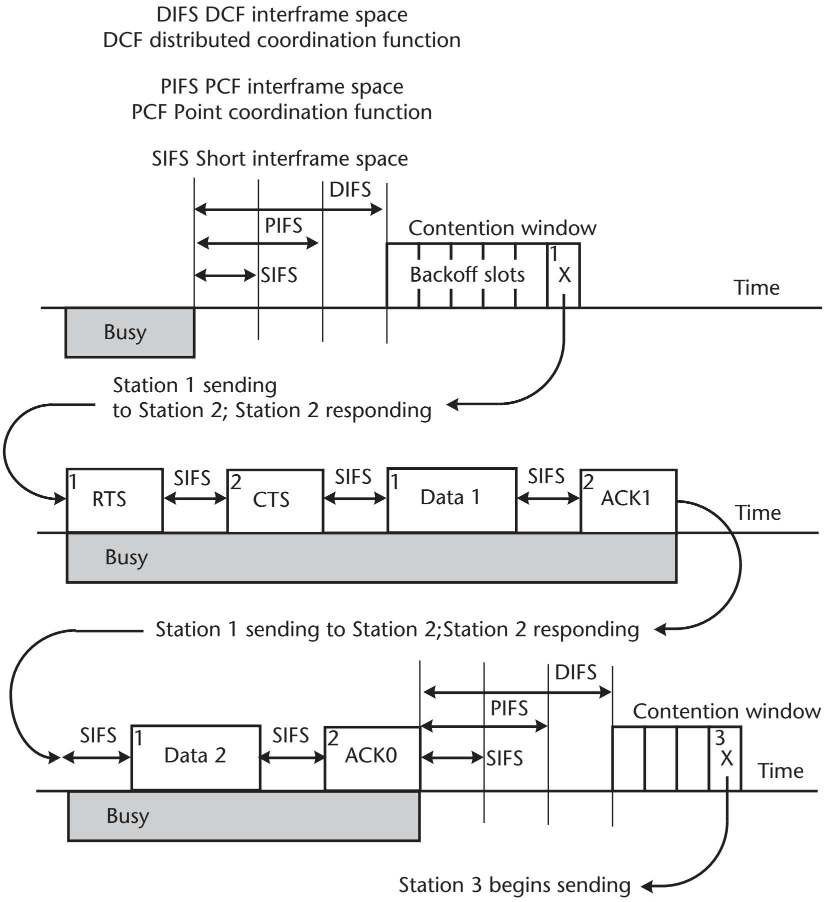

DIFS DCF interframe space
DCF distributed coordination function
PIFS PCF interframe space
PCF Point coordination function
SIFS Short interframe space
DIFS
PIFS
SIFS
Contention window
Backoff slots
1
X
Time
Busy
Station 1 sending
to Station 2; Station 2 responding
1
RTS
SIFS
2
CTS
SIFS
1
Data 1
SIFS
2
ACK1
Time
Busy
Station 1 sending to Station 2;Station 2 responding
SIFS
1
Data 2
SIFS
2
ACK0
DIFS
PIFS
SIFS
Contention window
3
X
Time
Busy
Station 3 begins sending

Figure 7.17 Illustrating collision avoidance.

waits for the receiver to reply with *clear-to-send* (CTS). As soon as the other movable stations hear the beginning of this exchange, they may not transmit. When the sender receives the CTS signal, it waits a short time then commences sending data. At the beginning of this action, all other stations in the BSS received a NAV time. They know they cannot transmit until it expires. When it does, stations with something to send wait a specific interframe time then back off a random number of slots. If no carrier is sensed, the station with the earliest backoff slot begins with an RTS/CTS routine and sets the NAV value to the estimated time of its transaction.

IEEE 802.11 specifies three interframe times, also shown in Figure 7.17:

- *DCF interframe space (DIFS):* The minimum idle time for contention-based services. If the channel has been idle for DIFS or longer, stations may have access to it subject only to random backoff (DCF: distributed coordination function).
- *PCF interframe space (PIFS):* An interval used during contention-free operation. Station with permission to transmit contention-free may begin after PIFS has elapsed and preempt contention-based traffic (PCF: point coordination function).

- *Short interframe space (SIFS):* An interval used for high-priority transmissions such as RTS/CTS frames and ACKs. SIFS is less than DIFS. Once a multiframe transmission has begun, subsequent frames are sent after SIFS interval. This preempts other frames that must wait for DIFS.

By using SIFS and extending the NAV as required, stations occupy the channel as long as necessary.

7.5.2.4 Security

Wireless signals are relatively easy to intercept. In the days when mobile radio used analog FM, many people though it fair game to listen in to other peoples' conversations. With the move to digital signals and spread spectrum modulation, eavesdropping is more difficult, but still can be done by determined listeners using more complex equipment. The question arises: How secure should IEEE 802.11 operations be? Like all questions of this kind, the answer is: It depends! It depends on the value of the information being passed over the link, and whether it must be protected for an hour, a day, a year, or forever. The quicker the information ages, the less concern there will be over keeping it secure, and it can never be completely protected. Given enough time and a fast enough computer, even state secrets are made known to the competition.

In truth, to be effective, several layers of security are needed. Starting with the weakest, which guards against casual compromise, and ending with the strongest, which guards against determined, well-prepared adversaries, they should be invoked according to the priority afforded security. IEEE 802.11 includes a symmetric key security procedure called *Wired Equivalent Privacy* (WEP). Its effectiveness depends on the length of the secret key used in ciphering and deciphering, and the size of the community with which each secret key is shared. Too large, and the probability of compromise is certain. Too small, and the problems of generating numbers of keys and distributing them in a timely (and regular) fashion becomes an administrative nightmare. Characterized by some as weak, WEP provides security against casual-compromise and not very determined adversaries. The 802.11 Committee is investigating opportunities to strengthen it. The strongest performance will always be given by encryption at the source using a one-time-only random key before entering the communication system.

CHAPTER 8

The Convergence of Voice and Data

In this final chapter, I tackle several topics that mix voice technology, wideband transmission, and data. I describe the local loop, which most of us use to bridge the last mile between our homes and to access the Internet; digital subscriber lines and cable television facilities that many of us use to obtain faster access to Internet; and the use of IP techniques to send voice over Internet. *Voice over IP* (VoIP) appeals to many as the application that will integrate data and voice services.

8.1 The Last Mile

The *last mile* is a descriptive term of art used by communicators. It is a somewhat inaccurate name for the connection between subscribers and a telephone central office or a remote terminal. In the United States, the average length of the connection between a subscriber and the central office is around 12,000 feet (i.e., 2.3 miles). A remote terminal serving area may extend up to 9,000 feet (i.e., 1.7 miles) from the terminal. A twisted pair of insulated copper wires makes the connection. Over it, users obtain voice and data services, and, for many, it is their connection to the Internet.

8.1.1 The Local Loop

In the public telephone network, all wiring and facilities between the customers' premises and the central office are known as *outside plant*. They make up the *local loop*.

8.1.1.1 Twisted Pairs in the Local Loop

Traditionally, the local loop has been composed of three levels of paired cables.

- *Feeder cables:* Bundles of twisted wire pairs contained in cables that connect the *main distributing frame* (MDF) in the CO to *feeder distribution interfaces* (FDIs).
- *Distribution cables:* Smaller cables made up of bundles of twisted pairs that extend the dedicated connections from FDIs to pedestals or cabinets close to individual service users.
- *Drop wires:* A final connection that is made by a multipair wire to the user's premises.

Figure 8.1 shows the arrangement of cables in a traditional local loop. The feeder/distribution cable topology can be described as *star-star*. The feeder cables and the CO form one star, and the distribution cables and each of the FDIs form a second ring of stars. In this environment, digital subscriber lines home on DSLAMs located in the CO, and optical fibers are laid to *optical network interfaces* (ONIs) located close to residential neighborhoods.

Taking advantage of improvements in technology, central offices are being consolidated into large wire centers with switches that support smaller, remote switches. Usually, connections between these satellite switches and the main switch employ optical fibers. Any DSLs served by remote switches home on DSLAMs located at the remote switch.

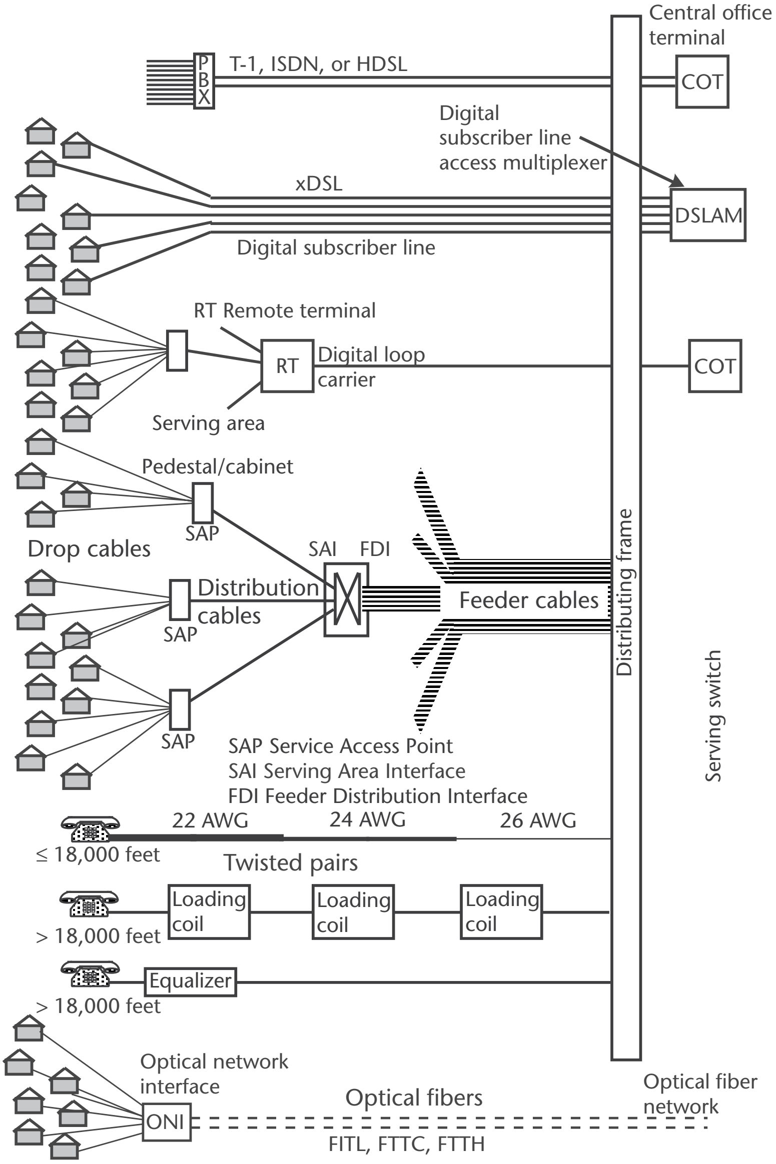


Figure 8.1 Types of local loop facilities.

In many loops, *remote terminals* (RTs) are set up at some distance from the wire center. Here 96, 672, or some other number of channels are aggregated and transmitted over optical fibers between the MDF and the remote terminals. Called *digital loop carrier* (DLC), the channels are distributed from the RTs to customers in the *carrier serving area* (CSA) over distribution and drop cables. The carrier serving area is limited to 9,000 feet from the RT. Any DSLs home on DSLAMs located at the RT.

8.1.1.2 Optical Fibers in the Local Loop

In the local loop, carriers have installed fiber to carry multiplexed signal streams close to their destination. They terminate in *optical network interfaces* (ONIs) where twisted pairs are used to complete the connection to residences or small businesses. Several acronyms are used to identify such installations:

- FITL: fiber in the loop;
- FTTC: fiber to the curb;
- FTTH: fiber to the home.

They are used without precision to indicate various levels of fiber availability. Most carriers are awaiting the development of demand for residential wideband services before making major commitments to these facilities.

SONET rings are employed to connect the main switching center, remote switches, remote terminals, distribution interfaces, and other traffic collection points. Figure 8.2 illustrates the principle of applying SONET in the local communication environment to replace feeder cables. In the figure, a *star-star* arrangement is compared to ring-based structures that employ SONETs. The *ring-bus* structure is constructed from the combination of cable television and incumbent local exchange

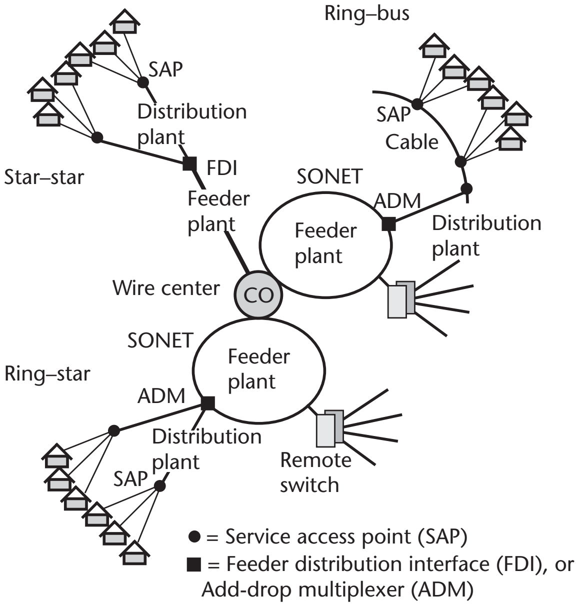


Figure 8.2 Alternative architectures for loop plant.

carrier (ILEC) facilities. The *ring-star* structure is constructed from ILEC facilities. Both arrangements can provide voice, video, and data services.

8.1.2 Modems and Digital Subscriber Lines

For residential applications such as working-at-home and Internet, the bandwidth of the data stream signals must be compatible with the bandwidth of the twisted pair cable that links the user to the network. Substantial processing is required to match the characteristics of the data signals to the line.

8.1.2.1 V.34 and V.90 Modems

Over the years, modem speeds have become faster and faster as designers have found ways to achieve more bits per symbol, and more symbols per second. Standardized by ITU, V.34 and V.90 are the latest in a long line of modems used on two-wire (twisted pair) telephone lines. Adjusted at the time of use to yield reliable performance, V.34 uses a symbol rate between 2,400 baud and 3,429 baud. Employing QAM on both channels of a duplex circuit, it can achieve bit rates of over 30 kbit/s. To prepare for data transfer, V.34 executes a four-part setup routine. Users of V.34 modems who listen during setup can hear them. The following is the four-part setup routine:

1. *Network interaction:* Exchange of signals with receiving modem to establish that the circuit is ready.
2. *Ranging and probing:* Exchange of signals to establish symbol rate, round trip delay, channel distortion, noise level, and final symbol rate selection.
3. *Equalizer and echo canceler training:* Exchange of signals designed to optimize performance of the equalizers and echo cancellers in the send and receive modem.
4. *Final training:* Exchange of known signals to establish setup is complete.

The V.90 modem makes use of V.34 technology in the upstream direction. In the downstream direction it uses 128 special symbols to send at 56 kbit/s. Should the line be unable to support this rate, the number of symbols is reduced with a consequent reduction in bit rate.

8.1.2.2 Digital Subscriber Lines

Digital subscriber lines (DSLs) provide a way to meet demands for high-speed services over existing telephone cable pairs. Moreover, DSLs can be used as alternatives to traditional digital lines (such as T-1 and ISDN PRI). Figure 8.3 shows the concept of using DSLs for residential and small business connections. In the central office, *DSL access multiplexers* (DSLAMs) connect individual DSLs on twisted pairs to a regional high-speed network that provides access to content providers and the Internet. At the CO, POTS services are split from the data signals and directed to the PSTN. In the home, a similar splitting function is performed to separate telephone traffic from data traffic. Taking advantage of significant advances in signal processing and solid-state technology, several types of DSLs have been deployed, and more are in active development. The following sections give some indication of the equipment that is available.

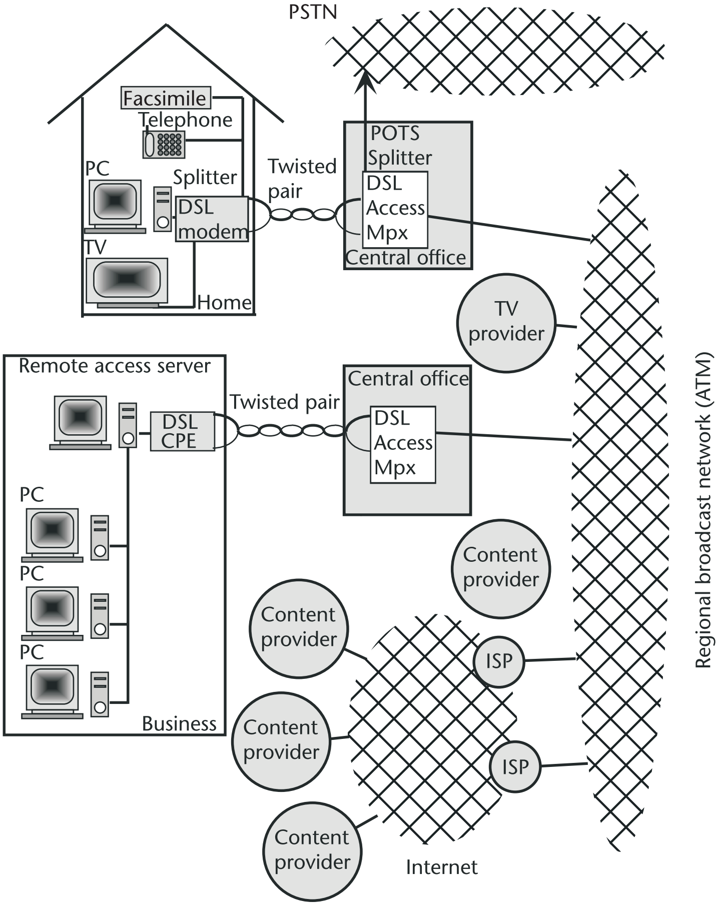
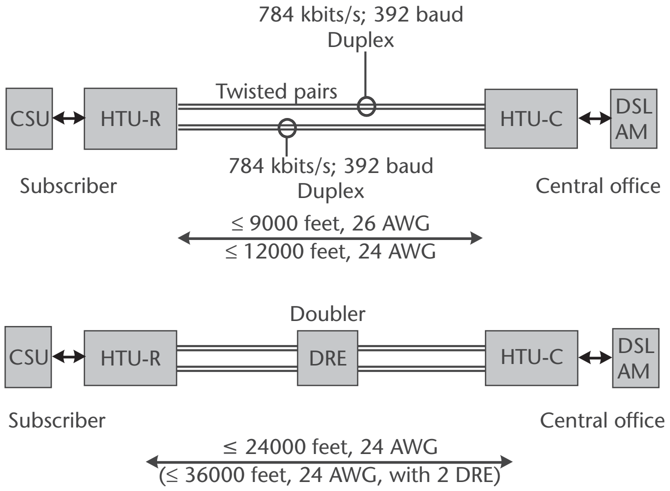


Figure 8.3 DSL network architecture.

8.1.2.3 High-Bit-Rate Digital Subscriber Line

Before the ITU Recommendations for ISDN were formally adopted, attempts were underway to simplify the provisioning of ISDN PRI services for local access. The goal was operation over 26 AWG wire up to 9,000 feet, or 24 AWG wire up to 12,000 feet, without repeaters. Called *high-bit-rate digital subscriber line* (HDSL), the DS-1 stream is split into two streams of 784 kbit/s (768 kbit/s for data, 8 kbit/s for signaling, and 8 Kbits for control). Each is transported over a cable pair giving rise to the term *dual-duplex transmission*. The elimination of repeaters results in bit-error rates of approximately 10^{-10}. This is equivalent to the error performance of fiber optic systems.

For installations greater than 12,000 feet, repeaters (known as *doublers*) are employed. With 24 AWG cable pairs, up to 24,000 feet can be reached with one repeater, and up to 36,000 feet with two repeaters. For installations less than 3,000 feet and greater than 36,000 feet, T-1 is used. Figure 8.4 shows the implementation of HDSL with and without *doublers*. HDSL circuits are designed to assure one-way signal transfer delay is less than 0.5 ms. With one mid-span repeater, the delay is less than 1 ms. Delay is important because some upper layer protocols may time out due to the total end-to-end delay.



HTU-CHDSL Transceiver unit–central office
HTU-RHDSL Transceiver unit–remote
CSU Channel service unit
DSLAM Digital subscriber line access multiplexer
DREHDSL Range extender

Figure 8.4 HDSL implementation.

8.1.2.4 HDSL2

HDSL2 complements HDSL. Sometimes, HDSL2 is called S–HDSL. S–HDSL is also used to refer to the implementation of one-half HDSL (duplex 784 kbit/s on a single pair). Operating over a single pair, HDSL2 provides T-1 speed over 26 AWG up to 12,000 feet. Transmission over a single pair of wires required the development of an efficient spectral shaping signaling technique to minimize crosstalk between adjacent pairs that might be running ISDN, T-1, HDSL, or HDSL2. Known as *overlapped pulse–amplitude modulation with interlocked space* (OPTIS), it supports PAM, QAM, CAP, and DMT (see Appendix A) with overlapping downstream and upstream bit streams. The current modulation format uses trellis-coded PAM with 3 bits per symbol and a 16-level constellation. The signaling rate is 517.3 kbaud.

8.1.2.5 Single-Pair High-Data-Rate Digital Subscriber Line

Single-pair high-data-rate digital subscriber line provides symmetrical services between 192 kbit/s and 2.3 Mbps. Intended for applications such as ISDN, T-1, POTS, frame relay, and ATM, it operates up to 24 kft on a 24 AWG loop. Called G.shdsl, the modulation scheme is similar to HDSL2—trellis-coded PAM with 3 information bits per symbol (a 16-level constellation) and OPTIS spectrum shaping. G.shdsl was standardized by ITU and ANSI.

8.1.2.6 Asymmetrical DSL (ADSL)

ADSL provides unequal data rates in downstream and upstream directions. In addition, the lowest portion of the bandwidth is used for analog voice. ADSL modems use two techniques to achieve downstream and upstream operation.

- *Frequency division multiplexing (FDM):* By dividing the operating spectrum into separate, nonoverlapping frequency bands, a voice channel and upstream and downstream data channels are created. This eliminates self-crosstalk as an impairment.
- *Echo cancellation (EC):* The upstream and downstream channels overlap. This necessitates using echo cancellers and retains self-crosstalk as an impairment.

ANSI specifies the use of DMT and two sets of operating rates for ADSL:

- Downstream 6.14 Mbps, upstream 224 kbit/s, over 24 AWG cable pairs up to 12,000 feet;
- Downstream 4 Mbps, upstream 512 kbit/s, over 24 AWG cable pairs up to 12,000 feet.

A later specification increased the downstream rate to 8.192 Mbps and the upstream rate 640 kbit/s. These speeds are achievable over relatively new copper installations. Available products use either DMT or CAP modulation.

Separating the voice channel from the data channels is achieved with highpass and lowpass filters. The lowpass filter prevents the data streams from adversely affecting the voice service, and the highpass filter prevents voice signals from adversely affecting the data streams. The combination of filters is known as a *splitter*. They are installed at both ends of the subscriber line.

8.1.2.7 Spliterless ADSL (G.lite)

G.lite is a scaled-down version of ADSL that does not require splitters to separate voice from data. This simplification makes installation by subscribers possible. However, installation *does* require lowpass filters (microsplitters) on each telephone. Spliterless ADSL is described as a best-effort transmission system. Achievable downstream/upstream data rates are 640/160 kbit/s to 18,000 feet, 1,024/256 kbit/s to 15,000 feet, and 1,512/510 kbit/s to 12,000 feet.

Ringing signals directed to a telephone connected to G.lite, and off-hook/on-hook activity, can result in impedance changes that unbalance the DSL modem operation and require modem retraining. During retraining, the modems are unable to transmit data. To make retraining as fast as possible, G.lite modems store up to 16 operating profiles.

8.1.2.8 Very-High-Bit-Rate DSL (VDSL)

VDSL is an extension of ADSL technology to rates up to 52 Mbps downstream. The configuration includes twisted pairs between subscribers and an *optical network unit* (ONU). In turn the ONU is connected by fiber to the CO.

As stated earlier in this chapter, the differences between the performance of DSLs reflects the year in which each was standardized and the capability of digital electronics at the time. They represent the determination of owners of existing wire plant to make it usable by those who want high-speed data capability.

8.1.3 Cable Television

The demand for faster response over Internet has provided an opportunity for cable companies to use part of their capacity for Internet access. Using MPEG compression and QAM modulation, modern cable television systems can offer 10 digital video channels in the 6-MHz bandwidth used by one analog television channel. With a cable bandwidth of 550 MHz, they can provide around 900 separate video channels to their customers. Assuming they have difficulty filling more than 500 channels with analog television, digital television, music, pay channels, and the like, up to half of the cable can be used for data transport.

A unique feature of cable connections is they are *always on*. The user does not have to wait for a connection to be established. To send data upstream from individual users to the *cable modem termination system* (CMTS), time division multiplex over a 2-MHz channel is employed. Each user has a private channel. The signals are placed in the frequency band 5 to 42 MHz. To receive data from the Internet, a community of as many as several hundred users shares one 6-MHz channel, Ethernet-style, placed in the frequency band 42 to 850 MHz. Since the channel is capable of up to 40 Mbps, if there are 10 users downloading data simultaneously, each can expect to have an average downloading speed of up to 4 Mbps. With 100 users downloading simultaneously, the average speed drops to 400 kbit/s. Like Ethernet, throughput drops as the number of simultaneous users increases.

8.2 Voice over IP (VoIP)

Most of us employ two networks to meet our communication needs—the PSTN for voice and Internet for data. In fact, many of us use the last mile of telephone company facilities to connect to an ISP to gain access to Internet. The PSTN and Internet are quite different. Making one carry traffic more properly carried by the other ignores the design and economic factors used to implement them and strains their resources. For instance, Internet users expect the local telephone company to support connections for many hours of Web browsing, and VoIP users expect the Internet to provide a steady, uniform stream of voice packets to support satisfactory voice quality. The telephone company has designed its network around average calls of a few minutes duration in the busy hour. It provides high-quality service and numerous features. The Internet is a best-effort network that mixes packets from many users and does not guarantee timely delivery. Indeed, they may not deliver some packets at all.

Since the early 1970s, voice transmission has been the subject of experiments mounted by ARPAnet users. They quickly showed that a virtual duplex circuit could carry intelligible voice in packets. More recently, the Internet has been used to carry voice between terminals operated by enthusiastic Web surfers. Such experiments have stimulated activity in the communications vendor community. The next step, implementation over enterprise IP networks (intranets), is underway. What remains to be done to emulate the telephone companies is provide toll-quality voice with intelligent network features all over the nation. However, carrying millions of calls per hour and providing the kind of quality, features, security, and reliability that telephone customers have come to expect causes the difficulties explode. Unfortunately, providing good voice quality and extensive features is only an aspect of the problem. It is much more difficult to create a signaling system that provides the complex features needed by multimedia communications and interface them to the international world. In this section, I discuss VoIP as a precursor of more exotic services using Internet and PSTN.

8.2.1 Packet Voice

The output of a microphone, the transducer that converts sounds to electrical signals, is a continuous value proportional to the air pressure exerted by the audio source. Voice signals, then, are naturally analog signals. Before packet voice is created, the voice signal must be conditioned and digitized.

The quality of reconstructed coded voice is evaluated by a number of participants in structured listening tests. The results are expressed as a *mean opinion score* (MOS). Reconstructed speech that is not distinguishable from natural speech is rated 5.0 (excellent). Other scores are 4 (good), 3 (fair), 2 (poor), and 1 (bad). Studio quality voice has an MOS between 4.5 and 5.0. Sixty-four-kbit/s PCM voice is known as *toll quality* voice and has an MOS of 4.3. Communication quality voice (i.e., quality acceptable to professional communicators such as airline pilots, military personnel) has an MOS between 3.5 and 4.0. A score below approximately 3.5 is considered unacceptable for most applications.

8.2.1.1 Lower Bit Rate Coding

Sixty four-kbit/s PCM voice is robust and fully up to the exigencies of global telephone service in which it may have to be coded and decoded a number of times before reaching the final destination. Newer voice coding techniques encode PCM samples to produce almost the same quality with far fewer bits per second. These lower bit rate voice coders are complex devices. Most of them are hosted on specialized *digital signal processors* (DSPs). The additional processing means that they impose significant delays on the coded voice stream. This may be troubling to some users. Standardized by ITU, some of these voice coders are:

- *G 726:* Uses *adaptive differential PCM* (ADPCM). Encodes voice to 32 kbit/s with MOS of 4.0 and processing delay of 0.125 ms.
- *G 728:* Uses *low-delay code-excited linear prediction* (LD-CELP). Encodes voice to 16 kbit/s with MOS of 4.0 and processing delay of 0.625 ms.
- *G 729:* Uses *conjugate-structure algebraic-CELP* (CSA-CELP). Encodes voice to 8 kbit/s with MOS of 4.0 and processing delay of 15 ms.
- *G 723.1:* Uses *algebraic-CELP* (ACELP). Encodes voice to 6.3 kbit/s with MOS of 3.8 and processing delay of 37.5 ms.

