1

Overview of Data Communication Networks

1.1 Introduction

In a very broad sense, a network is any interconnected group of people or devices that are capable of sharing meaningful information with one another. In the telecommunication sense, a data communication network is a collection of two or more computing devices that are interconnected to enable them to share data. Data communication networking arose in response to the need to share data in a timely manner. Data sharing and information dissemination are critical to the success of any business. Thus, data communication networks are important to all contemporary organizations.

As discussed earlier, data communication networks deal with the transfer of data between two points. Data originates at the *source* and is finally delivered to the destination, which is also called a *sink*. Sometimes, the source and destination are interconnected by one link; at other times, the data must traverse multiple links to reach the destination. A typical communication environment includes multiple sources and sinks that are interconnected by communication links to build a network. Thus, a communication network is essentially an arrangement of hardware and software that allows users to exchange information.

1.2 Data Communication Network Model

A communication model is necessary to enable us to introduce the main elements of a communication system as well as to define some of the terminology used in the remainder of this book. A communication system consists of the following:

- A *source* that generates the information.
- A *source encoder* that converts the information into an electrical form called *message signal* $m(t)$.

Fundamentals of Data Communication Networks, First Edition. Oliver C. Ibe.

- A *transmitter* that is used to convert the message signal into a form acceptable to the channel.
- The *channel* is the path or link that connects the transmitter and the receiver; it can be metallic, optical fiber, or air.
- A *receiver* performs an inverse function of that of the transmitter to recover the message signal.
- A *source decoder* converts the electrical signal back to a form acceptable to the receiver.
- A *sink* is the user of the information generated by the source.

The model is illustrated in Figure 1.1.

Note that information flow can be bidirectional because what is a source at one time can be a sink at another time. Thus, Figure 1.1 shows the basic blocks used to process information as it flows in one direction.

The simplest data communication network consists of a source that is directly connected to a sink, as shown in Figure 1.2.

In a more complex network, the two communicating nodes are interconnected by a complex structure, which is usually represented by a cloud as shown in Figure 1.3.

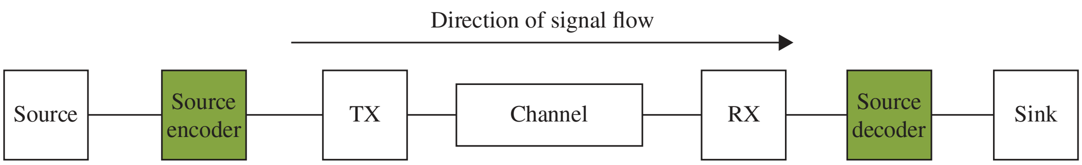

Figure 1.1 A Data Communication Network Model.

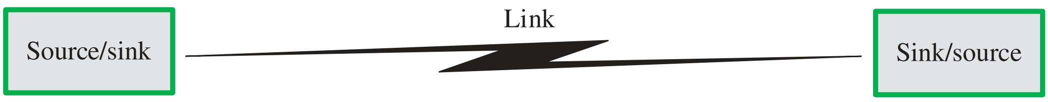

Figure 1.2 A Simple Data Communication Network.

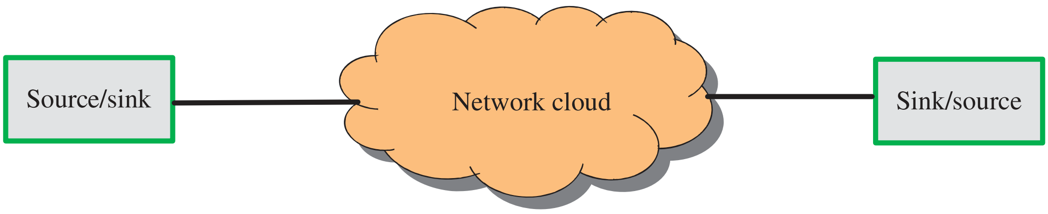

Figure 1.3 Representation of a Complex Network Structure.

1.3 Classification of Data Communication Networks

There are different ways to classify a data communication network, which are as follows:

- Transmission method
- Data flow direction
- Network topology
- Geographical coverage
- Transmission medium
- Data transfer technique
- Network access technique
- Media sharing technique.

In this section, we describe each of these methods.

