Internetworking LANs and WANs: Concepts, Techniques and Methods.
Second Edition. Gilbert Held

Print ISBN 0-471-97514-1 Online ISBN 0-470-84155-9

1

NETWORK CONCEPTS

In this introductory chapter, we will focus our attention upon the key concepts behind the construction of wide area networks (WANs) and local area networks (LANs). In doing so we will first examine each type of network to obtain an understanding of its primary design goal. Next, we will compare and contrast their operation and utilization as well as examine the primary *de facto* and *de jure* standards that govern the operation of different types of networks. As this is an introductory chapter, we will cover LAN and WAN networking concepts without concern for specific details which are presented in later chapters in this book.

1.1 WIDE AREA NETWORKS

The evolution of wide area networks can be considered to have begun in the mid to late 1950s, commensurate with the development of the first generation of computers. Based upon the use of vacuum tube technology, the first generation of computers were physically relatively large, power-hungry devices whose placement resulted in a focal point for data processing and the coining of the term 'data center'.

1.1.1 Computer-communications evolution

Originally, access to the computational capability of first generation computers was through the use of punched cards. After an employee of the organization used a keypunch to create a deck of cards, that card deck was submitted to a window in the data center, typically labeled input/output (I/O) control. An employee behind the window would accept the card deck and complete a form which contained instructions for running the submitted job. The card deck and instructions would then be sent to a person in production control who would schedule the job and turn it over to operations for execution at a predefined time. Once the job was completed, the card deck and any resulting output would be sent back to I/O control, enabling the job originator to return to the window in the data center to retrieve his or her card deck and the resulting output. With a little bit of luck, programmers might see the results of their efforts on the same day that they submitted their job. Since the computer represented a considerable financial investment for most organizations, it was understandable that they would be receptive to methods that would enable an extension of access to its computational capability. By the mid-1960s, several computer manufacturers had added remote access capabilities to one or more of their computers.

Remote batch transmission

One method of providing remote access was obtained by the installation of a batch terminal at a remote location. That terminal was connected via a telephone company supplied analog leased line and a pair of modems to the computer in the corporate data center.

The first type of batch terminal developed to communicate with a data center computer contained a card reader, printer, serial communications adapter, and hard-wired logic in one housing. The serial communications adapter converted the parallel bits of each internal byte read from the card reader into a serial data stream for transmission. Similarly, the adapter performed a reverse conversion process by converting a sequence of received serial bits into an appropriate number of parallel bits to represent a character internally within the batch terminal. Since the batch terminal was located remotely from the data center, it was often referred to as a remote batch terminal, while the process of transmitting data was referred to as remote batch transmission. In addition, the use of a remote terminal as a mechanism to group a number of card decks representing individual jobs to be executed at the remote data center resulted in the term 'remote job entry terminal' being used as a synonym to reference this device.

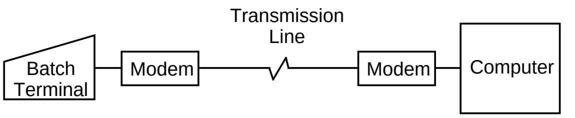

Figure 1.1 Remote batch transmission. The transmission of data from a remote batch terminal represents one of the first examples of wide area data communications networks

Figure 1.1 illustrates in schematic form the relationship between a batch terminal, transmission line, modems, and the data center computer. Since the transmission line connected a remote batch terminal in one geographic area to a computer located in a different geographic area, Figure 1.1 represents one of the earliest types of wide area data communications networks (WAN).

Paralleling the introduction of remote batch terminals was the development of a series of terminal devices, control units, and specialized communications equipment which resulted in the rapid expansion of interactive computer applications. One of the most prominent collections of products was introduced by the IBM Corporation under the trade name 3270 Information Display System.

IBM 3270 Information Display System

The IBM 3270 Information Display System was a term used to originally describe a collection of products ranging from interactive terminals, referred to as display stations that communicate with a computer, through several types of control units and communications controllers. Later, through the introduction of additional communications products from IBM and numerous third party vendors and the replacement of previously introduced products, the IBM 3270 Information Display System became more of a networking architecture and strategy rather than a simple collection of products. First introduced in 1971, the IBM 3270 Information Display System was designed to extend the processing power of the data center computer to remote locations. Since the data center computer typically represented the organization's main or primary computer, the term 'mainframe' was coined to refer to a computer with a large processing capability. As the mainframe was primarily designed for data processing, its utilization for supporting communications degraded its performance.

Communications controller

To offload communications functions from the mainframe, IBM and other computer manufacturers developed hardware whose primary function was to sample communications lines for incoming bits, group bits into bytes, and pass a group of bytes to the mainframe for processing as well as performing a reverse function for data destined from the mainframe to remote devices. When first introduced, such hardware was designed using fixed logic circuitry and the resulting device was referred to as a communications controller. Later, minicomputers were developed to execute communications programs; the ability to change the functionality of communications support by the modification of software was a considerable enhancement to the capabilities of this series of products. Because both hard-wired communications controllers and programmed minicomputers performing communications offloaded communications processing from the mainframe, the term 'front-end processor' evolved to refer to this category of communications equipment. Although most vendors refer to a minicomputer used to offload communications processing from the mainframe as a front- end processor, IBM has retained the term communications controller', even though their fixed logic hardware products were replaced over 20 years ago by programmable minicomputers.

Control units

To reduce the number of controller ports required to support terminals as well as the cabling between controller ports and terminals, IBM developed 'poll and select' software to support its 3270 Information Display System. Doing so enabled the communications controller to transmit messages from one port that could be destined to one or more terminals in a predefined group of devices. To share the communications controller port IBM developed a product called a control unit which acts as an interface between the communications controller and a group of terminals.

