DATA COMMUNICATIONS

Data communications are the exchange of data between two devices via some form of transmission medium such as a wire cable. For data communications to occur, the communicating devices must be part of a communication system made up of a combination of hardware (physical equipment) and software (programs). The effectiveness of a data communications system depends on four fundamental characteristics: delivery, accuracy, timeliness, and jitter.

1. **Delivery.** The system must deliver data to the correct destination. Data must be received by the intended device or user and only by that device or user.
2. **Accuracy.** The system must deliver the data accurately. Data that have been altered in transmission and left uncorrected are unusable.
3. **Timeliness.** The system must deliver data in a timely manner. Data delivered late are useless. In the case of video and audio, timely delivery means delivering data as they are produced, in the same order that they are produced, and without significant delay. This kind of delivery is called real-time transmission.
4. **Jitter.** Jitter refers to the variation in the packet arrival time. It is the uneven delay in the delivery of audio or video packets. For example, let us assume that video packets are sent every 30 ms. If some of the packets arrive with 30-ms delay and others with 40-ms delay, an uneven quality in the video is the result.

COMPONENTS DATA COMMUNICATIONS

A data communications system has five components, as shown figure 1.1.

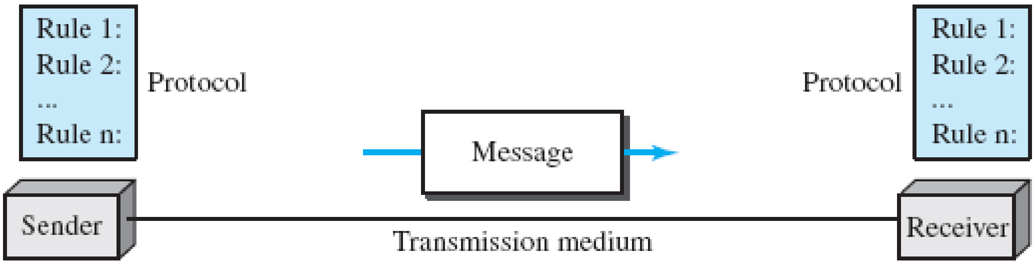


Figure 1.1 Components of Data Communication

1. **Message.** The **message** is the information (data) to be communicated. Popular forms of information include text, numbers, pictures, audio, and video.
2. **Sender.** The **sender** is the device that sends the data message. It can be a computer, workstation, telephone handset, video camera, and so on.
3. **Receiver.** The **receiver** is the device that receives the message. It can be a computer, workstation, telephone handset, television, and so on.
4. **Transmission medium.** The **transmission medium** is the physical path by which a message travels from sender to receiver. Some examples of transmission media include twisted-pair wire, coaxial cable, fiber-optic cable, and radio waves.
5. **Protocol.** A protocol is a set of rules that govern data communications. It represents an agreement between the communicating devices. Without a protocol, two devices may be connected but not communicating, just as a person speaking French cannot be understood by a person who speaks only Japanese.

DATA REPRESENTATION

1. Text

- In data communications, text is represented as a bit pattern, a sequence of bits (0s or 1s). The **American Standard Code for Information Interchange (ASCII),** developed some decades ago in the United States, now constitutes the first 127 characters in Unicode and is also referred to as **Basic Latin.** Appendix A includes part of the Unicode.

2. Numbers

- Numbers are also represented by bit patterns. However, a code such as ASCII is not used to represent numbers; the number is directly converted to a binary number to simplify mathematical operations.

3. Images

- **Images** are also represented by bit patterns. In its simplest form, an image is composed of a matrix of pixels (picture elements), where each pixel is a small dot.
- The size of the pixel depends on the resolution. For example, an image can be divided into 1000 pixels or 10,000 pixels. In the second case, there is a better representation of the image (better resolution), but more memory is needed to store the image.
- After an image is divided into pixels, each pixel is assigned a bit pattern. The size and the value of the pattern depend on the image. For an image made of only black-and-white dots (e.g., a chessboard), a 1-bit pattern is enough to represent a pixel. If an image is not made of pure white and pure black pixels, we can increase the size of the bit pattern to include gray scale. For example, to show four levels of gray scale, we can use 2-bit patterns. A black pixel can be represented by 00, a dark gray pixel by 01, a light graypixel by 10, and a white pixel by 11.
- There are several methods to represent color images. One method is called **RGB,** so called because each color is made of a combination of three primary colors: red, green,and blue. The intensity of each color is measured, and a bit pattern is assigned to it. Another method is called **YCM,** in which a color is made of a combination of three other primary colors: yellow, cyan, and magenta.

4. Audio

- **Audio** refers to the recording or broadcasting of sound or music. Audio is by nature different from text, numbers, or images. It is continuous, not discrete. Even when we use a microphone to change voice or music to an electric signal, we create a continuous signal.

5. Video

- **Video** refers to the recording or broadcasting of a picture or movie. Video can either be produced as a continuous entity (e.g., by a TV camera), or it can be a combination of images, each a discrete entity, arranged to convey the idea of motion.

DATA FLOW

Communication between two devices can be simplex, half-duplex, or full-duplex as shown in Figure 1.2.

1. Simplex

- In **simplex mode,** the communication is unidirectional, as on a one-way street. Only one of the two devices on a link can transmit; the other can only receive (see Figure 1.2a).
- Keyboards and traditional monitors are examples of simplex devices. The keyboard can only introduce input; the monitor can only accept output.
- The simplex mode can use the entire capacity of the channel to send data in one direction.

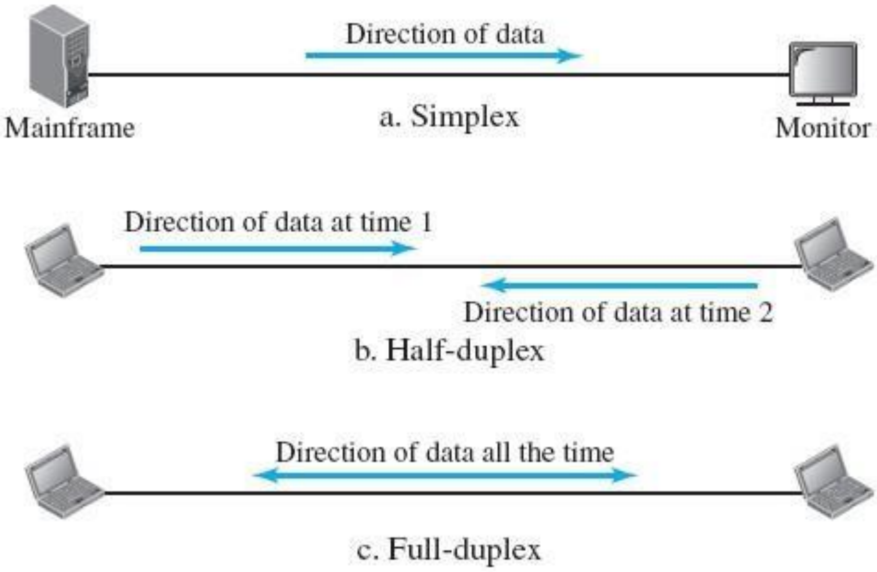

Direction of data
Mainframe
a. Simplex
Monitor
Direction of data at time 1
Direction of data at time 2
b. Half-duplex
Direction of data all the time
c. Full-duplex

Figure 1.2 Data Flow Mode

2. Half-Duplex

- In **half-duplex mode,** each station can both transmit and receive, but not at the same time. When one device is sending, the other can only receive, and vice versa.
- The half-duplex mode is like a one-lane road with traffic allowed in both directions. When cars are traveling in one direction, cars going the other way must wait.
- In a half-duplex transmission, the entire capacity of a channel is taken over by whichever of the two devices is transmitting at the time. Walkie-talkies and CB (citizens band) radios are both half-duplex systems.

3. Full-Duplex

- In **full-duplex mode** (also called duplex), both stations can transmit and receive simultaneously.
- The full-duplex mode is like a two-way street with traffic flowing in both directions at the same time. In full-duplex mode, signals going in one direction share the capacity of the link with signals going in the other direction.
- One common example of full-duplex communication is the telephone network. When two people are communicating by a telephone line, both can talk and listen at the same time. The full-duplex mode is used when communication in both directions is required all the time. The capacity of the channel, however, must be divided between the two directions.

NETWORKS

- A **network** is the interconnection of a set of devices capable of communication. In this definition, a device can be a **host** (or an end system as it is sometimes called) such as a large computer, desktop, laptop, workstation, cellular phone, or security system.
- A device in this definition can also be a **connecting device** such as a router, which connects the network to other networks, a switch, which connects devices together, a modem (modulator-demodulator), which changes the form of data, and so on.

Network Criteria

- A network must be able to meet a certain number of criteria. The most important of these are performance, reliability, and security.

1. Performance

- **Performance** can be measured in many ways, including transit time and response time. Performance is often evaluated by two networking metrics: **throughput** and **delay.**
- We often need more throughput and less delay. However, these two criteria are often contradictory. If we try to send more data to the network, we may increase throughput but we increase the delay because of traffic congestion in the network.

2. Reliability

- Network **reliability** is measured by the frequency of failure, the time it takes a link to recover from a failure, and the network's robustness in a catastrophe.

