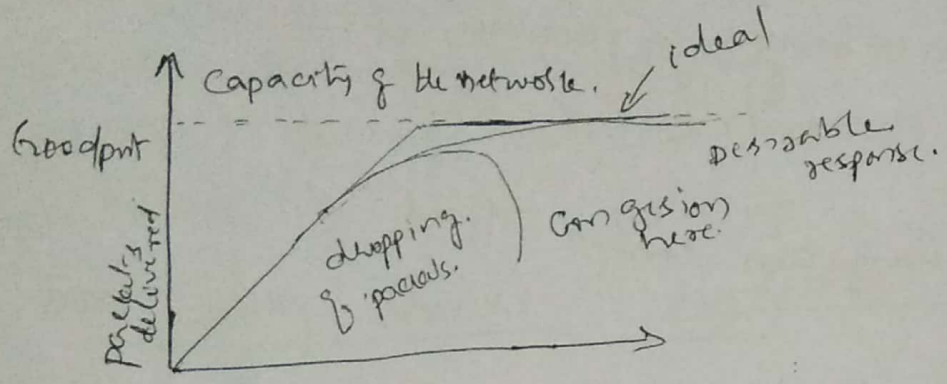
Congestion: Too many packets present in the network causes packet delay and packet loss that degrades the performance. This situation is called Congestion.

The responsibility of handling congestion is shared among network and Transport layer.



Reasons for Congestion

- Sudden arrival of packets on a particular o/p line from multiple i/p source
- Insufficient memory in a router to hold packets
- Slow processor can also cause congestion
- Low bandwidth channel can also cause congestion

Congestion Control: ensures that the subnet is able to carry the offered traffic and not overloaded.

Flow control: Ensures fast sender does not overload ~~packets~~ receiver.

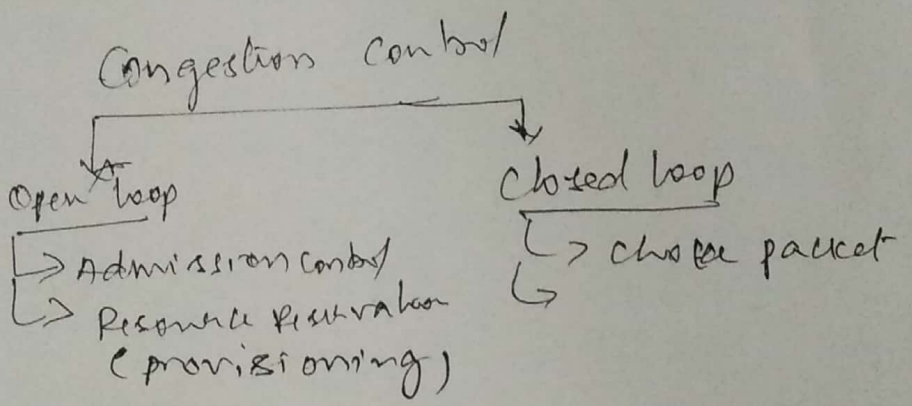
Congestion control required when the buffers at the switches overflows

* Congestion involves all hosts, links, routers & it is global issue

* flow control is needed when the buffers at receiver overflows i.e when Receiver is slow in processing

* flow is b/w Sender & receiver, it is local issue or can be done on end-to-end basis.

Approaches to Congestion control



Open loop congestion control : policies are designed and applied to prevent Congestion before it happens rather than dealing with it after it has occured.

* Admission Control : In this method the ~~load~~ load is decreased to avoid Congestion. In virtual networks new connections ~~are~~ can be refused if they would leads to congestion. i.e if the congestion is suppose to happen then the network can intimate the Sources responsible for the problem to throttle their traffic or slow down the traffic itself.

Problems are : How to identify the onset of Congestion

and how to inform source that needs to slow down

So router can monitor the average load and Queueing delay or packet loss. so rising number in packet loss indicates growing congestion

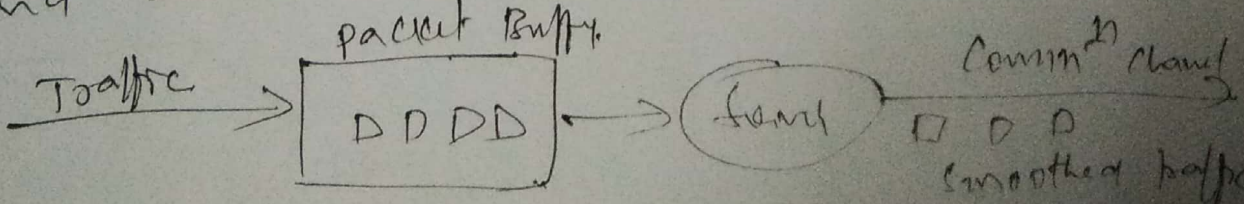
Traffic policing: process of monitoring and ~~The network may~~ enforcing the traffic flow of packets during connection period. The most implementation of traffic policing is done through

~~Leaky~~ Leaky Bucket algorithm:

In this each router will have buffer to hold the incoming packets, so that no matter at what rate the packet arrives at the i/p line of router, the router pass to its outgoing line at fixed rate:

If there is no overflow at the buffer. then the traffic is within the agreed norms all packet will be transported.

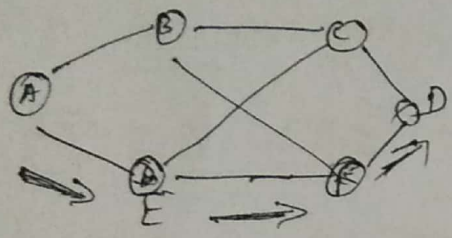
If buffer spillover. it may lead to congestion and additional packets will be lost.

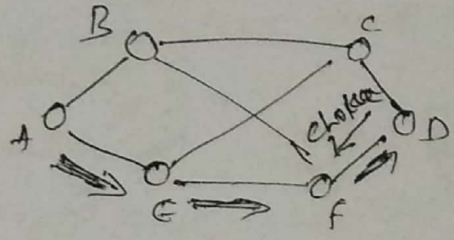


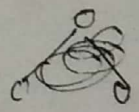
-Loop Congestion Control: These methods try to remove congestion after it happens.

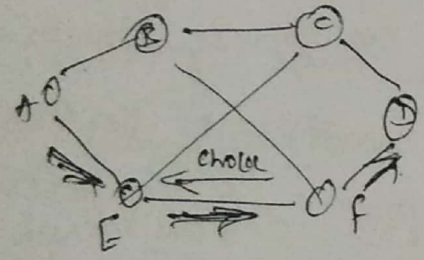
Ex Choke packets, Hop-by-Hop Backpressure, Random Early Detection (RED)

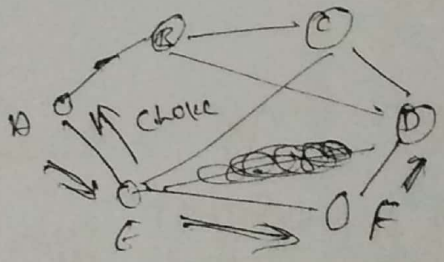
Choke packets: In this method the router selects a congested packet and sends a choke packet back to the source host, when the source host gets the choke packet, it is required to reduce the traffic sent to specified destination.

