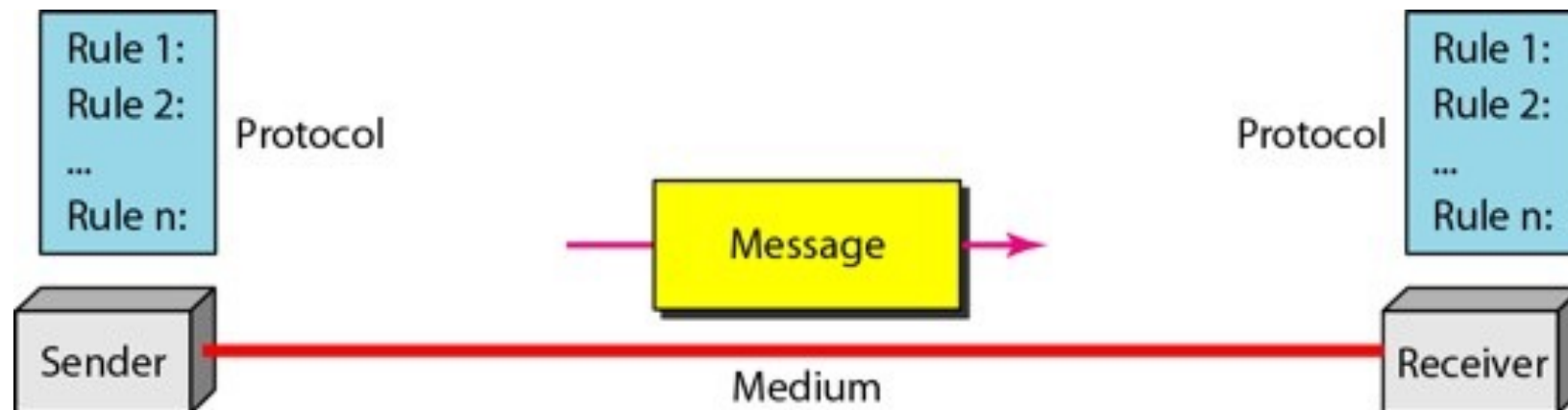


Computer Networks

UNIT 1

Data communications

- Data communications are the **exchange of data between two devices via some form of transmission medium such as a wire cable.**
- For data communications to occur, the **communicating devices must be part of a communication system** made up of a combination of hardware (physical equipment) and software (programs).



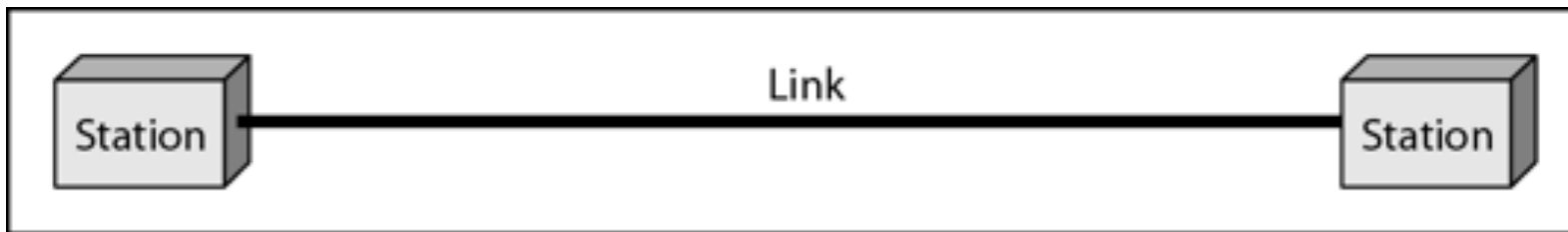
Components

A data communications system has five components

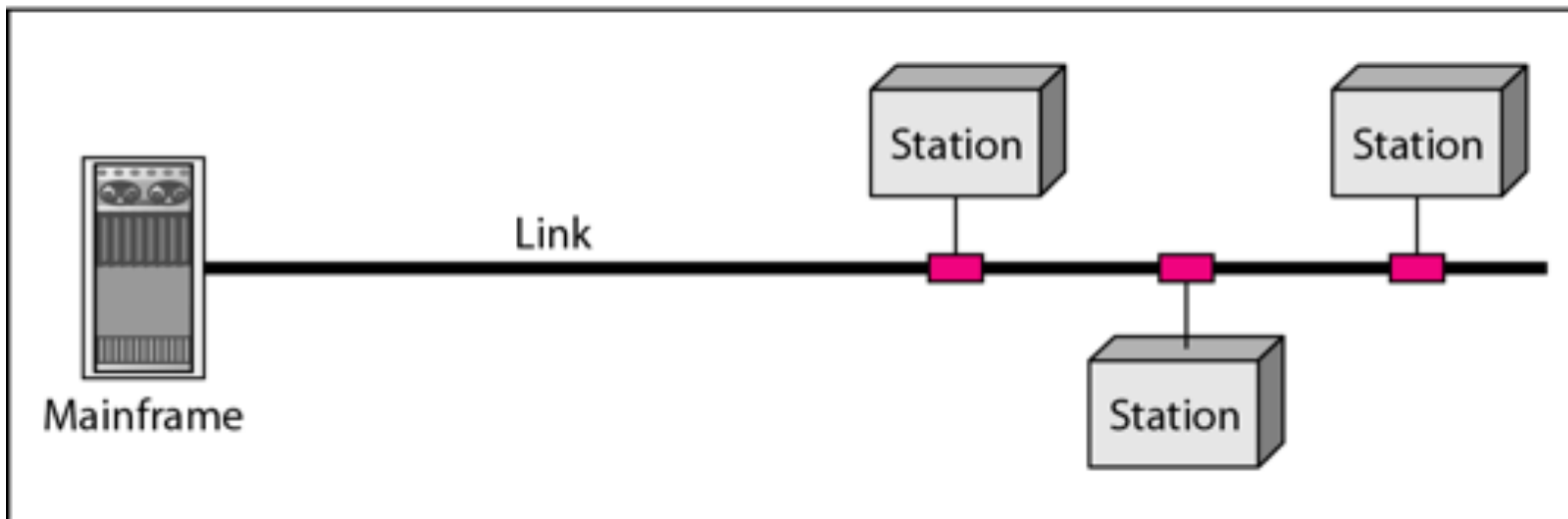
- 1. Message.** The message is the **information (data) to be communicated.**
- 2. Sender.** The sender is the **device that sends the data message.** It can be a computer, workstation, telephone handset, video camera, and so on.
- 3. Receiver.** The receiver is **the device that receives the message.** It can be a computer, workstation, telephone handset, television, and so on.
- 4. Transmission medium.** The transmission medium is the **physical path by which a message travels from sender to receiver.** Some examples of transmission media include twisted-pair wire, coaxial cable, fiber-optic cable, and radio waves.
- 5. Protocol.** A protocol is a set of rules that govern data communications. It represents an agreement between the communicating devices.

NETWORKS

- A network is a **set of devices (often referred to as *nodes*) interconnected 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
- **Type of Connection**
 - **Point to Point** - dedicated connection between 2 devices
 - **Multipoint** - multiple devices share the same communication link



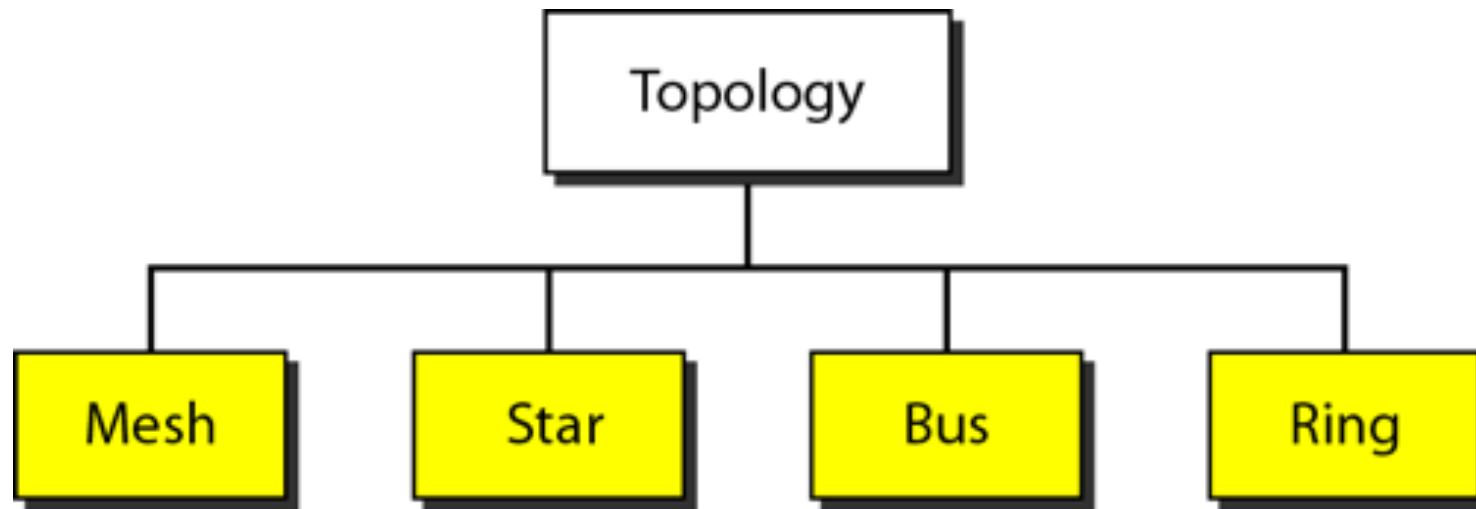
a. Point-to-point



b. Multipoint

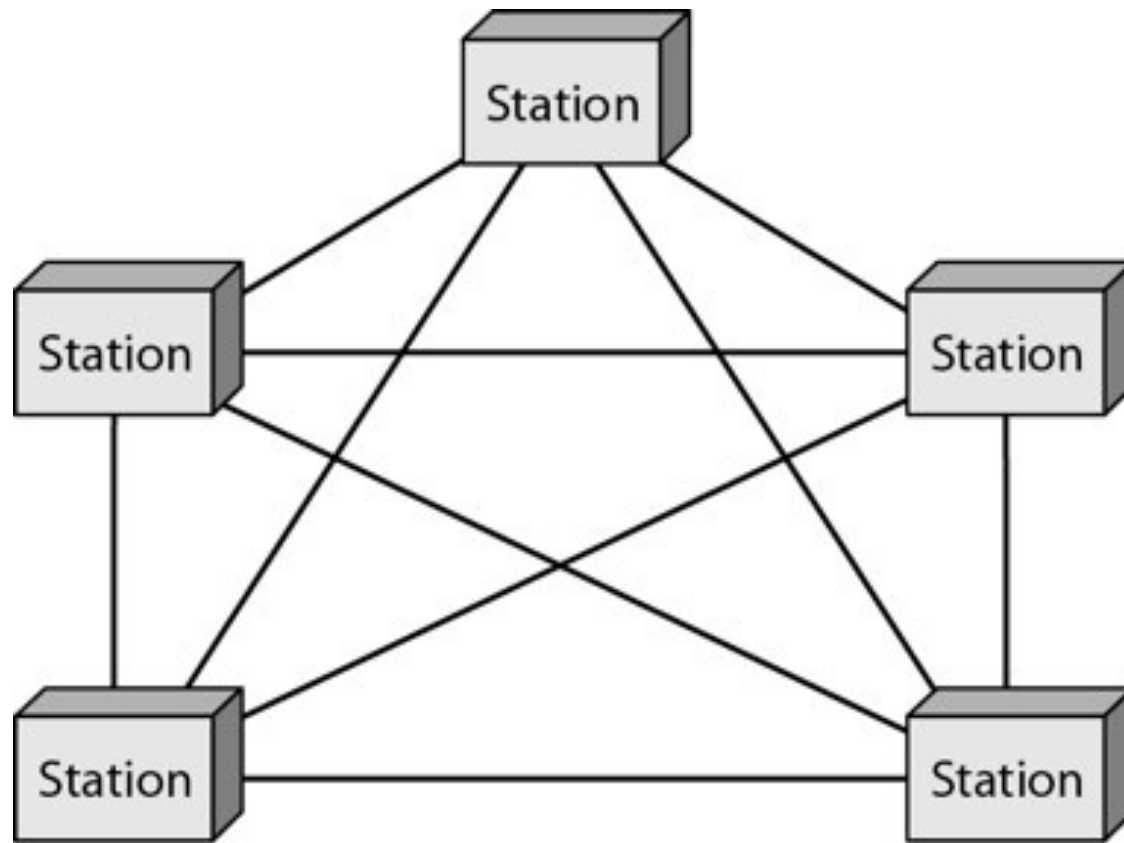
Physical Topology

- refers to the physical arrangement of nodes(stations) in a network.
- There are four basic topologies possible: mesh, star, bus, and ring



Mesh Topology

- every device has a **dedicated point-to-point link to every other**
dedicated means that the link carries traffic only between the two devices it connects.
- To find the number of physical links in a fully connected mesh network with n nodes, we first consider that each node must be connected to every other node.
- Node 1 must be connected to $n - 1$ nodes, node 2 must be connected to $n - 1$ nodes, node n must be connected to $n - 1$ nodes. ❖ need $n(n - 1)$ physical links.
- However, if each physical link allows communication in both directions (duplex mode), can do ❖ $n(n - 1) / 2$ duplex-mode links. / 2.
- To accommodate that many links, every device on the network must have **$n - 1$ input/output ports to be connected to the other $n - 1$ stations.**



Advantages

1. the use of dedicated links guarantees that each **connection can carry its own data load, thus eliminating the traffic problems that can occur when links must be shared by multiple devices.**

2. a mesh topology **is robust.** If one link becomes unusable, it does not incapacitate the entire system.

3. privacy or security

When every message travels along a dedicated line, only the intended recipient sees it. Physical boundaries prevent other users from gaining access to messages.

4. point-to-point links **make fault identification and fault isolation easy.** Traffic can be routed to avoid links with suspected problems. This facility enables the network manager to discover the precise location of the fault and aids in finding its cause and solution.

Disadvantages

1. every device must be connected to every other device, **installation and reconnection are difficult.**

2. **the sheer bulk of the wiring can be greater than the available space (in walls, ceilings, or floors) can accommodate.**

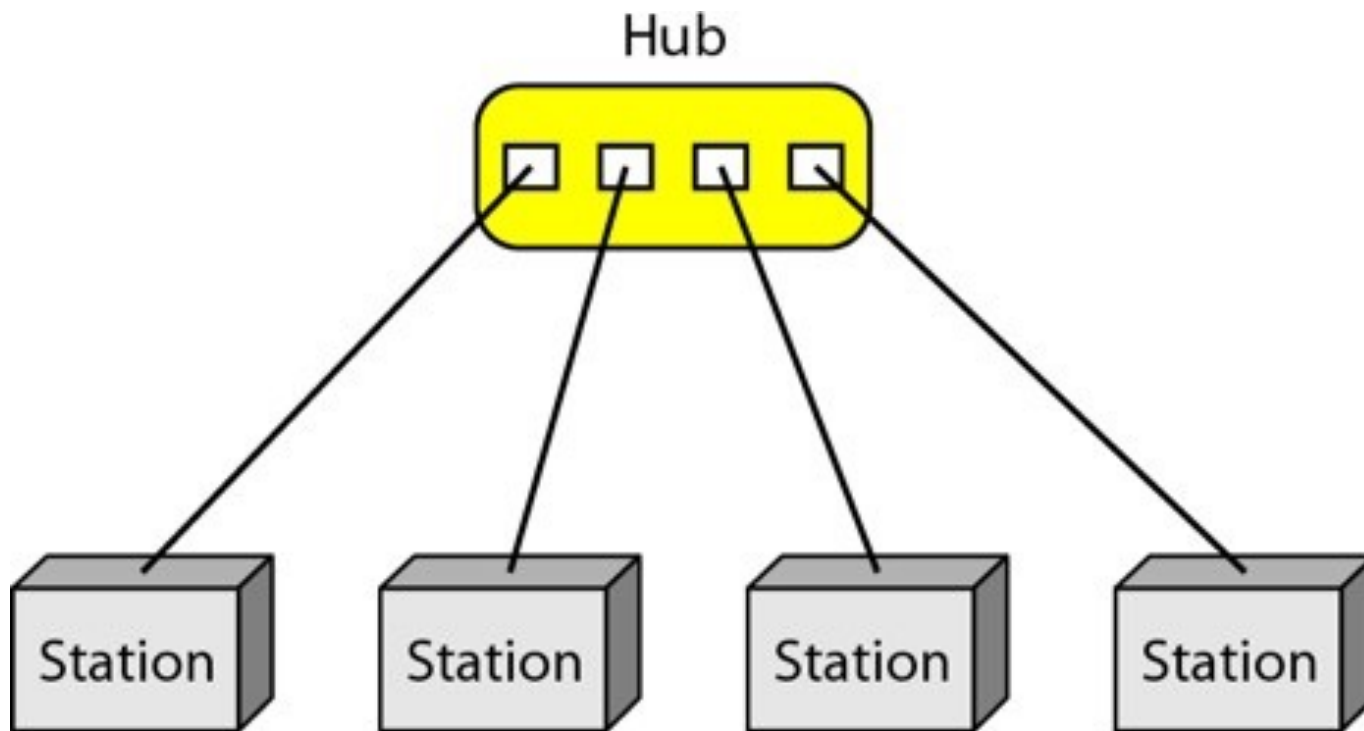
3. the hardware required to connect each link (I/O ports and cable) can be **prohibitively expensive.**

For these reasons a mesh topology is **usually implemented in a limited fashion**, for example, as a **backbone connecting the main computers of a hybrid network** that can include several other topologies.

Star Topology contd..

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

Star Topology contd..



advantage

- A star topology is **less expensive** than a mesh topology.
- In a star, each device **needs only one link and one I/O port to connect it to any number of others.**
- **easy to install and reconfigure.**
 - Far less cabling needs to be housed, and additions, moves, and deletions involve only one connection between that device and the hub.
- **robustness.**
 - If one link fails, only that link is affected. All other links remain active.
- **easy fault identification and fault isolation**
 - As long as the hub is working, it can be used to monitor link problems and bypass defective links.

disadvantage of a star topology

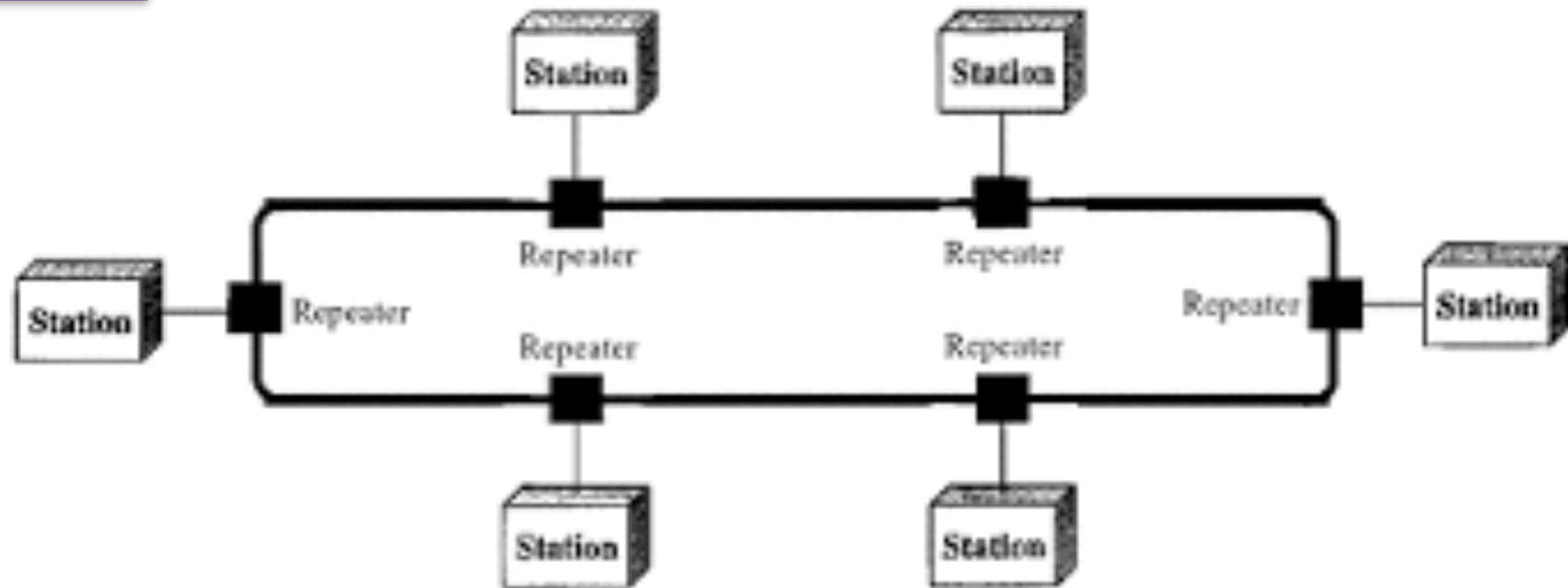
- the dependency of the whole topology on one single point, the hub. If **the hub goes down, the whole system is dead.**

- Although a star requires far less cable than a mesh, each node must be linked to a central hub. For this reason, **often more cabling is required in a star than in some other topologies**

Use

- **high speed LAN**

Ring Topology



- each device has a **dedicated point-to-point connection with only the two devices** on either side of it.
- **A signal is passed along the ring in one direction**, from device to device, until it reaches its destination.

- Each device in the ring incorporates a **repeater**.
- When a device receives a signal intended for another device, its **repeater regenerates the bits and passes them along**

Advantages

- A ring is relatively **easy to install and reconfigure**.
 - Each device is linked to only its immediate neighbors (either physically or logically). To add or delete a device requires changing only two connections, the only constraints are media and traffic considerations (maximum ring length and number of devices).
- **fault isolation is simplified**.
 - Generally in a ring, a signal is circulating at all times. 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**

Disadvantage

- **unidirectional traffic**
 - In a simple ring, a break in the ring (such as a disabled station) can disable the entire network.
 - This weakness can be solved by using a dual ring or a switch capable of closing off the break

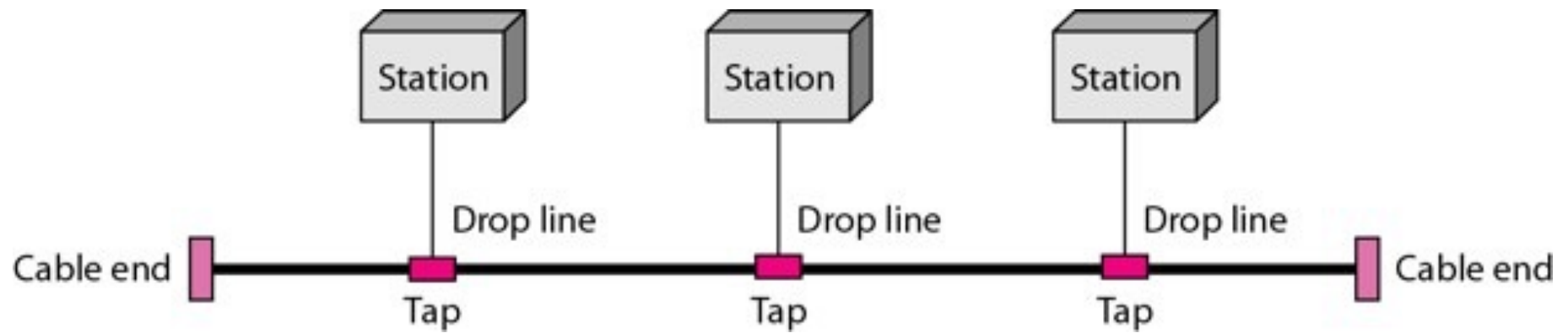
Use

- Ring topology was prevalent when IBM introduced its **local-area network Token Ring**

Bus topology

- **multipoint.**
- One long cable acts as a **backbone** to link all the devices in a network
- Nodes are connected to the bus cable by **drop lines and taps.**
- A **drop line** is a connection running between the device and the main cable.
- A **tap** is a connector that either splices into the main cable or punctures the sheathing of a cable to create a contact with the metallic core.
 - ♣ As a signal travels along the backbone, some of its energy is transformed into heat. Therefore, it becomes weaker and weaker as it travels farther and farther. For this reason there is a limit on the number of taps a bus can support and on the distance between those taps.

Bus topology contd..



Advantages of a bus topology

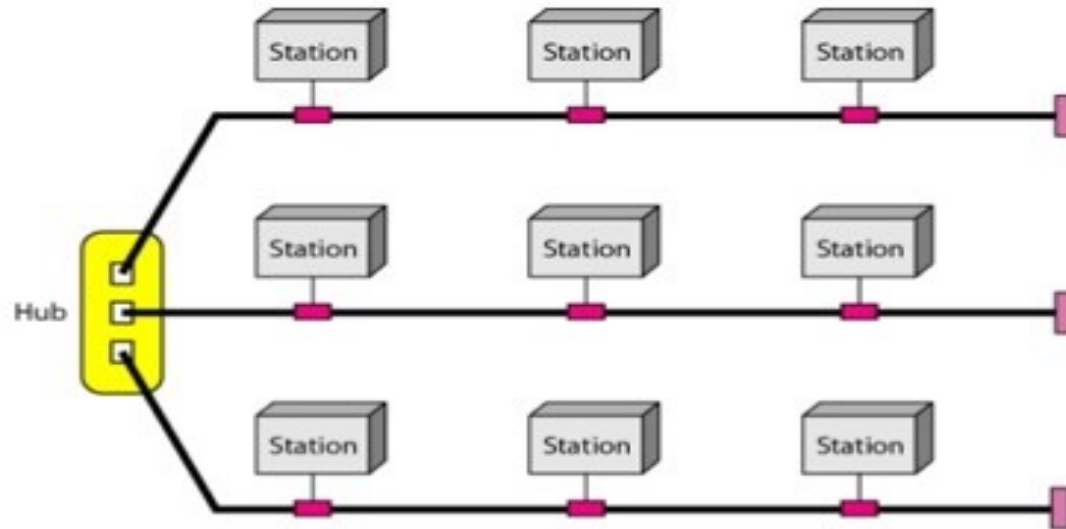
- **ease of installation**
 - ❖ Backbone cable can be laid along the most efficient path, then connected to the nodes by drop lines of various lengths.
- a bus uses **less cabling** than mesh or star topologies.
 - ❖ In a star, for example, four network devices in the same room require four lengths of cable reaching

Disadvantages

- difficult reconnection and fault isolation.
- ❖ A bus is usually designed to be optimally efficient at installation. It can therefore be difficult to add new devices. Signal reflection at the taps can cause degradation in quality. This degradation can be controlled by limiting the number and spacing of devices connected to a given length of cable.
- **Adding new devices may therefore require modification or replacement of the backbone.**
- ❖ a fault or break in the bus cable stops all transmission, even between devices on the same side of the problem.
- ❖ The damaged area reflects signals back in the direction of origin, creating noise in both directions.

- Bus topology was the one of the first topologies used in the design of early local area networks.
- Ethernet LANs can use a bus topology

HYBRID TOPOLOGY: A STAR BACKBONE WITH THREE BUS NETWORKS



- A network can be hybrid.
- For example, we can have a main star topology with each branch connecting several stations in a bus topology

Categories of Networks

- The category into which a network falls is determined by its size
 - LAN -Local Area Network
 - MAN-Metropolitan Area Network
 - WAN - Wide Area Network

Local Area Network (LAN)

- *privately owned and links the devices in a single office, building, or campus*
- Depending on the needs of an organization and the type of technology used, a LAN can be as simple as two PCs and a printer in someone's home office;
- Currently, LAN size is limited to a few kilometers

- LANs are designed to **allow resources to be shared between personal computers or workstations.**
- The resources to be shared can include **hardware (e.g., a printer), software(e.g., an application program), or data.**
- In addition to size, LANs are distinguished from other types of networks by their **transmission media and topology.**

In general, a given LAN will use only **one type of transmission medium.**

most common LAN topologies **are bus, ring, and star.**

Early LANs had **data rates in the 4 to 16 megabits per second (Mbps) range.**

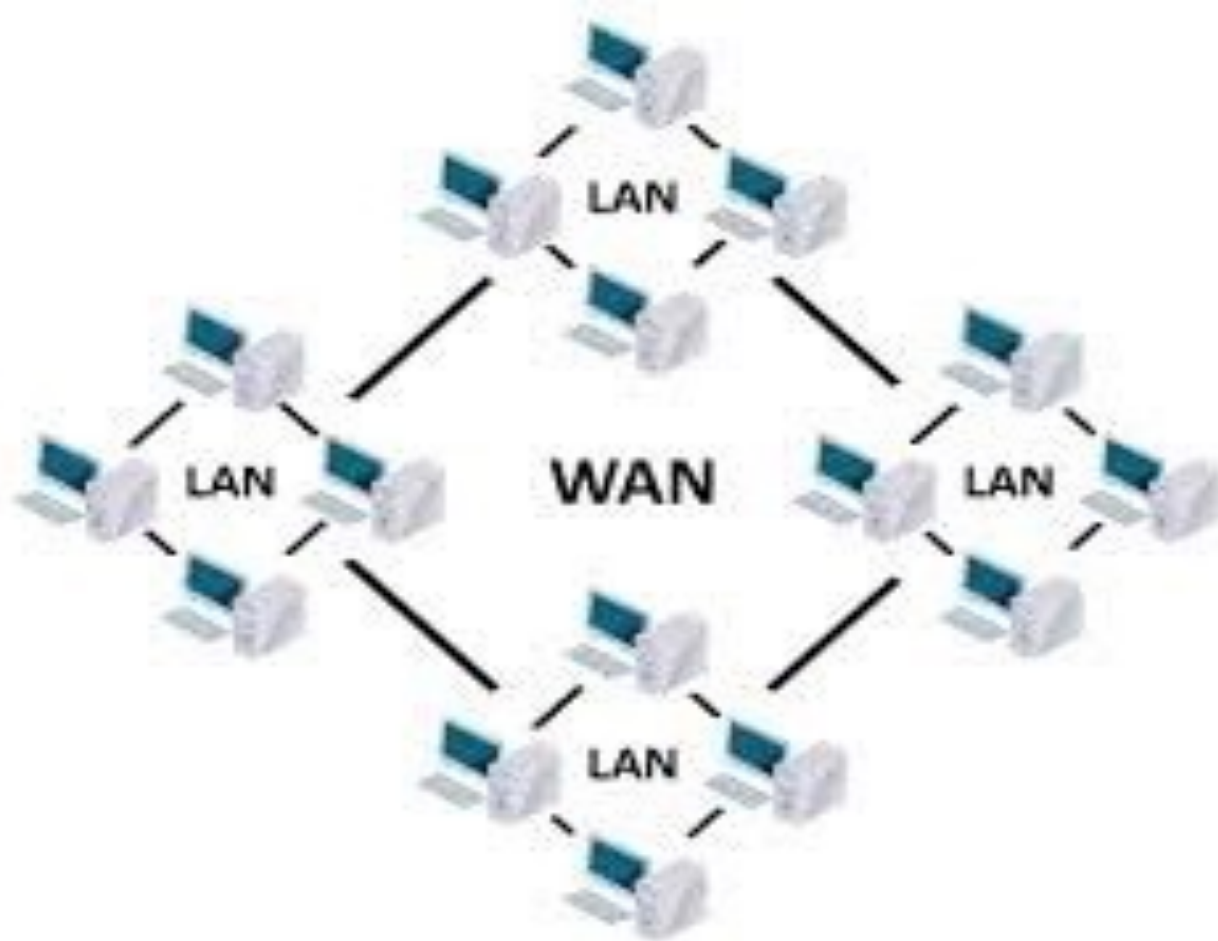
Today, speeds are normally 100 or 1000 Mbps

- provides **long-distance transmission of data, image, audio, and video information over large geographic areas** that may comprise a country, a continent, or even the whole world
- A WAN can be as **complex** as the backbones that connect the Internet (switched WAN) or as **simple** as a dial-up line that connects a home computer to the Internet (point-to-point WAN).
- The **switched WAN** connects the end systems, which usually comprise a router that connects to another LAN or WAN.

- The **point-to-point WAN** is normally
 - ◆ a line leased from a telephone or cable TV provider that connects a home computer or
 - ◆ a small LAN to an Internet service provider (ISP). This type of WAN is often used to provide Internet access.

- WAN that is **wholly owned and used by a single company is enterprise network**

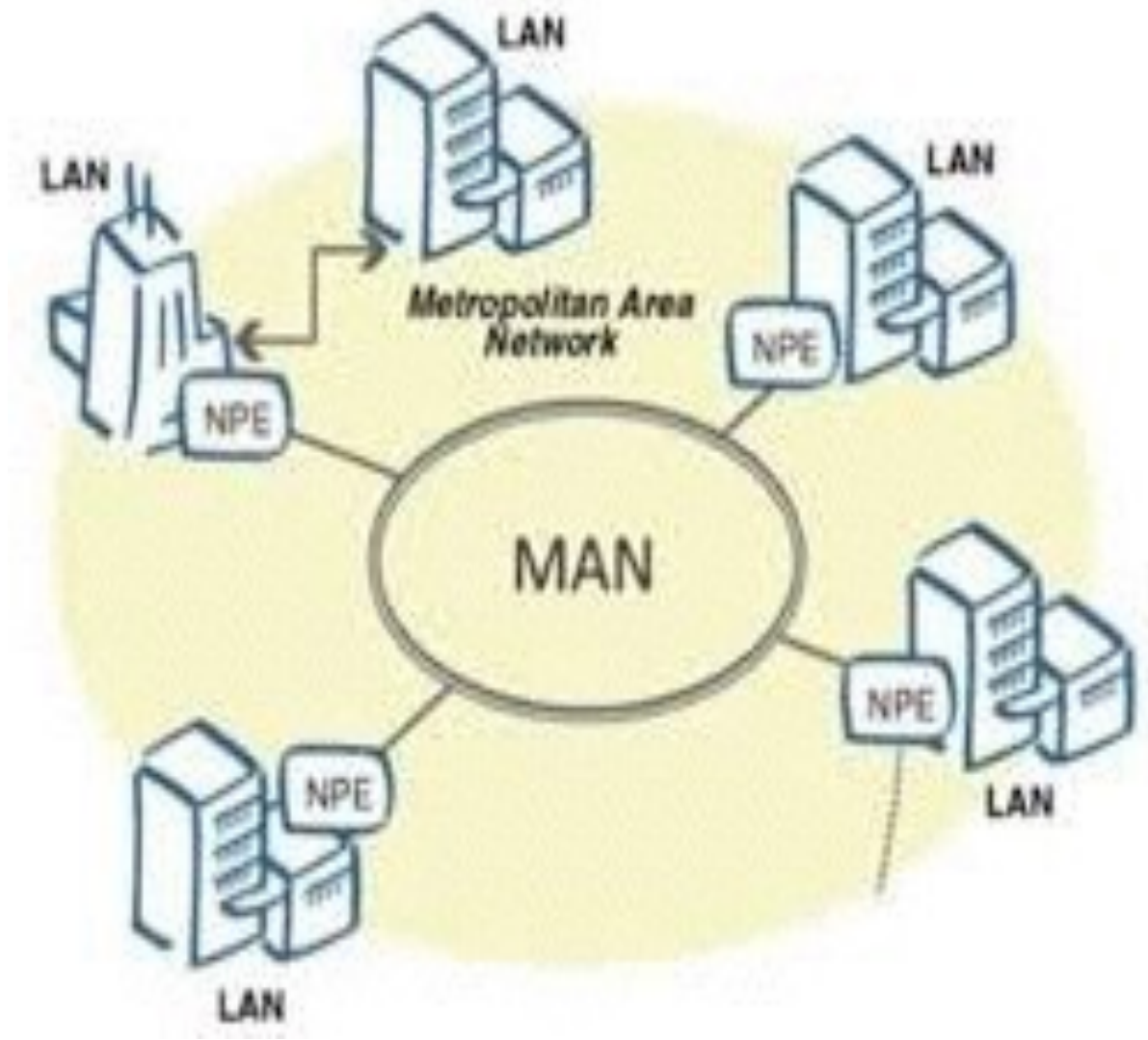
Wide Area Network (WAN) contd..



Metropolitan Area Networks (MAN)

- A metropolitan area network (MAN) is a network with a size between a LAN and a WAN.
- It normally covers the area inside a town or a city. It extends over an entire city
- It is designed for customers who need a high-speed connectivity, normally to the Internet, and have endpoints spread over a city or part of city.
- A good example of a MAN is the part of the telephone company network that can provide a high-speed DSL line to the customer.
- Another example is the cable TV network that originally was designed for cable TV, but can also be used for high-speed data connection to the Internet.

Metropolitan Area Networks (MAN) contd..



- **MAN may be wholly owned and operated by a private company or may be a service provided by a public company**

internet

- **When two or more networks are connected, they become an internetwork, or internet.**
- **An internet is a network of networks.**
- **The Internet is a collection of many separate networks**

A protocol is a set of rules that govern data communication

- A protocol defines
 - what is communicated,
 - how it is communicated, and
 - when it is communicated.
- The key elements of a protocol are
 - syntax,
 - semantics,
 - timing.
-

Syntax.

- The term *syntax* refers to the **structure or format of the data, meaning the order in which they are presented.**
- **For example,** a simple protocol might expect the first 8 bits of data to be the address of the sender, the second 8 bits to be the address of the receiver, and the rest of the stream to be the message itself

- The word *semantics* refers to ***the meaning of each section of bits.***
- How is a particular pattern to be interpreted, and what action is to be taken based on that interpretation?
- **For example**, does an address identify the route to be taken or the final destination of the message?

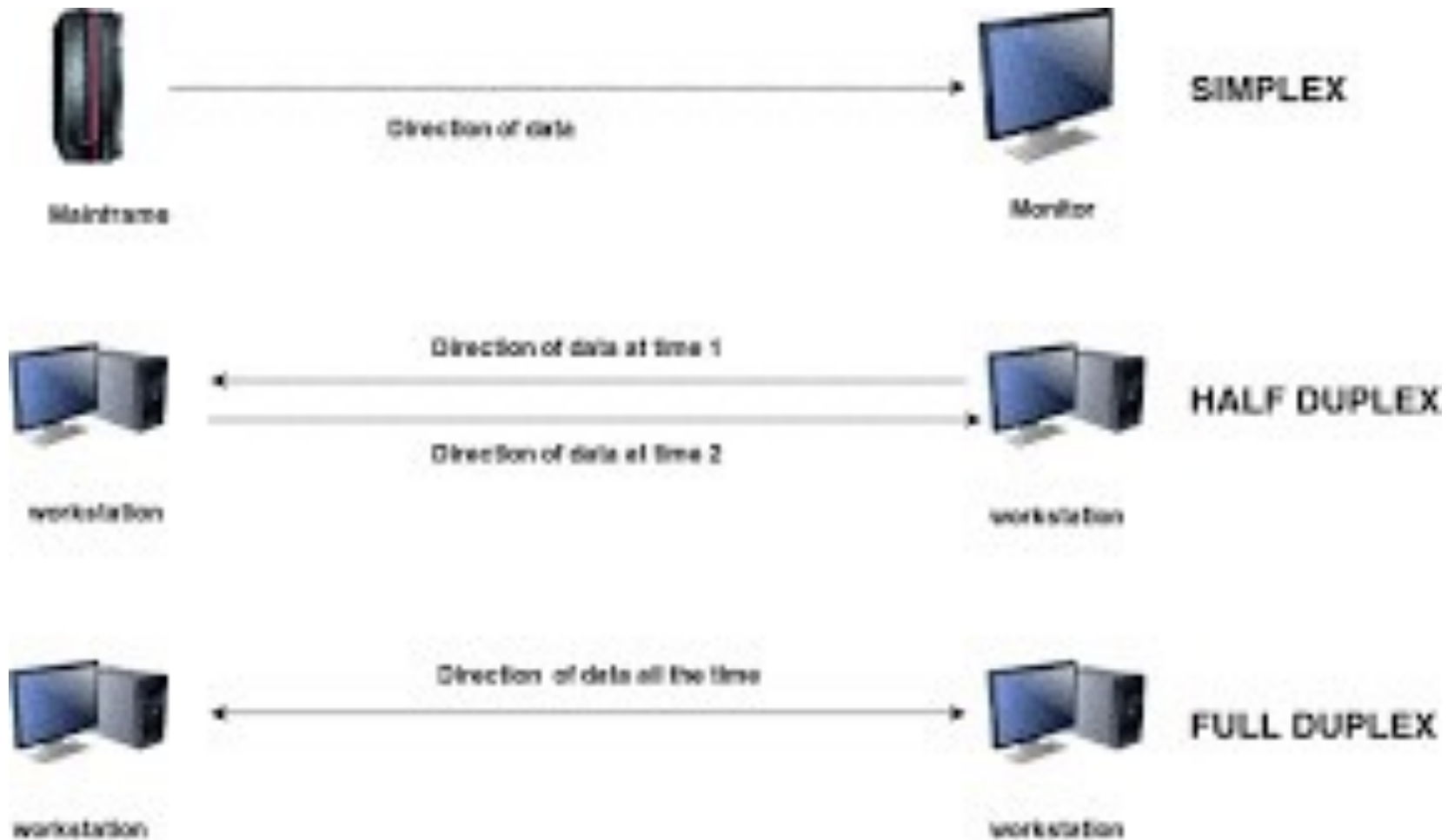
Timing.

- The term *timing* refers to ***two characteristics:***
 - *when data should be sent* and
 - *how fast they can be sent.*
- **For example**, if a sender produces data at 100 Mbps but the receiver can process data at only 1 Mbps, the transmission will overload the receiver and some data will be lost

Data Transmission modes

- Communication between two devices can be
 - simplex,
 - half-duplex,
 - or full-duplex

Data Transmission modes contd..



Simplex

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

Half-Duplex

- In half-duplex mode, **each station can both transmit and receive, but not at the same time.**
- When one device is sending, the other can only receive, and vice versa
- The half-duplex mode is like a one-lane road with traffic allowed in both directions.
- the entire capacity of a channel is taken over by whichever of the two devices is transmitting at the time.
- **Walkie-talkies and CB (citizens band) radios**
- The half-duplex mode is used in cases where there is no need for communication in both directions at the same time

Full-Duplex

- In full-duplex mode(also called duplex), both stations can transmit and receive simultaneously
- The full-duplex mode is like a two way street with traffic flowing in both directions at the same time.
- In full-duplex mode, signals going in one direction share the capacity of the link: with signals going in the other direction.

- Either the link must contain **two physically separate transmission paths**, one for sending and the other for receiving; or **the capacity of the channel is divided** between signals traveling in both directions.
- **eg telephone network.**
- When two people are communicating by a telephone line, both can talk and listen at the same time.
-

The full-duplex mode is used when communication in both directions is required all the time.

- Computer networks are created by different entities.
- **Standards are needed so that these heterogeneous networks can communicate with one another.**
- The two best-known standards
 - **ISO- OSI model**
 - **Internet model(TCP/IP)**

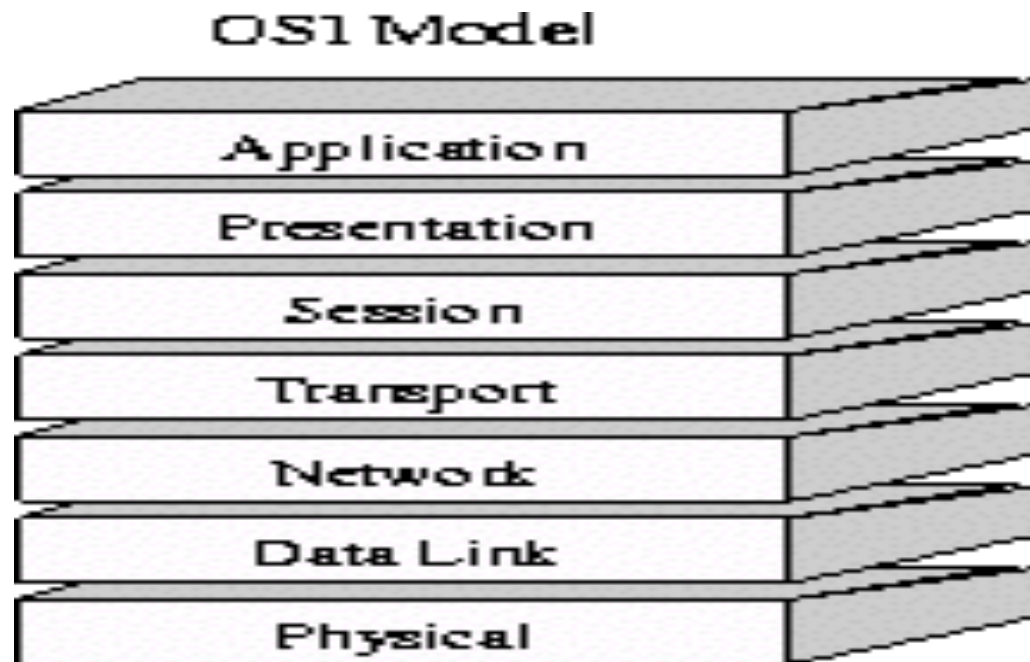
ISO -OSI MODEL

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

- **An open system is a set of protocols that allows any two different systems to communicate regardless of their underlying architecture.**

The OSI model is not a protocol;

- **It consists of seven separate but related layers, each of which defines a part of the process of moving information across a network**



The principles that were applied to arrive at the seven layers can be briefly summarized as follows:

1. A layer should be created where a different level of abstraction is needed.
2. Each layer should perform a well-defined function.
3. The function of each layer should be chosen based on the internationally standardized protocols.

4. The layer **boundaries should be chosen** to minimize the information flow across the interfaces.

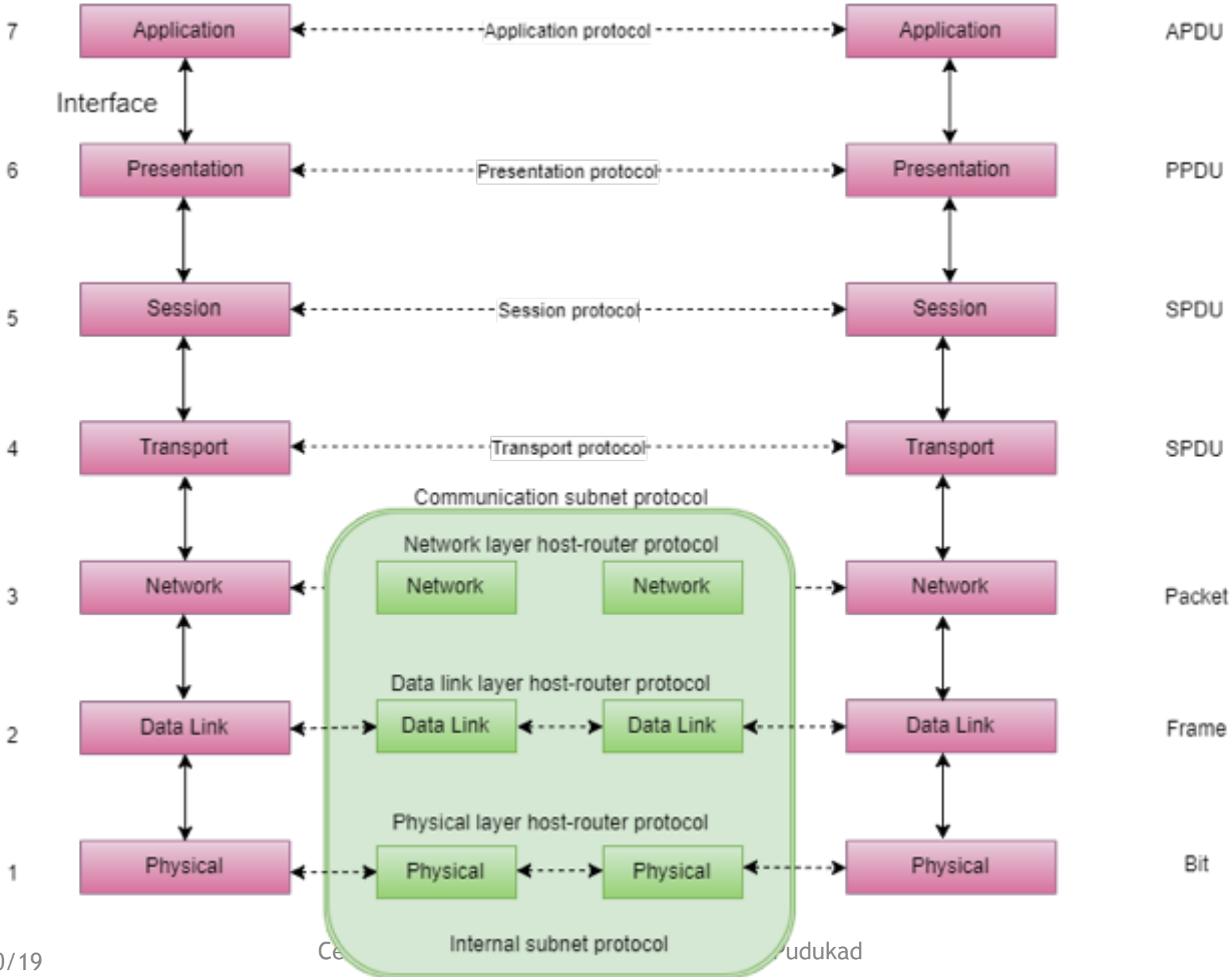
5. The number of layers **should be large enough** that distinct functions need not be thrown together in the same layer out of necessity and **small enough** that architecture does not become unmanageable

Host A

ISO -OSI MODEL contd..

Host B

Name of Unit Exchanged



Layer	Name of Protocol	Name of Unit exchanged
Application	Application Protocol	APDU - Application Protocol Data Unit
Presentation	Presentation Protocol	PPDU - Presentation Protocol Data Unit
Session	Session Protocol	SPDU - Session Protocol Data Unit
Transport	Transport Protocol	TPDU - Transport Protocol Data Unit
Network	Network layer host-router Protocol	Packet
Data Link	Data link layer host-router Protocol	Frame
Physical	Physical layer host-router Protocol	Bit

Peer-to-Peer Processes

- At the physical layer, **communication is direct**, device A sends a stream of bits to device B (through intermediate nodes).
- At the higher layers, **communication must move down through the layers on device A, over to device B, and then back up through the layers.**
- **Each layer in the sending device adds its own information to the message it receives from the layer just above it and passes the whole package to the layer just below it.**

- At layer 1 the **entire package is converted to a form that can be transmitted to the receiving device.**
- At the **receiving machine, the message is unwrapped layer by layer**, with each process receiving and removing the data meant for it.
- For **example**, layer 2 removes the data meant for it, then passes the rest to layer 3. Layer 3 then removes the data meant for it and passes the rest to layer 4, and so on.

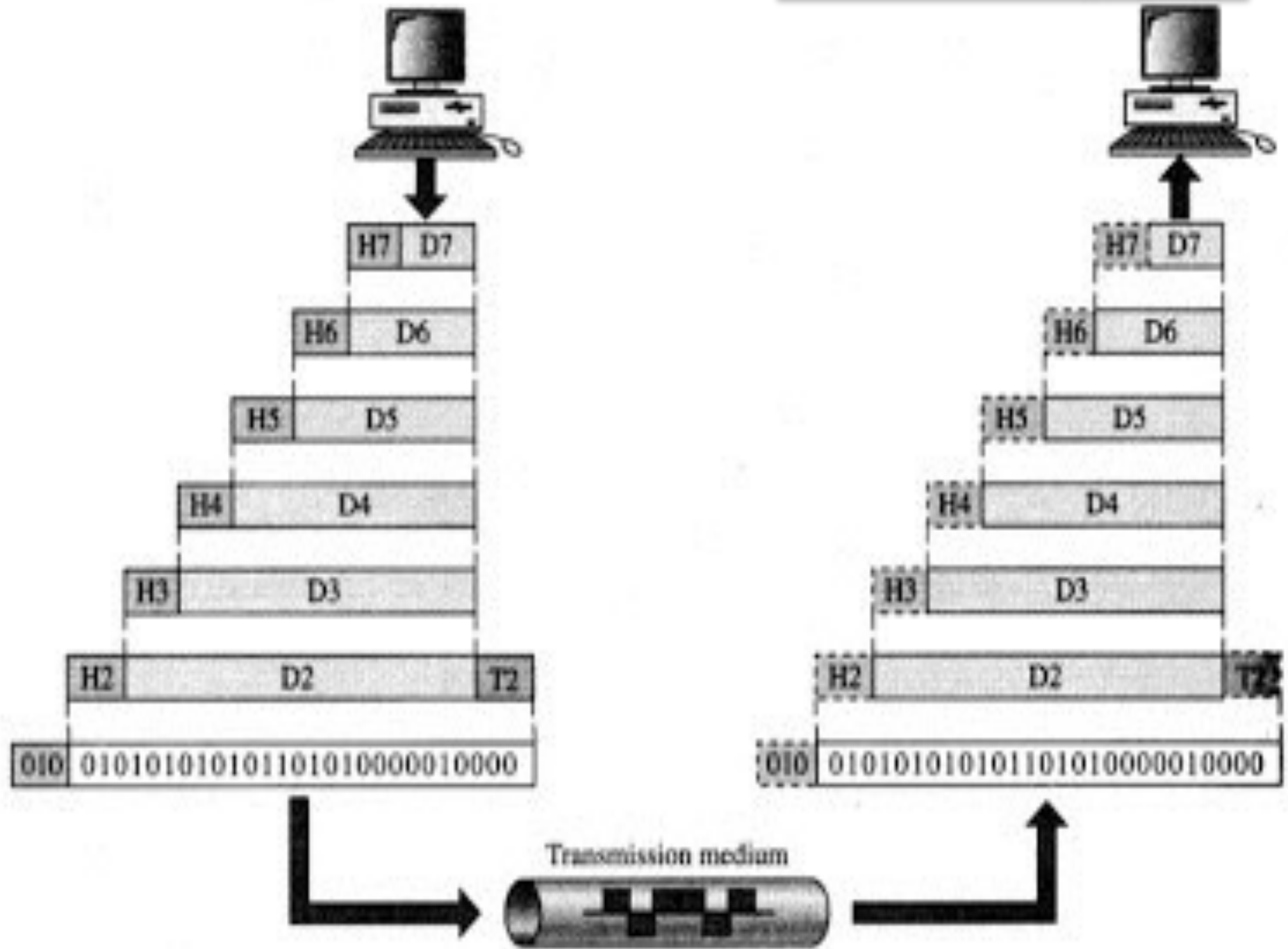
ISO -OSI MODEL contd..

- **The passing of the data and network information down through the layers of the sending device and back up through the layers of the receiving device is made possible by an interface between each pair of adjacent layers.**
- **Each interface also defines the information and services a layer must provide for the layer above it.**
- **Well-defined interfaces and layer functions provide modularity to a network.**

Organization of Layers

- The seven layers can be thought of as belonging to **three subgroups**.
- Layers 1, 2, and 3-physical, data link, and network-are the **network support layers(chained layers)**
 - they deal with the physical aspects of moving data from one device to another (such as electrical specifications, physical connections, physical addressing, and transport timing and reliability).

- Layers 5, 6, and 7-session, presentation, and application- can be thought of as the **user support layers(end to end layers)**
 - they allow interoperability among unrelated software systems.
- Layer 4, the transport layer, links the two subgroups and ensures that **what the lower layers have transmitted is in a form that the upper layers can use.**
- The upper OSI layers are almost always implemented in **software**
- **lower layers are a combination of hardware and software, except for the physical layer, which is mostly hardware.**



- D7 means the data unit at layer 7, D6 means the data unit at layer 6, and so on.
- The process starts at layer 7 (the application layer), then moves from layer to layer in descending, sequential order.
- At each layer, a **header, or possibly a trailer, can be added to the data unit.**
- Commonly, the trailer is added only at layer 2.
- When the formatted data unit passes through the physical layer (layer 1), it is changed into an electromagnetic signal and transported along a physical link.

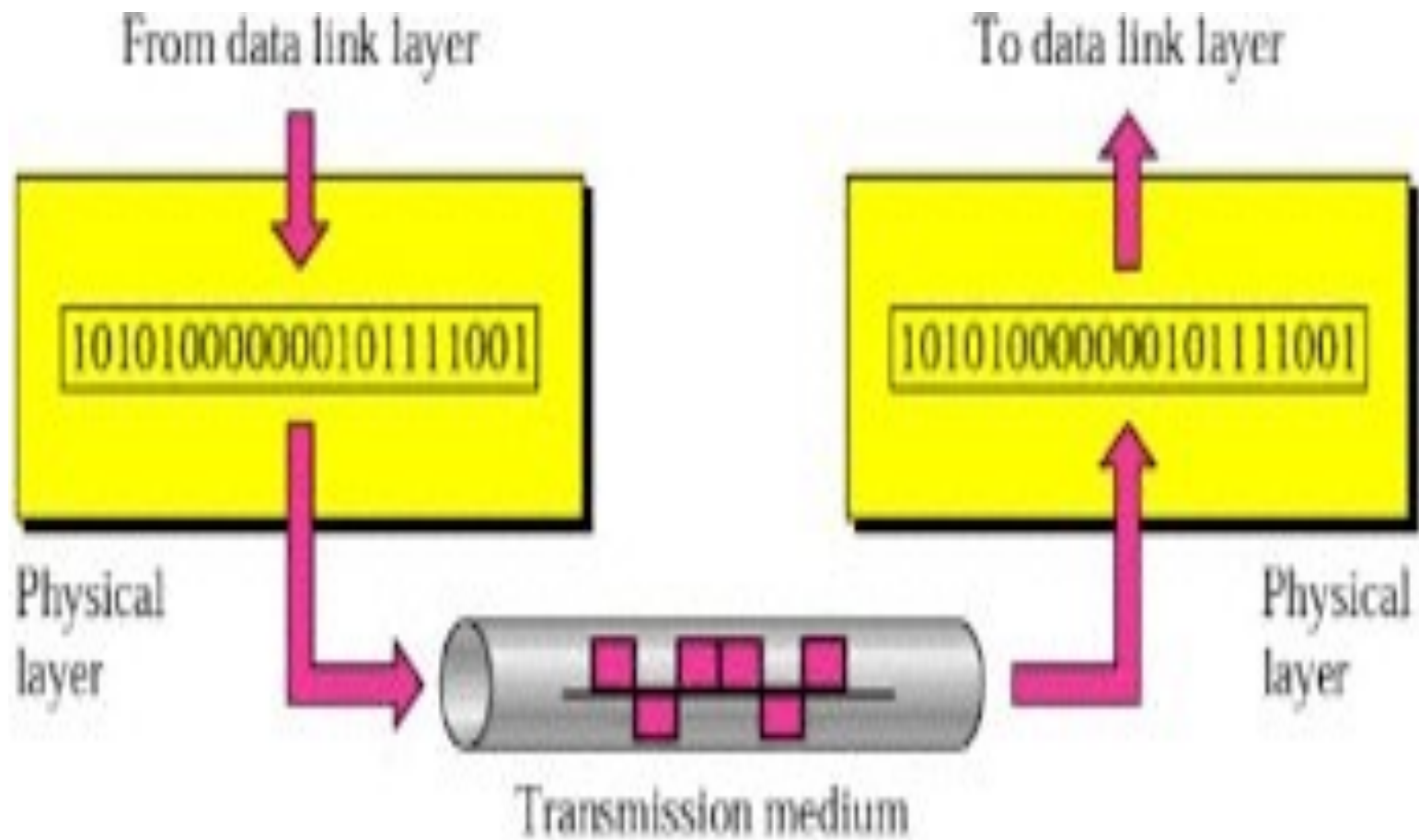
ISO -OSI MODEL contd..