3. Security

- Network **security** issues include protecting data from unauthorized access, protecting data from damage and development, and implementing policies and procedures for recovery from breaches and data losses.

PHYSICAL STRUCTURE

Type of Connection

- A network is two or more devices connected through links. A link is a communications pathway that transfers data from one device to another.
- There are two possible types of connections: point-to-point and multipoint.

1. Point-to-Point

- A **point-to-point connection** provides a dedicated link between two devices.

- The entire capacity of the link is reserved for transmission between those two devices. Most point-to-point connections use an actual length of wire or cable to connect the two ends, but other options, such as microwave or satellite links, are also possible.

- When we change television channels by infrared remote control, we are establishing a point-to-point connection between the remote control and the television's control system.

2. Multipoint

- A **multipoint** (also called **multidrop**) **connection** is one in which more than two specific devices share a single link.
- In a multipoint environment, the capacity of the channel is shared, either spatially or temporally. If several devices can use the link simultaneously, it is a spatially shared connection. If users must take turns, it is a timeshared connection.

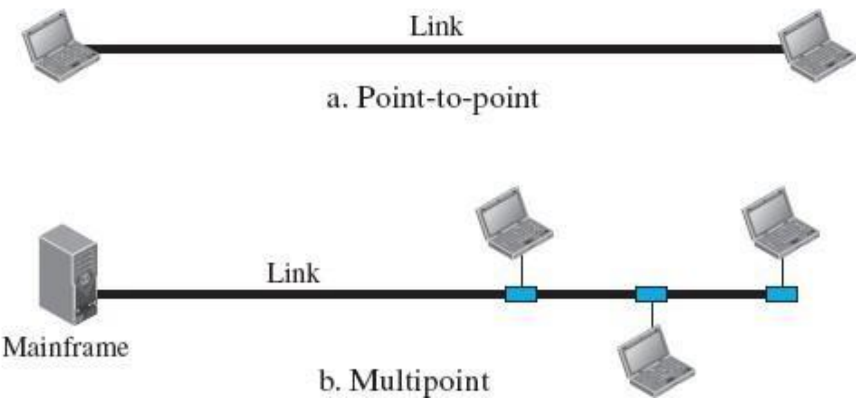


Figure 1.3 Type of Connection

- The term **physical topology** refers to the way in which a network is laid out physically. Two or more devices connect to a link; two or more links form a topology. The topology of a network is the geometric representation of the relationship of all the links and linking devices (usually called **nodes**) to one another.

1. Mesh Topology

- In a **mesh topology,** every device has a dedicated point-to-point link to every other device. The term dedicated means that the link carries traffic only between the twodevices it connects.
- In mesh topology we need n (n – 1) physical links if links are of type simplex and we need n (n – 1) /2 physical links if links are of type duplex.

Physical Topology

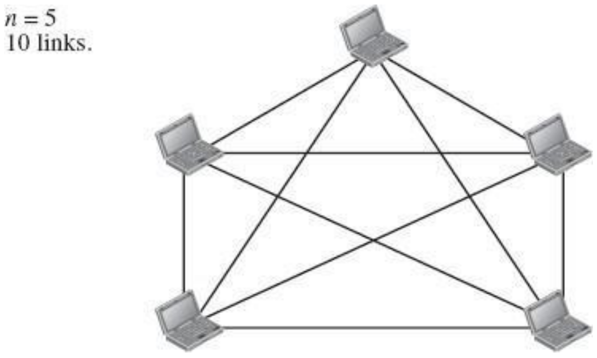


Figure 1.4 Mesh Topology

Advantages

- First, the use of dedicated links guarantees that each connection can carry its own data load.
- Second, a mesh topology is robust. If one link becomes unusable, it does not incapacitate the entire system. Third, there is the advantage of privacy or security.
- Finally, point-to-point links make fault identification and fault isolation easy.

Disadvantages

- First, because every device must be connected to every other device, installation and reconnection are difficult.
- Second, the sheer bulk of the wiring can be greater than the available space (in walls, ceilings, or floors) can accommodate.
- Finally, the hardware required to connect each link (I/O ports and cable) can be prohibitively expensive.

2. Star Topology

- In a star topology, each device has a dedicated point-to-point link only to a central controller, usually called a hub.
- The devices are not directly linked to one another. Unlike a mesh topology, a star topology does not allow direct traffic between devices.
- The controller acts as an exchange: If one device wants to send data to another, it sends the data to the controller, which then relays the data to the other connected device.

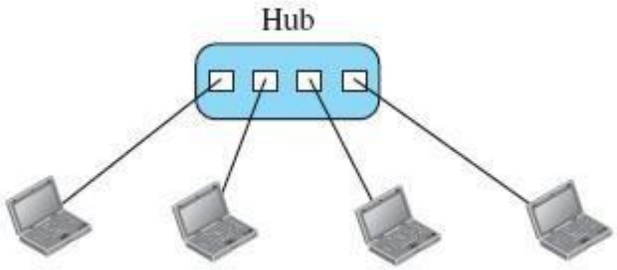


Figure 1.5 Star Topology

Advantages

- A star topology is less expensive than a mesh topology. As it needs farr less cable compared to mesh topology.
- In a star, each device needs only one link and one I/O port to connect it to any number of others. This factor also makes it easy to install and reconfigure.
- Robustness. If one link fails, only that link is affected. All other links remain active.
- Easy fault identification and fault isolation.

Disadvantages

- One big disadvantage of a star topology is the dependency of the whole topology on one single point, the hub. If the hub goes down, the whole system is dead.
- Although a star requires far less cable than a mesh, each node must be linked to a central hub.

3. Bus Topology

- The preceding examples all describe point-to-point connections. A **bus topology,** on the other hand, is multipoint. One long cable acts as a **backbone** to link all the devices in a network.
- Nodes are connected to the bus cable by drop lines and taps. A drop line is a connection running between the device and the main cable. A tap is a connector that either splices into the main cable or punctures the sheathing of a cable to create a contact with the metallic core.

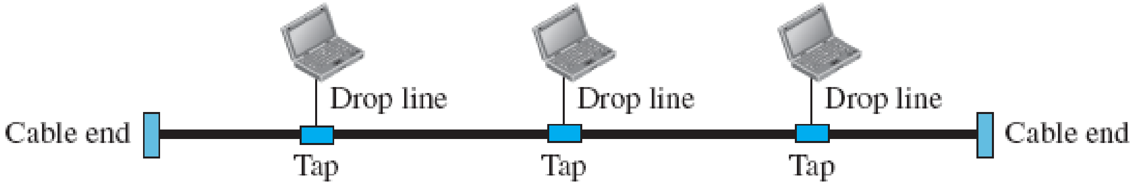


Figure 1.6 Bus Topology

Advantages

- Backbone cable can be laid along the most efficient path, then connected to the nodes by drop lines of various lengths. In this way, a bus uses less cabling than mesh or star topologies.
- In a star, for example, four network devices in the same room require four lengths of cable reaching all the way to the hub. In a bus, this redundancy is eliminated. Only the backbone cable stretches through the entire facility. Each drop line has to reach only as far as the nearest point on the backbone.
- Easy Installation.

Disadvantages

- Difficult reconnection and fault isolation.
- A bus is usually designed to be optimally efficient at installation. It can therefore be difficult to add new devices.
- A fault or break in the bus cable stops all transmission, even between devices on the same side of the problem.

4. Ring Topology

- In a **ring topology,** each device has a dedicated point-to-point connection with only the two devices on either side of it.
- A signal is passed along the ring in one direction, from device to device, until it reaches its destination. Each device in the ring incorporates a repeater. When a device receives a signal intended for another device, its repeater regenerates the bits and passes them along.

Advantages

- A ring is relatively easy to install and reconfigure.
- Fault Identification and isolation is simplified.

Disadvantages

- Unidirectional traffic can be a disadvantage. In a simple ring, a break in the ring can disable the entire network.

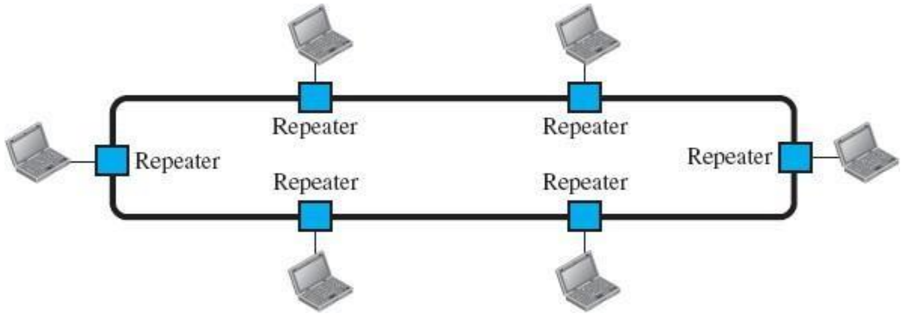


Figure 1.7 Ring Topology

NETWORK TYPES

- The criteria of distinguishing one type of network from another are difficult and sometimes confusing. We use a few criteria such as size, geographical coverage, and ownership to make this distinction.

