DATA COMMUNICATION

DATA COMMUNICATION

(Common to CSE & ISE)

Subject Code: 15CS46	**I.A. Marks : 20**
Hours/Week : 04	**Exam Hours: 03**
Total Hours : 50	**Exam Marks: 80**

MODULE – 1 **10 Hours**

Introduction: Data Communications, Networks, Network Types, Internet History, Standards and Administration
Networks Models: Protocol Layering, TCP/IP Protocol suite, The OSI model
Introduction to Physical Layer-1: Data and Signals, Digital Signals, Transmission Impairment, Data Rate limits, Performance
Digital Transmission: Digital to digital conversion (Only Line coding: Polar, Bipolar and Manchester coding)

MODULE – 2 **10 Hours**

Physical Layer-2: Analog to digital conversion (only PCM), Transmission Modes
Analog Transmission: Digital to analog conversion
Bandwidth Utilization: Multiplexing and Spread Spectrum
Switching: Introduction, Circuit Switched Networks and Packet switching

MODULE – 3 **10 Hours**

Error Detection and Correction: Introduction, Block coding, Cyclic codes, Checksum, Forward error correction.
Data link control: DLC services, Data link layer protocols, HDLC, and Point to Point protocol (Framing, Transition phases only).

MODULE – 4 **10 Hours**

Media Access control: Random Access, Controlled Access and Channelization
Wired LANs Ethernet: Ethernet Protocol, Standard Ethernet, Fast Ethernet, Gigabit Ethernet and 10 Gigabit Ethernet
Wireless LANs: Introduction, IEEE 802.11 Project and Bluetooth

MODULE – 5 **10 Hours**

Other wireless Networks: WIMAX, Cellular Telephony, Satellite networks
Network layer Protocols: Internet Protocol, ICMPv4, Mobile IP
Next generation IP: IPv6 addressing, The IPv6 Protocol, The ICMPv6 Protocol and Transition from IPv4 to IPv6.

Text Book:

Behrouz A. Forouzan, Data Communications and Networking 5E, 5th Edition, Tata McGraw-Hill, 2013. (Chapters 1.1 to 1.5, 2.1 to 2.3, 3.1, 3.3 to 3.6, 4.1 to 4.3, 5.1, 6.1, 6.2, 8.1 to 8.3, 10.1 to 10.5, 11.1 to 11.4, 12.1 to 12.3, 13.1 to 13.5, 15.1 to 15.3, 16.1 to 16.3, 19.1 to 19.3, 22.1 to 22.4)

MODULE 1: TABLE OF CONTENTS

MODULE 1: INTRODUCTION

1.1 DATA COMMUNICATIONS

- Data communication is defined as exchange of data between 2 devices over a transmission-medium.
- A communication-system is made up of
  - → hardware (physical equipment) and
  - → software (programs)
- For data-communication, the communicating-devices must be part of a communication-system.
- Four attributes of a communication-system:

1) Delivery

- The system must deliver data to the correct destination.

2) Accuracy

- The system must deliver the data accurately.
- Normally, the corrupted-data are unusable.

3) Timeliness

- The system must deliver audio/video data in a timely manner.
- This kind of delivery is called real-time transmission.
- Data delivered late are useless.

4) Jitter

- Jitter refers to the variation in the packet arrival-time.
- In other words, jitter is the uneven delay in the delivery of audio/video packets.

1.1.1 Components of Communication System

• Five components of a communication-system (Figure 1.1):

1) Message
2) Sender
3) Receiver
4) Transmission-Medium
5) Protocol

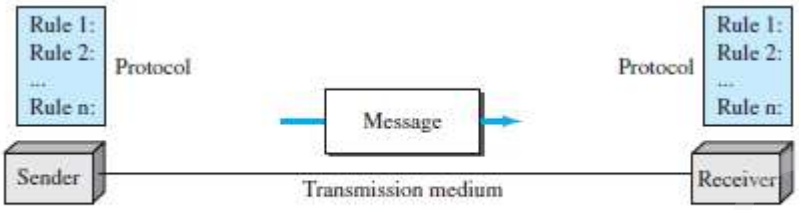


Figure 1.1 *Five components of data communication*

1) Message

- Message is the information (or data) to be communicated.
- Message may consist of
  - → number/text
  - → picture or
  - → audio/video

2) Sender

- Sender is the device that sends the data-message.
- Sender can be
  - → computer and
  - → mobile phone

3) Receiver

- Receiver is the device that receives the message.
- Receiver can be
  - → computer and
  - → mobile phone

4) Transmission Medium

- Transmission-medium is physical-path by which a message travels from sender to receiver.
- Transmission-medium can be wired or wireless.
- Examples of wired medium:
  - → twisted-pair wire (used in landline telephone)
  - → coaxial cable (used in cable TV network)
  - → fiber-optic cable
- Examples of wireless medium:
  - → radio waves
  - → microwaves
  - → infrared waves (ex: operating TV using remote control)

5) Protocol

- A protocol is a set of rules that govern data-communications.
- In other words, a protocol represents an agreement between the communicating-devices.
- Without a protocol, 2 devices may be connected but not communicating.

1.1.2 Data Representation

• Five different forms of information:

1) Text

- Text is represented as a bit-pattern. (Bit-pattern → sequence of bits: 0s or 1s).
- Different sets of bit-patterns are used to represent symbols (or characters).
- Each set is called a code.
- The process of representing symbols is called encoding.
- Popular encoding system: ASCII, Unicode.

2) Number

- Number is also represented as a bit-pattern.
- ASCII is not used to represent number. Instead, number is directly converted to binary-form.

3) Image

- Image is also represented as a bit-pattern.
- An image is divided into a matrix of pixels (picture-elements).
- A pixel is the smallest element of an image. (Pixel → Small dot)
- The size of an image depends upon number of pixels (also called resolution).
  For example: An image can be divided into 1000 pixels or 10,000 pixels.
- Two types of images:

  i) Black & White Image

  ¤ If an image is black & white, each pixel can be represented by a value either 0 or 1.

  ¤ For example: Chessboard

  ii) Color Image

  ¤ There are many methods to represent color images.

  ¤ RGB is one of the methods to represent color images.

  ¤ RGB is called so called '.' each color is combination of 3 colors: red, green & blue.

4) Audio

- Audio is a representation of sound.
- By nature, audio is different from text, numbers, or images. Audio is continuous, not discrete.

5) Video

- Video is a representation of movie.
- Video can either

  → be produced as a continuous entity (e.g., by a TV camera), or

  → be a combination of images arranged to convey the idea of motion.

1.1.3 Direction of Data Flow

• Three ways of data-flow between 2 devices (Figure 1.2):
1) Simplex
2) Half-duplex
3) Full-duplex

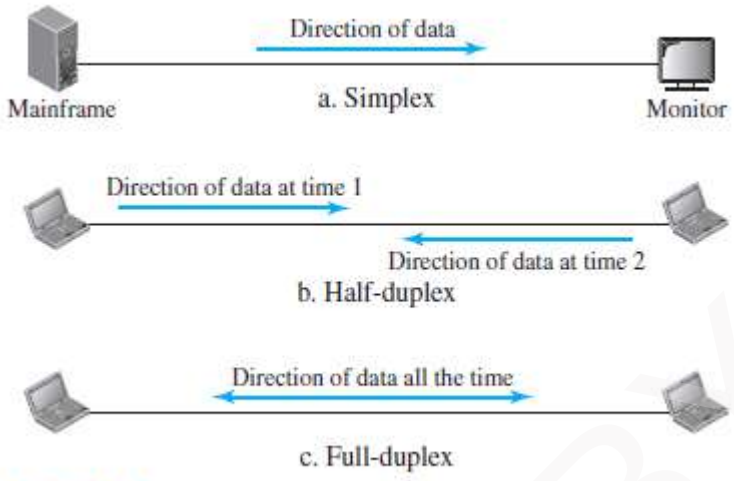


Figure 1.2 *Data flow (simplex, half-duplex, and full-duplex)*

1) Simplex
- The communication is unidirectional
  (For ex: The simplex mode is like a one-way street).
- On a link, out of 2 devices:
  i) Only one device can transmit.
  ii) Another device can only receive.
- For example (Figure 1.2a):
  The monitor can only accept output.
- Entire-capacity of channel is used to send the data in one direction.

2) Half Duplex
- Both the stations can transmit as well as receive but not at the same time.
  (For ex: The half-duplex mode is like a one-lane road with 2 directional traffic).
- When one station is sending, the other can only receive and vice-versa.
- For example (Figure 1.2b): Walkie-talkies
- Entire-capacity of a channel is used by one of the 2 stations that are transmitting the data.

3) Full Duplex
- Both stations can transmit and receive at the same time.
  (For ex: The full-duplex is like a 2-way street with traffic flowing in both directions at the same time).
- For example (Figure 1.2c):
  Mobile phones (When 2 people are communicating by a telephone line, both can listen and talk at the same time)
- Entire-capacity of a channel is shared by both the stations that are transmitting the data.

1.2 NETWORKS

- A network is defined as a set of devices interconnected by communication-links.
- This interconnection among computers facilitates information sharing among them.
- Computers may connect to each other by either wired or wireless media.
- Often, devices are referred to as nodes.
- A node can be any device capable of sending/receiving data in the network.
- For example: Computer & Printer
- The best-known computer network is the Internet.

1.2.1 Network Criteria

- A network must meet following 3 criteria's:

1) Performance

- Performance can be measured using i) Transit-time or ii) Response-time.
  - **i) Transit Time** is defined as time taken to travel a message from one device to another.
  - **ii) Response Time** is defined as the time elapsed between enquiry and response.
- The network-performance depends on following factors:
  - i) Number of users
  - ii) Type of transmission-medium
  - iii) Efficiency of software
- Often, performance is evaluated by 2 networking-metrics: i) throughput and ii) delay.
- Good performance can be obtained by achieving higher throughput and smaller delay times

2) Reliability

- Reliability is measured by
  - → frequency of network-failure
  - → time taken to recover from a network-failure
  - → network's robustness in a disaster
- More the failures are, less is the network's reliability.

3) Security

- Security refers to the protection of data from the unauthorized access or damage.
- It also involves implementing policies for recovery from data-losses.

1.2.2 Physical Structures

1.2.2.1 Type of Connection

- Two types of connections (Figure 1.3):

1) Point-to-Point

- Only two devices are connected by a dedicated-link (Figure 1.3a).
- Entire-capacity of the link is reserved for transmission between those two devices.
- For example: Point-to-Point connection b/w remote-control & TV for changing the channels.

2) Multipoint (Multi-Drop)

- Three or more devices share a single link.
- The capacity of the channel is shared, either spatially or temporally (Figure 1.3b).
  - i) If link is used simultaneously by many devices, then it is spatially shared connection.
  - ii) If user takes turns while using the link, then it is time shared (temporal) connection.
    (spatially→space or temporally→time)

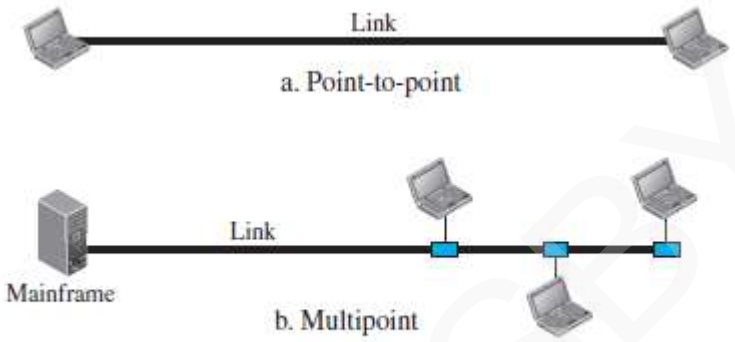


Figure 1.3 *Types of connections: point-to-point and multipoint*

1.2.2.2 Physical Topology

- The physical-topology defines how devices are connected to make a network.
- Four basic topologies are:
    1) Mesh
    2) Star
    3) Bus and
    4) Ring

1.2.2.2.1 Bus Topology

- All the devices are connected to the single cable called bus (Figure 1.4).
- Every device communicates with the other device through this bus.
- A data from the source is broadcasted to all devices connected to the bus.
- Only the intended-receiver, whose physical-address matches, accepts the data.

Figure 1.4 *A bus topology connecting three stations*

- Devices are connected to the bus by drop-lines and taps.
- A drop-line is a connection running between the device and the bus.
- A tap is a connector that links to the bus or
- Advantages:
    1) Easy installation.
    2) Cable required is the least compared to mesh/star topologies.
    3) Redundancy is eliminated.
    4) Costs less (Compared to mesh/star topologies).
    5) Mostly used in small networks. Good for LAN.
- Disadvantages:
    1) Difficult to detect and troubleshoot fault.
    2) Signal reflection at the taps can cause degradation in quality.
    3) A fault/break in the cable stops all transmission.
    4) There is a limit on
        i) Cable length
        ii) Number of nodes that can be connected.
    5) Security is very low because all the devices receive the data sent from the source.

1.2.2.2.2 Star Topology

- All the devices are connected to a central controller called a hub (Figure 1.5).
- There exists a dedicated point-to-point link between a device & a hub.
- The devices are not directly linked to one another. Thus, there is no direct traffic between devices.
- The hub acts as a junction:

  If device-1 wants to send data to device-2,
  the device-1 sends the data to the hub,
  then the hub relays the data to the device-2.

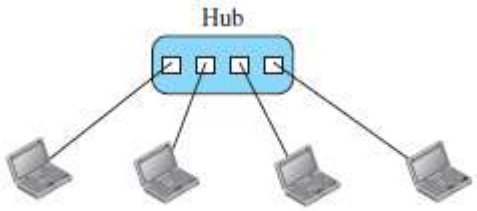


Figure 1.5 *A star topology connecting four stations*

- Advantages:
  1) Less expensive: Each device needs only one link & one I/O port to connect it to any devices.
  2) Easy installation & reconfiguration: Nodes can be added/removed w/o affecting the network.
  3) Robustness: If one link fails, it does not affect the entire system.
  4) Easy to detect and troubleshoot fault.
  5) Centralized management: The hub manages and controls the whole network.
- Disadvantages:
  1) Single point of failure: If the hub goes down, the whole network is dead.
  2) Cable length required is the more compared to bus/ring topologies.
  3) Number of nodes in network depends on capacity of hub.

1.2.2.2.3 Ring Topology

- Each device is connected to the next, forming a ring (Figure 1.6).
- There are only two neighbors for each device.
- Data travels around the network in one direction till the destination is reached.
- Sending and receiving of data takes place by the help of token.
- Each device has a repeater.
- A repeater
  → receives a signal on transmission-medium &
  → regenerates & passes the signal to next device.

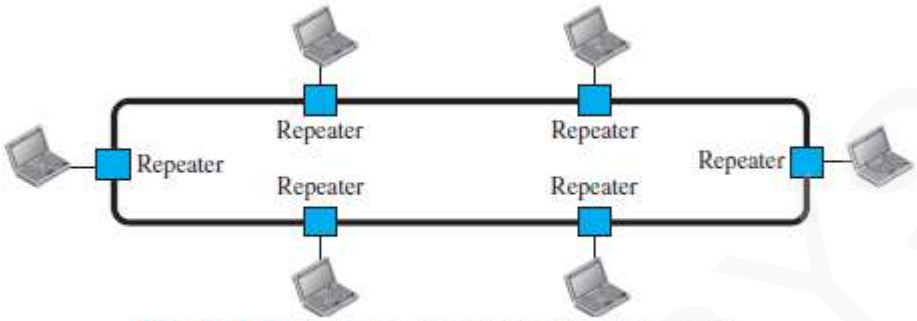


Figure 1.6 *A ring topology connecting six stations*

- Advantages:
  1) Easy installation and reconfiguration.
     To add/delete a device, requires changing only 2 connections.
  3) Fault isolation is simplified.
     If one device does not receive a signal within a specified period, it can issue an alarm.
     The alarm alerts the network-operator to the problem and its location.
  3) Congestion reduced: Because all the traffic flows in only one direction.
- Disadvantages:
  1) Unidirectional traffic.
  2) A fault in the ring/device stops all transmission.
     The above 2 drawbacks can be overcome by using dual ring.
  3) There is a limit on
     i) Cable length &
     ii) Number of nodes that can be connected.
  4) Slower: Each data must pass through all the devices between source and destination.

1.2.2.2.4 Mesh Topology

- All the devices are connected to each other (Figure 1.7).
- There exists a dedicated point-to-point link between all devices.
- There are n(n-1) physical channels to link n devices.
- Every device not only sends its own data but also relays data from other nodes.
- For 'n' nodes,
  - → there are n(n-1) physical-links
  - → there are n(n-1)/2 duplex-mode links
- Every device must have (n–1) I/O ports to be connected to the other (n-1) devices.

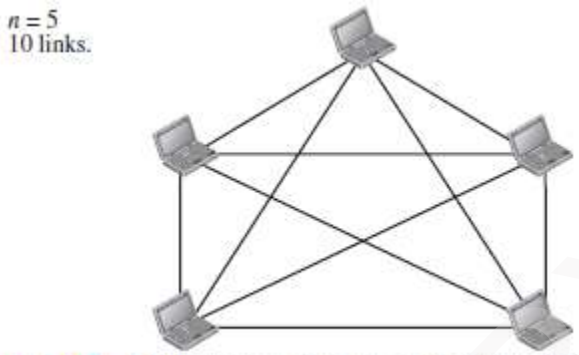


Figure 1.7 *A fully connected mesh topology (five devices)*

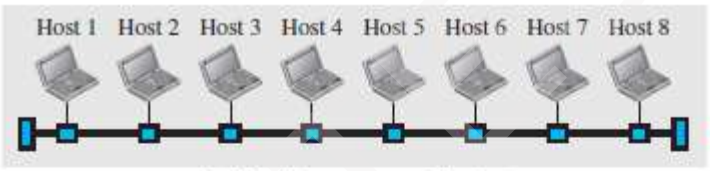
- Advantages:
  1) Congestion reduced: Each connection can carry its own data load.
  2) Robustness: If one link fails, it does not affect the entire system.
  3) Security: When a data travels on a dedicated-line, only intended-receiver can see the data.
  4) Easy fault identification & fault isolation: Traffic can be re-routed to avoid problematic links.
- Disadvantages:
  1) Difficult installation and reconfiguration.
  2) Bulk of wiring occupies more space than available space.
  3) Very expensive: as there are many redundant connections.
  4) Not mostly used in computer networks. It is commonly used in wireless networks.
  5) High redundancy of the network-connections.

1.3 Network Types

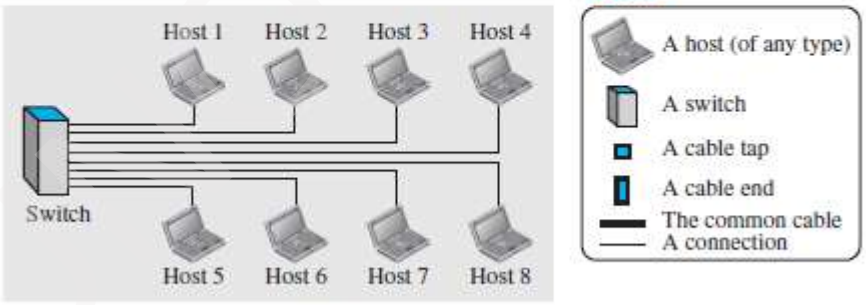
- Two popular types of networks:
  1) LAN (Local Area Network) &
  2) WAN (Wide Area Network)

1.3.1 LAN

- LAN is used to connect computers in a single office, building or campus (Figure 1.8).
- LAN is usually privately owned network.
- A LAN can be simple or complex.
  1) Simple: LAN may contain 2 PCs and a printer.
  2) Complex: LAN can extend throughout a company.
- Each host in a LAN has an address that uniquely defines the host in the LAN.
- A packet sent by a host to another host carries both source host's and destination host's addresses.
- LANs use a smart connecting switch.
- The switch is able to
  → recognize the destination address of the packet &
  → guide the packet to its destination.
- The switch
  → reduces the traffic in the LAN &
  → allows more than one pair to communicate with each other at the same time.
- Advantages:
  1) Resource Sharing
  ➢ Computer resources like printers and hard disks can be shared by all devices on the network.
  2) Expansion
  ➢ Nowadays, LANs are connected to WANs to create communication at a wider level.



a. LAN with a common cable (past)



b. LAN with a switch (today)

Figure 1.8 *An isolated LAN in the past and today*

1.3.2 WAN

- WAN is used to connect computers anywhere in the world.
- WAN can cover larger geographical area. It can cover cities, countries and even continents.
- WAN interconnects connecting devices such as switches, routers, or modems.
- Normally, WAN is
  → created & run by communication companies (Ex: BSNL, Airtel)
  → leased by an organization that uses it.
- A WAN can be of 2 types:

1) Point-to-Point WAN

➢ A point-to-point WAN is a network that connects 2 communicating devices through a transmission media (Figure 1.9).

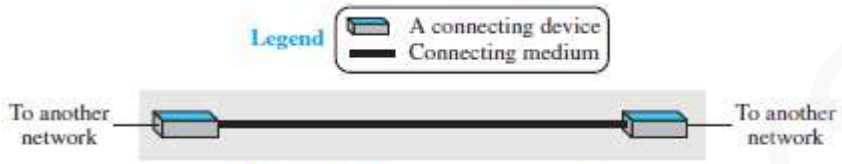


Figure 1.9 *A point-to-point WAN*

2) Switched WAN

➢ A switched WAN is a network with more than two ends.
➢ The switched WAN can be the backbones that connect the Internet.
➢ A switched WAN is a combination of several point-to-point WANs that are connected by switches (Figure 1.10).

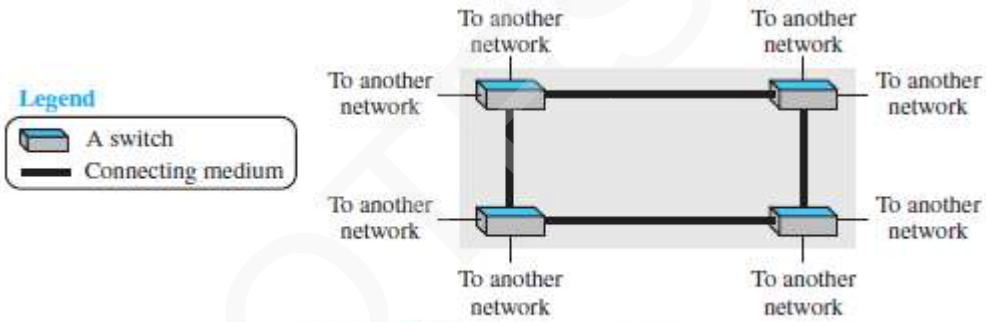


Figure 1.10 *A switched WAN*

1.3.2.1 Internetwork

- A network of networks is called an internet. (Internet → inter-network) (Figure 1.12).
- For example (Figure 1.11):
  - Assume that an organization has two offices,
    - i) First office is on the east coast &
    - ii) Second office is on the west coast.
  - Each office has a LAN that allows all employees in the office to communicate with each other.
  - To allow communication between employees at different offices, the management leases a point-to-point dedicated WAN from a ISP and connects the two LANs.
    (ISP → Internet service provider such as a telephone company ex: BSNL).

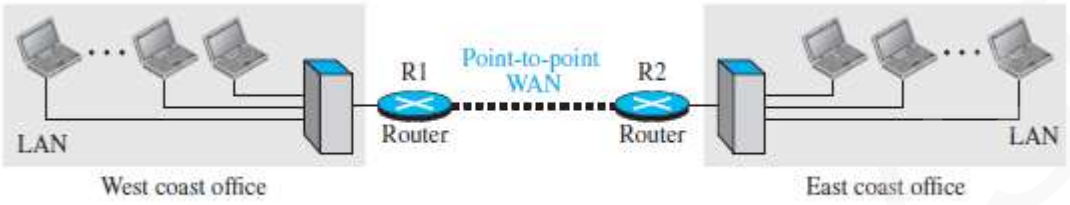


Figure 1.11 *An internetwork made of two LANs and one point-to-point WAN*

- 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.

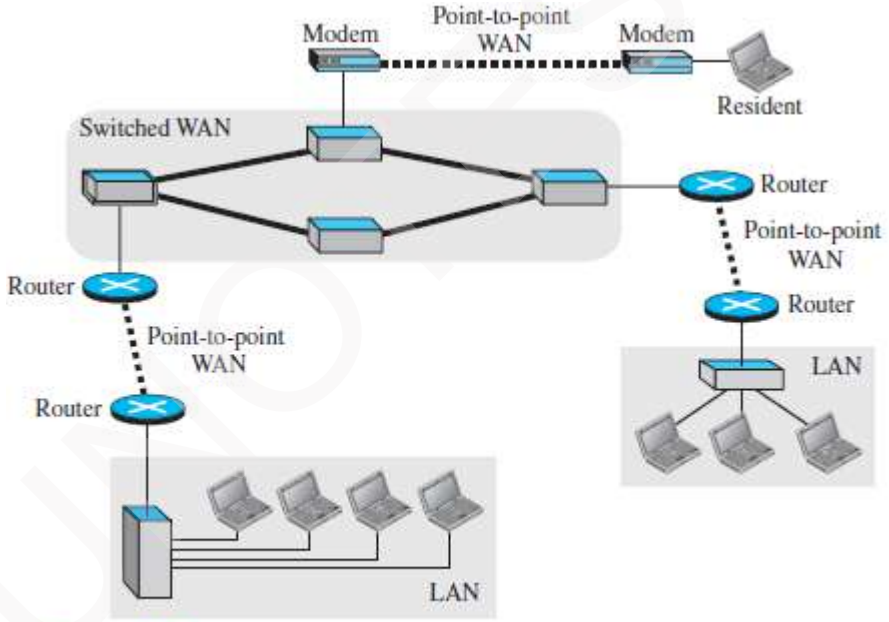


Figure 1.12 *A heterogeneous network made of four WANs and three LANs*

1.3.3 LAN vs. WAN

Parameters	LAN	WAN
Expands to	Local Area Network	Wide Area Network
Meaning	LAN is used to connect computers in a single office, building or campus	WAN is used to connect computers in a large geographical area such as countries
Ownership of network	Private	Private or public
Range	Small: up to 10 km	Large: Beyond 100 km
Speed	High: Typically 10, 100 and 1000 Mbps	Low: Typically 1.5 Mbps
Propagation Delay	Short	Long
Cost	Low	High
Congestion	Less	More
Design & maintenance	Easy	Difficult
Fault Tolerance	More Tolerant	Less Tolerant
Media used	Twisted pair	Optical fiber or radio waves
Used for	College, Hospital	Internet
Interconnects	LAN interconnects hosts	WAN interconnects connecting devices such as switches, routers, or modems

1.3.4 Switching

- An internet is a switched network in which a switch connects at least two links together.
- A switch needs to forward data from a network to another network when required.
- Two types of switched networks are 1) circuit-switched and 2) packet-switched networks.

1.3.4.1 Circuit Switched Network

- A dedicated connection, called a circuit, is always available between the two end systems.
- The switch can only make it active or inactive.

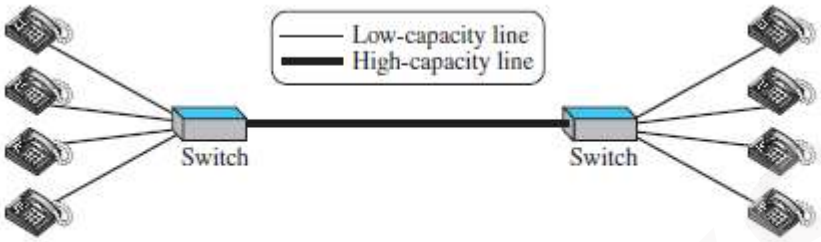


Figure 1.13 *A circuit-switched network*

- As shown in Figure 1.13, the 4 telephones at each side are connected to a switch.
- The switch connects a telephone at one side to a telephone at the other side.
- A high-capacity line can handle 4 voice communications at the same time.
- The capacity of high line can be shared between all pairs of telephones.
- The switch is used for only forwarding.

- Advantage:

  A circuit-switched network is efficient only when it is working at its full capacity.

- Disadvantage:

  Most of the time, the network is inefficient because it is working at partial capacity.

1.3.4.2 Packet Switched Network

- In a computer network, the communication between the 2 ends is done in blocks of data called packets.
- The switch is used for both storing and forwarding because a packet is an independent entity that can be stored and sent later.

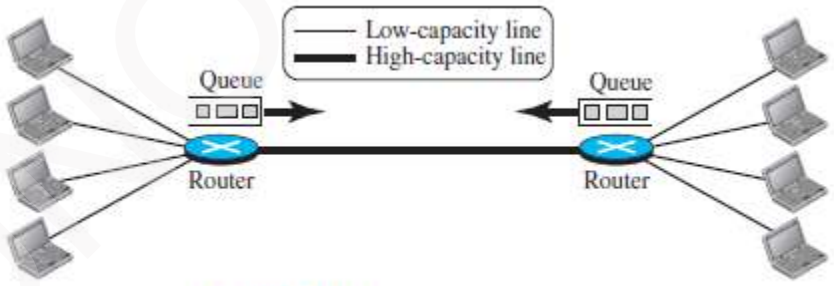


Figure 1.14 *A packet-switched network*

- As shown in Figure 1.14, the 4 computers at each side are connected to a router.
- A router has a queue that can store and forward the packet.
- The high-capacity line has twice the capacity of the low-capacity line.
- If only 2 computers (one at each site) need to communicate with each other, there is no waiting for the packets.
- However, if packets arrive at one router when high-capacity line is at its full capacity, the packets should be stored and forwarded.

- Advantages:

  A packet-switched network is more efficient than a circuit switched network.

- Disadvantage:

  The packets may encounter some delays.

1.3.5 The Internet Today

- A network of networks is called an internet. (Internet → inter-network)
- Internet is made up of (Figure 1.15)
  1) Backbones
  2) Provider networks &
  3) Customer networks

1) Backbones

- Backbones are large networks owned by communication companies such as BSNL and Airtel.
- The backbone networks are connected through switching systems, called peering points.

2) Provider Networks

- Provider networks use the services of the backbones for a fee.
- Provider networks are connected to backbones and sometimes to other provider networks.

3) Customer Networks

- Customer networks actually use the services provided by the Internet.
- Customer networks pay fees to provider networks for receiving services.

- Backbones and provider networks are also called Internet Service Providers (ISPs).
- The backbones are often referred to as international ISPs.
  The provider networks are often referred to as national or regional ISPs.

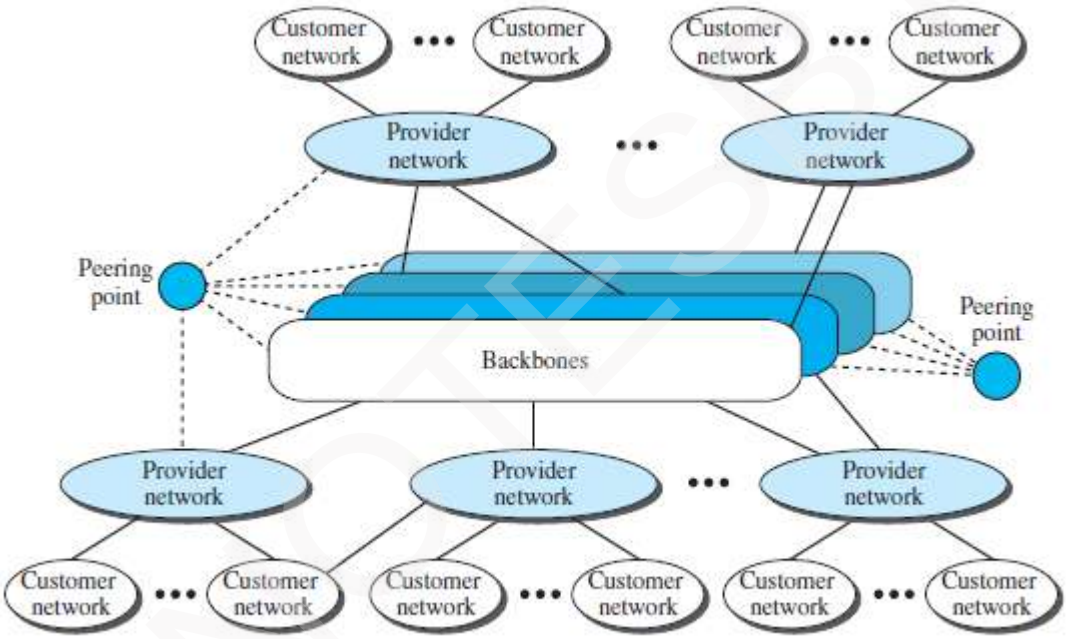


Figure 1.15 *The Internet today*

1.3.6 Accessing the Internet

- The Internet today is an internetwork that allows any user to become part of it.
- However, the user needs to be physically connected to an ISP.
- The physical connection is normally done through a point-to-point WAN.

1) Using Telephone Networks

➢ Most residences have telephone service, which means they are connected to a telephone network.

➢ Most telephone networks have already connected themselves to the Internet.

➢ Thus, residences can connect to the Internet using a point-to-point WAN.

➢ This can be done in two ways:

A) Dial-up service

¤ A modem can be added to the telephone line.

¤ A modem converts data to voice.

¤ The software installed on the computer

→ dials the ISP &

→ imitates making a telephone connection.

¤ Disadvantages:

i) The dial-up service is very slow.

ii) When line is used for Internet connection, it cannot be used for voice connection.

iii) It is only useful for small residences.

B) DSL Service

¤ DSL service also allows the line to be used simultaneously for voice & data communication.

¤ Some telephone companies have upgraded their telephone lines to provide higher speed Internet services to residences.

2) Using Cable Networks

➢ A residence can be connected to the Internet by using cable service.

➢ Cable service provides a higher speed connection.

➢ The speed varies depending on the number of neighbors that use the same cable.

3) Using Wireless Networks

➢ A residence can use a combination of wireless and wired connections to access the Internet.

➢ A residence can be connected to the Internet through a wireless WAN.

4) Direct Connection to the Internet

➢ A large organization can itself become a local ISP and be connected to the Internet.

➢ The organization

→ leases a high-speed WAN from a carrier provider and

→ connects itself to a regional ISP.

1.4 STANDARDS AND ADMINISTRATION

1.4.1 Internet Standards

- An Internet standard is a thoroughly tested specification useful to those who work with the Internet.
- The Internet standard is a formalized-regulation that must be followed.
- There is a strict procedure by which a specification attains Internet standard status.
- A specification begins as an Internet draft.
- An Internet draft is a working document with no official status and a 6-month lifetime.
- Upon recommendation from the Internet authorities, a draft may be published as a RFC.
- Each RFC is edited, assigned a number, and made available to all interested parties.
- RFCs go through maturity levels and are categorized according to their requirement level.

(working document → a work in progress RFC → Request for Comment)

1.4.1.1 Maturity Levels

- An RFC, during its lifetime, falls into one of 6 maturity levels (Figure 1.16):

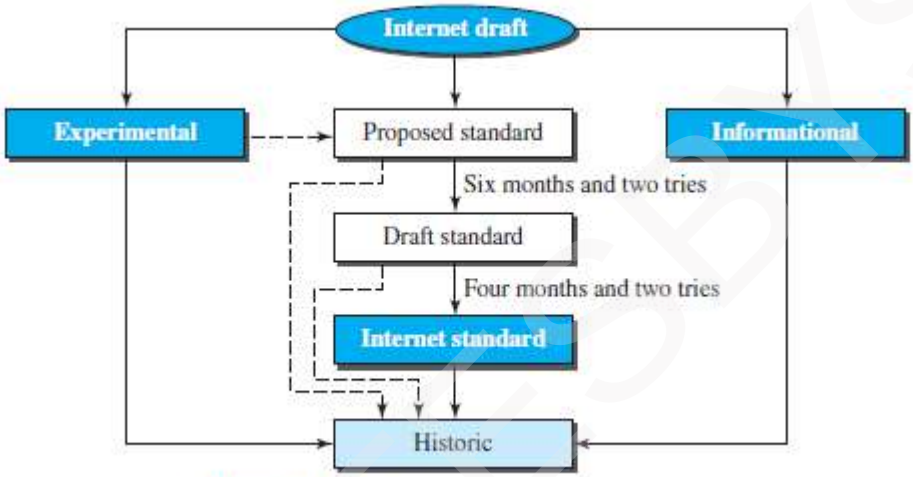


Figure 1.16 *Maturity levels of an RFC*

1) Proposed Standard

➢ Proposed standard is specification that is stable, well-understood & of interest to Internet community.

➢ Specification is usually tested and implemented by several different groups.

2) Draft Standard

➢ A proposed standard is elevated to draft standard status after at least 2 successful independent and interoperable implementations.

3) Internet Standard

➢ A draft standard reaches Internet standard status after demonstrations of successful implementation.

4) Historic

➢ The historic RFCs are significant from a historical perspective.

➢ They either

→ have been superseded by later specifications or

→ have never passed the necessary maturity levels to become an Internet standard.

5) Experimental

➢ An RFC classified as experimental describes work related to an experimental situation.

➢ Such an RFC should not be implemented in any functional Internet service.

6) Informational

➢ An RFC classified as informational contains general, historical, or tutorial information related to the Internet.

➢ Usually, it is written by a vendor.

(ISOC → Internet Society IAB → Internet Architecture Board)
(IETF → Internet Engineering Task Force IRTF → Internet Research Task Force)
(IESG → Internet Engineering Steering Group IRSG → Internet Research Steering Group)

1.4.1.2 Requirement Levels

• RFCs are classified into 5 requirement levels:

1) Required

➢ An RFC labeled required must be implemented by all Internet systems to achieve minimum conformance.

➢ For example, IP and ICMP are required protocols.

2) Recommended

➢ An RFC labeled recommended is not required for minimum conformance.

➢ It is recommended because of its usefulness.

➢ For example, FTP and TELNET are recommended protocols.

3) Elective

➢ An RFC labeled elective is not required and not recommended.

➢ However, a system can use it for its own benefit.

4) Limited Use

➢ An RFC labeled limited use should be used only in limited situations.

➢ Most of the experimental RFCs fall under this category.

5) Not Recommended

➢ An RFC labeled not recommended is inappropriate for general use.

➢ Normally a historic RFC may fall under this category.

1.4.2 Internet Administration

1) ISOC

• ISOC is a nonprofit organization formed to provide support for Internet standards process (Fig 1.17).

• ISOC maintains and supports other Internet administrative bodies such as IAB, IETF, IRTF, and IANA.

2) IAB

• IAB is the technical advisor to the ISOC.

• Two main purposes of IAB:

i) To oversee the continuing development of the TCP/IP Protocol Suite

ii) To serve in a technical advisory capacity to research members of the Internet community.

• Another responsibility of the IAB is the editorial management of the RFCs.

• IAB is also the external liaison between the Internet and other standards organizations and forums.

• IAB has 2 primary components: i) IETF and ii) IRTF.

i) IETF

➢ IETF is a forum of working groups managed by the IESG.

➢ IETF is responsible for identifying operational problems & proposing solutions to the problems

➢ IETF also develops and reviews specifications intended as Internet standards.

➢ The working groups are collected into areas, and each area concentrates on a specific topic.

➢ Currently 9 areas have been defined. The areas include applications, protocols, routing, network management next generation (IPng), and security.

ii) IRTF

➢ IRTF is a forum of working groups managed by the IRSG.

➢ IRTF focuses on long-term research topics related to Internet protocols, applications, architecture, and technology.

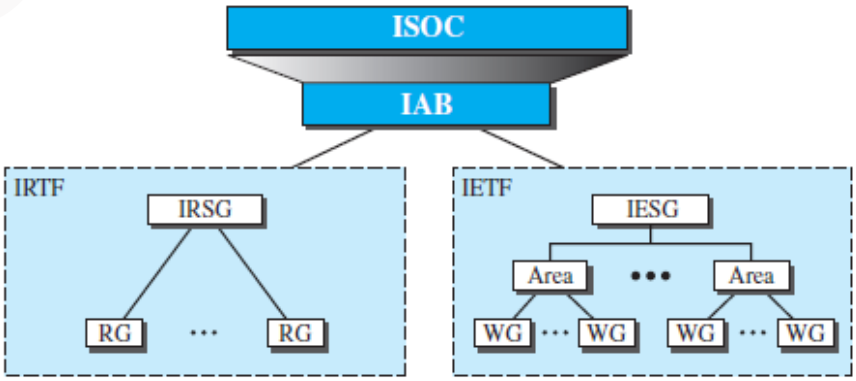


Figure 1.17 *Internet administration*

MODULE 1(CONT.): NETWORK MODELS

1.5 PROTOCOL LAYERING

• 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 communication is complex, we need to divide the task b/w different layers. We need a protocol at each layer, or protocol layering.

1.5.1 Scenarios

First Scenario

• In the first scenario, communication is so simple that it can occur in only one layer (Figure 2.1).
• 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

Figure 2.1 *A single-layer protocol*

Second Scenario

• Maria and Ann communicate using regular mail through the post office (Figure 2.2).
• 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.

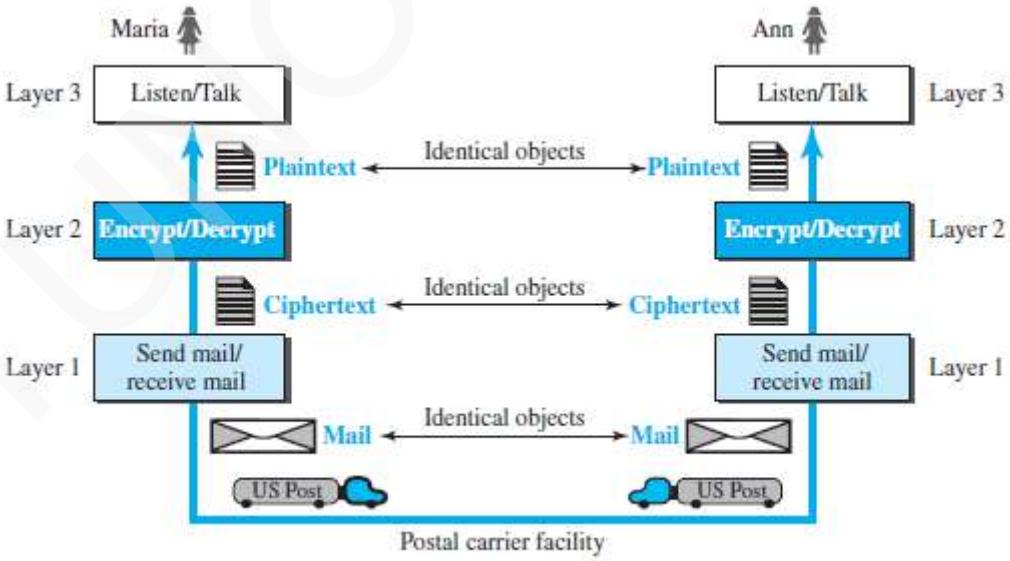


Figure 2.2 *A three-layer protocol*

1.5.1.1 Protocol Layering

• Protocol layering enables us to divide a complex task into several smaller and simpler tasks.
• Modularity means independent layers.
• A layer (module) can be defined as a black box with inputs and outputs, without concern about how inputs are changed to outputs.
• If two machines provide the same outputs when given the same inputs, they can replace each other.
• Advantages:
    1) It allows us to separate the services from the implementation.
    2) There are intermediate systems that need only some layers, but not all layers.
• Disadvantage:
    1) Having a single layer makes the job easier. There is no need for each layer to provide a service to the upper layer and give service to the lower layer.

1.5.2 Principles of Protocol Layering

1) First Principle

• If we want bidirectional communication, we need to make each layer able to perform 2 opposite tasks, one in each direction.
• For example, the third layer task is to listen (in one direction) and talk (in the other direction).

2) Second Principle

• The two objects under each layer at both sites should be identical.
• For example, the object under layer 3 at both sites should be a plaintext letter.

1.5.3 Logical Connections

• We have layer-to-layer communication (Figure 2.3).
• There is a logical connection at each layer through which 2 end systems can send the object created from that layer.

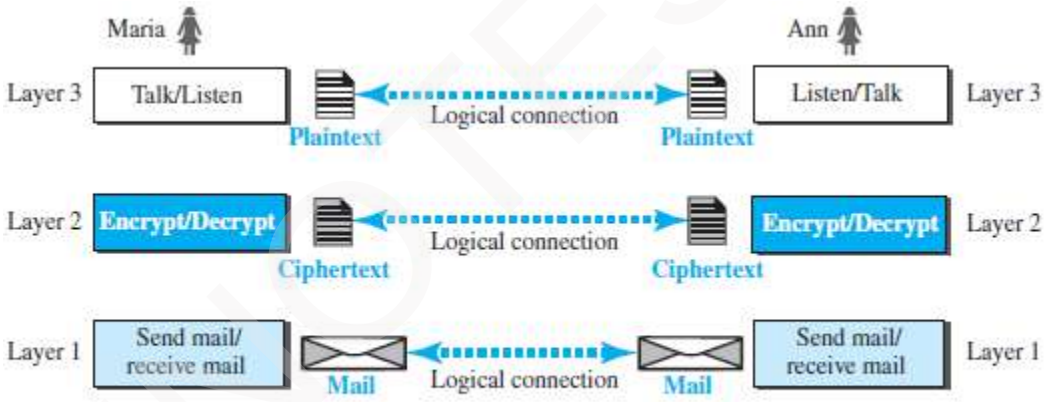


Figure 2.3 *Logical connection between peer layers*

1.6 TCP/IP PROTOCOL SUITE

• TCP/IP is a protocol-suite used in the Internet today.
• Protocol-suite refers a set of protocols organized in different layers.
• It is a hierarchical protocol made up of interactive modules, each of which provides a specific functionality.
• The term hierarchical means that each upper level protocol is supported by the services provided by one or more lower level protocols.

1.6.1 Layered Architecture

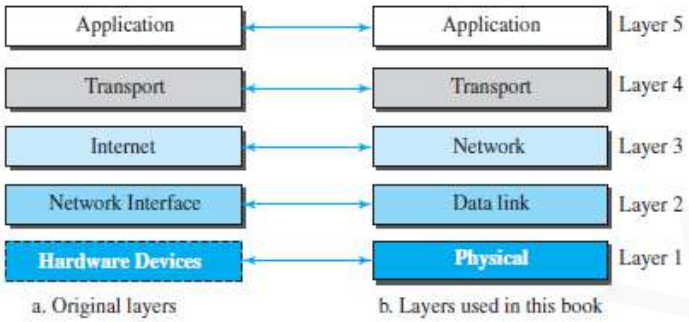


Figure 2.4 *Layers in the TCP/IP protocol suite*

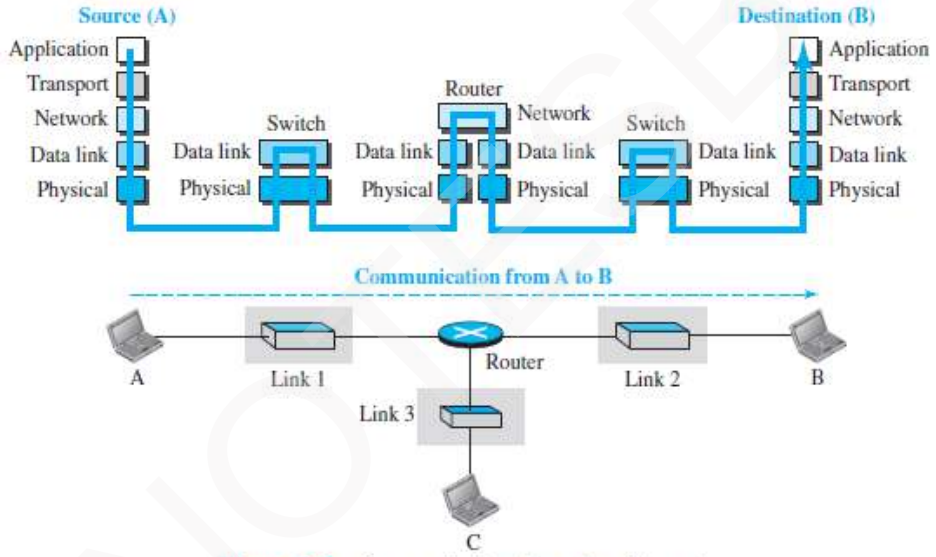


Figure 2.5 *Communication through an internet*

• Let us assume that computer A communicates with computer B (Figure 2.4).
• As the Figure 2.5 shows, we have five communicating devices:
  1) Source host(computer A)
  2) Link-layer switch in link 1
  3) Router
  4) Link-layer switch in link 2
  5) Destination host (computer B).
• Each device is involved with a set of layers depending on the role of the device in the internet.
• The two hosts are involved in all five layers.
• The source host
  → creates a message in the application layer and
  → sends the message down the layers so that it is physically sent to the destination host.
• The destination host
  → receives the message at the physical layer and
  → then deliver the message through the other layers to the application layer.
• The router is involved in only three layers; there is no transport or application layer.
• A router is involved in n combinations of link and physical layers.
  where n = number of links the router is connected to.
• The reason is that each link may use its own data-link or physical protocol.
• A link-layer switch is involved only in two layers: i) data-link and ii) physical.

1.6.2 Layers in the TCP/IP Protocol Suite

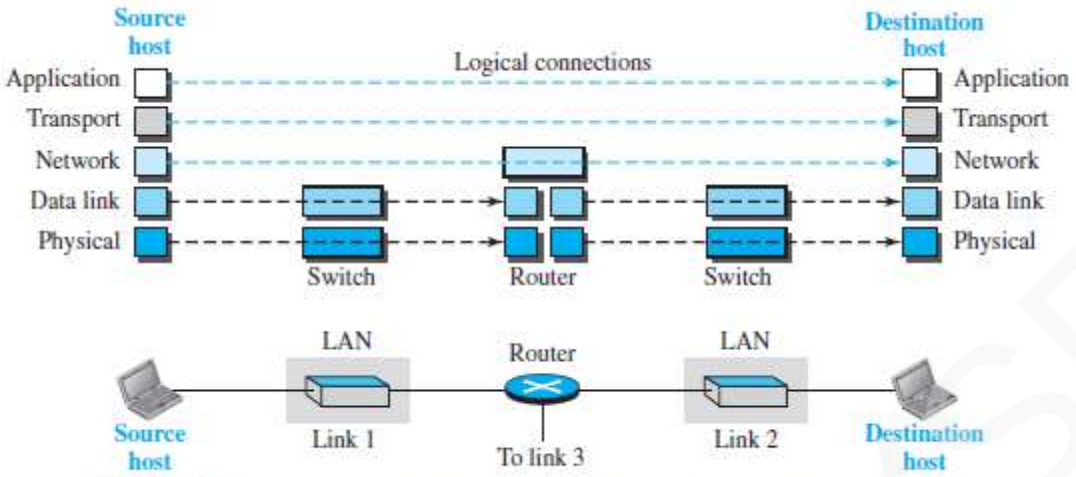


Figure 2.6 *Logical connections between layers of the TCP/IP protocol suite*

- As shown in the figure 2.6, the duty of the application, transport, and network layers is end-to-end.
- However, the duty of the data-link and physical layers is hop-to-hop. A hop is a host or router.
- The domain of duty of the top three layers is the internet.
  The domain of duty of the two lower layers is the link.
- In top 3 layers, the data unit should not be changed by any router or link-layer switch.
  In bottom 2 layers, the data unit is changed only by the routers, not by the link-layer switches.

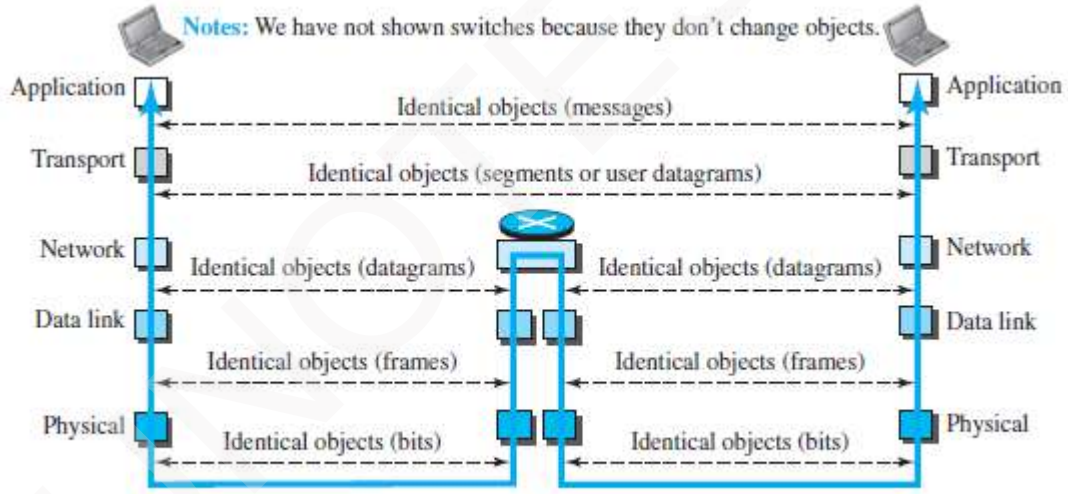


Figure 2.7 *Identical objects in the TCP/IP protocol suite*

- Identical objects exist between two hops. Because router may fragment the packet at the network layer and send more packets than received (Figure 2.7).
- The link between two hops does not change the object.

1.6.3 Description of Each Layer

Physical Layer

- The physical layer is responsible for movements of individual bits from one node to another node.
- Transmission media is another hidden layer under the physical layer.
- Two devices are connected by a transmission medium (cable or air).
- The transmission medium does not carry bits; it carries electrical or optical signals.
- The physical layer
  - → receives bits from the data-link layer &
  - → sends through the transmission media.

Data Link Layer

- Data-link-layer (DLL) is responsible for moving frames from one node to another node over a link.
- The link can be wired LAN/WAN or wireless LAN/WAN.
- The data-link layer
  - → gets the datagram from network layer
  - → encapsulates the datagram in a packet called a frame.
  - → sends the frame to physical layer.
- TCP/IP model does not define any specific protocol.
- DLL supports all the standard and proprietary protocols.
- Each protocol may provide a different service.
- Some protocols provide complete error detection and correction; some protocols provide only error correction.

Network Layer

- The network layer is responsible for source-to-destination transmission of data.
- The network layer is also responsible for routing the packet.
- The routers choose the best route for each packet.
- Why we need the separate network layer?
  1) The separation of different tasks between different layers.
  2) The routers do not need the application and transport layers.
- TCP/IP model defines 5 protocols:
  1) IP (Internetworking Protocol)
  2) ARP (Address Resolution Protocol)
  3) ICMP (Internet Control Message Protocol)
  4) IGMP (Internet Group Message Protocol)

1) IP

- IP is the main protocol of the network layer.
- IP defines the format and the structure of addresses.
- IP is also responsible for routing a packet from its source to its destination.
- It is a connection-less & unreliable protocol.
  - i) Connection-less means there is no connection setup b/w the sender and the receiver.
  - ii) Unreliable protocol means
    - → IP does not make any guarantee about delivery of the data.
    - → Packets may get dropped during transmission.
- It provides a best-effort delivery service.
- Best effort means IP does its best to get the packet to its destination, but with no guarantees.
- IP does not provide following services
  - → flow control
  - → error control
  - → congestion control services.
- If an application requires above services, the application should rely only on the transport-layer protocol.

2) ARP

- ARP is used to find the physical-address of the node when its Internet-address is known.
- Physical address is the 48-bit address that is imprinted on the NIC or LAN card.
- Internet address (IP address) is used to uniquely & universally identify a device in the internet.

3) ICMP

- ICMP is used to inform the sender about datagram-problems that occur during transit.

4) IGMP

- IGMP is used to send the same message to a group of recipients.

Transport Layer

- TL protocols are responsible for delivery of a message from a process to another process.
- The transport layer
  - → gets the message from the application layer
  - → encapsulates the message in a packet called a segment and
  - → sends the segment to network layer.
- TCP/IP model defines 3 protocols: 1) TCP (Transmission Control Protocol)
  2) UDP (User Datagram Protocol) &
  3) SCTP (Stream Control Transmission Protocol)

1) TCP

- TCP is a reliable connection-oriented protocol.
- A connection is established b/w the sender and receiver before the data can be transmitted.
- TCP provides
  - → flow control
  - → error control and
  - → congestion control

2) UDP

- UDP is the simplest of the 3 transport protocols.
- It is an unreliable, connectionless protocol.
- It does not provide flow, error, or congestion control.
- Each datagram is transported separately & independently.
- It is suitable for application program that
  - → needs to send short messages &
  - → cannot afford the retransmission.

3) SCTP

- SCTP provides support for newer applications such as voice over the Internet.
- It combines the best features of UDP and TCP.

Application Layer

- The two application layers exchange messages between each other.
- 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 of the application layer.
- TCP/IP model defines following protocols:
  1) SMTP is used to transport email between a source and destination.
  2) TELNET is used for accessing a site remotely.
  3) FTP is used for transferring files from one host to another.
  4) DNS is used to find the IP address of a computer.
  5) SNMP is used to manage the Internet at global and local levels.
  6) HTTP is used for accessing the World Wide Web (WWW).

(FTP → File Transfer Protocol
(DNS → Domain Name System
(SNMP → Simple Network Management Protocol
SMTP → Simple Mail Transfer Protocol)
HTTP → Hyper Text Transfer Protocol)
TELNET → Terminal Network)

1.6.4 Encapsulation and Decapsulation

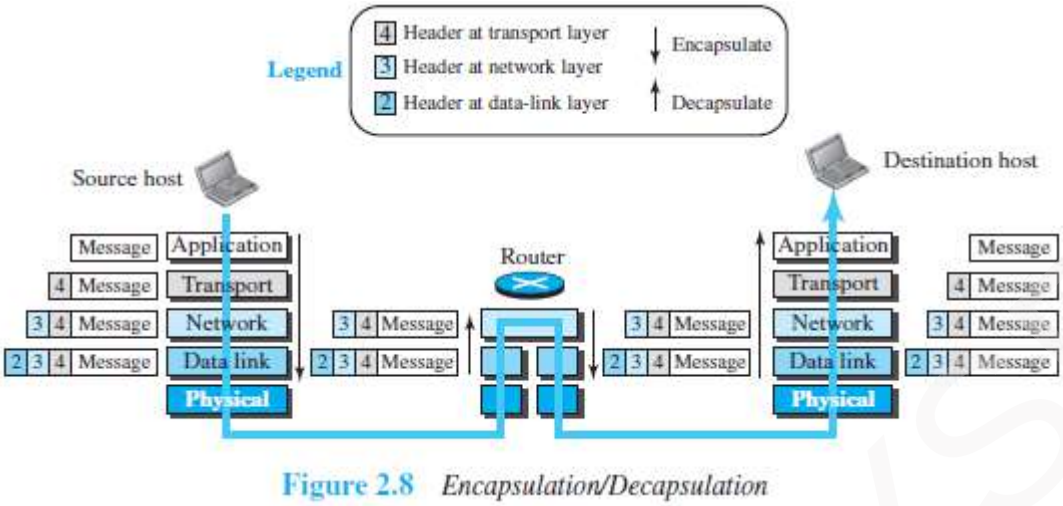


Figure 2.8 *Encapsulation/Decapsulation*

A) Encapsulation at the Source Host

• At the source, we have only encapsulation (Figure 2.8).

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.
- The message is passed to the transport layer.

2) The transport layer takes the message as the payload.
- TL adds its own header to the payload.
- The header contains
  - → identifiers of the source and destination application programs
  - → information needed for flow, error control, or congestion control.
- The transport-layer packet is called the segment (in TCP) and the user datagram (in UDP).
- The segment is passed to the network layer.

3) The network layer takes the transport-layer packet as payload.
- NL adds its own header to the payload.
- The header contains
  - → addresses of the source and destination hosts
  - → some information used for error checking of the header &
  - → fragmentation information.
- The network-layer packet is called a datagram.
- The datagram is passed to the data-link layer.

4) The data-link layer takes the network-layer packet as payload.
- DLL adds its own header to the payload.
- The header contains the physical addresses of the host or the next hop (the router).
- The link-layer packet is called a frame.
- The frame is passed to the physical layer for transmission

B) Decapsulation and Encapsulation at the Router

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

1) Data-link layer
- → receives frame from physical layer
- → decapsulates the datagram from the frame and
- → passes the datagram to the network layer.

2) The network layer
- → inspects the source and destination addresses in the datagram header and
- → consults forwarding table to find next hop to which the datagram is to be delivered.

- 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 the frame to the physical layer for transmission.

C) Decapsulation at the Destination Host

• At the destination host, each layer
→ decapsulates the packet received from lower layer
→ removes the payload and
→ delivers the payload to the next-higher layer

1.6.5 Addressing

• We have logical communication between pairs of layers.
• Any communication that involves 2 parties needs 2 addresses: source address and destination address.
• We need 4 pairs of addresses (Figure 2.9):

1) At the application layer, we normally use names to define
→ site that provides services, such as vtunotesbysri.com, or
→ e-mail address, such as vtunotesbysree@gmail.com.

2) At the transport layer, addresses are called port numbers.
➢ Port numbers define the application-layer programs at the source and destination.
➢ Port numbers are local addresses that distinguish between several programs running at the same time.

3) At the network-layer, addresses are called IP addresses.
➢ IP address uniquely defines the connection of a device to the Internet.
➢ The IP addresses are global, with the whole Internet as the scope.

4) At the data link-layer, addresses are called MAC addresses
➢ The MAC addresses defines a specific host or router in a network (LAN or WAN).
➢ The MAC addresses are locally defined addresses.

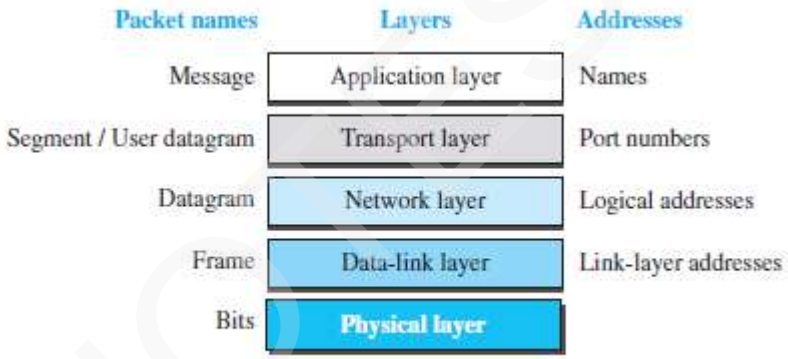


Figure 2.9 *Addressing in the TCP/IP protocol suite*

1.6.6 Multiplexing and Demultiplexing

• Multiplexing means a protocol at a layer can encapsulate a packet from several next-higher layer protocols (one at a time) (Figure 2.10).

• Demultiplexing means a protocol can decapsulate and deliver a packet to several next-higher layer protocols (one at a time).

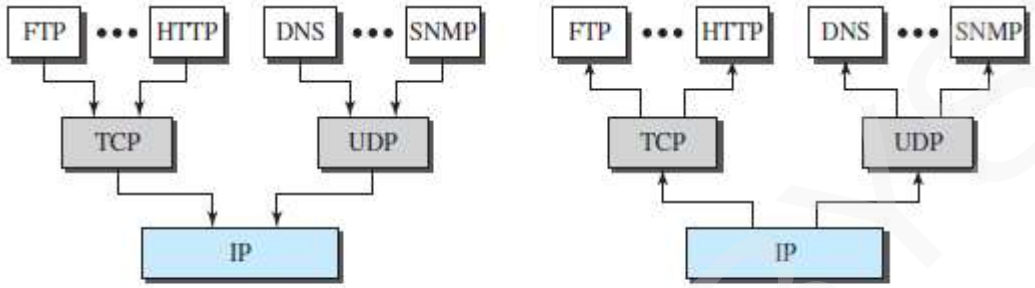
1) At transport layer, either UDP or TCP can accept a message from several application-layer protocols.

2) At network layer, IP can accept

→ a segment from TCP or a user datagram from UDP.

→ a packet from ICMP or IGMP.

3) At data-link layer, a frame may carry the payload coming from IP or ARP.



a. Multiplexing at source b. Demultiplexing at destination

Figure 2.10 *Multiplexing and demultiplexing*

1.7 OSI MODEL

- OSI model was developed by ISO.
- ISO is the organization, OSI is the model.
- Purpose: OSI was developed to allow systems with diff. platforms to communicate with each other.
- Platform means hardware, software or operating system.
- OSI is a network-model that defines the protocols for network communications.
- OSI has 7 layers as follows (Figure 2.11):
    1) Application Layer
    2) Presentation Layer
    3) Session Layer
    4) Transport Layer
    5) Network Layer
    6) Data Link Layer
    7) Physical Layer
- Each layer has specific duties to perform and has to co-operate with the layers above & below it.

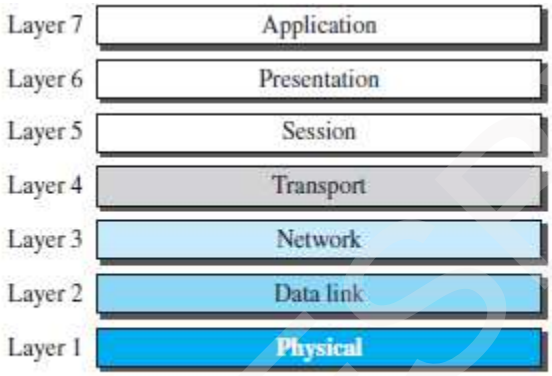


Figure 2.11 *The OSI model*

1.7.1 OSI vs. TCP/IP

1) The four bottommost layers in the OSI model & the TCP/IP model are same (Figure 2.12). However, the Application-layer of TCP/IP model corresponds to the Session, Presentation & Application Layer of OSI model.
    Two reasons for this are:
    1) TCP/IP has more than one transport-layer protocol.
    2) Many applications can be developed at Application layer
2) The OSI model specifies which functions belong to each of its layers.
    In TCP/IP model, the layers contain relatively independent protocols that can be mixed and matched depending on the needs of the system.

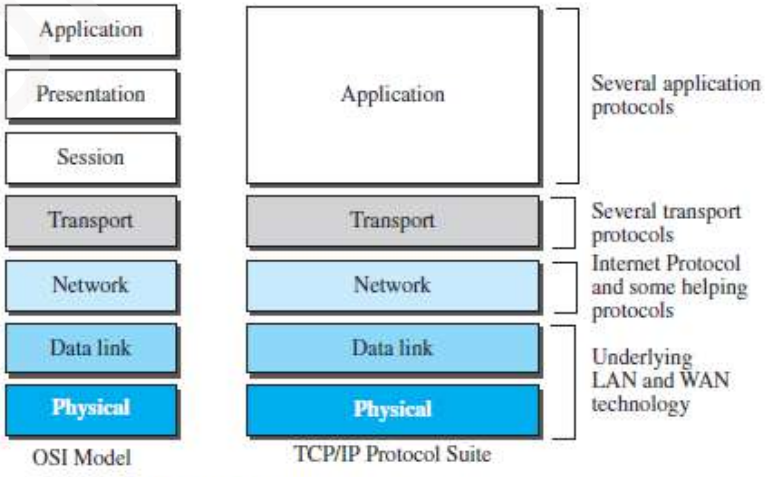


Figure 2.12 *TCP/IP and OSI model*

1.7.2 Lack of OSI Model's Success

• 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.
• Some layers in the OSI model were never fully defined.
• When OSI was implemented by an organization in a different application, it did not show a high enough level of performance

LAYERS IN THE OSI MODEL (Detailed OSI layers not in syllabus, it's for your reference)

Physical Layer

• Main Responsibility:
Physical-layer (PL) is responsible for movements of individual bits from one node to another node.

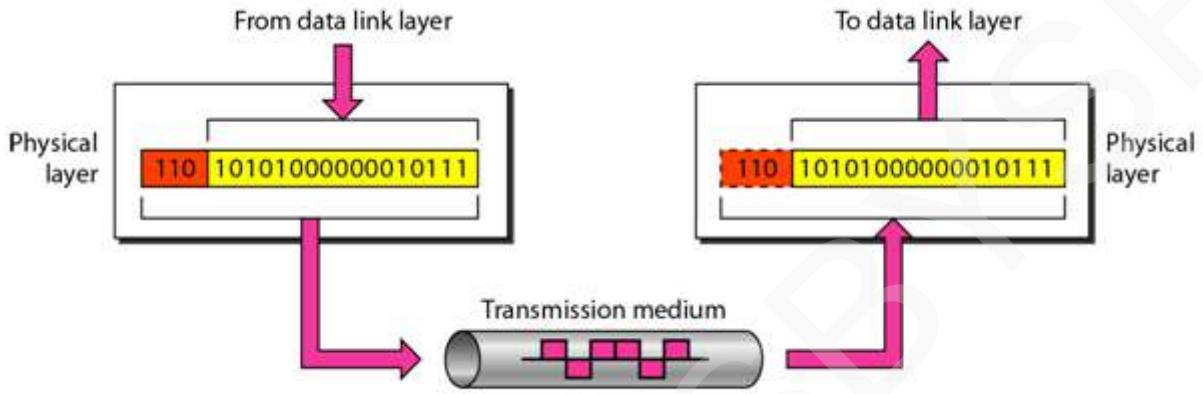


Figure 2.5 *Physical layer*

• Other responsibilities of Physical-layer (Figure 2.5):

1) Physical Characteristics of Interfaces and Medium

➢ PL defines the mechanical/electrical characteristics of the interface & transmission-medium
i.e. Mechanical → cable, plugs, pins
Electrical → modulation, signal strength, voltage levels
➢ PL also defines the type of transmission-medium. (Wired or wireless).

2) Representation of Bits

➢ PL defines the type of encoding i.e. how 0s and 1s are changed to signals.
➢ Data consists of a stream of bits: 0s or 1s.
➢ Bits must be encoded into signals for transmission.

3) Data Rate

➢ PL defines the transmission-rate.
➢ Transmission-rate refers to the number of bits sent per second.

4) Synchronization of Bits

➢ PL deals with the synchronization of the transmitter and receiver.
➢ The sender and receiver are synchronized at bit-level.

5) Line Configuration

➢ PL defines the nature of the connection.
i) In a point-to-point configuration, a dedicated-link is used to connect between 2 devices
ii) In a multipoint configuration, a shared-link is used to connect between 2 or more devices.

6) Physical Topology

➢ PL defines the type of topology used for connecting the devices in the network.
➢ Topologies can be mesh, star, ring or bus.

7) Transmission Mode

➢ PL defines the direction of data-transfer between 2 devices.
i) Simplex: Only one device can send; the other device can only receive.
ii) Half-duplex: Two devices can send and receive, but not at the same time.
iii) Full-duplex: Two devices can send and receive at the same time.

Data Link Layer

• Main Responsibility:

Data-link-layer (DLL) is responsible for moving frames from one node to another node.

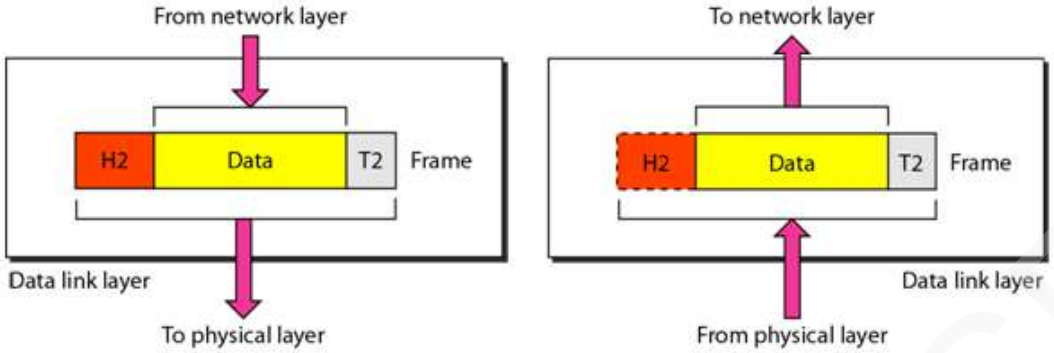


Figure 2.6 *Data link layer*

• Other responsibilities of data-link-layer (Figure 2.6 & 2.7):

1) Framing

➢ DLL receives & divides the stream of bits from network-layer into frames.

2) Physical-addressing

➢ DLL appends a header to the frame coming from the network-layer.

➢ Header contains the physical-address of sender & receiver of the frame.

3) Flow Control

➢ DLL provides flow-control.

➢ Flow-control ensures that source sends the data at a speed at which destination can receive it

➢ If there is an overflow at the receiver-side, the data will be lost.

4) Error Control

➢ DLL provides error-control.

➢ Error-control is process of identification or correction of error occurred in the transmitted data

➢ Error-control uses mechanisms to

→ detect damaged-frames

→ retransmit lost-frames

→ recognize duplicate frames.

➢ Normally, error control information is present in the trailer of a frame.

5) Access Control

➢ DLL provides access-control.

➢ Access-control determines which device has right to send the data in a multipoint connection.

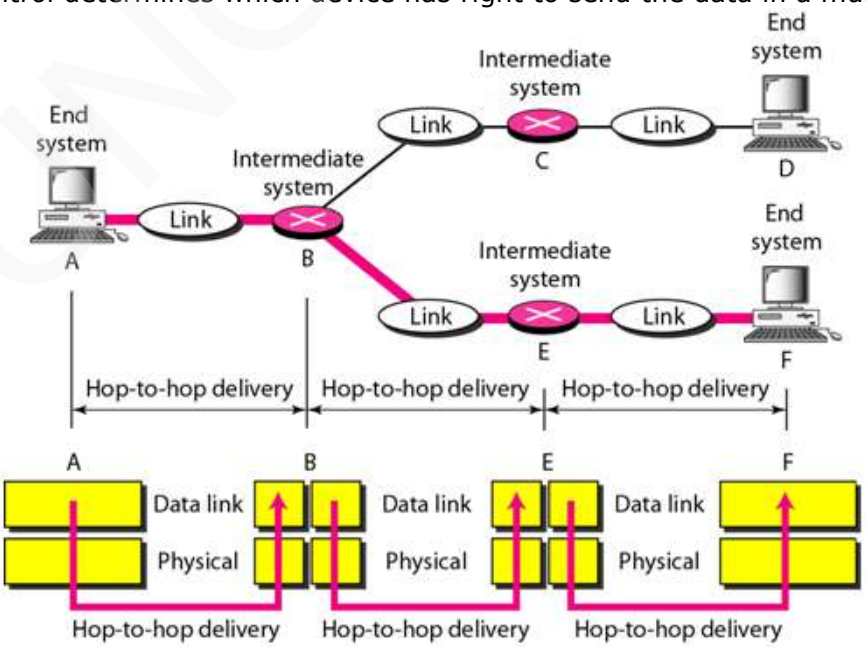


Figure 2.7 *Hop-to-hop delivery*

Network Layer

- Main Responsibility:
  Network-layer (NL) is responsible for source-to-destination delivery of a packet, possibly across multiple-networks.
- Data-link-layer vs. Network-layer:
  1) The data-link-layer ensures the delivery of the packet between 2 systems on the same link.
  2) The network-layer ensures that each packet gets from the source to the final destination.
- If 2 systems are connected to the same link, there is no need for a network-layer.
  However, if the 2 systems are attached to different links, there is often a need for the network-layer to accomplish source-to-destination delivery.

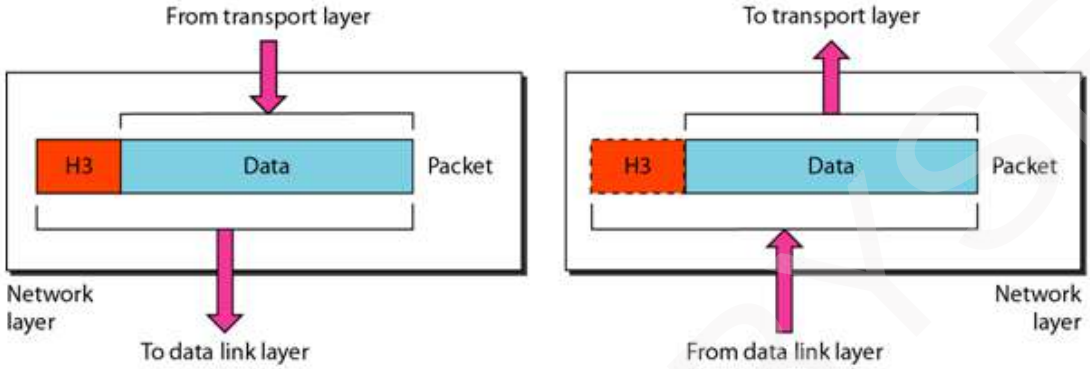


Figure 2.8 *Network layer*

- Other responsibilities of network-layer (Figure 2.8 & 2.9):

1) Logical Addressing

- NL appends a header to the packet coming from the transport-layer.
- The header contains the IP addresses of the sender and receiver.
- An IP address is a universally unique address in the network.
- NL uses IP address to recognize devices on the network.

2) Routing

- NL provides routing of packets.
- Routing is the process of finding the best path from a source to a destination.
- Routers/gateways are used for routing the packets to their final destination.
- NL is concerned with circuit, message or packet switching.

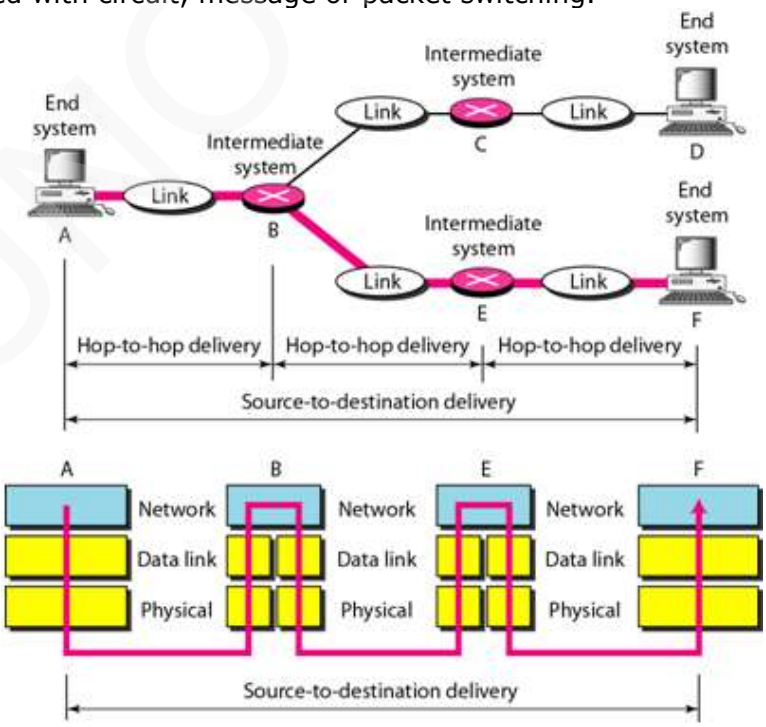


Figure 2.9 *Source-to-destination delivery*

Transport Layer

• Main Responsibility:

Transport-layer (TL) is responsible for process-to-process delivery of the entire message.

• Process-to-process delivery means delivery from a specific process on one computer to a specific process on the other computer.

• A process is an application program running on a host.

• Network-layer vs. Transport-layer:

1) Network-layer ensures source-to-destination delivery of individual packets.

2) Transport-layer ensures that the whole message arrives in order

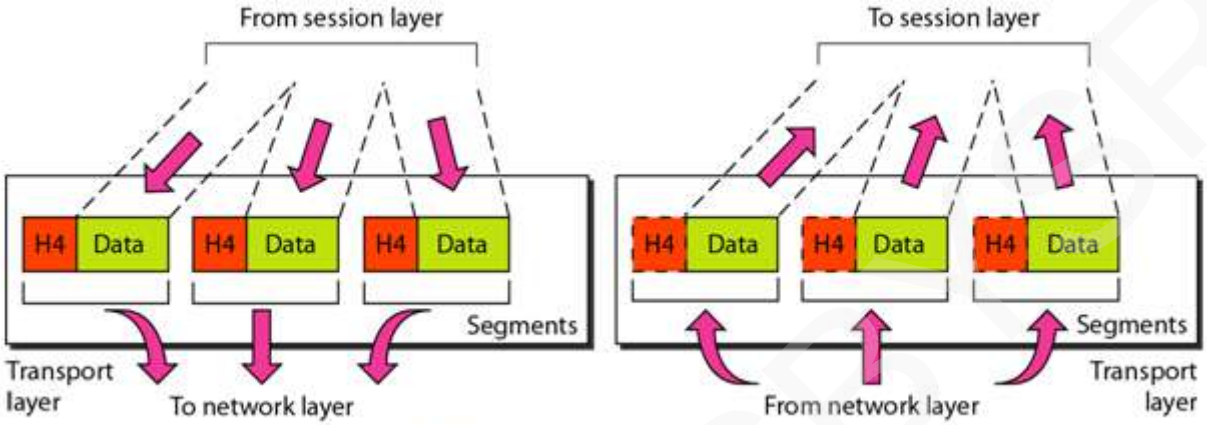


Figure 2.10 *Transport layer*

• Other responsibilities of transport-layer (Figure 2.10 & 2.11):

1) Service Point Addressing

- NL appends a header to the segments coming from the network-layer.
- Header contains the port-address of the sender and receiver.
- Network-layer vs. Transport-layer:

  i) The network-layer gets each packet to the correct computer.

  ii) The transport-layer gets the entire message to the correct process on that computer.

2) Segmentation & Reassembly

- A message is divided into segments.
- Each segment contains a sequence-number.
- At receiver, the sequence-numbers are used to

  → rearrange the segments in proper order

  → identify lost/duplicate segments

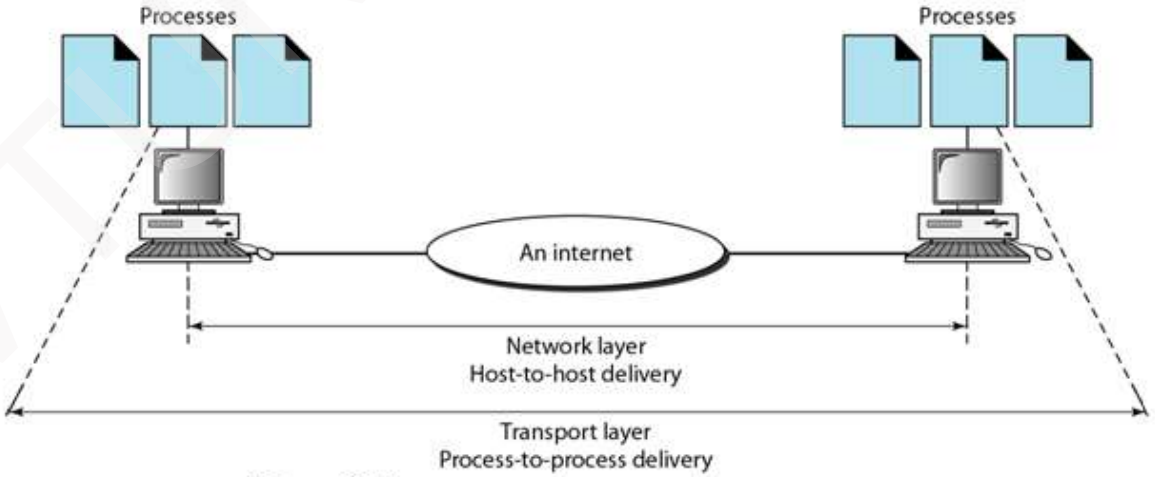


Figure 2.11 *Reliable process-to-process delivery of a message*

3) Connection Control

- TL can be either i) connectionless or ii) connection-oriented.
  - i) In connectionless, TL
    - → treats each segment as an independent packet and
    - → delivers the segment to the transport-layer at the destination-machine.
  - ii) In connection-oriented, TL
    - → first, makes a connection with the destination-machine.
    - → then, delivers the packets to the destination-machine.

