TCP/IP.

Application Layer. Message.
Transport Layer. Segments.
Network Layer. Packets.
Data Link Layer. Frames.
Physical Layer. Bits.

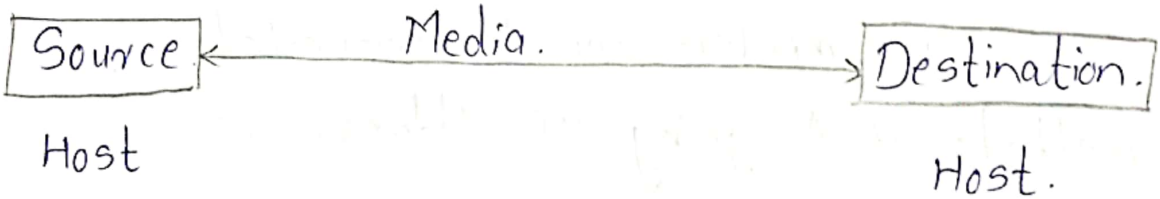
OSI Model.

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

Routing Protocols:

Static
Dynamic.

Devices are connected together to form a network.



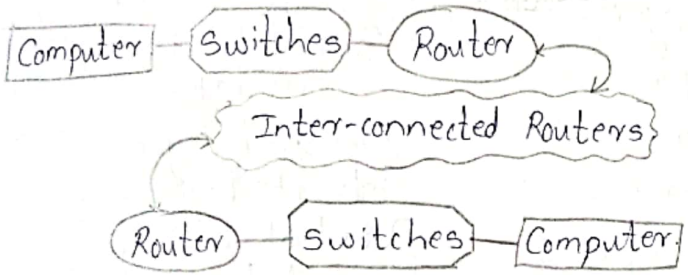
Host
Host.

Apart from media, address is required.

- Types of Addresses:- (ip a) (ipconfig -all) (ifconfig -all).

IPV4	IPV6.	MAC Address.
32 bits	128 bits	48-bits.
		Ethernet Address.
192.168.5.18	50b2:6400:::::1009.	Inbuilt in Card.
		(NIC Card, LAN Card, Ethernet Card).
		In HEX. Unique.

IP: ___ ___ ___ ___
8 bits, each.
written in decimal. (0-255).



Gatway is required, to send packet from network.

Subnet is used to Connect/Make/Establish the Network.

Subnet Mask: /24 Network. 255.255.255.___
8 + 8 + 8 = 24 bits.
Hence, the name.

Remaining 8 bits, $\longrightarrow 2^8 = 256$ Addresses.
2 Addresses are reserved for:
- Network Address
- Broadcast Address.

$\therefore 2^8 - 2 = 254$ Computers can be connected.

Two methods of Assigning IP Addresses:
- Static
- Dynamic.

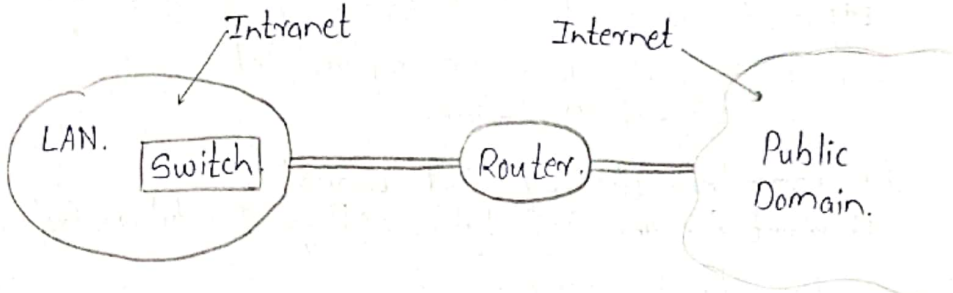
MAC Address is given by hardware.

Devices, Subnet, Gateway } → LAN. # We can ping to different levels to check whether they are working fine or not...

DNS → To go outside the network.

To connect two different Network, "Router" is used.

Switch is used inside the LAN.



Gateway is Connected to Router.
Then we have ISP Gateway.
Between ISP and the Router, Fibre optic Cable is used. Not CAT 6. (CAT 6 length ≤ 100 m)
Then we have DNS.

* Static IP's are used for Servers.
* Dynamic IP's are used for Devices.

To allocate Dyramic IP's, we require DHCP Servers.
We can assign Static IP for devices using class C (on our own, manually).

DHCP will assign IP on Lease, basis, to device.

eg. 210.212.183.43 Public Static IP → Mapped to phdadmissions.org.in
172.160. .... Private Static IP.

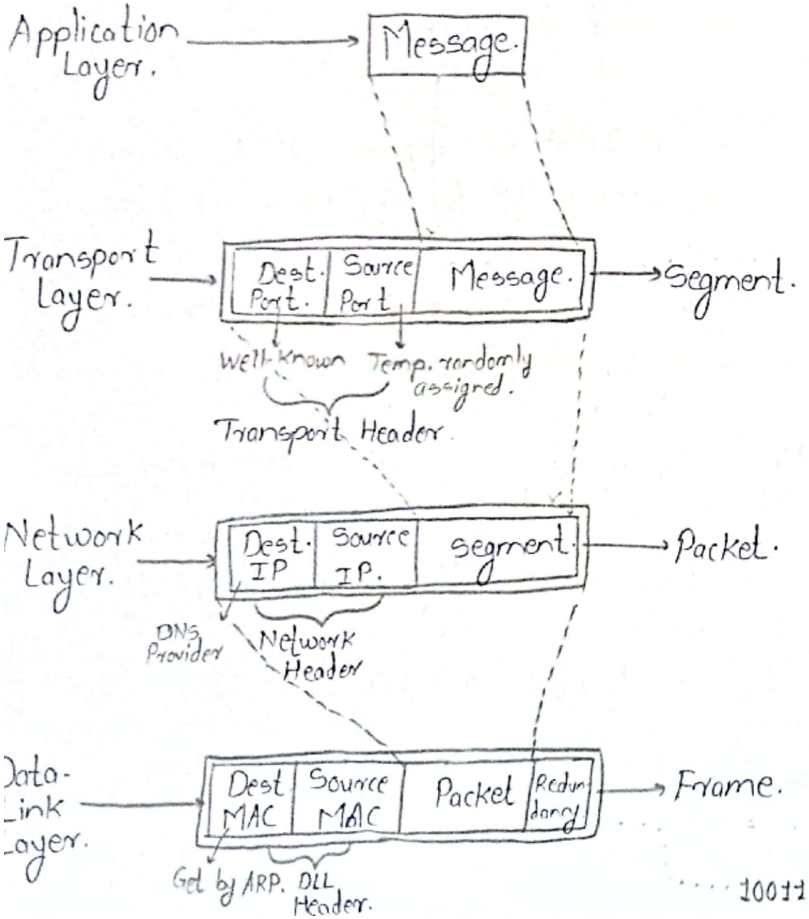
Fibre Optic: Multimode. (OM1, OM2, OM3, OM4).
Single Mode.

In College we have 1200 m Cable, from College to Hostel.
(Single mode fibre optic).
Cable terminates at D-block (D-109/109).

UTP Cable. Cat 6 → range 100 m.

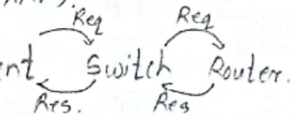
5 Layers Reside inside End devices (inside OS).
In switch, only 2 layers are implimented.
Physical and Data-link.
Switch Remembers the ~~IP~~ MAC. It "Learns."
It maintains an ARP table. (Address Resolution Protocol).
Switch forwards the frame.
Router routes the packets.

TCP/IP Model.



Communication between two devices is through MAC Address.

To get the MAC Address of gateway, we use Address Resolution Protocol. (ARP).



Client sends the ARP request.
Switch has it's ~~ARP~~ MAC table.
Switch learns the MAC Address of Client:
(Maintains the table).

Router maintains it's own ARP table and Routing table.
Router writes the MAC address of the switch/client.
It sends back the ARP response.
Switch writes the MAC address of gateway/Router.
Then the response is sent back to client.

Now client has MAC address of Gateway, and now it will generate the frame.

Client sends the frame. → Switch → Router → Internet.

If router don't have MAC of destination, then again we make ARP request to the next Routers.
The Process Continues.

Thus, we as (Client) don't have the MAC address of the Destination Server.

→ Hub Broadcasts.
→ Router Routes Packets.
→ Switch forwards Frames.

Host → ARP Table.
Switch → MAC Table.
Router → ARP + Routing Table.

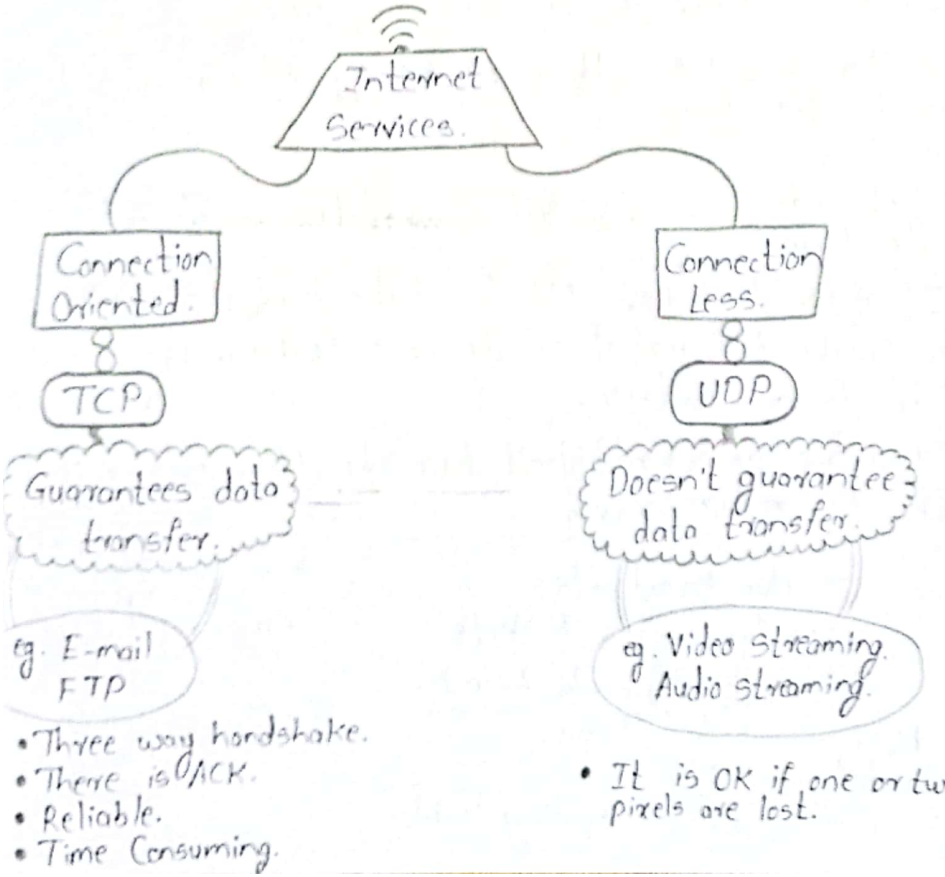
* All such tables are stored in Cache.