1. Local Area Network

- A **local area network** (**LAN**) is usually privately owned and connects some hosts in a single office, building, or campus.
- Depending on the needs of an organization, a LAN can be as simple as two PCs and a printer in someone's home office, or it can extend throughout a company and include audio and video devices.
- Each host in a LAN has an identifier, an address that uniquely defines the host in the LAN. A packet sent by a host to another host carries both the source host's and the destination host's addresses.

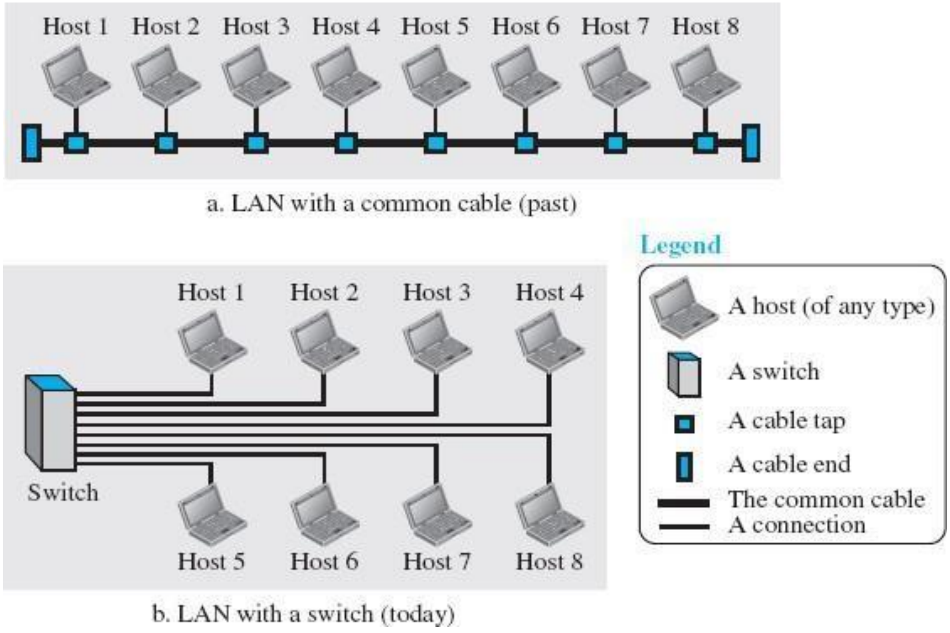


Figure 1.8 LAN

- In the past, all hosts in a network were connected through a common cable, which meant that a packet sent from one host to another was received by all hosts. The intended recipient kept the packet; the others dropped the packet.
- Today, most LANs use a smart connecting switch, which is able to recognize the destination address of the packet and guide the packet to its destination without sending it to all other hosts.

2. Wide Area Network

- A **wide area network (WAN)** is also an interconnection of devices capable of communication. However, there are some differences between a LAN and a WAN.
- A LAN is normally limited in size, spanning an office, a building, or a campus; a WAN has a wider geographical span, spanning a town, a state, a country, or even the world.
- A LAN interconnects hosts; a WAN interconnects connecting devices such as switches, routers, or modems.
- A LAN is normally privately owned by the organization that uses it; a WAN is normally created and run by communication companies and leased by an organization that uses it.
- We see two distinct examples of WANs today: point-to-point WANs and switched WANs.

Point-to-Point WAN

- A point-to-point WAN is a network that connects two communicating devices through a transmission media (cable or air). Figure 1.9 shows an example of a point-to-point WAN.

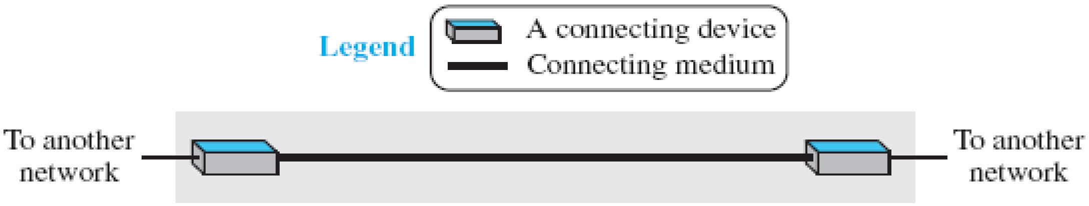


Figure 1.9 Point to Point to WAN

Switched WAN

- A switched WAN is a network with more than two ends. We can say that a switched WAN is a combination of several point-to-point WANs that are connected by switches. Figure 1.10 shows an example of a switched WAN.

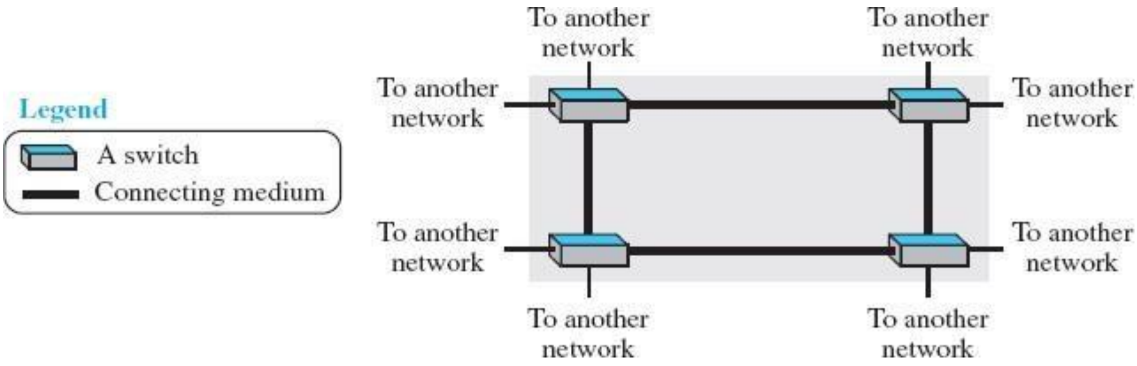


Figure 1.10 Switched WAN

Internetwork

- Today, it is very rare to see a LAN or a WAN in isolation; they are connected to one another. When two or more networks are connected, they make an **internetwork,** or **internet.**
- As an example, assume that an organization has two offices, one on the east coast and the other on the west coast. Each office has a LAN that allows all employees in the office to communicate with each other.
- To make the communication between employees at different offices possible, the management leases a point-to-point dedicated WAN from a service provider, such as a telephone company, and connects the two LANs.
- Now the company has an internetwork, or a private internet (with lowercase i). Communication between offices is now possible. Figure 1.11 shows this internet.

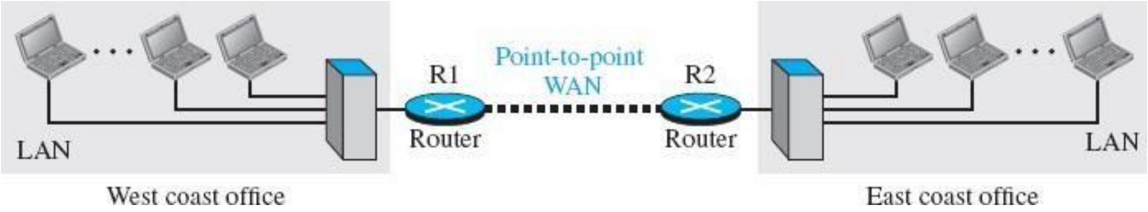


Figure 1.11 Internet

- When a host in the west coast office sends a message to another host in the same office, the router blocks the message, but the switch directs the message to the destination. On the other hand, when a host on the west coast sends a message to a host on the east coast, router R1 routes the packet to router R2, and the packet reaches the destination.

PROTOCOL LAYERING

- In data communication and networking, a **protocol** defines the rules that both the sender and receiver and all intermediate devices need to follow to be able to communicate effectively.
- When communication is simple, we may need only one simple protocol; when the communication is complex, we may need to divide the task between different layers, in which case we need a protocol at each layer, or **protocol layering.**

Scenarios

- Let us develop two simple scenarios to better understand the need for protocol layering.

First Scenario

- In the first scenario, communication is so simple that it can occur in only one layer. Assume Maria and Ann are neighbors with a lot of common ideas. Communication between Maria and Ann takes place in one layer, face to face, in the same language, as shown in Figure 1.11.

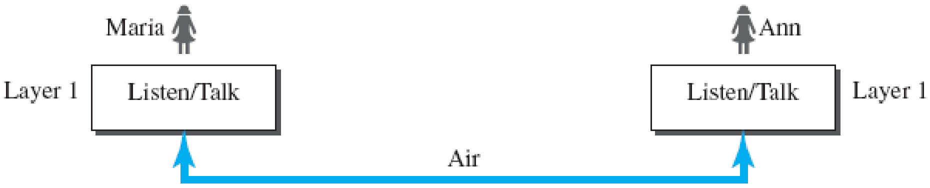


Figure 1.12 Single layer protocol

Even in this simple scenario, we can see that a set of rules needs to be followed.

- First, Maria and Ann know that they should greet each other when they meet.
- Second, they know that they should confine their vocabulary to the level of their friendship.
- Third, each party knows that she should refrain from speaking when the other party is speaking.
- Fourth, each party knows that the conversation should be a dialog, not a monolog: both should have the opportunity to talk about the issue.
- Fifth, they should exchange some nice words when they leave.

