Module 1

Introduction to Computer Networks

Dr. Sunandita Debnath, IIIT Vadodara

Evaluation Policy

Evaluation Breakdown

1. **Mid-Term Assessment - 30%**

a. Pre-Mid-Sem Examination (**10%**)

b. Online Mid-Sem Examination (**10%**)

c. Remote Mid-Sem Examination (**10%**)

2. **End-Term Assessment – 45%**

a. Pre-End-Sem Examination (**15%**)

b. Online End-Sem Examination (**15%**)

c. Remote End-Sem Examination (**15%**)

3. Other components (Assignments, Quizzes and Viva) - 25%

Course Content of Computer Networks

MODULE 1

Overview of an internet, internet as a service, internet architecture, circuit switching, packet switching, network performance metrics (delay, packet loss, and throughput), layered approach (TCP/IP and OSI models)

Course Instructor
Dr. Sunandita Debnath
Assistant Professor, IIIT Vadodara
Contact No. 9508370853, WhatsApp No. 9485145373
Google classroom Code: d3w7vqi

Reference Books

1. Computer Networking: A Top-Down Approach (Fifth Ed. by J. F. Kurose and K. W. Ross, publisher: Pearson

2. Data Communications and Networking (Fourth Ed. by B. A. Forouzan, publisher: McGraw Hill Education

3. Computer Networks (Fifth Ed. by A. S. Tanenbaum and D. J. Wetherall, publisher: Pearson

What is Internet?

The internet is a computer network that interconnects ***zillions*** *of computing devices throughout the world.*

"Internet is a network of networks"

Computer networks

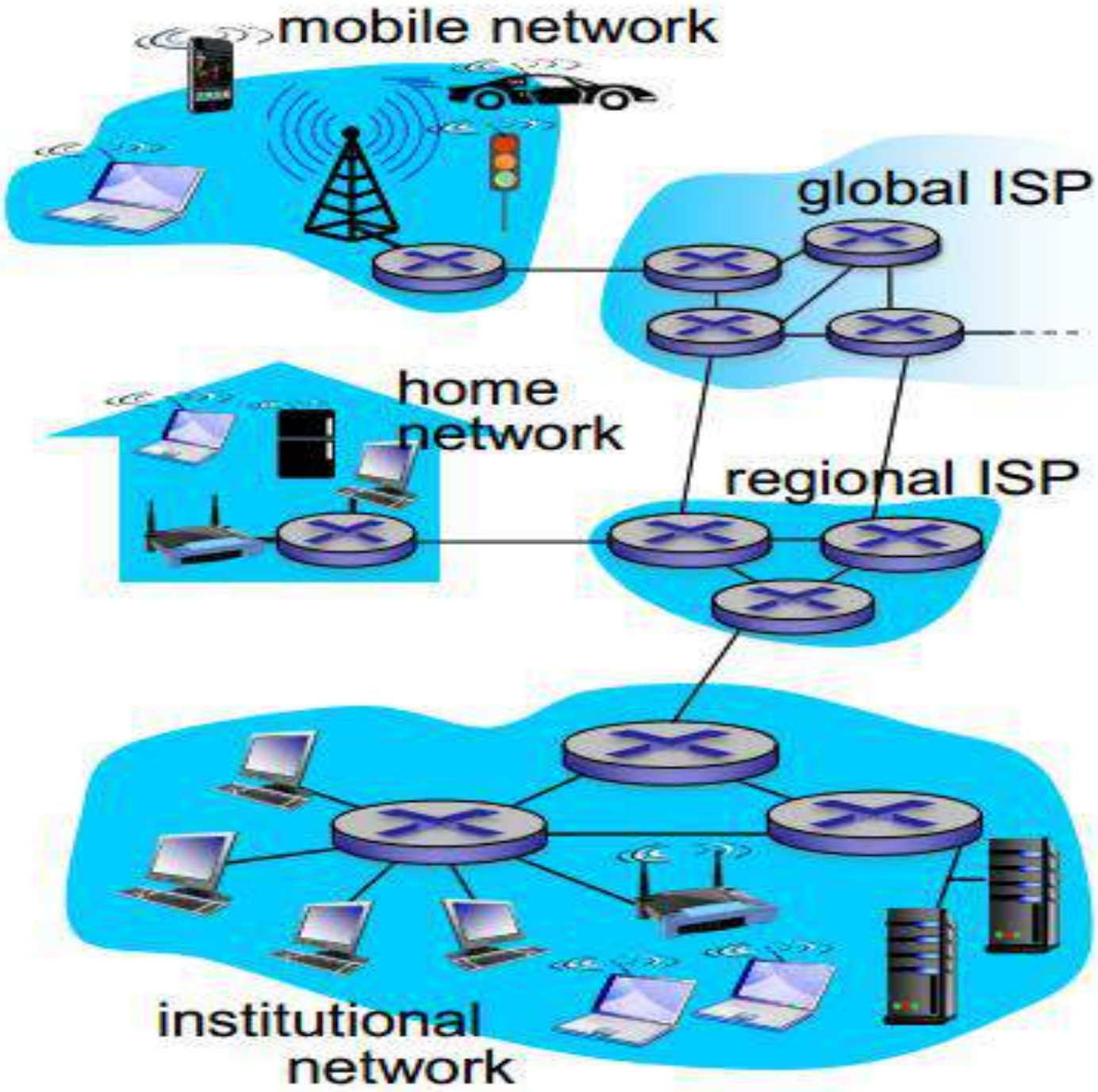
- *A computer network is a set of nodes or specifically end systems connected by communication links .*

- **Three important terms in computer networking is:**
- *Nodes*
- *Communication links*
- *Packet switches*

“Internet is everywhere”

Application includes:

- *Electronics mail*
- *Web surfing*
- *Social networks*
- *Instant Messaging*
- *Voice over IP (VoIP)*
- *Distributed games*
- *Peer-to-peer*
- *File sharing*
- *Television over internet*
- *Remote login*


mobile network
global ISP
home network
regional ISP
institutional network

Nodes

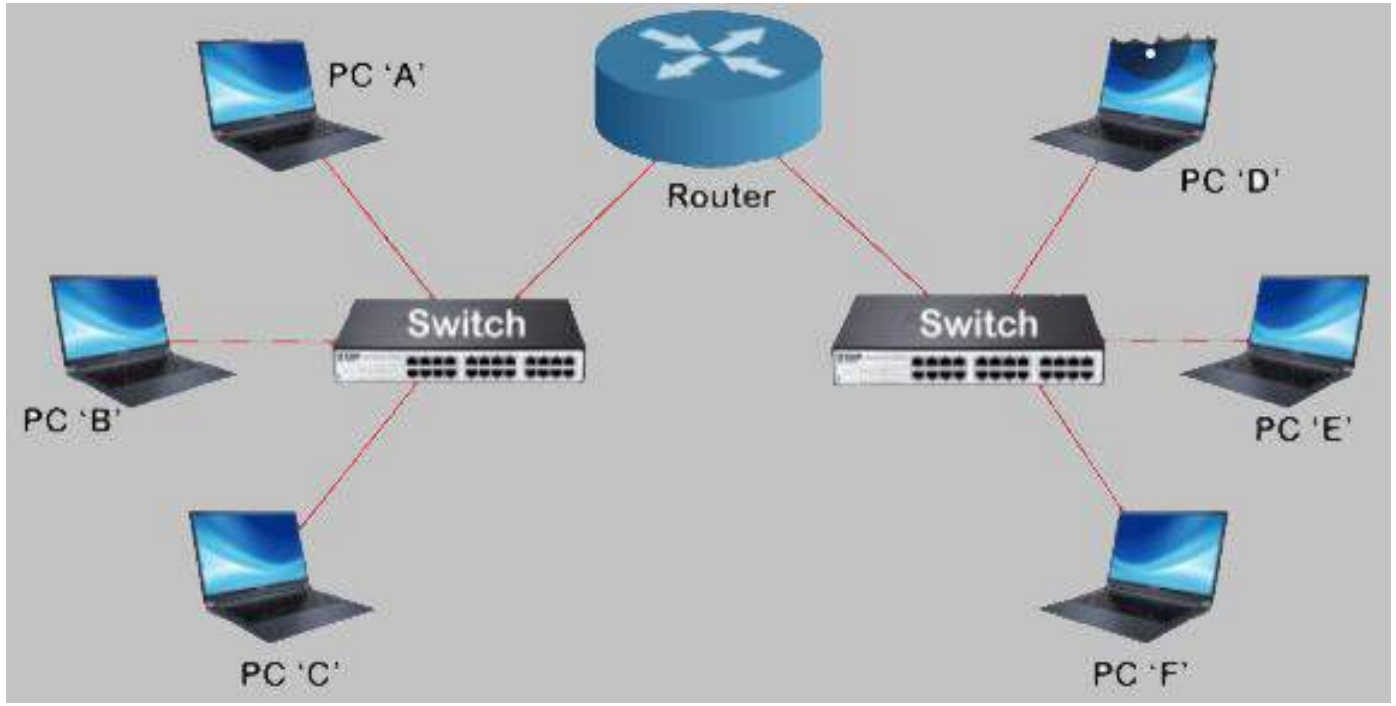
End Nodes/Host

A end node can be a computer, printer, mobile phones, PDAs (Personal Digital Assistants) or any other devices capable of sending/ receiving data generated by the other nodes in the network.

Intermediary nodes:

A intermediary node can be a hub, cell phone tower, repeaters, routers, switches etc.

- **Important terms in computer networking is:**
- *Hosts/End users*
- *Communication links (Wired or Wireless)*
- *Routers*
- *Switches*
- *Hub*
- *Servers*
- *Client*



- **Intermediary Devices**
- *Routers*
- *Switches*
- *Hub*

Network	**Routers**
Data Link	**Switches**
Physical	**Hubs**

Communication links

A commination link can be **wired** *or* **wireless** *links. This link or channel carries information.*

- **Co-axial cables, Optical Fiber, and RF spectrum (air)**

Wired Media

- Ethernet straight (for different type of devices)
- Ethernet cross (for same type of devices)
- Co-axial cables
- Optical Fibers
- USB cables

Wireless Media

- Infrared (for short rang e.g. TV/AC remote controls)
- Radio (e.g. Bluetooth, WiFi)
- Microwave links (e.g. Cell phone tower)
- Satellite links (for GPS)

Packet switches

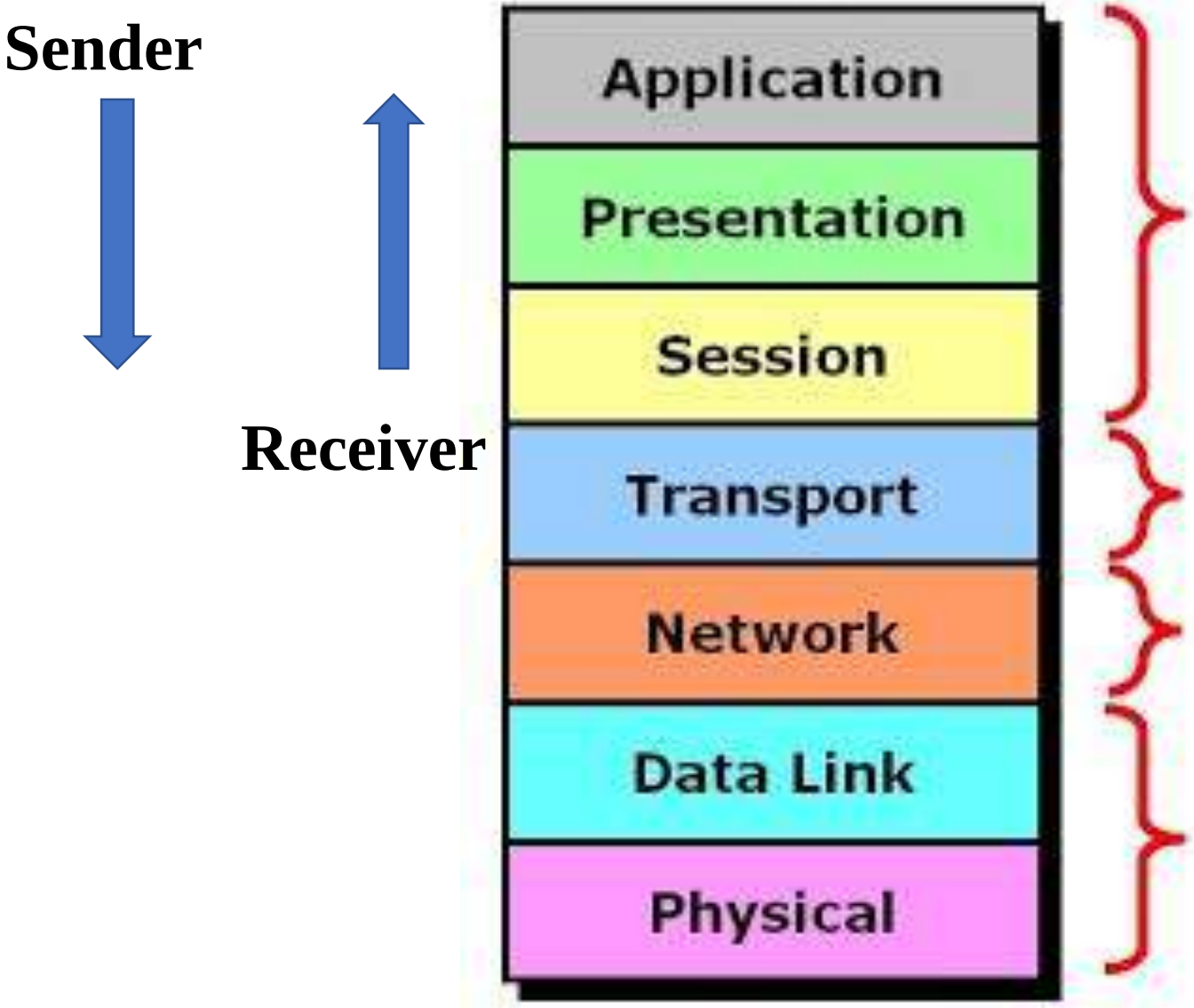
Packet switches receives a packet arriving on one of its incoming communication links and forwards that packet on one of its outgoing communication links.

- **Routers** *(Core Networks)*
- **Link layer switches** (*Access Networks*)

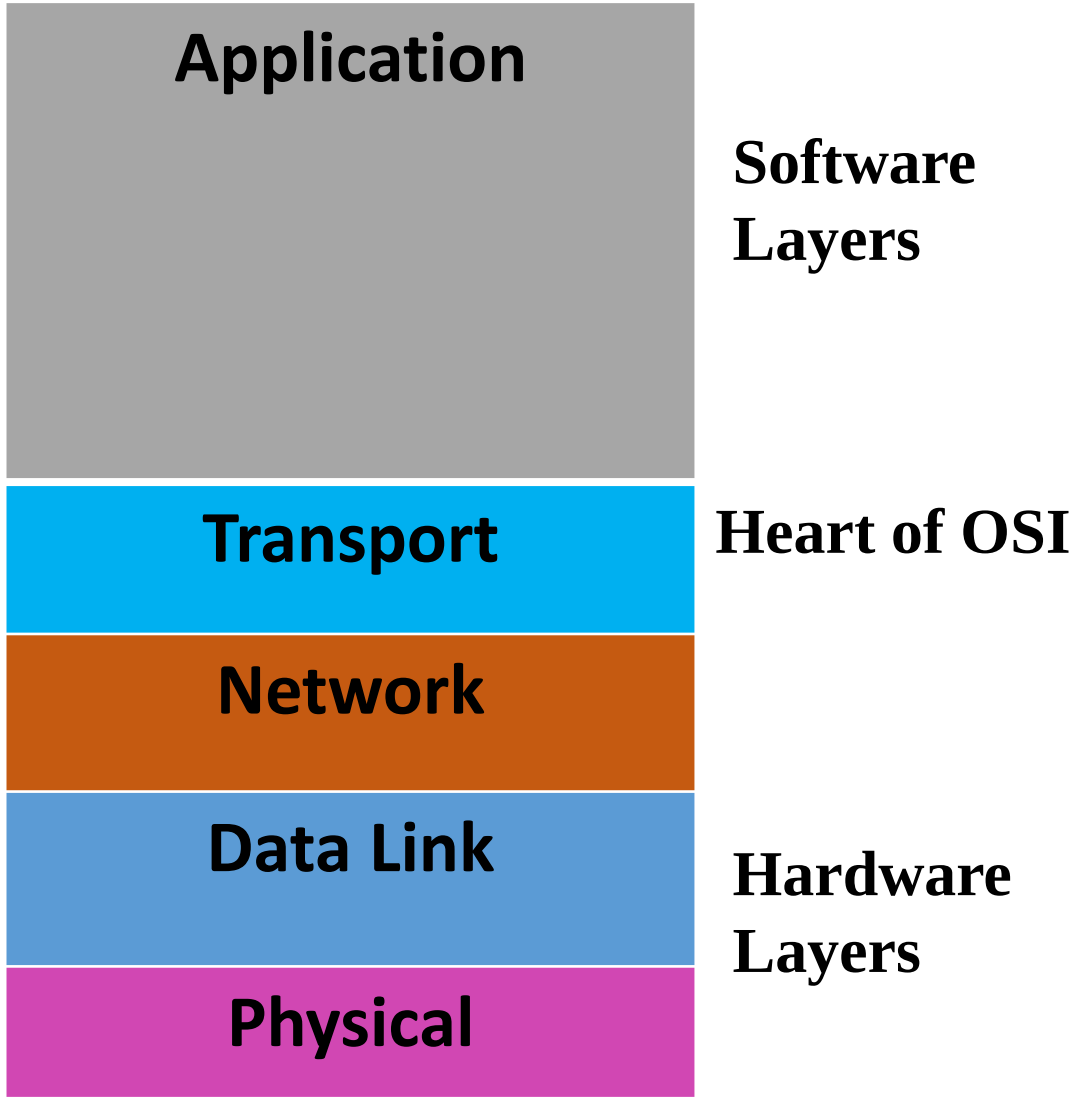
End systems access the internet through ISPs (Internet service providers). ISPs can be different type:

- **Residential ISPs (e.g. Local cables or Telephone companies)**
- **Corporate ISPs (e.g. Offices)**
- **University and Colleges ISPSs (IIIT Vadodara)**
- **ISPs that provides WiFi access in airports, coffee shops and railway stations and other public places**

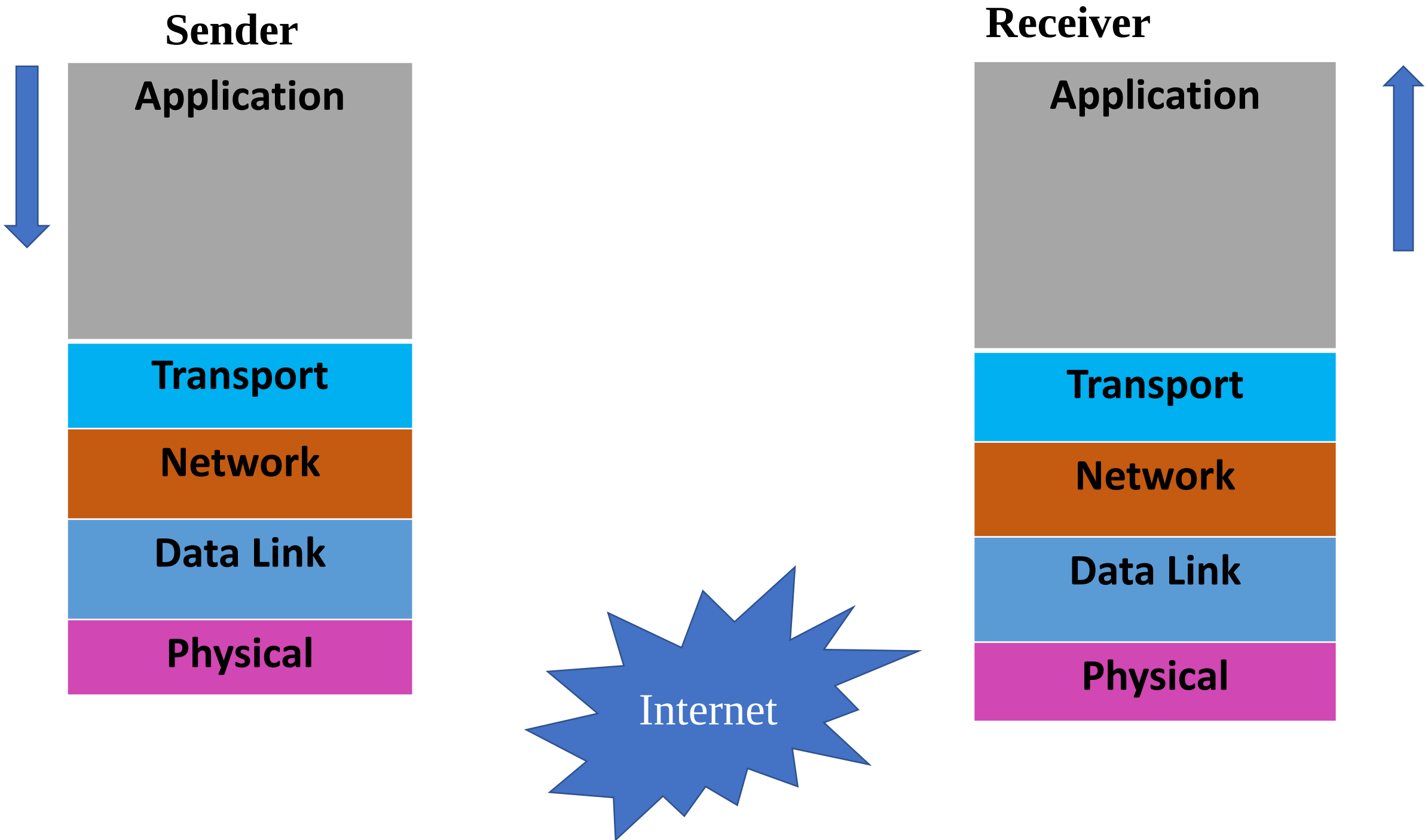
❖ *Each ISP n itself a network of packet switches and communication links.*



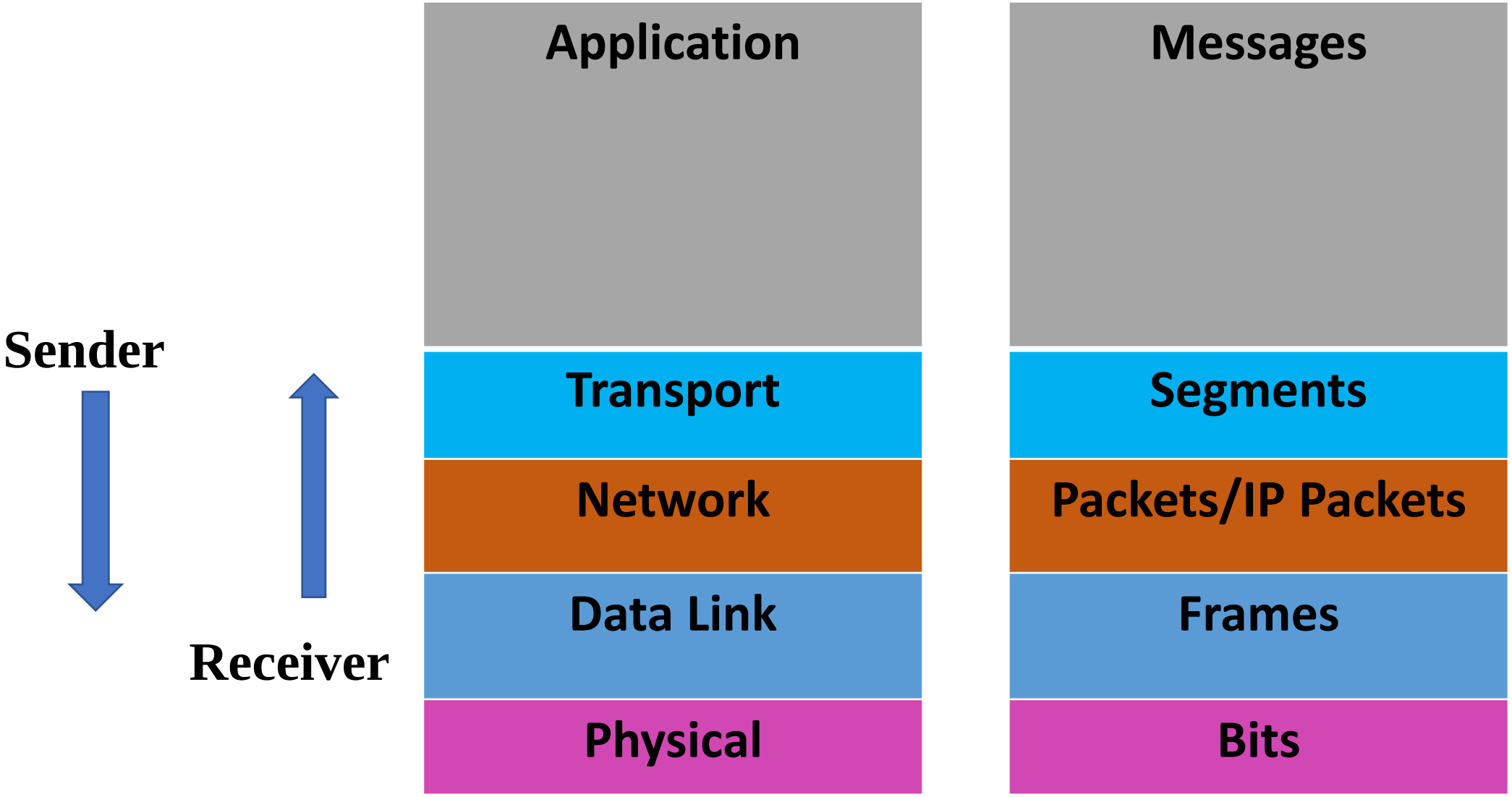
OSI Layers



TCP/IP Layers


Sender
Application
Transport
Network
Data Link
Physical
Internet
Receiver
Application
Transport
Network
Data Link
Physical

Packet data Units (PDUs)



TCP/IP Layers

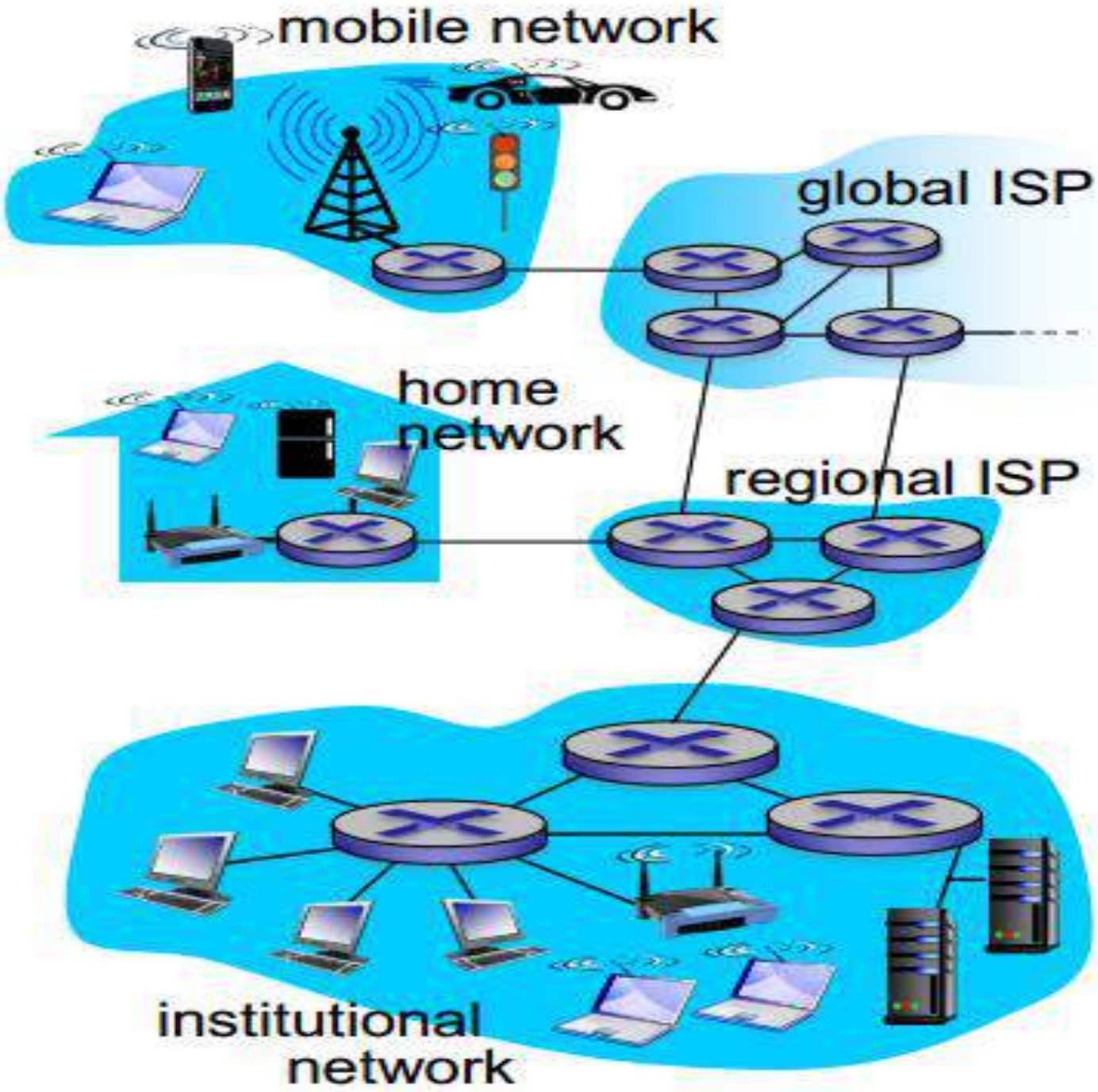
- **Intermediary Devices**

Network	**Routers: Data-Packets, Devices: between Intelligent devices, Full duplex, Addressing Type: IP address, Between two different networks, Routing Device**
Data Link	**Switches: frames, Devices: between Intelligent devices, Half/Full duplex, Addressing Type: MAC address, Between devices on same network, Multicasting Device**
Physical	**Hubs: Data-Electrical signals/bits, Devices: between Non-intelligent devices, Half duplex, Addressing Type: MAC address, Between devices on same network, Broadcasting device**

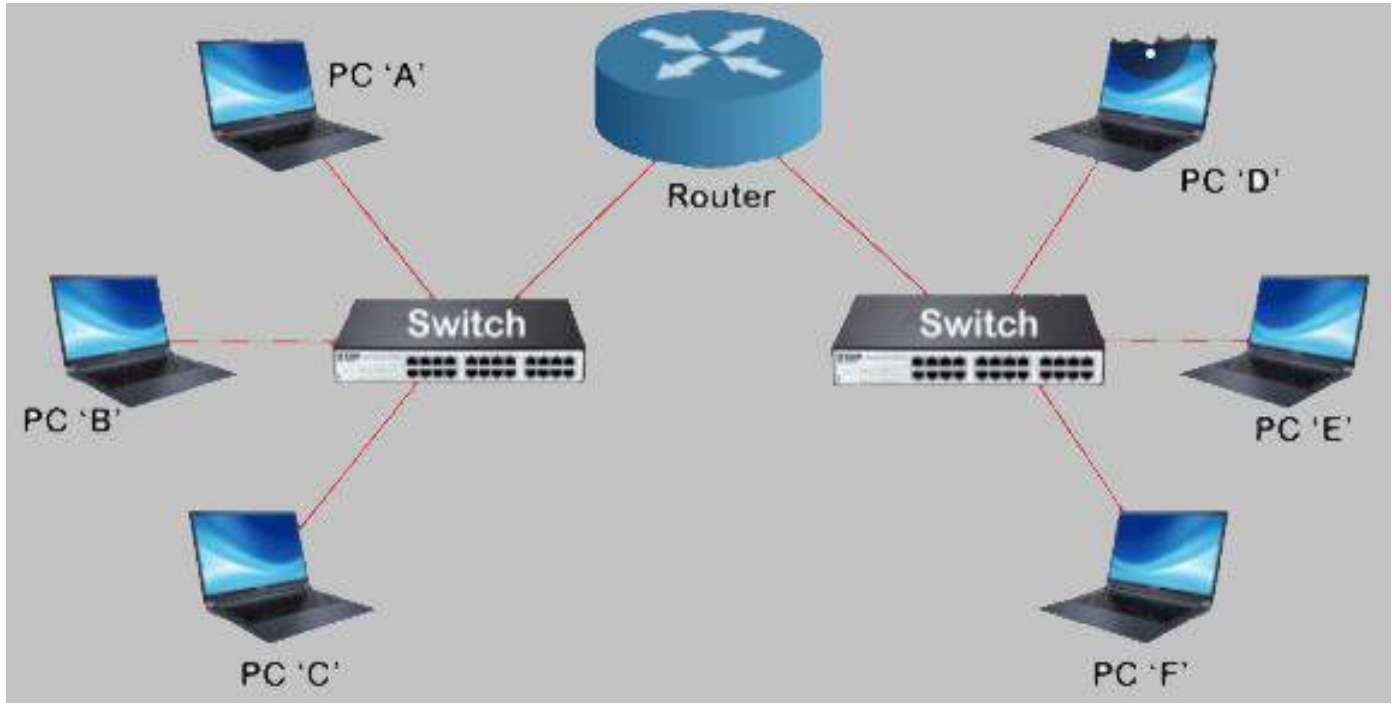
Module 1

Introduction to Computer Networks

Dr. Sunandita Debnath, IIIT Vadodara


mobile network
global ISP
home network
regional ISP
institutional network

- **Intermediary Devices**
- *Routers*
- *Switches*
- *Hub*



Network	**Routers**
Data Link	**Switches**
Physical	**Hubs**

- **Intermediary Devices**

Network	**Routers: Data-Packets, Devices: between Intelligent devices, Full duplex, Addressing Type: IP address, Between two different networks, Routing Devices**
Data Link	**Switches: frames, Devices: between Intelligent devices, Half/Full duplex, Addressing Type: MAC address, Between devices on same network, Multicasting Device**
Physical	**Hubs: Data-Electrical signals/bits, Devices: between Non-intelligent devices, Half duplex, Addressing Type: MAC address, Between devices on same network, Broadcasting device**

Communication links

A commination link can be **wired** *or* **wireless** *links. This link or channel carries information.*

- **Co-axial cables, Optical Fiber, and RF spectrum (air)**

Wired Media

- Ethernet straight (for different type of devices)
- Ethernet cross (for same type of devices)
- Co-axial cables
- Optical Fibers
- USB cables

Wireless Media

- Infrared (for short rang e.g. TV/AC remote controls)
- Radio (e.g. Bluetooth, WiFi)
- Microwave links (e.g. Cell phone tower)
- Satellite links (for GPS)

Packet switches

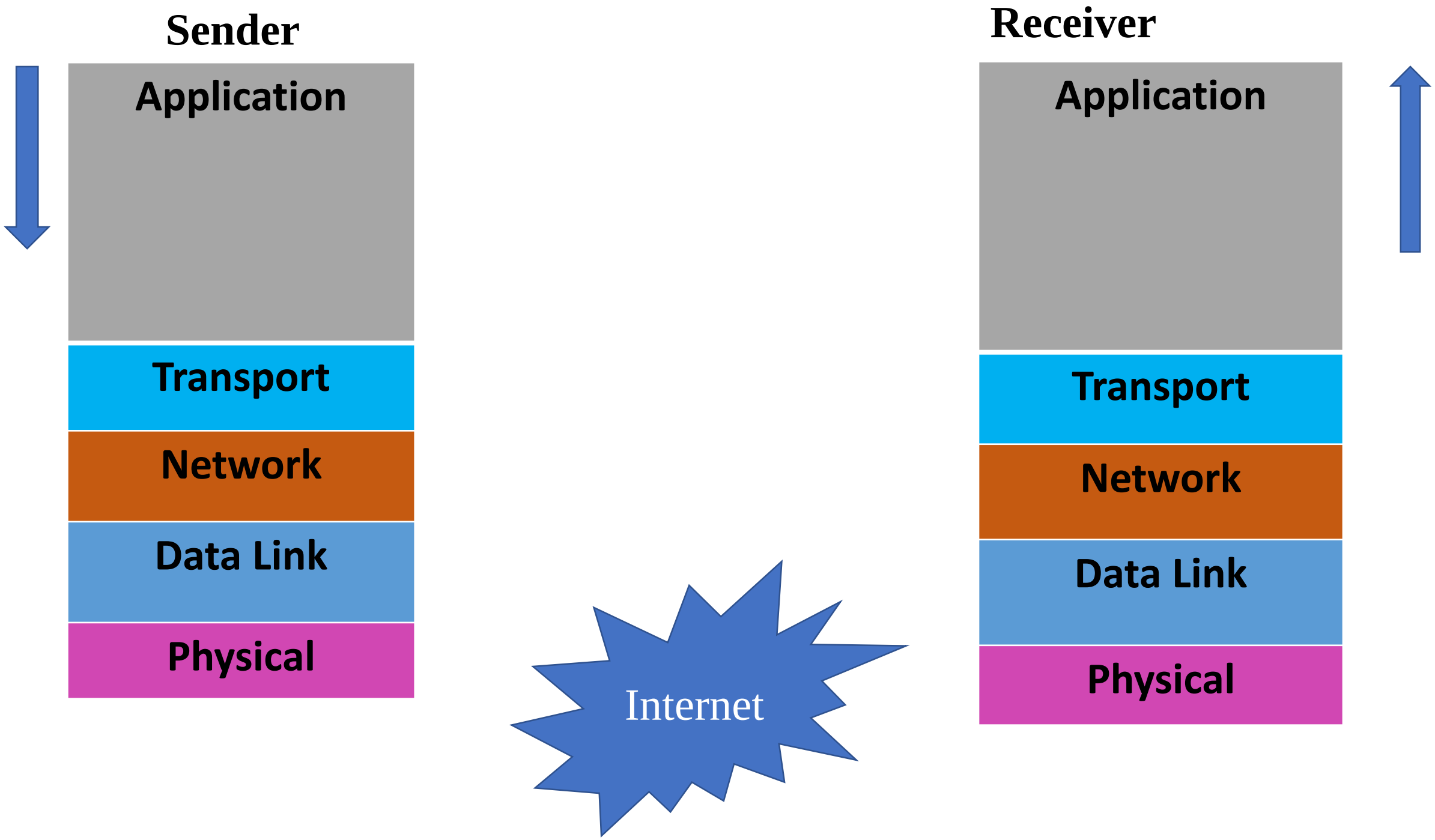
Packet switches receives a packet arriving on one of its incoming communication links and forwards that packet on one of its outgoing communication links.

- **Routers** *(Core Networks)*
- **Link layer switches** (*Access Networks*)

End systems access the internet through ISPs (Internet service providers). ISPs can be different type:

- **Residential ISPs (e.g. Local cables or Telephone companies)**
- **Corporate ISPs (e.g. Offices)**
- **University and Colleges ISPSs (IIIT Vadodara)**
- **ISPs that provides WiFi access in airports, coffee shops and railway stations and other public places**

❖ *Each ISP n itself a network of packet switches and communication links.*


Sender
Application
Transport
Network
Data Link
Physical
Internet
Receiver
Application
Transport
Network
Data Link
Physical

Peer to Peer Network	Client server network
❑ No centralized administration ❑ All peers are equal ❑ Not scalable ❑ Simple sharing applications	❑ Centralized administration (one will be master and others will be slave) ❑ Request response model ❑ Scalable ❑ Server may be overloaded

Data communication

Data communication is the exchange of data between two nodes via some from of links such as cable.

Data flow in this data communication can be of three types

- **Simplex (E.g. Keyboard, Printers etc. one way communication)**
- **Half-Duplex (E.g. Walkie-Talkies. Two way communication but both not simultaneously)**
- **Duplex (E.g. Telecommunication. Two way communication at the same time)**

Protocols

Protocol is set of rules which governs all types of communication, like

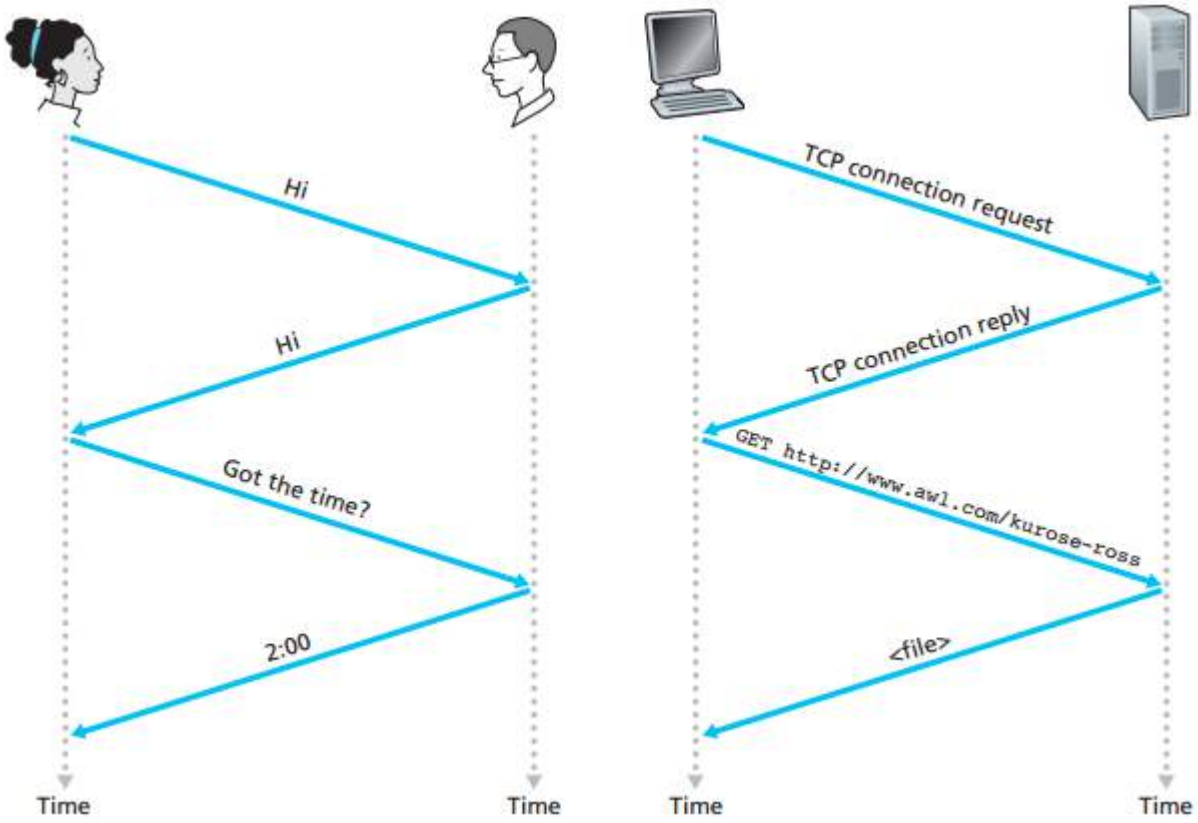
What is communicated?

How it is communicated?

When it is communicated?

Protocol

- *This is comparison of a human protocol and computer network protocol*



Important points:

- *If the system runs two different protocols then the protocols do not inter operate and no useful work can be accomplished.*
- *The two system should run the same protocol in order to accomplish the task.*

- ***A protocol defines the format and the order of message exchanged between two or more communicating entities, as well as the action taken on the transmission or reception of a message or other event.***

- ***A protocol defines the format and the order of message exchanged between two or more communicating entities, as well as the action taken on the transmission or reception of a message or other event.***

Classification of computer networks:

- ***LAN (Local Area Network)***
- ***MAN (Metropolitan Area Network)***
- ***WAN (Wide Area Network)***

Access Network	Core Network
❑ The access network is the network or part of communication systems which provide the user access to the internet services. ❑ Typically connecting devices like phones and laptops to a access point in a café or home network.	❑ The core network is the part of a network that connects different access networks. ❑ Typically cover wide ranges like network connecting two different cities.

LAN	MAN	WAN
❑ LAN (Local Area Network) is a computer network covering a small geographic area, like a home, office, school, or group of buildings, University campus. ❑ High speed (1000 Mbps) ❑ High bandwidth as lesser number of devices are connected. ❑ Range can be almost kms.	❑ A metropolitan area network (MAN) is a network with a size bigger than a LAN but smaller than a WAN. It normally covers the area inside a town or a city. ❑ moderate speed(44 to 155 Mbps) ❑ Less bandwidth than LAN as more devices are connected. ❑ 100 to 1000 of kms.	❑ WAN (Wide Area Network) is a computer network that covers a broad area e.g. country or continent. ❑ Less speed (150 Mbps) ❑ Low bandwidth as many MAN are connected. ❑ Satellite is used to manage WAN.

Module 1

Introduction to Computer Networks

Dr. Sunandita Debnath, IIIT Vadodara

Functions of each layer

Application Layer: *It enables the users to access the network resources. This layer also serves as a window for the application services to access the network and for displaying the received information to the user. It is also called as desktop layer.*

- *Mail services*
- *Directory services*
- *File transfer and Access management*

Protocols: *HTTP, SMTP, FTP, TFTP, DNS, DHCP*

Presentation Layer

It is concerned with the syntax and semantics of the information exchanges between the two systems. It is also called the as translation layer.

- *Translation (converting data into common format)*
- *Encryption (security purpose)*
- *Compression (compressing long duration videos and high resolution images and videos)*

Functions of each layer

Session Layer: *It establishes the connection and maintains the session between the communicating devices. It is also responsible for authentication and security.*

- *Dialog control*
- *Synchronization*
- *Session establishment, maintenance and termination*

Transport Layer: *It is responsible for process-to-process or end-to-end delivery of entire message/data. It provides services to the application layer and takes services from the network layer. This layer is responsible for acknowledgement of successful data transmission and re-transmit the data if error is found. Transport layer adds source and destination port numbers to the segmented data.*

- *Port addressing*
- *Segmentation and reassembly*
- *Connection control*
- *End to end flow control*
- *Error control*

Protocols: *TCP (connection oriented, used for reliable data transmission), UDP (connection less, used for real time traffic)*

Functions of each layer

Network Layer: *It is responsible for delivery of data from original source to destination with the help of IP addresses. It also takes cares of packet routing i.e. selection of the shortest route/path to transmit the packet, from the number of routes available.*

- *Logical addressing*
- *Routing*

Protocols: *ICMP (Internet Control Message Protocol), IPV4/IPV6 (Internet Protocol Version 4, Version 6, RIP (routing Information Protocol)*

Data link Layer: *It is responsible for the node-to-node delivery of frames. The main function of this layer is to make sure that the data transfer is error-free from one node to other, over the physical layer. This layer is divided into LLC(Logical Link Control) and MAC (Media Access Control).*

- *Framing*
- *Physical addressing*
- *Flow control*
- *Error control*
- *Access control*

Protocols: *LLDP, PPP (Point-to-Point), Spanning tree protocol*

Functions of each layer

Physical Layer: *It is responsible for the actual physical connection between the devices , the information are in form of bits. It also provides electrical and mechanical specifications.*

- *Physical characteristics of the medium (either wired or wireless)*
- *Synchronization of bits*
- *Bit rate control*
- *Physical topologies (star, bus or mesh connection)*
- *Transmission mode (duplex, half duplex, full duplex)*

IP Address	MAC Address
❑ Every node in the computer network is identified with the help of IP address or logical address. ❑ Can change based on the location of the device ❑ Assigned manually or dynamically. ❑ Represented in decimal and it has 4 octets (x.x.x.x). ❑ 0.0.0.0 to 255.255.255.255. (each octet is represented by 8 bit representation) total 32 bits long. ❑ Routers need IP address to forward data	❑ Every node in LAN is identified with the help of MAC address. ❑ Physical address or hardware address imprinted in the NIC (*Network Interface Controller*)card of the device. ❑ Cannot be changed it is a unique number, assigned by the manufacturer of the device. ❑ Represented in Hexadecimal. ❑ E.g. 70-20-84-00-ED-FC (48bits) ❑ Separated by, hyphen (-), dot(.), colon (:). ❑ Switches need MAC address to forward data.

Module 1

Introduction to Computer Networks

Dr. Sunandita Debnath, IIIT Vadodara

Network structure

Network Edge

- ***Access Networks***
- ***Physical Media***

Network core

- ***Packet switching***
- ***Circuit Switching***

Network structure

Access Networks-*The network that physically connects an end system to the first router (also known as edge router) on a path from the end system to any other distant end system.*

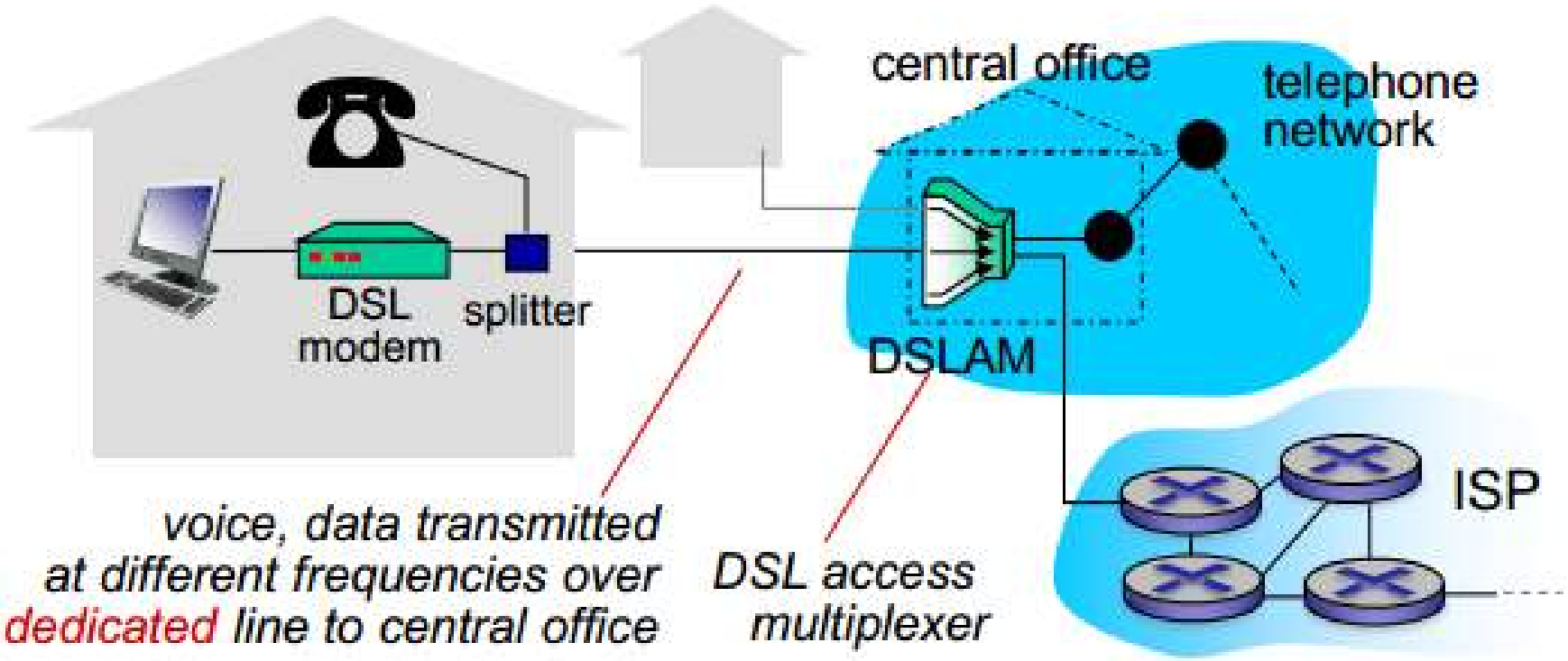
There are three types of access networks

- ***Home access***
- ***Enterprise access***
- ***Wide area mobile wireless***

Home Access structure

- *DSL (Digital subscriber line), Cable, FTTH (Fiber to the Home), Dial-Up*

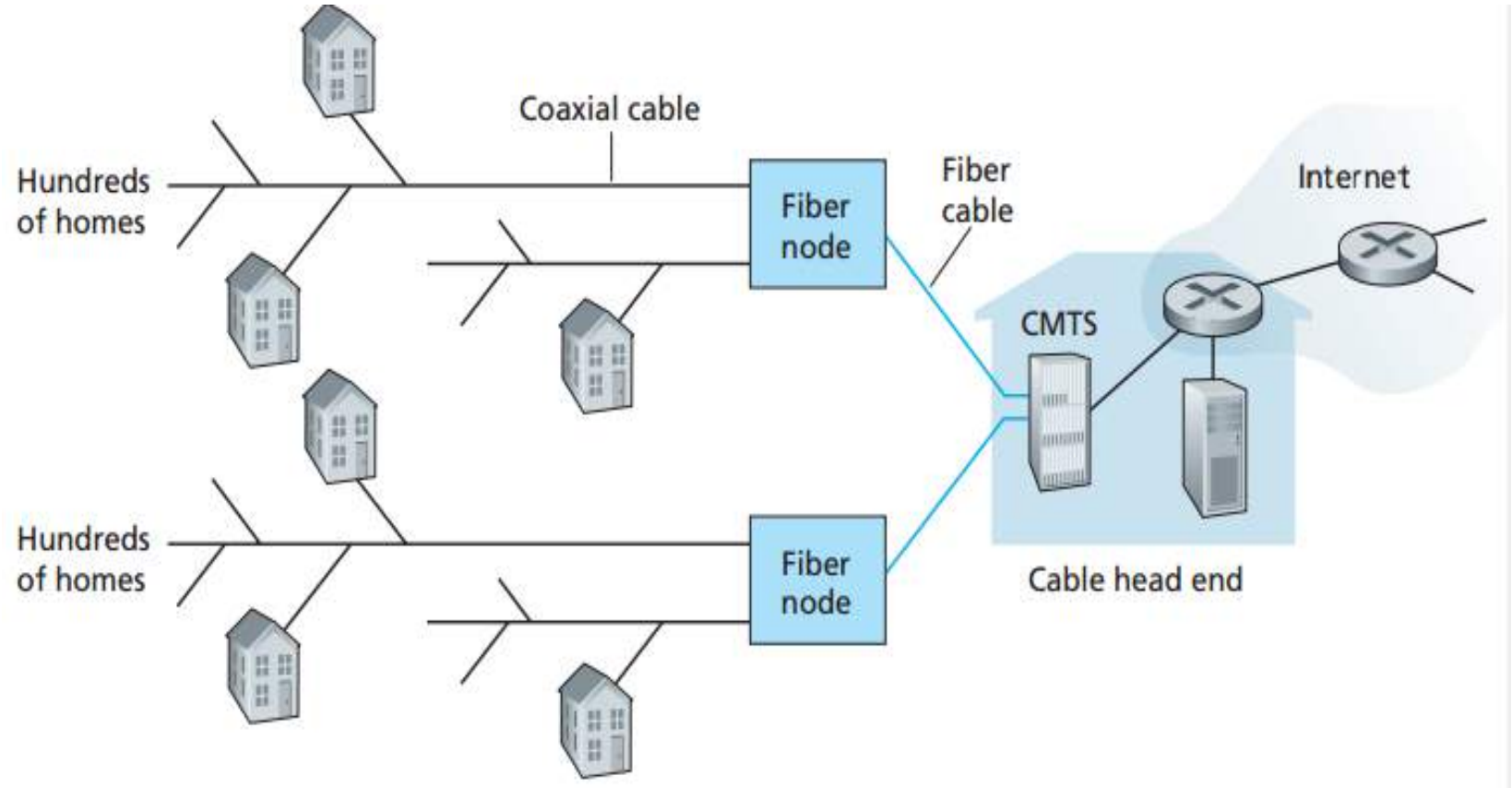
- *Each customer's DSL modem uses the existing telephone line (twisted wire) to exchange data with a DSLAM located in the telecom Central Office (CO).*
- *The home's DSL modem takes digital data and translated it to high frequency tones for transmission over telephone wires to CO.*



- *The analog signal from many such houses are translated back into digital format at the DSLAM.*

A Hybrid Fiber co-axial access network

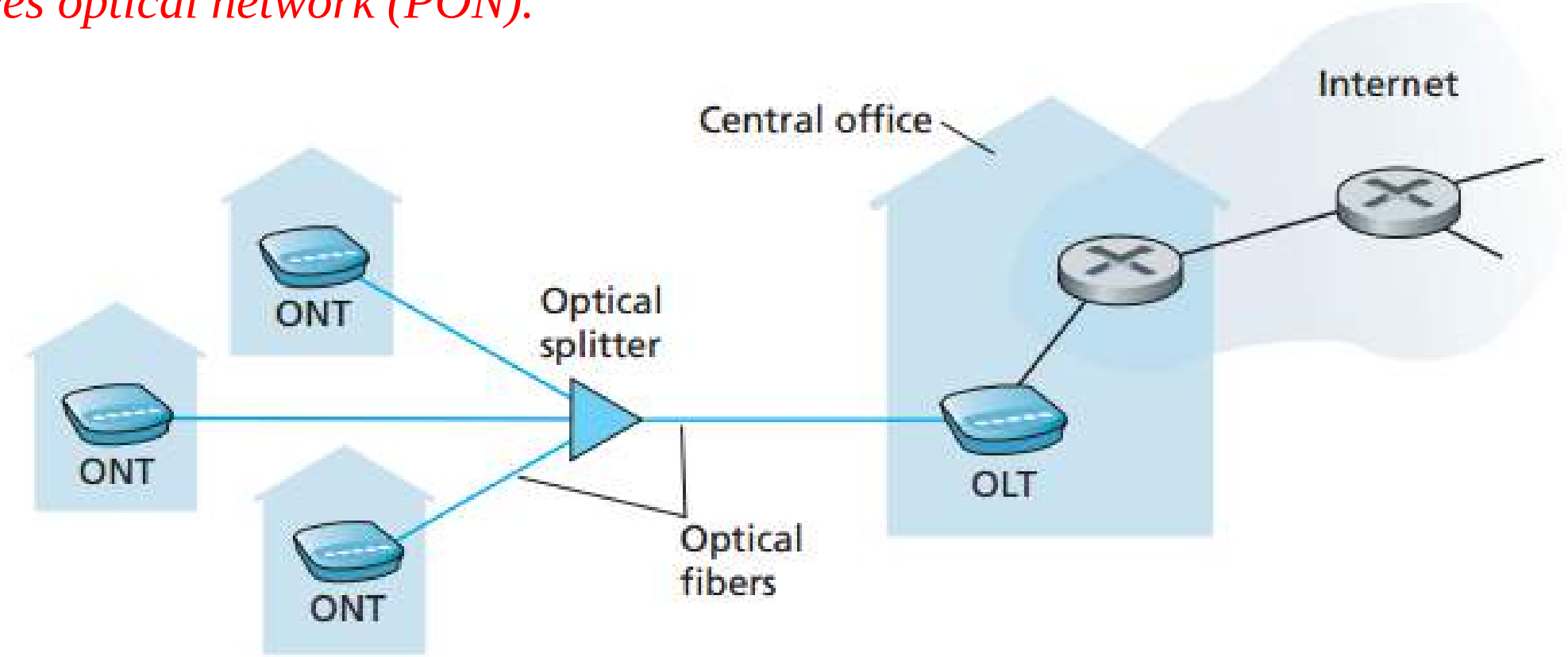
- *The fiber optics connect the cable head end to the neighbourhood level junctions, from which the traditional co-axial cable is then used to reach individual homes and apartments. Each neighbourhood junction typically supports 500 to 5000 homes.*
- *As both fiber optics and co-axial cables are involved it is often referred to as Hybrid Fiber Coax.*



- *At the cable head end, the CMT (Cable Modem Terminator Systems) serves a similar function as DSLM in the DSL networks-such as turning the analog signal sent from from several cable modems in many downstream home back into digital format.*

Fiber to The Home (FTTH)

- *It provides an dedicated optical fiber path from the central office (CO) directly to home.*
- *There are two types of Optical distribution network: Active optical network (AON) and Passives optical network (PON).*



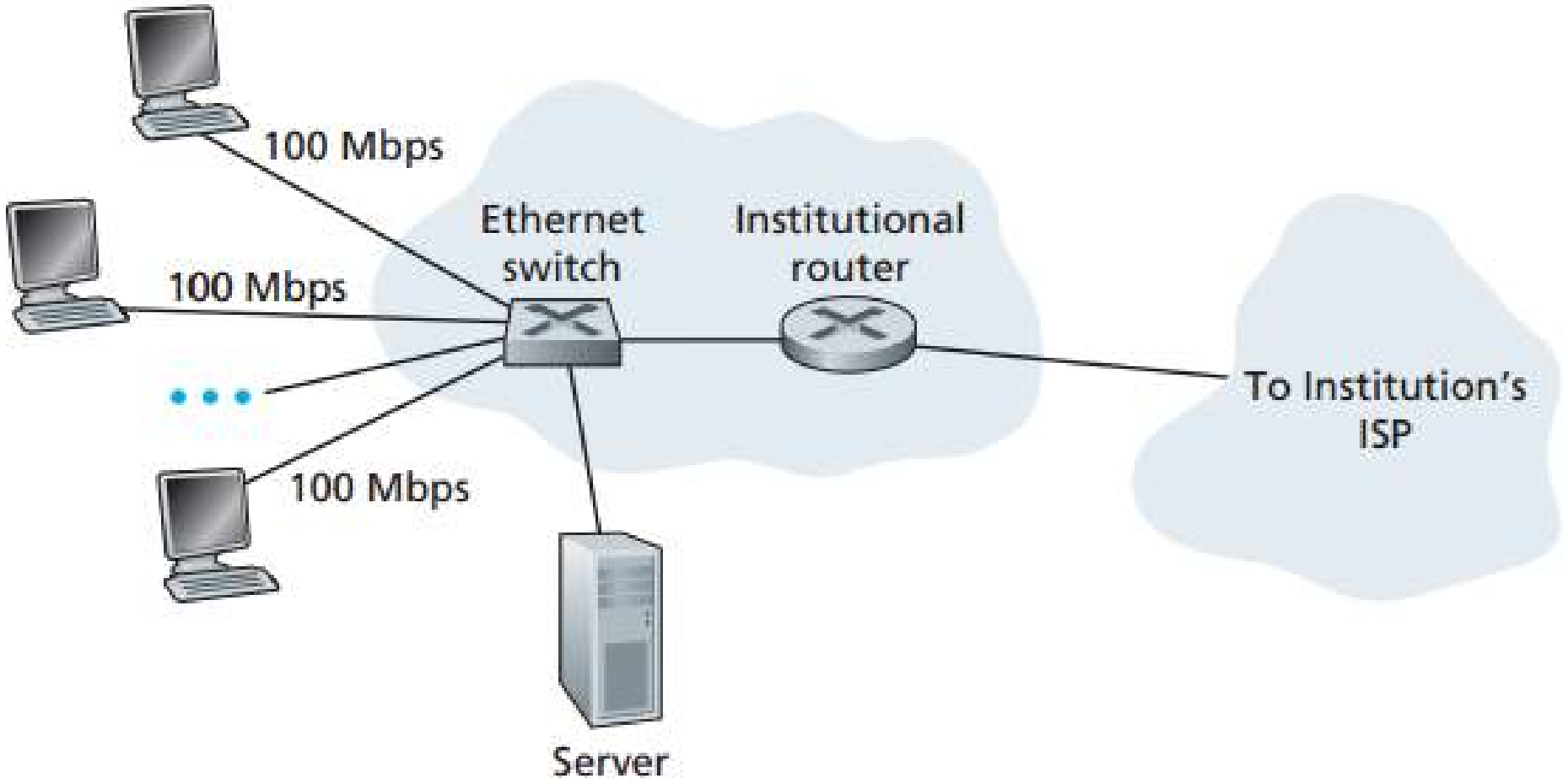
- *The analog signal from many such houses are translated back into digital format at the DSLAM.*

Fiber to The Home (FTTH)

- *Each home has an ONT (optical network terminator , which is connected to a optical splitter via a dedicated optical fiber link.*
- *The splitter combine a number of home (typically less then 100). The OLT (optical line terminator) provides the conversion between optical and electrical signals, connects to the internet via telco router*
- *FTTH can provide internet access rates in the Gbps range.*

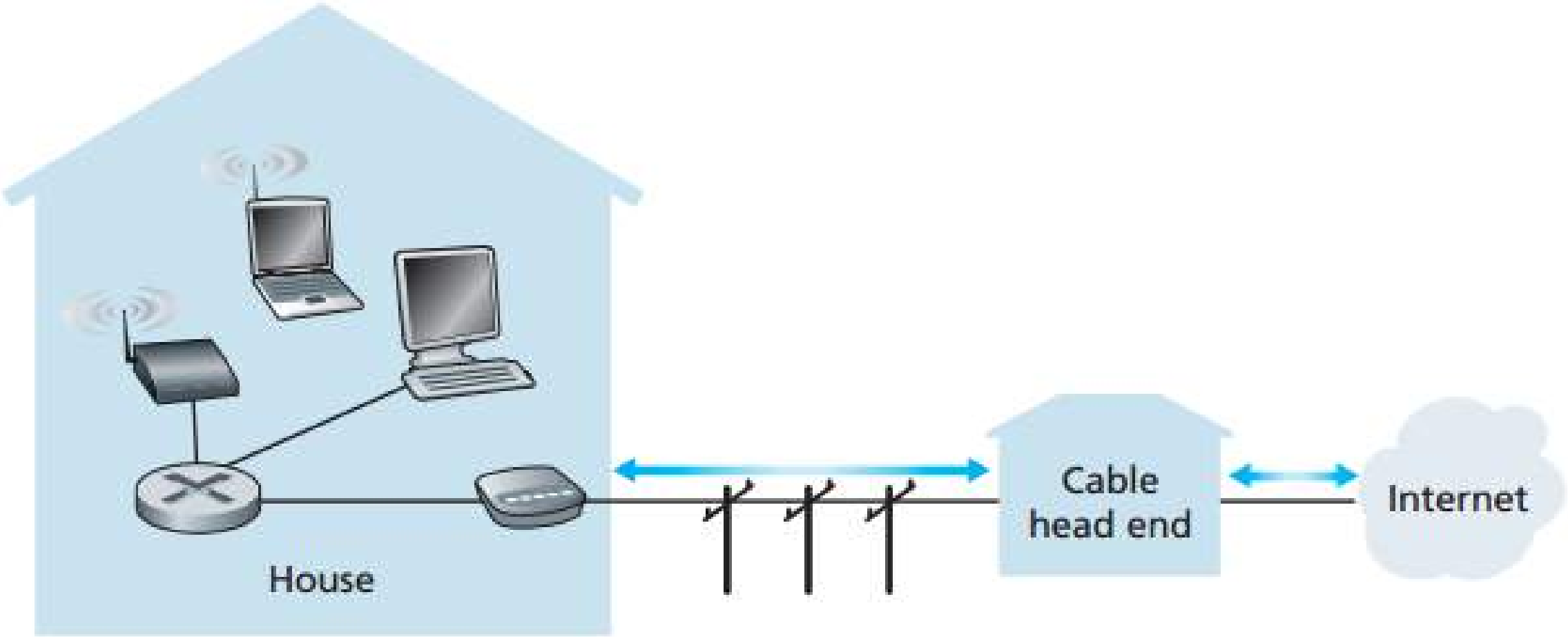
Access in Enterprise (and home): Ethernet and WiFi

- *It is typically use in university campus, companies and home settings where many devices and IOT related applications are connected.*
- *Ethernet users use twisted pair copper wire to connect to an Ethernet switch.*



Access in Enterprise (and home): Ethernet and WiFi

- *The Ethernet switch, or network of such interconnected Ethernet switches are then in turn connected to the larger internet.*
- *With Ethernet switch the user have 100 Mbps of data rate whereas the servers have 1 Gbps to 10 Gbps of data rate.*
- *The LAN (local area network) users must be in the range of 10 meters from the access point.*
- *Wireless LAN access based on IEEE 802.11 are mostly know as WIFi.*



A typical Home network

Wide- Area Wireless Access : 3G and LTE

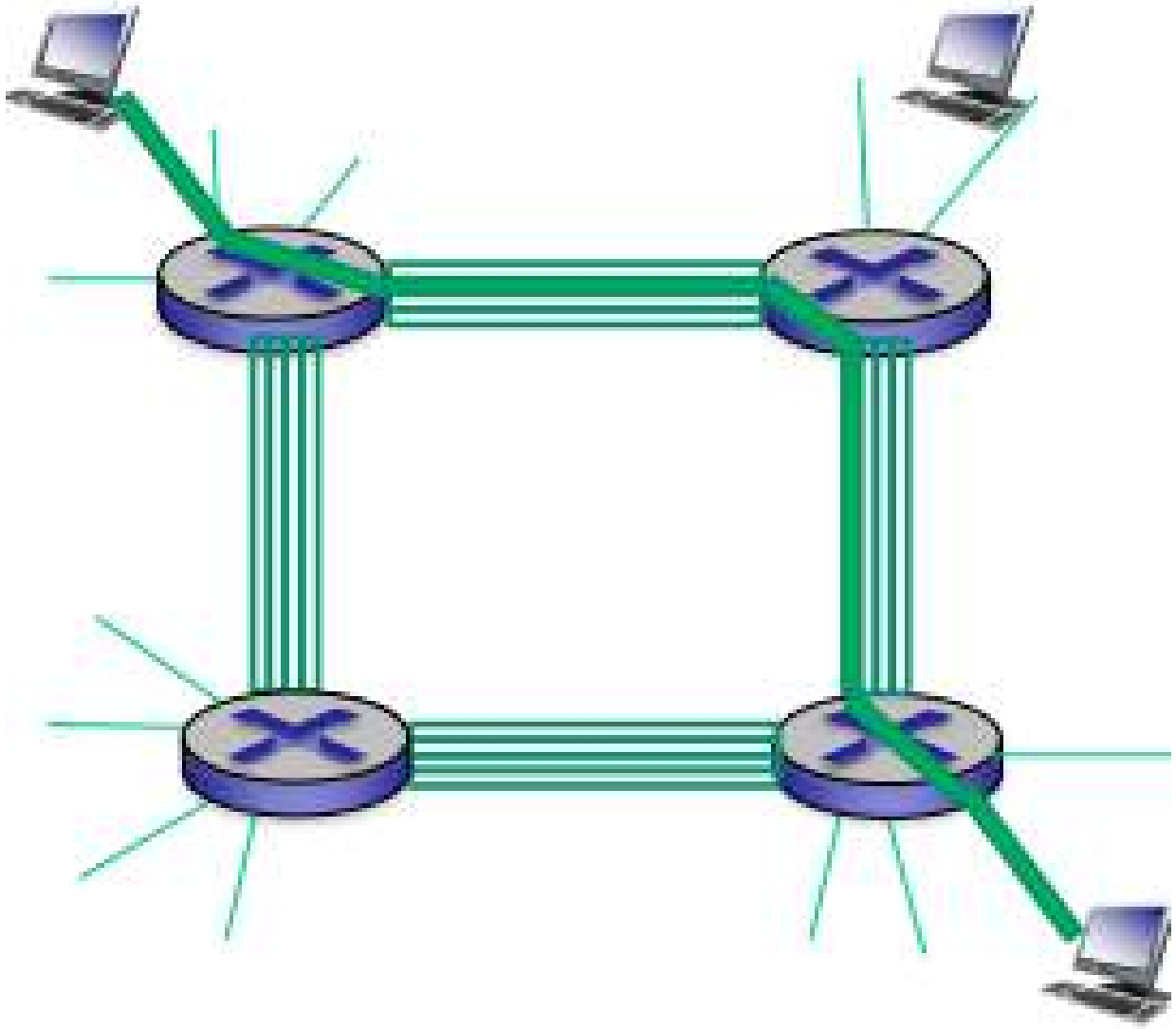
- *Here the wireless devices (e.g. mobile phones) need to be within few kms from the base station to avail the 3G services and a data rate of typically 100 Mbps.*
- *Here the range is in kms whereas in WiFi the user need to be in tens of meter range from the access point ..*

Physical Medium

Guided Medium	Unguided Medium
❑ Twisted copper wire (Telephone lines) ❑ Coaxial cable (television cable wire) ❑ Fiber optic	❑ Terrestrial radio networks • Very short distance over two or three meters (keyboard , mouse, headset) • Local area networks for ten meters to a few hundred meters (wires LANs) • Wide area network for tens of kms (cellular access technology e.g. cell phone tower) ❑ Satellite radio channels

Circuit Switching	Packet Switching
❑ A dedicated path is established between the sender and receiver. ❑ Before data transfer connection will be established first. ❑ The three phases of circuit switching are: ❑ connection establishment ❑ Data transfer ❑ Connection termination	❑ In packet switching, hosts breaks the messages into smaller packets. ❑ Forward packets from one router to the next, across links on the path from source to destination. ❑ Each packet is transmitted with full link capacity. ❑ Packet switching works with the principal of **Store-and –Forward.** ❑ **For sensing a packet of L bits over a link with transmission rate R bits/s, then the time required to transmit L bits is L/R secs.**

Circuit Switching



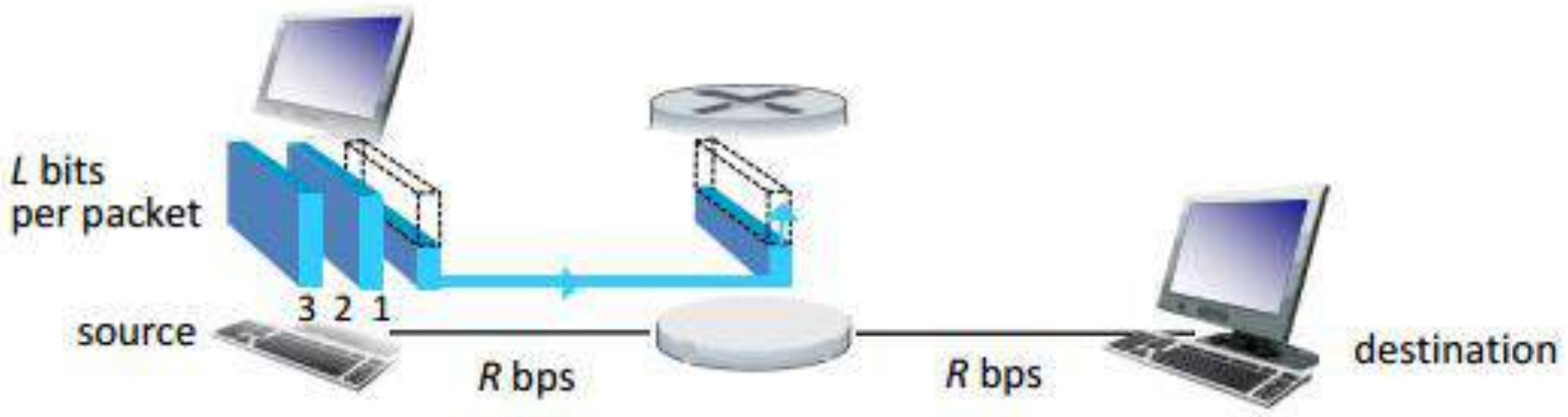
- A link consists of four circuits to support for connections.
- If each link between switches has a transmission rate of 1 Mbps, then each end-to-end circuit connection gets $\frac{1\ Mbps}{4}$= 250 Kbps of dedicated transmission rate.

Module 1

Introduction to Computer Networks

Dr. Sunandita Debnath, IIIT Vadodara

Store-and-Forward



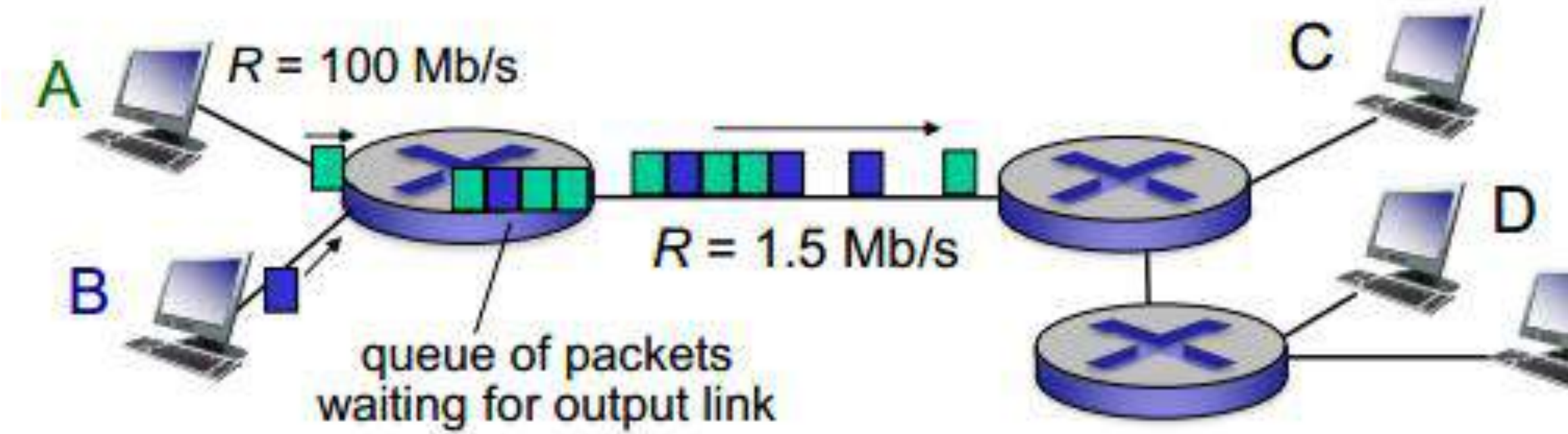
- The packet switch must receive the entire packet before it can *start* to transmit the first bit of the packet onto the outbound link. Instead it must *store* the packets' bits.
- Suppose the source is transmitting 3 packets of equal length of *L* bits and the link has a transmission rate of R bits/sec
- Until the entire bits of packet #1 will reaches at the router, the router will not begin transmission. So, the time required for the packet #1 to reach the destination is (L/R+L/R)secs i.e. 2L/R secs.

Store-and-Forward

- Now for reaching all the *three* packets for reaching the destination is (3L/R+L/R) secs.
- Here we have considered that the router has only one incoming and out going link.
- Now for sending one packet from source to destination over a path consists of N links (means N-1 routers have to crossed) each of transmission rate of R bits/sec.
- Then the total end to end delay is

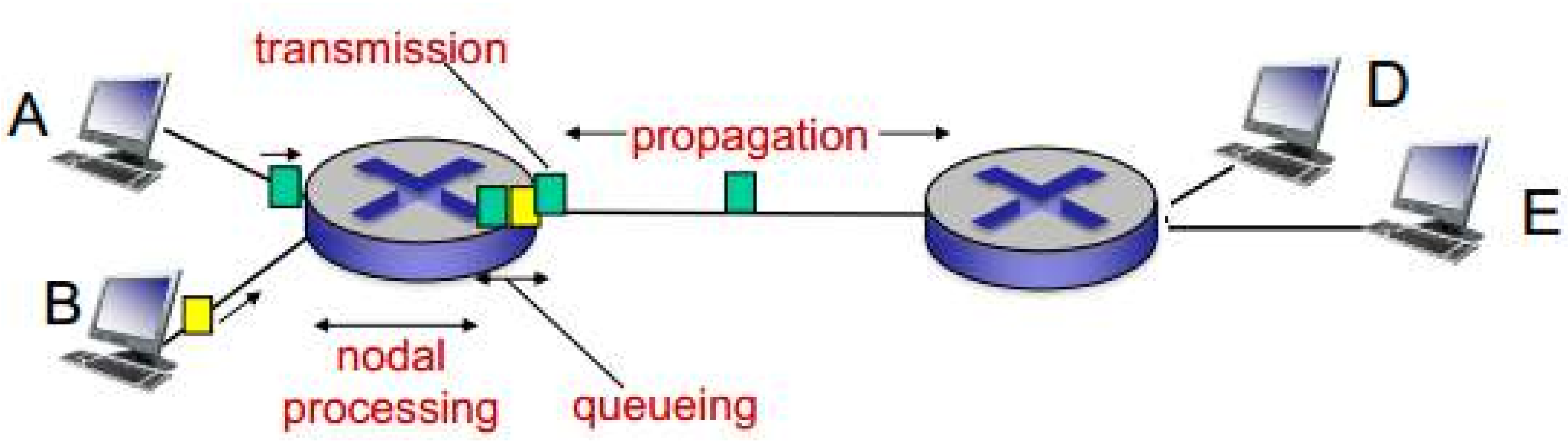
$$d_{end-to-end} = N\frac{L}{R}$$

Queuing delay and Packet loss



- For each attached link packet switch has an output buffer (also known as output queue) which stores packet which the router is about to send into that link.
- This queueing occurs when the arrival rate on the link exceeds the trasmmission rate of the link for a period of time.
- Two things happens in such situations:
- Packet will be queued, wait to be transmitted on the link
- Packets can be dropped (lost) if memory (buffer) fills

Four source of Packet delay



- A packet starts in a host (the source e.g. A or B), passes through a series of routers, ends its journey in another host (destination e.g. C or D). As packet travels from a node (host or router) to the subsequent node (host or router) along this path, suffers from several types of delay.

$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

Processing Delay (d_{proc}) *Processing delay in high speed routers are typically on the order of microseconds or less.*

Queueing Delay (d_{queue}) *packets will be transmitted in a first-come-first-served manner. waiting Time at the output link for transmission depends on congestion level of routers.*

Transmission Delay (d_{trans}) *If the length f packet is L bits and the transmission rate of the link from router A to B is R bits/sec then the transmission delay is*

$$d_{trans} = \frac{L}{R}\ sec$$

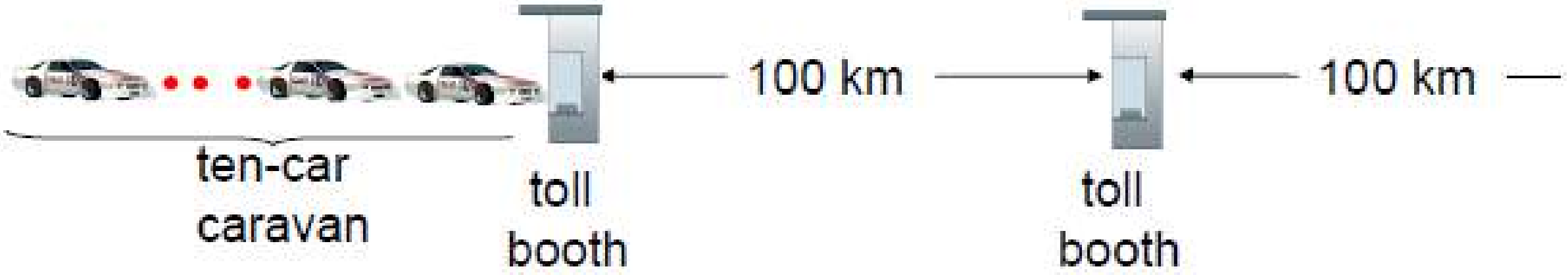
It is the amount of delay or time required to push all the bits of a packet into the link.
It is order of milisecs or microsecs.

Propagation Delay (d_{prop})

The propagation delay is given by

$$d_{prop} = \frac{Distance\ between\ the\ two\ router\ (phyisical\ Length)}{Propagation\ speed\ of\ the\ link} = \frac{d}{s}$$

Caravan Analogy

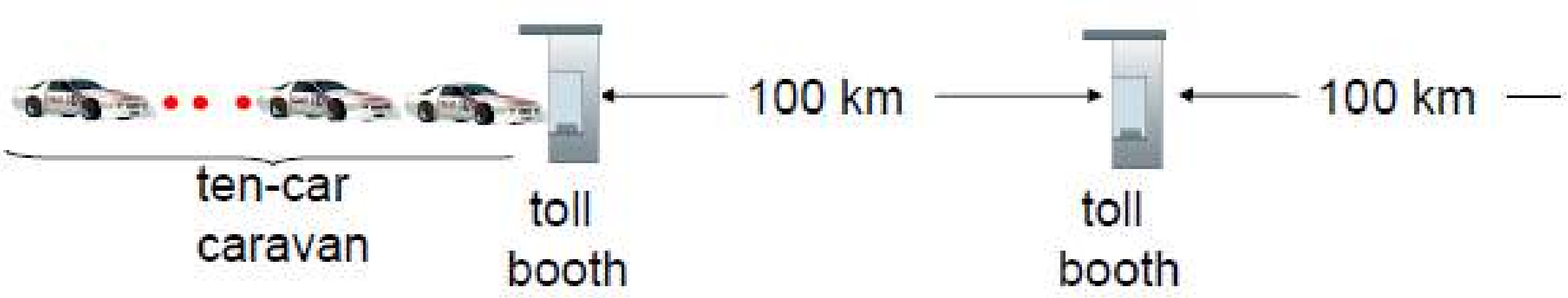


- *In a highway each of the tollbooths are 100 kms away.*
- *Suppose there are 10 cars travelling together as caravan, follow each other in a fixed order.*
- *Assume that the car travels on the highway at a rate of 100kms/hours.*

$$d_{prop} = \frac{Distance\ between\ the\ two\ router\ (phyisical\ Length)}{Propagation\ speed\ of\ the\ link} = \frac{d}{s} = \frac{100\ Kms}{100 kms/hour}$$

$$= 60\ mins$$

Caravan Analogy Contd…

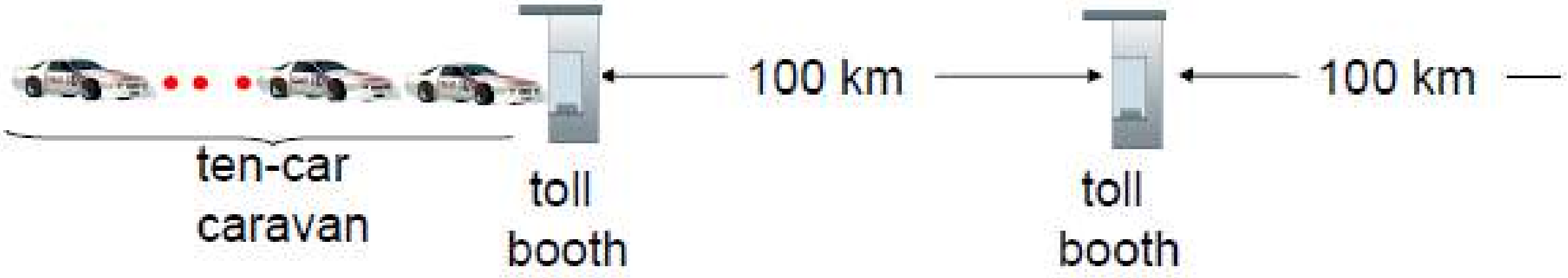


- *Let the tollbooth services the one car at 12 secs. On the other way in 60 secs it will be able to service all the 10 cars.*
- *So the transmission delay of the caravan*

$$d_{trans} = \frac{Length\ of\ packet}{Transmission\ rate\ of\ the\ link} = \frac{L}{R} = \frac{10\ cars}{5\ cars/min} = 2mins$$

- *Also assume that whenever the first car arrives at the tollbooth it waits at the entrance of the tollbooth until the other 9 cars arrives and lined up in sequence.*

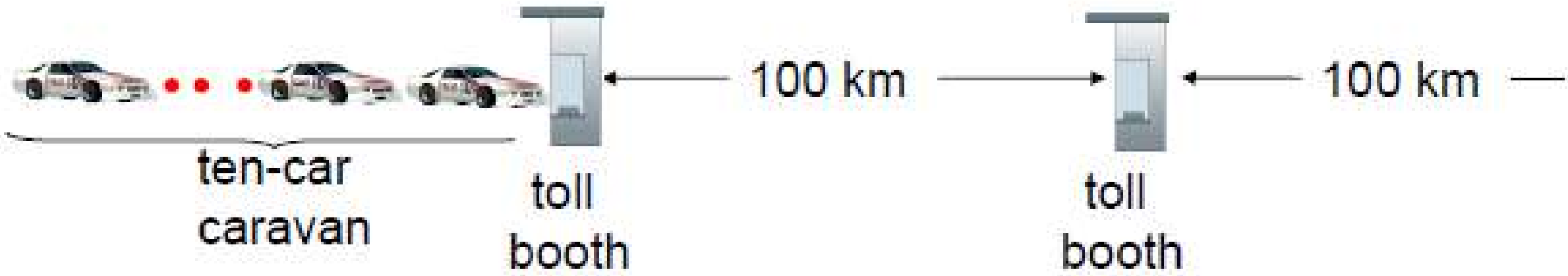
Caravan Analogy Contd…



- *So finally the time from when the caravan is stored infront of a tollbooth until the caravan is stored infront of the next tollbooth is the sum of the*

$d_{trans} + d_{prop} = 60 + 2 = 62mins$

Caravan Analogy



- *In a highway each of the tollbooths are 100 kms away.*
- *Suppose there are 10 cars travelling together as caravan, follow each other in a fixed order.*
- *Assume that the car travels on the highway at a rate of 1000kms/hours.*

$$d_{prop} = \frac{Distance\ between\ the\ two\ router\ (phyisical\ Length)}{Propagation\ speed\ of\ the\ link} = \frac{d}{s} = \frac{100\ Kms}{1000kms/hour}$$

$$= 6\ mins$$

Caravan Analogy Case2 Contd…

- *Let the tollbooth services the one car at 1 mins secs.*
- *So the transmission delay of the caravan*

$$d_{trans} = \frac{Length\ of\ packet}{Transmission\ rate\ of\ the\ link} = \frac{L}{R} = \frac{10\ cars}{1\ cars/min} = 10\ mins$$

- *So the* ***transmission delay*** > ***propagation delay****. In this case the first car will reach the next tollbooth before the last care in the caravan leaves the first tollbooth.*

Module 1

Introduction to Computer Networks

Dr. Sunandita Debnath, IIIT Vadodara

Queuing Delay and Packet Loss

Queuing Delay:

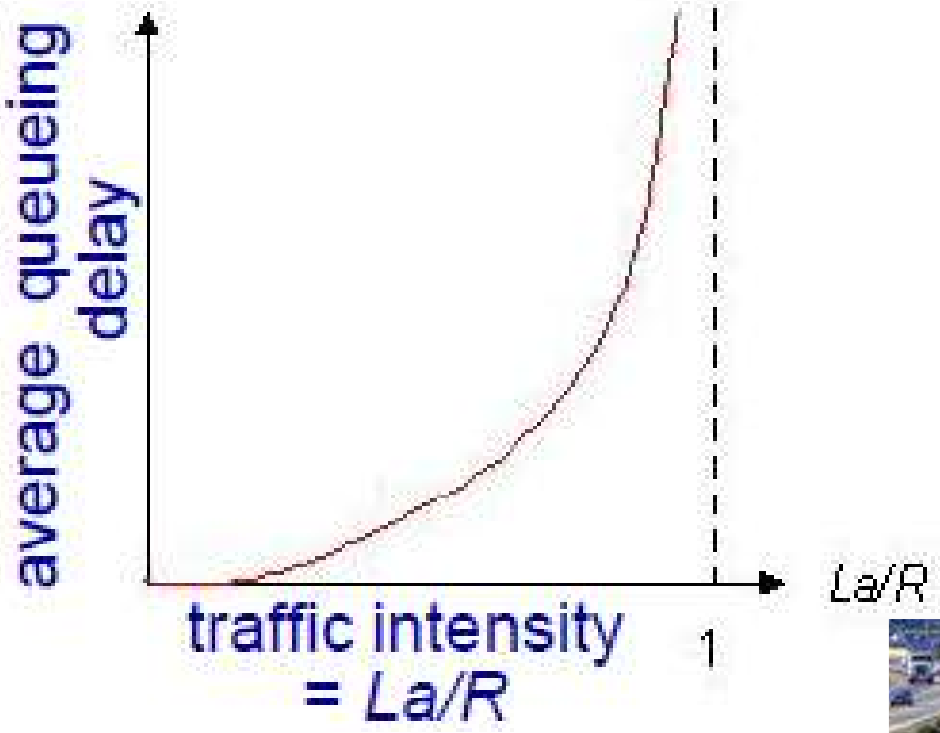
- *Let a = average rate at which packet arrives at the queue a packets/sec*
- *R = transmission rate (in bits/sec) at which the bits are pushed out of the queue and loaded to the outbound link of the router.*
- *Assume each of the packet is consists of L bits. So, the total size of traffic in bits La.*

$$Tarffic\ Intensity = \frac{La}{R}\ (the\ desired\ traffic\ intensity\ should\ not\ be\ greater\ than\ 1)$$

- $If \frac{La}{R} > 1,\ the\ arrival\ rate\ is\ greater\ than\ transmiison\ rate.$
- $If \frac{La}{R} < 1,\ the\ arrival\ rate\ is\ lesser\ than\ transmiison\ rate.$
- *If packets arrive in bursts but periodically, there can be a significant average queuing delay. For example, suppose N packets arrive simultaneously every (L/R)N seconds. Then the first packet transmitted has no queuing delay; the second packet transmitted has a queuing delay of L/R seconds; and more generally, the Nth packet transmitted has a queuing delay of (N-1)L/R seconds.*

Queuing Delay and Packet Loss

Queuing Delay



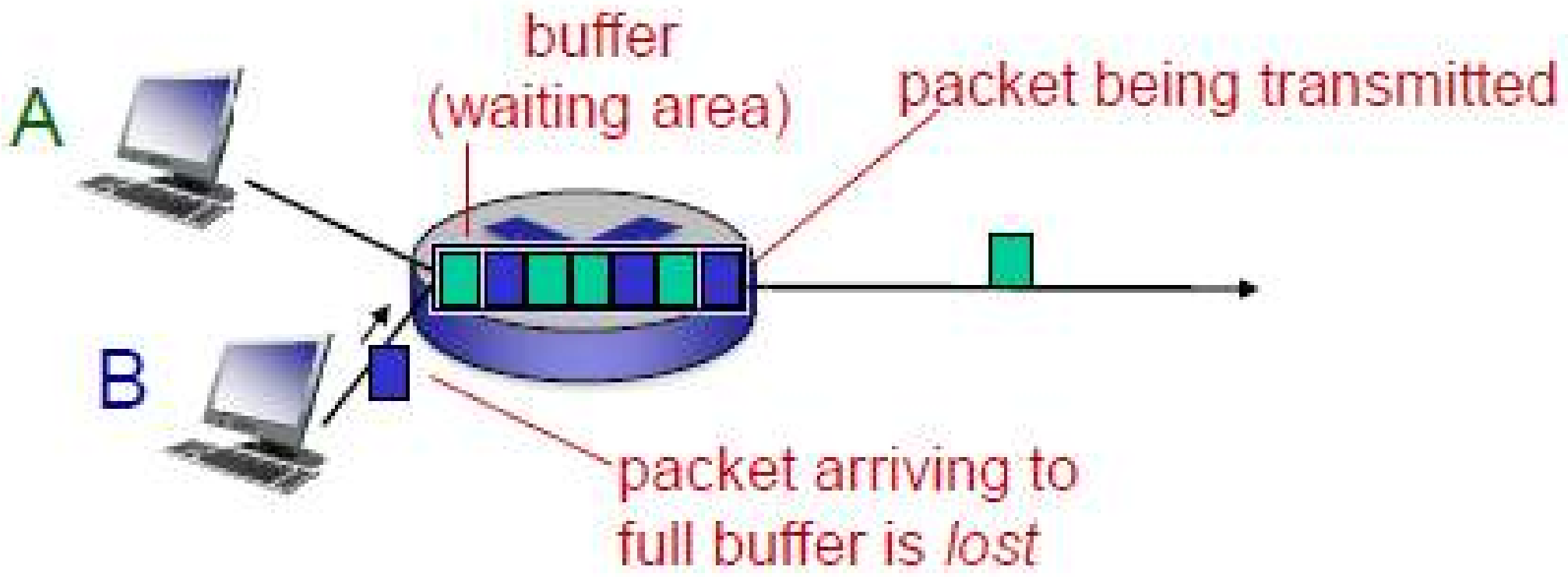
La/R~0

La/R> 1

Queuing Delay and Packet Loss

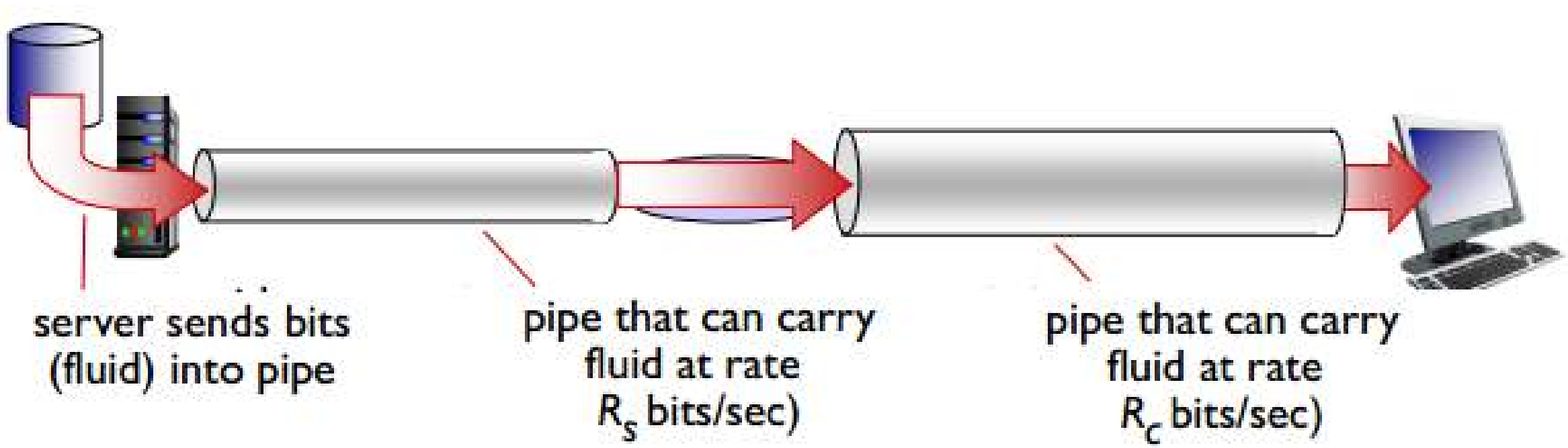
Packet Loss:

- *Queue (buffer) has finite capacity*
- *If the queue is full, then any new packet arriving at the router will be lost or dropped.*
- *Lost packet may be retransmitted*
- *The fraction of lost packet increasing as the traffic intensity increases.*



Throughput

- *The rate (bits/sec) at which bits transferred from sender to receiver.*
- *suppose the rate of the link between server and router is R_s $bits/sec$, and the rate of the link between the router and the client is R_c $bits/sec$.*
- *Two cases and occur $R_s < R_c$ and $R_s > R_c$*



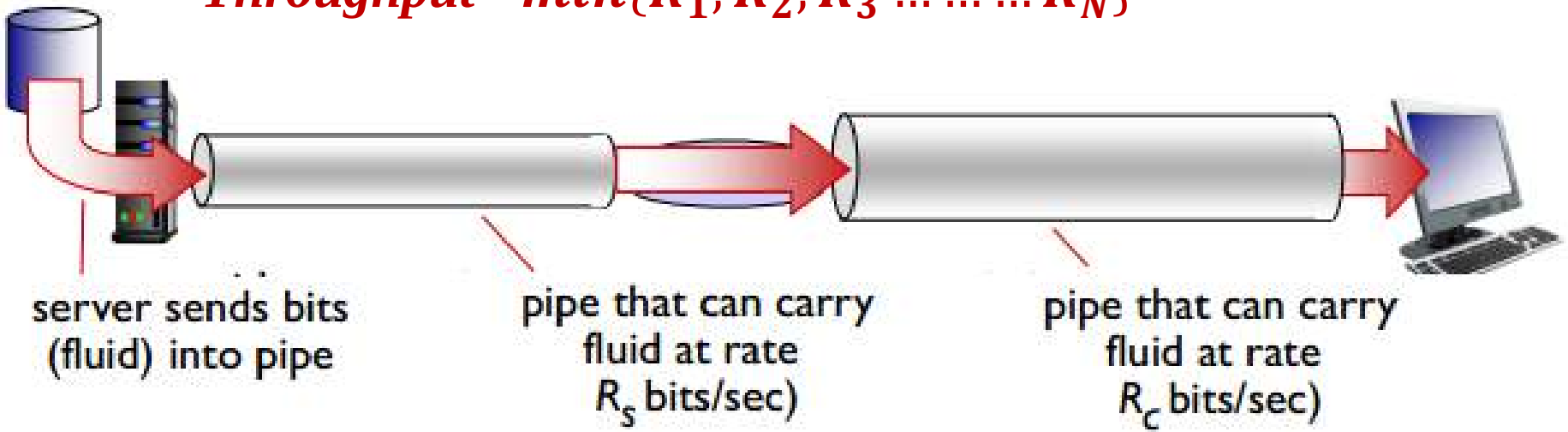
Throughput contd..

- $R_s < R_c$: *the bits pumped by the server will flow right through the router and arrive at the client at a rate of* R_s *bits/sec, giving a throughput of* R_s *bps.*
- $R_s > R_c$: *the router will not be able to forward bits as quickly as it receives them from the server. In this case bits will only leave the router at a rate of* R_c *bits/sec, giving a throughput of* R_c *bps.*
- *Because of this condition the backlog of bits at the router waiting for transmission to the client will grow and grow –which is a most undesirable condition*

$$\textbf{Throughput} = min\{R_s, R_c\}$$

If there are N no. of links and transmission rates of $R_1, \ldots\ldots R_N$***, then the throughput will be***

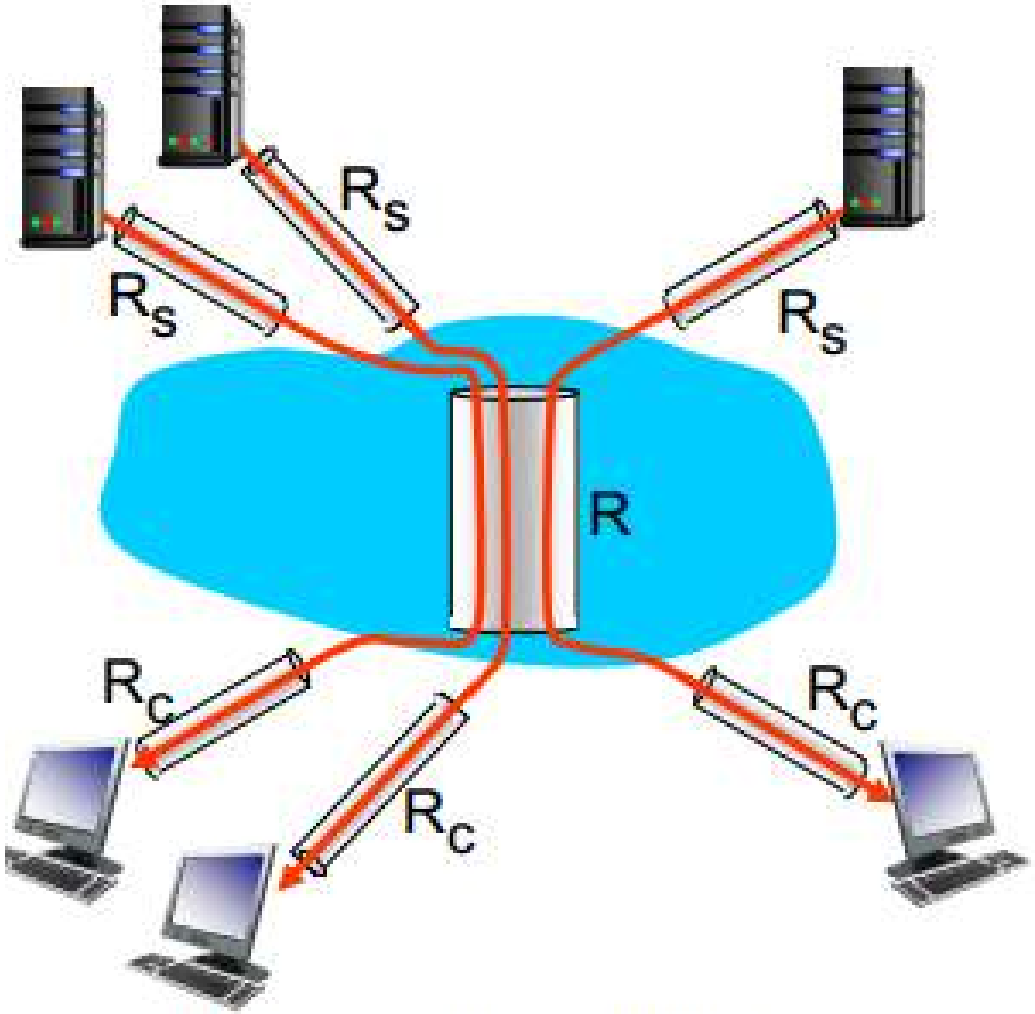
$$\textbf{Throughput} = min\{R_1, R_2, R_3 \ldots\ldots\ldots R_N\}$$



Throughput contd..

- *Suppose there are 10 downloads from the server to client and these are the only traffic in the network.*
- *All server access link have a rate of R_s, and all client access links a have a rate of R_c bps.*
- *The common link has a transmission rate of R bps.*
- *Suppose* $\boldsymbol{R_s}$ *= 2 Mbps,* $\boldsymbol{R_c}$ *= 1 Mbps and* $\boldsymbol{R}$ *= 5 Mbps. The R will be equally shared among the 10 links* ***5Mbps/10 = 500 Kbps***.

Throughput= 500 Kbps.

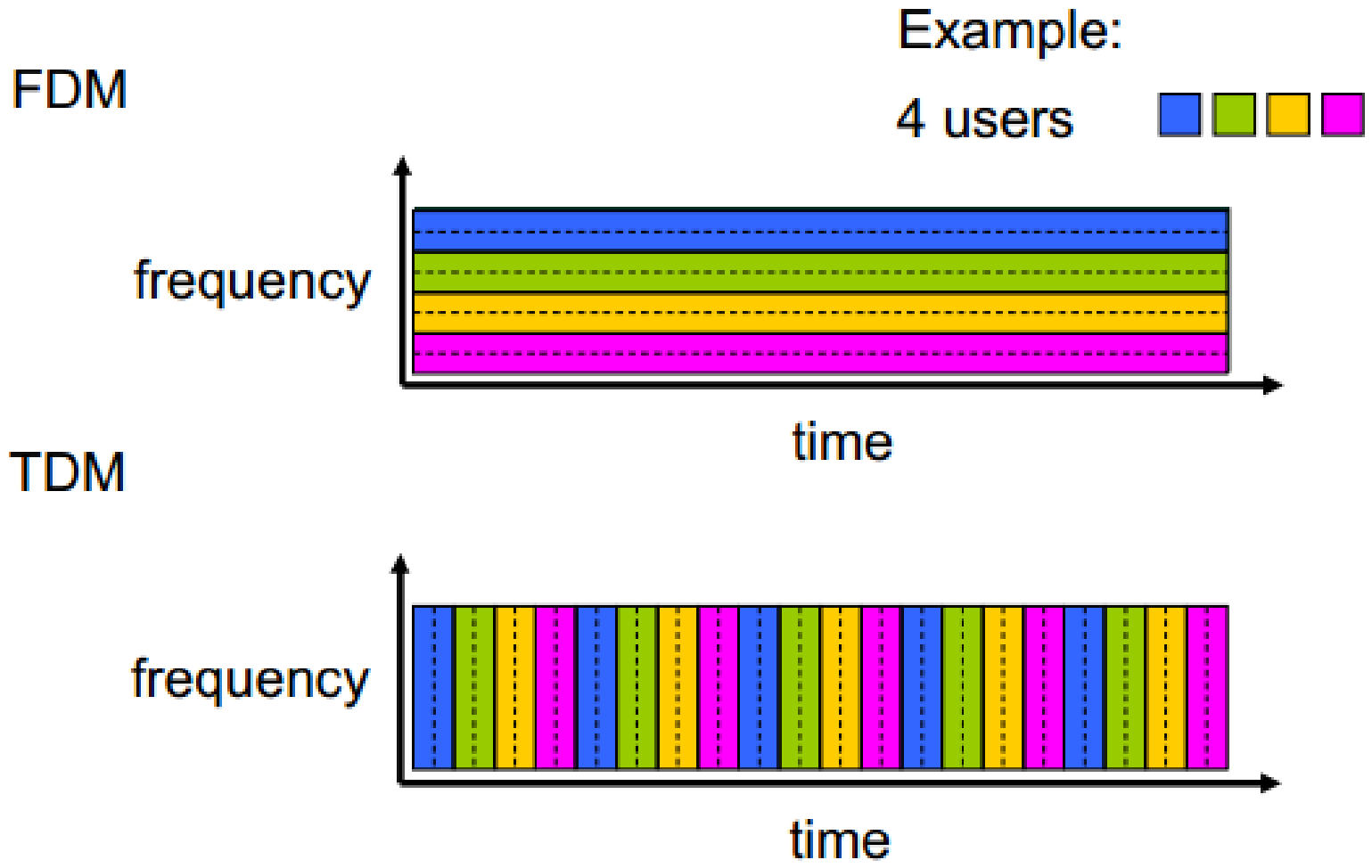


10 connections (fairly) share backbone bottleneck link *R* bits/sec

Note: Packet switches can be both routers and link layer switches.

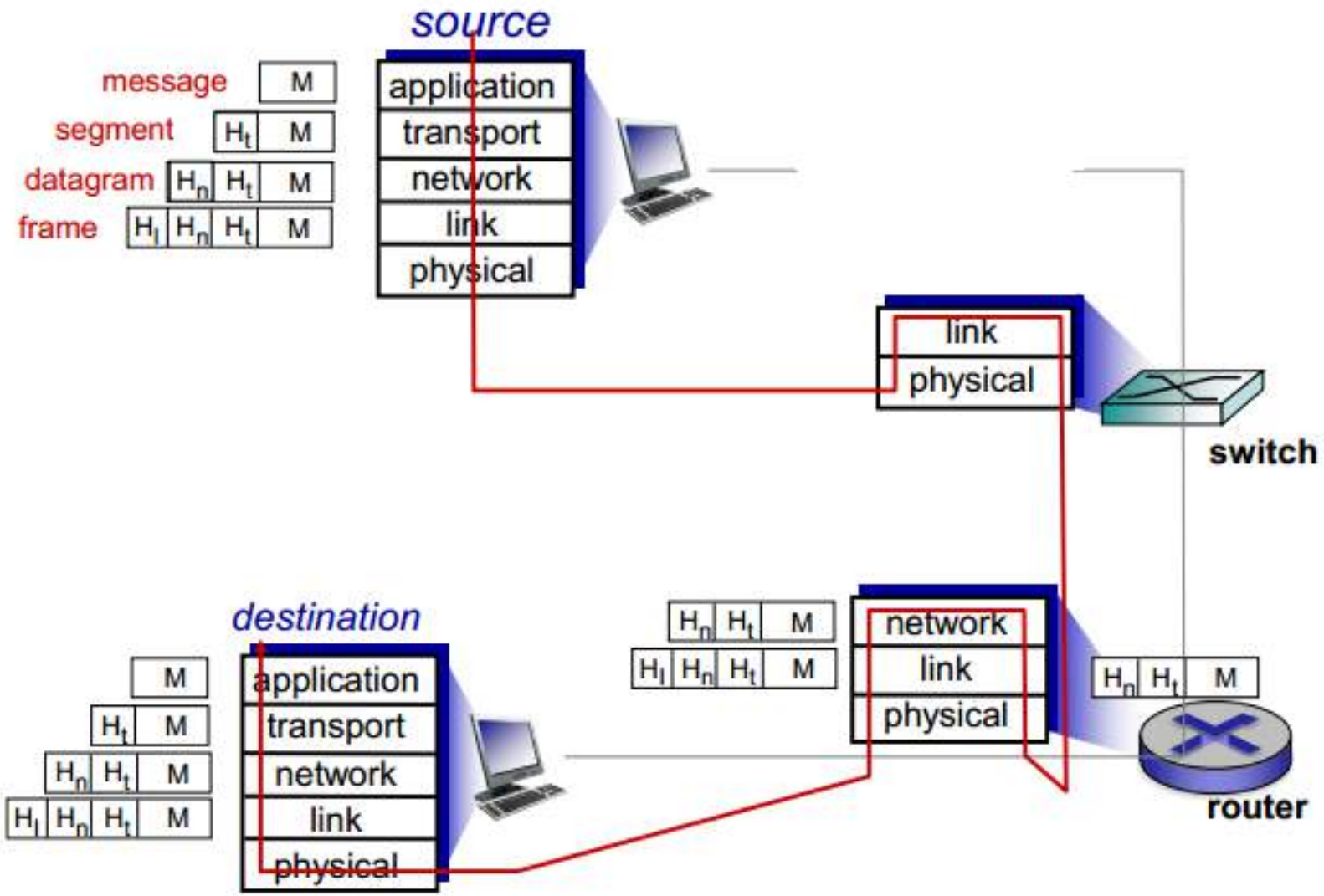
Circuit Switching

- ❑ TDM (Time Division Multiplexing)
- ❑ FDM (Frequency Division Multiplexing)



Encapsulation

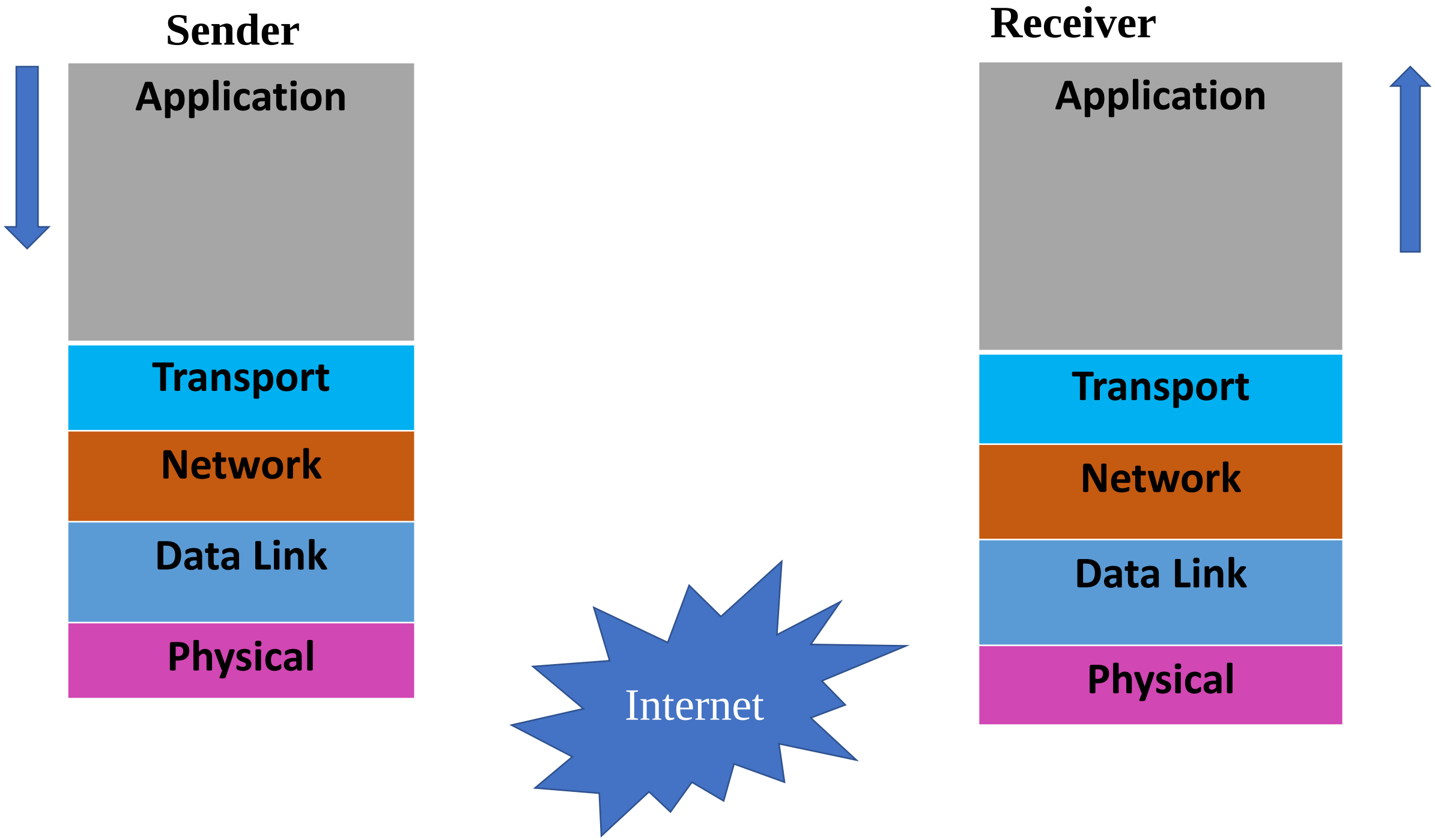
- At each layer the a packet has two fields one is header fields and one is payload field.
- Payload is the typically a packet received from the above layer.