Bandwidth is measured in bits/sec.

Class A ISP ⟶ Reliance, BSNL, etc.
Class B ISP ⟶ They take bandwidth from Class A ISP's.
Class C ISP ⟶ Subnetting.

RFC ⟶ Request for Comments.
⟶ Produced by IETF (Internet Engineering Task Force)

→ Public Network ⟶ Internet
→ Private Network ⟶ Intranet.

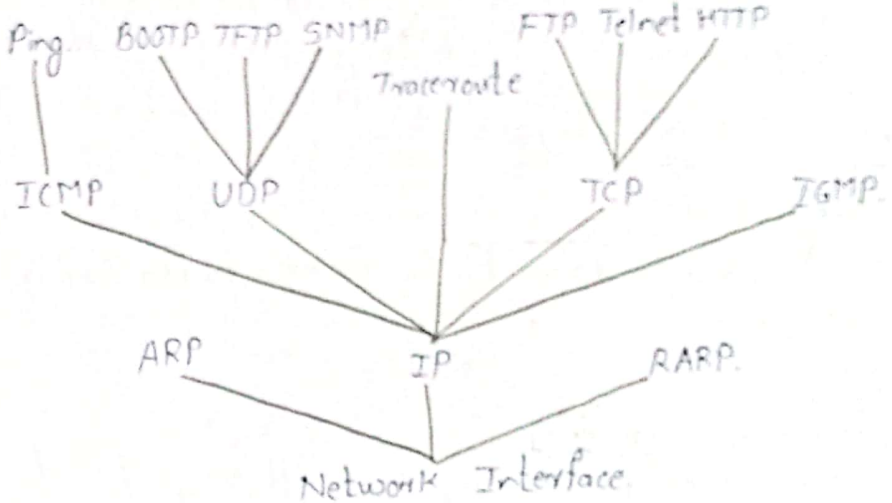


- Three way handshake.
- There is ACK.
- Reliable.
- Time Consuming.

- It is OK if one or two pixels are lost.

Protocol :-

It defines format, order of messages sent and received among network entity



* 2 RTTs (RTT_{HTTP}) are required to fetch a webpage. (Round Trip Time)
  → One to establish TCP.
  → One for data transfer.

- Hosts can be both Client and Servers.
- Network Edge :- Hosts.
- Network Core :- Routers.

Client - Server Architecture:

- FTP is used.
- eg. Printing Files.
- Draw-back : Connectivity / Server can fail.

▫ Types of Switches:-

Circuit Switch.	Packet Switch.
(i) Dedicated Path.	(i) It uses store and forward mechanism.
(ii) Path once allocated, others can't access it.	(ii) Delays are introduced.

▫ Packet Switching Delays:

(i) **Processing Delay:-**
Time required to examine the packet's header and determine the errors.

(ii) **Queuing Delay:**
Waiting in queue delay.
If there is no packet in queue, except single packet, then queuing delay is zero.

(iii) **Transmission Delay:**
Also known as store-and-forward delay.
Delay to place from router to the link.

(iv) **Propagation Delay:**
Delay in Propagation.

⊛ Application Layer - OSI and TCP/IP model.

(i) DNS Server. (Domain Name system).
(ii) Telnet Server. (SSH).
(iii) E-mail Server (SMTP, HTTP, POP3, IMAP).
(iv) DHCP Server. (Dynamic Host Configuration Protocol).
(v) Web Server. (HTTP, FTP).

• Peer-to-Peer (P2P) Networking.
→ Torrent is example of P2P.
→ It is decentralized.

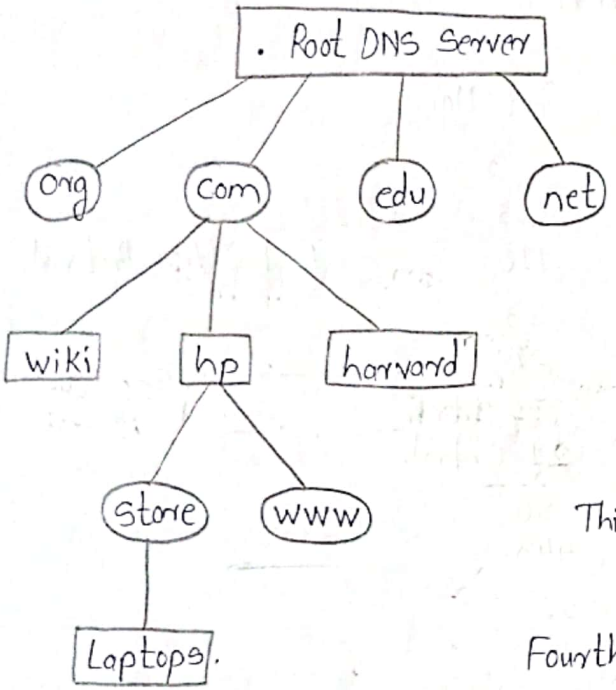
Application	Port No.
DNS	53
SMTP	25
POP3	110.
Telnet	23
DHCP	67
FTP	20 - Data 21 - Control.
HTTP.	80
HTTPS	443.
SSH	22.

Post Office Protocol.

▫ DNS (Domain Name System):

- It is used to translate the Host Name into IP Address
- Comprise of a hierarchy so that names are unique, yet easy to remember.
- HOST.TXT file mapps numerical addresses to name.
- Most commonly used implementation of the DNS, both resolver and name server - BIND.
  (Berkeley Internet Domain Server)
- Forward DNS: Host Name $\xrightarrow{\text{Mapping}}$ IP Address.
- Reverse DNS: IP Address $\xrightarrow{\text{Mapping}}$ Host Name.

There are 13 DNS root servers in the world.



Root level.

Top Level Domain.

Second level Domain/ Sub-Domain.

Third level Domain.

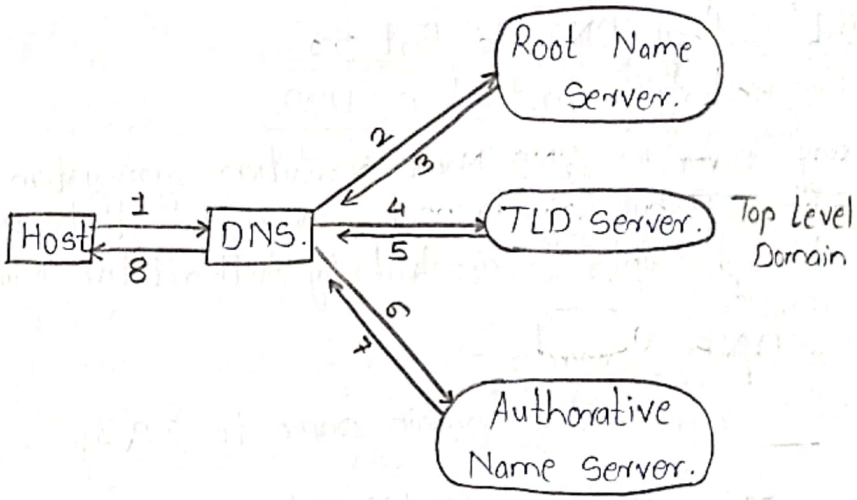
Fourth level Domain.

- 3 Character (Organizational Hierarchy):
  .com, .org, .edu, etc.
- 2 Character (Geographical Hierarchy):
  .in, .uk, .tv, .fr, etc.

- nslookup (net server lookup).

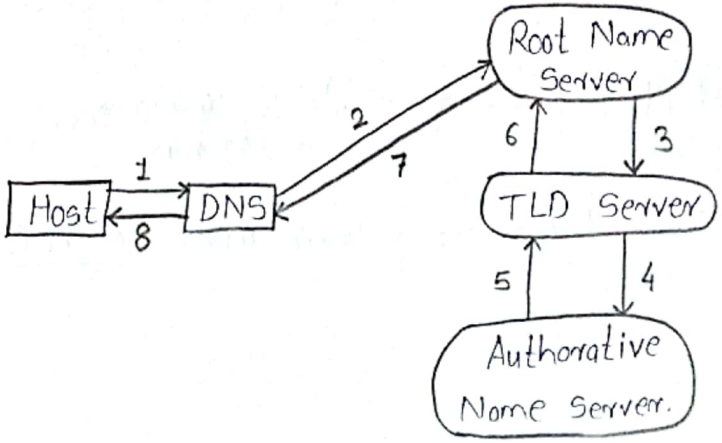
Getting Domain Resolved from DNS:-

1. Iterative Query.



Top Level Domain

2. Recursive Query:-



④ Authoritative Name Server actually returns the IP. (Last Server in DNS).

3. Non-Recursive Query:-

The IP is in Cache.

- Port used by DNS is Port 53.
- The underlying protocol is UDP.
- Every server in DNS Name Resolution sym system resolves IP, but only to next server. The actual IP which host wants is resolved by Authoritative Name Server.

DNS Records:-

- DNS A record match Domain name to IPv4.

eg. ns1 IN A 192.168.29.100.

- Time To Live (TTL):-

The default TTL for A records is 14,400 sec. ~ 240 min.

- DNS AAAA record matches a domain name to IPv6.

eg.

- CNAME:-
  - CNAME record points from an alias domain to "Canonical" domain.
  - It never points to the IP.
  - TTL 32600.sec.

- MX Record:
  - DNS 'mail exchange' record directs email to a mail server
  - It indicates how email messages should be routed in accordance with SMTP.
  - MTA (Mail Transfer Agent) is responsible for quering MX record.
  - MX record should point to A or AAAA record.

- NS record:- (Name Server)
  - It indicates ~~the~~ which DNS server is authoritative for that domain.
  - Tell the internet where to find IP of that domain.

NS Record.
3 W IN NS ns1.example.com.

eg. SOA Record.

@ IN. SOA ns1.example.com hostmaster.example.com (2024072900: serial no, 12h: refresh time ....)

Domain Name. Blank means TTL 2d; default.

Mail id. of master. @ is not used because @ refers to Domain name.

Zone file transfer between Master and Slave Servers.

* Serial No: YYYY:MM:DD:Sr.No.
2024 07 29 0 0

- Is transfered with SOA record.
- It it does not match, then Zone file is asked by slave to update the Zone file.