Second Scenario

- In the second scenario, we assume that Ann is offered a higher-level position in her company, but needs to move to another branch located in a city very far from Maria.
- They decide to continue their conversation using regular mail through the post office. However, they do not want their ideas to be revealed by other people if the letters are intercepted.
- They agree on an encryption/decryption technique. The sender of the letter encrypts it to make it unreadable by an intruder; the receiver of the letter decrypts it to get the original letter.
- Now we can say that the communication between Maria and Ann takes place in three layers, as shown in Figure 1.13. We assume that Ann and Maria each have three machines(or robots) that can perform the task at each layer.

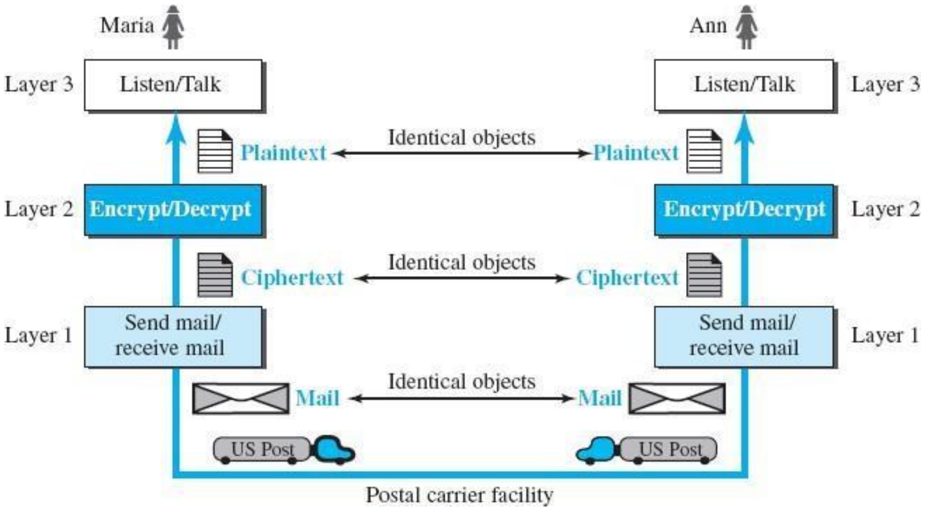


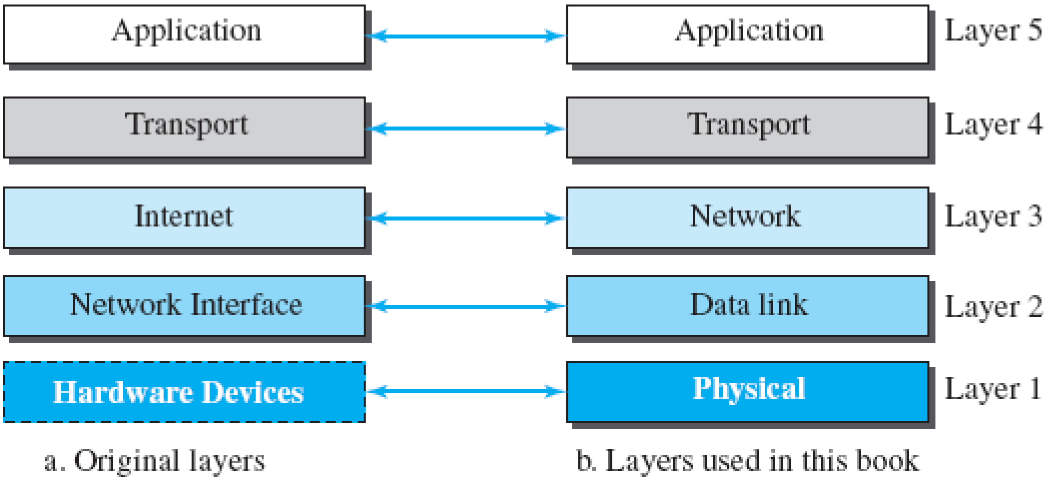
Figure 1.13 Three layer Protocol

- Let us assume that Maria sends the first letter to Ann.
- Maria talks to the machine at the third layer as though the machine is Ann and is listening to her. The third layer machine listens to what Maria says and creates the plaintext (a letter in English), which is passed to the second layer machine.
- The second layer machine takes the plaintext, encrypts it, and creates the ciphertext, which is passed to the first layer machine.
- The first layer machine, presumably a robot, takes the ciphertext, puts it in an envelope, adds the sender and receiver addresses, and mails it.
- At Ann's side, the first layer machine picks up the letter from Ann's mail box, recognizing the letter from Maria by the sender address. The machine takes out the ciphertext from the envelope and delivers it to the second layer machine.
- The second layer machine decrypts the message, creates the plaintext, and passes the plaintext to the third-layer machine.
- The third layer machine takes the plaintext and reads it as though Maria is speaking.
- Advantages of Protocol Layering
- Protocol layering enables us to divide a complex task into several smaller and simpler tasks.
- One of the advantages of protocol layering is that it allows us to separate the services from the implementation. A layer needs to be able to receive a set of services from the lower layer and to give the services to the upper layer; we don't care about how the layer is implemented.
- If we did not use protocol layering, we would have to make each intermediate system as complex as the end systems, which makes the whole system more expensive.

TCP/IP PROTOCOL SUITE

Layers in the TCP/IP Protocol Suite

- TCP/IP is a protocol suite (a set of protocols organized in different layers) used in the Internet today. It is a hierarchical protocol made up of interactive modules, each of which provides a specific functionality. The term hierarchical means that each upper levelprotocol is supported by the services provided by one or more lower level protocols.
- The original TCP/IP protocol suite was defined as four software layers built upon the hardware. Today, however, TCP/IP is thought of as a five-layer model. Figure 1.14 shows both configurations.



a. Original layers

b. Layers used in this book

Figure 1.14 Layers in TCP/IP protocol Suite

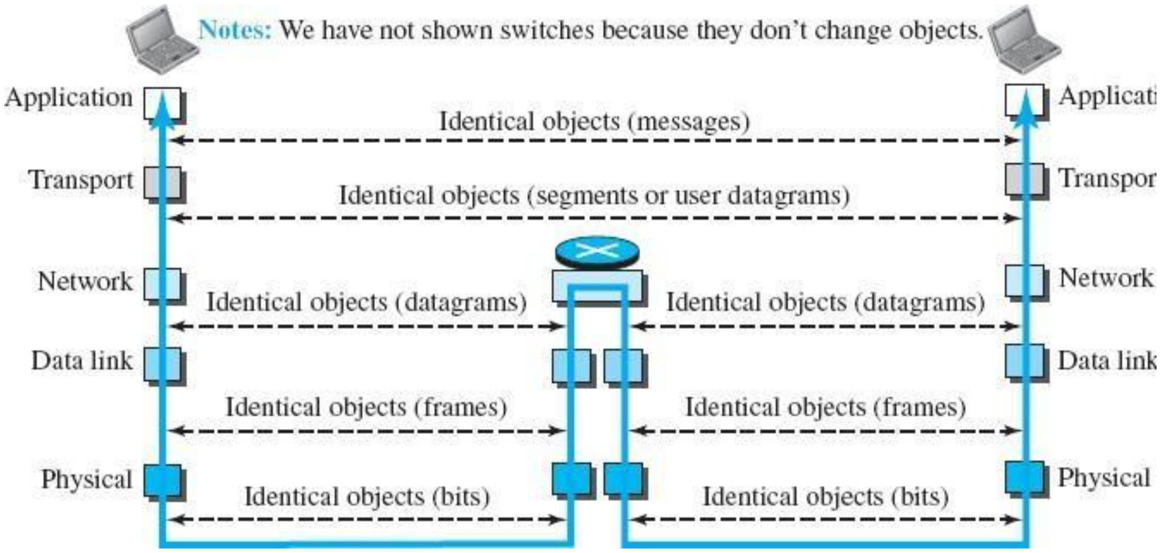

Notes: We have not shown switches because they don't change objects.
Application
Transport
Network
Data link
Physical
Applicati
Transpor
Network
Data link
Physical
Identical objects (messages)
Identical objects (segments or user datagrams)
Identical objects (datagrams)
Identical objects (datagrams)
Identical objects (frames)
Identical objects (frames)
Identical objects (bits)
Identical objects (bits)

Figure 1.15 Identical objects in the TCP/IP protocol suite

1. Physical Layer

- We can say that the physical layer is responsible for carrying individual bits in a frame across the link.

- We need to know that the transmission medium does not carry bits; it carries electrical or optical signals. So the bits received in a frame from the data-link layer are transformed and sent through the transmission media, but we can think that the logical unit between two physical layers in two devices is a bit.

2. Data-link Layer

- We have seen that an internet is made up of several links (LANs and WANs) connected by routers.
- The routers are responsible for choosing the best links. However, when the next link to travel is determined by the router, the data-link layer is responsible for taking the datagram and moving it across the link.
- The link can be a wired LAN with a link-layer switch, a wireless LAN, a wired WAN, or a wireless WAN.
- In each case, the data-link layer is responsible for moving the packet through the link.
- The data-link layer takes a datagram and encapsulates it in a packet called a **frame**.

