UNIT – 3

NETWORK LAYER

Network Layer Services – Packet switching – Performance – IPV4 Addresses – Forwarding of IP Packets – Network Layer Protocols: IP, ICMP v4 – Unicast Routing Algorithms – Protocols – Multicasting Basics – IPV6 Addressing – IPV6 Protocol.

3.1 INTRODUCTION

- This is the third layer in OSI model in computer networking.
- Network layer provides support for end to end communication (helps to forward the packets from source to destination) by using routers and switches.
- Network layer manages the Quality of Services (QoS).
- The service provided by the network layer to the transport layer is called as network service.
- Support for Inter-networking: It is possible as the physical and data link layer work together to deliver the data packets from one node to another node in the network.

The three important functions of Network Layer are:

(1) **Path determination**

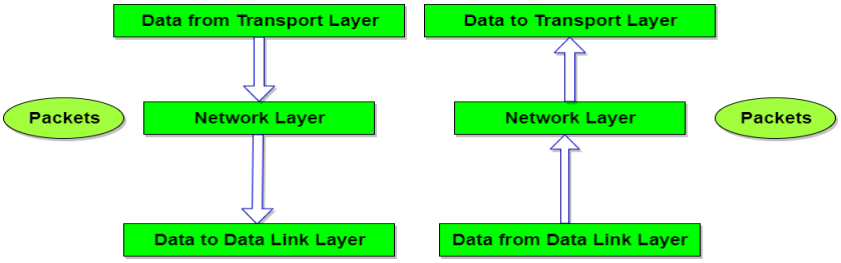
- When packet moves from source to destination, the network layer must determine the route or path taken by the packets.

(2) **Forwarding function**

- When packet reaches at the input of the router, then the router should move the packet to an appropriate link.

(3) **Call setup**

- Some networks require the router call setup along with the path, before the data flow.



Internet as a connectionless protocol:

- The connectionless network services are also known as datagrams.
- The internet at the network layer works as packet-switched network.
- Internet routes the packets by using universal address defined in the network layer.
- In connectionless service, the network layer protocol operates each packet independently.
- The internet is built from several heterogeneous networks. So, it is not possible to create a connection between the source and destination, before knowing the nature of the network.

IPv4

- The Internet Protocol version 4 (IPv4) is a connectionless protocol which is used for delivery mechanism (used by TCP/ IP protocols).
- The IPv4 is an unreliable protocol, but to make it reliable IPv4 is paired with a reliable protocol such as TCP.
- The IPv4 uses the datagram approach which means, that each datagram is handled independently and each datagram can follow the different route to the destination. Due to this, the datagrams sent from the same source to the same destination can reach at any order while some may get lost.

Datagrams

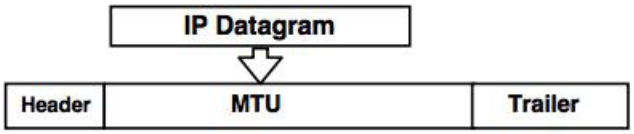
- Packets in the IPv4 layer are called as datagrams.
- A datagram is a variable- length packet consists of header and data.
- The size of header is 20 to 60 bytes, which is essential for routing and delivery.

Fragmentation

- The data travels through the different networks. Each router first decapsulates the IPv4 datagram from the received frame, then process it and again encapsulates in the another frame.
- The format and the size depends on the protocol used by the physical network through which it is going to travel.

Maximum Transfer Unit

- When a datagram is encapsulated in a frame, the total size of the datagram should be less than maximum size, which is defined or restricted by the hardware and software used in the network.



Maximum Transfer Unit (MTU)

3.2 NETWORK LAYER SERVICES:

(1) Guaranteed delivery

(2) Guaranteed delivery with bounded delay

(3) In-order packet delivery

(4) Guaranteed minimal bandwidth

(5) Guaranteed maximum jitter

(6) Security services

Guaranteed Delivery:

This service guarantees that the packet will eventually arrive at its destination.

Guaranteed Delivery with Bounded Delay:

This service not only guarantees delivery of the packet, but delivery within a specified host-to-host delay bound (for example, within 100 msec).

In-order Packet Delivery:

This service guarantees that packets arrive at the destination in the order that they were sent.

Guaranteed Minimal Bandwidth:

This network layer service emulates the behavior of a transmission link of a specified bit rate (for example, 1 Mbps) between sending and receiving hosts. As long as the sending host transmits bits (as part of packets) at a rate below the specified bit rate, then no packet is lost and each packet arrives within a prespecified host-to-host delay (for example, within 40 msec).

Guaranteed Maximum Jitter:

These service guarantees that the amount of time between the transmission of two successive packets at the sender is equal to the amount of time between their receipt at the destination (or that this spacing changes by no more than some specified value).

Security Services:

Using a secret session key known only by a source and destination host, the network layer in the source host could encrypt the payloads of all datagram being sent to the destination host.

The network layer in the destination host would then be responsible for decrypting the payloads. With such a service, confidentiality would be provided to all transport-layer segments (TCP and UDP) between the source and destination hosts. In addition to confidentiality, the network layer could provide data integrity and source authentication services.

The internet's network layer provides a single service, known as best-effort service. With best-effort service, timing between packets is not guaranteed to be preserved, packets are not guaranteed to be received in the order in which they were sent, nor is the eventual delivery of transmitted packets guaranteed

Design Issues with Network Layer

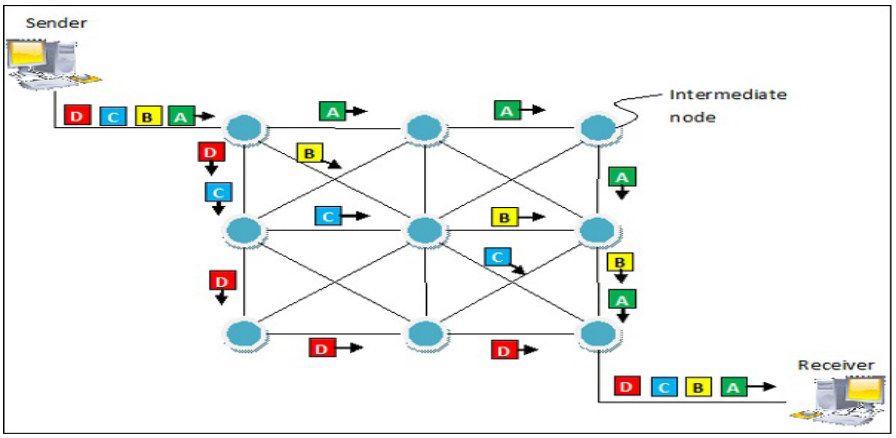
- A key design issue is determining how packets are routed from source to destination. Routes can be based on static tables that are wired into the network and rarely changed. They can also be highly dynamic, being determined anew for each packet, to reflect the current network load.
- If too many packets are present in the subnet at the same time, they will get into one another's way, forming bottlenecks. The control of such congestion also belongs to the network layer.
- Moreover, the quality of service provided (delay, transmit time, jitter, etc) is also a network layer issue.
- When a packet has to travel from one network to another to get to its destination, many problems can arise such as:
- The addressing used by the second network may be different from the first one.
- The second one may not accept the packet at all because it is too large.
- The protocols may differ, and so on.
- It is up to the network layer to overcome all these problems to allow heterogeneous networks to be interconnected.