* Refresh Time: 12 hr.
- After each refresh time, the slave asks for the SOA record to master.

* Retry Time:- 15 m.
- How much time to wait, if the master server fails to reply, to make a new SOA request.

Expire Time:- 3w.
- If Master is not replying, after how much time should we stop trying. After this time, it will stop working as slave.

Negative Cache TTL:- 2h.
- If query cannot be resolved, i.e. it is not present, then it it will not ask the master again for the above time period.

DNS:
. Explain concept of DNS.
. Types of DNS Servers. (Explain). → Resolver, Root, TLD, Authorative.
. DNS ↓ Records.
Resource.

(*) What is RFC? Who produces RFC?

(*) We don't need to buy Private IP. DHCP gives DNS server's IP. Find where to configure DNS in laptop.

- Intro. to CN.
- Internet.
- ISP.
- IP Address. (Public, Private). (cables)
- Devices used in CN. (Routers, switch, hub, repeater) (their working layer)
- Host, server, medium, etc.
- end system, distributed appl^ns
- Diff^n bet^n TCP/IP and OSI.
- ARP protocol (How MAC is obtained?)

- Cables
- UDP Cables.
- Fibreoptic.
- Client-Server Archi.
- Peer-to-Peer Archi.
- TCP/IP protocol. OSI.
- Wired - Wireless → Single mode, Multimode.
- Layers. (TCP/IP + OSI).

HTTP :- (Hyper Text Transfer Protocol):

→ Request + Response.
→ Communication bet^n Server and Client.
→ It is Web's application-layer protocol.
→ Client Server Archi.
→ It defines the structure of message.
→ Underlying Protocol is TCP.
→ Port 80.
→ It is stateless. Every request is independent.
→ GET, POST, PUT, DELETE Methods. (Client uses this to make request).
→ Status Code. (Server uses to give response).

1xx : Informational.
2xx : Success.
3xx : Redirect.
4xx : Client Error.
5xx : Server Error.

200 - ok.
400 - Bard request.
404 - Page not found.
500 - Internal Server Error.
403 - Forbidden.

→ Requests: method, path, Version of protocol.
eg. GET / HTTP/1.1 → Request Line
Host: → Header.

Body.

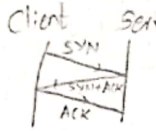
Response:-

Version. status Code. status message.

HTTP/1.1 200 OK. → Response Line.

Header. → Headers.

Body.

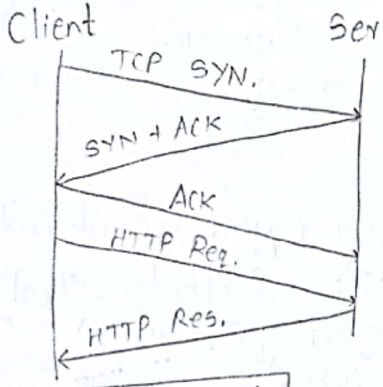


Web Client Web Server

① Establish TCP Connection ← 3-way Handshaking.

② Send HTTP Req. and wait for res.

Response.



TCP is underlying protocol for HTTP.

Socket → IP + Port.

Advantage of Layered Archi. is:
HTTP need not worry about data loss, or recovery.
TCP takes care of reordering, and all things.

HTTP is stateless.

Two methods of sending req:- (HTTP Connection type)

Persistent. Non-persistent.

Also,

Pipelining Non-pipelining.

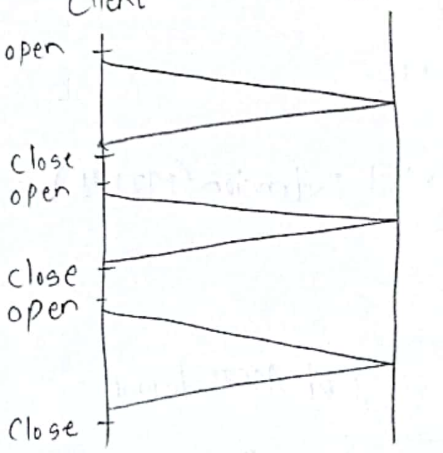
* RTT : (Round trip time).

HTTP response time = No. of RTT + file transmit time.

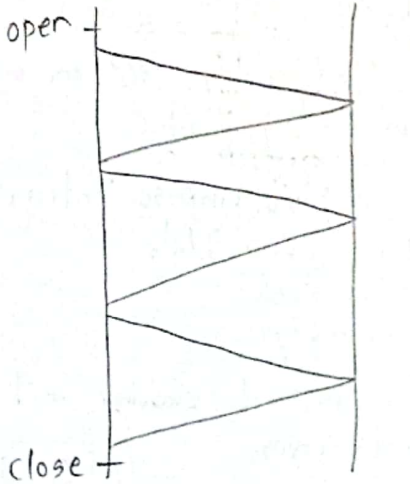
1 RTT to initiate TCP.
1 RTT for HTTP.
File transfer time.

* In non-persistent Connection, TCP is closed after every request. (RTT = 2 × no. of connections/req).
* In Persistent Connection, TCP is not closed after every request. (RTT = 1 + no. of connections/req).

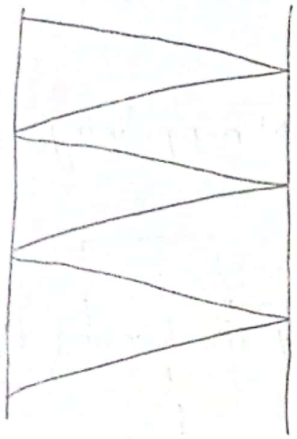
Non-Persistent



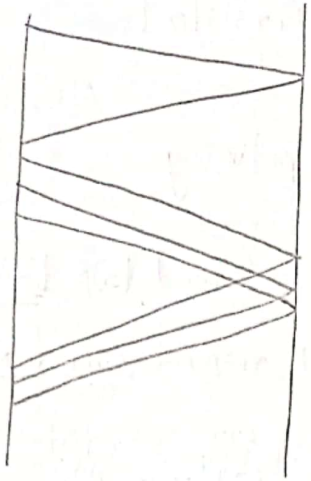
Persistent.



Without Pipelining.



With Pipelining.



g.

Persistent = RTT × 24.
Non-Persistent = 46 × RTT
∴ Extra time = 22 × RTT.
= 22 × 8.4 ms = 184.8 ms.

SMTP :- Port 25.

Electronic Mail :-

Most widely used on internet.

→ For Sending Mails:
- SMTP
- Multi-purpose Internet Mail Extension (MIME).

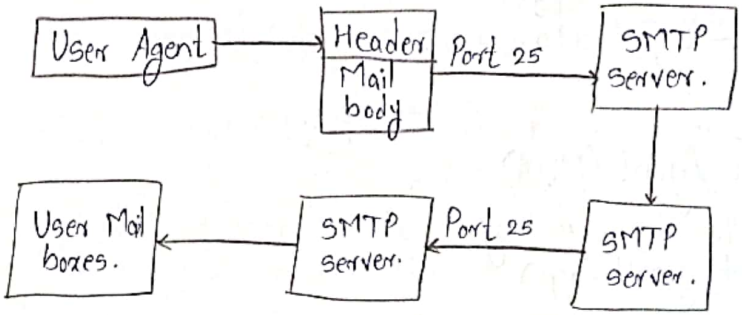
→ For receiving Mails:
- POP3
- IMAP.

- Allow users to exchange mail. 7-bit ASCII format.
- Client-Server.

RFC 821. → SMTP.

- SMTP is purely text based protocol.
- MIME extends limited capabilities of email.

Port 25: Non-encrypted.
Port 465: Secure SMTP.



- SMTP is Push Protocol.
- One or More TCP connections are established on Port 25.

Commands Issued by Client:
HELO - identifies SMTP sender to SMTP receiver.
MAIL FROM -
RCPT TO -
DATA -
QUIT -

Reponses by Server:-

* Main task of SMTP is to send main bet^n SMTP servers.
- Message is stored in Message queue, while transferring.

◈ Commands used by SMTP:

- HELO
- EHLO.
- MAIL FROM
- RCPT TO
- SIZE
- DATA.
- QUIT
- VRFY.

◈ Response from Server:-

- 101
- 211
- 214
- 220.
- 500
- 510.

□ Mail User Agent (MUA):

- Used to compose message.

eg. outlook, gmail (app).

□ Mail Transfer Agent (MTA):

- Used to forward email.
- Recieves message from other MUA or MTA.
- resides on SMTP server.

□ Mail Delivery Agent (MDA):

- Accept's mail from MTA.
- Place it into appropriate user's mailbox.
- On server only.

□ Post Office Protocol (POP3):

- Uses ~~port~~ TCP port 110.
- Retrive messages from the server.
- It is Pull Protocol.
- Two Modes : (i) Download and delete mode.
  (ii) Download and keep mode.

⊛ Company Mails (laptops) Uses POP3, for security purpose.

- POP3 allows us to read mail offline.
- Cannot access some mail from multiple location.

• POP3:

Port 110 → Non-encrypted.
Port 995 → Encrypted : Secure.

⊛ Download and keep is rare, as data redundency increase

□ IMAP:- (Internet Message Access Protocol).

- Email - not downloaded, kept on servers.
- HTTP can be used for web based mails, to ~~store~~ share from client to server, and from server to recipient.
- But from server to server SMTP is used.
- It is Pull Protocol.

Port 143 - Non-encrypted.
Port 993 - Secure

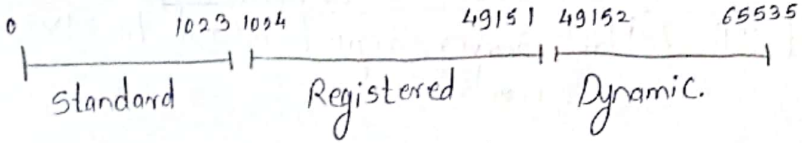
• To get IP of Mail Server, request is put to MX records.

Ports :-

- **Port:** A 16-bit no. that identifies the application process that receives an incomming message.
  ∴ $2^{16} \longrightarrow 65536$. Ports are available.
- **Standard / Well known Ports:-** 0 - 1023.
  or Reserved ports.
  /etc/services contains list of ports.

1024 ⟶ 49151 :- Registered Port. (By IANA).

Dynamic (or Private): 49152 ⟶ 65535
(temp. assigned by OS). [for Client].



* Remember the range of Ports.