- Upon reaching its destination, the signal passes into layer 1 and is transformed back into digital form.
- The data units then move back up through the OSI layers.
- As each block of data reaches the next higher layer, the headers and trailers attached to it at the corresponding sending layer are removed, and actions appropriate to that layer are taken.
- By the time it reaches layer 7, the message is again in a form appropriate to the application and is made available to the recipient.

Physical Layer

- Main function is **transmission of raw bits (0s and 1s) from one node to the next.**
- It deals with the following
 1. mechanical specifications
 2. electrical specifications
 3. procedural specifications
 4. functional specifications

Any system or device connected to a network is called a node
eg.
computer,
router,
printer



Other functions of Physical layer are

1. Physical characteristics of interfaces and medium.

- The physical layer defines the characteristics of the interface between the devices and the transmission medium.
- It also defines the **type of transmission medium**.

2. Representation of bits.

- The physical layer data consists of a stream of bits
- To be transmitted, bits must be encoded into signals--electrical or optical.
- The physical layer defines the type of encoding (how 0s and 1s are changed to signals).

3. Data rate(transmission rate)

- **number of bits sent each second** is also defined by the physical layer.
- the physical layer defines the **duration of a bit(bit period)**, which is how long it lasts.

4. Synchronization of bits.

- the sender and the receiver clocks must be synchronized.

5. Line configuration: the connection of devices to the media.

- In a **point-to-point configuration**, two devices are connected through a dedicated link.
- In a **multipoint configuration**, a link is shared among several devices.

6. Physical topology.

- The physical topology defines how devices are connected to make a network.

7. Transmission mode.

- The physical layer also defines the direction of transmission between two devices: simplex, half-duplex, or full-duplex

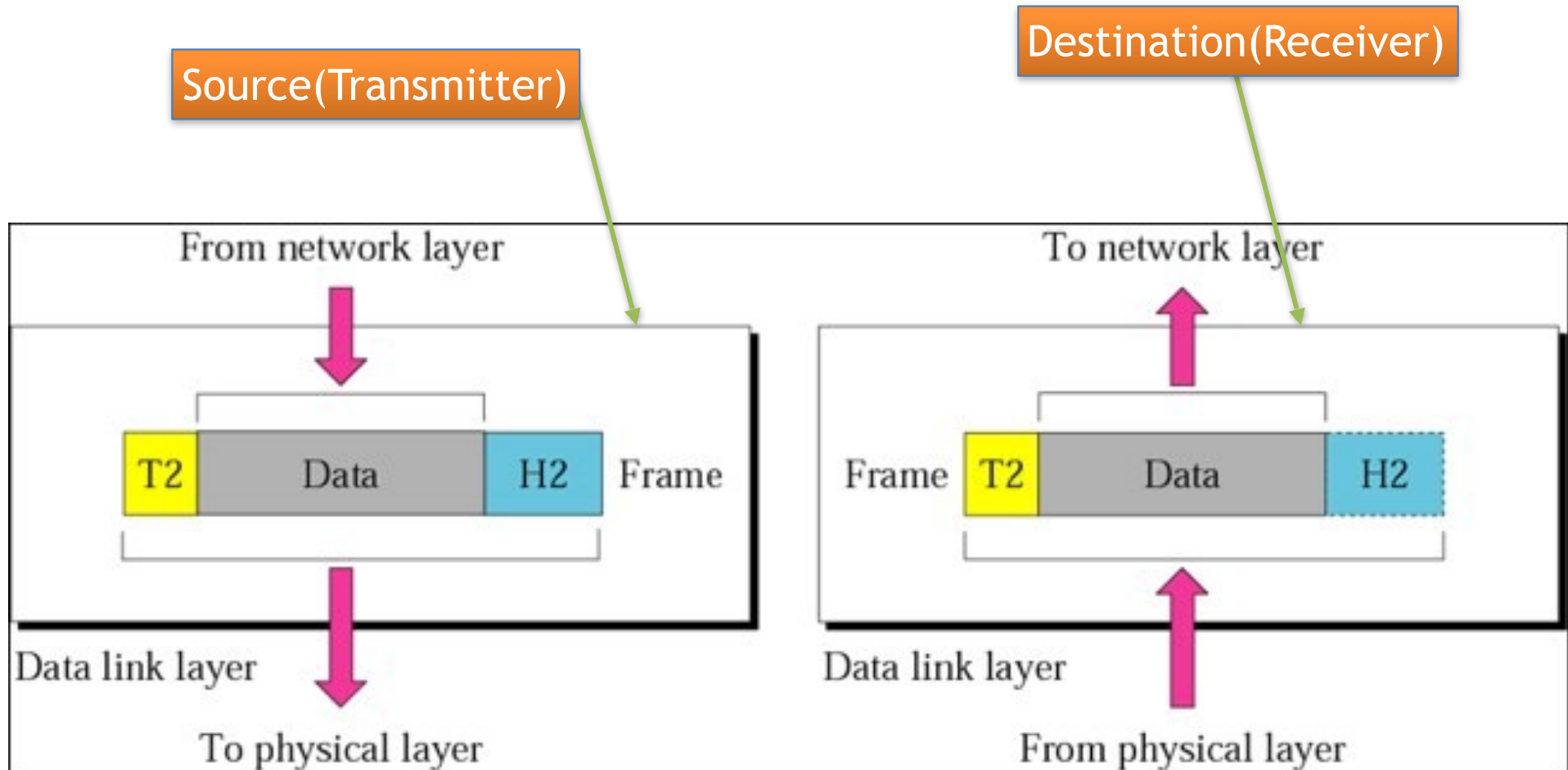
- **It transforms raw bits from physical layer, to a reliable link.**
- **It makes the physical layer appear error-free to the upper layer (network layer).**

Its responsibilities of the data link layer are

1. Framing

- **The data link layer divides the stream of bits received from the network layer into manageable data units called frames.**

relationship of the data link layer to the network and physical layers.



2. Physical addressing

- If frames are to be sent to different systems on the network, the data link layer **adds a header to the frame containing sender address and/or receiver address**

3. Flow control

- **Flow control means the source should not send data at a rate faster than the receiver can absorb it**

Data link layer has a flow control mechanism

4. Error control

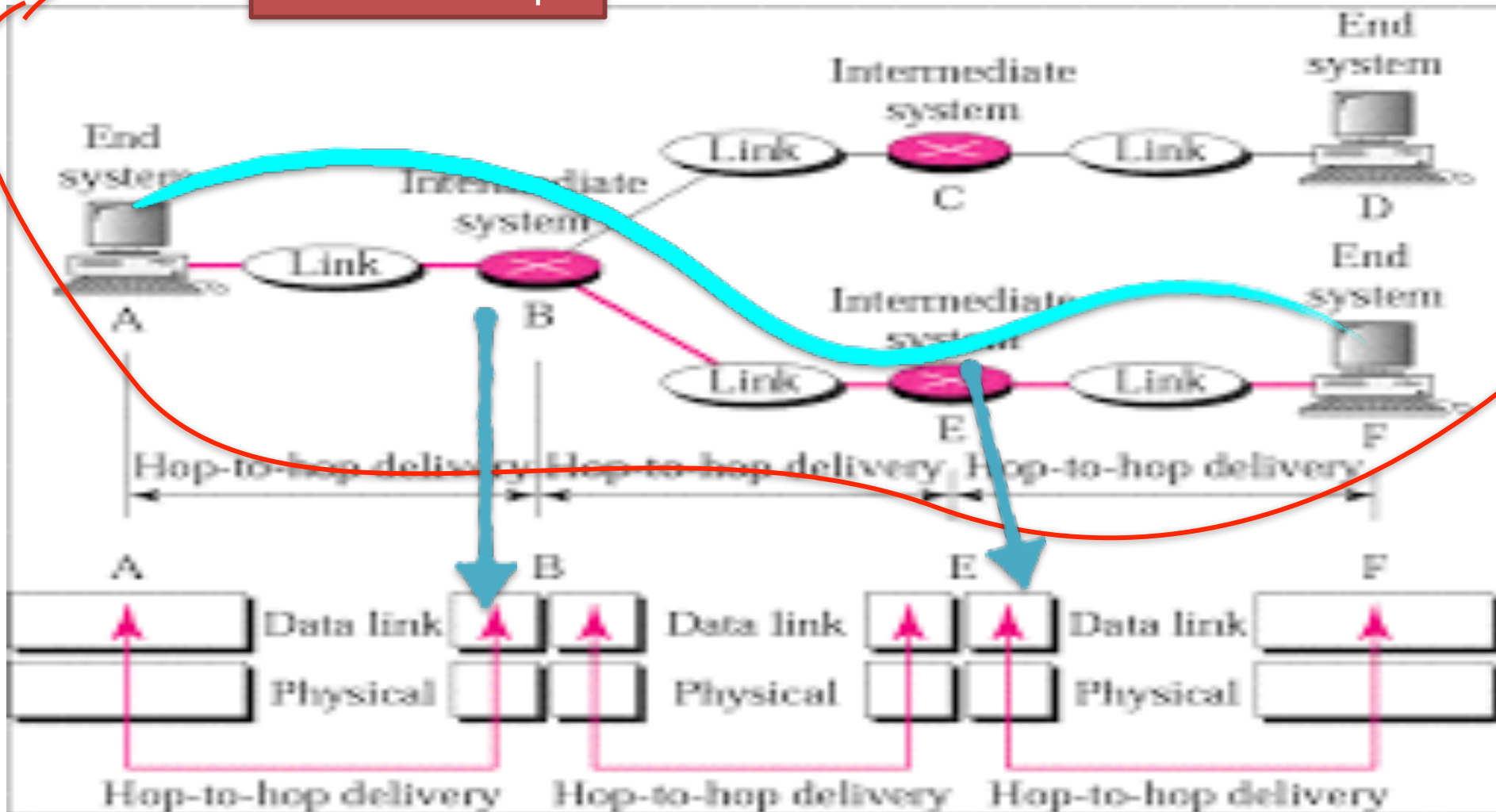
- mechanisms to **detect errors and retransmit damaged or lost frames.**
- also has mechanism to **recognize duplicate frames.**
- Error control is added in the trailer of the frame

5. Access control

- When **two or more devices are connected to the same link**, data link layer determines **which device has control over the link at any given time (ie who can use the transmission link at a time)**

Hop to hop delivery in data link layer

Network example



Network Layer

- The network layer is responsible for **the source-to-destination delivery of a packet, possibly across multiple networks (links).**

- the network layer ensures that **each packet gets from its point of origin to its final destination.**

Other responsibilities of the network layer include the following:

1. Logical addressing(IP address)

- If a packet passes the network boundary, we need another addressing system ie logical addressing to help distinguish the source and destination systems.

2. Routing.

- When independent networks or links are connected to create *intemetworks* (network of networks) or a large network

the connecting devices (called *routers* or *switches*)
route or switch the packets to their final destination.

3. **Congestion control**

- congestion occurs when there too many packets in the network- will cause bottleneck
- Includes mechanisms to control congestion

4. **Internetworking**

- Includes mechanisms to connect between heterogeneous networks

- It is responsible for **process-to-process delivery of the entire message.**

A process is an application program running on a host

- The transport layer, ensures that the **whole message arrives intact and in order, overseeing both error control and flow control at the source-to-destination level.**

Responsibilities of the transport layer include the following:

1. Service-point addressing.

- Computers often run several programs at the same time.
- For this reason, **source-to-destination delivery** means delivery not only from one computer to the next but also from a specific process (running program) on one computer to a specific process (running program) on the other.
- The transport layer header must therefore include a type of address called a **service-point address (or port address)**.
- The network layer gets each packet to the correct computer; the transport layer **gets the entire message to the correct process on that computer.**

2. Segmentation and reassembly

- A message may be **divided into segments**, with each segment containing a sequence number.
- These numbers **is used to reassemble the message correctly upon arriving at the destination** and to identify and replace packets that were lost in transmission.

3. Connection control

- Transport layer can be
 - **connectionless or**
 - **connection oriented.**

Connection control contd..

- **A connectionless transport layer treats each segment as an independent packet and delivers it to the transport layer at the destination machine.**
- **A connection oriented transport layer makes a connection with the transport layer at the destination machine first before delivering the packets.**
 - After all the data are transferred, the connection is terminated.

4. Flow control.

- Like the data link layer, the transport layer is responsible for flow control.
- flow control at this layer is **performed end to end** rather than across a single link

5. Error control.

- Like the data link layer, the transport layer is responsible for error control.
- However, error control at this layer is performed **process-to-process** rather than across a single link.
- The sending transport layer makes **sure that the entire message arrives at the receiving transport layer without error** (damage, loss, or duplication).
- Error correction is usually achieved through **retransmission**.

Session Layer

- The session layer is the **network dialog controller**.
- It **establishes, maintains, and synchronizes the interaction** among communicating systems.
- The session layer is responsible for **dialog control and synchronization**.

Specific responsibilities of the session layer include the following:

1. Dialog control

- The session layer allows two systems to enter into a dialog.
- It allows the communication between two processes to take place in either half duplex (one way at a time) or full-duplex (two ways at a time) mode.

2. Synchronization.

- The session layer allows a **process to add checkpoints, or synchronization points**, to a stream of data.
- For example
 - if a system is sending a file of 2000 pages, it is advisable to insert checkpoints after every 100 pages to ensure that each 100-page unit is received and acknowledged independently.
 - In this case, if a crash happens during the transmission of page 523, the only pages that need to be resent after system recovery are pages 501 to 523.
 - Pages previous to 501 need not be resent.

- It is concerned with the **syntax and semantics of the information exchanged between two systems**

1. Translation (encoding)

- The processes (running programs) in two systems are usually exchanging information in the form of character strings, numbers, and so on.
- The information **must be changed to bit streams** before being transmitted.

- Because different computers use different encoding systems, the presentation layer is **responsible for interoperability between these different encoding methods.**
- The presentation layer at the sender changes the information from its sender-dependent format **into a common format.**
- The presentation layer at the receiving machine changes the common format into its **receiver-dependent format.**

2. Encryption

To carry sensitive information, a system must be able to ensure privacy.

- Encryption **means** that the sender transforms the original information to another form and sends the resulting message out over the network.
- **Decryption** reverses the original process to transform the message back to its original form.

3. **Compression**

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

- Data compression becomes particularly important in the transmission of multimedia such as text, audio, and video.

Application Layer

- **The application layer enables the user, to access the network.**
- It provides user interfaces and support for services such as **electronic mail, remote file access and transfer, shared database management, and other types of distributed information services.**

Specific services provided by the application layer include the following:

1. Network virtual terminal

- It is a software version of a physical terminal, and it allows a user to log on to a remote host.
- To do so, the application creates a software emulation of a terminal at the remote host.
- The user's computer talks to the software terminal which, in turn, talks to the host, and vice versa.
- The remote host believes it is communicating with one of its own terminals and allows the user to log on.

2. File transfer, access, and management

- This application allows a **user to access files in a remote host (to make changes or read data), to retrieve files from a remote computer** for use in the local computer, and to manage or control files in a remote computer locally.

3. Mail services

This application provides the basis for e-mail forwarding and storage

4. Directory services

- This application provides distributed database sources and access for global information about various objects and services.

THANK YOU



Unit 1 contd..

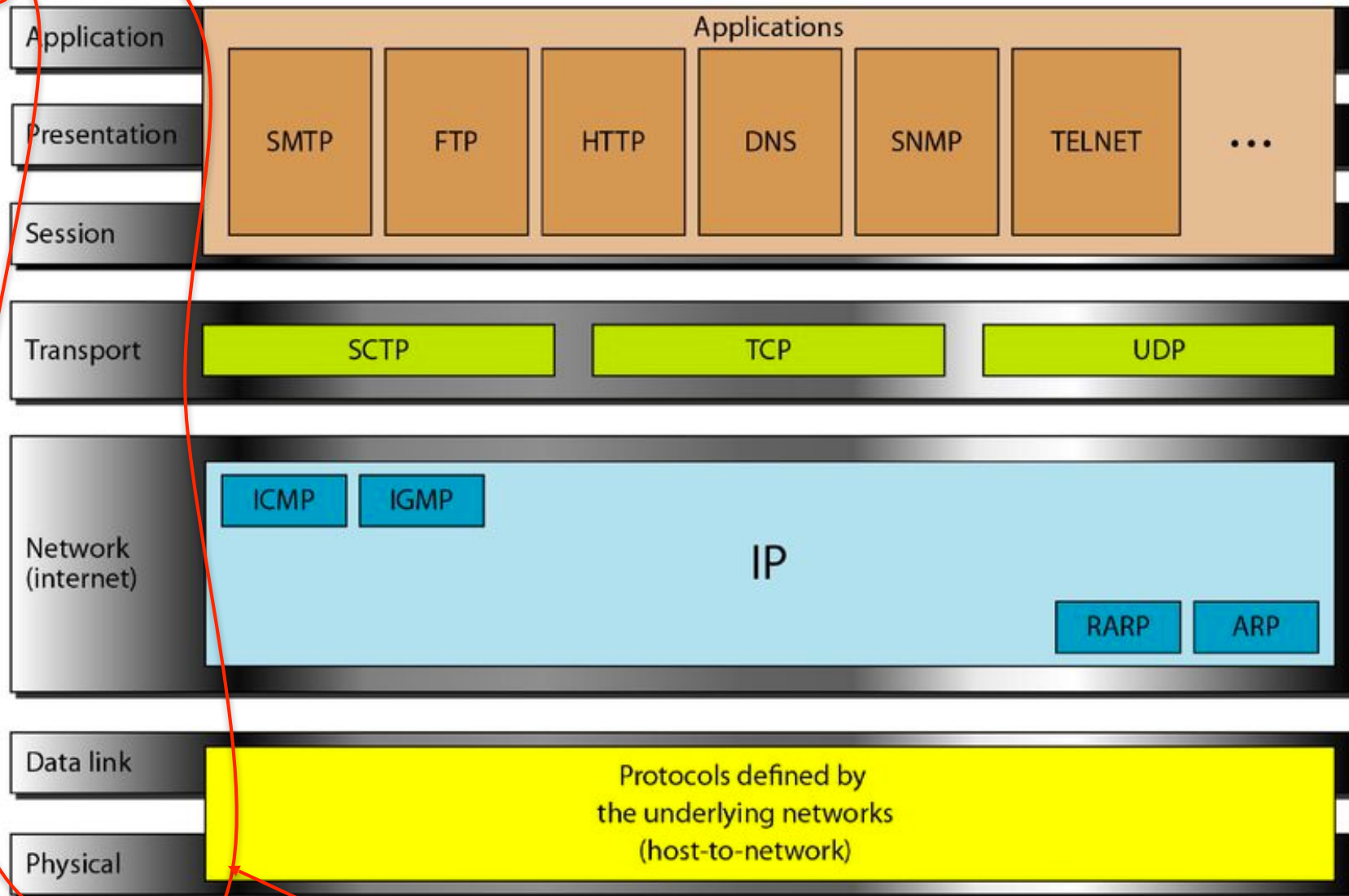
TCP/IP Protocol suite
Switching Techniques

Source: The TCP/IP Protocol Suite, Behrouz A. Forouzan

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**.
- When TCP/IP is compared to OSI, we can say that the TCP/IP protocol suite is made of five layers:
 1. **Physical layer**
 2. **Data link layer**
 3. **Network layer**
 4. **Transport layer**
 5. **Application layer**

TCP/IP Protocol Suite



ISO - OSI

TCP/IP PROTOCOL SUITE

- 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 one or more lower-level protocols.

- The layers contain relatively independent protocols.

TCP/IP PROTOCOL contd..

Physical and Data Link Layers

- No specific protocol is defined
- TCP/IP model supports all the standard and proprietary protocols.
- For instance, network in a TCP/IP internetwork can be a local-area network or a wide-area network.

Network layer (internetwork layer)

- ❖ supports the Internetworking Protocol (IP)

- ❖ some other protocols that support data movement
 - ◆ ARP
 - ◆ RARP
 - ◆ ICMP
 - ◆ IGMP.

Internetworking Protocol (IP)

- Most important protocol of the TCP/IP network
- Implements internetworking
- IP is an unreliable and connectionless protocol- a best-effort delivery.
- It is host-to-host protocol.

TCP/IP PROTOCOL contd..

Address Resolution Protocol (ARP)

- It is used to find the physical address (NIC) of the node after its internet address is known.

Reverse Address Resolution Protocol (RARP)

- It is used to find the Internet address of the node after its physical address is known.

Internet Control Message Protocol (ICMP)

- It is used by hosts and gateways to send notification of datagrams (packets) with a problem back to the sender.

Internet Group Message Protocol (IGMP)

- It is used to facilitate the simultaneous transmission of messages to a group of recipients.

TCP/IP PROTOCOL contd..

Transport Layer

- the protocol is responsible for delivery of message from a process to another process.

Protocols used at this layer

1. TCP

2. UDP

3. SCTP

TCP/IP PROTOCOL contd..

- **User Datagram Protocol (UDP)**

- It adds port addresses, checksum error control, and length information to the data from the upper layer.

Transmission Control Protocol (TCP)

- It is reliable and connection-oriented.



Stream Control Transmission Protocol (STCP)

- It supports the **newer application e.g. voice over the Internet.**
- It combine **best features of UDP and TCP.**

Application Layer

- The application layer in TCP/IP is equivalent to the combined session, presentation, and application.

Switching Techniques

Switching is process to forward packets coming in from one port to a port leading towards the destination.

3 types of switching techniques

1. Circuit switching
2. Message switching
3. Packet switching

Switching Methods

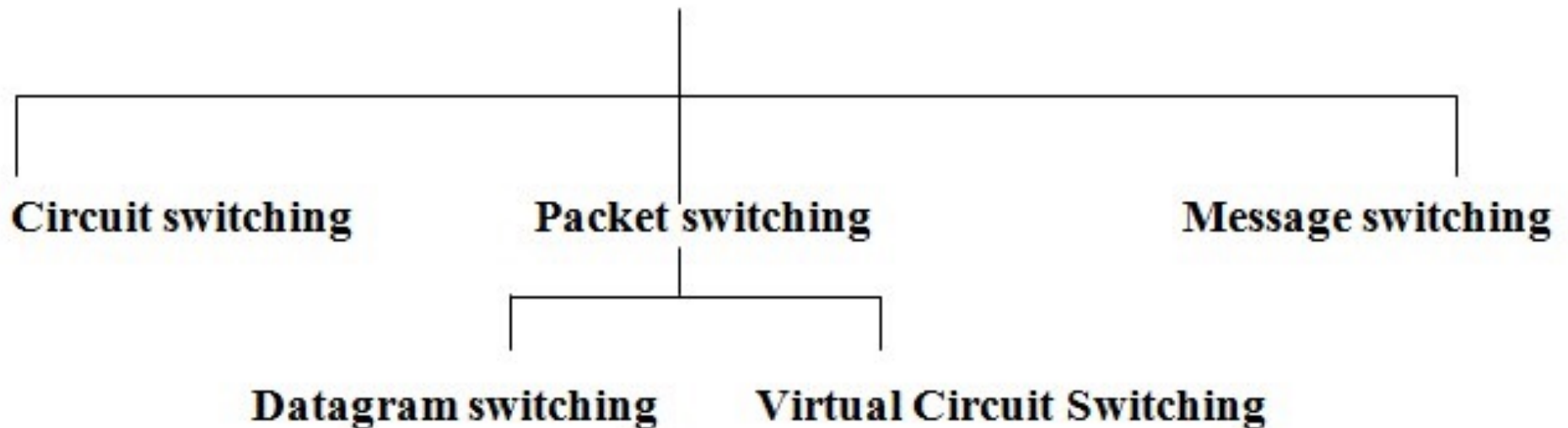
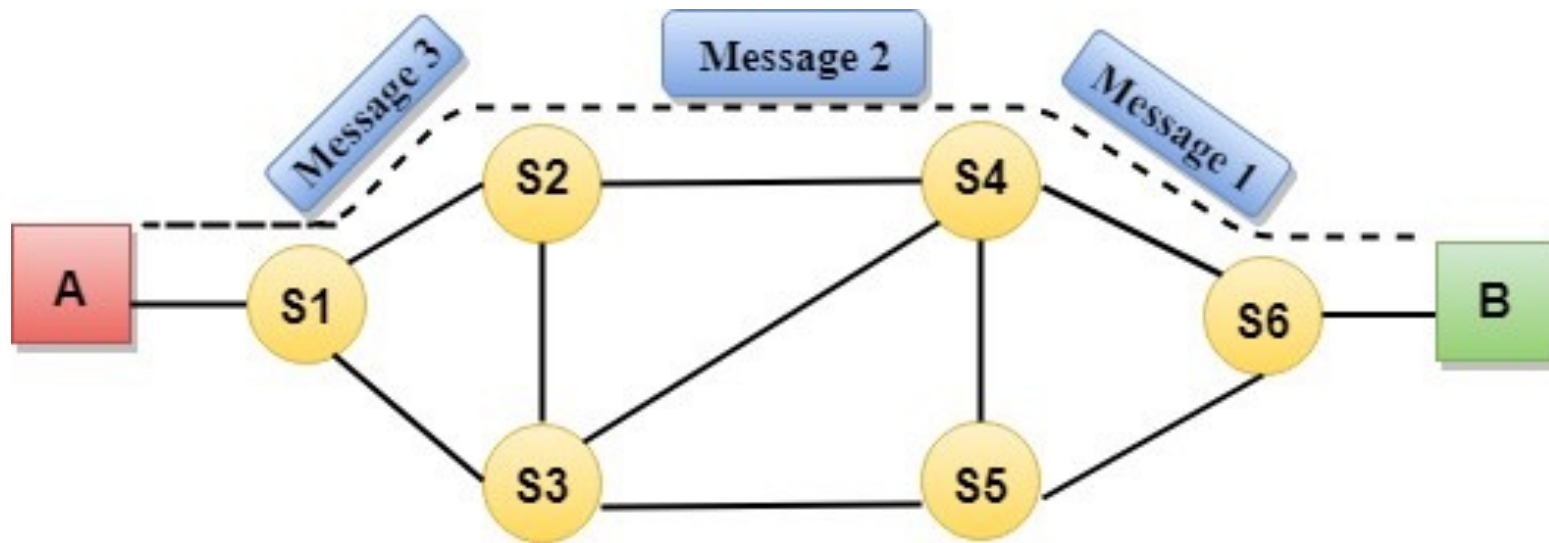


Fig- Types of switching methods

Circuit Switching

- When two nodes communicate with each other over a dedicated communication path, it is called **circuit switching**.
- **In** this, to transfer the data, circuit must be established before data transfer
- Circuits can be permanent or temporary.
- It has three phases:
 - **Establish a circuit**
 - **Transfer the data**
 - **Disconnect the circuit**

Circuit switching is primarily used in Telephone networks and not in Computer networks



Advantages:

The communication channel is a dedicated link

Disadvantages:

- **More bandwidth is required.**
- **Connection setup time is more.**
- **More expensive than any other switching techniques because a dedicated path is required for each connection.**
- **Inefficient use of communication channel(wastage of channel capacity)**

Message Switching(store & forward)

- **whole message is treated as a data unit and is transferred in its entirety.**
- A node, first **receives the whole message and buffers** it until there are resources available to transfer it to the next node.
- If the next node is not having enough resource to accommodate large size message, the message is stored temporarily and node waits.

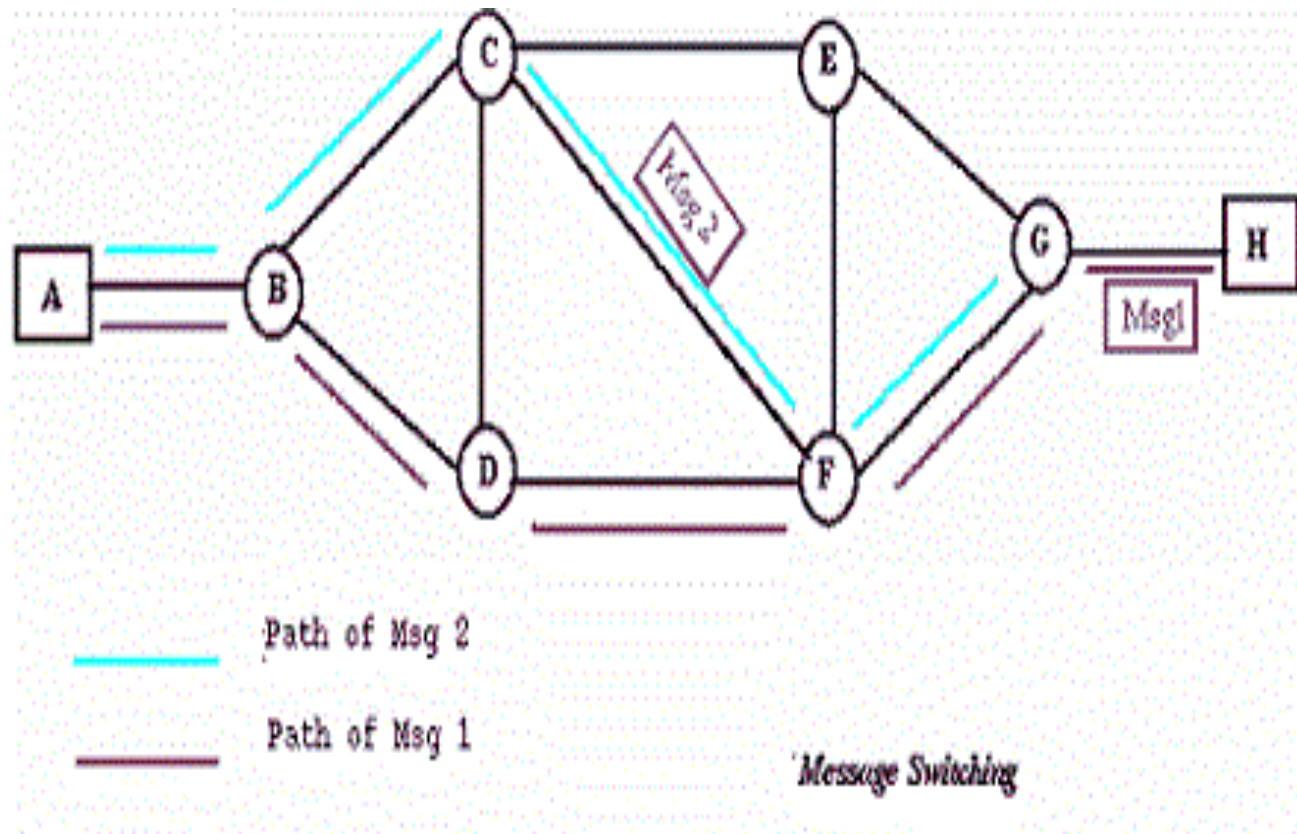


Fig- Message Switching



Data channels are shared among the communicating devices that improve the **efficiency of using available bandwidth.**



Traffic congestion can be reduced because the message is temporarily stored in the nodes.



Message priority can be used to manage the network.



The size of the message which is sent over the network can be varied. Therefore, it supports the data of unlimited size.

Message switching has the **following drawbacks:**

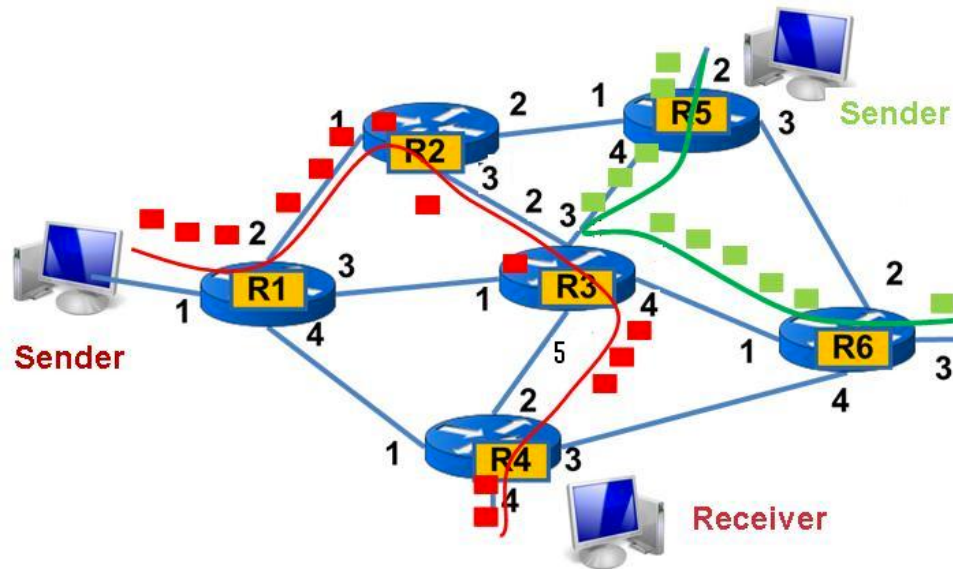
- Every node in transit path **needs enough storage to accommodate entire message.**
- Because of store-and-forward technique and waits included until resources are available, message switching **is very slow.**
- Message switching was **not suitable for streaming media and real-time applications.**

Packet switching

- The entire message is broken down into smaller chunks called packets.
- The switching information is added in the header of each packet and transmitted independently.
- It is easier for intermediate networking devices to store small size packets and they do not take much resources either on carrier path or in the memory of switches.
- Packet switching is the switching method used in data networks for computer communication.
- packet switching has two approaches:
 1. *Virtual Circuit*
 2. *Datagram*

Virtual Circuit Packet Switching

- virtual circuit approach has a **set up, data transfer and disconnect phases**.
- **Resources can be allocated during the set up phase, as in circuit switched networks or on demand, as in a datagram network.**
- All the packets of a message follow the same path established during the connection
- Every packet contains the **virtual circuit number(or label)**.
- As in circuit switching, virtual circuit needs call setup before actual transmission can be started.
- Routing is based on the virtual circuit number.
- Once such a path is identified, all packets of the data session/flow **MUST** follow the same path.
- But this path is not reserved for this session alone and multiple sessions can share the links in this path.



a sample VC based switched WAN topology involving six routers(R1 to R6), with two different flows.The first flow is shown in red colour and is through the path R1-R2-R3-R4. The second flow is shown in green colour and is through the path R5-R3-R6.

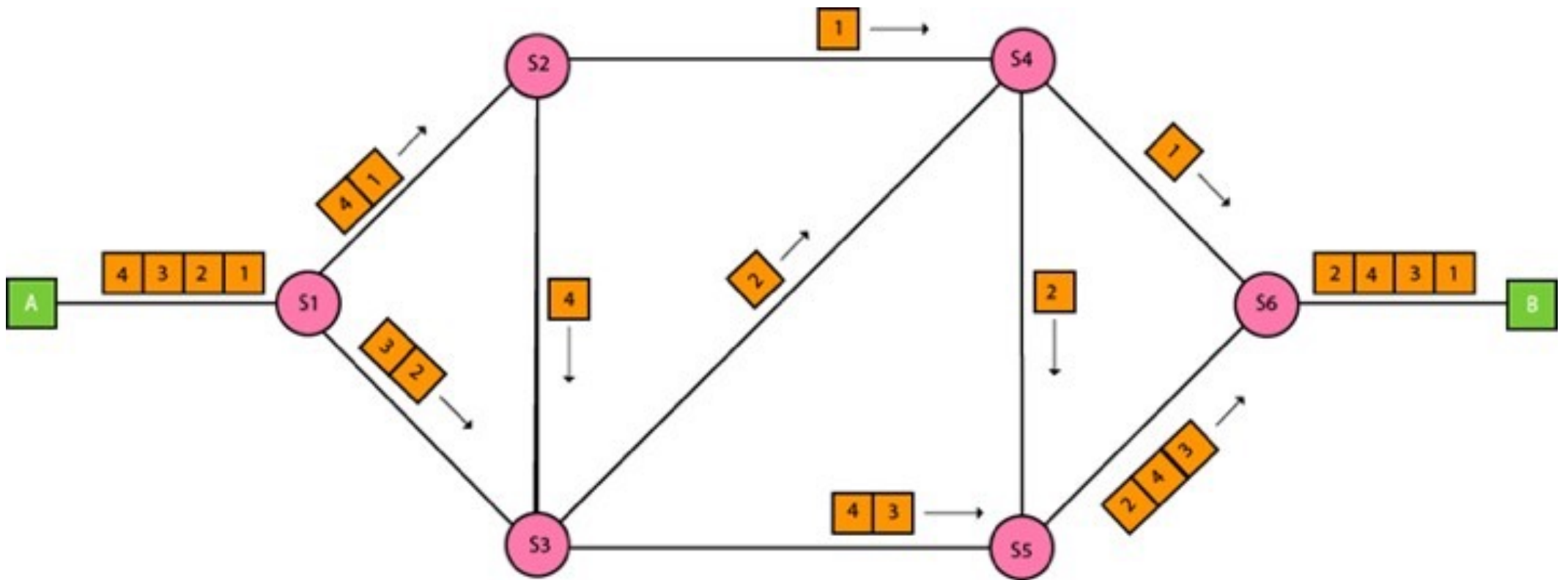
In VC table, the first entry is for the flow shown in red colour. It states that if a packet comes to R3 via. link 2 and with label 45, then R3 has to change the label in the packet to value 33 and send it via. its link 5. Similarly, the second entry in the table is for the flow shown in green colour. It states that if a packet comes to R3 via. link 3 and with label 22, then R3 has to change the label in the packet to value 24 and send it via. its link 4.

Incoming Link	Incoming Label	Outgoing Link	Outgoing Label
2 ■	45 ■	5 ■	33 ■
3 ■	22 ■	4 ■	24 ■

SAMPLE VC TABLE AT ROUTER R3

Datagram Packet Switching

- Every packet is treated as **individual, independent transmission**
- Even if a packet is a part of multi-packet transmission, the network treats it as though it existed alone.
- Packets in this approach are called **datagrams**.



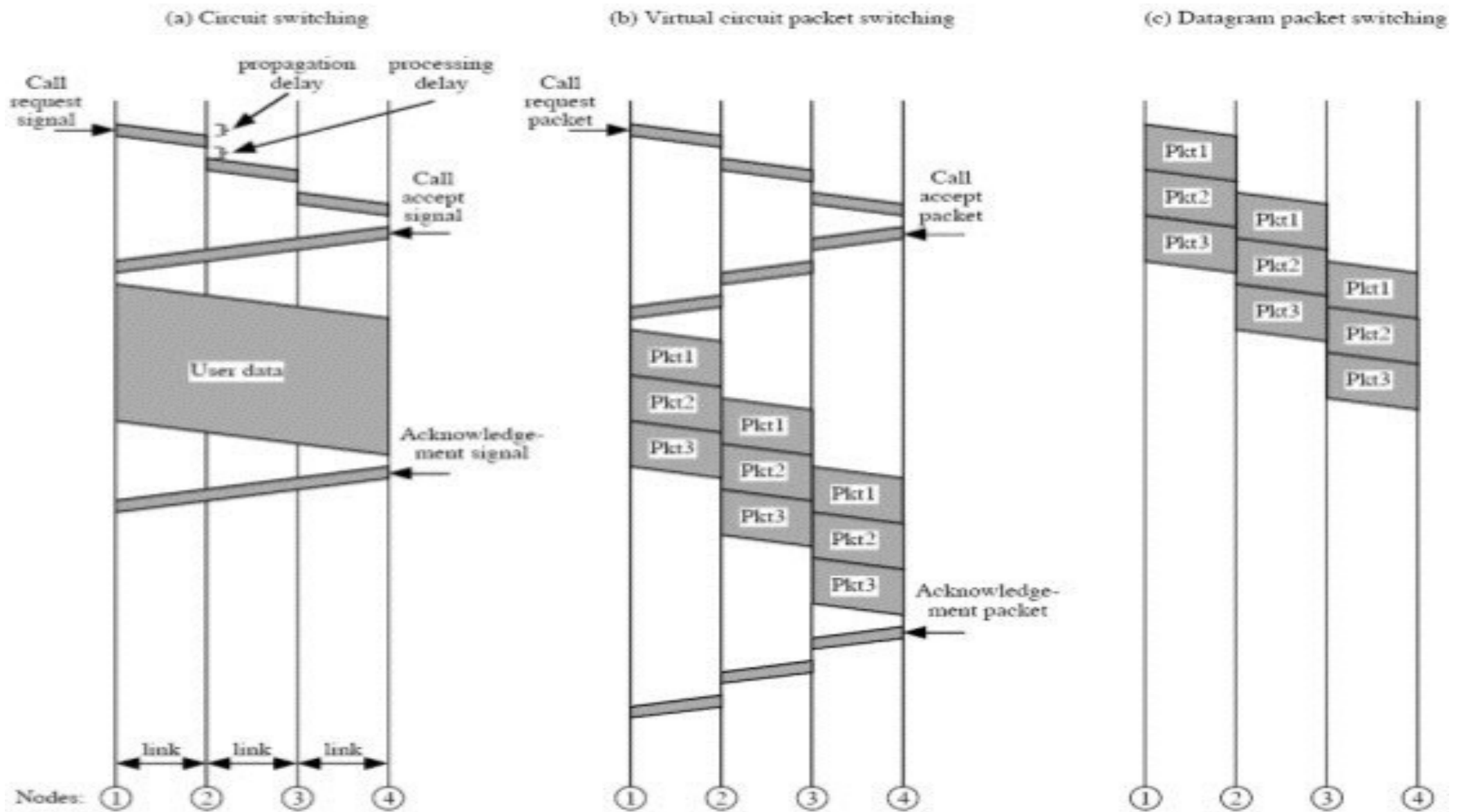


Figure 10.15 Event Timing for Circuit Switching and Packet Switching

Circuit Switching

Dedicated Transmission Path

Continous transmission of data

Fast enough for interactive

Messages are not stored

The path is established for entire conversation

Call setup delay, negligible transmission delay

Busy signal if called party busy

Overload may block call setup

Computerised switching nodes

User responsible for message loss protection

No speed or code conversion

Fixed bandwidth

No overhead bits after call setup

Datagram Packet Switching

No Dedicated Path

Transmission of packets

Fast enough for interactive

Packets may be stored until delivered

Route established for each packet

Packet transmission delay

Sender may be notified if packet not delivered

Overload increases packet delay

Small switching nodes

Network may be responsible for individual packets

Speed and code conversion

Dynamic use of bandwidth

Overhead bits in each packet

Virtual Circuit Packet Switching

No Dedicated Path

Transmission of packets

Fast enough for interactive

Packets stored until delivered

Route established for entire conversation

Call setup delay, packet transmission delay

Sender notified of connection denial

Overload may block call setup

Small switching nodes

Networks may responsible for packet sequences

Speed and code conversion

Dynamic use of bandwidth

Overhead bits in each packet

UNIT 1 (PART3)

INTERFACE STANDARDS

Ceena Mathews
PNC

Interface standards

- An **interface** is used between dissimilar (non-peer) entities and involves the direct physical transfer of data.
- **interconnection media** provides the physical transmission path for electrical or optical signals and is not concerned with the communications protocol.
- It must have an **agreed interface which specifies the design, dimensions and pin assignments of the connector plus the signalling voltages and signalling sequences** for data transmission and control.

Physical interface standards commonly have **four parts**:

- 1) Mechanical specifications** for the cable and connectors.
- 2) Electrical specifications** including voltages impedances and waveforms.
- 3) Functional specifications** including signal-pin assignments and signal definitions.
- 4) Procedural specifications** for control and data transfer.

RS-232-C or EIA 232

- The most popular **serial interface standard, RS-232**, was published in 1962 and was revised in 1969, 1972 and 1987.

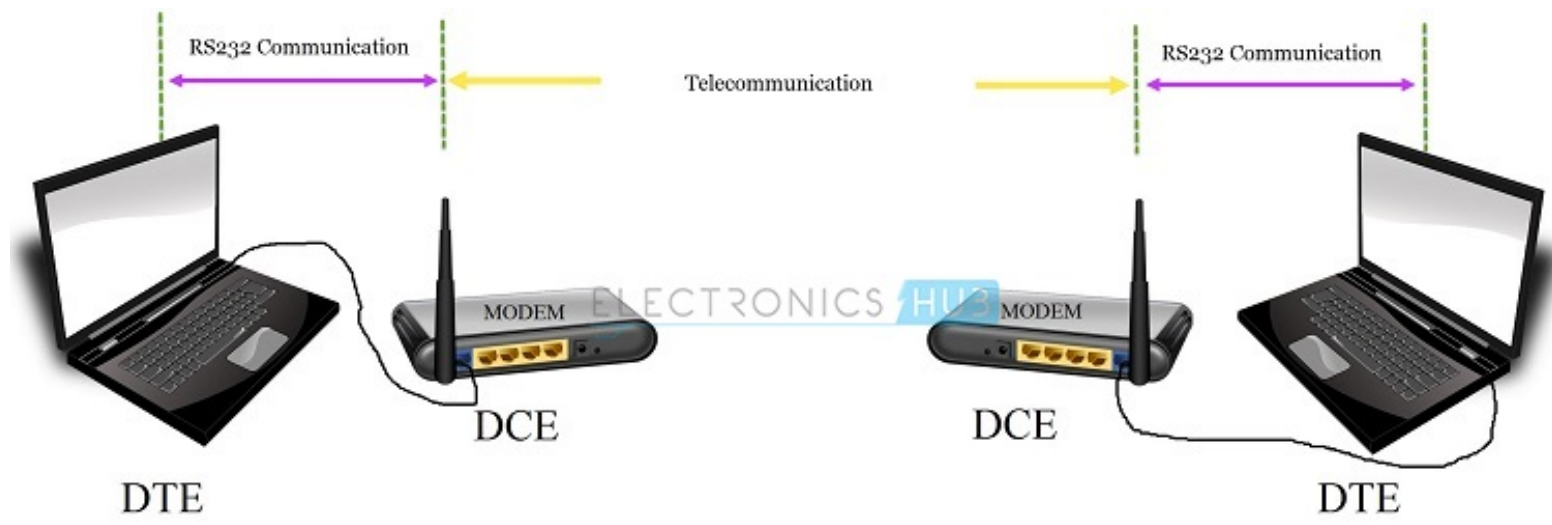
- RS stands for 'recommended standard'

was published by the **American Electronic Industries Association** to **standardise the interface between data terminal equipment (DTE) and data circuit-terminating (DCE) equipment employing serial binary data interchange.**

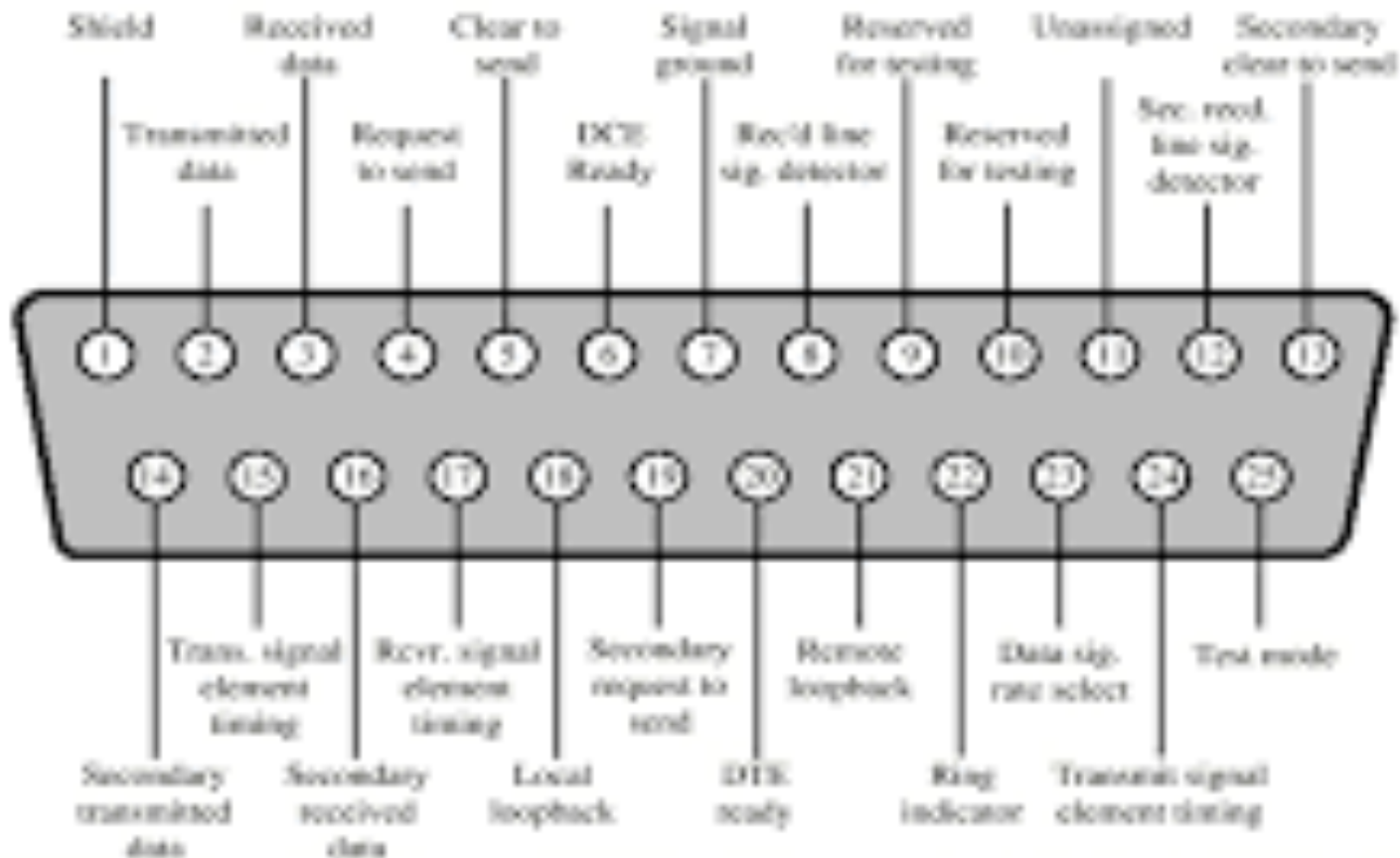
-

Connector

- The connector is a 25-pin D-type
- Only 21 pins are assigned to interchange circuits.



RS-232 DB-25 Pinouts

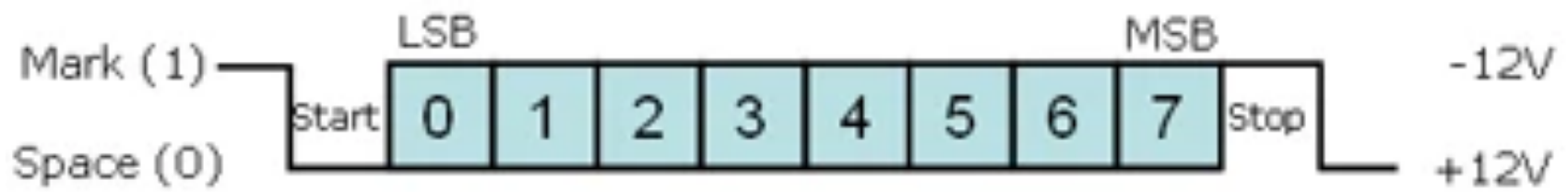


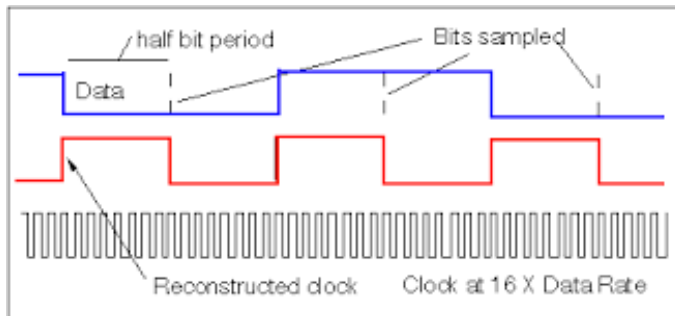


<i>DTE PIN NO</i>	<i>CIRCUIT NAME</i>	<i>ABBREVIATION</i>	<i>SOURCE</i>
1	Protective earth		
2	Transmitted data	TxD	DTE
3	Received data	RxD	DCE
4	Request to send	RTS	DTE
5	Clear to send	CTS	DCE
6	Data set ready	DSR	DCE
7	Signal ground		
8	Data carrier detect	DCD	DCE
9	Reserved for data set test		
10	Reserved for data set test		
11	Unassigned		
12	Secondary received signal detector		DCE
13	Secondary clear to send		DCE
14	Secondary transmitted data		DTE
15	Transmission signal element timing		DCE
16	Secondary received data		DCE
17	Receive signal element timing		DCE
18	Unassigned		
19	Secondary request to send		DTE
20	Data terminal ready	DTR	DTE
21	Signal quality detector		DCE
22	Ring indicator		DCE
23	Data signal rate selector		DCE/DTE
24	Transmit signal element timing		DTE
25	Unassigned		

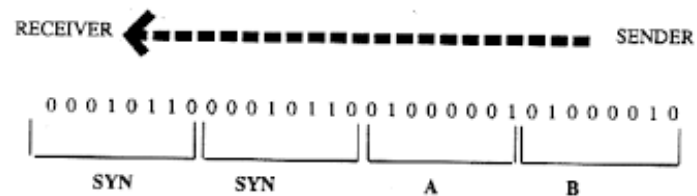
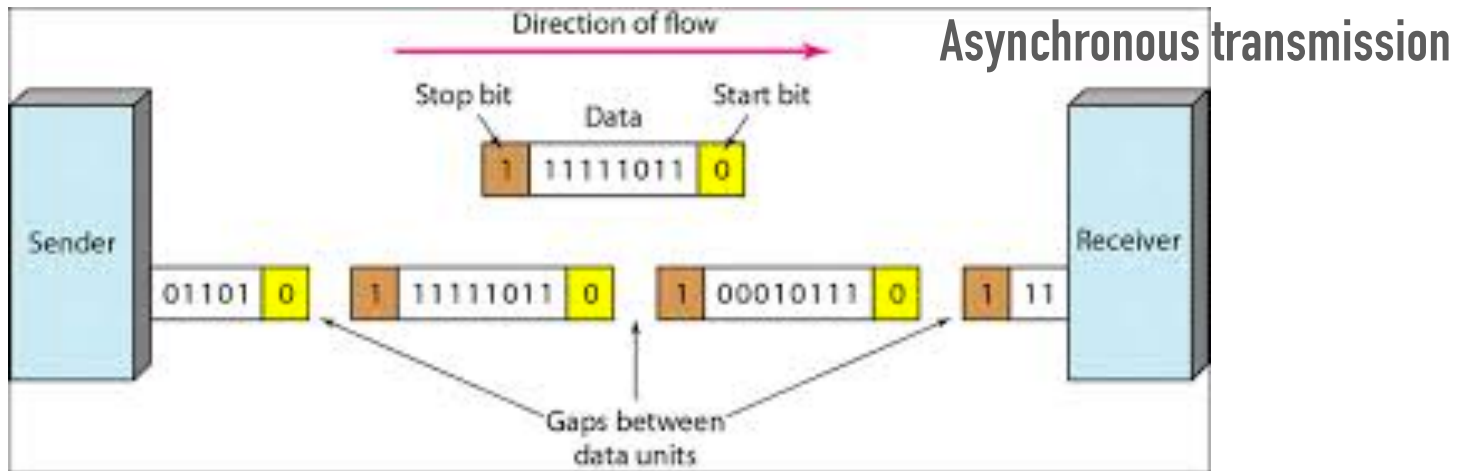
The electrical interface

- uses unbalanced circuits, all referenced to the signal ground.
- A positive signal between **3 and 15 V** is interpreted as **on for control circuits** or as **binary 0 (space) for data circuits**.
- A negative voltage between **- 3 and - 15 V** is interpreted as **off or binary 1 (mark)**.
- **The use of relatively high voltages and unbalanced circuits** makes RS-232-C highly susceptible to crosstalk, common mode noise (noise which is injected into the signal line and the signal ground), and differences in ground potential at the transmitter and receiver.
- The **maximum data rate** specified is 20 kbit/s
- **maximum cable length** specified is 50 ft.





For reference only



SYNCHRONOUS SERIAL TRANSMISSION

- RS-232-C is **only suitable for point-to-point interfaces**, although the DCE may support multipoint operation.
- **Synchronous or asynchronous baseband signalling** may be used.
- **Synchronous operation** is supported by **Transmitter Signal Element Timing signal (clock from the DTE)**, or **Transmitter Signal Element Timing (clock from DCE)** and **Receiver Signal Element Timing (clock from DCE)** interchange signals.
- Modems with out-of-band signalling channels may use **Secondary Transmitted Data and Secondary Received Data**.
- Other circuits are provided to **change the modem operating speed and to control automatic calling equipment**.

There are four types of line defined in the RS232 specification.

Data
Control
Timing
Ground

1. data lines

- There are two of these, one for data travelling in each direction.
- Transmit data is carried on pin 2 and the receive data is carried on pin 3.
- ✻ **Transmitted Data (TxD):** One of two separate data signals, this signal is generated by the DTE and received by the DCE.
- ✻ **Received Data (RxD):** second of two separate data signals, this signals is generated by the DCE and received by the DTE.

2. Control lines

1. RLSD (Received Line Signal Detector) or Data Carrier Detect:

- when the modem has detected a carrier on the telephone line and a connection appears to have been made.
- It produces a high, which is maintained until the connection is lost.

2. Data Terminal Ready (DTR):

- DTR indicates the readiness of the DTE.
- This signal is turned ON by the DTE when it is ready to transmit or receive data from the DCE.
- DTR must be ON before the DCE can assert DSR.

3. Data Set Ready (DSR):

- This signal is turned on by the DCE to indicate that it is connected to the telecommunications line

4. Ring Indicator (RI):

when asserted, indicates that a ringing signal is being received on the communications channel

5. Request to Send (RTS):

When the host system (DTE) is ready to transmit data to the peripheral system (DCE), RTS is turned ON.

After RTS is asserted, the DCE must assert CTS before communication can commence.

6. Clear to Send (CTS):

CTS is used along with RTS to provide handshaking between the DTE and the DCE. After the DCE sees an asserted RTS, it turns CTS ON when it is ready to begin communication.