3.3 PACKET SWITCHING

- Packet switching is a digital network transmission process in which data is broken into suitably-sized pieces or blocks for fast and efficient transfer via different network devices.
- When a computer attempts to send a file to another computer, the file is broken into packets so that it can be sent across the network in the most efficient way. These packets are then routed by network devices to the destination.

Process

- Each packet in a packet switching technique has two parts: a header and a payload. The header contains the addressing information of the packet and is used by the intermediate routers to direct it towards its destination. The payload carries the actual data.
- A packet is transmitted as soon as it is available in a node, based upon its header information. The packets of a message are not routed via the same path. So, the packets in the message arrive in the destination out of order. It is the responsibility of the destination to reorder the packets in order to retrieve the original message.



- The process is diagrammatically represented in the following figure. Here the message comprises of four packets, A, B, C and D, which may follow different routes from the sender to the receiver.

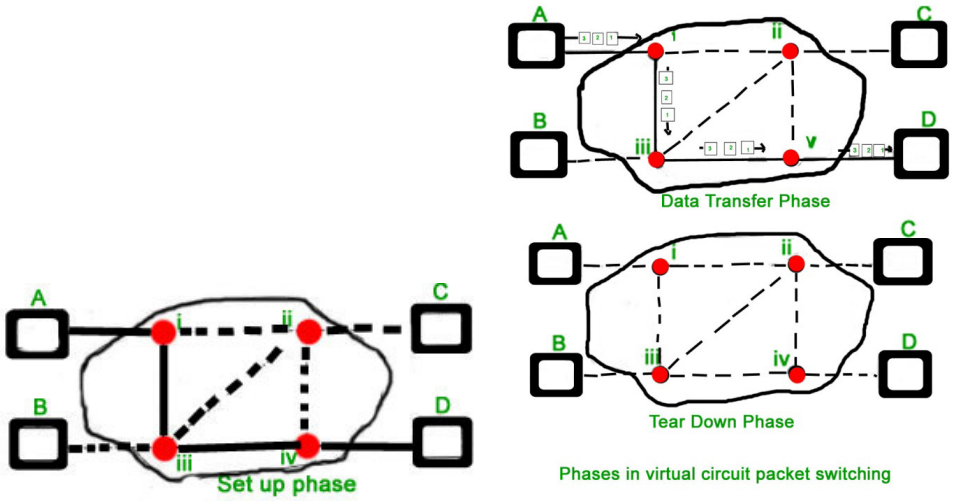
Packet Switching uses Store and Forward technique while switching the packets; while forwarding the packet each hop first store that packet then forward. This technique is very beneficial because packets may get discarded at any hop due to some reason.

More than one path is possible between a pair of source and destination. Each packet contains Source and destination address using which they independently travel through the network. In other words, packets belonging to the same file may or may not travel through the same path. If there is congestion at some path, packets are allowed to choose different path possible over existing network.

Packet-Switched networks were designed to overcome the weaknesses of Circuit-Switched networks since circuit-switched networks were not very effective for small messages.

Modes of Packet Switching:

Connection-oriented Packet Switching (Virtual Circuit):- Before starting the transmission, it establishes a logical path or virtual connection using signaling protocol, between sender and receiver and all packets belongs to this flow will follow this predefined route. Virtual Circuit ID is provided by switches/routers to uniquely identify this virtual connection. Data is divided into small units and all these small units are appended with help of sequence number. Overall, three phases takes place here- Setup, data transfer and tear down phase.



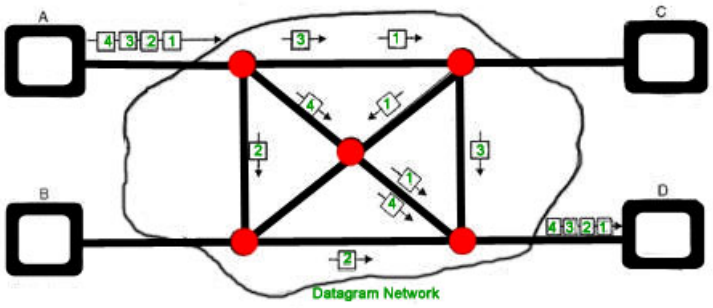
Phases in virtual circuit packet switching

All address information is only transferred during setup phase. Once the route to destination is discovered, entry is added to switching table of each intermediate node. During data transfer, packet header (local header) may contain information such as length, timestamp, sequence number etc.

Connection-oriented switching is very useful in switched WAN. Some popular protocols which use Virtual Circuit Switching approach are X.25, Frame-Relay, ATM and MPLS (Multi-Protocol Label Switching).

Connectionless Packet Switching (Datagram) :-

Unlike Connection-oriented packet switching, In Connectionless Packet Switching each packet contains all necessary addressing information such as source address, destination address and port numbers etc. In Datagram Packet Switching, each packet is treated independently. Packets belonging to one flow may take different routes because routing decisions are made dynamically, so the packets arrived at destination might be out of order. It has no connection setup and teardown phase, like Virtual Circuits.



Datagram Packet Switching

Packet delivery is not guaranteed in connectionless packet switching, so the reliable delivery must be provided by end systems using additional protocols.

3.4 PERFORMANCE IN PACKET SWITCHING

To send a packet from A to B there are delays since this is a Store and Forward network.

Delays in Packet switching:

- Transmission Delay
- Propagation Delay
- Queuing Delay
- Processing Delay

Transmission Delay:

Time taken to put a packet onto link. In other words, it is simply time required to put data bits on the wire/communication medium. It depends on length of packet and bandwidth of network.

Transmission Delay = Data size / bandwidth = (L/B) second

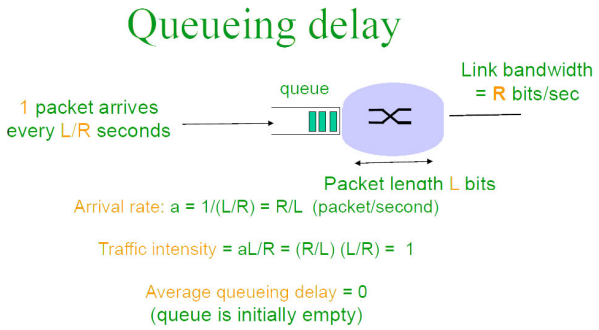
Propagation delay:

Time taken by the first bit to travel from sender to receiver end of the link. In other words, it is simply the time required for bits to reach the destination from the start point. Factors on which Propagation delay depends are Distance and propagation speed.

Propagation delay = distance/transmission speed = d/s

Queuing Delay:

Queuing delay is the time a job waits in a queue until it can be executed. It depends on congestion. It is the time difference between when the packet arrived Destination and when the packet data was processed or executed. It may be caused by mainly three reasons i.e. originating switches, intermediate switches or call receiver servicing switches.