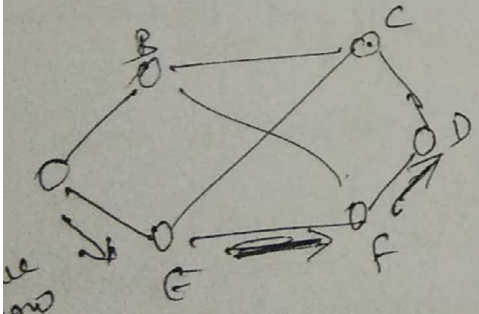


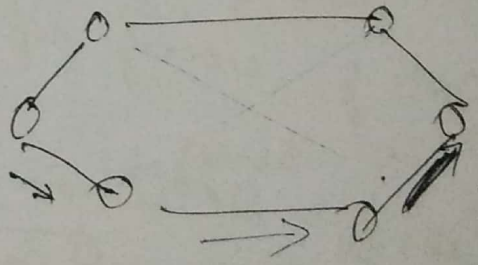




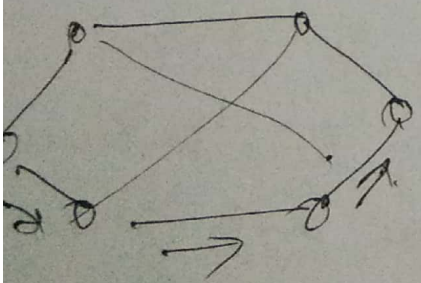






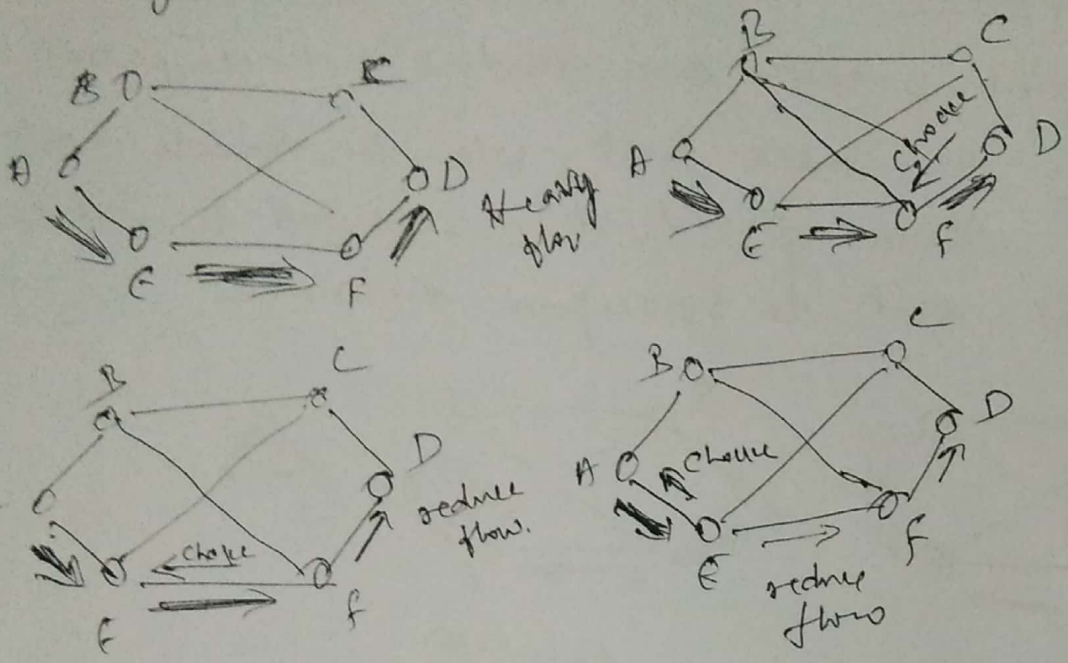


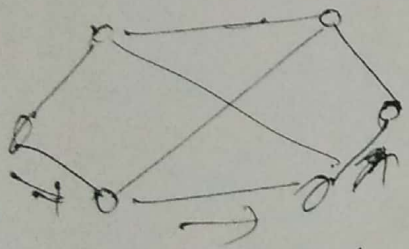
flow is still @ max rate



flow is reduced

Hop by Hop Backpressure: In choke packet method, high speed or over long distances many new packets may be transmitted even after the congestion has been identified. Therefore an alternative method is have the choke packet take effect at every hop it passes through. as shown below





The advantage of Hop by Hop scheme is it provides the quick relief at the point of congestion at the cost of using up more buffers upstream.

Random Early Detection (RED)

This is based on the dropping of packets at routers early.

To determine when to discard packets, routers maintains a running average of their queue lengths, when the average queue length exceeds a threshold, a small fraction of the packets are dropped at Random.

ACK.

RED routers improve performance compare to routers that drop packets only when their buffers are full.

QoS. Mechanisms make the network with less capacity also to meet the applications requirements at a lower cost.

The needs of each flow can be characterized by 4 primary parameters and together determine QoS.

① Reliability: Lack of reliability means losing a packet ⑪ ACK which initiates retransmission.

② Bandwidth: Different apps need different bandwidths. Ex video conferencing needs millions of bits/sec while an E-mail may not reach even a million.

③ Jitter: Jitter is defined as the variation in the packet delay. i.e if the ~~packets~~ 4 packets depart at time 0, 1, 2, 3 and arrives at 20, 21, 22, 23 all have the same delay i.e 20 units of time. ⑪ if the same packets arrive at 21, 23, 21 & 28 then have different delays i.e 21, 22, 19 and 24 i.e variation in delay.

Audio & Video

Delay: Source to destination delay is another flow characteristics, interactive applications like web surfing and remote login are more delay sensitive.

E mail, file transfer etc are not delay sensitive

E-mail, file transfer have more stringent requirements for reliability (loss/error) than audio & video.

Internetworking: Interconnection of several networks with different protocols.

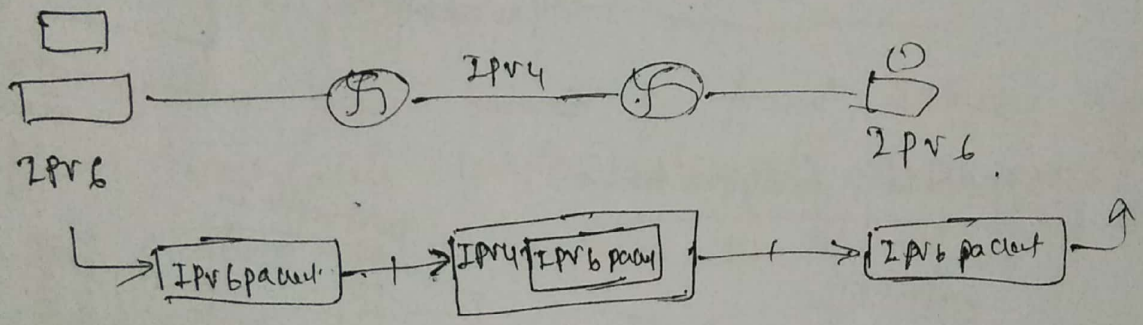
How does network differs?

service offered ⇒ Connectionless / Connection oriented.
Addressing ⇒ Different size, flat-/hierarchical
packet size ⇒ Every network has its own maximum
ordering ⇒ ordered & unordered delivery
QoS ⇒ present / absent; different kinds (parameters)
Reliability ⇒ Different levels & loss.
Accounting ⇒ By Connect time, packet, Byte, & not at all

Tunneling: Tunneling concept is used when a source and destination hosts are on the same type of networks but there is a different network in between. So the solution to connect source & destination is tunneling.

Tunneling encapsulates the data unit of one protocol inside a different kind of protocol.

Tunneling is widely used to connect isolated hosts and networks using other networks.



