COMPUTER NETWORKS

Computer networks today play a central role in our lives by enabling communication for the exchange of information between various computers. Standalone computers today are a rarity. Indeed, most personal computers in peoples' homes are connected to the giant computer network termed the *Internet.* Service providers provide access to the Internet for a modest fee. A basic facility provided by networks is electronic mail, which allows sending information over the network to any user whose electronic address on the Internet can be specified. One can send text, graphical images, and video.

Sitting at home, one can access the entire wealth of information that has been made publicly available. One can monitor the latest swings on the stock exchange, shop in electronic malls and purchase items, read the news, search for and collect all the electronically accessible information on any topic, and provide children access to educational resources for their projects. Even people who do not have access to a computer are heavily reliant on computer networks. For example, when a person calls in to make an airplane reservation over the phone, there is an airline reservation agent at the other end using the computer network to coordinate the reservation with other agents elsewhere. And when a person withdraws cash at a bank teller machine late at night, he or she is indirectly using the computer network to update the available bank balance.

Businesses, particularly large ones with branch offices in distant locations, depend heavily on the computer network for the exchange of critical business information, as well as routine but still confidential information for the day-to-day administration of the business, among the various branches. In fact, a lot of businesses have installed their own private networks, which only their employees can access, for security and efficiency. Many small (including home-based) businesses are also developing a presence on the computer network for advertising and for reaching out to their customers in the increasingly competitive marketplace. Books, journals, and magazines are all accessible over the network, perhaps for some access fee. Electronic banking and commerce allow the manipulation of financial accounts and the conduct of business transactions over the computer network, instantaneously. Just as the advent and proliferation of computers made offices "paperless," so also the proliferation of computer networks eliminates the need for paper mail and makes information exchange almost instantaneous.

Computer networks also enable distance learning and multimedia teleconferences, wherein several participants at different sites receive instruction or can participate in a conference live, using audio as well as video information and images transmitted over the network. Computer networks are also being used in unexpected fields such as medicine. Surgeons in small hospitals can perform emergency surgery by using the on-line advice of expert surgeons at distant sites who are observing a detailed video close-up of the surgery being transmitted live over the computer network. Radiologists can read images of computer tomography scans and magnetic resonance images that are sent to them from a remote site and perform diagnosis remotely. This allows the expertise of highly accomplished doctors to be made available to patients in inaccessible and distant regions.

The World Wide Web (WWW), which uses a computer network protocol, is extending the reach of information repositories on computers into our daily lives. Coupled with the growth of wireless networks that allow ubiquitous access to information, and the proliferation of fiber-optic cables and satellite technology that are capable of delivering data at very high speeds, computer networks are poised to reshape our lives in unimagined ways.

TYPES OF NETWORKS

A computer network, simply stated, is a collection of computers connected together by a communication medium and following a consistent network protocol for sending and receiving information between any two computers on that network. Today, there exist numerous computer networks, each of which follows one of a small suite of standardized network protocols. A computer network that connects geographically dispersed sites is classified as a *wide area network* (WAN). A computer network that connects computers within a building or a campus is classified as a *local area network* (LAN). A network that connects computers across an intermediate range is classified as a *metropolitan* (or *medium*) *area network* (MAN). This classification is important because the requirements and problems of transmitting data (henceforth, the term data implicitly includes not just textual data but also video and audio—in short, any bitstream) differ according to the geographical extent of the network, and hence the solutions differ greatly. Thus, for example, in a WAN, there are numerous intermediate nodes in the network that perform switching functions that perform the point-to-point or store-and-forward data transfer between the endpoints. The WAN topology itself can be viewed as a huge and complex graph; the links, which are represented by edges in the WAN topology, may be over telephone lines, microwave links, or even satellites. If all the computers of a WAN belong to the same organization, then the organization may lease transmission lines from public carriers and install a private switching system at each site to create an enterprisewide private network. Alternatively, public carrier networks can be used—previously, these were the public switched telephone networks geared toward carrying voice, but recently, they have converted to provide data service using public switched data networks. In fact, several have even converted to the all-digital mode, to form ISDNs (*integrated services digital networks*). In contrast to WANs, LANs have to deal specifically with high data transmission rates and low error rates. They are usually organized as a token ring, a bus, or a tree. Traditional LANs operate by broadcasting data on the common communication medium, although switched LANs are currently an emerging trend. The functionality provided in WANs subsumes much of the functionality of LANs and MANs, and WANs are more complex. A more elaborate distinction between LANs, MANs, and WANs, and their network protocols, is made later, and we will give a somewhat more detailed coverage of issues in WANs than in MANs and LANs, in view of their relative importance.

Individual computer networks, whether LANs or MANs or WANs, can elect to connect with the other computer networks through bridges, gateways, and routers, to form a giant conglomerate network known as the Internet. Likewise, individual users can connect their personal home computer to the Internet by using a modem to dial in to a service provider (such as America Online or CompuServe) that provides connectivity to the Internet for a fee. For the layperson, "the Internet" is synonymous with "the great computer network." Actually, there are two senses in which the word "internet" is used. In this article the internet (lowercase) is a WAN that is the vast web of all kinds of networks that interconnect all the computers in the world. The Internet (capital) is a collection of those specific networks that use a specific protocol, TCP/IP (*Transmission Control Protocol / Internet Protocol*), for communication.

EVOLUTION OF NETWORKS

The Internet evolved from ARPANET, which was one of the pioneering WANs along with TYMENET, DECnet from Digital Equipment Corporation, and SNA (Systems Network Architecture) from IBM in the early 1970s. The ARPANET was initially designed for use by the United States Department of Defense through its Defense Advanced Research Projects Agency (DARPA), and its evolved form, the Internet, began to be a de facto standard by the early 1980s. The ARPANET protocol TCP/IP was made publicly available early on, whereas the protocols for DECnet and IBM's SNA were proprietary. DECnet and SNA were widely used to form private networks of most of the large and multinational corporations in the 1970s and early 1980s. However, primarily due to the proprietary nature of these protocols, new customers and network designers stayed away from them from the mid-1980s on. Although the size and number of SNA, DECnet (and other proprietary networks) has barely grown since the mid-1980s, and their market share has considerably shrunk, they still have a very wide installed base as WANs within corporations. Corporations are slow to move over to more recent, more sophisticated, and more popular protocols due to the heavy investment involved, and it is expected that these proprietary networks will continue to exist into the early part of the twenty-first century even though their market share keeps shrinking drastically.

STANDARDS

For any two computers to talk to each other, they must share the same network protocol for data exchange. Although customers shied away from SNA and DECnet primarily because they were proprietary, a more compelling reason to stay away from them would have been that they hindered easy solutions to interconnectivity with networks based on open, standardized protocols. In the 1980s, two dominant protocol models for WANs became standardized. The protocol suite used with the Internet was TCP/IP—all its specifications were in the public domain, available without a fee. Due to the widespread use of the Internet and the openness of TCP/IP, it soon became a de facto standard. The other standard was the *Open Systems Interconnection* (OSI), designed explicitly by the *International Standards Organization* (ISO). The ISO, the IEEE, and the *International Telecommunications Union–Telecommunications Sector* (ITU-T), formerly known as the *International Telegraph and Telephone Consultative Committee* (CCITT), are the major traditional standards bodies for computer networks.

OSI Reference Model

As a computer network is composed of complex software and hardware, the ISO OSI reference model organizes the network protocol into seven layers and defines the function of each layer. Figure 1 shows the seven layers at the endpoint of a connection. As is seen, intermediate network switching nodes have only the lower three layers. The layers of the OSI protocol stack are above the physical hardware of the network terminating equipment. The study of computer network protocols does not deal directly with the actual media such as cable (twisted or coaxial), optical fiber, satellite, or microwave, beyond the extent to which they determine the available bandwidth, transmission speeds, and loss ratios for transmission. The functions of the seven layers of ISO OSI

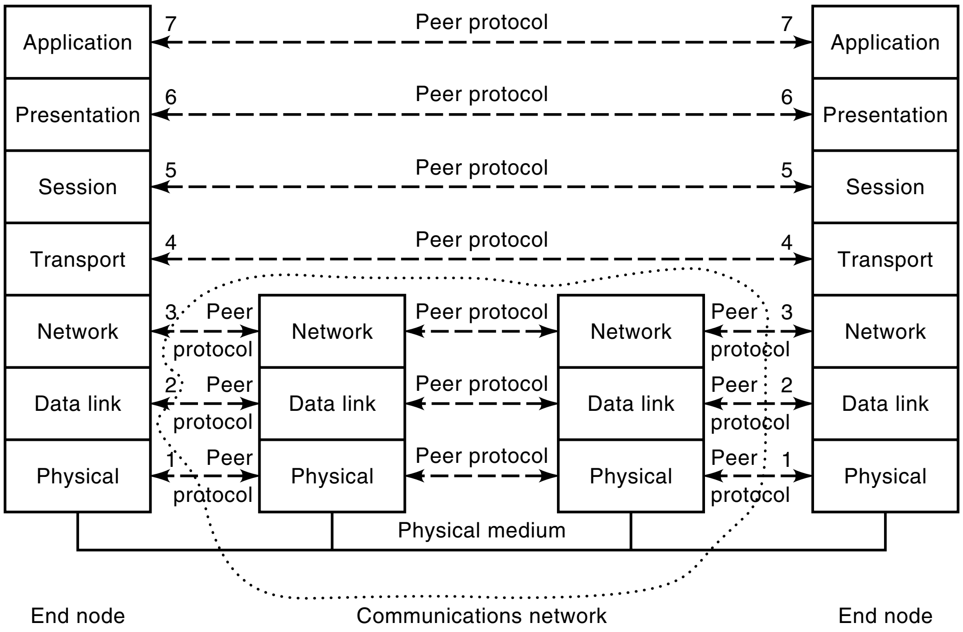

Figure 1. The OSI reference model. There are seven layers at the endpoint of a connection. Intermediate nodes have the three lower network-dependent layers.

were heavily influenced by the design of SNA and DECnet, and are as follows.

The lowest layer is the *physical* layer. Its function is to allow the data to be sent and received as raw bitstreams over the communication network hardware. It is concerned with the electrical and physical interfaces between the network equipment and the local computer's equipment. It is not concerned with transmission errors, how bits are organized, or what they mean. Rather, it needs to handle issues like the voltage levels used for representing 0 and 1 bits, the number of pins and their assignments in network connectors, the type of network (telephone system, optical-fiber interface, digital transmission, etc.), and the nature of switching used.

The *data-link* layer makes the communication facility provided by the physical layer reliable for the network layer above it. The data-link layer is responsible for recovering from transmission errors and for flow control. Handling transmission errors may involve retransmission of messages using various versions of *automatic repeat request* (ARQ) protocols that request retransmission when acknowledgements of earlier packets are not received. Flow control involves handling any disparity between the speeds at which bits can be sent and received. It involves controlling the rate of transmission of data on a link so that the receiver always has sufficient buffer storage to accept the data prior to its processing. Some common protocols for flow control are stop-and-wait and sliding-window. The *high-level data link control* (HDLC) protocol is one of the more popular data-link layer protocols.

The *network* layer is responsible for establishing an end-to-end connection between the two transport layers above it at the two endpoints of the connection. The network layer is responsible mainly for routing and congestion control. For routing, the network layer segments a message received from the transport layer into packets and decides which outgoing link will carry the packets to the destination. Congestion control concerns controlling the problem that arises when the composite rate at which data packets enter the network, or any part of the network, exceeds the rate at which the data packets leave the network or that part of the network. In the case of internetworking between two networks, the network layer provides compensating functions to enable the interconnectivity.

The physical layer, data-link layer, and network layer are network-dependent layers in that their detailed design depends on the characteristics of the network.

The *transport* layer's primary function is to hide all the details of the communication network, embodied in the network-dependent layers below, from the application-oriented layers above it. Thus, it provides a network-independent end-to-end communication facility. The transport layer can provide the network with the ability to inform the host computer that the network has crashed or has lost certain packets, thereby also providing improved reliability. The transport layer defines five classes of service, from class 0, which offers very basic functions, to class 4, which provides sophisticated error recovery and flow control functions. The choice of class of service accommodates the differing *qualities of service* (QoS) provided by the different networks.

The three application-oriented layers that use the services of the transport layer are as follows.