Average Queuing delay = (N-1)L/(2*R)

where N = no. of packets

L=size of packet

R=bandwidth

Processing Delay:

Processing delay is the time it takes routers to process the packet header. Processing of packets helps in detecting bit-level errors that occur during transmission of a packet to the destination. Processing delays in high-speed routers are typically on the order of microseconds or less.

In simple words, it is just the time taken to process packets.

Total time or End-to-End time = Transmission delay + Propagation delay+ Queuing delay +Processing delay

For M hops and N packets –**Total delay** = M*(Transmission delay + propagation delay) + (M-1)*(Processing delay + Queuing delay) + (N-1)*(Transmission delay)

For N connecting link in the circuit –

Transmission delay = N*L/R

Propagation delay = N*(d/s)

Question :

How much time will it take to send a packet of size L bits from A to B in given setup if Bandwidth is R bps, propagation speed is t meter/sec and distance b/w any two points is d meters (ignore processing and queuing delay) ?

A---R1---R2---B

Ans:

N = no. of links = no. of hops = no. of routers +1 = 3

File size = L bits

Bandwidth = R bps

Propagation speed = t meter/sec

Distance = d meters

Transmission delay = (N*L)/R = (3*L)/R sec

Propagation delay = N*(d/t) = (3*d)/t sec

Total time = 3*(L/R + d/t) *sec*

In a Packet switch network having Hops= 4, transfer 10 packets from A to B given packet size is L bits. Bandwidth to transfer data is R Mbps and speed of propagation is S meter/sec. Assume processing delay= P seconds and distance between two point is D meters. Find total time required for 10 packets to reach A from B.

A---R1---R2---R3---B

Explanation :

No. of hops= No. of links = M= 4

Here we send 10 packets, also since there is no acknowledgement of packet received required we perform parallel processing. When the 1st packet reaches R2, the second packet reaches R1.

Formulas used-

R is in Mbps so convert to bps by multiplying 10^6.

Bandwidth=R*(10^6) bps

Packet size =L bits

Transmission delay= Packet size/Bandwidth =L/(R*(10^6))

Propagation Delay = Distance / Speed = D /S

Processing delay is in seconds no change

Delay can also be calculated as : Delay for 1st packet to reach + delay for (N-1) packets

Delay for 1st packet = M*(Propagation delay + Transmission delay) +
(M-1)*(Processing delay + Queuing delay)

Delay for N-1 remaining packets = (N-1)*(Transmission delay)

So finally applying the formula and putting the values we get-

Total delay = 4*(L/(R*(10^6) + D/S) + (4-1)*(P + 0) + (10-1)*(L/(R*(10^6))

Advantages and Disadvantages of Packet Switching

Advantages:

- Delay in delivery of packets is less, since packets are sent as soon as they are available.
- Switching devices don't require massive storage, since they don't have to store the entire messages before forwarding them to the next node.
- Data delivery can continue even if some parts of the network faces link failure. Packets can be routed via other paths.
- It allows simultaneous usage of the same channel by multiple users.
- It ensures better bandwidth usage as a number of packets from multiple sources can be transferred via the same link.

Disadvantages:

- They are unsuitable for applications that cannot afford delays in communication like high quality voice calls.
- Packet switching high installation costs.
- They require complex protocols for delivery.
- Network problems may introduce errors in packets, delay in delivery of packets or loss of packets. If not properly handled, this may lead to loss of critical information.

3.5 IPV4 ADDRESSES

- IP stands for Internet Protocol and describes a set of standards and requirements for creating and transmitting data packets, or datagram, across networks.
- The Internet Protocol (IP) is part of the Internet layer of the Internet protocol suite. In the OSI model, IP would be considered part of the network layer. IP is traditionally used in conjunction with a higher-level protocol, most notably TCP. The IP standard is governed by RFC 791.

How IP works

- IP is designed to work over a dynamic network. This means that IP must work without a central directory or monitor, and that it cannot rely upon specific links or nodes existing. IP is a connectionless protocol that is datagram-oriented., so each packet must contain the source IP address, destination IP address, and other data in the header to be successfully delivered.
- Combined, these factors make IP an unreliable, best effort delivery protocol. Error correction is handled by upper level protocols instead. These protocols include TCP, which is a connection-oriented protocol, and UDP, which is a connectionless protocol.

IP versions

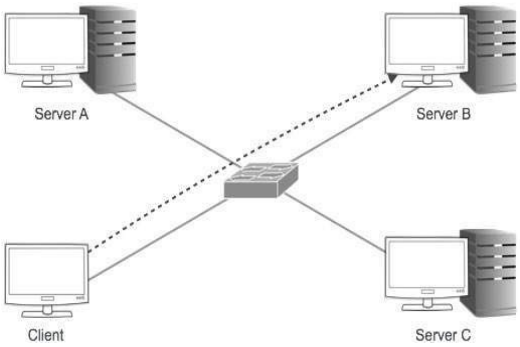
- There are two versions of IP in use today, IPv4 and IPv6. The original IPv4 protocol is still used today on both the internet, and many corporate networks. However, the IPv4 protocol only allowed for 232 addresses. This, coupled with how addresses were allocated, led to a situation where there would not be enough unique addresses for all devices connected to the internet.
- IPv6 was developed by the Internet Engineering Task Force (IETF), and was formalized in 1998. This upgrade substantially increased the available address space and allowed for 2128 addresses. In addition, there were changes to improve the efficiency of IP packet headers, as well as improvements to routing and security.

IPv4 addresses

- IPv4 addresses are actually 32-bit binary numbers, consisting of the two subaddresses (identifiers) mentioned above which, respectively, identify the network and the host to the network, with an imaginary boundary separating the two. An IP address is, as such, generally shown as 4 octets of numbers from 0-255 represented in decimal form instead of binary form.
- For example, the address 168.212.226.204 represents the 32-bit binary number 101010 00.11010100.11100010.11001100.
- The binary number is important because that will determine which class of network the IP address belongs to.
- An IPv4 address is typically expressed in dotted-decimal notation, with every eight bits (octet) represented by a number from one to 255, each separated by a dot. An example IPv4 address would look like this:

192.168.17.43

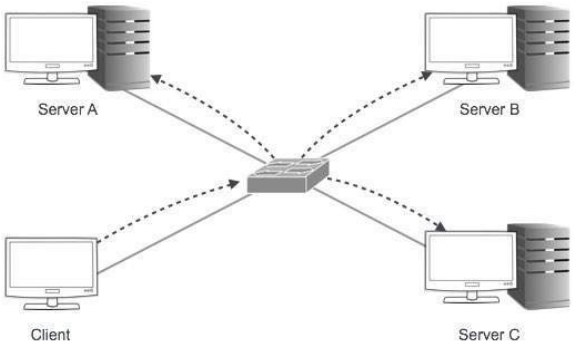
- IPv4 addresses are composed of two parts. The first numbers in the address specify the network, while the latter numbers specify the specific host. A subnet mask specifies which part of an address is the network part, and which part addresses the specific host.
- A packet with a destination address that is not on the same network as the source address will be forwarded, or routed, to the appropriate network. Once on the correct network, the host part of the address determines which interface the packet gets delivered to.