3. Network Layer

- The network layer is responsible for creating a connection between the source computer and the destination computer.
- he communication at the network layer is host-to-host. However, since there can be several routers from the source to the destination, the routers in the path are responsiblefor choosing the best route for each packet.
- We can say that the network layer is responsible for host-to-host communication and routing the packet through possible routes.
- The network layer in the Internet includes the main protocol, Internet Protocol (IP) that defines the format of the packet, called a datagram at the network layer. IP also defines the format and the structure of addresses used in this layer. IP is also responsible for routing a packet from its source to its destination, which is achieved by each router forwarding the datagram to the next router in its path. IP is a connectionless protocol that provides no flow control, no error control, and no congestion control services. This meansthat if any of these services is required for an application, the application should rely only on the transport-layer protocol.
- The network layer also has some auxiliary protocols such as Dynamic Host Configuration Protocol (DHCP), Address Resolution Protocol (ARP) these help IP in its delivery and routing tasks.

4. Transport Layer

- The logical connection at the transport layer is end-to-end. The transport layer at the source host gets the message from the application layer, encapsulates it in a transportlayer packet (called a segment or a user datagram) and sends it, through the logical (imaginary) connection, to the transport layer at the destination host.

- In other words, the transport layer is responsible for giving services to the application layer: to get a message from an application program running on the source host anddeliver it to the corresponding application program on the destination host.

- The transport layer should be independent of the application layer.

- In addition, we will see that we have more than one protocol in the transport layer, which means that each application program can use the protocol that best matches its requirement.
- The main protocol, Transmission Control Protocol (TCP), is a connection-oriented protocol that first establishes a logical connection between transport layers at two hosts before transferring data. It creates a logical pipe between two TCPs for transferring a stream of bytes. TCP provides flow control (matching the sending data rate of the source host with the receiving data rate of the destination host to prevent overwhelming the destination), error control (to guarantee that the segments arrive at the destination without error and resending the corrupted ones), and congestion control to reduce the loss of segments due to congestion in the network.
- The other common protocol, User Datagram Protocol (UDP), is a connectionless protocol that transmits user datagrams without first creating a logical connection. In UDP, each user datagram is an independent entity without being related to the previous or the next one. UDP is a simple protocol that does not provide flow, error, or congestion control.

5. Application Layer

- The two application layers exchange messages between each other as though there were a bridge between the two layers. However, we should know that the communication is done through all the layers. Communication at the application layer is between two processes (two programs running at this layer). To communicate, a process sends a request to the other process and receives a response. Process-to-process communication is the duty ofthe application layer.
- The Hypertext Transfer Protocol (HTTP) is a vehicle for accessing the World Wide Web (WWW). The Simple Mail Transfer Protocol (SMTP) is the main protocol used in electronic mail (e-mail) service. The File Transfer Protocol (FTP) is used for transferring files from one host to another.

Encapsulation and Decapsulation

One of the important concepts in protocol layering in the Internet is encapsulation / decapsulation. Figure 1.16 shows this concept for the small internet in Figure 1.15.

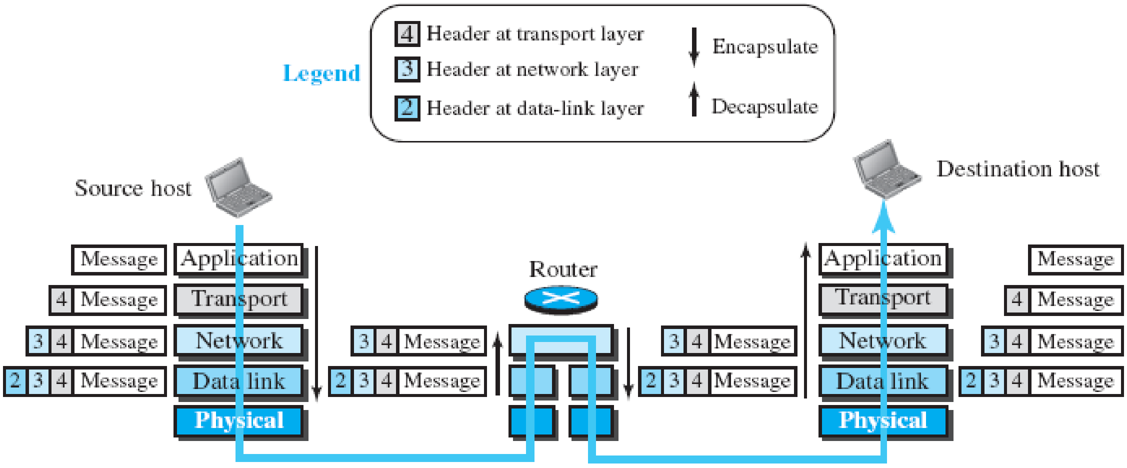


Figure 1.16 Encapsulation/Decapsulation

Encapsulation at the Source Host

At the source, we have only encapsulation.

1. At the application layer, the data to be exchanged is referred to as a **message**. A message normally does not contain any header or trailer, but if it does, we refer to the whole as the message. The message is passed to the transport layer.
2. The transport layer takes the message as the payload, the load that the transport layer should take care of. It adds the transport layer header to the payload, which contains the identifiers of the source and destination application programs that want to communicate plus some more information that is needed for the end-to-end delivery of the message, such as information needed for flow, error control, or congestion control
3. The network layer takes the transport-layer packet as data or payload and adds its own header to the payload. The header contains the addresses of the source and destination hosts and some more information used for error checking of the header, fragmentation information, and so on. The result is the network-layer packet, called a datagram. The network layer then passes the packet to the data-link layer.
4. The data-link layer takes the network-layer packet as data or payload and adds its own header, which contains the link-layer addresses of the host or the next hop (the router). The result is the link-layer packet, which is called a **frame**. The frame is passed to the physical layer for transmission.

Decapsulation and Encapsulation at the Router

At the router, we have both decapsulation and encapsulation because the router is connected to two or more links.

1. After the set of bits are delivered to the data-link layer, this layer decapsulates thedatagram from the frame and passes it to the network layer.
2. The network layer only inspects the source and destination addresses in the datagram header and consults its forwarding table to find the next hop to which the datagram is tobe delivered. The contents of the datagram should not be changed by the network layer in the router unless there is a need to fragment the datagram if it is too big to be passed through the next link. The datagram is then passed to the data-link layer of the next link.
3. The data-link layer of the next link encapsulates the datagram in a frame and passes it to the physical layer for transmission.

Decapsulation at the Destination Host

1. At the destination host, each layer only decapsulates the packet received, removes the payload, and delivers the payload to the next-higher layer protocol until the message reaches the application layer. It is necessary to say that decapsulation in the host involves error checking.

Addressing

- It is worth mentioning another concept related to protocol layering in the Internet, **addressing**. As we discussed before, we have logical communication between pairs of layers in this model.
- Any communication that involves two parties needs two addresses: source address and destination address. Although it looks as if we need five pairs of addresses, one pair per layer, we normally have only four because the physical layer does not need addresses; the unit of data exchange at the physical layer is a bit, which definitely cannot have an address. Figure 1.17 shows the addressing at each layer.

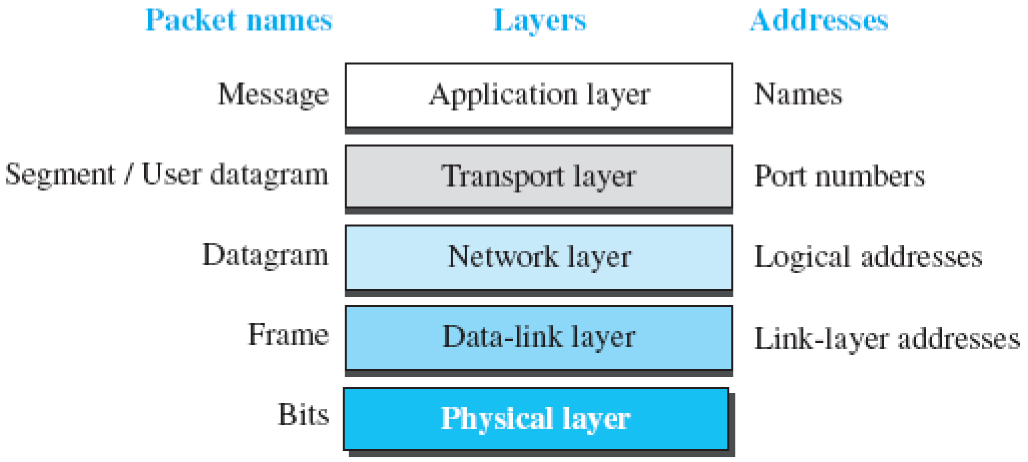


Figure 1.17 Addressing in TCP/IP protocol suite

THE OSI MODEL

- Although, when speaking of the Internet, everyone talks about the TCP/IP protocol suite, this suite is not the only suite of protocols defined.
- Established in 1947, the **International Organization for Standardization (ISO)** is a multinational body dedicated to worldwide agreement on international standards. Almost three-fourths of the countries in the world are represented in the ISO.
- An ISO standard that covers all aspects of network communications is the **Open Systems Interconnection (OSI) model.** It was first introduced in the late 1970s. An open system is a set of protocols that allows any two different systems to communicate regardless of their underlying architecture.
- The purpose of the OSI model is to show how to facilitate communication between different systems without requiring changes to the logic of the underlying hardware and software.
- The OSI model is not a protocol; it is a model for understanding and designing a network architecture that is flexible, robust, and interoperable.
- The OSI model was intended to be the basis for the creation of the protocols in the OSI stack.