The *session* layer is responsible for establishing a connection (also known as a session), maintaining it during the lifetime of the session, and then taking it down. Establishing the connection may involve authentication and the selection of the right transport service. The session layer's responsibilities include providing a mechanism to synchronize transfer of the data units depending on the duplex mode of communication (half-duplex or full-duplex), providing support to establish periodic checkpoints for transaction-processing support, and reporting nonrecoverable exceptions to the presentation and application layers above it. The synchronization function also involves keeping track of outstanding requests and replies from the two endpoints, and ordering them so as to simplify the design of user programs.

The *presentation* layer is the interface between the application layer above and the rest of the network services below it.

This layer is concerned with the representation (syntax) of data transfer, and uses the abstract data syntax of the application along with one of the standardized transfer syntaxes which it negotiates with the presentation layer at the other end of the connection. The use of the common transfer syntax may require the use of data transformation utilities to take care of the differences in data representation at the source and the destination. The presentation layer also performs data compression, encryption, and conversion to and from network standards for files and terminals.

The *application* layer provides a facility for the application processes to use the OSI protocol stack and provides a user interface to networkwide information services and operating system functions. The information services include document management and information interchange services, and file-transfer access and management. The content of the application is left to users. Thus, an application program for banking or airline reservations will have its own standards for this layer. Typical services that are used in this layer include identification of the intended partners by name or address, determining the availability of the intended communication partner, agreement on responsibility for error recovery, and identifying constraints on data syntax such as character sets.

The ISO OSI does not specify how the layers should be implemented. Each layer is aware only of its function and the formats of data that it handles; that is, data that it receives from the layer above (below) it are transformed in a standardized way, and then passed on to the layer below (above) it. Each layer runs a protocol having defined header formats with its corresponding layer at the other end of the connection. The layer is a *service user* of the layer below it, which is said to be a *service provider* to this layer. A layer does not understand the header or protocols used by the other layers. This makes each layer independent, so any layer can change its protocol without affecting other layers, as long as the interfaces between the layers remain unchanged.

To send a message from endpoint *A* to endpoint *B*, the message of the application at endpoint *A* gets successively transformed as follows. Starting with the topmost layer, each successive layer at *A* transforms the data of the message, adds a header containing control information meant for its peer layer at endpoint *B* to the message, and passes the transformed message, termed a *protocol data unit* (PDU), to the layer below it (the service provider to it). Finally, the physical layer at computer *A* transmits the raw bits to the physical layer at computer *B* using the transmission medium. Starting with the physical layer, each successive layer at *B* transforms the received message (the PDU for that layer) by stripping the header added by its peer layer at endpoint *A*, performs the necessary processing identified by the header, and passes the message to the next layer above it. Eventually, the application layer at *B* receives the message. OSI defines four classes of service primitives—*request, indication, response,* and *confirmation.* A service user at an endpoint *A* generates the request primitive, in response to which the service provider at endpoint *B* of the connection issues the indication primitive to the peer service user at that endpoint *B*. The peer service user then issues the response primitive to its service provider, in response to which the service provider at endpoint *A* generates a confirmation primitive for the service user at endpoint *A*. Normal data transfer involves only the request and indication primitives (*unconfirmed* service), whereas certain forms of synchronization between the two endpoints, such as when a connection has to be established, use all four primitives (*confirmed* service).

Other Standards

For each of the OSI layers, a set of standards has been established to form what is known as a open systems interconnection profile. These standards have been developed either individually or jointly by the standards bodies such as ISO, CCITT (now ITU-T), and IEEE. Some of the more commonly used profiles include GOSIP for use in US and UK government projects, TOP for use in technical and office environments, MAP for use in manufacturing automation, and CEN standards for use in European government projects. There are no optional profiles for the presentation layer and the session layer, which have to use the ISO 8822 and 8823 profiles, and the ISO 8326 and 8327 profiles, respectively. However, there are numerous options for the application layer. Some popular examples are file transfer, access, and management (FTAM) (ISO 8571), electronic mail (CCITT X.400), and directory services (CCITT X.500). There are a few options for the transport layer, such as ISO 8072 and 8073, which, when combined with the five different protocol classes for this layer, yield a variety of options. At the network layer, X.25 is the standard for WANs, whereas ISO 8348 and 8473 are the choices for LANs. At the physical and data-link layers, there are several options. The data-link layer can be viewed as being composed of two sublayers—the upper sublayer, *logical link control* (LLC), and the lower sublayer, *medium access control* (MAC). The OSI profiles refer to the IEEE standards for the above layers. The LLC is defined by IEEE 802.2, whereas the combination of MAC and the physical layer is defined by either IEEE 802.3 [Ethernet, or *carrier sense multiple access with collision detection* (CSMA/CD)], IEEE 802.4 (token bus), or IEEE 802.5 (token ring). IEEE 802.3, 802.4, and 802.5 are typically used for LANs. For WANs, the X.25 protocol is used.

The OSI reference model is widely used to demarcate the various functions that need to be performed in a computer network. IBM's SNA and DECnet come closest to following the OSI model; unfortunately, these networks have not been growing since the mid 1980s. TCP/IP, frame relay (FR), and asynchronous transfer mode (ATM) are some of the most popular current network protocols, and their layering structure differs from that suggested by the OSI model even though they perform similar functions. Thus for example, TCP/IP has the following five layers: application layer, transport layer, Internet layer, network access layer, and physical layer, whereas the layers of ATM, which performs cell switching, pertain only to functions of the network-dependant layers. Some of these protocols have been standardized by specific forums or by de facto popularity. Thus, TCP/IP has become a de facto standard and is being refined at the Internet Engineering Task Force, whereas ATM and FR are being standardized in dedicated forums in collaboration with ITU-T. ATM, TCP/IP, and other popular protocols will be examined in more detail subsequently.

OVERVIEW OF INTERNETWORKING

Intermediate nodes along the end-to-end route have only their network-dependent layers participate in the transmission of data between the endpoints. Such nodes simply provide a switching facility for forwarding the data to the destination. Special nodes in the networked system that connect together networks of different types to form larger composite networks are routers, repeaters, bridges, and gateways. *Routers* connect two similar or dissimilar networks and play the role of the network layer, i.e., they implement the physical, data-link, and network layers of the OSI model. A *repeater* is a physical layer device that relays the signals with amplification. A *bridge* is a data-link layer node that connects two nodes with different physical layers, but with the same data-link layers, that support identical packet size and addressing schemes. Typically, a bridge interconnects LANs with each other or with WANs. The functions of a bridge require protocols for forwarding packets and protocols for exchanging routing information. A *gateway* is a node that connects a OSI-profile-compliant network to a non-OSI-profile-compliant network. A gateway contains all the seven layers of functionality and is therefore a complex intermediate node connecting two very different networks with different protocol stacks.

MULTIPLEXING

Multiplexing is a technique that allows multiple signals to be transmitted simultaneously over a shared data link (or other shared resource), and is almost universally used in computer networks. The earliest forms of multiplexing were used in telephone networks. *Frequency division multiplexing* (FDM) is an analog form of multiplexing that uses modulation of different-frequency signals. FDM was used for the older telephone networks; it is used currently in cable television, which has a bandwidth of about 500 MHz; and a new variant called wavelength division multiplexing (WDM) is coming into use over wide-bandwidth fiber-optic cables. *Time division multiplexing* (TDM) is a digital technique wherein time slots are allocated among the different signals. In *synchronous* TDM, the various multiplexed signals are allocated time slots in a round-robin manner, with all the signals getting time slots at the same periodicity. Synchronous TDM can waste significant bandwidth if some of the multiplexed signals do not have any data to send and their allotted time slots consequently go unused. Synchronous TDM also causes bandwidth wastage due to bandwidth fragmentation, described later. *Asynchronous* TDM, also known as statistical TDM, overcomes the drawbacks of synchronous TDM by not preassigning time slots to input lines. If an input line has no data to send, then the time slot or frame is allocated to the next input line that has data to send. Thus, the time slots are allocated dynamically, and the ratio of time slots to input lines that are actively sending is higher, thus enabling a greater usage of bandwidth. In fact, with statistical TDM, the sum of the speeds of the input lines can be greater than the line speed of the path, according to a statistical analysis of the number of input lines that are likely to be actually transmitting at any given time.

Recently, telephone companies have been offering digital services using digital transmission, which has lower cost, higher speed, better quality, and less susceptibility to noise than previous analog services, to transmit data. Digital lines can be analog or switched, much like the earlier analog telephone lines, which were analog or switched. As high-speed digital transmission technology became more common, there was a need to standardize a hierarchy of digital services. DS-0, DS-1, DS-2, DS-3 and DS-4 service provides 64 kbit/s, 1.544 Mbit/s, 6.312 Mbit/s, 44.376 Mbit/s, and 274.376 Mbit/s digital service, respectively. Telephone companies in North America use standard T-1, T-2, T-3, and T-4 lines, whose speeds match the data rates of DS-1, DS-2, DS-3, and DS-4 service, respectively. Although T lines are digital and carry digital data, they can also carry analog signals, such as telephone signals, that are sampled and time-division multiplexed. For example, 24 telephone lines, each at 8000 samples/s and 8 bits/sample, along with a net overhead of 8 kbit/s, give an effective data rate of 1.544 Mbit/s, which is the bandwidth of a T-1 line. (Each sample of the 24 telephone lines needs 24 × 8 = 192 bits, and 1 bit of overhead. 8000 samples/s × 193 = 1.544 Mbit/s.) In Europe, the E series of line speeds is used instead of the T series. E-1, E-2, E-3, and E-4 lines have speeds of 2.048 Mbit/s, 8.448 Mbit/s, 34.368 Mbit/s, and 139.264 Mbit/s, respectively. Conceptually, the two systems are identical.

Although multiplexing was initially used in telephone networks, it is now being used almost universally for data networks. For example, asynchronous TDM is used in high-speed digital technologies such as ISDN, ATM, and SONET.

SWITCHING TECHNOLOGIES

WANs have traditionally been implemented with one of two technologies—circuit switching and packet switching. In circuit switching, a dedicated communication path is established between the two endpoints. This path is a connected sequence of links between the network nodes such that a logical channel is used for the connection on each of the links. Circuit switching is analogous to the operation of the telephone network. The path has to be established using a control signaling protocol before the data can be transferred, and the path is usually taken down after the data transfer. Circuit switching was driven by telephony, which required that there should be minimal transmission delay without any variation. As transmission and reception occur at the same rate, a constant signal transmission rate must be maintained. Routing in a circuit-switched network such as a telephone network must be efficient, adaptive in the face of varying traffic loads, and resilient to failures. Circuit-switched telephone networks have been very widely deployed and are ideally suited for voice traffic.

Circuit switching requires that the transmitting and receiving ends of a connection use the same data transfer rate, thus limiting the interconnectivity of different computers. Moreover, if the connection is used only intermittently, as in typical client–host interactions, the bandwidth is grossly underutilized, as time is typically allocated using synchronous TDM. In contrast, packet switching breaks up the data to be transmitted into small packets, which are individually transmitted after being enhanced with some overhead control information concerning the routing of the packet to the destination. At each intermediate node, the packet is received, its control information is examined if necessary for routing, and then the packet is forwarded to the next node. Multiple packets from different logical connections can share a link, thereby using bandwidth dynamically and efficiently instead of being confined to a fixed transmission bandwidth, and nodes can perform data-rate conversion to allow two nodes with different data rates to exchange packets. Under heavy traffic load, circuit-switched networks block calls, whereas packet-switched networks accept packets, which may encounter an increased delivery delay.

Unlike circuit-switched networks, in which a dedicated path is established for a connection, there is no dedicated path in packet-switched networks, in which there are two ways in which packets may be routed—datagram and virtual circuit. In datagram service, each packet is routed independently by the intermediate nodes, according to factors such as network load. Each packet encounters a packet transmission delay, and packets may be delivered out of order at the destination, hence, they need to be reordered there. However, there is no connection setup delay. In virtual circuit service, a preplanned route is established before packets are sent (thus incurring a connection setup delay) but this is not a dedicated path. Packets are individually sent on this virtual circuit and are buffered along the intermediate nodes. But individual nodes along the virtual circuit make no routing decisions, or simple routing decisions if multiple virtual circuits share the link. This simplifies the routing and reduces the routing delay as compared to datagram service, in which each node along the route of each individual packet is involved in making a routing decision for that packet. Also, the network can be relied upon for sequencing and error control when virtual circuit is used. Datagram service is used in the Internet protocol of TCP/IP, whereas virtual-circuit service is used in ATM networks. Datagram service may be preferable for short messages, whereas virtual circuit service may be preferable for long messages. The header information is greater in datagram service, thus consuming more bandwidth than in virtual circuit service. Likewise, flow control and congestion control are more difficult in datagram service because there is less information about what information is flowing for what connection. Datagram service also introduces more variation in the transit delay time. At the network layer, OSI defines two types of switching services: connectionless service and connection-oriented service, which use datagram and virtual circuit packet switching, respectively. OSI allows a choice between connection-oriented service and connectionless service at each of the three networking layers.