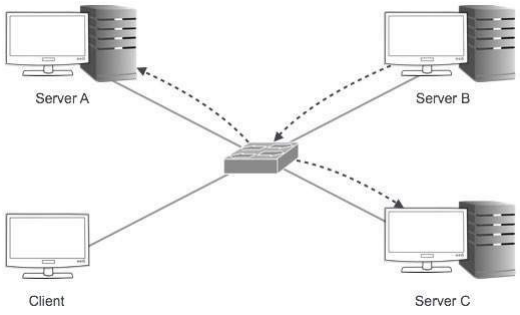
Unicast Addressing Mode:

In this mode, data is sent only to one destined host. The Destination Address field contains 32- bit IP address of the destination host. Here the client sends data to the targeted server:

Broadcast Addressing Mode:



In this mode, the packet is addressed to all the hosts in a network segment. The Destination Address field contains a special broadcast address, i.e. 255.255.255.255. When a host sees this packet on the network, it is bound to process it. Here the client sends a packet, which is entertained by all the Servers:



Multicast Addressing Mode:

This mode is a mix of the previous two modes, i.e. the packet sent is neither destined to a single host nor all the hosts on the segment. In this packet, the Destination Address contains a special address which starts with 224.x.x.x and can be entertained by more than one host.

Here a server sends packets which are entertained by more than one servers. Every network has one IP address reserved for the Network Number which represents the network and one IP address reserved for the Broadcast Address, which represents all the hosts in that network.

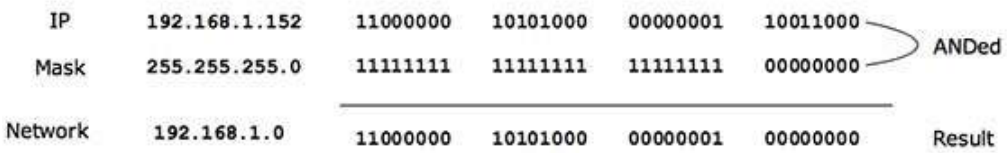
Hierarchical Addressing Scheme

IPv4 uses hierarchical addressing scheme. An IP address, which is 32-bits in length, is divided into two or three parts as depicted:

A single IP address can contain information about the network and its sub-network and ultimately the host. This scheme enables the IP Address to be hierarchical where a network can have many sub-networks which in turn can have many hosts.

Subnet masks

- A single IP address identifies both a network, and a unique interface on that network. A subnet mask can also be written in dotted decimal notation and determines where the network part of an IP address ends, and the host portion of the address begins.
- When expressed in binary, any bit set to one means the corresponding bit in the IP address is part of the network address. All the bits set to zero mark the corresponding bits in the IP address as part of the host address.
- The bits marking the subnet mask must be consecutive ones. Most subnet masks start with 255. and continue on until the network mask ends. A Class C subnet mask would be 255.255.255.0.



- This way the Subnet Mask helps extract the Network ID and the Host from an IP Address. It can be identified now that 192.168.1.0 is the Network number and 192.168.1.152 is the host on that network.

Binary Representation

- The positional value method is the simplest form of converting binary from decimal value. IP address is 32 bit value which is divided into 4 octets. A binary octet contains 8 bits and the value of each bit can be determined by the position of bit value '1' in the octet.

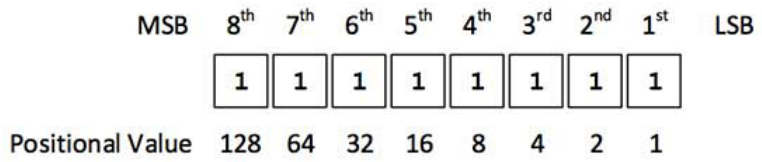


Fig: Binary Representation

- Positional value of bits is determined by 2 raised to power (position – 1), that is the value of a bit 1 at position 6 is 2^(6-1) that is 2^5 that is 32. The total value of the octet is determined by adding up the positional value of bits. The value of 11000000 is 128+64 = 192

IPv4 Classful Addressing

- To provide flexibility in the number of addresses distributed to networks of different sizes, 4-octet (32-bit) IP addresses were originally divided into three different categories or classes:
  - class A,
  - class B,
  - class C.

  Each address class specifies a different number of bits for its network prefix and host number:
- **Class A** addresses use only the first byte (octet) to specify the network prefix, leaving 3 bytes to define individual host numbers.
- **Class B** addresses use the first 2 bytes to specify the network prefix, leaving 2 bytes to define host addresses.
- **Class C** addresses use the first 3 bytes to specify the network prefix, leaving only the last byte to identify hosts.
- In binary format, with an x representing each bit in the host number, the three address classes can be represented as follows:

  00000000 xxxxxxxx xxxxxxxx xxxxxxxx (Class A)

  00000000 00000000 xxxxxxxx xxxxxxxx (Class B)

  00000000 00000000 00000000 xxxxxxxx (Class C)
- Because each bit (x) in a host number can have a 0 or 1 value, each represents a power of 2. For example, if only 3 bits are available for specifying the host number, only the following host numbers are possible: 111 110 101 100 011 010 001 000

- In each IP address class, the number of host-number bits raised to the power of 2 indicates how many host numbers can be created for a particular network prefix. Class A addresses have 224 (or 16,777,216) possible host numbers, class B addresses have 216 (or 65,536) host numbers, and class C addresses have 28 (or 256) possible host numbers.

Historical classful network architecture

Class	Leading bits	Size of *network number* bit field	Size of *rest* bit field	Number of networks	Addresses per network	Start address	End address
A	0	8	24	128 (2^7)	16,777,216 (2^{24})	0.0.0.0	127.255.255.255
B	10	16	16	16,384 (2^{14})	65,536 (2^{16})	128.0.0.0	191.255.255.255
C	110	24	8	2,097,152 (2^{21})	256 (2^8)	192.0.0.0	223.255.255.255

IPv4 Dotted Decimal Notation

- The 32-bit IPv4 addresses are most often expressed in dotted decimal notation, in which each octet (or byte) is treated as a separate number.
- Within an octet, the rightmost bit represents 20 (or 1), increasing to the left until the first bit in the octet is 27 (or 128). Following are IP addresses in binary format and their dotted decimal equivalents:

11010000 01100010 11000000 10101010 = 208.98.192.170

01110110 00001111 11110000 01010101 = 118.15.240.85

00110011 11001100 00111100 00111011 = 51.204.60.59

A

Destination	Route
Host B	R1, R2, Host B

R1

Destination	Route
Host B	R2, Host B

R2

Destination	Route
Host B	Host B

a. Routing tables based on route

A

Destination	Next Hop
Host B	R1

R1

Destination	Next Hop
Host B	R2

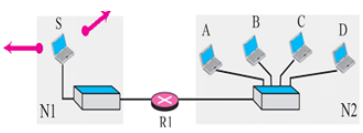
R2

Destination	Next Hop
Host B	---

b. Routing tables based on next hop

Destination	Next Hop
N2	R1

Destination	Next Hop
A	R1
B	R1
C	R1
D	R1

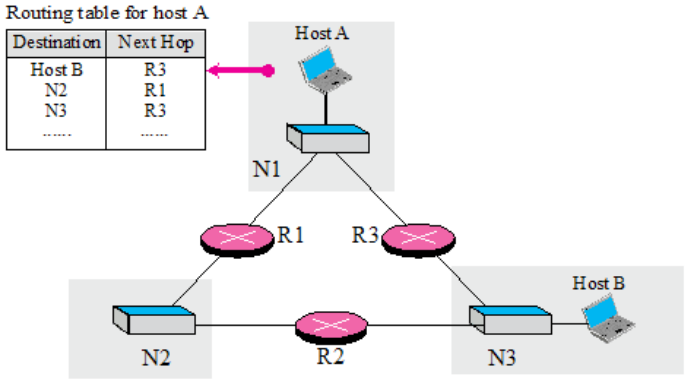


Network Specific Method

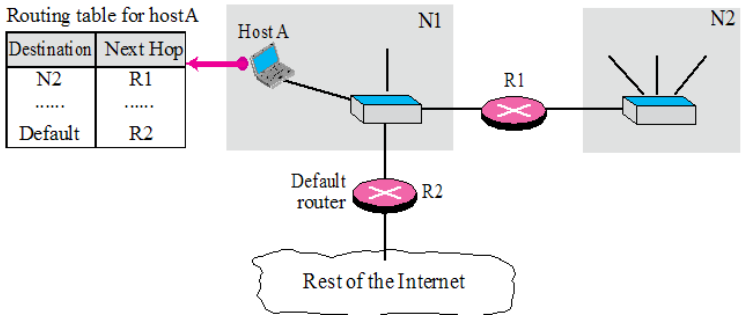
- Reduce the routing table and simplify the searching process
- The routing table has only one entry that defines the address of the destination network itself

Host-Specific Method

- The Destination host address is given in the routing table
- Inverse of network-specific method
- When administrator wants to have more control



Default Method



- Instead of listing all networks in the entire Internet host can just have one entry called the default

An Internet environment results from the interconnection of physical networks by routers. Each router is connected directly to two or more networks, hosts are generally connected to a single network, but this is not mandatory.

There are several types of routes:

- **Direct routing.**

  This is the case if the two machines who want to communicate are attached to the same network and therefore have the same IP network number. It may be two hosts or a router and a host. It is sufficient to perform the transport of the IP packet to determine the

recipient's physical address and encapsulate the datagram in a frame before sending it over the network.

- **Indirect Routing**.

  In this case, routing is more complex because it must determine the router to which datagram's are sent. The latter can thus be forwarded from router to router until they reach the destination host. The routing function is mainly based on the routing tables.

  Routing is done from the network number of the IP address of the destination host. The table contains for each network number to reach the IP address of the router to send the datagram. It may also include a default router address and direct routing indication. The difficulty comes from the routing initialization and updating of routing tables.

- **The subnetting.**

  This addressing technical and standardized routing to manage multiple physical networks from a single Internet IP address. The principle of subnetting is to divide the host number portion of an IP address subnet number and host number. Outside the site, the addresses are interpreted without taking account of subnetting, cutting being seen and treated from the inside.

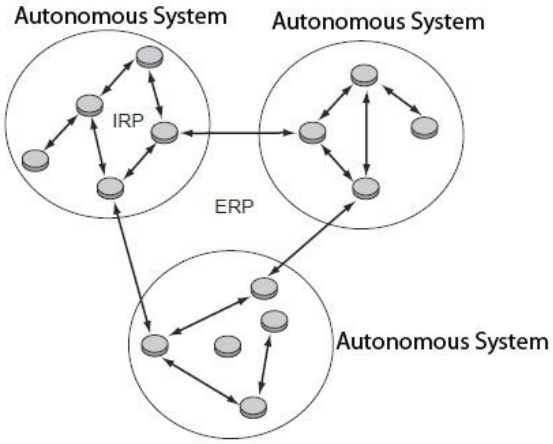
  Redistribution number allows the host to choose freely the number of machines based on the number of networks on the site. Conceptually, the addressing techniques and routing are the same. At the physical level, we use an address mask.

  The Internet has been so extensive that it had to be split into autonomous systems for easier management. Autonomous system called (AS) a set of routers and networks connected to each other, administered by an organization and exchanging packets through a same routing protocol.

  The routing protocol shared by all routers in an autonomous system is called interior routing protocol, or IRP (Interior Routing Protocol). An internal protocol does not need to be implemented outside the autonomous system. Therefore, one can choose its routing algorithm to optimize the internal routing. The interiors are also called protocols IGP (Interior Gateway Protocol).

  When an Internet network has multiple autonomous systems linked together, you have to use an external routing protocol, or ERP (Exterior Routing Protocol). ERP protocols must have knowledge of the various AS to accomplish their task. ERP protocols are also called EGP (Exterior Gateway Protocol).

  Figure shows an example of autonomous systems with PIR protocols interconnected by an ERP.



Interior and Exterior Routing Protocols

3.6 NETWORK LAYER PROTOCOLS

Network protocols provide what are called "link services". These protocols handle addressing and routing information, error checking, and retransmission requests. Network protocols also define rules for communicating in a particular networking environment such as Ethernet or Token Ring.

IP - This is short for Internet Protocol which works at the OSI network layer and is a routed protocol for forwarding layer 3 packets.

ICMP: ICMP stands for Internet Control Message Protocol. It is used to report some problem when routing a packet.

DHCP: DHCP stands for Dynamic Host Configuration Protocol. It helps to get the network layer address for a host.

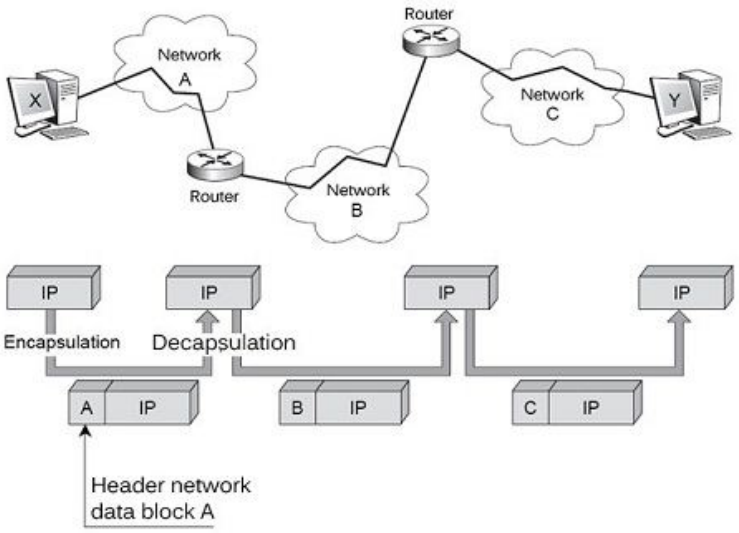
ARP: ARP acronym for Address Resolution Protocol. It helps to finds the link layer address of a host.

IP:

The Internet's basic protocol called IP for Internet Protocol. The objective of starting this protocol is assigned to interconnect networks do not have the same frame-level protocols or package level. The internet acronym comes from inter-networking and corresponds to an interconnection fashion: each independent network must transport in the weft or in the data area of the packet an IP packet, as shown in Figure.