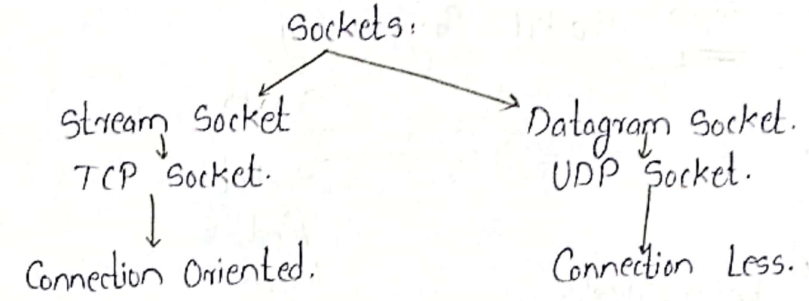
Socket:

It is Combination of 32-bit IP address and 16-bit Port address.

Socket = IP-address + Port.

- Socket is defined as an endpoint for Communication.

Port	Socket.
Different comp.s can use same port.	It is unique.
It identifies the program.	It is interface bet^n Application Layer and Transport Layer.



- **System functions: (Socket Programming).**

Functions in <sys/socket.h>.
- socket(), bind(), listen(), accept() —(Server).
- connet(), socket(). -(Client).

Both: send(), recv() (TCP) send to(), recvfrom(). (UDP)

Close() ⟶ Close Connection.

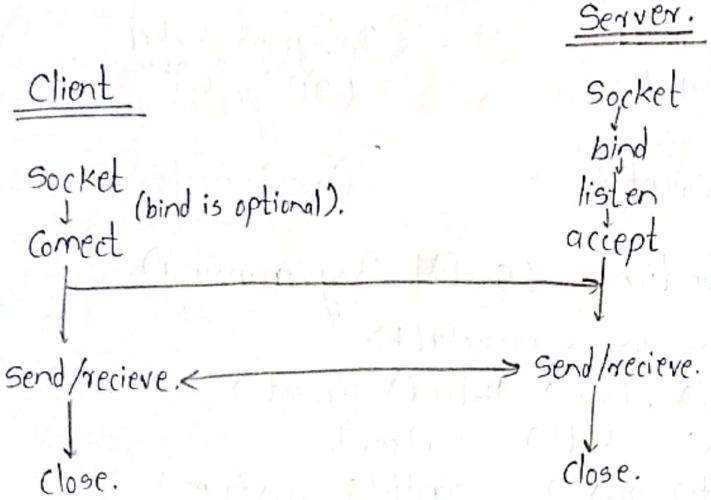
It is similar to telephonic Conversation.

Socket():

Syntax:- int socket(int family, int type, int protocol)

Family.		Type.
AF_INET	IPV4.	SOCK_STREAM
AF_INET6	IPV6	SOCK_DGRAM

TCP Socket Programming:



* Socket Interface is Above Transport Layer.

getaddrinfo (To get IP).

Socket: Creates a file descriptor.

Connection: Given fill descriptor and IP, establishes the connection. Protocol TCP/IP. (Transport Layer)

send and recv:
Given file descriptor (connected), asks os to send bits.

close.
Given file descriptor, closes the connection.

bind:
Given file descriptor, tell kernel to associate it with given IP and Port.

* Transport Layer is default model in OS.

127 - is used to loop back.

Class A:- 0 (7 : 2^7 networks). Host. Host. Host — 1–126

Class B:- 10..... Network. Host. Host — 128–191

Class C:- 110..... Network. Network. Host. — 192–223

Class D:- 1110..... Multicast. Multicast. Multicast group group group. — 224–239.

/28

8.8.8.④.

1111 0000.

No. of subnet masks:- $2^4 = 16$.

No. of Hosts per subnet :- $2^4 - 2 = 14$.

eg. Host IP: 192.168.0.250 — 255.255.255.240.

Network ID: 192.168.0.240 } 16th subnet.

Broadcast ID: 192.168.0.255.

25 - 8 = 17

ClassA: 10.0.0.10 /25 → 2^{17} subnets can be created. - Per subnet 2^7 Address, 2^7-2 Hosts.

25 - 16 = 9

Class B: 172.16.0.10 /25 → 2^9 subnets can be created. - Per subnet 2^7 Address. 2^7-2 usable Hosts.

25 - 24 = 1

Class C: 198.168.0.0 /25 → 2^1 subnets can be created. - Per subnet 2^7 Address 2^7-2 usable Hosts.

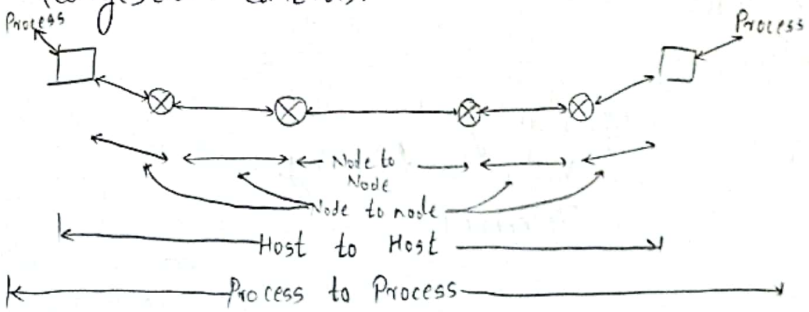
Transport Layer.

Services Provided by Transport Layer:-

- Process to Process Communication. (Transport Layer)
- Host to Host Communication. (Network Layer).
- Node to Node Communication. (DLL)

⊕ Different messages from various applications, will be sent through the same Link, using Multi-plexing, and differentiated based on Port number.

- Message is Divided into Segments.
- Multiplexing.
- Segmentation.
- Re-assembly.
- Connection Control. (Connection Establishment, management and Termination).
- Flow Control.
- Error Control. (Acknowledgement)
- Process level addressing.
- (Congestion Control).



* Only 1 Application per segment.
  Every segment will have only one Port number.

- Ordered delivery is also service of Transport Layer. (In-order Delivery).
- Reliability.
- (Both TCP/IP can be implemented here).
- This Layer Provides Logical Communication between processes.
- Multiplexing (while sending [Both Client+Server]).
- De-multiplexing (while Recieving [Both Client+Server])
- MSS (Maximum Segment Size).
- {MTU (Maximum Transfer Unit) in Network Layer}.

- **UDP**:-
  - No ACK.
  - Fast.

eg. DNS, SNMP, NFS.

eg. Goming, Video-Conferencing, Telephone.

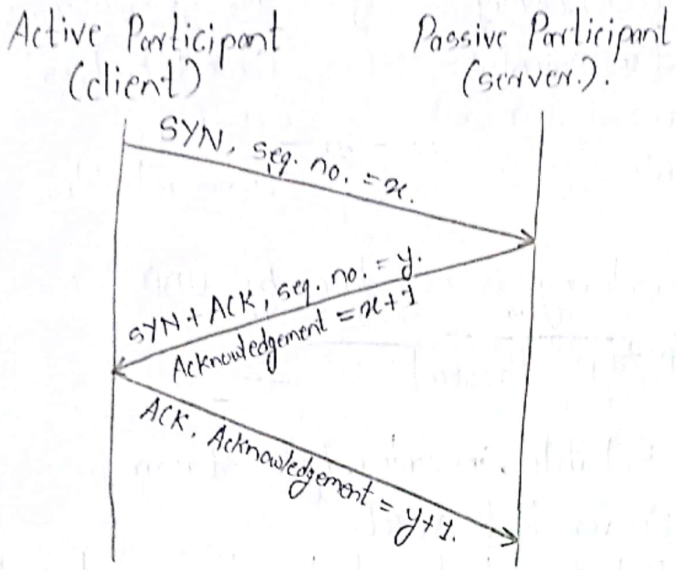
TCP	V/S	UDP
Garanteed Delivery.		No guarantee.
Connection Oriented.		Connection-Less.
Slow		fast.
Reliable.		Less-reliable.

* Re-ordering is not done by UDP.

- **TCP/IP Protocol**:-
  - Reliable, in-order, byte stream.
  - Recover lost packet.
  - Detect and drop duplicate/currupted packet.
  - Flow and Congestion Control.
  - Sliding window flow Control.
  - Connection Oriented.

Both **TCP/IP** and **UDP** are Protocols of The **Transport Layer**.

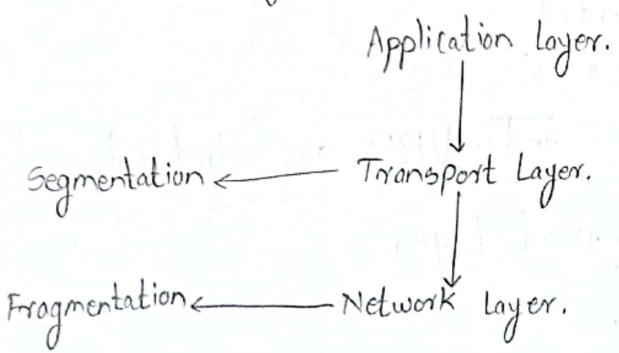
3 Way Hand-shaking:-



If seq. no. is x, the Ack. no. is x+1.

Maximum Segment Size:-

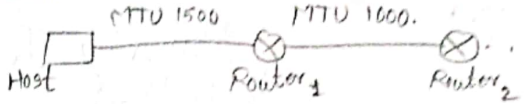
DF = No fragment. (If size is more, router will drop the packet).
MF = More fragment.



Q. Need of Both Segmentation and Fragmentation:
→ Segmentation at Transport Layer at Host (OS).
Fragmentation at Network Layer at Router.

- **MSS** (Maximum Segment Size) is the largest amount of payload (in bytes) that a communication device can handle in a single, unfragmented TCP segment.
- **MTU** (Maximum Transfer Unit):

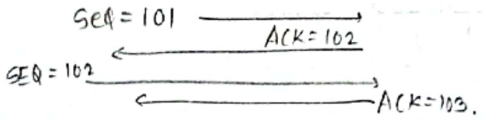
- Example Case:-

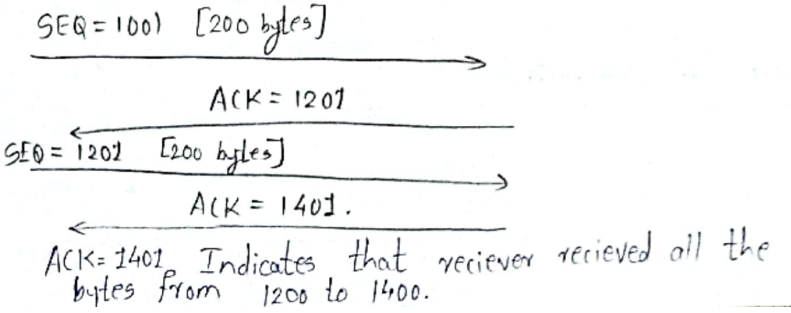


If Network layer data = 20 bytes.
Transport layer data = 20 bytes.
∴ MSS = 1460 bytes.