Module 1

Introduction to Computer Networks

Dr. Sunandita Debnath, IIIT Vadodara


Sender
Application
Transport
Network
Data Link
Physical
Internet
Receiver
Application
Transport
Network
Data Link
Physical

Data Link Layer

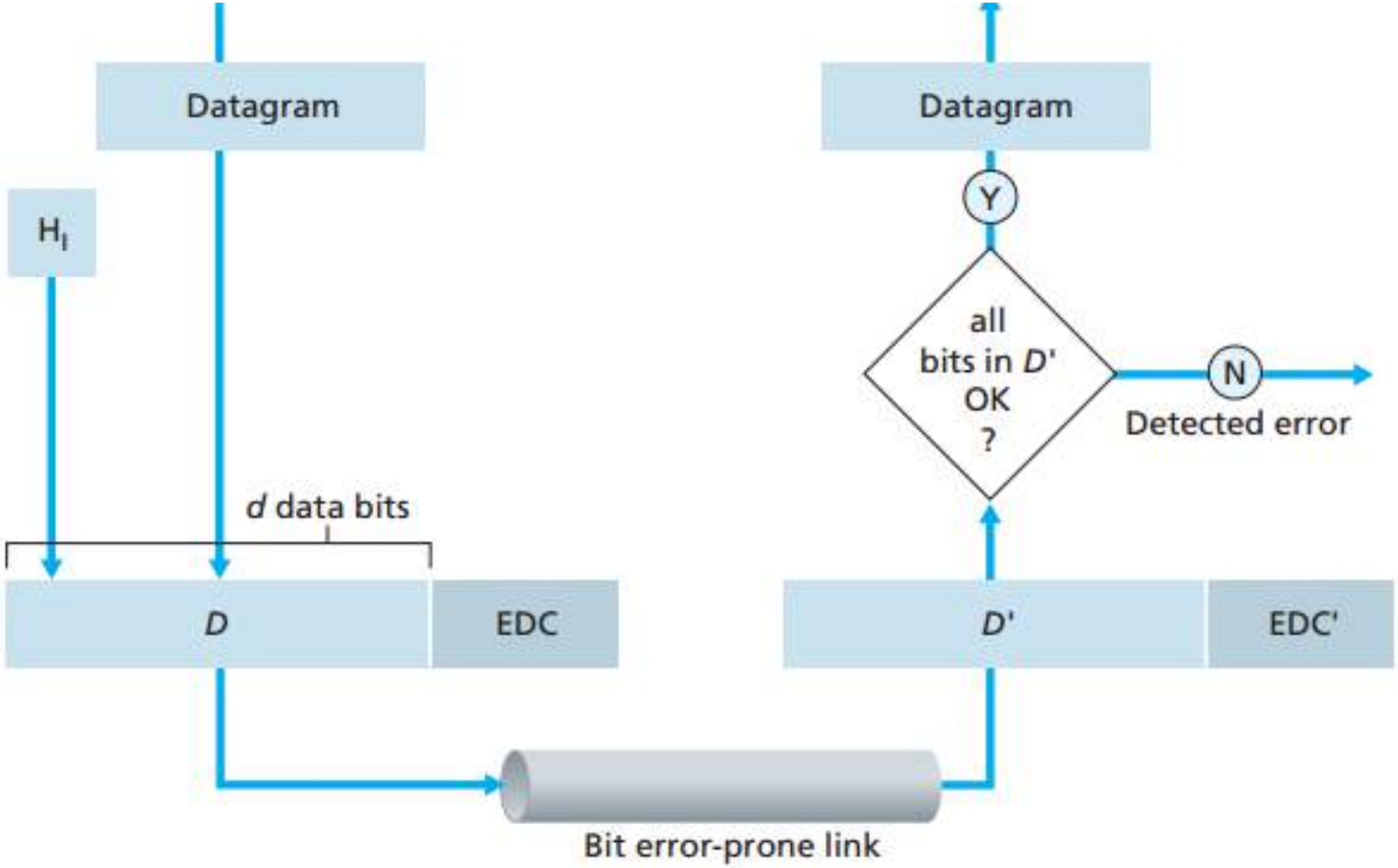
- **Data link layer is also known as Layer 2. Data link resides between the network layer and physical layer.**
- **Datalink layer are responsible for transferring datagram from one node to physically adjacent node over a link.**
- **Switch is a layer 2 device.**

Functions of Data Link Layer

- **Framing:** *encapsulate datagram into frame, i.e. adding header and trailer.*
- **Link access:** *Channel access if share media. A MAC protocol specifics the rules by which a frame is transmitted onto the link.*
  - *When there is a point-to-point connection between a sender ad receiver i.e. a single sender at one end of the link and a single receiver at another end of the link, then the MAC protocol is simple means whenever the link is idle the sender can send.*
  - *When multiple nodes share a single broadcast link-the multiple access problem occurs and the MAC protocol has to coordinate the frame transmissions of the many nodes.*
- **Reliable delivery: Reliable delivery services are offered by both transport layer and data link layer.**
  - *It is important for high error rate links such as wireless links, where the goal of correcting an error locally is more desirable than retransmission of the datagram.*
  - *Seldom used in low bit error links such as fiber and twisted pair wires.*

Functions of Data Link Layer Contd…

- **Error detection and Error correction:** *The transmitting node include error-detection bits in the frame, and having the receiving node perform an error check .*

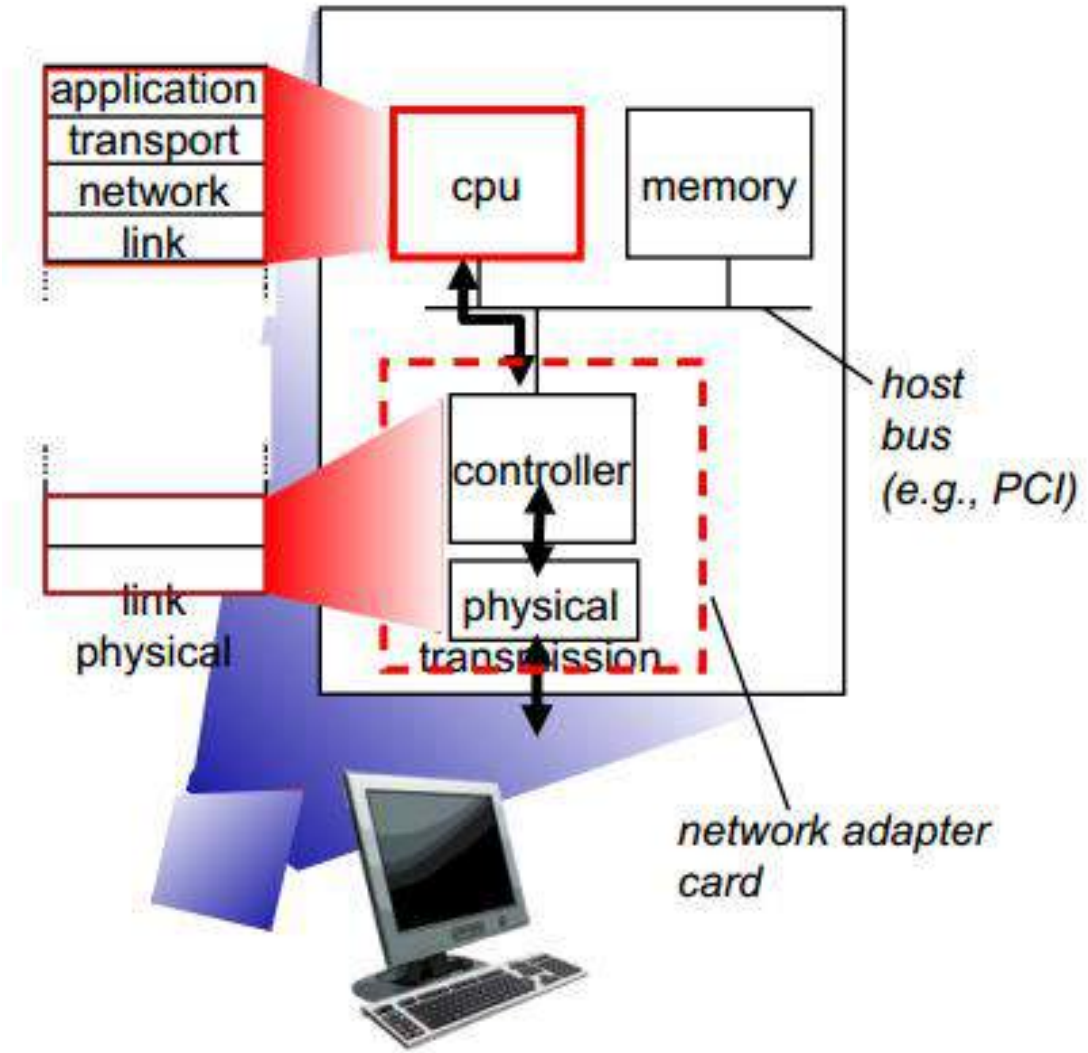


Where is this link layer implemented

- *Link layer is mostly implemented in the network adapter i.e. in the NIC (Network Interface Card) card*

- *At the heart of the network adapter is the link layer controller that implement many of the link layer services such as framing, link access, error detection and so on.*

- *Now a days network adapters are being integrated onto the hosts motherboard (LAN on motherboard configuration).*



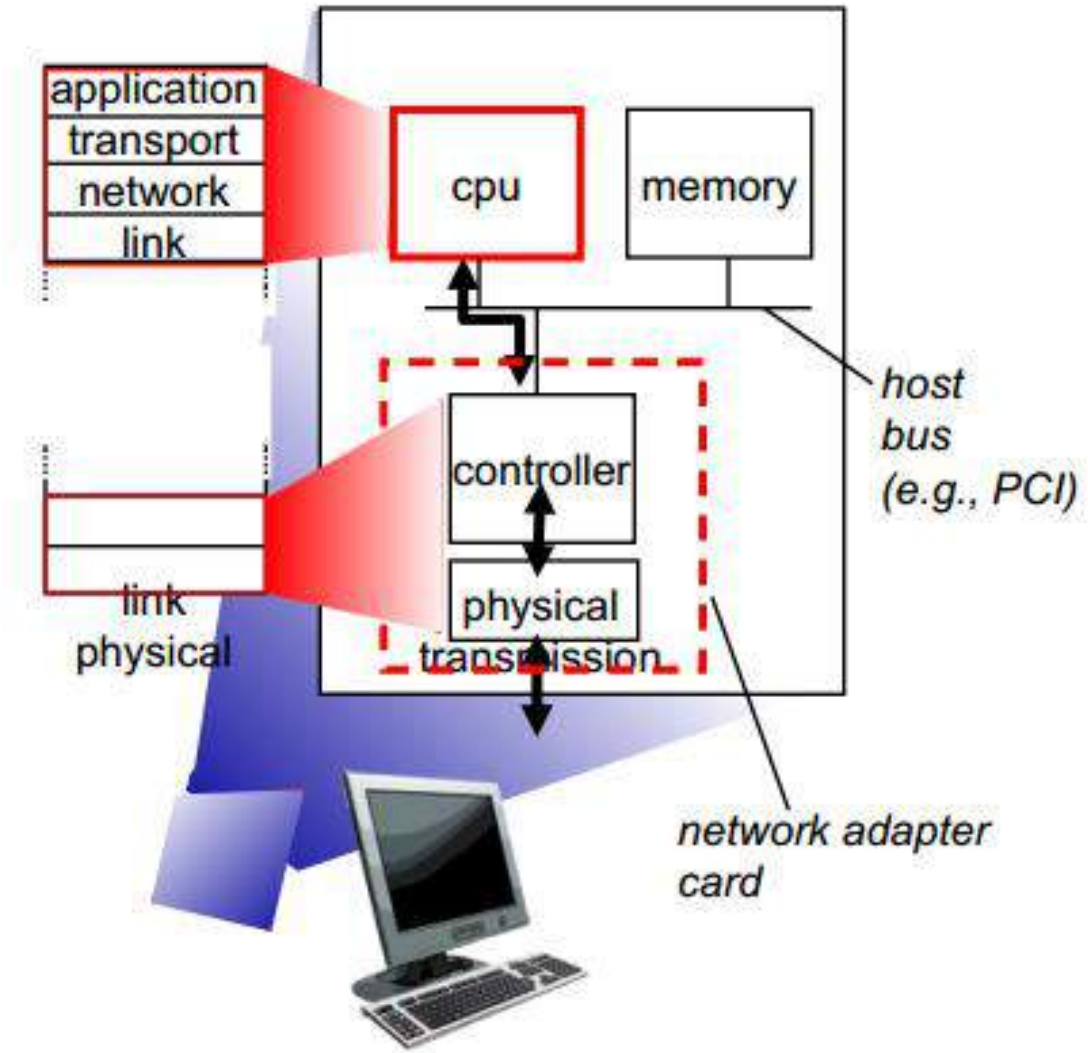
Where is this link layer implemented contd..

- *On the sender side, the controller takes the datagram that has been created and stored in the host memory by the higher layers of protocol stack, encapsulate the datagram in a link layer frame and transmit the frame into the communication link, following the link access protocol.*

- *On the receiver side, a controller receives the entire frame and extracts the network layer datagram.*

- *If the link layer performs error detection, then the sending controller that sets the error-detection bits in the frame header and is the receiving controller that performs error detection.*

- *Link layer is a combination of hardware and software-the place where software meets hardware*



Module 1

Introduction to Computer Networks

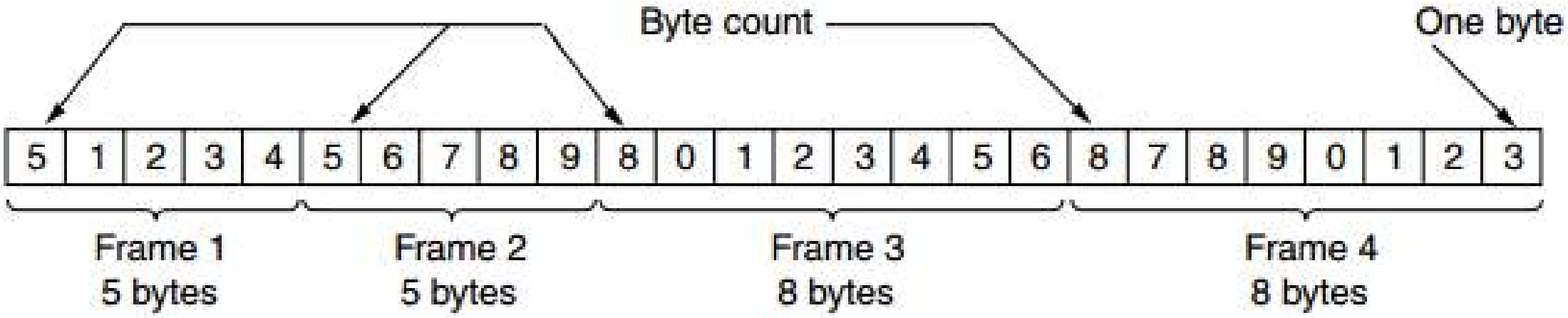
Dr. Sunandita Debnath, IIIT Vadodara

Framing

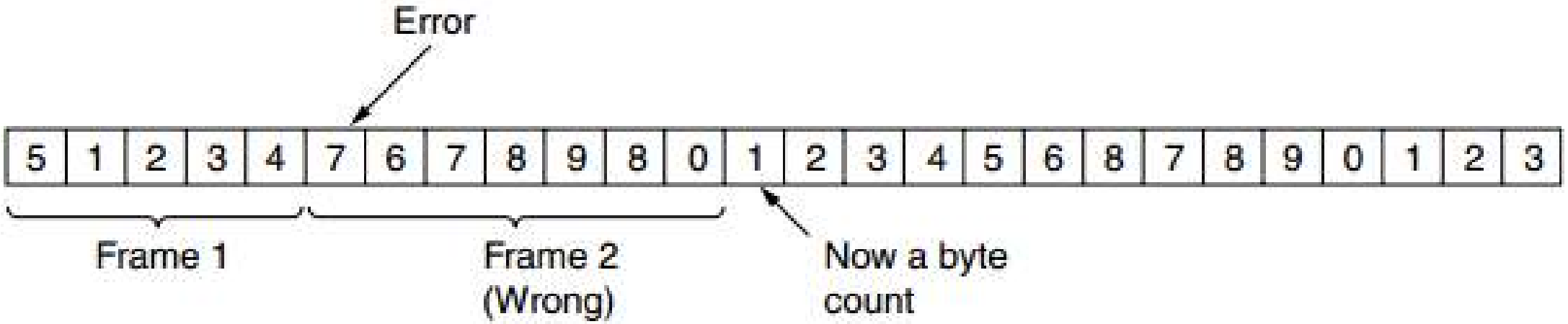
- Byte count
- Flag bytes with bytes stuffing
- Flag bits with bits stuffing
- Physical layer coding violation

Framing

- **Byte count**



(a)



Framing

➢ Flag Byte with Byte Stuffing

FLAG	Header	Payload field	Trailer	FLAG

- *Here the resynchronization problem in byte count is overcomed by inserted a special bytes at the starting and ending of a frame as delimiters, this special bytes are called Flag Byte*
- *Thus if the receiver ever loses synchronization then it will search for two consecutive Flag bytes to find the end of one frame and start of the next frame.*

Framing

- **Flag Byte with Byte Stuffing contd.**

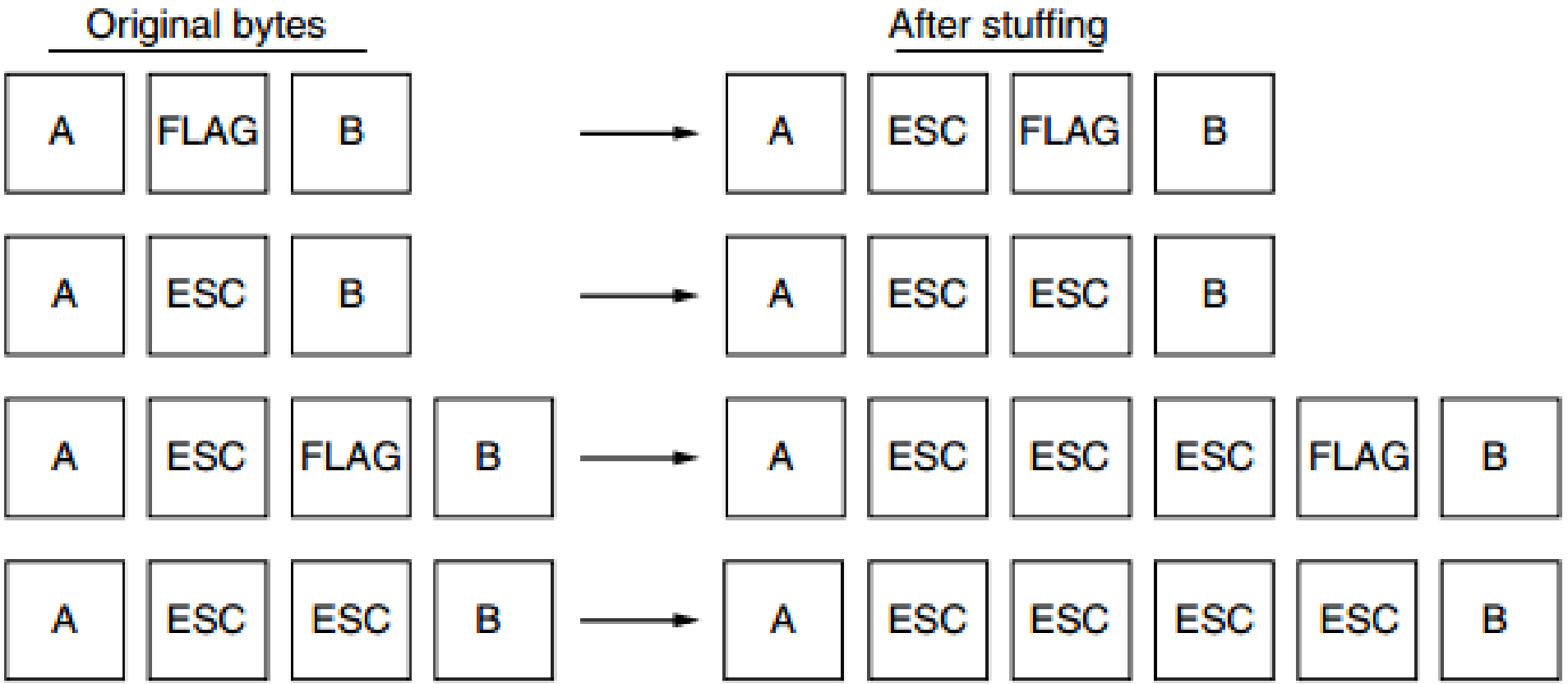
1100010011001100

- *If the same sequence of '0' and '1' which indicates the Flag bytes occurs in the middle of the data.*
- *To solve this problem is to have the sender's data link layer insert a special escape byte (ESC) just before each "accidental" flag byte in the data.*
- *Thus, a framing flag byte can be distinguished from one in the data by the absence or presence of an escape byte before it.*
- *The data link layer on the receiving end removes the escape bytes before giving the data to the network layer. This technique is called* ***byte stuffing****.*

Framing

Flag Byte with Byte Stuffing contd.

- *Now if the data itself consists of an ESC byte then what is the solution.*
- *It is again stuffed with an escape byte. At the receiver, the first escape byte is removed, leaving the data byte that follows it.*
- *the byte sequence delivered after **destuffing** is exactly the same as the original byte sequence.*



Framing

Flag Bit with Bit Stuffing

- *Byte stuffing has one disadvantage of that it is tied around to use of 8-bits byte.*
- *Frames can contain any arbitrary number of bits. It was for the protocol HDLC (High Level Data Link Control) protocol.*
- *Each of the frame begins and end with a special sequence of bits* **01111110** *or* **0x7E** *in Hex. This pattern is a Flag byte.*

01111110	**Data**	**01111110**

Framing

Flag Bit with Bit Stuffing contd…

- *If the sender data link layer finds the data consists of Flag byte then it automatically stuff one zero bit after five consecutive '1's.*
- *The reviver finds five consecutive 1 bits, followed by a 0 bit, it automatically dyestuffs (i.e.) delete the 0 bit.*

Flag	Data	Flag
	0101011111101010	
01111110	**01010111110101010**	**01111110**

Framing

Physical layer coding Violations

- *In bit stuffing and byte stuffing the length the data increases unnecessarily.*
- *This takes helps of physical layer where there are some reserved signal which will be used to indicate the starting and ending of a frame.*

Module 1

Introduction to Computer Networks

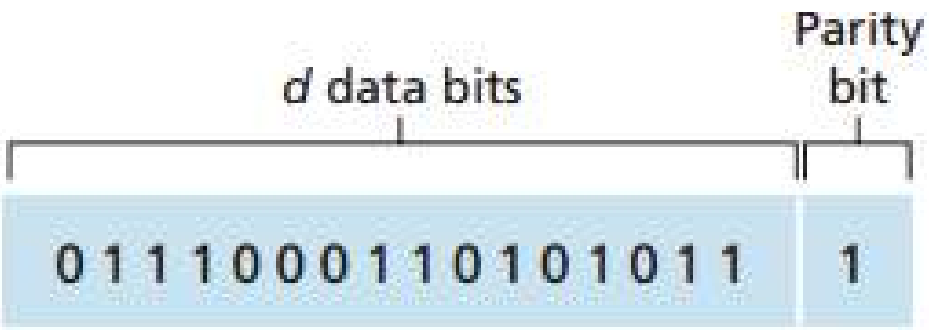
Dr. Sunandita Debnath, IIIT Vadodara

Error Detection and Correction Techniques

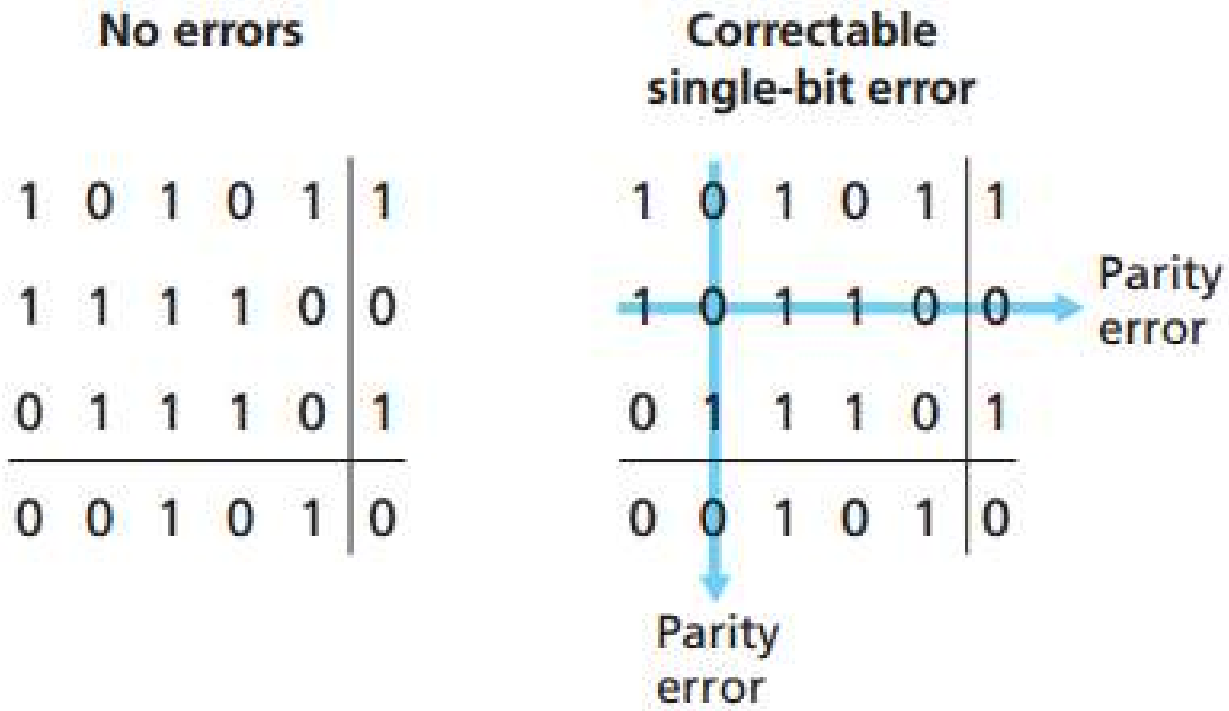
- *Parity Check*
- *Checksum*
- *Cyclic Redundancy Check (CRC)*

Error Detection and Correction Techniques

➢ Single Parity Check



➢ Two-dimensional Parity Check



Error Detection and Correction Techniques

Checksum(at the Sender side)

- *Break the original message into 'k' number of blocks with 'n' bits in each block.*
- *Sum all the 'k' data blocks*
- *Add the carry to the sum if any.*
- *Do 1's complement of the sum and this will be the checksum bit which will be appended with the transmitted data.*

Checksum(at the Receiver side)

- *Sum all the 'k' data blocks with the checksum bits*
- *If the resultant sum is all '1's then the received data bit is correct and error free.*

Error Detection and Correction Techniques

Cyclic Redundancy Check (CRC) at the Sender side

- *Find the length of the divisor 'L'.*
- *Append (L-1) bits to the original message*
- *Perform binary division operation*
- *Remainder of the divisor is the CRC*

Cyclic Redundancy Check (CRC) at the Receiver side

- *Divisor is common for both the transmitter and receiver.*
- *When there is no remainder or all 0 's at the remainder , the receiver can conclude that the transmitted bit was detected error free.*

Error Detection and Correction Techniques

CRC(at the Sender side)

Divisor 1001 (Length of Divisor L = 4)

Data 101110 (Append (L-1) zeros)

CRC bits 011

Transmitted data 101110011

```
G
1 0 0 1 | 1 0 1 1 1 0 0 0 0
          1 0 1 0 1 1
          1 0 0 1
            1 0 1
            0 0 0
            1 0 1 0
            1 0 0 1
              1 1 0
              0 0 0
              1 1 0 0
              1 0 0 1
                1 0 1 0
                1 0 0 1
                  0 1 1
                  R
D
```

CRC(at the Receiver side)

Divisor 1001 (Length of Divisor L = 4)

Received data 101110011

If no remainder (all ‘0’s)is there than received datagram is correct

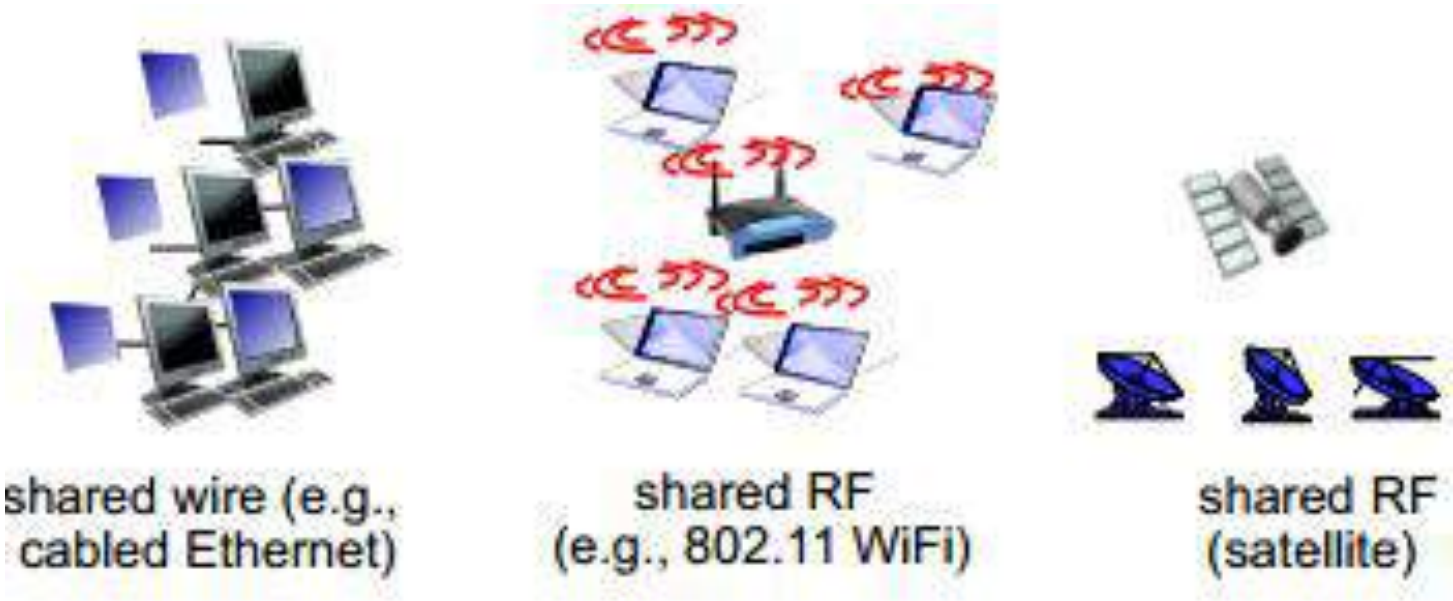
There are two types of network links in Data link layer

Point to Point link

It consists of a single sender at one end of the link and a single receiver at the other end of the link. Protocols for point-to-point link are

Point-to-point protocol (PPP)

High-level Data Link control (HDLC)



Broadcast link

It can have multiple sending and receiving nodes all connected to the same single shared broadcast channel..

- *E.g. Ethernet and Wireless LANs*

In computer networks broadcast channels can both send and receive . The main problem in this scenario is that of determining who gets to talk and when

- *Give everyone a chance to talk*
- *Don't speak until you are spoken to*
- *Don’t monopolize the conversation*
- *Raise a hand when you have a question*
- *Don’t interrupt when someone is speaking*

An ideal multiple access protocol should have the following desirable characteristics:

- *Suppose the broadcast channel is of rate R bps.*
- *When only one node has to send data that node has a throughput of Rbps*
- *When M nodes want to transmit each of M nodes has a throughput of R/M bps.*
- *Don't monopolize the conversation*
- *The protocol should be fully decentralized*
- *The protocol should be simple so that it is inexpensive to implement*

Three categories of multiple access protocols:

- *Channel portioning protocol*
- *Random access protocol*
- *Taking –turns protocol*

Module 2

Data Link Layer

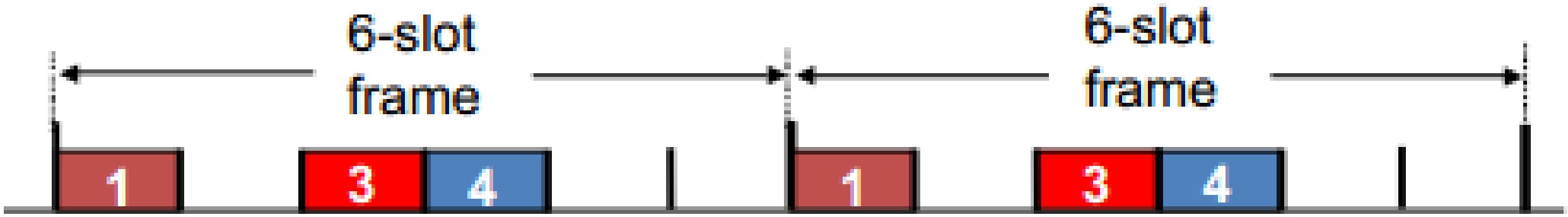
Dr. Sunandita Debnath, IIIT Vadodara

Channel partioning protocol

- *TDMA (Time Division Multiple Access)*
- *FDMA (Frequency Division Multiple Access)*
- *CDMA (Code Division Multiple access)*

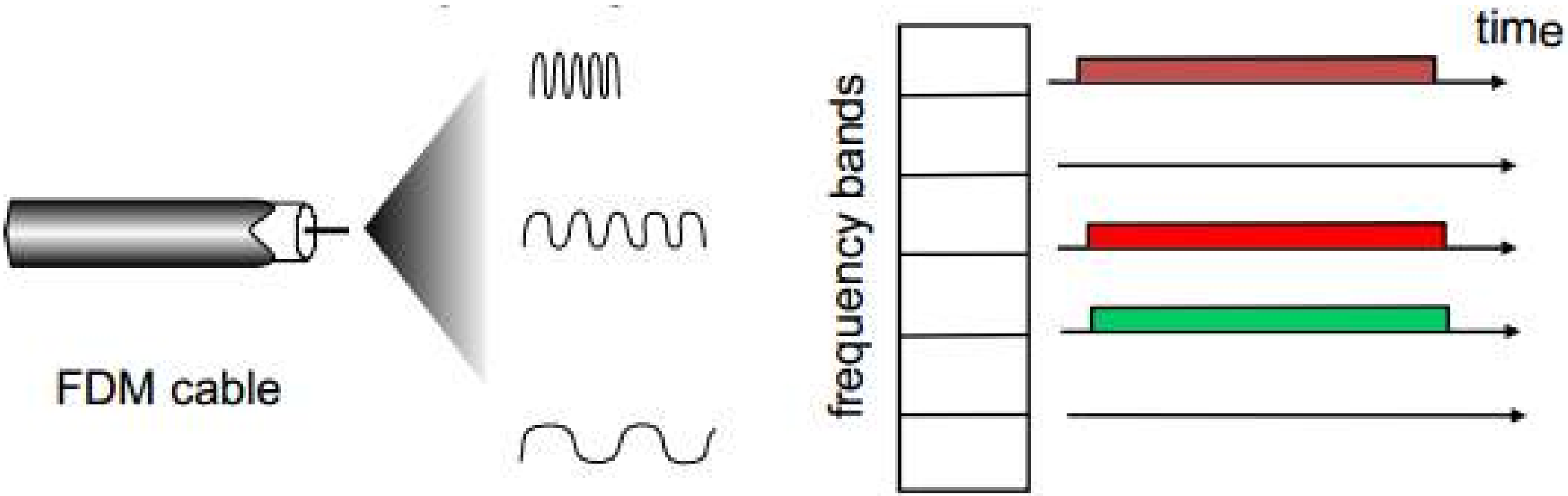
Channel portioning protocol

- *TDMA (Time Division Multiple Access)*



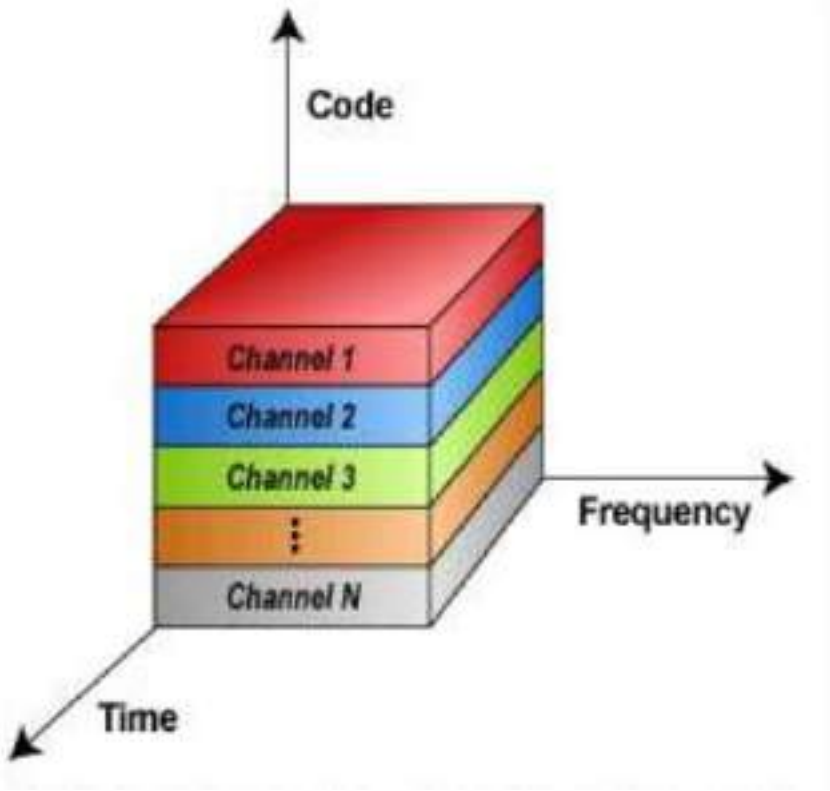
Channel portioning protocol

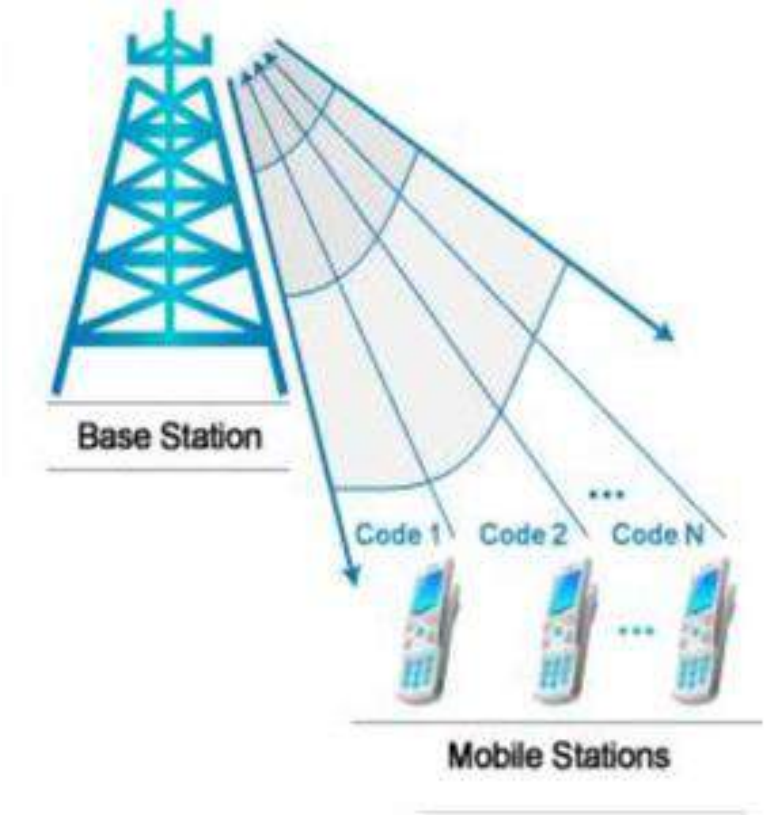
- *FDM (Frequency Division Multiple Access)*



Channel portioning protocol

- *CDMA (Code Division Multiple Access)*





Random access Protocol

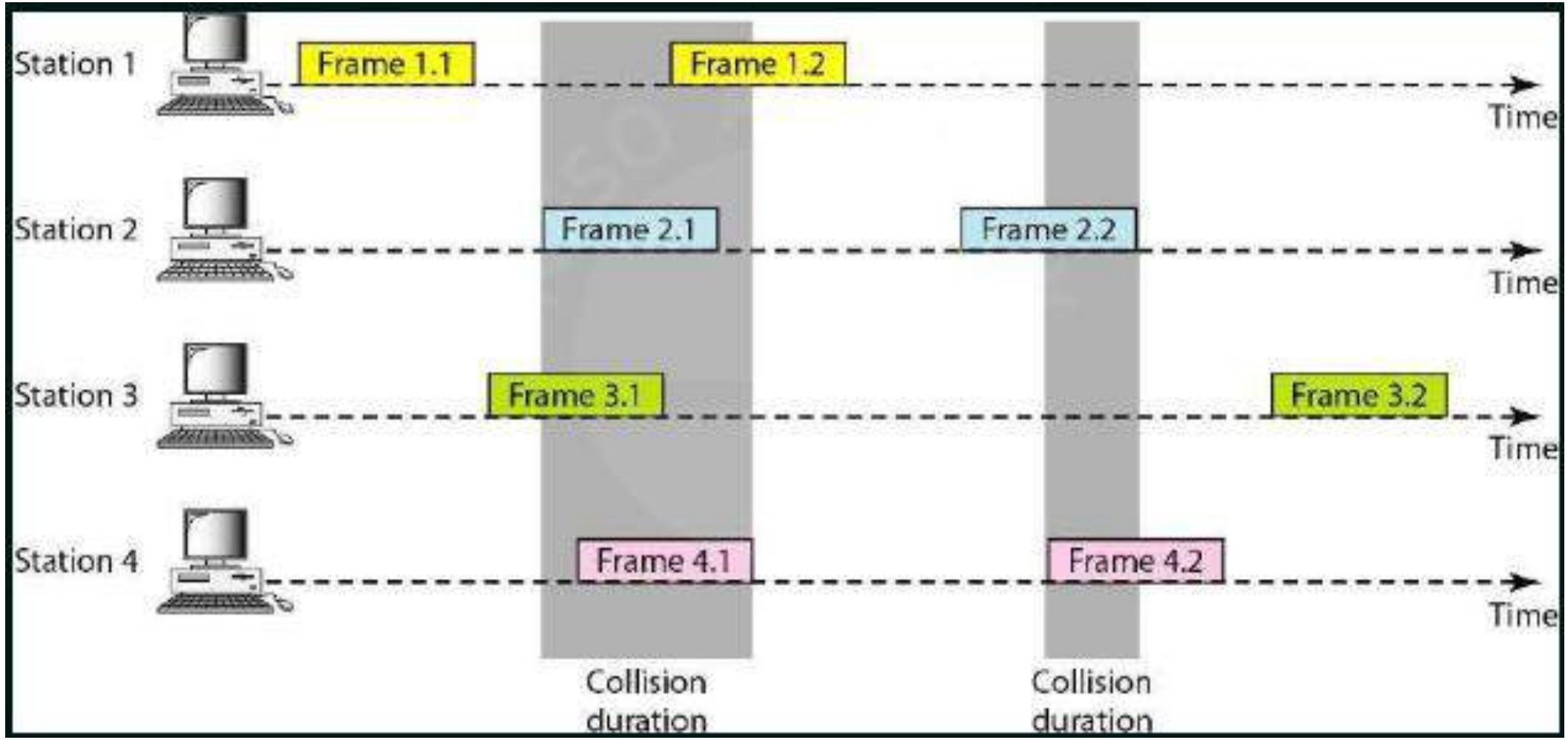
- ***ALOHA***
  - *Pure ALOHA*
  - *Slotted ALOHA*

- ***CSMA(Carrier Sense Multiple access)***
  - *1-persistent*
  - *non-persistent*
  - *P-persistent*
  - *O-persistent*

- ***CSMA/CA***
- ***CSMA/CD***

Random access Protocol

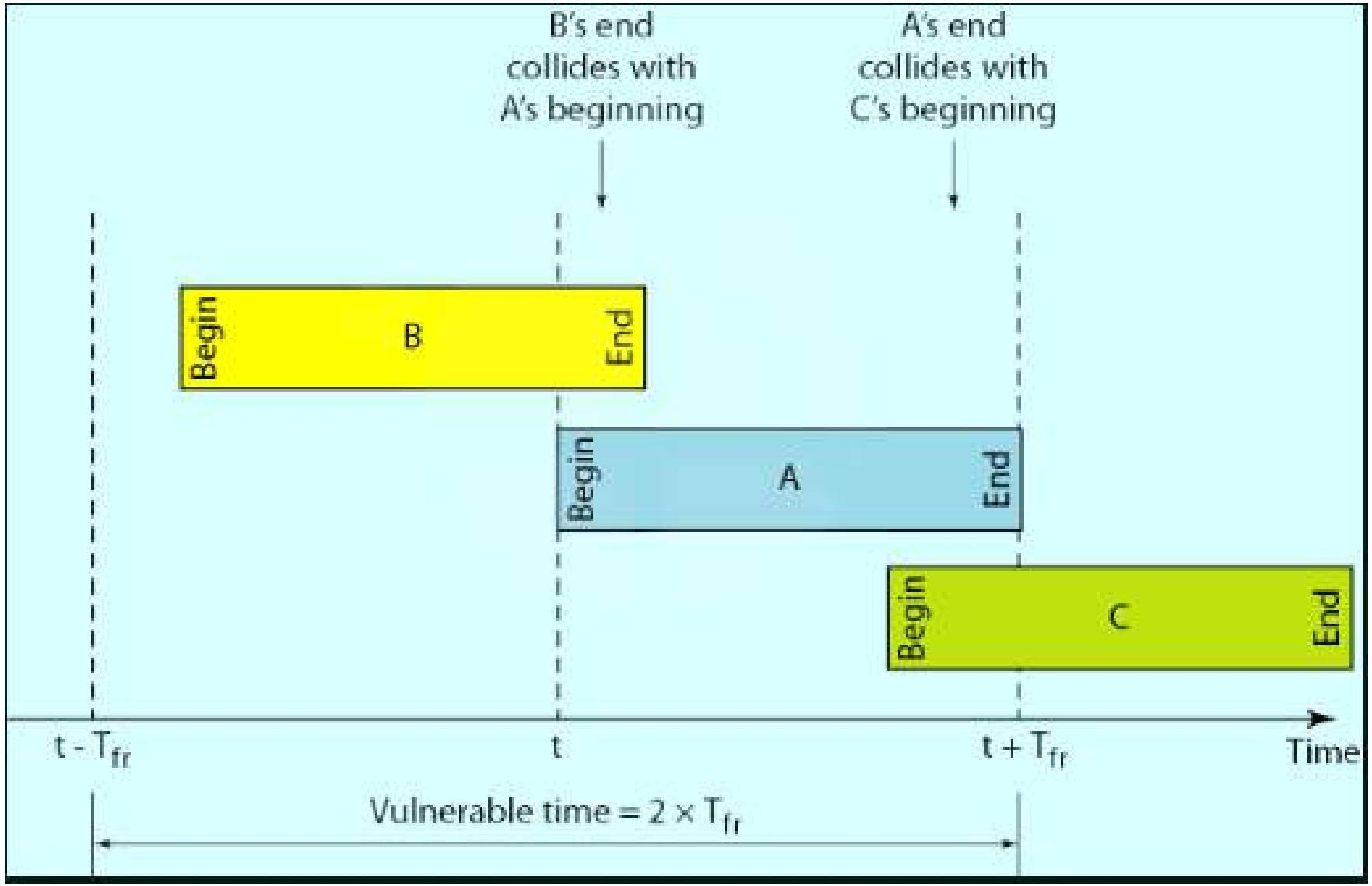
- ***ALOHA***
  - *Pure ALOHA*



Random access Protocol

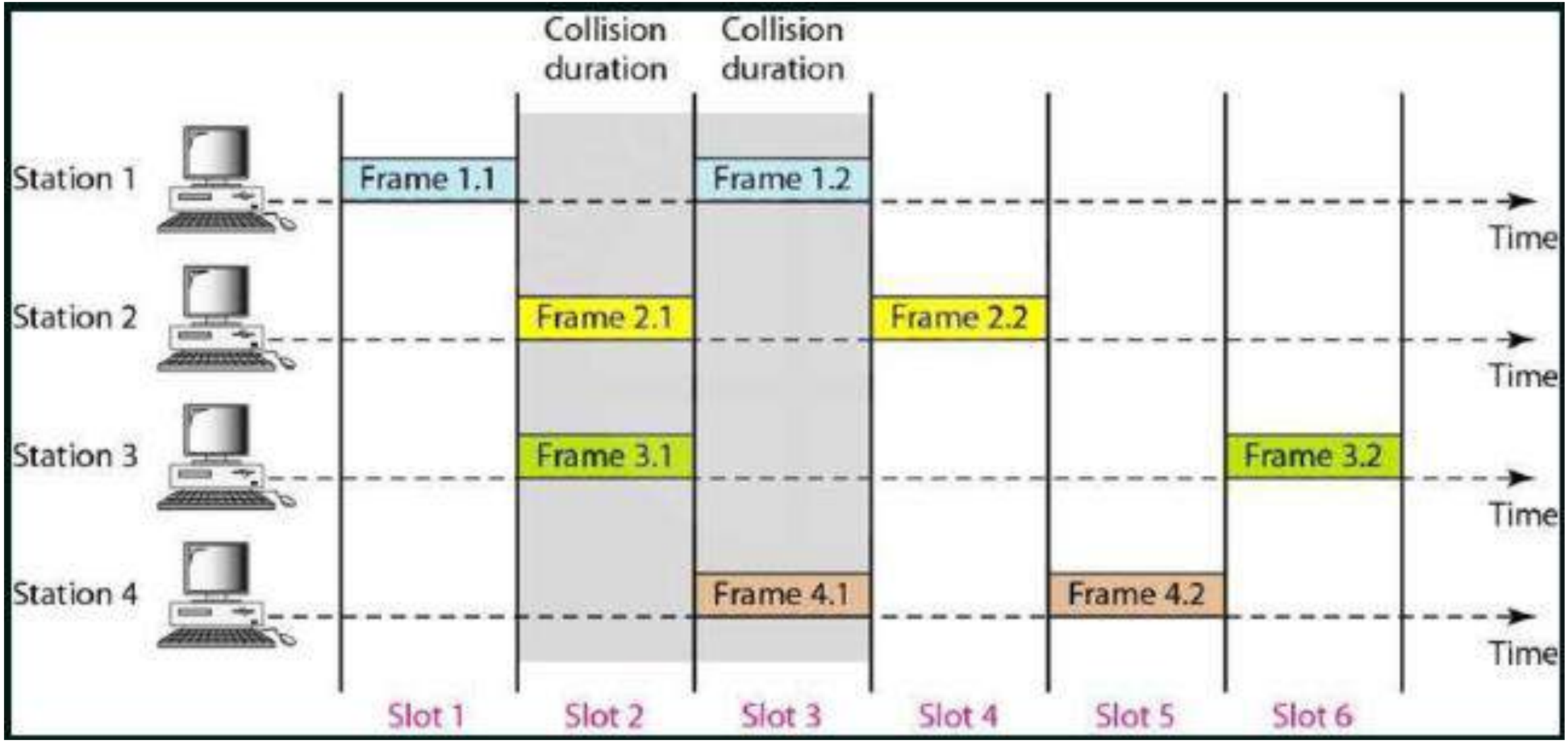
- *Pure ALOHA*

Vulnerable time $= 2 \times T_{fr}$



Random access Protocol

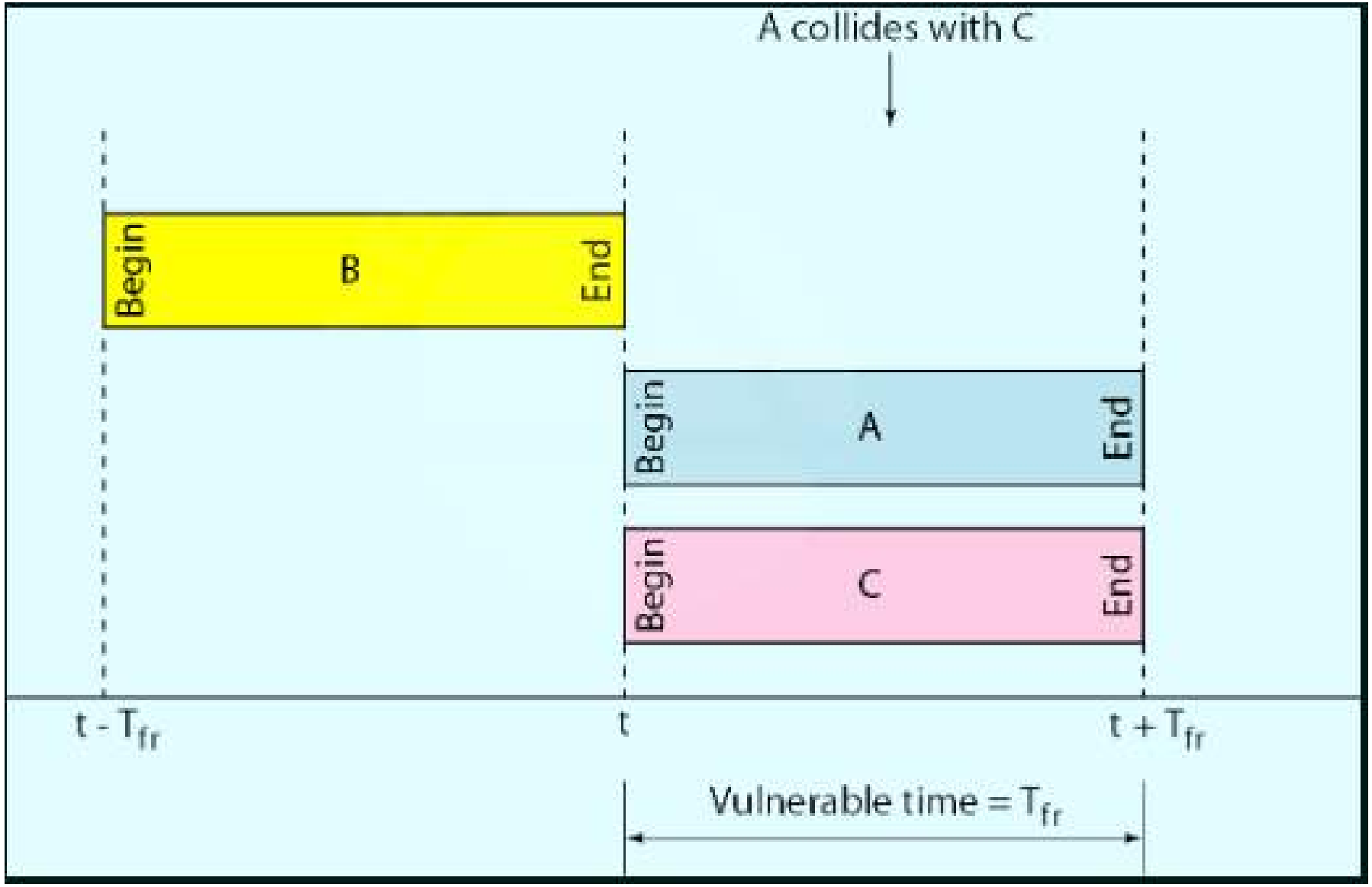
- *Slotted ALOHA*



Random access Protocol

- *Slotted ALOHA*

Vulnerable time $= T_{fr}$



Random access Protocol

Pure Aloha	Slotted Aloha
▪ Any station can transmit data at any time. ▪ The time is continuous and not globally synchronized. ▪ Vulnerable time in which Collison may occur = $2 \times T_{fr}$ ▪ Probability of successful transmission of data packet $= G \times e^{-2G}$ ▪ Maximum efficiency 18.4 % (G=1/2) ▪ Main advantage is simplicity in transmission.	▪ Any station can transmit the data at he beginning of any time slot. ▪ The time is discrete and globally synchronized. ▪ Vulnerable time in which Collison may occur = T_{fr} ▪ Probability of successful transmission of data packet $= G \times e^{-G}$ ▪ Maximum efficiency 36.8 % (G=1) ▪ Main advantage It reduces the number of collison to half and alost double the efficiency.

- ***Advantages of slotted Aloha over Channel partioning***
  - *In channel portioning (FDMA and TDMA) when only one active node to transmit , then also it will have access to apportion of the channel e.g. R/N bps not able to transmit at the full rate. In slotted Aloha if only one active user is there then it is able to transmit continuously at the full rate R bps.*

  - *But in slotted Aloha un-wasted slots are the only slot in which exactly one node transmits. The efficiency of slotted Aloha is when there are a large number of active users and each of them always having a large no. of frames to transmit.*

Module 2

Datalink Layer

Dr. Sunandita Debnath, IIIT Vadodara

- ***CSMA(Carrier Sense Multiple access)***
- *The disadvantages of pure and slotted Aloha has been overcomed in CSMA In Aloha the node's decision to transmit is made independently of the activity of the other nodes attached to the broadcast channel.*
- *CSMA is based on channel (carrier) sensing mechanism.*
  - *Listen before speaking-This is called carrier sensing, a node listens to the channel before transmitting. If a frame from another node is currently being transmitted into the channel, a node then waits until it detects no transmission for short amount of time and then begin transmission.*

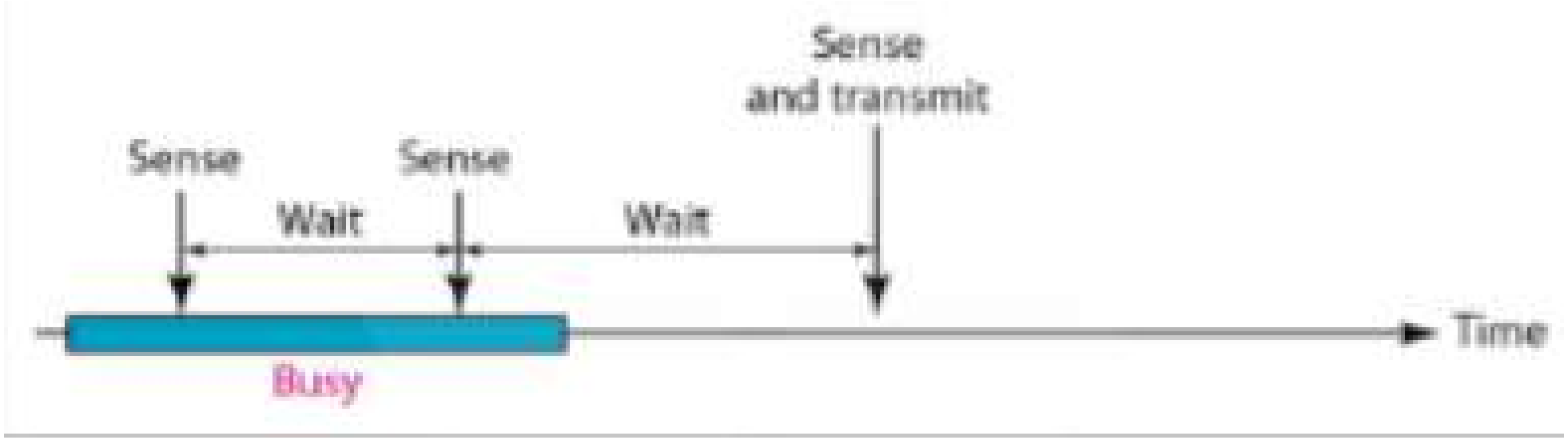
  - *If someone else begins talking at the same time, stop talking- This is called collision detection. A transmitting node listens to the channel while it is transmitting. If it detects that another node is transmitting an interfering frame, it stops transmitting and waits for a random amount of time before repeating the sense-and –transmit-when-idle cycle*

➢ *CSMA(Carrier Sense Multiple access)*

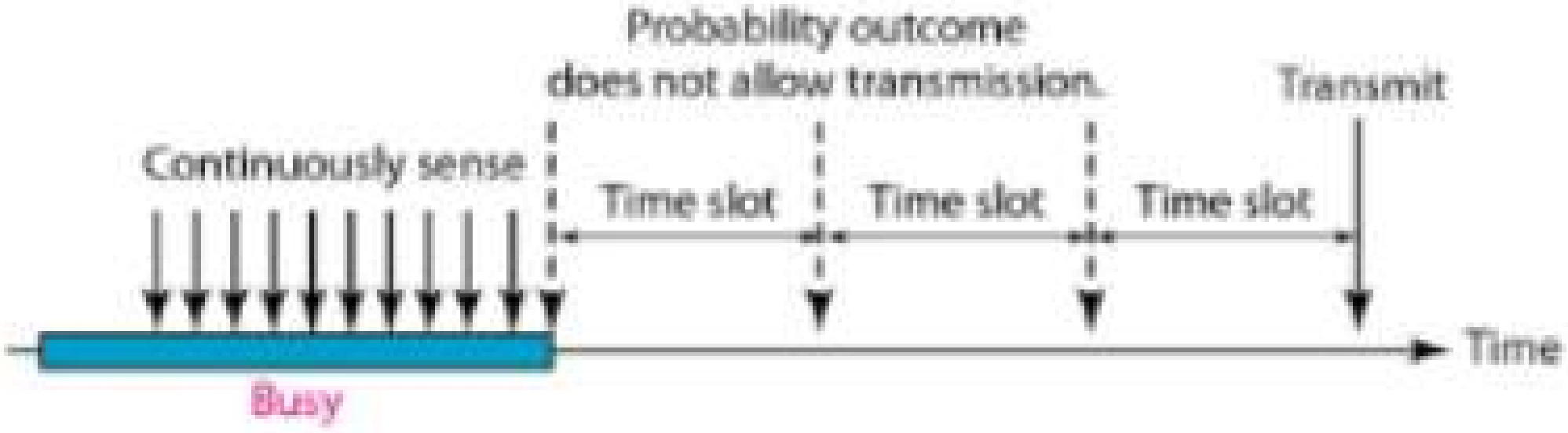
1-persistent	non-persistent	P-persistent
▪ Before sending the data, the station first listens to the channel to see if anyone else is transmitting at that moment. If the channel is sidle the station transmits a frame. ▪ If busy then it senses continuously until it becomes idle. ▪ Since the station transmits the frame with probability 1 when the channel is idle, the scheme is called 1-persistent CSMA.	▪ Before sending a frame, a station senses the channel . If no one else is sending the station begins sending. ▪ However, if a channel is busy, the station does not continuously sense it for the purpose of seizing immediately upon detecting the end of previous transmission. ▪ Instead it waits for a random amount of time and then repeats the algorithm. ▪ Consequently this algorithm leads to better channel utilization but longer delays than 1-persistent CSMA.	▪ It applies to a slotted channel. When a station becomes ready to send it senses the channel. ▪ If it is sidle it transmit with the probability P. and it defers with a probability $Q = (1 - P)$ until the next slot. ▪ If that slot is idle , it either transmits or defers again with probability P & Q. This process is repeated until either the frame is has been transmitted or any other station has begin transmission. ▪ In the later case the unlucky station acts like there had a collision and waits for random

1-persistent

non-persistent



P-persistent

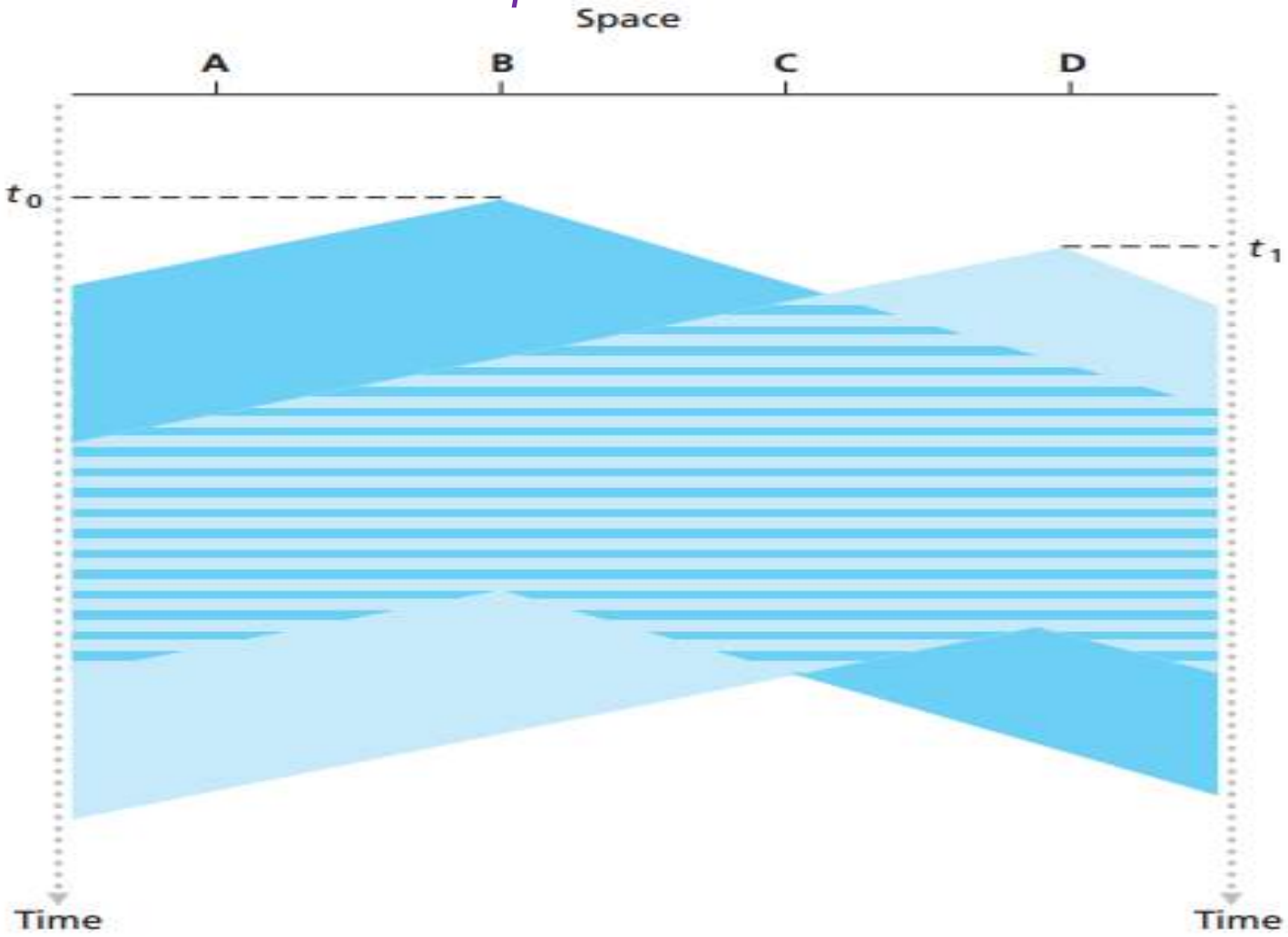


O-persistent

Each node is assigned a transmission order by a supervisory node

The effect of propagation delay in CSMA

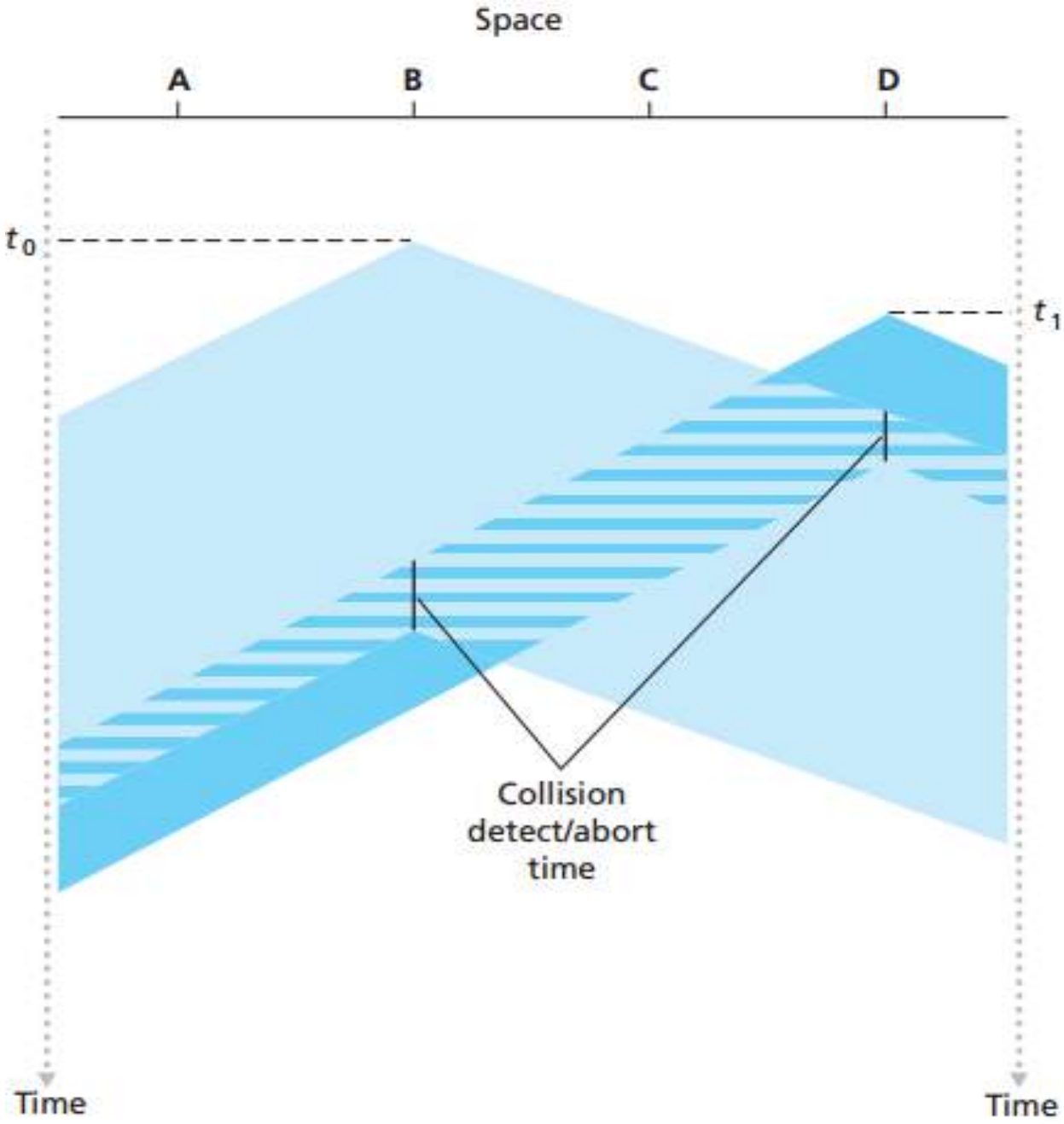
➢ *Sender B started transmission at t_0 as it senses the medium is idle. As the propagation delay is nonzero at time t_1 ($t_1 > t_0$), node D senses the medium is idle and begin its transmission and after a short interval of time B's frame interfere with D's frame. So longer the end-to-end propagation delay more will be the chances of collision.*



CSMA(Carrier Sense Multiple access)

CSMA/CD

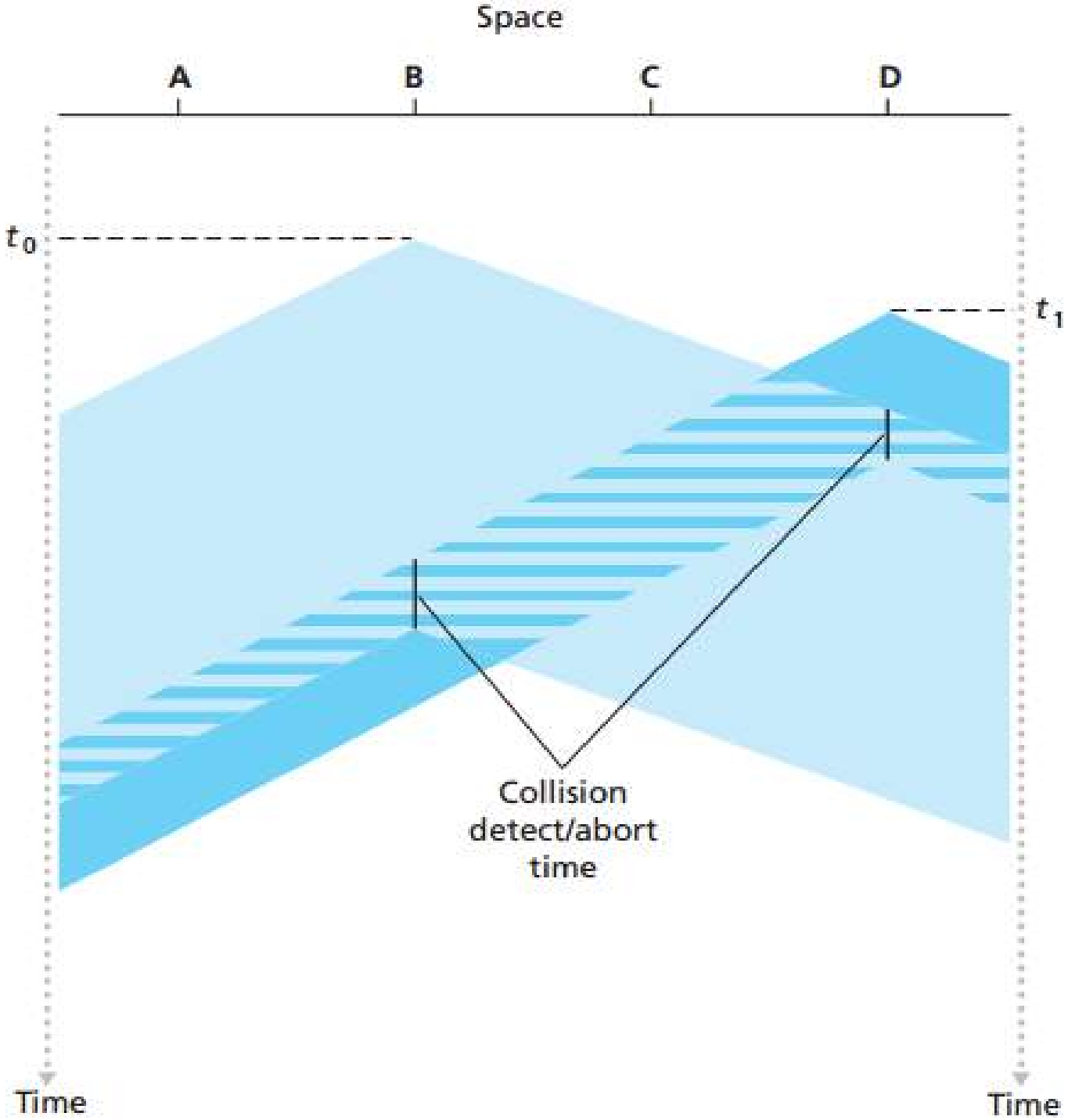
- *If two stations senses the channel is idle and begin transmitting simultaneously, then both will detect collision almost immediately.*
- *As soon as it detects a collision both the sender ceases transmission. Quickly terminating damaged and corrupted frames from transmitting saves time and bandwidth.*
- *CSMA/CD is used in LANs and Ethernet*



CSMA(Carrier Sense Multiple access)

CSMA/CD

- After aborting the transmission. Both the senders wait for a random amount of time.
- If both the stations waits for the same amount of time then again they will collide and continue colliding forever.
- Two cases to be considered:
  - If the interval is large and number of colliding node are small, nodes are likely to wait a large amount of time before repeating the transmission again.
  - If the interval time is small and large number of colliding nodes are there then the randomly chosen time to be nearly close and the nodes will collide again.



CSMA(Carrier Sense Multiple access)

Binary Back-off Algorithm for CSMA/CD

- *It is used in Ethernet.*
- *Suppose a frame has already experienced n collisions, then this sender will chooses the vale K at random K belongs to* $(0,1,2 \ldots\ldots 2^n - 1)$.
- *Thus the more numbers of collision the larger the interval from which the K value will be chosen.*
- *Suppose a sender attempts to frame for the first time and detects a collision then,* $n = 1, so\ 2^1 - 1 = 1$. *Therefore the range of* $k \in (0,1)$. *So it chooses K =0 with probability ½ and chooses 1 with probability ½.*
- *Thus, the size of the sets from which K is chosen grows exponentially with the number of collisions; for this reason this algorithm is referred to as binary exponential backoff.*
- *The efficiency of CDMA/CD is given by*

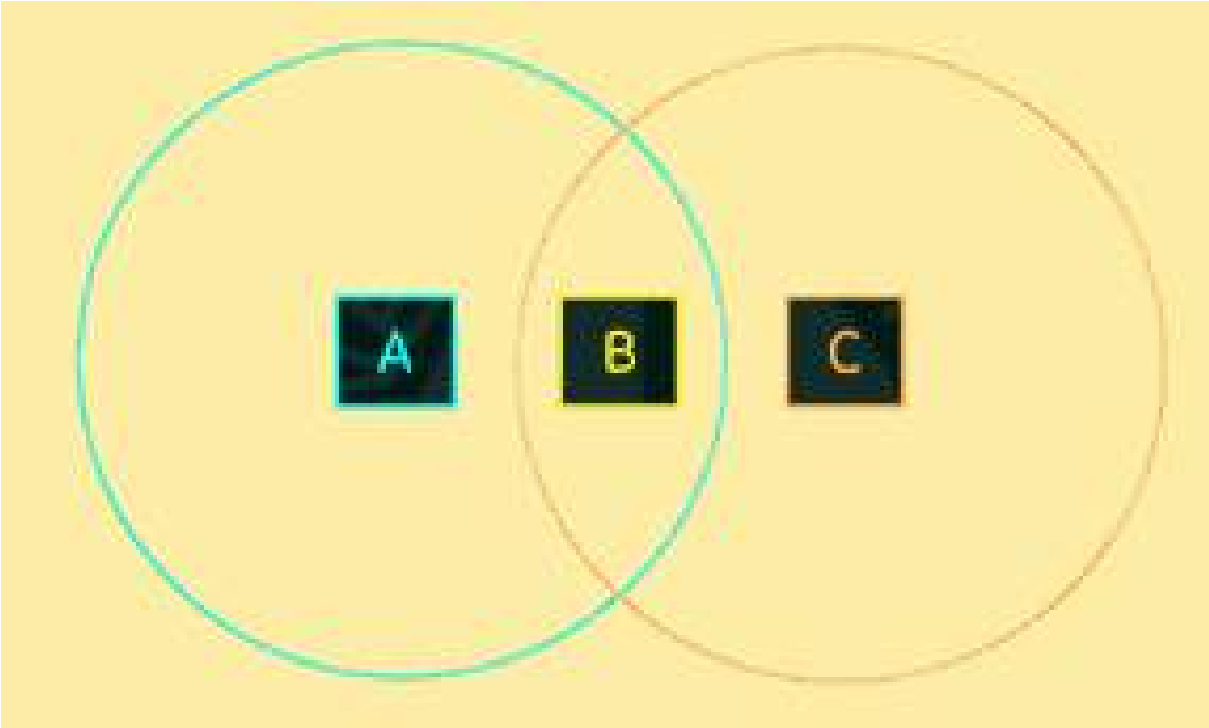
$$\boldsymbol{Efficency = \frac{1}{1 + 5d_{prop}/d_{trans}}}$$

CSMA(Carrier Sense Multiple access)

CSMA/CA (Collision Avoidance)

- *It is used for wireless communication.*
- *CSMA/CA is a network multiple access method in which carrier sensing is used, but nodes attempt to avoid collisions by beginning transmission only after channel is sensed to be idle.*
- *I is particularly important for wireless networks. Where the collision detection of the alternative CSMA/CD is not possible due to wireless transmitters desensing their receivers during packet transmission.*
- *CSMA/CA is unreliable because of hidden node problem and exposed terminal problem.*
- *Solution is RTS (request to send)/CTS (clear to send)*
- *The access method used in IEEE 802.11 (WiFi) is CSMA/CA.*

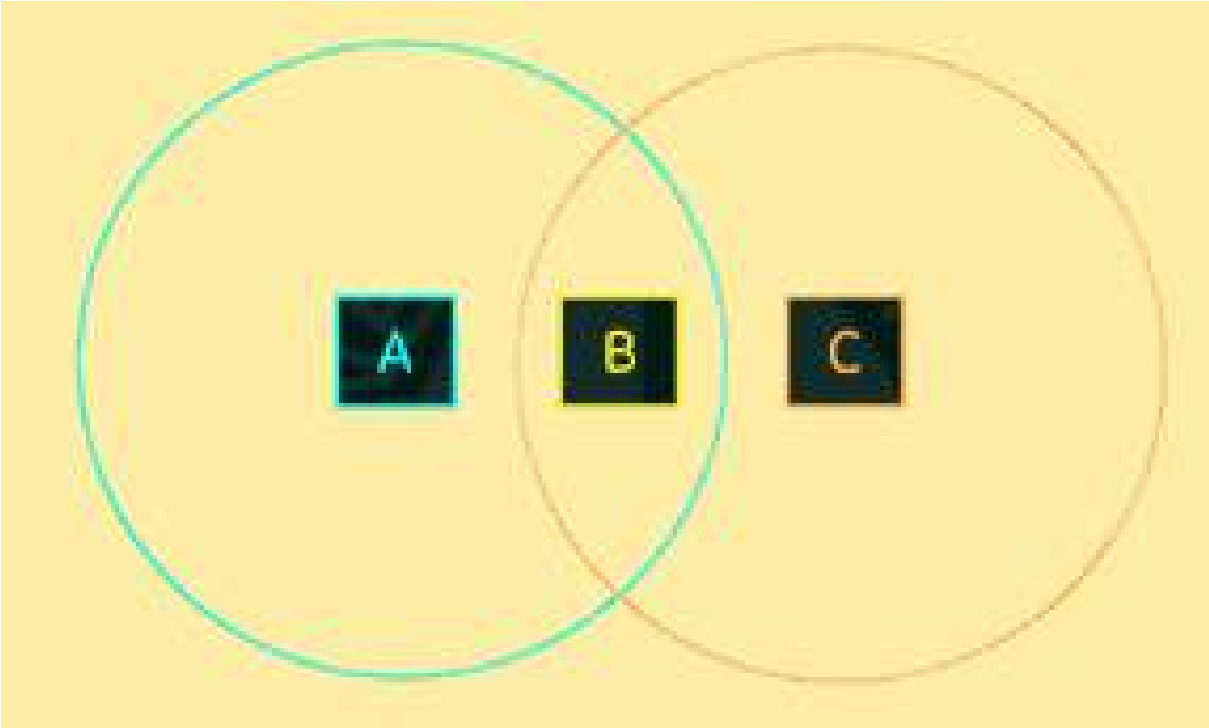
Hidden Terminal Problem



Suppose both A and C want to communicate with B and so they each can send it a frame

- *A and C are unware of each other since their signals do not reach that far in other words they are not in the sensing range of each other.*
- *These two frames collide with each other at B (but unlike Ethernet neither A or C is aware of this collision).*
- *A and C are hidden nodes with respect to each other.*

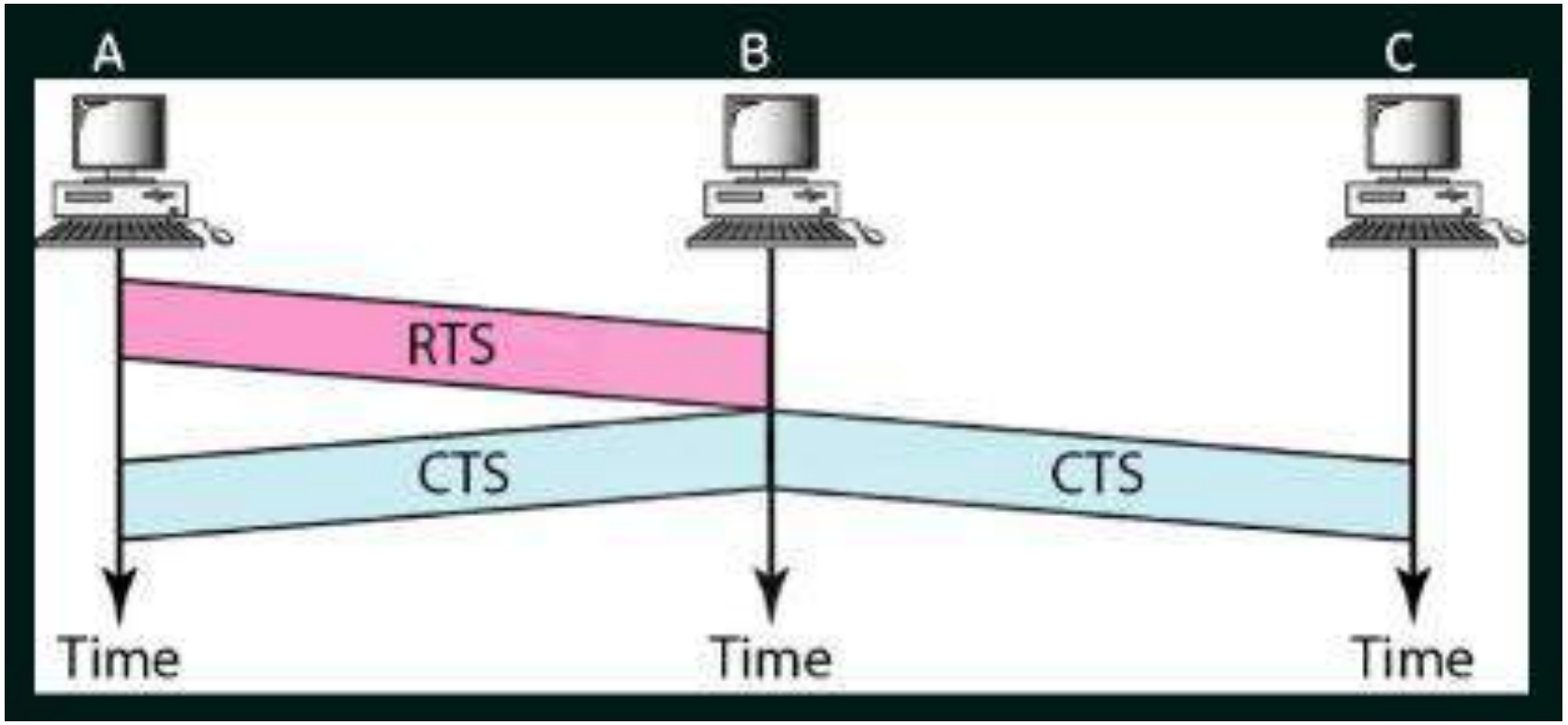
MACA Protocol



This hidden node problem is solved using MACA protocol. Suppose both A and C want to communicate with B and so they each can send it a frame

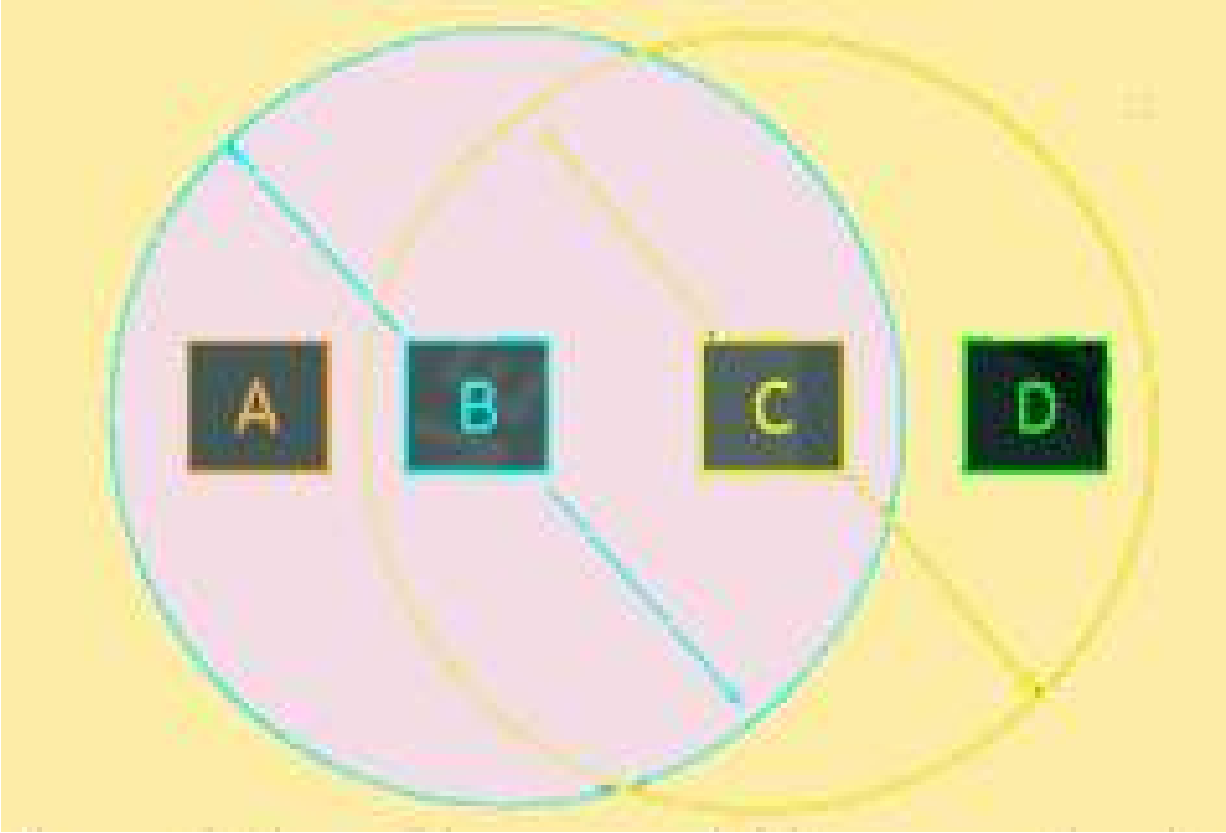
- *A and C are unware of each other since their signals do not reach that far in other words they are not in the sensing range of each other.*
- *These two frames collide with each other at B (but unlike Ethernet neither A or C is aware of this collision).*

MACA Protocol



Station C doesn't hear RTS from A, but it does hear CTS from B, so it know something is up.

Exposed Terminal Problem



Suppose B is sensing to A. Node C is aware of this communication because B is in range of C i.e. it can hear B's transmission.

- *It would be mistake to C to conclude that it cannot transmit to D or anyone just because it can hear B's transmission.*
- *IF C wants to transmit to D , this is not a problem since C's tranmissison to D will not interfere with A's ability to transmit or receive from B.*

CSMA/CA –Two algorithms are addressed in IEEE 802.11 with:

- ***MACA (Multiple Access with Collision Avoidance)***
- ***MACAW (MACA for Wireless LANs)***

MACA

- *In MACA sender and receiver exchange control frames with each other before the sender actually transmits any data.*
- *This exchange informs all nearby nodes that a transmission is about to begin*
- *Sender transmits a RTS (Request to Send) frame to the receiver*
  - *RTS frame includes a field that indicate how long the sender wants to hold the medium/ channel i.e. the length of the data frame be to transmitted.*
- *Receiver replies with a CTS (Clear to Send) frame*
  - *This CTS frame is broadcast by the receiver and also includes the length of the frame which is informed by the sender while asking for request to send.*
- *Any node which hears the CTS frame knows that it is close to the receiver, therefrom refrained from transmitting (i.e. cannot transmit) for the period of time it takes to send a frame of the specified length.*

MACA

- *Any node that hears the RTS frame but not the CTS frame it get to know that it is not close enough to the receiver to interfere with it and so free to transmit to any other node.*

MACAW

- *The idea of using CK (acknowledgement)A in MACA is proposed in MACAW: MACA foe Wireless LANs.*
- *Receiver sends a ACK to the sender after successfully receiving a frame.*
- *As this ACK is also a broadcast message, all nodes must wait till this ACK before trying to begin a new transmission.*
- *If two or more nodes detects an idle link and try to transmit an RTS at the same time*
  - *This RTS frame will collide with each other*
  - *802.11 does not support collision detection*
  - *So the sender realize the collision has happened when they are not receiving an CTS after a period of time from the receiver node.*
  - *In this case they both wait a random period of time before trying again*
  - *The amount of time a given node must wait is defined by same the exponential backoff algorithm used on the Ethernet.*

Flow control

Protocols

Noiseless channels

* Simplest
* Stop-and-wait

Noisy channels

*stop-and-wait ARQ

*Go-back-N-ARQ

*Selective Repeat ARQ

Stop and wait protocol

- *Stop-and-wait protocol is data link layer protocol for transmission of frames over noiseless channels.*
- *It provides unidirectional data transmission with flow control facilities but without error control facilities.*
- *After transmitting one frame, the sender waits for an acknowledgement before transmitting the next frame.*

- ***Sender side :***

Rule 1: send one data frame at a time

Rule 2: Send the next frame only after receiving ACK for the previous one.

- ***Receiver side:***

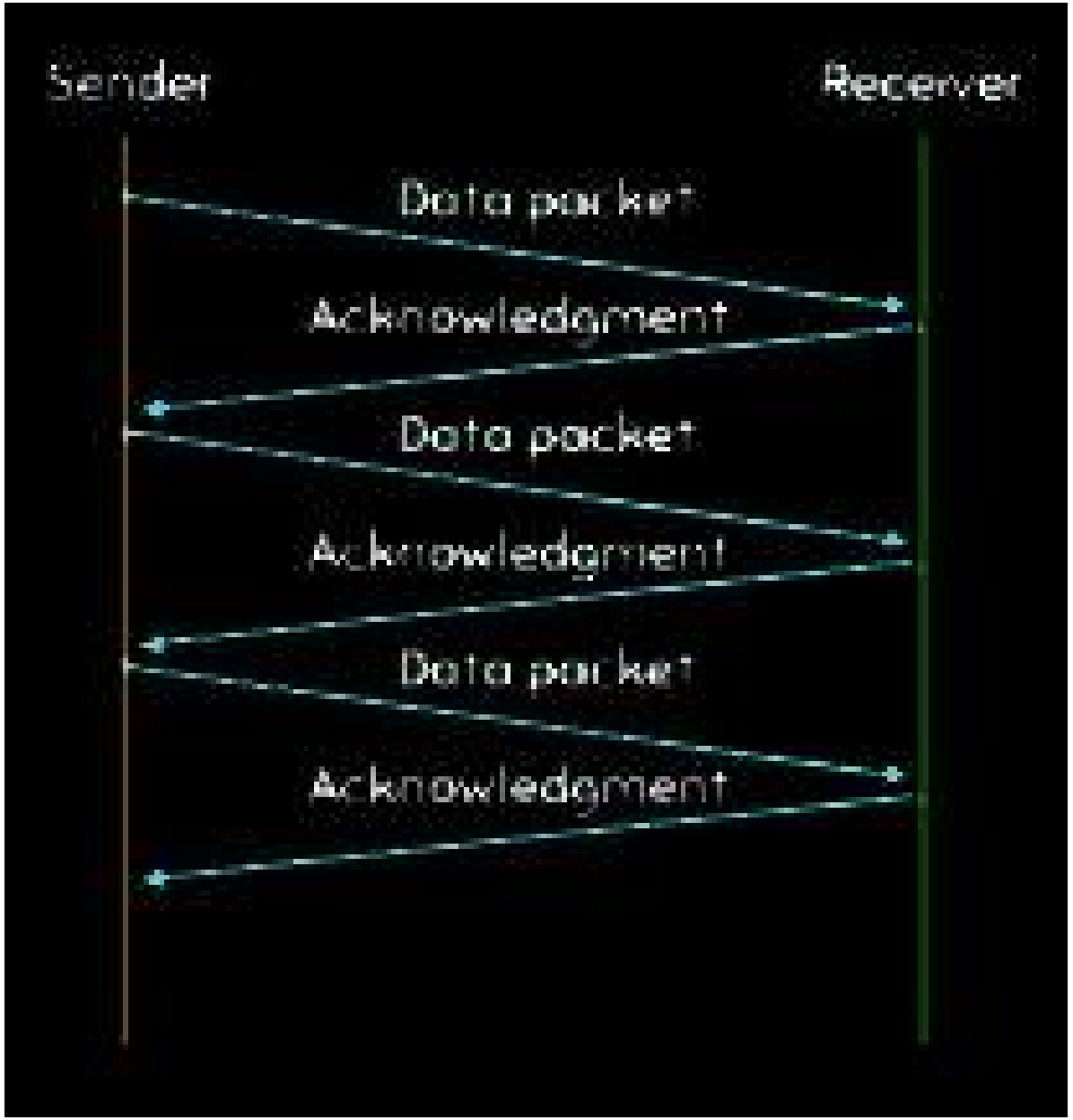
Rule 1: Receive and consume frame send one data frame at a time

Rule 2: After consuming frame ACK need to be sent (Flow control)

Stop and wait protocol

The disadvantage of this protocol is if the sender has a large number of frames say 1000 frames then it has to send only one frame at a time and waits for the ACK each time.

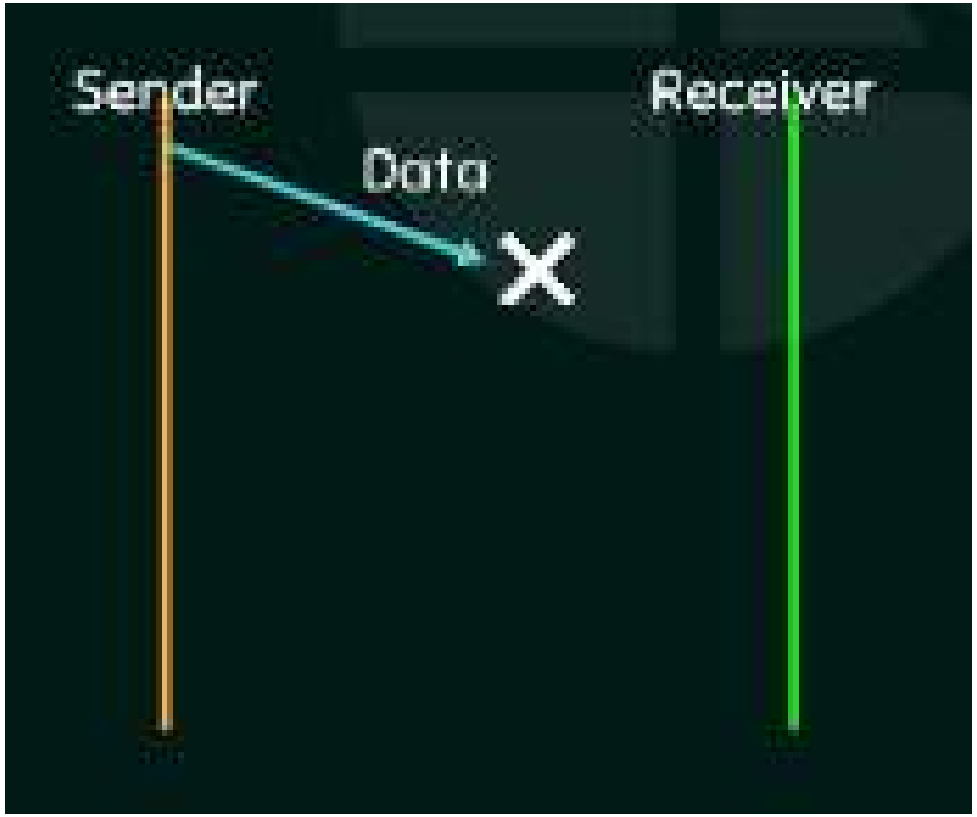
- *Problem due to lost frame*
- *Problem due to lost ACK*
- *Problem due to delayed frame/ACK*



Stop and wait protocol

Problem due to lost frame

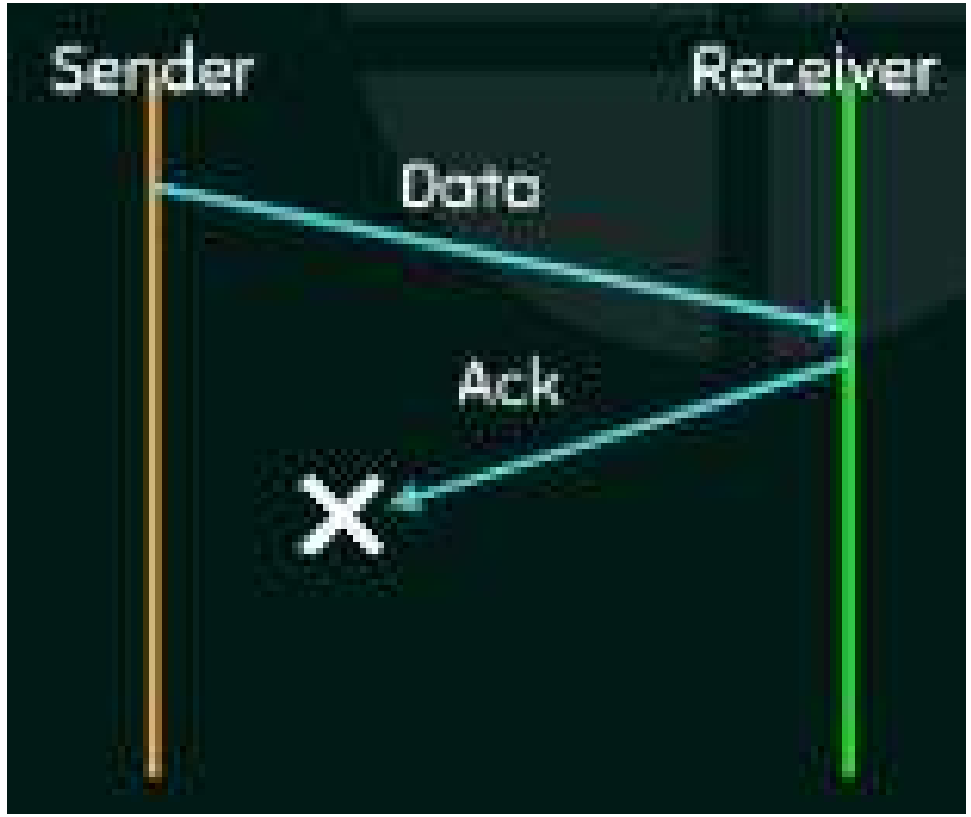
- *Sender waits for ACK for an infinite amount of time*
- *Receiver waits for data for an infinite amount of time*



Stop and wait protocol

Problem due to lost ACK

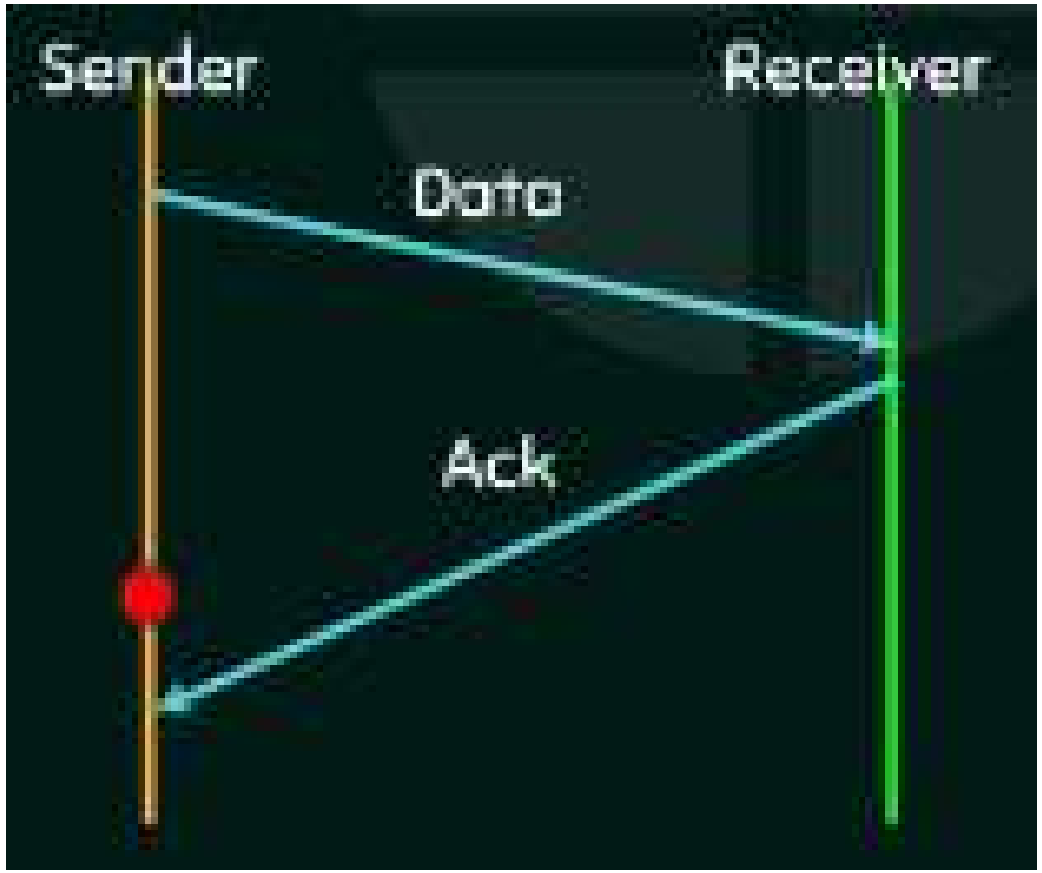
- *Sender waits for an infinite amount of time for ACK. So there is an infinite delay for the next frame in the queue to be transmitted.*



Stop and wait protocol

Problem due to delayed frame/ACK

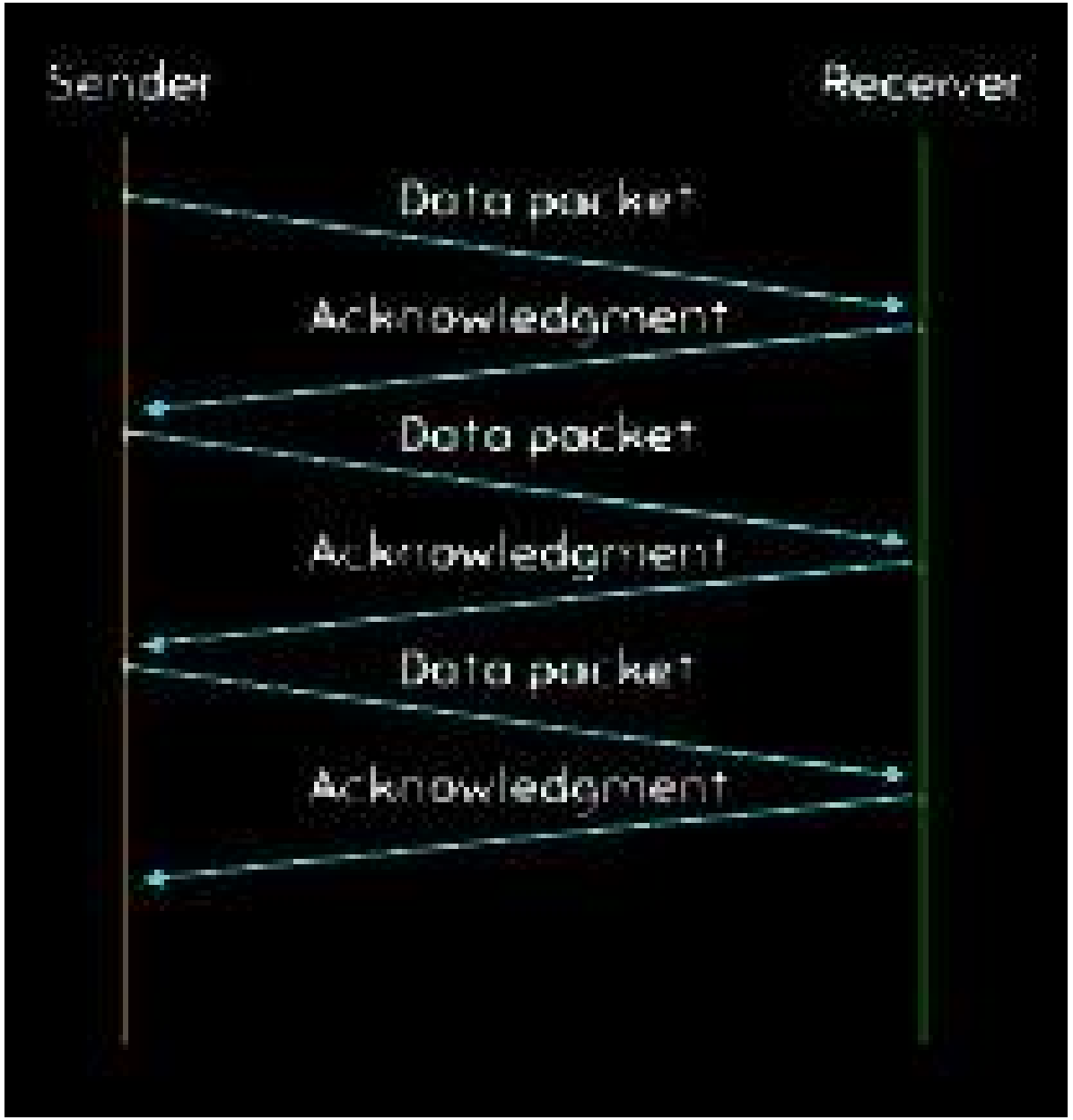
- *After timeout at sender side, a delayed ACK might be wrongly considered as ACK of some other data frames.*



Stop and wait protocol

The disadvantage of this protocol is if the sender has a large number of frames say 1000 frames then it has to send only one frame at a time and waits for the ACK each time.