For comparison, PCM voice is standardized as G711, which uses PCM and encodes voice to 64 kbit/s with an MOS of 4.3 and a processing delay of 0.125 ms.

By using lower bit rate coding, fewer packets are needed to contain a given amount of speech. At 64 kbit/s, each second of speech requires approximately 167 ATM cells (payload 48 bytes/cell). At 7 kbit/s, each second of speech requires approximately 18 cells. For VoIP, G 723.1 uses fewer packets than G 729 with lower voice quality and significantly more processing delay. G 729 uses some 13% more packets than G 723.1 with 5% better voice quality and less than one-half the processing delay. As a reference point, the one-way delay in a geostationary satellite channel is 250 ms. It is noticeable by everyone and is sufficient to cause users significant frustration unless echo cancellers are employed. Delays up to 100 ms are tolerated by most people. Presumably, we shall see further voice coder improvements in the future.

8.2.1.2 Packet Size, Delay, and Loss

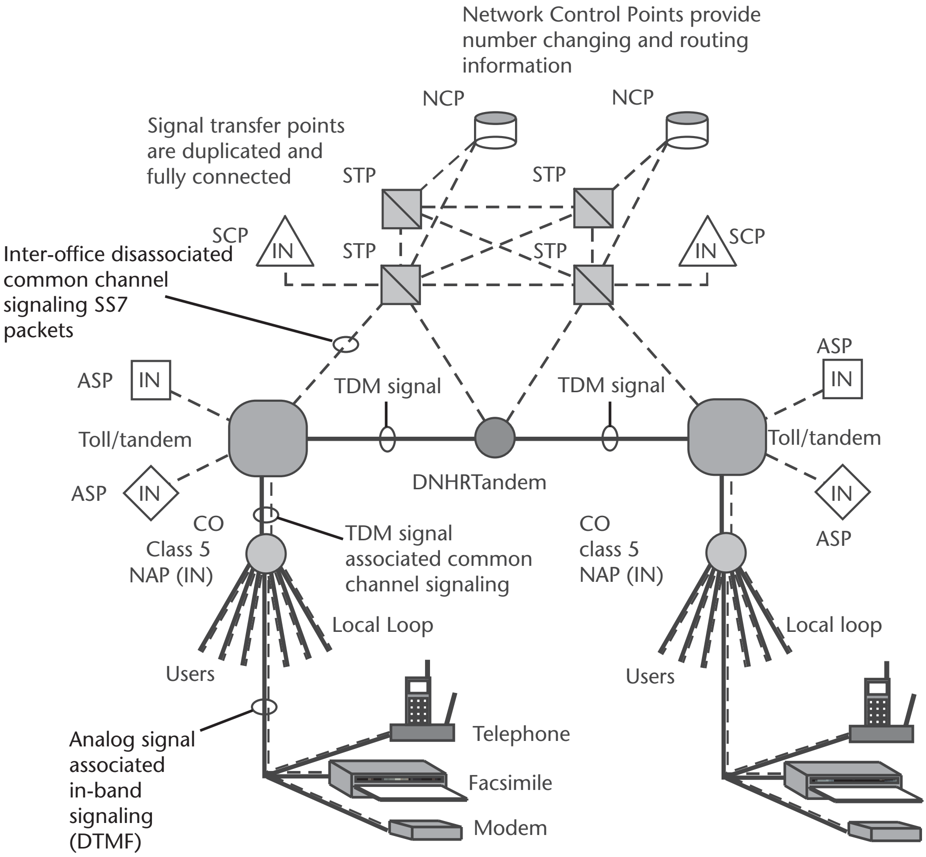
Interactive data requires two simplex channels. One links the send port on terminal 1 to the receive port on terminal 2; and the other links the send port on terminal 2 to the receive port on terminal 1. While one link may carry data in response to a command on the other link, the exact positioning of the response relative to the command is not important. The size of the packet affects the size of the buffer that has to be reserved (at both ends), and the delay incurred in receiving the packet. It does not affect the quality of the exchange. In addition, errored or lost packets are of little consequence since they can be retransmitted and folded into the sequence or used out of sequence.

VoIP is implemented on a duplex circuit. To support a conversation, the timing of the speech on both channels is important. The rhythm of the give and take of a conversation must not be compromised. In addition, packets must arrive on time so that the samples they carry can be used to reconstruct a waveform that contains something close to the original frequencies. If it does not, the participants will not feel *natural*, and their words may be unintelligible at times. Conversationalists have limited tolerance for delay, and fluctuations of delay. Both the end-to-end average delay, and the end-to-end variation of delay, should be small. The successful transmission of Vo IP depends on controlling the mean and variance of packet delay over each channel, and controlling the offset delay between the channels. Packet speech is particularly vulnerable to *tails* in the delay distribution (i.e., random occurrence of long delays). To mitigate their effect, the size of the receiver buffer can be increased. This increases mean delay, but reduces the variance.

Received speech is interrupted and distorted by losing or discarding (due to congestion, perhaps) packets. The severity depends on the packet size. It is generally believed that losses as high as 50% can be tolerated if they occur in very short intervals (less than 20 ms). Intelligibility of 80% is said to occur when the packet size is 20 ms and 10% when the packet size is 200 ms. The optimal packet length is generally accepted to be somewhere between 25 and 75 bytes. It is not just a coincidence that ATM cell relay employs payloads of 48 bytes.

8.2.2 Telephone Signaling

As pointed out earlier, the principle of VoIP is well established; on a private scale, it is implemented successfully. To implement VoIP on a public, national scale is a different matter. Figure 8.5 shows the equipment involved in setting up a long-distance voice call between parties using wire-line facilities. The calling party initiates call setup by signaling over the local loop with tones (DTMF). At the Class 5 central office, signaling is transferred to a digital, common-channel system that makes the


Network Control Points provide number changing and routing information
NCP
NCP
Signal transfer points are duplicated and fully connected
STP
STP
STP
STP
SCP
IN
SCP
IN
Inter-office disassociated common channel signaling SS7 packets
ASP
IN
ASP
IN
TDM signal
TDM signal
Toll/tandem
Toll/tandem
DNHRTandem
ASP
IN
ASP
IN
CO
Class 5
NAP (IN)
TDM signal associated common channel signaling
CO
class 5
NAP (IN)
Local Loop
Local loop
Users
Users
Analog signal associated in-band signaling (DTMF)
Telephone
Facsimile
Modem

ASP Adjunct Processor (IN)
CO Central Office
DNHR Dynamic Non-Hierarchical Routing
DTMF Dual-Tone Multi-Frequency Signaling
IN Intelligent Network
NAP Network Access Point (IN)
NCP Network Control Point
SCP Services Control Point (IN)
SS7 Signaling System #7
STP Signal Transfer Point

Figure 8.5 DTMF, common channel and SS7 signaling in telco network with intelligent network features.

request known to a toll/tandem office. Here, the signaling and calling paths are separated. The request moves into the *Signaling System #7* (SS7) network in packet form. The combination of *signal transfer points* (STPs) and *network control points* (NCPs) in SS7 find a path through the voice network to the toll/tandem serving the called party. Ideally, the available path includes a single, *dynamic nonhierarchical routing* (DNHR) tandem switch. If the called party's line is not in use, the voice connection is set up through the calling CO, the calling toll/tandem, the connecting DNHR tandem, the called toll/tandem, and the called CO. IN features such as calling number ID may be activated. If the called party's line is busy, IN features such as call waiting, call forwarding, and voicemail may be invoked. *Adjunct service points* (ASPs) and *signaling control points* (SCPs) in the intelligent network implement them as appropriate.

Transporting the caller's voice and the response of the called party between originating and terminating terminals is straightforward. Setting up and managing the call requires a significant amount of processing power; adding IN features requires even more. Multiply it by 100 or 200 million telephones, of which perhaps 10 million are active simultaneously, add many tens of carriers, and you begin to see the magnitude of a national VoIP network.

8.2.3 Real-Time Transport Protocols

Meanwhile, several protocols have been developed to support the real-time delivery of voice packets. They work in conjunction with signaling protocols (see Section 8.2.4). Once the connection has been made, they present (or receive) compressed voice segments to (from) the TCP/IP stack. Of note are:

- *Real-Time Transport Protocol (RTP):* Interfaces between the voice stream and existing transport protocols (UDP or TCP). RTP provides end-to-end delivery services for audio (and video) packets. Services include source and payload type identification (to determine payload contents), sequence numbering (to evaluate ordering at receiver), time stamping (to set timing at receiver during content playback), and delivery monitoring. RTP is run on top of UDP or TCP. RTP does not address resource reservation, or guarantee delivery, or prevent out-of-sequence delivery.
- *RTP Control Protocol (RTCP):* A protocol that monitors QoS based on the periodic transmission of control packets. RTCP provides feedback on the quality of packet distribution.
- *Real-Time Streaming Protocol (RTSP):* An application level protocol that compresses audio or video streams and passes them to transport layer protocols for transmission over the Internet. RTSP breaks up the compressed data stream into packets sized to match the bandwidth available between sender and receiver. At the receiver, the data stream is decompressed and reconstructed. Because of the compression and decompression actions, the received quality is unlikely to be equal to the original.

8.2.4 Major Signaling Protocols

The virtual circuit for VoIP is established by signaling protocols. They provide basic telephony features and IN items. Three signaling protocols are competing to provide VoIP services. They are ITU's Recommendation H.323, *Session Initiation Protocol* (SIP), and *Multimedia Gateway Control Protocol* (MGCP). Their relation and the relation of the media transport protocols to the IP stack are shown in Figure 8.6.

8.2.4.1 Recommendation H.323

H.323 is an ITU-developed multimedia communications recommendation that offers audio, video, and facsimile services over LANs. It does not guarantee QoS levels. Focusing on voice services, it provides connections for moderate numbers of users and is incorporated in commercial offerings. As an implementer of VoIP,

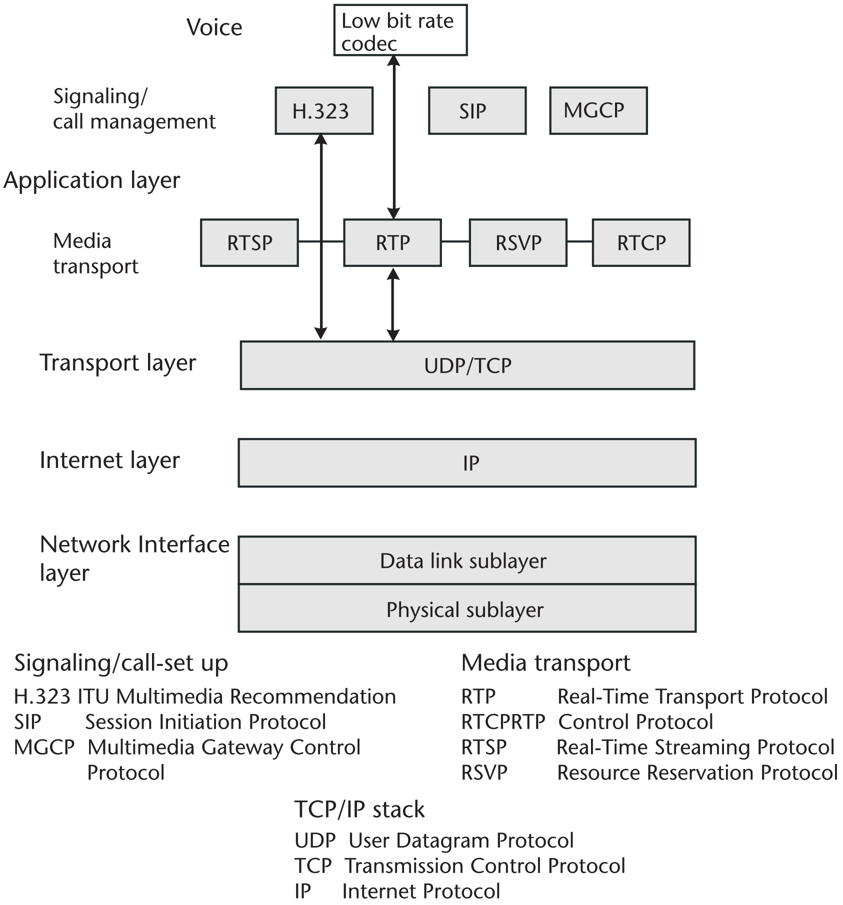

Voice
Low bit rate codec
Signaling/ call management
H.323
SIP
MGCP
Application layer
Media transport
RTSP
RTP
RSVP
RTCP
Transport layer
UDP/TCP
Internet layer
IP
Network Interface layer
Data link sublayer
Physical sublayer
Signaling/call-set up
H.323 ITU Multimedia Recommendation
SIP Session Initiation Protocol
MGCP Multimedia Gateway Control Protocol
Media transport
RTP Real-Time Transport Protocol
RTCPRTP Control Protocol
RTSP Real-Time Streaming Protocol
RSVP Resource Reservation Protocol
TCP/IP stack
UDP User Datagram Protocol
TCP Transmission Control Protocol
IP Internet Protocol

Figure 8.6 TCP/IP stack with VoIP protocols.

H.323 allows the calling and called parties to use their telephone experience including call forwarding, call waiting, and call hold. It is an application-level protocol that mediates between the calling and called parties and the end-to-end transport protocol layer. H.323 uses RTP and RTCP for transport. In Figure 8.6, I have tried to distinguish the domain of H.323 call set up functions and the domain of RTP call transport functions. The general flow of a two-party voice call is as follows:

1. The user goes *off-hook,* causing the call setup protocol of H.323 to issue a dial tone and wait for the caller to dial a telephone number.
2. The dialed numbers are accumulated and stored.
3. After the digits are received, the number is correlated with an IP host that has a direct connection to the destination telephone number or a PBX that will complete the call.
4. The call setup protocol establishes a duplex virtual circuit (using TCP) over the IP network.
5. If a PBX handles the call, the PBX forwards the call to its destination.
6. If RSVP is configured, resource reservations are made to achieve the desired QoS.

7. Call-progress indications (ringing, busy, and other signals that are carried in-band) are carried over the IP network encapsulated in RTCP.
8. Codecs are invoked at both ends of the circuit to provide low bit rate voice, and the call begins.
9. RTCP monitors performance and provides feedback to RTP.
10. When the parties go *on-hook*, the RSVP resource reservations are canceled and the session ends. H.323 becomes idle waiting for the next off-hook signal.

Originally developed to facilitate multimedia communications over local area networks, H.323 operates independently of network topology. Today, most implementations use H.323 with RTP/UDP/IP for speed and simplicity over any IP network. H.323 was an early starter in the VoIP race. Because it is sponsored by ITU, it has experienced wide dissemination and exploitation.

8.2.4.2 Session Initiation Protocol (SIP)

SIP is a signaling protocol developed to facilitate telephone sessions and multimedia conferences in a unicast or multicast private network environment. Through gateways, SIP communicates with public terminals, and provides a limited menu of IN services. In addition, it can connect with private networks that employ H.323, or other signaling protocols. In VoIP use, SIP operates much like the scenario given for H.323. It is claimed to be faster, simpler, and more scalable than H.323.

Developed by a committee of the IETF, SIP uses text-like messages. It does not use other protocols such as RTP, RSVP, and so forth. SIP responds to telephone numbers or URLs and negotiates the features and capabilities of a call prior to setup. It can modify them during the course of a session.

8.2.4.3 Media Gateway Control Protocol (MGCP)

MGCP is a commercial/IETF development designed to facilitate multimedia sessions between the Internet and the PSTN. The *media gateway* (MG) acts between the two networks to translate media streams from circuit-switched networks into packet-based streams, and *vice versa*. MG components may be distributed among several network devices. MGCP employs a series of commands written in ASCII code that contain an action verb (e.g., create, modify, delete, and so forth) and supporting data. The destination station acknowledges each command and may respond with information; the sender correlates any response with the enabling command.

8.3 Final Word

The needs in business and residential markets to have both voice and data (and limited video services) have produced the concept of the *convergence* of voice and data networks into one that offers *multimedia broadband services*. Data enthusiasts see the eventual triumph of packet techniques and the replacement of the PSTN by an expanded and improved Internet. For this to happen, their *technology push* must be converted into *market pull*. Meanwhile, the owners of hundreds of billions of dollars worth of *legacy* systems—the PSTN companies—will develop counter strategies that continue to recoup their investments and provide competing services. It is likely that multimedia broadband services will evolve from the combination of the two networks rather than by one replacing the other.

Communication by electrical, electronic, and optical means is an important, and essential, part of modern life. Global commerce depends on it. Take away the ability to generate data in one place, process it into information in another, and use it anywhere, immediately, and the world economy will slow dramatically. So, too, will the lives of the Internet generation. E-mail, the Web, and pervasive communications from the computer keyboard have permeated the very core of humankind. Between the more than 200 million computers connected to Internet, TCP/IP is the only suite of communication protocols in use. Does anyone doubt its dominance over all others? It makes the Internet what it is, an immensely successful, worldwide, digital communication network.

APPENDIX A

Connections, Codes, Signals, and Error Control

Throughout this book, I have assumed a certain amount of communication knowledge on the part of the reader. For those who need a refresher, several topics are discussed in this appendix.

A.1 Connections

A connection may provide one- or two-way message transport. The former is known as a *channel* and the latter is known as a *circuit*.

- *Channel:* A unidirectional communication path;
- *Circuit:* A bidirectional communication path. Can be considered to be two channels operating simultaneously (one in each direction).

Furthermore, communication can occur in three ways:

- It can be in the style of an *announcement* with information flowing in one direction and no reply possible.
- It can be *interactive* with the participants exchanging information as necessary (sometimes at the same time).
- It can be in the style of a *debate* with the participants addressing each other in turn.

While these examples are personal, they are close matches to the ways in which machines communicate. The connections that support them are identified as follows:

- *Simplex:* Supports announcement-style communication. Messages flow in one direction only—from sender to receiver. Simplex employs a channel.
- *Duplex (sometimes called full-duplex):* Supports interactive communications. Messages can flow in two directions at the same time. Duplex employs a circuit. The term *full-duplex* is used to distinguish a full-time, two-way circuit from a *half-duplex* connection.
- *Half-duplex:* Supports debate-style communication. Messages can flow in both directions, but only in one direction at a time. Many local area networks

are half-duplex—stations receive and transmit, but only one action can occur at a time. Half-duplex employs a single channel if it can be used in either direction, or a circuit in which only one side is used at a time.

In addition, other arrangements in which multiple circuits are operated in parallel, have been implemented, for example, *dual-duplex*, which is a connection with two duplex circuits on which signals are divided by frequency. The composite provides twice the bandwidth of a single circuit. Dual-duplex is used to provide 1.544 Mbps over two twisted pairs for ISDN and HDSL.

A.1.1 Addresses

Addresses are described as:

- *Unicast:* The address of a single station. Used in point-to-point communication.
- *Multicast:* An address that is shared by several stations. Used in point-to-many communication.
- *Broadcast:* An address that is processed by every station on the same segment of the network. Routers do not pass broadcast messages to other networks.

A.2 Codes, Code Words, and Code Sets

Binary symbols are known as *bits*, and sometimes as *binits*. Bits and binits are contractions of the words binary digits. When necessary, the term *binit* is used to distinguish between a binary digit and a symbol in information theory that has a 50% probability of being sent (and is therefore invested with 1 *bit* of self-information). Because a binary symbol can have only two values, it is used in groups of *n* bits. Each *n*-bit group (called a *code* or *code word*) contains a *code set* of 2^n unique codes (bit patterns). For transmission between originating (sending) and terminating (receiving) equipment, the code words are assembled in a stream that contains message, control, and perhaps padding, code words. To communicate, any devices in the communication path must know the meanings of the control codes, and the originating and terminating devices must know the meanings of the message, control, and padding, codes.

A.2.1 Code Word Length

With a set in which the code words are of equal length, the receiver's task of breaking the stream into words is as easy as counting groups of *n* bits. As long as the receiver can count accurately and a reliable start indication is available, it can divide the stream into code words for processing. In applications where the codes occur randomly and all the code words in the code table (i.e., 2^n) are in use, equal length code words achieve maximum efficiency in terms of *bits/character*. Alphanumeric codes do not meet these conditions. For instance, there will be one or more vowels in every text word so that the use of codes that represent vowels far exceeds those that represent consonants. Furthermore, since uppercase letters occur mostly at the beginning of sentences, uppercase letter codes will be used infrequently. In addition, punctuation marks and other text symbols are relatively rare. Nevertheless, equal length codes are used in all general-purpose applications.

A.2.2 Some Popular Codes

Some popular codes are the following:

- *ASCII code:* A 7-bit code standardized by ITU as International Telegraph Alphabet #5 (ITA#5), ASCII contains 128 (i.e., 2^7) code words. They permit the designation of code words as letters (uppercase and lowercase), numbers, punctuation, and control. In Table A.1 72 ASCII codes are shown. The remaining 56 codes are used for punctuation and for additional control purposes. ASCII is the coding scheme used almost universally with personal computers and other devices such as keyboards, printers, and the like. Most often, 7-bit ASCII code is converted to 8-bit code by the addition of a *parity* bit to check the correctness of transmission.
- *EBCDIC:* An 8-bit code developed and used by IBM in all of its larger computers. Table A.2 shows 72 of 256 (i.e., 2^8) EBCDIC characters. The remaining 184 are used for punctuation, other text-related functions, and special functions defined by the user.

Table A.1 Some Members of American Standard Code for Information Interchange

Alphas	*ASCII*	*Alphas*	*ASCII*	*Numerics*	*ASCII*
a	1100001	A	1000001	0	0110000
b	1100010	B	1000010	1	0110001
c	1100011	C	1000011	2	0110010
d	1100100	D	1000100	3	0110011
e	1100101	E	1000101	4	0110100
f	1100110	F	1000110	5	0110101
g	1100111	G	1000111	6	0110110
h	1101000	H	1001000	7	0110111
i	1101001	I	1001001	8	0111000
j	1101010	J	1001010	9	0111001
k	1101011	K	1001001		
l	1101100	L	1001100	*Control*	*ASCII*
m	1101101	M	1001101	SYN	0010110
n	1101110	N	1001110	SOH	0000001
o	1101111	O	1001111	STX	0000010
p	1110000	P	1010000	ETX	0000011
q	1110001	Q	101001	EOT	0000100
r	1110010	R	1010010	ENQ	0000101
s	1110011	S	1010011	ACK	0000110
t	1110100	T	1010100	NAK	0010101
u	1110101	U	1010101	DLE	0010000
v	1110110	V	1010110	ETB	0010111
w	1110111	W	1010111		
x	1111000	X	1011000		
y	1111001	Y	1011001		
z	1111010	Z	1011010		

Format ^{MSB}xxxxxxx^{LSB}

Table A.2 Some Members of Extended Binary Coded Digital Interface Code

Alphas	*EBCDIC*	*Alphas*	*EBCDIC*	*Numerics*	*EBCDIC*
a	10000001	A	11000001	0	11110000
b	10000010	B	11000010	1	11110001
c	10000011	C	11000011	2	11110010
d	10000100	D	11000100	3	11110011
e	10000101	E	11000101	4	11110100
f	10000110	F	11000110	5	11110101
g	10000111	G	11000111	6	11110110
h	10001000	H	11001000	7	11110111
i	10001001	I	11001001	8	11111000
j	10001010	J	11001010	9	11111001
k	10001011	K	11001011		
l	10001100	L	11001100	*Control*	*EBCDIC*
m	10001101	M	11001101	SYN	00110110
n	10001110	N	11001110	SOH	00000001
o	10001111	O	11001111	STX	00000010
p	10010000	P	11010000	ETX	00000011
q	10010001	Q	11010001	EOT	00110111
r	10010010	R	11010010	ENQ	00101101
s	10010011	S	11010011	ACK	00101110
t	10010100	T	11010100	NAK	00111101
u	10010101	U	11010101	DLE	00010000
v	10010110	V	11010110	ETB	00100110
w	10010111	W	11010111		
x	10011000	X	11011000		
y	10011001	Y	11011001		
z	10011010	Z	11011010		

Format ^{MSB}xxxxxxxx^{LSB}

- *Universal character set (UCS):* Also known as unicode. A 16-bit code intended to support all world languages, particularly Chinese, Japanese, and Korean. 65,536 (i.e., 2^{16}) code words are available.

A.2.3 Parity Bits

To provide a check on the integrity of transmission, a parity bit may be added to ASCII characters. Its value is determined by the number of ones (odd or even) in the character and whether odd parity or even parity is employed:

- *Odd parity:* If the number of 1s in the character is odd, the parity bit is 0 so that the number of 1s in the character plus the parity bit remains odd. If the number of 1s in the character is even, the parity bit is 1 so that the number of 1s in the character plus parity bit is odd.
- *Even parity:* If the number of 1s in the character is odd, the parity bit is 1 so that the number of 1s in the character plus parity bit is even. If the number of 1s in the character is even, the parity bit is 0 so that the number of 1s in the character plus parity bit remains even.

Should a bit error occur subsequent to the addition of the parity bit, the wrong parity state will exist and the receiver will declare an error is present. In fact, the parity bit will detect one, three, five, or seven errors (i.e., all odd numbers of errors) in the character. However, the parity bit will not detect two, four, and six errors (i.e., all even numbers of errors) in the character. Parity checking is also known as *vertical redundancy checking* (VRC).

A.2.4 Bit Order

The code words in Tables A.1 and A.2 are treated as binary numbers. The bit order is important. The *least significant bit* (LSB) is on the right end of each word, and the *most significant bit* (MSB) is on the left end. For ASCII with parity and EBCDIC, the codes are 8-bit groups for which the bit positions are numbered as follows:

$$^{\mathrm{MSB}}76543210^{\mathrm{LSB}}$$

In ASCII with parity, position 7 contains the parity bit, and positions 0 through 6 contain the character. In common with computer usage, an 8-bit group is called a *byte*. How do we read bytes into a serial stream? There are two ways to do it. We may read from the LSB to the MSB or from the MSB to the LSB. Is one way better than the other? No, they are equally effective. In fact, both methods are in use. For instance, in an Ethernet local area network, the letter *a*, which, in ASCII is

$$^{\mathrm{MSB}}1100001^{\mathrm{LSB}}$$

will be read into the data stream as

$$\Leftarrow 1000011$$

In a Token Ring local area network, it will be read into the data stream as

$$\Leftarrow 1100001$$

Ethernet is said to employ *little Endian* or *canonical* format and Token Ring is said to employ *big Endian* format:

- *Little Endian or canonical format:* Bits are read in ascending order from the least significant bit to the most significant bit. Bytes are numbered left to right, from 0 to *N*, and are read in ascending order.
- *Big Endian format:* Bits are read in descending order from the most significant bit to the least significant bit. Bytes are numbered left to right, from 0 to *N*, and are read in ascending order.

Figure A.1 shows the difference between these formats for a group of 6 bytes. The little Endian strategy results in a stream consisting of bits:

$$\Leftarrow 0\rightarrow 7,\ 8\rightarrow 15,\ 16\rightarrow 23,\ 24\rightarrow 31,\ 32\rightarrow 39,\ 40\rightarrow 47$$

The big Endian strategy results in a stream consisting of bits:

$$\Leftarrow 7\rightarrow 0,\ 15\rightarrow 8,\ 23\rightarrow 16,\ 31\rightarrow 24,\ 39\rightarrow 32,\ 47\rightarrow 40$$

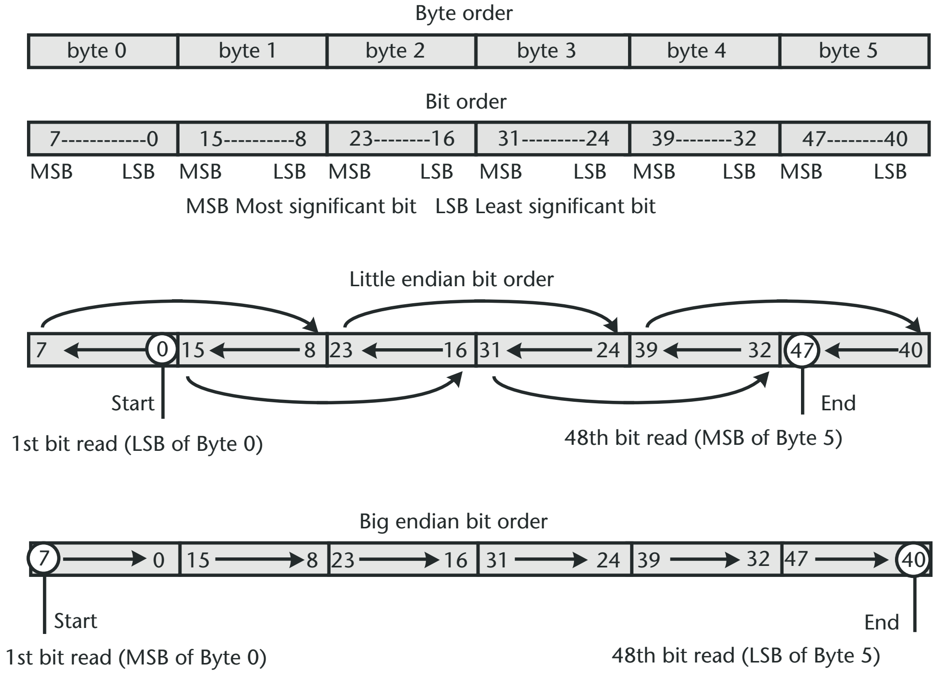


Figure A.1 Big Endian and little Endian bit order.