Now at router 1, MTU is 1000.
The 1460 bytes are fragmented and given Id.
Thus, both are essential.

SEQ / ACK:





ACK = 1401 Indicates that reciever recieved all the bytes from 1200 to 1400.

SEQ = 1401 [200 bytes] → ✕ — The router keeps these packets in cache, to re-transmit it if it gets lost (After timeout). After getting ACK, that packet is removed from the cache. (or buffer).

Packet loss

No ACK.

SEQ = 1401 [200 bytes].

ACK = 1601.

ACK lost:

SEQ = 1501 [200 bytes]

ACK = 1701

SEQ = 1701 [200 bytes]

ACK = 1901

✕ Timeout.

SEQ = 1701 [200 bytes]. → Droped.

ACK = 1901.

- Removing the duplicate packets.

Commulative ACK: (Delayed ACK)

SEQ = 1001 [200 bytes].

SEQ = 1201 [200 bytes].

ACK = 1401

SEQ = 1401 [100 bytes].

SEQ = 1501 [100 bytes].

ACK = 1601

SEQ = 1601 [200 bytes].

ACK = 1801.

Window Size:-

Window size limits how much ~~unacknowledged~~ ~~dff~~ unacknowledged data can be sent.

Reciever Advertises ~~this/her~~ it's remaining window size.

Flow Control

Window size

500.

300 — SEQ = 1001 [200 bytes]

100 — SEQ = 1201 [200 bytes]

STOP 0 — SEQ = 1401 [100 bytes]

500 — ACK = 1501 (Cumulative Ack).

300 — SEQ = 1501 [200 bytes]

Window size

500.

500

'Bytes in flight', the bytes whose ACK is yet to be recieved.

- Reciever will send back window size with ACK:-

ACK = 1501 [WIN 500].

ACK = 2001 [WIN 500].

SEQ=1001 [200 bytes] →

SEQ=1201 [200 bytes] →

SEQ=1401 [100 bytes] →

Data is too fast, please slow down!

← ACK=1501 [Win 0]

← ACK=1501 [Win 600].

SEQ=1501 [200 bytes] →

SEQ=1701 [200 bytes] →

SEQ=1901 [200 bytes] →

← ACK=2101 [Win 500].

TCP is Bidirectional:-

SEQ=1001 ACK=3001 [200 Bytes] →

← SEQ=3001 ACK=1201 [300 Bytes]

← SEQ=3301 ACK=1201 [300 Bytes]

SEQ=1201 ACK=3601 [200 bytes]. →

(*) Note:- Data link layer does the Flow control betwn two adjacent devices.

(*) Transport layer does the flow control between end devices.

▫ Initial Sequence Number (ISN) are randomly chosen by the sender, and is sent to other party while performing the three way handshaking.

20 byte TCP:-

TCP Header

Source Port	Dest. Port.
Initial seq. no. # = 1000	
Acknowledgement No. = 0000	
OFFSET | Reserved	CEUAPRSF | Window
Checksum	Urgent Pointer
TCP Options (Variable length).	

SYN.

ACK

·Three Way Handshaking:-

ACK=0 ACK#=0000
SYN=1 SEQ# = 1000 [0 bytes] →

← ACK# = 1001 ACK=1
SEQ# = 3000 [0 bytes] SYN=1

SYN=0
ACK=1 ACK# = 3001 [0 bytes]. →

IMP
⊛ See Flags and Numbers.

First Data Segment →
SEQ# = 1001 [500 bytes]
ACK # = 3001

Connection Teardown is 4-way process:-

DATA Transfer.

ACK=1 FIN=1 ACK=360 SEQ#=160 [0 bytes]

SEQ#=360 ACK#=160 [0 bytes] ACK=1

SEQ#=360 ACK#=160 [0 bytes] ACK=1 FIN=1

ACK=1 SEQ#=160 ACK#=360 [0 bytes].

Graceful Connection Termination.

4-way closure.

FIN (Finish) flag.

Ungraceful Connection Closing:-

One-way Termination.

Data Transfer.

RST (Reset).

RST=1

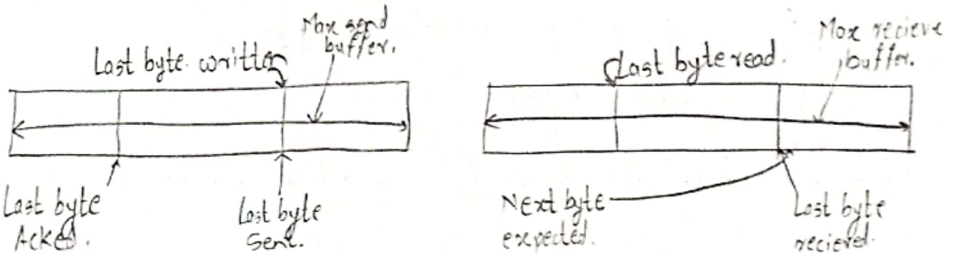
The window Principle:-

Send packet, wait for ACK, before sending next packet.

If a ACK is not recieved, retransmit the packet.

③ The sender can send all packets within the window without recieving the ACK.

[*] There is also minimum data size.



Advertise Window = Max Reciever Buffer − (Lastbyte recieved − Last byte read)

Advertise window = Max Reciever Buffer − Lastbyterecieved + Lastbyte read.

Sender Window = Advertised window − (Last byte sent − Last byte Acked)

= Advertised window − Lastbytesent + Last byte Acked.

Even if the Advertisement window is smaller than the lost packet, the sender can still send the packet because it is in it's queue /sequence.

eg. Max. reciever buffer size is 1500 bytes
Last read byte = 1700 byte.
Last recieved byte = 2400 bytes.
Next expected = 2401.
} Reciever.

Last byte written = 3000
Last byte sent = 2600.
Last byte acked = 2000.
} Sender.

Find effective window (i.e. Sender window):

→ Reciever Window = 1500 − (2400 − 1700)

Advertisement Window = 1500 − 700

= 800

Sender Window = 800 − (2600 − 2000)

= 800 − 600

= 200.

□ Error Control :-

RDT 1.0 / 2.0 / 2.1 / 3.0.
(Reliable Data Transfer).

TCP is a reliable transfer.

Error Control = Error detection + Error Control.

Error-free channel : No need for Error Control.

Noisy channel : Need for Error Control.

• Checksum :-

For Error Detection.

RDT 1.0 :-

Error free Channel.

RDT 2.0 :- ACK + NAK without SEQ.

No SEQ # is used.

ACK, NAK are used.

RDT 2.1 :- SEQ + ACK + NAK.

SEQ# is used to drop duplicate Packets.

If packet/ ACK is corrupted then there is re-transmission of packet.

RDT 2.2 :- SEQ + ACK without NAK.

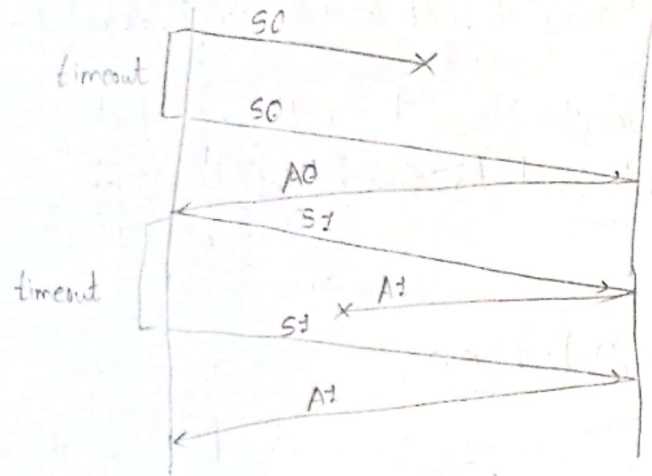
No NAK is used. Instead of NAK, send the ACK of previous packet. (Good strategy!)

☐ Lost Packet :-

Because intermediate device may drop it.
Buffer overflow.

RDT 3.0 :- SEQ + ACK + Timeout.

Stop and Wait Protocol.



RTT (Round Trip Time), T_T = Transmission delay.

$$\text{Utilization} = \frac{T_T}{T_T + RTT}$$

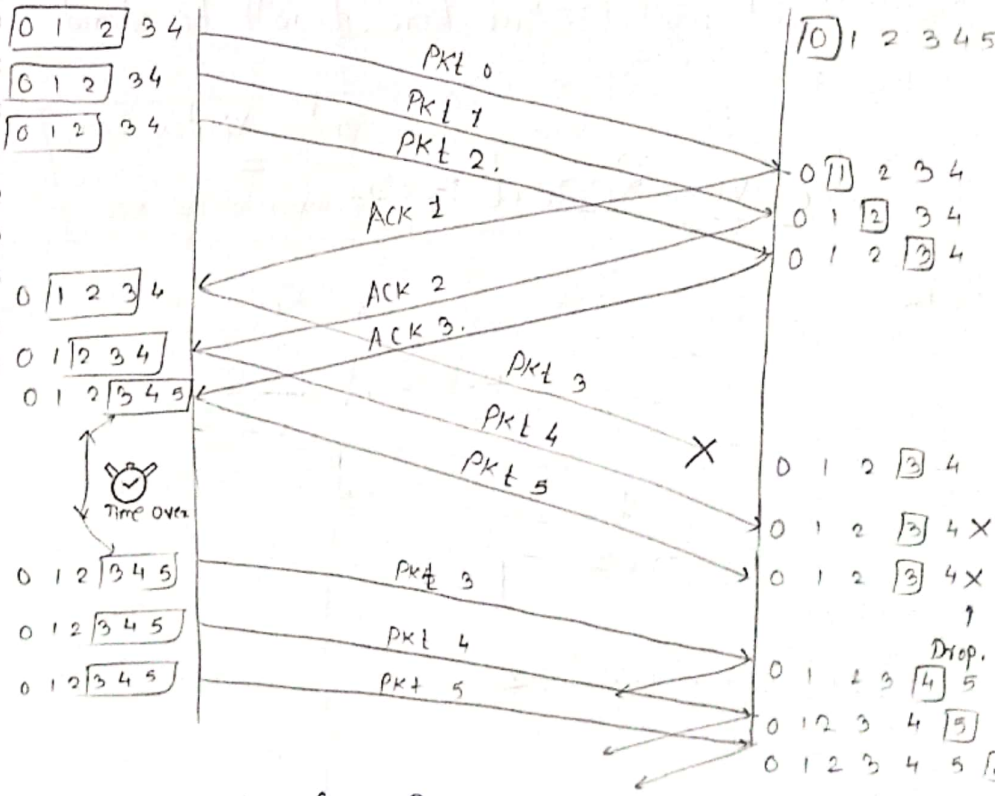
With pipelining, if no. of pipelines = N.

$$\text{Utilization} = \frac{N \times T_T}{T_T + RTT}$$

☐ Go-back N :- (Reciever Window size is 1).