1.3.1 Transmission Method

Data transmission method can be classified in two fundamental ways: *asynchronous* and *synchronous* transmissions. Asynchronous transmission is used when data is transmitted as individual characters. In this method, each character is preceded by one start bit and one or two stop bits that are used by the receiver for synchronization purposes. The need for synchronization arises from the fact that the interval between characters is random, which means the receiver that has been idle for some time needs to know when data is coming in.

Synchronous transmission is used to transmit large blocks of data at a time. In this scheme, data is usually organized in frames and each frame is preceded by a *flag* that consists of a few bits, and terminated by another flag. It is more efficient than asynchronous transmission because the overhead is smaller on a character-by-character basis. Figure 1.4 illustrates the difference between an asynchronous transmission scheme and a synchronous transmission scheme.

1.3.2 Data Flow Direction

Three ways are used to characterize the direction of data flow: *simplex*, *half duplex,* and *full duplex*. In a simplex transmission, data can only flow in one direction, which is usually from the source to the sink. This is illustrated in Figure 1.5.

In a half-duplex transmission (HDX) data can flow in both directions, but never simultaneously. It first flows in one direction, and then in the other direction. Thus, one station is the source and the other is the sink. Then the roles are interchanged such that the previous source becomes the sink and the previous sink becomes the source, and so on. This is illustrated in Figure 1.6.

In a full-duplex transmission (FDX), data can flow in both directions simultaneously. It can be viewed as a pair of simplex lines between the source and

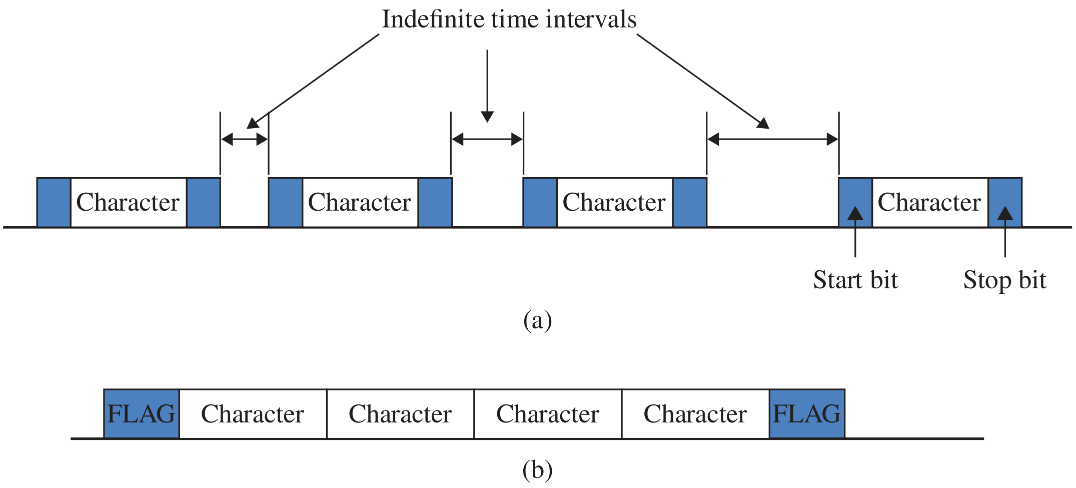

Figure 1.4 Asynchronous versus Synchronous Transmission. (a) Character-oriented asynchronous transmission and (b) frame-oriented synchronous transmission.

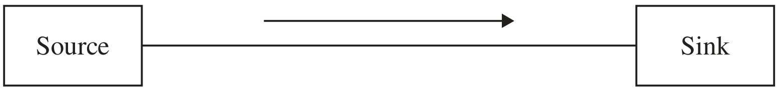

Figure 1.5 Simplex Transmission.

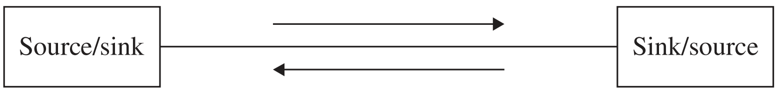

Figure 1.6 Half- Duplex Transmission.

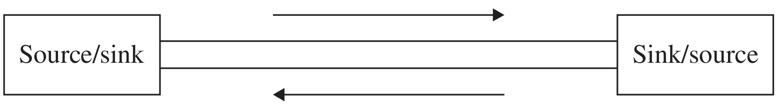

Figure 1.7 Full- Duplex Transmission.

sink with one line going from the source to the sink and the other going from the sink to the source. This is illustrated in Figure 1.7.