- *Problem due to lost frame*
- *Problem due to lost ACK*
- *Problem due to delayed frame/ACK*

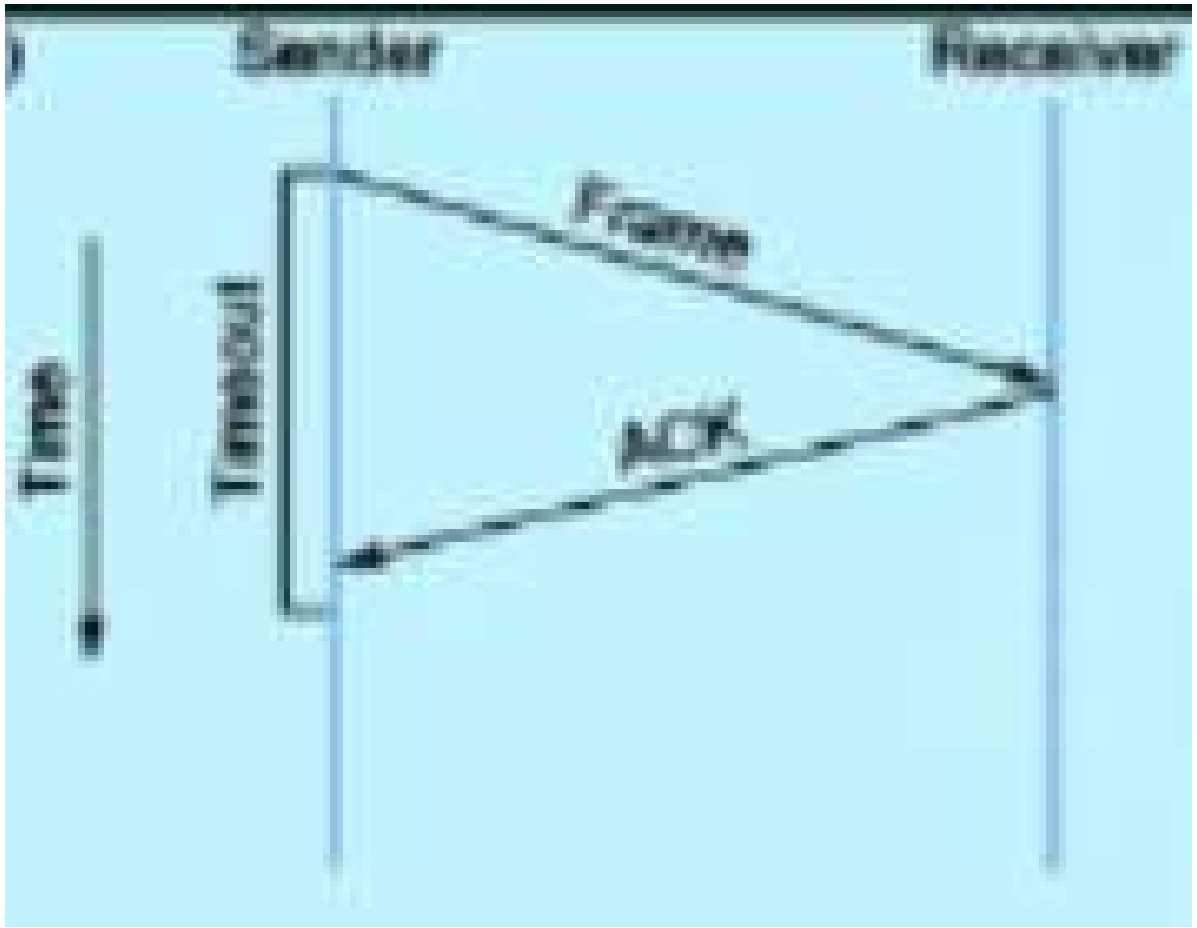


Stop and wait ARQ protocol

- *After transmitting one frame, the sender waits for an acknowledgement before transmitting the next frame.*
- *If the acknowledgement does not arrive after a certain period of time the sender times out and retransmits the original frame.*
- **Stop and wait ARQ = Stop and wait + Timeout Timer +Sequence number**

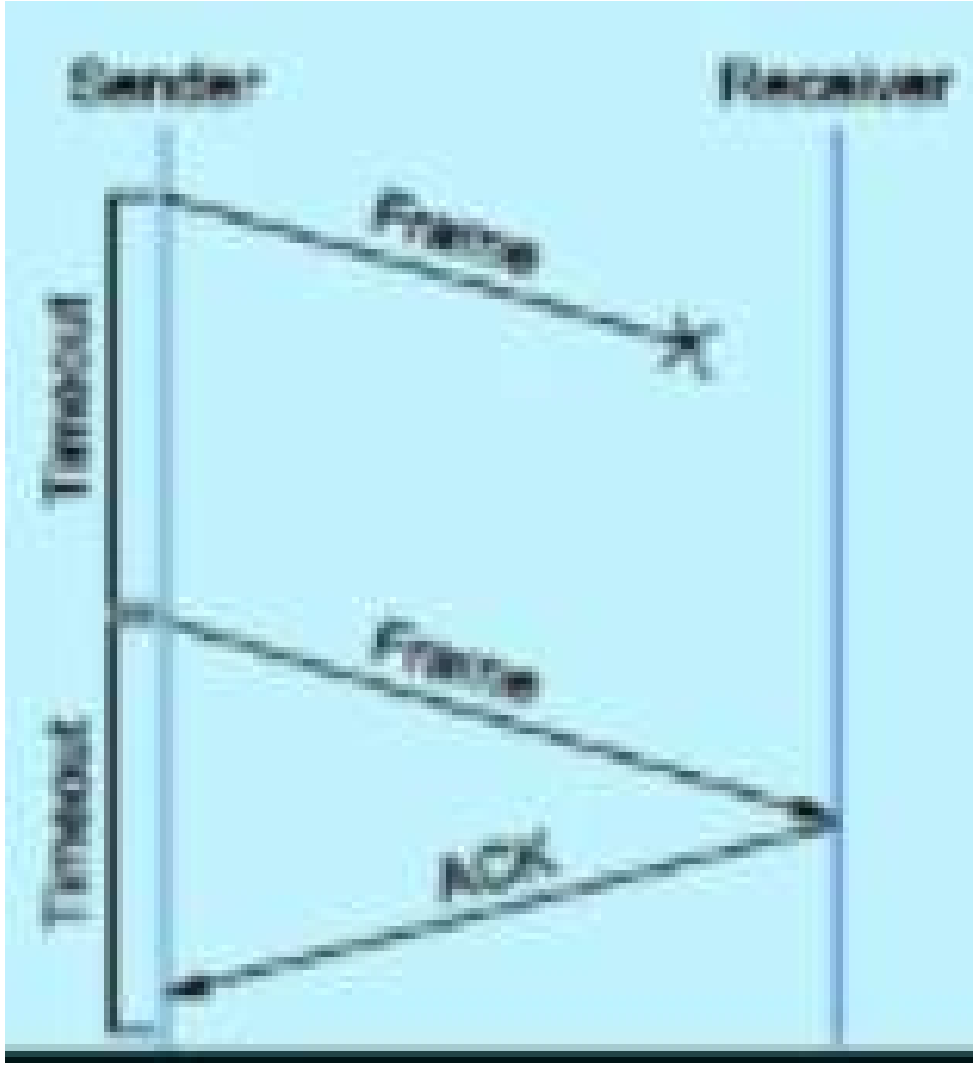
Stop and wait ARQ protocol

ACK is received before the timer expires



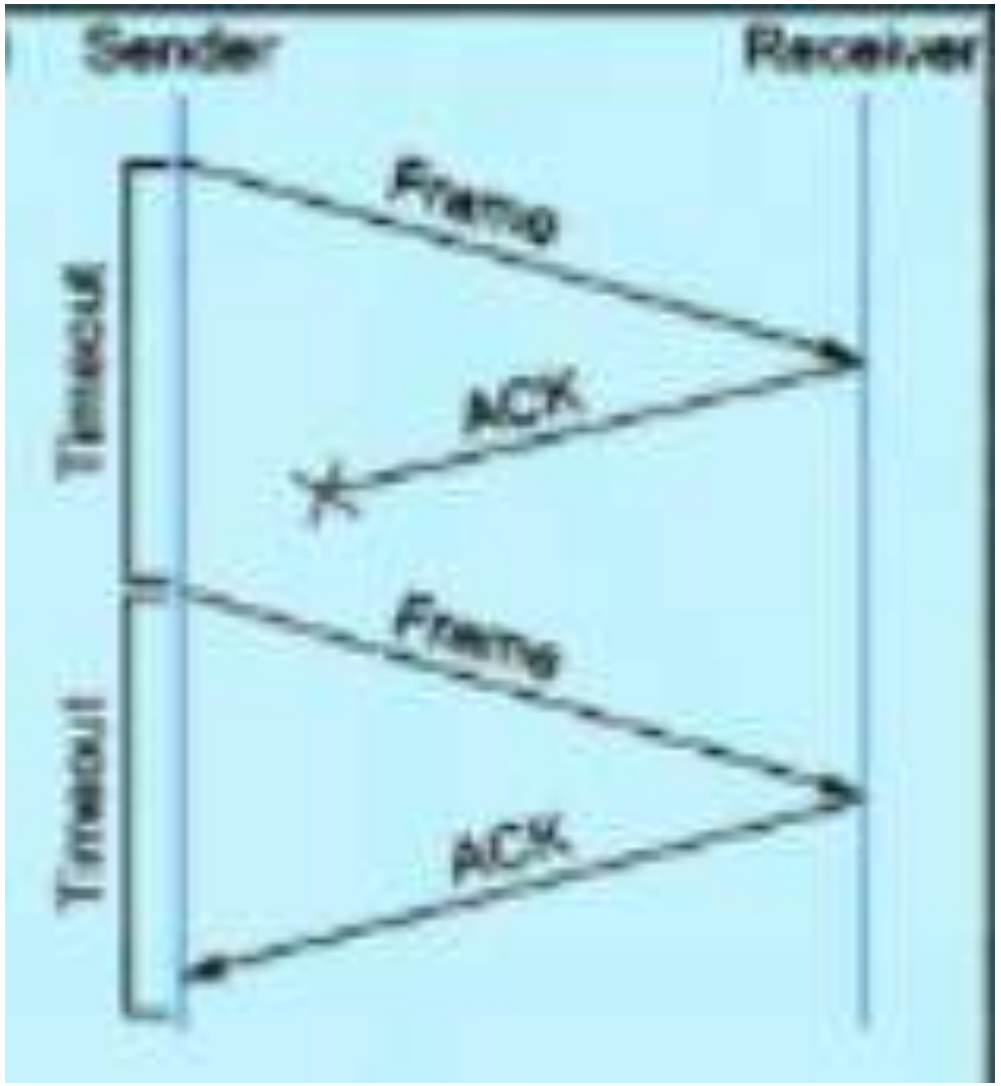
Stop and wait protocol

The original frame is lost



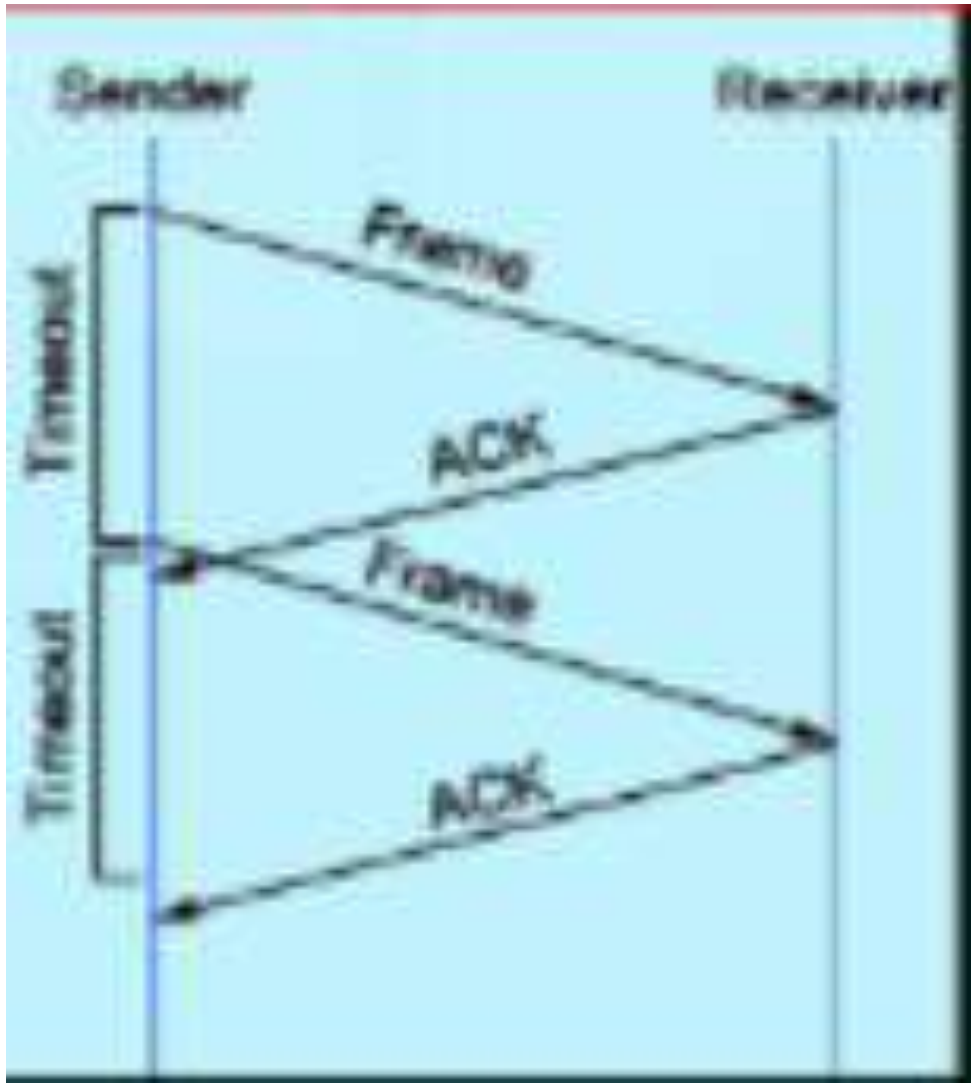
Stop and wait ARQ protocol

ACK is lost



Stop and wait protocol

The timer fires too soon



Sliding window Protocol

- *Go-back-N ARQ*
- *Selective Repeat Request*

The disadvantages of stop-and –wait ARQ protocol is

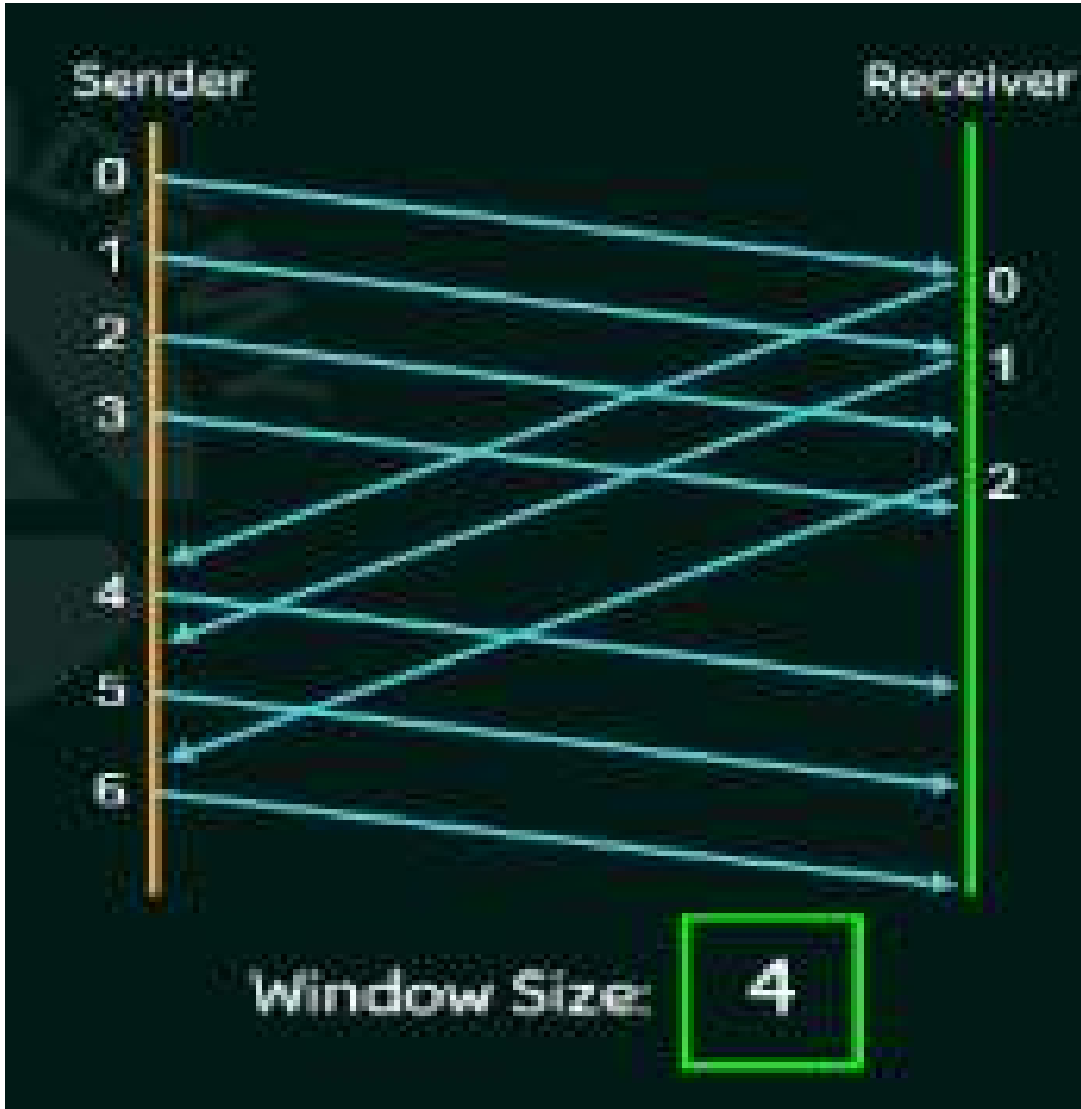
- *One frame at a time*
- *Poor utilization of bandwidth*
- *Poor performance*

The advantages of sliding window protocol is

- *Send multiples frames at a time*
- *Number of frames to be send is based on the window size*
- *Each frame is numbers i.e. each frame is associated with a sequence number.*

Working of sliding window protocol

- *Supposes sender has 11 no. of frames to transmit.*
- *Here 0 to 10 depicts the sequence nos. of the frames.*
- *Window size of the sender N = 4.*
- ***Step 1:** Sender will send first four frames at ones.*
- ***Step 2:** After getting the ACK from the 0th frame, the window will slide by one position on the left side. And the sender now transmits the 4th frame. Similarly after getting the acknowledgment from the 2nd frame the window slides on the left by one position*
- *So, there are three categories of frames in the buffer, 0 to 2nd are frames which are **sent and got acknowledged**, 3 to 6th frame are **sent and waiting for acknowledgement**, and 7 to 10th frame are **not yet sent and in the queue**.*



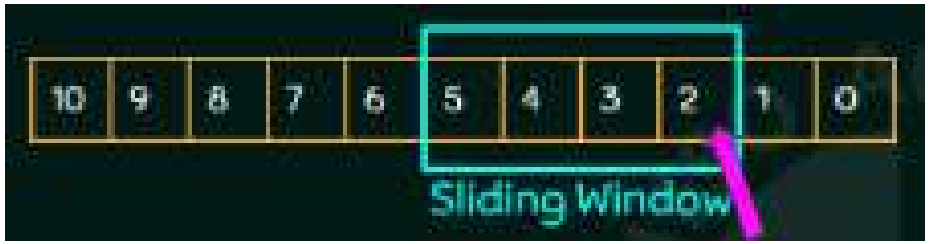
Go-Back-N ARQ protocol

- *Here N represents the window size, suppose N =4.*
- *Supposes sender has 11 no. of frames to transmit.*
- *Here sequence nos. of the frames depends on the window size*

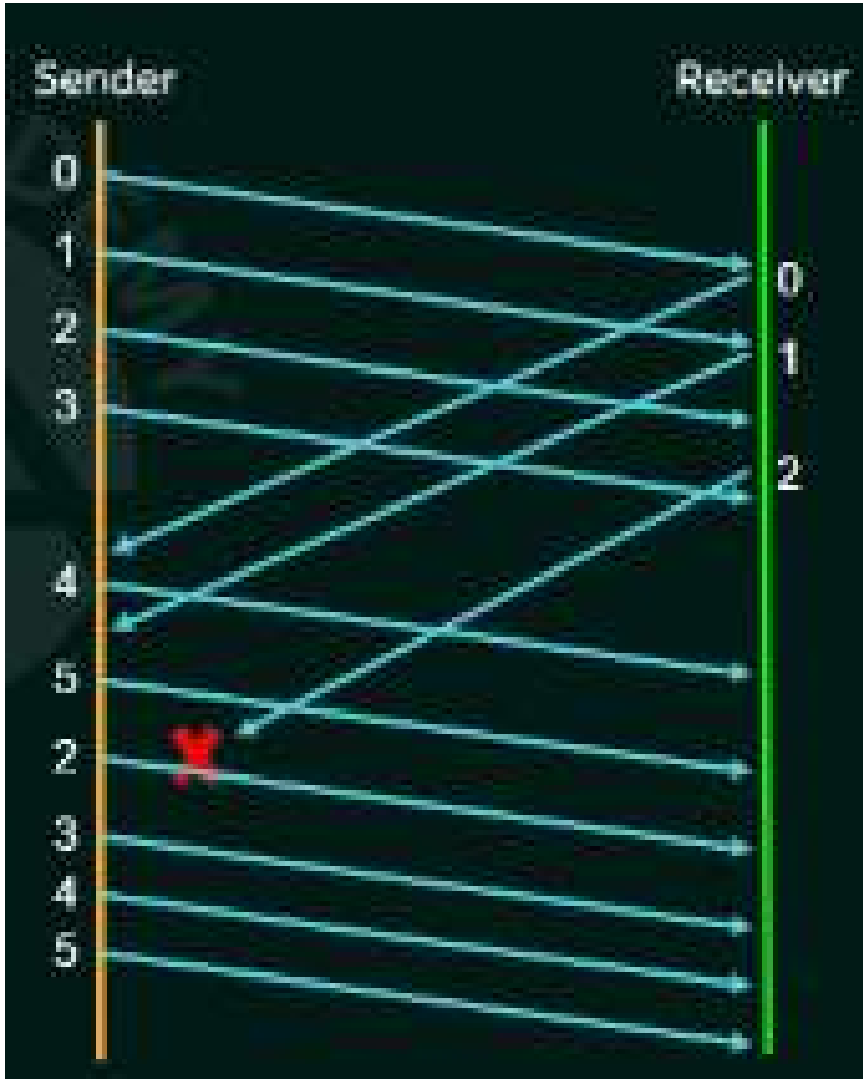
$$N = 4 = 2^2$$

Then the sequence nos. will be 0,1,2,3. (This power represents the number of binary bits required to represents the sequence nos. in its decimal equivalent). The max sequence no. can go up to $\{0,1,2 \ldots\ldots 2^n - 1\}$.

- *In Go-Back-N ARQ protocol the sender can send multiple frames before receiving acknowledgement from the first frame.*
- *The nos. of frames can be sent depends on the window size.*
- *If the ACK frames is not received within an agreed period of time, all frames in the current window is retransmitted*



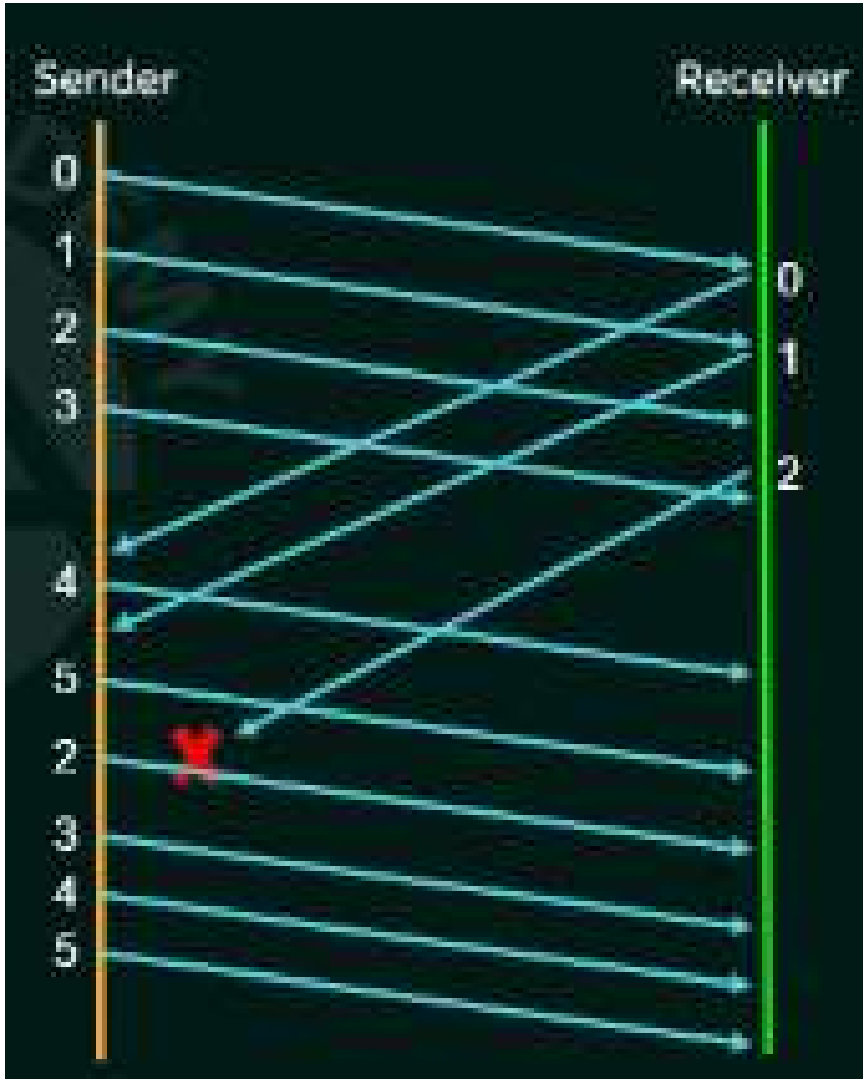
Go back to 2



Go-Back-N ARQ protocol

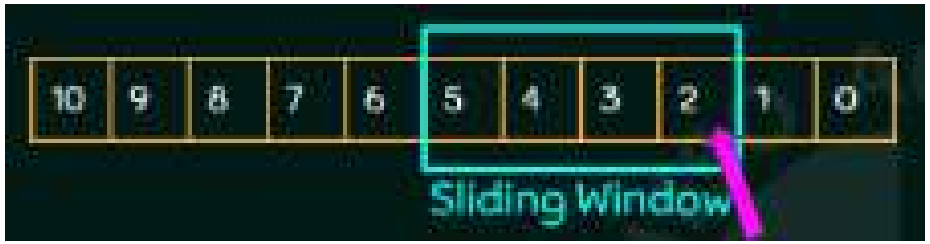
- *Supposes sender has sent 4 frames at ones 0,1,2, 3 and also got acknowledgment for the 0th frame.*
- *Now the window is holding 1, 2,3 and 4th frames. And the sender also got acknowledgment for the frame with sequence no. 1. Now the window of size 4 is containing frames 2, 3, 4 and 5.*
- *If the acknowledgement from the frame with sequence no.2 is not received then within a finite period of time. Then the sender times out again will transmit all the frames in the window starting from the frame with sequence no. 2. (i.e. 2,3,4,5) again.*
- *That's why it is called Go-Back ARQ protocol, because if one frame in the window is not acknowledged then the all the frame in the window will be retransmitted again.*

Go back to 2

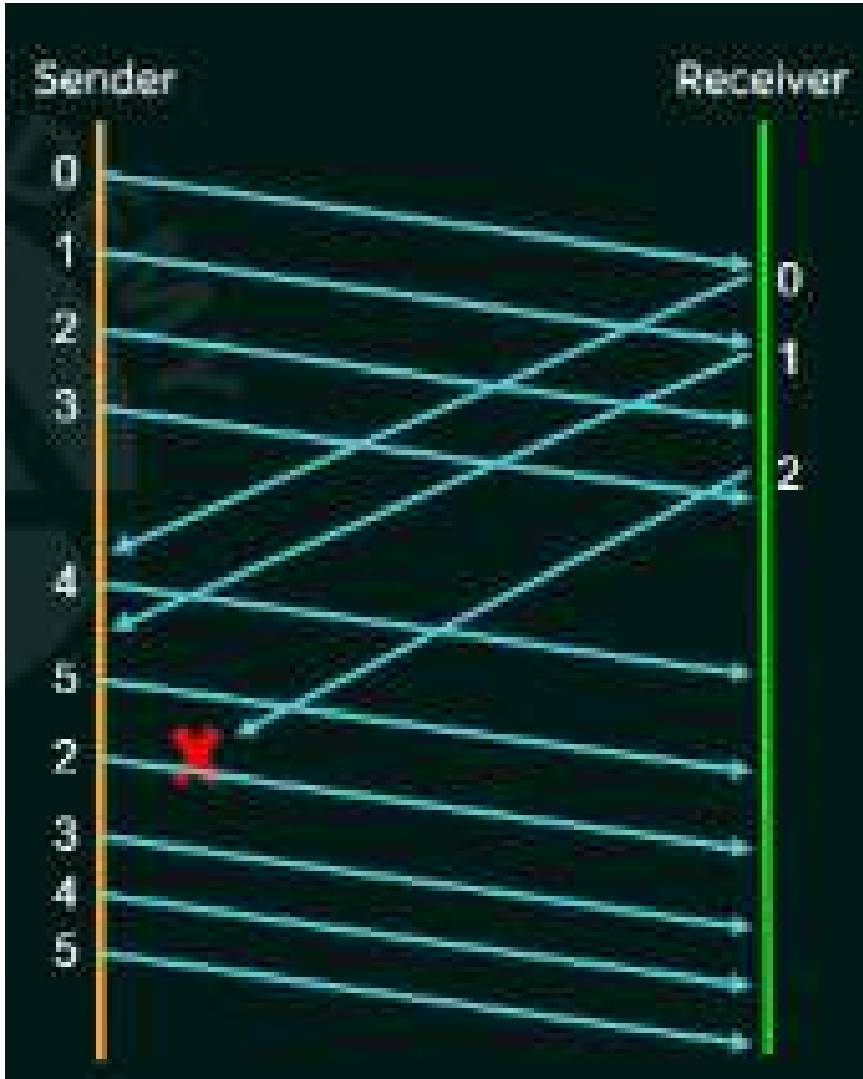


Go-Back-N ARQ protocol

- *Sender window size = N= (1+2a) for*
- *Receiver window size =1*
- $Efficiency = \frac{N\ (window\ size\ of\ sender)}{1+2a}$
- $Where\ a\ = \frac{Propagation\ delay\ T_p}{Transmission\ delay\ T_t}$
- *Maximum data rate possible/Throughput = Efficiency × Bandwidth*
- *Round Trip Time (RTT) = 2 × T_p*
- *Optimal Sender window size = N= (1+2a) for 100% efficency*



Go back to 2

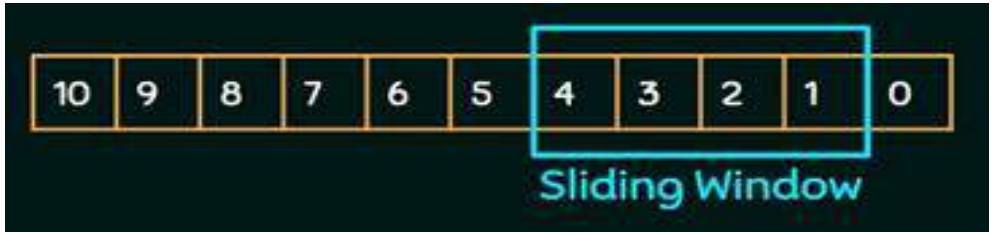


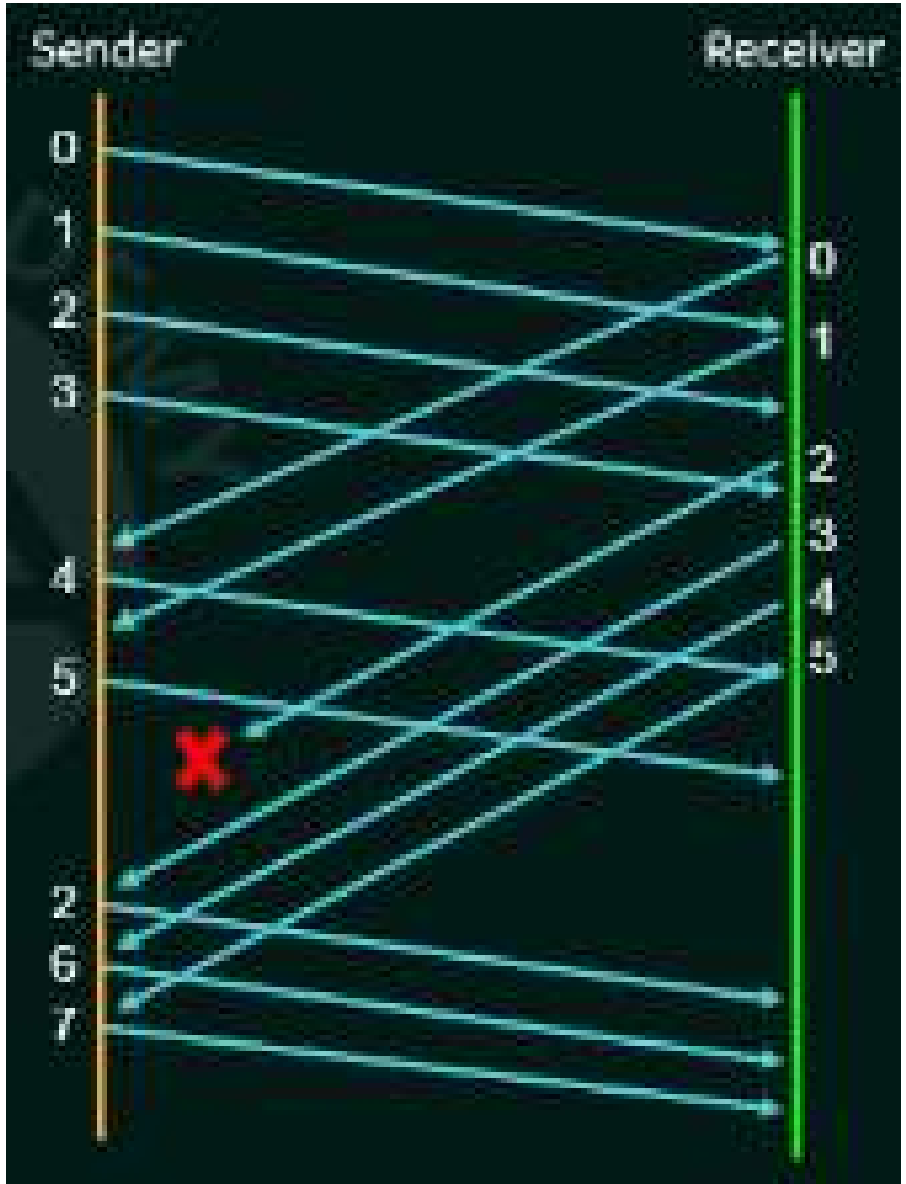
Selective Repeat Request ARQ Protocol

- *In selective repeat ARQ, only the erroneous or lost frames are retransmitted, while correct frames are received and buffered at the receiver side.*
- *The receiver while keeping track of sequence numbers, buffers the frames in memory and sends NACK (Not Acknowledged) for only frame which is missing or damaged.*
- *The sender will send/retransmits the frame for which NACK is received.*
- *In this protocol only the required frame is retransmitted and not the entire window.*

  - *Sender window size = N*
  - *Receiver window size =N*

Window Size = 4



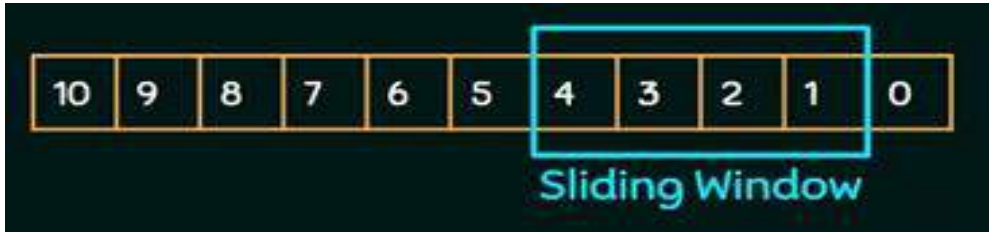


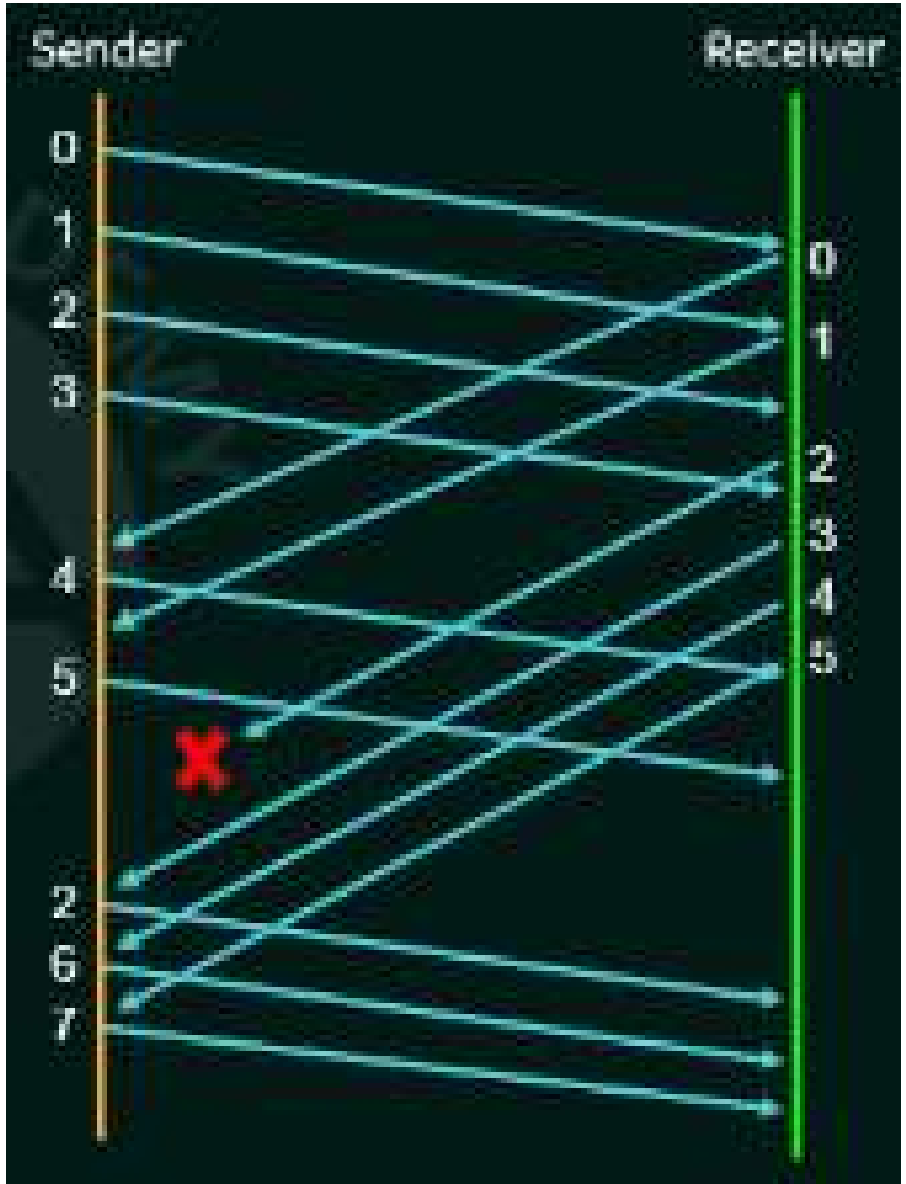
Selective Repeat Request ARQ Protocol

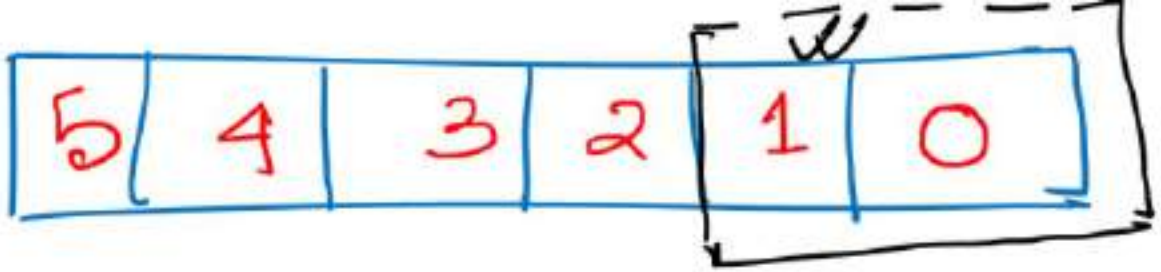
- *In selective repeat ARQ, only the erroneous or lost frames are retransmitted, while correct frames are received and buffered at the receiver side.*
- *The receiver while keeping track of sequence numbers, buffers the frames in memory and sends NACK (Not Acknowledged) for only frame which is missing or damaged.*
- *The sender will send/retransmits the frame for which NACK is received.*
- *In this protocol only the required frame is retransmitted and not the entire window.*

  - *Sender window size = N*
  - *Receiver window size =N*

Window Size = 4

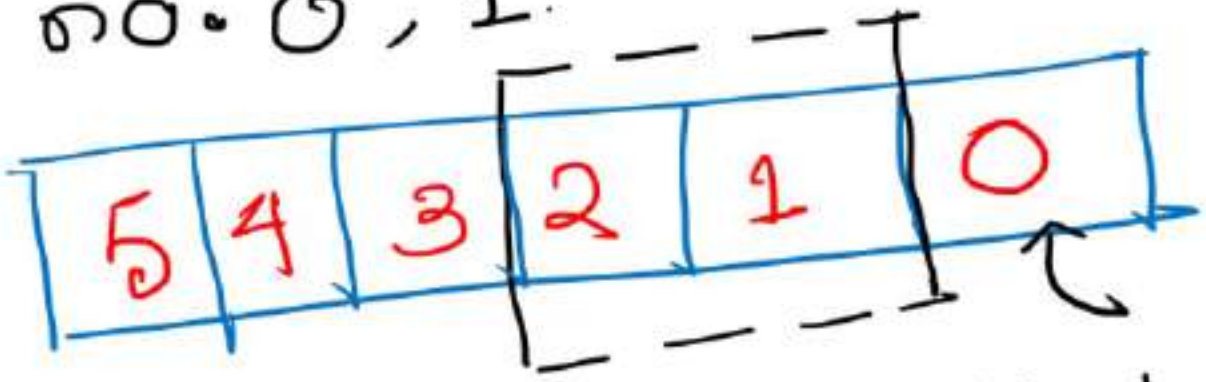






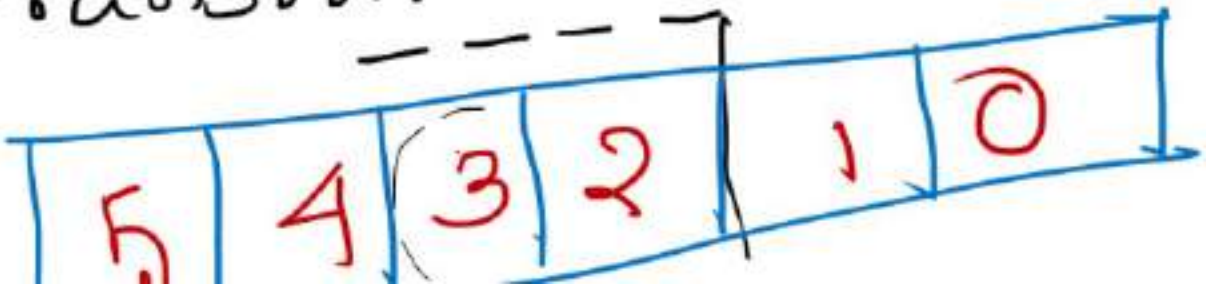
window size = 2

step 1 = sender will transmit Frame with seq. no. 0, 1.



after getting the ACK from the 0th Frame

step 2 = Frame seq no 2 will be transmitted



Go-back-N ARQ protocol:-

Sender window size = N w.

Receiver window size = 1 w

$N = 2^{n}$ → no. of bits required to represent the sequence now.

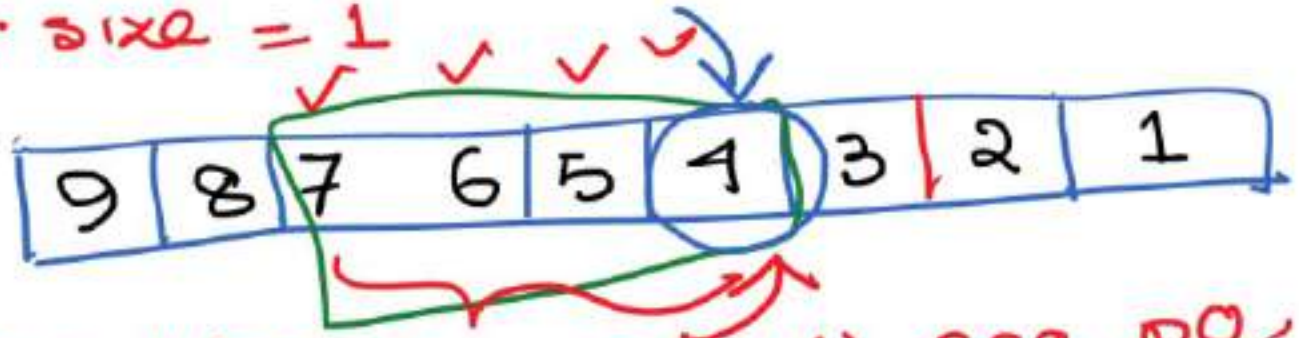
3-bit representation

2-bit

$N = 2^3 = 8$

$N = 2^2 = 4$

Go-back-4 ARQ → Receiver size = 1



:- ACK for 2, window content 3, 4, 5, 6,
6 will be transmitted

no. of Transmitted frame until now = 5 + 1 = 6

- ACK for 3, content of 4, 5, 6, 7 now 7
1 be transmit

Total no. of frame = (6 + 1) = 7

Step 1:- Frame will seq no. 1, 2, 3, 4 is transmi

Total no. of frames transmit until now =

Step 2:- ACK for frame seq. no 2 is received, now window content 2, 3, 4, 5

Total no. of frames transmit until now = 4 + 1

4th frame is in error

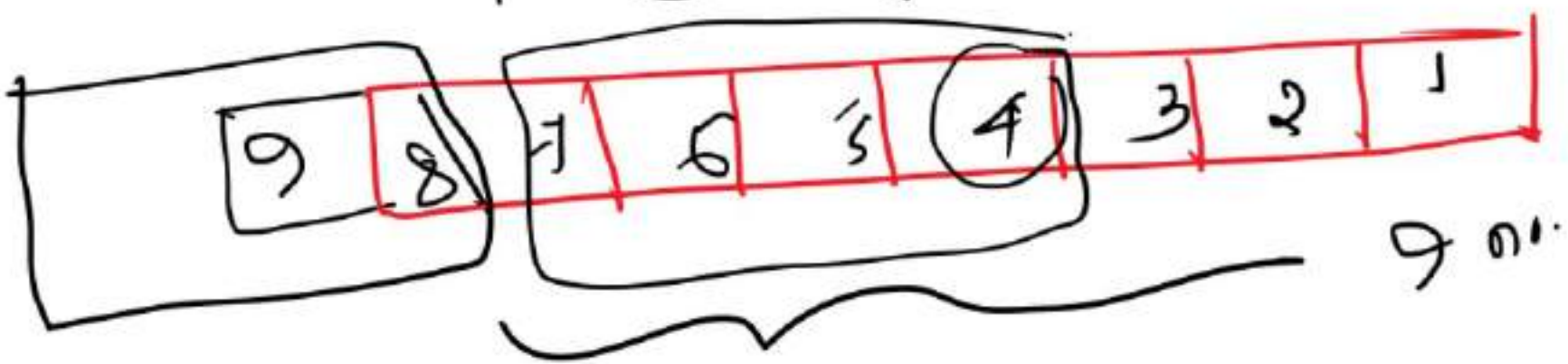
Step 5:- 4th frame all the frames in the window will be transmitted again

Total no. of frame have to be transmit

$4 = 4 + 7 = 12$

$11 + 2$

13



Go-back-N ARQ (Receiver window size = 1)

* window size = N (sender)

Stop-and-wait / Stop-and-wait ARQ

window size = 1

window size

Go-back-N ARQ

Go-back-8 ARQ :- sender window size N = 8

$N = 2^m$

3 bit representation

$N = 2^3 = 8$

00 → 0
01 → 1
10 → 2
11 → 3

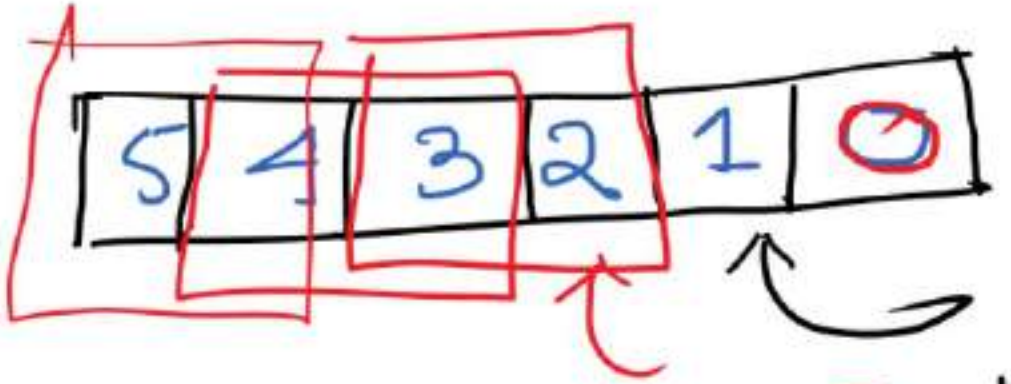
2 bit representation

$N = 2^2 = 4$

000

Sender :- message = 6 Frames

window = 2^w



w

step 3 :- ACK Frame seq. No. 1, 3 will b

step 2 :- receiver send ACK Frame no. Frame seq. no. 2 will be transmitted

Q. 9 Frame ⇒ window size ⇒ 4, that the 4th frame in the queue has not got and ACK (4th Frame in the queue is in error)

Now to successfully transmit my whole msg (9 no. Frame) how many no of frame we have to transmit?

$*w = 4 = 2^2$

8	7	6	5	4	3	2	1	0

ACK

2 no. of binary bit representa

0
1
2
3

#S1:- at a time we can transmit frame with seq No. 0, 1, 2, 3.

Total no. of frames sent until now = 4

#S2:- Frame seq. No. 0 got ACK, Frame with sequence no. 4 is transmitted now.

Total no. of frame sent until now = (4 + 1) = 5

#S3:- Frame seq no. 1 ACK, seq no. 5 will be transm

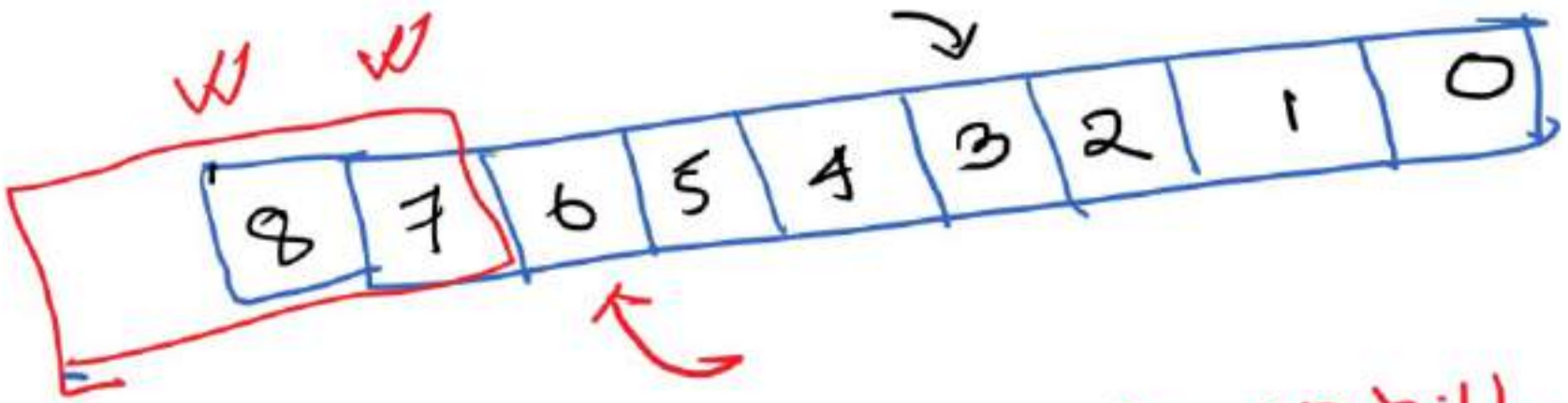
Total no. of frame sent until now = 5 + 1 = 6

#S4:- Frame seq. no. 2 ACK, seq no. 6 will

Total no. of frames sent until now =

#S5, again we have to transmit Frame with seq no. 3, 4, 5, 6

Total no. of frames sent until now = 7 + 4 = 11



9 Frame

Total no. of frame sent until now = 11 + 2 =

the queue is empty

W

Go-back-N ARQ

sender size $w = N$

receiver window is $= 1$

Selective Repeat ARQ

Efficency of Go-back-N ARQ

$$\eta = \frac{N}{(1+2a)} = \cdots \times 100\%$$

Sender window size $= N$

Receiver window size $= N$

$$a = \frac{T_p}{T_t}$$

window

$$\eta = \frac{N}{(1+2a)} = 1 = 1 \times 100\%$$

$$= 100\%$$

$$\frac{T_p}{T_t}$$

$$N = (1+2a)$$

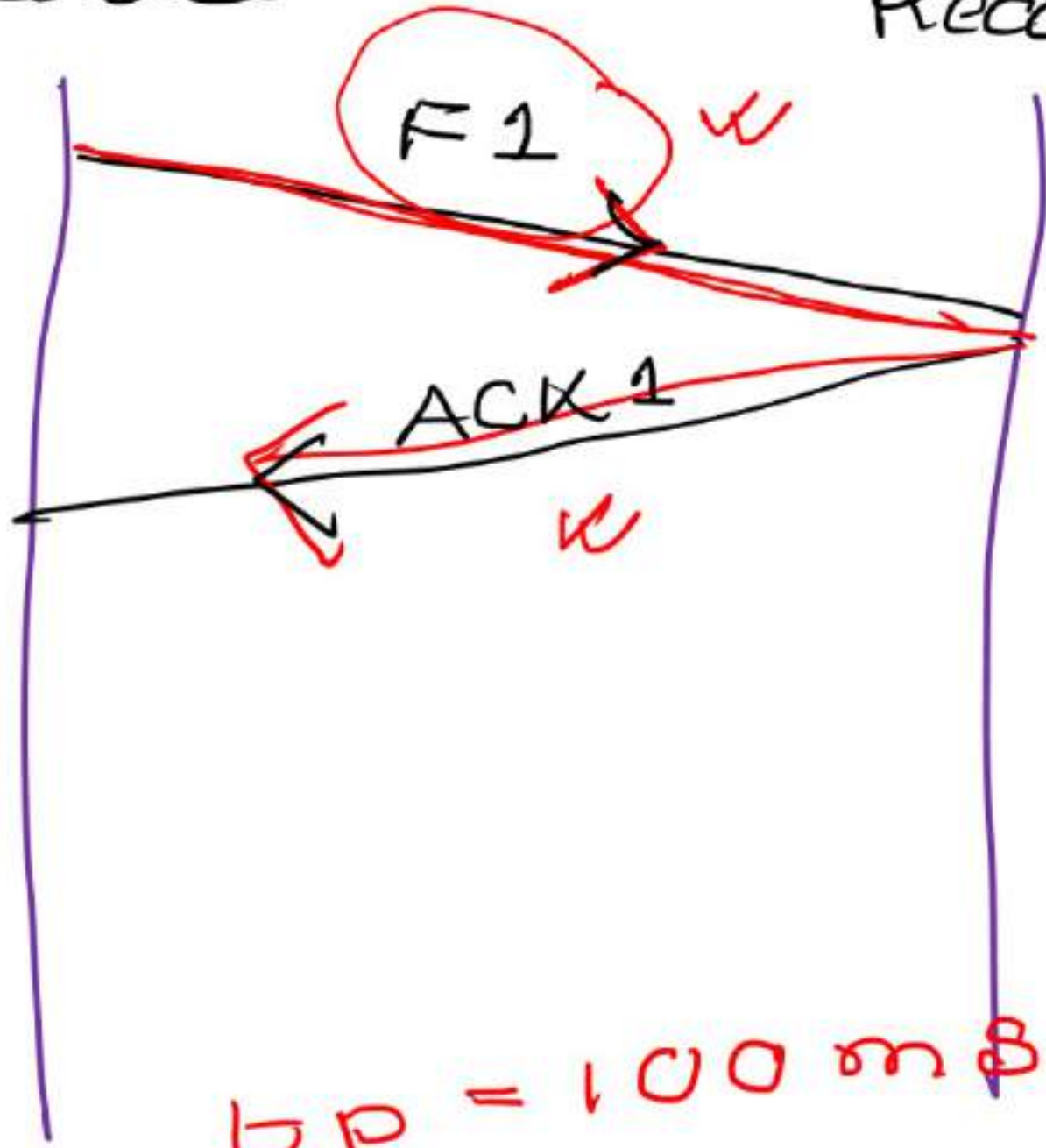
Sender Receiver

one-way Δ

$$= \frac{100\ \text{Km}}{\text{Speed of link}}$$

$$= \frac{100\ \text{Km}}{1000\ \text{Km/s}}$$

$$= 0.1\ \text{s}$$



$tp = 100\ \text{ms}$

Q. A 20 Kbps satellite link has a propagation delay of 400 ms. The transmitter employs a Go-back-N ARQ protocol. N = 10. The each frame is 100 bytes long. what will be the max data rate possible?

Given data :-

sender window size N = 10

Frame = 100 bytes
= 100 × 8 b

Tp = 400 ms

BW (Transmission) = 20 Kbps

Max data rate/Throughput

= BW × efficiency $\eta = \frac{1}{(1+2a)}$

$a = \frac{T_p}{T_t}$

= 20 Kbps × η

= 20 Kbps × 0.47

= 9.5 Kbps

$T_p = \underline{400\ ms}$, $\quad T_t = \dfrac{\text{Packet size (in bits)}}{\text{Transmission rate / B}}$

$$a = \frac{T_p}{T_t} = \frac{400\ \cancel{ms}}{40\ \cancel{ms}} = 10$$

$$= \frac{100 \text{ bytes}}{20 \text{ kbps}}$$

$$= \frac{\cancel{100} \times 8 \text{ bits}}{\cancel{20} \times 10^3 \text{ bits/s}}$$

$$= 40 \times 10^{-3} \text{ s}$$

$$T_t = 40\ ms$$

$$\eta = \frac{N}{(1+2a)}$$

$$= \frac{10}{(1+2\times 10)} = \frac{10}{21}$$

$$= 0.47$$

$$= 47\%$$

Q. Message sender size each frame 5000 bits o
data, Total msg size 1 Million (10^6 bits)

a. Stop-and-wait ARQ
b. Go-back-N ARQ.
c. Selective Repeat ARQ

All these ARQ are us
3 bit for representing
Seq. No. If the dis
betn sender & Receiv
5000 km & propag
speed 2×10^8 m/

How long all these ARQs will take send the entire msg.

* No Data & ACK is lost

Ignore Transmission
waiting & processing de
queueing delay

Given data :- Frame size = 5000 bits.

Message size = 10^6 bits

RTT delay.

Distance bet Sender & Receiver = 5000 K

Propagation speed = 2×10^8 m/s

nd Trip Time delay.

3-bit representing seq. no.

$\omega = 2^3 = 8$

l no. of frames

$$= \frac{10^6 \text{ bits}}{5000 \text{ bits}}$$

$$= \frac{1000\times10^3}{5\times10^3} = 200 \text{ Frames}$$

$Tp =$ Time required for transmitting one frame

$$= \frac{5000\times10^3}{2\times10^8 \text{ m/}}$$

$$= \frac{5\times10^6}{2\times10^8} \text{ s} = 2.5\times10^{-2}$$

RTT (Sent the Frame + getting the ACK) = 25 ms + 25 ms = 50

stop-and-wait ARQ = 50 ms × 200 = 10 ×

= 10

window

Go-back-N ARQ = $\frac{200}{8}$ × 50 ms = 125

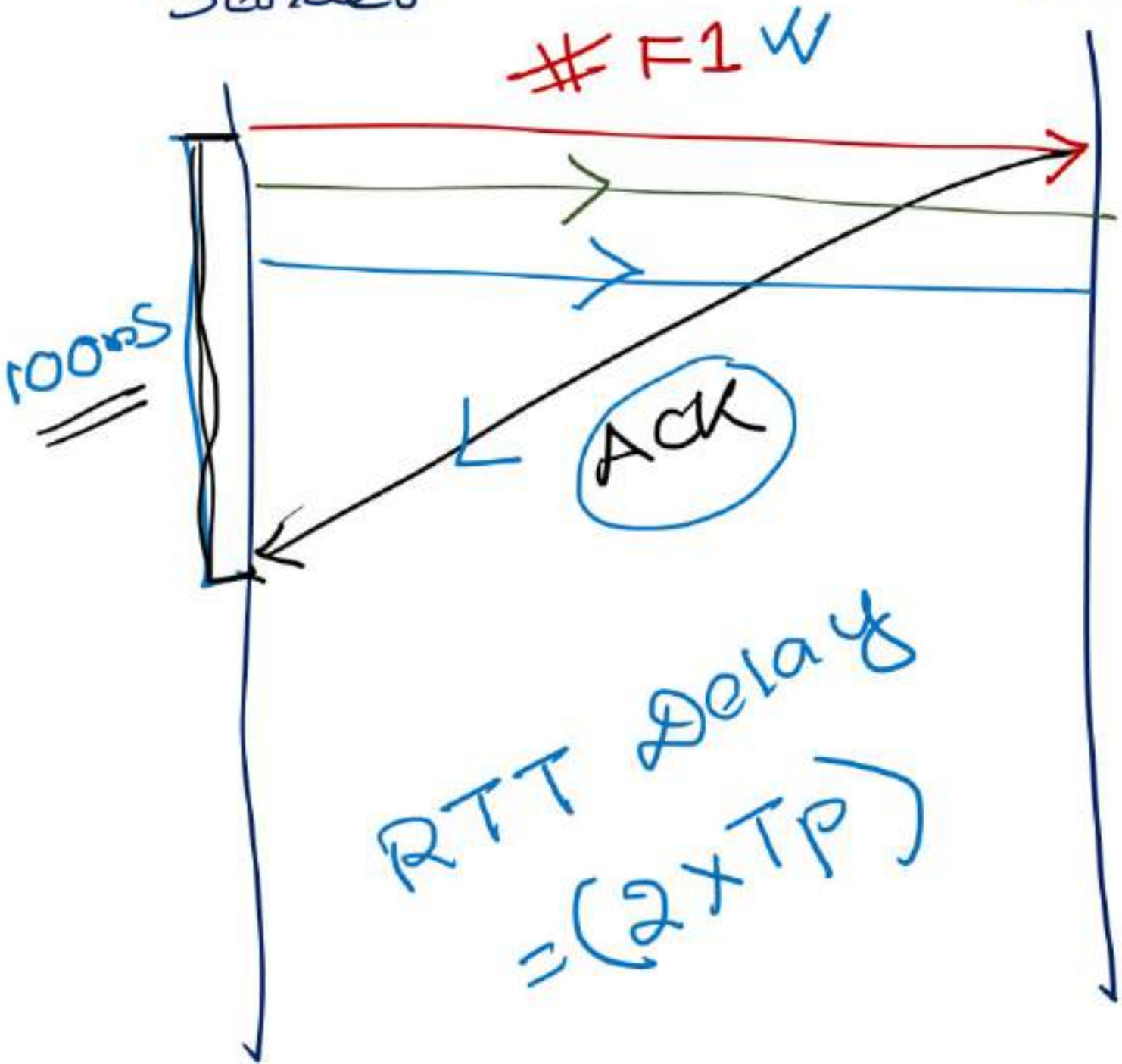
Selective-repeat ARQ = $\frac{200}{8}$ × 50 ms = 1250

* Disadvantage of Goback-N ARQ

→ if a frame is in error & didn't get acknowledgement then the entire window has to be transmitted again

→ waste

Sender



RTT Delay
$= (2 \times T_P)$

Dist = 100Km

propagation speed of Channel

$= 1000 Km/sec$

$$T_P = \frac{100 Km}{1000 Km/sec}$$

$= 0.1$ sec

$= 100ms$

$$T_t = \frac{\text{Packet size (bits)}}{\text{Transmission rate / BW. of the link}}$$

Go-back-N ARQ

$$\eta = \frac{N^{w}}{(1+2a)^{w}}$$

$w \rightarrow$ sender window size $= N^{w}$

receiver window size $= 1$

$$a = \frac{T_p}{T_t}$$

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

$$\eta = 100\%$$

$$\eta = \frac{N}{(1+2a)}$$

To achieve 100% efficiency what will be optimal window size w

$$\frac{N}{(1+2a)} = 1$$

$$[N = (1+2a)]$$

optimal window size

$N = 2^{m}$

m = no. of bits to represent msg

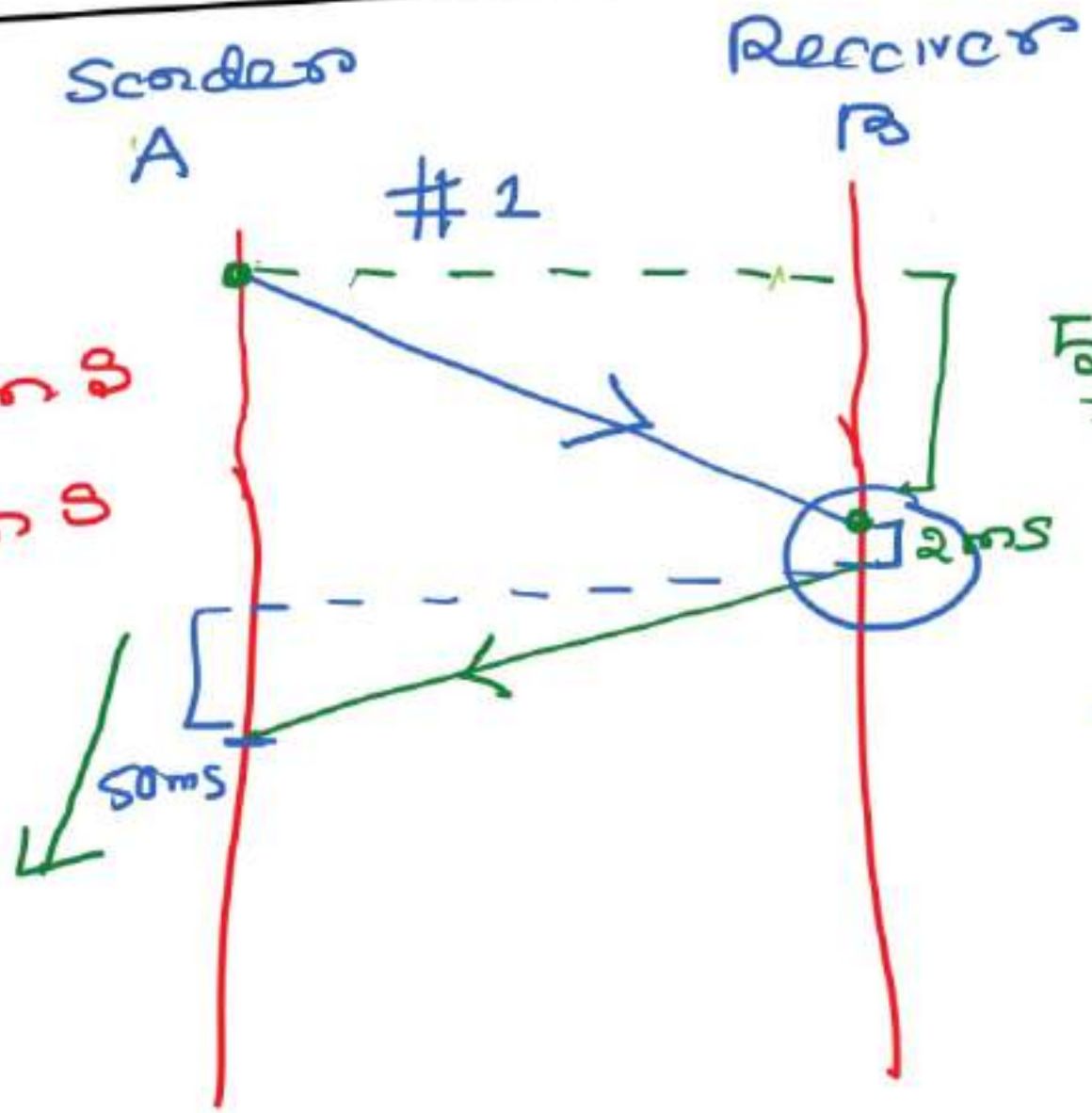
3-bit representation

$$N = 2^3 = 8$$

$$\{0, 1, \ldots 2^m - 1\}$$

max-sequence no.

* Selective Repeat request ARQ



$T_p = 50ms$

$T_t = 2ms$

Total

T_p +

RTT

$RTT = 2 \times T_p$

$T_p + T_t = 100ms$

$T_p + T_t + 0 = 10$

Q. A 20Kbps satellite link has a propagation delay of 400 ms. The transmitter employs a 'Go-back-N ARQ' N = 10. Assuming that each frame is 100 byte what is the maximum data rate possible?

* maximum data rate possible / Throughput

= Efficiency × Bandwidth.

$\eta \times 20\text{ Kbps}$

$\eta \rightarrow \frac{N}{(1+2a)}$

$a = \frac{T_p}{T_t}$

Given Data

BW = 20 Kbps

T_p = 400 ms

N = 10 (window s

Frame size = 100

$T_t = \frac{\text{packet size (bits)}}{\text{Transmission rate (BW)}} = \frac{100 \times 8}{20 \times 10^3}$

$T_t = 40$ ms

$$a = \frac{T_P}{T_t} = \frac{400\,ms}{40\,ms} = 10^{w}.$$

$$\eta = \frac{10}{(1+2\times 10)} = \frac{10}{21}$$

$$= 0.47$$

$$= 47\%$$

Max. data rate possible

$$= \eta \times BW$$

$$= 0.47 \times 20\,Kbps$$

$$= 9.4\,Kbps$$

$$\simeq 10\,Kbps$$

Q. each frame carries 5000 bits of data, how long it takes to send 1 million bit of data

a. stop-and-wait ARQ. b. Go-back-N ARQ

c. selective repeat ARQ.

Assume that all these three ARQ are using 3 bit representation for seq. No. The distance betn the sender & Receiver 5000 Km and propagation speed 2×10^8 m/s. Assuming no frame or ACK & Also assume there is no transmission delay, waiting & processing delay.

Given Data :- Frame size = 5000 bits

dist = 5000 km

propagation speed = 2×10^8 m/s

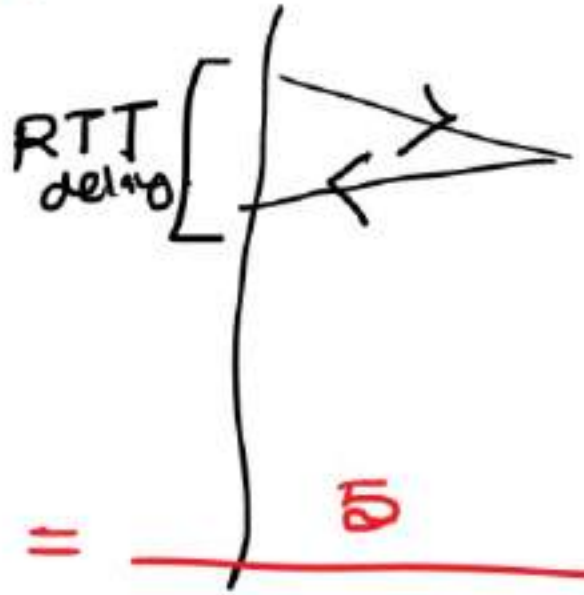
Total size of data = 10^6 bits

window size = $2^m = 2^3 = 8$

Total no. frames

$$= \frac{10^6 \text{ bits}}{5000 \text{ bits}}$$

$$= \frac{10 \times 10^5}{5 \times 10^3} = 200 \text{ Frames}$$



$$T_p = \frac{\text{distance}}{\text{propagation speed}} = \frac{5000 \text{ km}}{2 \times 10^8 \text{ m/s}} = \frac{5 \times 10^6}{2 \times 10^8}$$

$$= 2.5 \times 10^{-2} \text{ sec}$$

RTT delay $= 2 \times T_p = 2 \times 25\,ms = 50\,ms$.

* Stop-and-wait ARQ = (sender = Receiver = 1 window size)

time only one frame can be transmitted

$= 50\,ms \times 200 = 10$

→ • Go-back-N ARQ $= 50\,ms \times \frac{\cancel{200}\ 25 \text{ frame}}{\cancel{8}\ 2} = 12$

No error in frame/ACK

• Selective Repeat ARQ $= 50\,ms \times \frac{200}{8}$

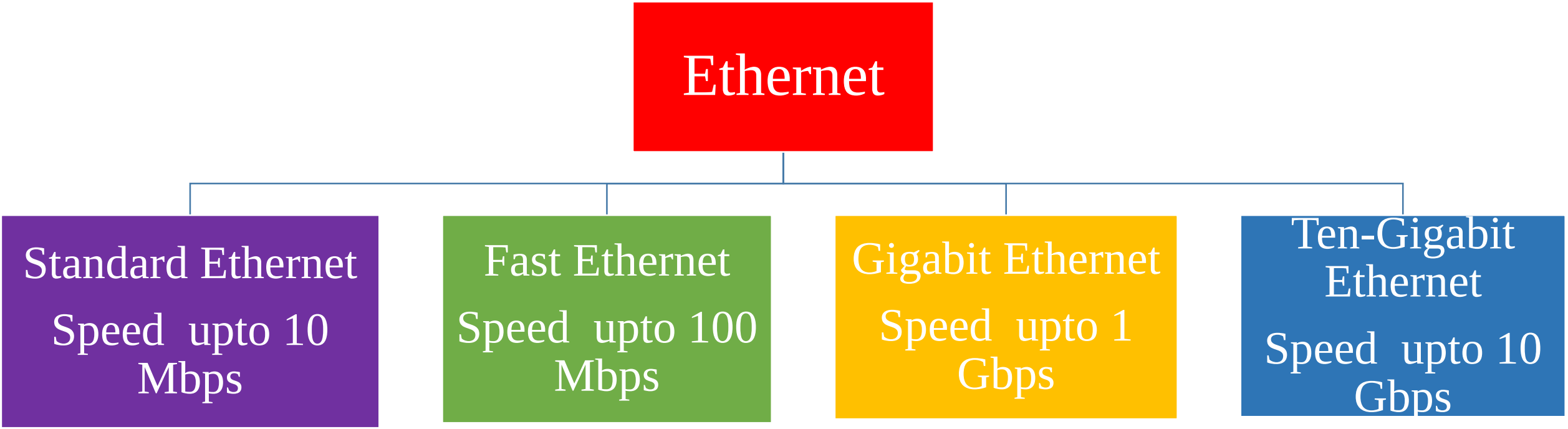
→

$= 1250\,ms$

Ethernet

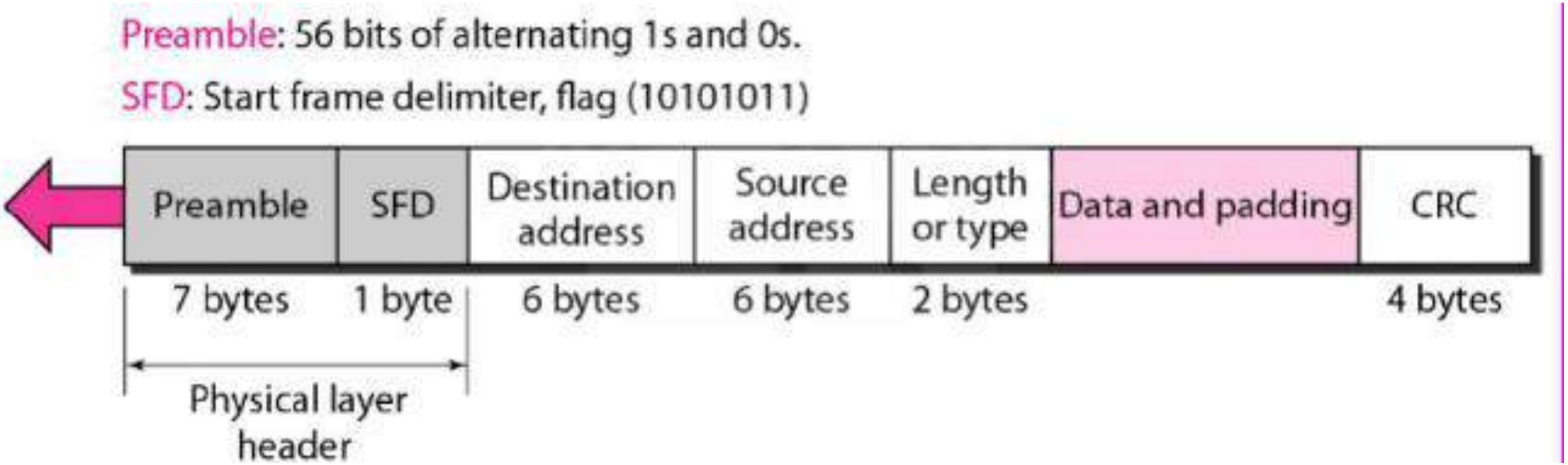
- *Ethernet is the most widely used wired LAN technologies.*
- *Ethernet protocol operates in data link layer and physical layer.*
- *Ethernet is in the family of networking technologies that are IEEE 802.2 and 802.3.*
- *Supports data bandwidth of 10 Mbps, 100 Mbps, 1000 Mbps, 40, 000 Mbps and 1 Gbps.*
- *In this protocol only the required frame is retransmitted and not the entire window.*
- *Ethernet standards define Layer 2 protocol and Layer 1 technologies.*
- *Ethernet operates in two separate sublayer of data link layer logical link layer and MAC sublayer.*

Evolution of Ethernet



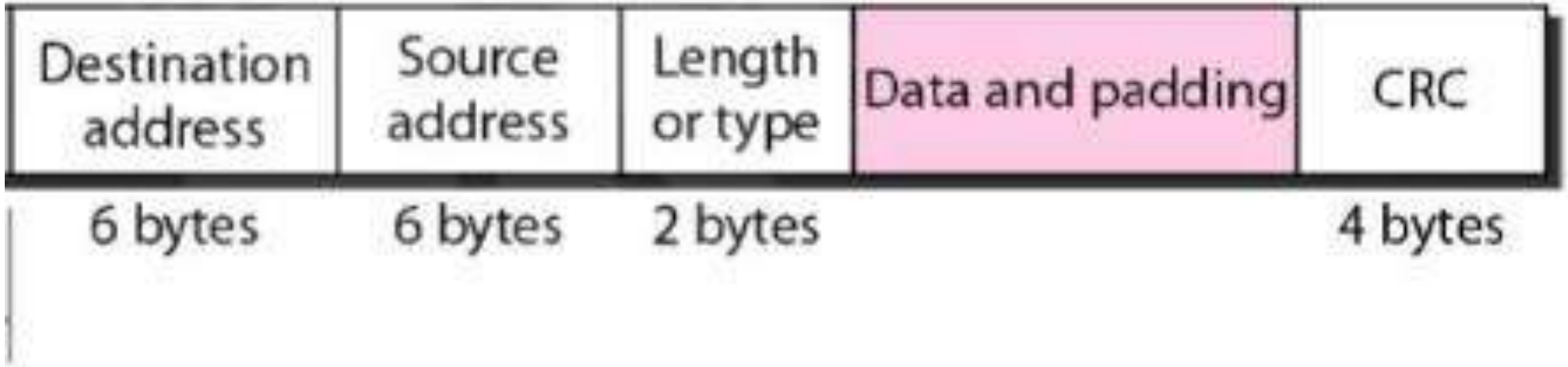
Ethernet Frame Format

- *Preamble: 56 bits of alternating '1's and '0's. It is needed for synchronization purpose.*
- *SDF (Start Frame Delimiter): Flag 10101011. This also needed for synchronization purpose. The last two bits of the Flag indicates the next field of the Ethernet frame is the destination address.*
- *Padding is required to make the received data from the upper layer (i.e network layer) a fixed size length.*



Minimum and Maximum Length of the Ethernet Frame

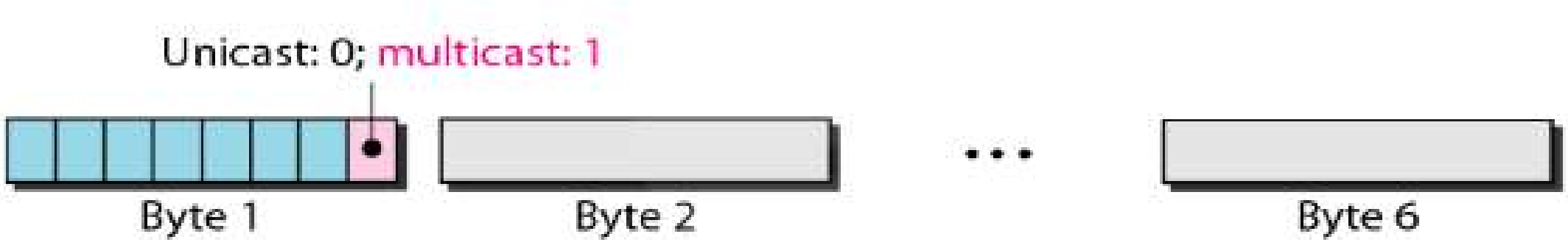
Payload from the network layer
Min payload = 46 bytes
Max. payload =1500 bytes



Min size of the Ethernet frame = 64bytes
Max. size of the Ethernet frame =1518bytes

Ethernet Address

Example: **06:01:02:01:2C:4B**



- *The least significant bit of the first byte defines the type of the address*
- *Last bit 0 defines Unicast, otherwise multicast*
- *If all bits in the frame is '1' then this is Broadcast.*

Module 3

Network Layer

Dr. Sunandita Debnath, IIIT Vadodara

IPV4 Address

- *IPV4 addresses are 32 bits long that uniquely and universally defines the connection of a device on the internet.*
- *IPV6 addresses are 128 bits long.*
- *Two devices on the internet can never have the same addresses at the same time.*
- *The addresses range of IPV4 addresses are* 2^{32} *i.e. more than 4 billions*
- *But we have more devices, the NAT (network address translation) technology will resolve this.*

IPV4 Address

- *IPV4 addresses have two notations*
  - *Binary notations.*
  - *Dotted Decimal notations.*

Binary notations: 01110101 10010101 00011101 00000010

Dotted Decimal notations: 117.149.29.2

- *IPV4 addresses consist of four octets each octet consists of 8 binary bits.*
- *Each octet ranges from 0 to* $2^8 - 1 = 255$

IPV4 Address Conversion

- Binary to Dotted Decimal conversion
- Dotted Decimal to Binary Conversion

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
128	64	32	16	8	4	2	1

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

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			

Decimal Notation

Address Class	1st Octet range in decimal	1st Octet bits (Blue Dots do not change)	Network (N) and Host (H) Portion	Default mask (Decimal)	Number of possible networks and hosts per network
A	0-127	00000000 - 01111111	N.H.H.H	255.0.0.0	128 Nets (2^7) 16,777,214 hosts ($2^{24}-2$)
B	128-191	10000000 - 10111111	N.N.H.H	255.255.0.0	16,384 Nets (2^{14}) 65,534 hosts ($2^{16}-2$)
C	192-223	11000000 - 11011111	N.N.N.H	255.255.255.0	2,09,150 Nets (2^{21}) 254 hosts (2^8-2)
D	224-239	11100000 - 11101111	NA (Multicast)	-	-
E	240-255	11110000 - 11111111	NA (Experimental)	-	-

Subnet Mask (Slash Notation)

Class	Subnet Mask (in Dotted Decimal)	Subnet Mask (in Binary)	Slash Notation
A	255.0.0.0	11111111.0000000.00000000.0000000	/8
B	255.255.0.0	11111111.11111111.00000000.0000000	/16
C	255.255.255.0	11111111.11111111.111111111.0000000	/24

IP address	Network portion of IP address ⟷ Host portion of IP address **192.168.10.2**
Subnet Mask	255.255.255.0

Ethernet protocol → Data link layer

Ethernet technologies → physical layer

IEEE 802.2 & 802.3

IEEE 802.11

bluetooth & WiFi

A Ethernet MAC sublayer received 42 bytes from the network layer, then how many padding bytes should be required?

→ min payload size = 46 bytes

42 bytes ⇒ 46 − 42 ⇒ 4 bytes

Data received from the upper layer 1530 bytes ?
Then how the data will be transmitted ?

→ max size payload = 1500 bytes

received data = 1530 bytes

↓

We can't send it in a single

Frame(1) = 1500 bytes

Frame(2) = 30 bytes < 46 bytes

min size of payload

46 − 30 = 16 bytes

padding in the frame 2 w

06:01:02: 1:2C:4B ⇒

Address of Ethernet Frame

06 ⇒ 00000110

LSB = 0 ⇒ unicast address

LSB ⇒ 1 ⇒ multicast

All bytes consist of 1's ⇒ broad

Hex →	decimal →	Binary
		8 4 2 1
0	0	0 0 0 0
1	1	0 0 0 1
⋮		
9	9	
A	10	1 0 1 0
B	11	
C	12	
F	15	1 1 1 1

FF : FF : FF : FF : FF : FF

47 : 20 : 1B : 2E : 08 : EE

8 4 2 1

47 → 0 1 0 0 0 1 1 1

multicast

47. 20 : 1B : 2E : 08 : EE

how it will sent out on the cable?

- address will be sent left to right
- now in byte we will right to left
- binary equivalent 8421 → 1248

B ⇒ 11

E – 14

format

74 : 02 : B1 : E2 : 80 : EE

1110 0010 0000 0100 1101 1000 0111 0100 0001 0000

0111 01

LAN → wireless LAN (802.11)

↓

bluetooth

WiFi

→ wired LAN (ethernet)

802.2 & 802.3

Ethernet cross wire

Ethernet — Layer 1 (physical layer)
technology

— Layer 2 (Data link Layer)
Protocol

Data and Padding ⇒
the data received from the upper layer (network layer) can be of variable
padding

Min ⇒ 46 byte

Max ⇒ 1500 byte

42 byres +

(46 - 42) = 4 byres

Maximum size of Ethernet Frame

$6+6+2+1500+4=1518$ byte

Minimum size of Ethernet-Frame.

$6+6+2+46+4=64$ bytes

64 bytes 1518 bytes

Q. An Ethernet MAC sublayer receives 1510 bytes of data from the upper layer. Can the data be encapsulated in one single frame? If not how many frame is required? what is the size of each frame?

network layer have send = 1510 bytes

Max = 1500 bytes

Two Frame to send this 1510 bytes

$1510 = 1500 + 10$

Frame 1

Frame 2

46 - 10

= 36 bytes

of padding

$(5B)_{16} \Rightarrow 5\ 11$

0101 1011 equivalent

Decimal	Hexadecimal
0	0
1	1
⋮	
9	9
10	A
11	B

Binary			
8	4	2	1
0	0	0	0
0	0	0	1
1	0	0	1

Ethernet Frame Address

47 : 20 : 1B : 2E : 08 : EE

Ethernet = 48 bits = (6 × 8)

= 48 bit

LSB

01000111

→ 0 ⇒ unicast

1 ⇒ multicast

11111111

broadcast

47 : 20 : 1B : 2E : 08 : EE

74 02 B1 E2 80 EE

10000001001101100010110100000100001011101 1

ernet cable (physical)

'1' '0'

Network layers

IP address

subnet mask

intelligent device

IPV4 ← 32 bit long

$2^{32} \Rightarrow$ A billion

→ MAC address (NIC card)

hardware address

unique

IPV6

← 128 bits long

$2^{128} \Rightarrow$ IP address

→ IP address

10.2.5.0

1 byte = 8 bits

cmd

ipconfig

124.255.255.255

0.0.0.0

255.255.255.255

IPV4 address 0.0.0.0

255.255.255.255

Valid address

99.88.67.89

1.2.3.6

100.200.230.6

Invalid Address (IPV4 ⇒ 4 octets)

56.89.1.2.4 ⟶ 5 octets

232.2.257.0 ⟶ invalid

↑ > 255

10.065.42.98 ⟶ no prefix with zero

↑ 0 to 255

invalid address

IPV4

Binary notation — Decimal notation

$2^7\ 2^6\ 2^5\ 2^4\ 2^3\ 2^2\ 2^1\ 2^0$ $2^7\ 2^6\ 2^5\ 2^4\ 2^3\ 2^2\ 2^1\ 2^0$ $2^7\ 2^6\ 2^5\ 2^4\ 2^3\ 2^2\ 2^1\ 2^0$

1110 1111 11110111 11000111 00011

239 . 247 . 199 . 29

A B C D

$0 < A, B, C, D < 255$

128 . 8 . 15 . 1

00001000 00001111 00000001

IPV4 ⇒ 0 . 0 . 0 . 0

255 . 255 . 255 . 255

2^{32} address

First bytes

MSB of First byte

Class A ⇒ 0 0000000 → 0

2^6 2^5 2^4 2^3 2^2 2^1 2^0

0 1111111 → 127

0 - 127

Class B

10000000 → 128

10111111 → 191

128-191

First byte

192 - 22

Class C

11000000 → 192

11011111 → 223

2^8 2^1 2^0

class D

First bytes

224 - 239

11100000] 224

11101111] 239

nal &

nclass E

rpose

First bytes

240 - 255

11110000] 240

11111111] 255

255.255.255

192.168.2.1 → class C

255.255.255.0 →

N.N.N.H

host of address

class A

125.168.2.1

First bytes Second bytes 3rd 4th

125 . 0 . 2 . 1

$0 < 125 < 127$

class A → subnet mask

255 . 0 . 0 . 0

N H H H

125 . 0 . 2 . 1

0 – 127

0 bit vs to classify class

0 0000000] 7 bits

0 1111111

$2^7 = 128$ networks

$2^{24} - 2 = 16,777,214$ hosts

128 . 8 . 15 . 1

00001000 00001111 00000001

128 . 8 . 15 . 1

00001000 00001111 00000001

128 . 8 . 15 . 1

00001000 00001111 00000001

128 . 8 . 15 . 1

00001000 00001111 00000001

128 . 8 . 15 . 1

00001000 00001111 00000001

128 . 8 . 15 . 1

00001000 00001111 00000001

128 . 8 . 15 . 1

00001000 00001111 00000001

IPV4 ⇒ 32 bits long

a	b	c	d
8 bits	8 bits	8 bits	8 bits

Four bytes

Four octets

Class C → 110 00000 → 192

110 11111 → 223

Network portion

Host portion of IP address

192.168.10.2 ← user 1

255.255.255.0

N N N H

192.168.10.2 → user 1

192.168.10.1 → user 2

192.168.20.3 → user 3

user 1 & user 2 belongs to same network

Class A ⟹ 0 to 127 ⟹ 255 · 0 · 0 · 0

N H H H

Class B ⟹ 128 to 191 ⟹ 255 · 255 · 0 · 0

N N H H

Class C ⟹ 192 to 223 ⟹ 255 · 255 · 255 · 0

N N N H

First byte/octet

00000000

11111111

128 network

16 host

First octet = 7

remaining bits are 3 octets

8 bits x 3 =

2^7 ⟹ combinations each network

24 bits ⟹ host ad

$2^{24} - 2$ host

Class C ⇒ 192 to 223 ⇒ 255 . 255 . 255 . 0

N N N H

First octet

110 00000

110 11111

First 3 bits are reserved to define the class

remaining 5 bits here

5 + 8 + 8 = 21

2^{21} ⇒ networks

$2^8 - 2$ hosts

gateway

broadc

Class A — 255.0.0.0 — 10.0.0.1 (0 to 127)

Class B → 255.255.0.0 → 132.16.0.2 (128 to 191)

Class C → 255.255.0.0 → 198.162.0.1 (192 to 23)

Class A ⇒ subnet mask of class B & C

Class B ⇒ subnet mask C

Class C ⇒ subnet mask G

10.0.0.1 → cl

255.255.0.0 →

255.255.255

IP ⟹ address For user 1 10 . 2 . 0 . 1

N

class A

255 . 0 .

Given subnet mask

255 . 255 . 0 . 0

class B

255 . 255 . 0 .

N N H

Q. user 1 have an IP address 172.16.200.1

user 2 " " "

172.16.165.2

Subnet mask used is

255.255.255.0

N N N

Can we connect user 1 & user 2 with a swi

→ So to belong to same network user 1 & u

ers the First 3 byt/octets should be same

Not belongs to same network we

cannot connected 1 & 2 usin

user 1 → 172.16.200.1

user 2 → 172.16.165.2

same network

user 1 & user

subnet mask _ 255.255.0.0

used vs N N H H

can be conne
through su
or not ?

→

8 bits 8 bits 8 bits

0 →

Class A → 0 00000000 → 0

255.255

Broadcast

0 1111111 → 127

128 blo

Network = 2^7 = 128 net

in each network

Host = $(2^{24} - 2)$

Default subnet mask

N H H H

= 255 . 0 . 0 . 0

≈ 16 crores

Service A = 10 . 2 . 1 . 0 (class A)

IPV4 → 32 bits → 4 octets / 4 bytes

N N N H

192.168.2.1 → Class C (192-223 First octe

255.255.255.0 Default subnet mask

N N N H

192.168.2.3

N H

192.168.2.1 → u1

192.168.1.3 — u2

255.255.255.0

u1 and u2 using switch?

No → u1 & u2 belong different network

Class C :- 110 00000 → 192

5 bits

110 11111 → 223

8 ✓ ✓

N N N H

8 bits

Default subnet mask = 255.255.255.0

Network address = (5 + 8 + 8) = 21 = 2^{21}

Host address = $2^8 - 2$ =

Class A → Default subnet mask → class B & Class C

Class B → Default subnet mask → Class C

Higher class subnet mask can be used for lower

NV H

10 . 1 . 1 . 2 → class A ip

255 . 255 . 255 . 0 → subnet mask for class C.

Dotted Decimal

Class A → 255.0.0.0 /8

Binary

/18

11111111.0000 . . . . . .

1111111.0000000

1111.1111.0000.1111

N N N H

192 . 168 . 10 . 1 / 26

mask

128
+ 64
19

subnet mask

192 . 168 . 10 . 1

11111111 . 11111111 . 11111111 . 11000

255 . 255 . 255 . 192 ←

su

130 . 34 . 12 . 64 / 28

8 8

4

8 8

11 -- 11 . 11 -- 11 . 111 -- 11 . 11110000

255 . 255 . 255 . 240

class C

128
64
32
16

2

$2^8 - 2$

5 + 8 + 8

$100 \cdot 10 \cdot 5 \cdot 2 \longrightarrow$ "1"

$100 \cdot 10 \cdot 5 \cdot 5 \longrightarrow$ "2"

$100 \cdot 10 \cdot 5 \cdot 6 \longrightarrow$ "3"

$\underbrace{255 \cdot 255 \cdot 255}_{N} \cdot \underbrace{252}_{+1} \longrightarrow$

subnet mask

$u_1 \rightarrow 100 \cdot 10 \cdot 5 \cdot 2 \rightarrow 01100100 \quad 00001010 \cdot 00000$

00000

$255 \cdot 255 \cdot 255 \cdot 252$

$\rightarrow 11111111 \cdot 11111112 \cdot 11111112 \cdot 111$

AND operation =

$01100100 \cdot 00001010 \cdot 00000101 \cdot 000000$

$= 100 \cdot 10 \cdot 5 \cdot 0 \rightarrow u_1 \rightarrow$ differe

$100 \cdot 10 \cdot 5 \cdot 4 \rightarrow u_2$

$100 \cdot 10 \cdot 5 \cdot 4 \rightarrow u_3$

same

0 ← Network (Gateway)

2.255 → broadcast

→ Subnet mask of class B & class C

Class B → Subnet mask of class C

First byte/octet

Class C → 110 00000 → 192

110 11111 → 223

8 8 5

N N N H

255.255.255.0

8 bits each network no. of possible

$2^8 - 2$ = no.

class C

$(5+8+8) =$

$2^{21} \rightarrow N$

255.255.255.255

N H

192.168.2.1 → u1 w

192.168.2.2 → u2 w

255.255.255.0

N N N H

u1 & u2 with switch

192.168.2.

192.168.2.

Host w

0 w

255.

Subnet mask

$U_1 \rightarrow 100 \cdot 10 \cdot 5 \cdot 2$ ✓ 3 4

$U_2 \rightarrow 100 \cdot 10 \cdot 5 \cdot 5$ ✓

$U_3 \rightarrow 100 \cdot 10 \cdot 5 \cdot 6$ ✓

$255 \cdot 255 \cdot 255 \cdot 252$

N		N	N			
2^7	2^6	2^5	2^4	2^3	2^2	2^1
128	64	32	16	8	4	2

$255 \cdot 255 \cdot 255 \cdot 0$

$10 \cdot 5 \cdot 2 \rightarrow$ 01100100 00001010 00000101 0000

mask $\rightarrow$ 11111111 11111111 11111111 1111

ahon $\rightarrow$ 01100100 00001010 00000101 000

$U1 \rightarrow 100 \cdot 10 \cdot 5 \cdot 0$

130 . 34 . 12 . 64 / 28 ← subnet mask

H

128

address

e accompanied IPV4 address

ont mask

/28

128+64+32+16

11111111 . 11111111 . 11111111 . 11110000

→ 255 . 255 . 255 . 240

N N N H

$2^8 - 2$ = no. host can be possible

130.34.12.0 ← network address (gateway

130.34.12.255 ← Broadcast address

6 8

N N H H

192.168.20.1 → class B

default subnet mask

255.255.0.0

10 000000

network = 2^{14}

Module 3

Network Layer

- **Subnetting**

Dr. Sunandita Debnath, IIIT Vadodara

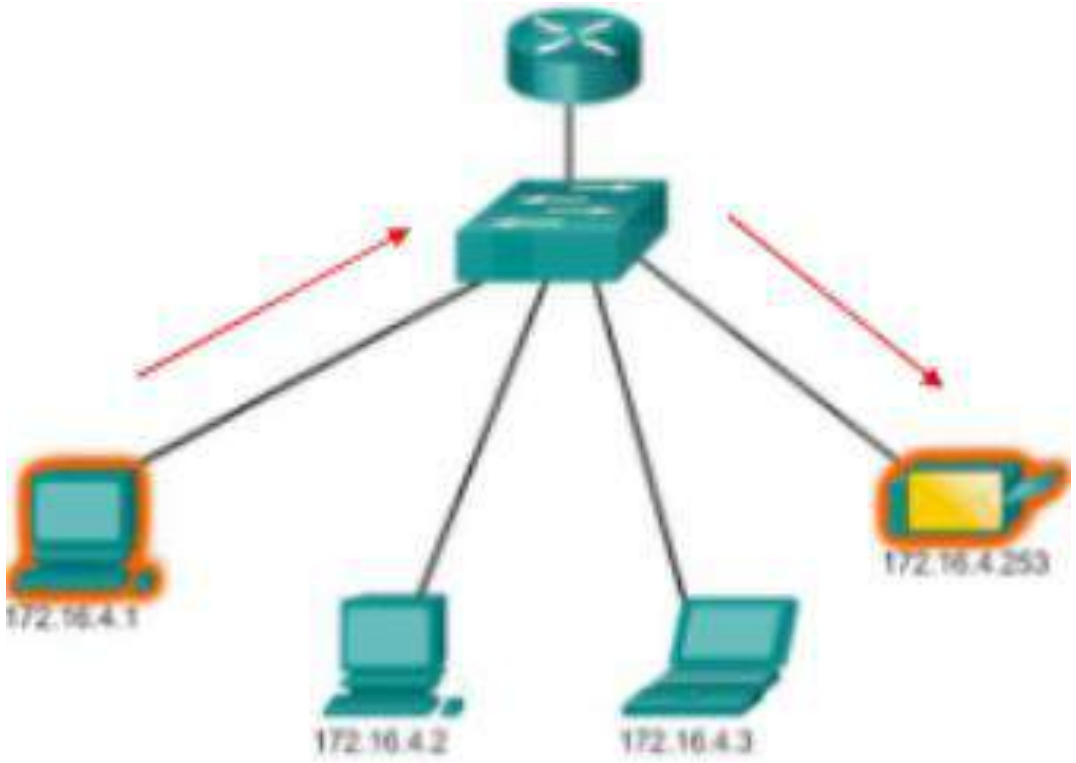
Different ways of Transmission in IPV4

- *In a IPV4 network, the hosts communicate one of three different ways*
  - *Unicast*
  - *Broadcast*
  - *Multicast*

Unicast Transmission: *The process of sending a packet from one individual host to another host.*

Source: 172.16.4.1

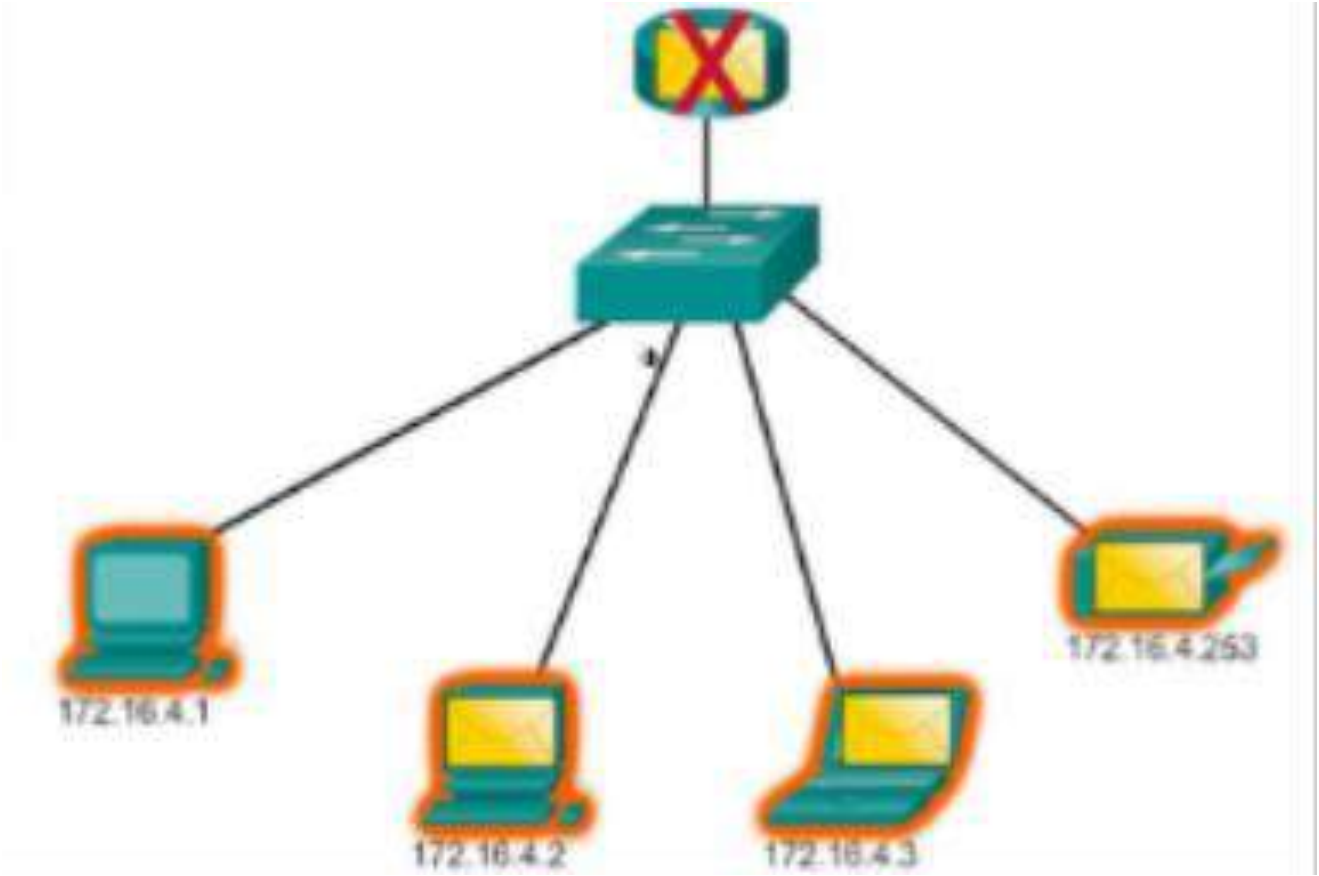
Destination: 172.16.4.253



Different ways of Transmission in IPV4

Broadcast Transmission: *The process of sending a packet from one host to all hosts in the network.*

- *Limited Broadcast*
- *Directed Broadcast*



Limited Broadcast

Source: 172.16.4.1

Destination: 255.255.255.255

Router will not forward a limited broadcast.

Directed Broadcast

Destination: 172.16.4.255

Host within the 172.16.4/24 network

Different ways of Transmission in IPV4

Multicast Transmission: *The process of sending a packet from one host to a selected group of hosts, possibly in different networks.*

- *Multicast transmission reduces traffic*
- *The multicast address range 224.0.0.0 to 239.255.255.255.*
- *Link Local-224.0.0.0 to 224.0.0.255 (E.g. routing information exchanged by routing protocols)*
- *Globally scoped address-224.0.1.0 to 238.255.255.255 (E.g. 224.0.1.1 has been reserved for Network Time Protocol)*

Public and Private IPV4 Addresses

Private IP addresses:

- *IPV4 address ranges from 0.0.0.0 to 255.255.255.255, where 0.0.0.0 is the network address and 25.255.255.255 is the broadcast address.*
- *According to early network design, IP addresses should be unique, but its is not true with the introduction of private and public IP addresses.*
- *Two computers can use the same private address using NAT (Network Address Translation).*
- *Hosts that do not require access to the internet can also use private addresses. E.g. many computers in a company do not need internet connectivity but can communicate internally among each other using private IP addresses.*

Private IPV4 Addresses

- *Ranges of Private IP address*
  - **10.0.0.0 to 10.255.255.255 (10/8) Class A**
  - **172.16.0.0. to 172.31.255.255 (172.16.0.0/12) Class B**
  - **192.168.0.0 to 192.168.255.255 (192.168.0.0/16) Class C**

- *The aforementioned are the three non overlapping ranges of IPV4 addresses for private networks are reserved.*

- *Public IP address is used to communicate outside the network. Public IP addresses are basically assigned by the Internet Service Provider (ISP)*

Special use IPV4 Addresses

- **Network and Broadcast addresses**

Within each network the last and first address can not be assigned to hosts.

- **Loopback addresses 127.0.0.1 (127.0.0.0 to 127.255.255.255 are reserved)**

127.0.0.1 *is a special address which host s use to direct traffic to themselves*

Link-Local addresses 169.254.0.0 to 169.254.255.255 (169.254.0.0/16)

Can automatically assigned to the local hosts.

Test-Net addresses 192.0.2.0 to 192.0.255.255 (192.0.2.0/24)

Set aside for teaching and learning purposes, used in documentation and network example.

- **Experimental addresses 240.0.0.0 to 255.255.255.254**

Are listed as reserved for experimental purpose.

Drawbacks of Classful Addressing

Address Class	1st Octet range in decimal	1st Octet bits (Blue Dots do not change)	Network (N) and Host (H) Portion	Default mask (Decimal)	Number of possible networks and hosts per network
A	0-127	00000000 - 01111111	N.H.H.H	255.0.0.0	128 Nets (2^7) 16,777,214 hosts ($2^{24}-2$)
B	128-191	10000000 - 10111111	N.N.H.H	255.255.0.0	16,384 Nets (2^{14}) 65,534 hosts ($2^{16}-2$)
C	192-223	11000000 - 11011111	N.N.N.H	255.255.255.0	2,09,150 Nets (2^{21}) 254 hosts (2^8-2)
D	224-239	11100000 - 11101111	NA (Multicast)	-	-
E	240-255	11110000 - 11111111	NA (Experimental)	-	-

Drawbacks of Classful Addressing

- *Lack of Internal Address Flexibility*
- *Inefficient use of address space*
- *Proliferation of Router Table entries*

Classless Addressing

- *Formal name is Classless Inter-Domain Routing (CIDR).*

- *Created a new set of standards that allowed service providers to allocate IPV4 addresses on nay address bit boundary (prefix length) instead of only Class A, B, and C addresses.*

- *Classless addressing is possible with the help of subnetting.*

Valid Subnet Masks

Subnet Value	Bit Value							
	128	64	32	16	8	4	2	1
255	1	1	1	1	1	1	1	1
254	1	1	1	1	1	1	1	0
252	1	1	1	1	1	1	0	0
248	1	1	1	1	1	0	0	0
240	1	1	1	1	0	0	0	0
224	1	1	1	0	0	0	0	0
192	1	1	0	0	0	0	0	0
128	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

Valid Subnet Masks

/n	Mask	/n	Mask	/n	Mask	/n	Mask
/1	128.0.0.0	/9	255.128.0.0	/17	255.255.128.0	/25	255.255.255.128
/2	192.0.0.0	/10	255.192.0.0	/18	255.255.192.0	/26	255.255.255.192
/3	224.0.0.0	/11	255.224.0.0	/19	255.255.224.0	/27	255.255.255.224
/4	240.0.0.0	/12	255.240.0.0	/20	255.255.240.0	/28	255.255.255.240
/5	248.0.0.0	/13	255.248.0.0	/21	255.255.248.0	/29	255.255.255.248
/6	252.0.0.0	/14	255.252.0.0	/22	255.255.252.0	/30	255.255.255.252
/7	254.0.0.0	/15	255.254.0.0	/23	255.255.254.0	/31	255.255.255.254
/8	255.0.0.0	/16	255.255.0.0	/24	255.255.255.0	/32	255.255.255.255

Subnetting

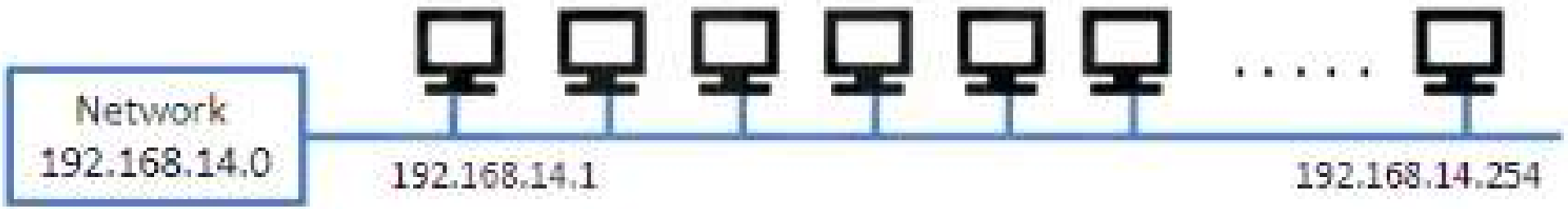
- *A subnetwork or subnet is a logical subdivision of an IP addresses.*

- *The practice of dividing a network into two or more networks is called subnetting.*

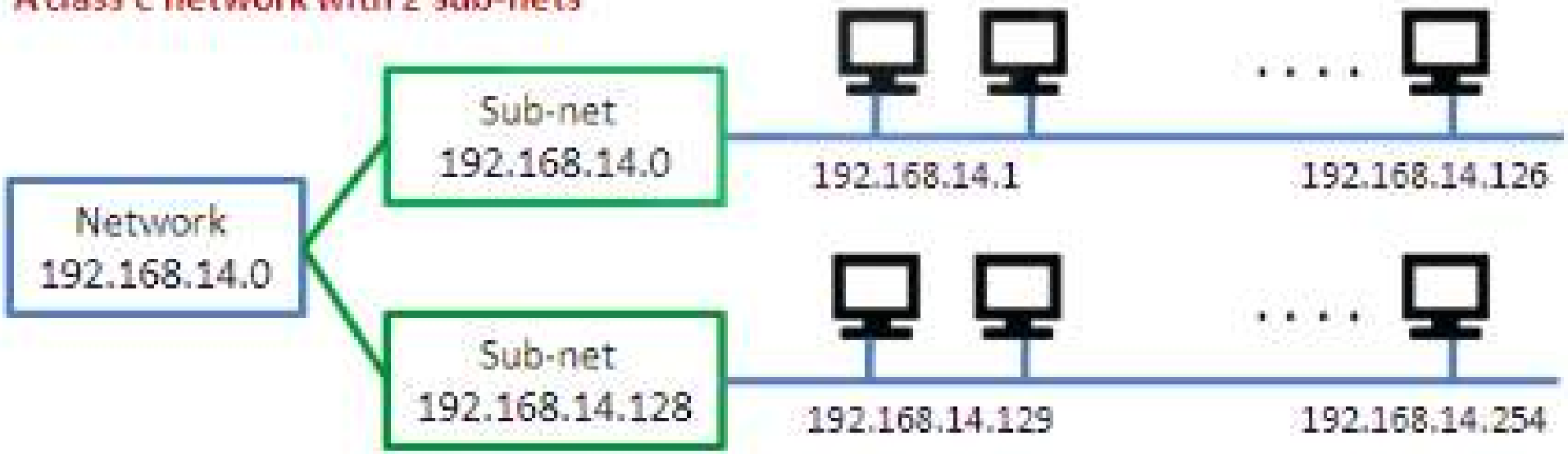
- *Computers that belongs to a subnet are addressed with an identical most significant bit-group in their IP addresses.*

IP Subnetting Example

A class C network without sub-netting



A class C network with 2 sub-nets



- *The first address and the last address is reserved for network and broadcast respectively. The default subnet mask for Class C which is 255.255.255.0 cannot be used for subnetworks. But the total number of host supported by Class C ie.e* $2^8 - 2 = 254$ *should be maintained.*

Subnetting

Subnetting can be done both according to Host requirement and network requirement.

