PART FOUR

Local Area Networks

The trend in local area networks (LANs) involves the use of shared transmission media or shared switching capacity to achieve high data rates over relatively short distances. Several key issues present themselves. One is the choice of transmission medium. Whereas coaxial cable was commonly used in traditional LANs, contemporary LAN installations emphasize the use of twisted pair or optical fiber. In the case of twisted pair, efficient encoding schemes are needed to enable high data rates over the medium. Wireless LANs have also assumed increased importance. Another design issue is that of access control.

ROAD MAP FOR PART FOUR

Chapter 15 Local Area Network Overview

The essential technology underlying all forms of LANs comprises topology, transmission medium, and medium access control technique. Chapter 15 examines the first two of these elements. Four topologies are in common use: bus, tree, ring, and star. The most common transmission media for local networking are twisted pair (unshielded and shielded), coaxial cable (baseband and broadband), optical fiber, and wireless (microwave and infrared). These topologies and transmission media are discussed, with the exception of wireless, which is covered in Chapter 17.

The increasing deployment of LANs has led to an increased need to interconnect LANs with each other and with WANs. Chapter 15 also discusses a key device used in interconnecting LANs: the bridge.

Chapter 16 High-Speed LANs

Chapter 16 looks in detail at the topologies, transmission media, and MAC protocols of the most important LAN systems in current use; all of these have been defined in standards documents. The most important of these is Ethernet, which has been deployed in versions at 10 Mbps, 100 Mbps, 1 Gbps, and 10 Gbps. Then the chapter looks at Fibre Channel.

Chapter 17 Wireless LANs

Wireless LANs use one of three transmission techniques: spread spectrum, narrowband microwave, and infrared. Chapter 17 provides an overview wireless LAN technology and applications. The most significant set of standards defining wireless LANs are those defined by the IEEE 802.11 committee. Chapter 17 examines this set of standards in depth.

CHAPTER 15

Local Area Network Overview

The whole of this operation is described in minute detail in the official British Naval History, and should be studied with its excellent charts by those who are interested in its technical aspect. So complicated is the full story that the lay reader cannot see the wood for the trees. I have endeavored to render intelligible the broad effects.

—*The World Crisis*, Winston Churchill

KEY POINTS

- A LAN consists of a shared transmission medium and a set of hardware and software for interfacing devices to the medium and regulating the orderly access to the medium.
- The topologies that have been used for LANs are ring, bus, tree, and star. A ring LAN consists of a closed loop of repeaters that allow data to circulate around the ring. A repeater may also function as a device attachment point. Transmission is generally in the form of frames. The bus and tree topologies are passive sections of cable to which stations are attached. A transmission of a frame by any one station can be heard by any other station. A star LAN includes a central node to which stations are attached.
- A set of standards has been defined for LANs that specifies a range of data rates and encompasses a variety of topologies and transmission media.
- In most cases, an organization will have multiple LANs that need to be interconnected. The simplest approach to meeting this requirement is the bridge.
- Hubs and switches form the basic building blocks of most LANs.

We turn now to a discussion of **local area networks** (LANs). Whereas wide area networks may be public or private, LANs usually are owned by the organization that is using the network to interconnect equipment. LANs have much greater capacity than wide area networks, to carry what is generally a greater internal communications load.

In this chapter we look at the underlying technology and protocol architecture of LANs. Chapters 16 and 17 are devoted to a discussion of specific LAN systems.

15.1 BACKGROUND

The variety of applications for LANs is wide. To provide some insight into the types of requirements that LANs are intended to meet, this section provides a brief discussion of some of the most important general application areas for these networks.

Personal Computer LANs

A common LAN configuration is one that supports personal computers. With the relatively low cost of such systems, individual managers within organizations often independently procure personal computers for departmental applications, such as spreadsheet and project management tools, and Internet access.

But a collection of department-level processors will not meet all of an organization's needs; central processing facilities are still required. Some programs, such as econometric forecasting models, are too big to run on a small computer. Corporate-wide data files, such as accounting and payroll, require a centralized facility but should be accessible to a number of users. In addition, there are other kinds of files that, although specialized, must be shared by a number of users. Further, there are sound reasons for connecting individual intelligent workstations not only to a central facility but to each other as well. Members of a project or organization team need to share work and information. By far the most efficient way to do so is digitally.

Certain expensive resources, such as a disk or a laser printer, can be shared by all users of the departmental LAN. In addition, the network can tie into larger corporate network facilities. For example, the corporation may have a building-wide LAN and a wide area private network. A communications server can provide controlled access to these resources.

LANs for the support of personal computers and workstations have become nearly universal in organizations of all sizes. Even those sites that still depend heavily on the mainframe have transferred much of the processing load to networks of personal computers. Perhaps the prime example of the way in which personal computers are being used is to implement client/server applications.

For personal computer networks, a key requirement is low cost. In particular, the cost of attachment to the network must be significantly less than the cost of the attached device. Thus, for the ordinary personal computer, an attachment cost in the hundreds of dollars is desirable. For more expensive, high-performance workstations, higher attachment costs can be tolerated.

Backend Networks and Storage Area Networks

Backend networks are used to interconnect large systems such as mainframes, supercomputers, and mass storage devices. The key requirement here is for bulk data transfer among a limited number of devices in a small area. High reliability is generally also a requirement. Typical characteristics include the following:

- **High data rate:** To satisfy the high-volume demand, data rates of 100 Mbps or more are required.

- **High-speed interface:** Data transfer operations between a large host system and a mass storage device are typically performed through high-speed parallel I/O interfaces, rather than slower communications interfaces. Thus, the physical link between station and network must be high speed.
- **Distributed access:** Some sort of distributed medium access control (MAC) technique is needed to enable a number of devices to share the transmission medium with efficient and reliable access.
- **Limited distance:** Typically, a backend network will be employed in a computer room or a small number of contiguous rooms.
- **Limited number of devices:** The number of expensive mainframes and mass storage devices found in the computer room generally numbers in the tens of devices.

Typically, backend networks are found at sites of large companies or research installations with large data processing budgets. Because of the scale involved, a small difference in productivity can translate into a sizable difference in cost.

Consider a site that uses a dedicated mainframe computer. This implies a fairly large application or set of applications. As the load at the site grows, the existing mainframe may be replaced by a more powerful one, perhaps a multiprocessor system. At some sites, a single-system replacement will not be able to keep up; equipment performance growth rates will be exceeded by demand growth rates. The facility will eventually require multiple independent computers. Again, there are compelling reasons for interconnecting these systems. The cost of system interrupt is high, so it should be possible, easily and quickly, to shift applications to backup systems. It must be possible to test new procedures and applications without degrading the production system. Large bulk storage files must be accessible from more than one computer. Load leveling should be possible to maximize utilization and performance.

It can be seen that some key requirements for backend networks differ from those for personal computer LANs. High data rates are required to keep up with the work, which typically involves the transfer of large blocks of data. The equipment for achieving high speeds is expensive. Fortunately, given the much higher cost of the attached devices, such costs are reasonable.

A concept related to that of the backend network is the **storage area network** (SAN). A SAN is a separate network to handle storage needs. The SAN detaches storage tasks from specific servers and creates a shared storage facility across a high-speed network. The collection of networked storage devices can include hard disks, tape libraries, and CD arrays. Most SANs use Fibre Channel, which is described in Chapter 16. Figure 15.1 contrasts the SAN with the traditional server-based means of supporting shared storage. In a typical large LAN installation, a number of servers and perhaps mainframes each has its own dedicated storage devices. If a client needs access to a particular storage device, it must go through the server that controls that device. In a SAN, no server sits between the storage devices and the network; instead, the storage devices and servers are linked directly to the network. The SAN arrangement improves client-to-storage access efficiency, as well as direct storage-to-storage communications for backup and replication functions.

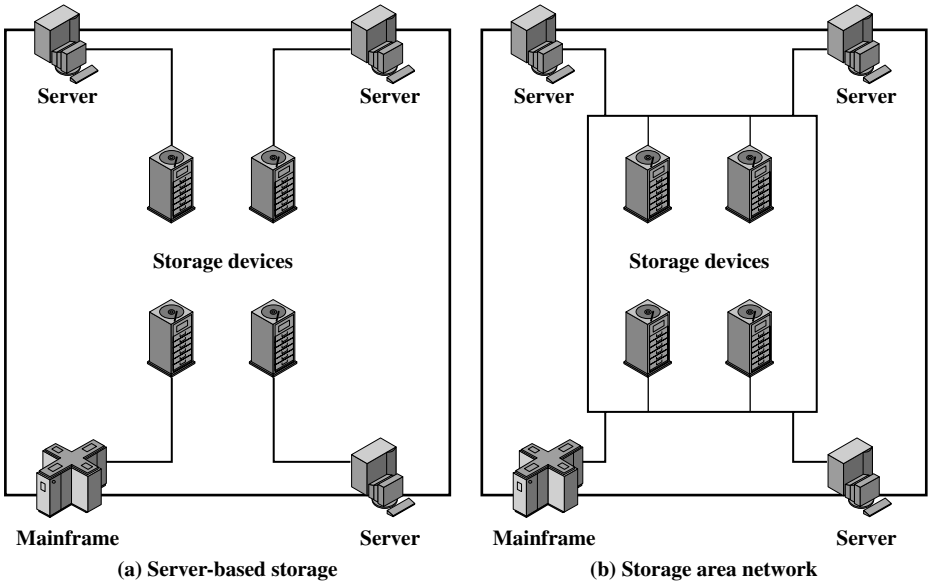

Figure 15.1 The Use of Storage Area Networks [HURW98]

High-Speed Office Networks

Traditionally, the office environment has included a variety of devices with low- to medium-speed data transfer requirements. However, applications in today's office environment would overwhelm the limited speeds (up to 10 Mbps) of traditional LAN. Desktop image processors have increased network data flow by an unprecedented amount. Examples of these applications include fax machines, document image processors, and graphics programs on personal computers and workstations. Consider that a typical page with 200 picture elements, or pels[1] (black or white points), per inch resolution (which is adequate but not high resolution) generates 3,740,000 bits (8.5 inches $\times$ 11 inches $\times$ 40,000 pels per square inch). Even with compression techniques, this will generate a tremendous load. In addition, disk technology and price/performance have evolved so that desktop storage capacities of multiple gigabytes are common. These new demands require LANs with high speed that can support the larger numbers and greater geographic extent of office systems as compared to backend systems.

Backbone LANs

The increasing use of distributed processing applications and personal computers has led to a need for a flexible strategy for local networking. Support of premises-wide data communications requires a networking service that is capable of spanning the distances involved and that interconnects equipment in a single (perhaps large) building

[1]A *picture element*, or *pel*, is the smallest discrete scanning-line sample of a facsimile system, which contains only black-white information (no gray scales). A *pixel* is a picture element that contains gray-scale information.

or a cluster of buildings. Although it is possible to develop a single LAN to interconnect all the data processing equipment of a premises, this is probably not a practical alternative in most cases. There are several drawbacks to a single-LAN strategy:

- **Reliability:** With a single LAN, a service interruption, even of short duration, could result in a major disruption for users.
- **Capacity:** A single LAN could be saturated as the number of devices attached to the network grows over time.
- **Cost:** A single LAN technology is not optimized for the diverse requirements of interconnection and communication. The presence of large numbers of low-cost microcomputers dictates that network support for these devices be provided at low cost. LANs that support very-low-cost attachment will not be suitable for meeting the overall requirement.

A more attractive alternative is to employ lower-cost, lower-capacity LANs within buildings or departments and to interconnect these networks with a higher-capacity LAN. This latter network is referred to as a backbone LAN. If confined to a single building or cluster of buildings, a high-capacity LAN can perform the backbone function.