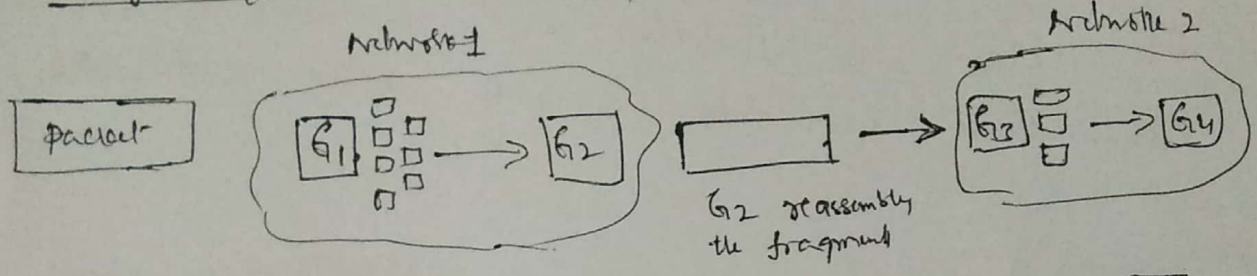
Fragmentation:

Every network (or) link imposes some maximum size on its packet due to

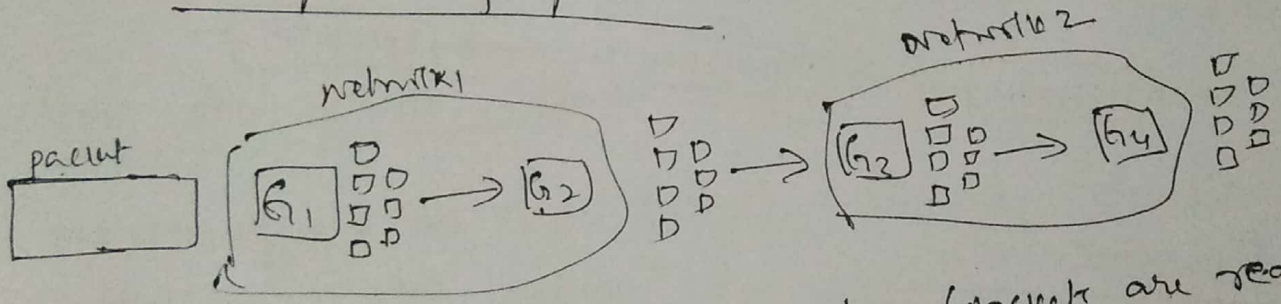
① Hardware (Ethernet frame) ② Operating System (: buffer size 512 bytes) ③ protocols (no. of bits in the packet length field) ④ Compliance with some national / international standard ⑤ Desire to reduce error-induced retransmissions ⑥ Desire to prevent one packet from occupying the channel too long.

when the packet size of networks
then the internetworking problem arises
the large packet needs to travel through a
network whose maximum packet size is small.
one solution to this problem is Fragmentation
i.e allow the routers to break up packets
into fragments

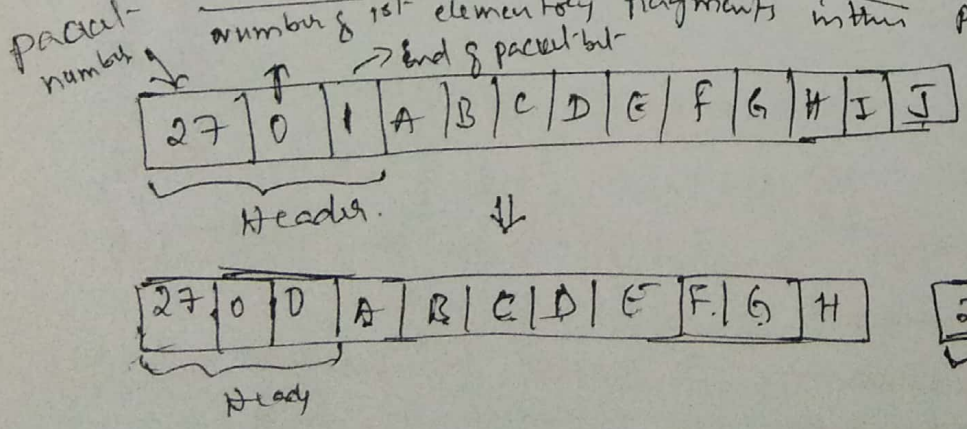
2 types of fragmentation → Transparent
→ non-transparent



Transparent fragmentation



Non Transparent fragmentation (packets are reassembled only at the destination)



fragments after passing through a network with packet size of 8 bytes payload plus header

Inet is an interconnected Collection of many networks. The glue that holds the whole Internet together is network layer protocol IP (Internet protocol)

The IPv4 protocol

To study of network layer in the Internet it starts with its format of IP datagrams.

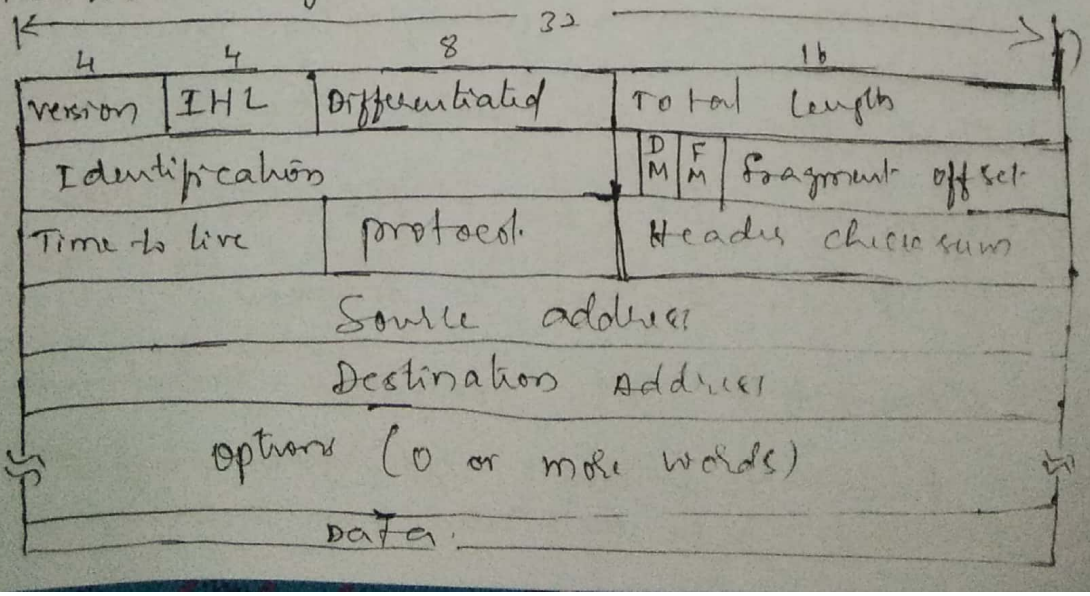
An IPv4 datagram Consists of.

1. Header part.
2. payload part

The Header part is fixed 20 bytes apart from variable length optional part.

The payload part Contains the actual data to transmitted

The bits in the datagram are transmitted from left to right and Top to bottom i.e the High-order bit of the Version field going first (Big-endian network byte order)



Version : 4 bit field keeps track of the version of protocol. This field needed for the transition between the versions over a long period of time

IHL : (Header length) : It tells how long the header is, The min is 5 when no options are present. The max value of this 4 bit field is 15 which limits the header to 60 bytes i.e (20 + 40 option bytes)

Differentiated Services (Type of Service) : It is used to distinguish the different classes of services i.e for digitalized voice transmission fast delivery beats accurate delivery. For file transfer error free transmission is imp than speed. But these bits have no importance in routing ∴ left unused.

Total length : It includes the length of both Header and data : The maximum length is 65,535 bytes,

Identification field (Sequence ID) (16bits) : It is used by destination machine to determine to which packet the fragment belongs to. All fragments of a packet contains same identification number.

DF & MF These two 1-bit fields are related to fragmentation. DF (Don't fragment). It is an order to Router not to fragment

MF (More fragments) : All fragments except the last one have this bit set. It is required. When all fragments of the datagram have arrived.

calculates the position of the fragment of data gram in the original IP datagram.

Fragment offset: tells where, in the current packet this ~~packet~~ fragment belongs. All fragments except the last one in data gram must be multiple of 8 bytes

Note: Identification, MF and fragment offset fields are used to implement fragmentation.

Time to live (TtL): It is counter used to limit the packet lifetime. In practice it is a hop count decremented at each Router and ~~dis~~ packet is discarded when it reaches 0 and warning packet is sent back to source host. The main ~~protocol field~~ ~~purp~~ purpose of TTL is to prevent IP datagrams from looping around forever in routing loop

protocol field: It tells the network layer at the destination host to which protocol the IP datagram belongs to i.e ICMP ⇒ 1, IGMP ⇒ 2 TCP ⇒ 6 UDP ⇒ 17 etc

Header checksum: 16 bit field which contains the checksum value of the entire header. It is used for error checking the of the Header.