RS232 Protocol defines four signals for the purpose of Handshaking:

1. Ready to Send (RTS)
2. Clear to Send (CTS)
3. Data Terminal Ready (DTR) and
4. Data Set Ready (DSR)

there are two types of channels that are specified in the RS-232 specification.

- **primary channels** :are normally used, and operate at the normal or higher data rates.
- **secondary channel** for providing control information. If it is used it will usually send data at a much slower rate than the primary channel.

- There is a secondary channel which includes a duplicate set of flow-control signals.

- This secondary channel provides for management of the remote modem, enabling baud rates to be changed on the fly, retransmission to be requested if a parity error is detected, and other control functions.

-

- This secondary channel, when used, is typically set to operate at a very low baud rate in comparison with the primary channel to ensure reliability in the control path.

- In addition, it may operate as either a simplex, half-duplex, or full-duplex channel, depending on the capabilities of the modem.

➤ **Secondary Communications Channel**

1. **Secondary Transmitted Data (STxD)**
2. **Secondary Received Data (SRxD)**
3. **Secondary Request to Send (SRTS)**
4. **Secondary Clear to Send (SCTS)**

➤ These signals are equivalent to the corresponding signals in the primary communications channel.

Ground connections

- are also important.
- There are two ground connections
 1. **protective ground**
 - ensures that both equipments are at the same earth potential.
 - This is very useful when there is a possibility that either equipment is not earthed.
 2. **signal ground** is used as the return for the digital signals travelling along the data link.

➤ Transmitter and Receiver Timing Signals

1. Transmitter Signal Element Timing(also called Transmitter Clock)

- This signal is relevant only when the DCE device is a modem and is operating with a synchronous protocol.
- The **modem generates this clock signal** to control exactly the rate at which data is sent on Transmitted Data (pin 2) from the DTE device to the DCE device.
- The **logic '1' to logic '0' (negative voltage to positive voltage) transition** on this line causes a corresponding transition to the next data element on the Transmitted Data line.
- The modem generates this signal continuously, except when it is performing internal diagnostic functions.

2. Receiver Signal Element Timing (also called Receiver Clock)

This signal is similar to TC described above, except that it provides timing information for the DTE receiver.

3. Transmitter Signal Element Timing (also called External Transmitter Clock)

- **Timing signals are provided by the DTE device for use by a modem.**
- This signal is used only when TC and RC (pins 15 and 17) are not in use.
- The logic '1' to logic '0' transition (negative voltage to positive voltage) indicates the time-center of the data element.
- Timing signals will be provided whenever the DTE is turned on, regardless of other signal conditions.

➤ Limitations of RS232

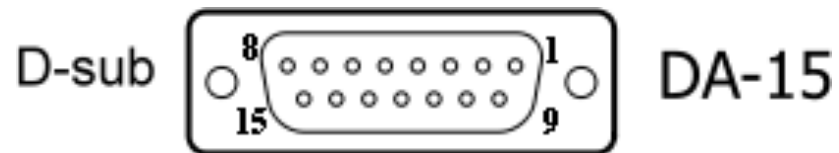
- RS232 Protocol **requires a common ground** between the transmitter (DTE) and receiver (DCE). Hence, the reason for shorter cables between DTE and DCE in RS232 Protocol.
- The signal in the line is **highly susceptible to noise**. The noise can be either internal or external.
- If there is an increase in baud rate and length of the cable, there is a chance of **cross talk introduced** by the capacitance between the cables.
- The voltage levels in **RS232 are not compatible with** modern TTL or CMOS logics. We need an external level converter.

➤ Applications

- Though RS232 is a very famous serial communication protocol, it is now has been replaced with advanced protocols like USB.
- Previously they we used for serial terminals like Mouse, Modem etc.
- But, RS232 is still being used in some Servo Controllers, CNC Machines, PLC machines and some microcontroller boards use RS232 Protocol.

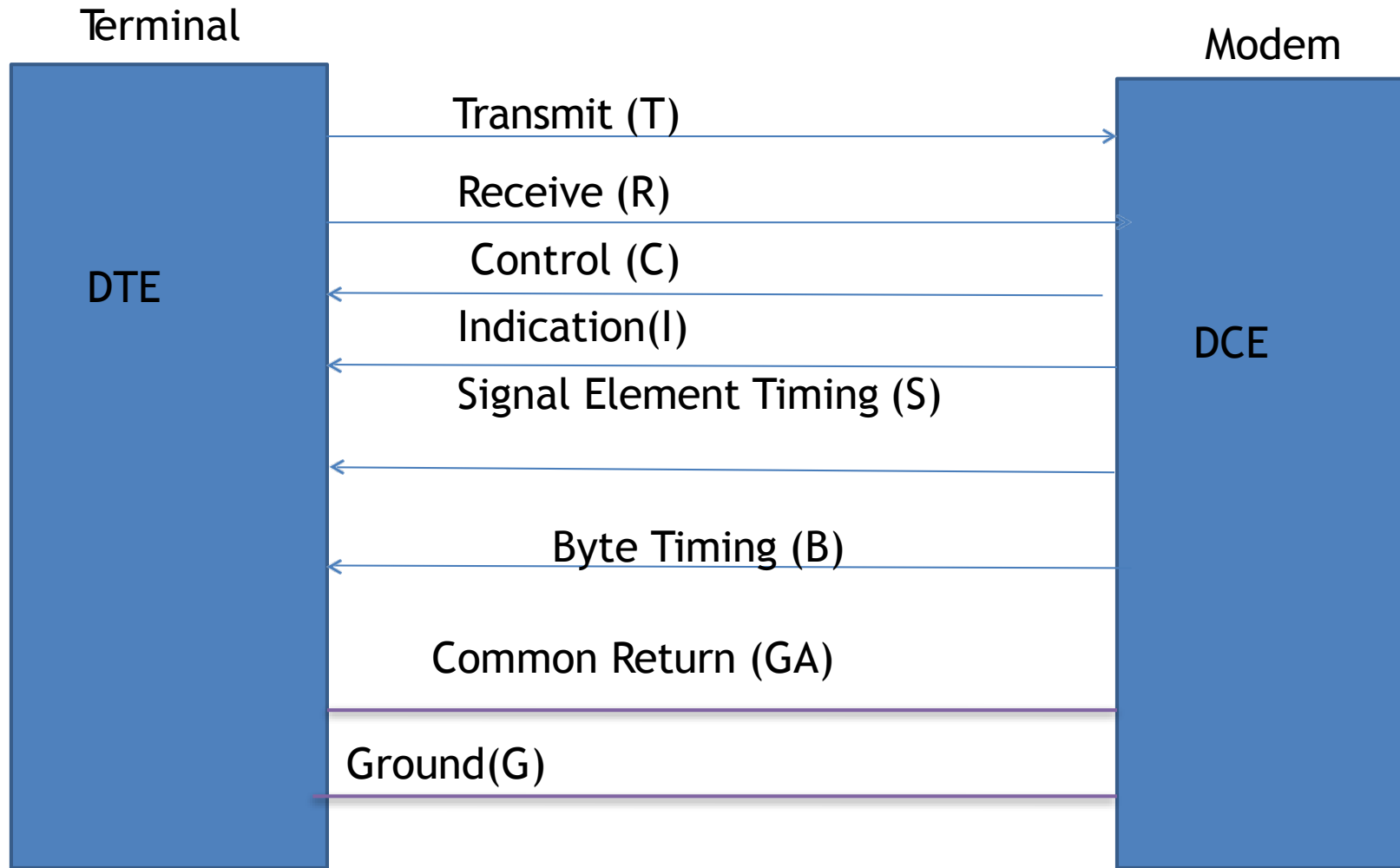
X.21

- **CCITT** (Consultative Committee for International Telephony and Telegraphy) standard X.21
- is a general-purpose **interface between data terminal equipment and data circuit-terminating equipment for synchronous operation on public data networks**
- use a 15-pin connector



- **control signals use the same circuits as user data**, rather than separate pins.
- **A separate control circuit identifies data or control signals.**
- This facility **reduces the number of pins and permits future extension of control signals.**
- X.21 not only **specifies the physical interface** but also specifies **data link and network layer functions for circuit switched networks;**
- X.21 has been adopted by many **circuit switched public digital networks throughout the world.**

Functions of the x.21 pins



- **Signal ground (G)**
 - Reference signal used to evaluate the logic states of the other signals.
 - This signal can be connected to the protective earth (ground).
- **DTE common return**
 - *Reference ground signal for the DCE interface.*
 - This signal is used only in **unbalanced mode**.
- **Transmit (T)**
 - The DTE sends data and control information on the Transmit line

- **Receive (R)**
 - DCE uses the Receive line for data and control with the Indication line to differentiate data from control

- **Control (C)**
 - DTE-controlled signal that controls the transmission on an X.21 link.
 - This signal must be on during data transfer, and can be on or off during call-control phases

- Indication (I)
 - DCE-controlled signal that controls the transmission on an X.21 link.
 - This signal must be on during data transfer, and can be on or off during call-control phases.

- Signal Element Timing (S)
 - Clocking signal that is generated by the DCE.
 - This signal specifies when sampling on the line must occur.
 - It provides bit timing from the DCE

- **Byte Timing (B)**
 - Binary signal that is on when data or call-control information is being sampled.
 - When an 8-byte transmission is over, this signal switches to off.
 - may optionally be used to group bits into 8-bit frames.
 - **The DTE must begin each character on a frame boundary.**
 - **If the Byte Timing line is not used, each control sequence must be preceded by at least two SYN characters to identify frame boundaries.**
- **Transmissions across an X.21 link require both the DCE and DTE devices to be in a ready state, indicated by an all 1s transmission on the T and R signals**

Each call on a circuit switched network goes through four phases:

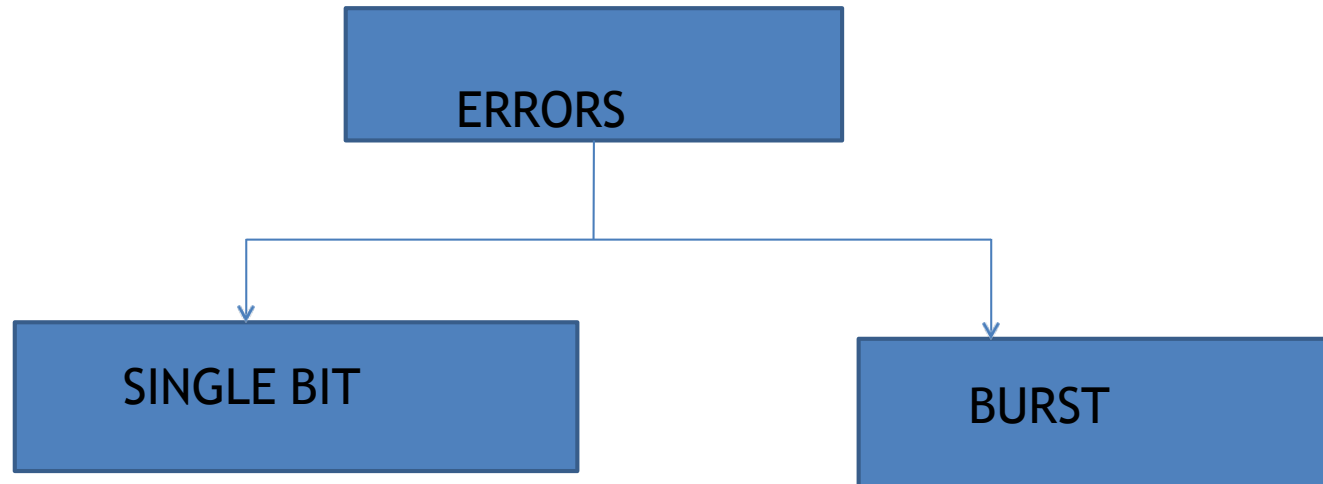
- 1) The **quiescent phase** before transmission or reception.
- 2) The **call establishment phase** in which the DTE establishes a circuit switched connection.
- 3) The **data transfer phase** in which a full-duplex transmission path is maintained.
- 4) The **clearing phase**, initiated by either the DTE or the network, in which the connection is released.

Unit II :

Error Detection and Correction

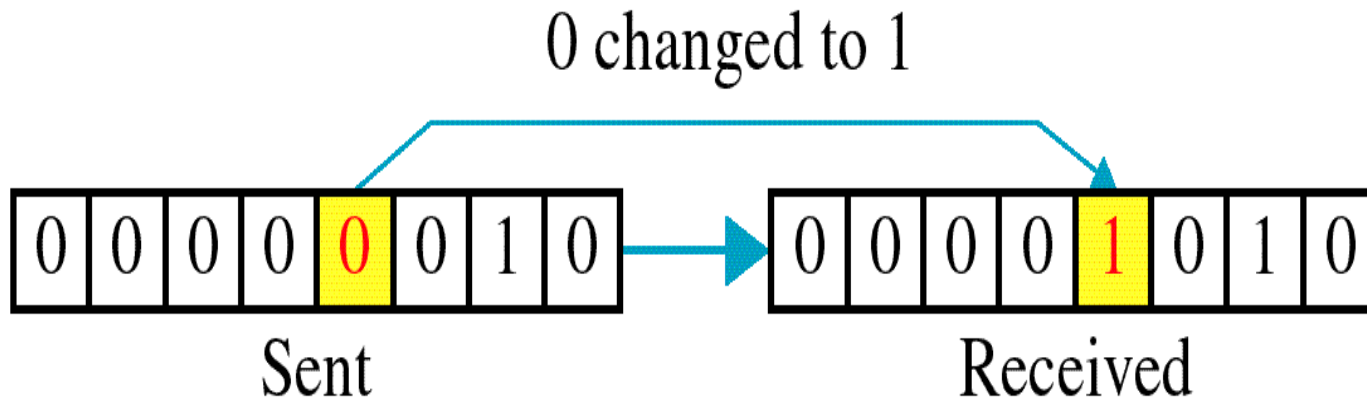
- Process of detecting and correcting errors
- Implemented at the data link layer and transport layer

Types of Errors



Single-bit error

Only 1 bit in the data unit is changed from 1 to 0 or 0 to 1



Single bit errors

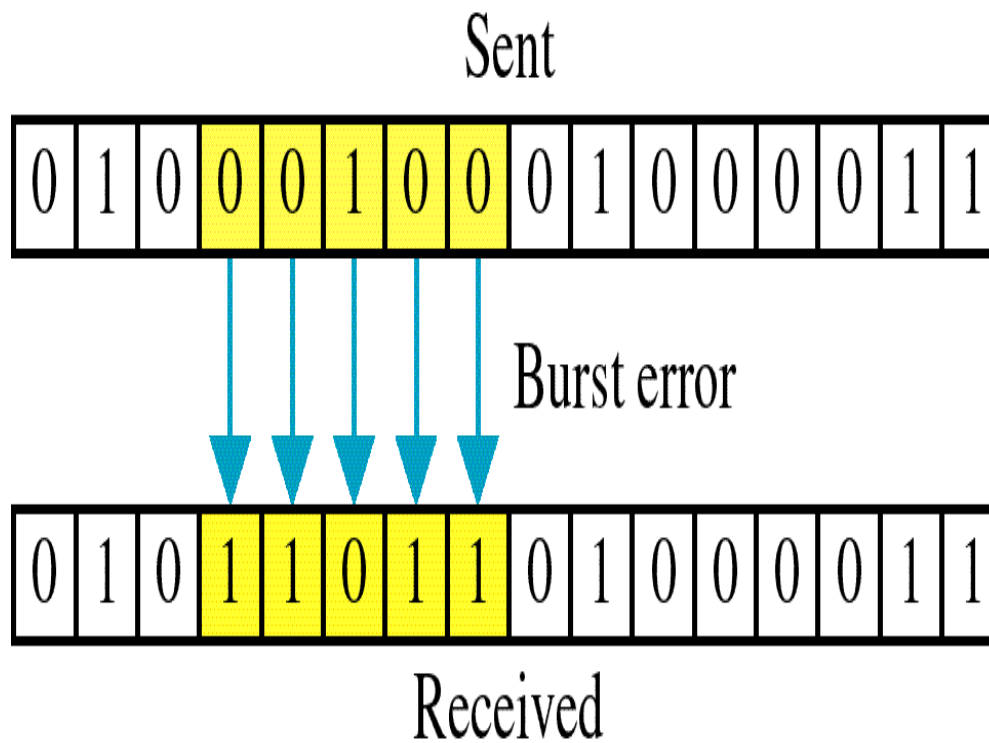
least likely type of errors in serial data transmission because the noise must have a very short duration which is very rare.

- this kind of errors can happen in **parallel transmission**

BURST ERROR

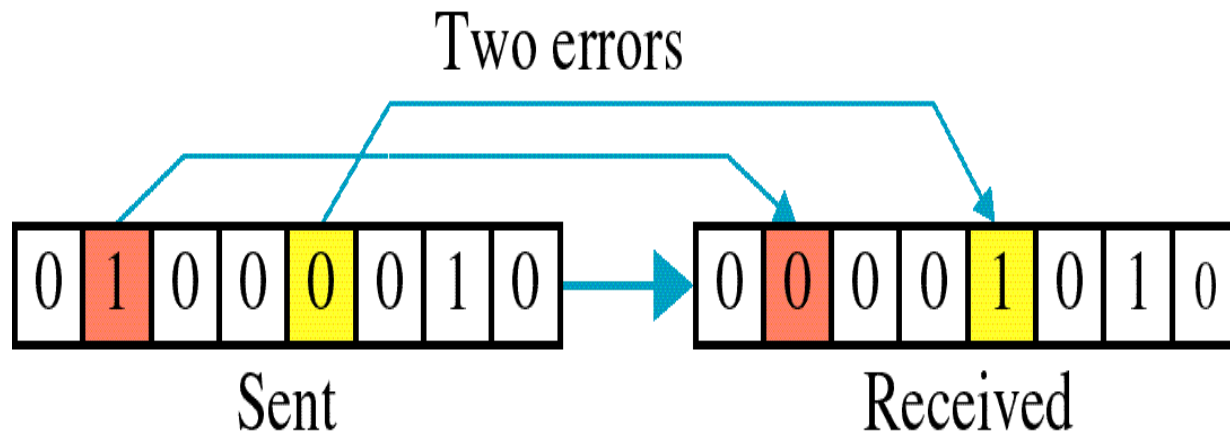
- **burst error** means that **two or more bits in the data unit have changed from 1 to 0 or from 0 to 1.**
- **does not mean that the errors occur in consecutive bits,** the length of the burst is measured from the first corrupted bit to the last corrupted bit.
- **Some bits in between may not have been corrupted.**

Burst error



- **Burst error is most likely to happen in serial transmission** since the duration of noise is normally longer than the duration of a bit.
- The number of bits affected depends on the data rate and duration of noise.

Multiple bit errors



Error Detection

1. VRC (Vertical Redundancy Check)
2. LRC (Longitudinal Redundancy Check)
3. CRC (Cyclic Redundancy Check)

Redundancy

- some extra bits are sent with data.
- These bits are added by the sender and removed by the receiver

Error Correction

Process of correcting the errors

Two types of error correction

- i. Forward error correction(FEC)
- ii. Backward Error Correction(BEC) -
Retransmit data (Automatic Repeat
reQuest[ARQ])

Forward Error Correction

- This is done where retransmission is impractical
- To correct an error, the receiver reverses the value of the altered bit.
- To do so, it must know which bit is in error.
- Number of redundancy bits needed
 - Let data bits = m
 - Redundancy bits = r
 - ∴ Total message sent, $n = m+r$

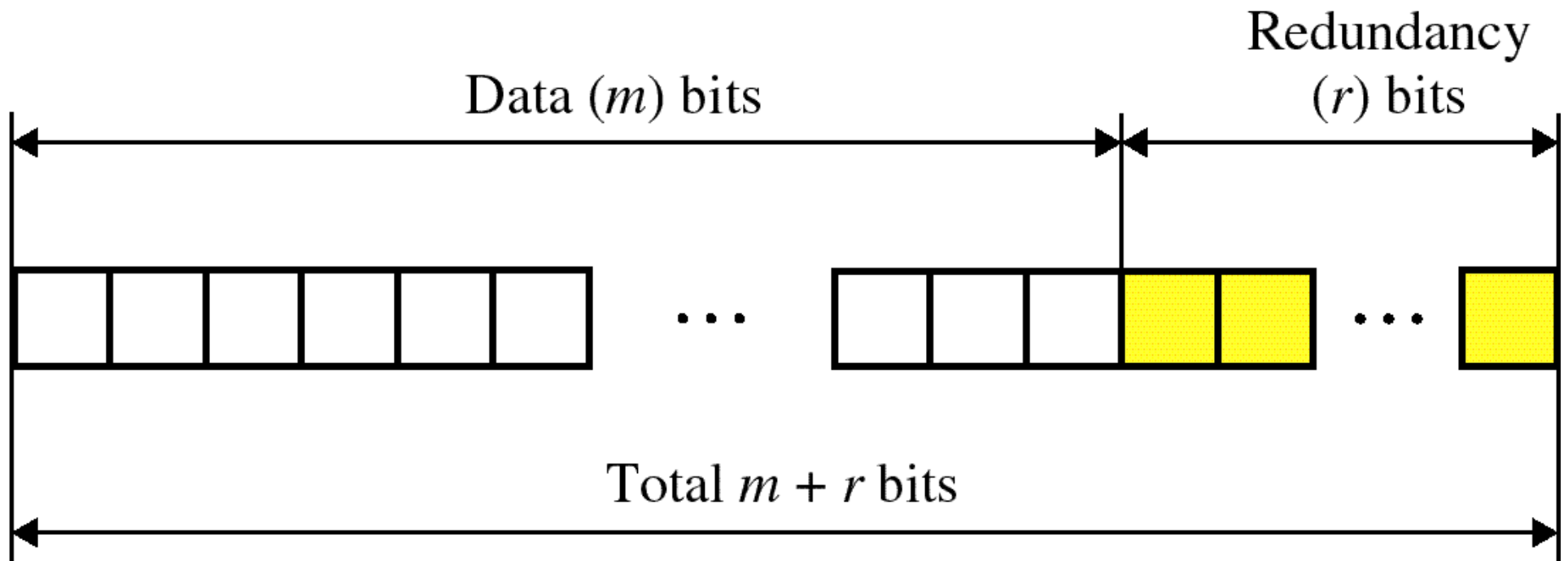
N bit unit containing data and checkbits is called n bit codeword

The value of r must satisfy the following relation:

$$2^r \geq m+r+1$$

Eg. Hamming codes

*2^r ≥ m+r+1
Eg. Hamming codes*

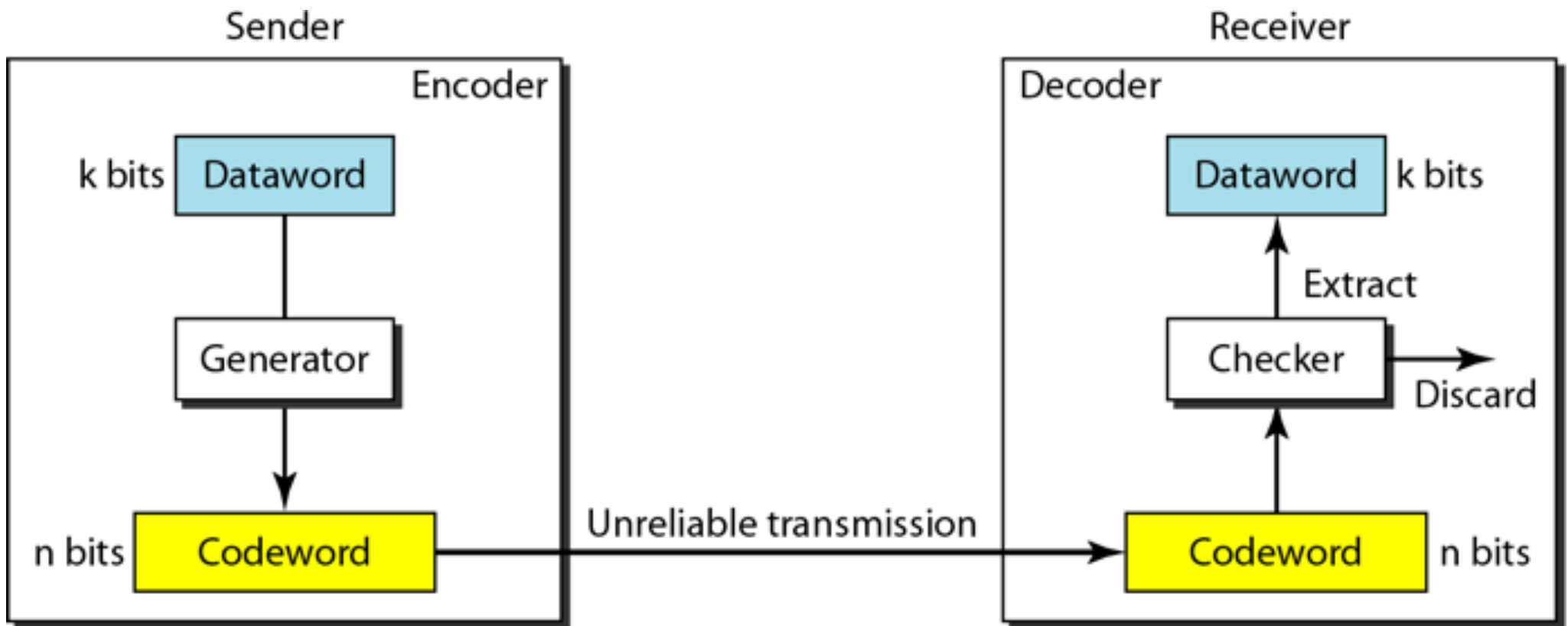


- Redundancy is achieved thru various coding schemes
- Coding schemes are categorised into 2
 1. Block coding
 2. Convolution coding

Block Coding

- Divide message into blocks, each of **k bits** called **datawords**
- Add **r redundant bits** to each block , **$n = k + r$**
- Resultant **n-bit blocks** are called **codewords**
- Block coding process is one to one, same dataword is encoded as same codeword which means **$2^n - 2^k$** codewords are not used (illegal codewords)

Figure 10.3 *The structure of encoder and decoder*



To detect or correct errors, we need to send redundant bits

Hamming Distance

- No of bit positions in which 2 code words differ is called **Hamming Distance**

Eg 1000011100, 1111100100

Hamming distance for the above 2 codewords is

--

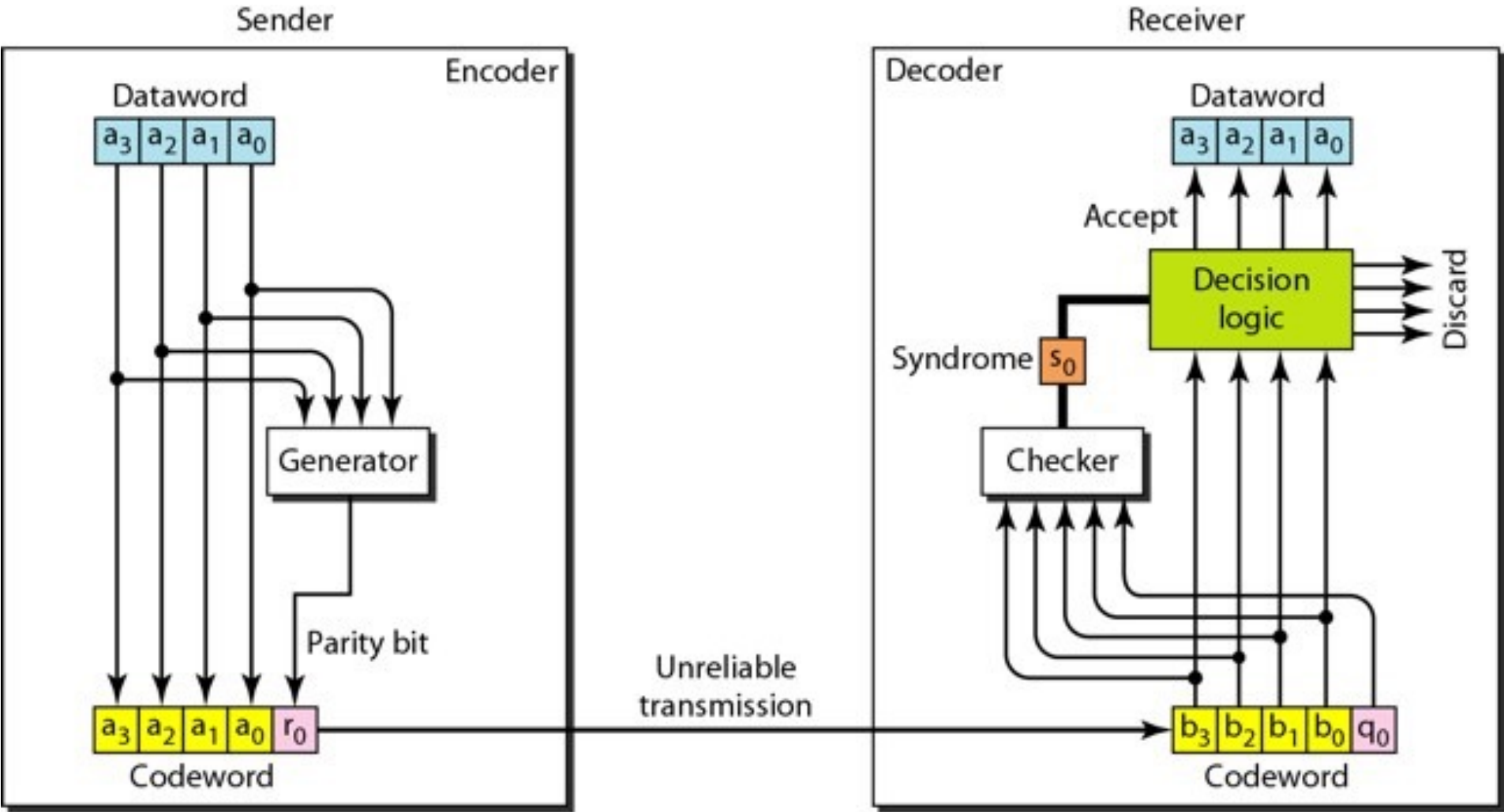
- If 2 codewords are hamming distance d apart, it will require d single bit errors to convert one into other
- To detect d errors, there must be $d+1$ distance
- To correct d errors, there must be $2d+1$ distance

Hamming code

- These codes were originally designed with $d_{min} = 3$, *which means that they can detect up to two errors or correct one single error.*
- the relationship between n and k in a Hamming code.
 - choose an integer $m \geq 3$.
 - *The values of n and k are then calculated from m as $n = 2^m - 1$ and $k = n - m$. [codeword $n = m + r$]*
 - *The number of check bits $r = m$.*
 - *For example, if $m = 3$, then $n = 7$ and $k = 4$. This is a Hamming code $C(7, 4)$ with $d_{min} = 3$.*

<i>Datawords</i>	<i>Codewords</i>	<i>Datawords</i>	<i>Codewords</i>
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

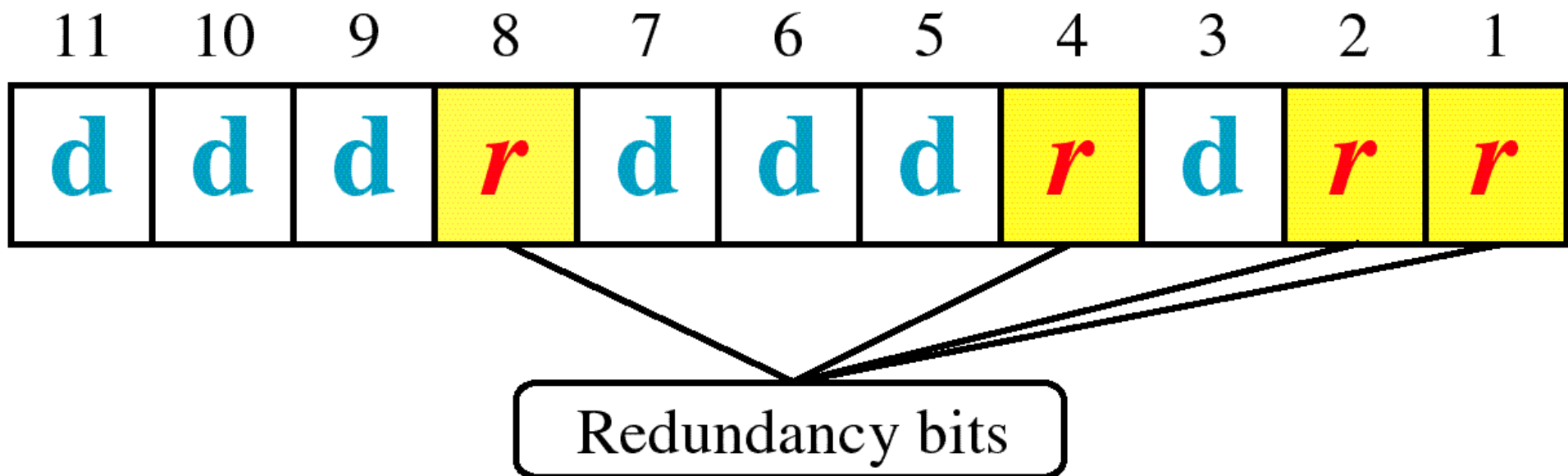
Structure of encoder and decoder - hamming code



- A copy of a 4-bit dataword is fed into the generator that creates three parity checks
 - *r0, r1 and r2 as shown below:*
 - $r0 = a2 + a1 + a0$
 - $r1 = a3 + a2 + a1$
 - $r2 = a1 + a0 + a3$
- checker in the decoder creates a 3-bit syndrome ($s2s1s0$) in which each bit is the parity check for 4 out of the 7 bits in the received codeword:
 - $S0 = b2 + b1 + b0 + q0$
 - $S1 = b3 + b2 + b1 + q1$
 - $S2 = b1 + b0 + b3 + q2$

Hamming Code - example

- Bits of the codeword are numbered consecutively starting with bit 1 at the left end
- Checkbits (redundant) are **bit positions that are powers of 2**
- Rest are filled with m data bits



Hamming code

- Each check bit is the parity of some collection of bits including itself
- Write K (data bit position) as a sum of powers of 2

$$\text{Eg } 11 = 1 + 2 + 8$$

$$10 = 2 + 8$$

$$9 = 1 + 8$$

- Compute redundant bits, find the parity of the data bits that has the redundant bits in its sum
- Eg check bit position 1 occurs in 11,9,7,5,3
- e.g. in notebook

Unit II - Flow Control

What is flow control?

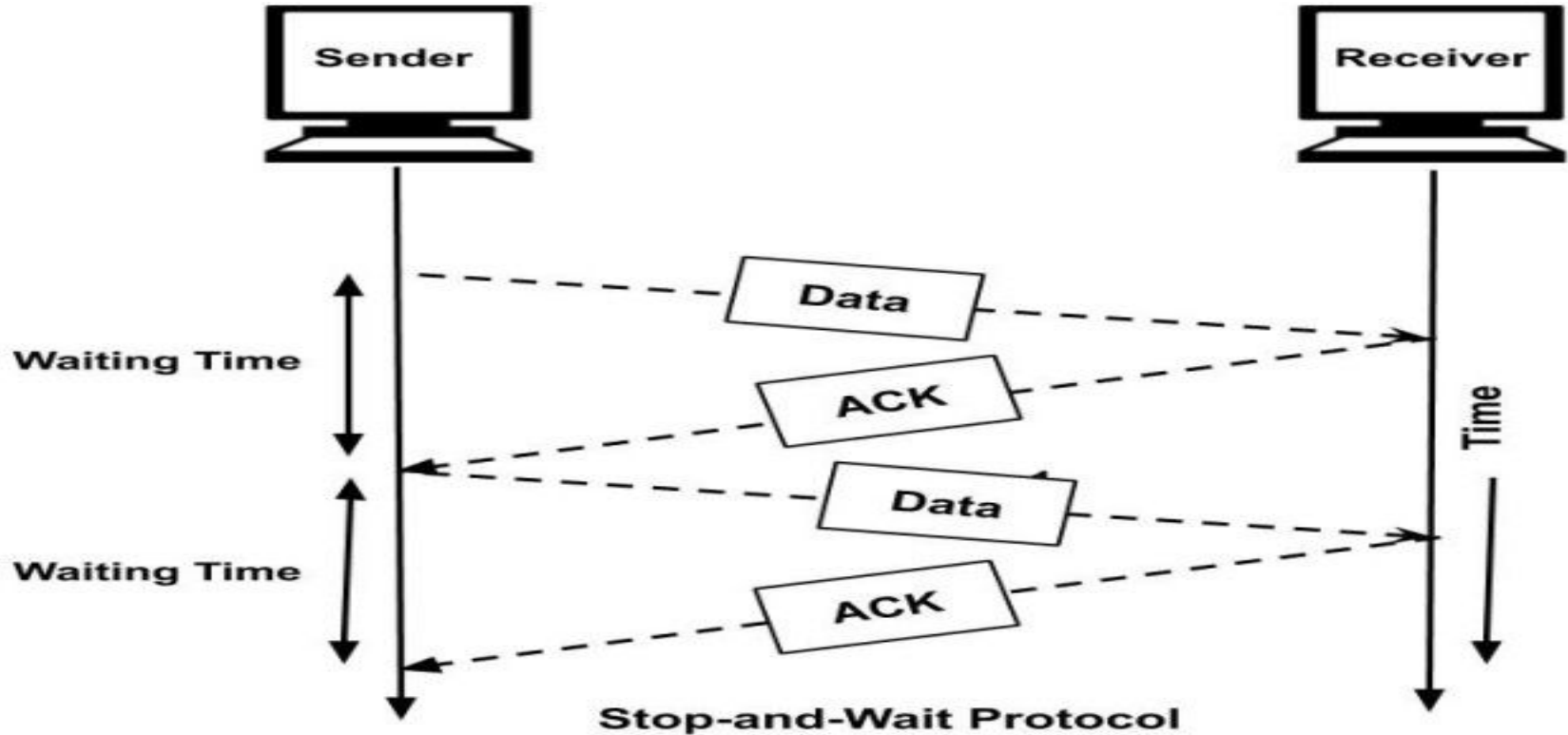
Flow control is a technique for assuring that a transmitting entity does not overwhelm a receiving entity with data.

ie sender should not send data at a rate faster than the receiver can receive it

Two techniques

1. **Stop and wait protocol**
2. **Sliding Window protocol**

Stop and wait Protocol



- It is the **simplest flow control** method.
- In this, the sender **will send one frame at a time** to the receiver.
- The sender will **stop and wait** for the acknowledgment from the receiver.
- This time(i.e. the time between message sending and acknowledgement receiving) is the waiting time for the sender and the sender is totally idle during this time.
- When the **sender gets the acknowledgment(ACK), then it will send the next data packet to the receiver**
- and wait for the acknowledgment again
- this process will continue as long as the sender has the data to send.

→ procedure works fine if a message is sent in a few large frames. often a source will break up a large block of data into smaller blocks and transmit the data in many frames.

→

→ **This is done for the following reasons:**

- **The buffer size of the receiver may be limited.**
- **The longer the transmission, the more likely that there will be an error.** With smaller frames, errors are detected sooner, and a smaller amount of data needs to be retransmitted.
- **On a shared medium it is not desirable to permit one station to occupy the medium for a long period(causing long delays to the other sending stations).**

STOP AND WAIT FLOW CONTROL

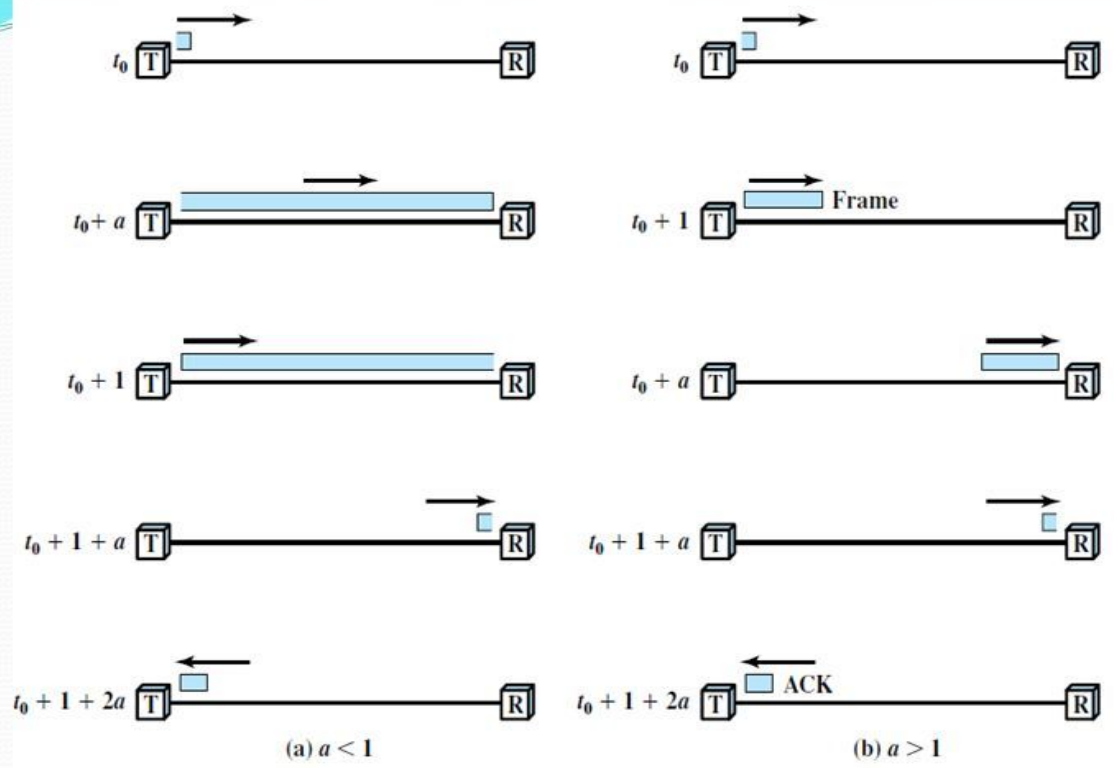


Figure 7.2 Stop-and-Wait Link Utilization (transmission time = 1; propagation time = a)

When a is less than 1, the propagation time is less than the transmission time. In this case, the frame is sufficiently long that the first bits of the frame have arrived at the destination before the source has completed the transmission of the frame.

When a is greater than 1, the propagation time is greater than the transmission time. In this case, the sender completes transmission of the entire frame before the leading bits of that frame arrive at the receiver..

Both parts of Figure consist of a sequence of snapshots of the transmission process over time. In both cases, the first four snapshots show the process of transmitting a frame containing data, and the last snapshot shows the return of a small acknowledgment frame. **Note that for $a > 1$, the line is always underutilized and even for $a < 1$, the line is inefficiently utilized.**

for very high data rates, for very long distances between sender and receiver, stop-and- wait flow control provides **inefficient line utilization**

In both cases $a > 1$ and $a < 1$

$$\text{Utilization} = \frac{1}{1 + 2a}$$

2. Sliding Window Protocol

- **Efficiency can be greatly improved by allowing multiple frames to be in transit at the same time.**
- **two stations, A and B,** are connected via a full-duplex link.
- **A maintains a list of sequence numbers** that it is allowed to send
- **B maintains a list of sequence numbers** that it is prepared to receive.
- Each of these lists can be thought of as a *window* of frames. The operation is referred to as **sliding-window flow control**.

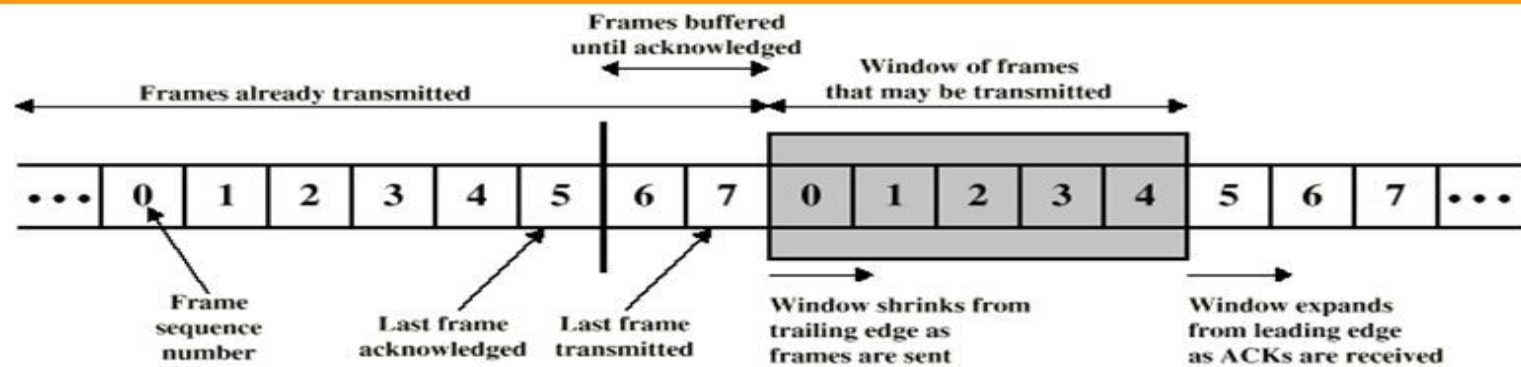
- **Station B** allocates buffer space for **W frames**.
- Thus, B can accept W frames, and **A is allowed to send W frames without waiting for any acknowledgments**.
- To keep track of which frames have been acknowledged, **each is labeled with a sequence number**.
- **B acknowledges a frame** by sending an acknowledgment that includes the **sequence number of the next frame expected**.

This acknowledgment specifies that B is prepared to receive the next W frames, beginning with the number specified. This scheme can also be used to acknowledge multiple frames.

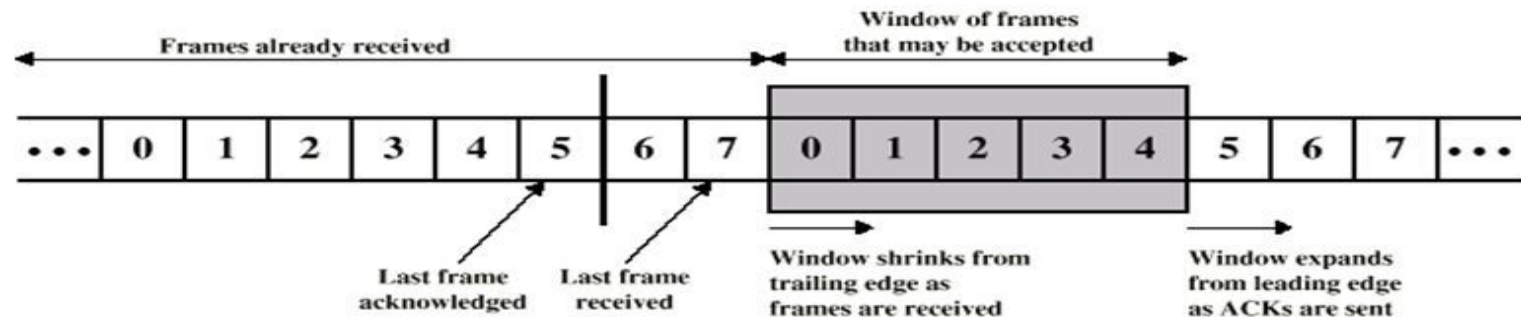
For example, B could receive frames 2, 3, and 4 but withhold acknowledgment until frame 4 has arrived.

By then returning an acknowledgment with sequence number 5, B acknowledges frames 2, 3, and 4 at one time.

Sliding Window Flow Control ($W = 7$)



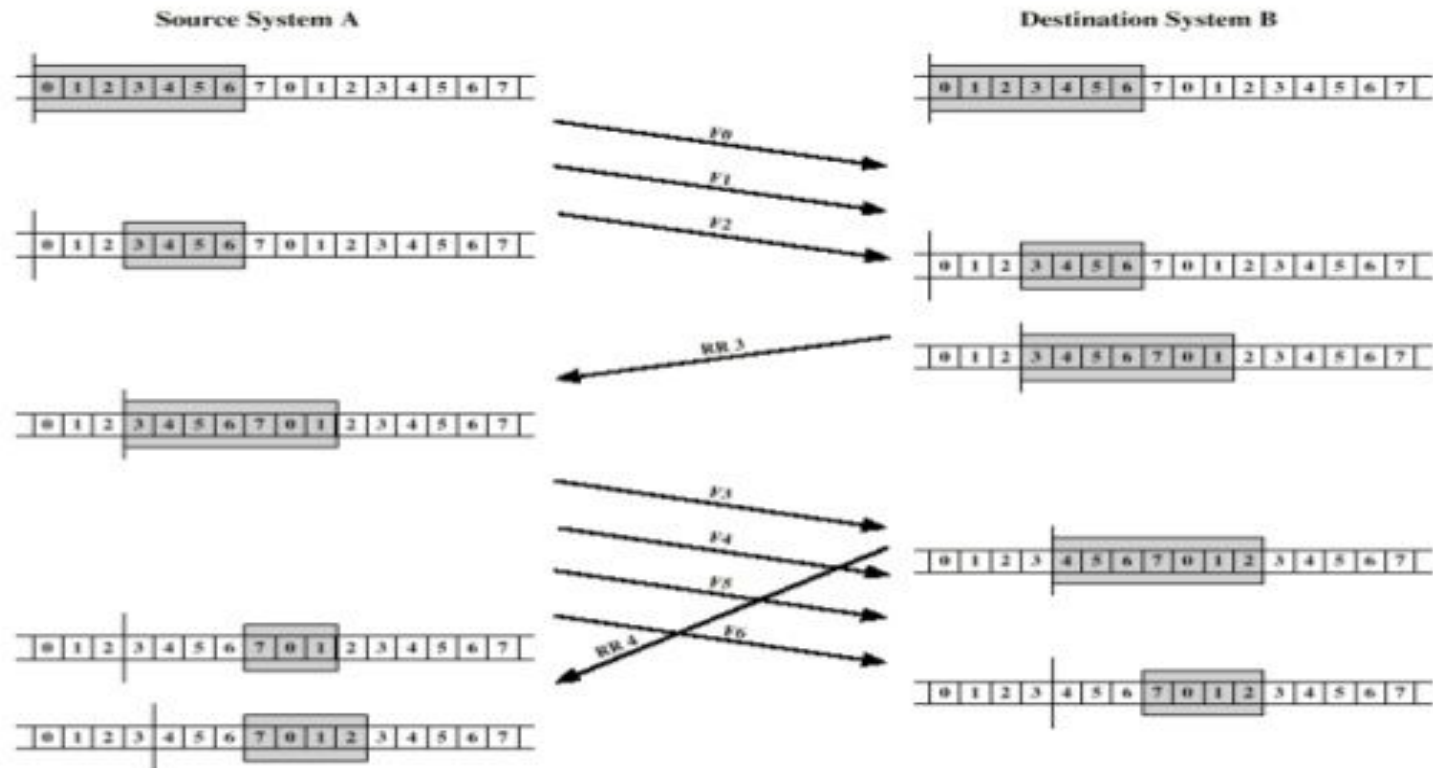
(a) Sender's perspective



(b) Receiver's perspective

- **sequence number to be used occupies a field** in the frame, and it is limited to a range of values.
- **For example**, for a 3-bit field, the sequence number can range from 0 to 7.
- **frames are numbered modulo 8**; that is, after sequence number 7, the next number is 0.
- for a k -bit field the range of sequence numbers is **0 through $2^k - 1$** ,
- frames are numbered modulo 2^k .
- **the maximum window size is $2^k - 1$** .

Example Sliding Window



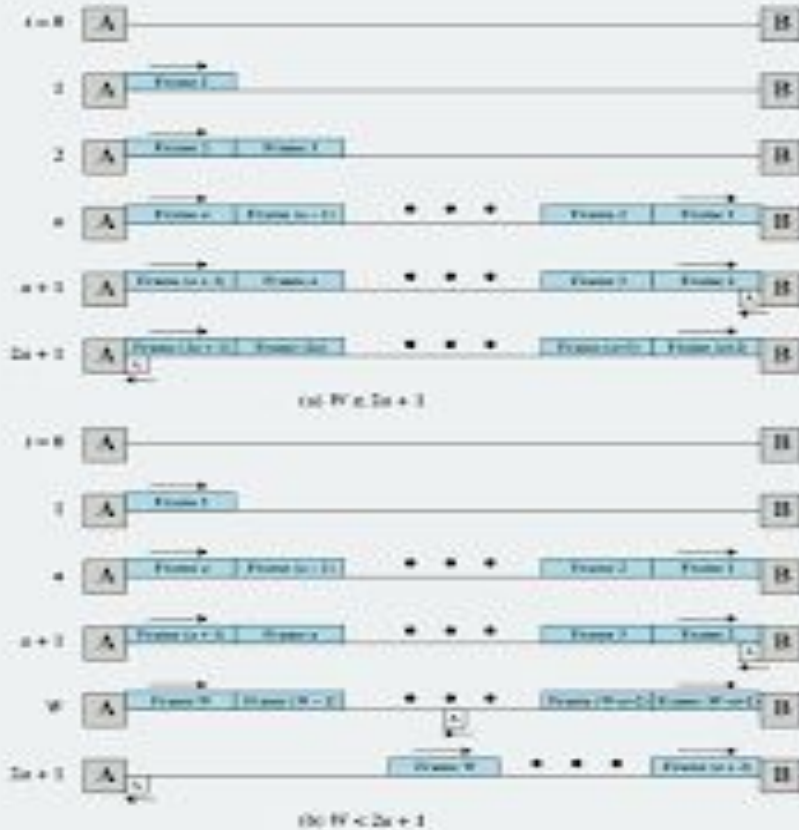


Figure 16.9 Timing of Sliding-Window Protocol

Two cases :

1. $N > 1+2a$

Utilization = 1

2. $N < 1+2a$

Utilization = $\frac{N}{1 + 2a}$

Unit II - Error Detection

Longitudinal Redundancy Check (LRC)

In this method, message which the user want to send is organised into **block of characters**

A block of bits is divided into **table or matrix of rows and columns.**

Parity check bits are calculated for each row, which is equivalent to a simple parity check bit.

Parity check bits are also calculated for all columns, then both are sent along with the data.

At the receiving end these are compared with the parity bits calculated on the received data.

Original Data

10011001	11100010	00100100	10000100
----------	----------	----------	----------

Row parities

10011001	0
11100010	0
00100100	0
10000100	0
11011011	0

Column
parities



Parity Check Character



100110010	111000100	001001000	100001000	110110110
-----------	-----------	-----------	-----------	-----------

Data to be sent

Disadvantage :

The main problem with LRC is that, it is not able to detect error if two bits in rows are damaged and two bits in exactly the same position in other columns are also damaged.

Example: 4 characters each of 7 bits

1010100 1010111 1111000 1011111

	Parity bit
1010100	0
1010111	0
1111000	1
<u>1011111</u>	<u>1</u>
<u>1011011</u>	<u>1</u>

Parity Check Character

At the Receiver

1110100	0
1110111	0
1111000	1
<u>1011111</u>	<u>1</u>
<u>1011011</u>	<u>1</u>

Error detected

Error goes undetected

Cyclic Redundancy Check

- 1) Given a **k-bit block of bits**, or message, the **transmitter generates an sequence**, known as a frame check sequence (FCS), consisting of n bits, is exactly **divisible by some predetermined number**.
- 2) The receiver then divides the incoming frame by that number and, if there is **no remainder, assumes there was no error**.
- 3) procedure can be implemented in three ways:
 - i) **modulo 2 arithmetic**
 - ii) **polynomials**
 - iii) **digital logic**.

1. Modulo 2 Arithmetic

It uses binary addition with no carries, which is just the **exclusive-OR (XOR) operation**.

M = k -bit block of data, or message

F = n bits FCS

T = $k + n$ bit frame to be transmitted

P = pattern of $n + 1$ bits; this is the predetermined divisor

The pattern P is chosen to be one bit longer than the desired FCS, and the exact bit pattern chosen depends on the type of errors expected. At minimum, both the high- and low-order bits of P must be 1.

$$T = 2^n M + F$$

Where $2^n M$ shifts message , left by n bits and padded out the result with zeroes

We want T to be exactly divisible by P.

$$2^n M / P = Q + R / P \quad \text{-----} \rightarrow (1)$$

There is a quotient and a remainder.

Because division is modulo 2, the remainder is always at least one bit shorter than the divisor.

We will use this remainder as our FCS. Then

$$T = 2^n M + R$$

Does this R satisfy our condition that T/P have no remainder? To see that it does, consider

$$T / P = \frac{2^n M + R}{P} = 2^n M / P + R / P$$

Substituting Equation (1), we have

$$T / P = Q + R / P + R / P$$

However, any binary number added to itself modulo 2 yields zero. Thus

$$\mathbf{T / P = Q + (R + R) / P = Q}$$

There is no remainder, and therefore T is exactly divisible by P.

Thus, the FCS is easily generated: Simply divide by P and use the remainder as the FCS.

EXAMPLE 1

Given

Message M = 1010001101 (10 bits)

Pattern P = 110101 (6 bits)

FCS R = to be calculated (5 bits)

k = 10 bits, n= 5 bits T = 15 bits .

The message is multiplied 2^5 by yielding 101000110100000.

This product is divided by P:

	1101010110	Q
P	110101	$2^n M$
	<u>110101</u>	
	111011	
	<u>110101</u>	
	111010	
	<u>110101</u>	
	111110	
	<u>110101</u>	
	101100	
	<u>110101</u>	
	110010	
	<u>110101</u>	
	01110	R

An error results in the reversal of a bit.

This is equivalent to taking the XOR of the bit and 1 (modulo 2 addition of 1 to the bit):

$$0 + 1 = 1; 1 + 1 = 0.$$

Thus, the errors in an n -bit frame can be represented by an n -bit field with 1s in each error position.

The resulting frame Tr can be expressed

$$Tr = T \oplus E$$

T = transmitted frame

E = error pattern with 1s in positions where errors occur

Tr = received frame

- If there is an error,
 - the receiver will fail to detect the error if and only if T_r is divisible by P , which is equivalent to E divisible by P .
- Intuitively, this seems an unlikely occurrence.

2. Polynomials (CRC)

- A second way of viewing the CRC process is to **express all values as polynomials in a dummy variable X , with binary coefficients.**
- The coefficients correspond to the bits in the binary number.
 - for $M = 110011$, we have $M(X) = X^5 + X^4 + X + 1$
 - for $P = 11001$, we have $P(X) = X^4 + X^3 + 1$.
 - Arithmetic operations are again modulo 2.

- The CRC process can now be described as

$$\frac{X^n M(X)}{P(X)} = Q(X) + \frac{R(X)}{P(X)}$$

Transmitted as

$$T(X) = X^n M(X) + R(X)$$

EXAMPLE 6.8 The architecture of a CRC circuit is best explained by first considering an example, which is illustrated in Figure 6.5. In this example, we use

$$\begin{aligned} \text{Data } D &= 1010001101; & D(X) &= X^9 + X^7 + X^3 + X^2 + 1 \\ \text{Divisor } P &= 110101; & P(X) &= X^5 + X^4 + X^2 + 1 \end{aligned}$$

which were used earlier in the discussion.