15.2 TOPOLOGIES AND TRANSMISSION MEDIA

The key elements of a LAN are

- Topology
- Transmission medium
- Wiring layout
- Medium access control

Together, these elements determine not only the cost and capacity of the LAN, but also the type of data that may be transmitted, the speed and efficiency of communications, and even the kinds of applications that can be supported.

This section provides a survey of the major technologies in the first two of these categories. It will be seen that there is an interdependence among the choices in different categories. Accordingly, a discussion of pros and cons relative to specific applications is best done by looking at preferred combinations. This, in turn, is best done in the context of standards, which is a subject of a later section.

Topologies

In the context of a communication network, the term *topology* refers to the way in which the end points, or stations, attached to the network are interconnected. The common topologies for LANs are bus, tree, ring, and star (Figure 15.2). The bus is a special case of the tree, with only one trunk and no branches.

Bus and Tree Topologies Both bus and tree topologies are characterized by the use of a multipoint medium. For the **bus**, all stations attach, through appropriate hardware interfacing known as a tap, directly to a linear transmission medium, or bus. Full-duplex operation between the station and the tap allows data to be transmitted onto

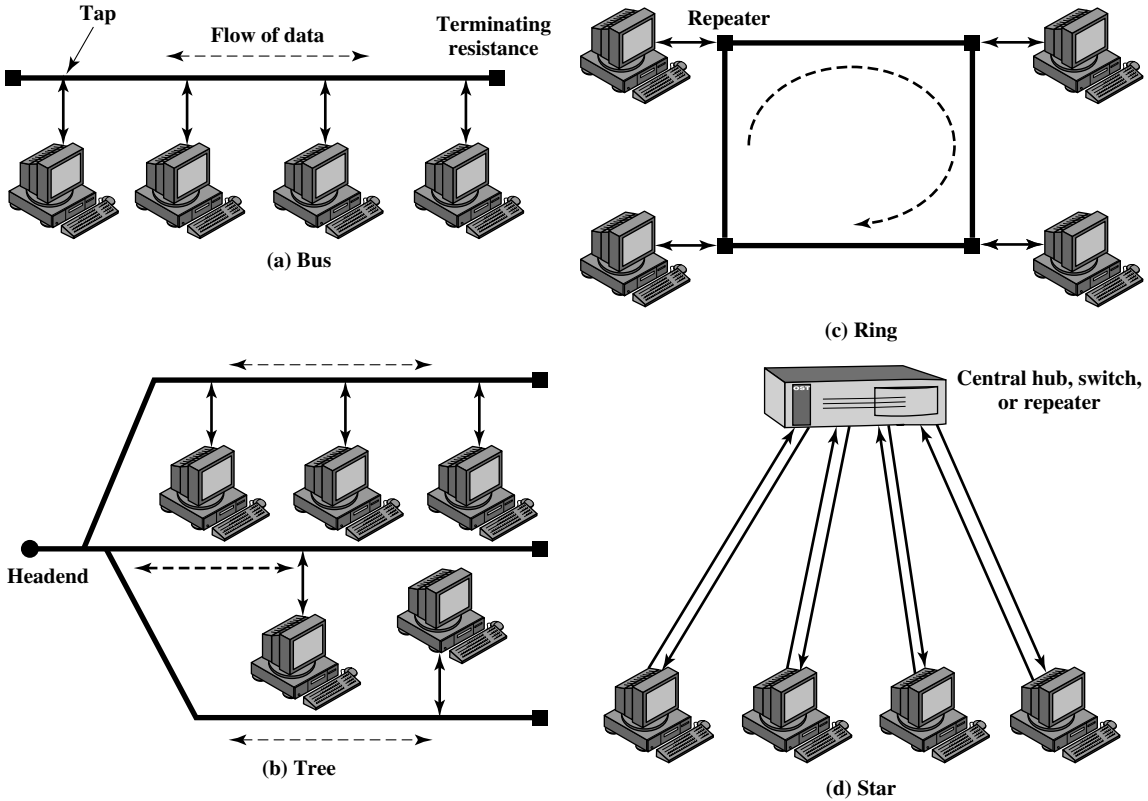

Figure 15.2 LAN Topologies

the bus and received from the bus. A transmission from any station propagates the length of the medium in both directions and can be received by all other stations. At each end of the bus is a terminator, which absorbs any signal, removing it from the bus.

The **tree topology** is a generalization of the bus topology. The transmission medium is a branching cable with no closed loops. The tree layout begins at a point known as the *headend.* One or more cables start at the headend, and each of these may have branches. The branches in turn may have additional branches to allow quite complex layouts. Again, a transmission from any station propagates throughout the medium and can be received by all other stations.

Two problems present themselves in this arrangement. First, because a transmission from any one station can be received by all other stations, there needs to be some way of indicating for whom the transmission is intended. Second, a mechanism is needed to regulate transmission. To see the reason for this, consider that if two stations on the bus attempt to transmit at the same time, their signals will overlap and become garbled. Or consider that one station decides to transmit continuously for a long period of time.

To solve these problems, stations transmit data in small blocks, known as frames. Each frame consists of a portion of the data that a station wishes to transmit, plus a frame header that contains control information. Each station on the bus is assigned a unique address, or identifier, and the destination address for a frame is included in its header.

Figure 15.3 illustrates the scheme. In this example, station C wishes to transmit a frame of data to A. The frame header includes A's address. As the frame propagates along the bus, it passes B. B observes the address and ignores the frame. A, on the other hand, sees that the frame is addressed to itself and therefore copies the data from the frame as it goes by.

So the frame structure solves the first problem mentioned previously: It provides a mechanism for indicating the intended recipient of data. It also provides the basic tool for solving the second problem, the regulation of access. In particular, the stations take turns sending frames in some cooperative fashion. This involves putting additional control information into the frame header, as discussed later.

With the bus or tree, no special action needs to be taken to remove frames from the medium. When a signal reaches the end of the medium, it is absorbed by the terminator.

Ring Topology In the **ring** topology, the network consists of a set of *repeaters* joined by point-to-point links in a closed loop. The repeater is a comparatively simple device, capable of receiving data on one link and transmitting them, bit by bit, on the other link as fast as they are received. The links are unidirectional; that is, data are transmitted in one direction only, so that data circulate around the ring in one direction (clockwise or counterclockwise).

Each station attaches to the network at a repeater and can transmit data onto the network through the repeater. As with the bus and tree, data are transmitted in frames. As a frame circulates past all the other stations, the destination station recognizes its address and copies the frame into a local buffer as it goes by. The frame continues to circulate until it returns to the source station, where it is removed (Figure 15.4). Because multiple stations share the ring, medium access control is needed to determine at what time each station may insert frames.

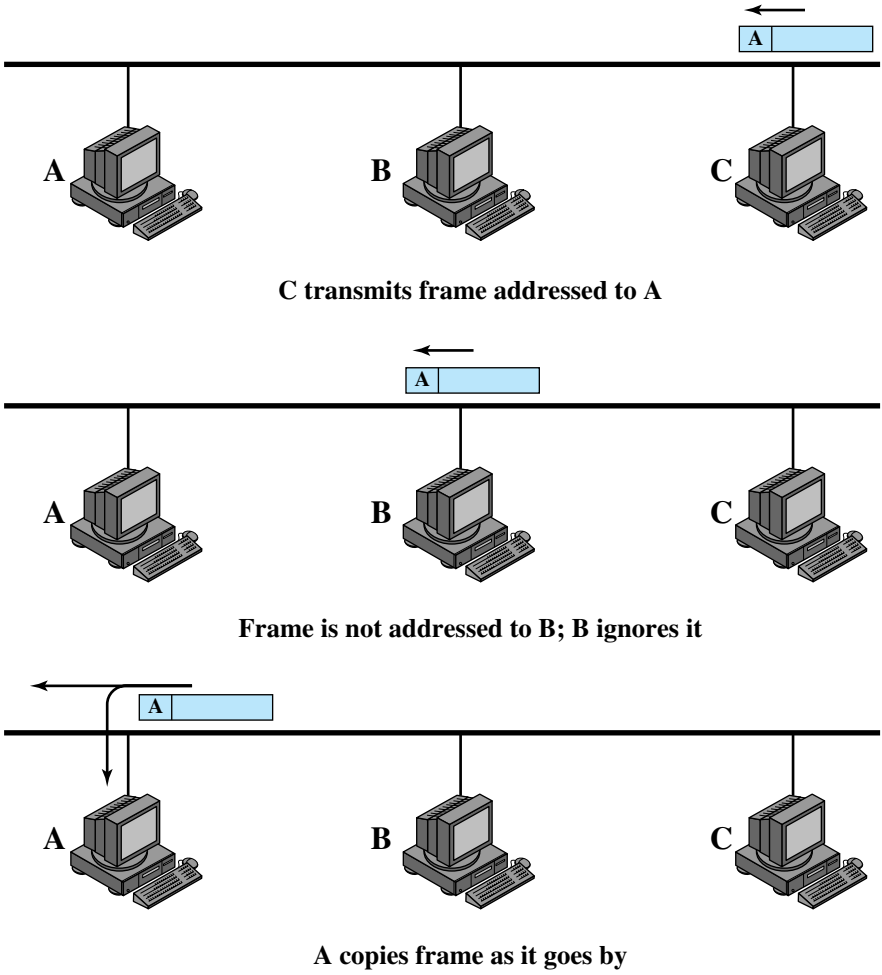

Figure 15.3 Frame Transmission on a Bus LAN

Star Topology In the **star** LAN topology, each station is directly connected to a common central node. Typically, each station attaches to a central node via two point-to-point links, one for transmission and one for reception.

In general, there are two alternatives for the operation of the central node. One approach is for the central node to operate in a broadcast fashion. A transmission of a frame from one station to the node is retransmitted on all of the outgoing links. In this case, although the arrangement is physically a star, it is logically a bus: A transmission from any station is received by all other stations, and only one station at a time may successfully transmit. In this case, the central element is referred to as a **hub**. Another approach is for the central node to act as a frame-switching device. An incoming frame is buffered in the node and then retransmitted on an outgoing link to the destination station. These approaches are explored in Section 15.5.