The steps for subnet the given network according to host requirements:

- *Identify the Class of the IP address and note the default subnet mask.*
- *Convert the default subnet mask into binary format.*
- *Note the number of hosts required per subnet and find the binary equivalent of it.*
- *Find the subnet generator (SG) and octet position.*
- *Generate the new subnet mask.*
- *Use the SG and generate the network ranges in the appropriate Octet positions.*

Q. Subnet the IP address 216.21.5.0 into 30 hosts in each subnet.

Q. Subnet the IP address 196.10.20.0 into 52 hosts in each subnet.

Q. Subnet the IP address 150.15.0.0 into 500 hosts in each subnet.

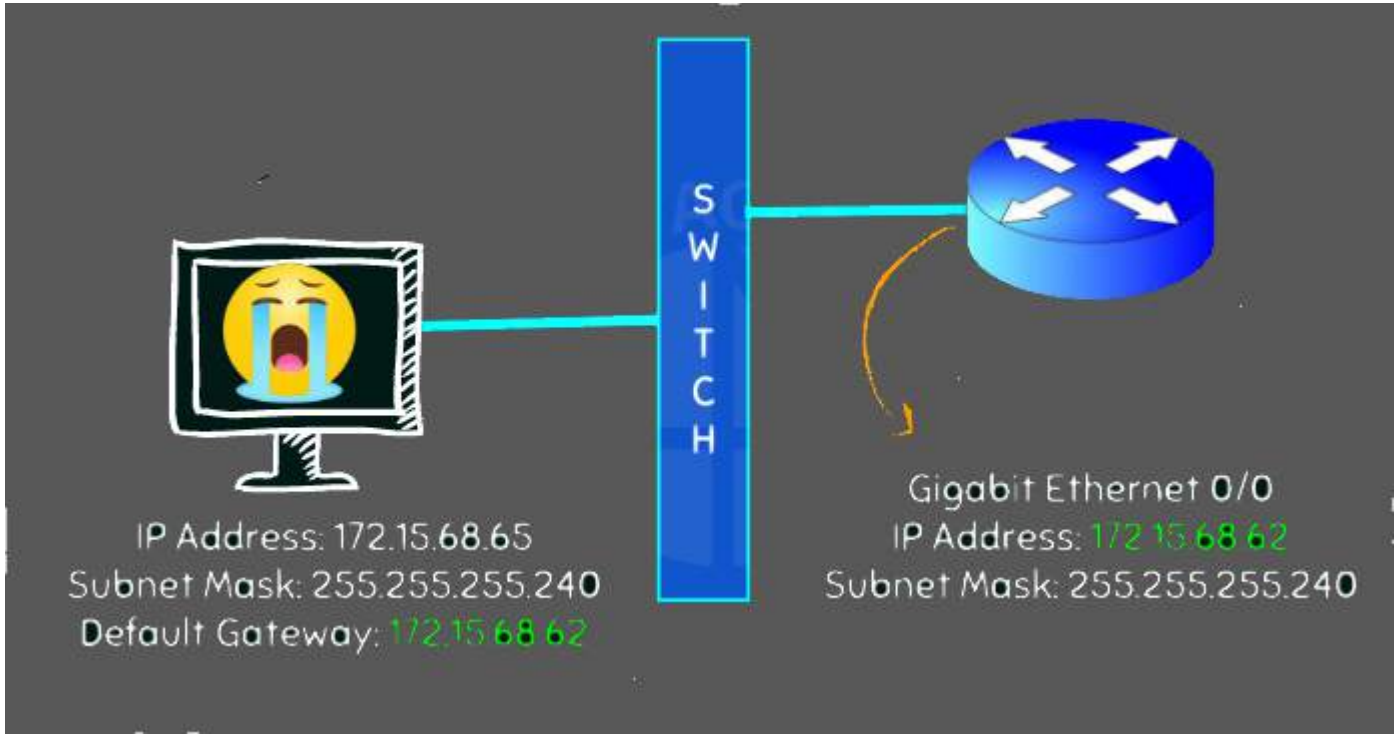
Q. Subnet the IP address 10.0.0.0 into 100 hosts in each subnet.

Trouble shooting in Subnetting

A computer is assigned with an IP address 192.168.1.127 and subnet mask 255.255.255.224. But the computer is not able to communicate with the PCs in the network, why?

Trouble shooting in Subnetting

A computer is assigned with an IP address 172.15.68.65 and subnet mask 255.255.255.240 and the default gateway for connecting to the internet or other computer in different is 172.15.68.62. But the computer is not able to communicate with the PCs in the other network, why?



Module 3

Network Layer protocols

Dr. Sunandita Debnath, IIIT Vadodara

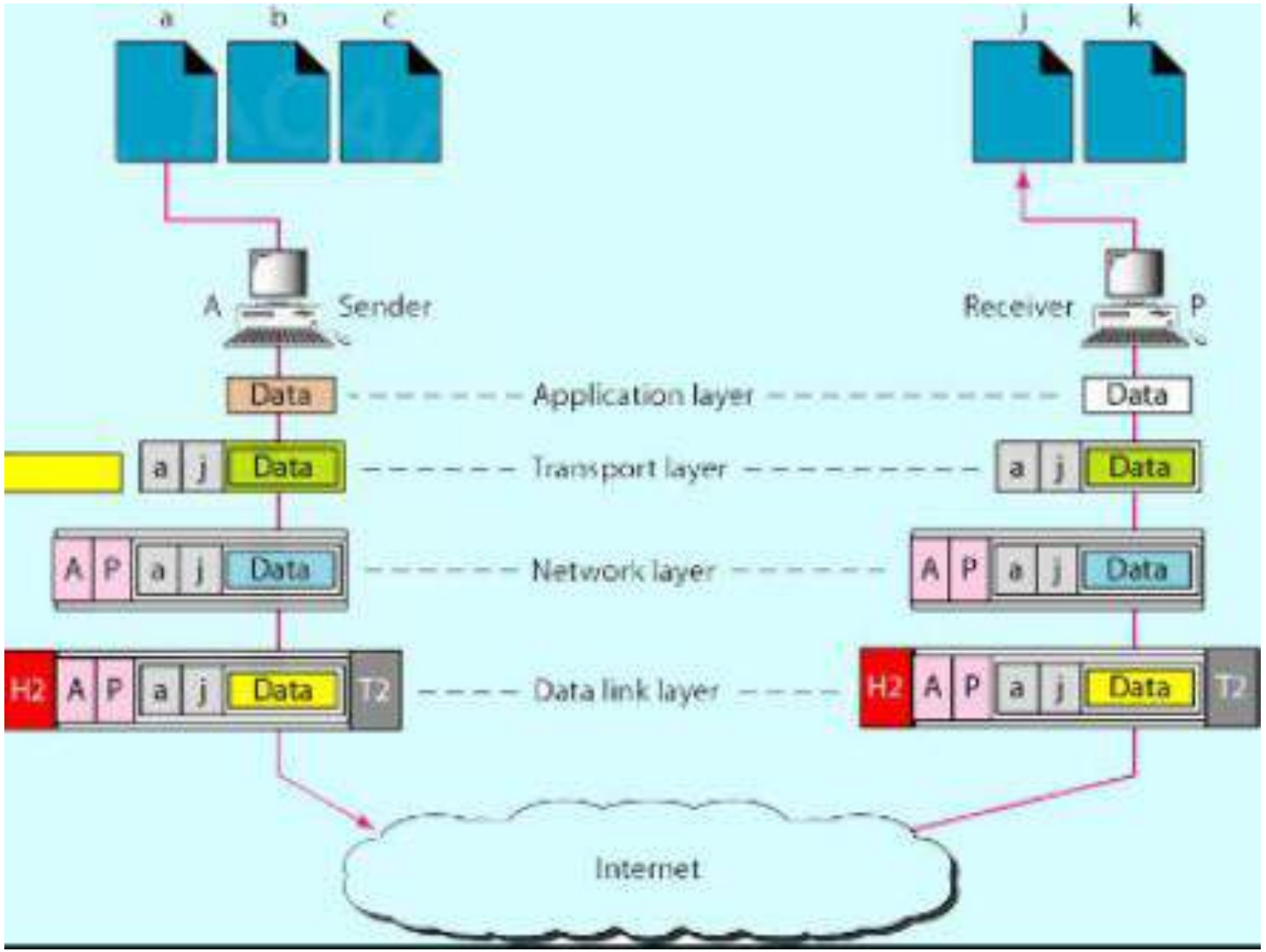
Different ways for Communication

- *Transport layer deals with port addressing, Port no is of 16 bits (0 to 65535)*

- *Well known ports (0 to 1023)*

- *Registered ports (1024 to 49151)*

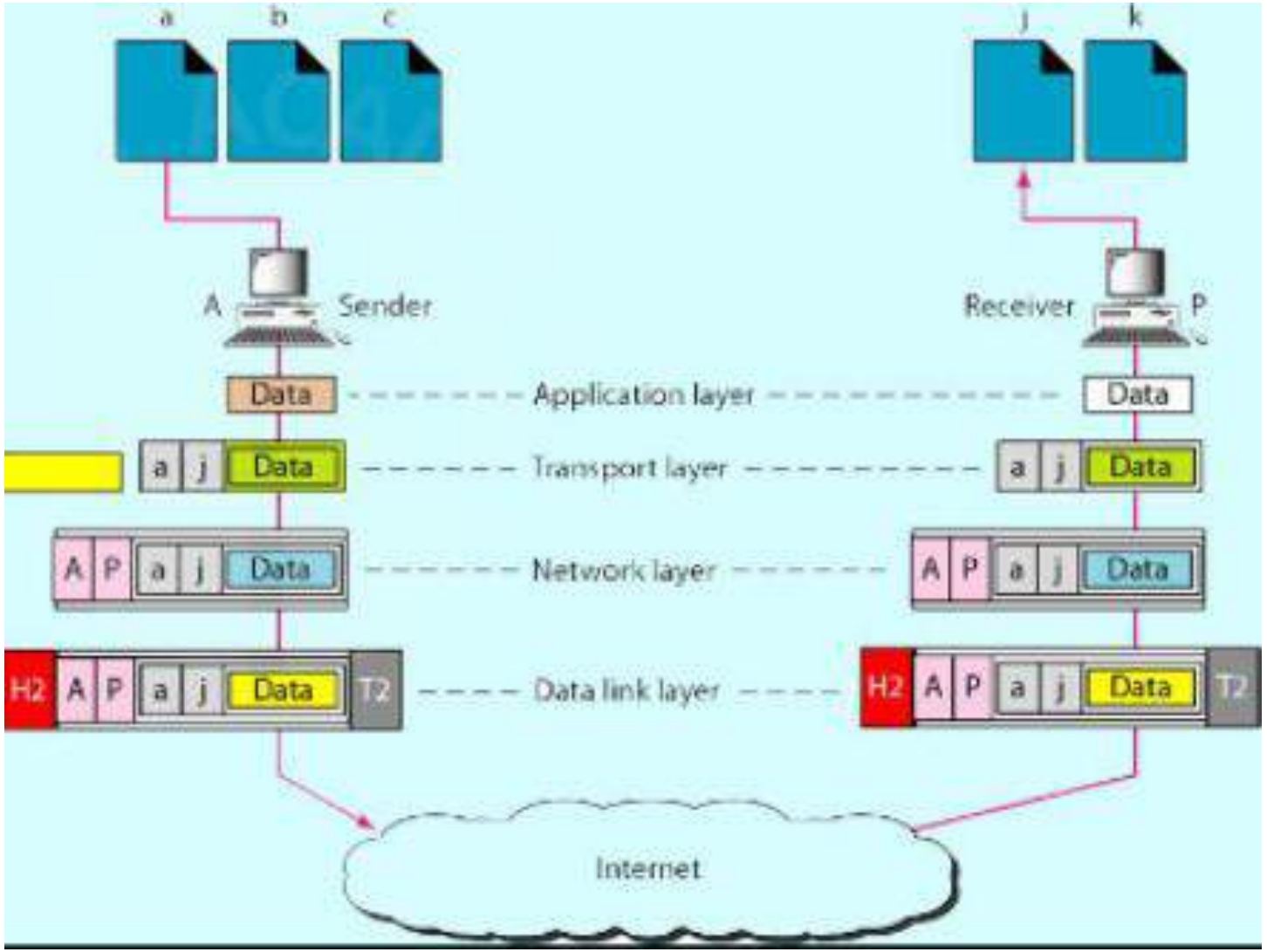
- *Dynamic or private ports (49152 to 65535)*



Different ways for Communication

- *Network layer or Internet layer deals with IP addresses IPV4 addresses are 32 bits and IPV6 addresses are 128 bits long (Hexadecimal).*

- *Either user provides the IP addresses or DNS resolves the names into IP addresses.*

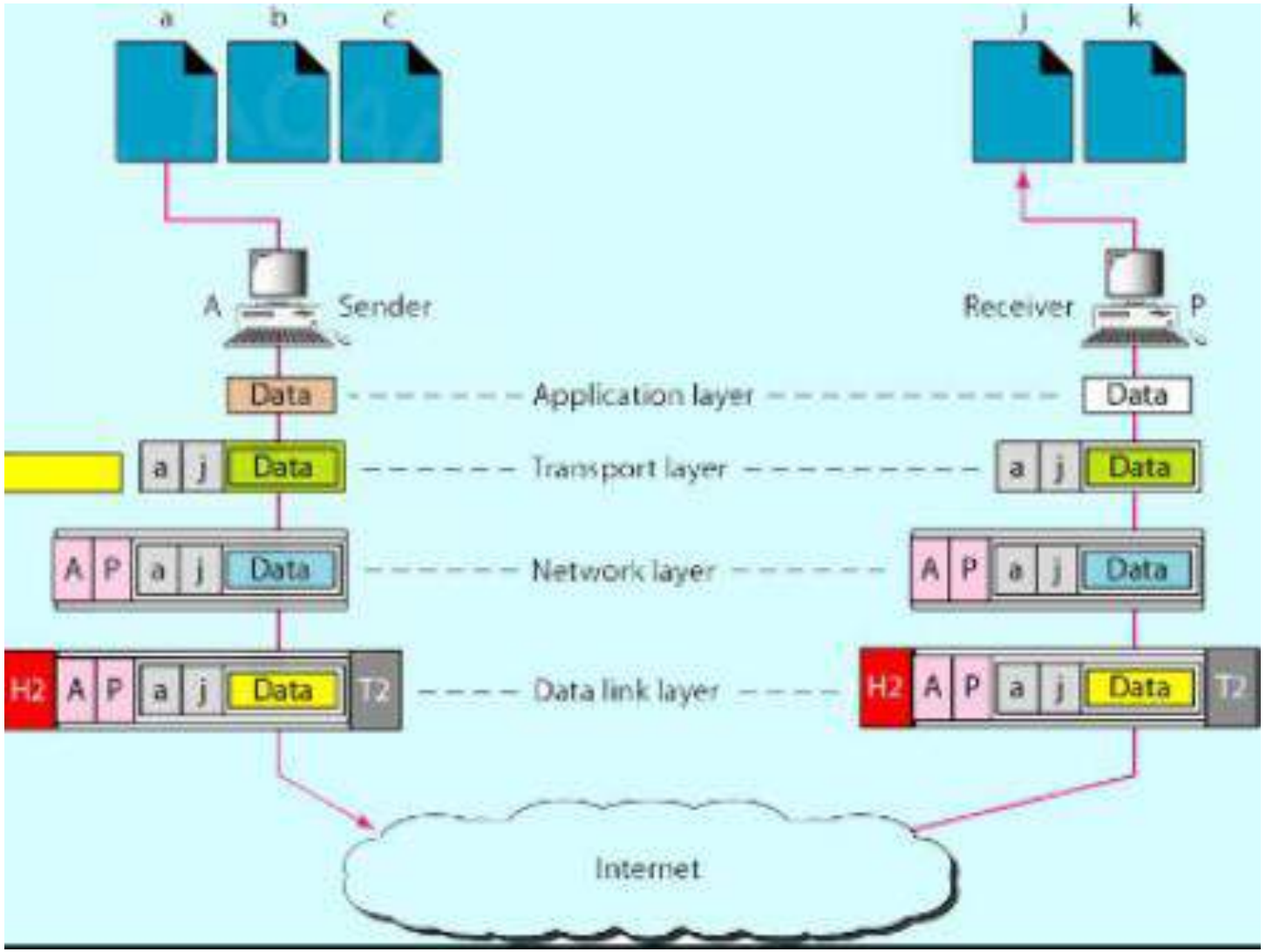


Different ways for Communication

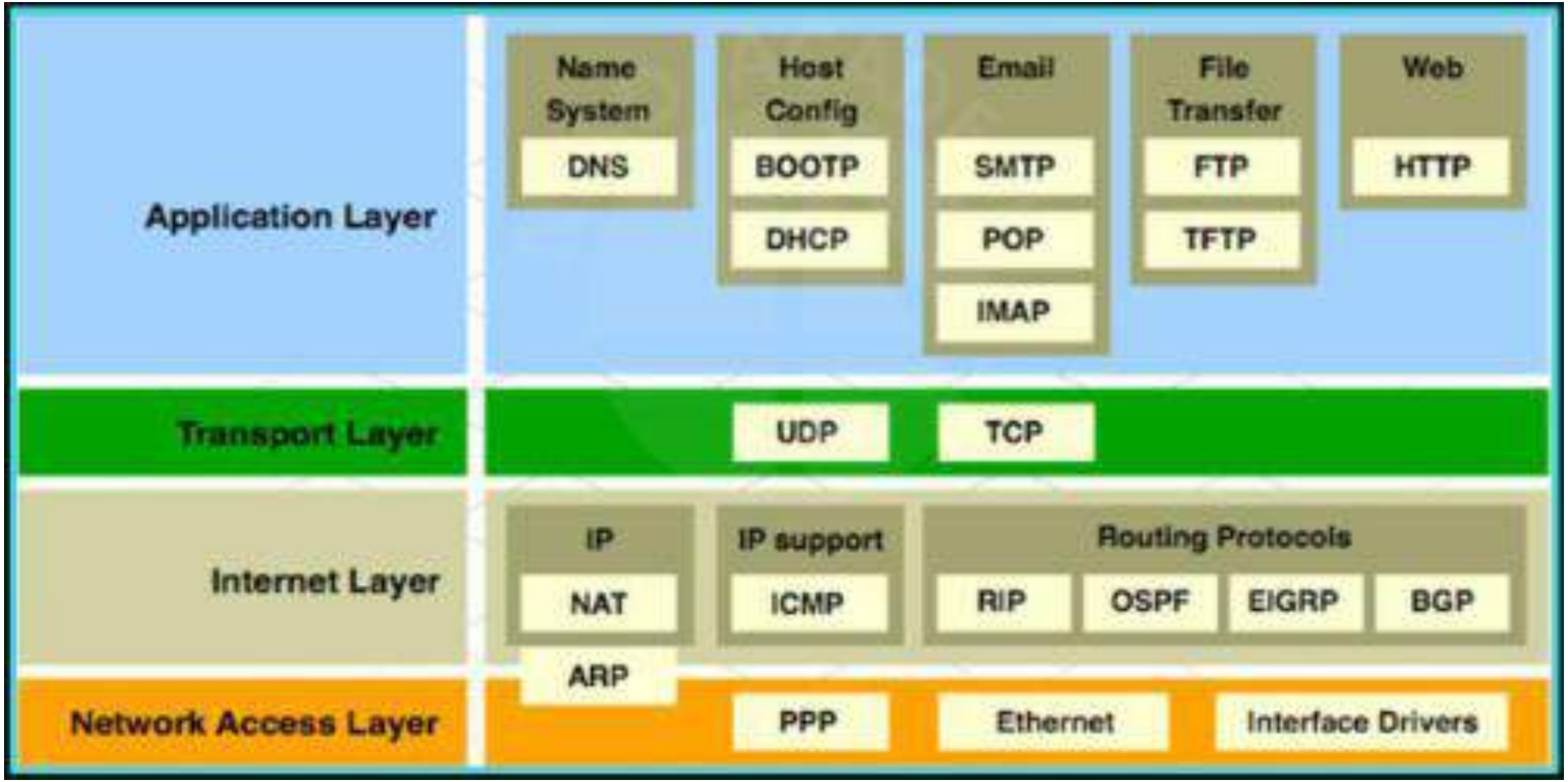
- *Data link deals with MAC addresses or hardware addresses. MAC addresses are 48 bits long (Hexadecimal).*

- *Assigned by manufacturer and usually wont be changed.*

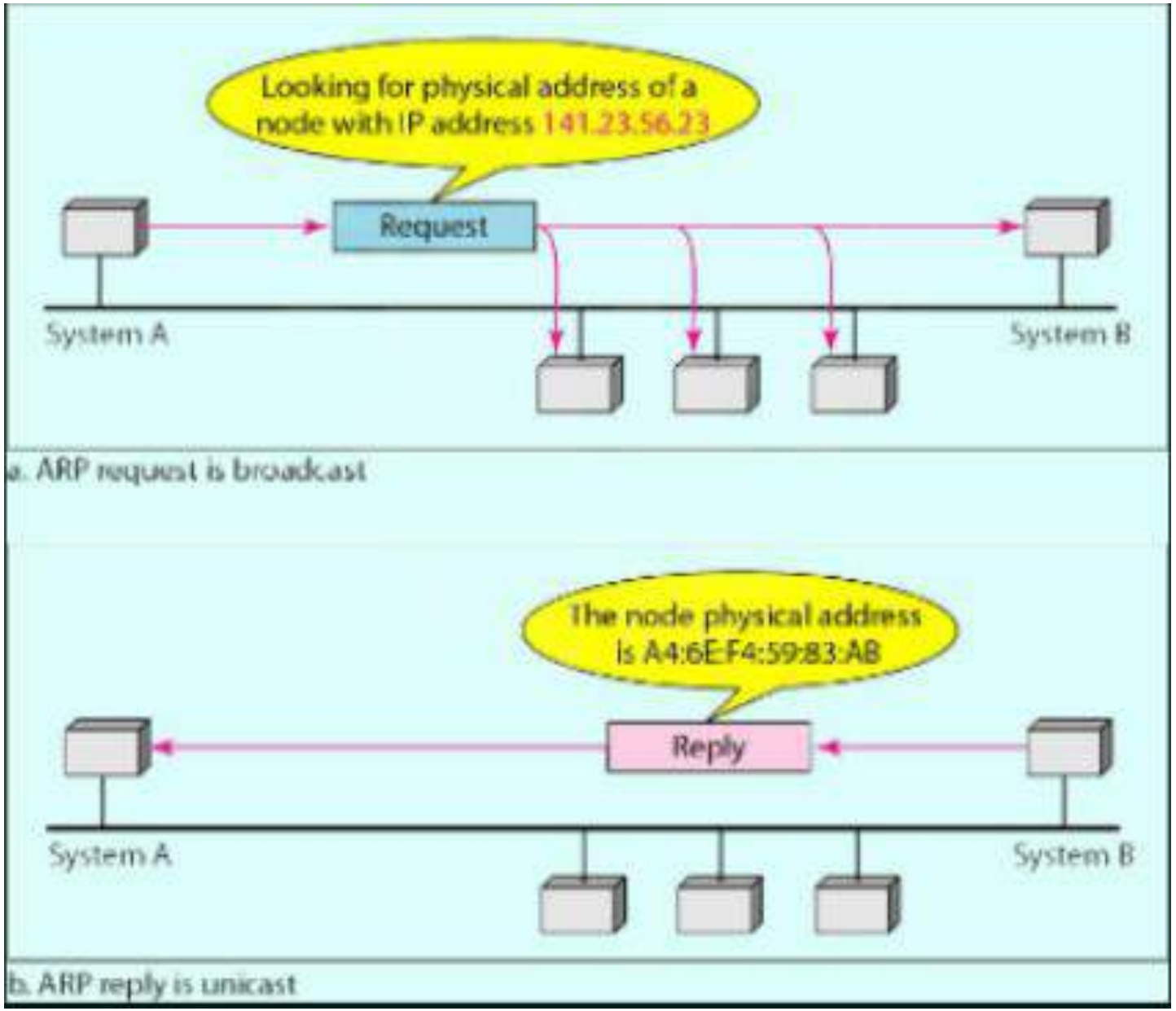
- ***The mapping between the IP address and the MAC address is done by the ARP (Address Resolution Protocol).***



TCP/IP suite


Application Layer
Name System
DNS
Host Config
BOOTP
DHCP
Email
SMTP
POP
IMAP
File Transfer
FTP
TFTP
Web
HTTP
Transport Layer
UDP
TCP
Internet Layer
IP
NAT
ARP
IP support
ICMP
Routing Protocols
RIP
OSPF
EIGRP
BGP
Network Access Layer
PPP
Ethernet
Interface Drivers

Address Resolution Protocol (ARP)



ARP Protocol

The ARP protocol provides two basic functions:

- *Resolving IPV4 address to MAC address.*
- *Maintaining a table of mapping.*

ARP Protocol

- *ARP maintains a cache of IP addresses and MAC addresses.*
- *ARP request is sent with an broadcast MAC address FFFF.FFFF.FFFF if the IP address ins not in the table.*
- *Target machine responds (unicast) with its physical address.*
- *Table entries are discarded if not refreshed.*

Header of ARP Packet

0	8	16	31
Hardware type = 1		ProtocolType = 0x0800	
HLen = 48	PLen = 32	Operation	
SourceHardwareAddr (bytes 0–3)			
SourceHardwareAddr (bytes 4–5)		SourceProtocolAddr (bytes 0–1)	
SourceProtocolAddr (bytes 2–3)		TargetHardwareAddr (bytes 0–1)	
TargetHardwareAddr (bytes 2–5)			
TargetProtocolAddr (bytes 0–3)			

Header of ARP Packet

- ***Hardware Type:*** *Type of Physical network (i.e. Ethernet)*
- ***Protocol Type:*** *Type of upper layer protocol (e.g. IP)*
- ***HLEN and PLEN:*** *Length of physical address and Protocol address*
- ***Operation:*** *ARP Request or a ARP Reply*
- *Source Physical address /Protocol address & Target Physical address /Protocol addresses*

RARP Protocol

The ARP protocol provides two basic functions:

- *Reverse ARP (RARP) is an obsolete computer networking protocol used by a cleuicent computer to request its internet protocol address from a computer network, when all it has available is its link layer or hardware address, such as MAC address.*

- *The protocol can use the known MAC address to retrieve its IP addrwss. Functionality wise, RARP is the complete opposite of ARP. The RAP uses the known IP address to determine the MAC address of the computer.*

- *It has been rendered obsolete by the Bootstrap protocol (BOOTP) and the modern dynamic host configuration protocol (DHCP), which both support a much greater feature set than ARP.*

- *The RARP cannot handle subnetting because no subnet masks are sent. If multiple subnets are deployed, an RARP server must be available in each subnet.*

DHCP

- *Dynamic Host Configuration Protocol .*

- *The MAC addresses are configured into network by manufacturer and they are expected to be unique.*

- *IP addresses must be unique in a inter-network. Most operating system provide away to manually configure the IP information for the host.*

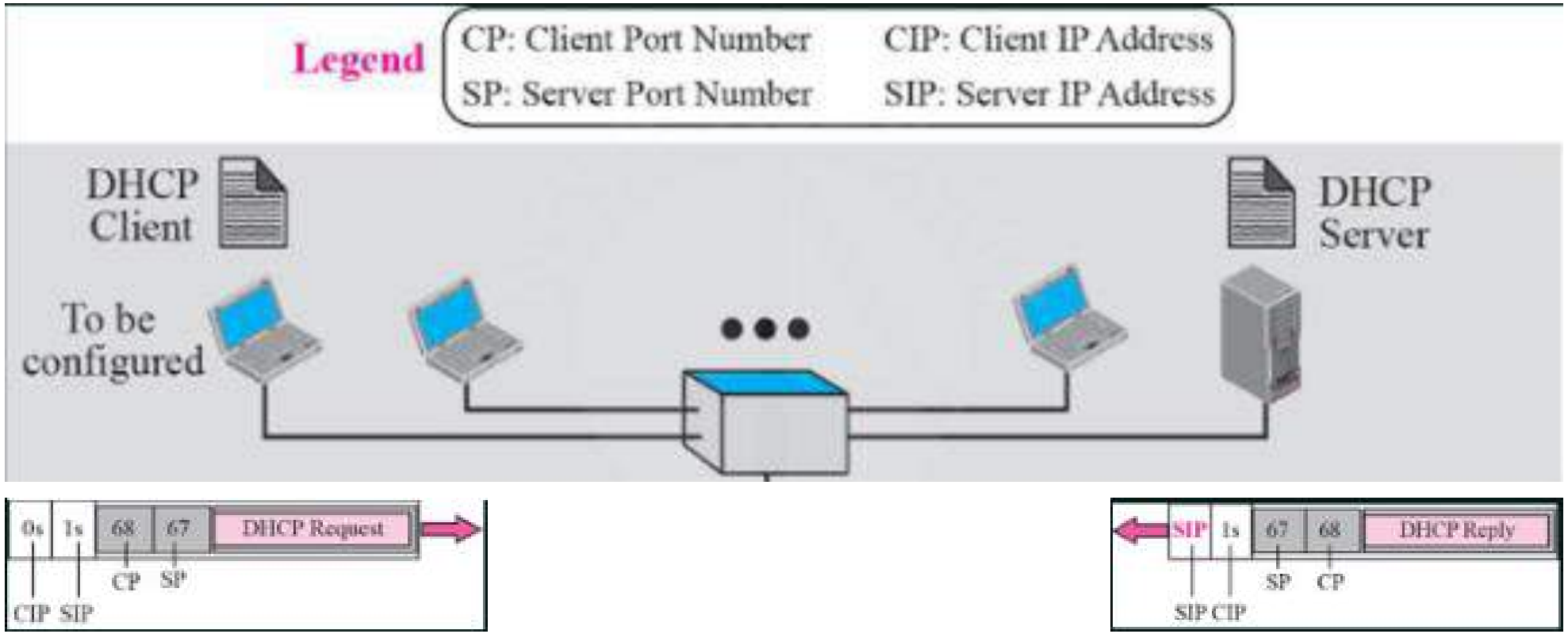
- *Drawback of manual configuration:*

A lot of work to configure all the hosts in a large network.
Configuration process in error prone.

- *Hence automated IP configuration is required.*

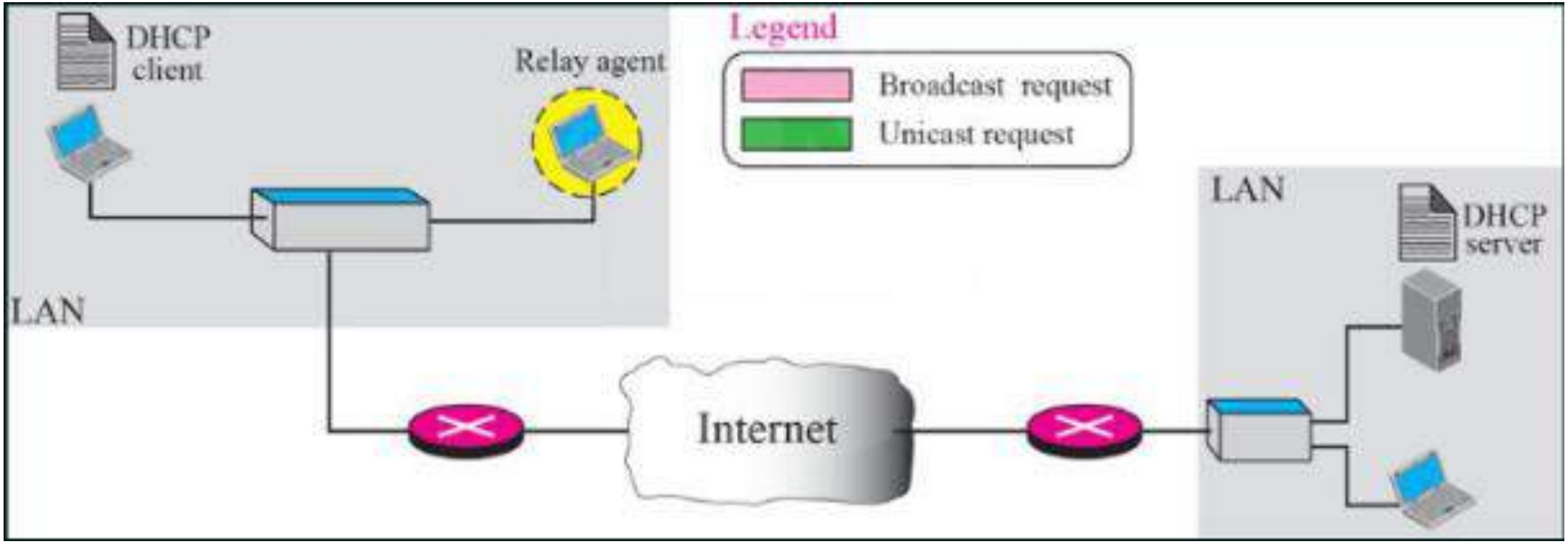
DHCP

❑*DHCP server and Host in same network*



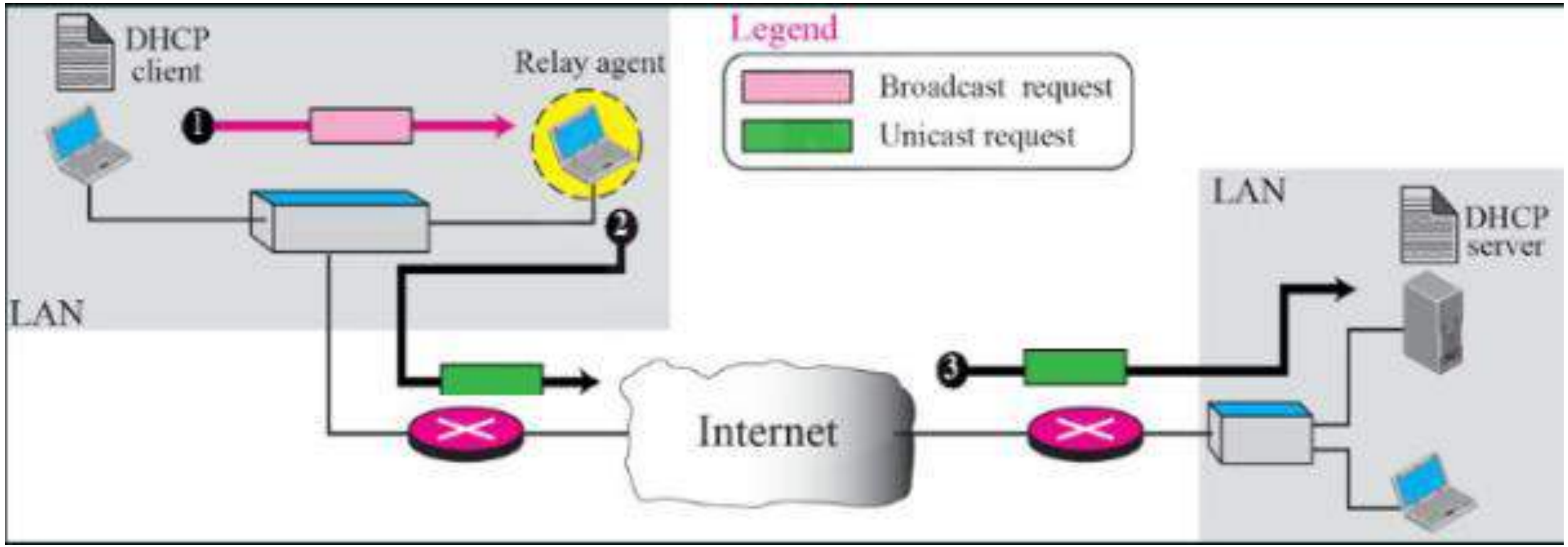
DHCP

❑*DHCP server and Host in Different network*



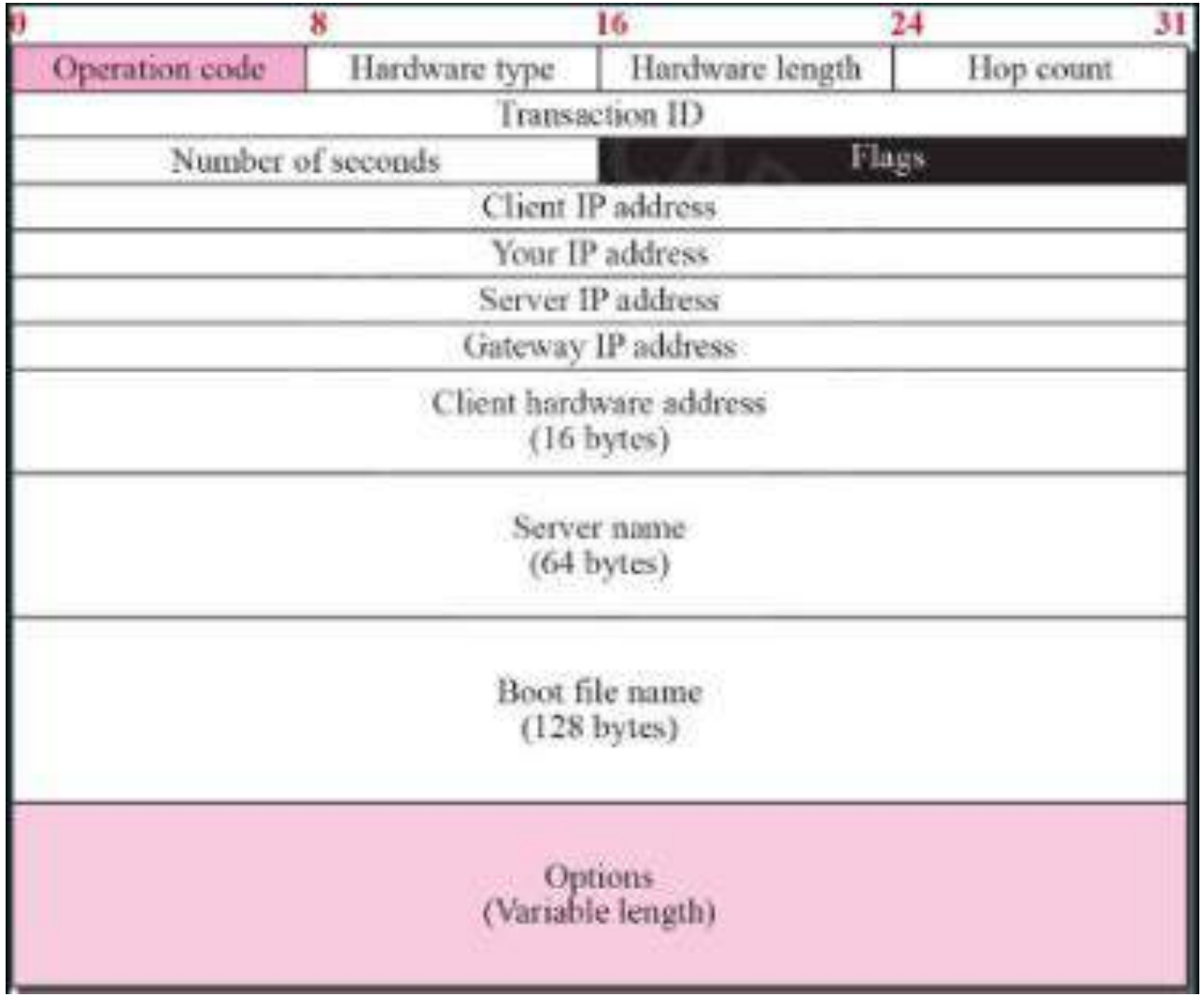
DHCP

- *Dynamic Host Configuration Protocol .*



DHCP

❑ *DHCP Packet Format*

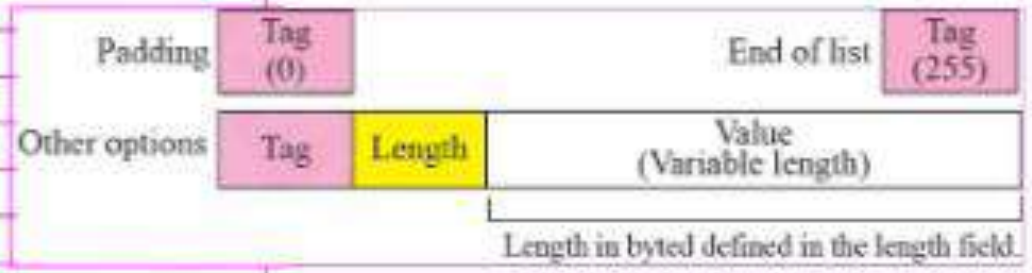


DHCP

❑ *DHCP Packet Format*

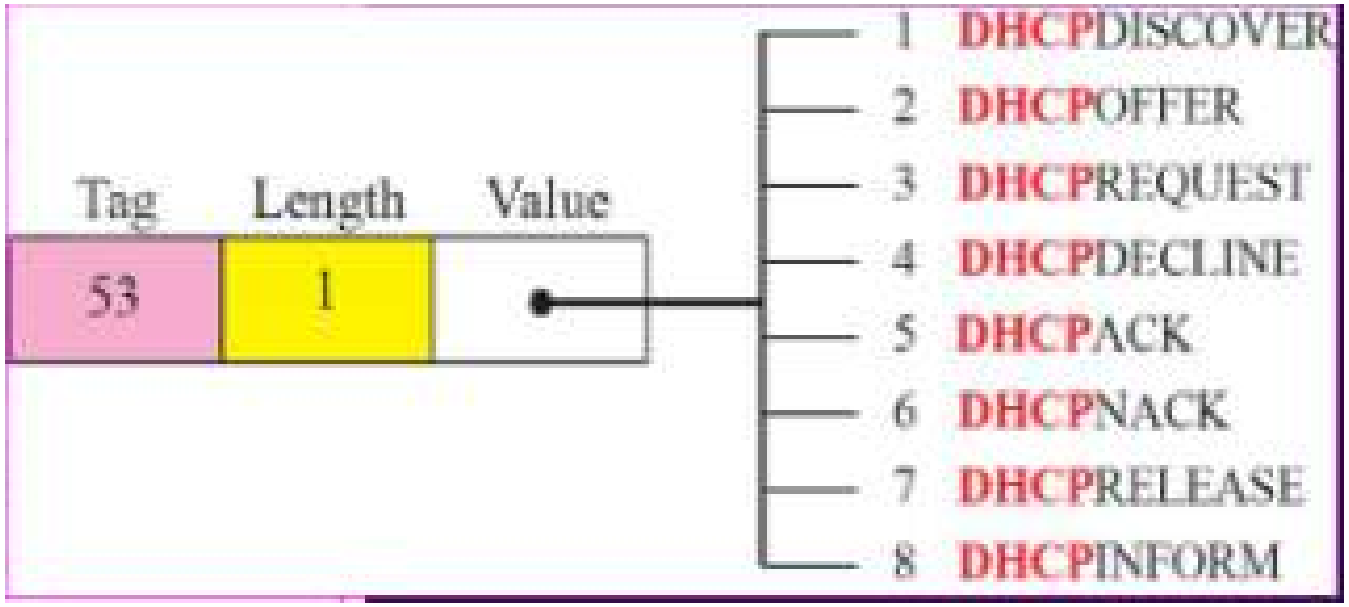
Options for DHCP

Tag	Length	Value	Description
0			Padding
1	4	Subnet mask	Subnet mask
2	4	Time of the day	Time offset
3	Variable	IP addresses	Default router
4	Variable	IP addresses	Time server
5	Variable	IP addresses	IEN 16 server
6	Variable	IP addresses	DNS server
7	Variable	IP addresses	Log server
8	Variable	IP addresses	Quote server
9	Variable	IP addresses	Print server
10	Variable	IP addresses	Impress
11	Variable	IP addresses	RLP server
12	Variable	DNS name	Host name
13	2	Integer	Boot file size
53	1	Discussed later	Used for dynamic configuration
128–254	Variable	Specific information	Vendor specific
255			End of list



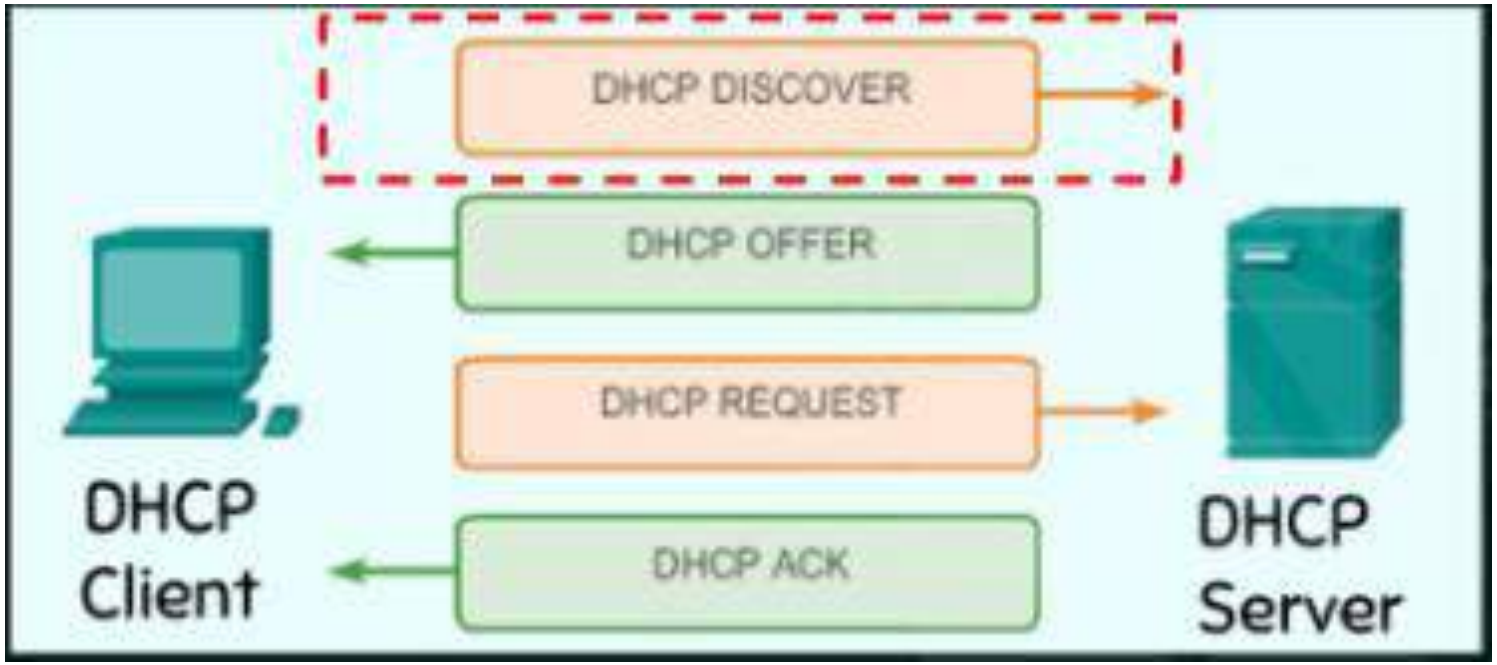
DHCP

❑*DHCP Packet Format*



DHCP

❑*DHCP Operation*



ICMP

- *Internet Control Message Protocol.*

- *ICMP is a supporting protocol in the internet protocol suite. It is used by network devices, including routers, to send error messages and operational information indicating success or failure when communicating with another IP address.*

- *Example :*
  - *Destination host unreachable due to link/node failure*
  - *Reassembly process failed*
  - *TTL (Time To Live) had reached 0 (So datagram don't cycle forever)*
  - *IP header Checksum failed.*

- *ICMP messages are typically used for diagnostic or control purpose or generated in response to errors in IP operation.*

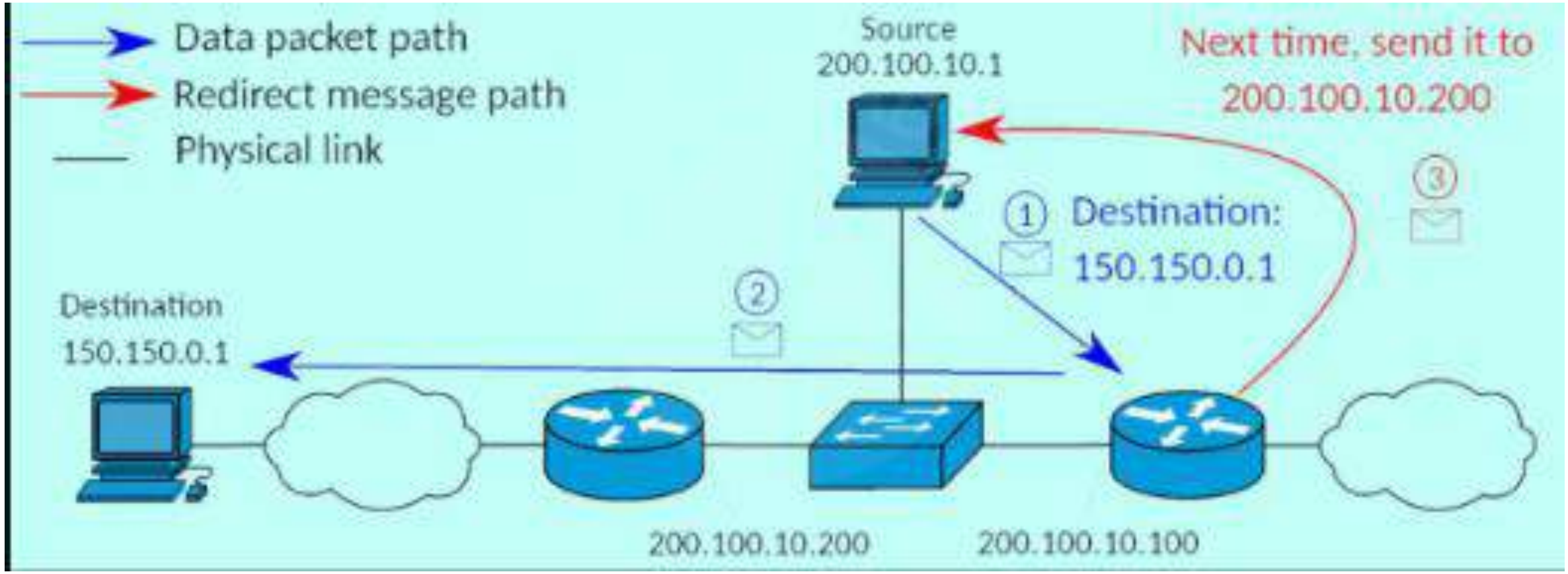
ICMP

- *ICMP errors are directed to the source IP address of the originating packet*

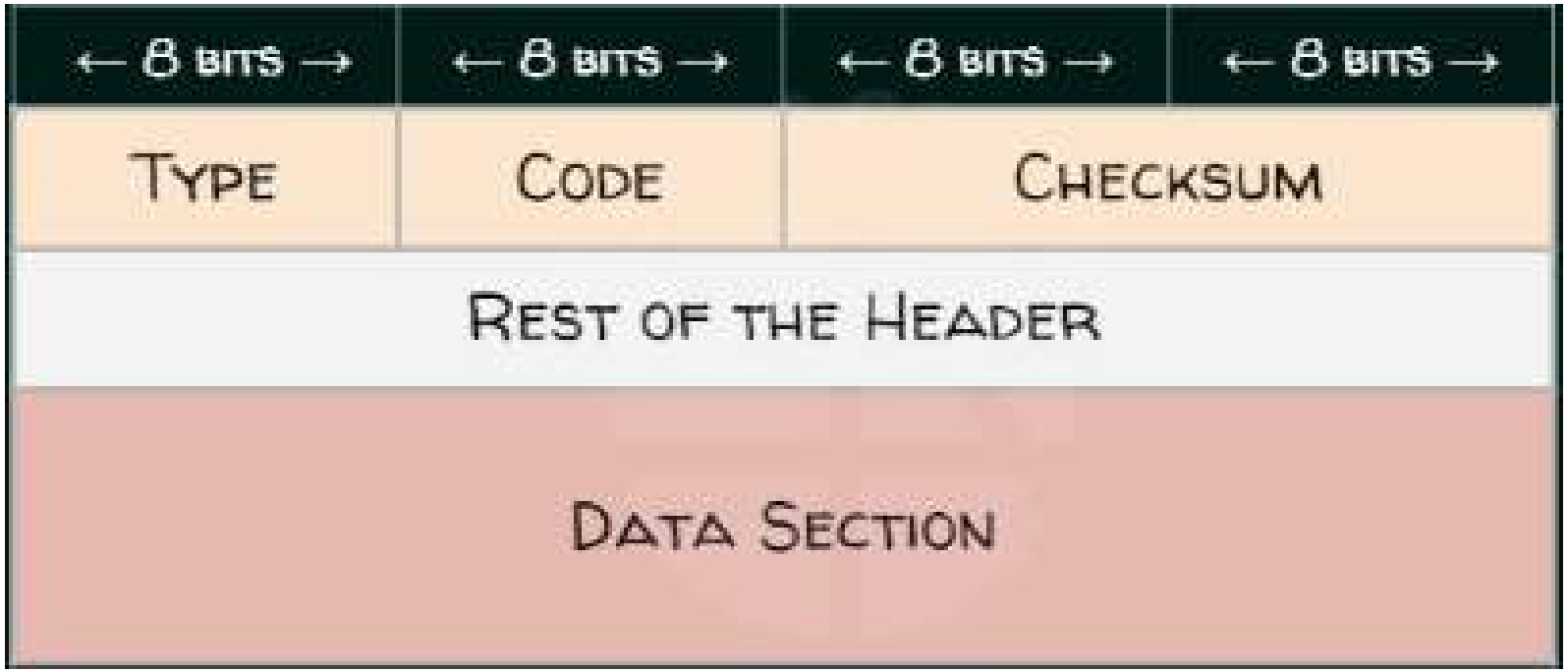
- *This protocol defines a collection of error messages that are sent back to the source host whenever a router or host is unable to process an IP packet successfully.*

- *ICMP redirect*
  - *From router to the source host*
  - *With a better route information*

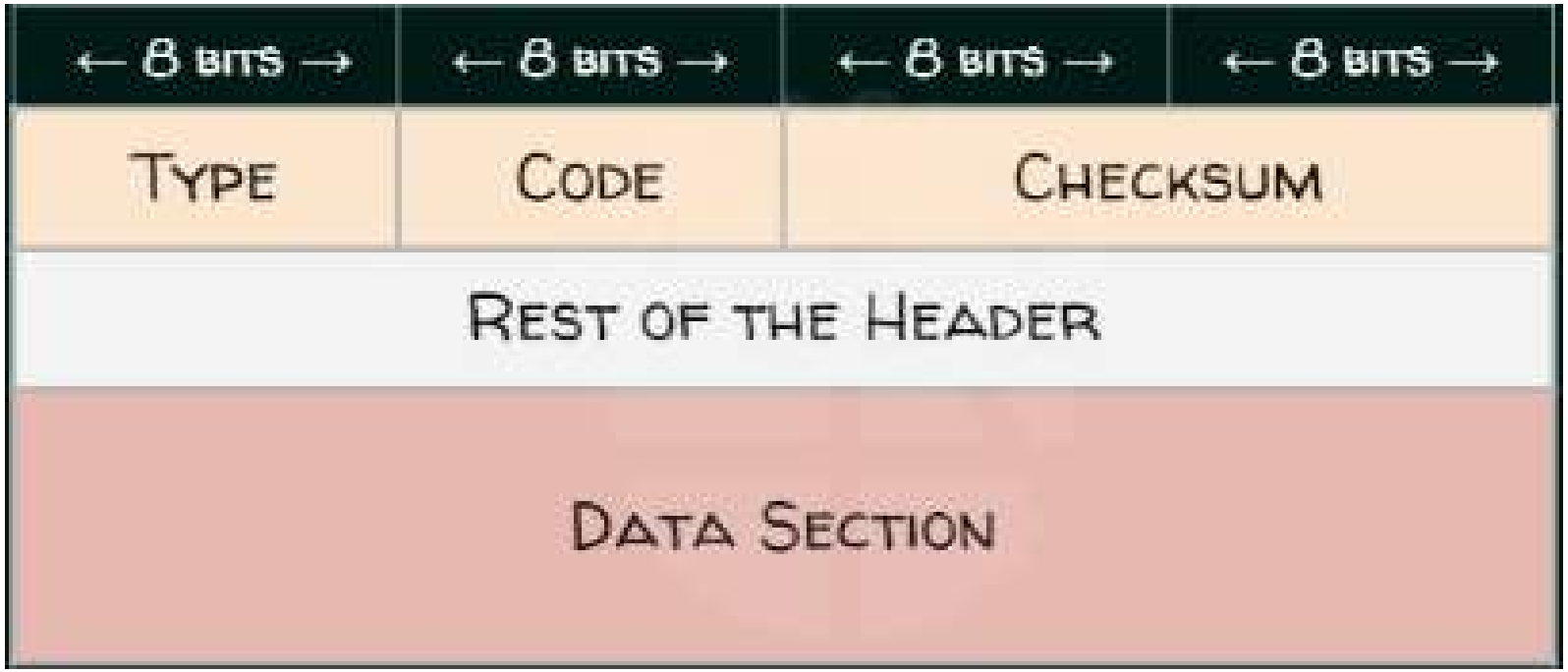
ICMP Redirect



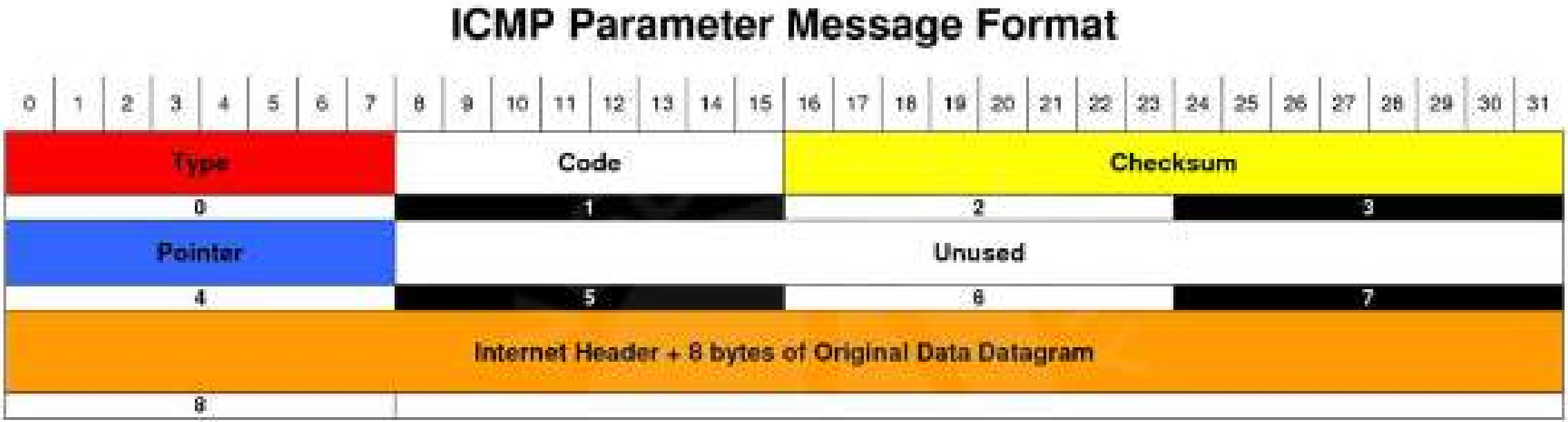
ICMP Packet Format



ICMPv4 Packet Format



ICMPv4 Packet Format



ICMPv4 Packet Format

Type	Code	Meaning
0	0	Echo Reply
3	0	Net Unreachable
	1	Host Unreachable
	2	Protocol Unreachable
	3	Port
	4	Frag needed and DF set
	5	Source route failed
	6	Destination network unknown
	7	Destination Host unknown
	8	Source host isolated
	9	
	.	
	13	Communication admin prohibited

ICMPv4 Packet Format

Type	Code	Meaning
4	0	Source Quench
5	0	Redirect data gram from the network
	1	
	2	
	3	
8	0	Echo
9	0	Router advertisement
10	0	Router selection
11	0	TTL exceeded in transit

ICMP

- *Calculate the checksum for the following ICMPv4 packet*

8	0	0
1		9
TEST		

ICMP

IP protocol (Internet Protocol)

- *The responsibility of network later or Layer 3 is delivery of data from the original source to the destination network.*

- *Service provided by the network layer*
  - *Logical addressing*
  - *Routing*

Module 3

Routing Protocol

Dr. Sunandita Debnath, IIIT Vadodara

Various method to build a Routing Table

Static Routing	Dynamic Routing
A static routing table is created, maintained and updated by a network administrator.	*A dynamic routing table is created , maintained and updated by a routing protocol running on the router.*
A static route to every network must be configured on every router for full connectivity.	*The choice of the best route is in the hands of the routing protocol and not the network administrator.*
This provides a granular level of control over routing, but quickly becomes impractical on large networks.	*Example of routing protocols: RIP (Routing Information Protocol, OSPF (Open Shortest Path First) etc.*
Router will not share static routes with each other, thus reducing CPU/RAM overhead and saving bandwidth.	

Various method to build a Routing Table

Static Routing	Dynamic Routing
However, static routing is not fault tolerant , as any change to the routing infrastructure (such as a link going down, or a new network added) requires manual intervention.	*Routers do share dynamic routing information with each other, which increases CPU, RAM and bandwidth usage.*
Routers operating in a purely static environment cannot seamlessly choose a better route if a link becomes unavailable.	*However, Routing protocols are capable of dynamically choosing a different (or better) path when there is a change to the routing infrastructure.*

Various method to build a Routing Table

A static approach has several shortcomings:

- *It does not deal with node or link failures.*
- *It dies not consider the addition of new nodes or links.*
- *It implies that edge costs cannot change.*

What is the solution?

- *Need a distributed and dynamic protocol*

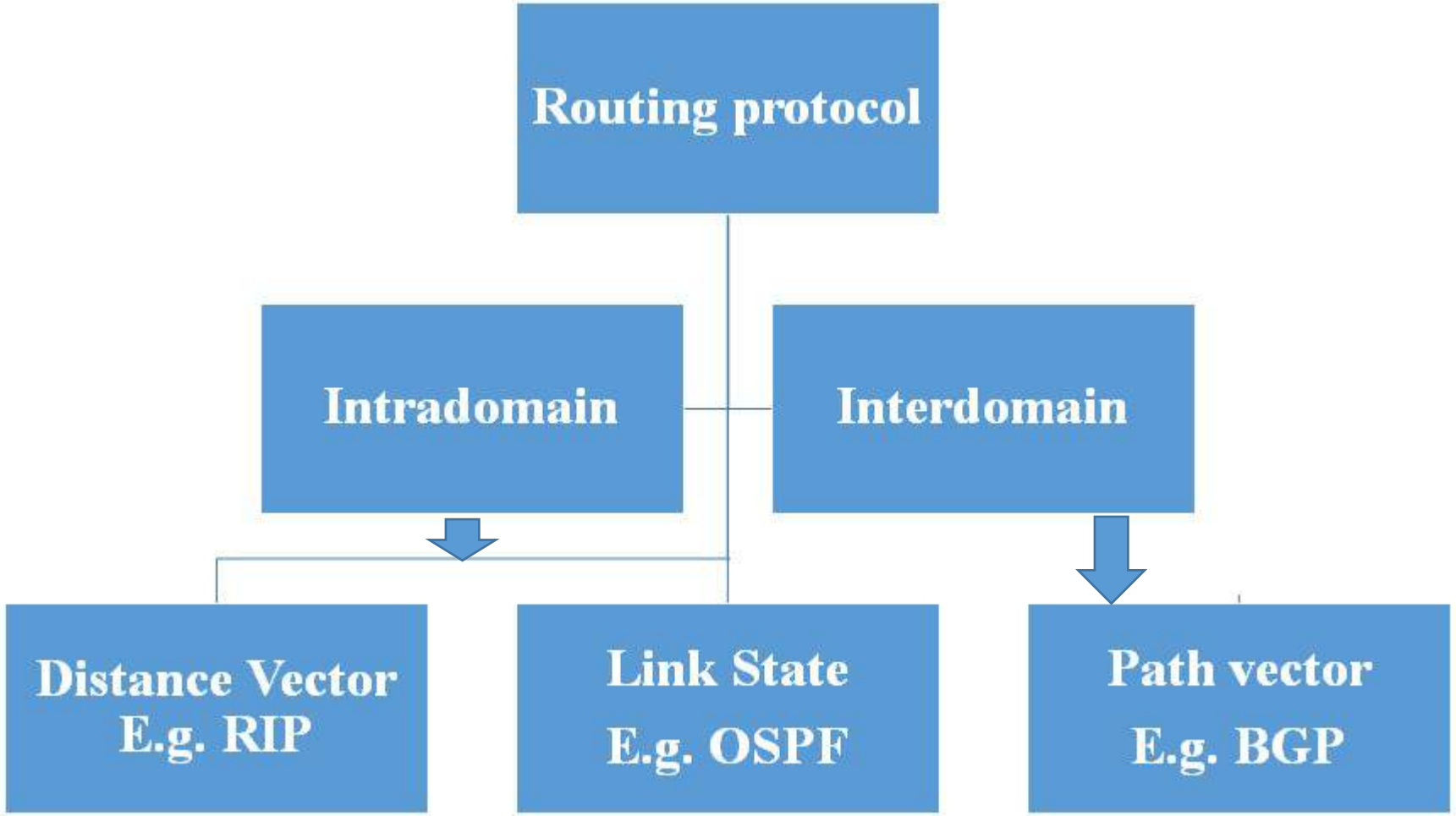
- ***Two main courses of protocols***

1. *Distance Vector*
2. *Link State*

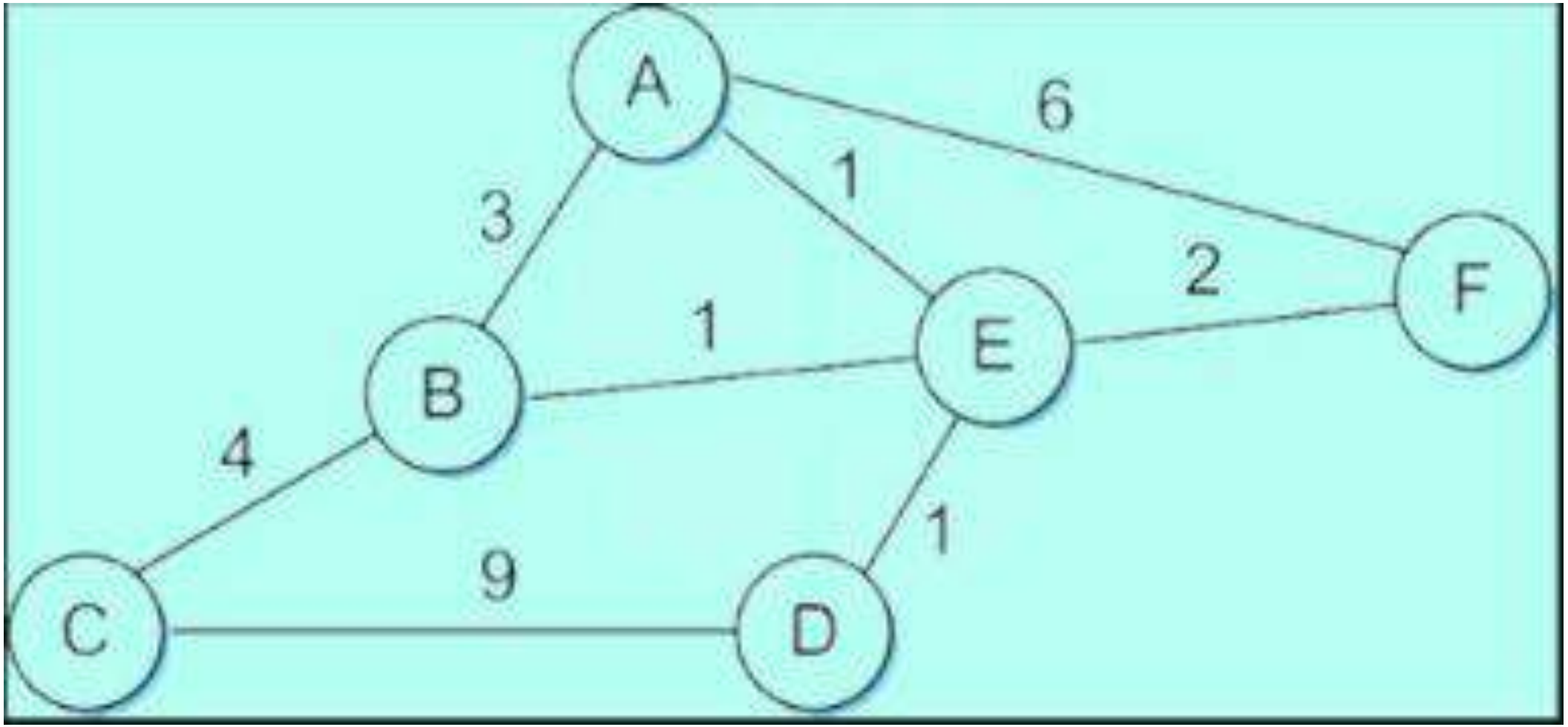
Dynamic Routing Protocols

- *A routing table can be either static and dynamic*
- *A static table is one with manual entries.*
- *A dynamic table is one that is updated automatically when there is a change somewhere in the internet.*
- *A routing protocol is a combination of rules and procedures that lets router in the internet inform each other of changes.*

Popular Routing Protocols



Network as a graph



Distance Vector Routing

- *Each node construct a one dimensional array (a vector) containing the distances (costs) to all other nodes and distributes that vector to its immediate neighbours.*

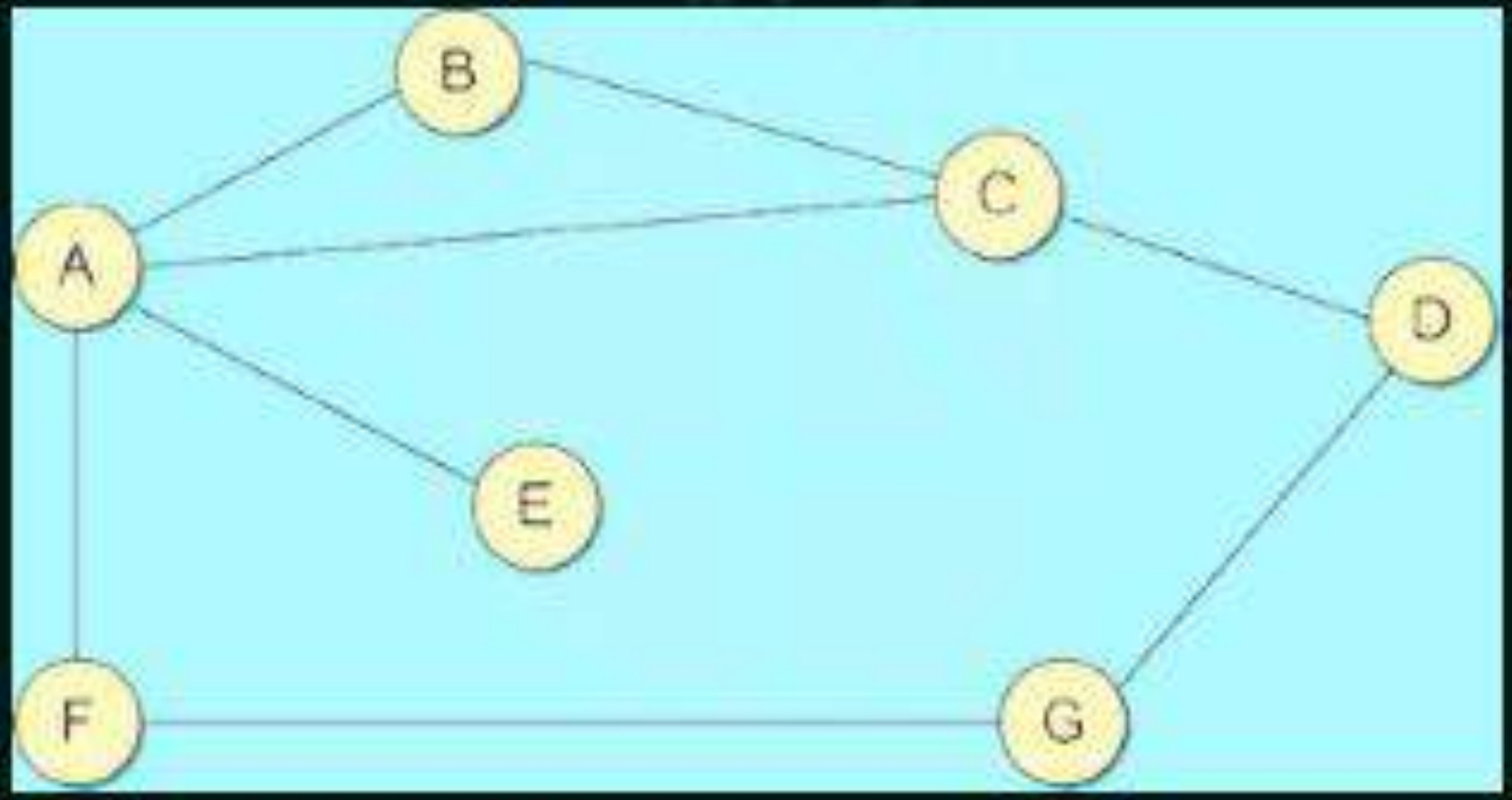
- *Starting assumption is that each node knows the cost of the link to each of its directly connected neighbours i.e. the immediate neighbours.*

- *The distance vector routing algorithm is sometimes called as Bellman-Ford algorithm.*

- *Every T seconds each router sends its table to its neighbour and each router then updates its table based on the new information.*

- *Problem includes fast response to good new and slow response to bad news.*

- *Also too many messages to update.*

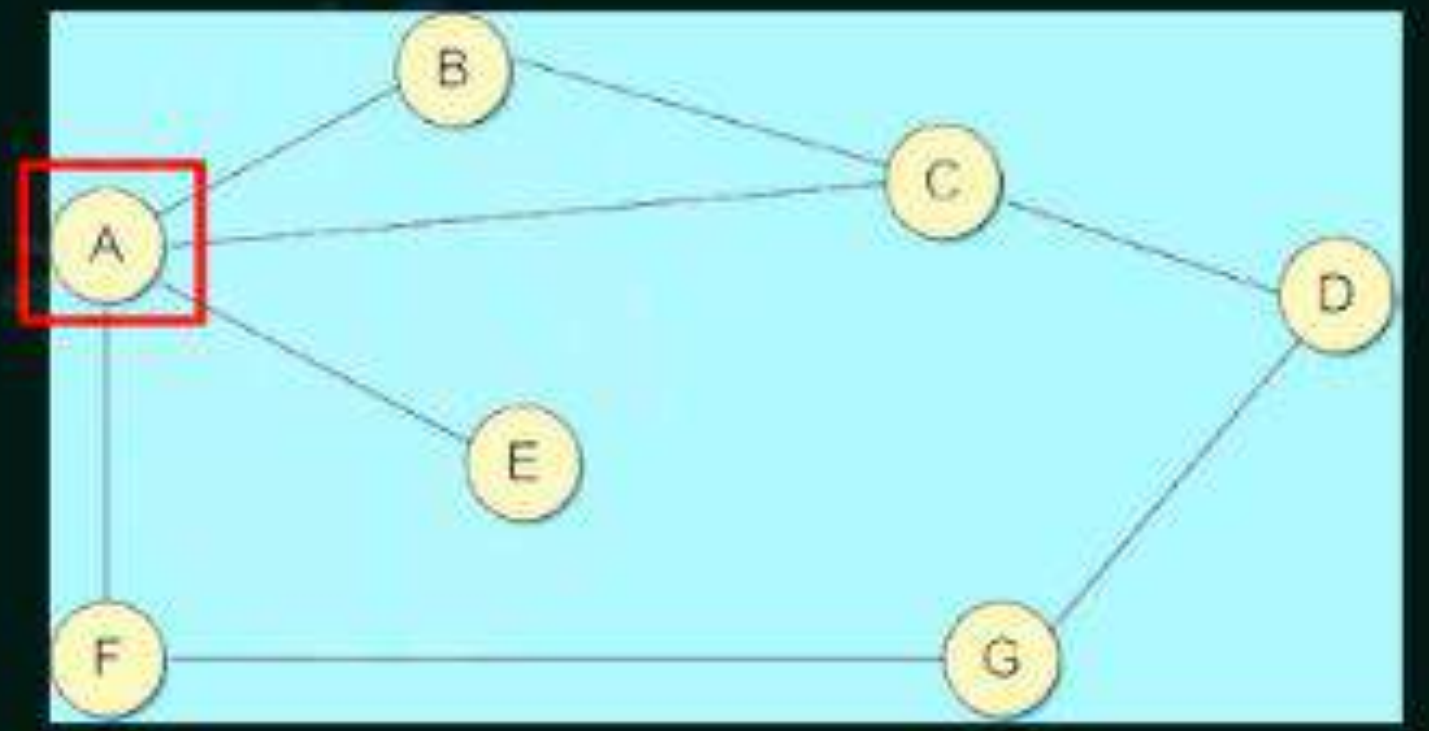


Routing Table Entries: (Destination, Distance/Cost, Next Hop)

Node A: (B,1,B), (C,1,C), (F,1,F), (E,1,E), (D,2,C) etc.,

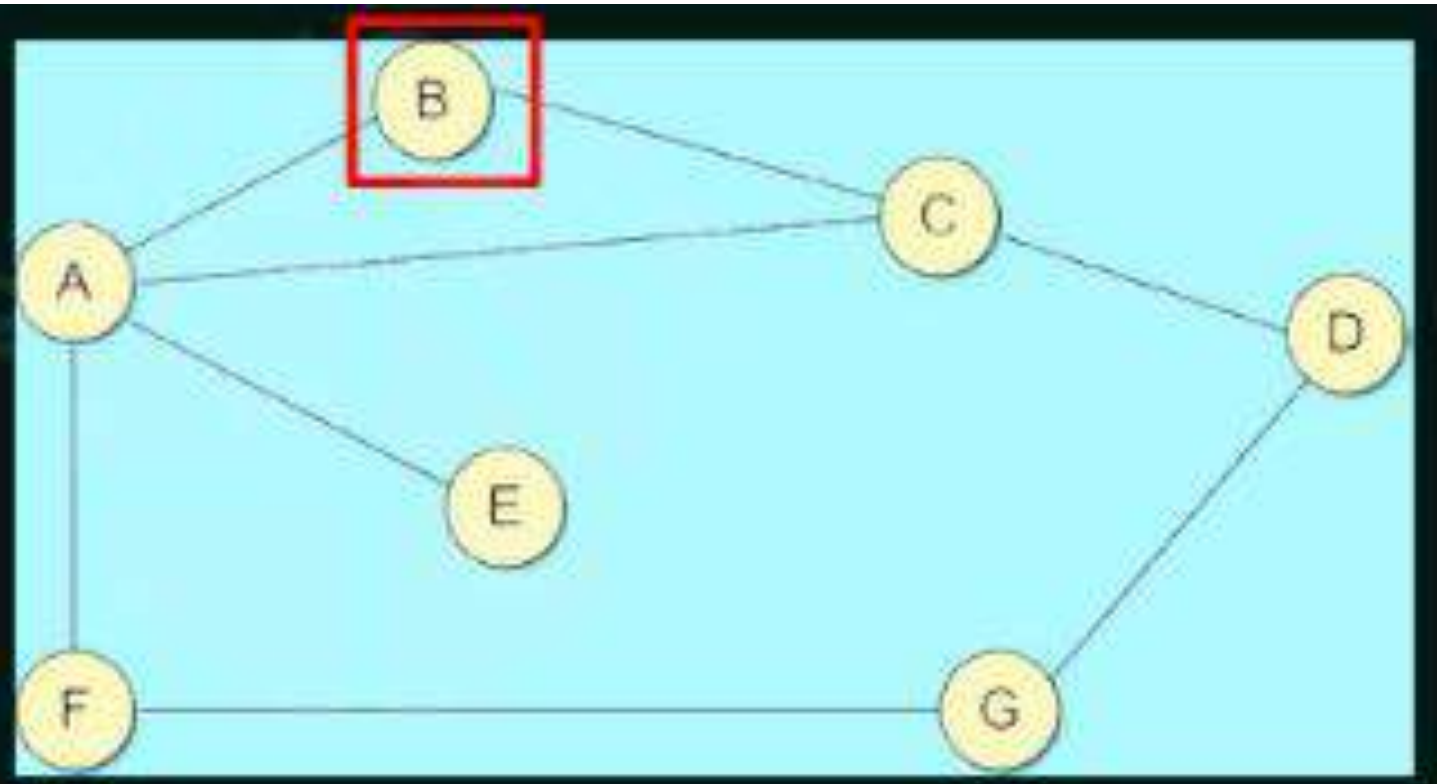
Initial Routing Table of A

Destination	Cost	Next Hop
A	0	-
B	1	B
C	1	C
D	∞	-
E	1	E
F	1	F
G	∞	-



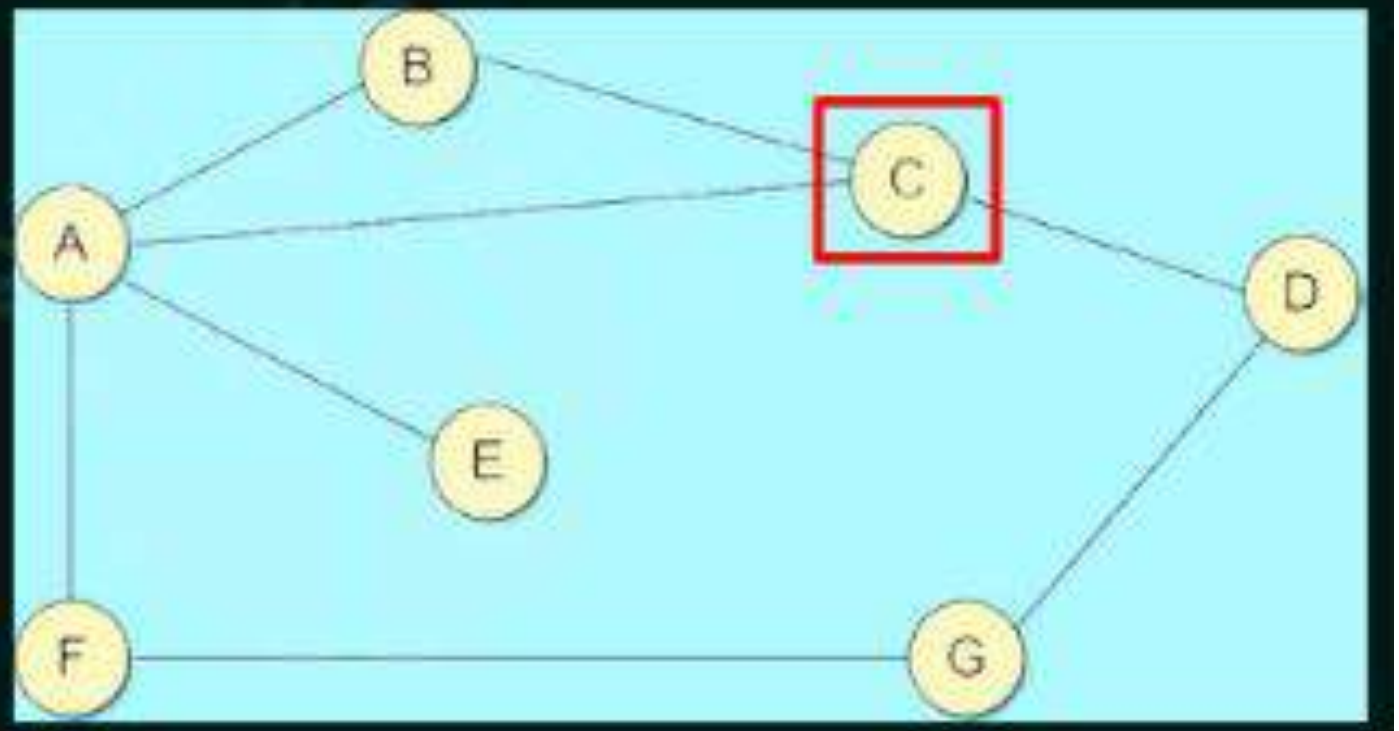
Initial Routing Table of B

Destination	Cost	Next Hop
A	1	A
B	0	–
C	1	C
D	∞	–
E	∞	–
F	∞	–
G	∞	–



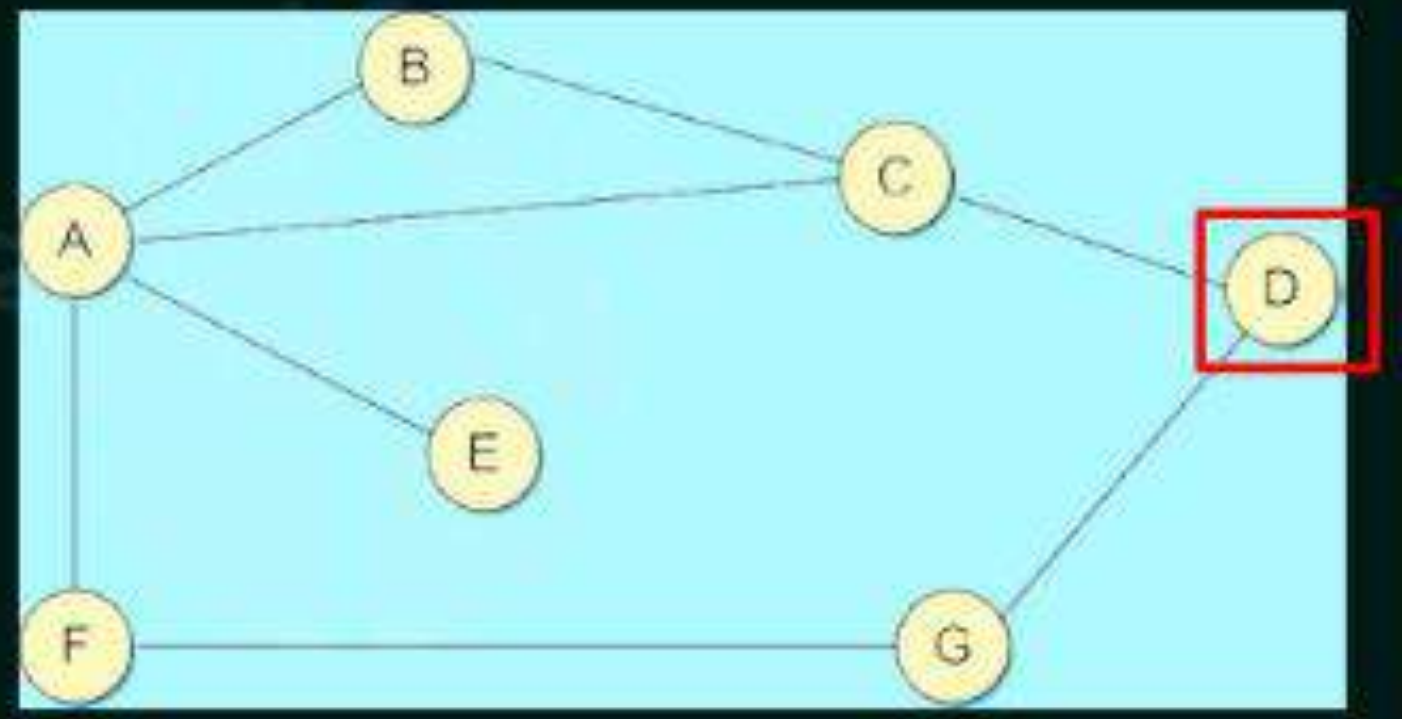
Initial Routing Table of C

Destination	Cost	Next Hop
A	1	A
B	1	B
C	0	-
D	1	D
E	∞	-
F	∞	-
G	∞	-



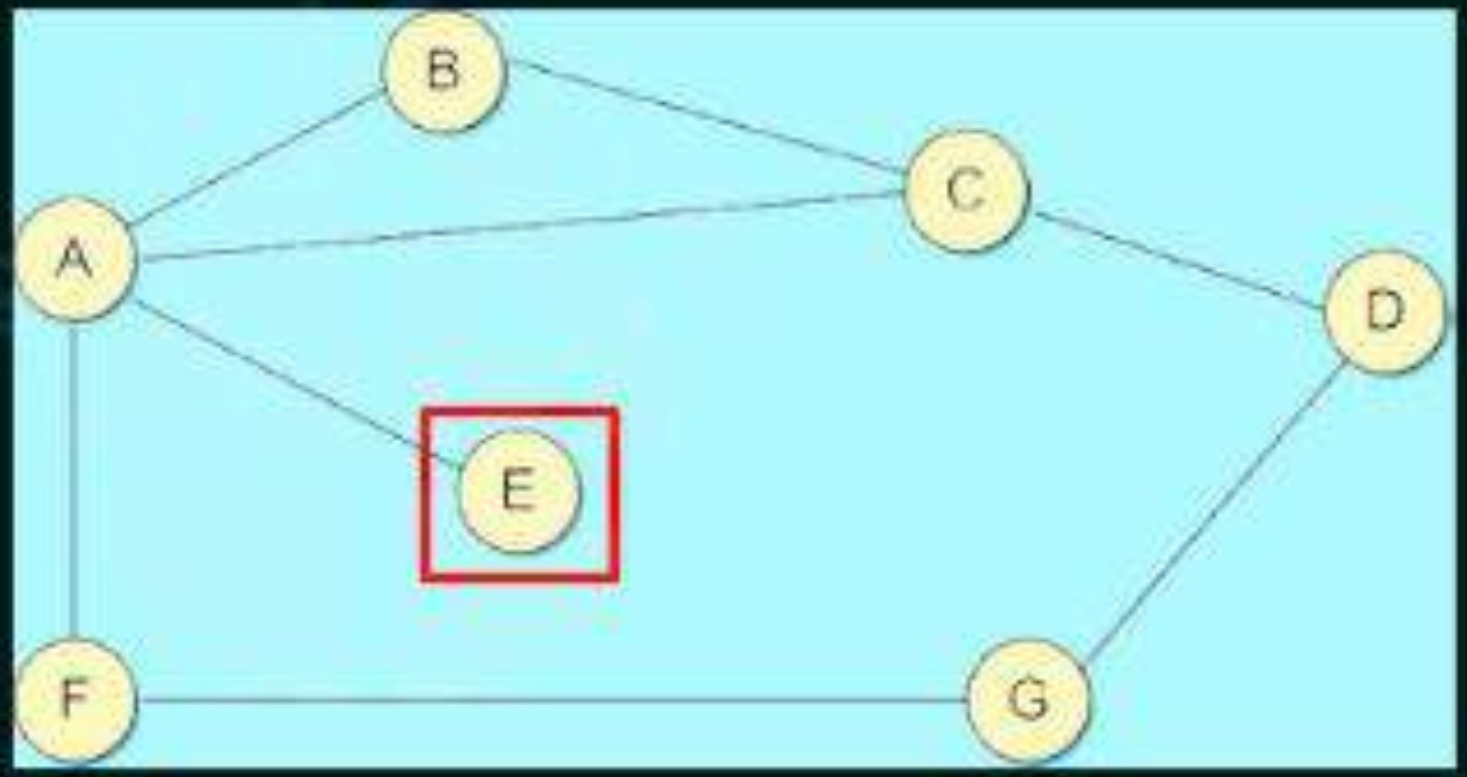
Initial Routing Table of D

Destination	Cost	Next Hop
A	∞	-
B	∞	-
C	1	C
D	0	-
E	∞	-
F	∞	-
G	1	G



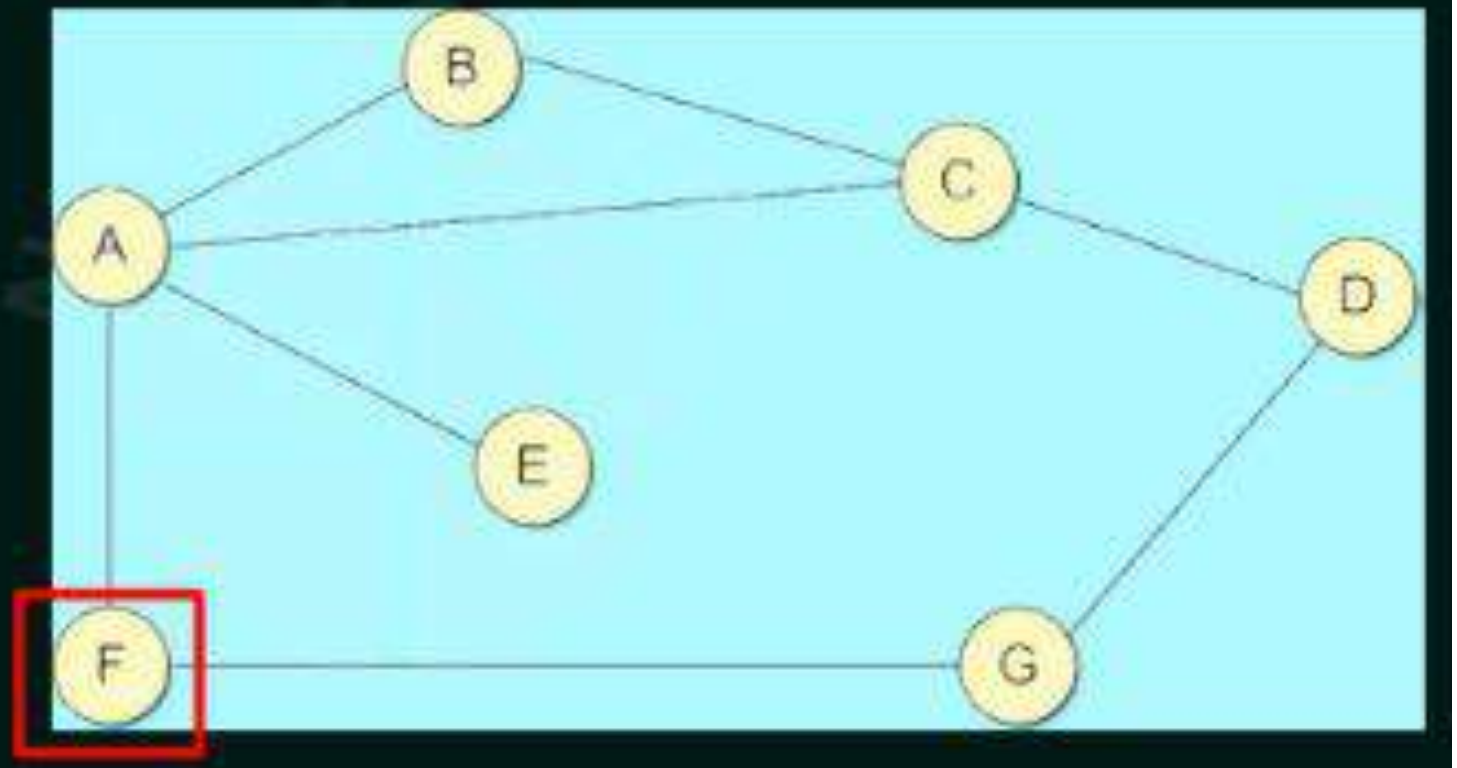
Initial Routing Table of E

Destination	Cost	Next Hop
A	1	A
B	∞	-
C	∞	-
D	∞	-
E	0	-
F	∞	-
G	∞	-



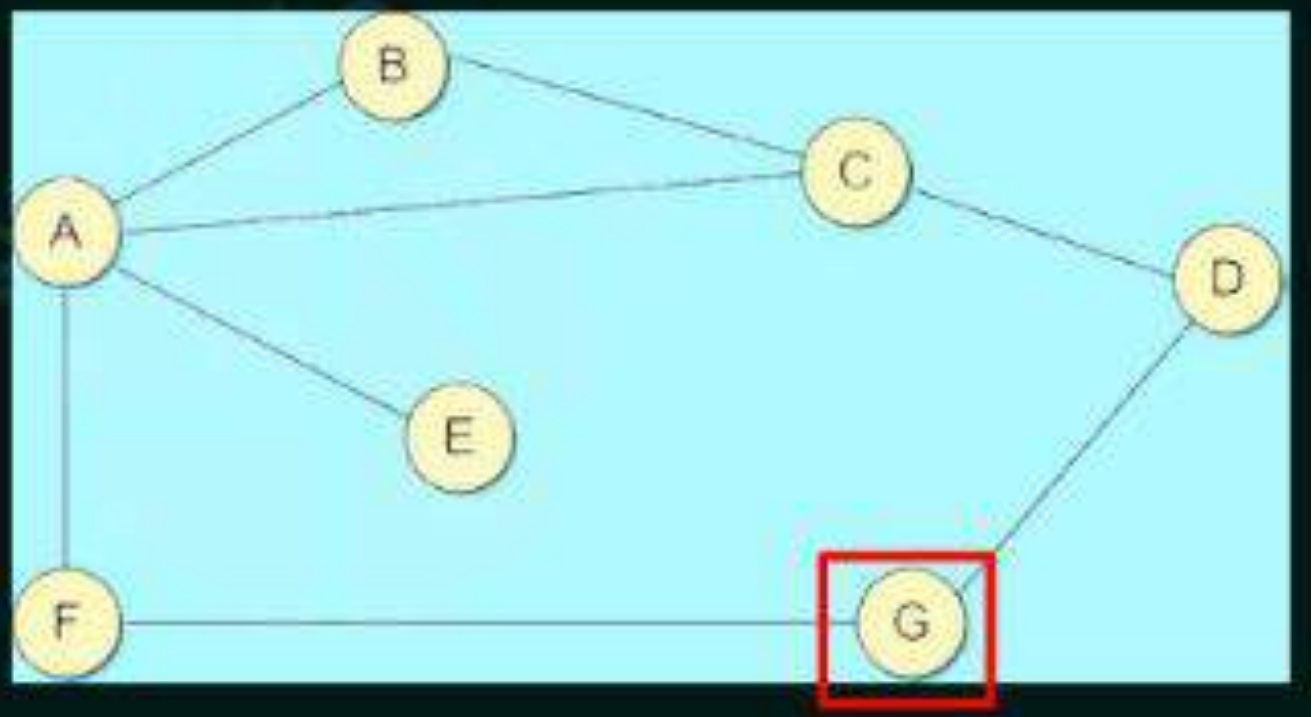
Initial Routing Table of F

Destination	Cost	Next Hop
A	1	A
B	∞	-
C	∞	-
D	∞	-
E	∞	-
F	0	-
G	1	G



Initial Routing Table of G

Destination	Cost	Next Hop
A	∞	-
B	∞	-
C	∞	-
D	1	1
E	∞	-
F	1	1
G	0	-



Initial Distance Vector stored at Each Node

Information stored at Node	Distance to reach node						
	A	B	C	D	E	F	G
A	0	1	1	∞	1	1	∞
B	1	0	1	∞	∞	∞	∞
C	1	1	0	1	∞	∞	∞
D	∞	∞	1	0	∞	∞	1
E	1	∞	∞	∞	0	∞	∞
F	1	∞	∞	∞	∞	0	1
G	∞	∞	∞	1	∞	1	0

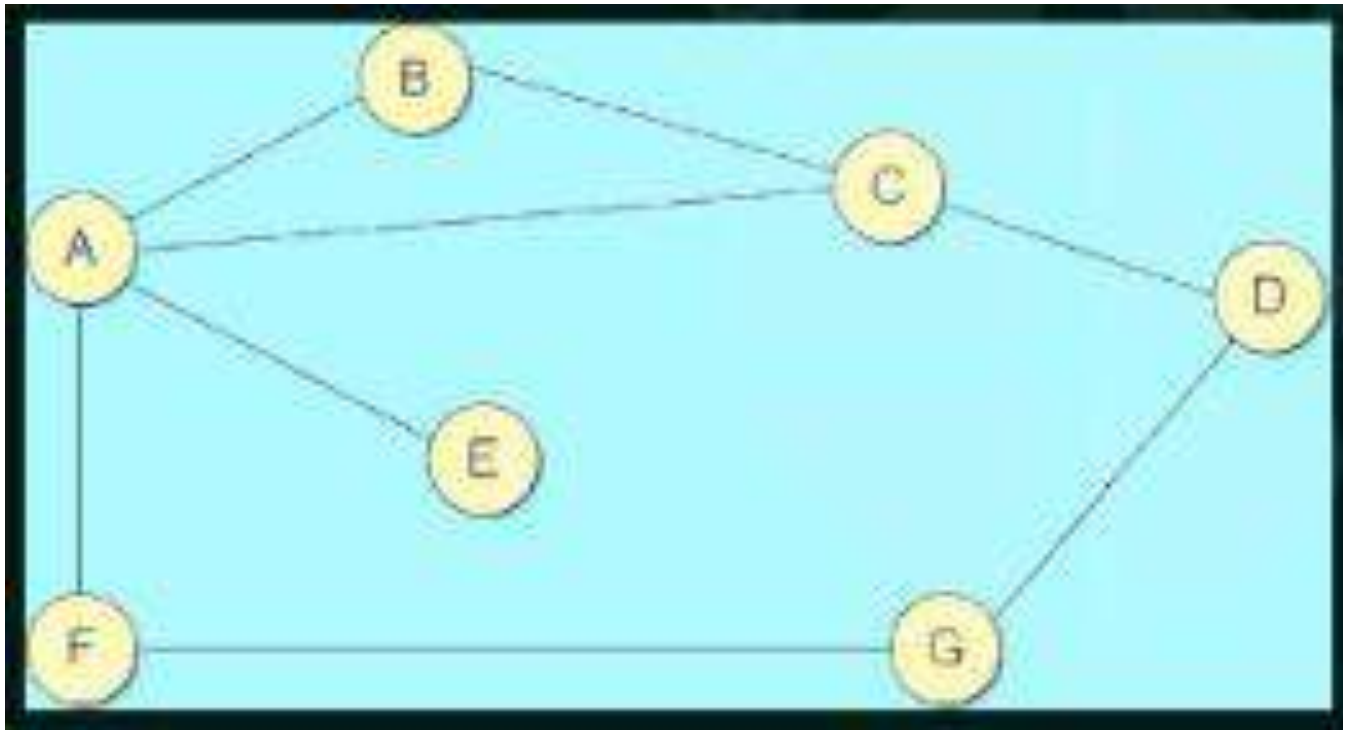
Exchange of Routing table at A from directly connected Neighbours