Figure 6.5a shows the shift register implementation. The process begins with the shift register cleared (all zeros). The message, or dividend, is then entered, one bit at a time, starting with the most significant bit. Figure 6.5b is a table that shows the step-by-step operation as the input is applied one bit at a time. Each row of the table shows the values currently stored in the five shift-register elements. In addition, the row shows the values that appear at the outputs of the three XOR circuits. Finally, the row shows the value of the next input bit, which is available for the operation of the next step.

Note that the XOR operation affects C_4 , C_2 , and C_0 on the next shift. This is identical to the binary long division process illustrated earlier. The process continues through all the bits of the message. To produce the proper output, two switches are used. The input data bits are fed in with both switches in the A position. As a result, for the first 10 steps, the input bits are fed into the shift register and also used as output bits. After the last data bit is processed, the shift register contains the remainder (FCS) (shown shaded). As soon as the last data bit is provided to the shift register, both switches are set to the B position. This has two effects: (1) All of the XOR gates become simple pass-throughs; no bits are changed, and (2) as the shifting process continues, the 5 CRC bits are output.

At the receiver, the same logic is used. As each bit of M arrives, it is inserted into the shift register. If there have been no errors, the shift register should contain the bit pattern for R at the conclusion of M . The transmitted bits of R now begin to arrive, and the effect is to zero out the register so that, at the conclusion of reception, the register contains all 0s.

$$\begin{array}{r}
 P(X) \rightarrow X^5 + X^4 + X^2 + 1 \overline{) X^9 + X^8 + X^6 + X^4 + X^2 + X} \quad \leftarrow Q(X) \\
 \underline{X^{14} \phantom{+ X^{13}} + X^{12} \phantom{+ X^{11}} + X^8 + X^7 + X^5} \quad \leftarrow X^5 D(X) \\
 X^{14} + X^{13} + \phantom{X^{12}} + X^{11} + X^9 \\
 \underline{\phantom{X^{14}} + X^{13} + X^{12} + X^{11} + X^9 + X^8} \\
 X^{13} + X^{12} + \phantom{X^{11}} + X^{10} + X^8 \\
 \underline{\phantom{X^{13}} + X^{12} + X^{10} + X^7} \\
 X^{11} + X^{10} + X^9 + X^6 \\
 \underline{\phantom{X^{11}} + X^{10} + X^8 + X^6} \\
 X^9 + X^8 + X^7 + X^6 + X^5 \\
 \underline{ + X^8 + X^6 + X^4} \\
 X^7 + X^5 + X^4 \\
 \underline{ + X^6 + X^4 + X^2} \\
 X^6 + X^5 + X^2 \\
 \underline{ + X^5 + X^3 + X} \\
 X^3 + X^2 + X \leftarrow R(X)
 \end{array}$$

Figure 6.4 Example of Polynomial Division

- Using the preceding example, for $M = 1010001101$, we have

$$M(X) = X^9 + X^7 + X^3 + X^2 + 1$$

- for $P=110101$, we have $P(X) = X^5 + X^4 + X^2 + 1$.

- We should end up with Remainder $R = 01110$, which corresponds to

$$R(X) = X^3 + X^2 + X.$$

- ★ An error $E(X)$ will only be undetectable if it is divisible by $P(X)$.

- ★ It can be shown that all of the **following errors are not divisible by a suitably chosen $P(X)$ and hence are detectable:**
 - All **single-bit errors**, if $P(X)$ has more than **one non-zero term**
 - All **double-bit errors**, as long as $P(X)$ has at least 3 terms
 - Any **odd number of errors**, as long as $P(X)$ contains a factor $(X + 1)$

 - Any burst error for which the length of the **burst is less than or equal to n** ie less than or equal to the length of the FCS
 - A fraction of error bursts of length $n + 1$; the fraction equals $1 - 2^{-(n-1)}$
 - A fraction of error bursts of length greater than $n + 1$; the fraction equals $1 - 2^{-n}$

- It can be shown that if all error patterns are considered equally likely, then
 - for a burst error of length $r + 1$, the probability of an undetected error ($E(X)$ is divisible by $P(X)$) is $1/2^{r-1}$, **r is the length of the FCS.**
 - for a longer burst, the probability is $1/2^r$

Four versions of $P(X)$ are widely used:

1) CRC-12= $X^{12} + X^{11} + X^3 + X^2 + X + 1$

2) CRC-16= $X^{16} + X^{15} + X^2 + 1$

3) CRC-CCITT= $X^{16} + X^{12} + X^5 + 1$

4) CRC-32= $X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + X + 1$

- **CRC-12** system is used for transmission of streams of 6-bit characters and generates a 12-bit FCS.
- Both **CRC-16** and **CRC-CCITT** are popular for 8-bit characters, in the United States and Europe, respectively, and both result in a 16-bit FCS.
- This would seem adequate for most applications, although **CRC-32 is specified as an option in some point-to-point synchronous transmission standards and is used in IEEE 802 LAN standards.**



3. Digital Logic (CRC)

- implemented as a **circuit consisting of XOR gates and a shift register.**
- The **shift register** is a string of **1-bit storage devices.**
- Each **device has an output line**, which indicates the value currently stored, and an **input line.**
- At discrete time instants, known as **clock times**, the **value in the storage device is replaced by the value indicated by its input line.**
- The **entire register is clocked simultaneously**, causing a 1-bit shift along the entire register.

○

■

circuit is implemented as follows:

1. The register contains n bits, equal to the length of the FCS.
2. There are upto n XOR gates.
3. The presence or absence of a gate corresponds to the presence or absence of a term in the divisor polynomial, $P(X)$, excluding the terms 1 and X^n

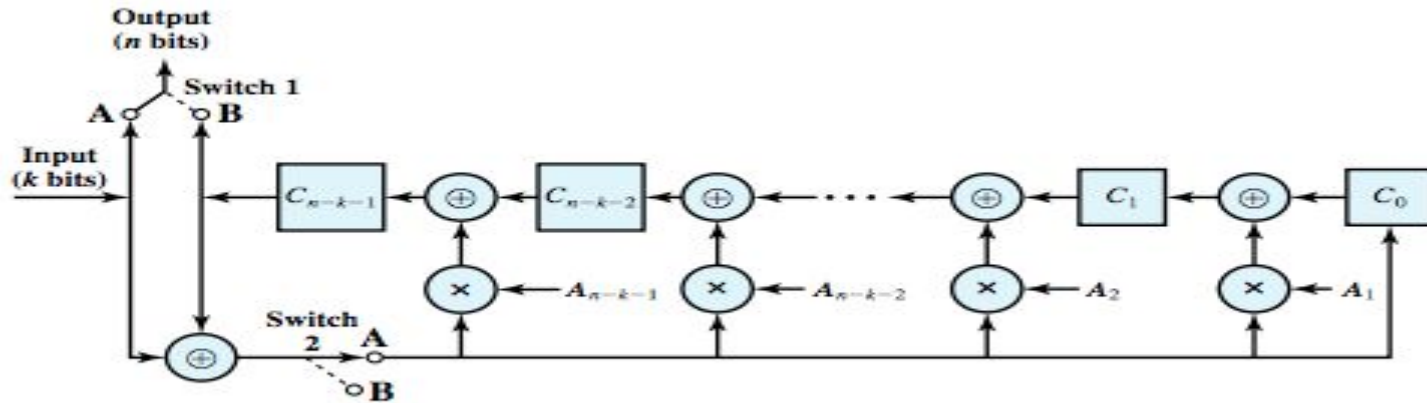
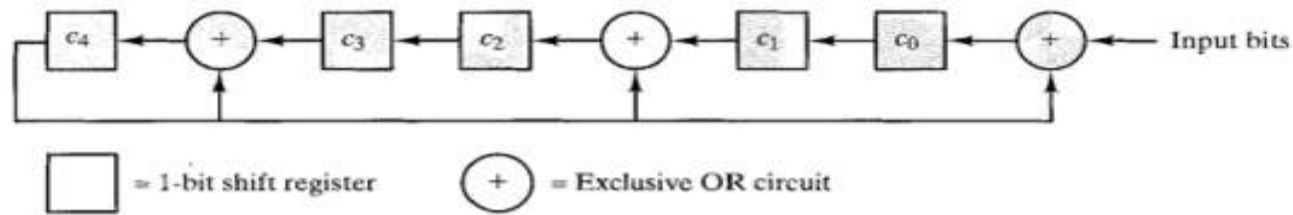


Figure 6.6 General CRC Architecture to Implement Divisor $(1 + A_1X + A_2X^2 + \dots + A_{n-1}X^{n-k-1} + X^{n-k})$

indicates the general architecture of the shift register implementation of a CRC for the polynomial $P(X) = \sum A_i X^i$, where $A_0 = A_n = 1$ and all other A_i equal either 0 or 1



(a) Shift-register implementation

	c_4	c_3	c_2	c_1	c_0	$c_4 \oplus c_3$	$c_4 \oplus c_1$	$c_4 \oplus \text{input}$	input	
Initial	0	0	0	0	0	0	0	1	1	} Message to be sent
Step 1	0	0	0	0	1	0	0	0	0	
Step 2	0	0	0	1	0	0	1	1	1	
Step 3	0	0	1	0	1	0	0	0	0	
Step 4	0	1	0	1	0	1	1	0	0	
Step 5	1	0	1	0	0	1	1	1	0	
Step 6	1	1	1	0	1	0	1	0	1	
Step 7	0	1	1	1	0	1	1	1	1	
Step 8	1	1	1	0	1	0	1	1	0	
Step 9	0	1	1	1	1	1	1	1	1	
Step 10	1	1	1	1	1	0	0	1	0	} Five zeros added
Step 11	0	1	0	1	1	1	1	0	0	
Step 12	1	0	1	1	0	1	0	1	0	
Step 13	1	1	0	0	1	0	1	1	0	
Step 14	0	0	1	1	1	0	1	0	0	
Step 15	0	1	1	1	0	1	1	0	—	

FIGURE 6.6 Circuit with shift registers for dividing by the polynomial $X^5 + X^4 + X^2 + 1$.

Error Correction

- **Process of correcting the errors**

Two types of error correction

- i. **Forward error correction (FEC)**
- ii. **Backward Error Correction (BEC) or called Automatic Repeat reQuest [ARQ]**

Forward Error Correction

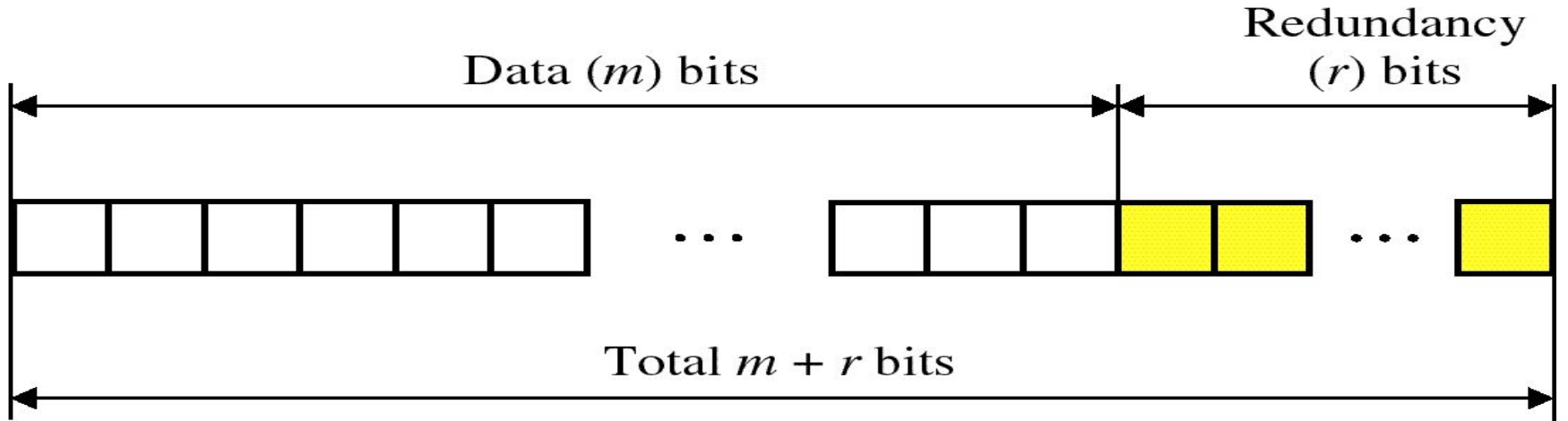
- This is done where retransmission is impractical
- To correct an error, the receiver reverses the value of the altered bit.
- To do so, it must know which bit is in error.
- Number of redundancy bits needed
 - Let data bits = m
 - Redundancy bits = r
 - ∴ Total message sent, $n = m+r$

n bit unit containing data and checkbits is called n bit codeword

The value of r must satisfy the following relation:

$$2^r \geq m+r+1$$

Eg. Hamming codes



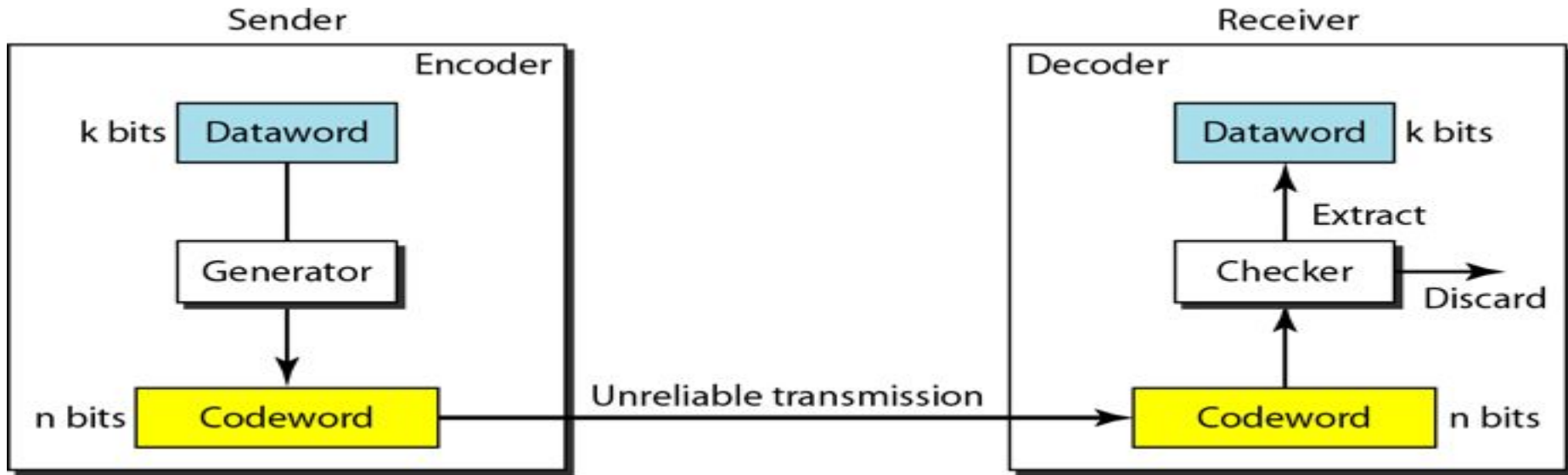
- Redundancy is achieved thru various coding schemes
- Coding schemes are categorised into 2
 1. **Block coding**
 2. **Convolution coding**

Block Coding

- Divide message into blocks, each of **k bits** called **datawords**
- **Add r redundant bits** to each block , $n = k + r$
- Resultant n-bit blocks are called **codewords**
- **Block coding process is one to one, same dataword is encoded as same codeword which means $2^n - 2^k$ codewords are not used(illegal codewords)**

<i>Datawords</i>	<i>Codewords</i>	<i>Datawords</i>	<i>Codewords</i>
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

The structure of encoder and decoder



To detect or correct errors, we need to send redundant bits

Hamming Distance

- No of bit positions in which 2 code words differ is called **Hamming Distance**

Eg 1000011100, 1111100100

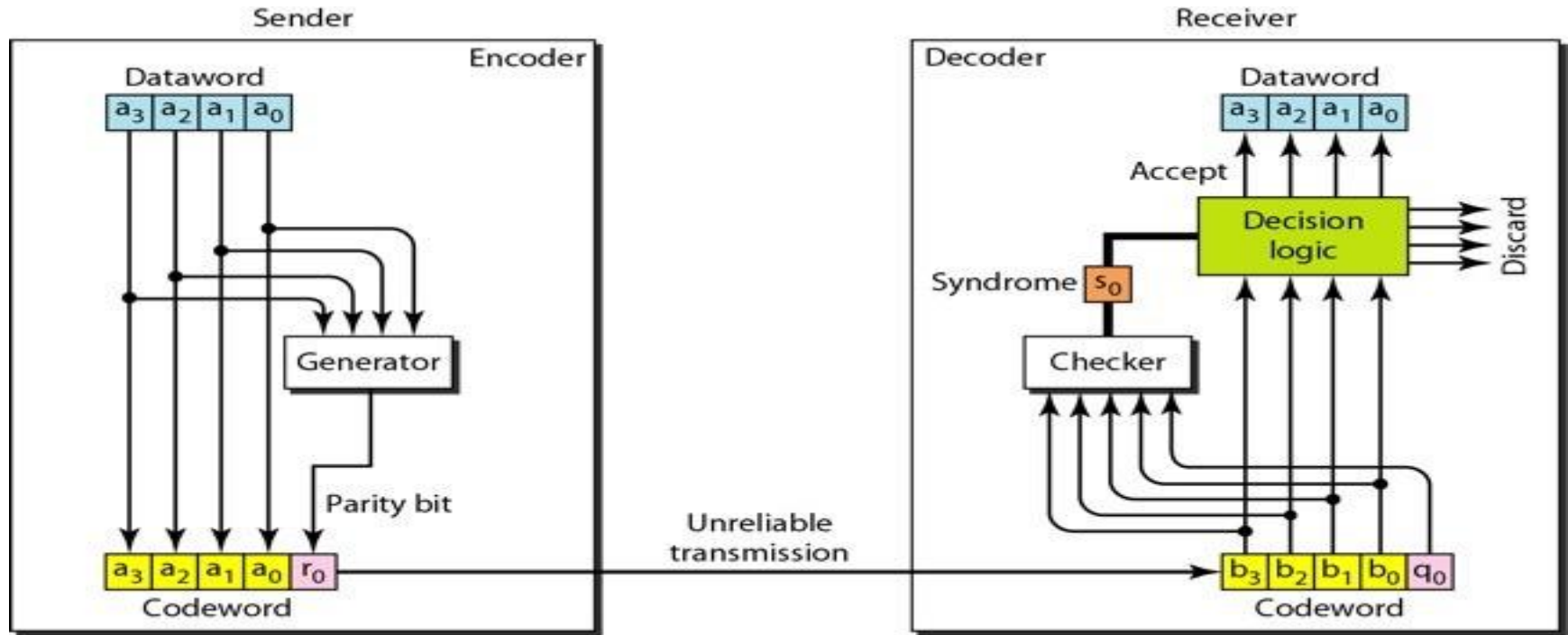
Hamming distance for the above 2 codewords is ?

- If 2 codewords are hamming distance **d** apart, it will require **d single bit errors to convert one into other**
- To **detect d errors**, there must be **$d+1$ distance**
- To **correct d errors**, there must be **$2d+1$ distance**

Hamming code

- These codes were originally designed with $d_{min} = 3$, which means that they can detect up to two errors or correct one single error.
- the relationship between n and k in a Hamming code.
 - choose an integer $r \geq 3$.
 - The values of n and k are then calculated from m as $n = 2^r - 1$ and $k = n - r$. [codeword $n = k + r$]
 - The number of check bits is r
 - For example, if $r = 3$, then $n = 7$ and $k = 4$. This is a Hamming code $C(7, 4)$ with $d_{min} = 3$.

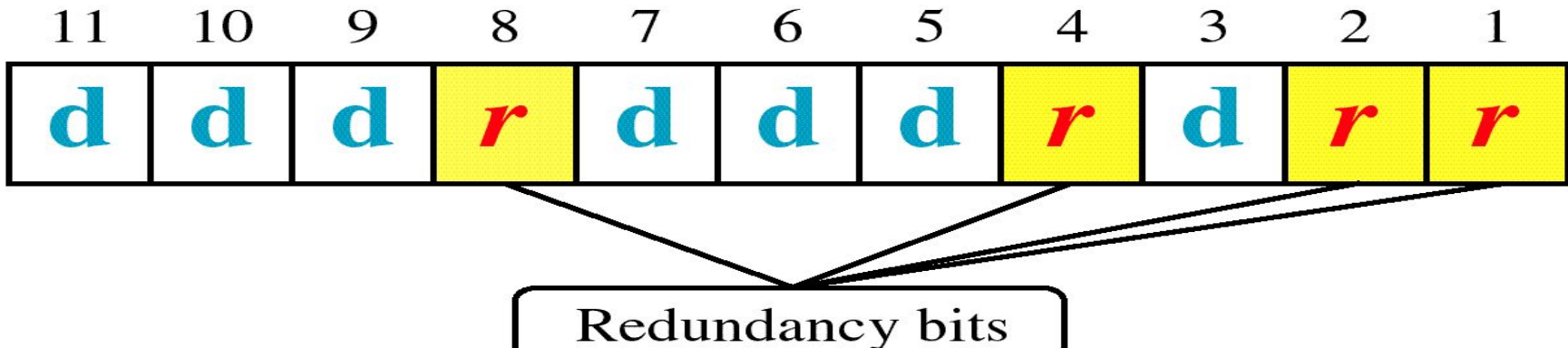
Structure of encoder and decoder - hamming code



- A copy of a 4-bit dataword is fed into the generator that creates three parity checks
 - *r0, r1 and r2 as shown below:*
 - $r0 = a2 + a1 + a0$
 - $r1 = a3 + a2 + a1$
 - $r2 = a1 + a0 + a3$
- checker in the decoder creates a 3-bit syndrome ($s2\ s1\ s0$) in which each bit is the parity check for 4 out of the 7 bits in the received codeword:
 - $S0 = b2 + b1 + b0 + q0$
 - $S1 = b3 + b2 + b1 + q1$
 - $S2 = b1 + b0 + b3 + q2$

Hamming Code - example

- Bits of the codeword are numbered consecutively starting with bit 1 at the left end
- Checkbits (redundant) are bit positions that are powers of 2
- Rest are filled with k data bits



Hamming code

- Each check bit is the parity of some collection of bits including itself
- Write K (data bit position) as a sum of powers of 2

$$\text{Eg } 11 = 1 + 2 + 8$$

$$10 = 2 + 8$$

$$9 = 1 + 8$$

.....

- Compute redundant bits, find the parity of the data bits that has the redundant bits in its sum
- Eg check bit position 1 occurs in
11,9,7,5,3
- e.g. in notebook

Backword Error Correction (BEC)

Ceena Mathews, PNC



Automatic Repeat reQuest(ARQ)

effect of ARQ is to turn an unreliable data link into a reliable one.

It includes mechanisms such as

1. **Error detection: As discussed earlier.**
2. **Positive acknowledgment(ACK): The destination returns a positive acknowledgment to successfully received, error-free frames.**
3. **Retransmission after timeout: The source retransmits a**



Three versions of ARQ have been standardized:

1. Stop-and-wait ARQ
2. Go-back-N ARQ
3. Selective-reject ARQ / Selective Repeat ARQ




Stop-and-Wait ARQ

based on the stop-and-wait flow control technique

source station transmits a single frame and then must await an acknowledgment (ACK).

No other data frames can be sent until the destination station's reply arrives at the source station

Two sorts of errors could occur.

- 
- **First**, the frame that arrives at the destination could be damaged.
 - The receiver **detects this by using the error-detection technique** and simply **discards the frame**.
 - To account for this possibility, the source **station is equipped with a timer**.
 - After a frame is transmitted, the source station waits for an acknowledgment.
 - If no acknowledgment is received by the time that the **timer expires**, then the same frame is sent again.
 - this method requires that the **transmitter maintain a copy of a transmitted frame until an acknowledgment** is received for that frame.

The **second sort of error** is a damaged acknowledgment.

Consider the following situation.

- Station A sends a frame.
- The frame is received correctly by station B, which responds with an acknowledgment (ACK).
- The **ACK is damaged in transit** and is not recognizable by A, which will therefore **time out and resend the same frame**.
- This **duplicate frame arrives and is accepted by B**.
- B has therefore **accepted two copies of the same frame as if they were separate**.
- To avoid this problem, **frames are alternately labeled with 0 or 1**, and positive acknowledgments are of the form ACK0 and ACK1.
- In keeping with the sliding-window convention, an ACK0 acknowledges receipt of a frame numbered 1 and indicates that the receiver is ready for a frame numbered 0.

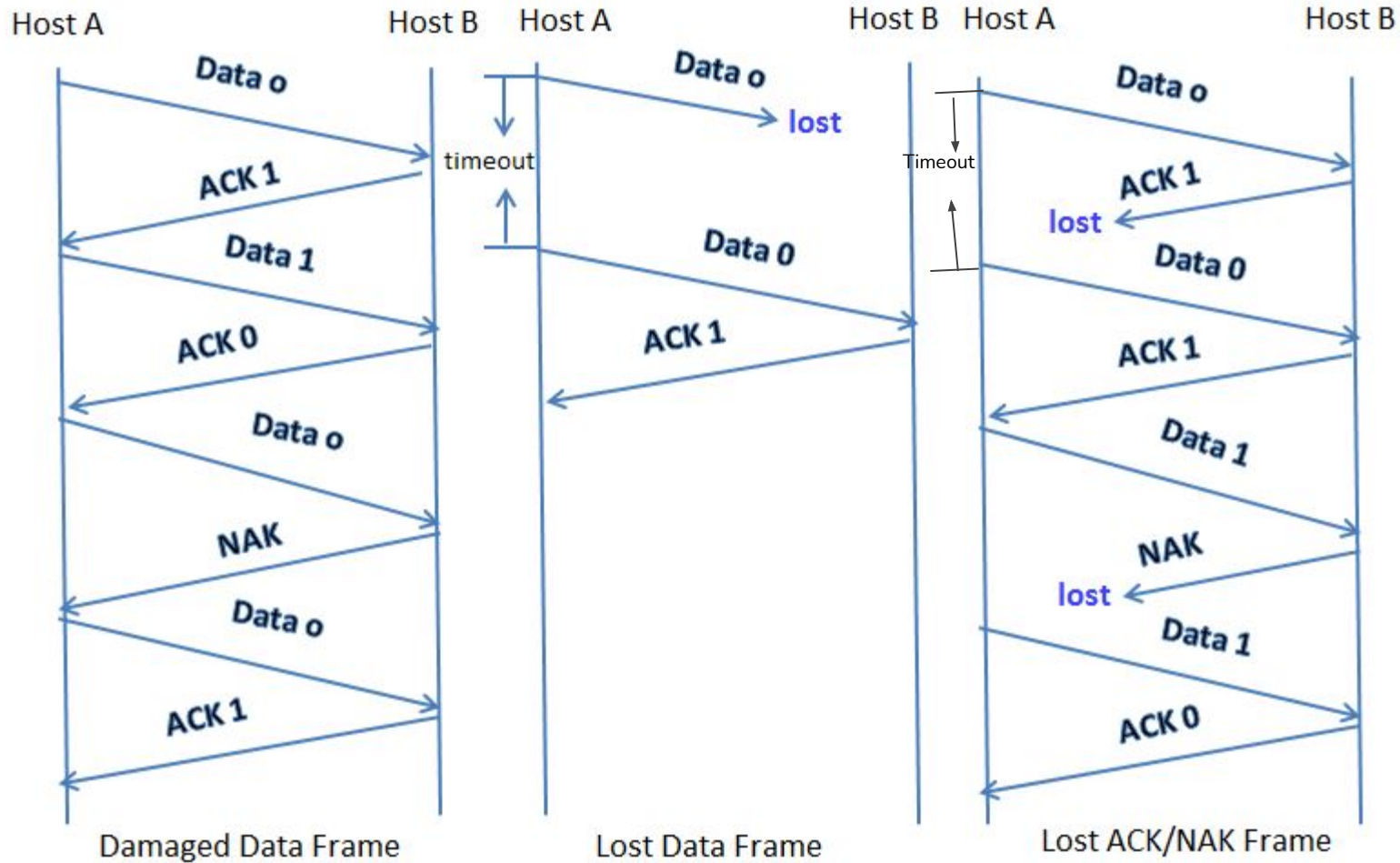


Fig: Stop and Wait ARQ



- **principal advantage** of stop-and-wait ARQ is its simplicity.
- Its **principal disadvantage** is that it is an inefficient mechanism.
- The **sliding-window flow control technique** can be adapted to provide more efficient line use
- it is sometimes referred to as ***continuous ARQ***.



Go-Back-N ARQ

- In this method, a station may send a series of frames sequentially numbered modulo some maximum value.
- The **number of unacknowledged frames outstanding** is determined by window size,
- when **no errors occur**, the destination will acknowledge incoming frames as usual (**RR = receive ready, or piggybacked acknowledgment**).
- If the destination station **detects an error** in a frame, it may send a **negative acknowledgment** (REJ = reject) for that frame,
- The **destination station will discard that frame and all future incoming frames until the frame in error is correctly received.**

- Thus, the source station, when **it receives a REJ, must retransmit the frame in error plus all succeeding frames that were transmitted in the interim.**
- After each transmission, A sets an acknowledgment timer for the frame just transmitted.
- Suppose that B has previously successfully received frame $(i - 1)$ and A has just transmitted frame i .
- The go-back-N technique takes into account the following contingencies:
 - 1. Damaged frame.** If the received frame is invalid (i.e., B detects an error, or the frame is so damaged that B does not even perceive that it has received a frame), B discards the frame and takes no further action as the result of that frame. There are two subcases:



(a) Within a reasonable period of time, A subsequently sends frame $(i + 1)$. B receives frame $(i + 1)$ out of order and sends a REJ i . A must retransmit frame i and all subsequent frames.

b) A does not soon send additional frames. B receives nothing and returns neither an RR nor a REJ. When A's timer expires, it transmits an RR frame that includes a bit known as the P bit, which is set to 1. B interprets the RR frame with a P bit of 1 as a command that must be acknowledged by sending an RR indicating the next frame that it expects, which is frame i . When A receives the RR, it retransmits frame i . Alternatively, A could just retransmit frame i when its timer expires.

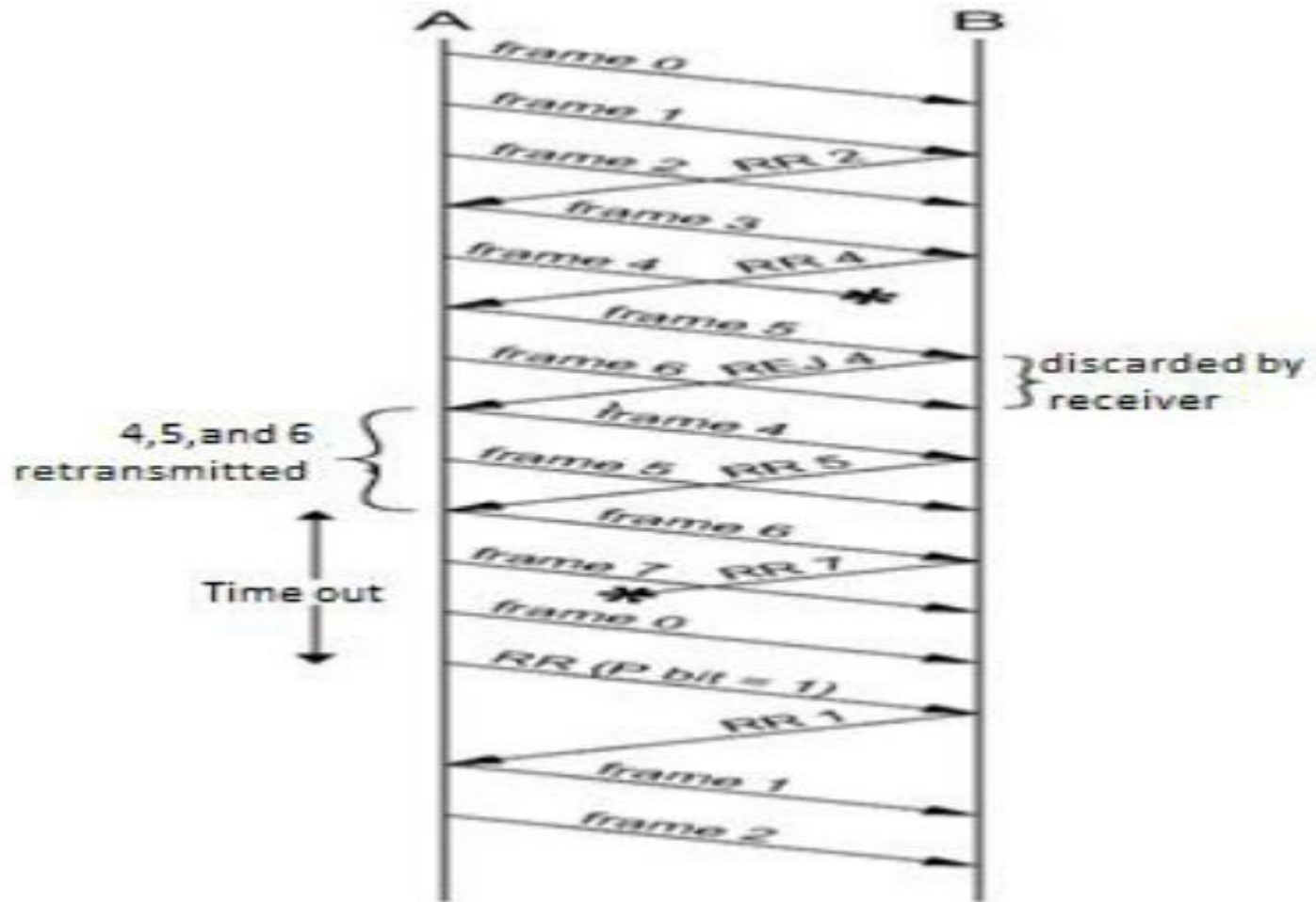
2. Damaged RR. There are two subcases:

(a) B receives frame i and sends RR ($i + 1$), which suffers an error in transit. Because acknowledgments are cumulative (e.g., RR 6 means that all frames through 5 are acknowledged), it may be that A will receive a subsequent RR to a subsequent frame and that it will arrive before the timer associated with frame i expires.

(b) If A's timer expires, it transmits an RR command as in Case 1b. It sets another timer, called the P-bit timer. If B fails to respond to the RR command, or if its response suffers an error in transit, then A's P-bit timer will expire. At this point, A will try again by issuing a new RR command and restarting the P-bit timer. This procedure is tried for a number of iterations. If A fails to obtain an acknowledgment after some maximum number of attempts, it initiates a reset procedure.



3. Damaged REJ. If a REJ is lost, this is equivalent to Case 1b.

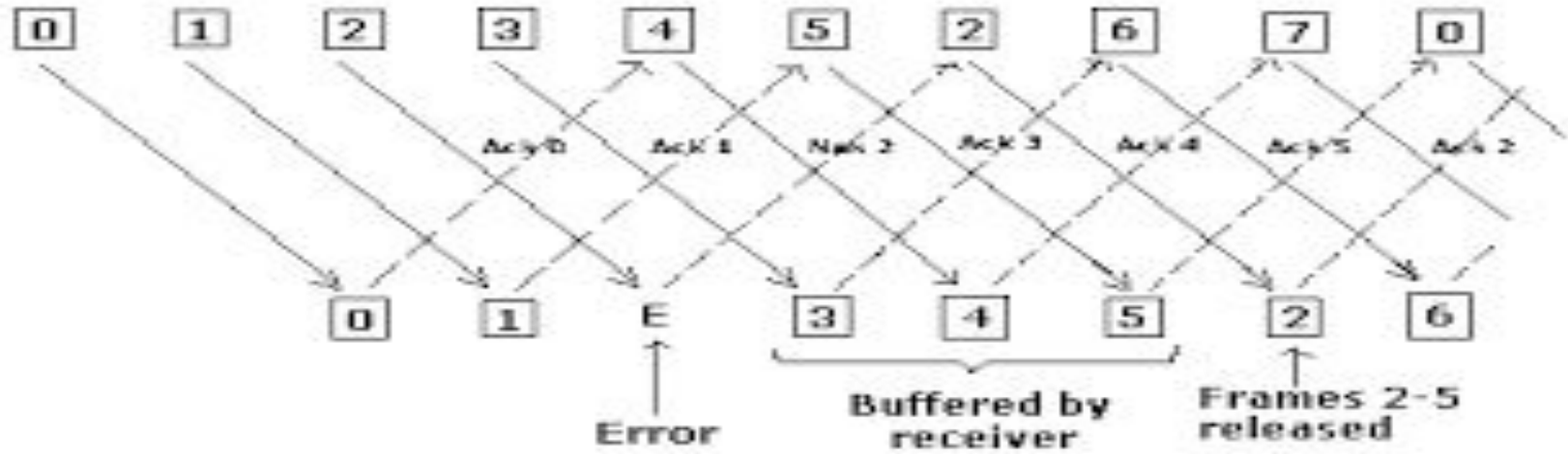


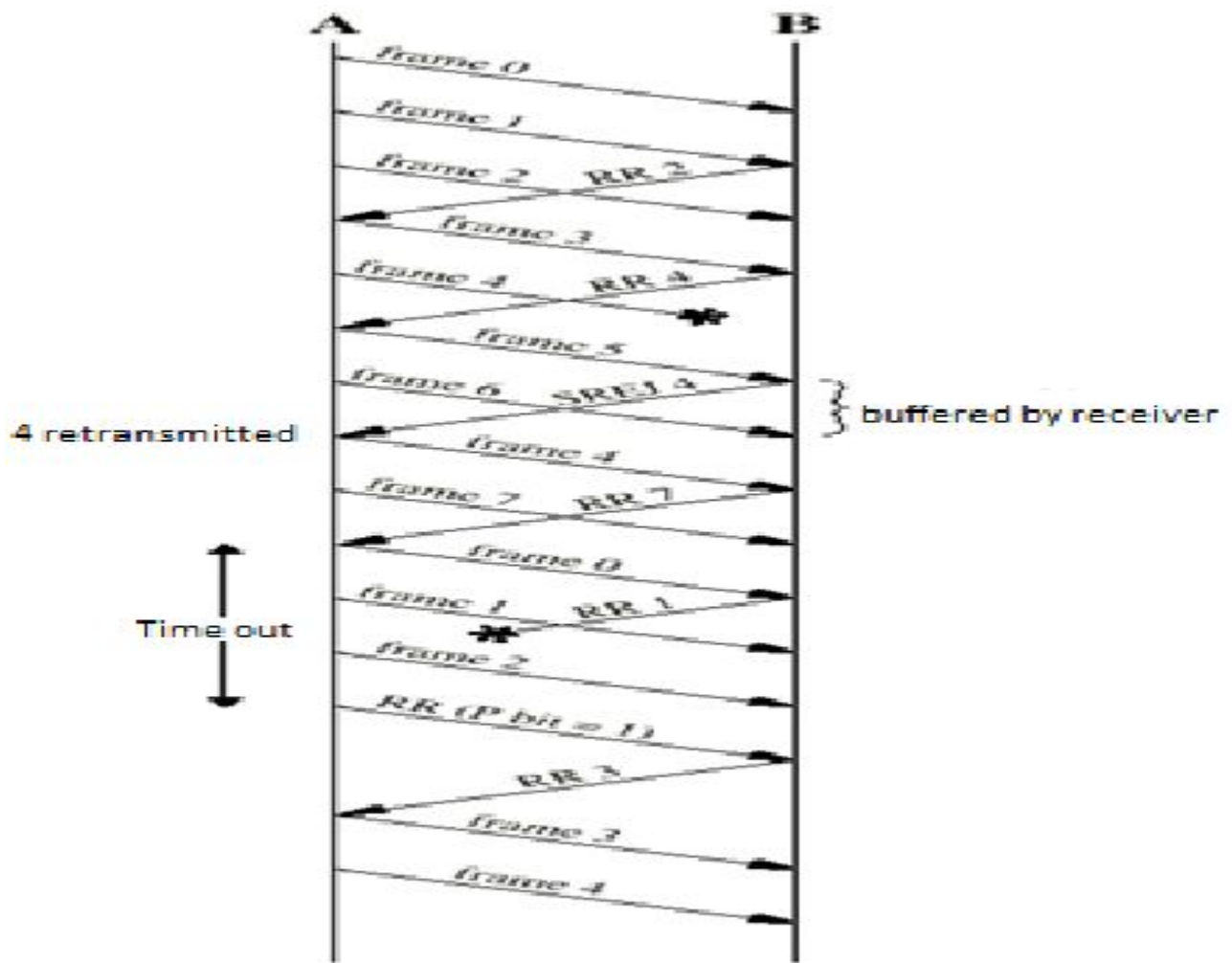
Selective-Reject ARQ

the only frames retransmitted are those that receive a negative acknowledgment, in this case called SREJ, or those that time out.

Selective reject would **appear to be more efficient than go-back-N**, because it minimizes the amount of retransmission.

But in **Selective-Reject ARQ** the receiver must maintain a buffer large enough to **save post-SREJ frames until the frame in error is retransmitted and must contain logic for reinserting that frame in the proper sequence.**







In transmitter, too, **requires more complex logic to be able to send a frame out of sequence.**

Because of such complications, **select-reject ARQ is much less widely used than go-back-N ARQ.**

Selective reject is a **useful choice for a satellite link because of the long propagation delay involved.**

- The window size limitation is more restrictive for selective-reject than for go-back-N.
- Consider the case of a 3-bit sequence number size for selective-reject. Allow a window size of seven, and consider the following scenario
 - **Station A sends frames 0 through 6 to station B.**
 - **Station B receives all seven frames and cumulatively acknowledges with RR 7.**
 - **Because of a noise burst, the RR7 is lost.**
 - **A times out and retransmits frame 0.**

B has already advanced its receive window to accept frames 7, 0, 1, 2, 3, 4, and 5. Thus it assumes that frame 7 has been lost and that this is a new frame 0, which it accepts.

- problem with the foregoing scenario is that **there is an overlap between the sending and receiving windows.**
- To overcome the problem, **the maximum window size should be no more than half the range of sequence numbers.**
- In the preceding scenario, **if only four unacknowledged frames may be outstanding, no confusion can result.**
- In general, for a k -bit sequence number field, which provides a $2^k - 1$ sequence number range of 2^k , the maximum window size is limited to 2^{k-1} .

Chapter 12

Multiple Access

DATA LINK CONTROL

- Data Link Control include the following functions
 - Line Discipline
 - Coordinate link systems
 - "Who should send now?"
 - Flow Control
 - Coordinate the amount of data that can be sent before receiving acknowledgement
 - "How much data may be sent?"
 - Error Control
 - Allows the receiver to inform the sender of any frames lost or damaged in transmission
 - "How can errors be detected and corrected?"

ernitong

Line Discipline

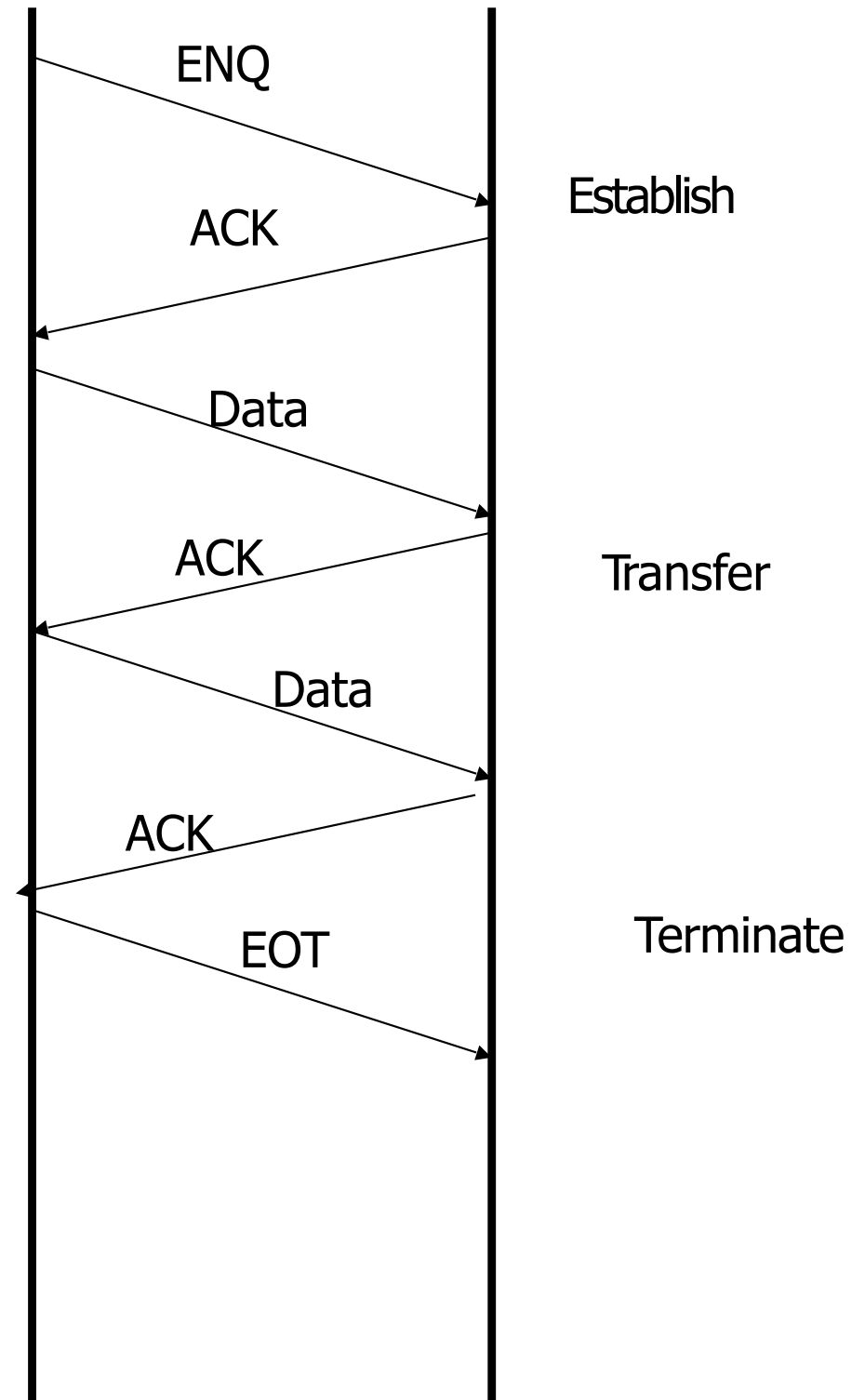
- This function of the data link layer oversees the establishment of links and the right of a particular device to transmit at a given time
- Categories:
 - ENQ / ACK
 - Used on peer to peer communication
 - Poll / Select
 - Used in a primary - secondary communication

ernitong

- SIMPLE ENQ/ACK protocol
- (e.g., "peer-to-peer", dedicated line)

A session may be initiated by any device of equal status

ENQ - enquiry
ACK - Acknowledgement
EOT - end of transmission

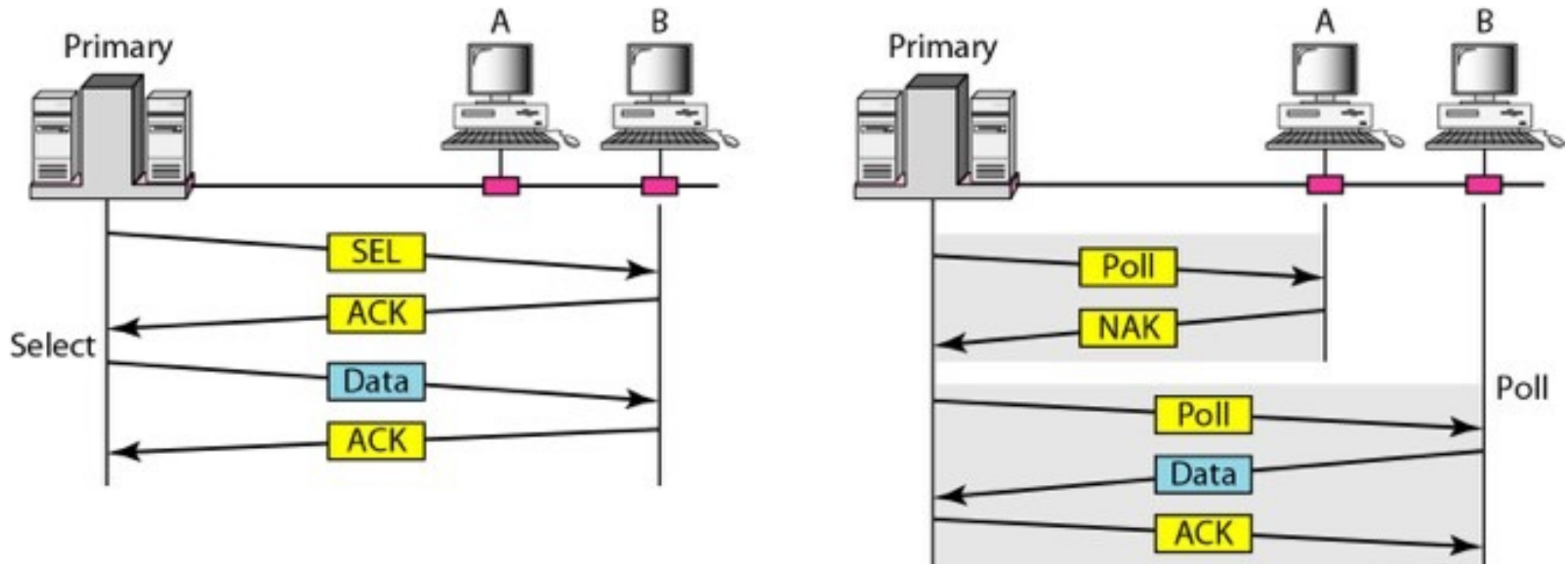


Polling

- One device is designated as a primary station.
- Other devices are secondary stations
- All data exchanges are done thru primary when ultimate destination is secondary
- **Primary device controls the link**
- **Secondary follows the instructions**
- Primary device decides **which device is allowed to use the channel at a given time**

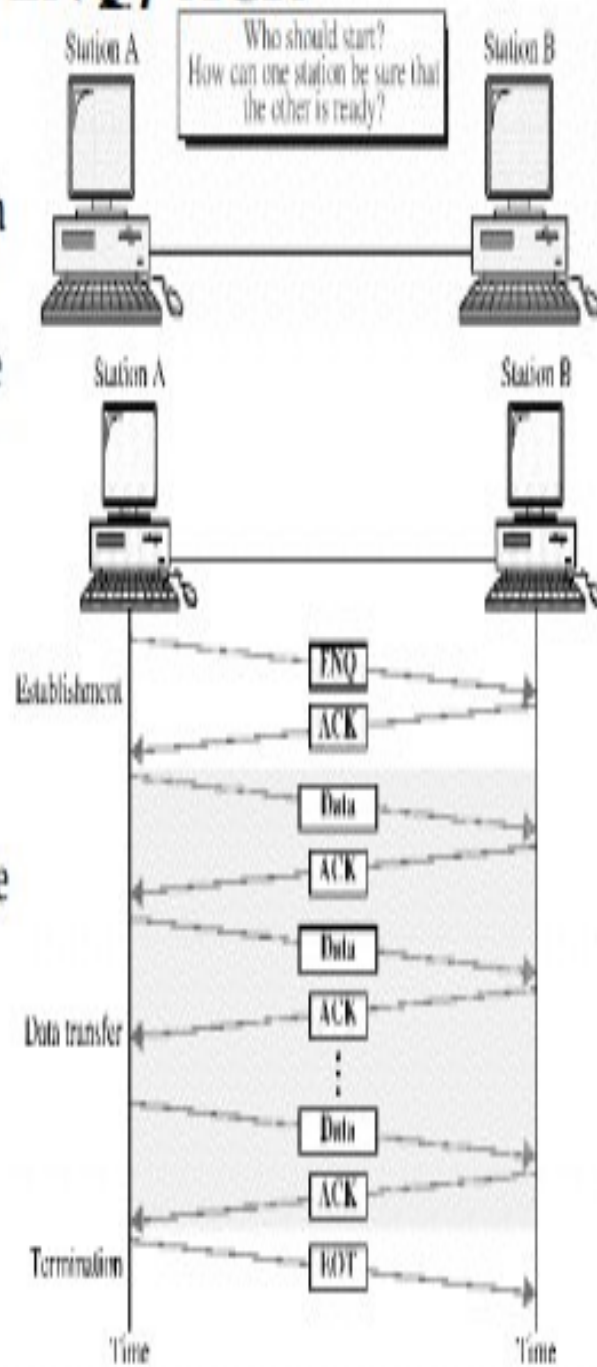
- **Primary is the initiator of the session**
- If primary is ready to receive data, it asks the secondary if they have anything to send **-poll function**
- If primary wants to send data it tells the secondary to get ready to receive – **select**

Figure 12.19 Select and poll functions in polling access method



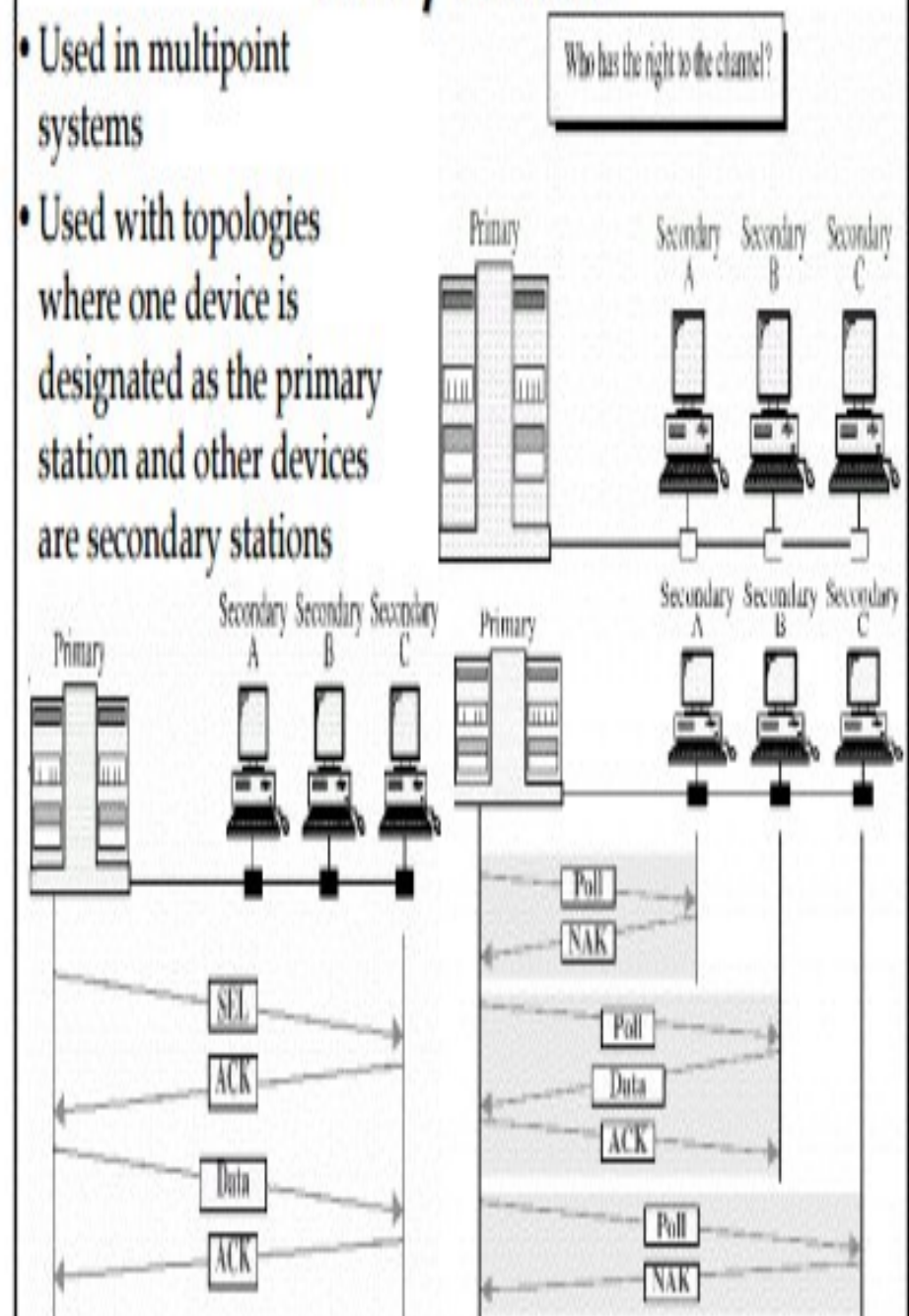
ENQ/ACK

- Used primarily in systems where there is no question of the wrong receiver getting the transmission
- A session can be initiated by any device on a link of equal rank
- An initiating device ordinarily makes three attempts to establish a link before giving up



Poll/Select

- Used in multipoint systems
- Used with topologies where one device is designated as the primary station and other devices are secondary stations