There are two generations of IP packets, called IPv4 (IP version 4) and IPv6 (IP version 6). IPv4 has been dominant so far. The transition to IPv6 could accelerate due to its adoption in many Asian countries. The transition is however difficult and will last many years.

Internet Protocol (IP) of network layer contains addressing information and some control information that enables the packets to be routed.



How IP works

- IP is designed to work over a dynamic network. This means that IP must work without a central directory or monitor, and that it cannot rely upon specific links or nodes existing. IP is a connectionless protocol that is datagram-oriented., so each packet must contain the source IP address, destination IP address, and other data in the header to be successfully delivered.
- Combined, these factors make IP an unreliable, best effort delivery protocol. Error correction is handled by upper level protocols instead. These protocols include TCP, which is a connection-oriented protocol, and UDP, which is a connectionless protocol.

IP has two primary responsibilities:

(1) Providing connectionless, best effort delivery of datagrams through a internetwork. The term best effort delivery means that IP does not provides any error control or flow control. The term connectionless means that each datagram is handled independently, and each datagram can follow different route to the destination. This implies that datagrams sent by the same source to the same destination could arrive out of order.

(2) Providing fragmentation and reassembly of datagrams to support data links with different maximum transmission unit (MTU) sizes.

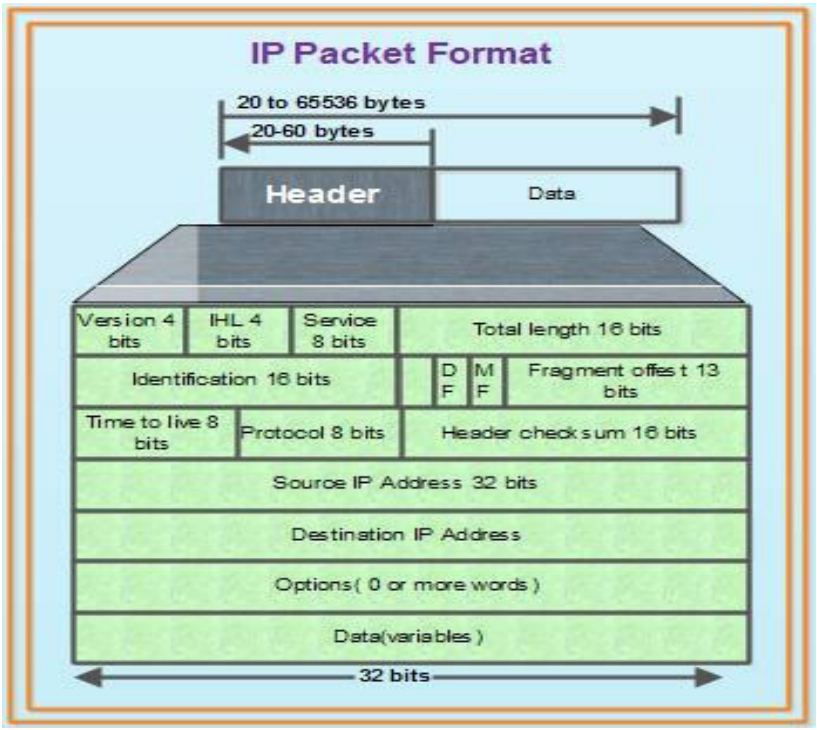
IP packet format

- Packets in the network layer are called datagrams.
- A datagram is a variable length packet consisting of two parts: header and data.
- The header is 20 to 60 bytes in length and contains information essential to routing and delivery

The various fields in IP header are:

(1) Version: It is a 4-bit field that specifies the version of IP currently being used. Two different versions of protocols are IPV4 (Internet Protocol Version 4) and IPV6 (Internet Protocol Version 6).

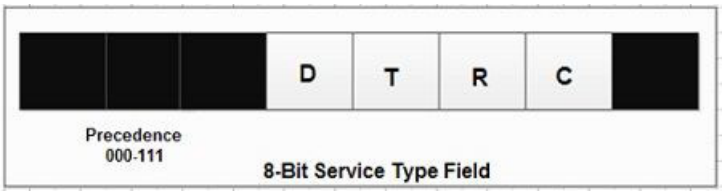
(2) IP Header Length (IHL): This 4-bit field indicates the datagram header length in 32 bit word. The header length i8 not constant in IP. It may vary from 20 to 60 bytes. When there are no options, the header length is 20 bytes, and the value of this field is 5. When the option field is at its maximum size, the value of this field is 15.



(3) Services: This 8 hit field was previously called services type but is now called differentiated services.

The various bits in service type are:

A 3-bit precedence field that defines the priority of datagram in issues such as congestion. This 3-bit subfield ranges from 0 (000 in binary) to 7 (111 in binary).



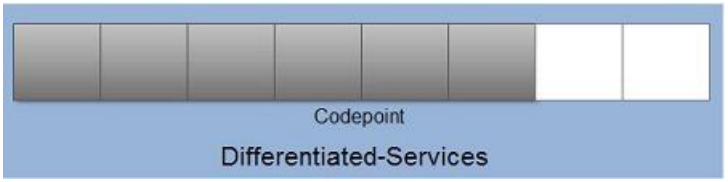
after 3-bit precedence there are four flag bits. These bits can be either 0 or 1 and only one of the bits can have value of 1 in each datagram.

The various flag bits are:

- D : Minimize delay
- T : Maximize throughout
- R : Maximize reliability
- C : Minimize Cost

The various bits in differentiated services are:

The first 6 bits defined a codepoint and last two bits are not used. If the 3 rightmost bits are 0s, the 3 leftmost bits are interpreted the same as the precedence bits in the service type interpretation.



(4) **Total length**: This 16 bit field specifies the total length of entire IP datagram including data and header in bytes. As there are 16 bits, the total length of IP datagram is limited to 65,535 (216 - 1) bytes.

(5) **Identification**: This 16 bit field is used in fragmentation. A datagram when passing through different networks may be divided into fragments to match the network frame size. Therefore, this field contains an integer that identifies the current datagram. This field is used to help piece together datagram fragments.

(6) **Flags:** Consists' of a 3-bit field of which the two low order bit DF, MF control fragmentation. DF stands for Don't Fragment. DF specifies whether the packet can be fragmented MF stands for more fragments. MF specifies whether the packet is the last fragment in a series of fragmented packets. The third or high order but is not used.

(7) **Fragment Offset:** This 13-bit field indicates the position of the fragment's data relative to the beginning of the data in the original datagram, which allows the destination IP process to properly reconstruct the original datagram.

(8) **Time to Live:** It is 8 bit field that maintain a counter that gradually decrements down to zero, at which point the datagram is discarded. This keeps the packet from looping endlessly.

(9) **Protocol:** This 8 bit field indicates which upper layer protocol receives incoming packets after IP processing is complete.

(10) **Header Checksum**: This 16 bit field contains a checksum that covers only the header and not the data.

(11) **Source IP address:** These 32-bit field contains the IP address of source machine.

(12) **Destination IP address**: This 32-bit field contains the IP address of destination machine.

(13) **Options**: This field allows IP to support various options such as security, routing, timing management and alignment.