In general terms, the communications controller transmits a message to the control unit. The control unit examines the terminal address and retransmits the message to the appropriate terminal connected to the control unit. Thus, control

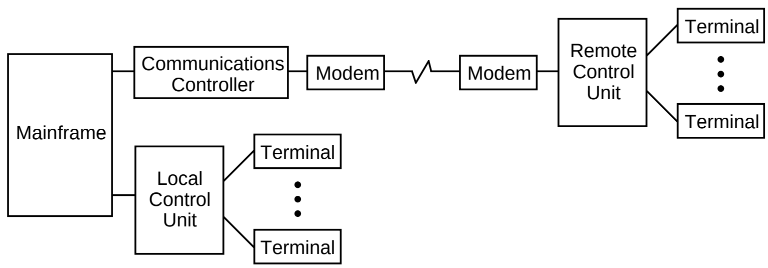

Figure 1.2 Relationship of 3270 Information Display products

units can be considered as devices which economize on the number of lines required to link display stations to mainframe computers. Both local and remote control units are available, with the key difference between the two primarily pertaining to the method of attachment to the mainframe computer and the use of intermediate devices between the control unit and the mainframe. Local control units are usually attached to a channel on the mainframe, whereas remote control units are connected to the mainframe's front-end processor, which is also known as a communications controller in the IBM environment. Since a local control unit is within a limited distance of the mainframe, no intermediate communications devices, such as modems, are required to connect a local control unit to the mainframe. In comparison, a remote control unit can be located in another building or in a different city, and it normally requires the utilization of intermediate communications devices, such as a pair of modems, for communications to occur between the control unit and the communications controller. The relationship of local and remote control units to display stations, mainframes, and a communications controller is illustrated in Figure 1.2.

1.1.2 Modern mainframe access

The introduction of the original IBM PC in 1981 resulted in the gradual displacement of terminals by PCs that provided a local processing capability and used coaxial cable adapters and special terminal emulation software to support the use of a personal computer to access a mainframe. Within a few years Ethernet and Token-Ring LANs were being used in organizations to provide client-server processing; however, client work-

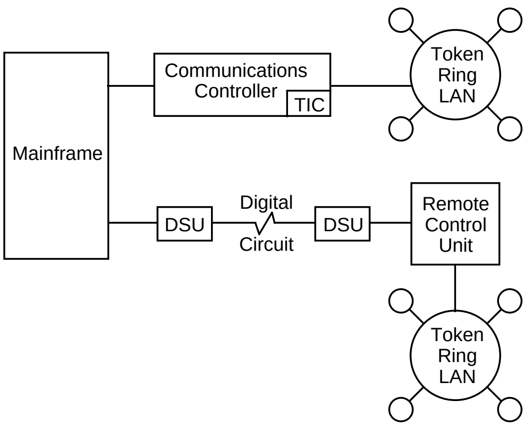

Figure 1.3 Initial methods developed to link 3270 networks to LANs

stations still required access to corporate mainframes. At first the primary method of linking LAN-based workstations to a mainframe was accomplished through the use of special adapter cards installed in communications controllers and control units. Figure 1.3 illustrates the initial method used to provide the integration of LAN-based workstations into a 3270 type network, and it is still in use by many organizations.

In the upper portion of Figure 1.3 a Token-Ring Interface Coupler (TIC) is shown installed in the communications controller which allows it to become a participant on a Token-Ring LAN. Under the communications controller a pair of Data Service Units (DSUs) which can be considered to represent digital modems required for transmission on digital circuits enable a remote control unit to be connected at a high data rate to the mainframe. The remote control unit uses a network adapter card similar to the TIC in the communication controller to become a participant on the remote Token-Ring network.

In addition to special adapter cards allowing communications controllers and control units to become participants on a LAN, special software is required on each workstation. Another key difference between the initial 3270 network and more modern LAN connections to mainframes is the reliance of the latter upon digital circuits that operate at 56 Kbps and above. In comparison, the initial 3270 networks were primarily constructed using analog transmission facilities limited to data rates up to 19.2 Kbps.

As LANs proliferated, a specialized communications device known as a router was developed to transport data between

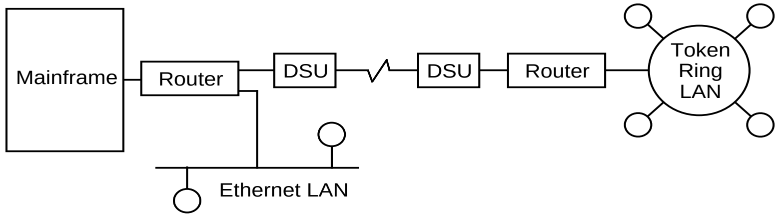

Figure 1.4 Modern LAN workstations' access to mainframes

networks. Routers were also developed with a channel interface to mainframes, enabling router-based networks and LAN workstations to replace the use of communications controllers and control units as illustrated in Figure 1.4.

In the example shown in Figure 1.4, a router with a mainframe interface is shown replacing both the communications controller and control unit. By supporting local and wide area interfaces, a single router can be used to provide both local and remote LAN access. As we will note later in this book, through the use of routers and leased lines we can create complex private networks or use routers with public networks to move traffic between LANs and from workstations on LANs to distant mainframes.

In addition, the basic operation of a router, which is to move data originating on one network and destined to another network off the source network, provides the routing capability which the device obtains its name. As we will note later in this book, the router is a key communications device which enables geographically separated LANs to be interconnected and whose operation enables the network of interconnected networks known as the Internet to function as intended.

Regardless of the method used to link geographically separated locations to one another, data are transported over wide area networks. Thus, let us turn our attention to WAN network construction and the general characteristics of this type of network.

1.1.3 Network construction

To provide batch and interactive access to the corporate mainframe from remote locations, organizations began to build sophisticated networks. At first, communications equipment such as modems and transmission lines was only obtainable from AT & T and other telephone companies. Commencing in 1974 in the United States with the well-known Carterphone decision, competitive non-telephone company sources for the supply of communications equipment became available. The divestiture of AT & T during the 1980s and the emergence of many local and long distance communications carriers paved the way for networking personnel being able to select from among two to hundreds of vendors for transmission lines and communications equipment.

As organizations began to link additional remote locations to their mainframes, the cost of providing communications began to escalate rapidly. This in turn provided the rationale for the development of a series of line sharing products referred to as multiplexers and concentrators. Although most organizations operated separate data and voice networks, beginning in the mid-1980s communications carriers began to make available for commercial use high-capacity circuits known as T1 in North America and E1 in Europe. Through the development of T1 and E1 multiplexers, voice, data, and video transmission can share the use of common high-speed circuits. Since the interconnection of corporate offices via the use of communications equipment and facilities normally covers a wide geographical area outside the boundary of one metropolitan area, the resulting network is known as a wide area network.

Figure 1.5 illustrates an example of a wide area network spanning the continental United States. In this example, regional offices in San Francisco and New York are connected with the corporate headquarters located in Atlanta via the use of T1 multiplexers and T1 transmission lines operating at 1.544 Mbps. Assuming each T1 multiplexer is capable of supporting the direct attachment of a private branch exchange (PBX), both voice and data are carried by the T1 circuits between the two regional offices and corporate headquarters. In addition, the San Francisco and Atlanta location have LANs that require communications connectivity with one another. Rather than install a separate leased line to interconnect the two locations, routers are connected to input ports on the T1 multiplexers in each location. This enables the T1 circuit between San Francisco and Atlanta to support inter-LAN connectivity as well as voice and data terminal traffic from tail circuits which we will shortly discuss. Thus, the three T1 circuits can be considered as the primary data highway or backbone of the corporate network.

