DATA COMMUNICATION
18CS46
4TH SEM

Table of Contents

MODULE – 1

INTRODUCTION

DATA COMMUNICATIONS

Telecommunication-communication at a distance.

Data -information presented in whatever form is agreed upon by the parties creating and using the data.

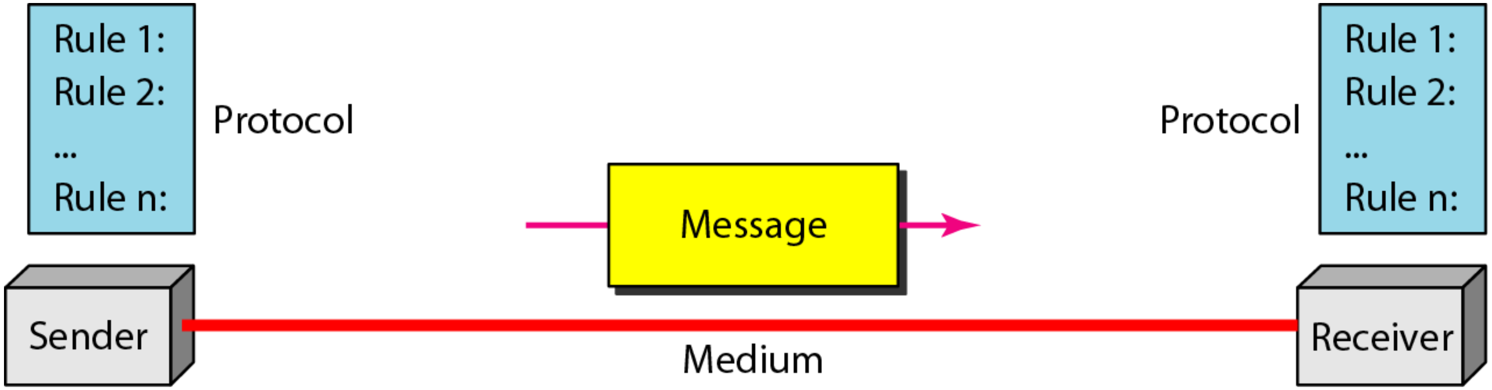
Data communications are the exchange of data between two devices via some form of transmission medium such as a wire cable.

Fundamental characteristics[Effectiveness of DC]:

1. *Delivery-correct destination*
2. *Accuracy-deliver data accurately*
3. *Timeliness- delivery data in timely manner*
4. *Jitter-variation in packet arrival time.*

Components of a data communication system

- **Message-** information to be communicated
- **Sender-** send the data message
- **Receiver –** receives the message
- **Transmission Medium-** physical path by which message travels from sender to receiver.
- **Protocol-** set of rules govern the data communications



Data Representation

- Text
    - Represented as bit pattern (sequence of bits 0s or 1s)
    - Different set of bit pattern used to represent symbols or characters.
    - Each set is called code
    - Process of representing symbols is called encoding
    - Ex: ASCII,UNICODE

- Numbers
    - Represented as bit pattern
    - Directly converted to binary form

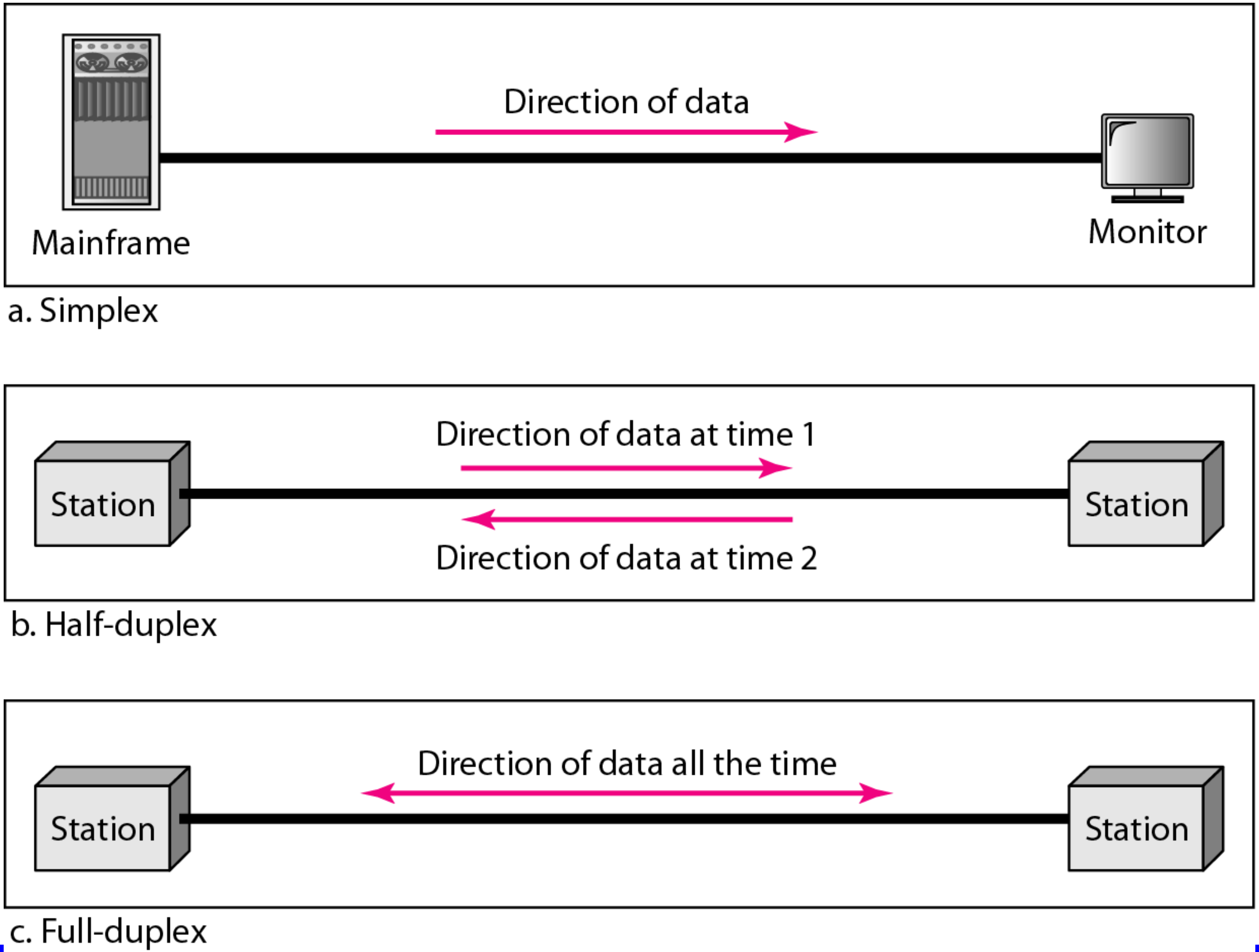
- Audio
    - Recording or broadcasting of sound or music.
    - Continuous not discrete

Data Representation (Cont.…)

- Video
  - Recording or broadcasting of picture or a movie
  - Produced as :
    - Continuous entity [TV camera]
    - Combination of images-discrete entity

- Images
  - Represented as bit pattern
  - Image is divided into matrix of pixels(smallest element of an image)
  - Each pixel is assigned a bit pattern (size and value of pattern depend on image)
  - Ex: black and white dots (chessboard) -1 bit pattern is enough to represent a pixel, gray scale- 2 bit pattern.
  - Several methods to represent colour images : RGB,YCM

Data flow (simplex, half-duplex, and full-duplex)


Mainframe
Direction of data
Monitor
Station
Direction of data at time 1
Direction of data at time 2
Station
Station
Direction of data all the time
Station

a. Simplex

b. Half-duplex

c. Full-duplex

NETWORKS

A network is a set of devices (often referred to as nodes) connected by communication links.

A node can be a computer, printer, or any other device capable of sending and/or receiving data generated by other nodes on the network.

A link can be a cable, air, optical fiber, or any medium which can transport a signal carrying information.

Topics discussed in this section:

- Network Criteria
- Physical Structures
- Categories of Networks

Network Criteria

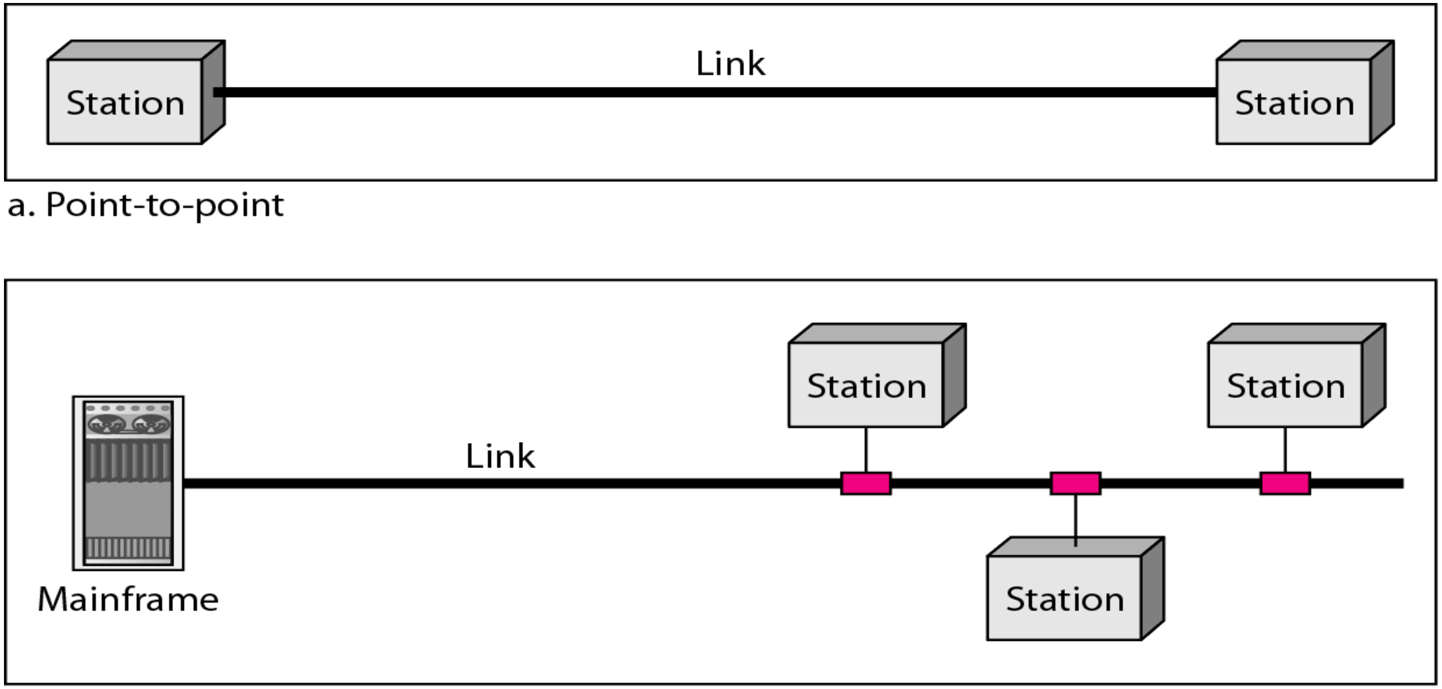
- **Performance**
  - **Measured using:**
    - **Transit time**: time taken to travel a message from one device to another.
    - **Response time**: time elapsed between enquiry and response.
  - **Depends on following factors:**
    - Number of users
    - Type of transmission medium
    - Efficiency of software
  - **Evaluated by 2 networking metrics:**
    - Throughput (high)
    - Delay (small)

Network Criteria (cont..)

- **Reliability**
  - **Measured by**
    - Frequency of failure.
    - Time taken to recover from a network failure.
    - Network robustness in a disaster.
- **Security**
  - **Protecting** data from unauthorized access, damage and development.
  - **Implementing** policies and procedures for recovery from breaches and data losses.

□Physical Structures

- Type of Connection
  - Point to Point - single transmitter and receiver
  - Multipoint - multiple recipients of single transmission



a. Point-to-point

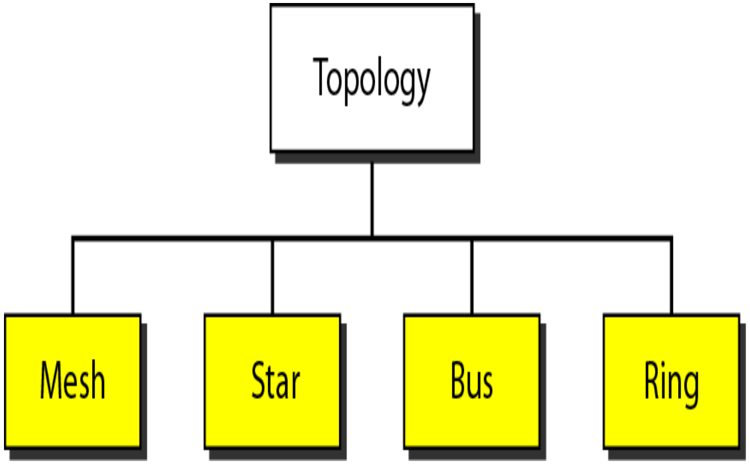
b. Multipoint

Physical Structures (Cont..)

- **Physical Topology:**

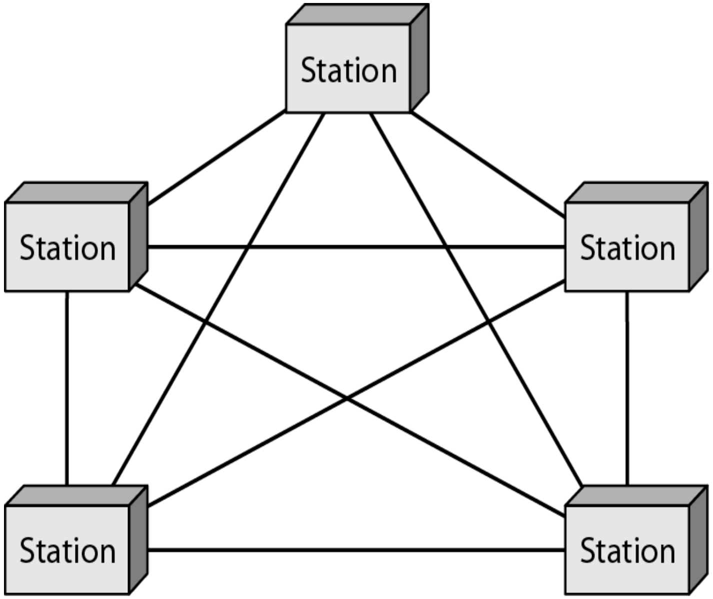
  - **Way in which network is laid out physically.**
  - **Two or more links form a topology.**
  - **“ *Topology of network is the geometric representation of all links and linking devices to one another”.***

  - ***Basic topologies:***
    - *Mesh*
    - *Star*
    - *Bus and*
    - *Ring*



Mesh Topology

- All devices are connected to each other
- Dedicated point-point link between all devices
- “n(n-1)” physical channels to link “n” device.
- For ‘n’ nodes
  - n(n-1) physical links
  - n(n-1)/2 duplex mode links
- Every device have (n-1) I/O ports to be connected to other (n-1) devices.



Fully connected Mesh Topology

Mesh Topology (cont..)

Advantages:

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.

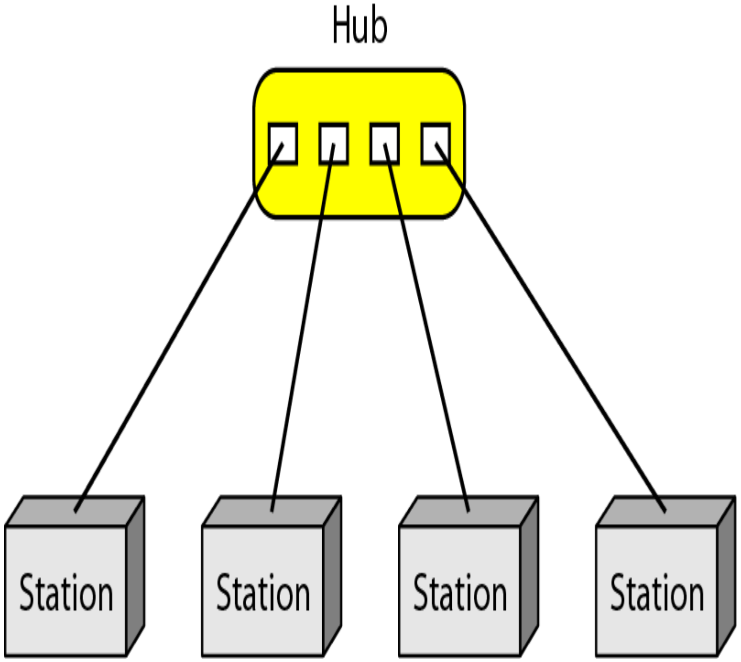
Disadvantages:

1) Difficult installation and reconfiguration.
2) Bulk of wiring occupies more space than available space.
3) Very expensive: hardware required to connect each link is expensive.

practical example: connection of telephone regional offices in which each regional office needs to be connected to every other regional office.

Star Topology

- Point to Point connection
- All the devices are connected to a central controller called a hub
- 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.



A star topology connecting four stations

Star Topology (cont..)

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.

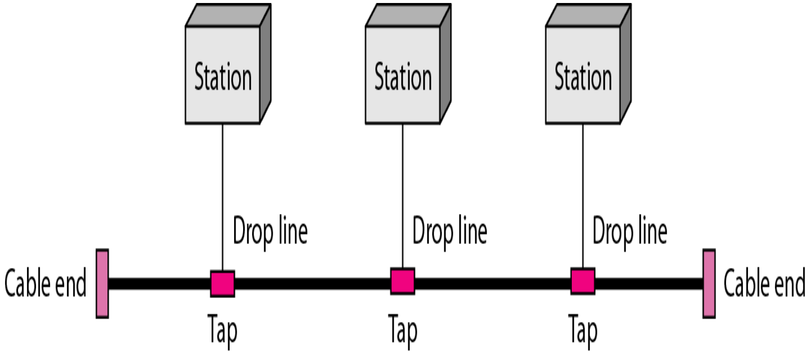
Disadvantages:

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**.

Example: Local area network

Bus Topology

- Multipoint connection
- All the devices are connected to the single cable called bus (backbone)
- Devices are connected to the bus by drop-lines and taps.
- A drop-line is a connection running between the device and the bus (main cable).
- A tap is a connector that links to the bus
- Limit on number of taps a bus can support and distance between those taps-(as signal travels along backbone some energy is transferred to heat which makes it weaker as it travels farther)



A bus topology connecting three stations

Bus Topology

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:

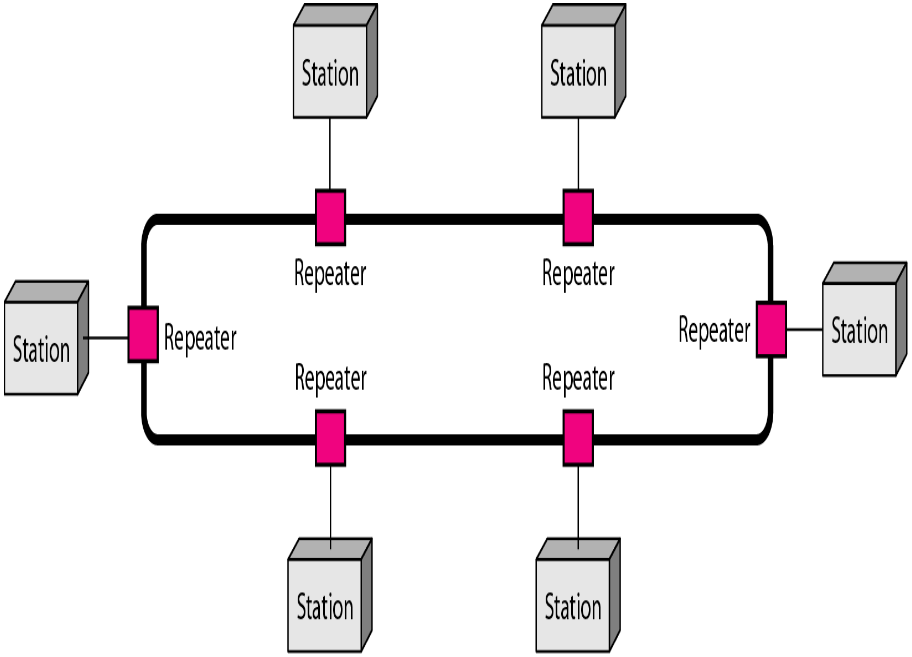
1) Difficult to detect and troubleshoot fault.
2) Signal reflection at the taps can cause degradation in quality.
3) A fault/break in the cable stops all transmission.
4) There is a limit on
i) Cable length
ii) Number of nodes that can be connected.
5) Security is very low because all the devices receive the data sent from the source.

Ring Topology

- Each device is connected to the next, forming a ring
- There are only two neighbours 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.



A ring topology connecting six stations

Ring Topology

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.

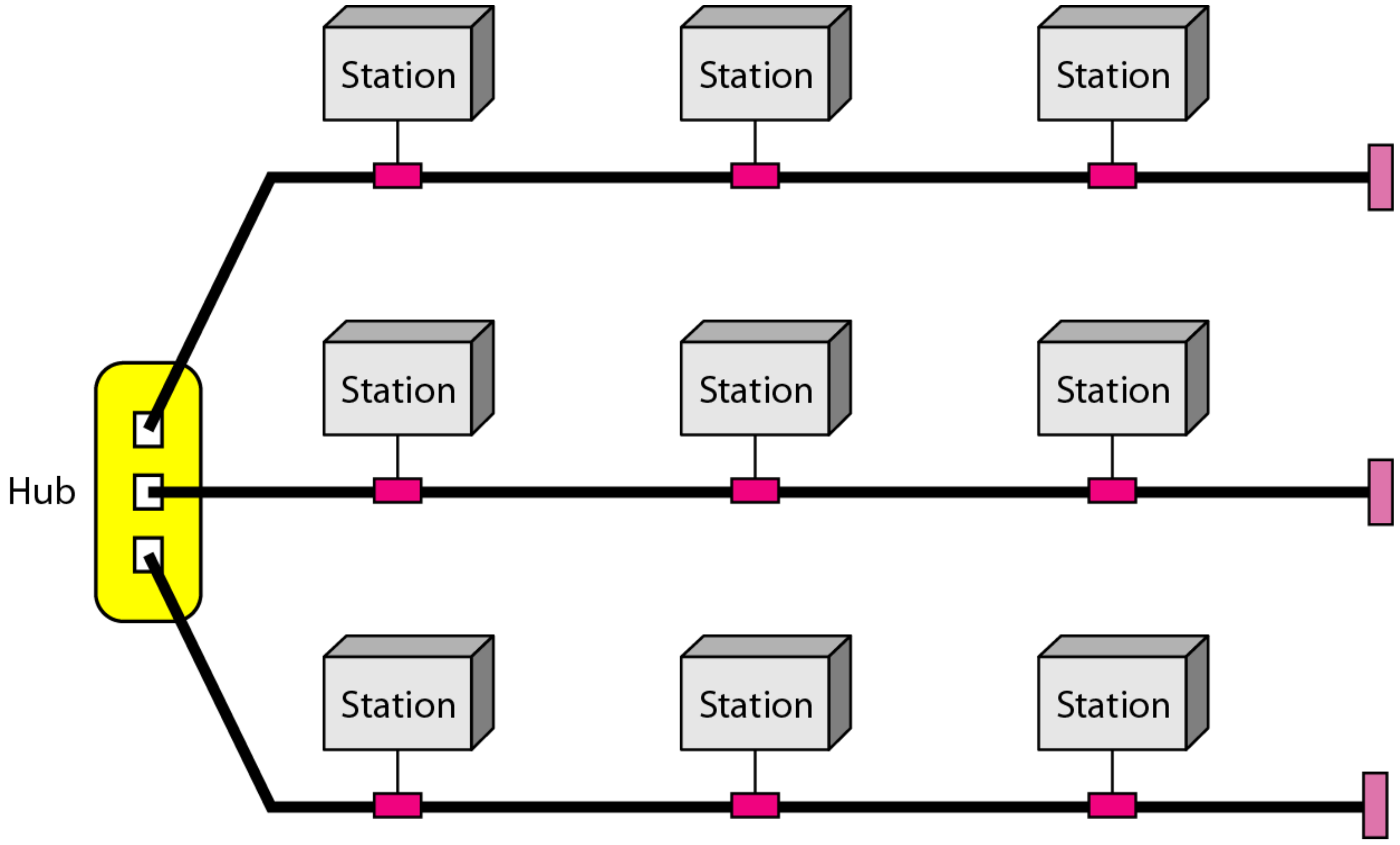
3) There is a limit on

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.

A hybrid topology: a star backbone with three bus networks



Categories of Networks

Few criteria –size, geographical coverage and ownership to make this distinction.

- **Local Area Networks (LANs)**
  - Short distances
  - Designed to provide local interconnectivity
- **Wide Area Networks (WANs)**
  - Long distances
  - Provide connectivity over large areas
- **Metropolitan Area Networks (MANs)**
  - Provide connectivity over areas such as a city, a campus

Local Area Networks (LANs)

- It is used to connect computers in a single office, building or campus.
- 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.(extend to audio and video devices.)
- 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.

Local Area Networks (LANs) (cont…)

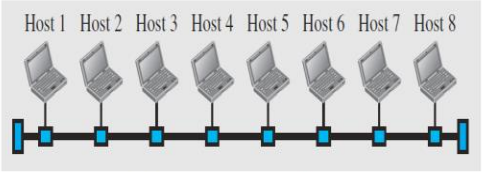
- LAN with common cable : packet sent by host is received by all hosts. Intended host keep the packet other drop the packet.

- LAN use a smart connecting switch.
- The switch is able to

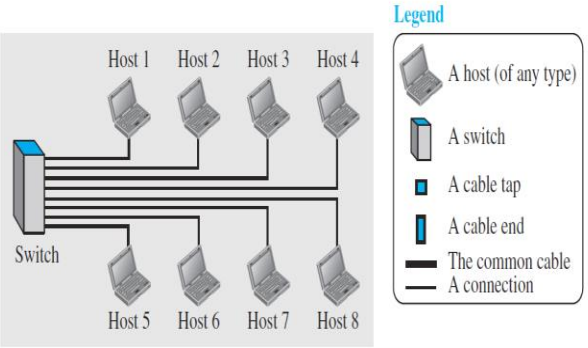
→ recognize the destination address of the packet & guide the packet to its destination.

→ reduces the traffic in the LAN

→ allows more than one pair to communicate with each other at the same time.



a. LAN with a common cable (past)



b. LAN with a switch (today)

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.

Wide Area Networks (WAN)

- WAN can cover wider geographical area. It can cover cities, states, countries and even world.
- 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.

Two distinct examples of WANs:

- Point-to-point WAN
- Switched WAN

Wide Area Networks (WAN) (cont…)

Two distinct examples of WANs:

- Point-to-point WAN:A point-to-point WAN is a network that connects 2 communicating devices through a transmission media

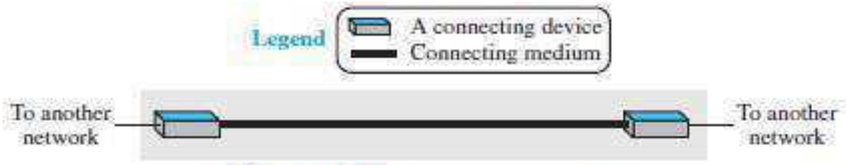


Figure 1.9 *A point-to-point WAN*

- 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

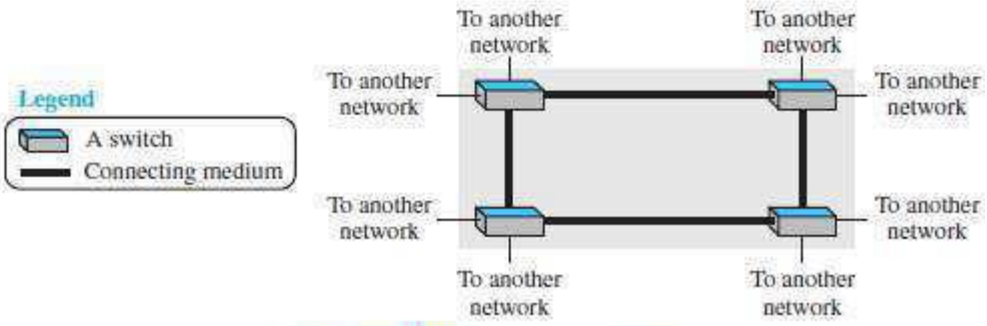


Figure 1.10 *A switched WAN*

Wide Area Networks (WAN) (cont…)

Internetwork:

- A network of networks is called an **internet**. (inter-network)
- EX: Assume that an organization has two offices, First office is on the **east coast** & 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.
- 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.

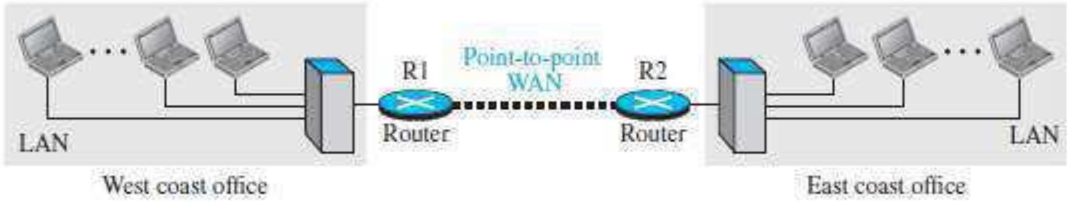


Figure 1.11 *An internetwork made of two LANs and one point-to-point WAN*

Wide Area Networks (WAN) (cont…)

Internetwork:

Internet with several LANs and WANs connected. One of the WANs is a switched WAN with four switches.

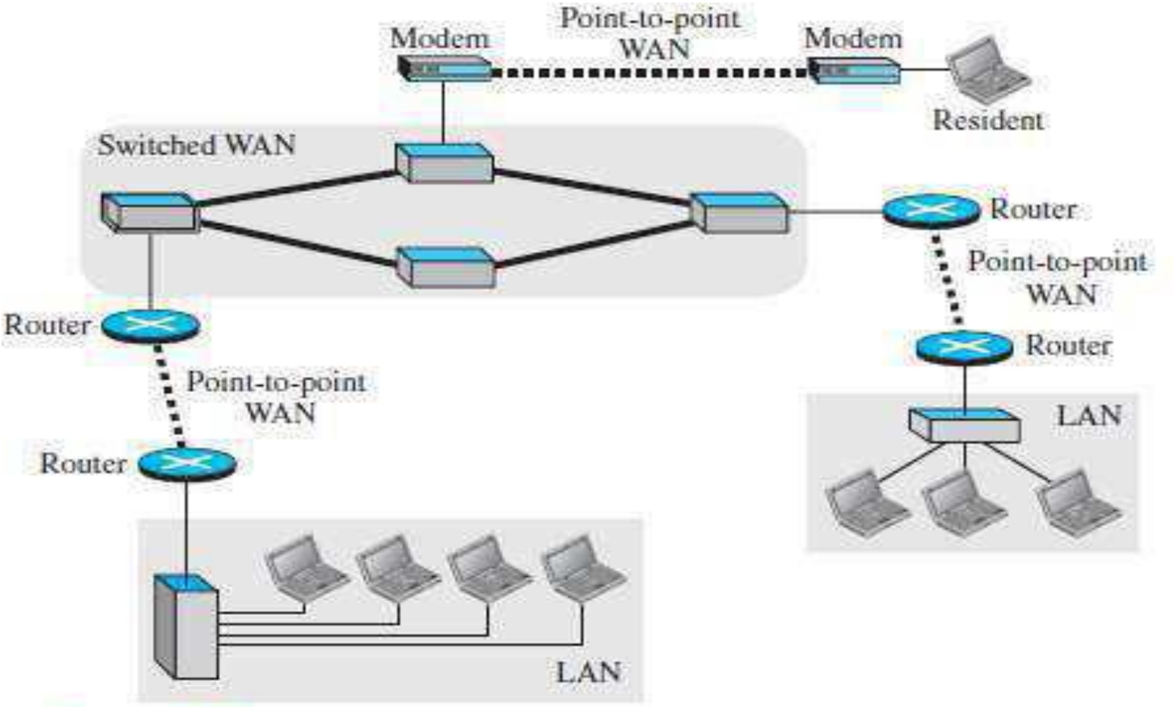


Figure 1.12 *A heterogeneous network made of four WANs and three LANs*

Wide Area Networks (WAN) (cont…)

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.

Wide Area Networks (WAN) (cont…)

- **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**.
  - Ex: In the figure below 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.

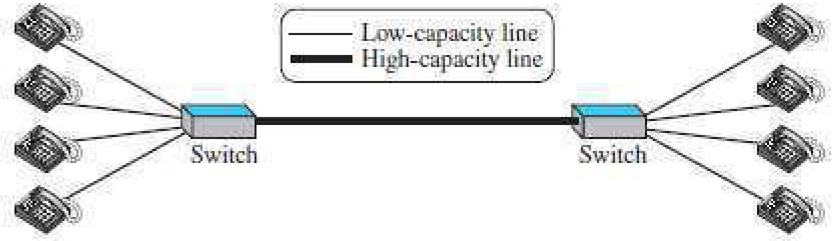


Figure 1.13 *A circuit-switched network*

Wide Area Networks (WAN) (cont…)

- **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..
  - As shown in Figure below, 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.

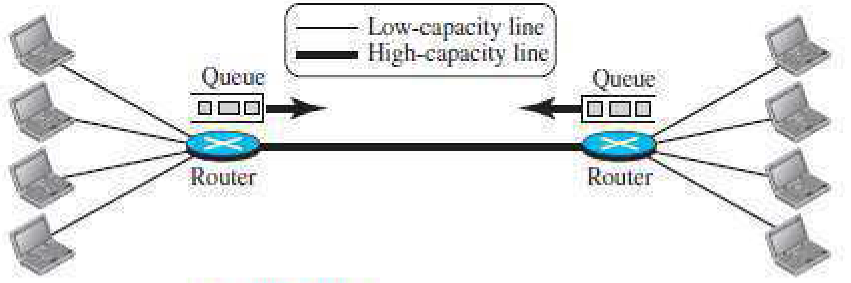


Figure 1.14 *A packet-switched network*

THE INTERNET

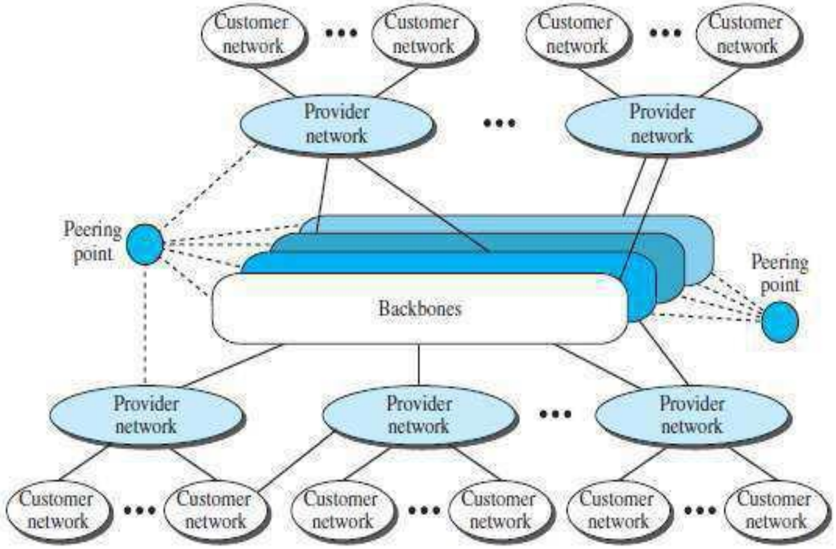


Figure 1.15 *The Internet today*

- The most notable internet(lowercase i) is called the Internet (uppercase I) and is composed of thousands of interconnected networks.
- Internet is made up of

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 (smaller network)

- 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.

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: Dial-up service and DSL Service
- 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. speed Internet services to residences

Accessing the Internet (cont..)

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 neighbours 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.

- For example, a large university with several campuses can create an internetwork and then connect to the Internet.

Internet History

- Internet has evolved from a private network to a global one in less than 40 years.

Early History

- Before 1960 there were some communication networks such as telegraph and telephone networks.
- Suitable for constant-rate communication (after a connection was made between two users, the encoded message(telegraphy) or voice (telephony) could be exchanged).
- Computer network should be able to handle bursty data (data received at variable rates at different times)

Birth of packet switched networks:

- The theory of packet switching for bursty traffic was first presented by Leonard Kleinrock in 1961 at MIT.
- At the same time, two other researchers, Paul Baran at Rand Institute and Donald Davies at National Physical Laboratory in England, published some papers about packet-switched networks.

Internet History (cont…)

ARPANET:

- In the mid-1960s, mainframe computers in research organizations were stand-alone devices.
- Computers from different manufacturers were unable to communicate with one another.
- The Advanced Research Projects Agency (ARPA) in the Department of Defence (DOD) was interested in finding a way to connect computers so that the researchers share their findings, thereby reducing costs and eliminating duplication of effort.
- In 1967, at an Association for Computing Machinery (ACM) meeting, ARPA presented its ideas for the Advanced Research Projects Agency Network (ARPANET), a small network of connected computers.
- The idea was that each host computer (not necessarily from the same manufacturer) would be attached to a specialized computer, called an *interface message processor* (IMP).
- The IMPs, in turn, would be connected to each other. Each IMP had to be able to communicate with other IMPs as well as with its own attached host.
- By 1969, ARPANET was a reality.
- Four nodes, at the University of California at Los Angeles (UCLA), the University of California at Santa Barbara (UCSB), Stanford Research Institute (SRI), and the University of Utah, were connected via the IMPs to form a network.
- Software called the *Network Control Protocol* (NCP) provided communication between the hosts.

Internet History (cont…)

Birth of the Internet

- In 1972, Vint Cerf and Bob Kahn, both of whom were part of the core ARPANET group, collaborated on what they called the *Internetting Project*.
- They wanted to link dissimilar networks so that a host on one network could communicate with a host on another.
- There were many problems to overcome: diverse packet sizes, diverse interfaces, and diverse transmission rates, as well as differing reliability requirements.
- Cerf and Kahn devised the idea of a device called a *gateway* to serve as the intermediary hardware to transfer data from one network to another.

TCP/IP

- New version of NCP-transmission control protocol (TCP) included concepts such as encapsulation, the datagram, and the functions of a gateway.
- In October 1977, an internet consisting of three different networks (ARPANET, packet radio, and packet satellite) was successfully demonstrated. Communication between networks was now possible.
- Authorities made a decision to split TCP into two protocols: Transmission Control Protocol (TCP) and Internet Protocol (IP). IP would handle datagram routing while TCP would be responsible for higher level functions such as segmentation, reassembly, and error detection. The new combination became known as TCP/IP.

Internet History (cont…)

MILNET:
In 1983, ARPANET split into two networks: Military Network (MILNET) for military users and ARPANET for non-military users.

CSNET:

- Creation of CSNET in 1981.
- Computer Science Network (CSNET) was a network sponsored by the National Science Foundation (NSF).
- The network was conceived by universities that were ineligible to join ARPANET due to an absence of ties to the Department of Defense.
- CSNET was a
  - Less expensive network;
  - there were no redundant links and
  - the transmission rate was slower.
- By the mid-1980s, most U.S. universities with computer science departments were part of CSNET.

Internet History (cont…)

NSFNET

- With the success of CSNET, the NSF in 1986 sponsored the National Science Foundation Network (NSFNET), a backbone that connected five supercomputer centres located throughout the United States.
- Community networks were allowed access to this backbone, a T-1 line with a 1.544-Mbps data rate, thus providing connectivity throughout the United States.
- In 1990, ARPANET was officially retired and replaced by NSFNET. In 1995, NSFNET reverted back to its original concept of a research network.

ANSNET

- In 1991, the U.S. government decided that NSFNET was not capable of supporting the rapidly increasing Internet traffic.
- Three companies, IBM, Merit, and Verizon, filled the void by forming a non-profit organization called Advanced Network & Services (ANS) to build a new, high-speed Internet backbone called Advanced Network Services Network (ANSNET).

Internet Today

- Rapid growth both in the infrastructure and new applications.
- The Internet today is a set of pier networks that provide services to the whole world.
- What has made the Internet so popular is the invention of new applications.

World Wide Web

- The 1990s saw the explosion of Internet applications due to the emergence of the World Wide Web (WWW).
- The Web was invented at CERN by Tim Berners-Lee. This invention has added the commercial applications to the Internet.

Multimedia

Recent developments in the multimedia applications such as

- voice over IP (telephony),
- video over IP (Skype),
- view sharing (YouTube), and
- television over IP (PPLive)

increased the number of users and the amount of time each user spends on the network.

Peer-to-Peer Applications (P2P)

Peer-to-peer networking is also a new area of communication

Ex: BitTorrent

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).

Standards and Administration (cont..)

Maturity Levels

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

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.

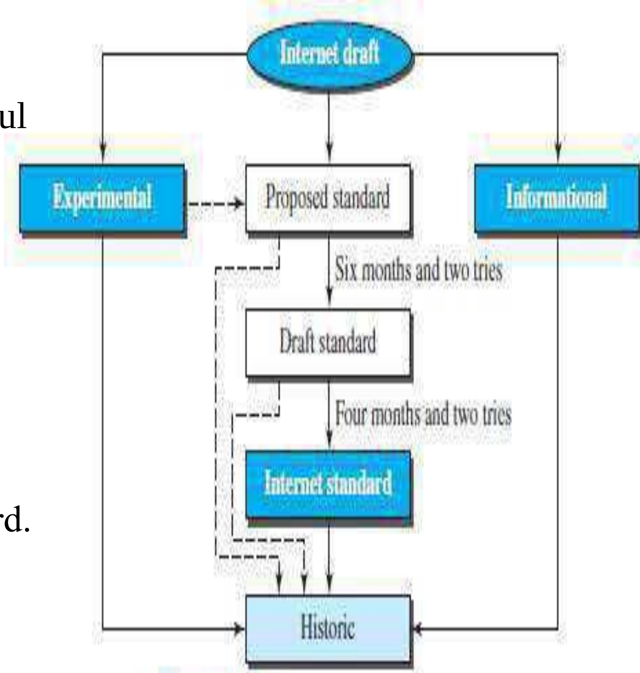


Figure 1.16 *Maturity levels of an RFC*

Standards and Administration (cont..)

Requirement Levels

•RFCs are classified into 5 requirement levels:

1) Required

□ An RFC labelled 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.

Internet Administration

1) ISOC

- ISOC is a non-profit organization formed to provide support for Internet standards process
- 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.

Standards and Administration (cont..)

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.

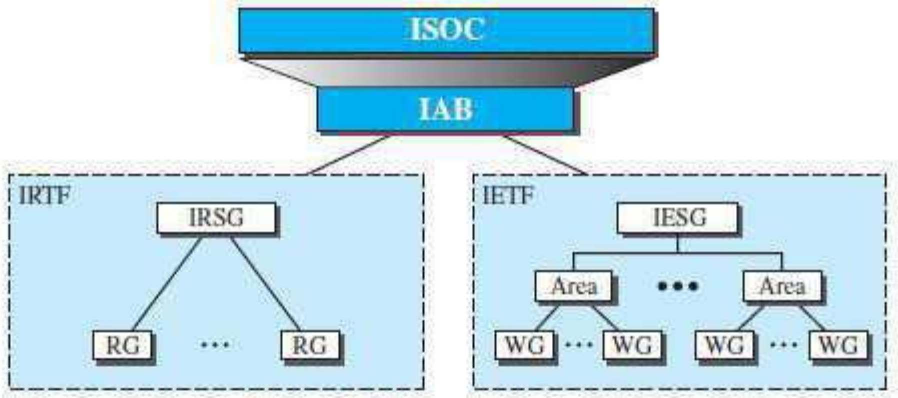


Figure 1.17 *Internet administration*

Chapter 2

Network Models

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 -only one simple protocol.
  - complex, we need to divide the task b/w different layers. We need a protocol at each layer, or protocol layering.

Scenarios

First Scenario

- Communication is so simple that it can occur in only one layer.
- Assume Maria and Ann are neighbours with a lot of common ideas.
- Communication between Maria and Ann takes place in one layer, face to face, in the same language

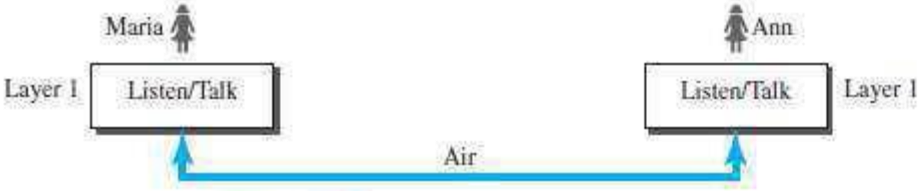


Figure 2.1 *A single-layer protocol*

PROTOCOL LAYERING (Cont..)

Second Scenario

- Maria and Ann communicate using regular mail through the post office .
- However, they do not want their ideas to be revealed by other people if the letters are intercepted.
- They agree on an encryption/decryption technique.
- The sender of the letter encrypts it to make it unreadable by an intruder; the receiver of the letter decrypts it to get the original letter.

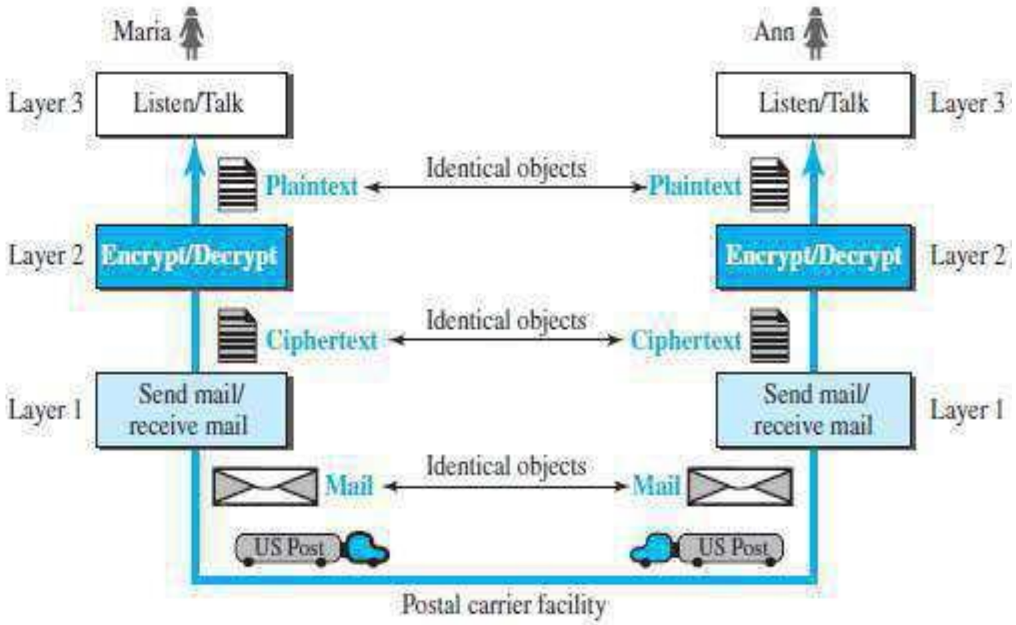


Figure 2.2 A three-layer protocol

PROTOCOL LAYERING (Cont..)

- 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.

PROTOCOL LAYERING (Cont..)

Logical Connections

- There is a logical (imaginary) connection at each layer through which two end systems can send the object created from that layer..

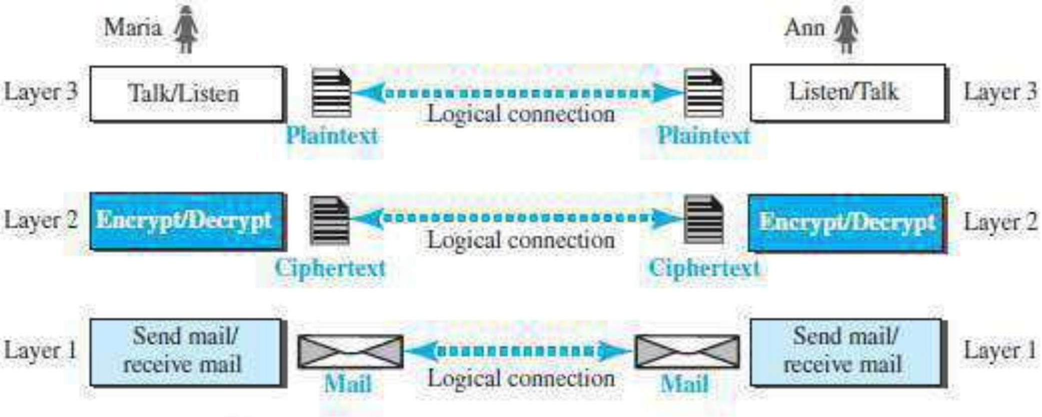


Figure 2.3 *Logical connection between peer layers*

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.
- TCP/IP is thought of as a five-layer model.

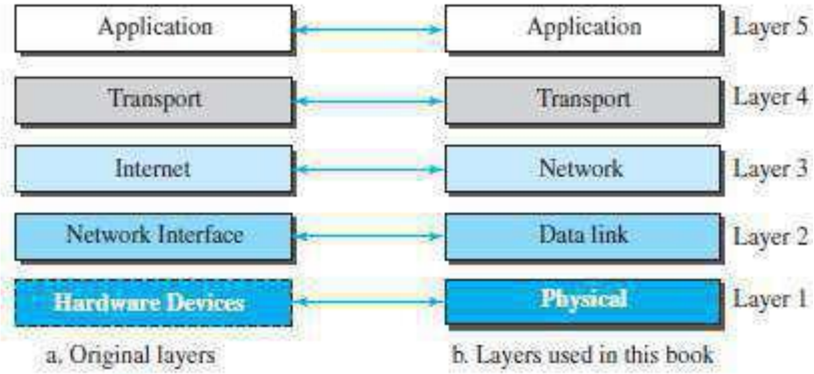


Figure 2.4 *Layers in the TCP/IP protocol suite*

TCP/IP PROTOCOL SUITE (Cont..)

Layered Architecture

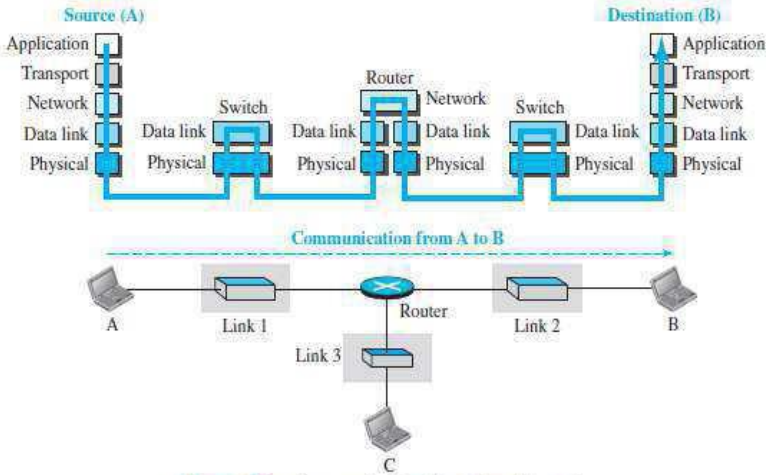


Figure 2.5 *Communication through an internet*

Let us assume that computer A communicates with computer B

- As shown in the Figure, 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.

TCP/IP PROTOCOL SUITE (Cont..)

Layers in TCP/IP Protocol suite

- The duty of the
  - Application, Transport, and Network layers is end-to-end.
  - Data-link and Physical layers is hop-to-hop (hop is a host or router).
- The domain of duty of the
  - top three layers is the internet.
  - 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.

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

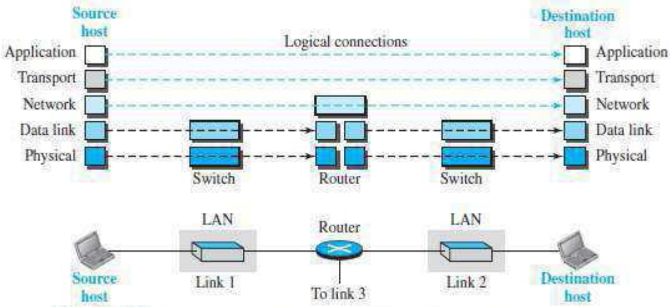
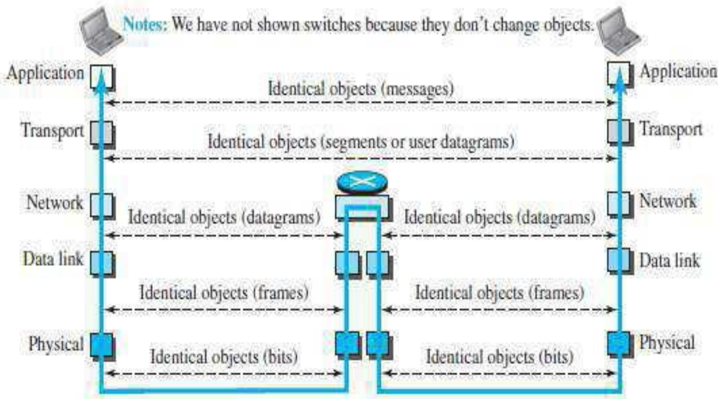


Figure 2.6 *Logical connections between layers of the TCP/IP protocol suite*



Application
Transport
Network
Data link
Physical
Identical objects (messages)
Identical objects (segments or user datagrams)
Identical objects (datagrams)
Identical objects (datagrams)
Identical objects (frames)
Identical objects (frames)
Identical objects (bits)
Identical objects (bits)
Application
Transport
Network
Data link
Physical

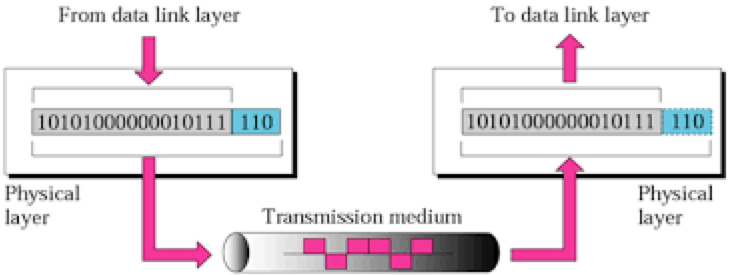
Notes: We have not shown switches because they don't change objects.

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

Description of Each layer

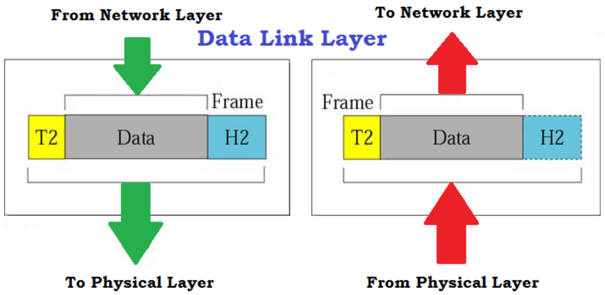
Physical Layer

- 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

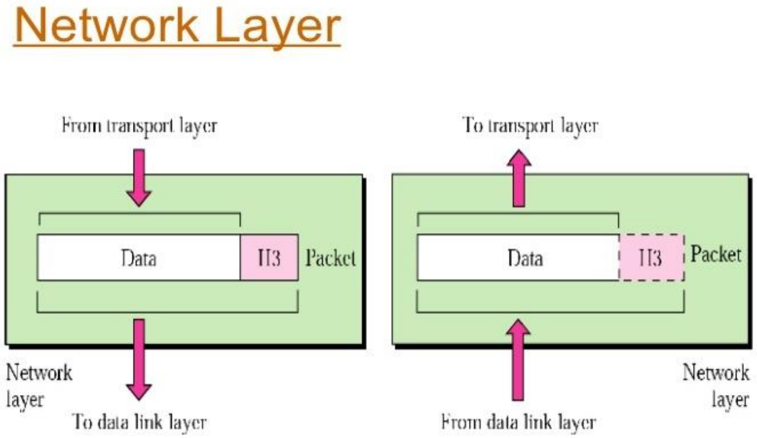
- Data-link-layer (DLL) is responsible for moving frames from one node to another node over a link.
- The link can be wired LAN/WAN or wireless LAN/WAN.
- The data-link layer
  - → gets the datagram from network layer
  - → encapsulates the datagram in a packet called a frame.
  - → sends the frame to 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..



Description of Each layer

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 4 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 transportlayer 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: (as listed below)

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.

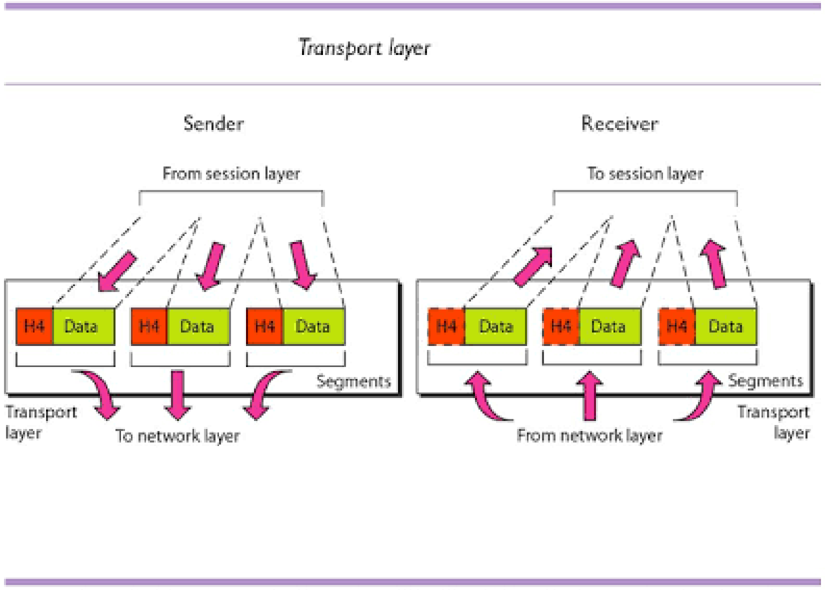
□ It does not provide flow, error, or congestion control.

□ 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 anothe

4) DNS is used to find the IP address of a computer.

5) SNMP is used to manage the Internet at global and local

6) HTTP is used for accessing the World Wide Web (WWW

(FTP □ File Transfer Protocol

SMTP □ Simple Mail Transfer Protocol)

(DNS □ Domain Name System

HTTP □ Hyper Text Transfer Protocol)

(SNMP □ Simple Network Management Protocol

TELNET □ Terminal Network)

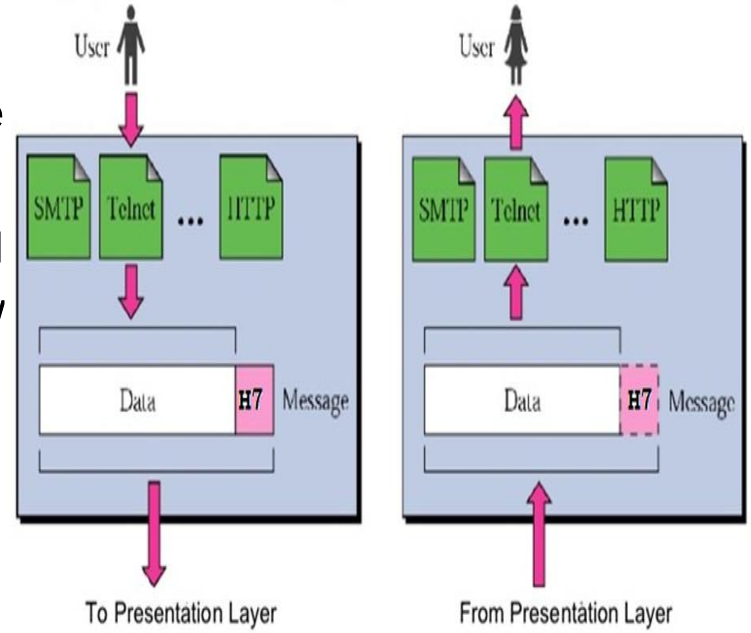


Fig: Application Layer

Encapsulation and Decapsulation

- One of the important concepts in protocol layering in the Internet.
- Figure below shows this concept for the small internet.

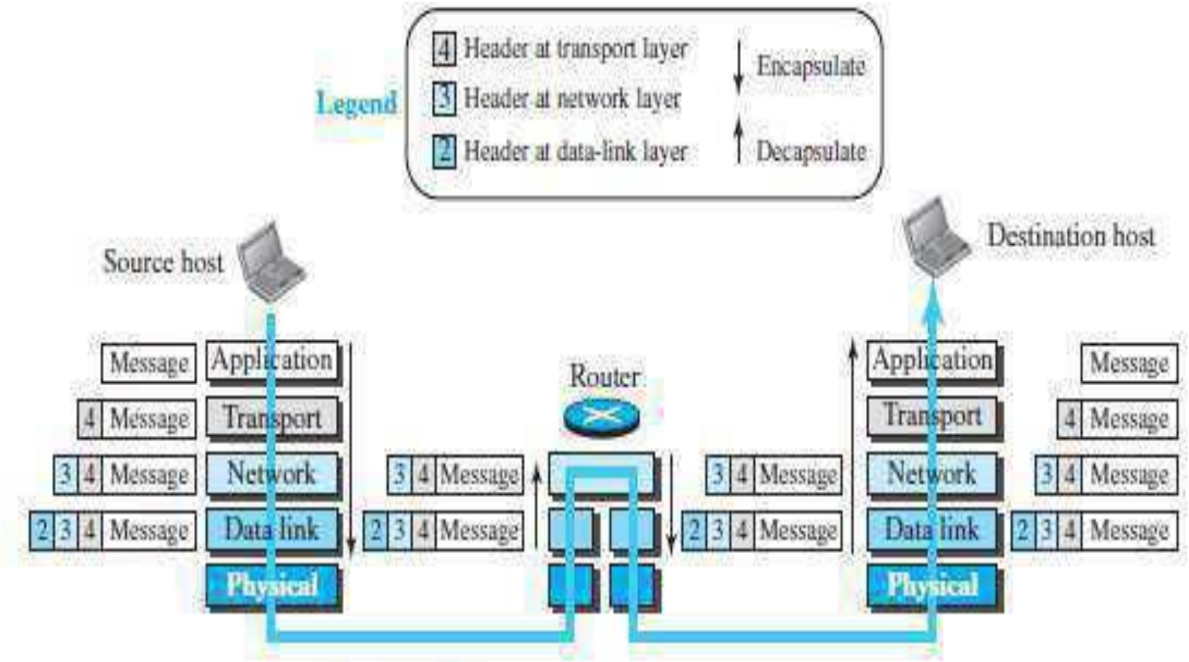


Figure 2.8 *Encapsulation/Decapsulation*

Encapsulation at the Source Host

At the source, we have only encapsulation

1) At the application layer, the data to be exchanged is referred to as a message.

- □ A message normally does not contain any header or trailer.
- □ The message is passed to the transport layer.

2) The transport layer takes the message as the payload.

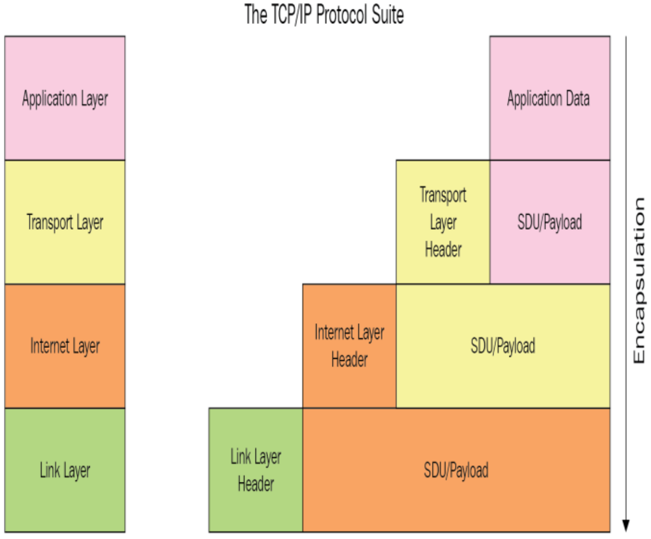
- □ TL adds its own header to the payload.
- □ The header contains
  - → identifiers of the source and destination application programs
  - → information needed for flow, error control, or congestion control.
- □ The transport-layer packet is called the segment (in TCP) and the user datagram (in UDP).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



1. After the set of bits are delivered to the data-link layer, this layer decapsulates the datagram from the frame and passes it to the network layer.

2. The network layer only inspects the source and destination addresses in the datagram
header and consults its forwarding table to find the next hop to which the datagram is to
be delivered. The contents of the datagram should not be changed by the network layer
in the router unless there is a need to fragment the datagram if it is too big to be passed
through the next link. The datagram is then passed to the data-link layer of the next link.

3. The data-link layer of the next link encapsulates the datagram in a frame and passes it to the physical layer for transmission.

Decapsulation at the Destination Host

• At the destination host, each layer
→ decapsulates the packet received from lower layer
→ removes the payload and

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 spec N). □ The
MAC addresses are locally defined ad

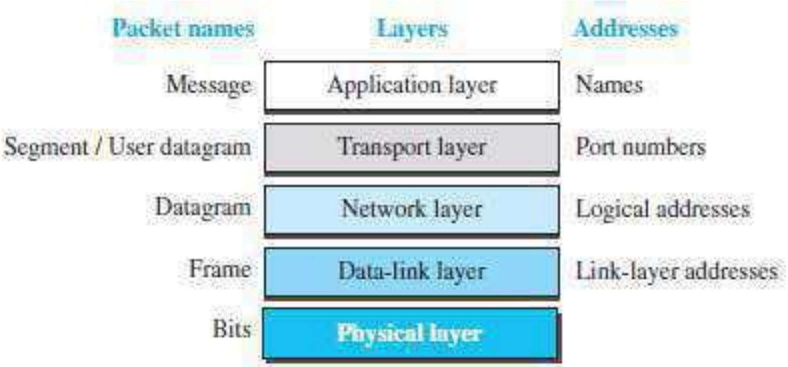


Figure 2.9 *Addressing in the TCP/IP protocol suite*

Multiplexing and Demultiplexing

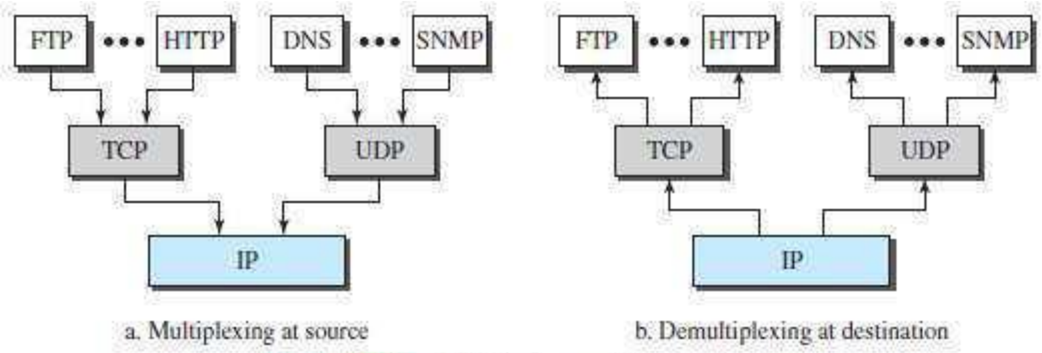
• Multiplexing means a protocol at a layer can encapsulate a packet from several next-higher layer
protocols (one at a time) (Figure 2.10).

• 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*

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.

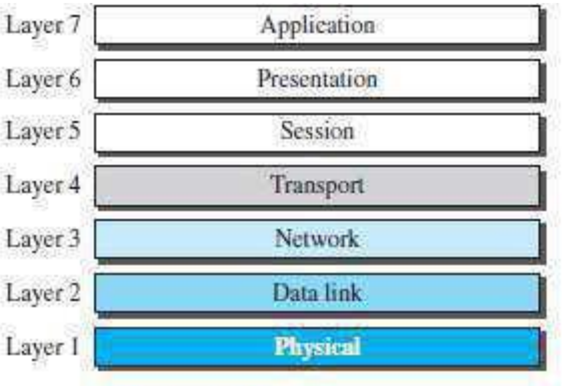


Figure 2.11 *The OSI model*

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..

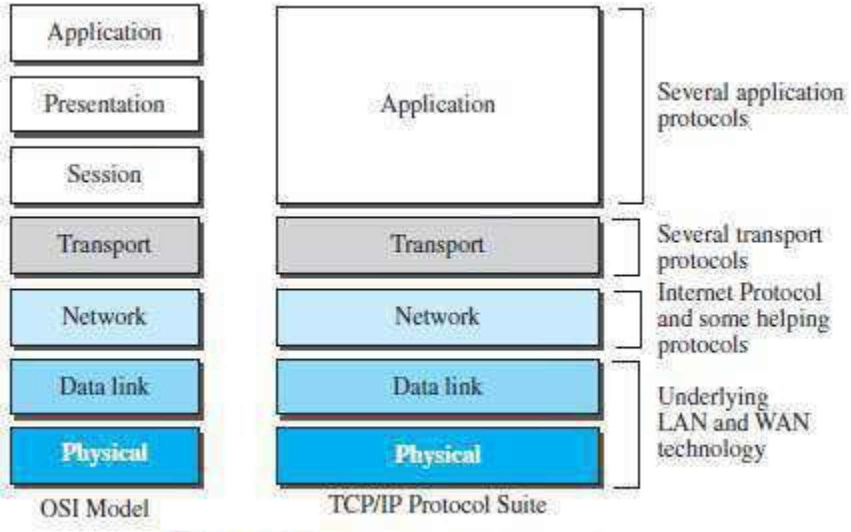


Figure 2.12 *TCP/IP and OSI model*

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

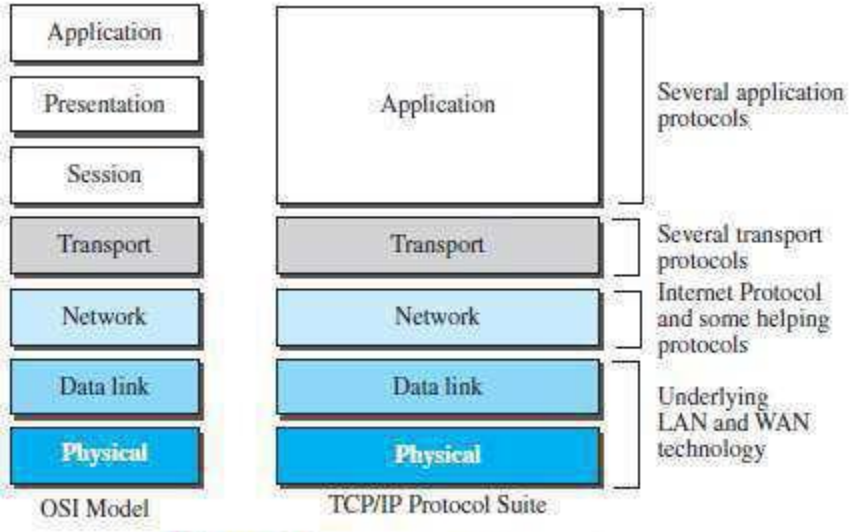


Figure 2.12 *TCP/IP and OSI model*

THE OSI MODEL

Established in 1947, the International Standards Organization (ISO) is a multinational body dedicated to worldwide agreement on international standards. An ISO standard that covers all aspects of network communications is the Open Systems Interconnection (OSI) model. It was first introduced in the late 1970s.

Note

ISO is the organization.
OSI is the model.

Figure 2.2 *Seven layers of the OSI model*

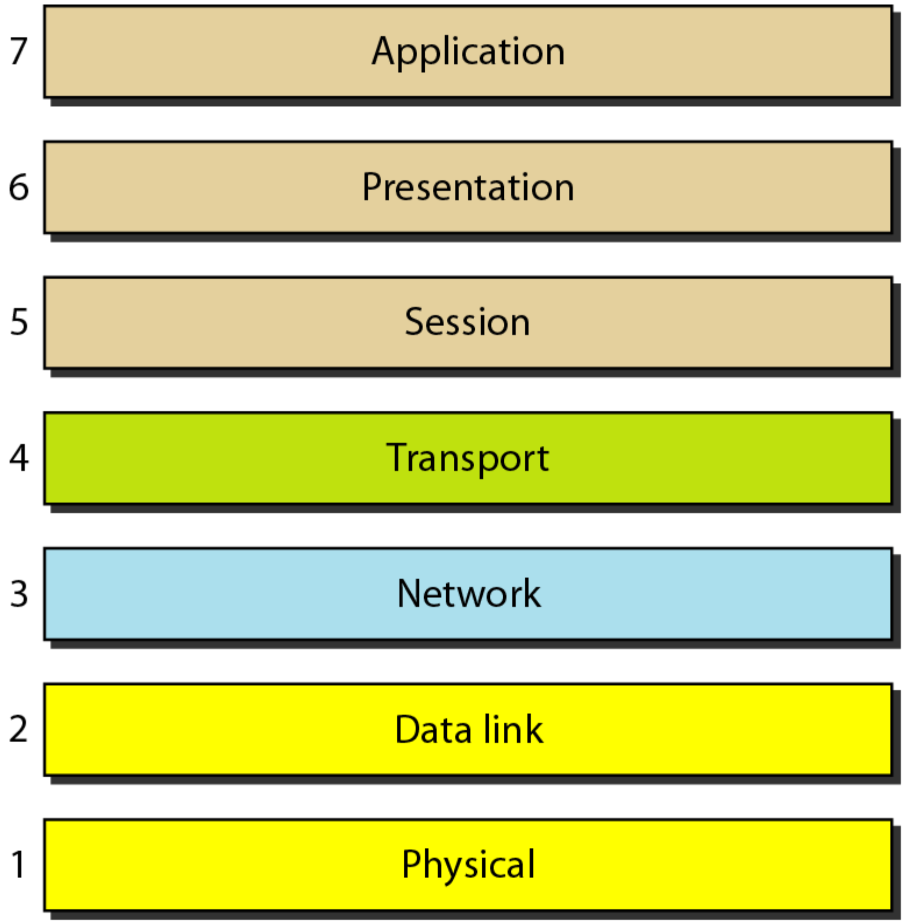


Figure 2.3 *The interaction between layers in the OSI model*

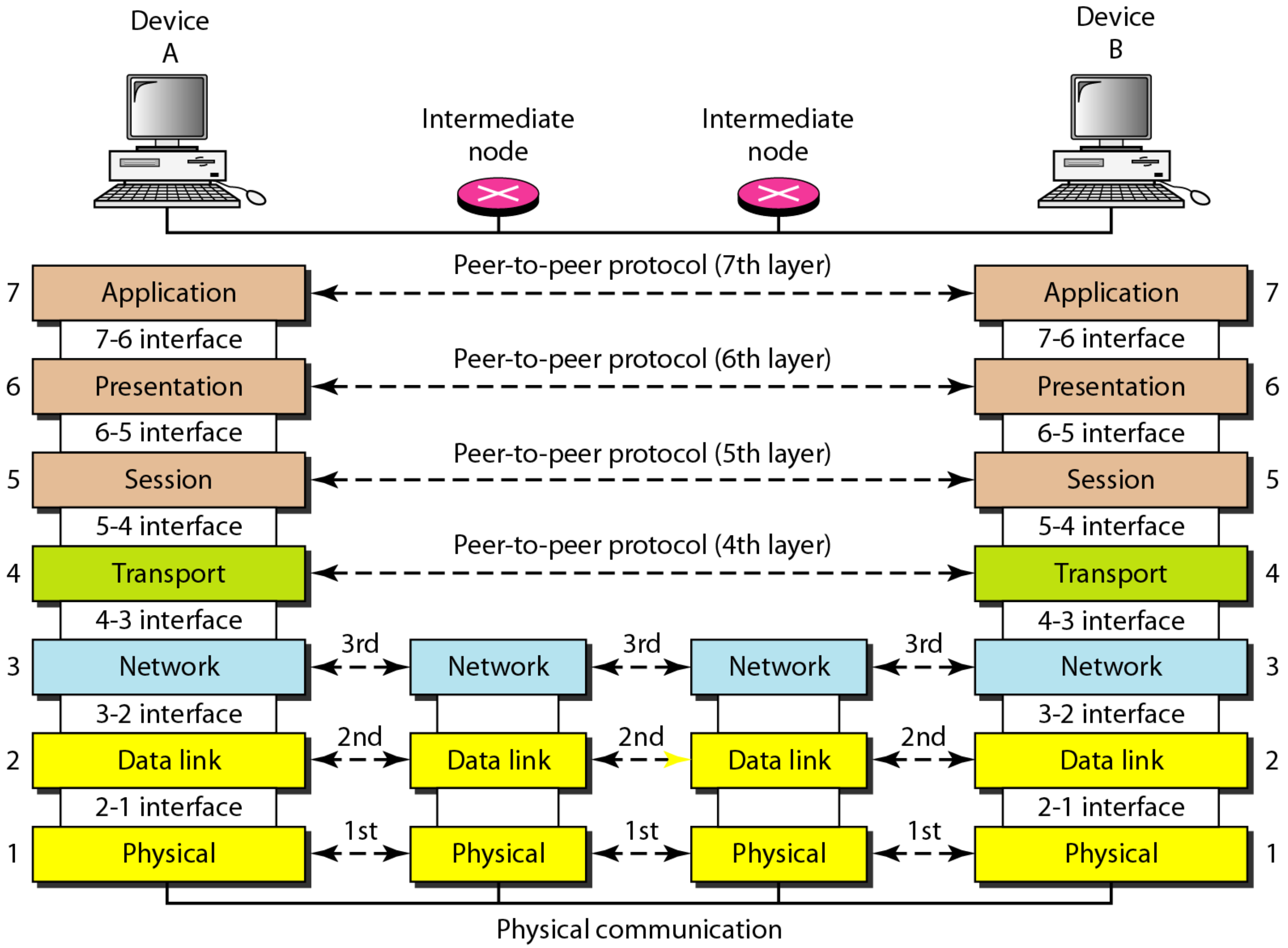
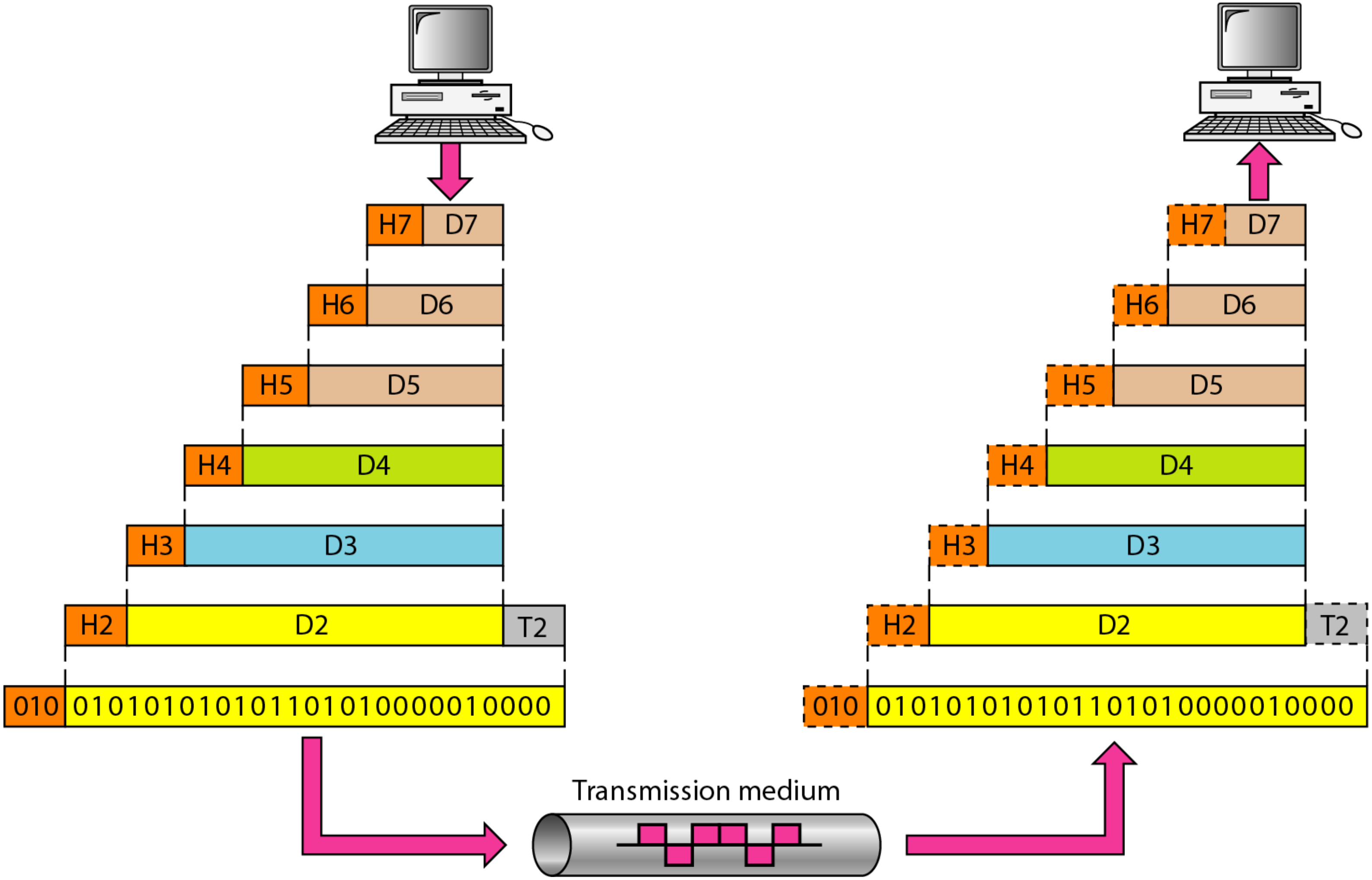


Figure 2.4 *An exchange using the OSI model*



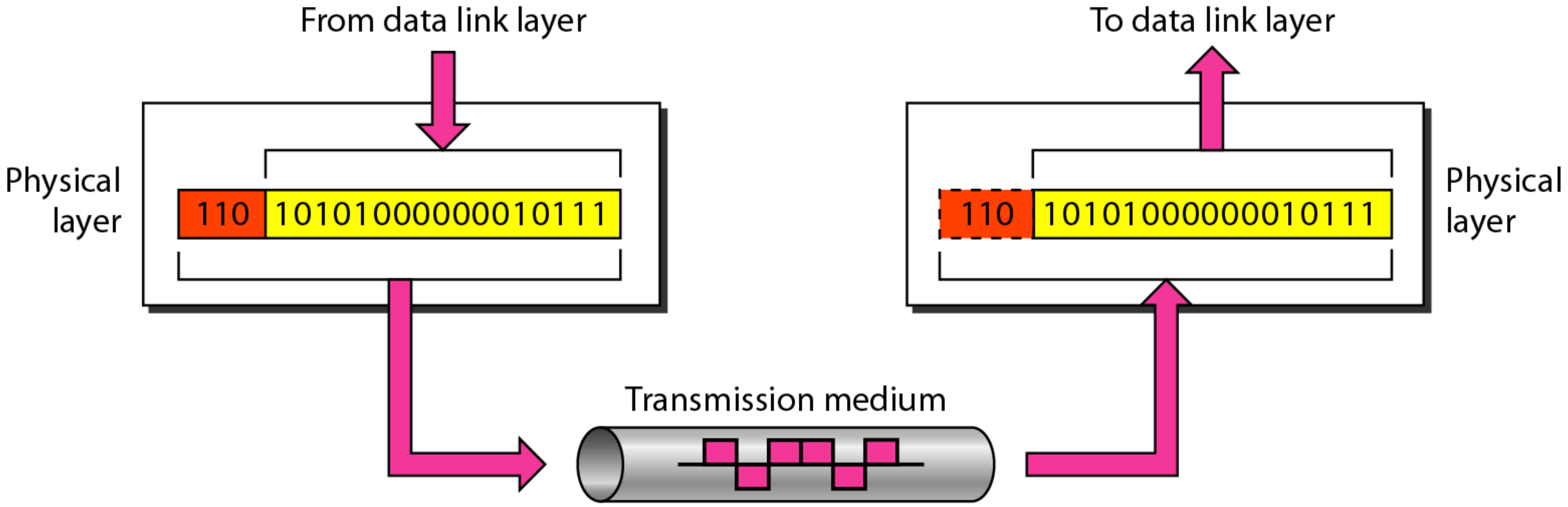
LAYERS IN THE OSI MODEL

In this section we briefly describe the functions of each layer in the OSI model.

Topics discussed in this section:

Physical Layer
Data Link Layer
Network Layer
Transport Layer
Session Layer
Presentation Layer
Application Layer

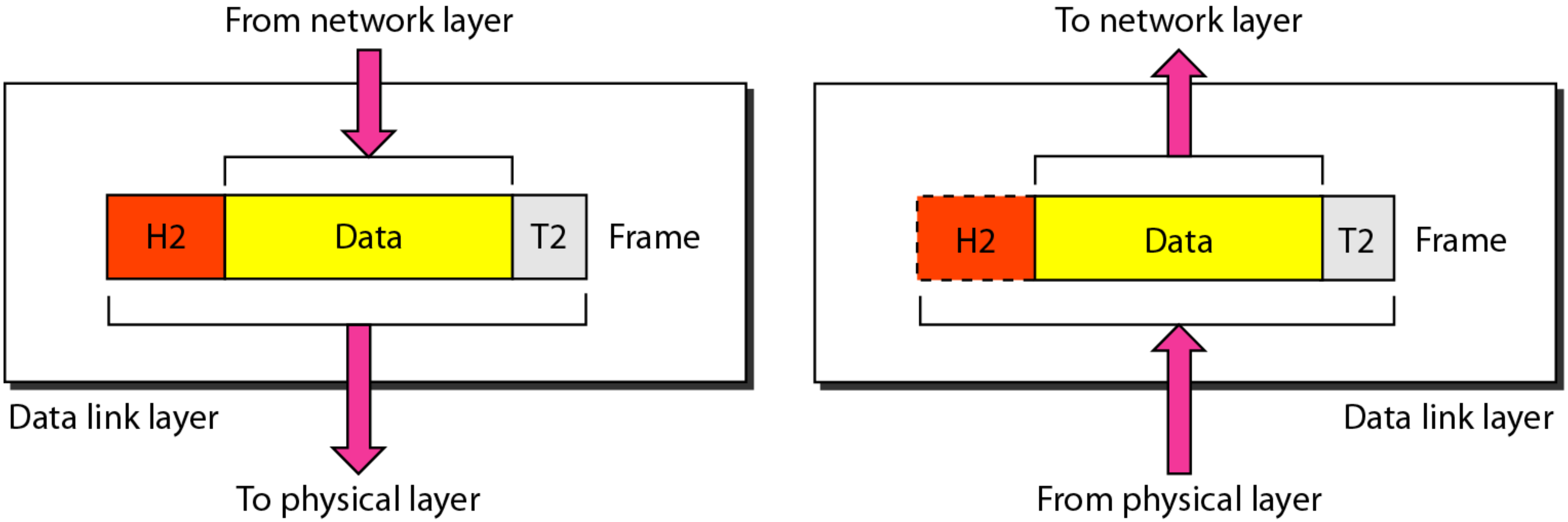
Figure 2.5 *Physical layer*



Note

The physical layer is responsible for movements of individual bits from one hop (node) to the next.

Figure 2.6 *Data link layer*



Note

The data link layer is responsible for moving frames from one hop (node) to the next.

Figure 2.7 *Hop-to-hop delivery*

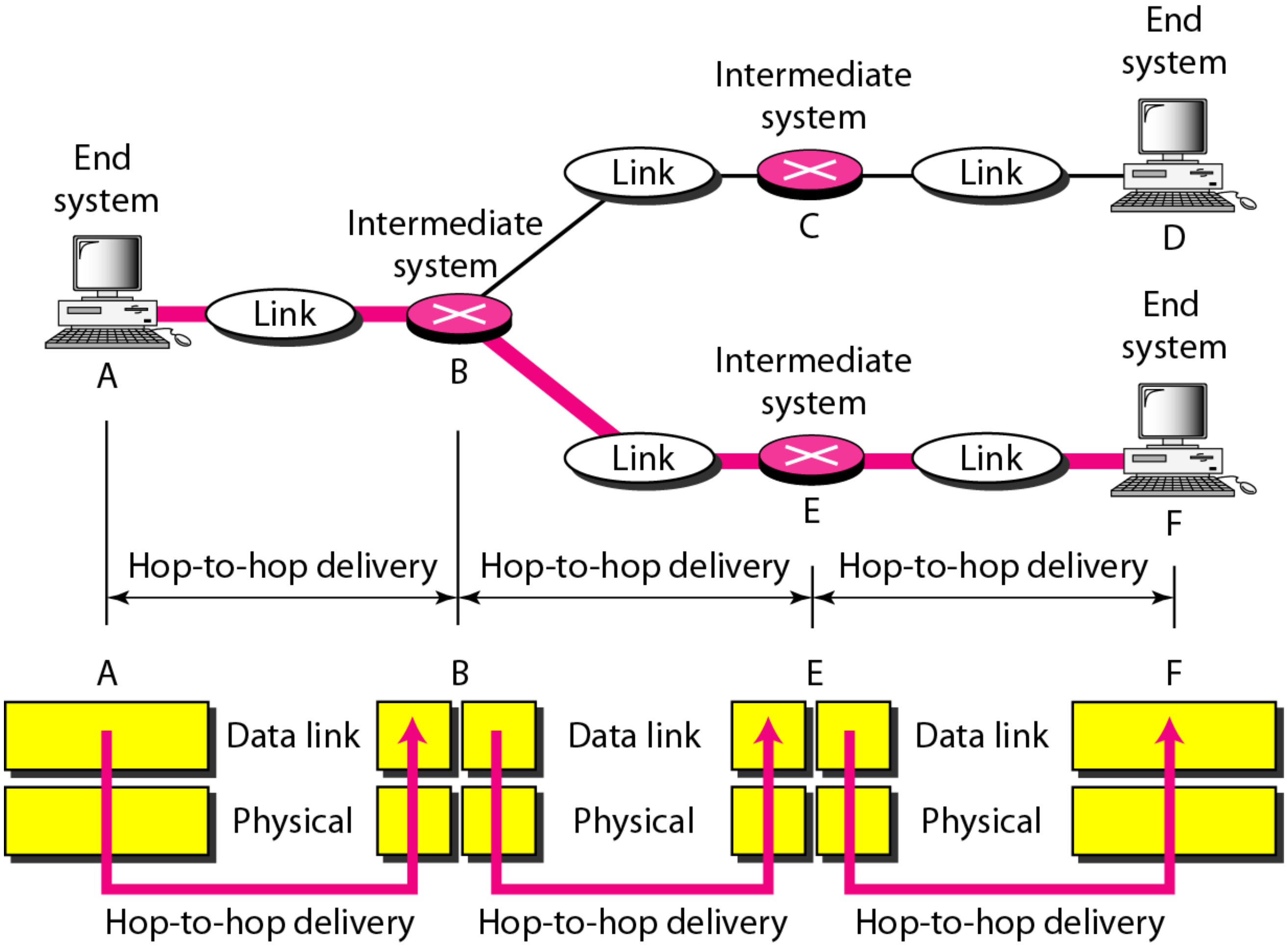
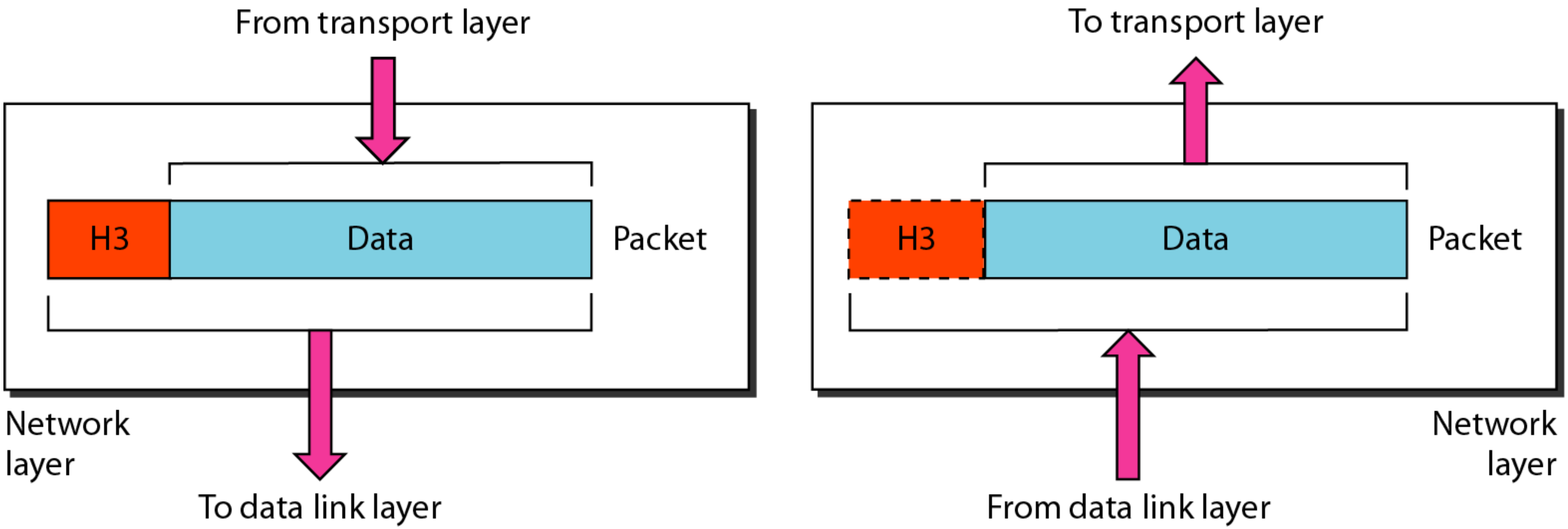


Figure 2.8 *Network layer*



Note

The network layer is responsible for the delivery of individual packets from the source host to the destination host.

Figure 2.9 *Source-to-destination delivery*

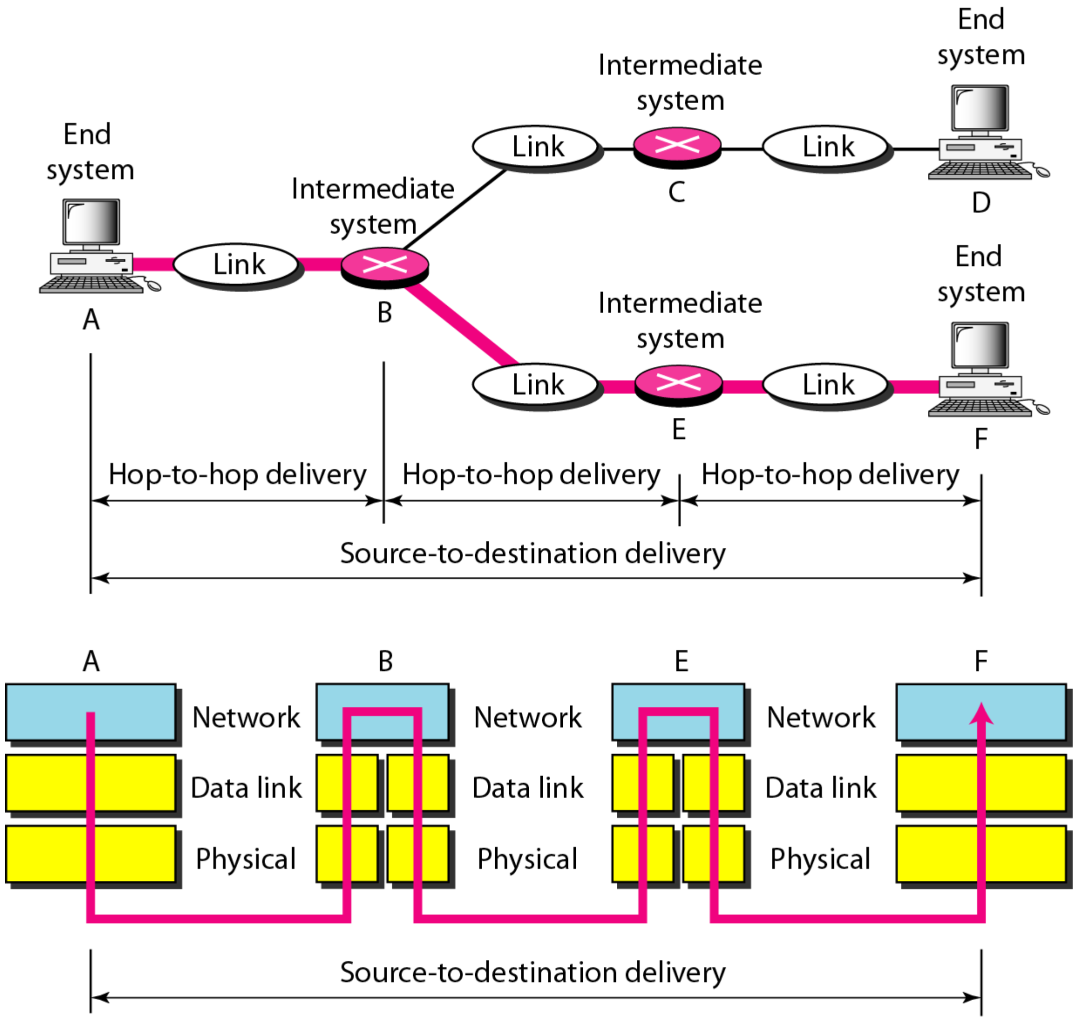
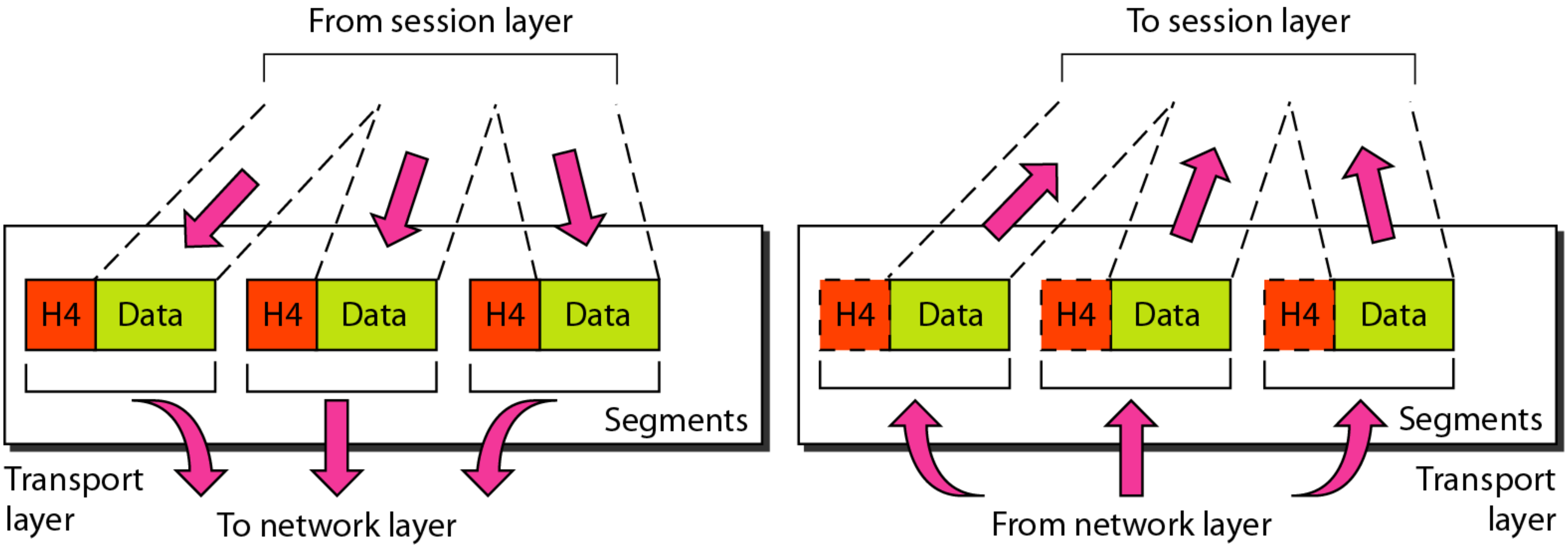


Figure 2.10 *Transport layer*



Note

The transport layer is responsible for the delivery of a message from one process to another.