Data link layer divided into two functionality-oriented sublayers

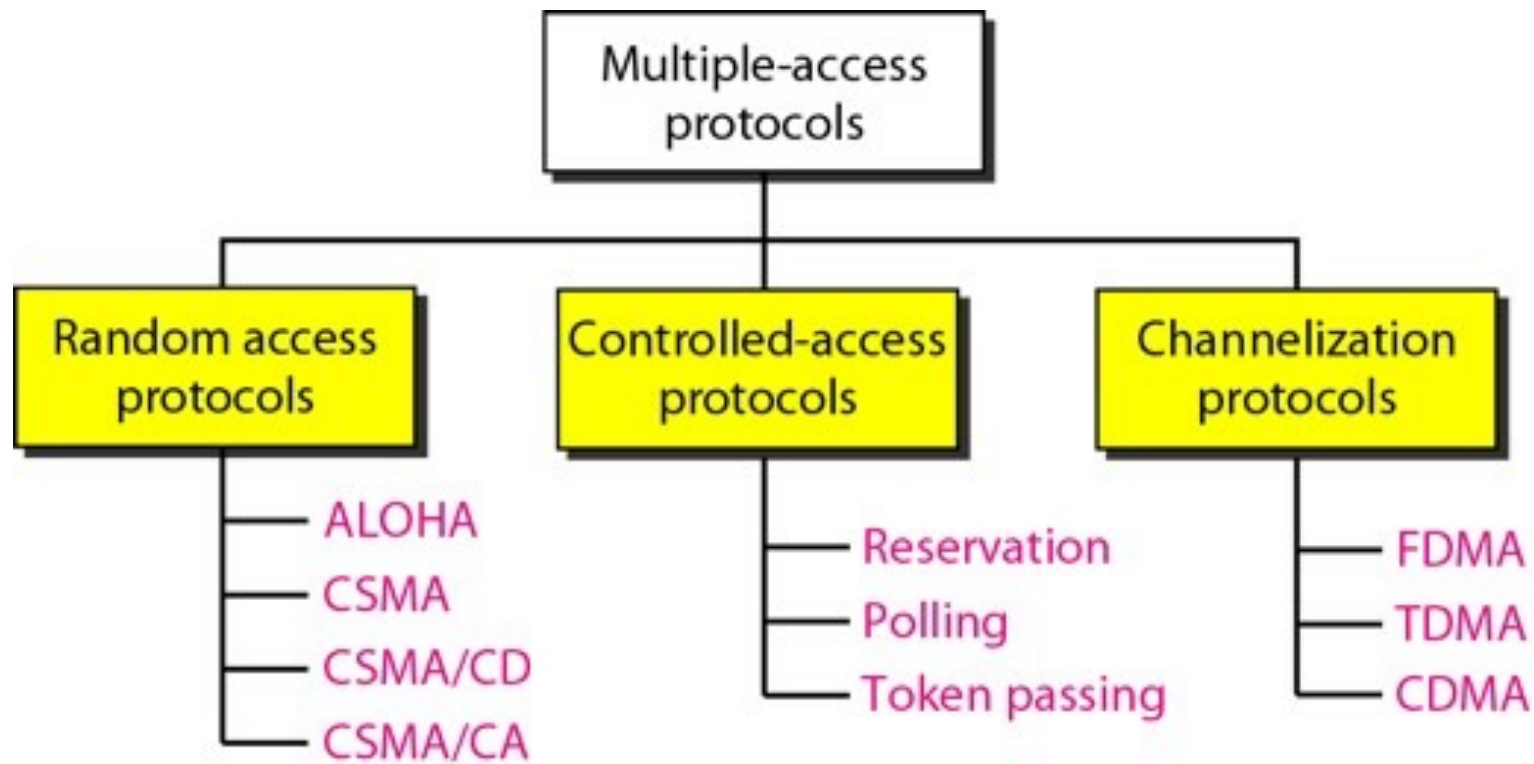
1. Logical link control (LLC)

- responsible for flow and error control

2. Medium Access Control(MAC)

- In multipoint and broadcast link, more stations share the communication medium
- Medium access resolution

Taxonomy of multiple-access protocols



RANDOM ACCESS

- random access or contention methods
- **no station is superior to another station and none is assigned the control over another.**
- a station that has data to send uses a procedure defined by the protocol to make a decision on whether or not to send.
- If more than one station tries to send, there will be collision

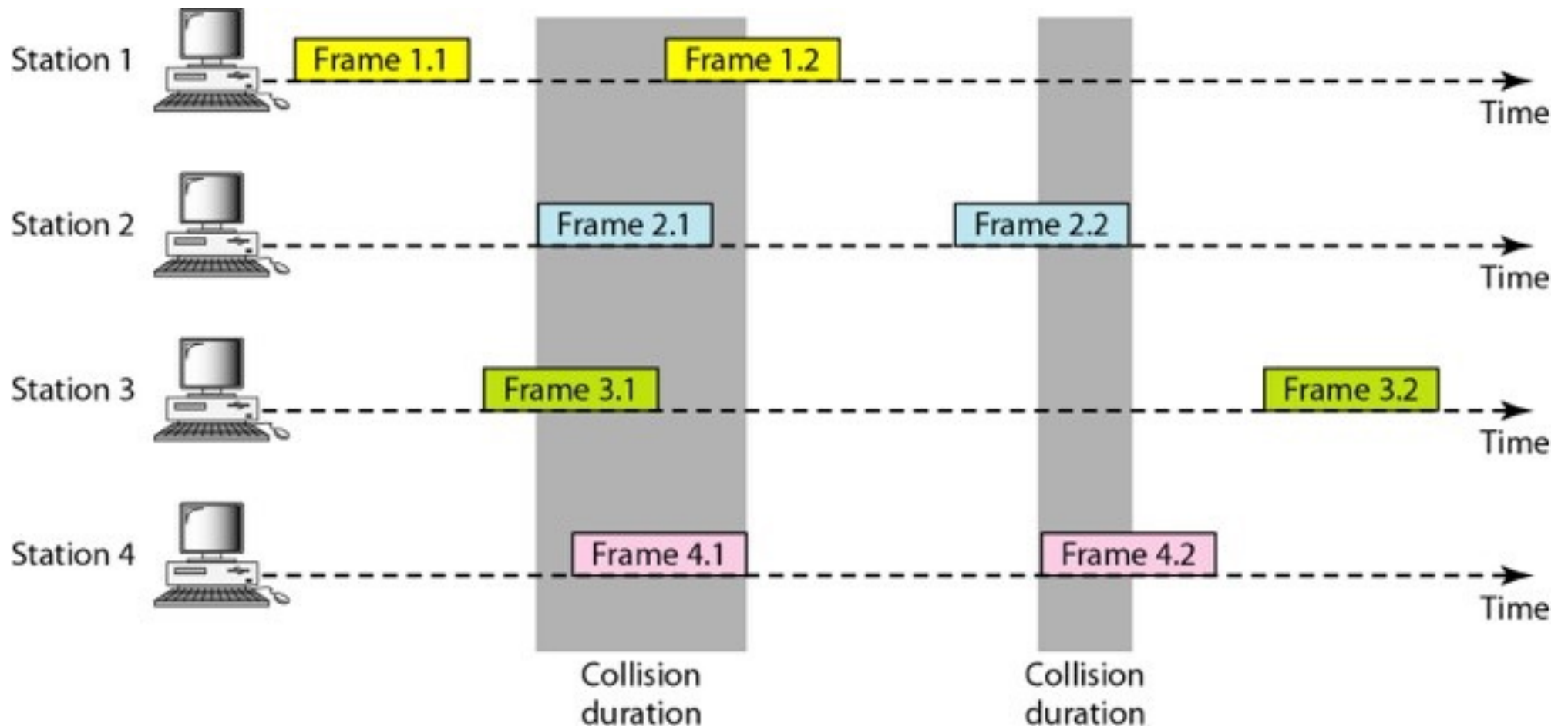
ALOHA Protocols

- Developed at **university of Hawaii in 1970**
- Was designed for **wireless LAN**
- can be used for **any shared medium**
- Uses **multiple access(MA) procedure**

- **Pure ALOHA(original Aloha)**

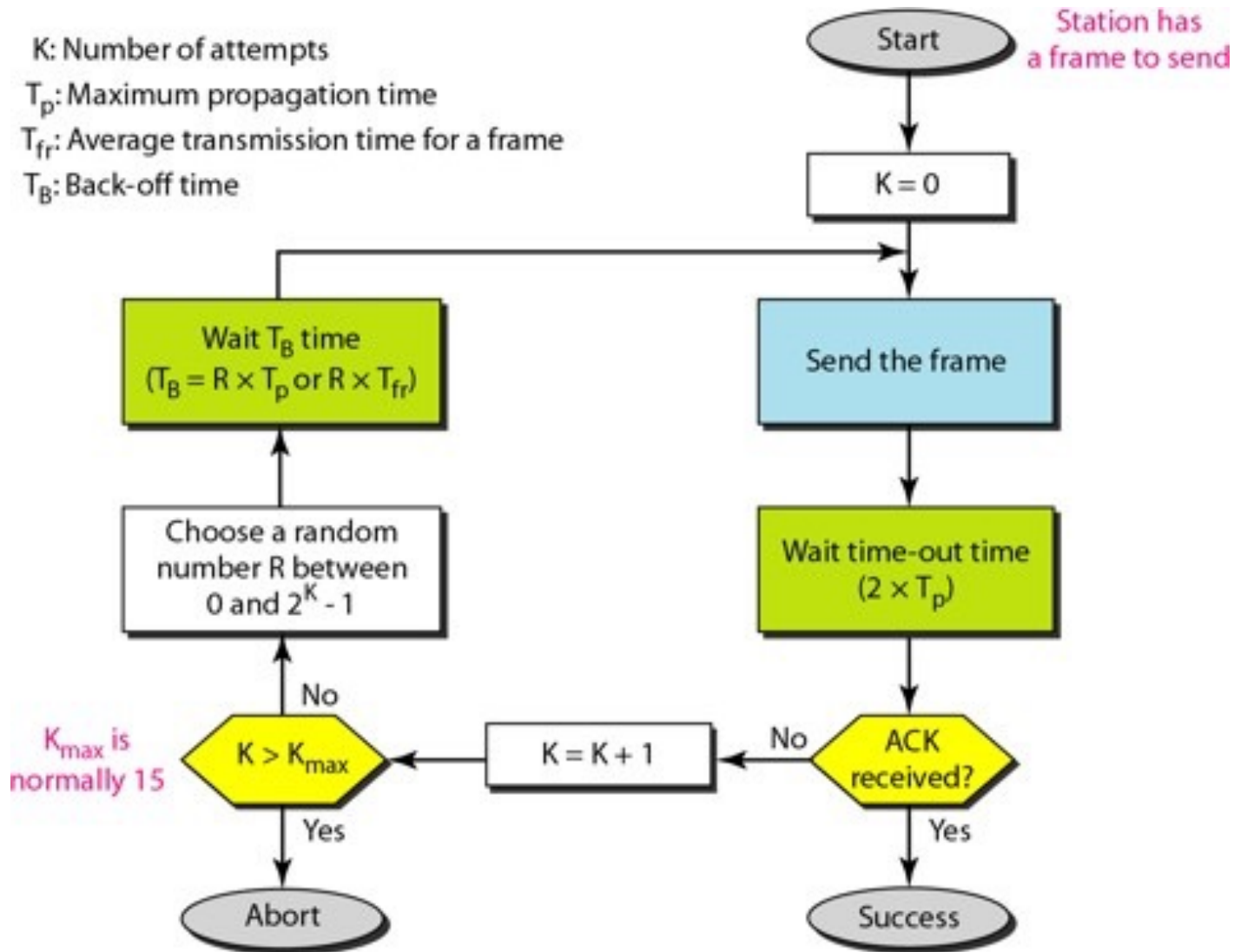
- Each station **sends a frame whenever it has a frame to send**
- There is **possibility of collision between frames from different stations**
- Stations transmit at equal **transmission time**
- **After transmitting a frame, the sender waits for an acknowledgment for an amount of time (time out) equal to the maximum round-trip propagation delay $= 2 * t_{prop}$**

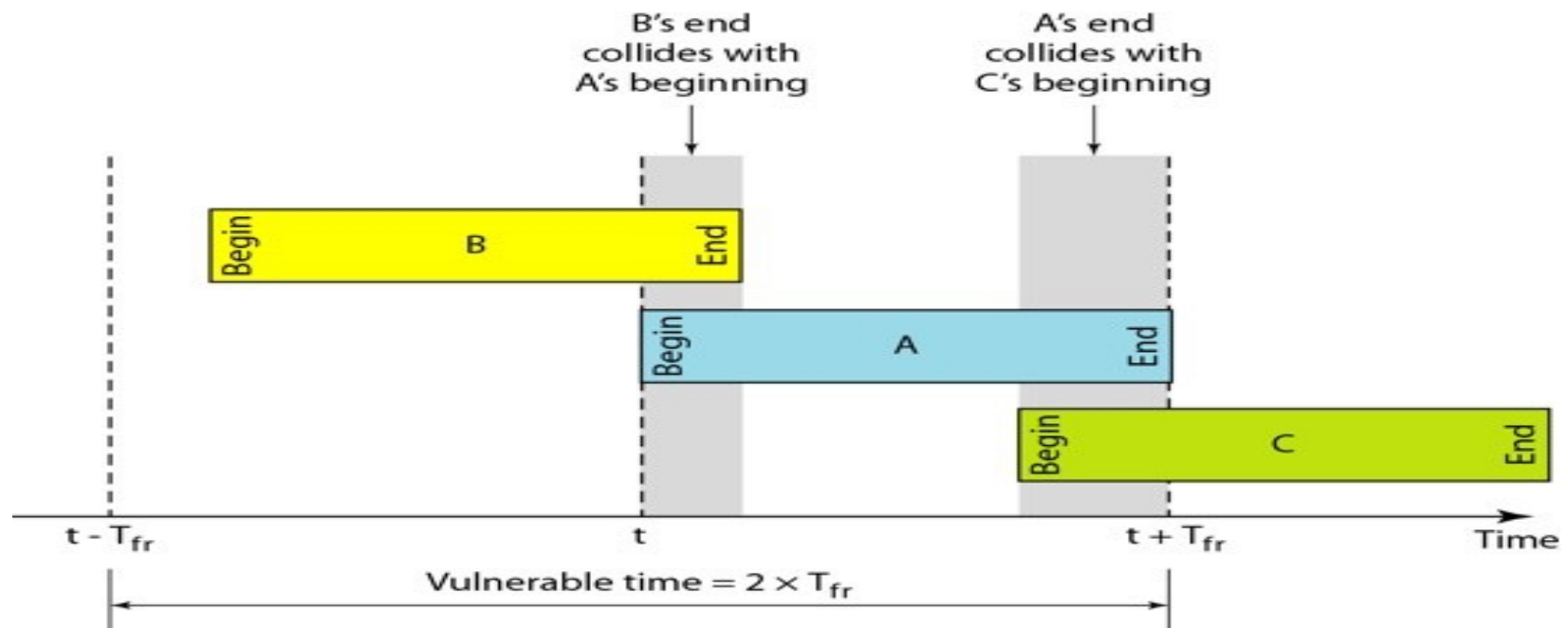
Frames in a pure ALOHA network



- If **no ACK** was received, sender assumes that the **frame or ACK** has been destroyed and **resends** that frame after it **waits for a *random amount of time* (back off time T_B)**
- T_B is computed using **binary exponential back-off**
- **Binary exponential back-off :**
 - ✳ a number in the range of 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)
- If station fails to receive an ACK after repeated transmissions (normally 15), it gives up

Procedure for pure ALOHA protocol





Vulnerable time : length of time in which there is possibility of collision

G – average number of frames generated during one frame transmission time

Average number of successful transmissions ie.,
The throughput (S) for pure ALOHA is

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

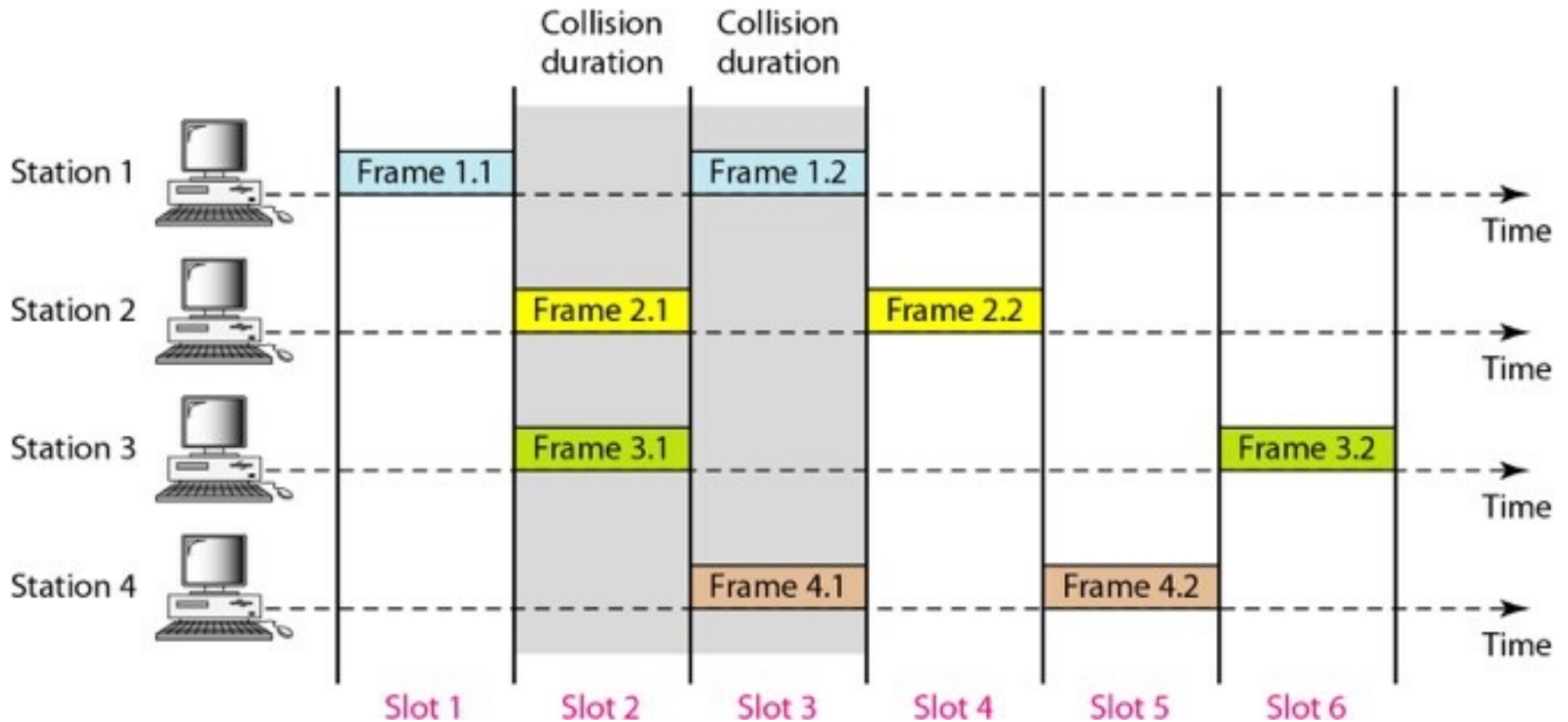
The maximum throughput

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

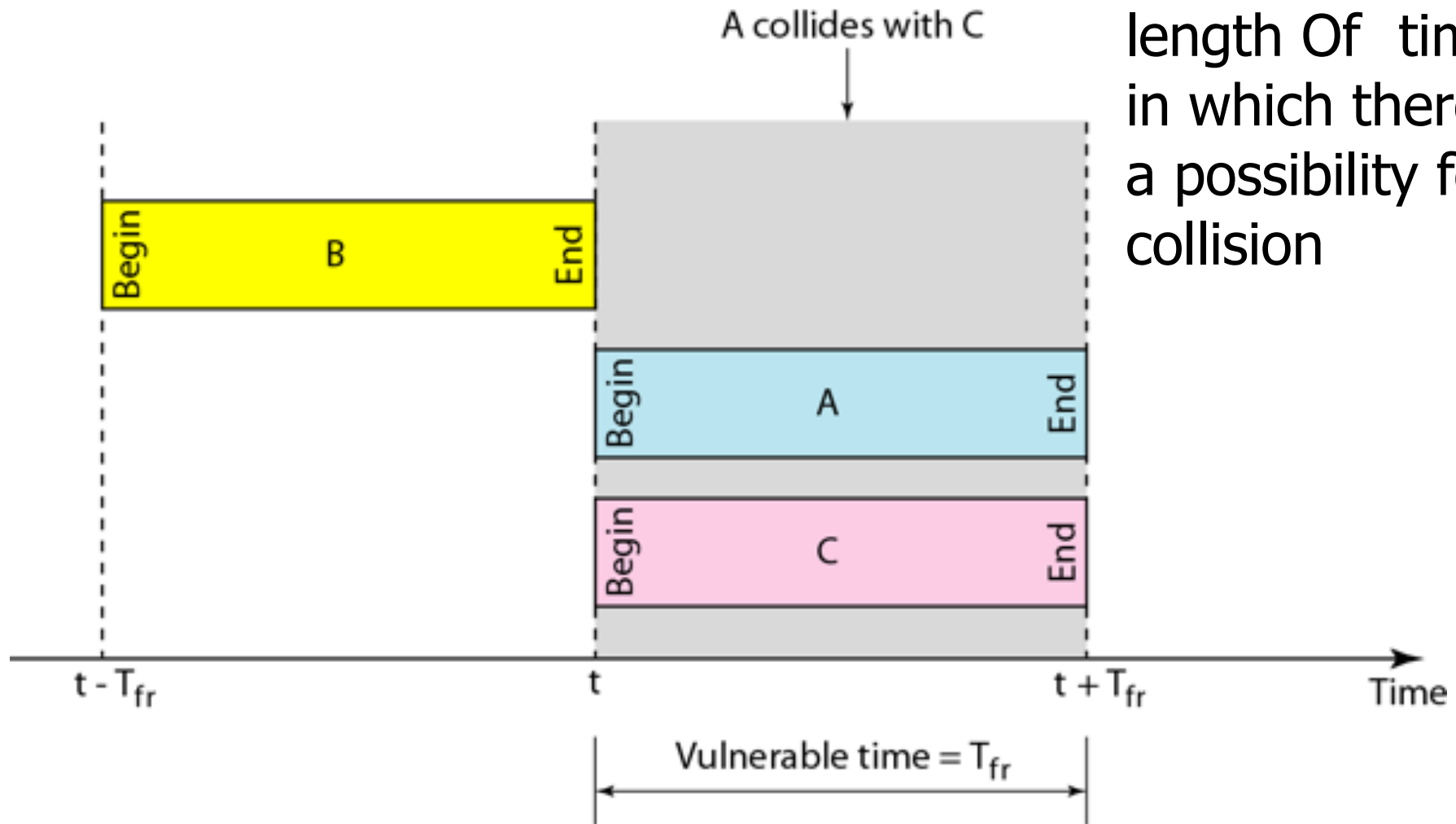
Slotted Aloha

- Time is divided into slots equal to a **frame transmission time (T_{fr})**
- A station can transmit at the beginning of a slot only
- If a station misses the beginning of a slot, it has to wait until the beginning of the next time slot.
- **A central clock** or station informs all stations about the start of a each slot
- Maximum channel utilization is **37%**

Frames in a slotted ALOHA network



Vulnerable time for slotted ALOHA protocol



Vulnerable time:
length Of time
in which there is
a possibility for
collision

The throughput for slotted ALOHA is

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

The maximum throughput

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

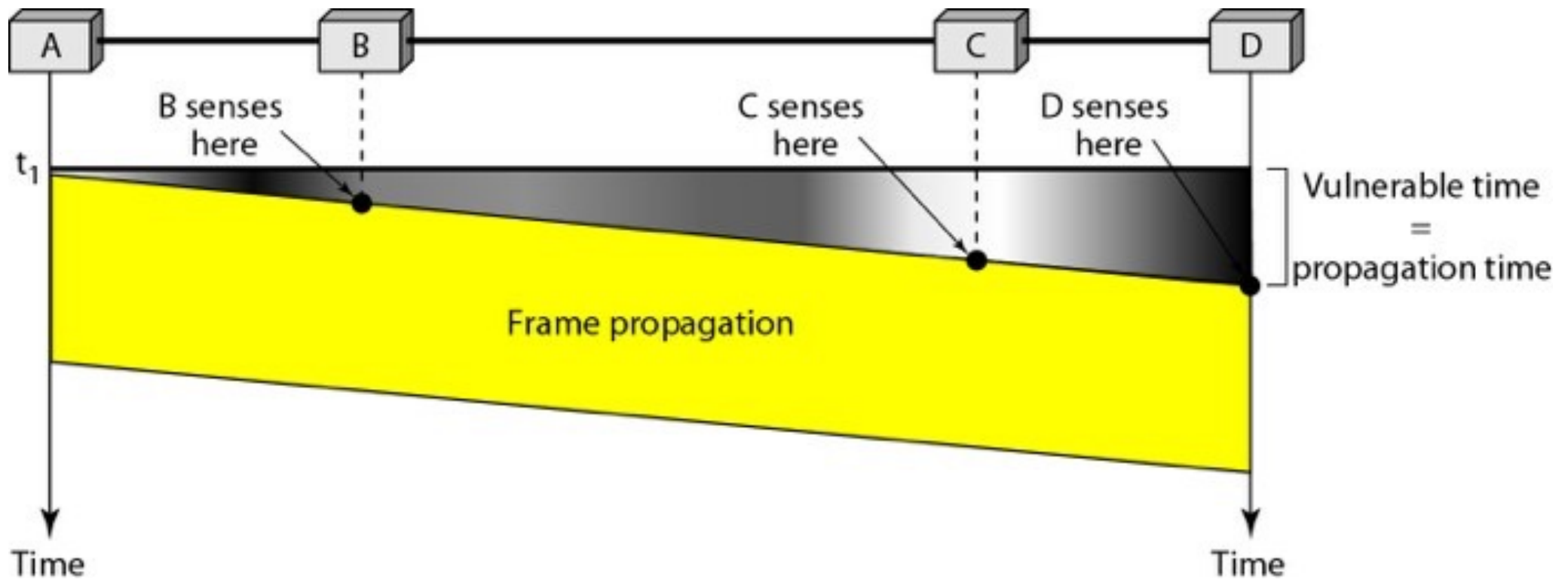
ie. 36.8% frame reach their destination successfully

Carrier Sense Multiple Access (CSMA)

- Based on the principle 'Sense before transmit'
- Each **station listen to the medium** to check the state of the medium before sending
- **Reduce the possibility of collision** but cannot eliminate it
- **possibility of collision because of propagation delay**

- A station senses the medium and find it idle only because the first bit sent by another station has not yet been received
- Vulnerable time = propagation time T_p

Figure 12.9 Vulnerable time in CSMA



Types of CSMA

1. **1-Persistent CSMA**
 - Sends the **frame immediately with probability 1** when the station finds the line idle
 - **Highest chance of collision** b/c 2 or more stations may find the line idle and transmit immediately
2. **Non-Persistent CSMA**
 - If a **station has frame to send**, it **senses** the line. Station **transmits frame if line is idle**
 - If line **is not idle**, it waits a random amount of time and senses the line again
 - **Reduces chance of collision** because it is unlikely that 2 or more stations waits same amount of time
3. **p-Persistent CSMA**

- **p-Persistent CSMA**
- Channel has **time slots with a slot duration equal to or greater than max propagation time**
- After the **station finds the line idle**, it follows the following steps
 1. Station sends frames with **propability p**
 2. Station then **waits for the beginning of the next time slot** with probability $q=1-p$ and senses the line again
- a) If line is idle repeats step 1
- b) If line is busy, it uses the back off pprocedure

Figure 12.11 Flow diagram for three persistence methods

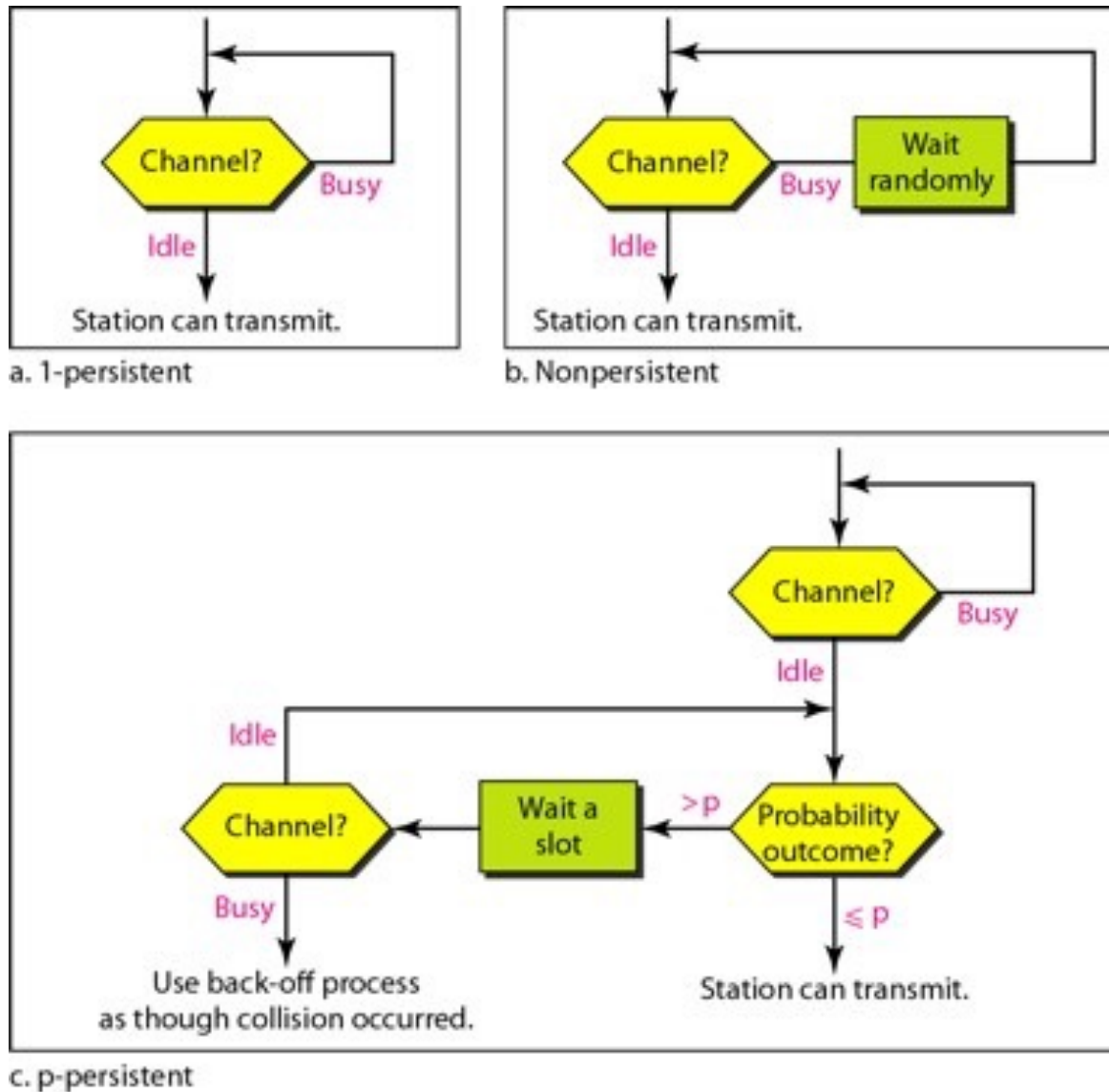
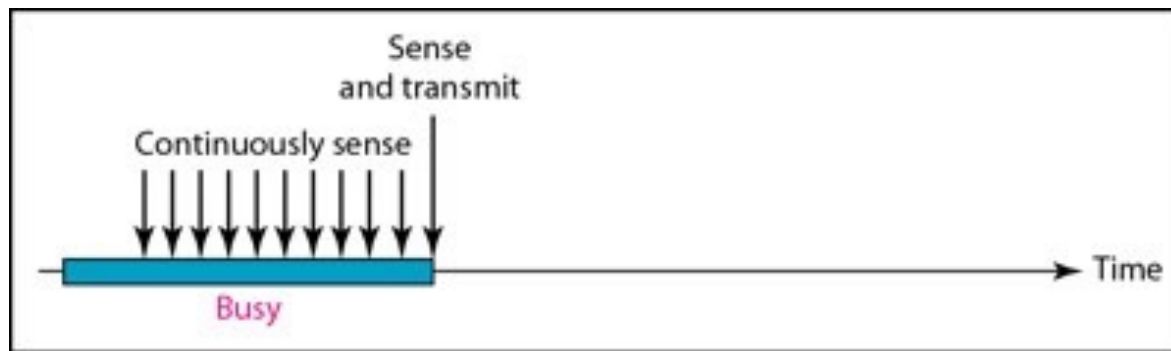
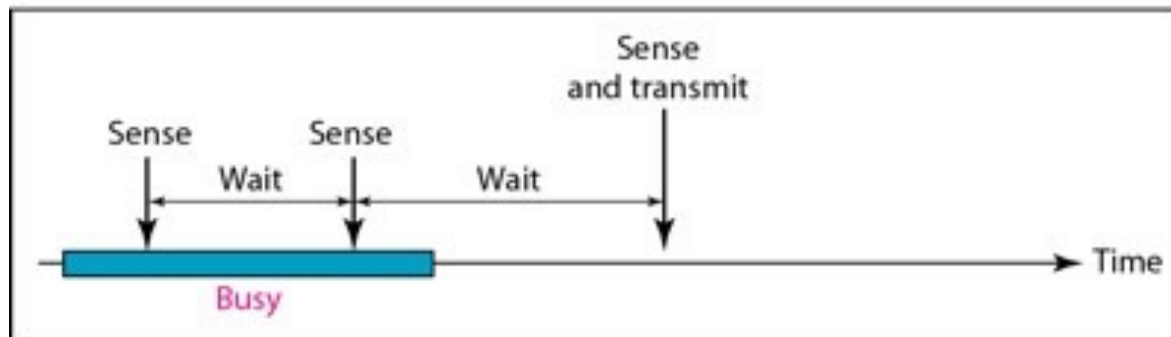


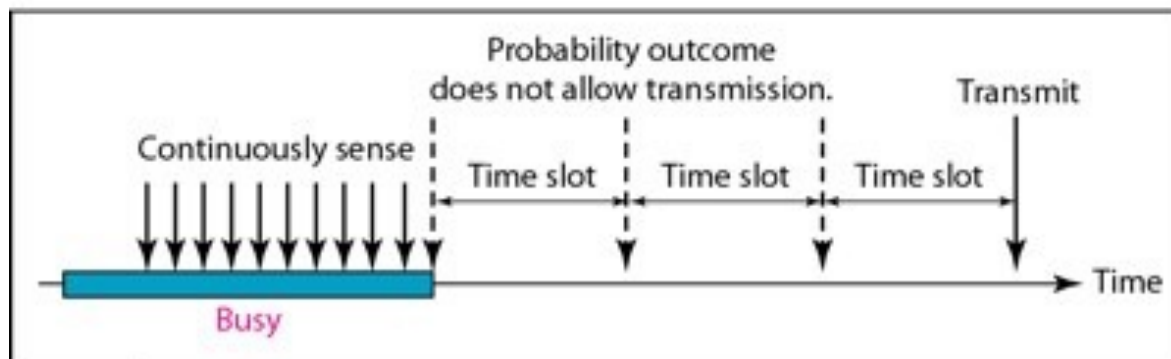
Figure 12.10 Behavior of three persistence methods



a. 1-persistent



b. Nonpersistent



c. p-persistent

Carrier Sense Multiple Access With Collision Detection(CSMA/CD)

- Station monitors the medium after it sends a frame to see if the transmission was successful.
- if there is collision the frame is sent again
- Otherwise it is successful
- Station transmits and receives continuously and simultaneously
- Frame transmission time must be atleast 2 times the max propagation time T_p

Carrier Sense Multiple Access With Collision Detection(CSMA/CD)

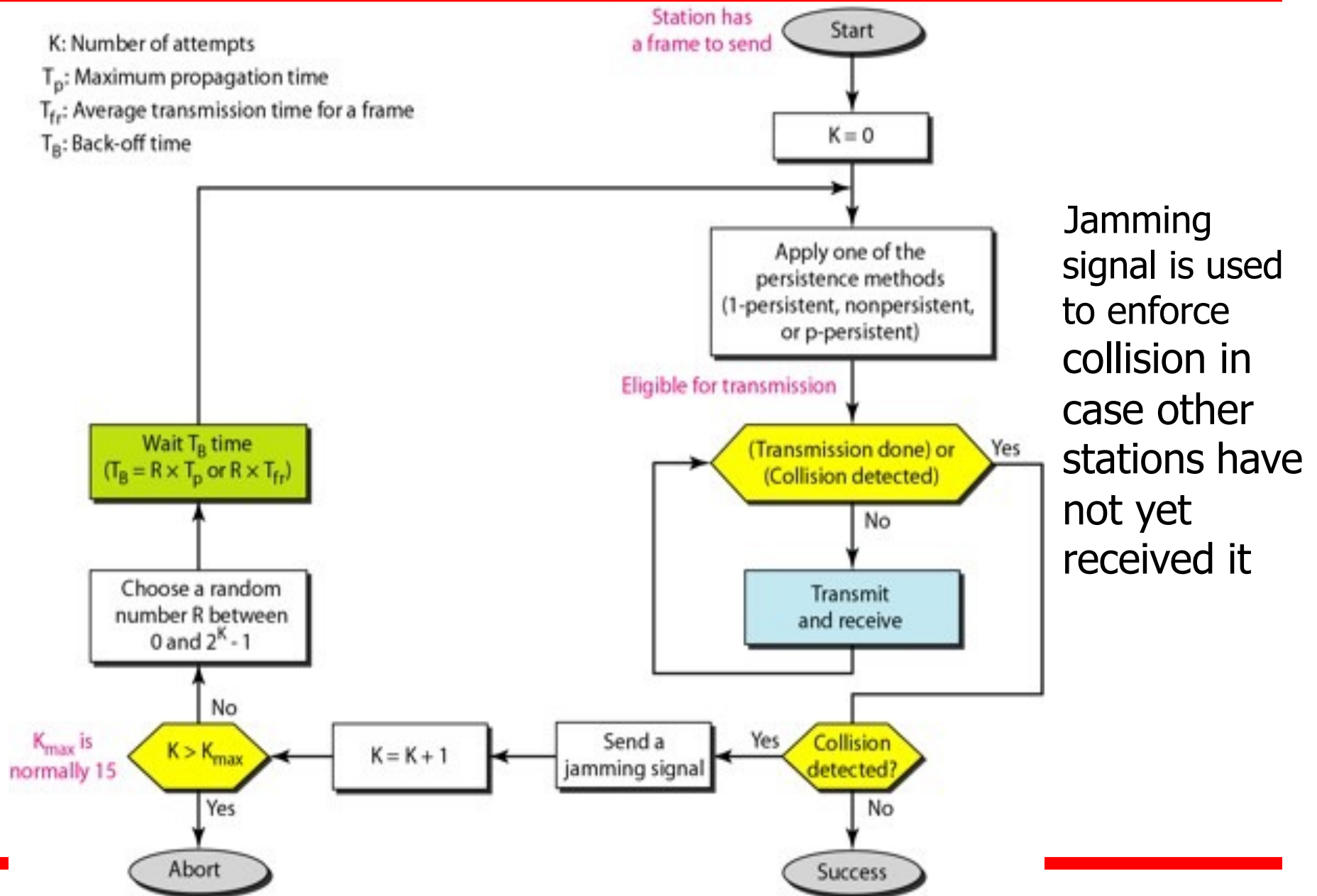


Figure 12.12 Collision of the first bit in CSMA/CD

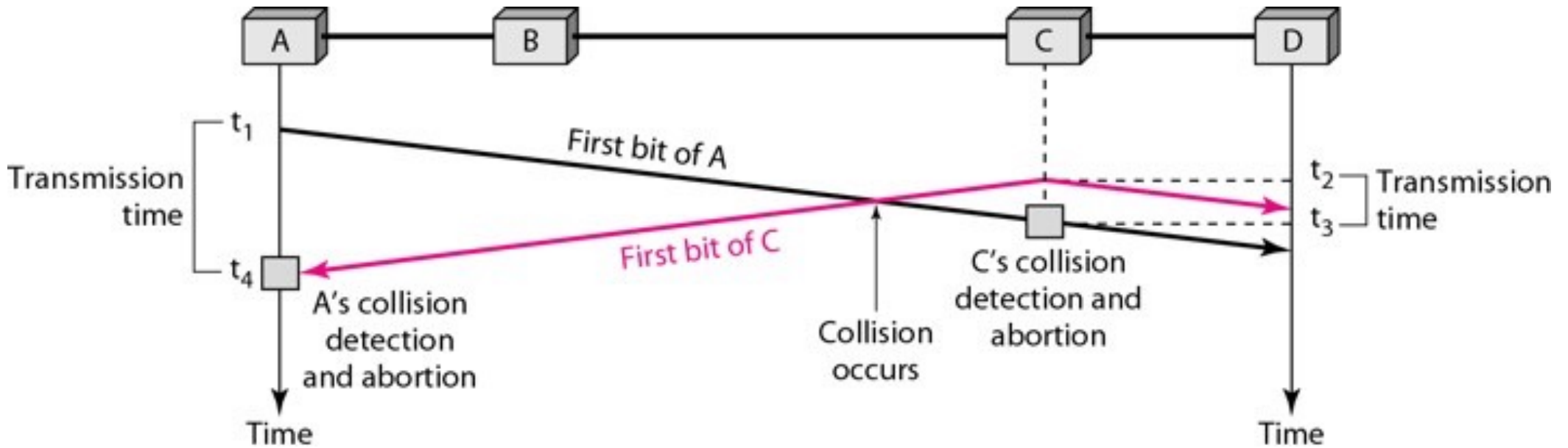
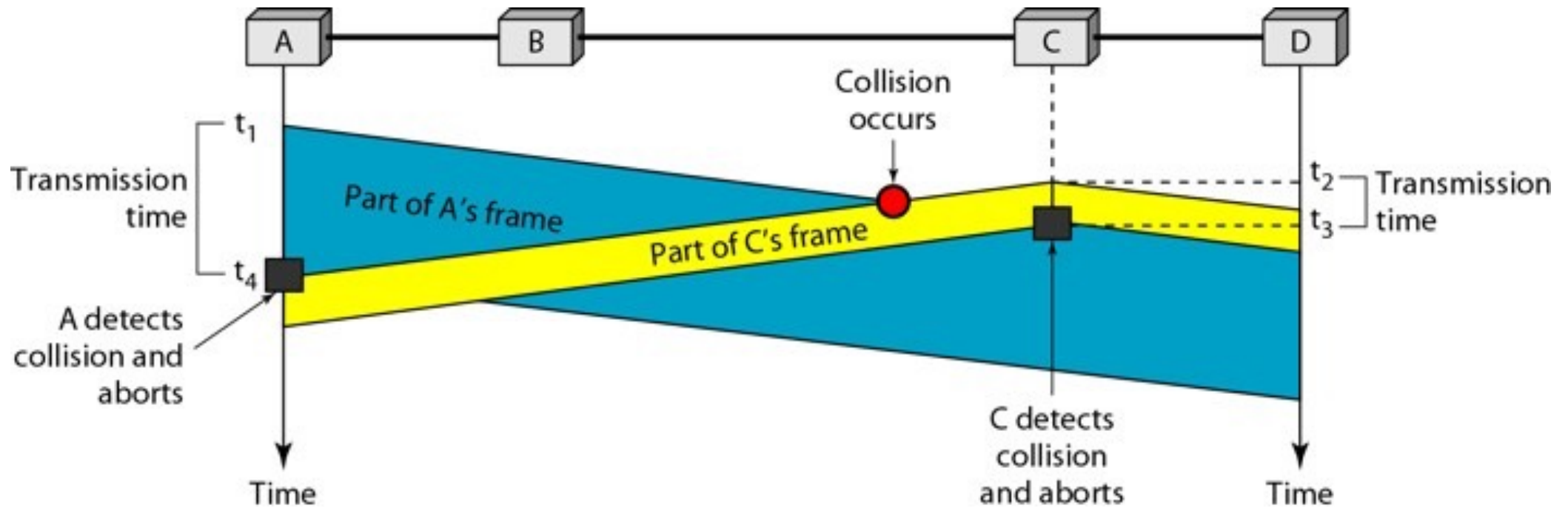


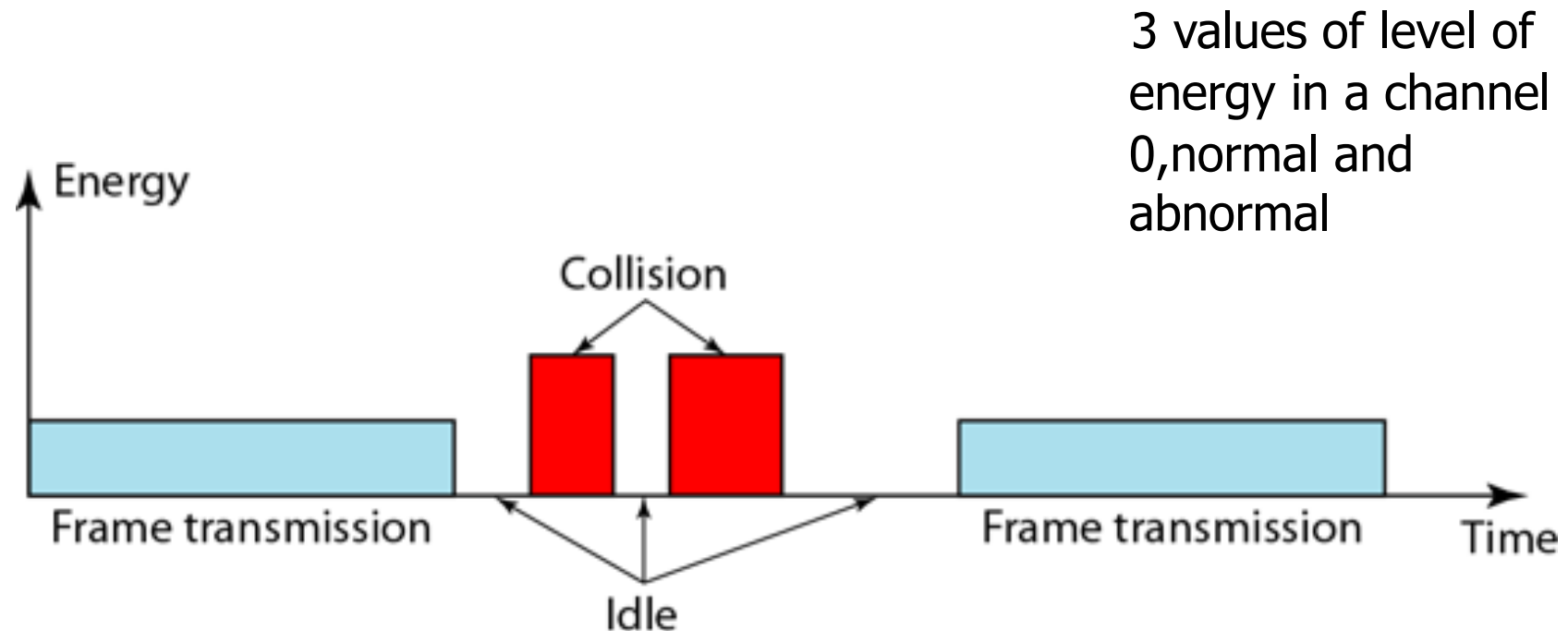
Figure 12.13 Collision and abortion in CSMA/CD



Minimum Frame size

- Before sending the last bit of the frame, sending station must detect a collision, if any and abort the transmission
- This is so because once the entire frame is sent , the station does not keep a copy of the frame and does not monitor the line for collision detection
- So the frame transmission time must be at least 2 times the maximum propagation time T_p

Figure 12.15 Energy level during transmission, idleness, or collision



On a wired network, energy level is almost double during a collision. This is how a receiver tells if there is a collision. But on a wireless network, energy level is not that high (barely 5-10% higher). So with wireless, we need to avoid collisions.

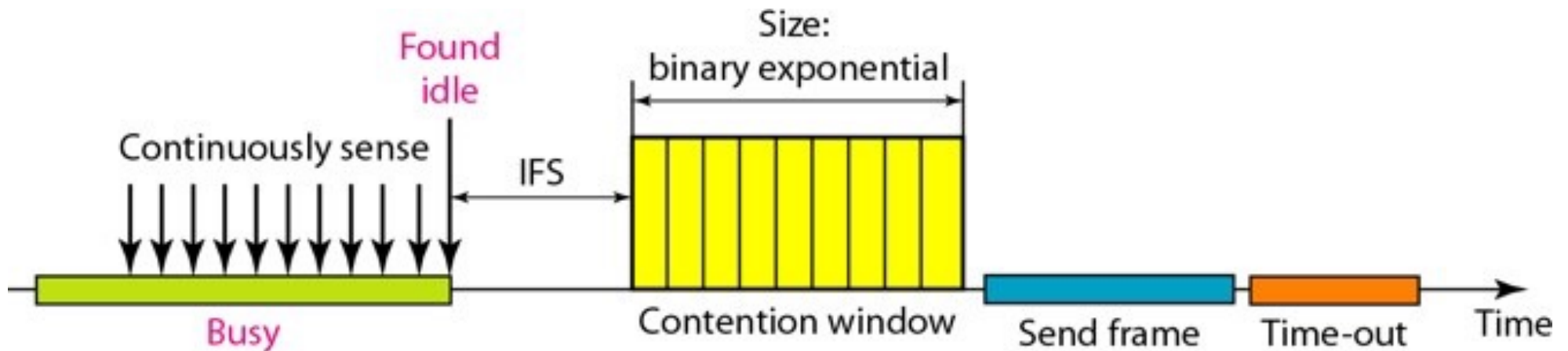
Throughput

- Greater than slotted aloha and pure aloha
- Max throughput occurs at different value of G and is based on the persistence method used and the value of p in p -persistent approach
 - Throughput is 50% for $G=1$ (1 persistent)
 - Throughput is 90% for G is between 3 and 8(non persistent)

Carrier Sense Multiple Access With Collision Avoidance(CSMA/CA)

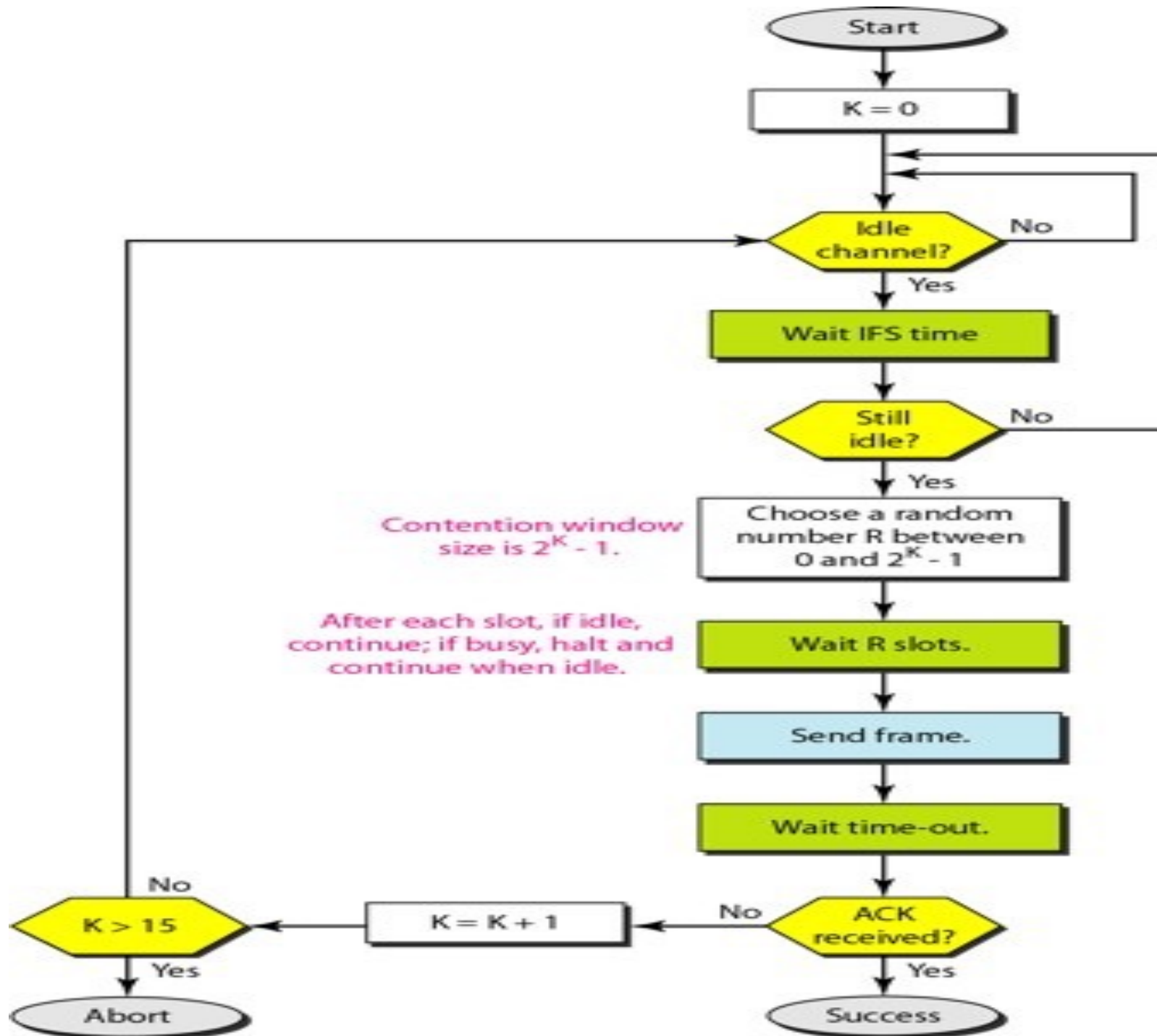
- Need to **avoid collisions on wireless networks because they cannot be detected**
- Collisions are avoided thru 3 strategies
 1. **Interframe Space(IFS)** (can be used to define the priority of a station)
 2. **Contention window** (amount of time divided into slots. A station chooses a random no. Of slots as its wait time.no of slots changes according to **binary exponential back off strategy**[it is set to one slot first time and doubles each time station cannot detect idle channel after IFS time])
 3. **Acknowledgements : positive acknowledgement and timeout guarantees**

Figure 12.16 Timing in CSMA/CA



IFS can be used to define the priority of a station or a frame. Higher priority station is assigned a shorter IFS

Flow diagram for CSMA/CA



- **COMPUTER NETWORKS**

UNIT #3

Chapter 15: Behrouz Forouzan

15-1 CONNECTING DEVICES

- *Connecting devices are used to connect LANs and segments of LAN.*
- *operate in different layers of the Internet model*
- *connecting devices are classified into five different categories based on the layer in which they operate in a network.*
 1. Passive hub
 2. Repeaters
 3. Bridges
 4. Routers
 5. Gateways

Passive hub

- It is just a connector
- It connects the wires coming from different branches**
- In star topology, a passive hub is just a point where the signals coming from different stations collide
- Hub is the collision point
- Its location in internet model is below the physical layer

Repeaters(active hub)

- **operate at the physical layer**
- A repeater **connects segments of a LAN.**

Attenuation : strength of the signal becomes weak after travelling a certain distance

- **Repeater receives a signal before it becomes weak or corrupted and regenerates a new signal and forwards every frame** (A repeater is a regenerator, not an amplifier.)
- Active hub is a **multiport repeater**



ege, Pudukad

A repeater connecting two segments of a LAN

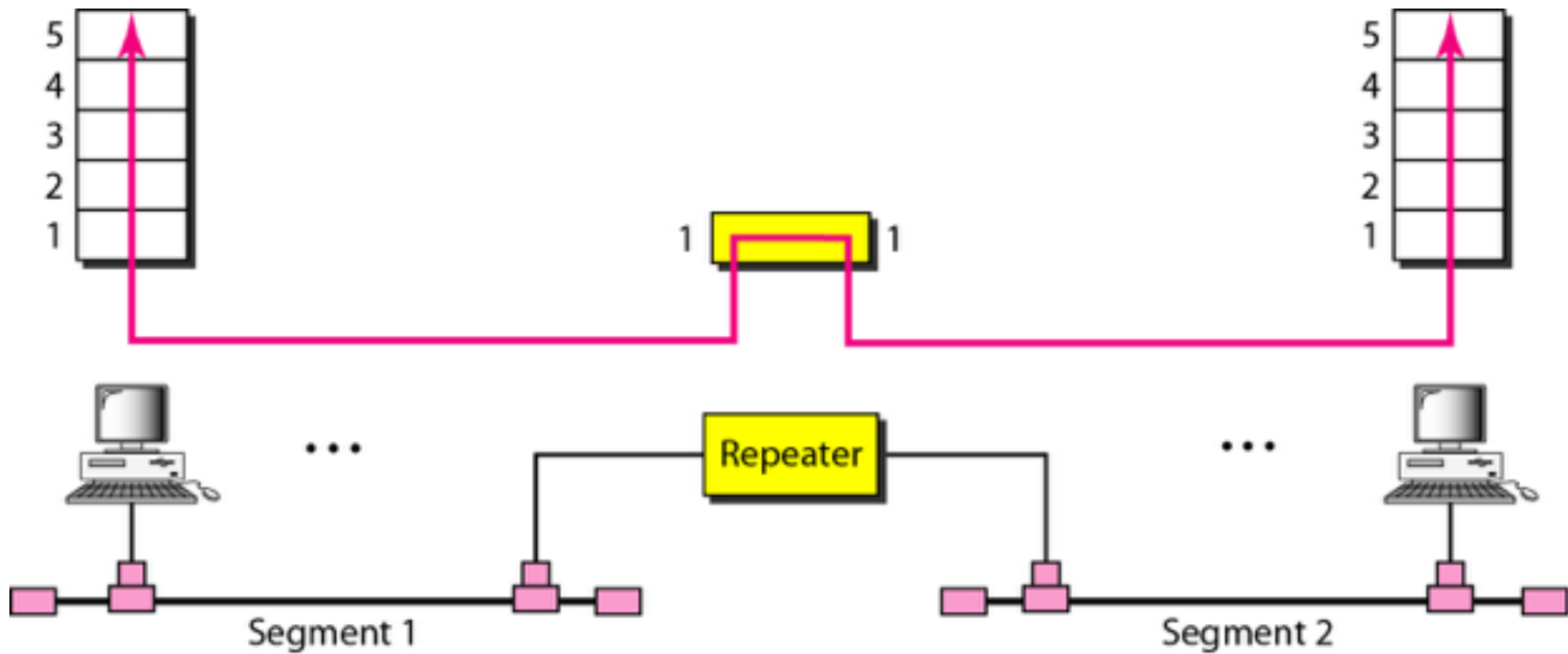
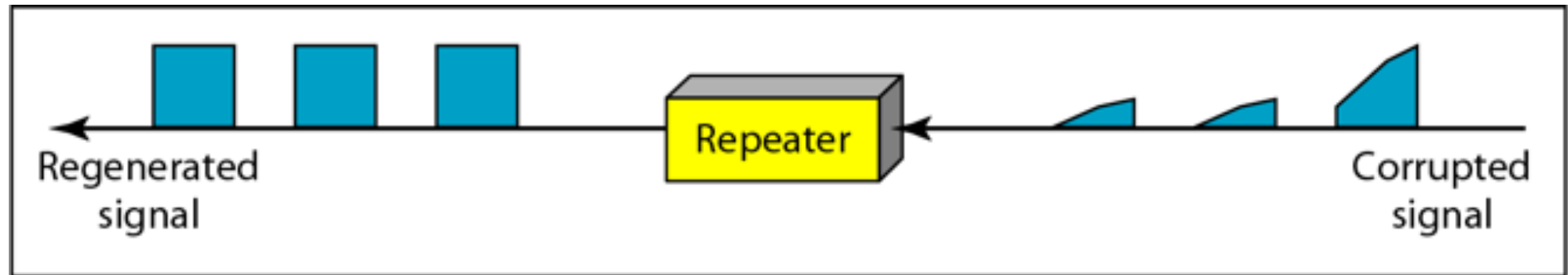
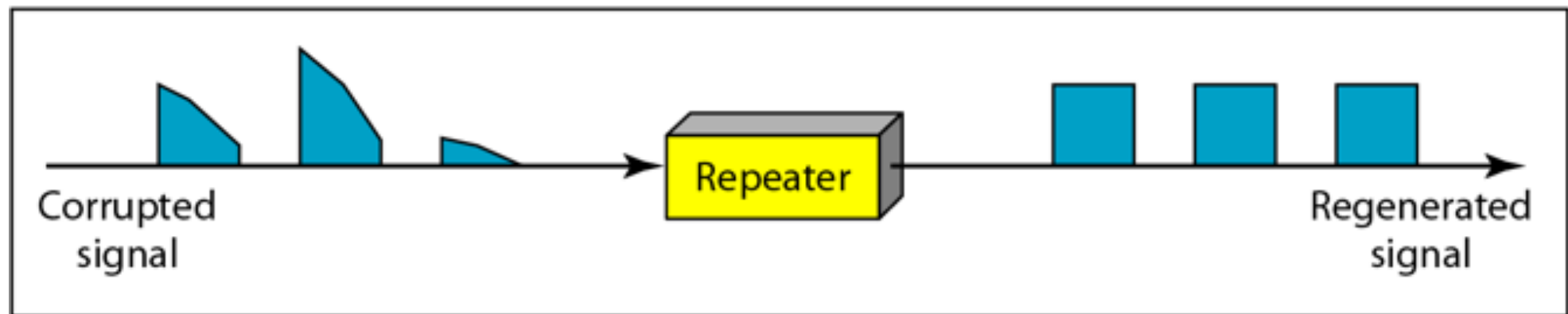


Figure 15.3 *Function of a repeater*



a. Right-to-left transmission.