4) Flow Control & Error Control

- Like DLL, TL is responsible for flow-control & error-control.
  However, flow-control & error-control are performed end-to-end rather than node-to-node.

Session Layer

- Main Responsibility:
  Session-layer (SL) establishes, maintains, and synchronizes the interaction between 2 systems.
- Other responsibilities of session-layer (Figure 2.12):

1) Dialog Control

- SL allows 2 systems to start communication with each other in half-duplex or full-duplex.

2) Synchronization

- SL allows a process to add checkpoints into stream of data.
- The checkpoint is a way of informing the status of the data transfer.
- For example:
  A checkpoint after first 500 bits of data will ensure that those 500 bits are not sent again in case of retransmission at 650th bit. (Checkpoints → Synchronization Points)

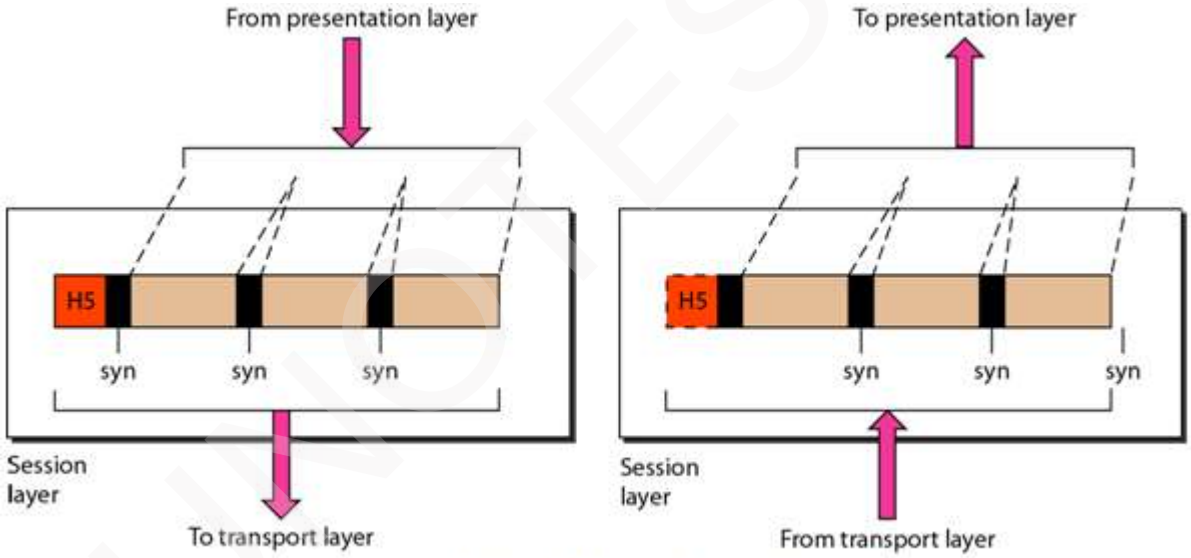


Figure 2.12 *Session layer*

Presentation Layer

• Main Responsibility:
Presentation-layer (PL) is concerned with syntax & semantics of the info. exchanged b/w 2 systems.

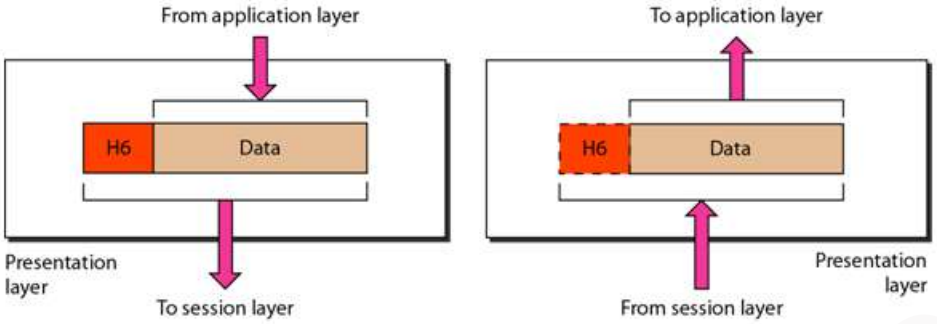


Figure 2.13 *Presentation layer*

• Other responsibilities of presentation-layer (Figure 2.13):

1) Translation

➢ PL translates data between
→ format the network requires and
→ format the computer understands.

➢ PL is responsible for interoperability between encoding methods as different computers use different encoding-methods.

2) Encryption

➢ PL performs
→ encryption at the sender and
→ decryption at the receiver.

➢ Encryption means the sender transforms the original information to another.

➢ Decryption means the receiver transforms the encrypted-message back to its original form.

3) Compression

➢ PL carries out data compression to reduce the size of the data to be transmitted.

➢ Data compression reduces the number of bits contained in the information.

➢ Data compression ensures faster data transfer.

➢ Data compression is important in transmitting multimedia such as audio, video, etc.

Application Layer

• Main Responsibility: The application-layer (AL)
→ provides services to the user
→ enables the user to access the network.

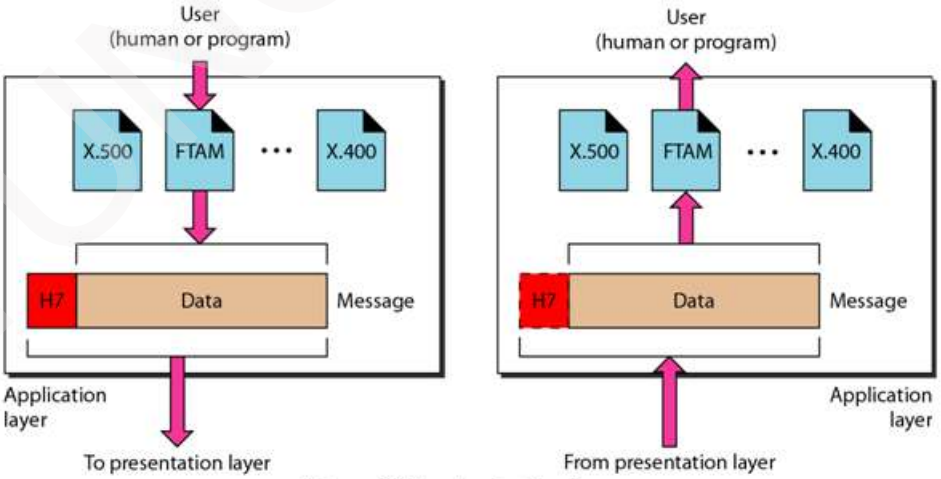


Figure 2.14 *Application layer*

• Other responsibilities of application-layer (Figure 2.14):
1) Mail Services
2) Directory Services
3) File Transfer, Access, and Management

MODULE 1(CONT.): DATA AND SIGNALS

1.8 DATA AND SIGNALS

1.8.1 Analog & Digital Data

- To be transmitted, data must be transformed to electromagnetic-signals.
- Data can be either analog or digital.

1) Analog Data refers to information that is continuous.

- For example:

The sounds made by a human voice.

2) Digital Data refers to information that has discrete states.

- For example:

Data are stored in computer-memory in the form of 0s and 1s.

1.8.2 Analog & Digital Signals

- Signals can be either analog or digital (Figure 3.2).

1) Analog Signal has infinitely many levels of intensity over a period of time.

2) Digital Signal can have only a limited number of defined values.

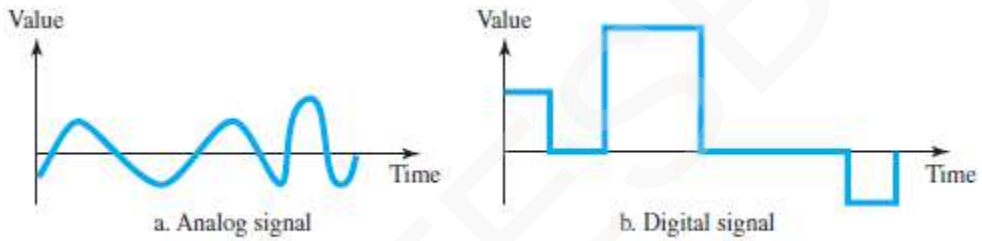


Figure 3.2 *Comparison of analog and digital signals*

1.8.3 Periodic & Non-Periodic Signals

- The signals can take one of 2 forms: periodic or non-periodic.

1) Periodic Signal

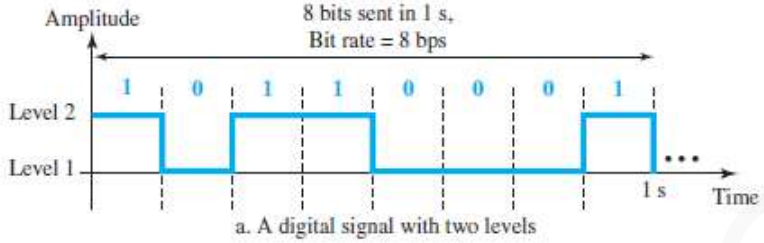
- Signals which repeat itself after a fixed time period are called Periodic Signals.
- The completion of one full pattern is called a cycle.

2) Non-Periodic Signal

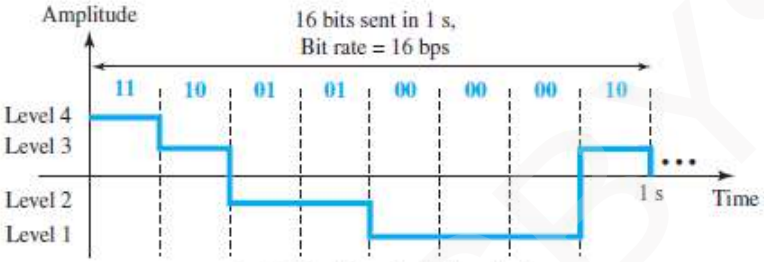
- Signals which do not repeat itself after a fixed time period are called Non-Periodic Signals.

1.9 DIGITAL SIGNALS

- Information can be represented by a digital signal.
- For example:
  1) 1 can be encoded as a positive voltage.
     0 can be encoded as a zero voltage (Figure 3.17a).
  2) A digital signal can have more than 2 levels (Figure 3.17b).



a. A digital signal with two levels



b. A digital signal with four levels

Figure 3.17 *Two digital signals: one with two signal levels and the other with four signal levels*

Example 1.1

A digital signal has eight levels. How many bits are needed per level? We calculate the number of bits from the following formula. Each signal level is represented by 3 bits.

$$\text{Number of bits per level} = \log_2 8 = 3$$

c

1.9.1 Bit Rate

- The bit rate is the number of bits sent in 1s.
- The bit rate is expressed in bits per second (bps).

Example 1.2

Assume we need to download text documents at the rate of 100 pages per second. What is the required bit rate of the channel?

Solution

A page is an average of 24 lines with 80 characters in each line. If we assume that one character requires 8 bits, the bit rate is

$$100 \times 24 \times 80 \times 8 = 1{,}536{,}000 \text{ bps} = 1.536 \text{ Mbps}$$

Example 1.3

A digitized voice channel, as we will see in Chapter 4, is made by digitizing a 4-kHz bandwidth analog voice signal. We need to sample the signal at twice the highest frequency (two samples per hertz). We assume that each sample requires 8 bits. What is the required bit rate?

Solution

The bit rate can be calculated as

$$2 \times 4000 \times 8 = 64{,}000 \text{ bps} = 64 \text{ kbps}$$

Example 1.4

What is the bit rate for high-definition TV (HDTV)?

Solution

HDTV uses digital signals to broadcast high quality video signals. The HDTV screen is normally a ratio of 16 : 9 (in contrast to 4 : 3 for regular TV), which means the screen is wider. There are 1920 by 1080 pixels per screen, and the screen is renewed 30 times per second. Twenty-four bits represents one color pixel. We can calculate the bit rate as

$$1920 \times 1080 \times 30 \times 24 = 1{,}492{,}992{,}000 \approx 1.5 \text{ Gbps}$$

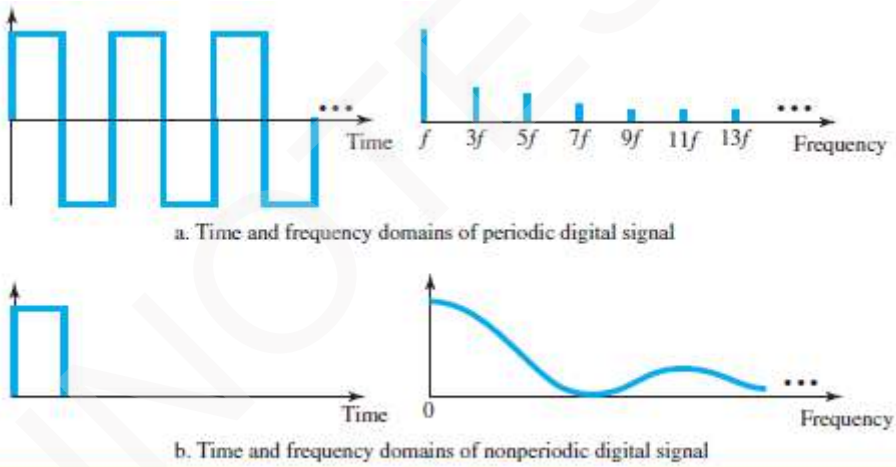
1.9.2 Bit Length

- The bit length is the distance one bit occupies on the transmission medium.

Bit length = propagation speed × bit duration

1.9.3 Digital Signal as a Composite Analog Signal

- A digital signal is a composite analog signal.
- A digital signal, in the time domain, comprises connected vertical and horizontal line segments.
  1) A vertical line in the time domain means a frequency of infinity (sudden change in time);
  2) A horizontal line in the time domain means a frequency of zero (no change in time).
- Fourier analysis can be used to decompose a digital signal.
  1) If the digital signal is periodic, the decomposed signal has a frequency domain representation with an infinite bandwidth and discrete frequencies (Figure 3.18a).
  2) If the digital signal is non-periodic, the decomposed signal has a frequency domain representation with an infinite bandwidth and continuous frequencies (Figure 3.18b).



a. Time and frequency domains of periodic digital signal

b. Time and frequency domains of nonperiodic digital signal

Figure 3.18 *The time and frequency domains of periodic and nonperiodic digital signals*

1.9.4 Transmission of Digital Signals

• Two methods for transmitting a digital signal:
1) Baseband transmission 2) Broadband transmission (using modulation).

1.9.4.1 Baseband Transmission

• Baseband transmission means sending a digital signal over a channel without changing the digital signal to an analog signal (Figure 3.19).

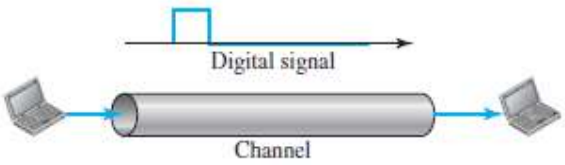


Figure 3.19 *Baseband transmission*

• Baseband transmission requires that we have a low-pass channel.
• Low-pass channel means a channel with a bandwidth that starts from zero.
• For example, we can have a dedicated medium with a bandwidth constituting only one channel.

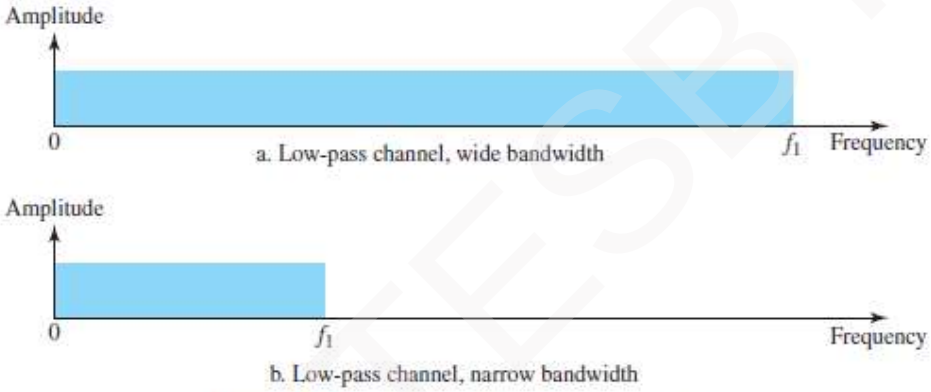


Figure 3.20 *Bandwidths of two low-pass channels*

• Two cases of a baseband communication:
Case 1: Low-pass channel with a wide bandwidth (Figure 3.20a)
Case 2: Low-pass channel with a limited bandwidth (Figure 3.20b)

Case 1: Low-Pass Channel with Wide Bandwidth

- To preserve the shape of a digital signal, we need to send the entire spectrum i.e. the continuous range of frequencies between zero and infinity.
- This is possible if we have a dedicated medium with an infinite bandwidth between the sender and receiver.
- If we have a medium with a very wide bandwidth, 2 stations can communicate by using digital signals with very good accuracy (Figure 3.21).
- Although the output signal is not an exact replica of the original signal, the data can still be deduced from the received signal.

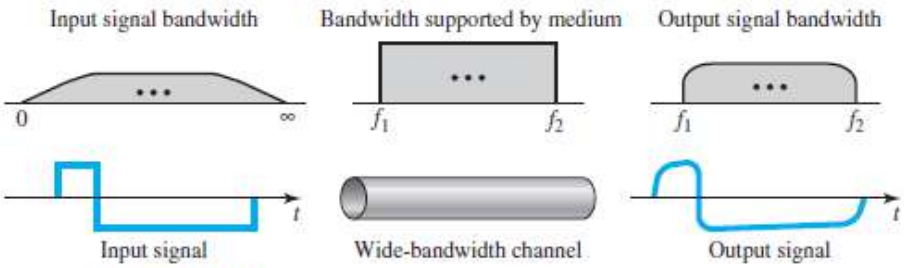


Figure 3.21 *Baseband transmission using a dedicated medium*

Case 2: Low-Pass Channel with Limited Bandwidth

- In a low-pass channel with limited bandwidth, we approximate the digital signal with an analog signal.
- The level of approximation depends on the bandwidth available.

A) Rough Approximation

- Assume that we have a digital signal of bit rate N (Figure 3.22).
- If we want to send analog signals to roughly simulate this signal, we need to consider the worst case, a maximum number of changes in the digital signal.
- This happens when the signal carries the sequence 01010101 . . or 10101010. . . .
- To simulate these two cases, we need an analog signal of frequency f = N/2.
- Let 1 be the positive peak value and 0 be the negative peak value.
- We send 2 bits in each cycle; the frequency of the analog signal is one-half of the bit rate, or N/2.
- This rough approximation is referred to as using the first harmonic (N/2) frequency. The required bandwidth is

$$\text{Bandwidth} = \frac{N}{2} - 0 = \frac{N}{2}$$

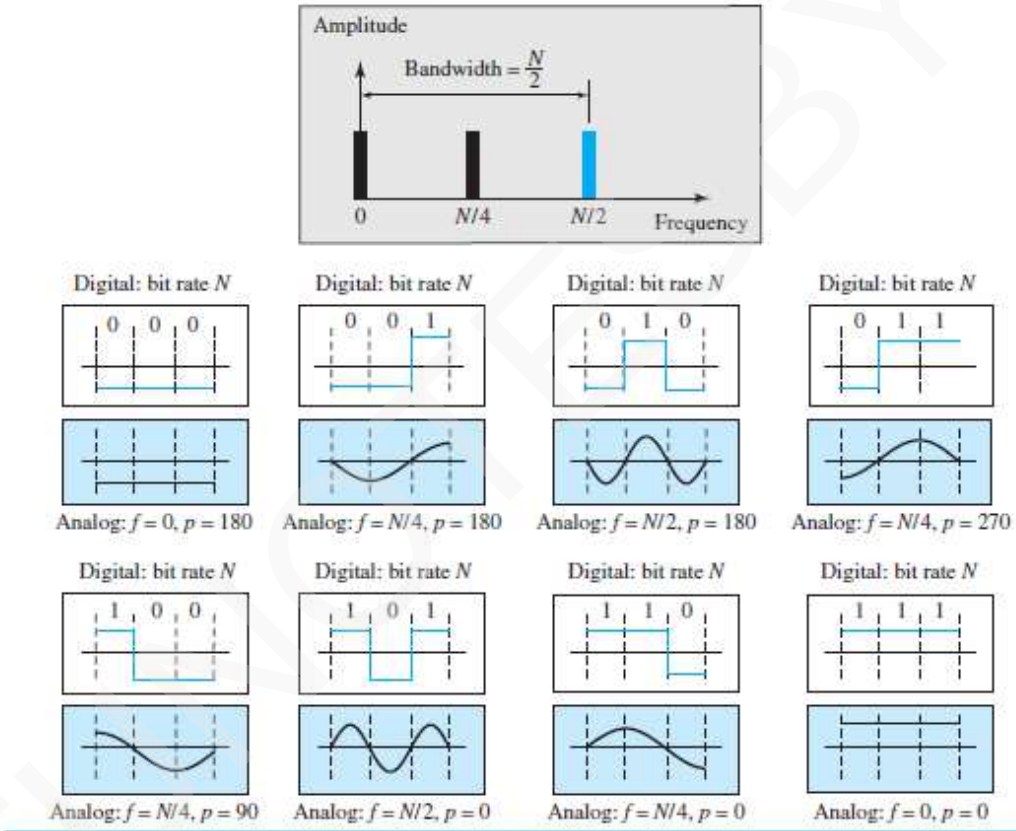


Figure 3.22 *Rough approximation of a digital signal using the first harmonic for worst case*

B) Better Approximation

- To make the shape of the analog signal look more like that of a digital signal, we need to add more harmonics of the frequencies (Figure 3.23).
- We can increase the bandwidth to 3N/2, 5N/2, 7N/2, and so on.
- In baseband transmission, the required bandwidth is proportional to the bit rate; If we need to send bits faster, we need more bandwidth.

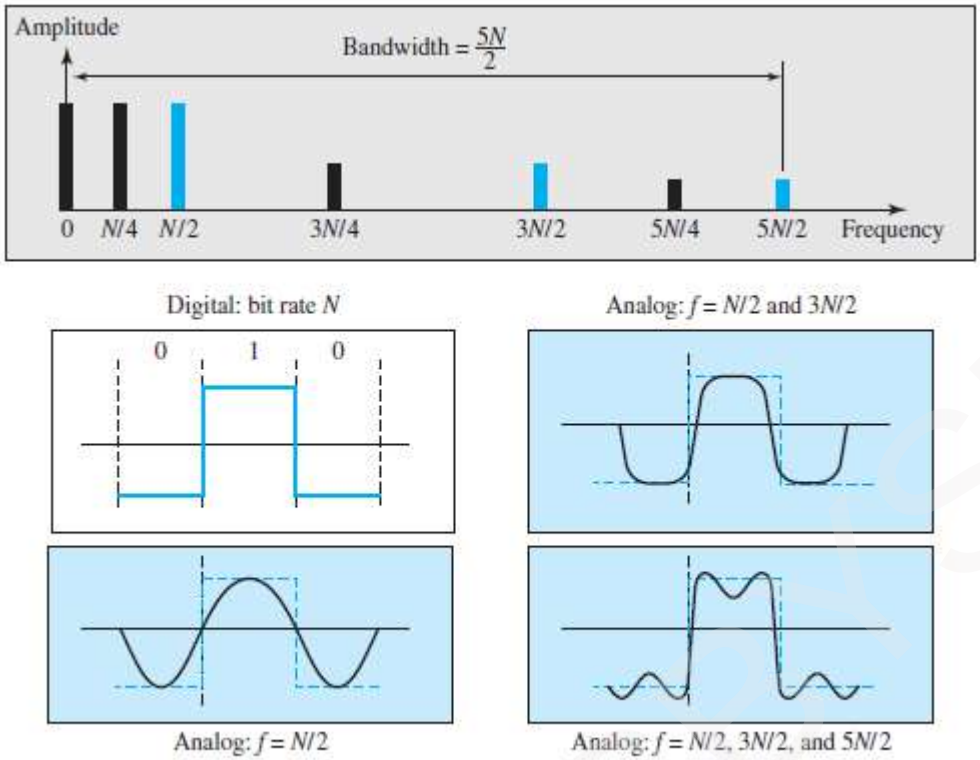


Figure 3.23 *Simulating a digital signal with first three harmonics*

Table 3.2 *Bandwidth requirements*

Bit Rate	*Harmonic 1*	*Harmonics 1, 3*	*Harmonics 1, 3, 5*
n = 1 kbps	B = 500 Hz	B = 1.5 kHz	B = 2.5 kHz
n = 10 kbps	B = 5 kHz	B = 15 kHz	B = 25 kHz
n = 100 kbps	B = 50 kHz	B = 150 kHz	B = 250 kHz

Example 1.5

What is the required bandwidth of a low-pass channel if we need to send 1 Mbps by using base-band transmission?

Solution

The answer depends on the accuracy desired.

a. The minimum bandwidth, a rough approximation, is B = bit rate /2, or 500 kHz. We need a low-pass channel with frequencies between 0 and 500 kHz.
b. A better result can be achieved by using the first and the third harmonics with the required bandwidth $B = 3 \times 500$ kHz = 1.5 MHz.
c. A still better result can be achieved by using the first, third, and fifth harmonics with $B = 5 \times 500$ kHz = 2.5 MHz.

Example 1.6

We have a low-pass channel with bandwidth 100 kHz. What is the maximum bit rate of this channel?

Solution

The maximum bit rate can be achieved if we use the first harmonic. The bit rate is 2 times the available bandwidth, or 200 kbps.

1.9.4.2 Broadband Transmission (Using Modulation)

• Broadband transmission or modulation means changing the digital signal to an analog signal for transmission.
• Modulation allows us to use a bandpass channel (Figure 3.24).
• Bandpass channel means a channel with a bandwidth that does not start from zero.
• This type of channel is more available than a low-pass channel.

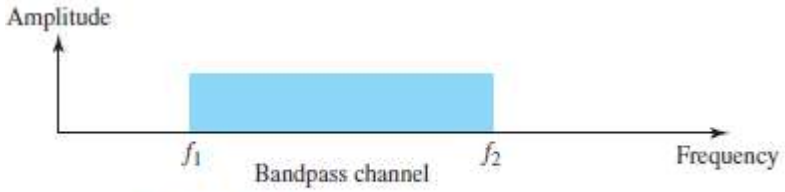


Figure 3.24 *Bandwidth of a bandpass channel*

• If the available channel is a bandpass channel,
We cannot send the digital signal directly to the channel;
We need to convert the digital signal to an analog signal before transmission (Figure 3.25).

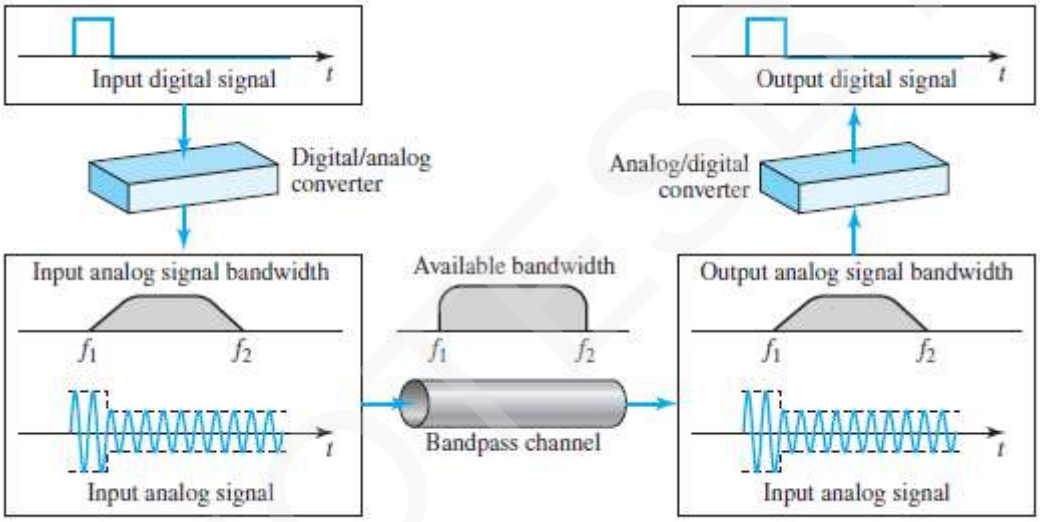


Figure 3.25 *Modulation of a digital signal for transmission on a bandpass channel*

1.10 TRANSMISSION IMPAIRMENT

- Signals travel through transmission media, which are not perfect.
- The imperfection causes signal-impairment.
- This means that signal at beginning of the medium is not the same as the signal at end of medium.
- What is sent is not what is received.
- Three causes of impairment are (Figure 3.26):
  1) Attenuation
  2) Distortion &
  3) Noise.

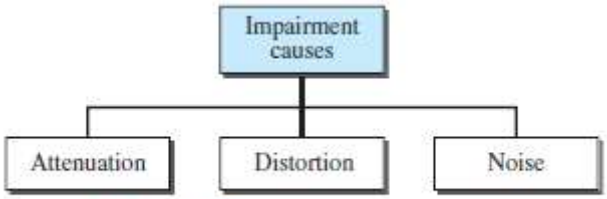


Figure 3.26 *Causes of impairment*

1.10.1 Attenuation

- As signal travels through the medium, its strength decreases as distance increases. This is called attenuation (Figure 3.27).
- As the distance increases, attenuation also increases.
- For example:
  Voice-data becomes weak over the distance & loses its contents beyond a certain distance.
- To compensate for this loss, amplifiers are used to amplify the signal.

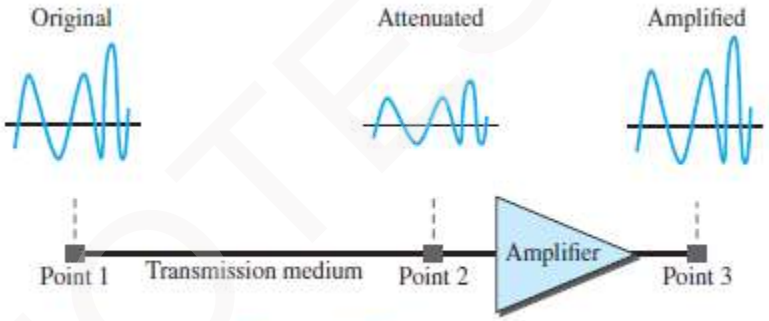


Figure 3.27 *Attenuation*

1.10.1.1 Decibel

- The decibel (dB) measures the relative strengths of
  → 2 signals or
  → one signal at 2 different points.
- The decibel is negative if a signal is attenuated.
  The decibel is positive if a signal is amplified.

$$dB = 10 \log_{10} \frac{P_2}{P_1}$$

- Variables P_1 and P_2 are the powers of a signal at points 1 and 2, respectively.
- To show that a signal has lost or gained strength, engineers use the unit of decibel.

Example 1.7

Suppose a signal travels through a transmission medium and its power is reduced to one-half. This means that $P_2 = \frac{1}{2} P_1$. In this case, the attenuation (loss of power) can be calculated as

$$10 \log_{10} \frac{P_2}{P_1} = 10 \log_{10} \frac{0.5P_1}{P_1} = 10 \log_{10} 0.5 = 10(-0.3) = -3 \text{ dB}$$

Example 1.8

A signal travels through an amplifier, and its power is increased 10 times. This means that $P_2 = 10P_1$. In this case, the amplification (gain of power) can be calculated as

$$10 \log_{10} \frac{P_2}{P_1} = 10 \log_{10} \frac{10P_1}{P_1} = 10 \log_{10} 10 = 10(1) = 10 \text{ dB}$$

Example 1.9

Sometimes the decibel is used to measure signal power in milliwatts. In this case, it is referred to as dB_m and is calculated as $dB_m = 10 \log_{10} P_m$, where P_m is the power in milliwatts. Calculate the power of a signal if its $dB_m = -30$.

Solution

We can calculate the power in the signal as

$$dB_m = 10 \log_{10} \longrightarrow dB_m = -30 \longrightarrow \log_{10} P_m = -3 \longrightarrow P_m = 10^{-3} \text{ mW}$$

Example 1.10

The loss in a cable is usually defined in decibels per kilometer (dB/km). If the signal at the beginning of a cable with –0.3 dB/km has a power of 2 mW, what is the power of the signal at 5 km?

Solution

The loss in the cable in decibels is $5 \times (-0.3) = -1.5$ dB. We can calculate the power as

$$dB = 10 \log_{10} (P_2 / P_1) = -1.5 \longrightarrow (P_2 / P_1) = 10^{-0.15} = 0.71$$

$$P_2 = 0.71P_1 = 0.7 \times 2 \text{ mW} = 1.4 \text{ mW}$$

1.10.2 Distortion

- Distortion means that the signal changes its form or shape (Figure 3.29).
- Distortion can occur in a composite signal made of different frequencies.
- Different signal-components
  - → have different propagation speed through a medium.
  - → have different delays in arriving at the final destination.
- Differences in delay create a difference in phase if delay is not same as the period-duration.
- Signal-components at the receiver have phases different from what they had at the sender.
- The shape of the composite signal is therefore not the same.

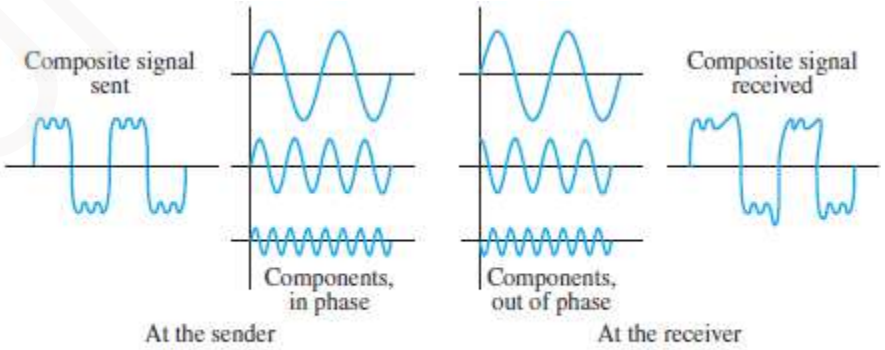


Figure 3.29 *Distortion*

1.10.3 Noise

- Noise is defined as an unwanted data (Figure 3.30).
- In other words, noise is the external energy that corrupts a signal.
- Due to noise, it is difficult to retrieve the original data/information.
- Four types of noise:

i) Thermal Noise

➢ It is random motion of electrons in wire which creates extra signal not originally sent by transmitter.

ii) Induced Noise

➢ Induced noise comes from sources such as motors & appliances.
➢ These devices act as a sending-antenna.
The transmission-medium acts as the receiving-antenna.

iii) Crosstalk

➢ Crosstalk is the effect of one wire on the other.
➢ One wire acts as a sending-antenna and the other as the receiving-antenna.

iv) Impulse Noise

➢ Impulse Noise is a spike that comes from power-lines, lightning, and so on.
(spike → a signal with high energy in a very short time)

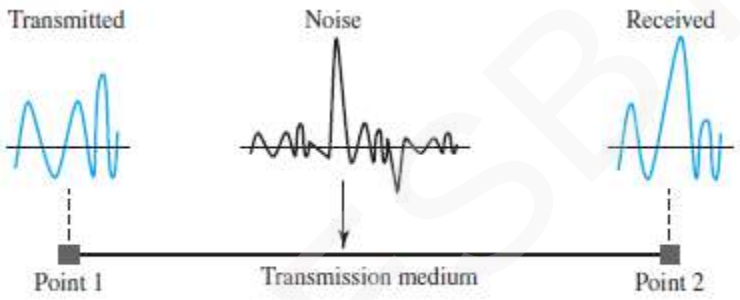


Figure 3.30 *Noise*

1.10.3.1 Signal-to-Noise Ratio (SNR)

- SNR is used to find the theoretical bit-rate limit.
- SNR is defined as

$$SNR = \frac{\text{average signal power}}{\text{average noise power}}$$

- SNR is actually the ratio of what is wanted (signal) to what is not wanted (noise).
- A high-SNR means the signal is less corrupted by noise.
A low-SNR means the signal is more corrupted by noise.
- Because SNR is the ratio of 2 powers, it is often described in decibel units, SNR_{dB}, defined as

$$SNR_{dB} = 10\log_{10} SNR$$

Example 1.11

The power of a signal is 10 mW and the power of the noise is 1 μW; what are the values of SNR and SNR_{dB}?

Solution

The values of SNR and SNR_{dB} can be calculated as follows:

$$SNR = (10{,}000\ \mu w) / (1\ \mu w) = 10{,}000 \qquad SNR_{dB} = 10 \log_{10} 10{,}000 = 10 \log_{10} 10^4 = 40$$

1.11 DATA RATE LIMITS

- Data-rate depends on 3 factors:
  1) Bandwidth available
  2) Level of the signals
  3) Quality of channel (the level of noise)
- Two theoretical formulas can be used to calculate the data-rate:
  1) Nyquist for a noiseless channel and
  2) Shannon for a noisy channel.

1.11.1 Noiseless Channel: Nyquist Bit Rate

- For a noiseless channel, the Nyquist bit-rate formula defines the theoretical maximum bit-rate

$$\text{Bitrate} = 2 \times \text{Bandwidth} \times \log_2 L$$

where bandwidth = bandwidth of the channel
L = number of signal-levels used to represent data
BitRate = bitrate of channel in bps

- According to the formula,
  - By increasing number of signal-levels, we can increase the bit-rate.
  - Although the idea is theoretically correct, practically there is a limit.
  - When we increase the number of signal-levels, we impose a burden on the receiver.
  - If no. of levels in a signal is 2, the receiver can easily distinguish b/w 0 and 1.
  - If no. of levels is 64, the receiver must be very sophisticated to distinguish b/w 64 different levels.
  - In other words, increasing the levels of a signal reduces the reliability of the system.

Example 1.12

Consider a noiseless channel with a bandwidth of 3000 Hz transmitting a signal with two signal levels. The maximum bit rate can be calculated as

$$\text{BitRate} = 2 \times 3000 \times \log_2 2 = 6000 \text{ bps}$$

Example 1.13

Consider the same noiseless channel transmitting a signal with four signal levels (for each level, we send 2 bits). The maximum bit rate can be calculated as

$$\text{BitRate} = 2 \times 3000 \times \log_2 4 = 12{,}000 \text{ bps}$$

Example 1.14

We need to send 265 kbps over a noiseless channel with a bandwidth of 20 kHz. How many signal levels do we need?

Solution

We can use the Nyquist formula as shown:

$$265{,}000 = 2 \times 20{,}000 \times \log_2 L \longrightarrow \log_2 L = 6.625 \longrightarrow L = 2^{6.625} = 98.7 \text{ levels}$$

1.11.2 Noisy Channel: Shannon Capacity

- In reality, we cannot have a noiseless channel; the channel is always noisy.
- For a noisy channel, the Shannon capacity formula defines the theoretical maximum bit-rate.

$$\text{Capacity} = \text{bandwidth} \times \log_2(1 + \text{SNR})$$

where bandwidth = bandwidth of channel in bps.
SNR = signal-to-noise ratio and
Capacity = capacity of channel in bps.

- This formula does not consider the no. of levels of signals being transmitted (as done in the Nyquist bit rate).

  This means that no matter how many levels we have, we cannot achieve a data-rate higher than the capacity of the channel.
- In other words, the formula defines a characteristic of the channel, not the method of transmission.

Example 1.15

We can calculate the theoretical highest bit rate of a regular telephone line. A telephone line normally has a bandwidth of 3000 Hz (300 to 3300 Hz) assigned for data communications. The signal-to-noise ratio is usually 3162. For this channel the capacity is calculated as

$$C = B \log_2(1 + \text{SNR}) = 3000 \log_2(1 + 3162) = 3000 \times 11.62 = 34{,}860 \text{ bps}$$

Example 1.16

The signal-to-noise ratio is often given in decibels. Assume that $\text{SNR}_{\text{dB}} = 36$ and the channel bandwidth is 2 MHz. The theoretical channel capacity can be calculated as

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \text{SNR} \longrightarrow \text{SNR} = 10^{\text{SNR}_{\text{dB}}/10} \longrightarrow \text{SNR} = 10^{3.6} = 3981$$

$$C = B \log_2(1 + \text{SNR}) = 2 \times 10^6 \times \log_2 3982 = 24 \text{ Mbps}$$

Example 1.17

We have a channel with a 1-MHz bandwidth. The SNR for this channel is 63. What are the appropriate bit rate and signal level?

Solution

First, we use the Shannon formula to find the upper limit.

$$C = B \log_2(1 + \text{SNR}) = 10^6 \log_2(1 + 63) = 10^6 \log_2 64 = 6 \text{ Mbps}$$

The Shannon formula gives us 6 Mbps, the upper limit. For better performance we choose something lower, 4 Mbps, for example. Then we use the Nyquist formula to find the number of signal levels.

$$4 \text{ Mbps} = 2 \times 1 \text{ MHz} \times \log_2 L \longrightarrow L = 4$$

1.12 PERFORMANCE

1.12.1 Bandwidth

- One characteristic that measures network-performance is bandwidth.
- Bandwidth of analog and digital signals is calculated in separate ways:

(1) Bandwidth of an Analog Signal (in hz)

- Bandwidth of an analog signal is expressed in terms of its frequencies.
- Bandwidth is defined as the range of frequencies that the channel can carry.
- It is calculated by the difference b/w the maximum frequency and the minimum frequency.

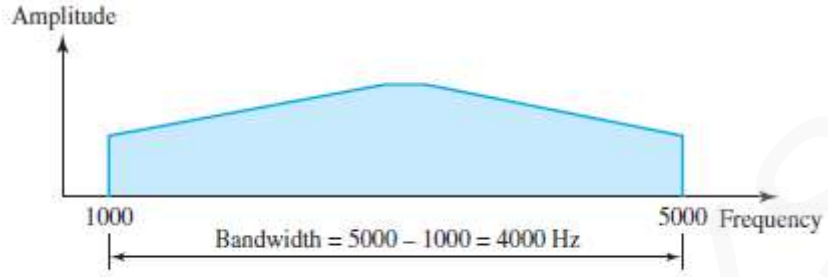


Figure 3.13 *The bandwidth of signals*

In figure 3.13, the signal has a minimum frequency of F1 = 1000Hz and maximum frequency of F2 = 5000Hz.

Hence, the bandwidth is given by F2 – F1= 5000 - 1000 = 4000 Hz

(2) Bandwidth of a Digital Signal (in bps)

- Bandwidth refers to the number of bits transmitted in one second in a channel (or link).
- For example:

The bandwidth of a Fast Ethernet is a maximum of 100 Mbps. (This means that this network can send 100 Mbps).

Relationship between (1) and (2)

- There is an explicit relationship between the bandwidth in hertz and bandwidth in bits per seconds.
- Basically, an increase in bandwidth in hertz means an increase in bandwidth in bits per second.
- The relationship depends on
  - → baseband transmission or
  - → transmission with modulation.

1.12.2 Throughput

- The throughput is a measure of how fast we can actually send data through a network.
- Although, bandwidth in bits per second and throughput seem the same, they are actually different.
- A link may have a bandwidth of B bps, but we can only send T bps through this link with T always less than B.
- In other words,
    1) The bandwidth is a potential measurement of a link.
    2) The throughput is an actual measurement of how fast we can send data.
        For example:
        - ¤ We may have a link with a bandwidth of 1 Mbps, but the devices connected to the end of the link may handle only 200 kbps.
        - ¤ This means that we cannot send more than 200 kbps through this link.

Example 1.18

A network with bandwidth of 10 Mbps can pass only an average of 12,000 frames per minute with each frame carrying an average of 10,000 bits. What is the throughput of this network?

Solution

We can calculate the throughput as

$$\text{Throughput} = (12{,}000 \times 10{,}000) / 60 = 2 \text{ Mbps}$$

1.12.3 Latency (Delay)

• The latency defines how long it takes for an entire message to completely arrive at the destination from the time the first bit is sent out from the source.

Latency = propagation time + transmission time + queuing time + processing delay

1) Propagation Time

- Propagation time is defined as the time required for a bit to travel from source to destination.
- Propagation time is given by

Propagation time = Distance / (Propagation Speed)

- Propagation speed of electromagnetic signals depends on
  - → medium and
  - → frequency of the signal.

Example 1.19

What is the propagation time if the distance between the two points is 12,000 km? Assume the propagation speed to be 2.4×10^8 m/s in cable.

Solution

We can calculate the propagation time as

$$\text{Propagation time} = (12{,}000 \times 10{,}000) / (2.4 \times 2^8) = 50 \text{ ms}$$

b

2) Transmission Time

- The time required for transmission of a message depends on
  - → size of the message and
  - → bandwidth of the channel.
- The transmission time is given by

Transmission time = (Message size) / Bandwidth

Example 1.20

What are the propagation time and the transmission time for a 2.5-KB (kilobyte) message (an e-mail) if the bandwidth of the network is 1 Gbps? Assume that the distance between the sender and the receiver is 12,000 km and that light travels at 2.4×10^8 m/s.

Solution

We can calculate the propagation and transmission time as

$$\text{Propagation time} = (12{,}000 \times 1000) / (2.4 \times 10^8) = 50 \text{ ms}$$

$$\text{Transmission time} = (2500 \times 8) / 10^9 = 0.020 \text{ ms}$$

c

Example 1.21

What are the propagation time and the transmission time for a 5-MB (megabyte) message (an image) if the bandwidth of the network is 1 Mbps? Assume that the distance between the sender and the receiver is 12,000 km and that light travels at 2.4×10^8 m/s.

Solution

We can calculate the propagation and transmission times as

$$\text{Propagation time} = (12{,}000 \times 1000) / (2.4 \times 10^8) = 50 \text{ ms}$$

$$\text{Transmission time} = (5{,}000{,}000 \times 8) / 10^6 = 40 \text{ s}$$

3) Queuing Time

- Queuing-time is the time needed for each intermediate-device to hold the message before it can be processed.
  (Intermediate device may be a router or a switch)
- The queuing-time is not a fixed factor. This is because
  i) Queuing-time changes with the load imposed on the network.
  ii) When there is heavy traffic on the network, the queuing-time increases.
- An intermediate-device
  → queues the arrived messages and
  → processes the messages one by one.
- If there are many messages, each message will have to wait.

4) Processing Delay

- Processing delay is the time taken by the routers to process the packet header.

1.12.4 Bandwidth Delay Product

- Two performance-metrics of a link are 1) Bandwidth and 2) Delay
- The bandwidth-delay product is very important in data-communications.
- Let us elaborate on this issue, using 2 hypothetical cases as examples.

Case 1: The following figure shows case 1 (Figure 3.32).

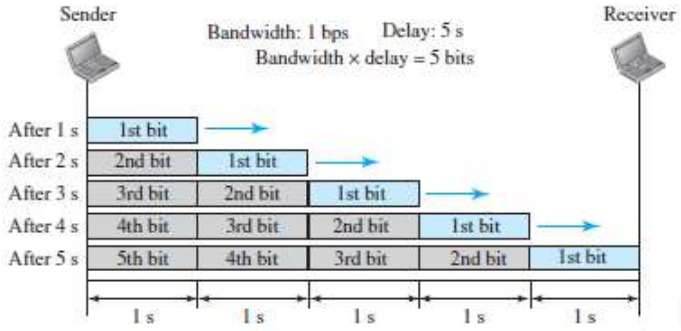


Figure 3.32 *Filling the link with bits for case 1*

- Let us assume,
  Bandwidth of the link = 1 bps Delay of the link = 5s.
- From the figure 3.32, bandwidth-delay product is 1 x 5 = 5. Thus, there can be maximum 5 bits on the line.
- There can be no more than 5 bits at any time on the link.

Case 2: The following figure shows case 2 (Figure 3.33).

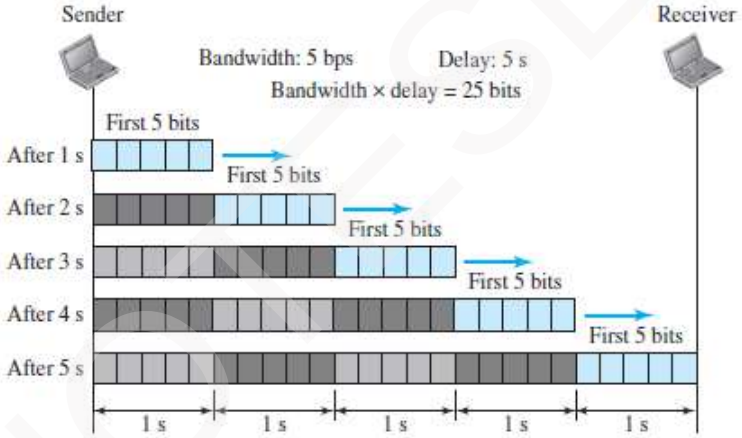


Figure 3.33 *Filling the link with bits in case 2*

- Let us assume,
  Bandwidth of the link = 4 bps Delay of the link = 5s.
- From the figure 3.33, bandwidth-delay product is 5 x 5 = 25. Thus, there can be maximum 25 bits on the line.
- At each second, there are 5 bits on the line, thus the duration of each bit is 0.20s.

- The above 2 cases show that the (bandwidth X delay) is the number of bits that can fill the link.
- This measurement is important if we need to
  → send data in bursts and
  → wait for the acknowledgment of each burst.
- To use the maximum capability of the link
  → We need to make the burst-size as (2 x bandwidth x delay).
  → We need to fill up the full-duplex channel (two directions).
- Amount (2x bandwidth x delay) is the number of bits that can be in transition at any time (Fig 3.34).

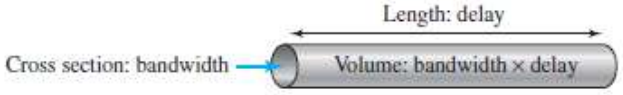


Figure 3.34 *Concept of bandwidth-delay product*

1.12.5 Jitter

- Another performance issue that is related to delay is jitter.
- We can say that jitter is a problem
  → if different packets of data encounter different delays and
  → if the application using the data at the receiver site is time-sensitive (for ex: audio/video).
- For example:
  If the delay for the first packet is 20 ms
  the delay for the second is 45 ms and
  the delay for the third is 40 ms
  then the real-time application that uses the packets suffers from jitter.

MODULE 1(CONT.): DIGITAL TRANSMISSION

1.13 DIGITAL TO DIGITAL CONVERSION

- Data can be analog or digital, so can be the signal that represents it.
- Signal encoding is the conversion from analog/digital data to analog/digital signal.
- The possible encodings are:
    1) Digital data to digital signal
    2) Digital data to analog signal
    3) Analog data to digital signal
    4) Analog data to analog signal

1.13.1 LINE CODING

- Line-coding is the process of converting digital-data to digital-signals (Figure 4.1).
- The data may be in the form of text, numbers, graphical images, audio, or video
- The data are stored in computer memory as sequences of bits (0s or 1s).
- Line-coding converts a sequence of bits to a digital-signal.
- At the sender, digital-data is encoded into a digital-signal.
  At the receiver, digital-signal is decoded into a digital-data.

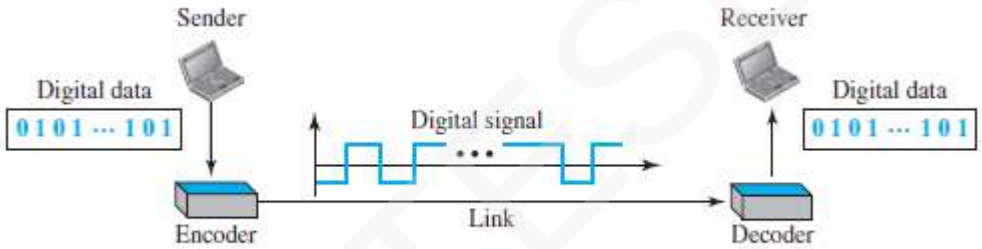


Figure 4.1 *Line coding and decoding*

1.13.1.1 Characteristics

- Different characteristics of digital signal are
  1) Signal Element Vs Data Element
  2) Data Rate Vs Signal Rate
  3) Bandwidth
  4) Baseline Wandering
  5) DC Components
  6) Built-in Error Detection
  7) Self-synchronization
  8) Immunity to Noise and Interference
  9) Complexity

1) Data Element vs. Signal Element

Data Element	Signal Element
A data-element is the smallest entity that can represent a piece of information (Figure 4.2).	A signal-element is shortest unit (timewise) of a digital-signal.
A data-element is the bit.	A signal-element carries data-elements.
Data-elements are being carried.	Signal-elements are the carriers.

➢ Ratio r is defined as number of data-elements carried by each signal-element.

Figure 4.2 *Signal element versus data element*

2) Data Rate vs. Signal Rate

Data Rate	Signal Rate
The data-rate defines the number of data-elements (bits) sent in 1 sec.	The signal-rate is the number of signal-elements sent in 1 sec.
The unit is bits per second (bps).	The unit is the baud.
The data-rate is sometimes called the bit-rate.	The signal-rate is sometimes called the pulse rate, the modulation rate, or the baud rate
Goal in data-communications: increase the data-rate.	Goal in data-communications: decrease the signal-rate.
Increasing the data-rate increases the speed of transmission.	Decreasing the signal-rate decreases the bandwidth requirement.

➢ The relationship between data-rate and signal-rate is given by

$$S_{ave} = c \times N \times (1/r) \quad \text{baud}$$

where N = data-rate (in bps)
c = case factor, which varies for each case
S = number of signal-elements and
r = previously defined factor.

➢ This relationship depends on
→ value of r.
→ data pattern.

(If we have a data pattern of all 1s or all 0s, the signal-rate may be different from a data pattern of alternating Os and 1s).

3) Bandwidth

- Digital signal that carries information is non-periodic.
- The bandwidth of a non-periodic signal is continuous with an infinite range.
- However, most digital-signals we encounter in real life have a bandwidth with finite values.
- The effective bandwidth is finite.
- The baud rate, not the bit-rate, determines the required bandwidth for a digital-signal.
- More changes in the signal mean injecting more frequencies into the signal.
  (Frequency means change and change means frequency.)
- The bandwidth refers to range of frequencies used for transmitting a signal.
- Relationship b/w baud rate (signal-rate) and the bandwidth (range of frequencies) is given as

$$B_{min} = c \times N \times (1/r)$$

where N = data-rate (in bps)
c = case factor, which varies for each case
r = previously defined factor
B_{min} = minimum bandwidth

4) Baseline Wandering

- While decoding, the receiver calculates a running-average of the received signal-power. This average is called the baseline.
- The incoming signal-power is estimated against this baseline to determine the value of the data-element.
- A long string of 0s or 1s can cause a drift in the baseline (baseline wandering).
  Thus, make it difficult for the receiver to decode correctly.
- A good line-coding scheme needs to prevent baseline wandering.

5) DC Components

- When the voltage-level in a digital-signal is constant for a while, the spectrum creates very low frequencies.
- These frequencies around zero are called DC (direct-current) components.
- DC components present problems for a system that cannot pass low frequencies.
- For example: Telephone line cannot pass frequencies below 200 Hz.
- For Telephone systems, we need a scheme with no DC component.

6) Built-in Error Detection

- Built-in error-detecting capability has to be provided to detect the errors that occurred during transmission.

7) Self Synchronization

- To correctly interpret the signals received from the sender, the receiver's bit intervals must correspond exactly to the sender's bit intervals.
- If the receiver clock is faster or slower, the bit intervals are not matched and the receiver might misinterpret the signals.

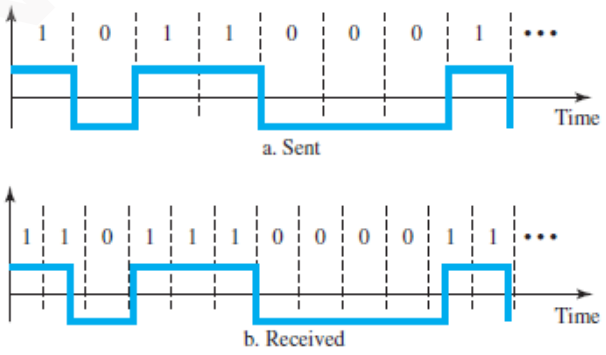


Figure 4.3 *Effect of lack of synchronization*

- As shown in figure 4.3, we have a situation where the receiver has shorter bit duration.
- The sender sends 10110001, while the receiver receives 110111000011.
- A self-synchronizing digital-signal includes timing-information in the data being transmitted.
  - This can be achieved if there are transitions in the signal that alert the receiver to the beginning, middle, or end of the pulse.
  - If the receiver's clock is out-of-synchronization, these points can reset the clock.

8) Immunity to Noise & Interference

- The code should be immune to noise and other interferences.

9) Complexity

- A complex scheme is more costly to implement than a simple one.
- For ex: A scheme that uses 4 signal-levels is more difficult to interpret than one that uses only 2 levels.

Example 1.22

A signal is carrying data in which one data element is encoded as one signal element ($r = 1$). If the bit rate is 100 kbps, what is the average value of the baud rate if c is between 0 and 1?

Solution

We assume that the average value of c is 1/2. The baud rate is then

$$S = c \times N \times (1 / r) = 1/2 \times 100{,}000 \times (1/1) = 50{,}000 = 50 \text{ kbaud}$$

Example 1.23

In a digital transmission, the receiver clock is 0.1 percent faster than the sender clock. How many extra bits per second does the receiver receive if the data rate is 1 kbps? How many if the data rate is 1 Mbps?

Solution

At 1 kbps, the receiver receives 1001 bps instead of 1000 bps.

1000 bits sent → 1001 bits received → 1 extra bps

At 1 Mbps, the receiver receives 1,001,000 bps instead of 1,000,000 bps.

1,000,000 bits sent → 1,001,000 bits received → 1000 extra bps

1.13.2 LINE CODING SCHEMES

• The Line Coding schemes are classified into 3 broad categories (Figure 4.4):

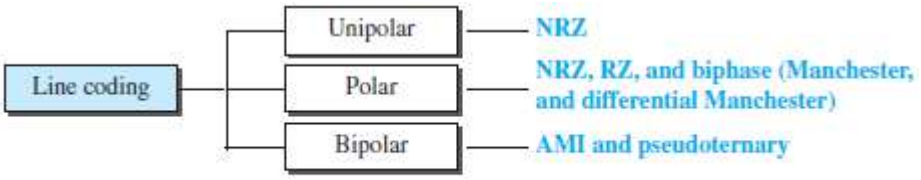


Figure 4.4 *Line coding schemes*

1.13.2.1 Unipolar Scheme

• All signal levels are either above or below the time axis.

NRZ (Non-Return-to-Zero)

- The positive voltage defines bit 1 and the zero voltage defines bit 0 (Figure 4.5).
- It is called NRZ because the signal does not return to 0 at the middle of the bit.

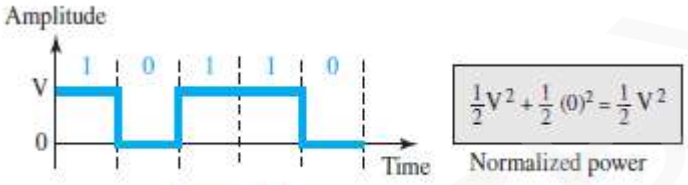


Figure 4.5 *Unipolar NRZ scheme*

- Disadvantages:
  1) Compared to polar scheme, this scheme is very costly.
  2) Also, the normalized power is double that for polar NRZ.
  3) Not suitable for transmission over channels with poor performance around zero frequency.

(Normalized power → power needed to send 1 bit per unit line resistance)

1.13.2.2 Polar Schemes

- The voltages are on the both sides of the time axis.
- Polar NRZ scheme can be implemented with two voltages (V).

For example: -V for bit 1
+V for bit 0.

a) Non-Return-to-Zero (NRZ)

- We use 2 levels of voltage amplitude.
- Two versions of polar NRZ (Figure 4.6):

i) NRZ-L (NRZ-Level)

¤ The level of the voltage determines the value of the bit.
¤ For example: i) Voltage-level for 0 can be positive and
ii) Voltage-level for 1 can be negative.

ii) NRZ-I (NRZ-Invert)

¤ The change or lack of change in the level of the voltage determines the value of the bit.
¤ If there is no change, the bit is 0;
If there is a change, the bit is 1.

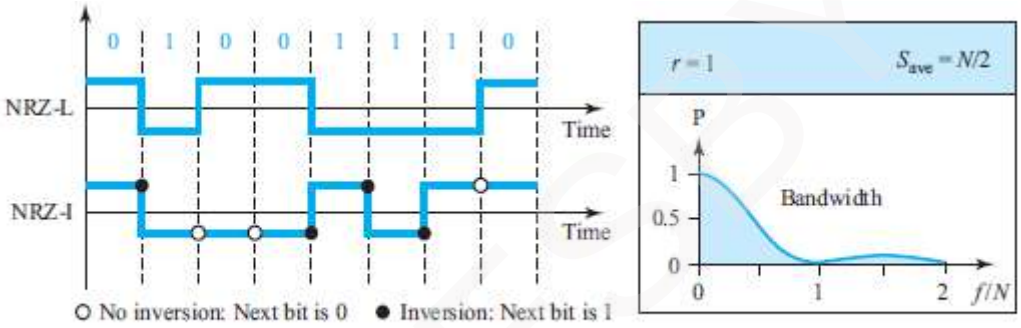


Figure 4.6 *Polar NRZ-L and NRZ-I schemes*

- Disadvantages:

1) **Baseline wandering** is a problem for both variations (NRZ-L NRZ-I).
i) In NRZ-L, if there is a long sequence of 0s or 1s, the average signal-power becomes skewed.
The receiver might have difficulty discerning the bit value.
ii) In NRZ-I, this problem occurs only for a long sequence of 0s.
If we eliminate the long sequence of 0s, we can avoid baseline wandering.
2) The **synchronization problem** also exists in both schemes.
→ A long sequence of 0s can cause a problem in both schemes.
→ A long sequence of 1s can cause a problem in only NRZ-L.
3) In NRZ-L, problem occurs when there is a sudden **change of polarity** in the system.
¤ For example:
In twisted-pair cable, a change in the polarity of the wire results in
→ all 0s interpreted as 1s and
→ all 1s interpreted as 0s.
¤ NRZ-I does not have this problem.
¤ Both schemes have an average signal-rate of N/2 Bd.
4) NRZ-L and NRZ-I both have a **DC component problem**.

Example 1.24

A system is using NRZ-I to transfer 10-Mbps data. What are the average signal rate and minimum bandwidth?

Solution

The average signal rate is $S = N/2 = 500$ kbaud. The minimum bandwidth for this average baud rate is $B_{min} = S = 500$ kHz.

b) Return-to-Zero (RZ)

- In NRZ encoding, problem occurs when the sender-clock and receiver-clock are not synchronized.
- Solution: Use return-to-zero (RZ) scheme (Figure 4.7).
- RZ scheme uses 3 voltages: positive, negative, and zero.
- There is always a transition at the middle of the bit. Either
  - i) from high to zero (for 1) or
  - ii) from low to zero (for 0)

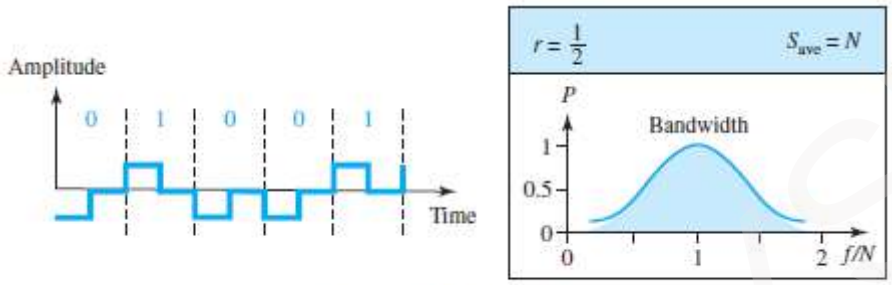


Figure 4.7 *Polar RZ scheme*

- Disadvantages:
  1) RZ encoding requires 2 signal-changes to encode a bit & .'. occupies greater bandwidth.
  2) Complexity: RZ uses 3 levels of voltage, which is more complex to create and detect.
  3) Problem occurs when there is a sudden change of polarity in the system. This result in
     → all 0s interpreted as 1s &
     → all 1s interpreted as 0s.

c) Biphase: Manchester & Differential Manchester

i) Manchester Encoding

- This is a combination of NRZ-L & RZ schemes (RZ→transition at the middle of the bit).
- There is always a transition at the middle of the bit. Either
  - i) from high to low (for 0) or
  - ii) from low to high (for 1).
- It uses only two voltage levels (Figure 4.8).
- The duration of the bit is divided into 2 halves.
- The voltage
  - → remains at one level during the first half &
  - → moves to the other level in the second half.
- The transition at the middle of the bit provides synchronization.

ii) Differential Manchester

- This is a combination of NRZ-I and RZ schemes.
- There is always a transition at the middle of the bit, but the bit-values are determined at the beginning of the bit.
- If the next bit is 0, there is a transition. If the next bit is 1, there is none.

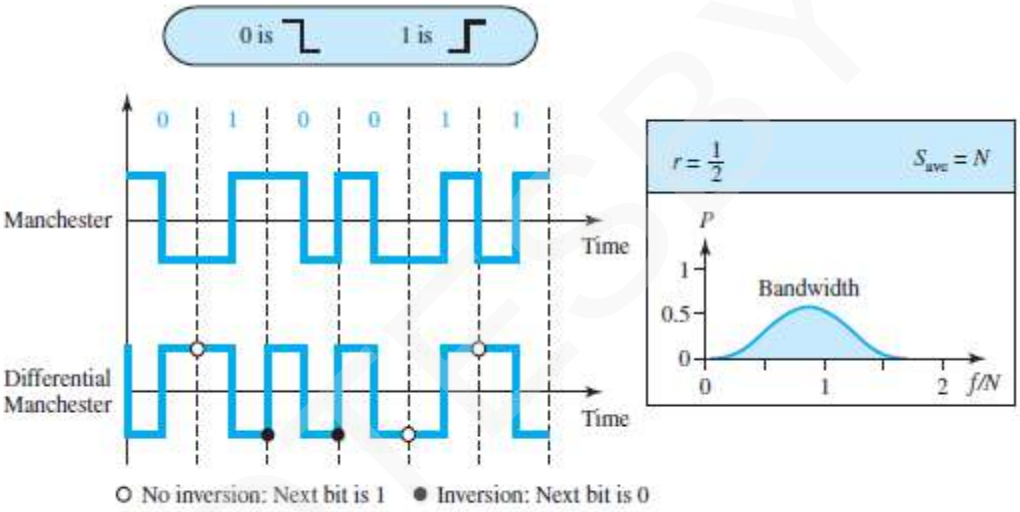


Figure 4.8 *Polar biphase: Manchester and differential Manchester schemes*

- Advantages:
  1) The Manchester scheme overcomes problems associated with NRZ-L.
     Differential Manchester overcomes problems associated with NRZ-I.
  2) There is no baseline wandering.
  3) There is no DC component '.' each bit has a positive & negative voltage contribution.
- Disadvantage:
  1) Signal-rate: Signal-rate for Manchester & diff. Manchester is double that for NRZ.

1.13.2.3 Bipolar Schemes (or Multilevel Binary)

- This coding scheme uses 3 voltage levels (Figure 4.9):
  - i) positive
  - ii) negative &
  - iii) zero.
- Two variations of bipolar encoding:
  - i) AMI (Alternate Mark Inversion)
  - ii) Pseudoternary

i) AMI

- Binary 0 is represented by a neutral 0 voltage (AMI → Alternate 1 Inversion).
- Binary 1s are represented by alternating positive and negative voltages.

ii) Pseudoternary

- Binary 1 is represented by a neutral 0 voltage.
- Binary 0s are represented by alternating positive and negative voltages.

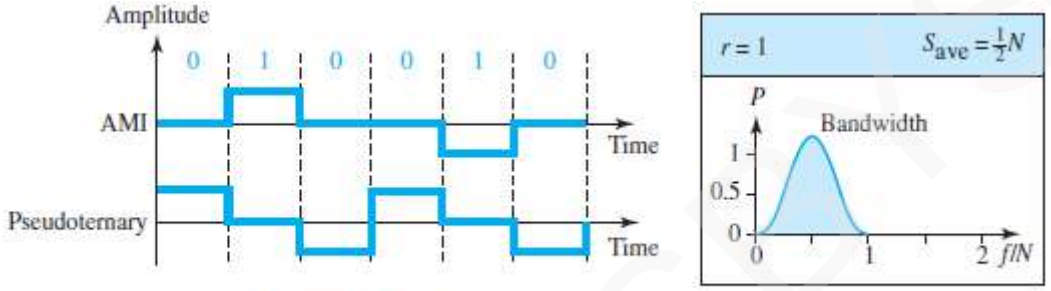


Figure 4.9 *Bipolar schemes: AMI and pseudoternary*

- Advantages:
  1) The bipolar scheme has the same signal-rate as NRZ.
  2) There is no DC component '.' each bit has a positive & negative voltage contribution.
  3) The concentration of the energy is around frequency N/2.
- Disadvantage:
  1) AMI has a synchronization problem when a long sequence of 0s is present in the data.

Table 4.1 *Summary of line coding schemes*

Category	*Scheme*	*Bandwidth (average)*	*Characteristics*
Unipolar	NRZ	$B = N/2$	Costly, no self-synchronization if long 0s or 1s, DC
Polar	NRZ-L	$B = N/2$	No self-synchronization if long 0s or 1s, DC
	NRZ-I	$B = N/2$	No self-synchronization for long 0s, DC
	Biphase	$B = N$	Self-synchronization, no DC, high bandwidth
Bipolar	AMI	$B = N/2$	No self-synchronization for long 0s, DC

MODULE-WISE QUESTIONS

MODULE 1: INTRODUCTION

1) Define data communications. Explain its 4 fundamental characteristics. (4*)
2) Explain different components of data communication system. (6*)
3) Explain different forms of information. (4)
4) Describe simplex, half-duplex and full duplex methods of data flow. (6*)
5) Explain the 3 criteria necessary for an effective and efficient network. (4*)
6) Explain point to point and multipoint connection. (4*)
7) Explain the following topologies:
   i) Mesh ii) Star iii) Bus iv) Ring (12*)
8) Explain in detail LAN & WAN. List the differences between LAN & WAN. (10*)
9) Explain circuit-switched and packet-switched networks. (6*)
10) Explain components of Internet. (4*)
11) Explain different method of accessing the Internet. (6*)
12) With respect to RFC, explain the following:
   i) Maturity Levels ii) Requirement Levels. (6)
13) What are internet standards? Explain the functions of standard organizations. (4*)

MODULE 1(CONT.): NETWORK MODELS

1) Explain TCP/IP architecture with a layer diagram. (4*)
2) List the 5 layers and its functionality in TCP/IP model. (8*)
3) With respect to in TCP/IP model, explain the following:
   i) Encapsulation and decapsulation. ii) Multiplexing and demultiplexing. (8)
4) Explain four levels of addressing employed in TCP/IP protocol. (6*)
5) What are the uses of a layered network model? Compare OSI and TCP/IP models. (4)

MODULE 1(CONT.): DATA AND SIGNALS

1) Compare the following:
   i) Analog signal vs. Digital signal. ii) Periodic signal vs. Non-periodic signal. (4)
2) Describe digital signal as a composite analog signal. (4)
3) Explain 2 methods for transmitting a digital signal (8*)
4) What do you mean by transmission impairment? Explain causes of transmission impairment? (6*)
5) What are the three factors data rate is dependent on? Explain the theoretical formula which was developed to calculate the data rate. (8*)
6) Explain 4 performance parameters of network. (8*)

MODULE 1(CONT.): DIGITAL TRANSMISSION

1) Explain in detail any 6 characteristics of digital signal. (6*)
2) Compare the following:
   i) Data element vs. Signal element. ii) Data rate vs. Signal rate. (4)
3) Explain following encoding schemes with example:
   i) Unipolar Scheme ii) Polar Schemes iii) Bipolar Schemes (8*)
4) Represent the following sequences using different line coding schemes.
   i) 101011100. ii) 10110011. iii) 00110101. (6*)
5) Define the following:
   i) Network ii) Internet iii) Protocol iv) Decibel v) SNR vi) Line coding (6*)

MODULE 2: TABLE OF CONTENTS

MODULE 2: DIGITAL TRANSMISSION (CONT.)

2.1 ANALOG TO DIGITAL CONVERSION

- An analog-signal may created by a microphone or camera.
- To change an analog-signal to digital-data, we use PCM (pulse code modulation).
- After the digital-data are created (digitization), then we convert the digital-data to a digital-signal.

2.1.1 PCM

- PCM is a technique used to change an analog signal to digital data (digitization).
- PCM has encoder at the sender and decoder at the receiver.
- The encoder has 3 processes (Figure 4.21):
  1) Sampling
  2) Quantization &
  3) Encoding.

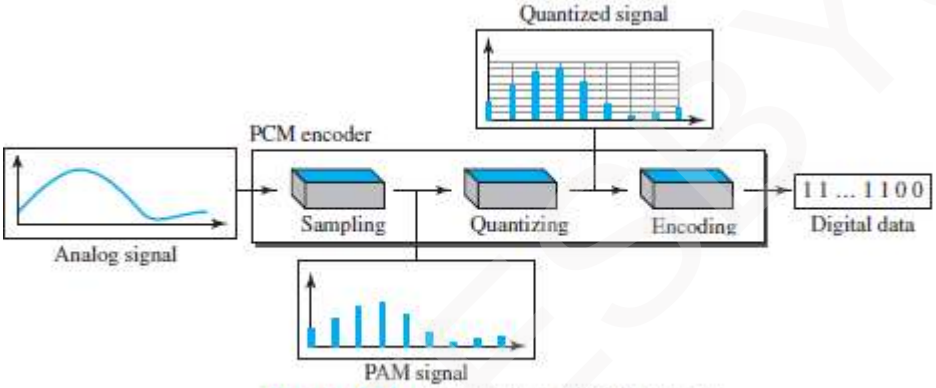


Figure 4.21 *Components of PCM encoder*

2.1.1.1 Sampling

- We convert the continuous time signal (analog) into the discrete time signal (digital).
- Pulses from the analog-signal are sampled every T_s sec
  where T_s is the sample-interval or period.
- The inverse of the sampling-interval is called the sampling-frequency (or sampling-rate).
- Sampling-frequency is given by

$$f_s = 1/T_s$$

- Three sampling methods (Figure 4.22):
  1) Ideal Sampling
  - This method is difficult to implement.

  2) Natural Sampling
  - A high-speed switch is turned ON for only the small period of time when the sampling occurs.
  - The result is a sequence of samples that retains the shape of the analog-signal.

  3) Flat Top Sampling
  - The most common sampling method is sample and hold.
  - Sample and hold method creates flat-top samples.
  - This method is sometimes referred to as *PAM* (pulse amplitude modulation).

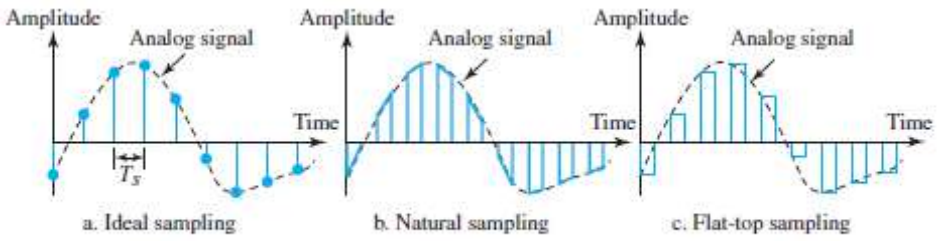


Figure 4.22 *Three different sampling methods for PCM*

2.1.1.1.1 Sampling Rate

- According to Nyquist theorem,

"The sampling-rate must be at least 2 times the highest frequency, not the bandwidth".

i) If the analog-signal is **low-pass**, the bandwidth and the highest frequency are the same value (Figure 4.23a).

ii) If the analog-signal is **bandpass**, the bandwidth value is lower than the value of the maximum frequency (Figure 4.23b).

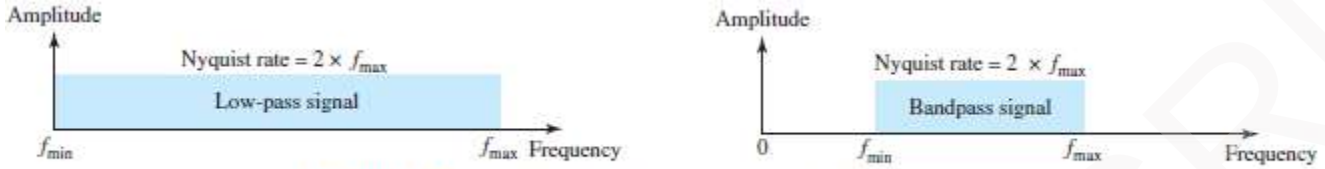


Figure 4.23 *Nyquist sampling rate for low-pass and bandpass signals*

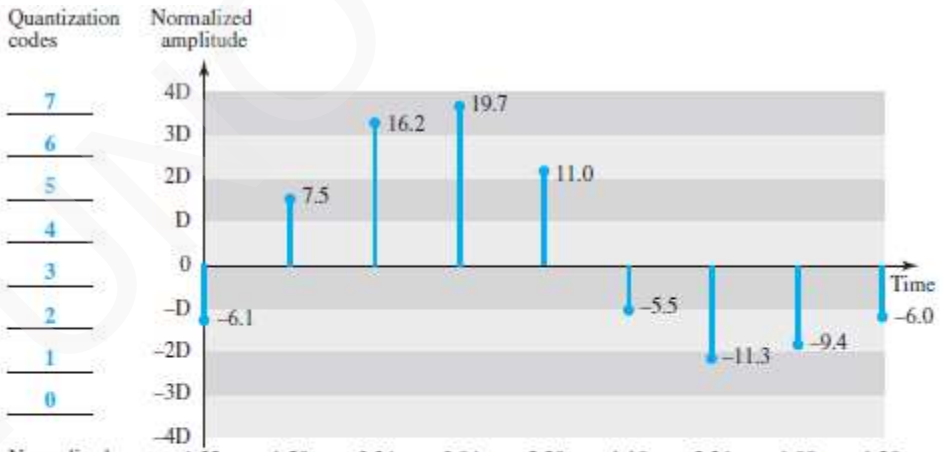
2.1.2 Quantization

- The sampled-signal is quantized.
- Result of sampling is a set of pulses with amplitude-values b/w max & min amplitudes of the signal.
- Four steps in quantization:
  1) We assume that the original analog-signal has amplitudes between V_{min} & V_{max}.
  2) We divide the range into L zones, each of height Δ(delta).

$$\Delta = \frac{V_{max} - V_{min}}{L}$$

where L = number of levels.

  3) We assign quantized values of 0 to (L-1) to the midpoint of each zone.
  4) We approximate the value of the sample amplitude to the quantized values.
- For example: Let V_{min}=-20 V_{max} =+20 V L = 8 Therefore, Δ = [+20-(-20)]/8= 5 V
- In the chart (Figure 4.26),
  1) First row is normalized-PAM-value for each sample.
  2) Second row is normalized-quantized-value for each sample.
  3) Third row is normalized error (which is diff. b/w normalized PAM value & quantized values).
  4) Fourth row is quantization code for each sample.
  5) Fifth row is the encoded words (which are the final products of the conversion).



Normalized PAM values	–1.22	1.50	3.24	3.94	2.20	–1.10	–2.26	–1.88	–1.20
Normalized quantized values	–1.50	1.50	3.50	3.50	2.50	–1.50	–2.50	–1.50	–1.50
Normalized error	–0.38	0	+0.26	–0.44	+0.30	–0.40	–0.24	+0.38	–0.30
Quantization code	2	5	7	7	6	2	1	2	2
Encoded words	010	101	111	111	110	010	001	010	010

Figure 4.26 *Quantization and encoding of a sampled signal*

2.1.2.1 Quantization Level

- Let L = number of levels.
- The choice of L depends on
  → range of the amplitudes of the analog-signal and
  → how accurately we need to recover the signal.
- If the signal has only 2 amplitude values, we need only 2 quantization-levels.
  If the signal (like voice) has many amplitude values, we need more quantization-levels.
- In audio digitizing, L is normally chosen to be 256.
  In video digitizing, L is normally thousands.
- Choosing lower values of L increases the quantization-error.

2.1.2.2 Quantization Error

- Quantization-error is the difference b/w normalized PAM value & quantized values
- Quantization is an approximation process.
- The input values to the quantizer are the real values.
  The output values from the quantizer are the approximated values.
- The output values are chosen to be the middle value in the zone.
- If the input value is also at the middle of the zone,
  Then, there is no error.
  Otherwise, there is an error.
- In the previous example,
  The normalized amplitude of the third sample is 3.24, but the normalized quantized value is 3.50. This means that there is an error of +0.26.

2.1.2.3 Uniform vs. Non Uniform Quantization

- Non-uniform quantization can be done by using a process called companding and expanding.
  1) The signal is companded at the sender before conversion.
  2) The signal is expanded at the receiver after conversion.
- Companding means reducing the instantaneous voltage amplitude for large values.
  Expanding means increasing the instantaneous voltage amplitude for small values.
- It has been proved that non-uniform quantization effectively reduces the SNR_{dB} of quantization.

2.1.3 Encoding

- The quantized values are encoded as n-bit code word.
- In the previous example,
  A quantized value 2 is encoded as 010.
  A quantized value 5 is encoded as 101.
- Relationship between number of quantization-levels (L) & number of bits (n) is given by
  $n=\log_2 L$ or $2^n=L$
- The bit-rate is given by:

$$\text{Bit rate} = \text{sampling rate} \times \text{number of bits per sample} = f_s \times n$$

Example 2.1

A complex low-pass signal has a bandwidth of 200 kHz. What is the minimum sampling rate for this signal?

Solution

The bandwidth of a low-pass signal is between 0 and f, where f is the maximum frequency in the signal. Therefore, we can sample this signal at 2 times the highest frequency (200 kHz). The sampling rate is therefore 400,000 samples per second.

Example 2.2

What is the SNR_{dB} in the example of Figure 4.26?

Solution

We can use the formula to find the quantization. We have eight levels and 3 bits per sample, so $SNR_{dB} = 6.02(3) + 1.76 = 19.82$ dB. Increasing the number of levels increases the SNR.

Example 2.3

A telephone subscriber line must have an SNR_{dB} above 40. What is the minimum number of bits per sample?