Sender Window: N = 3.

Reciever Window N = 1.



- **Sequence No. (SEQ) :-**

k-bits : 0 to $2^k - 1$.

k = 2, $0 \le SEQ \le 3$

0 1 2 3 0 1 2 3 0 ...

Window Size :- 2^k

(*) If ACK is lost, next ACK will work.

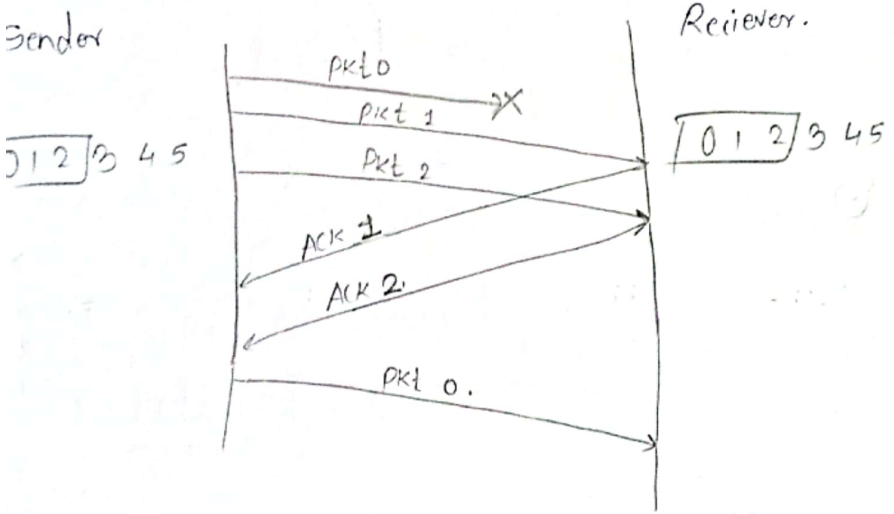
k is length of sequence number in binary (bits)

- Now, Sender Window = N.
  Reciever Window = N.

This is required, because in previous, the reciever will discard packets. But here, it will be stored in the buffer.

Selective Repeat :-

> Sender Window Size = Reciever window size.

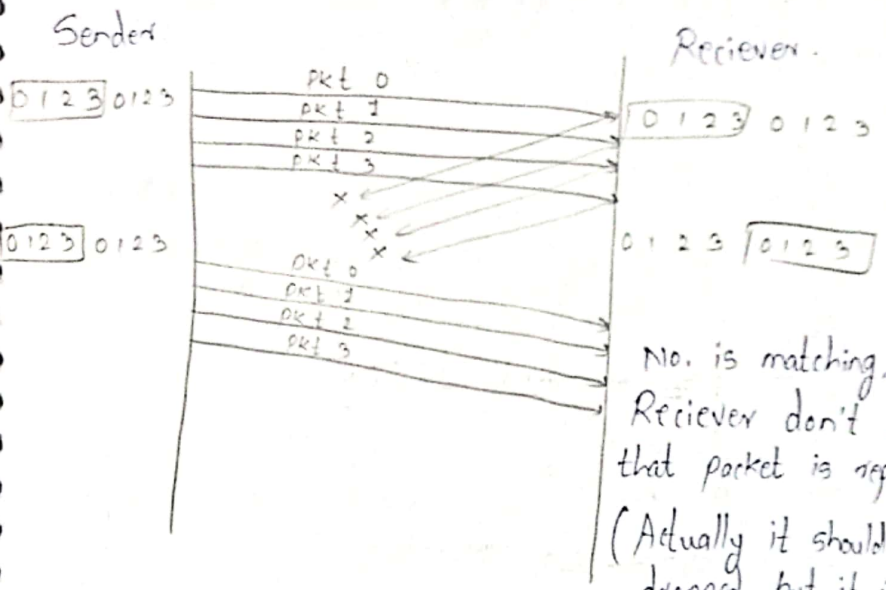


Here, if ACK is lost, Packet will be re-transmitted.

> ACK = SEQ + 1 will also work.
> ACK = SEQ will also work.

Will wait for the packet ACK, one by one.

- Problem / Bug in Selective Repeat:-

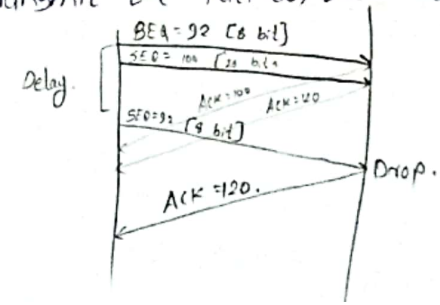


No. is matching. Reciever don't know that packet is repeating. (Actually it should be dropped, but it is not).

- Solution:-
  Keep K larger. (Larger SEQ number)

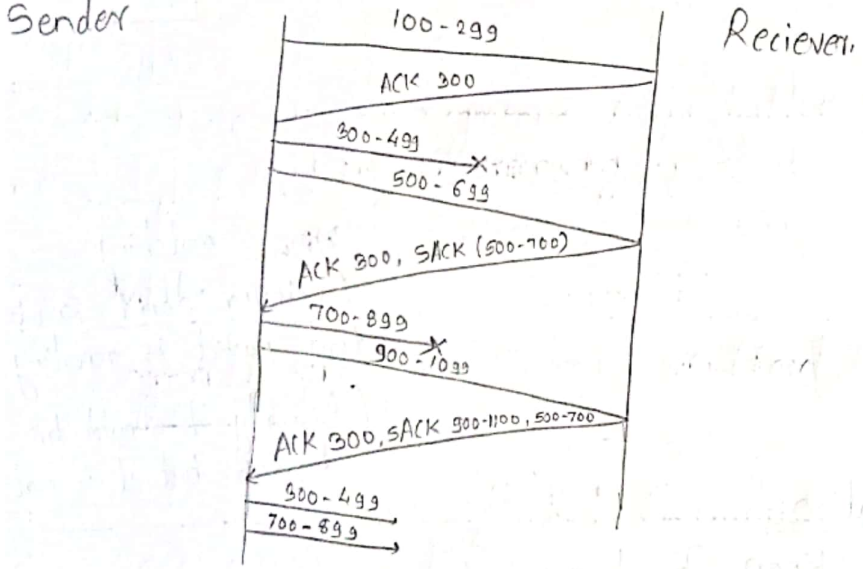
Delayed ACK :-

If delayed ACK is recieved by sender, it will retransmit the Packet, the reciever will discard it.



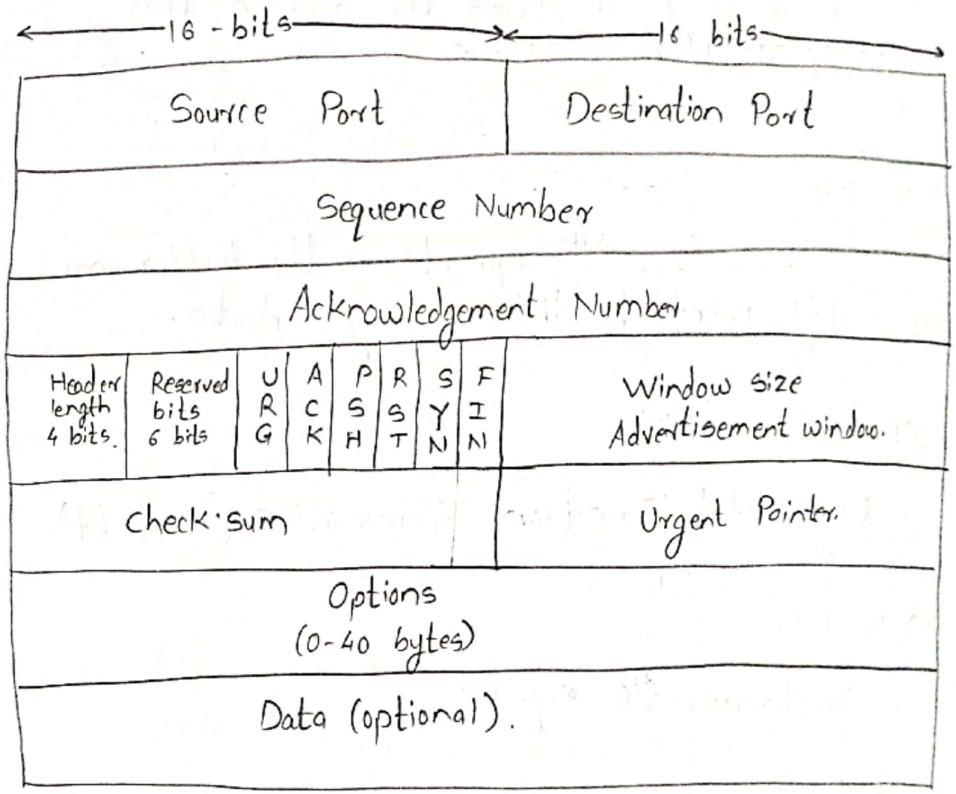
Selective Acknowledgement Option:-

Selective ACK :- (SACK)



This prevents sending the duplicate packets by the sender.

TCP Header.



- Source Port, Destination Port :- 16 bits.
- Seq No., ACK No. :- 32 bits.

Header length: 4 bits.

Scaling factor :- 4 bytes.

Length of TCP always lies bet^n : [20 bytes, 60 bytes].
(Source Port - Urgent pointer) (optional + 40)

20 bytes = 0101. × 4 bytes.

40 bytes = 1010 × 4 bytes

60 bytes = 1111 × 4 bytes.

- URG Bit:-

If it is '1', it means the data in this segment is 'Urgent'.

PSH Bit:-

If it is '1', All segments in the buffer are immediatdy pushed to the receiving applicaton.

RST Bit:-

Ungraceful Connection Termination. (Reset bit).

SYN Bit:-

SYnchronize the seq. no.

FIN Bit:-

It is used to terminate the TCP Connection.
(4-way handshaking).
(Graceful Connection Termination).

Window Size:-

Advertisement Window.

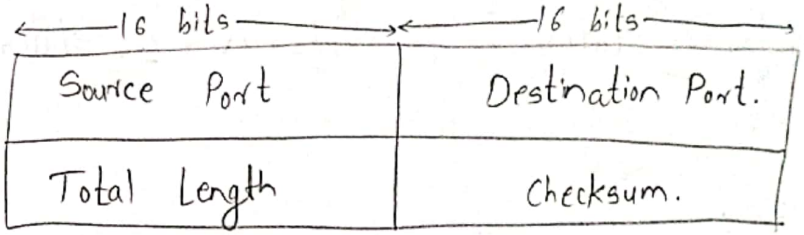
- Check Sum:-

Error Control and checking.