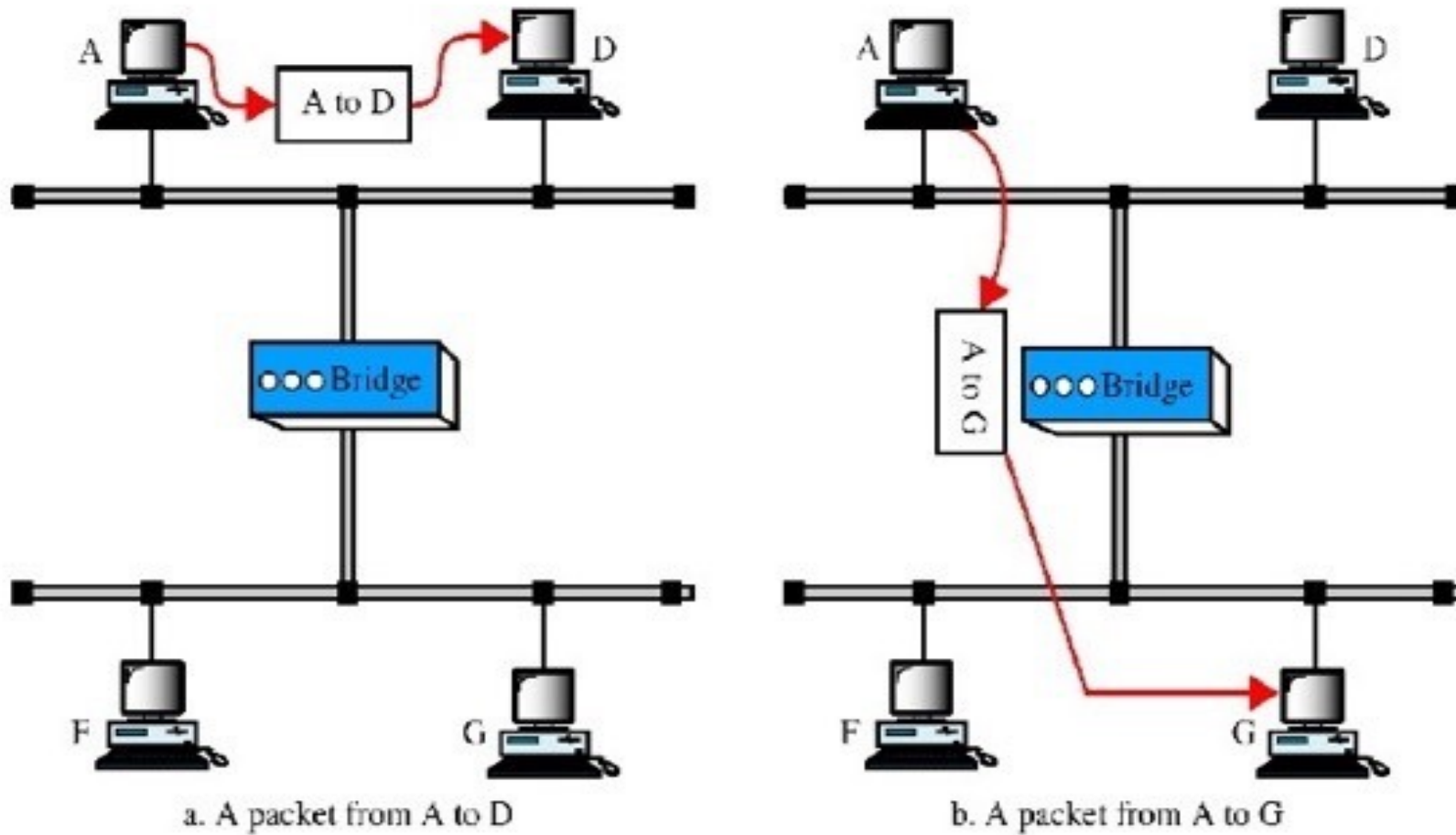


b. Left-to-right transmission.

BRIDGE

- device that operates at physical and data link layer
- it regenerates the signal it receives.
- As a data link layer device, the bridge can check physical (MAC) addresses (source and destination) and decide if frame should be **forwarded or discarded** {filtering capability}
- [repeater has no filtering capability]
- It can check the destination address of a frame and decide if the frame should be forwarded or dropped. If the frame is to be forwarded, the decision must specify the port. A bridge has a table that maps addresses to ports.

Function of Bridge



2 types of bridges

1. Transparent bridge
2. Source routing bridge

1. Transparent bridges

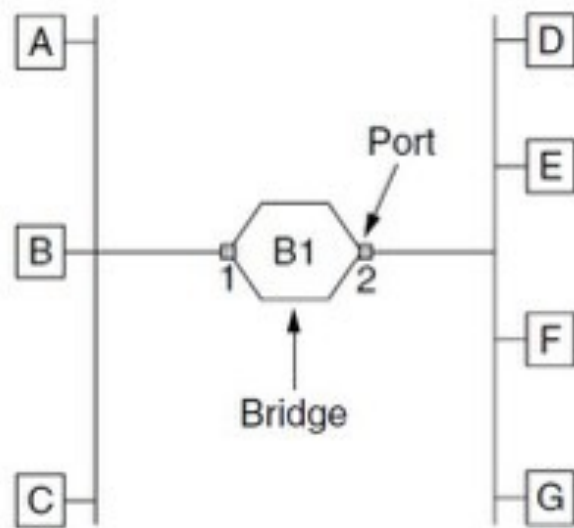
a system equipped with transparent bridges must meet three criteria:

1. Frames must be forwarded from one station to another.
 2. The forwarding table is automatically made by learning frame movements in the network.
 3. Loops in the system must be prevented.
- Two algorithms are used
 - ✓ backward learning algorithm
 - ✓ spanning tree algorithm

BACKWARD LEARNING ALGORITHM

- Each bridge operates in **promiscuous mode**, that is, it accepts every frame transmitted by the stations attached to each of its ports.
- The bridge must decide whether to **forward or discard each frame**
- **Forward** , on which port to output the frame. This decision is made by using the destination address.

Learning Bridges (1)



Bridge connecting two multidrop LANs

- There is a big (hash) table inside the bridge.
- This table list each possible destination and which output port it belongs on.
- When the bridges are first plugged in, all the hash tables are empty.

The **learning algorithm procedure** is as follows.

1. If the port for the destination address is the **same** as the source port, discard the frame.
2. If the port for the destination address and the source port **are different**, forward the frame on to the destination port.
3. If the destination port is unknown, use **flooding algorithm** and send the frame on all ports except the source port.

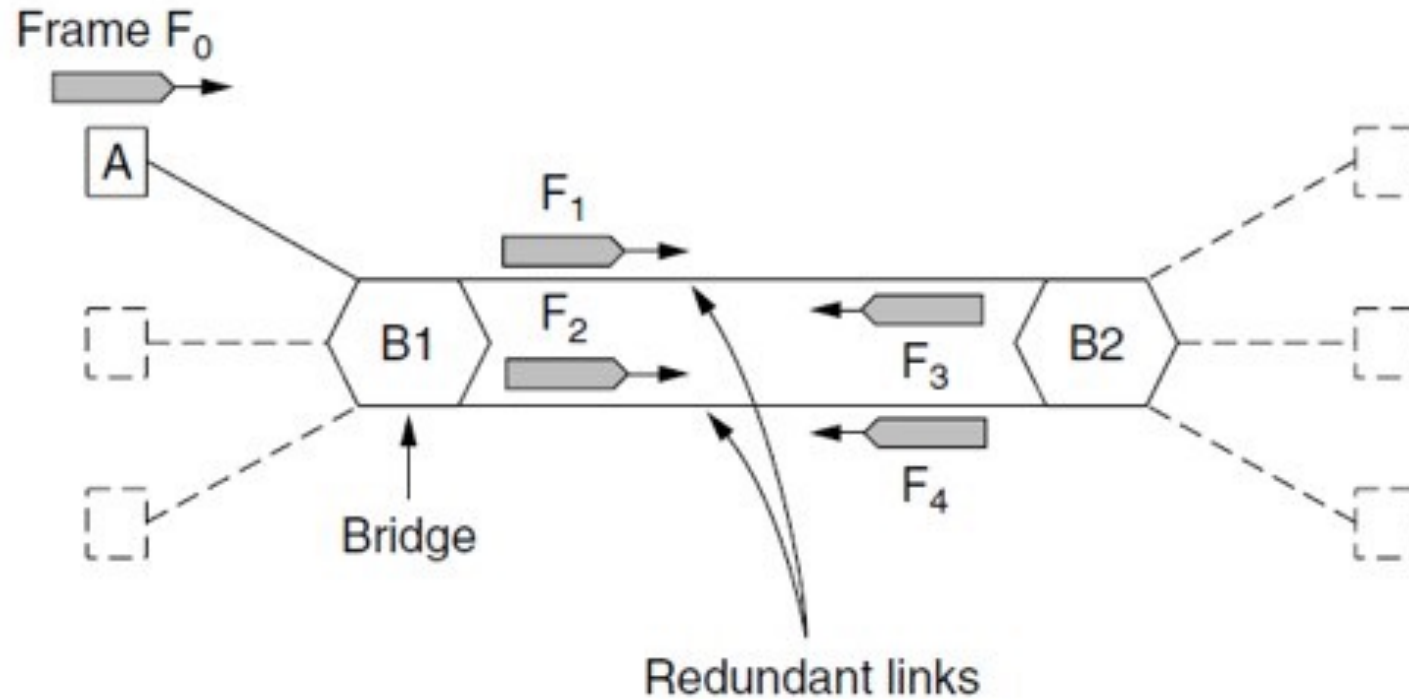
learning algorithm procedure contd..

- As time goes on, the bridges learn where destinations are.
- Once a destination is known, frames destined for it are put only on the proper port; they are not flooded

Spanning tree algorithm

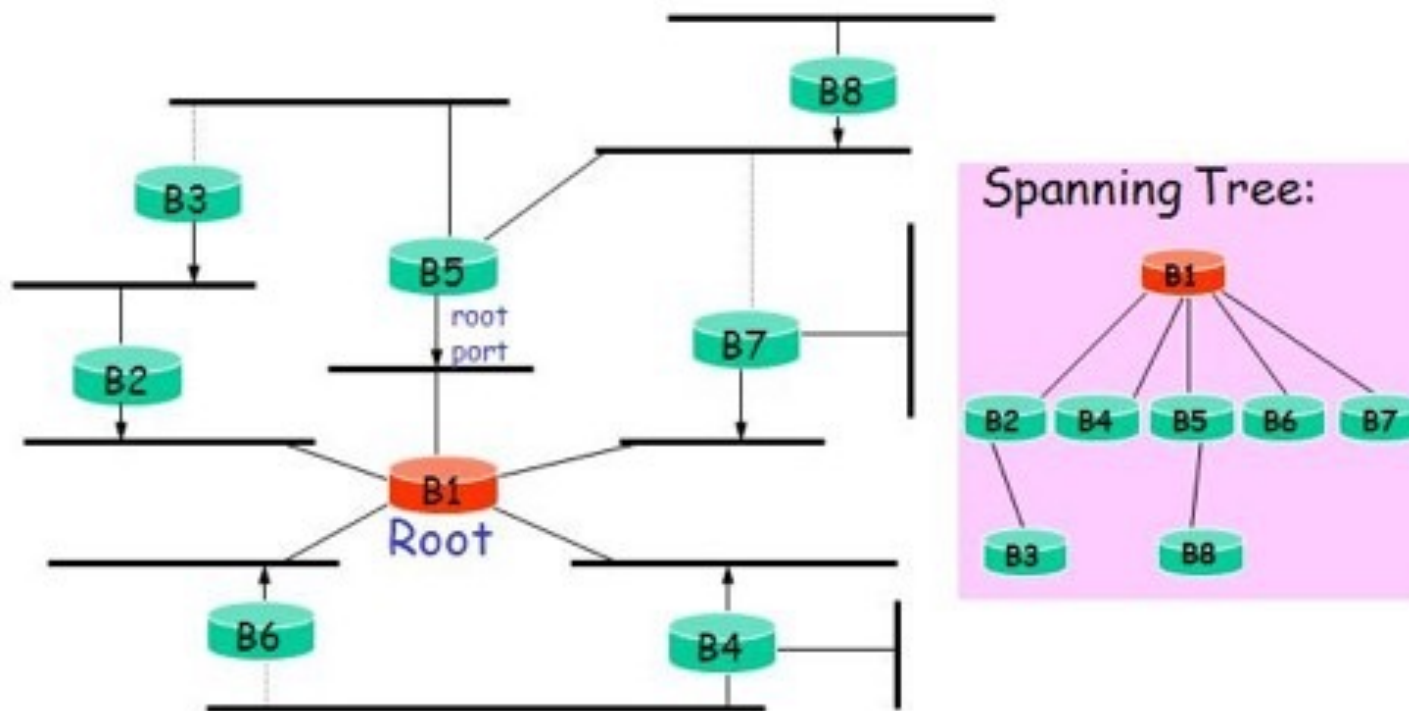
- To increase reliability, **redundant links can be used between bridges**
- this redundancy **introduces some additional problems because it creates loops in the topology.**
- To avoid loops, spanning tree algorithm is used
- **Spanning tree is a tree in which there is no loop.**
- To build the spanning tree, the bridges run a **distributed algorithm**(uses BPDUs to update the spanning tree on failure of a bridge or on addition or deletion of bridges)

Bridges with two parallel links



- Creating a topology in which each LAN can be reached from any other LAN thru **one path only** (no loop)
- Physical topology is not changed, creates a **logical topology**

Example Spanning Tree



To find the spanning tree, assign a cost (metric) to each arc. Metric may be the path with minimum hops (nodes), the path with minimum delay, or the path with maximum bandwidth.

Based on the spanning tree, we mark the ports that are part of the spanning tree, the forwarding ports, which forward a frame that the bridge receives

2. Source routing bridges

- Prevent loops in a system with redundant bridges
- Functions
 - Filtering
 - Forwarding(logical links that are part of spanning tree are marked as forwarding ports)
 - blocking (ports that are not of the spanning tree are marked as blocking ports)
- Sending station contains source address and destination address and the address of the bridges it should visit
- Sending station gets the address of the bridges using a discovery frame(sent before data frame)
- Used in IEEE token ring LAN

Router

- Operates at the network layer
- connects LANs and WANs in the Internet
- has a routing table that is used for making decisions about the route.
- The routing tables are normally dynamic and are updated using routing protocol

Gateway

- gateway is normally a computer that operates in all five layers of the Internet or seven layers of OSI model.
- A gateway takes an application message, reads it, and interprets it.
- This means that it can be used as a connecting device between two internetworks that use different models.

Network layer

Logical addressing

- Global addressing scheme called logical addressing
IP address (logical address) are unique and universal

Available in 2 versions

- IPv4
 - 32 bits in length.
 - address space is 2^N where N –no of bits in the address ie 2^{32} (4 billion)
- IPv6
 - uses 128-bit addresses

Two Notations :there are two notations to show an IPv4 address:

- **binary notation**

 - 32 bits are divided into 4 octets each of 8 bits

 - Each octet is often referred to as a byte.

 - Eg.01110101 10010101 00011101 00000010

- **Dotted-Decimal Notation**

 - written in decimal form with a decimal point (dot) separating the bytes

 -

 - each number in dotted-decimal notation is a value ranging from 0 to 255.

eg. 117.149.29.2

IPv4 addressing, at its inception, used the concept of classes.

Classful Addressing

- the address space here is divided into five classes: A, B, C, D, and E.

	First byte	Second byte	Third byte	Fourth byte
Class A	0			
Class B	10			
Class C	110			
Class D	1110			
Class E	1111			

a. Binary notation

	First byte	Second byte	Third byte	Fourth byte
Class A	0–127			
Class B	128–191			
Class C	192–223			
Class D	224–239			
Class E	240–255			

b. Dotted-decimal notation

- Class A addresses were designed for large organizations
- Class B addresses were designed for midsize organizations
- Class C addresses were designed for small organizations
- Class D addresses were designed for multicasting
- the class E addresses were reserved for future use

- In classful addressing, a large part of the available addresses were wasted.
- Each class is divided into fixed number of blocks each block of fixed size

<i>Class</i>	<i>Number of Blocks</i>	<i>Block Size</i>	<i>Application</i>
A	128	16,777,216	Unicast
B	16,384	65,536	Unicast
C	2,097,152	256	Unicast
D	1	268,435,456	Multicast
E	1	268,435,456	Reserved

Block size indicates number of addresses

Ceena Mathews, Prajyoti Niketan College, Pudukad

- In classful addressing, an IP address in class A, B, or C is divided into **netid and hostid**.
 - In *class A*, one byte defines the netid and three bytes define the hostid.
 - In *class B*, two bytes define the netid and two bytes define the hostid.
 - In *class C*, three bytes define the netid and one byte defines the hostid.

Netid and hostid are of varying lengths, depending on the class of the address

- Although, length of the netid and hostid (in bits) is predetermined in classful addressing, we can also use a mask (also called the default mask)
- **mask** can help us to **find the netid and the hostid**.
 - a **mask is a 32-bit number in which the n leftmost bits are 1s and the 32 - n rightmost bits are 0s**
 - For example, *the mask for a class A address has eight 1s, which means the first 8 bits of any address in class A define the **netid**; the next 24 bits define the **hostid**.*

<i>Class</i>	<i>Binary</i>	<i>Dotted-Decimal</i>	<i>CIDR</i>
A	11111111 00000000 00000000 00000000	255.0.0.0	/8
B	11111111 11111111 00000000 00000000	255.255.0.0	/16
C	11111111 11111111 11111111 00000000	255.255.255.0	/24

- **mask is in the form /n** where n can be 8, 16, or 24 in classful addressing.
- This notation is also called *slash notation* or *Classless Interdomain Routing (CIDR) notation*

Subnet

- If an organization was granted a **large block in class A or B**, it could **divide the addresses into several contiguous groups** and assign each group to smaller networks (called **subnets**) or, in rare cases, share part of the addresses with neighbors.

Supernetting

- when most of the **class A and class B** addresses were depleted, there was still a huge demand for midsize blocks.
- The size of a **class C** block with a maximum number of **256** addresses did not satisfy the needs of most organizations
- an organization can **combine several class C blocks** to create a larger range of addresses

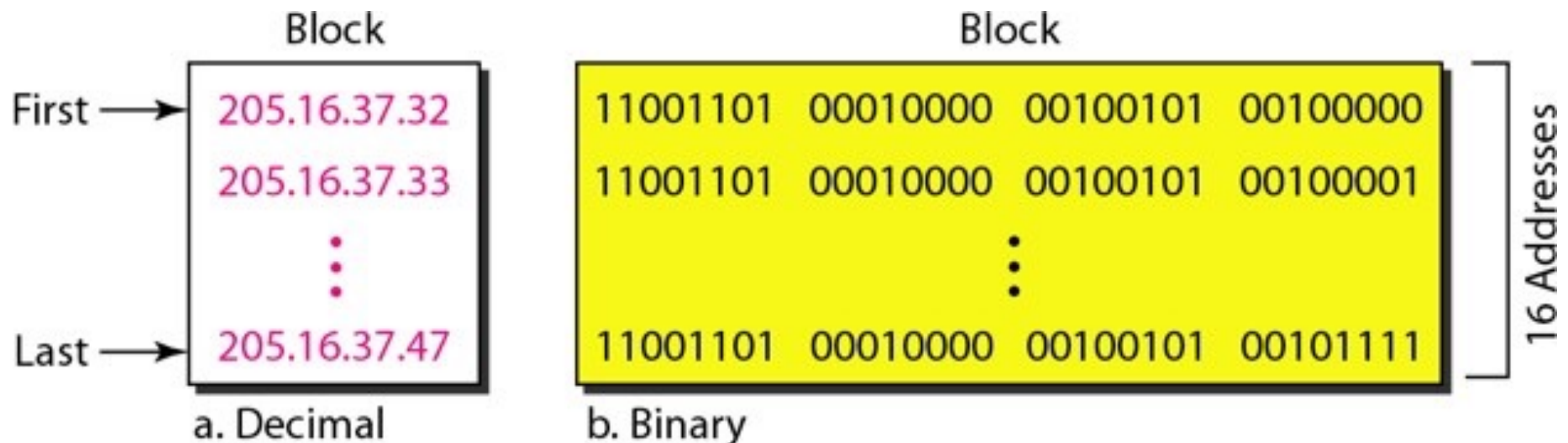
Ceena Mathews, Prajyoti Niketan College, Pudukad

Classless Addressing

- flaws in classful addressing scheme combined with the fast growth of the Internet led to the **near depletion of the available addresses.**
- the above problem of address depletion is **reduced by Classless Addressing**
- No classes but **organized as blocks**
- **The size of the block based on the nature and size of the entity**

three restrictions on classless address blocks:

1. The addresses in a block **must be contiguous**, one after another.
2. The number of addresses in a block **must be a power of 2** (1, 2, 4, 8, ...).
3. The **first address must be evenly divisible** by the number of addresses



- a mask is a **32-bit number in which the n leftmost bits are 1s and the 32 - n rightmost bits are 0s**

- a block of addresses can be defined as **x.y.z.t/n**
 - **x.y.z.t -> addresses**
 - **/n defines the mask.**

- The **first address** in the block can be found by setting the rightmost **32- n bits to 0s.**

- The **last address** in the block can be found by setting the rightmost **32- n bits to 1s.**

- The **number of addresses** in the block is the difference between the last and first address. It can easily be found using the formula 2^{32-n} .

example

- *A block of addresses is granted to a small organization. We know that one of the addresses is 205.16.37.39/28. What is the first address in the block?*

- ***Solution*** *The binary representation of the given address is 11001101 00010000 00100101 00100111.*

- *set 32 - 28 rightmost bits to 0, we get*

11001101 00010000 00100101 00010000 or 205.16.37.32.

Network address

- The first address in a block is normally not assigned to any device;
- It is **used as the network address** that represents the organization, to the rest of the world.

Hierarchy

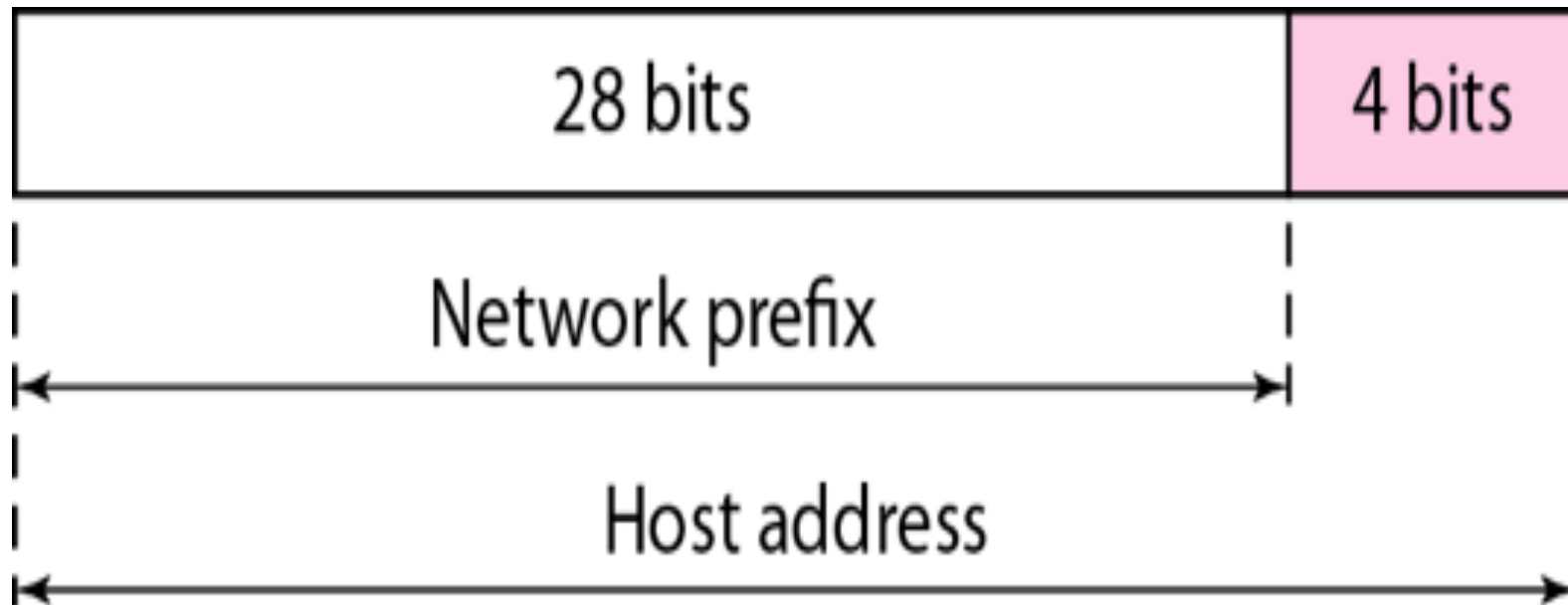
- IP addresses have levels of hierarchy.

2 level Hierarchy

- Each address in the block can be considered as a two-level hierarchical **structure without subnetting**
 - the leftmost n bits (prefix) define the network;
 - the rightmost $32 - n$ bits define the host(suffix)

Two level hierarchy

Ceena Mathews, Prajyoti Niketan College, Pudukad



3 level hierarchy

- An organization that is granted a large block of addresses may create clusters of networks (**called subnets**)
- divide the addresses between the different subnets
- **rest of the world sees the organization as one entity**; however, internally there are several subnets.
- **All messages are sent to the router address** that connects the organization to the rest of the Internet
- the **router then routes the message** to the appropriate subnets
- The **organization has its own mask**; each subnet must also have its own

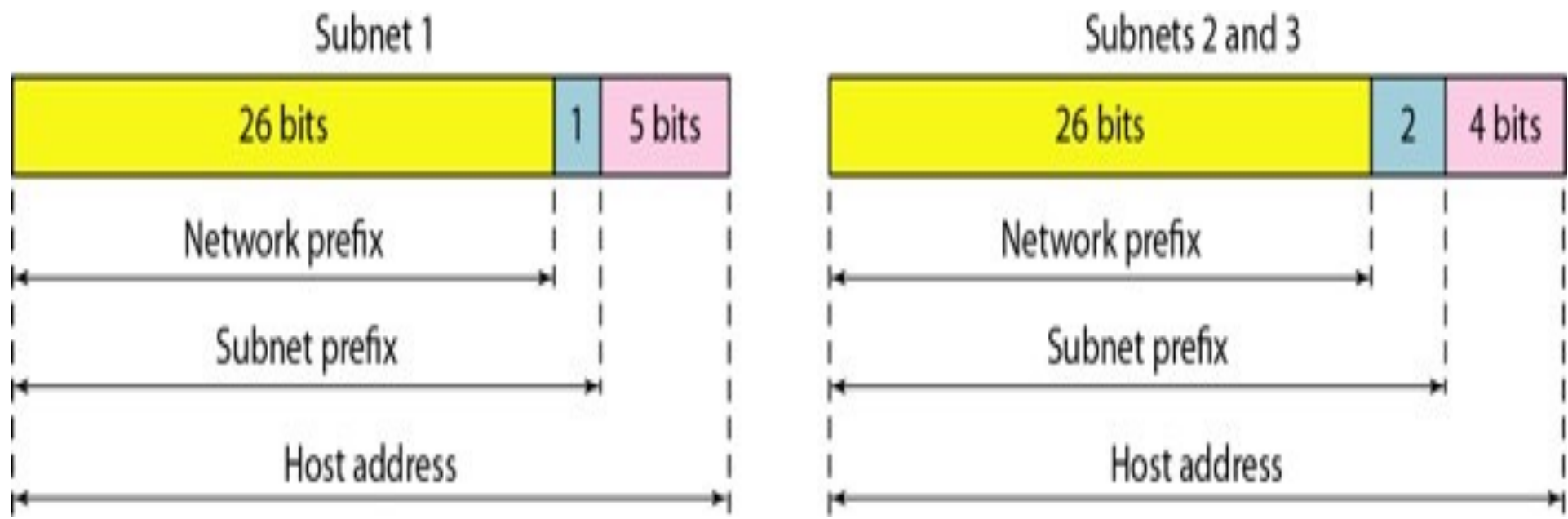
Ceena Mathews, Prajyoti Niketan College, Pudukad

- suppose an organization is given the block **17.12.40.0/26**, which contains **64 addresses**. The organization has **three offices** and needs to divide the addresses into **three subblocks of 32, 16, and 16** addresses.

- We can find the **new masks** by using the following arguments:
 1. Suppose the mask for the first subnet is n_1 , then 2^{32-n_1} must be 32, which means that $n_1 = 27$.
 2. Suppose the mask for the second subnet is n_2 , then 2^{32-n_2} must be 16, which means that $n_2 = 28$.
 3. Suppose the mask for the third subnet is n_3 , then 2^{32-n_3} must be 16, which means that $n_3 = 28$.

- This means that we have the masks 27, 28, 28 with the organization mask being 26

Subnet prefix length can differ for the subnets



More Levels of Hierarchy

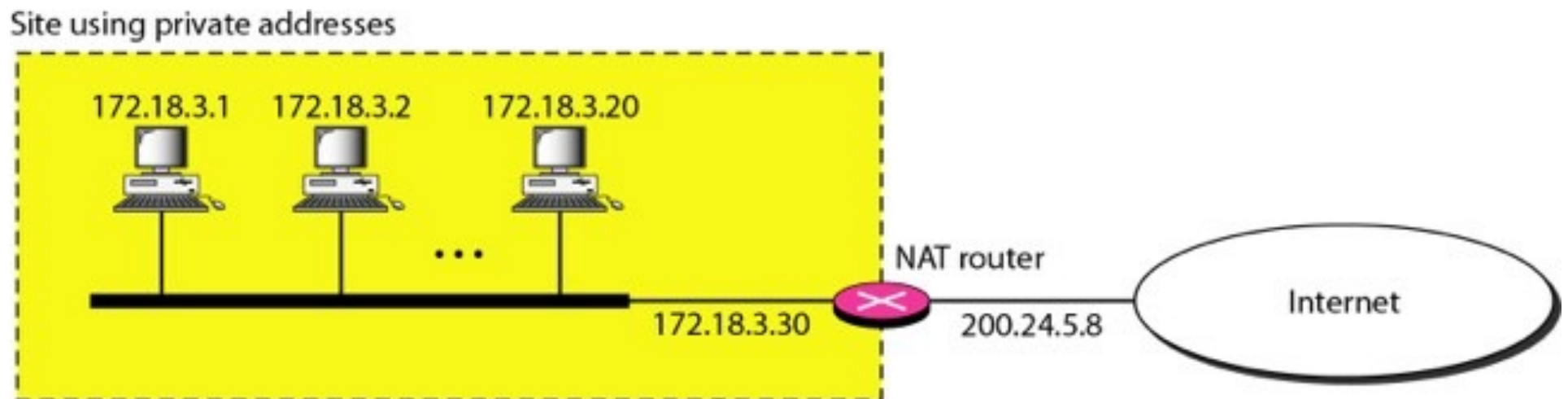
- The structure of classless addressing **does not restrict the number of hierarchical levels.**
- An organization can divide the granted block of addresses into subblocks.
- Each **subblock can in turn be divided into smaller subblocks.**
- Eg. A **national ISP** can divide a granted large block into smaller blocks and assign each of them to a **regional ISP**. A regional ISP can divide the block received from the national ISP into smaller blocks and assign each one to a **local ISP**. A **local ISP** can divide the block received from the regional ISP into smaller blocks and assign each one to a **different organization**. Finally, an organization can divide the received block and make **several subnets** out of it

Address allocation

- ultimate responsibility of address allocation is given to a global authority called the **Internet Corporation for Assigned Names and Addresses (ICANN)**
- It assigns a large block of addresses to an ISP.
- an ISP receives one large block to be distributed to its Internet users. This is called **address aggregation**

Network Address Translation (NAT).

- NAT enables a user to have a large set of addresses for internal traffic(communication) and **one address, or a small set of addresses for global communications**
- For external communication , it uses a router.(called NAT router)
- NAT router uses the NAT software



Ceena Mathews, Prajyoti Niketan College, Pudukad

private addresses

- To separate the addresses used inside the home or business and the ones used for the Internet, the **Internet authorities have reserved three sets of addresses as private addresses**
 - These are **for private networks**
 - These addresses **can be used by organisation without permission from internet authorities**

<i>Range</i>			<i>Total</i>
10.0.0.0	to	10.255.255.255	2^{24}
172.16.0.0	to	172.31.255.255	2^{20}
192.168.0.0	to	192.168.255.255	2^{16}

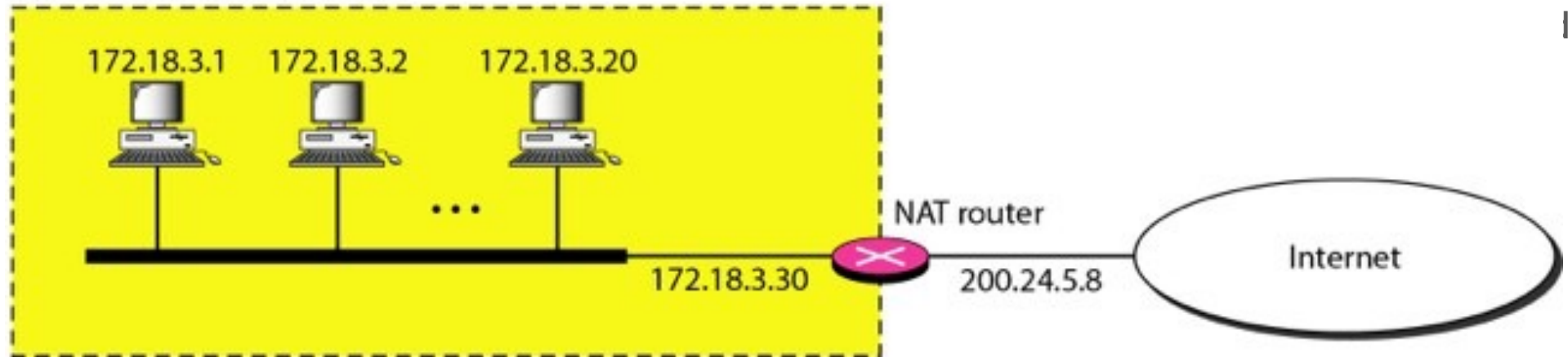
Private address

- They are **unique inside** the organization, but they are not unique globally.
 - No router will forward a packet that has one of these addresses as the destination address.
-
- **Router** that connects the network to the global address uses **one private address and one global address**
-
- the **source address** in the outgoing packet that go thru the NAT router will **be replaced** with the global **NAT address**.
 - incoming packets also pass through the NAT router, which **replaces the destination address** in the packet (the NAT router global address) with the appropriate **private address**

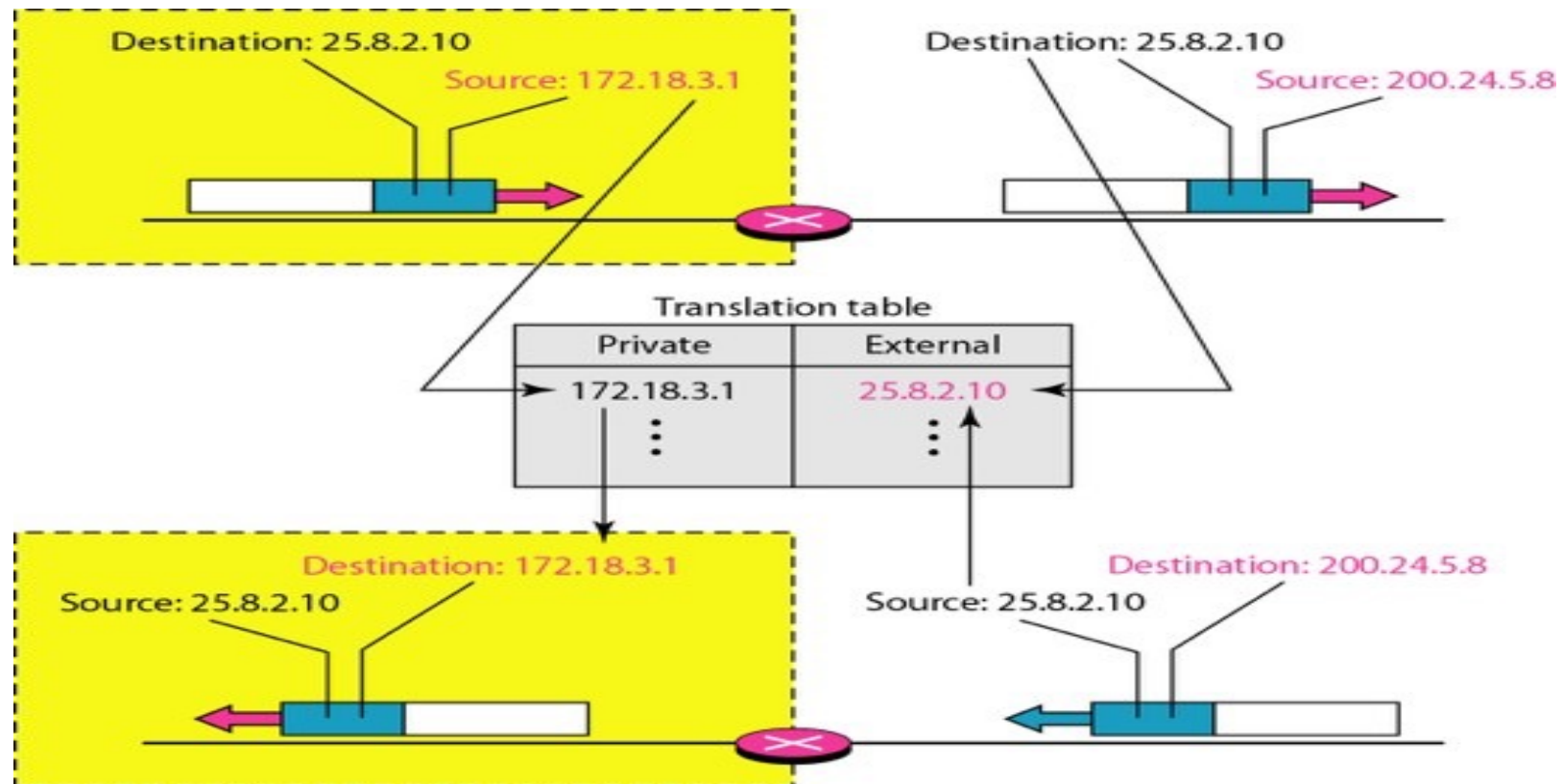
NAT contd..

- NAT router has a **translation table**.
- **In the case of using one IP address**
 - translation table has only **two columns**: the private address and the external address (destination address of the packet).
 - When the router translates the source address of the outgoing packet, it also makes note of the destination address-where the packet is going.

Site using private addresses



HOW NAT translation is done



- Since the **NAT router has only one global address**, only one private network host can access the same external host.
- **To remove this restriction**, the NAT router uses a pool of global addresses.
- For example, instead of using only one global address (200.24.5.8), the NAT router can use four addresses (200.24.5.8, 200.24.5.9, 200.24.5.10, and 200.24.5.11)
- using only IP addresses , **no more than 4 connections** can be made to the same destination and also no private network host can **access 2 external server programs**

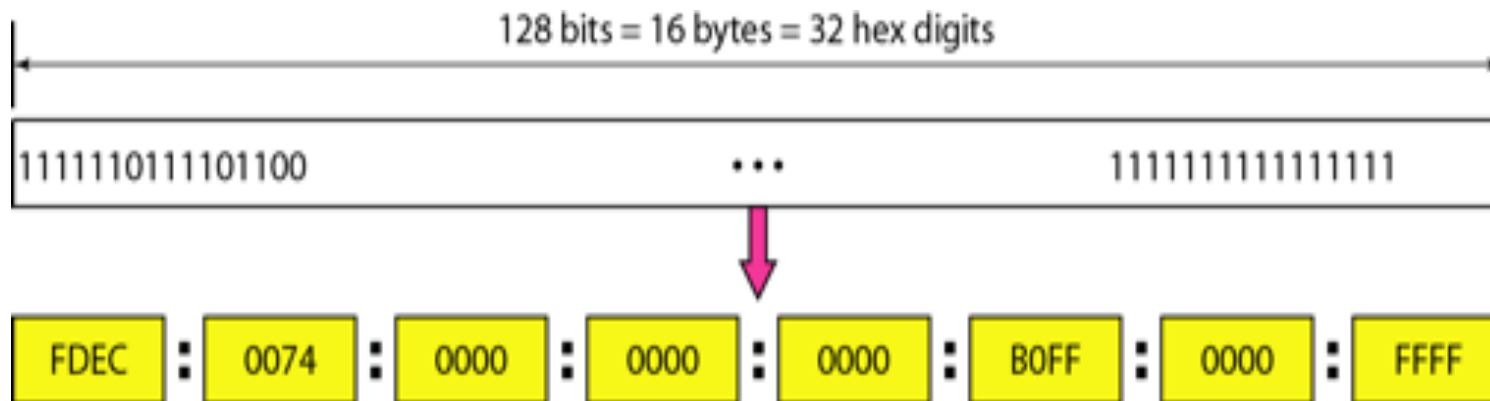
Using Both IP Addresses and Port Numbers

- To allow a **many-to-many relationship between private-network hosts and external server programs**, we need more information in the translation table.
- 5 column NAT

<i>Private Address</i>	<i>Private Port</i>	<i>External Address</i>	<i>External Port</i>	<i>Transport Protocol</i>
172.18.3.1	1400	25.8.3.2	80	TCP
172.18.3.2	1401	25.8.3.2	80	TCP
...

IPv6

- 128 bits long(16bytes)
- uses *hexadecimal colon notation*.
 - 128 bits is divided into **eight sections, each 2 bytes in length.**
 - **Two bytes** in hexadecimal notation requires **four hexadecimal digits.**
 - Therefore, the address consists of **32 hexadecimal digits**, with every four digits separated by a colon



ABBREVIATED form

Original

FDEC ■ 0074 ■ 0000 ■ 0000 ■ 0000 ■ BOFF ■ 0000 ■ FFF0



Abbreviated

FDEC ■ 74 ■ 0 ■ 0 ■ 0 ■ BOFF ■ 0 ■ FFF0



More abbreviated

FDEC ■ 74 ■ ■ BOFF ■ 0 ■ FFF0

Gap

- Expand the address 0:15::1:12:1213 to its original.

Address Space

- 2^{128} addresses are available.
-
- address is divided into several categories.
 - A few leftmost bits, called the **type prefix**, in each address define its **category. Or purpose**
 - The type prefix is variable in length.
 - there is no ambiguity; when an address is given, the type prefix can easily be determined
 - eg 11111111 Multicast addresses
 - 010 Provider-based unicast addresses

Table 19.5 *Type prefixes for IPv6 addresses*

<i>Type Prefix</i>	<i>Type</i>	<i>Fraction</i>
0000 0000	Reserved	1/256
0000 0001	Unassigned	1/256
0000 001	ISO network addresses	1/128
0000 010	IPX (Novell) network addresses	1/128
0000 011	Unassigned	1/128
0000 1	Unassigned	1/32
0001	Reserved	1/16
001	Reserved	1/8
010	Provider-based unicast addresses	1/8

Unicast address

- Defines a single computer

- Pkt sent to a unicast address must be delivered to that computer

- 2 types of unicast address
 1. Geographic based (future definition)
 2. Provider based (used by a normal host)

Figure 19.16 *Prefixes for provider-based unicast address*

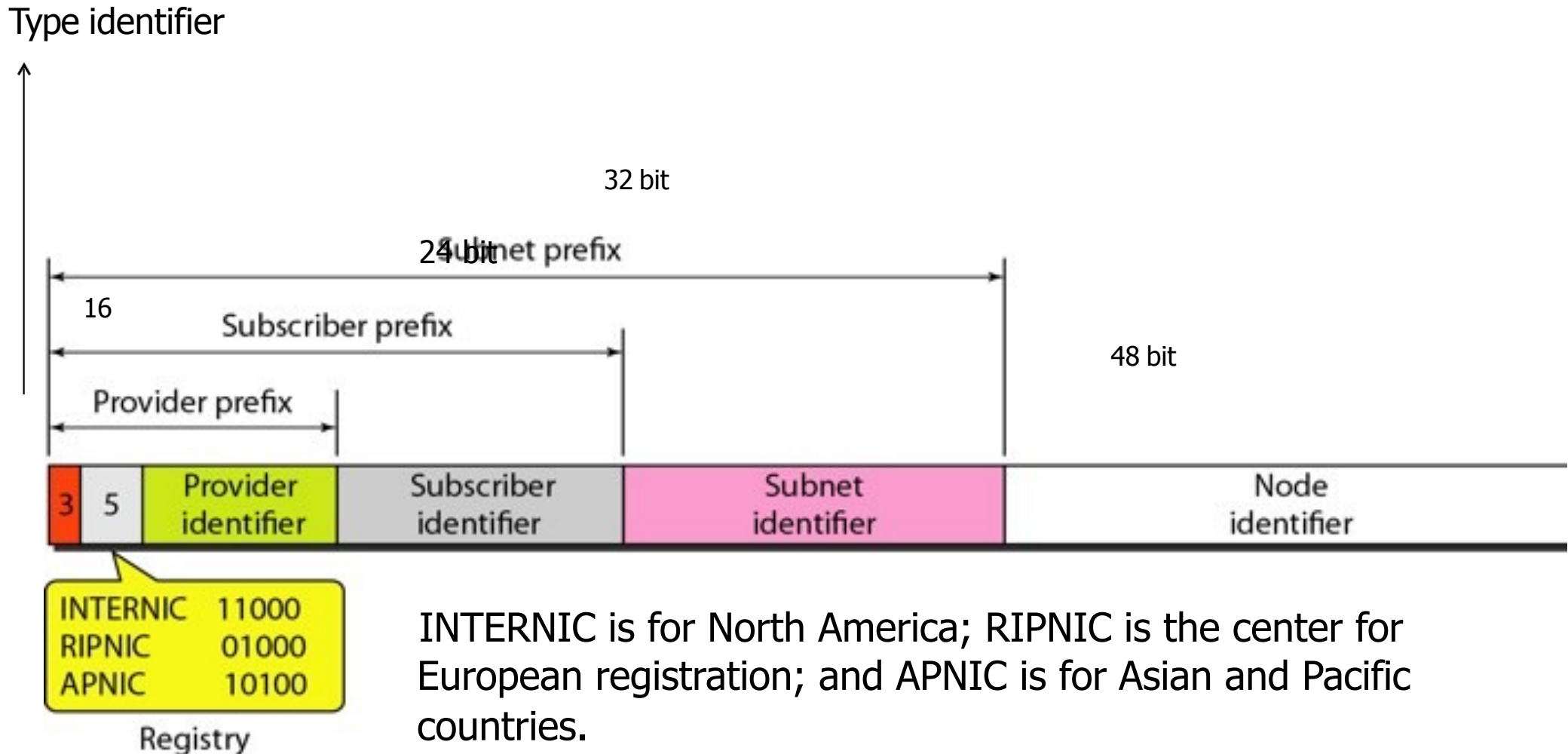
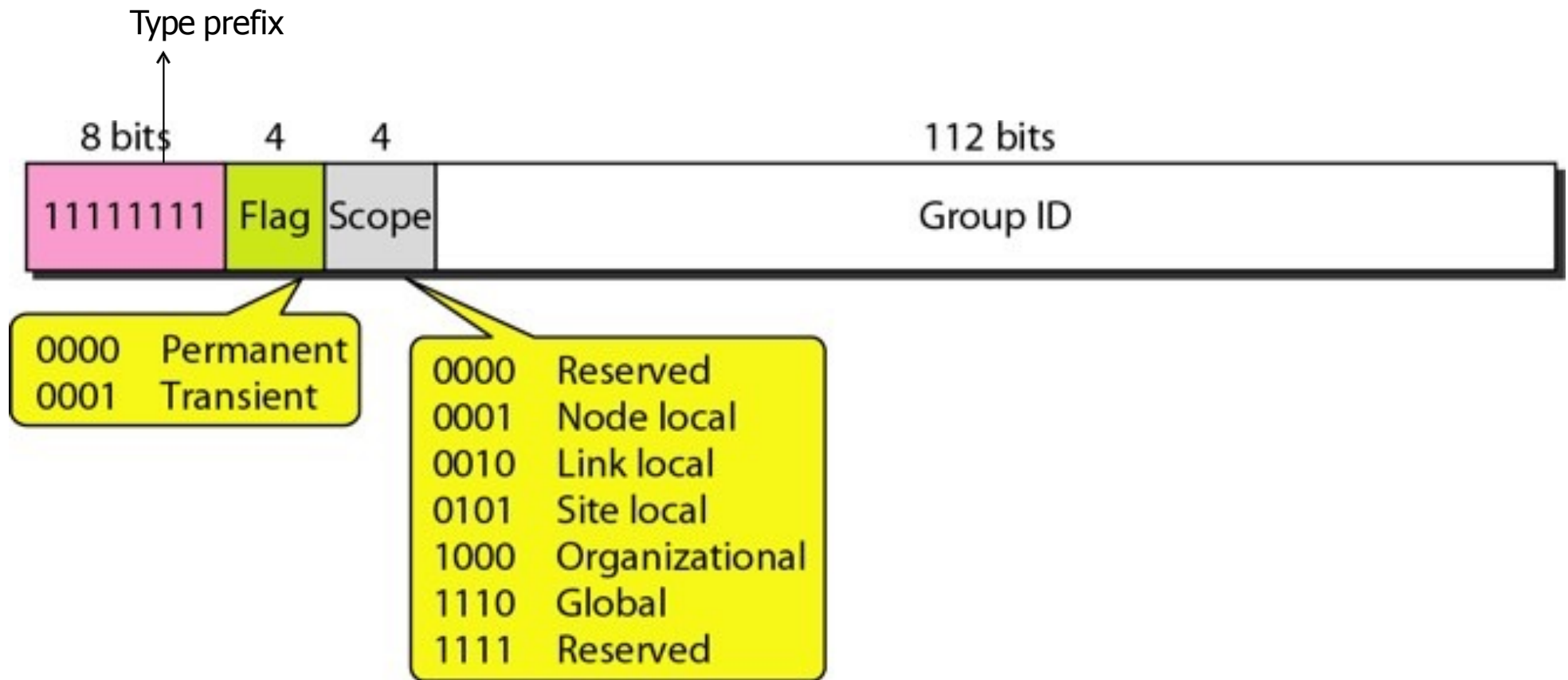


Figure 19.17 *Multicast address in IPv6*



Anycast addresses

- *like a multicast* address, also defines a group of nodes.
- a packet destined for an anycast address is delivered to only one of the members of the anycast group, the nearest one (the one with the shortest route)

Reserved addresses

Subcategories

- **Unspecified address**: when a **host does not know its own address, it sends an inquiry to find its address**
- **Loopback address** : used by a **host to test itself without going into network**
- **Compatible address**: It is used when a computer using IPv6 wants to send a message to another computer using IPv6, but the message needs to pass through a part of the network that still operates in IPv4.
- **Mapped address**: is used when a computer that has migrated to IPv6 wants to send a packet to a computer still using IPv4.

Figure 19.19 *Local addresses in IPv6*

Used when an organisation wants to use IPv6 protocol without being connected to global address



- Link local used in isolated subnet
- Site local used in isolated site with several subnets

Chapter 20

Network Layer: Internet Protocol

Internet Protocol

- In internet model , main network protocol is internet protocol
-
- *connecting networks together make an internetwork or an internet.*

Network layer/Internetwork layer

- *The physical and data link layers operate locally*
- *Network layer was designed to solve the problem of delivery thru several links*
- *N/w layer is responsible for*
 - *Create packets containing Logical address*
 - *Routing*
 - *Fragmentation & reassembly*
 - *Address verification*

- *Internet as a Datagram Network*
 - *The Internet is a packet-switched network.*
 - *Uses datagram approach (uses universal addresses to route packets from source to destination)*
- the **current version** of the Internet Protocol, version 4, or IPv4.
- the next generation of this protocol, IPv6,

IPv4

- IPv4 is an **unreliable and connectionless datagram protocol**-a best-effort delivery service.
- The term **best-effort** means that IPv4 provides **no error control or flow control** (except for error detection on the header)
- **each datagram is handled independently**, ie each datagram can follow a different route to the destination.
- datagrams sent by the same source to the same destination could arrive out of order.
- Also, some could be **lost or corrupted** during transmission.

IPv4 Datagram format

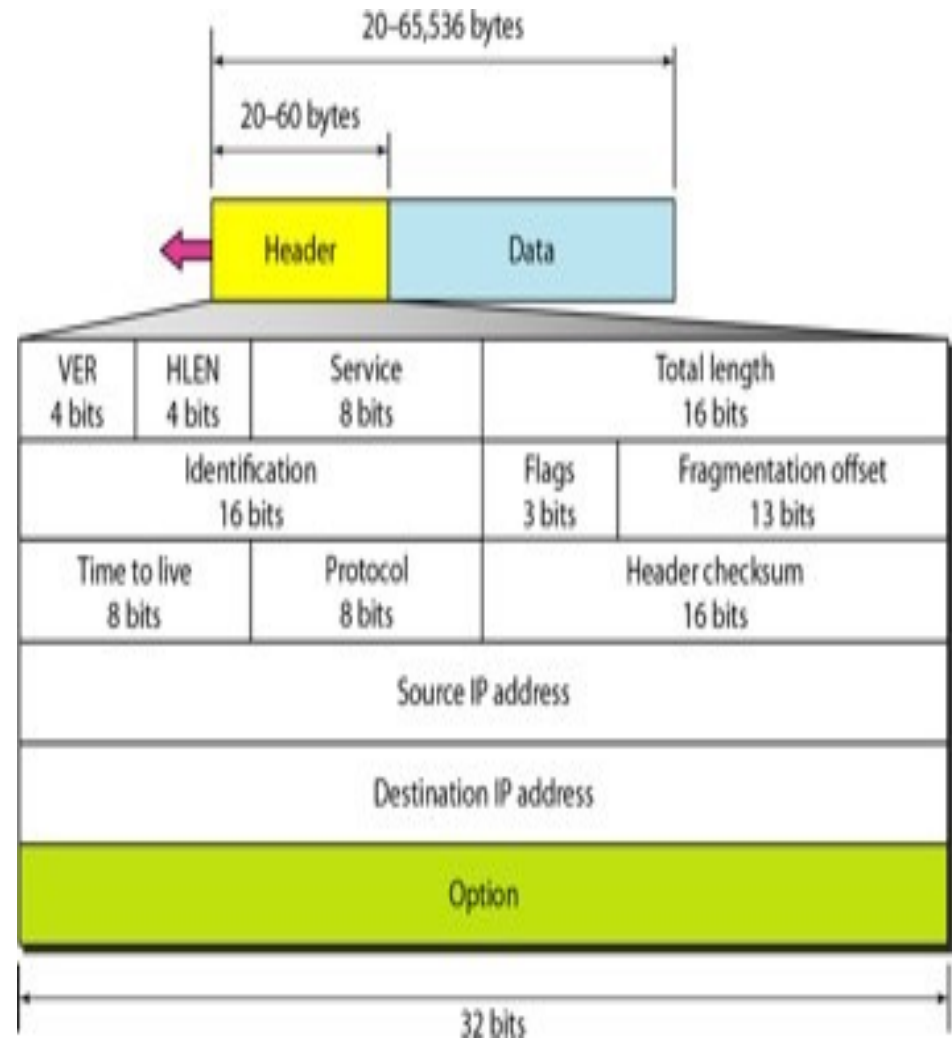
Datagram

Packets in the IPv4 layer are called datagrams

- A datagram is a variable-length packet consisting of two parts:
 - Header
 - Data

Header

- 20 to 60 bytes in length
- contains information essential to routing and delivery



1. Version (VER)

version of the IPv4 protocol. the version is 4.

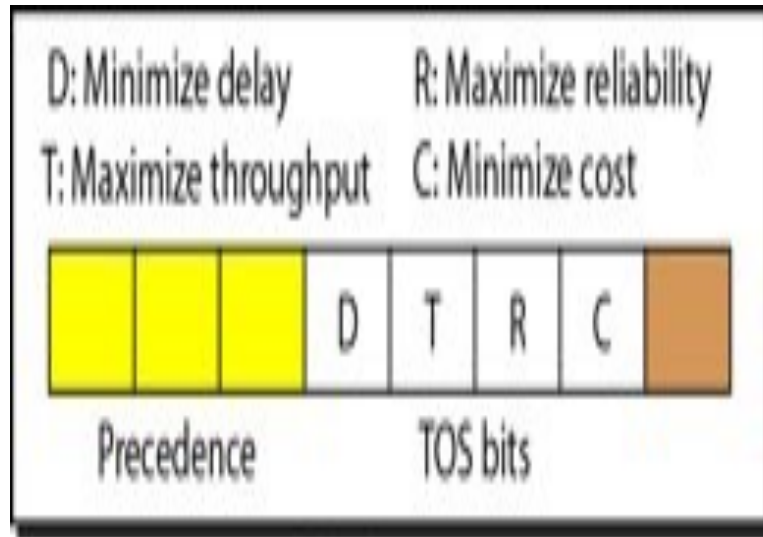
2. Header length (HLEN).