Solution

We can calculate the number of bits as

$$SNR_{dB} = 6.02n_b + 1.76 = 40 \rightarrow n = 6.35$$

Telephone companies usually assign 7 or 8 bits per sample.

Example 2.4

We want to digitize the human voice. What is the bit rate, assuming 8 bits per sample?

Solution

The human voice normally contains frequencies from 0 to 4000 Hz. So the sampling rate and bit rate are calculated as follows:

$$\text{Sampling rate} = 4000 \times 2 = 8000 \text{ samples/s}$$

$$\text{Bit rate} = 8000 \times 8 = 64{,}000 \text{ bps} = 64 \text{ kbps}$$

2.1.3.1 Original Signal Recovery

- PCM decoder is used for recovery of the original signal.
- Here is how it works (Figure 4.27):
  1) The decoder first uses circuitry to convert the code words into a pulse that holds the amplitude until the next pulse.
  2) Next, the staircase-signal is passed through a low-pass filter to smooth the staircase signal into an analog-signal.
- The filter has the same cut-off frequency as the original signal at the sender.
- If the signal is sampled at the Nyquist sampling-rate, then the original signal will be re-created.
- The maximum and minimum values of the original signal can be achieved by using amplification.

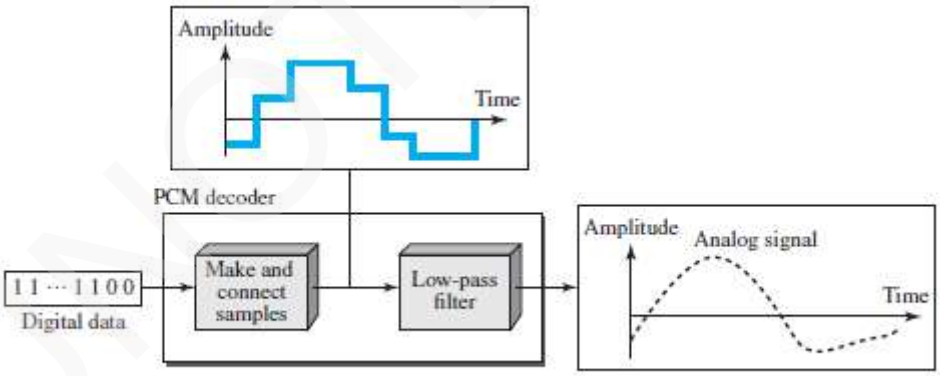


Figure 4.27 *Components of a PCM decoder*

2.1.3.2 PCM Bandwidth

- The minimum bandwidth of a line-encoded signal is

$$B_{min} = c \times N \times \frac{1}{r}$$

- We substitute the value of N in above formula:

$$B_{min} = c \times N \times \frac{1}{r} = c \times n_b \times f_s \times \frac{1}{r} = c \times n_b \times 2 \times B_{analog} \times \frac{1}{r}$$

- When 1/r = 1 (for a NRZ or bipolar signal) and c = (1/2) (the average situation), the minimum bandwidth is

$$B_{min} = n_b \times B_{analog}$$

- This means the minimum bandwidth of the digital-signal is n_b times greater than the bandwidth of the analog-signal.

2.1.3.3 Maximum Data Rate of a Channel

• The Nyquist theorem gives the data-rate of a channel as

$$N_{max} = 2 \times B \times \log_2 L$$

• We can deduce above data-rate from the Nyquist sampling theorem by using the following arguments.

1) We assume that the available channel is low-pass with bandwidth B.
2) We assume that the digital-signal we want to send has L levels, where each level is a signal-element. This means r = $1/\log_2 L$.
3) We first pass digital-signal through a low-pass filter to cut off the frequencies above B Hz.
4) We treat the resulting signal as an analog-signal and sample it at 2 x B samples per second and quantize it using L levels.
5) The resulting bit-rate is

$$N = f_s \times n_b = 2 \times B \times \log_2 L$$

This is the maximum bandwidth; if the case factor c increases, the data-rate is reduced.

$$N_{max} = 2 \times B \times \log_2 L \quad \text{bps}$$

2.1.3.4 Minimum Required Bandwidth

• The previous argument can give us the minimum bandwidth if the data-rate and the number of signal-levels are fixed. We can say

$$B_{min} = \frac{N}{(2 \times \log_2) L} \quad \text{Hz}$$

2.2 TRANSMISSION MODES

• Two ways of transmitting data over a link (Figure 4.31): 1) Parallel mode & 2) Serial mode.

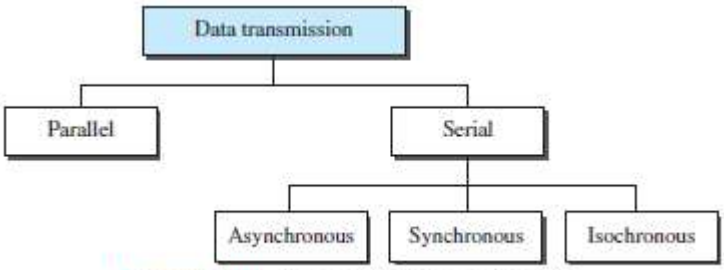


Figure 4.31 *Data transmission and modes*

2.1.1 PARALLEL TRANSMISSION

• Multiple bits are sent with each clock-tick (Figure 4.32).
• 'n' bits in a group are sent simultaneously.
• 'n' wires are used to send 'n' bits at one time.
• Each bit has its own wire.
• Typically, the 8 wires are bundled in a cable with a connector at each end.

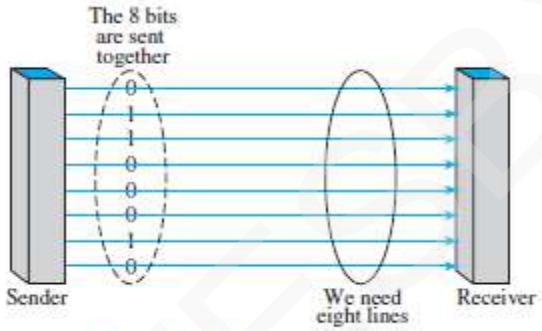


Figure 4.32 *Parallel transmission*

• Advantage:
  1) Speed: Parallel transmission can increase the transfer speed by a factor of n over serial transmission.
• Disadvantage:
  1) Cost: Parallel transmission requires n communication lines just to transmit the data-stream. Because this is expensive, parallel transmission is usually limited to short distances.

2.2.2 SERIAL TRANSMISSION

• One bit is sent with each clock-tick using only a single link (Figure 4.33).

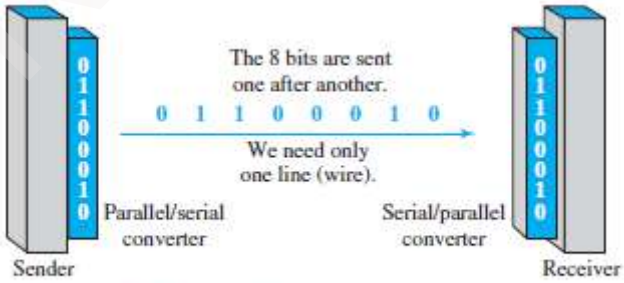


Figure 4.33 *Serial transmission*

• Advantage:
  1) Cost: Serial transmission reduces cost of transmission over parallel by a factor of n.
• Disadvantage:
  1) Since communication within devices is parallel, following 2 converters are required at interface:
    i) Parallel-to-serial converter
    ii) Serial-to-parallel converter
• Three types of serial transmission: asynchronous, synchronous, and isochronous.

2.2.2.1 Asynchronous Transmission

- Asynchronous transmission is so named because the timing of a signal is not important (Figure 4.34).
- Prior to data transfer, both sender & receiver agree on pattern of information to be exchanged.
- Normally, patterns are based on grouping the bit-stream into bytes.
- The sender transmits each group to the link without regard to a timer.
- As long as those patterns are followed, the receiver can retrieve the info. without regard to a timer.
- There may be a gap between bytes.
- We send
  → 1 start bit (0) at the beginning of each byte
  → 1 stop bit (1) at the end of each byte.
- Start bit alerts the receiver to the arrival of a new group.
  Stop bit lets the receiver know that the byte is finished.
- Here, the term asynchronous means "asynchronous at the byte level".
  However, the bits are still synchronized & bit-durations are the same.

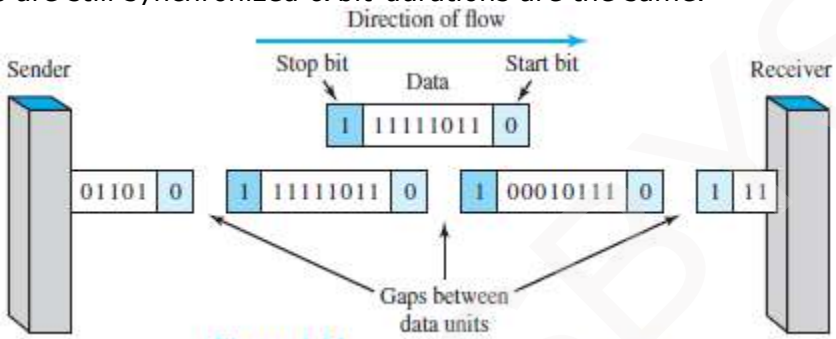


Figure 4.34 *Asynchronous transmission*

- Disadvantage:
  1) Slower than synchronous transmission. (Because of stop bit, start bit and gaps)
- Advantages:
  1) Cheap & effective.
  2) Useful for low-speed communication.

2.2.2.2 Synchronous Transmission

- We send bits one after another without start or stop bits or gaps (Figure 4.35).
- The receiver is responsible for grouping the bits.
- The bit-stream is combined into longer "frames," which may contain multiple bytes.
- If the sender wants to send data in separate bursts, the gaps b/w bursts must be filled with a special sequence of 0s & 1s (that means idle).

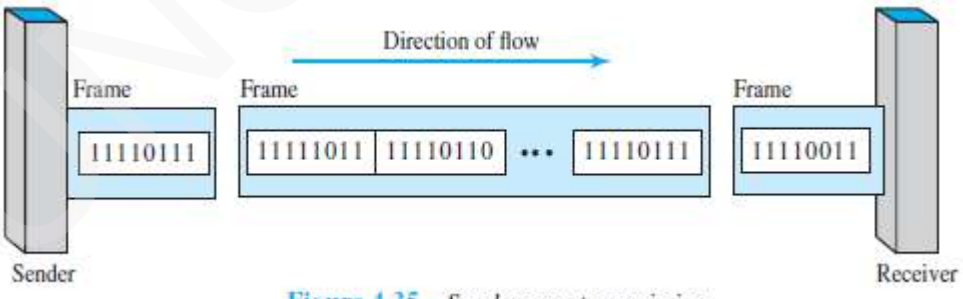


Figure 4.35 *Synchronous transmission*

- Advantages:
  1) Speed: Faster than asynchronous transmission. ('.' of no stop bit, start bit and gaps).
  2) Useful for high-speed applications such as transmission of data from one computer to another.

2.2.2.3 Isochronous

- Synchronization between characters is not enough; the entire stream of bits must be synchronized.
- The isochronous transmission guarantees that the data arrive at a fixed rate.
- In real-time audio/video, jitter is not acceptable. Therefore, synchronous transmission fails.
- For example: TV images are broadcast at the rate of 30 images per second. The images must be viewed at the same rate.

MODULE 2(CONT.): ANALOG TRANSMISSION

2.3 DIGITAL TO ANALOG CONVERSION

• Digital-to-analog conversion is the process of changing one of the characteristics of an analog-signal based on the information in digital-data (Figure 5.1).

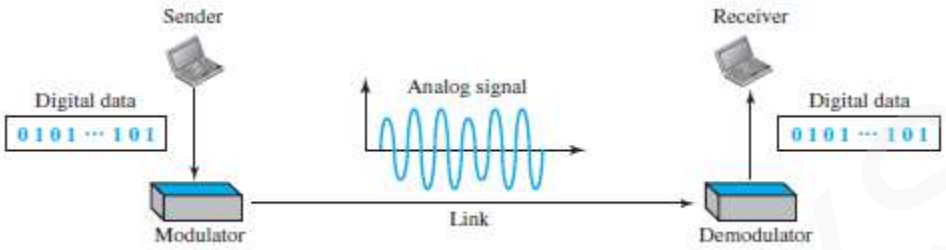


Figure 5.1 *Digital-to-analog conversion*

• A sine wave can be defined by 3 attributes:
  1) Amplitude
  2) Frequency &
  3) Phase.
• When anyone of the 3 attributes of a wave is varied, a different version of the wave will be created.
• So, by changing one attribute of an analog signal, we can use it to represent digital-data.
• Four methods of digital to analog conversion (Figure 5.2):
  1) Amplitude shift keying (ASK)
  2) Frequency shift keying (FSK)
  3) Phase shift keying (PSK)
  4) Quadrature amplitude modulation (QAM).
• QAM is a combination of ASK and PSK i.e. QAM combines changing both the amplitude and phase.
  QAM is the most efficient of these 4 methods.
  QAM is the method commonly used today.

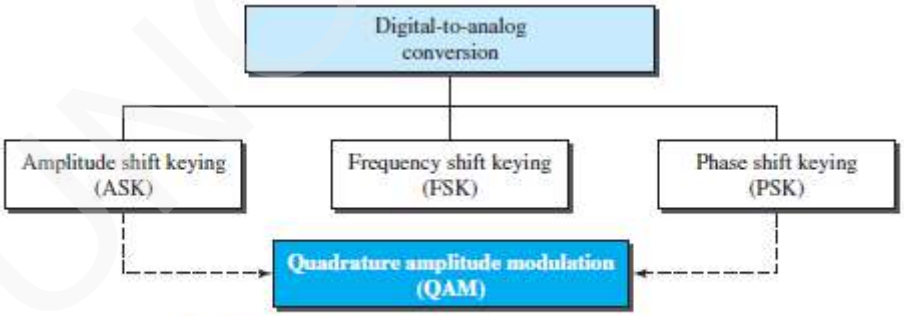


Figure 5.2 *Types of digital-to-analog conversion*

2.3.1 Aspects of Digital-to-Analog Conversion

1) Data Element vs. Signal Element

- A data-element is the smallest piece of information to be exchanged i.e. the bit.
- A signal-element is the smallest unit of a signal that is transmitted.

2) Data Rate vs. Signal Rate

- Data rate (Bit rate) is the number of bits per second.

  Signal-rate (Baud rate) is the number of signal elements per second.
- The relationship between data-rate(N) and the signal-rate(S) is

$$S = N \times \frac{1}{r} \quad \text{baud}$$

  where r = number of data-elements carried in one signal-element.
- The value of r is given by

  $r = \log_2 L$ or $2^r = L$

  where L = type of signal-element (not the level)

  (In transportation,

  → a baud is analogous to a vehicle, and

  → a bit is analogous to a passenger.

  We need to maximize the number of people per car to reduce the traffic).

3) Carrier Signal

- The sender produces a high-frequency signal that acts as a base for the information-signal.
- This base-signal is called the carrier-signal (or carrier-frequency).
- The receiver is tuned to the frequency of the carrier-signal that it expects from the sender.
- Then, digital-information changes the carrier-signal by modifying its attributes (amplitude, frequency, or phase). This kind of modification is called modulation (shift keying).

4) Bandwidth

- In both ASK & PSK, the bandwidth required for data transmission is proportional to the signal-rate.
- In FSK, the bandwidth required is the difference between the two carrier-frequencies.

Example 2.5

An analog signal carries 4 bits per signal element. If 1000 signal elements are sent per second, find the bit rate.

Solution

In this case, $r = 4$, $S = 1000$, and N is unknown. We can find the value of N from

$$S = N \times (1/r) \quad \text{or} \quad N = S \times r = 1000 \times 4 = 4000 \text{ bps}$$

Example 2.6

An analog signal has a bit rate of 8000 bps and a baud rate of 1000 baud. How many data elements are carried by each signal element? How many signal elements do we need?

Solution

In this example, $S = 1000$, $N = 8000$, and r and L are unknown. We first find the value of r and then the value of L.

$$S = N \times 1/r \longrightarrow r = N/S = 8000/10{,}000 = 8 \text{ bits/baud}$$

$$r = \log_2 L \longrightarrow L = 2^r = 2^8 = 256$$

2.3.2 Amplitude Shift Keying (ASK)

•The amplitude of the carrier-signal is varied to represent different signal-elements.
• Both frequency and phase remain constant for all signal-elements.

2.3.2.1 Binary ASK (BASK)

• BASK is implemented using only 2 levels. (Figure 5.3)
• This is also referred to as OOK (On-Off Keying).

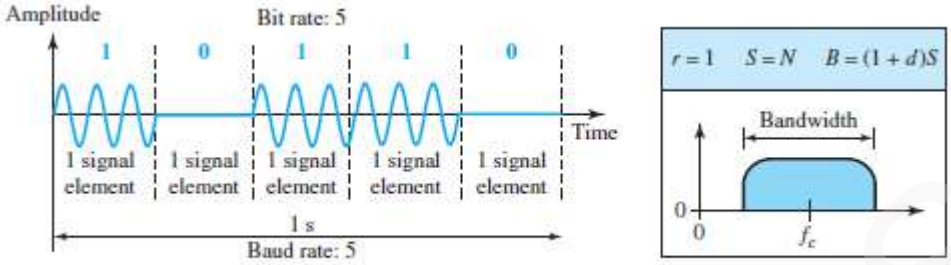


Figure 5.3 *Binary amplitude shift keying*

2.3.2.1.1 Implementation of BASK

• Here, line coding method used = unipolar NRZ (Figure 5.4).
• The unipolar NRZ signal is multiplied by the carrier-frequency coming from an oscillator.
1) When amplitude of the NRZ signal = 0, amplitude of the carrier-signal = 0.
2) When amplitude of the NRZ signal = 1, the amplitude of the carrier-signal is held.

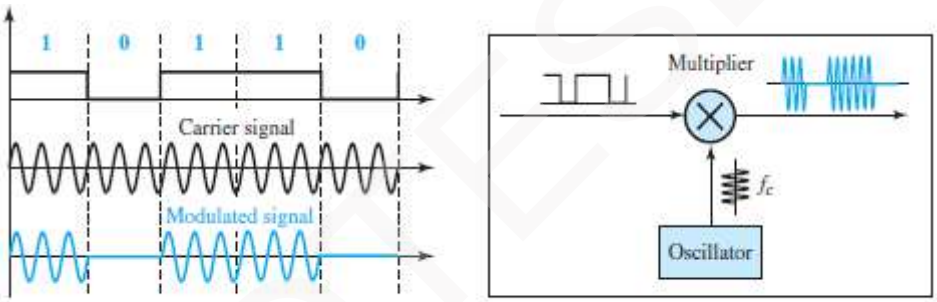


Figure 5.4 *Implementation of binary ASK*

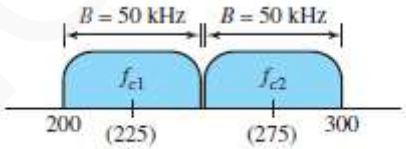


Figure 5.5 *Bandwidth of full-duplex ASK*

2.3.2.1.2 Bandwidth for ASK

• Here, the bandwidth (B) is proportional to the signal-rate (S) (Figure 5.5)
• The bandwidth is given by

$$B = (1 + d) \times S$$

where d(0<d<1)= this factor depends on modulation and filtering-process.

Example 2.7

We have an available bandwidth of 100 kHz which spans from 200 to 300 kHz. What are the carrier frequency and the bit rate if we modulated our data by using ASK with $d = 1$?

Solution

The middle of the bandwidth is located at 250 kHz. This means that our carrier frequency can be at $f_c = 250$ kHz. We can use the formula for bandwidth to find the bit rate (with $d = 1$ and $r = 1$).

$$B = (1 + d) \times S = 2 \times N \times (1/r) = 2 \times N = 100 \text{ kHz} \longrightarrow N = 50 \text{ kbps}$$

2.3.3 Frequency Shift Keying (FSK)

- The frequency of the carrier-signal is varied to represent different signal-elements.
- The frequency of the modulated-signal is constant for the duration of one signal-element, but changes for the next signal-element if the data-element changes.
- Both amplitude and phase remain constant for all signal-elements.

2.3.3.1 Binary FSK (BFSK)

- This uses 2 carrier-frequencies: f1 and f2. (Figure 5.6)
  1) When data-element = 1, first carrier frequency(f1) is used.
  2) When data-element = 0, second carrier frequency(f2) is used.

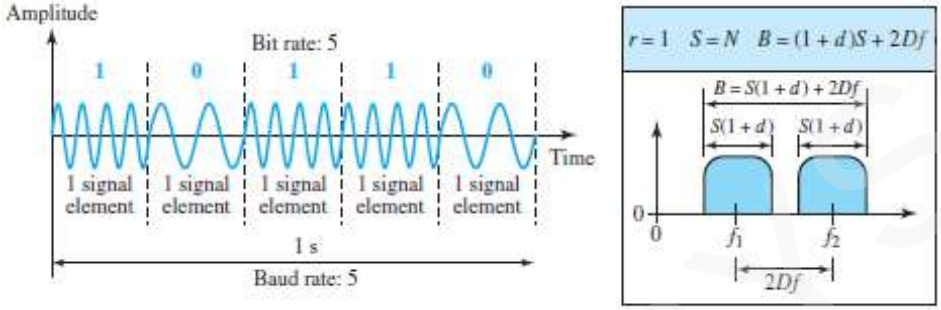


Figure 5.6 *Binary frequency shift keying*

2.3.3.1.1 Implementation

- Here, line coding method used = unipolar NRZ.
- Two implementations of BFSK: i) Coherent and ii) Non-Coherent.

Coherent BFSK	Non Coherent BFSK
The phase continues through the boundary of two signal-elements (Figure 5.7).	There may be discontinuity in the phase when one signal-element ends and the next begins.
This is implemented by using one voltage-controlled oscillator (VCO). VCO changes frequency according to the input voltage.	This is implemented by → treating BFSK as 2 ASK modulations and → using 2 carrier-frequencies
When the amplitude of NRZ signal = 0, the VCO keeps its regular frequency. When the amplitude of NRZ signal = 0, the VCO increases its frequency.	

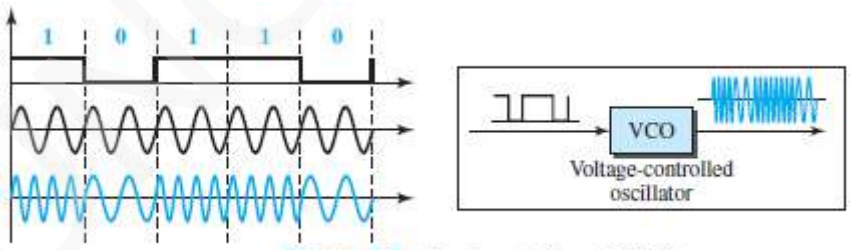


Figure 5.7 *Implementation of BFSK*

2.3.3.1.2 Bandwidth for BFSK

- FSK has two ASK signals, each with its own carrier-frequency f1 or f2. (Figure 5.6)
- The bandwidth is given by

$B = (1 + d) \times S + 2\Delta f$ where 2Δf is the difference between f1 and f2,

Example 2.8

We have an available bandwidth of 100 kHz which spans from 200 to 300 kHz. What should be the carrier frequency and the bit rate if we modulated our data by using FSK with $d = 1$?

Solution

This problem is similar to Example 5.3, but we are modulating by using FSK. The midpoint of the band is at 250 kHz. We choose $2\Delta_f$ to be 50 kHz; this means

$$B = (1 + d) \times S + 2\Delta_f = 100 \longrightarrow 2S = 50\,\text{kHz} \longrightarrow S = 25\,\text{kbaud} \longrightarrow N = 25\,\text{kbps}$$

Example 2.9

We need to send data 3 bits at a time at a bit rate of 3 Mbps. The carrier frequency is 10 MHz. Calculate the number of levels (different frequencies), the baud rate, and the bandwidth.

Solution

We can have $L = 2^3 = 8$. The baud rate is $S = 3$ MHz/3 = 1 Mbaud. This means that the carrier frequencies must be 1 MHz apart ($2\Delta_f = 1$ MHz). The bandwidth is $B = 8 \times 1 = 8$ MHz. Figure 5.8 shows the allocation of frequencies and bandwidth.

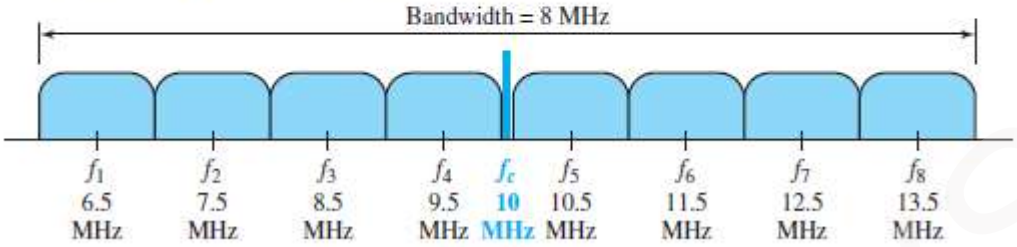


Figure 5.8 *Bandwidth of MFSK used*

2.3.4 Phase Shift Keying (PSK)

- The phase of the carrier-signal is varied to represent different signal-elements.
- Both amplitude and frequency remain constant for all signal-elements.

2.3.4.1 Binary PSK (BPSK)

- We have only two signal-elements:
  1) First signal-element with a phase of 0°.
  2) Second signal-element with a phase of 180° (Figure 5.9).
- ASK vs. PSK
  - In ASK, the criterion for bit detection is the amplitude of the signal.
  - In PSK, the criterion for bit detection is the phase.
- Advantages:
  1) PSK is less susceptible to noise than ASK.
  2) PSK is superior to FSK because we do not need 2 carrier-frequencies.
- Disadvantage:
  1) PSK is limited by the ability of the equipment to distinguish small differences in phase.

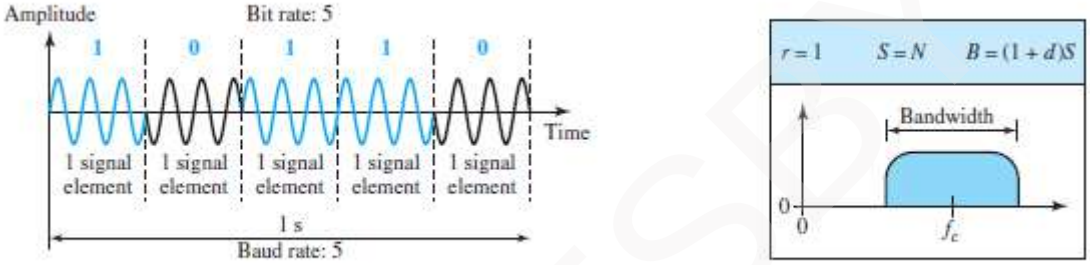


Figure 5.9 *Binary phase shift keying*

2.3.4.1.1 Implementation

- The implementation of BPSK is as simple as that for ASK. (Figure 5.10).
- The signal-element with phase 180° can be seen as the complement of the signal-element with phase 0°.
- Here, line coding method used: polar NRZ.
- The polar NRZ signal is multiplied by the carrier-frequency coming from an oscillator.
  1) When data-element = 1, the phase starts at 0°.
  2) When data-element = 0, the phase starts at 180°.

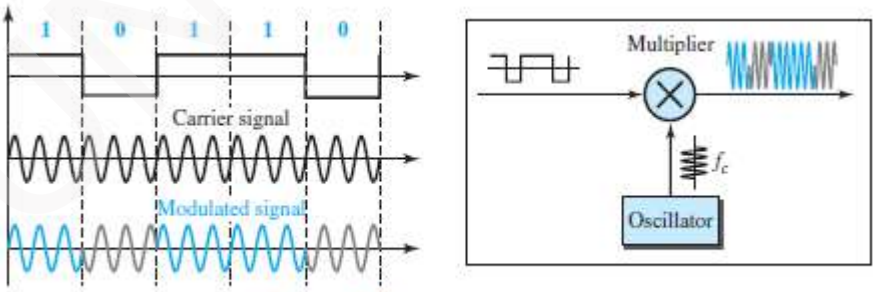


Figure 5.10 *Implementation of BASK*

2.3.4.1.2 Bandwidth for BPSK

- The bandwidth is the same as that for BASK, but less than that for BFSK. (Figure 5.9b)
- No bandwidth is wasted for separating 2 carrier-signals.

2.3.4.2 Quadrature PSK (QPSK)

- The scheme is called QPSK because it uses 2 separate BPSK modulations (Figure 5.11):
  1) First modulation is in-phase,
  2) Second modulation is quadrature (out-of-phase).
- A serial-to-parallel converter
  → accepts the incoming bits
  → sends first bit to first modulator and
  → sends second bit to second modulator.
- The bit to each BPSK signal has one-half the frequency of the original signal.
- Advantages:
  1) Decreases the baud rate.
  2) Decreases the required bandwidth.

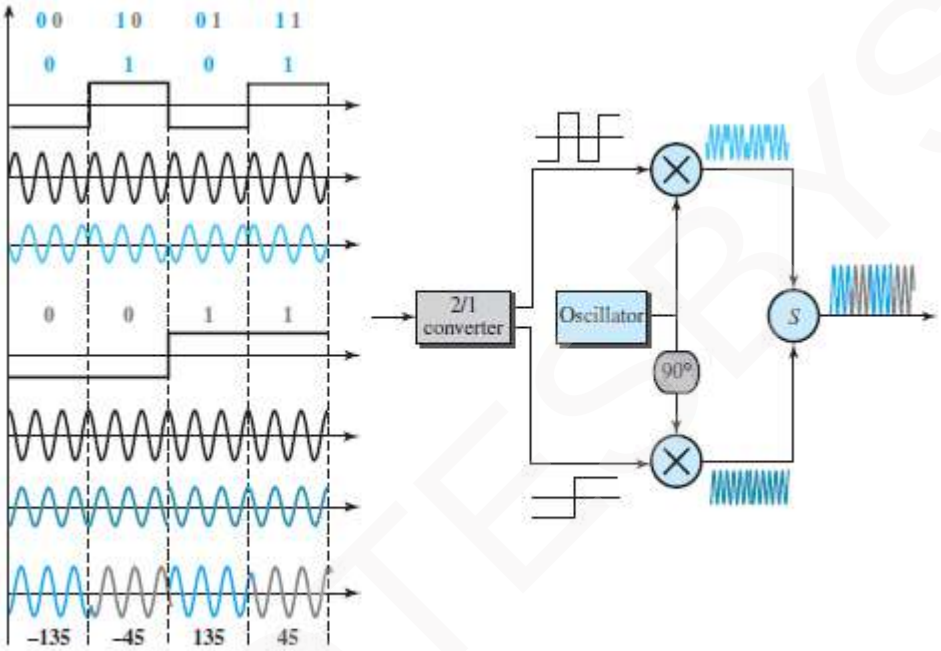


Figure 5.11 *QPSK and its implementation*

- As shown in Figure 5.11, the 2 composite-signals created by each multiplier are 2 sine waves with the same frequency, but different phases.
- When the 2 sine waves are added, the result is another sine wave, with 4 possible phases: 45°, -45°, 135°, and -135°.
- There are 4 kinds of signal-elements in the output signal (L=4), so we can send 2 bits per signal-element (r=2).

Example 2.10

Find the bandwidth for a signal transmitting at 12 Mbps for QPSK. The value of $d = 0$.

Solution

For QPSK, 2 bits are carried by one signal element. This means that $r = 2$. So the signal rate (baud rate) is $S = N \times (1/r) = 6$ Mbaud. With a value of $d = 0$, we have $B = S = 6$ MHz.

2.3.4.3 Constellation Diagram

- A constellation diagram can be used to define the amplitude and phase of a signal-element.
- This diagram is particularly useful
  - → when 2 carriers (one in-phase and one quadrature) are used.
  - → when dealing with multilevel ASK, PSK, or QAM.
- In a constellation diagram, a signal-element type is represented as a dot.
- The diagram has 2 axes (Figure 5.12):
  1) The horizontal X axis is related to the in-phase carrier.
  2) The vertical Y axis is related to the quadrature carrier.
- For each point on the diagram, 4 pieces of information can be deduced.
  1) The projection of point on the X axis defines the peak amplitude of the in-phase component.
  2) The projection of point on Y axis defines peak amplitude of the quadrature component.
  3) The length of the line that connects the point to the origin is the peak amplitude of the signal-element (combination of the X and Y components);
  4) The angle the line makes with the X axis is the phase of the signal-element.

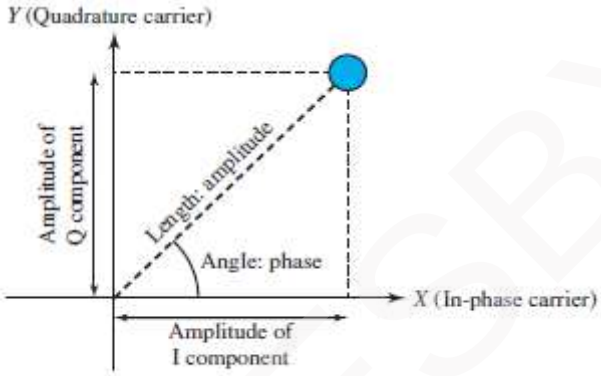


Figure 5.12 *Concept of a constellation diagram*

Example 2.11

Show the constellation diagrams for ASK (OOK), BPSK, and QPSK signals.

Solution

Figure 5.13 shows the three constellation diagrams. Let us analyze each case separately:

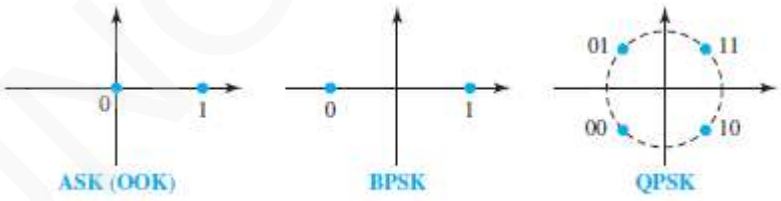


Figure 5.13 *Three constellation diagrams*

2.3.5 Quadrature Amplitude Modulation (QAM)

• This is a combination of ASK and PSK.

• Main idea: Using 2 carriers, one in-phase and the other quadrature, with different amplitude levels for each carrier.

• There are many variations of QAM (Figure 5.14).

A) Figure 5.14a shows the 4-QAM scheme using a unipolar NRZ signal. This is same as BASK.

B) Figure 5.14b shows another QAM using polar NRZ. This is the same as QPSK.

C) Figure 5.14c shows another 4-QAM in which we used a signal with 2 positive levels to modulate each of the 2 carriers.

D) Figure 5.14d shows a 16-QAM constellation of a signal with 8 levels, 4 positive & 4 negative.

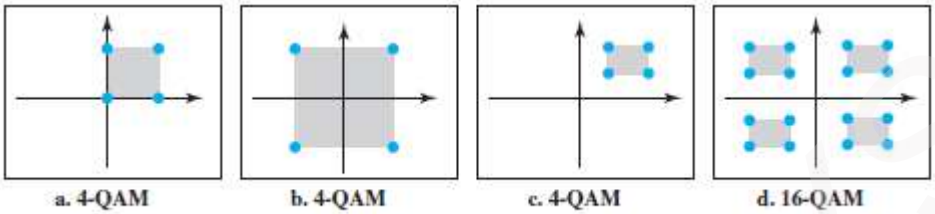


Figure 5.14 *Constellation diagrams for some QAMs*

2.3.5.1 Bandwidth for QAM

• The bandwidth is same as in ASK and PSK transmission.

• QAM has the same advantages as PSK over ASK.

MODULE 2(CONT.): BANDWIDTH UTILIZATION -- MULTIPLEXING AND SPREADING

2.4 MULTIPLEXING

- When bandwidth of a medium is greater than bandwidth needs of the devices, the link can be shared.
- *Multiplexing* allows simultaneous transmission of multiple signals across a single data-link (Fig 4.21).
- The traffic increases, as data/telecommunications use increases.
- We can accommodate this increase by
  - → adding individual links, each time a new channel is needed or
  - → installing higher-bandwidth links to carry multiple signals.
- Today's technology includes high-bandwidth media such as optical-fiber and satellite microwaves.
- Each has a bandwidth far in excess of that needed for the average transmission-signal.
- If the bandwidth of a link is greater than the bandwidth needs of the devices connected to it, the bandwidth is wasted.
- An efficient system maximizes the utilization of all resources; bandwidth is one of the most precious resources we have in data communications.

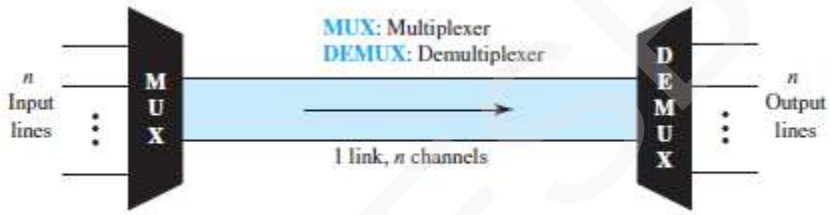


Figure 6.1 *Dividing a link into channels*

- In a multiplexed-system, 'n' lines share the bandwidth of one link.
- MUX combines transmission-streams from different input-lines into a single stream (many-to-one).
- At the receiving-end, that stream is fed into a demultiplexer (DEMUX).
- DEMUX
  - → separates the stream back into its component-transmissions (one-to-many) and
  - → directs the transmission-streams to different output-lines.
- Link vs. Channel:
  - 1) The link refers to the physical path.
  - 2) The channel refers to the portion of a link that carries a transmission between a given pair of lines. One link can have many channels.
- Three multiplexing techniques (Figure 6.2):
  - 1) Frequency-division multiplexing (FDM)
  - 2) Wavelength-division multiplexing (WDM) and
  - 3) Time-division multiplexing (TDM).
- The first two techniques are used for analog-signals.
  The third one technique is used for digital-signals.

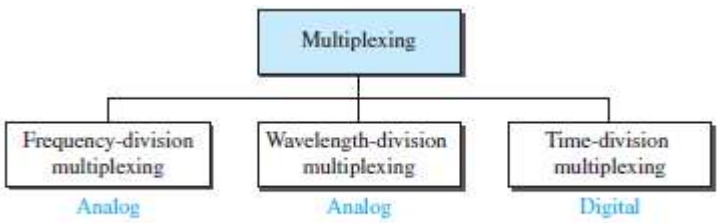


Figure 6.2 *Categories of multiplexing*

2.4.1 Frequency Division Multiplexing (FDM)

• FDM is an analog multiplexing technique that combines analog signals (Figure 6.3).
• FDM can be used when the bandwidth of a link is greater than the combined bandwidths of the signals to be transmitted. (Bandwidth measured in hertz).

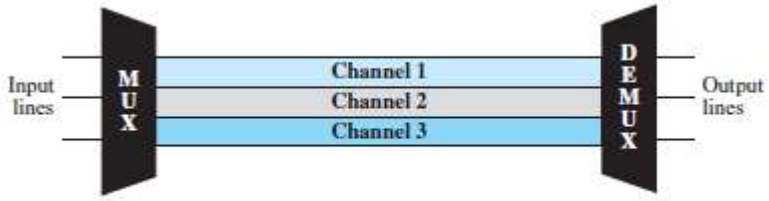


Figure 6.3 *Frequency-division multiplexing*

2.4.1.1 Multiplexing Process

• Here is how it works (Figure 6.4):
1) Each sending-device generates modulated-signals with different carrier-frequencies (f1, f2, & f3).
2) Then, these modulated-signals are combined into a single multiplexed-signal.
3) Finally, the multiplexed-signal is transported by the link.

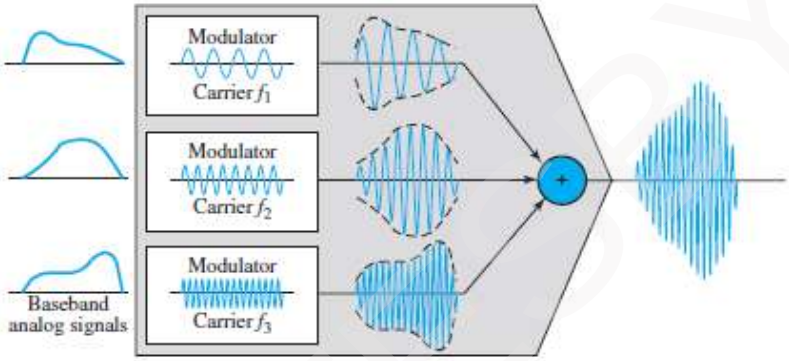


Figure 6.4 *FDM process*

• Carrier-frequencies are separated by sufficient bandwidth to accommodate the modulated-signal.
• Channels can be separated by strips of unused bandwidth called guard bands.
• Guard bands prevent signals from overlapping.
• In addition, carrier-frequencies must not interfere with the original data frequencies.
• Although FDM is considered as analog multiplexing technique, the sources can produce digital-signal.
• The digital-signal can be sampled, changed to analog-signal, and then multiplexed by using FDM.

2.4.1.2 Demultiplexing Process

• Here is how it works (Figure 6.5):
1) The demultiplexer uses filters to divide the multiplexed-signal into individual-signals.
2) Then, the individual signals are passed to a demodulator.
3) Finally, the demodulator
→ separates the individual signals from the carrier signals and
→ passes the individual signals to the output-lines.

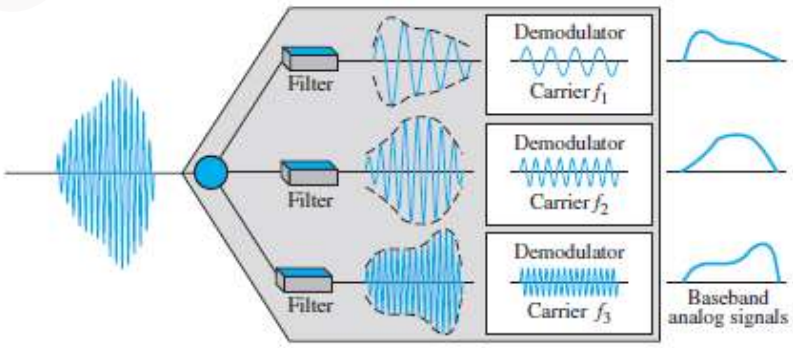


Figure 6.5 *FDM demultiplexing example*

Example 2.12

Assume that a voice channel occupies a bandwidth of 4 kHz. We need to combine three voice channels into a link with a bandwidth of 12 kHz, from 20 to 32 kHz. Show the configuration, using the frequency domain. Assume there are no guard bands.

Solution

We shift (modulate) each of the three voice channels to a different bandwidth, as shown in Figure 6.6. We use the 20- to 24-kHz bandwidth for the first channel, the 24- to 28-kHz bandwidth for the second channel, and the 28- to 32-kHz bandwidth for the third one. Then we combine them as shown in Figure 6.6.

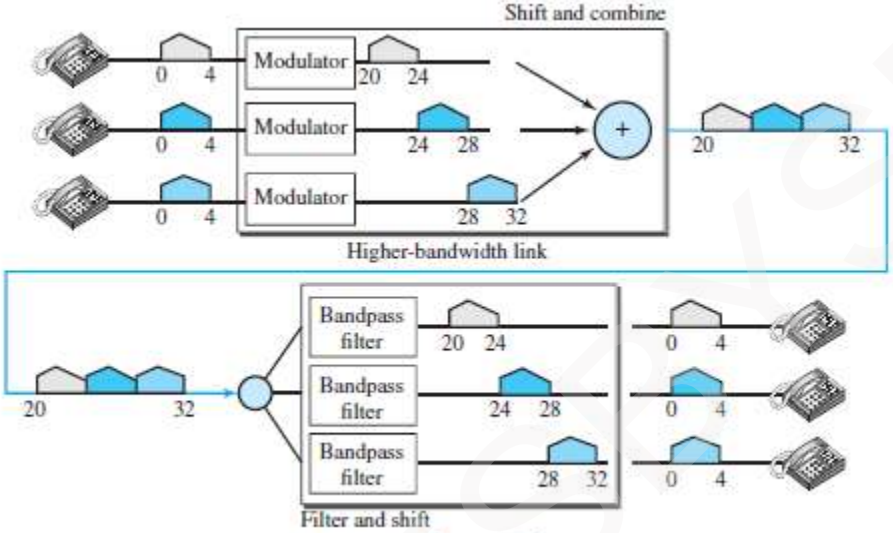


Figure 6.6

Example 2.13

Five channels, each with a 100-kHz bandwidth, are to be multiplexed together. What is the minimum bandwidth of the link if there is a need for a guard band of 10 kHz between the channels to prevent interference?

Solution

For five channels, we need at least four guard bands. This means that the required bandwidth is at least $5 \times 100 + 4 \times 10 = 540$ kHz, as shown in Figure 6.7.

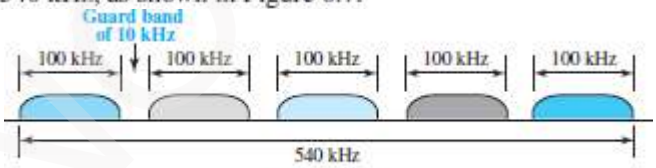


Figure 6.7

Example 2.14

Four data channels (digital), each transmitting at 1 Mbps, use a satellite channel of 1 MHz. Design an appropriate configuration, using FDM.

Solution

The satellite channel is analog. We divide it into four channels, each channel having a 250-kHz bandwidth. Each digital channel of 1 Mbps is modulated so that each 4 bits is modulated to 1 Hz. One solution is 16-QAM modulation. Figure 6.8 shows one possible configuration.

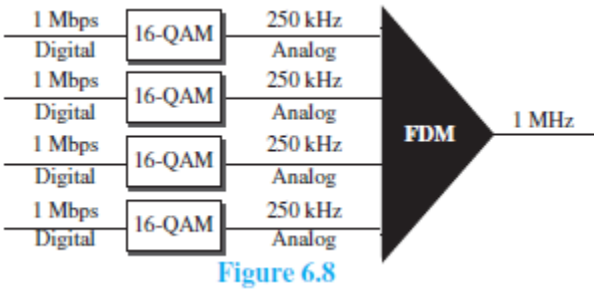


Figure 6.8

2.4.1.3 Applications of FDM

1) To maximize the efficiency of their infrastructure, Telephone-companies have traditionally multiplexed signals from lower-bandwidth lines onto higher-bandwidth lines.
2) A very common application of FDM is AM and FM radio broadcasting.
3) The first generation of cellular telephones (still in operation) also uses FDM.

2.4.1.4 Analog Carrier System

- To maximize the efficiency, telephone-companies have multiplexed-signals from lower-bandwidth lines onto higher-bandwidth lines.
- Many switched or leased lines are combined into bigger channels.
- For analog lines, FDM is used.
- One of these hierarchical systems used by AT&T is made up of (Figure 6.9):
  1) Groups
  2) Super groups
  3) Master groups, and
  4) Jumbo groups

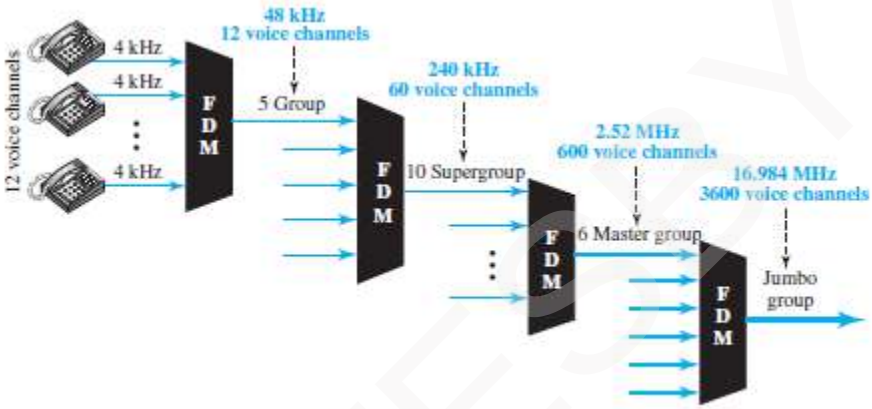


Figure 6.9 *Analog hierarchy*

1) Group: In the analog hierarchy, 12 voice channels are multiplexed onto a higher-bandwidth line to create a group.

- A group has 48 kHz of bandwidth and supports 12 voice channels.

2) Super Group: At the next level, up to five groups can be multiplexed to create a composite-signal called a supergroup.

- A supergroup has a bandwidth of 240 kHz and supports up to 60 voice channels.
- Supergroups can be made up of either five groups or 60 independent voice channels.

3) Master Groups: At the next level, 10 supergroups are multiplexed to create a master group.

- A master group must have 2.40 MHz of bandwidth, but the need for guard bands between the supergroups increases the necessary bandwidth to 2.52 MHz.
- Master groups support up to 600 voice channels.

4) Jumbo Group: Finally, six master groups can be combined into a jumbo group.

- A jumbo group must have 15.12 MHz (6 x 2.52 MHz) of bandwidth, but the need for guard bands b/w the master groups increases the necessary bandwidth to 16.984 MHz

Example 2.15

The Advanced Mobile Phone System (AMPS) uses two bands. The first band of 824 to 849 MHz is used for sending, and 869 to 894 MHz is used for receiving. Each user has a bandwidth of 30 kHz in each direction. The 3-kHz voice is modulated using FM, creating 30 kHz of modulated signal. How many people can use their cellular phones simultaneously?

Solution

Each band is 25 MHz. If we divide 25 MHz by 30 kHz, we get 833.33. In reality, the band is divided into 832 channels. Of these, 42 channels are used for control, which means only 790 channels are available for cellular phone users.

2.4.2 Wavelength Division Multiplexing (WDM)

- WDM is an analog multiplexing technique that combines analog signals (Figure 6.10).
- WDM is designed to use the high-data-rate capability of fiber optical-cable.
- The data-rate of optical-cable is higher than the data-rate of metallic-cable.
- Using an optical-cable for one single line wastes the available bandwidth.
- Multiplexing allows combining several lines into one line.
- WDM is same as FDM with 2 exceptions:
  1) Multiplexing & demultiplexing involve optical-signals transmitted through optical-cable.
  2) The frequencies are very high.

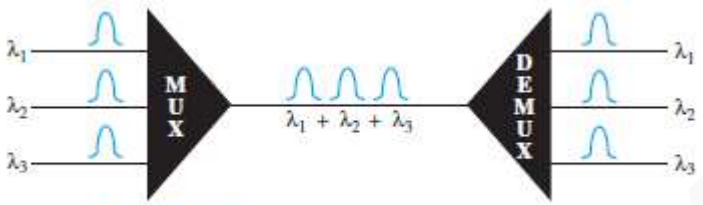


Figure 6.10 *Wavelength-division multiplexing*

- Here is how it works (Figure 6.11):
  - A multiplexer combines several narrow-bands of light into a wider-band of light.
  - A demultiplexer divides a wider-band of light into several narrow-bands of light.
  - A prism is used for combining and splitting of light sources
  - A prism bends a beam of light based on
    → angle of incidence and
    → frequency.

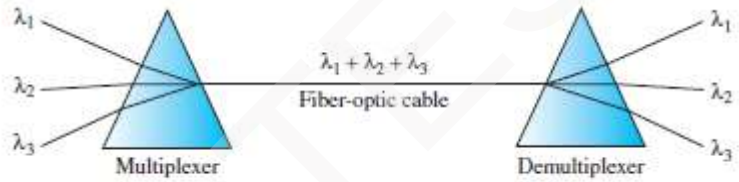


Figure 6.11 *Prisms in wavelength-division multiplexing and demultiplexing*

- Applications of WDM:
  1) SONET network: Multiple optical-fiber lines can be multiplexed and demultiplexed.
  2) Dense WDM (DWDM) can multiplex a very large number of channels by spacing channels very close to one another. DWDM achieves even greater efficiency

2.4.3 Time Division Multiplexing (TDM)

- TDM is a digital multiplexing technique that combines digital signals (Figure 6.12).
- TDM combines several low-rate channels into one high-rate one.
- FDM vs. TDM
    1) In FDM, a portion of the bandwidth is shared.
    2) In TDM, a portion of the time is shared.
- Each connection occupies a portion of time in the link.
- Several connections share the high bandwidth of a line.

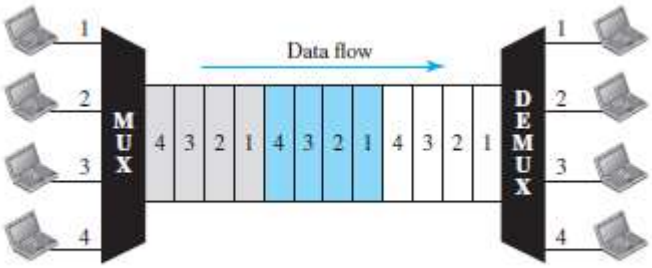


Figure 6.12 *TDM*

- As shown in Figure 6.12, the link is divided by time.
- Portions of signals 1, 2, 3, and 4 occupy the link sequentially.
- Digital-data from different sources are combined into one timeshared link.
- Although TDM is considered as digital multiplexing technique, the sources can produce analog-signal.
- The analog data can be sampled, changed to digital-data, and then multiplexed by using TDM.
- Two types of TDM:
    1) Synchronous and
    2) Statistical.

2.4.3.1 Synchronous TDM

2.4.3.1.1 Time Slots & Frames

- Each input-connection has an allotment in the output-connection even if it is not sending data.
- The data-flow of input-connection is divided into units (Figure 6.13).
- A unit can be 1 bit, 1 character, or 1 block of data.
- Each input-unit occupies one input-time-slot.
- Each input-unit
  - → becomes one output-unit and
  - → occupies one output-time-slot.
- However, duration of output-time-slot is 'n' times shorter than duration of input-time-slot.
- If an input-time-slot is T s, the output-time-slot is T/n s
  where n = No. of connections.
- In the output-connection, a unit has a shorter duration & therefore travels faster.

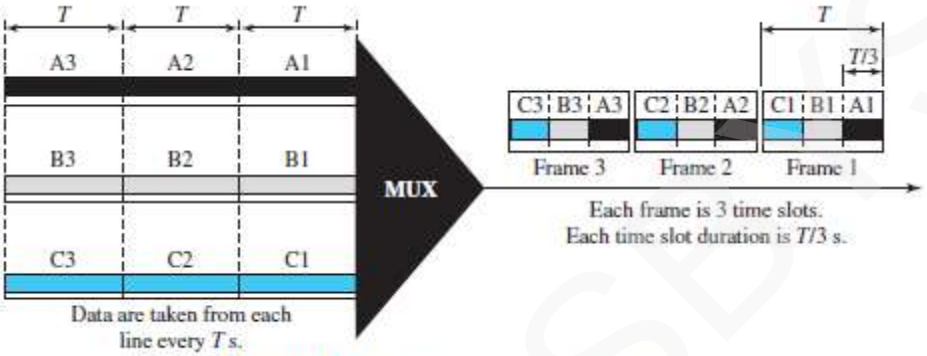


Figure 6.13 *Synchronous time-division multiplexing*

Figure 6.14

- In Figure 6.14, n = 3.
- A set of data-units from each input-connection is grouped into a frame.
- For example:
  If there are 3 connections, a frame is divided into 3 time-slots.
  One slot is allocated for each data-unit.
  One data-unit is used for each input-line.

Example 2.16

In Figure 6.13, the data rate for each input connection is 1 kbps. If 1 bit at a time is multiplexed (a unit is 1 bit), what is the duration of

1. each input slot,
2. each output slot, and
3. each frame?

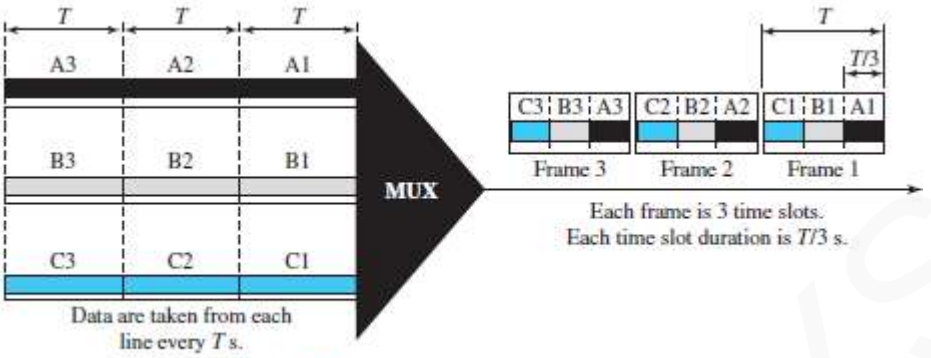


Figure 6.13 *Synchronous time-division multiplexing*

Solution

We can answer the questions as follows:

1. The data rate of each input connection is 1 kbps. This means that the bit duration is 1/1000 s or 1 ms. The duration of the input time slot is 1 ms (same as bit duration).
2. The duration of each output time slot is one-third of the input time slot. This means that the duration of the output time slot is 1/3 ms.
3. Each frame carries three output time slots. So the duration of a frame is 3 × 1/3 ms, or 1 ms. The duration of a frame is the same as the duration of an input unit.

Example 2.17

Figure 6.14 shows synchronous TDM with a data stream for each input and one data stream for the output. The unit of data is 1 bit. Find (1) the input bit duration, (2) the output bit duration, (3) the output bit rate, and (4) the output frame rate.

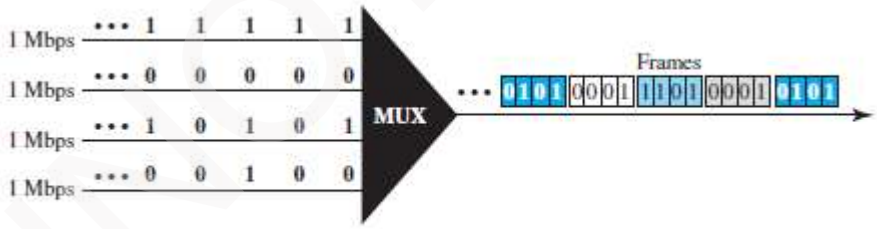


Figure 6.14

Solution

We can answer the questions as follows:

1. The input bit duration is the inverse of the bit rate: 1/1 Mbps = 1 µs.
2. The output bit duration is one-fourth of the input bit duration, or 1/4 µs.
3. The output bit rate is the inverse of the output bit duration, or 1/4 µs, or 4 Mbps. This can also be deduced from the fact that the output rate is 4 times as fast as any input rate; so the output rate = 4 × 1 Mbps = 4 Mbps.
4. The frame rate is always the same as any input rate. So the frame rate is 1,000,000 frames per second. Because we are sending 4 bits in each frame, we can verify the result of the previous question by multiplying the frame rate by the number of bits per frame.

Example 2.18

Four 1-kbps connections are multiplexed together. A unit is 1 bit. Find (1) the duration of 1 bit before multiplexing, (2) the transmission rate of the link, (3) the duration of a time slot, and (4) the duration of a frame.

Solution

We can answer the questions as follows:

1. The duration of 1 bit before multiplexing is 1/1 kbps, or 0.001 s (1 ms).
2. The rate of the link is 4 times the rate of a connection, or 4 kbps.
3. The duration of each time slot is one-fourth of the duration of each bit before multiplexing, or 1/4 ms or 250 μs. Note that we can also calculate this from the data rate of the link, 4 kbps. The bit duration is the inverse of the data rate, or 1/4 kbps or 250 μs.
4. The duration of a frame is always the same as the duration of a unit before multiplexing, or 1 ms. We can also calculate this in another way. Each frame in this case has four time slots. So the duration of a frame is 4 times 250 μs, or 1 ms.

2.4.3.1.2 Interleaving

- TDM can be seen as 2 fast-rotating switches (Figure 6.15):
  1) First switch on the multiplexing-side and
  2) Second switch on the demultiplexing-side.
- The switches are synchronized and rotate at the same speed, but in opposite directions.
  1) On the multiplexing-side (Figure 6.16)
  ➢ As the switch opens in front of a connection, that connection has the opportunity to send a unit onto the path. This process is called *interleaving*.
  2) On the demultiplexing-side
  ➢ As the switch opens in front of a connection, that connection has the opportunity to receive a unit from the path.

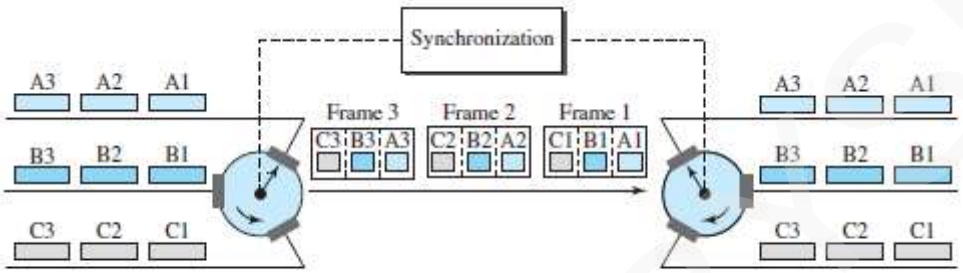


Figure 6.15 *Interleaving*

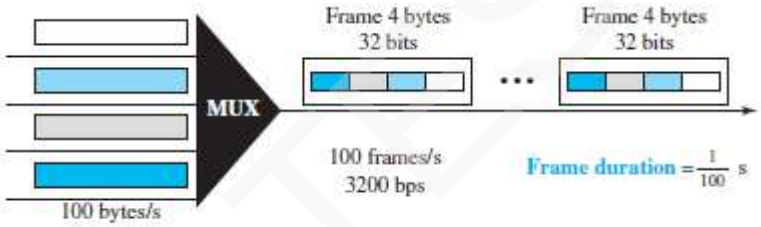


Figure 6.16

Example 2.19

Four channels are multiplexed using TDM. If each channel sends 100 bytes/s and we multiplex 1 byte per channel, show the frame traveling on the link, the size of the frame, the duration of a frame, the frame rate, and the bit rate for the link.

Solution

The multiplexer is shown in Figure 6.16. Each frame carries 1 byte from each channel; the size of each frame, therefore, is 4 bytes, or 32 bits. Because each channel is sending 100 bytes/s and a frame carries 1 byte from each channel, the frame rate must be 100 frames per second. The duration of a frame is therefore 1/100 s. The link is carrying 100 frames per second, and since each frame contains 32 bits, the bit rate is 100×32, or 3200 bps. This is actually 4 times the bit rate of each channel, which is $100 \times 8 = 800$ bps.

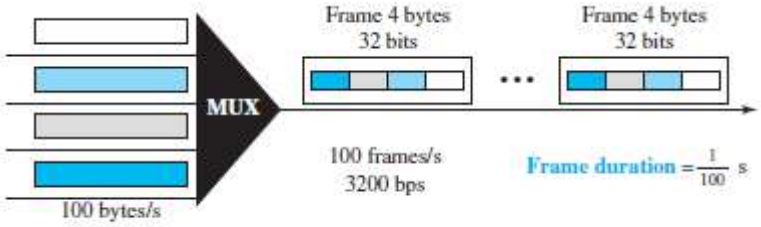


Figure 6.16

Example 2.20

A multiplexer combines four 100-kbps channels using a time slot of 2 bits. Show the output with four arbitrary inputs. What is the frame rate? What is the frame duration? What is the bit rate? What is the bit duration?

Solution

Figure 6.17 shows the output for four arbitrary inputs. The link carries 50,000 frames per second since each frame contains 2 bits per channel. The frame duration is therefore 1/50,000 s or 20 µs. The frame rate is 50,000 frames per second, and each frame carries 8 bits; the bit rate is 50,000 × 8 = 400,000 bits or 400 kbps. The bit duration is 1/400,000 s, or 2.5 µs. Note that the frame duration is 8 times the bit duration because each frame is carrying 8 bits.

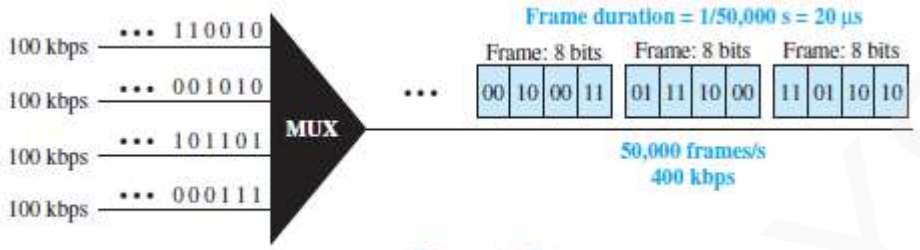


Figure 6.17

2.4.3.1.3 Empty Slots

- Problem: Synchronous TDM is not efficient.
  For example: If a source does not have data to send, the corresponding slot in the output-frame is empty.

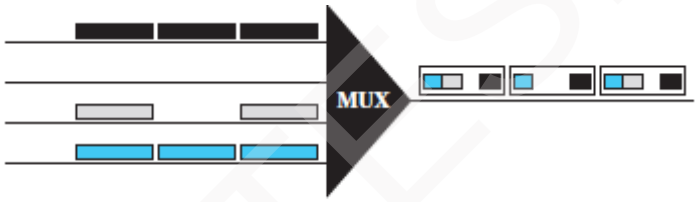


Figure 6.18 *Empty slots*

- As shown in Figure 6.18,
  Second input-line has no data to send
  Third input-line has discontinuous data.
- The first output-frame has 3 slots filled.
  The second frame has 2 slots filled.
  The third frame has 3 slots filled.
  No frame is full.
- Solution: Statistical TDM can improve the efficiency by removing the empty slots from the frame.

2.4.3.1.4 Data Rate Management

- Problem in TDM: How to handle differences in the input data-rates?
- If data-rates are not the same, three strategies can be used.
- Three different strategies: 1) Multilevel multiplexing 2) Multiple-slot allocation and 3) Pulse stuffing

1) Multilevel Multiplexing

➢ This technique is used when the data-rate of an input-line is a multiple of others.

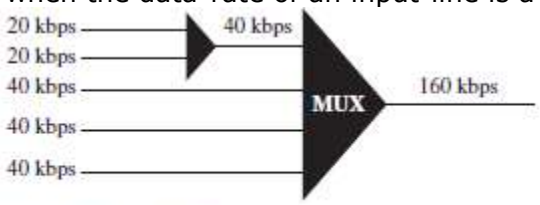


Figure 6.19 *Multilevel multiplexing*

➢ For example:
- ➢ As shown in Figure 6.19, we have 2 inputs of 20 kbps and 3 inputs of 40 kbps.
- ➢ The first 2 input-lines can be multiplexed to provide a data-rate of 40 kbps.

2) Multiple Slot Allocation

➢ Sometimes it is more efficient to allot more than 1 slot in a frame to a single input-line.

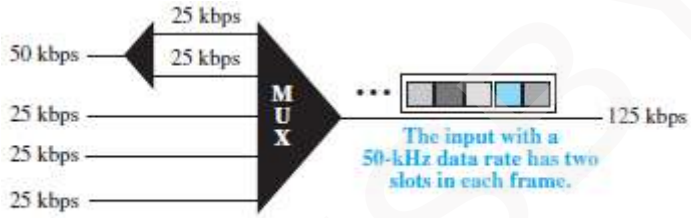


Figure 6.20 *Multiple-slot multiplexing*

➢ For example:

Data-rate of multiple input-lines can be data-rate of one input-line.

➢ As shown in Figure 6.20, the input-line with a 50-kbps data-rate can be given 2 slots in the output-line.

➢ In first input line, serial-to-parallel converter is used. The converter creates two 25 kbps input lines out of one 50 kbps input line.

3) Pulse Stuffing

➢ Sometimes the bit-rates of sources are not multiple integers of each other. .'. above 2 techniques cannot be used.

➢ Solution:

→ Make the highest input data-rate the dominant data-rate.
→ Then, add dummy bits to the input-lines with lower rates.
→ This will increase data rates of input-line.
→ This technique is called pulse stuffing, bit padding, or bit stuffing.

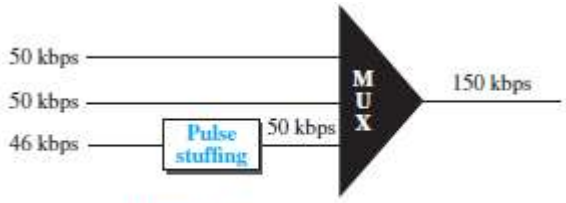


Figure 6.21 *Pulse stuffing*

➢ As shown in Figure 6.21, the input-line with a lower data-rate = 46kbps is pulse-stuffed to increase the data-rate to 50 kbps.

➢ Now, multiplexing can take place.

2.4.3.1.5 Frame Synchronizing

• Problem: Synchronization between the multiplexer and demultiplexer is a major issue.

If the multiplexer and the demultiplexer are not synchronized, a bit belonging to one channel may be received by the wrong channel.

Solution: Usually, one or more synchronization-bits are added to the beginning of each frame. These bits are called *framing-bits*.

The framing-bits follow a pattern (frame-to-frame) that allows multiplexer and demultiplexer to synchronize.

As shown in Figure 6.22, the synchronization-information

→ consists of 1 bit per frame and

→ alternates between 0 & 1.

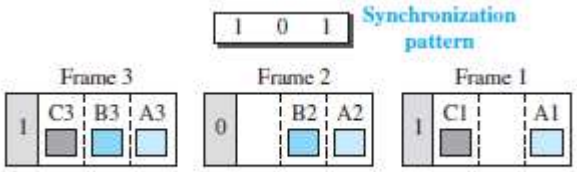


Figure 6.22 *Framing bits*

Example 2.21

We have four sources, each creating 250 characters per second. If the interleaved unit is a character and 1 synchronizing bit is added to each frame, find (1) the data rate of each source, (2) the duration of each character in each source, (3) the frame rate, (4) the duration of each frame, (5) the number of bits in each frame, and (6) the data rate of the link.

Solution

We can answer the questions as follows:

1. The data rate of each source is $250 \times 8 = 2000$ bps $= 2$ kbps.
2. Each source sends 250 characters per second; therefore, the duration of a character is 1/250 s, or 4 ms.
3. Each frame has one character from each source, which means the link needs to send 250 frames per second to keep the transmission rate of each source.
4. The duration of each frame is 1/250 s, or 4 ms. Note that the duration of each frame is the same as the duration of each character coming from each source.
5. Each frame carries 4 characters and 1 extra synchronizing bit. This means that each frame is $4 \times 8 + 1 = 33$ bits.
6. The link sends 250 frames per second, and each frame contains 33 bits. This means that the data rate of the link is 250×33, or 8250 bps. Note that the bit rate of the link is greater than the combined bit rates of the four channels. If we add the bit rates of four channels, we get 8000 bps. Because 250 frames are traveling per second and each contains 1 extra bit for synchronizing, we need to add 250 to the sum to get 8250 bps.

Example 2.22

Two channels, one with a bit rate of 100 kbps and another with a bit rate of 200 kbps, are to be multiplexed. How this can be achieved? What is the frame rate? What is the frame duration? What is the bit rate of the link?

Solution

We can allocate one slot to the first channel and two slots to the second channel. Each frame carries 3 bits. The frame rate is 100,000 frames per second because it carries 1 bit from the first channel. The frame duration is 1/100,000 s, or 10 ms. The bit rate is 100,000 frames/s $\times$ 3 bits per frame, or 300 kbps. Note that because each frame carries 1 bit from the first channel, the bit rate for the first channel is preserved. The bit rate for the second channel is also preserved because each frame carries 2 bits from the second channel.

2.4.3.2 Statistical TDM

• Problem: Synchronous TDM is not efficient.
For ex: If a source does not have data to send, the corresponding slot in the output-frame is empty.
Solution: Use statistical TDM.
Slots are dynamically allocated to improve bandwidth-efficiency.
Only when an input-line has data to send, the input-line is given a slot in the output-frame.
• The number of slots in each frame is less than the number of input-lines.
• The multiplexer checks each input-line in round robin fashion.
If the line has data to send;
Then, multiplexer allocates a slot for an input-line;
Otherwise, multiplexer skips the line and checks the next line.

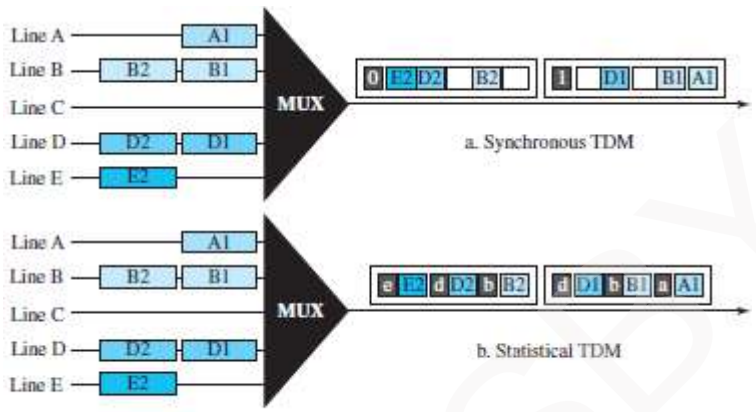


Figure 6.26 *TDM slot comparison*

• In synchronous TDM (Figure 6.26a), some slots are empty because the corresponding line does not have data to send.
• In statistical TDM (Figure 6.26b), no slot is left empty.

1) Addressing

Synchronous TDM	Statistical TDM
An output-slot needs to carry only data of the destination (Figure 6.26a).	An output-slot needs to carry both data & address of the destination (Figure 6.26b).
There is no need for addressing. Synchronization and pre-assigned relationships between the inputs and outputs serve as an address.	There is no fixed relationship between the inputs and outputs because there are no pre-assigned or reserved slots. We need to include the address of the receiver inside each slot to show where it is to be delivered.

2) Slot Size

➢ Usually, a block of data is many bytes while the address is just a few bytes.
➢ A slot carries both data and address.
➢ Therefore, address-size must be very small when compared to data-size.
This results in efficient transmission.
➢ For example:
It will be inefficient to send 1 bit per slot as data, when the address is 3 bits.
This means an overhead of 300%.

3) No Synchronization Bit

➢ In statistical TDM, the frames need not be synchronized, so synchronization-bits are not needed.

4) Bandwidth

➢ Normally, the capacity of the link is less than the sum of the capacities of each channel.
➢ The designers define the capacity of the link based on the statistics of the load for each channel.

2.5 SPREAD SPECTRUM

• Spread-spectrum is used in wireless applications (Figure 6.27).
• In wireless applications, all stations use air (or a vacuum) as the medium for communication.
• Goal: Stations must be able to share the air medium without interception by an attacker.

Solution: Spread-spectrum techniques add redundancy i.e. they spread the original spectrum needed for each station.

• If the required bandwidth for each station is B, spread-spectrum expands it to B_{ss} such that B_{ss}>>B.
• The expanded-bandwidth allows the source to place its message in a protective envelope for a more secure transmission.

(An analogy is the sending of a delicate, expensive gift. We can insert the gift in a special box to prevent it from being damaged during transportation).

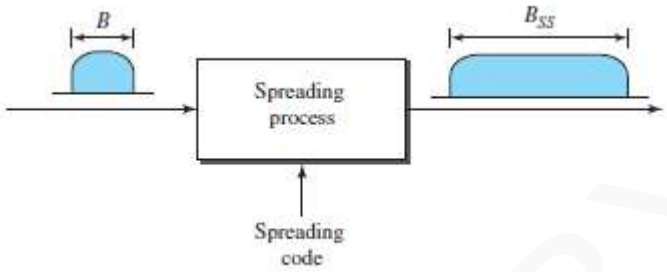


Figure 6.27 *Spread spectrum*

• Spread-spectrum achieves its goal through 2 principles:

1) The bandwidth allocated to each station needs to be, by far, larger than what is needed. This allows redundancy.

2) The spreading process must occur after the signal is created by the source.

• After the signal is created by the source, the spreading process
→ uses a spreading-code and
→ spreads the bandwidth.
• The spreading-code is a series of numbers that look random, but are actually a pattern.
• Two types of spread-spectrum:
1) Frequency hopping spread-spectrum (FHSS) and
2) Direct sequence spread-spectrum (DSSS).

2.5.1 Frequency Hopping Spread Spectrum (FHSS)

- This technique uses 'M' different carrier-frequencies that are modulated by the source-signal.
- At one moment, the signal modulates one carrier-frequency.
  At the next moment, the signal modulates another carrier-frequency.
- Although the modulation is done using one carrier-frequency at a time, 'M' frequencies are used in the long run.
- The bandwidth occupied by a source is given by

$$B_{FHSS} >> B$$

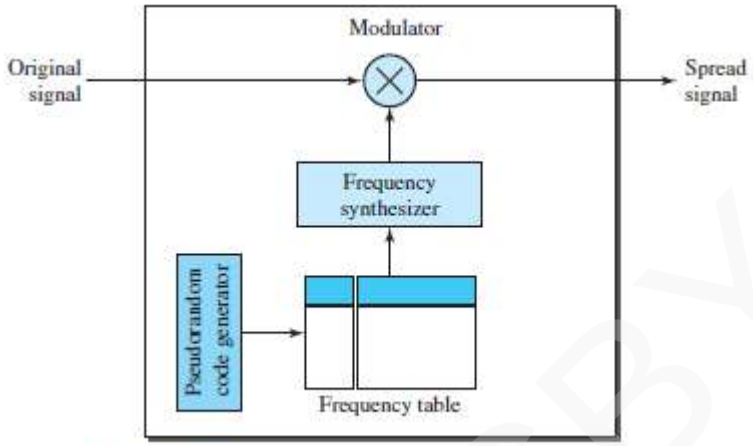


Figure 6.28 *Frequency hopping spread spectrum (FHSS)*

- As shown in Figure 6.28.
  - A pseudorandom code generator (PN) creates a k-bit pattern for every hopping period T_h.
  - The frequency-table
    → uses the pattern to find the frequency to be used for this hopping period and
    → passes the frequency to the frequency-synthesizer.
  - The frequency-synthesizer creates a carrier-signal of that frequency.
  - The source-signal modulates the carrier-signal.

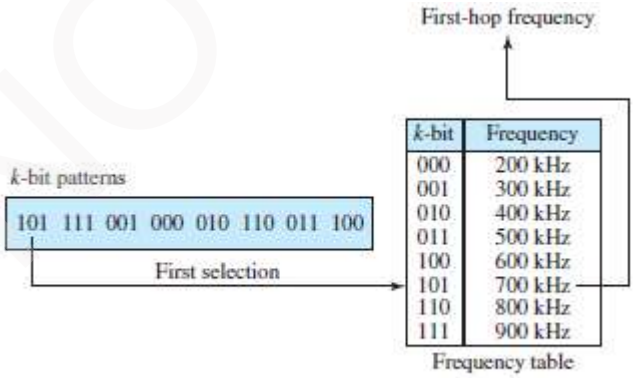


Figure 6.29 *Frequency selection in FHSS*

- As shown in Figure 6.29, assume we have 8 hopping frequencies.
  - Here, M = 8 and k = 3.
  - The pseudorandom code generator will create 8 different 3-bit patterns.
  - These are mapped to 8 different frequencies in the frequency table (see Figure 6.29).
  - The pattern for this station is 101, 111, 001, 000, 010, 111 & 100.
    1) At hopping-period 1, the pattern is 101.
       The frequency selected is 700 kHz; the source-signal modulates this carrier-frequency.
    2) At hopping-period 2, the pattern is 111.
       The frequency selected is 900 kHz; the source-signal modulates this carrier-frequency.

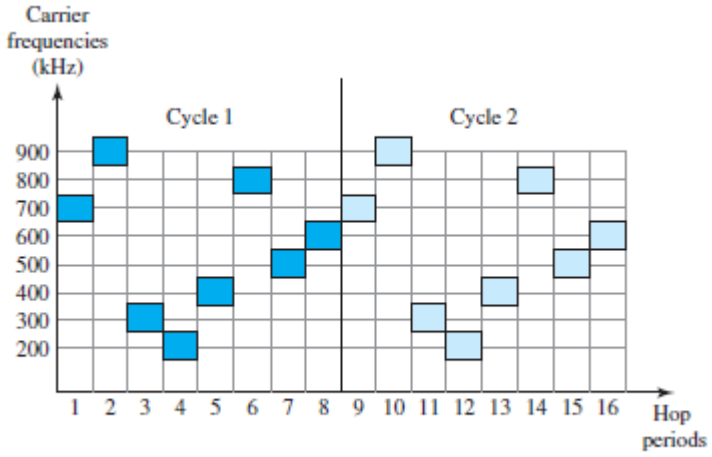


Figure 6.30 *FHSS cycles*

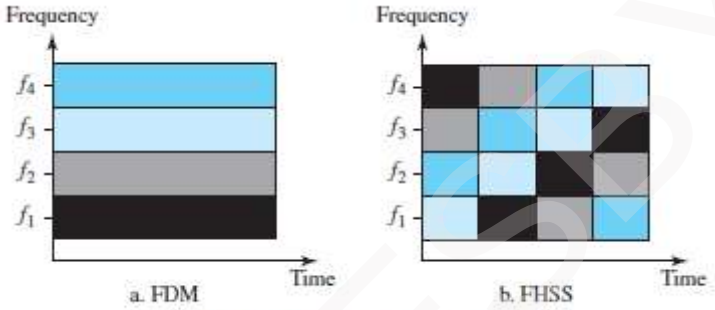


Figure 6.31 *Bandwidth sharing*

- If there are many k-bit patterns & the hopping period is short, a sender & receiver can have privacy.
  If an attacker tries to intercept the transmitted signal, he can only access a small piece of data because he does not know the spreading sequence to quickly adapt himself to the next hop.
- The scheme has also an anti-jamming effect.
  A malicious sender may be able to send noise to jam the signal for one hopping period (randomly), but not for the whole period.

2.5.1.1 Bandwidth Sharing

- If the number of hopping frequencies is M, we can multiplex M channels into one by using the same B_{ss} bandwidth.
- This is possible because
  1) A station uses just one frequency in each hopping period.
  2) Other M-1 stations uses other M-1 frequencies.
- In other words, M different stations can use the same B_{ss} if a multiple FSK (MFSK) is used.

2.5.2 Direct Sequence Spread Spectrum (DSSS)

- This technique expands the bandwidth of the original signal.
- Each data-bit is replaced with 'n' bits using a spreading-code.
- Each bit is assigned a code of 'n' bits called chips.
- The chip-rate is 'n' times that of the data-bit (Figure 6.32).

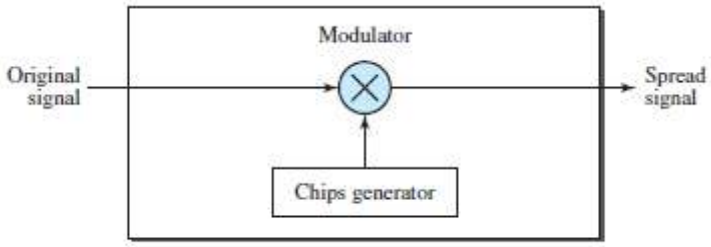


Figure 6.32 *DSSS*

- For example (Figure 6.33):
  - Consider the Barker sequence used in a wireless LAN. Here n =11.
  - Assume that the original signal and the chips in the chip-generator use polar NRZ encoding.
  - The spreading-code is 11 chips having the pattern 10110111000.
  - If the original signal-rate is N, the rate of the spread signal is 1/N.
  - This means that the required bandwidth for the spread signal is 11 times larger than the bandwidth of the original signal.