(14) **Data:** It contains upper layer information.

Example [IP Datagram fields]

An IP packet has arrived with the first 8 bits as shown: 01000010, The receiver discards the packet. Why?

Solution:

There is an error in this packet. The 4 left-most bits (0100) show the version, which is correct. The next 4 bits (0010) show the header length, which means ($2 \times 4 = 8$), which is wrong. The minimum number of bytes in the header must be 20. The packet has been corrupted in transmission.

In an IP packet, the value of HLEN is 1000 in binary. How many bytes of options? Are being carried by this packet?

Solution:

The HLEN value is 8, which means the total number of bytes in the header is 8x4 or 32 bytes. The first 20 bytes are the main header, the next 12 bytes are the options

3.7 ICMP (INTERNET CONTROL MESSAGE PROTOCOL)

ICMP (Internet Control Message Protocol) is an error-reporting protocol network device like routers use to generate error messages to the source IP address when network problems prevent delivery of IP packets. ICMP creates and sends messages to the source IP address indicating that a gateway to the Internet that a router, service or host cannot be reached for packet delivery. Any IP network device has the capability to send, receive or process ICMP messages.

ICMP is not a transport protocol that sends data between systems.

While ICMP is not used regularly in end-user applications, it is used by network administrators to troubleshoot Internet connections in diagnostic utilities including ping and traceroute.

One of the main protocols of the Internet Protocol suite, ICMP is used by routers, intermediary devices or hosts to communicate error information or updates to other routers, intermediary devices or hosts. The widely used IPv4 (Internet Protocol version 4) and the newer IPv6 use similar versions of the ICMP protocol (ICMPv4 and ICMPv6, respectively).

ICMP messages are transmitted as datagrams and consist of an IP header that encapsulates the ICMP data. ICMP packets are IP packets with ICMP in the IP data portion. ICMP messages also contain the entire IP header from the original message, so the end system knows which packet failed

The ICMP header appears after the IPv4 or IPv6 packet header and is identified as IP protocol number 1. The complex protocol contains three fields:

The major type that identifies the ICMP message;

The minor code that contains more information about the type field; and The checksum that helps detect errors introduced during transmission.

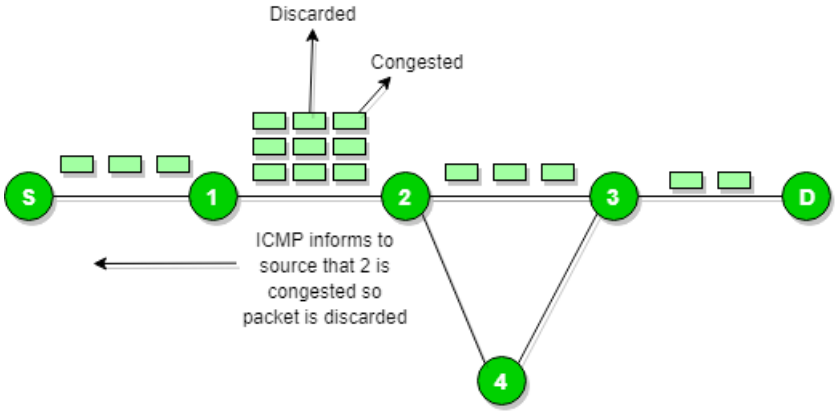
Following the three fields is the ICMP data and the original IP header to identify which packets actually failed.

- ICMP has been used to execute denial-of-service attacks (also called the ping of death) by sending an IP packet larger than the number of bytes allowed by the IP protocol.
- Since IP does not have a inbuilt mechanism for sending error and control messages. It depends on Internet Control Message Protocol (ICMP) to provide an error control. It is used for reporting errors and management queries. It is a supporting protocol and used by networks devices like routers for sending the error messages and operations information.

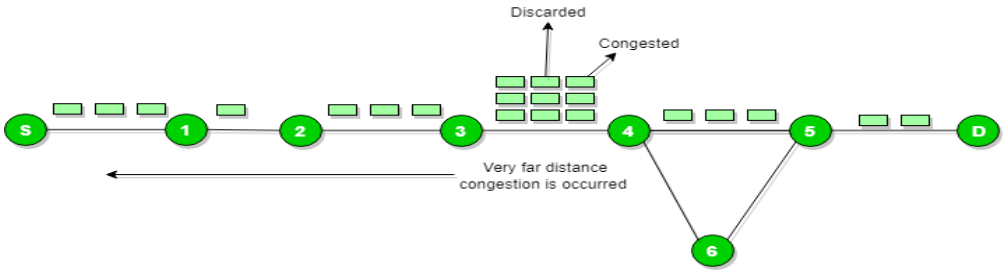
  e.g. the requested service is not available or that a host or router could not be reached.

Source quench message:

Source quench message is request to decrease traffic rate for messages sending to the host(destination). Or we can say, when receiving host detects that rate of sending packets (traffic rate) to it is too fast it sends the source quench message to the source to slow the pace down so that no packet can be lost.

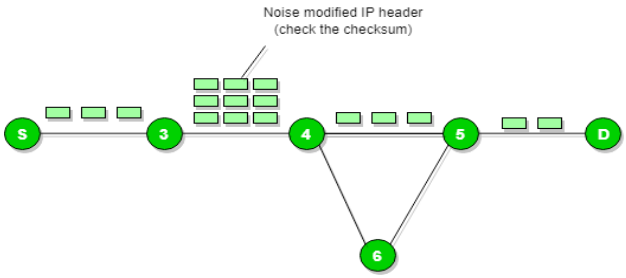


ICMP will take source IP from the discarded packet and informs to source by sending source quench message.Then source will reduce the speed of transmission so that router will free for congestion.



Parameter problem:

Whenever packets come to the router then calculated header checksum should be equal to received header checksum then only packet is accepted by the router.

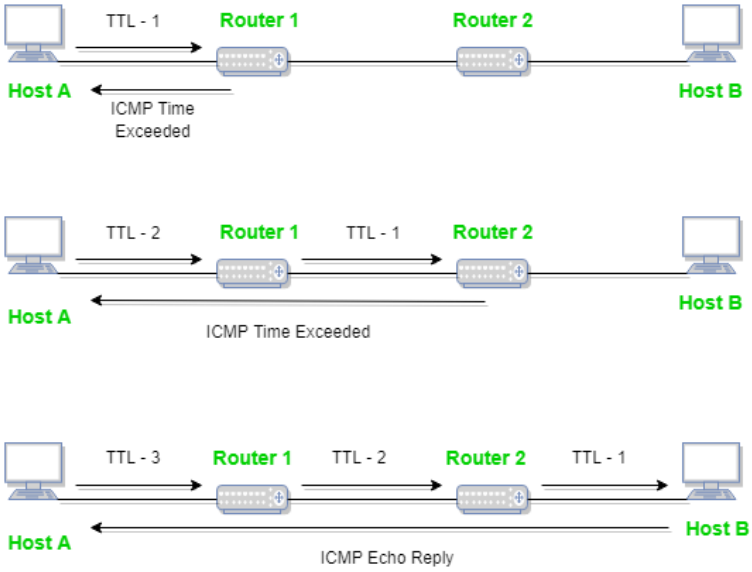


If there is mismatch packet will be dropped by the router.