- the total length of the datagram header in 4-byte words.
- This field is needed because the length of the header is variable (between 20 and 60 bytes).
- The **header length** can be found by multiplying the value in the HLEN field by 4.
- When there are **no options**, the header length is **20 bytes**, and the value of this field is **5** ($5 \times 4 = 20$).
- When the **option field is at its maximum size**, the value of this field is **15** ($15 \times 4 = 60$).

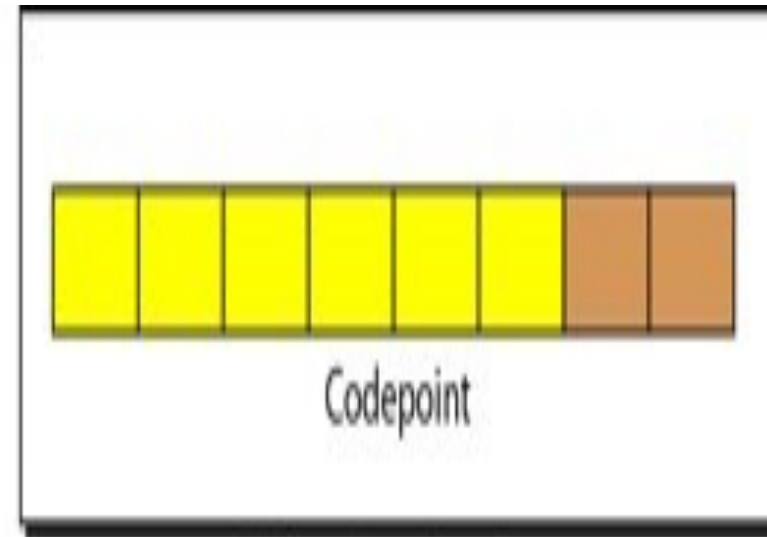
3. Services.

previously called service type, is now called differentiated services.

Service type or differentiated services



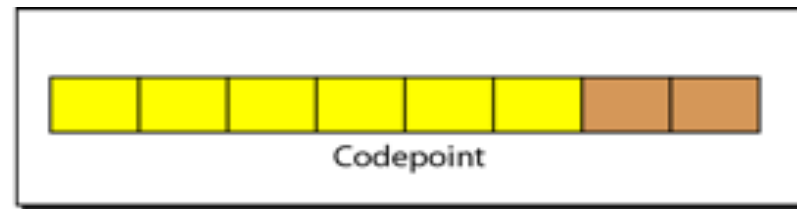
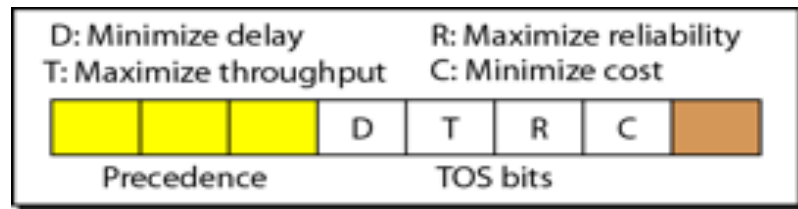
Service type



Differentiated services

Service Type

- the first 3 bits are called **precedence bits**.
- The next 4 bits are called **type of service (TOS) bits**
- last bit is not used.
- Precedence is a 3-bit subfield ranging from 0 (000 in binary) to 7 (111 in binary).



– The precedence defines the priority of the datagram in issues such as congestion.

- TOS bits is a 4-bit subfield with each bit having a special meaning.
- a bit can be either 0 or 1, one and only one of the bits can have the value of 1 in each datagram.
- With only 1 bit set at a time, we can have five different types of services

Types of service

<i>TOS Bits</i>	<i>Description</i>
0000	Normal (default)
0001	Minimize cost
0010	Maximize reliability
0100	Maximize throughput
1000	Minimize delay

Each Application programs can request a specific type of service based on service it needs.

<i>Protocol</i>	<i>TOS Bits</i>	<i>Description</i>
ICMP	0000	Normal
BOOTP	0000	Normal
NNTP	0001	Minimize cost
IGP	0010	Maximize reliability
SNMP	0010	Maximize reliability
TELNET	1000	Minimize delay
FTP (data)	0100	Maximize throughput
FTP (control)	1000	Minimize delay
TFTP	1000	Minimize delay
SMTP (command)	1000	Minimize delay
SMTP (data)	0100	Maximize throughput
DNS (UDP query)	1000	Minimize delay
DNS (TCP query)	0000	Normal
DNS (zone)	0100	Maximize throughput

- Activities requiring **immediate attention**, and activities requiring **immediate response** need **minimum delay**

- activities that send **bulk data** require **maximum throughput**

Differentiated Services

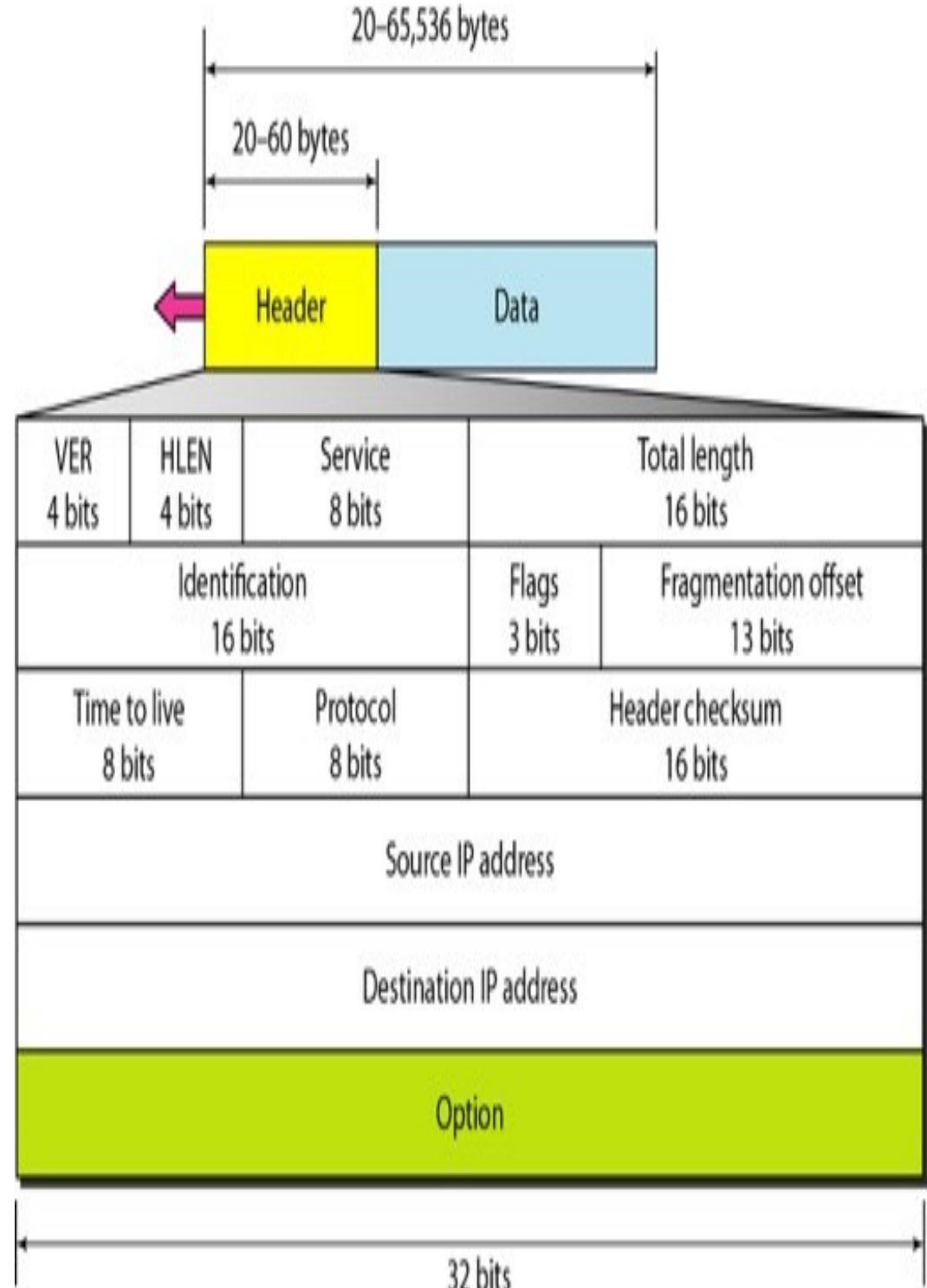
- the first 6 bits make up the code point subfield,
- the last 2 bits are not used.

codepoint subfield

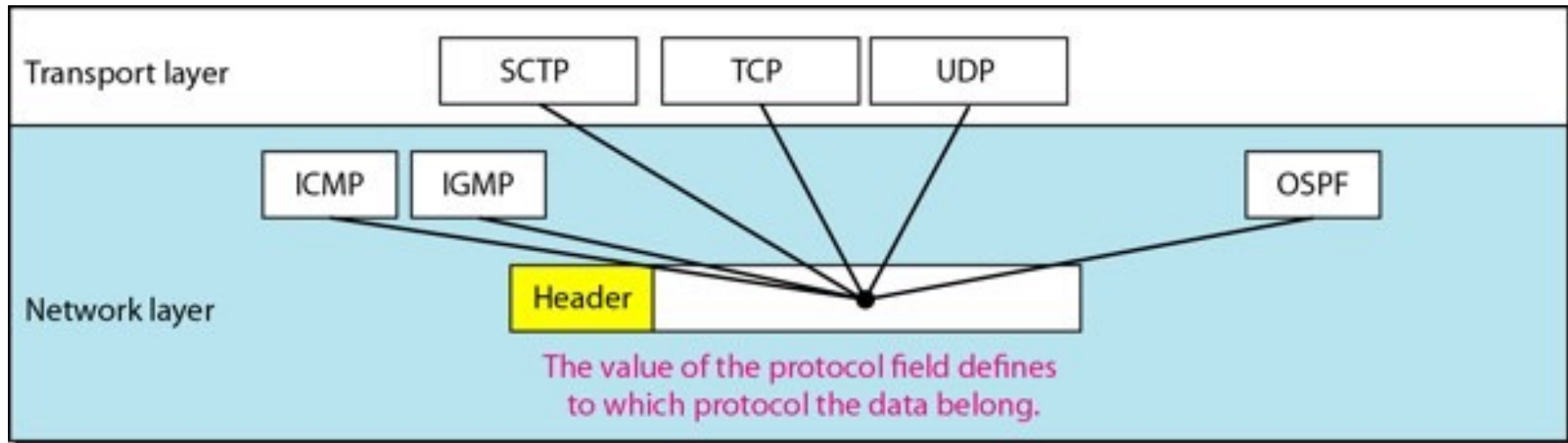
- When the 3 rightmost bits are 0s, the 3 leftmost bits are interpreted the same as the precedence bits in the service type interpretation.
- When the 3 rightmost bits are not all 0s, the 6 bits define 64 services based on the priority assignment by the Internet or local authorities

Total length.

- This is a **16-bit field** that **defines the total length** (header plus data).
- The **header length** can be found by multiplying the value in the HLEN field by 4.
- **Length of data**=total length - header length
- Since the field length is 16 bits, the **total length of the IPv4 datagram is limited to 65,535** ($2^{16} - 1$) bytes, of which 20 to 60 bytes are the header and the rest is data from the upper layer.



- **Identification.**
 - This field is used in fragmentation
- **Flags.**
 - This field is used in fragmentation
- **Fragmentation offset.**
 - This field is used in fragmentation
- **Time to live.**
 - A datagram has a **limited lifetime in its travel through an internet.**
 - This field was originally designed to hold a **timestamp**, which was **decremented by each visited router.**
 - it is used to **control the maximum number hops** visited by the datagram
 - Another use of this field is to intentionally **limit the journey of the packet.**
 - For example, if the source wants to confine the packet to the local network, it can store 1 in this field.
 - When the packet arrives at the first router, this value is decremented to 0, and the datagram is discarded



- **Protocol.**
 - This 8-bit field defines the higher-level protocol that uses the services of the IPv4 layer.
 - An IPv4 datagram can encapsulate data from several higher-level protocols such as TCP, UDP, ICMP, and IGMP.
 - This field specifies the final destination protocol to which the IPv4 datagram is delivered.

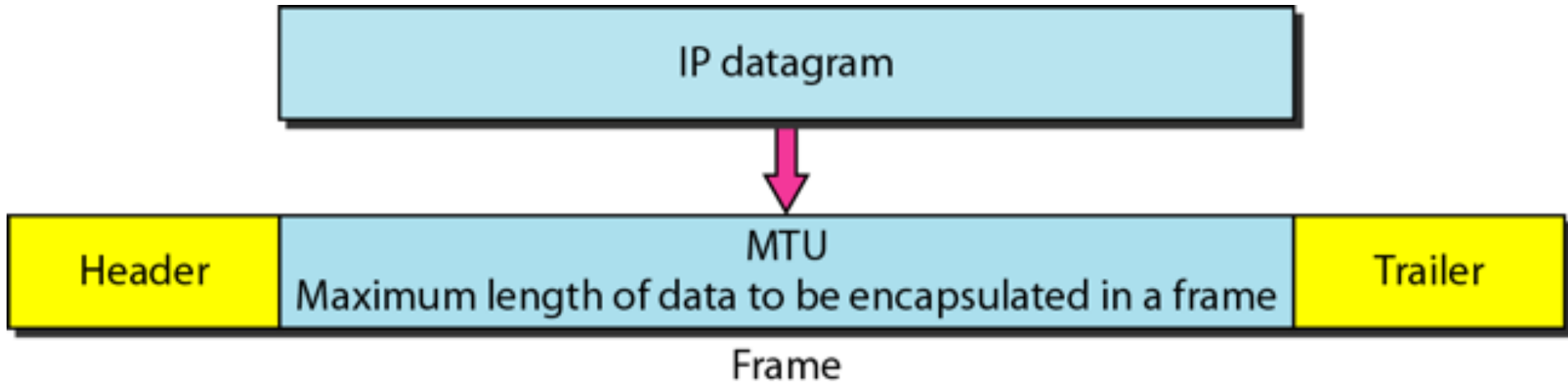
Protocol values

<i>Value</i>	<i>Protocol</i>
1	ICMP
2	IGMP
6	TCP
17	UDP
89	OSPF

- **Checksum.**
 - The checksum concept and its calculation
- **Source address.**
 - This 32-bit field defines the IPv4 address of the source.
 - This field must remain unchanged during the time the IPv4 datagram travels from the source host to the destination host.
- **Destination address.**
 - This 32-bit field defines the IPv4 address of the destination.
T
 - field must remain unchanged during the time the IPv4 datagram travels from the source host to the destination host.

Fragmentation

- A IP packet can travel through many different networks using different Data Link layers
- divide the datagram to make it possible to pass through other networks(whose packet size is smaller). This is called fragmentation
- when a datagram is encapsulated in a frame, the total size of the datagram must be less than this maximum size(**Maximum Transfer Unit (MTU)**)
- **MTU**
 - Each datalink layer has its own frame format and limitation.
 - One of such limitation is the maximum size of the frame, which is imposed by software, hardware, performance, and standards.



Each router decapsulates the IPv4 datagram from the frame it receives, processes it, and then encapsulates it in another frame

- For example, if a router connects a LAN to a WAN, it receives a frame in the LAN format and sends a frame in the WAN format

But, there is a possibility that a **packet travel thru a link whose MTU is smaller than one of the source node.**

- Then, the packet must be fragmented to go forward the next hop.
- Each fragment has its own header mostly repeated from the original packet.
- A fragmented packet can be further fragmented into even smaller packet.
- Fragmented packets will be re-assembled only by the final destination.

- a datagram can be **fragmented by the source host or any router in the path** although there is a tendency to limit fragmentation only at the source.
- **The reassembly of the datagram, however, is done only by the destination host** because each fragment becomes an independent datagram

- When a datagram is fragmented, **required parts of the header must be copied by all fragments**
- The host or router that fragments a datagram must **change the values of three fields:**
 1. flags
 2. fragmentation offset
 3. total length.
- The rest of the fields must be copied

- **Identification**
 - identifies a datagram originating from the source host.
 - A combination of the identification and source address must uniquely define a datagram as it leaves the source node.

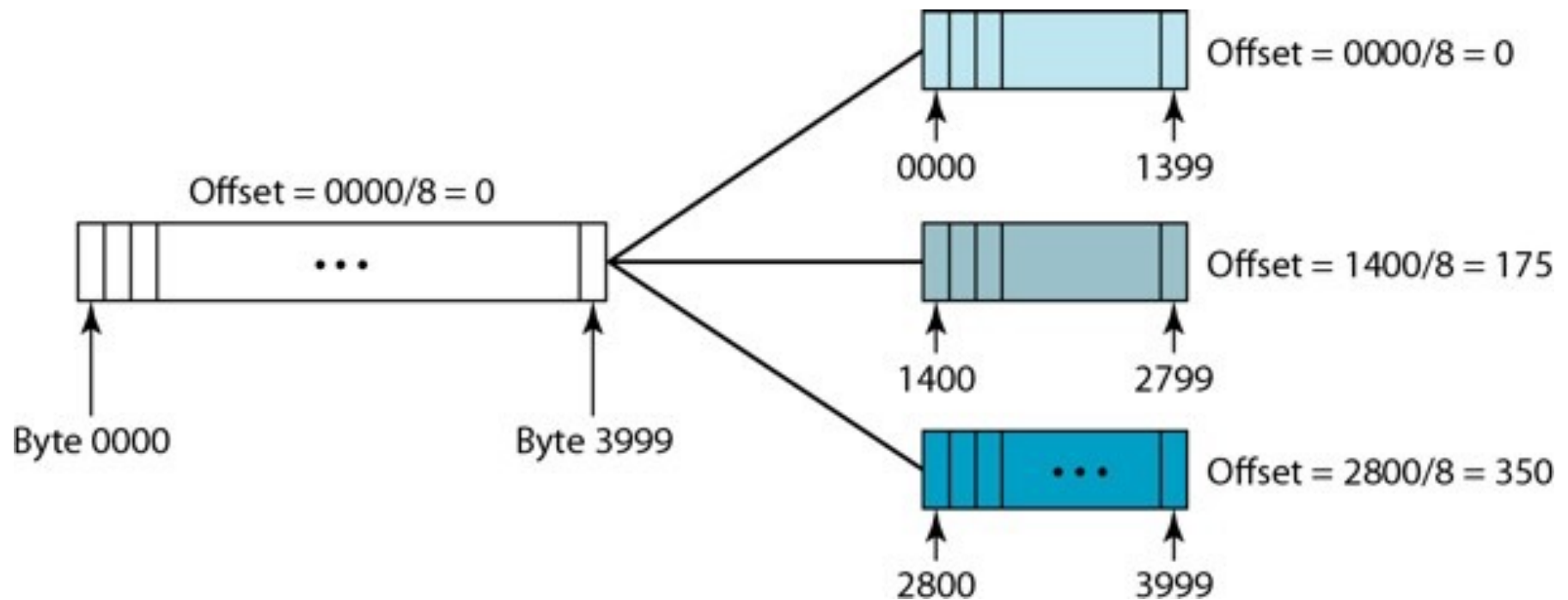
- **Fragmentation offset**
 - the relative position of this fragment with respect to the whole datagram
 - measured in units of 8 bytes.

FLAGS



- **first bit: reserved (not used)**
- **second bit:** = 1 means the packet should not be fragmented
 - A router on the path drops the packet if its size is $>$ MTU & sends an ICMP error message to the source host
- **third bit:** =1 -> tell the destination more fragmented packets follow this datagram
 - =0 the last fragmented packet

- shows a datagram with a data size of 4000 bytes fragmented into three fragments.



IP Fragmentation and Reassembly

Example

- ❑ 4000 byte datagram
- ❑ MTU = 1500 bytes

length	ID	fragflag	offset
=4000	=x	=0	=0

One large datagram becomes several smaller datagrams

1480 bytes in data field

offset = $1480/8$

length	ID	fragflag	offset
=1500	=x	=1	=0

length	ID	fragflag	offset
=1500	=x	=1	=185

length	ID	fragflag	offset
=1040	=x	=0	=370

IPv4 checksum

- use the 1's complement method
- The implementation of the checksum in the IPv4 packet follows the same principles.
 - ❑ First, the value of the **checksum field is set to 0.**
 - ❑ Then the entire header is **divided into 16-bit sections and added together.**
 - ❑ The result (sum) is **complemented and inserted into the checksum field.**

- The **checksum** in the IPv4 packet covers only the **header, not the data**.
- There are two good reasons for this.
 - First, **all higher-level protocols that encapsulate data in the IPv4 datagram have a checksum field that covers the whole packet**. Therefore, the checksum for the IPv4 datagram does not have to check the encapsulated data.
 - Second, **the header of the IPv4 packet changes with each visited router, but the data do not**

Options

- The header of the IPv4 datagram is made of two parts: a fixed part and a variable part.
- The fixed part is 20 bytes long .
- The variable part comprises the options that can be a maximum of 40 bytes.
- Options, are not required for a datagram.
- They can be used for **network testing and debugging.**

No Operation

- A 1-byte option used as a filler between options.

End of Option

- a 1-byte option used for padding at the end of the option field. It, however, can only be used as the last option.

Record Route

- is used to record the Internet routers that handle the datagram.
- It can list up to nine router addresses. It can be used for debugging and management purposes.

Strict Source Route

- A strict source route option is used by the source to predetermine a route for the datagram as it travels through the Internet.
- Dictation of a route by the source can be useful

- **Loose Source Route**

- A loose source route option is similar to the strict source route, but it is less rigid.
- Each router in the list must be visited, but the datagram can visit other routers as well.

- **Timestamp**

- A timestamp option is used to record the time of datagram processing by a router
- Knowing the time a datagram is processed can help users and managers track the behavior of the routers in the Internet.
- can estimate the time it takes for a datagram to go from one router to another

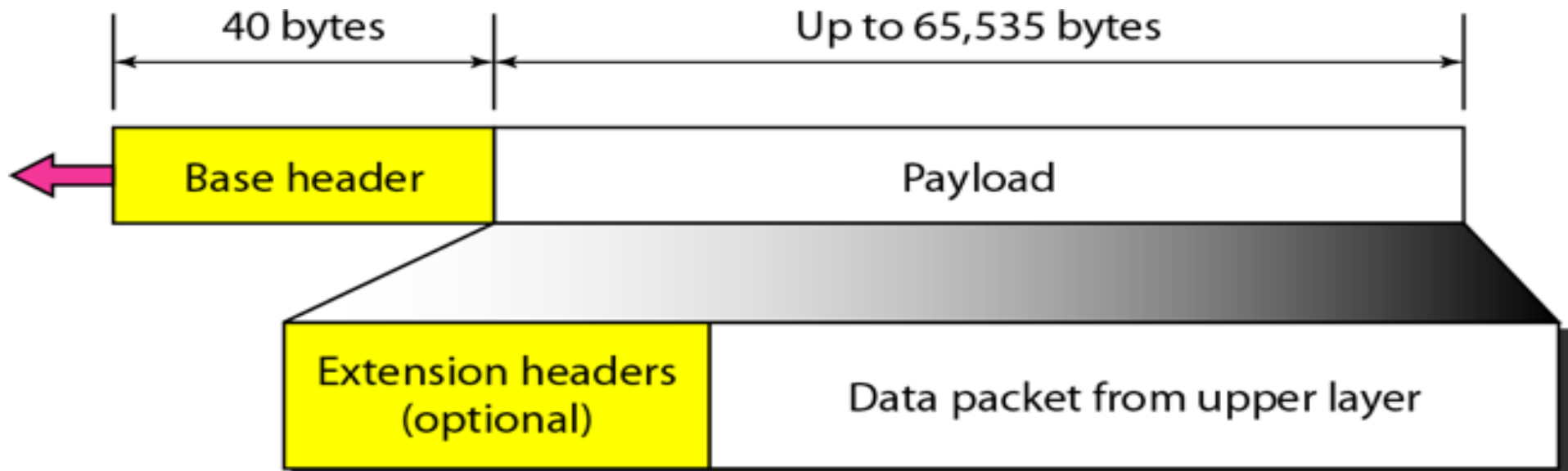
- **IPv4 has some deficiencies** that make it unsuitable for the fast-growing Internet.
 - Despite all short-term solutions, such as subnetting, classless addressing, and NAT, **address depletion is still a long-term problem in the Internet.**
 - The Internet must accommodate real-time audio and video transmission. This type of transmission requires **minimum delay strategies and reservation of resources not provided in the IPv4 design.**
 - The Internet must accommodate encryption and authentication of data for some applications. **No encryption or authentication is provided by IPv4.**
- To overcome these deficiencies, IPv6 (Internetworking Protocol, version 6), also known as IPng (Internetworking Protocol, next generation)

Advantages of IPv6

- **Larger address space.**
- **Better header format.** IPv6 uses a new header format in which **options are separated from the base header and inserted, when needed, between the base header and the upper-layer data.** This simplifies and speeds up the routing process because most of the options do not need to be checked by routers.
- **New options.** IPv6 has new options to allow for additional functionalities.
- **Allowance for extension.** IPv6 is designed to allow the extension of the protocol if required by new technologies or applications.
- **Support for resource allocation.** In IPv6, the type-of-service field has been removed, but a mechanism (called flow label) has been added to enable the source to request special handling of the packet.
 - This mechanism can be used to support traffic such as real-time audio and video.
- **Support for more security.** The encryption and authentication options in IPv6 provide confidentiality and integrity of the packet.

Packet Format

- Each packet is composed of **mandatory base header followed by the payload.**
- The payload consists of two parts: **optional extension headers and data from an upper layer.**
- The base header occupies **40 bytes,**
- the extension headers and data from the upper layer contain up to **65,535 bytes** of information.



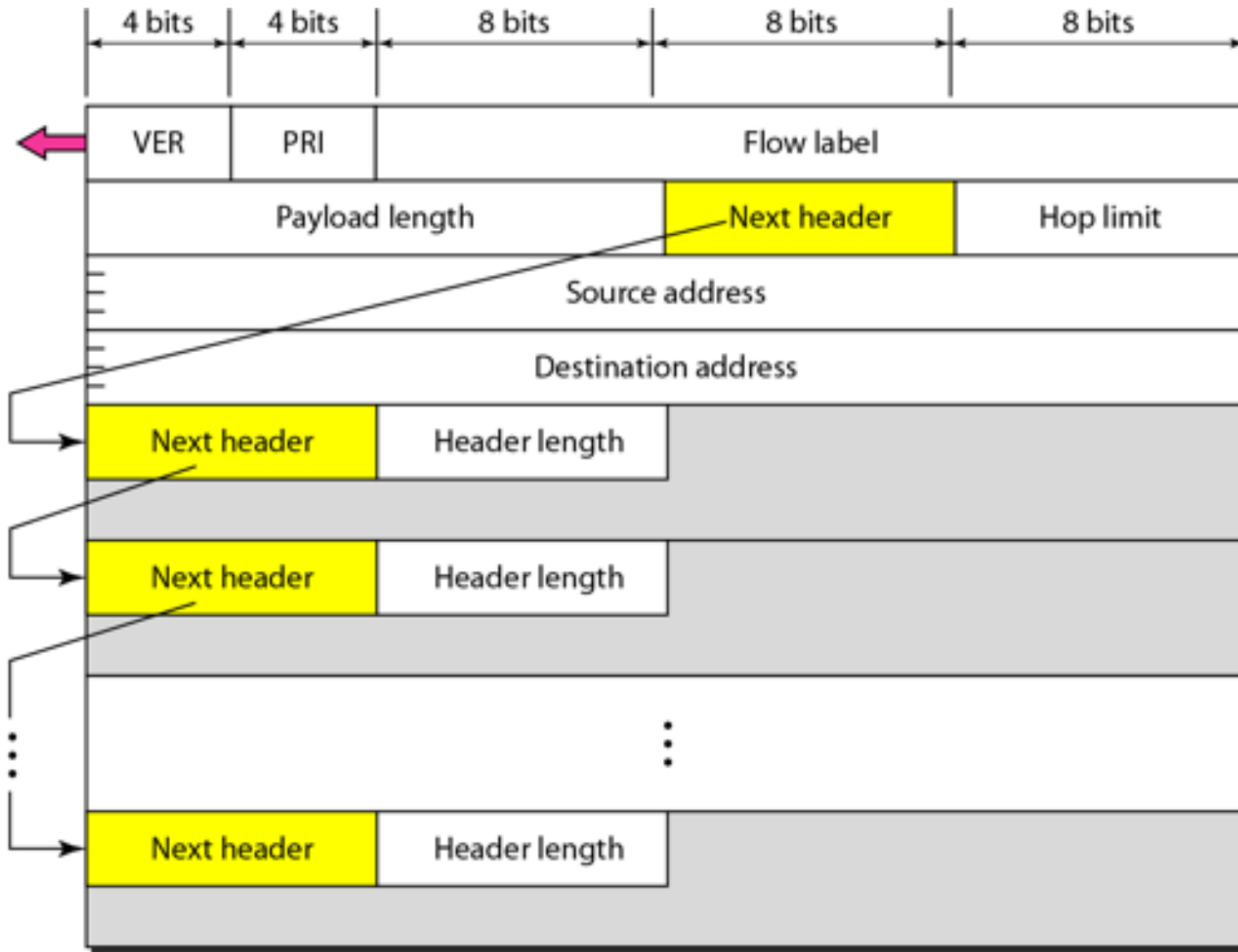
Base Header

- Has 8 fields
 - Version
 - the version number of the IP.
 - Priority.
 - the priority of the packet with respect to traffic congestion.
 - Priority numbers from 8 to 15 are assigned to noncongestion-controlled traffic
 - Congestion-controlled data are assigned priorities from 0 to 7
 - A priority of 0 is the lowest; a priority of 7 is the highest.
 - Flow label.
 - The flow label is a 3-byte (24-bit) field that is designed to provide special handling for a particular flow of data.
 - a flow label can be used to support the transmission of real-time audio and video.
 - Real-time audio or video, particularly in digital form, requires resources such as high bandwidth, large buffers, long processing time,

- **Payload length.**
 - The 2-byte payload length field defines the length of the IP datagram excluding the base header.
- **Next header.**
 - The next header is an 8-bit field defining the header that follows the base header in the datagram.
 - The next header is either one of the optional extension headers used by IP or the header of an encapsulated packet such as UDP or TCP. Each extension header also contains this field.
- **Hop limit.**
 - This 8-bit hop limit field serves the same purpose as the TTL field in IPv4.
- **Source address.**
 - The source address field is a 16-byte (128-bit) Internet address that identifies the original source of the datagram.
- **Destination address.**
 - The destination address field is a 16-byte (128-bit) Internet address that usually identifies the final destination of the datagram. However, if source routing is used, this field contains the address of the next router.

- **Extension Headers**

- to give greater functionality to the IP datagram, the base header can be followed by up to six extension headers



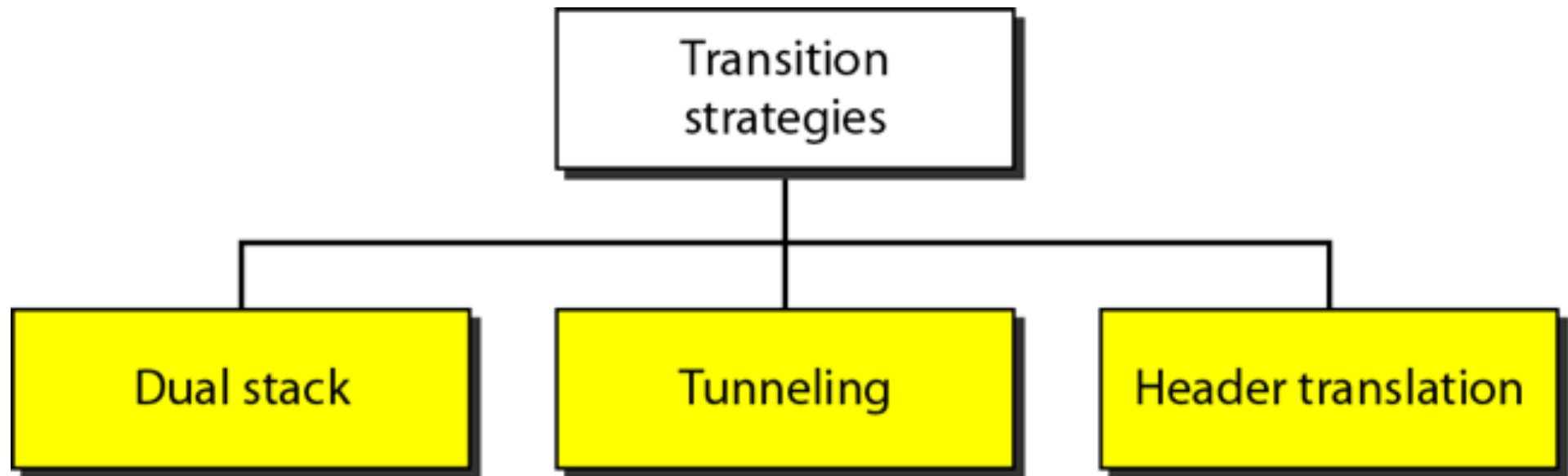
Comparison

1. The header length field is eliminated in IPv6 because the length of the header is fixed in this version.
2. The service type field is eliminated in IPv6. The priority and flow label fields together take over the function of the service type field.
3. The total length field is eliminated in IPv6 and replaced by the payload length field.
4. The identification, flag, and offset fields are eliminated from the base header in IPv6. They are included in the fragmentation extension header.
5. The TTL field is called hop limit in IPv6.
6. The protocol field is replaced by the next header field.
7. The header checksum is eliminated because the checksum is provided by upper-layer protocols; it is therefore not needed at this level.
8. The option fields in IPv4 are implemented as extension headers in IPv6.

TRANSITION FROM IPv4 TO IPv6

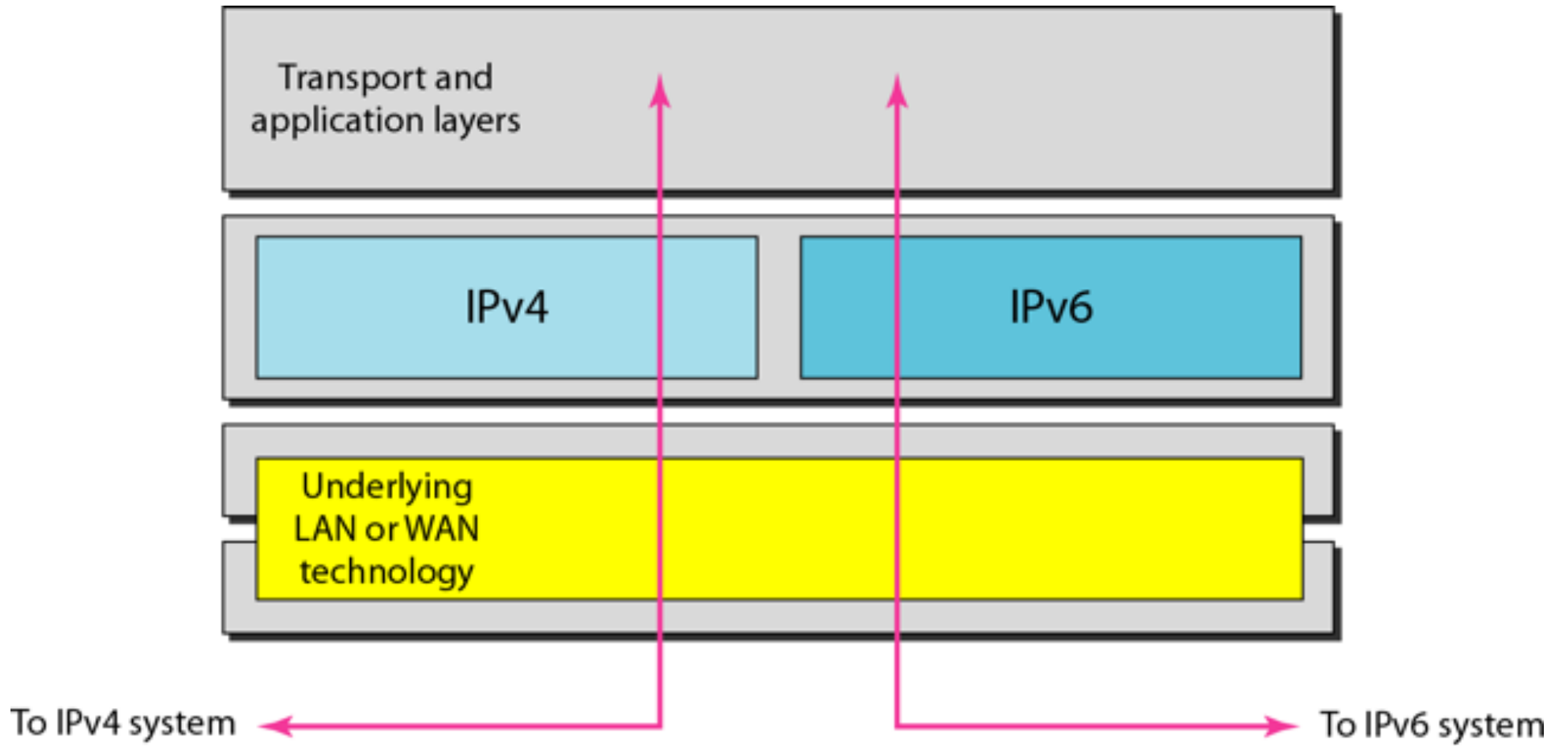
- *Because of the huge number of systems on the Internet, the transition from IPv4 to IPv6 cannot happen suddenly.*
- *It takes a considerable amount of time before every system in the Internet can move from IPv4 to IPv6.*
- *The transition must be smooth to prevent any problems between IPv4 and IPv6 systems*

TRANSITION FROM IPv4 TO IPv6



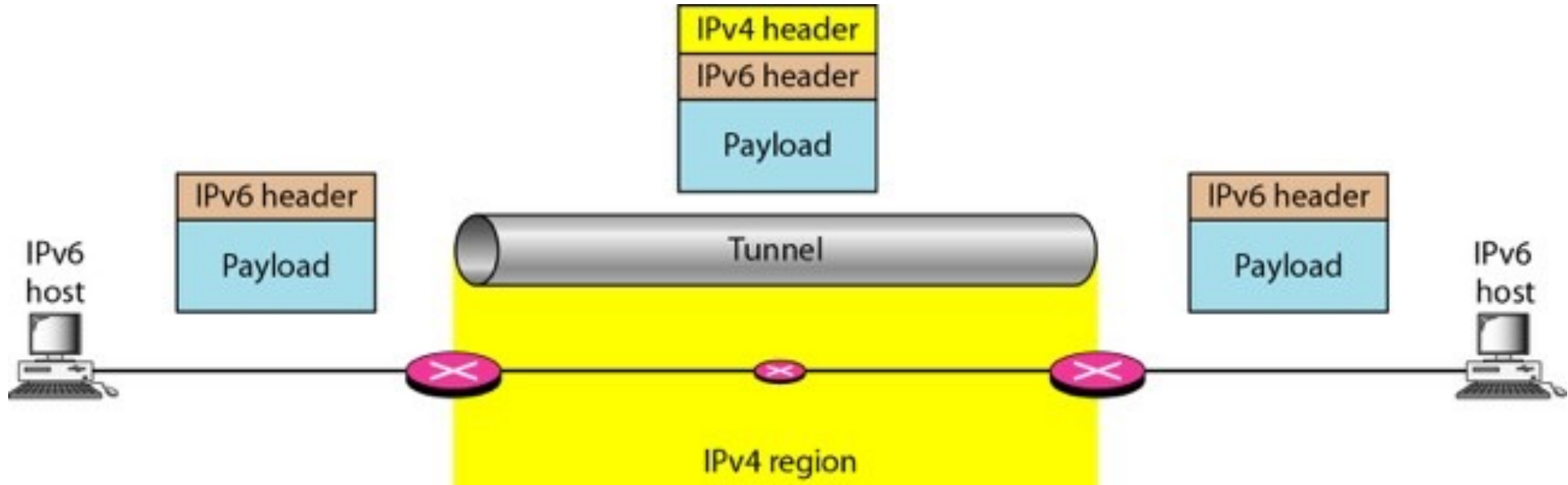
Dual stack

- recommended that all hosts, before migrating completely to version 6, have a dual stack of protocols
- station must run IPv4 and IPv6 simultaneously until all the Internet uses IPv6.
- To determine which version to use when sending a packet to a destination, the source host queries the DNS.
- If the DNS returns an IPv4 address, the source host sends an IPv4 packet.
- If the DNS returns an IPv6 address, the source host sends an IPv6 packet.



Tunneling

- a strategy used when two computers using IPv6 want to communicate with each other and the packet must pass through a region that uses IPv4.
- To pass through this region, the packet must have an IPv4 address.
- So the IPv6 packet is encapsulated in an IPv4 packet when it enters the region, and it leaves its capsule when it exits the region.
- It seems as if the IPv6 packet goes through a tunnel at one end and emerges at the other end.
- To make it clear that the IPv4 packet is carrying an IPv6 packet as data, the protocol value is set to 41



Header translation

- Header translation is necessary when the majority of the Internet has moved to IPv6 but some systems still use IPv4.
- The sender wants to use IPv6, but the receiver does not understand IPv6.
- Tunneling does not work in this situation because the packet must be in the IPv4 format to be understood by the receiver.
- In this case, the header format must be totally changed through header translation.
- The header of the IPv6 packet is converted to an IPv4 header

Header translation

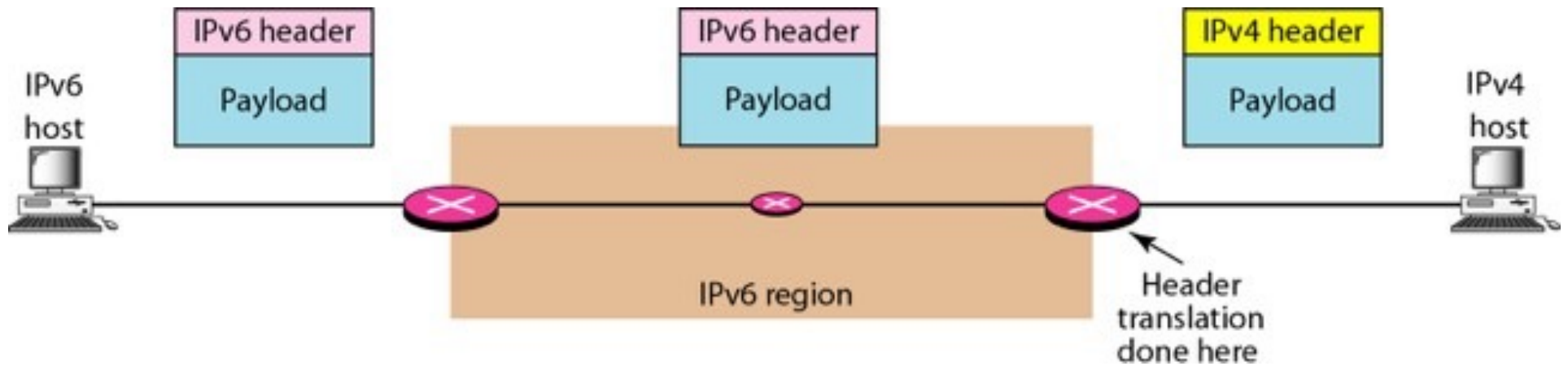
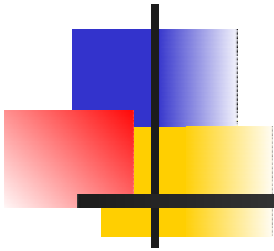


Table 20.11 *Header translation*

<i>Header Translation Procedure</i>
1. The IPv6 mapped address is changed to an IPv4 address by extracting the rightmost 32 bits.
2. The value of the IPv6 priority field is discarded.
3. The type of service field in IPv4 is set to zero.
4. The checksum for IPv4 is calculated and inserted in the corresponding field.
5. The IPv6 flow label is ignored.
6. Compatible extension headers are converted to options and inserted in the IPv4 header. Some may have to be dropped.
7. The length of IPv4 header is calculated and inserted into the corresponding field.
8. The total length of the IPv4 packet is calculated and inserted in the corresponding field.

Network Layer: Address Mapping, Error Reporting, and Multicasting



Address mapping

- *The delivery of a packet to a host or a router requires two levels of addressing: **logical(IP)** and **physical***
- *map a logical address to its corresponding physical address and vice versa.*
- *This can be done by using either **static** or **dynamic mapping**.*

Static mapping

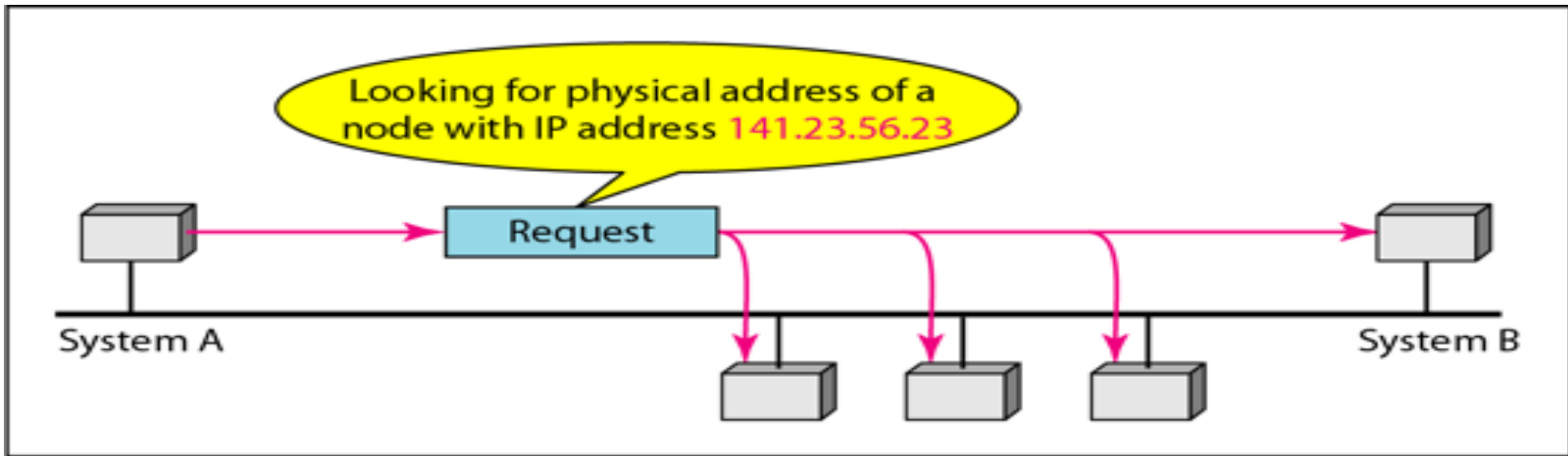
- Uses a table that associates a logical address with a physical address.
- This table is stored in each machine on the network.
- Each machine that knows, for example, the IP address of another machine but not its physical address can look it up in the table

Limitations of static mapping

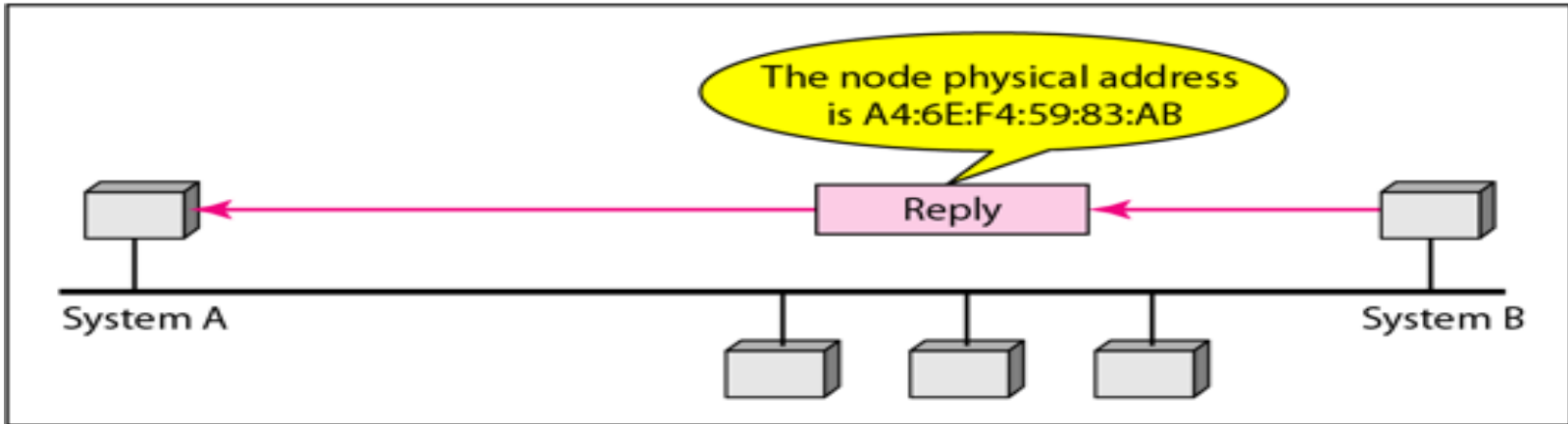
1. A machine could change its NIC, resulting in a new physical address.
 2. In some LANs, such as LocalTalk, the physical address changes every time the computer is turned on.
 3. A mobile computer can move from one physical network to another, resulting in a change in its physical address.
- to overcome these limitations, a **static mapping table must be updated periodically.**
 - This overhead could affect network performance.

Dynamic Mapping

- Uses **Address Resolution Protocol(ARP)**
- Anytime a host or a router has an IP datagram to send to another host or router, it has the logical (IP) address of the receiver.
- IP address is obtained from the DNS if sender is a host/ from the routing table if sender is router
- to pass the IP datagram through the physical network, sender needs the physical address of the receiver.
- The host or the router **sends an ARP query packet.**



a. ARP request is broadcast



b. ARP reply is unicast

These are the steps involved in an ARP process:

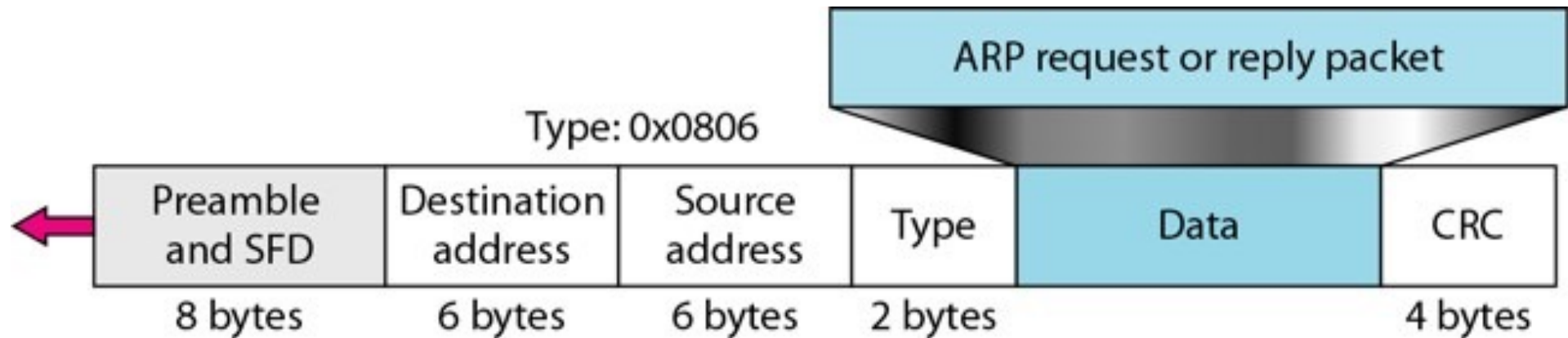
1. The sender knows **the IP address of the target.**

2. IP asks ARP to create an **ARP request** message, filling in the sender physical address, the sender IP address, and the target IP address.
 - The target physical address field is filled with 0s.

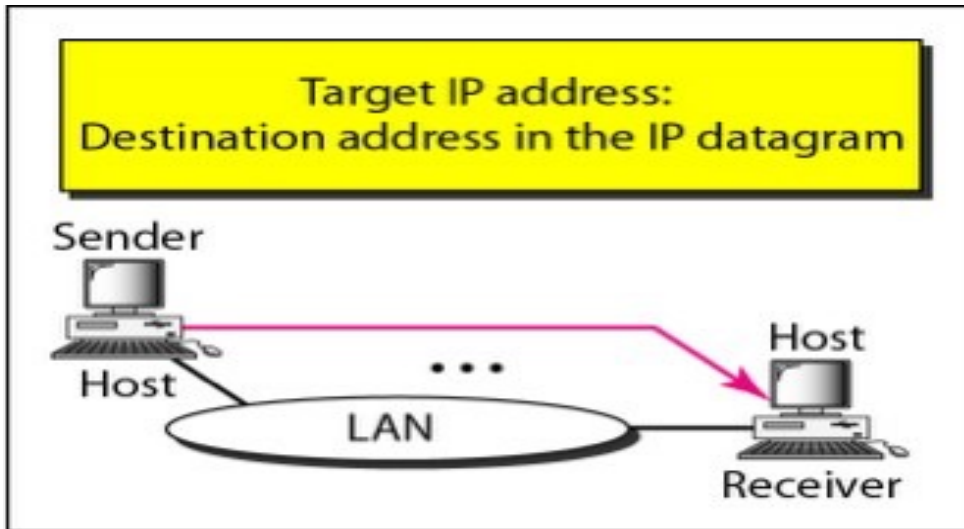
3. The message is passed to the data link layer where it is encapsulated in a frame by using **the physical address of the sender as the source address and the physical broadcast address as the destination address.**

4. Every host or router receives the frame. Because the frame contains **a broadcast destination address**, all stations accept the message and pass it to ARP. All machines except the one targeted **drop the packet**. The target machine recognizes its IP address.
5. The target machine replies with an **ARP reply message** that contains its physical address. The message is **unicast**.
6. The **sender receives the reply message**. It now knows the physical address of the target machine.
7. The IP datagram, which carries data for the target machine, is now encapsulated in a frame and is unicast to the destination.

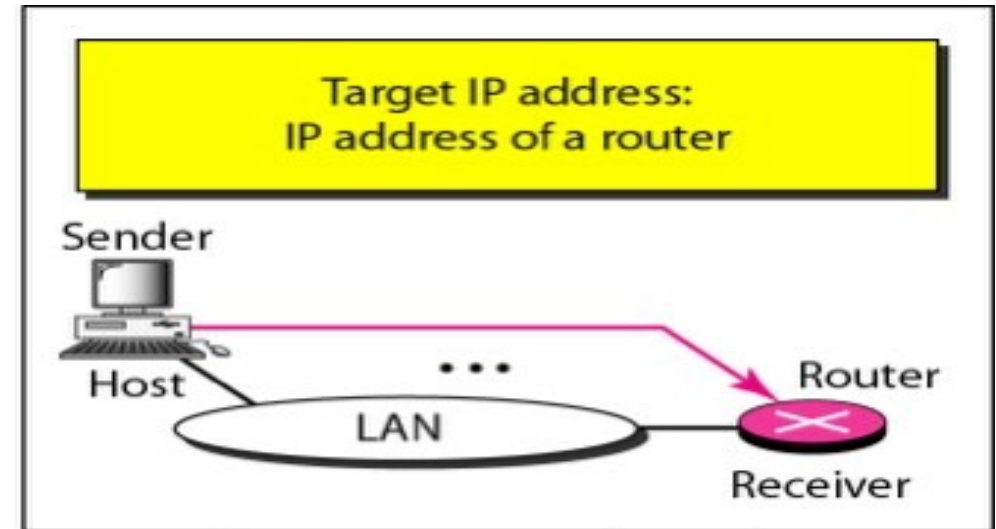
Encapsulation of ARP packet



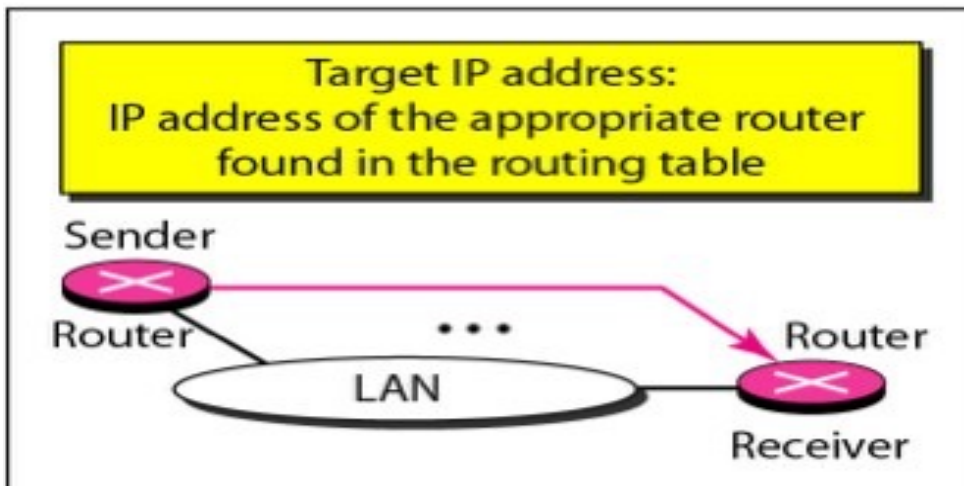
Four cases which uses ARP



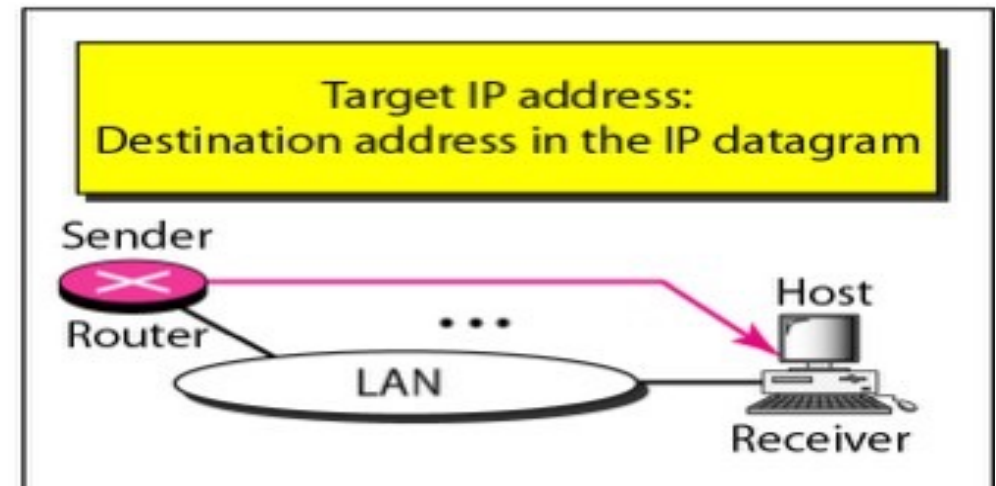
Case 1. A host has a packet to send to another host on the same network.



Case 2. A host wants to send a packet to another host on another network. It must first be delivered to a router.