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

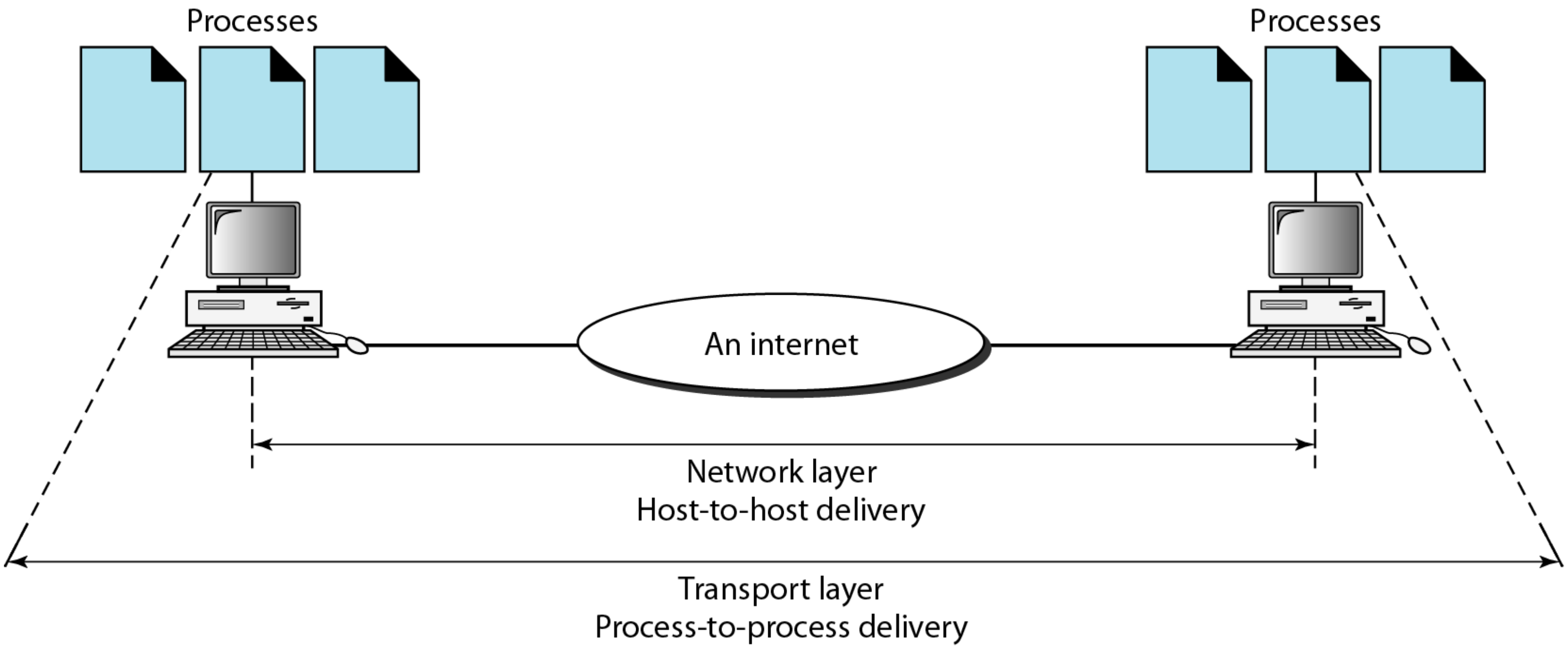
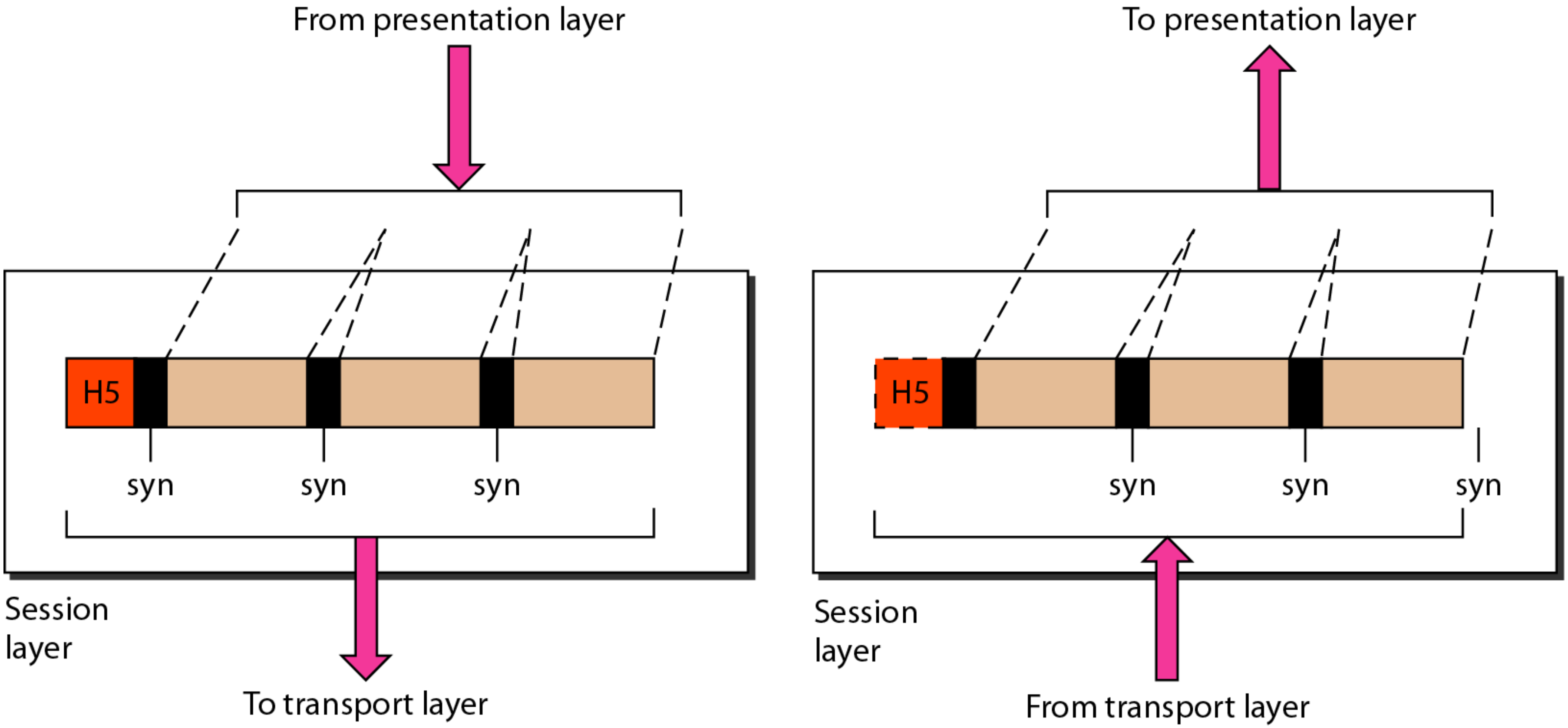


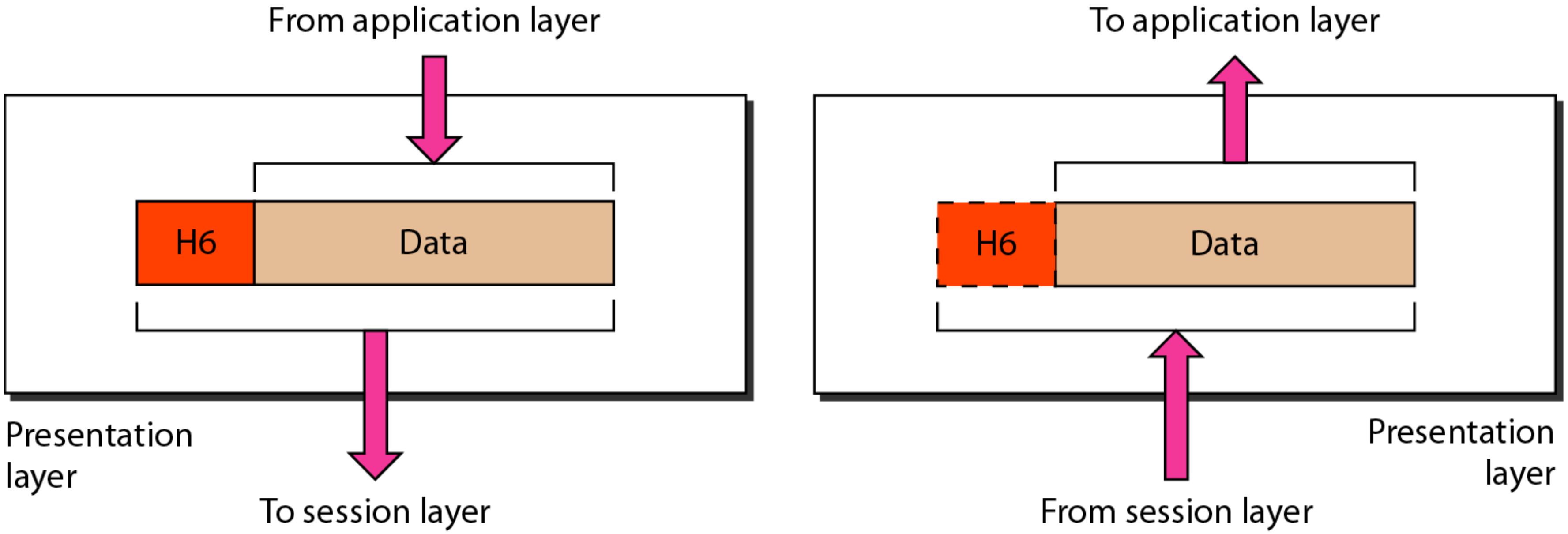
Figure 2.12 *Session layer*



Note

The session layer is responsible for dialog control and synchronization.

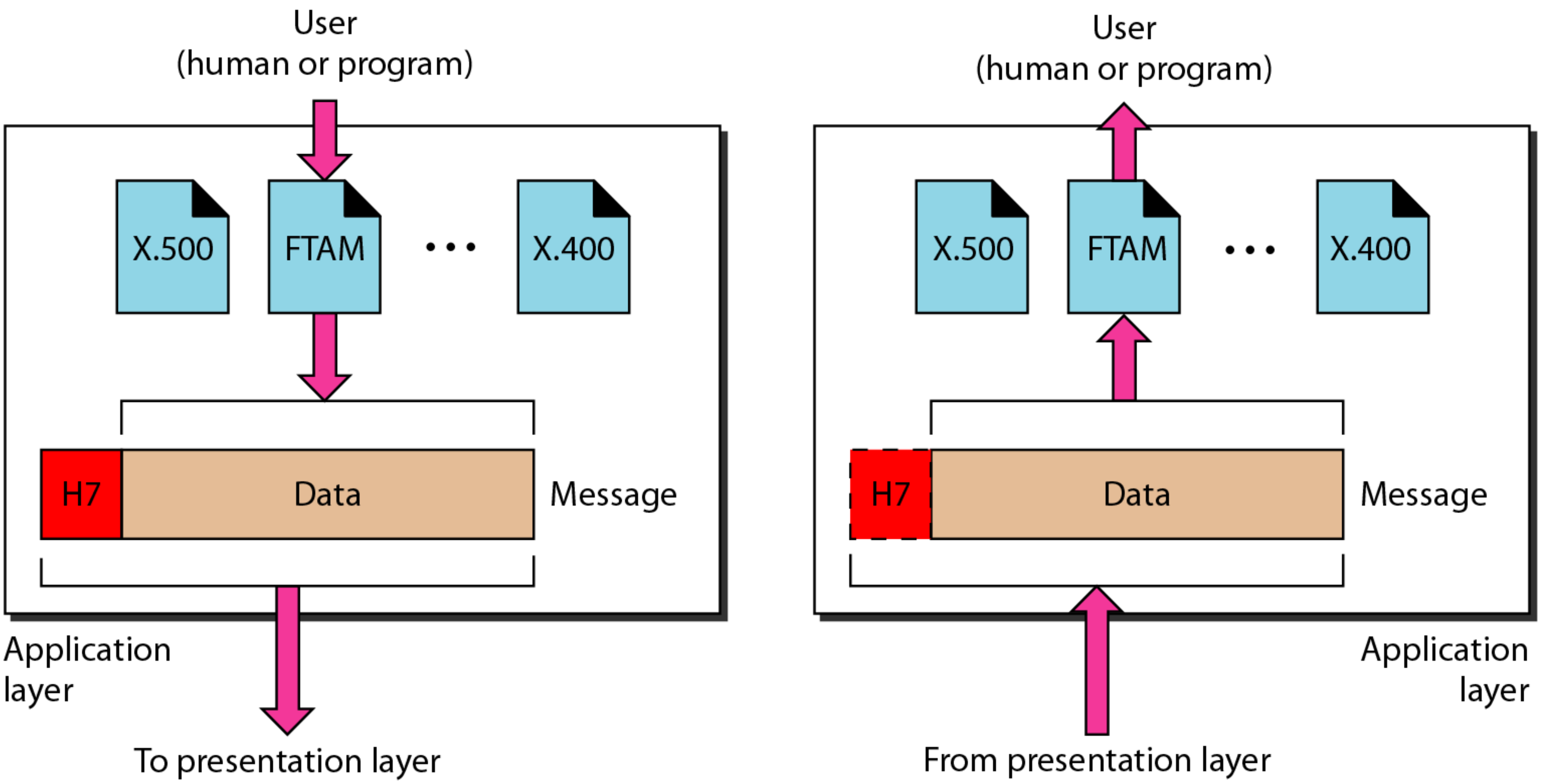
Figure 2.13 *Presentation layer*



Note

The presentation layer is responsible for translation, compression, and encryption.

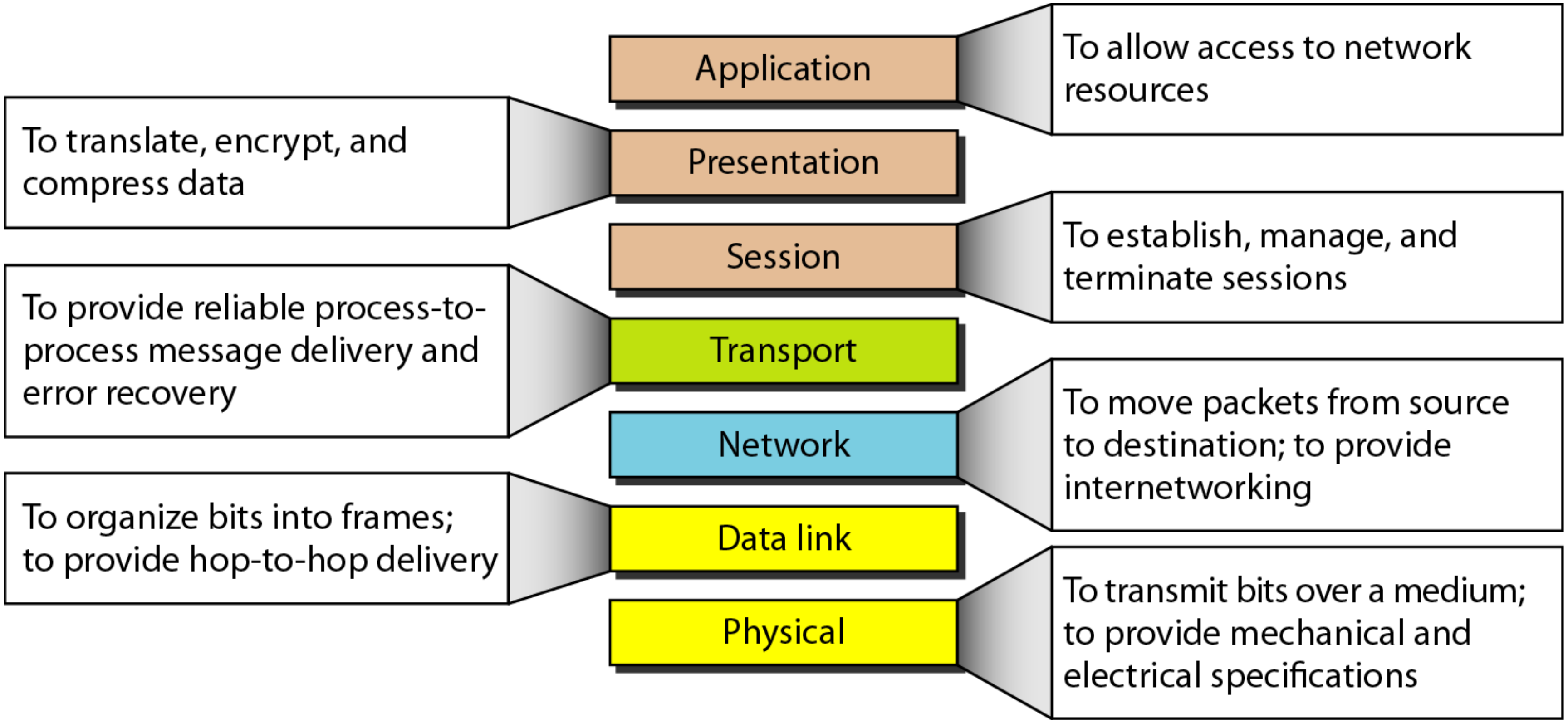
Figure 2.14 *Application layer*



Note

The application layer is responsible for providing services to the user.

Figure 2.15 *Summary of layers*



Application

Presentation

Session

Transport

Network

Data link

Physical

To allow access to network resources

To translate, encrypt, and compress data

To establish, manage, and terminate sessions

To provide reliable process-to-process message delivery and error recovery

To move packets from source to destination; to provide internetworking

To organize bits into frames; to provide hop-to-hop delivery

To transmit bits over a medium; to provide mechanical and electrical specifications

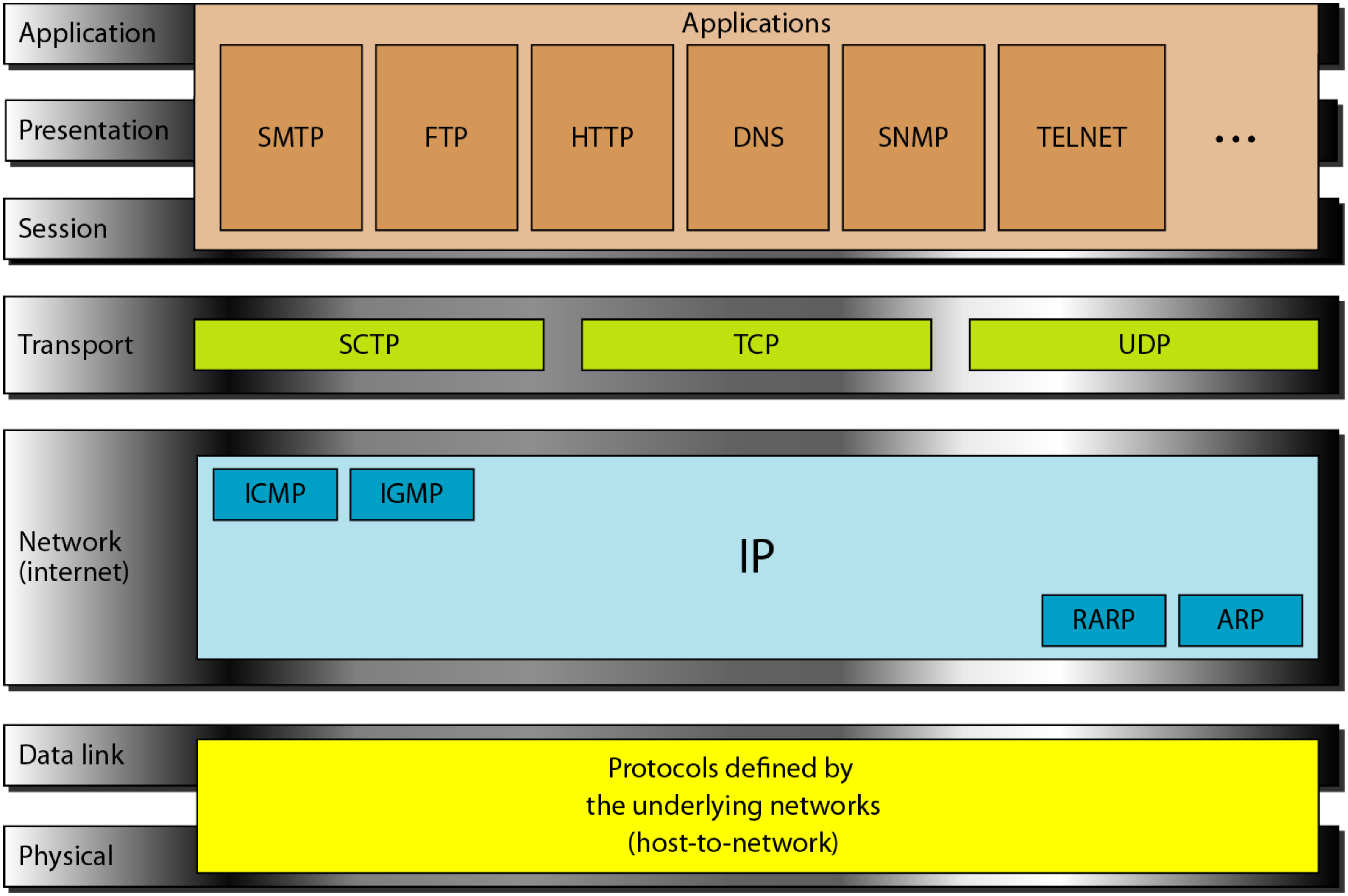
TCP/IP PROTOCOL SUITE

The layers in the TCP/IP protocol suite do not exactly match those in the OSI model. The original TCP/IP protocol suite was defined as having four layers: host-to-network, internet, transport, and application. However, when TCP/IP is compared to OSI, we can say that the TCP/IP protocol suite is made of five layers: physical, data link, network, transport, and application.

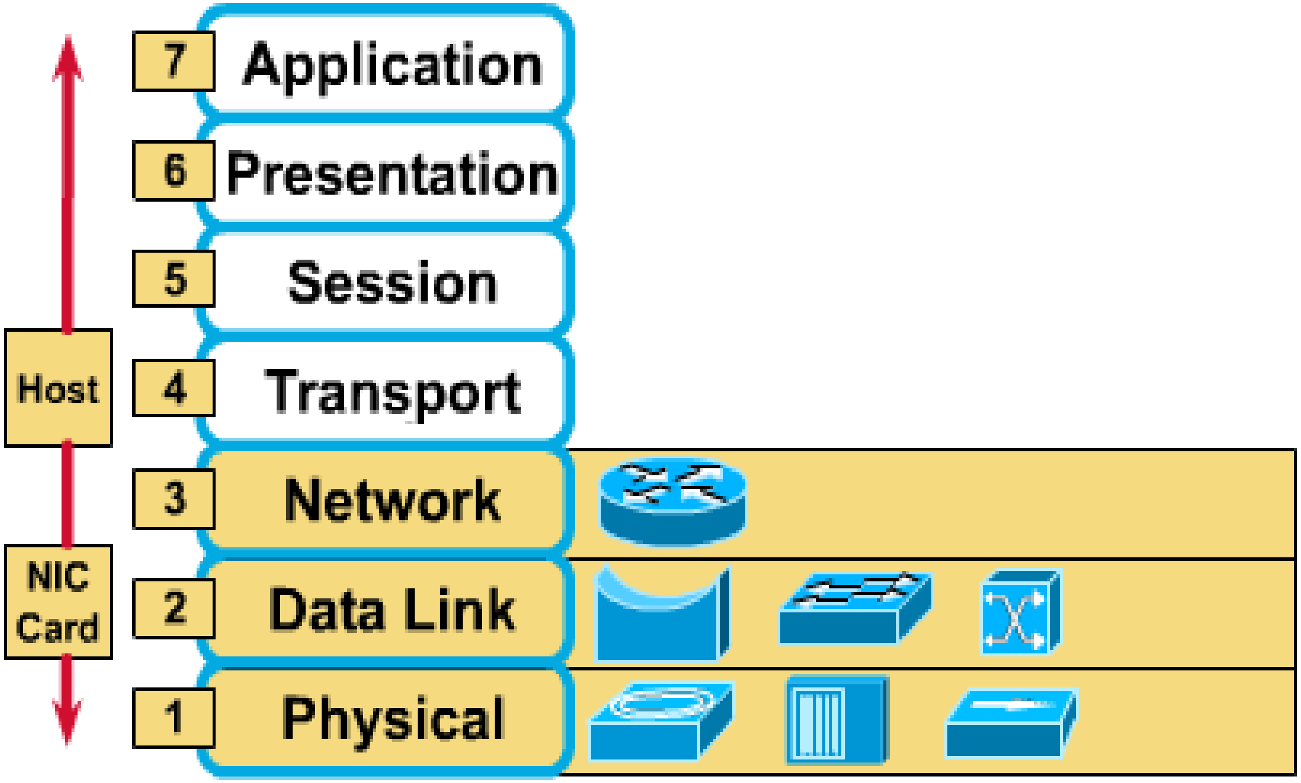
Topics discussed in this section:

Physical and Data Link Layers
Network Layer
Transport Layer
Application Layer

Figure 2.16 *TCP/IP and OSI model*



Network Devices



LAYER 1 - Transceiver, Repeater, & Hub.

LAYER 2 – Bridge, Switch, & NIC.

LAYER 3 - Router

ADDRESSING

Four levels of addresses are used in an internet employing the TCP/IP protocols: physical, logical, port, and specific.

Topics discussed in this section:

Physical Addresses
Logical Addresses
Port Addresses
Specific Addresses

Figure 2.17 *Addresses in TCP/IP*

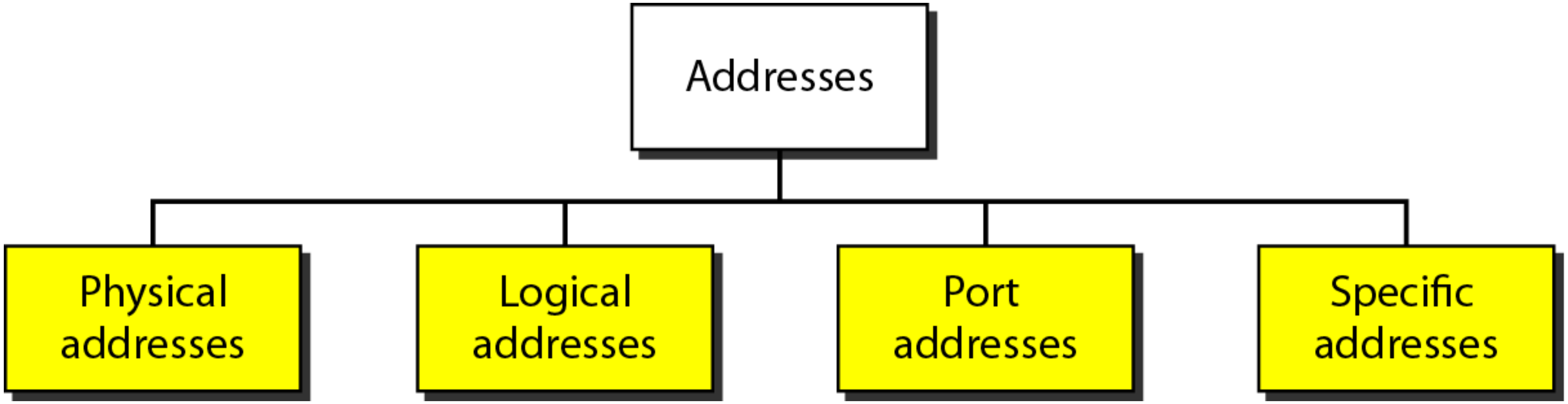
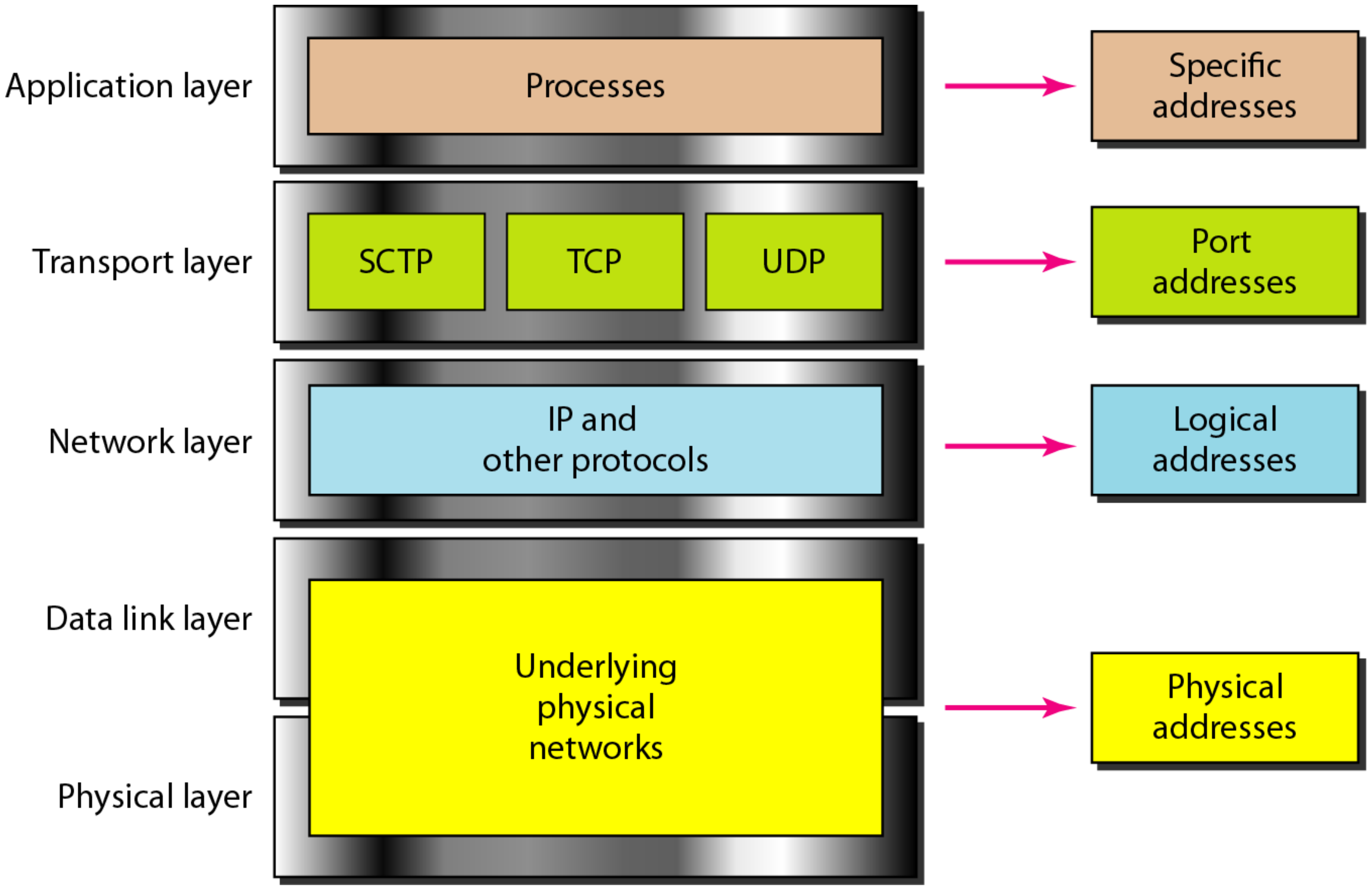


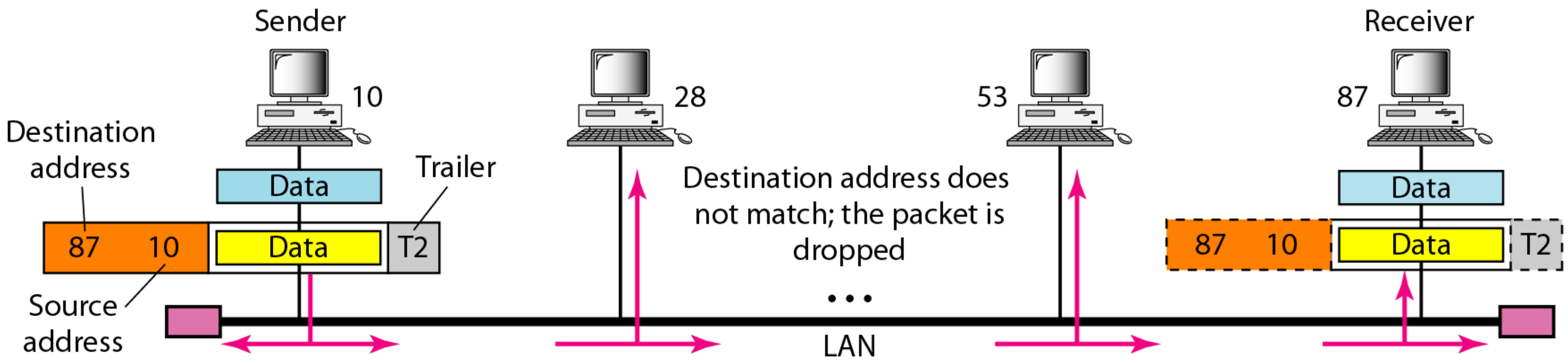
Figure 2.18 *Relationship of layers and addresses in TCP/IP*



Example 2.1

In Figure 2.19 a node with physical address 10 sends a frame to a node with physical address 87. The two nodes are connected by a link (bus topology LAN). As the figure shows, the computer with physical address 10 is the sender, and the computer with physical address 87 is the receiver.

Figure 2.19 *Physical addresses*



Example 2.2

Most local-area networks use a 48-bit (6-byte) physical address written as 12 hexadecimal digits; every byte (2 hexadecimal digits) is separated by a colon, as shown below:

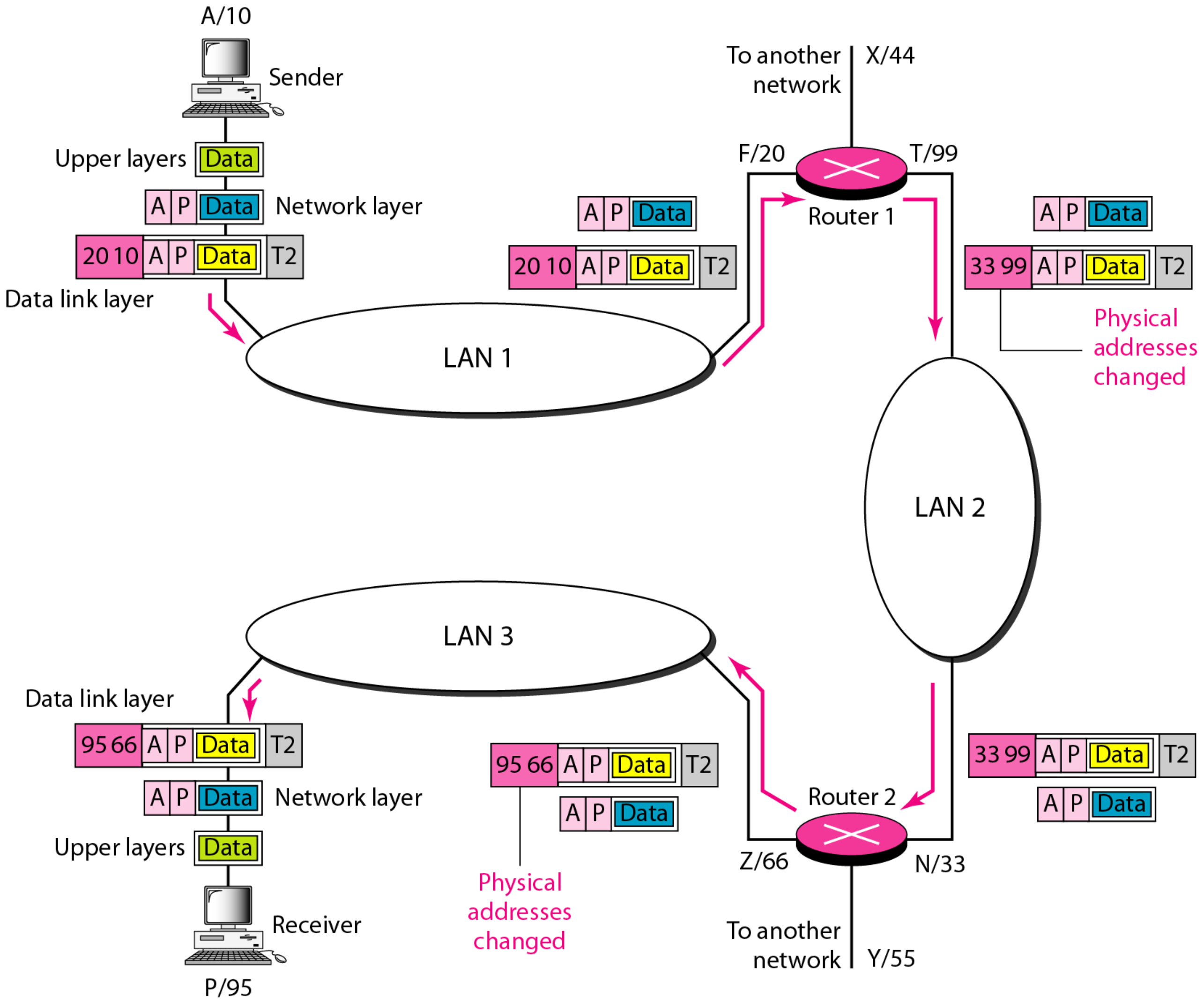
07:01:02:01:2C:4B

A 6-byte (12 hexadecimal digits) physical address.

Example 2.3

Figure 2.20 shows a part of an internet with two routers connecting three LANs. Each device (computer or router) has a pair of addresses (logical and physical) for each connection. In this case, each computer is connected to only one link and therefore has only one pair of addresses. Each router, however, is connected to three networks (only two are shown in the figure). So each router has three pairs of addresses, one for each connection.

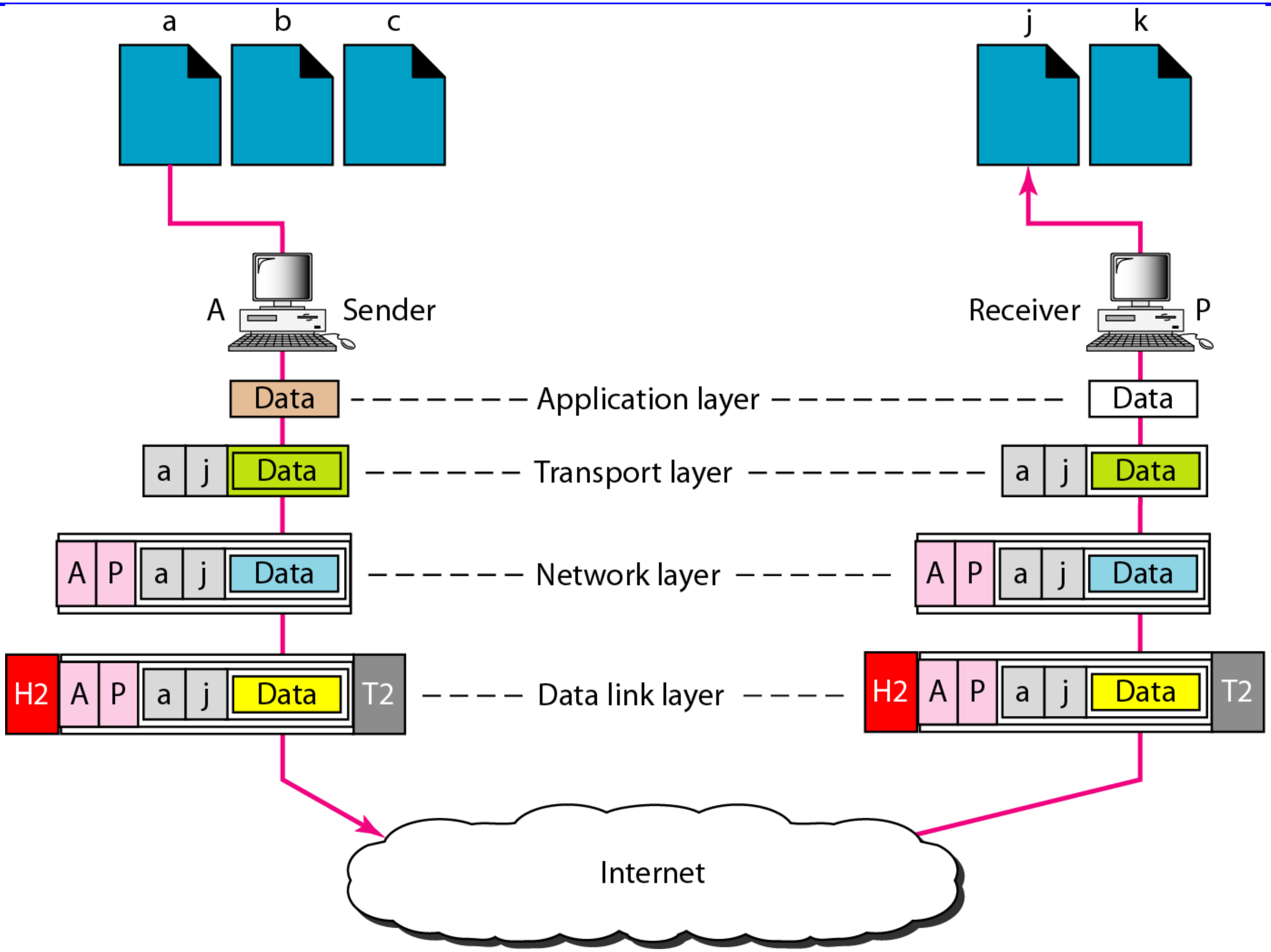
Figure 2.20 *IP addresses*



Example 2.4

Figure 2.21 shows two computers communicating via the Internet. The sending computer is running three processes at this time with port addresses a, b, and c. The receiving computer is running two processes at this time with port addresses j and k. Process a in the sending computer needs to communicate with process j in the receiving computer. Note that although physical addresses change from hop to hop, logical and port addresses remain the same from the source to destination.

Figure 2.21 *Port addresses*



Note

The physical addresses will change from hop to hop, but the logical addresses usually remain the same.

Example 2.5

A port address is a 16-bit address represented by one decimal number as shown.

753

A 16-bit port address represented as one single number.

Addressing

Four levels of addressing are used in an internet.

1) **Physical address** – It is used to identify the host on the network. It is a 48 bit size, represented in hexadecimal. It is a permanent address printed on the NIC.

   Example – 07:01:02:01:2C:4B

2) **Logical address** – It is also called as IP address or Network address. It is a 32 bits size. It is represented in 4 decimal dots. Example- 10. 25. 26.45

3) **Port address** - It defines a process running on a host. It is a 16 bits decimal representation.
   Example: TELNET -23, FTP-21, SMTP- 25, DNS – 53.

4) **specific address** – User friendly address are known as specific addresses. It is also called as URL (Universal Resource locator) or Domain Name System (DSN) address.

   Example – www.hotmail.com

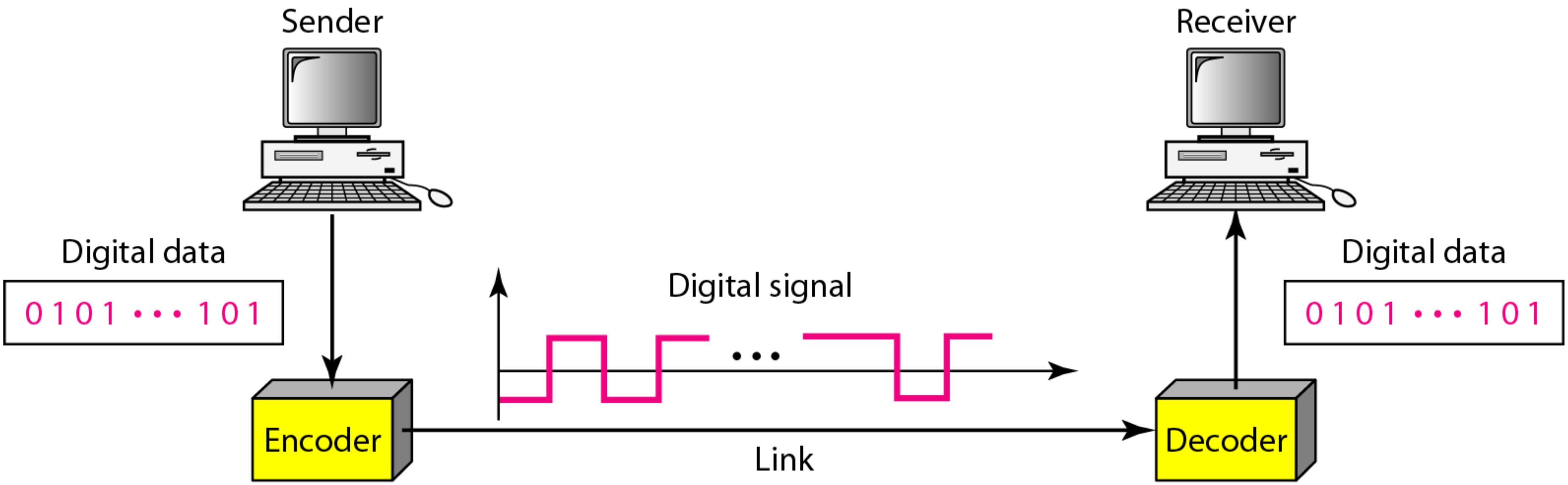
MODULE – 2

DIGITAL TRANSMISSION

Line Coding

- Converting a string of 1’s and 0’s (digital data) into a sequence of signals that denote the 1’s and 0’s.
- For example a high voltage level (+V) could represent a “1” and a low voltage level (0 or -V) could represent a “0”.

Figure 4.1 *Line coding and decoding*



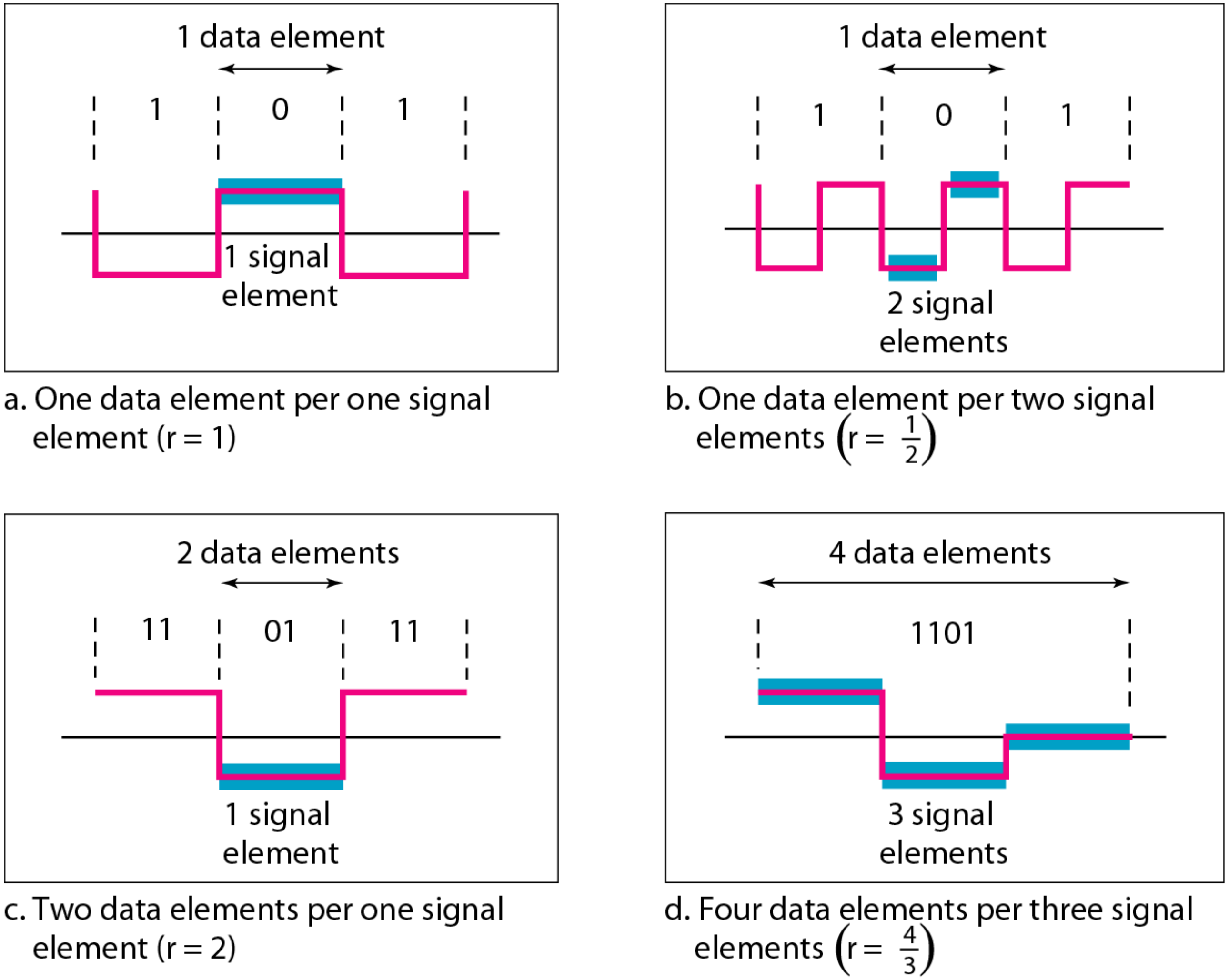
Mapping Data symbols onto Signal levels

- A data symbol (or element) can consist of a number of data bits:
  - 1 , 0 or
  - 11, 10, 01, ……
- A data symbol can be coded into a single signal element or multiple signal elements
  - 1 -> +V, 0 -> -V
  - 1 -> +V and -V, 0 -> -V and +V
- The ratio ‘r’ is the number of data elements carried by a signal element.

Relationship between data rate and signal rate

- The data rate defines the number of bits sent per sec - bps. It is often referred to the bit rate.
- The signal rate is the number of signal elements sent in a second and is measured in bauds. It is also referred to as the modulation rate OR baud rate.
- Goal is to increase the data rate whilst reducing the baud rate.

Figure 4.2 *Signal element versus data element*



a. One data element per one signal element (r = 1)

b. One data element per two signal elements $\left(r = \frac{1}{2}\right)$

c. Two data elements per one signal element (r = 2)

d. Four data elements per three signal elements $\left(r = \frac{4}{3}\right)$

Data rate and Baud rate

- The baud or signal rate can be expressed as:

$S = c \times N \times 1/r$ bauds

where N is data rate

c is the case factor (worst, best & avg.)

r is the ratio between data element & signal element

Example 4.1

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

$$\boxed{S = c \times N \times \frac{1}{r} = \frac{1}{2} \times 100{,}000 \times \frac{1}{1} = 50{,}000 = 50 \text{ kbaud}}$$

Example 4.2

The maximum data rate of a channel (see Chapter 3) is Nmax = 2 × B × $\log_2$ L (defined by the Nyquist formula). Does this agree with the previous formula for N_{max}?

Solution

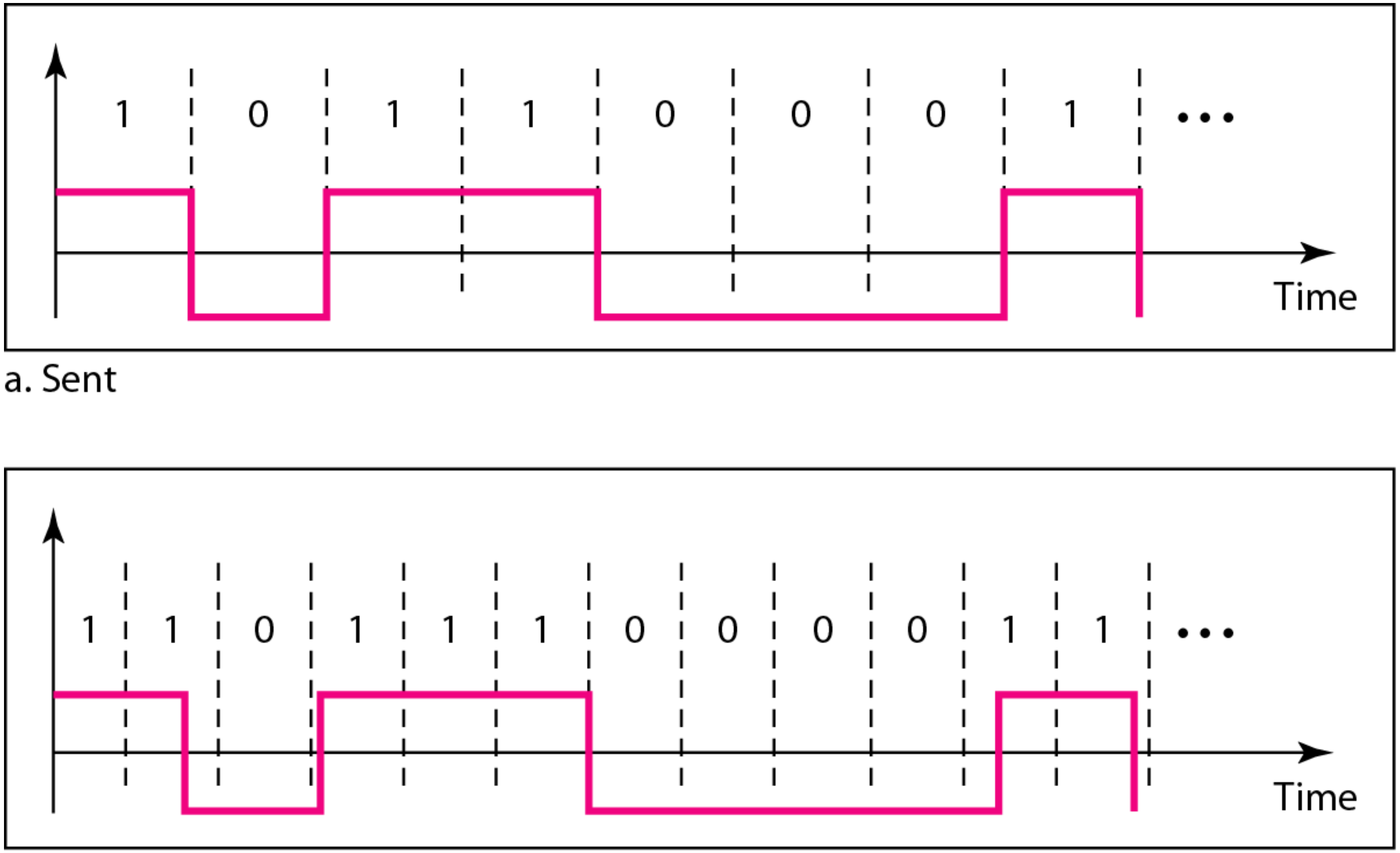
A signal with L levels actually can carry $\log_2 L$ bits per level. If each level corresponds to one signal element and we assume the average case (c = 1/2), then we have

$$\boxed{N_{\max} = \frac{1}{c} \times B \times r = 2 \times B \times \log_2 L}$$

Considerations for choosing a good signal element referred to as line encoding

- Self synchronization - the clocks at the sender and the receiver must have the same bit interval.
- If the receiver clock is faster or slower it will misinterpret the incoming bit stream.

Figure 4.3 *Effect of lack of synchronization*



a. Sent

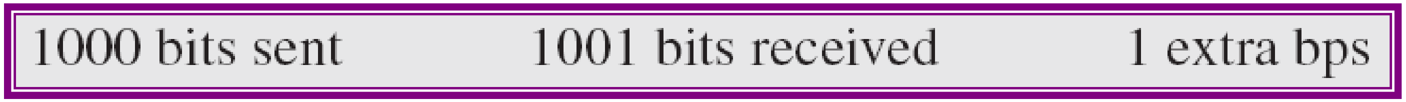
b. Received

Example 4.3

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?

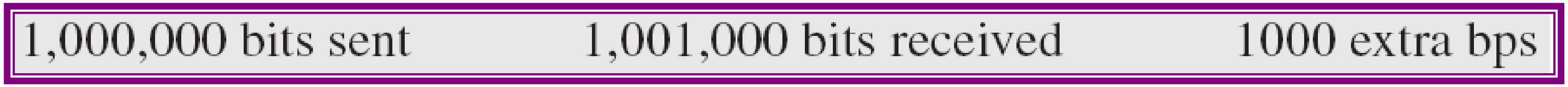
Solution

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

Line encoding C/Cs

- Error detection - errors occur during transmission due to line impairments.
- Some codes are constructed such that when an error occurs it can be detected.

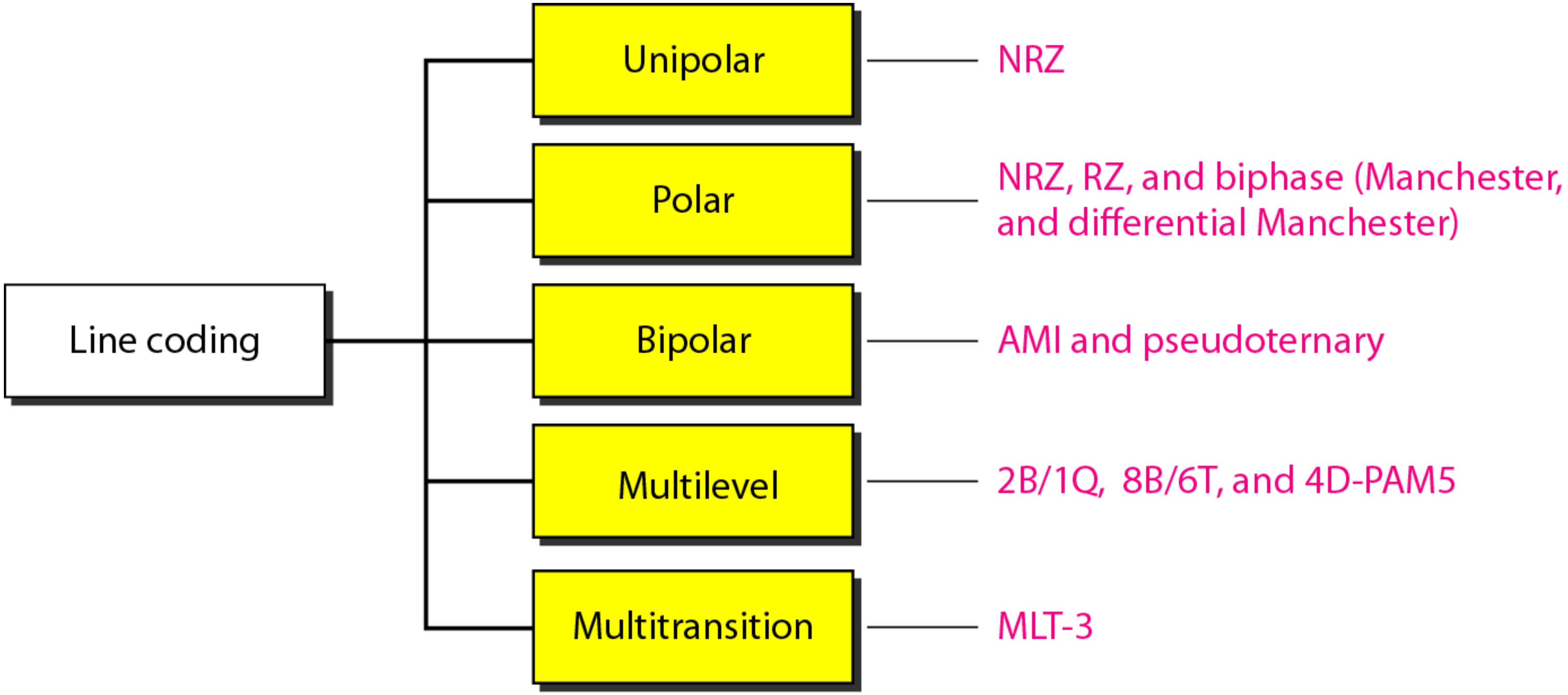
Line encoding C/Cs

- Noise and interference - there are line encoding techniques that make the transmitted signal “immune” to noise and interference.
- This means that the signal cannot be corrupted, it is stronger than error detection.

Line encoding C/Cs

- Complexity - the more robust and resilient the code, the more complex it is to implement and the price is often paid in baud rate or required bandwidth.

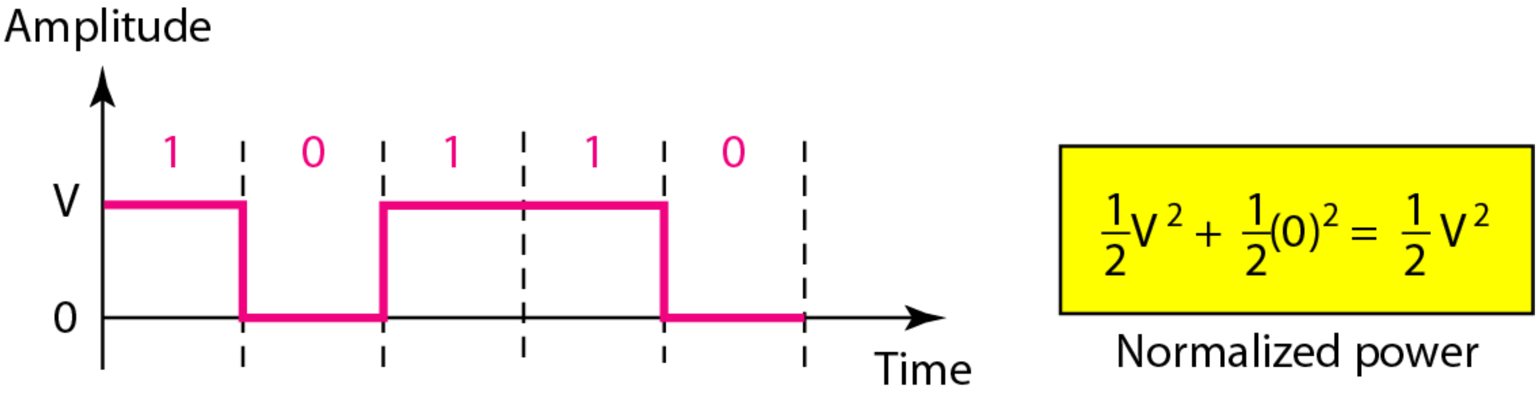
Figure 4.4 *Line coding schemes*


Line coding
Unipolar
NRZ
Polar
NRZ, RZ, and biphase (Manchester, and differential Manchester)
Bipolar
AMI and pseudoternary
Multilevel
2B/1Q, 8B/6T, and 4D-PAM5
Multitransition
MLT-3

Unipolar

- All signal levels are on one side of the time axis - either above or below
- NRZ - Non Return to Zero scheme is an example of this code. The signal level does not return to zero during a symbol transmission.
- Scheme is prone to baseline wandering and DC components. It has no synchronization or any error detection. It is simple but costly in power consumption.

Figure 4.5 *Unipolar NRZ scheme*

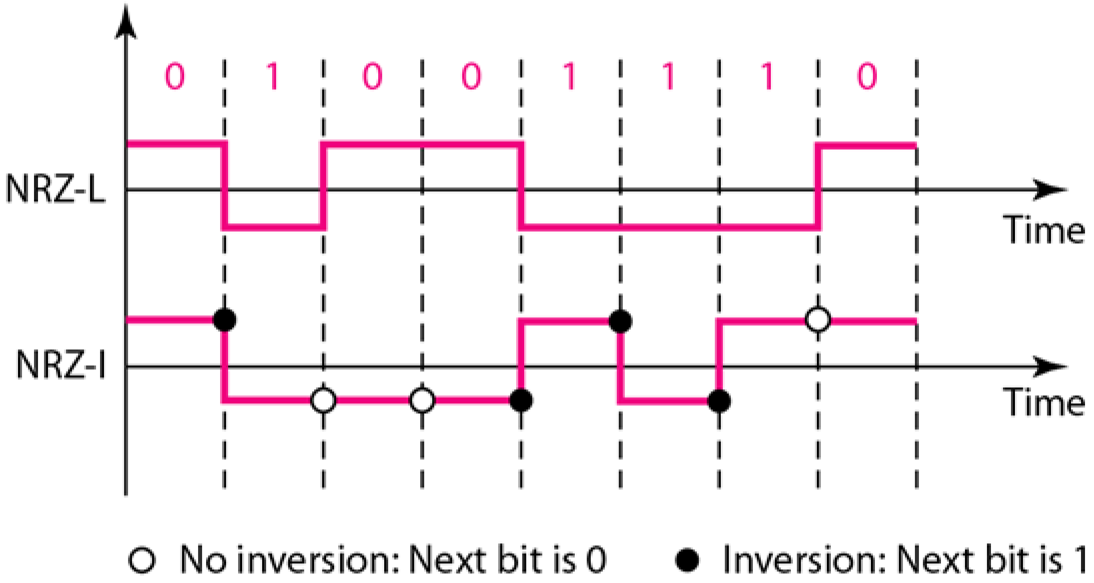


$$\frac{1}{2}V^2 + \frac{1}{2}(0)^2 = \frac{1}{2}V^2$$

Polar - NRZ

- The voltages are on both sides of the time axis.
- Polar NRZ scheme can be implemented with two voltages. E.g. +V for 1 and -V for 0.
- There are two versions:
  - NZR - Level (NRZ-L) - positive voltage for one symbol and negative for the other
  - NRZ - Inversion (NRZ-I) - the change or lack of change in polarity determines the value of a symbol. E.g. a “1” symbol inverts the polarity a “0” does not.

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



Note

In NRZ-L the level of the voltage determines the value of the bit.
In NRZ-I the inversion
or the lack of inversion
determines the value of the bit.

Note

NRZ-L and NRZ-I both have a DC component problem and baseline wandering, it is worse for NRZ-L. Both have no self synchronization &no error detection. Both are relatively simple to implement.

Example 4.4

A system is using NRZ-I to transfer 1-Mbps data. What are the average signal rate and minimum bandwidth?

Solution

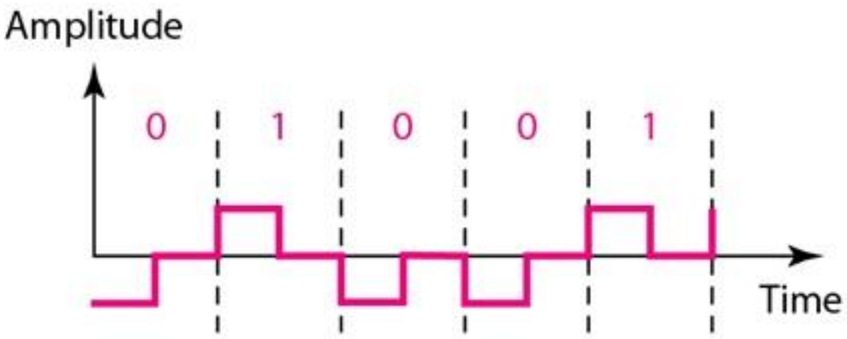
The average signal rate is S= c x N x R = 1/2 x N x 1 = 500 kbaud. The minimum bandwidth for this average baud rate is Bmin = S = 500 kHz.

Note c = 1/2 for the avg. case as worst case is 1 and best case is 0

Polar - RZ

- The Return to Zero (RZ) scheme uses three voltage values. +, 0, -.
- Each symbol has a transition in the middle. Either from high to zero or from low to zero.
- This scheme has more signal transitions (two per symbol) and therefore requires a wider bandwidth.
- No DC components or baseline wandering.
- Self synchronization - transition indicates symbol value.
- More complex as it uses three voltage level. It has no error detection capability.

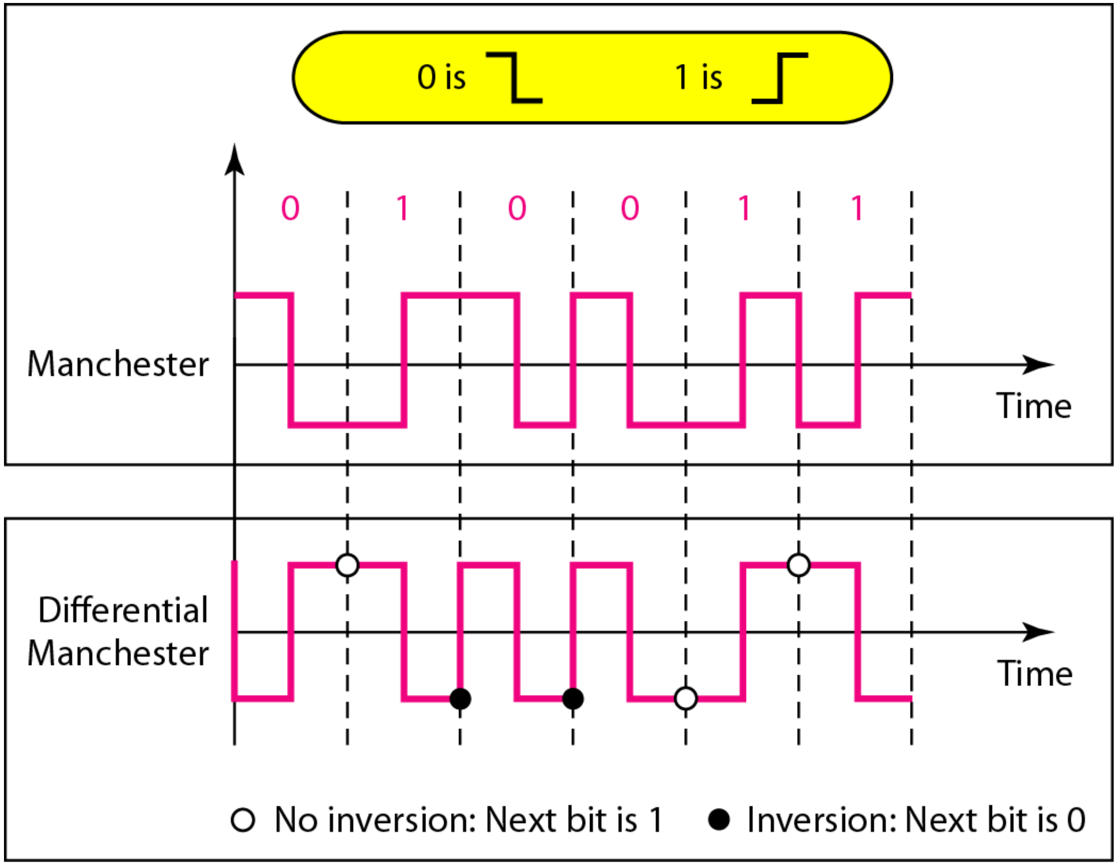
Figure 4.7 *Polar RZ scheme*

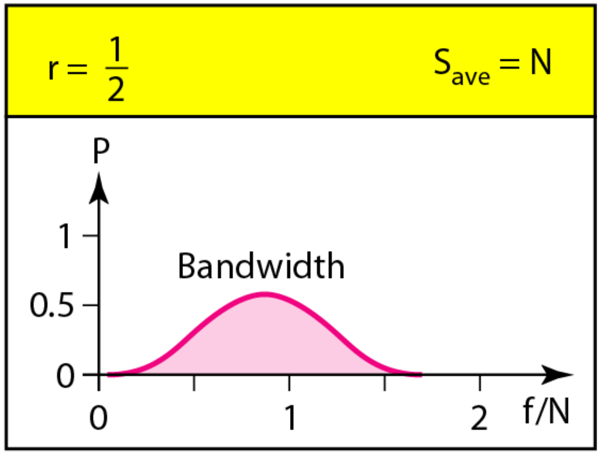


Polar - Biphase: Manchester and Differential Manchester

- Manchester coding consists of combining the NRZ-L and RZ schemes.
  - Every symbol has a level transition in the middle: from high to low or low to high. Uses only two voltage levels.
- Differential Manchester coding consists of combining the NRZ-I and RZ schemes.
  - Every symbol has a level transition in the middle. But the level at the beginning of the symbol is determined by the symbol value. One symbol causes a level change the other does not.

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





Note

In Manchester and differential Manchester encoding, the transition at the middle of the bit is used for synchronization.

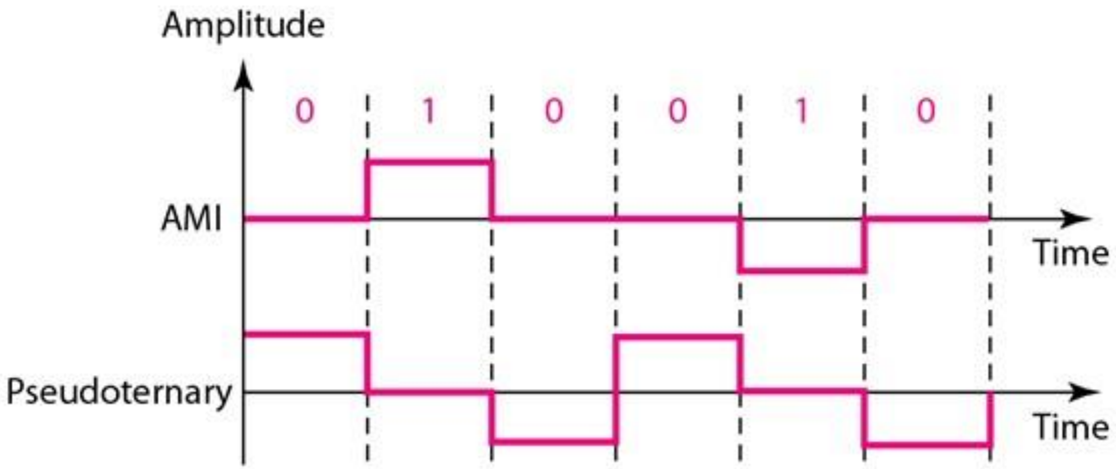
Note

The minimum bandwidth of Manchester and differential Manchester is 2 times that of NRZ. The is no DC component and no baseline wandering. None of these codes has error detection.

Bipolar - AMI and Pseudoternary

- Code uses 3 voltage levels: - +, 0, -, to represent the symbols (note not transitions to zero as in RZ).
- Voltage level for one symbol is at “0” and the other alternates between + & -.
- Bipolar Alternate Mark Inversion (AMI) - the “0” symbol is represented by zero voltage and the “1” symbol alternates between +V and -V.
- Pseudoternary is the reverse of AMI.

Figure 4.9 *Bipolar schemes: AMI and pseudoternary*



Bipolar C/Cs

- It is a better alternative to NRZ.
- Has no DC component or baseline wandering.
- Has no self synchronization because long runs of “0”s results in no signal transitions.
- No error detection.

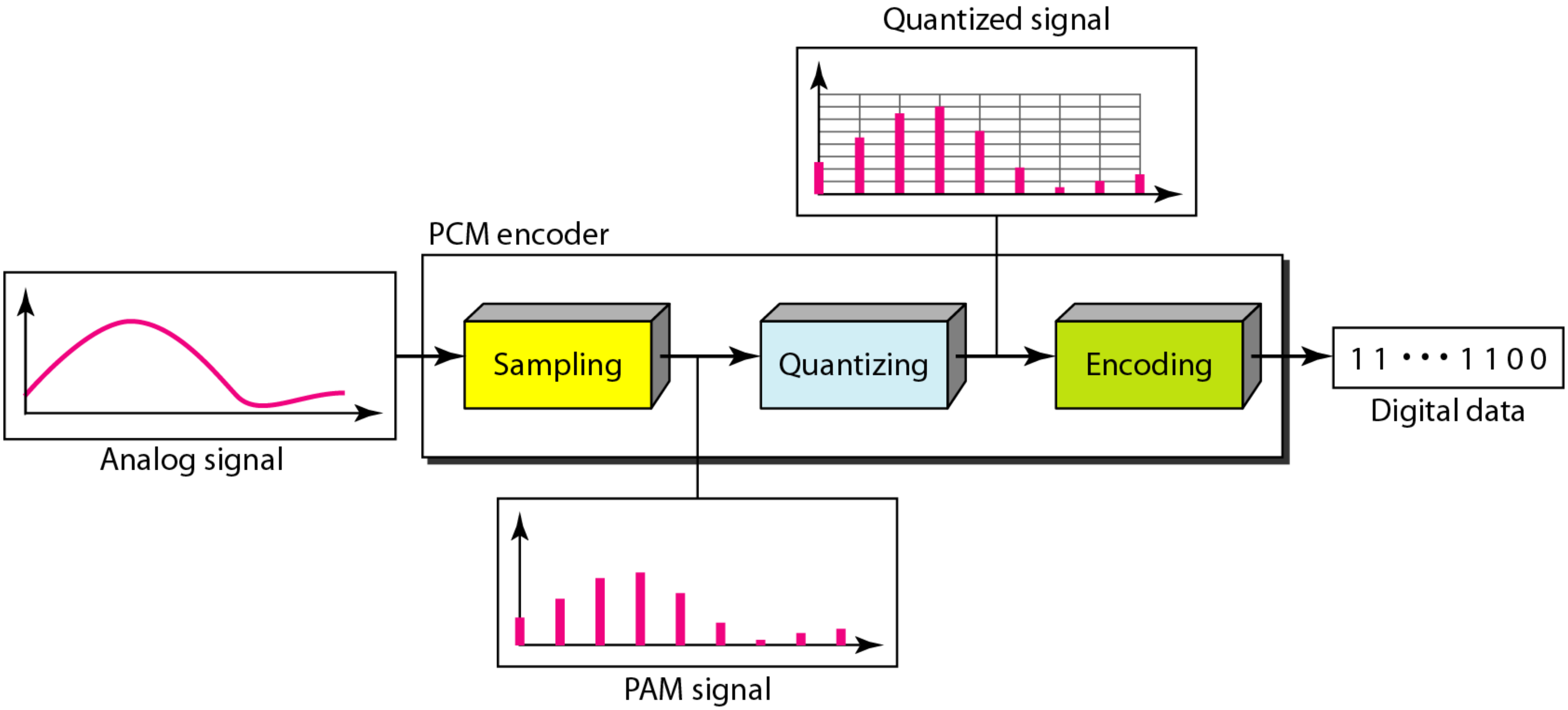
Digital Transmission

PCM

- PCM consists of three steps to digitize an analog signal:

1. Sampling
2. Quantization
3. Binary encoding

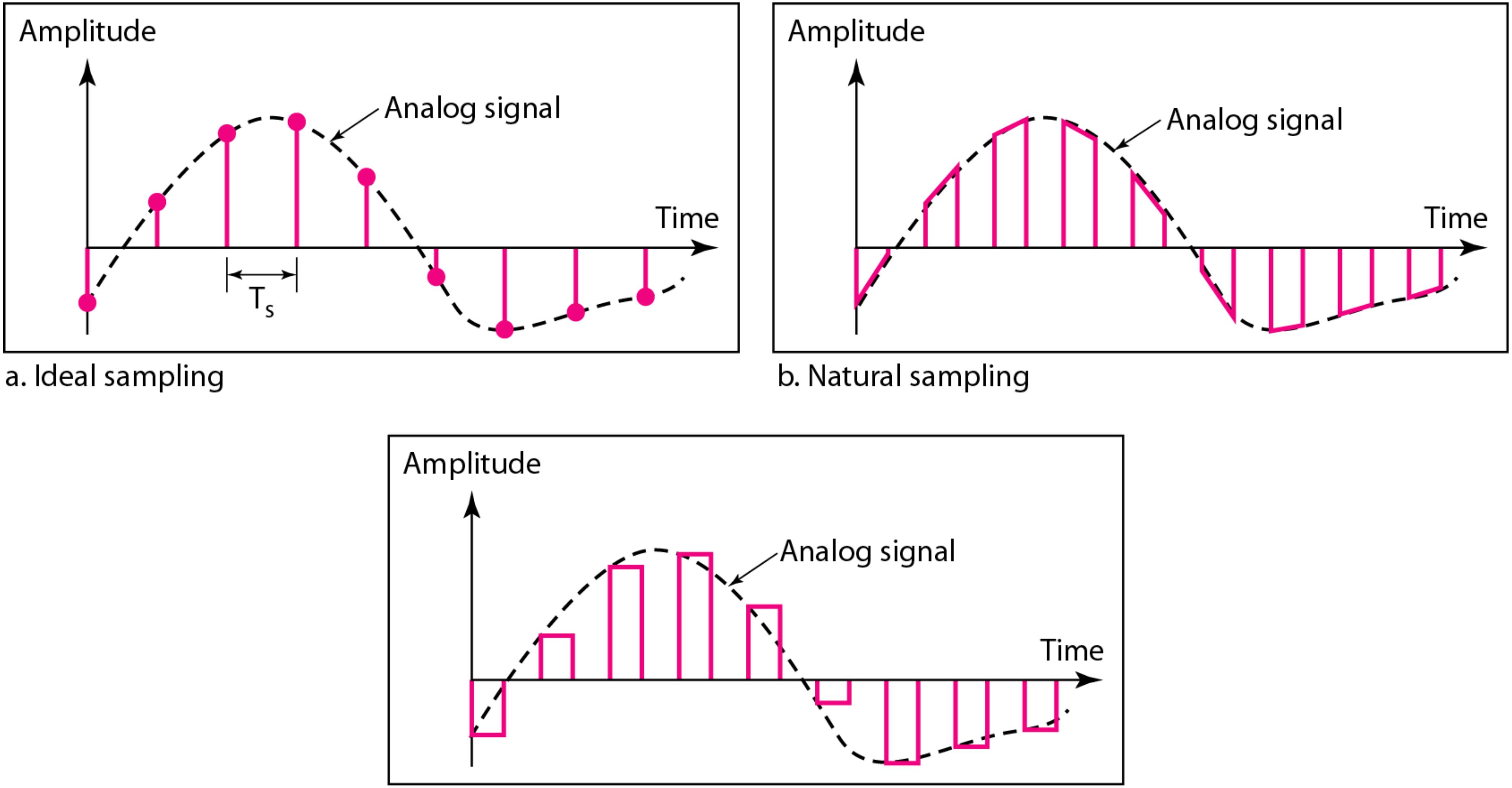
Figure 4.21 *Components of PCM encoder*



Sampling

- Analog signal is sampled every T_S secs.
- T_s is referred to as the sampling interval.
- $f_s = 1/T_s$ is called the sampling rate or sampling frequency.
- There are 3 sampling methods:
  - Ideal - an impulse at each sampling instant
  - Natural - a pulse of short width with varying amplitude
  - Flattop - sample and hold, like natural but with single amplitude value
- The process is referred to as pulse amplitude modulation PAM and the outcome is a signal with analog (non integer) values

Figure 4.22 *Three different sampling methods for PCM*



a. Ideal sampling

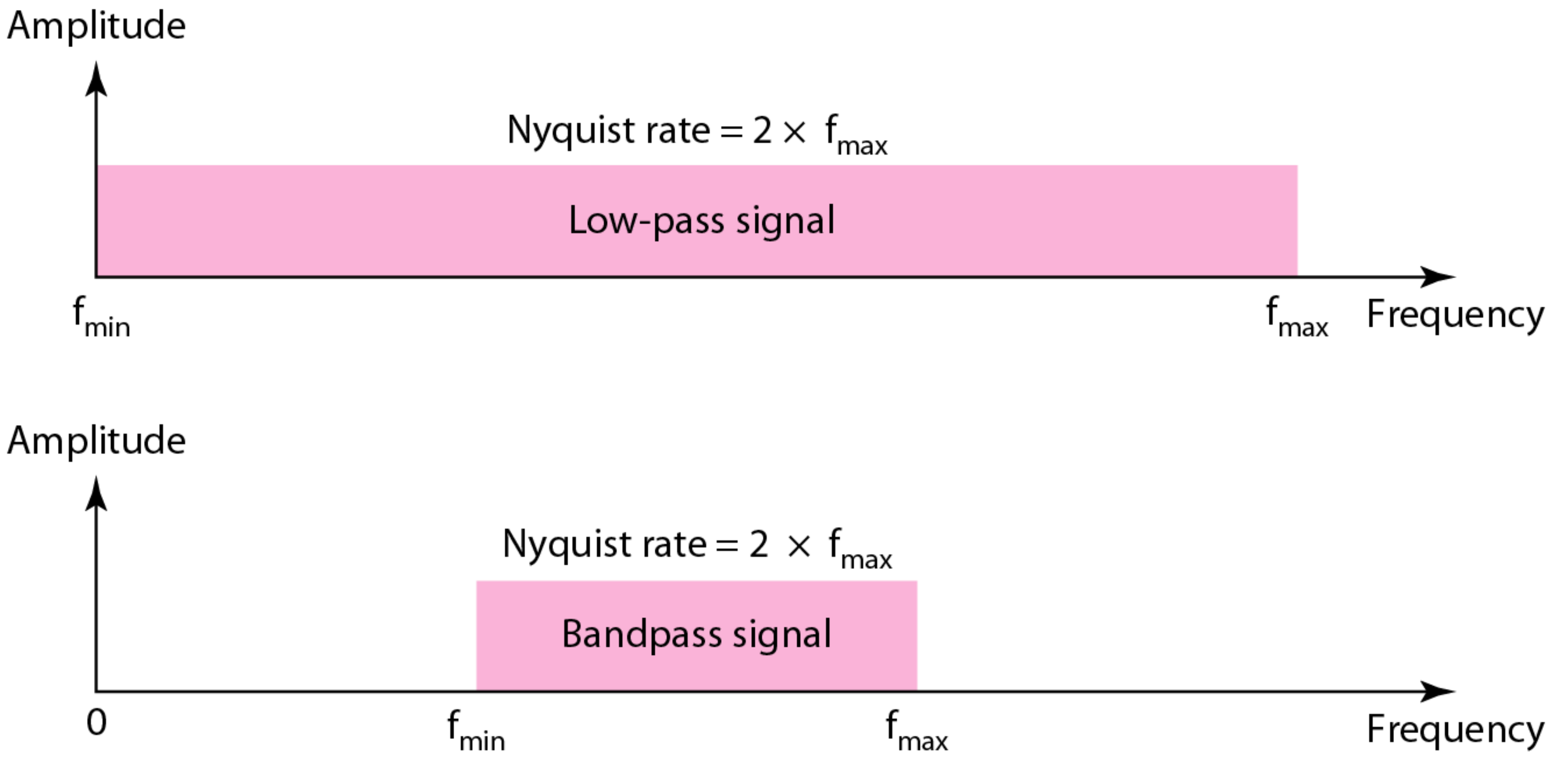
b. Natural sampling

c. Flat-top sampling

Note

According to the Nyquist theorem, the sampling rate must be at least 2 times the highest frequency contained in the signal.

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



Quantization

- Sampling results in a series of pulses of varying amplitude values ranging between two limits: a min and a max.
- The amplitude values are infinite between the two limits.
- We need to map the *infinite* amplitude values onto a finite set of known values.
- This is achieved by dividing the distance between min and max into L zones, each of height Δ.

$$\Delta = (max - min)/L$$

Quantization Levels

- The midpoint of each zone is assigned a value from 0 to L-1 (resulting in L values)
- Each sample falling in a zone is then approximated to the value of the midpoint.

Quantization Zones

- Assume we have a voltage signal with amplitutes V_{min}=-20V and V_{max}=+20V.
- We want to use L=8 quantization levels.
- Zone width Δ = (20 - -20)/8 = 5
- The 8 zones are: -20 to -15, -15 to -10, -10 to -5, -5 to 0, 0 to +5, +5 to +10, +10 to +15, +15 to +20
- The midpoints are: -17.5, -12.5, -7.5, -2.5, 2.5, 7.5, 12.5, 17.5

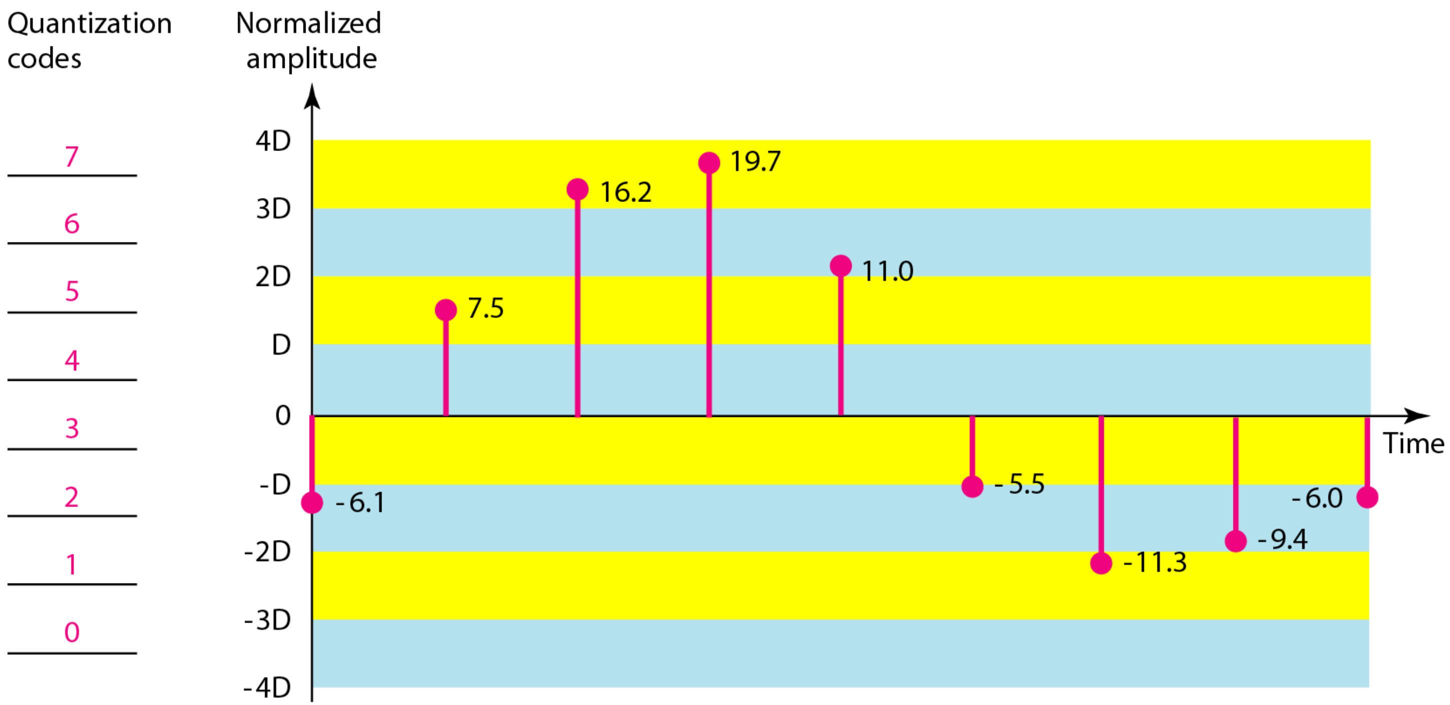
Assigning Codes to Zones

- Each zone is then assigned a binary code.
- The number of bits required to encode the zones, or the number of bits per sample as it is commonly referred to, is obtained as follows:

$$n_b = \log_2 L$$

- Given our example, $n_b = 3$
- The 8 zone (or level) codes are therefore: 000, 001, 010, 011, 100, 101, 110, and 111
- Assigning codes to zones:
  - 000 will refer to zone -20 to -15
  - 001 to zone -15 to -10, etc.

Figure 4.26 *Quantization and encoding of a sampled signal*



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

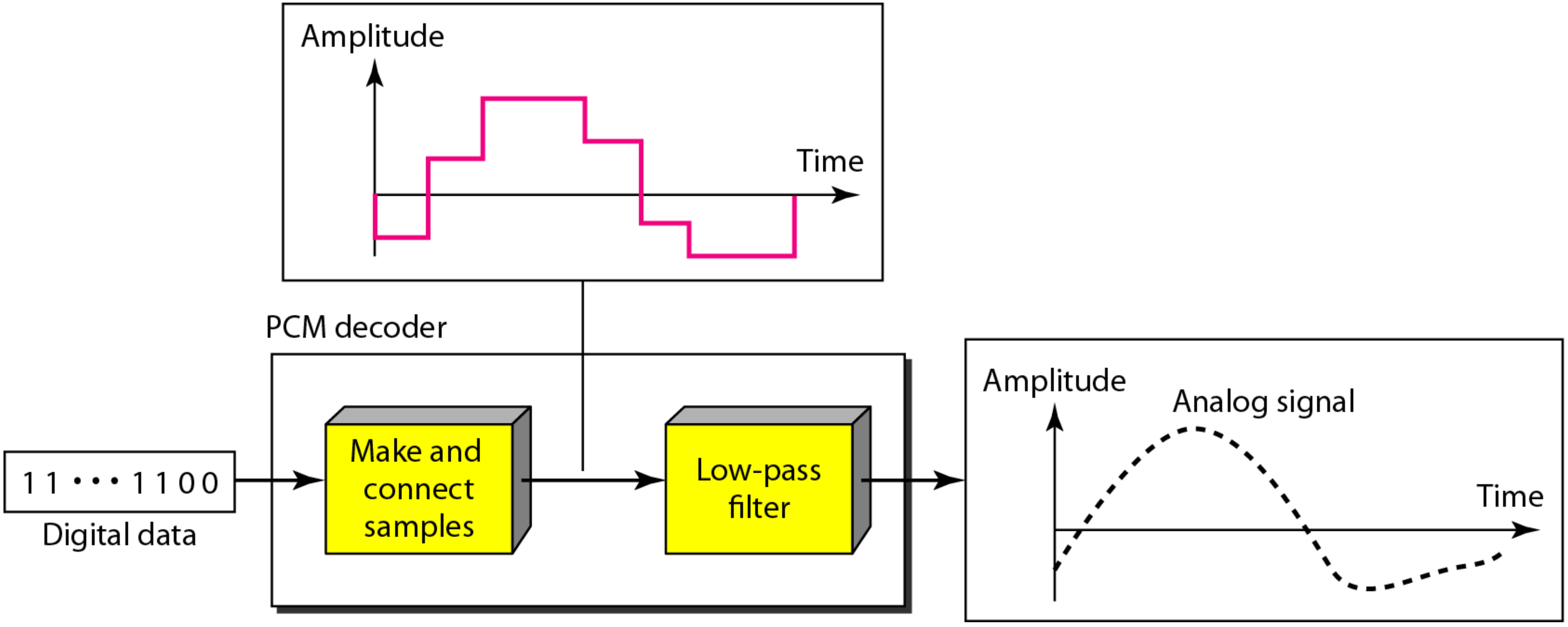
Quantization Error

- When a signal is quantized, we introduce an error - the coded signal is an approximation of the actual amplitude value.
- The difference between actual and coded value (midpoint) is referred to as the quantization error.
- The more zones, the smaller Δ which results in smaller errors.
- BUT, the more zones the more bits required to encode the samples -> higher bit rate

PCM Decoder

- To recover an analog signal from a digitized signal we follow the following steps:
  - We use a hold circuit that holds the amplitude value of a pulse till the next pulse arrives.
  - We pass this signal through a low pass filter with a cutoff frequency that is equal to the highest frequency in the pre-sampled signal.
- The higher the value of L, the less distorted a signal is recovered.

Figure 4.27 *Components of a PCM decoder*



4-3 TRANSMISSION MODES

The transmission of binary data across a link can be accomplished in either parallel or serial mode. In parallel mode, multiple bits are sent with each clock tick. In serial mode, 1 bit is sent with each clock tick. While there is only one way to send parallel data, there are three subclasses of serial transmission: asynchronous, synchronous, and isochronous.

Topics discussed in this section:

- **Parallel Transmission**
- **Serial Transmission**

Figure 4.31 *Data transmission and modes*

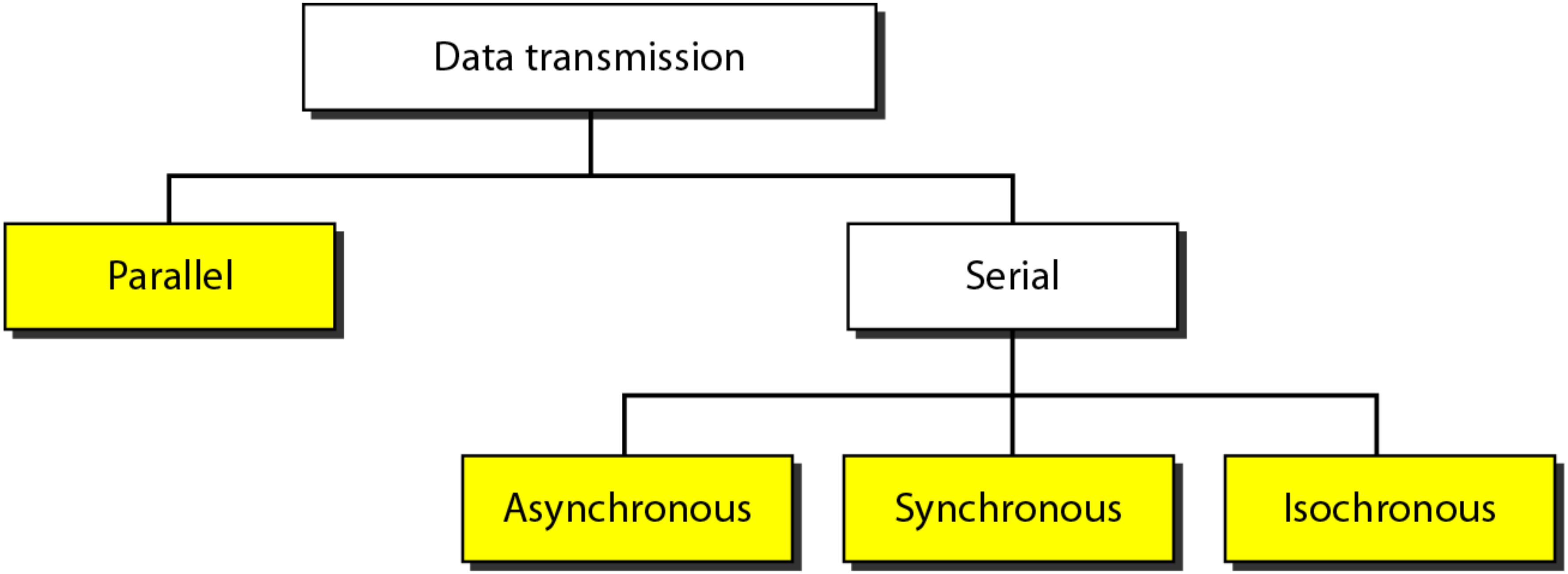


Figure 4.32 *Parallel transmission*

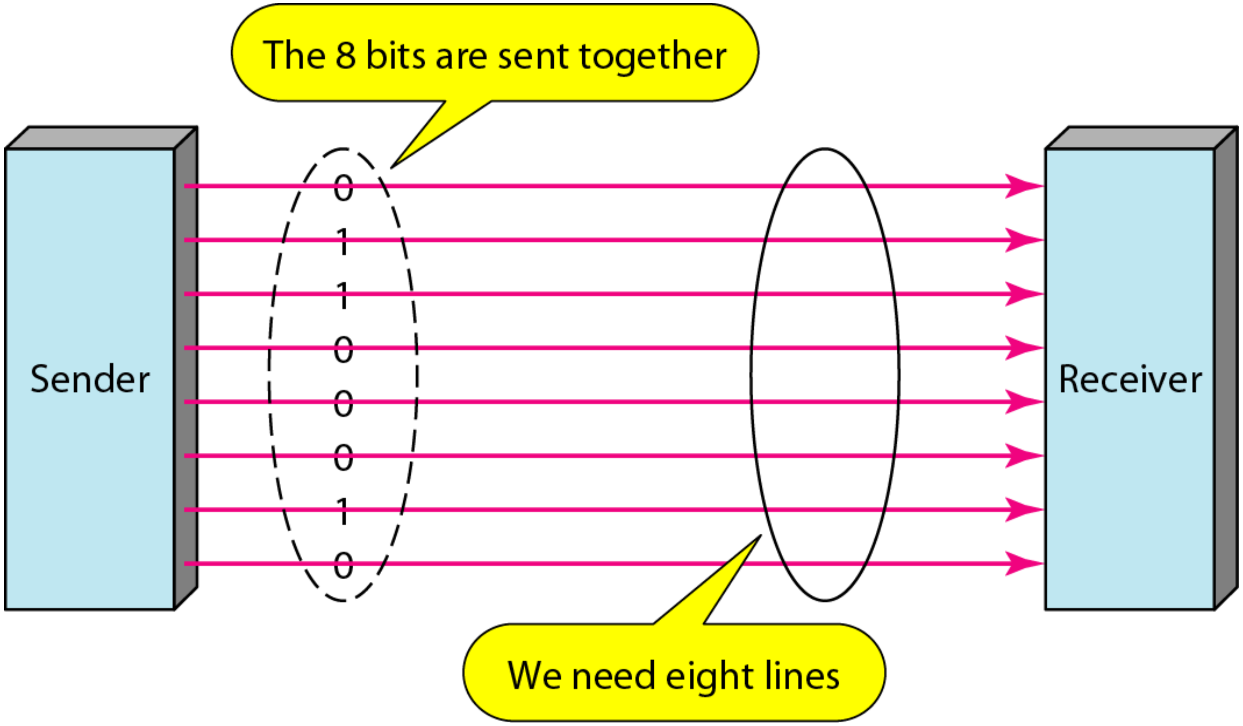
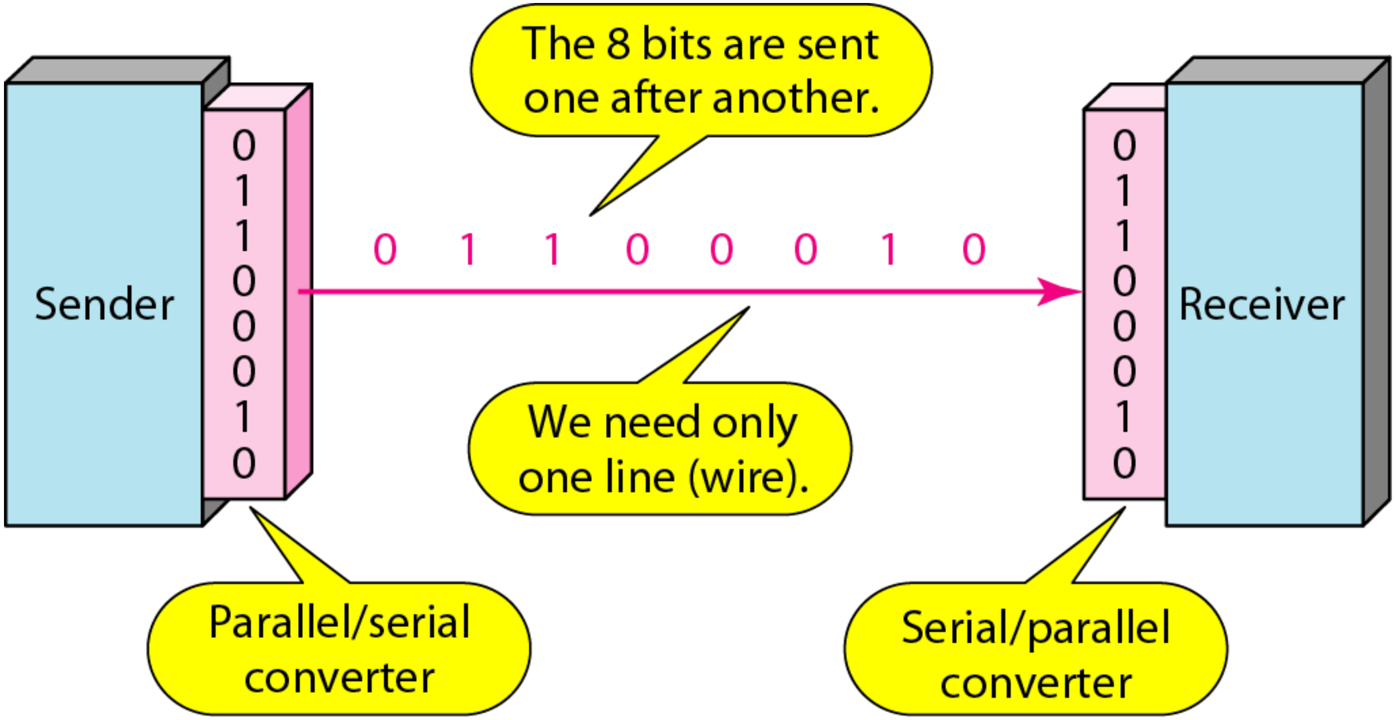


Figure 4.33 *Serial transmission*


The 8 bits are sent one after another.
Sender
0 1 1 0 0 0 1 0
0 1 1 0 0 0 1 0
0 1 1 0 0 0 1 0
Receiver
We need only one line (wire).
Parallel/serial converter
Serial/parallel converter

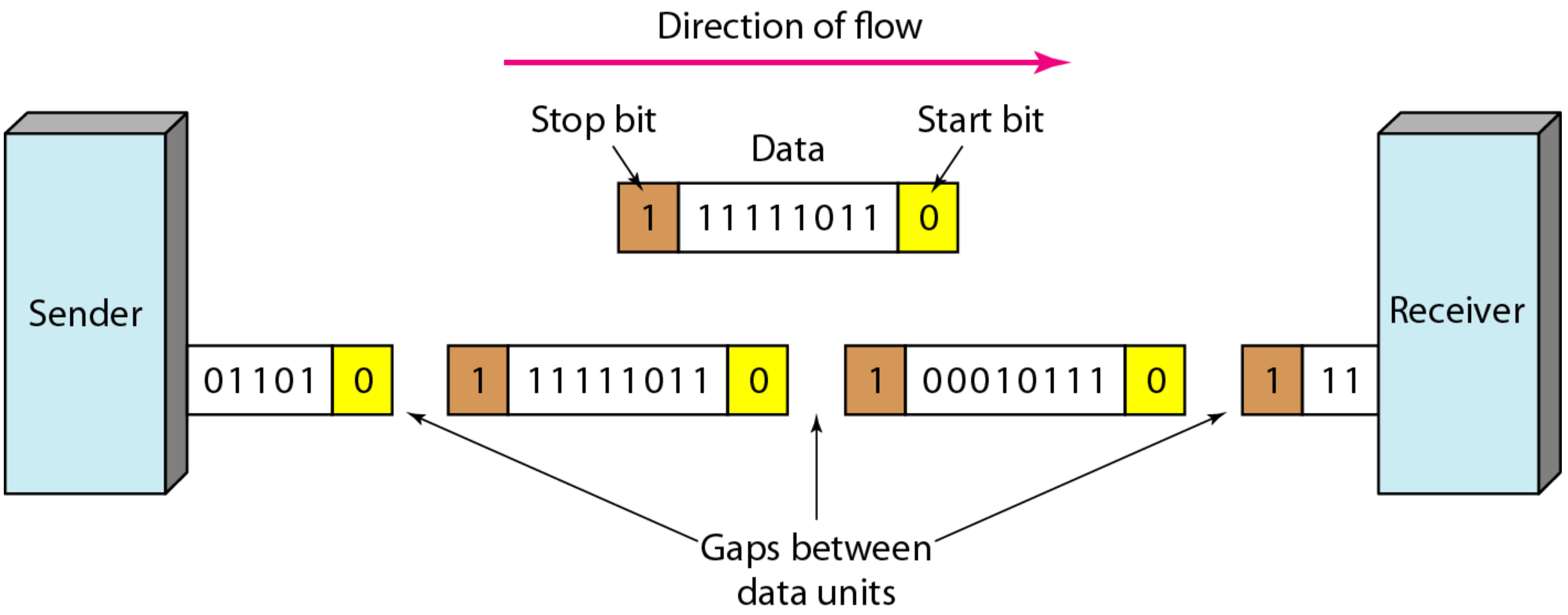
Note

In asynchronous transmission, we send 1 start bit (0) at the beginning and 1 or more stop bits (1s) at the end of each byte. There may be a gap between each byte.

Note

Asynchronous here means "asynchronous at the byte level," but the bits are still synchronized; their durations are the same.

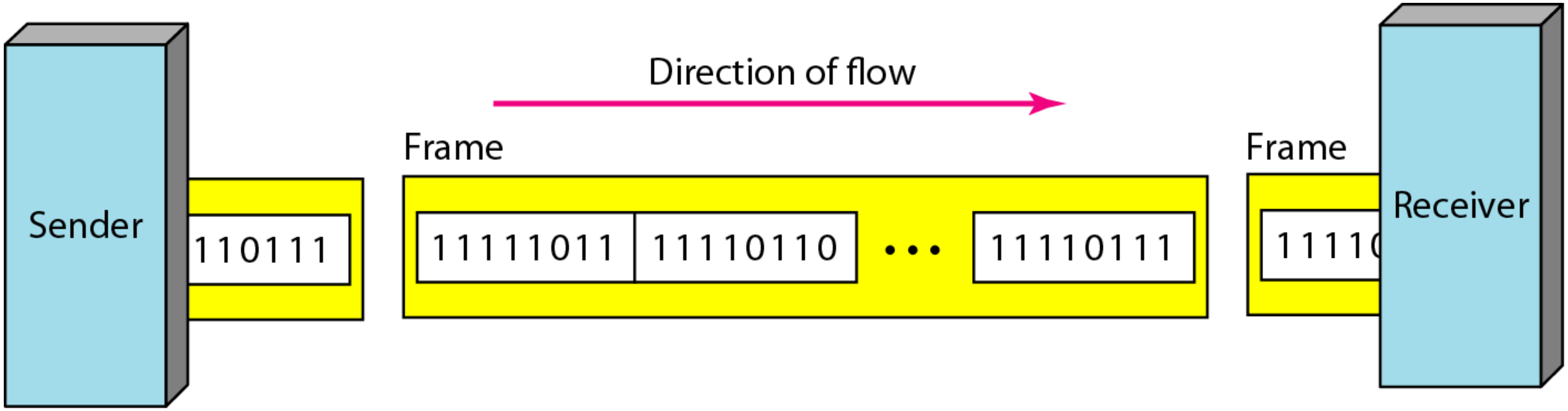
Figure 4.34 *Asynchronous transmission*



Note

In synchronous transmission, we send bits one after another without start or stop bits or gaps. It is the responsibility of the receiver to group the bits. The bits are usually sent as bytes and many bytes are grouped in a frame. A frame is identified with a start and an end byte.