Obviously, to decipher the data stream correctly, it is important to know which strategy has been employed.

In a digital voice network, an 8-bit group that represents the magnitude of a sample of a voice signal is called an *octet*. Bit #7 indicates whether the value defined by bits 0 through 6 is positive (1) or negative (0). Bit #7 is always transmitted first. In this book, to avoid making the distinction and bowing to general practice, all 8-bit words are called bytes.

A.2.5 Block Coding

To *fine-tune* the performance of the electronics and the data stream, block codes are used. For instance, 1000BASE-X Ethernet employs 8B/10B coding. Each byte is substituted by a 10-bit code word so that the 256 unique bytes are replaced by 256 of the 1,024, 10-bit code words. The words are chosen so that they never contain fewer than four 1s or four 0s and have a 1s/0s imbalance of no more than two. The code words consist of four 1s and six 0s, five 1s and five 0s, or six 1s and four 0s.

In addition to the first 256, 10-bit code words, a second set is defined. They are the bit inverse of the first set. Together, the first code word and its alternate contain ten 1s and ten 0s. To maintain a balance between 1s and 0s in the bit stream, the transmitter maintains a tally of whether more 1s than 0s or more 0s than 1s have been transmitted. Called the *running disparity* (RD), its value determines whether the transmitter selects the next code word as the one with more 1s than 0s, or the alternate with more 0s than 1s. Code words that contain five 1s and five 0s will not change RD. Its value remains constant until presented with the next unbalanced pair of code words. The remaining 512 10-bit code words in the 1,024-word *code space* are used to encode special functions.

A.2.6 Scrambling

Certain patterns of data produce constant level signals that can be troubling to transmission systems. For instance, strings of 0s may cause the terminals to lose synchrony. Other patterns can be equally as bad (e.g., strings of alternating 1s and 0s in the case of 2B1Q). To avoid these effects, many transmission systems *scramble* the bit stream before producing the physical signal. Figure A.2 shows the principle of scrambling. By performing logical operations on the bit stream at the transmitter, strings of the same symbol, or repeated patterns of symbols, are broken up and rendered *pseudorandom*. At the receiver, by repeating the logical changes, the scrambled sequence is *descrambled* and the original data stream is restored. Because it is automatic and completely reversible, scrambling is transparent to the sender and the receiver. It is widely used on long-distance connections.

A.2.7 Hexadecimal Representation

Because writing 8-bit bytes can be tedious and subject to errors, hexadecimal notation is used to represent them. Bytes are divided into two 4-bit binary words (4 bits, or half a byte, is known as a *nibble*), whose decimal values (0 to 15) are represented by the digits 0 through 9 and the letters A through F. Table A.3 shows the complete representation. As an example,

$$01111110 = 0111,1110 = 0 \times 7E$$

The symbols 0x are used to mean hexadecimal. Other examples are:

$$10101010 = 0 \times AA;\ 10101011 = 0 \times AB; \text{ and } 00100000 = 0 \times 20$$

A.3 Operating Modes

Code words are sent individually (asynchronously), or as part of a frame (synchronously). The former mode is generally employed with keyboards and other

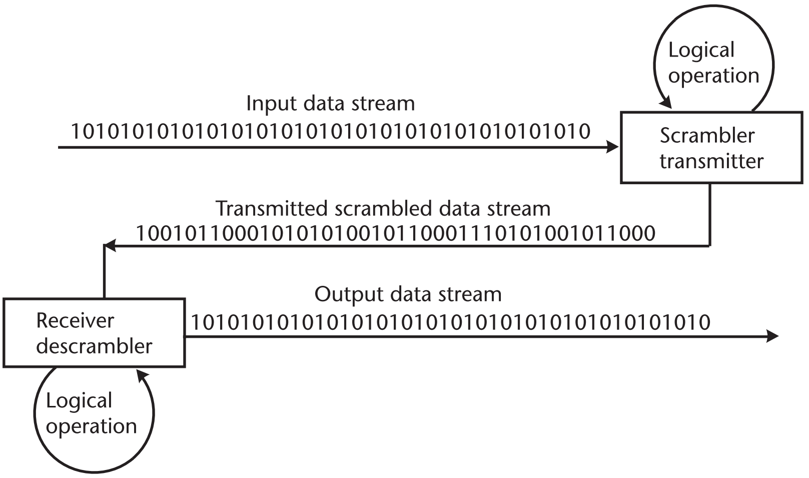


Figure A.2 Principle of scrambling.

Table A.3 Hexadecimal Codes

0 = 0000	1 = 0001	2 = 0010	3 = 0011
4 = 0100	5 = 0101	6 = 0110	7 = 0111
8 = 1000	9 = 1001	A = 1010	B = 1011
C = 1100	D = 1101	E = 1110	F = 1111

Format ^{MSB}xxxx^{LSB}

human/machine interaction devices at the edges of the network. The latter is employed universally by equipment within the network.

A.3.1 Asynchronous Operation

An *asynchronous operation* is an operation in which characters are framed by start and stop bits and sent as they are generated. A straightforward example of asynchronous operation is my use of a keyboard to input words into a data file in my *personal computer* (PC). As I type each character, use the space bar to separate words, or hit the enter key to form paragraphs, unique ASCII text and control codes are transmitted to my PC. Because I type at different speeds, the code words are generated at irregular intervals. Each word consists of 8 physical bits whose pulse shape and repetition rate is tightly controlled. To let the receiver know what is going on, a start bit is added to the beginning of the character, and a stop bit is added to the end. Traditionally, start bits are 0s and stop bits are 1s. In many cases, 2 stop bits are sent to emphasize the end of the word. Thus, ASCII *a* with parity bit *P* will be entered into a little Endian bit stream as:

⇐S1000011Pss

where S = start bit and s = stop bit.

A.3.2 Synchronous Operation

Synchronous operation is an operation in which a fixed number of characters are assembled in sequence without start and stop bits. To the sequence a header is added in front and a trailer is added at the rear to form a *frame*. (In some cases, the header or the trailer is omitted.)

Figure A.3 shows the arrangement of a simple frame. The header indicates the start of the frame and contains the address of the destination, if needed. The trailer contains information with which to check for errors and indicates the end of the frame. As noted earlier, the header and/or trailer fields may be omitted in some circumstances. In other modes of operation they will contain additional information needed to support the style of operation in progress. Synchronous operation is implemented in two ways depending on whether synchrony between the receiver and the incoming frame is achieved by internal or external means.

A.4 Signals

It is easy to get lost in the logic of digital communication and forget that communication cannot occur until signals are generated and dispatched. A basic understanding

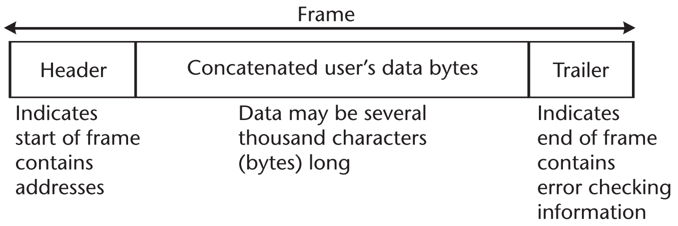

Frame
Header
Concatenated user's data bytes
Trailer
Indicates start of frame contains addresses
Data may be several thousand characters (bytes) long
Indicates end of frame contains error checking information

Figure A.3 Components of the frame.

of the types of signals can help explain some of the engineering mystery surrounding the physical layer.

A.4.1 Signal Classification

Signals are classified by the way in which their values vary over time, thus:

- *Analog:* A continuous signal that assumes positive, zero, or negative values. Changes occur smoothly and rates of change are finite.
- *Digital:* A disjoint signal that assumes a limited set of positive, zero, or negative values. Changes of value are instantaneous, and the rate of change at that instant is infinite—at all other times it is zero. In practice, they are pulse-type signals with finite rise and fall times. The peaks assume a limited set of positive, zero, or negative values.
- *Binary:* A digital signal that has two values.

Analog, digital, and binary are concepts that allow us to divide the communication world into classes that require different technical procedures. In addition, signals may be divided by the degree of certainty with which their behavior is known:

- *Deterministic:* At every instant, a deterministic signal exhibits a value that is related to values at neighboring times in a way that can be expressed exactly. Because determinism requires knowledge in the future as well as the past, deterministic signals only exist in academic exercises where they are analyzed using classical methods.
- *Probabilistic:* A signal whose future values are described in statistical terms based on past values. Probabilistic signals come closer to the real world. They include uncertainty, but still require some relation between the past and the future.
- *Random:* A probabilistic signal whose values are limited to a given range. Over a long time, each value within the range will occur as frequently as any other value. True randomness is almost impossible to attain. Nevertheless, many of the parameters contained in performance specifications are based on random occurrences.

Furthermore, signals are classified according to their *bandwidth*, which is a measure that is applied to frequency-limited signals (i.e., signals whose energy exists within a specific frequency band and nowhere else). Bandwidth is the difference (in hertz) between the highest frequency at which signal energy is detected and the lowest frequency at which signal energy is detected.

For analog signals, bandwidth is measured at some point such as 3 dB or 6 dB down from the signal peak. The decibel (dB) is a logarithmic measure of power ratio. Three dB corresponds to a power level that is one-half of maximum power. Six dB corresponds to a power level that is one-quarter of maximum power. For binary signals, the sharp changes in values give rise to energy throughout the frequency spectrum. Consequently, bandwidth is not easy to measure. Instead, it is stated in terms of the bit rate. Binary signals are loosely classified as follows:

- Narrow band(width): Up to 1.544 Mbps (T-1);
- Wide band(width): 1.544 Mbps to 44.736 Mbps (T-3);
- Broad band(width): Above 44.763 Mbps.

In addition, position with respect to the frequency axis is used to distinguish between signals:

- *Baseband signal:* An original unprocessed message signal. The energy it contains occupies a frequency band that may include 0 Hz (i.e., dc level). The energy of a baseband signal occupies a fixed, unchanging position in the frequency domain.
- *Passband signal:* A complex signal produced by using a baseband signal to modify a property of another signal (called the *carrier signal*). The energy of the passband signal occupies a range (the passband) that encompasses the frequency of the carrier signal, or is contiguous with it. The sideband components of the passband signal carry the information contained in the baseband signal. A passband signal may be moved in the frequency plane by changing the frequency of the carrier signal.

A.4.2 Baseband Signal Formats

Several digital baseband signal formats are in common use. Examples are shown in Figure A.4. They all represent the same bit sequence (101100111000). They are:

- *Nonreturn to zero (NRZ):* A positive current represents 1 and zero current represents 0. Sometimes called *unipolar* signaling, NRZ is used in integrated circuit chips and other circuits, as well as in Gigabit Ethernet. Reliable timing information can be obtained from the signal provided some minimum number of bit transitions occurs in the data stream.
- *Nonreturn to zero, invert on ones (NRZI):* Alternating positive and zero currents represent 1. The same current as the previous 1 represents 0. Put another way, the signal is unchanged for 0, and changes from its previous state for a 1. The strategy of inverting on ones produces a narrower frequency spectrum than NRZ. NRZI is used in FDDI and 100BASE-FX Ethernet.
- *Multilevel threshold-3 (MLT-3):* A sequence of positive, zero, and negative currents represent 1. The same current as the previous 1 represents 0. MLT-3 is a bipolar version of NRZI.

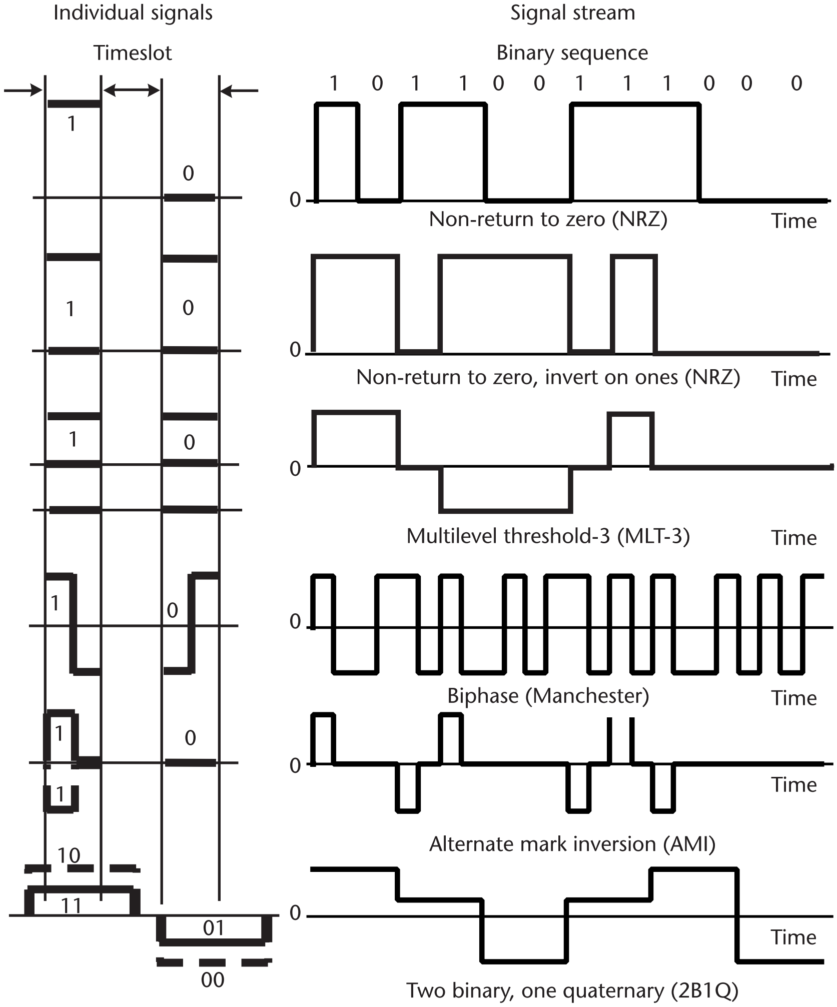


Figure A.4 Examples of binary signal formats.

- *Biphase or Manchester:* A positive current pulse of width one-half time slot, which changes to a negative current pulse of equal magnitude and width one-half time slot, represents 1. A negative current pulse of width one-half time slot, which changes to a positive current pulse of equal amplitude and width one-half time slot, represents 0. The changeover occurs exactly at the middle of the time slot, so that the signal is always *zero-mean*. Furthermore, because the level changes in each bit position, recovering a reliable timing signal is guaranteed. This convenience is bought at the expense of a frequency spectrum that is twice as wide as that of NRZ. Manchester is a popular signaling technique for short links between high-speed equipment, and is used in 10-Mbps Ethernet systems.
- *Alternate mark inversion (AMI):* Return-to-zero current pulses that alternate between positive and negative represent 1. The absence of current pulses represents 0. Thus, long strings of 0s produce no current. By changing the polarity of a pulse (from what it should be), a *violation* is created. In this way,

equipment can introduce *phantom* signals for special purposes and compensate for the occurrence of unusual conditions. AMI was developed for T-1 transmission facilities. To ensure reliable clock recovery, T-1 requires an average of at least 12.5% (i.e., 1 in 8) 1s in the data stream, and no more than 15 0s at once.

- *Two binary, one quaternary (2B1Q):* Four signal levels (± 3 and ± 1) each represent a pair of bits. Of each pair, the first bit determines whether the level is positive or negative (1 = +ve, 0 = –ve) and the second bit determines the magnitude of the level ($1 = |1|$, $0 = |3|$). For long sequences of 1s or 0s or alternating 1s and 0s (i.e., 101010...), 2B1Q signaling produces constant currents. 2B1Q was developed for ISDN transmission facilities.

The formats in Figure A.4 are theoretical. In practice, due to the presence of inductance and capacitance, it is impossible to obtain the sharp, square corners included in the diagrams. Templates define actual pulse shapes. They allow overshoots, undershoots, and finite rise and fall times.

A.4.2.1 Pulse Code Modulation

Two processes are required to digitize an analog signal:

- *Sampling:* Provides a series of discrete signals that represent the amplitude of the analog signal at the sampling time. Usually, sampling is done at regular intervals (such as 8,000 samples per second for PCM).
- *Quantizing:* Converts the sample values to the nearest digital level so that the digital number representing that level approximates the sample. Assigns the same byte value to samples that differ by less than the interval between contiguous levels.

Because low levels of energy are important to intelligibility, the quantizing process is adjusted so that more levels are devoted to low levels of signal than to high levels of signal. *Companding* means that to achieve an even distribution of samples over the range of quantizing levels, high-energy signals are compressed and low-energy signals are expanded. When reconstructing the voice signal, companding is reversed.

These procedures are used in the telephone network. Sampling the analog voice signal at 8,000 times per second produces a 64-kbit/s digital signal. The samples are quantized into +/– 128 levels identified by the 256 codes of an 8-bit byte (8,000 samples per second × 8 bits per sample = 64 kbit/s). Called *pulse code modulation* (PCM) voice, it is the basis for the speeds used throughout the parts of the PSTN that are digital, and limits the reconstructed signal to a 4-kHz bandwidth.

A.4.3 Passband Formats

The sideband components of the passband signal carry the information contained in the baseband signal. A passband signal may be moved on the frequency axis by changing the frequency of the carrier signal. A carrier wave of amplitude A signal units, frequency f hertz, and phase ϕ radians can be modulated by a message as follows:

- *Amplitude modulation:* The amplitude (A) of the carrier is varied based on the value of the modulating signal.
- *Frequency modulation:* The frequency (f) of the carrier is varied based on the value of the modulating signal.
- *Phase modulation:* The phase (ϕ) of the carrier is varied based on the value of the modulating signal.

Figure A.5 shows examples of amplitude, phase, and frequency modulation using a binary signal. In all diagrams, 1s are represented by two full cycles of the carrier signal. The representation of 0s depends on the modulation. In amplitude modulation, 0s take on zero signal level. In phase modulation, 0s are two full cycles of the carrier signal phase-shifted through 180°. In frequency modulation, 0s are formed from four cycles of a carrier signal at twice the frequency of the original carrier signal. These actions are called *keying*, and they are known as *amplitude shift keying* (ASK), *frequency shift keying* (FSK), and *phase shift keying* (PSK). In the examples, the magnitudes of the changes between 1s and 0s were chosen for diagrammatic simplicity; practical modulation schemes use many different values.

A.4.3.1 Symbols, Bauds, and Bits

A *symbol* is a repetitive signal element that represents a single binary element, or a group of binary elements. A *baud* is a signaling rate of one symbol per second.

In the examples of ASK, FSK, and PSK, in Figure A.5, each symbol is equivalent to one bit so that the signaling rate (in bauds) is equal to the bit rate (in bits per second). Symbols can be constructed so that they represent more than one bit. An example is shown in Figure A.6. By employing four signals of equal amplitude but phase-shifted by 90° from one another, four unique signals are created. This tech-

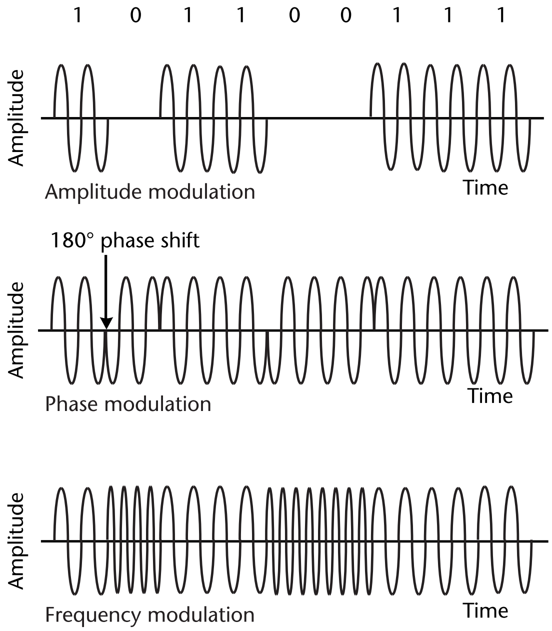


Figure A.5 Amplitude, phase, and frequency keying.

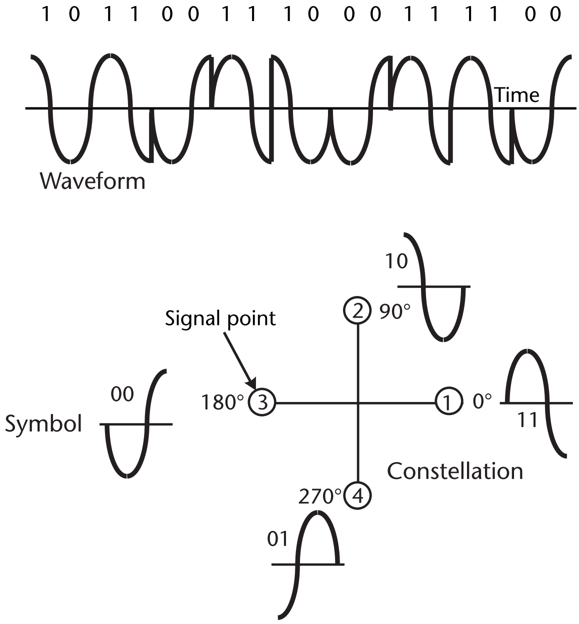


Figure A.6 Example of QAM to create a signal in which each symbol represents 2 bits.

nique is known as *quadrature amplitude modulation* (QAM). The parameters of the four symbols are shown in the center of Figure A.6. Such a diagram is known as a *constellation*. Each symbol is a 270° segment of the carrier signal that starts at carrier phase angles of 0°, 90°, 180°, and 270°. The assignment of codes to the signal points is arbitrary. Once made, however, they must be preserved for the receiver to interpret the received signal correctly. In the upper half of Figure A.6 the waveform corresponding to the data stream at the top of the figure is shown. A comparison with Figure A.5 reveals that twice as many bits are contained in the signal burst. With each symbol representing 2 bits, this was to be expected. Under these circumstances, the signal in Figure A.6 achieves a bit rate that is twice the baud rate.

In the 1920s, Harold Nyquist showed that the maximum signaling rate over a channel with a passband B Hz is 2B baud. This is known as the *Nyquist rate*.

The passband of a given signal is governed by the physical parameters of the transmitter, the transmission medium, and the receiver. In radio systems, filters at the transmitter and receiver establish the passband. They are tightly controlled to prevent one system interfering with another. In the telephone network, a passband (4 kHz) is established by the digital sampling rate (8 ksamples/sec). This gives an upper bound for the signaling rate of 8 kbauds, or 8 ksymbols/sec. In practice, the Nyquist limit cannot be achieved without complex processing of the signal stream.

A.4.3.2 Complex Modulation Techniques

Implementations of complex modulations may have constellations with as many as 256 or 512 signal points. They correspond to operating at 8 bits/baud and 16 bits/baud. Great care is taken to arrange the signal points so that they are equidistant from one another. This is necessary to provide an equal area around each point in which errored signals may fall. An example of a 16-point constellation (4 bits/baud) is given in Figure A.7. In the upper diagram, the signal points are formed from a minimum combination of two amplitudes and eight phase angles. The 16 signal points are not uniformly distributed over signal space and the inner ring of eight points has less signal space per point to cope with errors than the outer ring. To correct this, a practical 16-point constellation is formed out of the combination of three amplitudes and 12 phase angles shown in the lower diagram. The signal points are distributed uniformly, and each has the same signal space as its neighbors.

The successful deployment of various *flavors* of digital subscriber lines depends on the use of complex passband signal processing algorithms. Some of them are:

- *Pulse amplitude modulation:* A popular modulation format uses trellis-coded PAM with 3 bits per symbol and a 16-level constellation. The coding employs twice as many signal points in the constellation as are needed to represent the signal points. This redundancy is a form of forward error correction coding and is used to reduce errors.
- *Carrierless amplitude and phase (CAP) modulation:* A passband technology based on QAM. With a 256-point constellation (i.e., 8 bits per symbol) and a

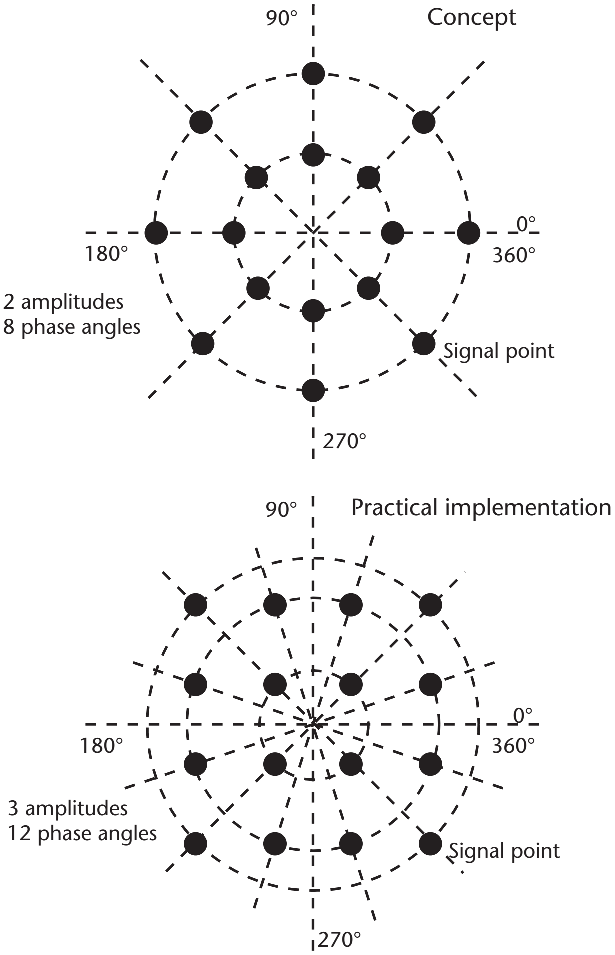


Figure A.7 16-point signal constellations.

signaling rate of 1,088 kbaud, bit rates of 8.704 Mbps are achieved. CAP employs trellis coding, Viterbi decoding, and Reed-Solomon forward error correction. Viterbi decoding implements maximum likelihood decoding of convolutional codes. Reed-Solomon codes employ groups of bits (known as *symbols*). With k information symbols, r parity symbols, and code words of length $n = k + r$, it is able to correct $r/2$ errors in a symbol.

- *Discrete multitone transmission (DMT):* A passband technology, DMT operates over a range of frequencies. The available frequency band is divided into parallel channels (4.3125 kHz wide). Known as *bins*, they employ QAM with a 4 kbaud symbol rate and up to 15 bits per symbol.

A.4.3.3 Spread Spectrum Modulation

Developed largely by the military as a means of hiding communications from adversaries, spread spectrum signals are hard to intercept and almost impossible to jam. Examples of their use are *global positioning systems* (GPSs), mobile telephones, *personal communication systems* (PCSs), and *very small aperture satellite systems* (VSATs).

Spread spectrum modulation is a technique in which the message-bearing modulated signal is processed (i.e., modulated again) to occupy a much greater bandwidth than the minimum required to transmit the information it carries.

The spectrum is spread in two ways:

- *Frequency hopping:* The frequency of the carrier of the narrowband-modulated message signal is caused to hop from one value to another in a high-speed, pseudorandom manner across the spread spectrum.
- *Direct sequence:* The narrowband-modulated message signal is modulated by a high-speed pseudorandom sequence to produce a signal that extends across the spread spectrum.

Because the spread spectrum signal has a lower power density (i.e., watts/hertz) than the original signal, it creates little interference in other signals in the same frequency band.

To generate a direct sequence spread spectrum signal requires remodulating the modulated message signal with a high-speed semirandom sequence of 1s and 0s. Each element (1 or 0) is called a *chip*, the bit speed is known as the *chipping rate*, and specific arrangements of 1s and 0s are a *chipping code*. If each user is assigned a chipping code that is orthogonal (a mathematical term meaning that the integral of the product of any two codes is zero) to others in use, each code stream can be distinguished from the codes of other users. Thus, many users can communicate in the same frequency space. This is known as CDMA. It is widely used in mobile telephone systems and PCSs.

Code division multiple access (CDMA) is a direct-sequence spread spectrum technique in which all stations in the network transmit on the same carrier and use the same chip rate to spread the signal spectrum over a wide frequency range. Each station employs a code that is orthogonal to the codes used by others. Each receiver sees the sum of the spread spectrum signals as uncorrelated noise. It can demodulate a specific signal if it has knowledge of the spreading code and the carrier frequency.