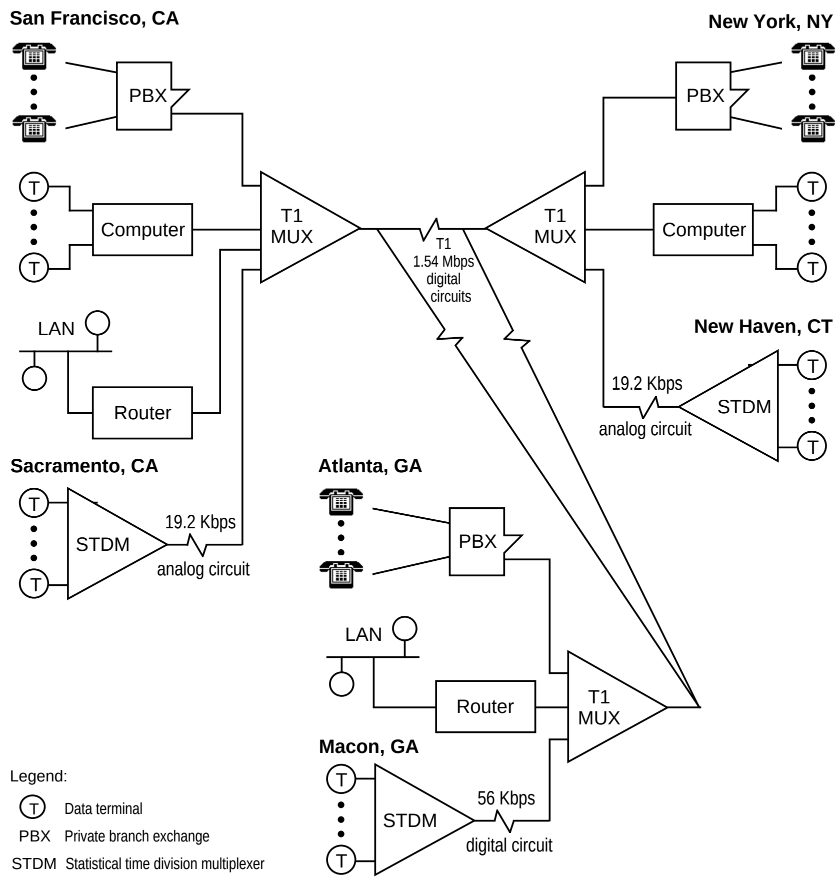

Figure 1.5 Wide area network example. A wide area network uses telecommunications lines obtained from one or more communications carriers to interconnect geographically dispersed locations

In addition to the three major corporate sites that require the ability to route voice calls and data between locations, let us assume that the corporation also has three smaller area offices located in Sacramento, CA; Macon, GA; and New Haven, CT. If those locations only require data terminals to access the corporate network for routing to the computers located in San Francisco and New York, one possible mechanism to provide network support is through the installation of tail circuits. Those tail circuits could be used to connect a statistical time division multiplexer (STDM) in each area office serving a group of data terminals to the nearest T1 multiplexer, using either analog or digital circuits. The T1 multiplexer would then be configured to route data terminal traffic over the corporate backbone portion of the network to its destination.

1.1.4 Network characteristics

There are certain characteristics we can associate with wide area networks. First, the WAN is typically designed to interconnect two or more dispersed geographical areas. This interconnection is accomplished by the lease of transmission facilities from one or more communications vendors. Secondly, most WAN transmission occurs at or under a data rate of 1.544 Mbps or 2.048Mbps, which are the operating rates of T1 and E1 transmission facilities. However, it should be noted that other popular WAN transmission facilities include T3 circuits that operate at approximately 45 Mbps and fractional T3 that operates at fractions of the T3 operating rate. In addition, communications carriers have deployed ATM as a transport mechanism that can be accessed at data rates from the T1 1.544 Mbps to direct ATM operating rates ranging up to 622 Mbps.

A third characteristic of WANs concerns the regulation of transmission facilities used for their construction. Most, if not all, transmission facilities marketed by communications carriers are subject to a degree of regulation at the federal, state, and possibly at the local government level. Even though we now live in an era of deregulation, carriers must seek approval for many offerings prior to making new facilities available for use. In addition, although many of the regulatory controls governing the pricing of services were removed, the communications market is still not a truly free market. Thus, regulatory agencies at the federal, state, and local level still maintain a degree of control over both the offering of new services as well as the pricing of existing and new services.

1.2 LOCAL AREA NETWORKS

The origin of local area networks can be traced, in part, to IBM terminal equipment introduced during 1974. At that time, IBM introduced a series of terminal devices designed for use in transaction processing applications for banking and retailing. What was unique about those terminals was their method of connection, with a common cable that formed a loop being used to provide a communications path within a localized geographical area. Unfortunately, limitations in the data transfer rate, incompatibility between each IBM loop system, and other problems precluded the widespread adoption of this method of networking. The economics of media sharing and the ability to provide common access to a centralized resource were, however, key advantages that resulted in IBM and other vendors investigating the use of different techniques to provide a localized communications capability between different devices. In 1977, Datapoint Corporation began selling its Attached Resource Computer Network (Arcnet), considered by most persons to be the first commercial local area networking product marketed. Since then, hundreds of companies have developed local area networking products, and the installed base of terminal devices connected to such networks has exponentially increased until they now number in the tens of millions.

1.2.1 Comparison to WANs

Local area networks can be distinguished from wide area networks by geographic area of coverage, data transmission and error rates, ownership, government regulation, and data routing, and, in many instances, by the type of information transmitted over the network.

Geographic area

Concerning the geographic area of coverage, the name of each network provides a general indication of the scope of the area in which it can support the interconnection of devices. As its name implies, a LAN is a communications network which covers a relatively small local area. This small local area can range in scope from a department located on a portion of a floor in an office building to the corporate staff located on several floors in the building, to several buildings on the campus of a university.

Regardless of the LAN area of coverage, its geographic boundary will be primarily restricted by the physical transmission limitations of the local area network. Those limitations are primarily in the area of cable distance between devices connected to the LAN and the total length of the LAN cable. In comparison, a wide area network can provide communications support to an area ranging in size from a town or city to a state, country, or even the entire world. Here the major factor governing transmission is the availability of communications facilities in different geographic areas that can be interconnected to route data from one location to another.