Figure 4.35 *Synchronous transmission*



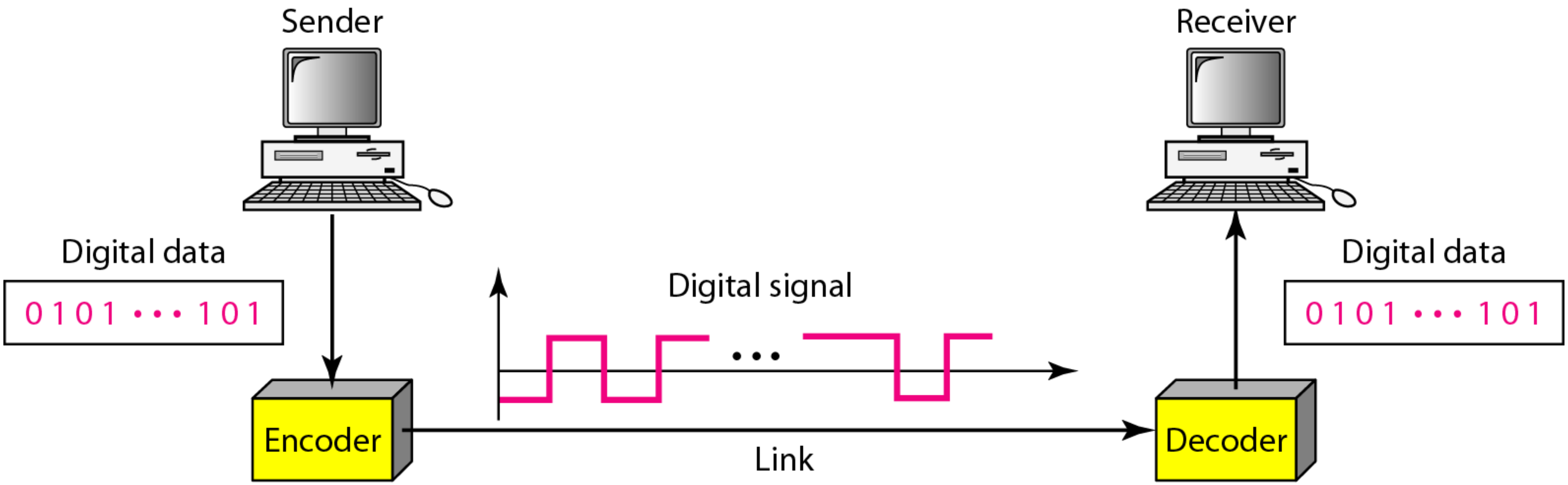
Isochronous

- In isochronous transmission we cannot have uneven gaps between frames.
- Transmission of bits is fixed with equal gaps.
- Example: Multimedia transmission

Line Coding

- Converting a string of 1’s and 0’s (digital data) into a sequence of signals that denote the 1’s and 0’s.
- For example a high voltage level (+V) could represent a “1” and a low voltage level (0 or -V) could represent a “0”.

Figure 4.1 *Line coding and decoding*



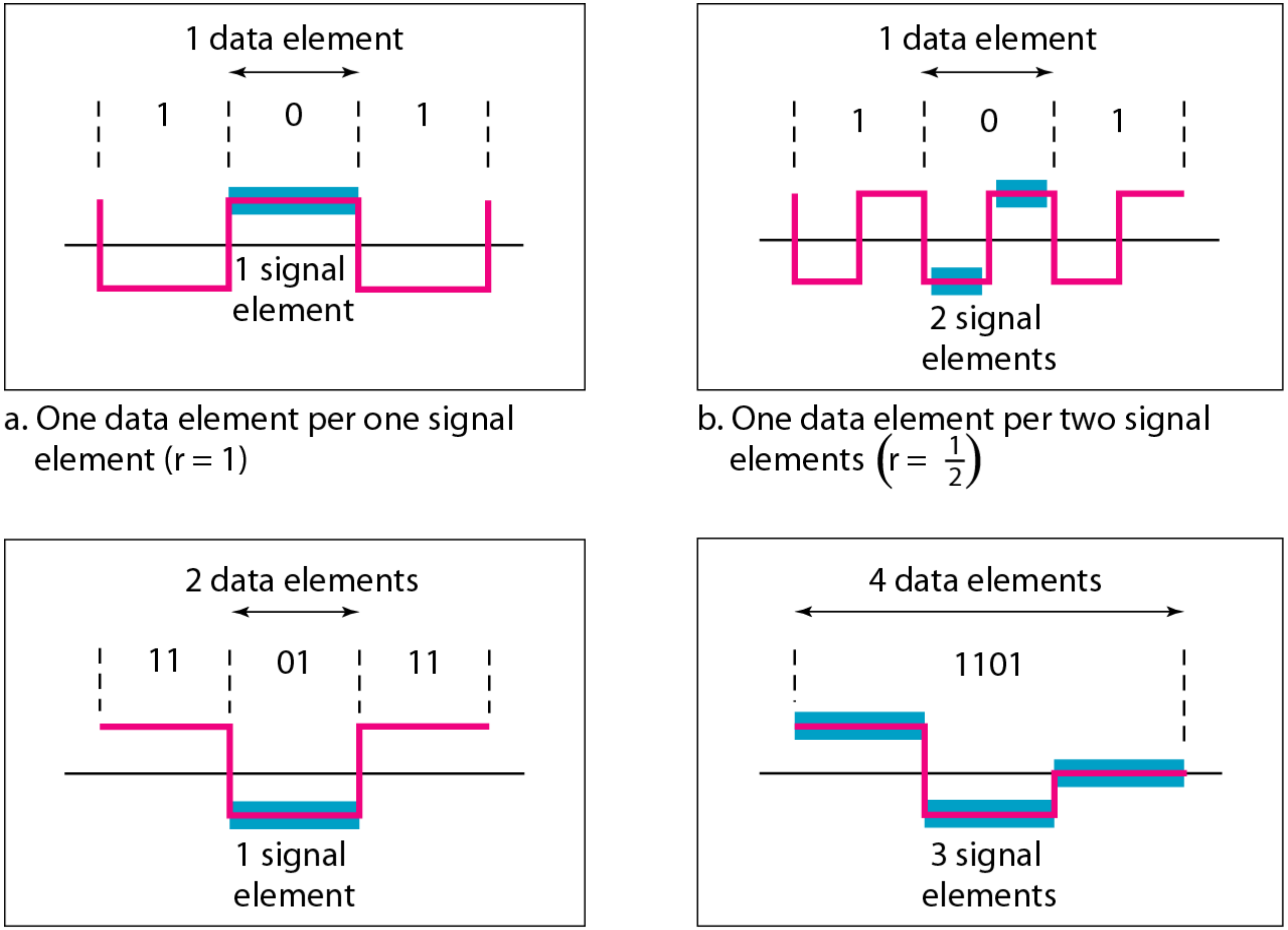
Mapping Data symbols onto Signal levels

- A data symbol (or element) can consist of a number of data bits:
  - 1 , 0 or
  - 11, 10, 01, ……
- A data symbol can be coded into a single signal element or multiple signal elements
  - 1 -> +V, 0 -> -V
  - 1 -> +V and -V, 0 -> -V and +V
- The ratio ‘r’ is the number of data elements carried by a signal element.

Relationship between data rate and signal rate

- The data rate defines the number of bits sent per sec - bps. It is often referred to the bit rate.
- The signal rate is the number of signal elements sent in a second and is measured in bauds. It is also referred to as the modulation rate OR baud rate.
- Goal is to increase the data rate whilst reducing the baud rate.

Figure 4.2 *Signal element versus data element*



a. One data element per one signal element (r = 1)

b. One data element per two signal elements $\left(r = \frac{1}{2}\right)$

c. Two data elements per one signal element (r = 2)

d. Four data elements per three signal elements $\left(r = \frac{4}{3}\right)$

Data rate and Baud rate

- The baud or signal rate can be expressed as:

  S = c x N x 1/r bauds

  where N is data rate

  c is the case factor (worst, best & avg.)

  r is the ratio between data element & signal element

Example 4.1

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

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

Example 4.2

The maximum data rate of a channel (see Chapter 3) is Nmax = 2 × B × $\log_2$ L (defined by the Nyquist formula). Does this agree with the previous formula for N_{max}?

Solution

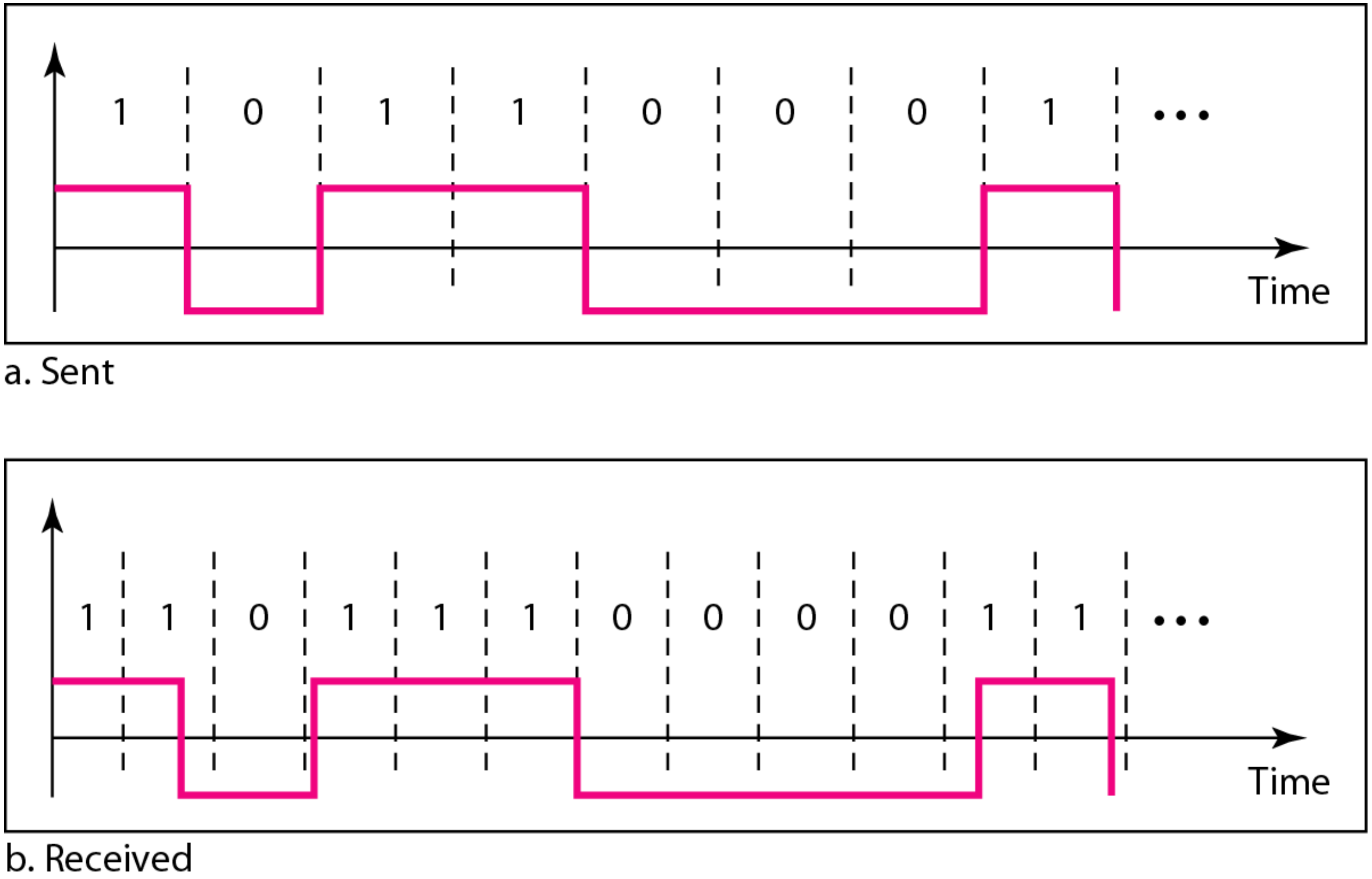
A signal with L levels actually can carry $\log_2 L$ bits per level. If each level corresponds to one signal element and we assume the average case (c = 1/2), then we have

$$\boxed{N_{\max} = \frac{1}{c} \times B \times r = 2 \times B \times \log_2 L}$$

Considerations for choosing a good signal element referred to as line encoding

- Self synchronization - the clocks at the sender and the receiver must have the same bit interval.
- If the receiver clock is faster or slower it will misinterpret the incoming bit stream.

Figure 4.3 *Effect of lack of synchronization*

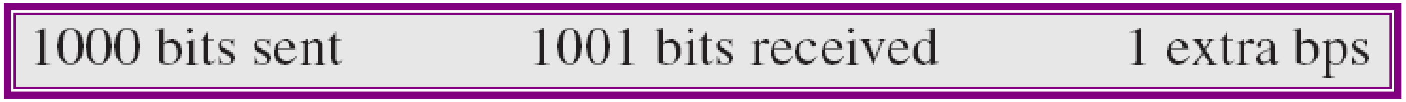


Example 4.3

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?

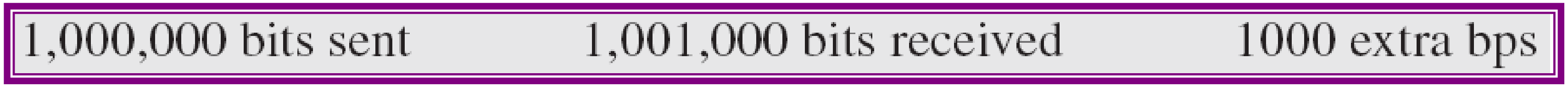
Solution

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

Line encoding C/Cs

- Error detection - errors occur during transmission due to line impairments.
- Some codes are constructed such that when an error occurs it can be detected.

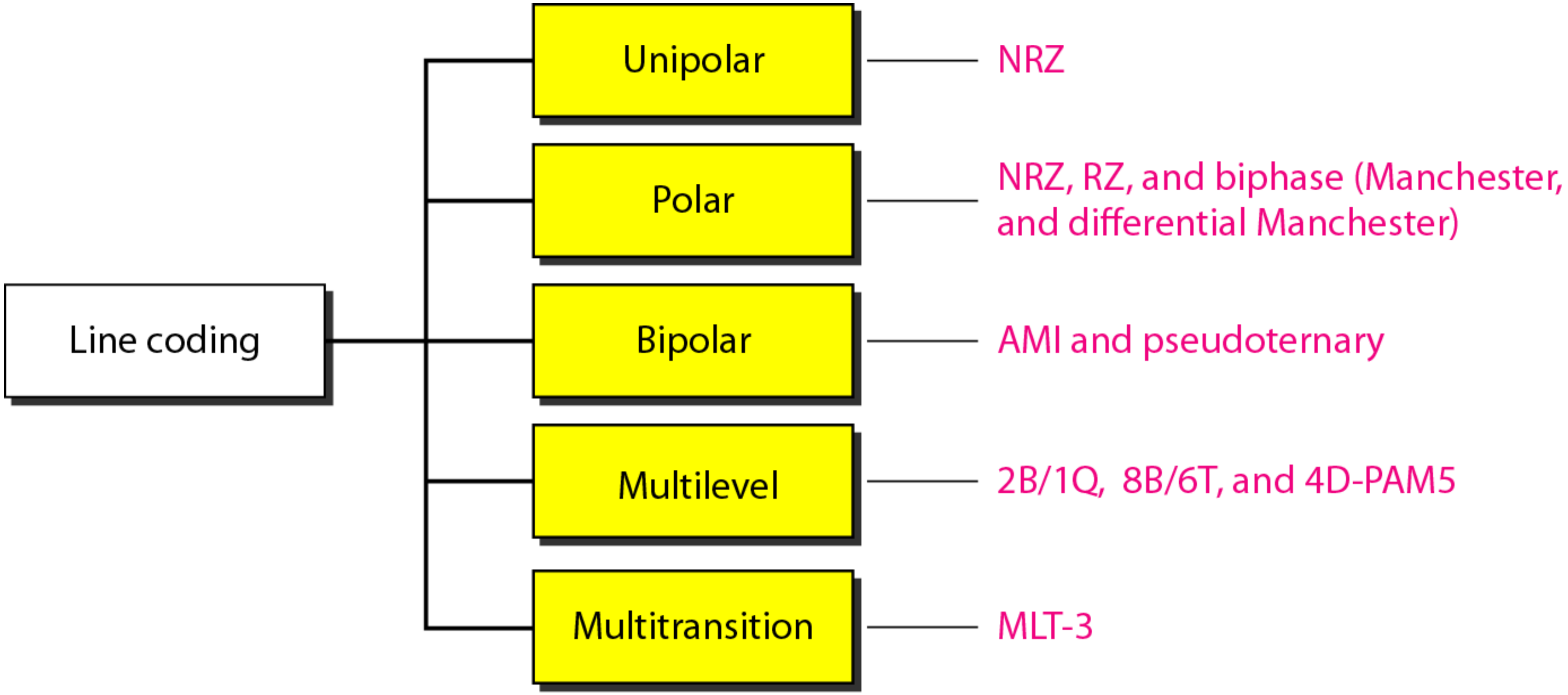
Line encoding C/Cs

- Noise and interference - there are line encoding techniques that make the transmitted signal “immune” to noise and interference.
- This means that the signal cannot be corrupted, it is stronger than error detection.

Line encoding C/Cs

- Complexity - the more robust and resilient the code, the more complex it is to implement and the price is often paid in baud rate or required bandwidth.

Figure 4.4 *Line coding schemes*



Line coding

Unipolar — NRZ

Polar — NRZ, RZ, and biphase (Manchester, and differential Manchester)

Bipolar — AMI and pseudoternary

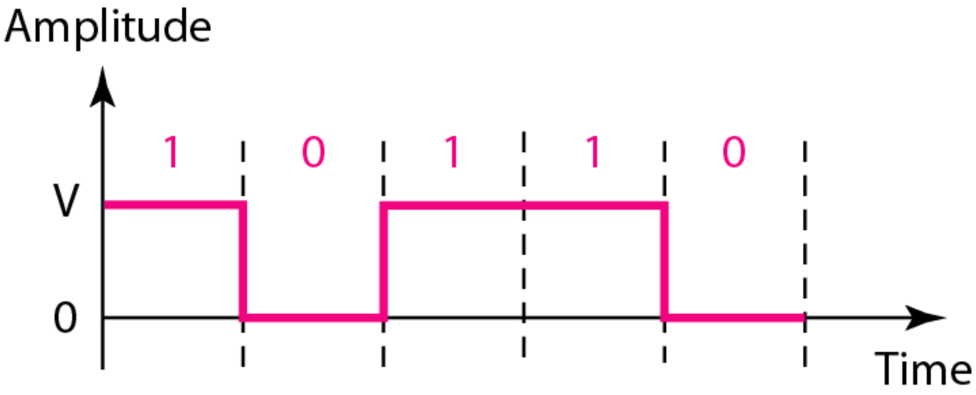
Multilevel — 2B/1Q, 8B/6T, and 4D-PAM5

Multitransition — MLT-3

Unipolar

- All signal levels are on one side of the time axis - either above or below
- NRZ - Non Return to Zero scheme is an example of this code. The signal level does not return to zero during a symbol transmission.
- Scheme is prone to baseline wandering and DC components. It has no synchronization or any error detection. It is simple but costly in power consumption.

Figure 4.5 *Unipolar NRZ scheme*



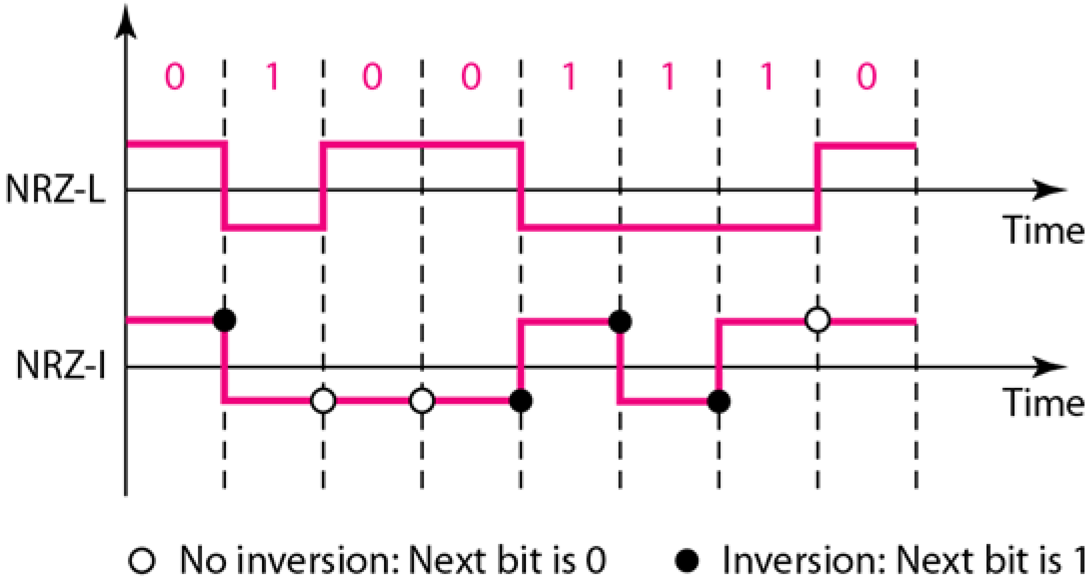
$$\frac{1}{2}V^2 + \frac{1}{2}(0)^2 = \frac{1}{2}V^2$$

Normalized power

Polar - NRZ

- The voltages are on both sides of the time axis.
- Polar NRZ scheme can be implemented with two voltages. E.g. +V for 1 and -V for 0.
- There are two versions:
  - NZR - Level (NRZ-L) - positive voltage for one symbol and negative for the other
  - NRZ - Inversion (NRZ-I) - the change or lack of change in polarity determines the value of a symbol. E.g. a “1” symbol inverts the polarity a “0” does not.

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



Note

In NRZ-L the level of the voltage determines the value of the bit.
In NRZ-I the inversion or the lack of inversion determines the value of the bit.

Note

NRZ-L and NRZ-I both have a DC component problem and baseline wandering, it is worse for NRZ-L. Both have no self synchronization &no error detection. Both are relatively simple to implement.

Example 4.4

A system is using NRZ-I to transfer 1-Mbps data. What are the average signal rate and minimum bandwidth?

Solution

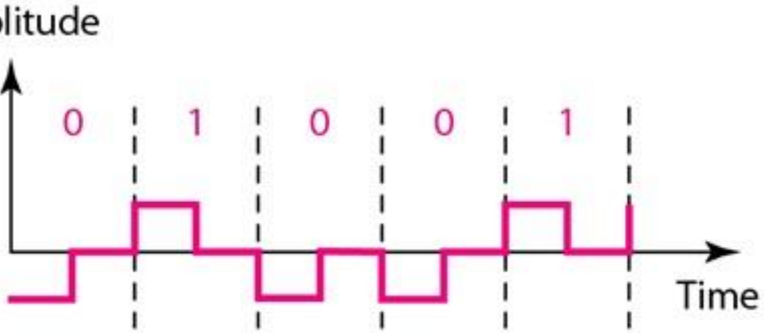
The average signal rate is S= c x N x R = 1/2 x N x 1 = 500 kbaud. The minimum bandwidth for this average baud rate is Bmin = S = 500 kHz.

Note c = 1/2 for the avg. case as worst case is 1 and best case is 0

Polar - RZ

- The Return to Zero (RZ) scheme uses three voltage values. +, 0, -.
- Each symbol has a transition in the middle. Either from high to zero or from low to zero.
- This scheme has more signal transitions (two per symbol) and therefore requires a wider bandwidth.
- No DC components or baseline wandering.
- Self synchronization - transition indicates symbol value.
- More complex as it uses three voltage level. It has no error detection capability.

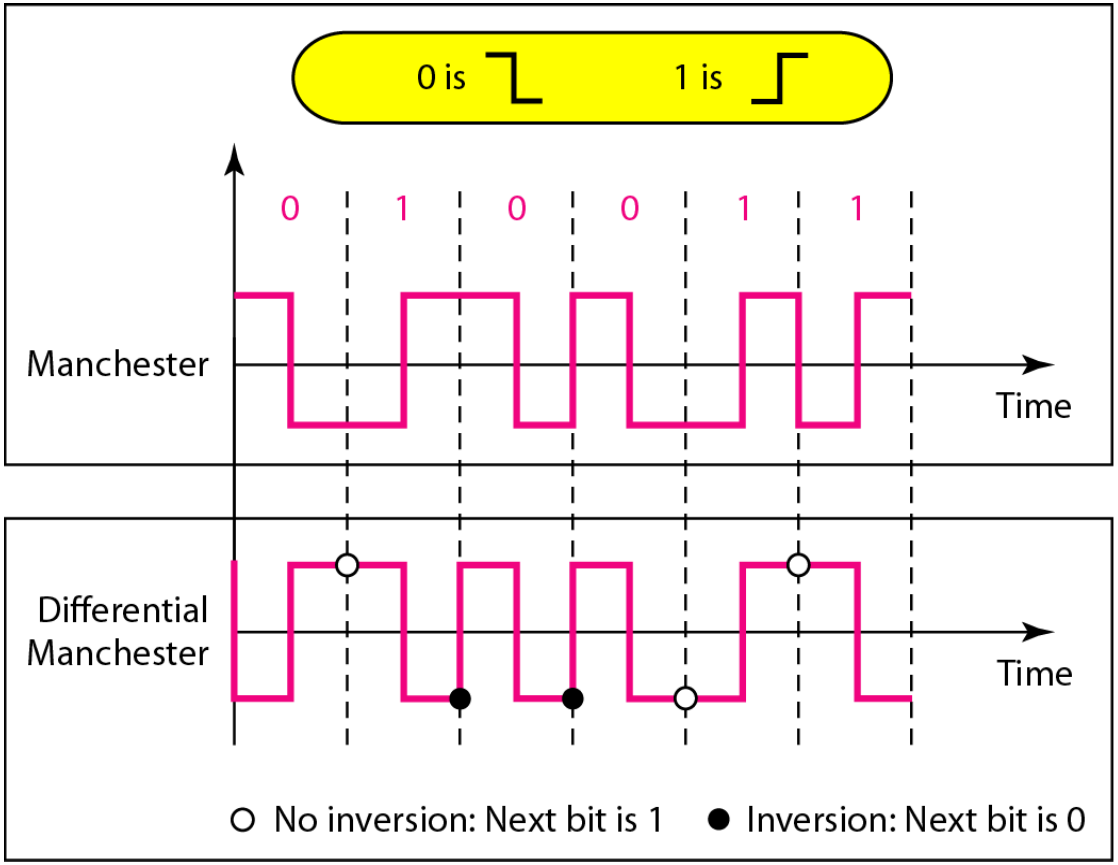
Figure 4.7 *Polar RZ scheme*

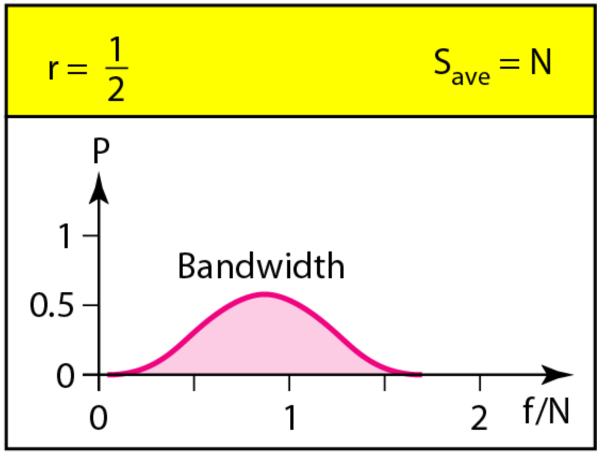


Polar - Biphase: Manchester and Differential Manchester

- Manchester coding consists of combining the NRZ-L and RZ schemes.
  - Every symbol has a level transition in the middle: from high to low or low to high. Uses only two voltage levels.
- Differential Manchester coding consists of combining the NRZ-I and RZ schemes.
  - Every symbol has a level transition in the middle. But the level at the beginning of the symbol is determined by the symbol value. One symbol causes a level change the other does not.

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





Note

In Manchester and differential Manchester encoding, the transition at the middle of the bit is used for synchronization.

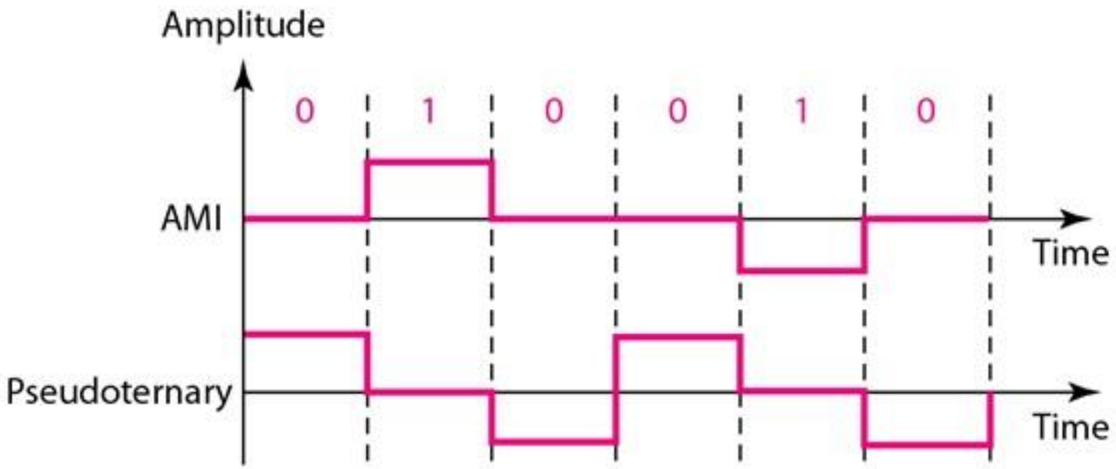
Note

The minimum bandwidth of Manchester and differential Manchester is 2 times that of NRZ. The is no DC component and no baseline wandering. None of these codes has error detection.

Bipolar - AMI and Pseudoternary

- Code uses 3 voltage levels: - +, 0, -, to represent the symbols (note not transitions to zero as in RZ).
- Voltage level for one symbol is at “0” and the other alternates between + & -.
- Bipolar Alternate Mark Inversion (AMI) - the “0” symbol is represented by zero voltage and the “1” symbol alternates between +V and -V.
- Pseudoternary is the reverse of AMI.

Figure 4.9 *Bipolar schemes: AMI and pseudoternary*



Bipolar C/Cs

- It is a better alternative to NRZ.
- Has no DC component or baseline wandering.
- Has no self synchronization because long runs of “0”s results in no signal transitions.
- No error detection.

MODULE – 3

BANDWIDTH UTILIZATION

MULTIPLEXING

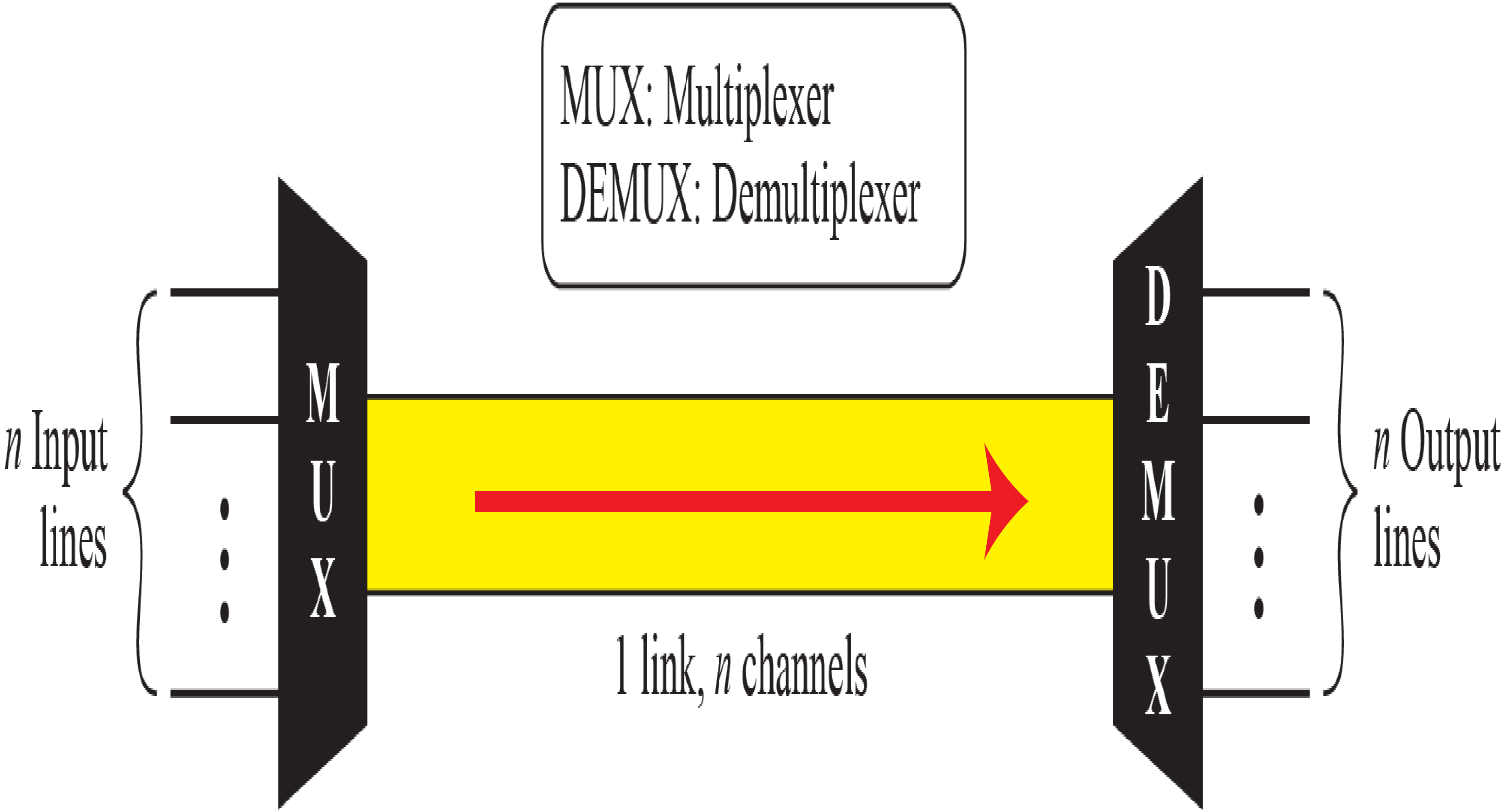
•Multiplexing is the set of techniques that allows the simultaneous transmission of multiple signals across a single data link

•As data and telecommunications use increases, so does traffic, this increase can be accommodated by continuing to add individual links each time a new channel is needed,

or

• Higher-bandwidth links are can be installed to carry multiple signals in each link.

Dividing a link into channels



Multiplexing

- Each has a excess bandwidth needed for average transmission of signal.
- If the bandwidth of a link is greater than the bandwidth needs of the devices connected to it,

  bandwidth is wasted.
- An efficient system maximizes the utilization of all resources.
- Bandwidth is one of the most precious resource in Data Communication.

Multiplexing

- In multiplexed system, n lines share the bandwidth of one link.
- The lines on the left direct their transmission stream into Multiplexer, which combines them into single stream.
- The stream is fed to demultiplexer at the receiving end, which separates the streams back into their corresponding lines.
- Link refers to physical path.

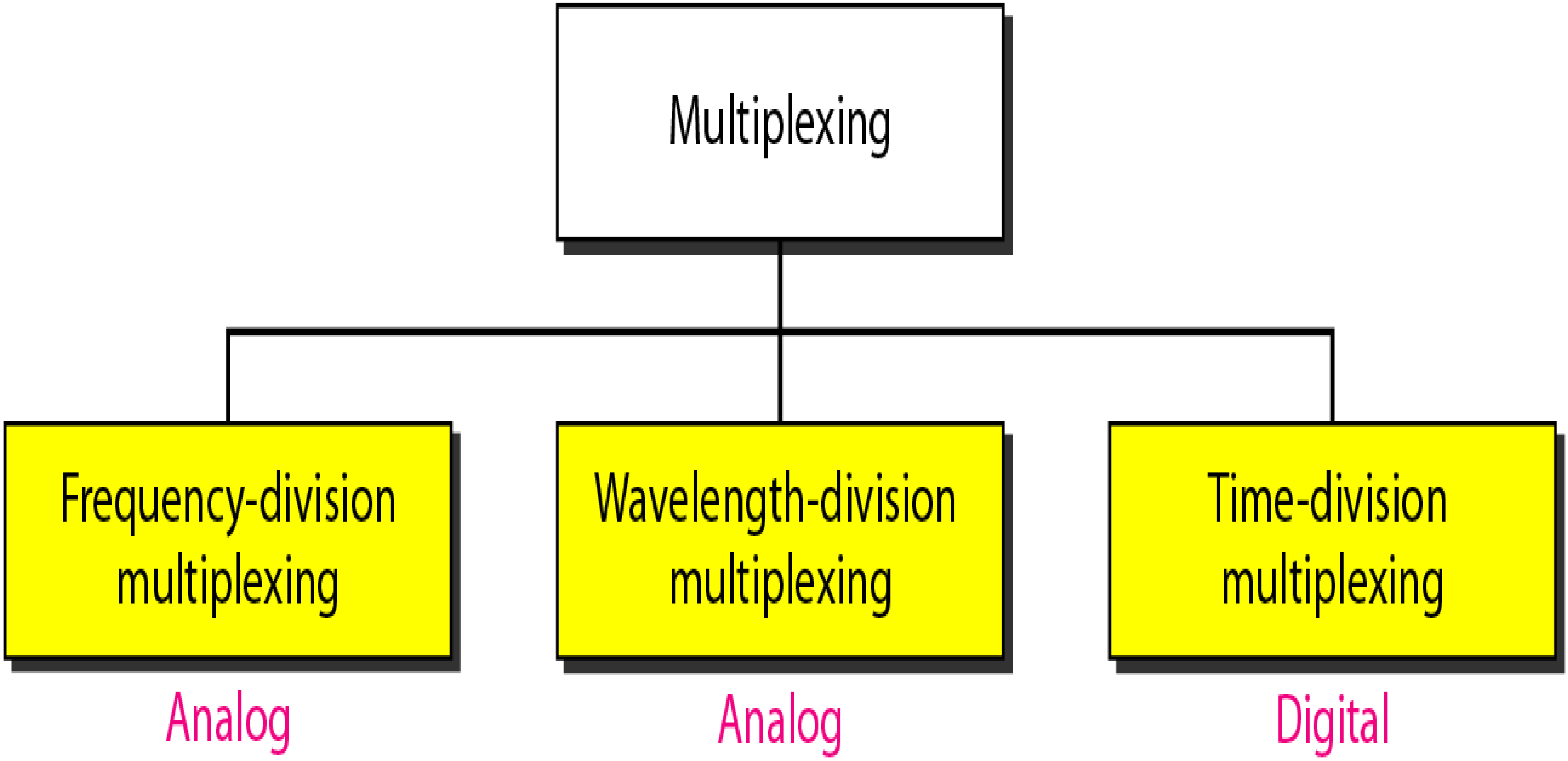
Multiplexing

- Channel refers to portion of a link that carries a transmission between a given pair of lines.
- One link can have n channels.

- There are three multiplexing techniques.

1) Frequency division multiplexing(FDM)
2) Wave length division multiplexing(WDM)
3) Time division multiplexing(TDM)

Categories of multiplexing


Multiplexing
Frequency-division multiplexing
Wavelength-division multiplexing
Time-division multiplexing
Analog
Analog
Digital

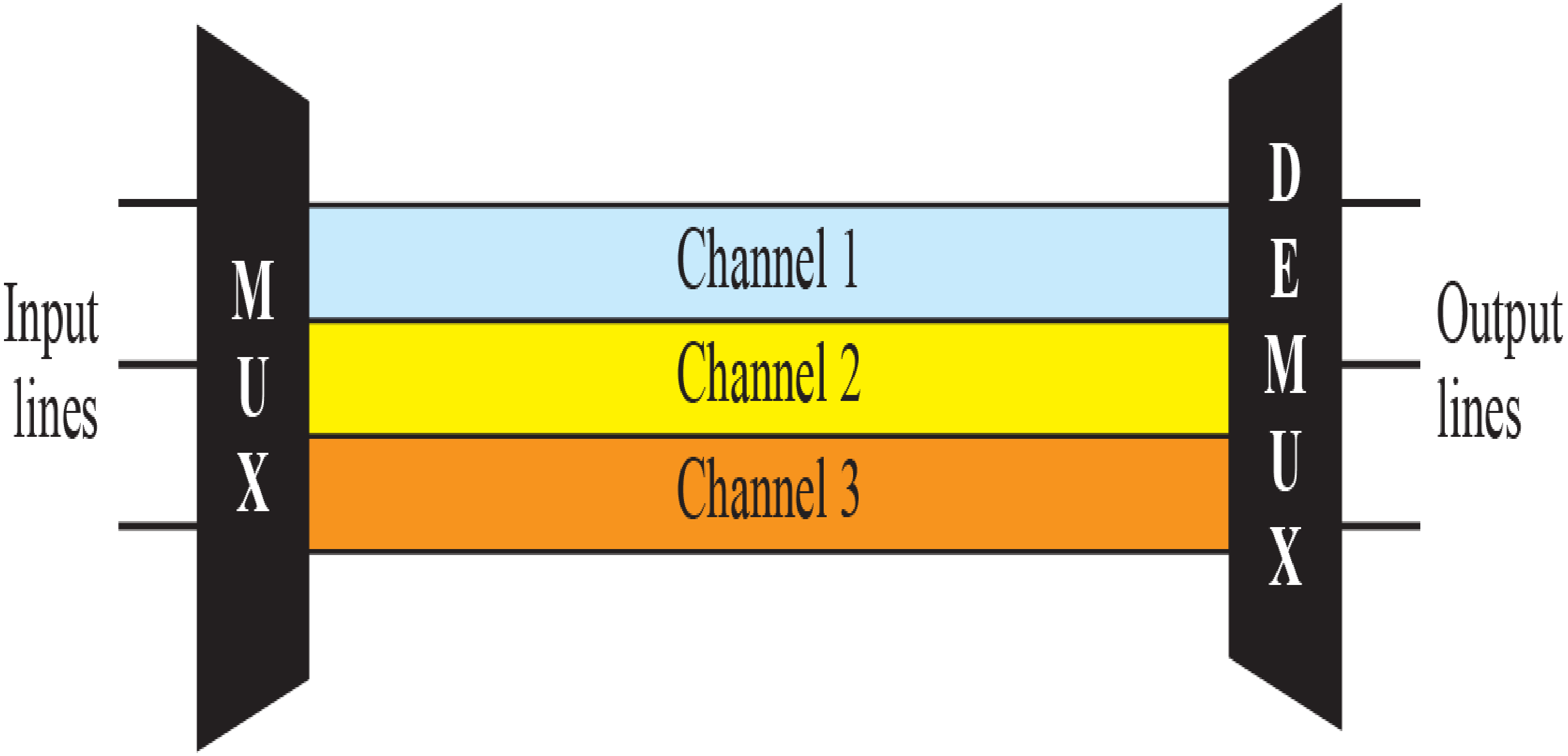
Frequency-Division Multiplexing

- Frequency-division multiplexing (FDM) is an analog technique that can be applied when the bandwidth of a link (in hertz) is greater than the combined bandwidths of the signals to be transmitted.

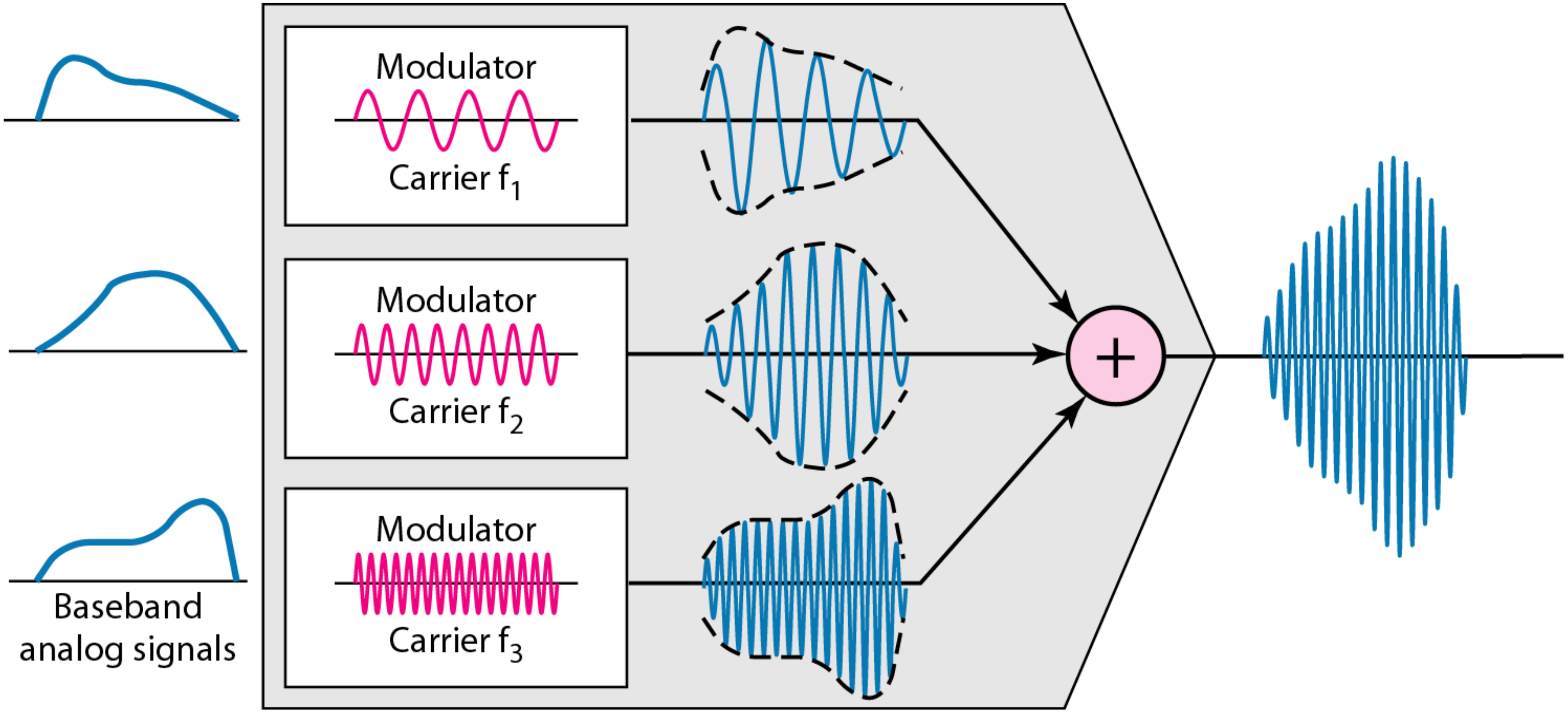
- In FDM, signals generated by each sending device are modulated at different carrier frequencies.

- These modulated signals are then combined into a single composite signal that can be transported by the link.

Frequency-division multiplexing



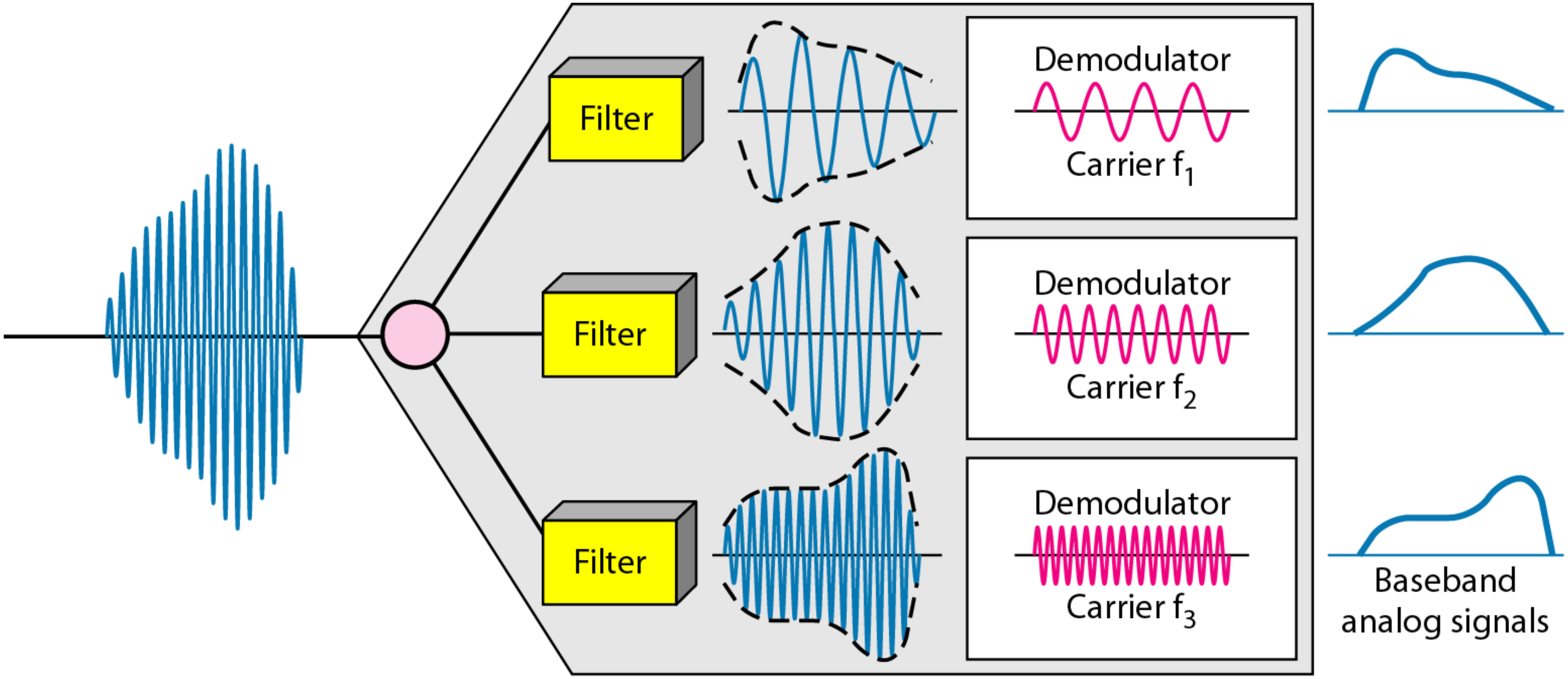
FDM Process


Modulator
Carrier f_1
Modulator
Carrier f_2
Modulator
Carrier f_3
+
Baseband analog signals

Multiplexing process

- Each source generates a signal of a similar frequency range.
- Inside the multiplexer, these similar signals modulates different carrier frequencies (f1,f2,f3).
- The resulting modulated signals are then combined into a single composite signal that is sent over the media which has enough bandwidth.

FDM Demultiplexing Process



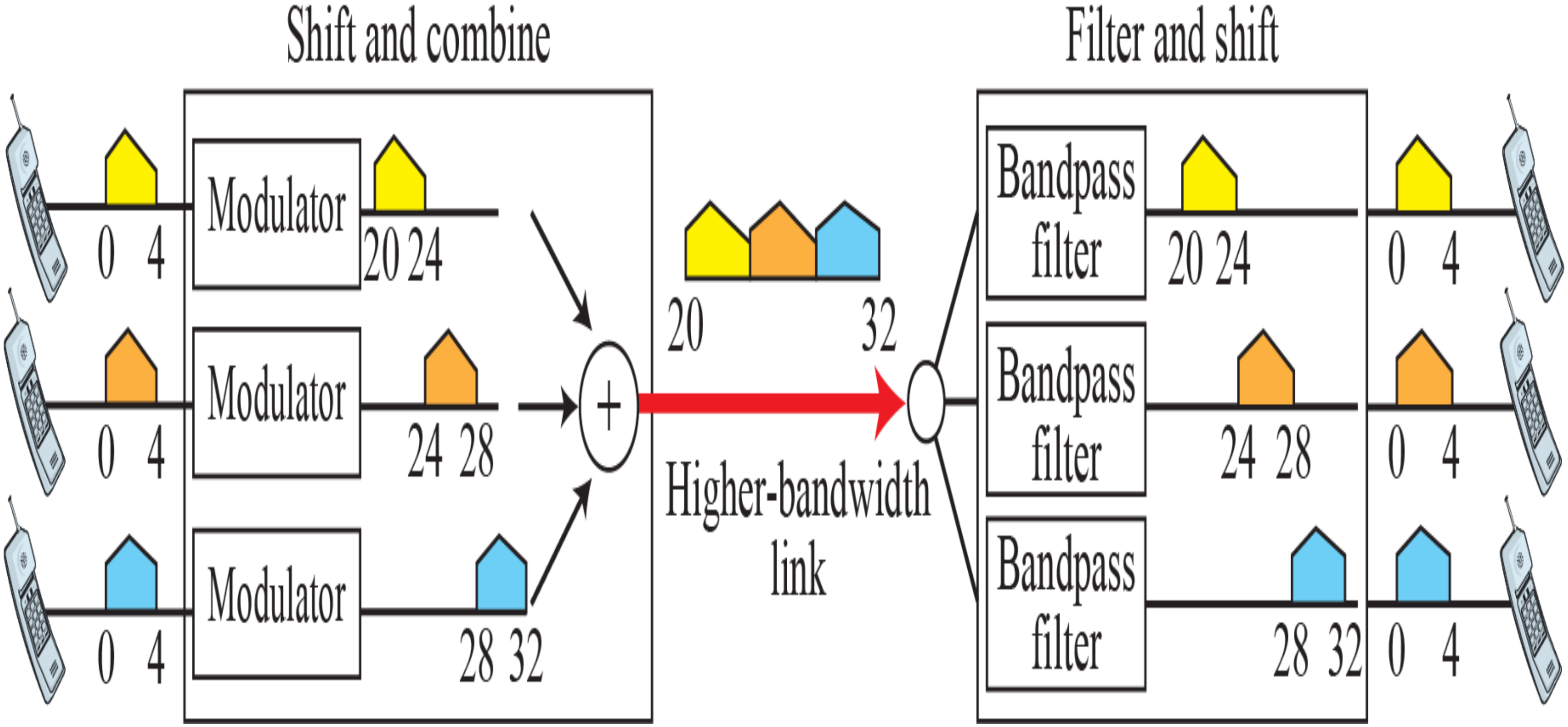
Example 6.1

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

→Shift (modulate) each of the three voice channels to a different bandwidth, as shown in the Figure below

Example 6.1


Shift and combine
Filter and shift
0 4
0 4
0 4
Modulator
Modulator
Modulator
20 24
24 28
28 32
20
32
Higher-bandwidth link
Bandpass filter
Bandpass filter
Bandpass filter
20 24
24 28
28 32
0 4
0 4
0 4

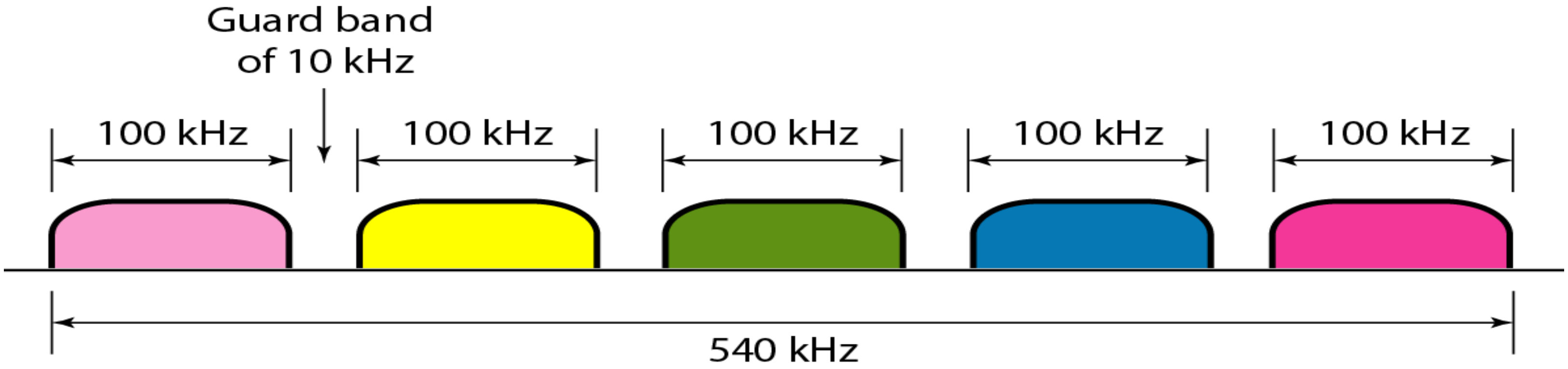
Example 6.2

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, at least four guard bands are required. This means that the required bandwidth is

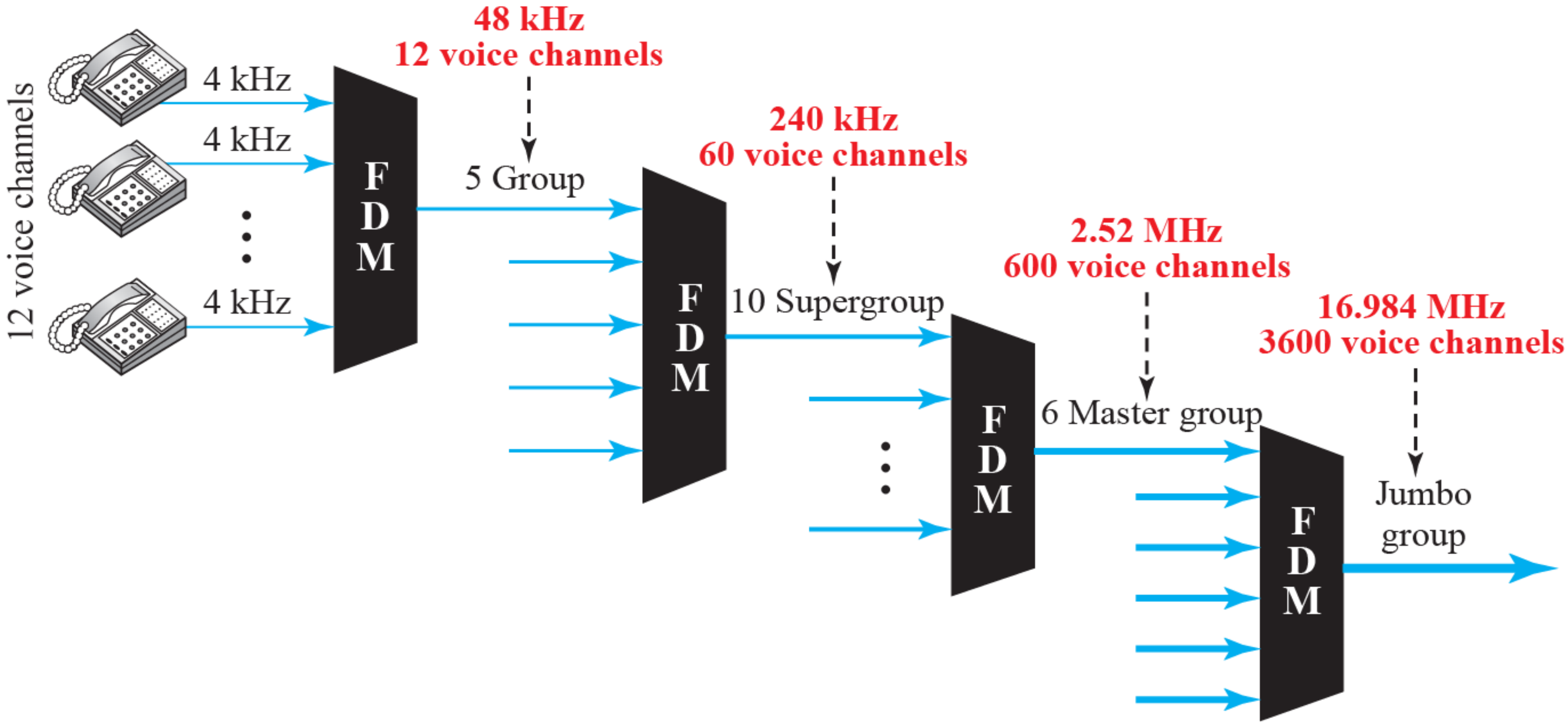
5 × 100 + 4 × 10 = 540 kHz, as shown below



Analog carrier System

- Telephone companies, to maximize efficiency of their infrastructure, multiplexed signals from lower bandwidth lines into higher-bandwidth lines.
- In analog hierarchy, 12 voice channels are multiplexed onto higher bandwidth line to create a group.
- The group has 48khz of BW and supports 12 voice channels and so on…..

Analog hierarchy



Applications of FDM

1) FDM is used in FM and AM broadcasting.

→ Radio uses air as transmission medium.

→ A special band from 530 to 1700 kHz is assigned to AM radio.

→ All radio channels need to share this band.

- AM stations needs 10kHz of BW. Each station uses different carrier frequency , which means it is shifting its signal and multiplexing.

Continued..

- The signal that goes to air is the combination of signals.
- The receiver receives all the signals but tunes only the station desired.
- In FM broadcasting, FM has wider band of 88 to 108 MHz because each station needs a BW of 200 KHz.

Uses:

1) FDM is used in television broadcasting. Each TV channel has its own BW of 6 MHz.

2) First generation of cellular telephones.

Example 6.4

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.

Divide 25 MHz by 30 kHz=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.

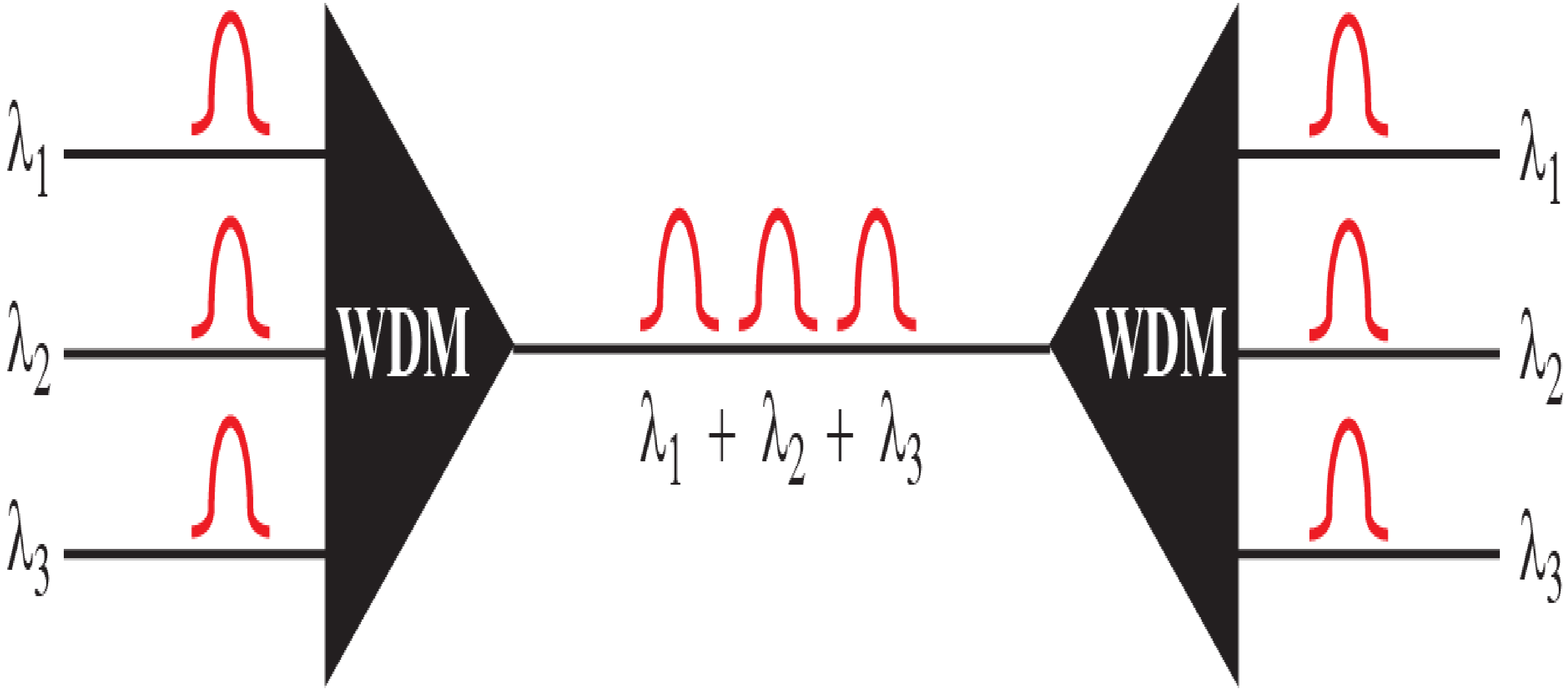
Wavelength-Division Multiplexing

- Wavelength-division multiplexing (WDM) is designed to use the high-data-rate capability of fiber-optic cable.

- The optical fiber data rate is higher than the data rate of metallic transmission cable, but using a fiber-optic cable for a single line wastes the available bandwidth.

- Multiplexing allows users to combine several lines into one.

Wavelength-division multiplexing



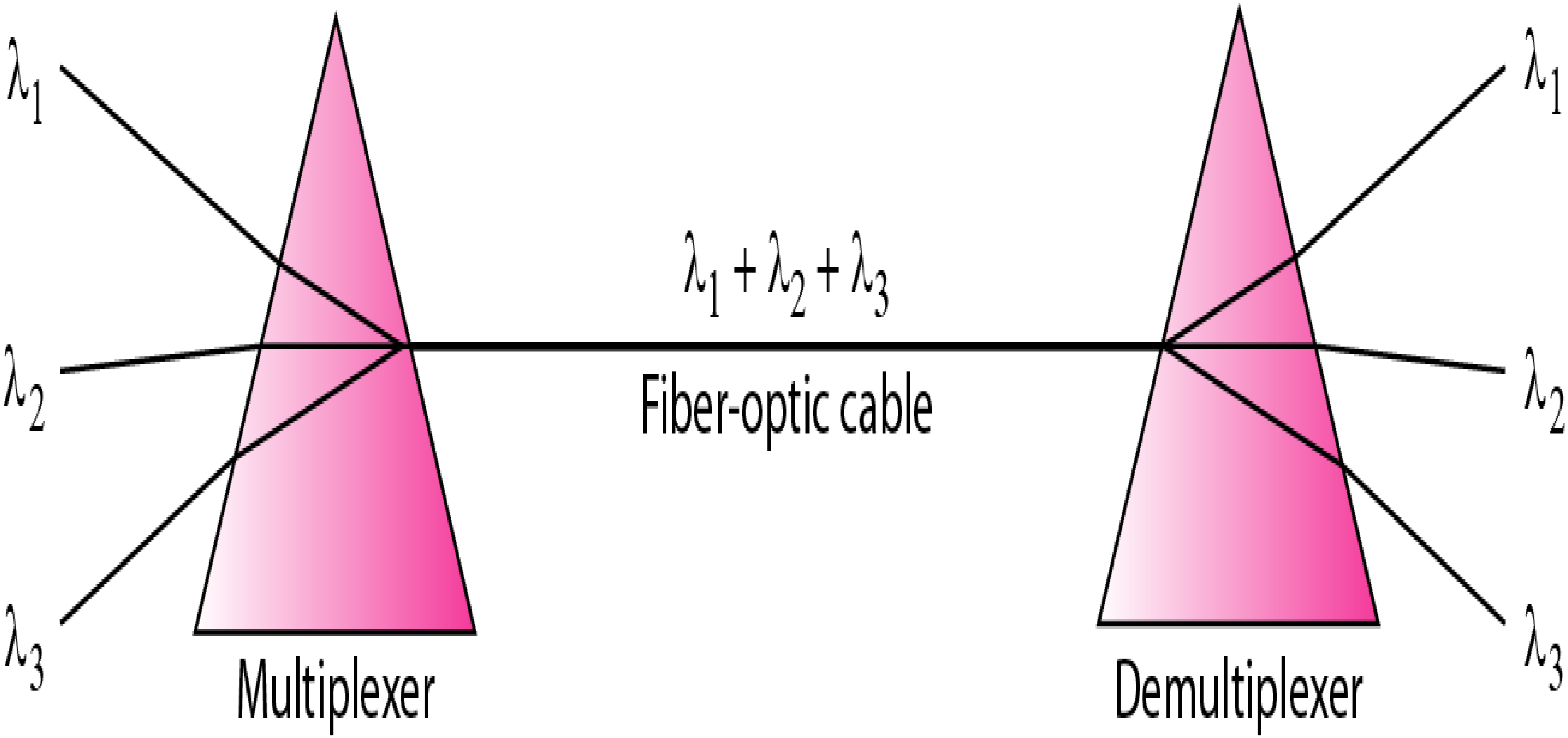
WDM

- Multiplexing and de multiplexing involves optical signals through fiber optic channels.
- In WDM, different high frequency signals of different frequencies are combined.
- A very narrow bands of light from different sources are combined to make a wider band of light.
- Signals are separated from de multiplexer at the receiver.

WDM

- Idea in WDM technology is, combining multiple light sources into one single light at multiplexer.
- *Combining and splitting of light sources are handled by prism.*
- Prism bends a beam of light based on angle of incidence and frequency.
- Using this technique, a multiplexer is made to combine several input beams of light.
- Each containing a narrow band of frequencies, into one output beam of wider band of frequencies.

Prisms in wave-length division multiplexing



Applications of WDM

- Used in SONET network in which multiple optical fiber lines are multiplexed and de multiplexed

- A new method called dense WDM(DWDM) is used to achieve greater efficiency

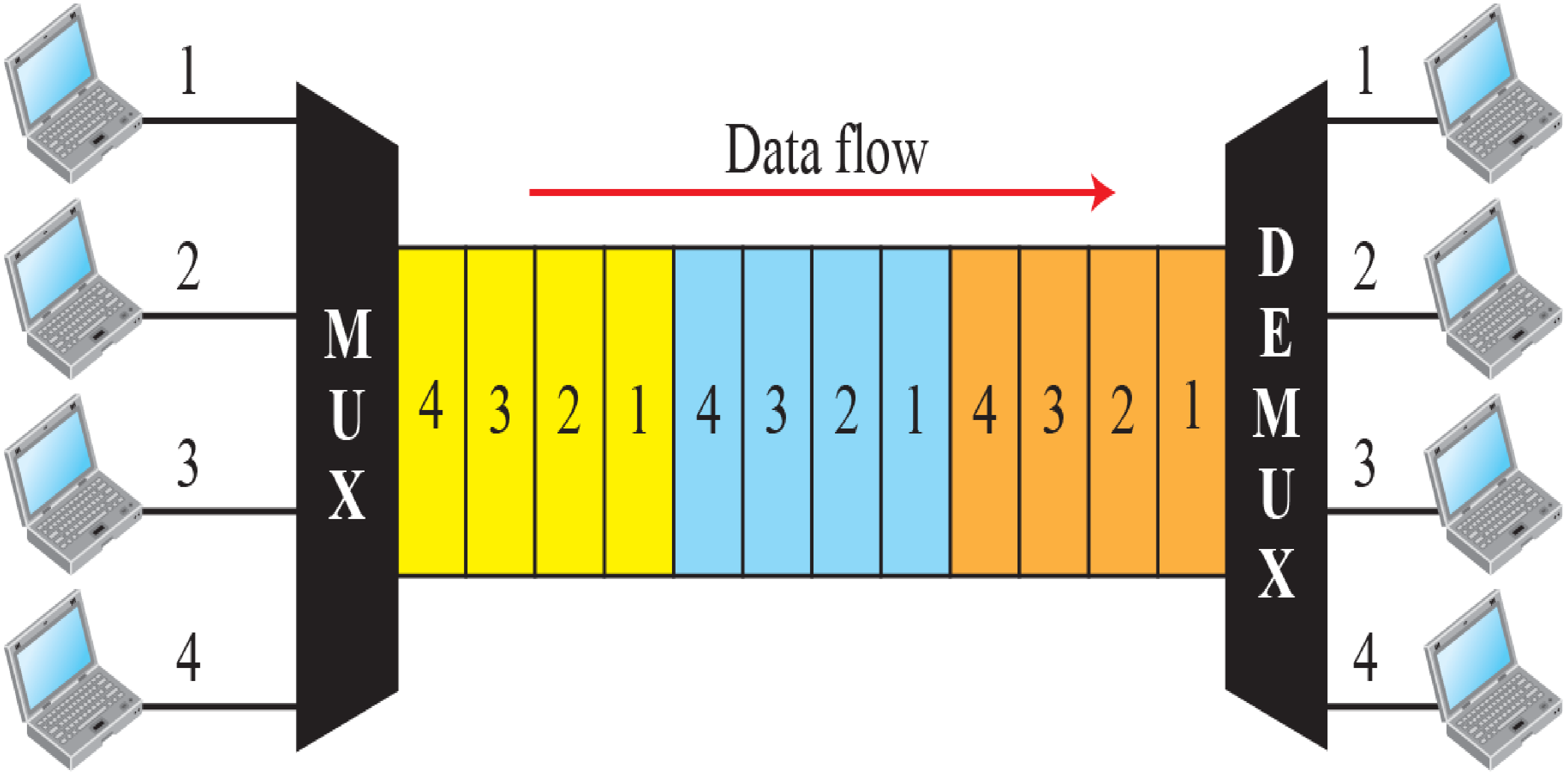
- DWDM can multiplex a very large number of channels by spacing channels very close to one another.

Time-Division Multiplexing

- Time-division multiplexing (TDM) is a digital process that allows several connections to share the high bandwidth of a link.

- Instead of sharing a portion of the bandwidth as in FDM, time is shared.
- Each connection occupies a portion of time in the link.
- Figure shows conceptual view of TDM. The link is sectioned by time rather than by frequency.
- In the figure below , portions of signals 1, 2, 3, and 4 occupy the link sequentially.

TDM



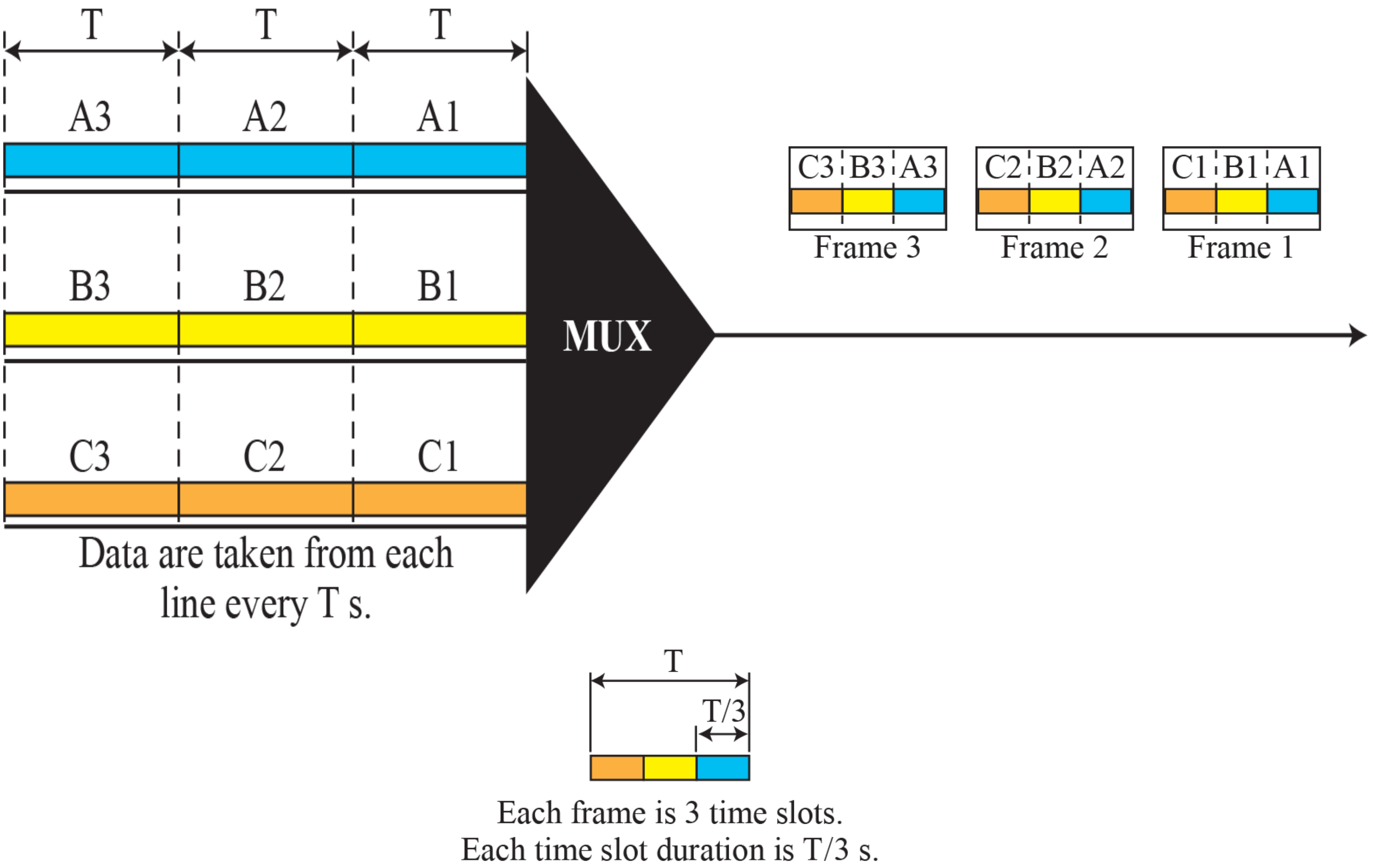
TDM

- In the fig., all the data in a message from source 1 always go to one specific destination 1,2,3 or 4.
- Digital data from different sources are combined into one timeshared link.
- If data is analog, data is sampled changed to digital data and then multiplexed using TDM.
- TDM is divided into Synchronous TDM and statistical TDM.
- In synchronous TDM, the data flow of each input connection is divided into units.
- Each input occupies one input time slot.
- A Unit can be one char, 1 bit or one block of data.

Time slots and frames

- Each i/p unit becomes one o/p unit and occupies one o/p time slot.
- Duration of Output time slot is n times shorter than the duration of an i/p time slot.
- TDM is a digital multiplexing technique for combining several low-rate channels into one high-rate one.
- *In synchronous TDM, the data rate of the link is n times faster, and the unit duration is n times shorter*

Synchronous time-division multiplexing



Data are taken from each line every T s.

Each frame is 3 time slots.
Each time slot duration is T/3 s.

STDM

- In STDM, a round of data units from each i/p connection is collected into a frame.
- If n connections , a frame is divided into n time slots and one slot is allocated for each unit, one for each i/p line.
- If the duration each unit is T sec, then duration of each slot is T/n.
- Duration of each frame is T.
- The data rate of the o/p link must be n times the data rate of a connection to guarantee the flow of data.

STDM

- In fig, the data rate of link is 3 times the data rate of a connection.
- Duration of a unit on a connection is 3 times that of the time slot.
- Time slots are grouped into frames.
- A frame consists of one complete cycle of time slots, with one slot dedicated to each sending device.

Example 6.5

In Figure above, 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

a. each input slot,

b. each output slot, and

c. each frame?

Example 6.5 (continued)

Solution

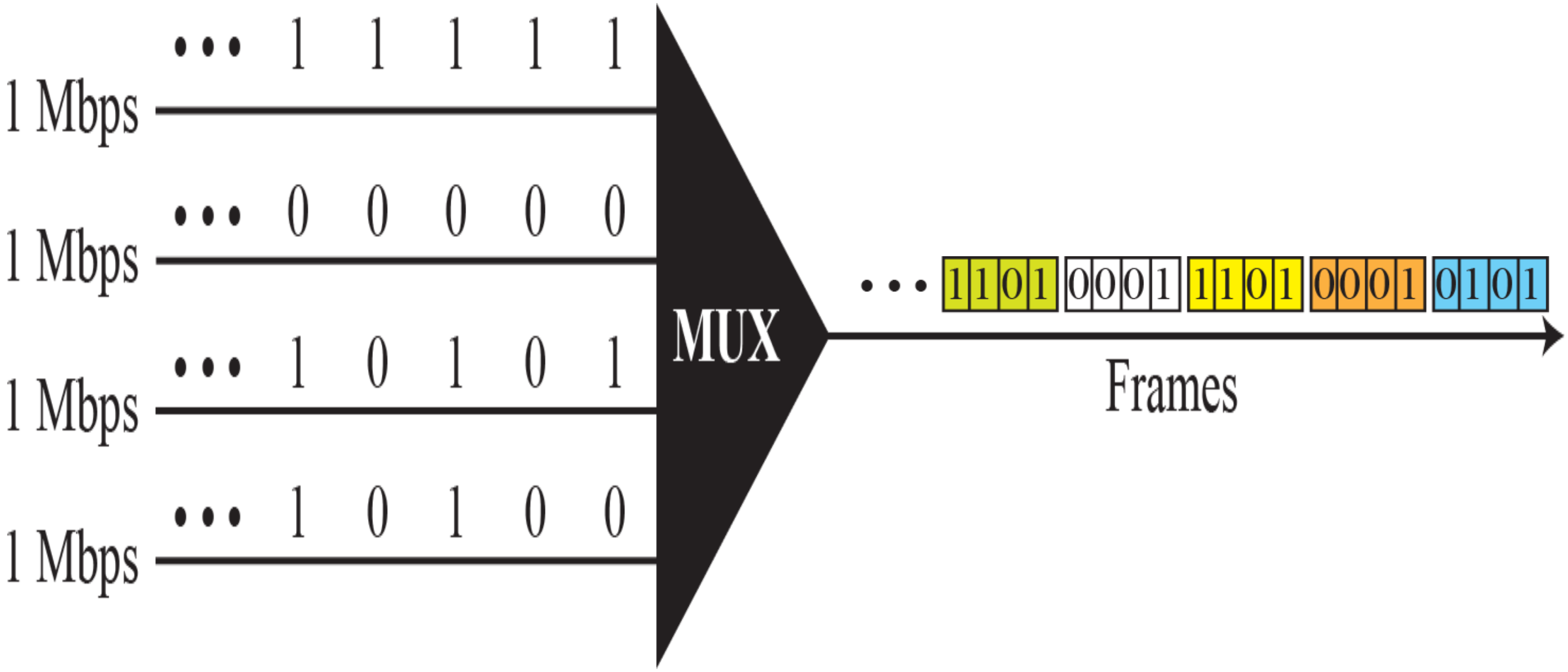
a. The data rate of each input connection is 1 kbps. This means that the bit duration is 1/1000 s or 1 ms. The duration of the input time slot is 1 ms (same as bit duration).

b 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.

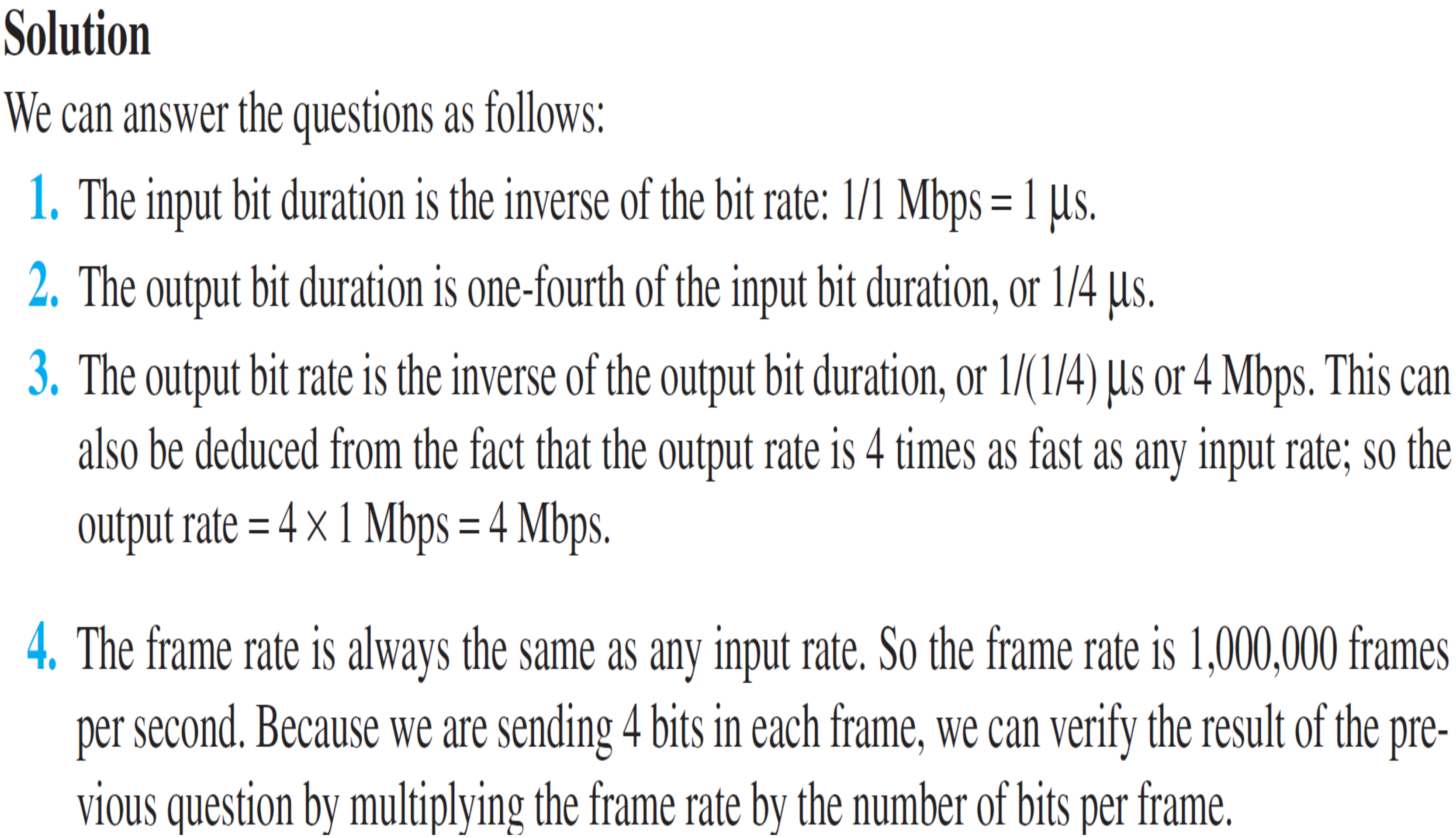
c.Each frame carries three output time slots. So the duration of a frame is 3 × (1/3) ms, or 1 ms. The duration of a frame is the same as the duration of an input unit.

Example 6.6

Figure below 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 (a) the input bit duration, (b) the output bit duration, (c) the output bit rate, and (d) the output frame rate.



Example 6.6



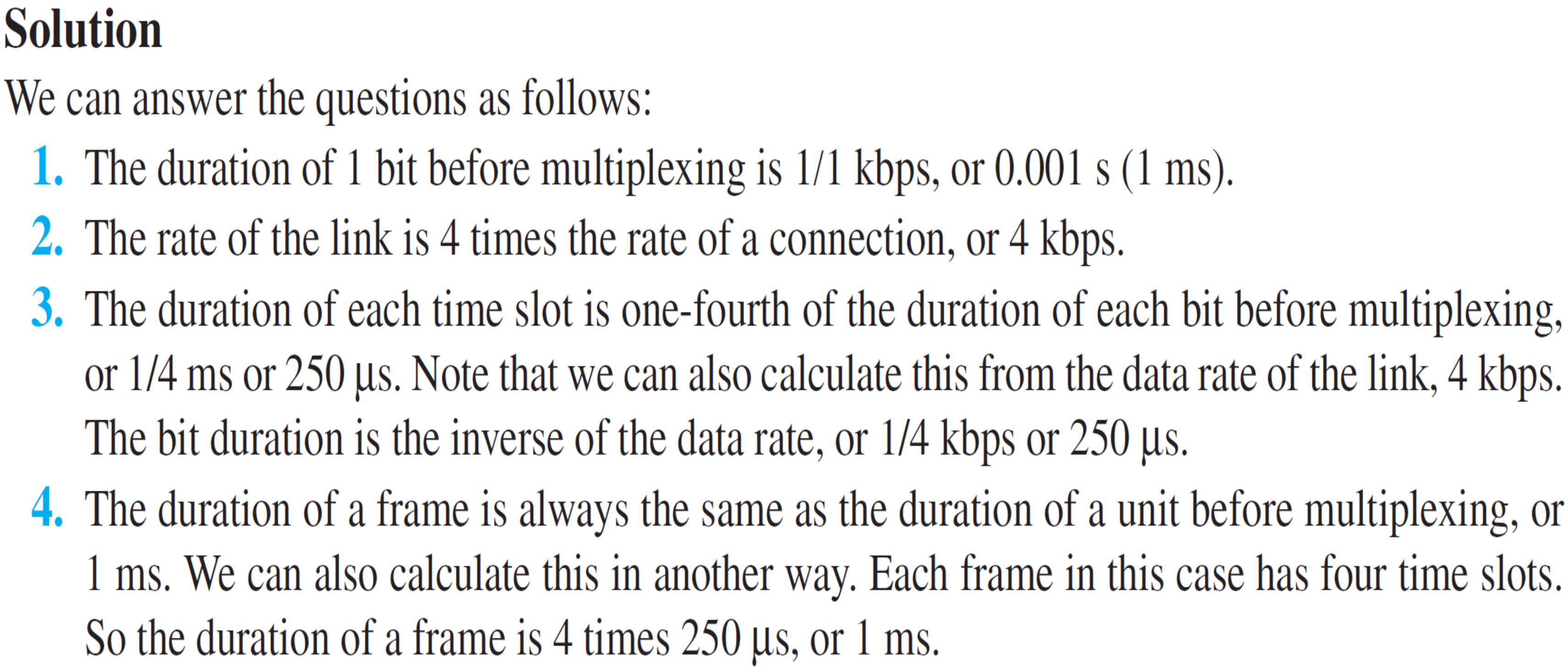
Solution

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/(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 6.7

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

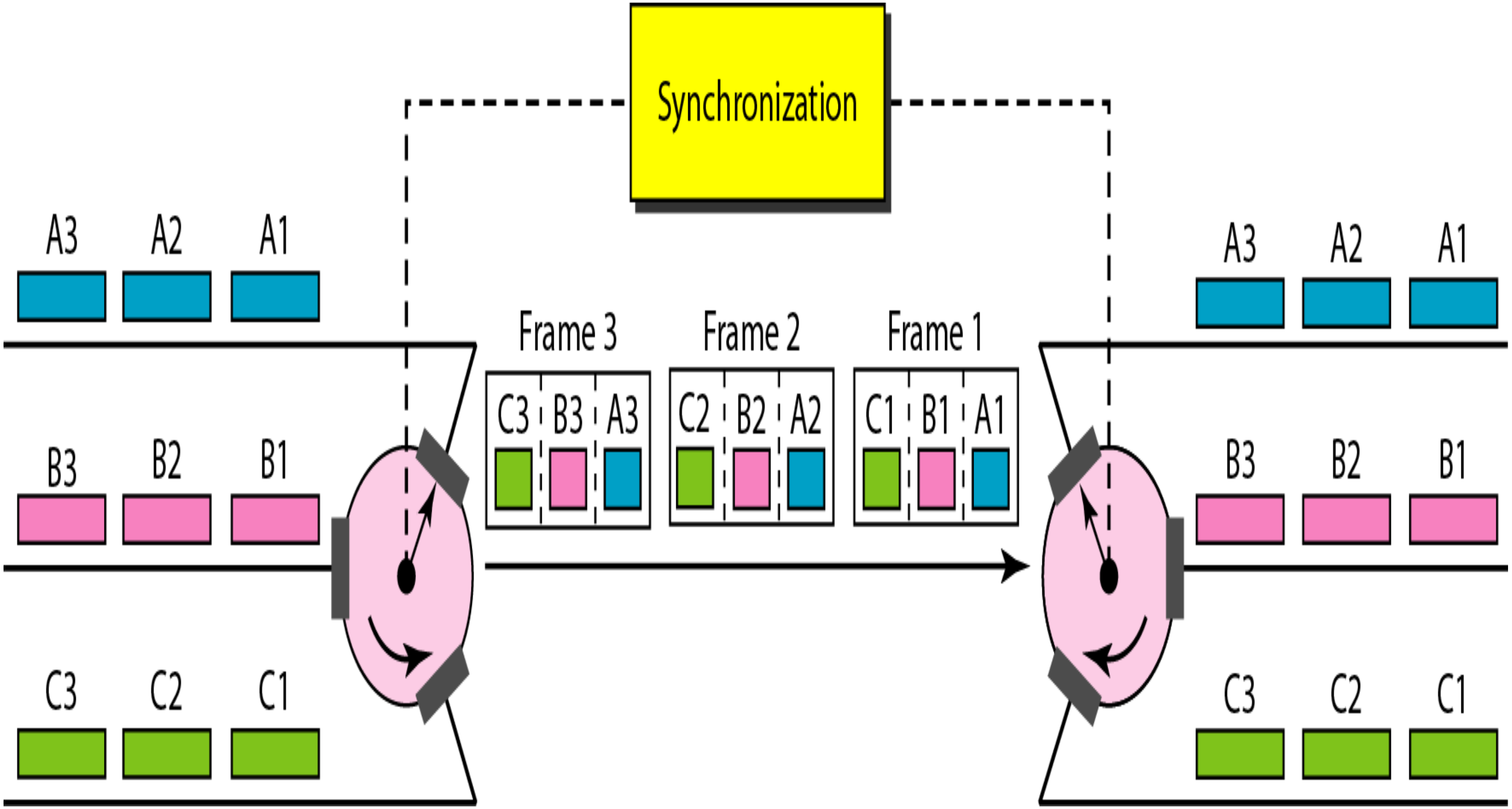
We can answer the questions as follows:

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.

Interleaving

- TDM can be visualized as two fast rotating switches, one on multiplexing side and other on de multiplexing side.
- The switches are synchronized and rotate at the same speed but in opposite direction.
- At the multiplexing side when switch opens in front of connection, the connection has an opportunity to send a unit onto the path.
- This process is called interleaving.
- At the de-multiplexing side when switch opens in front of connection, the connection has an opportunity to receive a unit from the path

Interleaving



Example 6.8

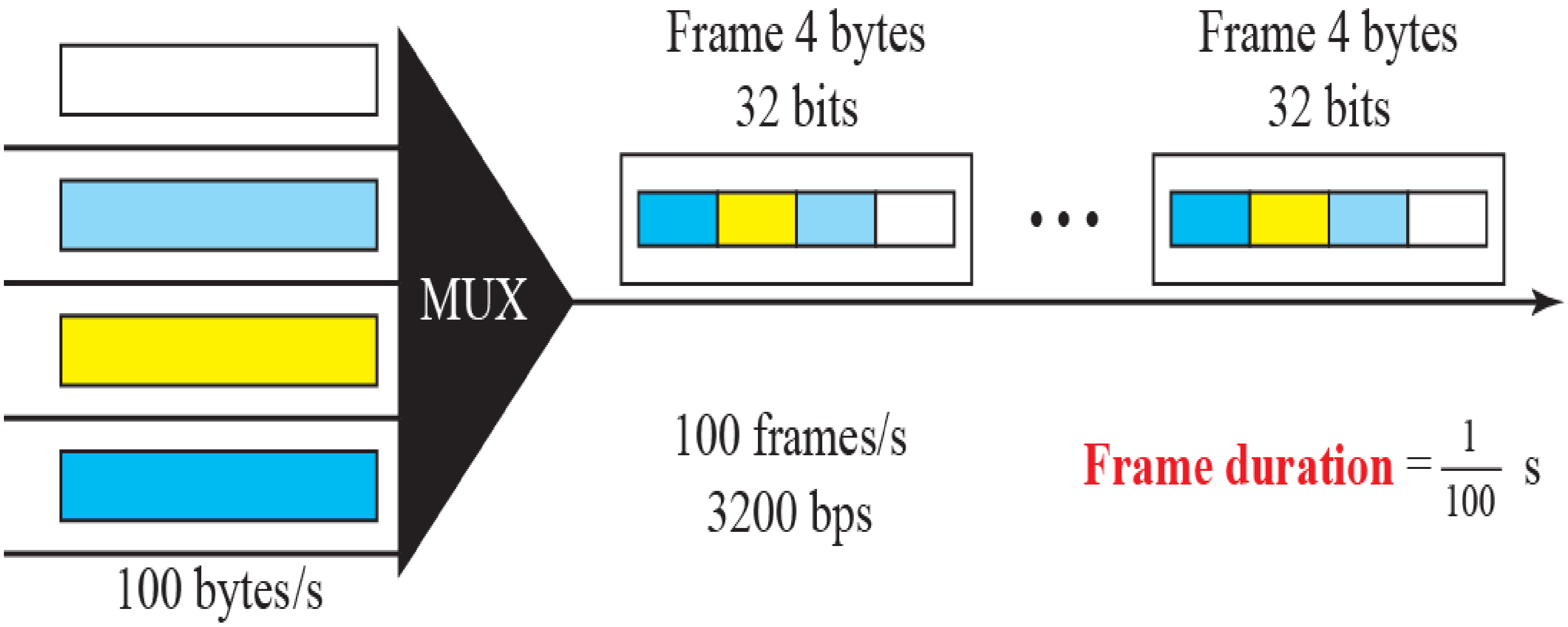
Four channels are multiplexed using TDM. If each channel sends 100 bytes/s and we multiplex 1 byte per channel, show the frame traveling on the link, the size of the frame, the duration of a frame, the frame rate, and the bit rate for the link.

Solution

The multiplexer is shown in the Figure below.

- Each frame carries 1 byte from each channel;
- Size of each frame=4 bytes, or 32 bits.
- The frame rate is 100 frames per second.
- The duration of a frame = 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.

Example



Frame duration $= \frac{1}{100}$ s

Example 6.9

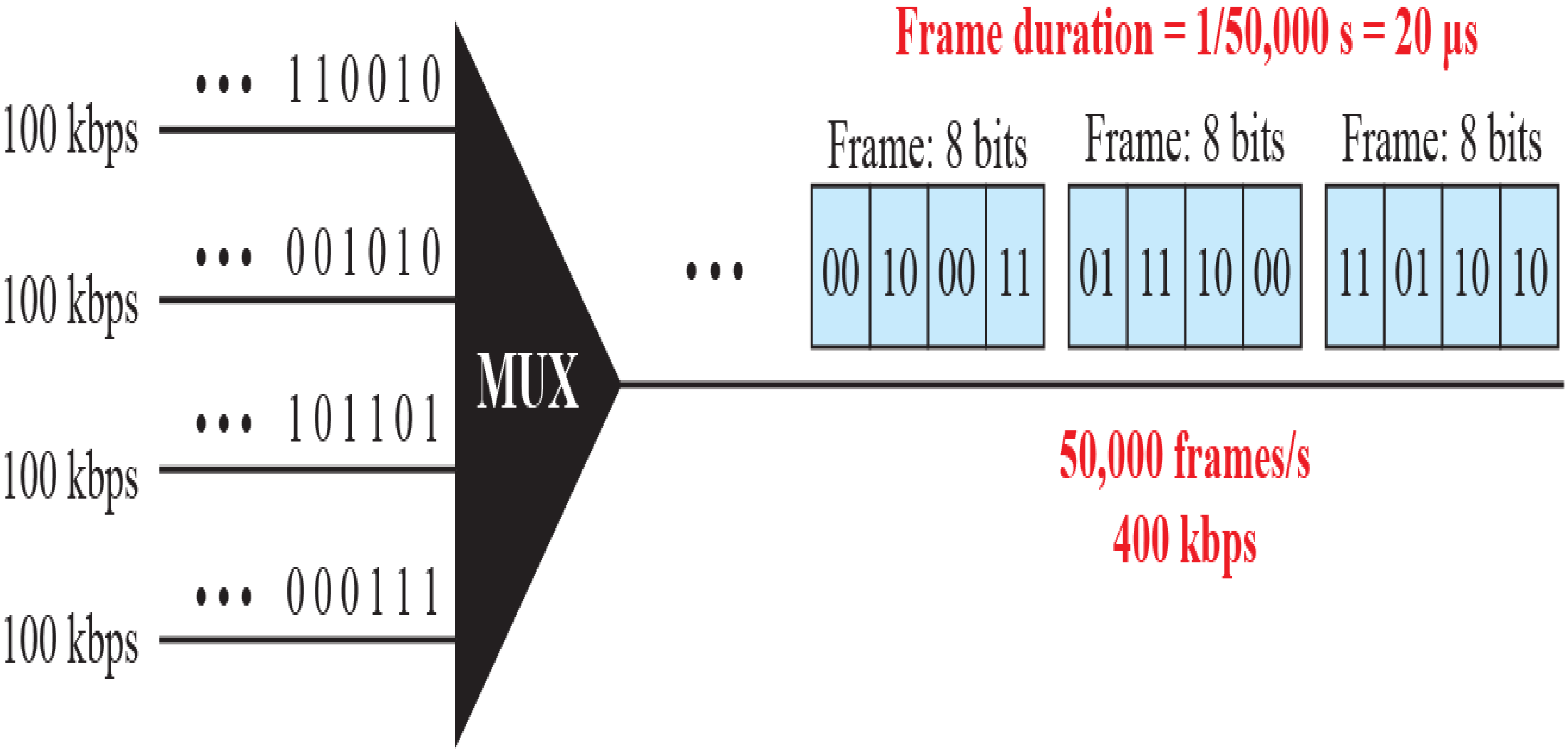
A multiplexer combines four 100-kbps channels using a time slot of 2 bits. Show the output with four arbitrary inputs. What is the frame rate? What is the frame duration? What is the bit rate? What is the bit duration?

Solution:

Figure below shows the output for four arbitrary inputs.

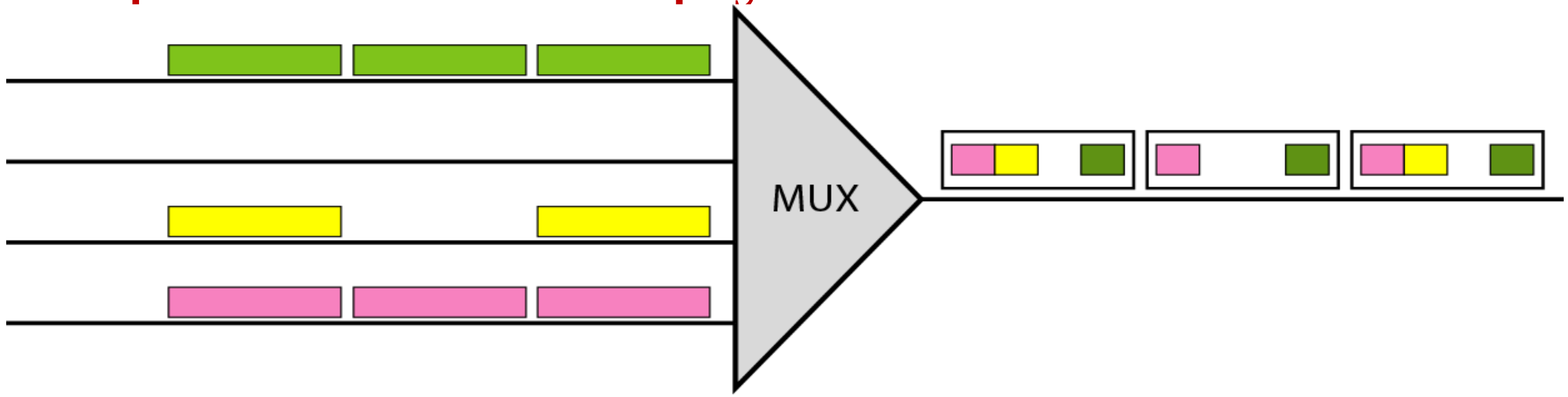
- The link carries 50,000 frames per second since each frame contains 2 bits per channel.
- The frame duration=1/50,000 s or 20 μs.
- The frame rate is 50,000 frames per second.
- Each frame carries 8 bits
- Bit rate = 50,000 ×8 = 400,000 bits or 400 kbps.
- The bit duration is 1/400,000 s, or 2.5 μs.

Example


Frame duration = 1/50,000 s = 20 µs
100 kbps
••• 110010
100 kbps
••• 001010
100 kbps
••• 101101
100 kbps
••• 000111
MUX
•••
Frame: 8 bits
00 10 00 11
Frame: 8 bits
01 11 10 00
Frame: 8 bits
11 01 10 10
50,000 frames/s
400 kbps

Empty slots

- STDM is not very efficient because if source does not have data to send, the corresponding o/p slot will be empty.