Source IP address: 32 bit which contains the logical address of the sender of the datagram,

Destination IP Address : 32 bit field which defines the logical address of the receiver of the data

Options field is used for several purposes such as

(*) Route Record : used to record the ip address of the routing through which the data gram passes on its way. In option field maximum 9 IP address can be recorded.

(*) Source Routing : used to specify the route that the datagram must take to reach the destination

(*) padding : Addition of dummy data to fill up unused space in the transmission unit.

IP Address - classes. & CIDR - classless InterDomain Routing (Discussed in Tutorial classes)

IPV6 : Addressing

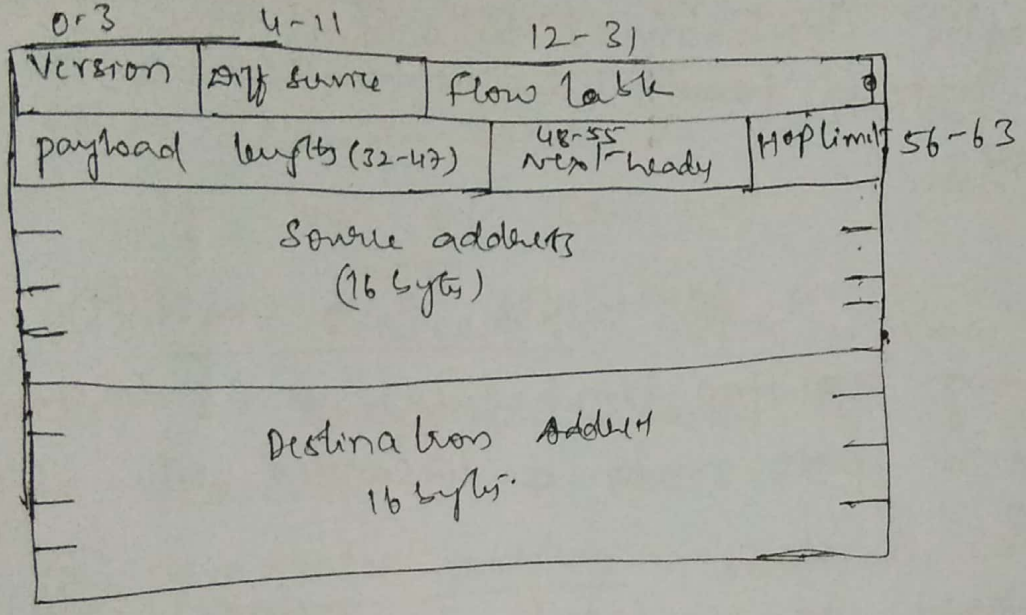
1) IPV6 has longer address than IPV4 i.e IPV6 uses 128 bits for logical addressing providing an unlimited Internet address.

2) IPV6 has the simplified header format it contains only seven fields. This allows the routers to process packet faster and thus improves throughput & delay

IPV6 provides a better and advanced features for security.

* In IPV6 more attention is given to QoS

IPV6 Header format



Version : Represents the version of the protocol 0110

Differentiated Services : Originally called Traffic class 8 bits field. The first 6 bits are used for type of services that the router should provide to this packet and LSB 2 bits are used for Explicit Congestion notification (ECN)

Flow Label : provides a way for a source and Destination to mark groups of packets that have the same requirement and should be treated in the same way by the network

Payload length : This field indicates the number of bytes of data follow the 40 byte header of IPv6

Next header : This field is used to indicate either the type of Extension Headers (six extension Headers) (Routing Header, Fragment Header, Authentication Header etc) or if the extension Header is not present then it tells which transport protocol handler (ex. TCP, UDP) to pass the packet.

Hop limit : It is same as TTL in IPv4. The value of Hop limit field is decremented by 1 as it pass router & discarded when it reaches 0

Source address : It indicates the address of the packet Sender. It contains 8 fields of 32 bits each separated by colon (:). Hence it is of 128 bit address

Destination Address : It indicates the address of the packet Receiver.

Internet control protocols (ICP)

In addition to IP which is used for data transfer the internet has several Companion protocols used in network layer.

are

(*) Internet Control Message protocol (ICMP) is used for reporting errors and management queries. It is a supporting protocol and used by network devices like routers for sending error messages and operation information. The most important messages are

(*) Destination Unreachable informs that the router cannot locate the destination or a packet with DF bit set can not be delivered.

(*) TIME Exceeded: This informs that the packet is dropped because its TTL counter has reached 0

(*) parameter problem: indicates an illegal value has been detected in the header field.

(*) Source quench: It is choke packet to decrease traffic rate when the congestion is detected

(*) Redirect message: is used by the routers to tell the sending host to update to better route & send packet in other route.

(*) Echo & Echo Reply: messages are sent by host to see if a given destination is reachable and currently alive, on receiving Echo destination sends ECHO REPLY message

These messages are used in ping utility

(*) Timestamp Request and Timestamp Reply similar to Echo & Echo replay except that the arival and replay time are recorded.

(*) Router Advertisement and Router Solicitation: messages are used to let hosts find nearby routers to send packets of the local network

ARP (Address Resolution protocol):

Computer programs / applications use logical address (IP address) to send / receive messages, but actual communication happens over the physical address (MAC). ARP protocol maps IP address to MAC address

Step 1 When a source machine want to communicate with an IP address, then it broadcast the packet on to the Ethernet asking who has the specified IP (Eg 192.168.1.20)

Step 2: The broadcasted message will arive at every machine on the Ethernet

Step 3: Each machine on the net will check its IP address and the only machine which matches its IP with the received IP will respond with its ethernet address (MAC address)

RARP (Reverse Address Resolution protocol): Allows the host to discover its IP address when it know only its physical address. It is used when the diskless computer is booted.

DHCP Dynamic Host Configuration protocol.

- It assigns dynamic IP address to the computers on the network.
- The computer broadcast a request for an IP on its network through DHCP DISCOVER packet.
- When ever the server receives the request, it allocates an IP address and sends to the host through DHCP OFFER.
- The Server identifies the host using MAC-ID
- The issue with DHCP is how long an IP address should be allocated. If a host leaves the network and does not return its IP address to DHCP server that address will be permanently lost. To prevent this the IP address assignment may be for fixed period of time and is called Leasing so just before the lease expires the host must request for DHCP renewal.
- DHCP is used by the ISPs to set the parameters of devices like IP address, network mask, default-gateway. Gateway. and the IP address of DNS and time servers.
- DHCP is a replacement for earlier RARP & BOOTP protocols.

The Transport Layer.

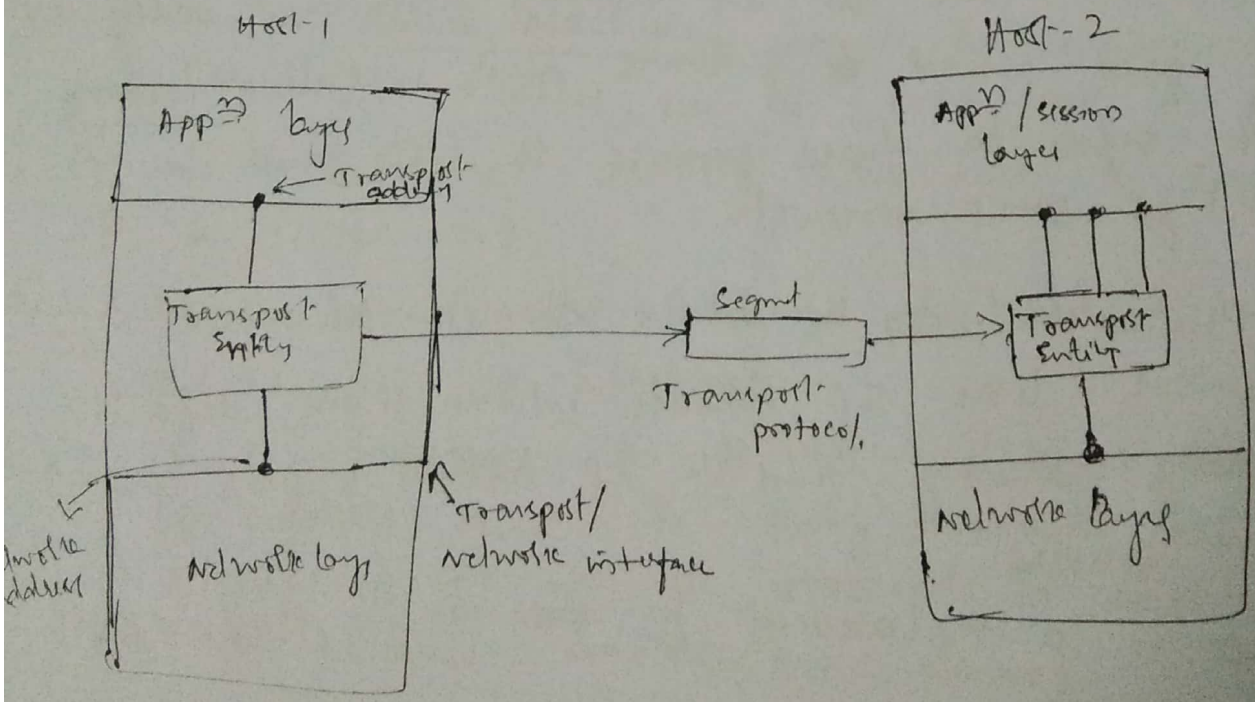
(*) The protocols at this layer - takes care & delivery & data flows an application program on one device to an application program on another device. &