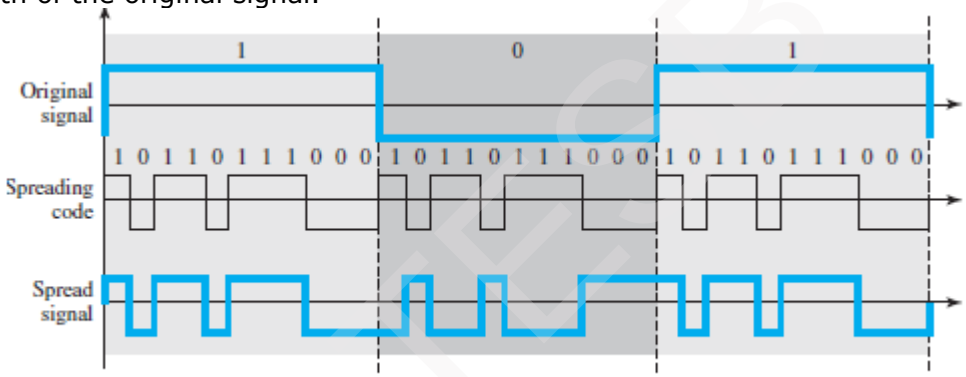


Figure 6.33 *DSSS example*

- The spread signal can provide privacy if the attacker does not know the code.
- It can also provide immunity against interference if each station uses a different code.

2.5.2.1 Bandwidth Sharing

- Can we share a bandwidth in DSSS?
- The answer is no and yes.

1) If we use a spreading-code that spreads signals that cannot be combined and separated, we cannot share a bandwidth.

- For example:

  Some wireless LANs use DSSS and the spread bandwidth cannot be shared.

2) If we use a special spreading-code that spreads signals that can be combined and separated, we can share a bandwidth.

- For example:

  Cellular telephony uses DSSS and the spread bandwidth is shared b/w several users.

MODULE 2(CONT.): SWITCHING

2.6 SWITCHING

• A network is a set of connected-devices.
• Problem: Whenever we have multiple-devices, we have the problem of how to connect them to make one-to-one communication possible.
• Solution: Use Switching.
• A switched-network consists of a series of interlinked-nodes, called switches.
• Switches are devices capable of creating temporary connections between two or more devices.
• In a switched-network,
    1) Some nodes are connected to the end-systems (For example: PC or TP).
    2) Some nodes are used only for routing.

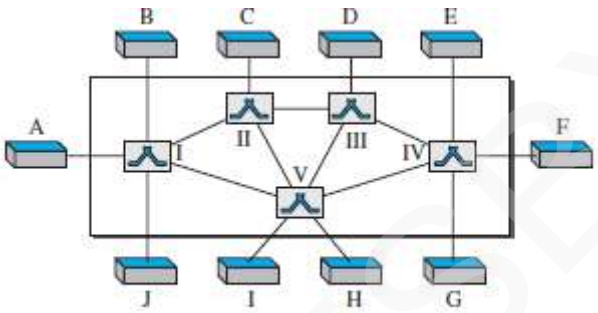


Figure 8.1 *Switched network*

• As shown in Figure 8.1,
    1) The end-systems are labeled A, B, C, D, and so on.
    2) The switches are labeled I, II, III, IV, and V. Each switch is connected to multiple links.

2.6.1 Three Methods of Switching

• Three methods of Switching are (Figure 8.2):
    1) Circuit Switching
    2) Packet Switching and
    3) Message Switching.
• The first two are commonly used today.
• The third has been phased out in general communications but still has networking applications.
• Packet switching can further be divided into two subcategories—virtual circuit approach and datagram approach

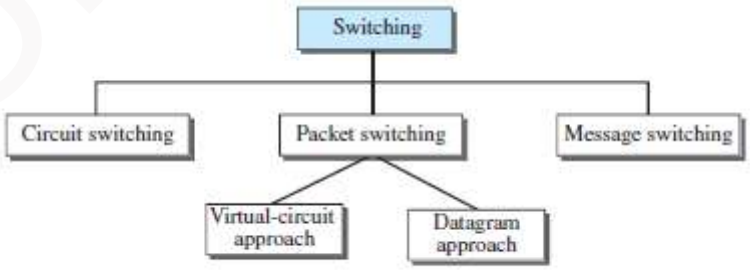


Figure 8.2 *Taxonomy of switched networks*

2.6.2 Switching and TCP/IP Layers

• Switching can happen at several layers of the TCP/IP protocol suite.

1) Switching at Physical Layer

- At the physical layer, we can have only circuit switching.
- There are no packets exchanged at the physical layer.
- The switches at the physical layer allow signals to travel in one path or another.

2) Switching at Data-Link Layer

- At the data-link layer, we can have packet switching.
- However, the term packet in this case means frames or cells.
- Packet switching at the data-link layer is normally done using a virtual-circuit approach.

3) Switching at Network Layer

- At the network layer, we can have packet switching.
- In this case, either a virtual-circuit approach or a datagram approach can be used.
- Currently the Internet uses a datagram approach, but the tendency is to move to a virtual-circuit approach.

4) Switching at Application Layer

- At the application layer, we can have only message switching.
- The communication at the application layer occurs by exchanging messages.
- Conceptually, we can say that communication using e-mail is a kind of message-switched communication, but we do not see any network that actually can be called a message-switched network.

2.7 CIRCUIT SWITCHED NETWORK

• This is similar to telephone system.
• Fixed path (connection) is established between a source and a destination prior to the transfer of packets.
• A circuit-switched-network consists of a set of switches connected by physical-links (Figure 8.3).
• A connection between 2 stations is a dedicated-path made of one or more links.
• However, each connection uses only one dedicated-channel on each link.
• Normally, each link is divided into 'n' channels by using FDM or TDM.
• The resources need to be reserved during the setup phase.
The resources remain dedicated for the entire duration of data transfer until the teardown phase.

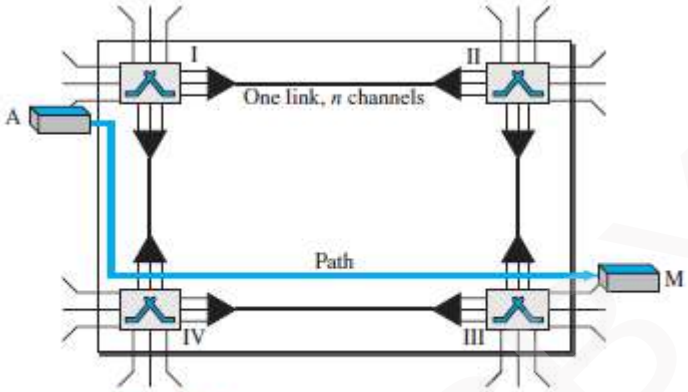


Figure 8.3 *A trivial circuit-switched network*

• The virtual-circuit setup procedure
→ first determines a path through the network &
→ sets parameters in the switches by exchanging connect-request & connect-confirm messages
• If a switch does not have enough resources to set up a virtual circuit, the switch responds with a connect-reject message and the setup procedure fails (Figure 7.15).
• A connection-release procedure may also be required to terminate the connection.

2.7.1 Three Phases

• The communication requires 3 phases: 1) Connection-setup
2) Data-transfer
3) Connection teardown.

1) Setup Phase

➢ Before the 2 parties can communicate, a dedicated-circuit needs to be established.
➢ Normally, the end-systems are connected through dedicated-lines to the switches.
So, connection-setup means creating dedicated-channels between the switches.
➢ For ex: Assume system-A needs to connect to system-M. For this, following events occur:
i) System-A sends a setup-request to switch-I.
ii) Switch-I finds a channel between itself and switch-IV that can be dedicated for this purpose.
iii) Switch-I then sends the request to switch-IV, which finds a dedicated-channel between itself and switch-III.
iv) Switch-III informs system-M of system-A's intention at this time.
v) Finally, an acknowledgment from system-M needs to be sent in the opposite direction to system-A.
➢ Only after system A receives this acknowledgment is the connection established.

2) Data Transfer Phase

➢ After the establishment of the dedicated-circuit (channels), the two parties can transfer data.

3) Teardown Phase

➢ When one of the parties needs to disconnect, a signal is sent to each switch to release the resources.

2.7.2 Efficiency

• Circuit-switched-networks are inefficient when compared to other two types of networks because
1) Resources are allocated during the entire duration of the connection.
2) These resources are unavailable to other connections.

2.7.3 Delay

• Circuit-switched-networks have minimum delay when compared to other two types of networks
• During data-transfer,
1) The data are not delayed at each switch.
2) The resources are allocated for the duration of the connection.

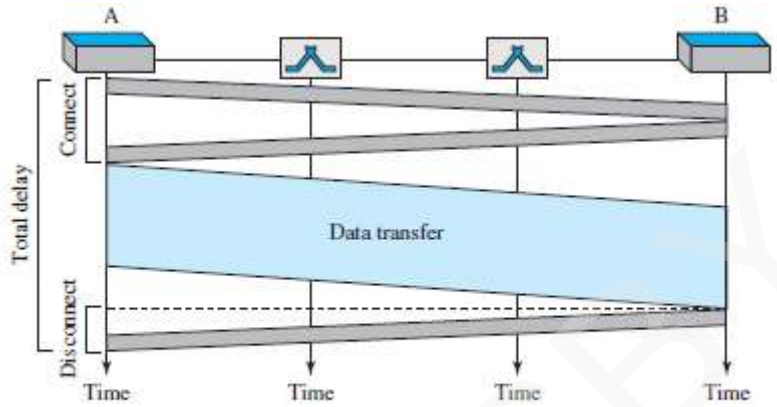


Figure 8.6 *Delay in a circuit-switched network*

- As in the above figure, there is no waiting time at each switch.
- The total delay is the time needed to
  1) Create the connection
  2) Transfer-data and
  3) Disconnect the circuit.
- The delay caused by the setup is the sum of 4 parts:
  1) The propagation time of the source-computer request.
  2) The request signal transfer time.
  3) The propagation time of the acknowledgment from the destination computer.
  4) The signal transfer time of the acknowledgment.
- The delay due to data-transfer is the sum of 2 parts:
  1) The propagation time.
  2) Data-transfer time which can be very long.

2.8 PACKET SWITCHED NETWORK

- The message is divided into packets of fixed or variable size.
- The packet-size is determined by
  → network and
  → governing protocol.
- There is no resource reservation; resources are allocated on-demand.

2.8.1 Datagram Networks

- This is analogous to postal system.
- Each packet is routed independently through the network.
- Each packet has a header that contains source and destination addresses.
- Each switch examines the header to determine the next hop in the path to the destination.
- If the transmission line is busy
  then the packet is placed in the queue until the line becomes free.
- Packets are referred to as datagrams.
- Datagram switching is normally done at the network layer.
- In Internet, switching is done by using the datagram switching.
- Advantage:
  1) High utilization of transmission-line can be achieved by sharing among multiple packets.
- Disadvantages:
  1) Packets may arrive out-of-order, and re-sequencing may be required at the destination
  2) Loss of packets may occur when a switch has insufficient buffer

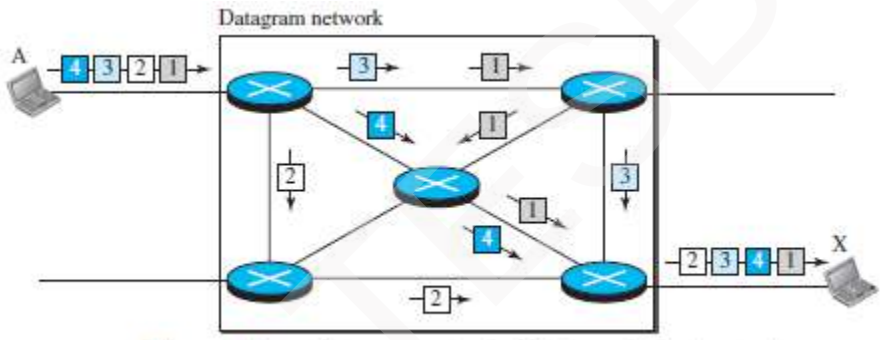


Figure 8.7 *A datagram network with four switches (routers)*

- The Figure 8.7 shows how the 4 packets are transferred from station-A to station-X.
- The switches are referred to as routers.
- All four packets (or datagrams) belong to the same message, but may travel different paths to reach their destination.
- This is so because the links may be involved in carrying packets from other sources and do not have the necessary bandwidth available to carry all the packets from A to X.
- This approach can cause the datagrams of a transmission to arrive at their destination out-of-order with different delays between the packets.
- Packets may also be lost or dropped because of a lack-of-resources.
- It is the responsibility of an upper-layer protocol to
  → reorder the datagrams or
  → ask for lost datagrams.
- The datagram-networks are referred to as connectionless networks. This is because
  1) The switch does not keep information about the connection state.
  2) There are no setup or teardown phases.
  3) Each packet is treated the same by a switch regardless of its source or destination.

2.8.1.1 Routing Table

• Each switch has a routing-table which is based on the destination-address.
• The routing-tables are dynamic & updated periodically.
• The destination-addresses and the corresponding forwarding output-ports are recorded in the tables.

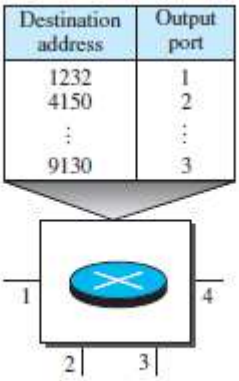


Figure 8.8 *Routing table in a datagram network*

2.8.1.1.1 Destination Address

• Every packet carries a header that contains the destination-address of the packet.
• When the switch receives the packet,
→ This destination-address is examined.
→ The routing-table is consulted to find the corresponding port through which the packet should be forwarded.
• The destination address in the header of a packet in remains the same during the entire journey of the packet.

2.8.1.1.2 Efficiency

• Datagram-networks are more efficient when compared to circuit-switched-network. This is because
1) Resources are allocated only when there are packets to be transferred.
2) If a source sends a packet and there is a delay of a few minutes before another packet can be sent, the resources can be re-allocated during these minutes for other packets from other sources.

2.8.1.1.3 Delay

• Datagram-networks may have greater delay when compared to circuit-switched-network. This is because
1) Each packet may experience a wait at a switch before it is forwarded.
2) Since not all packets in a message necessarily travel through the same switches, the delay is not uniform for the packets of a message.

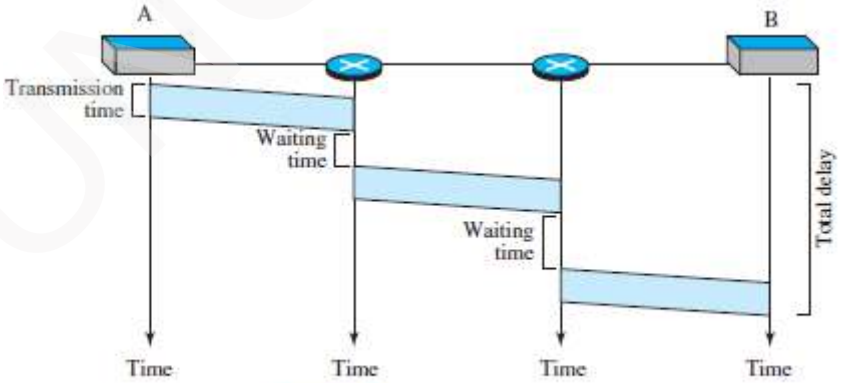


Figure 8.9 *Delay in a datagram network*

➢ The Figure 8.9 gives an example of delay for one single packet.
➢ The packet travels through two switches.
➢ There are three transmission times (3T), three propagation delays (slopes 3t of the lines), and two waiting times (W1+ W2).

$$\text{Total delay} = 3T + 3\tau + w_1 + w_2$$

2.8.2 Virtual Circuit Network (VCN)

- This is similar to telephone system.
- A virtual-circuit network is a combination of circuit-switched-network and datagram-network.
- Five characteristics of VCN:

1) As in a circuit-switched-network, there are setup & teardown phases in addition to the data transfer phase.

2) As in a circuit-switched-network, resources can be allocated during the setup phase.
As in a datagram-network, resources can also be allocated on-demand.

3) As in a datagram-network, data is divided into packets.
Each packet carries an address in the header.
However, the address in the header has local jurisdiction, not end-to-end jurisdiction.

4) As in a circuit-switched-network, all packets follow the same path established during the connection.

5) A virtual-circuit network is implemented in the data link layer.
A circuit-switched-network is implemented in the physical layer.
A datagram-network is implemented in the network layer.

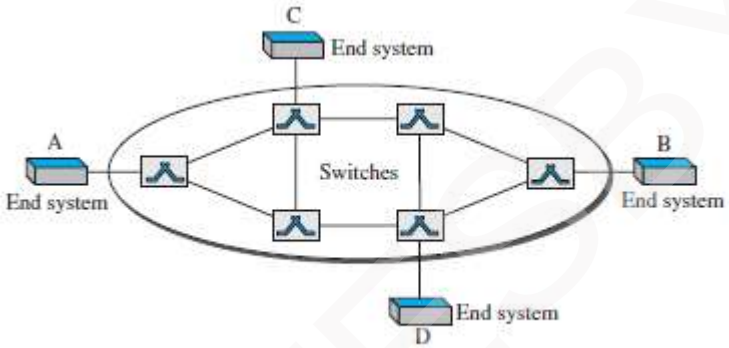


Figure 8.10 *Virtual-circuit network*

- The Figure 8.10 is an example of a virtual-circuit network.
- The network has switches that allow traffic from sources to destinations.
- A source or destination can be a computer, packet switch, bridge, or any other device that connects other networks.

2.8.2.1 Addressing

- Two types of addressing: 1) Global and 2) Local (virtual-circuit identifier).

1) Global Addressing

- A source or a destination needs to have a global address.
- Global address is an address that can be unique in the scope of the network or internationally if the network is part of an international network.

2) Virtual Circuit Identifier

- The identifier used for data-transfer is called the virtual-circuit identifier (VCI).
- A VCI, unlike a global address, is a small number that has only switch scope.
- VCI is used by a frame between two switches.
- When a frame arrives at a switch, it has a VCI.
When the frame leaves, it has a different VCI.

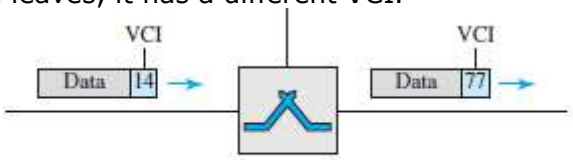


Figure 8.11 *Virtual-circuit identifier*

- Figure 8.11 show how the VCI in a data-frame changes from one switch to another.

2.8.2.2 Three Phases

• A source and destination need to go through 3 phases: setup, data-transfer, and teardown.

1) In setup phase, the source and destination use their global addresses to help switches make table entries for the connection.

2) In the teardown phase, the source and destination inform the switches to delete the corresponding entry.

3) Data-transfer occurs between these 2 phases.

2.8.2.2.1 Data Transfer Phase

• To transfer a frame from a source to its destination, all switches need to have a table-entry for this virtual-circuit.

• The table has four columns.

• The switch holds 4 pieces of information for each virtual-circuit that is already set up.

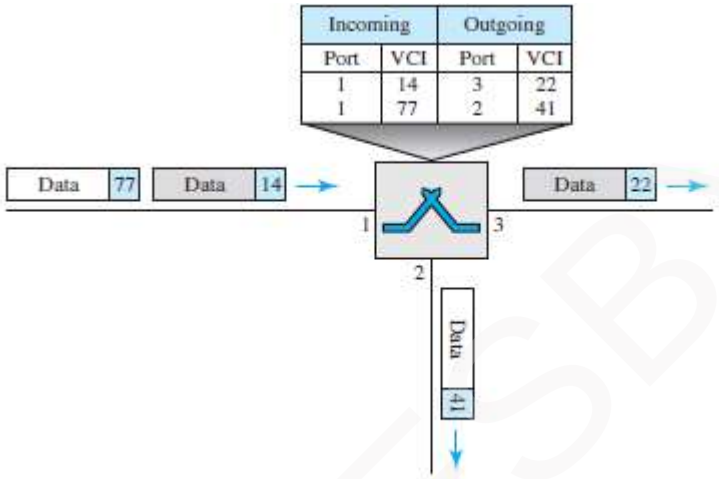


Figure 8.12 *Switch and tables in a virtual-circuit network*

- As shown in Figure 8.12, a frame arrives at port 1 with a VCI of 14.
- When the frame arrives, the switch looks in its table to find port 1 and a VCI of 14.
- When it is found, the switch knows to change the VCI to 22 & send out the frame from port 3.

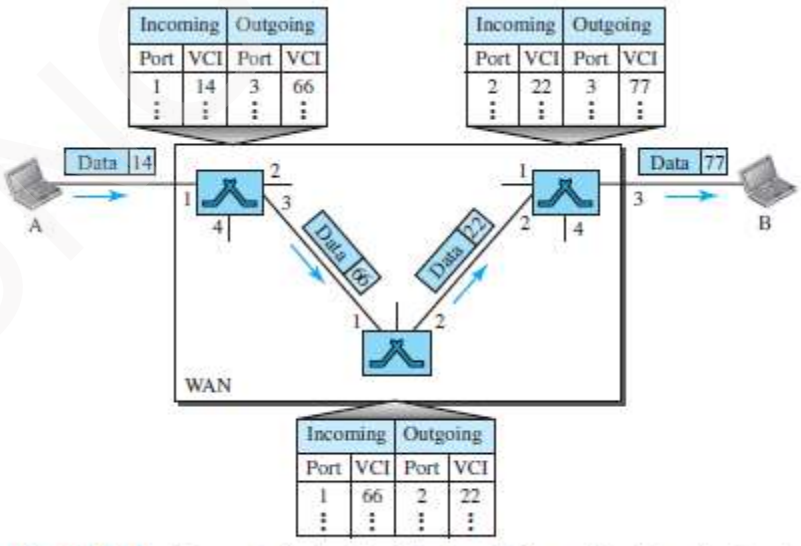


Figure 8.13 *Source-to-destination data transfer in a virtual-circuit network*

- As shown in Figure 8.13, each switch changes the VCI and routes the frame.
- The data-transfer phase is active until the source sends all its frames to the destination.
- The procedure at the switch is the same for each frame of a message.
- The process creates a virtual circuit, not a real circuit, between the source and destination.

2.8.2.2.2 Setup Phase

- A switch creates an entry for a virtual-circuit.
- For example, suppose source A needs to create a virtual-circuit to B.
- Two steps are required: 1) Setup-request and
  2) Acknowledgment.

1) Setup Request

➢ A setup-request frame is sent from the source to the destination (Figure 8.14).

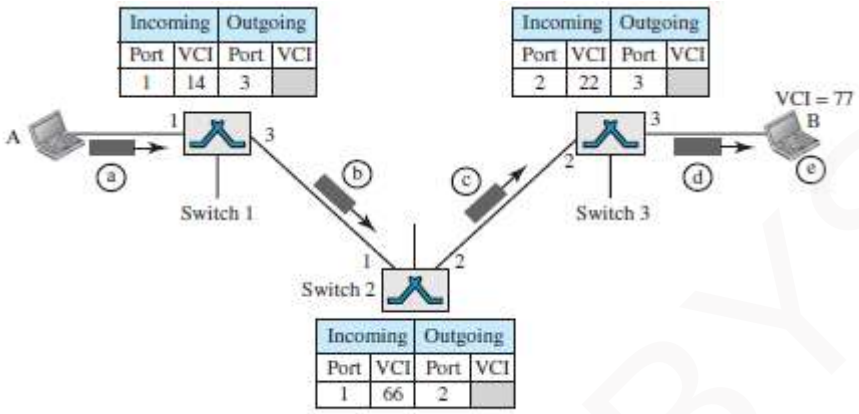


Figure 8.14 *Setup request in a virtual-circuit network*

➢ Following events occurs:

a) Source-A sends a setup-frame to switch-1.

b) Switch-1 receives the setup-frame.

- ¤ Switch-1 knows that a frame going from A to B goes out through port 3.
- ¤ The switch-1 has a routing table.
- ¤ The switch
  - → creates an entry in its table for this virtual-circuit
  - → is only able to fill 3 of the 4 columns.
- ¤ The switch
  - → assigns the incoming port (1) and
  - → chooses an available incoming-VCI (14) and the outgoing-port (3).
  - → does not yet know the outgoing VCI, which will be found during the acknowledgment step.
- ¤ The switch then forwards the frame through port-3 to switch-2.

c) Switch-2 receives the setup-request frame.

- ¤ The same events happen here as at switch-1.
- ¤ Three columns of the table are completed: In this case, incoming port (1), incoming-VCI (66), and outgoing port (2).

d) Switch-3 receives the setup-request frame.

- ¤ Again, three columns are completed: incoming port (2), incoming-VCI (22), and outgoing-port (3).

e) Destination-B

- → receives the setup-frame
- → assigns a VCI to the incoming frames that come from A, in this case 77.
- ¤ This VCI lets the destination know that the frames come from A, and no other sources.

2) Acknowledgment

➢ A special frame, called the acknowledgment-frame, completes the entries in the switching-tables (Figure 8.15).

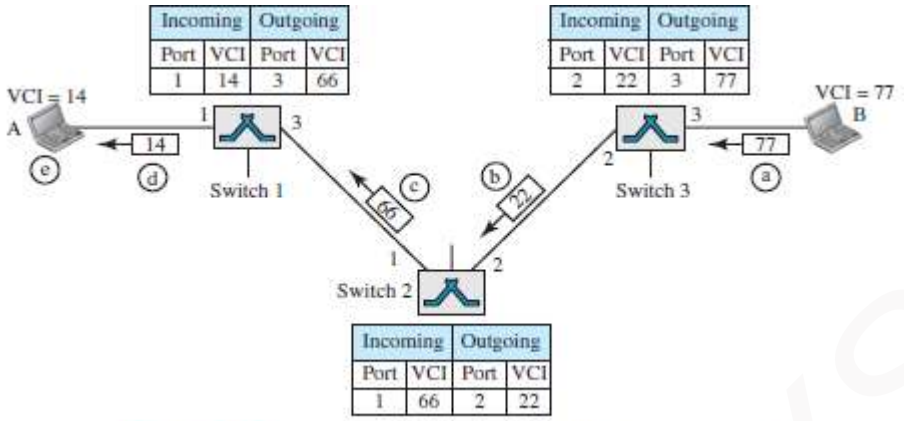


Figure 8.15 *Setup acknowledgment in a virtual-circuit network*

a) The destination sends an acknowledgment to switch-3.

- ¤ The acknowledgment carries the global source and destination-addresses so the switch knows which entry in the table is to be completed.
- ¤ The frame also carries VCI 77, chosen by the destination as the incoming-VCI for frames from A.
- ¤ Switch 3 uses this VCI to complete the outgoing VCI column for this entry.

b) Switch 3 sends an acknowledgment to switch-2 that contains its incoming-VCI in the table, chosen in the previous step.

- ¤ Switch-2 uses this as the outgoing VCI in the table.

c) Switch-2 sends an acknowledgment to switch-1 that contains its incoming-VCI in the table, chosen in the previous step.

- ¤ Switch-1 uses this as the outgoing VCI in the table.

d) Finally switch-1 sends an acknowledgment to source-A that contains its incoming-VCI in the table, chosen in the previous step.

e) The source uses this as the outgoing VCI for the data-frames to be sent to destination-B.

2.8.2.3 Teardown Phase

- Source-A, after sending all frames to B, sends a special frame called a teardown request.
- Destination-B responds with a teardown confirmation frame.
- All switches delete the corresponding entry from their tables.

2.8.2.4 Efficiency

- Resource reservation can be made in 2 cases:
  - 1) During the setup: Here, the delay for each packet is the same.
  - 2) On demand: Here, each packet may encounter different delays.
- Advantage of on demand resource allocation:
  - The source can check the availability of the resources, without actually reserving it.

2.8.2.5 Delay in Virtual Circuit Networks

- There is a one-time delay for setup and a one-time delay for teardown (Figure 8.16).
- If resources are allocated during the setup phase, there is no wait time for individual packets.
- The packet is traveling through two switches (routers).
- There are three transmission times (3T), three propagation times (3τ), data transfer delay, a setup delay and a teardown delay.
- The total delay time is

$$\text{Total delay} + 3T + 3\tau + \text{setup delay} + \text{teardown delay}$$

Circuit Switching	Datagram Packet Switching	Virtual circuit Packet switching
Dedicate transmission path	No dedicate path	No dedicate path
Continuous transmission of data	Transmission of packets	Transmission of packets
Fast enough for interactive	Fast enough for interactive	Fast enough for interactive
Message are not stored	Packets may be stored until delivered	Packets stored until delivered
The path is established for entire conversation	Route established for each packet	Route established for entire conversation
Call setup delay; negligible transmission delay	Packet transmission delay	Call setup delay; Packet transmission delay
Busy signal if called party busy	Sender may be notified if packet not delivered	Sender notified of connection denial
Overload may block call setup; no delay for established calls	Overload increases packet delay	Overload may block call setup; increases packet delay
Electromechanical or computerized switching nodes	Small switching nodes	Small switching nodes
User responsible for message loss protection	Network may be responsible for individual packets	Network may be responsible for packet sequences
Usually no speed or code conversion	Speed and code conversion	Speed and code conversion
Fixed bandwidth	Dynamic use of bandwidth	Dynamic use of bandwidth
No overhead bits after call setup	Overhead bits in each packet	Overhead bits in each packet

MODULE-WISE QUESTIONS

MODULE 2: DIGITAL TRANSMISSION (CONT.)

1) Explain the PCM encoder with neat diagram. (8*)
2) What do you mean by Sampling? Explain three sampling methods with a neat diagram. (4)
3) Explain non-uniform quantization and how to recover original signal using PCM decoder. (4)
4) Explain different types of transmission modes. (8*)
5) What is sampling and quantization? Explain briefly. (6)

MODULE 2(CONT.): ANALOG TRANSMISSION

1) Define digital to analog conversion? List different types of digital to analog conversion. (2)
2) Describe ASK, FSK and PSK mechanisms and apply them over the digital data 101101. (4)
3) Discuss the bandwidth requirement for ASK, FSK and PSK. (4*)
4) Explain different aspects of digital-to-analog conversion? (6*)
5) Define ASK. Explain BASK. (6*)
6) Define FSK. Explain BFSK. (6*)
7) Define PSK. Explain BPSK. (6*)
8) Explain QPSK. (6)
9) Explain the concept of constellation diagram. (6)
10) Explain QAM. (6)

MODULE 2(CONT.): BANDWIDTH UTILIZATION -- MULTIPLEXING AND SPREADING

1) Explain the concepts of multiplexing and list the categories of multiplexing? (4)
2) Define FDM? Explain the FDM multiplexing and demultiplexing process with neat diagrams. (6*)
3) Define and explain the concept of WDM. (6*)
4) Explain in detail synchronous TDM. (6*)
5) What do you mean by interleaving? Explain (4)
6) Explain Data Rate Management in Multi-level Multiplexing. (4*)
7) Explain the concept of empty-slots and frame-synchronizing in Multi-level Multiplexing. (6)
8) Explain in detail Statistical TDM. (6*)
9) Define FHSS and explain how it achieves bandwidth multiplexing. (8*)
10) Define DSSS and explain how it achieves bandwidth multiplexing. (8*)
11) Explain the analog hierarchy used by the telephone companies. (6)

MODULE 2(CONT.): SWITCHING

1) Explain in detail circuit-switched-network. (6*)
2) Explain switching with reference to TCP/IP Layers. (4)
3) Explain in detail datagram networks (8*)
4) What is Virtual-circuit Network? List five characteristics of VCN. (6*)
5) With relevant diagrams, explain the data transfer phase in a virtual circuit network. (8*)
6) Explain in detail setup Phase in VCN. (6)
7) Explain in detail acknowledgment Phase in VCN. (6)
8) Compare circuit-switched-network, Datagram & Virtual-circuit. (5*)

MODULE 3: TABLE OF CONTENTS

MODULE 3: ERROR-DETECTION AND CORRECTION

3.1 INTRODUCTION

3.1.1 Types of Errors

- When bits flow from 1 point to another, they are subject to unpredictable-changes '.' of interference.
- The interference can change the shape of the signal.
- Two types of errors: 1) Single-bit error 2) Burst-error.

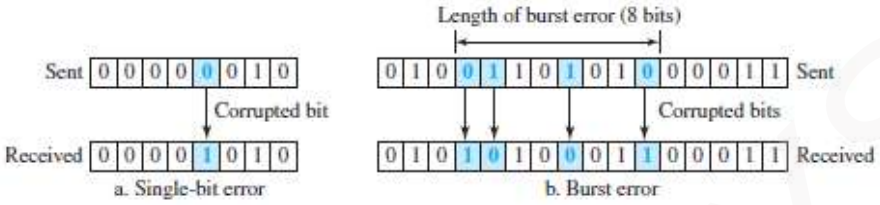


Figure 10.1 *Single-bit and burst error*

1) Single-Bit Error

- Only 1 bit of a given data is changed
  - → from 1 to 0 or
  - → from 0 to 1 (Figure 10.1a).

2) Burst Error

- Two or more bits in the data have changed
  - → from 1 to 0 or
  - → from 0 to 1 (Figure 10.1b).
- A burst-error occurs more than a single-bit error. This is because:
  - Normally, the duration of noise is longer than the duration of 1-bit.
- When noise affects data, the noise also affects the bits.
- The no. of corrupted-bits depends on
  - → data-rate and
  - → duration of noise.

3.1.2 Redundancy

- The central concept in detecting/correcting errors is *redundancy*.
- Some extra-bits along with the data have to be sent to detect/correct errors. These extra bits are called redundant-bits.
- The redundant-bits are
  - → added by the sender and
  - → removed by the receiver.
- The presence of redundant-bits allows the receiver to detect/correct errors.

3.1.3 Error Detection vs. Error Correction

- Error-correction is more difficult than error-detection.

1) Error Detection

- Here, we are checking whether any error has occurred or not.
- The answer is a simple YES or NO.
- We are not interested in the number of corrupted-bits.

2) Error Correction

- Here, we need to know
  - → exact number of corrupted-bits and
  - → location of bits in the message.
- Two important factors to be considered:
  - 1) Number of errors and
  - 2) Message-size.

3.1.4 Coding

- Redundancy is achieved through various coding-schemes.
  1) Sender adds redundant-bits to the data-bits. This process creates a relationship between
     → redundant-bits and
     → data-bits.
  2) Receiver checks the relationship between redundant-bits & data-bits to detect/correct errors.
- Two important factors to be considered:
  1) Ratio of redundant-bits to the data-bits and
  2) Robustness of the process.
- Two broad categories of coding schemes: 1) Block-coding and 2) Convolution coding.

3.2 Block Coding

- The message is divided into k-bit blocks. These blocks are called data-words.
- Here, r-redundant-bits are added to each block to make the length n=k+r.
- The resulting n-bit blocks are called code-words.
- Since n>k, the number of possible code-words is larger than the number of possible data-words.
- Block-coding process is 1-to-1; the same data-word is always encoded as the same code-word.
- Thus, we have 2^n-2^k code-words that are not used. These code-words are invalid or illegal.

3.2.1 Error Detection

- If the following 2 conditions are met, the receiver can detect a change in the original code-word:
  1) The receiver has a list of valid code-words.
  2) The original code-word has changed to an invalid code-words.

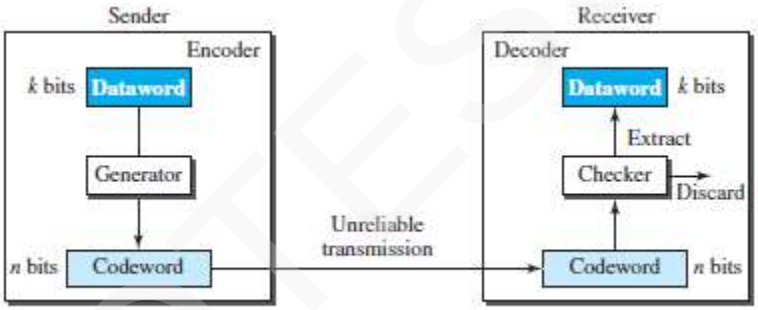


Figure 10.2 *Process of error detection in block coding*

- Here is how it works (Figure 10.2):

  1) At Sender

  i) The sender creates code-words out of data-words by using a generator.
  The generator applies the rules and procedures of encoding.

  ii) During transmission, each code-word sent to the receiver may change.

  2) At Receiver

  i) a) If the received code-word is the same as one of the valid code-words,
  the code-word is accepted;
  the corresponding data-word is extracted for use.

  b) If the received code-word is invalid, the code-word is discarded.

  ii) However, if the code-word is corrupted but the received code-word still matches a valid code-word, the error remains undetected.

- An error-detecting code can detect only the types of errors for which it is designed; other types of errors may remain undetected.

Example 3.1

Let us assume that $k = 2$ and $n = 3$. Table 10.1 shows the list of datawords and codewords.

Table 10.1 *A code for error detection*

Dataword	*Codeword*	*Dataword*	*Codeword*
00	000	10	101
01	011	11	110

Assume the sender encodes the dataword 01 as 011 and sends it to the receiver. Consider the following cases:

1. The receiver receives 011. It is a valid codeword. The receiver extracts the dataword 01 from it.
2. The codeword is corrupted during transmission, and 111 is received (the leftmost bit is corrupted). This is not a valid codeword and is discarded.
3. The codeword is corrupted during transmission, and 000 is received (the right two bits are corrupted). This is a valid codeword. The receiver incorrectly extracts the dataword 00. Two corrupted bits have made the error undetectable.

3.2.1.1 Hamming Distance

- The main concept for error-control: Hamming distance.
- The Hamming distance b/w 2 words is the number of differences between the corresponding bits.
- Let d(x,y) = Hamming distance b/w 2 words x and y.
- Hamming distance can be found by
  - → applying the XOR operation on the 2 words and
  - → counting the number of 1s in the result.
- For example:
  1) The Hamming distance d(000, 011) is 2 because 000⊕011= 011 (two 1s).
  2) The Hamming distance d(10101, 11110) is 3 because 10101⊕11110 = 01011 (three 1s).

Hamming Distance and Error

➢ Hamming distance between the received word and the sent code-word is the number of bits that are corrupted during transmission.

➢ For example: Let Sent code-word = 00000
Received word = 01101
Hamming distance = d(00000, 01101) =3. Thus, 3 bits are in error.

3.2.1.1.1 Minimum Hamming Distance for Error Detection

- Minimum Hamming distance is the smallest Hamming distance b/w all possible pairs of code-words.
- Let d_{min} = minimum Hamming distance.
- To find d_{min} value, we find the Hamming distances between all words and select the smallest one.

Minimum-distance for Error-detection

➢ If 's' errors occur during transmission, the Hamming distance b/w the sent code-word and received code-word is 's' (Figure 10.3).

➢ If code has to detect upto 's' errors, the minimum-distance b/w the valid codes must be 's+1' i.e. $d_{min}=s+1$.

➢ We use a geometric approach to define $d_{min}=s+1$.

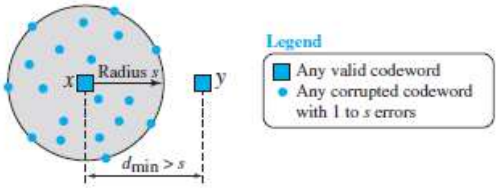


Figure 10.3 *Geometric concept explaining d_{min} in error detection*

- Let us assume that the sent code-word x is at the center of a circle with radius s.
- All received code-words that are created by 0 to s errors are points inside the circle or on the perimeter of the circle.
- All other valid code-words must be outside the circle

- For example: A code scheme has a Hamming distance d_{min} = 4.
  This code guarantees the detection of upto 3 errors (d = s + 1 or s = 3).

3.2.1.2 Linear Block Codes

- Almost all block codes belong to a subset of block codes called linear block codes.
- A linear block code is a code in which the XOR of 2 valid code-words creates another valid code-word.
  (XOR → Addition modulo-2)

Table 10.1 *A code for error detection*

Dataword	*Codeword*	*Dataword*	*Codeword*
00	000	10	101
01	011	11	110

- The code in Table 10.1 is a linear block code because the result of XORing any code-word with any other code-word is a valid code-word.
  For example, the XORing of the 2nd and 3rd code-words creates the 4th one.

3.2.1.2.1 Minimum Distance for Linear Block Codes

- Minimum Hamming distance is no. of 1s in the nonzero valid code-word with the smallest no. of 1s.
- In Table 10.1,
  The numbers of 1s in the nonzero code-words are 2, 2, and 2.
  So the minimum Hamming distance is $d_{min} = 2$.

3.2.1.3 Parity Check Code

- This code is a linear block code. This code can detect an odd number of errors.
- A k-bit data-word is changed to an n-bit code-word where n=k+1.
- One extra bit is called the parity-bit.
- The parity-bit is selected to make the total number of 1s in the code-word even.
- Minimum hamming distance d_{min} = 2. This means the code is a single-bit error-detecting code.

Table 10.2 *Simple parity-check code C(5, 4)*

Dataword	Codeword	Dataword	Codeword
0000	**00000**	1000	**10001**
0001	**00011**	1001	**10010**
0010	**00101**	1010	**10100**
0011	**00110**	1011	**10111**
0100	**01001**	1100	**11000**
0101	**01010**	1101	**11011**
0110	**01100**	1110	**11101**
0111	**01111**	1111	**11110**

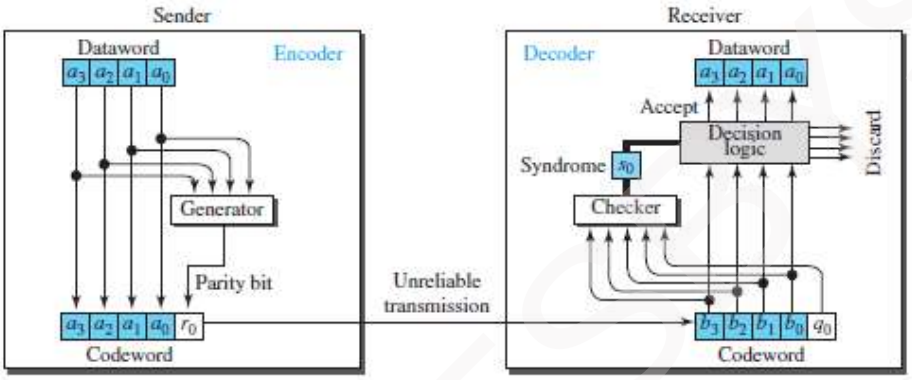


Figure 10.4 *Encoder and decoder for simple parity-check code*

- Here is how it works (Figure 10.4):

1) At Sender

- The encoder uses a generator that takes a copy of a 4-bit data-word (a_0, a_1, a_2, and a_3) and generates a parity-bit r_0.
- The encoder
  - → accepts a copy of a 4-bit data-word (a_0, a_1, a_2, and a_3) and
  - → generates a parity-bit r_0 using a generator
  - → generates a 5-bit code-word
- The parity-bit & 4-bit data-word are added to make the number of 1s in the code-word even.
- The addition is done by using the following:

$$r_0 = a_3 + a_2 + a_1 + a_0 \quad \text{(modulo-2)}$$

- The result of addition is the parity-bit.
  1) If the no. of 1s in data-word = even, result = 0. (r_0=0)
  2) If the no. of 1s in data-word = odd, result = 1. (r_0=1)
  3) In both cases, the total number of 1s in the code-word is even.
- The sender sends the code-word, which may be corrupted during transmission.

2) At Receiver

- The receiver receives a 5-bit word.
- The checker performs the same operation as the generator with one exception:
  The addition is done over all 5 bits.

$$s_0 = b_3 + b_2 + b_1 + b_0 + q_0 \quad \text{(modulo-2)}$$

- The result is called the syndrome bit (s_o).
- Syndrome bit = 0 when the no. of 1s in the received code-word is even; otherwise, it is 1.
- The syndrome is passed to the decision logic analyzer.
  1) If s_0=0, there is no error in the received code-word. The data portion of the received code-word is accepted as the data-word.
  2) If s_0=1, there is error in the received code-word. The data portion of the received code-word is discarded. The data-word is not created.

Example 3.2

Let us look at some transmission scenarios. Assume the sender sends the dataword 1011. The codeword created from this dataword is 10111, which is sent to the receiver. We examine five cases:

1. No error occurs; the received codeword is 10111. The syndrome is 0. The dataword 1011 is created.
2. One single-bit error changes a_1. The received codeword is 10011. The syndrome is 1. No dataword is created.
3. One single-bit error changes r_0. The received codeword is 10110. The syndrome is 1. No dataword is created. Note that although none of the dataword bits are corrupted, no dataword is created because the code is not sophisticated enough to show the position of the corrupted bit.
4. An error changes r_0 and a second error changes a_3. The received codeword is 00110. The syndrome is 0. The dataword 0011 is created at the receiver. Note that here the dataword is wrongly created due to the syndrome value. The simple parity-check decoder cannot detect an even number of errors. The errors cancel each other out and give the syndrome a value of 0.
5. Three bits—a_3, a_2, and a_1—are changed by errors. The received codeword is 01011. The syndrome is 1. The dataword is not created. This shows that the simple parity check, guaranteed to detect one single error, can also find any odd number of errors.

3.3 Cyclic Codes

- Cyclic codes are special linear block codes with one extra property:
  If a code-word is cyclically shifted (rotated), the result is another code-word.
  For ex: if code-word = 1011000 and we cyclically left-shift, then another code-word = 0110001.
- Let First-word = a_0 to a_6 and Second-word = b_0 to b_6, we can shift the bits by using the following:

$$b_1 = a_0 \quad b_2 = a_1 \quad b_3 = a_2 \quad b_4 = a_3 \quad b_5 = a_4 \quad b_6 = a_5 \quad b_0 = a_6$$

3.3.1 Cyclic Redundancy Check (CRC)

- CRC is a cyclic code that is used in networks such as LANs and WANs.

Table 10.3 *A CRC code with C(7, 4)*

Dataword	Codeword	Dataword	Codeword
0000	0000000	1000	1000101
0001	0001011	1001	1001110
0010	0010110	1010	1010011
0011	0011101	1011	1011000
0100	0100111	1100	1100010
0101	0101100	1101	1101001
0110	0110001	1110	1110100
0111	0111010	1111	1111111

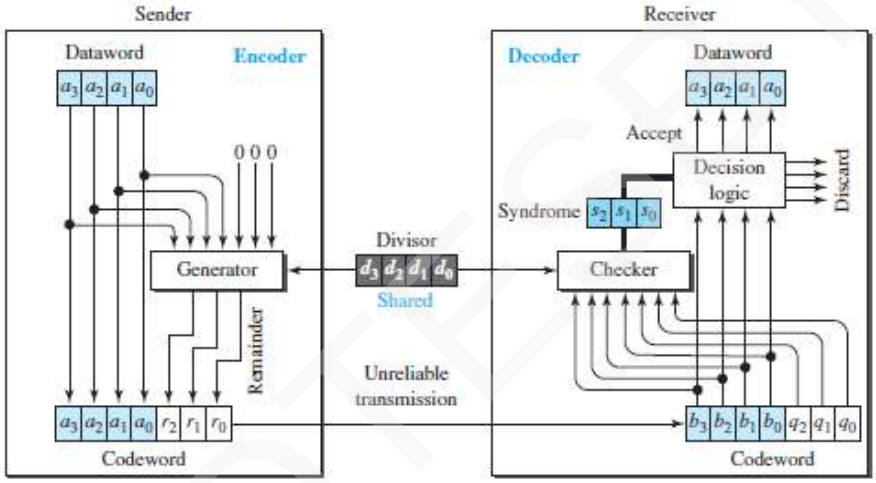


Figure 10.5 *CRC encoder and decoder*

- Let Size of data-word = k bits (here k=4).
  Size of code-word = n bits (here n=7).
  Size of divisor = n-k+1 bits (here n-k+1=4). (Augmented → increased)
- Here is how it works (Figure 10.5):

1) At Sender

- n-k 0s is appended to the data-word to create augmented data-word. (here n-k=3).
- The augmented data-word is fed into the generator (Figure 10.6).
- The generator divides the augmented data-word by the divisor.
- The remainder is called check-bits ($r_2r_1r_0$).
- The check-bits ($r_2r_1r_0$) are appended to the data-word to create the code-word.

2) At Receiver

- The possibly corrupted code-word is fed into the checker.
- The checker is a replica of the generator.
- The checker divides the code-word by the divisor.
- The remainder is called syndrome bits ($r_2r_1r_0$).
- The syndrome bits are fed to the decision-logic-analyzer.
- The decision-logic-analyzer performs following functions:

i) For No Error

- If all syndrome-bits are 0s, the received code-word is accepted.
- Data-word is extracted from received code-word (Figure 10.7a).

ii) For Error

- If all syndrome-bits are not 0s, the received code-word is discarded (Figure 10.7b).

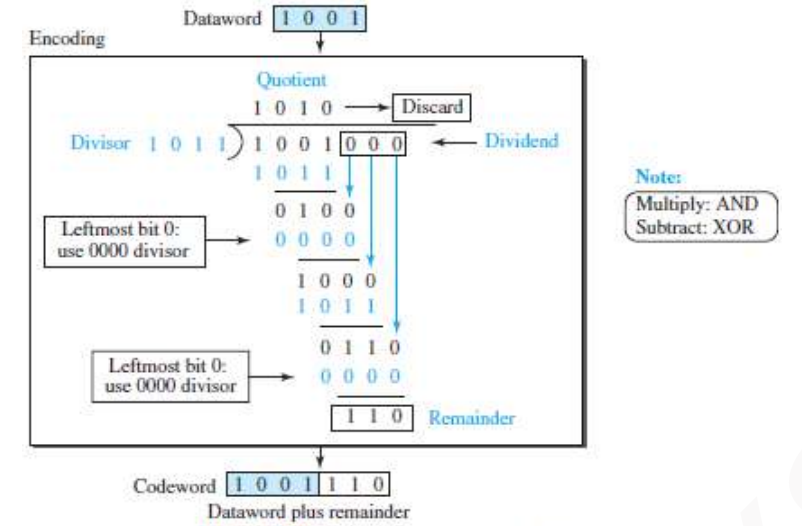


Figure 10.6 *Division in CRC encoder*

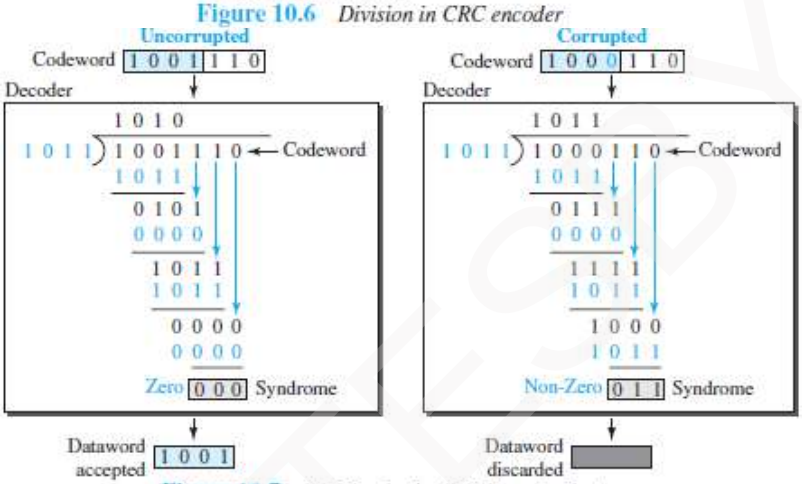
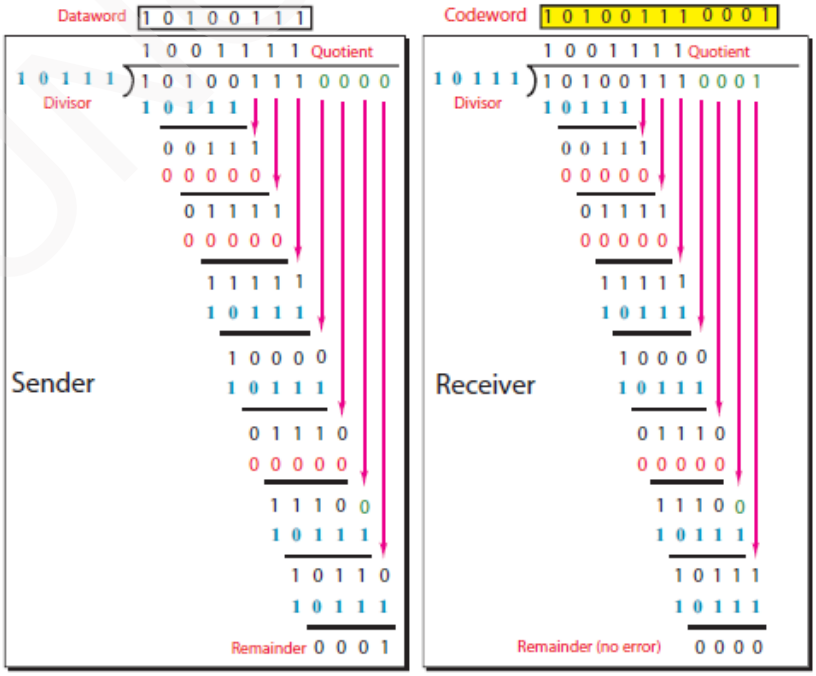


Figure 10.7 *Division in the CRC decoder for two cases*

Example 3.3

Given the dataword 10100111 and the divisor 10111, show the generation of the CRC codeword at the sender site (using binary division).



3.3.2 Polynomials

- A pattern of 0s and 1s can be represented as a polynomial with coefficients of 0 and 1 (Figure 10.8).
- The power of each term shows the position of the bit; the coefficient shows the value of the bit.

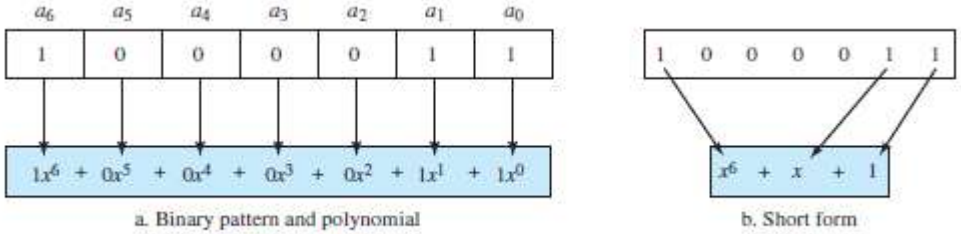


Figure 10.8 *A polynomial to represent a binary word*

3.3.3 Cyclic Code Encoder Using Polynomials

- Let Data-word = 1001 = x^3+1.
  Divisor = 1011 = x^3+x+1.
- In polynomial representation, the divisor is referred to as generator polynomial t(x) (Figure 10.9).

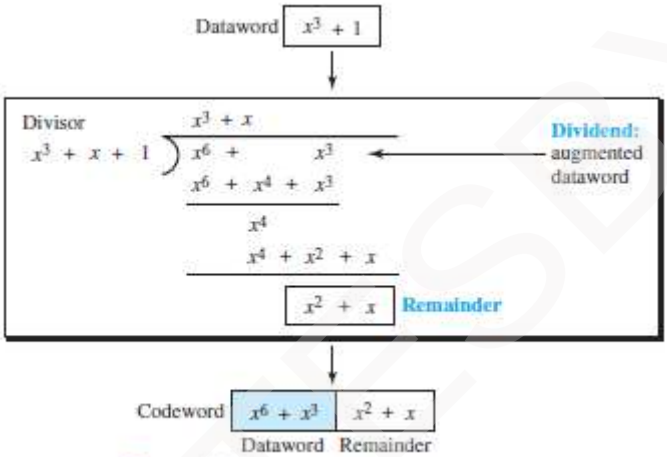


Figure 10.9 *CRC division using polynomials*

3.3.4 Cyclic Code Analysis

- We define the following, where *f*(*x*) is a polynomial with binary coefficients:

Dataword: *d*(*x*) Codeword: *c*(*x*) Generator: *g*(*x*) Syndrome: *s*(*x*) Error: *e*(*x*)

In a cyclic code,

1. If $s(x) \neq 0$, one or more bits is corrupted.
2. If $s(x) = 0$, either
   a. No bit is corrupted, or
   b. Some bits are corrupted, but the decoder failed to detect them.

Single Bit Error

- If the generator has more than one term and the coefficient of x^0 is 1, all single-bit errors can be caught.

Two Isolated Single-Bit Errors

- If a generator cannot divide x^i+1 (t between 0 & n-1), then all isolated double errors can be detected (Figure 10.10).

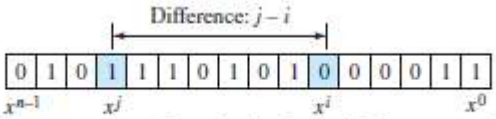


Figure 10.10 *Representation of two isolated single-bit errors using polynomials*

Odd Numbers of Errors

• A generator that contains a factor of x+1 can detect all odd-numbered errors.

A good polynomial generator needs to have the following characteristics:

1. **It should have at least two terms.**
2. **The coefficient of the term x^0 should be 1.**
3. **It should not divide $x^t + 1$, for t between 2 and $n - 1$.**
4. **It should have the factor $x + 1$.**

Standard Polynomials

Table 10.4 *Standard polynomials*

Name	*Polynomial*	*Used in*
CRC-8	x^8+x^2+x+1 **100000111**	ATM header
CRC-10	$x^{10}+x^9+x^5+x^4+x^2+1$ **11000110101**	ATM AAL
CRC-16	$x^{16}+x^{12}+x^5+1$ **10001000000100001**	HDLC
CRC-32	$x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^8+x^7+x^5+x^4+x^2+x+1$ **10000010011000001000111011011 0111**	LANs

3.3.5 Advantages of Cyclic Codes

• The cyclic codes have a very good performance in detecting
- → single-bit errors
- → double errors
- → odd number of errors and
- → burst-errors.

• They can easily be implemented in hardware and software. They are fast when implemented in hardware.

3.4 Checksum

- Checksum is an error-detecting technique.
- In the Internet,
  - → The checksum is mostly used at the network and transport layer.
  - → The checksum is not used in the data link layer.
- Like linear and cyclic codes, the checksum is based on the concept of redundancy.
- Here is how it works (Figure 10.15):

1) At Source

- Firstly the message is divided into m-bit units.
- Then, the generator creates an extra m-bit unit called the checksum.
- The checksum is sent with the message.

2) At Destination

- The checker creates a new checksum from the combination of the message and sent-checksum.
  - i) If the new checksum is all 0s, the message is accepted.
  - ii) If the new checksum is not all 0s, the message is discarded.

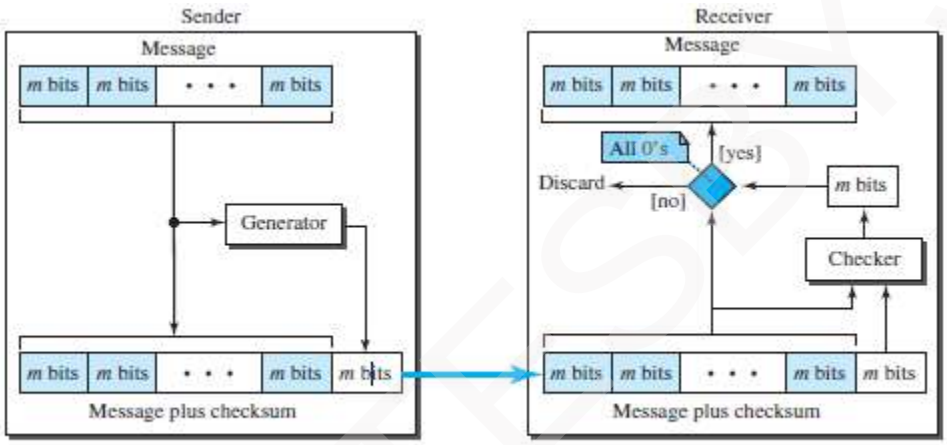


Figure 10.15 *Checksum*

3.4.1 Concept of Checksum

Consider the following example:

Example 3.4

- Our data is a list of five 4-bit numbers that we want to send to a destination.
- In addition to sending these numbers, we send the sum of the numbers.
- For example:
  - Let set of numbers = (7, 11, 12, 0, 6).
  - We send (7, 11, 12, 0, 6, 36), where 36 is the sum of the original numbers.
- The receiver adds the five numbers and compares the result with the sum.
- If the result & the sum are the same,
  - The receiver assumes no error, accepts the five numbers, and discards the sum.
  - Otherwise, there is an error somewhere and the data are not accepted.

Example 3.5

- To make the job of the receiver easy if we send the negative (complement) of the sum, called the checksum.
- In this case, we send (7, 11, 12, 0, 6, -36).
- The receiver can add all the numbers received (including the checksum).
- If the result is 0, it assumes no error; otherwise, there is an error.

3.4.1.1 One's Complement

- The previous example has one major drawback.
  - All of our data can be written as a 4-bit word (they are less than 15) except for the checksum.
- Solution: Use one's complement arithmetic.
  - ➢ We can represent unsigned numbers between 0 and 2^n-1 using only n bits.
  - ➢ If the number has more than n bits, the extra leftmost bits need to be added to the n rightmost bits (wrapping).
  - ➢ A negative number can be represented by inverting all bits (changing 0 to 1 and 1 to 0).
  - ➢ This is the same as subtracting the number from 2^n-1.

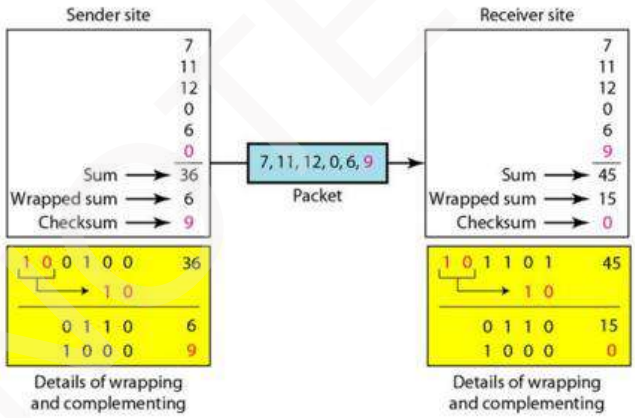


Figure 10.16

- Here is how it works (Figure 10.16):

1) At Sender

- ➢ The sender initializes the checksum to 0 and adds all data items and the checksum.
- ➢ The result is 36.
- ➢ However, 36 cannot be expressed in 4 bits.
- ➢ The extra two bits are wrapped and added with the sum to create the wrapped sum value 6.
- ➢ The sum is then complemented, resulting in the checksum value 9 (15 - 6 = 9).
- ➢ The sender now sends six data items to the receiver including the checksum 9.

2) At Receiver

- ➢ The receiver follows the same procedure as the sender.
- ➢ It adds all data items (including the checksum); the result is 45.
- ➢ The sum is wrapped and becomes 15. The wrapped sum is complemented and becomes 0.
- ➢ Since the value of the checksum is 0, this means that the data is not corrupted. The receiver drops the checksum and keeps the other data items.
- ➢ If the checksum is not zero, the entire packet is dropped.

3.4.1.2 Internet Checksum

- Traditionally, the Internet has been using a 16-bit checksum.
- The sender or the receiver uses five steps.

Table 10.5 *Procedure to calculate the traditional checksum*

Sender	*Receiver*
1. The message is divided into 16-bit words.	1. The message and the checksum are received.
2. The value of the checksum word is initially set to zero.	2. The message is divided into 16-bit words.
3. All words including the checksum are added using one's complement addition.	3. All words are added using one's complement addition.
4. The sum is complemented and becomes the checksum.	4. The sum is complemented and becomes the new checksum.
5. The checksum is sent with the data.	5. If the value of the checksum is 0, the message is accepted; otherwise, it is rejected.

3.4.1.3 Algorithm

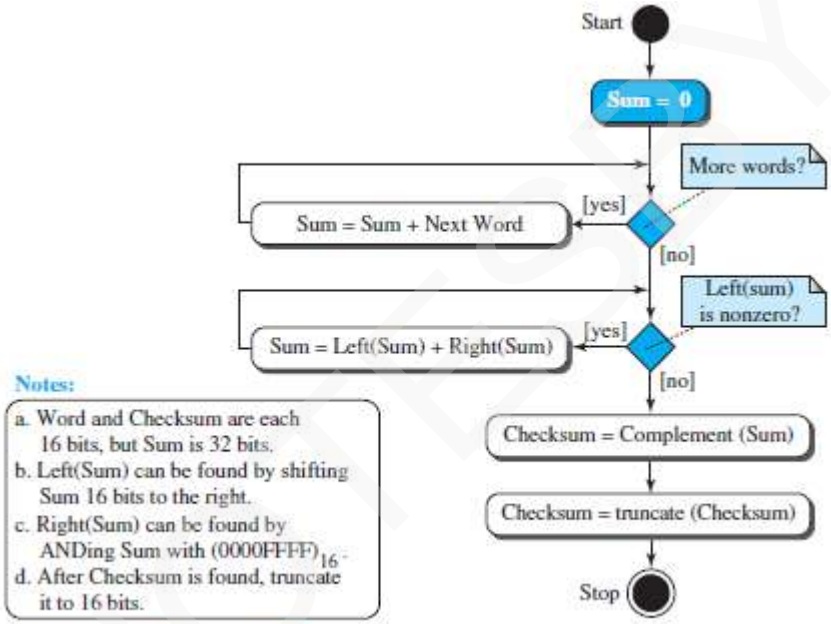


Figure 10.17 *Algorithm to calculate a traditional checksum*

3.4.2 Other Approaches to the Checksum

- If two 16-bit items are transposed in transmission, the checksum cannot catch this error.
- The reason is that the traditional checksum is not weighted: it treats each data item equally.
- In other words, the order of data items is immaterial to the calculation.
- Two approaches have been used to prevent this problem: 1)Fletcher and 2)Adler

3.4.2.1 Fletcher Checksum

- The Fletcher checksum was devised to weight each data item according to its position.
- Fletcher has proposed two algorithms: 8-bit and 16-bit (Figure 10.18).
- The first, 8-bit Fletcher, calculates on 8-bit data items and creates a 16-bit checksum.
  The second, 16-bit Fletcher, calculates on 16-bit data items and creates a 32-bit checksum.
- The 8-bit Fletcher is calculated over data octets (bytes) and creates a 16-bit checksum.
- The calculation is done modulo 256 (2^8), which means the intermediate results are divided by 256 and the remainder is kept.
- The algorithm uses two accumulators, L and R.
- The first simply adds data items together;
  The second adds a weight to the calculation.

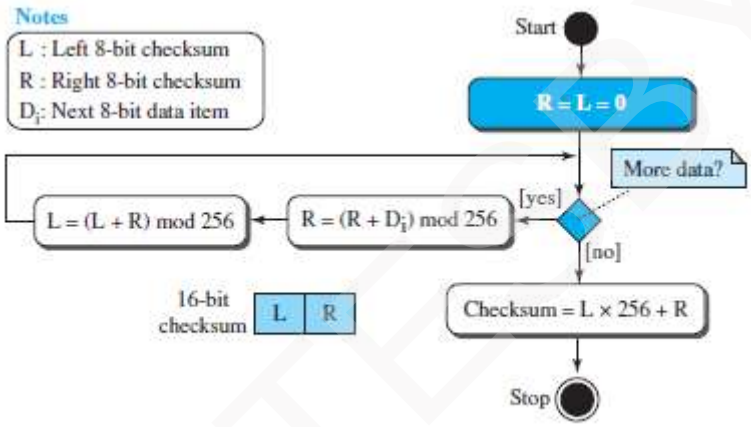


Figure 10.18 *Algorithm to calculate an 8-bit Fletcher checksum*

3.4.2.2 Adler Checksum

- The Adler checksum is a 32-bit checksum.
- It is similar to the 16-bit Fletcher with three differences (Figure 10.19).
  1) Calculation is done on single bytes instead of 2 bytes at a time.
  2) The modulus is a prime number (65,521) instead of 65,536.
  3) L is initialized to 1 instead of 0.
- A prime modulo has a better detecting capability in some combinations of data.

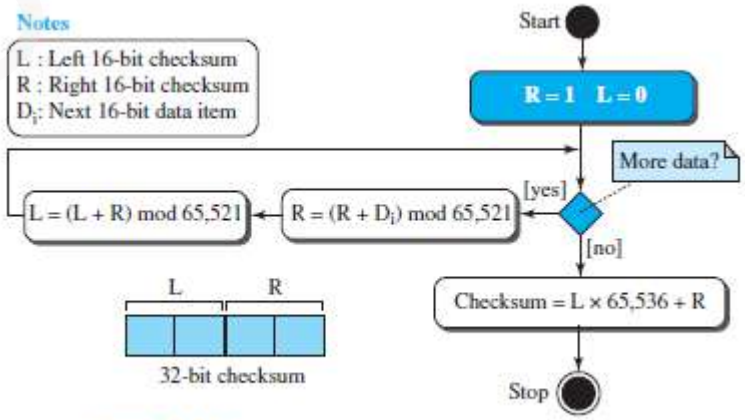


Figure 10.19 *Algorithm to calculate an Adler checksum*

3.5 FORWARD ERROR CORRECTION

• Retransmission of corrupted and lost packets is not useful for real-time multimedia transmission because it creates an unacceptable delay in reproducing: we need to wait until the lost or corrupted packet is resent.
• We need to correct the error or reproduce the packet immediately.
• Several schemes have been designed and used that are collectively referred to as forward error correction (FEC) techniques.

3.5.1 Using Hamming Distance

• To detect t errors, we need to have $d_{min} = 2t + 1$ (Figure 10.20).
• In other words, if we want to correct 10 bits in a packet, we need to make the minimum hamming distance 21 bits, which means a lot of redundant bits need to be sent with the data.

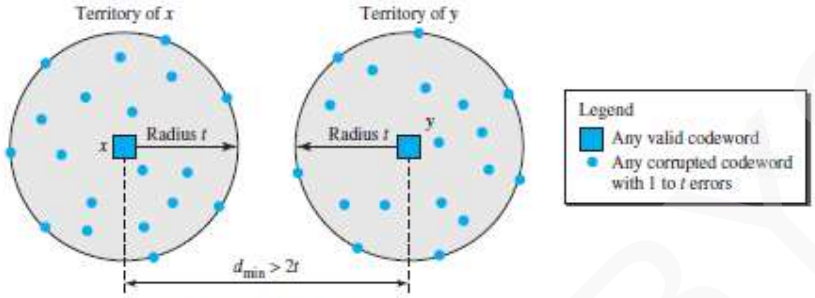


Figure 10.20 *Hamming distance for error correction*

3.5.2 Using XOR

• Use the property of the exclusive OR operation as shown below.

$$R = P_1 \oplus P_2 \oplus \ldots \oplus P_i \oplus \ldots \oplus P_N \quad \rightarrow \quad P_i = P_1 \oplus P_2 \oplus \ldots \oplus R \oplus \ldots \oplus P_N$$

• We divide a packet into N chunks, create the exclusive OR of all the chunks and send $N + 1$ chunks.
• If any chunk is lost or corrupted, it can be created at the receiver site.
• If N = 4, it means that we need to send 25 percent extra data and be able to correct the data if only one out of four chunks is lost.

3.5.3 Chunk Interleaving

• Another way to achieve FEC in multimedia is to allow some small chunks to be missing at the receiver.
• We cannot afford to let all the chunks belonging to the same packet be missing.
However, we can afford to let one chunk be missing in each packet.

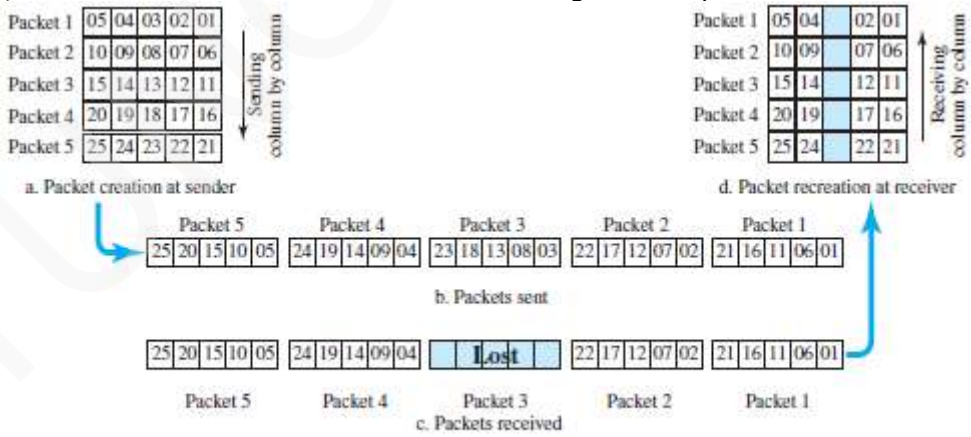


Figure 10.21 *Interleaving*

• In Figure 10.21, each packet is divided into 5 chunks (normally the number is much larger).
• Then, we can create data chunk-by-chunk (horizontally), but combine the chunks into packets vertically.
• In this case, each packet sent carries a chunk from several original packets.
• If the packet is lost, we miss only one chunk in each packet.
• Normally, missing of a chunk is acceptable in multimedia communication.

3.5.4 Combining Hamming Distance and Interleaving

- Hamming distance and interleaving can be combined.
- Firstly, we can create n-bit packets that can correct t-bit errors.
- Then, we interleave m rows and send the bits column-by-column.
- In this way, we can automatically correct burst-errors up to m×t bit errors.

3.5.5 Compounding High- and Low-Resolution Packets

- Another solution is to create a duplicate of each packet with a low-resolution redundancy and combine the redundant version with the next packet.
- For example (Figure 10.22):

  We can create 4 low-resolution packets out of 5 high-resolution packets and send them (Fig 10.22).
- If a packet is lost, we can use the low-resolution version from the next packet.
- In this method, if the last packet is lost, it cannot be recovered, but we use the low-resolution version of a packet if the lost packet is not the last one.
- The audio and video reproduction does not have the same quality, but the lack of quality is not recognized most of the time.

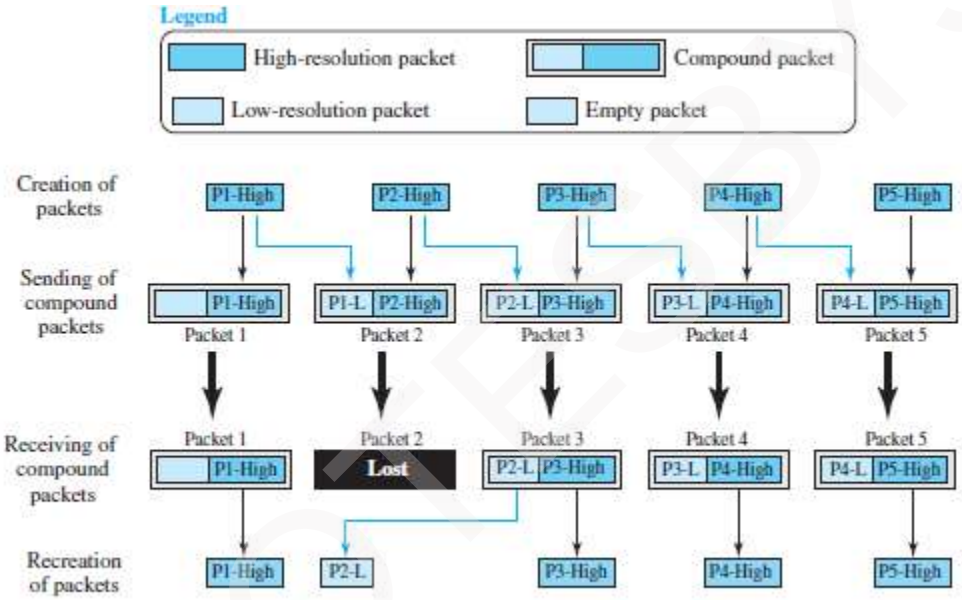


Figure 10.22 *Compounding high- and low-resolution packets*

MODULE 3(CONT.): DATA LINK CONTROL

3.6 DLC SERVICES

• The data link control (DLC) deals with procedures for communication between two adjacent nodes i.e. node-to-node communication.
• Data link control functions include 1) Framing and 2) Flow control and 3) Error control.

3.6.1 Framing

• A frame is a group of bits.
• Framing means organizing the bits into a frame that are carried by the physical layer.
• The data-link-layer needs to form frames, so that each frame is distinguishable from another.
• Framing separates a message from other messages by adding sender-address & destination-address.
• The destination-address defines where the packet is to go.
The sender-address helps the recipient acknowledge the receipt.
• Q: Why the whole message is not packed in one frame?
Ans: Large frame makes flow and error-control very inefficient.
Even a single-bit error requires the re-transmission of the whole message.
When a message is divided into smaller frames, a single-bit error affects only that small frame.

(Our postal system practices a type of framing. The simple act of inserting a letter into an envelope separates one piece of information from another; the envelope serves as the delimiter. In addition, each envelope defines the sender and receiver addresses since the postal system is a many-to-many carrier facility).

3.6.1.1 Frame Size

• Two types of frames:

1) Fixed Size Framing

- There is no need for defining boundaries of frames; the size itself can be used as a delimiter.
- For example: ATM WAN uses frames of fixed size called cells.

2) Variable Size Framing

- We need to define the end of the frame and the beginning of the next frame.
- Two approaches are used: 1) Character-oriented approach
  2) Bit-oriented approach.

3.6.1.2 Character Oriented Framing

• Data to be carried are 8-bit characters from a coding system such as ASCII (Figure 11.1).
• The header and the trailer are also multiples of 8 bits.
  1) Header carries the source and destination-addresses and other control information.
  2) Trailer carries error-detection or error-correction redundant bits.
• To separate one frame from the next frame, an 8- bit (I-byte) flag is added at the beginning and the end of a frame.
• The flag is composed of protocol-dependent special characters.
• The flag signals the start or end of a frame.

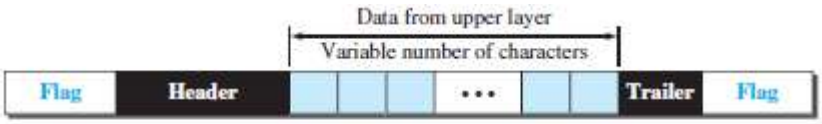


Figure 11.1 *A frame in a character-oriented protocol*

• Problem:
  ➢ Character-oriented framing is suitable when only text is exchanged by the data-link-layers.
  ➢ However, if we send other type of information (say audio/video), then any pattern used for the flag can also be part of the information.
  ➢ If the flag-pattern appears in the data-section, the receiver might think that it has reached the end of the frame.

Solution: A byte-stuffing is used.
  (Byte stuffing → character stuffing)

• In byte stuffing, a special byte is added to the data-section of the frame when there is a character with the same pattern as the flag.
• The data-section is stuffed with an extra byte. This byte is called the escape character (ESC), which has a predefined bit pattern.
• When a receiver encounters the ESC character, the receiver
  → removes ESC character from the data-section and
  → treats the next character as data, not a delimiting flag.
• Problem:
  ➢ What happens if the text contains one or more escape characters followed by a flag?
  ➢ The receiver removes the escape character, but keeps the flag, which is incorrectly interpreted as the end of the frame.

Solution:
  ➢ Escape characters part of the text must also be marked by another escape character (Fig 11.2).

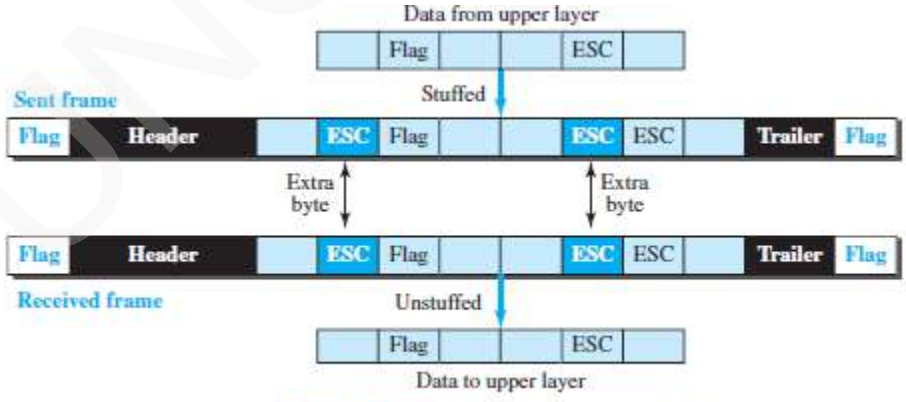


Figure 11.2 *Byte stuffing and unstuffing*

• In short, byte stuffing is the process of adding one extra byte whenever there is a flag or escape character in the text.

3.6.1.3 Bit Oriented Framing

• The data-section of a frame is a sequence of bits to be interpreted by the upper layer as text, audio, video, and so on.

• However, in addition to headers and trailers, we need a delimiter to separate one frame from the other.

• Most protocols use a special 8-bit pattern flag 01111110 as the delimiter to define the beginning and the end of the frame (Figure 11.3).

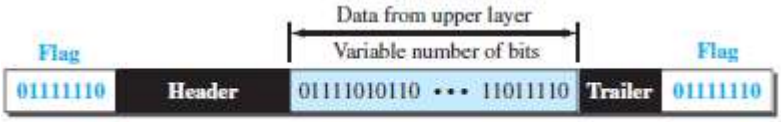


Figure 11.3 *A frame in a bit-oriented protocol*

• Problem:

- If the flag-pattern appears in the data-section, the receiver might think that it has reached the end of the frame.

Solution: A bit-stuffing is used.

• In bit stuffing, if a 0 and five consecutive 1 bits are encountered, an extra 0 is added. This extra stuffed bit is eventually removed from the data by the receiver. (Figure 11.4).

• This guarantees that the flag field sequence does not inadvertently appear in the frame.

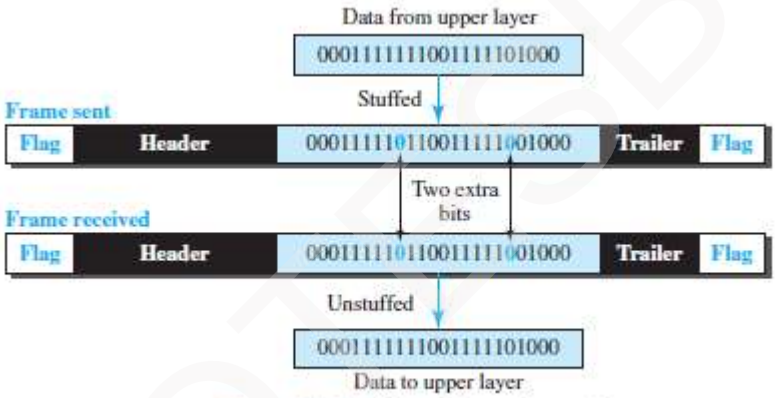


Figure 11.4 *Bit stuffing and unstuffing*

• In short, bit stuffing is the process of adding one extra 0 whenever five consecutive 1s follow a 0 in the data, so that the receiver does not mistake the pattern 0111110 for a flag.