- The o/p frame is not full. We can overcome that problem using statistical TDM

Data rate management

- How to handle disparity in the i/p data rates in TDM??
- If i/p data rates are not same, three strategies are used

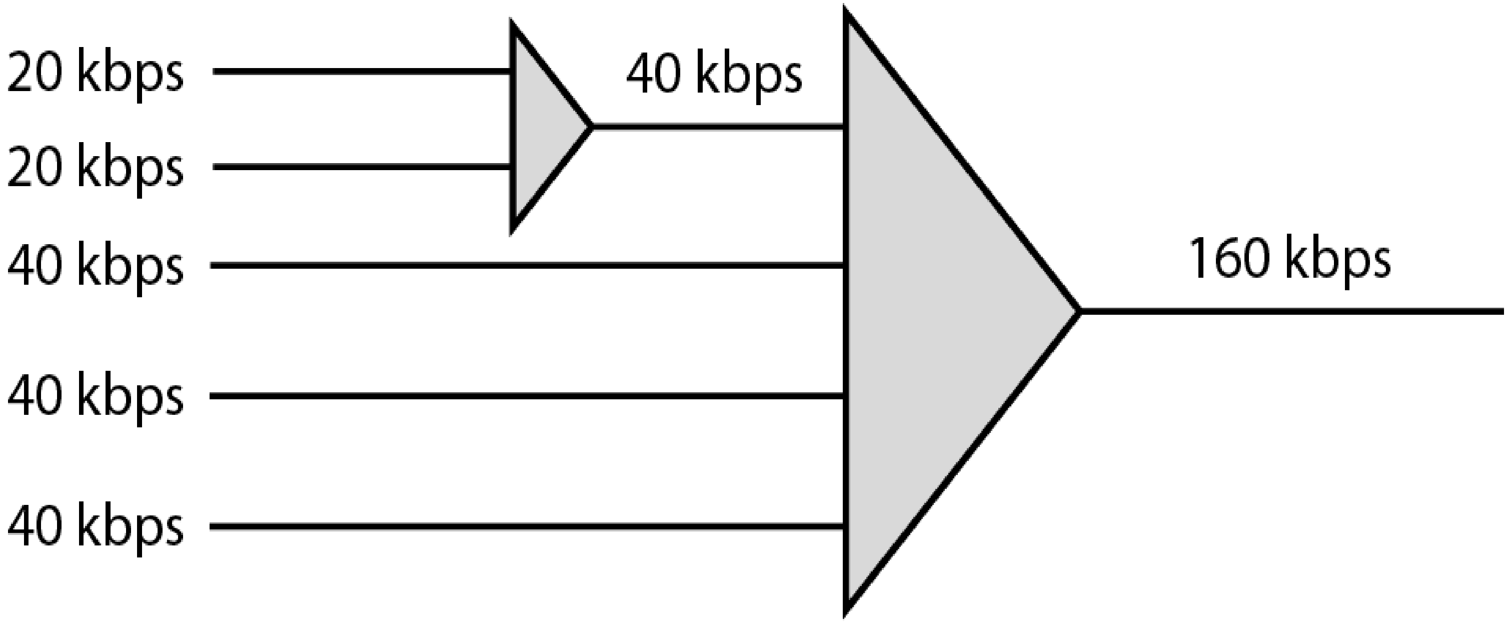
1) Multilevel Multiplexing

2) Multiple-slot allocation

3) Pulse stuffing.

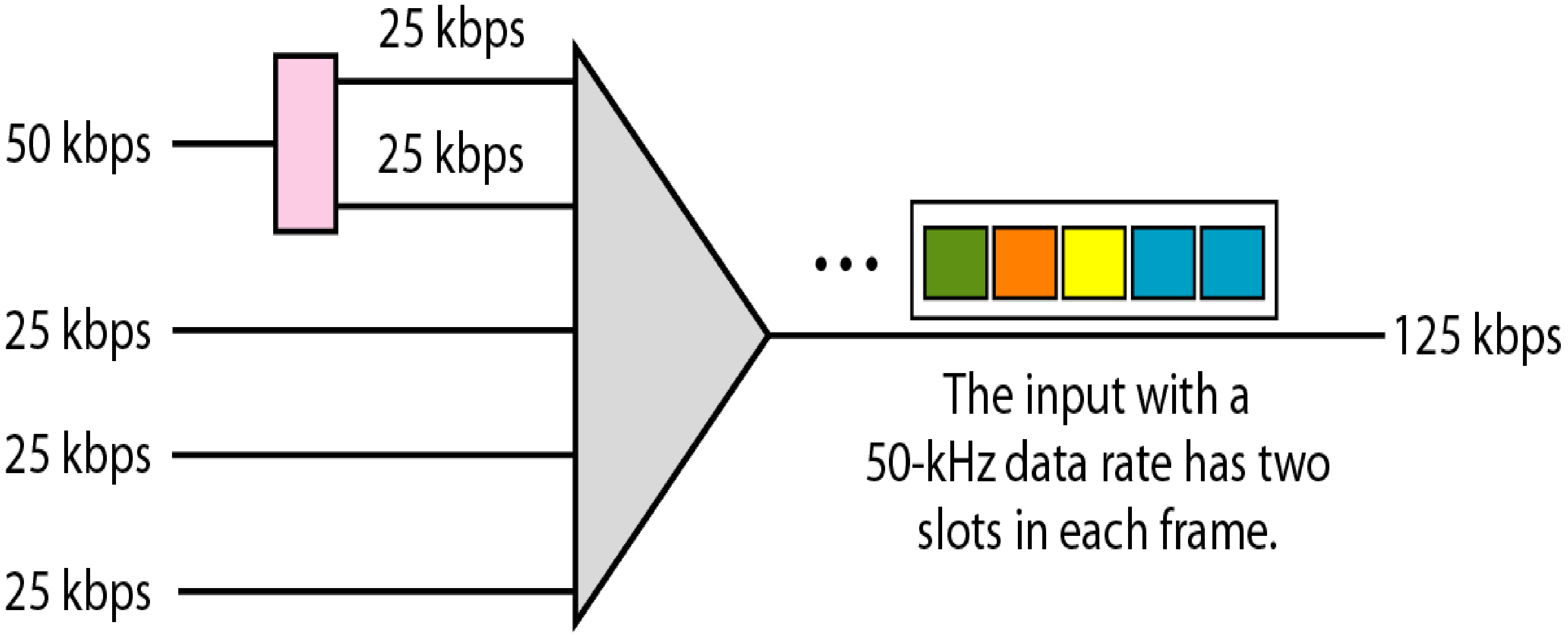
Multi level Multiplexing

- Is used when i/p data rate of an i/p line is multiple of others.



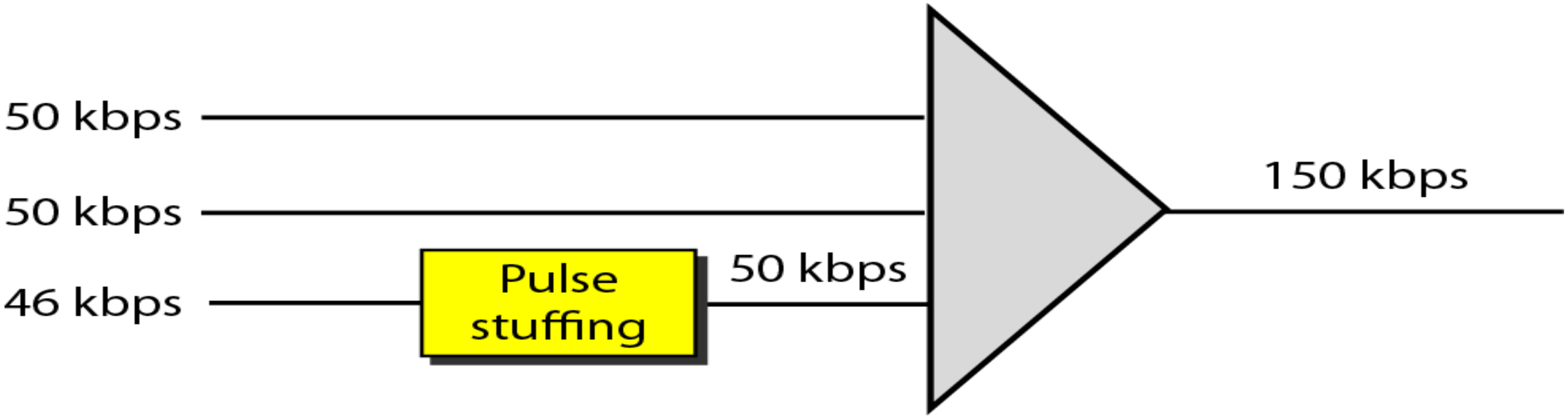
Multiple slot Multiplexing

- Allotting more than one slot in a frame to a single i/p line.



Pulse stuffing

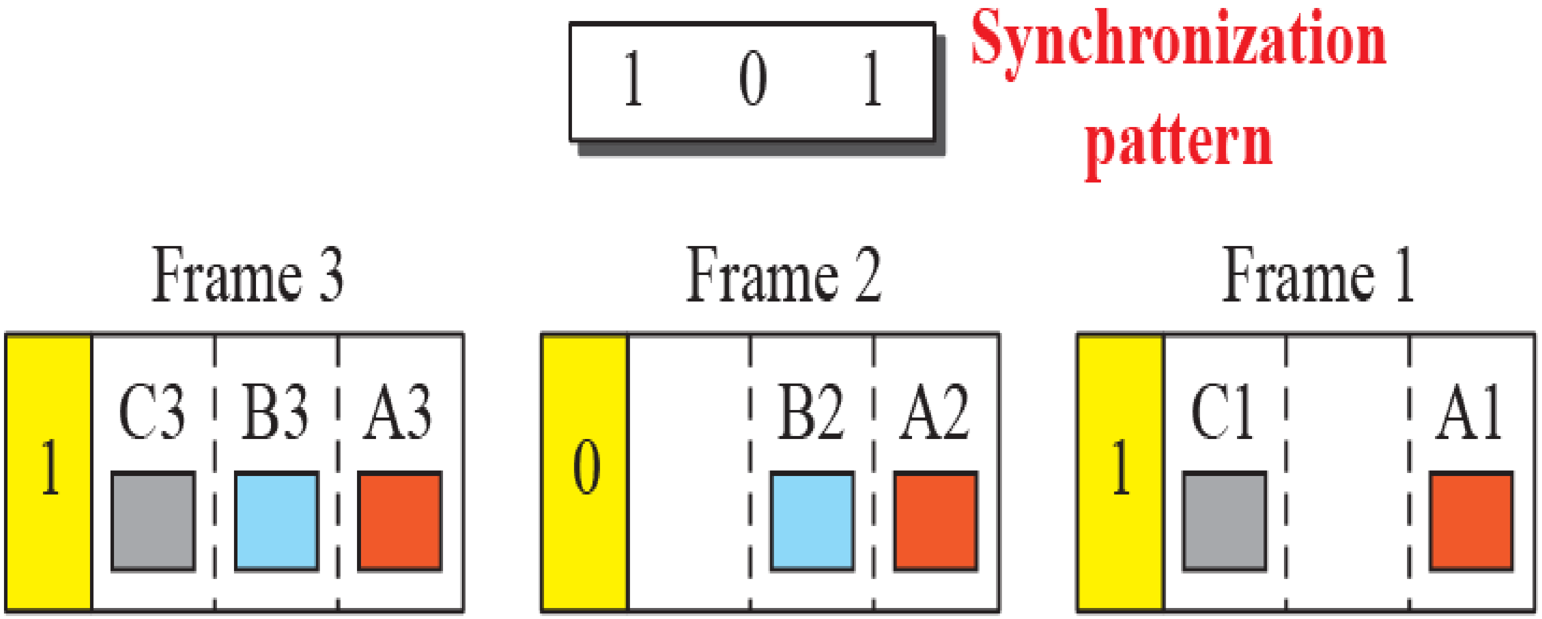
- Some times bit rates of source are not multiple integers of each other.
- Solution is to make highest i/p data rate the dominant data rate and add dummy bits to the i/p lines with lower rates.



Frame synchronizing

- In TDM synchronization between multiplexer and de multiplexer is a major issue.
- If not synchronized, the bits belonging one channel will be received by wrong channel.
- Therefore, one or more synchronization bits are added to the beginning of each frame,
- That allows de multiplexer to synchronize with the incoming stream , so that it can separate the time slots accurately.
- The Synchronization information consists of 1 bit per frame , alternating between 0 and 1.

Framing bits



Example 6.10

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

1. The data rate of each source is 250 × 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.

(Continued)

3. Each frame has one character from each source, which means the link needs to send 250 frames per second.

4. The duration of each frame is 1/250 s, or 4 ms.

5. Each frame carries 4 characters and 1 extra synchronizing bit. This means that each frame is 4 × 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.

Example 6.11

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

→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 × 3 bits per frame, or 300 kbps.

Applications

1) Second generation cellular telephone companies.

Statistical TDM

Limitations of Synchronous TDM are

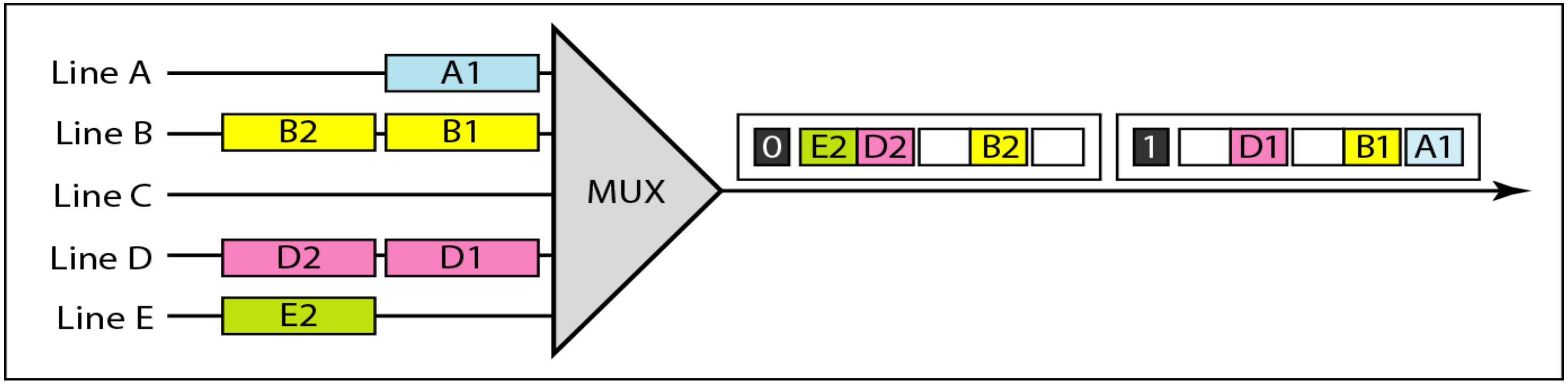
1) I/P time slots are wasted if some of the i/p lines does not have data to send.
2) Clock synchronization is critical.
3) In Statistical TDM, slots are dynamically allocated to improve the bandwidth efficiency.
4) Only when an i/p line has a slot worth of data to send, it is given a slot in the o/p frame.

Statistical TDM

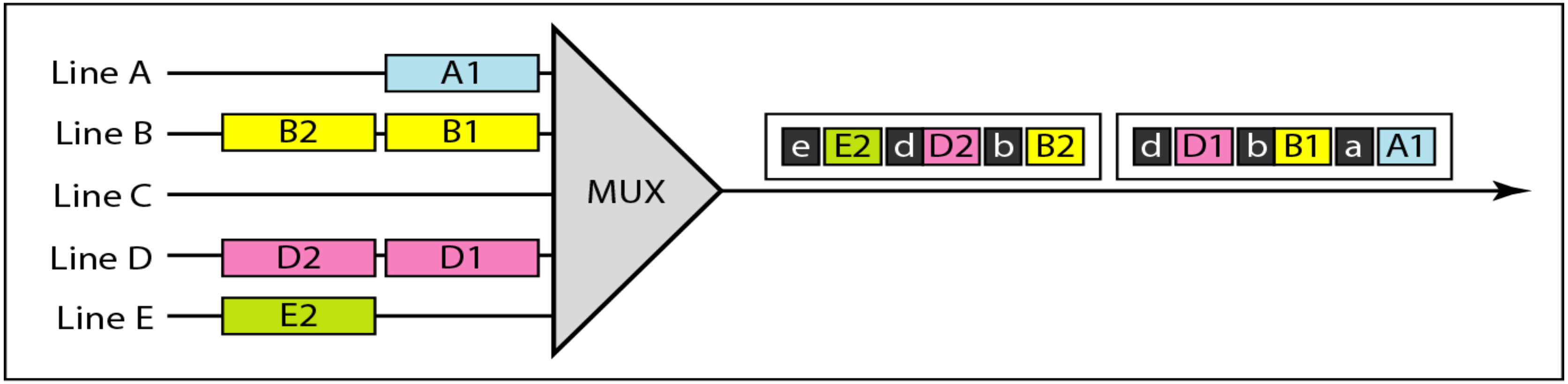
- In Statistical TDM, the no. of slots in each frame is less than number of i/p lines.
- The multiplexer checks the i/p lines in round robin fashion, and allocates slot for the i/p line only if it has data to send.

    if no data, it skips and checks the next line.
- In the fig 1. some slots are empty and therefore no data to send.
- Fig.2 no slots are left empty as long as there are data to be send by any i/p line.

TDM slot comparison



a. Synchronous TDM



b. Statistical TDM

Addressing: Differences

Synchronous TDM

1.O/P slot is occupied by data

2.No need of addressing : Synchronization and pre assigned relationships between the i/p and o/p's serve as address.

Statistical TDM

1.Slot needs to carry data as well as address of destination

2.No fixed relationship b/w i/p's and o/p's because there are no pre assigned or reserved slots. Need to include address of receiver inside the slot

Statistical TDM

- Slot size: Slots carry both data and address. There fore the ratio of data size to address size must be reasonable to make transmission efficient.
- No need of Synchronization bit.
- The capacity of channel is normally less than sum capacity of each channel.
- Capacity of link is based on statistics of load for each channel

SPREAD SPECTRUM

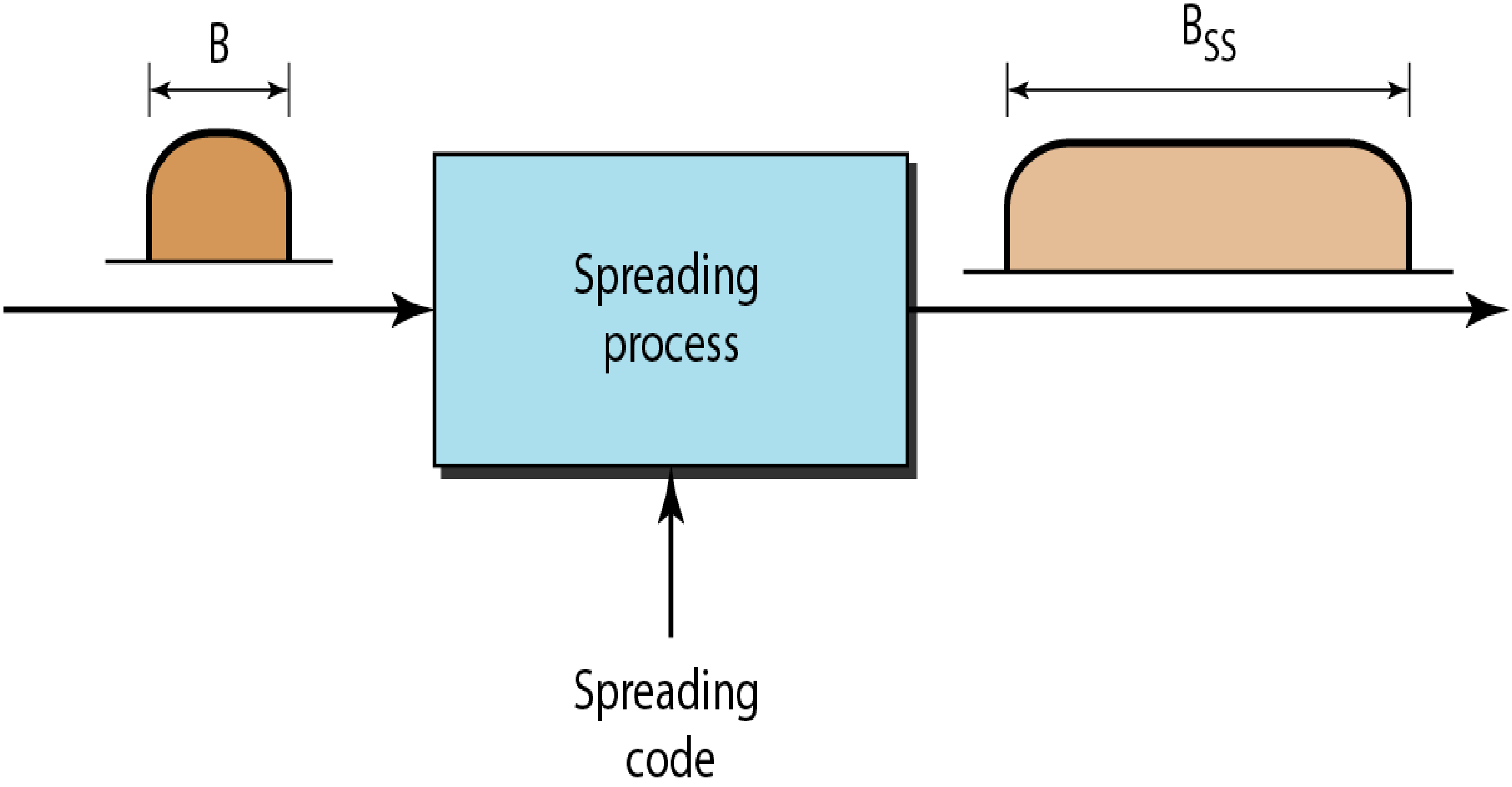
- In wireless applications, stations must be able to share this medium without interception by an eavesdropper and without being subject to jamming from a malicious intruder.

- To achieve these goals, spread spectrum techniques add redundancy

Spread spectrum

- Multiplexing combines signals from several sources to achieve BW efficiency.
- Spread spectrum (SS) also combines signals from different sources to fit into larger BW.
- SS is designed to be used in wireless applications.
- As all stations use air as the medium in wireless applications, stations must share this medium without interception by eaves droppers and jamming from malicious intruders.

Spread spectrum



Spread spectrum

- The bandwidth allocated to each station needs to be larger than what is needed. This allows redundancy.
- The expanding of original B to BW BSS must be done by a process that is independent of original signal.
- After the signal is created by the source, the spreading process uses a spreads code and spreads the BW.
- Two techniques to spread the badwidth

  1. Frequency hopping Spread Spectrum(FHSS)
  2. Direct sequence Spread Spectrum (DSSS)

FHHS

- Uses M different frequencies that are modulated by source signal.

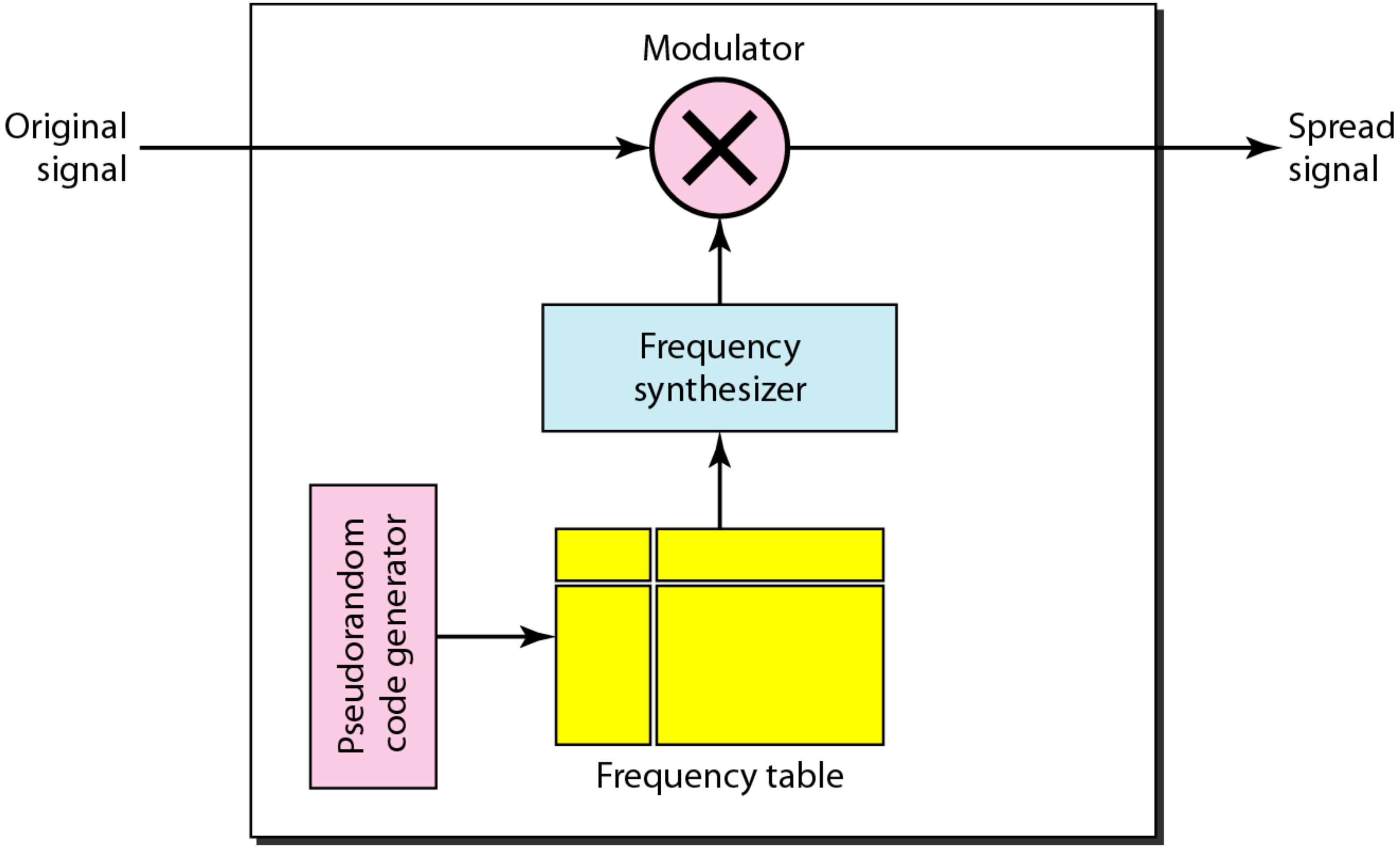
- Signal modulates carrier frequencies one after another.

- Modulation is done using one carrier frequency at a time and M frequencies are used.

- Bandwidth occupied by a source after spreading

$$B_{FHSS} > B$$

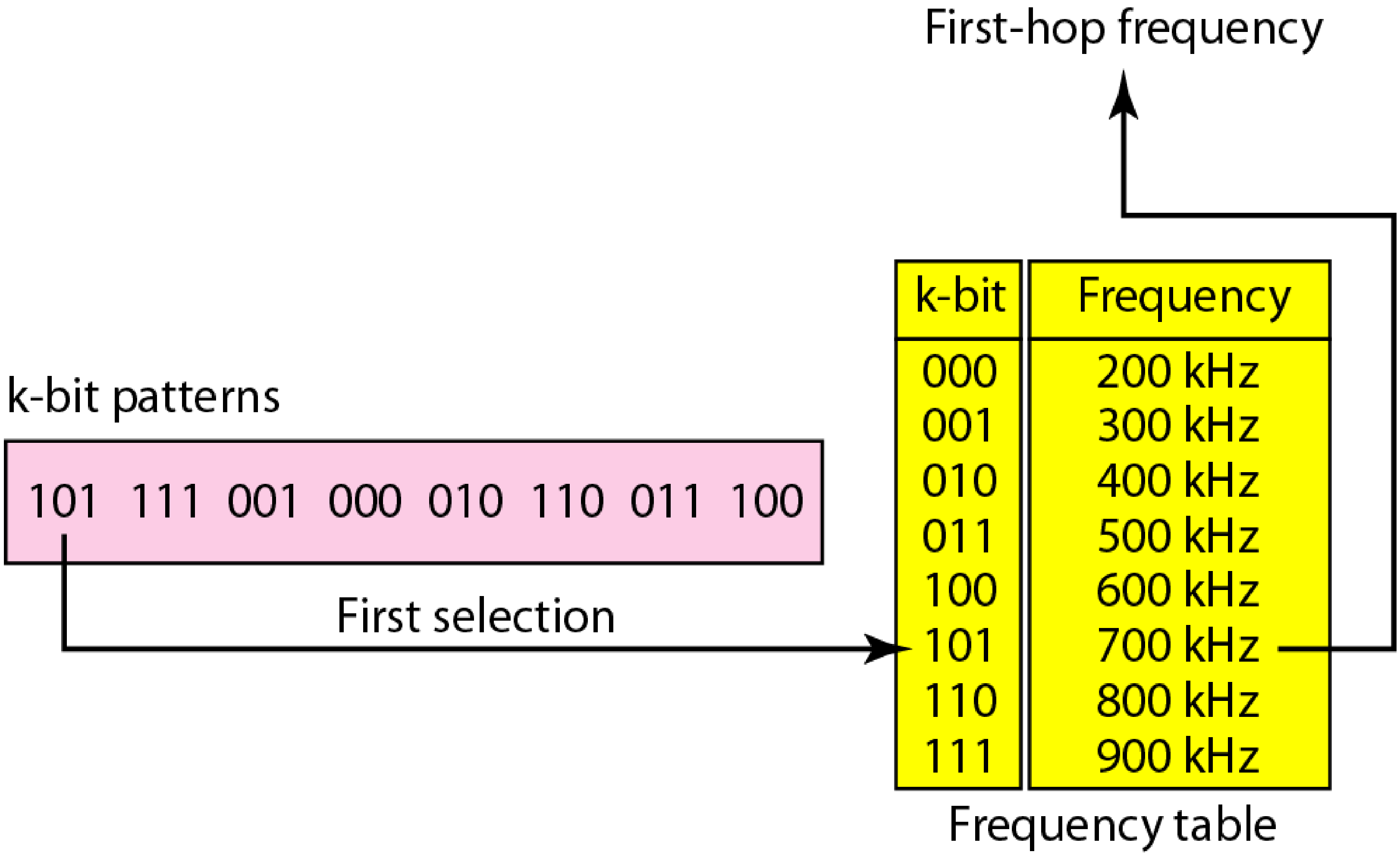
Frequency Hopping Spread Spectrum(FHSS)



Frequency Hopping Spread Spectrum(FHSS)

- Pseudo random code generator called pseudorandom noise(PN), creates K bit pattern for every hopping period T_h
- The frequency table uses a pattern to find the frequency to be used for this hopping period and passes it to frequency synthesizer.
- The frequency synthesizer creates a carrier signal of that frequency and
- source signal modulates the carrier frequency.

Frequency selection in FHSS



First-hop frequency

k-bit patterns

101 111 001 000 010 110 011 100

First selection

k-bit	Frequency
000	200 kHz
001	300 kHz
010	400 kHz
011	500 kHz
100	600 kHz
101	700 kHz
110	800 kHz
111	900 kHz

Frequency table

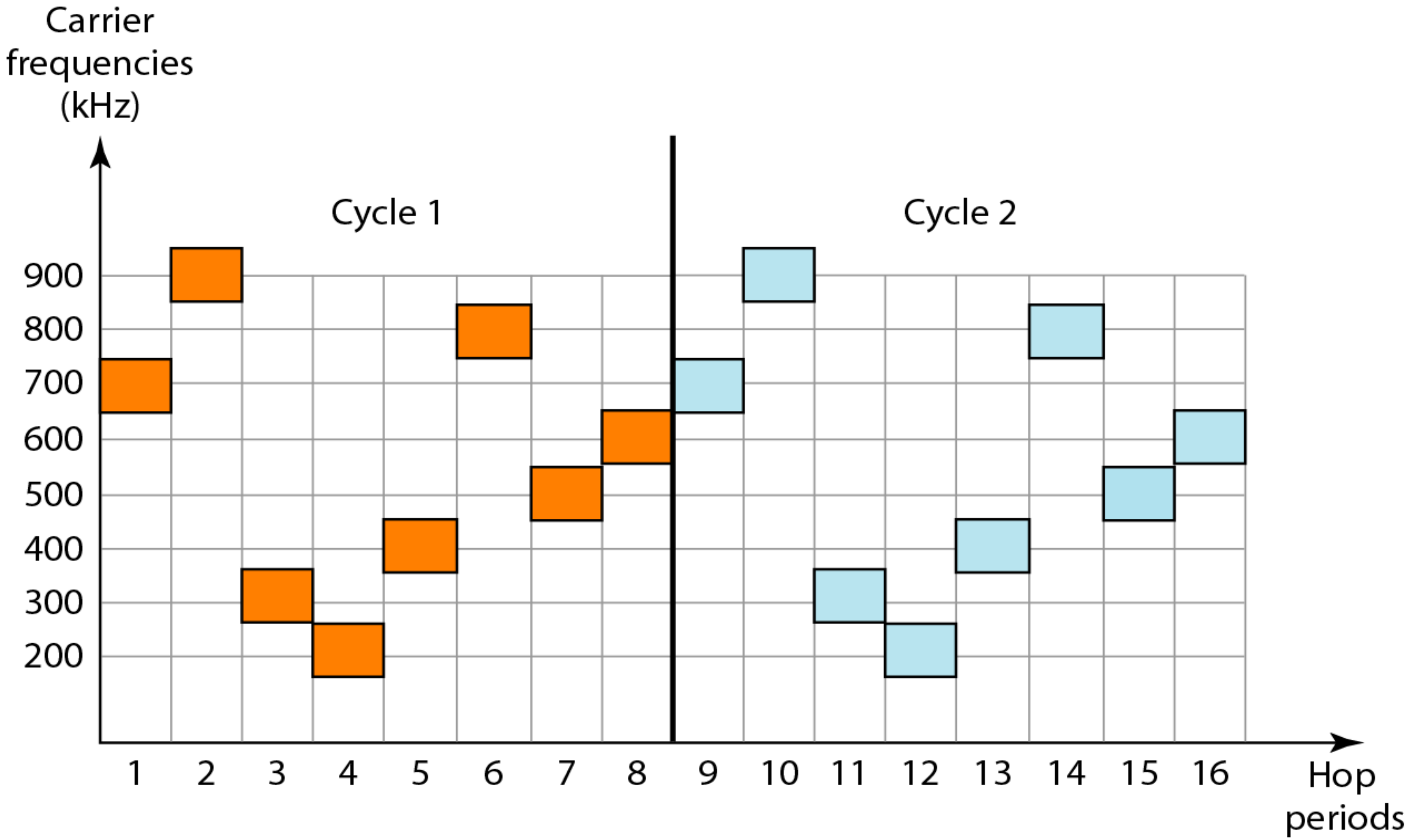
Frequency selection in FHSS

- Assume we have 8 hopping frequencies

  M=8 and K=3.

- Pseudo random generator will create 8 different 3 bit patterns.
- These are mapped to 8 different frequencies in frequency table.
- The pattern Pseudo random and it is repeated after 8 hoppings .

FHSS cycles


Carrier frequencies (kHz)
Cycle 1
Cycle 2
900
800
700
600
500
400
300
200
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Hop periods

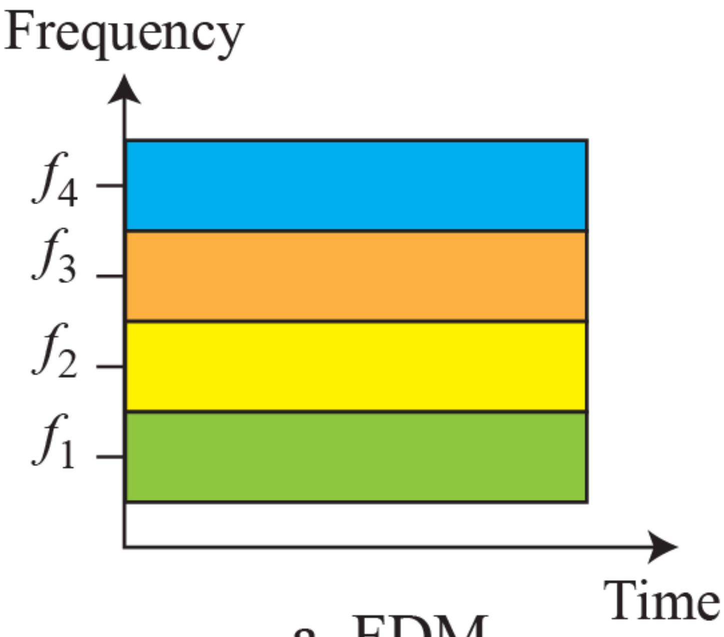
FHSS

- If there are k-bit patterns and hopping period is short, a sender and receiver can have privacy.
- If an intruder tries to intercept the transmitted signal, she can access only small piece of data because she does not know the spreading sequence to adapt quickly to next hop.
- Scheme also has an anti jamming effect. Malicious intruder may send noise to jam the signal for one hopping period , but not for the whole period.

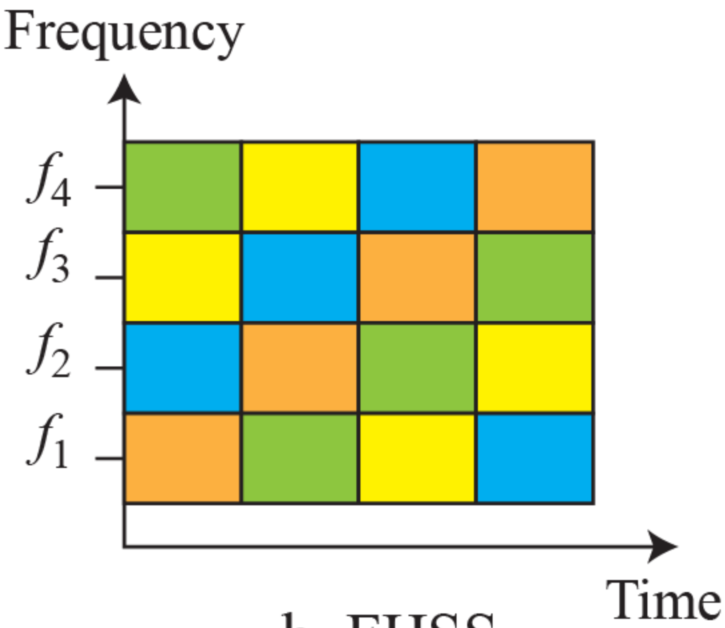
Bandwidth sharing

- If no. of hopping frequencies is M, we can multiplex M channels into one by using same B_{SS} BW.
- This is possible because station uses just one frequency in each hopping period.
- M-1 other frequencies can be used by M-1 stations.
- In FDM, each station uses 1/M of the BW, but allocation is fixed.
- In FHSS, each station uses 1/M of BW, but the allocation changes hop to hop.

Bandwidth sharing



a. FDM



b. FHSS

Direct sequence spread spectrum(DSSS)

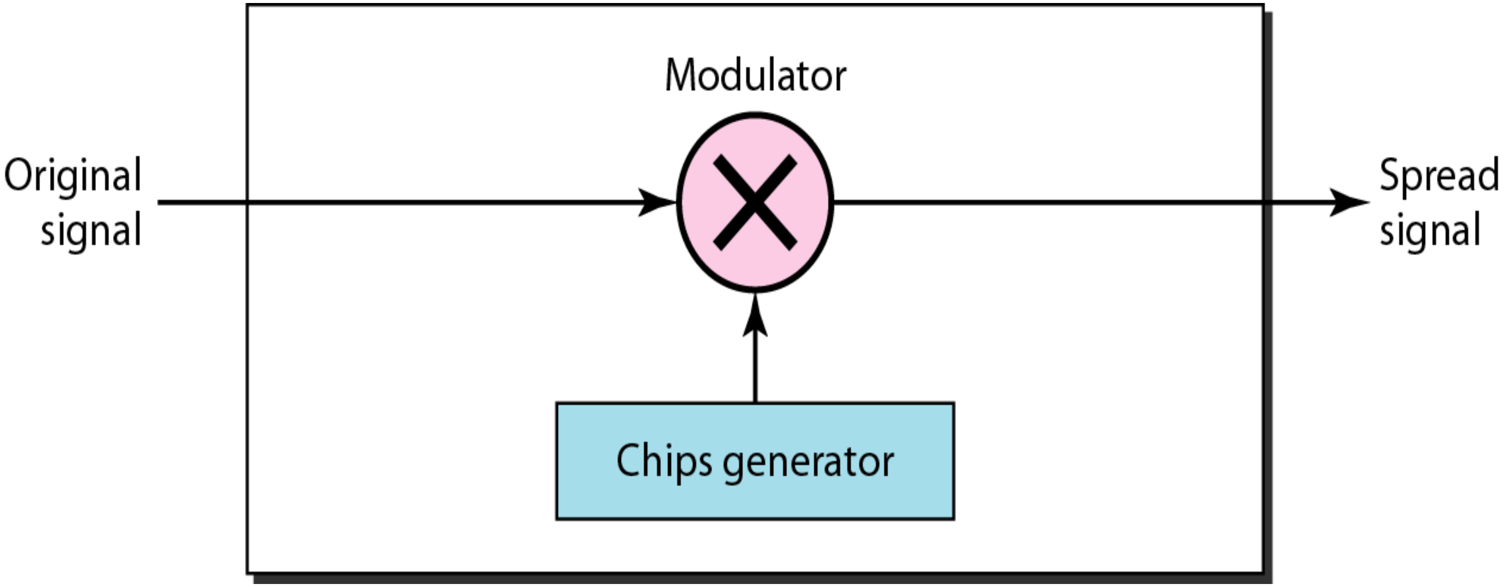
•The direct sequence spread spectrum (DSSS) technique also expands the bandwidth of the original signal, but the process is different.

•In DSSS, each data bit is replaced with n bits using a spreading code.

•In other words, each bit is assigned a code of n bits, called chips, where the chip rate is n times that of the data bit.

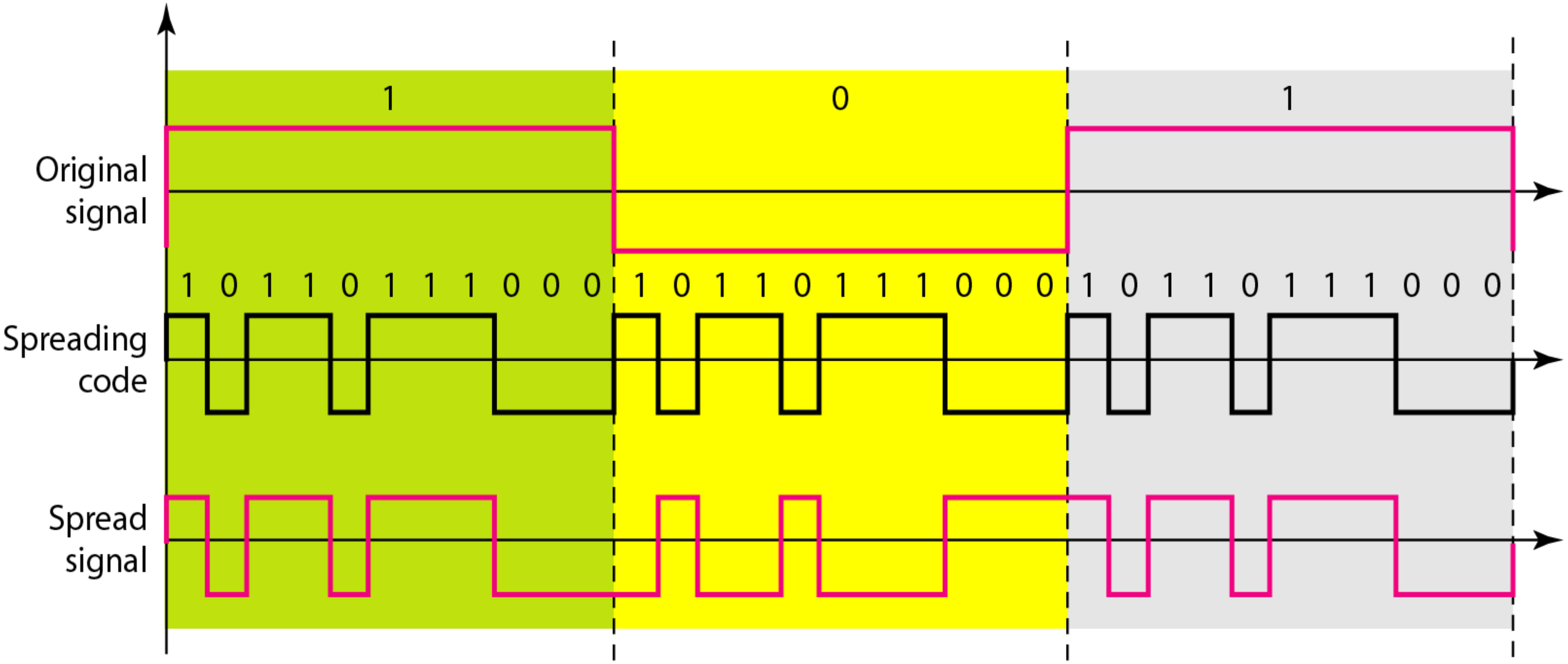
Direct sequence spread spectrum(DSSS)

- Also expands the BW of the original signal.
- Here we replace each data bit with n bits using a spreading code.



Direct sequence spread spectrum(DSSS)

- Each bit is assigned a code of n bits, called chips, Where chip rate is n times that of data rate.



DSSS

Eg: Consider a sequence used in wireless LAN, the famous bankers sequence where n=11.

→Assume that original signal and chips in the chip generator use NRZ polar encoding.

→ The spreading code is 11 chips having a pattern 10 110111000.

→ If original signal rate is N, the rate of spread signal is 11N means that the required BW for Spread signal is 11 times larger than the BW of the original signal.

→The spread signal can provide privacy if the intruder does not know the code.

DSSS

- The barker sequence provides good immunity against interference and noise as well as some protection against multipath propagation.
- Data bit 0 → -1v and 1 by +1V.
- To transmit 1 bit , sequence is 10110111000
- To transmit 0 bit , sequence is 01001000111

SWITCHING - INTRODUCTION

- A network is a set of connected devices. Whenever we have multiple devices, the problem is how to connect them to make one-to-one communication possible.

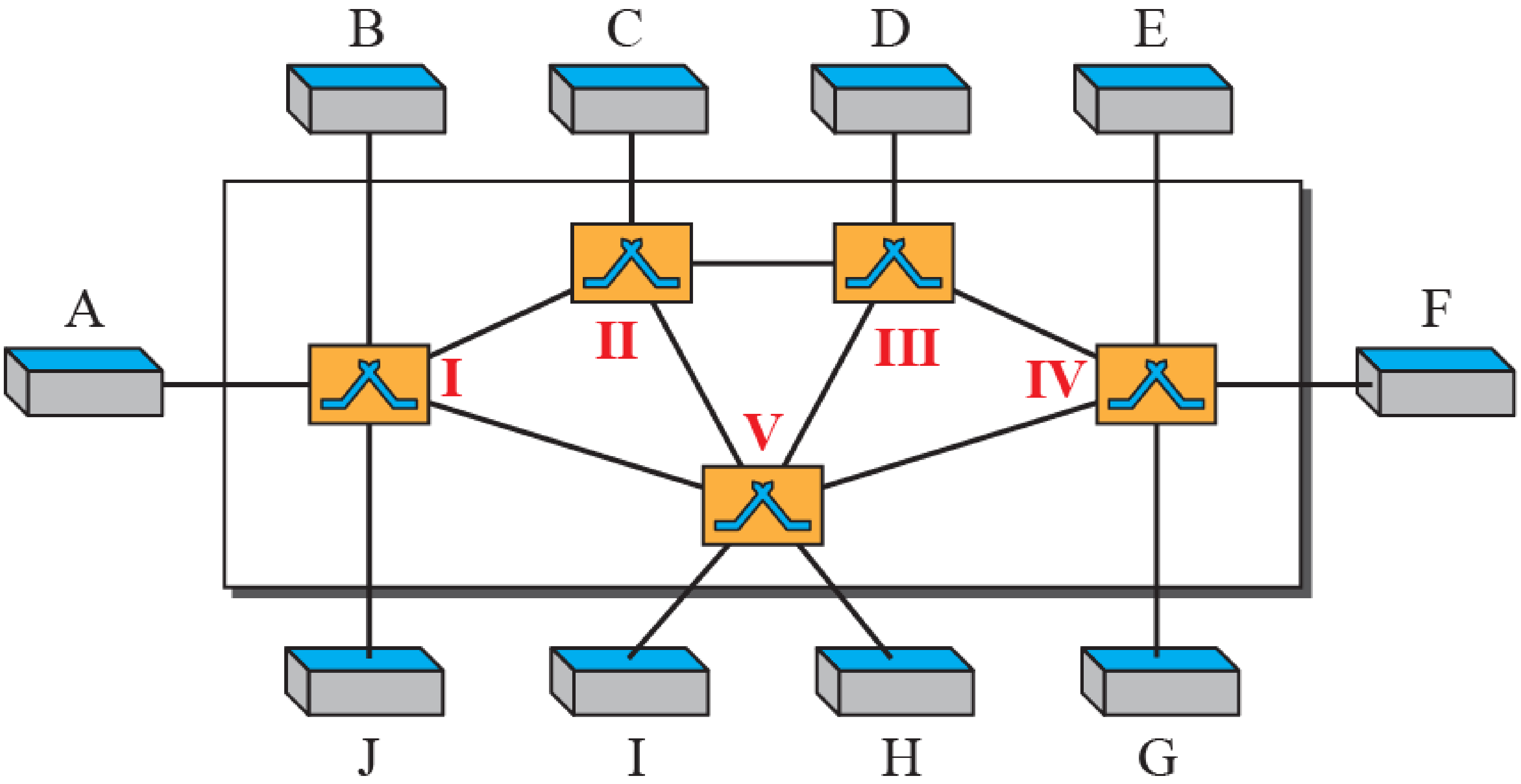
- The solution is switching.

- A switched network consists of a series of interlinked nodes, called switches.

- Switches are devices capable of creating temporary connections between the devices connected to it.

- Some of these node are connected to the end systems(Computer or Telephones)

Switched network



Three Methods of Switching

- Traditionally, three methods of switching have been discussed:

i)Circuit switching,

ii)Packet switching, and

iii)Message switching.

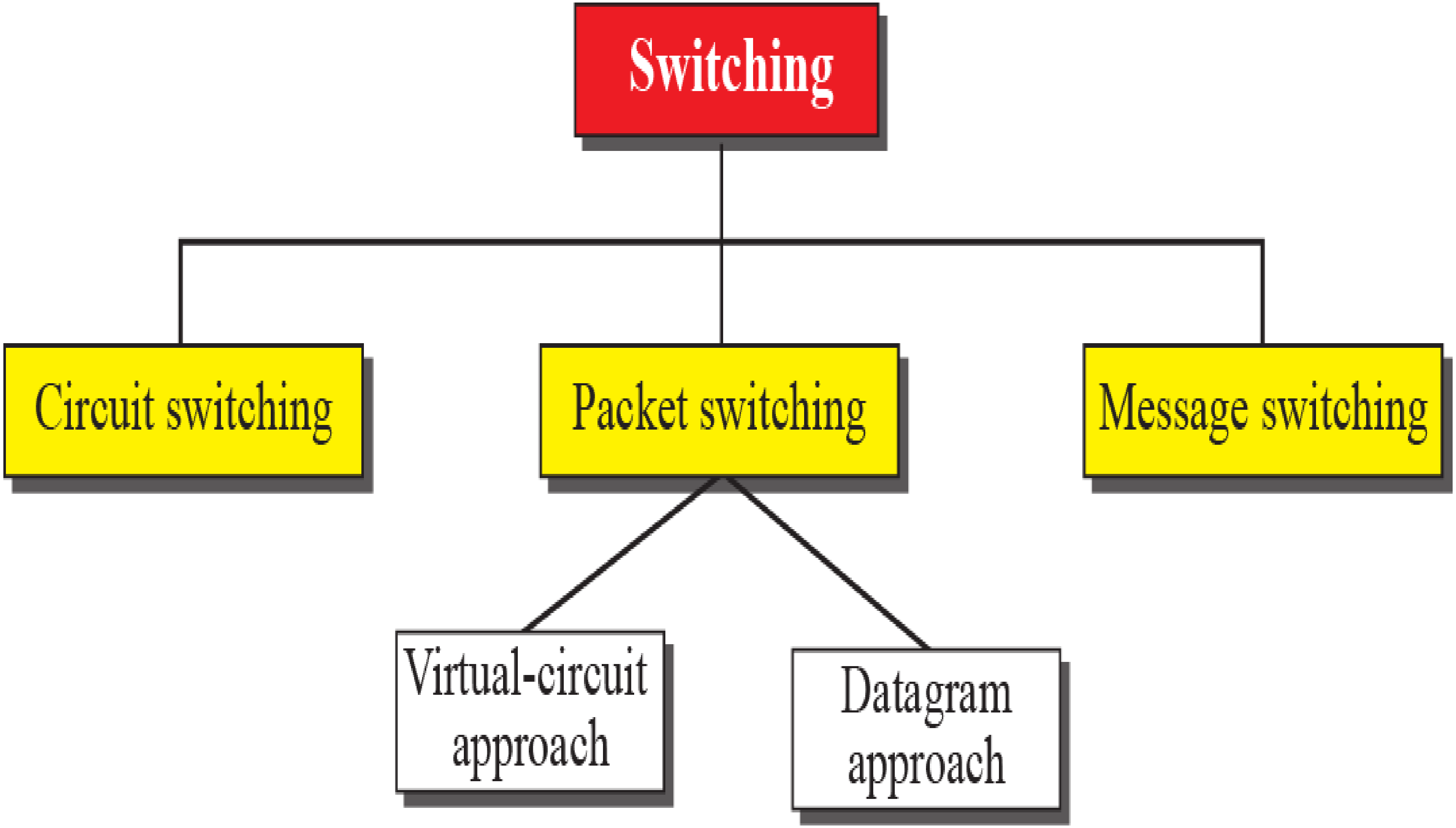
- The first two are commonly used today. The third has been phased out in general communications but still has applications.

➢Packet switching can further be divided into two subcategories virtual-circuit approach and datagram approach.

Switching and TCP/IP Layers

- Switching can happen at several layers of the TCP/IP protocol suite:
  - At the physical layer
  - At the data-link layer
  - At the network layer

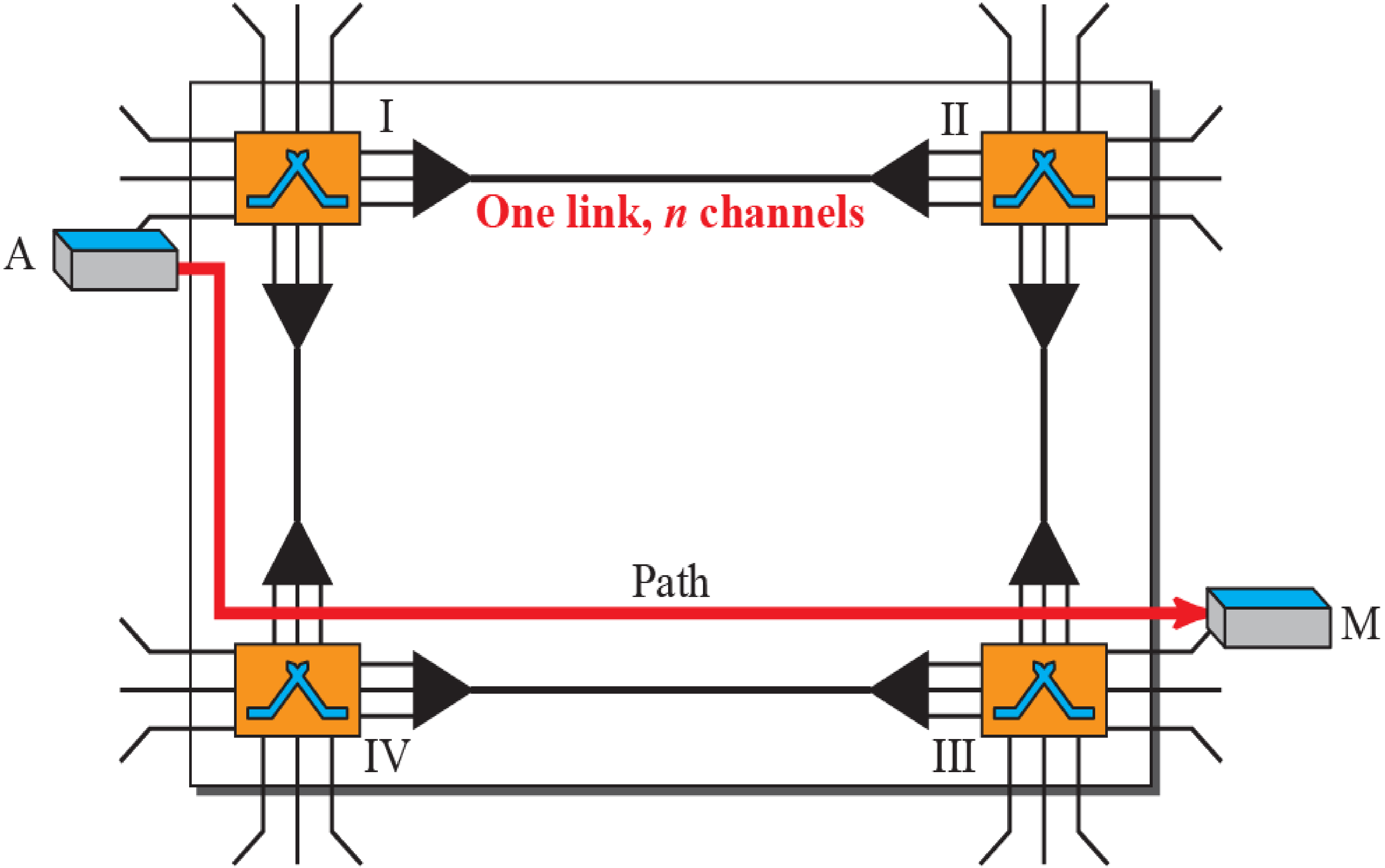
Taxonomy of switched networks



CIRCUIT-SWITCHED NETWORKS

- A circuit-switched network consists of a set of switches connected by physical links.
- Connection between two stations has a dedicated path made of one or more links.
- However, each connection uses only one dedicated channel on each link.
- Each link is normally divided into n channels by using FDM or TDM

A trivial circuit-switched network



Circuit switched network

- When end system A need to communicate with end system M, Connection request must be accepted by all switches and M itself.(Setup phase)
- A circuit is reserved on each link.
- A combination of circuits or channels defines dedicated paths.
- Next Phase is data transfer .
- Finally tearing down the connection.

Circuit switched network

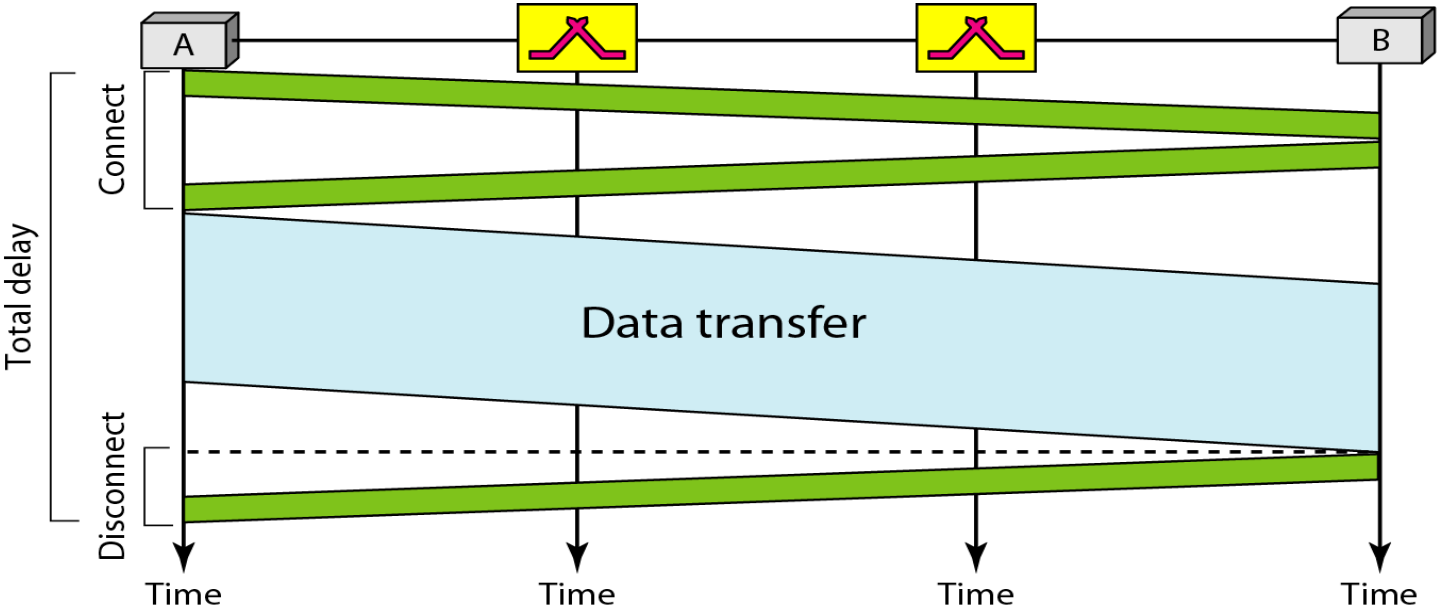
- Ckt. switching takes place at Physical Layer.
- Before starting communication, the stations must make a reservation for the resources to be used during communication.
- These resources :BW in FDM, time slots in TDM, switch buffers, processing time, switch i/o ports, must remain dedicate for the entire duration of data transfer until tear down phase.
- No addressing involved in data transfer.

Three phases

1) Connection Setup phase

2)Data transfer phase

3)Tear down phase



Three phases

- Before two parties can communicate, a dedicated channel needs to be established.
- Connection set up means creating dedicated channel.
- Eg: System A would like to communicate to M.
- The setup request must pass through all the switches between A and M.
- Switches have to find a dedicated channel.
- Destination should send back the Ack. to source in opposite direction to system A.

Three phases

- Only after System A receives the ACK. , connection is established.
- End-to-end addressing is required to set up a connection.
- Data transfer phase: After dedicated channel is established, two parties can transfer the data.
- Tear down phase: When no data to transfer, a signal is sent to all switches in between src. and dest. To release the resources.

Efficiency

- As resources are allocated before data transfer for the entire duration of connection,
- There is an argument that the CSN is not efficient.
- Because the resources are unavailable for other connections.

❑ In telephone N/W, the moment conversation is over, connection is released, Whereas in Computer N/Ws, a computer can be connected to other even if no activity for a long time.

Delay

- Though CSN has low efficiency, delay is minimal.
- During data transfer, data are not delayed at switches.
- The total delay will be due to time needed to setup connection, data transfer and tear down of connection.
- Delay caused by set up is sum of four parts

Propagation time + Request signal transfer time+ Propagation time of Ack. From the destination +Signal transfer time of ACK.

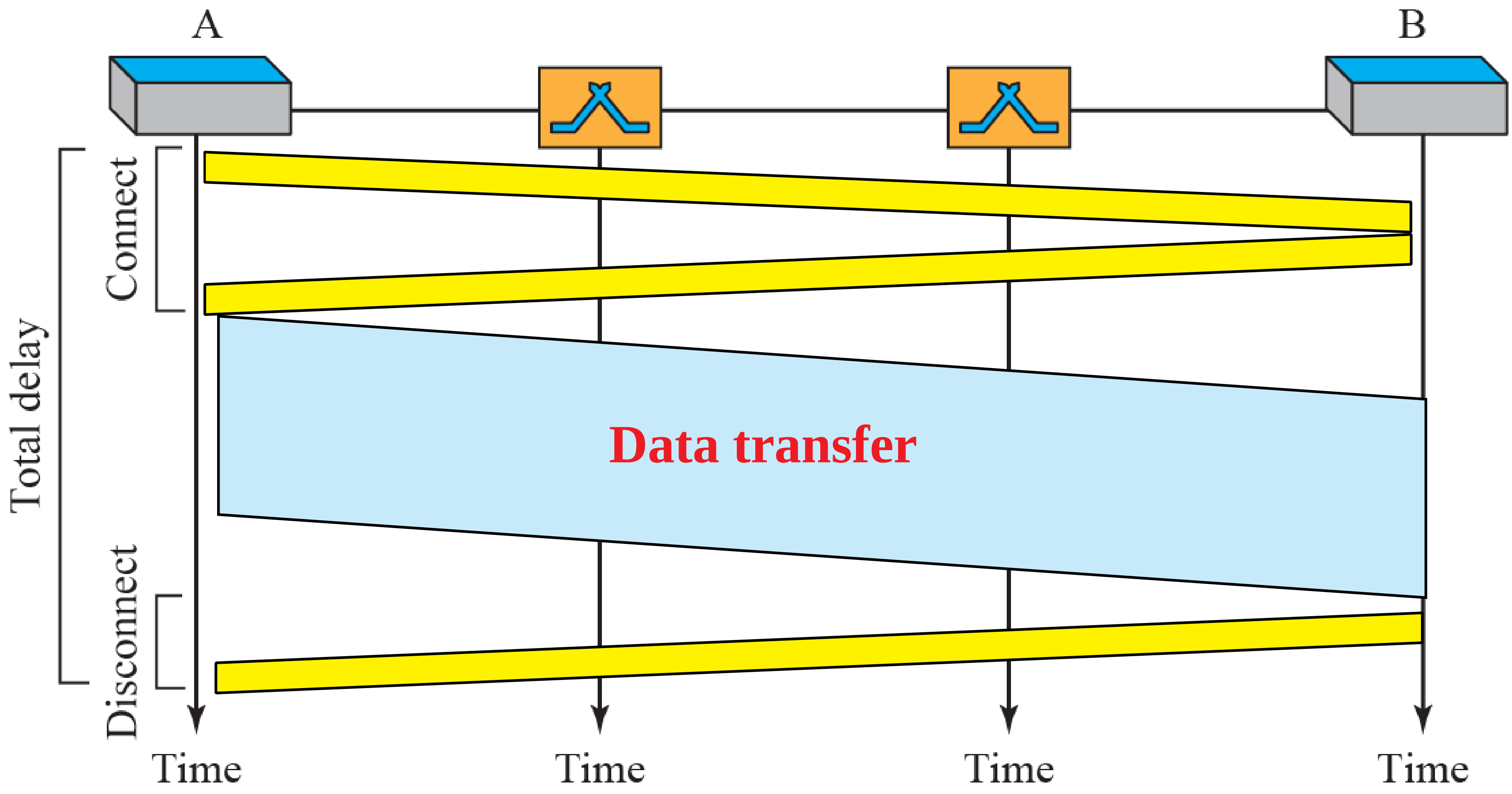
Delay

- The delay due to data transfer is sum of two parts

  Propagation time+ Data transfer time.

- Finally time required to tear down the connection.

Delay in a circuit-switched network



PACKET SWITCHING

- In data communications, we need to send messages from one end system to another.

- If the message is going to pass through a packet-switched network, it needs to be divided into packets of fixed or variable size.

- The size of the packet is determined by the network and the governing protocol.

Datagram Networks

- In a datagram network, each packet is treated independently of all others.

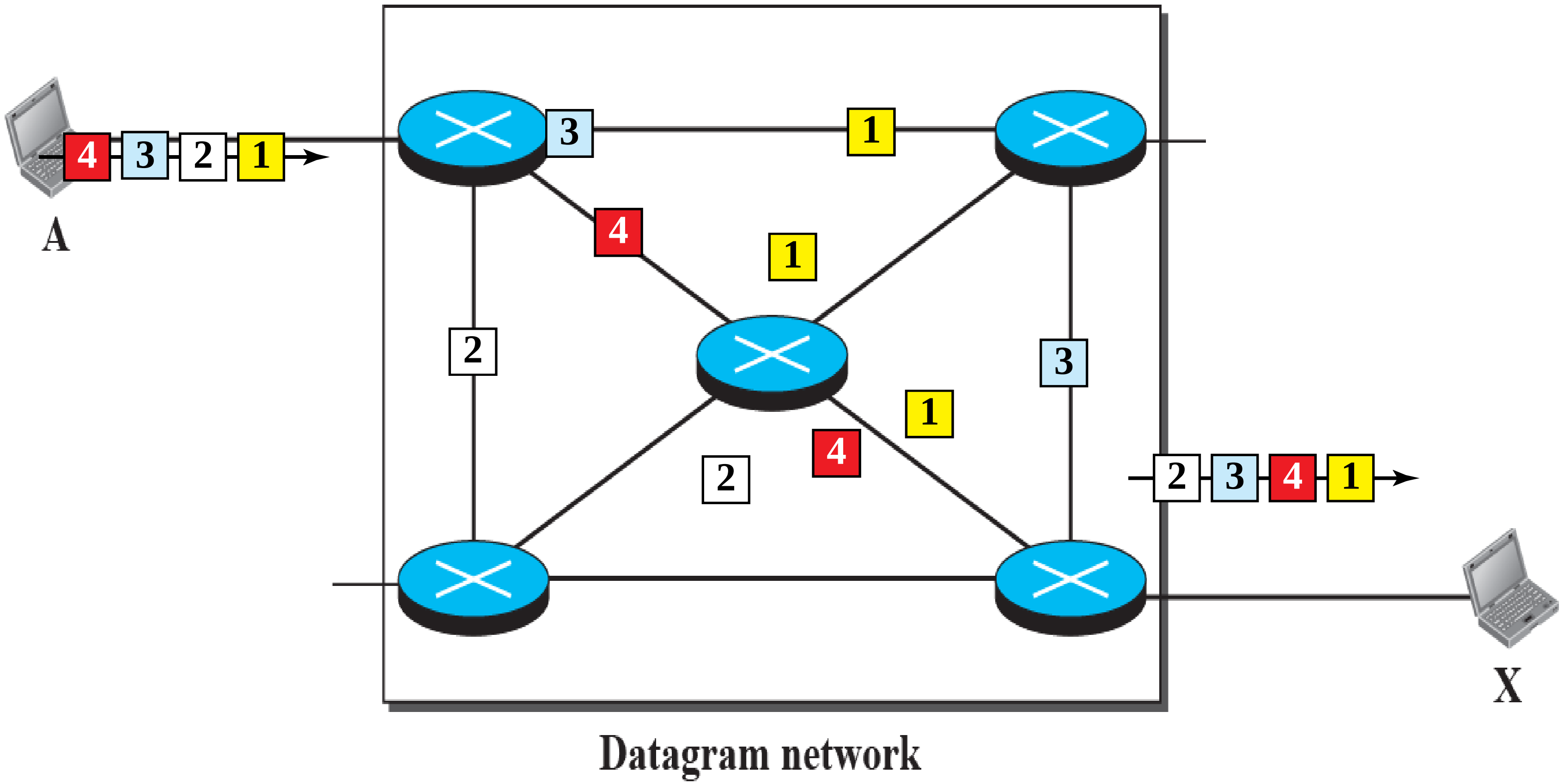
- Even if a packet is part of a multi packet transmission, the network treats it as though it existed alone.

Packets in this approach are referred to as data grams.

- Data gram switching is done at the N/W Layer.

- Data gram N/w's are sometimes referred to as connectionless N/W's

A Datagram network with four switches (routers)



Routing table

- Connection less means the switch does not keep information about the connection state.
- There is no setup or teardown phase Each switch has a routing table which is based on destination address.
- Routing tables are dynamic and updated periodically.
- The destination addresses and the corresponding forwarding o/p ports are recorded in the routing table.
- Every packet in the datagram N/W, carries the destination address along with the other header information.

Routing table

- When switch receives a packet,

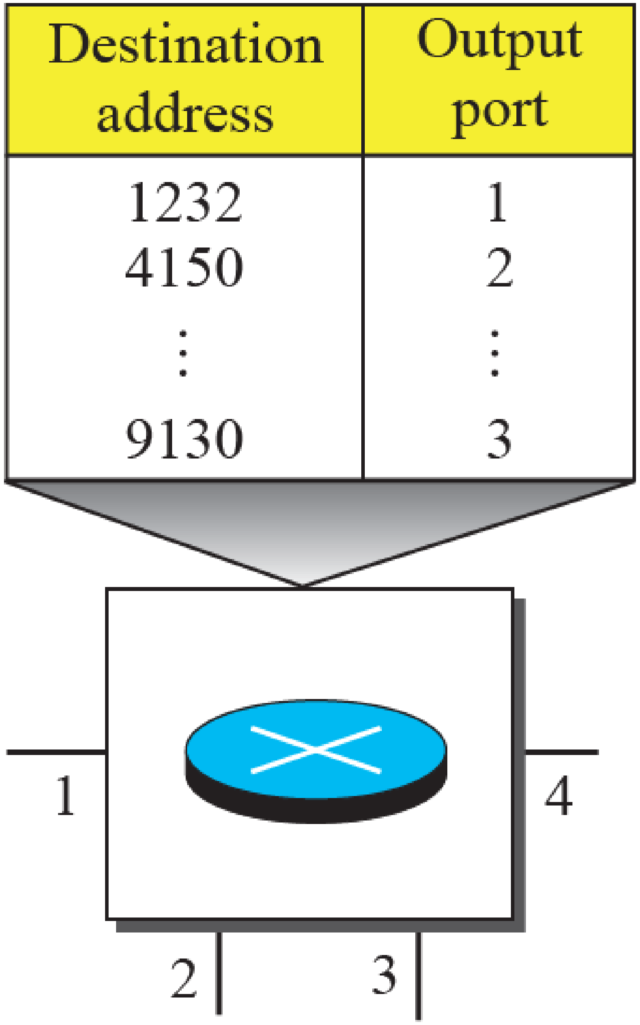
  i)It examines the destination address and

  ii) RT to find the corresponding port through which packet should be forwarded.

- Efficiency of datagram N/W:

i) Resources are allocated when the packet is to be transferred.

Routing table in a datagram network



Destination address	Output port
1232	1
4150	2
⋮	⋮
9130	3

Delay

- Greater delay and it is not uniform for every packet in the message
- As packet passes through 2 switches,

i)There are three transmission times(3T)

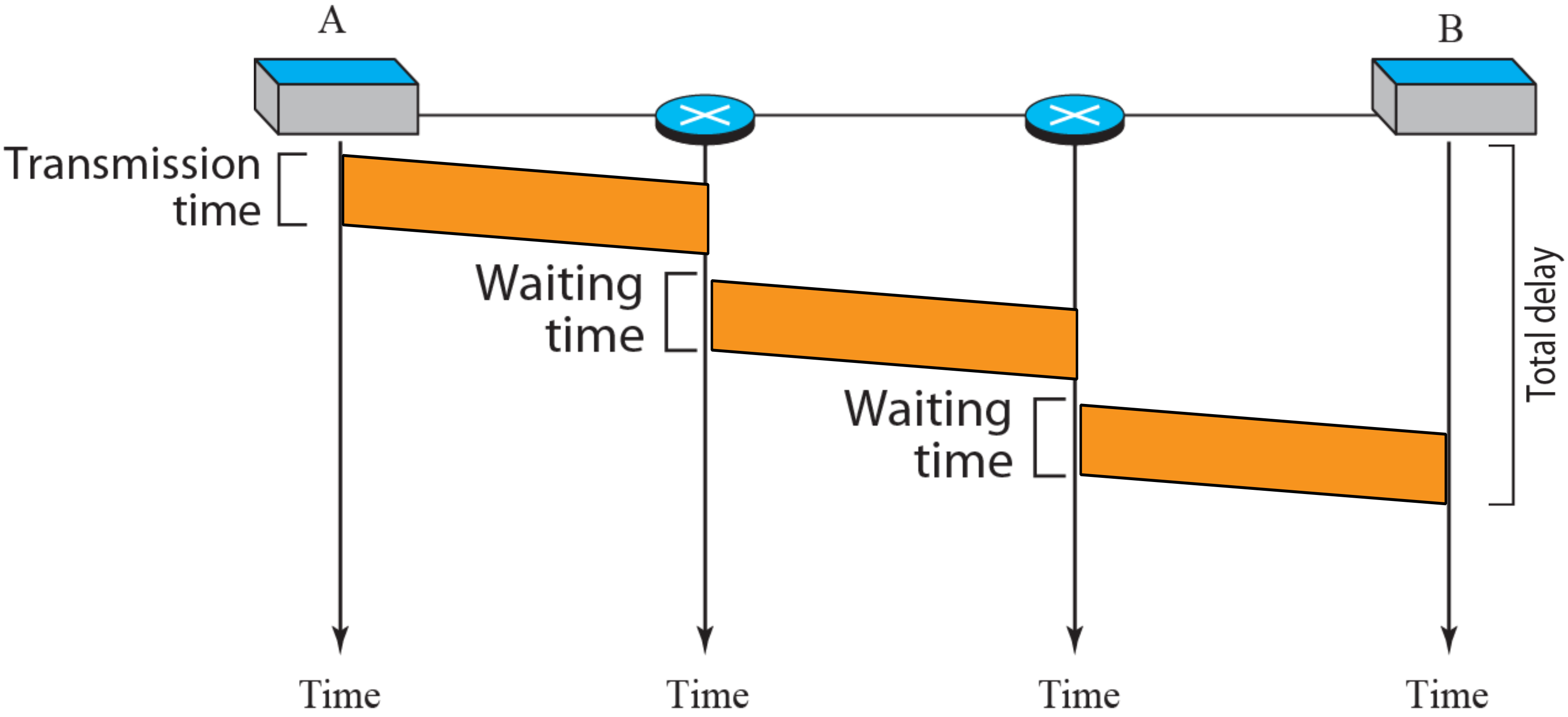
ii)Three propagation delays (3τ of the lines)

iii)Two waiting times(w1+w2)

Total delay = 3T+ 3τ+w1+w2.

- Internet has chosen Datagram approach to switch packets at the N/W Layer.
- It uses universal addresses defined in the network layer to route packet from source to destination.

Delays in a datagram network



Virtual-Circuit Networks

- A virtual-circuit network is a cross between a circuit-switched network and a datagram network.

- It has some characteristics of both.

Virtual –Circuit Networks (VCN)

- A virtual-circuit network is a cross between a circuit-switched network and a datagram network. It has some characteristics of both.

1)There is setup and tear down phase in addition to the data transfer phase.

2)Resources are allocated during setup phase.

3)Data are packetized and each packet carries an address in the header.

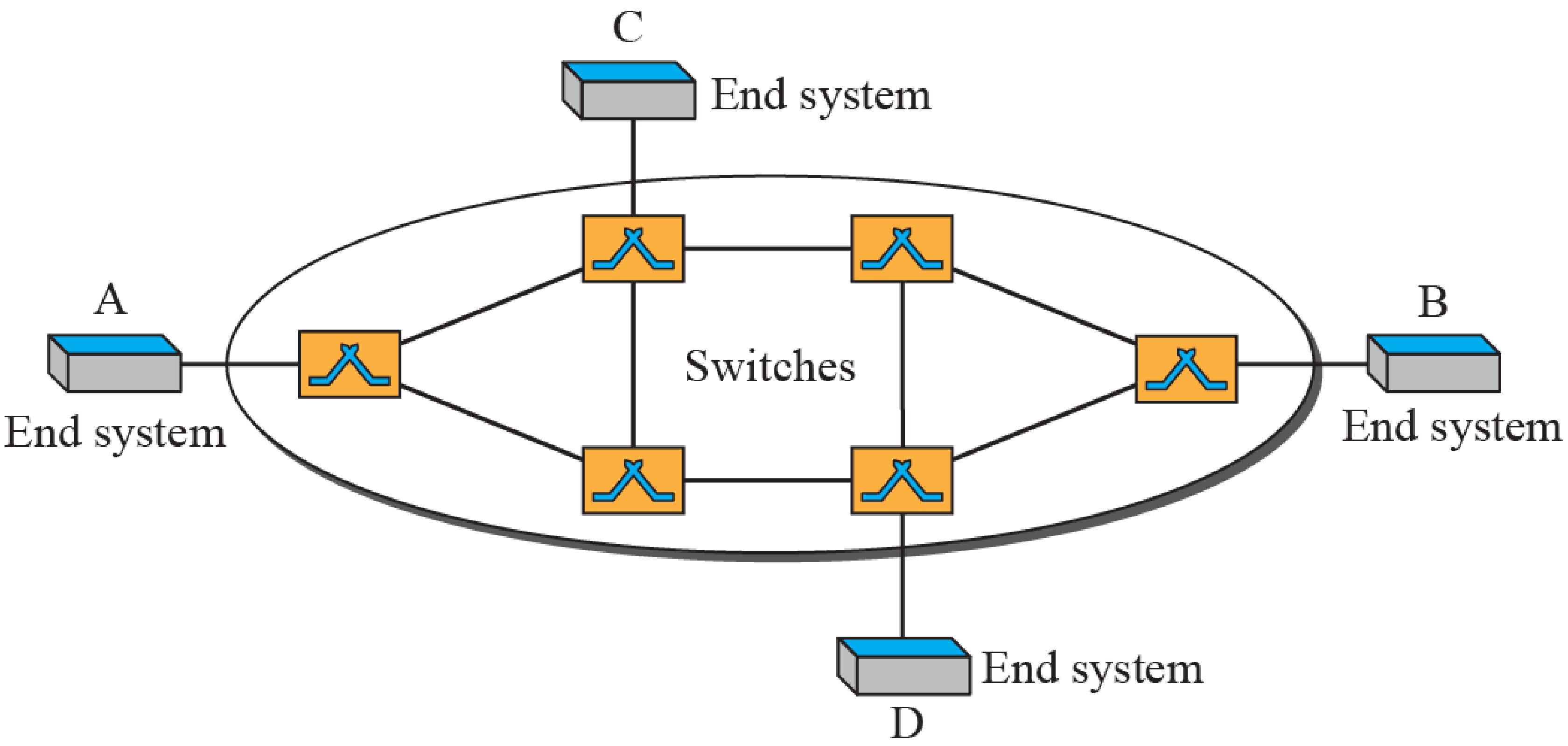
Address in the packet has local jurisdiction which defines which is the next switch and the channel to forward the packet.

Virtual –Circuit Networks

4)All packets follow the same path established during the connection setup phase

5) Implemented in Data Link Layer

Virtual-circuit network



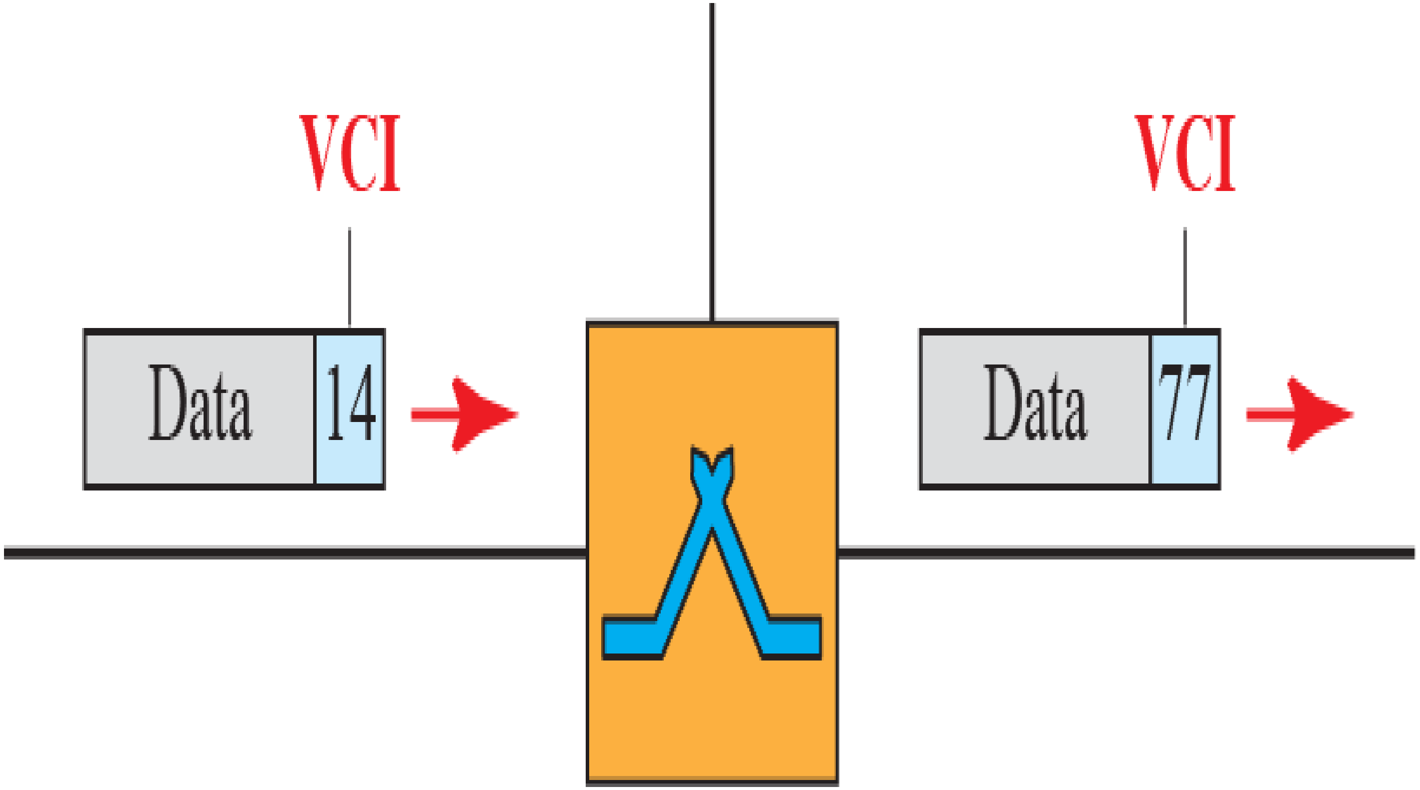
Addressing

- Two types of addresses are used:

  Global and Local

- Source and destination need to have global addresses i.e address which is unique through out the Internet.
- Global addresses in VCN are used only to create a virtual circuit identifier(VCI).
- The identifier which is used for data transfer is called VCI.
- VCI is a small number that has only switch scope.
- It is used by a frame b/w two switches and changes as it moves from switch to switch.

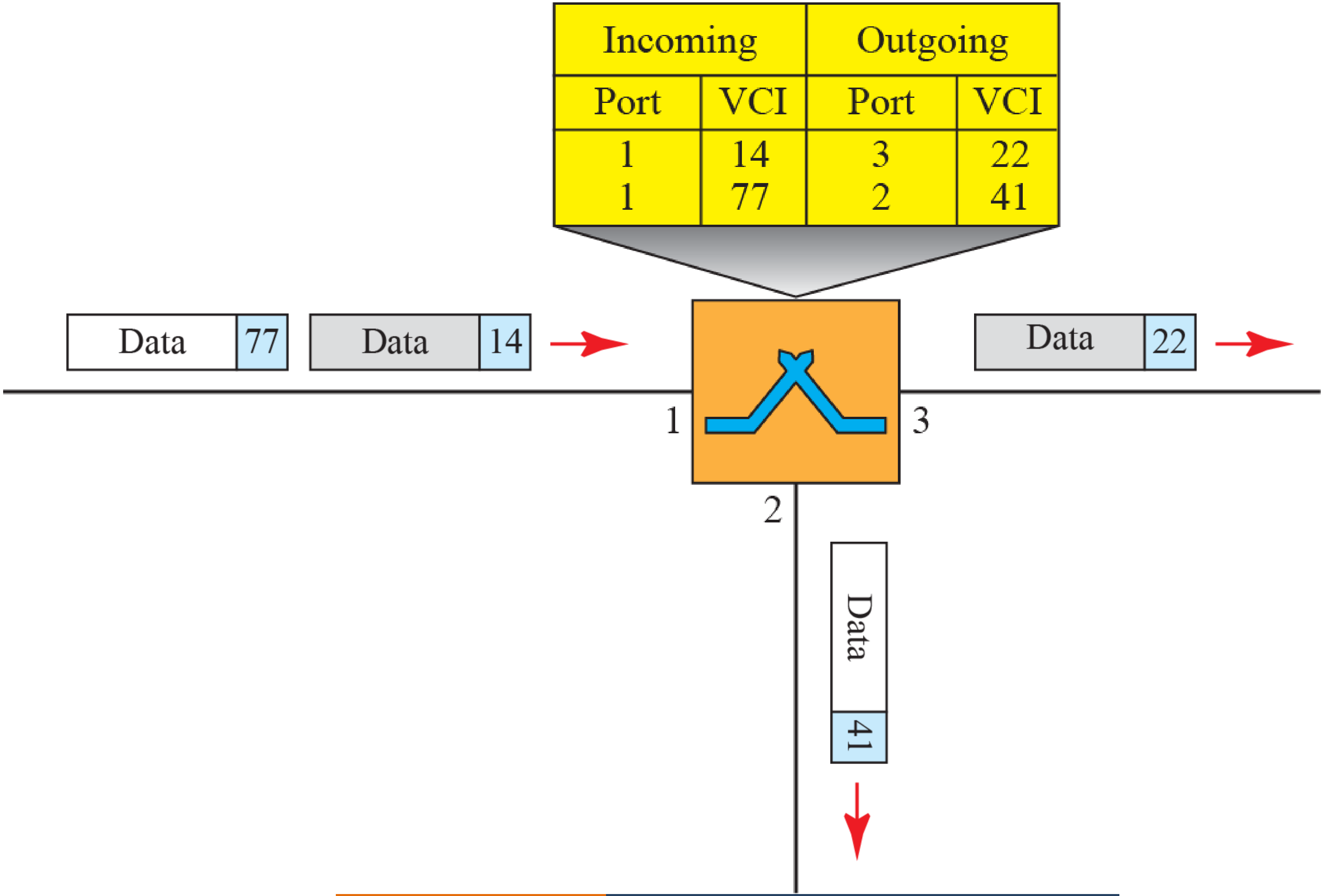
Virtual-circuit identifier



Three phases

- Setup, data transfer and tear down are 3phases.
- In setup phase, source and destination uses their global addresses to help switch to make table entries for the connection.
- In tear down phase, source and destination informs switch to remove the corresponding entry from the table.
- In Data transfer phase, data transfer occurs.

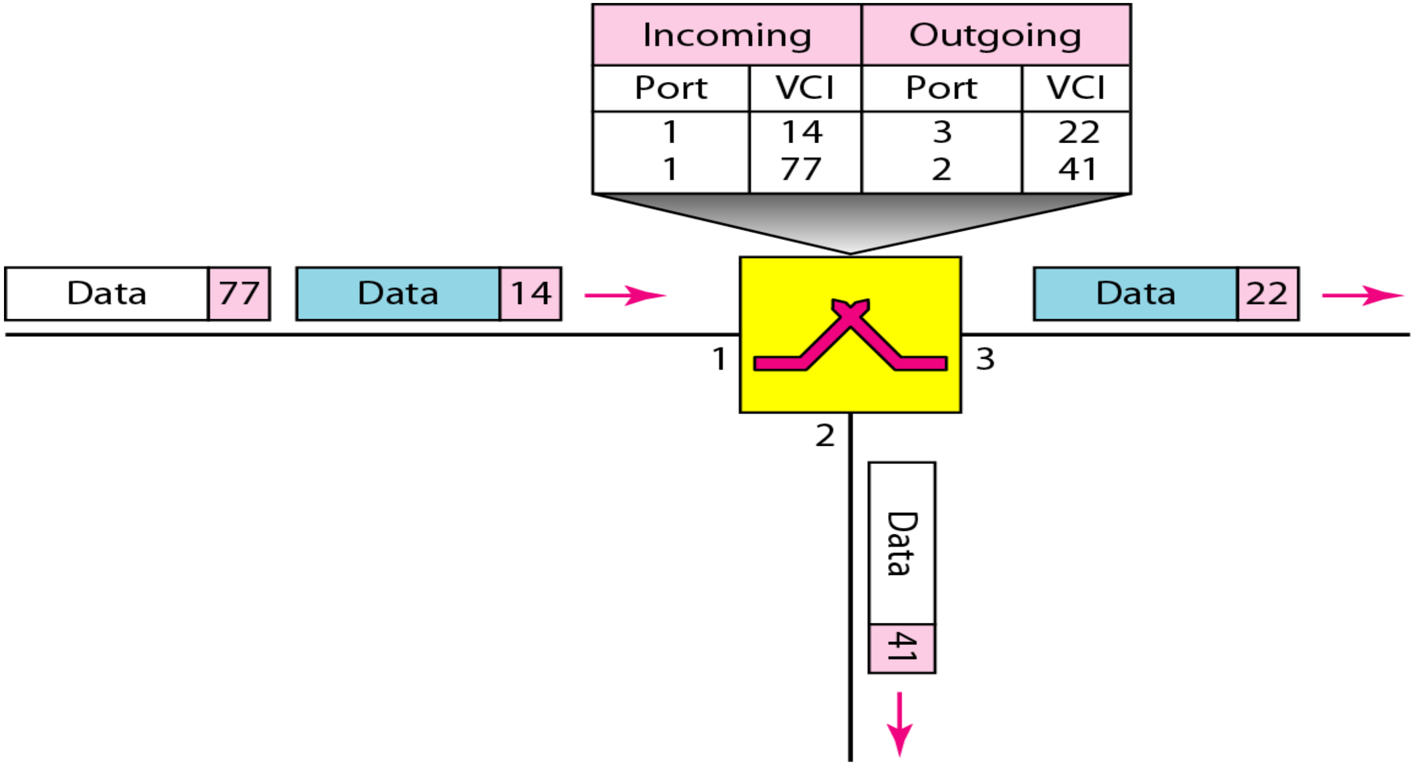
Switch and table for a virtual-circuit network



Incoming		Outgoing	
Port	VCI	Port	VCI
1	14	3	22
1	77	2	41

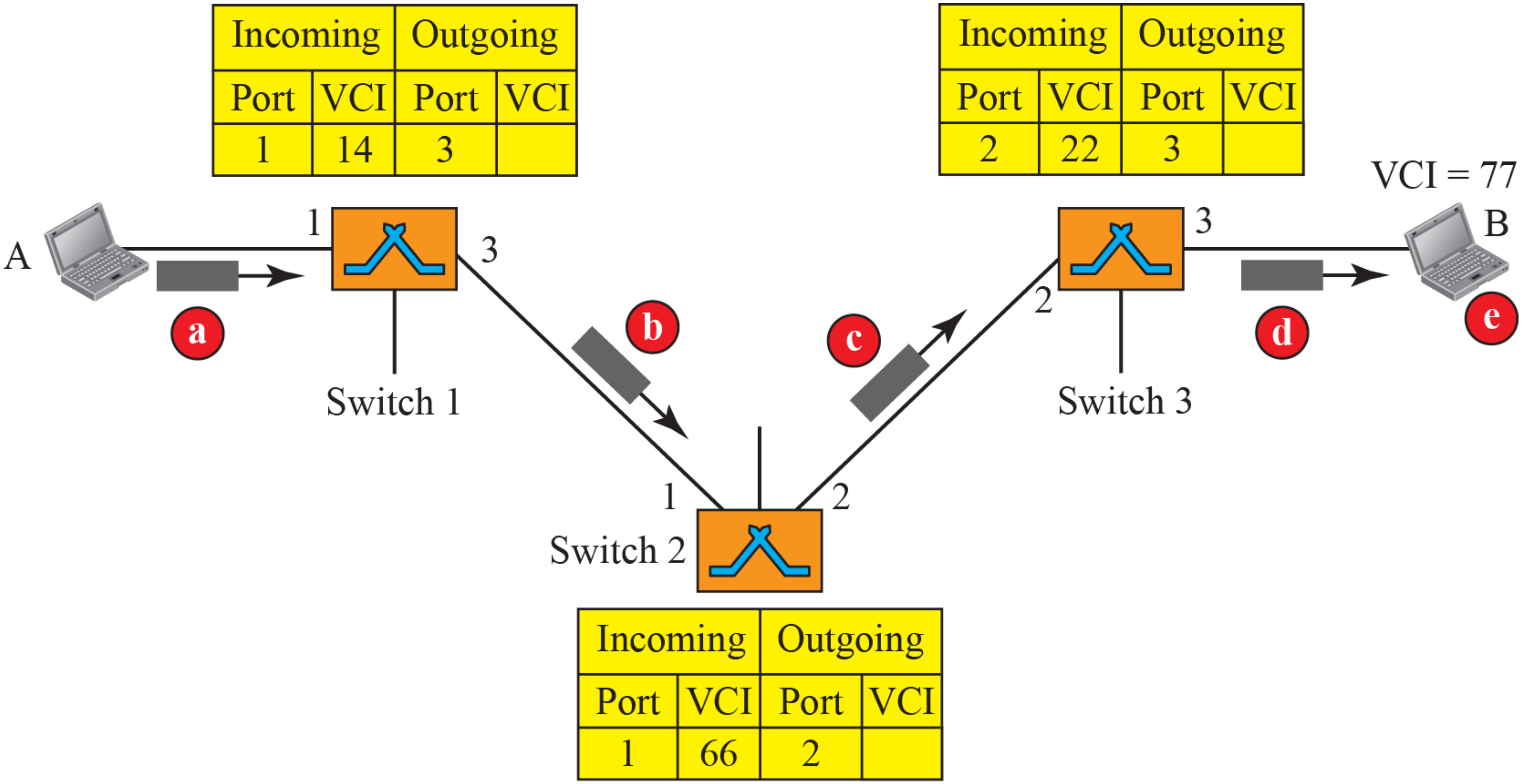
Setup phase

- Switch creates an entry for a virtual ckt.



Incoming		Outgoing	
Port	VCI	Port	VCI
1	14	3	22
1	77	2	41

Setup request in a virtual-circuit network

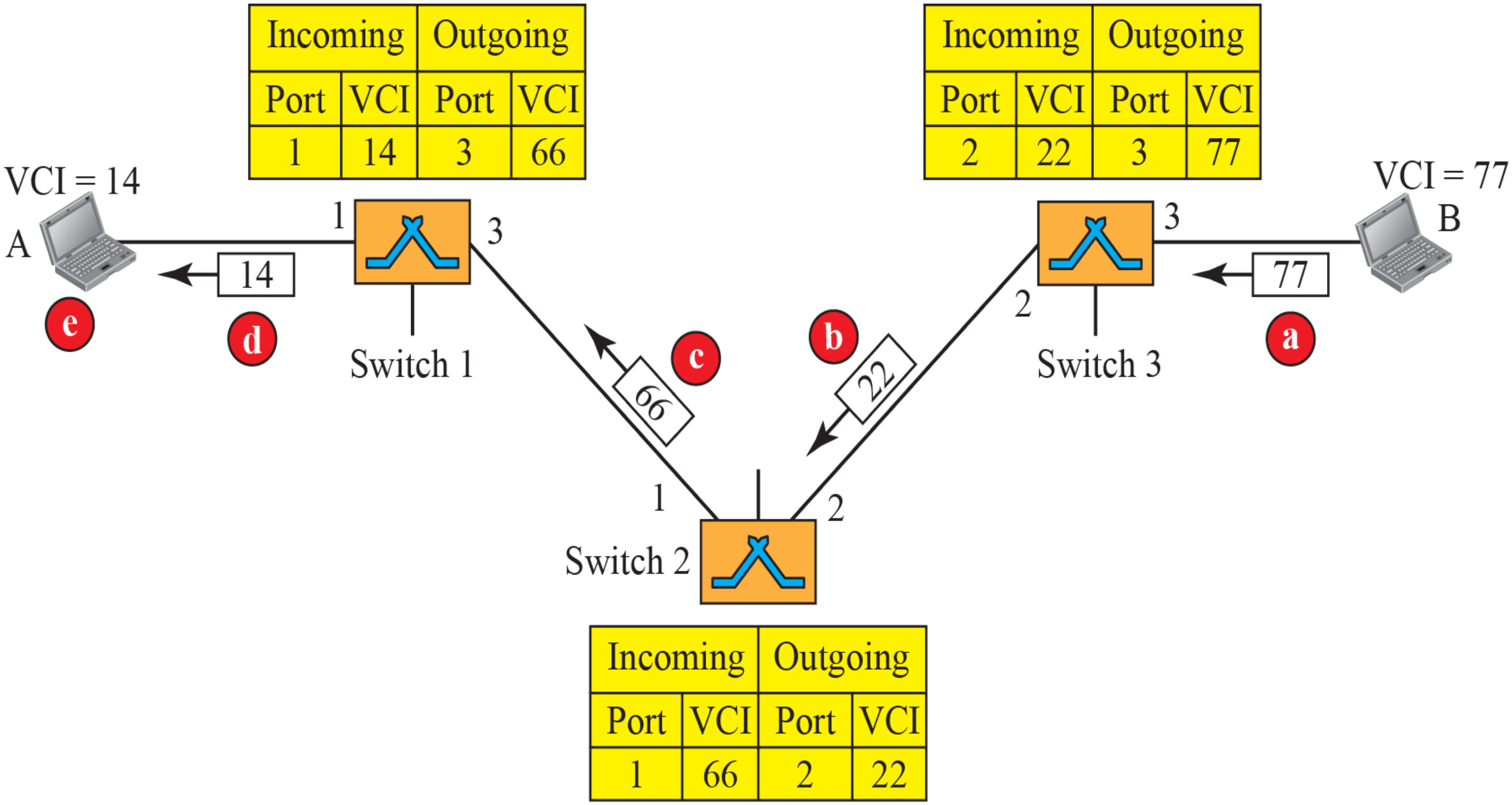


Incoming		Outgoing	
Port	VCI	Port	VCI
1	14	3	

Incoming		Outgoing	
Port	VCI	Port	VCI
1	66	2	

Incoming		Outgoing	
Port	VCI	Port	VCI
2	22	3	

Setup acknowledgment in a virtual-circuit network

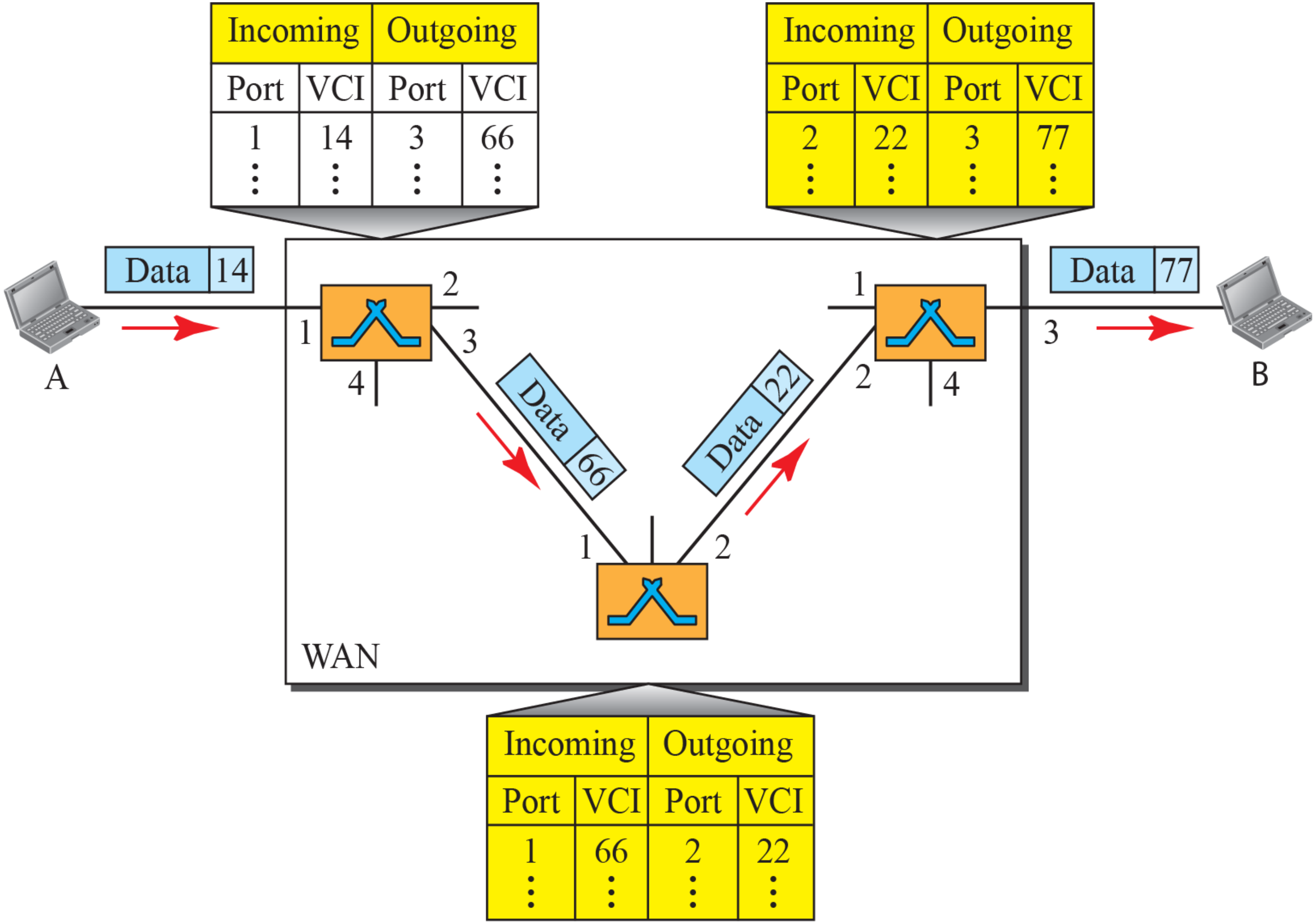


Incoming		Outgoing	
Port	VCI	Port	VCI
1	14	3	66

Incoming		Outgoing	
Port	VCI	Port	VCI
1	66	2	22

Incoming		Outgoing	
Port	VCI	Port	VCI
2	22	3	77

Source-to-destination data transfer in a circuit-switch network



Incoming		Outgoing	
Port	VCI	Port	VCI
1 ⋮	14 ⋮	3 ⋮	66 ⋮

Incoming		Outgoing	
Port	VCI	Port	VCI
2 ⋮	22 ⋮	3 ⋮	77 ⋮

Incoming		Outgoing	
Port	VCI	Port	VCI
1 ⋮	66 ⋮	2 ⋮	22 ⋮

Efficiency

- As resources are reserved in VCN during connection setup or can be on demand during data transfer phase.

i) If connection set up, delay for each packet is same.

ii) If during data transfer phase, each packet may encounter different delays.