- The OSI model is a layered framework for the design of network systems that allows communication between all types of computer systems.
- It consists of seven separate but related layers, each of which defines a part of the process of moving information across a network (see Figure 1.18).

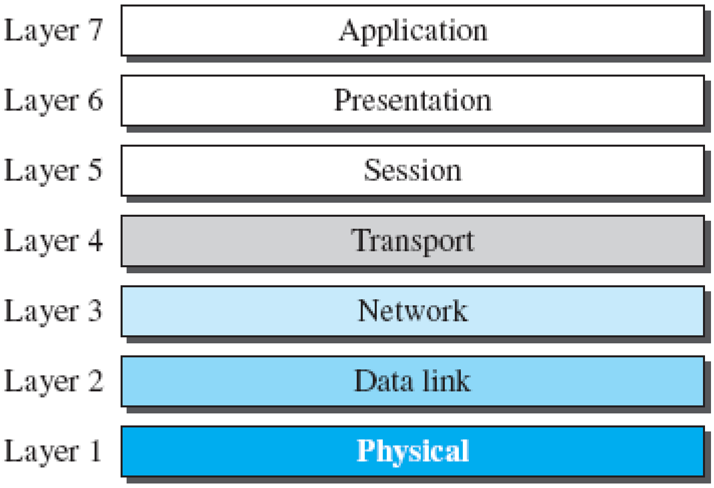


Figure 1.18 OSI Model

of OSI Model’s Success

- The OSI model appeared after the TCP/IP protocol suite. Most experts were at first excited and thought that the TCP/IP protocol would be fully replaced by the OSI model.
- This did not happen for several reasons, but we describe only three, which are agreed upon by all experts in the field.
- **First**, OSI was completed when TCP/IP was fully in place and a lot of time and money had been spent on the suite; changing it would cost a lot.
- **Second**, some layers in the OSI model were never fully defined. For example, although the services provided by the presentation and the session layers were listed in the document, actual protocols for these two layers were not fully defined, nor were theyfully described, and the corresponding software was not fully developed.
- **Third**, when OSI was implemented by an organization in a different application, it did not show a high enough level of performance to entice the Internet authority to switch from the TCP/IP protocol suite to the OSI model.

CHAPTER 2
Physical Layer

Transmission media

Transmission media are actually located below the physical layer and are directly controlled by the physical layer. We could say that transmission media belong to layer zero. Figure 2.1 shows the position of transmission media in relation to the physical layer.

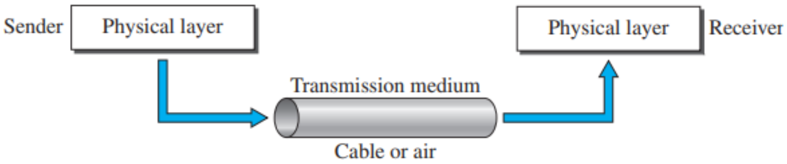


Figure 2.1 Transmission medium and physical layer

A **transmission medium** can be broadly defined as anything that can carry information from a source to a destination. For example, the transmission medium for two people having a dinner conversation is the air. The air can also be used to convey the message in a smoke signal or semaphore. For a written message, the transmission medium might be a mail carrier, a truck, or an airplane.

In data communications the definition of the information and the transmission medium is more specific. The transmission medium is usually free space, metallic cable, or fiber-optic cable. The information is usually a signal that is the result of a conversion of data from another form.

In telecommunications, transmission media can be divided into two broad categories: guided and unguided. Guided media include twisted-pair cable, coaxial cable, and fiber-optic cable. Unguided medium is free space. Figure 2.2 shows this taxonomy.

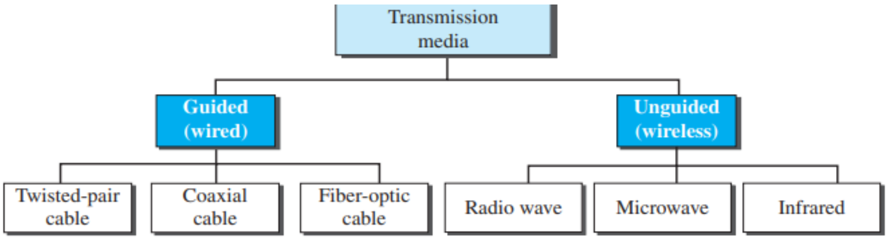


Figure 2.2 Classes of transmission media

GUIDED MEDIA

Guided media, which are those that provide a conduit from one device to another, include twisted-pair cable, coaxial cable, and fiber-optic cable. A signal traveling along any of these media is directed and contained by the physical limits of the medium. Twisted-pair and coaxial cable use metallic (copper) conductors that accept and transport signals in the form of electric current. Optical fiber is a cable that accepts and transports signals in the form of light.

Twisted-Pair Cable

A twisted pair consists of two conductors (normally copper), each with its own plastic insulation, twisted together, as shown in Figure 2.3.

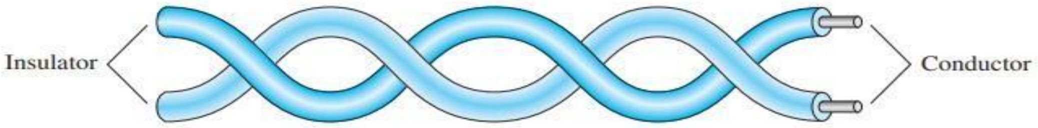


Figure 2.3. Twisted-pair cable

A twisted pair consists of two insulated copper wires, typically about 1 mm thick. The wires are twisted together in a helical form, just like a DNA molecule. Twisting is done because two parallel wires constitute a fine antenna.

- When the wires are twisted, the waves from different twists cancel out, so the wire radiates less effectively. A signal is usually carried as the difference in voltage between the two wires in the pair. This provides better immunity to external noise because the noise tends to affect both wires the same, leaving the differential unchanged.
- The most common application of the twisted pair is the telephone system. Nearly all telephones are connected to the telephone company (telco) office by a twisted pair. Both telephone calls and ADSL Internet access run over these lines.
- Twisted pairs can be used for transmitting either analog or digital information. The bandwidth depends on the thickness of the wire and the distance traveled, but several megabits/sec can be achieved for a few kilometer's.

The garden variety deployed in many office buildings is called Category 5 cabling, or ''Cat 5.

A category 5 twisted pair consists of two insulated wires gently twisted together. Four such pairs are typically grouped in a plastic sheath to protect the wires and keep them together. To reach higher speeds, 1-Gbps Ethernet uses all four pairs in both directions simultaneously.

Full duplex Links that can be used in both directions at the same time, like a two-lane road, are called full-duplex links.

Half duplex Links that can be used in either direction, but only one way at a time, like a single-track railroad line, called half duplex links.

Simplex Links that allow traffic in only one direction, like a one-way street are called simplex links.

Category 3

Cable uses a same connector, but has more twists per meter. More twists result in less crosstalk and a better quality signal over longer distances, making the cables more suitable for high-speed computer communication, especially 100-Mbps and 1-Gbps Ethernet LANs.

Coaxial Cable

- Coaxial cable has better shielding and greater bandwidth than unshielded twisted pairs, so it can span longer distances at higher speeds.
- Two kinds of coaxial cable are widely used. One kind, 50-ohm cable, is commonly used when it is intended for digital transmission from the start. The other kind, 75-ohm cable, is commonly used for analog transmission and cable television.
- A coaxial cable consists of a stiff copper wire as the core, surrounded by an insulating material. The insulator is encased by a cylindrical conductor, often as a closely woven braided mesh. The outer conductor is covered in a protective plastic sheath.

. The outer metallic wrapping serves both as a shield against noise and as the second conductor, which completes the circuit. This outer conductor is also enclosed in an insulating sheath, and the whole cable is protected by a plastic cover (see Figure 2.4).

- The bandwidth possible depends on the cable quality and length. Coaxial cables used to be widely used within the telephone system for long-distance lines but have now largely been replaced by fiber optics on long haul routes. Coax is still widely used for cable television and metropolitan areanetworks.

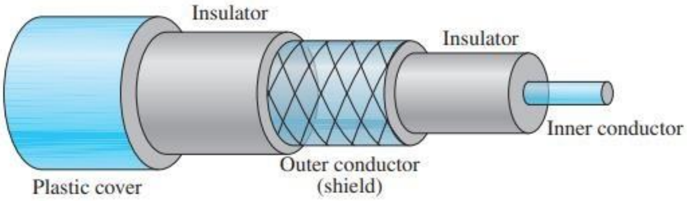


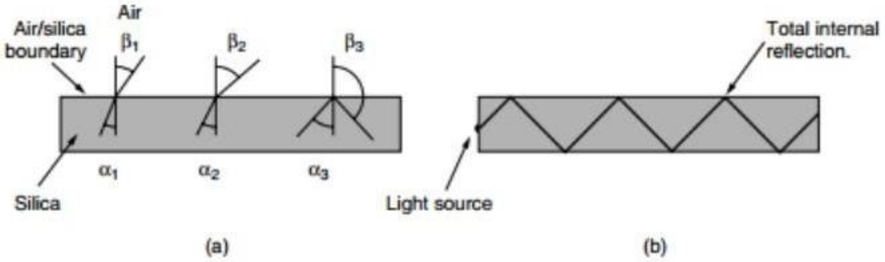
Figure 2.4 Coaxial cable

Fiber Optics

• Fiber optics are used for long-haul transmission in network backbones, highspeed LANs (although so far, copper has always managed catch up eventually), and high-speed Internet access such as FttH (Fiber to the Home).