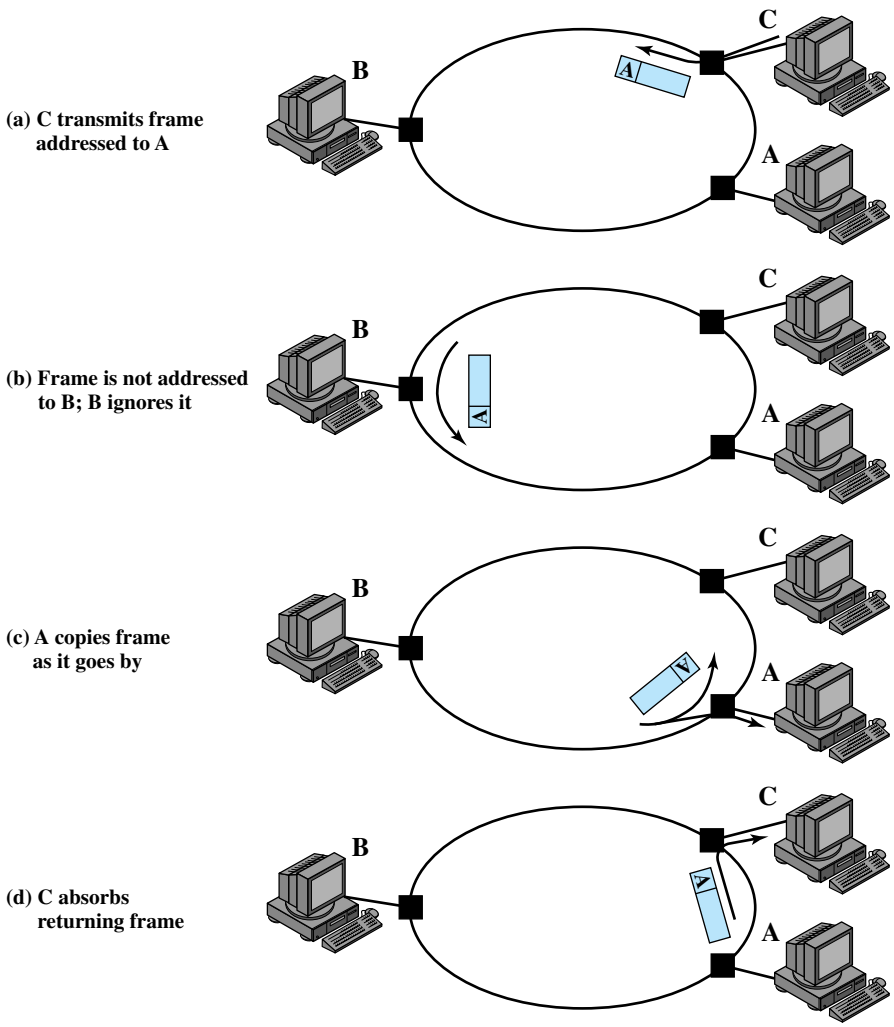

Figure 15.4 Frame Transmission on a Ring LAN

Choice of Topology The choice of topology depends on a variety of factors, including reliability, expandability, and performance. This choice is part of the overall task of designing a LAN and thus cannot be made in isolation, independent of the choice of transmission medium, wiring layout, and access control technique. A few general remarks can be made at this point. There are four alternative media that can be used for a bus LAN:

- **Twisted pair:** In the early days of LAN development, voice-grade twisted pair was used to provide an inexpensive, easily installed bus LAN. A number of systems operating at 1 Mbps were implemented. Scaling twisted pair up to higher data rates in a shared-medium bus configuration is not practical, so this approach was dropped long ago.

- **Baseband coaxial cable:** A baseband coaxial cable is one that makes use of digital signaling. The original Ethernet scheme makes use of baseband coaxial cable.
- **Broadband coaxial cable:** Broadband coaxial cable is the type of cable used in cable television systems. Analog signaling is used at radio and television frequencies. This type of system is more expensive and more difficult to install and maintain than baseband coaxial cable. This approach never achieved popularity and such LANs are no longer made.
- **Optical fiber:** There has been considerable research relating to this alternative over the years, but the expense of the optical fiber taps and the availability of better alternatives have resulted in the demise of this option as well.

Thus, for a bus topology, only baseband coaxial cable has achieved widespread use, primarily for Ethernet systems. Compared to a star-topology twisted pair or optical fiber installation, the bus topology using baseband coaxial cable is difficult to work with. Even simple changes may require access to the coaxial cable, movement of taps, and rerouting of cable segments. Accordingly, few if any new installations are being attempted. Despite its limitations, there is a considerable installed base of baseband coaxial cable bus LANs.

Very-high-speed links over considerable distances can be used for the ring topology. Hence, the ring has the potential of providing the best throughput of any topology. One disadvantage of the ring is that a single link or repeater failure could disable the entire network.

The star topology takes advantage of the natural layout of wiring in a building. It is generally best for short distances and can support a small number of devices at high data rates.

Choice of Transmission Medium The choice of transmission medium is determined by a number of factors. It is, we shall see, constrained by the topology of the LAN. Other factors come into play, including

- **Capacity:** to support the expected network traffic
- **Reliability:** to meet requirements for availability
- **Types of data supported:** tailored to the application
- **Environmental scope:** to provide service over the range of environments required

The choice is part of the overall task of designing a local network, which is addressed in Chapter 16. Here we can make a few general observations.

Voice-grade unshielded twisted pair (UTP) is an inexpensive, well-understood medium; this is the Category 3 UTP referred to in Chapter 4. Typically, office buildings are wired to meet the anticipated telephone system demand plus a healthy margin; thus, there are no cable installation costs in the use of Category 3 UTP. However, the data rate that can be supported is generally quite limited, with the exception of very small LAN. Category 3 UTP is likely to be the most cost-effective for a single-building, low-traffic LAN installation.

Shielded twisted pair and baseband coaxial cable are more expensive than Category 3 UTP but provide greater capacity. Broadband cable is even more expensive but provides even greater capacity. However, in recent years, the trend has been

toward the use of high-performance UTP, especially Category 5 UTP. Category 5 UTP supports high data rates for a small number of devices, but larger installations can be supported by the use of the star topology and the interconnection of the switching elements in multiple star-topology configurations. We discuss this point in Chapter 16.

Optical fiber has a number of attractive features, such as electromagnetic isolation, high capacity, and small size, which have attracted a great deal of interest. As yet the market penetration of fiber LANs is low; this is primarily due to the high cost of fiber components and the lack of skilled personnel to install and maintain fiber systems. This situation is beginning to change rapidly as more products using fiber are introduced.

15.3 LAN PROTOCOL ARCHITECTURE

The architecture of a LAN is best described in terms of a layering of protocols that organize the basic functions of a LAN. This section opens with a description of the standardized protocol architecture for LANs, which encompasses physical, medium access control (MAC), and logical link control (LLC) layers. The physical layer encompasses topology and transmission medium, and is covered in Section 15.2. This section provides an overview of the MAC and LLC layers.

IEEE 802 Reference Model

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

Figure 15.5 relates the LAN protocols to the OSI architecture (Figure 2.11). This architecture was developed by the IEEE 802 LAN standards committee[2] and has been adopted by all organizations working on the specification of LAN standards. It is generally referred to as the IEEE 802 reference model.

Working from the bottom up, the lowest layer of the IEEE 802 reference model corresponds to the **physical layer** of the OSI model and includes such functions as

- Encoding/decoding of signals
- Preamble generation/removal (for synchronization)
- Bit transmission/reception

In addition, the physical layer of the 802 model includes a specification of the transmission medium and the topology. Generally, this is considered "below" the lowest layer of the OSI model. However, the choice of transmission medium and topology is critical in LAN design, and so a specification of the medium is included.

Above the physical layer are the functions associated with providing service to LAN users. These include

[2]This committee has developed standards for a wide range of LANs. See Appendix D for details.

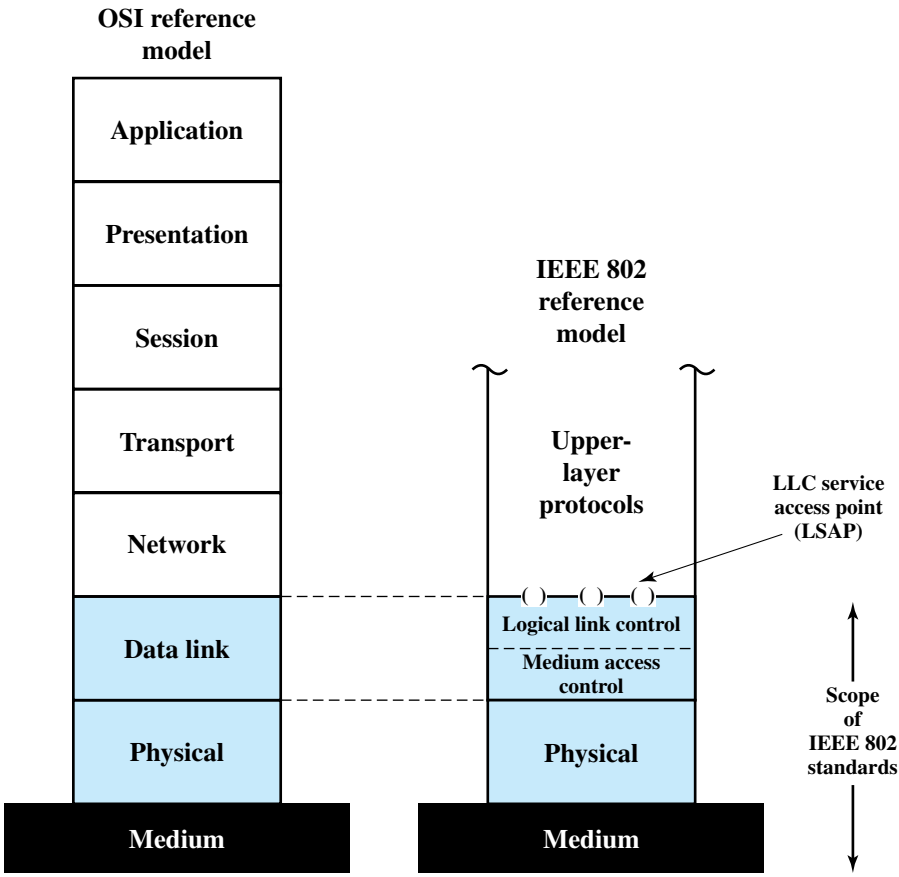

Figure 15.5 IEEE 802 Protocol Layers Compared to OSI Model

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

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

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

Figure 15.6 illustrates the relationship between the levels of the architecture (compare Figure 2.9). Higher-level data are passed down to LLC, which appends

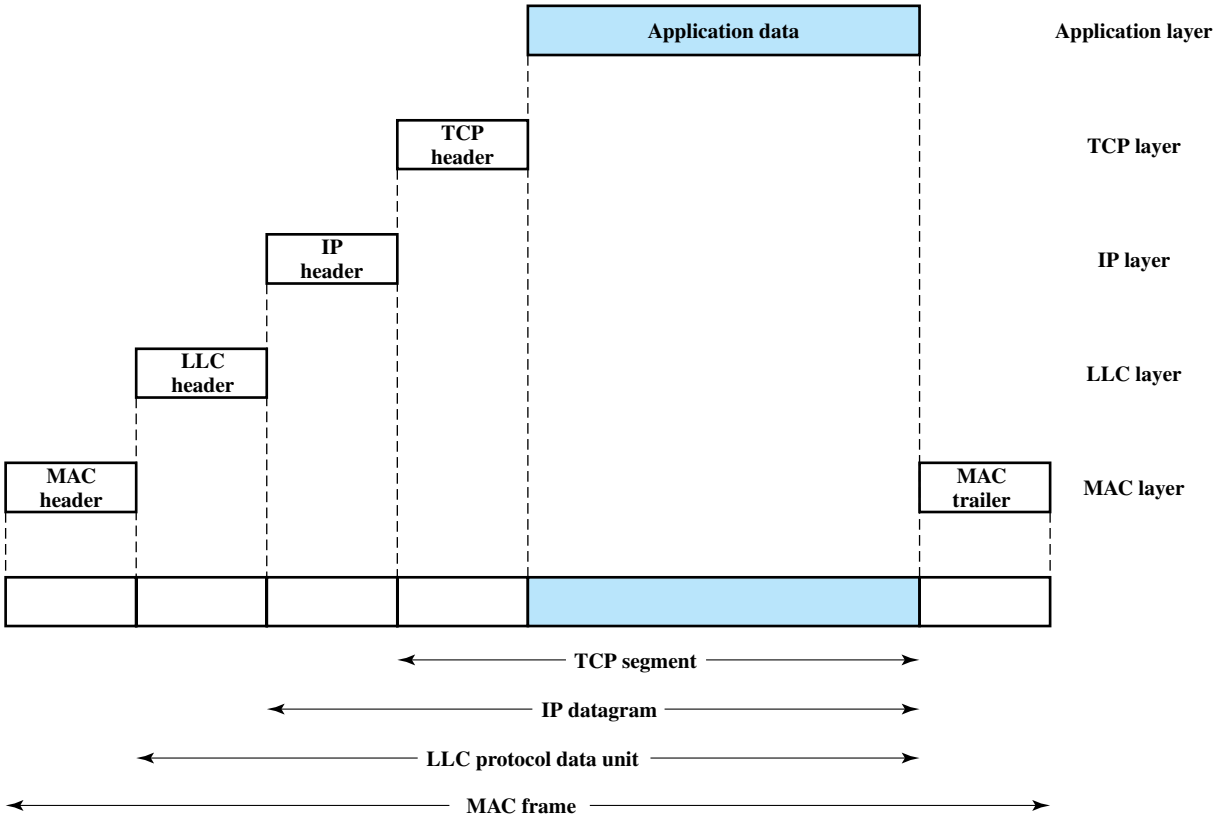

Figure 15.6 LAN Protocols in Context

control information as a header, creating an *LLC protocol data unit (PDU)*. This control information is used in the operation of the LLC protocol. The entire LLC PDU is then passed down to the MAC layer, which appends control information at the front and back of the packet, forming a *MAC frame*. Again, the control information in the frame is needed for the operation of the MAC protocol. For context, the figure also shows the use of TCP/IP and an application layer above the LAN protocols.