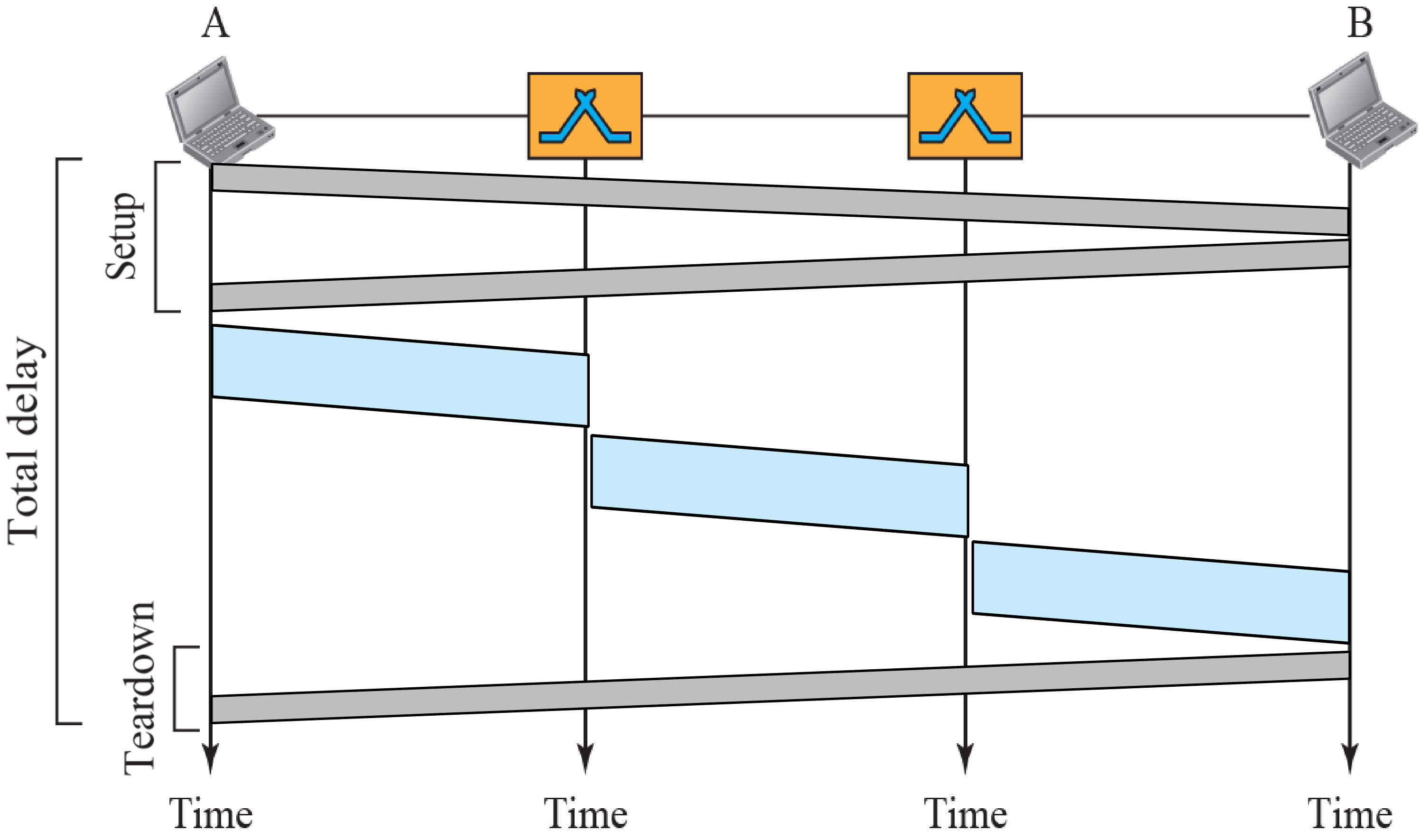
Delay

- One time delay only for connection set up and one time delay for tear down.
- If resources allocated during set up phase, no wait time for individual packets.

Total delay = 3T+ 3τ+setup delay and tear down delay.

- VCN’s are used in switched WAN’s such as frame relay and ATM N/W.

Delay in a virtual-circuit network



ERROR DETECTION AND CORRECTION

- TCP/IP protocol does not define any protocol in DLL or PL.
- These two layers are territories of networks that when connected make up the Internet.
- These networks (wired/wireless), provide services to upper three layers of TCP/IP suite.

- Networks must be able to transfer data from one device to another with acceptable accuracy.
- For most of the applications, system must guarantee that data received are identical to the one transmitted .

INTRODUCTION

- In transmission, there are possibilities that any time message gets corrupted.
- Many factors alter one or more bits of a message.
- Some applications require a mechanism for detecting and correcting errors.
- Some applications can tolerate some errors.

Eg: A/V (random errors)

but transfer of text, expect high level of accuracy.

- Discuss some issues related, directly or indirectly, to error detection and correction.

Types of Errors

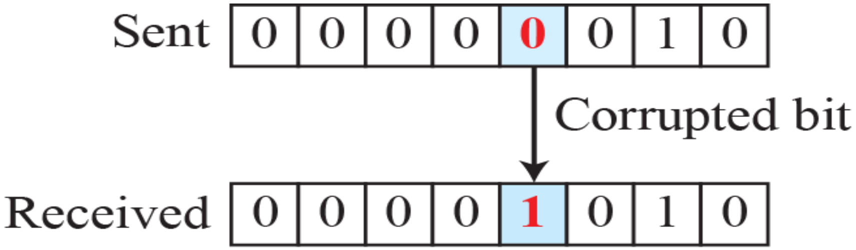
•Whenever bits flow from one point to another, they are subject to unpredictable changes because of interference.

•This interference can change the shape of the signal.

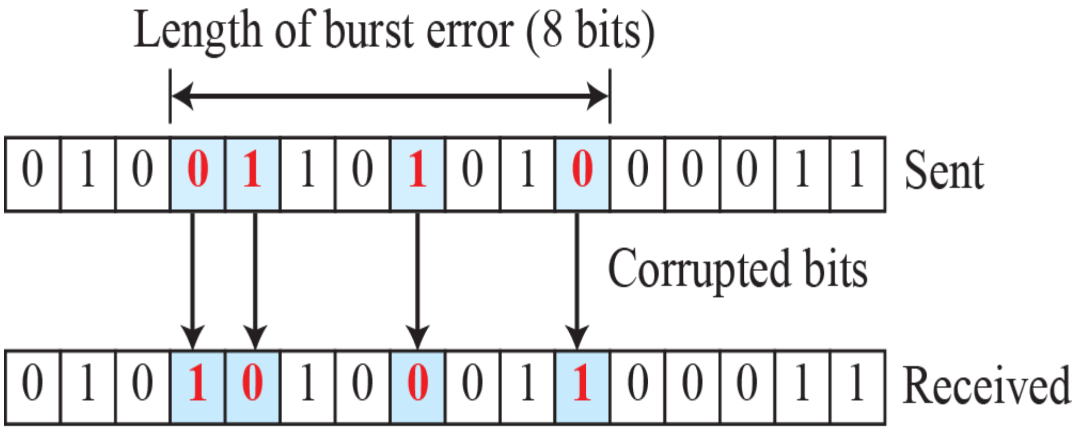
• The term single-bit error means that only 1 bit of a given data unit (such as a byte, character, or packet) is changed from 1 to 0 or from 0 to 1.

• The term burst error means that 2 or more bits in the data unit have changed from 1 to 0 or from 0 to 1. Figure below shows the effect of a single-bit and a burst error on a data unit.

Single-bit and burst error



a. Single-bit error



b. Burst error

Types of Errors

- Burst error is more likely to occur than single bit error.

- Reason is duration of noise signal is longer than the duration of a bit.

- When noise affects data, it affects set of bits.

- Number of bits affected depends on the duration of the noise and data rate.

Eg: If we send 1 kbps of data, a noise 1/100 s can affect 10 bits.
If 1 Mbps data rate, noise can affect 10,000 bits.

Redundancy

- The central concept in detecting or correcting errors is redundancy.

- To be able to detect or correct errors, need to send some extra bits with our data.

- These redundant bits are added by the sender and removed by the receiver.

- Presence of redundant bits allows the receiver to detect or correct corrupted bits.

Detection versus Correction

- The correction of errors is more difficult than the detection.

- In error detection, check if any error has occurred. The answer is a simple yes or no.Not interested in the number of corrupted bits.

- A single-bit error is the same for us as a burst error.

- In error correction,

  Need to know the exact number of bits that are corrupted and, more importantly, their location in the message.

Coding

- Redundancy is achieved through various coding schemes.

- The sender adds redundant bits through a process that creates a relationship between the redundant bits and the actual data bits.

- The receiver checks the relationships between the two sets of bits to detect errors.

- The ratio of redundant bits to data bits and the robustness of the process are important factors in any coding scheme.

Coding

- Two coding schemes are

  i) Block coding ii) Convolution coding
- Concentrate on block codes.
- In modulo-N arithmetic, the integers in the range 0 to N −1inclusive is used,

$0 \oplus 0 = 0 \qquad 1 \oplus 1 = 0$

a. Two bits are the same, the result is 0.

$0 \oplus 1 = 1 \qquad 1 \oplus 0 = 1$

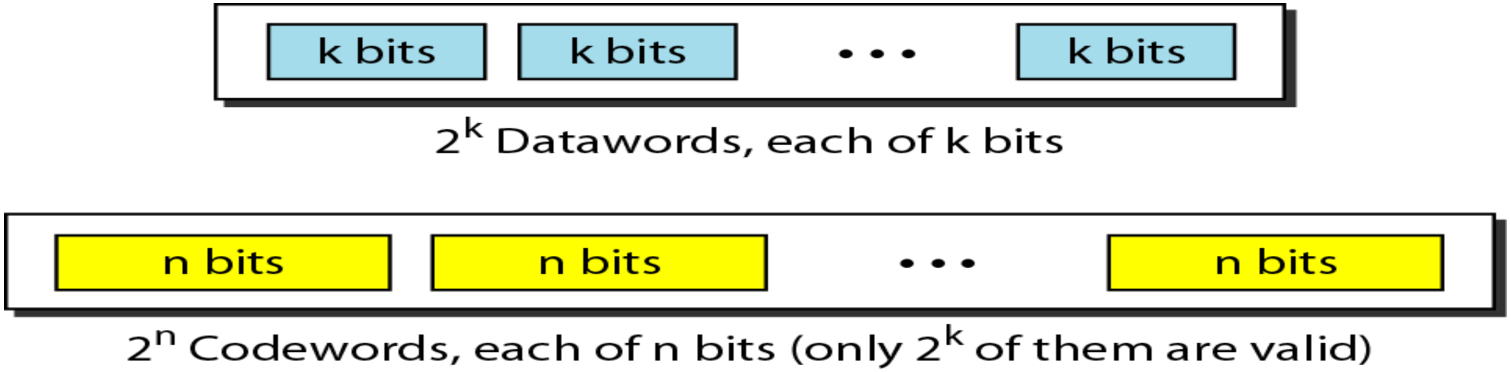
b. Two bits are different, the result is 1.

$$\begin{array}{cccccc} & 1 & 0 & 1 & 1 & 0 \\ \oplus & 1 & 1 & 1 & 0 & 0 \\ \hline & 0 & 1 & 0 & 1 & 0 \end{array}$$

c. Result of XORing two patterns

BLOCK CODING

- In block coding, message is divided into blocks, each of k bits, called data words.
- Add r redundant bits to each block to make the length n = k + r.
- The resulting n-bit blocks are called code words.


k bits
k bits
• • •
k bits
2^k Datawords, each of k bits
n bits
n bits
• • •
n bits
2^n Codewords, each of n bits (only 2^k of them are valid)

BLOCK CODING

- In block coding, we have set of data words, each of size k, code words, each of size n.
- Since n>k, the possible number of code words are larger than the data words.
- Block coding process is one-to-one, same data word is encoded as code word.
- $2^n - 2^K$ code words are not used.
- If receiver receives invalid codeword, it indicates data is corrupted during transmission.

Error Detection

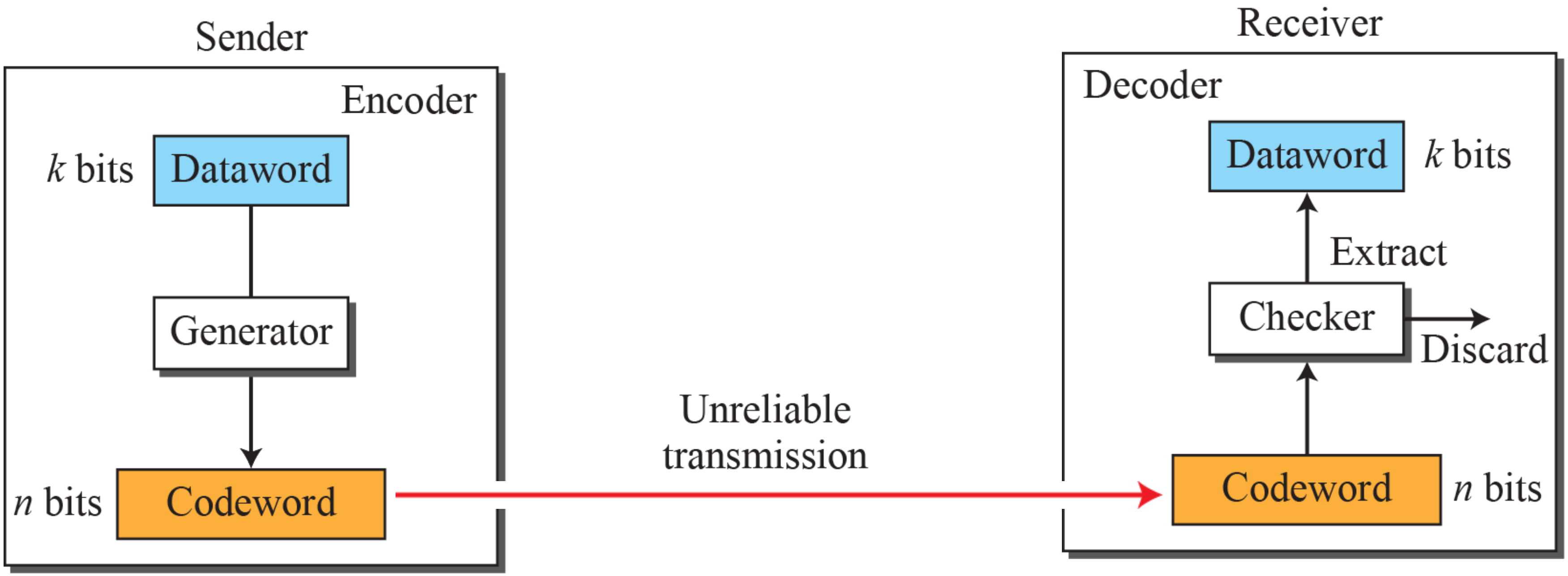
- How can errors be detected by using block coding?

- If the following two conditions are met, the receiver can detect a change in the original codeword.

a. The receiver has (or can find) a list of valid code words.

b. The original codeword has changed to an invalid one.

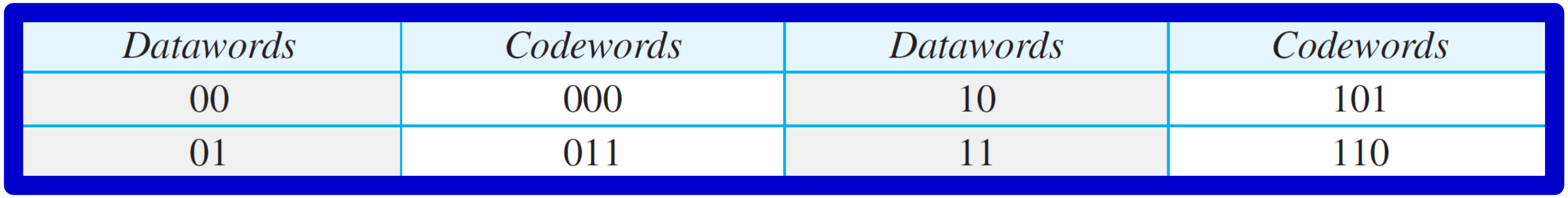
Process of error detection in block coding



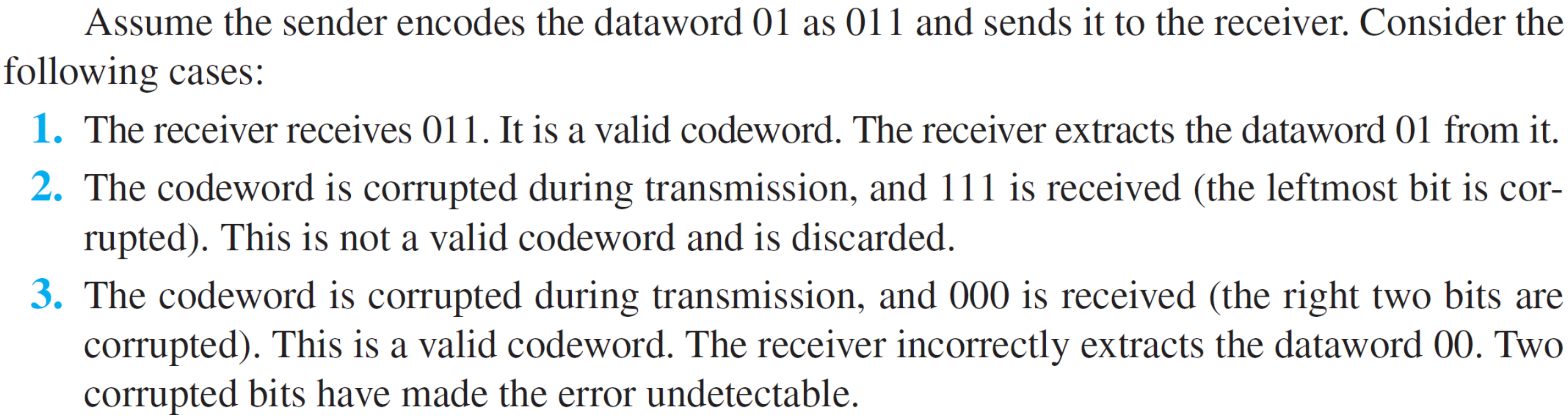
Example 10.1

Let us assume that k = 2 and n = 3. Table below shows the list of data words and code words.

A code for error detection



Datawords	*Codewords*	*Datawords*	*Codewords*
00	000	10	101
01	011	11	110



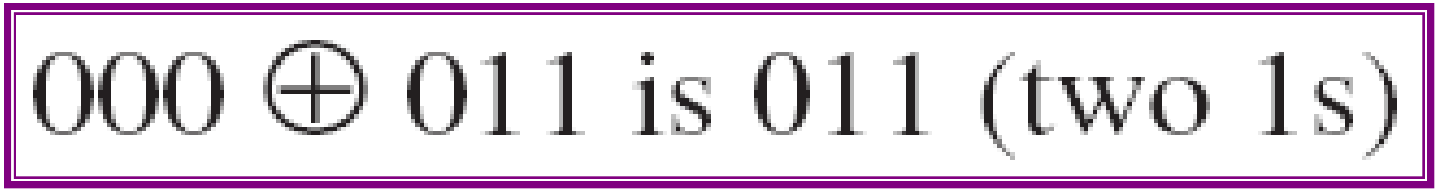
Assume the sender encodes the dataword 01 as 011 and sends it to the receiver. Consider the following cases:

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.

Hamming distance

- The Hamming distance between two words(of same size) is the number of differences between corresponding bits. *(total number of 1's).*
- *Why HD in Error detection????*
- *HD between the sent code word & received codeword is the number of bits that got corrupted during transmission.*

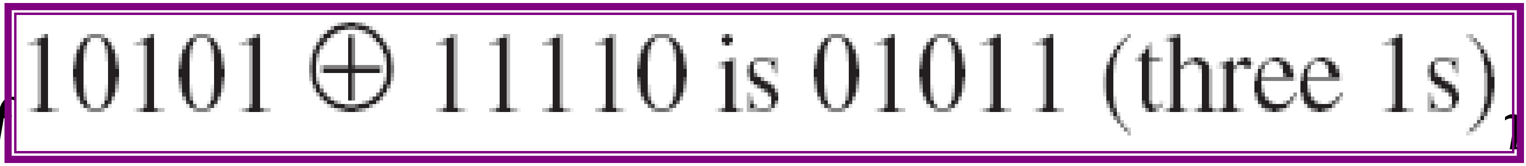
Eg: Let us find the Hamming distance between two pairs of words.



$000 \oplus 011$ is 011 (two 1s)

Hamming Distance

- The Hamming distance d(000, 011) is 2.
- The Hamming distance d(10101, 11110) is 3 because

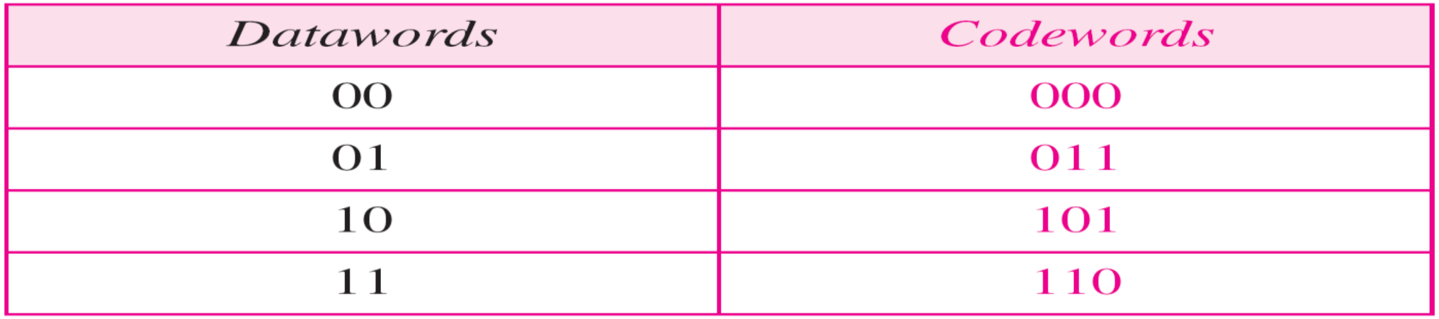


10101 ⊕ 11110 is 01011 (three 1s)

- *If not zero, then code word has been corrupted during transmission.*
- *The minimum Hamming distance is the smallest Hamming distance between all possible pairs in a set of words.*

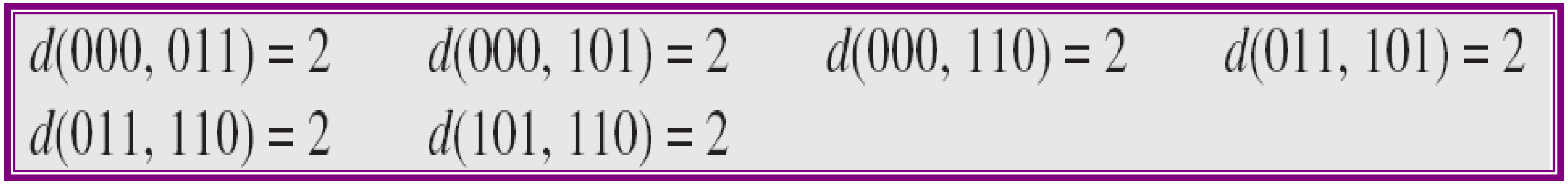
Hamming distance

- Find the minimum Hamming distance of the coding scheme in



Datawords	*Codewords*
00	000
01	011
10	101
11	110

We first find all Hamming distances



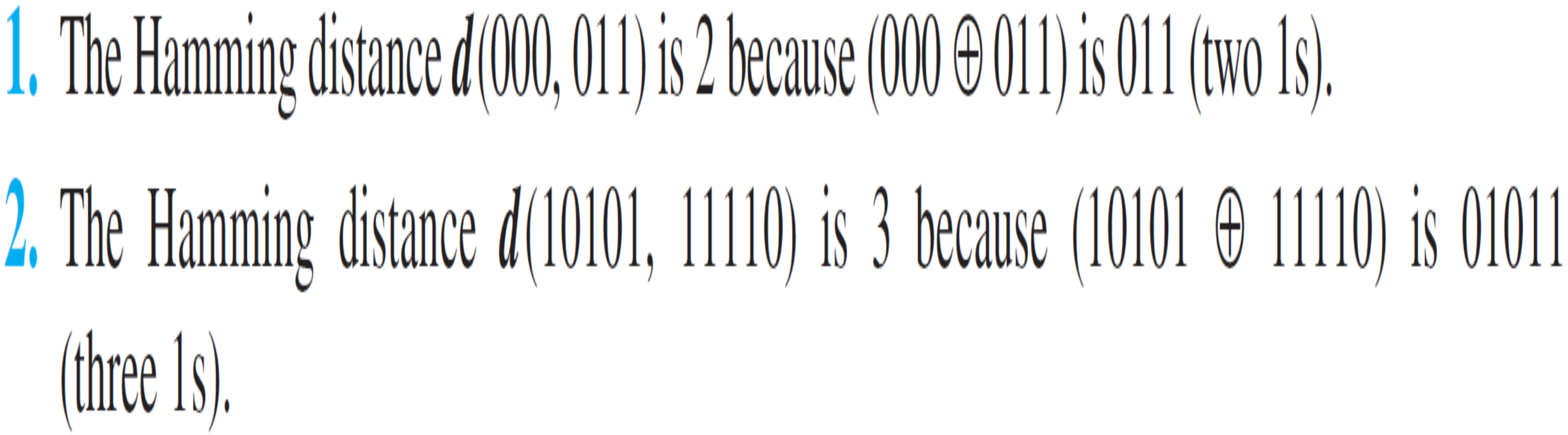
$d(000, 011) = 2$ $d(000, 101) = 2$ $d(000, 110) = 2$ $d(011, 101) = 2$

$d(011, 110) = 2$ $d(101, 110) = 2$

The d_{min} in this case is 2

Example 10.2

Let us find the Hamming distance between two pairs of words.



1. The Hamming distance $d(000, 011)$ is 2 because $(000 \oplus 011)$ is 011 (two 1s).
2. The Hamming distance $d(10101, 11110)$ is 3 because $(10101 \oplus 11110)$ is 01011 (three 1s).

Example 10.3

The minimum Hamming distance for our first code scheme (Table 10.1) is 2.

This code guarantees detection of only a single error. For example, if the third code word (101) is sent and one error occurs, the received code word does not match any valid code word.

If two errors occur, however, the received code word may match a valid code word and the errors are not detected.

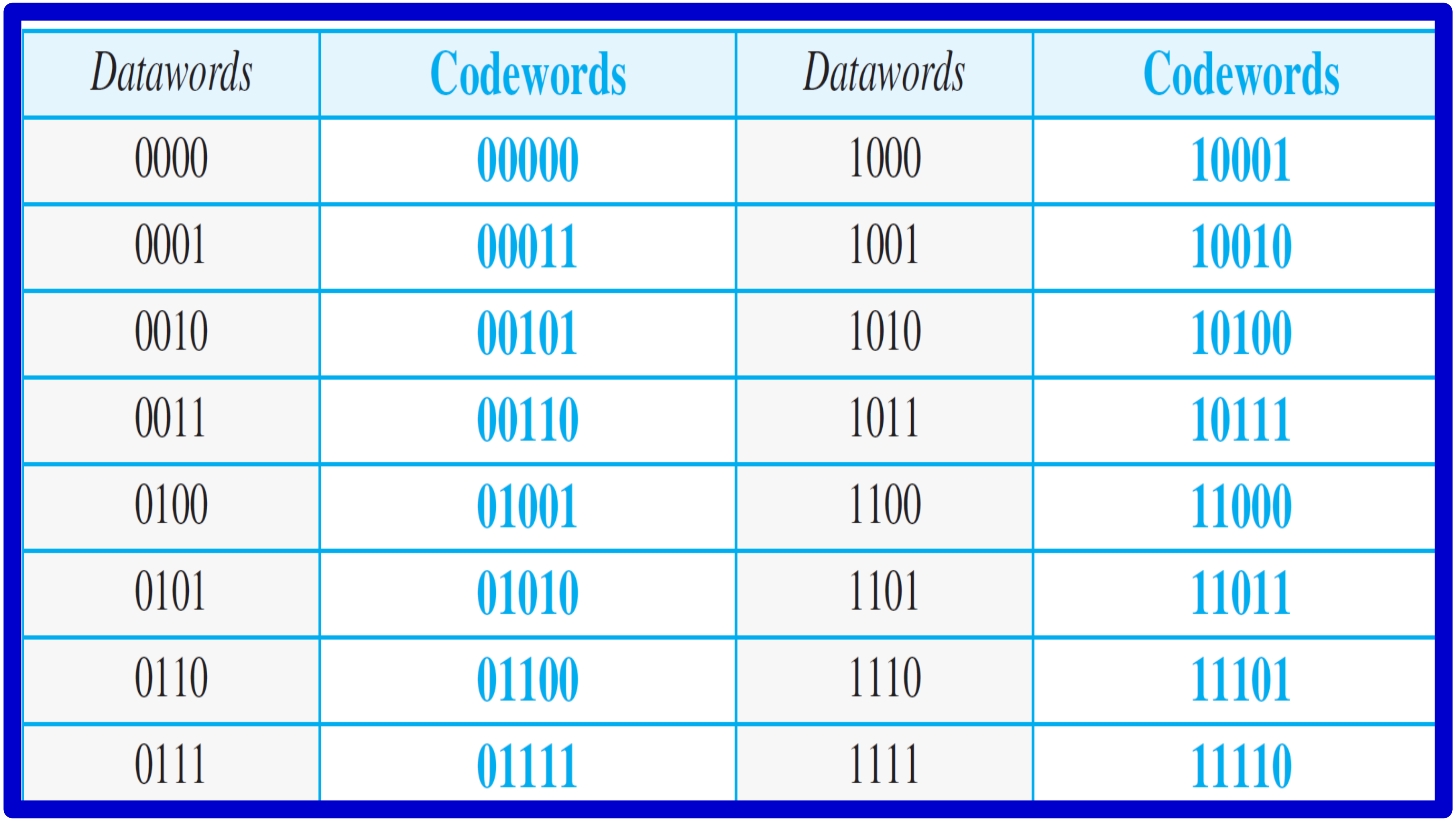
Example 10.6

In our first code (Table 10.1), the numbers of 1s in the nonzero codewords are 2, 2, and 2. So the minimum Hamming distance is $\boldsymbol{d}_{\min} = 2$.

LINEAR BLOCK CODES

- Almost all block codes used today belong to a subset called *linear block codes*.
- A linear block code is a code *in which the* exclusive OR (addition modulo-2) of two valid code words creates another valid codeword.

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



Datawords	Codewords	Datawords	Codewords
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

Example

- Let us see if the two codes defined in Tables above, belong to the class of linear block codes.

1. The scheme in the above Table is a linear block code because the result of XORing any codeword with any other codeword is a valid codeword.

   So the minimum Hamming distance is $d_{min} = 2$.

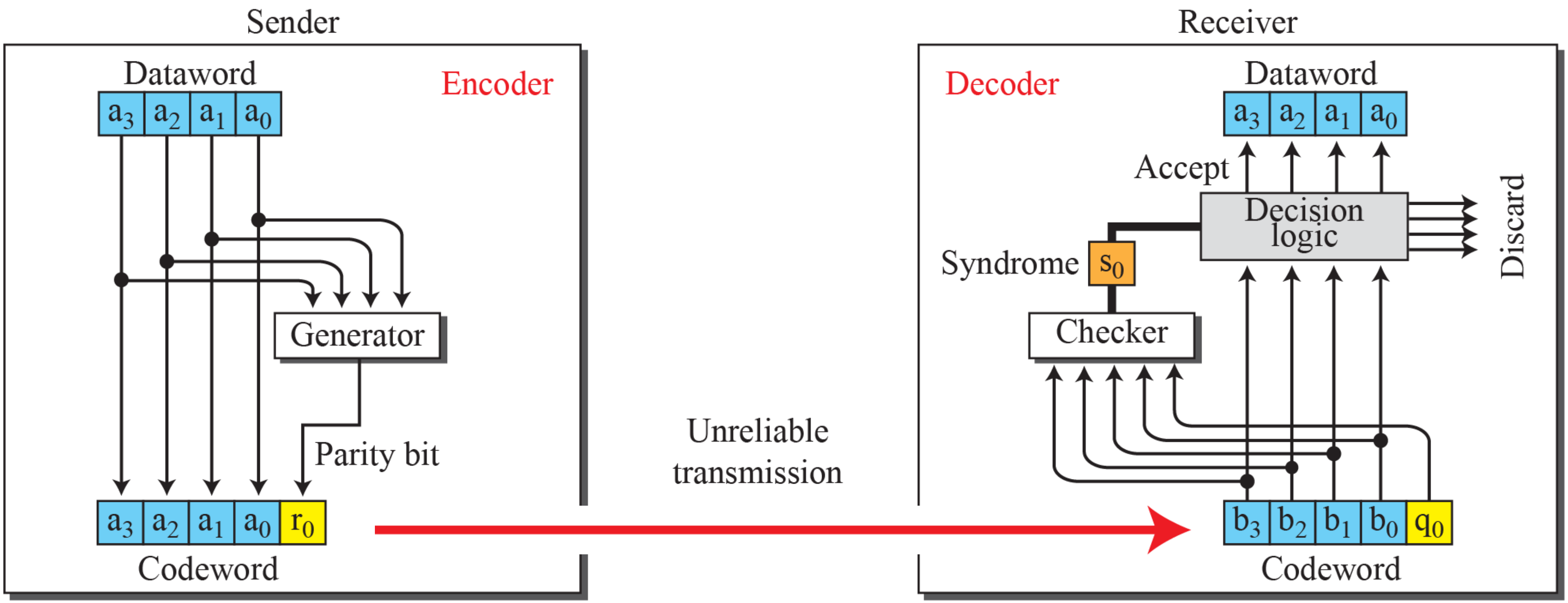
2. The scheme in Table second is also a linear block code. We can create all four codewords by XORing two other codewords.

- *So the minimum hamming distance is $d_{min} = 3$.*

Simple Parity check code

- A simple parity-check code is a single-bit error-detecting code in which $n = k + 1$ with $d_{min} = 2$.
- The extra bit, called parity bit , is selected to make number of 1's in the code word even.
- The code is single-bit error detecting code.
- It can not correct any error.
- In the fig. encoder takes a copy of 4 bit data word($a_{0,}a_{1,}a_{2,}a_3$) and generates a parity bit $r_{0.}$
- Parity bit creates the 5-bit code word.

Encoder and decoder for simple parity-check code



Simple Parity check code

- The parity bit that is added makes the number of 1's in the codeword even. i.e

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

- If number of 1s is even, the result is 0.
- If number of 1s is odd, the result is 1.

→The sender sends the code word which is corrupted during transmission.

→The receiver receives 5 bit code word.

→The checker performs modulo-2 of all the bits received. The result is called syndrome.

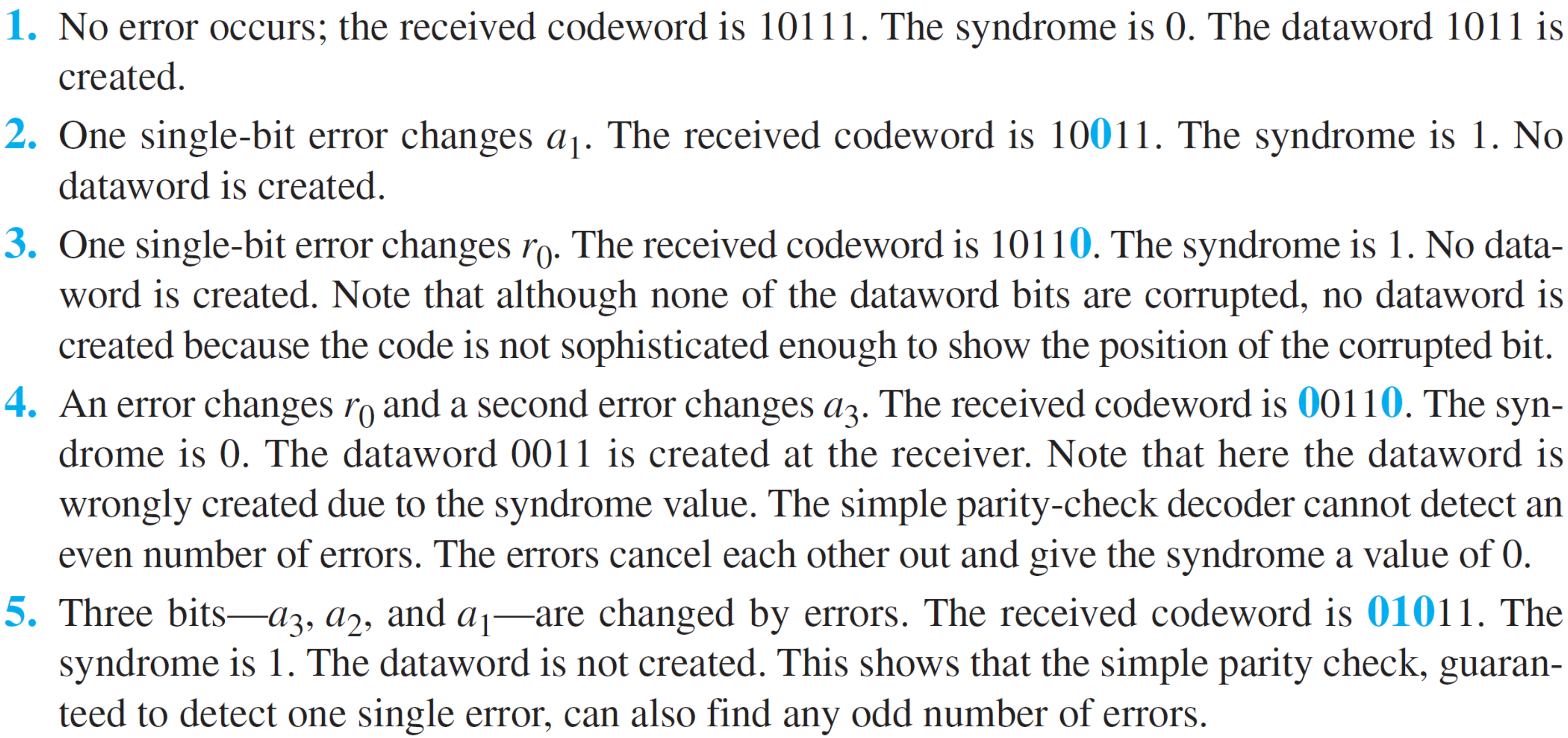
$$s_0 = b_0 + b_1 + b_2 + b_3 + q_0 \quad (\text{modulo-2})$$

Continued..

- If number of 1s in the codeword is even, the syndrome is 0.
- If number of 1s in the codeword is odd, the syndrome is 1.
- The syndrome is passed to decision logic analyzer. If Syndrome is 0, no error in the code word received, data portion is extracted.
- If syndrome is 1, error in the received codeword and discarded.

Example 10.7

Let us look at some transmission scenarios. Assume the sender sends the data word 1011. The codeword created from this data word is 10111, which is sent to the receiver. Five cases can be examined:



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 10**0**11. The syndrome is 1. No dataword is created.
3. One single-bit error changes r_0. The received codeword is 1011**0**. 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 **0**011**0**. 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 **010**11. 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.

CYCLIC CODES

- Cyclic codes are special linear block codes with one extra property.
- In a cyclic code, if a codeword is cyclically shifted (rotated), the result is another codeword.

Eg:1011000 is a codeword and if it is cyclically left-shifted, then 0110001 is also a codeword.

We can shift the bits using

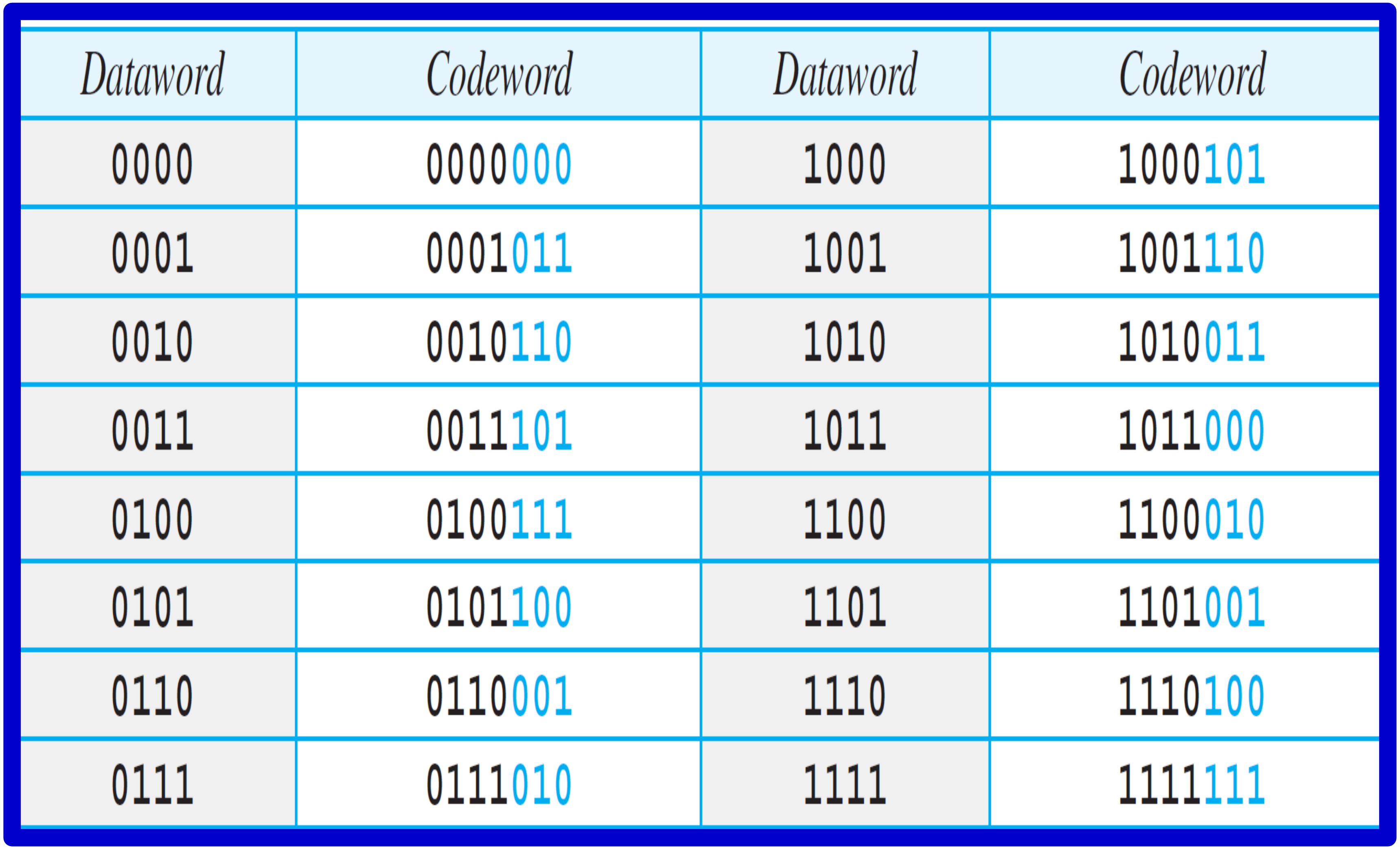
$b_1=a_{0,} b_2=a_{1,}$ $b_3=a_{2,}$ $b_4=a_{3,}$ $b_5=a_{4,}$ $b_6=a_{5,}$ $b_0=a_6$

Cyclic Redundancy Check

- Cyclic codes are created to correct errors.

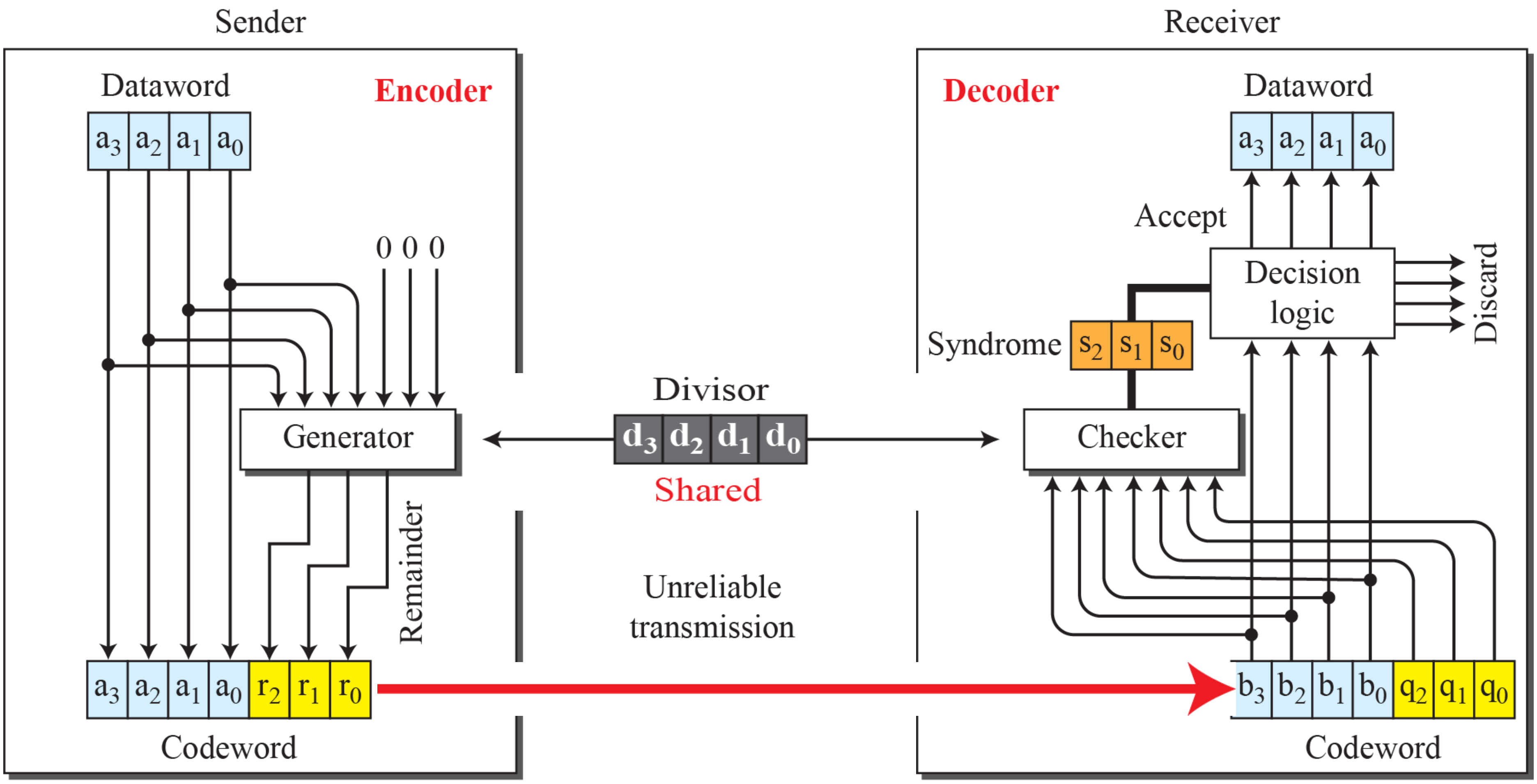
- Discuss a subset of cyclic codes called the cyclic redundancy check (CRC), which is used in networks such as LANs and WANs.

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

CRC encoder and decoder



CRC

- In the encoder, the data word has k bits(4 bits), code word has n bits(7 here).
- The data word is augmented by adding n-k 0s to the right hand side of the word.
- The n-bit result is fed to the generator.
- The generator uses a divisor of size

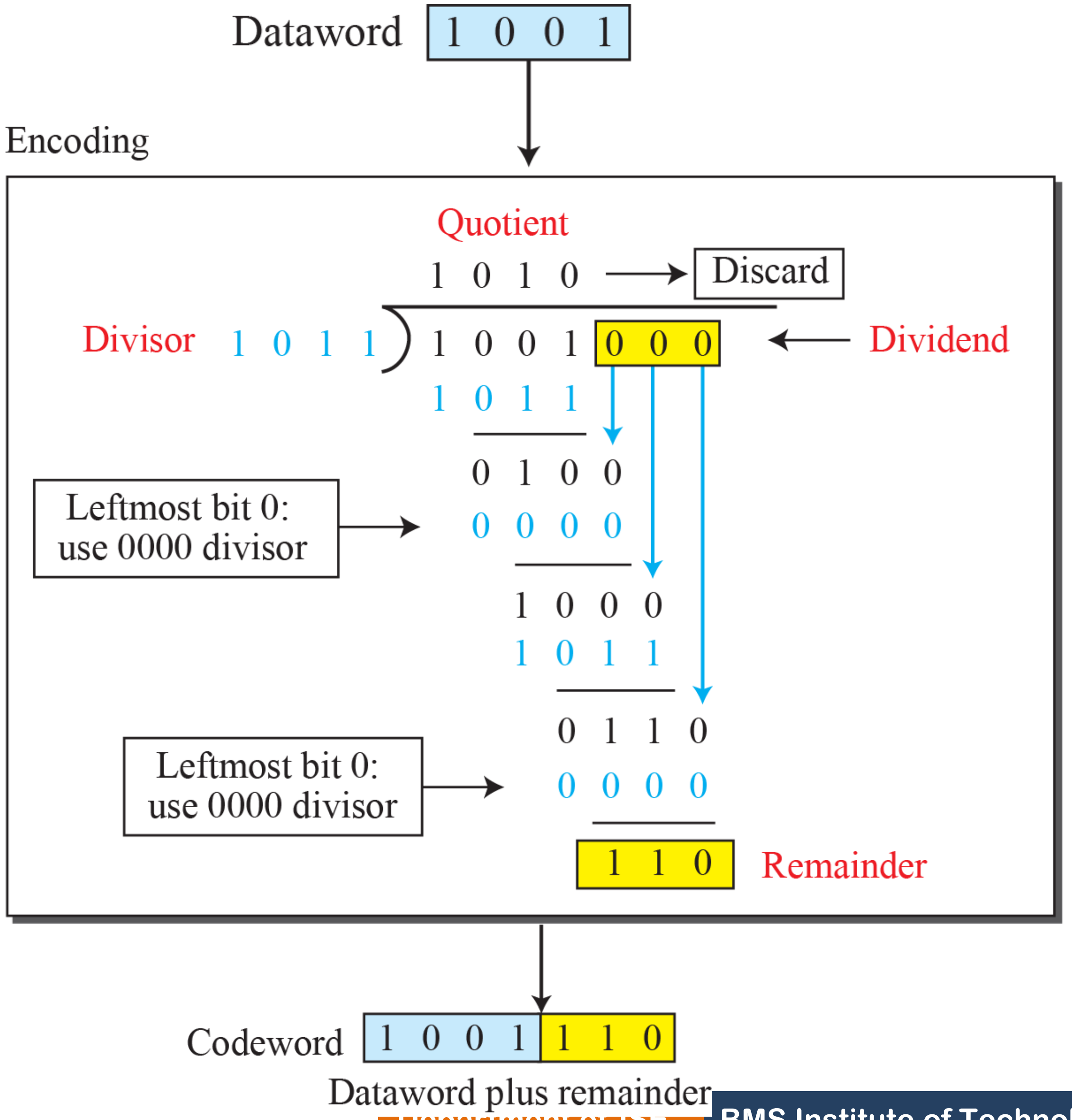
  n-k+1which is predefined and agreed upon.
- The generator divides the augmented dataword

CRC

by divisor (modulo-2 division).

- The quotient of the division is discarded.
- The remainder $r_2r_1\ r_0$ is appended to the data word to create the codeword.
- The decoder receives the corrupted codeword.
- A copy of all n bits are fed to the checker.
- The remainder produced by checker is a syndrome of n-k bits, which is fed to the decision logic analyzer.
- If syndrome bits are 0, the 4 left most bits of the code word are accepted as a data word.

Division in CRC encoder



Note:

Multiply: AND
Subtract: XOR

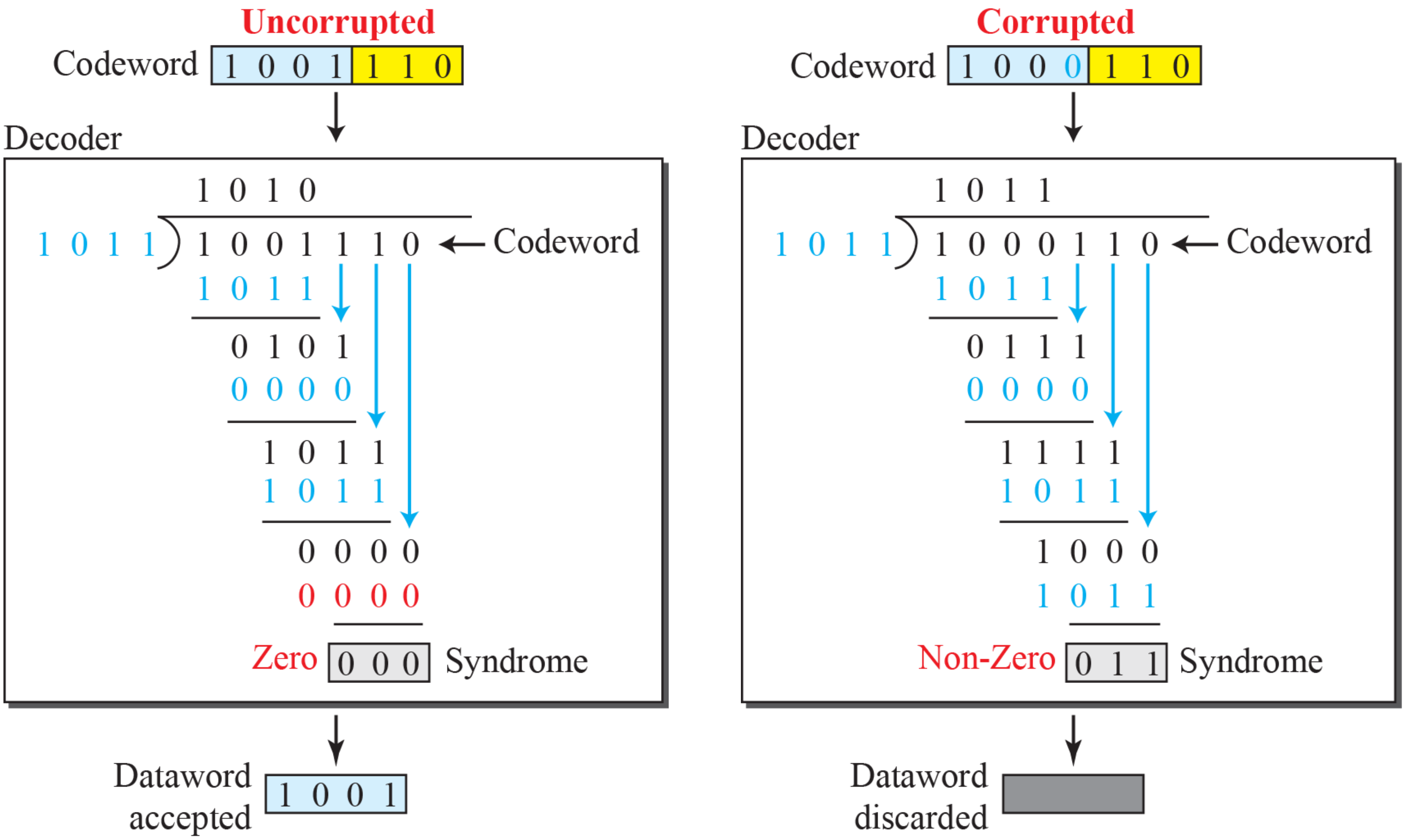
CRC-Encoder

- Encoder takes data word and augments it with n-k number of 0s.
- It then divides the augmented data word by a divisor.
- Modulo-2 binary division is used. For Addition and subtraction, XOR operations is used.
- In each step, copy of the divisor is XOR ed with 4 bits of dividend. The result of XOR operation is 3 bits.

CRC-Decoder

- Code word can change during transmission.
- The decoder same thing as encoder and remainder of the division id the syndrome.
- If syndrome is all 0s, no error, data word is separated from the code word and accepted.
- Otherwise , everything is discarded.
- Question is how divisor 1011 is chosen??

Division in the CRC decoder for two cases



Uncorrupted

Codeword 1 0 0 1 1 1 0

Decoder

```
        1 0 1 0
1 0 1 1)1 0 0 1 1 1 0  ← Codeword
        1 0 1 1
          0 1 0 1
          0 0 0 0
            1 0 1 1
            1 0 1 1
              0 0 0 0
              0 0 0 0
```

Zero 0 0 0 Syndrome

Dataword accepted 1 0 0 1

Corrupted

Codeword 1 0 0 0 1 1 0

Decoder

```
        1 0 1 1
1 0 1 1)1 0 0 0 1 1 0  ← Codeword
        1 0 1 1
          0 1 1 1
          0 0 0 0
            1 1 1 1
            1 0 1 1
              1 0 0 0
              1 0 1 1
```

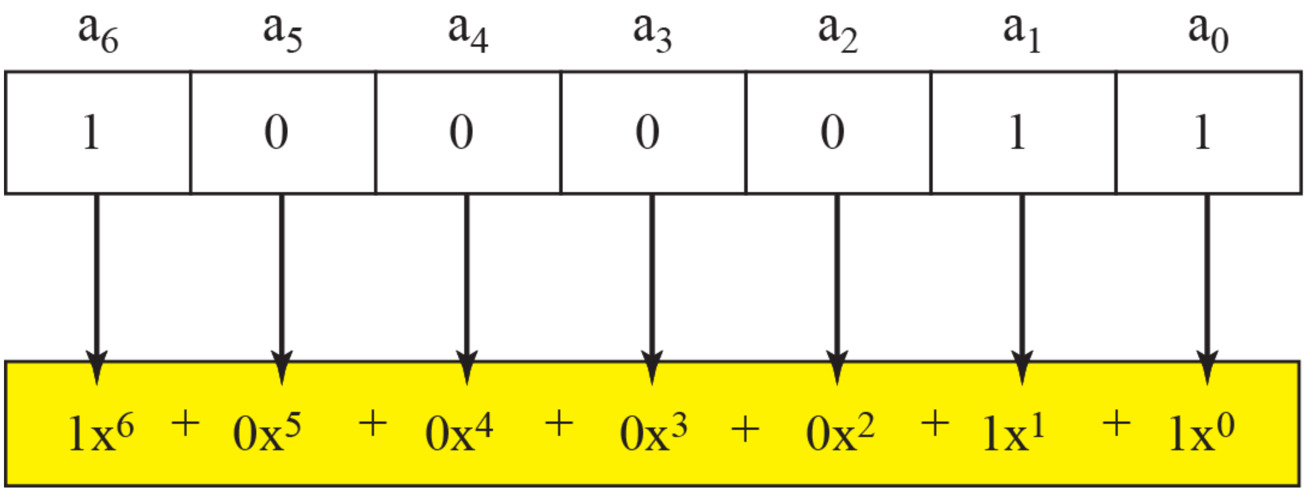
Non-Zero 0 1 1 Syndrome

Dataword discarded

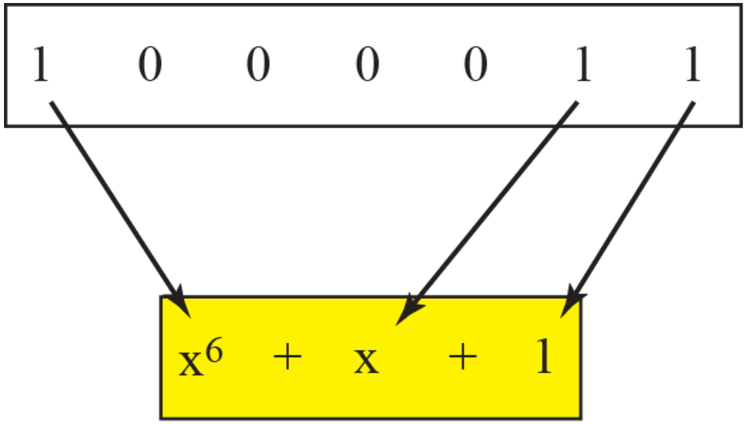
Polynomials

- A better way to understand cyclic codes, and how they can be analyzed is to represent them as polynomials.
- A pattern of 0s and 1s can be represented as polynomial with coefficient of 0 and 1.
- The power of each term shows the position of the bit.
- The coefficient shows the value of the bit.

A polynomial to represent a binary word



a. Binary pattern and polynomial



b. Short form

Degree of Polynomials

- Is the highest power of the polynomial.

Eg: Degree of $x^6 + x+1$ is 6.

- Degree of polynomial is one less than that the number of bits in the pattern.
- Addition and subtraction of polynomials is done by adding and subtracting the coefficient terms with the same power.
- Coefficients are only 0 and 1, and adding is in modulo-2

Adding and subtracting of polynomials

- Addition and subtraction is done by combining terms and deleting pair of identical terms.
- Eg: $x^5 + x^4 + x^2$

$$x^6 \quad\quad\quad + x^4 + x^2$$

$$x^6 + x^5$$

Multiplication and Division of polynomials

- Multiplication is adding powers.

Eg: $x^4 * x^2 = x^6$

- Division, subtract the power of second term from the first term.

Eg: $x^5 / x^2 = x^3$

- Multiplication of two polynomials

$(x^5 + x^3 + x^2)\ (x^2 + x + 1) = x^7 + x^6 + x^5 + x^5 + x^4 + x^3 + x^4 + x^3 + x^2$

$\rightarrow x^7 + x^6 + x^2$

Shifting of polynomials

- A binary pattern is shifted left or right by number of bits.
- Shifting to the left means adding extra 0s as right most bits.
- Shifting to the right means deleting some right most bits.
- Shifting to the left is accomplished by multiplying each term of the polynomial by x^m where m is the number of shifted bits

Continued..

- Shifting to the right is accomplished by dividing each term of the polynomial by x^m where m is the number of shifted bits.

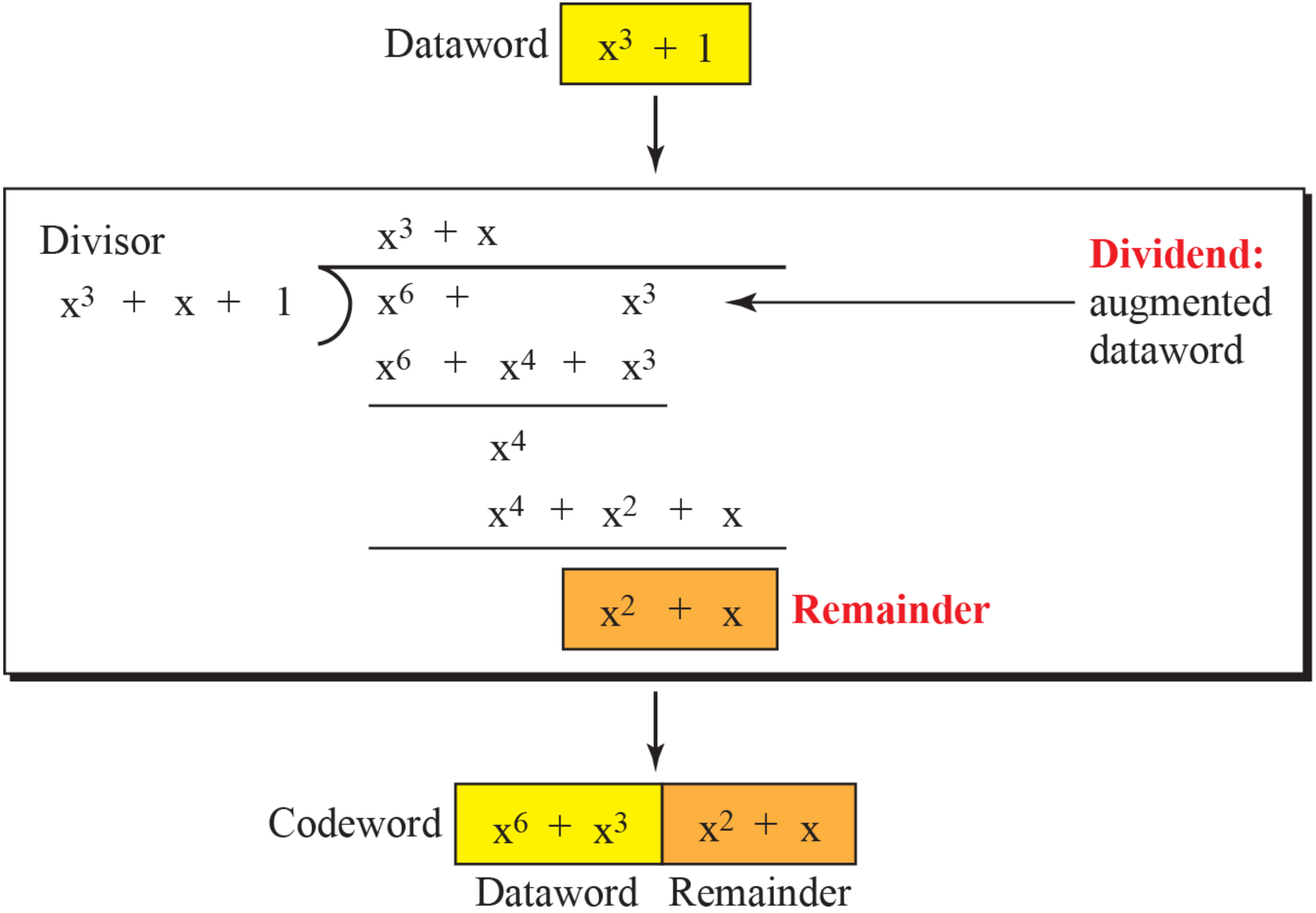
- Shift left by 3 bits :10011 becomes 10011000

  x^4+x+1 becomes $x^7 + x^4+x^3$

- Shift right by 3 bits: 10011 becomes 10

  x^4+x+1 becomes x

CRC division using polynomials



Dataword $x^3 + 1$

Divisor $x^3 + x + 1$

$x^3 + x$

$x^6 + x^3$

$x^6 + x^4 + x^3$

x^4

$x^4 + x^2 + x$

$x^2 + x$ Remainder

Dividend: augmented dataword

Codeword $x^6 + x^3$ $x^2 + x$

Dataword Remainder

Cyclic code analysis using polynomials

- In a polynomial representation, divisor is normally referred to as generator polynomial t(x).
- f(x) is the polynomial with binary coefficients

data word d(x), code word c(x), Generator g(x)

syndrome s(x) and error e(x).

→ Is s(x) is not zero, one or more bits are corrupted. If s(x) is zero, either no error or decoder failed to detect any error.

Cyclic Code Analysis

- Analyze a cyclic code to find its capabilities by using polynomials.

- Define the following, where f(x) is a polynomial with binary coefficients.

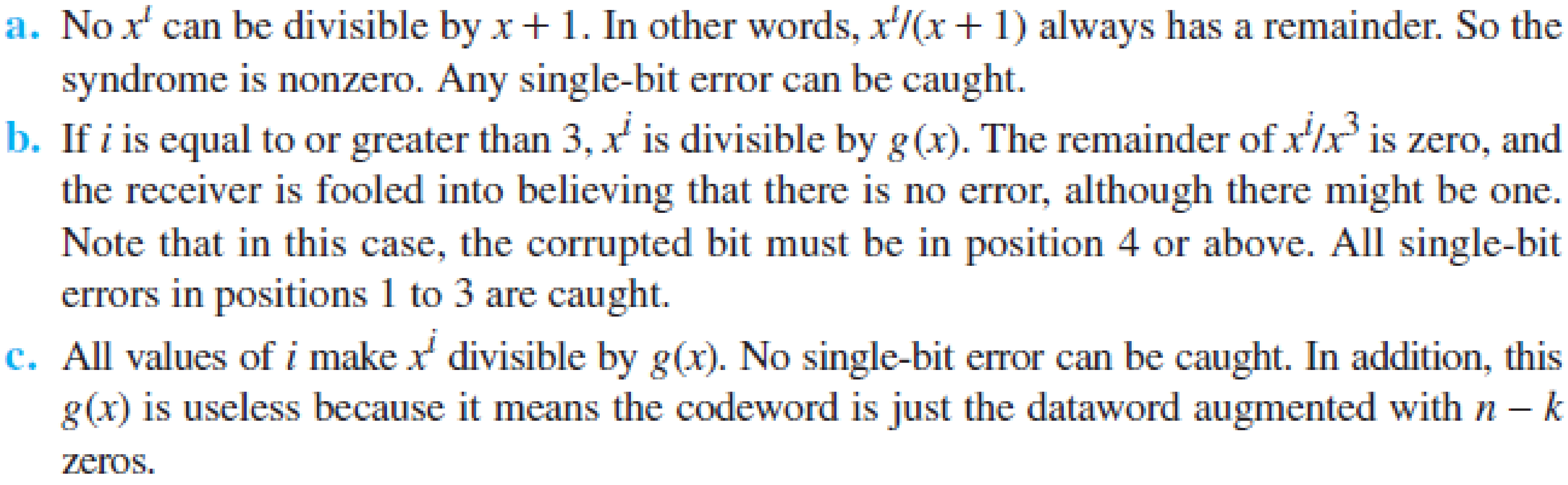
Dataword: $d(x)$ **Codeword:** $c(x)$

Generator: $g(x)$ **Syndrome:** $s(x)$ **Error:** $e(x)$

Example 10.8

Which of the following g(x) values guarantees that a single-bit error is caught? x + 1, x^3 and 1

Solution

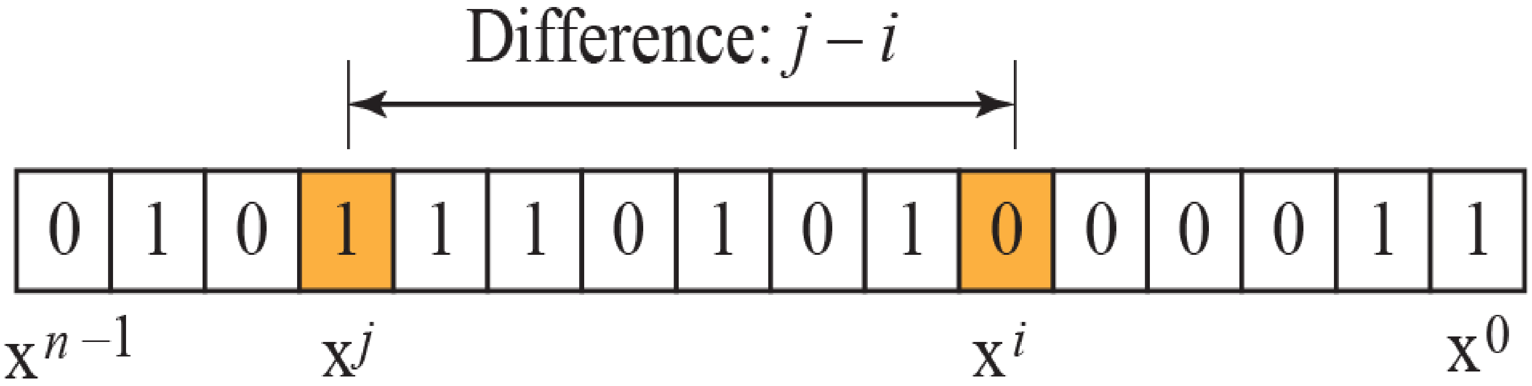


a. No x^i can be divisible by $x + 1$. In other words, $x^i/(x + 1)$ always has a remainder. So the syndrome is nonzero. Any single-bit error can be caught.

b. If i is equal to or greater than 3, x^i is divisible by $g(x)$. The remainder of x^i/x^3 is zero, and the receiver is fooled into believing that there is no error, although there might be one. Note that in this case, the corrupted bit must be in position 4 or above. All single-bit errors in positions 1 to 3 are caught.

c. All values of i make x^i divisible by $g(x)$. No single-bit error can be caught. In addition, this $g(x)$ is useless because it means the codeword is just the dataword augmented with $n - k$ zeros.

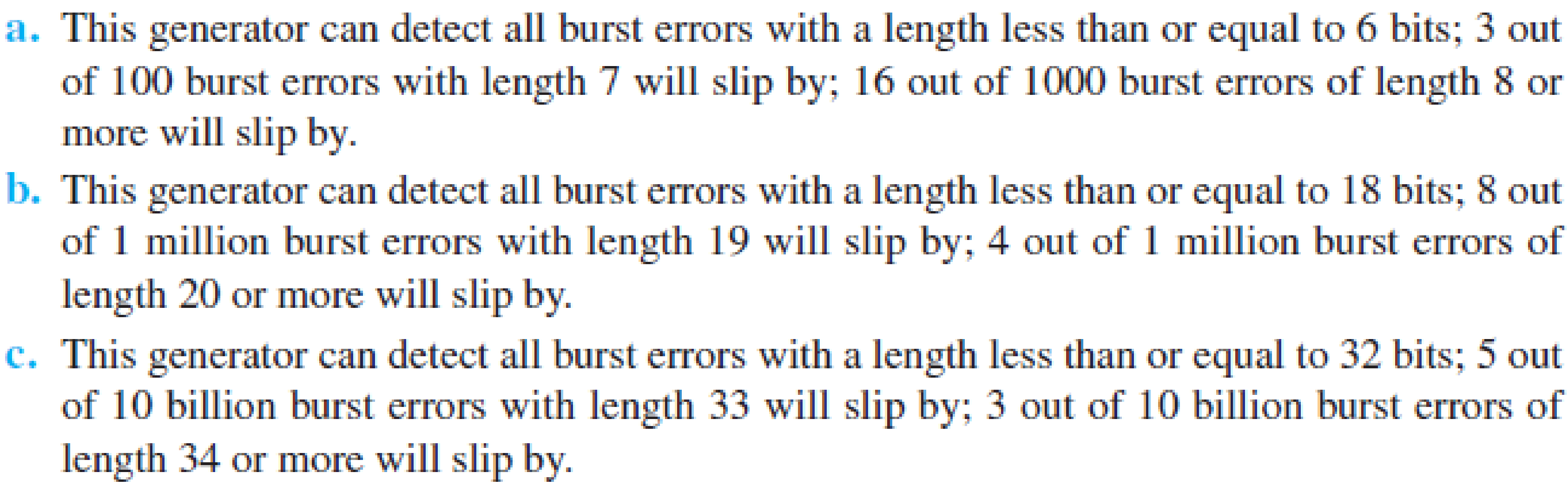
Representation of isolated single-bit errors



Example 10.9

Find the suitability of the following generators in relation to burst errors of different lengths: $x^6 + 1$, $x^{18} + x^7 + x + 1$, and $x^{32} + x^{23} + x^7 + 10$.

Solution



a. This generator can detect all burst errors with a length less than or equal to 6 bits; 3 out of 100 burst errors with length 7 will slip by; 16 out of 1000 burst errors of length 8 or more will slip by.

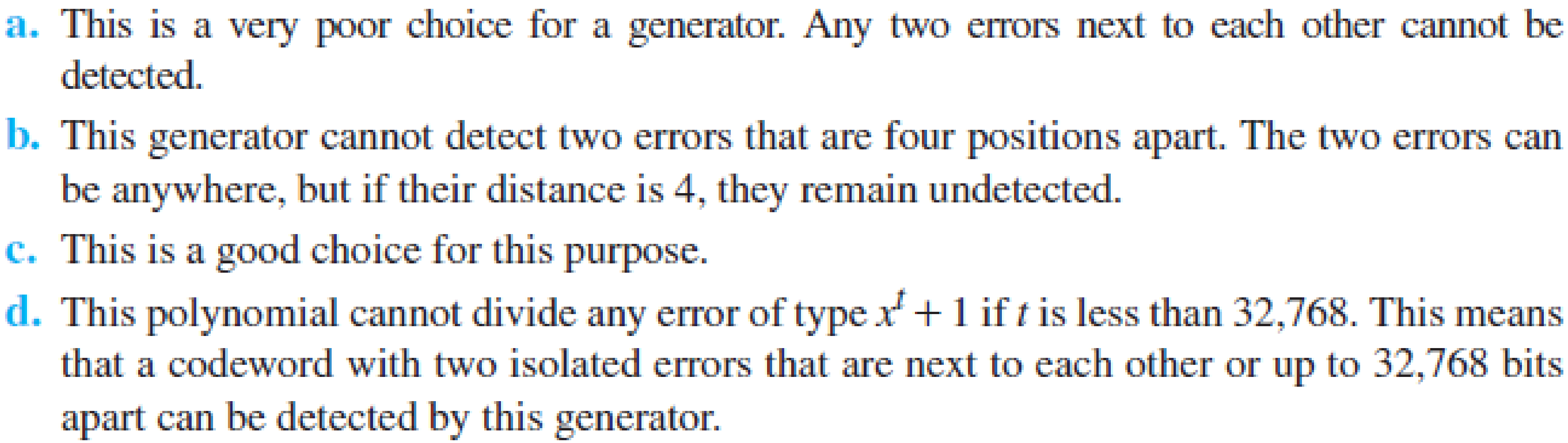
b. This generator can detect all burst errors with a length less than or equal to 18 bits; 8 out of 1 million burst errors with length 19 will slip by; 4 out of 1 million burst errors of length 20 or more will slip by.

c. This generator can detect all burst errors with a length less than or equal to 32 bits; 5 out of 10 billion burst errors with length 33 will slip by; 3 out of 10 billion burst errors of length 34 or more will slip by.

Example 10.10

Find the status of the following generators related to two isolated, single-bit errors: $x + 1$, $x^4 + 1$, $x^7 + x^6 + 1$, and $x^{15} + x^{14} + 1$

Solution



a. This is a very poor choice for a generator. Any two errors next to each other cannot be detected.

b. This generator cannot detect two errors that are four positions apart. The two errors can be anywhere, but if their distance is 4, they remain undetected.

c. This is a good choice for this purpose.

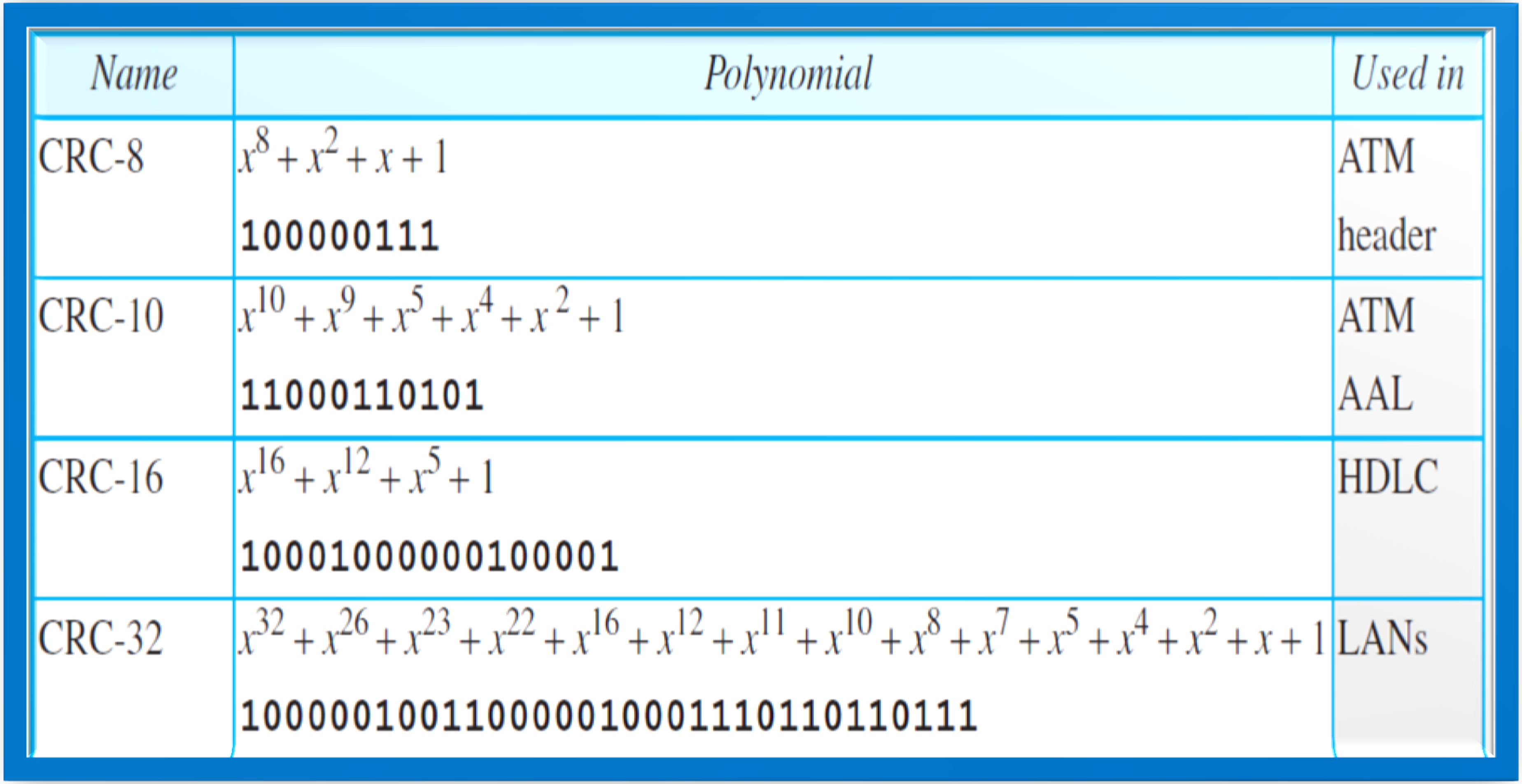
d. This polynomial cannot divide any error of type $x^t + 1$ if t is less than 32,768. This means that a codeword with two isolated errors that are next to each other or up to 32,768 bits apart can be detected by this generator.

Summary

- 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 $x + 1$.

Table :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	$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$ 100000100110000010001110110110111	LANs

Advantages of Cyclic Codes

- Cyclic codes have a very good performance in detecting single-bit errors, double errors, an odd number of errors, and burst errors.

- They can easily be implemented in hardware and software.

- They are especially fast when implemented in hardware which has made cyclic codes a good candidate for many networks.

Other Cyclic Codes

- The cyclic codes we have discussed are very simple.

- The check bits and syndromes can be calculated by simple algebra.

- There are, however, more powerful polynomials that are based on abstract algebra involving Galois fields.

- One of the most interesting of these codes is the Reed-Solomon code used today for both detection and correction.

CHECKSUM

- Checksum is an error-detecting technique that can be applied to a message of any length.

- In the Internet, the checksum technique is mostly used at the network and transport layer rather than the data-link layer.
- At source,

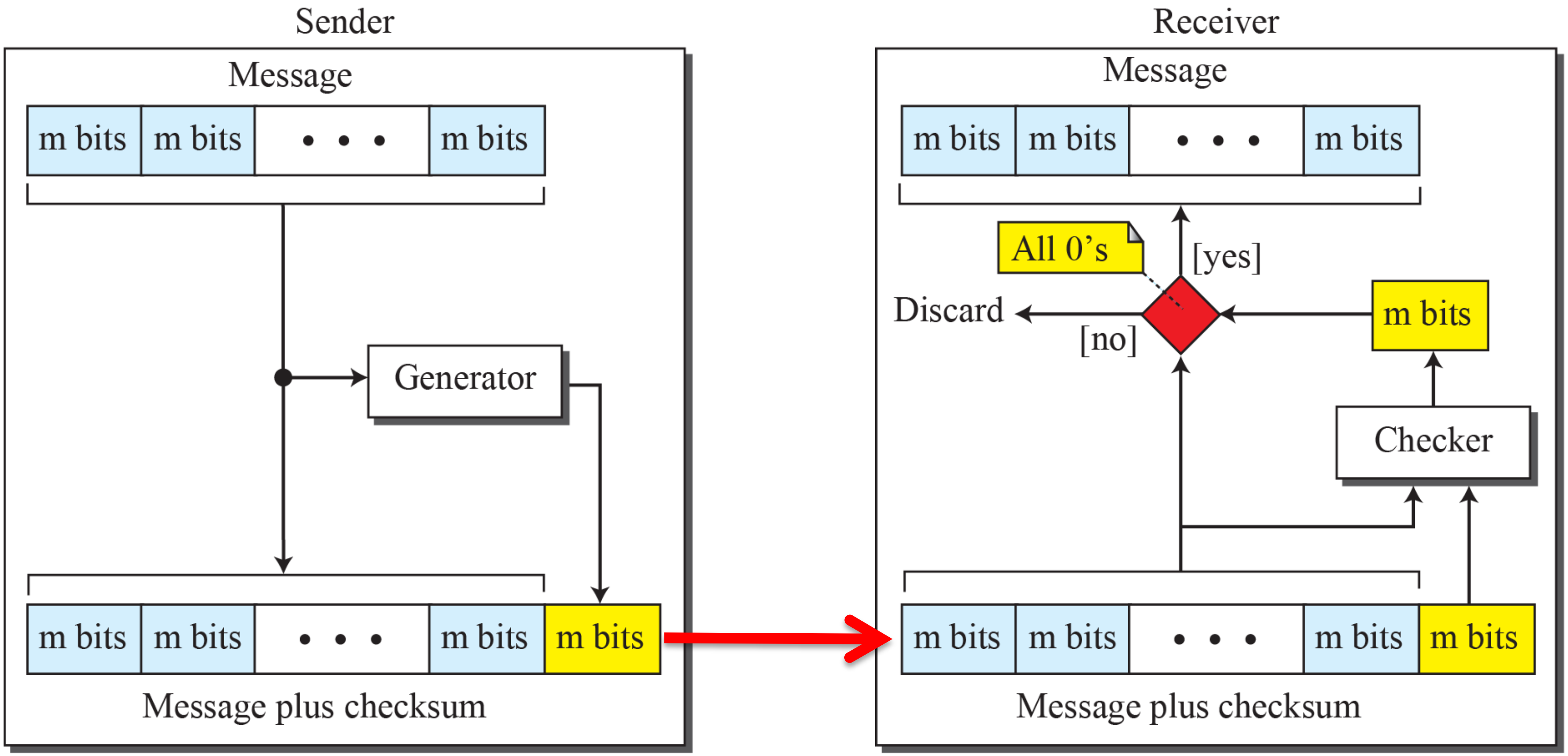
➢ Message is divided into m-bit units.

➢ Generator then creates an extra m-bit unit called checksum, which is sent with message.

CHECKSUM

- At Destination,
  - ➢Checker creates a new checksum from combination of the message and the sent checksum.
  - ➢If new checksum is all 0’s, the message is accepted else discarded.

Checksum


Sender
Message
m bits
m bits
• • •
m bits
Generator
m bits
m bits
• • •
m bits
m bits
Message plus checksum
Receiver
Message
m bits
m bits
• • •
m bits
All 0's
[yes]
Discard
[no]
m bits
Checker
m bits
m bits
• • •
m bits
m bits
Message plus checksum

Concept

- The idea of the traditional checksum is simple.

Show this using a simple example.

Suppose the message 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.

Eg: if the set of numbers is (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 two are the same, the receiver assumes no error, accepts the five numbers, and discards the sum.

Otherwise, there is an error somewhere and the message not accepted.

Ones compliment Addition

- Each number can be written as a 4 bit word except sum.
- This drawback we can over come using ones compliment arithmetic.
- Represent unsigned numbers between 0 and 2^m -1 Using only m bits.
- If the number has more than m bits, the extra leftmost bits need to be added to the m rightmost bits.

Example 10.12

In the previous example, the decimal number 36 in binary is (100100)2. To change it to a 4-bit number we add the extra leftmost bit to the right four bits as shown below.

$$\boxed{(10)_2 + (0100)_2 = (0110)_2 \rightarrow (6)_{10}}$$

Instead of sending 36 as the sum, we can send 6 as the sum (7, 11, 12, 0, 6, 6).

The receiver can add the first five numbers in one’s complement arithmetic.

If the result is 6, the numbers are accepted; otherwise, they are rejected.

Example 10.13

Let us use the idea of the checksum in Example 10.12. The sender adds all five numbers in one's complement to get the sum = 6.

The sender then complements the result to get the checksum = **9**, which is 15 − 6.

➢Note that 6 = (0110)2 and **9** = (1001)2;they are complements of each other.

➢The sender sends the five data numbers and the checksum (7, 11, 12, 0, 6, **9**).

➢If there is no corruption in transmission, the receiver receives (7, 11, 12, 0, 6, **9**) and adds them in one's complement to get 15.

Example 10.13

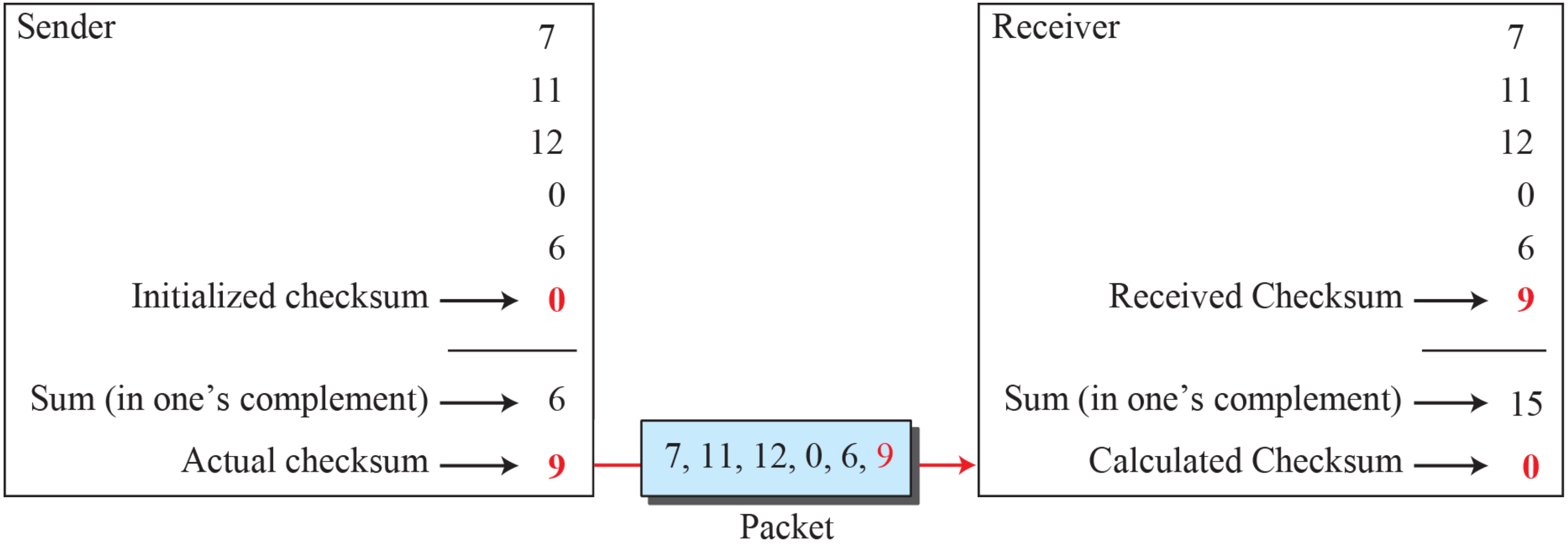
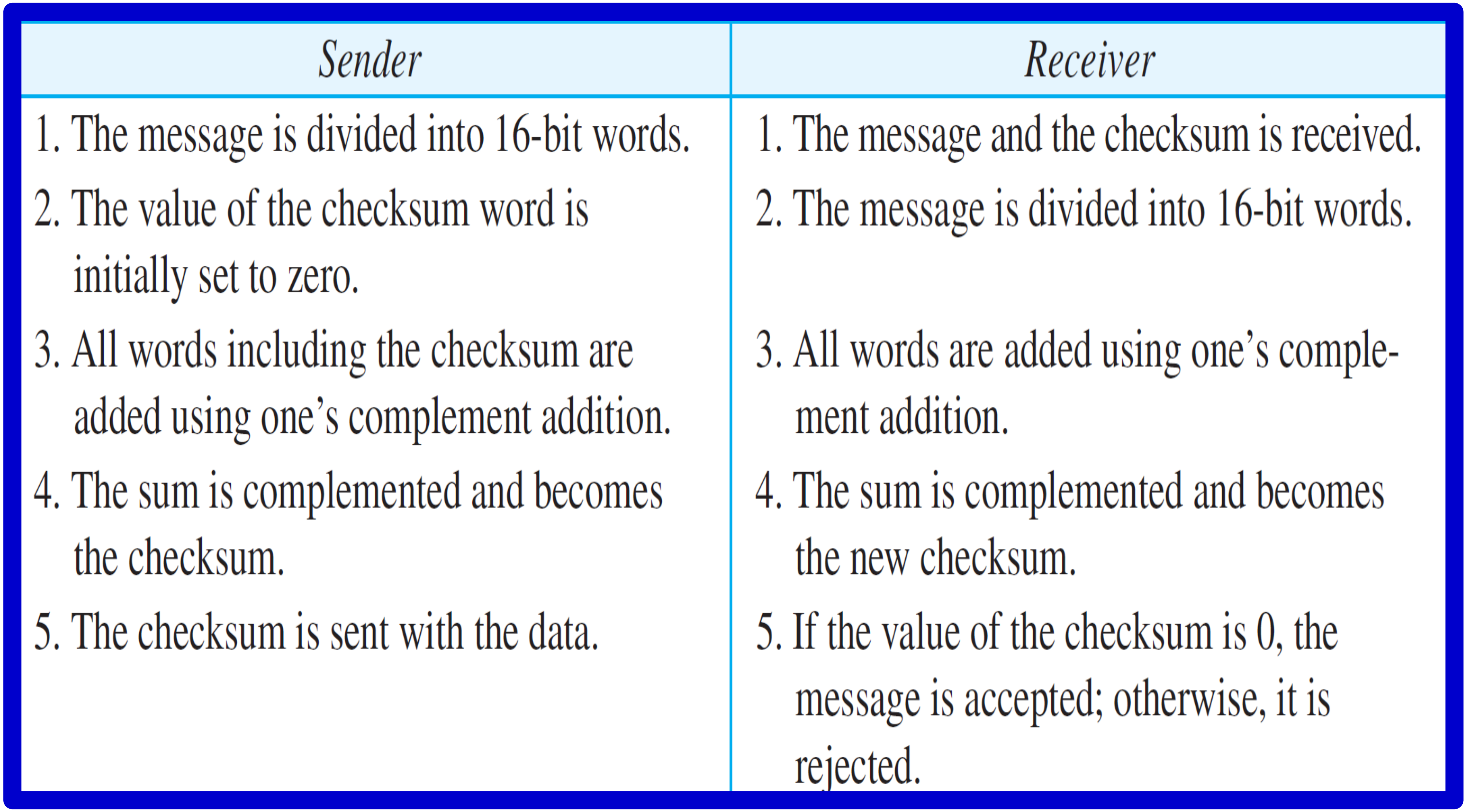
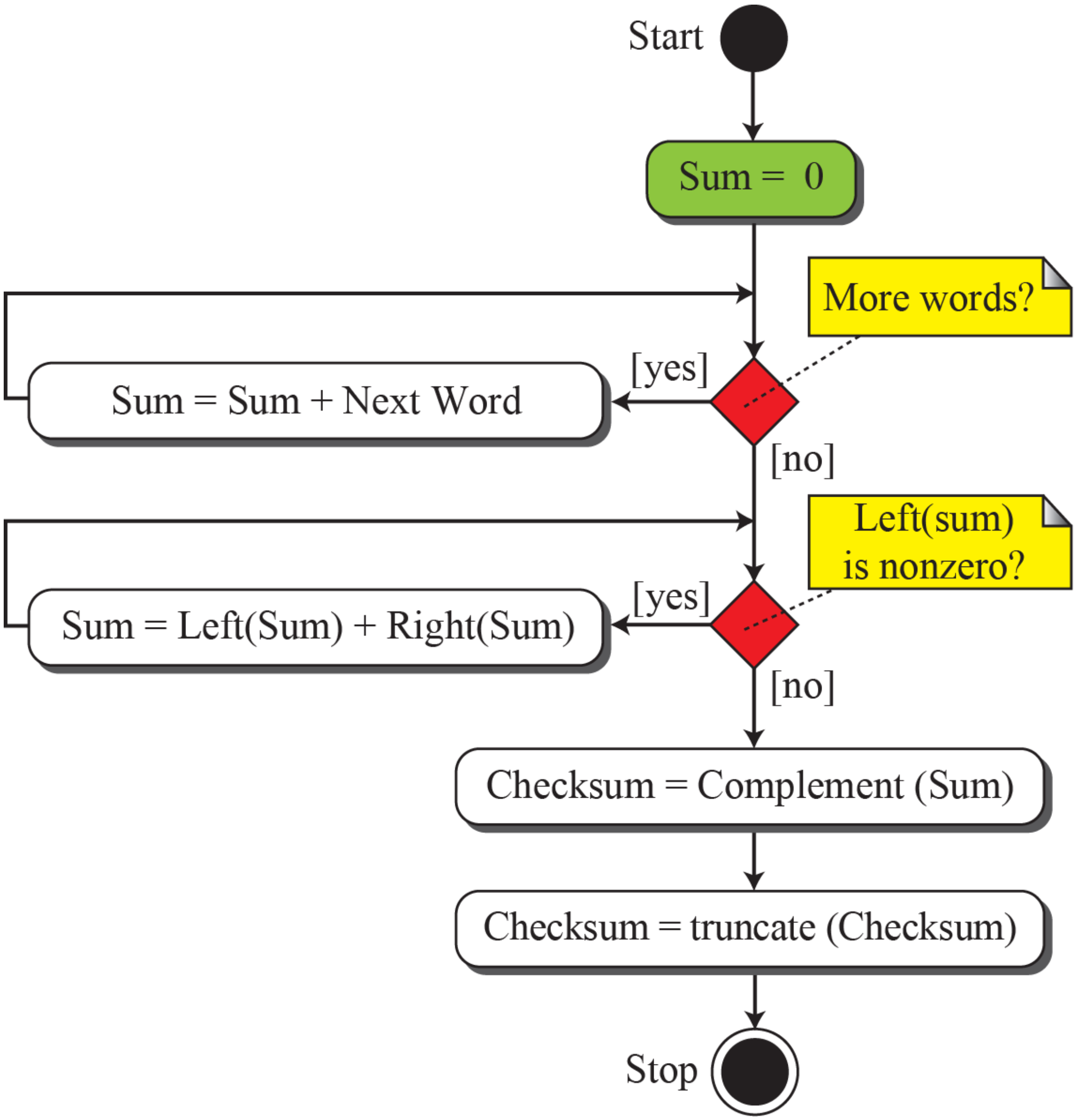


Table : Procedure to calculate the traditional checksum



Sender	*Receiver*
1. The message is divided into 16-bit words.	1. The message and the checksum is 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.

Algorithm to calculate a traditional checksum



Notes:

a. Word and Checksum are each 16 bits, but Sum is 32 bits.

b. Left(Sum) can be found by shifting Sum 16 bits to the right.

c. Right(Sum) can be found by ANDing Sum with $(0000FFFF)_{16}$.

d. After Checksum is found, truncate it to 16 bits.

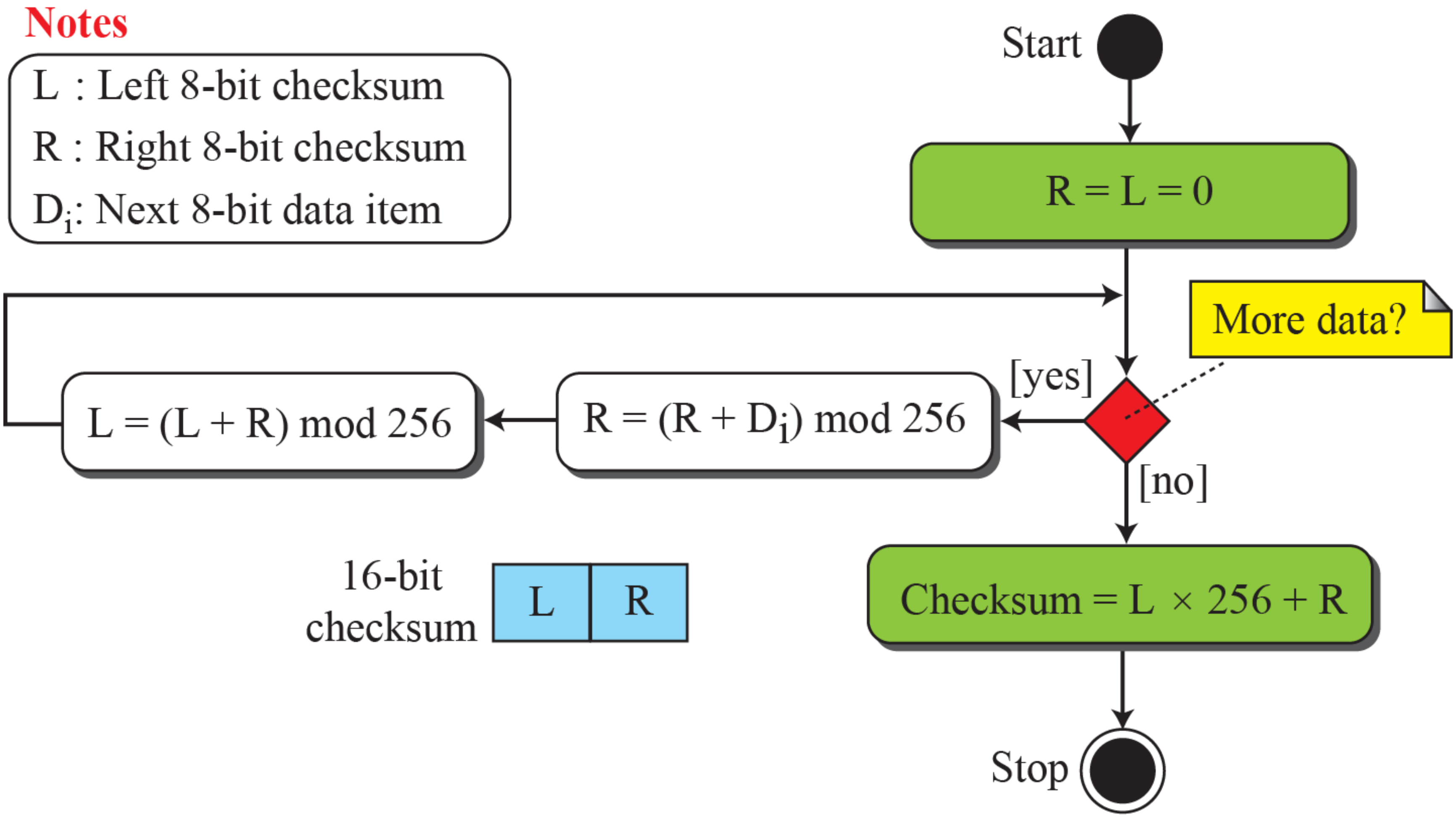
Other Approaches

- One major problem with the traditional checksum calculation.
- 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.
- Several approaches have been used to prevent this problem.
- Two of them are: Fletcher and Adler.