In the act of despreading the direct sequence spread spectrum signal, the receiver spreads any interfering signals, thereby improving the signal-to-noise ratio. Figure A.8 illustrates the relationships among: the original modulated message-bearing signal; the direct sequence, spread spectrum, message-bearing signal; interfering noise; and the despread spread spectrum message-bearing signal at the receiver. CDMA is a proven method of accommodating a large number of users in limited spectrum space without mutual interference.

A.4.3.4 Orthogonal Frequency Division Multiplex (OFDM)

In some ways, OFDM is the antithesis of CDMA. Instead of spreading all users on a single carrier using individual chipping codes, OFDM encodes a single user on several carriers. It splits a wide frequency band into narrow channels and inverse multiplexes a user's data signal on the subcarriers occupying a channel. Inverse

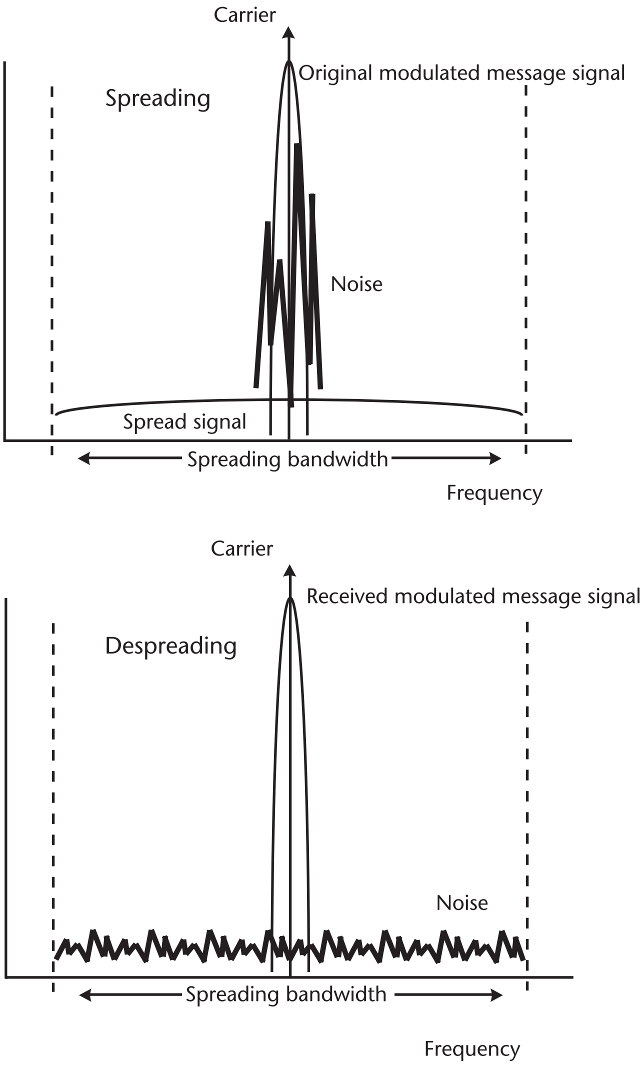


Figure A.8 Illustrating the spreading of a message signal and the despreading of a spread spectrum signal to yield the message signal and mitigate noise.

multiplexing is the action of splitting a higher-speed data stream into several slower-speed streams that are carried on separate channels and recombined at the terminating point. The channels are selected so that they overlap but the carriers do not interfere with each other (i.e., they are orthogonal). OFDM uses the *inverse fast Fourier transform* (IFFT) to create a composite signal from the inverse multiplexed data signal. In signal analysis, the Fourier transform provides a means of transforming a time-varying signal into its equivalent frequency components. The *fast Fourier transform* (FFT) is an implementation of the Fourier transform that produces a signal waveform from a finite number of sine and cosine waves. The inverse Fourier transform provides a means of transforming frequency components into an equivalent time-varying signal. At the receiver, the data stream is reconstructed using FFT.

A.5 Error Control

Noise corrupts the wanted signal and can produce errors in digital signals. Because the noise signal is random, it may add to, or subtract from, the signal pulse train and destroy the certainty of which level is present. Arguably, error control—the detection and correction of errors—is the most important value-added service performed by sending and receiving equipment.

Error control is a cooperative activity between a sender and receiver in which the sender adds information to the code words and/or within the frame to assist the receiver to determine whether an error has occurred. If it has, the sender and/or receiver work together to correct it.

Figure A.9 shows the principle of error control. It is divided into error detection and error correction.

A.5.1 Error Detection

Several techniques are available that detect the presence of an error or errors in the frame received. They have different capabilities.

A.5.1.1 Vertical Redundancy Checking

One method of error detection adds parity bits to individual codes. I discussed this technique with respect to ASCII code in Section A.2.

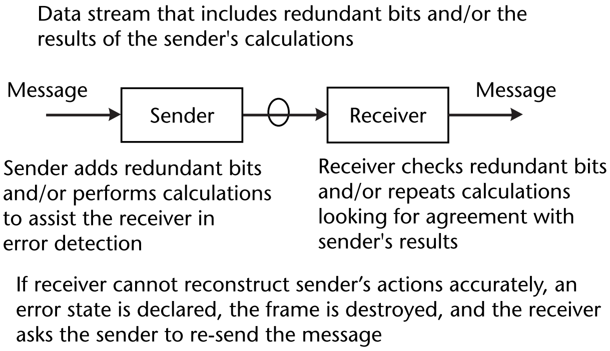

Data stream that includes redundant bits and/or the results of the sender's calculations
Message
Sender
Receiver
Message
Sender adds redundant bits and/or performs calculations to assist the receiver in error detection
Receiver checks redundant bits and/or repeats calculations looking for agreement with sender's results
If receiver cannot reconstruct sender’s actions accurately, an error state is declared, the frame is destroyed, and the receiver asks the sender to re-send the message

Figure A.9 Principle of error control.

A.5.1.2 Longitudinal Redundancy Checking

Bit-level error detection can be extended to check the entire sequence of bits between the header and trailer in a frame. The sender calculates parity bits for the sequences of bit positions #0, #1, ..., #7. They are placed in a byte located in the trailer. This byte is known as the *block check character* (BCC). At the receiver, the same calculations are run on the received frame. If the received BCC is the same as that calculated by the receiver, the receiver has some assurance that the transmission does not contain errors. By using the combination of VRC and LRC, it is possible to locate the bit position of single errors. Like VRC, LRC only detects odd numbers of errors.

A.5.1.3 Checksum

By treating the entire bit stream or segments of the bit stream as binary numbers, error detection can be based on calculations. One process adds them together as 8-bit or 16-bit numbers and determines the ones complement of the result. The sender attaches it to the bit stream it sends to the receiver. The receiver performs the same addition and includes the ones complement. If the result is all 1s, the data stream is likely to have been received without error.

A.5.1.4 Cyclic Redundancy Checking

In another process called *cyclic redundancy checking* (CRC), the sender calculates an *n*-bit sequence. When attached to the *k*-bit sequence in the frame, it produces a *k* + *n* bit binary number that is exactly divisible by a given binary prime number called the *generating function*. Known as the *frame check sequence* (FCS), the *n*-bit sequence is placed in the trailer of the frame. Upon receipt, the receiver divides the *k* + *n* bit stream by the generating function used by the sender. If the remainder is zero, the frame has been received without error. Figure A.10 shows the principle of cyclic redundancy checking and lists some representative generating functions. CRC is a powerful technique. It assures the receiver of detecting as few as 1 error in 10^{14} bits.

A.5.2 Error Correction

Once detected, an error must be corrected. Two basic approaches to error correction are:

- *Automatic-repeat-request (ARQ):* Upon request from the receiver, the transmitter resends portions of the exchange in which errors have been detected.
- *Forward error correction (FEC):* Employs special codes that allow the receiver to detect and correct a limited number of errors without referring to the transmitter.

A.5.2.1 ARQ Techniques

Three different procedures can be used to resend the portions of the exchange in which errors are detected.

- *Stop-and-wait:* The sender sends a frame and waits for acknowledgment from the receiver. If no error is detected, the receiver sends a *positive acknowledg-*

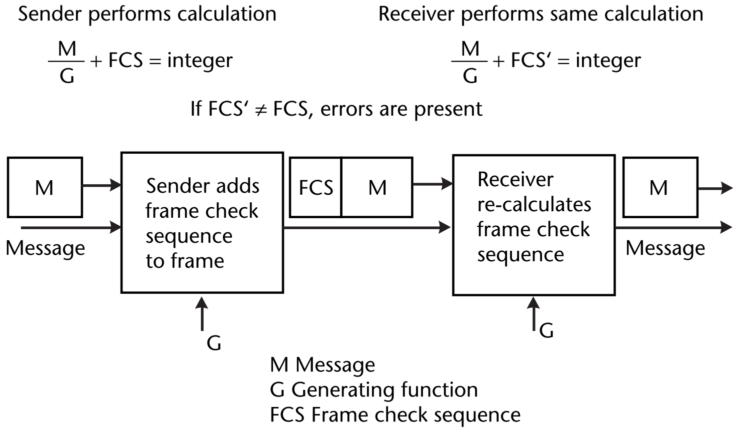

Sender performs calculation
$\frac{M}{G}$ + FCS = integer
Receiver performs same calculation
$\frac{M}{G}$ + FCS′ = integer
If FCS′ ≠ FCS, errors are present
M
Message
Sender adds frame check sequence to frame
G
FCS
M
Receiver re-calculates frame check sequence
G
M
Message
M Message
G Generating function
FCS Frame check sequence

Figure A.10 Principle of cyclic redundancy check.

ment (ACK). The sender responds with the next frame. If an error is detected, the receiver returns a *negative acknowledgment* (NAK). The sender repeats the frame.

- *Go-back-n:* The sender sends a sequence of frames and receives an acknowledgment from the receiver. On detecting an error, the receiver discards the corrupted frame and ignores all further frames in the sequence. The receiver notifies the sender of the number of the frame it expects to receive to replace the first frame discarded. The sender begins resending the sequence starting with that frame.
- *Selective-repeat:* Used on duplex connections only. On the return channel, the receiver returns negative acknowledgments for the individual frames found to have errors. The sender repeats the frames for which NAKs are received.

A.5.2.2 Forward Error Correction

Forward error correction (FEC) requires the sender to add additional coding to segments of the frame. Provided the number of errors is less than a value determined by the coding, the receiver can detect and correct errors without reference to the sender. In one technique (linear block coding), the sender adds check bits to information bits in a known way building on the principle of parity checking. In another technique (convolutional coding), the sender adds bits on the basis of logical operations performed on a moving string of information bits. In general, in an error environment of less than one error in 10,000 information bits (1 in 10^4), ARQ techniques are superior to FEC. In an error environment of more than one in 1,000 (1 in 10^3), FEC must be employed.

Most of the early FEC codes assumed errors were randomly distributed. In many instances, errors occur in bursts. They can be corrected to some extent by interleaving the bits in a frame so that a burst of errors is spread out when the frame is reassembled. In addition, complex block coding (e.g., Reed-Solomon codes) can be used.

APPENDIX B

Frames and Headers

Because there are more details to the frames and headers than it is possible to include in the chapter narratives, I have listed their fields and described their contents in this appendix. Each is entered in the order it is discussed. The entries are divided by chapter. Capitals show the major divisions of each frame (namely, IEEE 802.3 MAC HEADER, IEEE 802.5 TRAILER, and so forth), small capitals are used for field names (namely, SOURCE PORT, DESTINATION PORT, LENGTH, and so forth), and italics are used for subfields (namely, *Precedence*, *Delay*, and so forth).

B.1 Chapter 1: A TCP/IP World?

B.1.1 UDP Header

SOURCE PORT (2 bytes): Number of port in source from which message is sent. Identifies the application layer protocol sending the UDP message. If no reply is expected, the field may be set to 0×00–00.

DESTINATION PORT (2 bytes): Number of port in destination to which message is sent. Identifies the destination application layer protocol receiving the UDP message.

LENGTH (2 bytes): Length in bytes of the UDP Header + Data.

CHECKSUM (2 bytes): Provides integrity check of UDP message. Calculated over UDP Pseudo Header + UDP Header + Payload.

B.1.2 TCP Header

SOURCE PORT (2 bytes): Number of port in source from which message is sent. Identifies the application layer protocol sending the TCP segment.

DESTINATION PORT (2 bytes): Number of port in destination to which message is sent. Indicates the destination application layer protocol receiving the TCP segment.

SEQUENCE NUMBER (4 bytes): Number of outgoing segment's first byte.

ACKNOWLEDGMENT NUMBER (4 bytes): Sequence number of the next frame in the incoming byte stream that the receiver expects to receive. The acknowledgment number provides a positive acknowledgment of all frames in the incoming stream up to, but not including, the frame whose sequence number is the acknowledgement number.

DATA OFFSET (4 bits): Number of 4-byte words in header. Used to indicate where data begins. For the smallest header, the Data Offset field is set to 0x5 meaning the TCP segment data begins with the 20th byte offset from the beginning of the TCP segment. For the maximum TCP header (i.e., with Options and Padding), the Data Offset field is set to 0 × F, meaning the TCP segment data begins with the 60th byte offset from the beginning of the TCP segment.

RESERVED (6 bits): Set to 0. Reserved for future use.

FLAGS (6 bits): Individual bits are designated URG Urgent; ACK Acknowledgment; PSH Push; RST Reset; SYN Synchronize; FIN Finish.

WINDOW (2 bytes): Number of bytes available in the receive buffer of the sender of this segment.

CHECKSUM (2 bytes): Checks TCP segment (TCP Header + Payload). Calculated over TCP pseudo header, TCP header, Payload, and any padding.

URGENT POINTER (2 bytes): Indicates the location of urgent data in the segment.

OPTIONS AND PADDING ($n \times 4$ bytes): Variable size, but must be in 4-byte increments. Used for negotiating maximum segment sizes, scaling window sizes, performing selective acknowledgments, recording timestamps, and providing padding to 4-byte boundaries. The presence of TCP options is indicated by a Data Offset value greater than 5 (i.e., a TCP Header with a size greater than 20 bytes contains options).

B.1.3 IPv4 Header

VERSION (4 bits): Indicates version 4 in use (i.e., 0 × 4)

HEADER LENGTH (4 bits): Length of Header counted in 4-byte blocks. Used to find beginning of payload.

TYPE OF SERVICE (1 byte): Usually set to 0×00. Indicates the quality of service with which the datagram is to be delivered.

Precedence: A 3-bit subfield used to indicate the importance of the datagram;
Delay: A flag set to 0 for normal delay or to 1 for low delay;
Throughput: A flag set to 0 for normal throughput or to 1 for high throughput;
Reliability: A flag set to 0 for normal reliability or to 1 for high reliability;
Cost: A flag set to 0 for normal cost or to 1 for low cost;
Reserved: The last bit is reserved for future use. It is set to 0.

TOTAL LENGTH (2 bytes): Length of the datagram (header + payload) in bytes.

IDENTIFIER (2 bytes): Number that identifies a specific packet sent between a specific source and specific destination

FLAGS (3 bits): Contains flag to indicate whether datagram can be fragmented and another flag to indicate whether more fragments follow.

FRAGMENT OFFSET (13 bits): Indicates where this fragment belongs relative to the original datagram.

TIME TO LIVE (1 byte): Indicates number of links this datagram can travel before it is destroyed. Each node decrements the TTL count by one when forwarding the datagram. Prevents defective datagrams from circulating forever.

PROTOCOL (1 byte): Indicates the upper layer protocol contained within the IP payload. Common values are ICMP, 0×01; IGMP, 0×02; TCP, 0×06; and UDP, 0×11.

HEADER CHECKSUM (2 bytes): Checks IP header only; payload is not included.

SOURCE IP ADDRESS (4 bytes): Contains the IP address of the source host (or Network Address Translator).

DESTINATION ADDRESS (4 bytes): Contains the IP address of the destination host (or Network Address Translator).

OPTIONS AND PADDING (*n*×4 bytes): Options can be added to the IP header. It may have to be padded to bring the length to a multiple of 4 bytes. Some options are:

Record Route: Used to trace a route through an IP internetwork;
Loose Source Routing: Used to route a datagram along a specified path with alternate routes;
Strict Source Routing: Used to route a datagram along a specific path *without* alternate routes;
Internet Timestamp: Used to record a series of timestamps (e.g., time at each hop).

B.1.4 IPv6 Header

VERSION (4 bits): Indicates version 6 in use, (i.e., 0×6).

TRAFFIC CLASS (8 bits): Identifies traffic priority needed to meet QoS objectives.

FLOW LABEL (20 bits): Indicates the length of the remainder of the packet, in bytes.

PAYLOAD LENGTH (2 bytes): Indicates the length of the remainder of the packet, in bytes.

NEXT HEADER (1 byte): Identifies header immediately following this header. Same as protocol field in IPv4. Common values are ICMP, 0×01; IGMP, 0×02; TCP, 0×06; and UDP, 0×11.

HOP LIMIT (8 bits): Number of links to go before packet is discarded.

SOURCE ADDRESS (16 bytes): Unicast address of sending node.

DESTINATION ADDRESS (16 bytes): Address of final destination or NAT.

EXTENSION HEADERS (n×8 bytes): Up to eight extension headers: Hop-by-Hop; Destinations; Routing; Fragment; Authentication; Encapsulating Security Payload; Destination; TCP Header and Data.

B.1.5 ICMP Frame

NETWORK INTERFACE HEADER
IP HEADER

ICMP HEADER

TYPE (1 byte): 0, Echo Reply; 3, Destination Unreachable; 4, Source Quench; 5, Redirect; 8, Echo Request; 9, Router Advertisement; 10, Router Selection; 11, Time Exceeded; 12, Parameter Problem.

CODE (1 byte): Indicates a specific ICMP message within the message type in the type field. If there is only one ICMP message within an ICMP message type, it is set to 0.

CHECKSUM (2 bytes): Checks ICMP header only.

PAYLOAD

TYPE SPECIFIC DATA (n bytes): Variable to accommodate data for each type of message.

NETWORK INTERFACE TRAILER

B.1.6 Echo Request and Reply Messages

TYPE (1 byte): Set to 8 for Echo Request and 0 for Echo Reply.

CODE (1 byte): Set to 0 for both messages. There are no specific ICMP messages within the message type.

CHECKSUM (2 bytes): 16-bit sum that checks ICMP header and ICMP message data.

IDENTIFIER (2 bytes): Number generated by sender used to match Echo Reply with its Echo Request.

SEQUENCE NUMBER (2 bytes): Contains additional number used to match the Echo Reply with its Echo Request.

OPTIONAL DATA (*n* bytes): Variable; explanatory data can be added to the frame.

B.1.7 Destination Unreachable Message

TYPE (1 byte): Set to 3

CODE (1 byte): Some values are: 1, Host unreachable; 2, Protocol unreachable; 4, Fragmentation needed; 5, Source Route failed; 7, Destination Host unknown; 9, Communication with Destination Network administratively prohibited.

CHECKSUM (2 bytes): 16-bit sum that checks ICMP header and message data.

UNUSED (4 bytes): For future use.

DATA (variable): IP header and first 8 bytes of datagram payload.

B.1.8 ARP Request and Reply Messages

HARDWARE TYPE (1 byte): Length in bytes of hardware address in Sender's Hardware Address and Target Hardware Address fields.

PROTOCOL ADDRESS LENGTH (1 byte): Length in bytes of protocol address in Sender's Protocol Address and Target Protocol Address fields.

OPERATION (2 bytes): Indicates type of ARP frame: 1, ARP Request; 2, ARP Reply; 8, Inverse ARP Request; 9, Inverse ARP Reply.

SENDER HARDWARE ADDRESS (6 bytes): Contains hardware address of node sending ARP frame.

SENDER PROTOCOL ADDRESS (6 bytes): For IP, SPA field is 4 bytes. Contains the IP address of the node sending the ARP frame.

TARGET HARDWARE ADDRESS (6 bytes): Set to 0×00–00–00–00–00–00 for ARP Request frames and to hardware address of ARP requester for ARP Reply frames.

TARGET PROTOCOL ADDRESS (6 bytes): For IP, TPA field is 4 bytes. In ARP Request frame it is set to IP address being resolved. In ARP Reply frame it is set to address of IP requester.

B.2 Chapter 3: Local Area Networks

B.2.1 Classic Ethernet Frame

HEADER

PREAMBLE (8 bytes): 0×AA-AA-AA-AA-AA-AA-AA-AB

DESTINATION ADDRESS (6 bytes): If address is unicast, contains the hardware address of a specific station. If address is multicast, carries a code that identifies a group of stations. If address is broadcast, contains code 0×FF-FF-FF-FF-FF-FF.

SOURCE ADDRESS (6 bytes): Unicast address of station where frame originated.

ETHERTYPE (2 bytes): Code indicating upper layer protocol contained in frame. For IP datagram set to 0×08-00; for ARP set to 0×08-06.

PAYLOAD

IP DATAGRAM (46 to 1,500 bytes): Contains Internet layer header, transport layer header, and application PDU.

TRAILER

FRAME CHECK SEQUENCE (4 bytes): Remainder from dividing the data stream between the Preamble and FCS by a 33-bit prime number.

B.2.2 IEEE 802.3 Ethernet Frame

IEEE 802.3 MAC HEADER

PREAMBLE (7 bytes): 0×AA-AA-AA-AA-AA-AA-AA

START DELIMITER (1 byte): 0AB

DESTINATION ADDRESS (2 or 6 bytes): If address is unicast, contains the hardware address of a specific station. If address is multicast, carries a code that identifies a group of stations. If address is broadcast, contains code 0×FF-FF-FF-FF-FF-FF. Bits 1 and 2 of byte 1 are used to identify Universal/Local and Individual/Group addresses.

SOURCE ADDRESS (2 or 6 bytes): Unicast address of station whence frame originated. Bit 1 of byte 1 is used to indicate whether Token Ring MAC-level routing information is present.

LENGTH (2 bytes): Number of bytes from first byte of 802.2 LLC Header to last byte of Payload. Number is 1,500 (0×05-DC). Distinguishes MAC Header from Classic Ethernet header.

IEEE 802.2 LLC HEADER

DESTINATION SAP (1 byte): Identifies point to which payload is delivered. For IP, DSAP = 0×06. Set to 0×AA when combined with SNAP header.

SOURCE SAP (1 byte): Identifies point from which payload originated. For IP, SSAP = 0×06. Set to 0×AA when combined with SNAP header.

CONTROL (1 or 2 bytes): Type 1: If encapsulated data is an IP datagram or ARP message, Control field is 1 byte and is set to 0×03 [Unnumbered Information (UI) frame]. Type 2: If encapsulated data is part of a connection-oriented session, the Control field is 2 bytes. IP datagrams and ARP messages are always sent as Type 1.

IEEE 802.3 SNAP HEADER

ORGANIZATION CODE (3 bytes): Identifies organization that maintains meaning of EtherType field. For IP datagrams and ARP messages, set to 0×00–00–00.

ETHERTYPE (2 bytes): Identifies upper layer protocol in frame. For IP datagrams, value is 0×08–00. For ARP messages, value is 0×08–06.

PAYLOAD

IP DATAGRAM (38 to 1,492 bytes): 8 bytes less than Classic Ethernet because of extra bytes in headers.

IEEE 802.3 TRAILER

FRAME CHECK SEQUENCE (4 bytes): Remainder from dividing the data stream between the Preamble and FCS by a 33-bit prime number.

B.2.3 IEEE 802.5 Token Ring Frame

IEEE 802.5 HEADER

STARTING DELIMITER (1 byte): 0×JK. Contains two nondata symbols called J and K symbols. The J symbol is an encoding violation of 1; the K symbol is an encoding violation of 0. The Starting Delimiter provides a synchronizing signal.

ACCESS CONTROL (1 byte):

Priority bits: 3 bits (7 levels) that establish the priority the receiving station must have in order to seize the token and send a frame.

Token bit: Set to 0, the frame is a token. Set to 1, the frame is in use.

Monitor bit: Set to 1, the frame has passed the monitor station. If it appears a second time at the monitor, the frame is destroyed, and the monitor station generates an empty token.

Reservation bits: 3 bits that record the priority of a station upstream that wants the token. If the station currently handling the frame has something to send and its allocated priority is greater than the level to which the present reservation bits are set, it upgrades the reservation level to equal its allocated priority. The reservation bits become the priority bits when the station that is currently using it releases the token.

FRAME CONTROL (1 byte): 2 bits are reserved for future use.

Frame Type: 2 bits indicating the frame is a Token Ring MAC management frame, or a Token Ring LLC frame.

MAC Management Frame Type: 4 bits indicating the type of MAC management frame.

DESTINATION ADDRESS (6 bytes): The address of the destination station. It may be: a universal or locally administered unicast address; the universal broadcast address 0×FF–FF–FF–FF–FF–FF; the Token Ring broadcast address 0×C0–00–FF–FF–FF–FF; a multicast address; or a Token Ring functional address used by Token Ring MAC management frames. A frame using the Token Ring broadcast address remains on a single ring. Token Ring source-route bridges do not forward it.

SOURCE ADDRESS (6 bytes): Unicast address of station where frame originated.

IEEE 802.2 LLC HEADER

DESTINATION SAP (1 byte): For IP, set to 0×AA.

SOURCE SAP (1 byte): For IP, set to 0×AA.

CONTROL (1 byte): For IP, set to 0×03 [Unnumbered Information (UI) frame].

IEEE 802.3 SNAP HEADER

ORGANIZATION CODE (3 bytes): For IP datagrams and ARP messages, the Organization code is set to 0×00–00–00.

ETHERTYPE (2 bytes): For IP datagrams, value is 0×08–00. For ARP messages, value is 0×08–06.

PAYLOAD

IP DATAGRAM: No minimum size. Maximum size depends on the bit rate and the token holding time. For a token holding time of 10 ms, the maximum sizes for IP datagrams are 4,464 bytes at 4 Mbps and 17,914 bytes for 16 Mbps.

IEEE 802.5 TRAILER

FRAME CHECK SEQUENCE (4 bytes): Remainder from dividing the data stream between the access control byte and FCS by a 33-bit prime number.

ENDING DELIMITER (1 byte): Identifies the end of the frame. Contains J and K nondata symbols. Also contains:

Intermediate frame indicator bit: 1 bit used to indicate whether this is the last frame of a sequence (0), or more frames are to follow (1);

Error detected indicator bit: 1 bit used to indicate whether the frame failed FCS checking. The FCS is checked at each node on the ring. If the FCS fails at any node, the error bit is set to 1. The receiving node does not copy the frame.

FRAME STATUS (1 byte):

Address recognized indicator bit (duplicate copies): 1 bit set by the destination node to indicate that the address was recognized.

Frame copied indicator bit (duplicate copies): 1 bit set by the destination node to indicate the frame was copied successfully. Because they are not checked by FCS, the bits are duplicated.

B.2.4 FDDI Frame

FDDI HEADER

PREAMBLE (2 bytes): Provides receiver synchronization. 0×AA-AA.

STARTING DELIMITER (1 byte): 0×JK. Contains two nondata symbols called J and K symbols. The J symbol is an encoding violation of 1; the K symbol is an encoding violation of 0.

FRAME CONTROL (1 byte):

Class:1 bit denoting synchronous frame (1), or asynchronous frame (0).

Address: 1 bit denoting source and destination addresses are 2 bytes (0), or 6 bytes (1).

Frame Type: 6 bits indicating the type of frame (i.e., token, MAC frame, LLC frame).

DESTINATION ADDRESS (2 or 6 bytes): Indicates the address of the destination station. 2 byte addressing is not used with IP/ARP. For interoperability, made the same as Ethernet destination addresses. Bits 1 and 2 of byte 1 are used to identify universal or local addresses, and individual or group addresses.

SOURCE ADDRESS (2 or 6 bytes): Unicast address of station whence frame originated. 2 byte addressing is not used with IP/ARP. Bit 1 of byte 1 identifies whether Token-Ring MAC level routing information is present.

IEEE 802.2 LLC HEADER

DESTINATION SAP (1 byte): Identifies point to which payload is delivered. For IP, DSAP = 0×06. Set to 0×AA when combined with SNAP.

SOURCE SAP (1 byte): Identifies point from which payload is sent. For IP, SSAP = 0×06. Set to 0×AA when combined with SNAP.

CONTROL (1 byte): For IP, set to 003 [Unnumbered Information (UI) frame].

IEEE 802.3 SNAP HEADER

ORGANIZATION CODE (3 bytes): For IP datagrams and ARP messages, the organization code is set to 0×00–00–00.

ETHERTYPE (2 bytes): For IP datagrams, value is 0×08–00. For ARP messages, value is 0×08–06.

PAYLOAD

IP DATAGRAM (up to 4,352 bytes): No minimum size. Maximum *frame* size from start of Preamble through Frame Status is 4,500 bytes. FDDI header and trailer are 22 bytes. LLC header is 3 bytes. SNAP header is 5 bytes. 117 bytes are reserved for future uses.

FDDI TRAILER

FRAME CHECK SEQUENCE (4 bytes): Remainder from dividing the data stream between the access control byte and FCS by a 33-bit prime number.