A

A	0
B	1
C	1
D	∞
E	1
F	1
G	∞

B

A	1
B	0
C	1
D	∞
E	∞
F	∞
G	∞

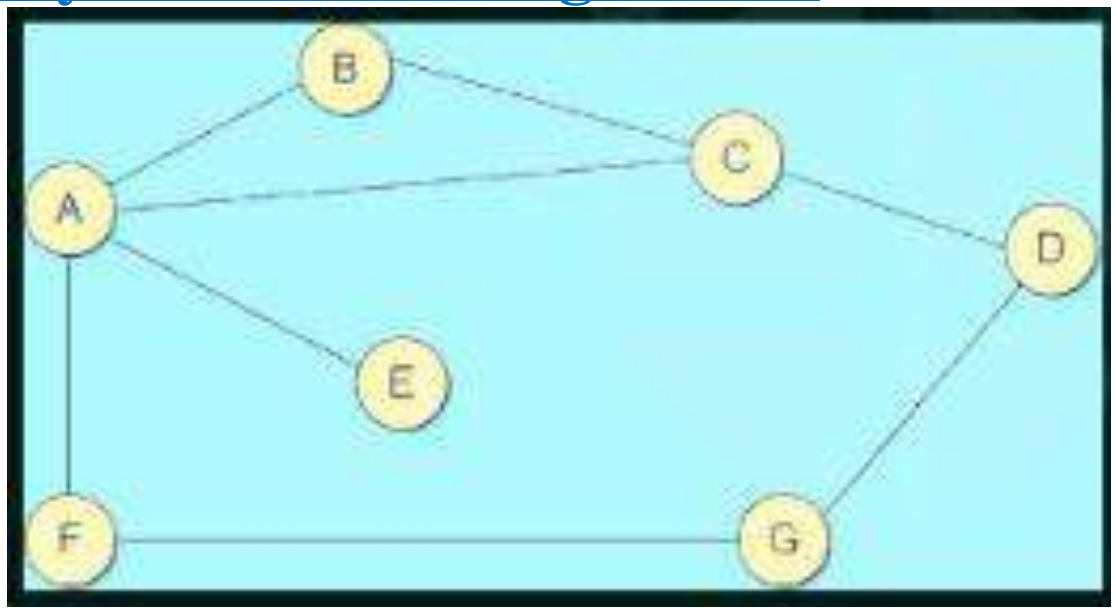


Exchange of Routing table at A from directly connected Neighbours

C	
A	1
B	1
C	0
D	1
E	∞
F	∞
G	∞

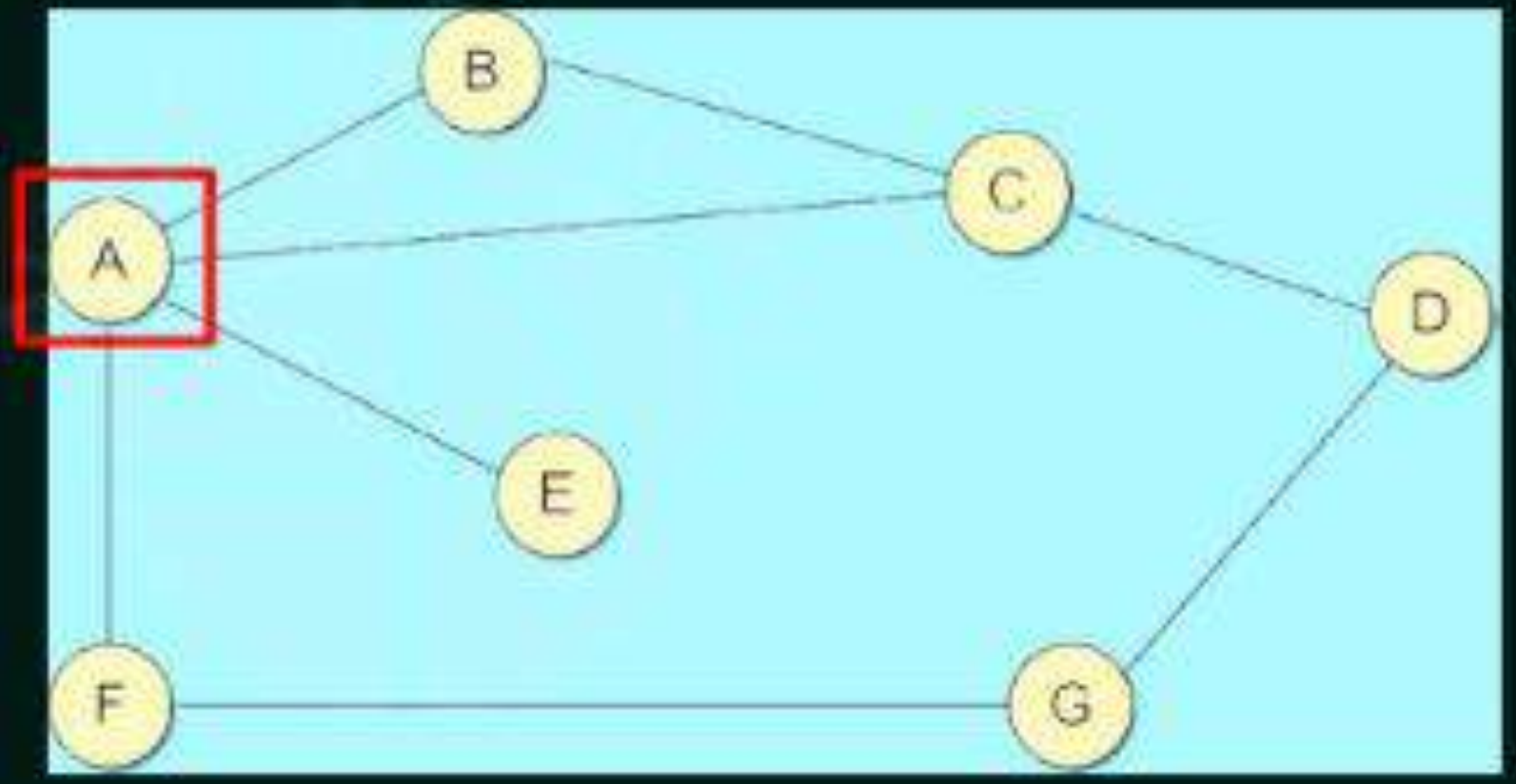
E	
A	1
B	∞
C	∞
D	∞
E	0
F	∞
G	∞

F	
A	1
B	∞
C	∞
D	∞
E	∞
F	0
G	1



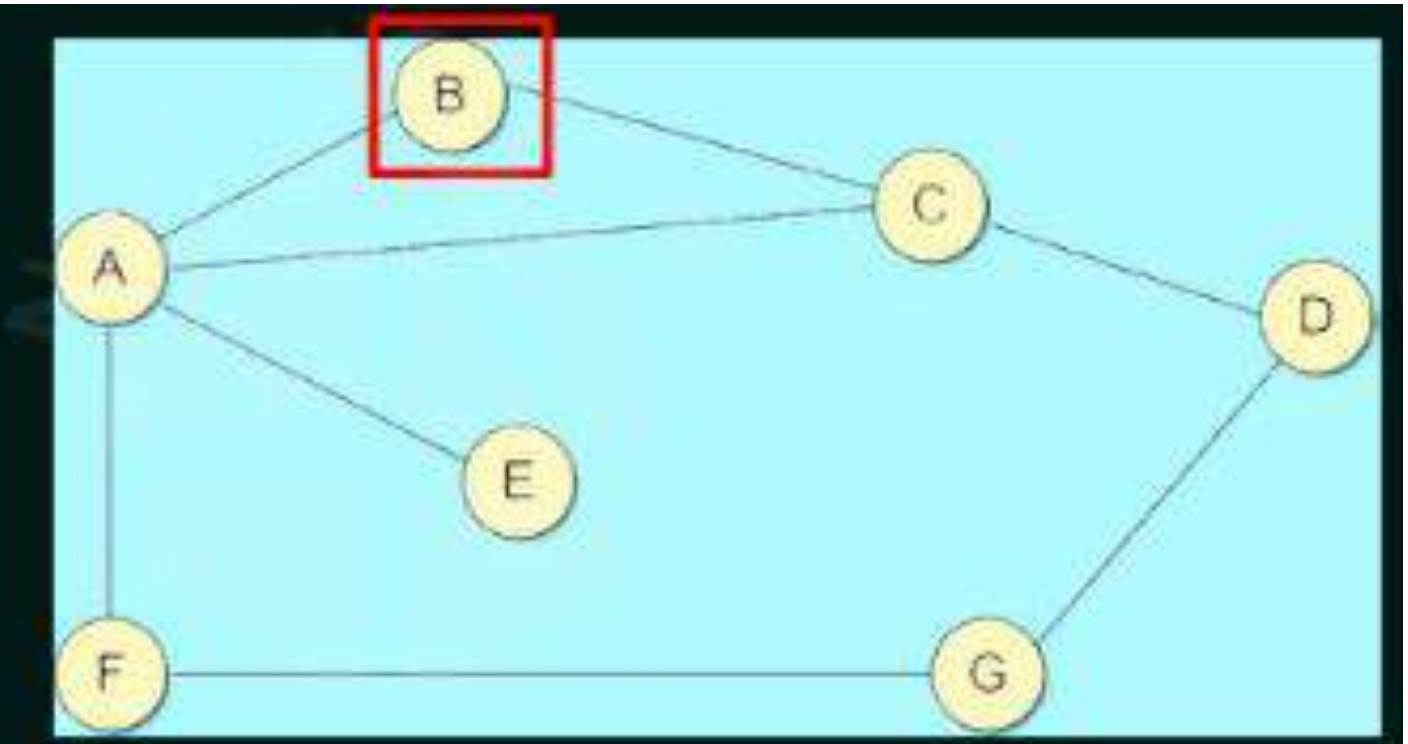
Final routing Table at A

Destination	Cost	Next Hop
A	0	-
B	1	B
C	1	C
D	2	C
E	1	E
F	1	F
G	2	F



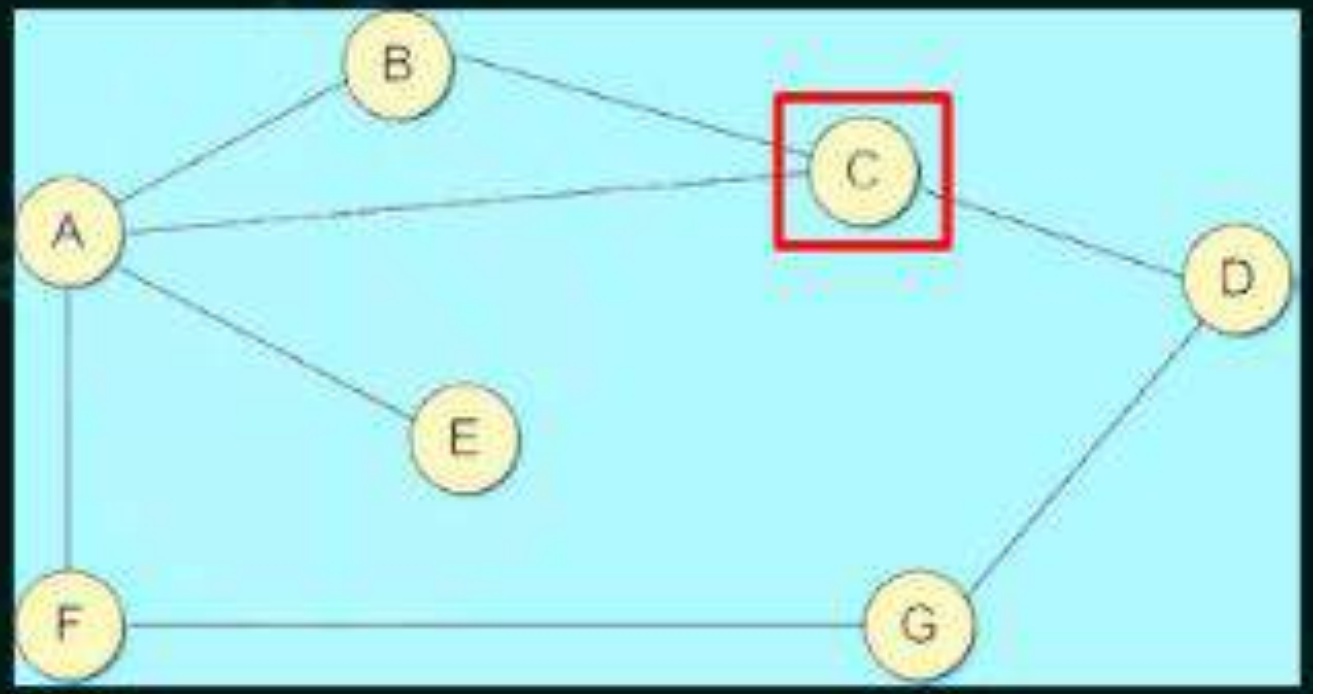
Final routing Table at B

Destination	Cost	Next Hop
A	1	A
B	0	–
C	1	C
D	2	C
E	2	A
F	2	F
G	3	F



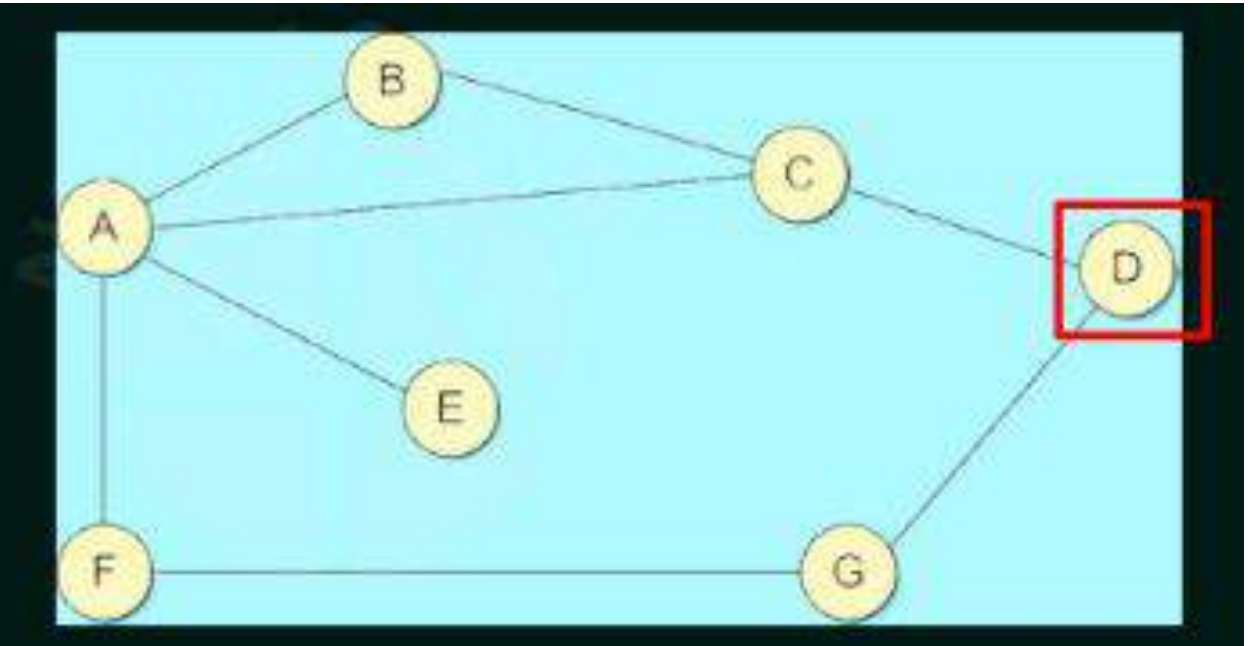
Final routing Table at C

Destination	Cost	Next Hop
A	1	A
B	1	B
C	0	-
D	1	D
E	2	A
F	2	A
G	2	D



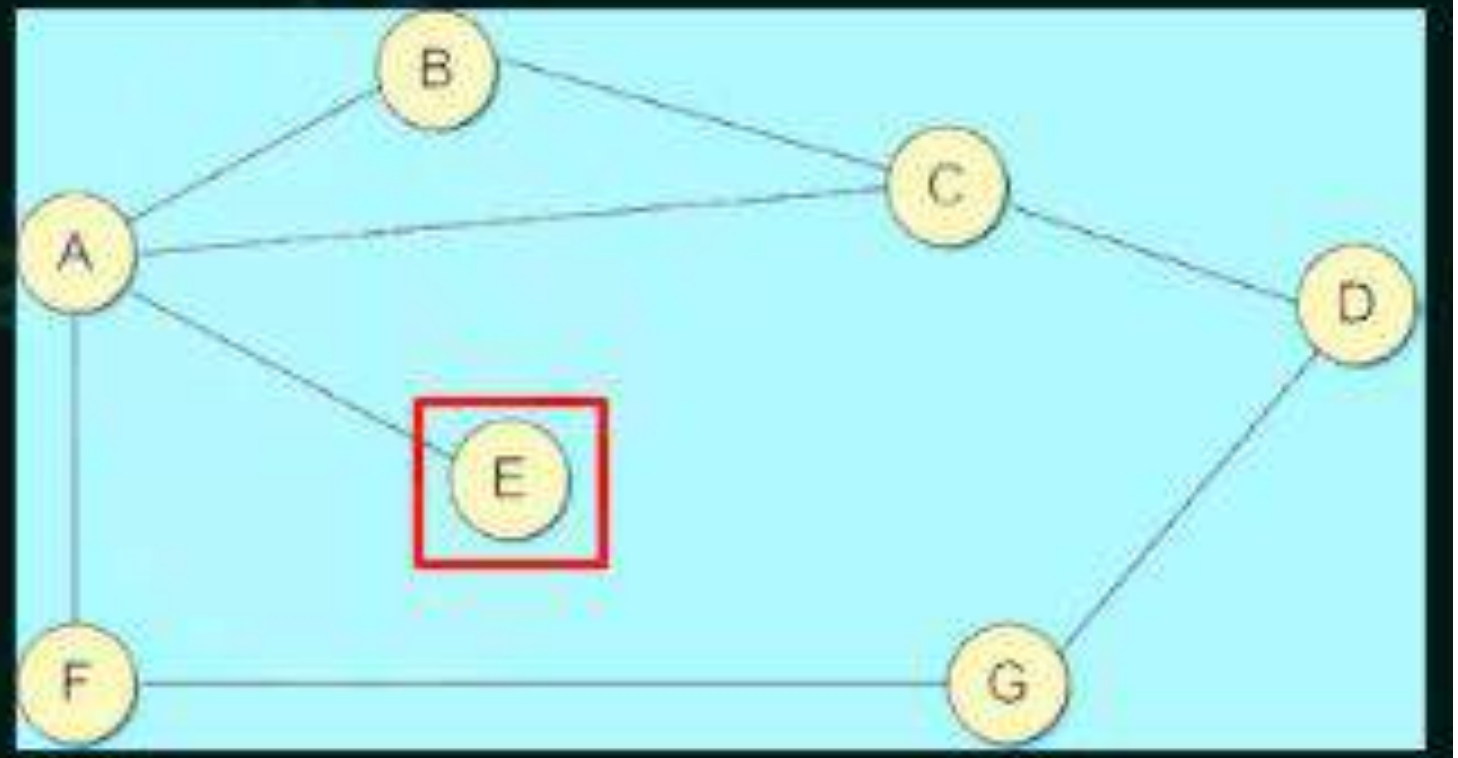
Final routing Table at D

Destination	Cost	Next Hop
A	2	C
B	2	C
C	1	C
D	0	-
E	3	C
F	2	G
G	1	G



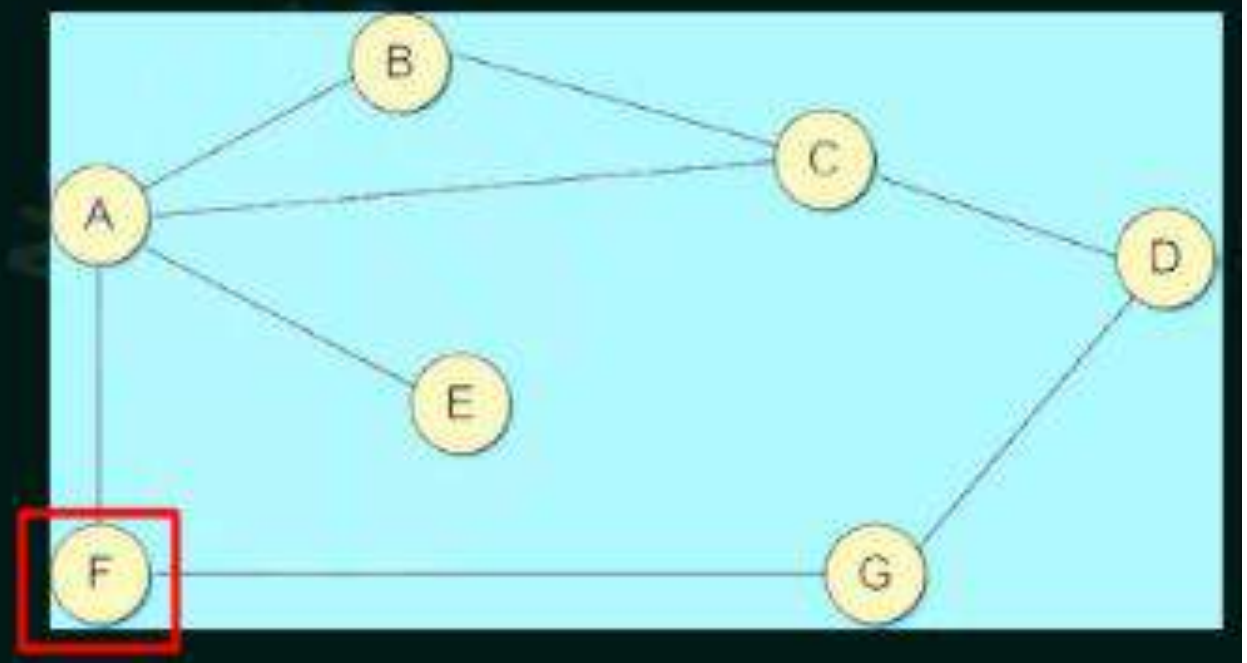
Final routing Table at E

Destination	Cost	Next Hop
A	1	A
B	2	A
C	2	A
D	3	A
E	0	-
F	2	A
G	3	A



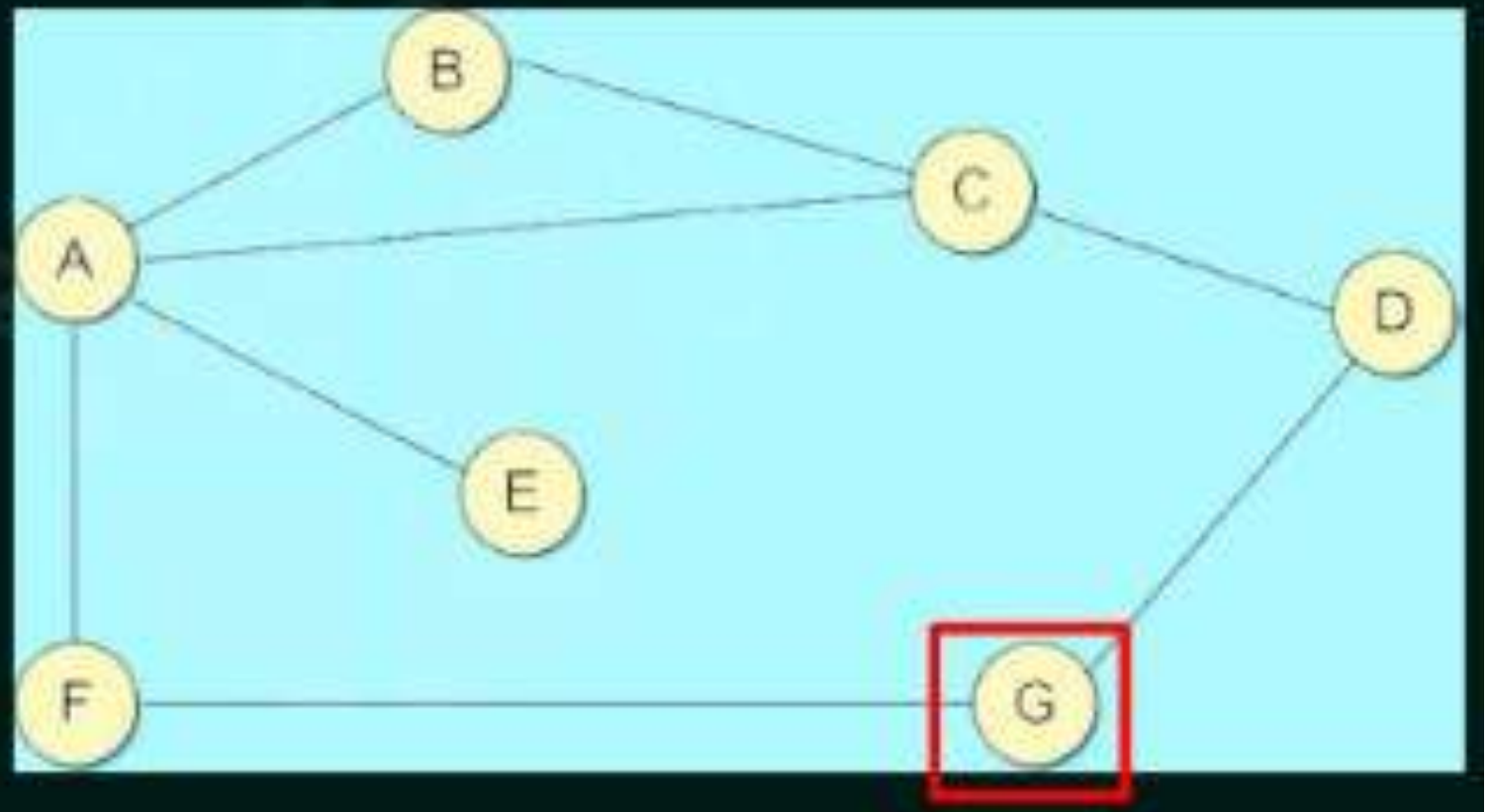
Final routing Table at F

Destination	Cost	Next Hop
A	1	A
B	2	A
C	2	A
D	2	G
E	2	A
F	0	-
G	1	G



Final routing Table at G

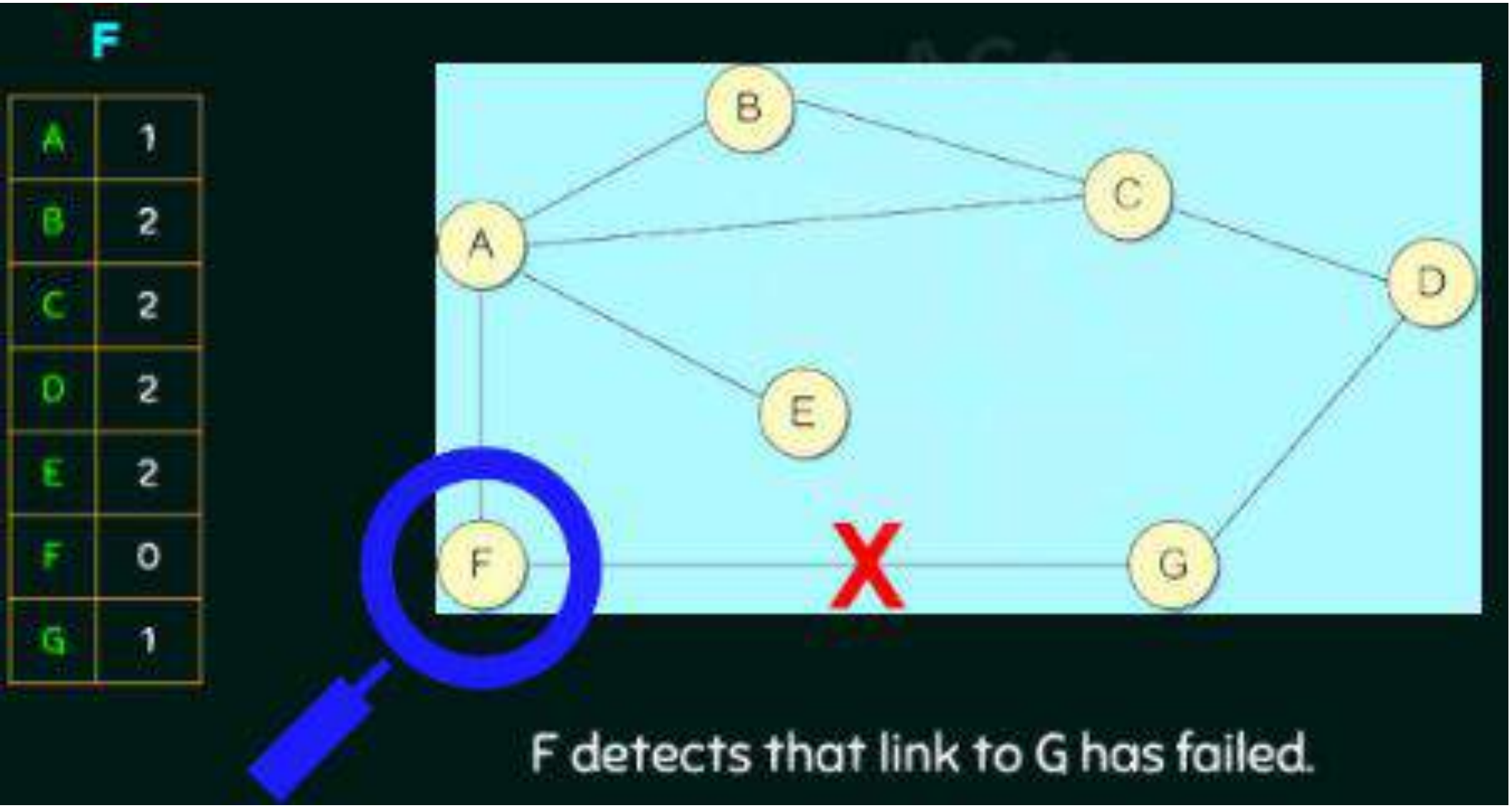
Destination	Cost	Next Hop
A	2	F
B	3	F
C	2	D
D	1	1
E	3	F
F	1	F
G	0	-

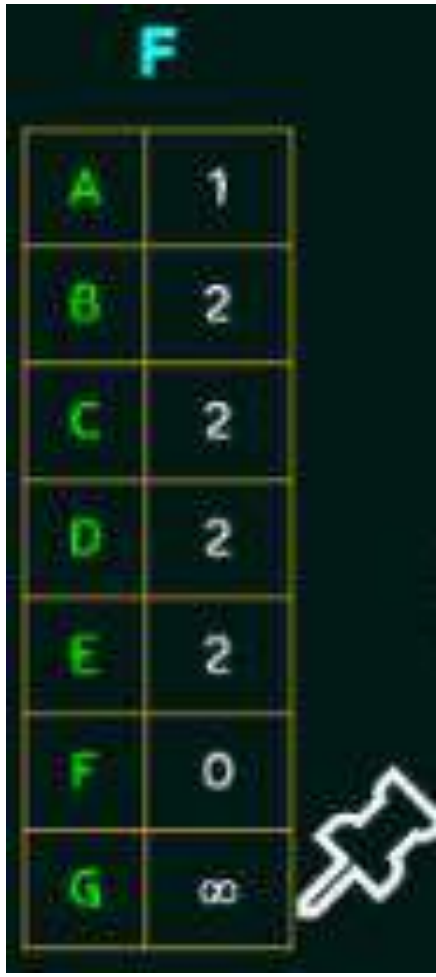


Final Distance stored at Each Node

Information stored at Node	Distance to reach node						
	A	B	C	D	E	F	G
A	0	1	1	2	1	1	2
B	1	0	1	2	2	2	3
C	1	1	0	1	2	2	2
D	2	2	1	0	3	2	1
E	1	2	2	3	0	2	3
F	1	2	2	2	2	0	1
G	2	3	2	1	3	1	0

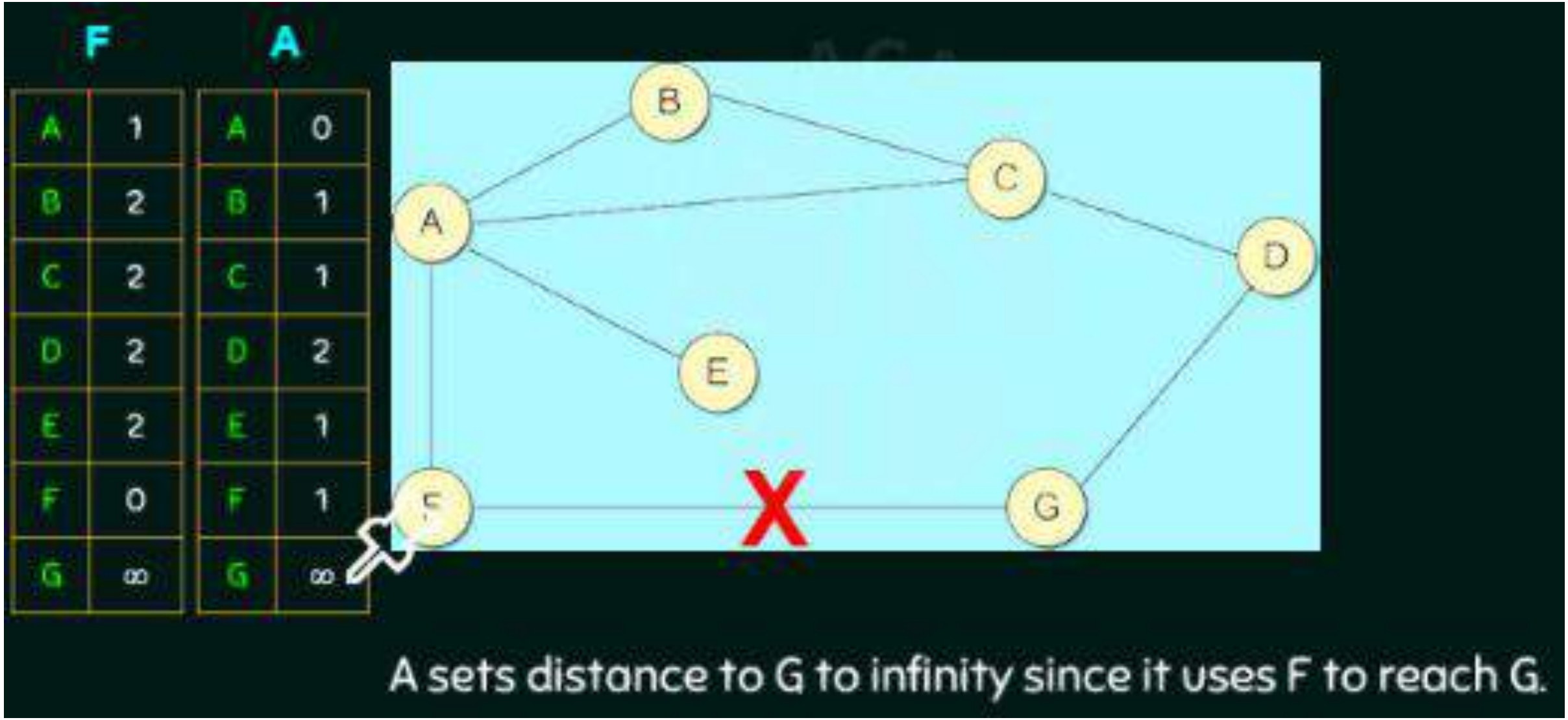
When a Node Detects Link Failure





F sets distance to G to infinity and sends update to A.

When a Node Detects Link Failure

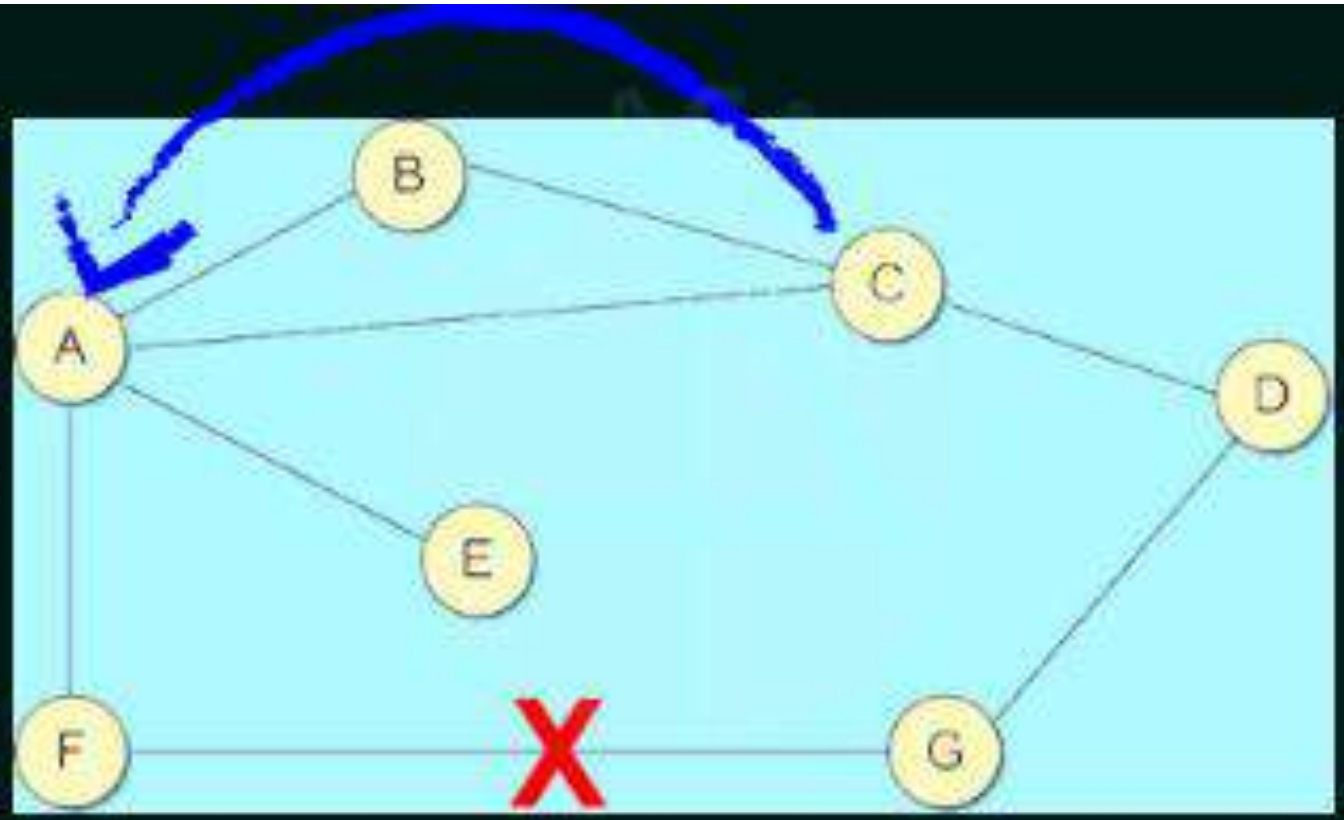


F

A	1
B	2
C	2
D	2
E	2
F	0
G	∞

A

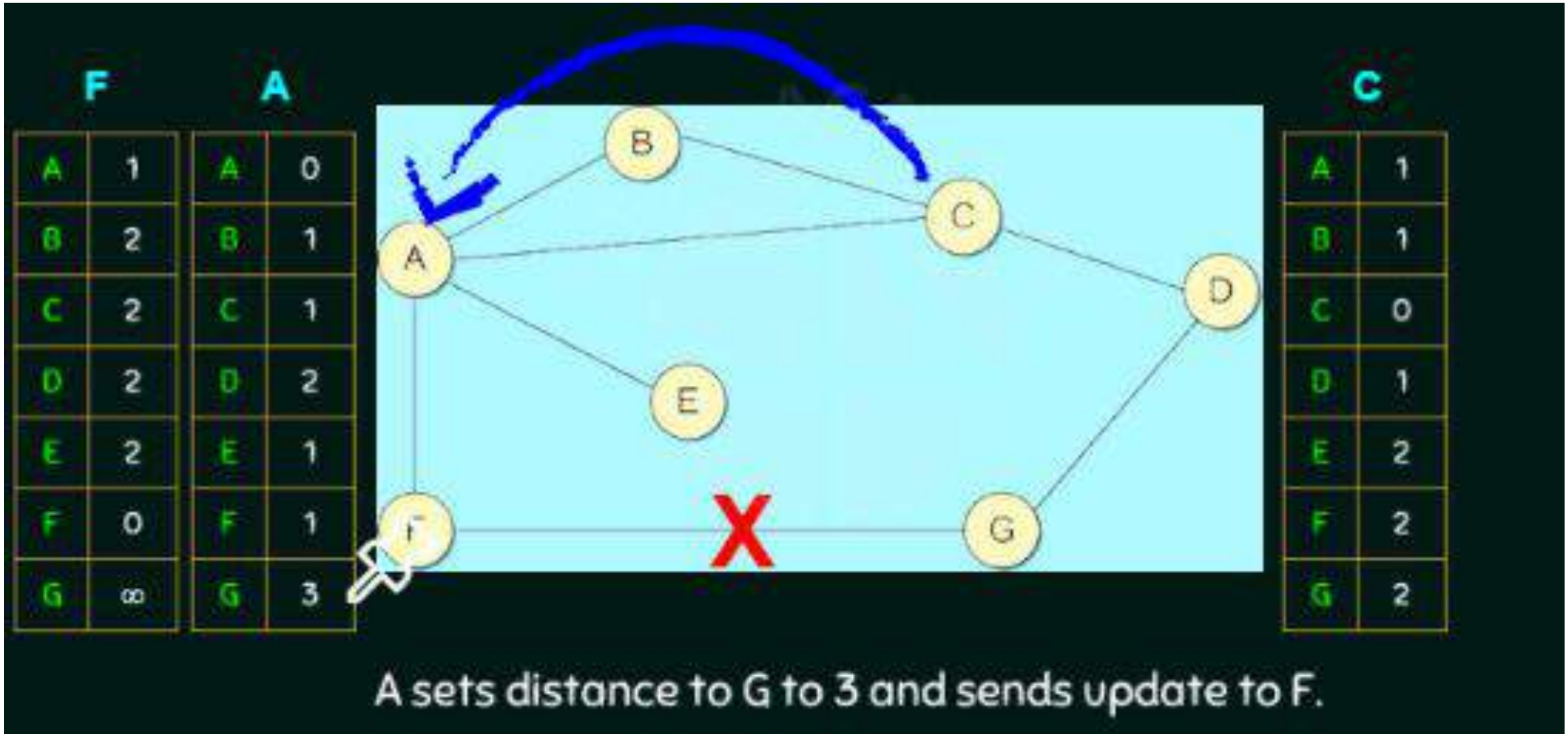
A	0
B	1
C	1
D	2
E	1
F	1
G	∞

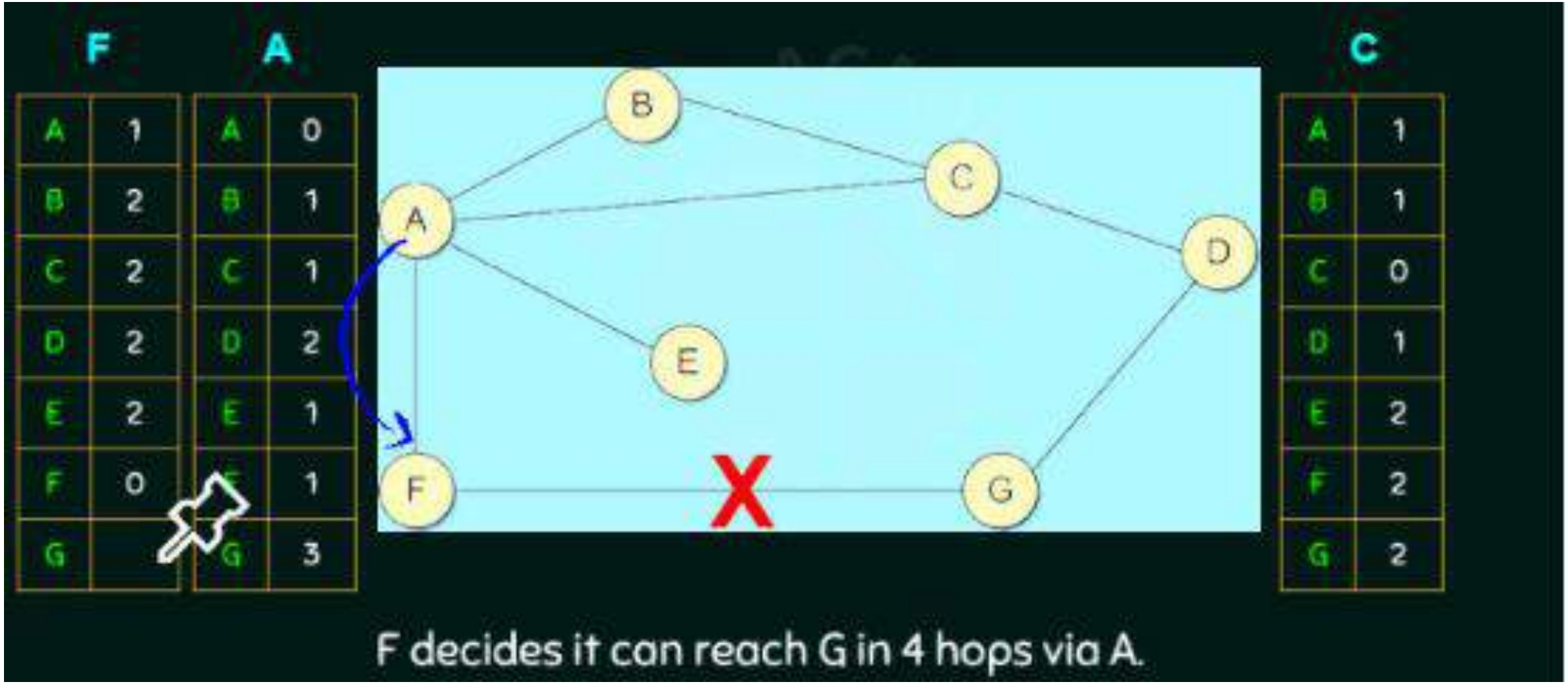


C

A	1
B	1
C	0
D	1
E	2
F	2
G	2

A receives periodic update from C with 2-hop path to G.


F
A
C
A 1
B 2
C 2
D 2
E 2
F 0
G ∞
A 0
B 1
C 1
D 2
E 1
F 1
G 3
A 1
B 1
C 0
D 1
E 2
F 2
G 2
A
B
C
D
E
F
G
X
A sets distance to G to 3 and sends update to F.

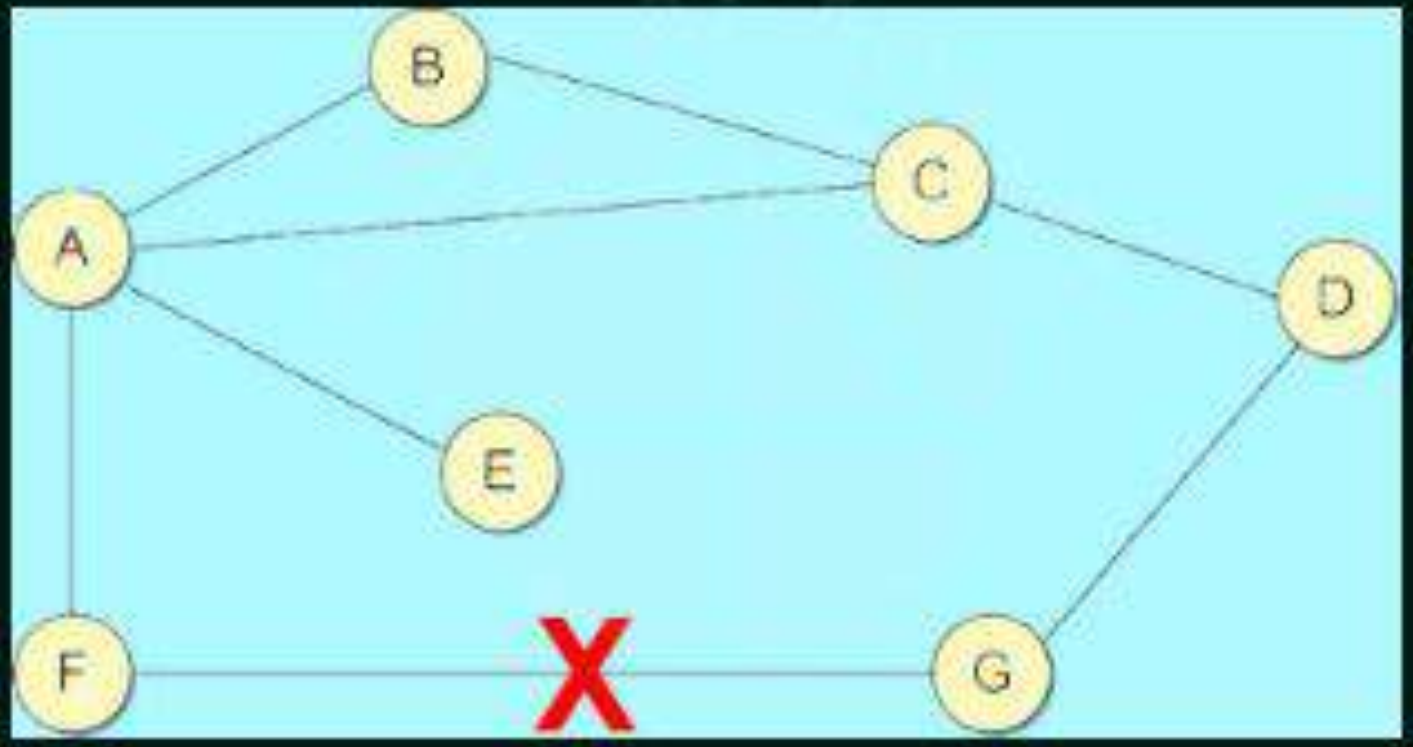

F
A
C
A 1
B 2
C 2
D 2
E 2
F 0
G
A 0
B 1
C 1
D 2
E 1
F 1
G 3
A 1
B 1
C 0
D 1
E 2
F 2
G 2
A
B
C
D
E
F
G
X
F decides it can reach G in 4 hops via A.

F

A	1
B	2
C	2
D	2
E	2
F	0
G	4

A

A	0
B	1
C	1
D	2
E	1
F	1
G	3



C

A	1
B	1
C	0
D	1
E	2
F	2
G	2

F decides it can reach G in 4 hops via A.

Count to Infinity problem

A's Routing Table

Dest	Cost	Next Hop
B	1	B
C	3	B

B's Routing Table

Dest	Cost	Next Hop
A	1	A
C	2	C

C's Routing Table

Dest	Cost	Next Hop
A	3	B
B	2	B

Count to Infinity problem

A's Routing Table

Dest	Cost	Next Hop
B	1	B
C	3	B

B's Routing Table

Dest	Cost	Next Hop
A	1	A
C	2	C

C's Routing Table

Dest	Cost	Next Hop
A		
B		

Count to Infinity problem

A's Routing Table

Dest	Cost	Next Hop
B	1	B
C	3	B

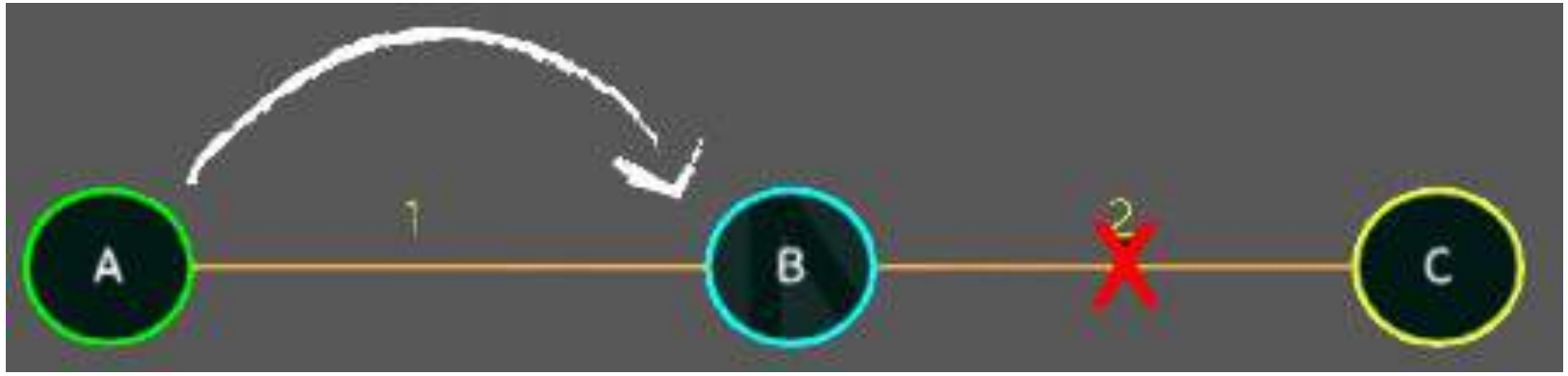
B's Routing Table

Dest	Cost	Next Hop
A	1	A
C		

C's Routing Table

Dest	Cost	Next Hop
A		
B		

Count to Infinity problem



A's Routing Table

Dest	Cost	Next Hop
B	1	B
C	3	B

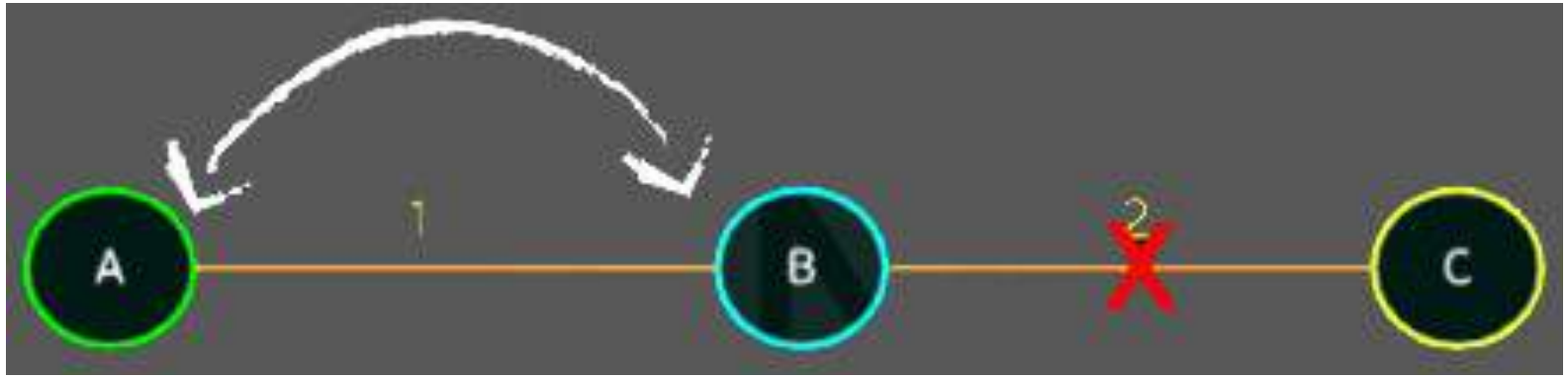
B's Routing Table

Dest	Cost	Next Hop
A	1	A
C	4	A

C's Routing Table

Dest	Cost	Next Hop
A		
B		

Count to Infinity problem



A's Routing Table		
Dest	Cost	Next Hop
B	1	B
C	15	B

B's Routing Table		
Dest	Cost	Next Hop
A	1	A
C	14	A

C's Routing Table		
Dest	Cost	Next Hop
A		
B		

Solution for Count to infinity

- **Split Horizon**
- **Split horizon with Poison Reverse**

Split Horizon

- *In computer network, split-horizon route advertisement is a method of preventing routing loops in distance-vector routing protocols by prohibiting a router from advertising a route back onto the interface from which it was learned.*

- *In other words, it is a method of preventing a routing loop in a network.*

- *The basic principle is simple, information about the routing for a particular packet is never sent back to the direction from which it was received.*

Solution for Count to infinity

Split Horizon

- *With the split-horizon rule in place , this particular loop scenario can not happen, improving convergence time in complex, highly redundant environments.*

Poison Reverse

- *Poisson reverse is an implemented algorithm that is often used within distance vector routing.*

- *The use of poison reverse is to solve the count-to-infinity problem. Practically, poison reverse can be thought of as the reverse of split horizon.*

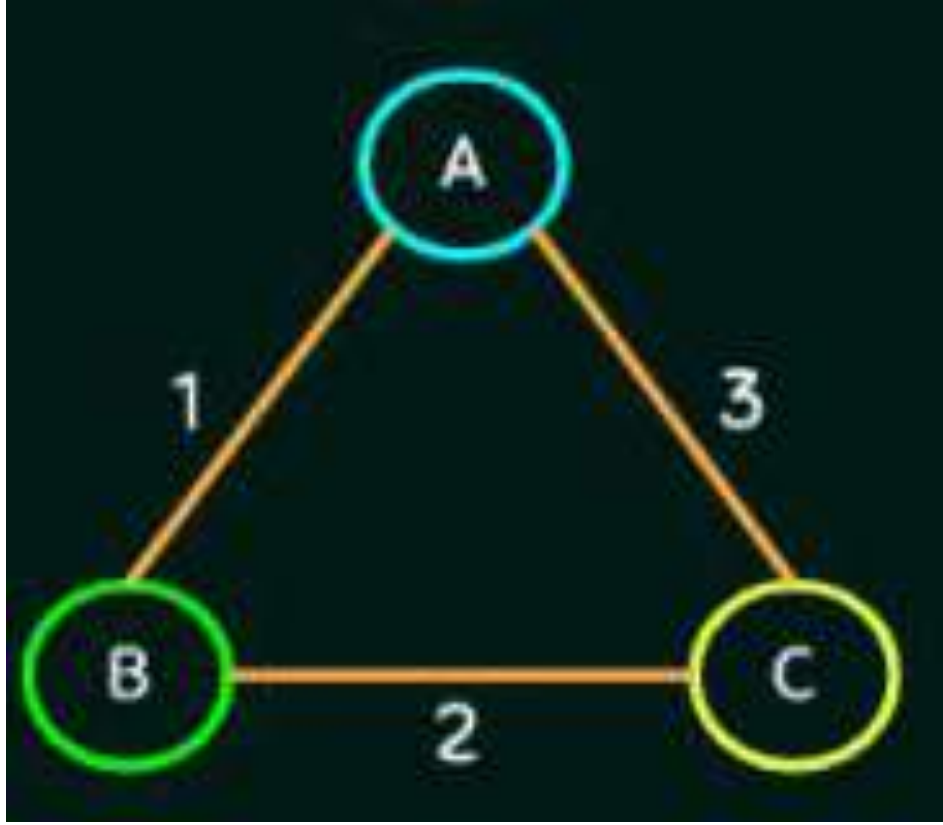
- *With poison reverse, route advertisements that would be supressed by split horizon are instead advertised with a distance of infinity.*

Poison Reverse

- *As A route via B to C and because of that have the cost 3. The poison reverse kicks in when we broadcast our distance vector to our neighbours. The distance table we broadcast is :*

  To B :[0,1,∞]

  To C: [0,1,3]



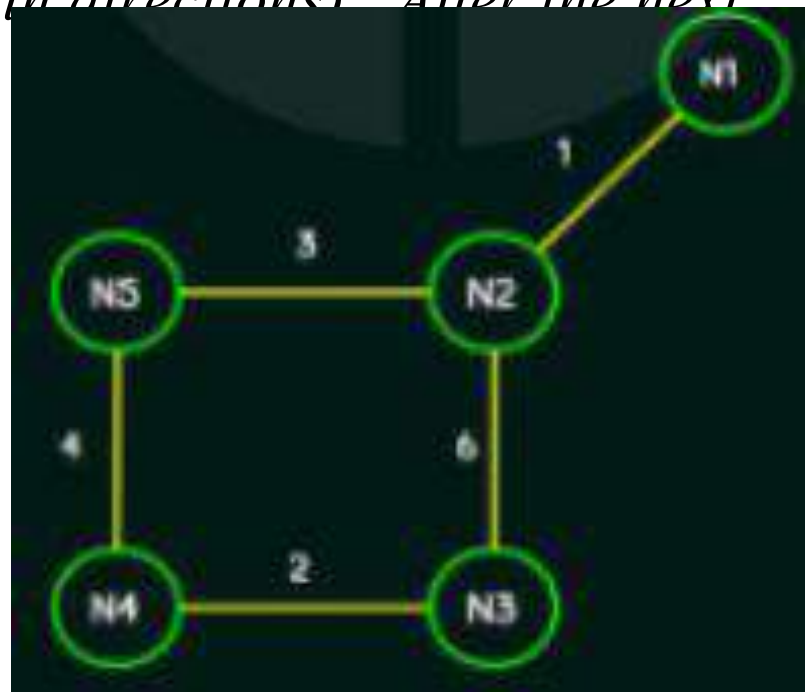
- *As we see in the distance vector that is broadcast to node B the end destination C has an infinity value. This solves the count-to-infinity problem since if the link between B and A will not bounce between each other and instead directly try another path.*

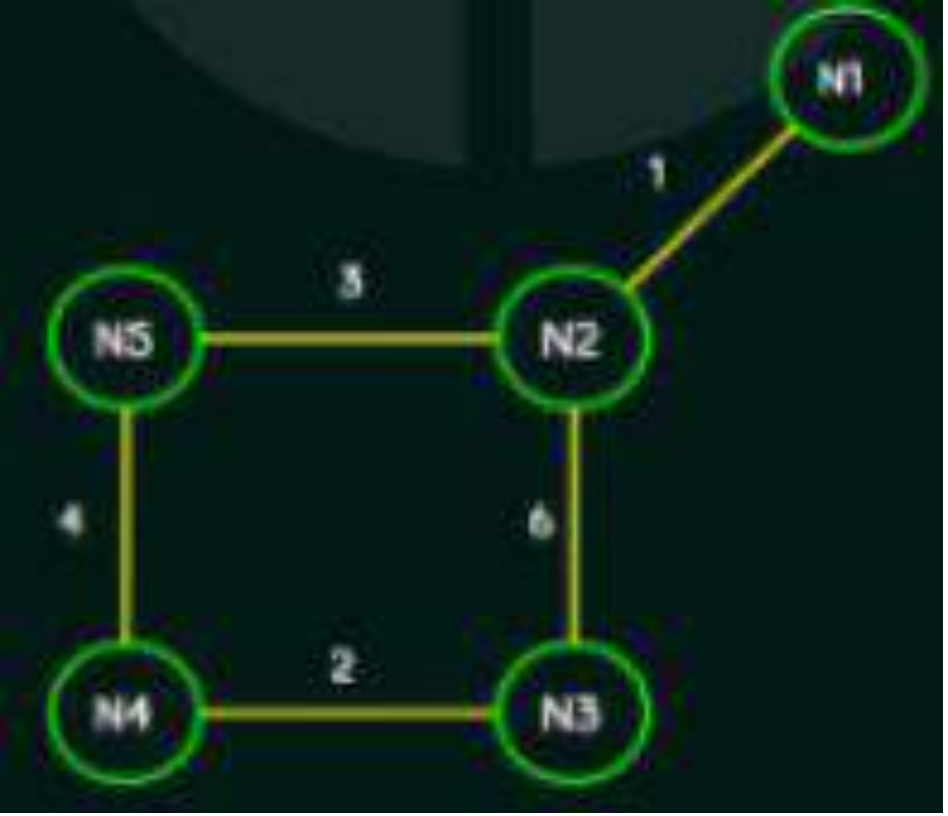
Distance	A	B	C
A	0	1	3
B	1	0	2
C	3	2	0

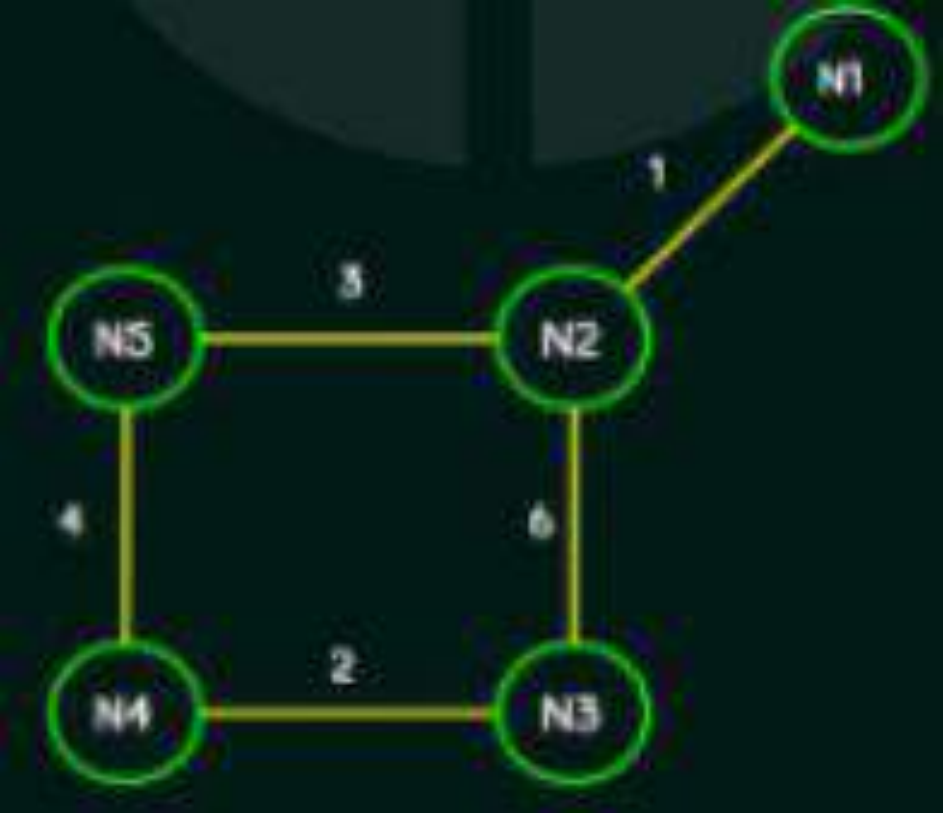
Questions:

Consider a network with five nodes N1 to N5, as shown below. The network uses a Distance Vector Routing protocol. Once the routes have stabilized, the distance vector at different nodes are as following: **N1: (0,1,7,8,4) N2: (1,0,6,7,3) N3:(7,6,0,2,6) N4:(8,7,2,0,4) N5:(4,3,6,4,0)**

Each distance vector is the distance of the best known path at that instance to nodes N1 to N5, where the distance to itself is 0. Also, all links are symmetric and the cost is identical in both directions. In each round, all node exchanges their distance vectors. In between two rounds, any change in cost of all links will cause the two incident nodes to change only that entry in their distance vectors. The cost of links N2-N3 reduces to 2 (in both directions). After the next round of update what will be the new distance vector at node N3?




N1
1
3
N5
N2
4
6
2
N4
N3


N1
1
3
N5
N2
4
6
2
N4
N3

RIP Messages

RIP defines two type of messages

- **Request Message:** *Asking a neighbouring RIPv1 enabled router to sends its routing table.*
- **Response Message:** *Carries the routing table of a router.*

RIP Version:

- **RIPv1**
- **RIPv2**
- **RIPng**

RIPv1	RIPv2	RIPng
When starting up, and every 30secs thereafter, a router with RIPv1 implementation broadcast to 255.255.255.255 a request message through every RIPv1 enabled interface. RIPv1 enabled router not only request the routing tables of other routers every 30 secs, they also listen to incoming requests from the neighbouring routers and send their own routing table in turns.	It includes the ability to carry subnet information thus supporting classless Inter-Domain Routing (CIDR). In an effort to avoid unnecessary load on hosts that do not participate in routing, RIPv2 multicast the entire routing table to all adjacent routers at the address 224.0.0.9, as opposed to RIPv1 which uses broadcast.	ng= Next Generation It is an extension of RIPv2 for support of IPV6, the next generation internet protocol. The main difference between RIPv2 and RIPng is: ▪ Support of IPV6 networking. ▪ While RIPv2 supports authentication, RIPng doe not. IPV6 routers where at time, supposed to use IPsec for authentication

RIPv1	RIPv2	RIPng
RIPv1 uses classful routing. The periodic routing updates do not carry subnet information, lacking support for variable length subnet mask (VLSM). This limitation makes it impossible to have different-sized subnets inside of the same network class. There is also no support for routers authentication, making RIP vulnerable to various attacks.	Unicast addressing is still allowed for special application. In RIPv2, MD5 authentication was introduced. Route tags were also added in RIPv2. Tis functionality allows a distinction between routes learned from the RIP protocol and routes learned from other protocols.	RIPv3 encodes the next-hop into each route entry, RIPng requires specific encoding of the next hop for a set of route entries. RIPng sends updates an UDP port 521 using the multicast group FF02::9

Link state Routing

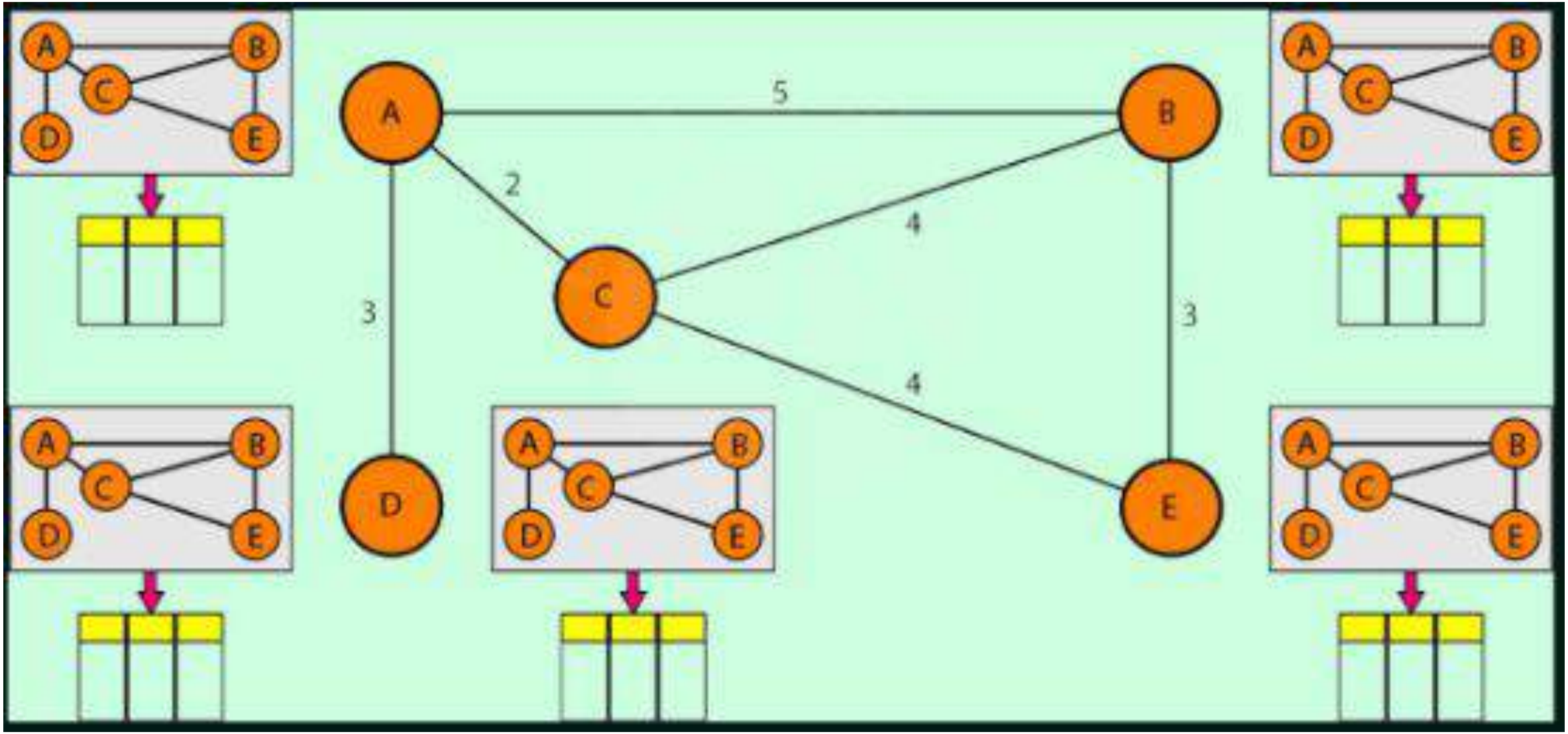
- **Link+State**
- **Link state routing (LSR) has a different philosophy from the DVR (Distance Vector Routing). In LSR, each node has the entire topology of the network.**

Link State	Distance Vector
❑ Complete topology.	❑ Local topology
❑ Flooding	❑ Exchange
❑ Global	❑ Decentralized
❑ “Tell the world about neighbours”	❑ “Tell the neighbours about the world”

Link state Routing

- A Link state *packet can carry a large amount of information*.
- **Link state packet contains:**
  - *ID of the node that created the LSP.*
  - *Cost of the link to each directly connected neighbour*
  - *Sequence number (SEQNO)*
  - *Time-to-Live (TTL) for this packet.*

Link state Routing



Link state Routing

- **Strategy:** *Send all nodes (not just neighbours) information about directly connected links (not entire routing table).*
- **Idea:** *LSR in two phases:*

- *Phase 1: Reliable flooding*
  - *Initial state: each node knows the cost of its neighbours*
  - *Final state: Each node knows the entire network topology*

- *Phase 2: Route calculation*
  - *Each node uses Dijkstra's algorithm on the graph to calculate optimal routes to all nodes.*

Link state Routing

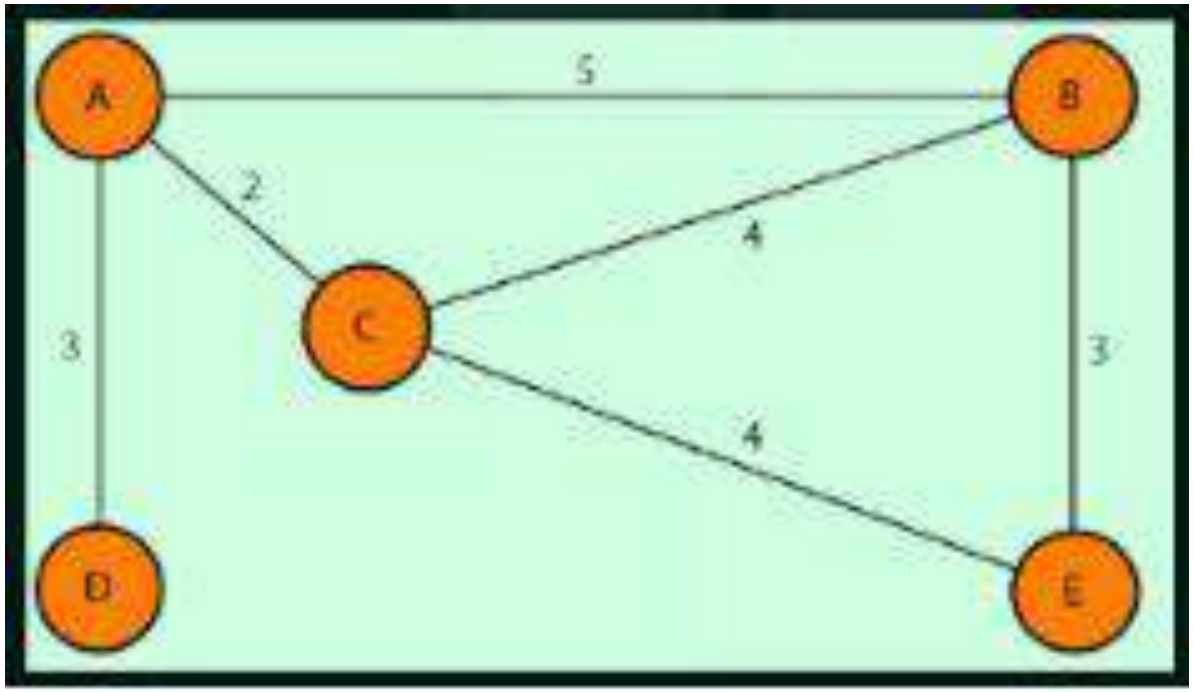
- **The Link state packet contains:**
  - *ID of the node that created the LSP.*
  - *Cost of the link to each directly connected neighbour*
  - *Sequence number (SEQNO)*
  - *Time-to-Live (TTL) for this packet.*

- **LSP are generated in two occasions:**
  - *Triggered*
  - *Periodic*

Link state Routing

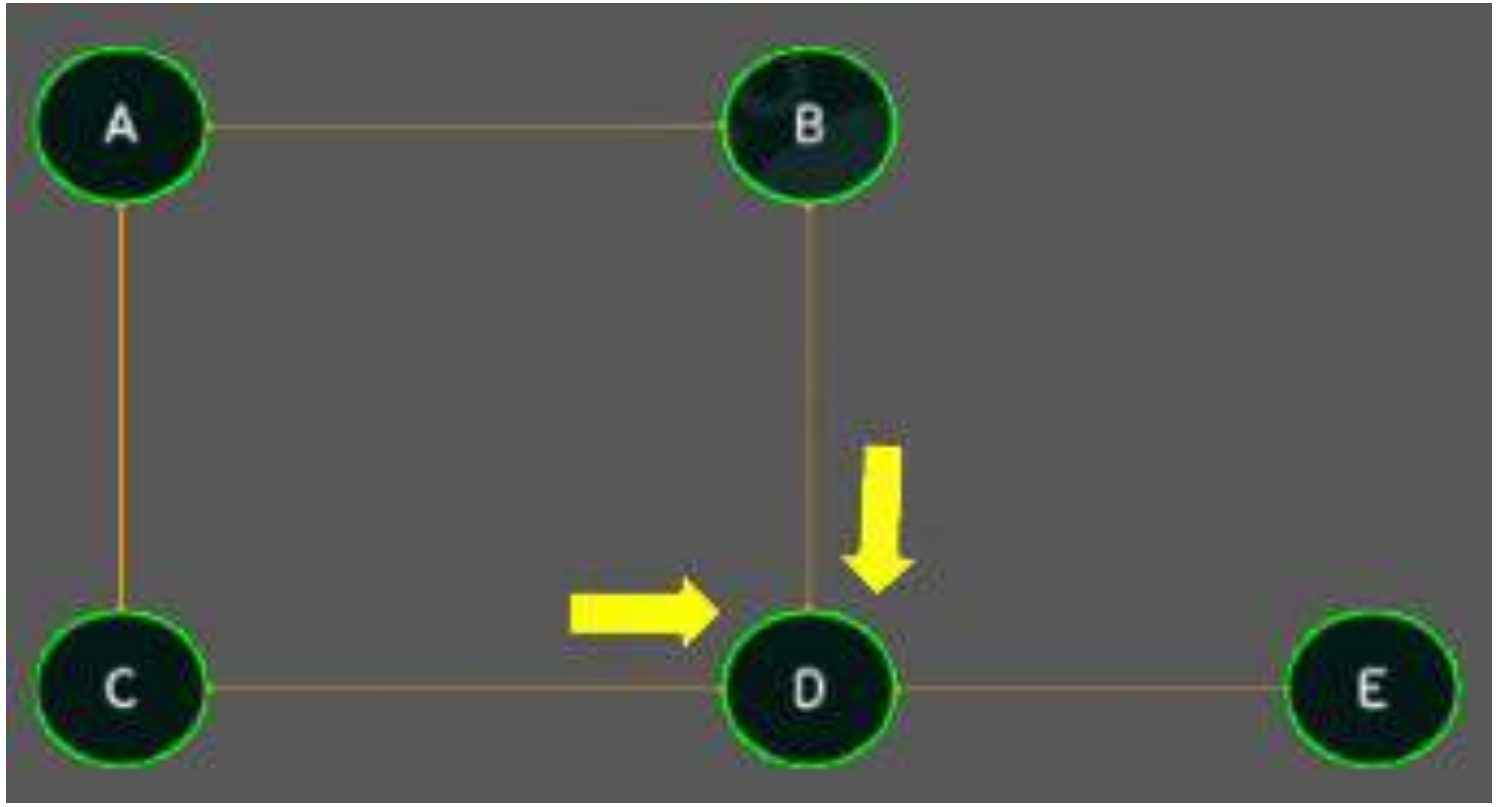
Phase 1: Reliable flooding

- *Store most recent LSP from each node.*
- *Forward LSP to all nodes but one that sent it.*
- *Generate new LSP periodically, increment SEQNO.*
- *Start SEQNO at o when reboot.*
- *Decrement TTL of each stored LSP, discard when TTL = 0.*



Reliable Flooding in Link state Routing

- *The node D will receive same LSP packet from both B and C but will discard the packet with old SEQNO number and retain the packet with new SEQNO.*



Features and Solution of LSP

Feature	Responsible field of the LSP
Minimise the number of messages and detect duplicated	ID field in the LSP.
Newer information should precede older info	Sequence number field in the LSP
Exhaust of sequence number	32 bits sequence number field in LSP
Looping of LSPs.	TTL field in LSP.

LSP vs DVR

Feature	Link State Routing	Distance Vector routing
Topology	Complete topology	Local topology
Communication type	Flooding	Exchange
Convergence	Fast	Slow
Routing Loops	Less Prone	Highly Prone
Algorithm	Dijkstra’s Algorithm	Bellman Ford Algorithm
Metrics	Cost	Hop count
Updates	Triggered, Partial	Frequent, Periodic
Scalability	Highly scalable	Limited
Protocols	OSPF, IS-IS	RIP, IGRP

Shortest Path Routing

- *In practice each router computes its routing table directly from LSP's it has collected using a realization of Dijkstra's algorithm.*

- *Specially each router maintains two lists:*

1. Tentative list

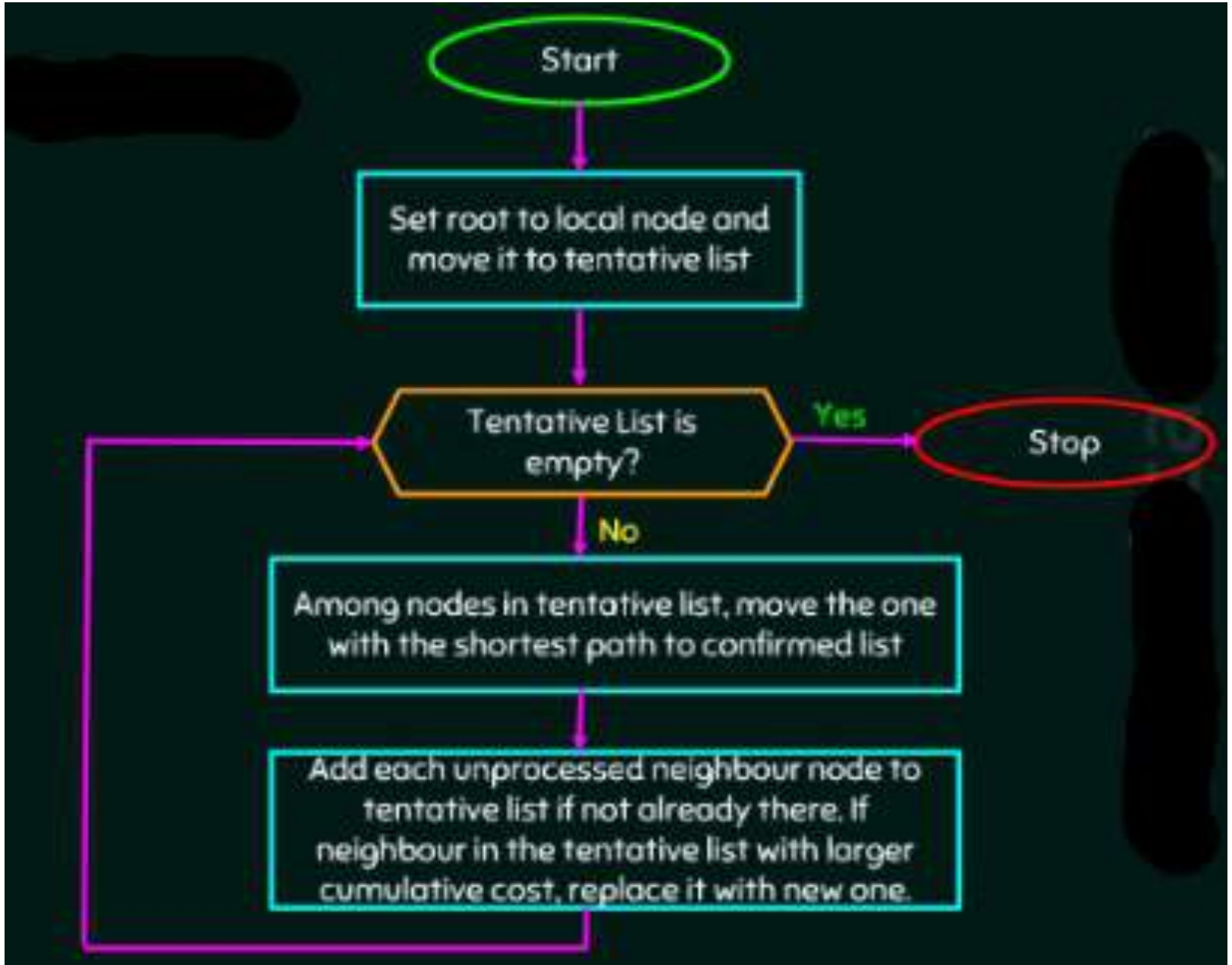
2. Confirmed List

- *Each of these lists contains a set of entries of the form*

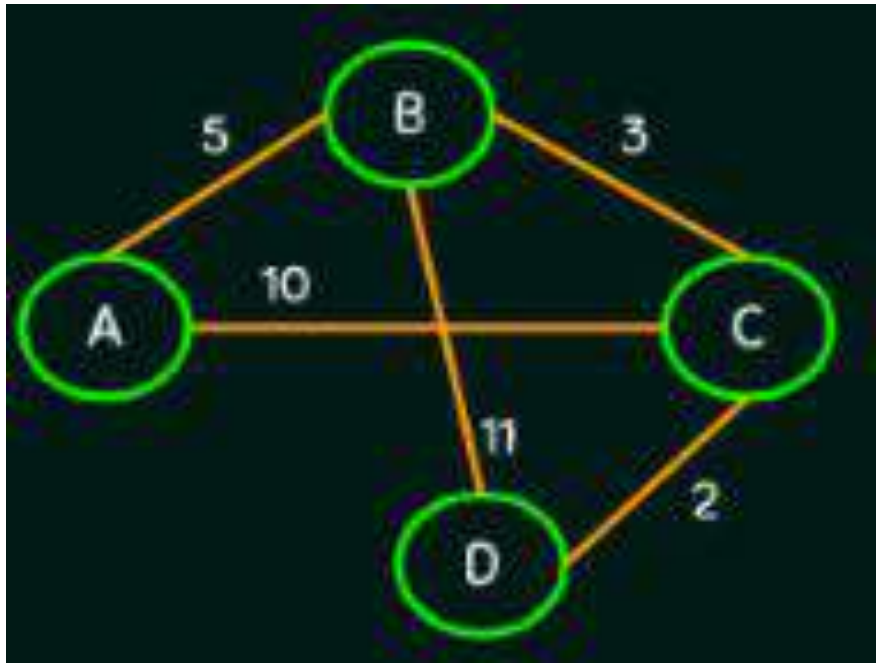
(Destination, Cost, Next hop)

E.g. (A, 5, C)

Dijkstra's Algorithm



Question: Find the shortest path routing for the following graph.



Open Shortest Path First (OSPF)

- *OSPF is a routing protocol for IP networks with used link state routing algorithm.*

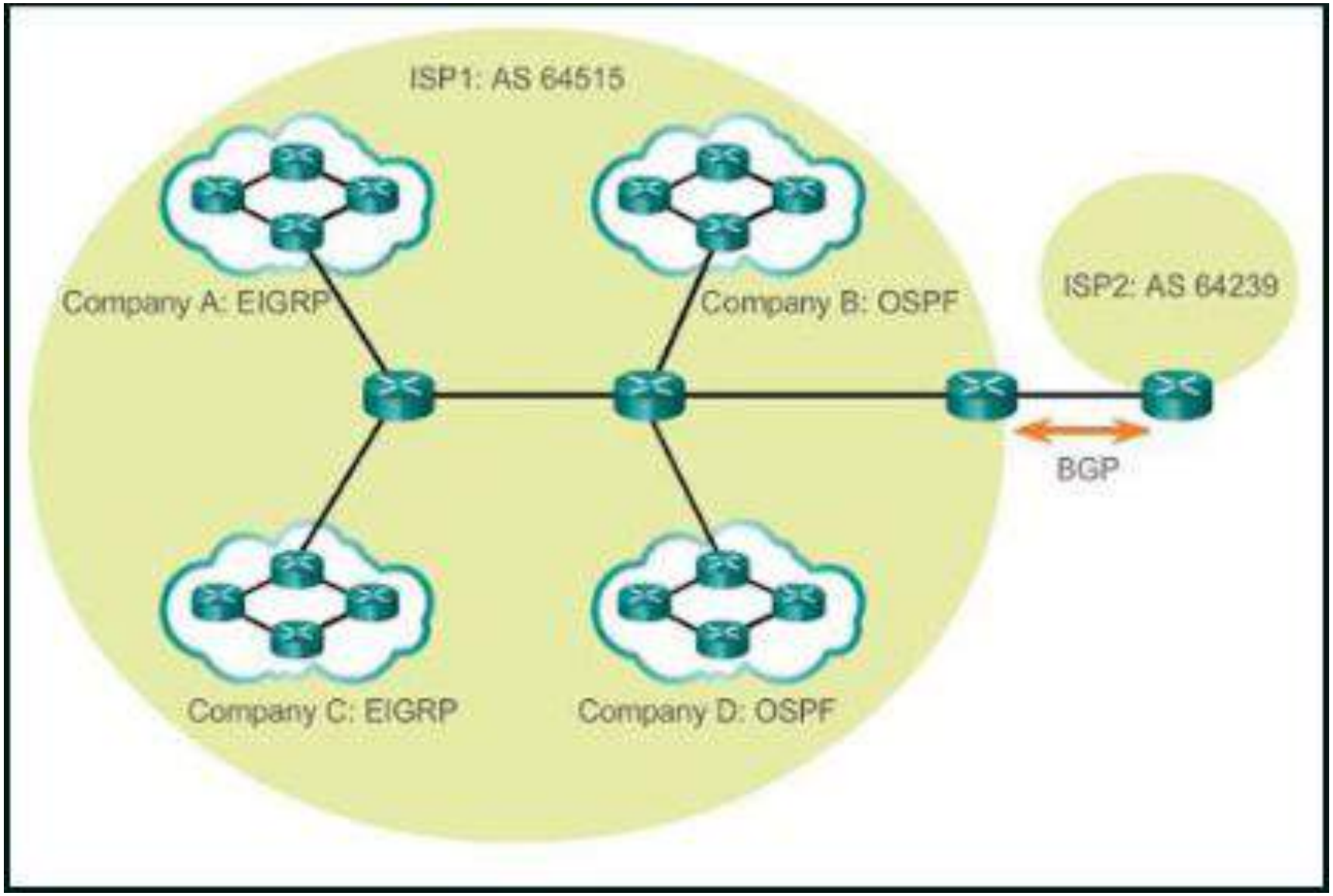
- *OSPF falls into the group of IGPs (Interior Gateway protocols, operating within a single Autonomous system (AS).*

- *RIP uses hop count as a metric whereas OSPF uses cost as the metric.*

- *OSPF was developed so that the shortest path through a network was calculated based on the cost of the route.*

- *The link cost is calculated by taking bandwidth, delay and load into account.*

Autonomous System



Open Shortest Path First (OSPF)

- *OSPF undertakes route cost calculation on the basis of link-cost parameters, which can be weighted by the administrator.*

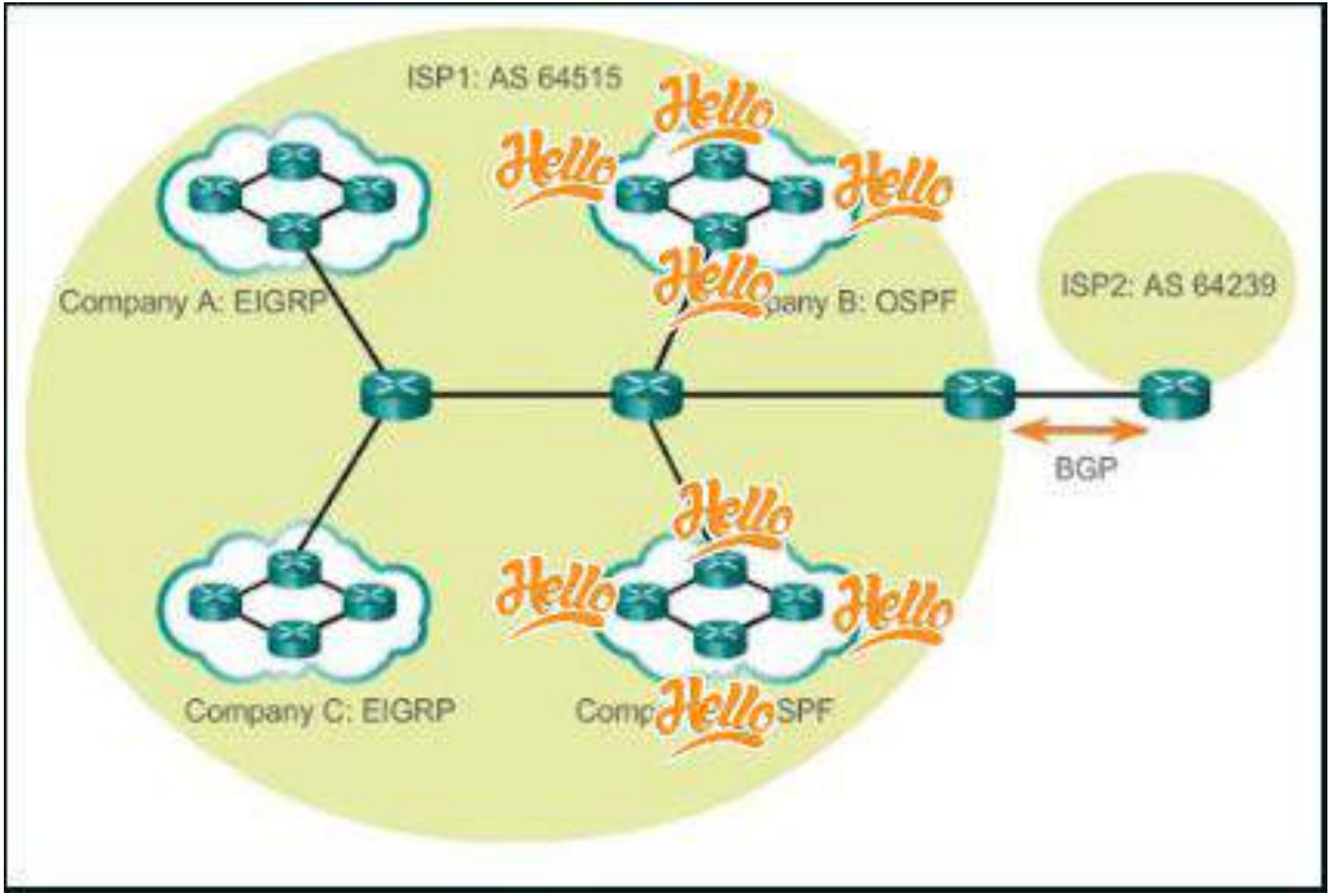
- *OSPF is reliable even for large and complex networks.*

- *As a link state routing OSPF maintains link-state databases. The state of a given route in the network is the cost, and OSPF algorithm allows every router to calculate the cost of the routes to any given reachable destination.*

- *An router interface with OSPF will then advertise its link cost to neighbouring routers through multicast, known as the HELLO procedure.*

- *The link cost is calculated by taking bandwidth, delay and load into account. All routers with OSPF implementation keep sending HELLO packets, and thus changes in the cost of their links become known to neighbouring outers.*

OSPF HELLO procedure



Open Shortest Path First (OSPF)

- *An OSPF network can be structured, or subdivided, into routing areas t simplify administration and optimize traffic and resource utilization.*

- *Area are identified by 32-bits numbers, expressed either simply in decimal, o often in the same dot-decimal notation used for IPV4 addresses.*

- *By convention, area 0 (zero), or 0.0.0.0 represents the core or backbone are of an OSPF network.*

- *While the identification of other areas may be chosen at will, administratio often select the IP address of a main router in an area as the area identifier.*

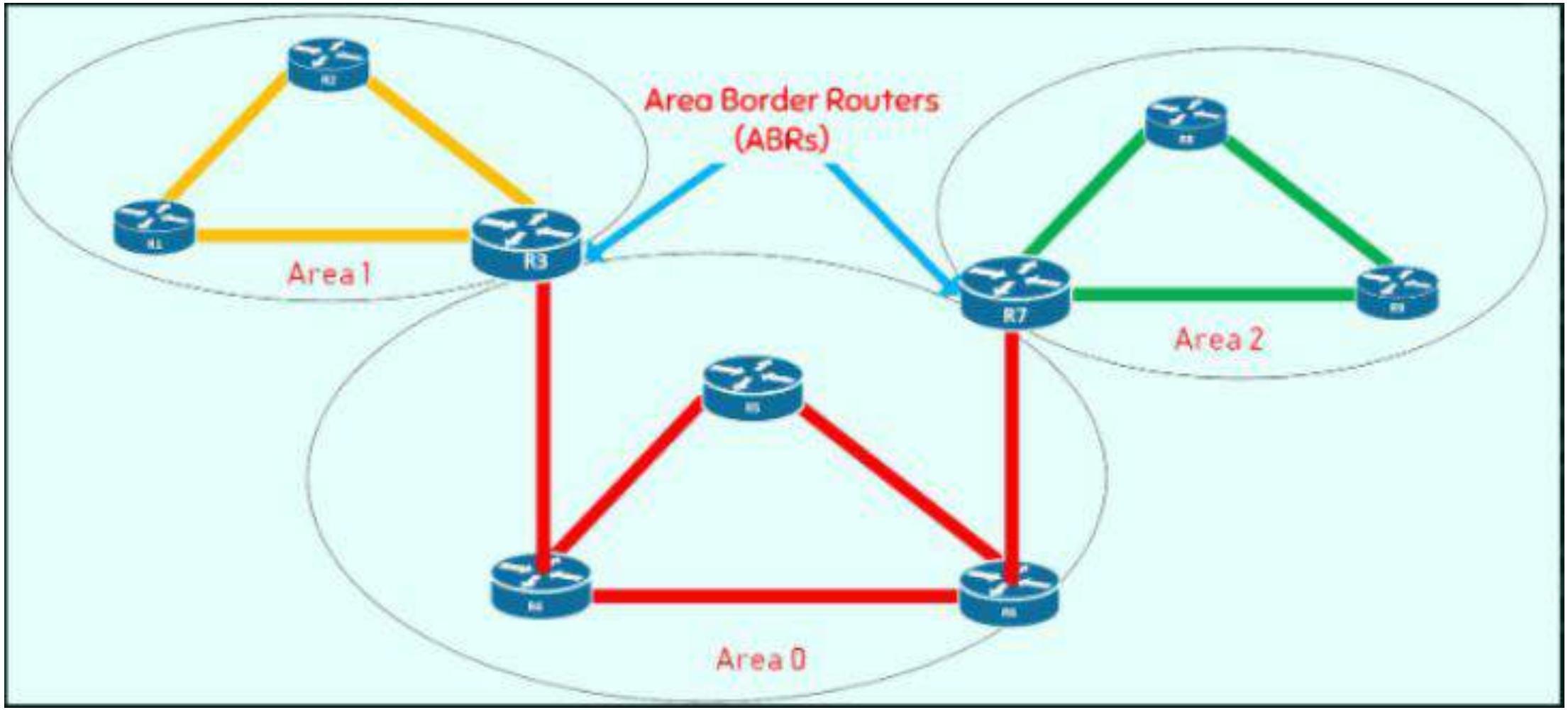
- *Each additional area must have a connection to the OSPF backbone are Such a connections are maintained by an interconnecting router, known a an Area Border Router (ABR).*

Open Shortest Path First (OSPF)

- *An ABR maintains separate link-state databases for each area it serves and maintain summarized routes for all areas in the network.*

- *OSPF detects changes in the topology, such as link failures, and converges on anew loop-free routing structure within seconds.*

- *OSPF supports complex networks with multiple routers, including backup routers, to balance traffic load on multiple links to other subnets.*


Area Border Routers
(ABRs)
R3
R7
Area 1
Area 2
Area 0

Open Shortest Path First (OSPF)

- *As a link state routing protocol, OSPF establishes and maintains neighbour relationships for exchanging routing updates with other routers.*

- *The neighbour relationship table is called adjacency database.*

- *Two OSPF routers are neighbours if they are members of the same subnet and share he same are IP, subnet mask, timers and authentication.*

- *In essence, OSPF neigbourship is a relationship between two routers that allow them to see and understand each other but nothing more. OSPF neighbours do not exchange any routing information –the only packets they exchange are HELLO packets.*

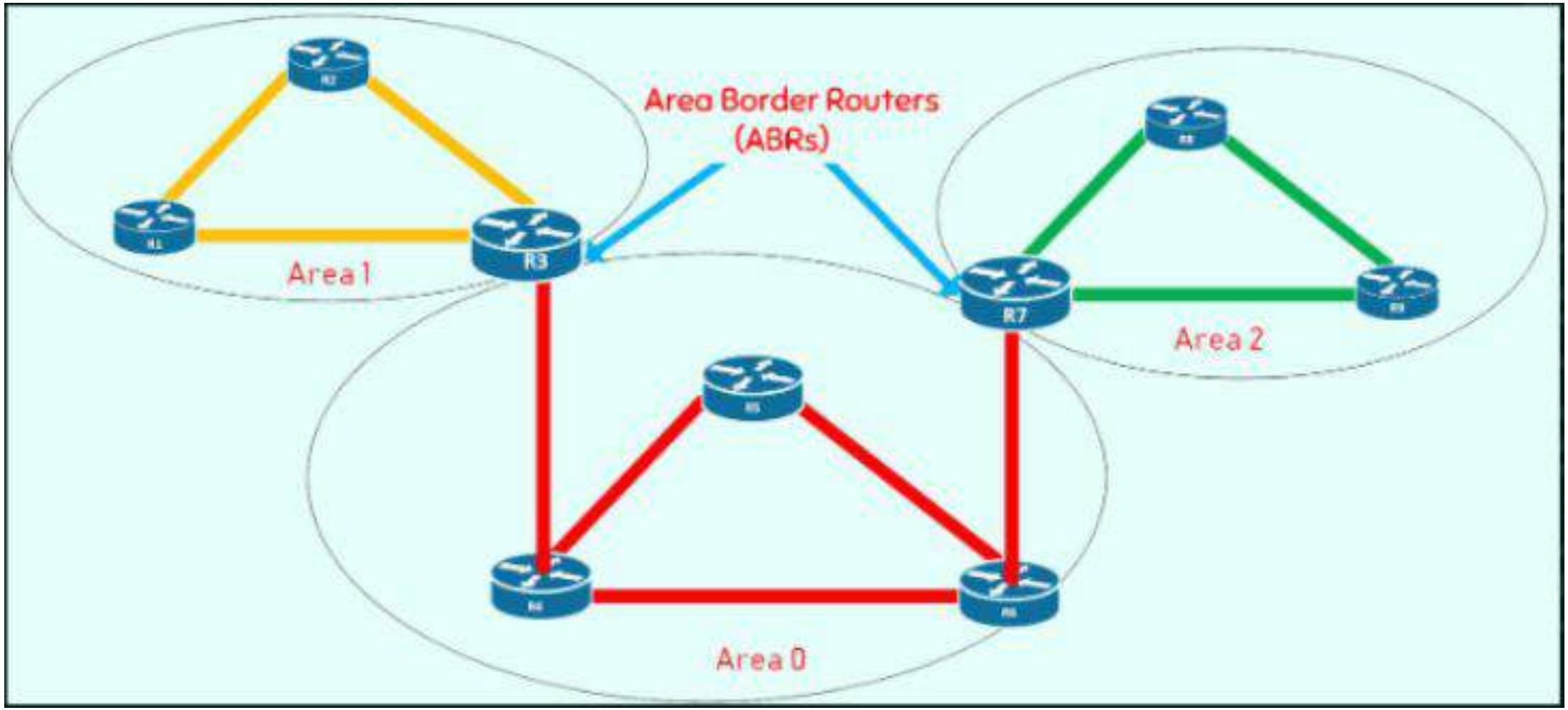
Open Shortest Path First (OSPF)

- *OSPF adjacencies are formed between selected neighbours and allow them to exchange routing information.*

- *Two routers must first be neighbours and then only then ,can they become adjacent.*

- *Two routers become adjacent if at least one of them is Designated router or Backup Designated Router (on multi-access type networks), or they are interconnected by a point-to-point or point-to-multipoint network type.*

- ***For forming a neighbour relationship between, the interfaces used to form the relationship must be in the same OSPF area.***


Area Border Routers
(ABRs)
R3
R7
Area 1
Area 2
Area 0

Open Shortest Path First (OSPF)

- *Each OSPF router within a network communicates with other neighbours routers on each connecting interface to establish the states of all adjacencies.*

- *Every such communication sequence is a separate conversation identified by the apir of router IDS of the communicating neighbours.*

- *During its course, each router conversation transitions through a maximum of eight conditions defined by a state machine*

OSPF States

1. **Down**
2. **Attempt**
3. **Init**
4. **2-way**
5. **Exstart**
6. **Exchange**
7. **Loading**
8. **Full**

Open Messages

- *OSPF defines five different message types, for communication:*

Type 1	Hello	Neighbors
Type 2	Database Description (DBD)	Adjacencies
Type 3	Link State Request (LSR)	
Type 4	Link State Update (LSU)	Reliable Update
Type 5	Link State Acknowledgment (LSAck)	

OSPF routers Type

OSPF defines the following overlapping categories of routers:

- **Internal Routers (IR):** *An internal router has all its interfaces belonging to the same area.*

- **Area Border Router (ABR):** *An ABR is a router that connects one or more areas to the main backbone network. It is considered a member of all area it is connected to.*

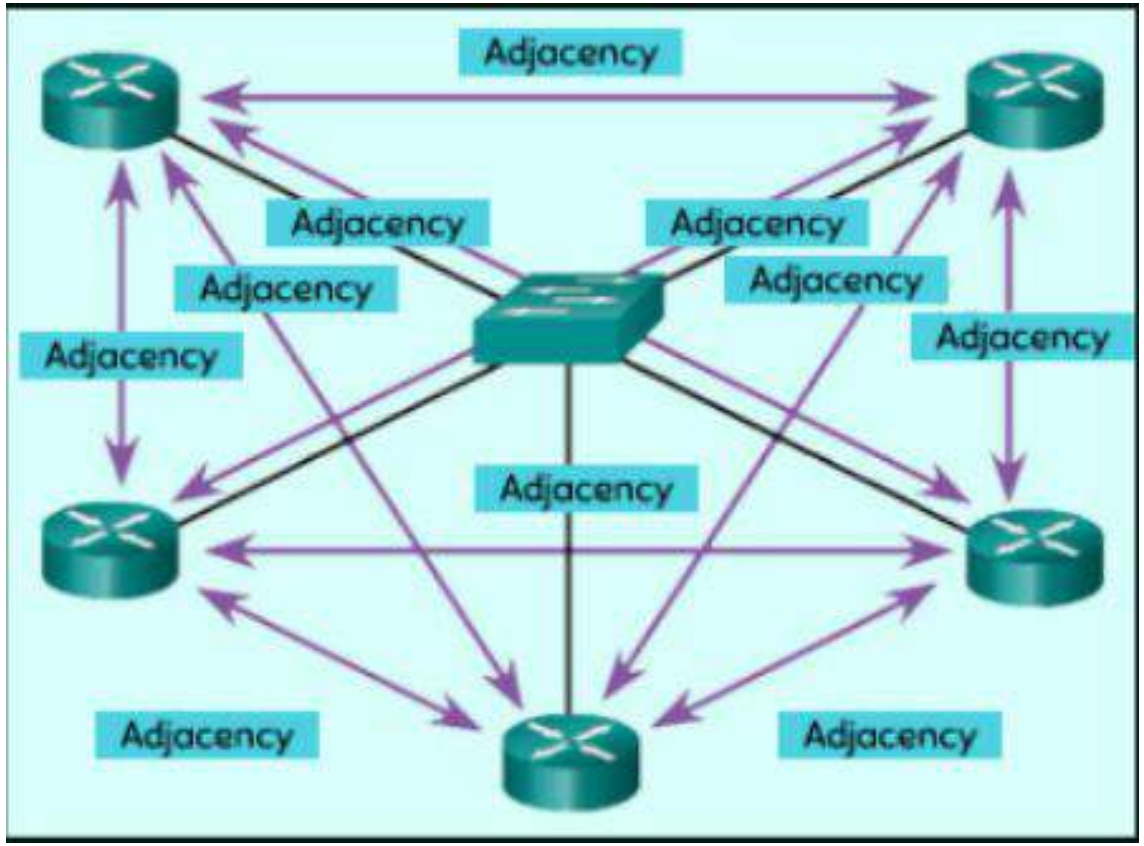
- **Backbone Router (BR):** *A backbone router has an interface to the backbone area. Backbone routers may also be area routers, but do not have to be.*

- **Autonomous System Boundary Router (ASBR):** *AN ASBR connects one AS with another or the internet.*

OSPF Routers Type

OSPF defines the following overlapping categories of routers:

- **Designated Router (DR)**
- **Backup Designated Router (DBR)**



- **If number of routers *N*, *then number of adjacencies*** $\frac{N(N-1)}{2}$
- **Example for 3 routers** $\frac{3(3-1)}{2}$ **= 3 number of adjacencies.**

Transport layer

Application → In this layer application related data is generated.

Presentation → Here data is being compressed, encrypted

Session → Dialog exchange (Dialoge control activity)

Transport → It is responsible for process to process delivery of the entire message.

Network

Datalink

Physical

Suppose at the host multiple processes are ~~generated~~ running.
Transport layer responsibility is to identify uniquely each ~~host~~ process and delivers the generated data to the intended ~~host~~ process.
To do this each process is associated with a unique port number.

Network layer or Internet protocol layer identifies networks with these unique IP address.
Host-to-host data delivery is done by Data link layer.

MAC address — Hardware address of the devices

Services provided by Transport layer :-

1. Port addressing (Source and Destination port no. are added to the data generated by the application layer and carried by the transport layer)

2. Segmentation and Reassembly

3. Connection control (connection oriented and connectionless).

4. Flow control (speed control)

5. Error control activity

Two important Transport layer protocols are
(1) TCP & (2) UDP

Scanned by TapScanner

Transport layer

Application → In this layer application related data is generated.

Presentation → Here data is being compressed, encrypted

Session → Dialog exchange (Dialoge control activities)

Transport → It is responsible for process to process delivery of the entire message.

Network

Datalink

Physical

Suppose at the host multiple processes are running generated.
Transport layer responsibility is to identify uniquely each ~~hos~~ process and delivers the generated data to the intended ~~host~~ process.
To do this each process is associated with a unique port number.

Network layer or Internet protocol layer identifies networks with these unique IP addresses.
Host-to-host data delivery is done by Data link layer.

MAC address — Hardware address of the devices

Services provided by Transport layer:-

1. Port addressing (Source and Destination port no. are added to the data generated by the application layer and carried by the transport layer)

2. Segmentation and Reassembly

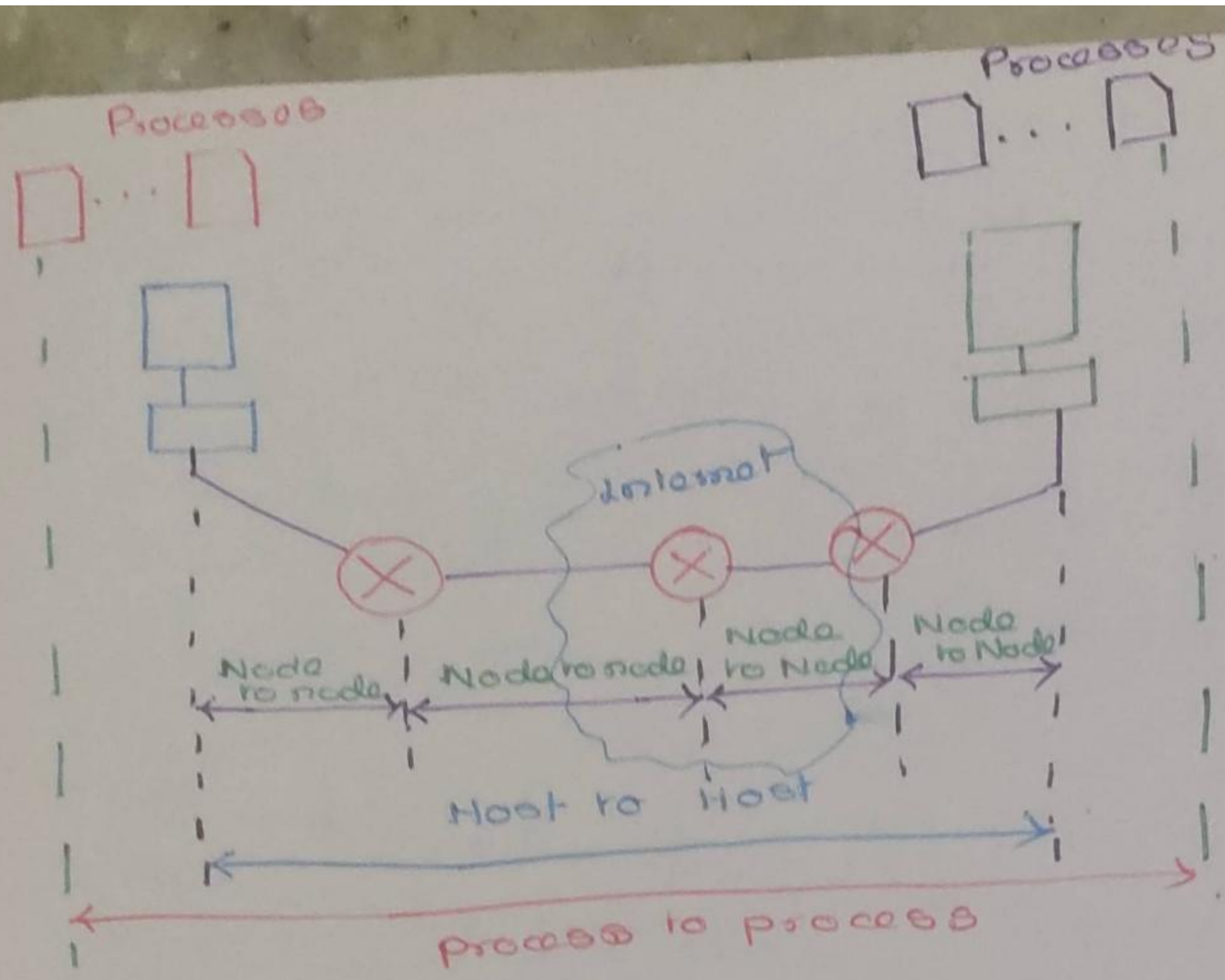
3. Connection control (connection oriented and connectionless).

4. Flow control (speed control)

5. Error control activity

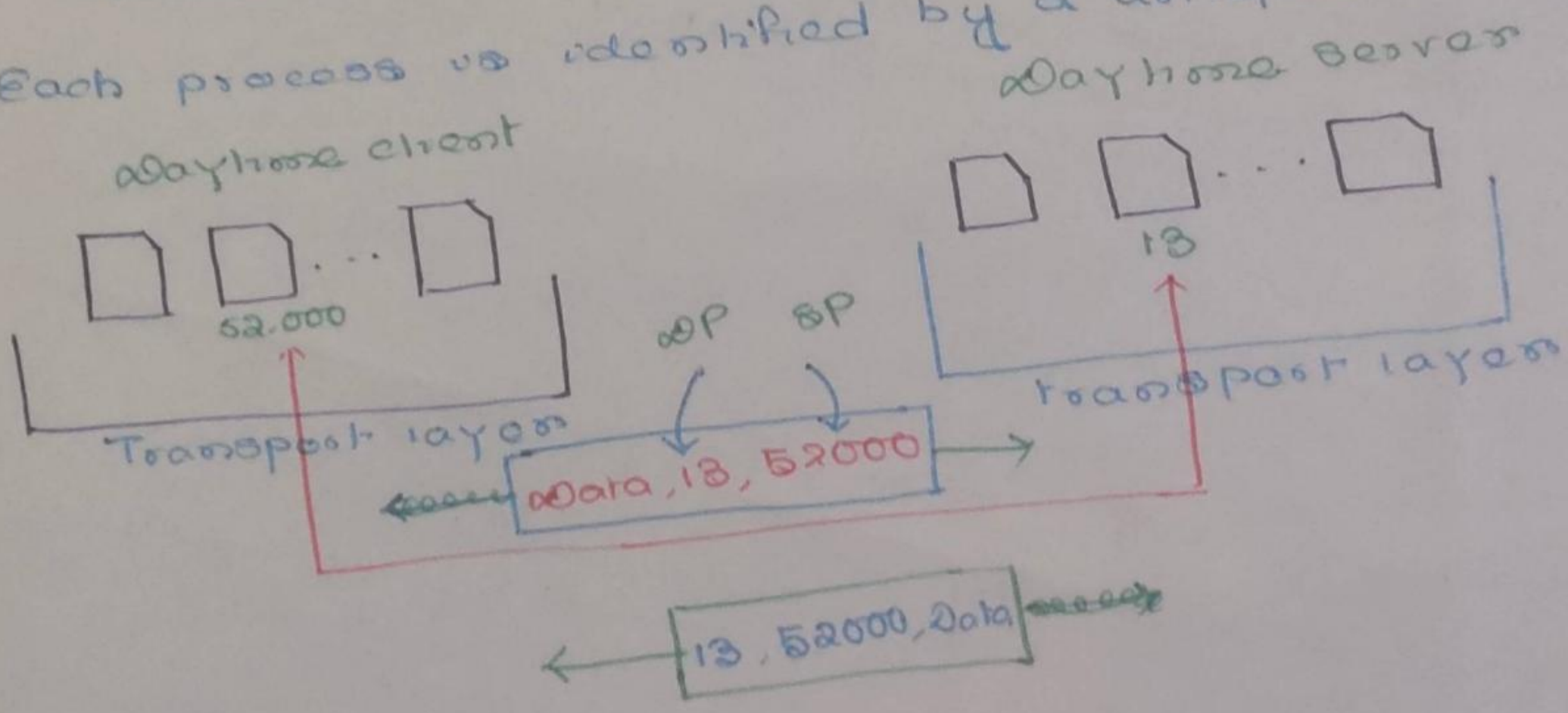
Two important Transport layer protocols are
(1) TCP & (2) UDP

Scanned by TapScanner



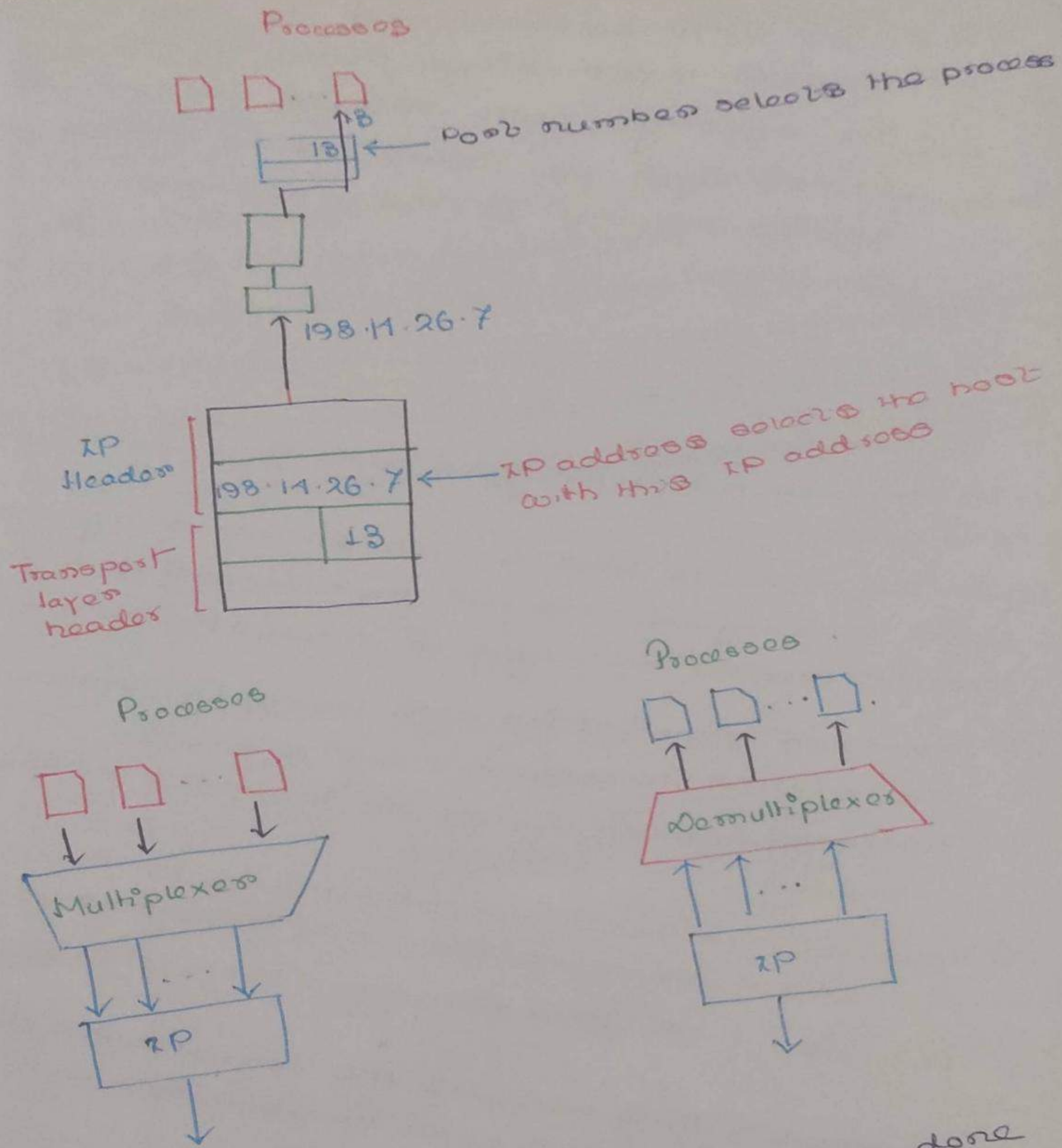
Node to Node : Data link layer
Host to Host : Network layer (IP address)
Process to Process : Transport layer (Port no)

Each process is identified by a unique port no.



DP = Destination Port NO
SP = source Port NO.