(*) ST is the first ~~end-end~~ - end-to-end layer in OSI model.

(*) To provide the end-to-end comm^{n}, transport layer makes use &. the services provided by the network layer.

(*) The Software / Hardware units in the transport layer that does the work is called & Transport Entity

The relationship & network, transport and application layer is

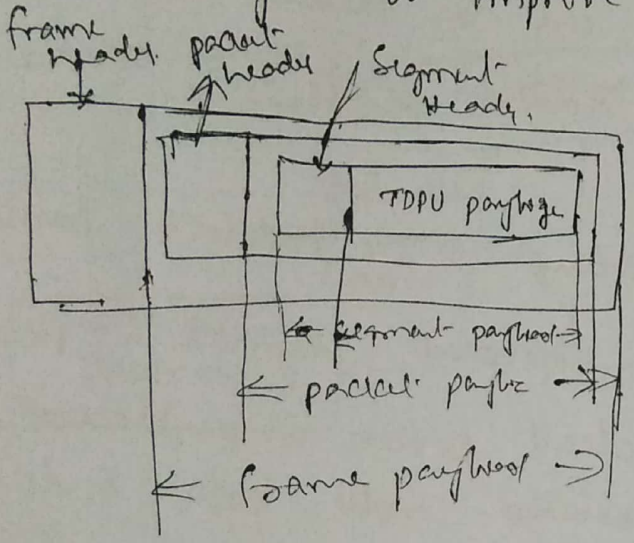


There are 2 types & transport service

(*) Connection oriented transport service &
(*) connection less transport service.

thoughts both Transport and network layer provide Connection oriented and connectionless Service, the Transport Code runs on the users machine but the network layer runs on the routers.

If the network layer offers inadequate service (loss of packets, router crash etc) then users will not have any control over the network layer. So it is better to put a layer on top of network layer to improve the QoS



Transport Service primitive :

To allow users to access the transport service, the transport layer provides some operation (interfaces) to application programs. The Transport Service primitives for to use the Transport Services are

SOCKET : creates a new communication endpoint and allocates table space for it.

BIND : newly created sockets do not have network address, so BIND primitive associates the local address with a socket.

LISTEN : Announces the willingness to accept connections and allocates space to queue incoming calls from several client try to connect at the same time.

ACCEPT: passively establish an incoming connection. When a request for connection arrives, the transport entity creates a new socket and returns file descriptor to client which can be used for reading and writing in the standard way.

At client side SOCKET will be created but BIND is not used, because this address does not matter for the server.

CONNECT : Actively attempt to establish connection process, and when the appropriate segment is received from the server, the client process is unblocked and the connection is established. Now both side can use SEND and RECEIVE to transmit and receive data over full-duplex communication.

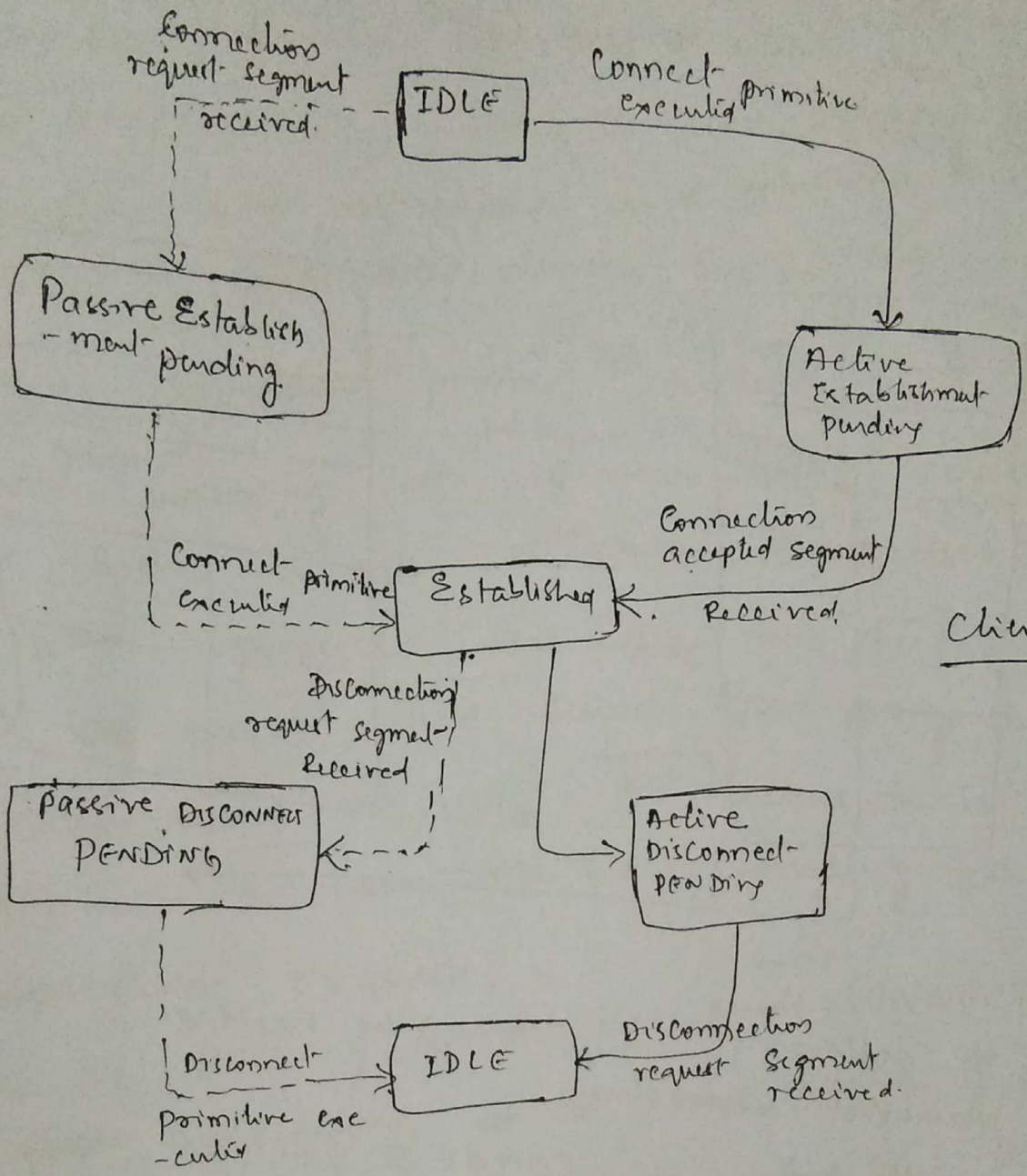
SEND : Send data over the connection

RECEIVE : Receive data from connection

CLOSE : Release the connection, when both sides executed ~~Close~~ CLOSE primitive, the connection is released.