To better grasp the differences between LANs and WANs, we can view the LAN as being analogous to our local telephone or cable TV company, while the WAN can be considered to be the long distance communications carrier. Then, communications support in different cities is provided by the local telephone company in each city. However, for calls between cities, the local telephone companies must interconnect to the long distance carrier. Similarly, we can have separate LANs in different cities or within different buildings in the same city; however, to interconnect those LANs we would require the use of a wide area network.

Data transmission and error rates

Two additional areas that both differentiate LANs from WANs as well as explain the physical limitation of the LAN geographic area of coverage are the data transmission and error rate for each type of network. LANs normally operate at the low megabit-per-second rate, typically ranging from 4 Mbps to 16 Mbps, with relatively short distance (100 meter) copper wire based LANs and extended distance fiber optic-based LANs operating at 100 Mbps. At data rates above 100 Mbps ATM based LANs can operate at 155 Mbps and 622 Mbps. In comparison, the communications facilities used to construct a major portion of most WANs provide a data transmission rate at or under the T1 and E1 data rates of 1.544 Mbps and 2.048 Mbps. Although most communications carriers now operate on ATM infrastructure the cost of access to that infrastructure at data rates above T1 or E1 rates can be prohibitive for many organizations as carriers will have to install high speed access lines to subscribers. Due to this, most organizations use T1 or E1 lines to connect their ATM networks via WANs; however, within a few years it is probably reasonable to expect the cost of ATM WAN transmission to decrease as the ATM infrastructure is extended to corporate access points via fiber into buildings, enabling greater use of high speed end-to-end ATM transmission.

Since LAN cabling is primarily within a building or extends over a small geographical area, it is less susceptible to the impairments of nature, such as thunderstorms, lightning, and other acts of God. This in turn enables transmission at a relatively high data rate resulting in a relatively low error rate. In comparison, since wide area networks are based upon the use of communications facilities that are much more distant in length and always exposed to the elements, they have a much higher probability of being disturbed by changes in the weather, or electronic emissions generated by equipment, as well as such unforeseen problems as construction workers tearing up a street or paving a highway and accidentally causing harm to a communications cable. It is due to the greater exposure to the elements, the higher probability of accidental eruptions, and the greater change of electrical interference that the error rate on WANs is considerably higher than the rate experienced on LANs. On most WANs you can expect to experience an error rate of 1 in a million to 1 in 10 million (1×10^6) to 1×10^7) bits. In comparison, the error rate on a typical LAN may exceed that range by one or more orders of magnitude, resulting in an error rate of 1 in 10 million to 1 in 100 million bits.

Ownership

The construction of a wide area network requires the leasing of transmission facilities from one or more communications carriers. Although your organization can elect to purchase or lease communications equipment, the transmission facilities used to connect diverse geographical locations are owned by the communications carrier. In comparison, an organization which installs a local area network normally owns all of the components used to form the network, including the cabling used to form the transmission path between devices.

Regulation

Since wide area networks require transmission facilities that may cross local, state, and national boundaries, they may be subject to a number of governmental regulations at the local, state, and national level. Most of those regulations will govern the services that communications carriers can provide customers and the rates they can charge for those services, with the latter referred to as a tariff. In comparison, regulations affecting local area networks are primarily in the areas of building codes.

Such codes regulate the type of wiring that can be installed in a building and whether or not the wiring must run in a conduit.

Data routing and topology

In a local area network, data are routed along a path which defines the network. That path is normally a bus, ring, tree, or star structure and data always flow on that structure. The topology of a wide area network can be much more complex than the network structure of a local area network. In fact, many wide area networks may resemble a mesh structure, with equipment used to reroute data in the event of the failure of a communications circuit or equipment, or when too much traffic flows between two locations. Thus, the data flow on a wide area network can dynamically change, while the data flow on a local area network primarily follows the same basic route.

One exception to the preceding involves the use of LAN switches to include those that support Ethernet, Token-Ring, FDDI, and ATM. Through the use of such switches it becomes possible to construct local mesh structured networks that can become as complex as mesh structured WANs; however, unlike wide area networks that can span the globe, mesh structured LANs developed through the use of switches are restricted to transmission distances measured in hundreds of feet between switches.

Type of information carried

The last major difference between local and wide area networks concerns the type of information carried by each network. Many wide area networks support the simultaneous transmission of voice, data, and video information. In comparison, most local area networks are limited to carrying data. In addition, although all wide area networks can be expanded to transport voice, data, and video, many local area networks are restricted by design to the transportation of data. The one exception to the preceding is ATM that can operate from the desktop onto LANs and from LANs to WANs. Although ATM is in use by communications carriers to support voice, data, and video transmission in a WAN environment, it is currently used as a backbone switching mechanism and is limited to transporting data. In the future as prices decline we can reasonably expect ATM acceptance for use at the desktop which will enable voice, data, and video support

Table 1.1 Comparing LANs and WANs

Characteristic	Local Area Network	Wide Area Network
Geographic area of coverage	Localized to a building, group of buildings, or campus	Can span an area ranging in size from a city to the globe
Data transmission rate	Typically 4 Mbps to 16 Mbps, with some limited distance copper pair based and extended distance fiber optic-based networks operating at 100 Mbps	Normally operate at or below T1 and E1 transmission rates of 1.54 Mbps and 2.048 Mbps and T3 at 45 Mbps
Error rate	1 in 10^7 to 1 in 10^6	1 in 10^6 to 1 in 10^7
Ownership	Usually with the implementor	Communications carrier retains ownership of the facilities
Data routing	Normally follows fixed route or mesh structure for limited distance	Switching capability of network allows dynamic alteration of data flow for long distance
Topology	Usually limited to bus, ring, tree, and star	Virtually unlimited design capability
Type of information carried	Primarily data	Voice, data, and video commonly integrated

from LANs onto WANs. Table 1.1 summarizes the similarities and differences between local and wide area networks.

1.2.2 Utilization benefits

In its simplest form, a local area network can be considered as a cable that provides an electronic highway for the transportation of information to and from different devices connected to the network. By providing the capability to route data between devices connected to a common network within a relatively limited distance, numerous benefits can accrue to users of the network. Such benefits can include the ability to share the use of peripheral devices, obtain common access to data files and programs, communicate with other people on the LAN by electronic mail, and obtain access to the larger processing capability of mainframes or minicomputers through common gateways that link a local area network to larger computer systems. Here the gateway can be directly cabled to the mainframe or minicomputer if they reside at the same location or it may be connected remotely via the use of the corporate wide area network.