ENDING DELIMITER (1 byte): Identifies the end of the frame. Contains J and K nondata symbols. Also contains:

Intermediate frame indicator bit, 1 bit used to indicate whether this is the last frame of a sequence (0), or more frames are to follow (1);

Error detected indicator bit, 1 bit used to indicate whether the frame failed FCS checking. (The FCS is checked at each node on the ring. If the FCS fails at any node, the error bit is set to 1. The receiving node does not copy the frame.)

FRAME STATUS (1 byte):

Address recognized indicator bit (duplicate copies): 2×1 bit set by the destination node to indicate that the address was recognized.

Frame copied indicator bit (duplicate copies): 2×1 bit set by the destination node to indicate the frame was copied successfully. Because they are not checked by FCS, the bits are duplicated.

B.3 Chapter 4: Wide Area Networks

B.3.1 Point-to-Point Protocol (PPP) Frame

HDLC HEADER

FLAG (1 byte): 0×7E

ADDRESS (1 byte): Because the connection is point-to-point, set to 0×FF. May be omitted.

CONTROL (1 byte): Set to 0×30 [i.e., Unnumbered Information (UI) frame with Poll/Final bit set to 0]. May be omitted.

PROTOCOL (2 bytes): For an IP datagram, set to 0×00–21.

PAYLOAD

IP DATAGRAM (1,500 bytes)

HDLC TRAILER

FRAME CHECK SEQUENCE (2 bytes): Remainder from dividing the data stream between the Begin Flag and FCS by a 17-bit prime number.

FLAG (1 byte): 0×7E

B.3.2 X.25 Data Frame

LINK ACCESS PROTOCOL – BALANCED (LAPB) HEADER

FLAG (1 byte): 0×7E

ADDRESS (1 byte): Indicate command or response frame.

CONTROL (1 byte): Provides further information on command and response frames and indicates frame format and function.

PACKET LAYER PROTOCOL (PLP) HEADER

GENERAL FORMAT INDICATOR (4 bits): Identifies the payload as user's data or an X.25 message. Specifies the packet numbering cycle is 7 or 127. Specifies whether delivery confirmation is required.

LOGICAL GROUP/ CHANNEL NUMBER (4 + 8 bits): Identifies virtual circuit over which frame will travel between DTE and DCE.

SEQUENCING (1 or 2 bytes): Provides number of this frame [N(S)], number of frame receiver expects [N(R)], and fragmentation information for user's segments.

PAYLOAD

NETWORK LAYER PROTOCOL IDENTIFIER (NLPID) (1 byte): For an IP datagram set to 0×CC. For a single protocol virtual circuit (e.g., only IP), NLPID is omitted.

IP DATAGRAM (≤ 4,096 bytes)

LAPB TRAILER

FRAME CHECK SEQUENCE (2 bytes); Remainder from dividing the data stream between the Begin Flag and FCS by a 17-bit prime number.

FLAG (1 byte): 0×7E

B.3.3 ATM Cell Structure

HEADER

GENERIC FLOW CONTROL (4 bits): User-node interface (UNI) only. Intended to support local connections. Little used.

VIRTUAL PATH IDENTIFIER (VPI) (UNI 1 byte, NNI 12 bits): Different for UNI and node-network interface (NNI). With VCI points to the location in switch tables that contains the actual route.

VIRTUAL CHANNEL IDENTIFIER (VCI) (2 bytes): With VPI points to the location in switch tables that contains the actual route.

PAYLOAD TYPE IDENTIFIER (PTI) (3 bits): Identifies payload as user payload or network management payload.

CELL LOSS PRIORITY (CPI) (1 bit): Guides cell discard in event of congestion. 1 signifies lower priority cell that should be discarded first. 0 signifies higher priority cell.

HEADER ERROR CONTROL (HEC) (1 byte): CRC computed over cell header.

PAYLOAD

SEGMENT (48 bytes): First 4 bytes may be used for AAL control information.

B.3.4 AAL5 Frame Containing IP Datagram

LLC HEADER: standard

SNAP HEADER: standard

PAYLOAD

IP DATAGRAM (38 to 1,492 bytes)

PAD (≤47 bytes)

AAL5 TRAILER

USER-TO-USER INDICATOR (1 byte): Transfers information between AAL users (not defined).

COMMON PART INDICATOR (1 byte): Aligns the AAL5 trailer on a 64-bit boundary.

LENGTH OF PAYLOAD (2 bytes): Length in bytes of the Payload so receiver can discard Pad.

FRAME CHECK SEQUENCE (4 bytes): Remainder from dividing the data stream formed by payload and trailer by a 33-bit prime number.

B.3.5 Frame Relay Frame with 2-Byte Addresses

FRAME RELAY HEADER

FLAG (1 byte): 0×7E

ADDRESS (2 bytes):

Data link connection identifier (DLCI): The first 6 bits of the first byte and the first 4 bits of the second byte comprise the 10-bit DLCI. It identifies the virtual circuit over which the frame relay (FR) frame is transported. The DLCI is only locally significant. Each FR switch changes the DLCI value as it forwards the FR frame.

Command/Response (C/R): The seventh bit in the first byte of the address field is the C/R bit. It is not used and is set to 0.

Extended address (EA): The last bit in each byte of the address field is the EA bit. If it is set to 1, the current byte is the last byte in the address field. Set to 0, there is at least one more address byte to follow.

Forward explicit congestion notification (FECN): The fifth bit in the second byte of the address field is the FECN bit. It is used to inform the destination node that congestion exists in the path from source to destination. The FECN bit is set to 1 by any FR node in the forward path that is becoming congested. When the destination node receives a frame with FECN set to 1, the information is passed to upper layer protocols that may initiate flow control procedures (receive side).

Backward explicit congestion notification (BECN): The sixth bit in the second byte of the address field is the BECN bit. It is used to inform the destination node that congestion exists in the path from destination to source. The BECN bit is set to 1 by any FR node that is becoming congested in the reverse path. When the destination node receives a frame with BECN set to 1, the information is passed to upper layer protocols that may initiate flow control procedures (send side).

Discard eligibility (DE): The seventh bit in the second byte of the address field is the DE bit. The first FR node sets the DE bit to 1 when the sender exceeds the *committed information rate* (CIR). Frames with DE = 1 are discarded first during periods of congestion.

CONTROL (1 byte): Set to 0×30

PAYLOAD

NETWORK LAYER PROTOCOL IDENTIFIER (1 byte): For an IP datagram set to 0×CC. For a single protocol virtual circuit, NLPID is omitted.

IP DATAGRAM (262 to 1,600 bytes)

FRAME RELAY TRAILER

FRAME CHECK SEQUENCE (2 bytes): Remainder from dividing the datastream between the Begin Flag and FCS by a 17-bit prime number.

FLAG (1 byte): 0×7E

B.4 Chapter 5: Connecting Networks Together

B.4.1 Source Routing Added to Token Ring Frame

IEEE 802.5 HEADER

STARTING DELIMITER: standard

ACCESS CONTROL: standard

FRAME CONTROL: standard

DESTINATION ADDRESS: standard

SOURCE ADDRESS (6 bytes): Bit 1: Set to 1, Source routed.

ROUTING CONTROL (2 bytes):

Routing Type (3 bits): 0xx, specifically routed frame; 11x, Spanning Tree Explorer; 10x, All Routes Explorer.

Length (5 bits): number of bytes in Routing Control and Route Descriptors.

Direction (1 bit): 0, read Route Descriptors left to right; 1, read Route Descriptors right to left.

Largest Frame (6 bits): indicates largest data payload field supported by route.

Reserved: 1 bit.

Route Descriptors (≤28 bytes): Route Descriptor #1 (2 bytes), Ring number (12 bits), Bridge number (4 bits). ... Route Descriptor #14 (2 bytes), Ring number (12 bits), Bridge number (4 bits).

IEEE 802.2 LLC HEADER: standard

PAYLOAD: IP Datagram

IEEE 802.5 TRAILER: standard

B.4.2 Tag for IEEE 802.3 (Ethernet) Frame Encapsulating an IP Datagram

IEEE 802.3 MAC HEADER: standard

IEEE 802.2 LLC HEADER: standard

IEEE 802.3 SNAP HEADER

ORGANIZATION CODE: Standard

ETHERTYPE (2 bytes): 0×81-00

TAG CONTROL INFORMATION FIELD (2 bytes):

Byte 1: bits 0 through 3, VLAN Identifier; bit 4, CFI, canonical format indicator; bits 5, 6, 7, priority information
Byte 2: bits 0 through 7, VLAN Identifier

PAYLOAD

IEEE 802.3 TRAILER: standard

B.4.3 IEEE 802.3 (Ethernet) Frame with Embedded Routing Information

IEEE 802.3 MAC HEADER: standard

IEEE 802.2 LLC HEADER: standard

IEEE 802.3 SNAP HEADER

ORGANIZATION CODE: Standard

ETHERTYPE: Standard

TAG CONTROL INFORMATION FIELD: Standard

ROUTING CONTROL (2 bytes):

Routing Type (3 bits): 00×, specifically routed frame; 01×, transparently bridged frame; 10×, All Routes Explorer; 11x, Spanning Tree Explorer frame.
Length (5 bits): number of bytes in Route Descriptor field.
Direction (1 bit): 0, read Route Descriptors left to right; 1, read Route Descriptors right to left.
Largest Frame (6 bits): indicates largest data payload field supported by route.
Noncanonical Format Indicator (1 bit): 0, Big Endian format; 1, Little Endian format

ROUTE DESCRIPTORS (≤ 28 bytes): Route Descriptor #1 (2 bytes): LAN Identifier (12 bits), Bridge number (4 bits). ... Route Descriptor #14 (2 bytes): LAN Identifier (12 bits), Bridge number (4 bits).

PAYLOAD: IP Datagram

IEEE 802.3 TRAILER: standard

B.5 Chapter 6: Protecting Enterprise Catenets

B.5.1 Authentication Header Fields in Datagrams in Figure 6.6

AUTHENTICATION HEADER

NEXT HEADER (1 byte): Identity of Header following AH. UDP = 0×11; TCP = 0×06.

LENGTH (2 bytes): Length of Authentication Header.

RESERVED (2 bytes): Set to 0×00-00, not allocated.

SECURITY PARAMETERS INDEX (4 bytes): In combination with destination address, identifies Security Association to be used.

SEQUENCE NUMBER (4 bytes): Datagram identifier. Begins at 0 when new Security Association is invoked. Counts by 1s. Prevents repetition of datagram.

AUTHENTICATION DATA (variable): Datagram identifier. Begins at 0 when new SA invoked. Counts by 1s. Prevents repetition of datagram.

B.5.2 Encapsulating Security Header and Trailer

IP HEADER: Protocol field is set to 0×32 to indicate ESP.

ENCAPSULATING SECURITY PAYLOAD (ESP) HEADER

SECURITY PARAMETERS INDEX (4 bytes): In combination with destination address, identifies security association to be used.

AUTHENTICATION DATA (variable): Hash integrity check from ESP header to ESP trailer. All mutable fields are set to 0s, and all immutable fields retain their values. The authentication data field is set to 0 during the calculation.

TCP HEADER: Authenticated, Encrypted.

PAYLOAD: Authenticated, Encrypted.

ESP TRAILER

PADDING (variable): Up to 255 bytes of padding.

PADDING LENGTH (1 byte): Number of bytes in padding field.

NEXT HEADER (1 byte): Identity of next header.

ESP AUTHENTICATION DATA (variable):

B.6 Chapter 7: Transmission Facilities

B.6.1 IEEE 802.11 Frame Containing IEEE 802.3 Payload

IEEE 802.11 HEADER

FRAME CONTROL (2 bytes):

Bits 0 and 1: indicate which version of 802.11 is in use. Set to 00 since only one version exists.

Bits 2 and 3: identify type of frame. Set to 00 for management frames; 01 control frames; 10 data frames.

Bits 4 through 7: identify subtype of frame (e.g., set to 1011 for RTS and 1100 for CTS control frames).

Bit 8: ToDS bit. Set to 1 for data frames transmitted from movable station to AP.

Bit 9: From DS bit. Set to 1 for data frames transmitted from AP to movable station.

Bit 10: More fragments bit. Set to 1 if fragments following. Set to 0 for final segment.

Bit 11: Retry bit. Set to 1 for retransmitted frames.

Bit 12: Power management bit. Set to 1 if movable station will enter power saving mode after this frame.

Bit 13: More data bit. Set to 1 by AP to alert movable station in power saving mode that AP has at least one frame for delivery.

Bit 14: WEP bit. Set to 1 when frame has been encrypted by Wired Equivalent Privacy (WEP) to protect data and authenticate sender.

Bit 15: Order bit. Set to 1 when frames must be delivered in sequence.

DURATION/ID (2 bytes): When bit 15 is set to 0, bits 0 through 14 (NAV) indicate the time (in microseconds) the medium is expected to remain busy for the transmission in progress. When bit 15 is set to 1, and bits 0 through 14 are set to 0, indicates a contention-free period of 32,768 microseconds. When bits 14 and 15 are set to 0, indicates a station has changed from power-saving mode to powered mode.

ADDRESS 1 (6 bytes): 48-bit MAC address of destination (from 802.3 frame).

ADDRESS 2 (6 bytes): 48-bit MAC address of source (from 802.3 frame).

ADDRESS 3 (6 bytes): 48-bit MAC address of AP/BSS hosting movable station.

SEQUENCE CONTROL (2 bytes): Used in reconstructing frames and discarding duplicate frames.

Fragment number: Bits 0 thru 3

Sequence number: Bits 4 thru 15, all fragments of a fragmented frame carry the same sequence number.

ADDRESS 4 (6 bytes): 48-bit MAC address for future use.

PAYLOAD Consists of 802.3 LLC and SNAP header and IP packet.

TRAILER

FRAME CHECK SEQUENCE (4 bytes): Checks entire IEEE 802.11 frame.

List of Acronyms and Abbreviations

4B/5B	4 binary/5 binary
8B/10B	8 binary/10 binary
AAL	ATM adaptation layer
ABM	asynchronous balanced mode
ABR	available bit rate
ACELP	Algebraic-Code-Excited-Linear-Prediction
ACK	acknowledge
ADM	add/drop multiplexer
ADPCM	adaptive differential PCM
ADSL	asymmetrical digital subscriber line
AMI	alternate mark inversion signal format
APDU	application protocol data unit
ARP	Address Resolution Protocol
ARPA	Advanced Research Projects Agency
ARPAnet	ARPA network
ARQ	await receiver request
ASCII	American Standard Code for Information Interchange
ASK	amplitude shift keying
ASP	adjunct service point
ATM	asynchronous transfer mode
B8ZS	bipolar with 8 zeros substitution
BCC	block check character
B-ISDN	broadband ISDN
BISYNC	Binary Synchronous Data Link Control Protocol
BS	bursty second
BSS	basic service set
BT	bridged tap
CA	certificate authority
CAP	carrierless amplitude and phase

CBR	constant bit rate
CDMA	code division multiple access
CELP	Code-Excited-Linear-Prediction
CI	congestion indicator
CIDR	classless interdomain routing
CIR	committed information rate
CLASS	custom local-area signaling services
CLEC	competitive local exchange carrier
CLP	cell loss priority
CLR	cell loss rate
CMR	cell misinsertion rate
CMTS	cable modem termination system
CO	central office
CORE	Council of Registrars
COT	central office terminal
CRC	cyclic redundancy check
CRS	cell relay service
CS	convergence sublayer
CSA	carrier serving area
CSA-CELP	Conjugate-Structure Algebraic-Code-Excited-Linear-Prediction
CSN	current sequence number
CSMA/CA	carrier sense multiple access with collision avoidance
CSMA/CD	carrier sense multiple access with collision detection
CSU	customer service unit; channel service unit
CTS	clear to send
dB	decibel
DCC	digital cross-connect
DCE	data circuit equipment
DCF	distributed coordination function
DHCP	Dynamic Host Configuration Protocol
DiffServ	differentiated services
DIFS	distributed coordination function interframe space
DLCI	data link connection identifier
DLE	data link escape
DNHR	dynamic nonhierarchical routing
DMT	discrete multitone transmission

DNS	domain name system, also domain name server
DS	differentiated services
DS-0	digital signal level 0
DS-n	digital signal level *n*
DSCP	differentiated services code point
DSL	digital subscriber line
DSLAM	digital subscriber line access multiplexer
DSP	digital signal processor
DSU	data service unit
DTE	data terminal equipment
DTMF	dual tone multifrequency
DUN	dial-up network
DWDM	dense wavelength division multiplexing
EBCDIC	extended binary coded decimal interchange code
EC	echo canceler
ECR	explicit cell rate
EDFA	Erbium-doped fiber amplifier
EIR	excess information rate
ENQ	enquiry
EOT	end of transmission
ERI	embedded routing information (Token Ring); explicit routing information (VLAN)
ESC	escape character
ESF	extended superframe
ESP	encapsulating security payload
ETB	end of text block
ETX	end of text
FCS	frame check sequence
FDI	feeder distribution interface
FDDI	fiber distributed data interface
FDM	frequency division multiplexing
FEC	forwarding equivalence class
FEXT	far-end crosstalk
FRAD	frame relay access device
FS	failed seconds
FSK	frequency shift keying
FSN	final sequence number

ft	foot
FTP	File Transfer Protocol
FTTC	fiber to the curb
FTTH	fiber to the home
Gbps	gigabits per second
GFC	generic flow control
gTLD	generic top level domain
H0	384-kbit/s channel
H11	1.536-Mbps channel
HDLC	High-Level Data Link Control Protocol
HDSL	high-bit-rate digital subscriber line
HDSL2	high-bit-rate digital subscriber line 2
HEC	header error control
HTTP	Hypertext Transfer Protocol
IANA	Internet Assigned Numbers Authority
ICANN	Internet Corporation for Assigned Names and Numbers
ICMP	Internet Control Message Protocol
IDU	interface data unit
IETF	Internet Engineering Task Force
I/G	individual/group (bit)
IGMP	Internet Group Management Protocol
IN	intelligent network
InvARP	inverse ARP
IP	Internet Protocol
IPsec	IP Security
IPv4	version 4 of Internet Protocol
IPv6	version 6 of Internet Protocol
IRTF	Internet Research Task Force
ISDN	integrated services digital network
ISM	industrial, scientific, and medical (radio bands)
ISN	initial sequence number
ISO	International Organization for Standardization
ISP	Internet service provider
ITB	end of intermediate text block
ITU	International Telecommunication Union
IXP	Internet exchange point

kbit/s	kilobits per second
km	kilometer
L2TP	Layer 2 Tunneling Protocol
LAN	Local Area Network
LAP-B	Link Access Protocol–Balanced
LAP-D	Link Access Protocol–Channel D
LAP-F	Link Access Procedure–Frame Mode
LCN	logical channel number
LD-CELP	Low-Delay-Code-Excited-Linear Prediction
LDP	Label Distribution Protocol
LEC	local exchange company
LLC	logical link control sublayer
LIB	label information base
LSA	link state advertisement
LSB	least significant bit
LSP	label switched path
MAC	medium access control
MAE	metropolitan area exchange
MAU	multistation access unit
Mbps	megabits per second
MBS	maximum burst size
MCR	minimum cell rate
MDF	main distributing frame
MG	media gateway
MGCP	Media Gateway Control Protocol
MLT-3	multilevel threshold-3 signal format
MN	matched node
MOS	mean opinion score
MPEG	Motion Picture Engineering Group
MPLS	multiprotocol label switching
MRU	maximum receive unit
MSB	most significant bit
MSS	maximum segment size
MTU	maximum transmission unit
NAP	network access point
NAK	negative Acknowledgment

NAV	network availability vector
NBMA	nonbroadcast multiple access
NCP	network control point
NEXT	near-end crosstalk
nm	nanometer
NNI	node–network interface
NRZ	nonreturn to zero
NRZI	nonreturn to zero, invert on ones
OC-1	optical carrier level 1
OC-N	optical carrier level *N*
OFDM	orthogonal frequency division multiplexing
OOF	out of frame (event)
ONU	optical network unit
OPTIS	overlapped pulse amplitude modulation with interlocked space
OSI	open systems interconnection
OSPF	open shortest path first
PAM	pulse amplitude modulation
PCF	point coordination function
PCM	pulse code modulation
PCR	peak cell rate
PDU	protocol data unit
P/F (bit)	poll/final bit
PI	protocol interpreter
PIC	polyolefin-insulated cable
PIFS	point coordination function interframe space
PLCP	physical layer convergence procedure
PLP	Packet Layer Protocol
POTS	plain old telephone service
PPP	Point-to-Point Protocol
PPTP	Point-to-Point Tunneling Protocol
PSK	phase shift keying
PSTN	public switched telephone network
PTI	payload type identifier
QAM	quadrature amplitude modulation
QoS	quality of service
RD	running disparity

REJ	reject
RER	residual error rate
RFC	Request for Comments
RIP	Routing Information Protocol
RM	resource management
RNR	receiver not ready
RR	receiver ready
RSVP	Resource Reservation Protocol
RT	remote terminal
RTCP	Real-Time Control Protocol
RTO	retransmission time out
RTP	Real-Time Transport Protocol
RTS	request to send
RTSP	Real-Time Streaming Protocol
RTT	round-trip time
SA	security association
SACK	selective acknowledgment
SAP	service access point
SAPI	service access point identifier
SAR	segmentation and reassembly
SCP	service control point
SCR	sustainable cell rate
SDH	synchronous digital hierarchy
SEAL	simple and efficient layer
SES	severely errored second
SF	superframe
SIFS	short interframe space
SLIP	Serial Line Internet Protocol
SMTP	Simple Mail Transfer Protocol
SNAP	Subnetwork Access Protocol
SNEXT	self near-end crosstalk
SOH	start of header
SONET	synchronous optical network
SPA	source protocol address
SPE	synchronous payload envelope
SPI	security parameters index

SS7	Signaling System #7
STM-1	synchronous transport module level 1
STM-N	synchronous transport module level *N*
STP	Spanning Tree Protocol; signal transfer point
STS-1	synchronous transport signal level 1
STS-N	synchronous transport signal level *N*
STX	start of text
SYN	synchronize
TCIF	tag control information field
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TEI	terminal endpoint identifier
TPA	target protocol address
TTL	time to live
UBR	unspecified bit rate
UDP	User Datagram Protocol
U/L	universal/local (bit)
UNI	user–network interface
URG	urgent (data)
URL	uniform resource locator
URN	uniform resource name
UTP	unshielded twisted pair
VBR	variable bit rate
VC	virtual circuit
VCI	virtual circuit identifier
VDSL	very-high bit-rate digital subscriber line
VLAN	virtual local area network
VoIP	voice over IP
VP	virtual path
VPI	virtual path identifier
VPN	virtual private network
WAN	wide area network
WDM	wavelength division multiplexing
WEP	wired equivalent privacy
www	World Wide Web
ZBTSI	zero-byte time slot interchange

Glossary

2B1Q Two binary, one quaternary; coding developed for the ISDN basic rate signal.

64-kbit/s clear channel A full 64-kbit/s channel that is available to the customer. This is achieved by introducing special coding that replaces all-0 bytes.

AAL1 Specialized ATM adaptation layer. Provides a connection-oriented, constant bit-rate voice service. Performs segmentation and reassembly, may detect lost or errored information, and recovers from simple errors.

AAL2 Specialized ATM adaptation layer. Provides a connection-oriented variable bit-rate video service. Performs segmentation and reassembly and detection and recovery from cell loss or wrong delivery.

AAL3/4 Specialized ATM adaptation layer. Supports connection-oriented and connectionless, variable bit-rate data services.

AAL5 Specialized ATM adaptation layer. Supports connection-oriented, variable bit-rate, bursty data services on a best-effort basis. Performs error detection, but does not pursue error recovery. Also known as the *simple and efficient layer* (SEAL).

Access point In IEEE 802.11, a fixed station that provides radio links to movable data terminals and hosts a basic service set.

Acknowledged connectionless service Message-handling feature of IEEE 802.3 LAN. The receiver acknowledges messages, but a logical connection is not established.

Acknowledge—Reset message Sent by receiver of TCP message when it cannot establish a connection. The ACK and RST flags are set.

Active OPEN function call Opens a port from the application layer to the transport layer.

Adaptive differential PCM A voice-encoding technique. Encodes voice to 32 kbit/s with a *mean opinion score* (MOS) of 4.0 and processing delay of 0.125 ms.

Add/drop multiplexer Aggregates or splits SONET traffic at various speeds so as to provide access to SONET channels without demultiplexing the signal stream.

Address mask See *subnet mask*.

Address Resolution Protocol In IPv4, used to resolve the IP address of a node and its hardware (MAC) address.

Adjunct service point In intelligent network, a unit that implements *custom local-area signaling services* (CLASS) features.

Aggregatable global unicast address In IPv6, address organized in three sections. Section 1 consists of address space managed by entities that provide public Internet services. Section 2 identifies an organization's internal routing paths. Section 3 identifies individual interfaces on the organization's physical links.

Algebraic-Code-Excited-Linear Prediction A low bit-rate voice-encoding technique. Encodes voice to 6.3 kbit/s with an MOS of 3.8 and processing delay of 37.5 ms.

Alternate mark inversion signal format 1s are represented by return-to-zero current pulses that alternate between positive and negative. 0s are represented by the absence of current pulses.

American Standard Code for Information Interchange Composed of 128 7-bit words that represent the alphabet, numbers, punctuation marks, and control symbols.

Amplitude modulation The amplitude of the carrier is varied based on the value of the modulating signal.

Amplitude-shift keying Digital modulating technique in which the carrier signal has two amplitude values.

Analog signal A continuous signal that assumes positive, zero, or negative values. Changes occur smoothly and rates of change are finite.

Application layer (1) Layer 7 in the OSI model; (2) Layer 4 in the Internet model. Invokes generic applications (e.g., mail, file transfer, terminal emulation) in support of data generated by specific user applications. Interfaces user processes with lower-level protocols.

Application-level filtering In VLAN, by testing the data contained in several frames with the characteristics of the application and the features of the destination, the filter determines whether to forward or destroy data frames.

Advanced Research Projects Agency An agency of the U.S. Department of Defense responsible for development of ARPAnet. Now called DARPA (Defense Advanced Research Projects Agency).

ARPAnet A U.S. government pioneering data communication network that was the forerunner of the Internet.

Asymmetrical digital subscriber line Provides unequal data rates in downstream and upstream directions. In addition, the lowest portion of the bandwidth is used for analog voice.

Asynchronous balanced mode The stations have equal status. Each station may initialize, supervise, recover from errors, and send frames at any time.

Asynchronous operation Not synchronous operation. The nodes operate with similar clocks, but their actions are not synchronized or coordinated. Actions are performed when nodes are ready without reference to the activities of other units. To alert the receiver that data is being transmitted and to synchronize the receiver with the bit stream, asynchronous operation requires the use of start and stop bits, preambles, flags, or other markers.

Asynchronous transfer mode (ATM) A packet-switching technology that uses 53-byte fixed-length cells to implement cell relay service.

ATM adaptation layer When sending, AAL converts messages into sequences of cells for use by the ATM layer. When receiving, AAL converts sequences of cells to messages for use by upper layers. Consists of the convergence sublayer and the segmentation and reassembly sublayer.

ATM layer Adds (deletes) a 5-byte header to 48 (from 53) byte cells. Multiplexes and demultiplexes cells to message streams identified by virtual channel identifiers and virtual path identifiers.

Authentication In IPsec, provides the receiver with the ability to check that the immutable fields in the received frame are identical to those in the frame that was sent.

Authentication header In IPv4, authentication information is carried in an authentication header inserted between the Internet layer header and the transport layer header in the IP datagram. In IPv6, the IP datagram consists of a base header, extension headers, transport layer header, and message. The authentication header is one of the extension headers.

Autonomous network In the Internet, an individual network operated by a single authority responsible for defining operating discipline.

Available bit rate service In ATM, to transfer cells as quickly as possible, the sender may try to use all of the bandwidth that is not allocated to other traffic. To do so without loss of data, the source must adjust its sending bit rate to match conditions as they fluctuate within the network. Resource management cells provide feedback for these changes.

Backbone network In an intranet, interconnects campus networks. The connection may be distributed or collapsed.

Background noise See *circuit noise*.

Backoff (time) In Ethernet, on ceasing to send, stations that have experienced a collision *backoff* for a random number of slot times before trying to send again.

Bandwidth A range of frequencies that just encompasses all of the energy present in a given signal. Digital signals $\leq$ 1.544 Mbps are referred to as narrow band(width), 1.544 Mbps < Mbps rate $\leq$ 44.736 Mbps are referred to as wide band(width), > 44.736 Mbps are referred to as broadband.

Baseband signal A message signal whose energy occupies a frequency band that may include or be contiguous with 0 Hz (i.e., dc level). The energy of a baseband signal occupies a fixed, unchanging position in the frequency domain.

Basic service set In wireless Ethernet, a grouping of movable terminals homing on a single access point.

Baud A signaling rate of 1 symbol per second.