1.3.3 Network Topology

Network topology refers to the different geometrical configurations that can be used to build a network. Different topologies exist and include the following:

- Point-to-point (P2P)
- Point-to-multipoint
- Multidrop
- Bus
- Ring (or loop)
- Star

- Tree
- Mesh.

In the *P2P topology*, a link permanently connects two nodes or network devices. The P2P topology is illustrated in Figure 1.8. Note that the link interconnecting two nodes can be either a wireless (or air) or a wired connection.

In the *point-to-multipoint topology*, one node is connected to multiple nodes, each in a P2P manner, as illustrated in Figure 1.9.

In the *multidrop* topology, all nodes are interconnected by a single link with one node that is the master node and the other nodes are secondary or slave nodes. The master node usually controls access to the link and is located at one end of the link as illustrated in Figure 1.10.

The *bus* topology is similar to the multidrop topology with the exception that there is no master–slave relationship; all nodes are peers. The topology is illustrated in Figure 1.11.

Figure 1.8 Point-to-Point Topology.

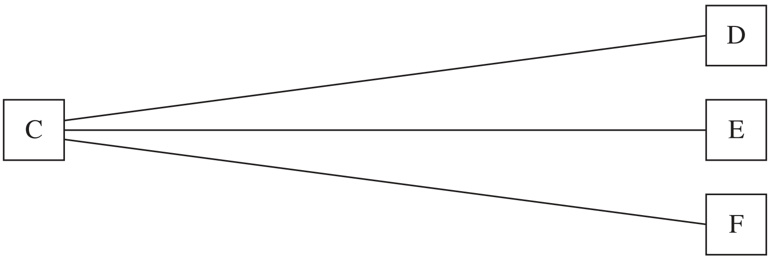

Figure 1.9 Point-to-Multipoint Topology.

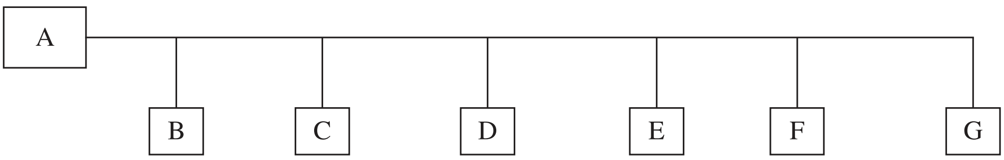

Figure 1.10 Mulitdrop Topology.

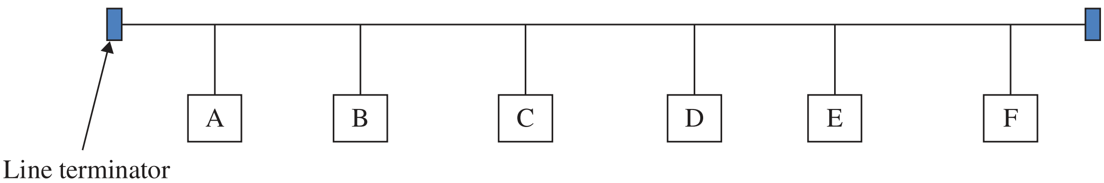

Figure 1.11 Bus Topology.

The *line terminator* in the figure is used to prevent a signal that comes to the end of a transmission line from bouncing back and corrupting other signals on the line. This “bouncing back” is called *signal reflection* and it has the tendency to interfere with and, therefore, corrupt the data on the line.

In *ring* topology, the nodes are connected serially in a P2P manner with the last node connected to the first node to form a loop. This is illustrated in Figure 1.12.

A *star* topology is a topology in which each node is connected in a P2P manner to a central node, called a *hub*. This is illustrated in Figure 1.13. Note that the star topology is similar to the point-to-multipoint topology. The difference between the two is that in the star topology, the hub is a passive device that does not control access to the network, while in the point-to-multipoint topology, the central node is an active device that controls communication in the network.

A *tree* topology is formed by connecting multiple buses together to form a system of branching links with no closed loop. It has a special node called the *headend* from which information flows to the other nodes. The topology is illustrated in Figure 1.14.