Peripheral sharing allows network users to access high speed color laser printers, CD-ROM jukebox systems, and other devices whose utilization may only be required a small portion of the time a workstation is in operation. Thus, users of a LAN can obtain access to resources that would probably be too expensive to justify for each individual workstation user.

The ability to commonly access data files and programs can substantially reduce the cost of software. In addition, shared access to database information allows network users to obtain access to updated files on a real-time basis.

One popular type of application program used on LANs enables users to transfer messages electronically. Commonly referred to as electronic mail, this type of application program can be used to supplement and, in many cases, eliminate the need for paper memorandums. Another of the more popular uses for LANs combines the ability to access data files and programs as well as the ability to transmit and receive electronic mail beyond the confines of the local area network. This is accomplished by connecting a LAN to the Internet, enabling each workstation on the network to access a variety of Internet sites and services through a common wide area network communications circuit routed to an Internet Service Provider.

For organizations with mainframe or minicomputers, a local area network gateway can provide a common method of access to those computers. In comparison, without the use of a LAN gateway, each personal computer requiring access to a mainframe or minicomputer would require a separate method of access. This might increase both the complexity of providing access as well as the cost of providing access.

1.3 STANDARDS ORGANIZATIONS AND THE OSI REFERENCE MODEL

The importance of standards and the work of standards organizations have proved essential for the growth of both local and worldwide communications. In the United States and many other countries, national standards organizations have defined physical and operational characteristics that enable vendors to manufacture equipment compatible with line facilities provided by communications carriers, as well as equipment produced by other vendors and which may be connected to a local or wide area network. At the international level, standards organizations have promulgated several series of communications-related recommendations. These recommendations, while not mandatory, have become highly influential on a worldwide basis for the development of equipment and facilities and have been adopted by hundreds of public companies and communications carriers.

In addition to national and international standards, a series of *de facto* standards has evolved through the licensing of technology among companies. Such *de facto* standards, as an example, have facilitated the development of communications software for use on personal computers. Today, consumers can purchase communications software that can control modems manufactured by hundreds of vendors since most modems are now constructed to respond to a core set of uniform control codes.

In this section we will focus our attention upon one national and three international standards organizations. The national standards organization we will discuss is the Institute of Electrical and Electronic Engineers (IEEE), whose work has been a guiding force in the rapid expansion in the use of local area networks due to a series of standards developed by that organization. In the international area, we will discuss the role of the International Telecommunications Union Telecommunications (ITU-T) standardization body, formerly known as the Consultative Committee for International Telephone and Telegraph (CCITT) and the International Standards Organization (ISO), both of which have developed numerous standards which facilitate the operation of wide area networks. In addition, we will discuss the role of the ATM Forum, a non-profit organization tasked with developing ATM standards.

Due to the importance of the ISO's Open Systems Interconnection (OSI) Reference Model in data communications, we will focus our attention upon the operation and general utilization of this model. This examination will provide an overview of the seven layers of the OSI model, and will serve as a foundation for a detailed investigation of several layers of that model presented in later chapters.

1.3.1 International standards organizations

Three important international standards organizations are the International Telecommunications Union Telecommunications

(ITU-T) standardization body formerly known as the Consultative Committee for International Telephone and Telegraph (CCITT), the International Standards Organization (ISO) and the ATM forum. The ITU-T can be considered as a governmental body as it functions under the auspices of an agency of the United Nations. Although the ISO is a non-governmental agency, its work in the field of data communications is well recognized. In comparison, the ATM Forum is an international non-profit organization that focuses its efforts on ATM issues to include standards.

ITU-T

The International Telecommunications Union Telecommunications (ITU-T) standardization body is a group within the International Telecommunications Union (ITU), the latter being a specialized agency of the United Nations headquartered in Geneva, Switzerland. The ITU-T is tasked with direct responsibility for developing data communications standards and consists of 15 Study Groups, each tasked with a specific area of responsibility.

The work of the ITU-T is performed on a four-year cycle which is known as a Study Period. At the conclusion of each Study Period, a Plenary Session occurs. During the Plenary Session, the work of the ITU-T during the previous four years is reviewed, proposed recommendations are considered for adoption, and items to be investigated during the next four-year cycle are considered.

The ITU-T's Tenth Plenary Session met in 1992 and its eleventh session occurred during 1996. Although approval of recommended standards is not intended to be mandatory, ITU-T recommendations have the effect of law in some Western European countries and many of its recommendations have been adopted by both communications carriers and vendors in the United States.

ISO

The International Standards Organization (ISO) is a non-governmental entity that has consultative status within the UN Economic and Social Council. The goal of the ISO is to

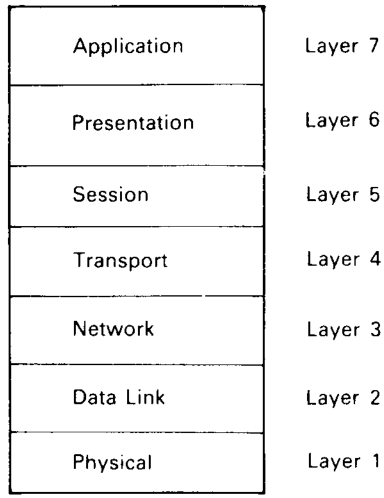

Figure 1.6 ISO Reference Model

'promote the development of standards in the world with a view to facilitating international exchange of goods and services'.

The membership of the ISO consists of the national standards organizations of most countries, with approximately 100 countries currently participating in its work.

Perhaps the most notable achievement of the ISO in the field of communications is its development of the seven-layer Open Systems Interconnection (OSI) Reference Model.

The ISO Reference Model

The International Standards Organization (ISO) established a framework for standardizing communications systems called the Open Systems Interconnection (OSI) Reference Model. The OSI architecture defines the communications process as a set of seven layers, with specific functions isolated and associated with each layer. Each layer, as illustrated in Figure 1.6, covers lower layer processes, effectively isolating them from higher layer functions. In this way, each layer performs a set of functions necessary to provide a set of services to the layer above it.

Layer isolation permits the characteristics of a given layer to change without impacting the remainder of the model, provided that the supporting services remain the same. One major advantage of this layered approach is that users can mix and match OSI conforming communications products to tailor their communications systems to satisfy a particular networking requirement.