Logical Link Control

The LLC layer for LANs is similar in many respects to other link layers in common use. Like all link layers, LLC is concerned with the transmission of a link-level PDU between two stations, without the necessity of an intermediate switching node. LLC has two characteristics not shared by most other link control protocols:

1. It must support the multiaccess, shared-medium nature of the link (this differs from a multidrop line in that there is no primary node).
2. It is relieved of some details of link access by the MAC layer.

Addressing in LLC involves specifying the source and destination LLC users. Typically, a user is a higher-layer protocol or a network management function in the station. These LLC user addresses are referred to as service access points (SAPs), in keeping with OSI terminology for the user of a protocol layer.

We look first at the services that LLC provides to a higher-level user, and then at the LLC protocol.

LLC Services LLC specifies the mechanisms for addressing stations across the medium and for controlling the exchange of data between two users. The operation and format of this standard is based on HDLC. Three services are provided as alternatives for attached devices using LLC:

- **Unacknowledged connectionless service:** This service is a datagram-style service. It is a very simple service that does not involve any of the flow- and error-control mechanisms. Thus, the delivery of data is not guaranteed. However, in most devices, there will be some higher layer of software that deals with reliability issues.
- **Connection-mode service:** This service is similar to that offered by HDLC. A logical connection is set up between two users exchanging data, and flow control and error control are provided.
- **Acknowledged connectionless service:** This is a cross between the previous two services. It provides that datagrams are to be acknowledged, but no prior logical connection is set up.

Typically, a vendor will provide these services as options that the customer can select when purchasing the equipment. Alternatively, the customer can purchase equipment that provides two or all three services and select a specific service based on application.

The **unacknowledged connectionless service** requires minimum logic and is useful in two contexts. First, it will often be the case that higher layers of software will provide the necessary reliability and flow-control mechanism, and it is efficient

to avoid duplicating them. For example, TCP could provide the mechanisms needed to ensure that data is delivered reliably. Second, there are instances in which the overhead of connection establishment and maintenance is unjustified or even counterproductive (for example, data collection activities that involve the periodic sampling of data sources, such as sensors and automatic self-test reports from security equipment or network components). In a monitoring application, the loss of an occasional data unit would not cause distress, as the next report should arrive shortly. Thus, in most cases, the unacknowledged connectionless service is the preferred option.

The **connection-mode service** could be used in very simple devices, such as terminal controllers, that have little software operating above this level. In these cases, it would provide the flow control and reliability mechanisms normally implemented at higher layers of the communications software.

The **acknowledged connectionless service** is useful in several contexts. With the connection-mode service, the logical link control software must maintain some sort of table for each active connection, to keep track of the status of that connection. If the user needs guaranteed delivery but there are a large number of destinations for data, then the connection-mode service may be impractical because of the large number of tables required. An example is a process control or automated factory environment where a central site may need to communicate with a large number of processors and programmable controllers. Another use of this is the handling of important and time-critical alarm or emergency control signals in a factory. Because of their importance, an acknowledgment is needed so that the sender can be assured that the signal got through. Because of the urgency of the signal, the user might not want to take the time first to establish a logical connection and then send the data.

LLC Protocol The basic LLC protocol is modeled after HDLC and has similar functions and formats. The differences between the two protocols can be summarized as follows:

- LLC makes use of the asynchronous balanced mode of operation of HDLC, to support connection-mode LLC service; this is referred to as type 2 operation. The other HDLC modes are not employed.
- LLC supports an unacknowledged connectionless service using the unnumbered information PDU; this is known as type 1 operation.
- LLC supports an acknowledged connectionless service by using two new unnumbered PDUs; this is known as type 3 operation.
- LLC permits multiplexing by the use of LLC service access points (LSAPs).

All three LLC protocols employ the same PDU format (Figure 15.7), which consists of four fields. The DSAP (Destination Service Access Point) and SSAP (Source Service Access Point) fields each contain a 7-bit address, which specify the destination and source users of LLC. One bit of the DSAP indicates whether the DSAP is an individual or group address. One bit of the SSAP indicates whether the PDU is a command or response PDU. The format of the LLC control field is identical to that of HDLC (Figure 7.7), using extended (7-bit) sequence numbers.

For **type 1 operation**, which supports the unacknowledged connectionless service, the unnumbered information (UI) PDU is used to transfer user data. There is

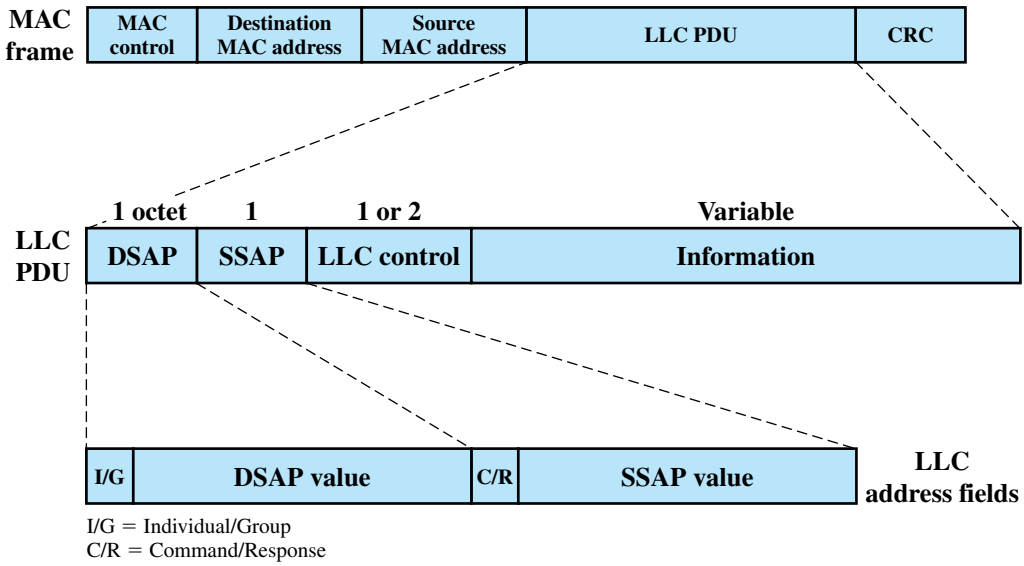

Figure 15.7 LLC PDU in a Generic MAC Frame Format

no acknowledgment, flow control, or error control. However, there is error detection and discard at the MAC level.

Two other PDUs are used to support management functions associated with all three types of operation. Both PDUs are used in the following fashion. An LLC entity may issue a command (C/R bit = 0) XID or TEST. The receiving LLC entity issues a corresponding XID or TEST in response. The XID PDU is used to exchange two types of information: types of operation supported and window size. The TEST PDU is used to conduct a loopback test of the transmission path between two LLC entities. Upon receipt of a TEST command PDU, the addressed LLC entity issues a TEST response PDU as soon as possible.

With **type 2 operation**, a data link connection is established between two LLC SAPs prior to data exchange. Connection establishment is attempted by the type 2 protocol in response to a request from a user. The LLC entity issues a SABME PDU[3] to request a logical connection with the other LLC entity. If the connection is accepted by the LLC user designated by the DSAP, then the destination LLC entity returns an unnumbered acknowledgment (UA) PDU. The connection is henceforth uniquely identified by the pair of user SAPs. If the destination LLC user rejects the connection request, its LLC entity returns a disconnected mode (DM) PDU.

Once the connection is established, data are exchanged using information PDUs, as in HDLC. The information PDUs include send and receive sequence numbers, for sequencing and flow control. The supervisory PDUs are used, as in HDLC,

[3]This stands for Set Asynchronous Balanced Mode Extended. It is used in HDLC to choose ABM and to select extended sequence numbers of seven bits. Both ABM and 7-bit sequence numbers are mandatory in type 2 operation.

for flow control and error control. Either LLC entity can terminate a logical LLC connection by issuing a disconnect (DISC) PDU.

With **type 3 operation**, each transmitted PDU is acknowledged. A new (not found in HDLC) unnumbered PDU, the Acknowledged Connectionless (AC) Information PDU, is defined. User data are sent in AC command PDUs and must be acknowledged using an AC response PDU. To guard against lost PDUs, a 1-bit sequence number is used. The sender alternates the use of 0 and 1 in its AC command PDU, and the receiver responds with an AC PDU with the opposite number of the corresponding command. Only one PDU in each direction may be outstanding at any time.

Medium Access Control

All LANs and MANs consist of collections of devices that must share the network's transmission capacity. Some means of controlling access to the transmission medium is needed to provide for an orderly and efficient use of that capacity. This is the function of a medium access control (MAC) protocol.

The key parameters in any medium access control technique are where and how. *Where* refers to whether control is exercised in a centralized or distributed fashion. In a centralized scheme, a controller is designated that has the authority to grant access to the network. A station wishing to transmit must wait until it receives permission from the controller. In a decentralized network, the stations collectively perform a medium access control function to determine dynamically the order in which stations transmit. A centralized scheme has certain advantages, including

- It may afford greater control over access for providing such things as priorities, overrides, and guaranteed capacity.
- It enables the use of relatively simple access logic at each station.
- It avoids problems of distributed coordination among peer entities.

The principal disadvantages of centralized schemes are

- It creates a single point of failure; that is, there is a point in the network that, if it fails, causes the entire network to fail.
- It may act as a bottleneck, reducing performance.

The pros and cons of distributed schemes are mirror images of the points just made.

The second parameter, *how*, is constrained by the topology and is a tradeoff among competing factors, including cost, performance, and complexity. In general, we can categorize access control techniques as being either synchronous or asynchronous. With synchronous techniques, a specific capacity is dedicated to a connection. This is the same approach used in circuit switching, frequency division multiplexing (FDM), and synchronous time division multiplexing (TDM). Such techniques are generally not optimal in LANs and MANs because the needs of the stations are unpredictable. It is preferable to be able to allocate capacity in an asynchronous (dynamic) fashion, more or less in response to immediate demand. The asynchronous approach can be further subdivided into three categories: round robin, reservation, and contention.

Round Robin With round robin, each station in turn is given the opportunity to transmit. During that opportunity, the station may decline to transmit or may transmit subject to a specified upper bound, usually expressed as a maximum amount of data transmitted or time for this opportunity. In any case, the station, when it is finished, relinquishes its turn, and the right to transmit passes to the next station in logical sequence. Control of sequence may be centralized or distributed. Polling is an example of a centralized technique.