In the *mesh* topology, the network nodes are interconnected in an arbitrary manner. Generally, users are connected to only a subset of the nodes and another set of internal nodes provides a switching facility that moves data

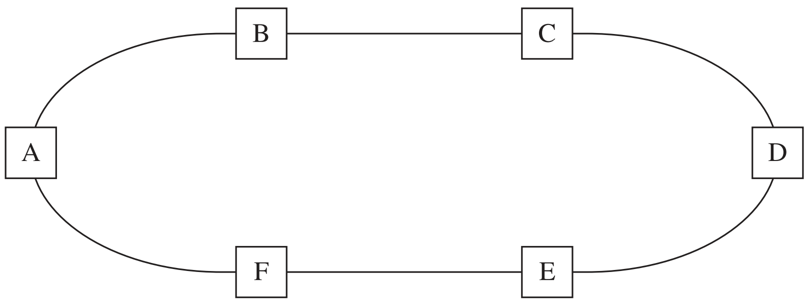

Figure 1.12 Ring Topology.

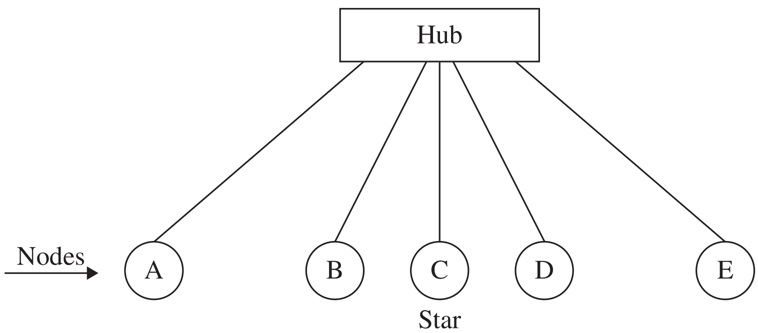

Figure 1.13 Star Topology.

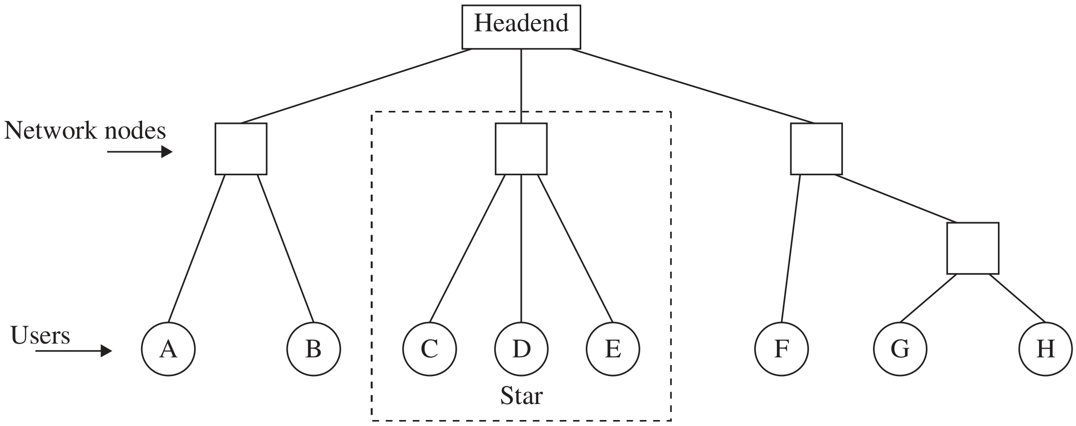

Figure 1.14 Tree Topology.