There are two flavors of virtual circuit packet switching. In switched virtual circuit (SVC) mode, the virtual circuit is established when needed, and is disconnected when the need is over. In contrast, a permanent virtual circuit (PVC) is dedicated to the pair of end users and exists even if it is not being used. Thus, two SVC users may get a different route each time they request a SVC, but two PVC users always use the same route.

Routing in Packet-Switched Networks

Routing in packet-switched networks greatly affects the efficiency of the networks. Multiple performance criteria such as the number of hops from the source to the destination, the cost of the route selected, its delay, and effective throughput must be considered. The route selection itself will occur at the source node for virtual circuit packet switching or at each intermediate node along the route for datagram service. The information used to compute a good route can be either locally available information, information from adjacent nodes, information from nodes along the route, or information from all the nodes. The topology and available bandwidth within the network are changing all the time; changed information can be distributed within the network either periodically, or when there is a major change in the load and/or the topology.

Some of the simpler routing algorithms used are *fixed routing,* in which a simple table lookup is performed to determine the route to the destination; flooding, which requires no network information and which is highly resilient to failures but which generates a very high traffic load; and *probabilistic routing,* in which an outgoing link from a node is chosen with fixed probability. However, *adaptive routing,* which is routing based on the present network conditions, is greatly preferred because it chooses a route that is close to optimal, can handle node or link failures by not including them in any route, and enforces congestion control by routing packets around regions of the network that are heavily congested. Adaptive routing has high overhead—status information has to be collected from throughout the network, and the choice of the route has high processing overhead. The ARPANET and its follow-on versions of the Internet used some versions of the classical Bellman–Ford algorithm and the classical Dijkstra algorithm, which are essentially graph algorithms that find the minimum-cost path (1).

X.25 is the protocol standard that specifies the interface between a host system and a packet-switching network. X.25 was initially proposed in 1976 and subsequently revised several times. X.25 defines the interface between the DTE (*data terminal equipment*) and the packet-switching network node DCE (*data circuit-terminating equipment*) for the physical layer, link access layer, and packet or network layer. One of the more common standards assumed by X.25 for the physical layer is known as X.21; it interoperates with others such as EIA-232. At the data-link layer, X.25 uses *link access procedure, balanced* (LAPB), which is a subset of HDLC, and at the network layer it uses a protocol named *packet layer protocol* (PLP). X.25 performs error detection and correction in both the data-link and network layers. Although X.25 is treated largely as a WAN technology in the Unites States, at the present time, Europe uses X.25 for both WANs and LANs.

LOCAL AREA NETWORKS AND METROPOLITAN AREA NETWORKS

A LAN connects computers spaced by a short distance such as a few miles. LANs find their applications in the academic world (to provide remote access to computing facilities), business offices (for sharing of expensive resources, files, and databases), and manufacturing (for distributed real-time control applications). A LAN is characterized by a variety of parameters such as topology, access control, architecture, and the transmission medium.

LAN Topology

The topology of a network defines in what physical configuration the components of the networks are interconnected. There are three popular LAN topologies: star, ring, and bus.

In the *star* topology, a central or common switching node (working as a primary or hub) provides direct connectivity to the other nodes (known as secondary). Each secondary node sends a data packet to the primary node, which passes it on to the destination secondary node. The data packet contains the destination node address. The communication between any two nodes takes place via circuit switching being used by the central node. Any node wishing to communicate with another node must request the central node to establish a connection or circuit between it and the destination node. The data transfer will take place only after this connection has been established. The central node basically acts as a repeater: it accepts data from a node and passes them to the destination node.

In the *ring* topology, all nodes are connected to a circular ring via repeaters. A repeater is a hardware device that accepts data from one node bit by bit and passes them to the next connected node. A node sends the data packet (frame) to another node on the ring via the repeater connected between them. As the data packet travels around the ring, the destination node recognizes its address defined in the frame and copies the data into its buffer. The data packet may have any number of bits (from one bit to several bits). The data packets are transmitted as blocks or frames, which are the PDUs. Each frame contains the user's data, control information, and destination address.

The ring topology supports the token-passing method for accessing data over the network. This method determines which node gets the access right to the ring for data communication. In this method, a token of unique fixed bit pattern is defined and this token always circulates around the ring. Any node that needs to send a frame must first capture the token. After it has done so, it attaches the data packet to the token and puts the token back onto the ring. The destination node receives this token along with the frame, copies the frame into its buffer, changes a certain bit in the token frame to indicate an acknowledgment to the sender node, and puts the token onto the ring. After the sender receives the acknowledgment, it sends the token back onto the ring. This topology is also known as a token-passing topology. The network interface devices are simply repeaters, and the nodes require less processing of functions needed to access the network (in contrast to the star topology, where complex processing of functions is needed to access the network at each node).

In the *bus* topology, the nodes are connected to a common link (transmission media) via connecting taps or hardware interfaces. Since only one link or bus is used as a transmission medium among connected nodes, each node that needs to transmit must first listen to the bus. If the bus is free, the node will send the data packet (frame) over it to the destination. If the bus is busy, the node will wait and try again to send the same data frame at a later time. All the nodes are treated equally, and any node can request to access the network at any time.

Access Control

There are two main types of access control techniques: *carrier sense multiple access* (CSMA) and *token passing.* The CSMA protocol is used by Ethernet over the bus topology, while the token-passing protocol works over either the ring or the bus topology.

ALOHA. The first carrier sense access technique was developed for the ALOHA system for packet radio transmission at the University of Hawaii. In this access method, a node waits for a data packet and sends it to the network immediately after it receives it from the user. Then, it listens for the acknowledgment (ACK) packet from the destination node. The duration of listening must be equal to the maximum round-trip propagation time delay of the network topology. The transmission is said to be *complete* when the node receives the ACK. If it does not receive the ACK or times out, it will retransmit the same packet after a random amount of time, typically greater than the round-trip propagation time.

In ALOHA, any station that wants to transmit a message uses the common channel, and it gets an acknowledgment from the receiving station. This works well for light traffic (offering fewer collisions), but for heavy traffic the maximum utilization of ALOHA degrades. The utilization in pure ALOHA was improved by introducing time slots and transmitting messages only at the beginning of slots. This type of network is known as slotted ALOHA.

Carrier Sense. In this technique, a node listens to the transmission medium (channel) before sending any data packet (message). If the channel is free, it can transmit the message in the form of a fixed-size frame containing control and data information. If the channel is not free, it will wait for a fixed amount of time, depending on the type of protocols used, and retry to see if the channel is free. It may happen that at the time of listening to the channel, a node finds the channel free and while it transmits a message, there is another message sent by another node also in the channel. A *collision* occurs when two (or more) stations after finding the channel free, simultaneously transmit their messages. In case of a collision, messages are withdrawn from the channel and must be transmitted again after some random amount of time, depending on the protocol used. This scheme on the bus topology uses a CSMA technique for accessing the media.

The carrier sense access control technique is mainly used in CSMA and token bus LANs for bus or tree topologies. The size of a data frame depends on the physical distance the frame has to travel from one end to another (propagation time). The propagation time is defined as the time taken by a frame to go from one end to another end of the LAN. The round-trip propagation delay is twice this. If the transmission time of the frame is too short compared to the propagation time, collisions may not be detected. In the event of collision, the first station to detect the collision sends a jam signal (a predefined bit pattern) on the channel. This unique signal indicates to all the connected stations that a collision has occurred. Then all the stations have to withdraw from the channel and retry after a random time.

CSMA/CD and CSMA/CA. There are versions of CSMA protocols where collisions can be *detected* by special hardware devices and appropriate utilities of protocols. Alternatively, collisions can be *avoided* by using special types of hardware and software. In these versions of protocols, the access control techniques are appropriately termed as CSMA/CD (collision detection) and CSMA/CA (collision avoidance).

In CSMA/CD, the network interface listens to the channel and, after it finds it free, puts the frame on the channel as discussed above. In CSMA/CA, the network interface avoids any attempt to detect a collision, and instead senses the channel twice and makes necessary arrangements before it sends the data packets. Thus, it may not be necessary to transmit a jam signal in the event of collision, as collisions are less likely to occur. Although the number of collisions is not reduced, the efficiency of this scheme is not lower than CSMA/CD.

Ethernet IEEE 802.3 is a well-known CSMA/CD bus system. Ethernet offers a data rate of 10 Mbit/s and is manufactured by a large number of manufacturers and vendors. The earlier versions of Ethernet used baseband transmission on coaxial cable with its bus topology. Currently, Ethernet uses broadband transmission techniques (fiber optics, etc.) and offers higher data rates.

Versions of Carrier Sense Protocols

There are different types of strategies for sensing and retrying the channel before and after the occurrence of the collision. These strategies can be classified as *nonpersistence* (NP) and *p-persistence* (PP) protocols. There exist a variety of vendor protocols for each of these classes. Several CSMA algorithms have been designed to handle the situation where the medium is busy by allowing the nodes to send their data messages later.

Nonpersistence Carrier Sense Protocols. In the NP class of access controls, the protocol senses the channel and, if it is free, transmits the message. If a collision occurs, a node waits for a random amount of time before it relistens or resenses the channel. If it finds the channel free, it will transmit; otherwise it keeps on retrying by waiting for a random amount of time until it is able to transmit the message. Thus, this access algorithm waits for some time after it finds the medium busy, before listening to the medium again. This type of protocol is not efficient and wastes bandwidth because during the time that a particular station that has experienced a collision is waiting, no stations, including that one, may be trying to sense the channel.

***p*-Persistence Carrier Sense Protocols.** The *p*-persistence (PP) class of access control techniques is mainly based on continuous monitoring of the channel to use its bandwidth effectively by reducing the number of collisions. There are two versions of PP protocols available: 1-persistence, and *p*-persistence proper. In the 1-persistence version, after finding that the medium is busy, a station listens to it continuously until it finds it free, and then transmits the data packets immediately. If a collision occurs, the station waits for a random amount of time and again senses the channel continuously. In this way, the station has to wait for a random amount of time only if a collision occurs, while in the other cases it is sensing the channel continuously. This protocol is better than the NP protocol in that it offers higher throughput and is more efficient. In the *p*-persistence version, the node transmits the data packet with probability p after it finds the medium free.

Token Passing

In the token-passing LAN where the stations are defined in a logical ring topology using a token bus or are physically connected in a ring, the network operates in a deterministic manner.

A token bus LAN on a bus topology (coaxial cable) defines a logical ring for the stations that need to transmit a message among themselves, irrespective of their physical locations. Each station is assigned an identification number, which is attached to the message. Each station knows the address of its predecessor and successor stations. It receives the data packet from its predecessor and, after performing the required functions on it, passes it to its successor station. A free token (defined by a unique bit pattern) always circulates around the logical ring. A station that needs to transmit any data packet waits for this free token. After it receives the token, it attaches the data packet to it and transmits it to the successor station, which forwards it on the logical ring. After the data packet has been received by the destination station, it marks the token as free and forwards it to the next station. This scheme of accessing the channel or network is based on the round-robin technique in which each station has an equal right to access the token. Further, the priority for each station can be defined at the beginning and a higher-priority station always get the token for transmission before a lower-priority station. The assignment of priority is independent of the physical locations of the stations.

A similar concept is used in the token ring LAN, where a token of fixed length circulates around the physical ring to which stations are connected. Any station needing to transmit the message data packet first captures the token, appends the data message to it, and then puts it back onto the ring. The destination station, after receiving the message, copies it and changes the status bit of the token, and sends it back onto the ring. The sending station looks at the bits in the frame status (FS) of the token ring MAC frame for the acknowledgment. If the status bit is not set, the token along with the message data packet keeps on circulating around the ring until it is either received or removed by the sender station.