When many stations have data to transmit over an extended period of time, round-robin techniques can be very efficient. If only a few stations have data to transmit over an extended period of time, then there is a considerable overhead in passing the turn from station to station, because most of the stations will not transmit but simply pass their turns. Under such circumstances other techniques may be preferable, largely depending on whether the data traffic has a stream or bursty characteristic. Stream traffic is characterized by lengthy and fairly continuous transmissions; examples are voice communication, telemetry, and bulk file transfer. Bursty traffic is characterized by short, sporadic transmissions; interactive terminal-host traffic fits this description.

Reservation For stream traffic, reservation techniques are well suited. In general, for these techniques, time on the medium is divided into slots, much as with synchronous TDM. A station wishing to transmit reserves future slots for an extended or even an indefinite period. Again, reservations may be made in a centralized or distributed fashion.

Contention For bursty traffic, contention techniques are usually appropriate. With these techniques, no control is exercised to determine whose turn it is; all stations contend for time in a way that can be, as we shall see, rather rough and tumble. These techniques are of necessity distributed in nature. Their principal advantage is that they are simple to implement and, under light to moderate load, efficient. For some of these techniques, however, performance tends to collapse under heavy load.

Although both centralized and distributed reservation techniques have been implemented in some LAN products, round-robin and contention techniques are the most common.

MAC Frame Format The MAC layer receives a block of data from the LLC layer and is responsible for performing functions related to medium access and for transmitting the data. As with other protocol layers, MAC implements these functions making use of a protocol data unit at its layer. In this case, the PDU is referred to as a MAC frame.

The exact format of the MAC frame differs somewhat for the various MAC protocols in use. In general, all of the MAC frames have a format similar to that of Figure 15.7. The fields of this frame are

- **MAC Control:** This field contains any protocol control information needed for the functioning of the MAC protocol. For example, a priority level could be indicated here.
- **Destination MAC Address:** The destination physical attachment point on the LAN for this frame.

- **Source MAC Address:** The source physical attachment point on the LAN for this frame.
- **LLC:** The LLC data from the next higher layer.
- **CRC:** The Cyclic Redundancy Check field (also known as the frame check sequence, FCS, field). This is an error-detecting code, as we have seen in HDLC and other data link control protocols (Chapter 7).

In most data link control protocols, the data link protocol entity is responsible not only for detecting errors using the CRC, but for recovering from those errors by retransmitting damaged frames. In the LAN protocol architecture, these two functions are split between the MAC and LLC layers. The MAC layer is responsible for detecting errors and discarding any frames that are in error. The LLC layer optionally keeps track of which frames have been successfully received and retransmits unsuccessful frames.

15.4 BRIDGES

In virtually all cases, there is a need to expand beyond the confines of a single LAN, to provide interconnection to other LANs and to wide area networks. Two general approaches are used for this purpose: bridges and routers. The bridge is the simpler of the two devices and provides a means of interconnecting similar LANs. The router is a more general-purpose device, capable of interconnecting a variety of LANs and WANs. We explore bridges in this section and look at routers in Part Five.

The bridge is designed for use between local area networks (LANs) that use identical protocols for the physical and link layers (e.g., all conforming to IEEE 802.3). Because the devices all use the same protocols, the amount of processing required at the bridge is minimal. More sophisticated bridges are capable of mapping from one MAC format to another (e.g., to interconnect an Ethernet and a token ring LAN).

Because the bridge is used in a situation in which all the LANs have the same characteristics, the reader may ask, why not simply have one large LAN? Depending on circumstance, there are several reasons for the use of multiple LANs connected by bridges:

- **Reliability:** The danger in connecting all data processing devices in an organization to one network is that a fault on the network may disable communication for all devices. By using bridges, the network can be partitioned into self-contained units.
- **Performance:** In general, performance on a LAN declines with an increase in the number of devices or the length of the wire. A number of smaller LANs will often give improved performance if devices can be clustered so that intranetwork traffic significantly exceeds internetwork traffic.
- **Security:** The establishment of multiple LANs may improve security of communications. It is desirable to keep different types of traffic (e.g., accounting,

personnel, strategic planning) that have different security needs on physically separate media. At the same time, the different types of users with different levels of security need to communicate through controlled and monitored mechanisms.

- **Geography:** Clearly, two separate LANs are needed to support devices clustered in two geographically distant locations. Even in the case of two buildings separated by a highway, it may be far easier to use a microwave bridge link than to attempt to string coaxial cable between the two buildings.

Functions of a Bridge

Figure 15.8 illustrates the action of a bridge connecting two LANs, A and B, using the same MAC protocol. In this example, a single bridge attaches to both LANs; frequently, the bridge function is performed by two "half-bridges," one on each LAN. The functions of the bridge are few and simple:

- Read all frames transmitted on A and accept those addressed to any station on B.
- Using the medium access control protocol for B, retransmit each frame on B.
- Do the same for B-to-A traffic.

Several design aspects of a bridge are worth highlighting:

- The bridge makes no modification to the content or format of the frames it receives, nor does it encapsulate them with an additional header. Each frame to be transferred is simply copied from one LAN and repeated with exactly the same bit pattern on the other LAN. Because the two LANs use the same LAN protocols, it is permissible to do this.
- The bridge should contain enough buffer space to meet peak demands. Over a short period of time, frames may arrive faster than they can be retransmitted.
- The bridge must contain addressing and routing intelligence. At a minimum, the bridge must know which addresses are on each network to know which frames to pass. Further, there may be more than two LANs interconnected by a number of bridges. In that case, a frame may have to be routed through several bridges in its journey from source to destination.
- A bridge may connect more than two LANs.

In summary, the bridge provides an extension to the LAN that requires no modification to the communications software in the stations attached to the LANs. It appears to all stations on the two (or more) LANs that there is a single LAN on which each station has a unique address. The station uses that unique address and need not explicitly discriminate between stations on the same LAN and stations on other LANs; the bridge takes care of that.

Bridge Protocol Architecture

The IEEE 802.1D specification defines the protocol architecture for MAC bridges. Within the 802 architecture, the endpoint or station address is designated at the

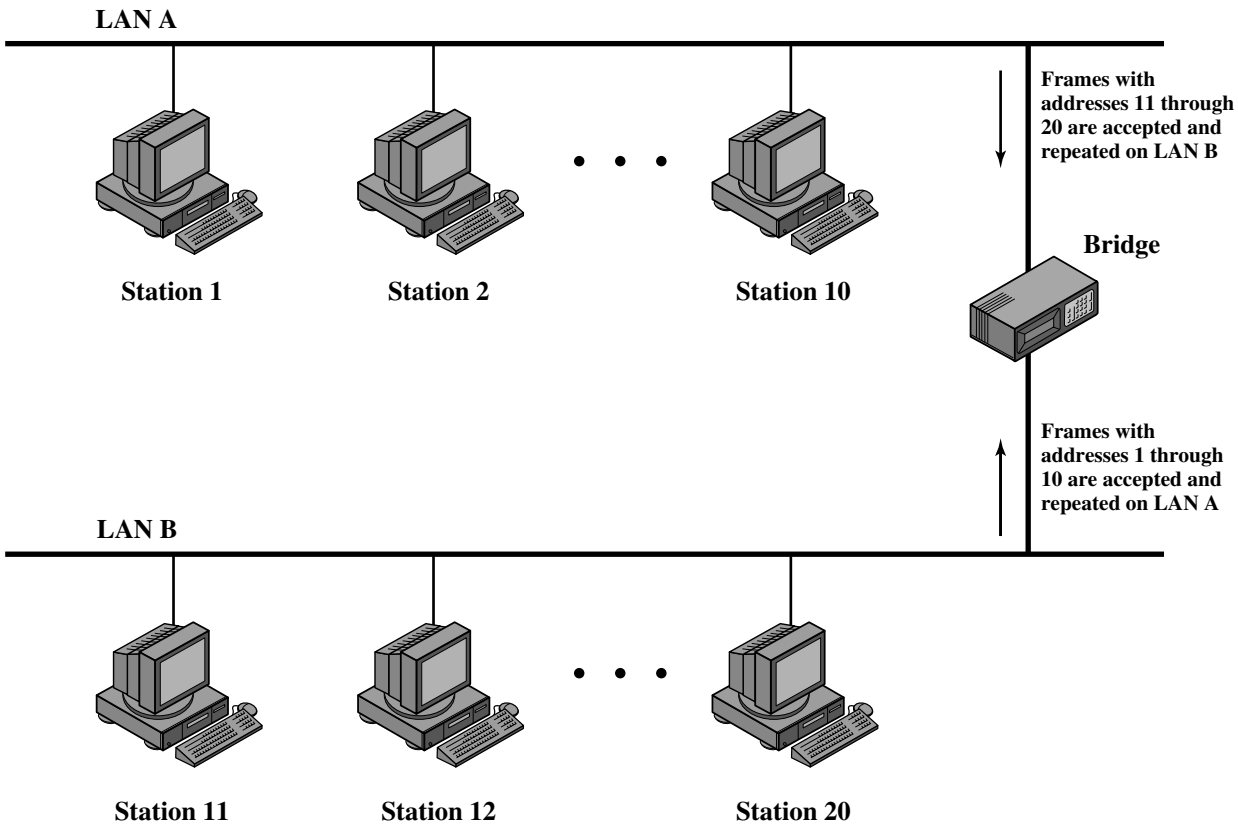

Figure 15.8 Bridge Operation

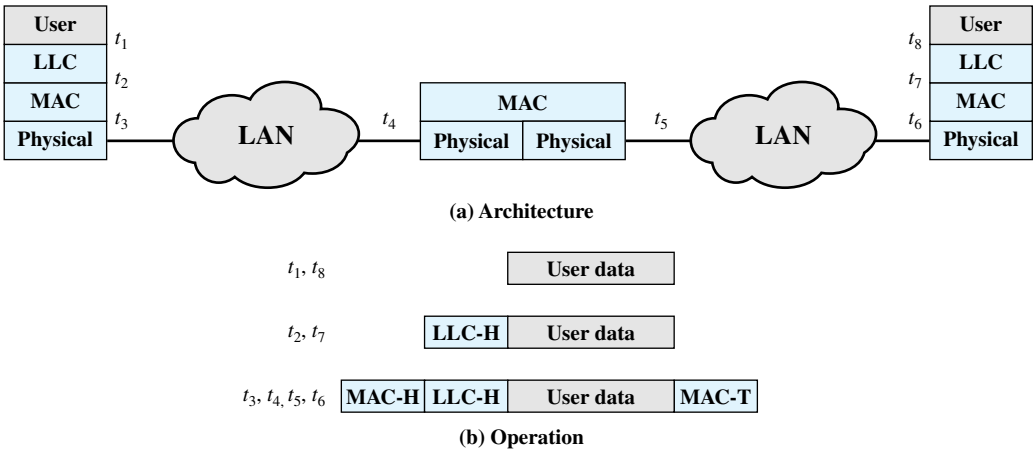

Figure 15.9 Connection of Two LANs by a Bridge

MAC level. Thus, it is at the MAC level that a bridge can function. Figure 15.9 shows the simplest case, which consists of two LANs connected by a single bridge. The LANs employ the same MAC and LLC protocols. The bridge operates as previously described. A MAC frame whose destination is not on the immediate LAN is captured by the bridge, buffered briefly, and then transmitted on the other LAN. As far as the LLC layer is concerned, there is a dialogue between peer LLC entities in the two endpoint stations. The bridge need not contain an LLC layer because it is merely serving to relay the MAC frames.