3.6.2 Flow Control and Error Control

• One of the responsibilities of the DLC sublayer is flow and error control at the data-link layer.

3.6.2.1 Flow Control

• Whenever an entity produces items and another entity consumes them, there should be a balance between production and consumption rates.

• If the items are produced faster than they can be consumed, the consumer can be overwhelmed and may need to discard some items.

• We need to prevent losing the data items at the consumer site.

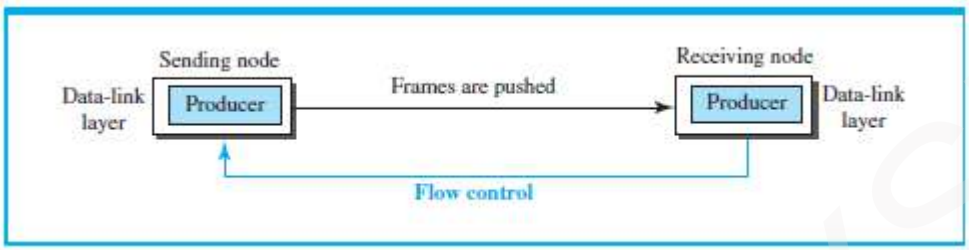


Figure 11.5 *Flow control at the data-link layer*

• At the sending node, the data-link layer tries to push frames toward the data-link layer at the receiving node (Figure 11.5).

• If the receiving node cannot process and deliver the packet to its network at the same rate that the frames arrive, it becomes overwhelmed with frames.

• Here, flow control can be feedback from the receiving node to the sending node to stop or slow down pushing frames.

3.6.2.1.1 Buffers

• Flow control can be implemented by using buffer.

• A buffer is a set of memory locations that can hold packets at the sender and receiver.

• Normally, two buffers can be used.

1) First buffer at the sender.
2) Second buffer at the receiver.

• The flow control communication can occur by sending signals from the consumer to the producer.

• When the buffer of the receiver is full, it informs the sender to stop pushing frames.

3.6.2.2 Error Control

• Error-control includes both error-detection and error-correction.

• Error-control allows the receiver to inform the sender of any frames lost/damaged in transmission.

• A CRC is

→ added to the frame header by the sender and
→ checked by the receiver.

• At the data-link layer, error control is normally implemented using one of the following two methods.

1) First method: If the frame is corrupted, it is discarded;
If the frame is not corrupted, the packet is delivered to the network layer.
This method is used mostly in wired LANs such as Ethernet.
2) Second method: If the frame is corrupted, it is discarded;
If the frame is not corrupted, an acknowledgment is sent to the sender.
Acknowledgment is used for the purpose of both flow and error control.

3.6.2.2.1 Combination of Flow and Error Control

• Flow and error control can be combined.

• The acknowledgment that is sent for flow control can also be used for error control to tell the sender the packet has arrived uncorrupted.

• The lack of acknowledgment means that there is a problem in the sent frame.

• A frame that carries an acknowledgment is normally called an ACK to distinguish it from the data frame.

3.6.3 Connectionless and Connection-Oriented

• A DLC protocol can be either connectionless or connection-oriented.

1) Connectionless Protocol

- Frames are sent from one node to the next without any relationship between the frames; each frame is independent.
- The term connectionless does not mean that there is no physical connection (transmission medium) between the nodes; it means that there is no connection between frames.
- The frames are not numbered and there is no sense of ordering.
- Most of the data-link protocols for LANs are connectionless protocols.

2) Connection Oriented Protocol

- A logical connection should first be established between the two nodes (setup phase).
- After all frames that are somehow related to each other are transmitted (transfer phase), the logical connection is terminated (teardown phase).
- The frames are numbered and sent in order.
- If the frames are not received in order, the receiver needs to wait until all frames belonging to the same set are received and then deliver them in order to the network layer.
- Connection oriented protocols are rare in wired LANs, but we can see them in some point-to-point protocols, some wireless LANs, and some WANs.

3.7 DATA LINK LAYER PROTOCOLS

• Traditionally 2 protocols have been defined for the data-link layer to deal with flow and error control:
1) Simple Protocol and 2) Stop-and-Wait Protocol.
• The behavior of a data-link-layer protocol can be better shown as a finite state machine (FSM).
• An FSM is a machine with a finite number of states (Figure 11.6).
• The machine is always in one of the states until an event occurs.
• Each event is associated with 2 reactions:
1) Defining the list (possibly empty) of actions to be performed.
2) Determining the next state (which can be the same as the current state).
• One of the states must be defined as the initial state, the state in which the machine starts when it turns on.

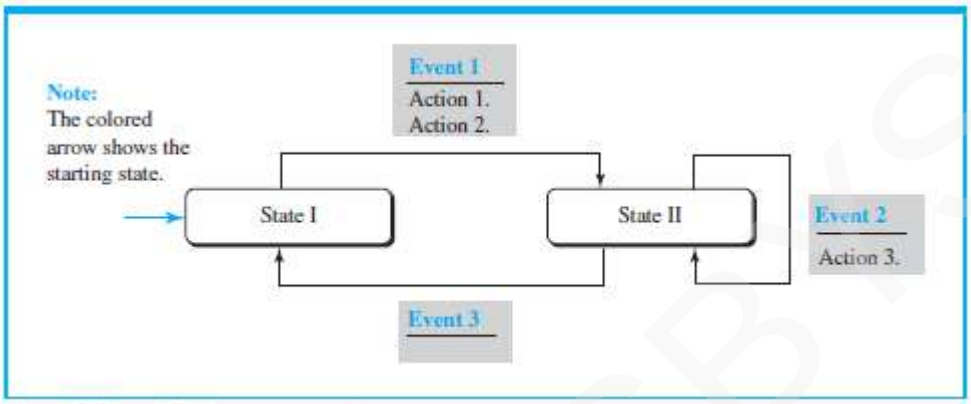


Figure 11.6 *Connectionless and connection-oriented service represented as FSMs*

3.7.1 Simplest Protocol

• Assumptions:
- The protocol has no flow-control or error-control.
- The protocol is a unidirectional protocol (in which frames are traveling in only one direction).
- The receiver can immediately handle any frame it receives.

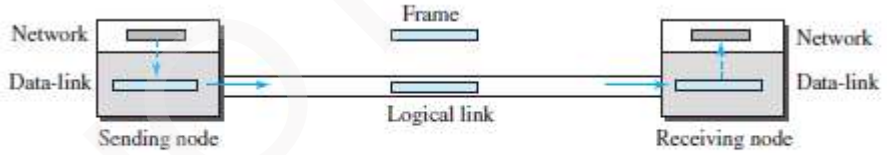


Figure 11.7 *Simple protocol*

3.7.1.1 Design

• Here is how it works (Figure 11.7):

1) At Sender

¤ The data-link-layer
→ gets data from its network-layer
→ makes a frame out of the data and
→ sends the frame.

2) At Receiver

¤ The data-link-layer
→ receives a frame from its physical layer
→ extracts data from the frame and
→ delivers the data to its network-layer.

- Data-link-layers of sender & receiver provide transmission services for their network-layers.
- Data-link-layers use the services provided by their physical layers for the physical transmission of bits.

3.7.1.2 FSMs

- Two main requirements:
  1) The sender-site cannot send a frame until its network-layer has a data packet to send.
  2) The receiver-site cannot deliver a data packet to its network-layer until a frame arrives.
- These 2 requirements are shown using two FSMs.
- Each FSM has only one state, the ready state.

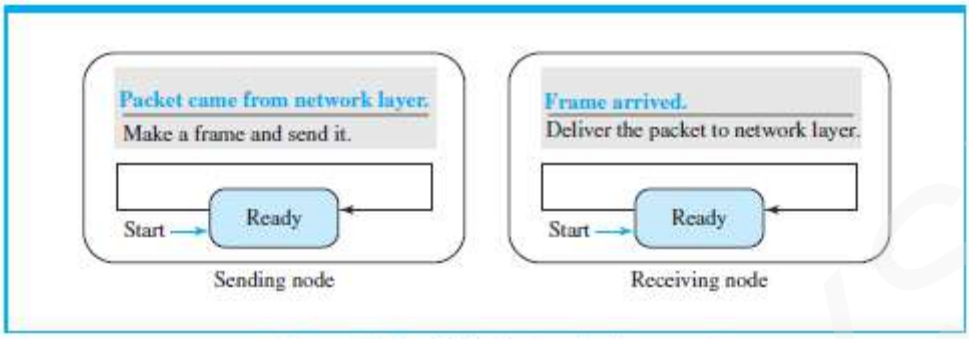


Figure 11.8 *FSMs for the simple protocol*

- Here is how it works (Figure 11.8):

1) At Sending Machine

➢ The sending machine remains in the ready state until a request comes from the process in the network layer.

➢ When this event occurs, the sending machine encapsulates the message in a frame and sends it to the receiving machine.

2) At Receiving Machine

➢ The receiving machine remains in the ready state until a frame arrives from the sending machine.

➢ When this event occurs, the receiving machine decapsulates the message out of the frame and delivers it to the process at the network layer.

Example 3.6

Figure 11.9 shows an example of communication using this protocol. It is very simple. The sender sends frames one after another without even thinking about the receiver.

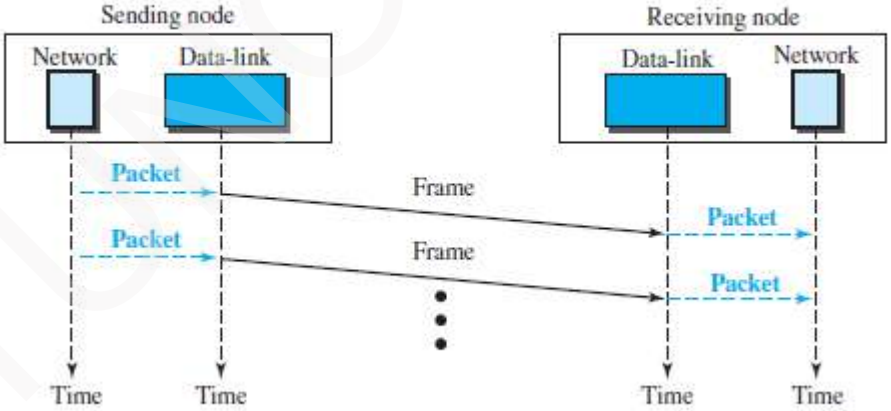


Figure 11.9 *Flow diagram*

3.7.2 Stop & Wait Protocol

- This uses both flow and error control.
- Normally, the receiver has limited storage-space.
- If the receiver is receiving data from many sources, the receiver may
  → be overloaded with frames &
  → discard the frames.
- To prevent the receiver from being overloaded with frames, we need to tell the sender to slow down.

3.7.2.1 Design

1) At Sender

- The sender
  → sends one frame & starts a timer
  → keeps a copy of the sent-frame and
  → waits for ACK-frame from the receiver (okay to go ahead).
- Then,
  1) If an ACK-frame arrives before the timer expires, the timer is stopped and the sender sends the next frame.
  Also, the sender discards the copy of the previous frame.
  2) If the timer expires before ACK-frame arrives, the sender resends the previous frame and restarts the timer

2) At Receiver

- To detect corrupted frames, a CRC is added to each data frame.
- When a frame arrives at the receiver-site, the frame is checked.
- If frame's CRC is incorrect, the frame is corrupted and discarded.
- The silence of the receiver is a signal for the sender that a frame was either corrupted or lost.

3.7.2.2 FSMs

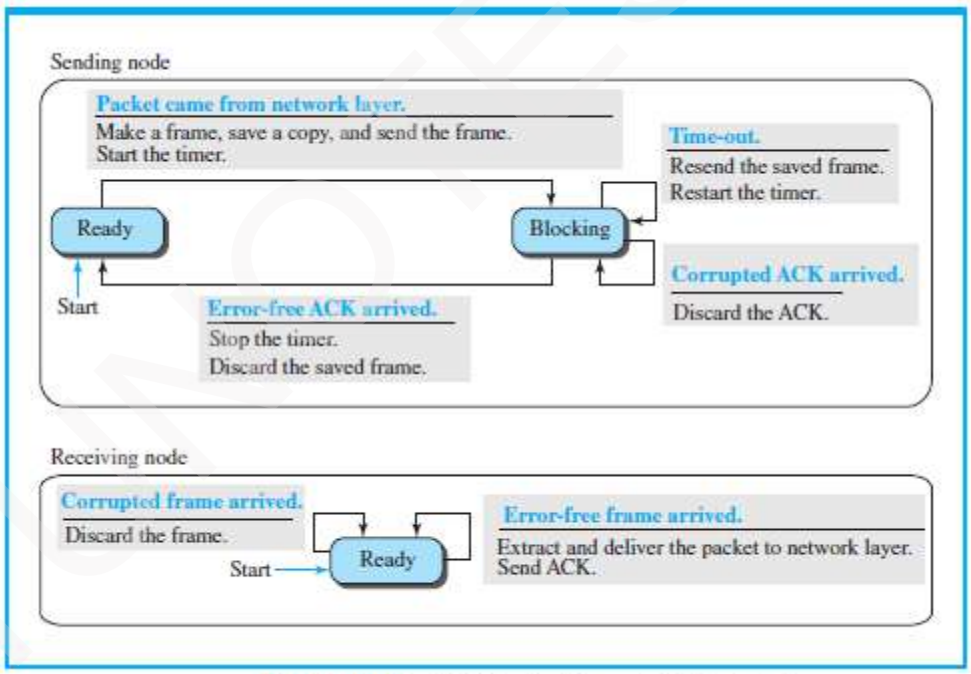


Figure 11.11 *FSM for the Stop-and-Wait protocol*

- Here is how it works (Figure 11.11):

1) Sender States

- Sender is initially in the ready state, but it can move between the ready and blocking state.

i) Ready State: When the sender is in this state, it is only waiting for a packet from the network layer.

If a packet comes from the network layer, the sender creates a frame, saves a copy of the frame, starts the only timer and sends the frame. The sender then moves to the blocking state.

ii) Blocking State: When the sender is in this state, three events can occur:

a) If a time-out occurs, the sender resends the saved copy of the frame and restarts the timer.

b) If a corrupted ACK arrives, it is discarded.

c) If an error-free ACK arrives, the sender stops the timer and discards the saved copy of the frame. It then moves to the ready state.

2) Receiver

- The receiver is always in the ready state. Two events may occur:

a) If an error-free frame arrives, the message in the frame is delivered to the network layer and an ACK is sent.

b) If a corrupted frame arrives, the frame is discarded.

Example 3.7

Figure 11.12 shows an example. The first frame is sent and acknowledged. The second frame is sent, but lost. After time-out, it is resent. The third frame is sent and acknowledged, but the acknowledgment is lost. The frame is resent. However, there is a problem with this scheme. The network layer at the receiver site receives two copies of the third packet, which is not right. In the next section, we will see how we can correct this problem using sequence numbers and acknowledgment numbers.

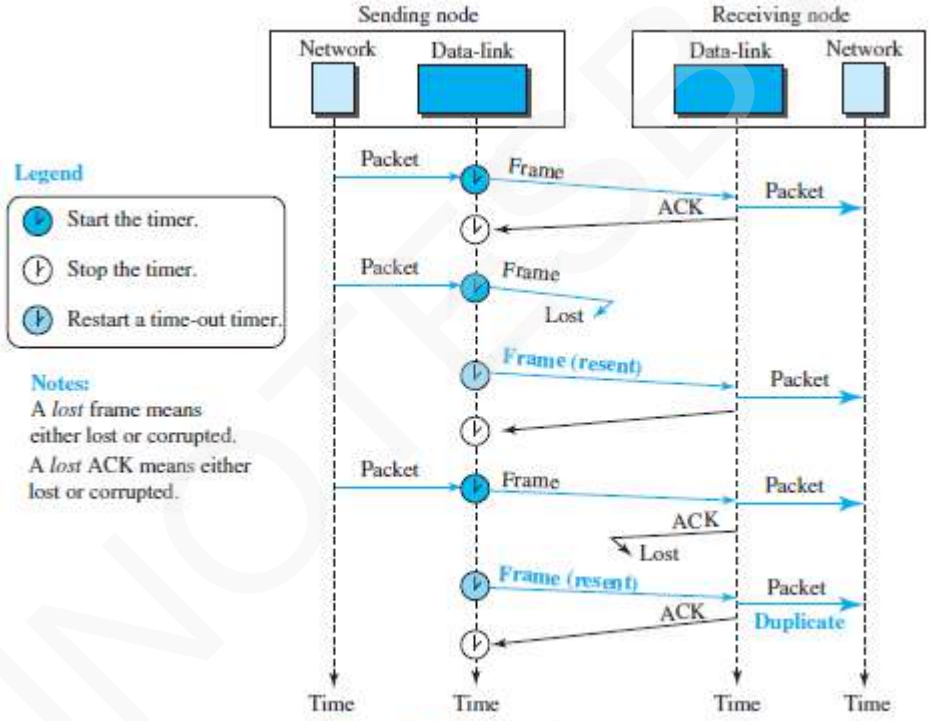


Figure 11.12 *Flow diagram*

3.7.2.3 Sequence and Acknowledgment Numbers

• Q: How to deal with corrupted-frame?

Ans: If the corrupted-frame arrives at the receiver-site, then the frame is simply discarded.

• Q: How to deal with lost-frames?

Ans: If the receiver receives out-of-order data-frame, then it means that frames were lost. .'. The lost-frames need to be resent.

• Problem in Stop and Wait protocols:

1) There is no way to identify a frame.

2) The received-frame could be the correct one, or a duplicate, or a frame out of order.

Solution: 1) Use sequence-number for each data frame.

2) Use Acknowledgment-number for each ACK frame.

Sequence Numbers

- Frames need to be numbered. This is done by using sequence-numbers.
- A sequence-number field is added to the data-frame.

Acknowledgment Numbers

- An acknowledgment-number field is added to the ACK-frame.
- Sequence numbers are 0, 1, 0, 1, 0, 1, . . .

  The acknowledgment numbers can also be 1, 0, 1, 0, 1, 0, …
- The acknowledgment-numbers always announce the sequence-number of the next frame expected by the receiver.
- For example,

  If frame-0 has arrived safely, the receiver sends an ACK-frame with acknowledgment-1 (meaning frame-1 is expected next).

Example 3.8

Figure 11.13 shows how adding sequence numbers and acknowledgment numbers can prevent duplicates. The first frame is sent and acknowledged. The second frame is sent, but lost. After time-out, it is resent. The third frame is sent and acknowledged, but the acknowledgment is lost. The frame is resent.

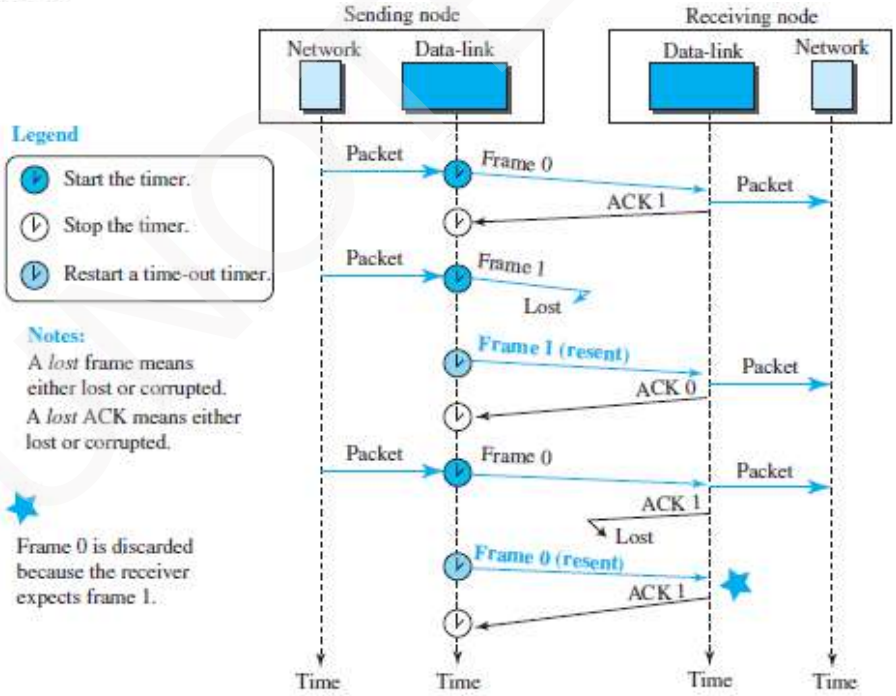


Figure 11.13 *Flow diagram*

3.7.3 Piggybacking

• A technique called piggybacking is used to improve the efficiency of the bidirectional protocols.

• The data in one direction is piggybacked with the acknowledgment in the other direction.

• In other words, when node A is sending data to node B, Node A also acknowledges the data received from node B.

3.8 High-Level Data Link Control (HDLC)

- HDLC is a bit-oriented protocol for communication over point-to-point and multipoint links.
- HDLC implements the ARQ mechanisms.

3.8.1 Configurations and Transfer Modes

- HDLC provides 2 common transfer modes that can be used in different configurations:
    1) Normal response mode (NRM)
    2) Asynchronous balanced mode (ABM).

NRM

- The station configuration is unbalanced (Figure 11.14).
- We have one primary station and multiple secondary stations.
- A primary station can send commands, a secondary station can only respond.
- The NRM is used for both point-to-point and multiple-point links.

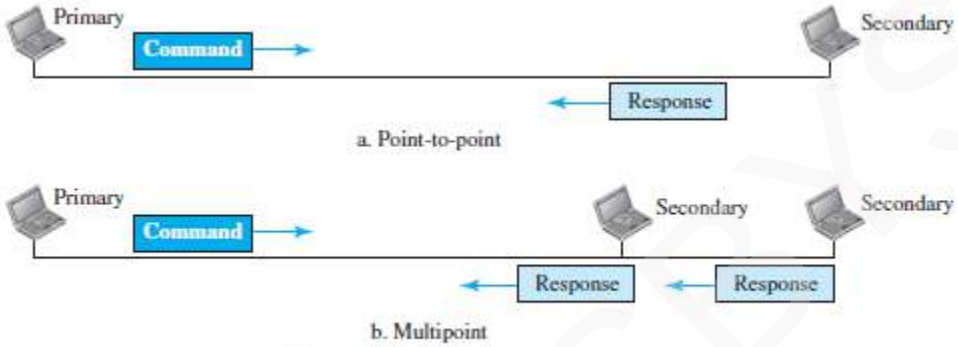


Figure 11.14 *Normal response mode*

ABM

- The configuration is balanced (Figure 11.15).
- Link is point-to-point, and each station can function as a primary and a secondary (acting as peers).
- This is the common mode today.

Figure 11.15 *Asynchronous balanced mode*

3.8.2 Framing

• To provide the flexibility necessary to support all the options possible in the modes and configurations, HDLC defines three types of frames:

1) Information frames (I-frames): are used to transport user data and control information relating to user data (piggybacking).
2) Supervisory frames (S-frames): are used only to transport control information.
3) Unnumbered frames (U-frames): are reserved for system management.
Information carried by U-frames is intended for managing the link itself.

• Each type of frame serves as an envelope for the transmission of a different type of message.

3.8.2.1 Frame Format

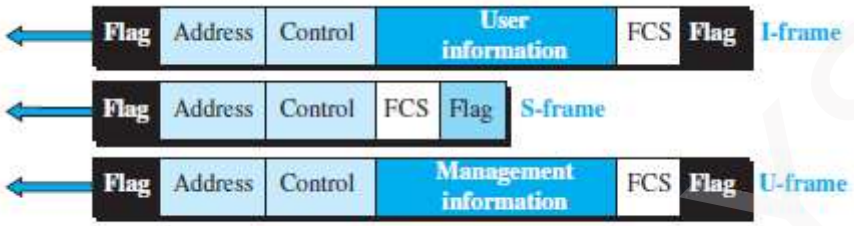


Figure 11.16 *HDLC frames*

• Various fields of HDLC frame are:

1) Flag Field
- This field has a synchronization pattern 01111110.
- This field identifies both the beginning and the end of a frame.

2) Address Field
- This field contains the address of the secondary station.
- If a primary station created the frame, it contains a to-address.
- If a secondary creates the frame, it contains a from-address.
- This field can be 1 byte or several bytes long, depending on the needs of the network.

3) Control Field
- This field is one or two bytes used for flow and error control.

4) Information Field
- This field contains the user's data from the network-layer or management information.
- Its length can vary from one network to another.

5) FCS Field
- This field is the error-detection field. (FCS → Frame Check Sequence)
- This field can contain either a 2- or 4-byte standard CRC.

3.8.2.1.1 Control Fields of HDLC Frames

• The control field determines the type of frame and defines its functionality (Figure 11.17).

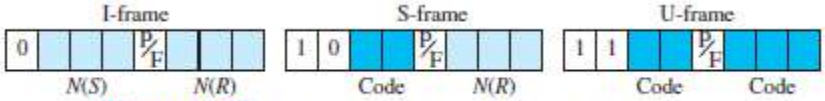


Figure 11.17 *Control field format for the different frame types*

1) Control Field for I-Frames

• I-frames are designed to carry user data from the network-layer.
• In addition, they can include flow and error-control information (piggybacking).
• The subfields in the control field are:

1) The first bit defines the type.
   If the first bit of the control field is 0, this means the frame is an I-frame.
2) The next 3 bits N(S) define the sequence-number of the frame.
   With 3 bits, we can define a sequence-number between 0 and 7
3) The last 3 bits N(R) correspond to the acknowledgment-number when piggybacking is used.
4) The single bit between N(S) and N(R) is called the P/F bit.
   The P/F field is a single bit with a dual purpose. It can mean poll or final.
   i) It means poll when the frame is sent by a primary station to a secondary (when the address field contains the address of the receiver).
   ii) It means final when the frame is sent by a secondary to a primary (when the address field contains the address of the sender).

2) Control Field for S-Frames

• Supervisory frames are used for flow and error-control whenever piggybacking is either impossible or inappropriate (e.g., when the station either has no data of its own to send or needs to send a command or response other than an acknowledgment).
• S-frames do not have information fields.
• The subfields in the control field are:

1) If the first 2 bits of the control field is 10, this means the frame is an S-frame.
2) The last 3 bits N(R) corresponds to the acknowledgment-number (ACK) or negative acknowledgment-number (NAK).
3) The 2 bits called code is used to define the type of S-frame itself.
   With 2 bits, we can have four types of S-frames:
   1) Receive Ready (RR) = 00
   ¤ This acknowledges the receipt of frame or group of frames.
   ¤ The value of N(R) is the acknowledgment-number.
   2) Receive Not Ready (RNR) = 10
   ¤ This is an RR frame with 1 additional function:
      i) It announces that the receiver is busy and cannot receive more frames.
   ¤ It acts as congestion control mechanism by asking the sender to slow down.
   ¤ The value of N(R) is the acknowledgment-number.
   3) ReJect (REJ) = 01
   ¤ It is a NAK frame used in Go-Back-N ARQ to improve the efficiency of the process.
   ¤ It informs the sender, before the sender time expires, that the last frame is lost or damaged.
   ¤ The value of N(R) is the negative acknowledgment-number.
   4) Selective REJect (SREJ) = 11
   ¤ This is a NAK frame used in Selective Repeat ARQ.
   ¤ The value of N(R) is the negative acknowledgment-number.

3) Control Field for U-Frames

• Unnumbered frames are used to exchange session management and control information between connected devices.
• U-frames contain an information field used for system management information, but not user data.
• Much of the information carried by U-frames is contained in codes included in the control field.
• U-frame codes are divided into 2 sections:
  i) A 2-bit prefix before the P/F bit
  ii) A 3-bit suffix after the P/F bit.
• Together, these two segments (5 bits) can be used to create up to 32 different types of U-frames.

Example 3.9

Figure 11.18 shows how U-frames can be used for connection establishment and connection release. Node A asks for a connection with a set asynchronous balanced mode (SABM) frame; node B gives a positive response with an unnumbered acknowledgment (UA) frame. After these two exchanges, data can be transferred between the two nodes (not shown in the figure). After data transfer, node A sends a DISC (disconnect) frame to release the connection; it is confirmed by node B responding with a UA (unnumbered acknowledgment).

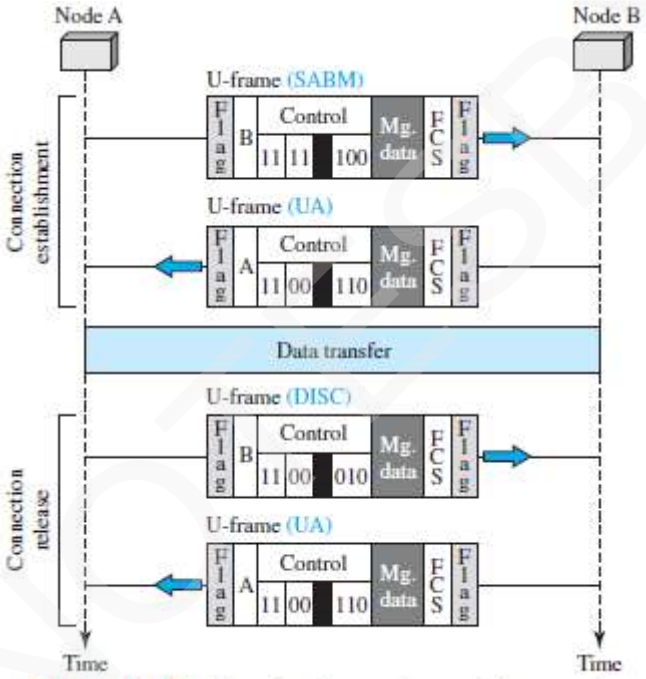


Figure 11.18 *Example of connection and disconnection*

3.9 POINT-TO-POINT PROTOCOL (PPP)

• PPP is one of the most common protocols for point-to-point access.
• Today, millions of Internet users who connect their home computers to the server of an ISP use PPP.

3.9.1 Framing

• PPP uses a character-oriented (or byte-oriented) frame (Figure 11.20).

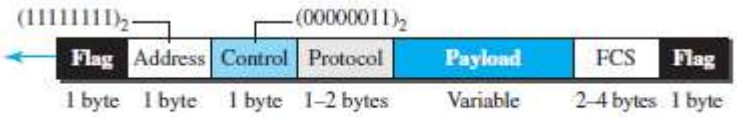


Figure 11.20 *PPP frame format*

• Various fields of PPP frame are:

1) Flag
- This field has a synchronization pattern 01111110.
- This field identifies both the beginning and the end of a frame.

2) Address
- This field is set to the constant value 11111111 (broadcast address).

3) Control
- This field is set to the constant value 00000011 (imitating unnumbered frames in HDLC).
- PPP does not provide any flow control.
- Error control is also limited to error detection.

4) Protocol
- This field defines what is being carried in the payload field.
- Payload field carries either i) user data or ii) other control information.
- By default, size of this field = 2 bytes.

5) Payload field
- This field carries either i) user data or ii) other control information.
- By default, maximum size of this field = 1500 bytes.
- This field is byte-stuffed if the flag-byte pattern appears in this field.
- Padding is needed if the payload-size is less than the maximum size.

6) FCS
- This field is the PPP error-detection field.
- This field can contain either a 2- or 4-byte standard CRC.

3.9.1.1 Byte Stuffing

• Since PPP is a byte-oriented protocol, the flag in PPP is a byte that needs to be escaped whenever it appears in the data section of the frame.
• The escape byte is 01111101, which means that every time the flag like pattern appears in the data, this extra byte is stuffed to tell the receiver that the next byte is not a flag.
• Obviously, the escape byte itself should be stuffed with another escape byte.

3.9.2 Transition Phases

• The transition diagram starts with the dead state (Figure 11.21).

1) Dead State

➢ In dead state, there is no active carrier and the line is quiet.

2) Establish State

➢ When 1 of the 2 nodes starts communication, the connection goes into the establish state.

➢ In establish state, options are negotiated between the two parties.

3) Authenticate State

➢ If the 2 parties agree that they need authentication,

Then the system needs to do authentication;

Otherwise, the parties can simply start communication.

4) Open State

➢ Data transfer takes place in the open state.

5) Terminate State

➢ When 1 of the endpoints wants to terminate connection, the system goes to terminate state.

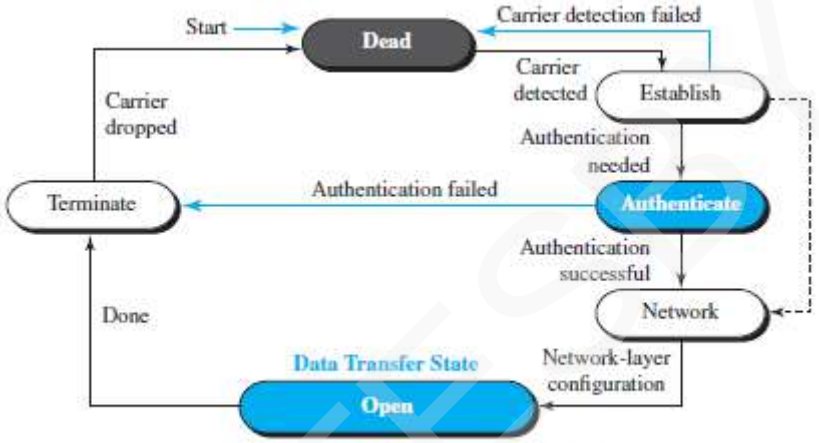


Figure 11.21 *Transition phases*

MODULE-WISE QUESTIONS

MODULE 3: ERROR-DETECTION AND CORRECTION

1) Explain two types of errors (4*)
2) Compare error detection vs. error correction (2)
3) Explain error detection using block coding technique. (10*)
4) Explain hamming distance for error detection (6*)
5) Explain parity-check code with block diagram. (6*)
6) Explain CRC with block diagram & an example. (10*)
7) Write short notes on polynomial codes. (5*)
8) Explain internet checksum algorithm along with an example. (6*)
9) Explain the following:
   i) Fletcher checksum and ii) Adler checksum (8)
10) Explain various FEC techniques. (6)

MODULE 3(CONT.): DATA LINK CONTROL

1) Explain two types of frames. (2)
2) Explain character oriented protocol. (6*)
3) Explain the concept of byte stuffing and unstuffing with example. (6*)
4) Explain bit oriented protocol. (6*)
5) Differentiate between character oriented and bit oriented format for Framing. (6*)
6) Compare flow control and error control. (4)
7) With a neat diagram, explain the design of the simplest protocol with no flow control. (6)
8) Write algorithm for sender site and receiver site for the simplest protocol. (6)
9) Explain Stop-and-Wait protocol (8*)
10) Explain the concept of Piggybacking (2*)
11) Explain in detail HDLC frame format. (8*)
12) Explain 3 type of frame used in HDLC (8*)
13) With a neat schematic, explain the frame structure of PPP protocol. (8*)
14) Explain framing and transition phases in Point-to-Point Protocol. (8*)

MODULE 4: TABLE OF CONTENTS

MODULE 4: MULTIPLE ACCESS

4.1 Introduction

- When nodes use shared-medium, we need multiple-access protocol to coordinate access to medium.
- Analogy:
  - This problem is similar to the rules of speaking in an assembly.
  - We need to ensure
    - → Each people has right to speak.
    - → Two people do not speak at the same time
    - → Two people do not interrupt each other (i.e. Collision Avoidance)
- Many protocols have been designed to handle access to a shared-link (Figure 12.1).
- These protocols belong to a sublayer in the data-link layer called Media Access Control (MAC).
  1) Four random-access protocols (or Contention Methods):
     i) ALOHA ii) CSMA iii) CSMA/CD iv) CSMA/CA
     These protocols are mostly used in LANs and WANs.
  2) Three controlled-access protocols:
     i) Reservation ii) Polling iii) Token-passing
     Some of these protocols are used in LANs.
  3) Three channelization protocols:
     i) FDMA ii) TDMA iii) CDMA
     These protocols are used in cellular telephony.

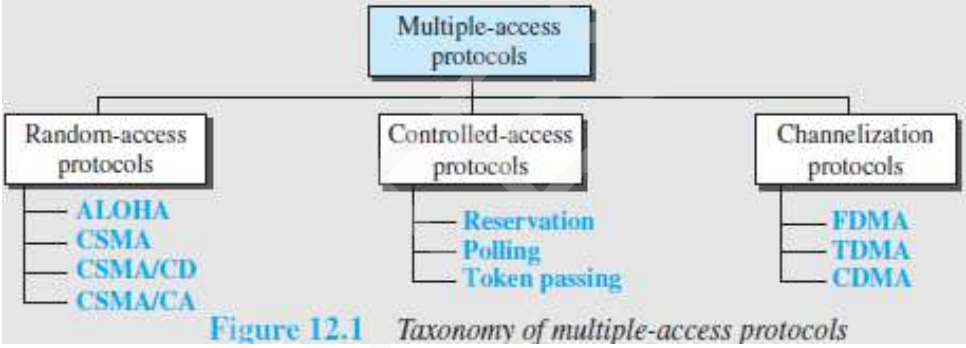


Figure 12.1 *Taxonomy of multiple-access protocols*

4.2 RANDOM ACCESS PROTOCOL

- No station is superior to another station.
- No station is assigned control over other station.
- To send the data, a station uses a procedure to make a decision on whether or not to send.
- This decision depends on the state of the medium: idle or busy.
- This is called Random Access because
  - → Transmission is random among the stations.
  - → There is no scheduled-time for a station to transmit.
- This is called Contention Method because
  - → Stations compete with one another to access the medium.
- If more than one station tries to send,
  there is an access-conflict (i.e. collision) and the frames will be destroyed.
- Each station follows a procedure that answers the following questions:
  1) When can the station access the medium?
  2) What can the station do if the medium is busy?
  3) How can the station determine the success or failure of the transmission?
  4) What can the station do if there is a collision?
- Four random-access protocols (or Contention methods):
  1) ALOHA
  2) CSMA (Carrier Sense Multiple Access)
  3) CSMA/CD (Carrier Sense Multiple Access with Collision-detection)
  4) CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance)

4.2.1 ALOHA

- ALOHA was designed for a wireless LAN, but it can be used on any shared medium.
- Since the medium is shared between the stations, there is possibility of collisions.
- When 2 or more stations send the data simultaneously, there is possibility of collision & data loss.

4.2.1.1 Pure ALOHA

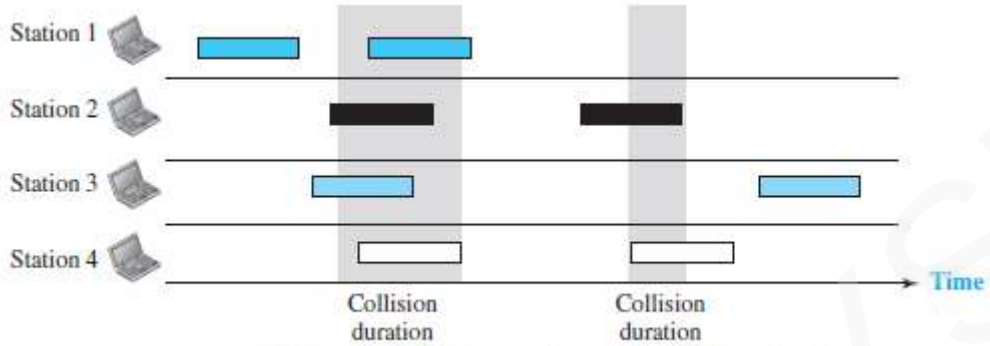


Figure 12.2 *Frames in a pure ALOHA network*

- Here is how it works (Figure 12.2):
  1) The sender sends a frame & starts the timer.
  2) The receiver receives the frame and responds with an acknowledgment.
  3) If the acknowledgment does not arrive after a time-out period, the sender resends the frame. The sender assumes that the frame (or the acknowledgment) has been destroyed.
  4) Since the medium is shared between the stations, there is possibility of collisions.
  5) If two stations try to resend the frames after the time-out, the frames will collide again.
  6) Two methods to deal with collision:
     1) Randomness
     ¤ When the time-out period passes, each station waits a random amount of time before resending the frame. This time is called back-off time T_B.
     ¤ The randomness will help avoid more collisions.
     2) Limit Maximum Retransmission
     ¤ This method prevents congestion by reducing the number of retransmitted frames.
     ¤ After a maximum number of retransmission-attempts K_{max}, a station must give up and try later (Figure 12.3).

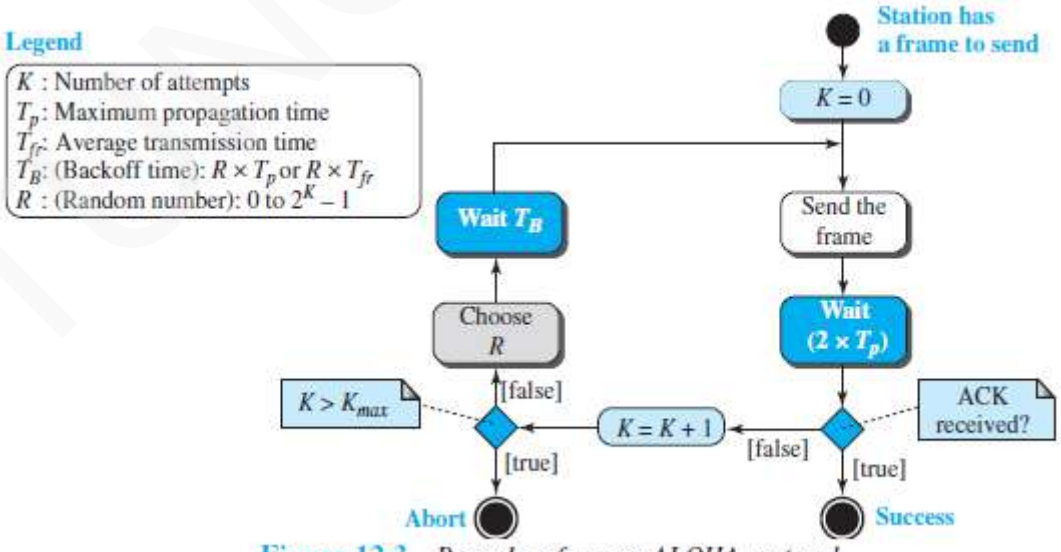


Figure 12.3 *Procedure for pure ALOHA protocol*

4.2.1.1.1 Vulnerable Time

- The vulnerable-time is defined as a time during which there is a possibility of collision.

Pure ALOHA vulnerable time = $2 \times T_{fr}$

where T_{fr} = Frame transmission time

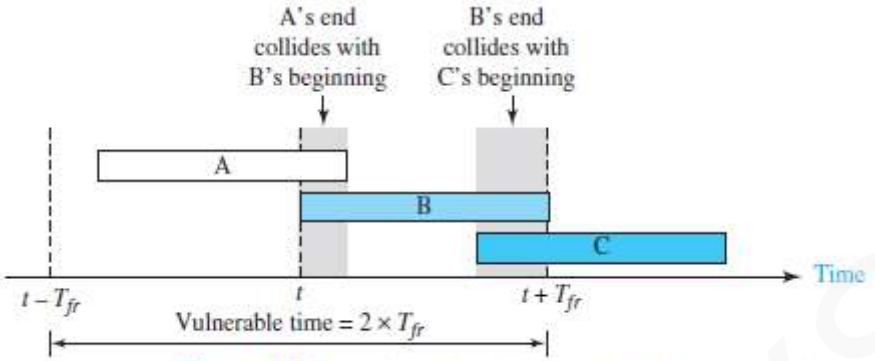


Figure 12.4 *Vulnerable time for pure ALOHA protocol*

- In Figure 12.4,
  - If station B sends a frame between t-T_{fr} and t, this leads to a collision between the frames from station A and station B.
  - If station C sends a frame between t and t+T_{fr}, this leads to a collision between the frames from station A and station C.

Example 4.1

A pure ALOHA network transmits 200-bit frames on a shared channel of 200 kbps. What is the requirement to make this frame collision-free?

Solution

Average frame transmission time T_{fr} is 200 bits/200 kbps or 1 ms. The vulnerable time is 2×1 ms = 2 ms. This means no station should send later than 1 ms before this station starts transmission and no station should start sending during the period (1 ms) that this station is sending.

4.2.1.1.2 Throughput

- The average number of successful transmissions is given by

$$S = G \times e^{-2G}$$

where G = average no. of frames in one frame transmission time (T_{fr})

- For G = 1, the maximum throughput S_{max} = 0.184.
- In other words, out of 100 frames, 18 frames reach their destination successfully.

Example 4.2

A pure ALOHA network transmits 200-bit frames on a shared channel of 200 kbps. What is the throughput if the system (all stations together) produces

a. 1000 frames per second? b. 500 frames per second? c. 250 frames per second?

Solution

The frame transmission time is 200/200 kbps or 1 ms.

a. If the system creates 1000 frames per second, or 1 frame per millisecond, then $G = 1$. In this case $S = G \times e^{-2G} = 0.135$ (13.5 percent). This means that the throughput is $1000 \times 0.135 = 135$ frames. Only 135 frames out of 1000 will probably survive.

b. If the system creates 500 frames per second, or 1/2 frames per millisecond, then $G = 1/2$. In this case $S = G \times e^{-2G} = 0.184$ (18.4 percent). This means that the throughput is $500 \times 0.184 = 92$ and that only 92 frames out of 500 will probably survive. Note that this is the *maximum* throughput case, percentagewise.

c. If the system creates 250 frames per second, or 1/4 frames per millisecond, then $G = 1/4$. In this case $S = G \times e^{-2G} = 0.152$ (15.2 percent). This means that the throughput is $250 \times 0.152 = 38$. Only 38 frames out of 250 will probably survive.

4.2.1.2 Slotted ALOHA

- Slotted ALOHA was invented to improve the efficiency of pure ALOHA.
- The time is divided into time-slots of T_{fr} seconds (Figure 12.5).
- The stations are allowed to send only at the beginning of the time-slot.

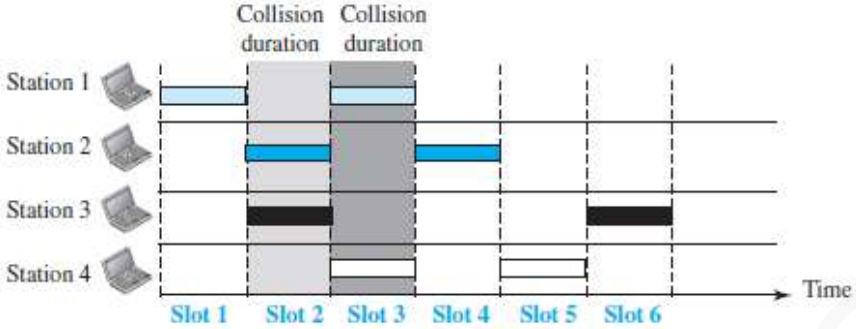


Figure 12.5 *Frames in a slotted ALOHA network*

- If a station misses the time-slot, the station must wait until the beginning of the next time-slot.
- If 2 stations try to resend at beginning of the same time-slot, the frames will collide again (Fig 12.6).

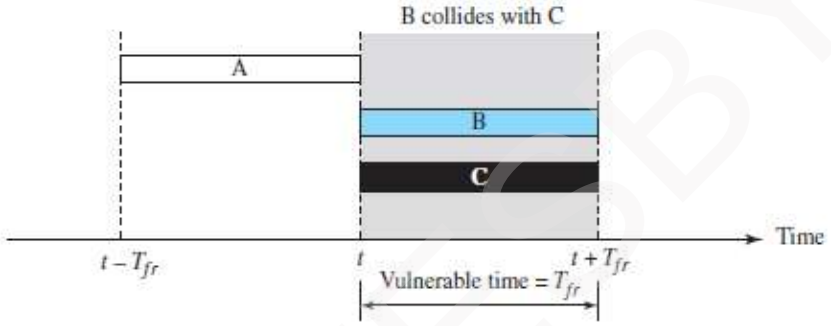


Figure 12.6 *Vulnerable time for slotted ALOHA protocol*

- The vulnerable time is given by:

vulnerable time = T_{fr}

4.2.1.2.1 Throughput

- The average number of successful transmissions is given by

$$S = G \times e^{-G}$$

- For G = 1, the maximum throughput S_{max} = 0.368.
- In other words, out of 100 frames, 36 frames reach their destination successfully.

Example 4.3

A slotted ALOHA network transmits 200-bit frames using a shared channel with a 200-kbps bandwidth. Find the throughput if the system (all stations together) produces

a. 1000 frames per second.
b. 500 frames per second.
c. 250 frames per second.

Solution

This situation is similar to the previous exercise except that the network is using slotted ALOHA instead of pure ALOHA. The frame transmission time is 200/200 kbps or 1 ms.

a. In this case G is 1. So $S = G \times e^{-G} = 0.368$ (36.8 percent). This means that the throughput is $1000 \times 0.0368 = 368$ frames. Only 368 out of 1000 frames will probably survive. Note that this is the maximum throughput case, percentagewise.
b. Here G is 1/2. In this case $S = G \times e^{-G} = 0.303$ (30.3 percent). This means that the throughput is $500 \times 0.0303 = 151$. Only 151 frames out of 500 will probably survive.
c. Now G is 1/4. In this case $S = G \times e^{-G} = 0.195$ (19.5 percent). This means that the throughput is $250 \times 0.195 = 49$. Only 49 frames out of 250 will probably survive.

4.2.2 CSMA

- CSMA was developed to minimize the chance of collision and, therefore, increase the performance.
- CSMA is based on the principle "sense before transmit" or "listen before talk."
- Here is how it works:
  1) Each station checks the state of the medium: idle or busy.
  2) i) If the medium is idle, the station sends the data.
     ii) If the medium is busy, the station defers sending.
- CSMA can reduce the possibility of collision, but it cannot eliminate it.

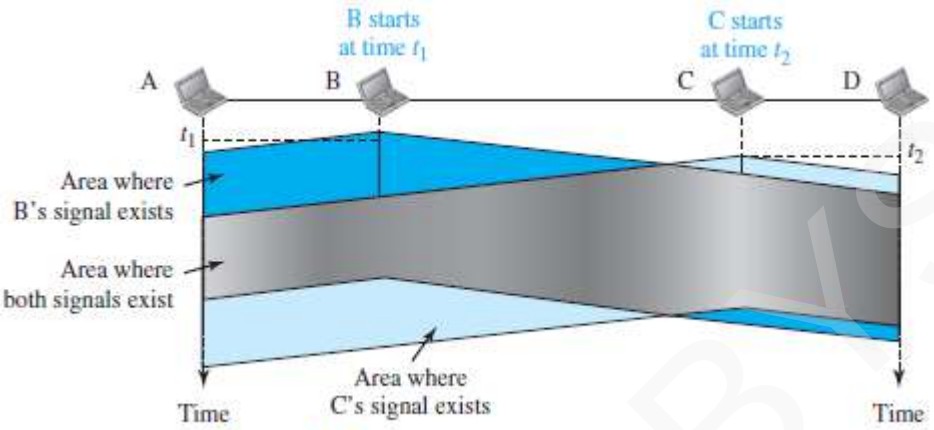


Figure 12.7 *Space/time model of a collision in CSMA*

- The possibility of collision still exists.
  For example:
  When a station sends a frame, it still takes time
  → for the first bit to reach every station and
  → for every station to sense it.
- For example: In Figure 12.7,
  - At time t1, station B senses & finds the medium idle, so sends a frame.
  - At time t2, station C senses & finds the medium idle, so sends a frame.
  - The 2 signals from both stations B & C collide and both frames are destroyed.

4.2.2.1 Vulnerable Time

- The vulnerable time is the propagation time T_p (Figure 12.8).
- The propagation time is the time needed for a signal to propagate from one end of the medium to the other.
- Collision occurs when
  → a station sends a frame, and
  → other station also sends a frame during propagation time
- If the first bit of the frame reaches the end of the medium, every station will refrain from sending.

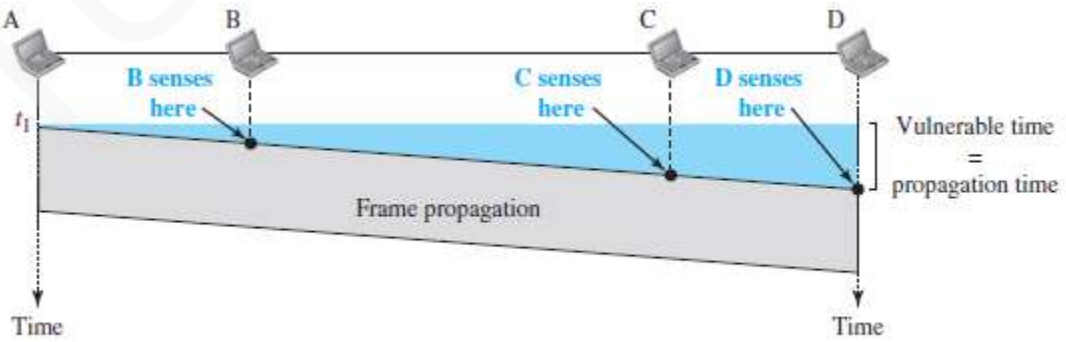


Figure 12.8 *Vulnerable time in CSMA*

4.2.2.2 Persistence Methods

• Q: What should a station do if the channel is busy or idle?
Three methods can be used to answer this question:
1) 1-persistent method 2) Non-persistent method and 3) p-persistent method

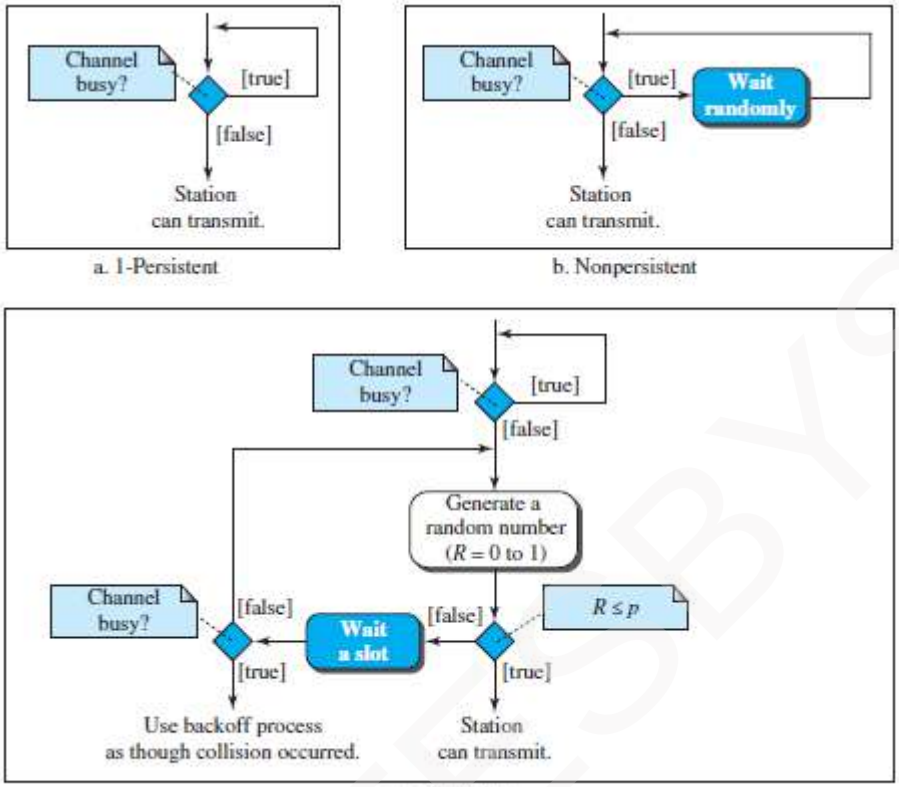


Figure 12.10 *Flow diagram for three persistence methods*

1) 1-Persistent

• Before sending a frame, a station senses the line (Figure 12.10a).
i) If the line is idle, the station sends immediately (with probability = 1).
ii) If the line is busy, the station continues sensing the line.
• This method has the highest chance of collision because 2 or more stations:
→ may find the line idle and
→ send the frames immediately.

2) Non-Persistent

• Before sending a frame, a station senses the line (Figure 12.10b).
i) If the line is idle, the station sends immediately.
ii) If the line is busy, the station waits a random amount of time and then senses the line again.
• This method reduces the chance of collision because 2 or more stations:
→ will not wait for the same amount of time and
→ will not retry to send simultaneously.

3) P-Persistent

• This method is used if the channel has time-slots with a slot-duration equal to or greater than the maximum propagation time (Figure 12.10c).
• Advantages:
i) It combines the advantages of the other 2 methods.
ii) It reduces the chance of collision and improves efficiency.
• After the station finds the line idle, it follows these steps:
1) With probability p, the station sends the frame.
2) With probability q=1-p, the station waits for the beginning of the next time-slot and checks the line again.
i) If line is idle, it goes to step 1.
ii) If line is busy, it assumes that collision has occurred and uses the back off procedure.

4.2.3 CSMA/CD

- Disadvantage of CSMA: CSMA does not specify the procedure after a collision has occurred.
  Solution: CSMA/CD enhances the CSMA to handle the collision.
- Here is how it works (Figure 12.12):
  1) A station
     → sends the frame &
     → then monitors the medium to see if the transmission was successful or not.
  2) If the transmission was unsuccessful (i.e. there is a collision), the frame is sent again.

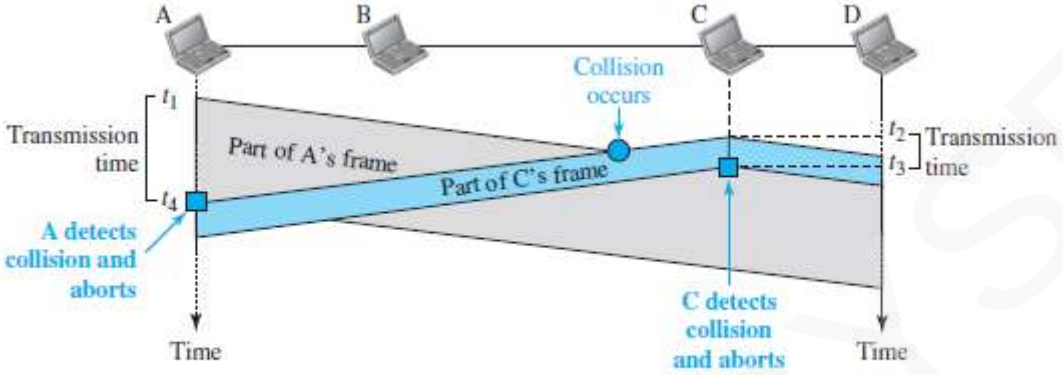


Figure 12.12 *Collision and abortion in CSMA/CD*

- In the Figure 12.11,
  - At time t_1, station A has executed its procedure and starts sending the bits of its frame.
  - At time t_2, station C has executed its procedure and starts sending the bits of its frame.
  - The collision occurs sometime after time t_2.
  - Station C detects a collision at time t_3 when it receives the first bit of A's frame.
    Station C immediately aborts transmission.
  - Station A detects collision at time t_4 when it receives the first bit of C's frame.
    Station A also immediately aborts transmission.
- Station A transmits for the duration $t_4 - t_1$.
  Station C transmits for the duration $t_3 - t_2$.
- For the protocol to work:
  The length of any frame divided by the bit rate must be more than either of these durations.

4.2.3.1 Minimum Frame Size

- For CSMA/CD to work, we need to restrict the frame-size.
- Before sending the last bit of the frame, the sender must
  → detect a collision and
  → abort the transmission.
- This is so because the sender
  → does not keep a copy of the frame and
  → does not monitor the line for collision-detection.
- Frame transmission time T_{fr} is given by
  $T_{fr} = 2T_p$ where T_p=maximum propagation time

Example 4.4

A network using CSMA/CD has a bandwidth of 10 Mbps. If the maximum propagation time (including the delays in the devices and ignoring the time needed to send a jamming signal, as we see later) is 25.6 μs, what is the minimum size of the frame?

Solution

The minimum frame transmission time is $T_{fr} = 2 \times T_p = 51.2$ μs. This means, in the worst case, a station needs to transmit for a period of 51.2 μs to detect the collision. The minimum size of the frame is 10 Mbps × 51.2 μs = 512 bits or 64 bytes. This is actually the minimum size of the frame for Standard Ethernet, as we will see later in the chapter.

4.2.3.2 Procedure

• CSMA/CD is similar to ALOHA with 2 differences (Figure 12.13):

1) Addition of the persistence process.
   ¤ We need to sense the channel before sending the frame by using non-persistent, 1-persistent or p-persistent.
2) Frame transmission.
   i) In ALOHA, first the entire frame is transmitted and then acknowledgment is waited for.
   ii) In CSMA/CD, transmission and collision-detection is a continuous process.

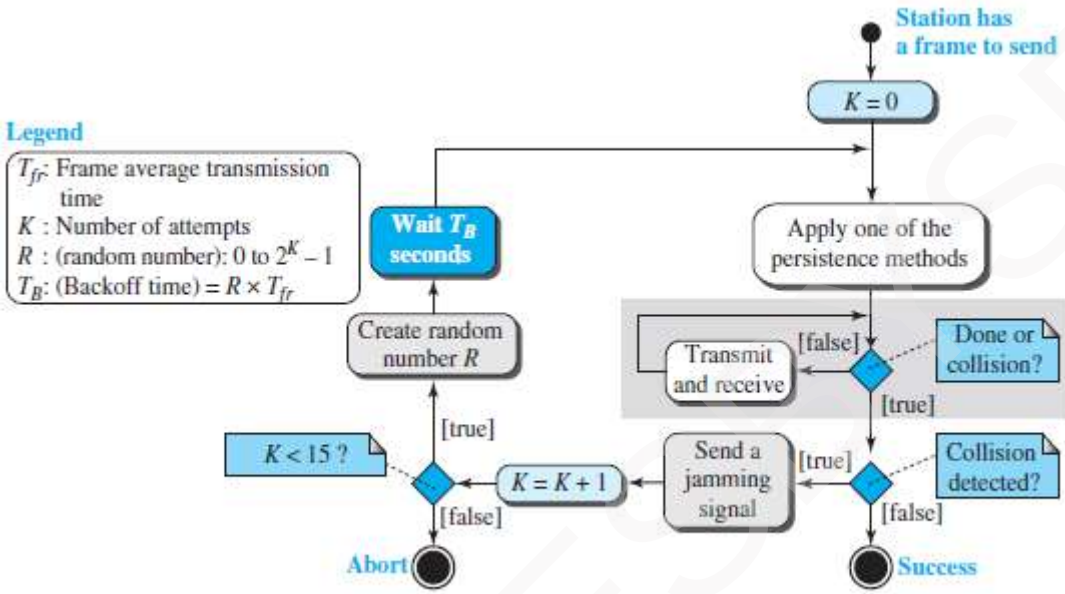


Figure 12.13 *Flow diagram for the CSMA/CD*

4.2.3.3 Energy Level

• In a channel, the energy-level can have 3 values: 1) Zero 2) Normal and 3) Abnormal.

1) At zero level, the channel is idle (Figure 12.14).
2) At normal level, a station has successfully captured the channel and is sending its frame.
3) At abnormal level, there is a collision and the level of the energy is twice the normal level.

• A sender needs to monitor the energy-level to determine if the channel is

→ Idle
→ Busy or
→ Collision mode

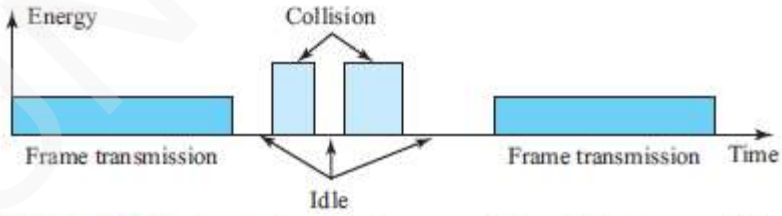


Figure 12.14 *Energy level during transmission, idleness, or collision*

4.2.3.4 Throughput

• The throughput of CSMA/CD is greater than pure or slotted ALOHA.
• The maximum throughput is based on

→ different value of G
→ persistence method used (non-persistent, 1-persistent, or p-persistent) and
→ 'p' value in the p-persistent method.

• For 1-persistent method, the maximum throughput is 50% when G =1.
• For non-persistent method, the maximum throughput is 90% when G is between 3 and 8.

4.2.4 CSMA/CA

• Here is how it works (Figure 12.15):

1) A station needs to be able to receive while transmitting to detect a collision.
   i) When there is no collision, the station receives one signal: its own signal.
   ii) When there is a collision, the station receives 2 signals:
      a) Its own signal and
      b) Signal transmitted by a second station.
2) To distinguish b/w these 2 cases, the received signals in these 2 cases must be different.

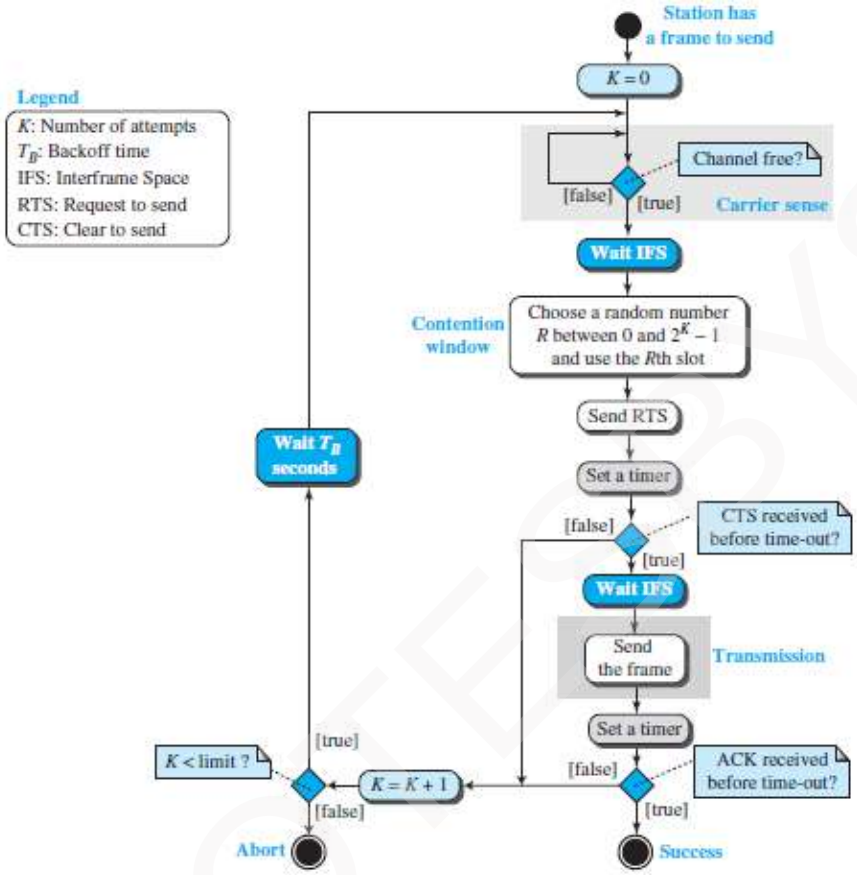


Figure 12.15 *Flow diagram of CSMA/CA*

• CSMA/CA was invented to avoid collisions on wireless networks.
• Three methods to avoid collisions (Figure 12.16):

1) Interframe space 2) Contention window and 3) Acknowledgments

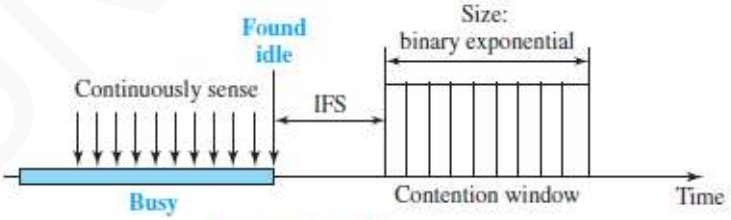


Figure 12.16 *Contention window*

1) Interframe Space (IFS)

• Collisions are avoided by deferring transmission even if the channel is found idle.
• When the channel is idle, the station does not send immediately.
   Rather, the station waits for a period of time called the inter-frame space or IFS.
• After the IFS time,
   if the channel is still idle,
      then, the station waits for the contention-time &
         finally, the station sends the frame.
• IFS variable can also be used to prioritize stations or frame types.
   For example, a station that is assigned a shorter IFS has a higher priority.

2) Contention Window

- The contention-window is an amount of time divided into time-slots.
- A ready-station chooses a random-number of slots as its wait time.
- In the window, the number of slots changes according to the binary exponential back-off strategy.
- For example:
  - At first time, number of slots is set to one slot and
  - Then, number of slots is doubled each time if the station cannot detect an idle channel.

3) Acknowledgment

- There may be a collision resulting in destroyed-data.
- In addition, the data may be corrupted during the transmission.
- To help guarantee that the receiver has received the frame, we can use
  - i) Positive acknowledgment and
  - ii) Time-out timer

4.2.4.1 Frame Exchange Time Line

- Two control frames are used:
  - 1) Request to send (RTS)
  - 2) Clear to send (CTS)
- The procedure for exchange of data and control frames in time (Figure 12.17):
  - 1) The source senses the medium by checking the energy level at the carrier frequency.
    - ii) If the medium is idle,
      - then the source waits for a period of time called the DCF interframe space (DIFS);
        - finally, the source sends a RTS.
  - 2) The destination
    - → receives the RTS
    - → waits a period of time called the short interframe space (SIFS)
    - → sends a control frame CTS to the source.
      - CTS indicates that the destination station is ready to receive data.
  - 3) The source
    - → receives the CTS
    - → waits a period of time SIFS
    - → sends a data to the destination
  - 4) The destination
    - → receives the data
    - → waits a period of time SIFS
    - → sends a acknowledgment ACK to the source.
      - ACK indicates that the destination has been received the frame.

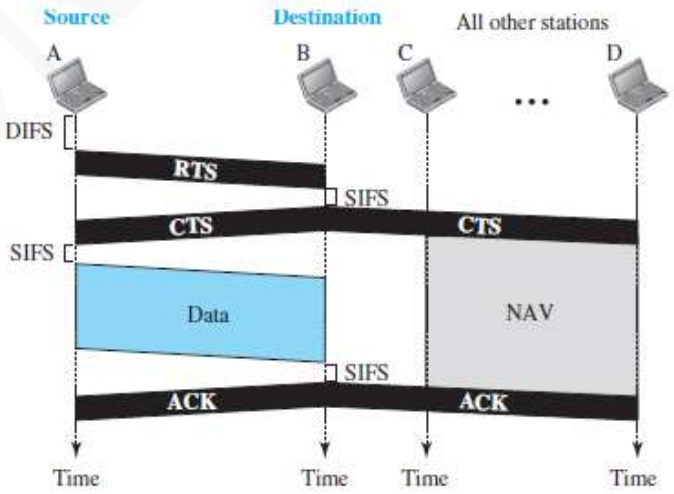


Figure 12.17 *CSMA/CA and NAV*

4.2.4.2 Network Allocation Vector

• When a source-station sends an RTS, it includes the duration of time that it needs to occupy the channel.
• The remaining stations create a timer called a network allocation vector (NAV).
• NAV indicates waiting time to check the channel for idleness.
• Each time a station accesses the system and sends an RTS frame, other stations start their NAV.

4.2.4.3 Collision during Handshaking

• Two or more stations may try to send RTS at the same time.
• These RTS may collide.
• The source assumes there has been a collision if it has not received CTS from the destination.
• The backoff strategy is employed, and the source tries again.

4.2.4.4 Hidden Station Problem

• Figure 12.17 also shows that the RTS from B reaches A, but not C.
• However, because both B and C are within the range of A, the CTS reaches C.
• Station C knows that some hidden station is using the channel and refrains from transmitting until that duration is over.

4.2.4.5 CSMA/CA and Wireless Networks

• CSMA/CA was mostly intended for use in wireless networks.
• However, it is not sophisticated enough to handle some particular issues related to wireless networks, such as hidden terminals or exposed terminals.

4.3 CONTROLLED ACCESS PROTOCOLS

- Here, the stations consult one another to find which station has the right to send.
- A station cannot send unless it has been authorized by other stations.
- Three popular controlled-access methods are: 1) Reservation 2) Polling 3) Token Passing

4.3.1 Reservation

- Before sending data, each station needs to make a reservation of the medium.
- Time is divided into intervals.
- In each interval, a reservation-frame precedes the data-frames.
- If no. of stations = N, then there are N reservation mini-slots in the reservation-frame.
- Each mini-slot belongs to a station.
- When a station wants to send a data-frame, it makes a reservation in its own minislot.
- The stations that have made reservations can send their data-frames.

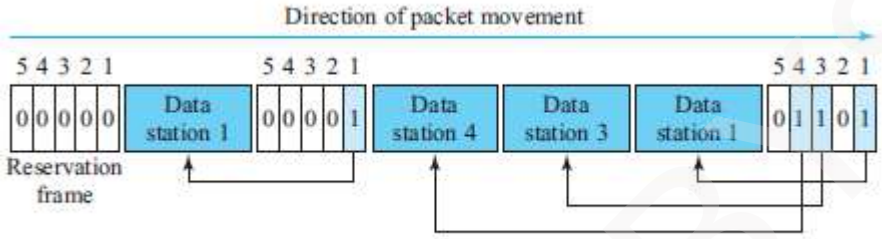


Figure 12.18 *Reservation access method*

- For example (Figure 12.18):
  - 5 stations have a 5-minislot reservation-frame.
  - In the first interval, only stations 1, 3, and 4 have made reservations.
  - In the second interval, only station-1 has made a reservation.

4.3.2 Polling

- In a network,
  - One device is designated as a primary station and
  - Other devices are designated as secondary stations.
- Functions of primary-device:
  1) The primary-device controls the link.
  2) The primary-device is always the initiator of a session.
  3) The primary-device is determines which device is allowed to use the channel at a given time.
  4) All data exchanges must be made through the primary-device.
- The secondary devices follow instructions of primary-device.
- Disadvantage: If the primary station fails, the system goes down.
- Poll and select functions are used to prevent collisions (Figure 12.19).

1) Select

➢ If the primary wants to send data, it tells the secondary to get ready to receive; this is called select function.

➢ The primary
- → alerts the secondary about upcoming transmission by sending select frame (SEL)
- → then waits for an acknowledgment (ACK) from secondary
- → then sends the data frame and
- → finally waits for an acknowledgment (ACK) from the secondary.

2) Poll

➢ If the primary wants to receive data, it asks the secondaries if they have anything to send; this is called poll function.

➢ When the first secondary is approached, it responds either
- → with a NAK frame if it has no data to send or
- → with data-frame if it has data to send.

i) If the response is negative (NAK frame), then the primary polls the next secondary in the same manner.

ii) When the response is positive (a data-frame), the primary
- → reads the frame and
- → returns an acknowledgment (ACK frame).

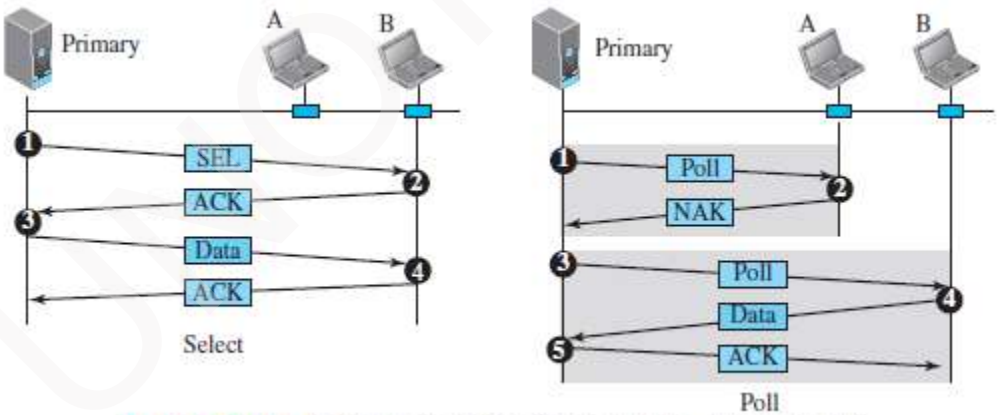


Figure 12.19 *Select and poll functions in polling-access method*

4.3.3 Token Passing

• In a network, the stations are organized in a ring fashion i.e. for each station; there is a predecessor and a successor.
  1) The predecessor is the station which is logically before the station in the ring.
  2) The successor is the station which is after the station in the ring.
• The current station is the one that is accessing the channel now.
• A token is a special packet that circulates through the ring.
• Here is how it works:
  - A station can send the data only if it has the token.
  - When a station wants to send the data, it waits until it receives the token from its predecessor.
  - Then, the station holds the token and sends its data.
  - When the station finishes sending the data, the station
    → releases the token
    → passes the token to the successor.
• Main functions of token management:
  1) Stations must be limited in the time they can hold the token.
  2) The token must be monitored to ensure it has not been lost or destroyed.
     For ex: if a station that is holding the token fails, the token will disappear from the network
  3) Assign priorities
     → to the stations and
     → to the types of data being transmitted.
  4) Make low-priority stations release the token to high priority stations.

4.3.3.1 Logical Ring

• In a token-passing network, stations do not have to be physically connected in a ring; the ring can be a logical one.
• Four physical topologies to create a logical ring (Figure 12.20):
  1) Physical ring
  2) Dual ring
  3) Bus ring
  4) Star ring

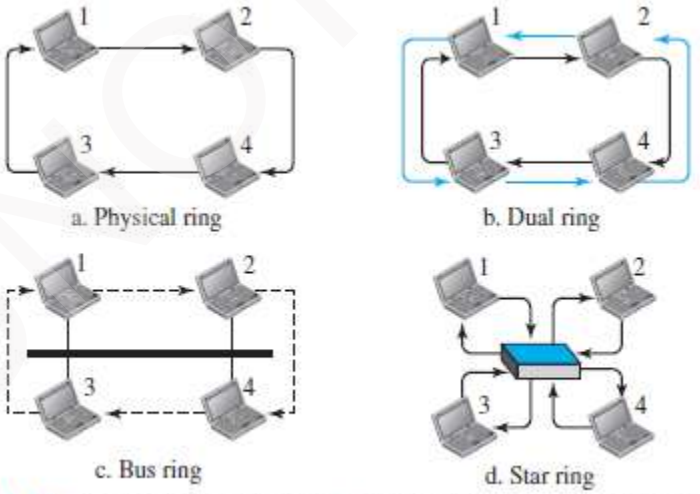


Figure 12.20 *Logical ring and physical topology in token-passing access method*

1) Physical Ring Topology

• When a station sends token to its successor, token cannot be seen by other stations. (Figure 12.20a)
• This means that the token does not have the address of the next successor.
• Disadvantage: If one of the links fails, the whole system fails.

2) Dual Ring Topology

- A second (auxiliary) ring
  - → is used along with the main ring (Figure 12.20b).
  - → operates in the reverse direction compared with the main ring.
  - → is used for emergencies only (such as a spare tire for a car).
- If the main ring fails, the system automatically combines the 2 rings to form a temporary ring.
- After the failed link is restored, the second ring becomes idle again.
- Each station needs to have 2 transmitter-ports and 2 receiver-ports.
- This topology is used in
  - i) FDDI (Fiber Distributed Data Interface) and
  - ii) CDDI (Copper Distributed Data Interface).

3) Bus Ring Topology

- The stations are connected to a single cable called a bus (Figure 12.20c).
- This makes a logical ring, because each station knows the address of its successor and predecessor.
- When a station has finished sending its data, the station
  - → releases the token and
  - → inserts the address of its successor in the token.
- Only the station gets the token to access the shared media.
- This topology is used in the Token Bus LAN.

4) Star Ring Topology

- The physical topology is a star (Figure 12.20d).
- There is a hub that acts as the connector.
- The wiring inside the hub makes the ring i.e. the stations are connected to the ring through the 2 wire connections.
- Disadvantages:
  1) This topology is less prone to failure because
     If a link goes down,
     then the link will be bypassed by the hub and
     the rest of the stations can operate.
  2) Also adding and removing stations from the ring is easier.
- This topology is used in the Token Ring LAN.

4.4 CHANNELIZATION PROTOCOLS

- Channelization is a multiple-access method.
- The available bandwidth of a link is shared b/w different stations in time, frequency, or through code.
- Three channelization protocols:
  1) FDMA (Frequency Division Multiple Access)
  2) TDMA (Time Division Multiple Access) and
  3) CDMA (Code Division Multiple Access)

4.4.1 FDMA

- The available bandwidth is divided into frequency-bands (Figure 12.21).
- Each band is reserved for a specific station.
- Each station can send the data in the allocated band.
- Each station also uses a bandpass filter to confine the transmitter frequencies.
- To prevent interferences, small guard bands are used to separate the allocated bands from one another.

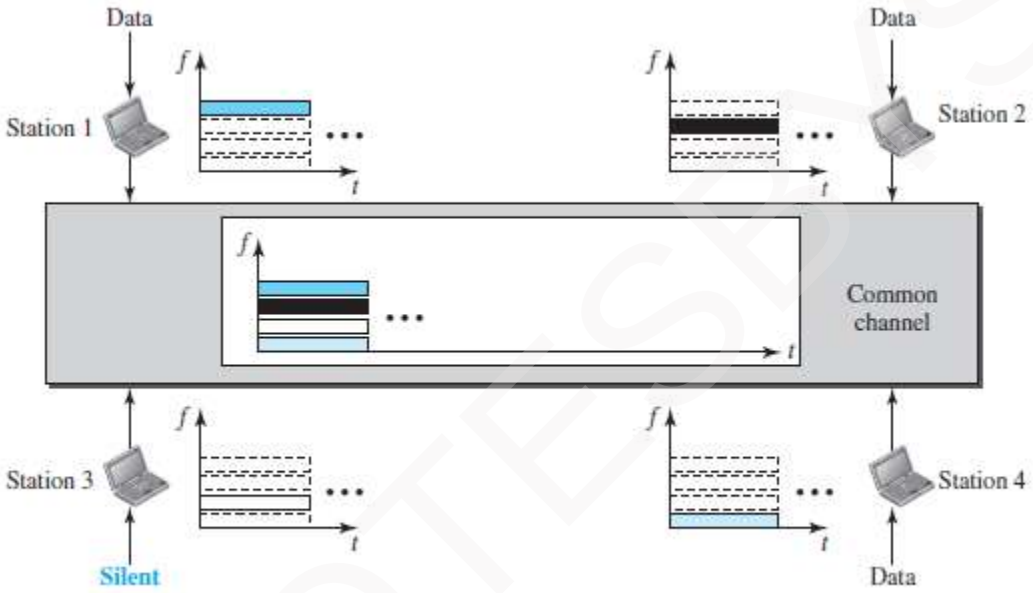


Figure 12.21 *Frequency-division multiple access (FDMA)*

- FDM vs. FDMA

1) FDM

- FDM is a multiplexing method in the physical layer.
- FDM
  → combines individual-loads from low-bandwidth channels and
  → transmits aggregated-load by using a high-bandwidth channel.
- The channels that are combined are low-pass.
- The multiplexer
  → modulates & combines the signals and
  → creates a bandpass signal.
- The bandwidth of each channel is shifted by the multiplexer.