Fletcher Checksum

- Was devised to weight each data item according to its position.
- Two algorithms i) 8-bit ii) 16-bit.
- In 8 bit Fletcher, calculation is done on 8-bit data and creates 16 bit checksum.
- Calculation is done using modulo256(2^8)i.e intermediate results are divided by 256 and remainder is kept.
- Algorithm uses two accumulators L & R
- First simply adds data items together,
- Second adds weights to the calculation.
- In 16 bit Fletecher, calculation is done on 16-bit data and creates 32 bit checksum.

Algorithm to calculate an 8-bit Fletcher checksum



Adler Checksum

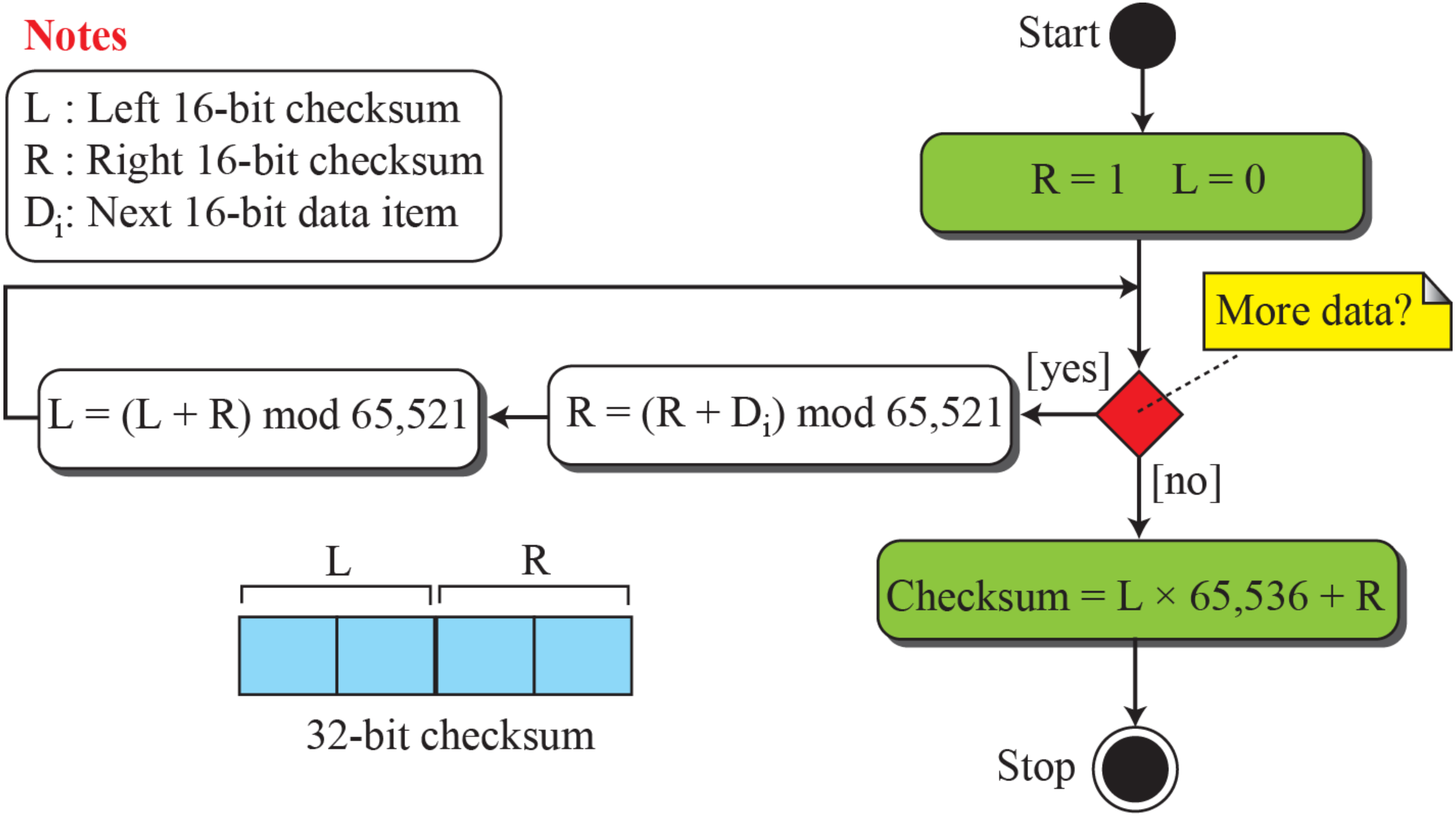
- Is a 32 bit checksum.
- Similar to Fletcher.
- 3 Major differences

i) Calculation is done on single byte instead of 2 bytes at a time

ii) Modulus is a prime number(65,521)instead of 65,536.

iii) L is initialized to 1 instead of 0

Algorithm to calculate an Adler checksum



Notes

L : Left 16-bit checksum
R : Right 16-bit checksum
D_i: Next 16-bit data item

Start

$R = 1 \quad L = 0$

More data?

[yes]

$R = (R + D_i) \bmod 65{,}521$

$L = (L + R) \bmod 65{,}521$

[no]

$\text{Checksum} = L \times 65{,}536 + R$

Stop

L

R

32-bit checksum

MODULE – 4

Data Link Control

DLC SERVICES

➢The data link control (DLC) deals with procedures for communication between two adjacent nodes no matter whether the link is dedicated or broadcast.
Data link control functions include framing and flow and error control.
➢Discuss framing, or how to organize the bits that are carried by the physical layer.
➢Discuss flow and error control.

Framing

- The data-link layer needs to pack bits into frames, so that each frame is distinguishable from another.
- 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.
- Framing in the data-link layer separates a message from one source to a destination by adding a sender address and a destination address.
- The destination address defines where the packet is to go; the sender address helps the recipient acknowledge the receipt.

Framing

- Whole message can be placed in a single frame, but that is not done because frame can be very large making Flow Control and Error Control very inefficient.

- When message is carried in one large frame, even single bit error would require the retransmission of the whole frame.

- When message is divided into small frames, single bit error affect only that small frame.

FrameSize

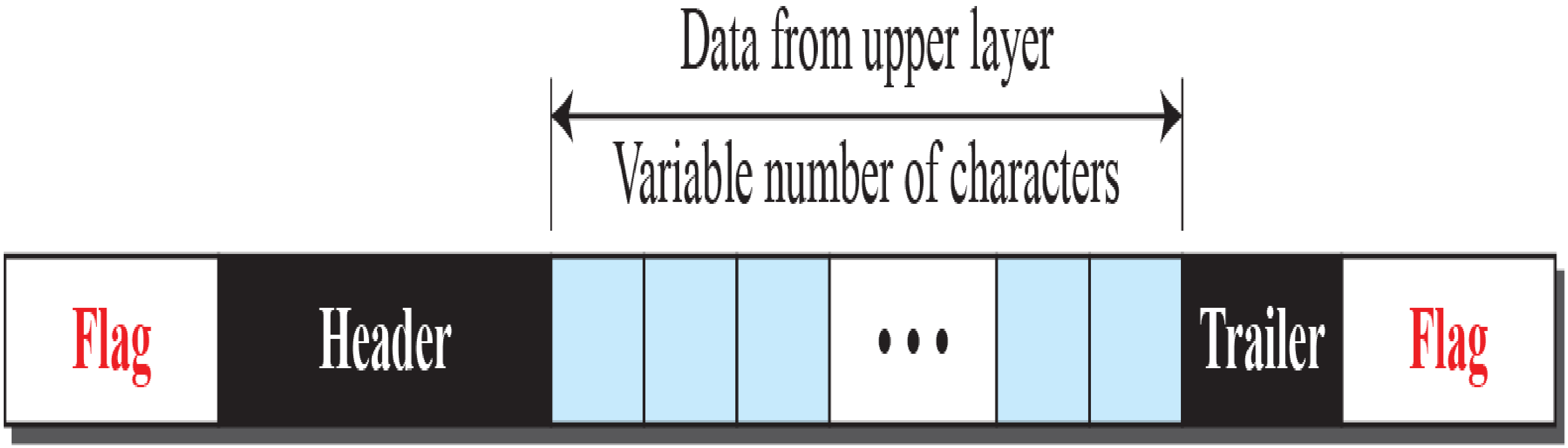
- Frame can be i) fixed size
  ii) Variable size.
- In fixed size framing, no need to define the frame boundaries. Size itself can be treated as a delimiter.
  Eg: ATM frames of fixed size called cells.

- In variable size framing, we need to define the frame boundaries to identify beginning and end of the frame.
- Two approaches are used for this purpose
  i) Character-oriented Framing
  ii) Bit-oriented Framing.

Character-oriented Framing

- Data to be carried are 8-bit characters from coding system such as ASCII.
- Header carries source, destination address along with some control information.
- Trailer carries error detection or correction redundant bits, are multiples of 8 bits.
- To separate one frame from other, a 1byte flag is added to the beginning and end of frame.
- The flag, composed of protocol dependent special characters which signals start or end of the frame.

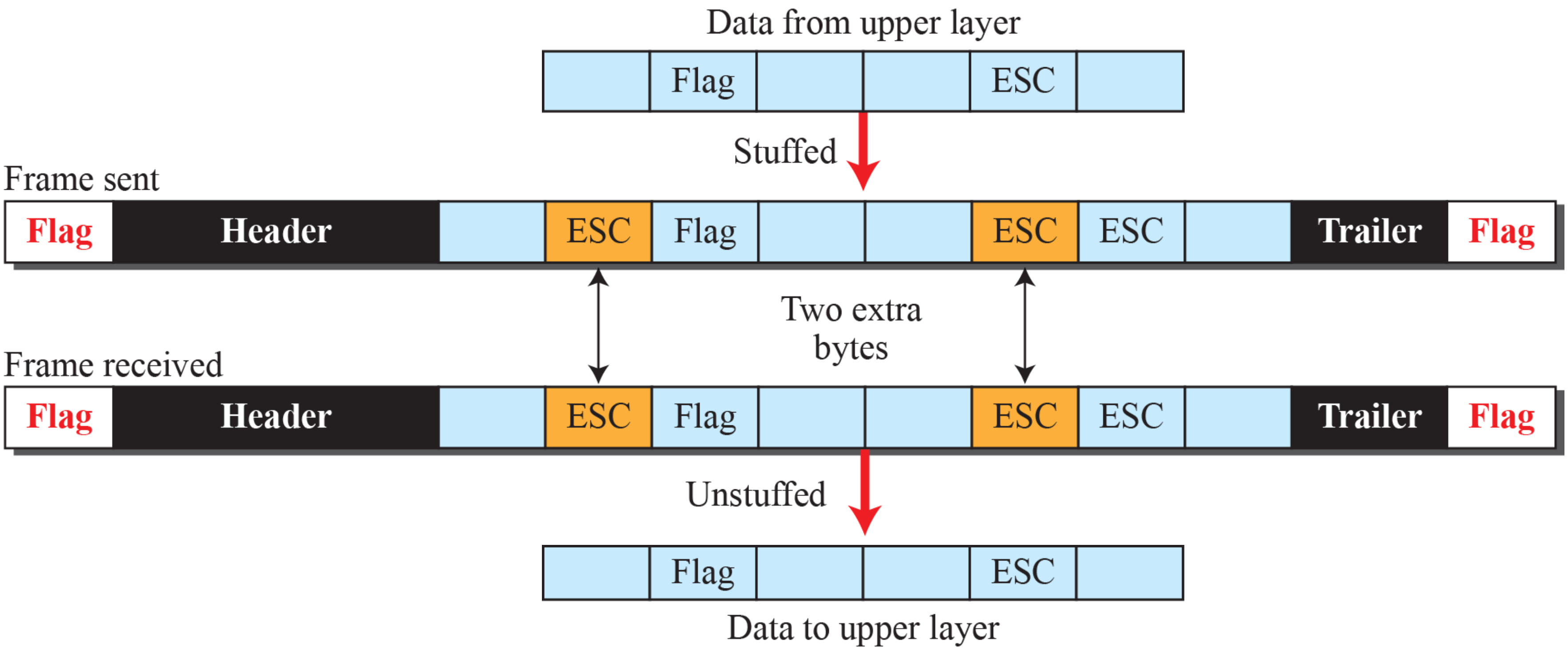
A frame in a character-oriented protocol



Character-oriented Framing

- Was popular only when text was exchanged by the DLL.
- Flag used, could be selected as any character which is not used for text communication. As we send now other types of information such graphs, audio and video, the pattern used for flag could be the part of information.
- If receiver encounters this pattern in the middle of the data, thinks it has reached the end of frame.
- Solution is byte-stuffing strategy was added in this approach.

Byte stuffing and unstuffing



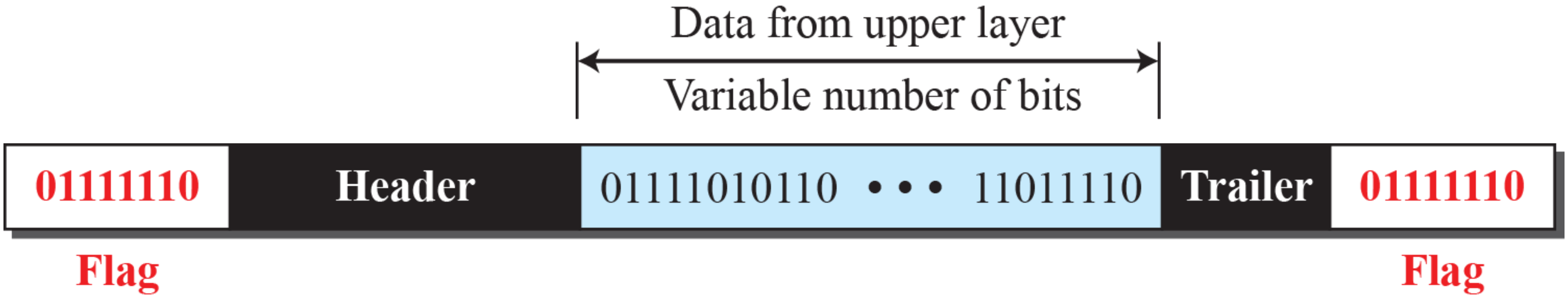
Character-oriented Framing

- 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 that of the flag.
- i.e Data section is stuffed with a special extra byte called escape character. This approach presents another problem in data communication.
- The universal coding systems, unicode, in use today, have 16 or 32 bit characters that conflict with 8-bit characters.
- Therefore, next approach is used

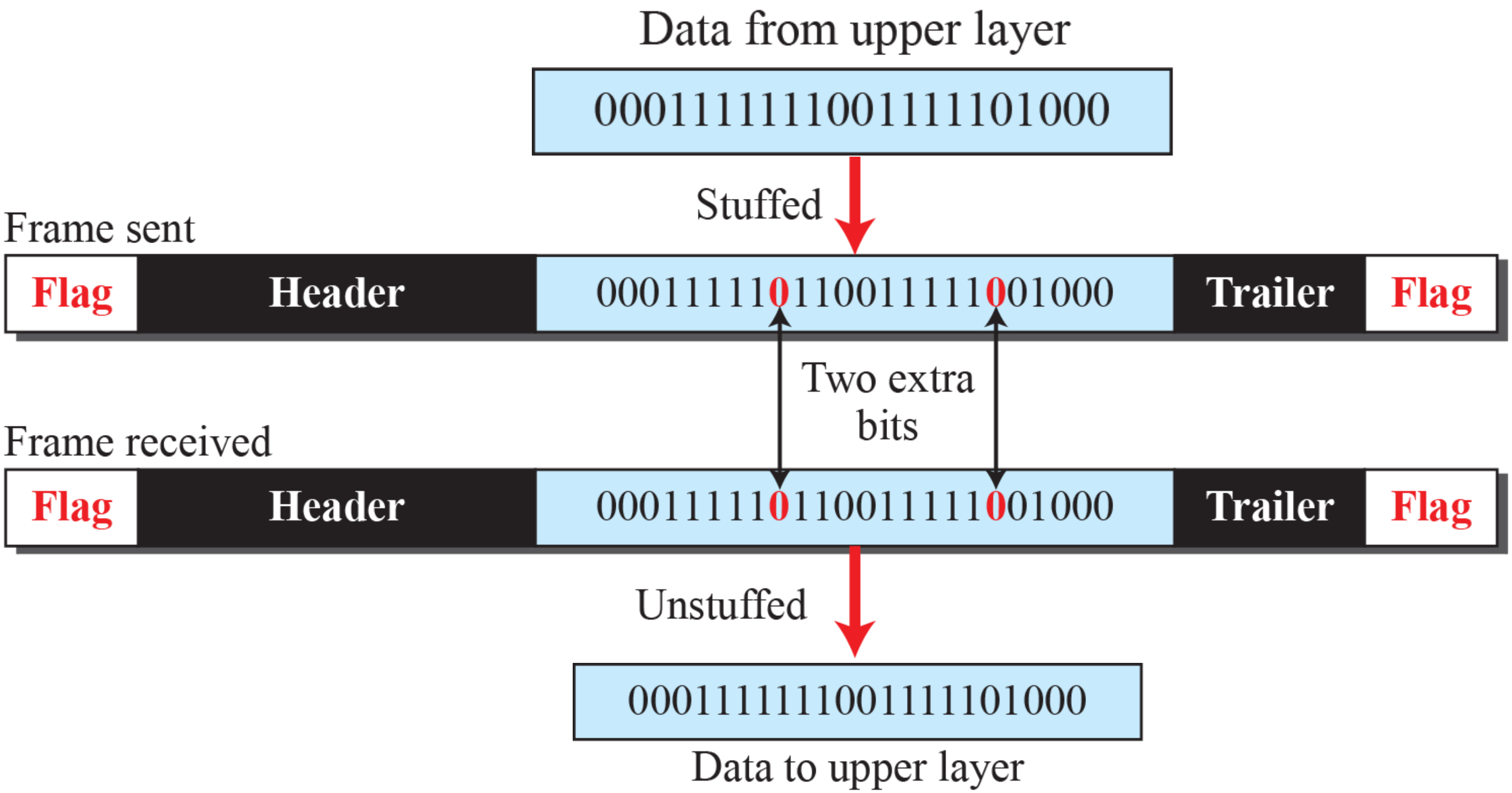
Bit-oriented Framing

- Data section of a frame is a sequence of bits to be interpreted by the upper layer as text, graphs, video and so on.
- Delimiter used to separate one frame from another and used in the beginning and end of frame is 8-bit pattern flag(01111110).
- Flag can create same problem. If flag pattern in present in the frame, need to inform the receiver that , this is data portion of the frame, not end of frame.
- This strategy is called bit stuffing.
- In bit stuffing, if a 0 and five consecutive 1 bits are encountered, an extra 0 bit added .
- Which is removed from the data by the receiver.

A frame in a bit-oriented protocol



Bit stuffing and unstuffing



Data from upper layer

0001111111001111101000

Stuffed

Frame sent

Flag | Header | 000111110110011111001000 | Trailer | Flag

Two extra bits

Frame received

Flag | Header | 000111110110011111001000 | Trailer | Flag

Unstuffed

0001111111001111101000

Data to upper layer

Flow and Error Control

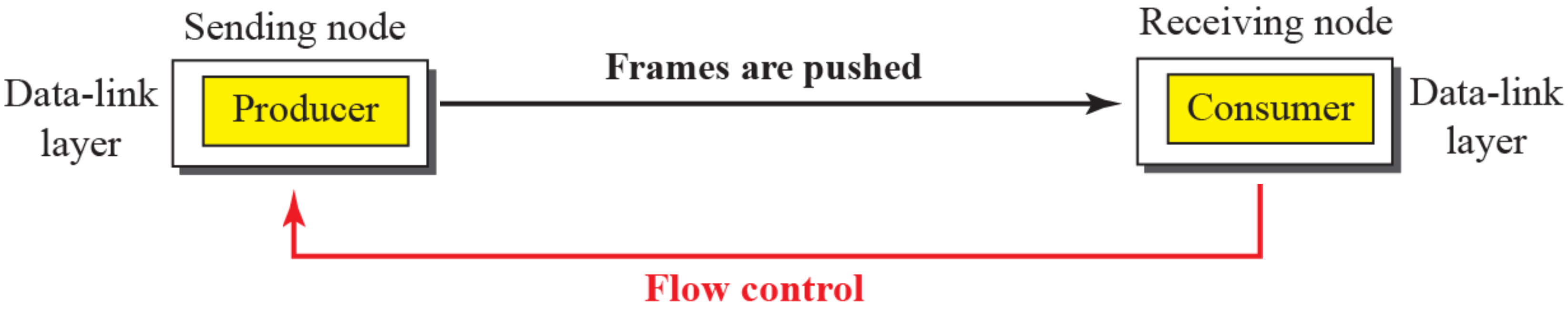
- Data Communication requires at least two devices working together, one to send and other to receive.

- For communication and exchange of information, a good coordination is must.

- Most important responsibilities of DLL are Flow Control and Error Control.

- These functions together are called data link control.

Flow control at the data link layer



Flow control

- When an entity produces items and another entity consumes them, there should be balance between them.
- If items produced faster or slower compared the rate at which it is consumed in faster/slower rate, the systems becomes less efficient.
- Flow Control is to prevent the data items at the consumer sites.
- In DC at DLL, we need 4 entities N/W & DLL at sending and receiving nodes.
- DLL at sending pushes frames towards the DLL in receiving nodes.

Flow control

- If the receiving node cannot deliver the frames in the same rate to the N/W Layer, it becomes overwhelmed with the frames.
- FC mechanism should send the feedback to the sending device to stop or slow down the transmission of frames.
- At sending and receiving nodes, buffers are used at DLL.
- A buffer is a set of memory locations that can hold packets at the sender and receiver.
- FC communication occur by sending signals from consumer to producer.
- When buffer is full, it informs sender at DLL to stop pushing frames.

Example 111.1

The above discussion requires that the consumers communicate with the producers on two occasions: when the buffer is full and when there are vacancies.

If the two parties use a buffer with only one slot, the communication can be easier.

Assume that each data-link layer uses one single memory slot to hold a frame.

When this single slot in the receiving data-link layer is empty, it sends a note to the network layer to send the next frame.

Error control (EC)

- Since Underlying physical N/w is not fully reliable, EC is required at the DLL to prevent receiving node from delivering corrupted packets to its N/W Layer.
- EC at DLL is implemented using one of the following methods

i) If frame is corrupted, it is silently discarded; if it is not corrupted, packet is delivered N/W layer.

ii) If frame is corrupted, it is silently discarded; if it is not corrupted, an acknowledgment is sent to the sender.

Combination of Error control & Flow control

i) Error Control & Flow Control can be combined.

ii) An Acknowledgment that is sent for flow control can also be used for error control to send the sender the packet has arrived uncorrupted.

iii) Lack of Acknowledgment means that there is a problem in the sent frame.

Connectionless and Connection-Oriented

- A DLC protocol can be either connectionless or connection-oriented.

Connectionless Protocol:

- Frames are sent from one node to another without any relationship between the frames, each frame is independent.
- Connectionless means there is no connection between the frames.
- Frames are not numbered and there is no sense of ordering.

Eg: Most of DLL protocols for LANs are connectionless protocols.

Connection-oriented Protocol

- A logical connection should be established between the two nodes (Setup phase).
- All frames that are related to each other are transmitted (Transfer phase)
- A Logical connection is terminated(teardown phase).
- Frames are numbered and sent in order.
- If they 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 Network Layer.

→ Rare in wired LANs, but we can see them in some PPP protocols, wireless LANs & WANs.

11-2 DATA-LINK LAYER PROTOCOLS

•Traditionally four protocols have been defined for the data-link layer to deal with flow and error control:

Simple, Stop-and-Wait, Go-Back-N, and Selective-Repeat.

Although the first two protocols still are used at the data-link layer, the last two have disappeared.

The behavior of DLL protocols can be better shown as a finite state machine(FSM)

Finite State Machines

- A FSM is thought of as a machine with a finite number of states.
- A machine is always in one of the state until an event occurs.
- Each event is associated with two reactions

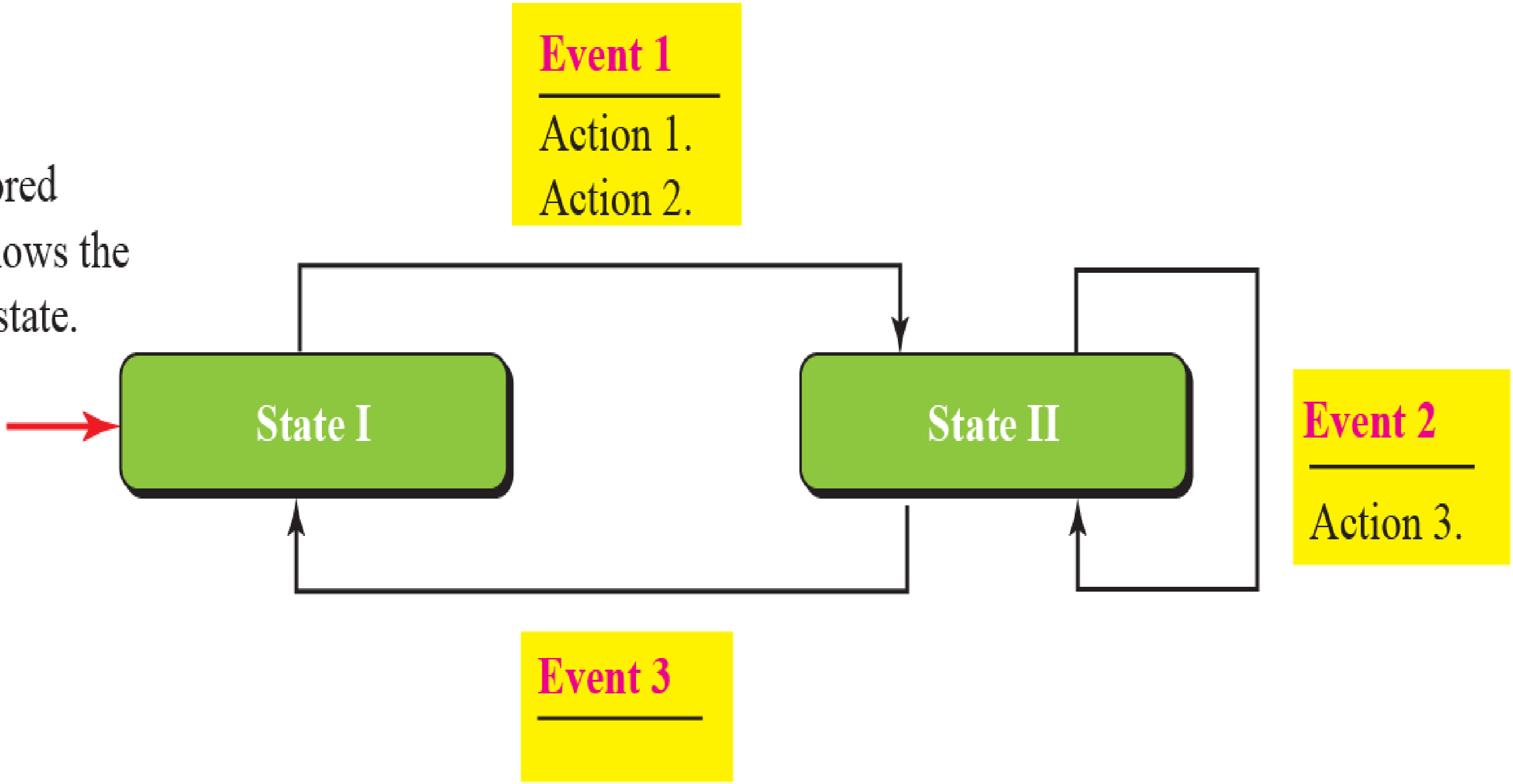
i) Defining the list of actions to be performed

ii) Determining the next state(Which can be same as current state).

- One of the state must be defined as the initial state, the state in which machine starts when turns on.

FSMs

Note:
The colored arrow shows the starting state.



State I

State II

Event 1

Action 1.
Action 2.

Event 2

Action 3.

Event 3

FSMs

- Machine with three states.
- There are only 3 possible events and 3 actions.
- Machine starts with state1.

→If event1 occurs, the machine performs action 1 & 2 & enters in to state2.

→When machine is in state2, two events may occur.

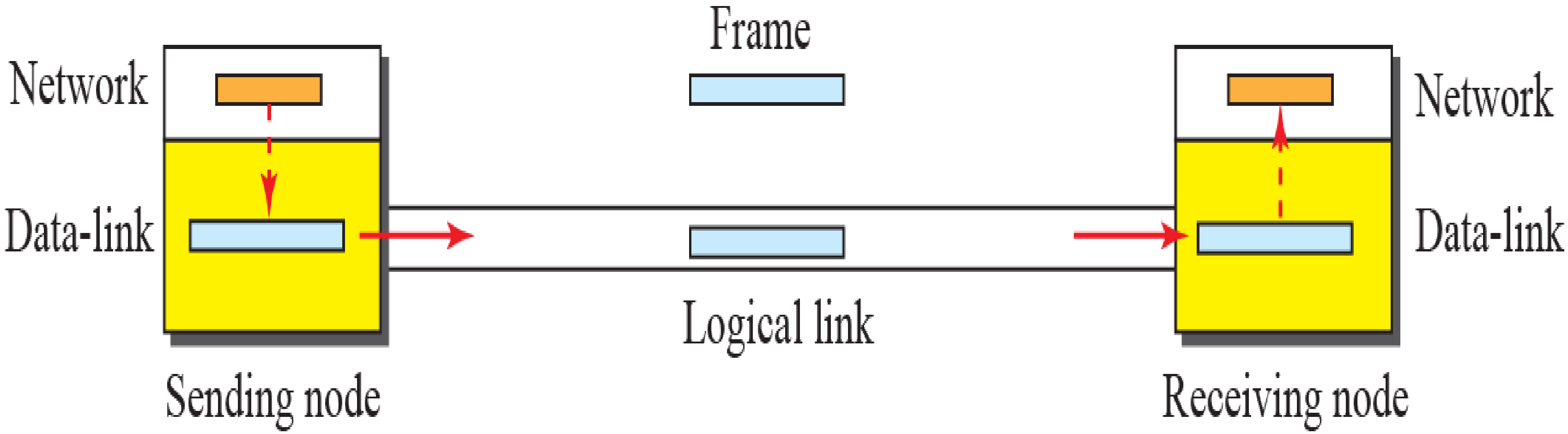
→ If event 2 occurs, the machine perform action3 and remains in the same state.

→If event 3 occurs, machine performs no action, but moves to state 1.

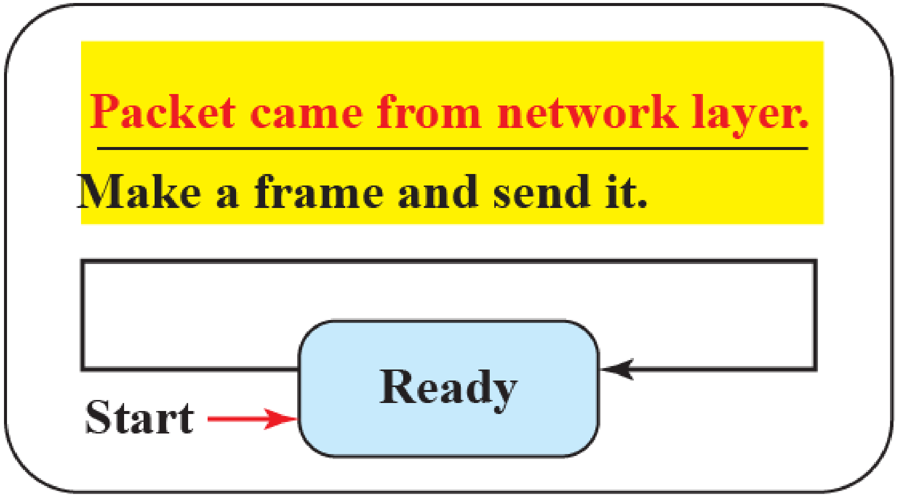
Simple Protocol

- First protocol is a simple protocol with neither flow nor error control.
- We assume that the receiver can immediately handle any frame it receives.
- In other words, the receiver can never be overwhelmed with incoming frames.
- Figure below shows the layout for this protocol.
- DLL at sender gets a packet from N/W Layer, makes a frames out of it and sends the frame.
- DLL at the receiver side, extracts the packet from the frame and delivers the packet to its N/W layer.

Simple protocol



FSM for the simple protocol



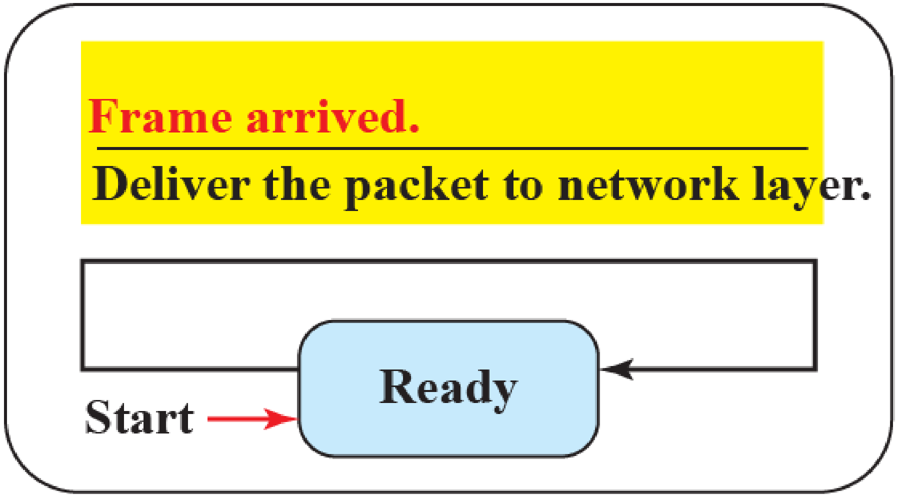
Packet came from network layer.

Make a frame and send it.

Start

Ready

Sending node



Frame arrived.

Deliver the packet to network layer.

Start

Ready

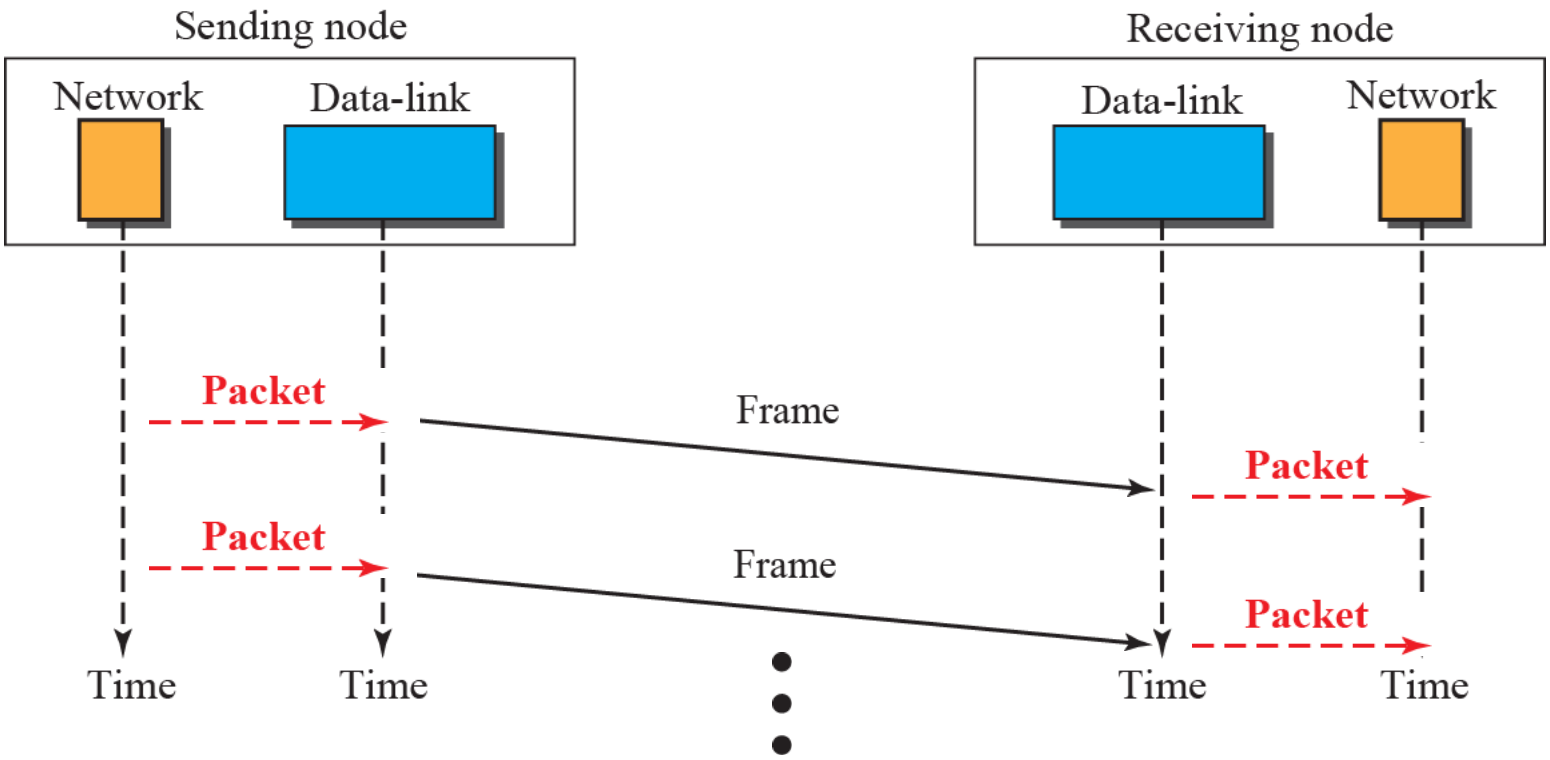
Receiving node

Example 11.2

Figure below 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.

Flow diagram for Example 11.2



Stop-and-Wait Protocol

- Second protocol is called the Stop-and-Wait protocol, which uses both flow and error control.
- In this protocol, the sender sends one frame at a time and waits for an acknowledgment before sending the next one.
- To detect corrupted frames, we need to add a CRC to each data frame.
- When a frames arrives at the receiver site, it is checked, if its CRC is incorrect, the frame is incorrect & silently discarded.
- Silence is an alert to the sender that frame is corrupted or lost.

Stop-and-Wait Protocol (Continued)

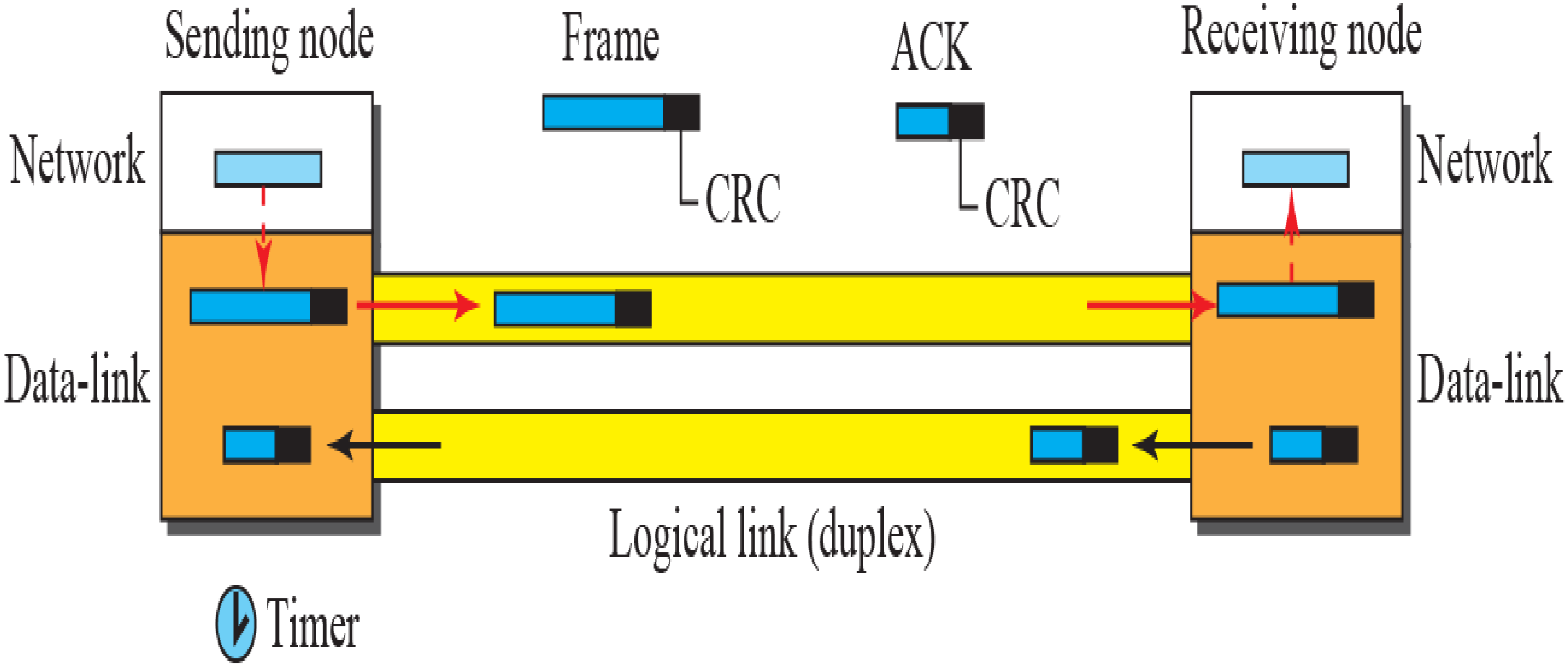
- Every time sender sends a frame, it starts a timer and if ack. arrives before timer expires, timer is stopped and next frame is transmitted.

- If timer expires before ack., frame is retransmitted because sender assumes that frame is corrupted or lost.

- This means that Sender needs to keep a copy of the sent frame until he receives the ack.

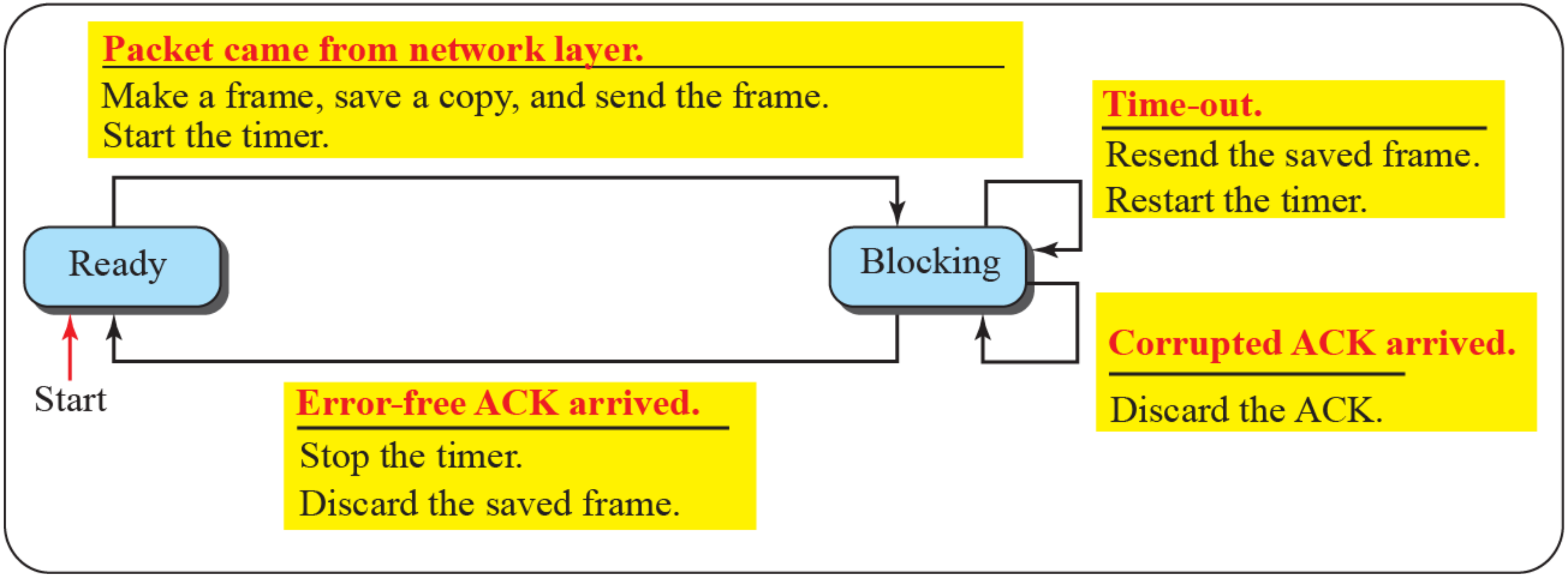
- When the corresponding acknowledgment arrives, the sender discards the frame and send the next frame if it is ready

Stop-and-wait Protocol



FSM for the stop-and-wait protocol

Sending node



Packet came from network layer.

Make a frame, save a copy, and send the frame.
Start the timer.

Time-out.

Resend the saved frame.
Restart the timer.

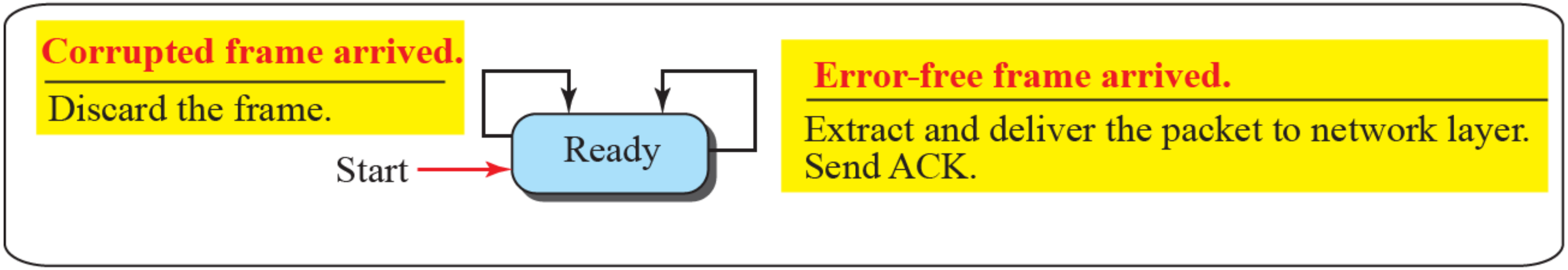
Corrupted ACK arrived.

Discard the ACK.

Error-free ACK arrived.

Stop the timer.
Discard the saved frame.

Receiving node



Corrupted frame arrived.

Discard the frame.

Error-free frame arrived.

Extract and deliver the packet to network layer.
Send ACK.

FSM for the stop-and-wait protocol

- At sender side:

The sender is initially in the ready state, but it can move between the ready state and blocking state

i) Ready state: It is waiting for a packet from the N/W layer.

→Sender creates a copy of the frame, starts the timer and sends the frame. The sender moves to the blocking state.

ii)Blocking state: 3 events occur

a) if time out occurs, sender resends the saved copy of the frame and starts 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 & moves to ready state.

FSM for the stop-and-wait protocol

At receiver side:

- ➢ The receiver is always in ready state. The two events may occur

a) If an error free frame arrives, message in the frame is delivered to the network layer and ACK is sent.

b) If a corrupted frame arrives, the frame is discarded.

Example 11.3

In the figure below 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.

Discuss how to correct this problem using sequence numbers and acknowledgment numbers.

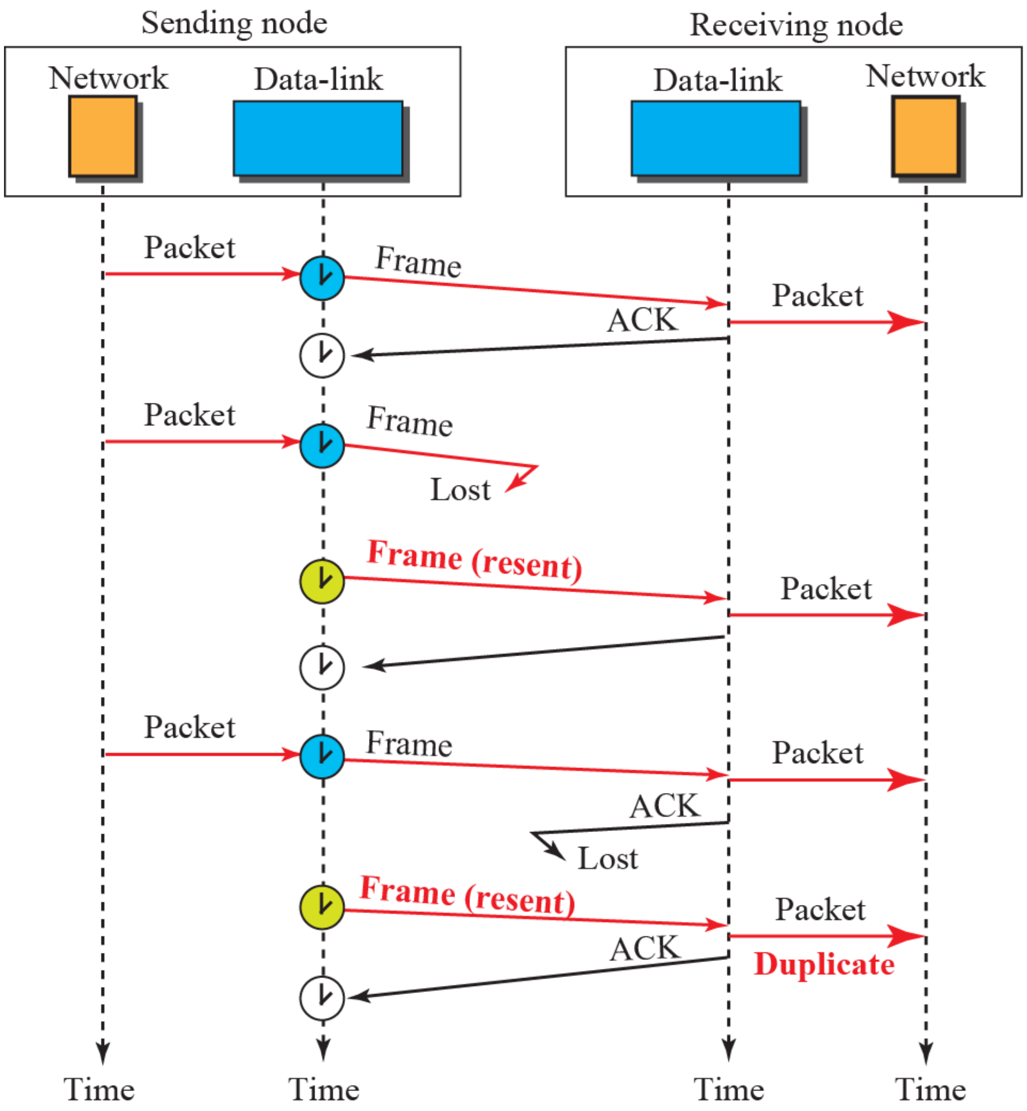
Flow diagram

Legend

◷ Start the timer.

◷ Stop the timer.

◷ Restart a time-out timer.



Notes:

A lost frame means either lost or corrupted.

A lost ACK means either lost or corrupted.

Example 11.4

Figure below 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.

Flow diagram

Legend

- ◷ Start the timer.
- ◷ Stop the timer.
- ◷ Restart a time-out timer.

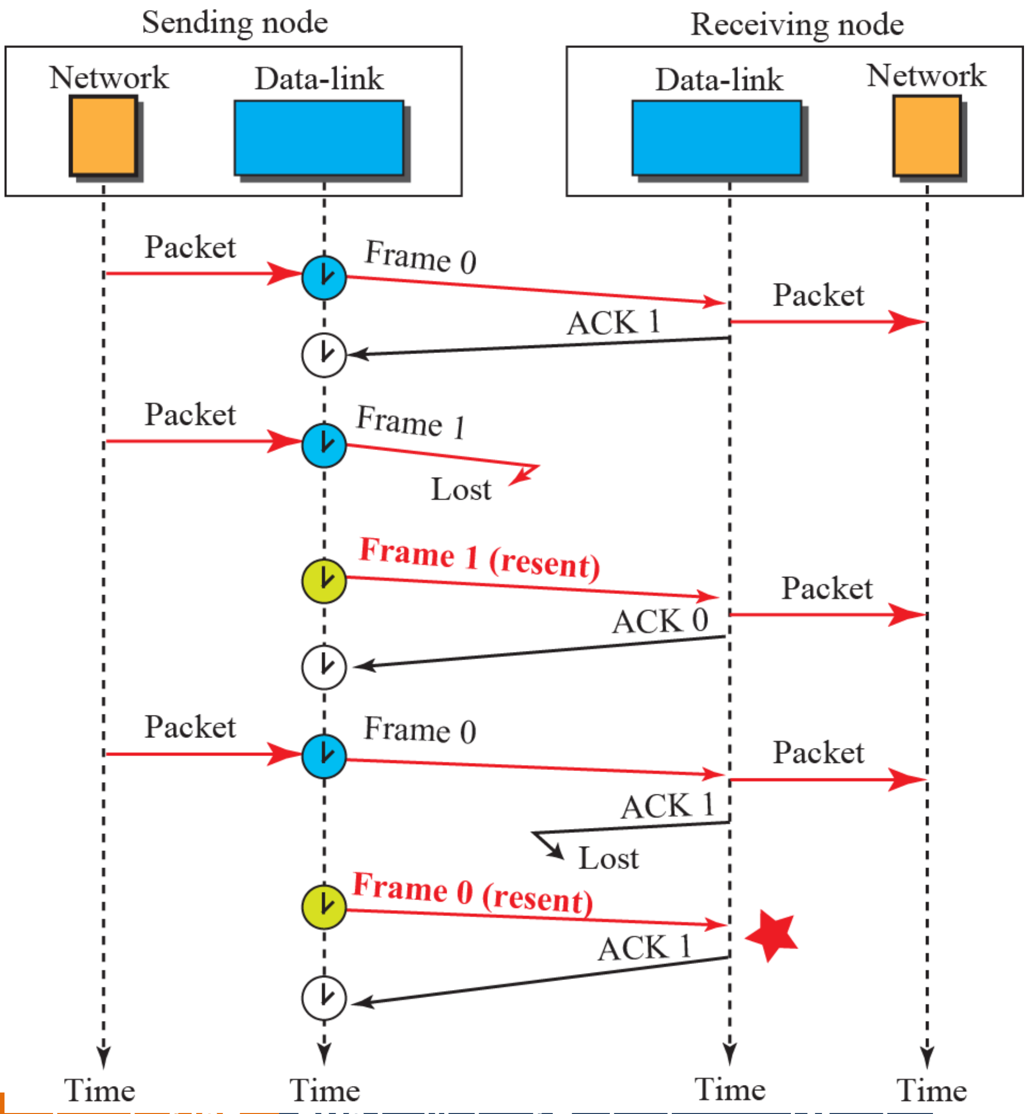
Notes:

A lost frame means either lost or corrupted.

A lost ACK means either lost or corrupted.

★

Frame 0 is discarded because the receiver expects frame 1.



Piggybacking

- The two protocols we discussed in this section are designed for unidirectional communication, in which data is flowing only in one direction although the acknowledgment may travel in the other direction.

- Protocols have been designed in the past to allow data to flow in both directions.

- However, to make the communication more efficient, the data in one direction is piggybacked with the acknowledgment in the other direction.

Point to Point Protocol (PPP)

- One of the most common protocols for point-to-point access is the Point-to-Point Protocol (PPP).

- Today, millions of Internet users who need to connect their home computers to the server of an Internet service provider use PPP.

- To control and manage the transfer of data, there is a need for a point-to-point protocol at the data-link layer. PPP is by far the most common.

Services

- The designers of PPP have included several services to make it suitable for a point-to-point protocol, but have ignored some traditional services to make it simple.

Services provided by Point-point protocol

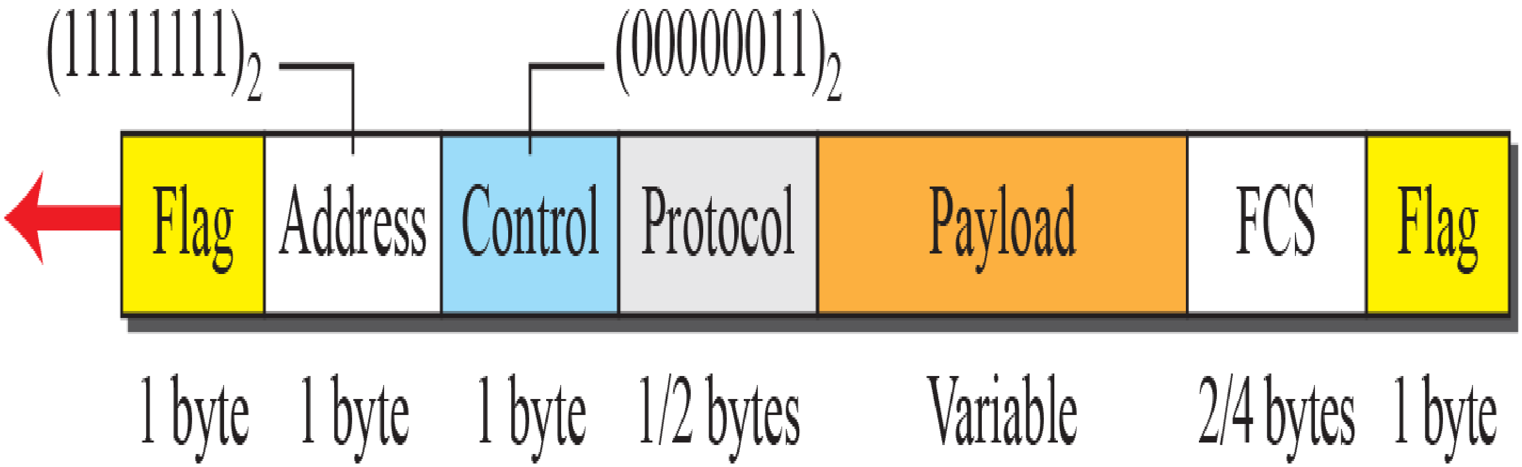
Defines

1. format of the frame to be exchanged between the devices.
2. How two devices can negotiate the establishment of the link and the exchange of data.
3. How N/W layer data are encapsulated in the data link frame.
4. How two devices can authenticate each other.
5. Provides Multiple N/W services supporting a variety of N/W Layer protocols.
6. Connection over multiple links is provided by new version of PPP.
7. N/W address configuration. Useful when a home user needs temporary N/W address to connect to Internet

Services not provided by ppp

- PPP does not provide flow control. Sender can send frames one after another w/o concerning about overwhelming receivers buffer.
- Simple mechanism for error control. A CRC field is used to detect errors. If frame is corrupted , frame is silently discarded. Upper Layer protocol need to take care of this.
- Does not provide addressing mechanism to handle frames in a multipoint configuration.

Framing



PPP frame format

Framing

- PPP uses a character-oriented (or byte-oriented) frame. Figure below shows the format of a PPP frame. PPP is a byte oriented protocol

Flag:→ A PPP frame starts and ends with a one byte flag with the bit pattern 0111110. The flag is treated as a byte.

Address:→ The address field is constant value and set to 11111111(broadcast address)

Control→ The field is set to a constant value 11000000. PPP doe not provide FC and limited EC. Two parties can negotiate and omit this byte.

Framing

Protocol→ defines what is being carried in the data field either user data or other information.
Default size is 2 bytes, two parties can negotiate to use only 1 byte.
Payload field→ carries use data or other info.

Data field is a sequence of bytes with the default of maximum of 1500 bytes, But can be changed during negotiation.
The data field is byte stuffed if the flag byte pattern appears in the field. Because there is no field defining the size of the data field, padding is needed if the size is less than the maximum default/negotiated value.

FCS:→ FCS is simply a 2/4 byte standard CRC

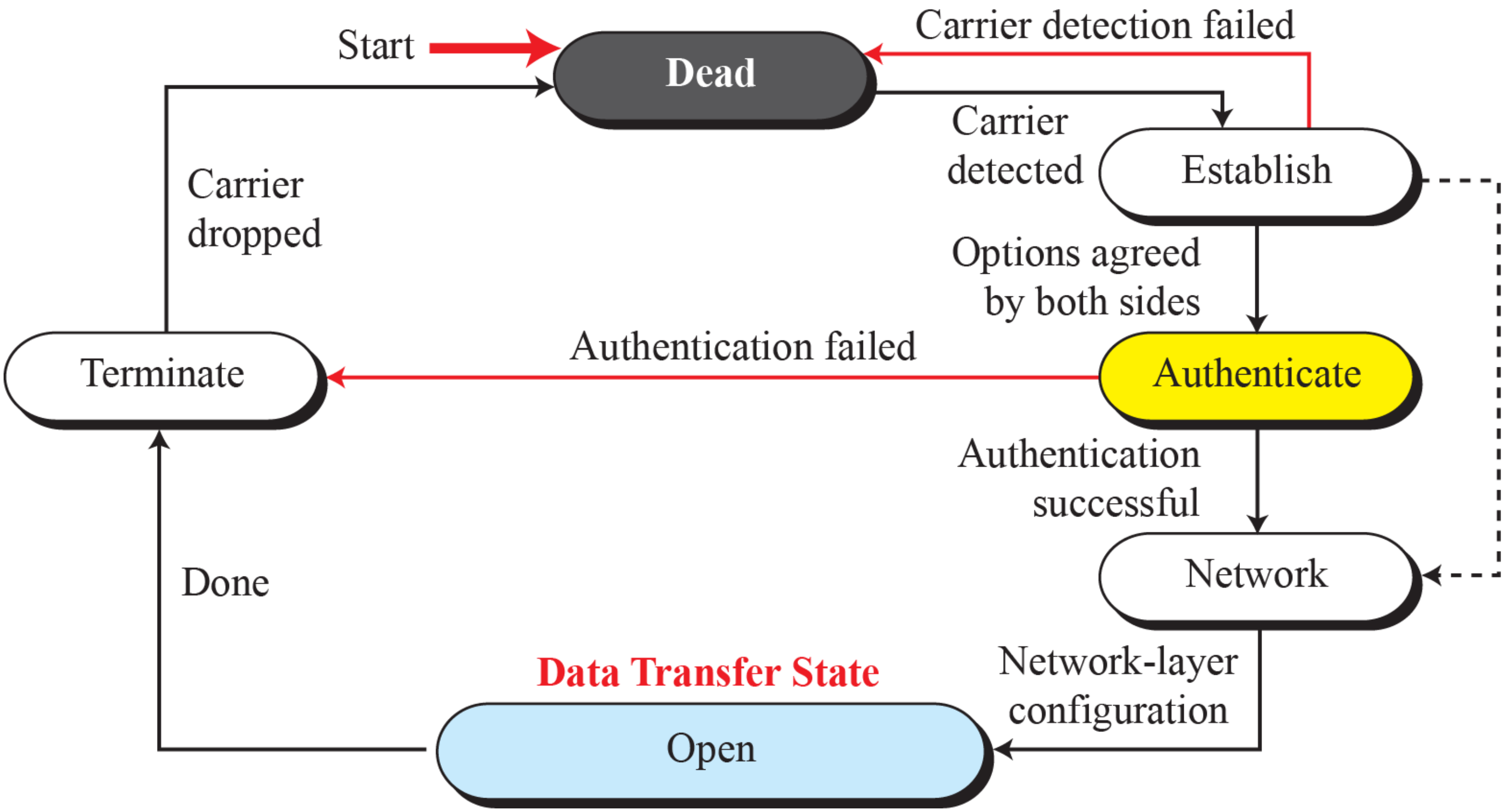
Byte stuffing

- In PPP flag is byte. The flag byte need 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 the next byte is not a flag.
- *PPP is a byte-oriented protocol using byte stuffing with the escape byte 01111101.*

Transition Phases

- A PPP connection goes through phases which can be shown in a transition phase diagram.

Transition phases



Transition phases

Dead: In this phase link is not used. There is no active carrier and line is quiet.

Establish: When one of the node starts communication, connection goes into establish phase. Options are negotiated between two parties.

If negotiation is successful, system goes to authentication phase or directly to networking phase. Link control packet are used in this phase.

Authenticate: Two nodes can negotiate and skip this phase as it is optional. If not to skip, authentication packets are used.

Transition phases

Network: Negotiation for the network layer protocols takes place. PPP specifies that two nodes establish a N/W agreement before data at the N/w can be exchanged.

PPP supports multi protocols at N/W Layer.

If a node is running multi protocol simultaneously at the N/W Layer, receiving node needs to know which protocol will receive the data.

Transition phases

Open: Data transfer takes place.

When a connection reaches this phase, exchange of data packets takes place

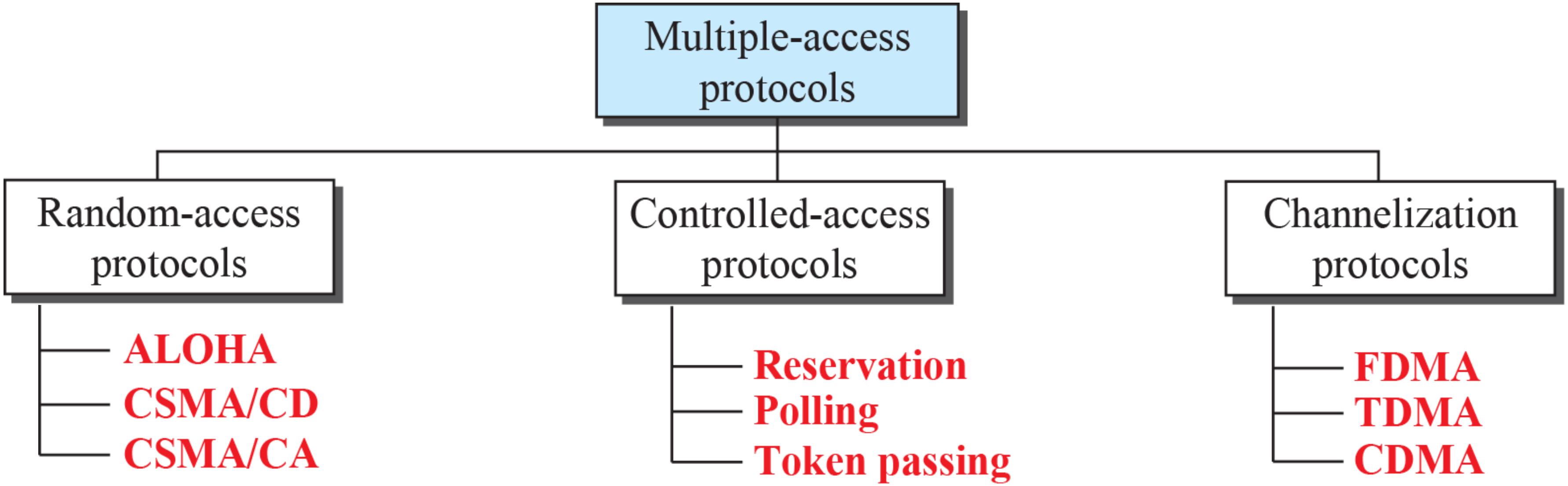
Connections remains in this phase until one of the endpoints wants to terminate the connection

Terminate: Connection is terminated. Several packets are exchanged b/w the two ends for house cleaning and closing of the link

Media Access Control(MAC)

- When stations are connected to a common link(multi point/broadcast link), multiple access protocol is required to coordinate access to the link.
- The problem of controlling the access to the medium should ensure that at the same time no two devices can interrupt or monopolize the link.
- Many protocols have been devised to handle access to a shared link.
- These protocols belong to a sub layer in DLL called media access control(MAC).

Taxonomy of multiple-access protocols



RANDOM ACCESS

- In random-access or contention no station is superior to another station and none is assigned control over another.
- At each instance, a station that has data to send uses a procedure defined by the protocol to make a decision on whether or not to send.
- This decision depends on the state of the medium (idle or busy).
- Each station can transmit when it desires on the condition that it follows the predefined procedure, testing the state of the medium.

Random access(Contention Method)

- Two features that gave this name

1) There is no scheduled time for the station to transmit. Transmission is random among the stations.

→ So, the name *Random access method*

2)No rules specify which station should send next. Stations compete with one another to access the medium.

→So, the method is called Contention methods

Random access

- As each station has right to access the medium, if more than one station tries to send, there is an access conflict and frames will be either modified or destroyed.
- To avoid this access conflict, each station follows the procedures which answers questions

1)When can the station access the medium?

2)What can station do if medium is busy.

3)How can the station determine the success or failure of the transmission

4) What can station do if there is access conflict

Random access

- Methods

1) ALOHA
2) CSMA/CD
3) CSMA/CA

ALOHA

- ALOHA, the earliest random access method, was developed at the University of Hawaii in early 1970.
- It was designed for a radio (wireless) LAN, but it can be used on any shared medium.
- It is obvious that there are potential collisions in this arrangement.
- The medium is shared between the stations. When a station sends data, another station may attempt to do so at the same time.
- The data from the two stations collide and become garbled.

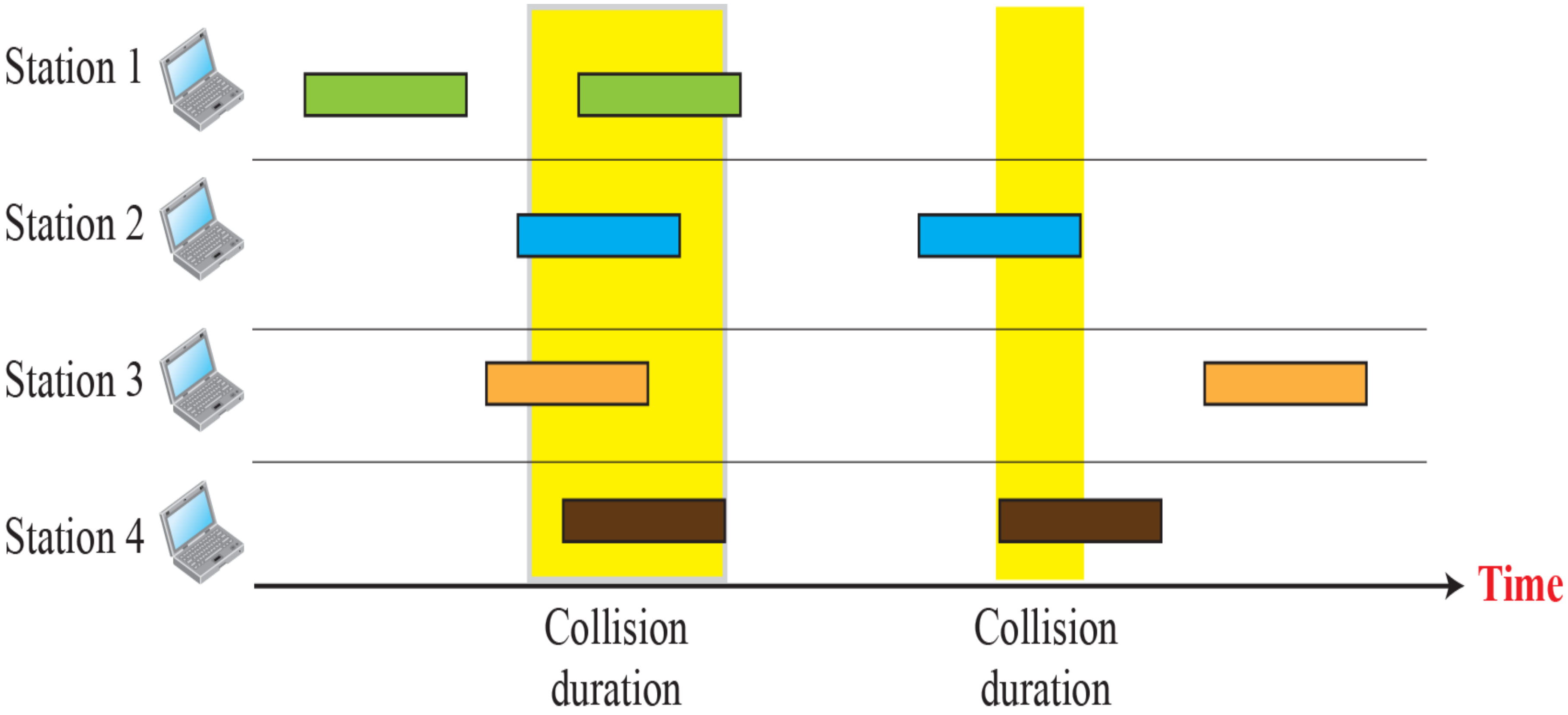
Pure ALOHA

- Original ALOHA protocol is called Pure ALOHA.

Idea is

1) Each station sends a frame whenever it has frame to send.
2) Since there is only one channel to share, there is possibility of collision between frames from different stations.

Frames in a pure ALOHA network



ALOHA

Eg: There are 4 stations that content with one another for access to the shared channel.

→Each station sends two frames, there are totally 8 frames on this medium.

→Some of these frames collide because multiple access frames are in the contention for the shared channel.

→Only two frames, frame 1.1 from station 1 and frame 3.2 from station 3 survived.

ALOHA

- Even if 1 bit of a frame coexists on the channel with one bit from another frame, there is a collision and both will be destroyed.
- Need to resend the frames that have been destroyed during transmission.
- Pure ALOHA protocol relies on Ack. from the receiver.
- If station does not receive ack. with in time-out period, station assumes that frames have been destroyed and resends it.

Pure ALOHA

- A Collision involves two or more stations. If all these stations resend frames again, frames collide.
- PURE ALOHA dictates that when the time out period passes, each station waits for random amount of time before resending the frames.
- Randomness avoid more collisions .
- This is called Back-off time T_B .

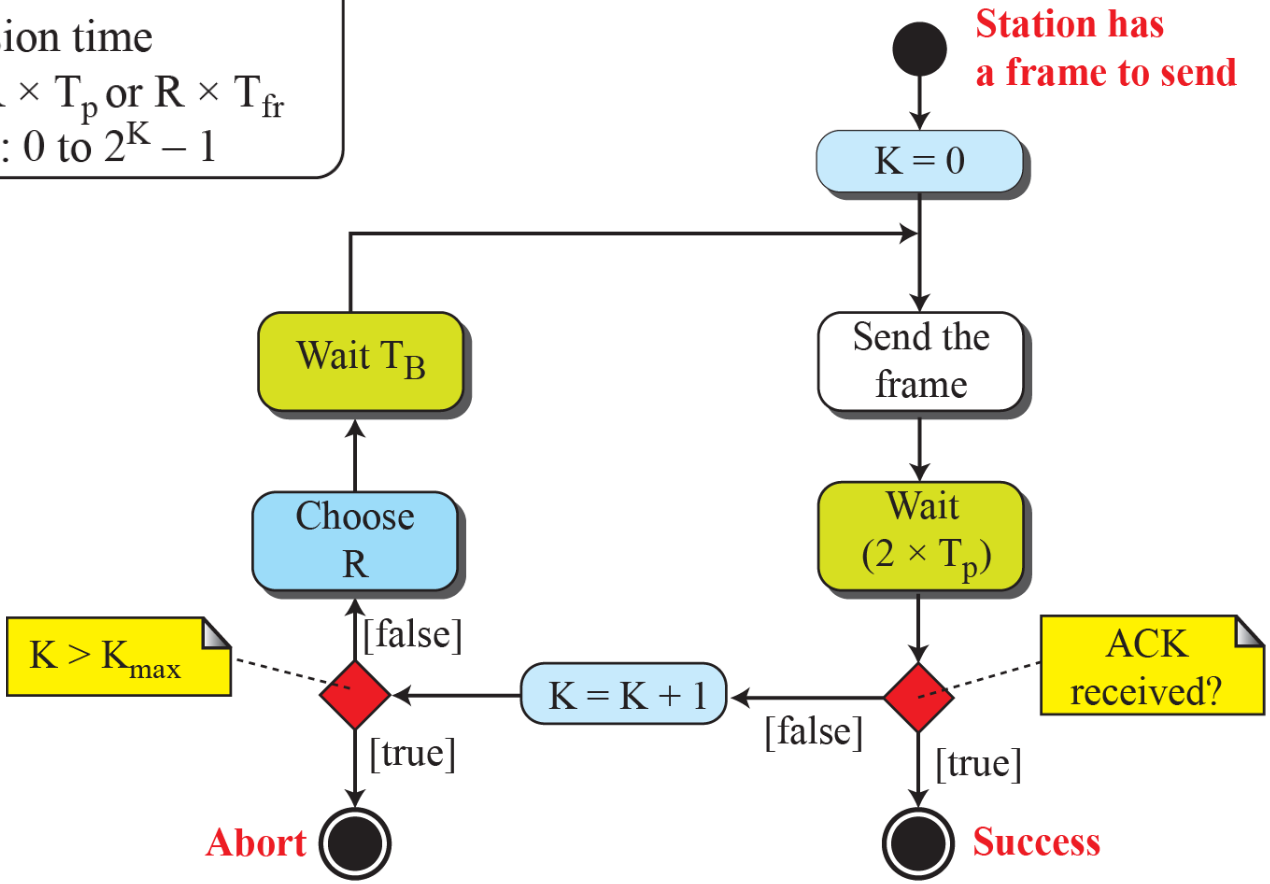
Pure ALOHA

- Second method to prevent congesting the channel with retransmitted frames is after maximum number of retransmission attempts K_{max} the stations must give up and try later.
- Time out period is equal to maximum possible RT propagation delay, which is twice the the amount of time required to send the frame between most widely separated stations(2*Tp).
- Back-off time T_B is a random value which depends on value of K.

Procedure for pure ALOHA protocol

Legend

K : Number of attempts
T_p : Maximum propagation time
T_{fr}: Average transmission time
T_B: (Back-off time): $R \times T_p$ or $R \times T_{fr}$
R : (Random number): 0 to $2^K - 1$



Example 12. 1

The stations on a wireless ALOHA network are a maximum of 600 km apart.

If we assume that signals propagate at 3×10^8 m/s, find Tp and T_B .

Solution:

To find $T_p = (600 \times 10^3) / (3 \times 10^8) = 2$ ms.

for K = 2, the range of R is {0, 1, 2, 3}.

This means that T_B can be 0, 2, 4, or 6 ms, based on the outcome of the random variable R.

Continued…

- Common formula is binary exponential back-off.

→ In this method, for each retransmission, a multiplier in the range 0 to 2^K-1 is randomly chosen and multiplied by T_p (Max. propagation time) or T_{fr}(Average time required to send out a frame) to find T_B

→The range of random numbers increases after each collision.

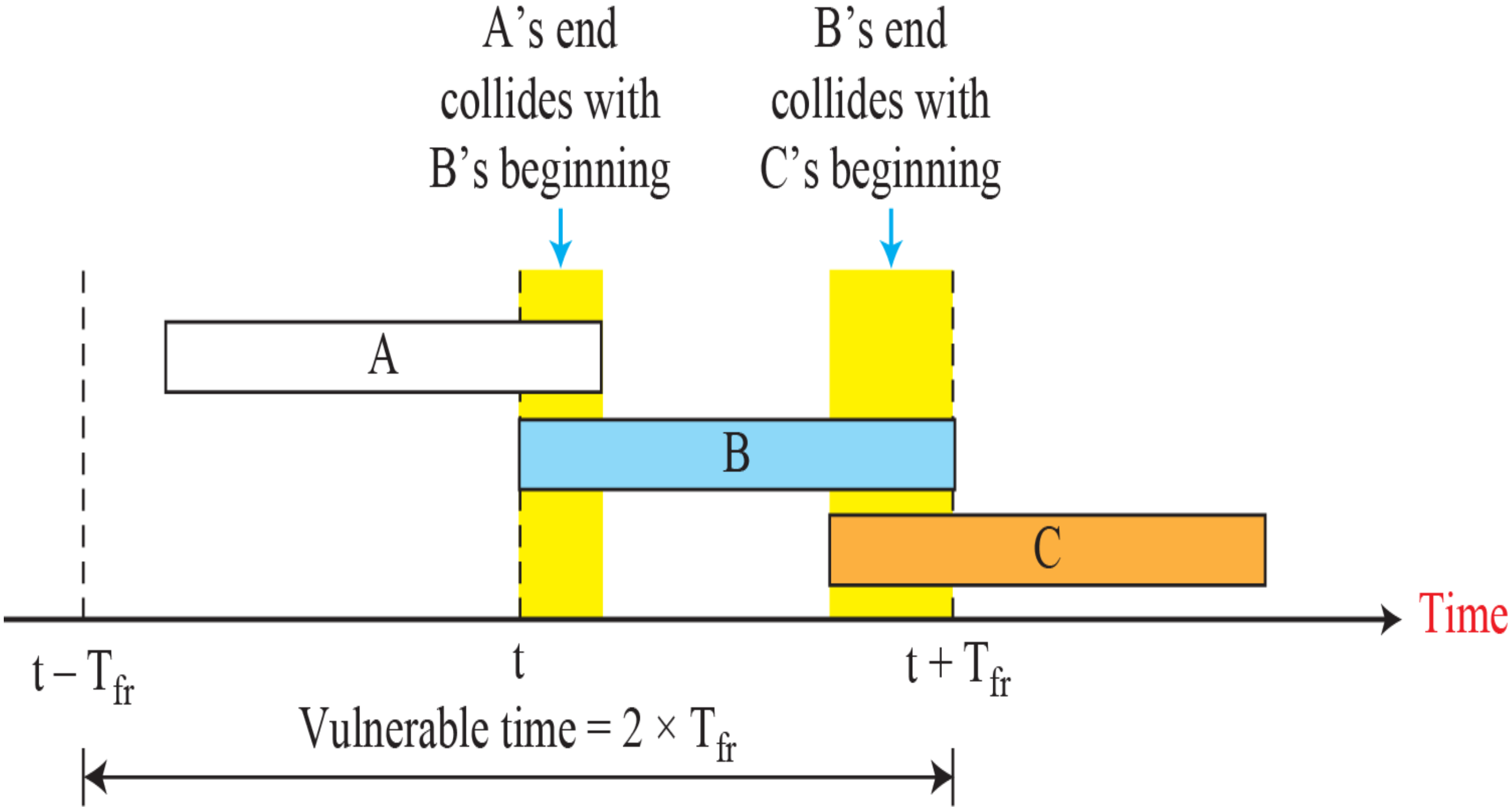
→Usually the value of Kmax chosen is 15

Vulnerable time

- Vulnerable time is length of a time, in which there is a possibility of collision.
- Assume that stations send fixed frames with each frame taking T_{fr} seconds to send .
- From the fig. the vulnerable time ,during which a collision may occur in PURE ALOHA, is 2 times the frame transmission time.

Pure Aloha Vulnerable time=2* T_{fr}

Vulnerable time for pure ALOHA protocol


A's end collides with B's beginning
B's end collides with C's beginning
A
B
C
Time
$t - T_{fr}$
t
$t + T_{fr}$
Vulnerable time = $2 \times T_{fr}$

Example 12.2

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.

Throughput

- Let G be the average number of frames generated by the system during one frame transmission time.
- It can be proved that, average number of successful transmission for pure ALOHA is

$$S = G * e^{-2G}$$

- The Maximum through put S_{max} =0.184 for G=1/2.

Throughput

- In other words, if one-half frame is generated during one frame transmission time, then 18.4% of these frames reach their destination successfully.
- *The throughput for pure ALOHA is*

$$S = G \times e^{-2G}.$$

- *The maximum throughput*

$$S_{max} = 0.184 \text{ when } G = (1/2).$$

Example 12. 3

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, then 1 frame per millisecond. Therefore G = 1.

Throughput $S = G \times e^{-2G} = 0.135$ (13.5 percent).

This means that the throughput is 1000 × 0.135 = 135 frames. Only 135 frames out of 1000 will probably survive.

Example 12. 3 (continued)

b. If the system creates 500 frames per second, or 1/2 frames per millisecond.

Therefore $G = 1/2$.

Throughput $S=G\times e^{-2G} =0.184$ (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.

c. If the system creates 250 frames per second, or 1/4 frames per millisecond.

Therefore $G = 1/4$.

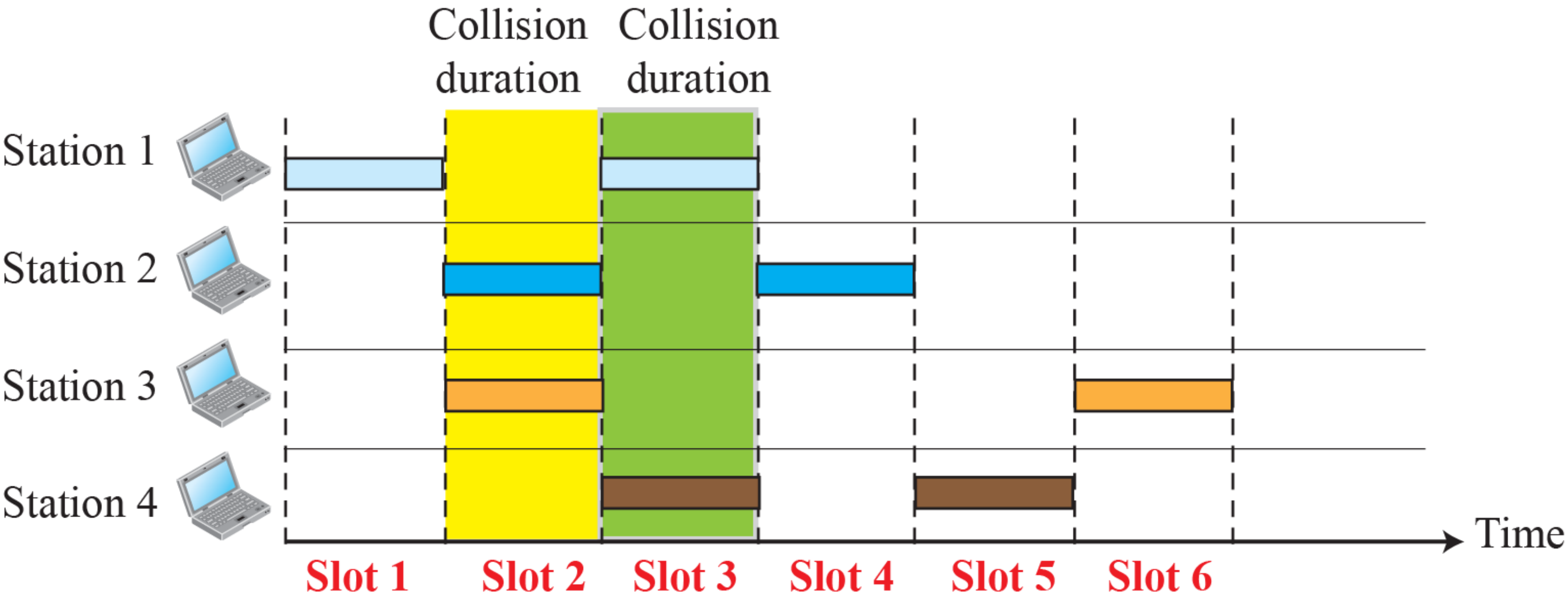
Throughput $S = G \times e^{-2G} = 0.152$ (15.2 percent).

This means that the throughput is $250 \times 0.152 = 38$. Only 38 frames out of 250 will probably survive

Slotted ALOHA

- Pure Aloha vulnerable time is 2*Tfr, because no rule that defines when the station is allowed to transmit.
- Station may send before other station has finished/another station has started.
- To improve the efficiency of Pure ALOHA, Slotted ALOHA was invented.
- In Slotted ALOHA, divide the time into slots of Tfr Sec and force the station to send only in the beginning of the time slot.

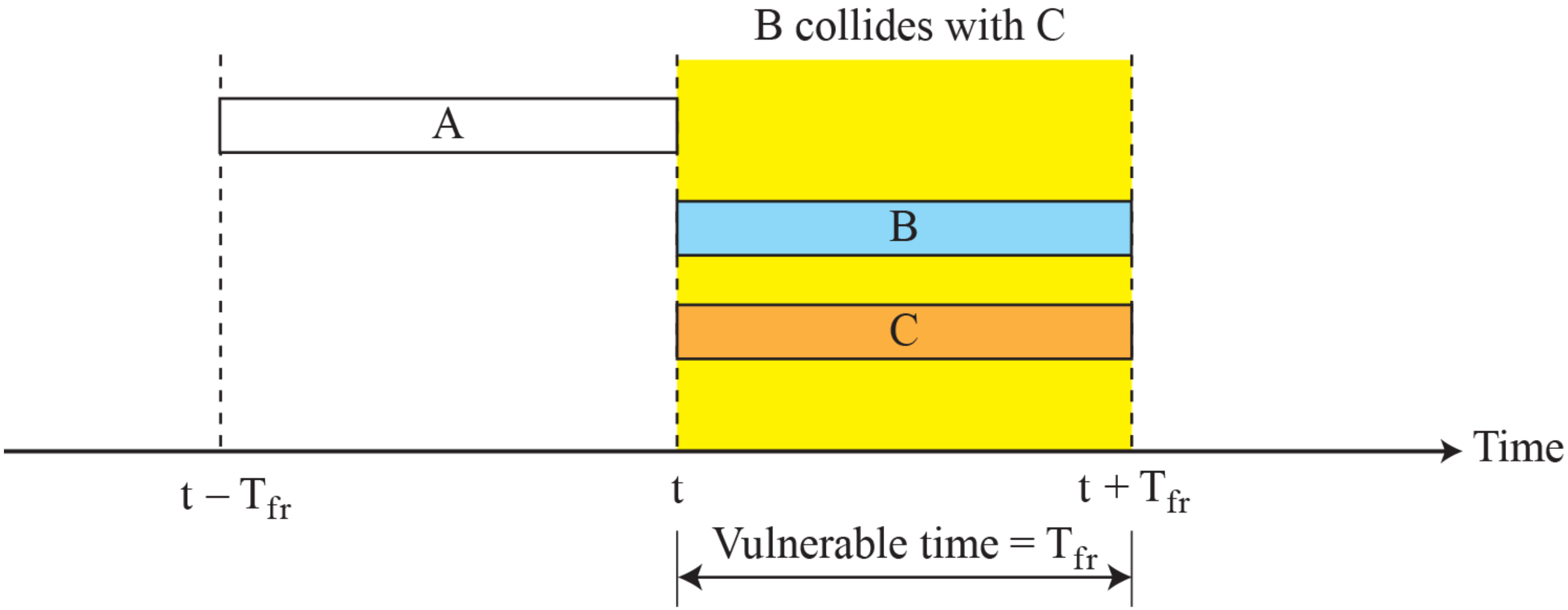
Frames in a slotted ALOHA network



Slotted ALOHA

- Because station is allowed to send only at the beginning of the synchronized time slot,
- if a station misses this moment, it must wait until the beginning of the next time slot.
- Still there is possibility of collision if two stations try to send at the beginning of the same time slot.
- Vulnerable time is one-half that of PURE ALOHA= T_{fr}.

Vulnerable time for slotted ALOHA protocol



Throughput

- The throughput for slotted ALOHA is

$$S = G \times e^{-G} .$$

- The maximum throughput

$$S_{max} = 0.368 \text{ when } G = 1.$$

- If one frame is generated during one frame transmission time, then 36.8% of these frames reach the destination successfully.
- Because vulnerable time is equal to frame transmission time., station generates only one frame in this time.

Example 12. 4

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

The frame transmission time is 200/200 kbps or 1 ms.

a. G =1. So throughput S = G × e^{-G} = 0.368 (36.8 percent).

This means that the throughput is 1000 × 0.368 = 368 frames. Only 368 out of 1000 frames will probably survive.

Example 12. 4 (continued)

b) G = 1/2.

In this case $S = G \times e^{-G} = 0.303$ (30.3 percent).

This means that the throughput $500 \times 0.303 = 151.5$. Only 151 frames out of 500 will probably survive.

c) G = 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.

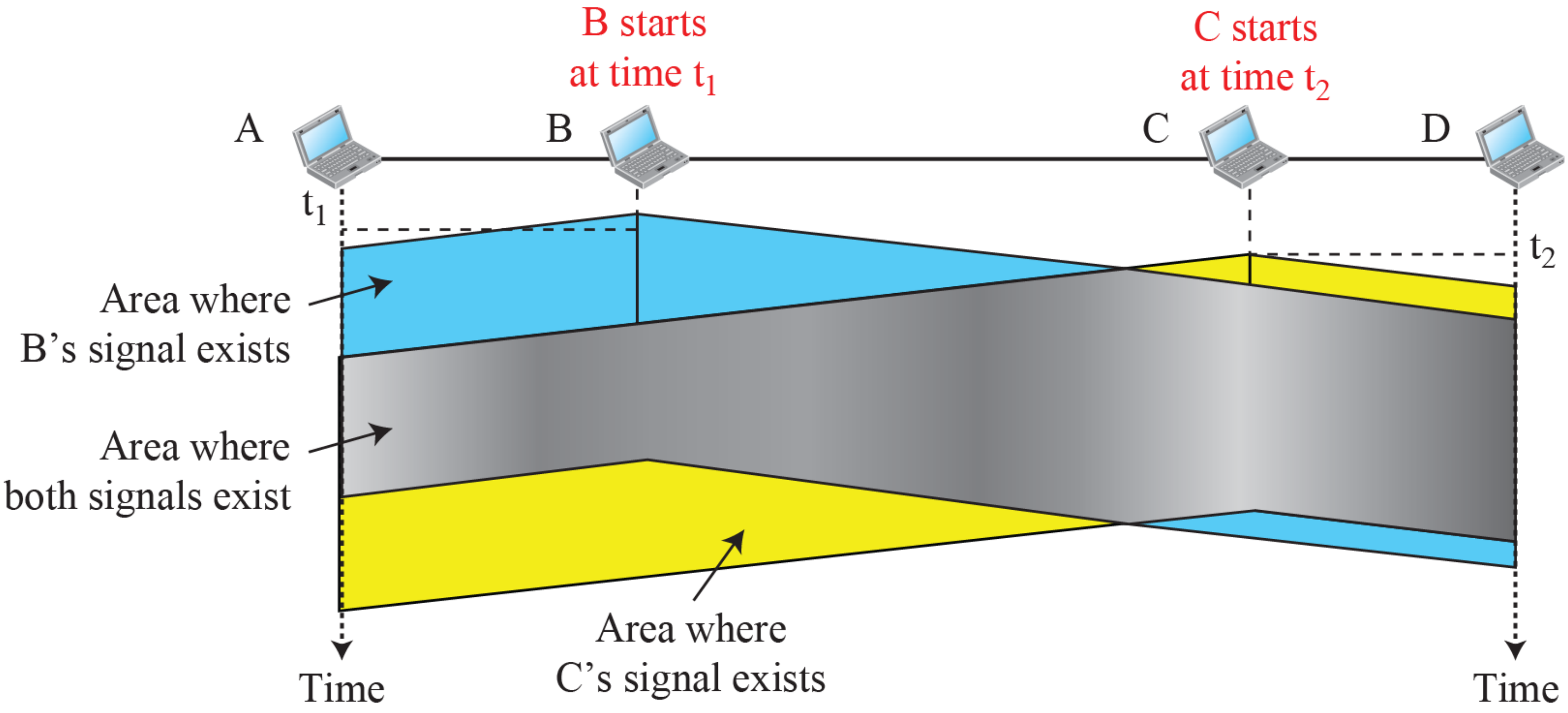
CSMA

- To minimize the chance of collision and, therefore, increase the performance, the CSMA method was developed.

- The chance of collision can be reduced if a station senses the medium before trying to use it.
- Carrier sense multiple access (CSMA) requires that each station first listen to the medium (or check the state of the medium) before sending.

- In other words, CSMA is based on the principle "sense before transmit" or "listen before talk."

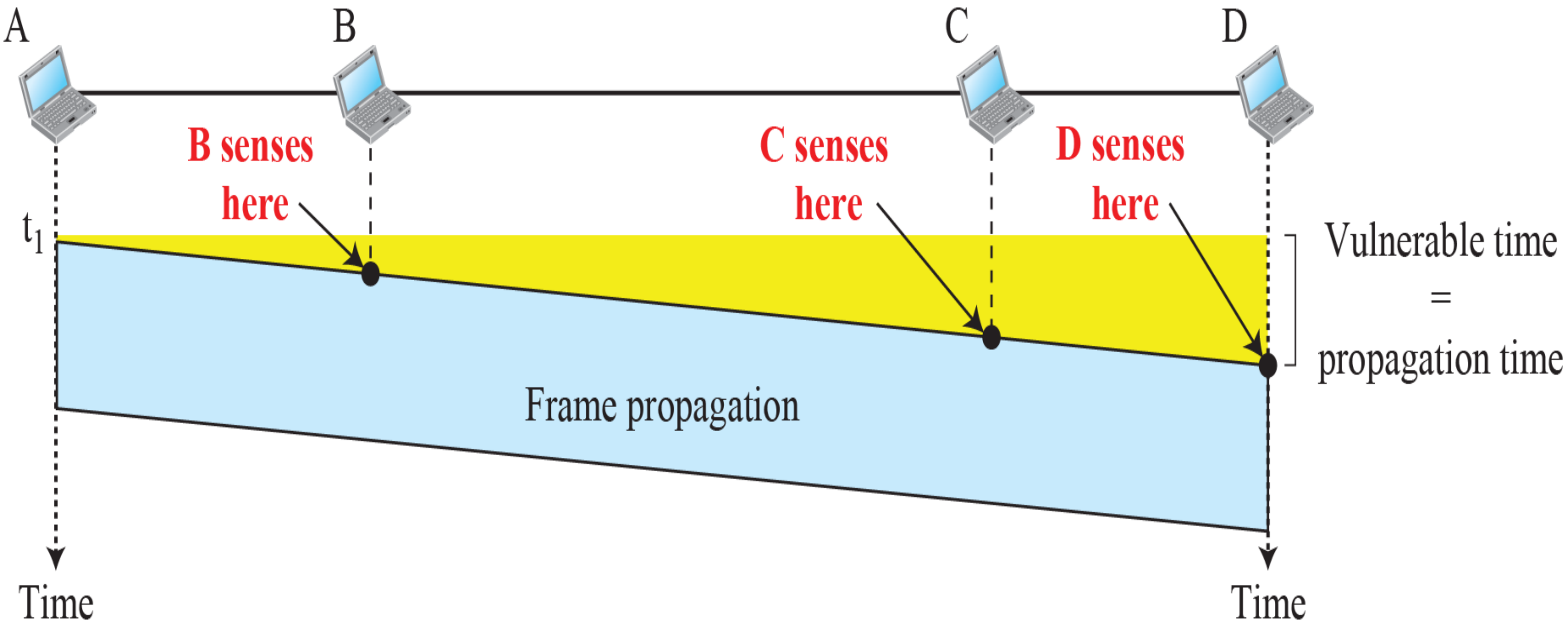
Space/time model of a collision in CSMA


B starts at time t_1
C starts at time t_2
A
B
C
D
t_1
t_2
Area where B's signal exists
Area where both signals exist
Area where C's signal exists
Time
Time

Vulnerable time

- Vulnerable time for CSMA is the Propagation time Tp.
- Tp is the time needed for the signal to propagate from one end of the medium to other.
- When any station sends a frame during this time ,collision will occur.
- If the first bit of the frame reaches the end of the medium, every station will already have heard the bit and will refrain from sending.

Vulnerable time in CSMA



Persistence Methods

What should station do if channel is busy/idle?

- Three methods to answer this

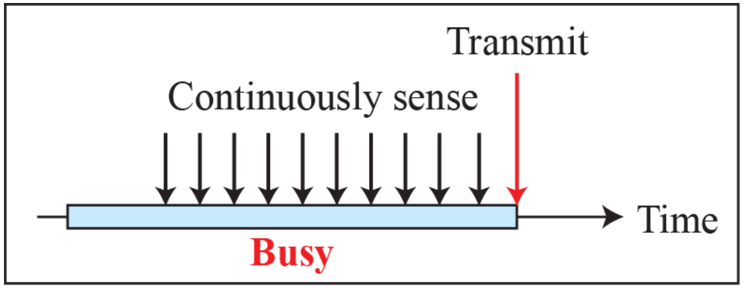
  1) 1-Persistent method

  2) Non persistent method.

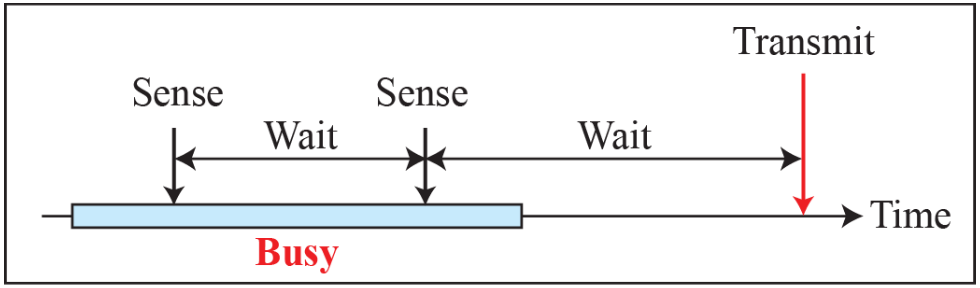
  3) p-Persistent method

- When a station finds channel is busy, behavior of three persistence methods are shown below

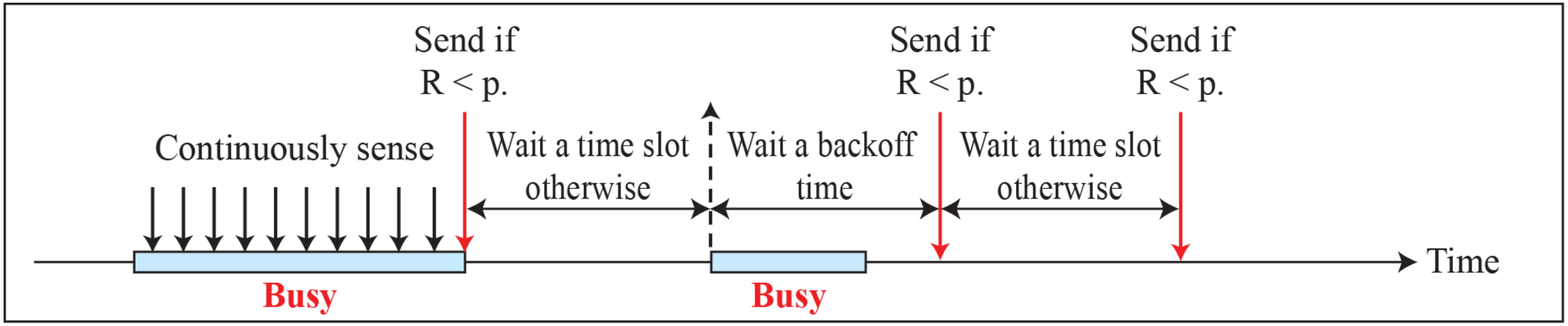
Behavior of three persistence methods



a. 1-persistent



b. Nonpersistent



c. *p*-persistent

Behavior of three persistence methods

1-Persistent : After station finds the line idle, it sends the frame immediately.(Probability 1)

→This methods has highest chance of collision because two or more stations may find the line idle and send their frames immediately.

Non-Persistent: Station that has frame senses the line. If line is idle, it sends immediately.

→If not, it waits for random amount of time and then senses the line again.

Continued..

P-persistent: This method is used if the channel has time slot with the slot duration equal to or greater than maximum propagation time.

→It reduces the chance of collision and improves the efficiency.

→ After station finds channel is idle,

it follows the following steps

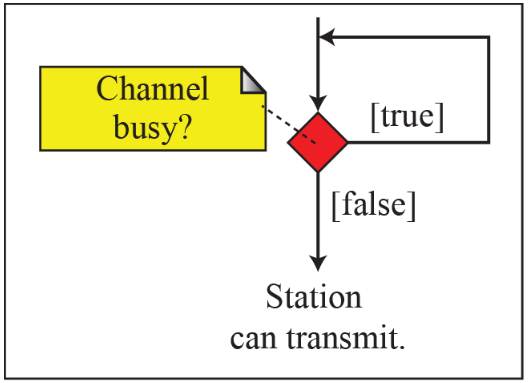
1) With probability p, the station sends its frame.
2) With probability q=1-p, the station waits for the beginning of the next time slot and checks the line again.
3) If the line is busy, it acts as though a collision has occurred and uses the back off procedure.

Continued..

→The Non-Persistent approach reduces the chance of collision because it is unlikely that two or more stations will wait for same amount of time.

→ This method reduces efficiency of the n/w because the medium remains idle when there may be stations with frames to send.

Flow diagram for three persistence methods



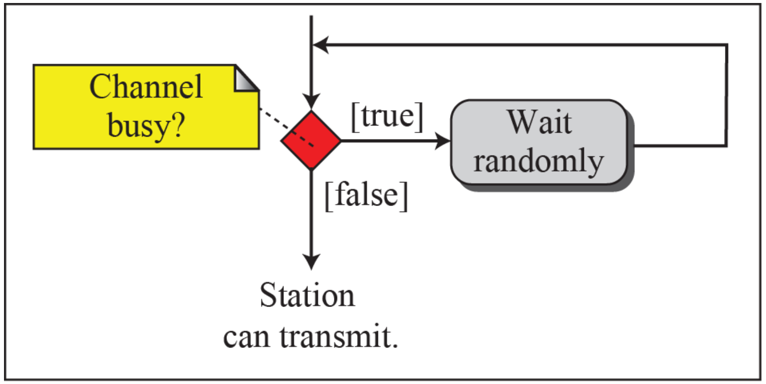
Channel busy?