Scanned by TapScanner



Multiplexing and Demultiplexing is done at the transport layer. As many processes are running at a time in the application layer. But only one transport layer is present. So multiplexing is done and port addresses are embedded and the packets are transferred to the next layer that is network layer and at the receiver side after receiving the network layer datagram the network layer header will be removed and

The packet will be transferred to the transport layer. After demultiplexing the message with specific port address will be delivered to the specific process at the receiver device.

Port Numbers:-

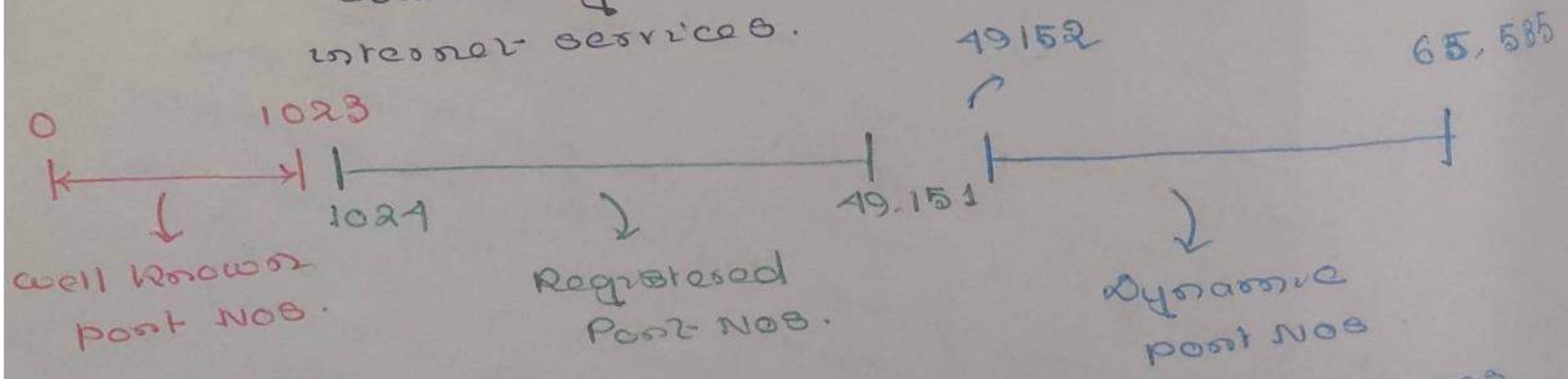
(1) In computer networking, a port is a communication end point.

(2) At the software level, within an operating system a port is a logical construct that identifies a specific process or a type of network service.

(3) A port is identified for each transport protocol and address combination by a 16-bit unsigned number, known as the port number
positive

(4) The most common transport protocols that use port number are the TCP (Transmission Control Protocol) and UDP (User Datagram Protocol).

Port numbers :- This includes the registration of commonly used port number for well-known internet services.



The port numbers are divided into three ranges

1. Well known
2. Registered
3. Dynamic or Private ports.

Scanned by TapScanner

2. Registered Port Numbers

The registered ports are those from 1024 through 49,151.

* IANA (Internet Assigned Numbering Authority) maintains the official list of well-known and registered ranges.

* IANA is responsible for the global coordination of the DNS root, IP addressing, and other internet protocol resources.

* This includes the registration of commonly used port numbers for well-known internet services.

3. Dynamic Port Numbers :- (Private port numbers)

* ranges from 49,152 to 65,535 (i.e. 216).

* Assigned by the operating system dynamically (suppose we are browsing www.google.com from our pc/laptop then this will assigned a dynamic port no to this process. As the destination port nos are very important so well-known port nos are assigned to the destinations)

*

Scanned by TapScanner

port numbers in URLs

* Port numbers are sometimes seen in web or other uniform resource locators (URLs).
* By default, HTTP uses port 80 and HTTPs uses port 443.
* URL - http://www.example.com:8080/path/ specifies the web browser to connect to port 8080 of the HTTP server.

In application layer many processes are running simultaneously. For process-to-process delivery of data we have seen port number are very important. When the application layer data is passed to transport layer, the transport layer headers are appended to the data. It may be TCP (Transmission control Protocol) and UDP (User Datagram Protocol) header depending on the information content.

Some information communication can allow some delay but cannot compromise on the reliability (i.e. data losses) and some information communication can allow some losses on the otherhand cannot allow any delay.

Data link layer - communication is done using MAC addresses i.e hardware address.

Network layer/Internet layer - communication is done using IP address or Internet protocol addresses these are often named as software addresses.

Scanned by TapScanner

Transport layer:- communication is done using port address of both the source process and the destination process.

The transport layer moves data between applications on devices in the network

~~Application layer:-~~

Application layer → Is the direct interface betn users and data are generated & retrieve at this layer

Two important protocols

- TCP (Reliable)

~~UDP~~

- No loss but delay
- e.g ⇒ Email
- Some delay can be allowed by no loss in information is allowed
- Retransmission is possible.

(1) HTTP (Website/webbrowsers)
(2) FTP
(3) SMTP
(4) Telnet (Remote login)

- UDP (unreliable)

- No delay but low loss
- Real time traffic
- Video ~~conferencing~~ conferencing, live call.
- Delay may introduce chaos in the communication and retransmission is also not possible.

(1) DHCP (Dynamic Host configuration Protocol)
(2) DNS
(3) SNMP
(4) TFTP
(5) VoIP

Scanned by TapScanner

UDP (User Datagram Protocol):-

- Simple protocol that provides the basic transport layer function.
- used by applications that can ~~tolerate~~ tolerate small loss of data (as no retransmission is possible)
- used by applications that cannot tolerate delay

used by :- (1) DNS
(2) VoIP
(3) ~~SMTP~~ SNMP (Simple Network Management Protocol).
(4) DHCP
(5) TFTP (Trivial File Transfer Protocol)
(6) Online games.

* connectionless and unreliable. → no connection establishment between source & destination
* Prior communications (or handshaking) are not required in order to setup communication channels or data paths.
* UDP-based servers applications are assigned well-known or registered port numbers.
* UDP client process randomly selects port numbers from range of dynamic port numbers (these are operating system assigned port numbers) as the source port
* UDP is suitable for purposes where error checking and correction are either not necessary or are performed in the applications. UDP avoids the overhead of such processing in the protocol stack.
* No error correction.

Scanned by TapScanner

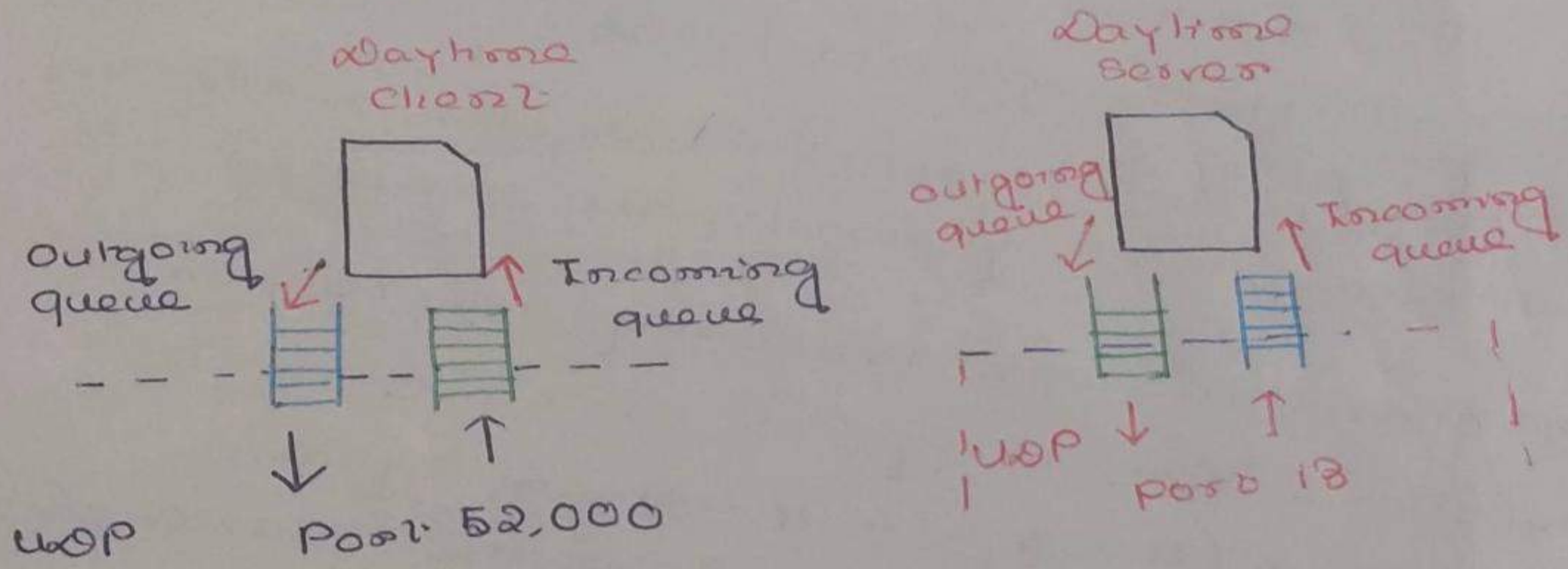
* Time-sensitive applications often use UDP because dropping packets is preferable than to waiting for packets due to retransmission which may not be an option in a real-time traffic on system.

* No flow control mechanism.

The UDP ~~Packets~~ headers are very simple and have only four fields. These UDP headers are kept simple because UDPs are used for real-time communication so when a router receives an UDP packet it should not take more ~~time~~ time to process it.

* Encapsulation and Decapsulation is done by the UDP. protocol.

* Queuing is done the UDP.



In application layer many processes are running simultaneously. If the client has two queue the server also has to maintain two queues. The UDP dequeue messages one after another an add UDP header and passes it to the next layer i.e. the network layer to add IP header. When the process are terminated at the application layer, the queues are also discarded.

Scanned by TapScanner

Uses of UDP :-

(1) It is a transaction-oriented, suitable for simple request-response protocols such as DNS (Domain Name System) or the NTP (Network Time Protocol).

(2) It provides datagrams, suitable for modeling other protocols such as IP tunneling or remote procedure call and the Network File system.

(3) It is simple, suitable for bootstraping or other purposes without a full protocol stack, such as the DHCP and TFTP.

(4) It is stateless, suitable for very large numbers of clients, such as in streaming media application such as IPTV. (because of simplied header).

(5) The lack of retransmission delays makes it suitable for real-time applications such as VoIP, online games and many real-time streaming applications.

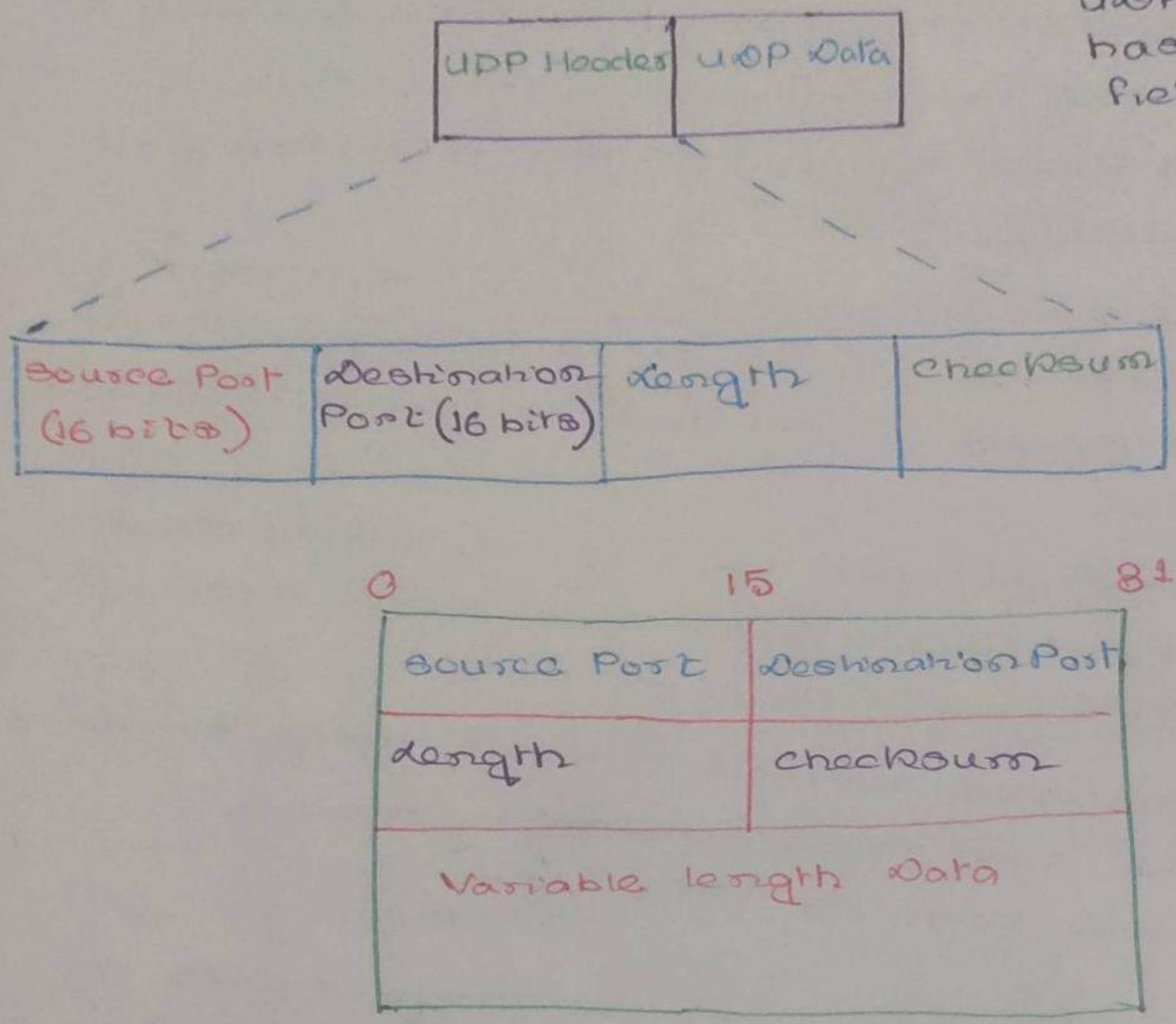
(6) UDP is a suitable transport protocol for multicasting. Multicasting capability is embedded in the UDP software but not in the TCP software

(7) UDP is used for some route updating protocols as routing information protocol (RIP)

Scanned by TapScanner

UDP Header Format :-

UDP Header has only Four fields



1024

All the four field are of 16 bits length.

Source port Number :-

- This is the port number used by the process running on the source host. usually this port number is assiged by the operating system.
- It is a 16-bits long. which means that the port number can range from 0 to $(2^{16}-1)$ i.e. 65.535.
- If the source host is the client (a client sending a request), the port number is most cases is an ephemeral port number i.e. random port number ~~greate~~ greater than 1024 (well known and registered) requested by the process and choosen by the UDP software running on the source host.

Scanned by TapScanner

- If the source host is the server (a server sending a response), the port number in most cases is a well-known port number

Destination Port number:-

- This is the port number used by the process running on the destination host
- It is also 16 bits long. If the destination host is the server (a client sending a request) the port number, in most cases, is a well-known port number.
- If the destination host is the client (a server sending a ~~request~~ response), the port number in most cases, is an ephemeral (i.e. random) port number
- In this case, the server copies the ephemeral port number it has received in the request packet.

Length Field:-

- This is a 16-bit field that defines the total length of the user datagram, header plus data. The 16 bits can define a total length of 0 to $(2^{16}-1)$ i.e. 65,535 bytes.
- However the total length of the user datagram need to be much less because a UDP user datagram is again stored in an IP datagram with a total length of 65,535 bytes.
- UDP length = IP length − IP header's length.

Scanned by TapScanner

UDP checksum

Pseudo-header in UDP:-

- The UDP checksum calculation is different from one for IP and ICMP.
- In UDP, the checksum includes three sections:
  - A Pseudo-header
  - The UDP header and
  - The data (coming from the application layer).

* The pseudo-header is the part of the header of the IP packet in which the user datagram is to be encapsulated with some fields filled with 0's (zeroes).

* If the checksum does not include the pseudo-header, a user datagram may arrive at destination with out errors.

* However, if the IP header is corrupted, it may be delivered to the wrong destination.

The data generated by the top-most layer that is the application layer pass to the transport layer where the UDP header is encapsulated with the data. Now in the network layer along with the IP header a header field containing 0's is

Scanned by TapScanner

added at the IPV4 field filled with 0's.
which is called as Pseudo-header.

- The protocal field is added to ensure that the packet belongs to UDP. and not to other transport layer protocols.
- The value of the protocol field for UDP is 17.
- If the value is changed during transmission the checksum calculation at the receiver will detect it and UDP drops the packet.
- Thus, the packet is not delivered to the wrong destination.

32-bit source IP address		
32-bit destination IP address		
All 0's (8 bits)	8-bit protocol (17)	16-bit UDP total lenght

Pseudo-Header

Source Port Address (16 bits)	Destination Port Address (16 bits)
UDP total lenght (16 bits)	Checksum 16 bits
Variable lenght data.	

UDP Header

Scanned by TapScanner

TCP (Transmission control Protocol) :-

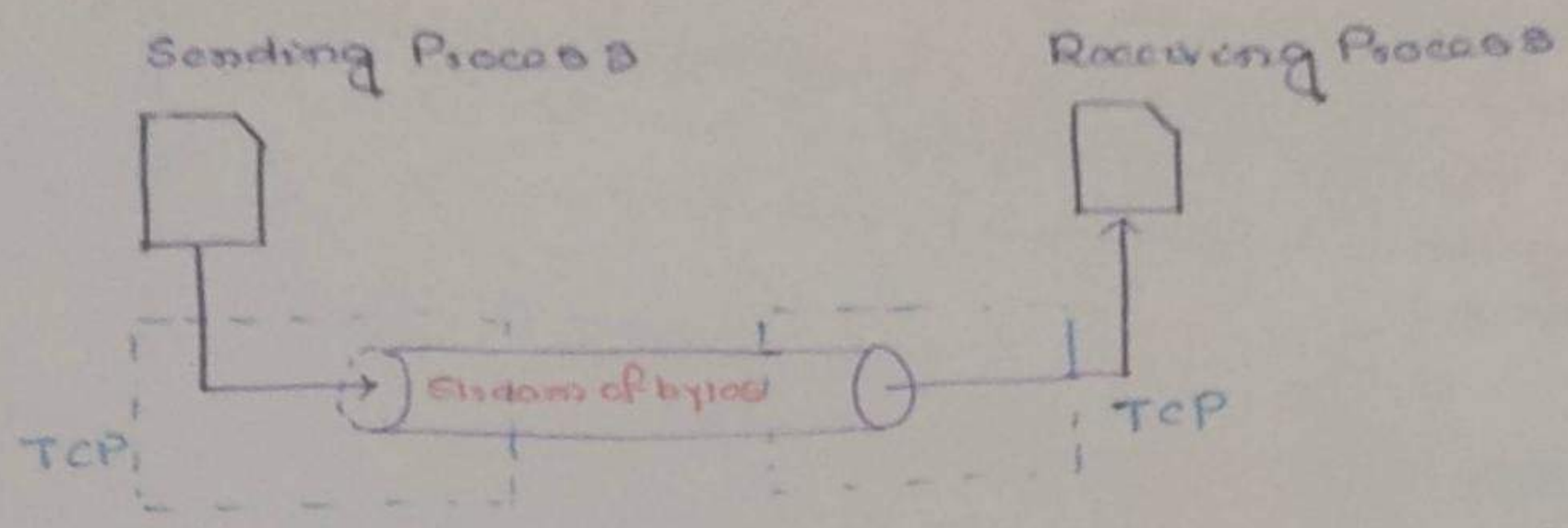
- TCP is the widely used Transport layer protocol.
- It creates a virtual connection between two TCPs to send data.
- It creates addition. TCP uses flow and error control mechanisms at the transport level
- Used by application that can tolerate delay but cannot tolerate loss.
- used by ~~proto~~ application such as:
  1. HTTP
  2. FTP
  3. Telnet
  4. SMTP
  5. world wide web (www)

Features of TCP

- TCP is a connection oriented protocol connection is established using three way handshaking before starting the communication
- Reliable delivery (sender will know receiver has received the data or retransmission will be done)
- Acknowledgment oriented.
- Retransmission.
- Flow control (Speed control mechanism)
- Error control
- Congestion control
- Segmentation and Reassembly
- Full Duplex Support.

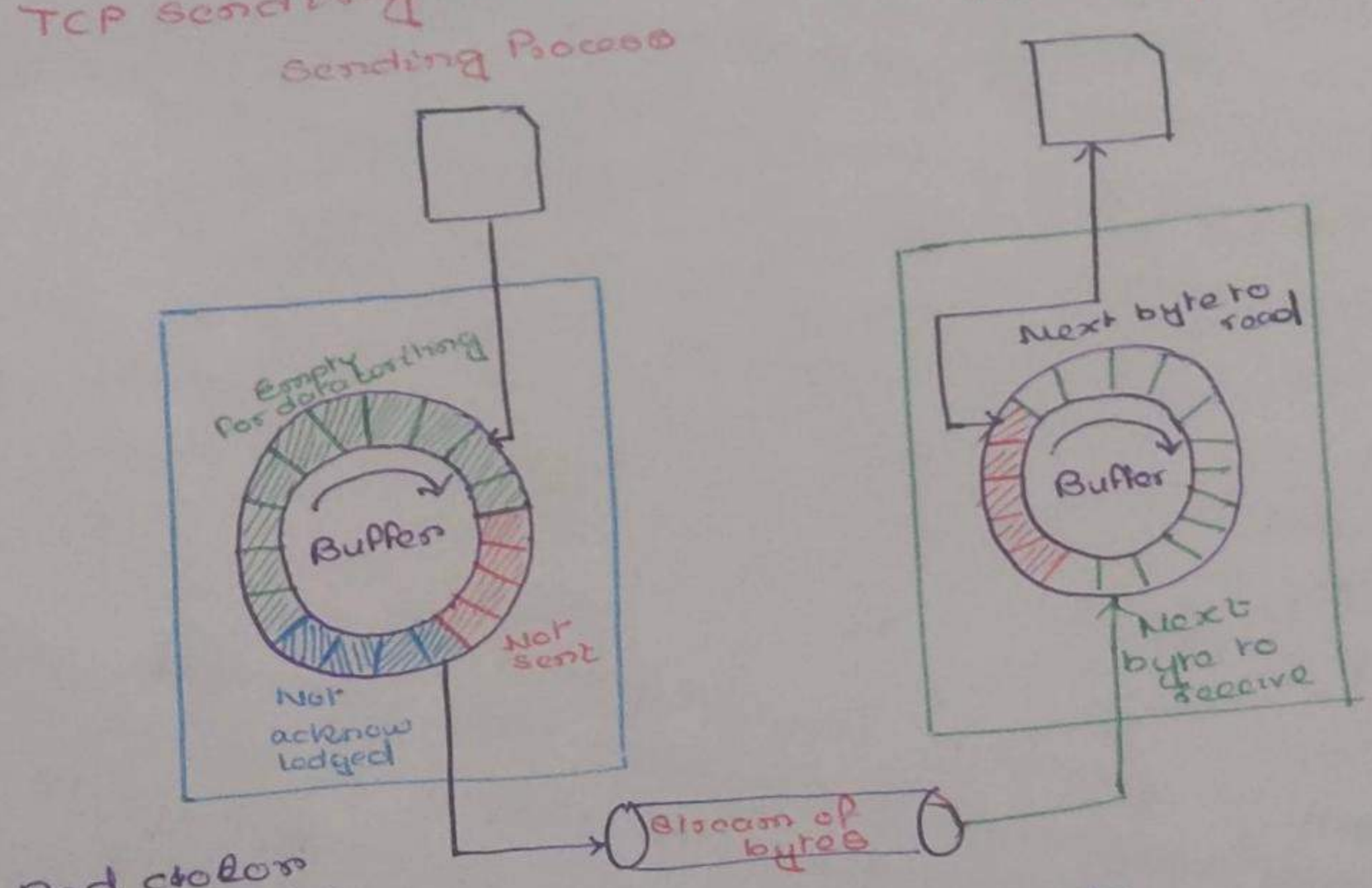
Scanned by TapScanner

TCP stream Delivery :-



This sending and receiving process are running in two different system (e.g. computer/laptop/ cellphones). This imaginary tube will deliver the data in form of stream of bytes.

TCP sending and Receiving Buffer :-



- Here Red color indicated data bytes which are not sent.
- Blue cells indicates data sent by not acknowledged
- Green indicated empty buffers when data will be stored
- After acknowledgment received the ~~chan~~ chambers will be recycled.

Scanned by TapScanner

TCP segments:-

Each segment can of of different lengths. Each segment carries hundred or thousand of bytes. The TCP adds TCP header to each of the segment for control and error detection activities.

Application layer generates data

↓

This data is passed to the transport layer The transport layer add TCP header and make segments consists of ~~data stream~~ streams ~~bytes~~ of data bytes. Transport layer ~~user~~ datagrams are called segment

↓

The segments will be passed to the network layer. And in this layer IP header will be embeded to the segments. This network layer datagrams are called packets. If TCP headers are added then it is called segment. If UDP headers are added

The IP packets are then sent to the data link layer for transmission. In this layer the communication link can be wired or wireless. The data link layer will add their header to the IP packets received from the network layer, and will make a frame.

↓

This will be pass to the physical layer and this layer will convert the received frames from the data link layer to electromagnetic signals or optical signals. The physical layer protocol data units are called bits.

Scanned by TapScanner

TCP header Format :-

Source Port Address 16 bits								Destination Port Address (16 bits)
Sequence No (32 bits)								
Acknowledgement No (32 bits								
HLEN 4 bits	RES 6 bits	URG	ACK	PSH	RST	SYN	FIN	window size (16 bits).
Checksum (16 bits)								Urgent Pointer (16 bits)
Options and Padding								

if options are added than 60 bytes (minimum Header)

The TCP header lenght is 20 bytes. The TCP header plus options and padding length can go upto 60 bytes. i.e. The options and padding lenght can maximum go upto (60-20) = 40 bytes.

We have seen that the UDP has only 4-fields. where as TCP/IP has more no. of fields. As TCP provides a reliable and acknowledgement oriented communication with error control and flow control mechanism.

Source Port Number :-

(1) It defines the port number of the application program in the host device that is sending the segment. We have seen in a client host many processes are running simultaneously. The process which is sending/requesting the data, this port number is associated with the process.

Scanned by TapScanner

(2) This serves the same purpose as the source port address in the UDP header.

Destination Port address :-

(1) It defines the port number of the application program in the host that is receiving the segment. (well known port numbers).

(2) This serves the same purpose as the destination port address in the UDP header

Sequence Number :-

This is a 32-bit field that defines the number assigned to the first byte of data contained in the segment. This sequence number helps us to know about missing packets, lost packet, out of order and delayed packets

Has a dual role :-

- If the SYN flag is set to 1, then this is the initial sequence number. The sequence number of the actual first data byte and the acknowledged number in the corresponding ACK are then this sequence number plus 1.

Suppose the sequence number of the first data byte is 1001 then the acknowledgment (ACK) number sent by the receiver is 1002 i.e. the receiver is expecting the second data byte

- If the SYN flag is 0, then this is the accumulated sequence no. of the first data byte of this segment for the current session.

Scanned by TapScanner

Acknowledgement number:-

(1) This is a 32 bit field.

(2) It defines the byte number that the receiver of the segment is expecting to receive from the other party.

(3) If the receiver of the segment has successfully received byte number x from the other party it defines (x+1) as the acknowledgement number. Acknowledgement and data can be piggybacked together.

HLEN (Header length):-

- It indicates the number 4-byte words in the TCP header.
- The length of header can be from 20 byte to 60 byte i.e. (60 - 20 = 40 bytes is options & padding)
- Therefore, the value of this field can be between 20 bytes i.e (5×4 = 20) 5 and 60 bytes i.e (15×4 = 60) 15. As the HLEN field contains 4-byte words.

RES (Reserved field)

- It is a 6-bit field.
- This is reserved for future use and usually filled with 0's.

Scanned by TapScanner

(6 Flag bits in TCP header :-

U	A	P	P	S	F
R	C	S	S	Y	I
G	K	H	T	N	N

G: Flags ~~bits~~ fields are consists of 6 different control bits or flags each of the 6 flags are of 1 bit. One or more of these bits can be set at a time

(2 URG (1 bit) :- when this is set i.e. 1 the urgent pointer field find its significance.

S ACK (1 bit) :- Indicates that the acknowledgement field is significant.

T PSH (1 bit) :- PUSH function. Ask to push the buffered data to the receiving application.

RST (1 bit) :- Reset the connection.

SYN (1 bit) :- Synchronize sequence numbers.

FIN (1 bit) : Last packet from sender.

Scanned by TapScanner

window size :-

- This field defines the size of the window in bytes, that the other party must maintain. Suppose the receiver window is of 20 byte receiver can handle 20 bytes at a time. But if the sender without knowing the receiver window size sends 50 bytes then it can only receives 20 bytes and other bytes will be discarded. So it is very important for the receiver to inform about its window size

This field defines the size of the wi.

- The length of this field is 16 bits, which means that the maximum size of the window is $(2^{16}-1) = 65,535$ bytes.

- This value is normally referred to as the receiving window (rwnd) and is determined by the receiver.

- The sender must obey the dictation of the receiver in this case.

checksum :-

- The calculation of the checksum for TCP follows the same procedure as the one described in UDP.

- However, the inclusion of the checksum in the UDP datagram is optional (i.e. if checksum is not added by the UDP it is the responsibility of application layer to calculate), whereas the inclusion of the checksum for TCP is mandatory.

- The same pseudo-header, serving the same purpose, is added to the segment. For the TCP pseudo-header, the value for the protocol field is 6. (for UDP it is 17)

Scanned by TapScanner

Urgent Pointer :-

① This 16-bit field, which is valid only if the urgent flag is set, is used when the segment contains urgent data.

We have seen the real-time traffic or urgent data is normally sent by the UDP transport protocol. But in some case we want a data to be sent urgently and also maintains its reliability, then this is used in TCP.

(2) It defines the number that must be added to the sequence number to obtain the number of the last urgent byte in the data section of the segment.

Options :-

This is field can be upto a maximum of 40 bytes. These can be upto 40 bytes of optional optional information in the TCP header.

- The length of this field is determined by the Data offset field. options have up to three fields :-

a. Option - Kind (1 byte)
b. Option - length (1 byte).
c. Option - data (Variable).

Scanned by TapScanner

Padding:-

The padding is composed of zeros. The TCP header padding is used to assure that the TCP header ends, and data begins, on a 32-bit boundary.

TCP connection:-

1. connection-oriented
2. virtual path (seq + acknowledgement).
3. Acknowledgment process
4. Retransmission of lost damaged frames
5. TCP connection- virtual not physical
6. IP- connectionless
7. Full-duplex mode
8. Approval from other party

Three phases of TCP connection.

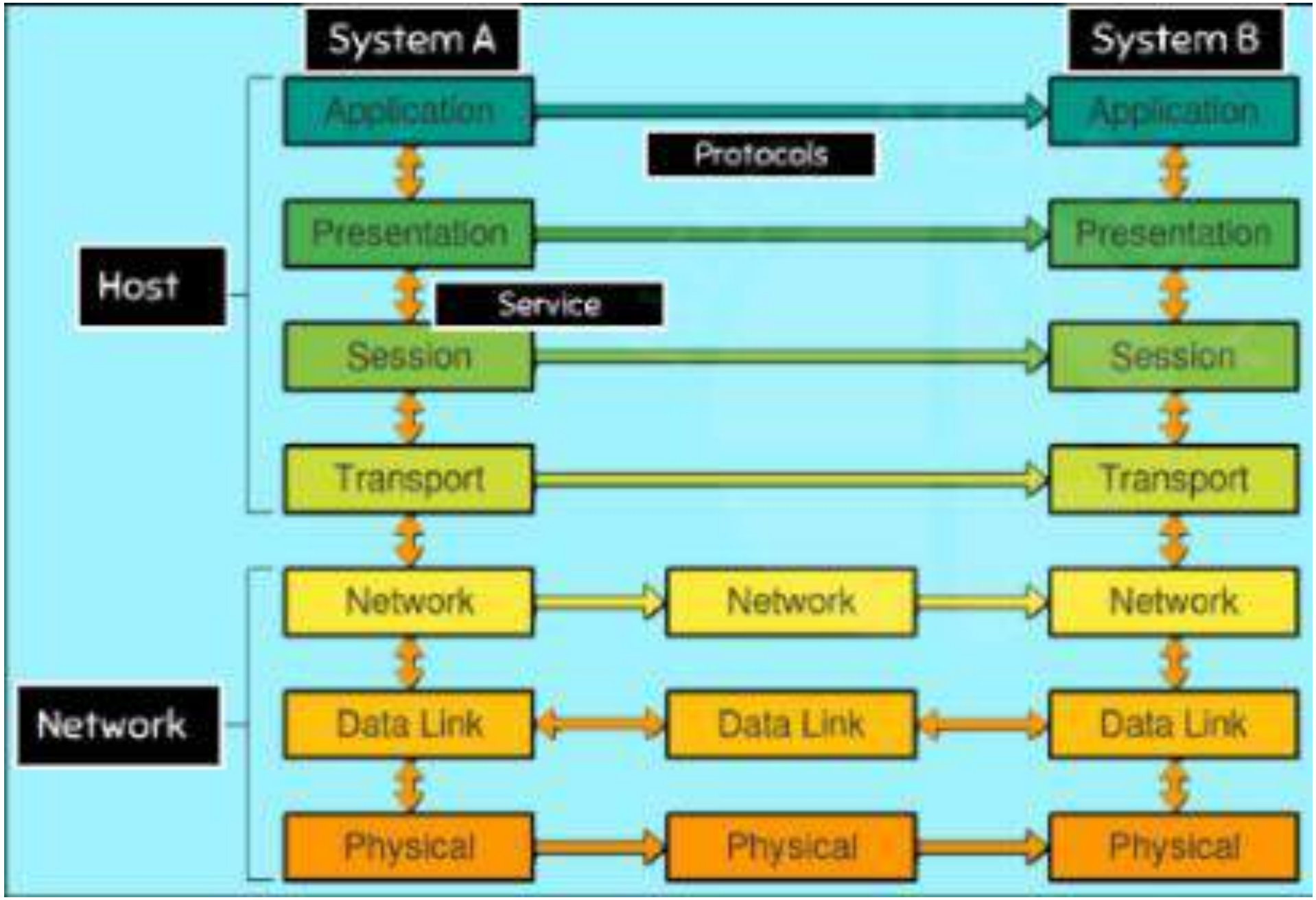
1. Connection Establishment
2. Data transfer
3. Connection Termination.

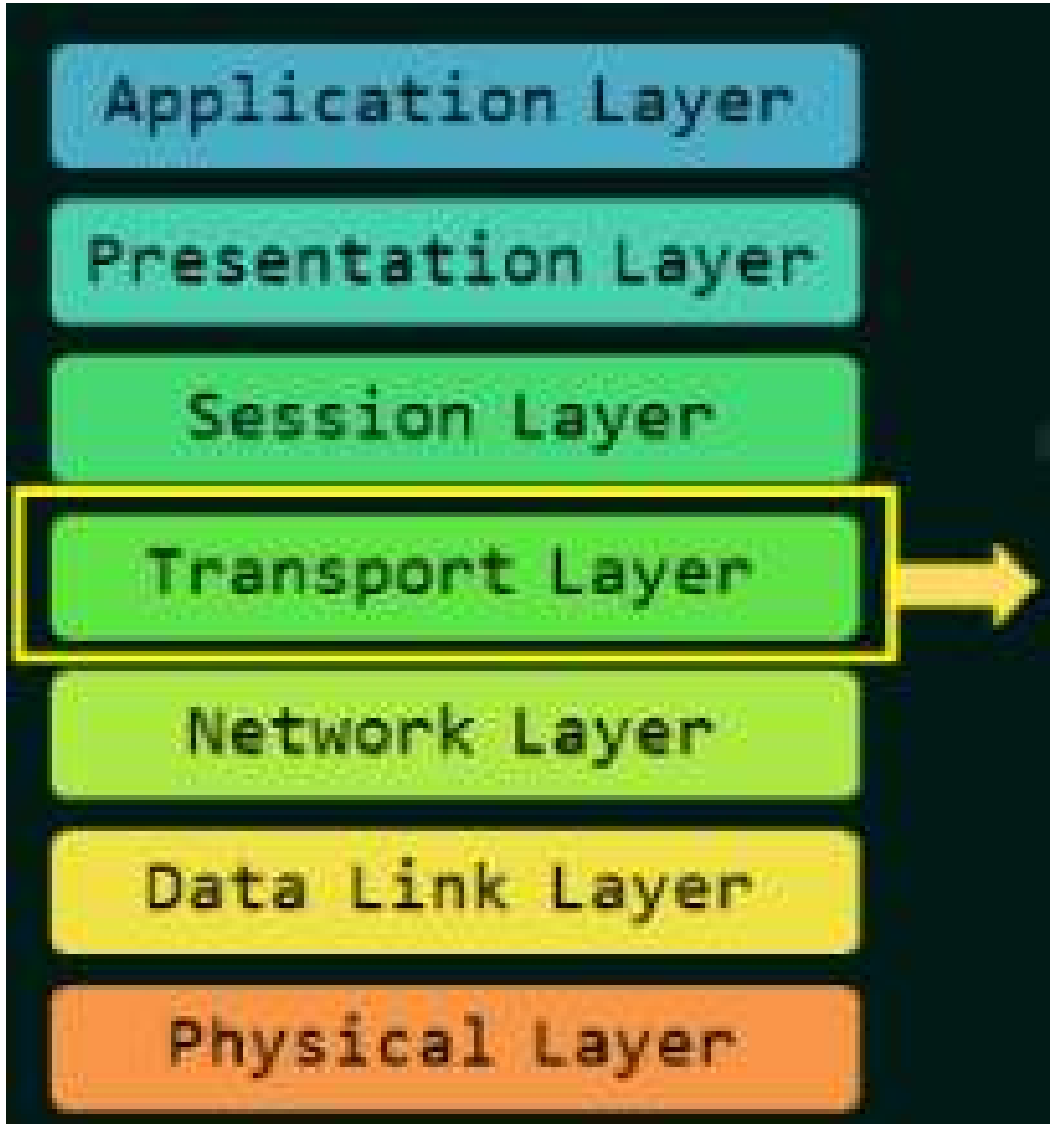
Scanned by TapScanner

Module 4

Transport Layer

Dr. Sunandita Debnath, IIIT Vadodara

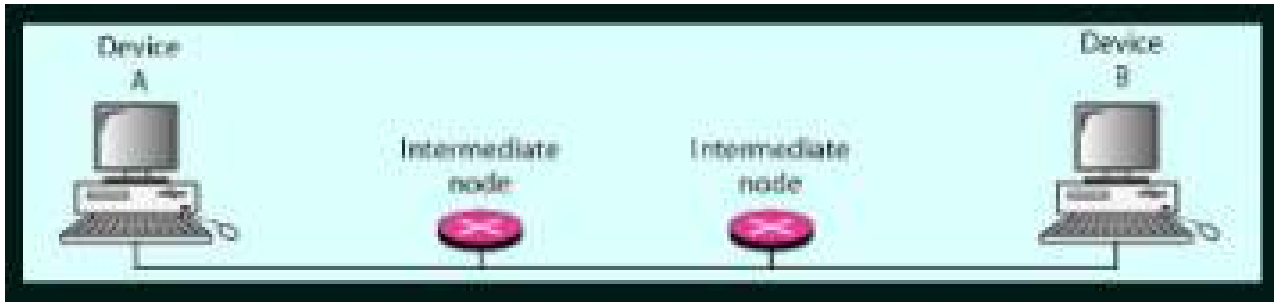

System A
System B
Application
Application
Protocols
Presentation
Presentation
Host
Service
Session
Session
Transport
Transport
Network
Network
Network
Network
Data Link
Data Link
Data Link
Physical
Physical
Physical

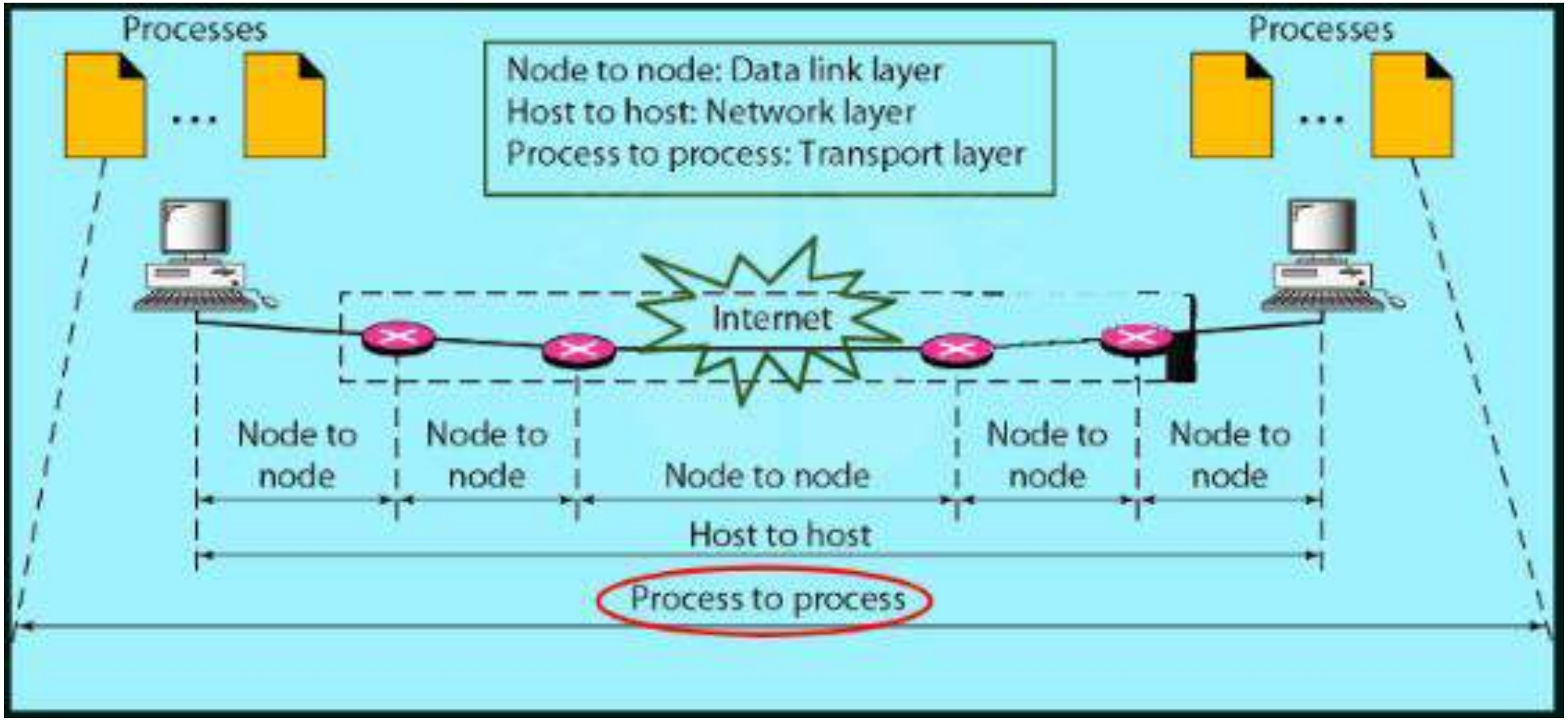


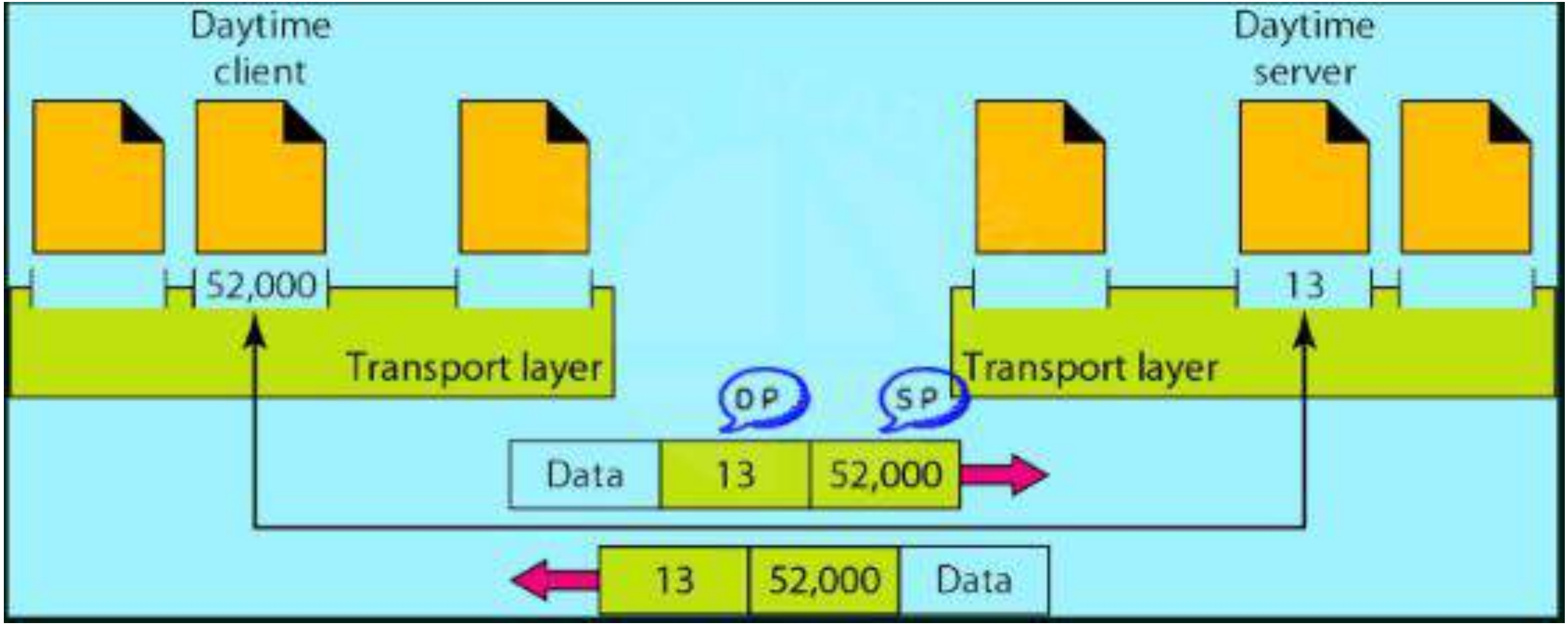
It is responsible for process to process delivery of the entire message.

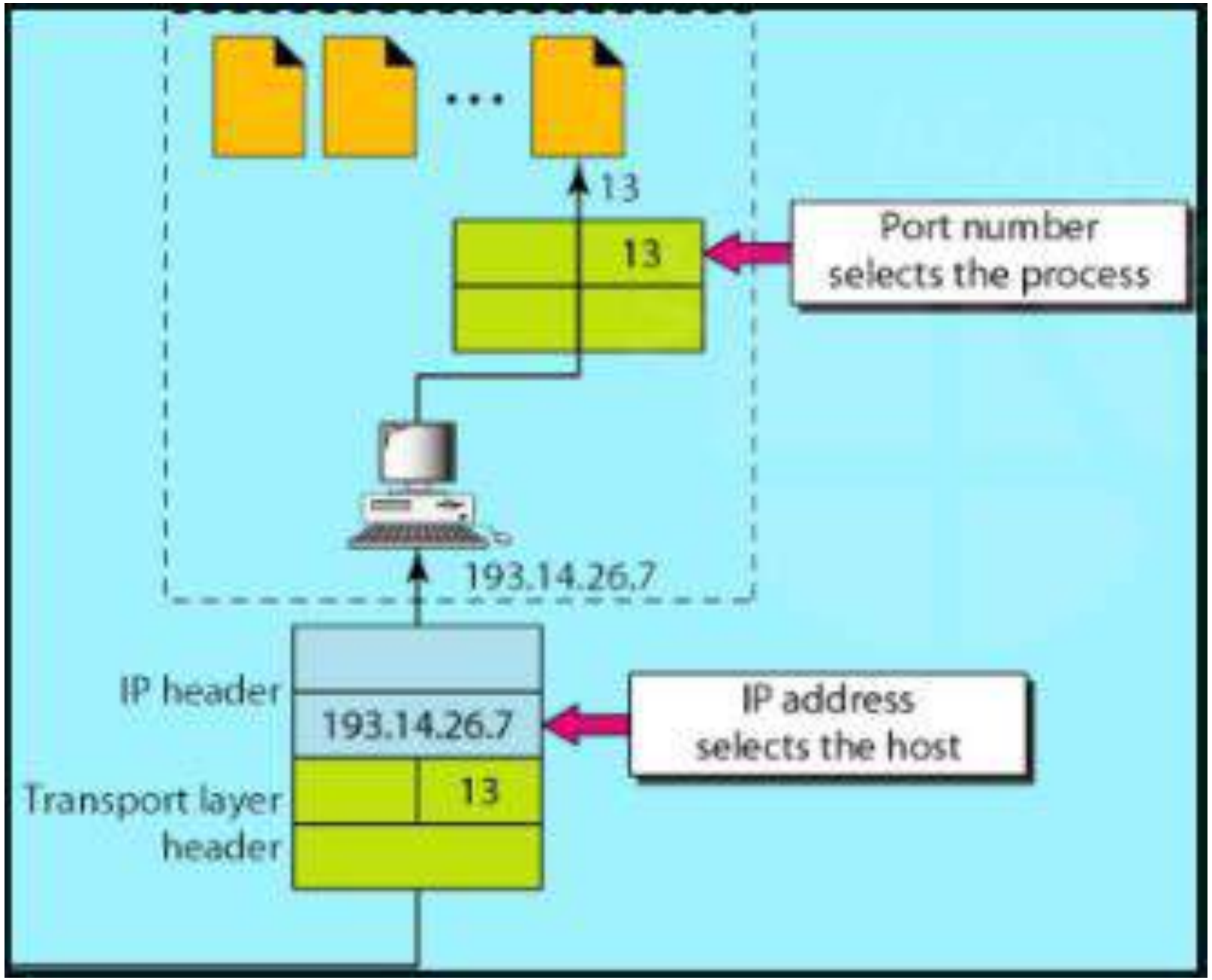
Service provided by Transport Layer

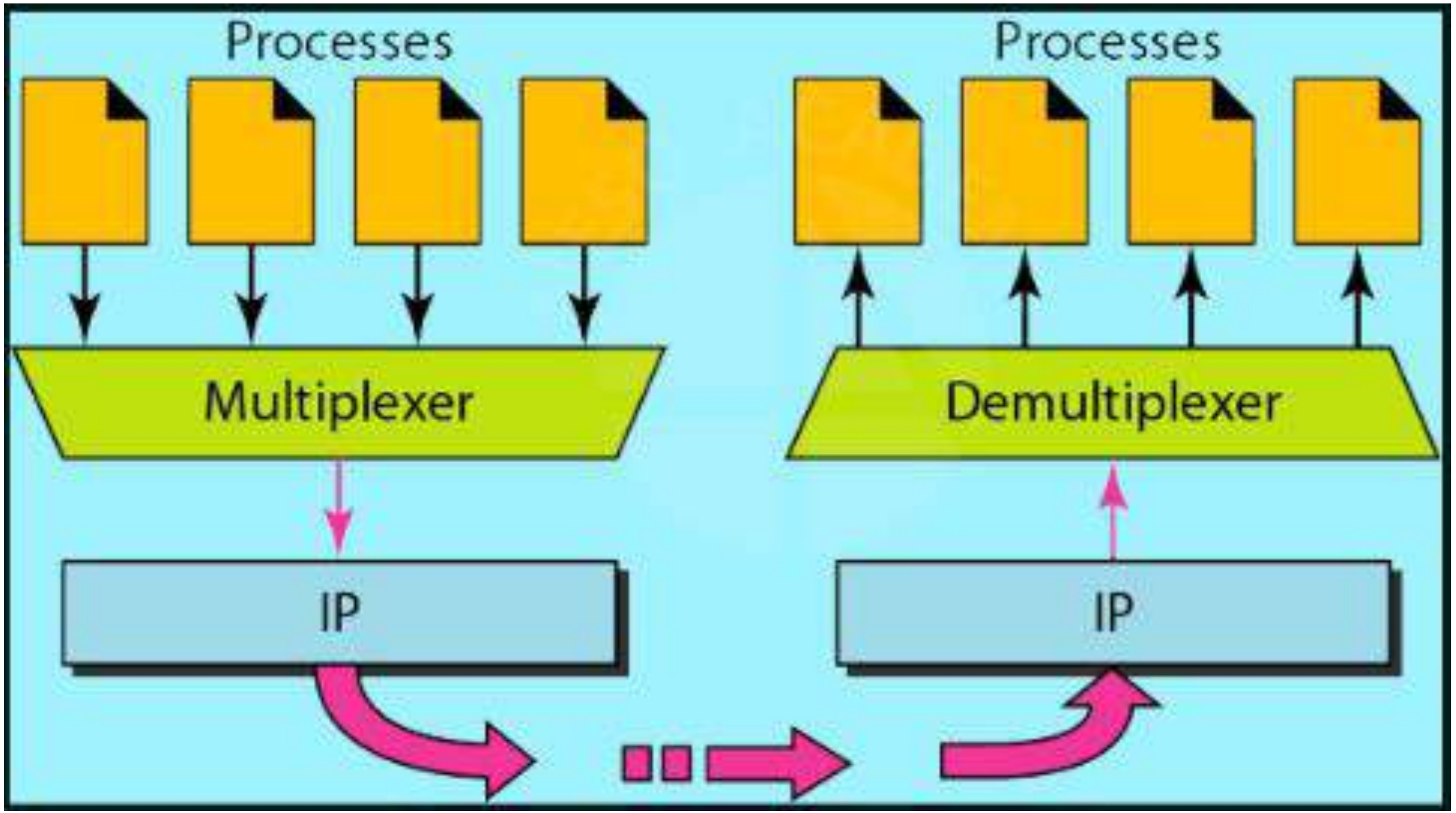
- **Port addressing**
- **Segmentation and Reassembly**
- **Connection Control**
- **Flow control**
- **Error control**

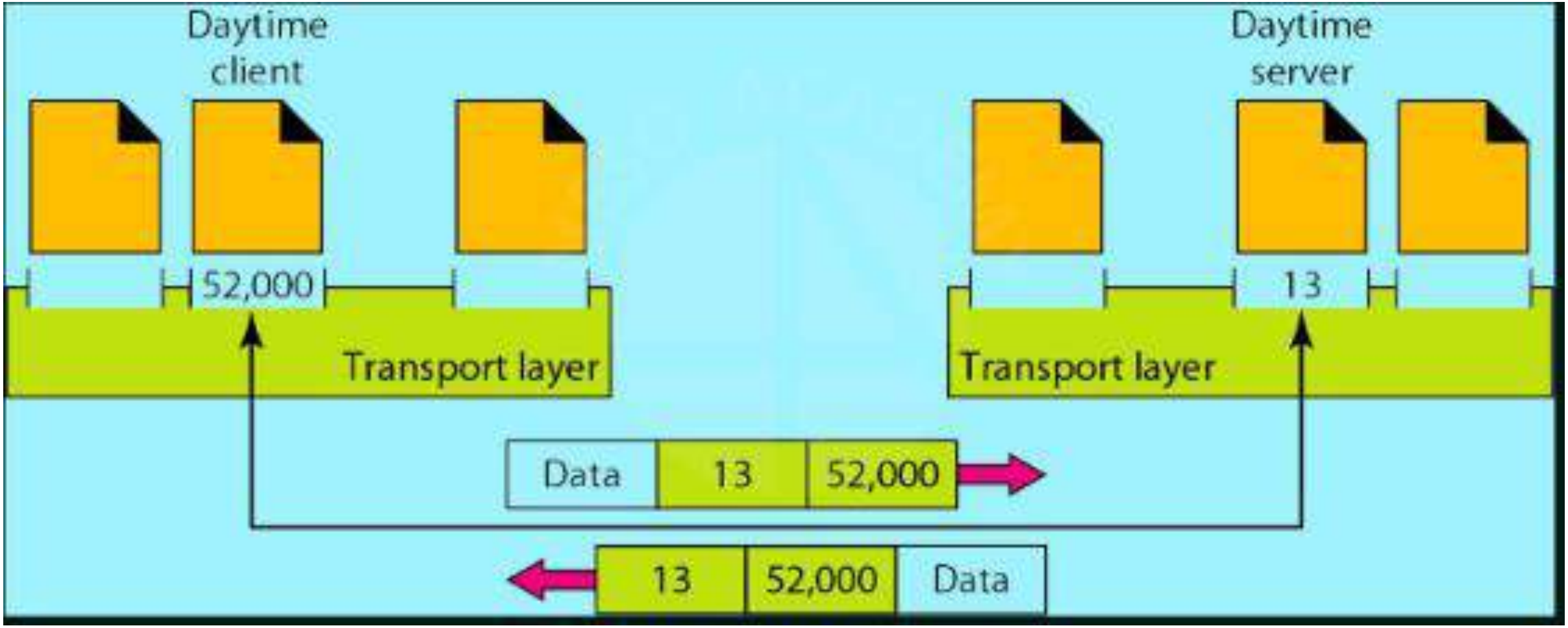



Processes
Processes
Node to node: Data link layer
Host to host: Network layer
Process to process: Transport layer
Internet
Node to node
Node to node
Node to node
Node to node
Node to node
Host to host
Process to process

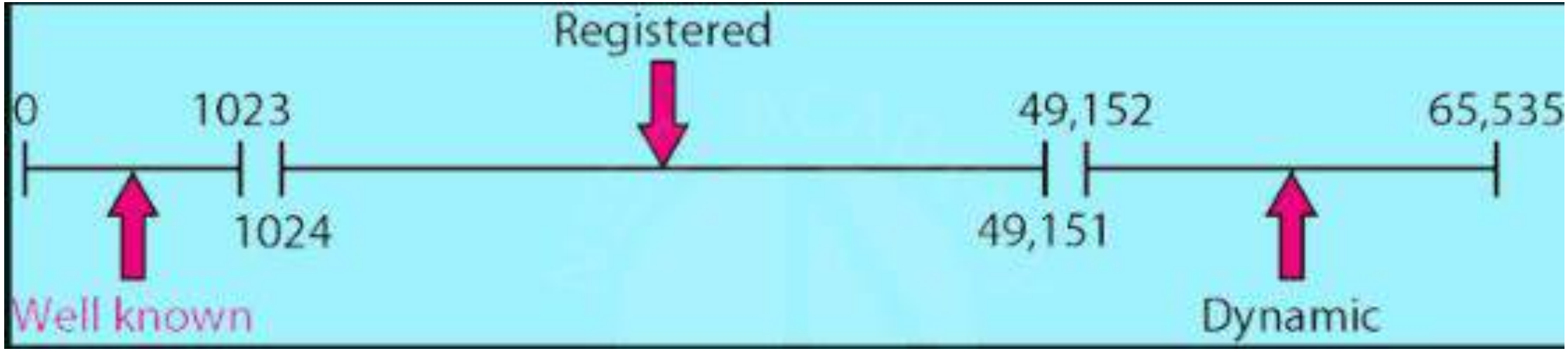

Daytime client
Daytime server
52,000
13
Transport layer
Transport layer
DP
SP
Data
13
52,000
13
52,000
Data


13
13
Port number
selects the process
193.14.26.7
IP header
193.14.26.7
IP address
selects the host
Transport layer
header
13


Processes
Processes
Multiplexer
Demultiplexer
IP
IP


Daytime client
Daytime server
52,000
13
Transport layer
Transport layer
Data
13
52,000
13
52,000
Data

Port Numbers

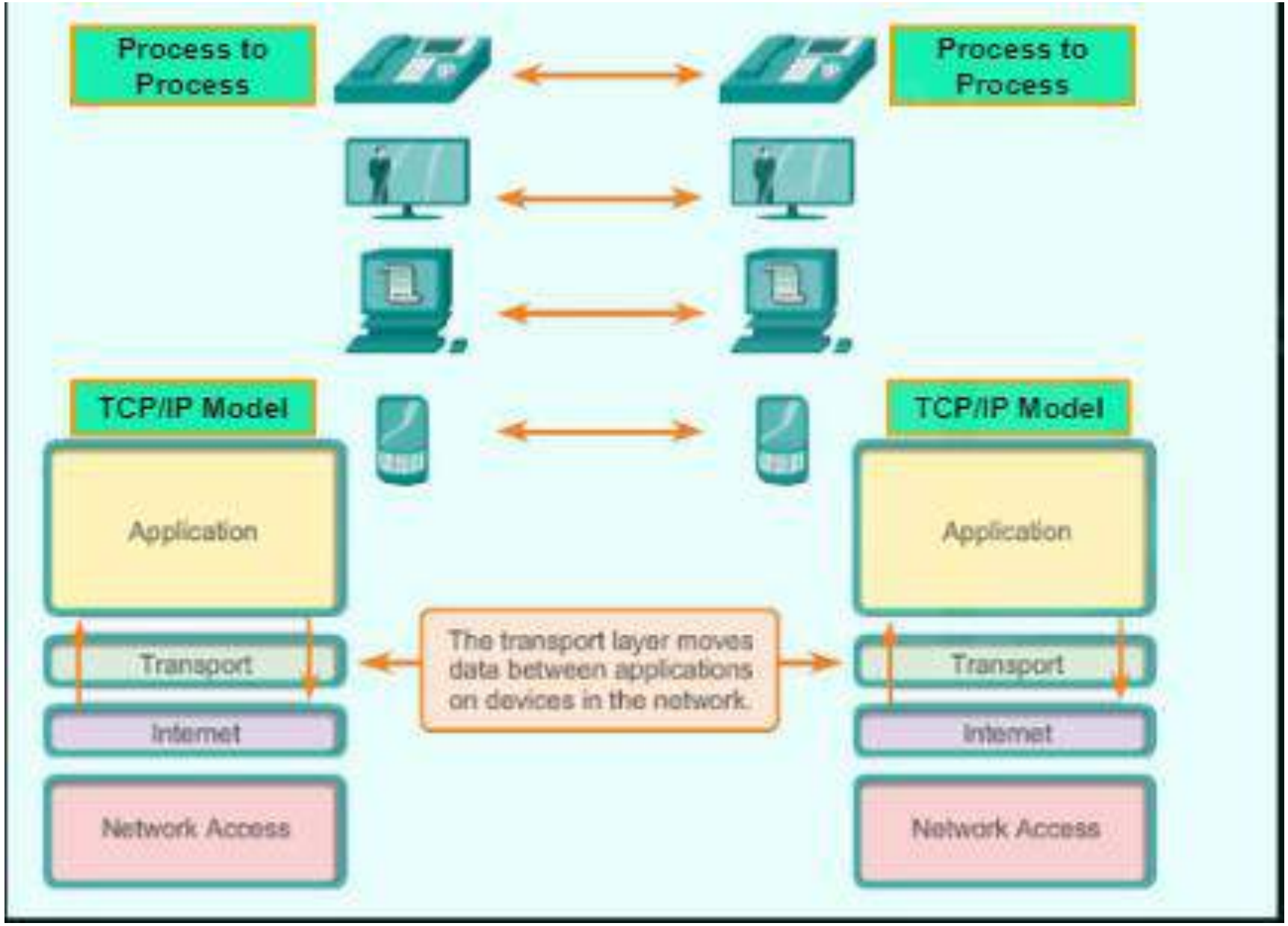


- This includes the registration of commonly used port numbers for well-known Internet services.

- The port numbers are divided into three categories
  - Well-known ports
  - Registration ports
  - Dynamic or private ports

Well Known Ports

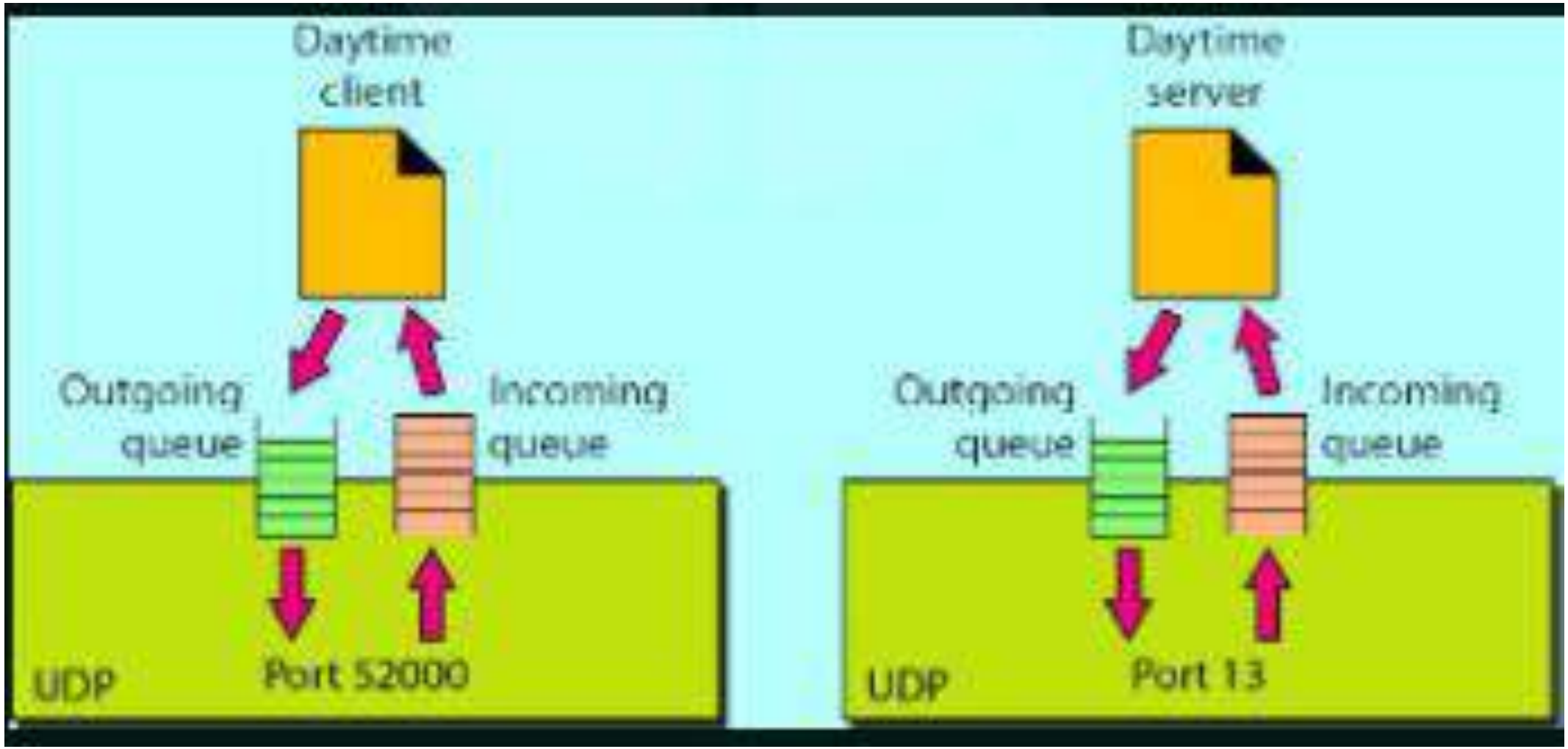
Ports Numbers	Assignments
20	File Transfer Protocol (FTP) Data Transfer
21	File Transfer Protocol (FTP) Command Control
22	Secure Shell (SSH) Secure Login
23	Telnet remote login service, unencrypted text messages
25	Simple Mail Transfer Protocol (SMTP) E-mail routing
53	Domain Name System (DNS) Service
67,68	Dynamic Host Configuration Protocol
80	Hypertext Transfer Protocol (HTTP) used in world wide web
110	Post Office Protocol (POP3)
123	Network Time Protocol (NTP)
143	Internet Message Access protocol (IMAP) Management of Digital mail
161	Simple Network Management Protocol (SNMP)
443	HTTP over SSL/ TLS (or) HTTP Secure


Process to Process
Process to Process
TCP/IP Model
TCP/IP Model
Application
Application
Transport
Transport
Internet
Internet
Network Access
Network Access
The transport layer moves data between applications on devices in the network.

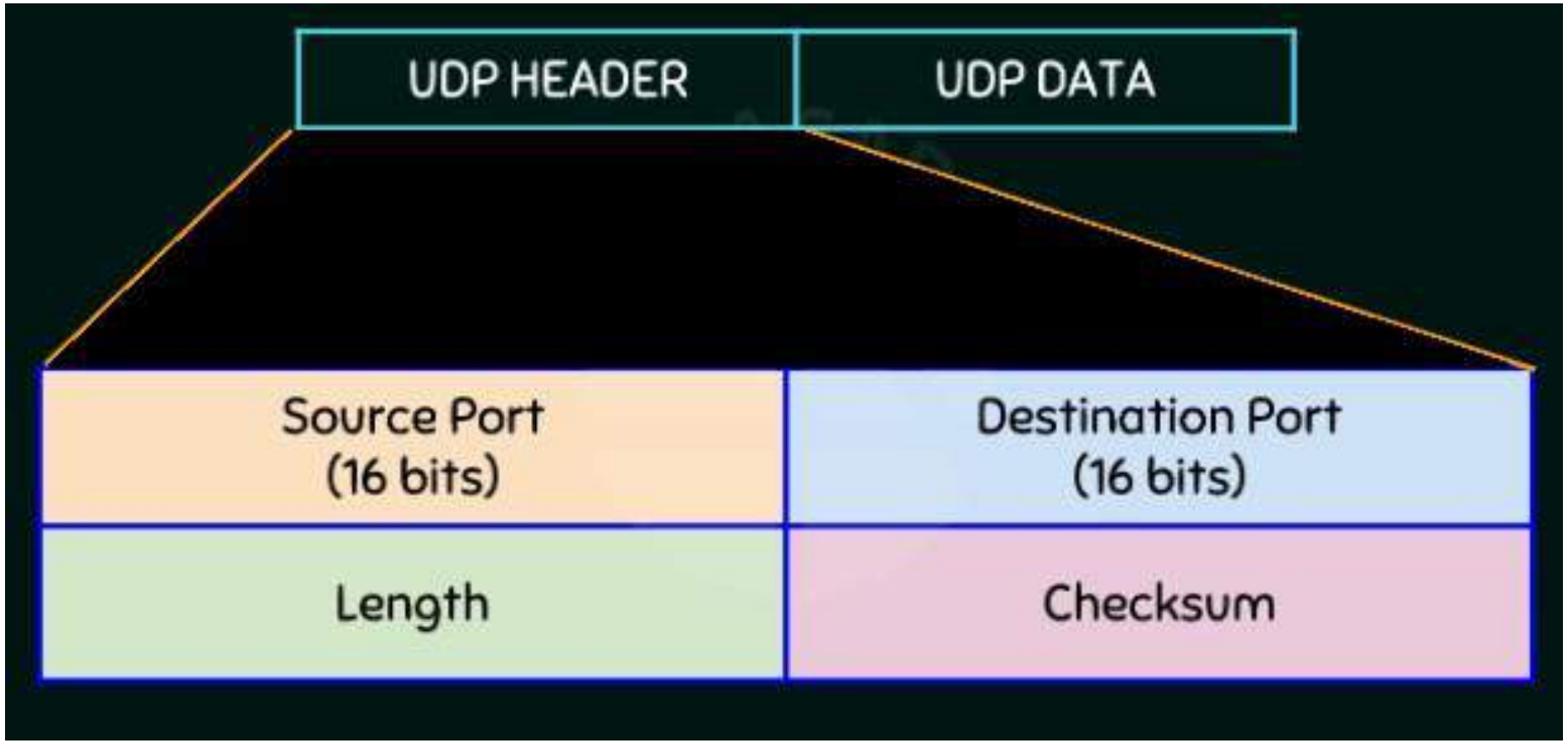
TCP (Transmission Control Protocol) vs UDP (User Datagram protocol))

TCP	UDP
➢ No loss but delay is allowed ➢ Retransmission is possible. ➢ Application uses TCP as transport layer protocol ❑ FTP ❑ SMTP ❑ HTTP ❑ Telnet	➢ No delay but low loss is allowed ➢ Retransmission is not possible ➢ Application uses UDP as transport layer protocol ❑ DNS ❑ DHCP ❑ SNMP ❑ TFTP ❑ VoIP

UDP


Daytime client
Outgoing queue
Incoming queue
UDP
Port 52000
Daytime server
Outgoing queue
Incoming queue
UDP
Port 13

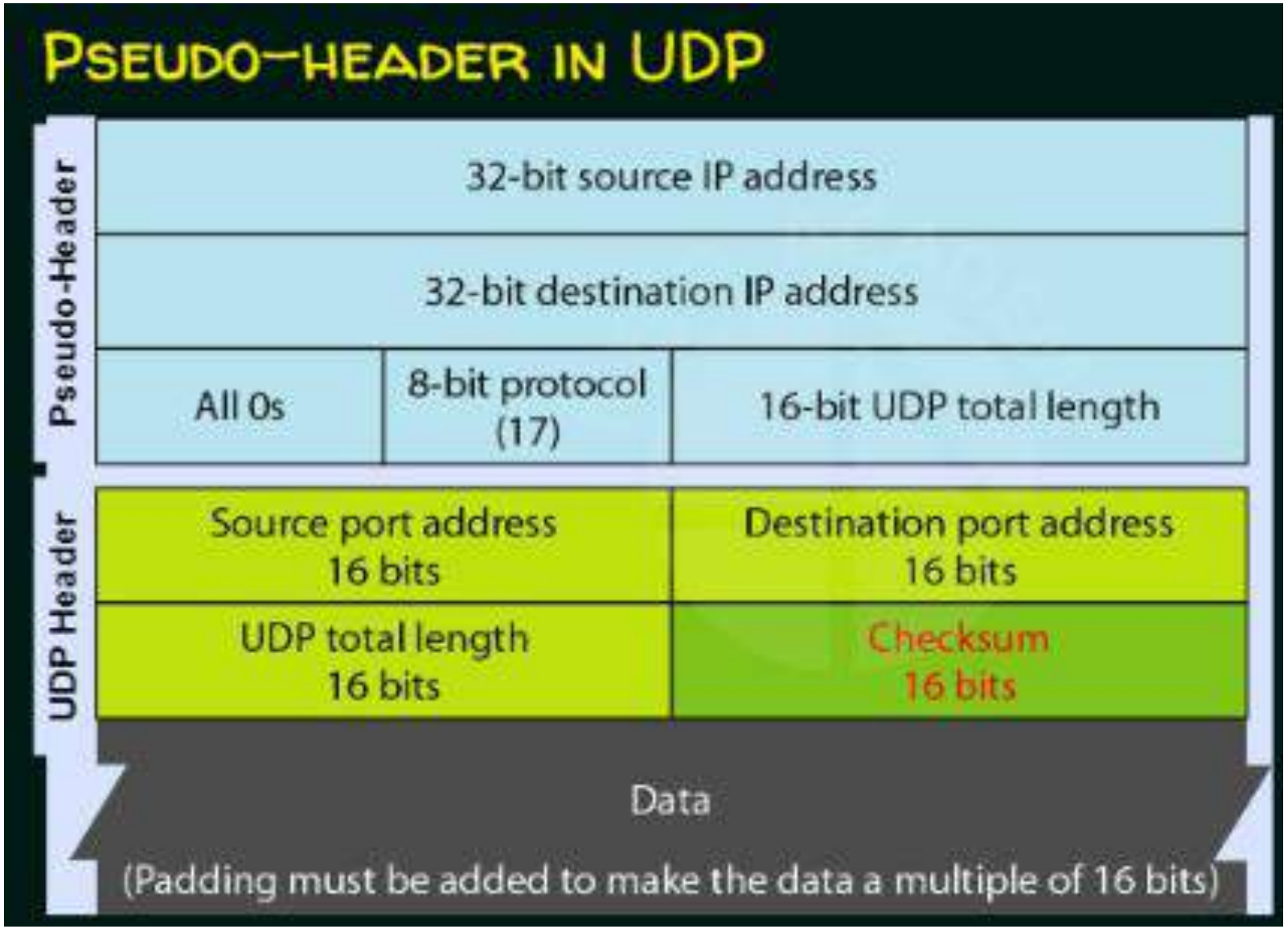
UDP (User Datagram protocol) Header Format



The following is a dump of a UDP header in hexadecimal format
0632000D001CE217

a. **What is the source port number?**
b. **What is the destination port number?**
c. **What is the length of the data?**
d. **What is the total length of the datagram?**
e. **Is the packet directed from a client to a server or vice versa?**

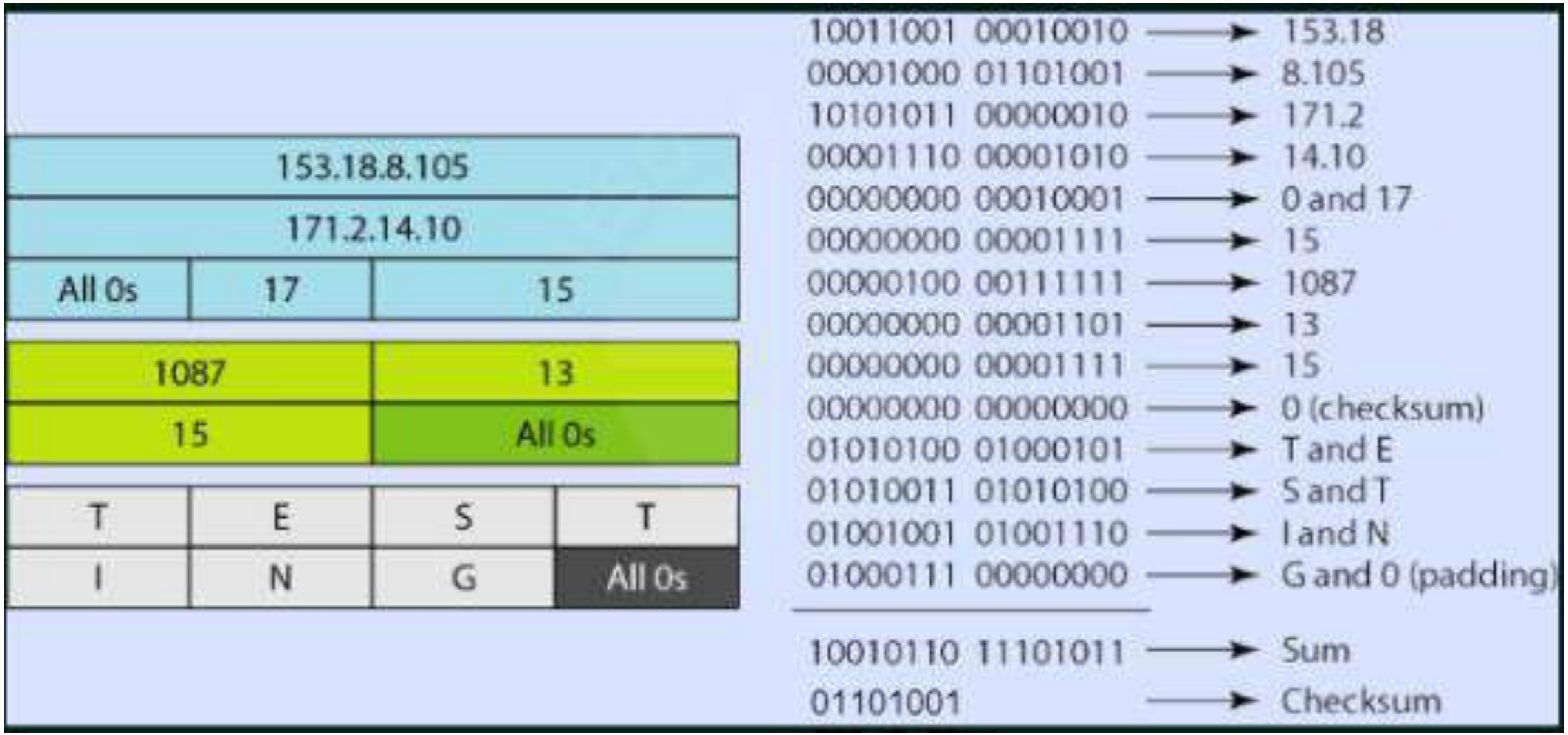
Pseudo Header in UDP (User Datagram protocol)



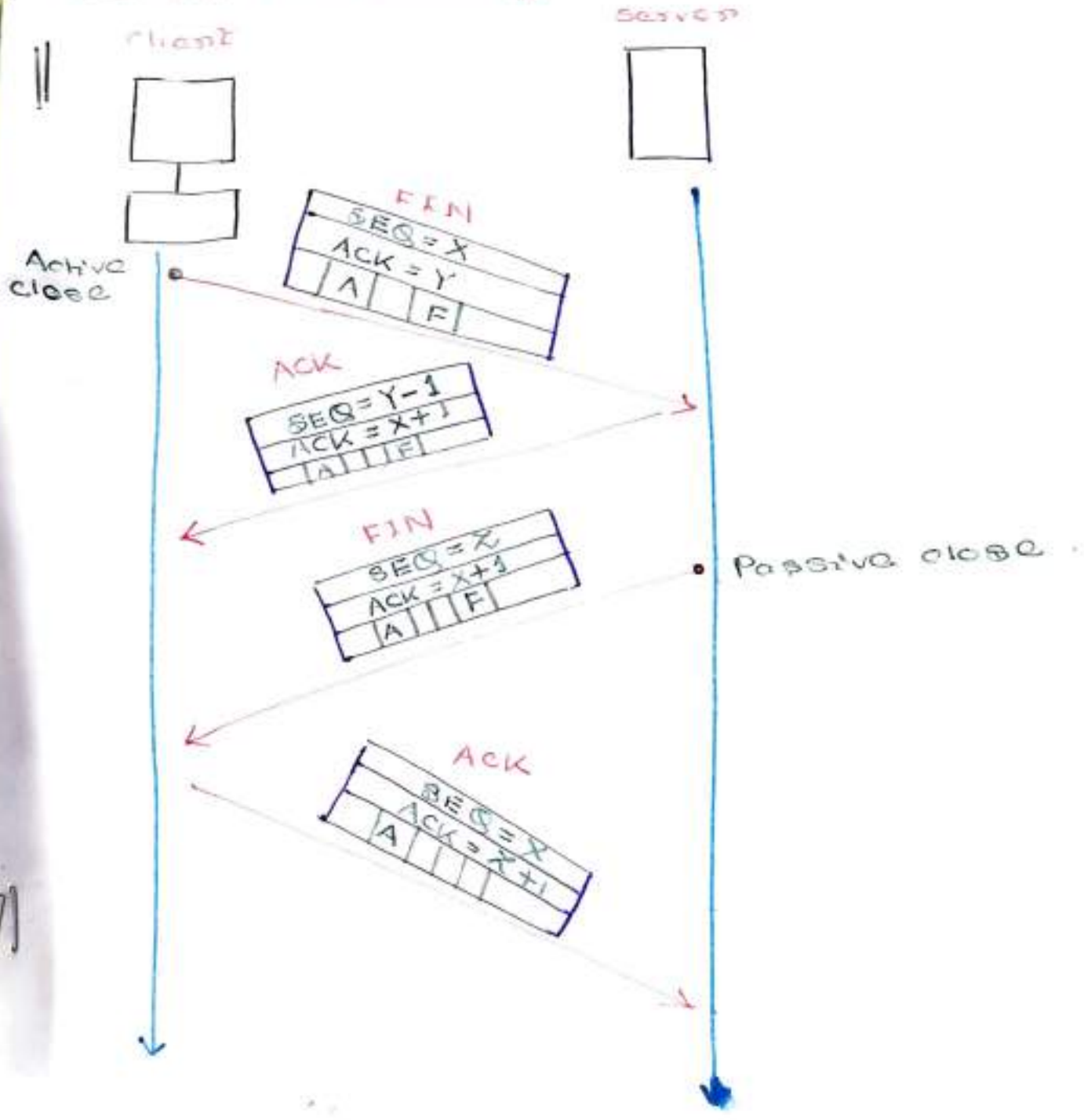
How to calculate Checksum in TCP Header?

- *Covert each pair composed of 16 bits into binary equivalent of the decimal. We are pairing in pair of 16 bits as the checksum value is of 16 bits.*
- *Same way find the ASCIII values of the characters, Suppose ASCII value of T is 84, now write the equivalent binary into 8 bits format.*
- *Add all binary 16 bits and if carry is generated do wrapping.*
- *Do 1's complement of the binary value resulted from the addition. S this will give you the checksum value.*

UDP (User Datagram protocol) Header Format



Four way handshaking:-



Sliding window:- window is used for flow control at the receiver.

Flow control is usually done in two layers:-

1. Transport layer.

2. Datalink layer.

Now we are in Transport layer:-

- TCP uses a sliding window to handle flow control.
- From Go-back-N
- does not use NAK.
- The receiver holds out-of-order segments

Selective Repeat

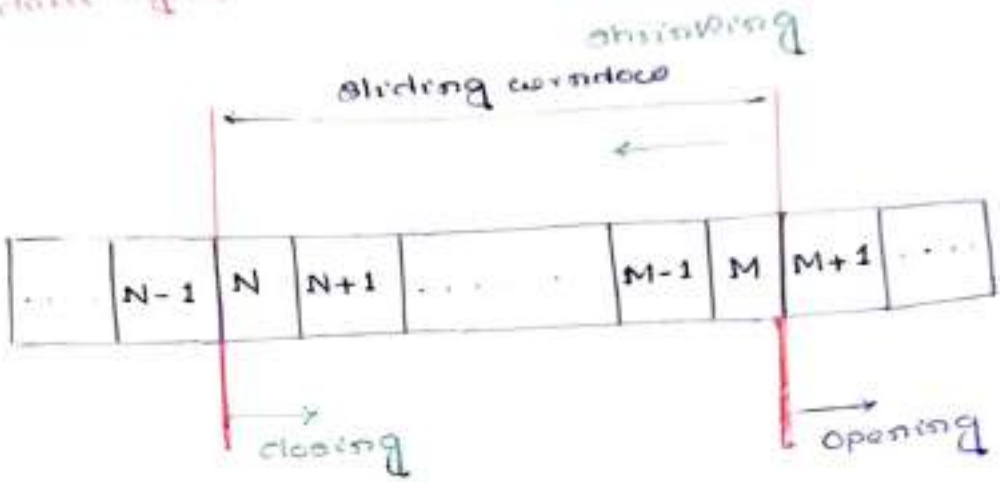
- TCP sliding window is of variable size. where as datalink layer sliding window is of fixed size.
- TCP sliding window is byte-oriented. whereas datalink layer sliding window is of frame oriented.

The window takes care about

① the data bytes which are sent

② the data bytes which are sent but not acknowledged

③ the data bytes which are yet to sent



window size = minimum (rwnd, cwnd)

rwnd → receiver window size

cwnd → congestion window size

The window size is of variable lenght.

The data which are sent but not have been acknowledge and the data which are to be sent immediately as will be in the window.

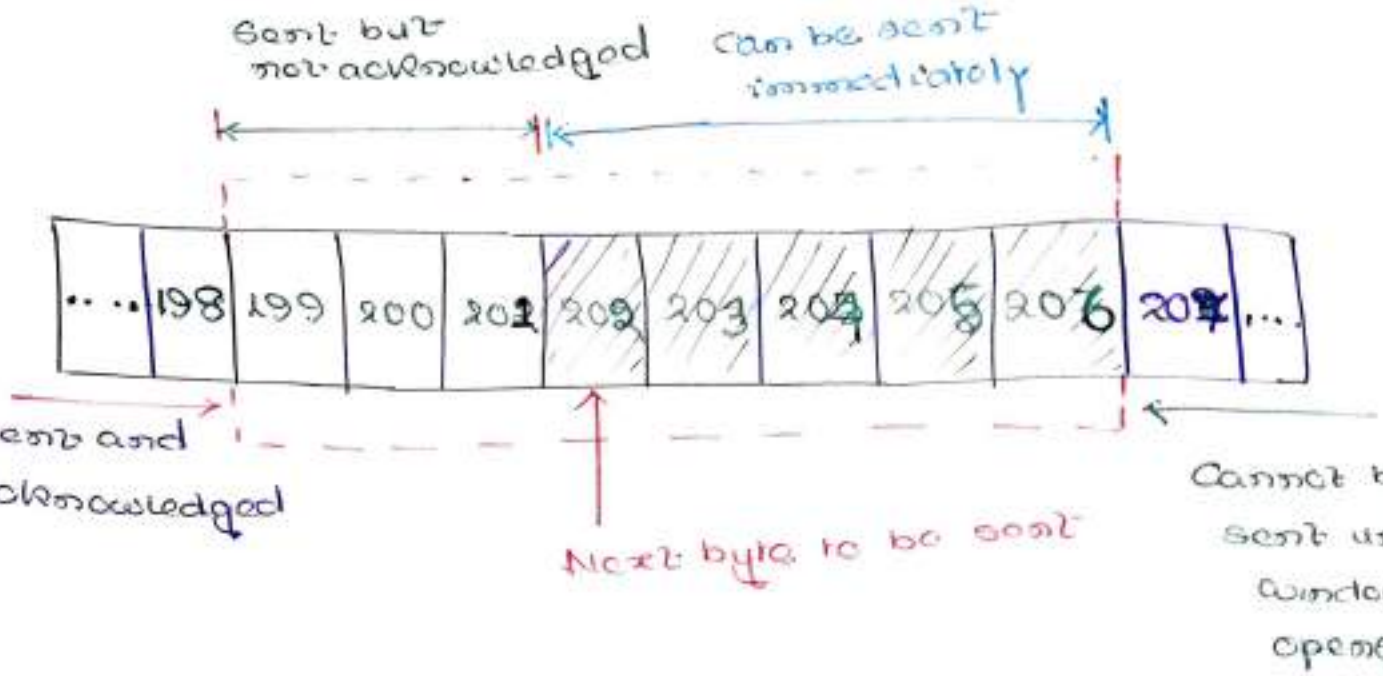
Suppose the N number segment is sent and get acknowledged now the window will move towards (N+1) and now the window can allow the (M+1) number segment to the the next segment to be sent in the queue.

• Congestion window is the ~~pre~~ feature of intermediary devices such as router or switches. suppose the receiver has a window size of 50. But the router through which many segments of other ~~ongoing~~ ongoing communications are processing, so there is a congestion.

window size = minimum (20, 8) = 8

receiver window size

congestion window size (size of network window)



- window size = min(rwnd, cwnd).
- The source does not have to send a full window's worth of data.
- The window can be opened or closed by the receiver but should not be shrunk.
- The destination/receiver can send an ACK at any time
- The sender can send 1-byte segment even after the window is shut down in the receiver side.

Q. What is the size of the window for host A if the value of rwnd (receiver window) is 3000 bytes and the value of cwnd (congestion window) is 3500 bytes?

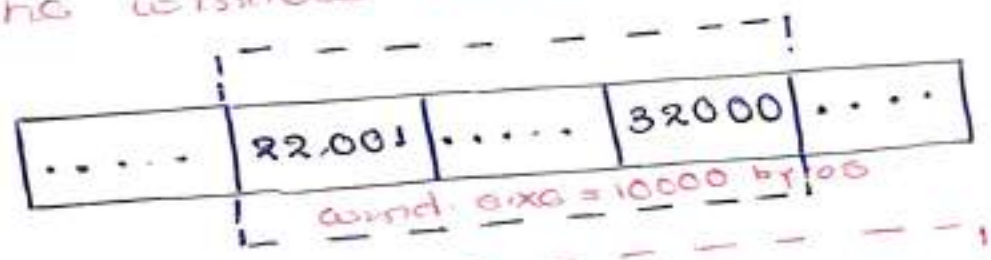
⇒ We know, window size = min (rwnd, cwnd)

= min (3000, 3500)

= 3000 bytes

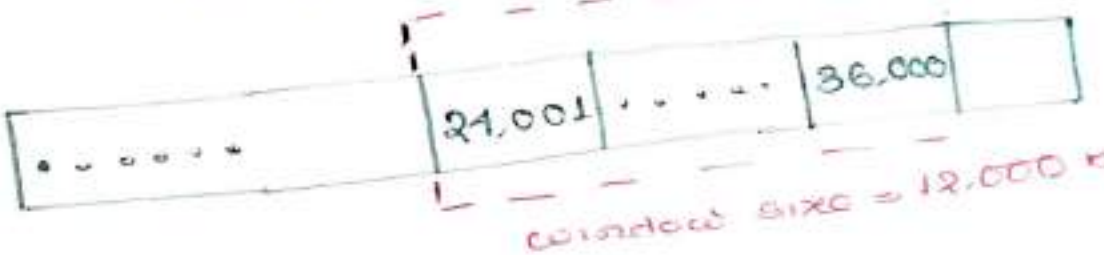
Q. A TCP connection is using a window size of 10,000 bytes and the previous acknowledgement number was 22,001. It receives a segment with acknowledgement number 21,001 and window size advertisement of 12000. Draw a diagram to show the situation of the window before and after.

$$\begin{array}{r} 22001 \\ +\ 10{,}000 \\ \hline 32{,}001 \\ -\ 1 \\ \hline 32{,}000 \end{array}$$



before receiving the segment

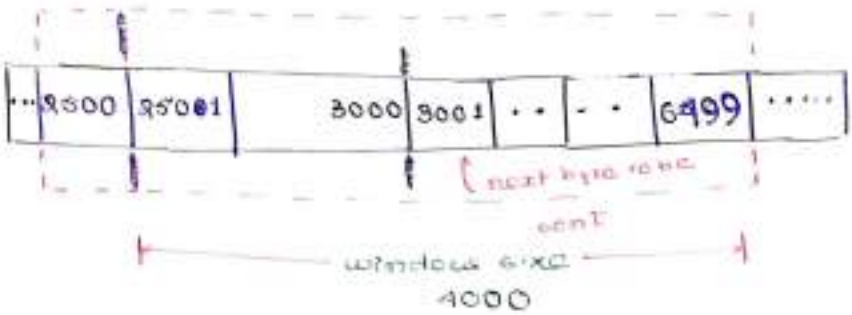
$$\begin{array}{r} 21{,}001 \\ +\ 12{,}000 \\ \hline 36{,}001 \\ -\ 1 \\ \hline 36{,}000 \end{array}$$



Q. A window holds bytes 2001 to 5000. The next byte to be sent is 3001. Draw a figure to show the situation of the window after the following two events:-

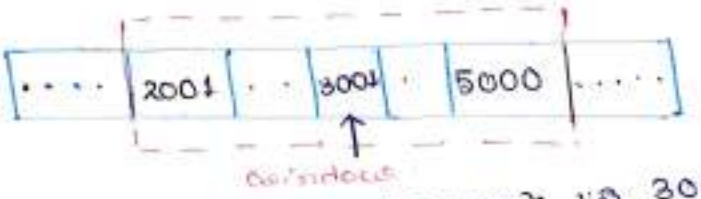
(a) An ACK segment with the acknowledgment number 2500 and window size advt. 1000 is received.

(b) A segment carrying 1000 byte is sent.



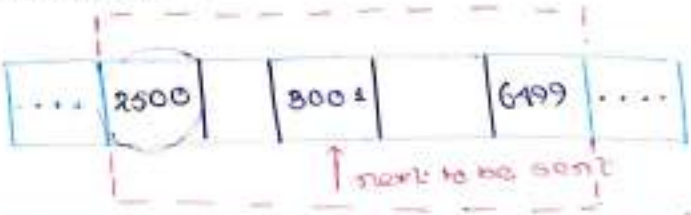
$$\begin{array}{r} 2501 \\ +\ 4000 \\ \hline 6501 \\ -\ 1 \\ \hline 6500 \end{array}$$

Initial figure:- The window holds byte 2001 to 5000



The next data byte to be sent is 3001

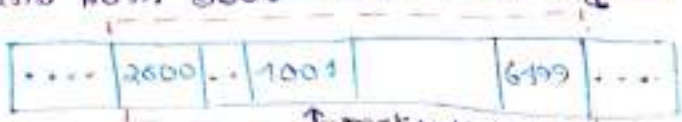
An acknowledgement for 2500 is received and the new window size is 4000.



$$\begin{array}{r} 2500\checkmark \\ +4000\checkmark \\ \hline 6500 \\ -\ 1 \\ \hline 6499 \end{array}$$

* A segment carrying 1000 byte is sent. means from 3001 next 1000 bytes is sent

$$\begin{array}{r} 3001 \\ +1000 \\ \hline 4000 \end{array}$$



From 3001 to 4000 next byte sent.

Transport layer

Application → In this layer application related data is generated.

Presentation → Here data is being compressed, encrypted

Session → Dialog exchange (Dialoge control activity)

Transport → It is responsible for process to process delivery of the entire message.

Network

Datalink

Physical

Suppose at the host multiple processes are ~~generated~~ running.
Transport layer responsibility is to identify uniquely each ~~host~~ process and delivers the generated data to the intended ~~host~~ process.
To do this each process is associated with a unique port number.

Network layer or Internet protocol layer identifies networks with these unique IP address.
Host-to-host data delivery is done by Data link layer.

MAC address — Hardware address of the devices

Services provided by Transport layer :-

1. Port addressing (Source and Destination port no. are added to the data generated by the application layer and carried by the transport layer)

2. Segmentation and Reassembly

3. Connection control (connection oriented and connectionless).

4. Flow control (speed control)

5. Error control activity

Two important Transport layer protocols are
(1) TCP & (2) UDP

Scanned by TapScanner

Transport layer

Application → In this layer application related data is generated.

Presentation → Here data is being compressed, encrypted

Session → Dialog exchange (Dialoge control activities)

Transport → It is responsible for process to process delivery of the entire message.

Network

Datalink

Physical

Suppose at the host multiple processes are running generated.
Transport layer responsibility is to identify uniquely each ~~hos~~ process and delivers the generated data to the intended ~~host~~ process.
To do this each process is associated with a unique port number.

Network layer or Internet protocol layer identifies networks with these unique IP addresses.
Host-to-host data delivery is done by Data link layer.

MAC address — Hardware address of the devices

Services provided by Transport layer:-

1. Port addressing (Source and Destination port no. are added to the data generated by the application layer and carried by the transport layer)

2. Segmentation and Reassembly

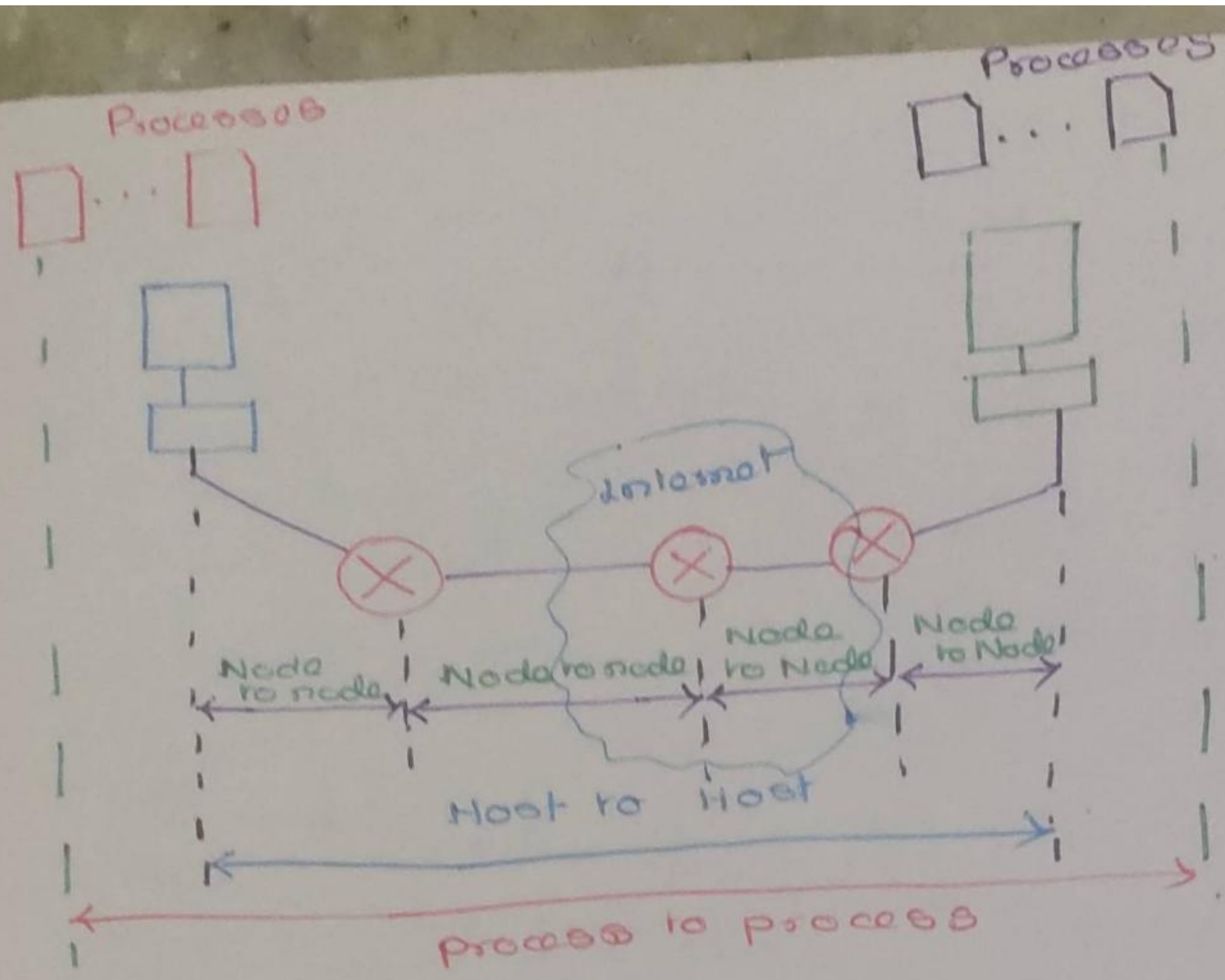
3. Connection control (connection oriented and connectionless).

4. Flow control (speed control)

5. Error control activity

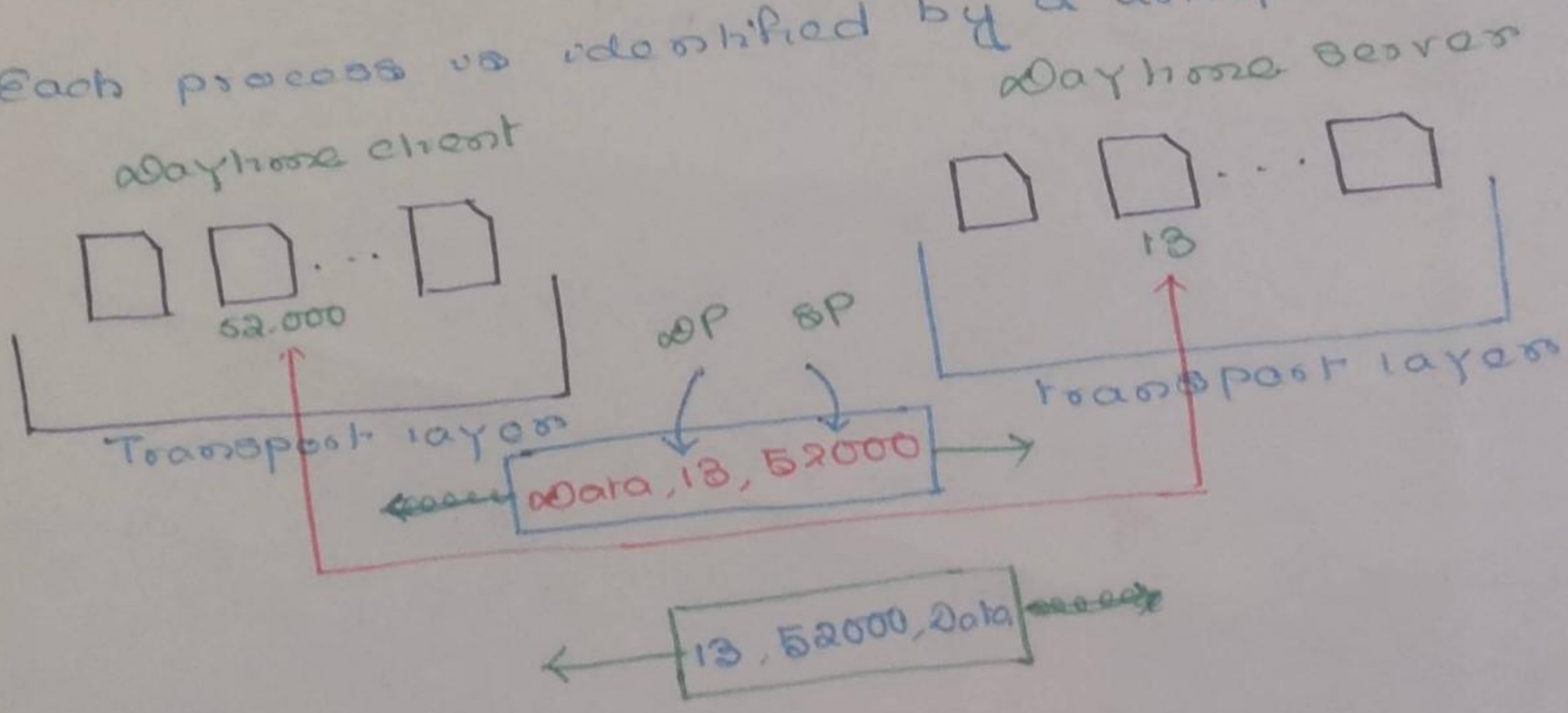
Two important Transport layer protocols are
(1) TCP & (2) UDP

Scanned by TapScanner



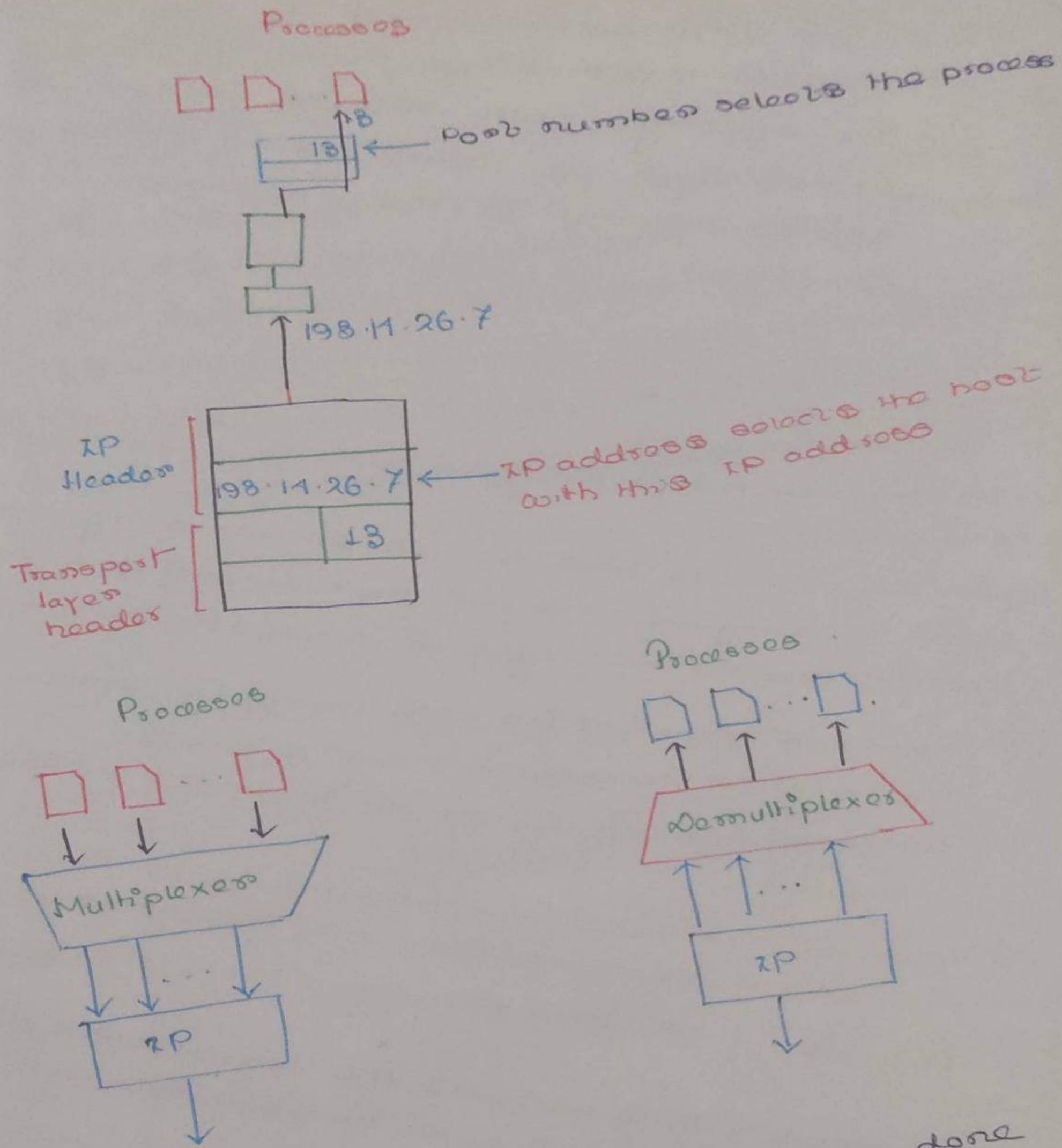
Node to Node : Data link layer
Host to Host : Network layer (IP address)
Process to Process : Transport layer (Port no)

Each process is identified by a unique port no.



DP = Destination Port NO
SP = source Port NO.

Scanned by TapScanner



Multiplexing and Demultiplexing is done at the transport layer. As many processes are running at a time in the application layer. But only one transport layer is present. So multiplexing is done and port addresses are embedded and the packets are transferred to the next layer that is network layer and at the receiver side after receiving the network layer datagram the network layer header will be removed and

The packet will be transferred to the transport layer. After demultiplexing the message with specific port address will be delivered to the specific process at the receiver device.

Port Numbers:-

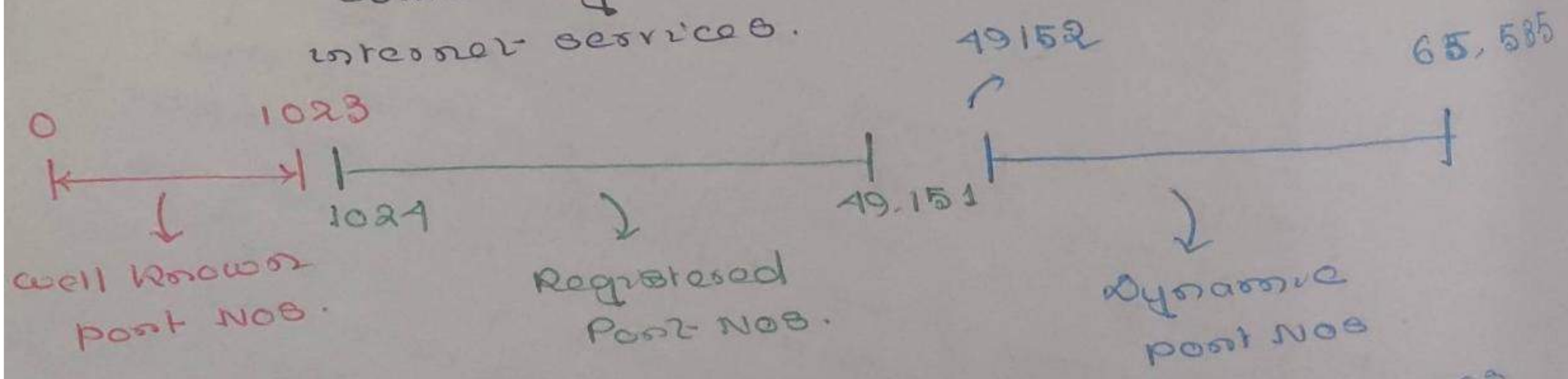
(1) In computer networking, a port is a communication end point.

(2) At the software level, within an operating system a port is a logical construct that identifies a specific process or a type of network service.

(3) A port is identified for each transport protocol and address combination by a 16-bit unsigned number, known as the port number
positive

(4) The most common transport protocols that use port number are the TCP (Transmission Control Protocol) and UDP (User Datagram Protocol).

Port numbers :- This includes the registration of commonly used port number for well-known internet services.



The port numbers are divided into three ranges

1. Well known
2. Registered
3. Dynamic or Private ports.

Scanned by TapScanner

2. Registered Port Numbers

The registered ports are those from 1024 through 49,151.

* IANA (Internet Assigned Numbering Authority) maintains the official list of well-known and registered ranges.

* IANA is responsible for the global coordination of the DNS root, IP addressing, and other internet protocol resources.

* This includes the registration of commonly used port numbers for well-known internet services.

3. Dynamic Port Numbers :- (Private port numbers)

* ranges from 49,152 to 65,535 (i.e. 216).

* Assigned by the operating system dynamically (suppose we are browsing www.google.com from our pc/laptop then this will assigned a dynamic port no to this process. As the destination port nos are very important so well-known port nos are assigned to the destinations)

*

Scanned by TapScanner

port numbers in URLs

* Port numbers are sometimes seen in web or other uniform resource locators (URLs).
* By default, HTTP uses port 80 and HTTPs uses port 443.
* URL - http://www.example.com:8080/path/ specifies the web browser to connect to port 8080 of the HTTP server.