2) FDMA

- FDMA is an access method in the data link layer.
- In each station, the data link layer tells the physical layer to make a bandpass signal from the data passed to it.
- The signal must be created in the allocated band.
- There is no physical multiplexer at the physical layer.
- The signals created at each station are automatically bandpass-filtered.
- They are mixed when they are sent to the common channel.

4.4.2 TDMA

- The stations share the bandwidth of the channel in time (Figure 12.22).
- Each time-slot is reserved for a specific station.
- Each station can send the data in the allocated time-slot.

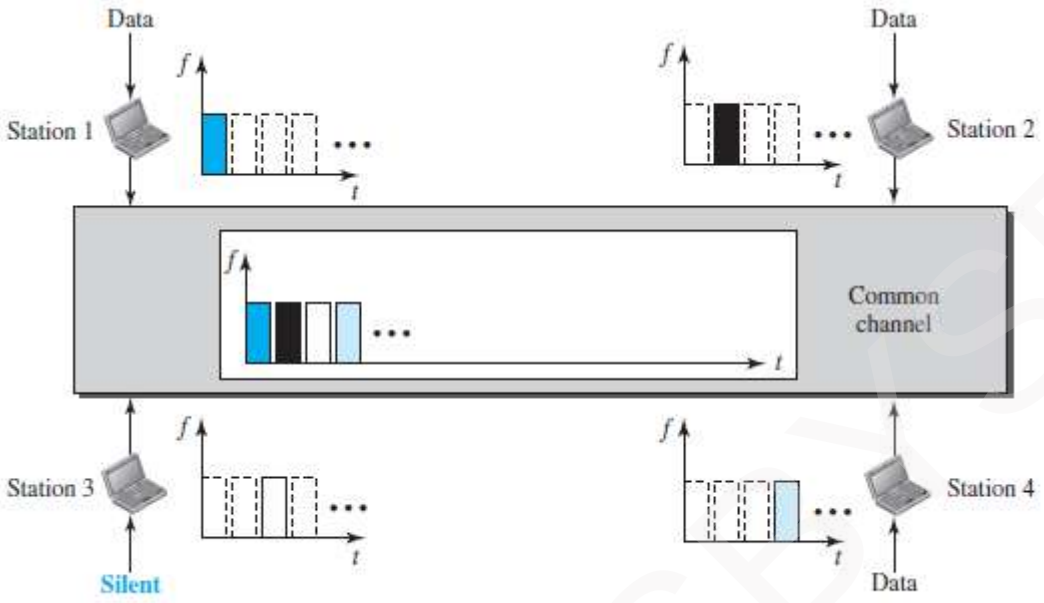


Figure 12.22 *Time-division multiple access (TDMA)*

- Main problem: Achieving synchronization between the different stations.
  i.e. each station needs to know the beginning of its slot and the location of its slot.
  This may be difficult because of propagation delays introduced in the system.
- To compensate for the delays, we can insert guard-times.
- Normally, synchronization is accomplished by having some synchronization bits at the beginning of each slot.
- TDMA vs. TDM

1) TDM

- TDM is a multiplexing method in the physical layer.
- TDM
  → combines the individual-data from slower channels and
  → transmits the aggregated- data by using a faster channel.
- The multiplexer interleaves data units from each channel.

2) TDMA

- TDMA is an access method in the data link layer.
- In each station, the data link layer tells the physical layer to use the allocated time-slot.
- There is no physical multiplexer at the physical layer.

4.4.3 CDMA

- CDMA simply means communication with different codes.
- CDMA differs from FDMA because
  → only one channel occupies the entire bandwidth of the link.
- CDMA differs from TDMA because
  → all stations can send data simultaneously; there is no timesharing.

(Analogy: CDMA simply means communication with different codes.
For example, in a large room with many people, 2 people can talk privately in English if nobody else understands English. Another 2 people can talk in Chinese if they are the only ones who understand Chinese, and so on).

4.4.3.1 Implementation

- Let us assume we have four stations 1, 2, 3, and 4 connected to the same channel.
- The data from station-1 are d_1, from station-2 are d_2, and so on.
- The code assigned to the first station is c_1, to the second is c_2, and so on.
- We assume that the assigned codes have 2 properties.
  1) If we multiply each code by another, we get 0.
  2) If we multiply each code by itself, we get 4 (the number of stations).

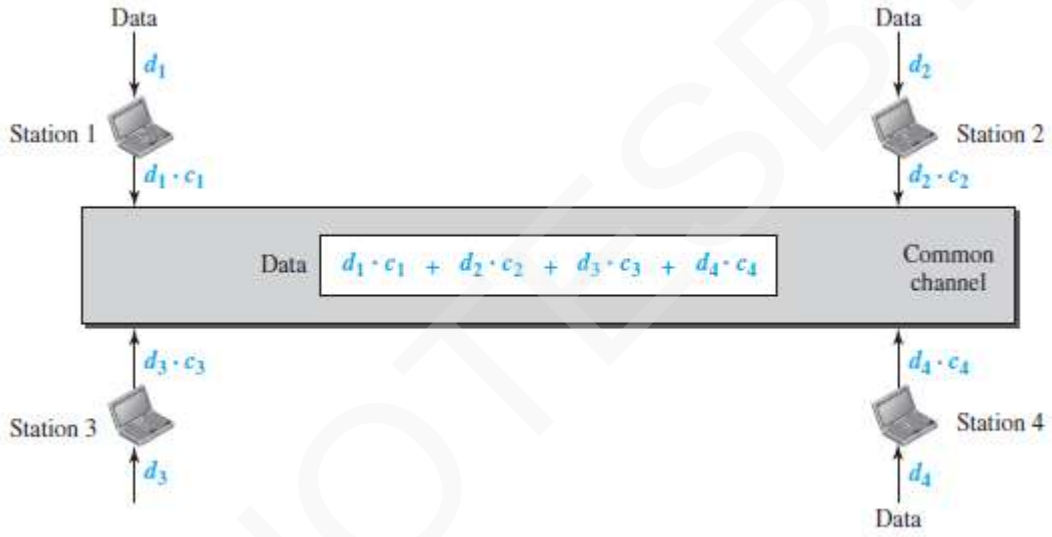


Figure 12.23 *Simple idea of communication with code*

- Here is how it works (Figure 12.23):
  - Station-1 multiplies the data by the code to get $d_1.c_1$.
  - Station-2 multiplies the data by the code to get $d_2.c_2$. And so on.
  - The data that go on the channel are the sum of all these terms.

$$d_1 \cdot c_1 + d_2 \cdot c_2 + d_3 \cdot c_3 + d_4 \cdot c_4$$

  - The receiver multiplies the data on the channel by the code of the sender.
  - For example, suppose stations 1 and 2 are talking to each other.
  - Station-2 wants to hear what station-1 is saying.
  - Station-2 multiplies the data on the channel by c_1 the code of station-1.
    $(c_1.c_1)=4$, $(c_2.c_1)=0$, $(c_3.c_1)=0$, and $(c_4.c_1)=0$,
    Therefore, station-2 divides the result by 4 to get the data from station-1.

$$\begin{aligned}\text{data} &= (d_1 \cdot c_1 + d_2 \cdot c_2 + d_3 \cdot c_3 + d_4 \cdot c_4) \cdot c_1 \\ &= d_1 \cdot c_1 \cdot c_1 + d_2 \cdot c_2 \cdot c_1 + d_3 \cdot c_3 \cdot c_1 + d_4 \cdot c_4 \cdot c_1 = 4 \times d_1\end{aligned}$$

4.4.3.2 Chips

- CDMA is based on coding theory.
- Each station is assigned a code, which is a sequence of numbers called chips (Figure 12.24).

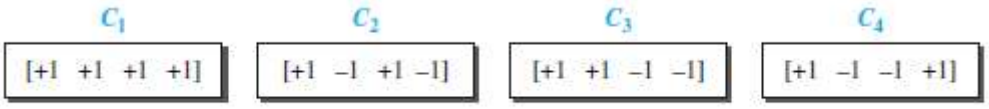


Figure 12.24 *Chip sequences*

- These sequences were carefully selected & are called orthogonal sequences
- These sequences have the following properties:
  1) Each sequence is made of N elements, where N is the number of stations.
  2) Multiplication of a sequence by a scalar:
     If we multiply a sequence by a number i.e. every element in the sequence is multiplied by that element.
     For example,
     $$2 \cdot [+1 +1 -1 -1] = [+2 +2 -2 -2]$$
  3) Inner product of 2 equal sequences:
     If we multiply 2 equal sequences, element by element, and add the results, we get N, where N is the number of elements in the each sequence.
     For example,
     $$[+1 +1 -1 -1] \cdot [+1 +1 -1 -1] = 1 + 1 + 1 + 1 = 4$$
  4) Inner product of 2 different sequences:
     If we multiply 2 different sequences, element by element, and add the results, we get 0.
     For example,
     $$[+1 +1 -1 -1] \cdot [+1 +1 +1 +1] = 1 + 1 - 1 - 1 = 0$$
  5) Adding 2 sequences means adding the corresponding elements. The result is another sequence.
     For example,
     $$[+1 +1 -1 -1] + [+1 +1 +1 +1] = [+2 +2 \ 0 \ 0]$$

4.4.3.3 Data Representation

- We follow the following rules for encoding:
  1) To send a 0 bit, a station encodes the bit as -1
  2) To send a 1 bit, a station encodes the bit as +1
  3) When a station is idle, it sends no signal, which is interpreted as a 0.

4.4.3.4 Encoding and Decoding

- We assume that
  - Stations 1 and 2 are sending a 0 bit.
  - Station-4 is sending a 1 bit.
  - Station-3 is silent.
- Here is how it works (Figure 12.26):
  - At the sender-site, the data are translated to -1, -1, 0, and +1.
  - Each station multiplies the corresponding number by its chip (its orthogonal sequence).
  - The result is a new sequence which is sent to the channel.
  - The sequence on the channel is the sum of all 4 sequences.
  - Now imagine station-3, which is silent, is listening to station-2.
  - Station-3 multiplies the total data on the channel by the code for station-2, which is [+1 -1 +1-1], to get

$$[-1\ -1\ -3\ +1] \bullet [+1\ -1\ +1\ -1] = -4/4 = -1 \rightarrow \text{bit 1}$$

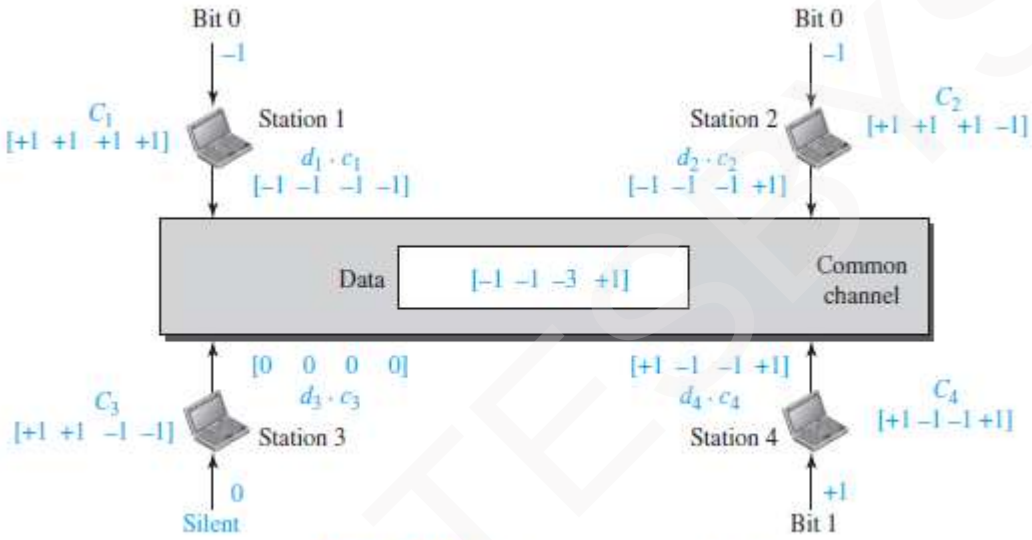
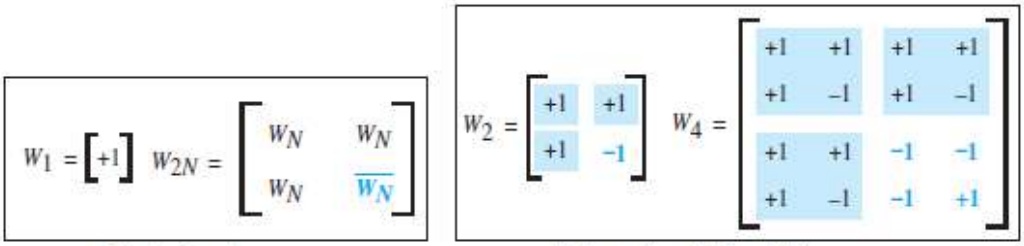


Figure 12.26 *Sharing channel in CDMA*

4.4.3.5 Sequence Generation

- To generate chip sequences, we use a Walsh table (Figure 12.29).
- Walsh table is a 2-dimensional table with an equal number of rows and columns.



a. Two basic rules b. Generation of W_2 and W_4

Figure 12.29 *General rule and examples of creating Walsh tables*

- In the Walsh table, each row is a sequence of chips.
- W_1 for a one-chip sequence has one row and one column. We can choose –1 or +1 for the chip for this trivial table (we chose +1).
- According to Walsh, if we know the table for N sequences W_N, we can create the table for 2N sequences W_{2N} (Figure 12.29).
- The W_N with the overbar W_N stands for the complement of W_N where each +1 is changed to -1 and vice versa.
- After we select W_1, W_2 can be made from four W_1's, with the last one the complement of W_1
- After W_2 is generated, W_4 can be made of four W_2's, with the last one the complement of W_2.
- The number of sequences in a Walsh table needs to be $N = 2^m$.

Example 4.5

Find the chips for a network with

a. Two stations
b. Four stations

Solution

We can use the rows of W_2 and W_4 in Figure 12.29:

a. For a two-station network, we have [+1 +1] and [+1 –1].
b. For a four-station network we have [+1 +1 +1 +1], [+1 –1 +1 –1], [+1 +1 –1 –1], and [+1 –1 –1 +1].

Example 4.6

What is the number of sequences if we have 90 stations in our network?

Solution

The number of sequences needs to be 2^m. We need to choose $m = 7$ and $N = 2^7$ or 128. We can then use 90 of the sequences as the chips.

Example 4.7

Prove that a receiving station can get the data sent by a specific sender if it multiplies the entire data on the channel by the sender's chip code and then divides it by the number of stations.

Solution

Let us prove this for the first station, using our previous four-station example. We can say that the data on the channel $D = (d_1 \cdot c_1 + d_2 \cdot c_2 + d_3 \cdot c_3 + d_4 \cdot c_4)$. The receiver that wants to get the data sent by station 1 multiplies these data by c_1.

$$
\begin{aligned}
D \cdot c_1 &= (d_1 \cdot c_1 + d_2 \cdot c_2 + d_3 \cdot c_3 + d_4 \cdot c_4) \quad c_1 \\
&= d_1 \cdot c_1 \cdot c_1 + d_2 \cdot c_2 \cdot c_1 + d_3 \cdot c_3 \cdot c_1 + d_4 \cdot c_4 \cdot c_1 \\
&= d_1 \times N + d_2 \times 0 + d_3 \times 0 + d_4 \times 0 \\
&= d_1 \times N
\end{aligned}
$$

When we divide the result by N, we get d_1.

MODULE 4(CONT.): WIRED LANS -- ETHERNET

4.5 ETHERNET PROTOCOL

4.5.1 IEEE Project 802

• The data-link-layer is divided into 2 sublayers (Figure 13.1):

1) LLC

- Flow-control, error-control, and framing duties are grouped into one sublayer called LLC.
- Framing is handled in both the LLC and the MAC.
- LLC vs. MAC
  - i) LLC provides one single data-link-control protocol for all IEEE LANs.
  - ii) MAC provides different protocols for different LANs.
- A single LLC protocol can provide interconnectivity between different LANs because
  → it makes the MAC sublayer transparent.

2) MAC

- This defines the specific access-method for each LAN.
- For example:
  - i) CSMA/CD is used for Ethernet LANs.
  - ii) Token-passing method is used for Token Ring and Token Bus LANs.
- The framing function is also handled by the MAC layer.
- The MAC contains a number of distinct modules.
- Each module defines the access-method and the framing-format specific to the corresponding LAN protocol.

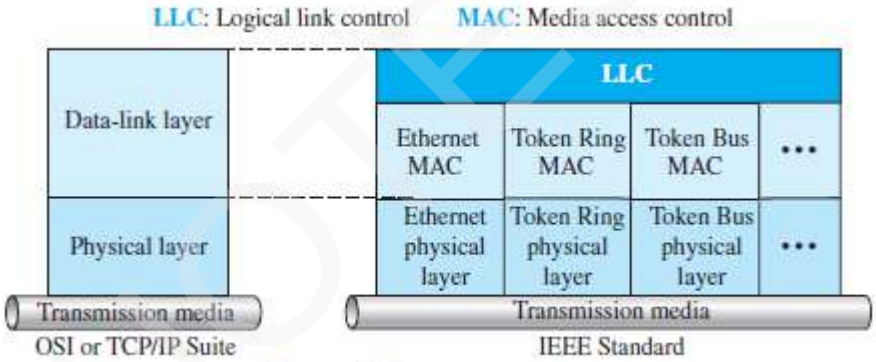


Figure 13.1 *IEEE standard for LANs*

4.5.2 Ethernet Evolution

• Four generations of Ethernet (Figure 13.2):

1) Standard-Ethernet (10 Mbps)
2) Fast-Ethernet (100 Mbps)
3) Gigabit-Ethernet (1 Gbps) and
4) Ten-Gigabit-Ethernet (10 Gbps)

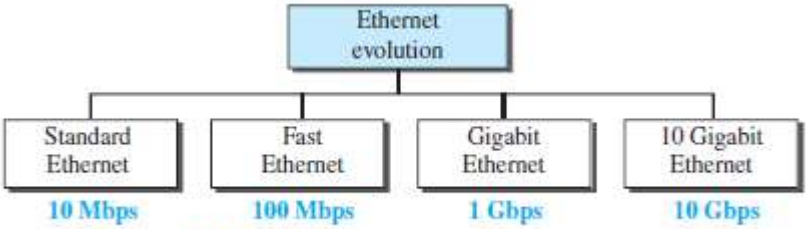


Figure 13.2 *Ethernet evolution through four generations*

4.6 STANDARD ETHERNET

• The original Ethernet technology with data-rate of 10 Mbps are referred to as the Standard Ethernet.

4.6.1 Characteristics

4.6.1.1 Connectionless and Unreliable Service

• Ethernet provides a connectionless service. Thus, each frame sent is independent of another frame.
• Ethernet has no connection establishment or connection termination phases.
• The sender sends a frame whenever it has it.
The receiver may or may not be ready for receiving the frame.
• The sender may overload the receiver with frames, which may result in dropping frames.
1) If a frame drops, the sender will not know about it.
2) If a frame is corrupted during transmission, the receiver drops the frame.
• Since IP is also connectionless, it will also not know about frame drops.
1) If the transport layer is UDP (connectionless protocol), the frame is lost.
2) If the transport layer is TCP, the sender-TCP does not receive acknowledgment for its segment and sends it again.
• Ethernet is also unreliable like IP and UDP.

4.6.1.2 Frame Format

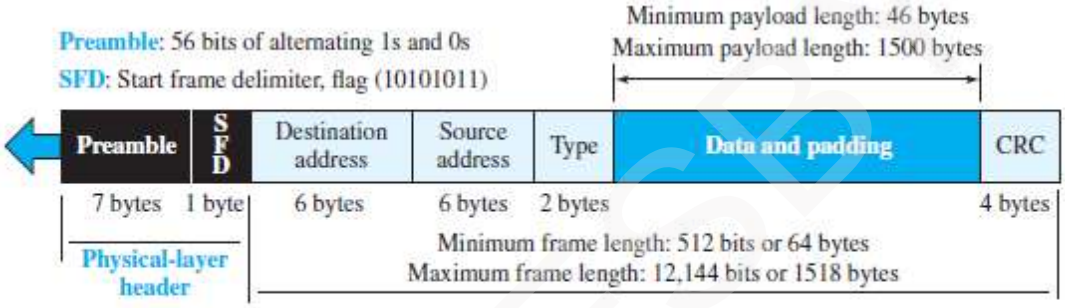


Figure 13.3 *Ethernet frame*

• The Ethernet frame contains 7 fields (Figure 13.3):

1) Preamble
- This field contains 7 bytes (56 bits) of alternating 0s and 1s.
- This field
  → alerts the receiving-system to the coming frame and
  → enables the receiving-system to synchronize its input timing.
- The preamble is actually added at the physical-layer and is not (formally) part of the frame.

2) Start Frame Delimiter (SFD)
- This field signals the beginning of the frame.
- The SFD warns the stations that this is the last chance for synchronization.
- This field contains the value: 10101011.
- The last 2 bits (11) alerts the receiver that the next field is the destination-address.

3) Destination Address (DA)
- This field contains the physical-address of the destination-station.

4) Source Address (SA)
- This field contains the physical-address of the sender-station.

5) Length or Type
- This field is defined as a i) type field or ii) length field.
  i) In original Ethernet, this field is used as the type field.
  ¤ Type field defines the upper-layer protocol using the MAC frame.
  ii) In IEEE standard, this field is used as the length field.
  ¤ Length field defines the number of bytes in the data-field.

6) Data
- This field carries data encapsulated from the upper-layer protocols.
- Minimum data size = 46 bytes. Maximum data size = 1500 bytes.

7) CRC
- This field contains error detection information such as a CRC-32.

4.6.1.3 Frame Length

• Ethernet has imposed restrictions on both minimum & maximum lengths of a frame (Figure 13.5).

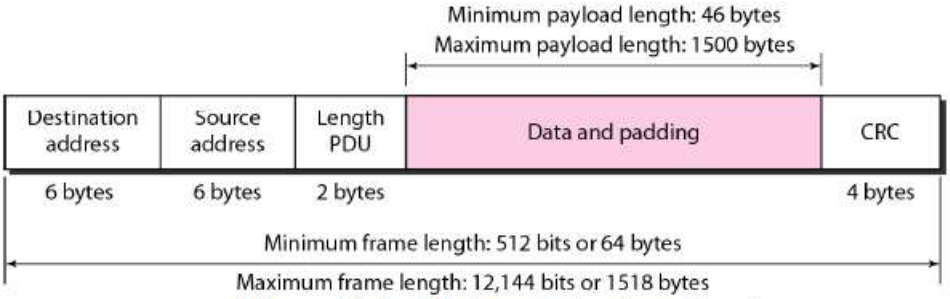


Figure 13.5 *Minimum and maximum lengths*

• The minimum length restriction is required for the correct operation of CSMA/CD.
• Minimum length of frame = 64 bytes.
    1) Minimum data size = 46 bytes.
    2) Header size + Trailer size = 14 + 4 = 18 bytes.
  (i.e. 18 bytes → 6 bytes source-address + 6 bytes dest-address + 2 bytes length + 4 bytes CRC).
• The minimum length of data from the upper layer = 46 bytes.
• If the upper-layer packet is less than 46 bytes, padding is added to make up the difference.
• Maximum length of frame =1518 bytes.
    1) Maximum data size = 1500 bytes.
    2) Header size + trailer size = 14 + 4 = 18 bytes.
• The maximum length restriction has 2 reasons:
    1) Memory was very expensive when Ethernet was designed.
        A maximum length restriction helped to reduce the size of the buffer.
    2) This restriction prevents one station from
        → monopolizing the shared medium
        → blocking other stations that have data to send.

4.6.2 Addressing

- In an Ethernet-network, each station has its own NIC (6-byte → 48 bits).
- The NIC provides the station with a 6-byte physical-address (or Ethernet-address).
- For example, the following shows an Ethernet MAC address:

06:01 :02:01:2C:4B
6 bytes = 12 hex digits = 48 bits

(NIC → network interface card)

Example 4.8

Show how the address 47:20:1B:2E:08:EE is sent out online.

Solution

The address is sent left to right, byte by byte; for each byte, it is sent right to left, bit by bit, as shown below

Hexadecimal	47	20	1B	2E	08	EE
Binary	01000111	00100000	00011011	00101110	00001000	11101110
Transmitted ←	11100010	00000100	11011000	01110100	00010000	01110111

4.6.2.1 Unicast, Multicast, and Broadcast Addresses

- A source-address is always a unicast address i.e. the frame comes from only one station.

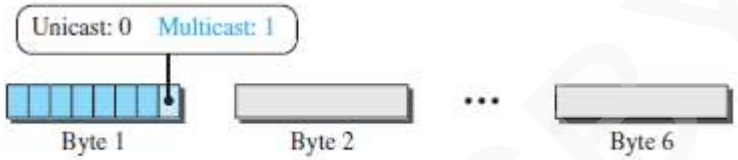


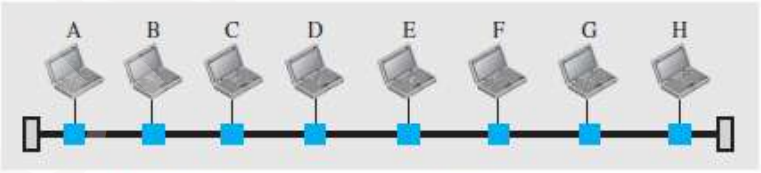
Figure 13.4 *Unicast and multicast addresses*

- However, the destination-address can be 1) Unicast 2) Multicast or 3) Broadcast.
- As shown in Figure 13.4,

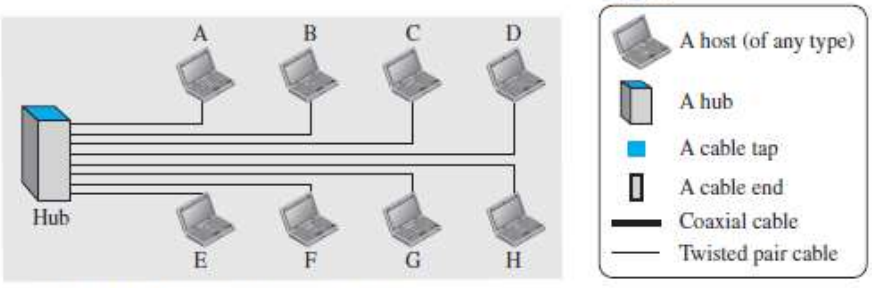
If LSB of first byte in a destination-address is 0,
Then, the address is unicast;
Otherwise, the address is multicast.

1) A unicast destination-address defines only one recipient.
   - ¤ The relationship between the sender and the receiver is one-to-one.
2) A multicast destination-address defines a group of addresses.
   - ¤ The relationship between the sender and the receivers is one-to-many.
3) The broadcast address is a special case of the multicast address.
   - ¤ The recipients are all the stations on the LAN.
   - ¤ A broadcast destination-address is 48 1s (6-byte → 48 bits).

- Standard Ethernet uses a coaxial cable (bus topology) or a set of twisted-pair cables with a hub (star topology) (Figure 13.5).



a. A LAN with a bus topology using a coaxial cable



b. A LAN with a star topology using a hub

Figure 13.5 *Implementation of standard Ethernet*

• Question: How actual unicast, multicast & broadcast transmissions are distinguished from each other?

Answer: The way the frames are kept or dropped.

1) In a unicast transmission, all stations will receive the frame, the intended recipient keeps and handles the frame; the rest discard it.

2) In a multicast transmission, all stations will receive the frame, the stations that are members of the group keep and handle it; the rest discard it.

3) In a broadcast transmission, all stations (except the sender) will receive the frame and all stations (except the sender) keep and handle it.

Example 4.9

Define the type of the following destination addresses:

a. 4A:30:10:21:10:1A
b. 47:20:1B:2E:08:EE
c. FF:FF:FF:FF:FF:FF

Solution

To find the type of the address, we need to look at the second hexadecimal digit from the left. If it is even, the address is unicast. If it is odd, the address is multicast. If all digits are Fs, the address is broadcast. Therefore, we have the following:

a. This is a unicast address because A in binary is 1010 (even).
b. This is a multicast address because 7 in binary is 0111 (odd).
c. This is a broadcast address because all digits are Fs in hexadecimal.

4.6.3 Access Method

• Standard-Ethernet uses 1-persistent CSMA/CD.

1) Slot Time

Slot time = round-trip time + time required to send the jam sequence.

• The RTT means time required for a frame to travel from one end of a maximum-length network to the other end (RTT → round-trip time).
• The slot time is defined in bits.
• The slot time is the time required for a station to send 512 bits.
• The actual slot time depends on the data-rate.

For example: 10-Mbps Ethernet has slot time of 51.2 μs.

2) Slot Time and Collision

• The choice of a 512-bit slot time was not accidental.
• It was chosen to allow the proper functioning of CSMA/CD.

3) Slot Time and Maximum Network Length

• There is a relationship between
→ slot time and
→ maximum length of the network (collision domain).
• This relationship is dependent on the propagation-speed of the signal in the particular medium.

i) In most transmission media, the signal propagates at 2×10^8 m/s (two-thirds of the rate for propagation in air).

ii) For traditional Ethernet, we calculate

$$\text{MaxLength} = \text{PropagationSpeed} \times \frac{\text{SlotTime}}{2}$$

MaxLength= $(2 \times 10^8) \times (51.2 \times 10^{-6})/2 = 5120$m

4.6.4 Efficiency of Standard Ethernet

• The efficiency is defined as the ratio of the time used by a station to send data to the time the medium is occupied by this station.

• The practical efficiency of standard Ethernet has been measured to be

$$\text{Efficiency} = 1 / (1 + 6.4 \times a)$$

where a = number of frames that can fit on the medium.

a = (propagation delay)/(transmission delay)

• As the value of parameter a decreases, the efficiency increases.

• If the length of the media is shorter or the frame size longer, the efficiency increases.

• In the ideal case, a = 0 and the efficiency is 1.

Example 4.10

In the Standard Ethernet with the transmission rate of 10 Mbps, we assume that the length of the medium is 2500 m and the size of the frame is 512 bits. The propagation speed of a signal in a cable is normally 2×10^8 m/s.

Propagation delay = $2500/(2 \times 10^8) = 12.5\ \mu s$	Transmission delay = $512/(10^7) = 51.2\ \mu s$
$a = 12.5/51.2 = 0.24$	Efficiency = 39%

4.6.5 Implementation

• The Standard-Ethernet defines several physical-layer implementations (Table 13.1).

Table 13.1 *Summary of Standard Ethernet implementations*

Implementation	*Medium*	*Medium Length*	*Encoding*
10Base5	Thick coax	500 m	Manchester
10Base2	Thin coax	185 m	Manchester
10Base-T	2 UTP	100 m	Manchester
10Base-F	2 Fiber	2000 m	Manchester

4.6.5.1 Encoding and Decoding

• All standard implementations use digital-signaling (baseband) at 10 Mbps (Figure 13.6).

1) At the sender, data are converted to a digital-signal using the Manchester scheme.
2) At the receiver, the received-signal is
→ interpreted as Manchester and
→ decoded into data.

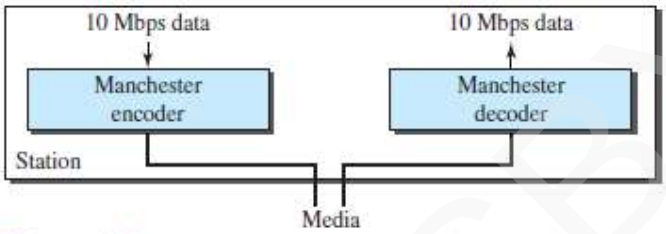


Figure 13.6 *Encoding in a Standard Ethernet implementation*

1) 10Base5: Thick Ethernet

➢ 10Base5 uses a bus topology (Figure 13.7).
➢ A external transceiver is connected to a thick coaxial-cable.
(transceiver → transmitter/receiver)
➢ The transceiver is responsible for
→ transmitting
→ receiving and
→ detecting collisions.
➢ The transceiver is connected to the station via a coaxial-cable.
The cable provides separate paths for sending and receiving.
The collision can only happen in the coaxial cable.

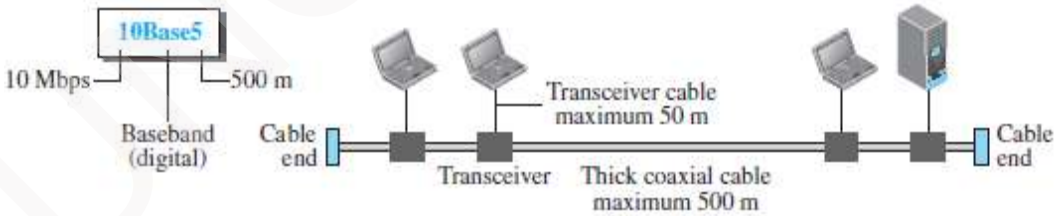


Figure 13.7 *10Base5 implementation*

➢ The maximum-length of the cable must not exceed 500m.
If maximum-length is exceeded, then there will be excessive degradation of the signal.
➢ If a cable-length of more than 500 m is needed, the total cable-length can be divided into up to 5 segments.
➢ Each segment of maximum length 500-meter, can be connected using repeaters.

2) 10Base2: Thin Ethernet

➢ 10Base2 uses a bus topology (Figure 13.8).
➢ The cable is much thinner and more flexible than 10Base5.
➢ Flexible means the cable can be bent to pass very close to the stations.
➢ The transceiver is part of the NIC, which is installed inside the station.
➢ The collision can only happen in the coaxial cable.

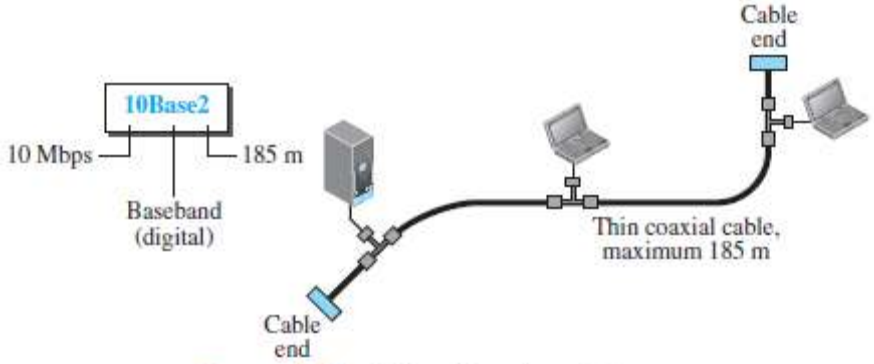


Figure 13.8 *10Base2 implementation*

- Advantages:
  1) Thin coaxial-cable is less expensive than thick coaxial-cable.
  2) Tee connections are much cheaper than taps.
  3) Installation is simpler because the thin coaxial cable is very flexible.
- Disadvantage:
  1) Length of each segment cannot exceed 185m due to the high attenuation in the cable.

3)10Base- T: Twisted Pair Ethernet

- 10Base-T uses a star topology to connect stations to a hub (Figure 13.9).
- The stations are connected to a hub using two pairs of twisted-cable.

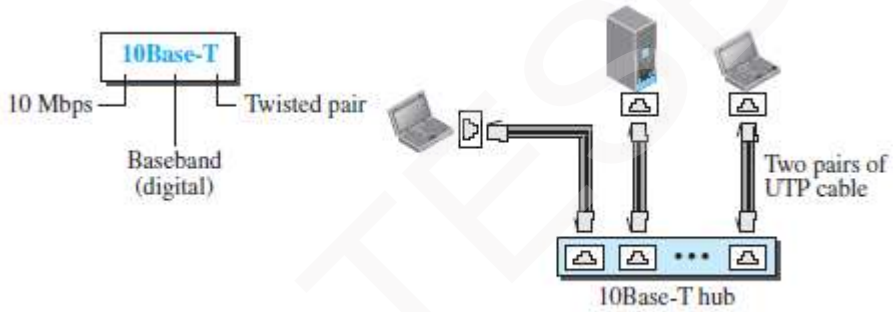


Figure 13.9 *10Base-T implementation*

- Two pairs of twisted cable create two paths between the station and the hub.
  1) First path for sending.
  2) Second path for receiving.
- The collision can happen in the hub.
- The maximum length of the cable is 100 m. This minimizes the effect of attenuation in the cable.

4) 10Base-F: Fiber Ethernet

- 10Base-F uses a star topology to connect stations to a hub (Figure 13.10).
- The stations are connected to the hub using two fiber-optic cables.

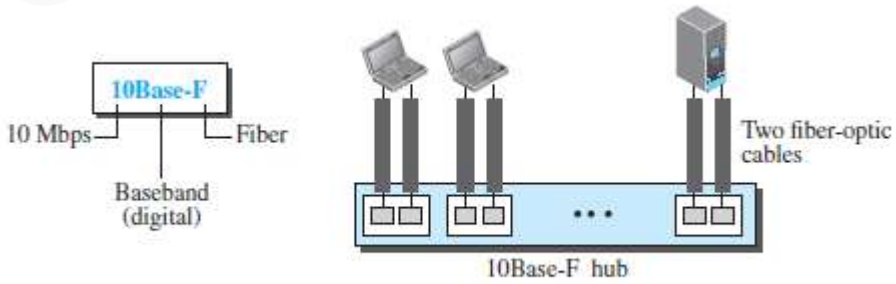


Figure 13.10 *10Base-F implementation*

4.6.6 Changes in the Standard

4.6.6.1 Bridged Ethernet

• Bridges have two effects on an Ethernet LAN:
  i) They raise the bandwidth &
  ii) They separate collision domains.

1) Raising the Bandwidth

• A bridge divides the network into two or more networks.
• Bandwidth-wise, each network is independent.

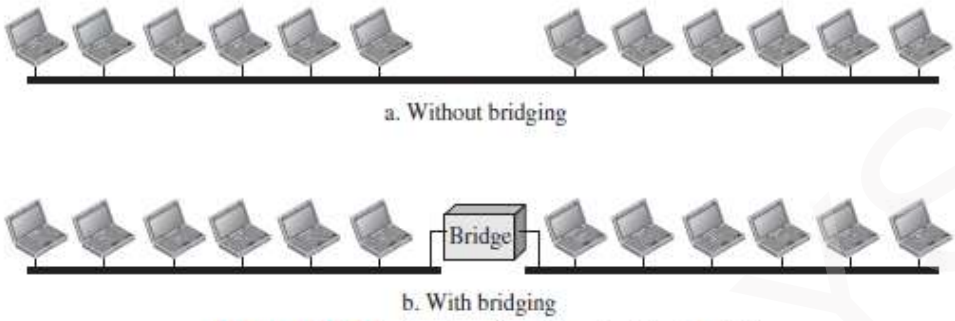


Figure 13.12 *A network with and without a bridge*

• For example (Figure 13.12):
  - A network with 12 stations is divided into two networks, each with 6 stations.
  - Now each network has a capacity of 10 Mbps.
  - The 10-Mbps capacity in each segment is now shared between 6 stations (actually 7 because the bridge acts as a station in each segment), not 12 stations.
  - In a network with a heavy load, each station theoretically is offered 10/7 Mbps instead of 10/12 Mbps.

2) Separating Collision Domains

• Another advantage of a bridge is the separation of the collision domain.
• Figure 13.13 shows the collision domains for an un-bridged and a bridged network.
• You can see that the collision domain becomes much smaller and the probability of collision is reduced tremendously.

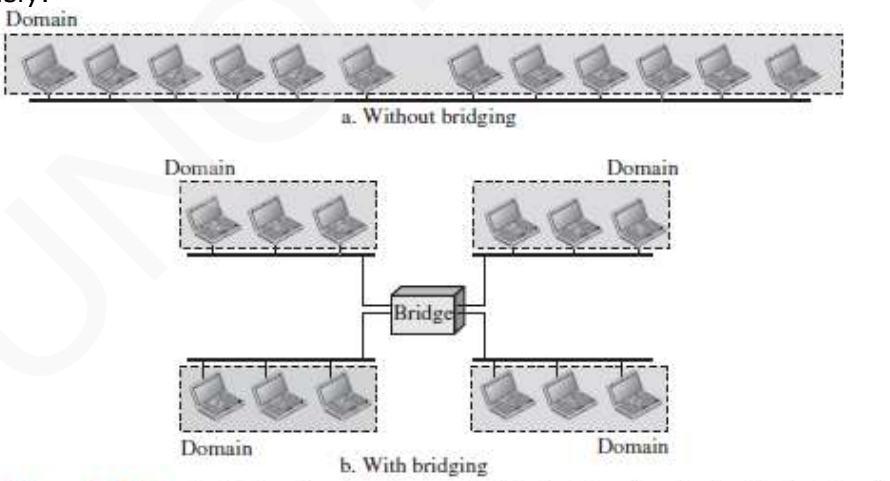


Figure 13.13 *Collision domains in an unbridged network and a bridged network*

4.6.6.2 Switched Ethernet

• The idea of a bridged LAN can be extended to a switched LAN (Figure 13.14).
• If we can have a multiple-port bridge, we can have an N-port switch.
• In this way, the bandwidth is shared only between the station and the switch.
• A layer-2 switch is an N-port bridge with additional sophistication that allows faster handling of the packets.

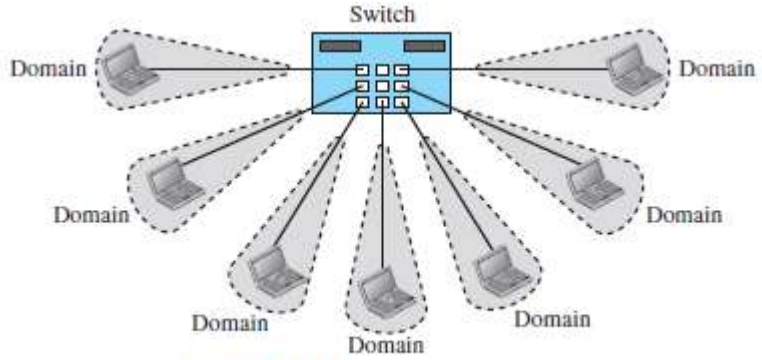


Figure 13.14 *Switched Ethernet*

4.6.6.3 Full-Duplex Ethernet

• The full-duplex mode increases the capacity of each domain from 10 to 20 Mbps.

• Instead of using one link between the station and the switch, the configuration uses two links: one to transmit and one to receive.

1) No Need for CSMA/CD

- In full-duplex switched Ethernet,
  - → There is no need for the CSMA/CD method.
  - → Each station is connected to the switch via two separate links.
- Each station or switch can send and receive independently without worrying about collision.
- Each link is a point-to-point dedicated path between the station and the switch.
- There is no longer a need for carrier sensing; there is no longer a need for collision-detection.
- The job of the MAC layer becomes much easier.
- Carrier sensing and collision-detection functionalities of the MAC sublayer can be turned off.

2) MAC Control Layer

- To provide for flow and error control in full-duplex switched Ethernet, a new sublayer, called the MAC control, is added between the LLC sublayer and the MAC sublayer.

4.7 FAST ETHERNET (100 MBPS)

- IEEE created Fast-Ethernet under the name 802.3u.
- Fast-Ethernet was designed to compete with LAN protocols such as FDDI or Fiber Channel.
- Goals of Fast-Ethernet:
    1) Upgrade the data-rate to 100 Mbps.
    2) Make it compatible with Standard-Ethernet.
    3) Keep the same 48-bit address.
    4) Keep the same frame format.
    5) Keep the same minimum and maximum frame-lengths.

4.7.1 Access Method

- Access method is same in Standard-Ethernet.
- Only the star topology is used.
- For the star topology, there are 2 choices:
    1) In the half-duplex approach, the stations are connected via a hub.
       CSMA/CD was used as access-method.
    2) In the full-duplex approach, the connection is made via a switch with buffers at each port.
       There is no need for CSMA/CD.

Autonegotiation

- A new feature added to Fast-Ethernet is called autonegotiation.
- It provides a station/hub with a range of capabilities.
- It was used for the following purposes:
    1) To allow 2 devices to negotiate the mode or data-rate of operation.
    2) To allow incompatible devices to connect to one another.
       For example: a device with a maximum capacity of 10 Mbps can communicate with a device with a 100 Mbps capacity.
    3) To allow one device to have multiple capabilities.
    4) To allow a station to check a hub's capabilities.

4.7.2 Physical Layer

• The physical-layer in Fast-Ethernet is more complicated than the one in Standard-Ethernet.
• Some of the features of this layer are as follows. 1) Topology 2) Implementation and 3) Encoding.

4.7.2.1 Topology

• Fast-Ethernet is used to connect two or more stations together (Figure 13.19).
1) If there are only 2 stations, they can be connected in point-to-point.
2) If there are 3 or more stations, they can be connected in star topology with a hub at the center.

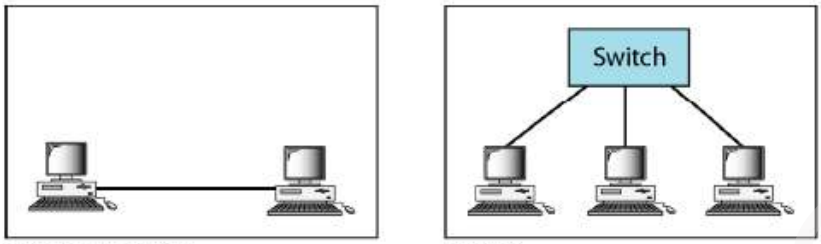


Figure 13.19 *Fast Ethernet topology*

4.7.2.2 Implementation

• Fast-Ethernet can be classified as either a two-wire or a four-wire implementation (Table 13.2).
1) The 2-wire implementations use
→ Category 5 UTP (100Base-TX) or
→ Fiber-optic cable (100Base-FX)
2) The 4-wire implementations use category 3 UTP (100Base-T4).

Table 13.2 *Summary of Fast Ethernet implementations*

Implementation	*Medium*	*Medium Length*	*Wires*	*Encoding*
100Base-TX	UTP or STP	100 m	2	4B5B + MLT-3
100Base-FX	Fiber	185 m	2	4B5B + NRZ-I
100Base-T4	UTP	100 m	4	Two 8B/6T

4.7.2.3 Encoding

• There are 3 different encoding schemes.

1) 100Base-TX

➢ This uses 2 pairs of twisted-pair cable (either category 5 UTP or STP) (Figure 13.16a).

➢ The MLT-3 encoding scheme is used for implementation.
This is because MLT-3 has good bandwidth performance.

➢ However, 4B/5B block-coding is used to provide bit synchronization.
This is because MLT-3 is not a self-synchronous line coding scheme.

➢ 4B/5B coding creates a data-rate of 125 Mbps, which is fed into MLT-3 for encoding.

2) 100Base-FX

➢ This uses 2 pairs of fiber-optic cables (Figure 13.16b).

➢ Optical fiber can easily handle high bandwidth requirements.

➢ The NRZ-I encoding scheme is used for implementation.

➢ However, 4B/5B block-coding is used to provide bit synchronization.
This is because NRZ-I is not a self-synchronous line coding scheme.

➢ 4B/5B encoding increases the bit rate from 100 to 125 Mbps, which can easily be handled by fiber-optic cable.

3) 100Base-T4

➢ This uses 4 pairs of UTP for transmitting 100 Mbps (Figure 13.16c).

➢ Each UTP cannot easily handle more than 25 Mbaud.

➢ One pair switches between sending and receiving.

➢ Three pairs of UTP can handle only 75 Mbaud (25 Mbaud) each.

➢ Encoding/decoding is more complicated.

➢ We need an encoding scheme that converts 100 Mbps to a 75 Mbaud signal. This requirement is satisfied by 8B/6T.

➢ The 8B/6T encoding scheme is used for implementation.
i) 8 data elements are encoded as 6 signal elements.
ii) This means that 100 Mbps uses only (6/8) x 100 Mbps, or 75 Mbaud.

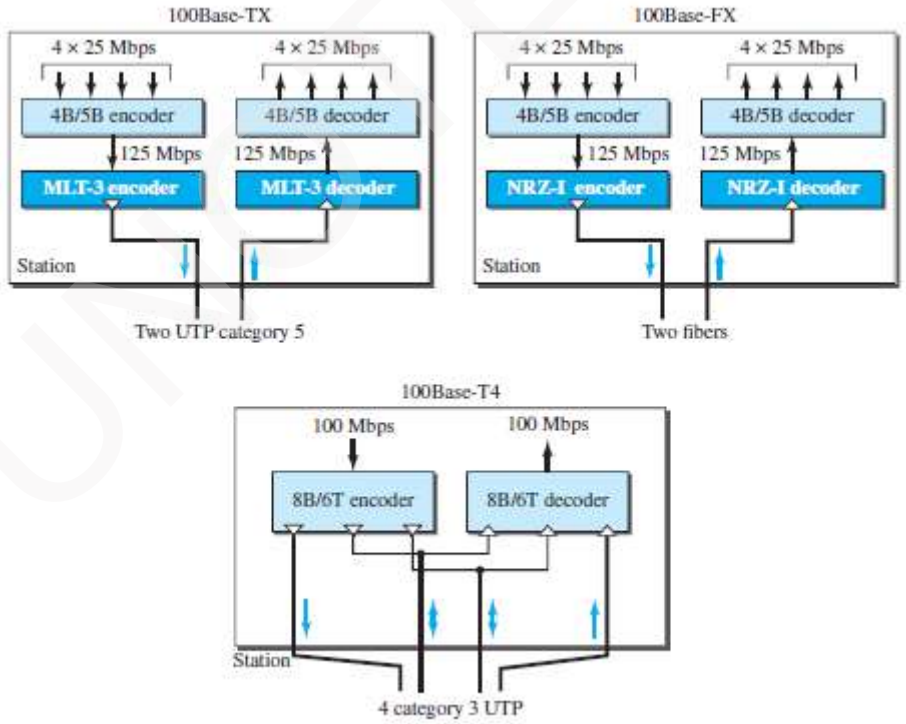


Figure 13.16 *Encoding for Fast Ethernet implementation*

4.8 GIGABIT ETHERNET

- IEEE created Gigabit-Ethernet under the name 802.3z.
- Goals of Gigabit-Ethernet:
  1) Upgrade the data-rate to 1 Gbps.
  2) Make it compatible with Standard or Fast-Ethernet.
  3) Use the same 48-bit address.
  4) Use the same frame format.
  5) Keep the same minimum and maximum frame-lengths.
  6) To support auto-negotiation as defined in Fast-Ethernet.

4.8.1 MAC Sublayer

- Gigabit-Ethernet has two distinctive approaches for medium access: half-duplex and full-duplex.
- Almost all implementations of Gigabit-Ethernet follow the full-duplex approach.

1) Full Duplex Mode

- There is a central switch connected to all computers or other switches.
- Each switch has buffers for each input-port in which data are stored until they are transmitted.
- There is no collision. This means that CSMA/CD is not used.
- Lack of collision implies that
  the maximum length of the cable is determined
  → by the signal attenuation in the cable &
  → not by the collision-detection process.

2) Half Duplex Mode

- A switch is replaced by a hub, which acts as the common cable in which a collision might occur.
- CSMA/CD is used.
- The maximum length of the network is totally dependent on the minimum frame size.
- Three methods have been defined: traditional, carder extension, and frame bursting.

i) Traditional

- Like traditional Ethernet, the minimum length of a frame is 512 bits.
- However, because the length of a bit is 1/100 shorter,
  Slot time is 512 bits x 1/1000 gs which is equal to 0.512 gs.
- The reduced slot time means that collision is detected 100 times earlier.
- The maximum length of the network is 25 m.
- This length may be suitable if all the stations are in one room.

ii) Carrier Extension

- To allow for a longer network, we increase the minimum frame-length.
- Minimum length of frame is 512 bytes (4096 bits). Thus, minimum length is 8 times longer.
- A station adds extension bits (padding) to any frame that is less than 4096 bits.
- The maximum length of the network is 200 m.
- A length from the hub to the station is 100 m.

iii) Frame Bursting

- Carrier extension is very inefficient if
  → we have a series of short frames to send
  → each frame carries redundant data.
- To improve efficiency, frame bursting was proposed.
- Instead of adding an extension to each frame, multiple frames are sent.
- However, to make these multiple frames look like one frame, padding is added between the frames. Thus, the channel is not idle.

4.8.2 Physical Layer

- The physical-layer in Gigabit-Ethernet is more complicated than that in Standard or Fast-Ethernet.
- Some of the features of this layer are as follows. 1) Topology 2) Implementation and 3) Encoding.

4.8.2.1 Topology

- Gigabit-Ethernet is used to connect two or more stations together.
  1) If there are only 2 stations, they can be connected in point-to-point.
  2) If there are 3 or more stations, they can be connected in star topology with a hub at center.

4.8.2.2 Implementation

- Gigabit-Ethernet can be classified as either a two-wire or a four-wire implementation (Table 13.3).
  1) The 2-wire implementations use
     → Fiber-optic cable (1000Base-SX, short-wave, or 1000Base-LX, long-wave) or
     → STP (1000Base-CX)
  2) The 4-wire implementations use category 5 twisted-pair cable (1000Base-T).

Table 13.3 *Summary of Gigabit Ethernet implementations*

Implementation	*Medium*	*Medium Length*	*Wires*	*Encoding*
1000Base-SX	Fiber S-W	550 m	2	8B/10B + NRZ
1000Base-LX	Fiber L-W	5000 m	2	8B/10B + NRZ
1000Base-CX	STP	25 m	2	8B/10B + NRZ
1000Base-T4	UTP	100 m	4	4D-PAM5

4.8.2.3 Encoding

1) Two Wire Implementation

- The NRZ encoding scheme is used for two-wire implementation (Figure 13.17a).
- However, 8B/10B block-coding is used to provide bit synchronization.
  This is because NRZ is not a self-synchronous line coding scheme.
- 8B/10B coding creates a data-rate of 1.25 Gbps.
- One wire (fiber or STP) is used for sending.
  Another wire is used for receiving.

2) Four Wire Implementation

- In this, it is not possible to have 2 wires for input and 2 for output (Figure 13.17b).
- This is '.' each wire would need to carry 500 Mbps, which exceeds the capacity for category 5 UTP.
- As a solution, 4D-PAM5 encoding is used to reduce the bandwidth.
- Thus, all four wires are involved in both input and output.
  Each wire carries 250 Mbps, which is in the range for category 5 UTP cable.

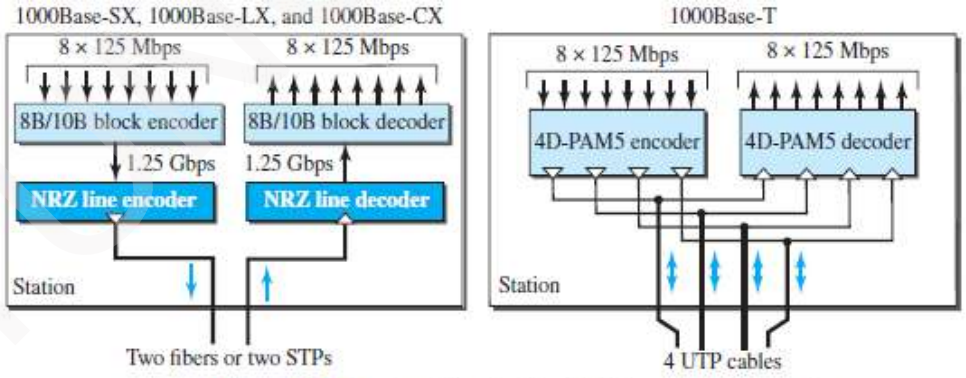


Figure 13.17 *Encoding in Gigabit Ethernet implementations*

4.9 TEN GIGABIT ETHERNET

- IEEE created Ten-Gigabit-Ethernet under the name 802.3ae.
- Goals of the Gigabit-Ethernet:
    1) Upgrade the data-rate to 10 Gbps.
    2) Make it compatible with Standard, Fast, and Gigabit-Ethernet.
    3) Use the same 48-bit address.
    4) Use the same frame format.
    5) Keep the same minimum and maximum frame-lengths.
    6) Allow the interconnection of existing LANs into a MAN or a WAN .
    7) Make Ethernet compatible with technologies such as Frame Relay and ATM.

4.9.1 Implementation

- Ten-Gigabit-Ethernet operates only in full duplex mode.
- This means there is no need for contention; CSMA/CD is not used.
- Four implementations are the most common (Table 13.4):
    1) 10GBase-SR
    2) 10GBase-LR
    3) 10GBase-EW and
    4) 10GBase-X4

Table 13.4 *Summary of 10 Gigabit Ethernet implementations*

Implementation	*Medium*	*Medium Length*	*Number of wires*	*Encoding*
10GBase-SR	Fiber 850 nm	300 m	2	64B66B
10GBase-LR	Fiber 1310 nm	10 Km	2	64B66B
10GBase-EW	Fiber 1350 nm	40 Km	2	SONET
10GBase-X4	Fiber 1310 nm	300 m to 10 Km	2	8B10B

MODULE 4(CONT.): WIRELESS-LANS

4.10 INTRODUCTION OF WIRELESS-LANS

4.10.1 Architectural Comparison

1) Medium

• In a wired LAN, we use wires to connect hosts.
• In a switched LAN, with a link-layer switch, the communication between the hosts is point-to-point and full-duplex (bidirectional).
• In a wireless LAN, the medium is air, the signal is generally broadcast.
• When hosts in a wireless LAN communicate with each other, they are sharing the same medium (multiple access).

2) Hosts

• In a wired LAN, a host is always connected to its network at a point with a fixed link layer address related to its network interface card (NIC).
• Of course, a host can move from one point in the Internet to another point.
• In this case, its link-layer address remains the same, but its network-layer address will change.
• In a wireless LAN, a host is not physically connected to the network; it can move freely and can use the services provided by the network.
• Therefore, mobility in a wired network and wireless network are totally different issues.

3) Isolated LANs

• A wired isolated LAN is a set of hosts connected via a link-layer switch (Figure 15.1).
• A wireless isolated LAN, called an ad hoc network in wireless LAN terminology, is a set of hosts that communicate freely with each other.
• The concept of a link-layer switch does not exist in wireless LANs.

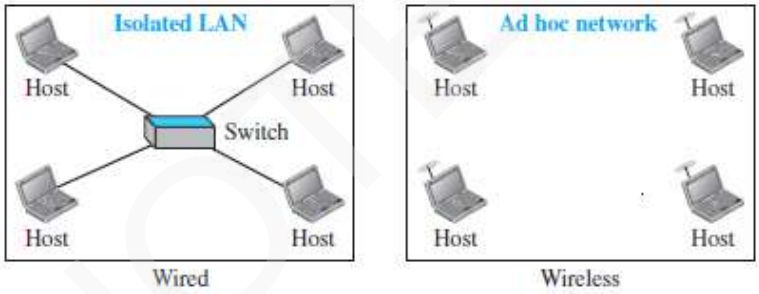


Figure 15.1 *Isolated LANs: wired versus wireless*

4) Connection to Other Networks

• A wired LAN can be connected to another network or the Internet using a router.
• A wireless LAN may be connected to a wired infrastructure network, to a wireless infrastructure network, or to another wireless LAN (Figure 15.2).
• In this case, the wireless LAN is referred to as an infrastructure network, and the connection to the wired infrastructure, such as the Internet, is done via a device called an access point (AP).
• An access point is gluing two different environments together: one wired and one wireless.
  1) Communication between the AP and the wireless host occurs in a wireless environment.
  2) Communication between the AP and the infrastructure occurs in a wired environment.

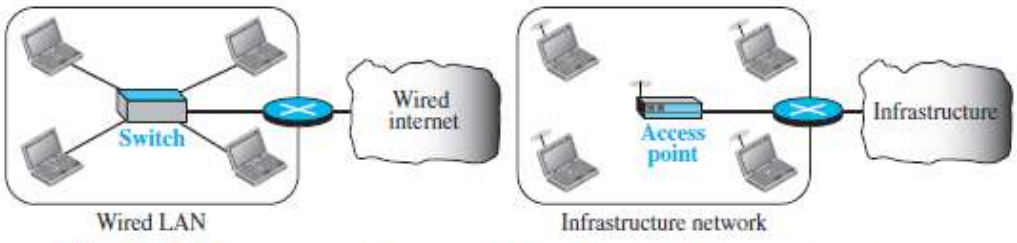


Figure 15.2 *Connection of a wired LAN and a wireless LAN to other networks*

4.10.2 Characteristics

1) Attenuation

• The strength of electromagnetic signals decreases rapidly because the signal disperses in all directions; only a small portion of it reaches the receiver.

• The situation becomes worse with mobile senders that operate on batteries and normally have small power supplies.

2) Interference

• Another issue is that a receiver may receive signals not only from the intended sender, but also from other senders if they are using the same frequency band.

3) Multipath Propagation

• A receiver may receive more than one signal from the same sender because electromagnetic waves can be reflected back from obstacles such as walls, the ground, or objects.

• The result is that the receiver receives some signals at different phases (because they travel different paths). This makes the signal less recognizable.

4) Error

• Error detection is more serious issues in a wireless network than in a wired network.

i) If SNR is high, it means that the signal is stronger than the noise (unwanted signal), so we may be able to convert the signal to actual data.

ii) When SNR is low, it means that the signal is corrupted by the noise and the data cannot be recovered.

4.10.3 Access Control

• The CSMA/CD algorithm does not work in wireless LANs for three reasons:

1) To detect a collision, a host needs to send and receive at the same time which means the host needs to work in a duplex mode. Wireless hosts do not have enough power to do so (the power is supplied by batteries).

¤ They can only send or receive at one time.

2) The distance between stations can be great.

¤ Signal fading could prevent a station at one end from hearing a collision at other end.

3) Because of the hidden station problem, in which a station may not be aware of another station's transmission due to some obstacles or range problems, collision may occur but not be detected.

Hidden Station Problem

¤ Figure 15.3 shows an example of the hidden station problem.

¤ Every station in transmission range of Station B can hear any signal transmitted by station B.

¤ Every station in transmission range of Station C can hear any signal transmitted by station C.

¤ Station C is outside the transmission range of B;

Likewise, station B is outside the transmission range of C.

¤ However, Station A is in the area covered by both B and C;

Therefore, Station A can hear any signal transmitted by B or C.

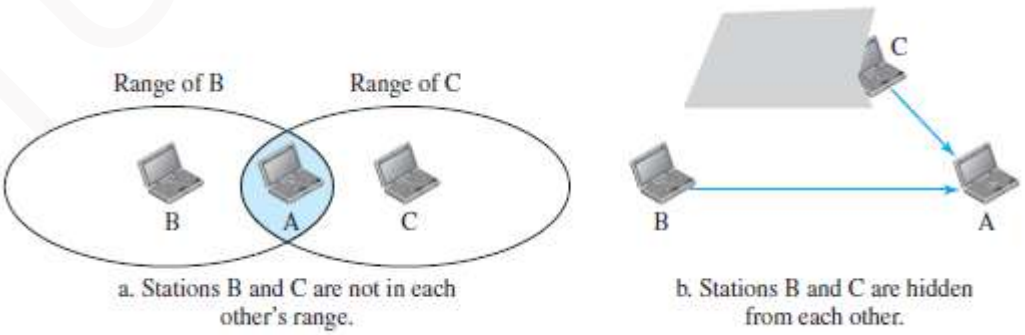


Figure 15.3 *Hidden station problem*

4.11 IEEE 802.11

4.11.1 Architecture

• The standard defines 2 kinds of services: 1) Basic service set (BSS) and
2) Extended service set (ESS).

4.11.1.1 BSS

• IEEE 802.11 defines the basic service set (BSS) as the building block of a wireless-LAN.
• A basic service set is made of (Figure 15.4):
→ stationary or mobile wireless stations and
→ optional central base station, known as the access point (AP).
• There are 2 types of architecture:

1) Ad hoc Architecture

- The BSS without an AP is a stand-alone network and cannot send data to other BSSs.
- Stations can form a network without the need of an AP.
- Stations can locate one another and agree to be part of a BSS.

2) Infrastructure Network

- A BSS with an AP is referred to as an infrastructure network.

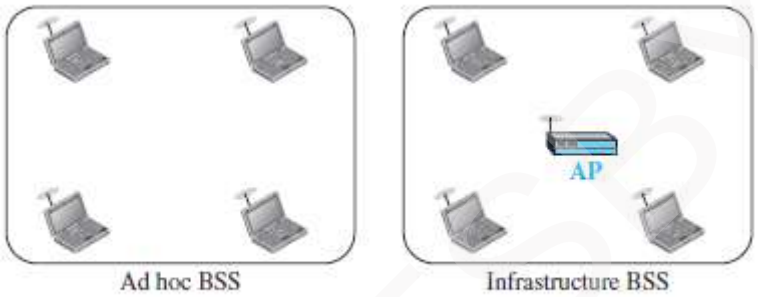


Figure 15.4 *Basic service sets (BSSs)*

4.11.1.2 ESS

• The ESS is made up of 2 or more BSSs with APs (Figure 15.5).
• The BSSs are connected through a distribution-system, which is usually a wired LAN.
• The distribution-system connects the APs in the BSSs.
• IEEE 802.11 does not restrict the distribution-system;
The distribution-system can be any IEEE LAN such as an Ethernet.
• The ESS uses 2 types of stations:

1) Mobile Stations are normal stations inside a BSS.
2) Stationary Stations are AP stations that are part of a wired LAN.

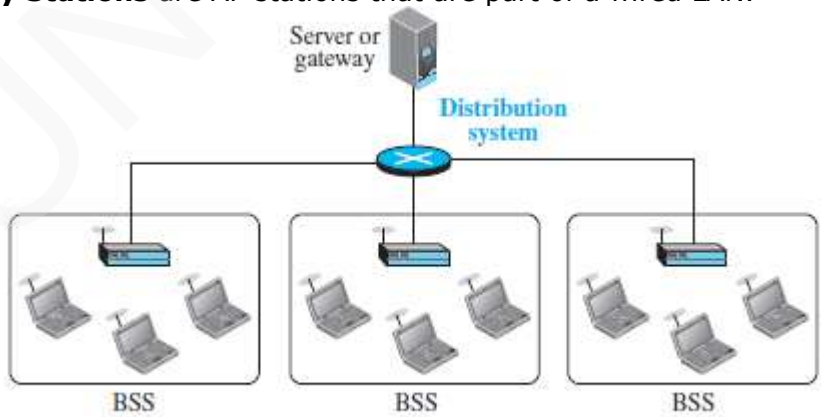


Figure 15.5 *Extended service set (ESS)*

• When BSSs are connected, the stations within reach of one another can communicate without the use of an AP.
• However, communication between two stations in two different BSSs usually occurs via two APs.

4.11.1.3 Station Types

• IEEE 802.11 defines three types of stations based on their mobility in a wireless-LAN:
1) No-transition 2) BSS-transition 3) ESS-transition mobility

1) A station with no-transition mobility is either
→ stationary (not moving) or
→ moving only inside a BSS.

2) A station with BSS-transition mobility can move from one BSS to another, but the movement is confined inside one ESS.

3) A station with ESS-transition mobility can move from one ESS to another.
However, IEEE 802.11 does not guarantee that communication is continuous during the move.

4.11.2 MAC Sublayer

• IEEE 802.11 defines 2 MAC sublayers:
1) Distributed coordination function (DCF) &
2) Point coordination function (PCF).

• The figure 15.6 shows the relationship between
1) Two MAC sublayers
2) LLC sublayer &
3) Physical layer.

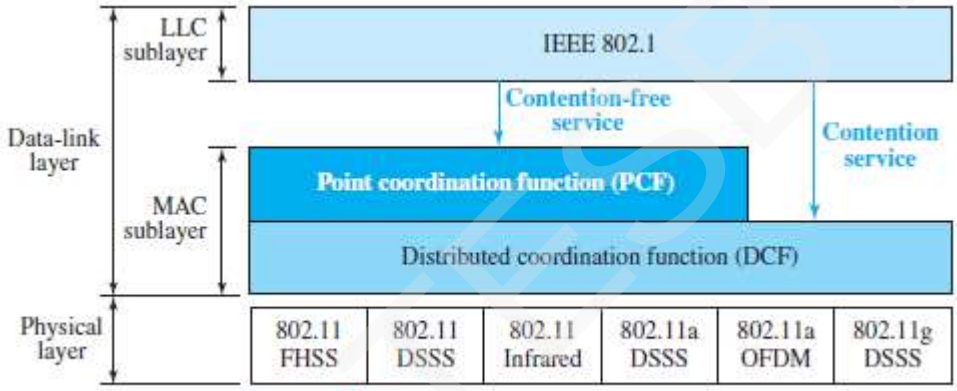


Figure 15.6 *MAC layers in IEEE 802.11 standard*

4.11.2.1 DCF

• One of the 2 protocols defined by IEEE at the MAC sublayer is called the distributed coordination function (DCF).
• DCF uses CSMA/CA as the access method.
• Wireless-LANs cannot implement CSMA/CD for 3 reasons:

1) For collision-detection, a station must be able to send data & receive collision-signals at the same time. This can mean costly stations and increased bandwidth requirements.
2) Collision may not be detected because of the hidden station problem.
3) The distance between stations can be great.
Signal fading could prevent a station at one end from hearing a collision at the other end.

• Process Flowchart: Figure 15.7 shows the process flowchart for CSMA/CA as used in wireless-LANs.

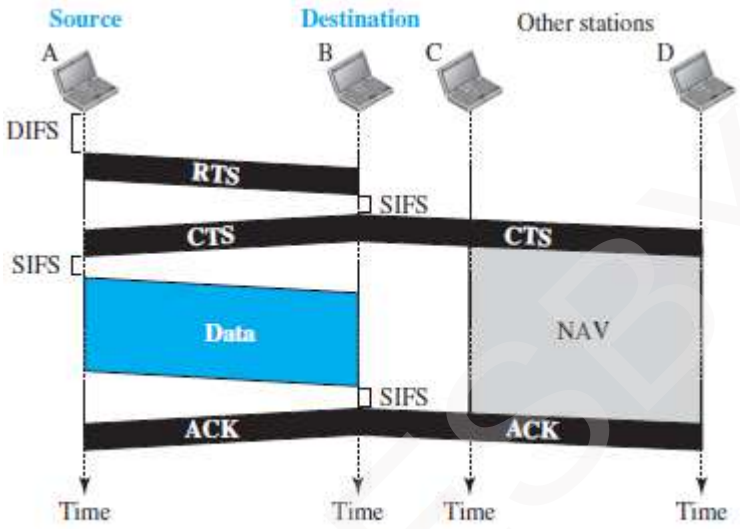


Figure 15.7 *CSMA/CA and NAV*

1) Before sending a frame, the source-station senses the medium by checking the energy-level at the carrier-frequency.
i) The channel uses a persistence strategy with back-off until the channel is idle.
ii) After the station is found to be idle,
→ the station waits for a period of time called the DIFS.
→ then the station sends a control frame called the RTS.

2) After receiving the RTS and waiting a period of time called the SIFS, the destination-station sends a control frame, called the CTS, to the source-station.
CTS frame indicate that the destination-station is ready to receive data.

3) The source-station sends data after waiting an amount of time equal to SIFS.

4) The destination-station, after waiting an amount of time equal to SIFS, sends an acknowledgment to show that the frame has been received.
Acknowledgment is needed in this protocol because the station does not have any means to check for the successful arrival of its data at the destination.
On the other hand, the lack of collision in CSMA/CD is a kind of indication to the source that data have arrived.

(DIFS → distributed inter frame space SIFS → short inter frame space)
(RTS → request to send CTS → clear to send)

4.11.2.1.1 Network Allocation Vector

• When a station sends an RTS frame, it includes the duration of time that it needs to occupy the channel (NAV → Network Allocation Vector).
• The stations that are affected by this transmission create a timer called a NAV.
• NAV shows how much time must pass before these stations are allowed to check the channel for idleness.
• Each time a station accesses the system and sends an RTS frame, other stations start their NAV.
• In other words, each station, before sensing the medium to see if it is idle, first checks its NAV to see if it has expired.

4.11.2.1.2 Collision During Handshaking

• Two or more stations may try to send RTS frames at the same time.
• These control frames may collide.
• However, because there is no mechanism for collision-detection, the sender assumes there has been a collision if it has not received a CTS frame from the receiver.
• The back-off strategy is employed, and the sender tries again.

4.11.2.2 PCF

• The PCF is an optional access method that can be implemented in an infrastructure network (not in an ad hoc network) (PCF → Point Coordination Function).
• The PCF is implemented on top of the DCF.
• The PCF is used mostly for time-sensitive transmission.
• PCF has a centralized, contention-free polling access method.
• The AP performs polling for stations that are capable of being polled.
• The stations are polled one after another, sending any data they have to the AP.
• To give priority to PCF over DCF, another set of inter-frame spaces has been defined: PIFS and SIFS.
    1) The SIFS is the same as that in DCF &
    2) PIFS (PCF IFS) is shorter than the DIFS.
• This means that if, at the same time, a station wants to use only DCF and an AP wants to use PCF, the AP has priority.
• Due to the priority of PCF over DCF, stations that only use DCF may not gain access to the medium.
• To prevent this, a repetition interval has been designed to cover both contention-free (PCF) and contention-based (DCF) traffic.
• The repetition interval, which is repeated continuously, starts with a special control frame, called a beacon frame.
• When the stations hear the beacon frame, they start their NAV for the duration of the contention-free period of the repetition interval.

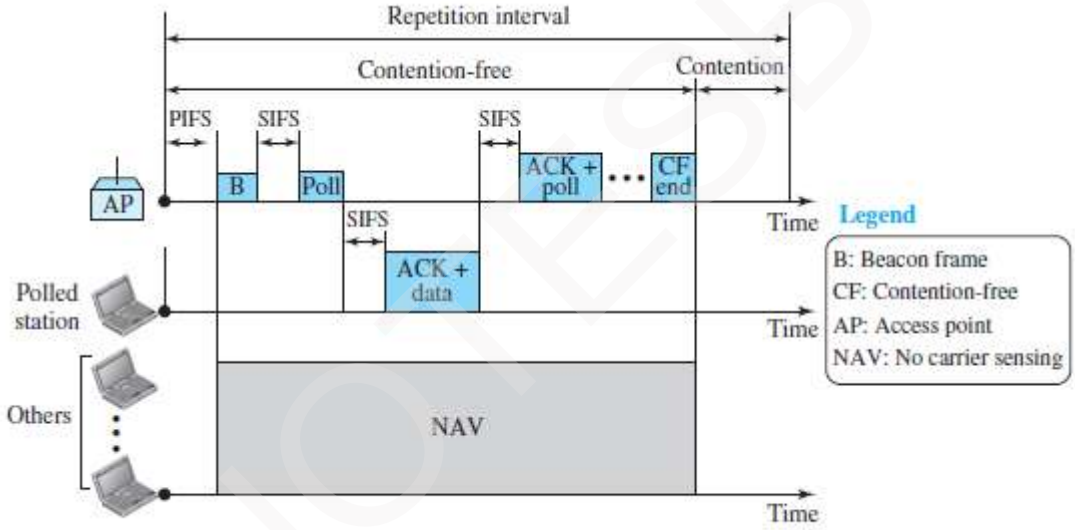


Figure 15.8 *Example of repetition interval*

• During the repetition interval, the PC (point controller) can send a poll frame, receive data, send an ACK, receive an ACK, or do any combination of these (802.11 uses piggybacking).
• At the end of the contention-free period, the PC sends a CF end (contention-free end) frame to allow the contention-based stations to use the medium.

4.11.2.2.1 Fragmentation

• The wireless environment is very noisy; a corrupt frame has to be retransmitted.
• The protocol, therefore, recommends fragmentation--the division of a large frame into smaller ones. It is more efficient to resend a small frame than a large one.

4.11.2.2.2 Frame Format

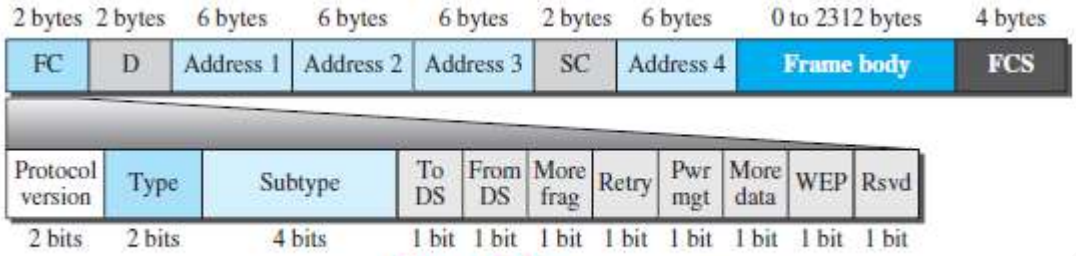


Figure 15.9 *Frame format*

• The MAC layer frame consists of nine fields (Figure 15.9):

1) Frame Control (FC)

➢ The FC field is 2 bytes long and defines the type of frame and some control information. The table describes the subfields.

Table 15.1 *Subfields in FC field*

Field	*Explanation*
Version	Current version is 0
Type	Type of information: management (00), control (01), or data (10)
Subtype	Subtype of each type (see Table 15.2)
To DS	Defined later
From DS	Defined later
More frag	When set to 1, means more fragments
Retry	When set to 1, means retransmitted frame
Pwr mgt	When set to 1, means station is in power management mode
More data	When set to 1, means station has more data to send
WEP	Wired equivalent privacy (encryption implemented)
Rsvd	Reserved

2) D

➢ In all frame types except one, this field defines the duration of the transmission that is used to set the value of NAV.

➢ In one control frame, this field defines the ID of the frame.

3) Addresses

➢ There are four address fields, each 6 bytes long.

➢ The meaning of each address field depends on the value of the ToDS and FromDS subfields.

4) Sequence Control

➢ This field defines the sequence number of the frame to be used in flow control.

5) Frame Body

➢ This field contains information based on the type and the subtype defined in the FC field.

➢ This field can be between 0 and 2312 bytes,

6) FCS

➢ The FCS contains a CRC-32 error detection sequence.

4.11.2.2.3 Frame Types

• A wireless-LAN defined by IEEE 802.11 has three categories of frames: 1.management frames, 2.control frames, and 3.data-frames.

1) Management Frames

➢ Management frames are used for the initial communication between stations and access points.

2) Control Frames

➢ Control frames are used for accessing the channel and acknowledging frames (Figure 15.10).

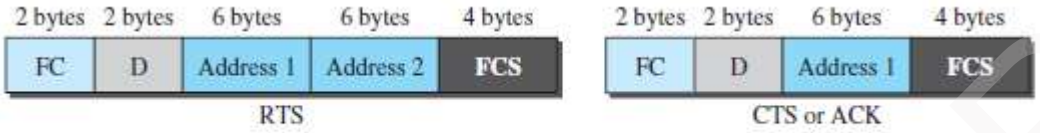


Figure 15.10 *Control frames*

➢ For control frames the value of the type field is 01; the values of the subtype fields for frames are shown in the table 14.2.

Table 15.2 *Values of subtype fields in control frames*

Subtype	*Meaning*
1011	Request to send (RTS)
1100	Clear to send (CTS)
1101	Acknowledgment (ACK)

3) Data Frames

➢ Data-frames are used for carrying data and control information.

4.11.3 Addressing Mechanism

• The IEEE 802.11 addressing mechanism specifies 4 cases, defined by the value of the 2 flags in the FC field, To DS and From DS.
• Each flag can be either 0 or 1, resulting in 4 different situations.
• The interpretation of the 4 addresses (address 1 to address 4) in the MAC frame depends on the value of these flags, as shown in the Table 15.3.

Table 15.3 *Addresses*

To DS	*From DS*	*Address 1*	*Address 2*	*Address 3*	*Address 4*
0	0	Destination	Source	BSS ID	N/A
0	1	Destination	Sending AP	Source	N/A
1	0	Receiving AP	Source	Destination	N/A
1	1	Receiving AP	Sending AP	Destination	Source

• Address 1 is always the address of the next device.
• Address 2 is always the address of the previous device.
• Address 3 is the address of the final destination-station if it is not defined by address 1.
• Address 4 is the address of the original source-station if it is not the same as address 2.

Case-1:00

- In this case, To DS = 0 and From DS = 0 (Figure 15.11a).
- This means that the frame is
  → not going to a distribution-system (To DS = 0) and i
  → not coming from a distribution-system (From DS = 0).
- The frame is going from one station in a BSS to another without passing through the distribution-system.
- The ACK frame should be sent to the original sender.

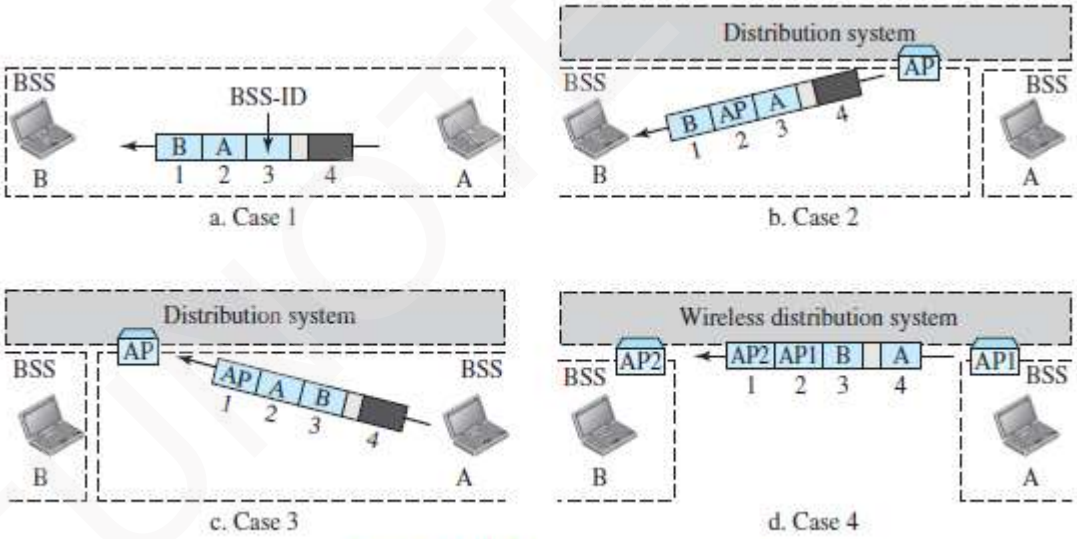


Figure 15.11 *Addressing mechanisms*

Case-2:01

- In this case, To DS = 0 and From DS = 1 (Figure 15.11b).
- This means that the frame is coming from a distribution-system (From DS = 1).
- The frame is coming from an AP and going to a station.
- The ACK should be sent to the AP.
- The address 3 contains the original sender of the frame (in another BSS).

Case-3:10

- In this case, To DS = 1 and From DS = 0 (Figure 15.11c).
- This means that the frame is going to a distribution-system (To DS = 1).
- The frame is going from a station to an AP. The ACK is sent to the original station.
- The address 3 contains the final destination of the frame (in another BSS).