...tate diagram scheme.

fig: Simple Connection Management



...ervers

Client-

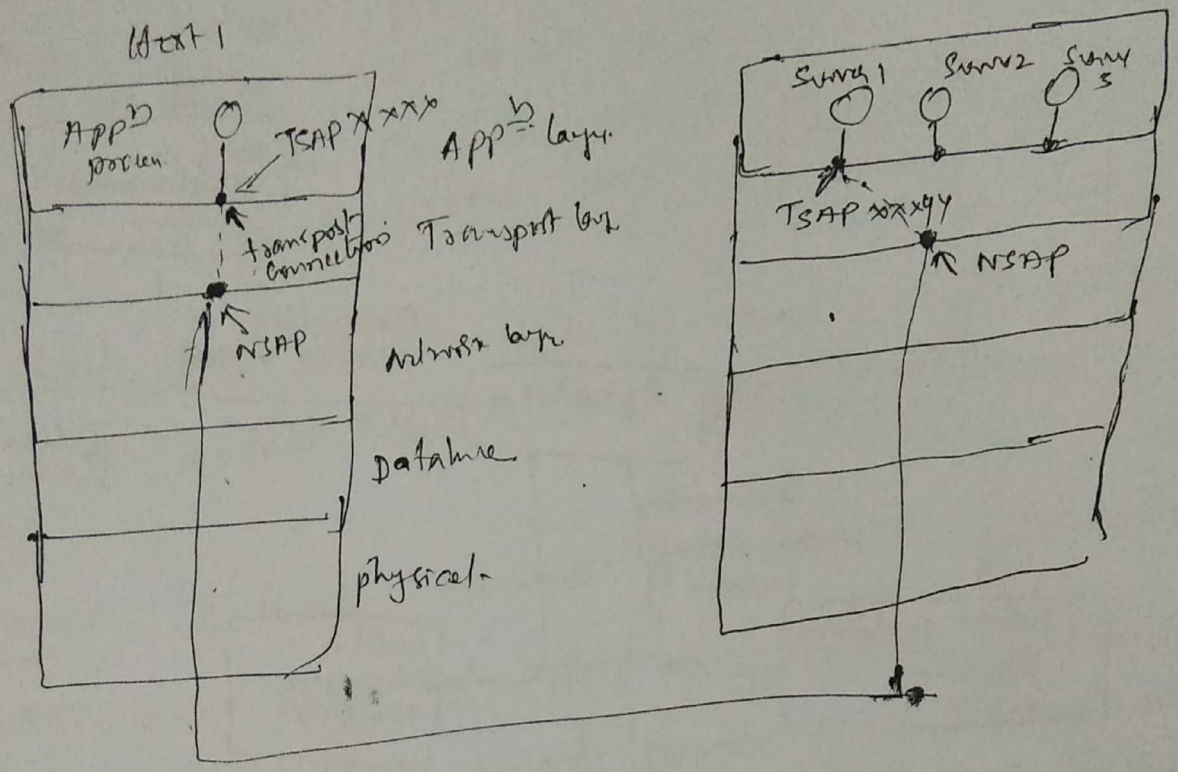
Elements of Transport protocols

(*) The transport service is implemented by transport protocol used b/n two transport entities.

(*) The important components are

Addressing: when a process needs to set up a connection to a process on remote system. To establish connection process need to define the Transport Address

also called port (or) TSAP (Transport service Access points) at transport layer and these end points at network layer are called NSAP (Network Service Access points)



So the following steps may be needed to establish a transport connection

1) client issues a CONNECT request with its TSAP Number (TSAP XXXX) and Server TSAP number (TSAP XXYY) to establish transport connection with the destination.

2) Application process (client) sends request for service (weather report)

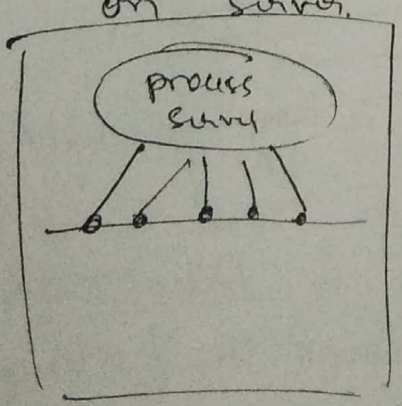
3) The server process responds with service (weather report)

4) The transport connection then released

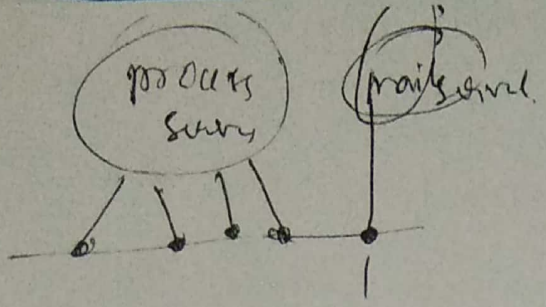
Mapper: Special proces that returns the TSAP (port) number in response to service name (/etc/services) Then the user releases the connection with portmapper and establishes a new connection with the desired service

Initial Connection protocol: Instead of every server listening at wellknown TSAP, each machine that wishes to offer services to remote users has a special process server that acts as a proxy for rarely used servers. This proxy server is called inetd on UNIX. It listens to set of ports at the same time. waiting for a connection request.

When the user of the service sends CONNECT request, specifying the TSAP address of the service, If no server is waiting for them, then it gets connection to the process server as shown in figure.



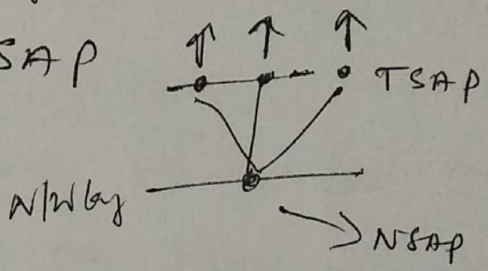
After it gets incoming, the process server forks the server process to inherit the existing connection with the user. The new server does the requested work, while the process server goes back to listening for new request as shown



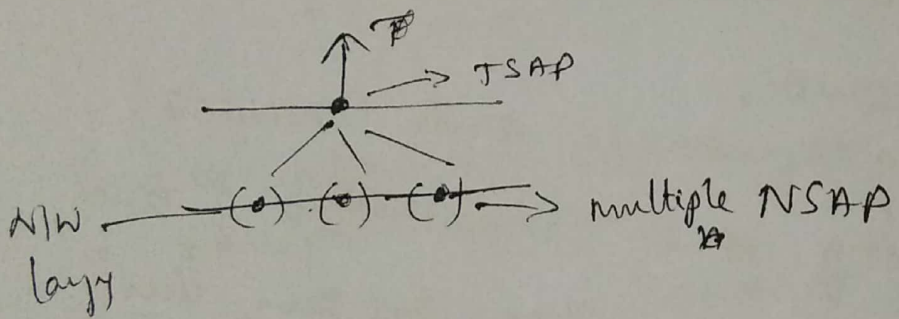
Multiplexing : There are two types of multiplexing

① Upward and ② Downward multiplexing

Upward multiplexing Through one NSAP request for multiple TSAP



Downward multiplexing



Internet ~~Transport protocols~~ :

Crash Recovery : In a network a host as well as a router can crash. Here for the issue is its recovery. The router crash recovery is taken care by network layer. But how to recover the host crashes.

partway through the transmission if source crashes

d. Comeback, is an attempt to recover. the ver. might send broadcast segment to all other hosts. The host can then send the last data segment which was never unacknowledge.

INTERNET TRANSPORT PROTOCOLS

The internet has two main protocols in transport layer. (1) connection less protocol UDP and (2) Connection oriented service TCP.