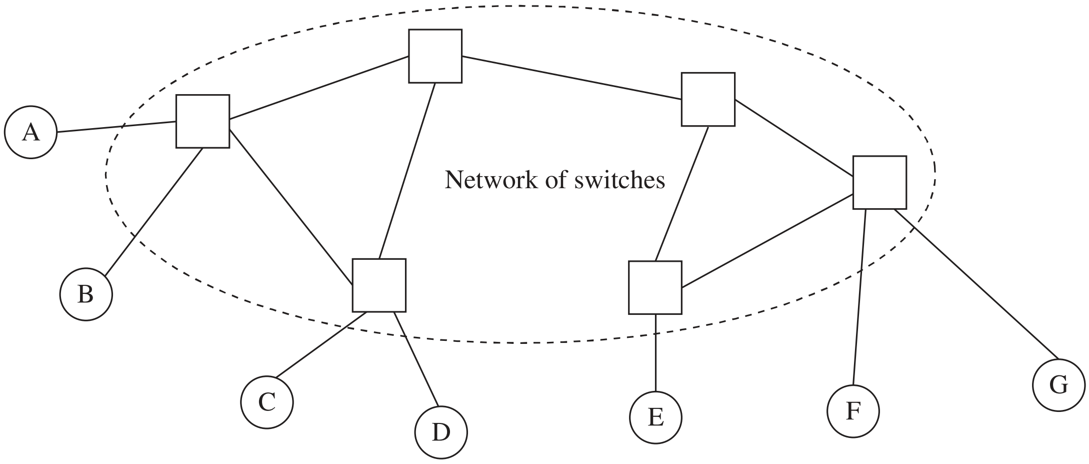

Figure 1.15 Mesh Topology.

from one node to another until it reaches its destination. An example of the mesh topology is shown in Figure 1.15.

1.3.4 Geographical Coverage

Networks are sometimes classified according to their geographical coverage. The following are examples of networks that are based on their geographical coverage:

- *Personal area networks* (PANs) are networks that interconnect devices that are within the reach of an individual, usually, within a range of 10 m. These devices are usually cell phones, tablets, and laptops.
- *Local area networks* (LANs) cover small geographical areas, typically one building, a floor, or a campus. Examples include the Ethernet and token ring networks.

- *Metropolitan area networks* (MANs) interconnect LANs in a campus or metropolitan area. Example includes the fiber distributed data interface (FDDI).
- *Wide area networks* (WANs) cover much larger areas such as a country (e.g., the public switched telephone network, or PSTN), or the globe (e.g., the Internet).

1.3.5 Transmission Medium

Data communication networks can also be classified according to the type of medium over which the signal propagates. In this case, there are two types of transmissions: *guided transmission* and *wireless transmission*; wireless transmission is also called *unguided transmission*.

In guided transmission, a physical path is provided along which the signal propagates. Guided transmission includes the twisted pair, coaxial cable, and optical fiber. In wireless transmission, the medium over which the signal propagates is mostly the air. Such networks use radio transmission.

1.3.6 Data Transfer Technique

There are two ways in which data can be transferred from source to destination: *switching* and *broadcasting*.

In a *switched network*, data is transferred from source to destination through a series of intermediate switching nodes. Data passes through a subset of the network nodes. There are two types of switched networks: *circuit-switched* networks and *packet-switched* networks. *Circuit switching* involves establishing a path from source to destination before the commencement of communication. The path is dedicated to the source–destination pair for the duration of communication session.

Packet switching involves organizing data in blocks called *packets* that are sent in a *store-and-forward* manner without prior establishment of communication path. By store-and-forward we mean that when a node receives a packet, it stores the packet and checks it for errors. If the packet is found to have an error, it is discarded. If it is found to be error-free, it is then scheduled for transmission to the next node on its way to the destination. There are two types of packet switching: *virtual circuit switching* and *datagram service*. Virtual circuit switching uses the same path for all packets belonging to the same session. Datagram service can use different paths for the different packets in a session.

In a *broadcast network*, a transmission from a source is received by all nodes in the network. Thus, unlike a switched network where data passes through only a subset of the nodes in the network, a broadcast network generally ensures that all the nodes in the network see the transmitted data.

1.3.7 Network Access Technique

There are two network access techniques that are closely related to the transfer technique used: *switched network access* and *broadcast network access*. Switched network access uses either circuit switching or packet switching. *Circuit switching* involves three phases:

- *Call setup phase* is used to establish communication path between the source and the sink.
- *Data transfer phase* is used to transmit the data after the path has been established.
- *Call teardown phase* is used to clear, tear down, or delete the communication path after the communication has been completed.

Packet switching sends packets of data into the network to be routed in a store-and-forward manner without prior establishment of the communication path.

Broadcast network access uses two access methods: *random access* in which users contend for control of the channel and *controlled access* where no contention is allowed. Controlled access uses one of two polling schemes: *centralized polling* (or *roll-call polling*) and *distributed polling*.

Centralized polling is used in master–slave systems where the master (or controller) uses a round-robin scheme to invite each station to transmit its data. Thus, no station can transmit until it is explicitly invited by the controller to transmit. *Distributed polling* generally uses a *token passing scheme* in peer-to-peer systems to control access. When a station receives the token, it transmits its data and after that it passes the token to the next node in a logical order. If a station that has no data to transmit receives the token, it simply passes the token to the next station in the logical order.