The OSI Reference Model, while not completely viable with many current network architectures, offers the potential to directly interconnect networks based upon the use of different vendor equipment. This interconnectivity potential will be of substantial benefit to both users and vendors. For users, interconnectivity will remove the shackles that in many instances tie them to a particular vendor. For vendors, the ability to easily interconnect their products will provide them with access to a larger market. The importance of the OSI model is such that it has been adopted by the ITU-T as Recommendation X.200.

Layered Architecture

As previously discussed, the OSI Reference Model is based upon the establishment of a layered, or partitioned, architecture. This partitioning effort can be considered as being derived from the scientific process whereby complex problems are subdivided into functional tasks that are easier to implement on an aggregate individual basis than as a whole.

As a result of the application of a partitioning approach to communications network architecture, the communications process was subdivided into seven distinct partitions, called layers. Each layer consists of a set of functions designed to provide a defined series of services which relate to the mission of that layer. For example, the functions associated with the physical connection of equipment to a network are referred to as the physical layer.

With the exception of layers 1 and 7, each layer is bounded by the layers above and below it. Layer 1, the physical layer, can be considered to be bound below by the interconnecting medium over which transmission flows, while layer 7 is the upper layer and has no upper boundary. Within each layer is a group of functions which can be viewed as providing a set of defined services to the layer which bounds it from above, resulting in layer n using the services of layer $n-1$. Thus, the design of a layered architecture enables the characteristics of a particular layer to change without affecting the rest of the system, assuming the services provided by the layer do not change.

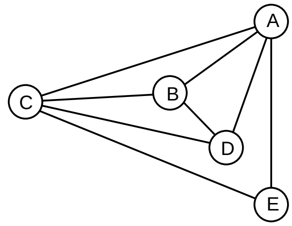

Figure 1.7 Network structure and components. The path between a source and destination node established on a temporary basis is called a logical connection

An understanding of the OSI layers is best obtained by first examining a possible network structure that illustrates the components of a typical wide area network. Figure 1.7 illustrates a network structure which is only typical in the sense that it will be used for a discussion of the components upon which networks are constructed.

The circles in Figure 1.7 represent nodes which are points where data enter or exit a network or is switched between two paths. Nodes are connected to other nodes via communications paths within the network where the communications paths can be established on any type of communications media, such as cable, microwave, or radio.

From a physical perspective, a node can be based upon the use of one of several types of computers, including a personal computer, minicomputer, mainframe computer, or specialized computer, such as a front-end processor gateway or router. Connections to network nodes into a wide area network can occur via the use of terminals directly connected to computers, computers connected to other computers, terminals connected to a node via the use of one or more intermediate communications devices, or paths linking one network to another network. Although terminals were originally fixed logic devices with little or no intelligence, the replacement of those devices by personal computers resulted in many persons continuing to refer to devices that interact with mainframes and minicomputers as terminals regardless of their processing capability. When used on a local area network we again modify terminology, referring to personal computers attached to LANs as workstations. When those workstations communicate with servers we refer to the communications method as client-server communications.

The routes between two nodes, such as C–E–A, C–D–A, C–A, and C–B–A, which could be used to route data between nodes A and C, are information paths. Due to the variability in the flow of information through a network, the shortest path between nodes may not be available for use or may represent a non-efficient path with respect to other paths constructed through intermediate nodes between a source and destination node. A temporary connection established to link two nodes whose route is based upon such parameters as current network activity is known as a logical connection. This logical connection represents the use of physical facilities, including paths and node switching capability on a temporary basis. Now that we have an appreciation of the general structure of networks, let us turn our attention to the layers of the OSI Reference Model.

The major functions of each of the seven OSI layers are described in the following seven paragraphs.

Layer 1 — the physical layer At the lowest or most basic level, the physical layer (level 1) is a set of rules that specifies the electrical and physical connection between devices. This level specifies the cable connections and the electrical rules necessary to transfer data between devices. Typically, the physical link corresponds to previously established interface standards, such as the RS-232/V.24 interface which governs the attachment of data terminal equipment, such as the serial port of personal computers, to data communications equipment, such as modems, at data rates below 19.2 Kbps. Included in the physical layer rules is the encoding method used to place data on the transmission medium. In a LAN environment common coding methods include Manchester coding, differential Manchester coding, and techniques that encode different numbers of data bits in a group of bits.

Layer 2 — the data link layer The next layer, which is known as the data link layer (level 2), denotes how a device gains access to the medium specified in the physical layer; it also defines data formats, including the framing of data within transmitted messages, error control procedures, and other link control activities. From defining data formats, including procedures to correct transmission errors, this layer becomes responsible for the reliable delivery of information. At the data link layer information is grouped into entities referred to as frames. As a minimum, each frame includes control information which enables the receiver to synchronize itself to an incoming frame, addressing information that identifies the source and destination, a field containing the actual information being transmitted from the source to the destination, and a field used for verifying the integrity of the data. Data link control protocols such as Binary Synchronous Communications (BSC) and High- level Data Link Control (HDLC) reside in this layer.

Since the development of OSI layers was originally targeted towards wide area networking, its applicability to local area networks required a degree of modification. Under the IEEE 802 standards, the data link layer was divided into two sublayers, logical link control (LLC) and media access control (MAC). The LLC layer is responsible for generating and interpreting commands which control the flow of data and perform recovery operations in the event of errors. In comparison, the MAC layer is responsible for providing access to the local area network, which enables a station on the network to transmit information. Later in this chapter we will examine the IEEE 802 standards in detail, including the functions and operation of the LLC and MAC layers.

Layer 3—the network layer The network layer (level 3) is responsible for arranging a logical connection between a source and destination on the network, including the selection and management of a route for the flow of information between source and destination based upon the available data paths in the network. Services provided by this layer are associated with the movement of data packets through a network, including addressing, routing, switching, sequencing, and flow control procedures. In a complex network, the source and destination may not be directly connected by a single path, but instead require a path to be established that consists of many subpaths. Thus, routing data through the network onto the correct paths is an important feature of this layer.

Several protocols have been defined for layer 3, including the ITU-T X.25 packet switching protocol and the ITU-T X.75 gateway protocol. X.25 governs the flow of information through a packet network, while X.75 governs the flow of information between packet networks. Other popular network layer protocols include the Internet Protocol (IP) and Novell's Internet Packet Exchange (IPX).