Big Endian format In each code word, the *least significant bit* (LSB) is on the right, and the *most significant bit* (MSB) is on the left. Bits are read in descending order from the MSB to the LSB. Bytes are numbered left to right, from 0 to *N*, and are read in ascending order. See *little Endian format*.

Binary search A technique for finding routing instructions in a large table. With the routing table sorted in numerical address order, the address for which routing instructions are to be found is compared to the address at the center of the table. If it is larger than the center value, the address must be in the bottom half of the table. If it is less than the center value, the address must be in the upper half of the table. The search proceeds to the center of the half in which the address is located. If the address is less than the new center value, it must be in the upper half of that half of the table. If the address is more than the new center value, it must be in the lower half of that half of the table. The search then divides the quarter in which the address is located into halves and repeats the procedure.

Binary signal A digital signal that has two values.

Binary Synchronous Data Link Control Protocol A Layer 2 protocol that uses control codes.

Binit An alternative name for bit. Used when it is necessary to distinguish between a logical bit (binit) and a symbol imbued with 1 bit of self-information.

Biphase signal format See *Manchester signal format*.

Bipolar with 8 zeros substitution Special coding that eliminates all-0 bytes to make the entire 64-kbit/s channel available to the customer.

Bit A contraction of binary digit. A two-valued symbol usually assigned the values 0 and 1.

Bit stuffing In asynchronous operations that employ flags (0×7E), bit stuffing is used to break up strings of 1s into segments of five 1s. Without regard to byte boundaries, 0 is stuffed after a sequence of five 1s. In this way, only the beginning and ending flags contain six consecutive 1s. The stuffed bits are removed by the receiver.

Block check character A character formed from parity bits created by *longitudinal redundancy check* (LRC) process. In LRC, parity bits are assigned to sequences formed by selecting bits in specific positions in a data block.

Blocking Setting up another signal path is not possible because an existing signal path blocks it.

Bridge (1) A matching device for interfaces that differ in the physical and/or data link layers. (2) A device that connects networks. It forwards messages between them based on a hardware address and a table of corresponding port numbers for the bridge. When implemented mostly in hardware, it is called a Layer 2 switch.

Bridged tap A cable pair continued beyond the point at which the pair is connected to a subscriber. An unterminated cable pair attached to an active cable pair.

Broadband-ISDN Broadband, multimedia, digital network. Provides interactive services up to 150 Mbps and distributed services up to 600 Mbps.

Broadcast address A terminating address (all 1s) for a frame that is processed by every station on the same segment of the network. The frame is not transferred by bridges and routers because the time-to-live field is set to 001.

Broadcast link A link with two or more nodes on the same network segment. Unicast, multicast, and broadcast packets sent by any one of these nodes are received by all nodes on the segment.

Browser Software that acquires pages from the World Wide Web. Translates digital streams into text and pictures for display on PCs.

Bursty second A second in which from 2 to 319 extended superframe (ESF) error events are present.

Byte A data word of 8 bits. See *octet* and *nibble*.

Campus network Interconnects workgroup networks within a single location. Consists of two or more workgroup networks bridged together. Facilitates communication to other campus networks.

Canonical format See *little Endian format*.

Canonical format indicator Bit 5 of first byte of the tag control information field. Indicates whether big Endian or little Endian format is employed.

Carrierless amplitude and phase modulation A passband technology based on *quadrature amplitude modulation* (QAM). With a 256-point constellation (i.e., 8 bits per symbol) and a signaling rate of 1,088 kbaud, bit rates of 8,704 kbit/s are achieved.

Carrier sense multiple access with collision avoidance In IEEE 802.11, frames employ stop-and-wait *await receiver request* (ARQ). Before transmitting data, the sender sends a *request-to-send* (RTS) control frame to the receiver and waits for the receiver to reply with *clear to send* (CTS). As soon as the other movable stations in the *basic service set* (BSS) hear the beginning of this exchange, they may not transmit. When the sender receives the CTS signal, it waits a short time then commences sending data. At the beginning of this action, all other stations in the BSS received a *network availability vector* (NAV) time. They cannot transmit until it expires.

Carrier sense multiple access with collision detection When activity on the common channel ceases, the station with a frame to send waits for a time equal to the Ethernet Interframe gap in case the frame just sent is one of a series. The station then waits a further time period that is a random multiple of the slot time. If there is still no activity, it may send the frame. Once any station has begun transmission, other stations should detect the activity and withhold their own messages. If two, or more, stations begin to transmit at the same time, a collision occurs. They will detect that they are interfering with each other, will jam one another for a short time, so that all stations can hear that a collision has occurred, will cease transmitting, and back off a random number of slot times. At the end of the backoff time, they will try again.

Carrier serving area In the local loop, an area limited to 9,000 feet from a remote terminal (carrier termination) in which twisted pairs are used

Catenet Several individual networks linked together to facilitate the execution of distributed data operations. An aggregate of networks that behaves like a single logical network

Cell In ATM, consists of 48 bytes of payload and 5 bytes of header information.

Cell relay service Transports voice, video, and data messages in streams of short, fixed-length cells.

Centralized routing A primary (and perhaps an alternate) path is (are) dedicated to a pair of stations at the time of need.

Central office A facility that contains the lowest node in the hierarchy that forms the network; used loosely to describe any facility at which significant switching or routing occurs.

Certificate authority A trusted third-party organization or company that issues digital keys (certificates) used to create digital signatures and public-private cryptographic keys.

Central office terminal Terminates line carrier equipment in telephone central office.

Channel A unidirectional communication path.

Channel service unit Part of the data circuit-terminating equipment (DCE) serving the digital line.

Character stuffing In the payload, to prevent misinterpretation of text, addition of a specific character ahead of a text character that mimics a command. For an IP datagram on an asynchronous link, a character that mimics the flag character (0×7E) is replaced by the sequence 0×7D–5E. 0×7D is the ESC character. At the receiving node, 0×7D–5E is replaced by 0×7E.

Checksum An error-detecting procedure. The sender treats the bytes in the datagram as numbers and adds them together to create a total number. The 1s-complement of the total is sent to the receiver. At the receiver, the bytes are summed with the transmitted 1s-complement. If the result is all-1s, it is likely that transmission was free of errors.

Circuit A bidirectional communication path. Can be considered to be two channels operating simultaneously (one in each direction).

Circuit-level filtering With respect to the actions of a proxy server, by observing the grouping of frames, a connection between client and server is detected. Using rules to determine whether the source and destination are compatible (i.e., are likely to have legitimate business to transact), the passage of information is permitted or denied.

Circuit noise On a twisted pair, using a band-limited weighting filter, the power measured between tip and ring when no message signal is present; also known as metallic, background, or differential noise.

Classic Ethernet Original Ethernet LAN. Consists of a common coaxial cable bus operating at 10 Mbps to which all stations are connected. Operation is half-duplex. Only one station can transmit at a time, and when transmitting, it cannot receive.

Class A address An IPv4 address that consists of an 8-bit network ID beginning with 0 and a 24-bit host ID. Identifies 126 networks, each of which can support 16,777,214 hosts. Because they have an 8-bit ID, networks with Class A IDs are called slash eights (written /8s) or eights.

Class B address An IPv4 address that consists of a 16-bit network ID beginning with 10 and a 16-bit host ID. Identifies 16,382 networks, each of which can support 65,534 hosts. Networks with Class B IDs are called slash sixteens (written /16s) or sixteens.

Class C address An IPv4 address that consists of a 24-bit network ID beginning with 110 and an 8-bit host ID. Identifies 2,097,150 networks, each of which can support 254 hosts. Networks with Class C IDs are called slash twenty-fours (written /24s) or twenty-fours.

Class D address An IPv4 address beginning with 1110. The remaining 28 bits ranging from 224.0.0.0 to 239.255.255.255 are used for individual IP multicast addresses.

Classful IP addresses In IPv4, three unicast address classes are defined for public use. See *Class A, Class B, and Class C addresses*.

Classless address See *classless interdomain routing*.

Classless interdomain routing A technique that expresses a group of Class C addresses as a single routing address. As long as the CIDR block is accompanied by the appropriate mask, the network ID for the CIDR block can be any number of bits.

Clear to send In IEEE 802.11, frames employ stop-and-wait ARQ. Before transmitting data, the sender sends a *request-to-send* (RTS) control frame to the receiver and waits for the receiver to reply with *clear to send* (CTS).

Client A terminal with significant computing and processing capability. Acquires data from a server and accomplishes its tasks without outside support. Provides an interface for users' instructions and graphical or textual outputs.

Code division multiple access Direct-sequence spread spectrum modulation technique in which all stations in the network transmit on the same carrier and use the same chip rate to spread the signal spectrum over a wide frequency range. Each station employs a code that is orthogonal to the codes used by others. Each receiver sees the sum of the spread spectrum signals as uncorrelated noise. It can demodulate a specific signal if it has knowledge of the spreading code and the carrier frequency.

Code point First 6 bits in type of service field of IP header. The 64 code points are mapped to a few service definitions.

Collapsed backbone A single core switch or router that interconnects all edge switches or routers in the enterprise catenet. Can provide very large aggregate bandwidth.

Collision domain With respect to Ethernet, a combination of hub/repeater and attached stations.

Command frame Requires a response from receiver.

Committed information rate In frame relay, the average rate, in bits per second, at which the network agrees to transfer data.

Common mode signals On a twisted pair, signals that occur between the two wires and ground. Also known as longitudinal signals. Common mode signals are created by outside interference (power influence and other noise).

Communication Activity associated with distributing or exchanging information.

Communication Protocol A procedure governing communication.

Communication quality voice Voice quality acceptable to professional communicators. It has an MOS of 3.5 to 4.0.

Conjugate-Structure Algebraic-Code-Excited-Linear Prediction A low bit-rate voice-encoding technique. Encodes voice to 8 kbit/s with an MOS of 4.0 and processing delay of 15 ms.

Connectionless service Commonly provided over packet networks for short data messages. Carrying originating and terminating addresses, they are sent without negotiating a connection, carry no sequence numbers, and their receipt goes unacknowledged. Messages sent in sequence are unlikely to follow the same path so that the times they take to reach the destination will vary, and they may arrive out of sequence.

Connection-oriented service A logical connection is set up between originating and terminating stations. Acknowledgments, error and flow controls, and other features are employed to ensure reliable data transfer. The delay between packets may vary, but they will arrive in sequence.

Constellation A polar plot of the combinations of amplitude and phase used to form symbols in a complex modulated wave.

Container Part of the payload in *synchronous digital hierarchy* (SDH). One or more tributary signals are carried in each container.

Content-addressable memory A silicon-intensive database-searching device that employs the content (hardware address of destination) as the key for retrieving associated data (port to which destination is attached).

Convergence sublayer (CS) Part of ATM adaptation layer. When sending (i.e., receiving a PDU from the Internet layer), the CS constructs a CS PDU that consists of the payload, a pad to maintain 48-byte alignment, and a trailer. When receiving, accepts CS PDU from SAR, strips off the trailer, reconstructs PDU received from the sending Internet layer, confirms error-free reception, and delivers PDU to the receiving Internet layer. If the reception is not error-free, the CS discards the CS PDU and notifies the Internet layer.

Core switch VLAN-aware switch. Selects paths based on the tag carried by each frame. Knowing the VLAN to which the frame belongs from the ID carried in the tag, the tag-aware switch makes its forwarding decision.

CRC-6 error event In a T-1 ESF operation, the condition when the *frame check sequence* (FCS) calculated by the receiver does not agree with the FCS delivered with the frame

Crosstalk Interfering signal induced in nearby pairs by signals on an unbalanced tip and ring. May be divided into near-end and far-end crosstalk. See *self-crosstalk* and *foreign crosstalk*.

Cumulative acknowledgment procedure The number in the TCP header acknowledgment field is the number of the first byte of the frame the receiver next expects to receive. Its presence explicitly acknowledges error-free receipt of all bytes up to, but not including, this byte.

Current sequence number See *initial sequence number*.

Cut-through In switching, as soon as the destination address is received in the input buffer, it is compared to the entries in the port-forwarding table. If a path through the switch to the designated port is available, the frame is fed to it. Should the port be busy with other traffic, the frame is stored in the input buffer to wait for the interfering traffic to clear.

Data circuit-terminating equipment A device that assists the *data terminal equipment* (DTE) to send or receive data messages over data circuits. DCEs condition (i.e., prepare) signals received from DTEs for transmission over communication connections, and restore signals received from the network so as to be compatible with receiving DTEs.

Data communication The act of sharing data among devices. The act of transferring data among data processing machines over communication links under the control of communication protocol(s).

Datagram A protocol data unit that is routed across a packet network by decisions made at each node (distributed routing) without establishing a connection or a call record (see *IP datagram*).

Data link connection identifier A combination of *terminal endpoint identifier* (TEI) and *service access point identifier* (SAPI) that identifies a unique logical connection to a Layer 3 protocol in a specific receiving device. A given SAP is connected by a single DLCI to the sending machine.

Data link layer Level 2 in OSI model. Delivers frames over each link in the communication path.

Data link sublayer Part of the network interface layer in the Internet protocol stack. Hardware addresses are discovered, conditions for access to the transport medium are accommodated, and a header and trailer are constructed. When added to the IP datagram, they form the IP frame.

Data service unit Part of *data circuit-terminating equipment* (DCE) serving a digital line.

Data terminal equipment A device that creates, sends, receives, and interprets data messages (i.e., the part that performs terminal, client, host, server, router, or switch functions).

Datum (pl. data) A value given or stipulated.

DCF interframe space (DIFS) See *distributed coordination function interframe space*.

Decapsulation When ascending the protocol stack of the receiving system, at each layer, a header and, in the case of the data link layer, a header and trailer are stripped from the frame. The procedure is known as *decapsulation*, and the user data is said to be decapsulated. See *encapsulation*.

Decryption The reverse of encryption. Through the application of one or more rules based on those employed to encrypt a packet, the restoration of an encrypted frame to its original meaning. See *encryption*.

Demodulation Action of converting a modulated signal to a baseband signal.

Desktop network Several interconnected clients, servers, and printer stations. Consists of individual stations connected by a local area network that employs a common bus or a repeatered or switched hub. A desktop network is the lowest level of the enterprise catenet hierarchy.

Deterministic signal At every instant, a deterministic signal exhibits a value that is related to values at neighboring times in a way that can be expressed exactly.

Differential mode signals Signals applied between the wires of a twisted pair. Also known as metallic signals. Message signals are always transmitted as differential signals.

Differential noise See *circuit noise*.

Differentiated Services Also called *DiffServ*. Technique that makes use of type of service field in IP header to offer limited number of services to IP frames in accordance with instructions from the sender.

Digital cross connect Redistributes (and adds or drops) individual SONET channels among several STS-N links. Consolidates and segregates STS-1s, and can be used to separate high-speed traffic from low-speed traffic.

Digital line carrier Used to link serving offices with *carrier serving area* (CSA) interface points.

Digital signal A signal that assumes a limited set of positive, zero, or negative values. Changes of value are instantaneous, and the rate of change at that instant is infinite. At all other times it is zero.

Digital signal level *n* When $n = 0$, rate = 64 kbit/s; when $n = 1$, rate = 1.544 Mbps; when $n = 2$, rate = 6.312 Mbps; when $n =3$, rate = 44.736 Mbps; when $n = 4$, rate = 274.176 Mbps.

Digital subscriber line (1) High bit-rate DSL, 1.544 Mbps; symmetrical channels; employs two pairs (dual-duplex); without repeater operates to 12,000 feet, with one repeater (doubler) operates to 24,000 feet; with two repeaters operates to 36,000 feet; uses 2B1Q line code. (2) Single-pair high-data-rate DSL (G.shdsl). Up to 2.32 Mbps; symmetrical channels; employs one pair; operates up to 24,000 feet without repeater. (3) Asymmetric DSL. Up to 8 Mbps downstream and up to 640 kbit/s upstream, employs one pair; operates to 12,000 feet without repeater.

Digital subscriber line access multiplexer Multiplexes high-speed DSL circuits for transport to a regional high-speed network that provides access to content providers and the Internet.

Direct delivery The destination IP address carries the same network ID as the router so that the packet is delivered directly to a station on the network.

Discrete multitone transmission (DMT) A passband technology, DMT operates over a range of frequencies. In one implementation, the available frequency band is divided into channels (4.3125 kHz wide). Known as bins, they employ QAM with a 4 kbaud symbol rate and up to 15 bits per symbol.

Distributed backbone A (wide area) network (e.g., frame relay or ATM network) that interconnects campus network edge switches to create an enterprise catenet. Provides moderate to high bandwidth over moderate to long distances.

Distributed coordination function interframe space In IEEE 802.11, the minimum idle time for contention-based services. If the channel has been idle for DIFS or longer, stations may have access to it subject only to random backoff.

Distributed routing On the basis of information about traffic conditions and equipment status (network map, port status), each node decides which path a frame shall take to its destination.

Distribution cables In the local loop, smaller cables (distribution cables) made up of bundles of twisted wire pairs extend the dedicated connections from feeder distribution interfaces to pedestals or cabinets close to individual service users.

Domain name server Provides IP addresses given host names and host names given IP addresses.

Domain name system A process that maps host names and IP addresses. It resolves names into numbers and numbers into names.

Dotted decimal notation 32-bit IPv4 addresses are divided into 4 bytes. They are written as four decimal numbers separated by dots.

Downstream Direction from the CO (central office) to the subscriber.

Drop-and-repeat node SONET devices configured to split SONET traffic and copy (repeat) individual channels on two or more output links. Applications include the distribution of residential video and alternate routing.

Drop wire In the local loop, makes the final connection to the customer.

Dual-duplex connection A connection with two duplex circuits on which signals are divided by frequency. The composite provides twice the bandwidth of a single circuit.

Dual tone multifrequency signaling A combination of audible frequencies used in the local loop to signal called number and other information.

Duplex connection Supports interactive communications. Messages can flow in two directions at the same time.

Duration/ID field In IEEE 802.11, gives the time in microseconds the originator expects to occupy the radio channel to complete transmission.

Dynamic Host Configuration Protocol A client-server protocol that manages client IP configurations and the assignment of IP configuration data.

Dynamic nonhierarchical routing tandem In the telephone long-distance network, a switch so connected that it can complete calls between toll offices by itself. The first-attempt calling path includes a single, DNHR tandem switch.

Dynamic routing Routing protocols are used to update routing tables. When a route becomes unreachable, it is removed from the routing table. When a router becomes unreachable, alternate routes are worked out and shared between routers. In a dynamic routing environment, routers are in regular touch with each other concerning the state and capabilities of the network.

E-1 First level in European digital hierarchy. A rate of 2.048 Mbps.

Edge switch A VLAN-aware switch that filters received frames individually and determines whether to forward them. If the frame is forwarded, the switch uses rules to find the VLAN for which it is intended and determines which of its ports connect with the LANs needed to transmit the frame to the VLAN members. In addition, it decides whether the frame will be sent in tagged or untagged format.

Egress process In VLANs, the final process used by edge and core switches to process frames. Uses egress rules and egress filter to determine whether, and in what format (tagged or untagged), to transmit the frames.

Embedded routing information In IEEE 802.3 Ethernet, a 2-byte routing control field followed by up to 14 route descriptors.

Encapsulating bridge Connects dissimilar LANs at the data link sublayer by encapsulating the original frame with a header and trailer that is understood by the intermediate nodes.

Encapsulating security payload An IPsec field used when authentication and privacy are required. ESP has three sections: a header that is positioned between the Internet header and the transport header, a trailer that follows the message, and an authentication field following the trailer.

Encapsulation When descending the protocol stack, a header and, in the case of the data link layer, a header and trailer, are attached by each layer to form a frame. The procedure is known as *encapsulation*, and the headers and trailer are said to encapsulate the user data. See *decapsulation*.

Encryption Through the application of one or more rules, the action of making readable (clear-text) data frames into not-readable (cipher-text) data frames. The rules for encryption are chosen so that the application of the same rules, or a set of rules based on them, will restore the not-readable frame to readability. See *decryption*.

END character Special character (0×C0) used by Serial Line Internet Protocol (SLIP) placed at the beginning and ending of each IP datagram.

Envelope In SONET, a synchronous payload envelope is generated 8,000 times a second. It contains $n \times 774$ bytes (where $n = 1, 3, 12, 24, 48, 96, \ldots, 792$).

Errored second event A second in which one or more ESF error conditions are present.

Escape character A character (0×7D) used to change the meaning of the following character.

ESF controller A device that maintains surveillance on a group of T-1 links and interrogates the CSUs on a routine basis. Depending on circumstances, the controller will report emergencies and prepare operating reports.

ESF error event An out of frame (OOF) event, or a 6-bit Cyclic Redundancy Check (CRC-6) error event, or both, has (have) occurred.

Ethernet A local area network defined by the IEEE 802.3 committee. Improved on speed and versatility of Classic Ethernet.

Ethernet header Contains a preamble, destination address, source address, and EtherType field.

Ethernet interframe gap The end of an Ethernet frame is not marked explicitly. Instead, a gap (96 bit times) is left between Ethernet frames.

Ethernet trailer Consists of a 4-byte *frame check sequence* (FCS) generated by the source.

Excess information rate In frame relay, the rate at which bits are sent minus the committed information rate.

Explicit tag A tag created by applying VLAN association rules to frame data. Explicit tags are created by VLAN-aware stations or by the first VLAN-aware switch. They must be removed before passing the frame to a tag-unaware device.

Extended binary coded decimal interchange code Composed of 256 8-bit patterns that represent the alphabet, numbers, punctuation marks, and control symbols.

Extended superframe (ESF) A block of 24 T-1 frames in which the framing bits are used to provide synchronization, error checking, and other functions.

Extranet VPN An Intranet VPN used by customers, suppliers, and vendors. Tunnels are established over Internet connections to a secure enterprise server.

Failed Seconds state In ESF, 10 consecutive *severely errored seconds* (SESs) have occurred. The state remains active until the facility transmits 10 consecutive seconds without an SES.

Far-end crosstalk A condition in which a signal transmitted over a twisted pair in a paired cable creates a disturbance in other twisted pairs at the far end of the cable.

Fast Ethernet Ethernet LANs that operate at 100/125 and 1,000/1,250 Mbps. They employ block coding.

Feeder cables In the local loop, bundles of twisted wire pairs contained in feeder cables connect the main distributing frame in the *central office* (CO) to feeder distribution interfaces.

Feeder distribution interface In the local loop, the interface between the feeder cable and distribution cables.

Fiber distributed data interface A local area network that employs a fiber ring. Can include a dual-fiber ring so that the system can recover from a single catastrophic fault. Uses block coding (4B/5B). The signaling rate is 125 Mbaud. Provides connectionless delivery using 48-bit addressing and token passing similar to Token Ring.

File Transfer Protocol Used to share and transfer files between computers, and use other computers for remote storage.

Final sequence number See *initial sequence number*.

Finish–Acknowledge message TCP message sent to terminate one side of an exchange. The ACK and FIN flags are set.

Firewall A software/hardware device that denies unauthorized callers access to a private network and controls calls from the private network to destinations reached over the public network.

Flow control A procedure for controlling the rate of transfer of packets between sender and receiver so that packets are not lost due to congestion at critical points along the path or overwhelm the receiver.

Foreign crosstalk Crosstalk from a different type of data signal running in the same binder. May be divided into near end and far end.

Format prefix In IPv6 address, a variable length field that identifies the type of address.

Forwarding equivalence class In MPLS, frames bearing the same label are known as a *forwarding equivalence class* (FEC). They follow the path established by the first frame.

Forwarding IP address For indirect deliveries, the IP address of a directly reachable router to which the IP datagram is being forwarded to facilitate eventual delivery to the destination IP address. The IP address to which the IP datagram is to be forwarded on its next hop.

Frame check sequence The result of performing a cyclic redundancy check on part or all of a frame. Usually, placed in the trailer.

Frame control field In IEEE 802.11, a 2-byte field that provides the version number and identifies the frame that follows as management, control, or data. Other bits perform specific alerting functions.

Frame filtering With respect to the actions of a proxy server, after checking the address fields and contents of the frame for keywords, passage of the frame to its destination is permitted or denied.

Frame relay A connection-oriented, data link layer packet-switching technology that transfers variable length frames (262 to 8,189 bytes).

Frequency division multiplexing Several signals operating at different frequencies are combined for transmission on a single bearer.

Frequency modulation The frequency of the carrier is varied based on the value of the modulating signal.

Frequency-shift keying A digital modulating technique in which the carrier signal is shifted between two frequencies.

Full-duplex connection Supports interactive communications. Messages can flow in two directions at the same time. The term *full-duplex* is used to distinguish a full-time, two-way circuit from a half-duplex connection.

Gateway A matching device for interfaces that differ above the network layer.

Generic Flow Control A field in ATM UNI (User-Network Interface) cell.

G.lite A scaled-down version of ADSL that does not require splitters to separate voice from data. Standardized by ANSI, a best-effort transmission system.

Global/local bit See *universal/local bit*.

G.shdsl Single-pair high-data-rate digital subscriber line standardized by ITU and ANSI.

Gratuitous ARP frame An *Address Resolution Protocol* (ARP) Request frame in which the *source protocol address* (SPA) and *target protocol address* (TPA) are set to the source's IP address. If no ARP reply frames are received, the node can assume its IP address is unique within its subnetwork.

H.323 An ITU Recommendation that offers audio, video, and facsimile services over local area networks. It does not guarantee quality of service (QoS) levels. Focusing on voice services, it provides connections for moderate numbers of users and is incorporated in commercial offerings.

Half-closed In TCP, a connection in which one station has notified the other that it has completed its transmission, but the partner still has frames to send.

Half-duplex connection Supports debate-style communication. Messages can flow in both directions, but only in one direction at a time. Many older local area networks are half-duplex. Stations receive and transmit, but only one action can occur at a time.

Hardswitch A circuit switch. See also *softswitch*.

Hashing A mathematical procedure that maps address space into a smaller pointer space so that an address search is started by searching the smaller pointer field. The hashing function must produce a consistent hash value for the same address, and, for any arbitrary set of addresses, produce an approximately uniform distribution of pointers.

Header Administrative information added at the beginning of the PDU.

Hexadecimal representation Because writing 8-bit bytes can be tedious and is subject to errors, hexadecimal notation is used to represent them. Bytes are divided into two 4-bit binary words (nibbles) whose decimal values (0 to 15) are represented by the digits 0 through 9 and the letters A through F.

High-bit-rate digital subscriber line The DS-1 stream is split into two streams of 768 kbit/s. Each stream is transported (duplex) over a cable pair (dual-duplex transmission) up to 12,000 feet. For installations greater than 12,000 feet, repeaters (known as doublers) are employed.

High-bit-rate digital subscriber line 2 Operating over a single pair, HDSL2 provides T-1 speed over 26 AWG up to 12,000 feet.

High-Level Data Link Control Protocol (HDLC) Makes use of a special character, the flag character (0×7E), to mark the beginning and ending of the frame. Between these markers, the header and the trailer fields are of predetermined lengths.

Hop The action of passing over a data link between contiguous nodes.

Host (1) Provides processing services and data support to terminals and may support clients (if required). Processes and stores data. (2) In IPv6, a node that does not forward packets.

HTTP server A process that accepts *Hypertext Transfer Protocol* (HTTP) requests for connections from client programs and provides data in response.

Hub In Ethernet, a common hub to which each station is attached by separate twisted pair cables. The hub is a combiner and a repeater. It provides a separate port for each station and uses CSMA/CD to allocate the channel capacity to individual stations.

Hybrid Mode Two-way operation over a twisted pair is achieved by the use of hybrid transformers, echo-canceling devices, and adaptive filters.

Hypertext Transfer Protocol (HTTP) A request/response protocol that transfers data between client computers and HTTP servers. Requests are likely to be submitted from browsers.

IEEE 802.2 LLC header In the Ethernet, the IEEE 802.2 *logical link control* (LLC) header follows the IEEE 802.3 *medium access control* (MAC) header. Consists of destination and source service access point (DSAP and SSAP) fields that identify the points to which the payload is to be delivered in order to reach the proper upper-layer protocol. When used in conjunction with a SNAP header, DSAP and SSAP are set to 0×AA. See *IEEE 802.3 SNAP header*.

IEEE 802.3 Ethernet LAN Classic Ethernet LAN with extended message handling capability.

IEEE 802.3 MAC header In the Ethernet, IEEE MAC precedes LLC and *Subnetwork Access Protocol* (SNAP) headers. Consists of preamble and start delimiter fields, destination and source address fields, a length field that indicates how many bytes are contained in the remaining two headers and the payload so that the receiver can detect the frame check sequence.

IEEE 802.3 SNAP header In the Ethernet, permits protocols designed to operate with Classic Ethernet to be used in IEEE 802.3 applications. Messages sent over an IEEE 802.3 LAN use SNAP headers to identify the upper level protocols in use. For IP datagrams and ARP messages, the organization code is set to 0×00-00-00. In Token Ring, for IP datagrams and ARP messages, the organization code is set to 0×00-00-00. For both LANs the EtherType code is set to 0×08-00 for IP datagrams and 0×08-06 for ARP messages.