Figure 15.9b indicates the way in which data are encapsulated using a bridge. Data are provided by some user to LLC. The LLC entity appends a header and passes the resulting data unit to the MAC entity, which appends a header and a trailer to form a MAC frame. On the basis of the destination MAC address in the frame, it is captured by the bridge. The bridge does not strip off the MAC fields; its function is to relay the MAC frame intact to the destination LAN. Thus, the frame is deposited on the destination LAN and captured by the destination station.

The concept of a MAC relay bridge is not limited to the use of a single bridge to connect two nearby LANs. If the LANs are some distance apart, then they can be connected by two bridges that are in turn connected by a communications facility. The intervening communications facility can be a network, such as a wide area packet-switching network, or a point-to-point link. In such cases, when a bridge captures a MAC frame, it must encapsulate the frame in the appropriate packaging and transmit it over the communications facility to a target bridge. The target bridge strips off these extra fields and transmits the original, unmodified MAC frame to the destination station.

Fixed Routing

There is a trend within many organizations to an increasing number of LANs interconnected by bridges. As the number of LANs grows, it becomes important to

provide alternate paths between LANs via bridges for load balancing and reconfiguration in response to failure. Thus, many organizations will find that static, preconfigured routing tables are inadequate and that some sort of dynamic routing is needed.

Consider the configuration of Figure 15.10. Suppose that station 1 transmits a frame on LAN A intended for station 6. The frame will be read by bridges 101, 102, and 107. For each bridge, the addressed station is not on a LAN to which the bridge is attached. Therefore, each bridge must make a decision whether or not to retransmit the frame on its other LAN, in order to move it closer to its intended destination. In this case, bridge 102 should repeat the frame on LAN C, whereas bridges 101 and 107 should refrain from retransmitting the frame. Once the frame has been transmitted on LAN C, it will be picked up by both bridges 105 and 106. Again, each must decide whether or not to forward the frame. In this case, bridge 105 should retransmit the frame on LAN F, where it will be received by the destination, station 6.

Thus we see that, in the general case, the bridge must be equipped with a routing capability. When a bridge receives a frame, it must decide whether or not to

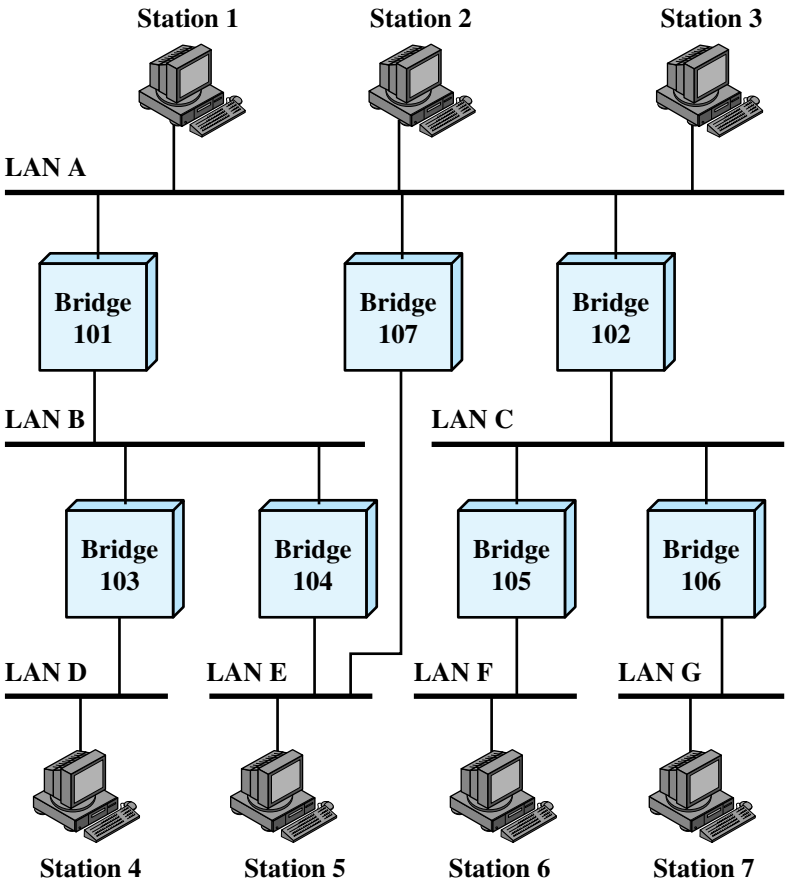

Figure 15.10 Configuration of Bridges and LANs, with Alternate Routes

forward it. If the bridge is attached to two or more networks, then it must decide whether or not to forward the frame and, if so, on which LAN the frame should be transmitted.

The routing decision may not always be a simple one. Figure 15.10 also shows that there are two routes between LAN A and LAN E. Such redundancy provides for higher overall Internet availability and creates the possibility for load balancing. In this case, if station 1 transmits a frame on LAN A intended for station 5 on LAN E, then either bridge 101 or bridge 107 could forward the frame. It would appear preferable for bridge 107 to forward the frame, since it will involve only one hop, whereas if the frame travels through bridge 101, it must suffer two hops. Another consideration is that there may be changes in the configuration. For example, bridge 107 may fail, in which case subsequent frames from station 1 to station 5 should go through bridge 101. So we can say that the routing capability must take into account the topology of the internet configuration and may need to be dynamically altered.

A variety of routing strategies have been proposed and implemented in recent years. The simplest and most common strategy is **fixed routing**. This strategy is suitable for small internets and for internets that are relatively stable. In addition, two groups within the IEEE 802 committee have developed specifications for routing strategies. The IEEE 802.1 group has issued a standard for routing based on the use of a **spanning tree** algorithm. The token ring committee, IEEE 802.5, has issued its own specification, referred to as **source routing**. In the remainder of this section, we look at fixed routing and the spanning tree algorithm, which is the most commonly used bridge routing algorithm.

For fixed routing, a route is selected for each source-destination pair of LANs in the configuration. If alternate routes are available between two LANs, then typically the route with the least number of hops is selected. The routes are fixed, or at least only change when there is a change in the topology of the internet.

The strategy for developing a fixed routing configuration for bridges is similar to that employed in a packet-switching network (Figure 12.2). A central routing matrix is created, to be stored perhaps at a network control center. The matrix shows, for each source-destination pair of LANs, the identity of the first bridge on the route. So, for example, the route from LAN E to LAN F begins by going through bridge 107 to LAN A. Again consulting the matrix, the route from LAN A to LAN F goes through bridge 102 to LAN C. Finally, the route from LAN C to LAN F is directly through bridge 105. Thus the complete route from LAN E to LAN F is bridge 107, LAN A, bridge 102, LAN C, bridge 105.

From this overall matrix, routing tables can be developed and stored at each bridge. Each bridge needs one table for each LAN to which it attaches. The information for each table is derived from a single row of the matrix. For example, bridge 105 has two tables, one for frames arriving from LAN C and one for frames arriving from LAN F. The table shows, for each possible destination MAC address, the identity of the LAN to which the bridge should forward the frame.

Once the directories have been established, routing is a simple matter. A bridge copies each incoming frame on each of its LANs. If the destination MAC address corresponds to an entry in its routing table, the frame is retransmitted on the appropriate LAN.

The fixed routing strategy is widely used in commercially available products. It requires that a network manager manually load the data into the routing tables. It has the advantage of simplicity and minimal processing requirements. However, in a complex internet, in which bridges may be dynamically added and in which failures must be allowed for, this strategy is too limited.

The Spanning Tree Approach

The spanning tree approach is a mechanism in which bridges automatically develop a routing table and update that table in response to changing topology. The algorithm consists of three mechanisms: frame forwarding, address learning, and loop resolution.

Frame Forwarding In this scheme, a bridge maintains a **forwarding database** for each port attached to a LAN. The database indicates the station addresses for which frames should be forwarded through that port. We can interpret this in the following fashion. For each port, a list of stations is maintained. A station is on the list if it is on the "same side" of the bridge as the port. For example, for bridge 102 of Figure 15.10, stations on LANs C, F, and G are on the same side of the bridge as the LAN C port, and stations on LANs A, B, D, and E are on the same side of the bridge as the LAN A port. When a frame is received on any port, the bridge must decide whether that frame is to be forwarded through the bridge and out through one of the bridge's other ports. Suppose that a bridge receives a MAC frame on port *x*. The following rules are applied:

1. Search the forwarding database to determine if the MAC address is listed for any port except port *x*.
2. If the destination MAC address is not found, forward frame out all ports except the one from which is was received. This is part of the learning process described subsequently.
3. If the destination address is in the forwarding database for some port *y,* then determine whether port *y* is in a blocking or forwarding state. For reasons explained later, a port may sometimes be blocked, which prevents it from receiving or transmitting frames.
4. If port *y* is not blocked, transmit the frame through port *y* onto the LAN to which that port attaches.

Address Learning The preceding scheme assumes that the bridge is already equipped with a forwarding database that indicates the direction, from the bridge, of each destination station. This information can be preloaded into the bridge, as in fixed routing. However, an effective automatic mechanism for learning the direction of each station is desirable. A simple scheme for acquiring this information is based on the use of the source address field in each MAC frame.

The strategy is this. When a frame arrives on a particular port, it clearly has come from the direction of the incoming LAN. The source address field of the frame indicates the source station. Thus, a bridge can update its forwarding database for that port on the basis of the source address field of each incoming frame. To allow for changes in topology, each element in the database is equipped with a

timer. When a new element is added to the database, its timer is set. If the timer expires, then the element is eliminated from the database, since the corresponding direction information may no longer be valid. Each time a frame is received, its source address is checked against the database. If the element is already in the database, the entry is updated (the direction may have changed) and the timer is reset. If the element is not in the database, a new entry is created, with its own timer.

Spanning Tree Algorithm The address learning mechanism described previously is effective if the topology of the internet is a tree; that is, if there are no alternate routes in the network. The existence of alternate routes means that there is a closed loop. For example in Figure 15.10, the following is a closed loop: LAN A, bridge 101, LAN B, bridge 104, LAN E, bridge 107, LAN A.

To see the problem created by a closed loop, consider Figure 15.11. At time t_0, station A transmits a frame addressed to station B. The frame is captured by both bridges. Each bridge updates its database to indicate that station A is in the direction of LAN X, and retransmits the frame on LAN Y. Say that bridge α retransmits at time t_1 and bridge β a short time later t_2. Thus B will receive two copies of the frame. Furthermore, each bridge will receive the other's transmission on LAN Y. Note that each transmission is a frame with a source address of A and a destination address of B. Thus each bridge will update its database to indicate that station A is in

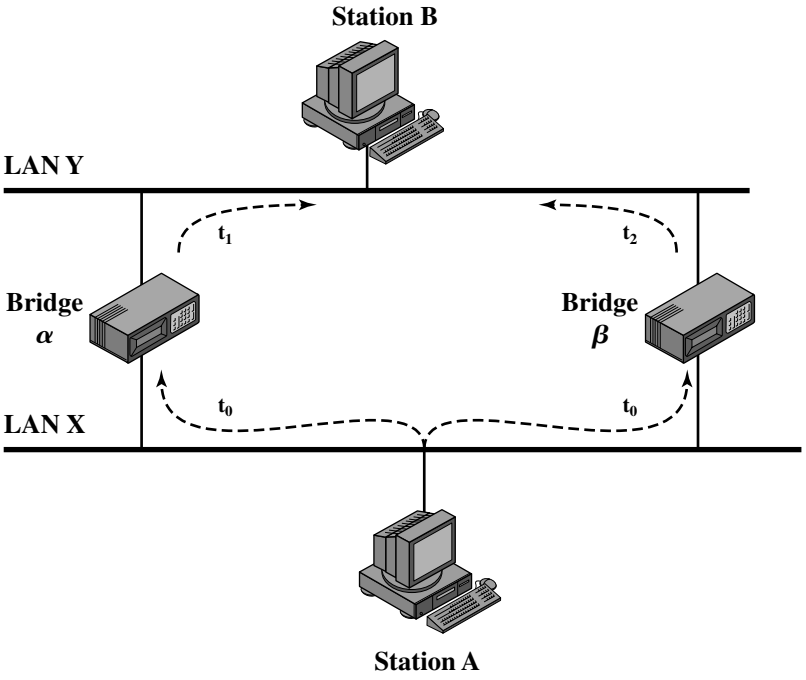

Figure 15.11 Loop of Bridges

the direction of LAN Y. Neither bridge is now capable of forwarding a frame addressed to station A.