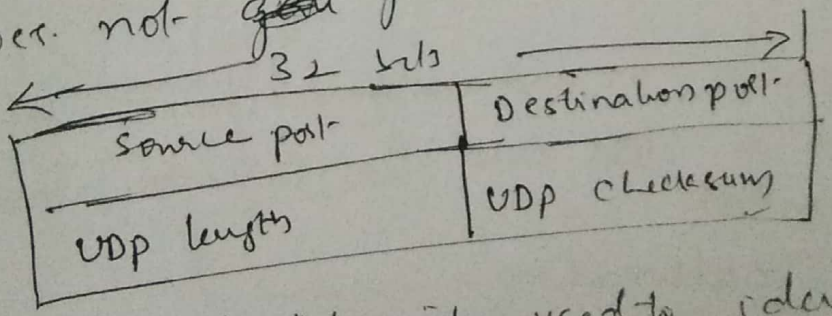
(1) UDP (User Datagram protocol)

(*) provides way for applications to send IP datagrams without having to establish a connection.

(*) It is suitable for streaming applications such as VoIP, multimedia streaming, computer gaming.

(*) UDP is more efficient in terms of both latency and band width.

(*) Does not guarantee ordered delivery of data.

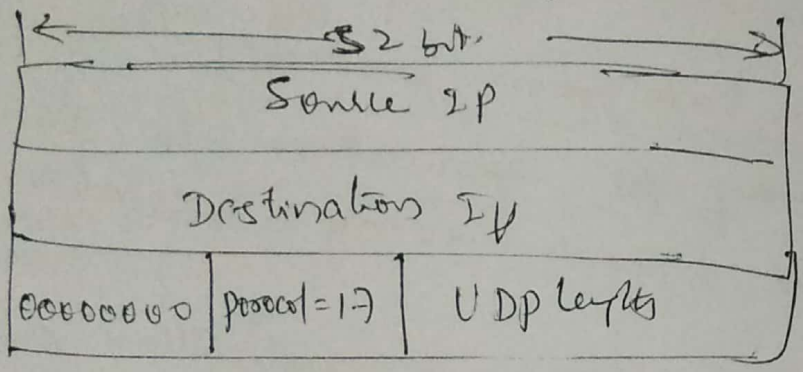


Source port : 16 bit info used to identify source port. just by copying Source port from the incoming segment to the destination port field of the outgoing segment.

UDP length: It shows the length of UDP header and data. The minimum is 8 bytes to cover header. The max length is 65,575 bytes

Checksum: It checksums the header, the header of the conceptual IP pseudoheader and the data, padded with zero octets at end (if necessary) to make multiple of two octets (16 bits because main group of 16 bits, adder and complements)

The pseudoheader for the case IPv4



It contains 32-bit IPv4 address of source and destination machine, protocol number (17) and the byte count for UDP segment

Applications of UDP

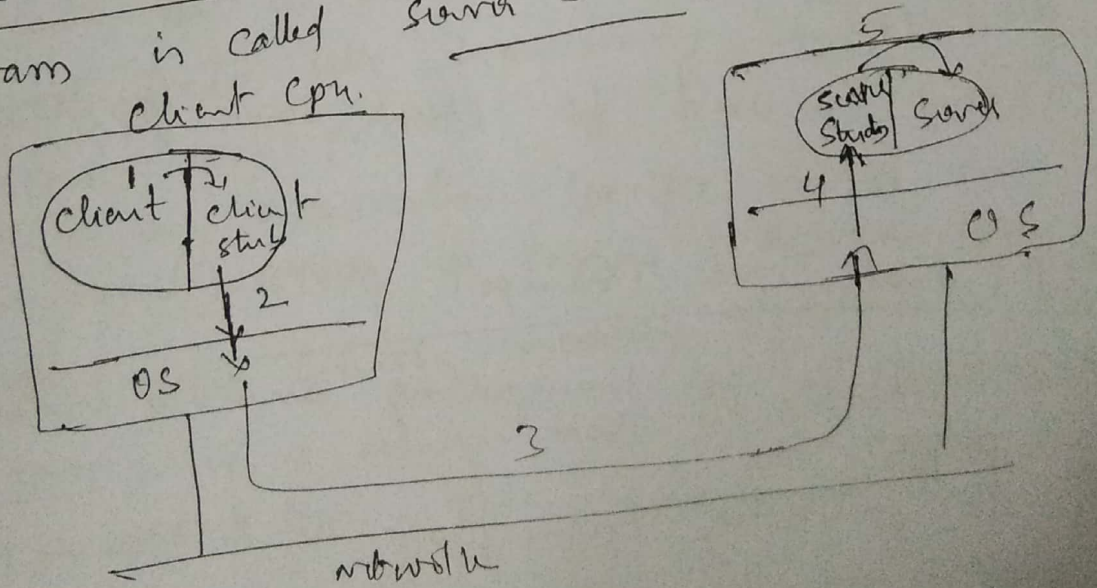
- used for simple request response communication when the size of data is less and less concern about flow and error control
- Suitable for multicasting
- used for some routing update protocol
- DNS, DHCP, NNP (network news protocol) implementations uses UDP as transport layer protocol

Remote procedure call (RPC)

machine 1 calls a procedure (function) on machine 2, the calling process on machine 1 is suspended and the execution of the called programs takes place on machine 2. The information from the caller to the callee is transported in parameters and can come back in procedure result and message passing is visible to application programmer. This technique is called RPC

Client stub: Small library procedure bound to client program. It represents the server procedure in the client's address space.

Server stub: The procedure bound with server program is called server stub.



Step 1: client calling the client stub, it is a local procedure call with the parameters

pushed onto stack in normal way.

Step 2: The client stub packing the parameters into a message and making a system call to send the message.

Step 3: Operating system sending the message from the client machine to server machine

Step 4: The operating system passing incoming packet to the server stub.

Step 5: Server stub calling the server procedure with unmarshaled parameters

Marshaling: Embedding the parameters into message is called marshing.

Note: ⊛ With RPC passing pointers is impossible because the client and server are in different address spa

⊛ UDP is used for client-server RPC.

RTP (Real Time Transport protocol)

⊛ RTP is for transporting audio & video data in packets.

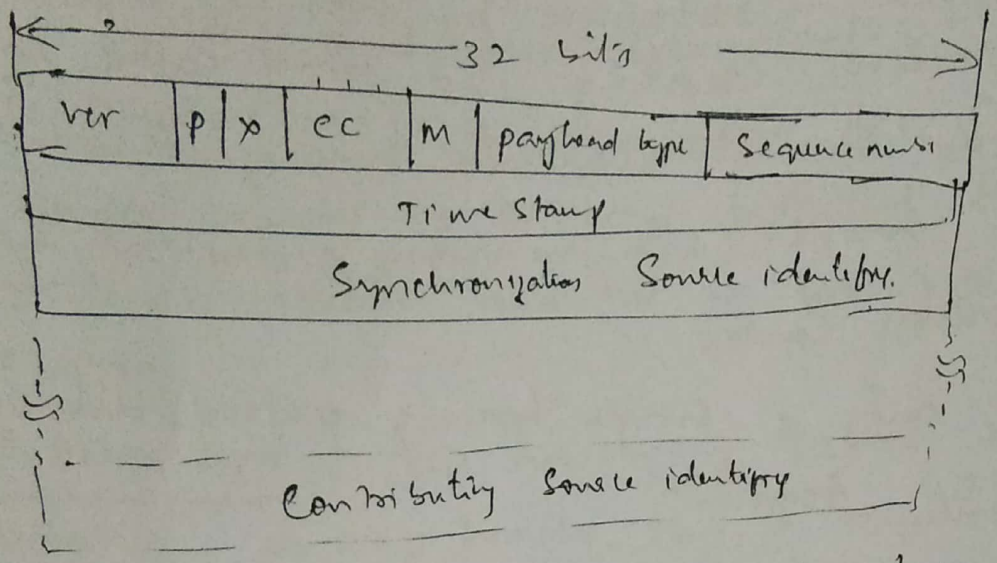
⊛ The processing takes place at the receiving to play out the audio & video

⊛ RTP runs in user space over UDP

⊛ The basic function of RTP is to multiplex

The basic function of RTP is to multiplex several real-time data streams onto a single stream of UDP packets. The UDP streams can be UNICASTED (or) Multicasted to multiple destinations.