Case-4:11

- In this case, To DS = 1 and From DS = 1 (Figure 15.11d).
- This is the case in which the distribution-system is also wireless.
- The frame is going from one AP to another AP in a wireless distribution-system.
- We do not need to define addresses if the distribution-system is a wired LAN because the frame in these cases has the format of a wired LAN frame (for example: Ethernet,).
- Here, we need four addresses to define
  - → original sender
  - → final destination, and
  - → two intermediate APs.

4.11.3.1 Exposed Station Problem

• In this problem, a station refrains from using a channel even when the channel is available for use.
• In the figure 14.12, station A is transmitting to station B.
• Station C has some data to send to station D, which can be sent without interfering with the transmission from A to B.
• However, station C is exposed to transmission from A i.e. station C hears what A is sending and thus refrains from sending.
• In other words, C is too conservative and wastes the capacity of the channel.

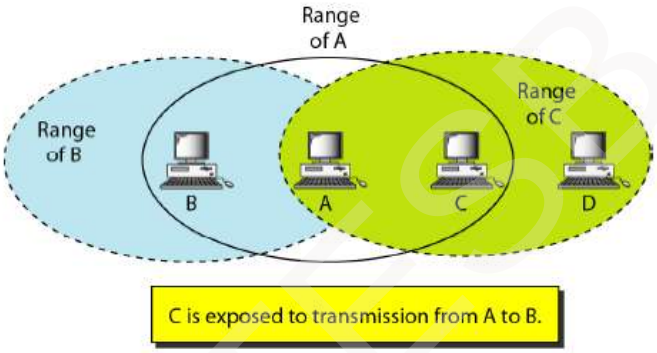


Figure 14.12 *Exposed station problem*

• The handshaking messages RTS and CTS cannot help in this case.
• Station C hears the RTS from A, but does not hear the CTS from B.
• Station C, after hearing the RTS from A, can wait for a time so that the CTS from B reaches A; it then sends an RTS to D to show that it needs to communicate with D.
• Both stations B and A may hear this RTS, but station A is in the sending state, not the receiving state.
• However, Station B responds with a CTS.
• The problem is here (Figure 15.12).

If station A has started sending its data, station C cannot hear the CTS from station D because of the collision; it cannot send its data to D. It remains exposed until A finishes sending its data as the figure shows.

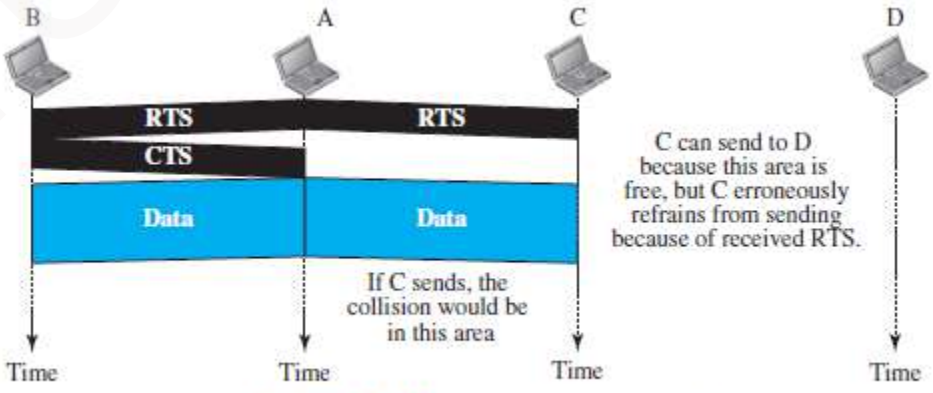


Figure 15.12 *Exposed station problem*

4.11.4 Physical Layer

Table 15.4 *Specifications*

IEEE	*Technique*	*Band*	*Modulation*	*Rate (Mbps)*
802.11	FHSS	2.400–4.835 GHz	FSK	1 and 2
	DSSS	2.400–4.835 GHz	PSK	1 and 2
	None	Infrared	PPM	1 and 2
802.11a	OFDM	5.725–5.850 GHz	PSK or QAM	6 to 54
802.11b	DSSS	2.400–4.835 GHz	PSK	5.5 and 11
802.11g	OFDM	2.400–4.835 GHz	Different	22 and 54
802.11n	OFDM	5.725–5.850 GHz	Different	600

- All implementations, except the infrared, operate in the ISM band.
- ISM band defines 3 unlicensed bands in the 3 ranges.
  - i) 902-928 MHz,
  - ii) 2.400--4.835 GHz, and
  - iii) 5.725-5.850 GHz. (ISM → industrial, scientific, and medical)

4.11.4.1 IEEE 802.11 FHSS

- IEEE 802.11 FHSS uses the FHSS method (Figure 15.13).
- FHSS uses the 2.4-GHz ISM band.
- The band is divided into 79 subbands of 1 MHz (and some guard bands).
- A pseudorandom number generator selects the hopping sequence.
- The modulation technique is either two-level FSK or four-level FSK with 1 or 2 bits/baud.
- This results in a data-rate of 1 or 2 Mbps

(FHSS → frequency-hopping spread spectrum DSSS → direct sequence spread spectrum)

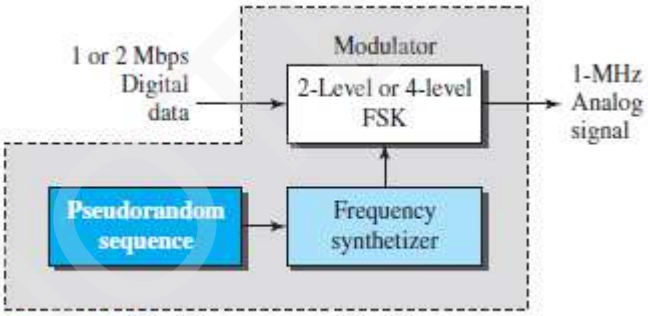


Figure 15.13 *Physical layer of IEEE 802.11 FHSS*

4.11.4.2 IEEE 802.11 DSSS

- IEEE 802.11 DSSS uses the DSSS method (Figure 15.14).
- DSSS uses the 2.4-GHz ISM band.
- The modulation technique in this specification is PSK at 1 Mbaud/s.
- The system allows 1 or 2 bits/baud (BPSK or QPSK).
- This results in a data-rate of 1 or 2 Mbps.

Figure 15.14 *Physical layer of IEEE 802.11 DSSS*

(HRDSSS → high-rate direct sequence spread spectrum CCK → complementary code keying)
(OFDM → orthogonal frequency-division multiplexing)

4.11.4.3 IEEE 802.11 Infrared

• IEEE 802.11 infrared uses infrared light in the range of 800 to 950 nm (Figure 15.15).
• The modulation technique is called pulse position modulation (PPM).
• For a 1-Mbps data-rate, a 4-bit sequence is first mapped into a 16-bit sequence in which only one bit is set to 1 and the rest are set to 0.
• For a 2-Mbps data-rate, a 2-bit sequence is first mapped into a 4-bit sequence in which only one bit is set to 1 and the rest are set to 0.
• The mapped sequences are then converted to optical signals; the presence of light specifies 1, the absence of light specifies 0.

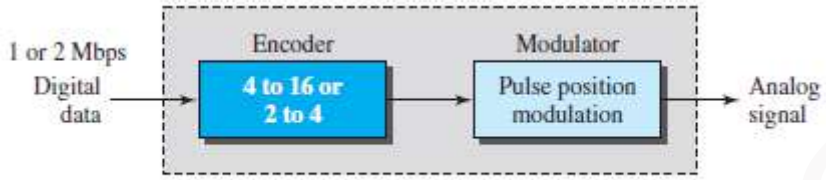


Figure 15.15 *Physical layer of IEEE 802.11 infrared*

4.11.4.4 IEEE 802.11a OFDM

• IEEE 802.11a OFDM describes the OFDM method for signal generation in a 5-GHz ISM band.
• OFDM is similar to FDM, with 2 major difference:
  1) All the subbands are used by one source at a given time.
  2) Sources contend with one another at the data-link-layer for access.
• The band is divided into 52 subbands. Out of which,
  1) 48 subbands are used for sending 48 groups of bits at a time.
  2) 4 subbands are used for sending control information.
• The scheme is similar to ADSL.
• Dividing the band into subbands diminishes the effects of interference.
• If the subbands are used randomly, security can also be increased.
• OFDM uses PSK and QAM for modulation.
• The common data-rates are
  i) 18 Mbps (PSK) and ii) 54 Mbps (QAM).

4.11.4.5 IEEE 802.11b DSSS

• IEEE 802.11 b DSSS describes the HRDSSS method for signal generation in the 2.4-GHz ISM band.
• HR-DSSS is similar to DSSS, with 1 major difference: HR-DSSS uses encoding method called CCK.
• CCK encodes 4 or 8 bits to one CCK symbol (Figure 15.16).
• To be backward compatible with DSSS, HR-DSSS defines 4 data-rates: 1, 2, 5.5, and 11 Mbps.
  1) The first two versions (1- & 2-Mbps) use the same modulation techniques as DSSS.
  2) The 5.5-Mbps version
    → uses BPSK and
    → transmits at 1.375 Mbaud/s with 4-bit CCK encoding.
  3) The 11-Mbps version
    → uses QPSK and
    → transmits at 1.375 Mbps with 8-bit CCK encoding.

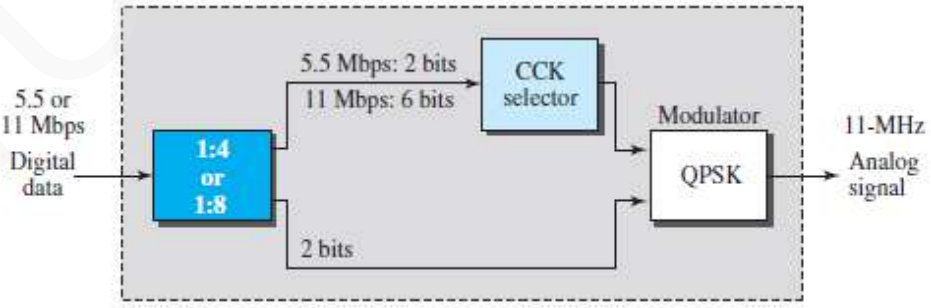


Figure 15.16 *Physical layer of IEEE 802.11b*

4.11.4.6 IEEE 802.11g

• This new specification defines forward error correction and OFDM using the 2.4-GHz ISM band.
• The modulation technique achieves a 22- or 54-Mbps data-rate.
• It is backward compatible with 802.11b, but the modulation technique is OFDM.

4.12 BLUETOOTH

• Bluetooth is a wireless-LAN technology designed to connect devices of different functions such as telephones, notebooks, computers, cameras, printers, coffee makers, and so on.
• A Bluetooth LAN is an ad hoc network. This means the network is formed spontaneously.
• The devices
→ find each other and
→ make a network called a piconet (Usually, devices are called gadgets)
• A Bluetooth LAN can even be connected to the Internet if one of the devices has this capability.
• By nature, a Bluetooth LAN cannot be large.
• If there are many devices that try to connect, there is confusion.
• Bluetooth technology has several applications.
1) Peripheral devices such as a wireless mouse/keyboard can communicate with the computer.
2) In a small health care center, monitoring-devices can communicate with sensor-devices.
3) Home security devices can connect different sensors to the main security controller.
4) Conference attendees can synchronize their laptop computers at a conference.
• Today, Bluetooth technology is the implementation of a protocol defined by the IEEE 802.15 standard.
• The standard defines a wireless PAN operable in an area the size of a room or a hall.
(PAN → personal-area network)

4.12.1 Architecture

• Bluetooth defines 2 types of networks: 1) Piconet and 2) Scatternet.

4.12.1.1 Piconets

➢ A Bluetooth network is called a piconet, or a small net. (Figure 15.17).
➢ A piconet can have up to 8 stations. Out of which
i) One of station is called the primary.
ii) The remaining stations are called secondaries.
➢ All the secondary-stations synchronize their clocks and hopping sequence with the primary station.
➢ A piconet can have only one primary station.
➢ The communication between the primary and the secondary can be one-to-one or one-to-many.

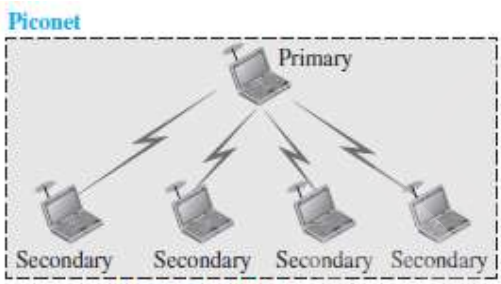


Figure 15.17 *Piconet*

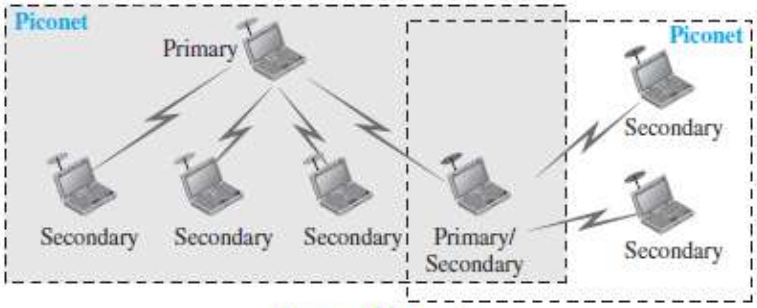


Figure 15.18 *Scatternet*

➢ Although a piconet can have a maximum of 7 secondaries, an additional 8 secondaries can be in the parked state.
➢ A secondary in a parked state is synchronized with the primary, but cannot take part in communication until it is moved from the parked state.
➢ Because only 8 stations can be active in a piconet, activating a station from the parked state means that an active station must go to the parked state.

4.12.1.2 Scatternet

➢ Piconets can be combined to form a scatternet (Figure 15.18).
➢ A station can be a member of 2 piconets.
➢ A secondary station in one piconet can be the primary in another piconet. This is called mediator station.
1) Acting as a secondary, mediator station can receive messages from the primary in the first piconet.
2) Acting as a primary, mediator station can deliver the message to secondaries in the second piconet.

4.12.1.3 Bluetooth Devices

• A Bluetooth device has a built-in short-range radio transmitter.
• The current data-rate is 1 Mbps with a 2.4-GHz bandwidth.
• This means that there is a possibility of interference between the IEEE 802.11b wireless-LANs and Bluetooth LANs.

4.12.2 Bluetooth Layers

• Bluetooth uses several layers that do not exactly match those of the Internet model (Figure 15.19).

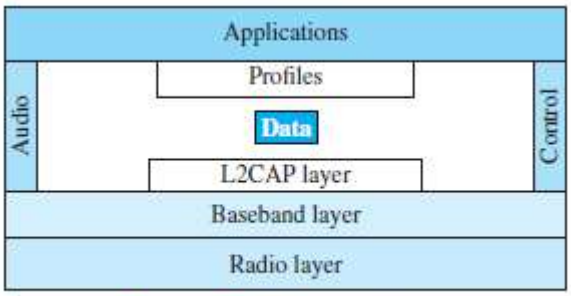


Figure 15.19 *Bluetooth layers*

4.12.2.1 Radio Layer

• The radio layer is roughly equivalent to the physical layer of the Internet model.
• Bluetooth devices are low-power and have a range of 10 m.

1) Band

➢ Bluetooth uses a 2.4-GHz ISM band divided into 79 channels of 1 MHz each.

2) FHSS

➢ Bluetooth uses the frequency-hopping spread spectrum (FHSS) method in the physical layer to avoid interference from other devices or other networks.
➢ Bluetooth hops 1600 times per second, which means that each device changes its modulation frequency 1600 times per second.
➢ A device uses a frequency for only 625 μs (1/1600 s) before it hops to another frequency; the dwell time is 625 μs.

3) Modulation

➢ To transform bits to a signal, Bluetooth uses a sophisticated version of FSK, called GFSK (FSK with Gaussian bandwidth filtering).
➢ GFSK has a carrier frequency.
➢ Bit 1 is represented by a frequency deviation above the carrier; bit 'a' is represented by a frequency deviation below the carrier.
➢ The frequencies, in megahertz, are defined according to the following formula for each channel:

$$f_c = 2402 + n \text{ MHz} \qquad n = 0, 1, 2, 3, \ldots, 78$$

➢ For example,
The first channel uses carrier frequency 2402 MHz (2.402 GHz).
The second channel uses carrier frequency 2403 MHz (2.403 GHz).

4.12.2.2 Baseband Layer

• The baseband layer is roughly equivalent to the MAC sublayer in LANs.
• The access method is TDMA.
• The primary and secondary communicate with each other using time slots.
• The length of a time slot is exactly the same as the dwell time, 625 μs.
• This means that during the time that one frequency is used, a sender sends a frame to a secondary, or a secondary sends a frame to the primary.
• The communication is only between the primary and a secondary; secondaries cannot communicate directly with one another.

4.12.2.2.1 TDMA

• Bluetooth uses a form of TDMA that is called TDD-TDMA (timedivision duplex TDMA).
• TDD-TDMA is a kind of half-duplex communication in which the secondary and receiver send and receive data, but not at the same time (halfduplex);
• However, the communication for each direction uses different hops.
• This is similar to walkie-talkies using different carrier frequencies.

Single-Secondary Communication

- If the piconet has only one secondary, the TDMA operation is very simple (Fig 15.21).
- The time is divided into slots of 625 μs.
- The primary uses even numbered slots (0, 2, 4, ...); the secondary uses odd-numbered slots (1, 3, 5,...).
- TDD-TDMA allows the primary and the secondary to communicate in half-duplex mode.
- In slot 0, the primary sends, and the secondary receives; in slot 1, the secondary sends, and the primary receives. The cycle is repeated.

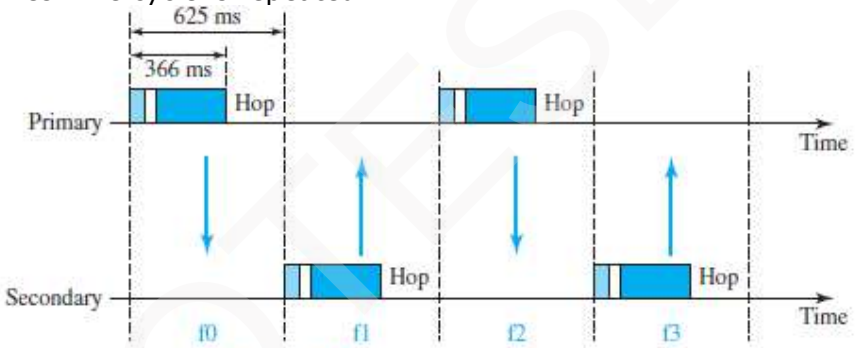


Figure 15.21 *Single-secondary communication*

4.12.2.2.2 Links

• Two types of links can be created between a primary and a secondary:
1) SCQ link (Synchronous Connection-oriented Link) and
2) ACL links (Asynchronous Connectionless Link).

1) SCA

➢ This link is used when avoiding latency is more important than data-integrity.
(Latency → delay in data delivery Integrity → error-free delivery)
➢ A physical-link is created between the primary and a secondary by reserving specific slots at regular intervals.
➢ The basic unit of connection is 2 slots. One slot is used for each direction.
➢ If a packet is damaged, it is never retransmitted.
➢ Application: Used for real-time audio where avoiding delay is all-important.
➢ A secondary
→ can create up to 3 SCQ links with the primary
→ can send digitized audio (PCM) at 64 kbps in each link.

2) ACL

➢ This link is used when data-integrity is more important than avoiding latency.
➢ If a payload encapsulated in the frame is corrupted, it is retransmitted.
➢ A secondary returns an ACL frame in the available odd-numbered slot if and only if the previous slot has been addressed to it.
➢ ACL can use one, three, or more slots and can achieve a maximum data-rate of 721 kbps.

4.12.2.2.3 Frame Types

• A frame in the baseband layer can be one of 3 types: 1) one-slot 2) three-slot or 3) five-slot.

1) One Slot Frame

➢ A slot is 625 μs.
➢ However, in a one-slot frame exchange, 259 μs is needed for hopping & control mechanisms.
➢ This means that a one-slot frame can last only 625 - 259, or 366 μs.
➢ With a 1-MHz bandwidth and 1 bit/Hz, the size of a one-slot frame is 366 bits.

2) Three Slot Frame

➢ A three-slot frame occupies 3 slots.
➢ However, since 259 μs is used for hopping, the length of the frame is 3 x 625 - 259 = 1616 μs or 1616 bits.
➢ A device that uses a three-slot frame remains at the same hop (at the same carrier frequency) for 3 slots.
➢ Even though only once hop number is used, 3 hop numbers are consumed.
➢ That means the hop number for each frame is equal to the first slot of the frame.

3) Five Slot Frame

➢ A five-slot frame also uses 259 bits for hopping, which means that the length of the frame is 5 x 625 - 259 =2866 bits.

4.12.2.2.4 Frame Format

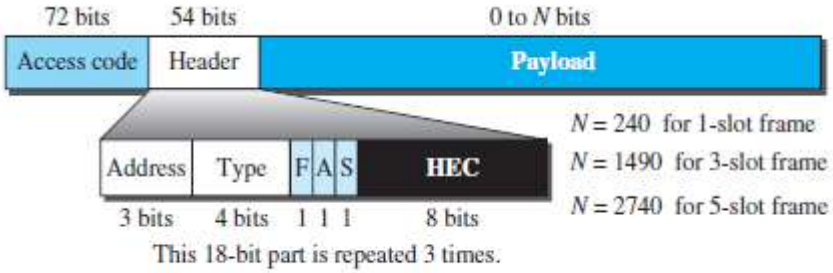


Figure 15.23 *Frame format types*

• The following describes each field (Figure 15.23):

1) Access Code

➢ This field contains synchronization bits and the identifier of the primary to distinguish the frame of one piconet from anothe

2) Header

➢ This field is a repeated I8-bit pattern. Each pattern has the following subfields:

i) Address

¤ This subfield can define up to 7 secondaries (1 to 7).

¤ If the address is zero, it is used for broadcast communication from the primary to all secondaries.

ii) Type

¤ This subfield defines the type of data coming from the upper layers.

iii) F

¤ This subfield is for flow control.

¤ When set (1), it indicates that the device is unable to receive more frames (buffer is full).

iv) A

¤ This subfield is for acknowledgment.

¤ Bluetooth uses Stop-and-Wait ARQ.

➢ 1 bit is sufficient for acknowledgment.

v) S

¤ This subfield holds a sequence number.

¤ Bluetooth uses Stop-and-Wait ARQ

¤ 1 bit is sufficient for sequence numbering.

vi) HEC (Header Error Correction)

¤ This subfield is a checksum to detect errors in each 18-bit header section.

➢ The header has three identical 18-bit sections.

➢ The receiver compares these three sections, bit by bit.

➢ If each of the corresponding bits is the same, the bit is accepted; if not, the majority opinion rules.

➢ This is a form of forward error correction (for the header only).

➢ This double error control is needed because the nature of the communication, via air, is very noisy.

➢ There is no retransmission in this sublayer.

3) Payload

➢ This subfield can be 0 to 2740 bits long.

➢ It contains data or control information coming from the upper layers.

4.12.2.3 L2CAP

- The L2CAP is roughly equivalent to the LLC sublayer in LANs (Figure 15.20).
- It is used for data exchange on an ACL link. (L2CAP → Logical Link Control and Adaptation Protocol)
- SCQ channels do not use L2CAP"(Figure 14.25)

Figure 15.20 *L2CAP data packet format*

- The following describes each field:

1) Length

- This field defines the size of the data, in bytes, coming from the upper layers.
- Data can be up to 65,535 bytes.

2) CID (Channel ID)

- This field defines a unique identifier for the virtual channel created at this level.

- The L2CAP has specific duties: 1) Multiplexing
  2) Segmentation and reassembly
  3) QoS (quality of service) and
  4) Group management.

1) Multiplexing

- The L2CAP can do multiplexing.
- At the sender site, L2CAP
  → accepts data from one of the upper-layer protocols
  → frames the data and
  → delivers the data to the baseband layer.
- At the receiver site, L2CAP
  → accepts a frame from the baseband layer
  → extracts the data, and
  → delivers the data to the appropriate protocol layer.
- It creates a kind of virtual channel.

2) Segmentation and Reassembly

- In the baseband layer, the maximum size of the payload field is 2774 bits, or 343 bytes.
- This includes 4 bytes to define the packet and packet-length.
- Therefore, the size of the packet that can arrive from an upper layer can only be 339 bytes.
- However, application layers sometimes need to send a data packet that can be up to 65,535 bytes (for example: an Internet packet).
- The L2CAP
  → divides the large packets into segments and
  → adds extra information to define the location of the segments in the original packet.
- The L2CAP segments the packet at the source and reassembles them at the destination.

3) QoS

- Bluetooth allows the stations to define a QoS level.
- If no QoS level is defined, Bluetooth defaults to best-effort service; it will do its best under the circumstances.

4) Group Management

- Another functionality of L2CAP is to allow devices to create a type of logical addressing between themselves.
- This is similar to multicasting.
- For example:
  2 or 3 secondary devices can be part of a multicast group to receive data from the primary.

MODULE-WISE QUESTIONS

MODULE 4: MULTIPLE ACCESS

1) Explain random access protocol. (4)
2) Explain pure ALOHA. (6*)
3) Explain slotted ALOHA. (4*)
4) Explain CSMA. (6*)
5) Explain different persistence methods of CSMA. (6*)
6) Explain CSMA/CA. (6*)
7) Explain CSMA/CD. (10*)
8) List & explain different controlled access protocols. (10*)
9) Explain reservation access method. (4*)
10) Explain polling access method. (6*)
11) Explain token passing access method. (6*)
12) List & explain channelization protocols. (10*)
13) Explain FDMA. (6*)
14) Explain TDMA. (6*)
15) Explain CDMA. (8*)

MODULE 4(CONT.): WIRED LANS -- ETHERNET

1) Explain frame format of standard ethernet. (8*)
2) Explain frame length of standard ethernet. (4*)
3) Explain addressing in standard ethernet. (6*)
4) Explain encoding in standard ethernet. (8)
5) Explain changes in the standard ethernet. (6)
6) List out 5 goals of fast-ethernet. Explain autonegotiation. (6*)
7) Explain encoding in fast-ethernet. (8)
8) Explain MAC Sublayer in gigabit-ethernet (6*)
9) Explain encoding in gigabit-ethernet. (4)

MODULE 4(CONT.): WIRELESS-LANS

1) Differentiate b/w wireless LAN & wired LAN with reference to architectural comparison (6)
2) Explain characteristics of wireless medium. (4)
3) Explain hidden station problem. (6*)
4) Explain architecture of IEEE 802.11 (10*)
5) Explain DCF in IEEE 802.11 (6)
6) Explain PCF in IEEE 802.11 (6)
7) Explain frame format of IEEE 802.11 (8*)
8) Explain frame types of IEEE 802.11 (8*)
9) Explain addressing in IEEE 802.11 (8*)
10) Explain exposed station problem. (6*)
11) Explain physical Layer of IEEE 802.11 (8)
12) Explain architecture of Bluetooth. (6*)
13) Explain layers of Bluetooth. (8*)
14) Explain radio Layer in Bluetooth. (6)
15) Explain baseband Layer in Bluetooth. (6)
16) Explain 2 types of links in Bluetooth. (6)
17) Explain frame format of Bluetooth. (6)
18) Explain L2CAP in Bluetooth. (6)

MODULE 5: TABLE OF CONTENTS

MODULE 5: OTHER WIRELESS NETWORKS

5.1 WiMAX

- WiMAX stands for Worldwide Interoperability for Microwave Access.
- Purpose of WiMAX:
  1) People want to have access to the Internet from home or office (fixed) where the wired access to the Internet is either not available or is expensive.
  2) People need to access the Internet when they are using their cellular phones (mobiles).
- WiMAX provides the "last mile" broadband wireless access.

5.1.1 Services

- WiMAX provides 2 types of services to subscribers: 1) Fixed and 2) Mobile.

1) Fixed WiMAX

- A base-station can use 3 different types of antenna to optimize the performance:
  1) Omni-directional 2) Sector or 3) Panel (Figure 16.1).
- WiMAX uses a beam-steering AAS (Adaptive Antenna System).
- While transmitting, antenna can focus its energy in the direction of the subscriber-station.
  While receiving, antenna can focus in the direction of the subscriber-station to receive maximum energy sent by the subscriber.

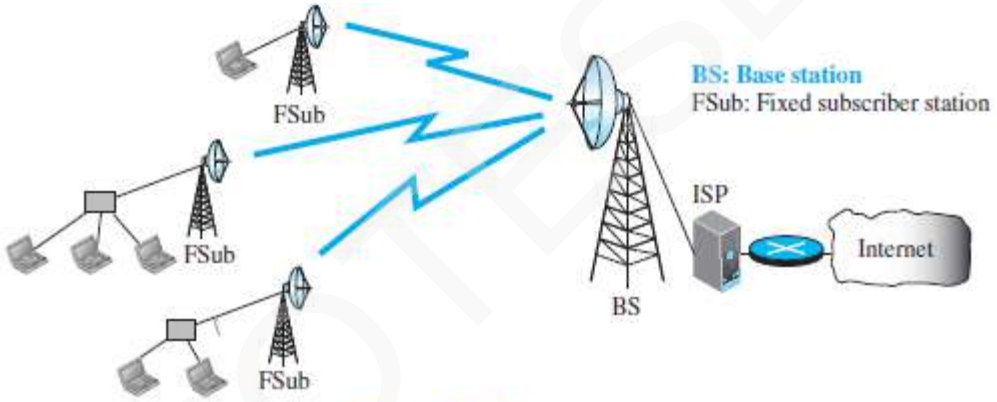


Figure 16.1 *Fixed WiMAX*

2) Mobile WiMAX

- The subscribers are mobile-stations that move from one place to another (Figure 16.2).

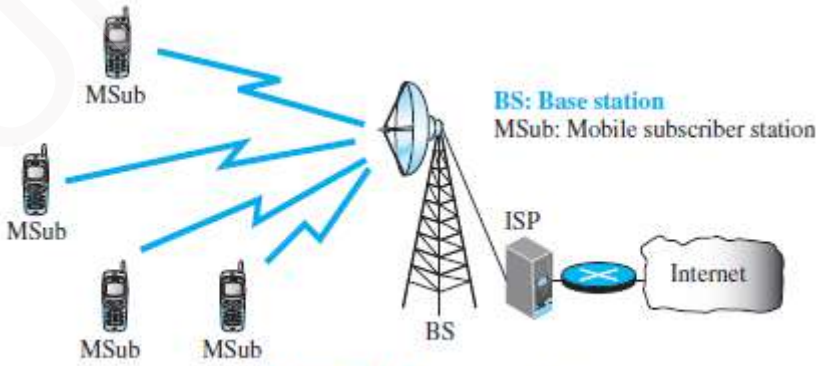


Figure 16.2 *Mobile WiMAX*

5.1.2 IEEE Project 802.16

- WiMAX is the result of the IEEE 802.16 project.
- The standard is also referred to as wireless local loop.

802.11 Projects	802.16 Projects
Standard for a wireless LAN	Standard for a wireless WAN
Defines a connectionless communication	Defines a connection-oriented service
Distance b/w base-station & host is very limited	Distance b/w base-station & host is above 10 km

- IEEE 802.16 was revised into 2 new standards:
  1) IEEE 802.16d which concentrates on the fixed WiMAX.
  2) IEEE 802.16e which defines the mobile WiMAX.

5.1.3 Layers in Project 802.16

- Here we discuss, following 2 layers (Figure 16.3):
  1) Data-link layer &
  2) Physical layer.
- The data-link layer is divided into 3 sublayers:
  1) Service Specific Convergence Sublayer
  2) Security Sublayer &
  3) MAC Sublayer.
- The physical layer is divided into 2 sublayers:
  1) Transmission Convergence Sublayer &
  2) Physical Medium Dependent Sublayer.

5.1.3.1 Data Link layer

1) Service Specific Convergence Sublayer

- This is actually the DLC sublayer revised for broadband wireless communication.
- It is devised for a connection-oriented service where each connection may benefit from a specific QoS

2) Security Sublayer

- This sublayer provides security for communication using WiMAX.
- The nature of wireless communication requires security.
- Security provided using encryption for information exchanged b/w subscriber-station & base-station.

3) MAC Sublayer

- The MAC sublayer defines the access method and the format of the frame.
- This sublayer is designed for connection-oriented service.
- The packets are routed from the base-station to the subscriber-station using a connection identifier.
- Connection identifier remains same during the duration of the communication.

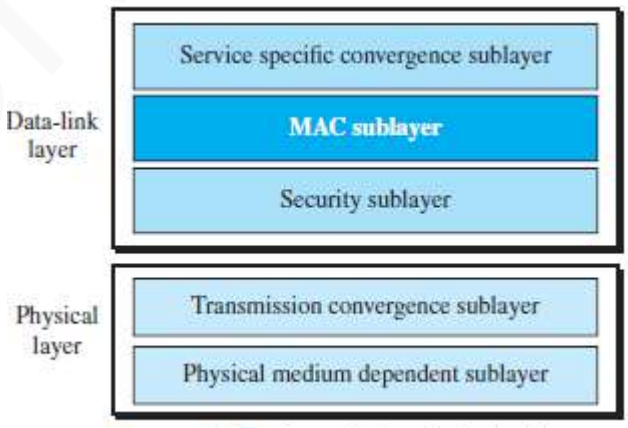


Figure 16.3 *Data-link and physical layers*

5.1.3.2 Physical Layer

1) Transmission Convergence Sublayer

- This sublayer uses TDD.
- TDD a variation of TDM designed for duplex (bidirectional) communication.
- Each frame is made of 2 subframes (Figure 16.5):
  - **1) Downstream Subframes:** carry data from the base-station to the subscribers.
  - **2) Upstream Subframes:** carry data from the subscribers to the base-station.
- Each subframe is divided into slots.

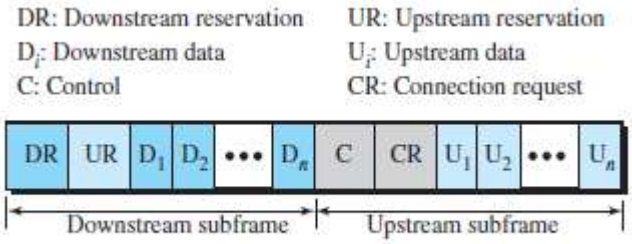


Figure 16.5 *WiMAX frame structure at the physical layer*

2) Physical Medium Dependent Sublayer

- This sublayer is in continuous revision.

i) Originally, 802.16 defined the band 10-66 GHz.

802.16 defined following modulations:

QPSK used for long-distance communication.
QAM-16 used for medium-distance communication.
QAM-64 used for short-distance communication.

ii) Later, IEEE defined 802.16d (fixed WiMAX), which added the band 2-11 GHz (compatible with wireless LANs) using the OFDM.

iii) Sometime later, IEEE defined 802.16e (mobile WiMAX) and added SOFDM.

(TDD → Time-Division Duplex FEC → forward error correction)
(OFDM → Orthogonal Frequency-Division Multiplexing)
(SOFDM → scalable orthogonal frequency division multiplexing)

5.1.3.3 MAC Sublayer

- The MAC sublayer defines the access method and the format of the frame.
- This sublayer is designed for connection-oriented service.
- The packets are routed from the base-station to the subscriber-station using a connection identifier.
- Connection identifier remains same during the duration of the communication.
- Here we discuss, following issues: 1) Access Method 2) Frame Format 3) Addressing

1) Access Method

- WiMAX uses the reservation (scheduling) access method.
- Base-station needs to make a slot-reservation before sending a data to a subscriber-station
  Each subscriber-station needs to make a reservation before sending a data to the base-station

2) Frame Format

- Two types of frames (Figure 16.4):
  1) Generic Frame is used to send and receive payload.
  2) Control Frame is used only during the connection establishment.
- Both frame-types use a 6-byte generic header.
  However, some bytes have different interpretations in different frame types.

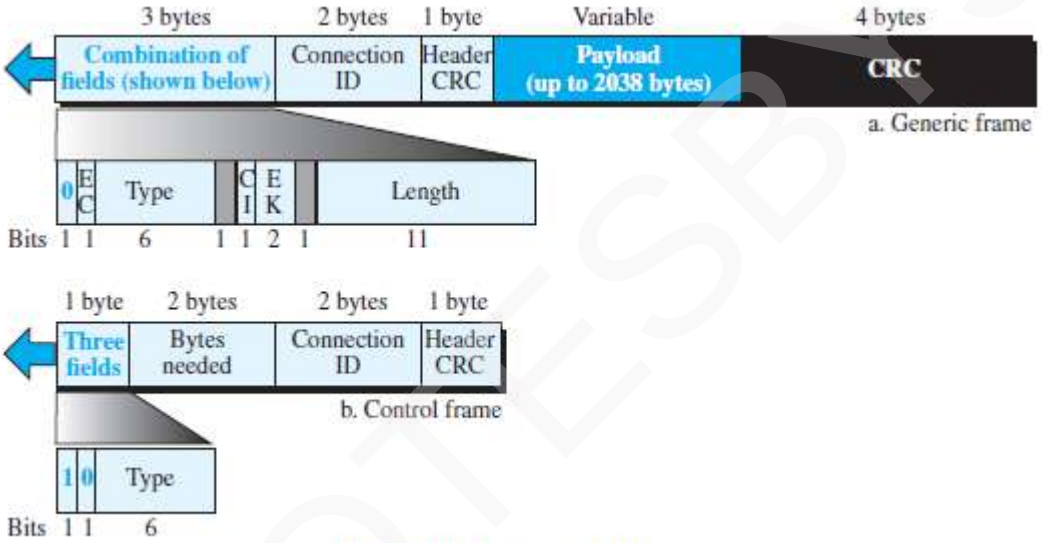


Figure 16.4 *WiMAX MAC frame format*

- The frame contains following fields:

1) First bit

- The first bit in a frame is the frame identifier.
- If first bit = 0, the frame is a generic frame.
  If first bit = 1, the frame is a control frame.

2) EC (Encryption Control)

- This field uses one bit to define whether the frame should be encrypted for security purpose.
- If EC = 0, it means no encryption.
  If EC = 1, it means the frame needs to be encrypted at the security sublayer.

3) Type

- This field is used to define the type of the payload.
- This field is only present in the generic frame.
- The payload can be a packed-load or a fragmented-load.

4) CI (Checksum ID)

- This field defines whether the checksum field should be present or not.
- If the payload is multimedia, FEC is applied to the frame and there is no need for checksum.

5) EK (Encryption Key)

- This field defines one of the 4 keys for encryption if encryption is required.

6) Length

- This field defines the total length of the frame.
- This field is only present in the generic frame.
- This field is replaced by the bytes needed field in the control frame.

7) Bytes Needed
- This field defines the number of bytes needed for allocated slots in the physical layer.

8) Connection ID
- This field defines the connection identifier for the current connection.

9) Header CRC
- Both types of frames need to have header CRC field.
- Header CRC is used to check whether the header itself is corrupted.
- This field uses the polynomial ($x^8 + x^2 + x + 1$) as the divisor.

10) Payload
- This field defines the payload.
- Payload is encapsulated in the frame from the service specific convergence sublayer.
- This field is not needed in the control frame.

11) CRC
- This field is used for error detection over the whole frame.

3) Addressing
• Each subscriber and base-station typically has a 48-bit MAC address.
• However, there is no source or destination address field.
• The reason is that the combination of source and destination addresses are mapped to a VCI during the connection-establishing phase.
• This protocol is a connection-oriented protocol that uses a VCI (Virtual Connection Identifier).
• Then, each frame uses the same connection identifier for the duration of data transfer

16.2 CELLULAR TELEPHONY

- Cellular telephony is designed to provide communications
  - → between two moving units called mobile-stations (MSs) or
  - → between one mobile-station and one stationary unit called a land unit (Figure 16.6).
- A service-provider is responsible for
  - → locating & tracking a caller
  - → assigning a channel to the call and
  - → transferring the channel from base-station to base-station as the caller moves out-of-range.
- Each cellular service-area is divided into small regions called cells.
- Each cell contains an antenna.
- Each cell is controlled by AC powered network-station called the base-station (BS).
- Each base-station is controlled by a switching office called a mobile-switching-center (MSC).
- MSC coordinates communication between all the base-stations and the telephone central office.
- MSC is a computerized center that is responsible for
  - → connecting calls
  - → recording call information and
  - → billing.
- Cell-size is not fixed; Cell-size can be increased or decreased depending on population of the area.
- Cell-radius = 1 to 12 mi.
- Compared to low-density areas, high-density areas require many smaller cells to meet traffic demands.
- Cell-size is optimized to prevent the interference of adjacent cell-signals.

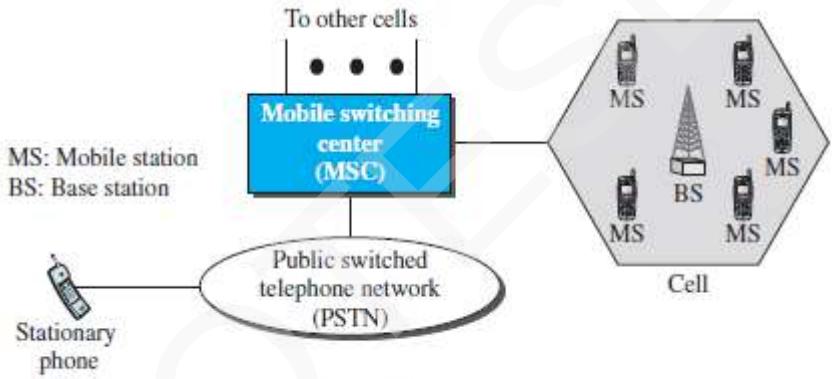


Figure 16.6 *Cellular system*

5.2.1 Operation

5.2.1.1 Frequency Reuse Principle

- In general, neighboring-cells cannot use the same set of frequencies for communication.
- Using same set of frequencies may create interference for the users located near the cell-boundaries.
- However,
  - → set of frequencies available is limited and
  - → frequencies need to be reused.
- A frequency reuse pattern is a configuration of N cells. Where N = reuse factor
- Each cell uses a unique set of frequencies.
- When the pattern is repeated, the frequencies can be reused.
- There are several different patterns (Figure 16.7).
- The numbers in the cells define the pattern.
- The cells with the same number in a pattern can use the same set of frequencies. These cells are called the reusing cells.

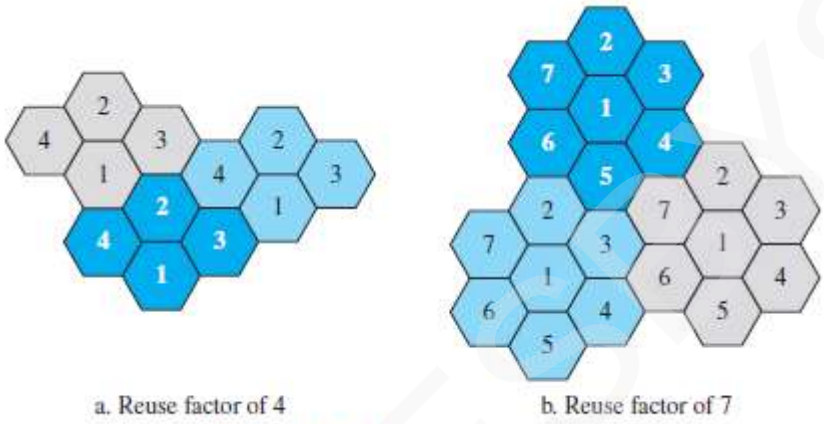


Figure 16.7 *Frequency reuse patterns*

5.2.1.2 Transmitting

- Procedure to place a call from a mobile-station:
  1) The caller
     - → enters a phone number and
     - → presses the send button.
  2) The mobile-station
     - → scans the band to determine setup channel with a strong signal and
     - → sends the data (phone number) to the closest base-station.
  3) The base-station sends the data to the MSC.
  4) The MSC sends the data on to the telephone central office.
  5) If called party is available, a connection is made and the result is relayed back to the MSC.
  6) The MSC assigns an unused voice channel to the call, and a connection is established.
  7) The mobile-station automatically adjusts its tuning to the new channel.
  8) Finally, voice communication can begin.

5.2.1.3 Receiving

- Procedure to receive a call from a mobile-station:
  1) When a mobile phone is called, the telephone central office sends phone number to the MSC.
  2) MSC searches for the location of the mobile-station by sending query-signals to each cell in a process. This is called paging.
  3) When the mobile-station is found, the MSC transmits a ringing signal.
  4) When the mobile-station answers, the MSC assigns a voice channel to the call.
  5) Finally, voice communication can begin.

5.2.1.4 Handoff

- During a conversation, the mobile-station may move from one cell to another.
- Problem: When the mobile-station goes to cell-boundary, the signal becomes weak.
- To solve this problem, the MSC monitors the level of the signal every few seconds.
- If signal-strength decreases, MSC determines a new cell to accommodate the communication.
- Then, MSC changes the channel carrying the call (hands signal off from old channel to a new one).
- Two types of Handoff: 1) Hard Handoff 2) Soft Handoff

1) Hard Handoff

- Early systems used a hard handoff.
- A mobile-station only communicates with one base-station.
- When the MS moves from one cell to another cell,
  - i) Firstly, communication must be broken with the old base-station.
  - ii) Then, communication can be established with the new base-station.
- This may create a rough transition.

2) Soft Handoff

- New systems use a soft handoff.
- A mobile-station can communicate with two base-stations at the same time.
- When the MS moves from one cell to another cell,
  - i) Firstly, communication must be broken with the old base-station.
  - ii) Then, the same communication may continue with the new base-station.

5.2.1.5 Roaming

- Roaming means that the user
  - → can have access to communication or
  - → can be reached where there is coverage.
- Usually, a service-provider has limited coverage.
- Neighboring service-providers can provide extended coverage through a roaming contract.

5.2.2 First Generation (1G)

- The first generation was designed for voice communication using analog signals.
- The main system evolved in the first generation: AMPS (Advanced Mobile Phone System).

5.2.2.1 AMPS

- This system is a 1G analog cellular system.
- The system uses FDMA to separate channels in a link.
- Here we discuss, two issues: 1) Bands 2) Transmission

1) Bands

- The system operates in the ISM 800-MHz band.
- The system uses 2 separate channels (Figure 16.8):
  - i) First channel is used for forward communication (base-station to mobile-station)
    Band range: 869 to 894 MHz
  - ii) Second channel is used for reverse communication (mobile-station to base-station).
    Band range: 824 to 849 MHz

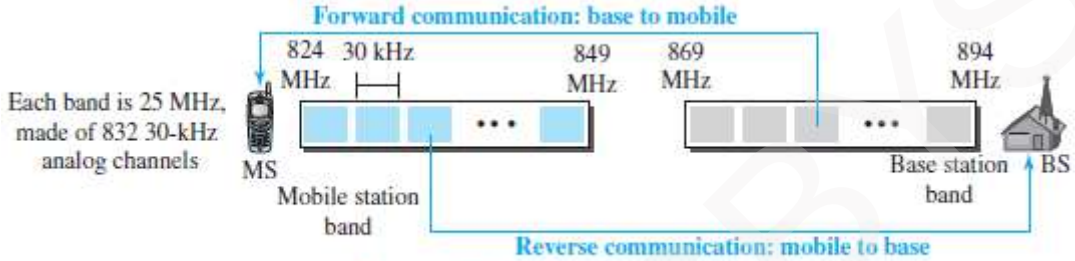


Figure 16.8 *Cellular bands for AMPS*

2) Transmission

- The system uses FM and FSK for modulation (Figure 16.9).
  - i) Voice channels are modulated using FM.
  - ii) Control channels are modulated using FSK to create 30-kHz analog signals.
- The system uses FDMA to divide each 25-MHz band into 30-kHz channels.

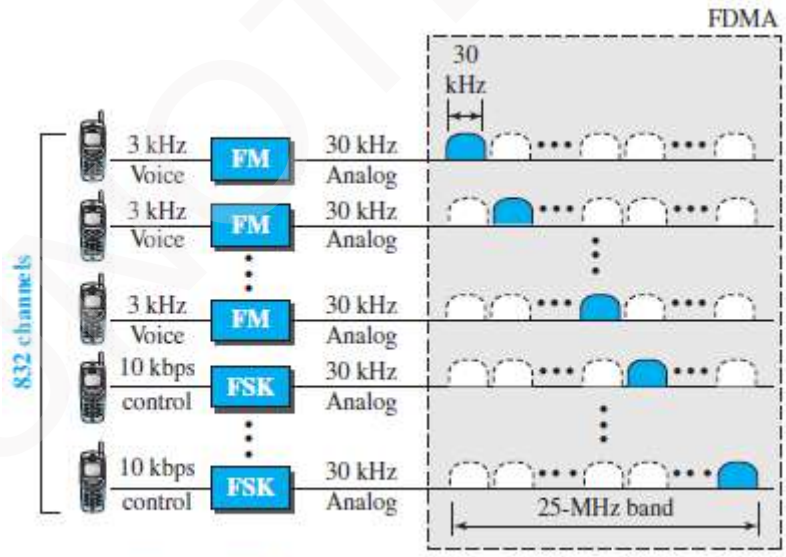


Figure 16.9 *AMPS reverse communication band*

5.2.3 Second Generation (2G)

- The second generation was designed for higher-quality voice communication using digital signals.
- 1G vs. 2G:
  1) The first generation was designed for analog voice communication.
  2) The second generation was mainly designed for digital voice communication.
- Three major systems evolved in the second generation:
  1) D-AMPS (digital AMPS)
  2) GSM (Global System for Mobile communication) and
  3) IS-95 (Interim Standard).

5.2.3.1 D-AMPS

- D-AMPS (Digital AMPS) was improved version of analog AMPS.
- D-AMPS was backward-compatible with AMPS.
- Thus, in a cell,
  1) First telephone may use AMPS and
  2) Second telephone may use D-AMPS.
- Here we discuss, two issues: 1) Bands 2) Transmission

1) Band

➢ The system uses the same bands and channels as AMPS (Figure 16.10).

2) Transmission

➢ Each voice channel is digitized using a very complex PCM and compression technique.

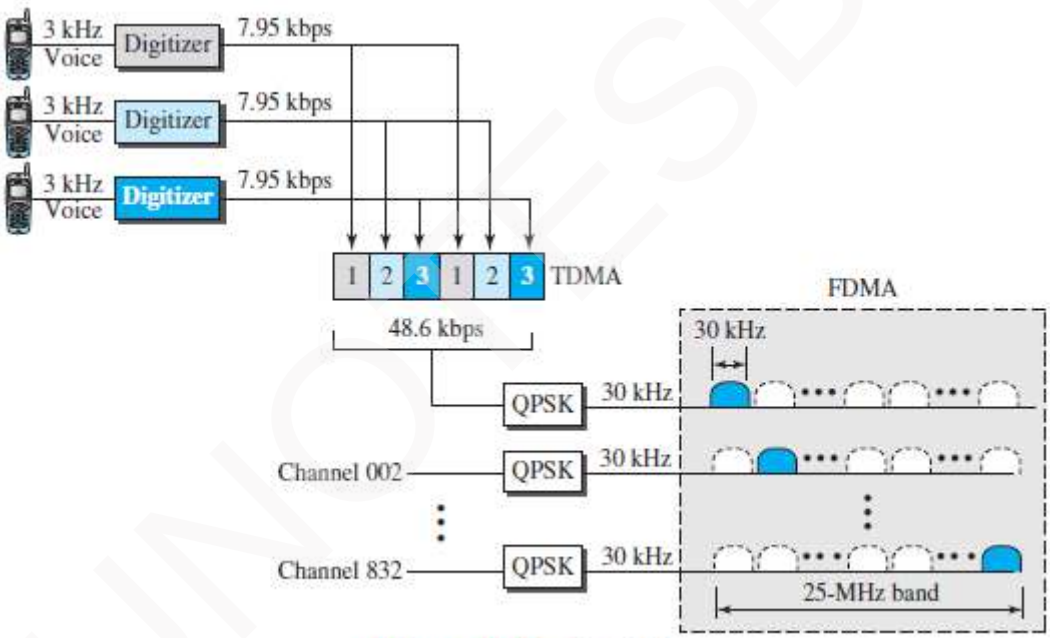


Figure 16.10 *D-AMPS*

5.2.3.2 GSM

• Aim of GSM: to replace a number of incompatible 1G technologies.

• Here we discuss, two issues: 1) Bands 2) Transmission

1) Bands

- The system uses two bands for duplex communication (Figure 16.11).
- Each band is 25 MHz in width.
- Each band is divided into 124 channels of 200 kHz.

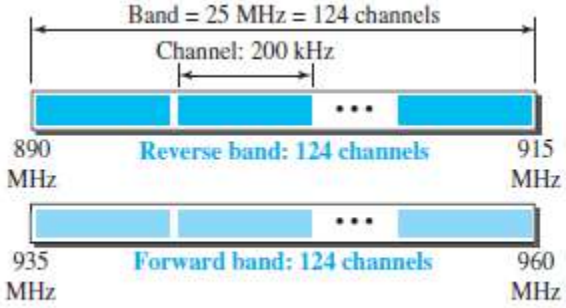


Figure 16.11 *GSM bands*

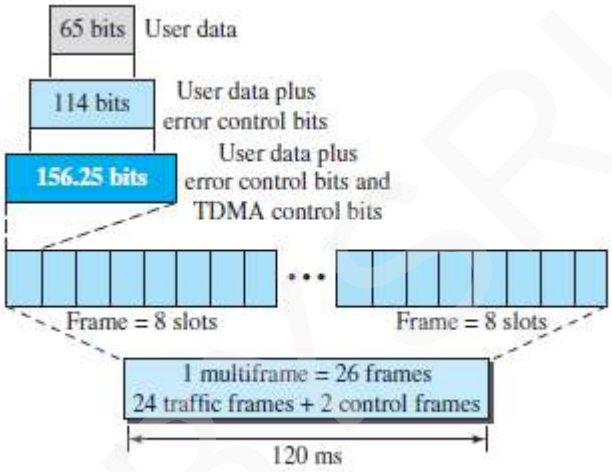


Figure 16.13 *Multiframe components*

2) Transmission

- Each voice channel is digitized and compressed to a 13-kbps digital signal (Figure 16.12).
- Each slot carries 156.25 bits.
- Eight slots share a frame (TDMA).
- 26 frames also share a multiframe (TDMA).
- We can calculate the bit rate of each channel as follows.

$$\text{Channel data rate} = (1/120\text{ ms}) \times 26 \times 8 \times 156.25 = 270.8\text{ kbps}$$

- Each 270.8-kbps digital channel modulates a carrier using GMSK (a form of FSK); the result is a 200-kHz analog signal.
- Finally, 124 analog channels of 200 kHz are combined using FDMA. The result is a 25-MHz band (Figure 16.13).

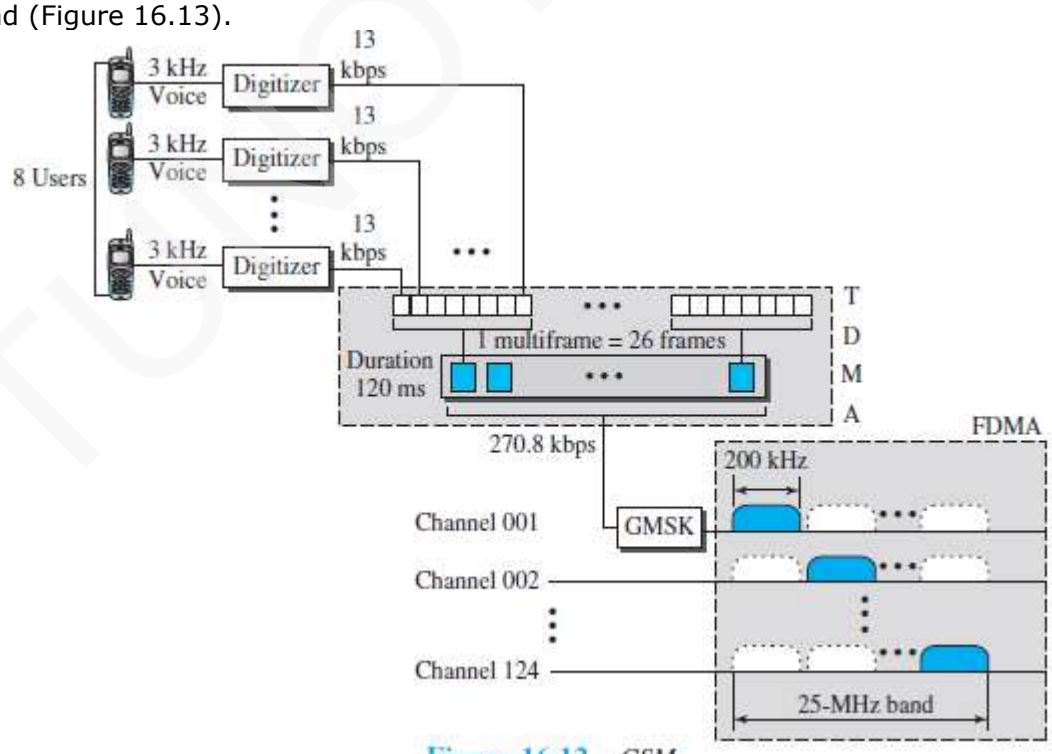


Figure 16.12 *GSM*

5.2.3.3 IS-95

- The system is based on CDMA and DSSS.
- Here we discuss, following 6 issues: 1) Bands 2) Transmission 3) Synchronization 4) Two Data-rate Sets 5) Frequency-Reuse Factor 6) Soft Handoff

1) Bands

- The system uses two bands for duplex communication.
- The bands can be ISM 800-MHz band or ISM 1900-MHz band.
- Each band is divided into 20 channels of 1.228 MHz.
- Each service-provider is allotted 10 channels.
- IS-95 can be used in parallel with AMPS.
- Each IS-95 channel is equivalent to 41 AMPS channels (41 × 30 kHz = 1.23 MHz).

2) Transmission

- Two types of Transmission:

i) Forward Transmission (base to mobile)

¤ Communications between the base and all mobiles are synchronized.

¤ The base sends synchronized data to all mobiles (Figure 16.14).

ii) Reverse Transmission (mobile to base)

¤ The use of CDMA in the forward direction is possible because the pilot channel sends a continuous sequence of 1s to synchronize transmission.

¤ The synchronization is not used in the reverse direction because we need an entity to do that, which is not feasible.

¤ Instead of CDMA, the reverse channels use DSSS (Figure 16.15).

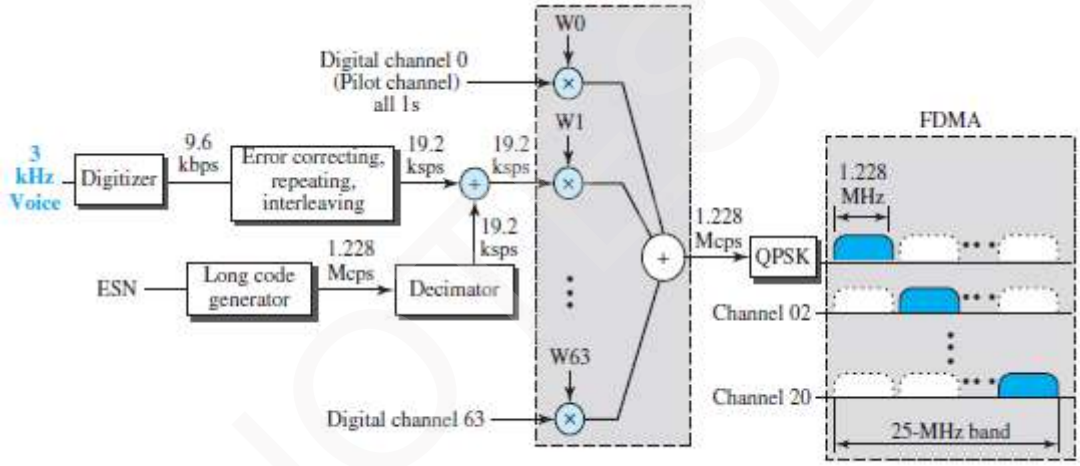


Figure 16.14 *IS-95 forward transmission*

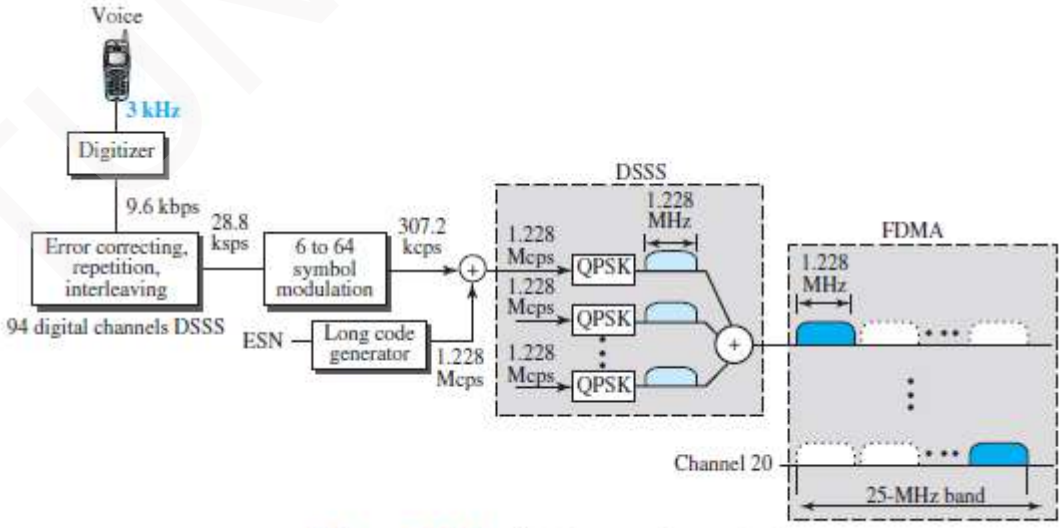


Figure 16.15 *IS-95 reverse transmission*

3) Synchronization

- All base channels need to be synchronized to use CDMA.
- To provide synchronization, bases use the services of a satellite system (GPS).

4) Two Data Rate Sets

- IS-95 defines two data-rate sets:
  - i) The first set defines 9600, 4800, 2400, and 1200 bps.
  - ii) The second set defines 14,400, 7200, 3600, and 1800 bps.

5) Frequency Reuse Factor

- The frequency-reuse factor is normally 1 because the interference from neighboring cells cannot affect CDMA or DSSS transmission.

6) Soft Handoff

- Every base-station continuously broadcasts signals using its pilot channel.
- Thus, a mobile-station can detect the pilot signal from its cell and neighboring cells.
- This enables a mobile-station to do a soft handoff.

5.2.4 Third Generation (3G)

• 3G cellular telephony provides both digital data and voice communication.
• For example: Using a Smartphone,
A person can talk to anyone else in the world.
A person can download a movie, surf the Internet or play games.
• Interesting characteristics: the Smartphone is always connected; we do not need to dial a number to connect to the Internet. (IMT → Internet Mobile Communication)
• Some objectives defined by the blueprint IMT-2000 (3G working group):
1) Voice quality comparable to that of the existing public telephone network.
2) Data-rate of
→ 144 kbps for access in a moving vehicle (car)
→ 384 kbps for access as the user walks (pedestrians) and
→ 2 Mbps for the stationary user (office or home).
3) Support for packet-switched and circuit-switched data services.
4) A band of 2 GHz.
5) Bandwidths of 2 MHz.
6) Interface to the Internet

5.2.4.1 IMT-2000 Radio Interfaces

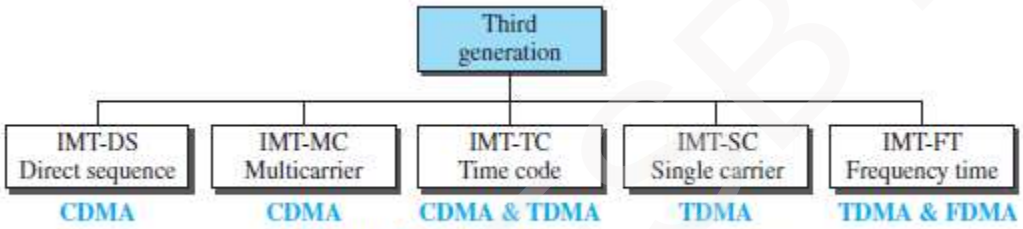


Figure 16.16 *IMT-2000 radio interfaces*

• Radio interfaces (wireless standards) adopted by IMT-2000 (Figure 16.16):

1) IMT-DS
- This uses a version of CDMA called W-CDMA (wideband CDMA).
- W-CDMA uses a 5-MHz bandwidth.
- It is compatible with the CDMA used in IS-95.

2) IMT-MC
- This was known as CDMA 2000.
- It is an evolution of CDMA technology used in IS-95 channels.
- It combines
  → new wideband (15-MHz) spread spectrum &
  → narrowband (1.25-MHz) CDMA of IS-95.
- It is backward-compatible with IS-95.
- It allows communication on multiple 1.25-MHz channels up to 15 MHz.

3) IMT-TC
- This uses a combination of W-CDMA and TDMA.
- It tries to reach the IMT-2000 goals by adding TDMA multiplexing to W-CDMA.

4) IMT-SC
- This uses only TDMA.

5) IMT-FT
- This uses a combination of FDMA and TDMA.

5.2.5 Fourth Generation (4G)

- 4G cellular telephony is expected to be a complete evolution in wireless communications.
- Some objectives defined by the 4G working group:
  1) A spectrally efficient system.
  2) High network capacity.
  3) Data-rate of
     → 100 Mbps for access in a moving vehicle
     → 1 Gbps for stationary users and
     → 100 Mbps between any two points in the world.
  4) Smooth handoff across heterogeneous networks.
  5) Seamless connectivity and global roaming across multiple networks.
  6) High quality of service for next generation multimedia support.
  7) Interoperability with existing wireless standards.
  8) All IP, packet-switched, networks.
- 4G is only packet-based networks.
- 4G supports IPv6.
- 4G provides better multicast, security, and route optimization capabilities.
- Here we discuss, following issues: 1) Access Scheme 2) Modulation 3) Radio System 4) Antenna 5) Applications

1) Access Scheme

- To increase efficiency,
  i) capacity, ii) scalability & iii) new access techniques are being considered for 4G.
- For example:
  i) OFDMA and IFDMA are being considered for the downlink & uplink of the next generation UMTS.
  ii) MC-CDMA is proposed for the IEEE 802.20 standard.

2) Modulation

- More efficient 64-QAM is being proposed for use with the LTE standards.

3) Radio System

- The 4G uses a SDR system.
- The components of an SDR are pieces of software and thus flexible.
- The SDR can change its program to shift its frequencies to mitigate frequency interference.

4) Antenna

- The MIMO and MU-MIMO antenna system is proposed for 4G.
- Using this antenna, 4G allows independent streams to be transmitted simultaneously from all the antennas to increase the data-rate.
- MIMO also allows the transmitter and receiver coordinates to move to an open frequency when interference occurs.

5) Applications

- At the present rates of 15-30 Mbps, 4G is capable of providing users with streaming high-definition television.
- At 100 Mbps, the content of a DVD-5 can be downloaded within about 5 minutes for offline access.

(OFDMA → Orthogonal FDMA IFDMA → interleaved FDMA)
(LTE → Long Term Evolution SDR → Software Defined Radio)
(MIMO → multiple-input multiple-output MU-MIMO →multiuser MIMO)
(UMTS → Universal Mobile Telecommunications System)
(MC-CDMA → multicarrier code division multiple access)

5.3 SATELLITE NETWORKS

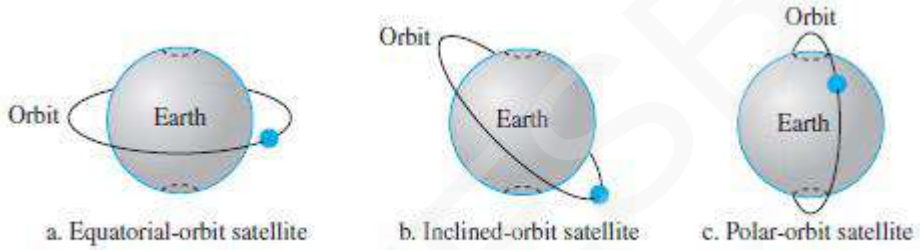
• A satellite network is a combination of nodes that provides communication from one point on the Earth to another.
• A node can be
→ Satellite
→ Earth station or
→ End-user terminal/telephone
• Like cellular networks, satellite networks divide the planet into cells.
• Satellites can provide transmission capability to and from any location on Earth.
• Advantages:
1) Satellite makes high-quality communication available to undeveloped parts of the world.
2) Cost effective: A huge investment in ground-based infrastructure is not required.

5.3.1 General Issues for Operation of Satellites

• Three issues related to the operation of satellites:
1) Orbits 2) Footprint 3) Frequency Bands for Satellite Communication

1) Orbits

• An artificial satellite needs to have an orbit.
• An orbit is the path in which the satellite travels around the Earth.
• The orbit can be equatorial, inclined, or polar (Figure 16.17.)



a. Equatorial-orbit satellite b. Inclined-orbit satellite c. Polar-orbit satellite

Figure 16.17 *Satellite orbits*

• The period means the time required for a satellite to make a complete trip around the Earth.
• The period of a satellite is determined by Kepler's law.
• Kepler's law defines period as a function of the distance of the satellite from the center of the Earth.

Example 5.1

What is the period of the moon, according to Kepler's law?

$$\text{Period} = C \times \text{distance}^{1.5}$$

Here C is a constant approximately equal to 1/100. The period is in seconds and the distance in kilometers.

Solution

The moon is located approximately 384,000 km above the Earth. The radius of the Earth is 6378 km. Applying the formula, we get the following.

$$\text{Period} = (1/100) \times (384{,}000 + 6378)^{1.5} = 2{,}439{,}090 \text{ s} = 1 \text{ month}$$

Example 5.2

According to Kepler's law, what is the period of a satellite that is located at an orbit approximately 35,786 km above the Earth?

Solution

Applying the formula, we get the following.

$$\text{Period} = (1/100) \times (35{,}786 + 6378)^{1.5} = 86{,}579 \text{ s} = 24 \text{ h}$$

This means that a satellite located at 35,786 km has a period of 24 h, which is the same as the rotation period of the Earth. A satellite like this is said to be *stationary* to the Earth. The orbit is called a *geostationary orbit*.

2) Footprint

- Satellites process microwaves with bidirectional antennas (line-of-sight).
- Normally, the signal from a satellite is aimed at a specific area called the footprint.
- The signal-power at the center of the footprint is maximum.
- The signal-power decreases, as we move out from the footprint-center.
- The boundary of the footprint is the location where the power-level is at a predefined threshold.

3) Frequency Bands for Satellite Communication

- For satellite communication, the frequencies reserved are in the GHz range.
- Each satellite sends and receives over 2 different bands (Table 16.1):
  - **1) Uplink:** refers to the transmission from the Earth to the satellite.
  - **2) Downlink:** refers to the transmission from the satellite to the Earth.

Table 16.1 *Satellite frequency bands*

Band	*Downlink, GHz*	*Uplink, GHz*	*Bandwidth, MHz*
L	1.5	1.6	15
S	1.9	2.2	70
C	4.0	6.0	500
Ku	11.0	14.0	500
Ka	20.0	30.0	3500

4) Three Categories of Satellites

- Three categories of satellites based on the location of the orbit (Figure 16.18):
  - **1) Geostationary Earth Orbit (GEO)**
    - There is only one orbit, at an altitude of 36000 km, for the GEO satellite.
  - **2) Low Earth Orbit (LEO)**
    - LEO satellites are below an altitude of 2000 km.
  - **3) Medium Earth Orbit (MEO)**
    - MEO satellites are located at altitudes between 5000 and 15,000 km.

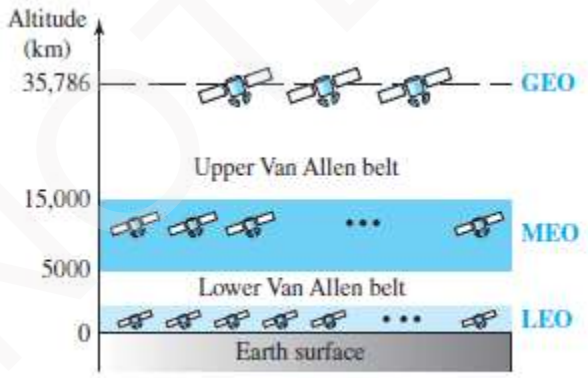


Figure 16.18 *Satellite orbit altitudes*

5.3.2 GEO Satellites

- There is only one orbit at an altitude of 36,000 km (Figure 16.19).
- Because orbital speed is based on the distance from the planet, only one orbit can be geostationary.
- The orbit occurs at the equatorial plane.
- Sending-antenna must have receiving-antenna in LOS (Line-of-sight).
- Problem: A satellite that moves faster/slower than Earth's rotation is useful only for short periods.
  Solution: To ensure constant communication, the satellite must move at same speed as the Earth.
  Thus, the satellite seems to remain fixed above a certain spot.

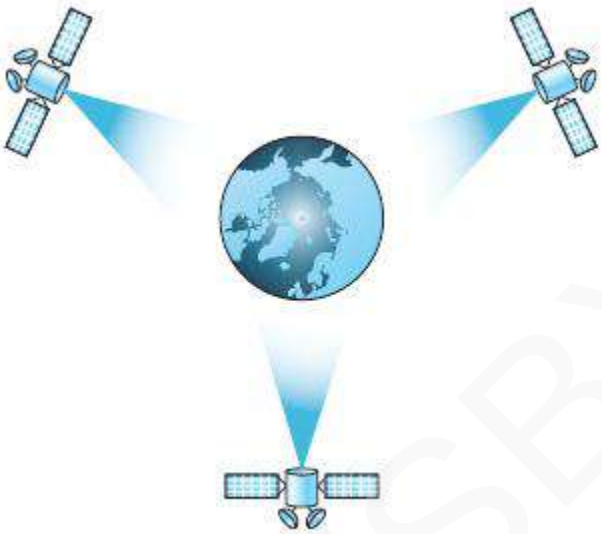


Figure 16.19 *Satellites in geostationary orbit*

5.3.3 MEO Satellites

- MEO satellites are located at altitudes between 5000 and 15,000 km.
- Example: Global Positioning System (GPS)

5.3.3.1 Global Positioning System

- GPS consists of 24 satellites in 6 orbits (Figure 16.20).
- GPS is used for land, sea, and air navigation to provide time and location for vehicles and ships.
- The orbits and the locations of the satellites in each orbit are designed systematically.
  For example: At any time, 4 satellites are visible from any point on Earth.
- A GPS receiver has an almanac (or calendar) that tells the current position of each satellite.
- Here we discuss, 4 issues: 1) Trilateration 2) Measuring the Distance 3) Synchronization 4) Application

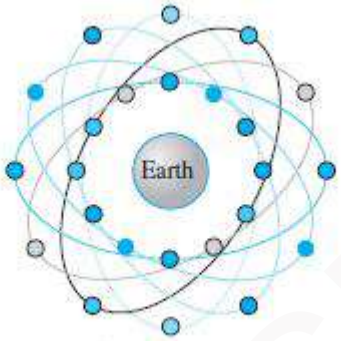
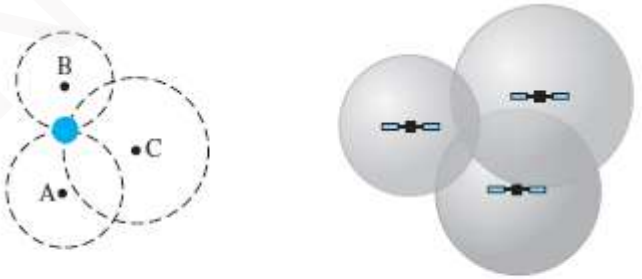


Figure 16.20 *Orbits for global positioning system (GPS) satellites*

1) Trilateration

- GPS is based on a principle called trilateration.
- Trilateration means using three distances.
- For example (Figure 16.21a):
  - On a plane, if we know our distance from three points, we know exactly where we are.
  - Assume we are 10 miles away from point A, 12 miles away from point B, and 15 miles away from point C.
  - If we draw three circles with the centers at A, B, and C, we must be somewhere on circle A, somewhere on circle B, and somewhere on circle C.
  - These three circles meet at one single point; this is our position (Figure 16.21a).



a. Two-dimensional trilateration b. Three-dimensional trilateration

Figure 16.21 *Trilateration on a plane*

  - In three-dimensional space, the situation is different.
  - Three spheres meet in two points, (Figure 16.21b).
  - We need at least four spheres to find our exact position in space (longitude, latitude, and altitude).

2) Measuring the Distance

- Trilateration principle can find our location on the Earth if we know
  → our distance from 3 satellites and
  → position of each satellite.
- The position of each satellite can be calculated by a GPS receiver.
- Then, the GPS receiver needs to find its distance from at least three GPS satellites.
- The distance is measured using a principle called one-way ranging.

3) Synchronization

- Satellites use atomic clocks, which are precise and can function synchronously with each other.
- The receiver's clock is a normal quartz clock.
- However, there is no way to synchronize receiver's clock with the satellite's clock.
- There is an unknown offset between the satellite-clocks and the receiver-clock.
- The unknown offset introduces a corresponding offset in the distance calculation.
- Because of the offset, the measured distance is called a pseudo-range.

4) Applications

1) GPS is used by military forces.
   - For example:
     Thousands of portable GPS receivers were used during the WW2 by foot soldiers, vehicles, and helicopters.

2) GPS is used in navigation.
   - For example:
     The driver of a car can find the location of the car.

3) GPS is used for clock synchronization.

5.3.4 LEO Satellites

- LEO satellites have polar orbits.
- Usually, a LEO system has a cellular type of access (similar to the cellular telephone system).
- Specifications:

  Altitude = 500 to 2000 km
  Rotation Period = 90 to 120 min
  Satellite Speed =20,000 to 25,000 km/h
  Footprint Diameter = 8000 km
  Round-Trip Time < 20 ms
- Because LEO satellites are close to Earth, 20 ms RTT is normally acceptable for audio communication.
- A LEO system is made of a group of satellites that work together as a network.
- Each satellite acts as a switch.
- Different types of links (Figure 16.22):

  1) ISLs (Inter-Satellite Links)
  - Satellites that are close to each other are connected through ISLs.

  2) UML (User Mobile Link)
  - A mobile system communicates with the satellite through a UML.

  3) GWL (Gate Way Link)
  - A satellite communicates with the Earth station (gateway) through a GWL.

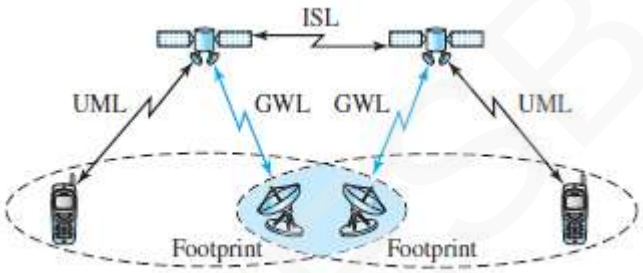


Figure 16.22 *LEO satellite system*

- LEO satellites can be divided into three categories:

  1) Little LEO
  - Operating frequency < 1 GHz.
  - They are mostly used for low data-rate messaging.

  2) Big LEO
  - Operating frequency = 1 to 3 GHz.
  - Examples: Globalstar & Iridium

	Globalstar	Iridium
Orbit-Altitude	1400 km	750 km
No. of Satellites	48	66
No. of Orbits	6	6
Satellites per Orbit	8	11

  3) Broadband LEO
  - Broadband LEO provides communication similar to fiber-optic networks.
  - Example: Teledesic
  - Teledesic satellite provides fiber-optic-like communication (broadband channels, low error rate, and low delay).
  - Main purpose: To provide broadband Internet access for users all over the world.

MODULE 5(CONT.): NETWORK LAYER PROTOCOLS

5.4 Network Layer Protocols

• The network layer contains following 4 protocols (Figure 19.1):

1) Internet Protocol (IP)

➢ IP is the main protocol responsible for packetizing, forwarding, and delivery of a packet at the network layer.

2) Internet Control Message Protocol (ICMP)

➢ ICMP helps IP to handle some errors that may occur in the network-layer delivery.

3) Internet Group Management Protocol (IGMP)

➢ IGMP is used to help IPv4 in multicasting.

4) Address Resolution Protocol (ARP)

➢ ARP is used to glue the network and data-link layers in mapping network-layer addresses to link-layer addresses.

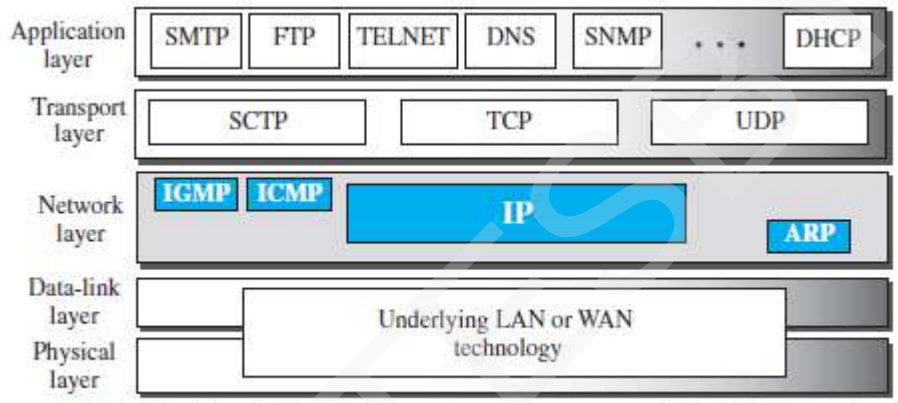


Figure 19.1 *Position of IP and other network-layer protocols in TCP/IP protocol suite*

5.5 INTERNET PROTOCOL (IP)

5.5.1 Internet Protocol (IP)

- IP is main protocol responsible for packetizing, forwarding & delivery of a packet at network layer.
- IP is an unreliable datagram protocol.
- IP provides a best-effort delivery service.
- The term best-effort means that the packets can
  - → be corrupted
  - → be lost or
  - → arrive out-of-order.
- If reliability is important, IP must be paired with a TCP which is reliable transport-layer protocol.
- IP is a connectionless protocol.
- IP uses the datagram approach.
  1) Each datagram is handled independently.
  2) Each datagram can follow a different route to the destination.
  3) Datagrams may arrive out-of-order at the destination.

5.5.2 Datagram Format

- IP uses the packets called datagrams.
- A datagram consist of 2 parts (Figure 19.2): 1) Payload 2) Header.