To overcome this problem, a simple result from graph theory is used: For any connected graph, consisting of nodes and edges connecting pairs of nodes, there is a spanning tree of edges that maintains the connectivity of the graph but contains no closed loops. In terms of internets, each LAN corresponds to a graph node, and each bridge corresponds to a graph edge. Thus, in Figure 15.10, the removal of one (and only one) of bridges 107, 101, and 104, results in a spanning tree. What is desired is to develop a simple algorithm by which the bridges of the internet can exchange sufficient information to automatically (without user intervention) derive a spanning tree. The algorithm must be dynamic. That is, when a topology change occurs, the bridges must be able to discover this fact and automatically derive a new spanning tree.

The spanning tree algorithm developed by IEEE 802.1, as the name suggests, is able to develop such a spanning tree. All that is required is that each bridge be assigned a unique identifier and that costs be assigned to each bridge port. In the absence of any special considerations, all costs could be set equal; this produces a minimum-hop tree. The algorithm involves a brief exchange of messages among all of the bridges to discover the minimum-cost spanning tree. Whenever there is a change in topology, the bridges automatically recalculate the spanning tree.

15.5 LAYER 2 AND LAYER 3 SWITCHES

In recent years, there has been a proliferation of types of devices for interconnecting LANs that goes beyond the bridges discussed in Section 15.4 and the routers discussed in Part Five. These devices can conveniently be grouped into the categories of layer 2 switches and layer 3 switches. We begin with a discussion of hubs and then explore these two concepts.

Hubs

Earlier, we used the term *hub* in reference to a star-topology LAN. The hub is the active central element of the star layout. Each station is connected to the hub by two lines (transmit and receive). The hub acts as a repeater: When a single station transmits, the hub repeats the signal on the outgoing line to each station. Ordinarily, the line consists of two unshielded twisted pairs. Because of the high data rate and the poor transmission qualities of unshielded twisted pair, the length of a line is limited to about 100 m. As an alternative, an optical fiber link may be used. In this case, the maximum length is about 500 m.

Note that although this scheme is physically a star, it is logically a bus: A transmission from any one station is received by all other stations, and if two stations transmit at the same time there will be a collision.

Multiple levels of hubs can be cascaded in a hierarchical configuration. Figure 15.12 illustrates a two-level configuration. There is one **header hub** (HHUB) and one or more **intermediate hubs** (IHUB). Each hub may have a

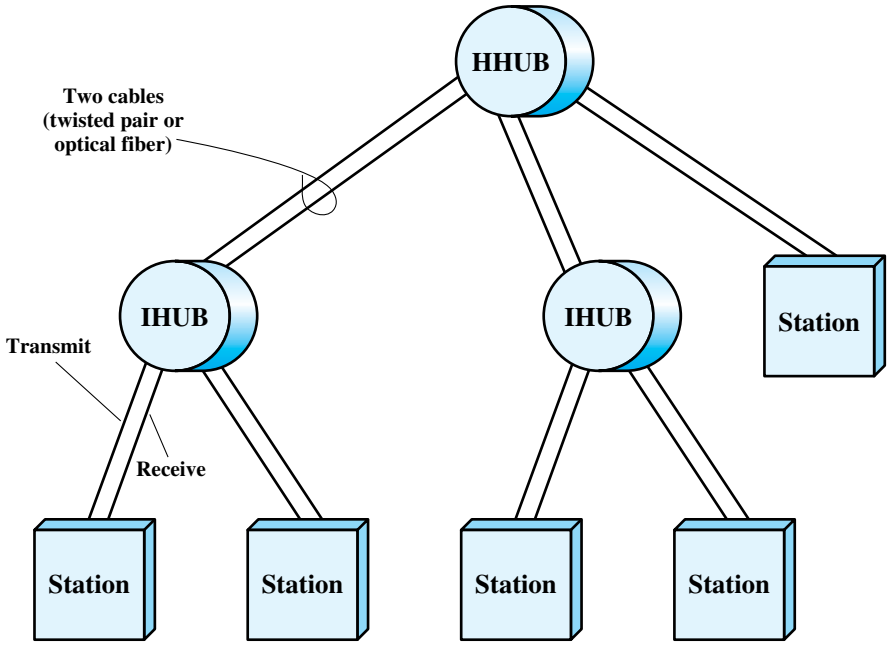

Figure 15.12 Two-Level Star Topology

mixture of stations and other hubs attached to it from below. This layout fits well with building wiring practices. Typically, there is a wiring closet on each floor of an office building, and a hub can be placed in each one. Each hub could service the stations on its floor.

Layer 2 Switches

In recent years, a new device, the layer 2 switch, has replaced the hub in popularity, particularly for high-speed LANs. The layer 2 switch is also sometimes referred to as a switching hub.

To clarify the distinction between hubs and switches, Figure 15.13a shows a typical bus layout of a traditional 10-Mbps LAN. A bus is installed that is laid out so that all the devices to be attached are in reasonable proximity to a point on the bus. In the figure, station B is transmitting. This transmission goes from B, across the lead from B to the bus, along the bus in both directions, and along the access lines of each of the other attached stations. In this configuration, all the stations must share the total capacity of the bus, which is 10 Mbps.

A hub, often in a building wiring closet, uses a star wiring arrangement to attach stations to the hub. In this arrangement, a transmission from any one station is received by the hub and retransmitted on all of the outgoing lines. Therefore, to avoid collision, only one station can transmit at a time. Again, the total capacity of the LAN is 10 Mbps. The hub has several advantages over the simple bus arrangement. It exploits standard building wiring practices in the layout of cable. In addition, the hub can be configured to recognize a malfunctioning station that is

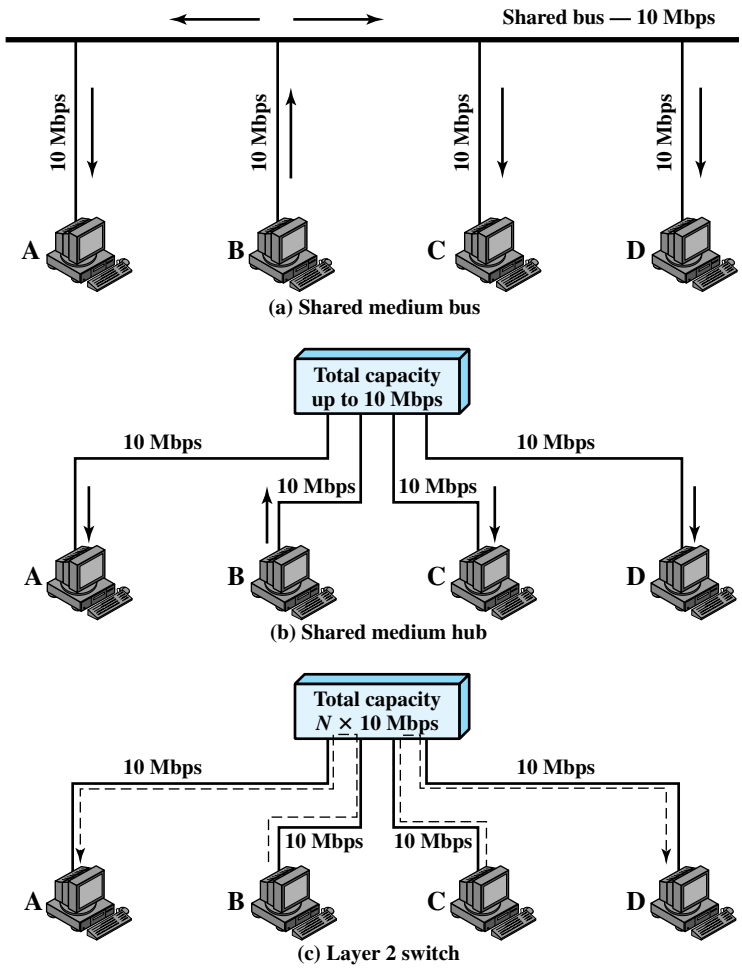

Figure 15.13 Lan Hubs and Switches

jamming the network and to cut that station out of the network. Figure 15.13b illustrates the operation of a hub. Here again, station B is transmitting. This transmission goes from B, across the transmit line from B to the hub, and from the hub along the receive lines of each of the other attached stations.

We can achieve greater performance with a layer 2 switch. In this case, the central hub acts as a switch, much as a packet switch or circuit switch. With a layer 2 switch, an incoming frame from a particular station is switched to the appropriate output line to be delivered to the intended destination. At the same time, other unused lines can be used for switching other traffic. Figure 15.13c shows an example in which B is transmitting a frame to A and at the same time C is transmitting a frame to D. So, in this example, the current throughput on the LAN is 20 Mbps, although each individual device is limited to 10 Mbps. The layer 2 switch has several attractive features:

1. No change is required to the software or hardware of the attached devices to convert a bus LAN or a hub LAN to a switched LAN. In the case of an Ethernet LAN, each attached device continues to use the Ethernet medium access control protocol to access the LAN. From the point of view of the attached devices, nothing has changed in the access logic.
2. Each attached device has a dedicated capacity equal to that of the entire original LAN, assuming that the layer 2 switch has sufficient capacity to keep up with all attached devices. For example, in Figure 15.13c, if the layer 2 switch can sustain a throughput of 20 Mbps, each attached device appears to have a dedicated capacity for either input or output of 10 Mbps.
3. The layer 2 switch scales easily. Additional devices can be attached to the layer 2 switch by increasing the capacity of the layer 2 switch correspondingly.

Two types of layer 2 switches are available as commercial products:

- **Store-and-forward switch:** The layer 2 switch accepts a frame on an input line, buffers it briefly, and then routes it to the appropriate output line.
- **Cut-through switch:** The layer 2 switch takes advantage of the fact that the destination address appears at the beginning of the MAC (medium access control) frame. The layer 2 switch begins repeating the incoming frame onto the appropriate output line as soon as the layer 2 switch recognizes the destination address.

The cut-through switch yields the highest possible throughput but at some risk of propagating bad frames, because the switch is not able to check the CRC prior to retransmission. The store-and-forward switch involves a delay between sender and receiver but boosts the overall integrity of the network.

A layer 2 switch can be viewed as a full-duplex version of the hub. It can also incorporate logic that allows it to function as a multiport bridge. [BREY99] lists the following differences between layer 2 switches and bridges:

- Bridge frame handling is done in software. A layer 2 switch performs the address recognition and frame forwarding functions in hardware.
- A bridge can typically only analyze and forward one frame at a time, whereas a layer 2 switch has multiple parallel data paths and can handle multiple frames at a time.
- A bridge uses store-and-forward operation. With a layer 2 switch, it is possible to have cut-through instead of store-and-forward operation.

Because a layer 2 switch has higher performance and can incorporate the functions of a bridge, the bridge has suffered commercially. New installations typically include layer 2 switches with bridge functionality rather than bridges.