In application layer many processes are running simultaneously. For process-to-process delivery of data we have seen port number are very important. When the application layer data is passed to transport layer, the transport layer headers are appended to the data. It may be TCP (Transmission control Protocol) and UDP (User Datagram Protocol) header depending on the information content.

Some information communication can allow some delay but cannot compromise on the reliability (i.e. data losses) and some information communication can allow some losses on the otherhand cannot allow any delay.

Data link layer - communication is done using MAC addresses i.e hardware address.

Network layer/Internet layer - communication is done using IP address or Internet protocol addresses these are often named as software addresses.

Scanned by TapScanner

Transport layer:- communication is done using port address of both the source process and the destination process.

The transport layer moves data between applications on devices in the network

~~Application layer:-~~

Application layer → Is the direct interface betn users and data are generated & retrieve at this layer

Two important protocols

- TCP (Reliable)

~~UDP~~

- No loss but delay
- e.g ⇒ Email
- Some delay can be allowed by no loss in information is allowed
- Retransmission is possible.

(1) HTTP (Website/webbrowsers)
(2) FTP
(3) SMTP
(4) Telnet (Remote login)

- UDP (unreliable)

- No delay but low loss
- Real time traffic
- Video ~~conferencing~~ conferencing, live call.
- Delay may introduce chaos in the communication and retransmission is also not possible.

(1) DHCP (Dynamic Host configuration Protocol)
(2) DNS
(3) SNMP
(4) TFTP
(5) VoIP

Scanned by TapScanner

UDP (User Datagram Protocol):-

- Simple protocol that provides the basic transport layer function.
- used by applications that can ~~tolerate~~ tolerate small loss of data (as no retransmission is possible)
- used by applications that cannot tolerate delay

used by :- (1) DNS
(2) VoIP
(3) ~~SMTP~~ SNMP (Simple Network Management Protocol).
(4) DHCP
(5) TFTP (Trivial File Transfer Protocol)
(6) Online games.

* connectionless and unreliable. → no connection establishment between source & destination
* Prior communications (or handshaking) are not required in order to setup communication channels or data paths.
* UDP-based servers applications are assigned well-known or registered port numbers.
* UDP client process randomly selects port numbers from range of dynamic port numbers (these are operating system assigned port numbers) as the source port
* UDP is suitable for purposes where error checking and correction are either not necessary or are performed in the applications. UDP avoids the overhead of such processing in the protocol stack.
* No error correction.

Scanned by TapScanner

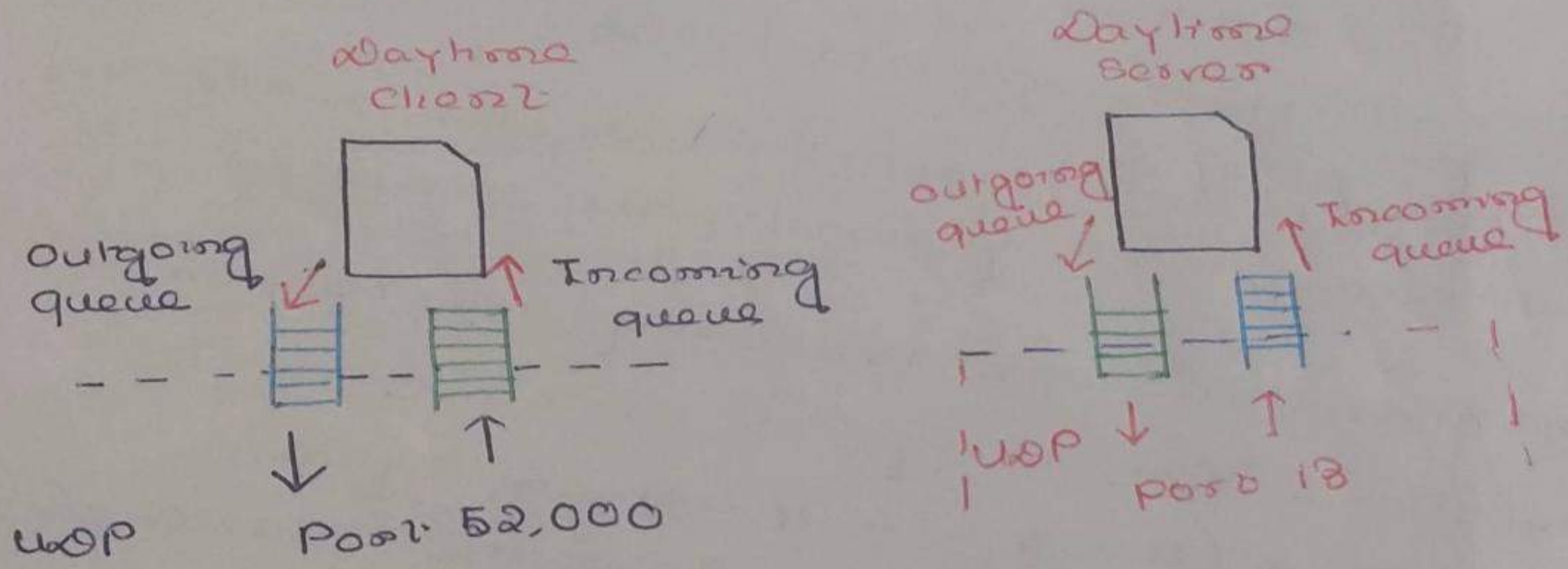
* Time-sensitive applications often use UDP because dropping packets is preferable than to waiting for packets due to retransmission which may not be an option in a real-time traffic on system.

* No flow control mechanism.

The UDP ~~Packets~~ headers are very simple and have only four fields. These UDP headers are kept simple because UDPs are used for real-time communication so when a router receives an UDP packet it should not take more ~~time~~ time to process it.

* Encapsulation and Decapsulation is done by the UDP. protocol.

* Queuing is done the UDP.



In application layer many processes are running simultaneously. If the client has two queue the server also has to maintain two queues. The UDP dequeue messages one after another an add UDP header and passes it to the next layer i.e. the network layer to add IP header. When the process are terminated at the application layer, the queues are also discarded.

Scanned by TapScanner

Uses of UDP :-

(1) It is a transaction-oriented, suitable for simple request-response protocols such as DNS (Domain Name System) or the NTP (Network Time Protocol).

(2) It provides datagrams, suitable for modeling other protocols such as IP tunneling or remote procedure call and the Network File system.

(3) It is simple, suitable for bootstraping or other purposes without a full protocol stack, such as the DHCP and TFTP.

(4) It is stateless, suitable for very large numbers of clients, such as in streaming media application such as IPTV. (because of simplied header).

(5) The lack of retransmission delays makes it suitable for real-time applications such as VoIP, online games and many real-time streaming applications.

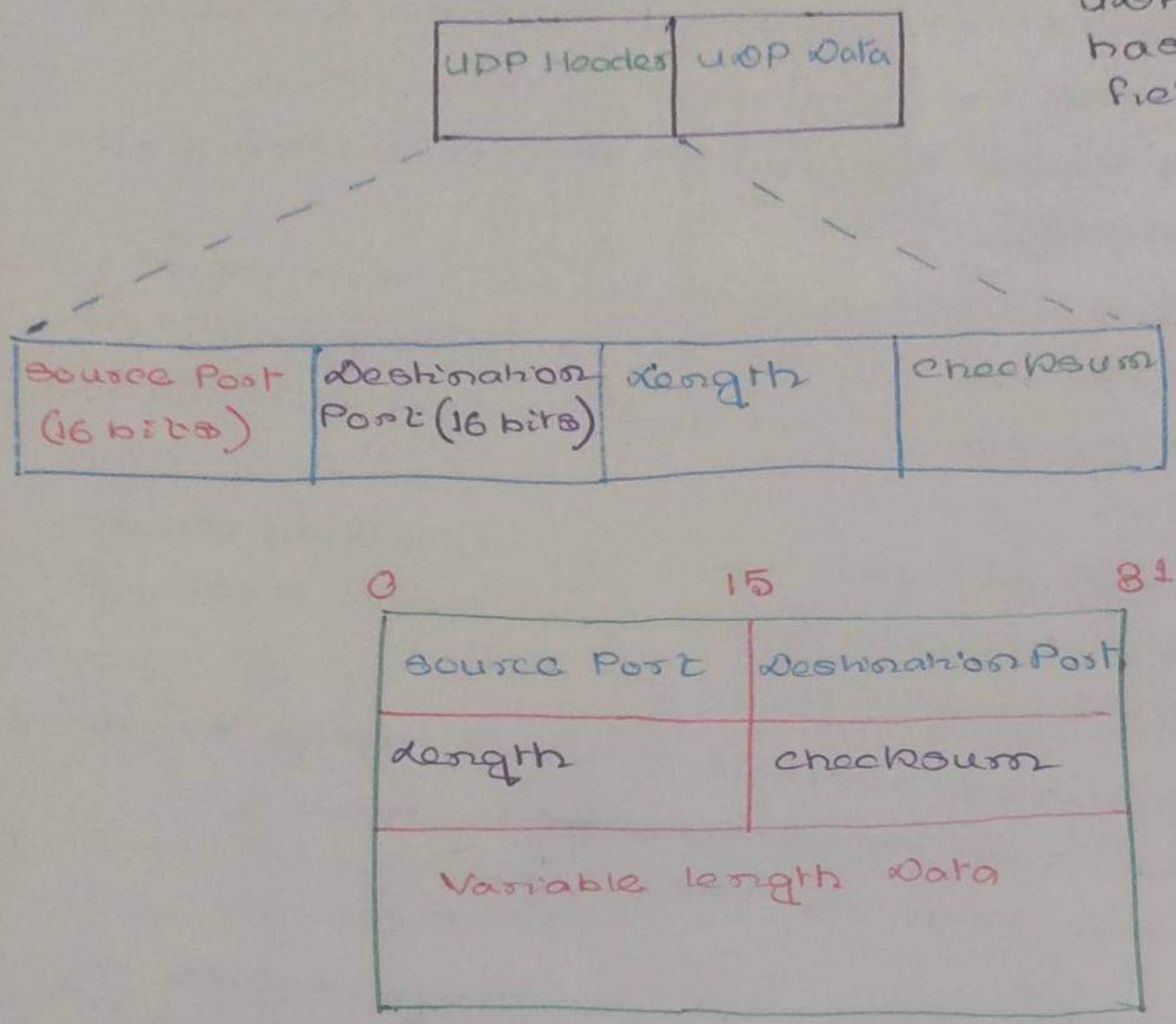
(6) UDP is a suitable transport protocol for multicasting. Multicasting capability is embedded in the UDP software but not in the TCP software

(7) UDP is used for some route updating protocols as routing information protocol (RIP)

Scanned by TapScanner

UDP Header Format :-

UDP Header has only Four fields



1024

All the four field are of 16 bits length.

Source port Number :-

- This is the port number used by the process running on the source host. usually this port number is assiged by the operating system.
- It is a 16-bits long. which means that the port number can range from 0 to $(2^{16}-1)$ i.e. 65.535.
- If the source host is the client (a client sending a request), the port number is most cases is an ephemeral port number i.e. random port number ~~greate~~ greater than 1024 (well known and registered) requested by the process and choosen by the UDP software running on the source host.

Scanned by TapScanner

- If the source host is the server (a server sending a response), the port number in most cases is a well-known port number

Destination Port number:-

- This is the port number used by the process running on the destination host
- It is also 16 bits long. If the destination host is the server (a client sending a request) the port number, in most cases, is a well-known port number.
- If the destination host is the client (a server sending a ~~request~~ response), the port number in most cases, is an ephemeral (i.e. random) port number
- In this case, the server copies the ephemeral port number it has received in the request packet.

Length Field:-

- This is a 16-bit field that defines the total length of the user datagram, header plus data. The 16 bits can define a total length of 0 to $(2^{16}-1)$ i.e. 65,535 bytes.
- However the total length of the user datagram need to be much less because a UDP user datagram is again stored in an IP datagram with a total length of 65,535 bytes.
- UDP length = IP length − IP header's length.

Scanned by TapScanner

UDP checksum

Pseudo-header in UDP:-

- The UDP checksum calculation is different from one for IP and ICMP.
- In UDP, the checksum includes three sections:
  - A Pseudo-header
  - The UDP header and
  - The data (coming from the application layer).

* The pseudo-header is the part of the header of the IP packet in which the user datagram is to be encapsulated with some fields filled with 0's (zeroes).

* If the checksum does not include the pseudo-header, a user datagram may arrive at destination with out errors.

* However, if the IP header is corrupted, it may be delivered to the wrong destination.

The data generated by the top-most layer that is the application layer pass to the transport layer where the UDP header is encapsulated with the data. Now in the network layer along with the IP header a header field containing 0's is

Scanned by TapScanner

added at the IPV4 field filled with 0's.
which is called as Pseudo-header.

- The protocal field is added to ensure that the packet belongs to UDP. and not to other transport layer protocols.
- The value of the protocol field for UDP is 17.
- If the value is changed during transmission the checksum calculation at the receiver will detect it and UDP drops the packet.
- Thus, the packet is not delivered to the wrong destination.

32-bit source IP address		
32-bit destination IP address		
All 0's (8 bits)	8-bit protocol (17)	16-bit UDP total lenght

Pseudo-Header

Source Port Address (16 bits)	Destination Port Address (16 bits)
UDP total lenght (16 bits)	Checksum 16 bits
Variable lenght data.	

UDP Header

Scanned by TapScanner

TCP (Transmission control Protocol) :-

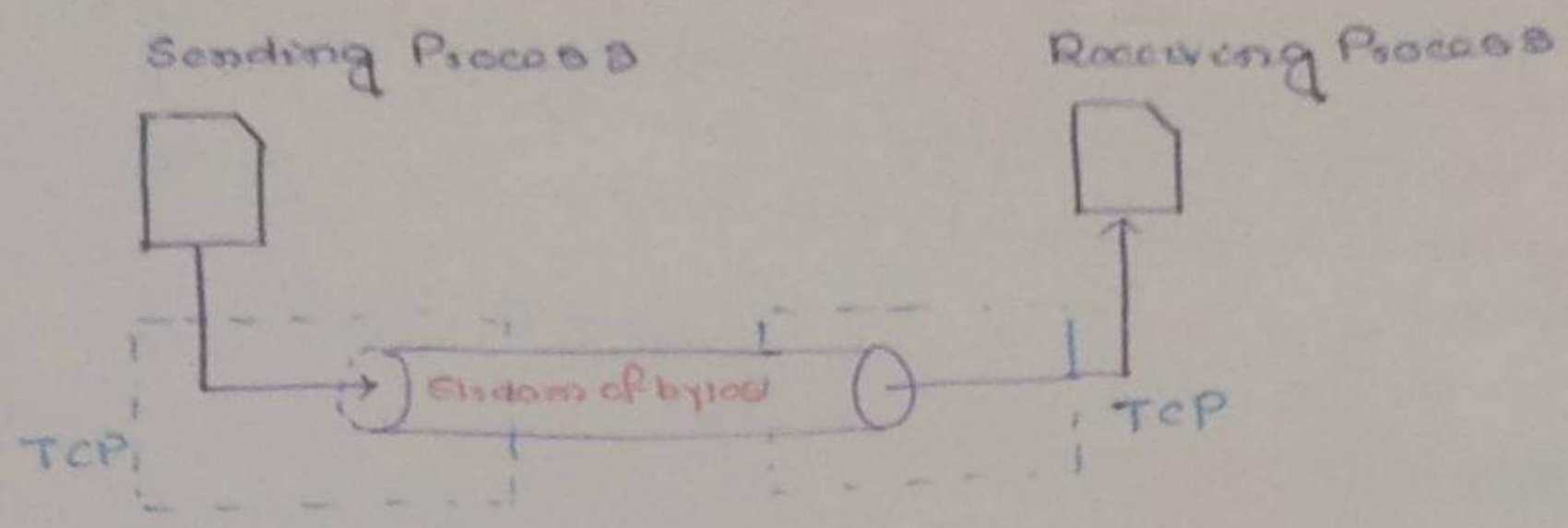
- TCP is the widely used Transport layer protocol.
- It creates a virtual connection between two TCPs to send data.
- It creates addition. TCP uses flow and error control mechanisms at the transport level
- Used by application that can tolerate delay but cannot tolerate loss.
- used by ~~proto~~ application such as:
  1. HTTP
  2. FTP
  3. Telnet
  4. SMTP
  5. world wide web (www)

Features of TCP

- TCP is a connection oriented protocol connection is established using three way handshaking before starting the communication
- Reliable delivery (sender will know receiver has received the data or retransmission will be done)
- Acknowledgment oriented.
- Retransmission.
- Flow control (Speed control mechanism)
- Error control
- Congestion control
- Segmentation and Reassembly
- Full Duplex Support.

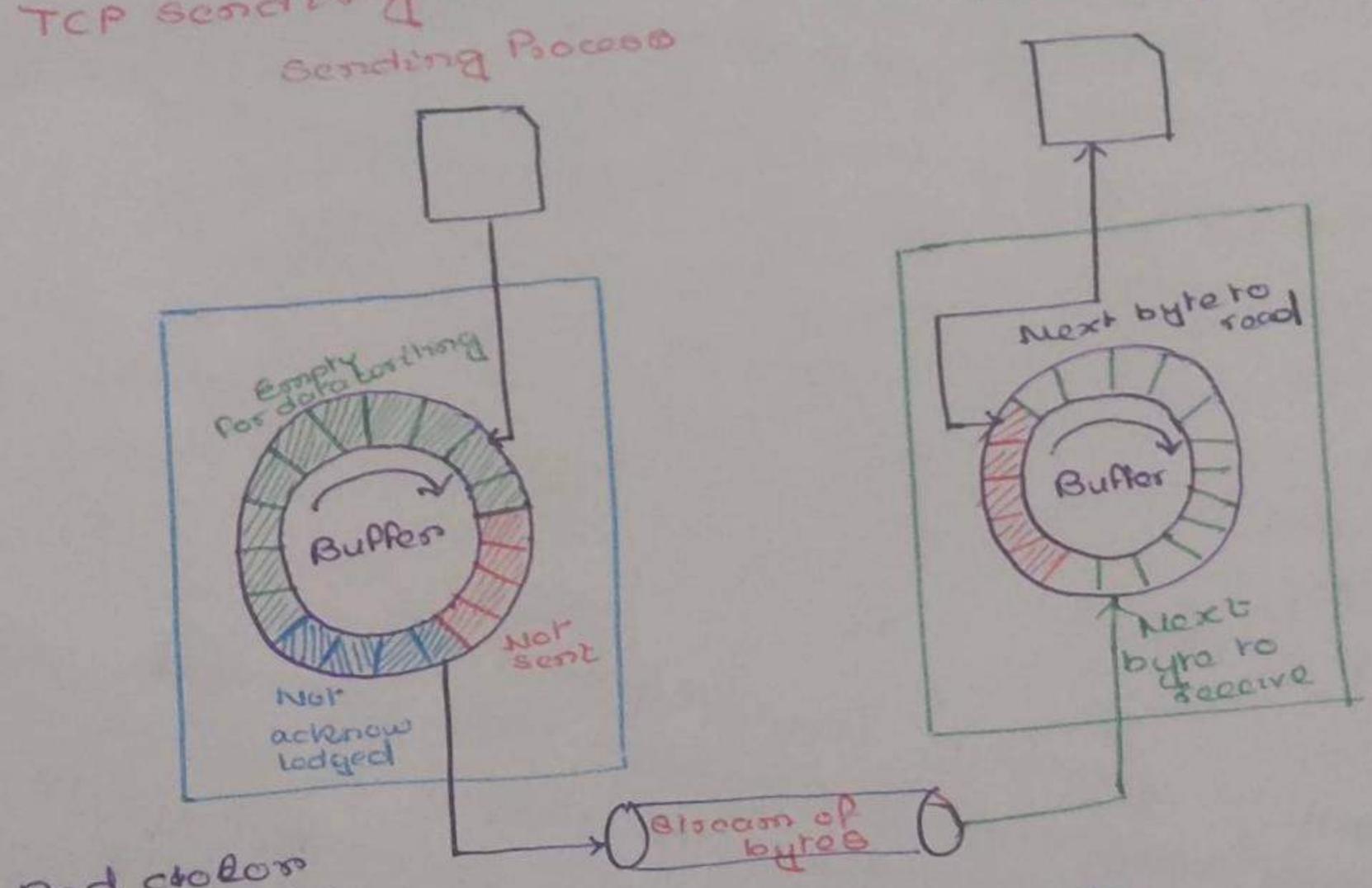
Scanned by TapScanner

TCP stream Delivery :-



This sending and receiving process are running in two different system (e.g. computer/laptop/ cellphones). This imaginary tube will deliver the data in form of stream of bytes.

TCP sending and Receiving Buffer :-



- Here Red color indicated data bytes which are not sent.
- Blue cells indicates data sent by not acknowledged
- Green indicated empty buffers when data will be stored
- After acknowledgment received the ~~chan~~ chambers will be recycled.

Scanned by TapScanner

TCP segments:-

Each segment can of of different lengths. Each segment carries hundred or thousand of bytes. The TCP adds TCP header to each of the segment for control and error detection activities.

Application layer generates data

↓

This data is passed to the transport layer The transport layer add TCP header and make segments consists of ~~data stream~~ streams ~~bytes~~ of data bytes. Transport layer ~~user~~ datagrams are called segment

↓

The segments will be passed to the network layer. And in this layer IP header will be embeded to the segments. This network layer datagrams are called packets. If TCP headers are added then it is called segment. If UDP headers are added

The IP packets are then sent to the data link layer for transmission. In this layer the communication link can be wired or wireless. The data link layer will add their header to the IP packets received from the network layer, and will make a frame.

↓

This will be pass to the physical layer and this layer will convert the received frames from the data link layer to electromagnetic signals or optical signals. The physical layer protocol data units are called bits.

Scanned by TapScanner

TCP header Format :-

Source Port Address 16 bits								Destination Port Address (16 bits)
Sequence No (32 bits)								
Acknowledgement No (32 bits								
HLEN 4 bits	RES 6 bits	URG	ACK	PSH	RST	SYN	FIN	window size (16 bits).
Checksum (16 bits)								Urgent Pointer (16 bits)
Options and Padding								

if options are added than 60 bytes (minimum Header)

The TCP header lenght is 20 bytes. The TCP header plus options and padding length can go upto 60 bytes. i.e. The options and padding lenght can maximum go upto (60-20) = 40 bytes.

We have seen that the UDP has only 4-fields. where as TCP/IP has more no. of fields. As TCP provides a reliable and acknowledgement oriented communication with error control and flow control mechanism.

Source Port Number :-

(1) It defines the port number of the application program in the host device that is sending the segment. We have seen in a client host many processes are running simultaneously. The process which is sending/requesting the data, this port number is associated with the process.

Scanned by TapScanner

(2) This serves the same purpose as the source port address in the UDP header.

Destination Port address :-

(1) It defines the port number of the application program in the host that is receiving the segment. (well known port numbers).

(2) This serves the same purpose as the destination port address in the UDP header

Sequence Number :-

This is a 32-bit field that defines the number assigned to the first byte of data contained in the segment. This sequence number helps us to know about missing packets, lost packet, out of order and delayed packets

Has a dual role :-

- If the SYN flag is set to 1, then this is the initial sequence number. The sequence number of the actual first data byte and the acknowledged number in the corresponding ACK are then this sequence number plus 1.

Suppose the sequence number of the first data byte is 1001 then the acknowledgment (ACK) number sent by the receiver is 1002 i.e. the receiver is expecting the second data byte

- If the SYN flag is 0, then this is the accumulated sequence no. of the first data byte of this segment for the current session.

Scanned by TapScanner

Acknowledgement number:-

(1) This is a 32 bit field.

(2) It defines the byte number that the receiver of the segment is expecting to receive from the other party.

(3) If the receiver of the segment has successfully received byte number x from the other party it defines (x+1) as the acknowledgement number. Acknowledgement and data can be piggybacked together.

HLEN (Header length):-

- It indicates the number 4-byte words in the TCP header.
- The length of header can be from 20 byte to 60 byte i.e. (60 - 20 = 40 bytes is options & padding)
- Therefore, the value of this field can be between 20 bytes i.e (5×4 = 20) 5 and 60 bytes i.e (15×4 = 60) 15. As the HLEN field contains 4-byte words.

RES (Reserved field)

- It is a 6-bit field.
- This is reserved for future use and usually filled with 0's.

Scanned by TapScanner

(6 Flag bits in TCP header :-

U	A	P	P	S	F
R	C	S	S	Y	I
G	K	H	T	N	N

G: Flags ~~bits~~ fields are consists of 6 different control bits or flags each of the 6 flags are of 1 bit. One or more of these bits can be set at a time

(2 URG (1 bit) :- when this is set i.e. 1 the urgent pointer field find its significance.

S ACK (1 bit) :- Indicates that the acknowledgement field is significant.

T PSH (1 bit) :- PUSH function. Ask to push the buffered data to the receiving application.

RST (1 bit) :- Reset the connection.

SYN (1 bit) :- Synchronize sequence numbers.

FIN (1 bit) : Last packet from sender.

Scanned by TapScanner

window size :-

- This field defines the size of the window in bytes, that the other party must maintain. Suppose the receiver window is of 20 byte receiver can handle 20 bytes at a time. But if the sender without knowing the receiver window size sends 50 bytes then it can only receives 20 bytes and other bytes will be discarded. So it is very important for the receiver to inform about its window size

This field defines the size of the wi.

- The length of this field is 16 bits, which means that the maximum size of the window is $(2^{16}-1) = 65,535$ bytes.

- This value is normally referred to as the receiving window (rwnd) and is determined by the receiver.

- The sender must obey the dictation of the receiver in this case.

checksum :-

- The calculation of the checksum for TCP follows the same procedure as the one described in UDP.

- However, the inclusion of the checksum in the UDP datagram is optional (i.e. if checksum is not added by the UDP it is the responsibility of application layer to calculate), whereas the inclusion of the checksum for TCP is mandatory.

- The same pseudo-header, serving the same purpose, is added to the segment. For the TCP pseudo-header, the value for the protocol field is 6. (for UDP it is 17)

Scanned by TapScanner

Urgent Pointer :-

① This 16-bit field, which is valid only if the urgent flag is set, is used when the segment contains urgent data.

We have seen the real-time traffic or urgent data is normally sent by the UDP transport protocol. But in some case we want a data to be sent urgently and also maintains its reliability, then this is used in TCP.

(2) It defines the number that must be added to the sequence number to obtain the number of the last urgent byte in the data section of the segment.

Options :-

This is field can be upto a maximum of 40 bytes. These can be upto 40 bytes of optional optional information in the TCP header.

- The length of this field is determined by the Data offset field. options have up to three fields :-

a. Option - Kind (1 byte)
b. Option - length (1 byte).
c. Option - data (Variable).

Scanned by TapScanner

Padding:-

The padding is composed of zeros. The TCP header padding is used to assure that the TCP header ends, and data begins, on a 32-bit boundary.

TCP connection:-

1. connection-oriented
2. virtual path (seq + acknowledgement).
3. Acknowledgment process
4. Retransmission of lost damaged frames
5. TCP connection- virtual not physical
6. IP- connectionless
7. Full-duplex mode
8. Approval from other party

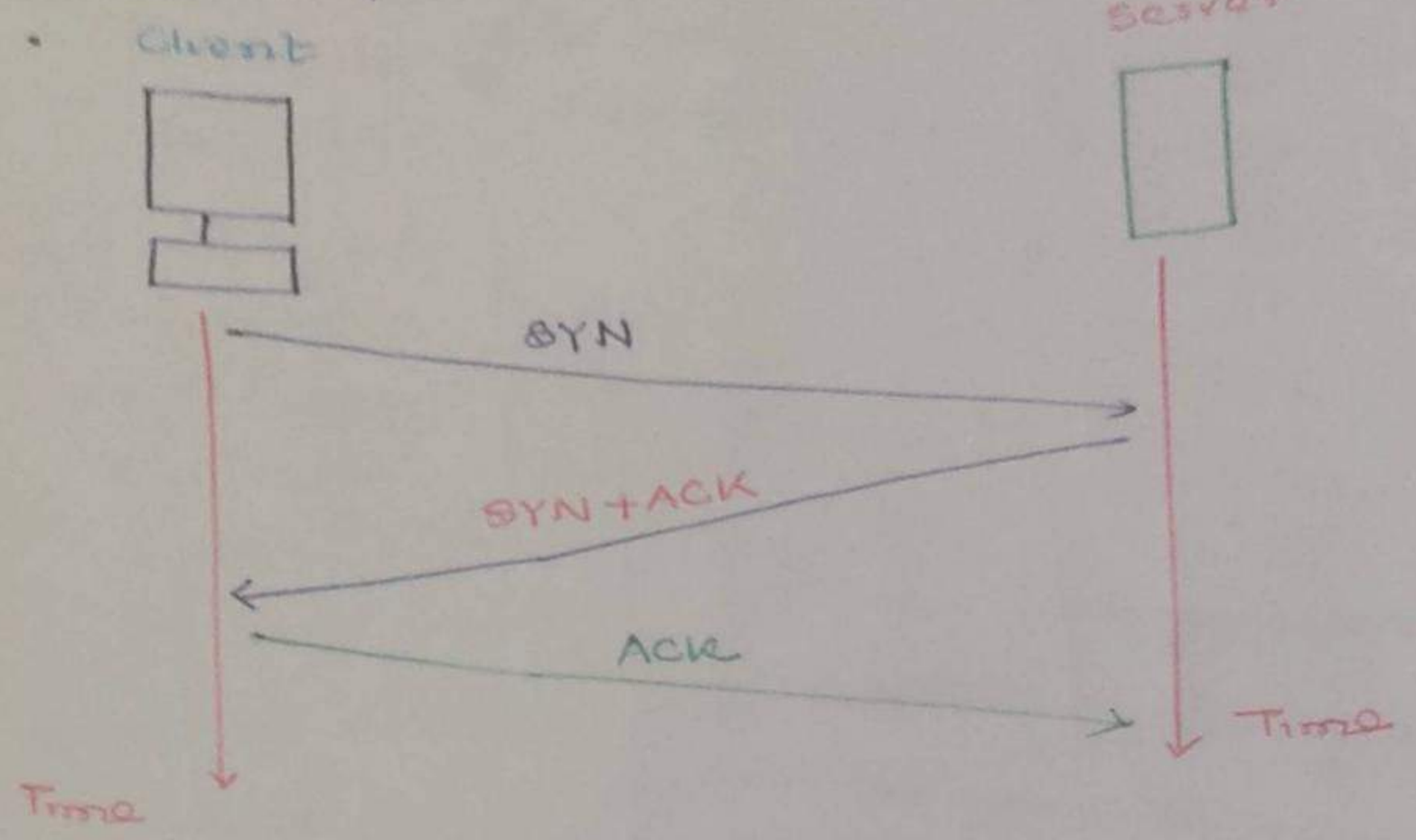
Three phases of TCP connection.

1. Connection Establishment
2. Data transfer
3. Connection Termination.

Scanned by TapScanner

TCP connection Establishment:-

Three-way-handshaking mechanism:-



Three segments are exchanged thats why it is called Three-way Hand shaking

- client will initiate the connection. ~~But in this case~~ the server has to also initiate the connection with passive open. Means the server has to inform that he is ready to accept messages.

SYN (from the client) Flag is set

SEQ = 8000
A | | S |
1

SYN + ACK (from the server) different sequence No

SEQ = 20.000
ACK = 8001
A | | S |
1 1

→ randomly choosen by the server

→ I received 8000 and I am ready to accept 8001.

ACK (from the client)

SEQ = 8001
ACK = 20001
A | | S |
1

→ it received the sequence no 20.000 and it is ready to accept 20001.

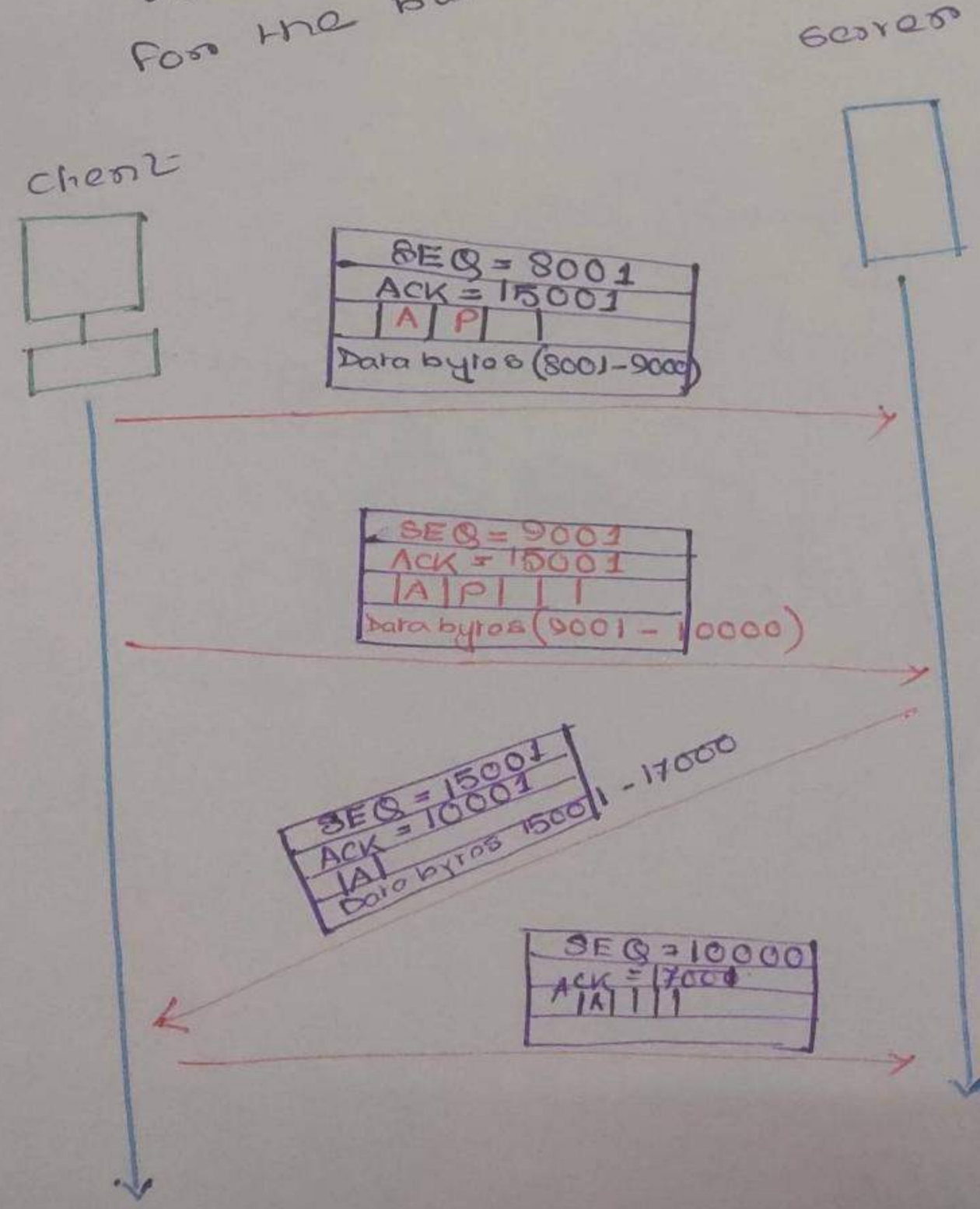
SEQ no will no change because it contains no data

TCP Data Transfer:-

2. Data Transfer

- Bidirectional Data Transfer
  (from client data and server sends ACK)
- Segment can carry both data + acknowledgement means the acknowledgement is piggybacked with data
- Push and urg flags → Data bytes which are urgent message as ~~suppose~~ (ctrl+c) means abort operation this message should be executed urgently.

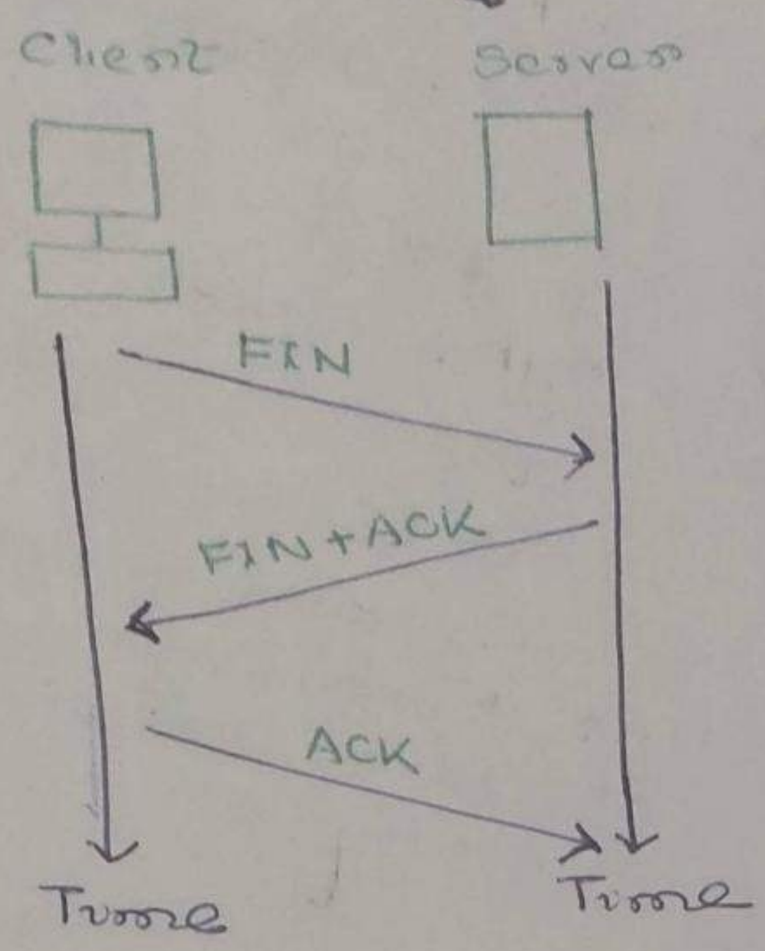
  ↓
  receiving TCP should handover the ~~pac~~ segment to the receiving application urgently/immediately.
  before the buffer overflows to be filled up.
  it should not wait for the buffer to be filled



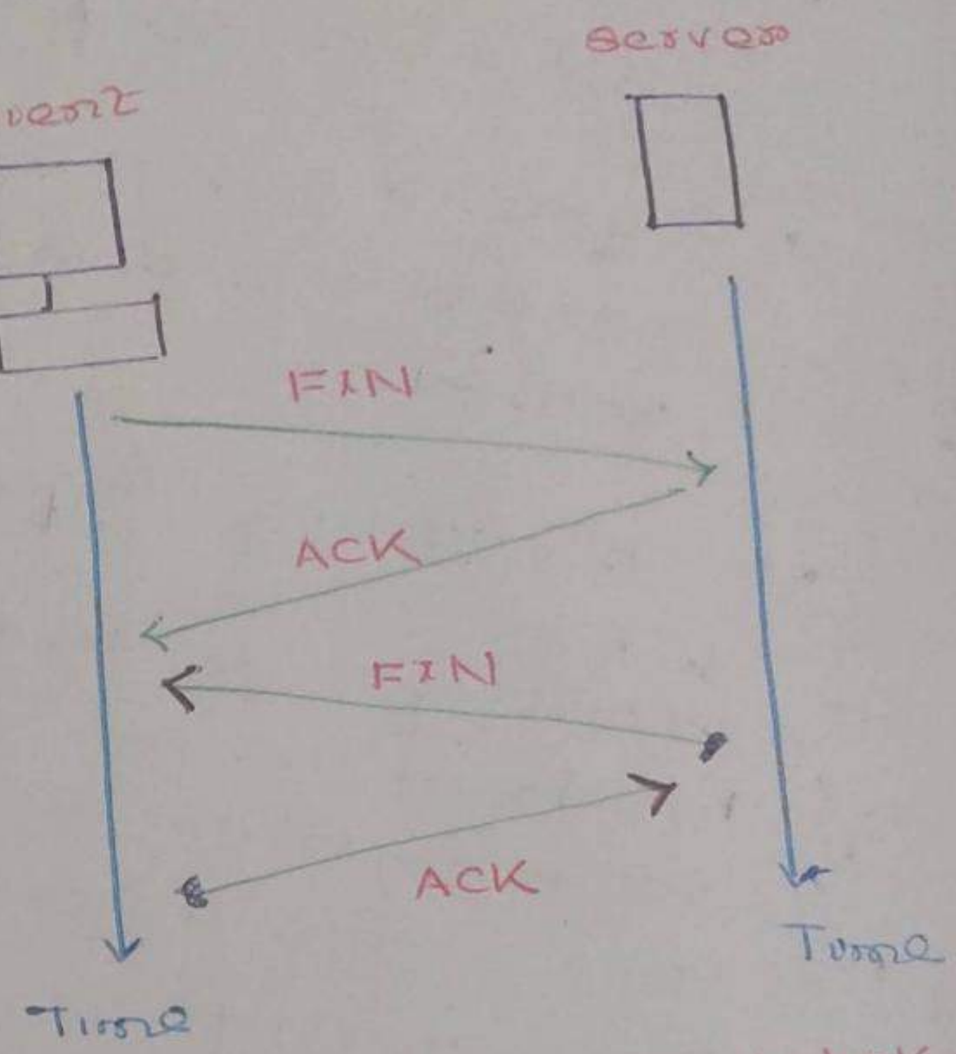
Assume that the three-way handshaking is done and now only data transmission is going on

TCP connection Termination:-

- Usually connection termination is initiated by the client.
- Two options for connection termination.
  - Three way handshaking.
  - Four way handshaking and a half-close option.



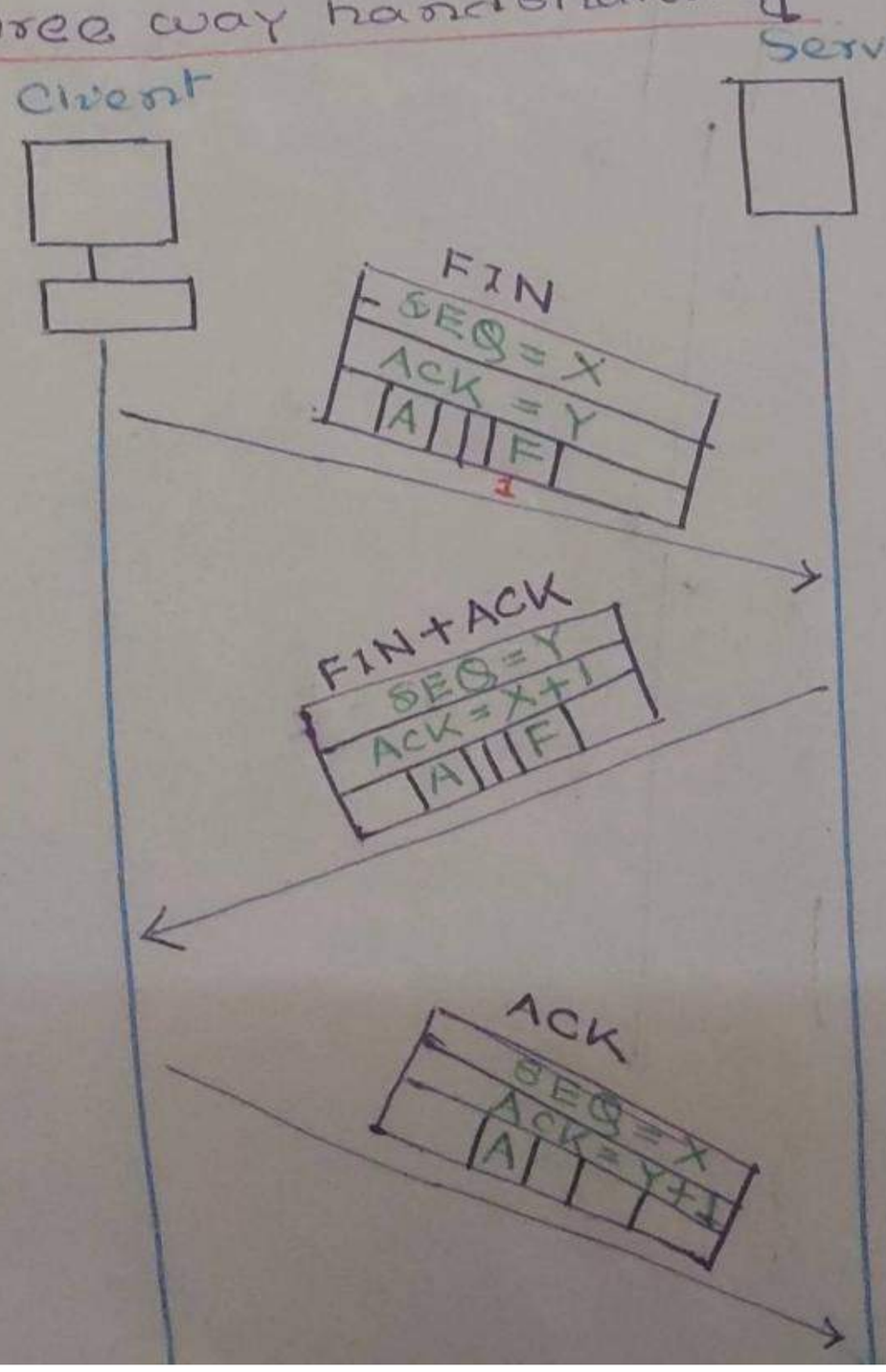
Three-way handshaking



Four-way handshaking

Both client and server can close the connection.

Three way handshaking:-



When server is initiating finish (FIN) then the server is no more allowed to send data. This FIN segment can be embedded with the last data segment. But in this time the server can send both data and acknowledgment till the server is not sending any FIN message.

Scanned by TapScanner

Module 4

Transport Layer

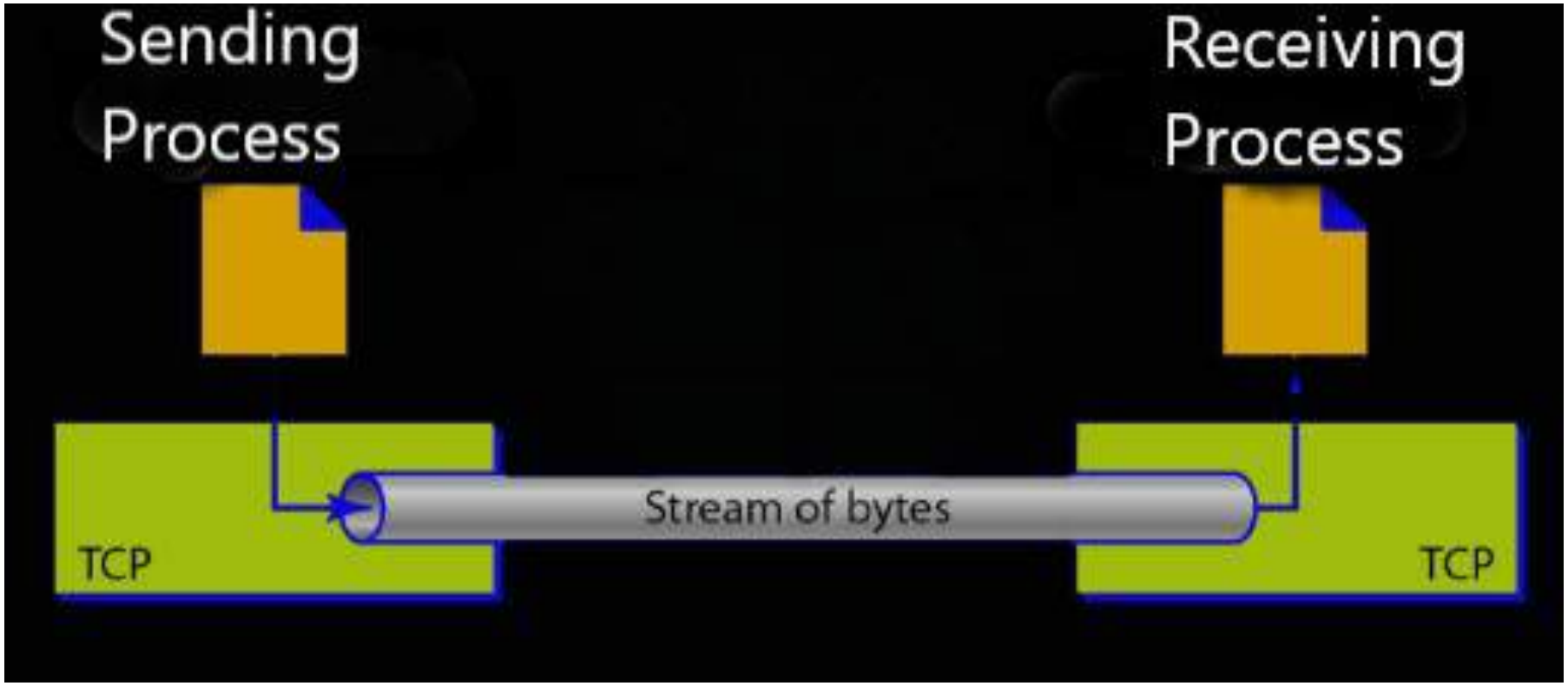
Dr. Sunandita Debnath, IIIT Vadodara

TCP

Features of TCP

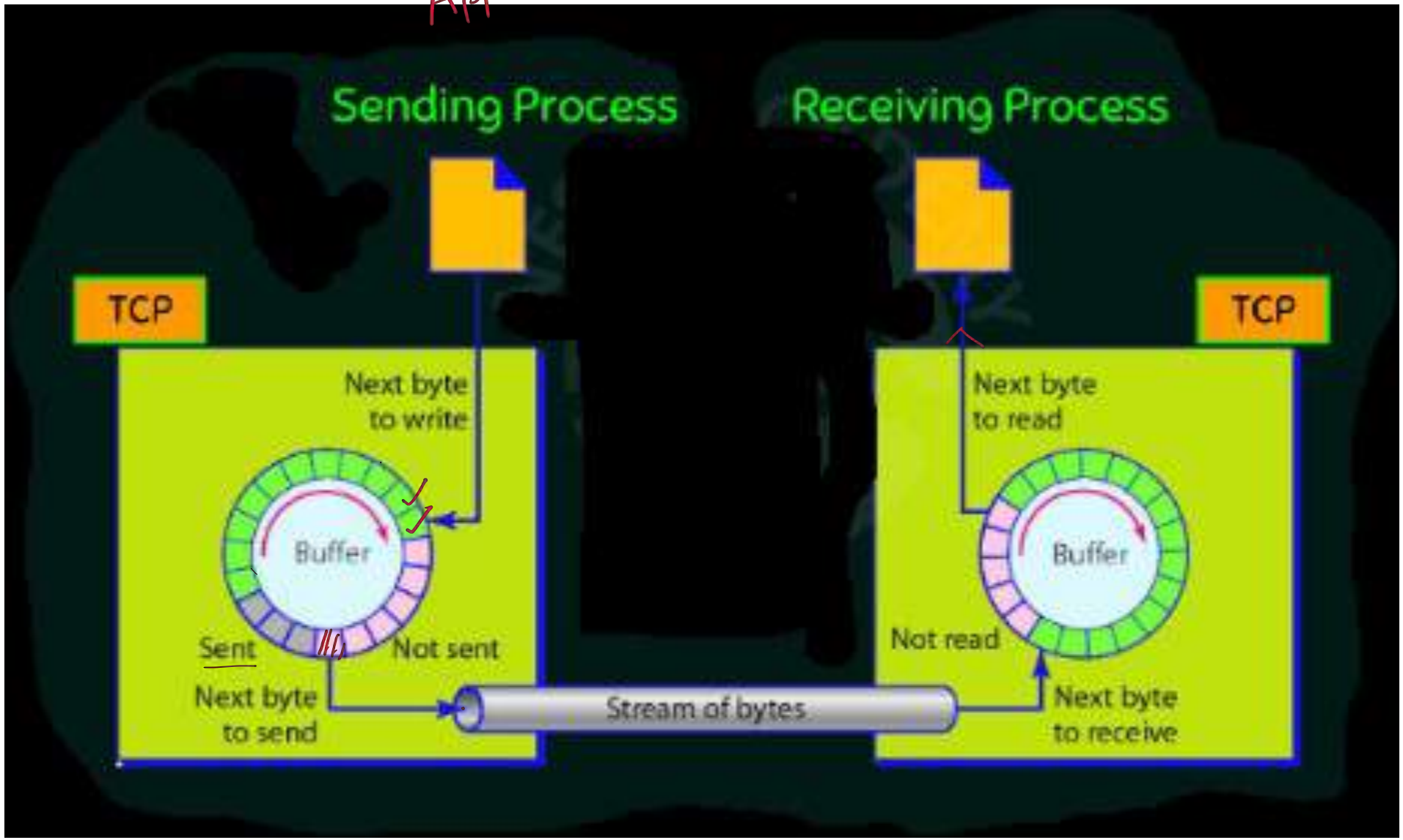
- *Connection oriented*
- *Reliable delivery*
- *Acknowledgement oriented*
- *Retransmission*
- *Flow control*
- *Error Control*
- *Congestion control*
- *Segmentation and Reassembly*
- *Full Duplex Support*

TCP Stream Delivery

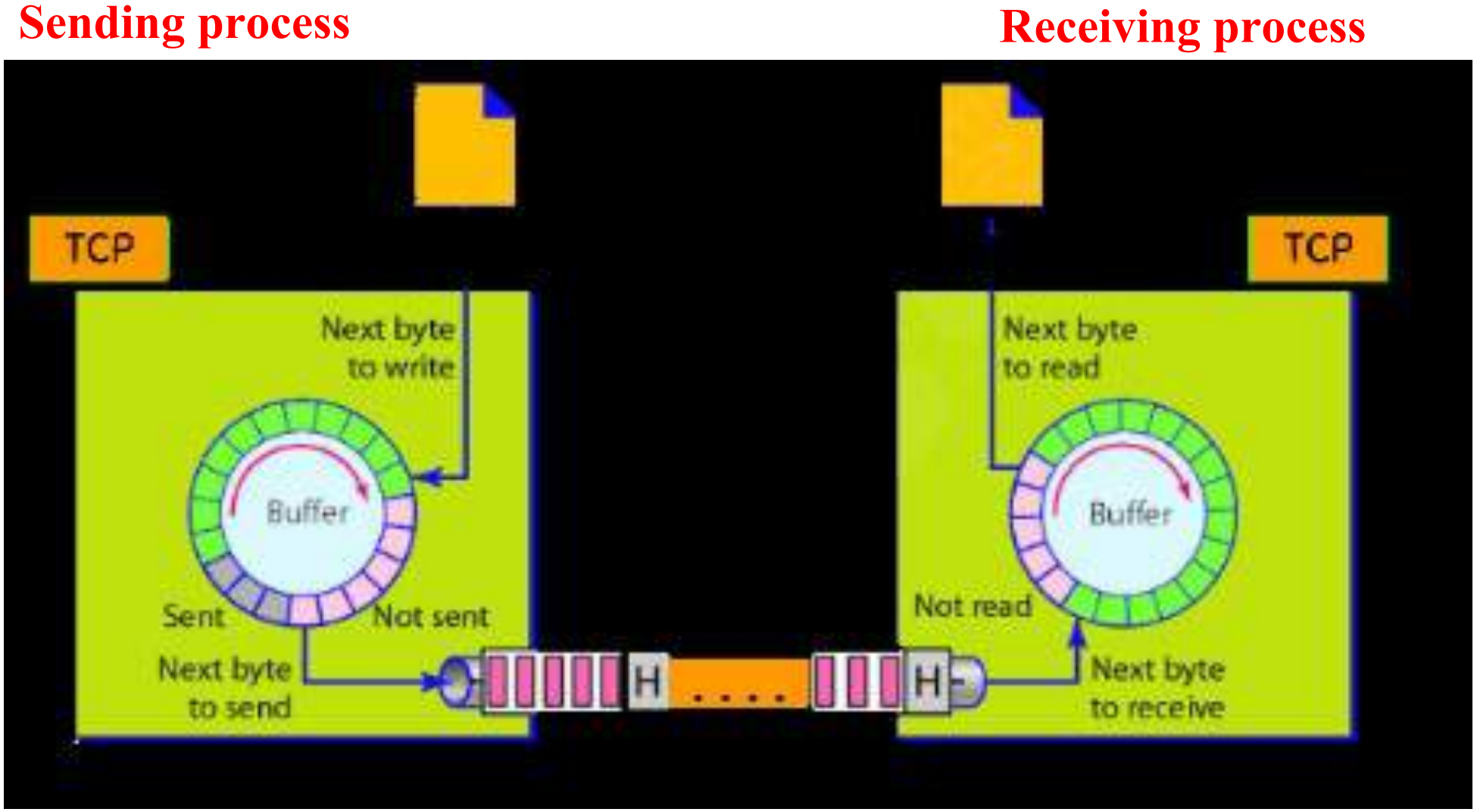


TCP Sending and Receiving Buffers

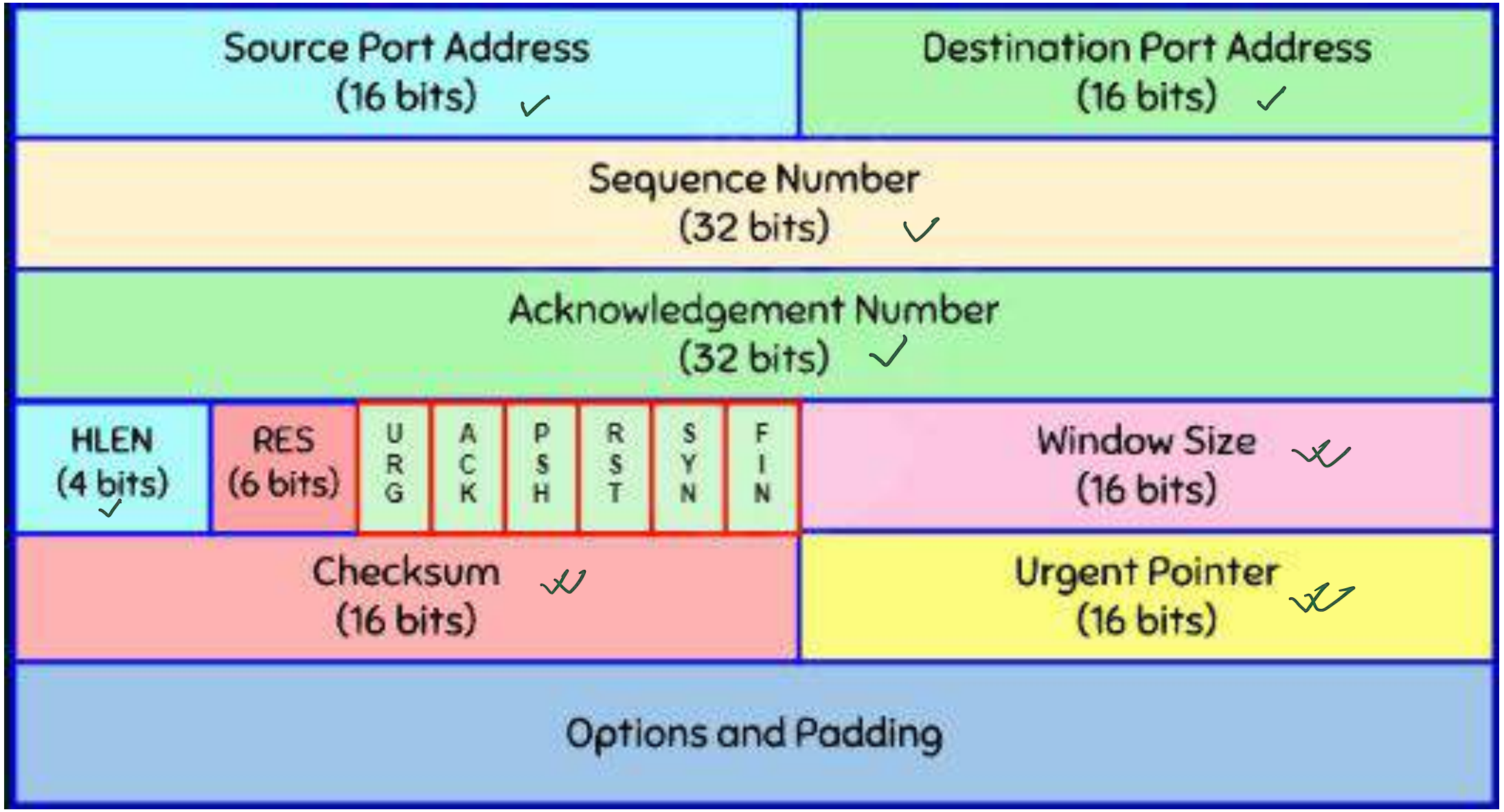
Application



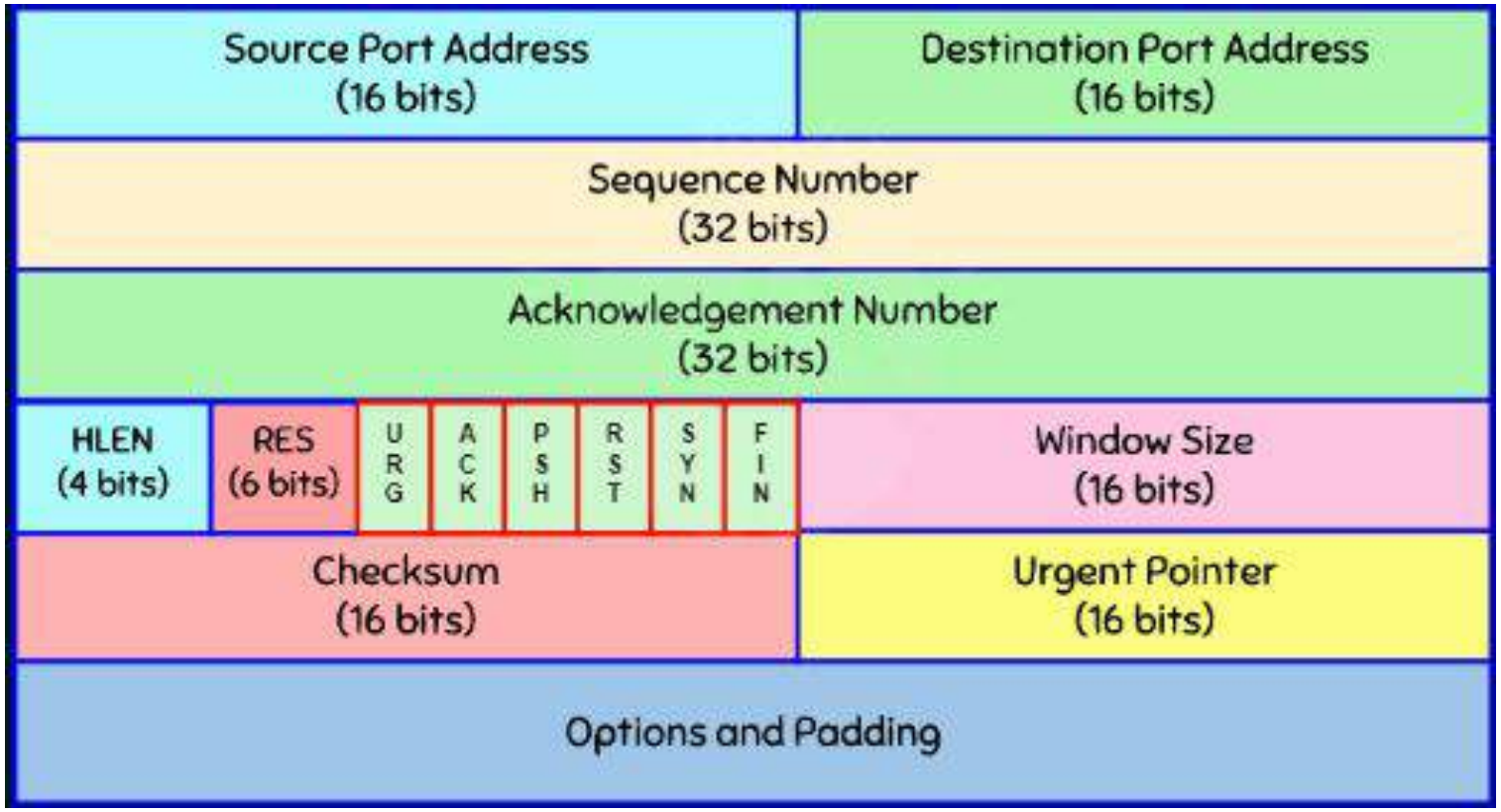
TCP Segments

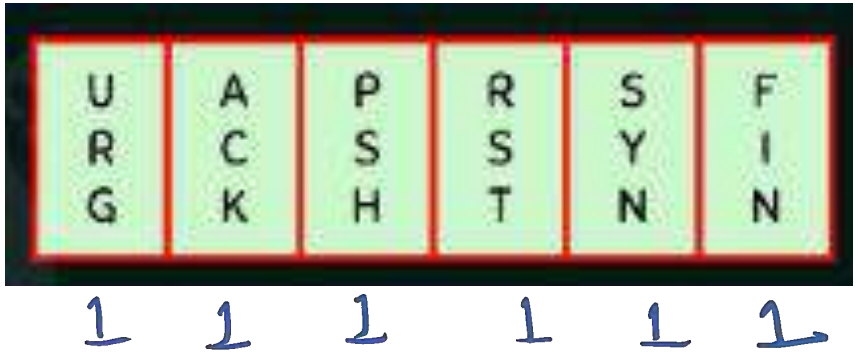


TCP Header Format



TCP Flags





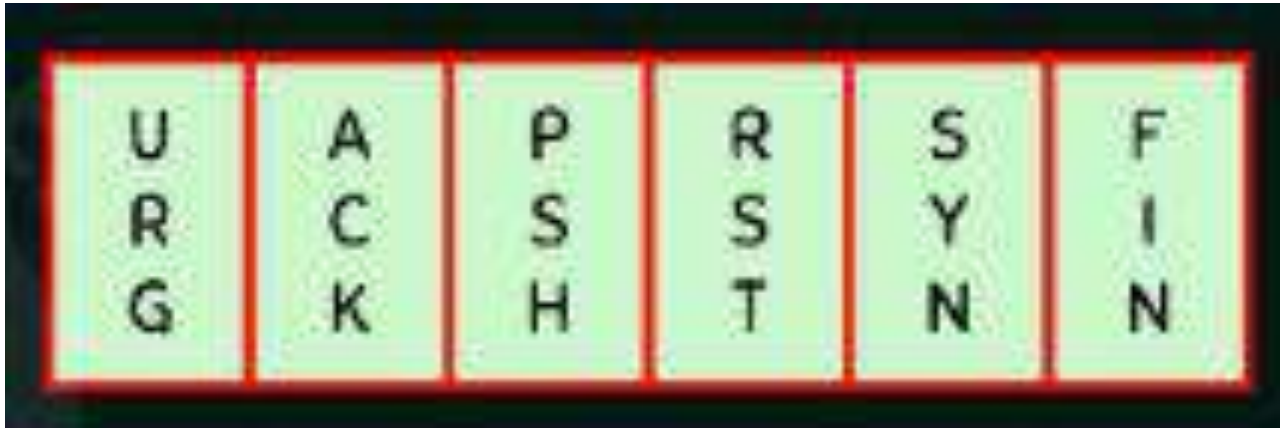
Data

seq 1 seq 2 seq 3

Flag:

- *This field defines 6 different control bits or flags.*
- *One or more of these bits can be set at a time.*

TCP Flags



Flag:

- *URG (1 bit): Indicates that the Urgent pointer field is significant.*
- *ACK (1 bit): Indicates that the Acknowledgement field is significant.*
- *PSH (1 bit): Push function. Asks to push the buffered data to the receiving application.*
- *RST (1 bit): Reset the connection*
- *SYN (1 bit): Synchronize sequence number.*
- *FIN (1 bit): Last packet from the sender.*

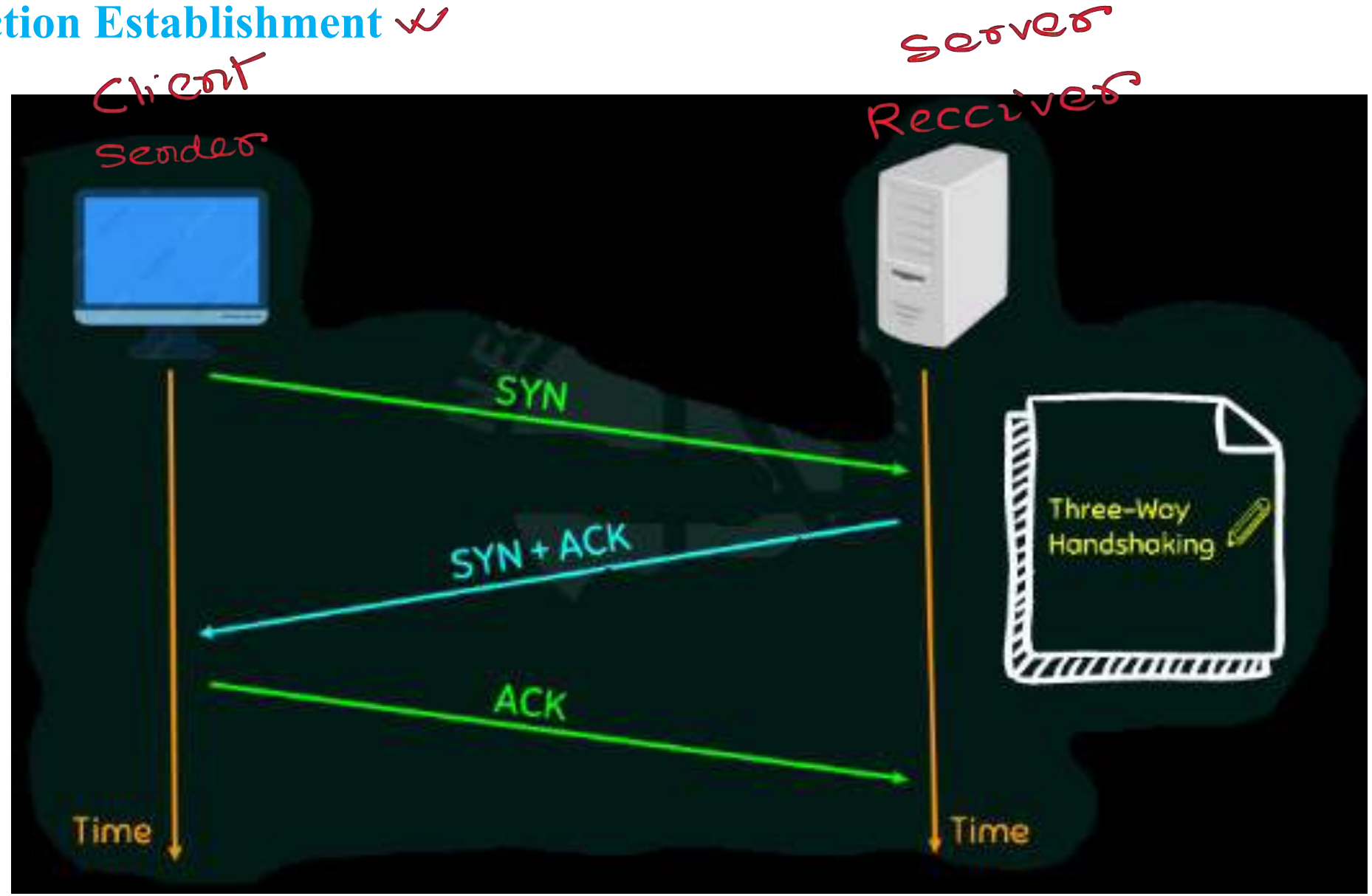
TCP connections

- *Connection oriented.*
- *Virtual path.*
- *Acknowledgement process.*
- *RST (1 bit): Reset the connection*
- *Retransmission of lost or damaged segments.*
- *TCP connection-Virtual not physical.*
- *IP –connection*
- *Full-duplex mode.*
- *Approval from other party.*

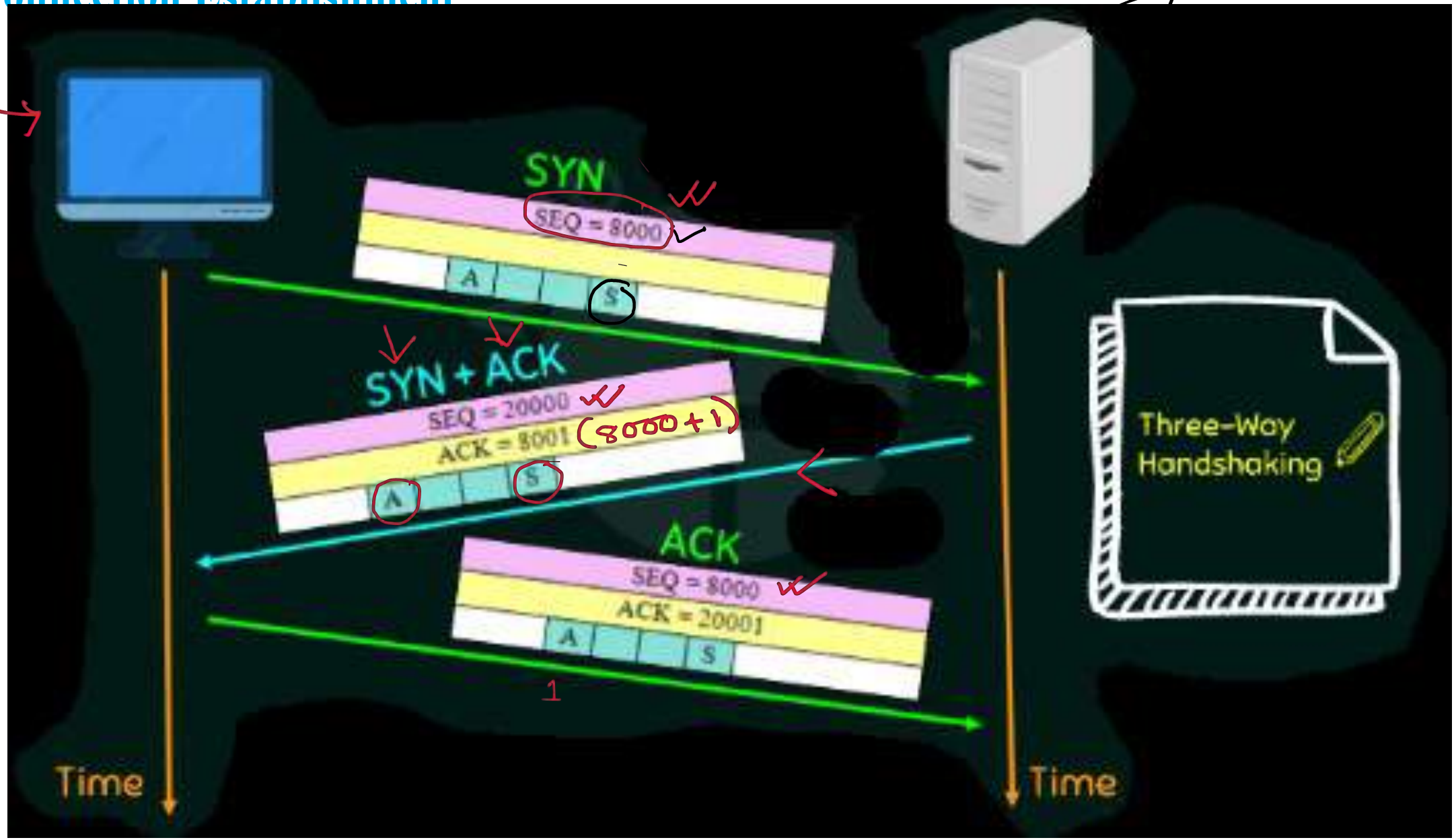
Three Phases of TCP connections

- *Connection Establishment*
- *Data Transfer*
- *Connection Termination*

TCP Connection Establishment

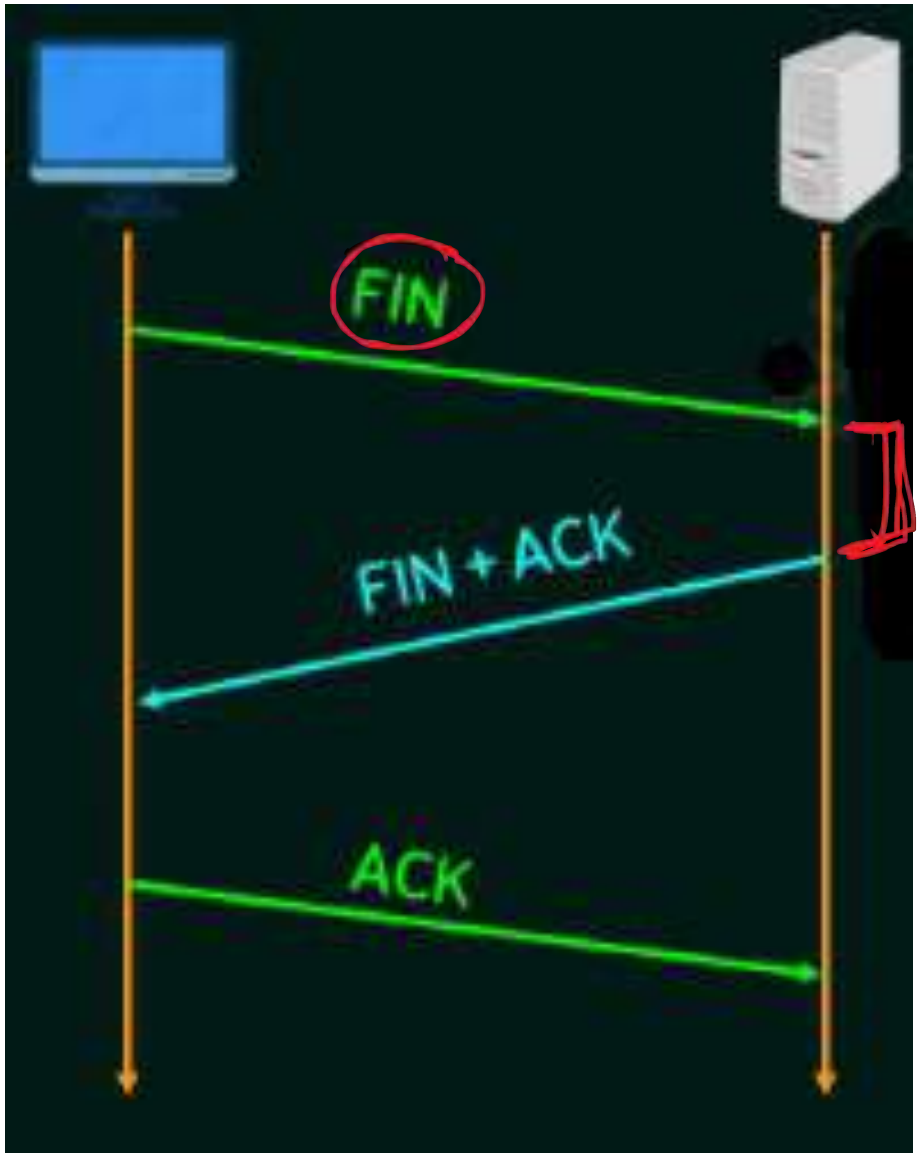


TCP Connection Establishment

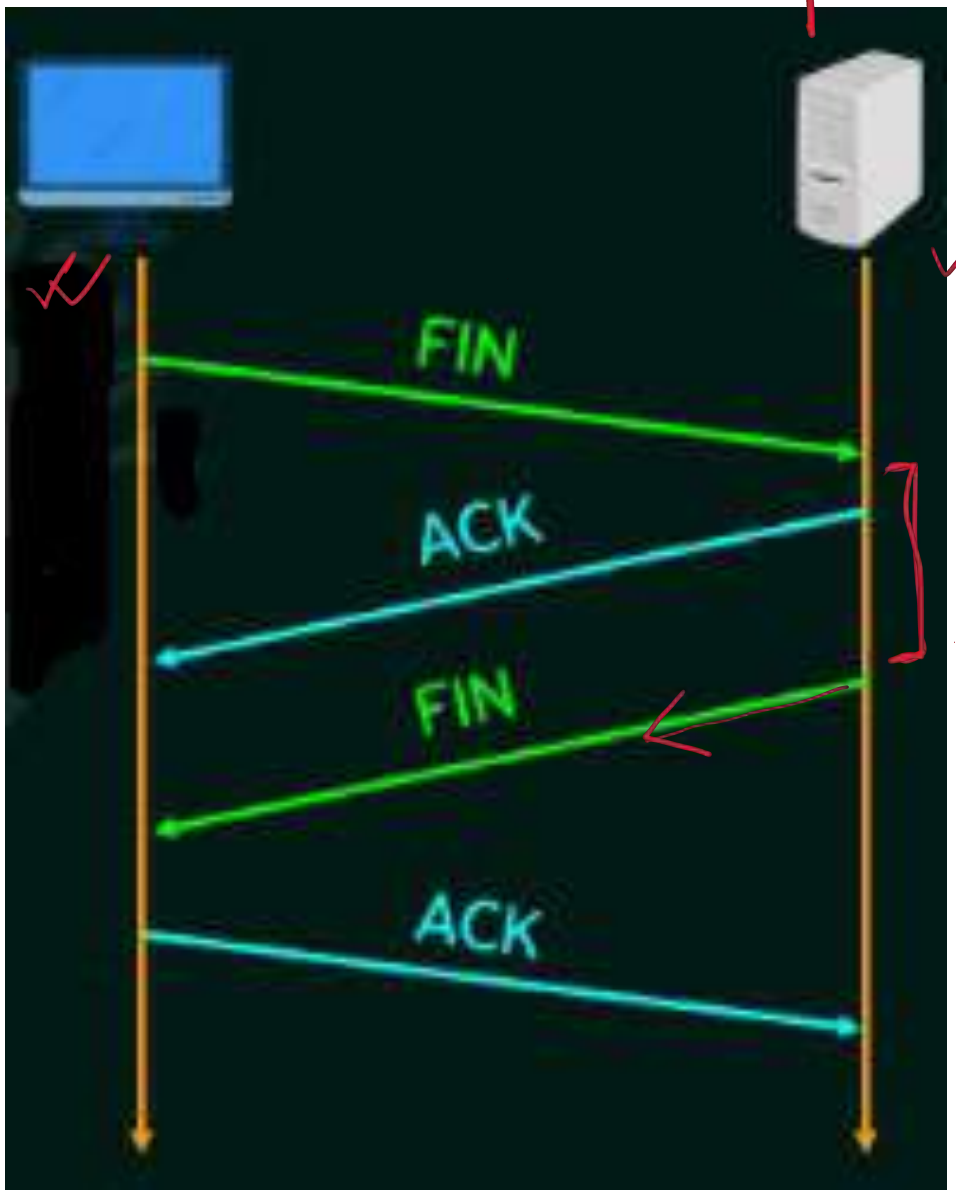


TCP Connection Termination

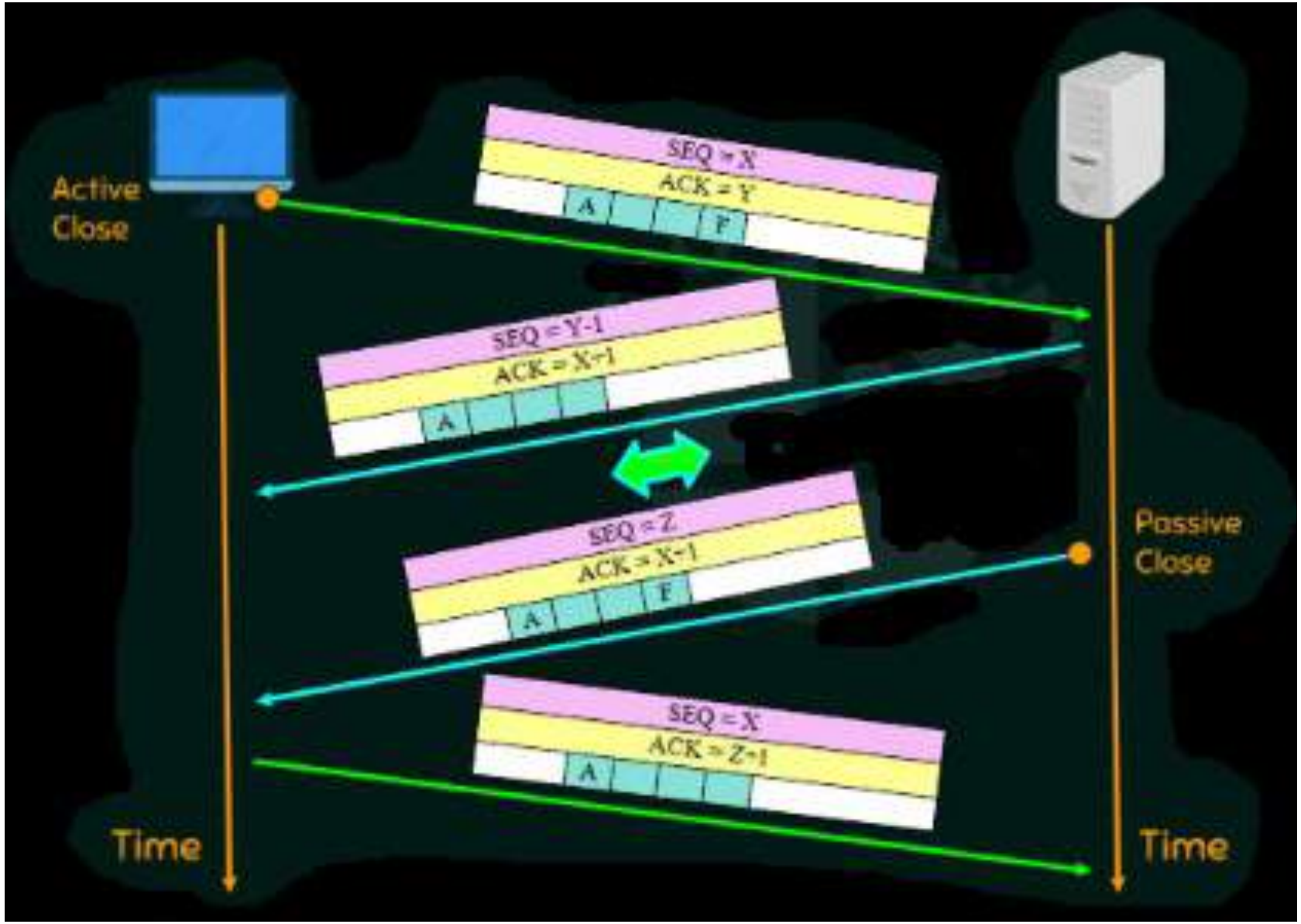
Three-way



Four-way



Connection Termination



TCP Flow Control

- *TCP uses a sliding window to handle flow control.*
- *Between Go-back and selective Repeat.*
- *Does not use NAK.*
- *The receiver holds out-of-order segments.*
- *TCP sliding window is byte-oriented.*
- *TCP's sliding window is of variable size.*
- *Imaginary window.*

Sender

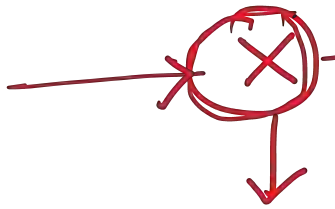
200 Kbps

Receiver

OK

TCP Sliding Window

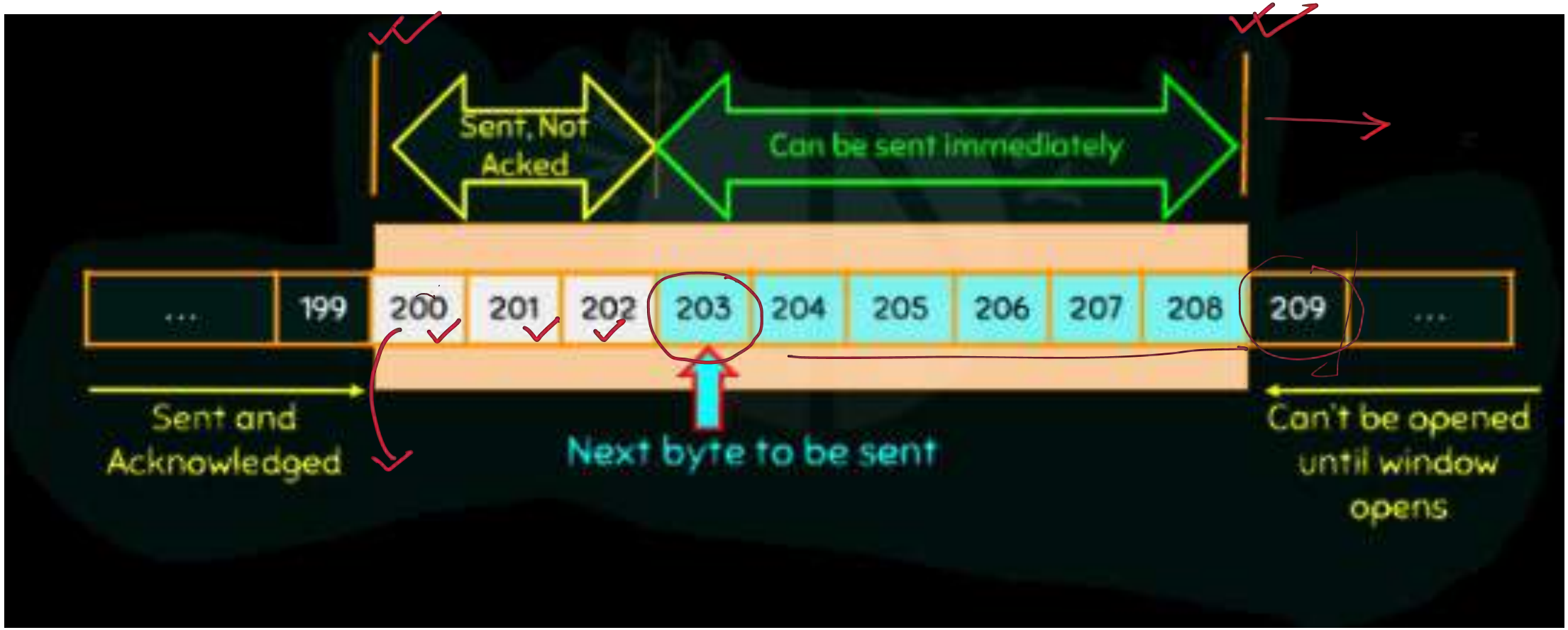
200 Kbps Sender



2.5 Kb

Cw

Window size = Minimum (rwnd, cwnd) = Minimum (20,9) = 9

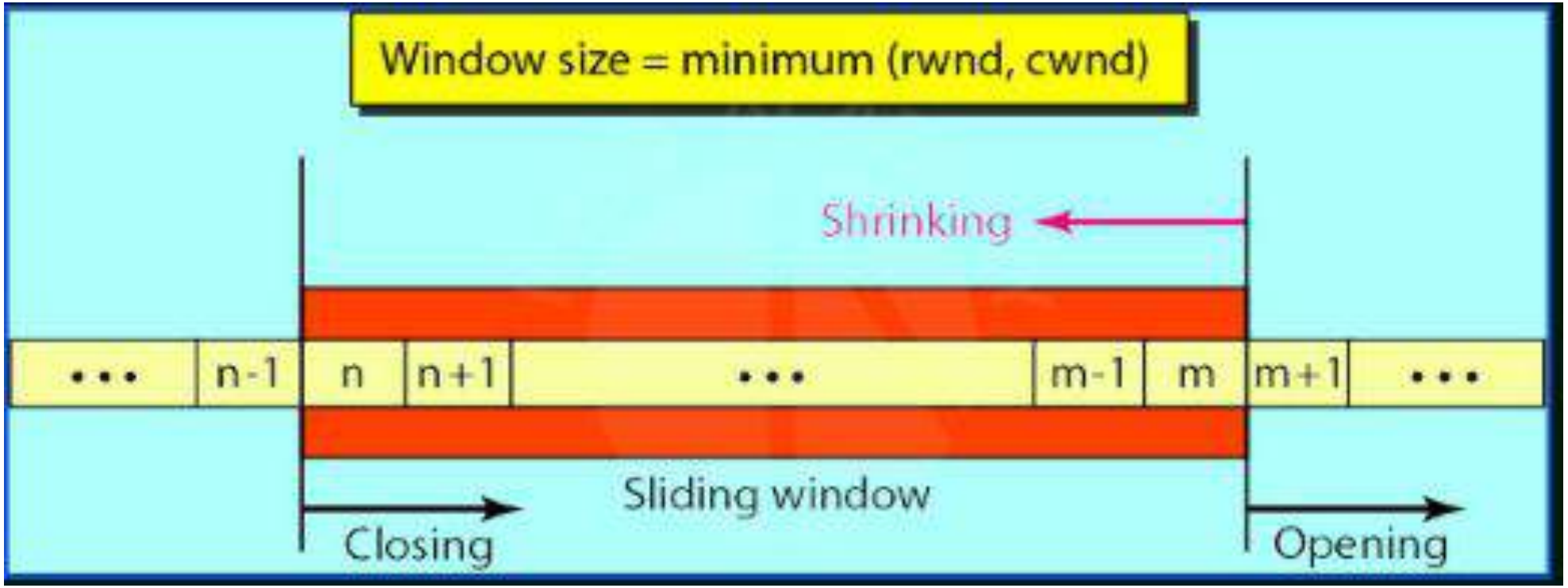


TCP Sliding Window

Points to remember

- *Window size = Minimum (rwnd, cwnd).*
- *The source does not have to send a full window's worth of data.*
- *The window can be opened or closed by the receiver, but should not be shrunk.*
- *The destination can send an ACK at any time*
- *The sender can send 1-byte segment even after the window is shut down in the receiver side.*

TCP Sliding Window



Questions. What is the value of the receiver window (rwnd) for host A if the receiver host B, has a buffer size of 5000 bytes and 1000 bytes of received and unprocessed data.

rwnd

Rwnd = 5000-1000=4000 bytes

Host B can only receive bytes of data before overflowing its buffer. Host B advertises this value 4000 in its next segment to Host A.

buffer size is

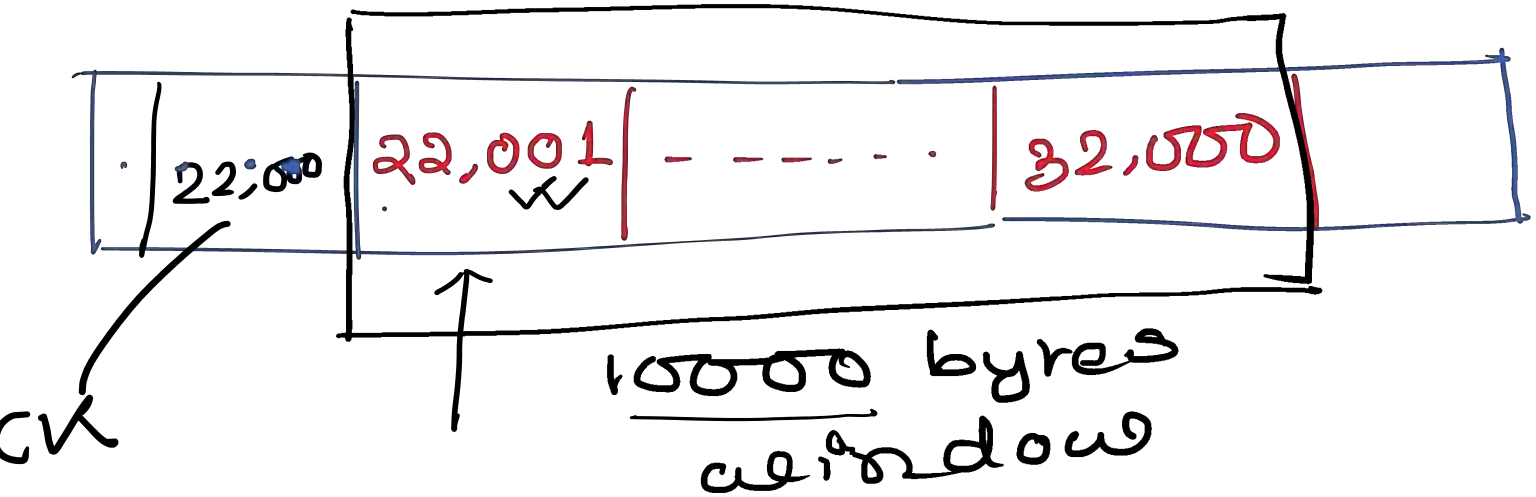
5000 byte
− 1000 bytes
4000 bytes

Questions. Suppose a TCP connection is transferring a file of 1000 bytes, The first byte is numbered 100001. What is the sequence number of the segment if all data is sent in only one segment?

File size = 1000 bytes

First byte of data = 100001 +
1000
101001 bytes
1
101000 bytes

Questions. A TCP connection is using a window size of 10,000 bytes, and the previous acknowledgement number was 22,001. It receives a segment with acknowledgement number 24,001 and window size advertisement of 12,000. Draw a diagram to show the situation of the window before and after.

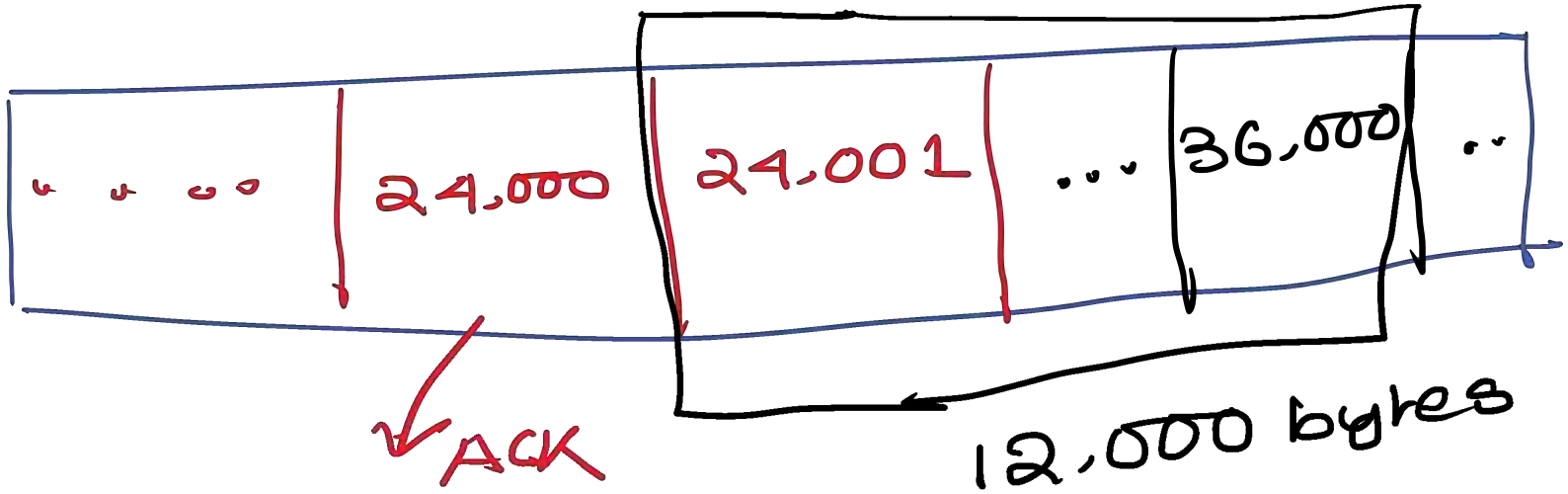


24,001
+ 12,000
36001
− 1

22,001
+10,00
32,00

22,000 @ 3

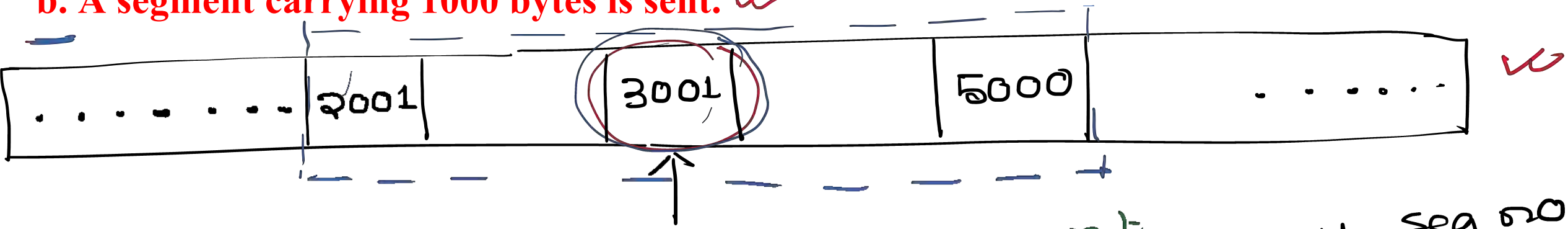
successful receive
now the receiver expect
(22,000



Questions. A window holds bytes 2001 to 5000. The next byte to be sent is 3001. Draw a figure to show the situation of the window after the following two events.

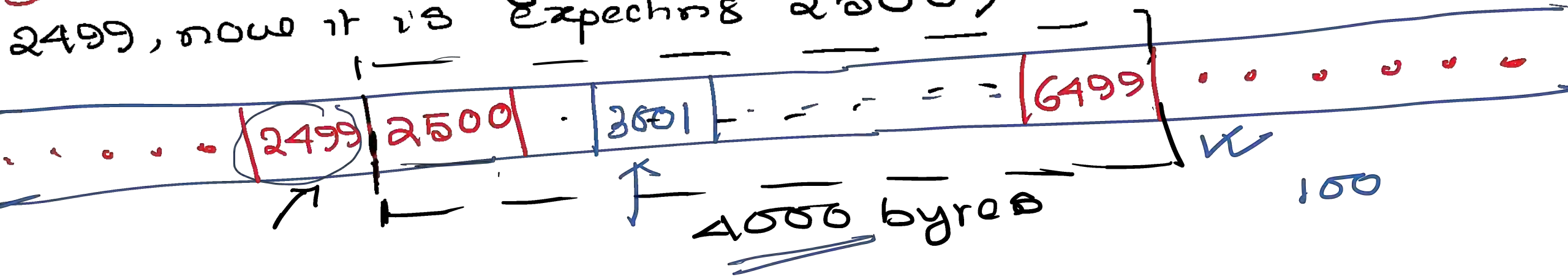
a. An ACK segment with the acknowledgement number 2500 and window size advertisement 4000 is received.

b. A segment carrying 1000 bytes is sent.

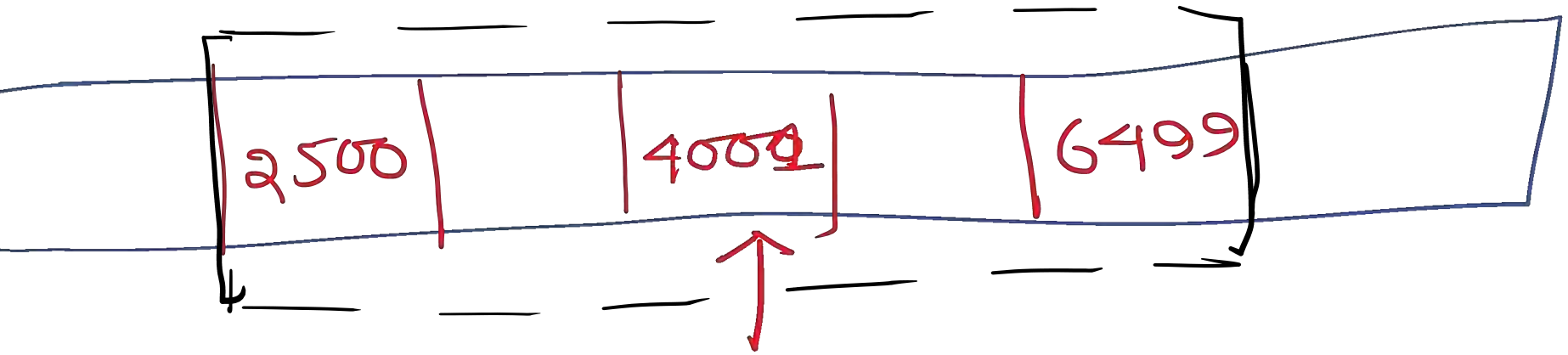


next byte to be sent

Case 1:- successfully received the segment with seq no 2499, now it is expecting 2500, window = 4000



100



next byte to be sent.