Contention Ring. In a traditional ring system, there is only one token circulating around the ring all the time, and this may lead to significant waiting time in the case of light network traffic. A different approach, known as the *contention ring,* wherein no token is circulating continuously, is adopted in ring-based LANs. If any node needs to transmit a data frame, it waits to see if any data frame with the token has passed over it. If there is no such frame, it will send the data and append a free token frame at the end of the data frame. The source node will either remove the data frame from the ring after it has been received by the destination node, or let the frame go around the ring few more times, depending on the protocols being used.

During the time the data frame and appended (new) token are circulating around the ring, the token may become busy or remain free. If it becomes busy, it will carry the data from the requesting node, and the LAN will behave like a token ring network. But if it is free, it will come back as a free token. The source node removes it from the ring along with the data frame. Now if any node needs to transmit a data frame, it has to redefine the token contention configuration by creating a token and appending it at the end of its data frame. This scheme is comparable to that of the token ring LAN except for light-network-traffic applications. The choice between token ring and contention ring depends on parameters such as the collision frequency, waiting time, and network traffic.

Slotted Ring. Another version of a ring-based system is the slotted ring, where time slots of fixed size (length) are defined.

These slots are regarded as carriers (like a conveyor belt) for the data frames around the ring. Each slot has its own status indicator, and if a slot is free, a data frame can be copied into it. If the length of data is greater than the length of a carrier, the data get partitioned into packets of the size of the carriers (slots) and then transmitted. For each slot, a bit can be set and reset to indicate acknowledgment, and the source node must make it free for further transfer of data frames.

Cambridge Token Ring. The cambridge token ring was developed from the slotted ring by Wilkies and Wheelerin in the mid-1970s in the Computer Laboratory at Cambridge University, England.

This LAN protocol, based on the ring topology, defines an empty slot (in contrast to a fixed-size token used in token ring and token bus LANs) and offers a data speed of 10 Mbit/s. Each node is assigned a slot for sending a single packet only; that is, the slot can be used by a node only once for sending one packet, after which it has to pass control to the next node. The second time slot must be passed unused. The third time slot will be used again. This means that a node can use at most one out of three slot rings at any time. In general, we can say that a node will use a ring slot once in $n + 2$ slots in an n-slot ring provided no other node is making a request to use it. If more than one node is requesting the ring slot, each will get a slot at least once in $m + n$ slots, where n is the number of slots and m the number of nodes making requests concurrently.

The MAC protocol used in this LAN is based on the empty-slot concept. In this technique, a small-sized packet continuously circulates around the ring. At the beginning of the packet, a bit is reserved for indicating the status of the packet (full or empty) and is similar to the bit used in the token ring LAN for indicating acknowledgment. Any station that needs to transmit the packet will capture the token, change the status of this bit to full, and append the data to the token. No station can use the token more than two times after capturing it once; it must send the token back onto the ring by making it an empty token. If the station wants to send more data packets, it has to capture the token again.

Metropolitan Area Networks

The standard LANs discussed above provide data communication within a limited range. The coverage of LANs can be enhanced by using bridges between identical LANs. LANs are usually defined for data communication, resource sharing, and so on within the bandwidth of 10 Mbit/s (Ethernet), 4 or 16 Mbit/s (token ring), or 100 Mbit/s [*fiber distributed data interface* (FDDI)]. The Ethernet and token-ring LANs are typically used for data communication and do not offer enough bandwidth and speed for video and voice communications. Although one may try to compress the data before sending over these LANs, this restricts their use for other data, besides the limited range of such LANs. To provide both audio and video services over a long distance in metropolitan cities, IEEE defined another standard LAN [ISO 88802.6 (2)], known as the *metropolitan area network* (MAN) (3). This LAN allows media access protocols to provide sharing of resources by users over a large distance within a metropolitan city using coaxial cable or fiber transmission media. The media access protocols of existing LANs (Ethernet, token ring, etc.) by themselves cannot cover a longer distance. Further, the speed of transmission of data over these links is limited to 10 Mbit/s or 16 Mbit/s only. The MAN provides a speed over 100 Mbit/s and supports data and voice over the same link.

LANs are usually controlled by a single user and so can be considered as dedicated to one user. They are limited in geographical distance to the premises of an organization such as a university. In order to make LANs cost-effective, the LANs of different organizations may be connected to each other via MANs, making the LANs public networks that offer services to all the interconnected organizations. This interconnection covers a larger geographical area such as a metropolis. The MAN consists of dedicated circuits, which are distributed throughout the metropolitan area at various organizational locations.

A MAN provides transfer services for medical images and data, graphics, computer-aided software engineering (CASE) tools such as CAD/CAM, and the compressed digital video signals used in teleconferencing and the like. For voice communication, the delay introduced by the network must be small; the IEEE 802.6 standard sets a maximum delay for voice communication of 2 ms in a MAN. The protocols defined for a MAN must deal with this delay and provide proper synchronization between the sending and receiving sites for voice and video signals, along with some kind of security.

Within each organization, various buildings and computer stations may be scattered over a metropolitan region. Distance and the speed of networks become a major problem for data communication across such stations. From the implementation point of view, it is efficient to have smaller networks connecting the computers within an immediate vicinity and then connect these networks by a high-speed network (MAN). The MAN connects these networks (typically LANs) via gateways, which provide a suitable set of protocols for internetworking between them.

Internetworking can also be achieved between WANs, mainframes, and so on. Based on distance and applications, the concept of the MAN (a high-speed network providing interconnections to LANs and WANs and supporting data, video, and audio traffic) has been extended to different types of networks defined lately. Some of the networks using MAN concepts include the FDDI, the *distributed queue dual bus* (DQDB), broadband integrated services digital networks (B-ISDN), and *switched multimegabit data services* (SMDS). The FDDI has been used as a backbone network in a majority of campus environments and as such can be considered as a private network providing high-speed interconnections to LANs and WANs.

Distributed Queue Dual Bus. The DQDB, as defined by the IEEE 802.6 Working Group (4), is based on the switching used in queued packet and synchronous circuit exchange (QPSX) and defined as a MAN/LAN standard. It supports isochronous, connection-oriented, and connectionless services simultaneously. The DQDB MAN uses two buses in such a way that two signals flow in opposite directions. It is independent of physical media and supports the transmission of *pleisochronous digital hierarchy* (PDH) systems (34 Mbit/s, 45 Mbit/s, and 145 Mbit/s), *synchronous digital hierarchy* (SDH) transmission, and data rates of up to several gigabits per second.

The communication switching in a DQDB MAN integrates video, voice, and data traffic and transmits it over circuit-switched and packet-switched links. The DQDB is mainly used for communication over a larger geographical area than the LANs. The DQDB also offers sharing of telecommunication resources and various facilities to its users, and is also known as a *high-speed public communication network* within that area. The information within this network is transmitted within slots. A slot consists of a header of 5 octets and information field of 48 octets (the ATM cell of 53 octets—48 octets of information and 5 octets of header—is derived from the DQDB slot). The slots are identified by inserting virtual circuit identifier (VCI) values in the slot header and are controlled by a slot generator. There are two types of slots defined in DQDB: prearbitrated and queue-arbitrated. Non-isochronous information is transported with queue-arbitrated slots.

IEEE Project 802 covers the physical layer and the data link layer for LANs and MANs, as well as part of the network layer for interconnectivity of LAN protocols. Project 802.2 deals with the LLC, 802.3 with CSMA/CD, 802.4 with the token bus, 802.5 with the token ring, and 802.6 with DQDB, as discussed above.

EMERGING NETWORKS AND PROTOCOLS

The earliest computer networks used public switched telephone networks for data communication. The evolving computing and communications technologies, along with the need to transfer not just voice but also other forms of data such video, images, and multimedia data at high speeds all across the globe, paved the way for an all-digital-mode network architecture termed the *integrated services digital network* (ISDN). Although ISDN is intended to be a worldwide integrated public telecommunications network, it has not yet been deployed as ubiquitously as was originally hoped. In part, this is due to the existing investment in TCP/IP and other network protocols, and the prohibitive cost of replacing all the existing public telephone and telecommunications networks at once.

ISDN is defined by a set of user interfaces allowing a broad range of traffic types and value-added processing services. All users get a uniform view of a single worldwide network, although there may really be multiple networks with digital switches. There are two parts to the ISDN specification—narrowband ISDN uses 64 kbit/s channels as the basic unit of switching, with a circuit-switching flavor, and broadband ISDN supports data rates of several hundred megabits per second with a virtual-circuit packet-switching flavor. FR falls under the narrowband ISDN category, whereas ATM falls under the broadband ISDN category.

Both FR and ATM have very little overhead of error checking and control built into their protocol stacks, compared to the OSI reference model. This saves a lot of bits that packet-switched protocols following the OSI model need for redundancy checking and error control at the intermediate nodes and at the end nodes, and the associated processing times at all these nodes. Such a protocol design for FR and ATM is possible because the error rates have fallen dramatically from those of the digital transmission facilities available when the early packet-based networks were being developed and the OSI reference model was formulated. The few occurring errors today can be detected and recovered from by the higher layers of protocols, without having to waste a significant percentage of the bandwidth for error detection and correction, and the associated processing times. Currently, FR networks are designed to operate efficiently around 2 Mbit/s and use 64 kbit/s channels as the basic unit of switching, which is about the effective data rate for the end user in early packet-switched networks. Broadly speaking, ATM is similar to FR except that FR uses variable-sized packets, whereas in ATM all packets have a fixed length of 53 bytes, thereby reducing the processing overhead even further as compared to FR. Currently, ATM networks operate at a rate from tens to hundreds of megabits per second, as compared to the 2 Mbit/s of FR.

INTERACTION WITH OTHER SYSTEM COMPONENTS

Although a computer network transfers data from one computer to another, the use of such a network, particularly a WAN, by an application is more involved. The application must be provided a way to indicate that it wants to communicate with a remote computer. This is usually done through an *application programming interface* (API), which allows an application to specify the manner in which it wants to send and receive information and the desired communication partners. Three broad communication paradigms exist:

1. *Remote procedure call* (RPC), by which an application makes a procedure call to a routine that resides at a remote site. RPC can be synchronous or asynchronous, and orthogonally, it can be blocking or nonblocking.
2. Messaging and queuing, by which communication is through mailboxes.
3. Connection-oriented communication.

Some examples of standardized APIs for networking are Common Programming Interface—Communications (CPI-C), which is standardized by X/Open (5), and Message Passing Interface (MPI).

The processing of the communication primitives of the application program involves considerable interaction with the operating system. For example, the physical location of the partner has to be determined using *directory services* to map the logical name of the partner to its network address. Also, a communication connection has to be established with the partner. If any transactional semantics are associated with the application, then interaction with the *transaction manager* at the site is required. If the application is involved in group communication, then interaction with group computing software is required to maintain a cohesive view of the distributed computation in the face of node or link failures, in addition to providing essential properties such as consistent and ordered message delivery across all messages to all recipients within the group.

HIGH-SPEED NETWORK PROTOCOLS

Traffic Characteristics

The earlier data networks were designed to handle interactive and batch data. Broadly speaking, voice, video, and images, which are being increasingly sent on computer networks, can also be regarded as data. However, there are significant differences between traditional data traffic and these other forms of traffic that need to be taken into account when designing high-speed networks that are sensitive to the type of traffic being transmitted. Voice transmission does not need any flow control, whereas traditional data need flow control because the source computer has a very high capacity to generate data traffic. Data transmission occurs in a very bursty manner; voice traffic, although occurring in spurts, statistically poses a less severe problem for the switching system. Voice traffic, on the other hand, has very stringent transmission delay constraints: even a satellite communication may impose only a 250 ms delay. In contrast, batch and interactive data can be more tolerant of transmission delays. Errors in data traffic must be detected and corrected, whereas voice traffic is tolerant of a significantly larger error rate because it is inherently very redundant, besides which voice traffic cannot tolerate any delays that would result if error recovery were attempted. Voice traffic is usually two-way, whereas most other data traffic is one-way. The volume of voice traffic at 64 kbit/s greatly exceeds the volume of other traffic, although this disparity is slowly decreasing. As a result, any feature added to the network design to improve the performance of data transmission must not increase the cost of or adversely affect voice transmission. Image transmission is similar to traditional data transmission, except that images are very voluminous.