RTP header



P => bit indicates that packet has been padded to a multiple of 4 bytes.

X => indicates that an extension header present.

CC => field tells how many contributing sources are present

M => bit is an application specific marker bit. It is used to mark the start of the video frame.

Payload type : tells which encoding algorithm is used.

Sequence number : is a count that is incremented on each RTP packet sent, It is used to detect lost packet.

Time stamp : produced by stream's source to note when the first sample in the packet was made

The Internet Transport protocol TCP

(5) TCP (Transmission Control protocol)

(x) provides reliable, sequenced, end-to-end byte stream

(x) TCP dynamically adapt to properties of the internet work

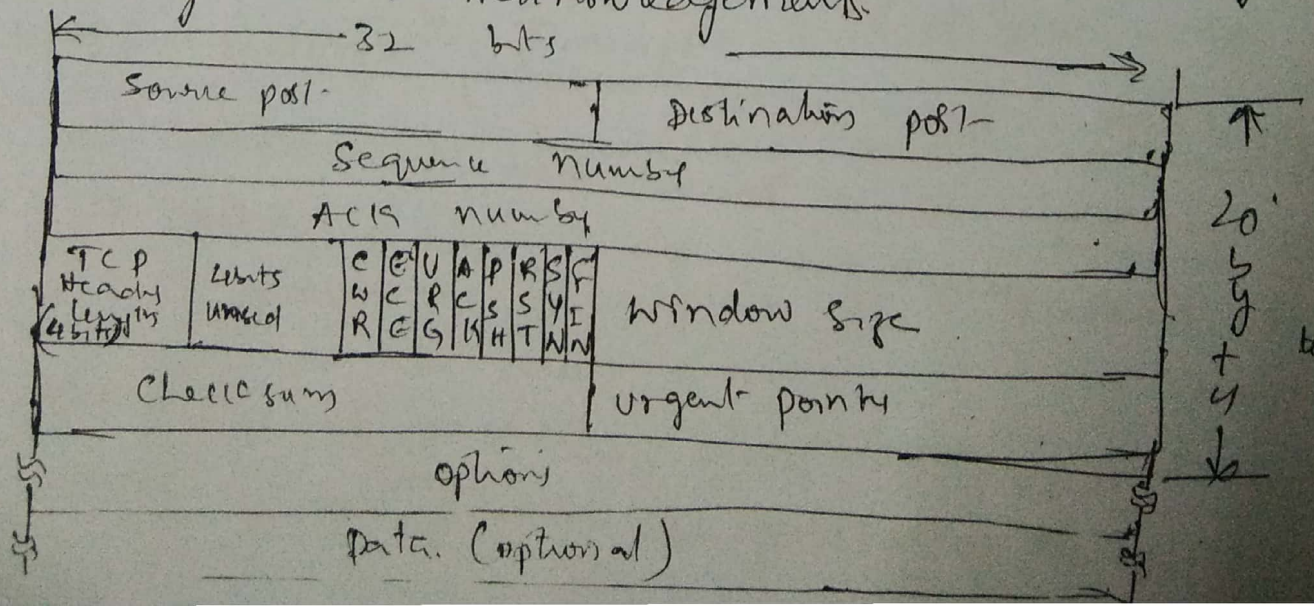
(x) TCP Service is obtained by both the sender and the receiver creating end points called sockets

(x) It provides Connection oriented Service, Reliable transport Service (with no missing & duplicate bytes) and Congestion Control

(x) TCP uses only a single type of protocol data unit called TCP segment

(x) The TCP segment begins with a fixed format & 20 byte header followed by options fields. After the options if any upto 65,535 − 20 (IP header) − 20 (TCP header) = 65 495 data bytes may follow

(x) A Segment with no data is used, for Controlling messages and ACKnowledgements



port & Destination port : 16 bit fields used identify end points at Source and destination

Sequence number (32 bit) Sequence number of the first data octet in this segment except when SYN is present. If SYN is present it is the initial sequence number (ISN) and first data octet is ISN+1

ACK number (32 bit) : A piggybacked ACK contains the Sequence number of the next byte that the TCP entity expects to receive and not for last byte correctly received.

TCP header length : Tells how many 32-bit words are contained in TCP header. This is required because the options field is of variable length

Flag bits : CWR and ECE are used to signal Congestion when ECN (Explicit Congestion Notification) is used. to tell the source to slow down the traffic.

CWR is set to signal (Congestion Window Reduced)

URG : used to indicate that there is a data in this segment which sender at the upper layer has marked urgent.

ACK : ACK field indicates that the value carried in ACK field is valid.

~~PSH~~ PSH : Push function : The receiver is requested to deliver the data to the application

upon arrival and not buffer it until a buffer has been received.

RST : Reset the Connection due to host crash or some other reason.

SYN : Synchronize the Sequence numbers.

FIN : No more data from Sender.

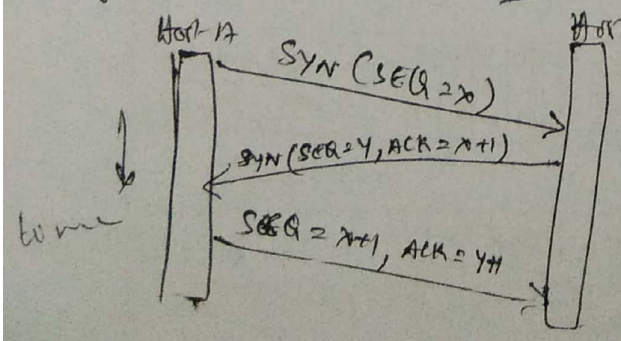
Receive Window Size (16 bits) : used for credit based flow control scheme, in bytes. at receiver side it shows the receiver is willing to accept.

Checksum (16 bits) : it checksums the header, the data, and Conceptual pseudo header.

Urgent Data pointer : points to the octet following the urgent data. This allows the receiver to know how much urgent data is coming.

TCP Connection Establishment :

Before the data transmission a connection must be established. In TCP the connection is established by means of Three-way handshake.



① Sender A initiates the connection to Host B by setting SYN bit.

② Host B the receiver Acknowledges the request and also initiates a Connection request.

Acknowledge to connection request & B

3-way hand shake procedure ensures that both transport entities agree upon their initial sequence numbers. which must be different.

TCP Connection Released

To release the connection either party can send. TCP segment with FIN bit set, telling that no more data to transmit

Each side will explicitly acknowledge the FIN & the other

Normally four. TCP segments are needed to release a connection One FIN & One ACK for each. direction

TCP flow control

flow control is the process of regulating the traffic b/n two machines to prevent the sender from flooding the receiver with too much data.

At the Receiver buffer. TCP uses sliding window with credit scheme to handle flow control

Assume that the sender wants to send a data of 4K (4096 B) to a receiver with buffer size of 4096 B

1) Initially buffer size = [Empty] 4KB

2) Sender transmits 2KB segment (data) with sequence number 0 as shown below.

[Sender] —Data = 2K SeqNo = 0→ [Receiver]

3) The packet is examined at the receiver and the it will be acknowledged by the receiver. The free buffer size is 2048B ∴ the credit value is ~~2048B~~ (window size) is 2048 Byte.

Receiver buffer is [2K data | Empty] 2048B

←ACK = 2048 SeqNo = 2048 Credit = 2048— [Receiver]

[Sender]

4) Now the sender transmits another 2048 bytes, which is acknowledged but the advertised window is (credit) is 0

Receiver Buffer = [Full]

[Sender] —Data = 2048 SeqNo = 2048→ [Receiver]

5)

Sender ←ACK = 4096 Credit = 0— Receiver.

5) The sender must stop sending data until the application process on the receiving host has removed removed some data from the buffer.

Receiver Buffer = [Application has read 2KB data. | 2K]

Sender. ←ACK = 4097 Credit = 2048— Receiver.

1D/29084/ Ess