- Urgent Pointer:

When URG Bit is '1', this pointer tell where is the end of urgent data.

UDP Header:-



It has Pseudo header, which is also used to Calculate the Checksum.

- FSM for Sender and Reciever:- IMP

Reliable Data Transfer:-

- RDT 1.0:- (Error-free channel). [Draw/Remember Diagrams.]

Sender 1 state. Reciever 1 state.

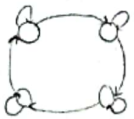
- RDT 2.0:- (ACK + NAK).

Sender 2 states. Reciever 1 state.

- RDT 2.1:- (ACK + NAK + SEQ).

Sender 4 states. Reciever 2 states.

For seq. no (0,1)

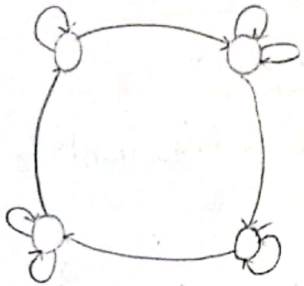


- RDT 2.2:- (ACK + SEQ).

Sender 4 states. Reciever 2 states.

- RDT 3.0 :- (ACK + SEQ + Timeout) :-

Sender 4 states. Reciever 2. states.



RDT 1.0
RDT 2.0
RDT 2.1
RDT 2.2
} → Packet Corrupted.

RDT 3.0 → Packet Lost.

Hub Operates at Physical Layer.
Switch at DLL.

Congestion Control :-

It is mechanism of adjusting the sending rate according to the resources (i.e. bandwidth and router queue size) available in the immediate networks.

* Effective Window.

= Min (Congestion window, Advertisement Window) - (Lastbytesent - Lastbyte Acked).

It occurs between nodes, i.e. between routers.

* Actual window size = min (reciever window size (Advertisement window), Congestion window size).

Example :- IMP

CW = 15000 bytes. (Congestion Window).

Reciever-max = 30000 bytes
Last byte Acked = 20,000 bytes.
Last byte sent = 30,000 bytes.
Last byte read = 15,000 bytes.
Last byte recieved = 20,000 bytes.

→ Advertisement Window = 30,000 - (20,000 - 15,000)

AW. = 25,000.

Actual window = min (AW, CW) = 15,000

∴ Effective Window = Actual window - (sent - acked)
= 15000 - (30000 - 20000)

Effective Window. = 5000.

- Ingrace :- Incomming Capacity.
- Egress :- Outcomming Capacity.

Congestion Detection :-

- If Packet loss :-
  - Time out. (RTO)
  - Three duplicate ACK.

- TCP Versions :-
  - Taho TCP
  - Reno TCP.

- Policies :-
  - Slow start
  - Congestion Avoidance.
  - Fast Recovery.

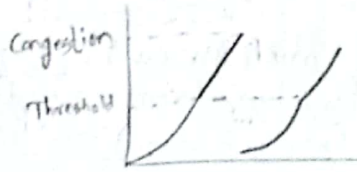
Slow Start :-

Here, the Congestion window size will grow exponentially.

Congestion Avoidance :-

From a particular threshold value, the window will increase by One. (Additively). [Additive Increase].

Whenever, there is packet loss, the congestion size is suddenly droped to '1' again above process is repeated.

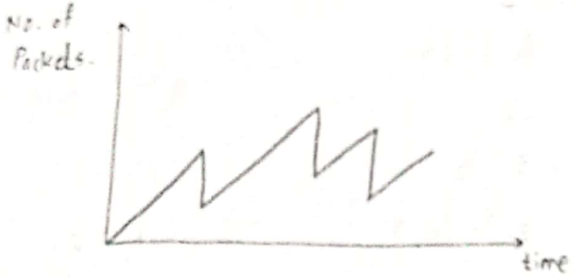


start: RTT 0 $\longrightarrow 2^0 = 1$

* The threshold is set, by current window divided by two. *

- AIMD (Additive Increase / Multiplicative Decrease)

For each packet loss, the sender decreases it's Congestion window by half of it's current value.



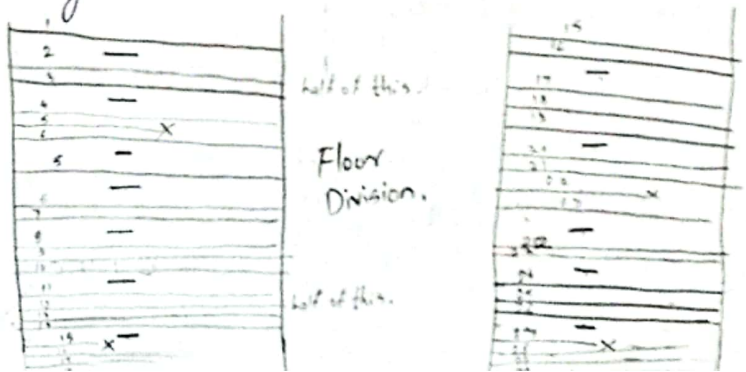
Fast Recovery :-

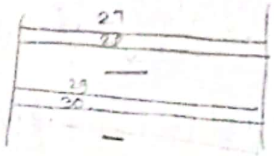
It start's with slow start, then when first time congestion occurs, it starts with AIMD.

The congestion window becomes half, start growing Additively.

Example :- IMP

30 Packets are to be sent. If packet 15th packet is lost, 15th, 22nd, & 27th packet is lost. By using AIMD, tell how many RTT's are used to send 30 Packets.





16 RTT's are there.

Do the same with slow start:-

RTT's	Packets.	
1	1	
2	2, 3	
3	4, [5], 6, 7.	Threshold = 2.
4.	4	
5.	5, 6. —	
6.	7, 8, 9.	
7.	10, 11, 12, 13.	
8.	14, [15], 16, 17, 18.	Threshold = 2.
9.	15	
10.	16, 17. —	
11.	18, 19, 20.	
12.	21, [22], 23, 24.	Threshold = 2.
13.	22.	
14.	23, 24. —	
15.	25, 26, [27]	Threshold = 1.
16.	27. —	
17.	28, 29.	
18.	30.	

18 RTTS.

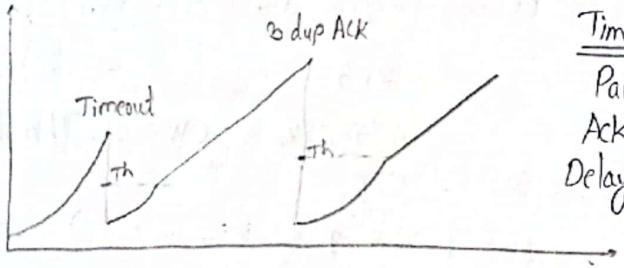
Do the same with fast recovery:-

RTT's	Packets.	
1.	1	
2.	2, 3	
3.	4, [5] 6, 7.	CW = 2. Threshold = 2.
4.	5, 6.	
5.	7, 8, 9	
6.	10, 11, 12, 13.	
7.	14, [15] 16, 17, 18	Threshold = 2 = 5/2.
8.	15, 16.	
9.	17, 18, 19.	
10.	20, 21 [22], 23.	Threshold = 2.
11.	22, 23.	
12.	24, 25, 26.	
13.	[27] 28, 29, 30.	Threshold = 2.
14.	27, 28.	
15.	29, 30.	

No. of RTTs = 15.

TCP Tahoe:-

Slow start and Congestion Avoidance.



Timeout:
Packet loss
Ack Loss
Delayed Ack.

TCP Reno:-

A newer version of TCP, called TCP Reno added a new state to the congestion-control, called the fast recovery state.

eg. TCP Tahoe: ss + AI + Fast Retransmit.

[Study this!!]

- DNS Caching
- Peer to Peer - Bit torrent.
- TCP state transition, diagram.
- Web Caching.
- CDN: Content Delivery Network.

Autonomous Systems:-

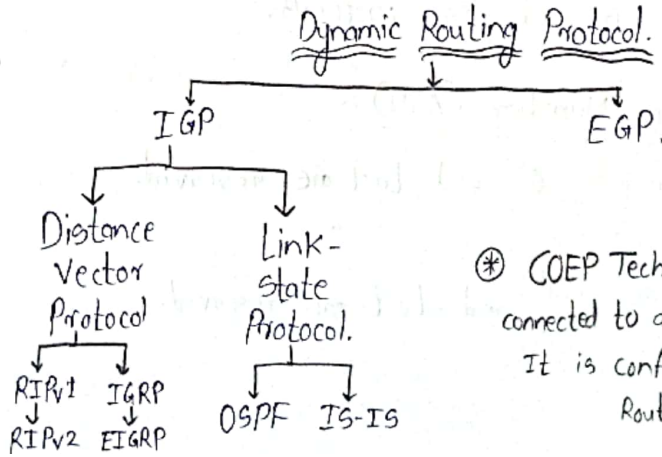
It is collection of routers whose prefixes and routing policies are under common administrative control

eg. NSP, large company, university, etc.

- **Interior Gateway Protocol (IGPs):**
  Distribute routing information within AS.
- **Exterior Gateway Protocol (EGPs):**
  Distribute routing information outside AS.

Static Routes V/s Dynamic Routes:-

- **Static Routes:**
  - Configured Manually into the router.
- **Dynamic Routes:-**



* COEP Tech. has only one Router, connected to a lot of switches. It is configured using static Routing. *

Distance Vector Routing Protocols:-

- Two Main Characteristics:-
  - Distance. (based on hop count, cost, etc).
  - Vector.

Link-state Routing Protocols:

It uses Cost. Cost $\propto \frac{1}{\text{Band-width.}}$

RIP-enabled routers send periodic updates of their routing information to their neighbours.

But, Link-state do not use periodic updates.
Only the changes will be sent, when there is a change in the topology.

Topology changes when link fails.

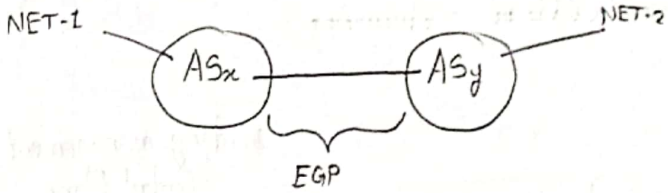
RIPv1 and IGRP are class full protocols.

RIPv2 and EIGRP are classless protocols.