Digital video transmission is similar to voice transmission in several respects: it is tolerant of errors on account of the redundancy of information, and it is isochronous, that is, frames that are delivered to the network at a constant rate must be played out at the other end at the same rate. However, video data rates are much higher—for example, the H.320 Motion Pictures Expert Group (MPEG) MPEG-4 specifies a rate of 3.04 Mbit/s for videophone, MPEG-1 PAL and MPEG-1 NTSC specify a rate of 30.4 Mbit/s for videoconferencing, MPEG-2 PAL and MPEG-2 NTSC specify rates of 124.4 Mbit/s and 124.3 Mbit/s, respectively, for broadcast television, and HDTV and MPEG-3 specify rates of 994.3 Mbit/s and 745.8 Mbit/s for high-quality television. These rates can be contrasted with 64 kbit/s for telephony, with 2.458 Mbit/s and 6.636 Mbit/s for SVGA and JPEG (Joint Photographic Experts Group), respectively, meant for normal-resolution images, and with 31.46 Mbit/s for very high-resolution images. Data compression typically gives up to an order-of-magnitude reduction in all these rates, but even after compression they are demanding by current network technology standards.

Although video transmission requires a constant bit rate, for less demanding applications such as videoconferencing and videophone one can exploit the fact that there are only marginal changes from one frame to another, and thus much of information content of each frame is redundant. Frames that are digitally encoded and then compressed can exhibit a variable bit rate.

Multimedia applications involve the simultaneous transfer of traditional data, images, video, and voice. Even outside such applications, networks are being expected to carry all these different kinds of data from unrelated connections.

Principles of High-Speed Networks

A network architecture that accommodates the varied and often conflicting requirements of traditional data, voice, video, and images must be very flexible, provide very wide bandwidth, provide a means of synchronization, minimize *latency* (the end-to-end transfer time, including the transmission time, propagation time, and buffering delays), minimize *jitter* (the variation in latency), and minimize *skew* (the lack of synchronization between multiple media being played back at the destination). In this section, we examine the characteristics of such a high-speed network, without confining ourselves to the specifics of any one network protocol.

The network architecture must clearly be of the packet-switching type, and perform the switching at intermediate nodes in hardware, which can be several orders of magnitude faster than software switching. Packet switching minimizes wasted bandwidth, and also allows the overlap of transmission times with propagation time and buffering delays. See also the discussion on switching technologies. (For a datagram service on packet-switching networks, the transit delay, jitter, and cost of extra headers are hard to control, and hence datagram service without special enhancements is not competitive in high-speed networks.) It is essential that no delays be introduced at intermediate nodes for switching. Recovery from errors conflicts with the requirement for uniform delivery rates. An error in the packet header will cause a packet to be lost and an extra packet to appear at a wrong destination. Fortunately, modern transmission media such as fiber-optic cables have a very low error rate; error detection in the data portion of the cells can be relegated to upper-layer protocols on an end-to-end basis, whereas there should be some method to cope with errors in the packet header.

Short fixed-length packets, called *cells,* have the following advantages over large variable-length packets. Their fixed length gives a uniform transmission time to the queuing within a node for an outgoing link, leading to a more uniform delay characteristic for the network. Short cells have a short assembly time and hence shorter delay. Buffering in link queues is easier and requires less logic. Buffers at intermediate nodes can be smaller and more easily managed. Also, when an error occurs on a link, there are less data to retransmit. The disadvantages are the following. The processing time increases, particularly at the endpoints, and at intermediate nodes if hardware switching is not used. Packet headers consume additional bandwidth, and this effect is aggravated if connection-oriented protocols (i.e., virtual circuit protocols) are not used. If end-to-end error recovery is done by retransmitting whole blocks even though congestion control discards individual packets, then significant overhead is incurred. ATM transfers data in 53-byte cells, whereas FR uses variable-sized packets.

In a high-speed packet-switching network, traditional flow control techniques based on acknowledgments are not practical. Rather, input rate regulation and bandwidth reservation provide the equivalent functionality of flow control. These can be agreed upon prior to the traffic flow for a connection, and enforced at the origin. In spite of having flow control, congestion can build up at intermediate nodes when the traffic arriving at a node along different connections overwhelms its buffers. An obvious solution to the congestion problem is to ensure that the maximum demand on any part of the network can always be met, but this solution is unacceptable because it requires utilizing links and nodes at a small fraction of their maximum bandwidth. Rather, the most expedient solution is to allocate bandwidth to various connections so that the sum of their average utilizations is close to the maximum bandwidth of the link or node. Relying on the statistics of large numbers, it is expected that at no time will all the connections be transmitting data at their peak capacity simultaneously, and hence all their traffic can be accommodated most of the time, with only a small chance that occasionally the net rate of traffic arriving at the node or link exceeds its capacity. On such rare occasions, when congestion sets in, packets can be discarded, perhaps based on some scheme that ranks the relative importance of the connections affected by congestion. The high-speed network needs some *adaptation* at its entry and exit points to create frames or cells at the source and assemble them at the destination, and to detect and handle errors or lost packets appropriately. Typically, cell networks do not use any other strategy for flow control or congestion control or error recovery, whereas packet networks may use some other strategy within the network, although at the cost of slightly degraded performance.

Flow control for high-speed packet networks is best done at the entry point to the network, not within the network, and congestion control, if it is still needed, is done by discarding packets. When a connection is set up, various quality-of-service parameters such as the minimum guaranteed throughout rate, priority, maximum guaranteed packet loss probability, guaranteed maximum allowed throughput rate, and security level are negotiated. Then a leaky bucket mechanism is used at the input, which allows packets to enter the network only at the negotiated rate. A small input queue or buffer queues up the extra packets, which get transmitted once the burst above the average rate dies out. However, for bursts of large duration, the input queue overflows and packets are lost. Thus, the leaky bucket limits the rate to the negotiated average rate, but allows packets sent during short bursts to enter the network.

The design of high-speed packet switches and cell switches is also important. A popular design is the serial multistage switch. Some of the functions performed by the switch on receiving data are: detecting boundaries of characters, recognizing characters and sequences, detecting and synchronizing boundaries between fields of data and blocks of data, transferring data and control information into memory, and processing the link control logic, the switching logic, and the management logic. Analogous functions must be performed when transmitting or forwarding the data. Some of the architectural issues concerning the switch design are the interconnection structure of the switch, the routing architecture for the packets, buffering strategies and strategies for resolving contention for concurrent access to the same output port, and the interconnection with the control processor of the switch.

TCP/IP

Transmission Control Protocol. This protocol was defined by the Department of Defense (DoD) for use in ARPANET, industries, unreliable subnets, vendor networks, universities, and research and development (R&D) divisions, and is also referenced as military standards. In the original version of ARPANET, the subnet was supposed to offer virtual circuit services (reliable) and the transport layer was rightly named Network Control Protocol (NCP). TCP was designed to work on unreliable subnets and was mainly concerned with the transmission of *transport PDUs* (TPDUs) over networks. It offers reliable and sequenced packet delivery at the destination site. This means that it provides error recovery for lost or damaged packets and duplicated packets, and also handles nonsequenced packets. The error recovery facility includes procedures for sequencing the data TPDUs (expressed in octets), an appropriate checksum algorithm for error detection, and methods for acknowledgment, retransmission, and avoidance of duplicate packets. The TCP resides in the transport layer under the upper-layer protocols. Another protocol defined by DoD for the network layer is Internet Protocol (IP) (4). By using TCP and IP together, users can transmit large amounts of data and large files over the network reliably and efficiently.

The TCP-and-IP suite is known as TCP/IP. TCP offers support for connection-oriented service, while another protocol, User Datagram Protocol (UDP), defined within TCP/IP, supports connectionless service.

TCP/IP offers a communication protocol that is independent of underlying networks (LANs, WANs, or any other interconnected network). It assumes that the network layer is using IP and that the protocols of the transport layer (TCP) and the network layer (IP) interact with each other via a defined set of primitives. It defines data and acknowledgment formats, and also defines procedures to ensure the orderly delivery of packets, the initialization of data stream transfer, and the indication of the completion of data transfer. This protocol does not specify application interfaces, and further it offers system-dependent implementation, thus offering flexibility to the users. It offers two important options to users: *data stream push* and *urgent data signaling.*

The first option allows the formation of a TPDU only after the node has received enough data for its transmission, which means the data identified up to the push flag boundary. The TCP user can also request TCP to transmit all outstanding data using this flag. In the second option, the TPDUs are sent as datagram packets by IP and as such are also known as IP datagrams. IP does not define or concern itself with any type of underlying networks. In contrast to this, the ISO protocols for the network and transport layers are network-dependent, that is, the protocols used by the network layer define the type of underlying network being used. In other words, the network layer protocol offers a connection-oriented interface for connection-oriented networks like X.25, while it offers a connectionless interface for connectionless networks that use ISO–IP protocol.

There is also a conceptual difference between TCP/IP and the ISO protocols in terms of the layered architecture. There is no layer over TCP that interfaces directly with the application, whereas the transport layer of the ISO protocols has to provide services to the higher layers.

TCP defines similar logically distinct processes (like that of ISO 8073) between which the data are transferred and offers a full-duplex line configuration between them. The TCP transport entity accepts a message from the user process, breaks the message into packets of fixed size (64 kbyte), and transmits each packet as a datagram. A sequence field of 32 bits is used to provide a sequence number to every packet transmitted by the transport layer (TCP). TCP offers services to the users through various application protocols being used above TCP, and these services are known as *reliable stream transport service.* This service is similar to the one provided by the class 4 version of the transport ISO protocol.

A request for connection establishment is issued from the local transport entity to the remote transport entity using certain services primitives. After the connection is established, the data transfer takes place. At the end of this session, a request for connection termination can be issued by either transport entity. The *higher layers protocols* (HLPs) issue a request for connection and assign a specific transport layer through which the connection is to be established. An HLP can also request a specific transport layer to wait for a connection request that it is expecting from the remote transport layer. This scheme is useful for allowing remote users to access a database, electronic bulletin board, or any other sharable files/programs.

During the data transfer, the TCP usually offers an error-free environment, that is, it makes sure that no packet is lost or duplicated and that the sequence of arrival of packets is maintained. All the packets corresponding to user data or control packets are treated as a sequence of message primitives. These primitives are *request, indication, response,* and *confirm,* which are the same as the ones in ISO protocols. These primitives are usually grouped into request and response packets. All the data and control messages are treated as units known as *segments* (in TCP/IP terminology), which are similar to PDUs defined in ISO protocols. The segments are transmitted by TCP and may include messages from one or more users.

The user data (UD) from HLPs are given to the transport entity in the form of stream-oriented as opposed to block-oriented CCITT TPDUs. TCP defines the data packets for stream-oriented data, using the user's data packets it receives from HLPs, to contain user data and control information. Each of the data segments (expressed in octets) defined by TCP is numbered sequentially and is acknowledged appropriately. These data segments are given to IP (the network layer), which transmits them to the remote TCP user process (the transport layer entity). When a packet arrives at the destination, it is stored in a memory buffer assigned to the application. Such packets are delivered to the destination when this buffer is full. TCP supports priority for small messages, and these messages are transmitted by bypassing the normal flow control.

The packets containing user data and control information are segmented into segments of fixed size. These segments are numbered sequentially and the acknowledgments are also received using these numbers. The interface between TCP and the user process is defined by a set of primitive calls including various parameters. For sending/receiving the data, commands such as OPEN, CLOSE, and SEND, similar to system calls in operating systems, are used. The segments are passed by TCP to IP, which transmits them via subnets to the remote TCP user process. Each transmitted octet is assigned a sequence number, which forces the remote TCP user process to send an acknowledgment. An acknowledgment can be requested for a block of octets, which implies that all the preceding octets have been received.

Internet Protocol. The DoD defined the Internet Protocol (IP) (4), which is a protocol to be used with TCP for internetworking. The development of IP took place during a DARPA internetworking research project. IP offers connectionless services to the user processes and does not require any connection establishment between them, thus reducing connection setup time. The structure of IP is somewhat similar to that for connection-oriented protocols and does not restrict any node from having both types of configurations (connectionless and connection-oriented services). As IP supports connectionless service, it constructs a datagram for each packet it receives from the transport layer entity.

The header of an IP datagram defines global addresses of distant sites. Different networks are connected via gateways, and IP datagrams are sent to appropriate IP gateways. A gateway examines the control information field in the datagram header, which defines the datagram's route within the adjacent network. If the packet does not belong to a network connected to the gateway, the gateway will send the packet to another IP gateway, which similarly forwards it via different gateways until the packet is delivered to the gateway to which is connected the network with the destination address.