[true]

[false]

Station can transmit.

a. 1-persistent



Channel busy?

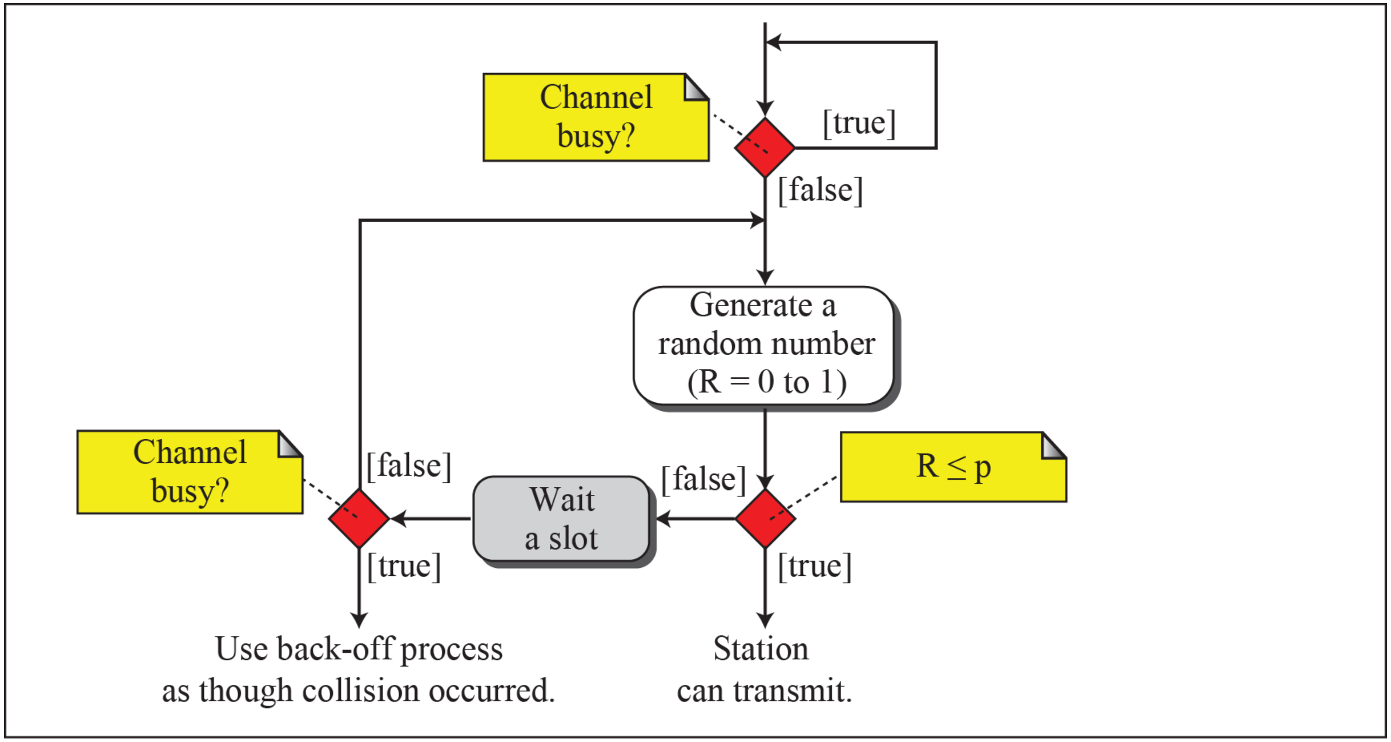
[true]

Wait randomly

[false]

Station can transmit.

b. Nonpersistent



Channel busy?

[true]

[false]

Generate a random number (R = 0 to 1)

Channel busy?

[false]

Wait a slot

[false]

$R \leq p$

[true]

[true]

Use back-off process as though collision occurred.

Station can transmit.

c. *p*-persistent

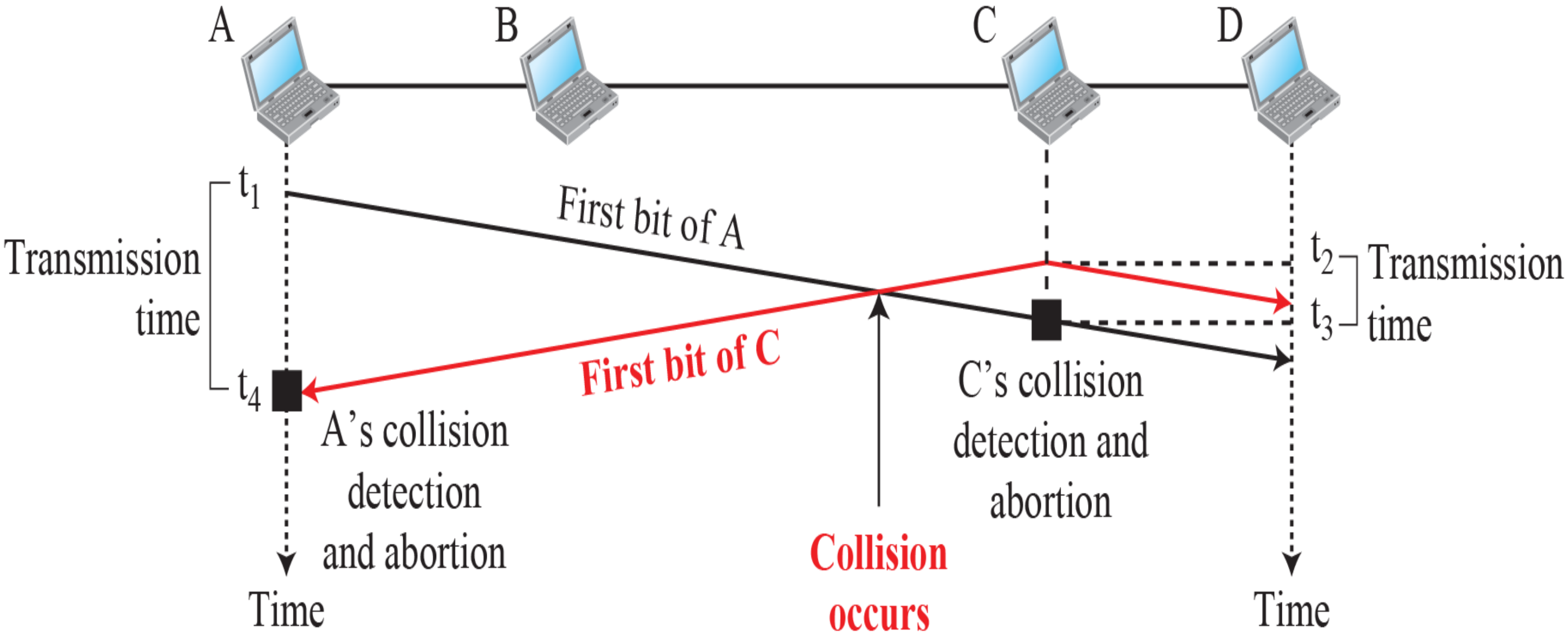
CSMA/CD

- The CSMA method does not specify the procedure following a collision.

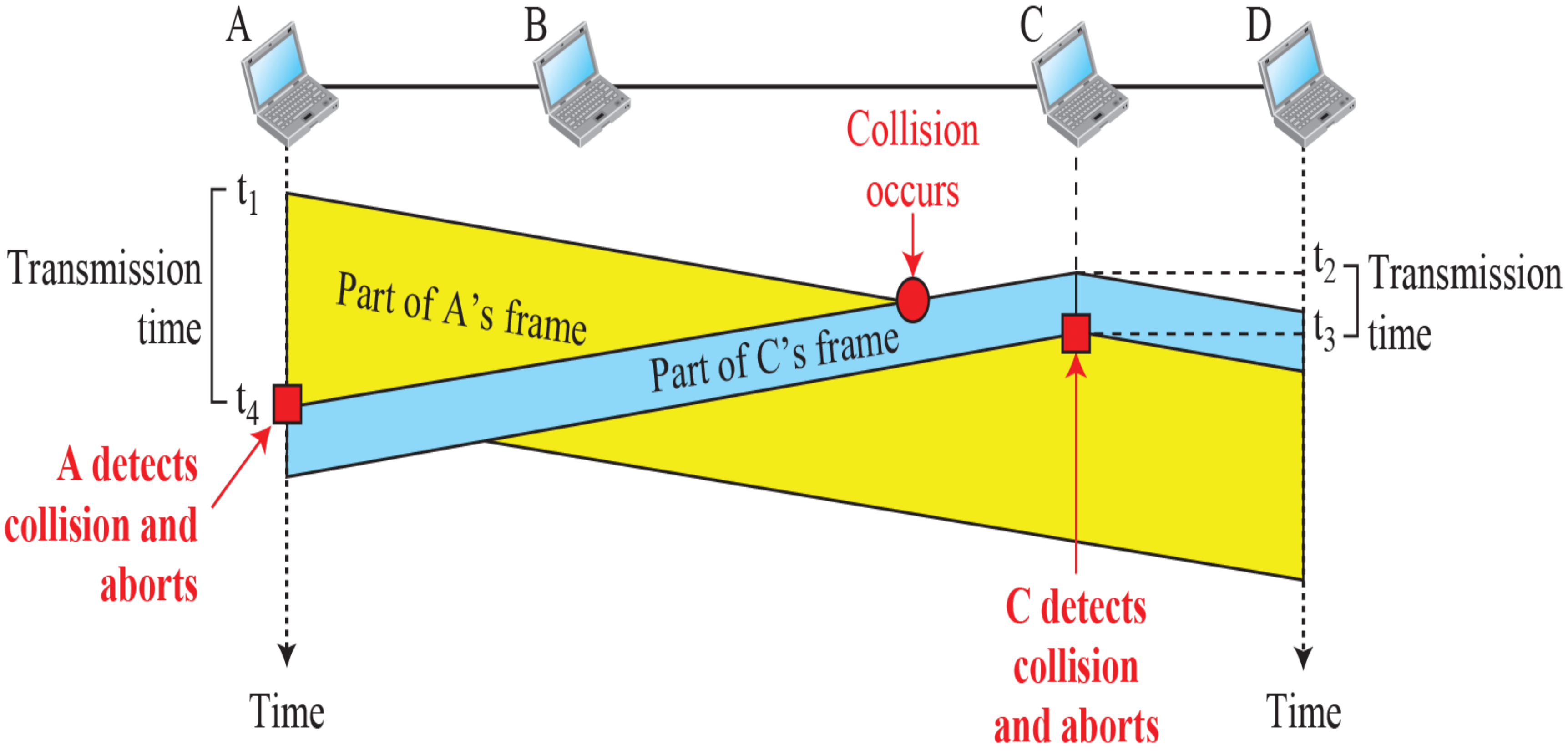
- Carrier sense multiple access with collision detection (CSMA/CD) augments the algorithm to handle the collision.

➢In this method, a station monitors the medium after it sends a frame to see if the transmission was successful. If so, the station is finished. If, however, there is a collision, the frame is sent again.

Collision of the first bits in CSMA/CD



Collision and abortion in CSMA/CD



Minimum Frame size

- CSMA/CD to work, need a restriction on the frame size.
- Before sending the last bit of the frame, if sending station detects a collision, if any, abort the transmission.
- Because station once sends entire frame, does not keep a copy of the frame and monitor the line for collision detection.
- There fore frame transmission time Tfr must be at least two times the maximum propagation time Tp(2Tp).

Example 12. 5

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

$$T_{fr} = 2 \times T_p = 51.2\ \mu s.$$

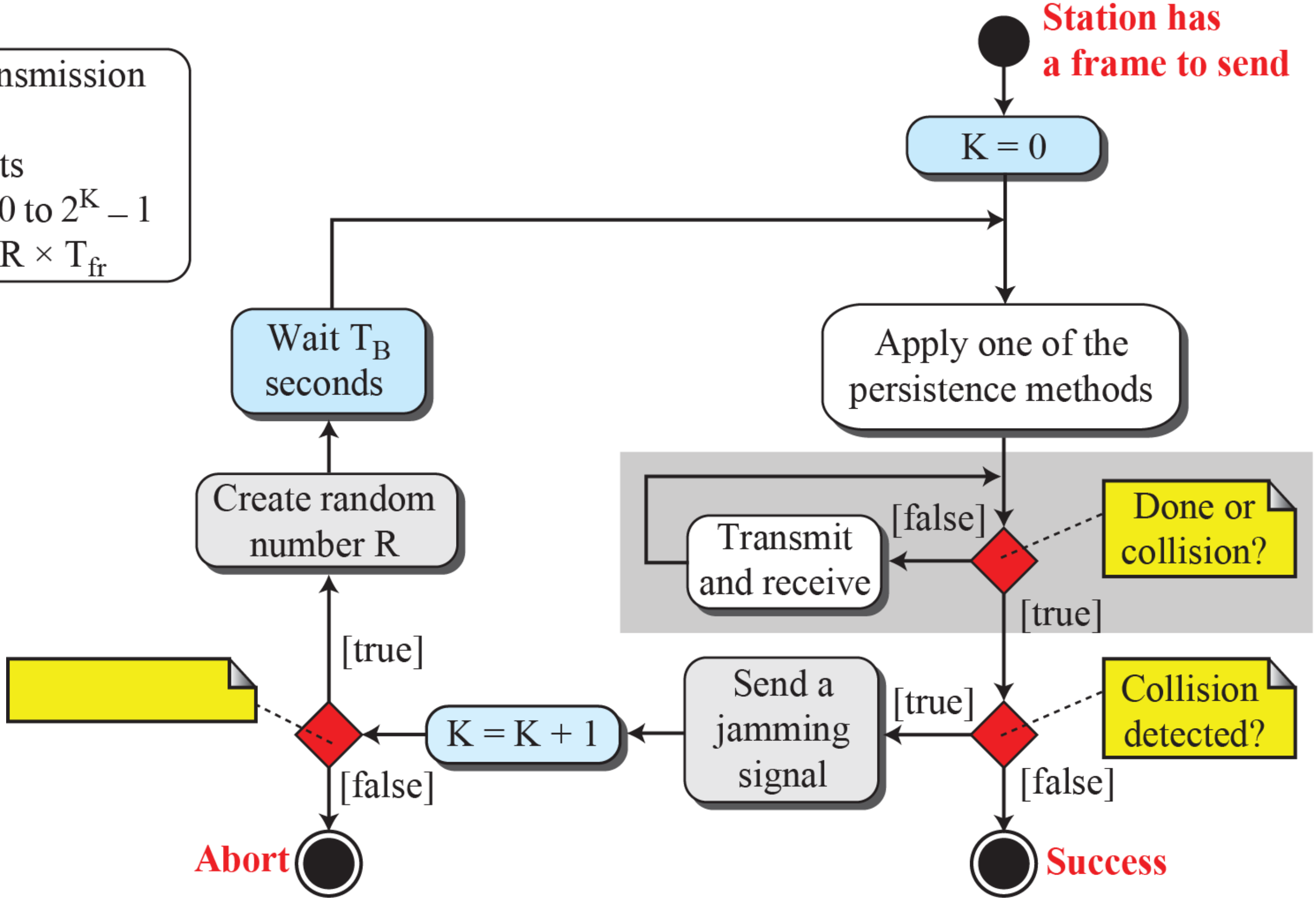
This means, in the worst case, a station needs to transmit for a period of 51.2 μs to detect the collision.

- The minimum size of the frame is 10 Mbps × 51.2 μs = 512 bits or 64 bytes Which is actually the minimum size of the frame for Standard Ethernet.

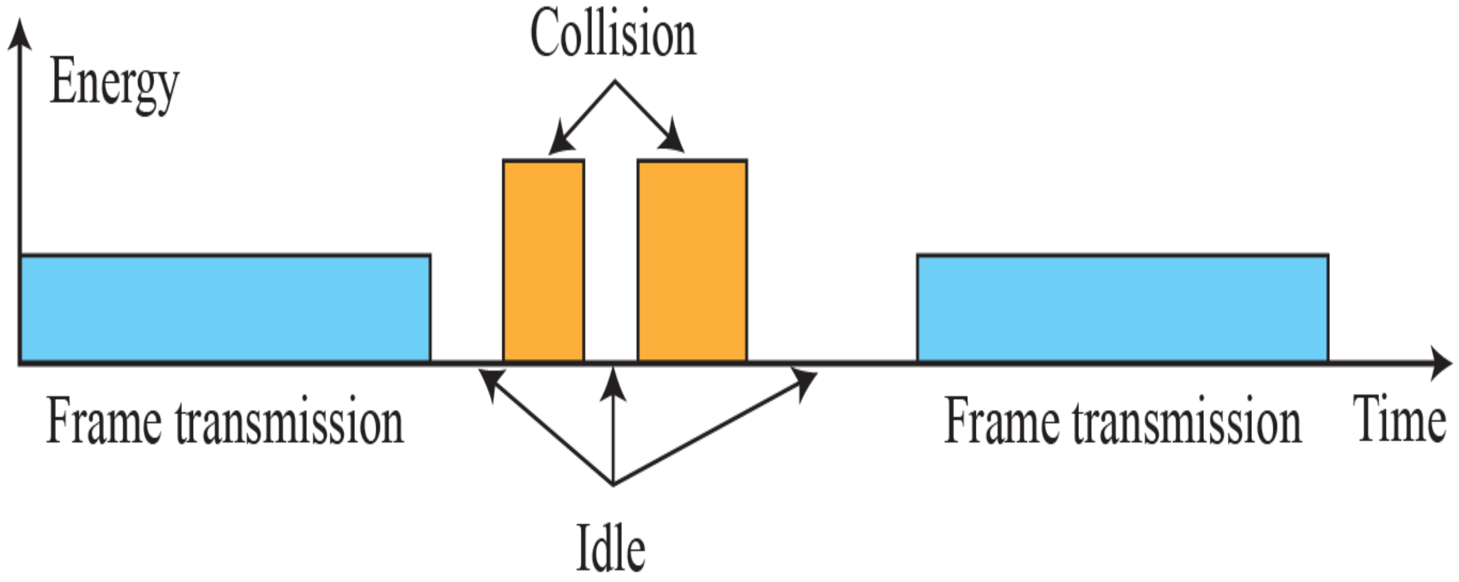
Flow diagram for the CSMA/CD

Legend

T_{fr}: Frame average transmission time
K : Number of attempts
R : (random number): 0 to $2^K - 1$
T_B: (Back-off time) = R × T_{fr}



Energy level during transmission, idleness, or collision



Energy level

- Level of energy in a channel can have three values zero, normal and abnormal.

→At Zero level, channel is idle.

→At Normal Level, station has successfully captured the channel and is sending its frame.

→At Abnormal level, there is a collision and level of energy is twice the normal level.

→The station that has frame to send, needs to monitor the energy level of the channel.

Throughput

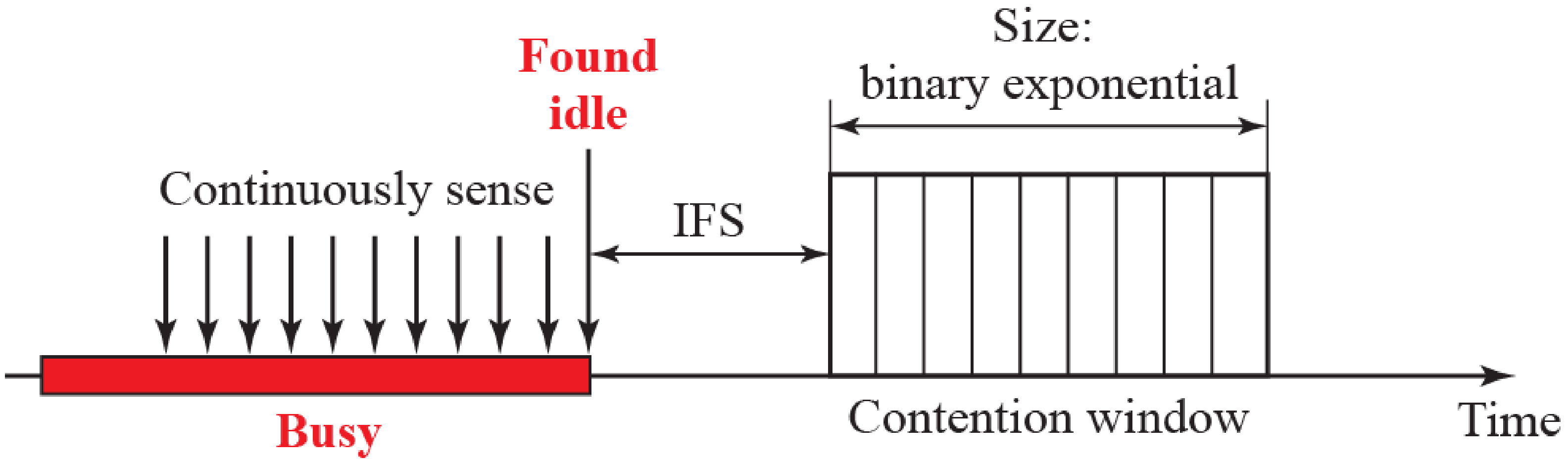
- Is greater than PURE ALOHA.
- Maximum throughput occurs at a different values of G and is based on Persistence method and value p in p-persistent approach.
- For 1-persistent method the maximum through put is around 50% when G=1.
- For non persistent method, the maximum through put can go up to 90% when G is between 3 and 8.

CSMA/CA

- Carrier sense multiple access with collision avoidance (CSMA/CA) was invented for wireless networks.
- Collisions are avoided through the use of CSMA/CA's three strategies:

1) Inter frame space,
2) Contention window
3) Acknowledgments

Contention window



Inter frame Space(IFS)

- Collisions are avoided by deferring the transmission even if channel is idle.
- Station does not send immediately, it waits for a period called IFS.
- Even though channel may be found idle when it is sensed, a distant station may have already started transmitting and the signal has not yet reached this station.
- IFS time allows signal from the distant station to reach this station.

Inter frame Space(IFS)

- After IFS time, if channel is found idle, the station can send, but still need to wait a time equal to contention time.
- IFS variable can be used to prioritize the stations.

In CSMA/CA, the IFS can also be used to define the priority of a station or a frame.

Contention window

- Is an amount of time divided into slots.
- A station that is ready to send chooses a random number of slots as its wait time.
- The number of time slots in the window changes according to binary exponential back off strategy.
- Means it is set to one slot first time and then doubles each time the station can not detect an idle channel after IFS time.

Contention window

- Station needs to sense the channel after each time slot.
- If station finds that channel is busy, it does not restart the process; stops the timer and restarts it when the channel is sensed as idle.
- This gives priority to the station with longest waiting time.

In CSMA/CA, if the station finds the channel busy, it does not restart the timer of the contention window;

it stops the timer and restarts it when the channel becomes idle.

Acknowledgment

- With all these precautions, data may get corrupted during transmission resulting in destroyed data.
- Positive ack. and the time-out timer guarantee that receiver has received the frame.

Frame exchange Time Line

1. Before sending the frame, the source station senses the medium by checking the energy level at the carrier frequency.

   a) Channel uses persistent strategy with back off until channel is idle.

   b) After the station is found idle, the station waits for a period of time called DCF inter frame space(DIFS), then frame sends a control frame called request to send(RTS).

2. After receiving the RTS & waiting a period time called short inter frame space(SIFS), the destination station sends the a control frame, called clear to send(CTS)

Frame exchange Time Line

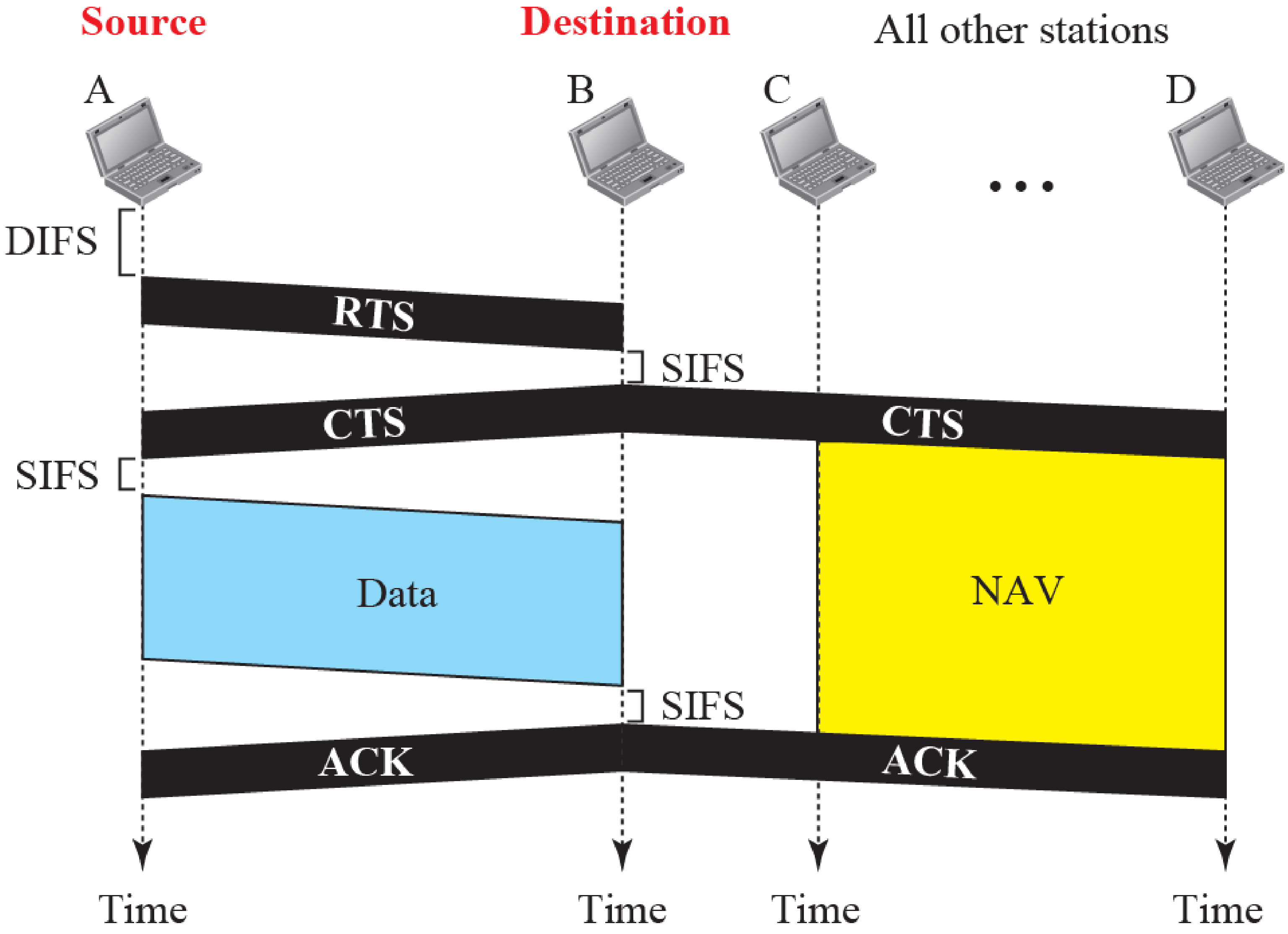
to the source station. The control frame indicates that the destination is ready to receive data.

3. Source station sends data after waiting an amount of time equal to SIFS.

4. Destination station after waiting an amount of time called SIFS, sends an acknowledgment to show that frame has been received.

Acknowledgment is needed in this protocol to check for the successful transmission of its data at the destination.

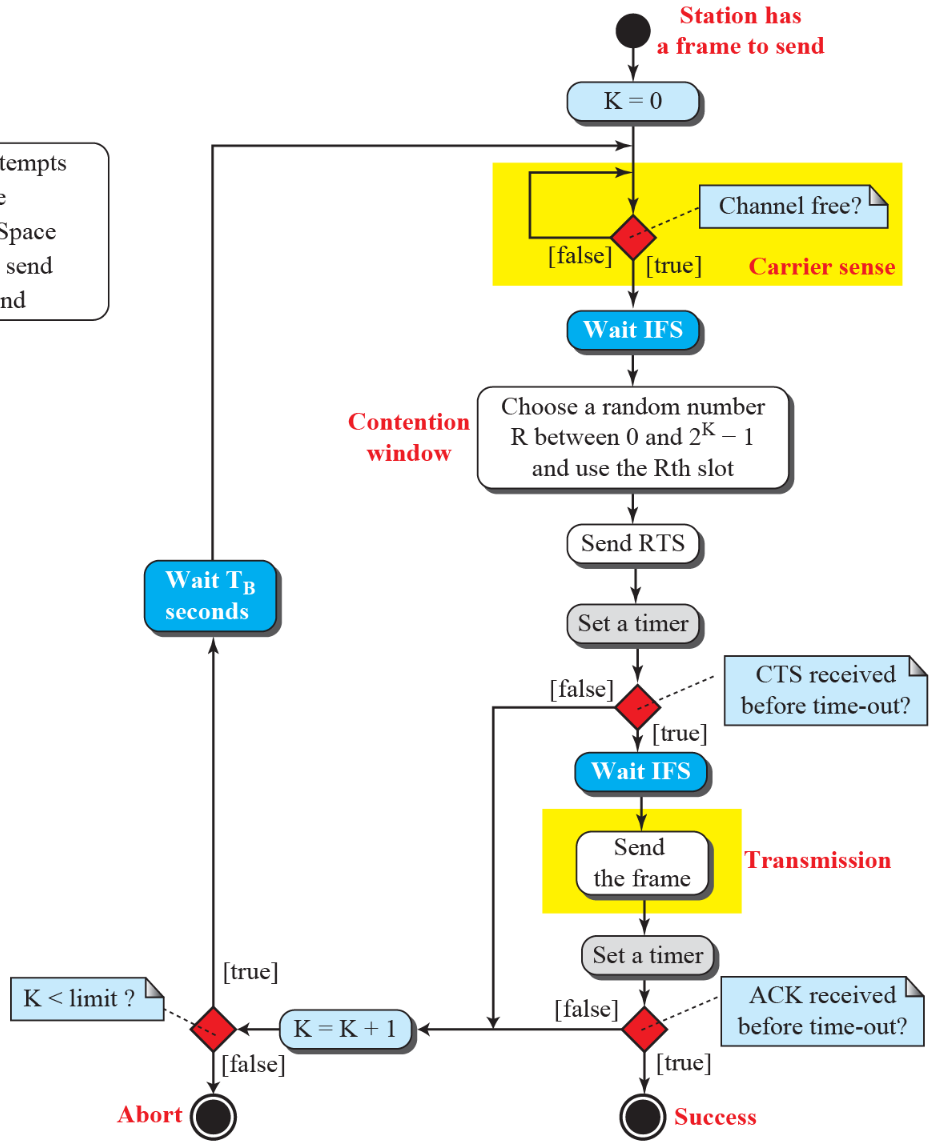
CSMA/CA and NAV



Flow diagram for CSMA/CA

Legend

K: Number of attempts
T_B: Backoff time
IFS: Interframe Space
RTS: Request to send
CTS: Clear to send



Network Allocation vector

- How do other stations defer sending data when one of the station acquire access i.e How collision avoidance aspect is accomplished??
- Feature is NAV

I. When station sends an RTS frame, it includes the duration of the time that it needs to occupy the channel.

➢ The stations that are affected by this transmission creates a times called Network Allocation Vector(NAV), that shows how much time is passed and when these stations are allowed to check the idleness of the medium.

➢ Each time a station access the system and sends an RTS frame, other stations starts their NAV.

➢ i.e Each station before sensing the medium to see if it is idle, it first checks NAV if it has expired.

Collision during the Handshaking

I. What happens if there is a collision during the time when RTS or CTS control frames are in transition(Handshaking period)

- Two or more stations try to sends RTS frame at the same time, the control frames may collide.
- Since there is no mechanism for collision detection, the sender assumes there has been a collision if CTS frame is not received from the receiver.
- The back off strategy is applied and sender tries again.

Hidden station Problem

- Solution is to use hand shake frames.
- When RTS message is sent from B to A, not C.
- Because both B & C are with in transmission range of A, the CTS message, which contains the duration of data transmission from B to A, reaches C.
- Station C knows that there is some hidden station is using the channel and refrains from transmitting until the duration is over.

CONTROLLED ACCESS

- In controlled access, 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 controlled-access methods:
  1) Reservation
  2) Polling
  3) Token passing

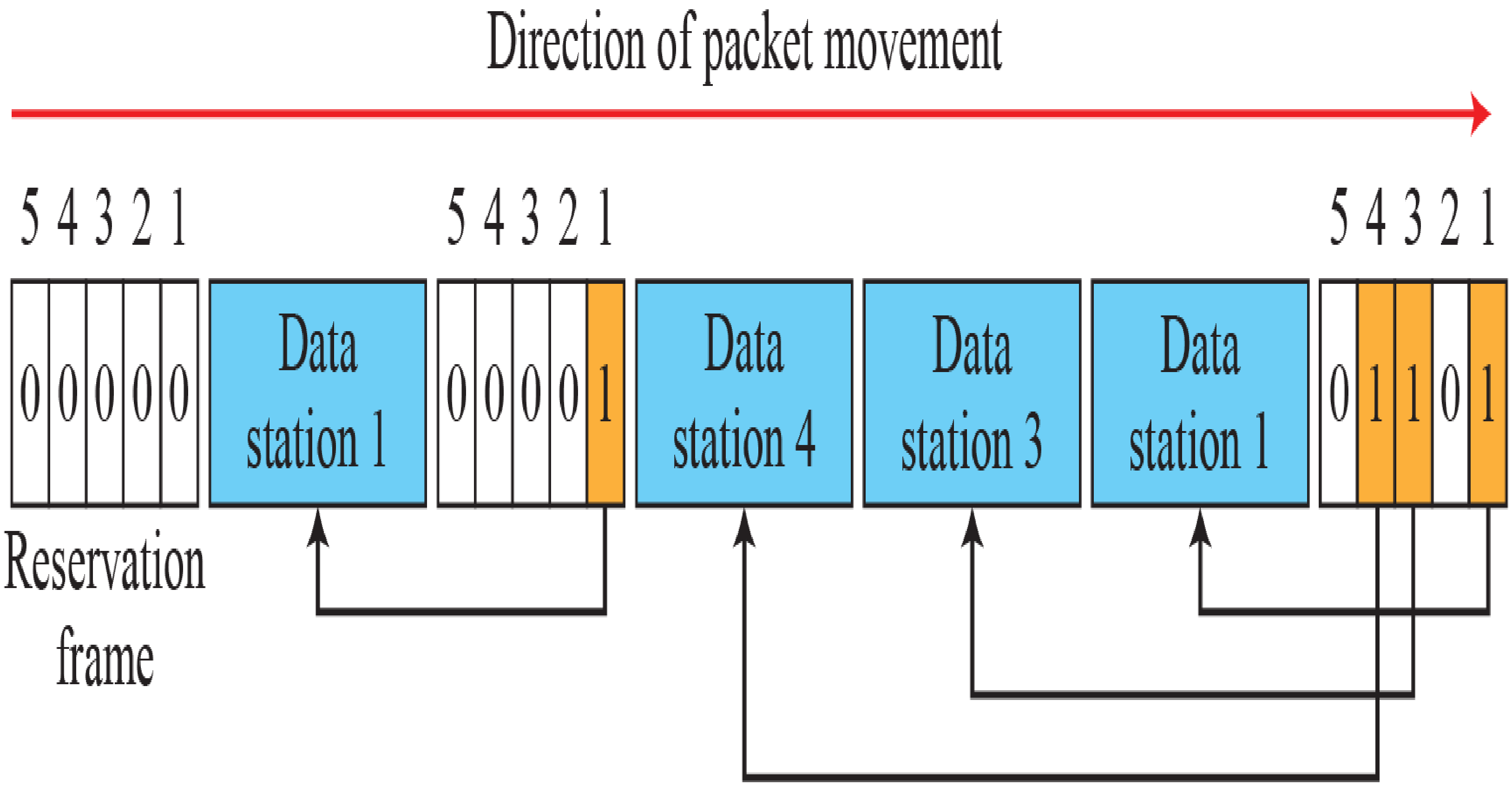
Reservation

- In the reservation method, a station needs to make a reservation before sending data.

- Time is divided into intervals.

- In each interval, a reservation frame precedes the data frames sent in that interval.

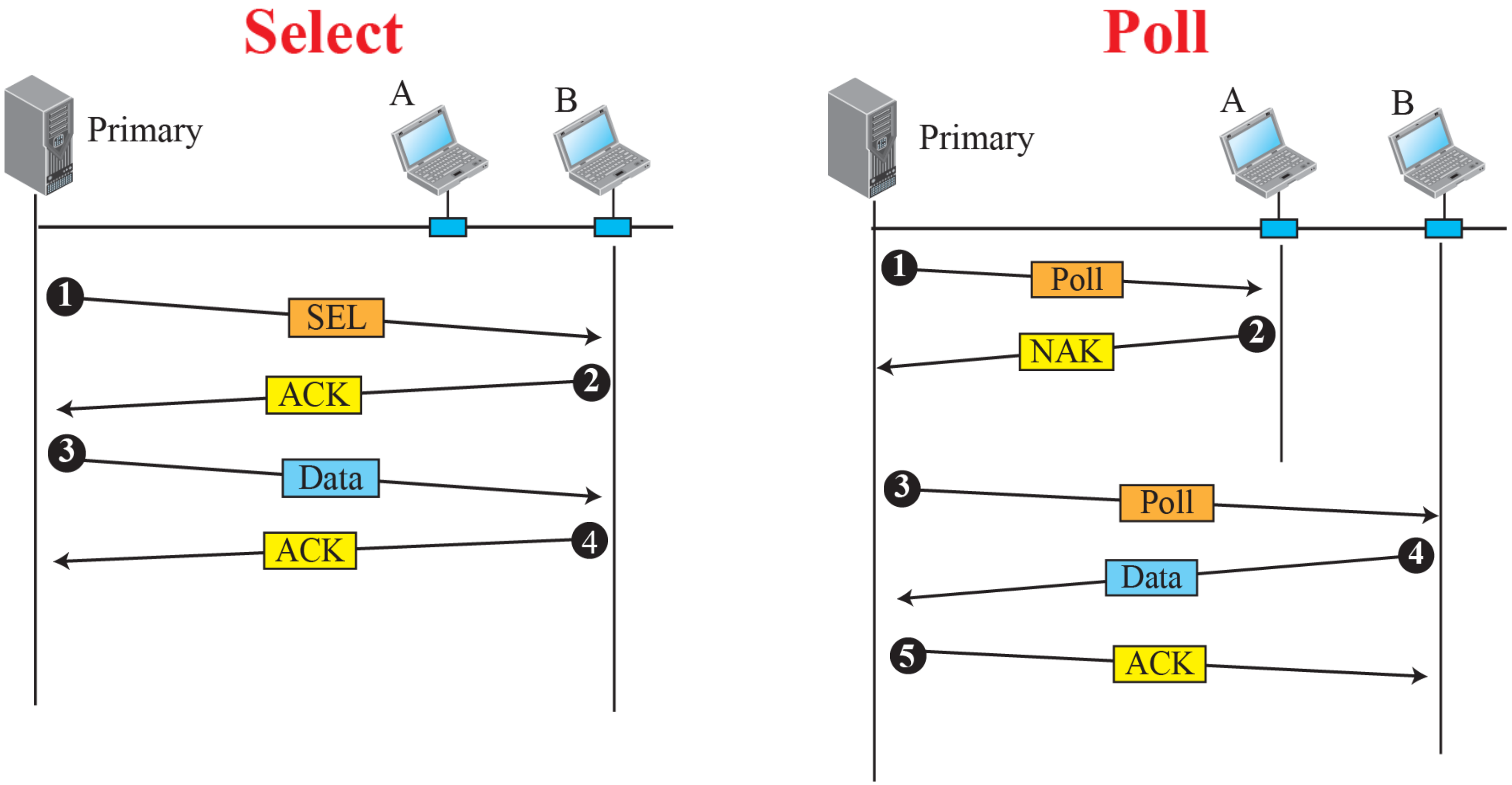
Reservation access method



Polling

- Polling works with topologies in which one device is designated as a primary station and the other devices are secondary stations.
- All data exchanges must be made through the primary device even when the ultimate destination is a secondary device.
- The primary device controls the link; the secondary devices follow its instructions.
- It is up to the primary device to determine which device is allowed to use the channel at a given time.

Select and poll functions in polling-access method



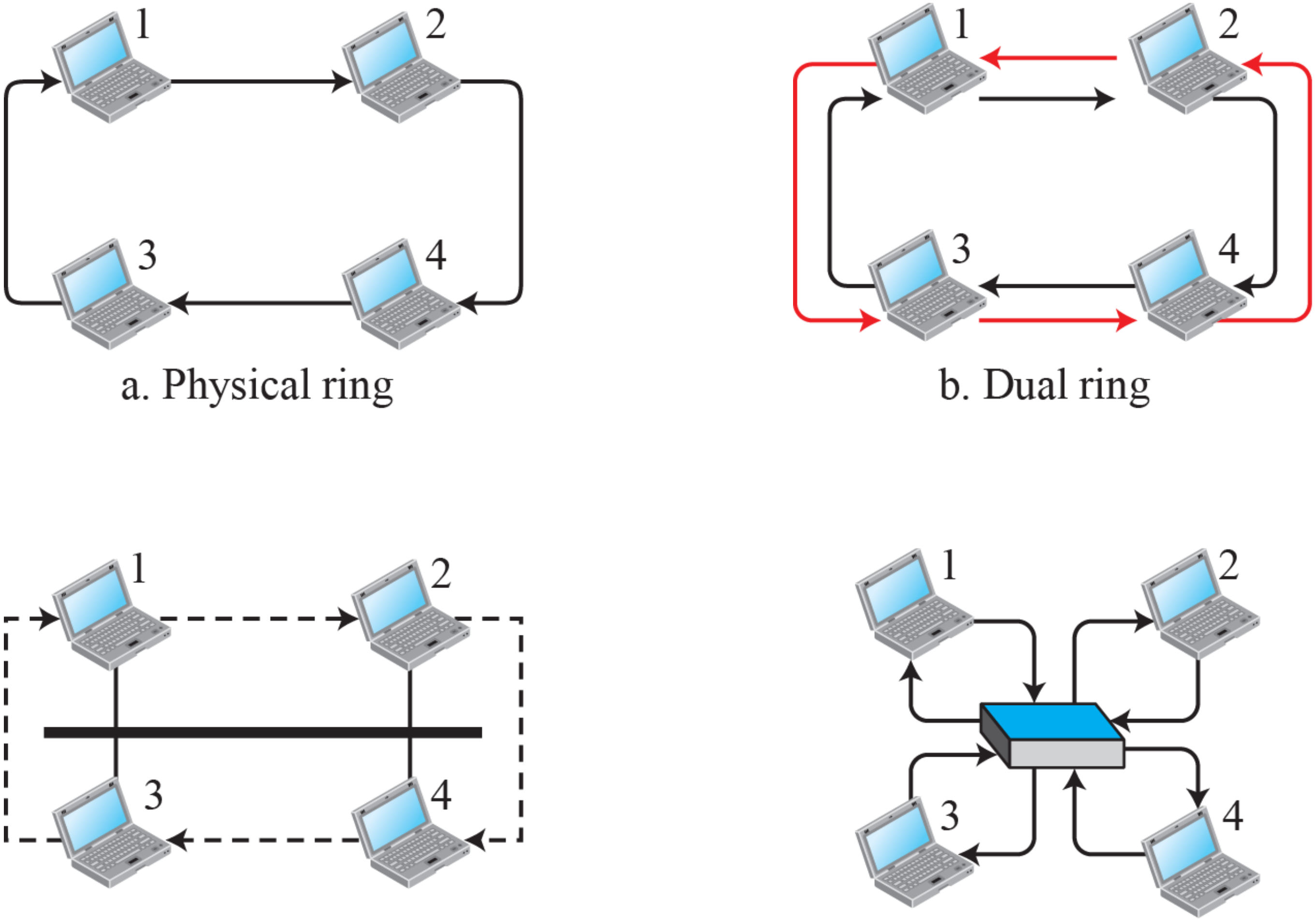
Polling

- If primary wants to receive data, it asks the secondary's if they have anything to send.
- This is called Polling.
- If primary wants to send data, it tells the secondary to get ready to receive.
- This is called Select function.

Token Passing

- In the token-passing method, the stations in a network are organized in a logical ring.
- In other words, for each station, there is a predecessor and a successor.
- The predecessor is the station which is logically before the station in the ring;
- The successor is the station which is after the station in the ring.

Logical ring and physical topology in token-passing access method



a. Physical ring

b. Dual ring

c. Bus ring

d. Star ring

CHANNELIZATION

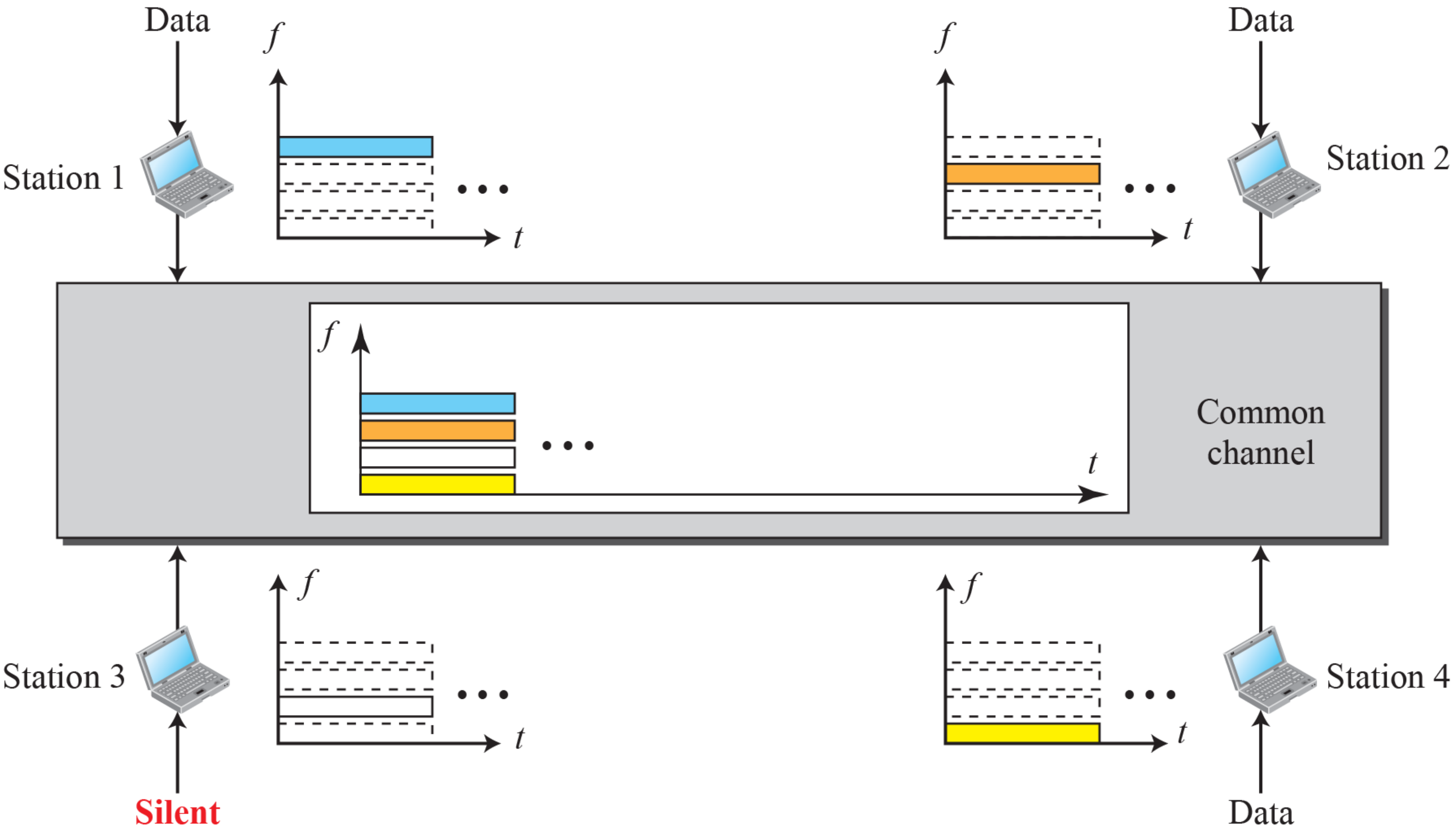
- Channelization (or channel partition, as it is sometimes called) is a multiple-access method in which the available bandwidth of a link is shared in time, frequency, or through code, among different stations.

- Discuss three protocols:
- FDMA, TDMA, and CDMA.

FDMA

- In frequency-division multiple access (FDMA), the available bandwidth is divided into frequency bands. Each station is allocated a band to send its data.
- In other words, each band is reserved for a specific station, and it belongs to the station all the time.
- Each station also uses a bandpass filter to confine the transmitter frequencies.
- To prevent station interferences, the allocates bands are separated from one another by guard bands. FDMA specifies a predetermined frequency band for the entire period of communication

Frequency-division multiple access (FDMA)



FDMA

- FDMA is access method in DLL.
- DLL in each station tells the PL to make band pass 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.
- Signals created at each stations are automatically band pass filtered and mixed when they are sent to the common channel.

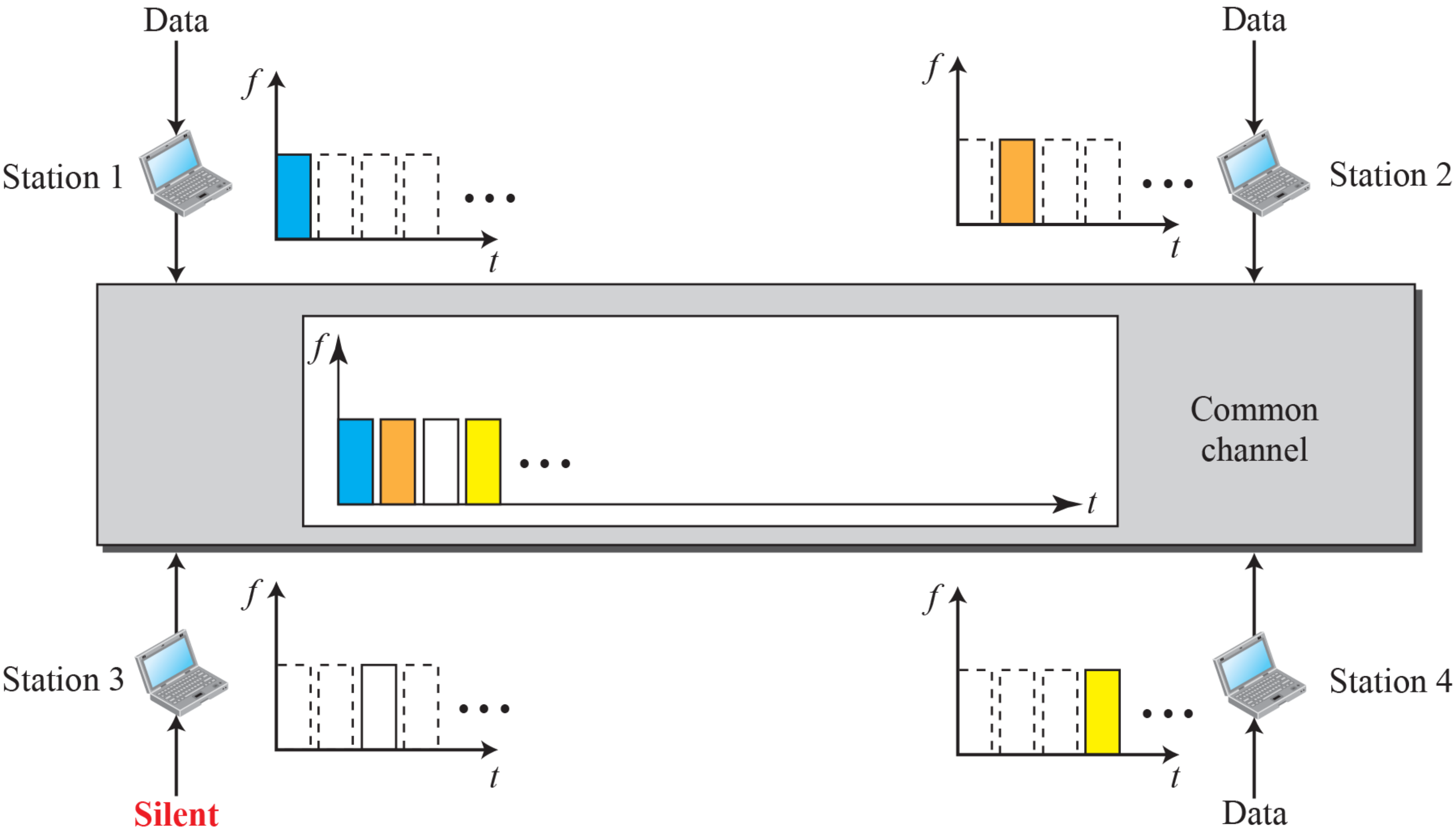
TDMA

- In time-division multiple access (TDMA), the stations share the bandwidth of the channel in time.
- Each station is allocated a time slot during which it can send data.

Each station transmits its data in its assigned time slot.

- Main problem of TDMA lies in achieving synchronization between different stations.
- Each station needs to know beginning of its time slot and location of its slot.

Time-division multiple access (TDMA)



TDMA

- This may be difficult because of propagation delays introduced in the system if the stations are spread over a large area.
- Synchronization is normally accomplished by having some synchronization bits at the beginning of each slot.
- TDMA is an access method in DLL.
- DLL in each station tells its physical layer to use allocated time slots.
- There is no physical multiplexer at the physical layer.

CDMA

➢Code-division multiple access (CDMA) was conceived several decades ago.

➢Recent advances in electronic technology have finally made its implementation possible.

➢CDMA differs from FDMA in that only one channel occupies the entire bandwidth of the link.

It differs from TDMA in that all stations can send data simultaneously; there is no timesharing.

Idea: Assume that there are 4 stations 1,2,3 and 4 connected to the same channel.

→The data from station1 are d1, station2 are d2, station 3 are d3 and station 4 are d4.

→ The code for station1 is C1, station2 is C2, station 3 is C3 and station 4 is C4.

CDMA

- Assume that assigned codes have two properties

1)If we multiply each code by another, we get 0.

2) if we multiply each code by itself, we get 4.

→ Station 1 multiplies its data by its code to get d1.c1.

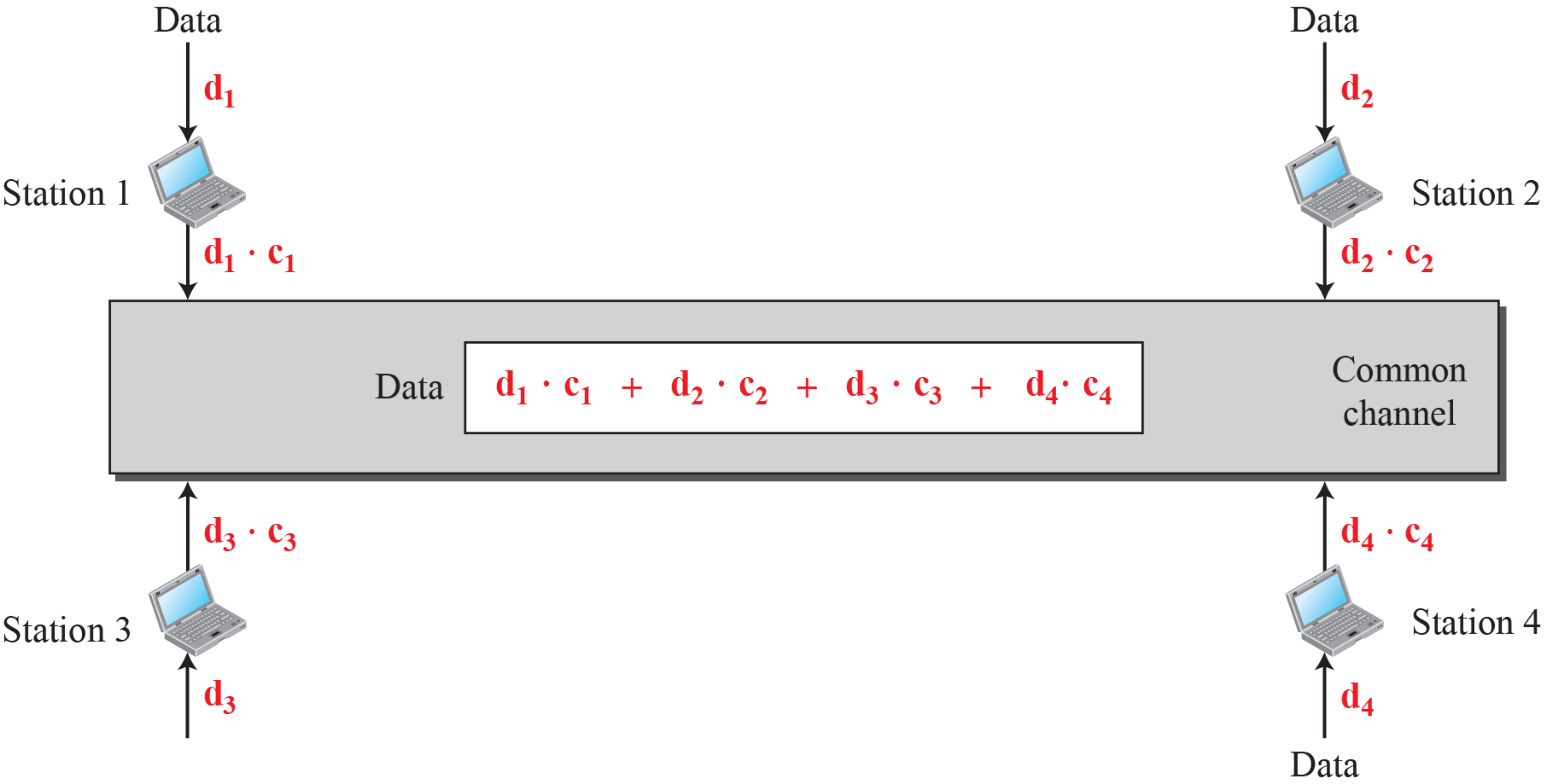
→ Station 2 multiplies its data by its code to get d2.c2.

→ Station 3 multiplies its data by its code to get d3.c3.

→ Station 4 multiplies its data by its code to get d4.c4.

- The data that go on channel is sum of all these terms.
- *CDMA is based on coding theory.*
- *Each station is assigned a code, which is a sequence of numbers called Chips.*

Simple idea of communication with code



Chip sequences

C_1

$[+1 \quad +1 \quad +1 \quad +1]$

C_2

$[+1 \quad -1 \quad +1 \quad -1]$

C_3

$[+1 \quad +1 \quad -1 \quad -1]$

C_4

$[+1 \quad -1 \quad -1 \quad +1]$

Chips

- Sequence is not choosen randomly.
- They are called orthogonal sequences and have following properties.

1) Each sequence is made of N elements, where N is number of stations.
2) If we multiply a sequence by a number, every element in the sequence is multiplied by that number. This is called multiplication of a sequence by a scalar.

   Eg: 2.[+1 +1 -1 -1] = [+2 +2 -2 -2]

3) If we multiply two same sequences element by element, and add the results, we get N. Where N is number of elements in each sequence. This is called Linear product of two equal sequences.

Chips

Eg: [+1 +1 -1 -1]. [+1 +1 -1 -1]= 1+1+1+1=4.

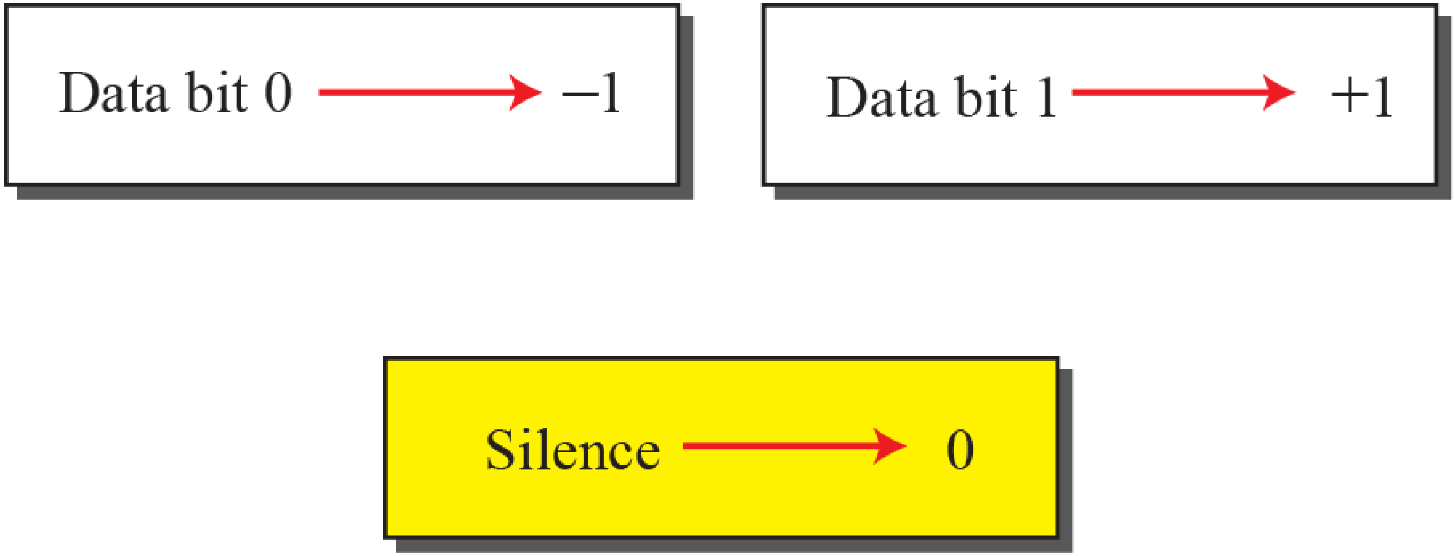
4). If we multiply two different sequences, element by element and add the results, we get 0.This is called inner product of two different sequences.

Eg: [+1 +1 -1 -1]. [+1 -1 -1+1]= +1-1+1-1=0

5) Adding two sequences means adding the corresponding elements. The result is another sequence.

Eg: [+1 +1 -1 -1]+ [+1 +1 -1 -1]= [+2+2 0 0]

Data representation in CDMA



Data bit 0 → −1

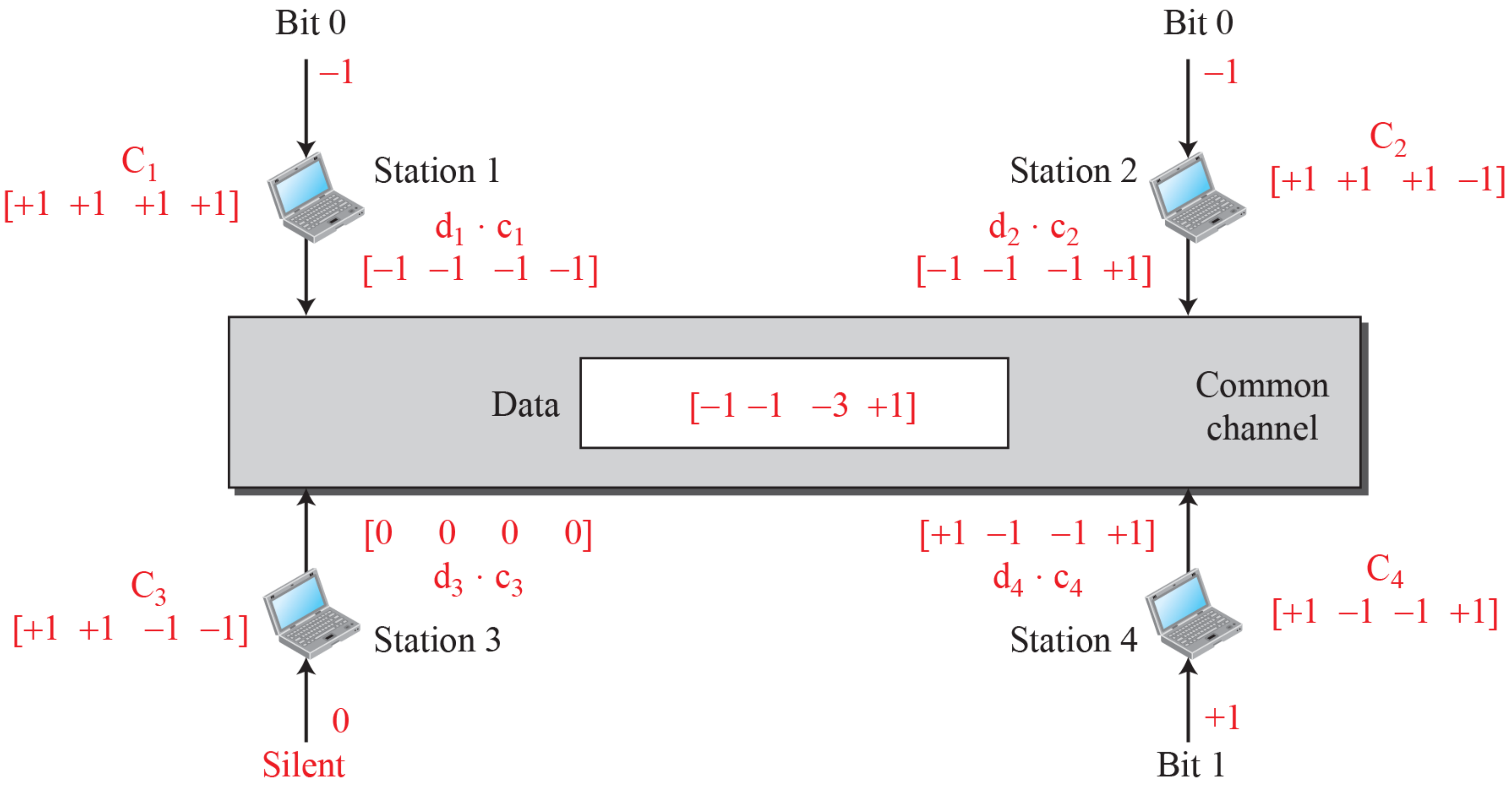
Data bit 1 → +1

Silence → 0

Data representation in CDMA

- Follow the following rules for encoding.
- If a station needs to send 0 bit, it encodes it as -1, to send 1 bit, it encodes it as +1.
- When station is idle, it sends no signal.

Sharing channel in CDMA



Encoding and decoding

- Assume that stations 1 and 2 are sending a 0 bit, station 4 is sending 1 bit. Station 3 is silent.
- The data at the sender site are translated to -1,-1,0 and +1.
- Each station multiplies corresponding number by its chip, which is unique for each station.
- The result is a new sequence which is sent to the channel.
- We Assume that all stations send the resulting sequences at the same time.
- The sequence on the channel is the sum of all four sequences.

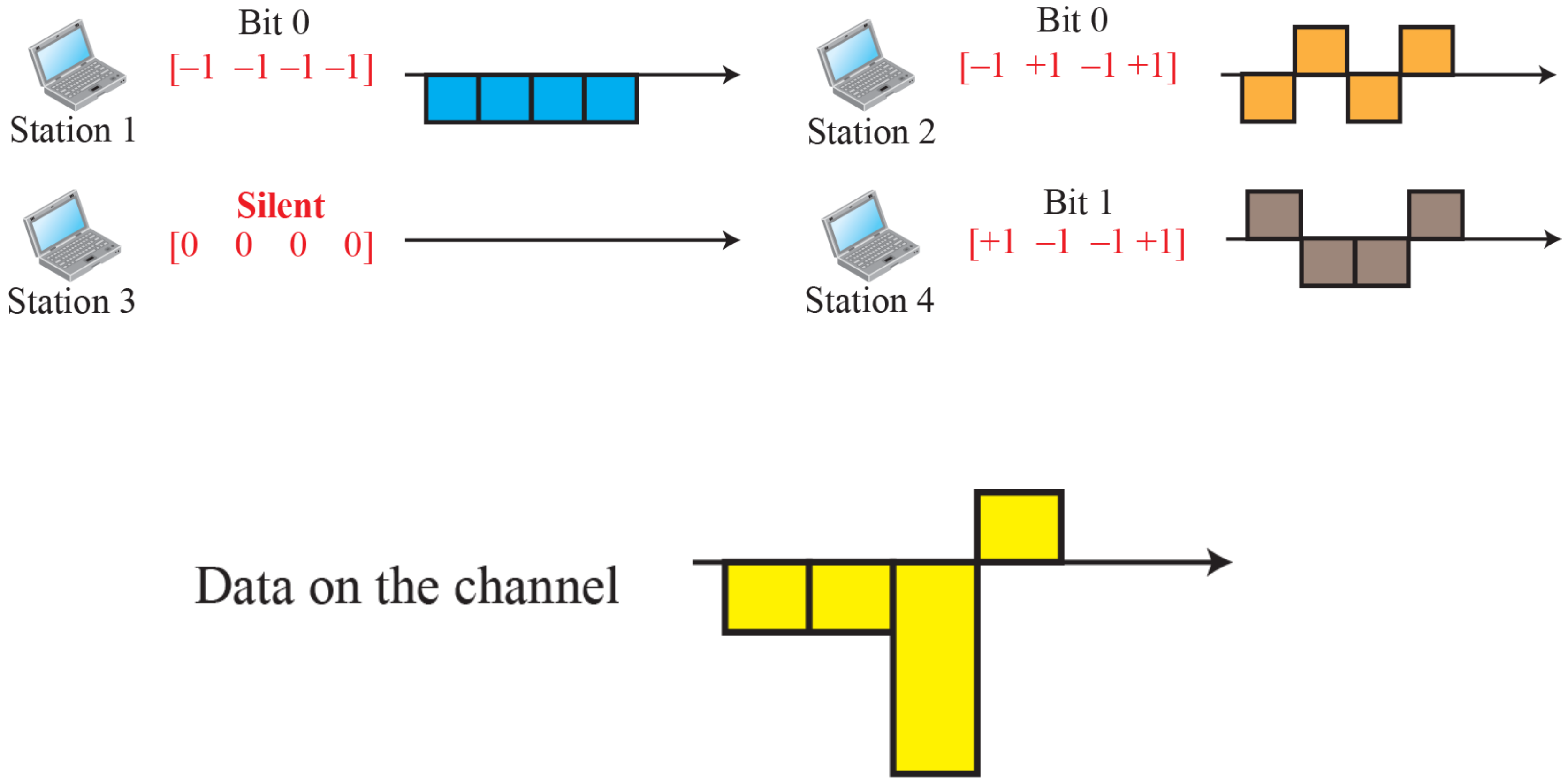
Continued..

- Imagine station 3, which is silent, listening to station 2.
- Station 3 multiplies the total data on the channel by the code for station 2, which is

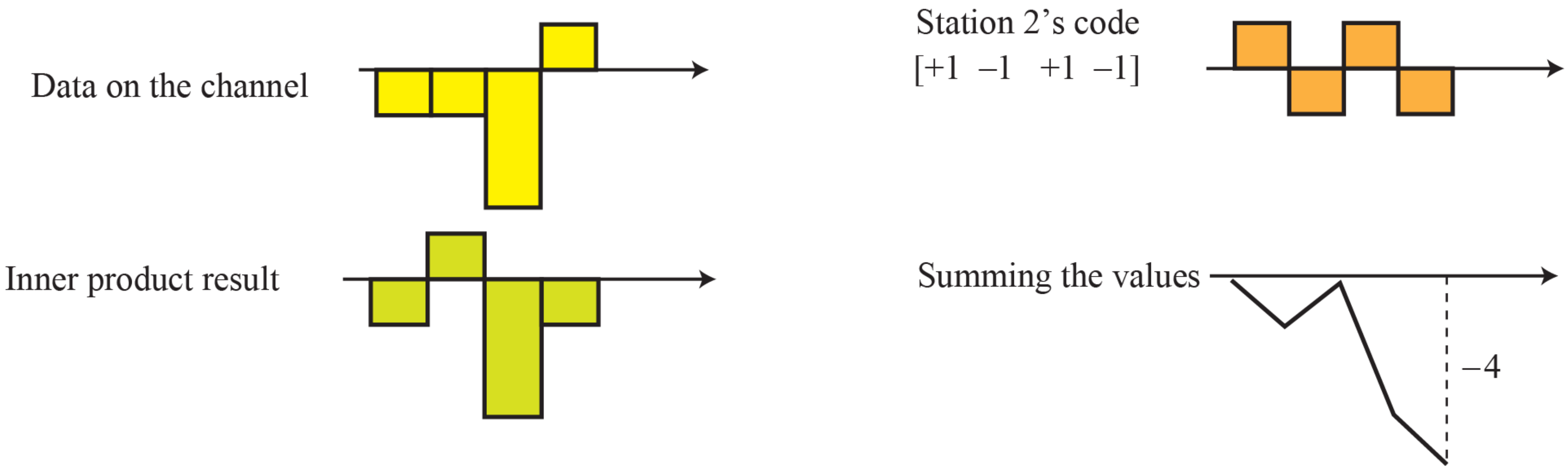
[+1 -1 +1 -1], to get

[-1 -1 -3 +1[.[+1 -1 +1 -1]=-4/4=-1→bit 0.

Digital signal created by four stations in CDMA



Decoding of the composite signal for one in CDMA



−4 → −4/4 → −1 → Bit 0

General rules and examples of creating Walsh tables

- To generate chip sequences, walsh table is used which is a two dimensional table with an equal number of rows and columns.

$$W_1 = \begin{bmatrix} +1 \end{bmatrix} \qquad W_{2N} = \begin{bmatrix} W_N & W_N \\ W_N & \overline{W_N} \end{bmatrix}$$

a. Two basic rules

$$W_1 = \begin{bmatrix} +1 \end{bmatrix} \qquad W_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \qquad W_4 = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix}$$

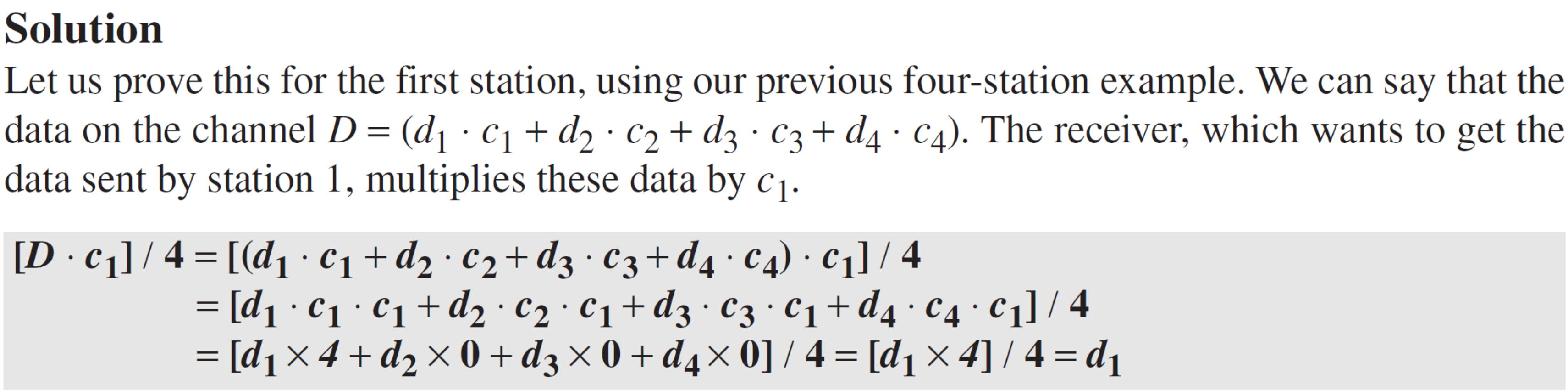
b. Generation of W_1, W_2, and W_4

Walsh table

- In walsh table, each row is a sequence of chips.
- W1 for a one-chip sequence has one row and one column.
- We choose -1 or +1 for the chip.
- According to Walsh, if we know the table for N sequences Wn, we can create a table for 2N sequences W2n.
- *The number of sequences in a Walsh table needs to be $N = 2^m$.*

Example 12.8

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, which wants to get the data sent by station 1, multiplies these data by c_1.

$$
\begin{aligned}
[D \cdot c_1] / 4 &= [(d_1 \cdot c_1 + d_2 \cdot c_2 + d_3 \cdot c_3 + d_4 \cdot c_4) \cdot c_1] / 4 \\
&= [d_1 \cdot c_1 \cdot c_1 + d_2 \cdot c_2 \cdot c_1 + d_3 \cdot c_3 \cdot c_1 + d_4 \cdot c_4 \cdot c_1] / 4 \\
&= [d_1 \times 4 + d_2 \times 0 + d_3 \times 0 + d_4 \times 0] / 4 = [d_1 \times 4] / 4 = d_1
\end{aligned}
$$

Example-2

- 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.

MODULE – 5

Wired LANs Ethernet

ETHERNET PROTOCOL

- The data-link layer and the physical layer are the territory of the local and wide area networks

- Discussed LAN whose sole purpose is to share the resources

- Most of the LANs today is connected to WAN or Internet

- Ethernet Protocol is most popular LAN technology because it was able to update itself to meet the needs of organization with the demand for higher data transmission.

IEEE Project 802

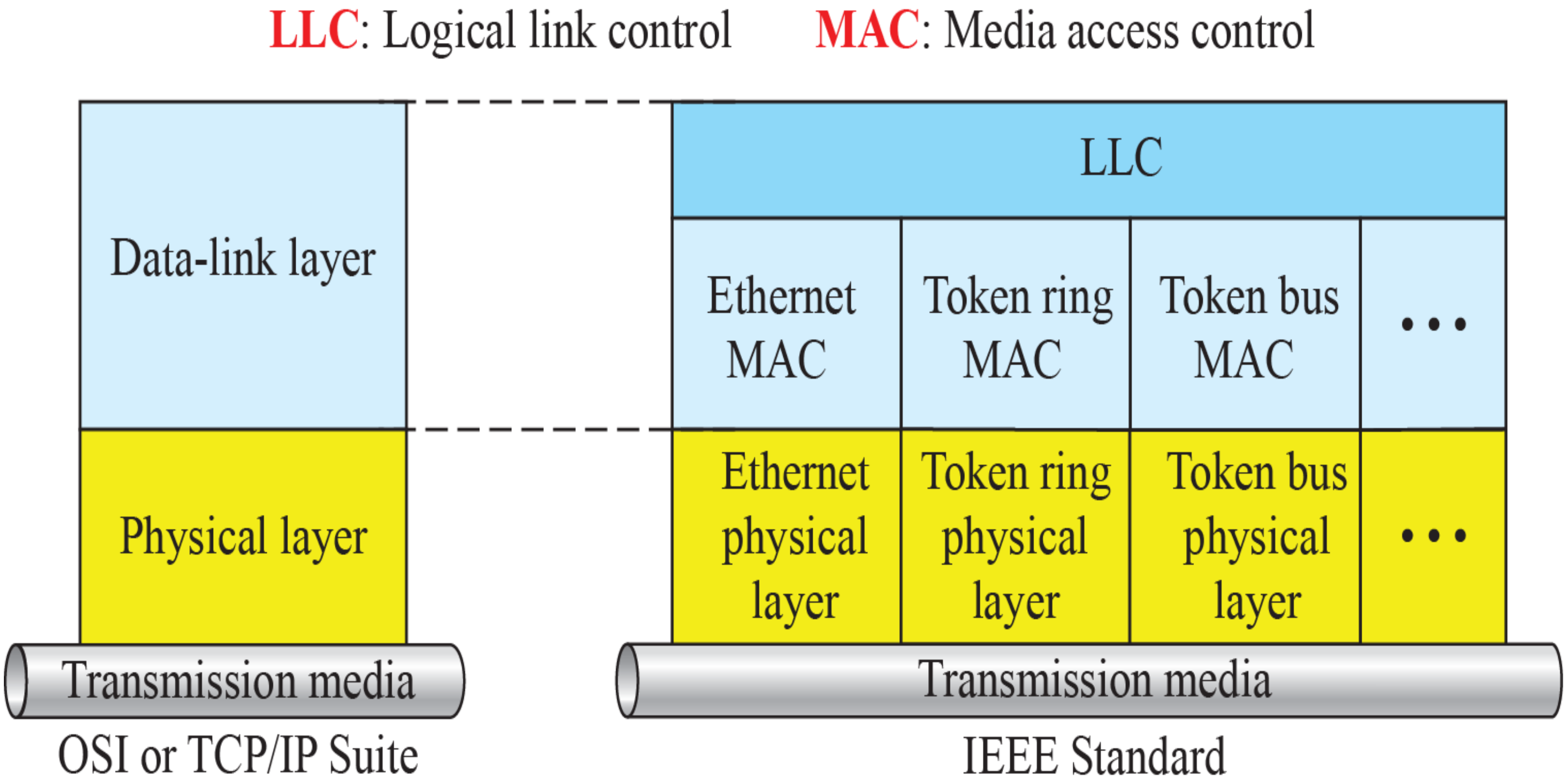
- In 1985, the Computer Society of the IEEE started a project, called Project 802, to set standards to enable intercommunication among equipment from a variety of manufacturers.

- Project 802 does not seek to replace any part of the OSI model or TCP/IP protocol suite

- Instead, it is a way of specifying functions of the physical layer and the data-link layer of major LAN protocols

- The relationship of the 802 Standard to the TCP/IP protocol suite is shown in the Figure below

IEEE standard for LANs



IEEE Project 802

- IEEE has subdivided the data link into two sub layers

  i) Logical Link Control(LLC)

  ii) Media Access Control(MAC)

- Logical Link Control: In IEEE 802 project, Flow Control, Error Control & part of framing duties are collected into this sub layer

- Framing is handled in both LLC & MAC sub layers

- LLC provides a single link layer control protocol for IEEE LANs. i.e LLC protocol can provide inter connectivity between different LANs

- Media access Control that defines the specific access method for each LAN

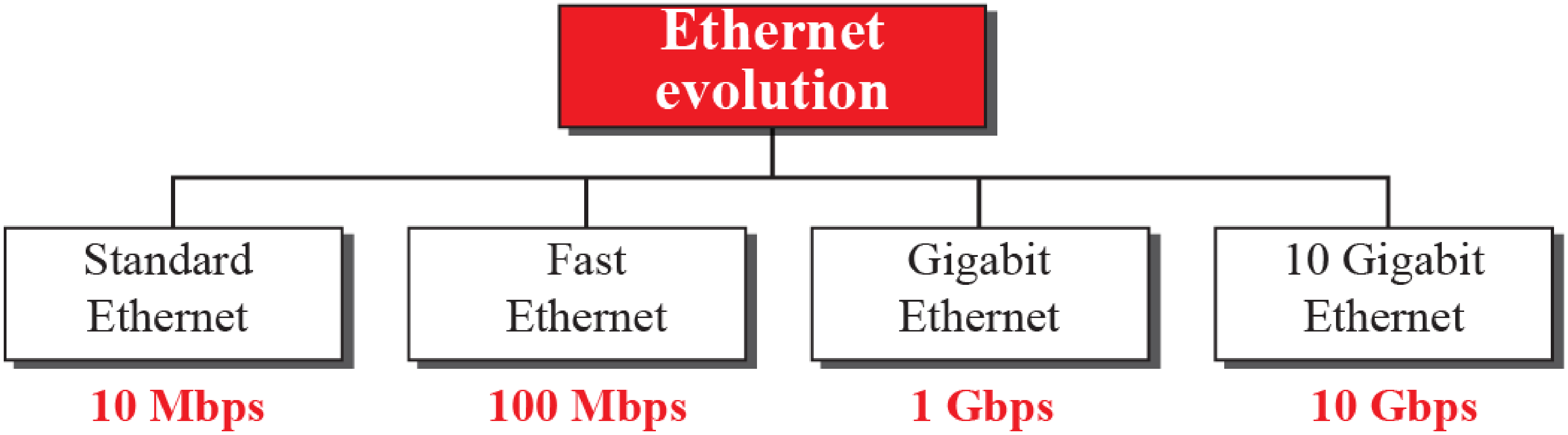
Ethernet Evolution

- The Ethernet LAN was developed in the 1970s by Robert Metcalfe and David Boggs

- It has gone through four generations:

1)Standard Ethernet (10 Mbps)
2)Fast Ethernet (100 Mbps)
3)Gigabit Ethernet (1 Gbps)
4)10 Gigabit Ethernet (10 Gbps)

Ethernet evolution



STANDARD ETHERNET

- Original Ethernet technology with the data rate of 10 Mbps as the Standard Ethernet.
- Although most implementations have moved to other technologies in the Ethernet evolution, there are some features of the Standard Ethernet that have not changed during the evolution.

Characteristics

- Discuss some characteristics of the Standard Ethernet.

Connectionless and unreliable service.

- Ethernet provides connectionless service i.e each frame sent is independent of the previous or next frame.
- Ethernet has no connection establishment or connection termination phases.
- Sender sends a frame whenever it has and receiver may not be ready for it.
- Sender may overwhelm the receiver buffer with frames, which may result in dropping frames.
- IP, connectionless protocol, which is using the services of the Ethernet, which is also connectionless, will not know if frame is dropped. If TL is also connectionless using UDP, frame is lost

Characteristics

- This loss will come to know only from Application Layer

- If TL is using TCP, Sender TCP segment if does not receive acknowledgment for its segment and sends it again

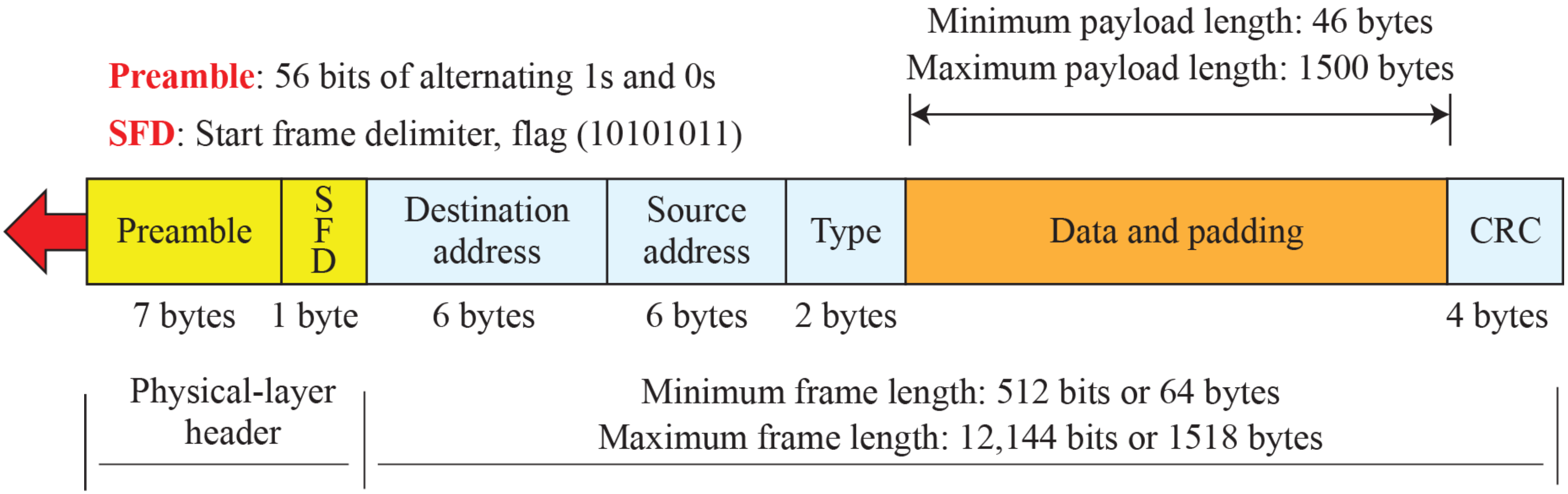
- Ethernet if connectionless, if frame is corrupted during transmission, and receiver finds out about the corruption because of CRC-32, the receiver drops the frame silently

- It is the duty of high level protocols to find out

Ethernet frame

Preamble: 56 bits of alternating 1s and 0s

SFD: Start frame delimiter, flag (10101011)



Minimum payload length: 46 bytes
Maximum payload length: 1500 bytes

Preamble | SFD | Destination address | Source address | Type | Data and padding | CRC

7 bytes | 1 byte | 6 bytes | 6 bytes | 2 bytes | 4 bytes

Physical-layer header

Minimum frame length: 512 bits or 64 bytes
Maximum frame length: 12,144 bits or 1518 bytes

802.3 MAC frame

- Ethernet frame contains 7 fields.
- Ethernet does not provide any mechanism for Ack.ing the received frames.
- Ack. must be implemented at the higher layers.

1. Preamble: contains 7 bytes(56 bits) of alternating 1s and 0s to alert the station that frame is arriving and enables it to synchronize its i/p timing.

→Pattern provides only an alert and a timing pulse.

→Preamble is actually added at the physical layer and is not part of the frame.

2.Start frame delimiter(SFD):1 byte10101011 signals the beginning of the frame.

802.3 MAC frame

→ Warns the station/stations that this is the last chance of synchronization.

→ Last 2 bits is 11 and alerts the receiver that the next field is the destination address.

3.Destination address: is 6 bytes and contains the physical address of the destination station/s to receive the packet.

4.Source address: is 6 bytes and contains the physical address of the destination station/s to receive the packet.

5.Type: This field defines the upper layer protocol whose packet is encapsulated in the frame. The protocols can be IP,ARP,OSPF etc.,

It is used for multiplexing and demultiplexing.

802.3 MAC frame

6.Data: This field carries data encapsulated from the upper layer protocols. It is minimum of 46 bytes and maximum of 1500 bytes.

If data coming from upper layers is more than 1500 bytes, it should be fragmented and encapsulated in more than one frame. If less than 46 bytes, it should be padded with 0's.

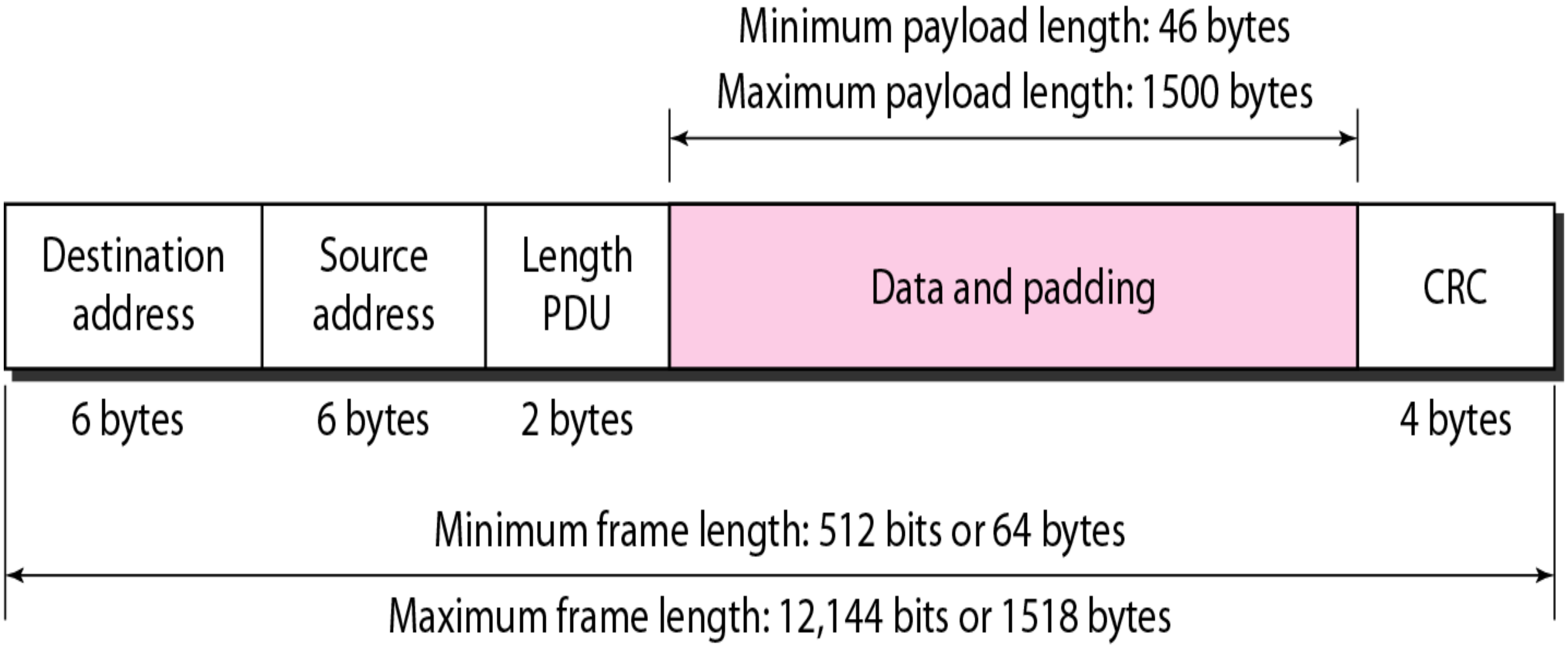
A padded data frame is delivered to upper layer protocol as it is and it is the responsibility of the upper layer protocol to remove the zeros.

7.CRC: Contains error detection mechanism(CRC-32). The CRC is calculated over address, types and data field.

if receiver calculates CRC and finds that it is not zero, it discards the frame.

Frame Length

- Ethernet imposed min and max lengths of a frame. The minimum length restriction is required for the correct operation of CSMA/CD.



Minimum payload length: 46 bytes
Maximum payload length: 1500 bytes
Destination address
Source address
Length PDU
Data and padding
CRC
6 bytes
6 bytes
2 bytes
4 bytes
Minimum frame length: 512 bits or 64 bytes
Maximum frame length: 12,144 bits or 1518 bytes

Frame Length

- Ethernet frame needs to have a minimum length of 512 bits or 64 bytes.
- Part of this length is header and the trailer.
- 18 bytes of header and trailer, length,CRC, then minimum length of data from upper layer protocols is 64-18=46 bytes.
- If upper layer packet is less than 46 bytes, padding is added to make up the difference.
- The standard defines maximum length of a frame without preamble and SFD is 1518 bytes.
- Max. length has two historical reasons

Frame Length

1) Memory was expensive when Ethernet was designed. Max. length restriction helped to reduce the size of the buffer.

2) Max. length restriction prevents one station from monopolizing the shared medium blocking other stations that have data to send.

Minimum Frame length:64 bytes (512 bits)

Minimum data length:46 bytes (368bits)

Maximum frame length: 1518 bytes (12,144 bits)

Maximum data length : 1500 bytes(12000 bits)

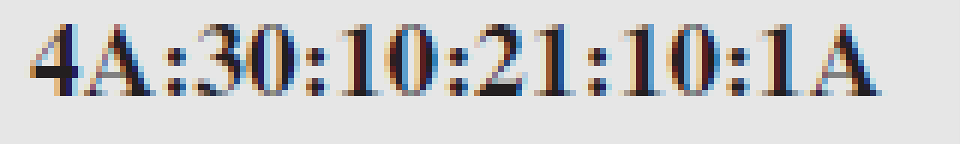
Addressing

- Each station on an Ethernet network (such as a PC, workstation, or printer) has its own network interface card (NIC)

- The NIC fits inside the station and provides the station with a link-layer address

- The Ethernet address is 6 bytes (48 bits), normally written in hexadecimal notation, with a colon between the bytes

- For example, the following shows an Ethernet MAC address:



4A:30:10:21:10:1A

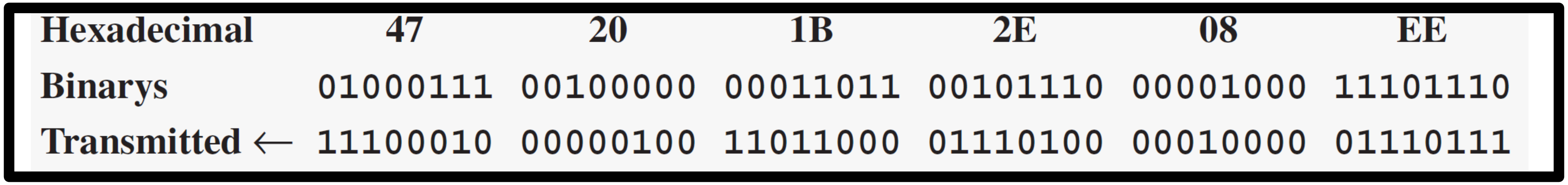
- The transmission is left to right, byte by byte. For each byte, the LSB is sent first & MSB is sent last.

Example 13.1

Show how the address 47:20:1B:2E:08:EE is sent out online.

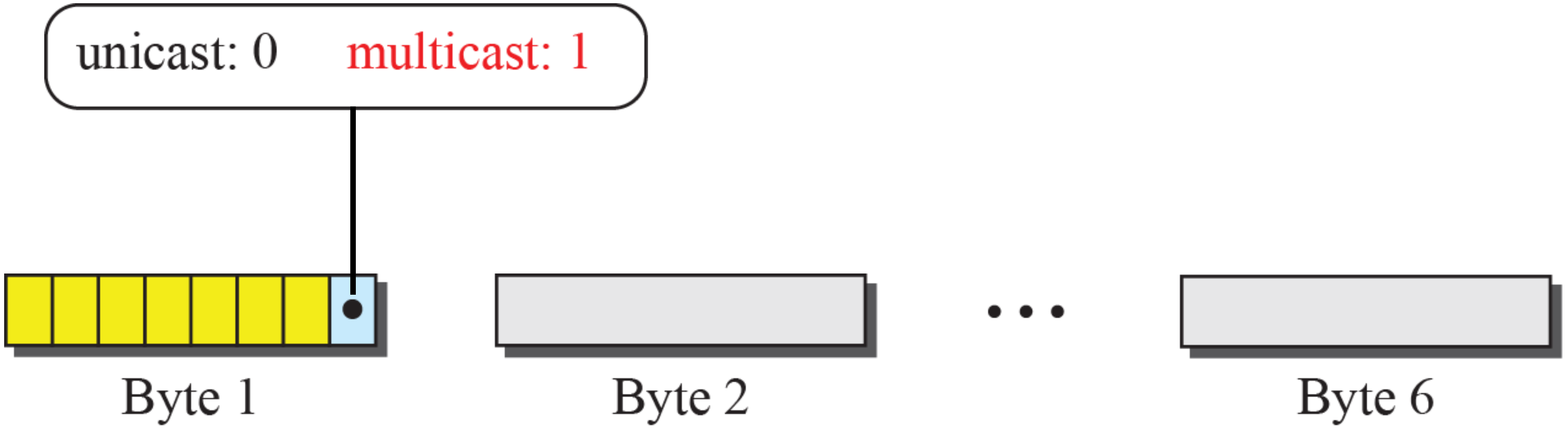
Solution

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
Binarys	01000111	00100000	00011011	00101110	00001000	11101110
Transmitted ←	11100010	00000100	11011000	01110100	00010000	01110111

Unicast and multicast addresses



Continued..

- Source address is always Uni cast address- frame comes from only one station
- Destination can be Uni cast, Multi cast or Broad cast
- The least significant bit of the first byte defines the type of address. If the bit is 0, the address is unicast; otherwise, it is multicast
- The broadcast destination address is a special case of the multicast address in which all bits are 1s.
- The recipients are all stations on the LAN. Broad cast destination address is forty eight 1s.

Example 13.2

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,

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:

Example 13.2 (continued)

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.

Distinguish between Unicast, multicast & Broadcast transmission

- Std. Ethernet uses coaxial cable(bus) or a set of twisted pair cables with hub topology
- Transmission in standard Ethernet is always broadcast, even if the intention is multicast, unicast or broadcast
- In bus topology, when station A sends frame to B, all stations will receive it
- In star topology, when station sends frame to B, hub receives it, since hub is a passive element, it does not check the destination address of the frame, it regenerates the bits and sends to all stations except A.
- Question is how unicast, broadcast and multicast transmission are distinguished from each other????

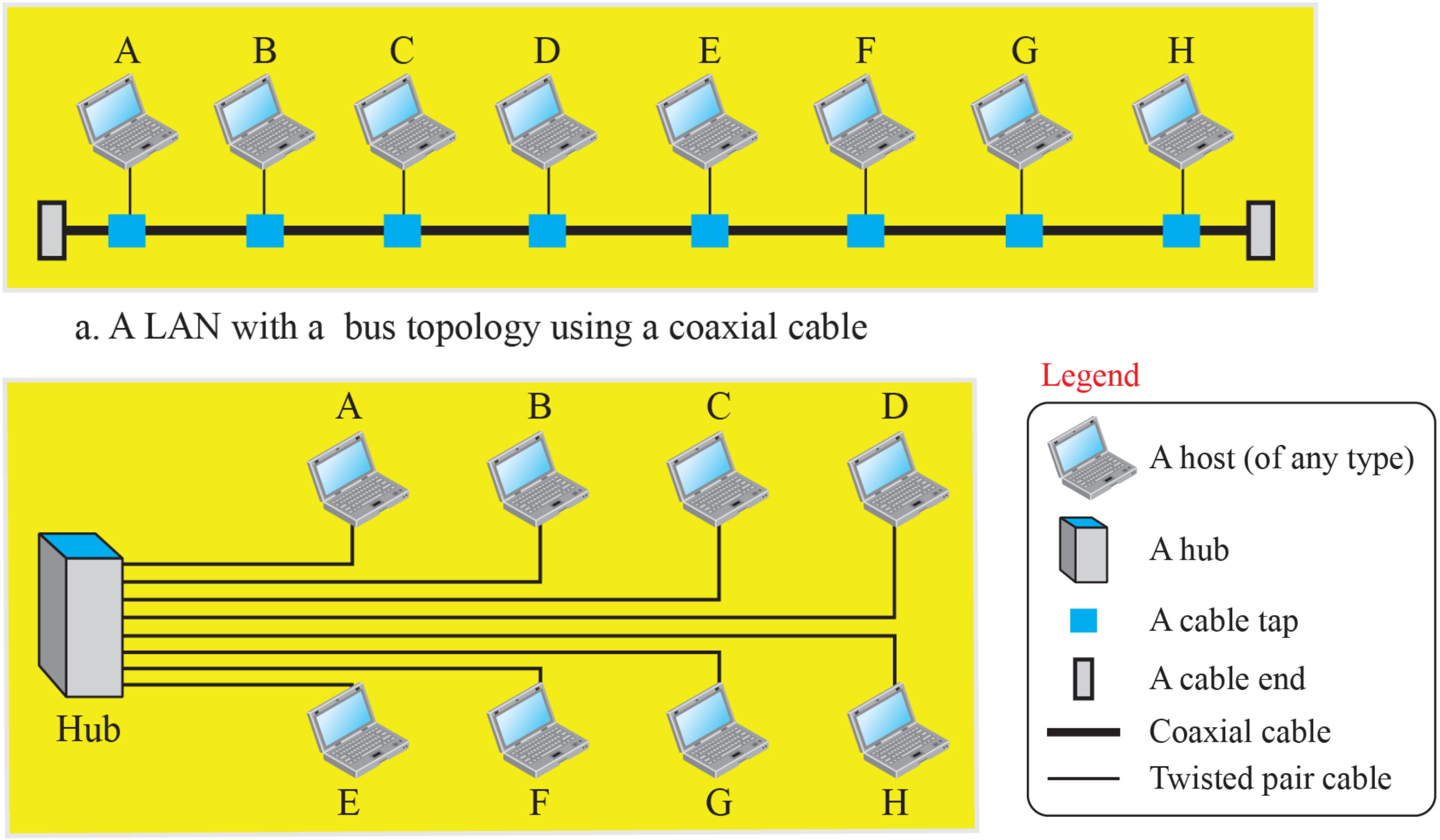
Distinguish between Unicast, multicast & Broadcast transmission

- In Unicast transmission, all stations will receive a frame, only the intended station will keep the frame rest will discard

- In Multi cast transmission, all stations will receive a frame, the stations that are part of group will keep the frame rest will discard

- In broad cast transmission, all stations will receive a frame except the sender and all stations will keep the frame

Implementation of standard Ethernet



a. A LAN with a bus topology using a coaxial cable

b. A LAN with a star topology using a hub

Access Method

➢ Network that uses the standard Ethernet protocol is a broadcast network, need to use an access method to control access to the sharing medium

➢The standard Ethernet choose CSMA/CD with 1-persistent method

➢Let us use a scenario to see how this method works for the Ethernet protocol

Scenario –I

➢Assume station A has a frame to send to station D

1) Station A first should sense(carrier) if any other station is sending. It measures the energy level of the medium for a short period of time Eg:100microsec.

   → If no signal energy on the medium, it means no station is sending data, station A interprets the medium as idle and starts sending the frame

Access Method(Continued..)

- If signal energy on the medium is not zero, it means medium is not idle. station A continuously monitors the medium until it becomes idle for 100 microsecs and starts sending the frame.
- Station A needs to keep a copy of the frame in its buffer to make sure that there is no collision.

Scenario-II

- Medium sensing does not stop after station A has started sending the frame. Station A needs to send & receive continuosly.