Layer 4—the transport layer The transport layer (level 4) is responsible for guaranteeing that the transfer of information occurs correctly after a route has been established through the network by the network level protocol. Thus, the primary function of this layer is to control the communications session between network nodes once a path has been established by the network control layer. Error control, sequence checking, and other end-to-end data reliability factors are the primary concern of this layer.

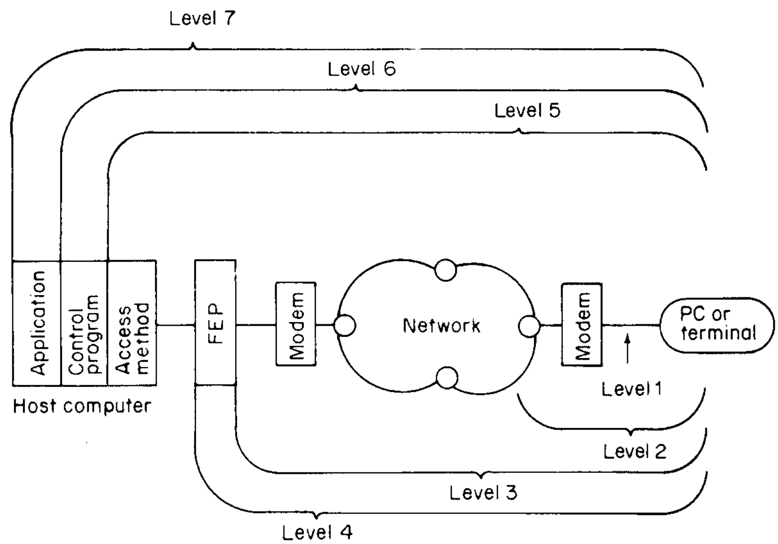

Figure 1.8 OSI model schematic

Two popular transport layer protocols are the Transport control Protocol (TCP) and the User Datagram Protocol (UDP).

Layer 5—the session layer The session layer (level 5) provides a set of rules for establishing and terminating data streams between nodes in a network. The services that this session layer can provide include establishing and terminating node connections, message flow control, dialogue control, and end-to-end data control.

Layer 6—the presentation layer The presentation layer (level 6) services are concerned with data transformation, formatting, and syntax. One of the primary functions performed by the presentation layer is the conversion of transmitted data into a display format appropriate for a receiving device. Data encryption/decryption and data compression and decompression are examples of the data transformation that could be handled by this layer.

Layer 7—the application layer Finally, the application layer (level 7) acts as a window through which the application gains access to all of the services provided by the model. Examples of functions performed at this level include file transfers, resource sharing, and database access. While the first four layers are fairly well defined, the top three layers may vary considerably, depending upon the network used. In fact, many protocols tend to group the higher

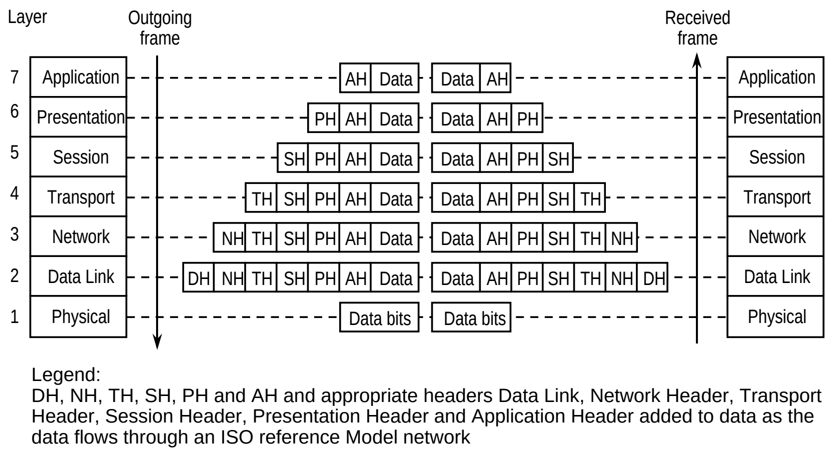

Figure 1.9 Data flow within an ISO Reference Model network

levels into one functional layer. Thus, such popular Internet protocols as the File Transfer Protocol (FTP), Telnet, and the HyperText Transport Protocol (HTTP) represent a blend of layer 5 through 7 functions. Later in this book we will examine the use of each of those Inernet protocols. Figure 1.8 illustrates the OSI model in schematic format, showing the various levels of the model with respect to a terminal accessing an application on a host computer system.

Data flow

As data flows within an ISO network each layer appends appropriate heading information to frames of information flowing within the network while removing the heading information added by a lower layer. In this manner, layer (n) interacts with layer ($n-1$) as data flow through an ISO network.

Figure 1.9 illustrates the appending and removal of frame header information as data flow through a network constructed according to the ISO Reference Model. Since each higher level removes the header appended by a lower level, the frame traversing the network arrives in its original form at its destination.

As the reader will surmise from the previous illustrations, the ISO Reference Model is designed to simplify the construction of data networks. This simplification is due to the eventual standardization of methods and procedures to append appropriate heading information to frames flowing through a network, permitting data to be routed to the appropriate destination following a uniform procedure.

1.3.2 IEEE

The Institute of Electrical and Electronic Engineers (IEEE) is a US-based engineering society that is very active in the development of data communications standards. In fact, the most prominent developer of local area networking standards is the IEEE, whose subcommittee 802 began its work in 1980 prior to the establishment of a viable market for the technology.

The IEEE Project 802 efforts are concentrated upon the physical interface of equipment and the procedures and functions required to establish, maintain, and release connections among network devices, including defining data formats, error control procedures, and other control activities governing the flow of information. This focus of the IEEE actually represents the lowest two layers of the ISO model, physical and link data.

IEEE 802 standards were originally limited to communications at data rates up to 20 Mbps within a local area up to 4 km, or a metropolitan area with a data path length up to 50 km. Additions to the 802 standards in the form of Fast Ethernet in all its flavors and 100Vg-AnyLAN resulted in specifications for data rates up to 100 Mbps, while work on a gigabit Ethernet can be expected to result in a specification for a LAN operating at 1000 Mbps.

802 Committees

Table 1.2 lists the organization of IEEE 802 committees involved in local area networks. In examining the lists of committees in Table 1.2 it is apparent that the IEEE early on noted that a number of different systems would be required to satisfy the requirements of a diverse end-user population. Accordingly, the IEEE adopted the CSMA/CD, Token Bus and Token-Ring as standards 802.3, 802.4 and 802.5, respectively.