The routing decision by an IP gateway is important during the implementation of the protocol, as it allows to check at each IP gateway if the destination site(s) is in the network connected directly to it (7). If the destination is in the adjacent network, then the packet will be delivered to it. If the destination is not present in any of the networks connected directly to the gateway, the gateway will find out an optimal route of gateways, and the packet will be routed over gateways and networks until it is delivered to the final destination site. In each routing decision, the objective is always to reduce the number of hops. A *hop* is defined as a simple path of length one; a *simple path* is a direct link between two nodes. A routing table containing information such as shortest routes, minimum-cost routes, and alternative routes is maintained at each site and also at each gateway. The decision tables may be static or dynamic, depending on the algorithms used to develop them.

IP defines a connectionless protocol and thus uses all the functions (e.g., routing, segmentation, reassembly) defined by ISO 8073. Further, the datagram may be lost during transmission for other reasons (e.g., insufficient buffer space, hardware failure, link failure, violation of other functions). In order to avoid the loss of datagrams, the transport layer protocols are expected to provide error control strategies. For some of the above-mentioned reasons, both standards, TCP and IP, have become very popular and are being used together around the world. The entire address for TCP/IP includes the Internet-wide IP address of the host and an additional protocol port address. The first field, *netid,* specifies the network address, while the *hostid* field specifies the host address.

The routing tables stored at each site and gateway help the gateways to identify the sequence of gateways over which the PDU should be transmitted. The routing tables adapt to any changes in the network topology, such as the failure of any site, link, or gateway. The neighboring gateways for a broken link or gateway transmit timed-out packets to all other gateways. IP datagram packets are segmented into packets (NPDUs) of mutually agreed-on size (within the maximum size); each NPDU is assigned a unique identifier. The field length indicator specifies the relative address of PDUs with respect to the IP datagram and is placed in the NPDU. Two primitives, SEND and DELIVER, are used in IP for providing communication between end-user processes. The services of networks are requested by SEND, while the DELIVER primitive is used to inform the destination end-user process about the received data packets. IP offers interfaces with a variety of networks. The IP network interface has few functions to perform, and the main task of IP is handling the routing functions.

The current version of IP is IPv4. The rapid growth of the Internet implies that the 32-bit addresses used in IPv4 will not be sufficient in the near future. In IPv4, the size of the routing tables at gateways is also very large. The Internet Engineering Task Force (IETF) is currently formulating a new version of IP, IPv6, that will address the above problems and also provide for network security and multicast support within IP.

Internetworking—Bridges, Routers, Gateways

Internetworking. In business organizations, voice and data communication are widely used. Voice and facsimile communication are handled by PBX, while data communication is handled by LANs. Business premises use several LANs, and different types of standard and nonstandard protocols are being used in these communication systems. The existing standard LANs typically offer data rates up to 16 Mbit/s over a typical distance of about 10 km. Typically, coaxial cable is used as a transmission medium in these LANs, and most of them use the ring or bus topology or the star configuration. The available LANs include CSMA/CD, token bus, and token ring. High-speed LANs (e.g., MAN, FDDI, DQDB) offer data rates of 100 Mbit/s. These high-speed LANs are used for interconnecting the existing LANs and also for high-speed data communication required by workstations and file servers. They are generally used over a large geographical region, such as a metropolitan area of 100 km diameter, and support a large number of users for data exchange. In these LANs, optical fiber is used as a transmission medium. There are two main classes of high-speed LANs: FDDI and DQDB.

Due to increasing communication requirements, interconnections must be defined between different types of LANs, between MANs and B-ISDNs, and between LANs and private MANs. A hardware device known as an *internetworking unit* (IWU) is designed to provide the interconnections between these networks. Two networks can be connected by this IWU if the distance between them is small. If the distance is large, then the networks are interconnected by intermediate subnetworks. IWUs have to deal with problems such as addressing, naming, routing, congestion control, flow control (due to different speeds), and segmentation and reassembly (due to different sizes).

The IWU is known as a *repeater* if two similar LANs are interconnected in layer 1 (physical). It is known as a *bridge* if different LANs are interconnected in layer 2 (data link), and a *router* if the networks are interconnected in layer 3 (network). If the networks are interconnected in a higher layer (normally the transport or the application layer), the IWU is known as a *gateway*.

LANs offer data rates of up to 16 Mbit/s, while ISDN currently offers only 64 kbit/s. If we have to interconnect LANs for higher-data-rate services, then they are connected to public MANs via dedicated links (attached with IWU). MANs provide connectionless services, but in the future we may expect connection-oriented and isochronous services. LANs can be interconnected with public MANs, which are then connected to B-ISDN nodes. The B-ISDN nodes are connected to each other. This type of interconnection offers LAN users access to wider areas with flexibility, low delay, and high throughput. TCP/IP protocol is generally used for LAN interoperability. A client application defined within TCP/IP allows user access to any application software on the host, transfer of files between any hosts irrespective of their locations, electronic mail facilities, and many other applications. TCP/IP allows the users to run their terminal sessions with telnet, transfer files with FTP, and use the electronic mail system with SMTP, among other functions. Many TCP/IP-based software applications for workstations for different network operating systems and servers are available in the market.

Broadband Ethernet LANs are simpler to implement and install, and are very reliable. They are available with bandwidths of 12 MHz or 18 MHz. These LANs operate at full 10 Mbit/s CSMA/CD capacity with 100% collision detection. They are transparent to nonstandard LANs (such as DECnets), TCP/IP, and many other configurations and other higher layer protocols. The baseband Ethernet usually operates at 120 Mbit/s with 100% collision detection and enforcement. Broadband Ethernet bridges are also available to provide interconnection between baseband and broadband Ethernet for increased coverage and capacity. These are available for bandwidths of 12 MHz and 18 MHz. These bridges offer high performance for 10 Mbit/s throughput, redundancy and loop detection for reliable network operations, and flexibility.

Internetworking Devices for LANs. Internetworking allows users working on different machines under different operating systems to interact with each other and use the services of remote networks as if these were local networks to the users. Internetworking can be implemented for both LANs and WANs. For internetworking between similar LANs, repeaters and bridges are used, while gateways are used to internetwork dissimilar LANs. The internetworking defined by higher layers is obtained by protocol converters.

Repeaters. As digital signals travel along the cable, their amplitude gradually decreases. If the communication stations are widely separated from each other, then the signals must be regenerated along the cable. The device that performs this regeneration is known as a *repeater*. A repeater copies the bits from one segment of a LAN and passes them on to another, connected segment. Obviously, both segments belong to same category of LAN, and the repeaters are used to enhance the length (cable length) of the LAN. For example, the transceiver chip used in CSMA/CD LANs covers a distance of 500 m, but the use of repeaters can extend the length of such a LAN up to 2.5 km. Repeaters provide internetworking at the physical layer. In some implementations, network stations themselves provide the operation of a repeater at their network interfaces (token-passing networks).

Bridges. A repeater provides interconnection between two identical LANs. A *bridge* is used when LANs have different physical transmission media and different protocols at the physical layer. The protocols higher than the physical layer must be same. A bridge may be connected between more than one LAN, but all the nodes must have the same address format, because address format conversion is not provided by bridges. A bridge also provides temporary storage for the messages that it has forwarded to another network, and fetches the messages from the storage if retransmission is requested. Multiple bridges may be used for connecting multiple LANs, and there must be one route to every node connected, as the bridges do not provide any routing.

LANs can be interconnected by bridges (at the data link layer), routers (at the network layer), and gateways (at higher layers). The LANs that utilize the same protocols for the physical layer and the IEEE MAC sublayer are connected by a device known as *MAC bridges* (8). Such a bridge can be considered as a store-and-forward internetworking unit between similar or different LANs.

A bridge listens to a LAN, accepts frames from the LAN, and passes the frames on to appropriate LANs. If a frame belongs to a LAN connected to the bridge, the bridge accepts the frame and passes it on to the destination node on that LAN. Otherwise, the bridge passes the frame on to another connected LAN. As the two LANs are identical, these MACs offer minimal processing. Thus, the function of a bridge is to transfer the frame data from one LAN to another LAN, and it may be thought of as being equivalent to a repeater that also provides a link between identical standard LANs. A repeater merely accepts the frame data from one LAN and transfers them to another LAN after regenerating the data signal (equivalent to amplification in analog signals), thus extending the length of LANs. It does not provide any mapping or routing. On the other hand, a bridge offers the following main functions to the frames during internetworking: address mapping, routing information, relaying of the frames, and buffer space for providing flow control.

The features offered by bridges include reliability (by inserting bridges at critical nodes) and security (by programming bridges not to forward sensitive traffic), connecting LANs of different buildings or floors, partitioning of the load on a single LAN, and connecting segments of LANs to avoid cabling.

Routers. Bridges provides interconnection between two similar LANs. A *router* provides interconnection between two different networks. This internetworking unit (device) is compatible with the lower three layers. Unlike bridges, it supports at least three physical links (in general, it supports other links too). A message frame transmitted over a LAN goes to all the nodes. Each node determines, by examining the address defined in the frame, if the frame belongs to it. If the frame belongs to it, the router accepts this frame and specifies the route for the frame to be transmitted to its destination. It is possible that a router allows more than one route for a frame, and the frame may have to go through a number of routers.

Each frame must contain two addresses: the destination address and the address of the next node along the route. The second address changes as the frame moves from one router to another. The routing strategies basically deal with the determination of the next node to which the frame must be sent. Routers are most commonly used for interconnecting networks from a single vendor or for interconnecting networks that are based on the same network architecture. The physical and data-link layer protocols may be different, but higher-layer protocols must be the same.

Gateways. This interconnecting device is used to interconnect different networks and must offer high-level protocol conversion. It must offer message format conversion, as the messages from different networks have different formats, different sizes, and different coding. It must provide address translation, as different networks use different addressing schemes. Finally, because these networks are using different sets of protocols at each of their layers, the gateway must provide conversions for different functions (implemented differently in different networks), flow control, error control, and error recovery. Gateways provide interconnection between different networks and are therefore flexible, expensive, and complex. The conversion of protocols has to be performed on the basis of layers.

For an incoming packet, a gateway determines the output link. It supports connection-oriented configuration-based protocols (e.g., X.25), and the decision to route the packets is made only after the connection is established. It defines an internal path during the duration of the call. In the case of a connectionless protocol, the address of every incoming packet is examined. As previously discussed, the overheads for connectionless protocols are higher than those for connection-oriented protocols. As the gateway operates at the network layer, it can easily transform or map the address of one LAN to that of another, but that makes it slower. This internetworking device is usually used in WANs where the response is slow and it is not required to handle more than 10,000 packets/s. Internetworking between dissimilar LANs can be accomplished for both connection-oriented and connectionless services.

The gateways for connection-oriented networks define virtual circuits at the network layer and are usually managed by different organizations (as opposed to bridges, which are managed by the same organizations). A gateway is partitioned into two parts, both of which are attached to each host, and which are connected by a communication link. Each part of a gateway consists of two sublayers: LAN to Internet and Internet to LAN. Each partition of a gateway is known as a *half gateway* (HG) and is controlled by a different organization. The HG uses CCITT's X.75 protocol for data communication over the network.

The X.25 protocol builds up an internetworking connection by concatenating a series of intranetworks and HG-to-HG virtual circuits. Each connection consists of five adjacent virtual circuits, called VC1 to VC5. The VC1 connection is between the source node and the HG (also known as the *signaling terminal*) in the local network. The VC2 connection is between the HG of the source and the HG of intermediate networks. The VC3 and VC5 connections are intranet just like VC1, while VC4 is another form of intranet just like VC2.

The internetworking supporting connectionless services implements datagrams. The gateways typically consist of Internet and transport packet format protocols and the formats of frames of networks that are connected by gateways. The formats of frames include data-link layer headers and trailers around the Internet packets. The Internet packet format and transport packet formats remain the same for all types of networks.