Layer 3 Switches

Layer 2 switches provide increased performance to meet the needs of high-volume traffic generated by personal computers, workstations, and servers. However, as the number of devices in a building or complex of buildings grows, layer 2 switches

reveal some inadequacies. Two problems in particular present themselves: broadcast overload and the lack of multiple links.

A set of devices and LANs connected by layer 2 switches is considered to have a flat address space. The term *flat* means that all users share a common MAC broadcast address. Thus, if any device issues a MAC frame with a broadcast address, that frame is to be delivered to all devices attached to the overall network connected by layer 2 switches and/or bridges. In a large network, frequent transmission of broadcast frames can create tremendous overhead. Worse, a malfunctioning device can create a *broadcast storm*, in which numerous broadcast frames clog the network and crowd out legitimate traffic.

A second performance-related problem with the use of bridges and/or layer 2 switches is that the current standards for bridge protocols dictate that there be no closed loops in the network. That is, there can only be one path between any two devices. Thus, it is impossible, in a standards-based implementation, to provide multiple paths through multiple switches between devices. This restriction limits both performance and reliability.

To overcome these problems, it seems logical to break up a large local network into a number of **subnetworks** connected by routers. A MAC broadcast frame is then limited to only the devices and switches contained in a single subnetwork. Furthermore, IP-based routers employ sophisticated routing algorithms that allow the use of multiple paths between subnetworks going through different routers.

However, the problem with using routers to overcome some of the inadequacies of bridges and layer 2 switches is that routers typically do all of the IP-level processing involved in the forwarding of IP traffic in software. High-speed LANs and high-performance layer 2 switches may pump millions of packets per second, whereas a software-based router may only be able to handle well under a million packets per second. To accommodate such a load, a number of vendors have developed layer 3 switches, which implement the packet-forwarding logic of the router in hardware.

There are a number of different layer 3 schemes on the market, but fundamentally they fall into two categories: packet by packet and flow based. The packet-by-packet switch operates in the identical fashion as a traditional router. Because the forwarding logic is in hardware, the packet-by-packet switch can achieve an order of magnitude increase in performance compared to the software-based router. A flow-based switch tries to enhance performance by identifying flows of IP packets that have the same source and destination. This can be done by observing ongoing traffic or by using a special flow label in the packet header (allowed in IPv6 but not IPv4). Once a flow is identified, a predefined route can be established through the network to speed up the forwarding process. Again, huge performance increases over a pure software-based router are achieved.

Figure 15.14 is a typical example of the approach taken to local networking in an organization with a large number of PCs and workstations (thousands to tens of thousands). Desktop systems have links of 10 Mbps to 100 Mbps into a LAN controlled by a layer 2 switch. Wireless LAN connectivity is also likely to be available for mobile users. Layer 3 switches are at the local network's core, forming a local backbone. Typically, these switches are interconnected at 1 Gbps and connect to layer 2 switches at from 100 Mbps to 1 Gbps. Servers connect directly to layer 2 or

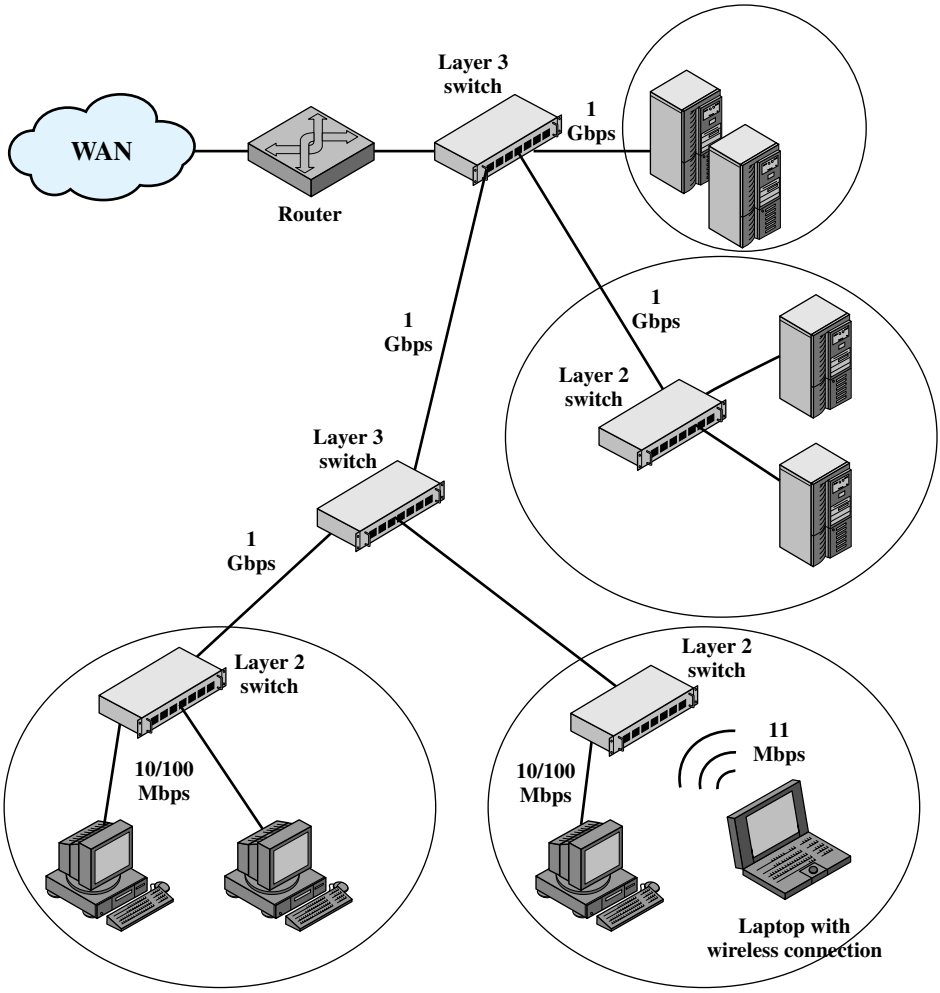

Figure 15.14 Typical Premises Network Configuration

layer 3 switches at 1 Gbps or possible 100 Mbps. A lower-cost software-based router provides WAN connection. The circles in the figure identify separate LAN subnetworks; a MAC broadcast frame is limited to its own subnetwork.

15.6 RECOMMENDED READING AND WEB SITE

The material in this chapter is covered in much more depth in [STAL00]. [REGA04] and [FORO02] also provides extensive coverage. [METZ99] is an excellent treatment of layer 2 and layer 3 switches, with a detailed discussion of products and case studies. Another comprehensive account is [SEIF00].

FORO02 Forouzan, B., and Chung, S. *Local Area Networks.* New York: McGraw-Hill, 2002.

METZ99 Metzler, J., and DeNoia, L. *Layer 2 Switching.* Upper Saddle River, NJ: Prentice Hall, 1999.

REGA04 Regan, P. *Local Area Networks.* Upper Saddle River, NJ: Prentice Hall, 2004.

SEIF00 Seifert, R. *The Switch Book.* New York: Wiley, 2000.

STAL00 Stallings, W. *Local and Metropolitan Area Networks, Sixth Edition.* Upper Saddle River, NJ: Prentice Hall, 2000.

Recommended Web site:

- **IEEE 802 LAN/MAN Standards Committee:** Status and documents for all of the working groups

15.7 KEY TERMS, REVIEW QUESTIONS, AND PROBLEMS

Key Terms

bridge	local area network (LAN)	star topology
bus topology	logical link control	tree topology
hub	medium access control (MAC)	switch
layer 2 switch	ring topology	storage area networks (SAN)
layer 3 switch	spanning tree	

Review Questions

15.1. How do the key requirements for computer room networks differ from those for personal computer local networks?

15.2. What are the differences among backend LANs, SANs, and backbone LANs?

15.3. What is network topology?

15.4. List four common LAN topologies and briefly describe their methods of operation.

15.5. What is the purpose of the IEEE 802 committee?

15.6. Why are there multiple LAN standards?

15.7. List and briefly define the services provided by LLC.

15.8. List and briefly define the types of operation provided by the LLC protocol.

15.9. List some basic functions performed at the MAC layer.

15.10. What functions are performed by a bridge?

15.11. What is a spanning tree?

15.12. What is the difference between a hub and a layer 2 switch?

15.13. What is the difference between a store-and forward switch and a cut-through switch?

Problems

15.1 Instead of LLC, could HDLC be used as a data link control protocol for a LAN? If not, what is lacking?

15.2 An asynchronous device, such as a teletype, transmits characters one at a time with unpredictable delays between characters. What problems, if any, do you foresee if such a device is connected to a LAN and allowed to transmit at will (subject to gaining access to the medium)? How might such problems be resolved?

15.3 Consider the transfer of a file containing one million 8-bit characters from one station to another. What is the total elapsed time and effective throughput for the following cases:

a. A circuit-switched, star-topology local network. Call setup time is negligible and the data rate on the medium is 64 kbps.

b. A bus topology local network with two stations a distance D apart, a data rate of B bps, and a frame size of P with 80 bits of overhead per frame. Each frame is acknowledged with an 88-bit frame before the next is sent. The propagation speed on the bus is 200 m/μs. Solve for:

1. $D = 1$ km, $B = 1$ Mbps, $P = 256$ bits
2. $D = 1$ km, $B = 10$ Mbps, $P = 256$ bits
3. $D = 10$ km, $B = 1$ Mbps, $P = 256$ bits
4. $D = 1$ km, $B = 50$ Mbps, $P = 10{,}000$ bits

c. A ring topology local network with a total circular length of $2D$, with the two stations a distance D apart. Acknowledgment is achieved by allowing a frame to circulate past the destination station, back to the source station, with an acknowledgment bit set by the destination. There are N repeaters on the ring, each of which introduces a delay of one bit time. Repeat the calculation for each of b1 through b4 for $N = 10; 100; 1000$.

15.4 Consider a baseband bus with a number of equally spaced stations with a data rate of 10 Mbps and a bus length of 1 km.

a. What is the mean time to send a frame of 1000 bits to another station, measured from the beginning of transmission to the end of reception? Assume a propagation speed of 200 m/μs.

b. If two stations begin to transmit at exactly the same time, their packets will interfere with each other. If each transmitting station monitors the bus during transmission, how long before it notices an interference, in seconds? In bit times?

15.5 Repeat Problem 15.4 for a data rate of 100 Mbps.

15.6 At a propagation speed of 200 m/μs, what is the effective length added to a ring by a bit delay at each repeater?

a. At 1 Mbps

b. At 40 Mbps

15.7 A tree topology is to be provided that spans two buildings. If permission can be obtained to string cable between the two buildings, one continuous tree layout will be used. Otherwise, each building will have an independent tree topology network and a point-to-point link will connect a special communications station on one network with a communications station on the other network. What functions must the communications stations perform? Repeat for ring and star.

15.8 System A consists of a single ring with 300 stations, one per repeater. System B consists of three 100-station rings linked by a bridge. If the probability of a link failure is P_1, a repeater failure is P_r, and a bridge failure is P_b, derive an expression for parts (a) through (d):

a. Probability of failure of system A

b. Probability of complete failure of system B

c. Probability that a particular station will find the network unavailable, for systems A and B

d. Probability that any two stations, selected at random, will be unable to communicate, for systems A and B

e. Compute values for parts (a) through (d) for $P_1 = P_b = P_r = 10^{-2}$.

15.9 Draw figures similar to Figure 15.9 for a configuration in which

a. Two LANs are connected via two bridges that are connected by a point-to-point link.

b. Two LANs are connected via two bridges that are connected by an X.25 packet-switching network.

15.10 For the configuration of Figure 15.10, show the central routing matrix and the routing tables at each bridge.