IEEE 802.5 header In Token Ring, the start delimiter field contains two nondata symbols (called J and K) that are violations of the signaling scheme. They alert the receiver to the incoming frame and provide a synchronizing signal. The access control field manages the token. The frame control field identifies the frame that follows as a Token Ring MAC management frame or a Token Ring data frame. The address fields contain the hardware addresses of the destination and source.

IEEE 802.5 LAN Token Rink LAN. Each station is connected to two others to form a single-thread loop that connects all the stations. The cabling system uses twisted-pairs with Manchester signaling. Data speeds of 4 Mbps, 16 Mbps, and 100 Mbps are in use. Access is provided by means of a token that circulates around the ring.

IEEE 802.5 Trailer The *frame check sequence* (FCS) is calculated over the data stream between the access control byte and the end of the payload. The FCS is checked at each node. The end delimiter contains J and K nondata symbols. In addition, it contains an intermediate frame indicator bit that identifies whether this frame is the last in a sequence (0), or there are more frames to follow (1). The end delimiter byte also contains an error detected indicator bit. Should the FCS fail, the node performing the check sets this bit and the destination node does not copy the frame. The frame status field contains duplicate address recognized indicator and frame copied indicator bits. They inform the sender that the node recognized its address and successfully copied the frame.

IEEE 802.11 Wireless Ethernet Employs the logical link sublayer of the data link layer; uses a unique MAC sublayer which includes collision *avoidance*; and has four physical layers that accommodate different implementations of the radio link.

Immutable field A field that is not changed during transport. The message, the transport header, and parts of the network header are immutable. Items such as time to live and network checksum vary with the number of nodes the frame passes. They are not immutable.

Implicit tag A tag implied by the contents of an untagged frame generated by a VLAN-unaware station or switch.

Impulse noise On a twisted-pair, short, intense bursts of noise that produce a voltage increase of 12 dB above the root-mean-squared (rms) background noise lasting 10 ms.

Indirect delivery The destination address does not carry the same network ID as the router, and the datagram is sent to the forwarding address contained in the table entry, for eventual delivery to its destination.

Individual/group bit Indicates whether the address is unicast (0) or multicast (1). For a broadcast address, the I/G bit is set to 1.

Industrial, scientific, and medical bands Unlicensed radio bands at 902 to 928 MHz (UHF), 2.4 to 2.5 GHz (S-band), and 5.725 to 5.875 GHz (C-band).

Information (I) frame One of three types of frame employed by LAP-D.

Ingress process In VLANs, initial process used by edge and core switches to process frames. Processors include: acceptable frame filter, rules module, and ingress filter. The edge switches use them to tag frames and discard those assigned to VLANs not recognized by the incoming port.

Initial sequence number A random number between 1 and 65,024 assigned to first byte of message. The sequence number is counted by bytes thereafter.

Integrated services digital network A switched digital network that provides voice, data, and image services through standard user interfaces based on 64-kbit/s clear channels.

Intelligent network A voice network with distributed call-processing capabilities. Implements *custom local area signaling services* (CLASS).

Interface In IPv6, the connection to a transmission medium over which packets are sent. In IPv6, all addressing is directed to interfaces.

Internet Assigned Numbers Authority An Internet agency responsible for the assignment and maintenance of well-known port numbers and other number codes.

Internet Control Message Protocol (ICMP) Reports errors and abnormal control conditions encountered by the first fragment of an IP datagram. ICMP messages are not sent for problems encountered by ICMP error messages or for problems encountered by multicast and broadcast datagrams.

Internet exchange point The lowest level of traffic exchange points between autonomous networks in the Internet.

Internet Group Management Protocol Manages multicast communications among a changing set of stations. To achieve one-to-many delivery, sends a single datagram to local nodes that forward it across routers to the distant nodes interested in receiving it.

Internet layer Layer 3 of the Internet model. Implements destination addressing, provides routing, and initiates advertising to build routing tables. The output of the Internet layer is a packet called an IP datagram.

Internet Protocol Adds addressing information necessary for routing the frame from source to destination.

Internet Protocol Datagram Consists of IP header, TCP or UDP header, and Payload.

Internet service provider Operator who provides access to the Internet for individuals and businesses.

Intracompany VPN A single campus network or an Intranet VPN in which encrypted communications are used to protect against security breaches within the enterprise.

Intranet VPN A VPN in which several enterprise campus networks are interconnected by tunnels over Internet connections (distributed backbone).

Inverse ARP For *nonbroadcast multiple access* (NBMA) WAN technologies (X.25, frame relay, and ATM) the network interface layer address is a virtual circuit identifier (not a MAC address). InvARP is used to determine the IP address of the interface at the other end of the virtual circuit.

IP Datagram A combination of the transport layer PDU and the Internet layer header.

IP multicast address A destination address associated with a group of hosts that receive the same packet(s) from a single source (one-to-many). Because routers forward IP multicast packets, the hosts can be located anywhere and may join or leave the group at will. Managing multicast groups is the purpose of the Internet Group Management Protocol.

IP Security A set of protocols that provides authentication and privacy services for IPv4 and IPv6.

IP version 6 Version 6 of the Internet Protocol. Increases the size of the address space from 4 bytes (IPv4) to 16 bytes and modifies other IPv4 header fields.

IPv6 address 128 bits long. In the preferred text representation, written as eight 16-bit hexadecimal sections separated by colons.

ISDN subscriber lines (1) Basic rate, 160 kbit/s; symmetrical channels; employs one pair; operates to 18,000 feet; uses 2B1Q line code. (2) Primary rate, 1.544 Mbps; symmetrical channels; operates over any existing DS-1 rate transmission systems (e.g., repeatered T-1 or HDSL).

Isochronous process A synchronizing process in which timing is embedded in the signal.

Jamming signal In Ethernet, in the event of a collision the colliding stations jam one another for a short time, so that all stations can hear that a collision has occurred. Then they cease transmitting. The jamming signal is 4 bytes long (usually 0×AA-AA-AA-AA).

Label In MPLS, edge routers insert labels describing the routing in the headers of IP frames. Labels are placed at the beginning of the packet so that, without having to consult switching tables, the receiving intermediate node can route the packet quickly to the next node. Labels are only locally significant and define one hop.

Label switched path In MPLS, labeling creates a virtual circuit for the transport of a burst of packets through the core switches called the *label switched path* (LSP).

Last mile A descriptive term of art used by communicators for the connection between subscribers and a telephone central office or a remote terminal.

Layer 2 Switch See *Bridge*.

Layer 2 Tunneling Protocol (L2TP) A Layer 2 protocol that encapsulates PPP frames for transmission over IP, X.25, frame relay, or ATM. L2TP supports multiple tunnels.

Layer 3 Switch see *Router*.

Link In IPv6, a bearer over which IPv6 is carried.

Link Access Protocol–Balanced A form of HDLC. First applied to the *user-network interface* (UNI) of X.25 packet switched networks. Works in *asynchronous balanced mode* (ABM). LAP-B served as the model for LAP-D, and LAP-F.

Link Access Protocol–D Channel A form of HDLC. First applied to the data channel (D-channel) in ISDN. Works in ABM.

Link Access Protocol–D core In frame relay, supports limited error detection (but not correction) on a link-by-link basis. It recognizes flags (to define frame limits), executes bit stuffing (to achieve bit-transparency), generates or confirms frame check sequences, destroys errored frames, and, using logical channel numbers, multiplexes frames over the links.

Link Access Protocol–D remainder In frame relay, acknowledges receipt of frames, requests retransmission of destroyed frames, repeats unacknowledged frames, and performs flow control.

Link Access Procedure–Frame Mode A form of HDLC. First applied to frame mode services over the ISDN user-network interface (UNI) on B-, D-, or H-channels. In frame relay, LAP-F is split in two parts that are applied separately. See *Link Access Protocol–D core* and *Link Access Protocol–D remainder*.

Link layer address In IPv6, the physical address of an interface.

Link state advertisement A routing message used by the Open Shortest Path First routing protocol.

Listening mode An application in the receiver is capable of receiving the connection request message that passes up the protocol stack to the port on which it is listening. To do this, applications issue passive OPEN function calls to specific port numbers or to ranges of port numbers.

Little Endian format In each code word, the *least significant bit* (LSB) is on the right end, and the *most significant bit* (MSB) is on the left end. Bits are read in ascending order from the least significant bit to the most significant bit. Bytes are numbered left to right, from 0 to *N*, and are read in ascending order. See *big Endian format*.

LLC header See *IEEE 802.2 LLC header*.

Loading coils On long connections (over 18 kft) it was standard practice to add loading coils to improve voice signal performance. Loading is used on 19, 22, and 24 gauge loops longer than 18,000 feet, or 26 gauge loops longer than 15,000 feet. D66 loading consists of 66 mH coils spaced 4,500 feet apart. H88 loading consists of 88 mH coils spaced 6,000 feet apart. The first load coil from the CO is located a half-section out.

Local area network Bus or ring connected, limited distance network that serves the data communication needs of a group of workers.

Local loop In the public telephone network, all wiring and facilities between the customers' premises and the central office.

Local-use unicast address In IPv6, address used for communication over a single link.

Logical link control sublayer Standardized in IEEE 802.2 as the upper sublayer of the data link layer. Defines the format and functions of the *protocol data unit* (PDU) passed between *service access points* (SAPs) in the source and destination stations. SAPs are associated with specific applications so that messages created by executing the applications can be identified and correlated.

Longitudinal signal See *common mode signal*.

Loopback address In IPv6, 0:0:0:0:0:0:0:1 or ::1. Used by a node to send a packet to itself.

MAC header See *IEEE 802.3 header*.

Manchester signal format A 1 is a positive current pulse of width one-half time slot that changes to a negative current pulse of equal magnitude and width one-half time slot. A 0 is a negative current pulse of width one-half time slot that changes to a positive current pulse of equal amplitude and width one-half time slot. The changeover occurs exactly at the middle of the time slot. The signal is always *zero-mean*.

Matched node Pairs of MNs are used to interconnect SONET rings and provide alternate paths for recovery in case of link failure. SONET traffic is duplicated and sent over two paths between the rings. One set of MNs provides the active path; the other set is on standby in case of failure of the active connection.

Maximum receive unit The maximum size frame that can be handled by a specific protocol.

Maximum segment size The greatest number of bytes that will be sent at any one time.

Maximum transmission unit The largest frame that can be sent to receiver.

Mean opinion score The subjective evaluation of speech quality. Reconstructed speech that is not distinguishable from natural speech is rated 5.0 (excellent). Studio quality voice has an MOS between 4.5 and 5.0. The 64-kbit/s PCM voice is known as *toll quality* voice and has an MOS of 4.3. Communication quality voice has an MOS between 3.5 and 4.0. A score below 3.5 is unacceptable for most applications.

Media Gateway Control Protocol An application-level protocol designed to facilitate multimedia sessions between the Internet and the *public switched telephone network* (PSTN). The media gateway acts between the two networks to translate media streams from circuit-switched networks into packet-based streams, and *vice versa*.

Medium access control address The hardware address of a node.

Medium access control sublayer Standardized in IEEE 802.3 as the lower sublayer of the data link layer. Defines the format and functions of headers and trailers that encapsulate the PDUs. The MAC sublayer contains the hardware addresses of source and destination.

Metallic noise See *circuit noise*.

Message In TCP/IP, the combination of application layer PDU and TCP or UDP header. Also called a segment.

Metropolitan area exchange In the Internet, a traffic exchange point between autonomous networks that serves a metropolitan area or region.

Microsplitter In ADSL, lowpass filter that stops data signals and passes voice signals.

Model A theoretical description of some aspect of the physical universe that identifies essential components and is amenable to analysis.

Modem A DCE that creates an analog signal for transmission over an analog circuit (e.g., telephone line). When sending, a modem converts the binary signals received from the DTE to analog signals that match the passband of the line. When receiving, a modem converts the analog signals to binary signals and passes them to the DTE.

Modulation A process that changes the amplitude, frequency, or phase of a carrier wave in sympathy with the instantaneous value of the modulating wave.

Movability Limited mobility.

Multicast address A terminating address that is shared by several stations. Used in point-to-many communication.

Multilevel threshold-3 signal format 1s are represented by a sequence of positive, zero, and negative currents. 0 is represented by the same current as the previous 1. MLT-3 is a bipolar version of NRZI.

Multiplexer A device that causes several similar signals to be carried on a single physical bearer.

Multiplexing The action of interleaving several signal streams so that they can be carried on a single bearer.

Multiprotocol Label Switching A project of IETF designed to address problems of scalability, speed, and quality of service in today and tomorrow's networks. Intended to extend to various packet-based technologies, the work has concentrated on speeding up the passage of IP frames across a network consisting of edge routers and core switches on *label switched paths* (LSPs).

Multistation access unit In Token Ring, provides the ability to connect stations by *unshielded twisted pair* (UTP) wiring to a central device in which the token ring is implemented. MAUs can be connected together in a ring so as to connect communities of stations. If the ring consists of dual cables (or fibers), or should a link fail, it can be made self-healing by arranging for one of the cables/fibers to reverse itself to provide loopback.

Near-end crosstalk A condition in which a signal transmitted over a twisted pair in a paired cable creates a disturbance in other pairs at the same end of the cable.

Neighbors In IPv6, nodes connected to the same link.

Network A (complex) tool that facilitates the execution of distributed data applications.

Network access point In the Internet, a highest-level traffic exchange point between autonomous networks. In the United States, four NAPs serve national and international traffic.

Network address translator A router that translates between private and public (Internet) addresses.

Network availability vector In IEEE 802.11, time in microseconds that the sender expects to occupy the radio channel.

Network control point An element in common-channel signaling network that contains databases needed to set up special services.

Network interface layer Layer 1 in the Internet model. Consists of two sublayers: the data link sublayer and the physical sublayer. Employs standard data link protocols. Determines and uses hardware addresses. Connects to LANs and WANs. The output of the network interface layer is a frame.

Network layer Layer 3 in the OSI model. Conditions packets to match the network(s) employed, and routes them over the network(s). If necessary, it will segment and reassemble the message to suit the maximum lengths the network(s) can accommodate.

Network mask A bit mask used to determine the network ID of the destination IP address (also see *subnet mask*).

Nibble Four contiguous bits. There are two nibbles in a byte.

Node In IPv6, any device that implements IPv6.

Noise The sum of all unwanted signals added to the message signal in the generation, transmission, and reception processes. The difference between the received signal and an ideal, attenuated, transmitted signal.

Nonblocking An existing (switch) path cannot prevent the setting up of another (switch) path.

Nonbroadcast multiple access links They connect more than two nodes, but do not provide multicast or broadcast services. The physical link supports multiple virtual circuits that connect to different nodes and service access points (SAPs). NBMA links include those that operate with X.25, frame relay, and cell relay or ATM. In an IP environment, *inverse ARP* (InvARP) is used to discover the IP addresses of the nodes on the other ends of the virtual circuits.

Nonreturn-to-zero-signal format 1 is represented by a positive current and 0 is represented by zero current. Sometimes called unipolar signaling, NRZ is used in integrated circuit chips, and other circuits, as well as in Gigabit Ethernet. Reliable timing information can be obtained from the signal provided some minimum number of bit transitions occurs in the data stream. Gigabit Ethernet uses an 8B/10B block code to guarantee the presence of sufficient 1s.

Nonreturn to zero, invert on ones 1 is represented by alternating a positive current and a zero current. 0 is represented by the same current as the previous 1. Put another way, the signal is unchanged for 0 and changes from its previous state for a 1. The strategy of inverting on ones produces a narrower frequency spectrum than NRZ. NRZI is used in FDDI and 100BASE-FX Ethernet. For reliable clock recovery, an adequate 1s density is guaranteed by the 4B/5B block code.

Nyquist Rate A signaling rate of 2B baud over a channel with a passband of *B* Hz.

Nyquist's theorem For a signal with bandwidth *B* Hz, sampling at a rate of 2*B* samples per second is sufficient to reconstruct the original signal.

Octet A word containing 8 bits whose values are derived from communication equipment. No matter how derived, common practice calls all 8-bit words bytes.

Open shortest path first (OSPF) A link state routing protocol. Routing information is disseminated as *link state advertisements* (LSAs) that contain the IDs of connected networks, network masks, and a cost figure. The LSA of each OSPF router is distributed throughout the network through logical relationships between neighboring routers known as adjacencies. When all current LSAs have been disseminated, the network is described as converged.

Open system A system defined by the parameters of the interfaces between its functional blocks.

Open systems interconnection reference model A model designed to guide the development of open systems so that they can communicate with each other. The model does not define the equipment that implements the communication functions, only the states that must exist between them. The model divides the actions of communicating hosts into seven independent activities that are invoked in sequence.

Optical carrier level 1 The optical equivalent of STS–1.

Optical carrier level N The optical equivalent of STS–N.

Orthogonal frequency division multiplexing A modulation technique that encodes a single user on several carriers. It splits a wide frequency band into narrow channels and inverse multiplexes a user's data signal on the subcarriers occupying a channel.

OSI model See *open systems interconnection reference model*.

Out-of-frame event In ESF, a condition when 2 out of 4 consecutive framing bits are incorrect (i.e., do not match the 101010 pattern).

Outside plant In the public telephone network, all wiring and facilities between the customers' premises and the central office.

Packet A sequence of as many as a few thousand bits. Some are users' data (the message) and some are control (overhead) data. In the control data is destination information that guides the packet across a network.

Passive OPEN function call See listening mode.

Packet Layer Protocol In the packet layer or X.25-3 layer, divides the user's data into fixed length segments and adds a 3-byte header.

Paired cable Cable that has twisted pairs as conductors.

Passband signal A complex signal produced by using a baseband signal to modify a property of another signal (called the carrier signal). The energy of the passband signal occupies a range (the passband) that encompasses the frequency of the carrier signal, or is contiguous with it. The sideband components of the passband signal carry the information contained in the baseband signal. A passband signal may be moved in the frequency plane by changing the frequency of the carrier signal.

PCF interframe space See point coordination function interframe space.

Peer-to-peer communication Communication between same layers of sending and receiving protocol stacks to set up and manage transfer of data.

Permanent virtual circuit A virtual connection that is permanently assigned between two stations.

Poll/final (bit) In LAP-D, the first bit of the second byte of the control field. In command frames, it is known as the poll (P) bit. When set to 1, it identifies this frame as requiring a response from the receiver. When set to 0, a response is not required. In response frames, the P/F bit is known as the final (F) bit. When set to 0, it identifies this frame as one of a continuing sequence. When set to 1, it is the final frame in the sequence.

Phase modulation The phase of the carrier is varied based on the value of the modulating signal.

Phase-shift keying Digital modulating technique in which the carrier signal may assume two phase values.

Physical layer Layer 1 of the OSI model. Converts the logical symbol stream into the physical symbol stream. Connects to transmission, routing, and switching facilities.

Physical layer convergence procedure In IEEE 802.11 Wireless Ethernet, adds fields to the frame for use on the radio link.

Physical sublayer Of the network interface layer in the Internet, is concerned with signals, wires, optical fibers, and individual transmission facilities.

Pinging Action to determine the status and reachability of a specific node. The message sent to the node is called an Internet Control Message Protocol (ICMP) echo request and the message returned is an ICMP echo reply.

Plain old telephone service (POTS) The services provided by the public switched telephone system.

Point coordination function interframe space In IEEE 802.11, interval between frames used during contention-free operation. Station with permission to transmit contention-free may begin after PIFS has elapsed and preempt contention-based traffic.

Point-to-point links They form a network segment with two terminal nodes. These links include telephone lines, ISDN circuits, digital subscriber lines, and T-carrier links. If the receiving node is the final destination, the IP address is irrelevant and ARP is not needed to resolve the destination MAC address. If the receiving node is not the final destination, the IP destination address will be required to facilitate further handoffs.

Point-to-Point Protocol Incorporates LAP-D. Provides full-duplex data link services between peers.

Point-to-Point Tunneling Protocol A Layer 2 protocol that encapsulates PPP frames in IP datagrams for transmission over an IP network. PPTP supports a single tunnel between client and server.

Port A message queue (or similar component) that connects one layer to the next to facilitate communication between them.

Port number Defines a location through which an application layer process sends a data segment to a transport layer process, or to which transport layer process delivers a data segment for an application layer process.

Power influence Noise caused by inductive interference from the public power system.

Presentation layer Layer 6 in the OSI model. Conditions the application PDU so as to compensate for local data formats in the sender and receiver.

Privacy Provides the sender and receiver with the assurance that, even if a message is intercepted, it is unlikely that it can be read.

Private IP address (1) An address space with 24 host ID bits. Contains a single network. Host IDs range from 0.0.0 to 255.255.255. (2) An address space with 20 host ID bits. Contains 16 network addresses that range from 172.16.0.0 through 172.31.0.0. Host IDs range from 0.0.0 through 15.255.255. (3) An address space with 16 host ID bits. Contains 256 network addresses that range from 192.168.0.0 through 192.168.255.0.

Probabilistic signal A signal whose future values are described in statistical terms based on past values.

Progress process In VLANs, an intermediate process used by edge and core switches to process frames. Forwards the tagged frame to the egress port and maintains the switching database. Frames are transported through a switching fabric and queued for transmission. The egress port is determined by the VLAN identifier and the MAC address of the destination. By observing traffic flow, the switch maps VLANs to ports to ensure an up-to-date database.

Protocol data unit Data exchanged between peer layers in a protocol stack.

Protocol interpreter When using File Transfer Protocol, the agent that sets up and controls the data exchange.

Proxy An entity that stands for another. A proxy is used to perform a function on the behalf of another.

Proxy ARP Software that allows a node other than the node whose IP address appears in an ARP request message to reply with the hardware address sought.

Proxy server (1) An application layer gateway that mediates between the private intranet and the public Internet. (2) A server that filters traffic according to rules formulated by administrators.

Pulse amplitude modulation A modulation format in which the amplitude of the carrier pulse is changed between a limited number of levels by the modulating data stream.

Pulse code modulation Encodes voice at 64 kbit/s with an MOS of 4.3 and processing delay of 0.125 ms.

Random signal A probabilistic signal whose values are limited to a given range. Over a long time, each value within the range will occur as frequently as any other value.

Real-Time Streaming Protocol An application-level protocol that compresses audio or video streams and passes them to transport layer protocols for transmission over the Internet.

Real-Time Transport Protocol (RTP) An application-level protocol that interfaces between the voice stream and existing transport protocols (UDP or TCP). RTP provides end-to-end delivery services for audio (and video) packets.

Receiver-side flow control Actions taken by the receiver so that the incoming byte stream does not overload the receiver's buffer storage.

Remote access VPN A VPN in which enterprise employees on the move can establish a dial-up connection to a remote ISP and create tunnels to enterprise campus networks.

Remote terminal In the local loop, a distribution terminal between the CO and subscriber serving area; may terminate a loop carrier system.

Repeater A device that regenerates, retimes, and reshapes signals. Extends the distance over which a signal is carried. Facilitates transport of packets across a network.

Request to send See clear to send.

Residual error rate In frame relay, the total number of frames sent minus the number of good frames received divided by the total number of frames sent.

Resource An object or service provided by a server. See uniform resource identifier.

Resource management cell To control the source bit rate when using the available bit rate (ABR) service, resource management (RM) cells are introduced periodically into the sender's stream. When an RM cell reaches the receiver, the receiver changes the direction bit to return the cell to the source. If the destination is congested, it sets the congestion indication bit and reduces the bit rate value to a rate it can support. On the return of the RM cell to the source, the sending rate is adjusted.

If the RM cell returns to the source without the congestion indication bit set, the sender can increase the sending rate.

Resource Reservation Protocol An application-level protocol that requests a path from a sender to a receiver (or multiple receivers) with given QoS features (i.e., bandwidth, delay less than).

Response frame Frame generated by receiver in response to a command frame.

Retransmission time-out In TCP, the amount of time within which an ACK is expected for the segment just sent. If the sender does not receive an ACK before the retransmission time-out (RTO) expires, the segment is retransmitted.

Round-trip time An interval from the time a message is sent to the time an ACK should be received. To prevent needless repetitions, round-trip time (RTT) is less than RTO (see retransmission time-out). Since RTT is likely to vary with traffic conditions, it must be monitored continually, and RTO must be adjusted accordingly.

Route descriptor Information inserted in Token Relay or VLAN-aware frames that describes a segment of the route to be followed between source and destination. Up to 14 segments are allowed.

Router (1) A device that interconnects networks. It forwards messages between them based on the destination network address and a table of possible routes. The path between sender and receiver is likely to contain numerous routers. When implemented mostly in hardware, it is called a Layer 3 switch. Each router advertises its status and capabilities and discovers the status and capabilities of its neighbors. (2) Using its up-to-date knowledge of the topology, an intelligent device that discovers routes across a network so as to guide frames towards their destination. (3) In IPv6, a node that forwards packets.

Routing The process of forwarding unicast or multicast packets from a sending host to (a) destination host(s).

Routing information indicator bit Indicates whether Token Ring source routing information is present. Token Ring source routing allows a Token Ring sending node to discover and specify a route to the destination in a Token Ring segment.

Routing Information Protocol A simple routing protocol with a periodic route-advertising routine that can be used in small- to medium-size networks. RIP is described as a distance vector routing protocol. The distance is the number of hops between the router and a specific network ID. Destinations with 16 or more hops are described as unreachable.

RTP Control Protocol Monitors QoS based on the periodic transmission of control packets. RTCP provides feedback on the quality of packet distribution.

Running disparity When using a two-set complementary block code, the receiver keeps track of whether more 1s than 0s, or more 0s than 1s, have been transmitted. The value of RD determines whether the transmitter selects the next code word as the one with more 1s than 0s, or the alternate with more 0s than 1s.

Scrambling By performing logical operations on the data stream at the transmitter, scrambling breaks up strings of the same symbol, or repeated patterns of symbols and makes the signal stream pseudorandom. At the receiver, by reversing the logical changes, the scrambled sequence is descrambled and the original data stream is restored.

Security association Lists the security parameters to be used in encrypted communication with a specific destination. The list includes: an identification number (security parameters index); a cryptographic algorithm; a key, or keys, that implement the algorithm; the lifetime of the key(s); and a list of sending stations that can use the security association.

Security parameters index Identifies the security association in use.

Segment The transport layer PDU.

Segmentation and reassembly sublayer (SAR) Part of the ATM adaptation layer. When sending, SAR divides CS PDU into 48-byte SAR PDUs and delivers them to the ATM layer. When receiving, receives 48-byte SAR PDUs from ATM layer, reconstructs CS PDUs, and sends them to CS.

Selective Acknowledgement Procedure The receiver sends acknowledgment for last good byte in series of good bytes and first good byte in next series of good bytes. The sender will repeat the bytes between the two numbers.

Self-crosstalk Crosstalk from the same type of data signal running in the same binder. May be divided into near end and far end.

Sender-side flow control Actions taken by the sender to send the byte stream as quickly as possible but without overloading the receiver or causing congestion on the links used.

Serial Line Internet Protocol (SLIP) A very simple packet-framing protocol that provides frame delimitation services only. To delimit IP datagrams, SLIP uses a special character. Called an END character (0×C0), it is placed at the beginning and ending of each IP datagram.

Server A device that stores data, organizes and maintains databases, and delivers copies of data files to clients on demand. A process that stores and distributes data.

Service access point A port within the sending or receiving device that permits PDUs to flow between contiguous protocol layers. May be a message queue that transfers PDUs to the upper level protocol agent identified by the EtherType entry.

Service access point identifier (SAPI) Each node may support several Internet layer protocols. SAPI values are assigned to identify the buffer/queue serving the specific protocol in the destination machine.

Service control point In intelligent network, unit with software to implement one or more custom local area signaling service (CLASS) features.

Session Initiation Protocol (SIP) A signaling protocol developed to facilitate telephone sessions and multimedia conferences in a unicast or multicast private network environment. Through gateways, SIP communicates with public terminals, and provides a limited menu of IN services.

Session layer Layer 5 in the OSI model. Manages the communication process.

Severely errored second In T-1, second in which from 320 to 333 ESF error events are present.

Short interframe space (SIFS) In IEEE 802.11, interval used for high-priority transmissions such as RTS/CTS frames and ACKs. SIFS is less than DIFS. Once a multiframe transmission has begun, subsequent frames are sent after SIFS interval. This preempts other frames that must wait for DIFS and a backoff time.

Signal transfer point A facility that performs as a link concentrator and message switcher to interconnect signaling end points. Routes signaling messages to the terminating switch or to the STP that serves the terminating switch.

Signaling rate One symbol per second is a signaling rate of 1 baud.

Simple and efficient layer In ATM, alternative name for AAL5.

Simple Mail Transfer Protocol (SMTP) A procedure that facilitates the transfer of electronic mail between computers. SMPT provides message transfer. It does not manage mailboxes or mail systems.

Simplex connection Supports announcement-style communication. Messages flow in one direction only, from sender to receiver.

Single-key cryptography Also known as secret-key cryptography, employs the same key for encryption and decryption. The key is a 64- or 128-bit-long bit pattern. To be effective, the key must be kept secret from everyone except the users.