Integrated Services Digital Network

ISDN specifies a digital user interface to a public digital communication network. ISDN is defined by ITU-T and is intended to provide universal end-to-end connectivity despite the different transmission and switching services in existence. ISDN services include bearer services, teleservices, and supplementary services. Bearer services transfer data between the end users without any modification by the network, and correspond to those of the network-dependent layers of the OSI model. Teleservices are services such as fax, videotex, and teleconferencing that process and modify the transmitted data transparently to the user and correspond to the application-dependent layers of the OSI reference model. Supplementary services provide additional functions such as call waiting and message handling. ISDN is intended to reach every home and office, replacing the currently prevalent analog local loops with digital subscriber loops, but using the same installed copper wiring currently used to reach most homes and offices. All such communication connections will use the same digital interface and involve digitizing data in the homes and offices.

The ISDN model includes a digital transmission over the local subscriber loop from the home or office to the local ISDN office. ISDN classifies the various types of equipment from the home or office to the local ISDN office. Thus, (1) an NT1 device controls the physical connection between the user's system and the ISDN at the user's premises, (2) an NT2 device (optional) performs data and signal processing, such as multiplexing, flow control, and packetizing, between the user's data-generating device and a NT1 device, (3) a TE1 device is digital subscriber equipment (e.g., digital telephone, digital fax machine, integrated voice–data terminal), and (4) a TE2 device is a nondigital subscriber equipment (e.g., terminal, host computer, regular telephone) that uses (5) a *terminal adapter* (TA) device, which converts the nondigital data from a TE2 to digital form usable by ISDN. Interfaces are defined between a TE2 device and a TA (R interface), between a TA or TE1 and NT1 or NT2 (if present) (S interface), between an NT2 (if present) and an NT1 (T interface), between an NT1 and the termination of the line at the ISDN office (U interface), and between the line termination at the exchange and the exchange termination (V interface).

For the logical connection between the home or office and the ISDN office, ISDN defines two types of digital subscriber loops: the *basic rate interface* (BRI) and the *primary rate interface* (PRI), both of which are defined in terms of individual channel types: the bearer (B) channel with a data rate of 64 kbit/s full duplex (chosen to match the rate of existing telephone lines), the data (D) channel with a rate of 16 kbit/s or 64 kbit/s, and the hybrid (H) channel with a data rate of 384 kbit/s (H0), 1536 kbit/s (H11), or 1920 kbit/s (H12). The B channel carries the end user's data and can carry multiplexed streams end to end, providing an end-to-end connection using TDM techniques. A D channel, despite its name, is intended to carry control information such as synchronization information, call establishment, and alarms, and transmits data in packets. An H channel is for high-data-rate applications such as video, multimedia, and teleconferencing. The BRI is defined to contain two B channels (64 kbit/s each) and one D channel (16 kbit/s), and can use the same twisted pair of copper wires (subscriber loop) to transmit digital information rather than the traditional analog data. The PRI is defined, based on the geographical region, as a combination of B and D channels whose net data rate matches the rates of lines used for the regional telephone service. Thus, in North America, where T-1 lines with a capacity of 1.544 Mbit/s are used, the PRI contains 23 B channels (64 kbit/s each) and one D channel (64 kbit/s) for a total of 1.536 Mbit/s. Combined with the 8 kbit/s overhead of the PRI service itself, the total rate becomes that of the T-1 line. However, other combinations that add up to 1.544 Mbit/s can be used, such as a single LAN signal of 1.544 Mbit/s. In Europe, where E-1 lines with rate 2.048 Mbit/s are used, the PRI is defined to have 20 B channels and 2 D channels, or 23 B channels and 1 D channel, with a 64 kbit/s rate.

The ISDN architecture varies from the OSI reference model even though the layering structure is similar. The ISDN layers are defined in three planes—the user plane for the B and H channels, the control plane for the D channels, and the management plane for network management. In the user plane, the data link and network layers use options similar to those for the OSI layers, and the physical layer is explicitly defined by ISDN, whereas the transport and the application-dependent layers are not defined by ISDN. In the control plane, the transport and application-dependent layers use standardized ITU-T protocols such as SS-7, whereas the lower layers are explicitly defined by ISDN.

The ISDN described above is narrowband ISDN (N-ISDN), so called because of its low data rate, which was set to the existing telephone line speeds of 64 kbit/s. This ISDN offers synchronous TDM channels. Using multiple channels of 64 kbit/s gives independent 64 kbit/s channels rather than a single higher-speed channel, unless special synchronizing equipment that is not part of the network is used. FR is a specific protocol that is based on the N-ISDN definition.

The data rates of narrowband ISDN are low. To enable the growing sophisticated applications such as video and image transfer and high-definition television, it is necessary to support much higher data rates within ISDN. The resulting ISDN is known as broadband ISDN (B-ISDN) and supports data rates of the order of 600 Mbit/s currently. B-ISDN is currently under definition, and technology exists to support such data rates. Unlike N-ISDN, B-ISDN is a cell-based packet-switching network. This is because cell-based switching addresses the following two problems much better than (variable-length) packet-based switching: (1) the problem of wasted or unused bandwidth for variable-rate traffic, and (2) the bandwidth fragmentation problem associated with TDM when there is demand for arbitrary amounts of bandwidth and variable bandwidth as in video signals; it is not possible to effectively reallocate time slots for new connections in this situation. For N-ISDN and traditional telecommunication backbone systems this is not a problem, because TDM is cost-effective and efficient when bandwidth is allotted in a small number of fixed amounts. Much of the current research is on bringing about an effective realization of B-ISDN, from the physical layer hardware to the software protocols and network management solutions. ATM is a specific protocol that is based on the B-ISDN definition, and its definition is currently evolving.

SDH/SONET

Synchronous digital hierarchy (SDH) and synchronous optical network (SONET) are, respectively, the ITU-T standards from Europe and the ANSI standards from North America for the high-bandwidth TDM systems services of fiber-optic networks. SDH/SONET is a synchronous network that uses a single network-wide clock for synchronization.

SDH/SONET addresses the problems, faced by telephone companies, of compatibility between multiplexers made by different manufacturers. These devices were used to multiplex several low-speed circuits onto a high-speed link, by using several stages of multiplexers. There is a big cost benefit to integrating the multiplexing function with the internal functioning of the telephone exchange and eliminating the multiplexers. Previously, if access to a single circuit was needed, the entire signal had to be demultiplexed and then remultiplexed. SDH/SONET specifies a single multiplexing scheme that standardizes the internal operation and management of equipment from different manufacturers, and allows worldwide compatibility through a single optical multiplexing hierarchy to accommodate various existing speeds. SDH/SONET specifications allow the TDM to carry broadband services traffic such as B-ISDN and ATM traffic, and also accommodate low-speed channels, termed *tributaries.* SDH/SONET enables individual channels to be switched without requiring the entire signal to be demultiplexed and then remultiplexed, reducing the delay and associated costs. It also allows multiple speeds of operation so as to be upwardly compatible as higher optical speeds are introduced, and allows different channel speed payloads.

SONET defines several line speeds, termed synchronous transport signals (STSs). STS-1 supports a rate of 51.84 Mbit/s, and STS-3 supports a rate of 155.52 Mbit/s. The corresponding physical links are optical carriers OC-1 and OC-3. SONET has also defined STS-9, STS-12, STS-18, STS-24, STS-36, STS-48, STS-96, and STS-192 so far, where STS-x has a rate x times the rate of STS-1, and analogously for OC-x. Observe that STS-1 is defined to have a rate slightly exceeding the rate of T-3 lines, which are the fastest commonly installed electrical lines at present. Analogous to the STSs of SONET, SDH has defined the synchronous transport module (STM), which is compatible with the European E-line rates. STM-1 is defined to have the rate of STS-3, and higher multiples of STM-1 are defined to correspond to the rates of STS-3x lines.

SDH/SONET transmission uses multiplexers, regenerators, and add/drop multiplexers. A *section* is the optical link connecting any two of the above devices that are adjacent. A *line* is the optical link connecting two multiplexers and/or add/drop multiplexers. A *path* is the end-to-end connection between two multiplexers. SONET defines the photonic layer (which corresponds to the physical layer of the OSI model), and the section layer, line layer, and path layer (which correspond to the data-link layer). The section layer, line layer, and path layer are responsible for the movement of the signal across a section, line, and path, respectively. A SONET STS-1 frame is 810 bytes, and is transmitted each 125 μs, resulting in a rate of 51.84 Mbit/s. The 810 bytes are logically organized as 9 rows of 90 columns. The data payload envelope can fit exactly in a SONET frame, but is allowed to start anywhere within the frame, thereby allowing the payload envelope to span two frames. This feature is necessary to allow for differences in clock rates within the synchronous operation of SONET. Different STS-1 frames can be byte-multiplexed to form higher-speed STS-x signals. Tributaries or lower-speed signals can also be carried by SONET by allocating contiguous columns to them. A T1 payload of 1.544 Mbit/s occupies three consecutive columns, whereas a E1 payload of 2.048 Mbit/s occupies four consecutive columns. It is expected that SONET/SDH will increase in popularity and will emerge as the dominant carrier for ATM.

Frame Relay

FR is a standard interface to a packet network and was initially intended as an optional service within N-ISDN. However, it has become popular independently of N-ISDN because it can be implemented easily on existing packet-switching equipment and can provide up to an order of magnitude of improvement of throughput over similar protocols such as X.25. FR specifications pertain only to the physical and data-link layers, as FR eliminates the functions of the network layer as well as some functions of the data-link layer. FR supports all of the protocols recognized by ANSI at the physical layer. At the data-link layer, FR uses a simplified version of HDLC, eliminating most of the error control and flow control fields of HDLC. In comparison with the X.25 interface, which performs extensive error checking and flow control, requiring much overhead by way of packet headers, acknowledgments, and large buffers, FR provides for very elementary flow control and error control. This is possible and indeed desirable because the error rates of recent transmission media such as fiber-optic cables are much lower than those of media used earlier. The few errors that might occur can be handled at the network or transport layers that use the services of FR. FR thus uses no acknowledgments at the data-link layer, and most of the error control fields and messages used in X.25 are not used, saving considerable bandwidth and processing time. Further, FR switches do not have to buffer transmitted frames until acknowledgments are received, further increasing efficiency.

FR establishes a *permanent virtual circuit* (PVC) connection across the network in the data-link layer. A PVC is a form of virtual circuit packet switching. Multiple PVCs can share some of the switches and physical links along their circuits. A PVC connection involves the pairing of a local address, called a *data-link connection identifier* (DLCI), on one port (link) with a local address on another port in the network. Thus, the DLCI has only local significance on a specific network link and changes as frames traverse the network along the PVC. When a frame is received during data transfer, the FR node uses the DLCI within the frame header as an index into a local table for the incoming link on which the frame arrived, and determines the outgoing link and DLCI for this PVC. The DLCI read from the table is used as the new DLCI within the frame header, and the frame is queued up for forwarding on the link identified by the table entry. This allows frames to be routed very quickly and in first-in, first-out (FIFO) order. This function is performed in the data link-layer, rather than in the network layer as in earlier virtual circuit protocols.

Besides the switching function described above, a FR switch also does some elementary error control and flow control. When a frame arrives at a switch, the switch examines the CRC field in the frame header. If an error is detected, the frame is simply discarded, and FR relies on the transport-layer protocol to deliver the dropped frame reliably. There is no provision for an explicit flow control mechanism. When congestion occurs at a FR switch in the network, as when the traffic load is high, the switch sets a forward explicit congestion notification (FECN) bit on the frame header to indicate congestion in the forward direction of the PVC. It also sets a backward explicit congestion notification (BECN) bit on frame headers headed in the reverse direction to indicate congestion. The FECN and BECN bits are used to inform the end devices of the congestion. However, there is no explicit protocol to prevent or control the end devices from sending further frames.

While FR is not regarded as a true high-speed technology, it is still popular because of its simplicity and advantages over X.25, which is already very widely deployed.

Asynchronous Transfer Mode

ATM is a cell-switching protocol currently being standardized by the ATM Forum and being adopted by ITU-T (9). ATM has been accepted as the basis of B-ISDN. Recall that as opposed to (variable-length) packet switching, (fixed-length) cell-switching simplifies the switching hardware, reduces the transit delay through the network, provides a low variance of the transit delay due to statistical multiplexing, and simplifies the buffering and queuing at intermediate nodes. The virtual circuit routing of ATM eliminates the disadvantages of datagram service. ATM is suitable for voice, video, traditional data, and image traffic, constant-rate as well as bursty wide-bandwidth, and utilizes the link–node bandwidth very efficiently. It eliminates the large delays caused by interleaving small-packet streams, such as for audio and video traffic, with larger-packet streams using statistical multiplexing. It utilizes the wide bandwidth and low error rates of modern media such as fiber-optic cables, is connection-oriented, performs switching in hardware for speed and reliability, and is intended for use in both WANs and LANs. ATM is also designed to provide automatic error correction for each individual cell.