1.3.8 Media Sharing Technique

Some transmission media provide more capacity than one user can use. The utilization of such media can be increased by allowing multiple users to transmit their data simultaneously, or close to simultaneously. Three methods exist for sharing such media among the users, which are as follows:

(a) *Frequency-division multiplexing* (FDM) where the frequency spectrum of the medium is partitioned into multiple frequency blocks called *channels* that are assigned to users who can use these channels simultaneously without interference from each other.

(b) *Time-division multiplexing* (TDM) where transmission time is divided into nonoverlapping time slots that are assigned to users. Transmissions are staggered using a round-robin method to schedule the transmissions. Specifically, when it is time for a user to transmit, he uses the entire transmission medium for the duration of the time slot and then relinquishes

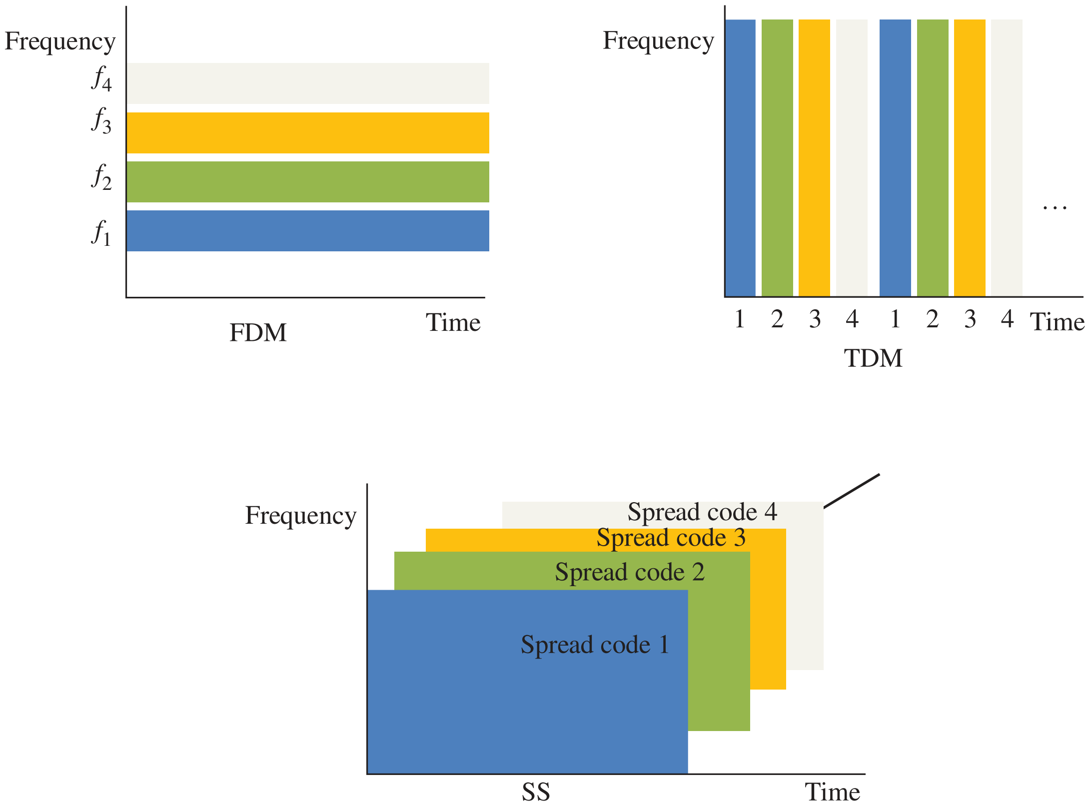

Figure 1.16 Media Shring Schemes.

control of the medium for the next user in the logical order. There are two types of TDM:

- *Synchronous TDM* where a time slot is dedicated to user whether or not the user has data to transmit.
- *Statistical multiplexing* (also called *asynchronous TDM*) where a user gets a slot only when he has data to transmit.

(c) *Spread spectrum* (SS) where the output signal, which appears like noise, occupies more bandwidth than original signal. There two types of SS:

- *Direct sequence* SS, where the frequency spectrum of a data signal is spread using a code that is uncorrelated with that signal and the codes are unique for every user and uncorrelated with other codes.
- *Frequency hopping* SS, where the transmitted frequency is pseudorandomly changed at a rate called the "hopping rate." The hopping pattern (i.e., the fixed order of frequencies that the user hops into) assigned to a user constitutes the channel for that user.