Single-mode fiber In such a fiber, the central glass core is ≤10 microns in diameter. A significant (and essential) fraction of the optical energy travels in the cladding glass. Because its velocity is slightly higher than the energy in the core, conditions are right to support single-mode propagation. With a refractive index of 1.475, the velocity of energy in the core is approximately 200,000 km/sec (i.e., approximately two-thirds of the velocity of light in free-space).

Slot time In the Ethernet, the round-trip transmission time between a node at one end of the network and a node at the other end of the network. Usually, a slot time is assumed to be 512 bit times (i.e., 51.2 μs for a 10-Mbps LAN).

Socket The globally unique address of the application. It comprises the combination of port number and network address of the host.

Softswitch A multimedia packet switch. See also hardswitch.

Source routing Before a communication session begins, the source station discovers the routes to each station with which it is likely to communicate. During the session the source station selects the least cost route and inserts this routing information into the frames immediately following the source address.

Spanning Tree Protocol A protocol invoked to ensure frames sent between one station and another use the single, most efficient (least cost) path.

Star-star Original topology of local loop. One star is formed by the feeder cables and the CO, and a second ring of stars is formed by the distribution cables and each of the feeder distribution interfaces (FDIs).

Spread spectrum modulation A technique in which the message-bearing modulated signal is processed (i.e., modulated again) to occupy a much greater bandwidth than the minimum required to transmit the information it carries.

Splitter In ADSL, filter that separates voice and high-speed data signals.

Static routing Employs manually configured routes. A static router cannot dynamically adjust its routing table so that it is unable to react to the state of contiguous routers, and neighboring routers cannot update the static router's table.

Stop-and-wait ARQ A procedure in which the sender sends a frame then waits for the receiver to acknowledge error-free (ACK) or errored (NACK) receipt.

Store-and-forward In switching, the entire frame is received and stored in the input buffer before being forwarded over a switch path to the buffer serving the port connected to the destination. In the process of storing the frame, the buffer logic may check for errors and perform other frame management functions.

Subnet mask In IPv4, contains 32 bits that are configured as follows. If the bit position in the mask corresponds to a bit in the network ID, it is set to 1. If the bit position in the mask corresponds to a bit in the host ID, it is set to 0. By performing ANDing between the address and the subnet mask, the network ID can be found. What is left is the host ID.

Subnetting Creating additional smaller subnets by robbing some of the bits that are reserved for host IDs to become parts of the network IDs.

Subnetwork Access Protocol (SNAP) See IEEE 802.3 SNAP header.

Subrate digital line 2.4-56 kbit/s; symmetrical channels; employs one pair.

Superframe A block of 12 T-1 frames in which the framing bits are used to provide synchronization and other functions.

Supernetting A technique that assigns one network address to several subnets. It reduces the number of network IDs and masks the routers must maintain in their routing tables.

Supervisory frame One of three types of frame employed by LAP-D.

Switch (1) A device that selects paths or circuits so as to make real connections between sender and receiver. Normally, a switch will implement a direct connection, or a connection that only transits one or two additional switches. (2) Facilitates transport of packets across a network. (3) A multiport device that makes and breaks circuits. (4) A multiport device that selects virtual paths and virtual circuits to transport frames to specific destination. May contain buffers to hold frames until transport capacity is available. (5) A device with a number of simplex or duplex physical ports that receive and/or transmit frames. Each frame may be tagged or untagged.

Switched Ethernet hub A common hub in which individual input channels are connected to output channels by a nonblocking switching fabric. Collisions are eliminated. CSMA/CD is no longer needed. Stations do not have to wait for the bus to be quiet, and they can operate at the full bit rate of the switching fabric.

Synchronize flag In TCP, informs receiving host that sending host wishes to synchronize counting the forward data stream and establish other parameters preparatory to communication.

Synchronous digital hierarchy A hierarchy of transport speeds standardized by ITU for B-ISDN. The speeds are exactly three times SONET speeds.

Synchronous operation The stations and nodes are disciplined by a common clock. Actions occur at specific times in synchrony with other units in the network.

Synchronous optical network An all-digital, optical fiber transport structure that operates from 51.84 Mbps to 40 Gbps and beyond. SONETs serve as very high-speed backbones in Internet, as high-speed distribution networks in local exchange and interoffice plant, and provide optical transport channels for private connections. Usually SONETs are employed in rings to connect traffic collection points.

Synchronous payload envelope In SONET, part of a frame consisting of payload and path overhead. An SPE is generated 8,000 times a second. It contains n × 774 bytes (where n = 1, 3, 12, 24, 48, 96, ..., 792).

Synchronous transport module level 1 In SDH, a frame of 2,430 bytes at 155.52 Mbps. STM-1 = 3 STS-1 = STS-3.

Synchronous transport module level N In SDH, a frame of N × 2,430 bytes at N ×155.52 Mbps. STM-N frames are created by byte multiplexing N STM-1 frames. STM-N = N STM-1 = 3N STS-1.

Synchronous transport signal level 1 With a basic speed of 51.84 Mbps, STS-1 signals are designed to carry T-3 signals, or a combination of T-1, T-1C and T-2 signals that is equivalent to DS-3.

Synchronous transport signal level N With speeds that are multiples of STS-1, that is, N × 51.84 Mbps (where N may assume any integer value), STS-N signals are created by byte multiplexing N STS-1 signals. For various reasons, the values N = 3 (155.52 Mbps), 12 (622.08 Mbps), 24 (1244.16 Mbps), 48 (2488.32 Mbps), 96 (4,976.64 Mbps), 192 (9,953.28 Mbps), and 768 (39,813.12 Mbps) are preferred.

SYN flag See synchronize flag.

T-1 First digital transmission equipment widely deployed in the Bell System. Multiplexes 24 DS-0 (64 kbit/s) signals into one DS-1 (1.544 Mbps) signal (DS-1 = 24 DS-0s+ framing bit).

T-1C Multiplexes two DS-1 signals into one DS-1C (3.152 Mbps) signal (DS-1C = 48 DS-0s).

T-1 carrier line 1.544 Mbps; symmetrical channels; employs two pairs, one for each direction; with repeaters every 6,000 feet, operates up to 50 miles; uses AMI line code.

T-1 data frame Consists of 23 bytes of payload, 1 byte of signaling data, and a framing bit (the 193rd bit). The last bit of every data byte is set to 1. This action reduces the per channel data throughput to 56 kbit/s. Thus, the data throughput becomes 1.288 Mbps per T-1 line.

T-2 Multiplexes four DS-1 signals into one DS-2 (6.312 Mbps) signal (DS-2 = 96 DS-0s).

T-3 Multiplexes seven DS-2 signals into one DS-3 (44.736 Mbps) signal (DS-3 = 672 DS-0s). A special version developed for enterprise networks known as T3 SYNTRAN (synchronous transmission), multiplexes 28 DS-1 signals directly to DS-3.

T-4 Multiplexes six DS-3 signals into one DS-4 (274.176 Mbps) signal (DS-4 = 4,032 DS-0s).

T-4NA Multiplexes three DS-3 signals into one DS-4NA (139.264 Mbps) signal (DS-4NA = 2076 DS-0s).

Tag A 2-byte field inserted between the EtherType field of the SNAP header and the payload. The EtherType field contains the VLAN protocol identifier¾0´81-00. It indicates the frame is VLAN-tagged, and the next 2 bytes contain tag control information.

TCP checksum Calculated by summing 16-bit words over a pseudoheader, the TCP header, and the payload. The pseudo header contains the source IP address, the destination IP address, a TCP identifier code (0´06), and the length (in bytes) of the segment. If the number of bytes in this stream is odd, a padding byte is added. The 1s-complement of the total is sent to the receiver. At the receiver, the bytes are summed with the transmitted 1s-complement. If the result is all-1s, it is likely that transmission was free of errors.

TCP header Consists of 11 fields. Contains entries necessary for the sender and receiver to establish a connection and implement reliable delivery.

TCP/IP Transmission Control Protocol/Internet Protocol. TCP and IP are major procedures contained in the transport and Internet layers and are common to all communications that employ the Internet model. The term used to describe the software implementing data communication in the Internet.

TELNET A remote terminal protocol that allows a user to log on to another host elsewhere on the Internet.

Terminal A device used to input and display data. May have native computing and data processing capabilities. Relies on a host for support to accomplish the more intensive data processing tasks. Provides an interface for users' instructions and graphical or textual outputs.

Terminal endpoint identifier In HDLC, each physical node is assigned an address identifier. Assignment may be manual or automatic. The values are 0 through 63, manual assignment, 64 through 126, automatic assignment, 127 for temporary use during automatic TEI assignment.

Terminal multiplexer An end point or terminating device that connects originating or terminating electrical traffic to SONET. Has only one network connection.

Time to live In IPv4, field that records the number of hops the datagram may make before being destroyed. Each node handling the datagram reduces the TTL number by one. When TTL reaches zero, unless the node handling it is the destination host, the datagram is destroyed.

Token In Token Ring, an access control byte with start and end delimiters. The byte contains three priority bits, a token bit, a monitor bit, and three reservation bits.

Token Ring LAN See IEEE 802.5 LAN.

Toll quality voice 64-kbit/s PCM voice. It has an MOS of 4.3.

Trailer Administrative information added at the end of the PDU.

Translating bridge Connects dissimilar LANs at the data link sublayer by translating different field entries.

Transmission Control Protocol (TCP) Provides connection-oriented services. Before data is transferred between processes running on two hosts, a duplex connection is negotiated (connection establishment process). At the end of the communication exchange, it is closed using a termination process. Provisions are made for recovery from untoward events. Data sent over a TCP connection are tracked by the sender and receiver to ensure reliable delivery service.

Transport layer (1) Layer 4 in the OSI model. Responsible for the sequenced delivery of the entire message including error control, flow control, and quality of service requirements, if they are invoked. (2) Layer 3 in Internet model. Establishes, controls and terminates network connections between ports on source and destination. Implements error control and flow control if required. The transport layer PDU is called a segment or message.

Trellis coding A coding that employs twice as many signal points in the constellation as are needed to represent the data. This redundancy is a form of forward error correction coding and is used to reduce errors.

Tunnel A secure temporary connection between two points in an insecure public network.

Tunneling The action of encapsulating an encrypted datagram inside another datagram so that it can be forwarded between two points over an insecure temporary connection without making use of its contents.

Twisted pair Two insulated wires twisted together. Also known as a cable pair.

Two binary, one quaternary signal format Four signal levels (± 3 and ± 1) each represent a pair of bits. Of each pair, the first bit determines whether the level is positive or negative (1 = +ve, 0 = –ve) and the second bit determines the magnitude of the level (1 = |1|, 0 = |3|).

Two-key cryptography Also known as public-key cryptography, employs two keys. One key is available to the public (public key); the other key is known only to its owner (private key). Either key can be used to create encrypted messages. They are decrypted by the other key.

UDP checksum Calculated by summing 16-bit words over the UDP datagram (Header + Application PDU) and a pseudoheader that consists of the source IP address, the destination IP address, an unused byte, a byte that identifies the UDP protocol, and the length (in bytes) of the segment. If the number of bytes in this stream is odd, a padding byte is added. (The padding byte is for computation only. It is not transmitted.) The 1s-complement of the total is sent to the receiver. At the receiver, the bytes are summed with the transmitted 1s-complement. If the result is all-1s, it is likely that transmission was free of errors.

UDP datagram Ideal carrier for short messages, such as requests, answers, and repetitive announcements, sent to single locations using IP unicast addresses. In addition, UDP is used whenever data is sent to multiple locations using IP multicast or broadcast addresses.

UDP data unit Application PDU encapsulated by a UDP header.

Unacknowledged connectionless service Message-handling feature of IEEE 802.3 Ethernet LAN. The receiver does not acknowledge messages. Error control and flow control are not employed. The service is used in applications where the occasional loss or corruption of a PDU can be corrected by procedures invoked by the upper layer communicating software entities.

Unicast address The originating or terminating address of a single station.

Uniform resource identifier includes two items, uniform resource locator (URL) and uniform resource name (URN). A resource is requested by location or name.

Universal/local bit Indicates whether the address is globally unique (0) or locally administered (1).

Unnumbered (U) frame One of three types of frame employed by LAP-D.

Unspecified address In IPv6, 0:0:0:0:0:0:0:0 or ::. Used by nodes in the initializing process before they learn their own addresses.

Unwrapped See decapsulation.

Upstream The direction from the subscriber to the CO.

Urgent pointer A field that records the number of bytes from the beginning of the TCP header to the last byte of urgent data in the payload.

User Datagram Protocol (UDP) A simple transport layer protocol for applications that do not require reliable delivery service. UDP is connectionless. UDP messages are sent without negotiating a connection. They carry no sequence number, and their receipt goes unacknowledged. UDP datagrams do not provide information on buffer storage available at the receiver or sender, they are not segmented, nor do they provide flow control information.

Very-high bit-rate digital subscriber line An extension of asynchronous digital subscriber line technology to rates up to 52 Mbps downstream.

Virtual circuit A circuit with a logical identifier. Several virtual circuits share a physical circuit. Known as nonbroadcast access links, the physical circuits connect Internet layer entities in the sending terminal with Internet layer entities in one or more receiving terminals. X.25 packet switching, frame relay, and ATM employ NBMA links.

Virtual local area network A logical network created from specific stations in a catenet so that they appear to occupy a private LAN.

Virtual path A group of virtual circuits that connect the same endpoints.

Virtual private network (VPN) A data network composed of private and public sections that permits sending protected data over unprotected public connections without the risk of compromise by eavesdroppers, thieves, or those who would sabotage information. To the users, a VPN appears as a private network.

Virtual tributary In SONET, a synchronous payload that occupies 9 rows × n columns in the SPE. Thus, the virtual tributary for DS-1 consists of 27 bytes (9 rows ×3 columns). Twenty-four of them are DS-0 bytes from the T1 frame, 2 bytes are overhead related to the virtual tributary, and 1 byte is framing information. A similar arrangement exists in synchronous digital hierarchy (SDH).

VLAN association rules Also known as ingress rules. Simple rules are based on port ID, MAC address, protocol type, and application. More complex rules parse the relevant information fields.

VLAN-aware station A station organized to generate, insert, or accept and interpret tags. The tag can be placed in the frame when the frame is first generated or it may be present in an arriving frame. In addition, source routing instructions can be attached to ensure the frame is forwarded by a specific route through the intervening catenet.

VLAN-unaware station One that is unable to accept tags. When presented with a tagged frame, the unaware station will most likely destroy the frame.

Wavelength division multiplex Several optical carriers are transmitted simultaneously in the same fiber.

Well-known port numbers Ports #0 through #1023 whose use is controlled by IANA.

Wide area network Consists of long-distance links joined together at various points by nodes that perform switching or routing functions. The nodes move frames from one link to another so as to guide them between the sending local network and the receiving local network. All links will carry several multiplexed channels. Operation is synchronous or asynchronous.

Wired equivalent privacy In IEEE 802.11, a symmetric key security procedure.

Wire speed At the speed of signals on a wire. Description intended to differentiate the speed of solid-state logic devices and logic derived from a software program.

Workgroup network Interconnected desktop networks (LANs) that may be situated in several areas (floors, bays) of a single building. Consists of two, or more, desktop networks bridged together.

X.25 ITU recommendation that describes the user-network interface of a packet switch. X.25 defines a three-layer protocol stack.

X.25-1 X.25 physical layer.

X.25-2 X.25 data link layer.

X.25-3 X.25 packet layer.

Zero-byte time slot interchange Coding that makes entire 64 kbit/s channel available to customer.

Selected Bibliography

Those of you who have reached this chapter may be wondering where to obtain specific information. After all, the rest of this book does little more than acquaint you with the field of knowledge that is modern data communications. That was my intention, to paint the scene, to chronicle what is involved. Because it seemed an impossible task, I soon realized I could not give references for all my statements, so I have given none.

Where can you get further information? First, I suggest asking questions of a good search engine. There are literally hundreds of pages available on the subjects I have discussed. Choose wisely and you will have the latest information. It will be more current than information contained in a book. Second, you may wish to consult some of the books listed here for greater depth and understanding of specific topics. They are included because I have found them useful in this endeavor.

Brown, S., *Implementing Virtual Private Networks*, New York: McGraw-Hill, 1999. Almost 600 pages of practical considerations for implementing VPNs.

Comer, D. A., *Internetworking with TCP/IP*, *Volume 1*, 4th ed., Upper Saddle River, NJ: Prentice Hall, 2000. Generally regarded as the bible on TCP/IP. It is very readable and thorough.

De Prycker, M., *Asynchronous Transfer Mode: Solution for Broadband ISDN*, 2nd ed., Hemel Hempstead, Hertfordshire, England: Ellis Horwood, 1993. The original book on ATM written by a pioneer in the field. Somewhat dated, but an interesting read nonetheless.

Douskalis, B., *IP Telephony: The Integration of Robust VoIP Services*, Upper Saddle River, NJ: Prentice-Hall, 2000. An impressive analysis of the problems involved in converting POTS to VoIP.

Gast, M. S., *802.11 Wireless Networks: The Definitive Guide*, Sebastopol, CA: O'Reilly, 2002. Covers all kinds of wireless networks, not just the last mile. A truly comprehensive, well-written book.

Ghosh, A. K., *E-Commerce Security:Weak Links, Best Defenses*, New York: John Wiley & Sons, 1998. A very practical book on protecting privacy even if you let everyone use your network.

Goralski, W., *SONET: A Guide to Synchronous Optical Networks*, New York: McGraw-Hill, 1997. Almost 500 pages on SONET. May be a little dated, but the principles are well spelled out.

Johnston, A. B., *SIP: Understanding the Session Initiation Protocol*, 2nd ed., Norwood, MA: Artech House, 2004. Describes the use of SIP for call signaling, IP telephony and wireless multimedia communications.

Kadambi, J., I. Crawford, and M. Kalkunte, *Giganet Ethernet*, Upper Saddle River, NJ: Prentice Hall, 1998. An outstanding description of Ethernet in all of its flavors.

Lee, T., and J. Davies, *Microsoft Windows 2000 TCP/IP Protocols and Services Technical Reference*, Redmond, WA: Microsoft Press, 2000. Directed specifically to Microsoft applications, it gives a comprehensive, bit-by-bit description of TCP/IP.

Minoli, D., *Enterprise Networking: Fractional T1 to SONET, Frame Relay to BISDN*, Norwood, MA: Artech House, 1993. A book I continue to use. It covers digital transmission systems in public and private networks in great detail. It is still available from Amazon.com.

Minoli, D., *Telecommunications Technology Handbook*, 2nd ed., Norwood, MA: Artech House, 2003. Describes optical networking and other advanced multimedia delivery systems.

Minoli, D., and A. Schmidt, *Internet Architectures*, New York: John Wiley & Sons, 1999. Gives an overview of Internet operations and technology in 500 pages. The enormity of the network of networks is apparent.

Muller, N. J., *LANs to WANs: The Complete Management Guide*, Norwood, MA: Artech House, 2003. Comprehensive guide to management of network reliability, storage resources, and so forth.

Radcom Ltd, *Telecom Protocol Finder*, New York: McGraw-Hill, 2001. A compilation of telecom protocols at the bit level. It is a useful reference to have.

Rauschmayer, D. J., *ADSL/VDSL Principles: A Practical and Precise Study of Asymmetric Digital Subscriber Lines and Very High Speed Digital Subscriber Lines*, Indianapolis, IN: Macmillan Technical Publishing, 1999. Gives a technical description of the operation of digital subscriber lines, particularly ADSL and VDSL. It contains good diagrams and the mathematics is explained well.

Reeve, W. D., *Subscriber Loop Signaling and Transmission Handbook:Digital*, New York: IEEE Press, 1995. Also, *Subscriber Loop Signaling and Transmission Handbook: Analog*, New York: IEEE Press, 1992. These are truly handbooks on the local loop. Well written and organized, they contain just about everything you need to know about it.

Seifert, R., *The Switch Book*, New York: John Wiley & Sons, 2000. An impressive book that, in more than 500 pages, addresses the operation of LANs, including bridging, routing, and tagging in great detail.

Sinnreich, H., and A. B. Johnston, *Internet Communications Using SIP: Delivering VoIP and Multimedia Services with Session Initiation Protocol*, New York: John Wiley & Sons, 2001. Describes the use of SIP to provide comprehensive multimedia services.

Smith, M., *Virtual LANs: A Guide to Construction, Operation and Utilization*, New York: McGraw-Hill, 1998. Does exactly what the title says. Contains 400 pages of principles and practice.

Viterbi, A. J., *CDMA: Principles of Spread Spectrum Communication*, Reading, MA: Addison-Wesley, 1995. The pioneer of spread spectrum communications explains it all. The mathematics is somewhat overpowering, but the book is well worth reading.

About the Author

E. Bryan Carne received a Ph.D in electrical engineering from the University of London. He began his professional career in the United States working on Univac computers and then pilot production and manufacturing of proprietary devices. Beginning in 1959, he worked as a manager, director, and general manager for contractors associated with military communications and intelligence collection programs.

In 1969, Dr. Carne completed the Advanced Management Program at Harvard University, in Cambridge, Massachusetts. He joined GTE Laboratories in Waltham, Massachusetts, to direct its telecommunications programs.

In 1986, Dr. Carne was appointed visiting professor of electrical engineering at Northeastern University in Boston, Massachusetts, and later, BellSouth distinguished visiting professor of telecommunications and information management at Christian Brothers University in Memphis, Tennessee.

Dr. Carne is the author of four books on telecommunications: *Telecommunications Primer: Data, Voice and Video Communications, Second Edition* (Prentice Hall, 1999), *Telecommunications Topics: Applications of Functions & Probabilities in Electronic Communications* (Prentice Hall, 1999), *Telecommunications Primer: Signals, Building Blocks and Networks* (Prentice Hall, 1995), and *Modern Telecommunication (Applications of Communications Theory)* (Plenum Press, 1984). He is a Life Senior Member of IEEE. Living in Peterborough, New Hampshire, he divides his time between writing, occasional teaching, hiking, and his grandchildren.

Index

Recent Titles in the Artech House Telecommunications Library

Vinton G. Cerf, Senior Series Editor

Access Networks: Technology and V5 Interfacing, Alex Gillespie

Achieving Global Information Networking, Eve L. Varma et al.

Advanced High-Frequency Radio Communications, Eric E. Johnson et al.

ATM Interworking in Broadband Wireless Applications, M. Sreetharan and S. Subramaniam

ATM Switches, Edwin R. Coover

ATM Switching Systems, Thomas M. Chen and Stephen S. Liu

Broadband Access Technology, Interfaces, and Management, Alex Gillespie

Broadband Local Loops for High-Speed Internet Access, Maurice Gagnaire

Broadband Networking: ATM, SDH, and SONET, Mike Sexton and Andy Reid

Broadband Telecommunications Technology, Second Edition, Byeong Lee, Minho Kang, and Jonghee Lee

The Business Case for Web-Based Training, Tammy Whalen and David Wright

Centrex or PBX: The Impact of IP, John R. Abrahams and Mauro Lollo

Chinese Telecommunications Policy, Xu Yan and Douglas Pitt

Communication and Computing for Distributed Multimedia Systems, Guojun Lu

Communications Technology Guide for Business, Richard Downey, Seán Boland, and Phillip Walsh

Community Networks: Lessons from Blacksburg, Virginia, Second Edition, Andrew M. Cohill and Andrea Kavanaugh, editors

Component-Based Network System Engineering, Mark Norris, Rob Davis, and Alan Pengelly

Computer Telephony Integration, Second Edition, Rob Walters

Customer-Centered Telecommunications Services Marketing, Karen G. Strouse

Deploying and Managing IP over WDM Networks, Joan Serrat and Alex Galis, editors

Desktop Encyclopedia of the Internet, Nathan J. Muller

Digital Clocks for Synchronization and Communications, Masami Kihara, Sadayasu Ono, and Pekka Eskelinen

Digital Modulation Techniques, Fuqin Xiong

E-Commerce Systems Architecture and Applications, Wasim E. Rajput

Engineering Internet QoS, Sanjay Jha and Mahbub Hassan

Error-Control Block Codes for Communications Engineers, L. H. Charles Lee

Essentials of Modern Telecommunications Systems, Nihal Kularatna and Dileeka Dias

FAX: Facsimile Technology and Systems, Third Edition, Kenneth R. McConnell, Dennis Bodson, and Stephen Urban

Fundamentals of Network Security, John E. Canavan

Gigabit Ethernet Technology and Applications, Mark Norris

Guide to ATM Systems and Technology, Mohammad A. Rahman

A Guide to the TCP/IP Protocol Suite, Floyd Wilder

Home Networking Technologies and Standards, Theodore B. Zahariadis

Information Superhighways Revisited: The Economics of Multimedia, Bruce Egan

Installation and Maintenance of SDH/SONET, ATM, xDSL, and Synchronization Networks, José M. Caballero et al.

Integrated Broadband Networks: TCP/IP, ATM, SDH/SONET, and WDM/Optics, Byeong Gi Lee and Woojune Kim

Internet E-mail: Protocols, Standards, and Implementation, Lawrence Hughes

Introduction to Telecommunications Network Engineering, Second Edition, Tarmo Anttalainen

Introduction to Telephones and Telephone Systems, Third Edition, A. Michael Noll

An Introduction to U.S. Telecommunications Law, Second Edition, Charles H. Kennedy

IP Convergence: The Next Revolution in Telecommunications, Nathan J. Muller

LANs to WANs: The Complete Management Guide, Nathan J. Muller

The Law and Regulation of Telecommunications Carriers, Henk Brands and Evan T. Leo

Managing Internet-Driven Change in International Telecommunications, Rob Frieden

Marketing Telecommunications Services: New Approaches for a Changing Environment, Karen G. Strouse

Mission-Critical Network Planning, Matthew Liotine

Multimedia Communications Networks: Technologies and Services, Mallikarjun Tatipamula and Bhumip Khashnabish, editors

Next Generation Intelligent Networks, Johan Zuidweg

Open Source Software Law, Rod Dixon

Performance Evaluation of Communication Networks, Gary N. Higginbottom

Performance of TCP/IP over ATM Networks, Mahbub Hassan and Mohammed Atiquzzaman

Practical Guide for Implementing Secure Intranets and Extranets, Kaustubh M. Phaltankar

Practical Internet Law for Business, Kurt M. Saunders

Practical Multiservice LANs: ATM and RF Broadband, Ernest O. Tunmann

Principles of Modern Communications Technology, A. Michael Noll

A Professional's Guide to Data Communication in a TCP/IP World, E. Bryan Carne

Programmable Networks for IP Service Deployment, Alex Galis et al., editors

Protocol Management in Computer Networking, Philippe Byrnes

Pulse Code Modulation Systems Design, William N. Waggener

Security, Rights, and Liabilities in E-Commerce, Jeffrey H. Matsuura

Service Level Management for Enterprise Networks, Lundy Lewis

SIP: Understanding the Session Initiation Protocol, Second Edition, Alan B. Johnston

Smart Card Security and Applications, Second Edition, Mike Hendry

SNMP-Based ATM Network Management, Heng Pan

Spectrum Wars: The Policy and Technology Debate, Jennifer A. Manner

Strategic Management in Telecommunications, James K. Shaw

Strategies for Success in the New Telecommunications Marketplace, Karen G. Strouse

Successful Business Strategies Using Telecommunications Services, Martin F. Bartholomew

Telecommunications Cost Management, S. C. Strother

Telecommunications Department Management, Robert A. Gable

Telecommunications Deregulation and the Information Economy, Second Edition, James K. Shaw

Telecommunications Technology Handbook, Second Edition, Daniel Minoli

Telemetry Systems Engineering, Frank Carden, Russell Jedlicka, and Robert Henry

Telephone Switching Systems, Richard A. Thompson

Understanding Modern Telecommunications and the Information Superhighway, John G. Nellist and Elliott M. Gilbert

Understanding Networking Technology: Concepts, Terms, and Trends, Second Edition, Mark Norris

Videoconferencing and Videotelephony: Technology and Standards, Second Edition, Richard Schaphorst

Visual Telephony, Edward A. Daly and Kathleen J. Hansell

Wide-Area Data Network Performance Engineering, Robert G. Cole and Ravi Ramaswamy

Winning Telco Customers Using Marketing Databases, Rob Mattison

WLANs and WPANs towards 4G Wireless, Ramjee Prasad and Luis Muñoz

World-Class Telecommunications Service Development, Ellen P. Ward

For further information on these and other Artech House titles,
including previously considered out-of-print books now available through our
In-Print-Forever® (IPF®) program, contact:

Artech House
685 Canton Street
Norwood, MA 02062
Phone: 781-769-9750
Fax: 781-769-6334
e-mail: artech@artechhouse.com

Artech House
46 Gillingham Street
London SW1V 1AH UK
Phone: +44 (0)20 7596-8750
Fax: +44 (0)20 7630-0166
e-mail: artech-uk@artechhouse.com

Find us on the World Wide Web at:
www.artechhouse.com