• An optical transmission system has three key components: the light source, the transmission medium, and the detector. The transmission medium is an ultra-thin fiber of glass. The detector generates an electrical pulse when light falls on it.

• By attaching a light source to one end of an optical fiber and a detector to the other, we have a unidirectional data transmission system that accepts an electrical signal, converts and transmits it by light pulses, and then reconverts the output to an electrical signal at the receiving end.



Fig(a): When a light ray passes from one medium to another—for example, from fused silica to air—the ray is refracted (bent) at the silica/air boundary. A light ray incident on the boundary at an angle α1 emerging at an angle β1. The amount of refraction depends on the properties of the two media (in particular, their indices of refraction)

Fig (b): The light is refracted back into the silica; none of it escapes into the air. Thus, a light ray incident at or above the critical angle is trapped inside the fiber. The fig shows the only one trappedray, but since any light ray incident on the boundary above the critical angle will be reflected internally, many different rays will be bouncing around at different angles. Each ray is said to have a different mode, so a fiber having this property is called a **multimode fiber.**

Single mode fiber

if the fiber's diameter is reduced to a few wavelengths of light the fiber acts like a wave guide and the light can propagate only in a straight line, without bouncing, yielding a single-mode fiber. Single-mode fibers are more expensive but are widely used for longer distances.

Transmission of Light Through Fiber

Optical fibers are made of glass, which, in turn, is made from sand, an inexpensive raw material available in unlimited amounts. The attenuation of light through glass depends on the wavelength of the light. It is defined as the ratio of input to output signal power.

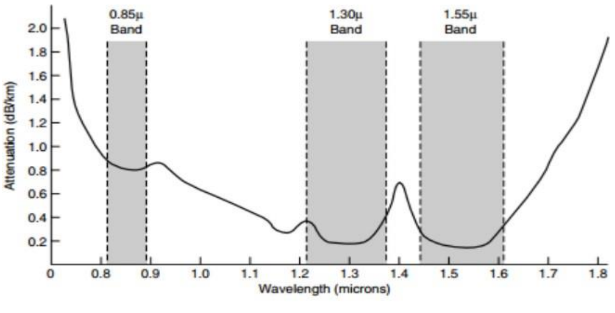


Figure 2-7. Attenuation of light through fiber in the infrared region.

The figure shows the near- infrared part of the spectrum. Visible light has slightly shorter wavelengths, from 0.4 to 0.7 microns. The true metric purist would refer to these wavelengths as 400 nm to 700 nm. Three wavelength bands are most commonly used at present for optical communication. They are centered at 0.85, 1.30, and 1.55 microns, respectively. All three bands are 25,000 to 30,000 GHz wide. The 0.85- micron band was used first. It has higher attenuation and so is used for shorter distances. The last two bands have good attenuation properties (less than 5% loss per kilometer). The 1.55- micron band is now widely used with erbium-doped amplifiers that work directly in the optical domain.

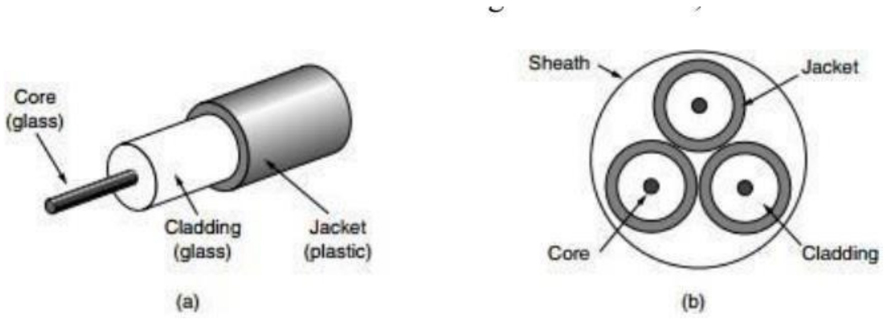
Chromatic Dispersion

Light pulses sent down a fiber spread out in length as they propagate.

Soliton: A soliton is a pulse that can collide with another similar pulse and still retain its shape after the collision, again in the presence of both dispersion and non-linearities.

Fiber Cables

• Fiber optic cables are similar to coax, except without the braid. At the center is the glass core through which the light propagates. In multimode fibers, the core is typically 50 microns in diameter, about the thickness of a human hair. In single-mode fibers, the core is 8 to 10 microns.

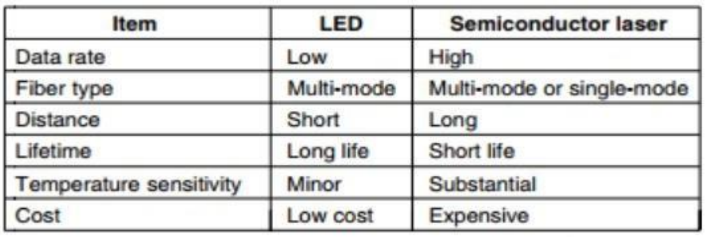


• The core is surrounded by a glass cladding with a lower index of refraction than the core, to keep all the light in the core. Next comes a thin plastic jacket to protect the cladding. Fibers are typically grouped in bundles, protected by an outer sheath.

• Fibers can be connected in three different ways. First, they can terminate in connectors and be plugged into fiber sockets. Connectors lose about 10 to 20% of the light, but they make it easy to reconfigure systems.

• Second, they can be spliced mechanically. Mechanical splices just lay the two carefully cut ends next to each other in a special sleeve and clamp them in place. Alignment can be improved by passing light through the junction and then making small adjustments to maximize the signal.

• Third, two pieces of fiber can be fused (melted) to form a solid connection. A fusion splice is almost as good as a single drawn fiber, but even here, a small amount of attenuation occurs. For all three kinds of splices, reflections can occur at the point of the splice, and the reflected energy can interfere with the signal.



Item	LED	Semiconductor laser
Data rate	Low	High
Fiber type	Multi-mode	Multi-mode or single-mode
Distance	Short	Long
Lifetime	Long life	Short life
Temperature sensitivity	Minor	Substantial
Cost	Low cost	Expensive

Figure 2-9. A comparison of semiconductor diodes and LEDs as light sources.

Comparison of Fiber Optics and Copper Wire

Advantages of Fiber Optics:

- It can handle higher bandwidths than copper.
- Fiber optic cables have low attenuation.
- Fiber optics cables are not affected by electromagnetic interferences and power fluctuations.
- Fiber optic cables are much more secured.
- Fiber cables are thin and light weight.
- The life cycle of fiber cables is 30 to 50 years, which is much higher than copper cables.

Disadvantages of Fiber Optics:

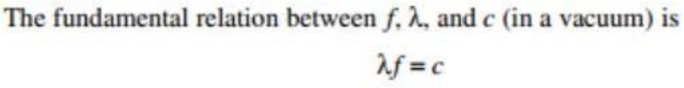
- It is a newer technology with not much expertise.
- Copper cables and connectors are much cheaper than fiber optic cables and connectors.
- Propagation ofsignals in fiber optic cables is unidirectional.

WIRELESS TRANSMISSION

Wireless communications have many important applications besides providing connectivity to users from any place.

The Electromagnetic Spectrum

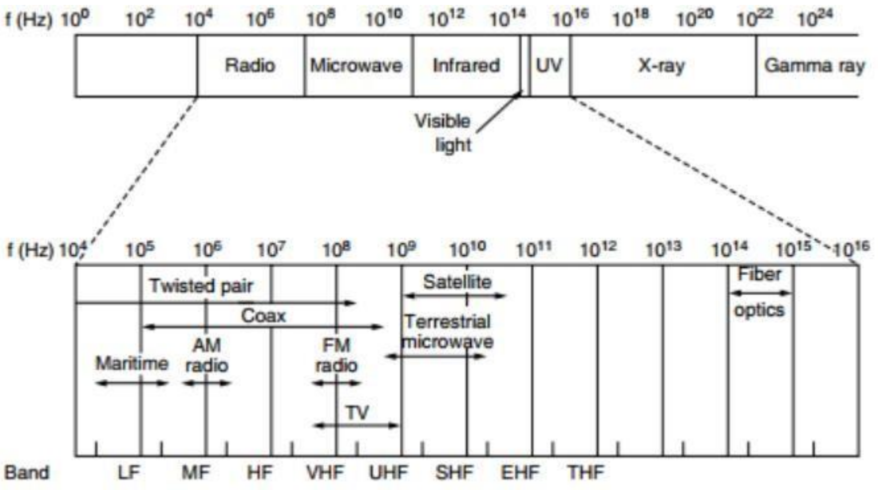
- When electrons move, they create electromagnetic waves that can propagate through space.
- The number of oscillations per second of a wave is called its frequency, f, and is measured in Hz. The distance between two consecutive maxima (or minima) is called the wavelength, represented by lambda.
- In a vacuum, all electromagnetic waves travel at the same speed, no matter what their frequency. This speed, usually called the **speed of light**, c, is approximately 3 × 108 m/sec, or about 1 foot (30 cm) per nanosecond.