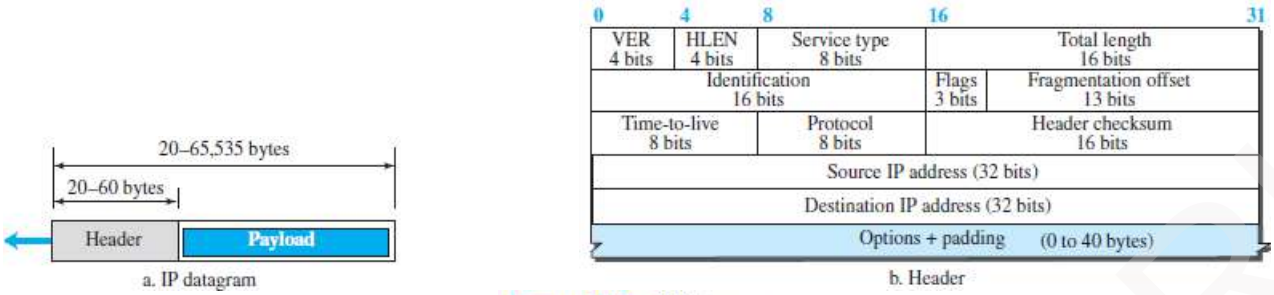


Figure 19.2 *IP datagram*

1) Payload

- Payload (or Data) is the main reason for creating a datagram.
- Payload is the packet coming from other protocols that use the service of IP.

2) Header

- Header contains information essential to routing and delivery.
- IP header contains following fields:

1) Version Number (VER)

➢ This field indicates version number used by the packet. Current version=4

2) Header Length (HLEN)

➢ This field specifies length of header.

➢ When a device receives a datagram, the device needs to know

→ when the header stops and

→ when the data starts.

3) Service Type

➢ This field specifies priority of packet based on delay, throughput, reliability & cost requirements.

4) Total Length

➢ This field specifies the total length of the datagram (header plus data).

➢ Maximum length=65535 bytes.

5) Identification, Flags, and Fragmentation Offset

➢ These 3 fields are used for fragmentation and reassembly of the datagram.

➢ Fragmentation occurs when the size of the datagram is larger than the MTU of the network.

6) Time-to-Live (TTL)

➢ This field is indicates amount of time, the packet is allowed to remain in the network.

➢ If TTL becomes 0 before packet reaches destination, the router

→ discards packet and

→ sends an error-message back to the source.

7) Protocol

➢ This field specifies upper-layer protocol that is to receive the packet at the destination-host.

➢ For example (Figure 19.3):

For TCP, protocol = 6 For UDP, protocol = 17

8) Header Checksum

➢ This field is used to verify integrity of header only.

➢ If the verification process fails, packet is discarded.

9) Source and Destination Addresses

➢ These 2 fields contain the IP addresses of source and destination hosts.

10) Options

➢ This field allows the packet to request special features such as

→ security level

→ route to be taken by packet and

→ timestamp at each router.

➢ This field can also be used for network testing and debugging.

11) Padding

➢ This field is used to make the header a multiple of 32-bit words.

Example 5.3

An IPv4 packet has arrived with the first 8 bits as $(01000010)_2$. The receiver discards the packet. Why?

Solution

There is an error in this packet. The 4 leftmost bits $(0100)_2$ show the version, which is correct. The next 4 bits $(0010)_2$ show an invalid header length $(2 \times 4 = 8)$. The minimum number of bytes in the header must be 20. The packet has been corrupted in transmission.

Example 5.4

In an IPv4 packet, the value of HLEN is $(1000)_2$. How many bytes of options are being carried by this packet?

Solution

The HLEN value is 8, which means the total number of bytes in the header is 8×4, or 32 bytes. The first 20 bytes are the base header, the next 12 bytes are the options.

Example 5.5

In an IPv4 packet, the value of HLEN is 5, and the value of the total length field is $(0028)_{16}$. How many bytes of data are being carried by this packet?

Solution

The HLEN value is 5, which means the total number of bytes in the header is 5×4, or 20 bytes (no options). The total length is $(0028)_{16}$ or 40 bytes, which means the packet is carrying 20 bytes of data $(40 - 20)$.

(Comparing a datagram to a postal package.

1) Payload is the content of the package.
2) Header is only the information written on the package).

5.5.3 Fragmentation

5.5.3.1 Maximum Transfer Unit (MTU)

• Each network imposes a restriction on maximum size of packet that can be carried. This is called the MTU (maximum transmission unit).

• For example:

For Ethernet, MTU = 1500 bytes

For FDDI, MTU = 4464 bytes

• When IP wants send a packet that is larger than MTU of physical-network, IP breaks packet into smaller fragments. This is called fragmentation (Figure 19.5).

• Designers have decided to make the maximum length of IP datagram = 65,535 bytes. This ensures that the IP protocol is independent of the physical network,

• When a datagram is fragmented, each fragment has its own header.

• A fragmented-datagram may itself be fragmented if it encounters a network with an even smaller MTU.

• Source host or router is responsible for fragmentation of original datagram into the fragments.

Destination host is responsible for reassembling the fragments into the original datagram.

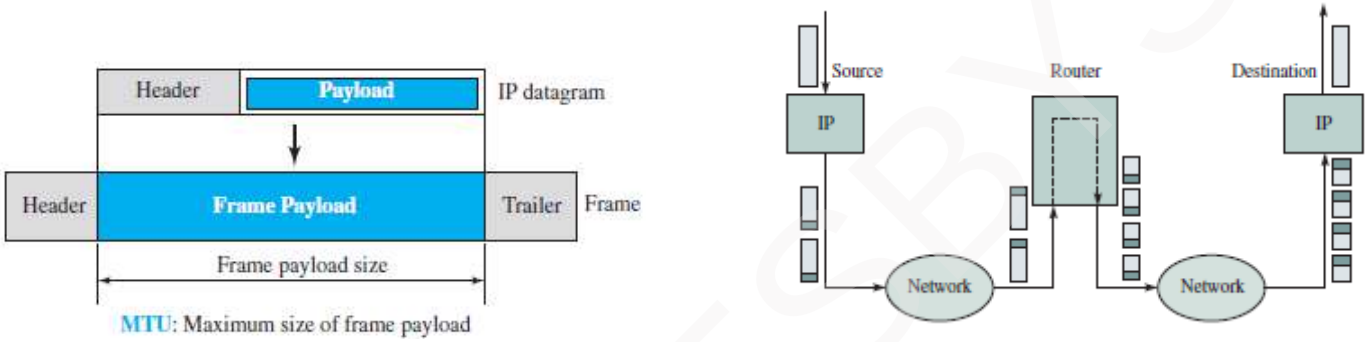


Figure 19.5b: Packet Fragmentation

5.5.3.2 Fields Related to Fragmentation & Reassembly

• Three fields in the IP header are used to manage fragmentation and reassembly:

1) Identification 2) Flags 3) Fragmentation offset.

1) Identification

• This field is used to identify to which datagram a particular fragment belongs to (so that fragments for different packets do not get mixed up).

• To guarantee uniqueness, the IP protocol uses an up-counter to label the datagrams.

• When the IP protocol sends a datagram, IP protocol

→ copies the current value of the counter to the identification field and

→ increments the up-counter by 1.

• When a datagram is fragmented, the value in the identification field is copied into all fragments.

• The identification number helps the destination in reassembling the datagram.

2) Flags

• This field has 3 bits.

1) The leftmost bit is not used.

2) DF bit (Don't Fragment):

i) If DF=1, the router should not fragment the datagram. Then, the router

→ discards the datagram and

→ sends an error-message to the source host.

ii) If DF=0, the router can fragment the datagram if necessary.

3) MF bit (More Fragment):

i) If MF=1, there are some more fragments to come.

ii) If MF=0, this is last fragment.

3) Fragmentation Offset

• This field identifies location of a fragment in a packet.

• This field is the offset of the data in the original datagram.

Example 5.6

A packet has arrived with an M bit value of 0. Is this the first fragment, the last fragment, or a middle fragment? Do we know if the packet was fragmented?

Solution

If the M bit is 0, it means that there are no more fragments; the fragment is the last one. However, we cannot say if the original packet was fragmented or not. A nonfragmented packet is considered the last fragment.

Example 5.7

A packet has arrived with an M bit value of 1. Is this the first fragment, the last fragment, or a middle fragment? Do we know if the packet was fragmented?

Solution

If the M bit is 1, it means that there is at least one more fragment. This fragment can be the first one or a middle one, but not the last one. We don't know if it is the first one or a middle one; we need more information (the value of the fragmentation offset).

Example 5.8

A packet has arrived with an M bit value of 1 and a fragmentation offset value of 0. Is this the first fragment, the last fragment, or a middle fragment?

Solution

Because the M bit is 1, it is either the first fragment or a middle one. Because the offset value is 0, it is the first fragment.

Example 5.9

A packet has arrived in which the offset value is 100. What is the number of the first byte? Do we know the number of the last byte?

Solution

To find the number of the first byte, we multiply the offset value by 8. This means that the first byte number is 800. We cannot determine the number of the last byte unless we know the length of the data.

Example 5.10

A packet has arrived in which the offset value is 100, the value of HLEN is 5, and the value of the total length field is 100. What are the numbers of the first byte and the last byte?

Solution

The first byte number is $100 \times 8 = 800$. The total length is 100 bytes, and the header length is 20 bytes (5×4), which means that there are 80 bytes in this datagram. If the first byte number is 800, the last byte number must be 879.

5.5.4 Options

- This field allows the packet to request special features such as
  - → security level
  - → route to be taken by packet and
  - → timestamp at each router.
- This field can also be used for network testing and debugging.
- As the name implies, options are not required for a datagram.
- The header is made of two parts: 1) Fixed part and 2) Variable part.
  - 1) Maximum size of Fixed part = 20 bytes.
  - 2) Maximum size of Variable part = 40 bytes
- Options are divided into two broad categories: 1) Single-byte options and 2) Multiple-byte options.

1) Single Byte Options

i) No Operation

- This option is used as filler between options.

ii) End of Option

- This option is used for padding at the end of the option field.

2) Multiple Byte Options

i) Record Route

- This option is used to record the routers that handle the datagram.
- This option can list up to 9 router-addresses.

ii) Strict Source Route

- This option is used by the source to pre-determine a route for the datagram.
- Useful purposes: The sender can choose a route with a specific type of service, such as
  - → minimum delay
  - → maximum throughput or
  - → more secure/reliable.
- All the defined-routers must be visited by the datagram.
- If the datagram visits a router that is not on the list, the datagram is discarded.

iii) Loose Source Route

- This option is similar to the strict source route, but it is less rigid.
- Each router in the list must be visited, but the datagram can visit other routers as well.

iv) Timestamp

- This option is used to record the time of datagram processing by a router.
- The time is expressed in milliseconds from midnight GMT (Greenwich Mean Time).
- The recorded-time can help the managers to track the behavior of the routers in the Internet.

5.5.5 Security of IPv4 Datagrams

- Nowadays, the Internet is not secure anymore.
- Three security issues applicable to the IP protocol:
  1) Packet sniffing
  2) Packet modification and
  3) IP spoofing.

1) Packet Sniffing

- Attackers may
  → capture certain packets
  → intercept the packets and
  → make a copy of the packets.
- Packet sniffing is a passive attack.
- Passive attack means the attacker does not modify the contents of the packet.
- The attack is difficult to detect '.' sender & receiver may never know that the packet has been copied.
- Solution:
  Although the attack cannot be stopped, encryption of packet may make the attacker's job difficult.
  The attacker may still sniff the packet, but the content is not detectable (or understandable).

2) Packet Modification

- Attackers may succeed in accessing the content of a packet.
- Then, the attacker can
  → change the address of the packet or
  → change the data of the packet
- Solution:
  The attack can be prevented by data integrity mechanism.
  Data integrity guarantees that the packet is not modified during the transmission.

3) IP Spoofing

- The attacker pretends as a trusted entity and obtains all the secret information.
- For example:
  An attacker sends an IP packet to a bank pretending as legitimate customers.
- Solution:
  The attack can be prevented using an origin-authentication mechanism.

5.5.5.1 IPSec (IP Security)

- IP packets can be protected from the various network-attacks using a protocol called IPSec.
- IPSec protocol & IP protocol can be used to create a connection-oriented service between 2 entities.
- Four services of IPSec:

  1) Defining Algorithms & Keys
  - To create a secure channel b/w two entities, the two entities can agree on some available algorithms and keys.

  2) Packet Encryption
  - To provide privacy, the packets exchanged b/w two parties can be encrypted using the encryption-algorithms and a shared key.
  - This prevents the packet sniffing attack.

  3) Data Integrity
  - Data integrity guarantees that the packet is not modified during the transmission.
  - If the received packet does not pass the data integrity test, the packet is discarded.
  - This prevents the packet modification attack.

  4) Origin Authentication
  - Origin Authentication guarantees that the packet is not created by a pretender.
  - This prevents the IP Spoofing attack.

5.6 ICMP

• ICMP is a network-layer protocol.
• This is used to handle error and other control messages.

5.6.1 MESSAGES

• ICMP messages are divided into 2 broad categories:

1) Error Reporting Messages

➢ These messages report problems that a router or a host may encounter during the processing of datagram.

2) Query Messages

➢ These messages help a host or a network manager get specific information from a router or another host.

➢ For example:

Nodes can discover their neighbors.
Hosts can discover and learn about routers on their network.
Routers can help a node redirect the messages.

• Fields of ICMP messages (Figure 19.8):

1) Type: This field identifies the type of message.

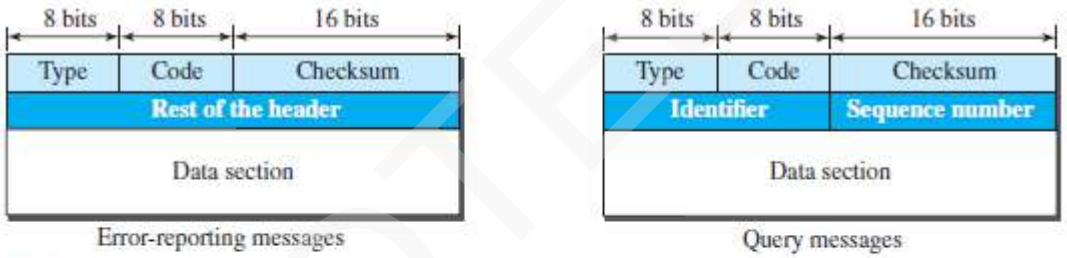
2) Code: This field specifies the reason for the particular message type.

For example,

Type 03 = problem reaching the destinations
Type 11 = problem related to time exceeded.

3) Checksum: This field is used to detect errors in the ICMP message.

4) Data Section: This field can be used for diagnostic purposes by matching the information in the ICMP message with the original data in the IP packet.



Type and code values

Error-reporting messages	Query messages
03: Destination unreachable (codes 0 to 15)	08 and 00: Echo request and reply (only code 0)
04: Source quench (only code 0)	13 and 14: Timestamp request and reply (only code 0)
05: Redirection (codes 0 to 3)	
11: Time exceeded (codes 0 and 1)	
12: Parameter problem (codes 0 and 1)	

Figure 19.8 *General format of ICMP messages*

5.6.1.1 Error Reporting Messages

• Main responsibility of ICMP: To report some errors that may occur during the processing of the datagram (Figure 19.9).
• These messages report problems that a router or a host may encounter during the processing of datagram.
• ICMP does not correct errors; ICMP simply reports the errors to the source.
• Error correction is left to the higher-level protocols (such as TCP or UDP)

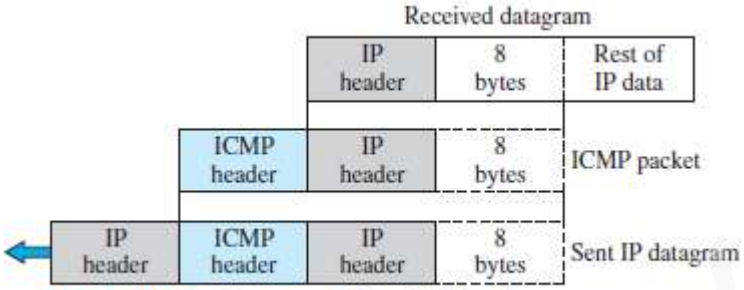


Figure 19.9 *Contents of data field for the error messages*

• Rules for reporting messages:
1) No error-message will be generated for a datagram having a multicast address (or special address).
2) No error-message will be generated in response to a datagram carrying an ICMP error-message.
3) No error-message will be generated for a fragmented datagram that is not the first fragment.

1) Destination Unreachable (Type=3)

• This message is related to problem reaching the destinations.
• This message uses different codes (0 to 15) to define type of error-message.
• Possible values for code field:
Code 0 = network unreachable
Code 1 = host unreachable
Code 2 = protocol unreachable
Code 3 = port unreachable

2) Source Quench (Type=4)

• This message informs the sender that
→ network has encountered congestion and
→ datagram has been dropped.
• The source needs to slow down sending more datagrams.
• In other words, ICMP adds a kind of congestion control mechanism to the IP protocol.

3) Redirection Message (Type=5)

• This is used when the source uses a wrong router to send out its message.
• The router
→ redirects the message to the appropriate router &
→ informs the source to change its default router in the future.
• The IP address of the default router is sent in the message.
• TTL prevents a datagram from being aimlessly circulated in the Internet.
• When TTL becomes 0,
→ the datagram is dropped by the visiting router and
→ a time exceeded message (type 11) is sent to the source.

4) Parameter Problem (Type=12)

• This message can be sent when either
→ there is a problem in the header of a datagram (code 0) or
→ some options are missing or cannot be interpreted (code 1).

5.6.1.2 Query Messages

- These messages help a network manager to get specific information from a router or host.
- Two types of query messages: request (type 8) and reply (type 0).

1) Echo Request & Echo Reply

- These messages are used to determine whether a remote-host is alive.
- A source-host sends an echo request message to destination-host;
  If the destination-host is alive, it responds with an echo reply message.
- Type=8 is used for echo request
  Type=0 is used for echo reply.
- These messages can be used in two debugging tools: ping and traceroute.

2) Timestamp Request & Timestamp Reply

- These messages are used to
  → find the round-trip time between two devices or
  → check whether the clocks in two devices are synchronized.
- The timestamp request sends a number, which defines the time the message is sent.
- The timestamp reply resends another number, which defines the time the message is sent.
- The timestamp reply also includes 2 new numbers representing
  i) the time the request was received and
  ii) the time the response was sent.
- Type=13 is used for timestamp request
  Type=14 is used for timestamp reply.

5.6.2 Debugging Tools

- There are several tools that can be used in the Internet for debugging.
- We can determine the viability of a host or router.
- We can trace the route of a packet.
- Two tools used for debugging: 1) Ping and 2) Traceroute.

5.6.2.1 Ping

- The ping program can be used to find if a host is alive and responding
- Here, ping is used to see how it uses ICMP packets
- The source host sends ICMP echo-request messages;
  The destination, if alive, responds with ICMP echo-reply messages.
- The ping program
  → sets the identifier field in the echo-request and echo-reply message and
  → starts the sequence number from 0; this number is incremented by 1 each time a new message is sent.
- Ping can calculate the round-trip time.
  It inserts the sending time in the data section of the message.
  When the packet arrives, it subtracts the arrival time from the departure time to get the round-trip time (RTT).

5.6.2.2 Traceroute

- The traceroute program can be used to trace the path of a packet from a source to the destination.
- It can find the IP addresses of all the routers that are visited along the path.
- The program is usually set to check for the maximum of 30 hops (routers) to be visited.

Traceroute

- The traceroute program is different from the ping program.
- The ping program gets help from 2 query messages;
  The traceroute program gets help from two error-reporting messages: time-exceeded and destination-unreachable.
- The traceroute is an application layer program, but only the client program is needed.
  In other words, there is no traceroute server program.
- The traceroute application program is encapsulated in a UDP user datagram, but traceroute intentionally uses a port number that is not available at the destination.

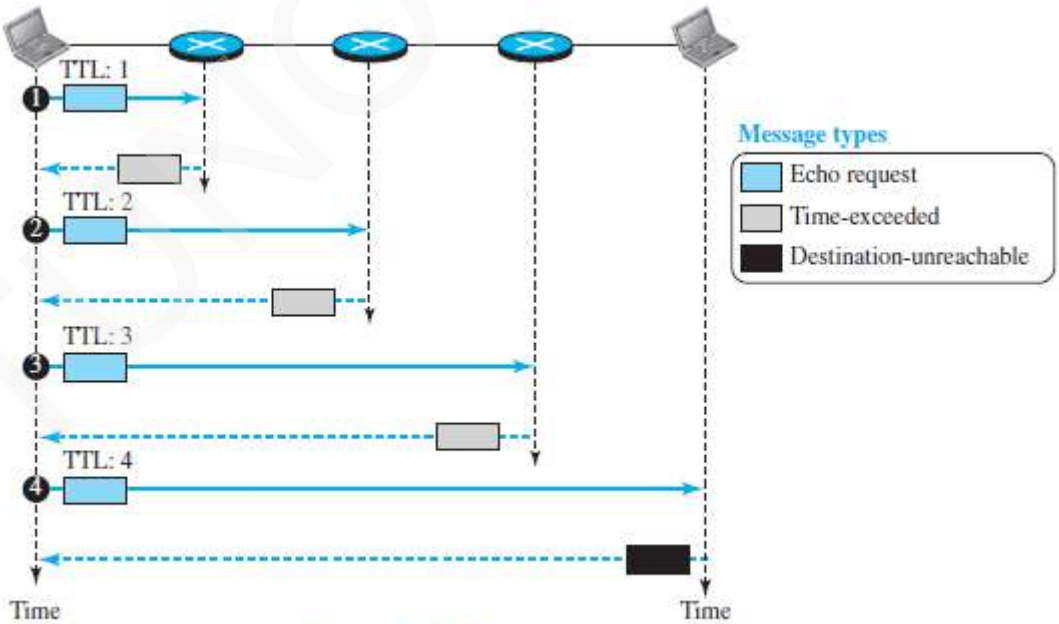


Figure 19.10 *Use of ICMPv4 in traceroute*

5.7 MOBILE IP

- Mobile IP is the extension of IP protocol.
- Mobile IP allows mobile computers to be connected to the Internet.

5.7.1 Addressing

- In Mobile IP, the main problem that must be solved is addressing.

5.7.1.1 Stationary Hosts

- The original IP addressing assumed that a host is stationary.
- A router uses an IP address to route an IP datagram.
- An IP address has two parts: a prefix and a suffix.
- The prefix associates a host with a network.

  For example, the IP address 10.3.4.24/8 defines a host attached to the network 10.0.0.0/8.
- The address is valid only when the host is attached to the network.
- If the network changes, the address is no longer valid.

5.7.1.2 Mobile Hosts

- When a host moves from one network to another, the IP addressing structure needs to be modified.
- The host has two addresses (Figure 19.12):
  1) Home address &
  2) Care-of address

1) Home Address

- Original address of host called the home address.
- The home address is permanent.
- The home address associates the host with its home network.
- Home network is a network that is the permanent home of the host.

2) Care-of-Address

- The care-of address is temporary.
- The care-of address changes as the mobile-host moves from one network to another.
- Care-of address is associated with the foreign network.
- Foreign network is a network to which the host moves.
- When a mobile-host visits a foreign network, it receives its care-of address during the agent discovery and registration phase.

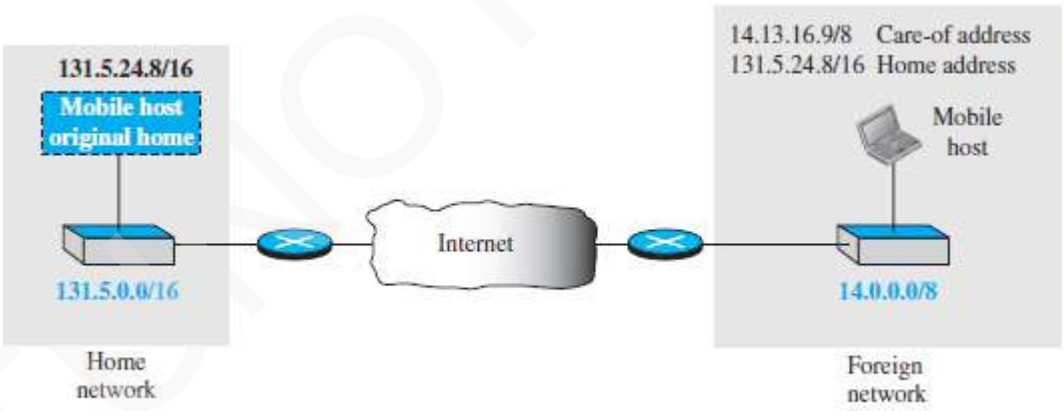


Figure 19.12 *Home address and care-of address*

5.7.2 Agents

• Two agents are required to make change of address transparent to rest of the Internet (Fig 19.13):
1) Home-agent and
2) Foreign-agent.

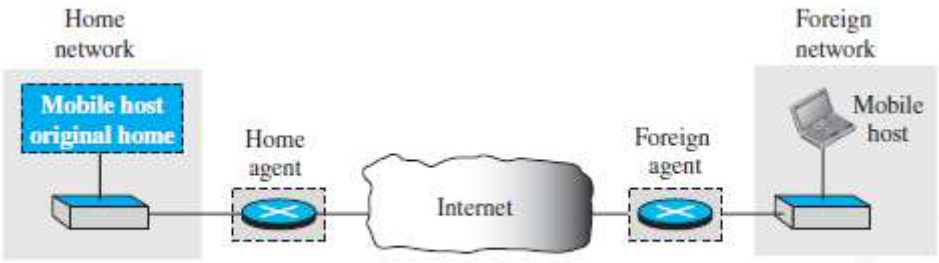


Figure 19.13 *Home agent and foreign agent*

1) Home Agent

• The home-agent is a router attached to the home network.
• The home-agent acts on behalf of mobile-host when a remote-host sends a packet to mobile-host.
• The home-agent receives and delivers packets sent by the remote-host to the foreign-agent.

2) Foreign Agent

• The foreign-agent is a router attached to the foreign network.
• The foreign-agent receives and delivers packets sent by the home-agent to the mobile-host.
• The mobile-host can also act as a foreign-agent i.e. mobile-host and foreign-agent can be the same.
• However, to do this, a mobile-host must be able to receive a care-of address by itself.
• In addition, the mobile-host needs the necessary software to allow it to communicate with the home-agent and to have two addresses: i) its home address and ii) its care-of address.
• This dual addressing must be transparent to the application programs.

Collocated Care-of-Address

➢ When the mobile-host and the foreign-agent are the same, the care-of-address is called a collocated care-of-address.
➢ Advantage:
1) mobile-host can move to any network w/o worrying about availability of a foreign-agent.
➢ Disadvantage:
1) The mobile-host needs extra software to act as its own foreign-agent.

5.7.3 Three Phases

• To communicate with a remote-host, a mobile-host goes through 3 phases (Figure 19.14):

1) Agent Discovery: involves the mobile-host, the foreign-agent, and the home-agent.

2) Registration: involves the mobile-host, the foreign-agent, and the home-agent.

3) Data Transfer: Here, the remote-host is also involved.

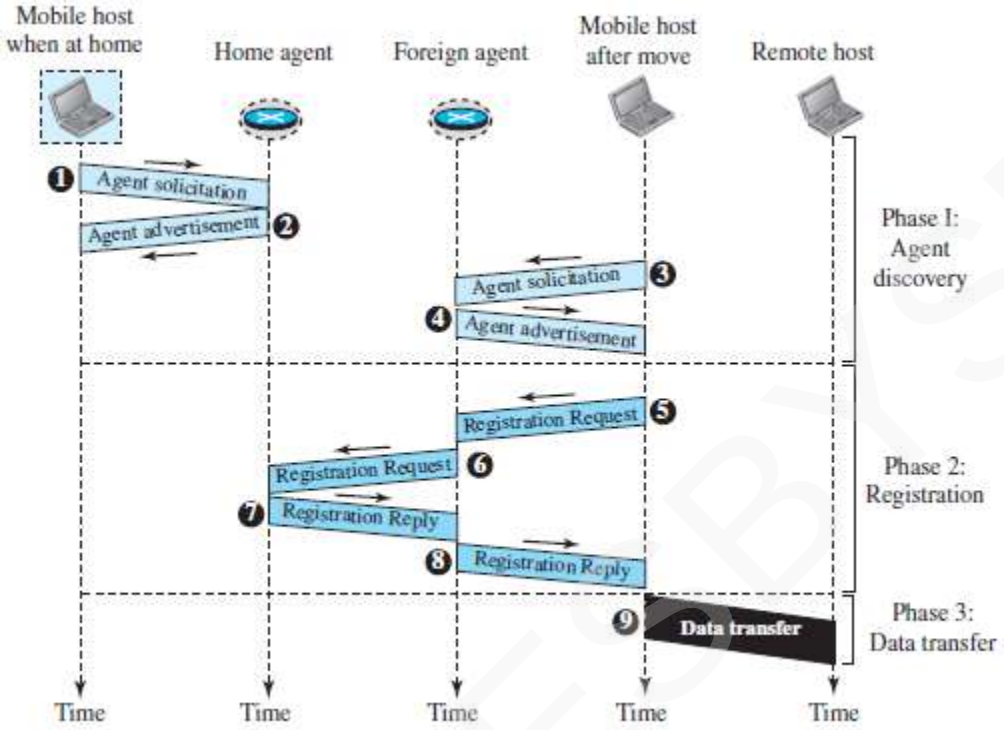


Figure 19.14 *Remote host and mobile host communication*

5.7.3.1 Agent Discovery

• Agent discovery consists of two subphases:

1) A mobile-host must discover (learn the address of) a home-agent before it leaves its home network.

2) A mobile-host must also discover a foreign-agent after it has moved to a foreign network.

• This discovery consists of learning the care-of address as well as the foreign-agent's address.

• Two types of messages are used: i) advertisement and ii) solicitation.

1) Agent Advertisement

• When a router advertises its presence on a network using an ICMP router advertisement, it can append an agent advertisement to the packet if it acts as an agent.

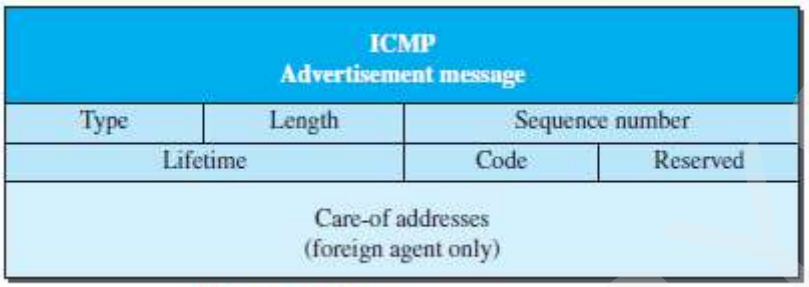


Figure 19.15 *Agent advertisement*

• Various fields are (Figure 19.15):

1) Type

➢ This field is set to 16.

2) Length

➢ This field defines the total length of the extension message.

3) Sequence Number

➢ This field holds the message number.

➢ The recipient can use the sequence number to determine if a message is lost.

4) Lifetime

➢ This field defines the number of seconds that the agent will accept requests.

➢ If the value is a string of 1s, the lifetime is infinite.

5) Code

➢ This field is a flag in which each bit is set (1) or unset (0) (Table 19.1).

Table 19.1 *Code Bits*

Bit	*Meaning*
0	Registration required. No collocated care-of address.
1	Agent is busy and does not accept registration at this moment.
2	Agent acts as a home agent.
3	Agent acts as a foreign agent.
4	Agent uses minimal encapsulation.
5	Agent uses generic routing encapsulation (GRE).
6	Agent supports header compression.
7	Unused (0).

6) Care-of Addresses

➢ This field contains a list of addresses available for use as care-of addresses.

➢ The mobile-host can choose one of these addresses.

➢ The selection of this care-of address is announced in the registration request.

2) Agent Solicitation

• When a mobile-host has moved to a new network and has not received agent advertisements, it can initiate an agent solicitation.

• It can use the ICMP solicitation message to inform an agent that it needs assistance

5.7.3.2 Registration

• After a mobile-host has moved to a foreign network and discovered the foreign-agent, it must register.

• Four aspects of registration:

1) The mobile-host must register itself with the foreign-agent.
2) The mobile-host must register itself with its home-agent. This is normally done by the foreign-agent on behalf of the mobile-host.
3) The mobile-host must renew registration if it has expired.
4) The mobile-host must cancel its registration (deregistration) when it returns home.

5.7.3.2.1 Request & Reply

• To register with the foreign-agent and the home-agent, the mobile-host uses a registration request and a registration reply.

1) Registration Request

• A registration request is sent from the mobile-host to the foreign-agent
→ to register its care-of address and
→ to announce its home address and home-agent address.

• Foreign-agent, after receiving and registering the request, relays the message to the home-agent.

• The home-agent now knows the address of the foreign-agent because the IP packet that is used for relaying has the IP address of the foreign-agent as the source address.

Figure 19.16 *Registration request format*

• Various fields are (Figure 19.16):

1) Type
- This field defines the type of message.
- For a request message the value of this field is 1.

2) Flag
- This field defines forwarding information.
- The value of each bit can be set or unset (Table 19.2).

Table 19.2 *Registration request flag field bits*

Bit	*Meaning*
0	Mobile host requests that home agent retain its prior care-of address.
1	Mobile host requests that home agent tunnel any broadcast message.
2	Mobile host is using collocated care-of address.
3	Mobile host requests that home agent use minimal encapsulation.
4	Mobile host requests generic routing encapsulation (GRE).
5	Mobile host requests header compression.
6–7	Reserved bits.

3) Lifetime
- This field defines the number of seconds the registration is valid.
  i) If the field is a string of 0s, the request message is asking for deregistration.
  ii) This field If the field is a string of 1s, the lifetime is infinite.

4) Home Address
- This field contains the permanent (first) address of the mobile-host.

5) Home Agent Address
- This field contains the address of the home-agent.

6) Care-of-Address
- This field is the temporary (second) address of the mobile-host.

7) Identification
- This field contains a 64-bit number that is inserted into the request by the mobile-host.
- This field matches a request with a reply.

8) Extensions
- This field is used for authentication.
- This field allows a home-agent to authenticate the mobile agent.

2) Registration Reply
- A registration reply is sent from home-agent to foreign-agent and then relayed to the mobile-host.
- The reply confirms or denies the registration request. (Figure 19.17)
- The fields are similar to registration request with the 3 exceptions:

1) The value of the type field is 3.
2) The code field replaces the flag field and shows the result of the registration request (acceptance or denial).
3) The care-of address field is not needed.

Figure 19.17 *Registration reply format*

5.7.3.3 Data Transfer

• After agent discovery & registration, a mobile-host can communicate with a remote-host (Fig 19.17).

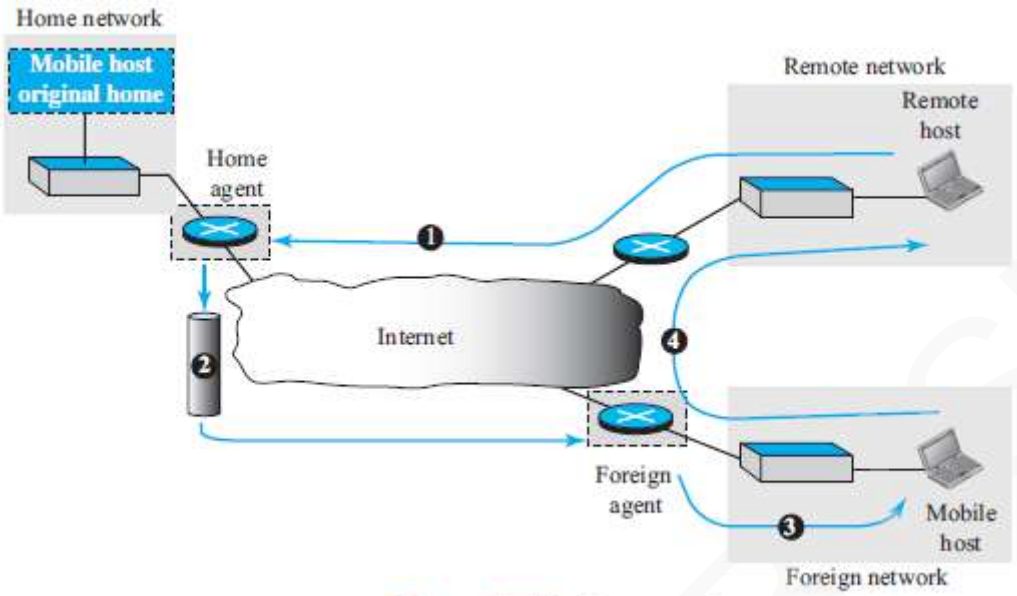


Figure 19.18 *Data transfer*

• Here we have 4 cases (Figure 19.18):

1) From Remote Host to Home Agent

• When a remote-host wants to send a packet to the mobile-host, the remote-host uses
→ address of itself as the source address and
→ home address of the mobile-host as the destination address.
• In other words, the remote-host sends a packet as though the mobile-host is at its home network.
• The packet is intercepted by the home-agent, which pretends it is the mobile-host.
• This is done using the proxy ARP technique (Path 1 of Figure 19.18).

2) From Home Agent to Foreign Agent

• After receiving the packet, the home-agent sends the packet to the foreign-agent, using the tunneling concept.
• The home-agent encapsulates the whole IP packet inside another IP packet using its address as the source and the foreign-agent's address as the destination. (Path 2 of Figure 19.18).

3) From Foreign Agent to Mobile Host

• When the foreign-agent receives the packet, it removes the original packet.
• However, since the destination address is the home address of the mobile-host, the foreign-agent consults a registry table to find the care-of address of the mobile-host. (Otherwise, the would just be sent back to the home network.)
• The packet is then sent to the care-of address (Path 3 of Figure 19.18).

4) From Mobile Host to Remote Host

• When a mobile-host wants to send a packet to a remote-host (for example, a response to the packet it has received), it sends as it does normally.
• The mobile-host prepares a packet with its home address as the source, and the address of the remote-host as the destination.
• Although the packet comes from the foreign network, it has the home address of the mobile-host (Path 4 of Figure 19.18).

5.7.4 Inefficiency in Mobile IP

• Communication involving mobile IP can be inefficient.
• The inefficiency can be severe or moderate.
1) The severe case is called double crossing or 2X.
2) The moderate case is called triangle routing or dog-leg routing.

5.7.4.1 Double Crossing

• Double crossing occurs when a remote-host communicates with a mobile-host that has moved to the same network (or site) as the remote-host (Figure 19.19).
• When the mobile-host sends a packet to the remote-host, there is no inefficiency; the communication is local.
• However, when remote-host sends a packet to mobile-host, the packet crosses the Internet twice.
• Since a computer usually communicates with other local computers (principle of locality), the inefficiency from double crossing is significant.

5.7.4.2 Triangle Routing

• Triangle routing occurs when the remote-host communicates with a mobile-host that is not attached to the same network (or site) as the mobile-host.
• When the mobile-host sends a packet to the remote-host, there is no inefficiency.

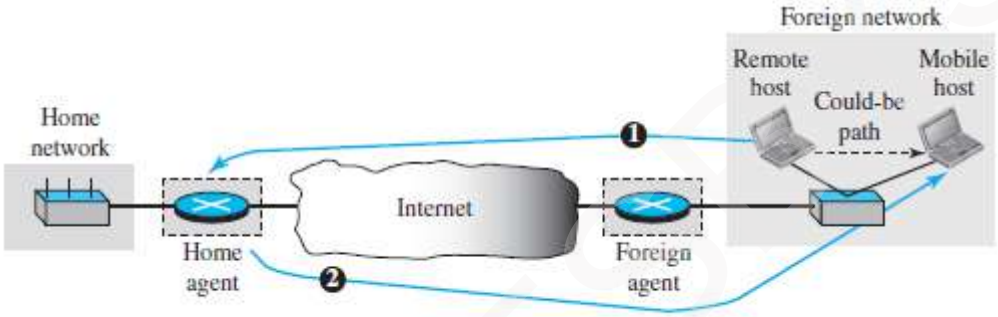


Figure 19.19 *Double crossing*

• However, when the remote-host sends a packet to the mobile-host, the packet goes from the remote-host to the home-agent and then to the mobile-host.
• The packet travels the two sides of a triangle, instead of just one side (Figure 19.20).

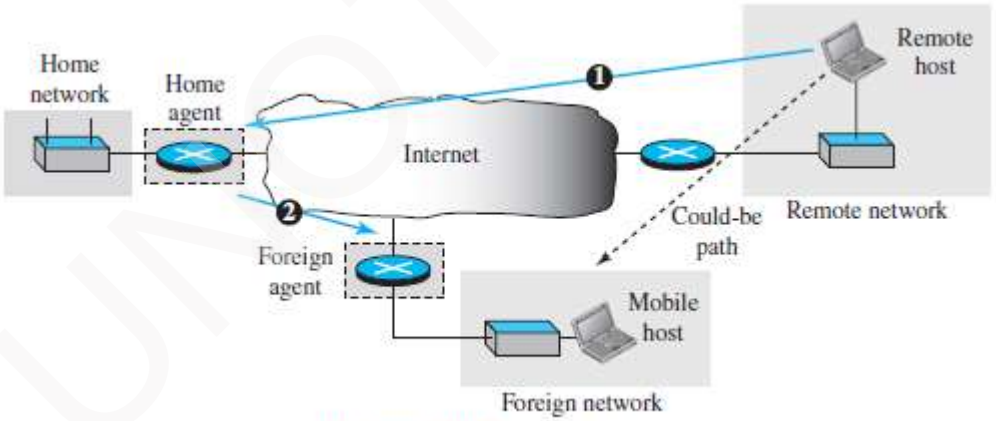


Figure 19.20 *Triangle routing*

Solution

• One solution to inefficiency is for the remote-host to bind the care-of address to the home address of a mobile-host.
• For example, when a home-agent receives the first packet for a mobile-host, it forwards the packet to the foreign-agent; it could also send an update binding packet to the remote-host so that future packets to this host could be sent to the care-of address.
• The remote-host can keep this information in a cache.
• The problem with this strategy is that the cache entry becomes outdated once the mobile-host moves.
• In this case, the home-agent needs to send a warning packet to the remote-host to inform it of the change.

MODULE 5(CONT.): NEXT GENERATION IP

5.8 IPv6 ADDRESSING

- The main reason for migration from IPv4 to IPv6 is the small size of the address-space in IPv4.
- Size of IPv6 address =128 bits (four times the address length in IPv4, which is 32 bits).

5.8.1 Representation

- Two notations can be used to represent IPv6 addresses: 1) binary and 2) colon hexadecimal.

Binary (128 bits)	1111111011110110 ... 1111111100000000
Colon Hexadecimal	FEF6:BA98:7654:3210:ADEF:BBFF:2922:FF00

5.8.2 Address Space

- The address-space of IPv6 contains 2^{128} addresses.

5.8.2.1 Three Address Types

- Three types of destination address: 1) Unicast 2) Anycast and 3) Multicast.

1) Unicast Address

- A unicast address defines a single interface (computer or router).
- The packet with a unicast address will be delivered to the intended recipient.

2) Anycast Address

- An anycast address defines a group of computers that all share a single address.
- A packet with an anycast address is delivered to only one member of the group.
- The member is the one who is first reachable.

3) Multicast Address

- A multicast address also defines a group of computers.
- Difference between anycasting and multicasting.
  - i) In anycasting, only one copy of the packet is sent to one of the members of the group.
  - ii) in multicasting each member of the group receives a copy.

5.8.3 Address Space Allocation

• The address-space is divided into several blocks of varying size.
• Each block is allocated for a special purpose.

Table 22.1 *Prefixes for assigned IPv6 addresses*

Block prefix	*CIDR*	*Block assignment*	*Fraction*
0000 0000	0000::/8	Special addresses	1/256
001	2000::/3	Global unicast	1/8
1111 110	FC00::/7	Unique local unicast	1/128
1111 1110 10	FE80::/10	Link local addresses	1/1024
1111 1111	FF00::/8	Multicast addresses	1/256

5.8.3.1 Global Unicast Addresses

• The block in the address-space used for unicast communication b/w 2 hosts in the Internet is called global unicast address block.
• CIDR for the block is 2000::/3. This means that the three leftmost bits are the same for all addresses in this block (001).
• The size of this block is 2^{125} bits, which is more than enough for Internet expansion for many years to come.
• An address in the block is divided into 3 parts (Figure 22.1):
    1) Global routing prefix (n bits)
    2) Subnet identifier (m bits) and
    3) Interface identifier (q bits).

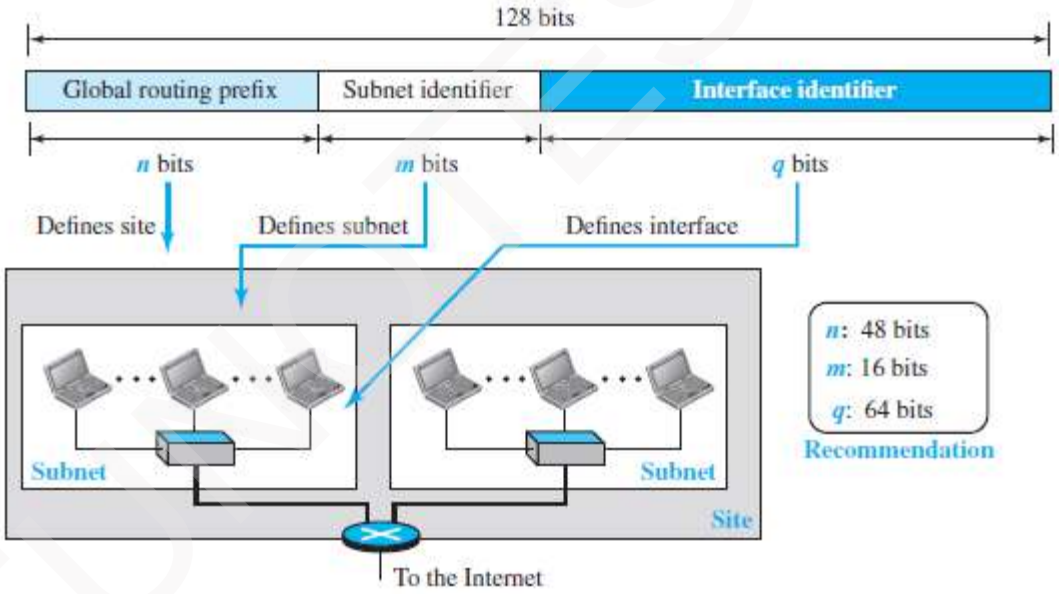


Figure 22.1 *Global unicast address*

• The global routing prefix is used to route the packet through the Internet to the organization site, such as the ISP that owns the block.
• Since the first 3 bits in this part are fixed (001), the rest of the 45 bits can be defined for up to 2^{45} sites (a private organization or an ISP).
    1) The global routers in Internet route a packet to its destination site based on the value of n.
    2) The next m bits define a subnet in an organization.
    3) The last q bits define the interface identifier.
• Two link layer addressing schemes:
    1) 64-bit extended unique identifier (EUI-64) defined by IEEE and
    2) 48-bit link-layer address defined by Ethernet.

1) Mapping EUI-64

• To map a 64-bit physical address, the global/local bit of this format needs to be changed from 0 to 1 (local to global) to define an interface address (Figure 22.2).

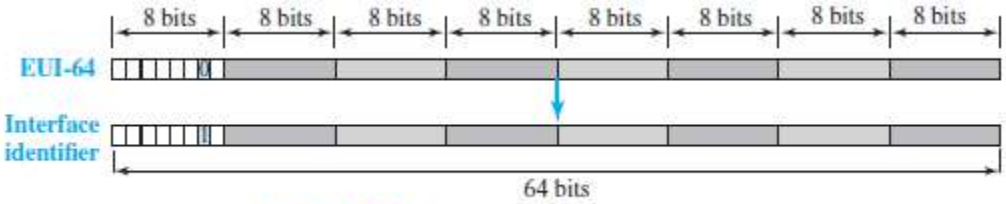


Figure 22.2 *Mapping for EUI-64*

2) Mapping Ethernet MAC Address

• Mapping a 48-bit Ethernet address into a 64-bit interface identifier is more involved.
• We need to change the local/global bit to 1 and insert an additional 16 bits.
• The additional 16 bits are defined as 15 ones followed by one zero, or $FFFE_{16}$ (Figure 22.3).

5.8.3.2 Special Addresses

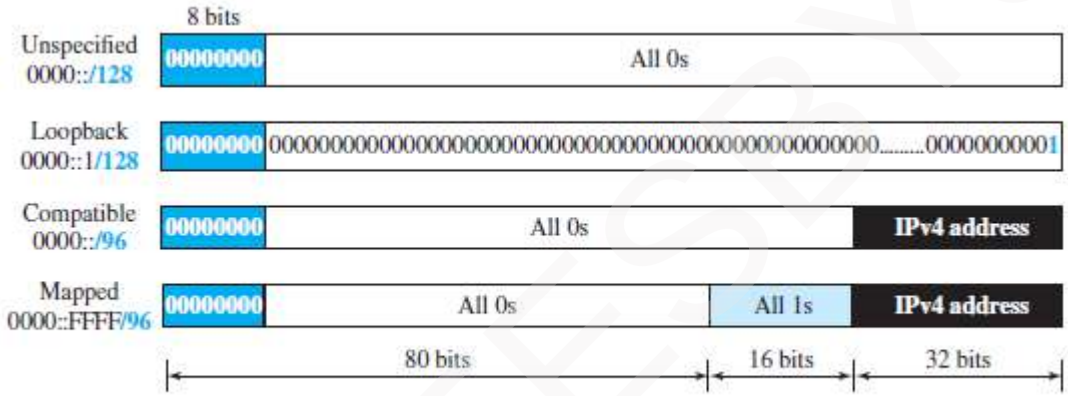


Figure 22.4 *Special addresses*

• Following are different special addresses (Figure 22.4):

1) Unspecified Address

• The unspecified address is a subblock containing only one address.
• This address is used during bootstrap when a host does not know its own address and wants to send an inquiry to find it.

2) Loopback Address

• The loopback address also consists of one address.

3) Transition Address

• During the transition from IPv4 to IPv6, hosts can use their IPv4 addresses embedded in IPv6 addresses.
• Two formats have been designed for this purpose: compatible and mapped.

1) Compatible Address

➢ A compatible address is an address of 96 bits of zero followed by 32 bits of IPv4 address.
➢ It is used when a computer using IPv6 wants to send a message to another computer using IPv6.

2) Mapped Address

➢ A mapped address is used when a computer already migrated to version 6 wants to send an address to a computer still using version 4.

5.8.3.3 Other Assigned Blocks

• IPv6 uses 2 large blocks for private addressing and one large block for multicasting (Figure 22.5).

1) Unique Local Unicast Block

• A subblock in a unique local unicast block can be privately created and used by a site.
• The packet carrying this type of address as the destination address is not expected to be routed.
• This type of address has the identifier 1111 110.
• The next bit can be 0 or 1 to define how the address is selected (locally or by an authority).

2) Link Local Block

• A subblock in link local block can be used as a private address in a network.
• This type of address has the block identifier 1111111010.
• The next 54 bits are set to zero.
• The last 64 bits can be changed to define the interface for each computer.

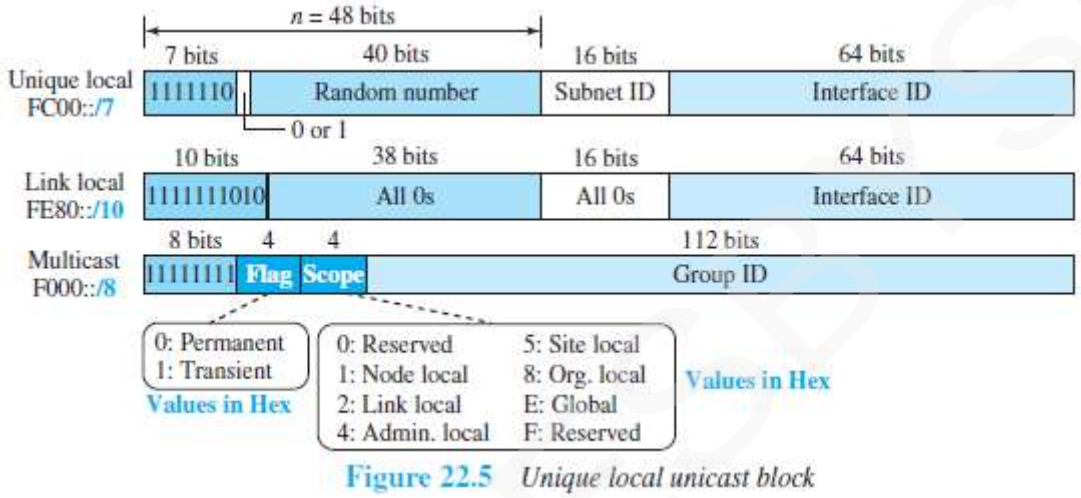


Figure 22.5 *Unique local unicast block*

5.8.4 Autoconfiguration

• When a host in IPv6 joins a network, it can configure itself using the following process:

1) The host first creates a link local address for itself.
- ➢ This is done by
  - → taking the 10-bit link local prefix (1111 1110 10)
  - → adding 54 zeros and
  - → adding the 64-bit interface identifier.
- ➢ The result is a 128-bit link local address.

2) The host then tests to see if this link local address is unique and not used by other hosts.
- ➢ Since the 64-bit interface identifier is supposed to be unique, the link local address generated is unique with a high probability.
- ➢ To check uniqueness, the host
  - → sends a neighbor solicitation message and
  - → waits for a neighbor advertisement message.
- ➢ If any host in the subnet is using this link local address, the process fails and the host cannot auto-configure itself.

3) If the uniqueness of the link local address is passed, the host stores this address as its link local address (for private communication), but it still needs a global unicast address.
- ➢ The host then sends a router solicitation message to a local router.
- ➢ If there is a router running on the network, the host receives a router advertisement message that includes
  - → global unicast prefix and
  - → subnet prefix that the host needs to add to its interface identifier to generate its global unicast address.
- ➢ If the router cannot help the host with the configuration, it informs the host in the router advertisement message (by setting a flag).

5.9 THE IPv6 PROTOCOL

5.9.1 Changes from IPv4 to IPv6 (Advantages of IPv6)

1) Header Format

- IPv6 uses a new header format.
- Options are
  - → separated from the base-header and
  - → inserted between the base-header and the data.
- This speeds up the routing process (because most of the options do not need to be checked by routers).

2) New Options

- IPv6 has new options to allow for additional functionalities.

3) Extension

- IPv6 is designed to allow the extension of the protocol if required by new technologies or applications.

4) Resource Allocation

- In IPv6,
  - → type-of-service (TOS) field has been removed
  - → two new fields: 1) traffic class and 2) flow label, are added to enable the source to request special handling of the packet.
- This mechanism can be used to support real-time audio and video.

5) Security

- The encryption option provides confidentiality of the packet.
- The authentication option provides integrity of the packet.

5.9.2 Packet Format

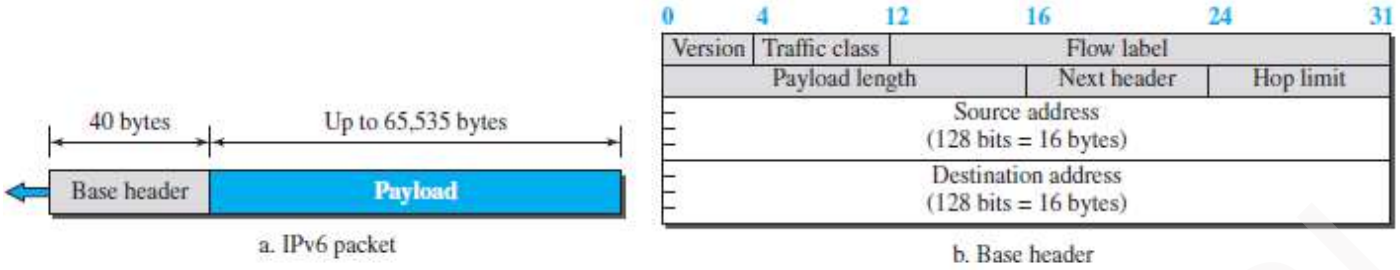


Figure 22.6 *IPv6 datagram*

• IP header contains following fields (Figure 22.6):

1) Version

➢ This specifies version number of protocol. For IPv6, version=6.

2) Traffic Class

➢ This field is used to distinguish different payloads with different delivery requirements.
(Traffic class replaces the type-of-service field in IPv4).

3) Flow Label

➢ This field is designed to provide special handling for a particular flow of data.

4) Payload Length

➢ This indicates length of data (excluding header). Maximum length=65535 bytes.

➢ The length of the base-header is fixed (40 bytes); only the length of the payload needs to be defined.

5) Next Header

➢ This identifies type of extension header that follows the basic header.

6) Hop Limit

➢ This specifies number of hops the packet can travel before being dropped by a router.
(Hop limit serves the same purpose as the TTL field in IPv4).

7) Source and Destination Addresses

➢ These identify source host and destination host respectively.

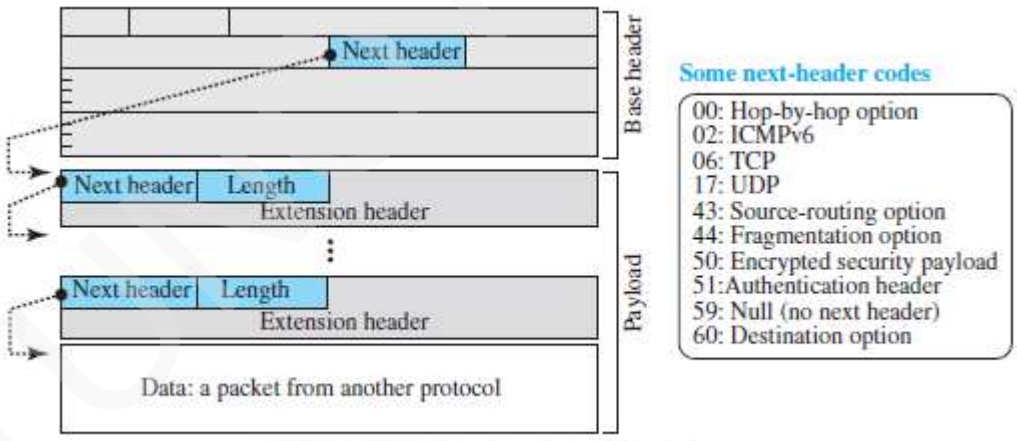


Figure 22.7 *Payload in an IPv6 datagram*

8) Payload

➢ The payload contains zero or more extension headers (options) followed by the data from other protocols (UDP, TCP, and so on).

➢ The payload can have many extension headers as required by the situation.

➢ Each extension header has 2 mandatory fields (Figure 22.7):
1) Next header and
2) Length

• Two mandatory fields are followed by information related to the particular option.

5.9.2.1 Concept of Flow & Priority in IPv6

- To a router, a flow is a sequence of packets that share the same characteristics such as
  - → traveling the same path
  - → using the same resources or
  - → having the same kind of security
- A router that supports the handling of flow labels has a flow label table.
- The table has an entry for each active flow label.
  Each entry defines the services required by the corresponding flow label.
- When a router receives a packet, the router consults its flow label table.
- Then, the router provides the packet with the services mentioned in the entry.
- A flow label can be used to support the transmission of real-time audio/video.
- Real-time audio/video requires resources such as
  - → high bandwidth
  - → large buffers or
  - → long processing time
- Resource reservation guarantees that real-time data will not be delayed due to a lack of resources.

5.9.2.2 Fragmentation & Reassembly

- Fragmentation of the packet is done only by the source, but not by the routers.
  The reassembling is done by the destination.
- At routers, the fragmentation is not allowed to speed up the processing in the router.
- Normally, the fragmentation of a packet in a router needs a lot of processing. This is because
  1) The packets need to be fragmented.
  2) All fields related to the fragmentation need to be recalculated.
- The source will
  - → check the size of the packet and
  - → make the decision to fragment the packet or not.
- If packet-size is greater than the MTU of the network, the router will drop the packet.
- Then, the router sends an error message to inform the source.

5.9.3 Extension Header

- An IP packet is made of
  - → base-header &
  - → some extension headers.
- Length of base header = 40 bytes.
- To support extra functionalities, extension headers can be placed b/w base header and payload.
- Extension headers act like options in IPv4.
- Six types of extension headers (Figure 22.8):
  1) Hop-by-hop option 2) Source routing 3) Fragmentation
  4) Authentication 5) Encrypted security payload 5) Destination option.

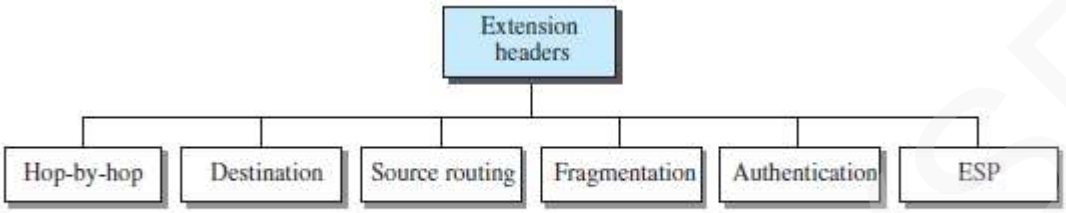


Figure 22.8 *Extension header types*

1) Hop-by-Hop Option

- This option is used when the source needs to pass information to all routers visited by the datagram.
- Three options are defined: i) Pad1, ii) PadN, and iii) Jumbo payload.

i) Pad1

- This option is designed for alignment purposes.
- Some options need to start at a specific bit of the 32-bit word.
- Pad1 is added, if one byte is needed for alignment.

ii) PadN

- PadN is similar in concept to Pad1.
- The difference is that PadN is used when 2 or more bytes are needed for alignment.

iii) Jumbo Payload

- This option is used when larger packet has to be sent. (> 65,535 bytes)
- Large packets are referred to as jumbo packets.
- Maximum length of payload = 65,535 bytes.

2) Destination Option

- This option is used when the source needs to pass information to the destination only.
- Intermediate routers are not allowed to access this information.
- Two options are defined: i) Pad1 & ii) PadN

3) Source Routing

- This option combines the concepts of
  - → strict source routing and
  - → loose source routing.

4) Fragmentation

- In IPv6, only the original source can fragment.
- A source must use a "Path MTU Discovery technique" to find the smallest MTU along the path from the source to the destination.
- Minimum size of MTU = 1280 bytes. This value is required for each network connected to the Internet.
- If a source does not use a Path MTU Discovery technique, the source fragments the datagram to a size of 1280 bytes.

5) Authentication

- This option has a dual purpose:
  - i) Validates the message sender: This is needed so the receiver can be sure that a message is from the genuine sender and not from an attacker.
  - ii) Ensures the integrity of data: This is needed to check that the data is not altered in transition by some attacker.

6) Encrypted Security Payload (ESP)

- This option provides confidentiality and guards against attacker.

5.9.3.1 Comparison of Options between IPv4 and IPv6

1) The no-operation and end-of-option options in IPv4 are replaced by Pad1 and PadN options in IPv6.
2) The record route option is not implemented in IPv6 because it was not used.
3) The timestamp option is not implemented because it was not used.
4) The source route option is called the source route extension header in IPv6.
5) The fragmentation fields in the base-header section of IPv4 have moved to the fragmentation extension header in IPv6.
6) The authentication extension header is new in IPv6.
7) The encrypted security payload extension header is new in IPv6.

5.10 THE ICMPv6 PROTOCOL

• ICMP, ARP & IGMP protocols in IPv4 are combined into one single protocol called ICMPv6 (Fig 22.9).

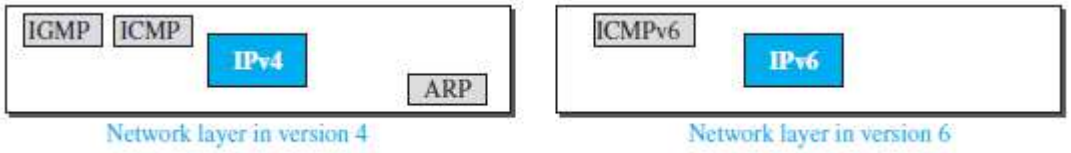


Figure 22.9 *Comparison of network layer in version 4 and version 6*

• Four groups of messages (Figure 22.10):
1) Error-reporting messages
2) Informational messages
3) Neighbor-discovery messages and
4) Group-membership messages.

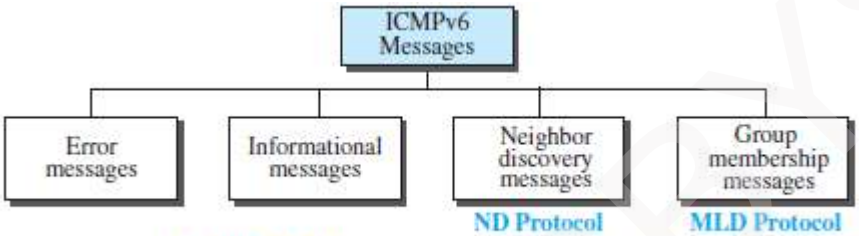


Figure 22.10 *Categories of ICMPv6 messages*

5.10.1 Error-Reporting Messages

• Main responsibility of ICMP: Report errors.
• ICMP forms an error packet, which is then encapsulated in the datagram.
• The encapsulated datagram is delivered to the original source.
• Four types of errors:
1) Destination unreachable
2) Packet too big
3) Time exceeded and
4) Parameter problems.

1) Destination Unreachable Message

➢ Here, a router cannot forward a datagram or a host cannot deliver the datagram to the upper layer protocol.
➢ So, the router/host
→ discards the datagram and
→ sends a destination-unreachable message to the source.

2) Packet Too Big Message

➢ Fragmentation of the packet is done only by the source, but not by the routers.
➢ If a router receives a datagram larger than MTU size of the network, the router
→ discards the datagram and
→ sends a packet-too-big message to the source.

3) Time Exceeded Message

➢ A time-exceeded error message is generated in 2 cases:
i) When the TTL value becomes zero and
ii) When not all fragments of a datagram have arrived in the time-limit.

4) Parameter Problem Message

➢ Any missing value in the datagram-header can create serious problems.
➢ If a router discovers any missing value in any field, the router
→ discards the datagram and
→ sends a parameter-problem message to the source.

5.10.2 Informational Messages

- Two types of messages: i) echo request and ii) echo reply.
- These 2 messages are used to check whether 2 devices can communicate with each other.
- A source-host can send an echo-request message to another host.
  The destination-host can respond with the echo-reply message to the source-host.

5.10.3 Neighbor Discovery Messages

- Two new protocols are used:
  1) Neighbor-Discovery (ND) protocol and
  2) Inverse-Neighbor-Discovery (IND) protocol.
- These 2 protocols are used by nodes on the same link for 3 main purposes:
  1) Hosts use the ND protocol to find routers in the neighborhood that will forward packets for them.
  2) Nodes use the ND protocol to find the link-layer addresses of neighbors.
  3) Nodes use the IND protocol to find the IPv6 addresses of neighbors.
- Seven types of errors:

1) Router Solicitation Message

- A host/router uses router-solicitation message to find a router in n/w that can forward a datagram.
- Physical address of the host/router is included to make the response easier for the router.

2) Router Advertisement Message

- A host/router sends the router-advertisement message in response to a router solicitation message.

3) Neighbor Solicitation Message

- The neighbor solicitation message has the same duty as the ARP request message.
- A host uses the neighbor solicitation message when the host has a message to send to a neighbor.
- The sender knows the IP address of the receiver, but needs the physical address of the receiver.
- The physical address is needed for the datagram to be encapsulated in a frame.

4) Neighbor Advertisement Message

- A host sends the neighbor-advertisement message in response to a neighbor solicitation message.

5) Redirection Message

- The purpose of the redirection message is the same as for version 4.
- However, the format of the packet now accommodates the size of the IP address in version 6.
- Also, an option is added to let the host know the physical address of the target router.

6) Inverse Neighbor Solicitation Message

- A host uses inverse-neighbor-solicitation message to know the physical address of a neighbor, but not the neighbor's IP address.
- The message is encapsulated in a datagram using a multicast address.
- The node must send the following 2 information in the option field:
  i) Physical address of the sender and
  ii) Physical address of the target node.
- The sender can also include its IP address and the MTU value for the link.

7) Inverse Neighbor Advertisement Message

- A host sends the inverse-neighbor-advertisement message in response to a inverse-neighbor-discovery message.

5.10.4 Group Membership Messages

• The management of multicast delivery handling in IPv4 is given to the IGMPv3 protocol.
• In IPv6, this responsibility is given to the Multicast Listener Delivery protocol.
• MLDv2 has 2 types of messages:
    1) Membership-query message and
    2) Membership-report message.
• The first type can be divided into 3 subtypes: i) General, ii) Group-specific, and iii) Group-and-source specific.

1) Membership Query Message

• A router sends a membership-query message to find active group-members in the network.
• The format of the membership-query in MLDv2 is exactly the same as the one in IGMPv3 three exceptions:
    i) Size of the multicast address & source address has been changed from 32 bits to 128 bits.
    ii) The field size is in the maximum response code field, in which the size has been changed from 8 bits to 16 bits.
    iii) The format of the first 8 bytes matches the format for other ICMPv6 packets because MLDv2 is considered to be part of ICMPv6.

2) Membership Report Message

• The format of the membership-report in MLDv2 is exactly the same as the one in IGMPv3 one exception:
    i) Size of the multicast address & source address has been changed from 32 bits to 128 bits.
• In particular, the record type is the same as the one defined for IGMPv3 (types 1 to 6).

5.11 TRANSITION FROM IPv4 TO IPv6

5.11.1 Strategies

• Three strategies have been devised for transition:
1) Dual stack
2) Tunneling and
3) Header translation.

1) Dual Stack

• Recommended: All hosts must run IPv4 and IPv6 (dual stack) simultaneously until all the Internet uses IPv6 (Figure 22.11).

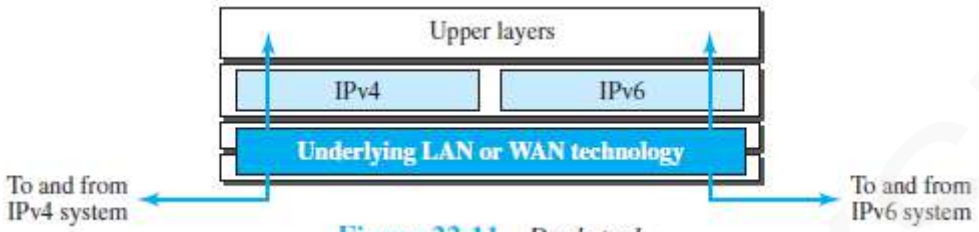


Figure 22.11 *Dual stack*

• To determine which version to use, the source queries the DNS.
i) If the DNS returns an IPv4 address, the source sends an IPv4 packet.
ii) If the DNS returns an IPv6 address, the source sends an IPv6 packet.

2) Tunneling

• Tunneling is a strategy used when
→ two computers using IPv6 want to communicate with each other and
→ the packet must pass through an IPv4 network.
• To pass through IPv4 network, the packet must have an IPv4 address (Figure 22.12).
• So,
i) IPv6 packet is encapsulated in an IPv4 packet when the packet enters the IPv4 network.
ii) IPv6 packet is decapsulated from an IPv4 packet when the packet exits the IPv4 network.

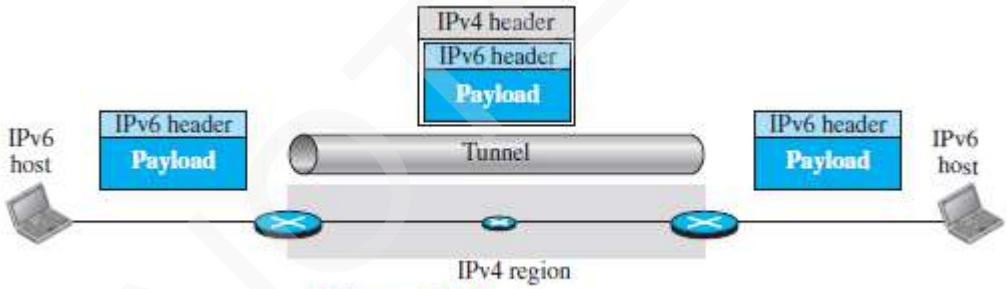


Figure 22.12 *Tunneling strategy*

3) Header Translation

• Header translation is necessary when the majority of the Internet has moved to IPv6 but some systems still use IPv4 (Figure 22.13).
• The sender wants to use IPv6, but the receiver does not understand IPv6.
• Tunneling does not work in this situation because
→ the packet must be in the IPv4 format to be understood by the receiver.
• In this case, the header format must be totally changed through header translation.
• The header of the IPv6 packet is converted to an IPv4 header/

Figure 22.13 *Header translation strategy*

MODULE-WISE QUESTIONS

MODULE 5: OTHER WIRELESS NETWORKS

1. Explain two types of services of WiMAX. (2)
2. Explain layers in Project 802.16. (6)
3. Explain WiMAX MAC frame format. (6*)
4. Discuss the operation of the cellular telephony. (6*)
5. Explain first generation 1G of cellular telephony. (6)
6. Explain second generation 2G of cellular telephony (6)
7. Explain third generation 3G of cellular telephony. (6*)
8. Explain fourth generation 4G of cellular telephony (6*)
9. Explain the following terms with reference to satellite (4*)
   i) Orbits ii) Footprint
10. Explain the 3 categories of satellites. (8*)

MODULE 5(CONT.): NETWORK LAYER PROTOCOLS

11. Explain various field of IPv4. (8*)
12. Explain fragmentation. Explain 3 fields related to fragmentation (6*)
13. Explain options of IPv4. (6*)
14. Explain three network attacks to IP protocol. Also, explain four services of IPSec. (8*)
15. With general format, explain various ICMPv4 messages. (6*)
16. Explain two tools that use ICMP for debugging. (6)
17. Explain the following term with reference to Mobile IP: (4*)
    i) Home address ii) Care-of address iii) Home-agent iv) Foreign-agent
18. Explain three phases for communication in Mobile IP. (8*)

MODULE 5(CONT.): NEXT GENERATION IP

19. Explain 3 address types of IPv6. (6)
20. Explain changes from IPv4 to IPv6. (4*)
21. Explain various field of IPv6. (8*)
22. Explain various extension header of IPv6. (8)
23. Explain various ICMPv6 messages. (6)
24. Explain various group membership messages. (6)
25. Explain 3 ways to make transition from IPv4 to IPv6. (6)

MODEL PAPER-1

1a) List the differences between LAN & WAN. (4 Marks)
1b) Explain the following topologies:
i) Mesh ii) Star (8 Marks)
1c) Explain 4 levels of addressing employed in TCP/IP protocol. (4 Marks)

2a) Explain the theoretical formula which was developed to calculate the data rate. What are the 3 factors on which data rate depends? (10 Marks)
2b) We need to send 265 kbps over a noiseless channel with a bandwidth of 20 kHz. How many signal levels do we need? (2 Marks)
2c) Explain manchester & differential-manchester encoding schemes. Represent the sequence 101011100 using the same encoding schemes. (4 Marks)

3a) Explain non-uniform quantization & how to recover original signal using PCM decoder. (6 Marks)
3b) An analog signal has a bit rate of 8000 bps & a baud rate of 1000 baud. How many data elements are carried by each signal element? How many signal elements do we need? (4 Marks)
3c) Describe ASK, FSK and PSK mechanisms and apply them over the digital data 101101. (6 Marks)

4a) What is multiplexing? Explain the FDM multiplexing and demultiplexing process with neat diagrams. (6 Marks)
4b) Four 1-kbps connections are multiplexed together. A unit is 1 bit. (4 Marks)
Find (i) the duration of 1 bit before multiplexing
(ii) the transmission rate of the link
(iii) the duration of a time slot and
(iv) the duration of a frame.
4c) Explain in detail circuit-switched-network. (6 Marks)

5a) Explain error detection using block coding technique. (10 Marks)
5b) Given the dataword 101001111 and the divisor 10111, show the generation of the CRC codeword at the sender site. (6 Marks)

6a) Differentiate between character oriented and bit oriented format for framing. (4 Marks)
6b) What is PPP? With a neat diagram, explain the frame structure of PPP. Also, explain framing and transition phases in PPP. (12 Marks)

7a) Explain reservation access, polling access & token passing access methods. (12 Marks)
7b) List out 5 goals of fast Ethernet. Explain auto-negotiation. (4 Marks)

8a) Explain MAC sublayer in gigabit-Ethernet (6 Marks)
8b) Explain architecture of IEEE 802.11 (10 Marks)

9a) Explain various components of cellular system with neat diagram. (6 Marks)
9b) Explain the following terms with reference to satellite (4 Marks)
i) Orbits ii) Footprint
9c) Explain various field of IPv4. (6 Marks)

10a) Explain the following term with reference to Mobile IP: (8 Marks)
i) Home address ii) Care-of address iii) Home-agent iv) Foreign-agent
10b) Explain 3 ways to make transition from IPv4 to IPv6. (8 Marks)

MODEL PAPER-2

1a) Explain the following topologies: (8 Marks)
i) Bus ii) Ring
1b) List the 5 layers and its functionality in TCP/IP model. (8 Marks)

2a) Explain 4 performance parameters of network. (8 Marks)
2b) i) A network with bandwidth of 10 Mbps can pass only an average of 12,000 frames per minute with each frame carrying an average of 10,000 bits. What is the throughput of this network? (2 Marks)
ii) Using Shannon theorem, calculate the maximum bit rate of the channel having bandwidth of 3100 Hz & SNR_{db} of 20 db. (2 Marks)
2c) Define the following: (4 Marks)
i) Line coding ii) Internet iii) SNR iv) Decibel

3a) A complex low-pass signal has a bandwidth of 200 kHz. What is the minimum sampling rate for this signal? (2 Marks)
3b) Explain different aspects of digital-to-analog conversion. (8 Marks)
3c) Define ASK. Explain BASK. (6 Marks)

4a) Five channels, each with a 100-kHz bandwidth, are to be multiplexed together. What is the minimum bandwidth of the link if there is a need for a guard band of 10 kHz between the channels to prevent interference? (2 Marks)
4b) Explain in detail synchronous TDM. (8 Marks)
4c) Explain in detail datagram networks. (6 Marks)

5a) Explain checksum with example. Also, write algorithm for Internet Checksum. (12 Marks)
5b) Explain two types of errors. (4 Marks)

6a) Explain HDLC frame format. Also, explain control fields in HDLC. (12 Marks)
6b) Explain the concept of piggybacking. (4 Marks)

7a) Explain CSMA/CA & CSMA/CD. (10 Marks)
7b) Explain frame format of standard Ethernet. (6 Marks)

8a) Explain frame format of IEEE 802.11. (6 Marks)
8b) Explain hidden station problem. (4 Marks)
8c) What is Bluetooth? Explain architecture of Bluetooth. (6 Marks)

9a) What is WiMAX? Explain WiMAX MAC frame format. (6 Marks)
9b) What is cellular telephony? Explain third generation 3G of cellular telephony. (4 Marks)
9c) What is ICMP? With general format, explain various ICMPv4 messages. (6 Marks)

10a) What is Mobile IP? Explain three phases for communication in Mobile IP. (8 Marks)
10b) Explain various field of IPv6. (8 Marks)

MODEL PAPER-3

1a) Define data communications. Explain its 4 fundamental characteristics. (4 Marks)
1b) Explain 3 diff. methods of data flow. Also, explain point to point & multipoint connection. (6 Marks)
1c) Explain in detail LAN. (6 Marks)

2a) What is transmission impairment? Explain causes of transmission impairment. (6 Marks)
2b) i) Consider a noiseless channel with a bandwidth of 3000 Hz transmitting a signal with two signal levels. What is the maximum bit rate? (2 Marks)
ii) The loss in a cable is usually defined in decibels per kilometer (dB/km). If the signal at the beginning of a cable with −0.3 dB/km has a power of 2 mW, what is the power of the signal at 5 km? (2 Marks)
2c) Explain following encoding schemes with example as the sequence 10110011: (8 Marks)
i) Unipolar Scheme ii) Polar Schemes iii) Bipolar Schemes

3a) Explain different types of transmission modes. (8 Marks)
3b) Define FSK. Explain BFSK. (6 Marks)
3c) An analog signal carries 4 bits per signal element. If 1000 signal elements are sent per second, find the bit rate. (2 Marks)

4a) Define and explain the concept of WDM. (4 Marks)
4b) What is spread spectrum? Explain in detail DSSS. (6 Marks)
4c) Explain data transfer phase in Virtual-circuit networks. (6 Marks)

5a) Explain hamming distance for error detection. (6 Marks)
5b) Explain CRC with block diagram & example. (10 Marks)

6a) Explain bit oriented protocol. (6 Marks)
6b) Explain Stop-and-Wait protocol. (10 Marks)

7a) Explain pure ALOHA. A pure ALOHA network transmits 200-bit frames on a shared channel of 200 kbps. What is the requirement to make this frame collision-free? (6 Marks)
7b) Explain TDMA & FDMA. (8 Marks)
7c) Explain addressing in standard Ethernet. (2 Marks)

8a) Explain frame types of IEEE 802.11. (4 Marks)
8b) Explain exposed station problem. (6 Marks)
8c) Explain frame format of Bluetooth. (6 Marks)

9a) What is WiMAX? Explain two types of services of WiMAX. (4 Marks)
9b) Discuss the operation of the cellular telephony. (8 Marks)
9c) Explain fragmentation. Explain 3 fields related to fragmentation (4 Marks)

10a) Explain 3 network attacks to IP protocol. Also, explain four services of IPSec. (8 Marks)
10b) Explain various extension header of IPv6. (8 Marks)

MODEL PAPER-4

1a) Define data communication. Explain different components of data communication system. Explain the 3 criteria necessary for an effective and efficient network. (6 Marks)
1b) Define network. Explain in detail WAN. (6 Marks)
1c) Explain different method of accessing the Internet. (4 Marks)

2a) Explain 2 methods for transmitting a digital signal (8 Marks)
2b) The power of a signal is 10 mW and the power of the noise is 1 μW; what are the values of SNR and SNR_{dB}? (2 Marks)
2c) Explain in detail any 6 characteristics of digital signal. (6 Marks)

3a) Explain the PCM encoder with neat diagram. (10 Marks)
3b) Define PSK. Explain BPSK. (6 Marks)

4a) Explain in detail Statistical TDM. (6 Marks)
4b) Explain in detail FHSS. (6 Marks)
4c) Compare circuit-switched-network, datagram & virtual-circuit. (4 Marks)

5a) Write short notes on polynomial codes. (6 Marks)
5b) Explain parity-check code with block diagram. (10 Marks)

6a) Explain character oriented protocol. (8 Marks)
6b) Explain 3 type of frames used in HDLC. (8 Marks)

7a) Explain slotted ALOHA & CSMA. (10 Marks)
7b) Explain briefly any 2 implementation of standard Ethernet. (6 Marks)

8a) Explain addressing in IEEE 802.11. (8 Marks)
8b) Explain layers of Bluetooth. (8 Marks)

9a) Explain the 3 categories of satellites. (10 Marks)
9b) Explain fourth generation 4G of cellular telephony. (6 Marks)

10a) Explain changes from IPv4 to IPv6. (4 Marks)
10b) Explain 3 address types of IPv6. (4 Marks)
10c) Explain various ICMPv6 messages. (6 Marks)