3001

1000 bo

30
+ 10
40

Module 4

Transport Layer

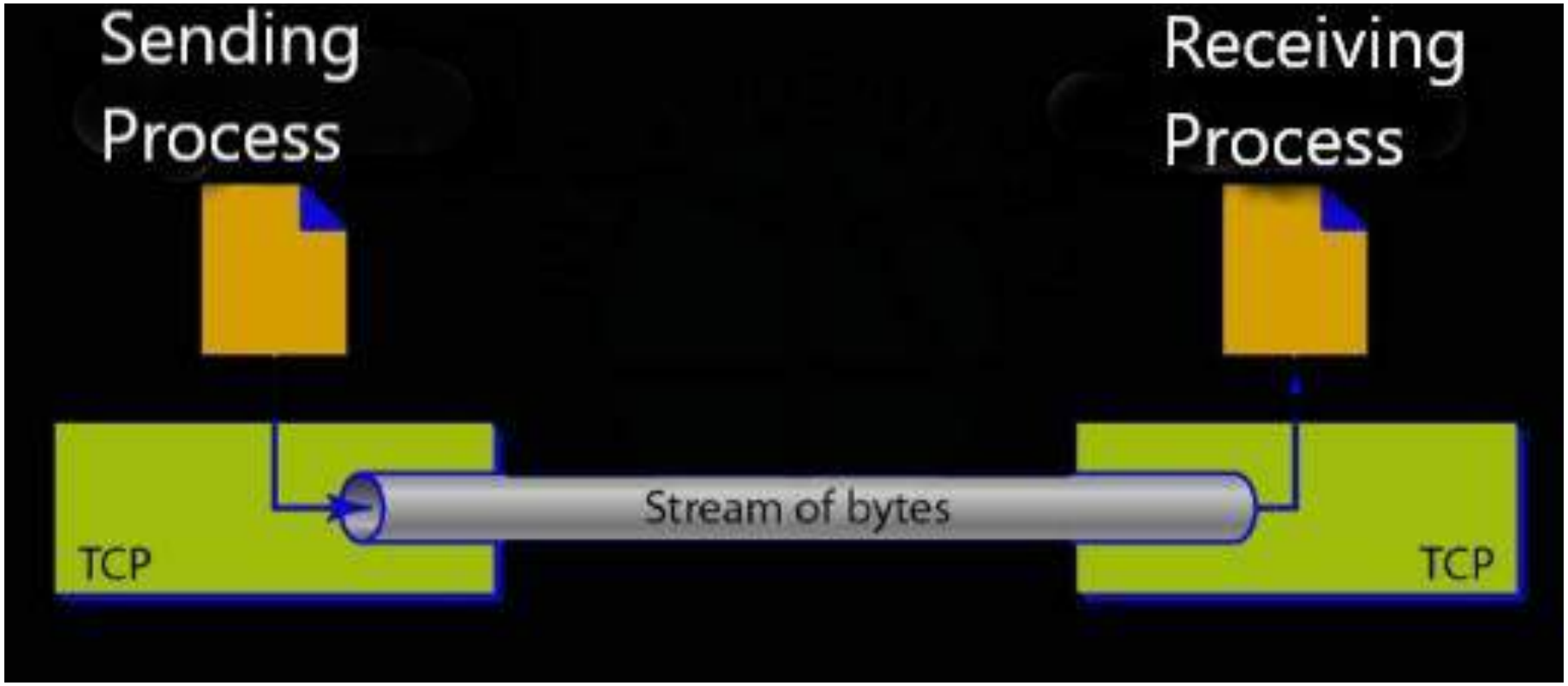
Dr. Sunandita Debnath, IIIT Vadodara

TCP

Features of TCP

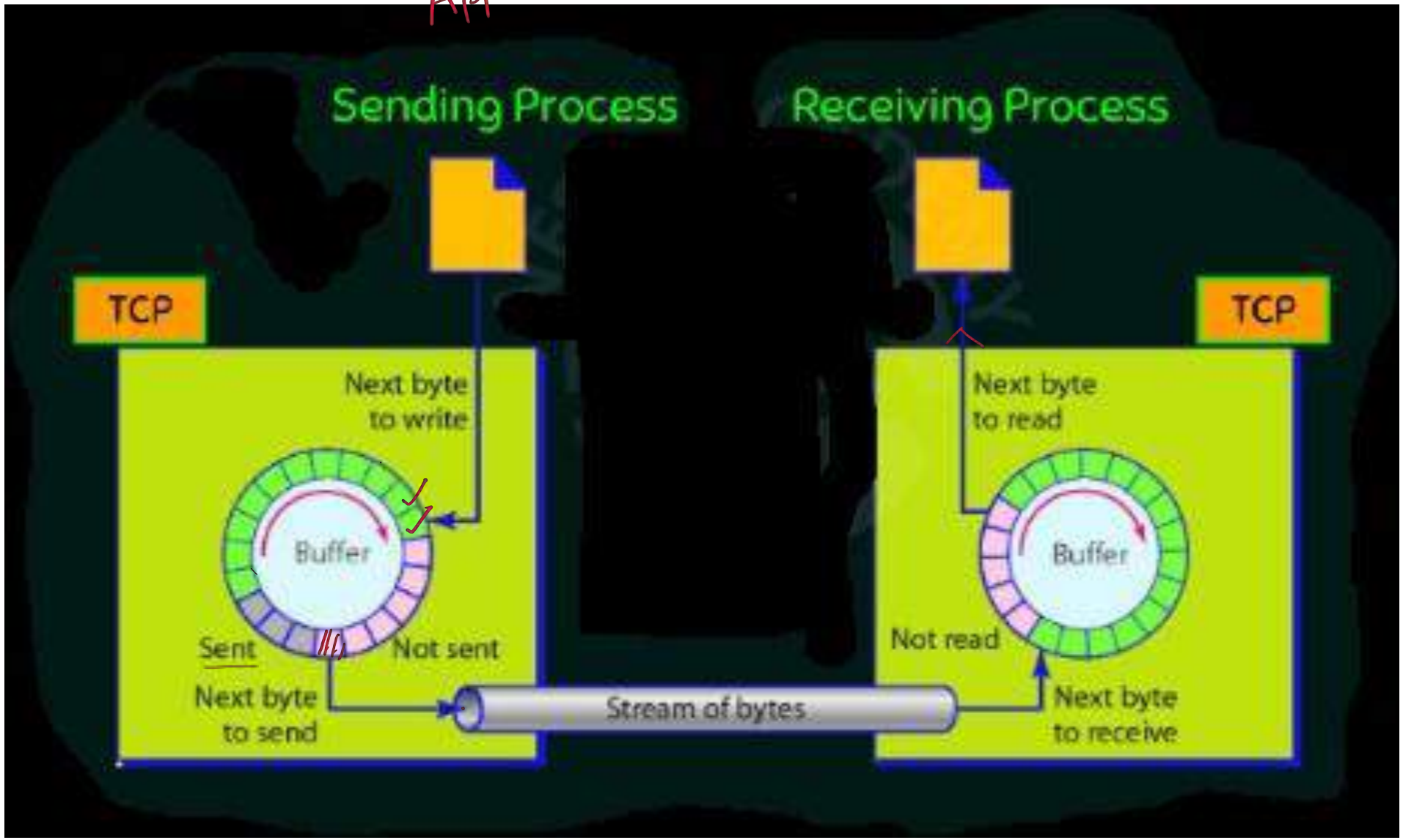
- *Connection oriented*
- *Reliable delivery*
- *Acknowledgement oriented*
- *Retransmission*
- *Flow control*
- *Error Control*
- *Congestion control*
- *Segmentation and Reassembly*
- *Full Duplex Support*

TCP Stream Delivery

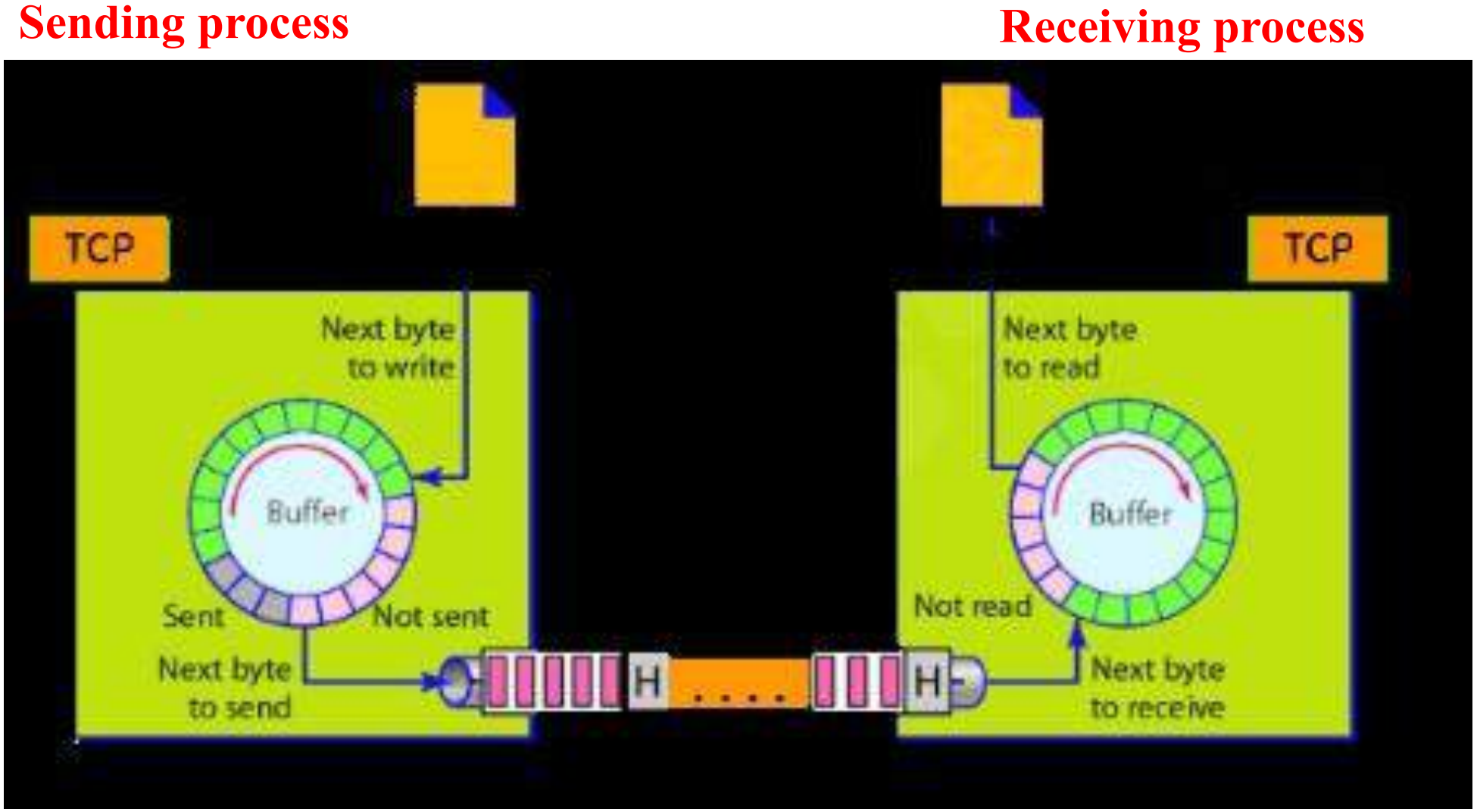


TCP Sending and Receiving Buffers

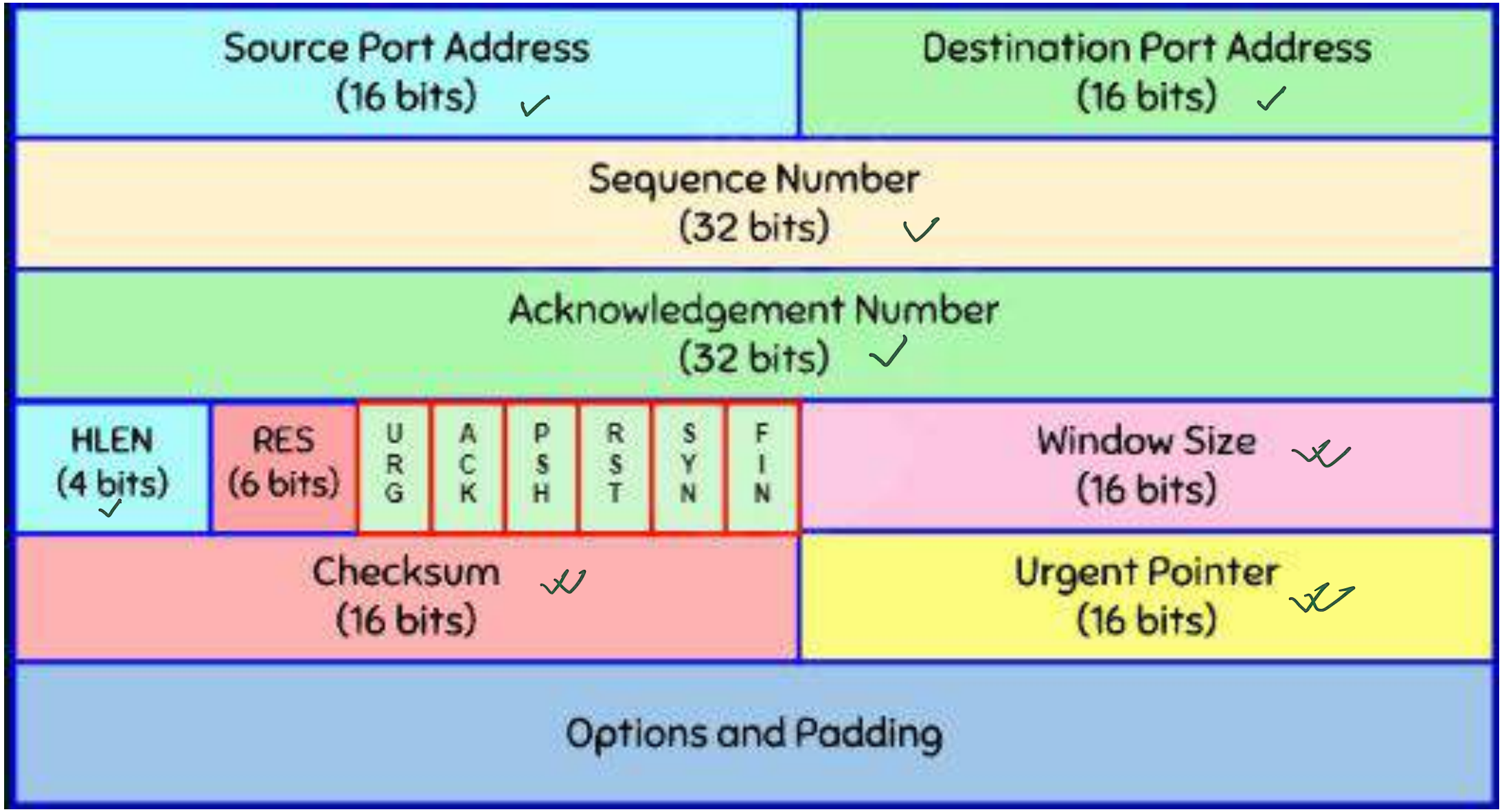
Application



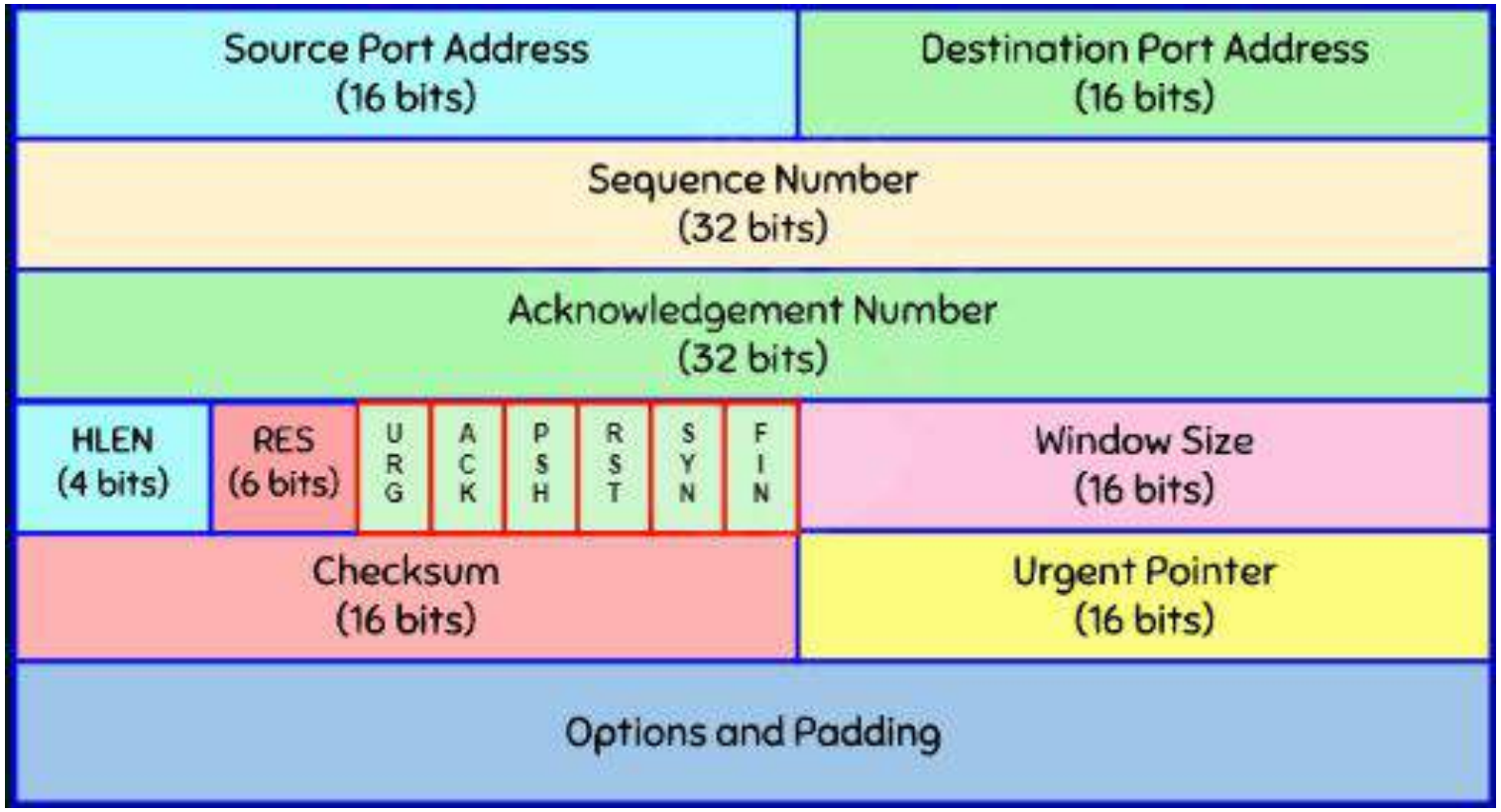
TCP Segments

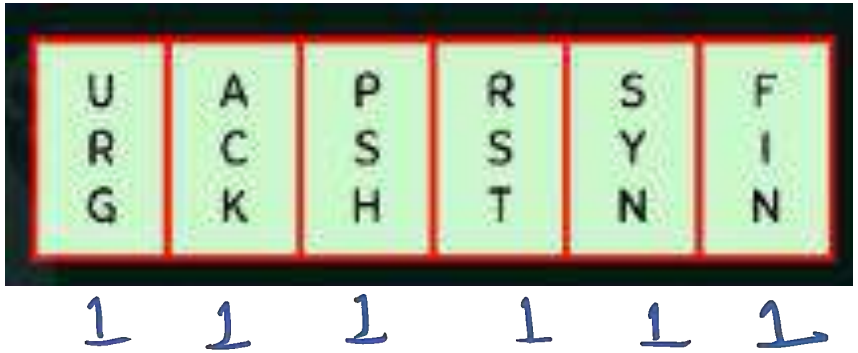


TCP Header Format



TCP Flags





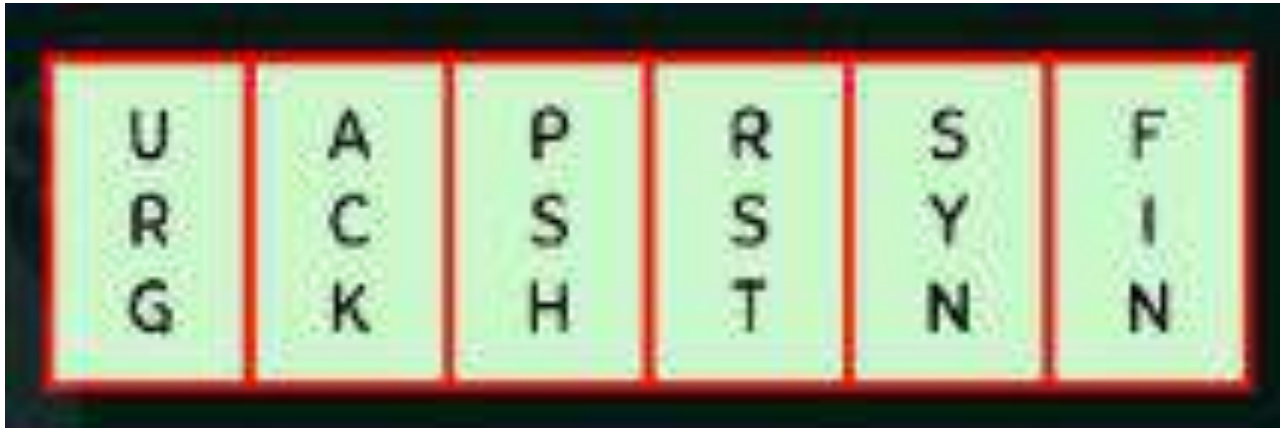
Data

seq 1 seq 2 seq 3

Flag:

- *This field defines 6 different control bits or flags.*
- *One or more of these bits can be set at a time.*

TCP Flags



Flag:

- *URG (1 bit): Indicates that the Urgent pointer field is significant.*
- *ACK (1 bit): Indicates that the Acknowledgement field is significant.*
- *PSH (1 bit): Push function. Asks to push the buffered data to the receiving application.*
- *RST (1 bit): Reset the connection*
- *SYN (1 bit): Synchronize sequence number.*
- *FIN (1 bit): Last packet from the sender.*

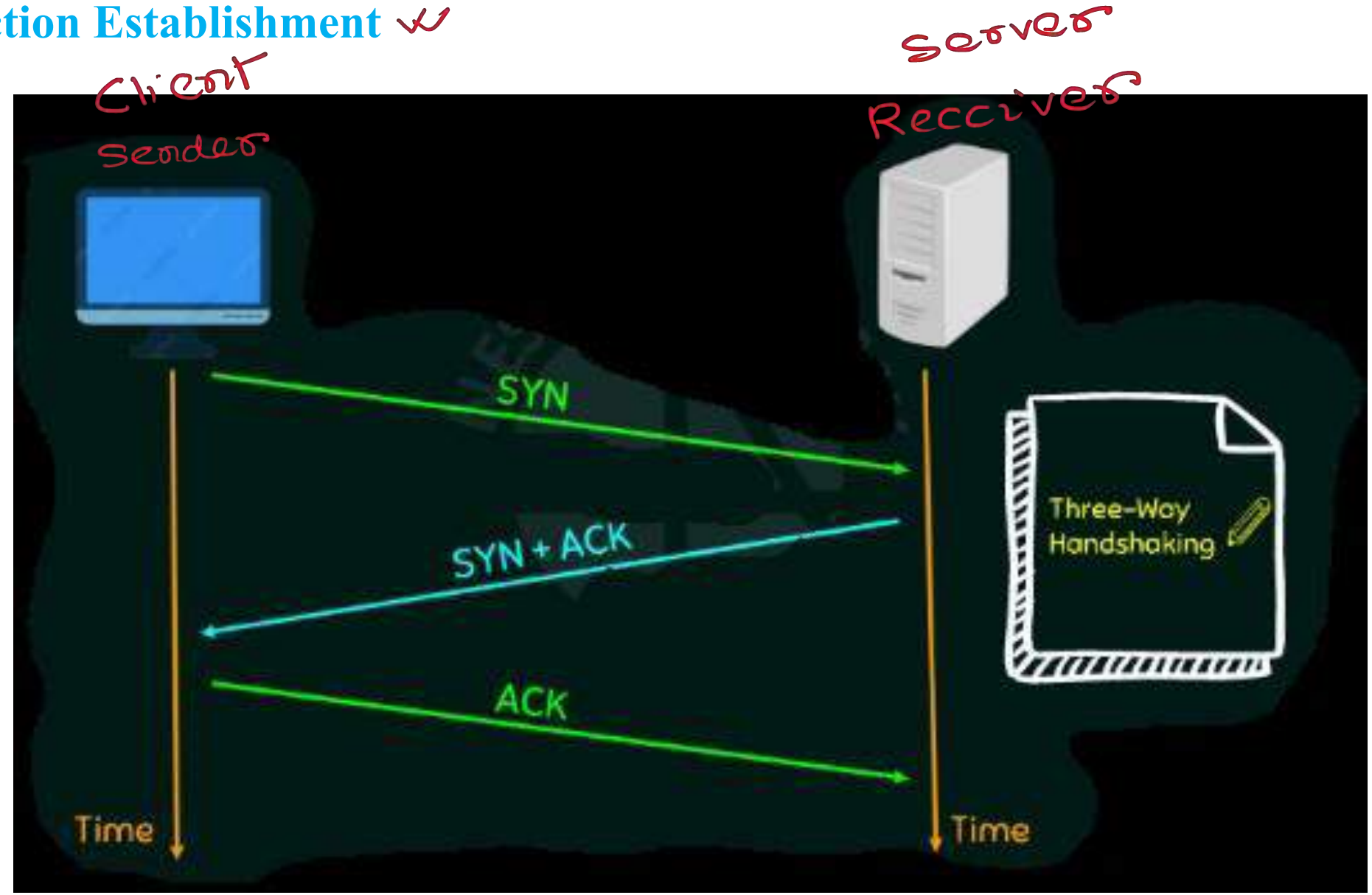
TCP connections

- *Connection oriented.*
- *Virtual path.*
- *Acknowledgement process.*
- *RST (1 bit): Reset the connection*
- *Retransmission of lost or damaged segments.*
- *TCP connection-Virtual not physical.*
- *IP –connection*
- *Full-duplex mode.*
- *Approval from other party.*

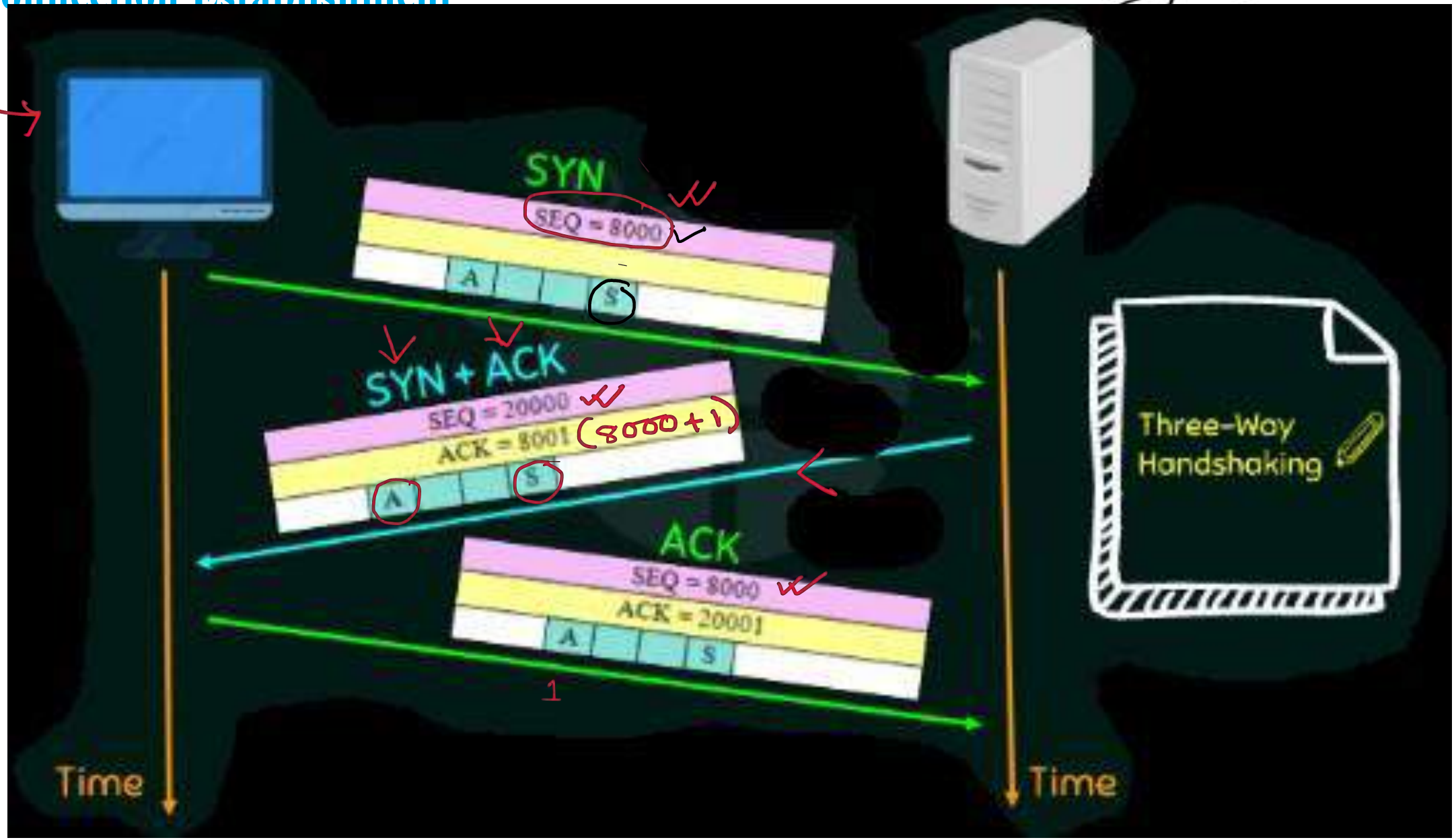
Three Phases of TCP connections

- *Connection Establishment*
- *Data Transfer*
- *Connection Termination*

TCP Connection Establishment

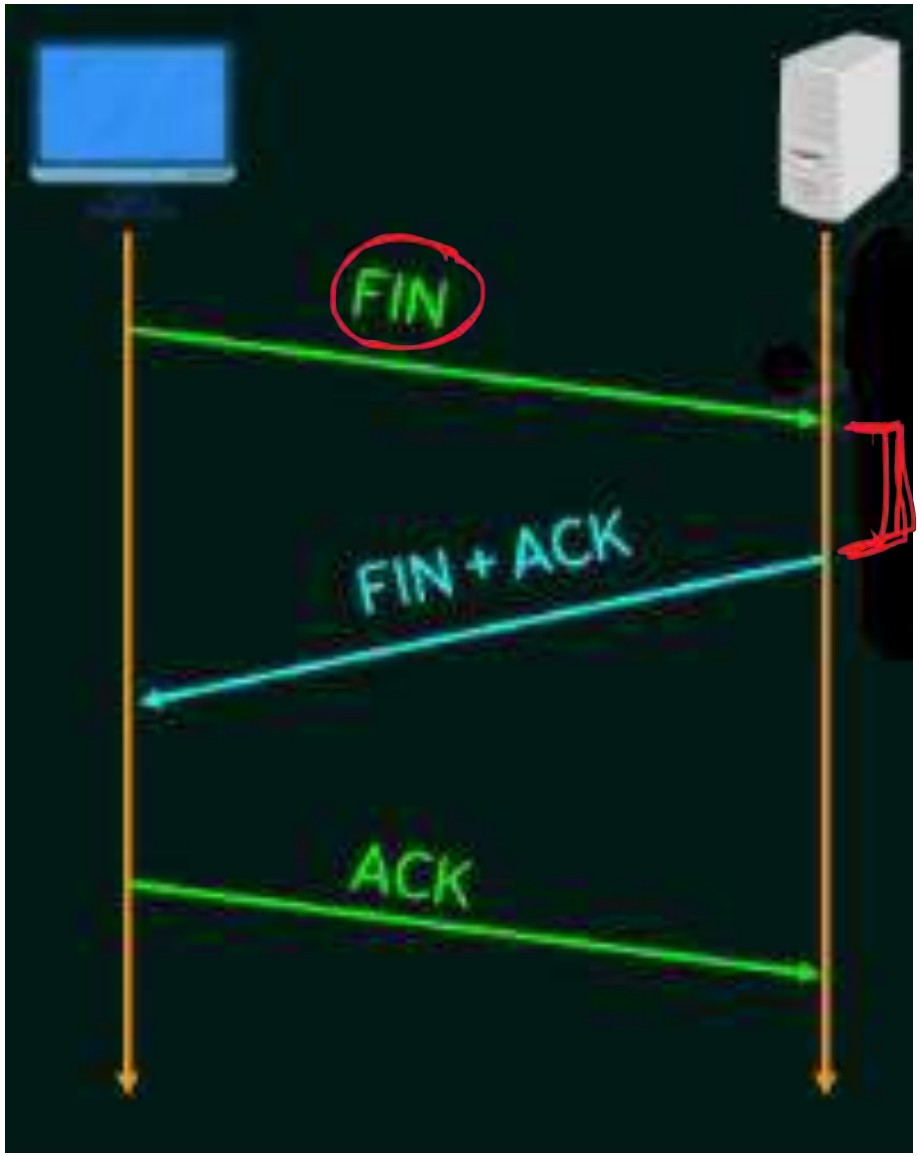


TCP Connection Establishment

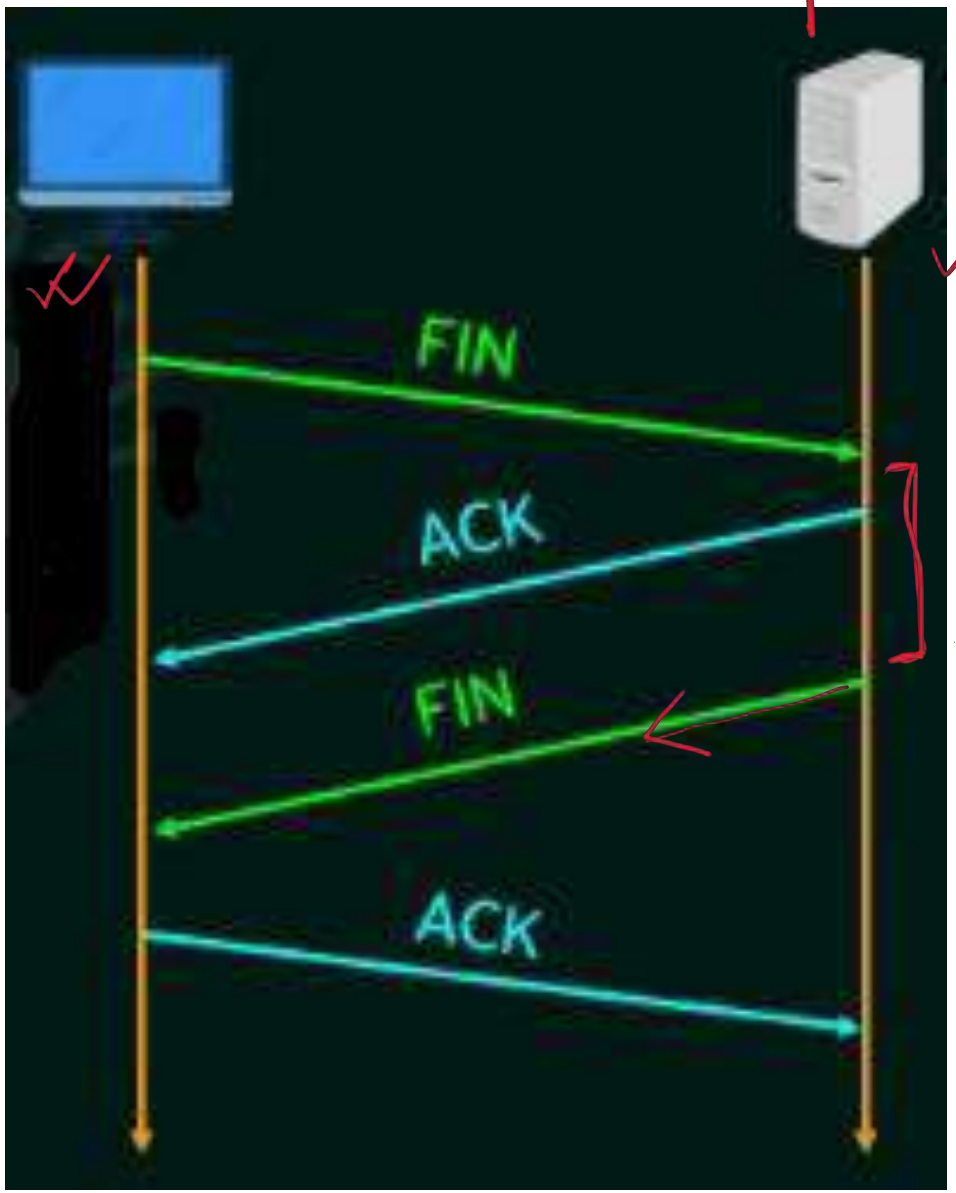


TCP Connection Termination

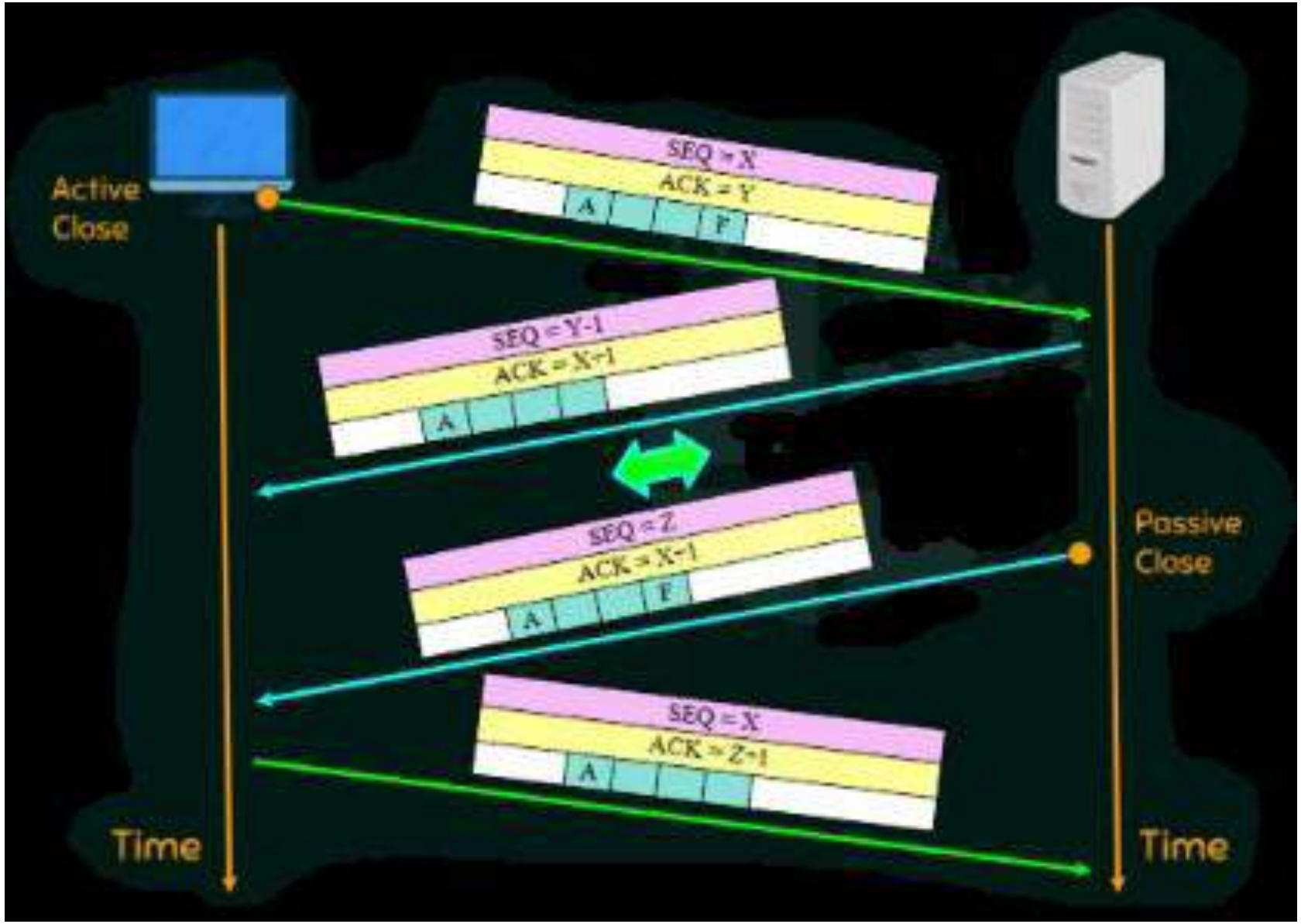
Three-way



Four-way



Connection Termination



TCP Flow Control

- *TCP uses a sliding window to handle flow control.*
- *Between Go-back and selective Repeat.*
- *Does not use NAK.*
- *The receiver holds out-of-order segments.*
- *TCP sliding window is byte-oriented.*
- *TCP's sliding window is of variable size.*
- *Imaginary window.*

Sender

200 kbps

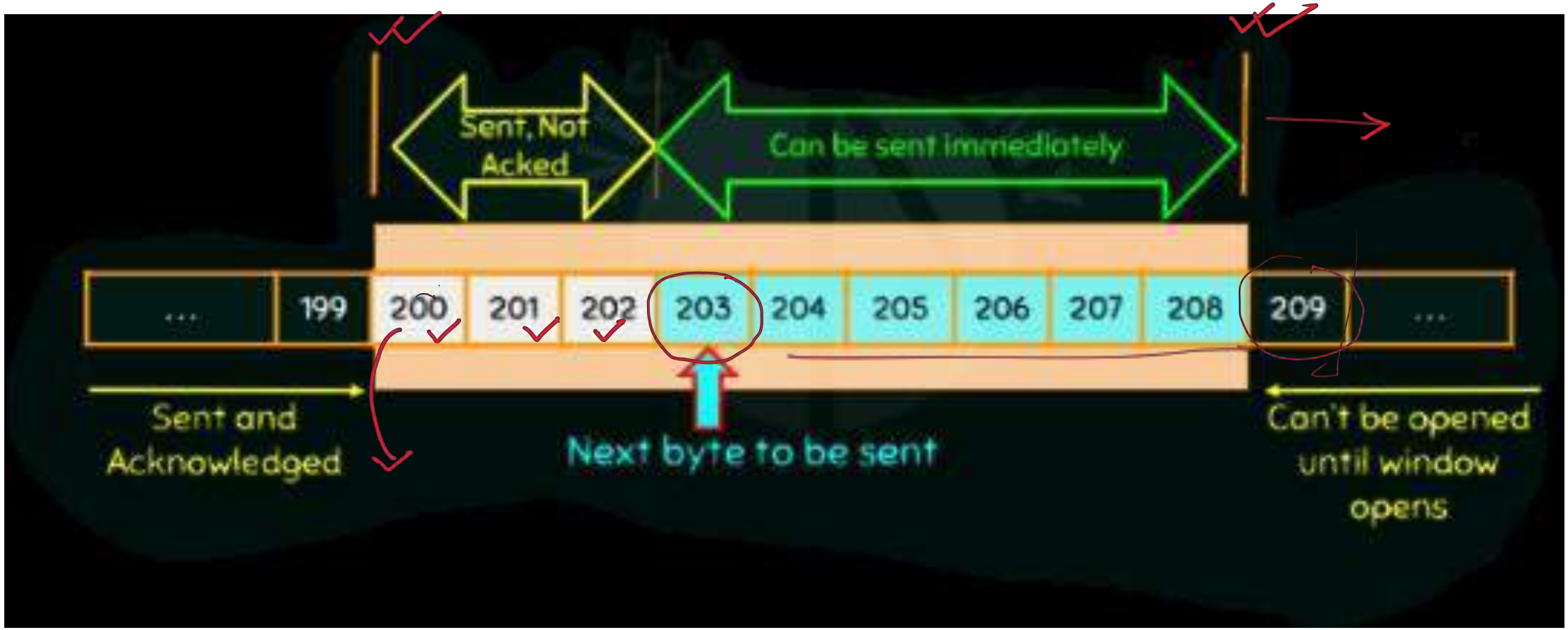
Receiver

50 kbps

TCP Sliding Window

200 Kbps Sender

Window size = Minimum (rwnd, cwnd) = Minimum (20,9) = 9

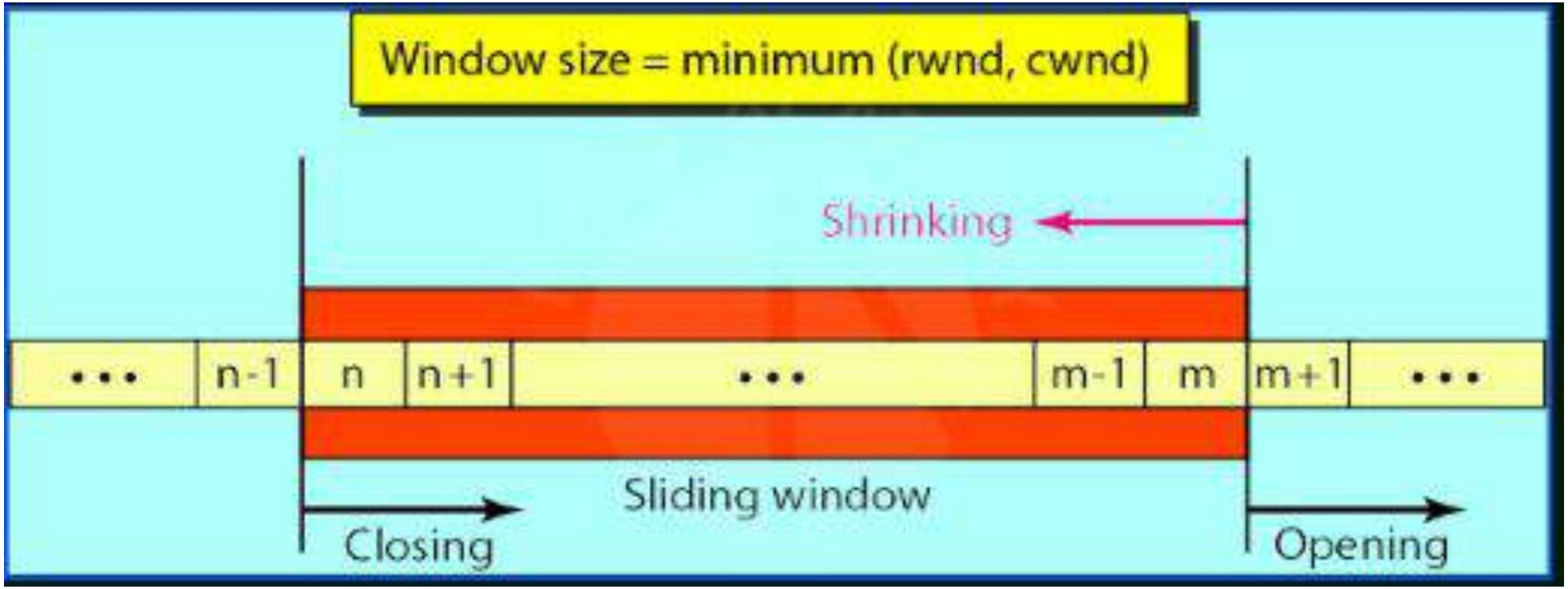


TCP Sliding Window

Points to remember

- *Window size = Minimum (rwnd, cwnd).*
- *The source does not have to send a full window's worth of data.*
- *The window can be opened or closed by the receiver, but should not be shrunk.*
- *The destination can send an ACK at any time*
- *The sender can send 1-byte segment even after the window is shut down in the receiver side.*

TCP Sliding Window



Questions. What is the value of the receiver window (rwnd) for host A if the receiver host B, has a buffer size of 5000 bytes and 1000 bytes of received and unprocessed data.

rwnd

Rwnd = 5000-1000=4000 bytes

Host B can only receive bytes of data before overflowing its buffer. Host B advertises this value 4000 in its next segment to Host A.

buffer size is

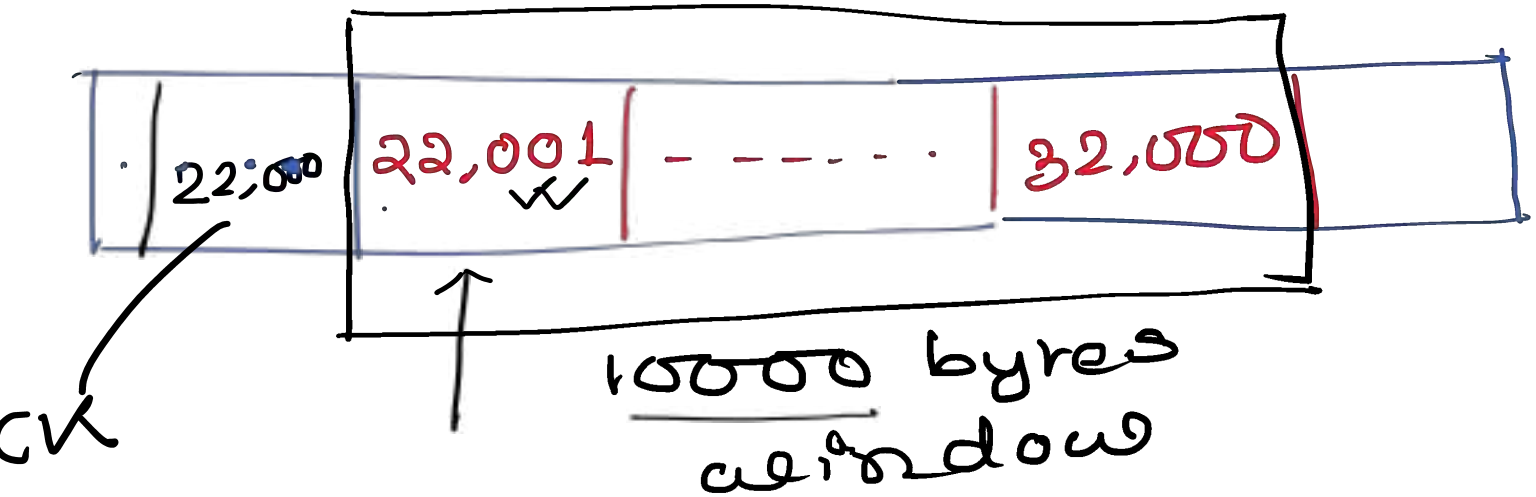
5000 byte
− 1000 bytes
4000 bytes

Questions. Suppose a TCP connection is transferring a file of 1000 bytes, The first byte is numbered 100001. What is the sequence number of the segment if all data is sent in only one segment?

File size = 1000 bytes

First byte of data = 100001 +
1000
101001 bytes
−1
101000 bytes

Questions. A TCP connection is using a window size of 10,000 bytes, and the previous acknowledgement number was 22,001. It receives a segment with acknowledgement number 24,001 and window size advertisement of 12,000. Draw a diagram to show the situation of the window before and after.

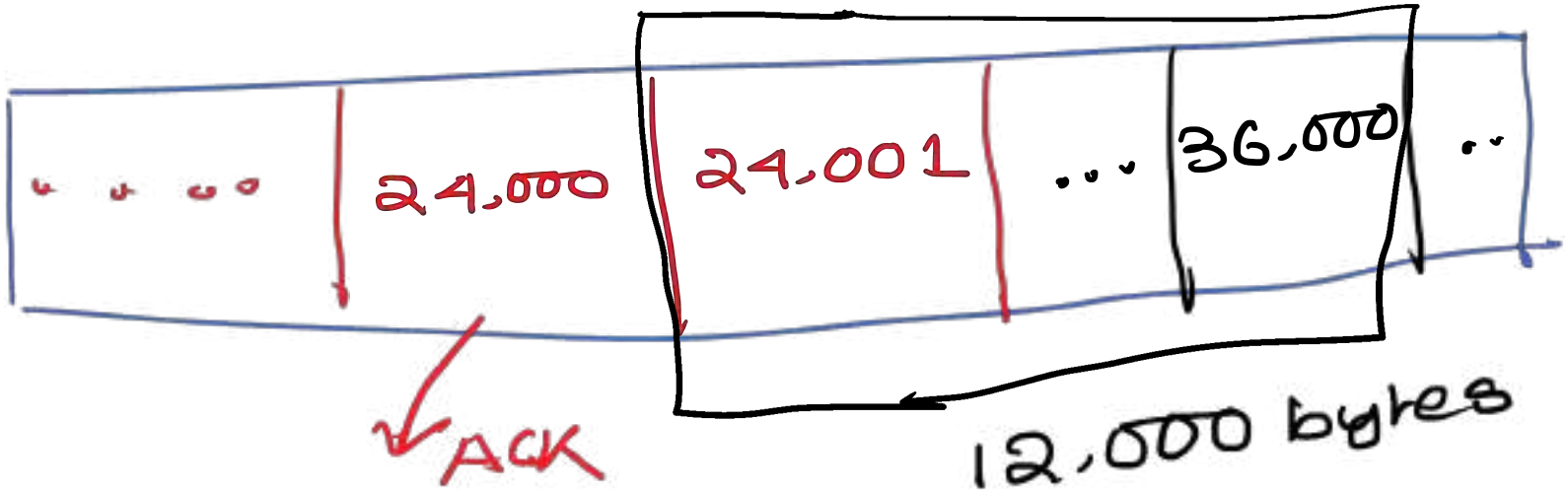


24,001
+ 12,000
36001
1

22,001
+10,00
32,00

22,000 @ 3

successful
receive
now the
receiver
expect
(22,000



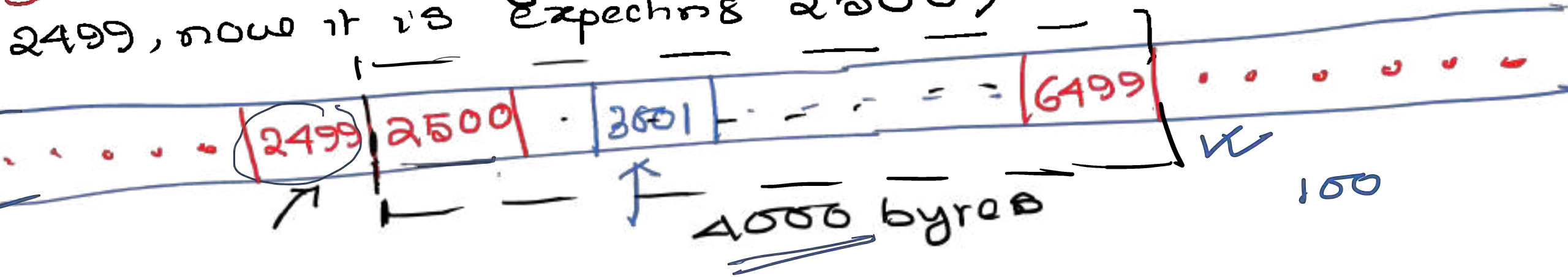
Questions. A window holds bytes 2001 to 5000. The next byte to be sent is 3001. Draw a figure to show the situation of the window after the following two events.

a. An ACK segment with the acknowledgement number 2500 and window size advertisement 4000 is received.

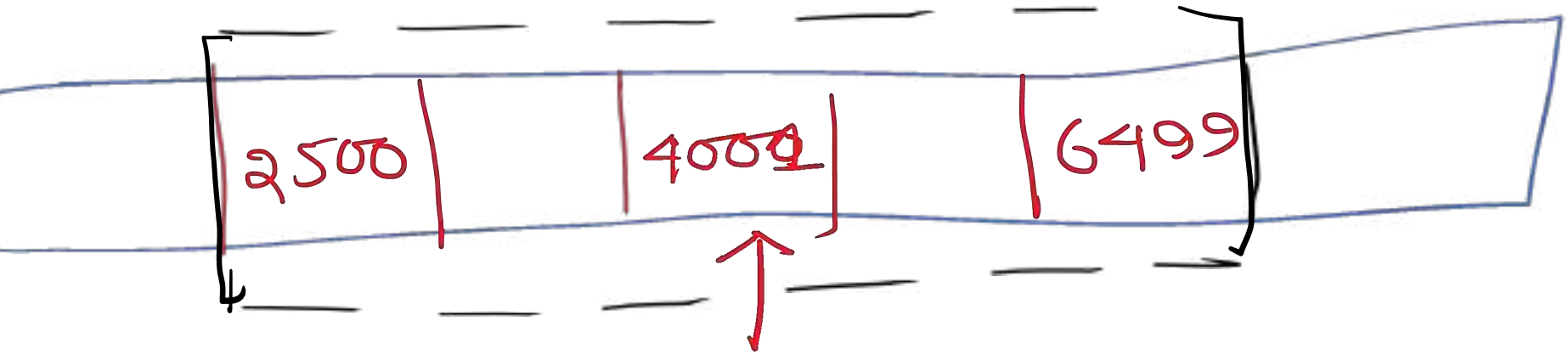
b. A segment carrying 1000 bytes is sent.

next byte to be sent

Case 1:- successfully received the segment with seq no 2499, now it is expecting 2500, window = 4000



100



next byte to be sent.

3001

1000 bo

30
+ 10
40

Congestion

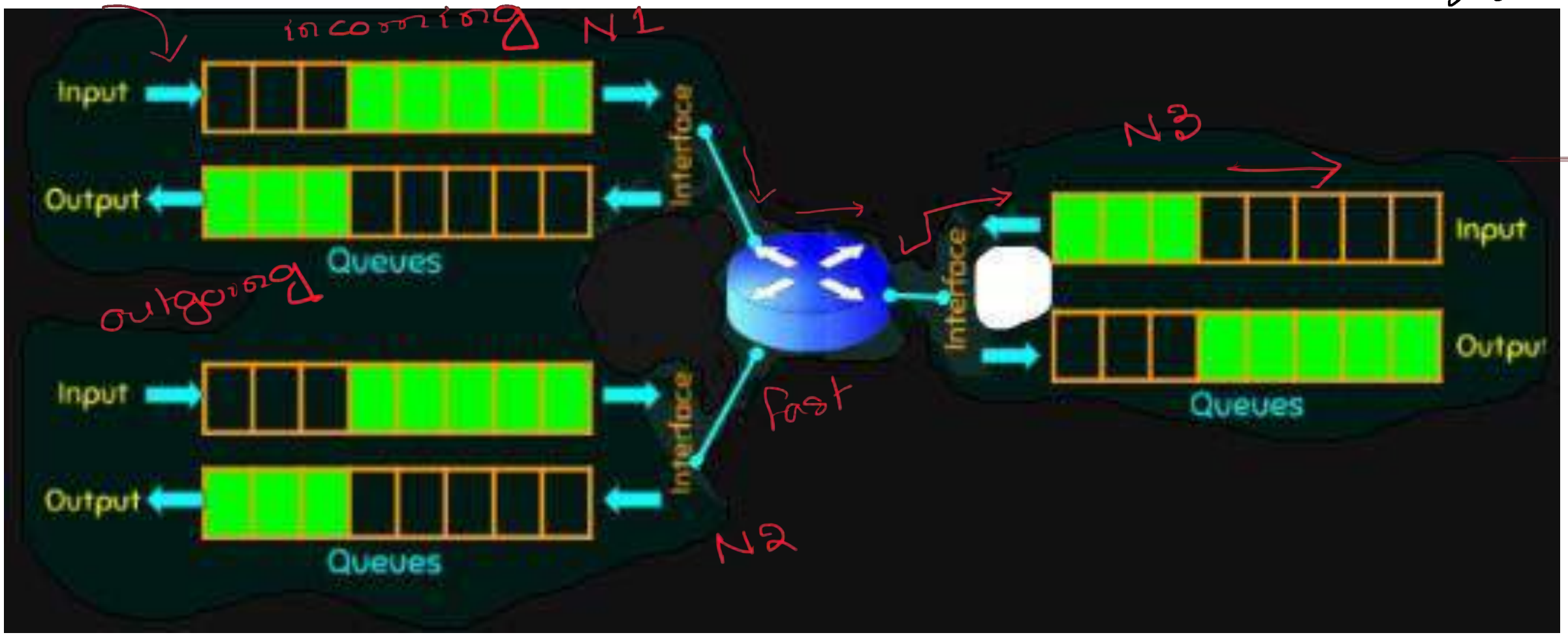
(cwnd, rwnd)

= 25 Kbps

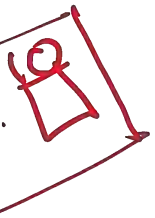
200Kbps

router

25 Kbp



Congestion Control



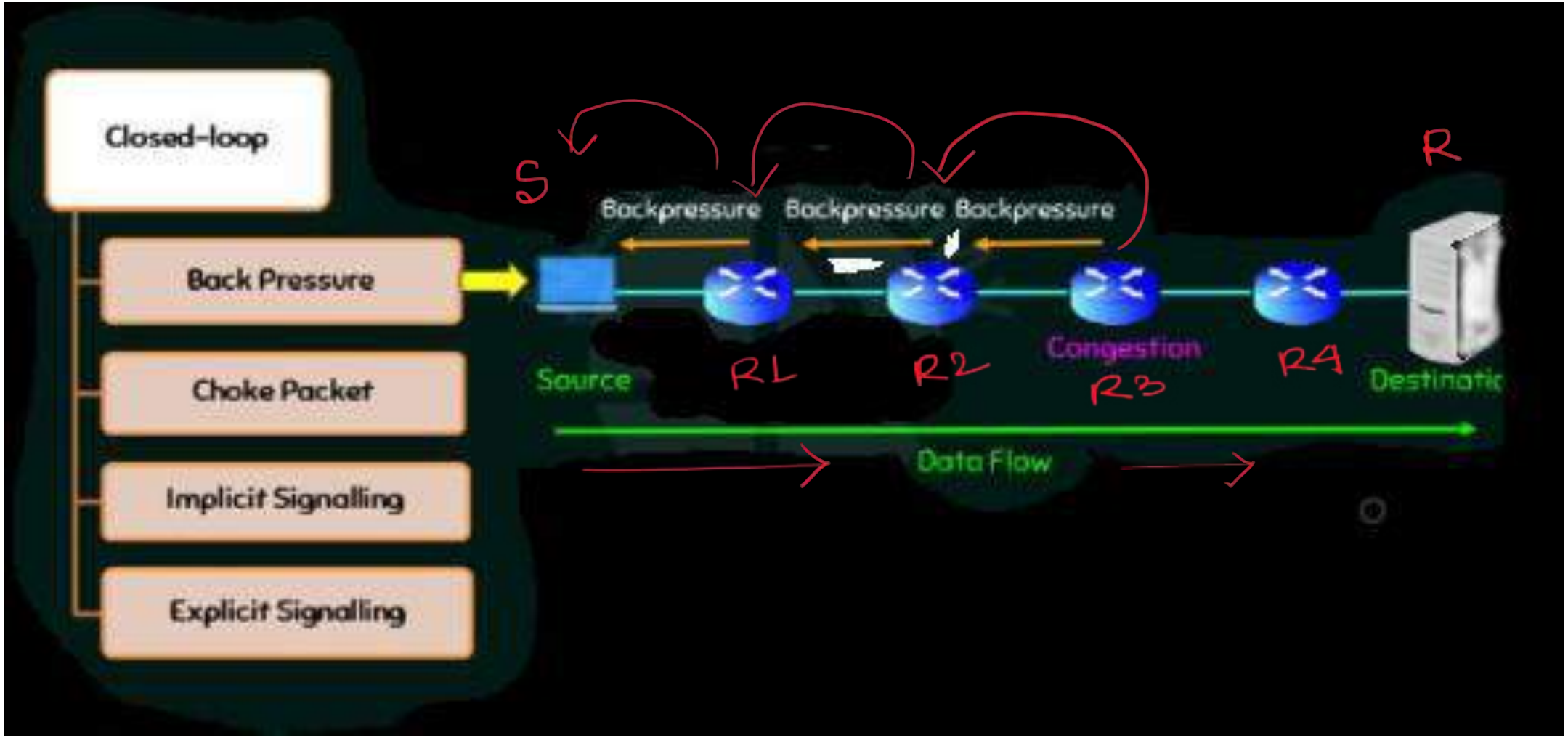
before congestion happens

prevention

(Removable methods congestion ha happ

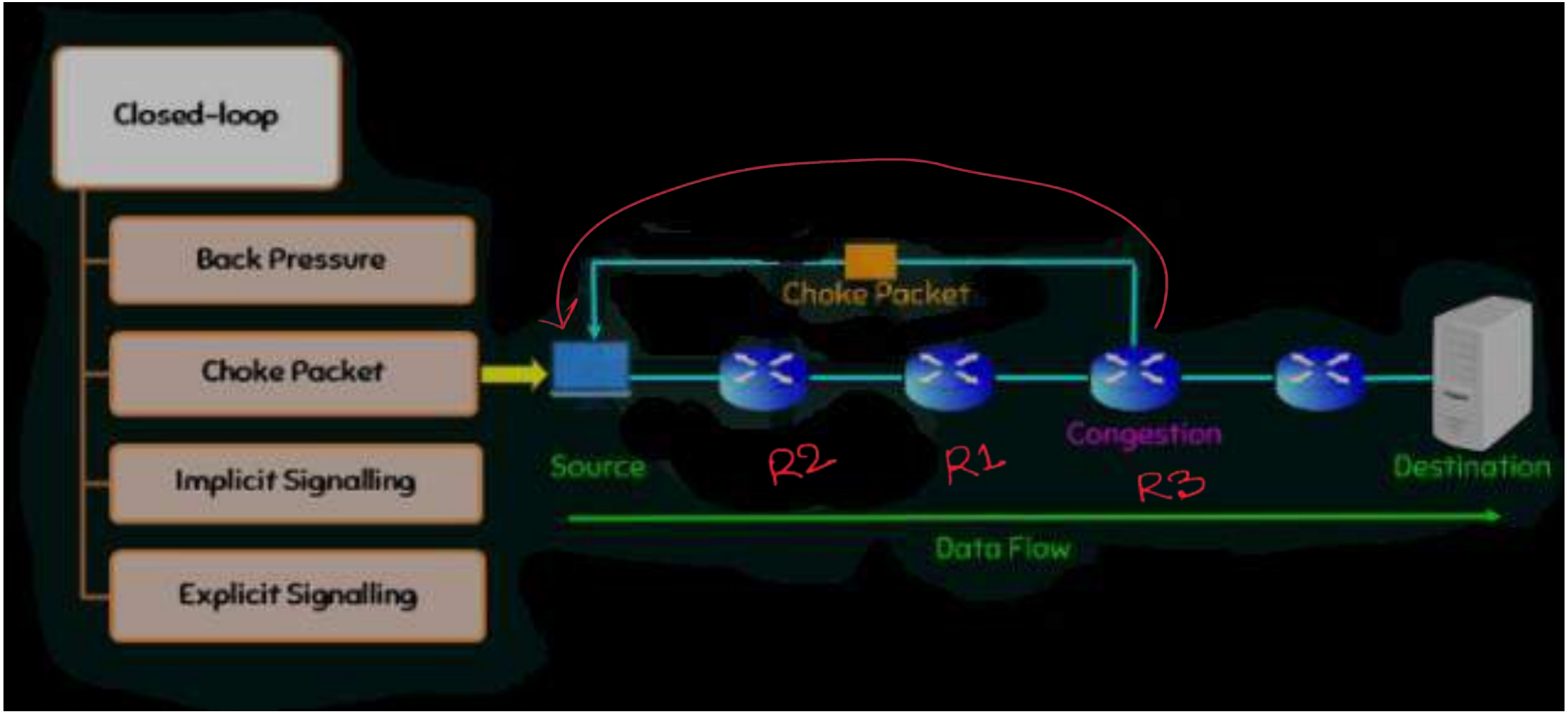
- Open-loop
  - Retransmission Policy
  - Window Policy
  - Acknowledgement Policy
  - Discarding Policy
  - Admission Policy
- Closed-loop
  - Back Pressure
  - Choke Packet
  - Implicit Signalling
  - Explicit Signalling

Congestion Control

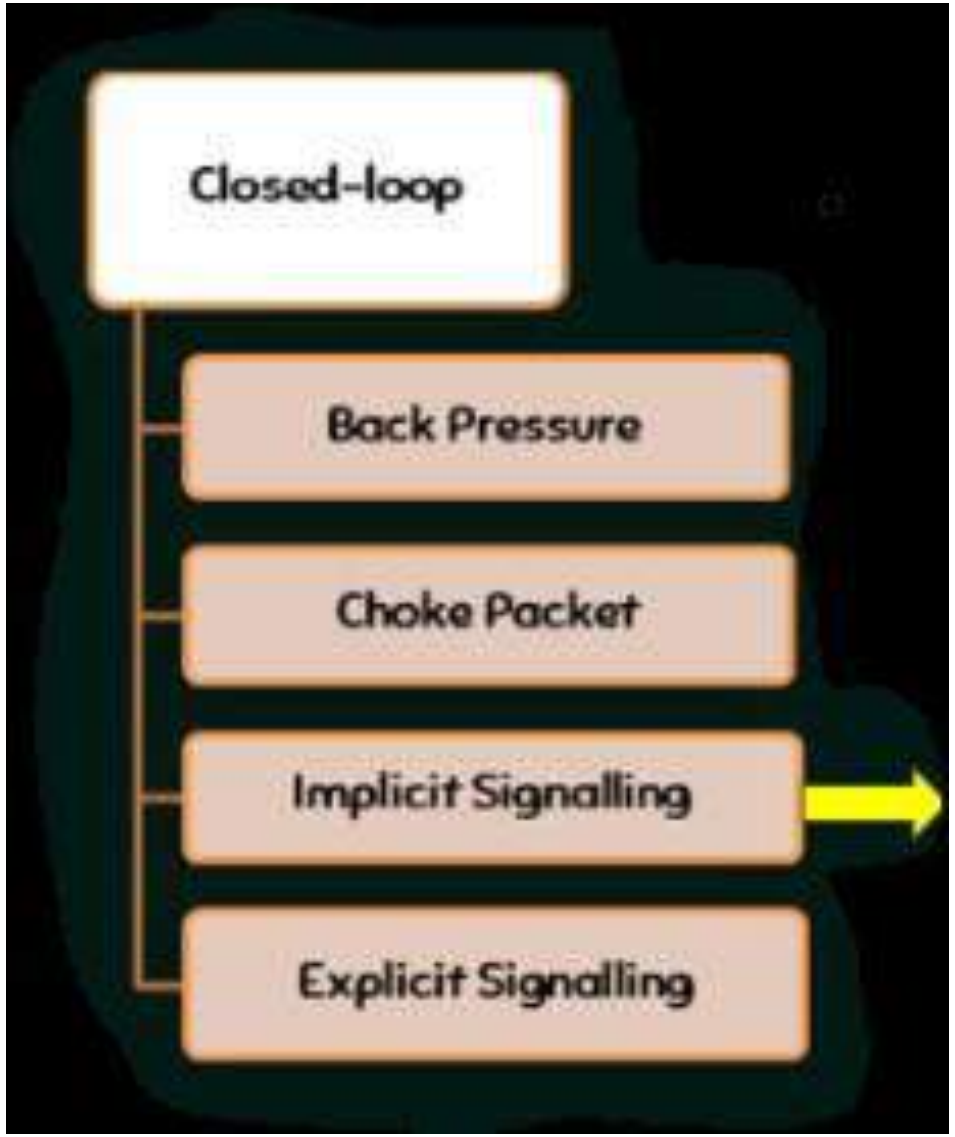


Congestion Control

ICMP

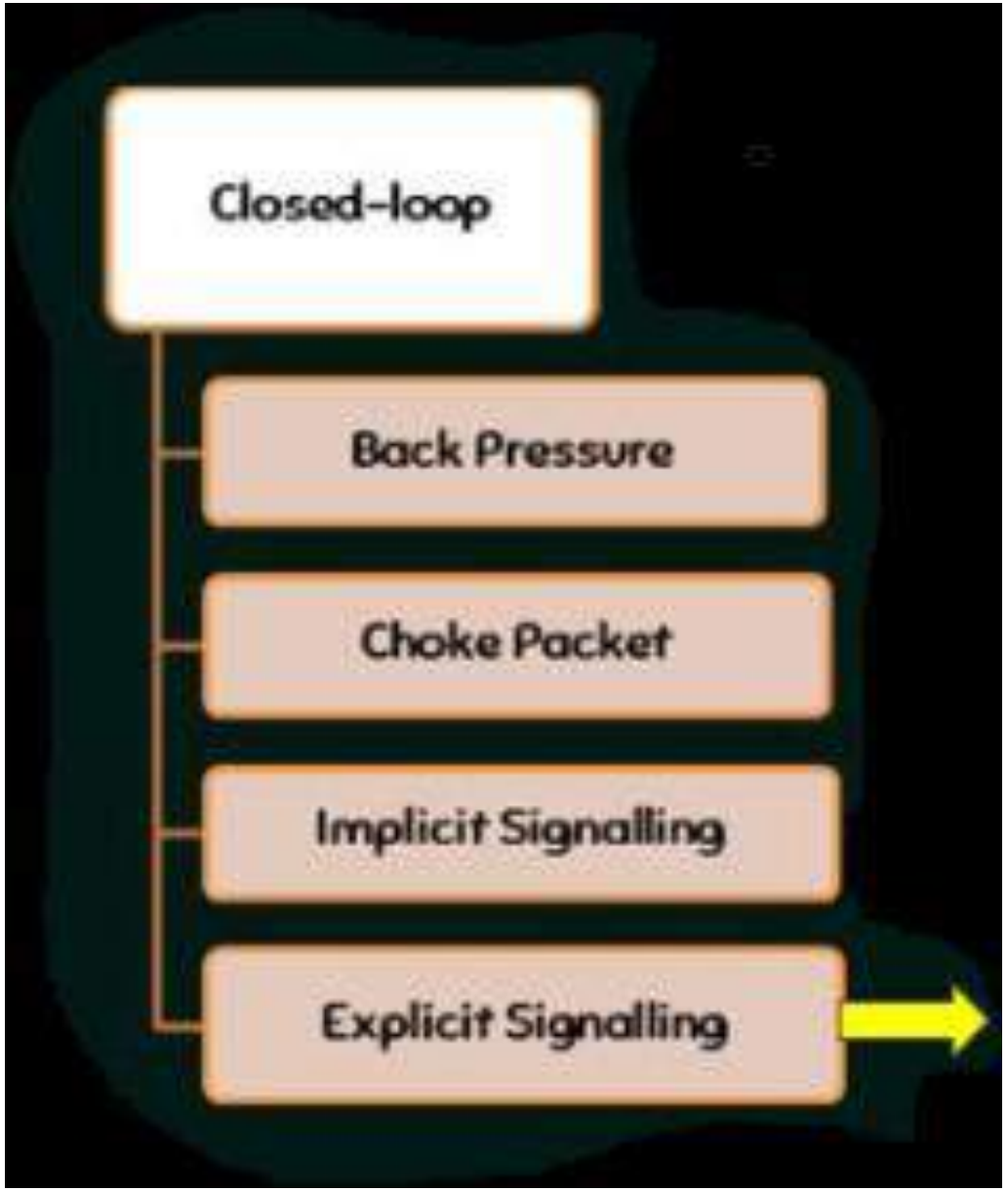


Congestion Control



- *There will be no implicit communication between the congested node and the source.*
- *The source will get to know about the congestion by no ACK packet received or delayed ACK.*

Congestion Control



- *There will be explicit communication between the congested node and the source.*
- *Here signals will be included in the data packet itself. Router can discard less sensitive packets that carry data.*

Congestion control in TCP

- *The sender window size is determined by the available buffer space in the receiver (rwnd).*
- *In other words, we assumed that it is only the receiver that can dictate to the sender the size of the sender's window.*
- *If the network cannot deliver the data as fast as they are created by the sender, it must tell the sender to slow down.*
- *The sender's window size is determined not only by the receiver but also by congestion in the network. TCP sliding window is byte-oriented.*
- *Actual window size = Minimum (rwnd, cwnd)*
- *TCP's general policy for handling congestion is based on three phases:*

- *Slow start Phase* ✓ → slow start
- *Congestion Avoidance* ✓ → Additive increase
- *Congestion Detection* ✓ → Multiplicative decrease

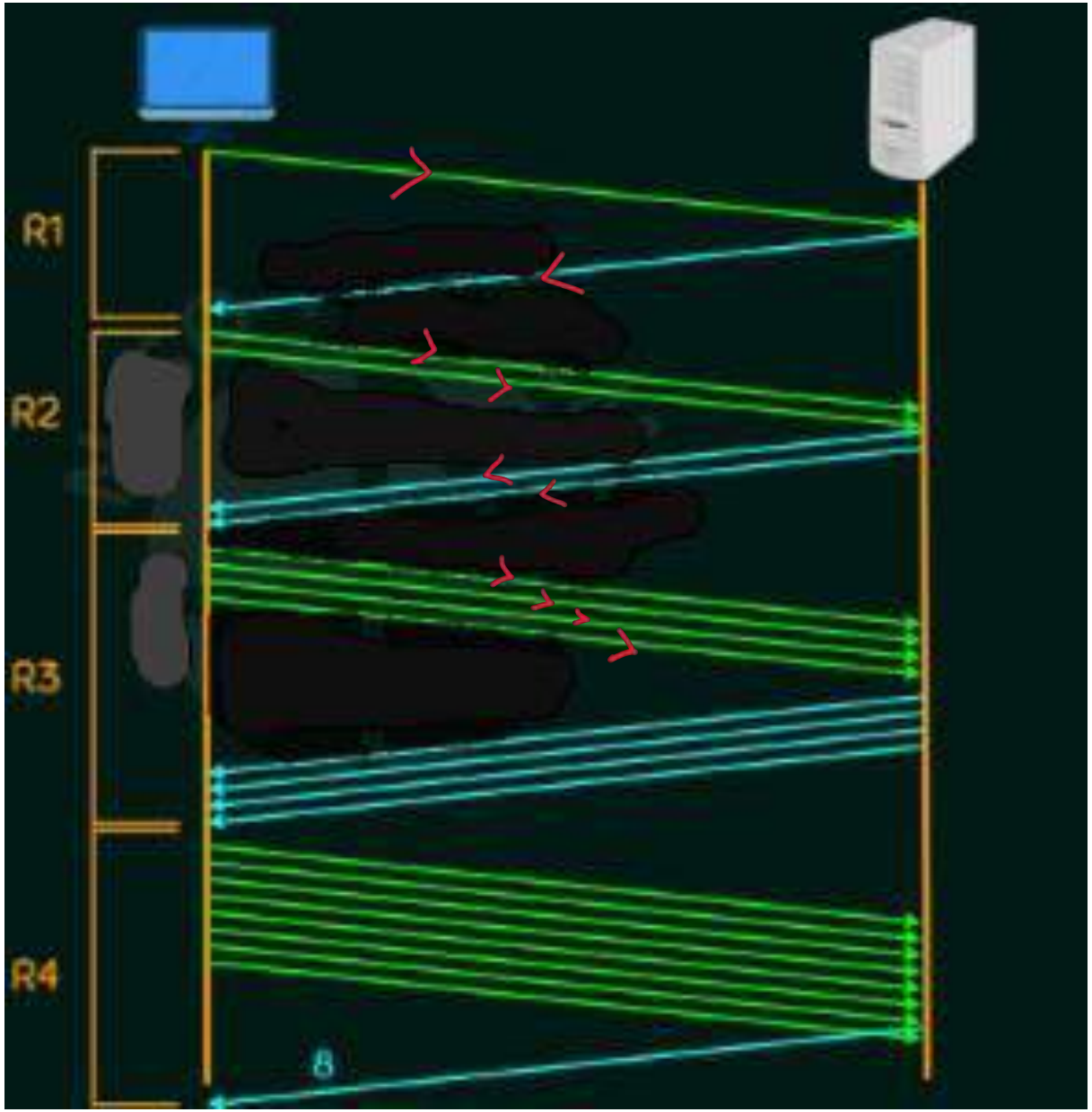
Congestion control in TCP

- *In the slow-start phase, the sender starts with a very slow rate of transmission, but increase the rate rapidly to reach a threshold.*
- *When the threshold is reached, the data rate is reduced to avoid congestion.*
- *Finally congestion is detected, the sender goes back to the slow-state or congestion avoidance phase based on how the congestion is detected.*

Congestion Policy

- *Slow start* (slow start phase)
- *Additive Increase* (avoidance)
- *Multiplicative decrease* (congestion detection)

Slow Start



Round	Value of cwnd
Start	$2^0 = 1$
After Round 1	$2^1 = 2$
After Round 2	$2^2 = 4$
After Round 3	$2^3 = 8$
After Round 4	$2^4 = 16$
After Round 5	$2^5 = 32$
After Round 6	$2^6 = 64$

ssthreshold

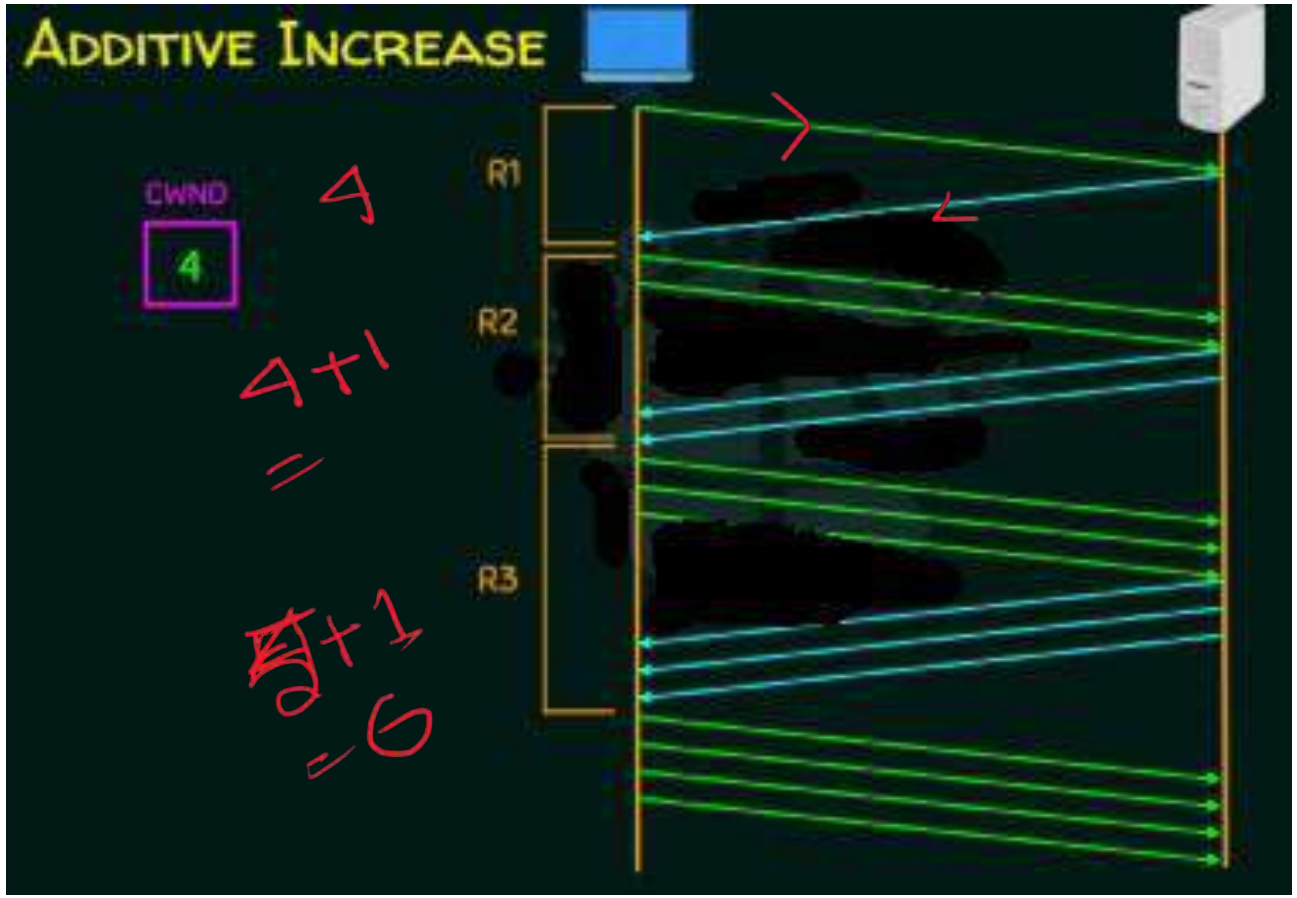
Congestion Policy

- *Slow start cannot continue indefinitely.*
- *There must be a threshold to stop this exponential growth.*
- *The sender keeps track a of a variable named ssthresh (slow-start-threshold).*
- *When the size of window in bytes reaches the threshold, slow start stops and the next phase starts.*
- *In most implementation, the value of ssthresh is 65,535 bytes.*

Congestion Policy

- *Slow start- cwnd increase exponentially.*
- *The exponential growth must be slowed down, to avoid congestion before it happens.*
- *Collison avoidance: Additive increase.*
- *When cwnd reaches the threshold, the slow-start phase stops and the additive phase begins.*
- *In this algorithm, each time the whole window of segments is acknowledged (one round), the size of the congestion window is increased by 1.*
- *In the congestion avoidance algorithm, the size of the congestion window increase additively until congestion is detected.*

Additive Increase



ADDITIVE INCREASE

Round	Value of cwnd
Start	1
After Round 1	2
After Round 2	3
After Round 3	4
After Round 4	5
After Round 5	6
After Round 6	7
	⋮

Multiplicative Decrease

- *Slow start- Exponential increase..*
- *Additive increase- Linear increase.*
- *The cwnd must be decreased , if there is a congestion.*
- *Sender can guess that congestion.*
- *Retransmission of segment.*
- *Retransmission can occur in one of two cases:*
  - *When a timer times out or*
  - *When three ACKs are received.*
- *Threshold is dropped to one-half, a multiplicative decrease.*

20

$\frac{1}{2} \times 2$

10

Multiplicative Decrease

TCP implementations have two reactions:

current cwnd = F

Reaction 1:

- *It sets the value of the threshold to one-half of the current window size..*
- *It sets cwnd to the size of one segment.*
- *. It starts the Slow-start phase again.*

25

cwnd =

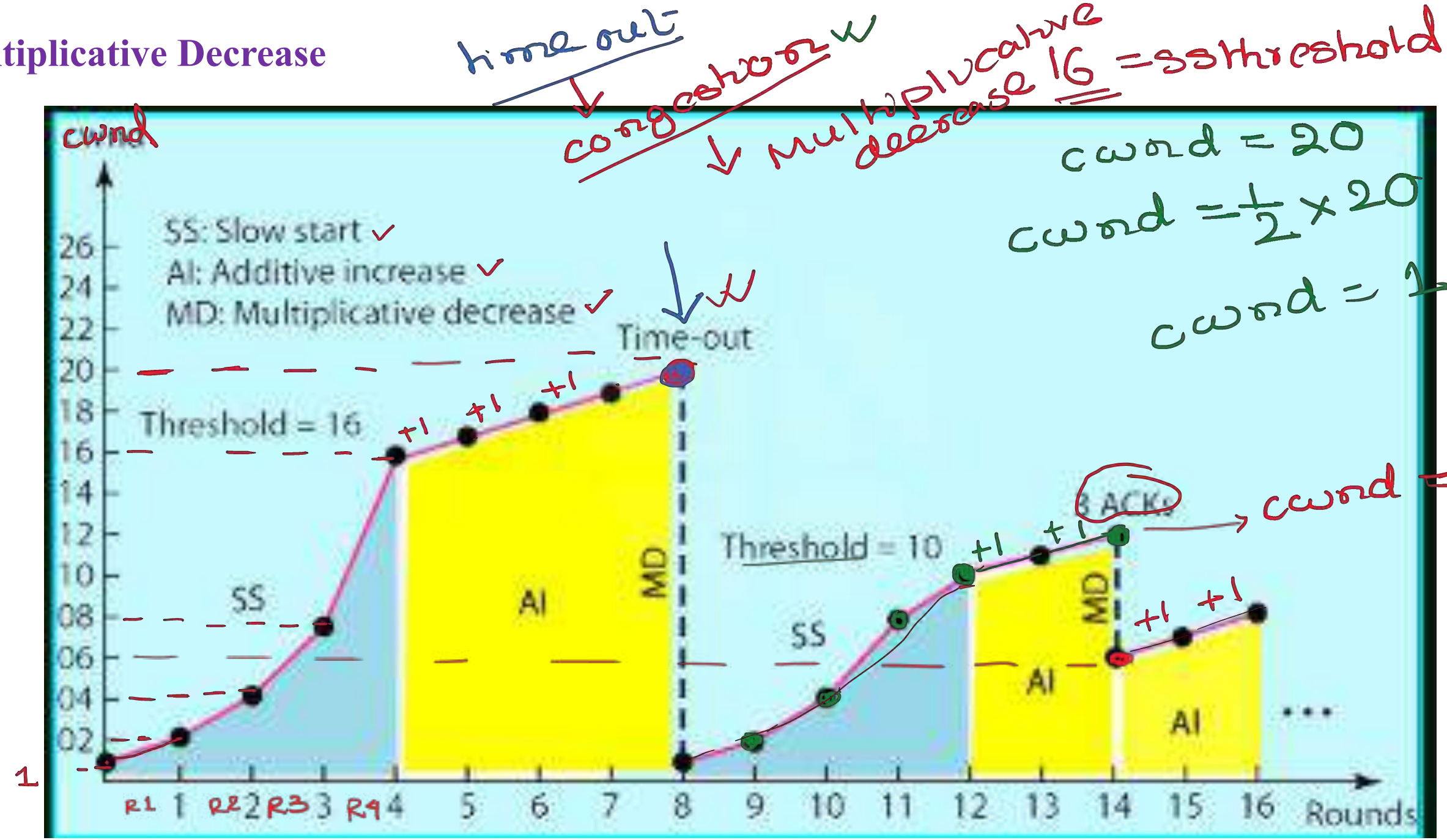
Reaction 2:

- *It sets the value of the threshold to one-half of the current window size.*
- *It sets cwnd to the value of the threshold.*
- *It starts the congestion avoidance phase.*

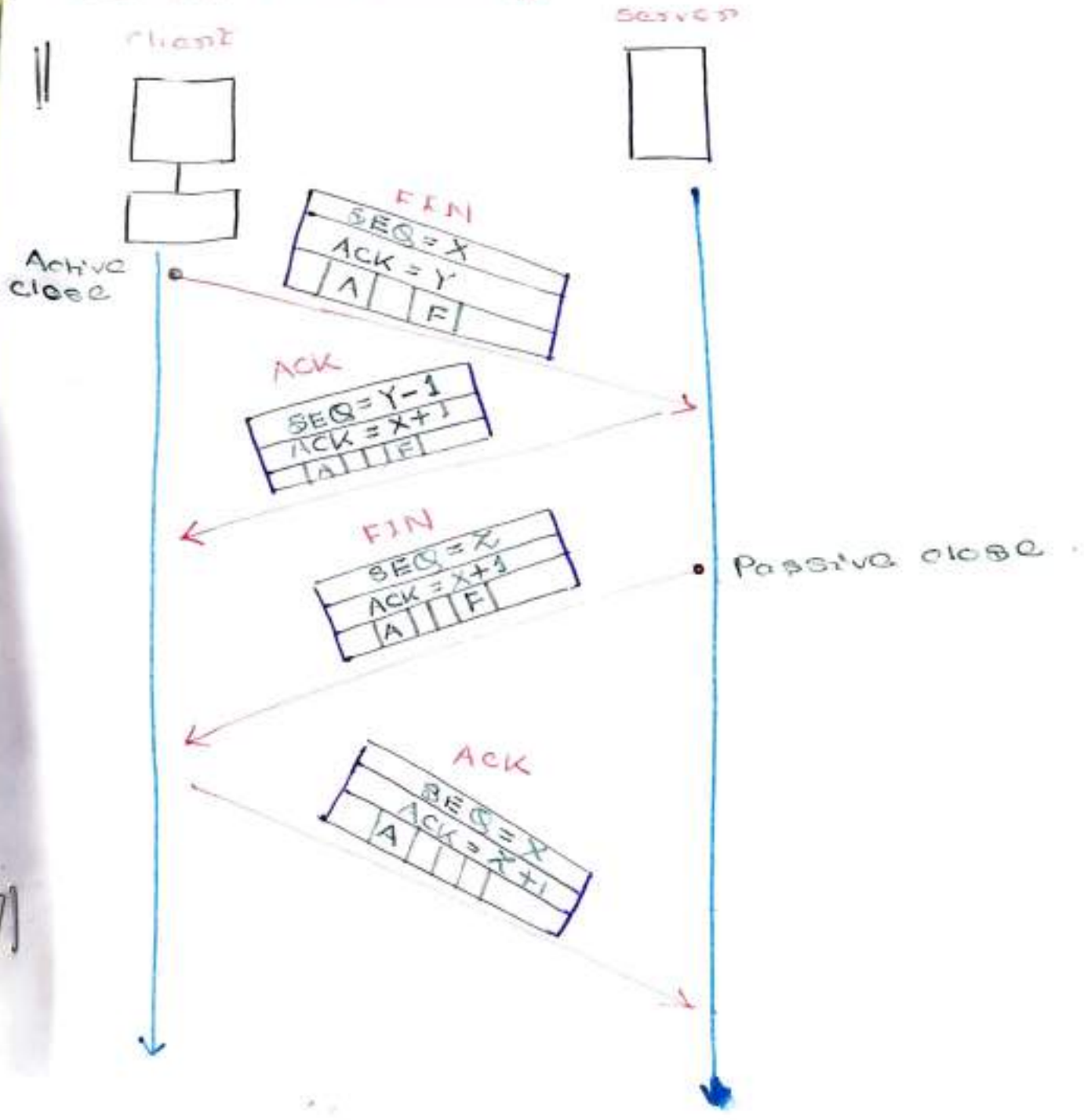
cwnd = 25 → +1 +1

additive increase

Multiplicative Decrease



Four way handshaking:-



Sliding window:- window is used for flow control at the receiver.

Flow control is usually done in two layers:-

1. Transport layer.

2. Datalink layer.

Now we are in Transport layer:-

- TCP uses a sliding window to handle flow control.
- From Go-back-N
- does not use NAK.
- The receiver holds out-of-order segments

Selective Repeat

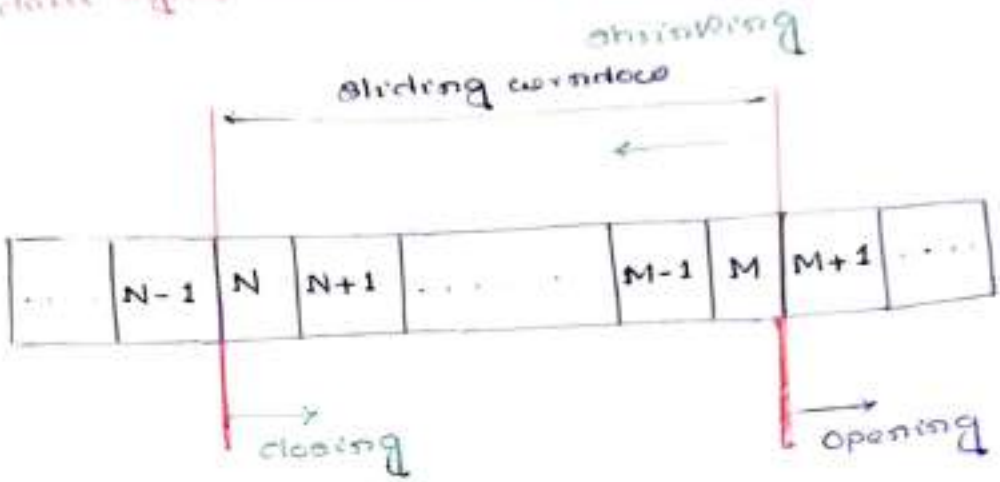
- TCP sliding window is of variable size. where as datalink layer sliding window is of fixed size.
- TCP sliding window is byte-oriented. whereas datalink layer sliding window is of frame oriented.

The window takes care about

① the data bytes which are sent

② the data bytes which are sent but not acknowledged

③ the data bytes which are yet to sent



window size = minimum (rwnd, cwnd)

rwnd → receiver window size

cwnd → congestion window size

The window size is of variable lenght.

The data which are sent but not have been acknowledge and the data which are to be sent immediately as will be in the window.

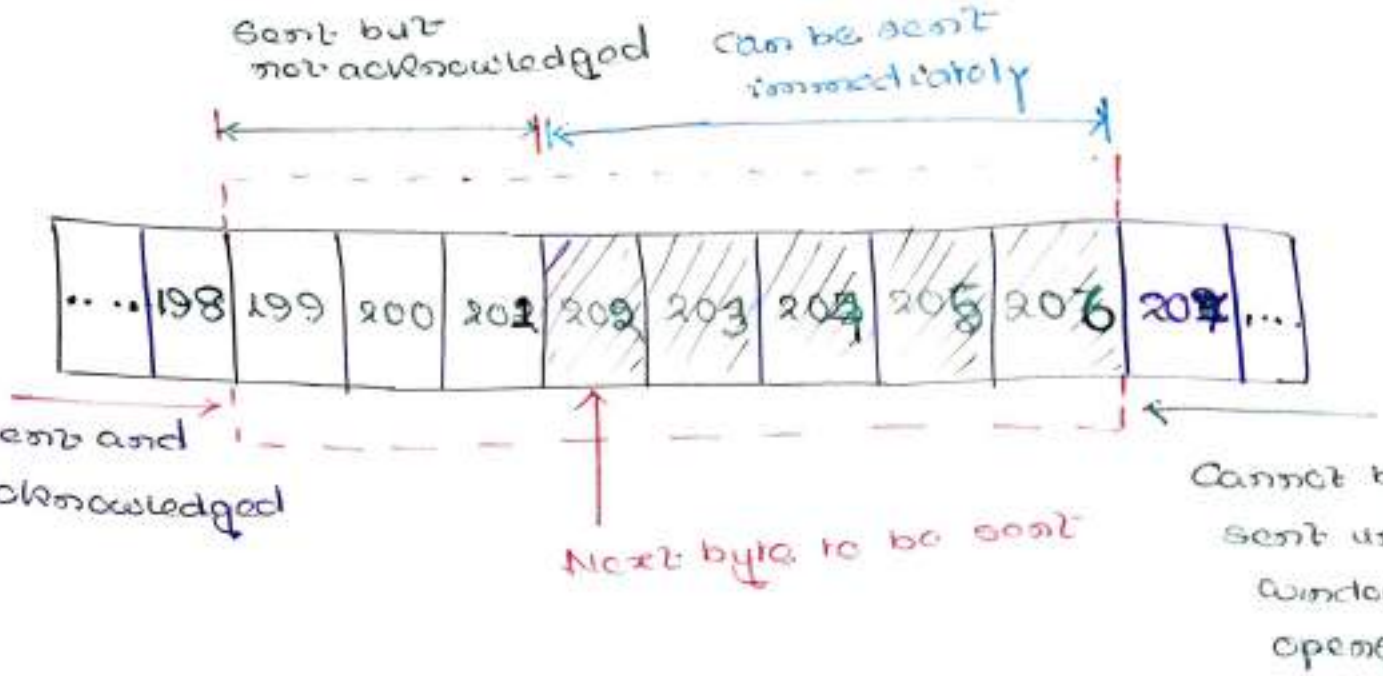
Suppose the N number segment is sent and get acknowledged now the window will move towards (N+1) and now the window can allow the (M+1) number segment to the the next segment to be sent in the queue.

• Congestion window is the ~~pre~~ feature of intermediary devices such as router or switches. suppose the receiver has a window size of 50. But the router through which many segments of other ~~ongoing~~ ongoing communications are processing, so there is a congestion.

window size = minimum (20, 8) = 8

receiver window size

congestion window size (size of network window)



- window size = min(rwnd, cwnd).
- The source does not have to send a full window's worth of data.
- The window can be opened or closed by the receiver but should not be shrunk.
- The destination/receiver can send an ACK at any time
- The sender can send 1-byte segment even after the window is shut down in the receiver side.

Q. What is the size of the window for host A if the value of rwnd (receiver window) is 3000 bytes and the value of cwnd (congestion window) is 3500 bytes?

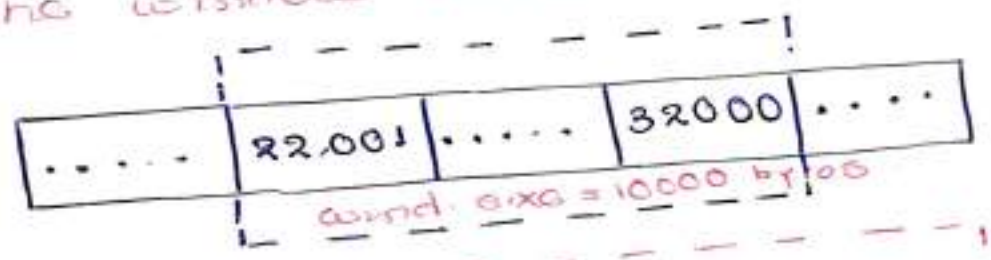
⇒ We know, window size = min (rwnd, cwnd)

= min (3000, 3500)

= 3000 bytes

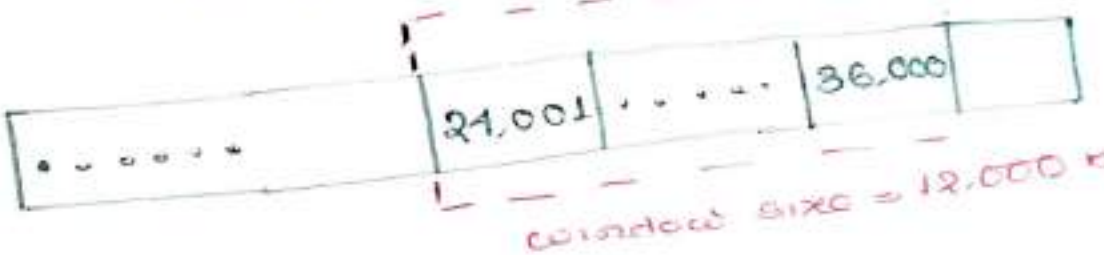
Q. A TCP connection is using a window size of 10,000 bytes and the previous acknowledgement number was 22,001. It receives a segment with acknowledgement number 21,001 and window size advertisement of 12000. Draw a diagram to show the situation of the window before and after.

$$\begin{array}{r} 22001 \\ +\ 10{,}000 \\ \hline 32{,}001 \\ -\ 1 \\ \hline 32{,}000 \end{array}$$



before receiving the segment

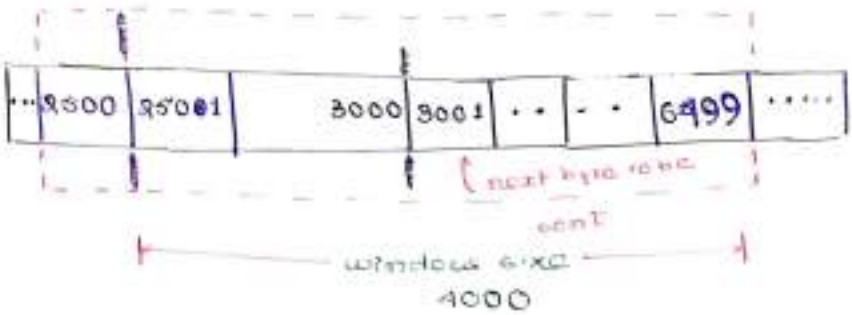
$$\begin{array}{r} 21{,}001 \\ +\ 12{,}000 \\ \hline 36{,}001 \\ -\ 1 \\ \hline 36{,}000 \end{array}$$



Q. A window holds bytes 2001 to 5000. The next byte to be sent is 3001. Draw a figure to show the situation of the window after the following two events:-

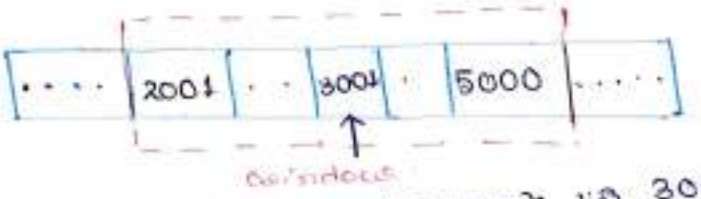
(a) An ACK segment with the acknowledgment number 2500 and window size advt. 1000 is received.

(b) A segment carrying 1000 byte is sent.



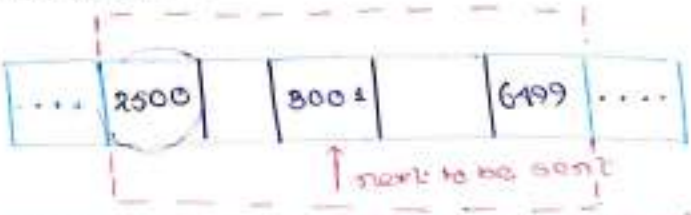
$$\begin{array}{r} 2501 \\ +\ 4000 \\ \hline 6501 \\ -\ 1 \\ \hline 6500 \end{array}$$

Initial figure:- The window holds byte 2001 to 5000



The next data byte to be sent is 3001

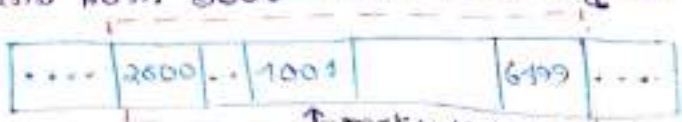
An acknowledgement for 2500 is received and the new window size is 4000.



$$\begin{array}{r} 2500\checkmark \\ +4000\checkmark \\ \hline 6500 \\ -\ 1 \\ \hline 6499 \end{array}$$

* A segment carrying 1000 byte is sent. means from 3001 next 1000 bytes is sent

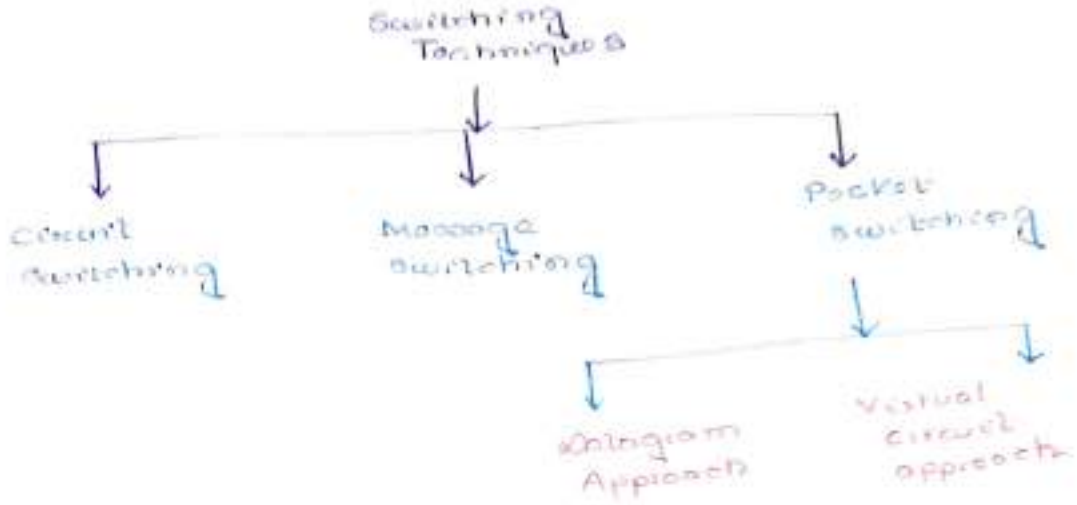
$$\begin{array}{r} 3001 \\ +1000 \\ \hline 4000 \end{array}$$



From 3001 to 4000 next byte sent.

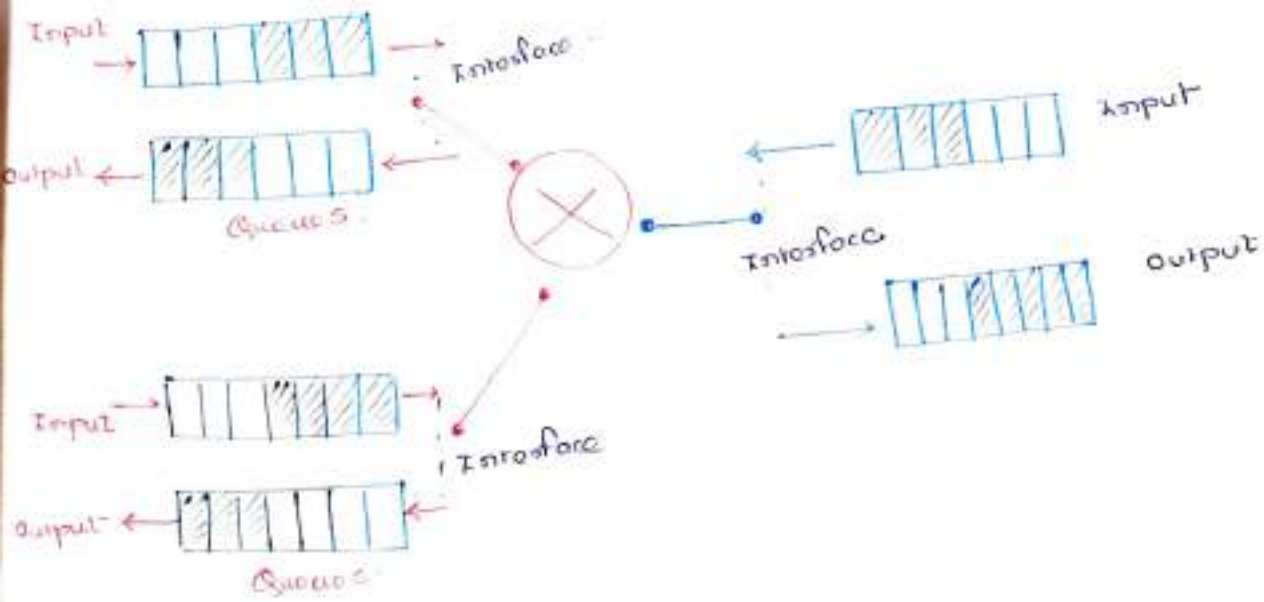
Congestion:-

We know there are three types of switching



Internetwork communication is not possible without routers. But routers has limited memory capacity. Routers has memories in form of queues.

Suppose a movie hall has limited amount of seats. As there are no prior seat booking is allowed. If the number of people arrived at the hall is large than the no. of available seats then the congestion occurs. Same way congestion in computer network occurs at the routers/switches as routers maintains buffering - queues before and after processing the data.



(1) The packets are put at the end of the output queue

(2) Processing the packets have the same. The packets must be sent in the order/ correct sequence. For this the routing has to follow the routing rules

(3) The number of packets in the output queue

- When packets are added at the output queue it has to find more resources. If the output queue is full then the packets processing will have the same resources. Expanding size of queue is impossible. Hence packet loss will occur. Hence more will be congestion at the output queue

Congestion control :-

(3)

Before happening (Preventive method)	After happening (Removing method)
→ Retransmission policy	→ Back pressure
→ Windows policy	→ Choke packet
→ Acknowledgement policy	→ Implicit signalling
→ Discarding policy	→ Explicit signalling
→ Admission policy	

Open loop congestion control mechanism :-

- Retransmission policy :-

- As reliability is ~~gu~~ guaranted in TCP so till the time acknowledgement is received by the receiver the data packets will be retransmitted.

- Retransmission will also be done incase of loss or corrupted or damaged packets

- But this will increase congestion.

- A good retransmission policy can help
- The TCP's retransmission policy is designed to prevent or alleviate congestion

Window Policy:-

① window at sender may also affect congestion
Suppose in a busy or congested network if the window size of the sender is larger than sender will be sending more no. of data packets which can again make the network more congested

② SR (Selective Repeat) windowing is more better than Go-back-N window
In Go-back-N if one segment is corrupted or missing or damage then the entire window is retransmitted, whereas in Selective Repeat window only the missing/corrupted segment is retransmitted again.

③ This duplication in Go-back-N may worse congestion.

④ Selective Repeat retransmits only lost or corrupted packets.

Acknowledgement Policy:-

① Ack policy imposed by the receiver may also affect congestion.
Suppose there are 10 (N=10) packet if the network is already congested and acknowledgment has to be sent for each of the 10 packets then the network will be loaded by both data and also with acknowledgment packets.

② so either then sending Ack for each individual packet if acknowledgement can be sent for N=10 packets together, than fewer Acks impose less load.

Discarding policy:

- A good discarding policy if less sensitive data packets can be discarded than by discarding focus less important or sensitive data packets where the integrity of the data sent is still acceptable, then this type of discarding policy can make the network less congested.

- Ex: audio transmission

Admission policy:-

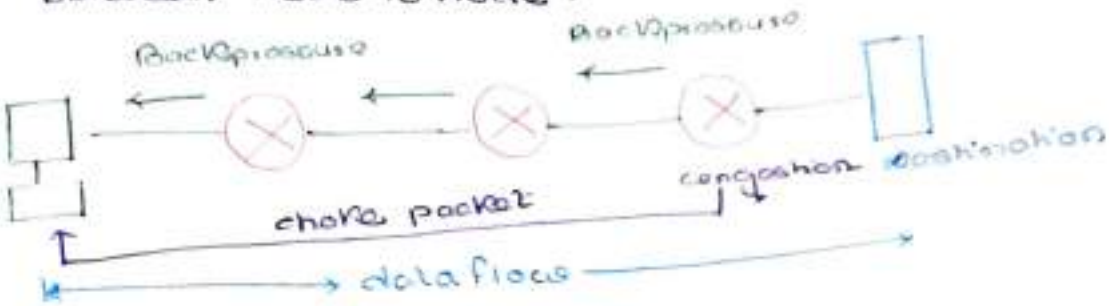
- QoS (Quality of service) mechanism prevents congestion in virtual circuit network. (virtual ckt network means no physical connection is established before transmission of data packets).

- Switches in a flow first check the resource requirement of a flow before admitting it in the network.

- Routers can also deny if there is a possibility of future congestion.

closed-loop congestion control:-

→ Backpressure mechanism is a virtual connection control mechanism ~~in node~~ between node to node.



Cannot be applied to datagram.
Opposite direction to the data flow.
It will inform the ~~up~~ previous upstream node about the congestion in one node.

choke packet:-

choke packet is not a node-to-node mechanism.
choke packet is directly sent between the congested node and the source. It is a special type of packet whose it will not pass through the intermediatory node and will be sent directly to the source for flow control. choke packet ~~ICMP message~~ will take help of ICMP messages.

Implicit signalling:-

There will be not implicit communication between the congested node and the source to between the congested node and intermediate nodes. The source will get ~~to~~ to know about the congestion by No ACK (acknowledgment) packet received or delayed ACK.

Explicit signalling :-

There will be ~~no~~ explicit between congested node between and source. Here signals will be included in the data packets itself.

- Routers can discard less sensitive packets that carry data.
- There are two types of Explicit signalling

(1) Backward signalling.

(2) Forward signalling.

Congestion control in TCP -

Scenario

Congestion occurs to the input queue when the packet arrival rate is larger than the processing rate of the router.

Congestion occurs at the output queue when the packet despatch rate is slower than the processing rate of the router.

In both the cases there will be congestion at the router.

Who decided the window size?

window size = min (rwnd, cwnd)

rwnd → receiver window size

cwnd → router window size

which has set by the sender after consulting to the intermediary device (e.g. router) and the receiver.

Congestion control in TCP:-

1. Slow start
2. Additive increase
3. Multiplicative decrease.

] Three congestion policy

Slow Start Algorithm :-

① The size of the congestion window (cwnd) starts with one maximum segment size (mss)

② The MSS is determined during TCP connection establishment.

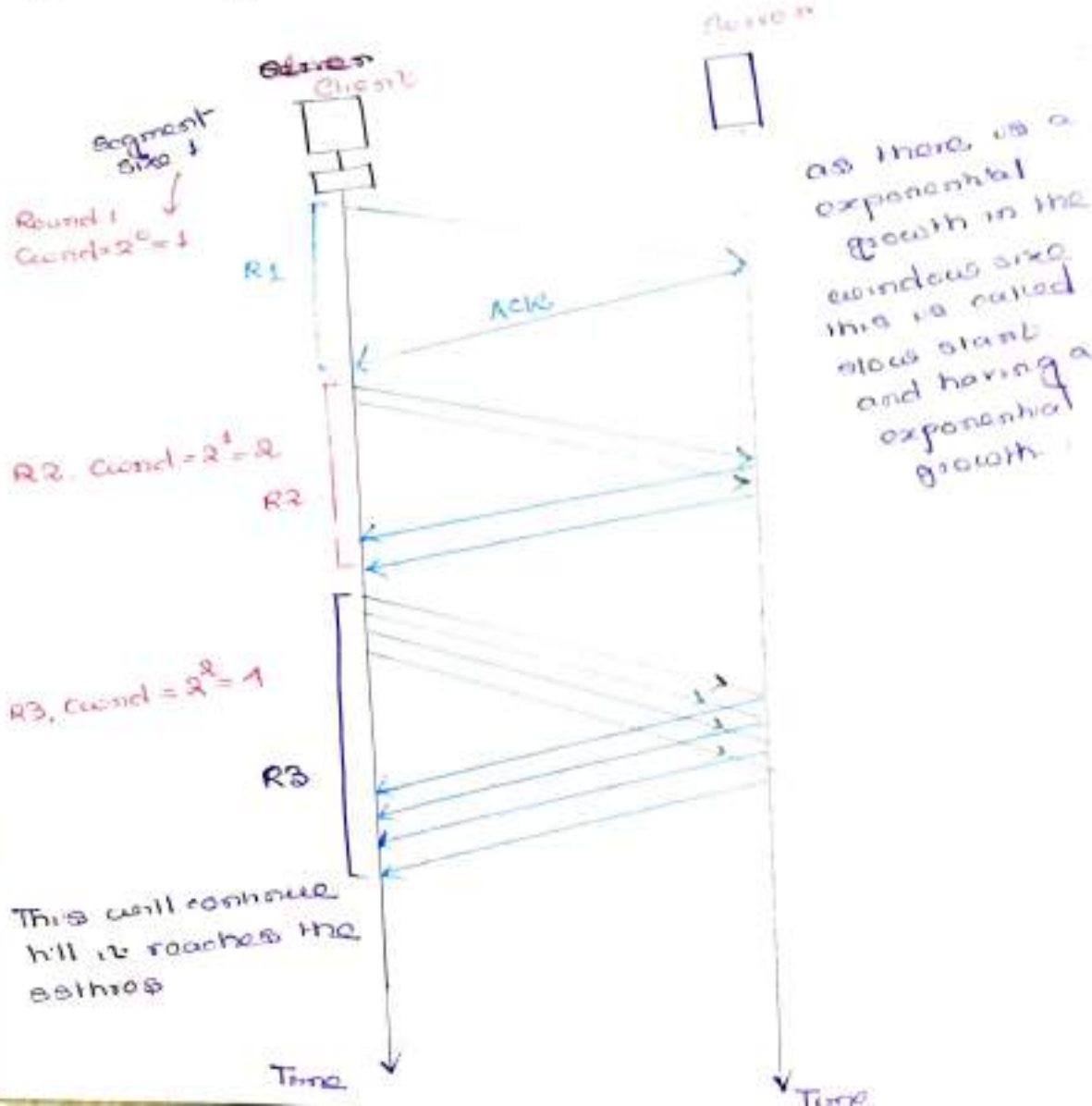
③ The window starts slowly but grows exponentially

(4) In the slow start algorithm, the size of the congestion window increases exponentially until it reaches a threshold.

Assumptions:-

① The value of rwnd > cwnd and hence the window size = cwnd

② Each segment is acknowledged individually



Case 1: Exponential growth may not

1. Slow start control continue exponential
There must be a threshold to stop this exponential growth.

2. The sender keeps track of a variable named ssthresh (slow-start threshold) which the size of the window in bytes reaches it's ssthresh. Now slow start stops and the slow phase starts (i.e. the Additive increase algorithm starts)

3. In most implementations the value of ssthresh is $(2^{16}-1)$ i.e. 65,535 bytes.

Additive Increase Algorithm :-

Congestion avoidance – Additive Increase

There are three phases in congestion control in TCP

1. Slow start
2. Additive Increase (Congestion Avoidance)
3. Multiplicative Decrease (Congestion Detection)

In slow-start process cwnd increases exponentially in power of 2. This exponential growth must be slowed down, to avoid congestion before it happens.

Now for congestion avoidance additive increase will be implemented. When cwnd reaches the threshold the slow start stops and the additive phase begins.

In this algorithm, each time the whole window of segments is acknowledged (one round) the size of the congestion window is increased by 1.

Suppose in the earlier diagram (time diagram). Suppose the sstheshold is 4, i.e. after receiving the acknowledgement of round 3 i.e. R3 now the next window size will not be 23 (i.e. 8) it will be $(2^2 = 4+1)$ i.e. 5.

In the congestion avoidance algorithm, the size of the congestion window increases additively until congestion is detected.

Congestion detection - multiplicative decrease

Slow start - exponential increase in cwnd

Additive increase - linear increase " "

The cwnd must be decreased, if there is a congestion

Now sender can guess if there is a congestion if the ack message is timed out or delayed

Retransmission of the segments have to be done

Retransmission can occur in two cases :-

1. when a timer times out or
2. when three ACKs are received.

- Threshold is dropped to one-half, a multiplicative decrease. when the cwnd reaches the sthres (suppose the value is 20) and congestion is sensed by the sender and retransmission has to be done, then this time the ~~cwn~~ cwnd size will be decrease to $\left(\frac{1}{2} \times 20 = 10\right)$

Now after congestion occurs TCP implementations have two reactions :-

Reaction 1 :-

1. It sets the value of the threshold i.e. sthres to one-half of the current window size
2. It sets cwnd to the size of one segment i.e $2^0 = 1$
3. It starts the slow-start phase again.

Reaction 2 :-

1. It sets the value of the threshold i.e. sthres to one half ~~of~~ of the current window size.
2. It sets the cwnd to the value of threshold
3. It starts the congestion avoidance phase i.e. the additive increase, that is the cwnd will be increase by 1 after each round.