The fundamental relation between f, λ, and c (in a vacuum) is

$$\lambda f = c$$

- The radio, microwave, infrared, and visible light portions of the spectrum can all be used for transmitting information by modulating the amplitude, frequency, or phase of the waves. Ultraviolet light, X-rays, and gamma rays would be even better, due to their higher frequencies, but they are hard to produce and modulate, do not propagate well. The terms LF, MF, and HF refer to Low, Medium,and High Frequency, respectively.

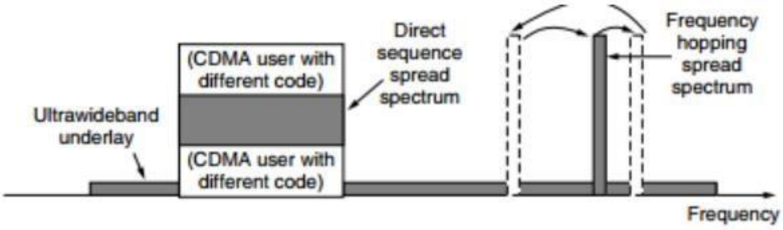


f (Hz) 10^0 10^2 10^4 10^6 10^8 10^{10} 10^{12} 10^{14} 10^{16} 10^{18} 10^{20} 10^{22} 10^{24}
Radio
Microwave
Infrared
UV
X-ray
Gamma ray
Visible light
f (Hz) 10^4 10^5 10^6 10^7 10^8 10^9 10^{10} 10^{11} 10^{12} 10^{13} 10^{14} 10^{15} 10^{16}
Twisted pair
Coax
Satellite
Terrestrial microwave
Fiber optics
Maritime
AM radio
FM radio
TV
Band
LF
MF
HF
VHF
UHF
SHF
EHF
THF

• Most transmissions use a relatively narrow frequency band (i.e., Δf / f << 1). In **frequency hopping spread** spectrum, the transmitter hops from frequency-to-frequency hundreds of times per second. It is pop ular for military communication because it makes transmissions hard to detect and next to impossible to jam. This technique is used commercially, for example, in Bluetooth and older versionsof 802.11.

Direct Sequence spread spectrum

A second form of spread spectrum, direct sequence spread spectrum, uses a code sequence to spread the data signal over a wider frequency band. It is widely used commercially as a spectrally efficient way to let multiple signals share the same frequency band. These signals can be given different codes, a method called **CDMA (Code Division Multiple Access).**



Ultrawideband underlay
(CDMA user with different code)
(CDMA user with different code)
Direct sequence spread spectrum
Frequency hopping spread spectrum
Frequency

It forms the basis of 3G mobile phone networks and is also used in GPS (Global Positioning System).

UWB (Ultra-Wide Band)

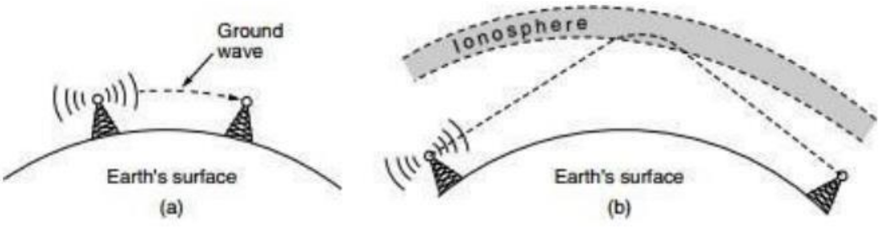
UWB sends a series of rapid pulses, varying their positions to communicate information. The rapid transitions lead to a signal that is spread thinly over a very wide frequency band. UWB is defined as signals that have a bandwidth of at least 500 MHz or at least 20% of the center frequency of their frequency band. It can tolerate a substantial amount of relatively strong interference from other narrowband signals, because it is spread across wide band of frequencies.

Radio Transmission

• Radio frequency (RF) waves are easy to generate, can travel long distances, and can penetrate buildings easily, so they are widely used for communication, both indoors and outdoors. Radio waves also are omni directional, meaning that they travel in all directions from the source, so thetransmitter and receiver do not have to be carefully aligned physically.

• The properties of radio waves are frequency dependent. At low frequencies, radio waves pass through obstacles well, but the power falls off sharply with distance from the source—at least as fast as 1/r 2 in air—asthe signal energy is spread more thinly over a larger surface. This attenuation is called path loss.

• At high frequencies, radio waves tend to travel in straight lines and bounce off obstacles. Path loss still reduces power, though the received signal can depend strongly on reflections as well. High-frequency radio waves are also absorbed by rain and other obstacles to a larger extent than are low-frequency ones. At all frequencies, radio waves are subject to interference from motors and other electrical equipment.



From fig(a): In the VLF, LF, and MF bands, radio waves follow the ground. These waves can be detected for perhaps 1000 km at the lower frequencies, less at the higher ones. AM radio broadcasting uses the MF band.

Fig(b) In the HF and VHF bands, the ground waves tend to be absorbed by the earth. However, the waves that reach the ionosphere, a layer of charged particles circling the earth at a height of 100 to 500 km, are refracted by it and sent back to earth.

Microwave Transmission

• Microwaves travel in a straight line. Thus, repeaters are needed periodically. The distance between repeaters is square root of the tower height. For 100- meter-high towers, repeaters can be 80 km apart. Microwaves do not pass through buildings well.

• Some waves may be refracted off low- lying atmospheric layers and may take slightlylonger to arrive than the direct waves. The delayed waves may arrive out of phase with the direct wave and thus cancel the signal. This effect is called multipath fading.

• Bands up to 10 GHz are now in routine use. These waves are only a few centimetres long and are absorbed by rain. Microwave communication is so widely used for long-distance telephone communication, mobile phones, television distribution. It does not require to lay down cables. By buying a small plot of ground every 50 km and putting a microwave tower on it, one can bypass the telephone system entirely. Microwave is also relatively inexpensive.

Politics of ElectroMagnetic Spectrum

National governments allocate spectrum for AM and FM radio, television, and mobile phones, as well as for telephone companies, police, maritime, navigation, military, government, and many other competing users. 3 algorithms are widely used.

Beauty Contest

Oldest algorithm requires each carrier to explain why its proposal serves the public interest best. Government officials then decide which of the nice stories they enjoy most.

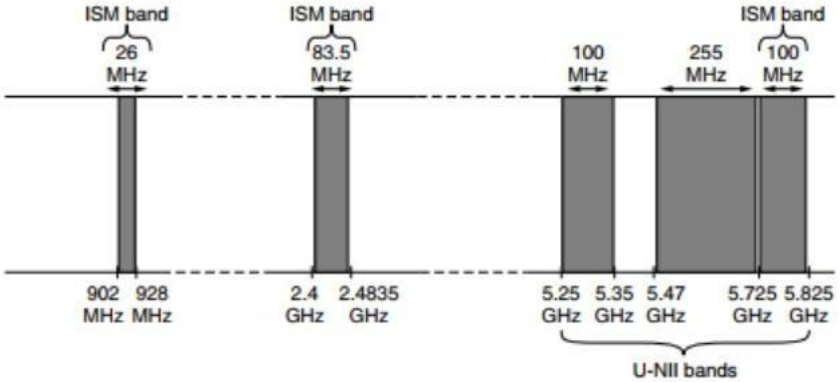
Lottery

Second algorithm which holds lottery among the interested companies. The problem with that idea is that companies with no interest in using the spectrum can enter the lottery.

Auction

Random companies has been severely criticized by many, which led to algorithm 3: auction off the bandwidth to the highest bidder.

A completely different approach to allocating frequencies is to not allocate them at all. Accordingly, most governments have set aside some frequency bands, called the ISM (Industrial, Scientific, Medical) bands for unlicensed usage. To minimize interference between these uncoordinated devices, the FCC mandates that all devices in the ISM bands limit their transmit power.



The 900-MHz band was used for early versions of 802.11, but it is crowded. The 2.4-GHz band is available in most countries and widely used for 802.11b/g and Bluetooth, though it is subject to interference from microwave ovens and radar installations. The 5-GHz part of the spectrum includes U-NII (**Unlicensed National Information Infrastructure**) bands.

One exciting development in the U.S. is the FCC decision in 2009 to allow unlicensed use of white spaces around 700 MHz. White spaces are frequency bands that have been allocated but are not being used locally. The only difficulty to use the **white spaces**, unlicensed devices must be able to detect any nearby licensed transmitters, including wireless microphones.

InfraRed Transmission

- Unguided infrared waves are widely used for short-range communication. The remotecontrols used for televisions, VCRs, and stereos all use infrared communication.
- They are relatively directional, cheap, and easy to build but have a major drawback: they do not pass through solid objects. It means that an infrared system in one room of a building will not interfere with a similar system in adjacent rooms or buildings.
- Infrared communication has a limited use on the desktop, for example, to connect notebook computers and printers with the **IrDA (Infrared Data Association)** standard, but it is not a major player in the communication game.

]