Case 3. A router receives a packet to be sent to a host on another network. It must first be delivered to the appropriate router.



Case 4. A router receives a packet to be sent to a host on the same network.

Case 1

- The sender is a host and wants to send a packet to another host on the same network.
- In this case, the logical address that must be mapped to a physical address is the destination IP address in the datagram header.

Case 2

- The sender is a host and wants to send a packet to another host on another network.
- In this case, the host looks at its routing table and finds the IP address of the next hop (router) for this destination.
- If it does not have a routing table, it looks for the IP address of the default router.
- The IP address of the router becomes the logical address that must be mapped to a physical address.

Case 3.

- The sender is a router that has received a datagram destined for a host on another network.
- It checks its routing table and finds the IP address of the next router.
- The IP address of the next router becomes the logical address that must be mapped to a physical address.

Case 4

- The sender is a router that has received a datagram destined for a host on the same network.
- The destination IP address of the datagram becomes the logical address that must be mapped to a physical address

using Cache Memory

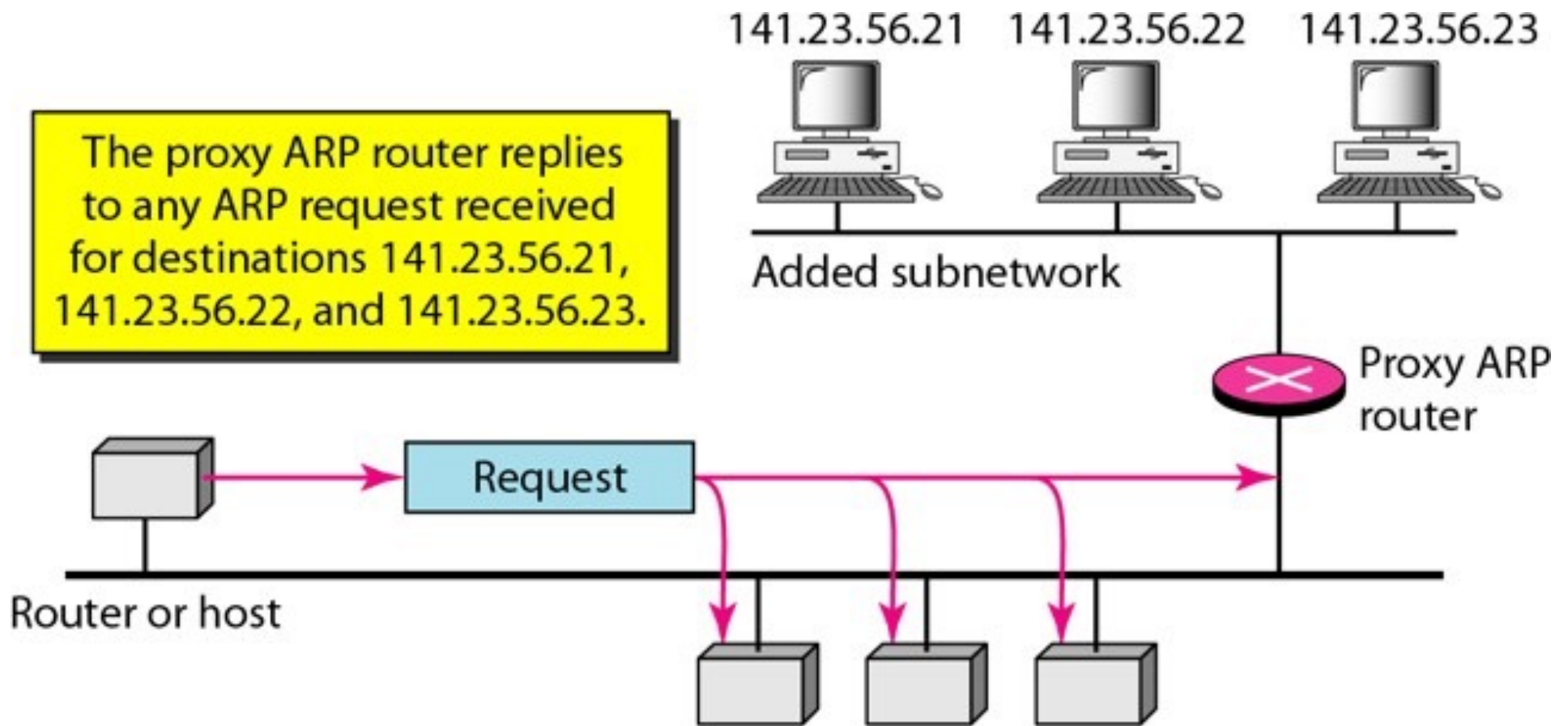
- Using ARP is inefficient if system A needs to broadcast an ARP request for each IP packet it needs to send to system B.
- It could have broadcast the IP packet itself.
- **ARP can be useful if the ARP reply is cached (kept in cache memory for a while) because a system normally sends several packets to the same destination.**
- **A system that receives an ARP reply stores the mapping in the cache memory and keeps it for 20 to 30 minutes unless the space in the cache is exhausted.**
- Before sending an ARP request, the system first checks its cache to see if it can find the mapping.

ProxyARP

- A proxy ARP is an ARP that acts on behalf of a set of hosts.
Whenever a router running a proxy ARP receives an ARP request
- looking for the IP address of one of these hosts, the router sends an ARP reply announcing its own hardware (physical) address.

- After the router receives the actual IP packet, it sends the packet to the appropriate host or router.

- Proxy ARP is used to create an subnetting effect



Mapping Physical to Logical Address:

Cases in which a host knows its physical address, but needs to know its logical address.

1. A **diskless station is just booted**. The station can find its physical address by checking its interface, but it does not know its IP address.
2. **An organization does not have enough IP addresses to assign to each station**; it needs to assign IP addresses **on demand**. The station can send its physical address and ask for a short time lease.

- **Mapping Physical to Logical Address:**

1. RARP 2. BOOTP 3. DHCP

RARP(Reverse ARP)

- Physical address is known (from NIC), does not know the IP address(diskless station)
- RARP uses **physical address to get the logical address by using RARP protocol**
- A RARP request is created by **the RARP client** and broadcast on the local network.
- **RARP Server machine** on the local network that knows all the IP addresses **will respond with a RARP reply.**

- **Disadvantage** : there should be a RARP server on each n/w because the **physical broadcast address**, all 1's in the case of Ethernet, **does not pass the boundaries of a network**
- **Hence it is obsolete**

Bootstrap Protocol (BOOTP)

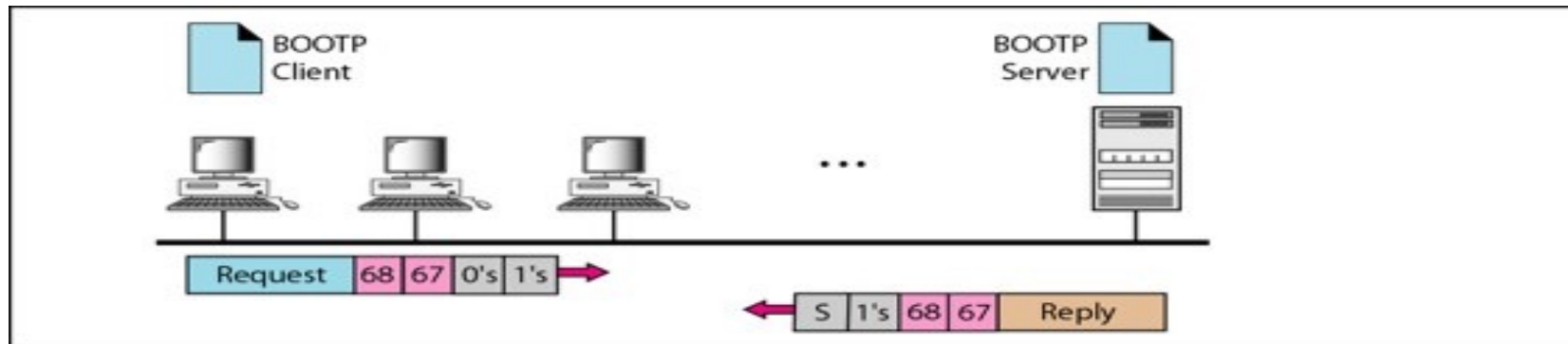
It is application layer protocol

- a **client/server protocol designed to provide physical address to logical address mapping.**
- The administrator may put the client and the server on the same network or on different networks
- BOOTP messages are encapsulated in UDP packet and the UDP packet is encapsulated in IP packet
- a **client** can send an IP datagram without knowing neither its own IP address (the source address) nor the server's IP address (the destination address) by using **all 0s as the source address and all 1s as the destination address.**

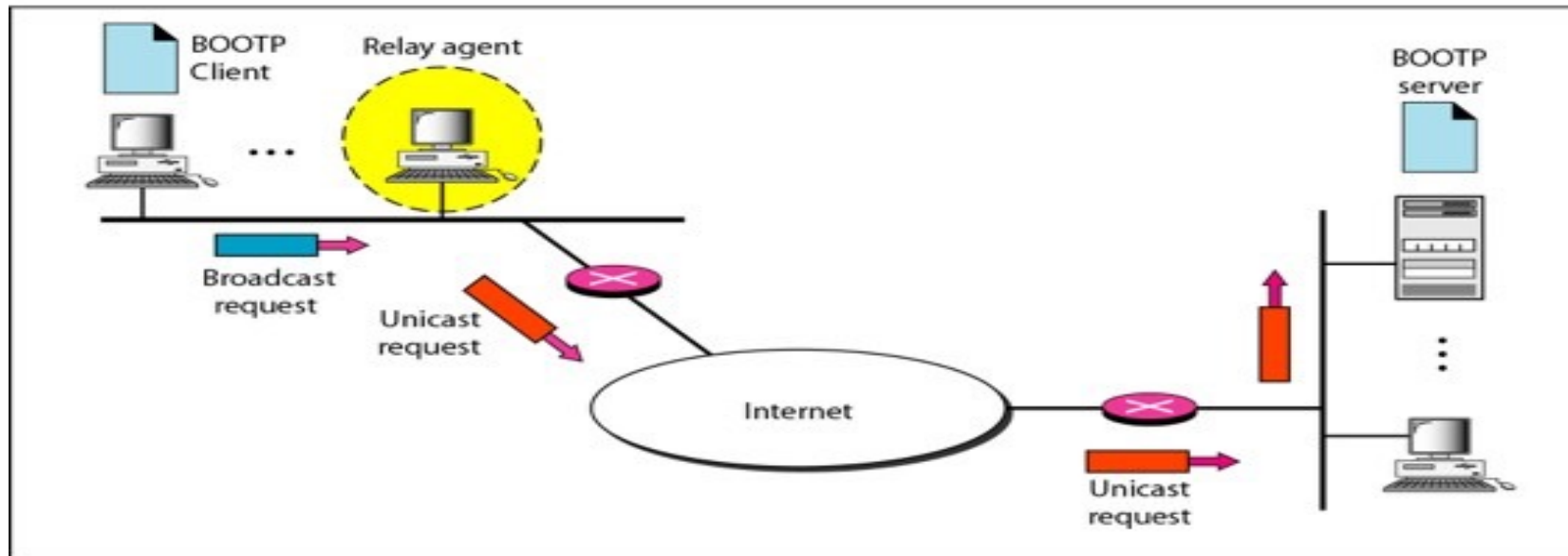
- The BOOTP **request is broadcast** because the client does not know the IP address of the server.
- **BOOTP server consults a table that matches the physical address of the client with its IP address**
- A broadcast IP datagram **cannot pass through any router.**
- To solve the problem, there is a need for an intermediary. One of the hosts/router is a relay agent

- **The relay agent knows the unicast address of a BOOTP server.**
- **When it receives this type of packet, it encapsulates the message in a unicast datagram and sends the request to the BOOTP server**
-
- **The BOOTP server knows the message comes from a relay agent because one of the fields in the request message defines the IP address of the relay agent.**
- **The relay agent, after receiving the reply, sends it to the BOOTP client**

- BOOTP is a static configuration protocol



a. Client and server on the same network



b. Client and server on different networks

Dynamic Host Configuration Protocol (DHCP)

- provides static and dynamic address allocation that can be **manual or automatic.**

- **Static Address Allocation**

- DHCP acts as BOOTP

- **Dynamic Address Allocation**

- Uses 2 databases -> static & dynamic

- DHCP has a second database with a pool of available

IP addresses. This second database makes DHCP dynamic

- When a **DHCP client sends a request to a DHCP server**, the server first checks its static database.

- If an entry with the requested physical address **exists** in the static database, **the permanent IP address of the client** is returned.
- if the entry **does not exist** in the static database, **the server selects an IP address from the available pool**, assigns the address to the client, and adds the entry to the dynamic database for a period of time(lease)

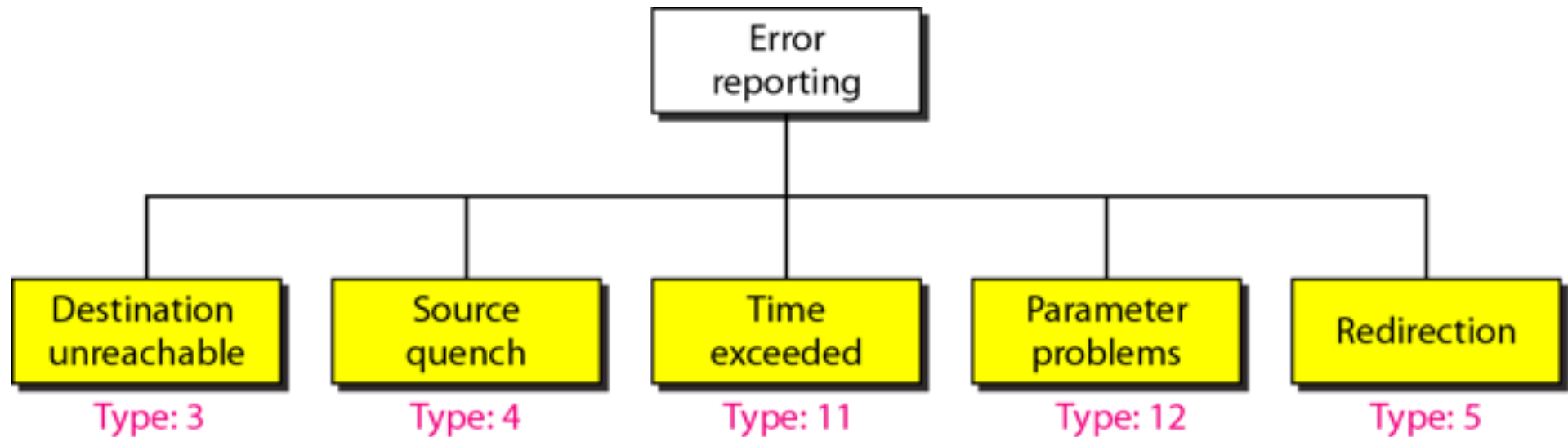
Manual and Automatic Configuration

- One major problem with the **BOOTP** protocol is that the table mapping the IP addresses to physical addresses needs to be **manually configured**.
- This means that **every time there is a change in a physical or IP address**, the administrator needs to manually enter the changes.
- DHCP, on the other hand, allows both manual and automatic configurations.
- Static databases are created manually~ dynamic databases are created automatically

ICMP

- **IP protocol has no error-reporting or error-correcting mechanism**
- **Internet Control Message Protocol (ICMP) has been designed to compensate for the above two deficiencies.**
- ICMP messages are divided into two broad categories
 - **error-reporting messages (problems encountered by host/router when a IP packet is processed)**
 - **query messages**
 - **help a host or a network manager get specific information from a router or another host**
 - **Eg nodes can discover their neighbours, can learn abt routers**

Error reporting messages



■ Destination Unreachable

- When a router cannot route a datagram or a host cannot deliver a datagram, the datagram **is discarded** and the router or the host **sends a destination-unreachable message back to the source host that initiated the datagram**

■ Source Quench

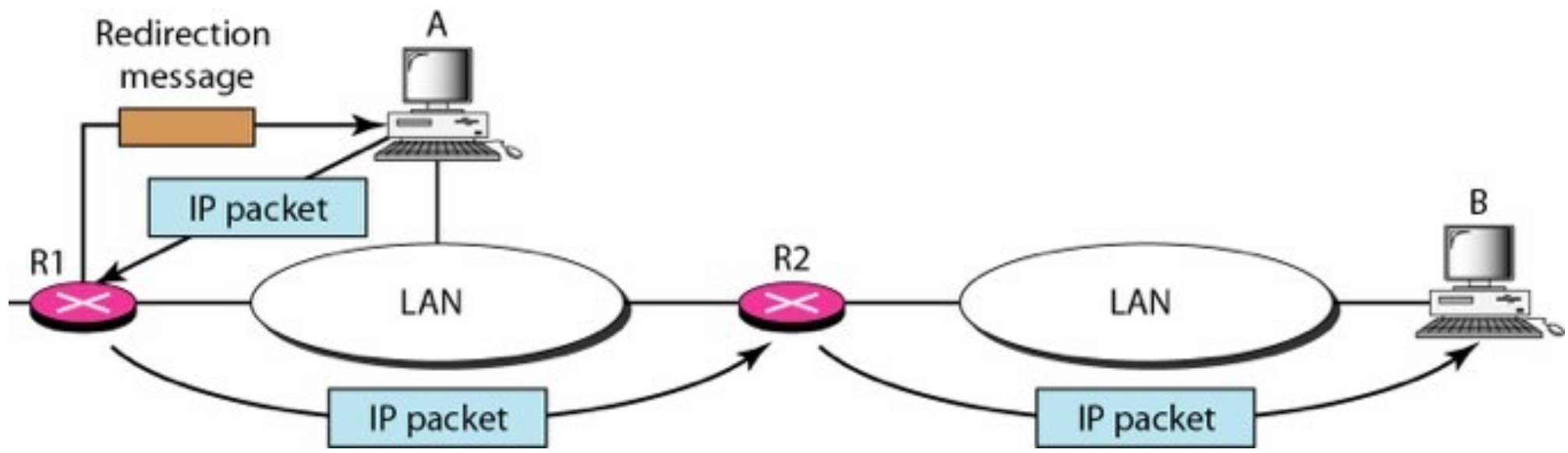
- If the datagrams are received much faster than they can be forwarded or processed, the **queue may overflow**.
- In this case, the **router or the host has no choice but to discard some of the datagrams**.
- The **source-quench message in ICMP was designed to add a kind of flow control to the IP**.
- When a router or host discards a datagram **due to congestion**, it sends a source-quench message to the sender of the datagram.

- This message has two purposes.
 - First, it **informs the source that the datagram has been discarded.**
 - Second, it warns the source that there is congestion somewhere in the path and that the source should **slow down (quench) the sending process**

- Time Exceeded
 - each datagram contains a field called **time to live** that controls this situation.
 - When a **datagram visits a router**, the value of this field is **decremented by 1**.
 - When the **time-to-live value reaches 0, after decrementing, the router discards the datagram**.
 - However, when the datagram is discarded, a **time-exceeded message must be sent by the router to the original source**.
 - Second, a **time-exceeded message is also generated when not all fragments that make up a message arrive at the destination host within a certain time limit**

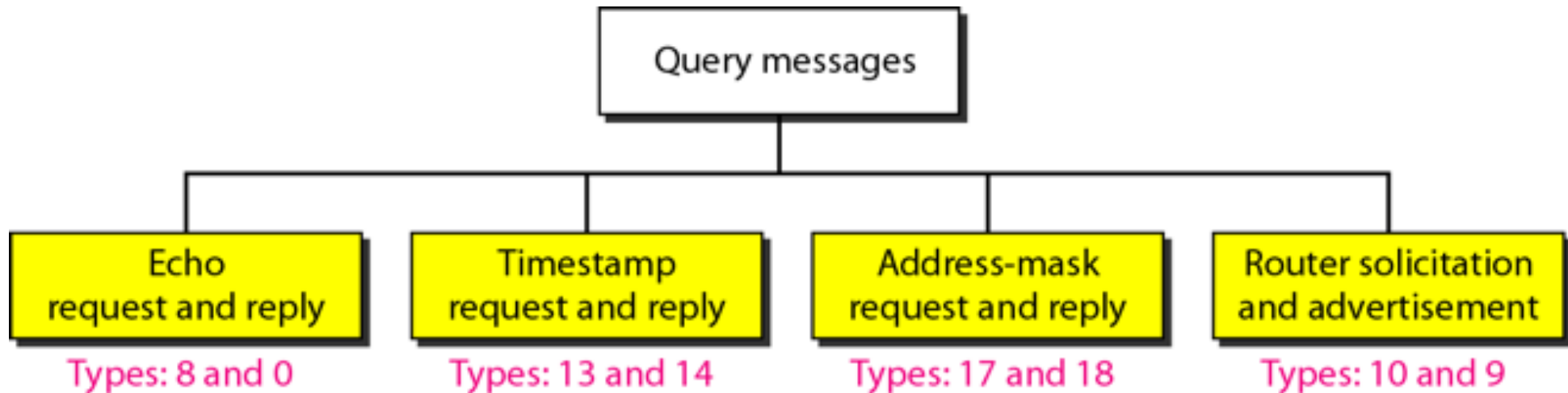
- Parameter Problem
 - If a router or the destination host discovers an **ambiguous or missing value in any field of the datagram**, it discards the datagram and sends a **parameter-problem message** back to the source.

- Redirection
 - to update the routing table of the host, it sends a redirection message to the host.



QUERY REPORTING

- ICMP can diagnose some network problems thru query messages



■ **Echo Request and Reply**

- The echo-request and echo-reply messages are designed for **diagnostic purposes.**
- **Network managers and users utilize this pair of messages to identify network problems.**
- The combination of echo-request and echo-reply messages determines whether two systems (hosts or routers) can communicate with each other. **e.g..ping command**

■ **Timestamp Request and Reply**

- Two machines (hosts or routers) can use the **timestamp request and timestamp reply messages to determine the round-trip time** needed for an IP datagram to travel between them.
- It can also be used to **synchronize the clocks in two machines**

- **Address-Mask Request and Reply**
 - A host may know its IP address but may not know its mask
 - To **obtain its mask**, a host sends an address-mask-request message to a router on the LAN.
 - If the **host knows the address of the router**, it sends the request **directly to the router**.
 - If it does not know, it **broadcasts the message**.
 - The router receiving the **address-mask-request** message responds with an **address-mask-reply** message, providing the necessary mask for the host

■ Router Solicitation and Advertisement

- a host that wants to send data to a host on another network needs to **know the address of routers connected to its own network.**
- Also, the host **must know if the routers are alive and functioning.**
- **The router-solicitation** and router-advertisement messages can help in this situation
- A host can **broadcast (or multicast) a router-solicitation message.**
- The router or routers that **receive the solicitation message, broadcast their routing information** using the **router-advertisement message.**
- A router can also **periodically send router-advertisement messages even if no host has solicited**

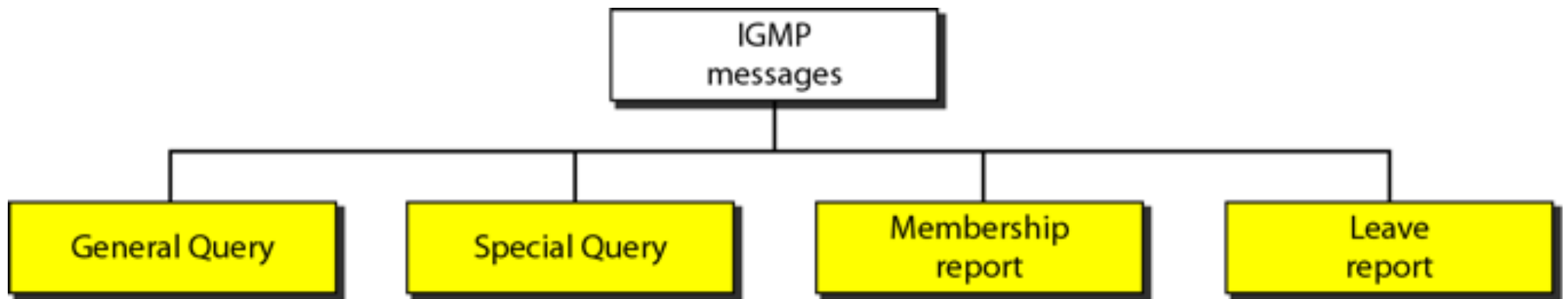
- ***Debugging Tool that uses ICMP***
 - **Packet InterNet Groper (ping)**
 - an application program that uses the services of ICMP to test the **reachability of a host**
 - We can use the ping program to find if a host is alive and responding.
 - **Traceroute**
 - traceroute program in UNIX or tracert in Windows can be used to trace the route of a packet from the source to the destination.

IGMP Internet Group Management Protocol

- The IP protocol can be involved in two types of communication: unicasting and multicasting.
- (IGMP) is involved in multicasting
- In any network, there are one or more *multicast routers that distribute multicast packets* to hosts or other routers.
- The IGMP protocol gives the multicast routers, *information about the membership status of hosts (routers)* connected to the network.

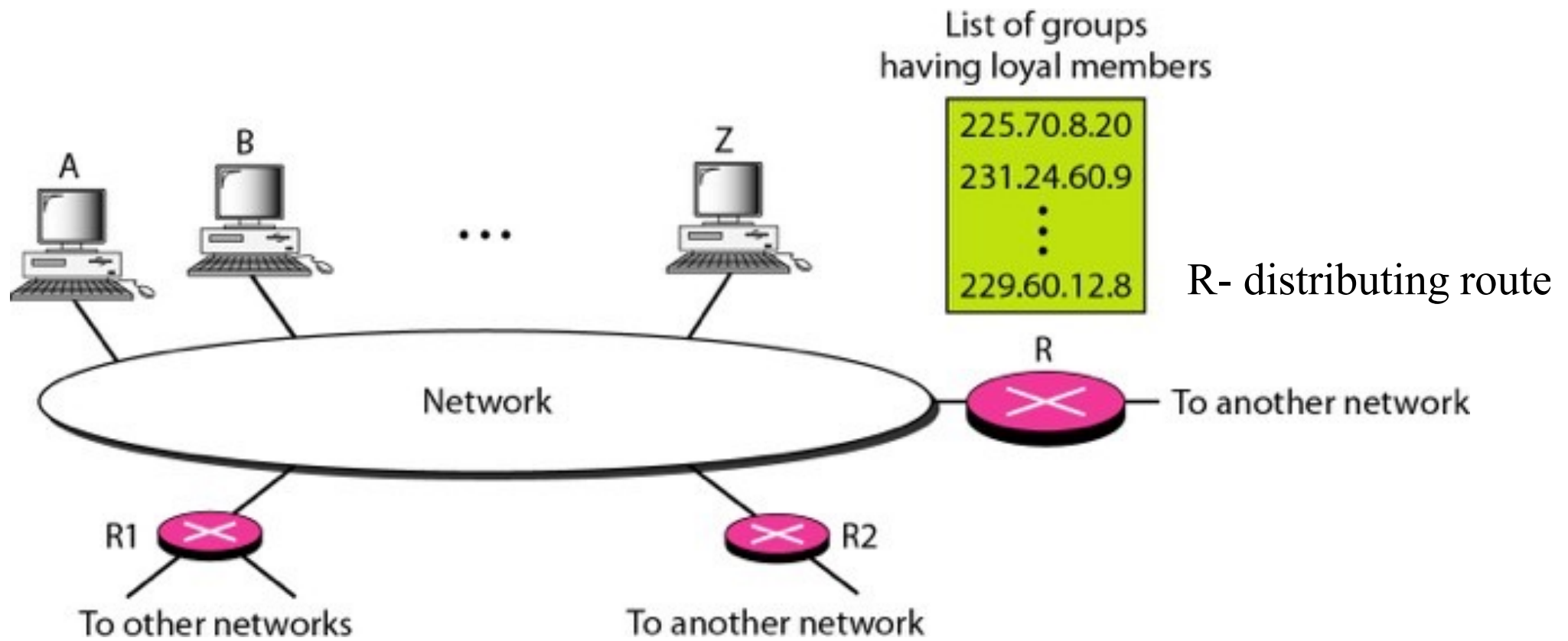
- A multicast router may receive thousands of multicast packets every day for different groups.
- If a router has no knowledge about the membership status of the hosts, it must broadcast all these packets. This creates a lot of traffic and consumes bandwidth.
- A better solution is to keep a list of groups in the network for which there is at least one loyal member.
- IGMP helps the multicast router create and update this list

- Internet Group Management Protocol (IGMP) helps multicast routers *create and update a list of loyal members* related to a network
- a protocol that *manages group membership*.
- *IGMP messages has 3 types of messages : general query, special query, membership report and leave report*



IGMP Operation

- IGMP operates locally.
- A **multicast router** connected to a network has a **list of multicast addresses of the groups with at least one loyal member** in that network
- For each group, there is **one router that has the duty of distributing the multicast packets** destined for that group.
(distributing router)



- A **host or multicast router** can have membership in a group.
- **When a host has membership**, it means that one of its processes (an application program) receives multicast packets from some group
- **When a router has membership**, it means that a network connected to one of its other interfaces receives these multicast packets
- host and the router keep a list of **groupids and relay** their interest to the distributing router.

Joining a group

- When a process wants to join a new group, it *sends its request* to the host.
- *host adds the name of the process* and the name of the requested group to its list
- If *the first entry* for this particular group, the *host sends a membership report message*.
- If this is **not the first entry**, there is **no need to send the membership** report.
- *a membership report is sent twice*, one after the other within a few moment, to ensure it's delivery.

Leaving a Group

- When a host sees that **no process is interested** in a specific group, it *sends a leave report*.
- when a router sees that none of the networks connected to its interfaces is interested in a specific group, it sends a leave report about that group.

Purge a group

- when a multicast router receives a leave report, it cannot immediately purge that group from the list
- the router sends a *special query message* and inserts the groupid, or multicast address, related to the group.
- allows *a specified time* for any host or router to respond
- **If no response, purge the group**

Monitoring Membership

- multicast router is responsible for **monitoring all the hosts or routers in a LAN**
- router **periodically (by default, every 125 s) sends a general query message**
- In this message, the *group address field is set to 0.0.0.0*. This means the query for *membership continuation is meant for all groups* in which a host is involved, not just one.
- query message has a *maximum response time of 10 s*
- *responds with a membership report* if it is interested in a group.
- common interest (two hosts, for example, are interested in the same group), only *one response is sent for that group to prevent unnecessary traffic(delayed response)*

- query message must be sent by only one router (normally called the *query router*), also to prevent unnecessary traffic

Delayed Response

- To prevent unnecessary traffic, IGMP uses a **delayed response strategy**.
- When a host or router receives a query message, it ***does not respond immediately; it delays the response.***
- Each host or router ***uses a random number to create a timer, which expires between 1 and 10seconds***
- expiration time can be ***in steps of 1s or less.***
- ***A timer is set for each group*** in the list.

- Each host or router waits until its timer has expired before sending a membership report message.
- During this waiting time, if the timer of another host or router, for the same group, expires earlier, that host or router sends a membership report.

Netstat Utility

The netstat utility can be used to find the multicast addresses supported by an interface

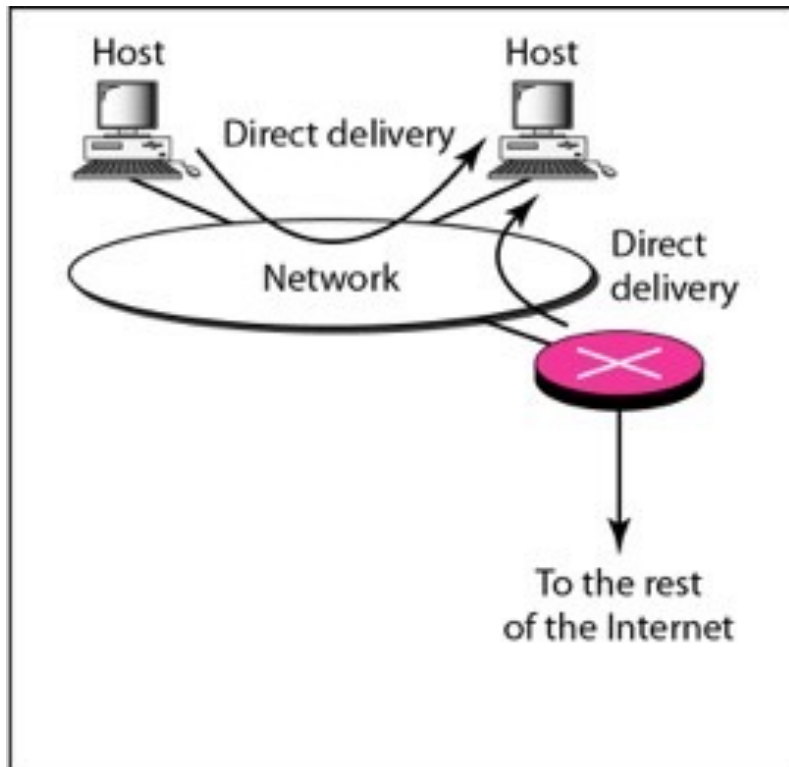
Chapter 22

Network Layer: Delivery, Forwarding, and Routing

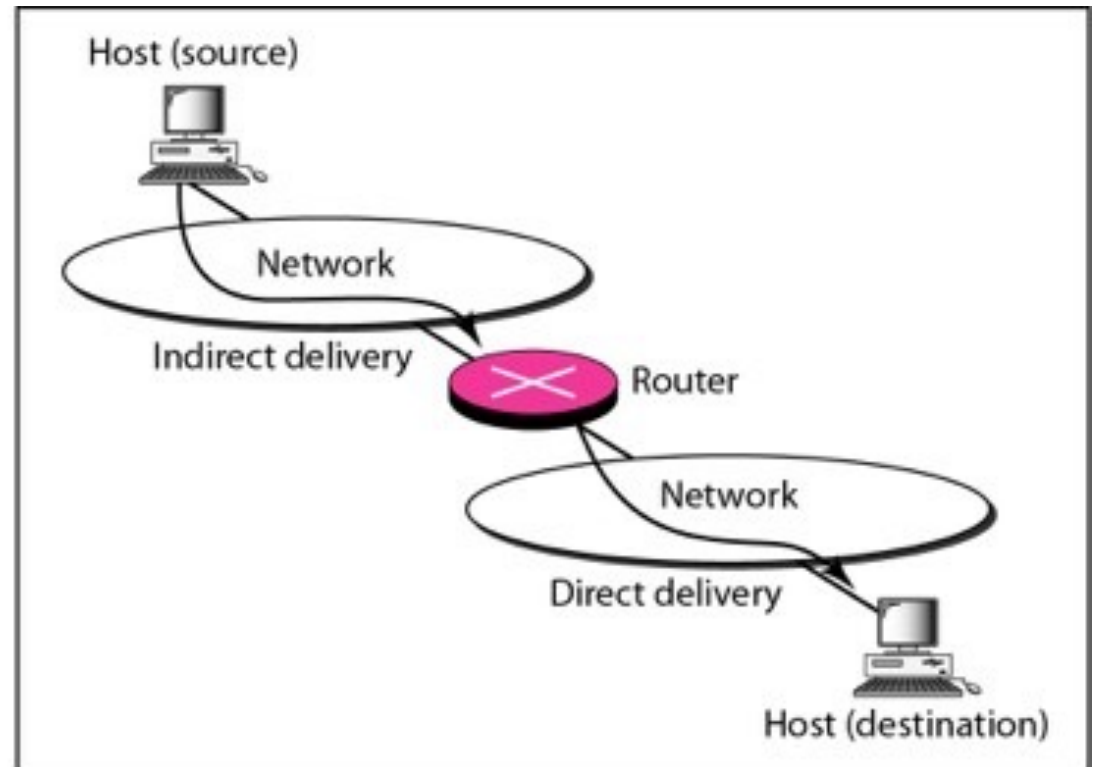
- **Delivery** refers to the way a packet is handled by the underlying networks under the control of the network layer.
- **Forwarding** refers to the way a packet is delivered to the next station.
- **Routing** refers to the way routing tables are created to help in forwarding.

Direct Versus Indirect Delivery

- *Direct delivery* occurs when the source and destination of the packet are located on *the same physical network* or when the delivery is between the last router and the destination host.
- If the destination host *is not on the same network* as the deliverer, the packet is *delivered indirectly*.
- In an indirect delivery, the packet goes *from router to router until it reaches the one connected to the same physical* network as its final destination



a. Direct delivery



b. Indirect and direct delivery

Forwarding

- **Forwarding** means to place the packet in its route to its destination.
- Forwarding requires a host or a router to have a routing table.
- When a host has a packet to send or when a router has received a packet to be forwarded, it looks at this table to find the route to the final destination.
- several techniques are used to make the size of routing table manageable

Forwarding Techniques

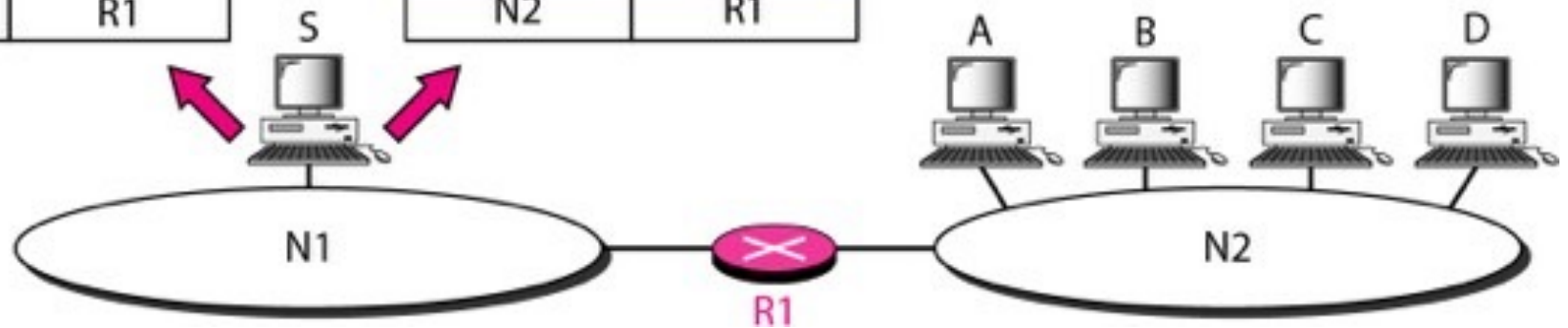
1. *Next-Hop Method (neighbour) Versus Route Method :*
routing table holds only the address of the next hop instead of information about the complete route(route method)
2. *Network-Specific Method Versus Host-Specific Method*
one entry that defines the address of the destination network

Routing table for host S based on host-specific method

Destination	Next hop
A	R1
B	R1
C	R1
D	R1

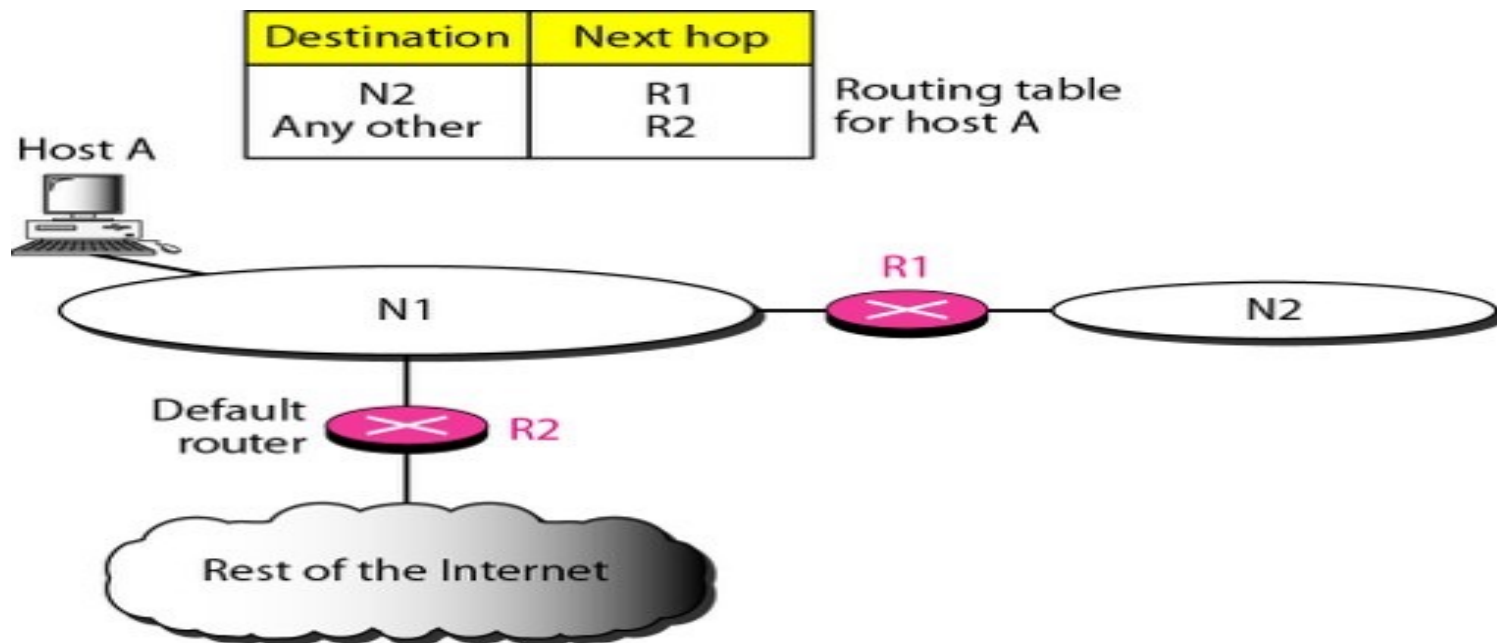
Routing table for host S based on network-specific method

Destination	Next hop
N2	R1



3. Default method

- host A can just have one entry called the default (normally defined as network address 0.0.0.0).



ROUTING ALGORIHMS

- **Main function of n/w layer** is routing packets from source to destination
- Algorithms that choose the routes are part of network layer design
- **Routing algorithm** - n/w layer software responsible for deciding on which output line, an incoming packet should be transmitted
- Grouped into 2 major classes
 1. Non adapitve
 2. Adaptive

Nonadaptive Algorithms(static routing)

- **Do not base** their routing decisions on measurements or estimates of **current traffic and topology**
- **choice of route to use** to get from I to J, is **computed in advance offline** and downloaded to the routers when the network is booted

Adaptive Algorithms(Dynamic routing)

- **Change their routing** decisions to reflect changes in **current traffic and topology**
- **They differ in where they get info, when they change routers, and what metric is used**

Routing table

- A host or a router has a routing table with an entry for each destination, or a combination of destinations, to route IP packets
- table can be either static or dynamic

Static Routing Table

- **contains information entered manually** by the administrator
- **it cannot update automatically when there is a change in the Internet.**
- The table must be **manually altered** by the administrator
- Small network
-

Dynamic Routing Table

- A dynamic routing table is updated periodically by using one of the dynamic routing protocols
 - Whenever there is a **change in the Internet, such as a shutdown of a router or breaking of a link**, the dynamic routing protocols **update all the tables in the routers** (and eventually in the host) automatically.
- Large network

Intradomain and Interdomain routing

- **Autonomous system** : group of networks and routers under the authority of a single administration
- routing inside an autonomous system - **intradomain routing**
 - e.g. distance vector routing and link state routing
 - RIP(Routing Information Protocol) is an implementation of distance vector
- routing between autonomous systems - **inter domain routing**
 - e.g. path vector routing
 - Open Shortest path First(OSPF) is an implementation of link state

Distance Vector Routing

- an adaptive routing algorithm
- Bellman-Ford Routing
- Ford Fulkerson Algorithm
- Eg.RIP(Routing Information Protocol)
- each **router maintain a table** giving the **best known distance** to each destination and which line to use to get there

- **Neighboring routers periodically exchange information from their routing tables.**

- **Routers replace routes in their own routing tables anytime that neighbors have found better routes.**

- Information provided from neighbours
 - **Outgoing line used for destination**
 - **Estimate of time or distance**
 - can be number of hops, time delay, packet queue length, etc.

Distance Vector Routing

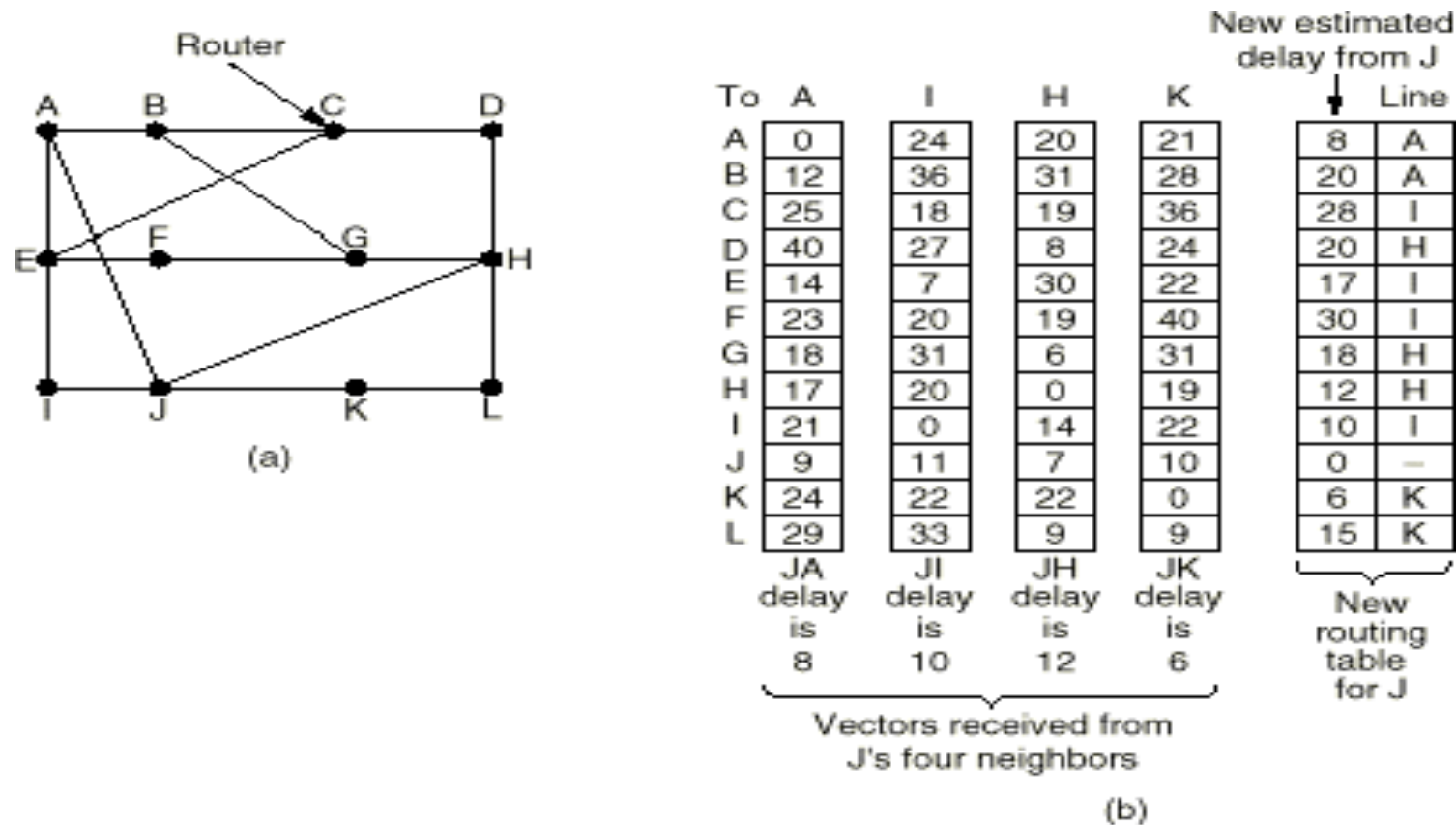


Fig. 5-10. (a) A subnet. (b) Input from A, I, H, K, and the new routing table for J.

Shortest Path Routing (**Dijkstra algorithm**)

- a **nonadaptive routing** algorithm
- Shortest path algorithm first developed by E. W. **Dijkstra**
- Find the **shortest path from a specified source to all other destinations in the network.**
- The **subnet** is represented as a **graph** with **nodes** in the network as **nodes** of the graph and **links as arcs** of the graph
- Given a network topology and a set of weights describing the cost to send data across each link in the network

- Mark the source node as **permanent**.
- Designate the source node as the **working node**.
- Set the **tentative** distance to all other nodes to **infinity**.
- While some nodes are not marked permanent
 - Compute the tentative distance from the source to all nodes adjacent to the working node. If this is shorter than the current tentative distance replace the tentative distance of the destination and record the label of the working node there.
 - Examine ALL tentatively labeled nodes in the graph. Select the node with the smallest value and make it the new working node. Designate the node permanent.

Example of Shortest Path Routing

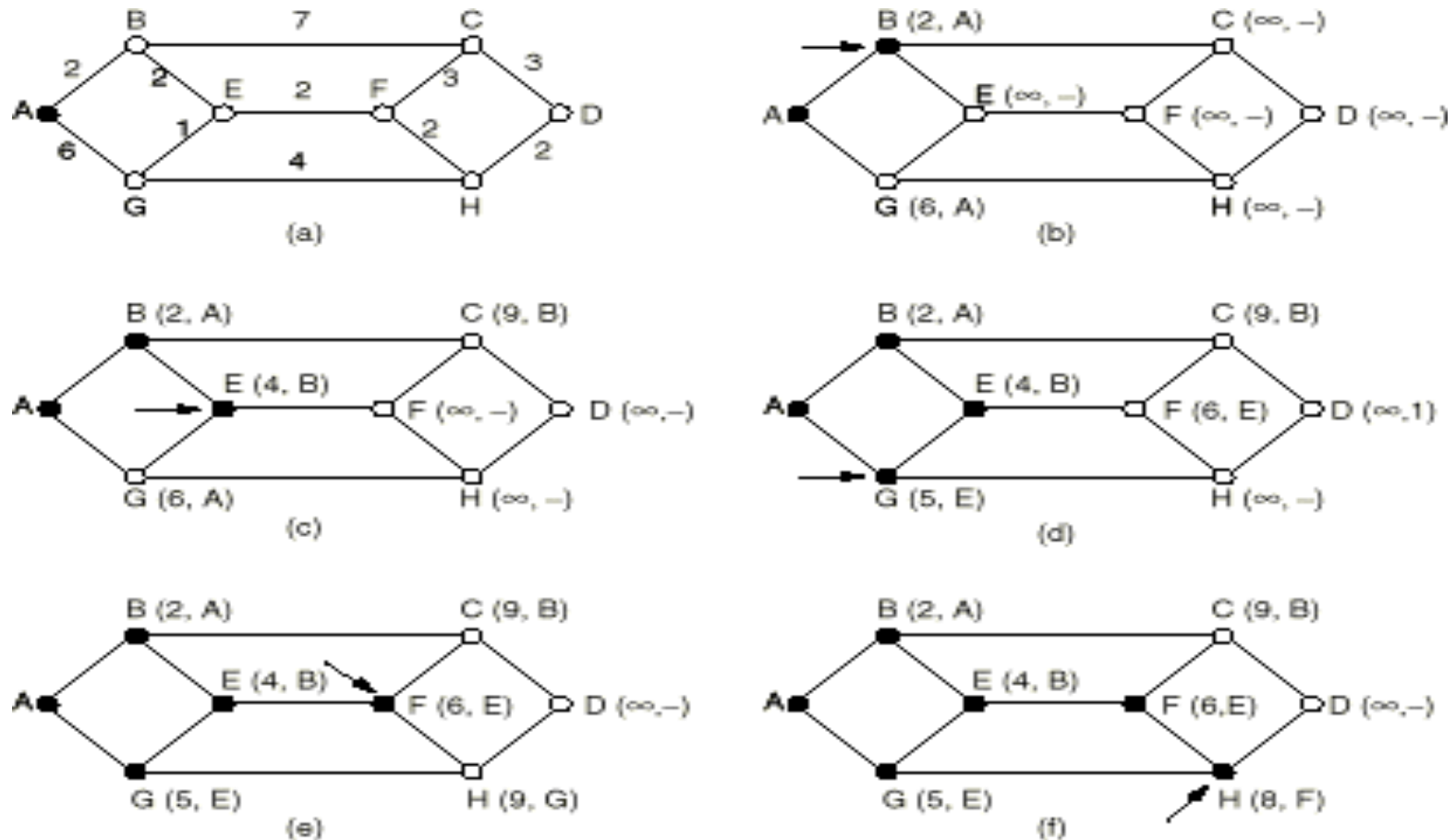


Fig. 5-6. The first five steps used in computing the shortest path from A to D. The arrows indicate the working node.

Link State Routing

(an adaptive routing algorithm)

- OSPF (open shortest path first)
- In link state routing, if **each node in the domain** has the **entire topology of the domain** the list of nodes and links, how they are connected including the type, cost (metric), and condition of the links (up or down)-the node can **use Dijkstra's algorithm to build a routing table**
- **Each node uses the same topology to create a routing table, but the routing table for each node is unique** because the calculations are based on different interpretations of the topology

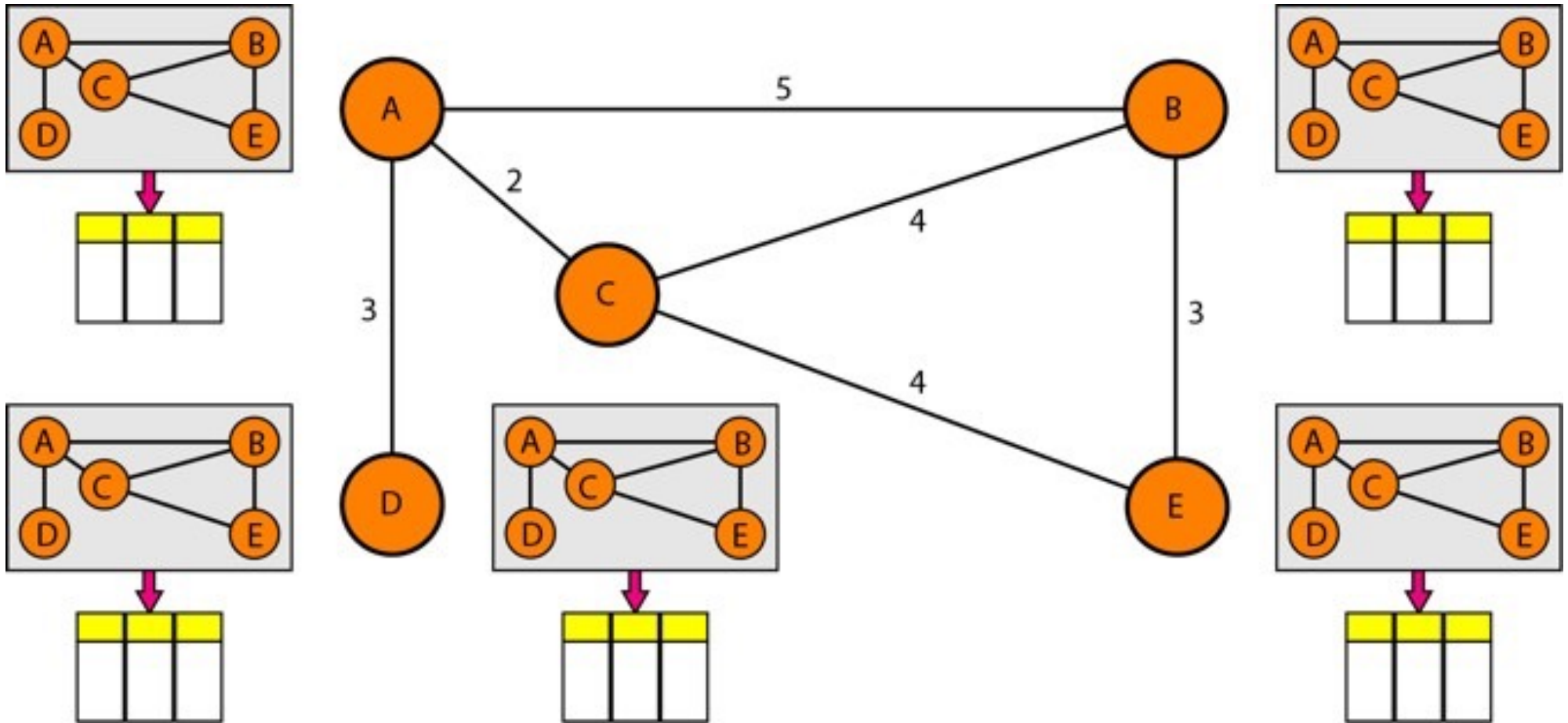
Building Routing Tables

- **four sets of actions** are required to ensure that each node has the routing table showing the **least-cost node to every other node**.
 1. Creation of the states of the links by each node, called **the link state packet (LSP)**.
 2. Dissemination of LSPs to every other router, called **flooding**, in an efficient and reliable way.
 3. **Formation of a shortest path tree (spanning tree)** for each node.
 4. **Calculation** of a routing table based on the shortest path tree.

Creation of Link State Packet (LSP)

- LSP carries a minimum amount of data: **the node identity, the list of links, a sequence number, and age.**
- The first **two, node identity and the list of links**, are needed to make the topology.
- The third, **sequence number**, facilitates flooding and distinguishes new LSPs from old ones.
- The fourth, **age**, prevents old LSPs from remaining in the domain for a long time.

Build Link State Packets



LSPs are generated on two occasions:

- When there is a change in the topology of the domain.
- On a periodic basis.
 - It is done to ensure that old information is removed from the domain.

Flooding of LSPs

- After a node has prepared an LSP, it must be disseminated to all other nodes, not only to its neighbors.
- The process is called **flooding** and based on the following:
 - The creating node sends a copy of the LSP out of each interface
 - A node that receives an LSP compares it with the copy it may already have.

If the newly arrived LSP is older than the one it has (found by checking the sequence number), it discards the LSP. If it is newer, the node does the following:

- a. It **discards the old LSP and keeps the new one.**
- b. It **sends a copy of it out of each interface except the one from which the packet arrived.**

Formation of Shortest Path Tree: After receiving all LSPs, each node will have a copy of the whole topology.

- However, the topology is not sufficient to find the shortest path to every other node; a shortest path tree is needed.
- A **tree** is a graph of nodes and links; one node is called the **root**. All other nodes can be reached from the root through only one single route.
- A **shortest path tree** is a tree in which the path between the root and every other node is the shortest.
- **for each node, there is a shortest path tree with that node as the root.**
- The **Dijkstra algorithm** creates a **shortest path tree** from a graph

Calculation of Routing Table from Shortest Path Tree

- Each node uses the shortest path tree protocol to construct its routing table.
- The routing table shows the cost of reaching each node from the root

<i>Node</i>	<i>Cost</i>	<i>Next Router</i>
A	0	—
B	5	—
C	2	—
D	3	—
E	6	C

THANK YOU