The different media sharing schemes are illustrated in Figure 1.16. They can be summarized as follows. In FDM, a user is assigned a part of the frequency spectrum that they use all the time. In TDM, the user is allowed to use the entire frequency spectrum but only part of the time. Finally, in SS, each user can use the entire spectrum all the time as long as they use a code that is uncorrelated to other codes to transmit their data.

1.4 Data Network Architecture

Because data communication deals with exchange of data messages between computers, transferring a message from one computer to another is not a trivial task as all conditions have to be anticipated and the necessary course of action exhaustively specified. To simplify intercomputer communication, the International Standards Organization (ISO) proposed a seven-layer architectural model called the *Open Systems Interconnection* (OSI) *Reference Model* for implementing data communication between cooperating systems. Each layer deals with a specific data communication function and provides services for the layer immediately above it while using the services of the layer immediately below it, except for the lowest layer that has no layer below it and the uppermost layer that has no layer above it.

One advantage of the model is that the implementation of one layer can change with technology without affecting the implementation of other layers as long as it provides the same services to the immediate upper layer as before. For example, we can change the network interface card (NIC) on a laptop from wired NIC to wireless NIC without affecting other layers; only the layer that is concerned with direct connection to the medium is affected.

1.4.1 The OSI Protocol Reference Model

In this book, we use the term "protocol" very often. A *protocol* is a set of rules that ensure the effective exchange of information. The seven-layer reference model is shown in Table 1.1.

- *Physical layer* defines the electrical and mechanical standards and signaling required to establish, maintain, and terminate connections. It deals with issues such as size and shape of connectors, signal strength, bit representation, and bit synchronization.

Table 1.1 The OSI model.

Layer 7	Application layer
Layer 6	Presentation layer
Layer 5	Session layer
Layer 4	Transport layer
Layer 3	Network layer
Layer 2	Data link layer
Layer 1	Physical layer

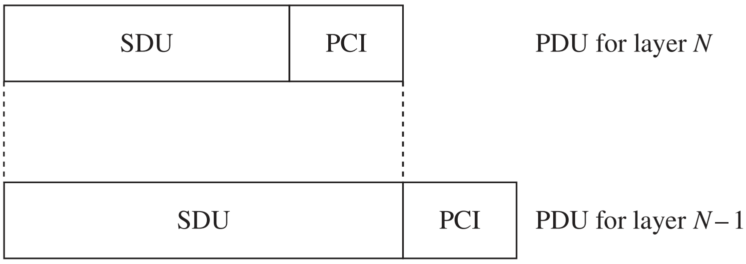

Figure 1.17 PDU Generation Process.

- *Data link layer* is responsible for organizing data in frames and for detecting errors that occur in a frame.
- *Network layer* is responsible for routing data to its destination and for providing unique addresses in the network.
- *Transport layer* is responsible for reliable end-to-end data transfer.
- *Session layer* is responsible for establishing, maintaining, and terminating sessions between applications.
- *Presentation layer* is responsible for translating data in a form that can be understood by the receiver.
- *Application layer* is responsible for providing services to end-user applications that lie outside the scope of the OSI model. It defines procedures by which end-user applications access network services.

At each layer, the protocol for that layer creates a *protocol data unit* (PDU) for transmission that includes the header information required by that protocol and the data to be transmitted. The header is called the *protocol control information* (PCI), and is the information exchanged between entities at a given layer. The PDU of layer N is passed down to layer $N-1$ where it becomes a unit called the *service data unit* (SDU) of that layer. Layer $N-1$ adds its PCI to the SDU to create its PDU that is passed down to layer $N-2$, and so on. The concept of PDU generation is illustrated in Figure 1.17.

Figure 1.18 shows how different layers add their overheads to an application until it gets down to the physical layer where it is forwarded to the next node on its way to the destination. Each layer, with the exception of the physical layer, appends its PCI that is manifested in the form of a layer header (LH). In addition to a header, layer 2 also appends a layer trailer (LT) to create a frame that is passed on to the physical layer. The physical layer interprets the frame as a sequence of zeros and ones that need to be transmitted over the medium.

1.4.2 The Internet Architecture

Internet architecture is also layered but has only four layers (or three, if the so-called *user process* layer is not included). The architecture is shown in Figure 1.19, where it is compared to the OSI model.