ATM uses virtual circuit routing and maintains the order of delivery of cells transmitted. Each physical link between adjacent nodes in the network can carry multiple virtual paths, and each such virtual path can contain multiple virtual circuits between that pair of nodes. The addressing information in each ATM cell header contains a *virtual path identifier* (VPI) for each virtual path link and a *virtual circuit identifier* (VCI) for each virtual channel link. Based on the VPI and VCI in the incoming cell header and the port number on which the cell arrives, the switch looks up a local table, which tells the output port number along which the cell should be routed, and the VPI and VCI using which the cell should be routed. Thus, VPIs and VCIs have only local significance. A *virtual path connection* (VPC) is a concatenation of virtual path links, each of which may have a different VPI. A *virtual channel connection* (VCC) is a concatenation of virtual channels, each of which may have a different VCI. When there exist multiple connections between the same endpoints, all the individual connections (VCCs) may possibly, but not necessarily, be multiplexed together on the same end-to-end VPC. When several calls from a source are made to different destinations, their corresponding VCCs may share the same VPC for part of the route, and then become a part of some other VPC. The initial determination of the route, which is a concatenation of VCCs, is done by using an explicit route selection algorithm at the source node. Once the path is determined, certain control signaling is used to reserve bandwidth resources along the path before traffic is allowed to flow on the path.

ATM defines two network interfaces. A *network node interface* (NNI) is the interface between two WANs, whereas a *user network interface* (UNI) is the interface between a user and a wide-area ATM network. An ATM cell itself is defined to be 53 bytes long, of which 48 bytes are for the payload, and 5 bytes are for the cell header. For a user payload cell or a cell sent across a NNI, one header byte is used for the header error checksum, $3\frac{1}{2}$ for the VPI/VCI labels, 1 bit to indicate the cell loss priority (CLP), which indicates the preference for whether the cell should be discarded during congestion, and 3 bits for the payload type indicator (PTI). The structure of the header of a cell sent across a UNI is similar except that the VPI/VCI label is 3 bytes and there is half a byte for flow control, which uses the generic flow control (GFC) field.

The ATM architecture defines application planes—the user plane (U), the control plane (C), and the management plane (M). The user plane protocols are used for end-to-end user communication, the control plane protocols are used for signaling that sets up and disconnects VCs, and the management plane protocols handle tasks such as reporting error conditions and dealing with VPIs and VCIs that get allocated. The application functions in these planes use the services of the following three layers defined by ATM.

The uppermost layer is the ATM adaptation layer (AAL), which provides the transparent interface between the upper application protocols and the actual ATM-specific switching and transmission. Thus, the AAL converts voice, video, traditional data, and image streams of fixed or variable rates into fixed-size (48 byte) payloads at the transmitting end, and vice versa at the receiving end. The AAL is composed of two sublayers—the convergence sublayer, which performs certain compensating functions between the service offered at the layer interface and that provided by the ATM layer below it, and the segmentation and reassembly sublayer, which converts the data packets into cells and vice versa. The exact definitions of these two layers depend on the type of service, which depends on the type of traffic.

ATM defines four types of service. AAL1 is *constant-bit-rate* (CBR) service such as for real-time voice calls and video transmission. AAL2 is *variable-bit-rate* (VBR) service such as for compressed video and compressed data. Both AAL1 and AAL2 require a strict timing relationship to be maintained between the sending and receiving ends, and are for connection-oriented traffic. AAL3/4 and AAL5 are for VBR service and do not require a timing relationship between source and destination. AAL3/4 handles the TCP as well as IP service of TCP/IP. AAL5 offers connectionless service and is broader in the scope of its service, while not offering either sequencing or error correction information.

The ATM layer is below the AAL. This layer is responsible for switching, multiplexing, routing, and traffic management services. At the sending end, it adds the 5 byte header to the 48 byte payload, and at the receiving end it strips it off. In the transmission direction, cells from multiple VP and VC streams are multiplexed to give a single noncontinuous stream of cells, and vice versa at the receiving end. ATM switch nodes perform VPI and VCI translation for routing, as described earlier. This layer also performs flow control across a UNI. The physical layer deals with the transformation of the cell flow into a bit flow and vice versa. It deals with the transmission medium and encoding. It is expected that fiber-optic cables will become more widely used for ATM transmission.

ATM is well suited not just for WANs but also for LANs. The initial CSMA/CD and token ring (4 Mbit/s and 16 Mbit/s rates) or token bus (1 Mbit/s, 5 Mbit/s, or 10 Mbit/s) LANs and the subsequent FDDI and DQDB are not as competitive as ATM when the following criteria are considered together: real-time information transfer, high bandwidth, interworking between LANs and WANs, and scalable throughput. Moreover, ATM is a standard. ATM LANs are largely star-configured with a central hub in order to use the full capacity to each end user. In order to provide this ATM to the desktop, the ATM Forum has standardized a 52 Mbit/s interface.

World Wide Web

The WWW is a vast distributed information base that is linked together by hypertext. A hypertext environment is a collection of documents that are linked together by pointers from one document to another. A reader who is reading one document can jump to other documents by choosing or clicking on hypertext in the document being read. A file of hypertext is called a *Web page* and can contain images, graphics, and voice besides plain text. Individuals can create their own Web pages through a hypertext markup language (HTML) and have the contents of the Web page, viewable through a Web browser, displayed in a very user-friendly, well-formatted manner. One can perform various powerful functions such as searching for all the information made electronically available in the world on any particular topic, and making any information available to the rest of the world. The WWW is a very recent phenomenon; since its marketplace appearance around 1990, it has become very popular because of its ease of use by the population at large.

Web pages are given unique worldwide addresses called *uniform resource locators* (URLs). A pointer in the hypertext is simply the URL of the Web page being pointed to. When one chooses or clicks on a pointer of hypertext, the associated URL is accessed over the network and displayed on one's terminal. The documents on the WWW are accessed by the hypertext transfer protocol (Http), which is an application layer protocol of TCP/IP.

ADVANCED TOPICS

Fiber Optic Networks and WDM

Lightwave networks differ from electronic networks in that light is used to transmit information instead of electrons. Lightwave networks are still in their infancy, and much research is currently being carried out to make them feasible. The principal advantage of a lightwave network is that the fiber-optic cable is capable of at least ten thousand times higher channel capacity than the current practical limit of 2 Gbit/s. Lightwave networks operate on the principle of wavelength division multiplexing (WDM), which is multiplexing light of multiple wavelengths (analogous to FDM). Provided that the wavelengths are somewhat apart, their interference is minimal. Communication on each channel can occur at full optical speed along the same optic fiber.

A simple optical communication system consists of a modulator that converts a serial bitstream into a format appropriate for optical transmission, a light source such as a laser, and a detector that converts the light into electrical form. Currently, transmission rates of 150 Mbit/s or 620 Mbit/s are commonly used. Optical fiber is smaller and lighter than electrical cables, is much cheaper per unit bandwidth, has much higher bandwidth, experiences no electrical or electromagnetic interference, offers better security, can have a longer distance between repeaters, and is such that the data transmission speed can be increased whenever newer technology becomes available. The few disadvantages are the difficulty of interconnecting the fibers and interference from gamma radiation and high-voltage electrical fields.

Wireless Networks and Wireless LANs

In most standards on LANs defined by IEEE or non-IEEE private or proprietary products, the transmission medium (coaxial cable, twisted pair, or even optical fiber) plays an important role, as it not only provides the physical link or circuit across LANs, but also determines the capacity and bandwidth (data rates) of data frames transmitted across the networks. For the LANs used on campuses or other organizational premises, the cabling or wiring sometimes becomes too messy and expensive and needs to be redone in the event of relocation of resources and other communication devices. Further, the cabling sometimes poses a serious problem in cases where lots of updating is done frequently within premises for relocation of offices, resources, and so on. Quite often, the cable installation and also the length of cables connecting these devices may become cumbersome.

The problems of cabling installation and connection have been alleviated in another category of LANs, which are based on data communication over a wireless transmission medium. The medium through which the communication takes place is air, and an air interface needs to be defined. This does not require any cabling to connect devices to LANs. Wireless transmission has been used in voice communication, for example, through radio frequency (RF) transmission, microwave links, and satellite links.

Wireless LANs offer the following advantages: (1) the cabling installation and cost are reduced; (2) they offer support for portability and flexibility; (3) the reorganization or relocation of office devices does not require any additional cost in the configuration or cabling or moving of devices.

In a typical wireless layout configuration, a terminal interacts with the hub node by using the RF band of 1.88 GHz to 1.90 GHz over a distance of less than 200 m. A typical hub may support a few radio devices or units. The wireless LAN configuration as defined above offers a new type of star topology, which supports radio connection to terminals and is based on an intelligent hub and several cell management units. The hub offers network management services, and the wireless workstations communicate with each other transparently. The hub topology follows a client–server implementation that allows it to offer high performance and efficient configuration. The hub is a wired LAN system, and the client unit comprises a network interface card (NIC) and software. The servers are connected to wired LANs, which define the hub. The servers must manage the wireless connections to all PCs via radio units and bridge the standard with the Ethernet segment of LAN. This defines a typical hub which contains the following components: radio unit, ethernet controller, board and other software units.

The clients' PCs have to be near the hub (typically within a few hundred meters) for a wireless communication link to be provided. The radio unit is typically a small board and has two antennas (omnidirectional) coming out through holes. A cable connects a controller and the PC bus provides electrical supply to this unit. The radio unit usually operates within a frequency band of 1880 MHz to 1900 MHz, which is divided into 10 channels with a spacing of 1.728 MHz between them.

The software required for client wireless PC includes three modules: LAN manager, network operating system, and installations. The LAN manager provides an NDIS-compliant interface for the LAN manager protocol stack (NETBEUI), and NetWare version protocol provides an OSI-compliant interface for the IPX protocol. The hub can be a dedicated PC that offers functions such as communication with wireless clients and communication between Ethernet and DECT (via a bridge).

BIBLIOGRAPHY

1. D. Bertsekas and R. Gallager, *Data Networks,* 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1987.
2. F. E. Ross, An overview of FDDI: The fiber distributed data interface, *IEEE J. Selected Areas Commun.,* **7**, 1043–1051, 1989.
3. C. F. Hemrick et al., Switched multi-megabit service and early availability via MAN technology, *IEEE Commun. Mag.,* **26** (4): 9–14, 1988.
4. IEEE Standard 802.6-1991, Distributed Queue Dual Bus (DQDB) Subnetwork of MAN.
5. *The X/Open CPI-C Specification, Version 2,* 2nd ed., Professional Technical Reference, Englewood Cliffs, NJ: Prentice-Hall, 1996.
6. IEEE Computer Society, IEEE Draft 802.1 part A, *Overview and architecture,* October 1990.
7. C. Huitema, *Routing in the Internet,* Englewood Cliffs, NJ: Prentice-Hall, 1991.
8. U. Black, *Data Networks: Concepts, Theory and Practice,* Englewood Cliffs, NJ: Prentice-Hall, 1989.
9. R. Handel, M. Huber, and S. Schroder, *ATM Networks: Concepts, Protocols, Applications,* 2nd ed., Reading, MA: Addison-Wesley, 1994.

Reading List

F. Halsall, *Data Communications, Open Networks, and Open Systems,* 4th ed., Reading, MA: Addison-Wesley, 1996.

G. S. Hura and M. Singhal, *Data and Computer Communications: Networking and Internetworking,* London: Oxford Univ. Press, 1999.

W. Stallings, *Data and Computer Communications,* 5th ed., Englewood Cliffs, NJ: Prentice-Hall, 1987.

A. Tanenbaum, *Computer Networks,* 3rd ed., Englewood Cliffs, NJ: Prentice-Hall, 1997.

AJAY D. KSHEMKALYANI
University of Cincinnati

MUKESH SINGHAL
The Ohio State University

COMPUTERS, ANALOG. See ANALOG COMPUTER CIRCUITS; ANALOG COMPUTERS.

COMPUTERS AND SOCIETY. See SOCIAL AND ETHICAL ASPECTS OF INFORMATION TECHNOLOGY.