Case 1:(With the Assumption of Length of cable 5120 mts)

- Station has send 512 bits and no collision sensed, station then makes sure that frame will go through and stops sensing.

Access Method(Continued..)

- ➢If transmission rate of Ethernet is 10 Mbps, it means it takes 51.2microsec to send 512 bits.
- ➢With speed of propagation in cable ($2*10^8$ mts), the first bit would have gone 10,240mts(one way) or 5120mts(Round trip)
- ➢If collision were to occur it should occur by the time the sender has sent out first 512 bits(worst case) and first bit has made a round trip of 5120 mts.
- ➢If collision happens in the middle of the cable, station A hears the collision earlier and abort the transmission.

Case 2:Station A has sensed a collision before sending 512 bits.

- ➢Means one of the bit ahs collided with bit from other station.
- ➢Both stations refrain from sending and keeps a copy of the same in the buffer for resending when the link will be free.
- ➢However station informs other stations that there is a collision by sending 48 bit Jam signal

Access Method(Continued..)

- Station needs to increment the value of K. if k=15, station thinks network is busy and abort its effort.
- If K <15, the station can wait for back off time and restarts process.
- Station creates random number between 0 to 2^k -1, which means each time the collision occurs, the range of random numbers increases exponentially.

- K=1, random numbers in the range are 0 and 1.
- K=2 range is 0,1,2,3 etc..

Efficiency of Standard Ethernet

- The efficiency of the Ethernet 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

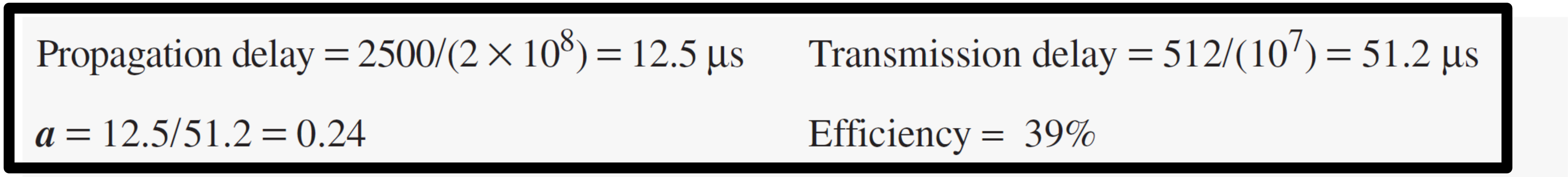
$$\text{Efficiency} = 1 / (1 + 6.4 \times a)$$

Where a is number of frames that can fit on the medium.
It can be calculated as

a=(Propagation delay/ Transmission delay)
PD→Time taken to reach end of medium
TD→Time taken to sent out an average size frame,
NOTE: If value of a decreases, efficiency increases.
Means if the length of the medium is shorter or frame length is longer, the efficiency increases.

Example 13.3

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 $= 512/(10^7) = 51.2\ \mu s$

$\boldsymbol{a} = 12.5/51.2 = 0.24$ Efficiency $=\ 39\%$

The example shows that a = 0.24, which means only 0.24 of a frame occupies the whole medium in this case. The efficiency is 39 percent, which is considered moderate; it means that only 61 percent of the time the medium is occupied but not used by a station.

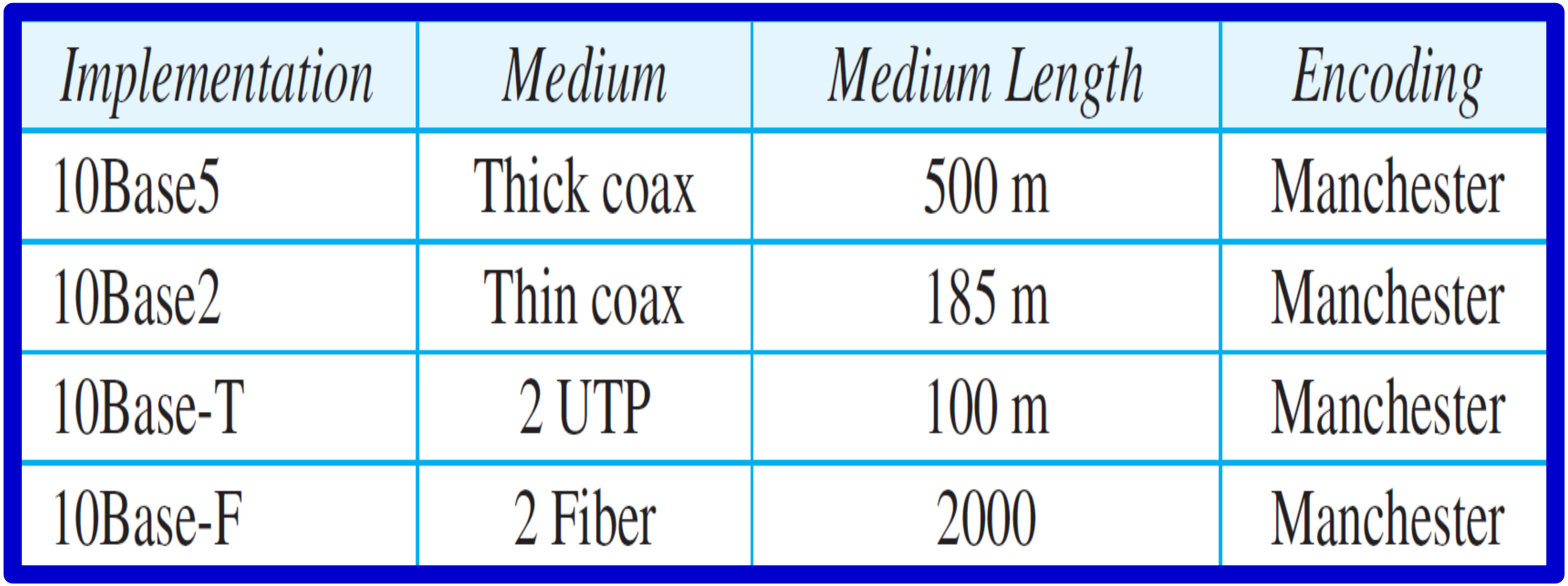
Implementation

- The Standard Ethernet defined several implementations, but only four of them became popular during the 1980s

- Table below shows a summary of Standard Ethernet implementations

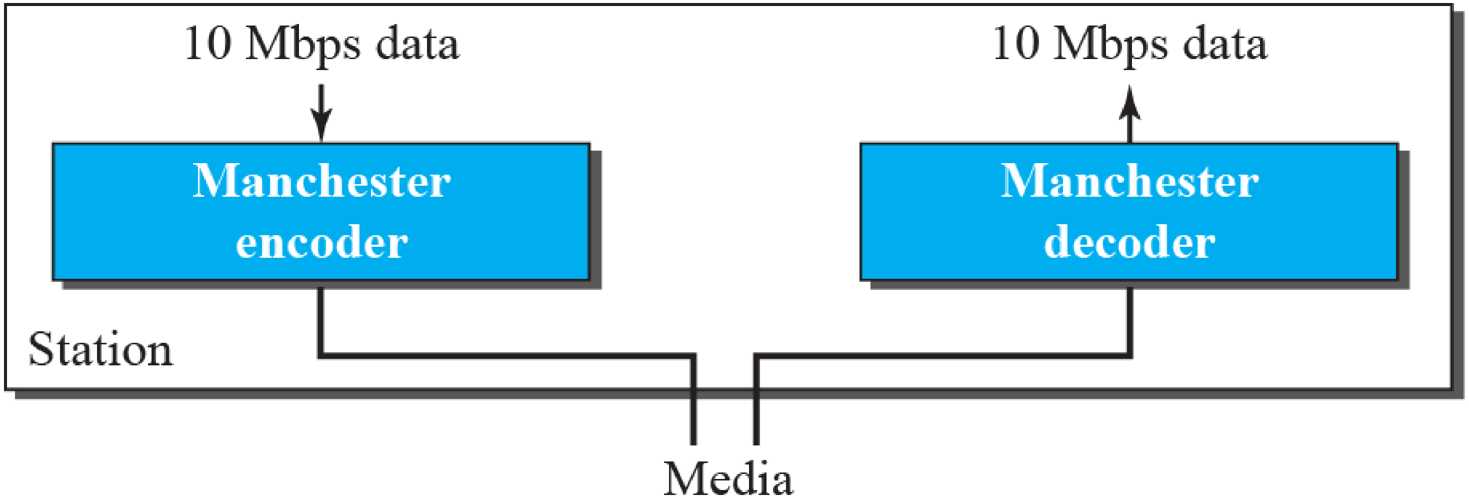
- In the nomenclature, 10 BaseX, number defines the data rate 10Mbps, Base means base band signal, X defines either maximum size of the cable in 100 mts or the type of the cable

Table : 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	Manchester

Encoding in a Standard Ethernet



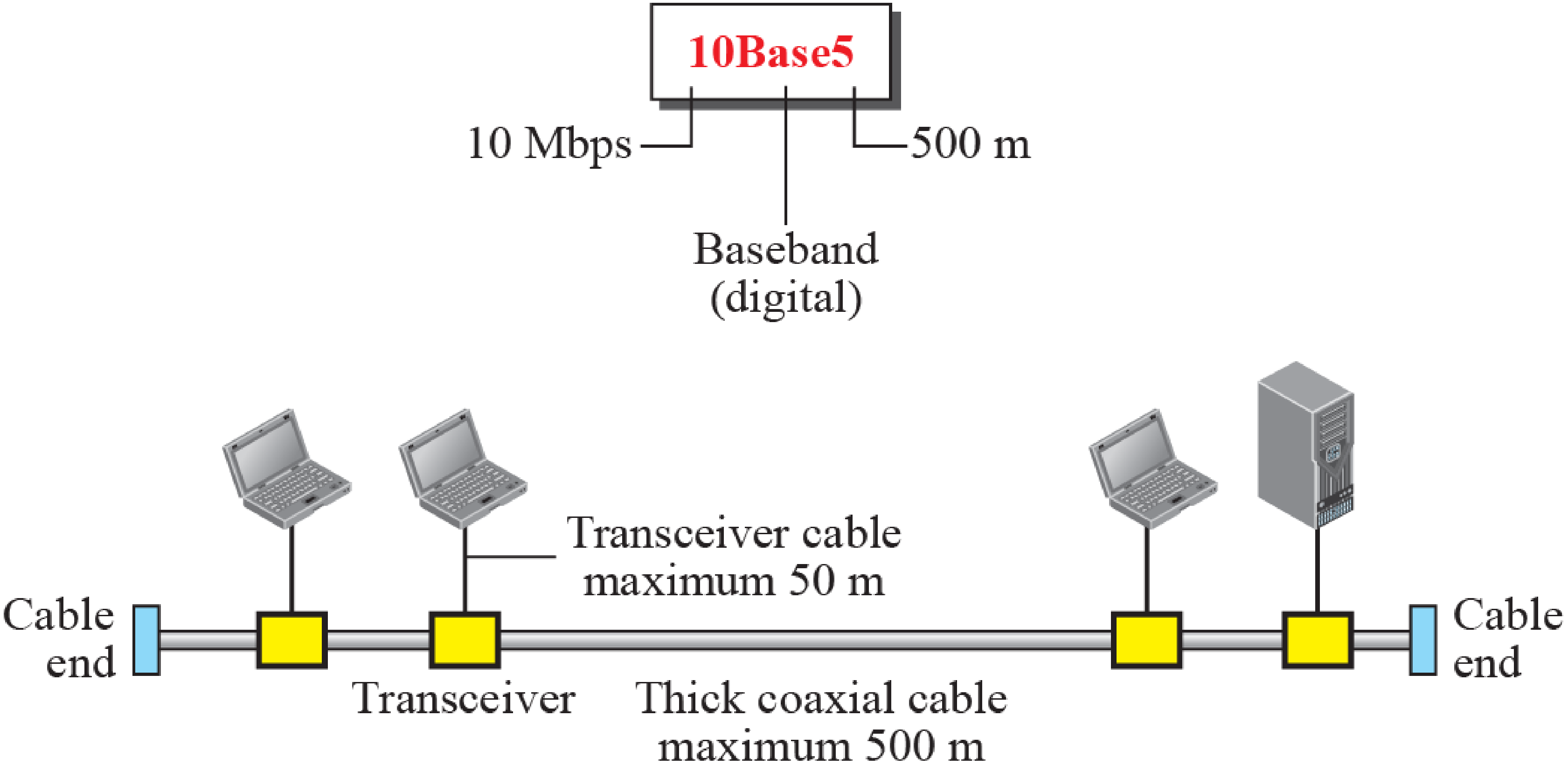
Encoding in a Standard Ethernet

- All standard implementations use digital signaling at 10 Mbps

- At the sender site, data is converted to a digital signal using Manchester scheme and at the receiver site, received signal is decoded into data

- Manchester encoding is self synchronous, providing a transition at each bit interval

10Base5 implementation



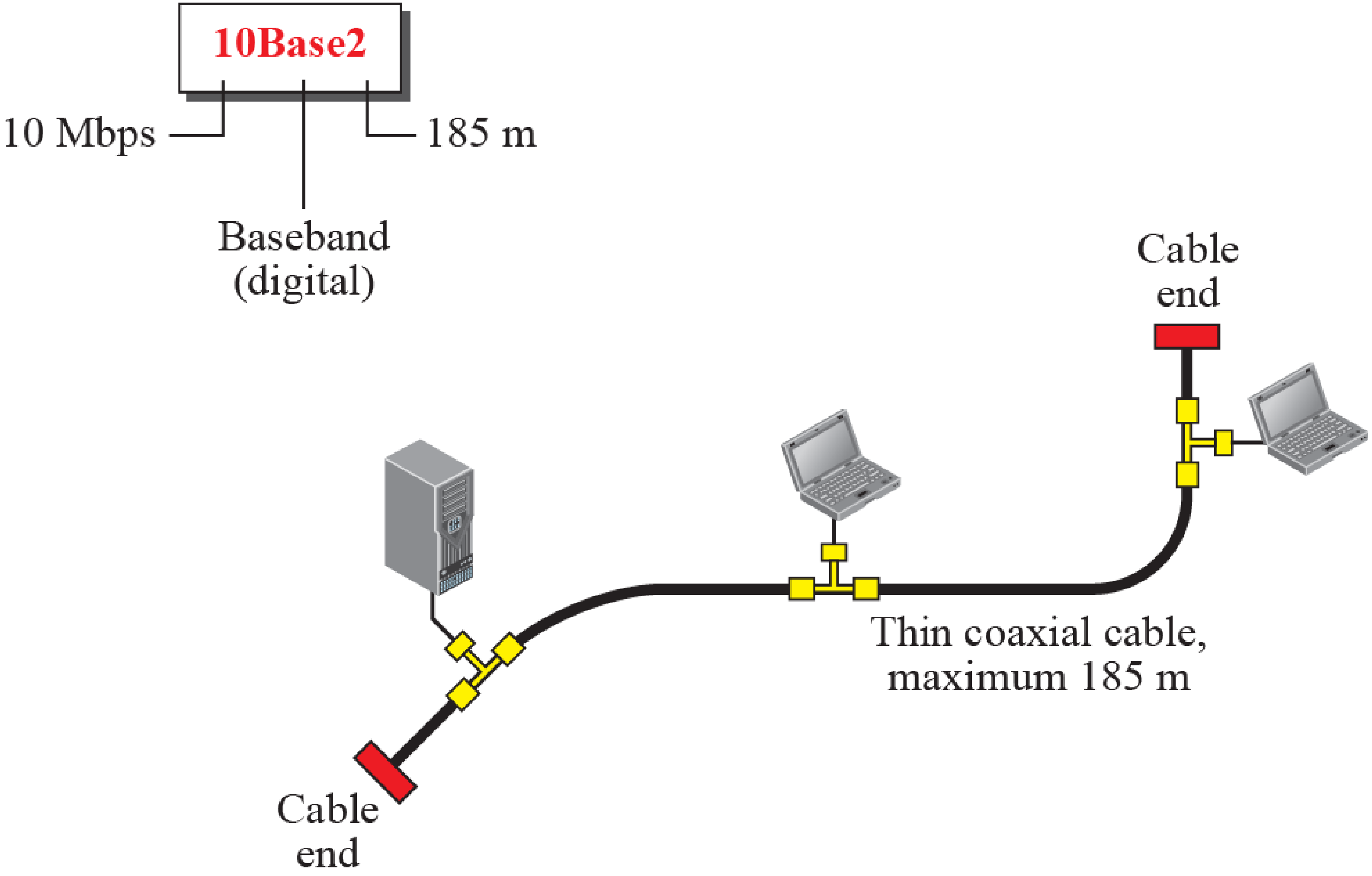
10Base5 implementation

- It is called 10Base5, thick Ethernet or Thick net
- Nick name derives from the size of the cable
- First Ethernet specification to use bus topology with an external trans receiver connected via tap to a thick coaxial cable
- Trans receiver is responsible for transmitting, receiving & detecting collision

❑ Trans receiver is connected to the station via a trans receiver cable that provides separate paths for sending and receiving

❑ Maximum length of co axial cable should not exceed 500 mts. If greater than 500mts, there will be excessive degradation of the signal. If length > 500mts is needed, five segments with a maximum of 500 mts can be connected using repeaters

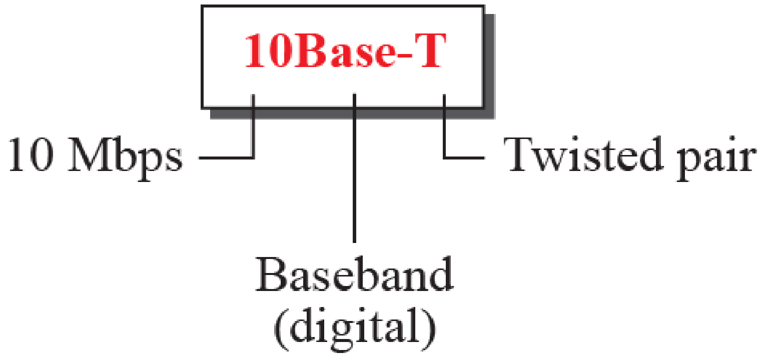
10Base2 implementation

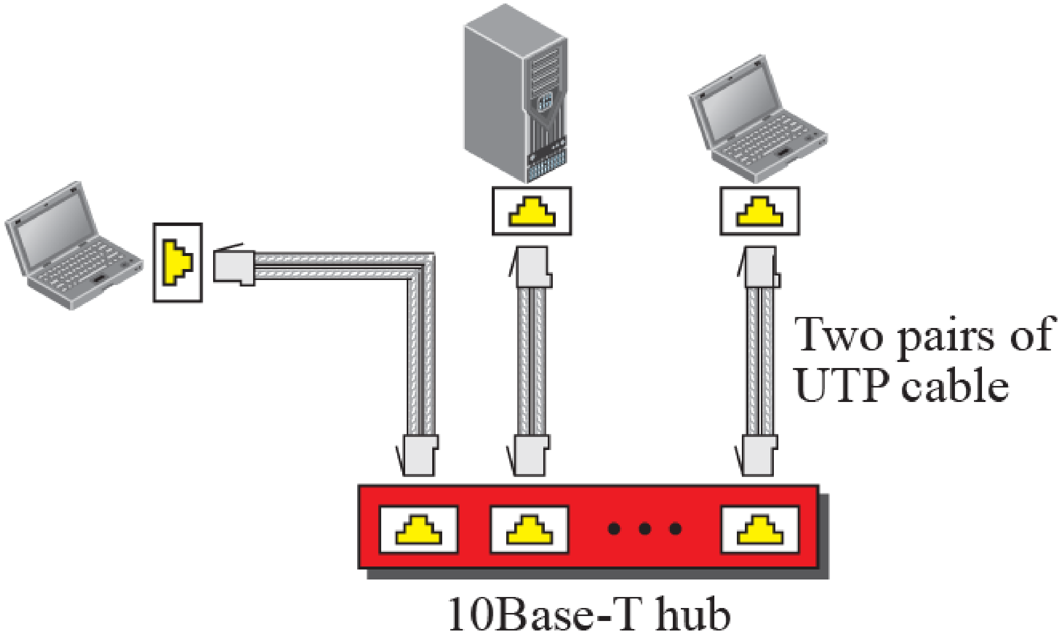


10Base2-Thin Ethernet

- It is called 10Base5, thin Ethernet or Cheap net
- Nick name derives from the size of the cable
- Uses bus topology , but the cable is much thinner and more flexible
- The cable can be bent to pass very close to stations
- Trans receiver is part of the NIC which is installed inside the station
- As Thin co axial cable is very flexible, installation is flexible.
- Length of segment can not exceed 185 mts due to high level of attention in thin coaxial cable

10Base-T implementation

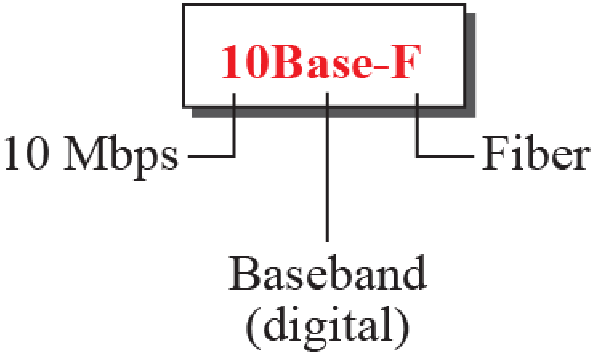




10Base-T Ethernet

- It is called 10BaseT, twisted pair Ethernet.
- Uses Star topology.
- Stations are connected to a hub via two pair of twisted cables.
- Two pairs creates two paths between station and the hub.
- Any collision happens in the hub.
- Maximum length of twisted pair cable here is defined as 10mts to minimize the effect of attenuation in the cable.

10Base-F implementation

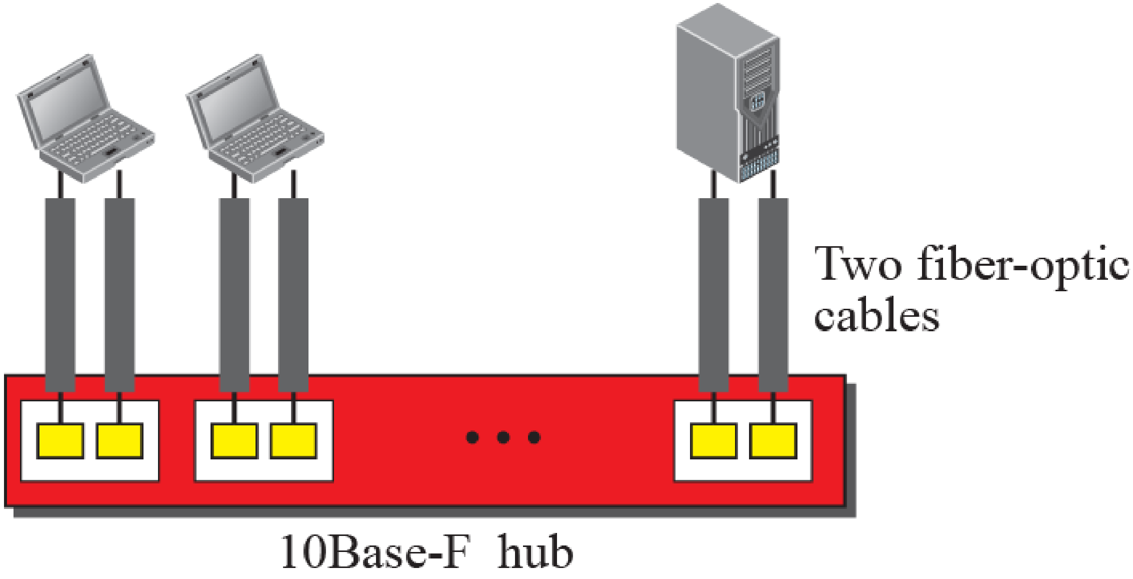


10Base-F

10 Mbps

Fiber

Baseband (digital)



Two fiber-optic cables

10Base-F hub

10Base-F Ethernet

- It is called 10Base-F, Optical fiber 10Mbps Ethernet.
- Uses Star topology.
- Stations are connected to a hub via two pair of fiber-optic cables.
- Two pairs creates two paths between station and the hub.

Changes in the Standard

- ➢Discuss the changes that occurred to the 10-Mbps Standard Ethernet.

- ➢These changes actually opened the road to the evolution of the Ethernet to become compatible with other high-data-rate LANs.

Ethernet Evolution:

1) Division of LAN by bridges:

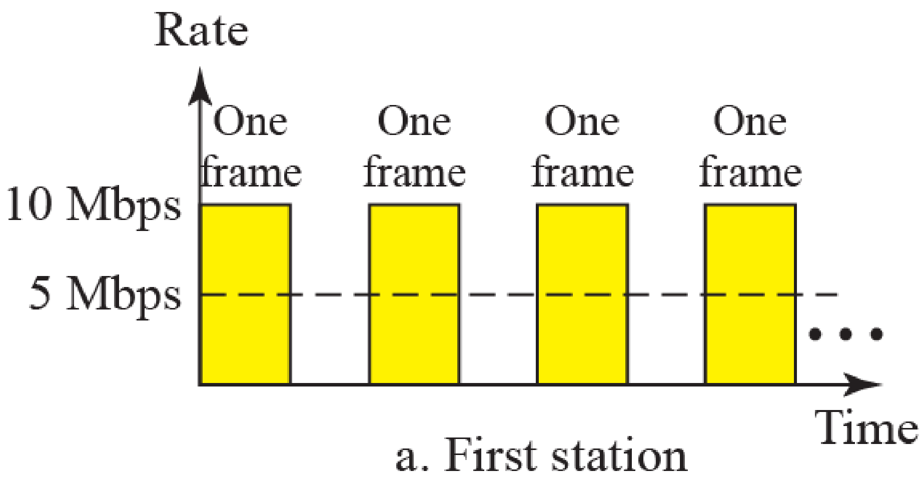
- ➢ Bridges has two Effects on an Ethernet LANs

  i) Raise the Bandwidth

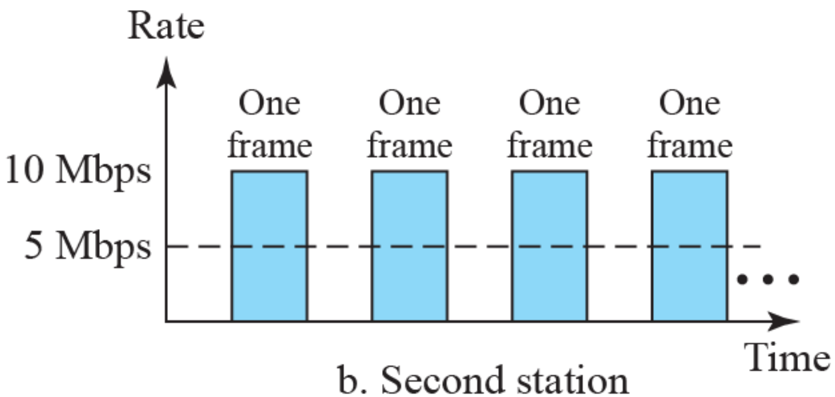
  ii) Separate Collision Domains.

- ➢ In an un bridged Ether N/w, total of 10 Mbps capacity is shared among all stations with a frame to send.

Sharing bandwidth



a. First station



b. Second station

Raising the Bandwidth

- The bridge divides the network into two or more networks. Bandwidth of each network is independent

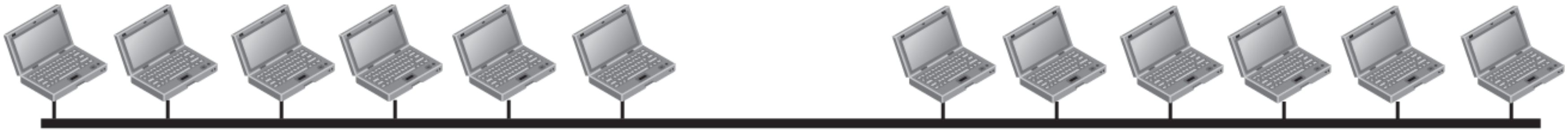
- Fig. below a network with 12 stations is divided into two networks each with 6 stations.

- Each network has a capacity of 10 Mbps which is shared between 6 stations and the bridge.

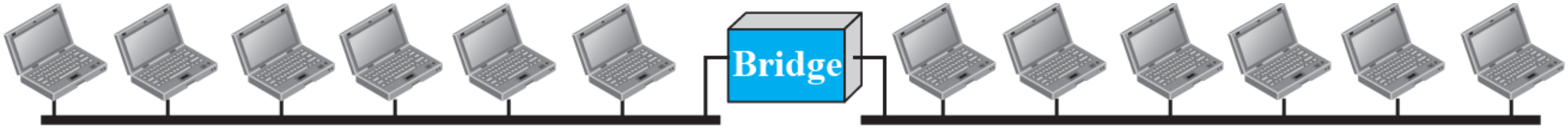
- In a network with a heavy load, each station theoretically is offered 10/7 Mbps instead of 10/12 devices.

- If four port bridge is used, each station is offered bandwidth of 10/4.

A network with and without bridging



a. Without bridging



b. With bridging

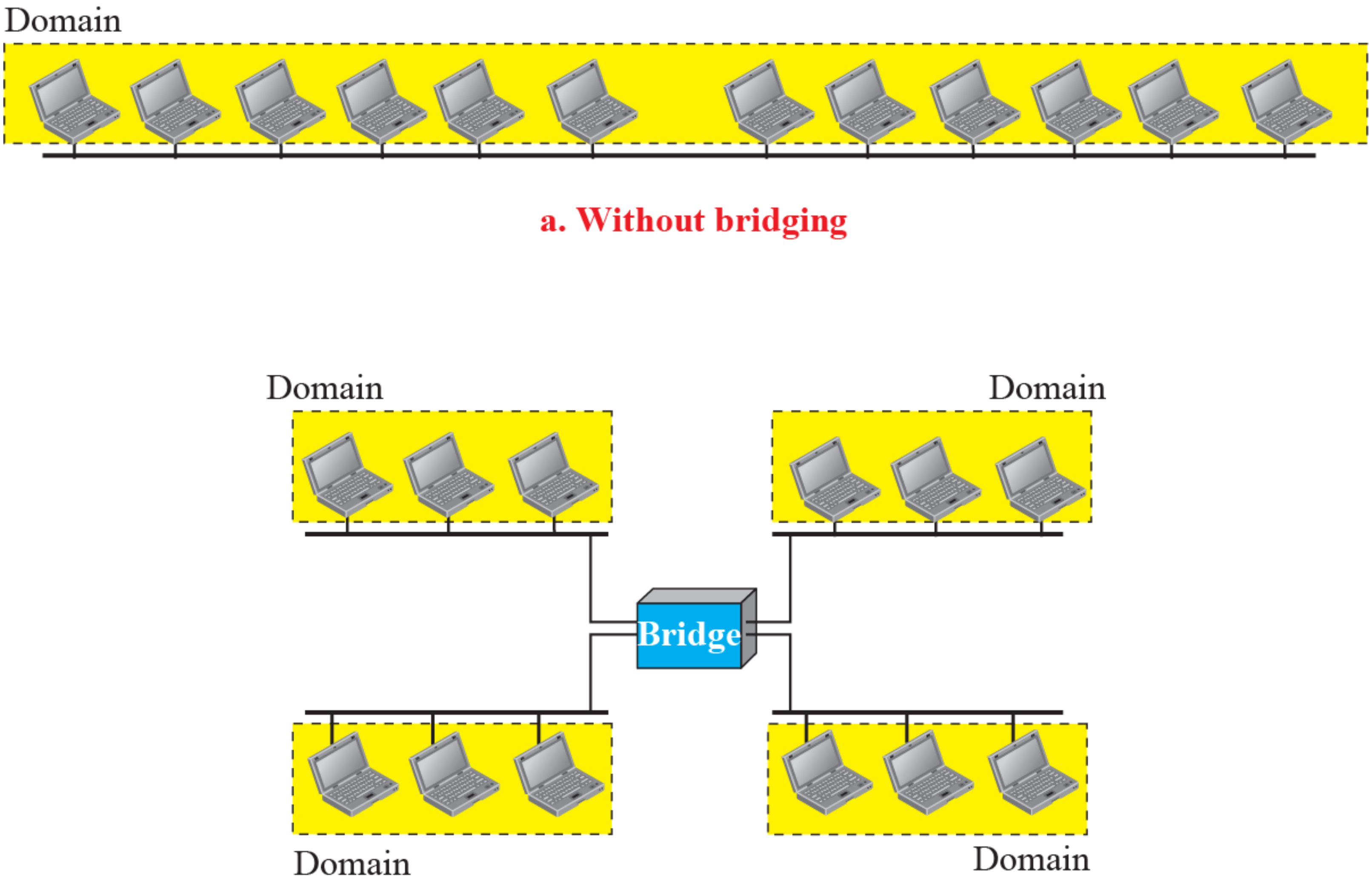
Separating collision domain

- Another advantage of bridge is the separation of collision domain

- From the figure below, it is observed that the probability of collision is reduced

- With bridging 3 stations contend for the access for the medium.

Collision domains



a. Without bridging

b. With bridging

Switched Ethernet

- The idea of bridged LAN can be extended to switched LAN

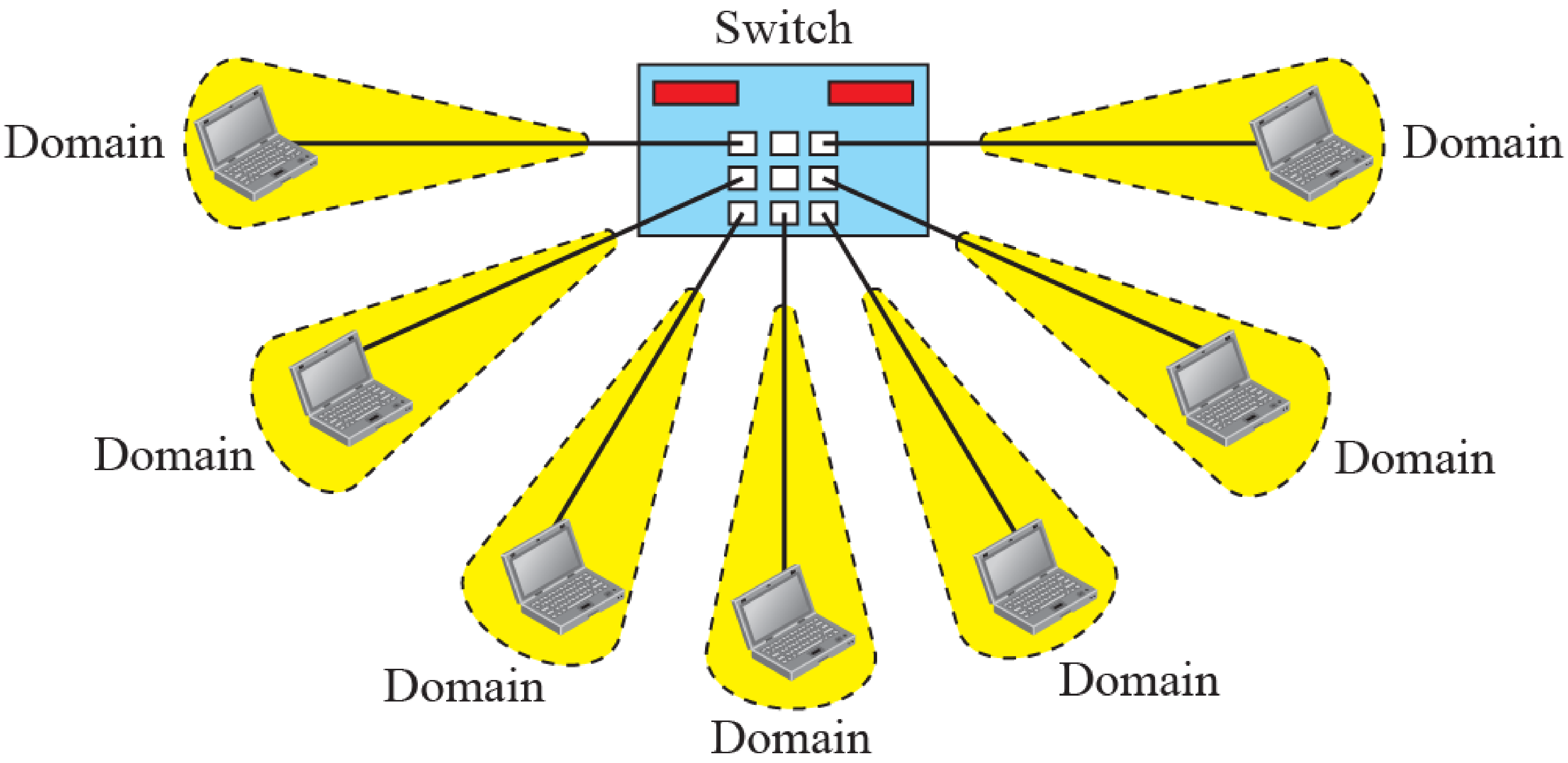
- In Switched LAN , instead of 4 networks, there will be N networks where N is the number of stations

- Here N port switch is used and the bandwidth is shared between the station and the switch

- Collision domain is divided in to N domains

- Evolution from a bridged Ethernet to switched Ethernet was a big step that opened the way to an even faster Ethernet

Switched Ethernet



Full duplex Ethernet

- One of the limitations of 10Base5 and 10Base2 is that communication is half duplex

- Full duplex mode increases the capacity of each domain from 10 to 20 Mbps

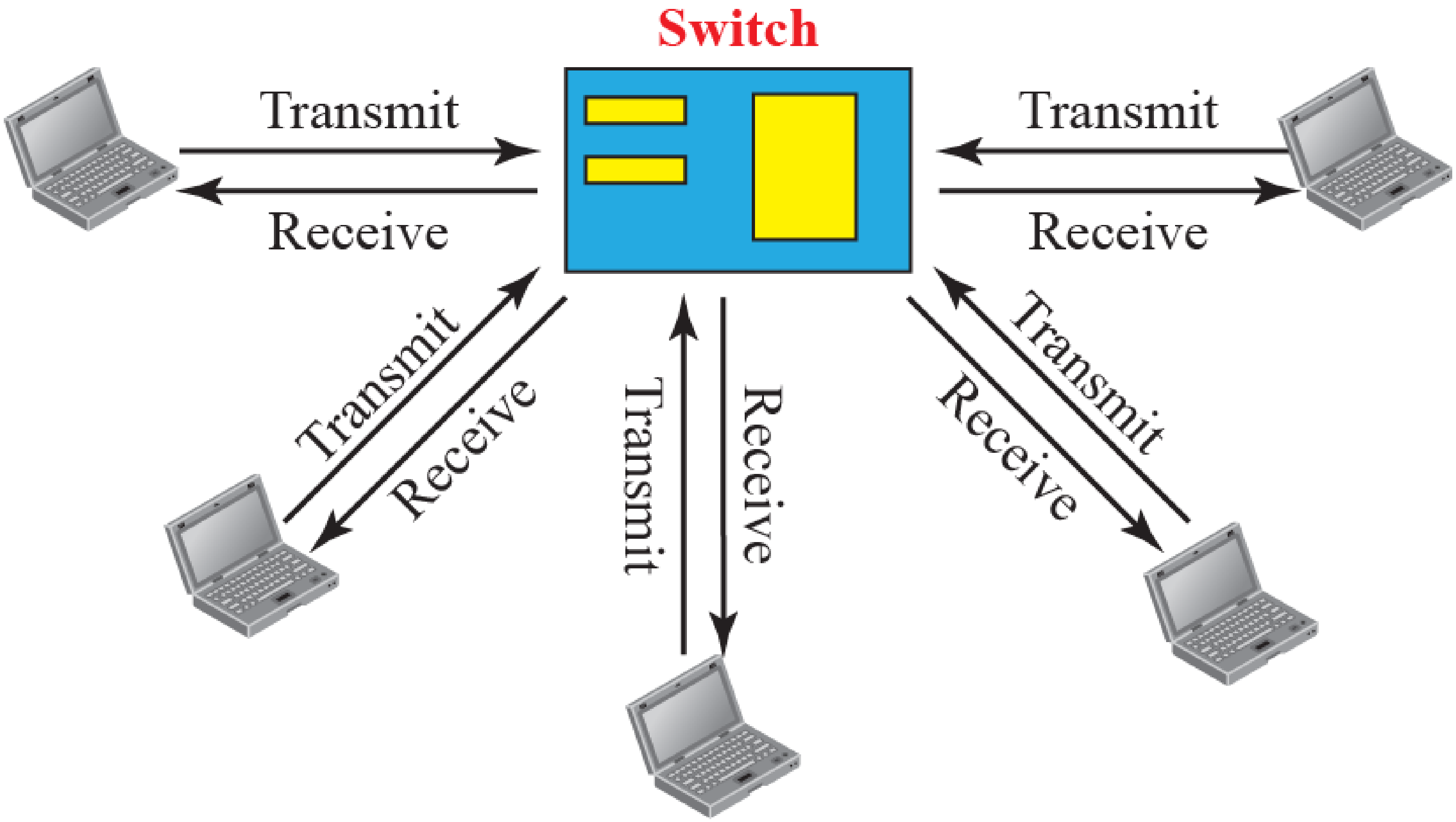
- In Full duplex switched Ethernet, each station is connected to a switch via two separate links, so there is no need for CSMA/CD

- Each link has a point-to-point dedicated path between the station and switch

- Job of MAC layer is easy as carrier sensing & collision detection functionalities can be turned off

- Std. Ethernet was designed as a connectionless protocol at MAC sub layer. To Provide FC & EC in full duplex switched Ethernet, a new sub layer, MAC control, is added between LLC & MAC

Full – duplex switched Ethernet



FAST ETHERNET(100 Mbps)

➢In the 1990s, Ethernet made a big jump by increasing the transmission rate to 100 Mbps, and the new generation was called the Fast Ethernet

➢The designers of the Fast Ethernet needed to make it compatible with the Standard Ethernet

➢ The MAC sub layer was left unchanged, means frame format, min & max size of frame remain unchanged

➢By increasing the transmission rate, the features of the Standard Ethernet that depend on the transmission rate, access method, implementation had to be changed.

Goals of Fast Ethernet

- Upgrade the data rate to 100 Mbps.
- Make it compatible with standard Ethernet.
- Keep the same 48 bit address.
- Keep the same frame format.

Access Method

- The proper operation of the CSMA/CD depends on the transmission rate, the minimum size of the frame, and the maximum network length.
- If we want to keep the minimum size of the frame, the maximum length of the network should be changed.
- In other words, if the minimum frame size is still 512 bits, and it is transmitted 10 times faster, the collision needs to be detected 10 times sooner, which means the maximum length of the network should be 10 times shorter (the propagation speed does not change).

Access Method

- The Fast Ethernet came with two solutions

i)Drop bus topology & use a passive hub or a star topology but make the maximum size of Ethernet 250 meters instead of 2500 meters.

ii) Use Link layer switch with a buffer to store frames and full duplex connection to each host to make the transmission medium private for each host.

No need of CSMA/CD.

Since the connection to switch is full duplex, the destination address can even send a frame to another station at the same time.

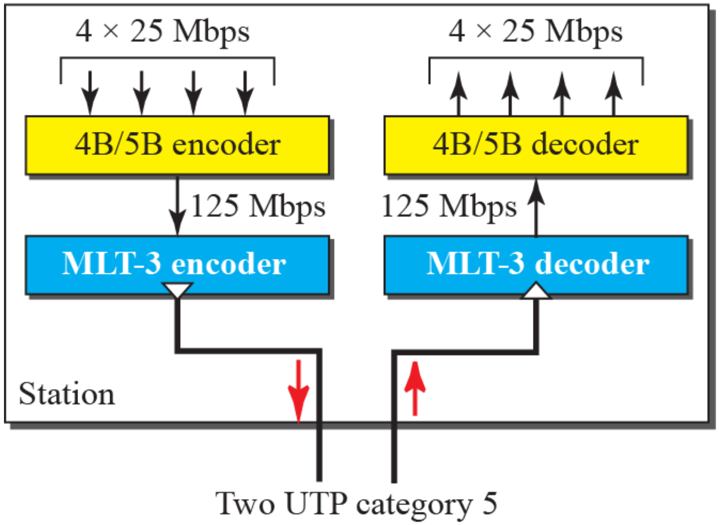
- New feature called Auto negotiation is added. Auto negotiation allows two devices to negotiate the mode or data rate of operation.

Physical Layer

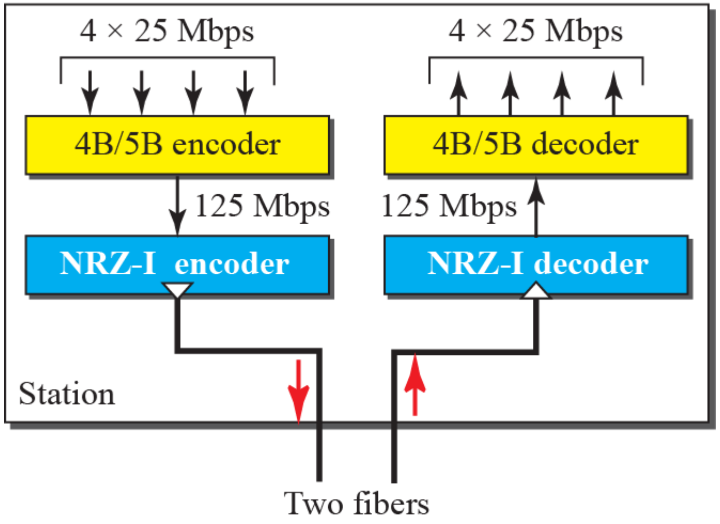
➢To be able to handle a 100 Mbps data rate, several changes need to be made at the physical layer.

Encoding for fast Ethernet

100Base-TX



100Base-FX



100Base-T4

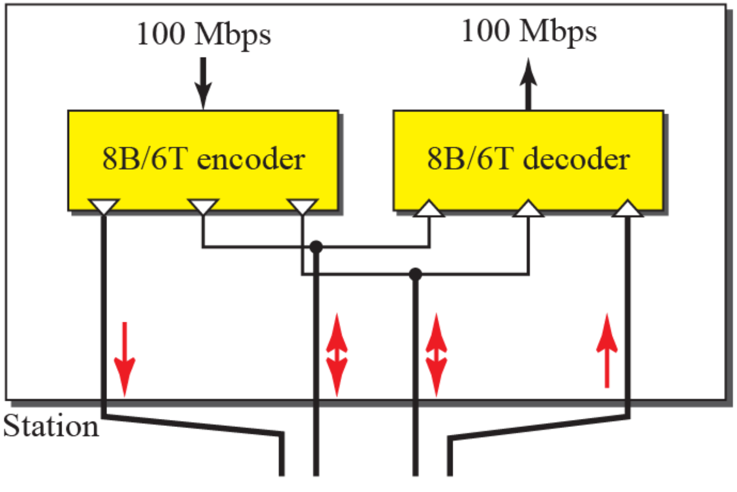
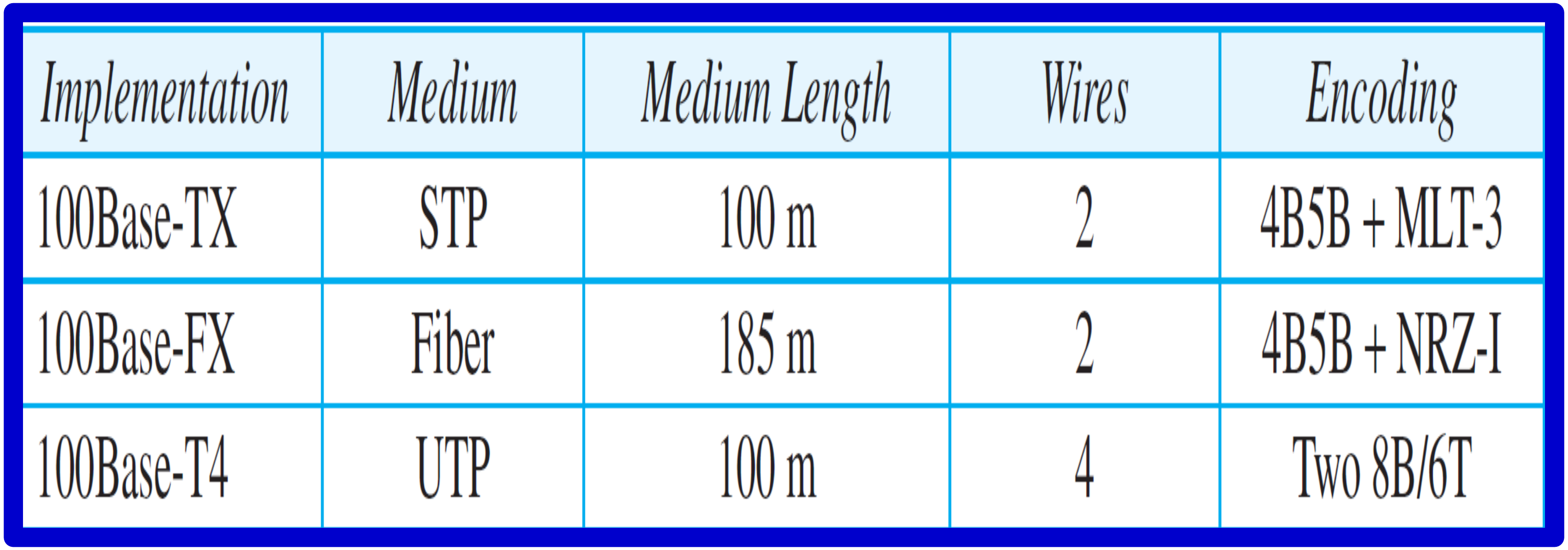


Table Summary of Fast Ethernet implementations



Implementation	*Medium*	*Medium Length*	*Wires*	*Encoding*
100Base-TX	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

GIGABIT ETHERNET

- The need for an even higher data rate resulted in the design of the Gigabit Ethernet Protocol (1000 Mbps).
- The IEEE committee calls it the Standard 802.3z.
- The goals of the Gigabit Ethernet were to upgrade the data rate to 1 Gbps, but keep the address length, the frame format, and the maximum and minimum frame length the same.

MAC Sublayer

- A main consideration in the evolution of Ethernet was to keep the MAC sub layer untouched

- However, to achieve a data rate of 1 Gbps, this was no longer possible

- 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.

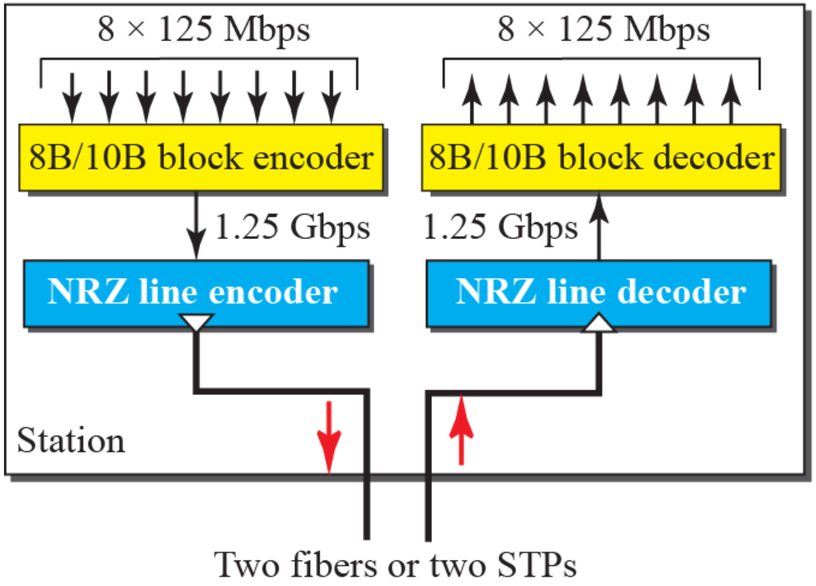
Physical Layer

- The physical layer in Gigabit Ethernet is more complicated than that in Standard or Fast Ethernet

- Discuss some features of this layer.

Encoding in Gigabit Ethernet

1000Base-SX, 1000Base-LX, and 1000Base-CX



1000Base-T

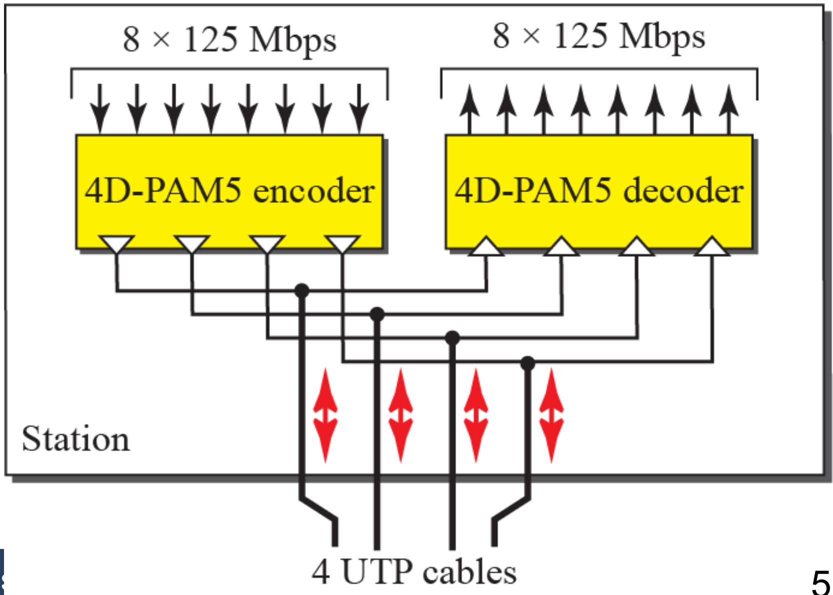
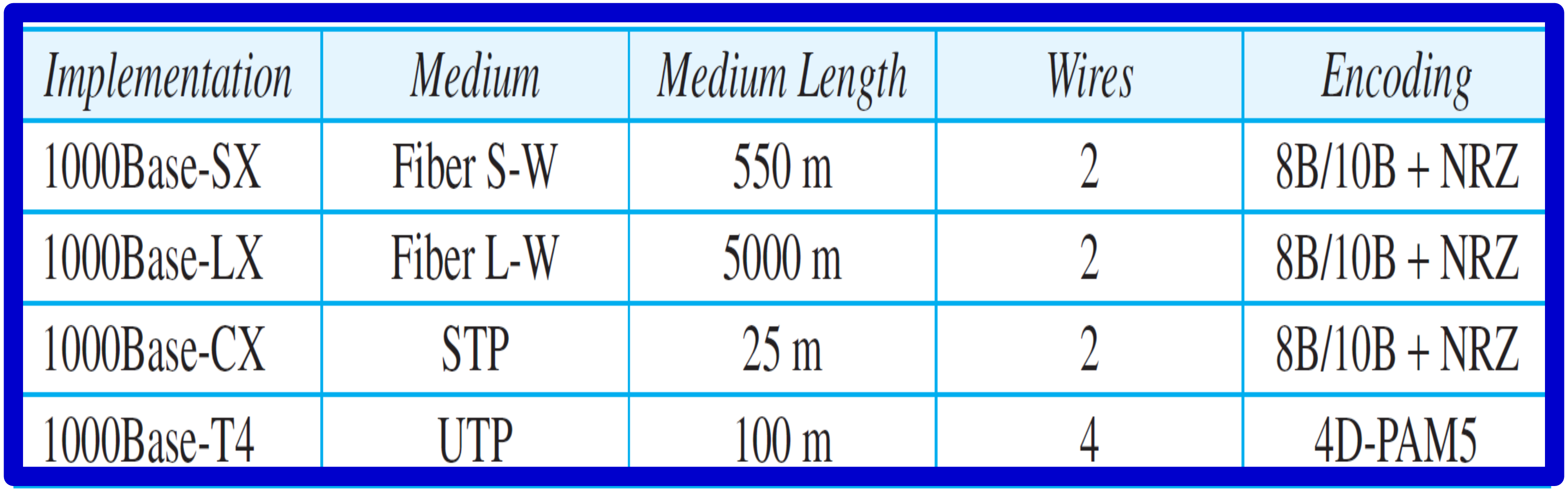


Table : Summary of Gigabit Ethernet implementations



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

10-GIGABIT EHTERNET

- In recent years, there has been another look into the Ethernet for use in metropolitan areas

- The idea is to extend the technology, the data rate, and the coverage distance so that the Ethernet can be used as LAN and MAN (metropolitan area network)

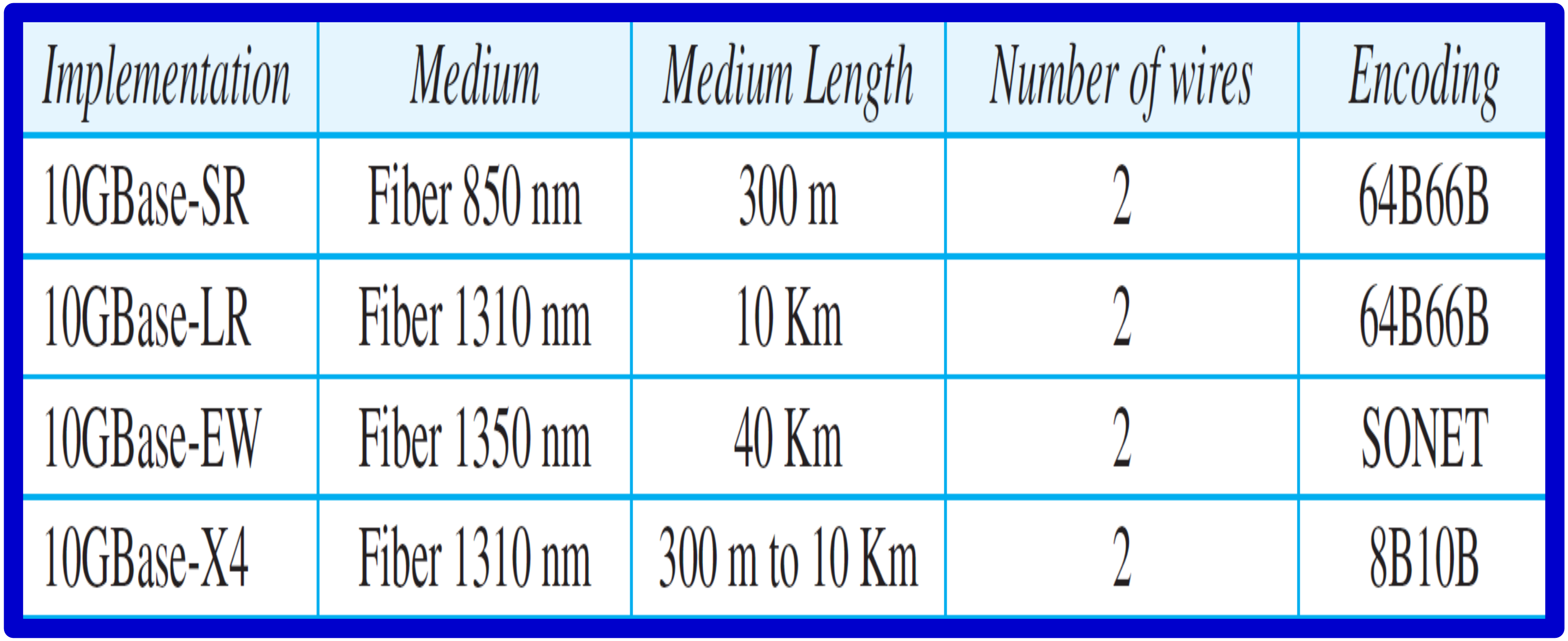
- The IEEE committee created 10 Gigabit Ethernet and called it Standard 802.3ae

Implementation

- 10 Gigabit Ethernet operates only in full-duplex mode, which means there is no need for contention; CSMA/CD is not used in 10 Gigabit Ethernet

Four implementations are the most common: 10GBase-SR, 10GBase-LR, 10GBase-EW, and 10GBase-X4.

Table 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

Cellular Telephony

- Cellular telephony is designed to provide communications between two moving units, called **mobile stations (MSs)** or between one mobile unit and one stationary unit, often called a **land unit**

- A service provider must be able to locate and track a caller, assign a channel to the call, and transfer the channel from base station to base station as the caller moves out of range.

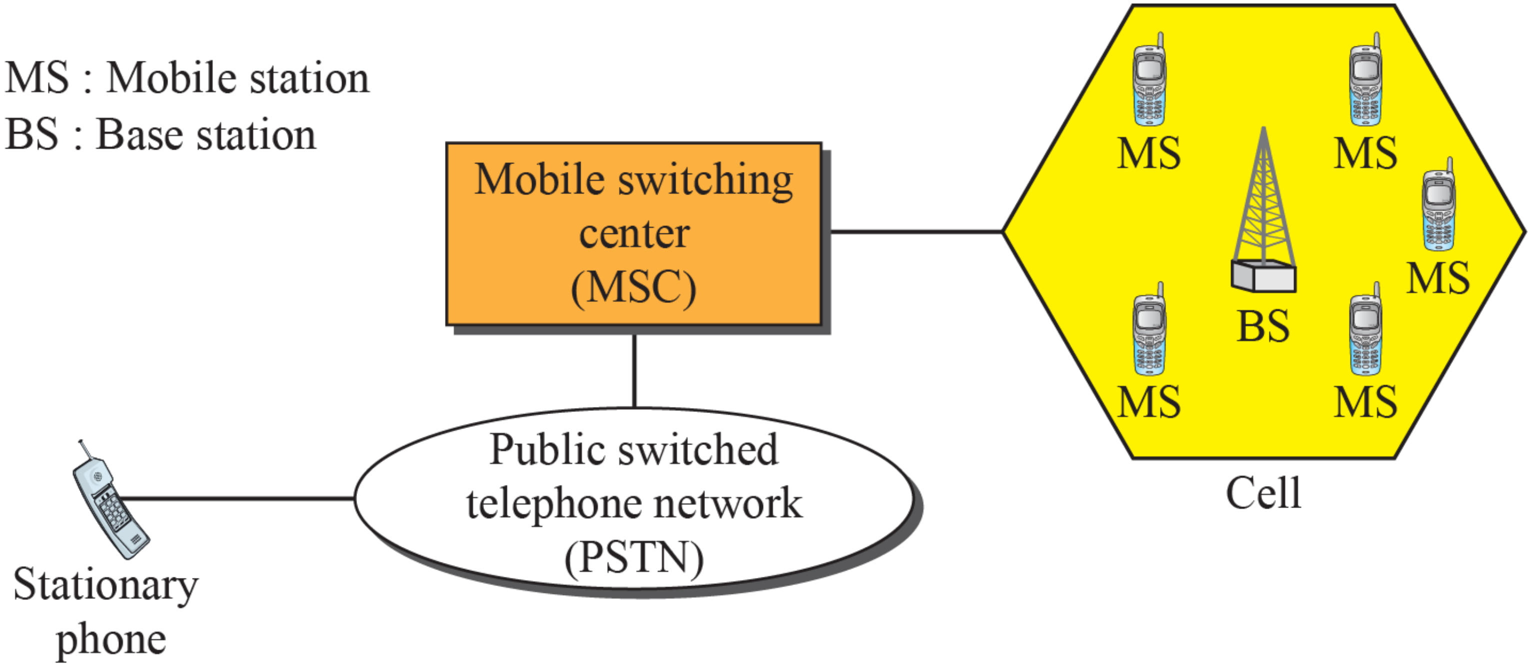
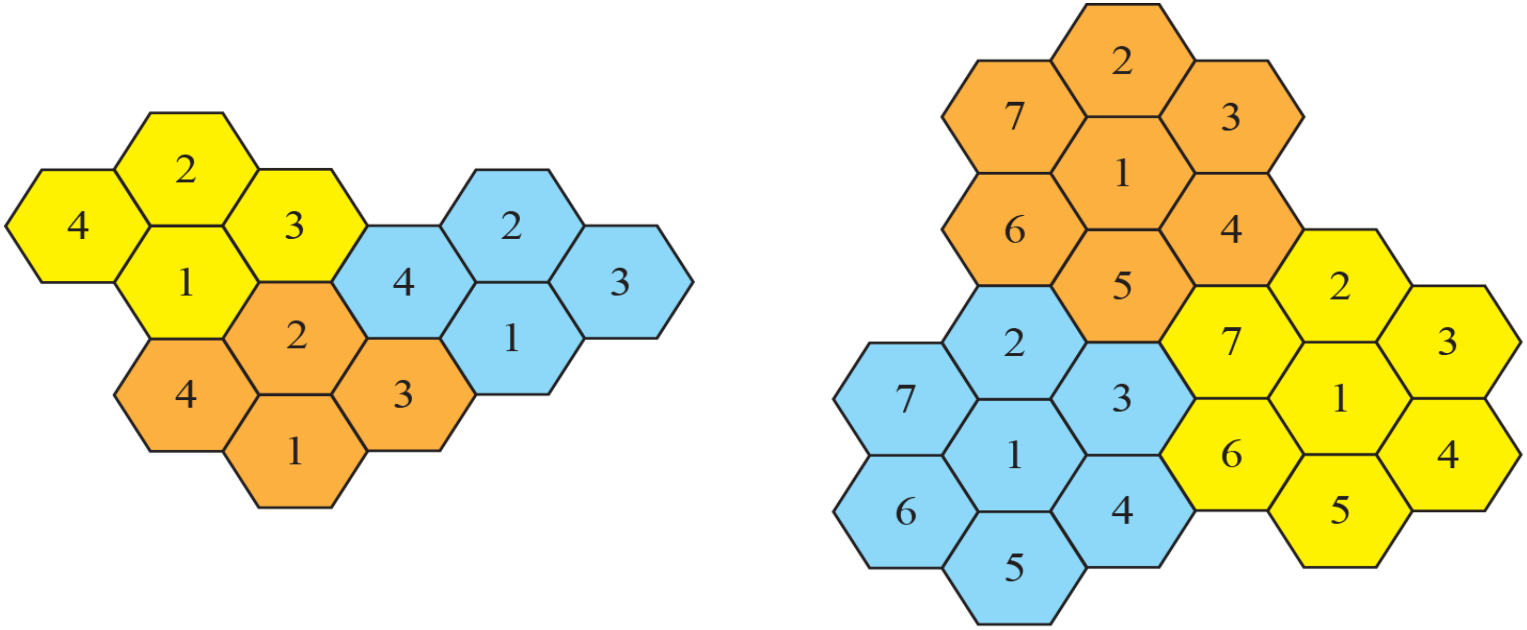

MS : Mobile station
BS : Base station
Mobile switching center (MSC)
Public switched telephone network (PSTN)
Stationary phone
MS
MS
MS
MS
MS
BS
Cell

Figure 16.6: Cellular system

Operation

Let us first briefly discuss the operation of the cellular telephony



a. Reuse factor of 4 b. Reuse factor of 7

Figure 16.7: Frequency reuse patterns

Frequency-Reuse Principle

- In general, neighboring cells cannot use the same set of frequencies for communication because doing so may create interference for the users located near the cell boundaries

- The set of frequencies available is limited, and frequencies need to be reuse

- A frequency reuse pattern is a configuration of N cells, N being the reuse factor, in which 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 shows two of them.

Frequency-Reuse Principle (Contd)

- The numbers in the cells define the pattern. The cells with the same number in a pattern can use the same set of frequencies. We call these cells the reusing cells

- As Figure 16.7 shows, in a pattern with reuse factor 4, only one cell separates the cells using the same set of frequencies

- In a pattern with reuse factor 7, two cells separate the reusing cells.

Transmitting

- To place a call from a mobile station, the caller enters a code of 7 or 10 digits (a phone number) and presses the send button

- The mobile station then scans the band, seeking a setup channel with a strong signal, and sends the data (phone number) to the closest base station using that channel. The base station relays the data to the MSC

- The MSC sends the data on to the telephone central office. If the called party is available, a connection is made and the result is relayed back to the MSC

- At this point, the MSC assigns an unused voice channel to the call, and a connection is established. The mobile station automatically adjusts its tuning to the new channel, and communication can begin.

Receiving

- When a mobile phone is called, the telephone central office sends the number to the MSC

- The MSC searches for the location of the mobile station by sending query signals to each cell in a process called paging

- Once the mobile station is found, the MSC transmits a ringing signal

- When the mobile station answers, assigns a voice channel to the call, allowing voice communication to begin.

Handoff

- It may happen that, during a conversation, the mobile station moves from one cell to another. The signal may become weak

- To solve this problem, the MSC monitors the level of the signal every few seconds. If the strength of the signal diminishes, the MSC seeks a new cell that can better accommodate the communication

- The MSC then changes the channel carrying the call (hands the signal off from the old channel to a new one).

Types of handoffs

- Hard Handoff

- Soft Handoff

Hard Handoff

- Early systems used a hard handoff

- In a hard handoff, a mobile station only communicates with one base station

- When the mobile station moves from one cell to another, communication must first be broken with the previous base station before communication can be established with the new one

- This may create a rough transition

Soft Handoff

- New systems use a soft handoff

- In this case, a mobile station can communicate with two base stations at the same time

- During handoff, a mobile station may continue with the new base station before breaking off from the old one

Roaming

- One feature of cellular telephony is called roaming

- Roaming means, in principle, that a user can have access to communication or can be reached where there is coverage. A service provider usually has limited coverage

- Neighboring service providers can provide extended coverage through a roaming contract

First Generation (1G)

- Cellular telephony is now in its fourth generation

- The first generation was designed for voice communication using analog signals

- We discuss one first-generation mobile system used in North America, AMPS.

AMPS

- Advanced Mobile Phone System (AMPS) is one of the leading analog cellular systems in North America
- It uses FDMA

Bands

AMPS operates in the ISM 800-MHz band

The system uses two separate analog channels, one for forward (base station to mobile station) communication and one for reverse (mobile station to base station) communication

The band between 824 and 849 MHz carries reverse communication; the band between 869 and 894 MHz carries forward communication

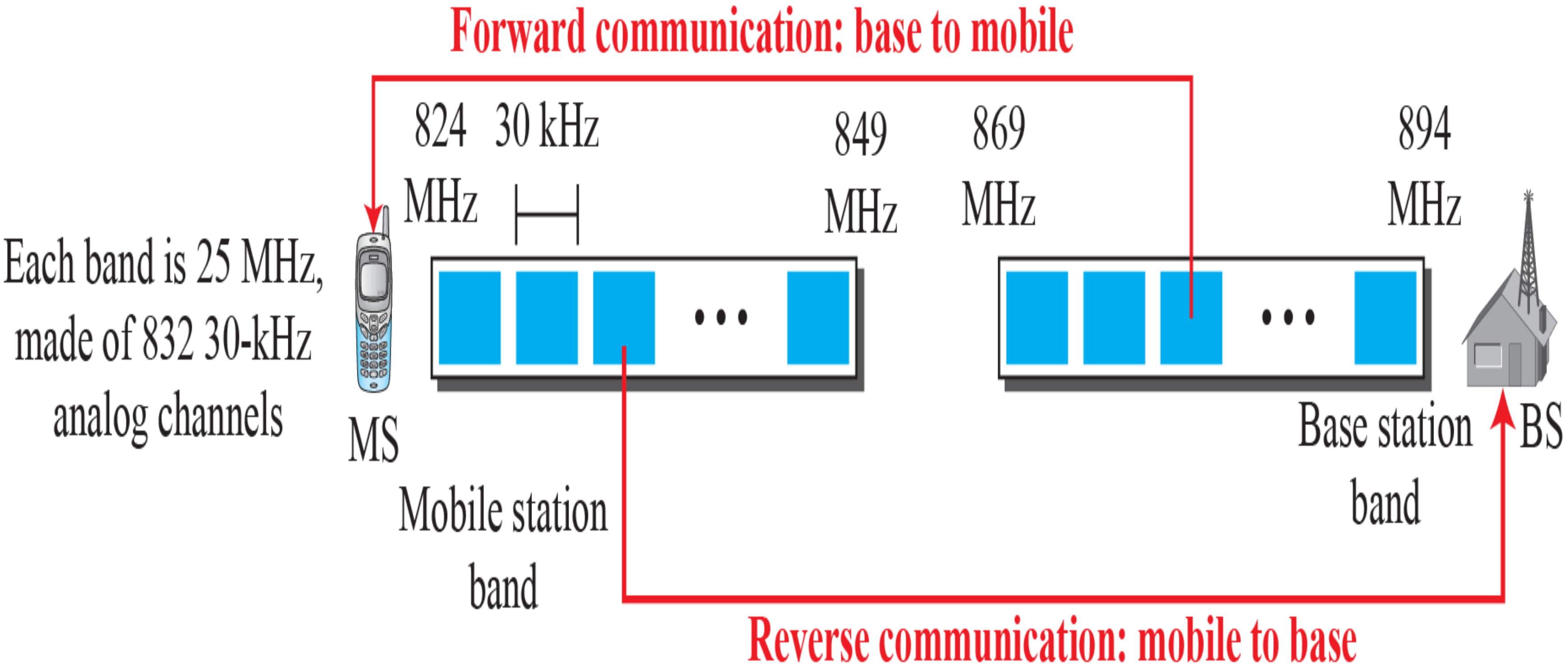

Forward communication: base to mobile
824 MHz
30 kHz
849 MHz
869 MHz
894 MHz
Each band is 25 MHz, made of 832 30-kHz analog channels
• • •
• • •
MS
Mobile station band
Base station band
BS
Reverse communication: mobile to base

Figure 16.8: Cellular bands for AMPS

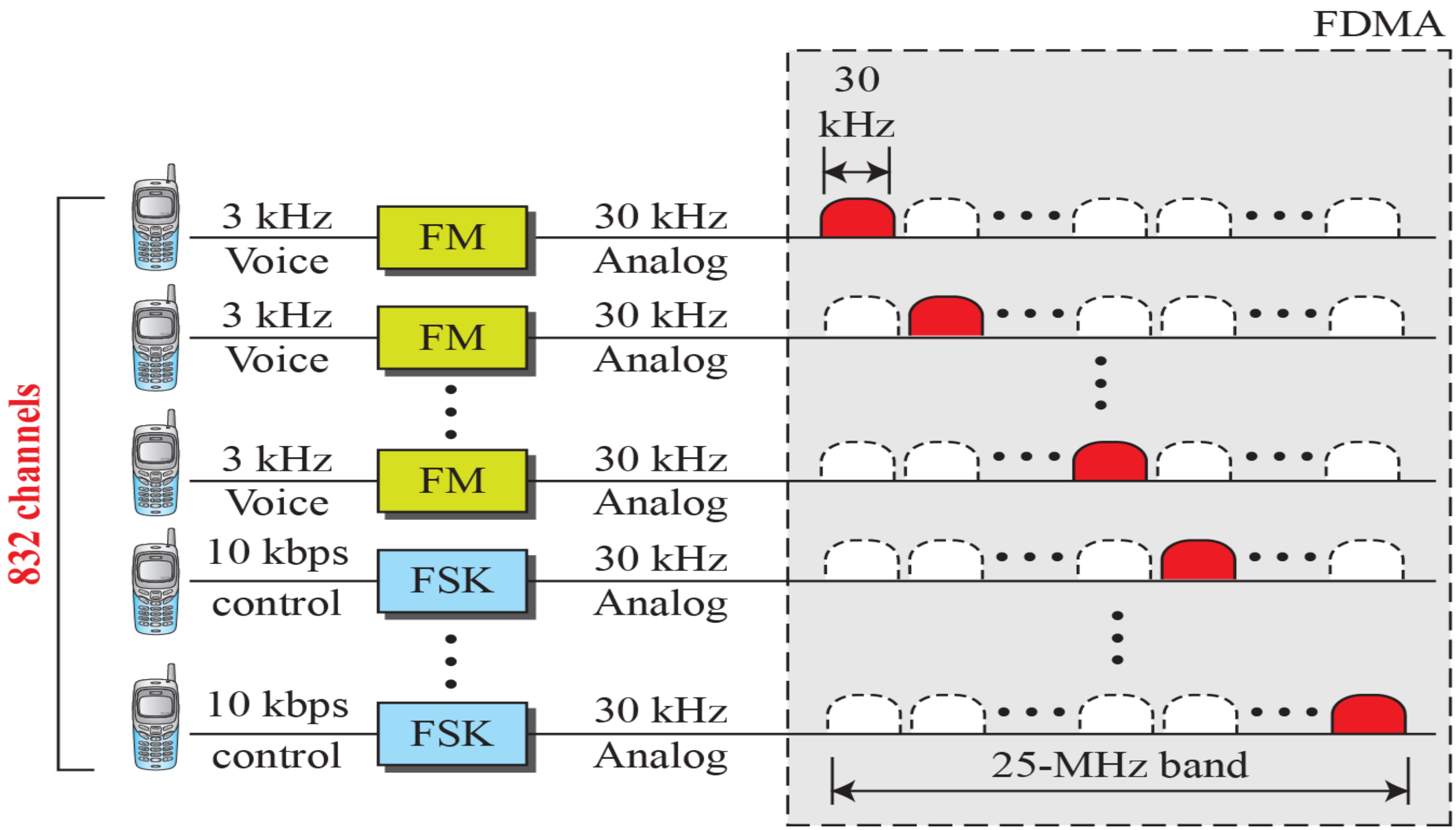


Figure 616.9: AMPS reverse communication band

Second Generation (2G)

- To provide higher-quality (less noise-prone) mobile voice communications, the second generation of the cellular phone network was developed

- The second generation was mainly designed for digitized voice

- Three major systems evolved in the second generation: D-AMPS, GSM, and IS-95

D-AMPS

D-AMPS was designed to be backward-compatible with AMPS. This means that in a cell, one telephone can use AMPS and another D-AMPS

D-AMPS was first defined by IS-54 (Interim Standard 54) and later revised by IS-136

Band

D-AMPS uses the same bands and channels as AMPS

Transmission

Each voice channel is digitized using a very complex PCM and compression technique. A voice channel is digitized to 7.95 kbps. Three 7.95-kbps digital voice channels are combined using TDMA. The result is 48.6 kbps of digital data

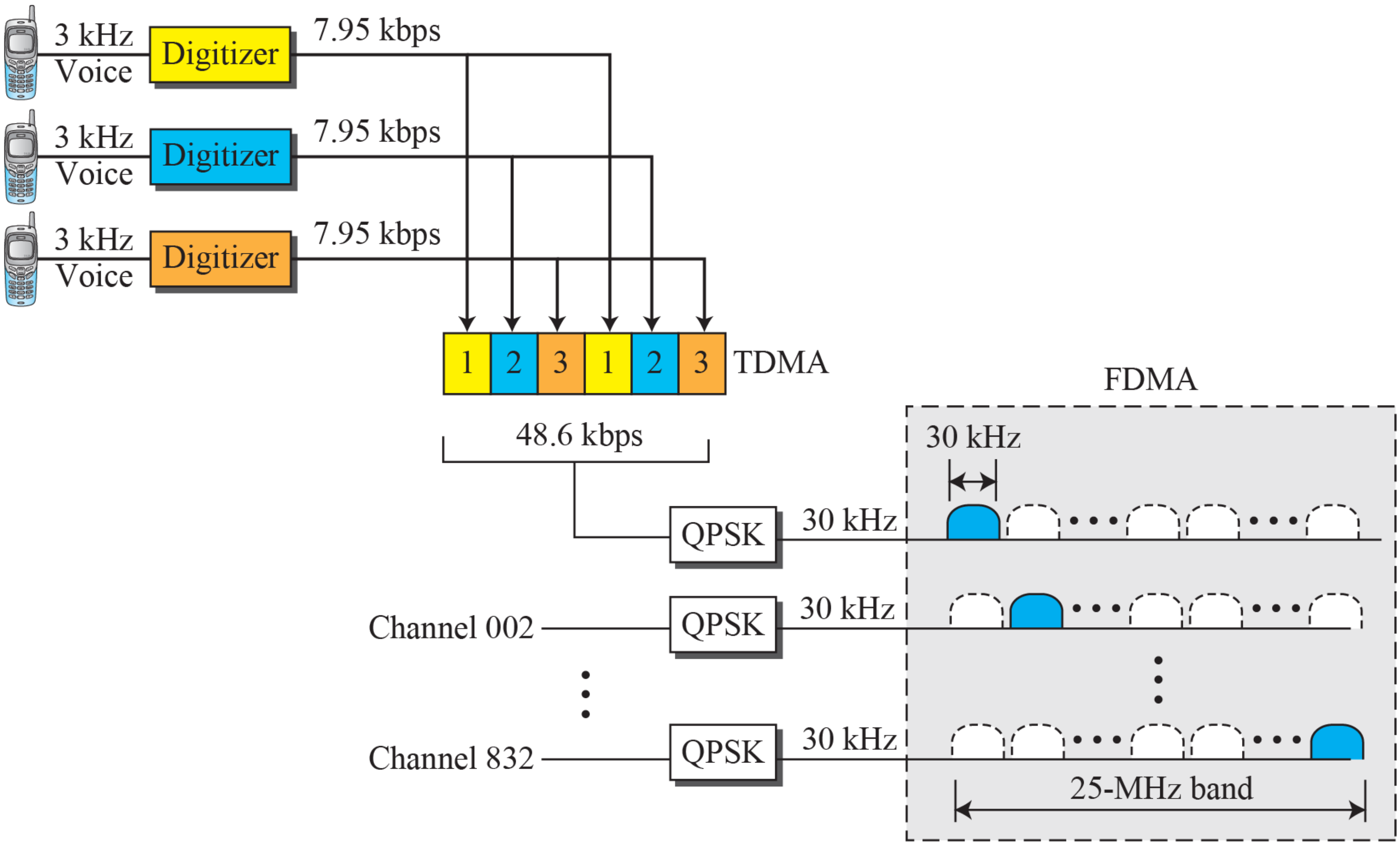


Figure 16.10: D-AMPS

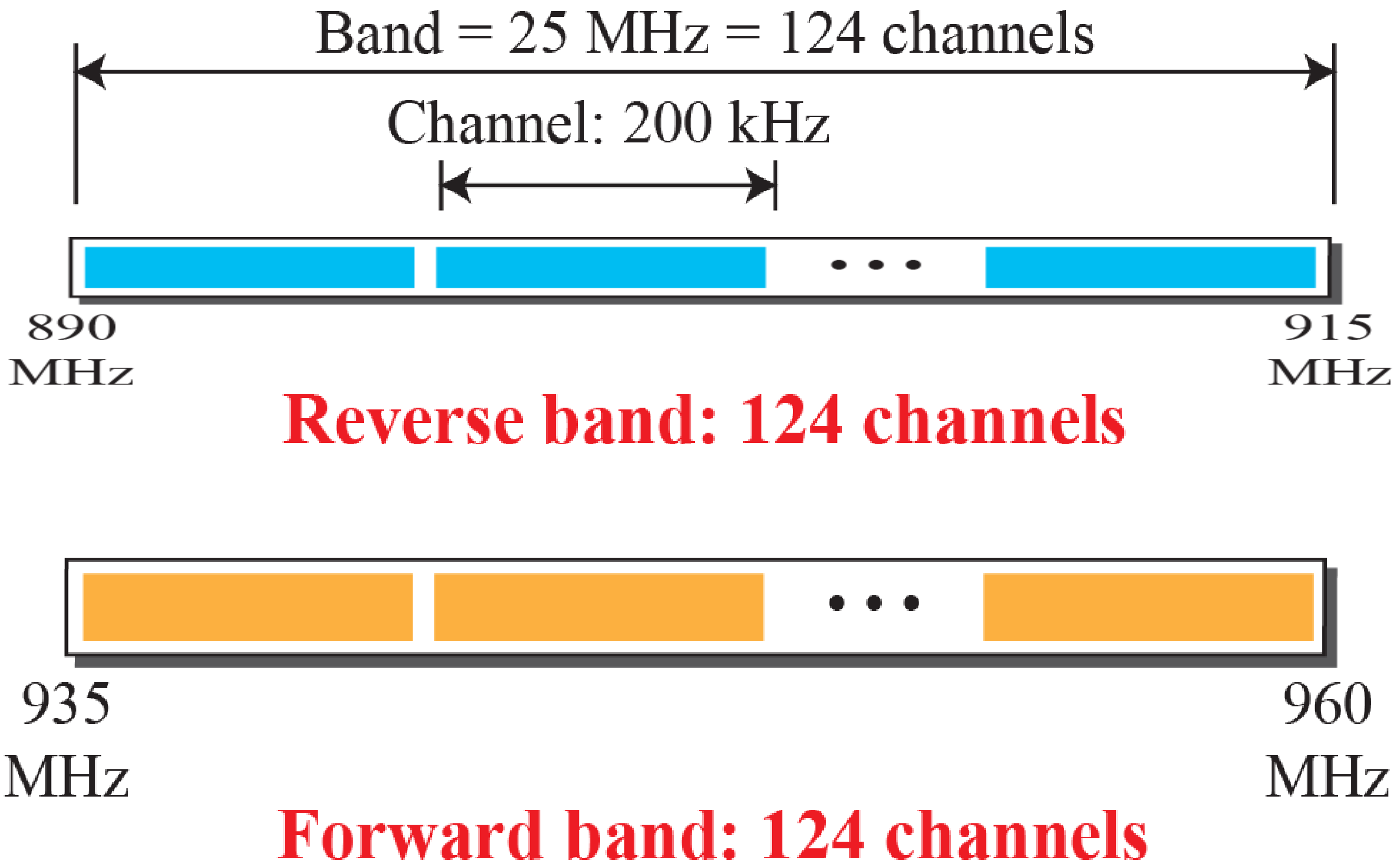

Band = 25 MHz = 124 channels
Channel: 200 kHz
890 MHz
915 MHz
Reverse band: 124 channels
935 MHz
960 MHz
Forward band: 124 channels

Figure 616.11: GSM bands

GSM

- The Global System for Mobile Communication (GSM) is a European standard that was developed to provide a common second-generation technology for all Europe
- The aim was to replace a number of incompatible first-generation technologies

Bands

- GSM uses two bands for duplex communication. Each band is 25 MHz in width, shifted toward 900 MHz
- Each band is divided into 124 channels of 200 kHz separated by guard bands

Transmission

- Each voice channel is digitized and compressed to a 13-kbps digital signal. Each slot carries 156.25 bits. Eight slots share a frame (TDMA).

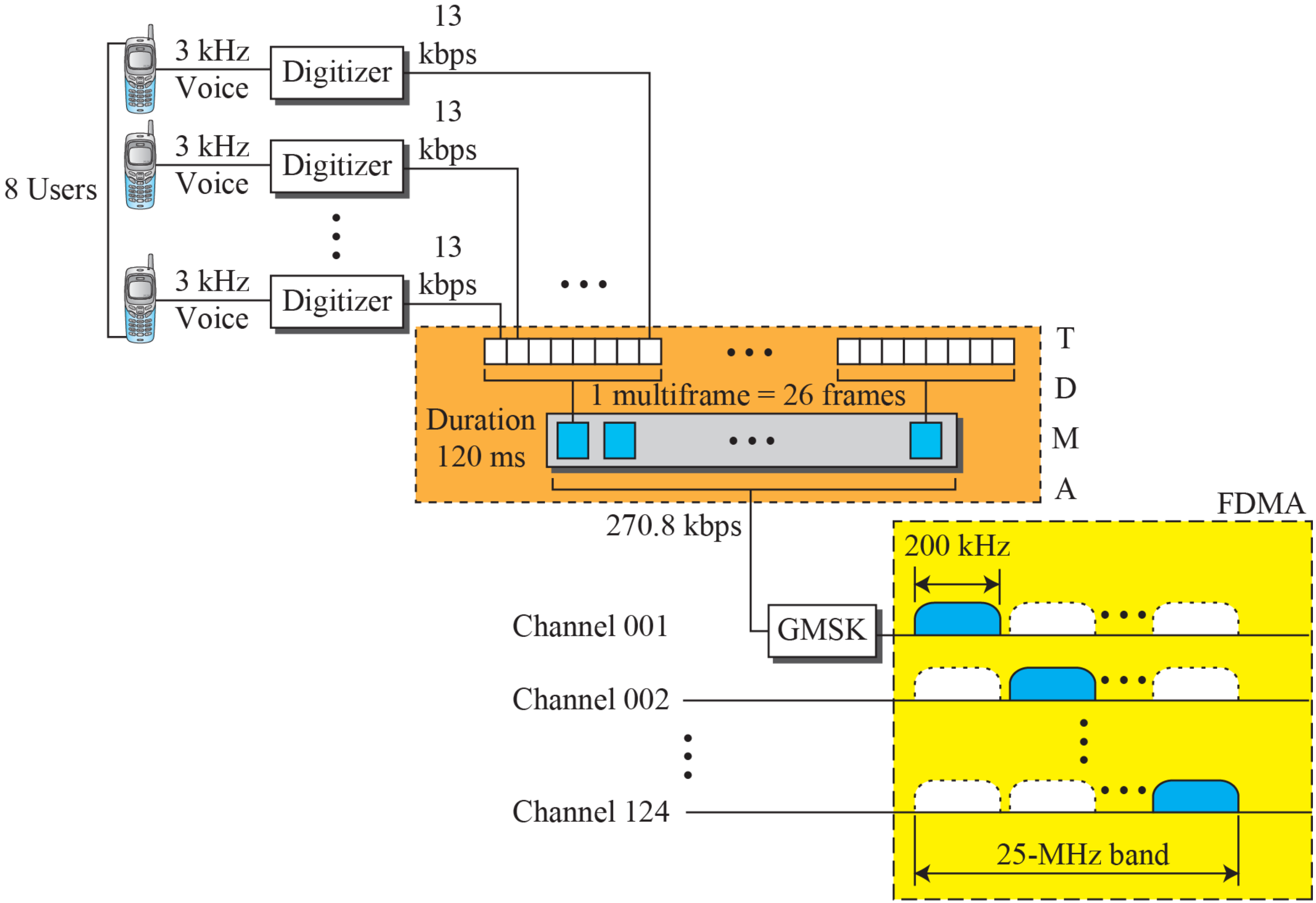


Figure 16.12: GSM

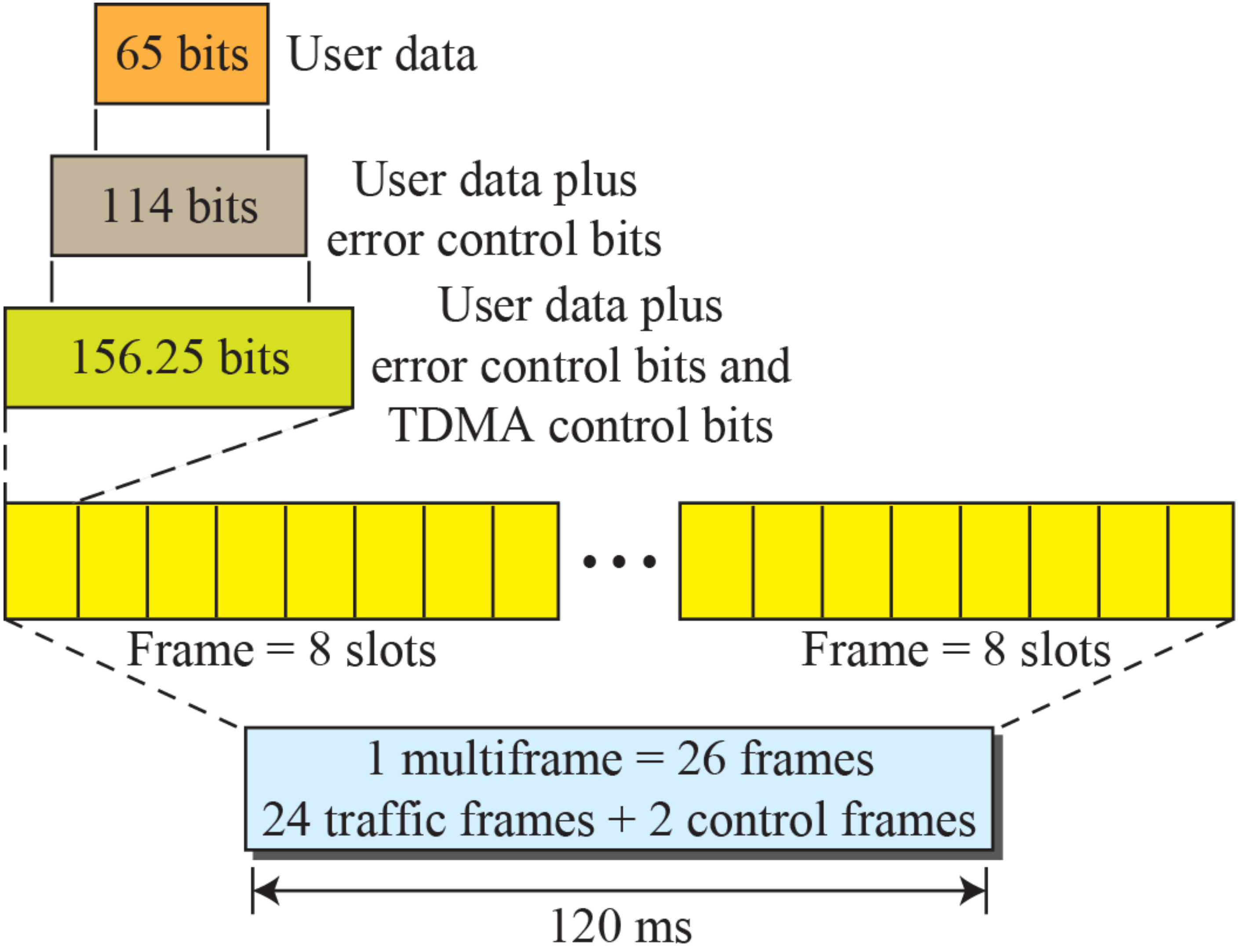

65 bits
User data
114 bits
User data plus error control bits
156.25 bits
User data plus error control bits and TDMA control bits
• • •
Frame = 8 slots
Frame = 8 slots
1 multiframe = 26 frames
24 traffic frames + 2 control frames
120 ms

Figure 16.13: Multiframe components

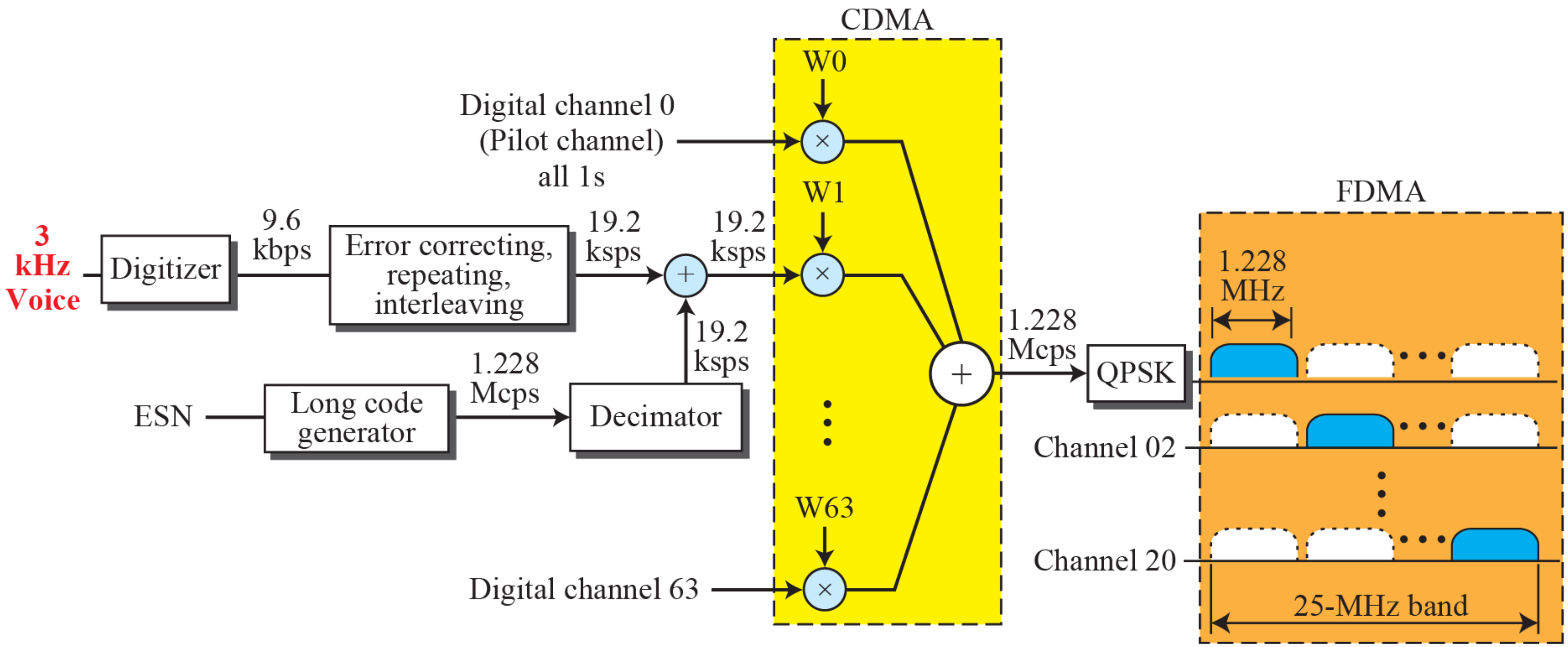
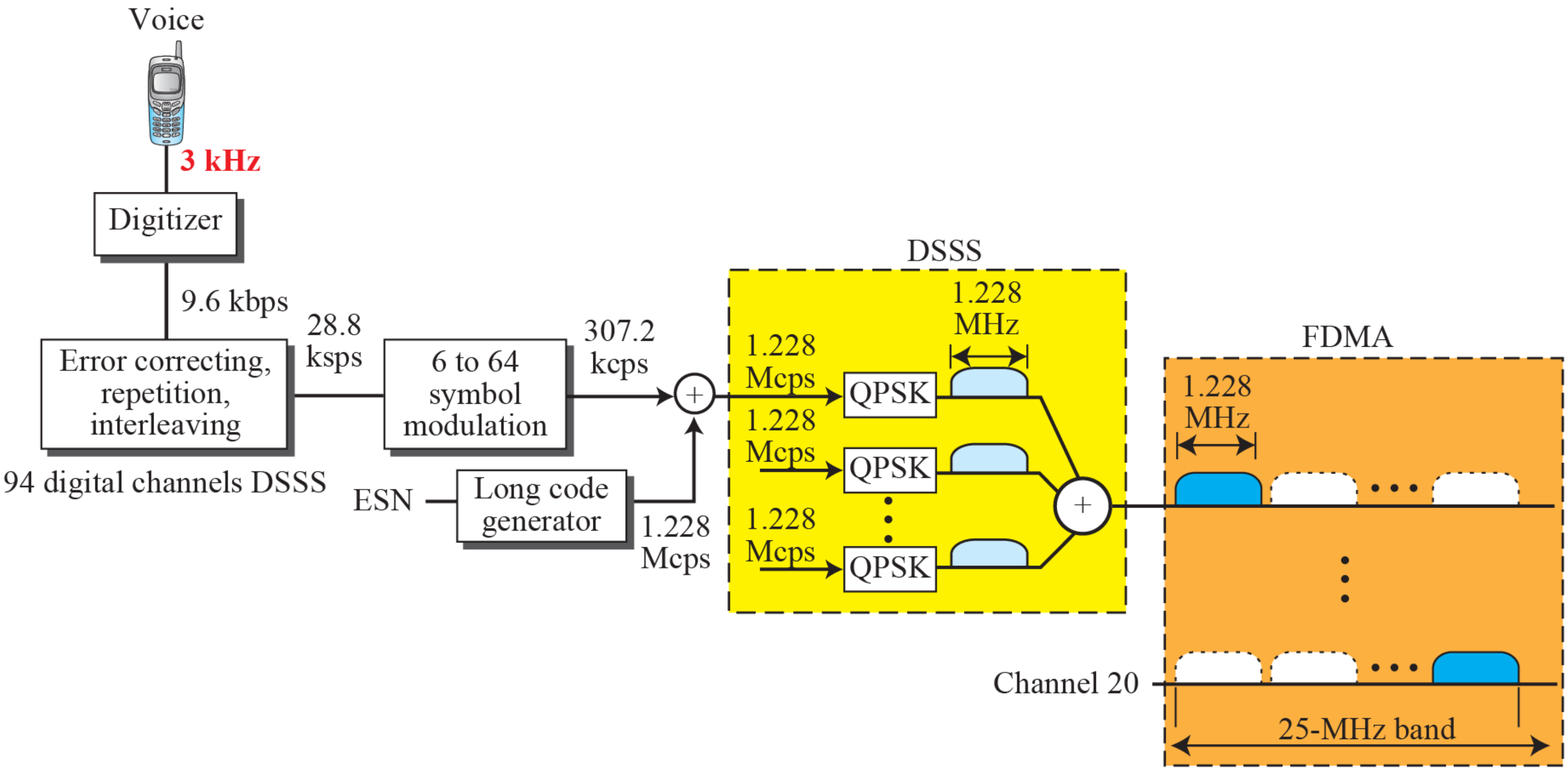


Figure 16.14: IS-95 forward transmission

Figure 16.15: S-95 reverse transmission



Third Generation (3G)

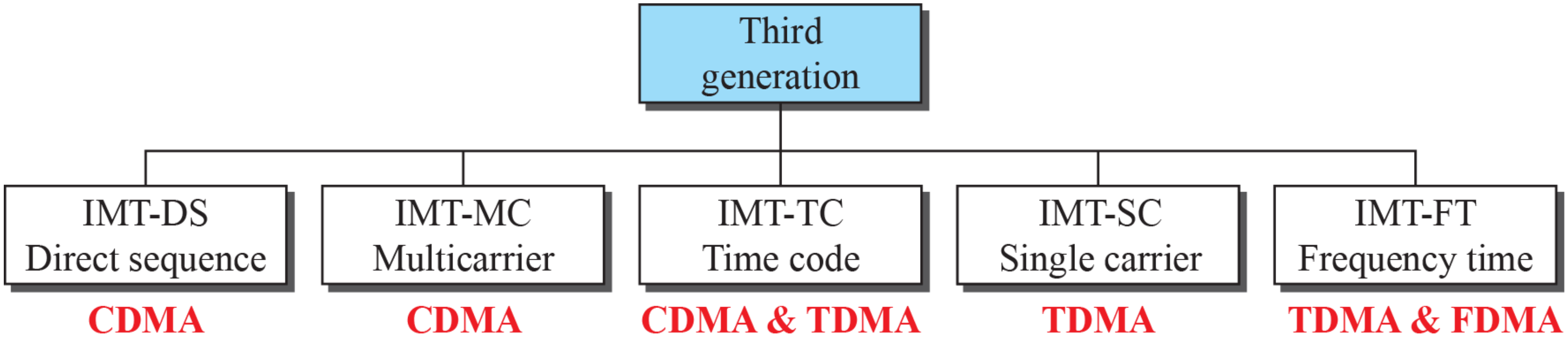
- The third generation of cellular telephony refers to a combination of technologies that provide both digital data and voice communication

- Using a small portable device, a person is able to talk to anyone else in the world with a voice quality similar to that of the existing fixed telephone network

- A person can download and watch a movie and listen to music, surf the Internet or play games

- The third-generation concept started in 1992, when ITU issued a blueprint called the Internet Mobile Communication 2000 (IMT-2000)

Figure 16.16: IMT-2000 radio interfaces



Fourth Generation (4G)

The fourth generation of cellular telephony is expected to be a complete evolution in wireless communications

- Some of the objectives defined by the 4G working group are as follows:

a. A spectrally efficient system.

b. High network capacity.

c. Data rate of 100 Mbit/s for access in a moving car and 1 Gbit/s for stationary users.

d. Data rate of at least 100 Mbit/s between any two points in the world.

e. Smooth handoff across heterogeneous networks.

f. Seamless connectivity and global roaming across multiple networks. g. High quality of service for next generation multimedia support

h. Interoperability with existing wireless standards.

i. All IP, packet-switched, networks

The fourth generation is only packet-based (unlike 3G) and supports IPv6. This provides better multicast, security, and route optimization capabilities.

- **Access Scheme**

To increase efficiency, capacity, and scalability, new access techniques are being considered for 4G. For example, orthogonal FDMA (OFDMA) and interleaved FDMA (IFDMA) are being considered respectively for the downlink and uplink of the next generation Universal Mobile Telecommunications System (UMTS). Similarly, multicarrier code division multiple access (MC-CDMA) is proposed for the IEEE 802.20 standard.

- **Modulation** More efficient quadrature amplitude modulation (64-QAM) is being proposed for use with the Long Term Evolution (LTE) standards.

.

Radio System The fourth generation uses a Software Defined Radio (SDR) system. Unlike a common radio, which uses hardware, 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.

Antenna The multiple-input multiple-output (MIMO) and multiuser MIMO (MU-MIMO) antenna system, a branch of intelligent antenna, is proposed for 4G. Using this antenna system together with special multiplexing, 4G allows independent streams to be transmitted simultaneously from all the antennas to increase the data rate into multiple folds. MIMO also allows the transmitter and receiver coordinates to move to an open frequency when interference occurs.