- Autonomous System Number (ASN):-

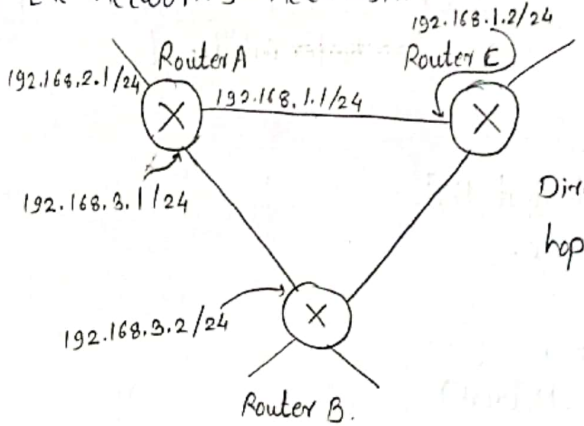
16-bit → 2^{16} no.s 0^{th} and Last are reserved.

32-bit → 2^{32} no.s 0^{th} and Last are reserved.



Distance bet^n directly connected routers is Zero.

Router is connected to different networks through different interfaces. Each interface has different IP given by the network's netmask.



Directly Connected networks' hop count is zero (0).

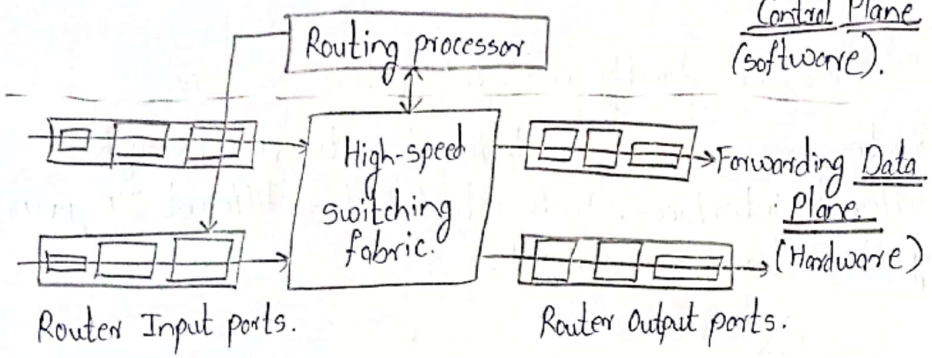
- Metric:-

There can be multiple paths to reach the same network.

Metric is a value assigned to each path, so that the router can calculate the best path for the packet.

- RIP defines "best" path as path with least hop count.
  - 16 hop count ≡ Unreachable. (In RIP).
- IGRP defines "best" path as path based on the combination of the bandwidth and the delay.

Router Architecture:-

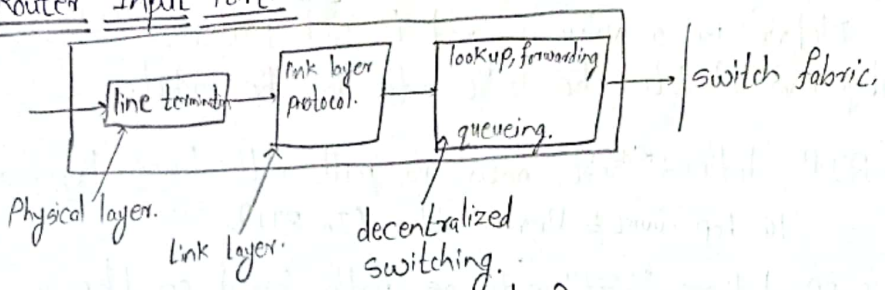


- Four Components:-
  1. Input Port. ~~/Output Port~~
  2. Switching Fabric.
  3. Output Port.
  4. Routing Processor.
     (Software Define Network)

* SDN only has Data Plane, Control Plane is present in cloud.
  It is high speed.

"Router is Traditional way"

Router Input Port:-



* Control Plane creates the routing table, and it takes the decision which route is best/to be used.
* Input Port decides which interface to use.

Control Plane → Routing Decision.
Data Plane → Forwarding Decision.

Match Plus Action Principle:-

- Operates on this principle.
- Uses header file.

Destination-based forwarding:

- There are 2^{32} addresses for destination.

To search the entry, longest prefix matching is used, for fast searching.

Destination Address Range.	Link Interface.	
11001000 00010111 00010*** ********	0	First prefix 21 bit
11001000 00010111 00011000 ********	1	24 bit
11001000 00010111 00011*** ********	2	21 bits
Otherwise.	3.	

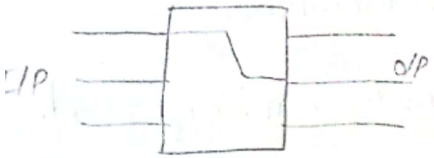
- Match-Plus-Action:-
  - TCAM memory is used
  - Ternary Content-Addressable Memories.
  - Process occurs in Hardware.
  - Takes place in "Input Port"

Switching Fabric:- (This determines the cost of router)

Switching Rate:-

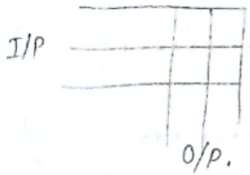
The rate at which packets can be transfer from input to output.

- **Major Components:- (Types of Switching fabrics)**
  - (i) Memory.
  - (ii) Bus.
  - (iii) Interconnection network.



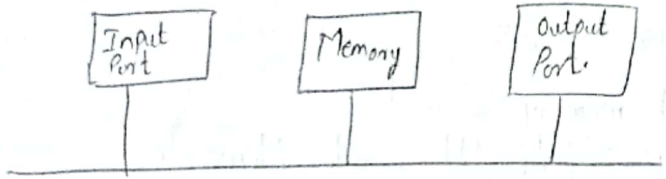
Memory.

Bus.



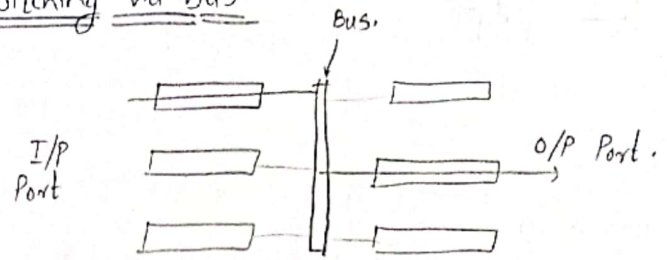
Interconnected Network.

- **Switching via Memory:-**



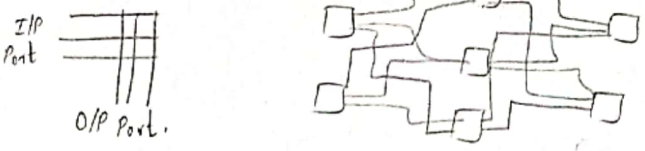
speed limited by Memory bandwidth.

- **Switching via bus:-**



Speed is limited by Bus Bandwidth.

- **Switching via Interconnected Network:-**

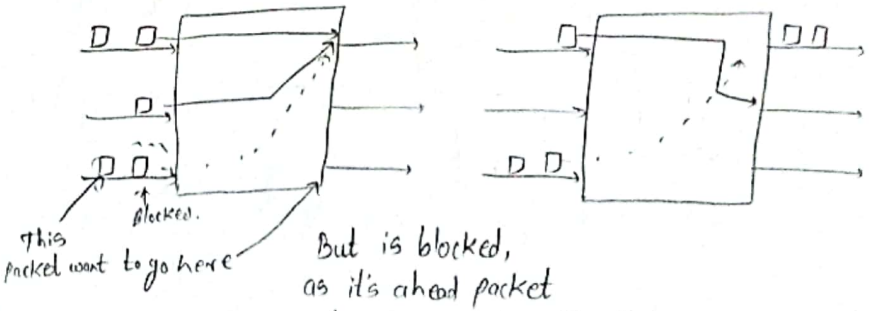


Parallel Data Transmission

Input Port Queueing:-

- If switch fabric is slower than input ports, queueing is required.
- queueing delay and loss can occur due to input buffer overflow.

- **Head-of-the-Line (HOL) blocking:-**



Output Port Queueing:-

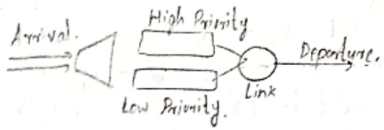
- **Buffer Management:-**
  - Drop: which packet to add when buffer is full.
    - Priority
    - Random.

Packet Scheduling:-

(i) FCFS. (FIFO). [Queue operation].

(ii) Round Robin. (RR).

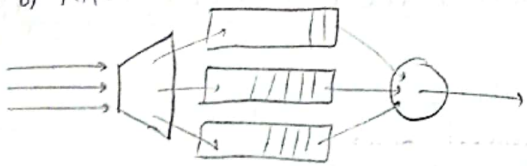
(iii) Priority Queueing.



Non-premptive priority queuing is used.

(iv) Weighted Fair Queue. (WFQ).

Generalized form of RR.



Network Neutrality:-

- Technical.
- Social, economic principles.
- Enforced legal rules and policies.

IPV4 Header Format:-

Version	IHL	TOS	Total length.
Identification		3 bits Flag bits	13 bits Fragment offset
Time to Live.	Protocol	Header Checksum. ← Only header Checksum	
Source IP			
Destination IP.			
Option.			
Data. [segment].			

- **Version:** 4 bits.
- **Header Length (IHL)** : 4 bits.
  (n × 4 bytes). eg. (0101 (5×4 = 20 bytes)).
- **Type of Service (TOS)** :- 8 bits.
- **Total Length:** 16 bits.
  Total length = Header length + Payload length.

- Identification:- 16 bits.

  Identification of the fragment of an original IP datagram.

- DF bit:- [Do not fragment bit].

  If 1 :- Drop packet. Don't fragment. Error is sent to sender.
  If 0 :- Fragment if required.

- MF bit:- [More fragments bit].

  If 0 :- This is last fragment or the only fragment.
  If 1 :- It is fragment of some larger datagram.

MF bit is always set to 1 except last fragment.

Fragment offset:- [13 bits]

It indicates the position of the fragment in the unfragmented data.

Scaling factor:-

Scaling factor is 8.

Time to Live:- (8 bits).

- Indicates the max. no. of hops a datagram can take to reach destination.
- It is used to prevent IP datagrams from looping around forever in routing loop.
- If value of TTL is found to be '0' before destination, the packet is dropped.
- At each hop it is decreamented by 1.

Protocol:- (8 bits).

ICMP 1, IGMP 2, TCP 6, UDP 17.

Header Checksum:- (16 bits)

Checksum w.r.t. header.
This checksum is calculated at each hop.
And changed, because fields are modified at each hop.
eg. TTL.
Options, etc.