The 802.1 committee responsible for Architecture and Overviews defines standards for enabling an IEEE 802 local or metropolitan area network station to communicate with another station on a different LAN or MAN or on a wide area network. Included under the 802.1 arena are such standards as 802.1D

Table 1.2 IEEE Series 802 committees

802.1	Architecture and Overview
802.2	Logical Link Control
802.3	CSMA/CD
802.3μ	Fast Ethernet
802.4	Token Bus
802.5	Token-Ring
802.6	Metropolitan Area Networks
802.7	Broadband Technical Advisory Group
802.8	Fiber Optic Technical Advisory Group
802.9	Integrated Voice and Data Networks
802.10	Network Security
802.11	Wireless LANs
802.12	100VG AnyLAN

which defines transparent bridging between two or more 802.X local area networks, 802.1H which defines Token-Ring bridging, and 802.li which defines the specifications for FDDI bridging.

The IEEE Committee 802 published draft standards for CSMA/CD and Token Bus local area networks in 1982. Standard 802.3, which describes a baseband CSMA/CD network similar to Ethernet, was published in 1983. Since then, several addenda to the 802.3 standard were adopted which govern the operation of CSMA/CD on different types of media. Those addenda include 10BASE-2 which defines a 10 Mbps baseband network operating on thin coaxial cable, 1BASE-5 which defines a 1 Mbps baseband network operating on twisted-pair, and 10BROAD-36, a broadband 10 Mbps network that operates on thick coaxial cable.

The next standard published by the IEEE was 802.4, which describes a token passing bus oriented network for both baseband and broadband transmission. This standard is similar to the Manufacturing Automation Protocol (MAP) standard developed by General Motors.

The third LAN standard published by the IEEE was based upon IBM's specifications for its Token-Ring network. Known as the 802.5 standard, it defines the operation of Token-Ring networks on shielded twisted-pair cable at data rates of 1 and 4 Mbps.

Once Token-Ring had been standardized for operation at 4 Mbps, the 802.5f committee specified changes to the 4 Mbps ring for operations at 16 Mbps. Recognizing the development of Fast Ethernet and a competitive technology that uses a demand priority access method instead of Ethernet's CSMA/CD protocol, the IEEE standardized both competitive methods after a

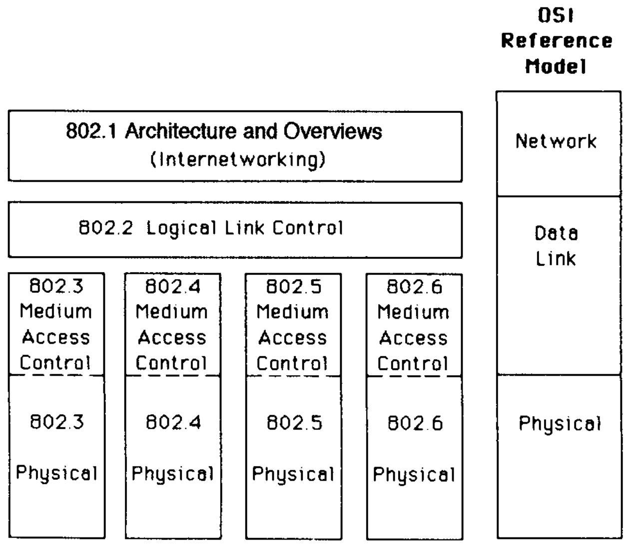

Figure 1.10 Relationship between IEEE standards and the OSI Reference Model

prolonged battle by various industry groups. As Fast Ethernet uses the CSMA/CD protocol similar to other versions of Ethernet, the resulting standard was promulgated as 802.3μ. The second 100 Mbps network, which also operates at 100 Mbps but uses a demand priority access protocol, was promulgated as 802.12. Since this standard supports either Ethernet or Token-Ring at 100 Mbps and operates on voice grade copper pair, its trade name became 100Vg-AnyLAN.

At the time this book was written the IEEE had approved Project Authorization Request (PAR) for the formation of a committee to develop standards for virtual LANs. Work was progressing under the direction of the 802.1Q committee, and an initial standard may be promulgated by the time you read this book.

Data link subdivision

One of the most interesting facets of IEEE 802 standards is the subdivision of the ISO Open Systems Interconnection Model's data link layer into two sublayers, logical link control and medium access control. This separation enables the mechanism for regulating access to the medium to be independent of the method for establishing, maintaining and terminating the

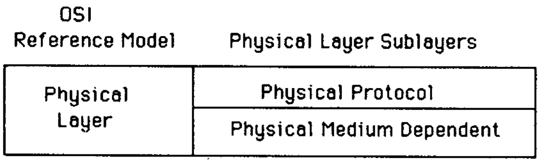

Figure 1.11 The physical layer is subdivided for high speed LANs and MANs developed to operate on different types of media

logical link between workstations. Here the method of regulating access to the medium is defined by the medium access control portion of each local area network standard. This enables the logical link control standard to be applicable to each type of network. Figure 1.10 illustrates the relationship between IEEE 802 local area network standards and the first three layers of the OSI Reference Model.

Physical layer subdivision

As local area networks gained acceptance it was soon recognized that the development of higher operating rate networks would require support for two or more distinct types of media such as unshielded twisted pair (UTP) and fiber, or even two or more types of fiber. This meant that a single physical layer would not be practical, and this resulted in the subdivision of the physical layer in a manner similar to the subdivision of the data link layer.

The actual method used to subdivide the physical layer is based upon the operating characteristics of the LAN or MAN and the media it will operate upon. One of the more common methods is to subdivide the physical layer into physical protocol and physical medium dependent sublayers as illustrated in Figure 1.11. The physical medium dependent layer will specify the type of connectors and their use with different types of media, while the physical protocol sublayer will define the signaling rate and method of encoding used with each type of physical medium dependent specification.

1.3.3 ATM Forum

The primary organization responsible for the development of ATM standards is the Technical Committee of the ATM Forum. The ATM Forum is a relatively new organization that was founded in 1991 as an international non-profit entity whose primary goal is to accelerate the use of ATM products and services via a rapid development of interoperability standards.

The ATM Forum's Technical Committee works with other standards bodies such as the American National Standards Institute (ANSI) and the International Telecommunications Union Telecommunications (ITU-T) standardization body to reconcile differences among standards as well as to recommend new standards when they are noted as being necessary or if existing ones are found inappropriate. Unlike the IEEE 802 committee which is tasked with LAN and MAN standards, the ATM Forum is responsible for both LAN and WAN standards, as ATM was developed as a universal technology that can run from an adapter in a workstation through LANs onto private and public WANs.