ICMP will take the source IP from the discarded packet and informs to source by sending parameter problem message.

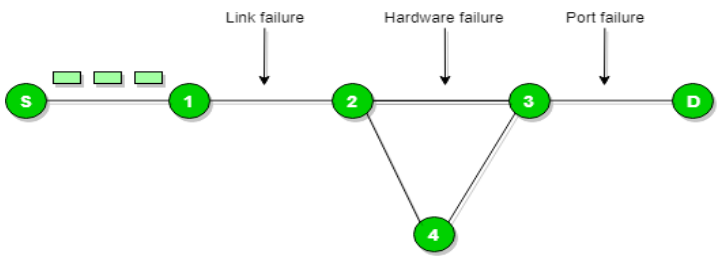
Time exceeded message:

When some fragments are lost in a network then the holding fragment by the router will be droped then ICMP will take source IP from discarded packet and informs to the source, of discarded datagram due to time to live field reaches to zero, by sending time exceeded message.



Destination un-reachable:

Destination unreachable is generated by the host or its inbound gateway to inform the client that the destination is unreachable for some reason.



There is no necessary condition that only router give the ICMP error message some time destination host send ICMP error message when any type of failure (link failure, hardware failure, port failure etc) happen in the network.

Redirection message:

- Redirect requests data packets be sent on an alternate route. The message informs to a host to update its routing information (to send packets on an alternate route).
- Ex. If host tries to send data through a router R1 and R1 sends data on a router R2 and there is a direct way from host to R2. Then R1 will send a redirect message to inform the host that there is a best way to the destination directly through R2 available. The host then sends data packets for the destination directly to R2.
- The router R2 will send the original datagram to the intended destination.
- But if datagram contains routing information then this message will not be sent even if a better route is available as redirects should only be sent by gateways and should not be sent by Internet hosts.

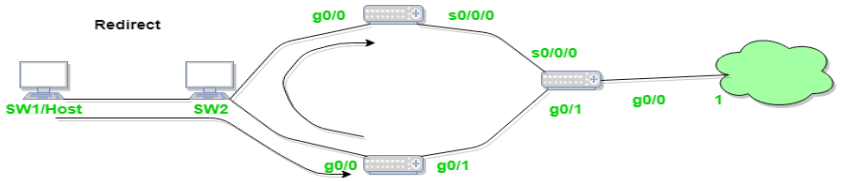


Figure - ICMP redirect Verification CCNP 2.0 100 - 101 (v - 71)

- ICMP Redirect
- ICMP Redirect for host
- ICMP Redirect for network
- How ICMP redirect work
- ICMP Redirect verification step by step

Whenever a packet is forwarded in a wrong direction later it is re-directed in a current direction then ICMP will send re-directed message.

Unicast Routing

Unicast means the transmission from a single sender to a single receiver. It is a point to point communication between sender and receiver. There are various unicast protocols such as TCP, HTTP, etc.

- TCP is the most commonly used unicast protocol. It is a connection oriented protocol that relay on acknowledgement from the receiver side.
- HTTP stands for Hyper Text Transfer Protocol. It is an object oriented protocol for communication.

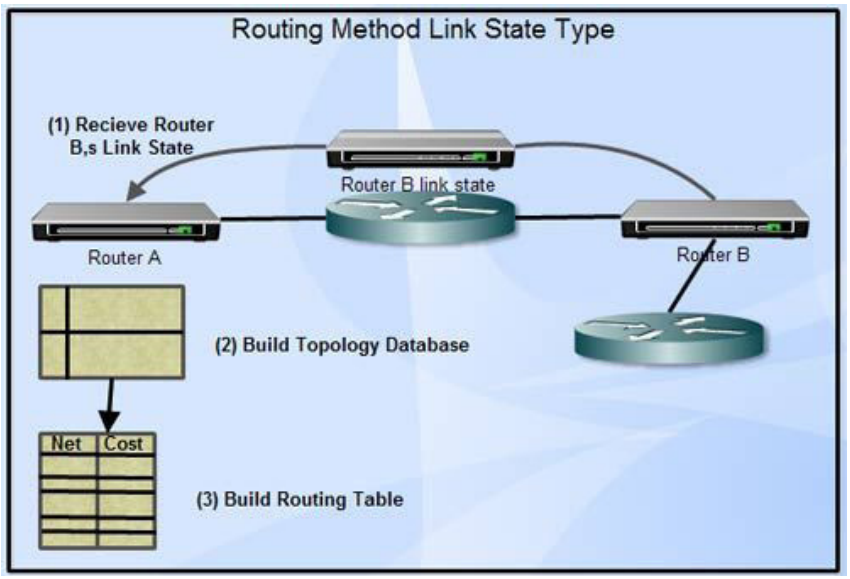
There are three major protocols for unicast routing:

- Distance Vector Routing
- Link State Routing
- Path-Vector Routing

Link State Routing –

Link state routing is the second family of routing protocols. While distance vector routers use a distributed algorithm to compute their routing tables, link-state routing uses link-state routers to exchange messages that allow each router to learn the entire network topology. Based on this learned topology, each router is then able to compute its routing table by using a shortest path computation.

In link-state protocols, there are no restrictions. in number of hope as in distance-vector protocols, and these are aimed at relatively large networks such as Internet backbones. The load on routers will be large however, since processing is complex.



Features of link state routing protocols –

- Link state packet – A small packet that contains routing information.
- Link state database – A collection information gathered from link state packet.
- Shortest path first algorithm (Dijkstra algorithm) – A calculation performed on the database results into shortest path
- Routing table – A list of known paths and interfaces.

Calculation of shortest path –

To find shortest path, each node needs to run the famous Dijkstra algorithm. This famous algorithm uses the following steps:

Step-1: The node is taken and chosen as a root node of the tree, this creates the tree with a single node, and now set the total cost of each node to some value based on the information in Link State Database

Step-2: Now the node selects one node, among all the nodes not in the tree like structure, which is nearest to the root, and adds this to the tree.The shape of the tree gets changed .

Step-3: After this node is added to the tree, the cost of all the nodes not in the tree needs to be updated because the paths may have been changed.

Step-4: The node repeats the Step 2. and Step 3. until all the nodes are added in the tree.

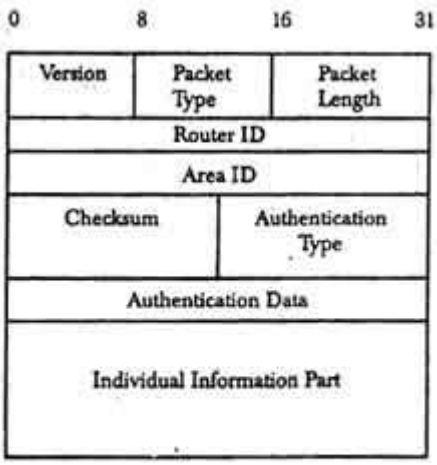
Link State protocols in comparison to Distance Vector protocols have:

- It requires large amount of memory.
- Shortest path computations require many CPU circles.
- If network use the little bandwidth; it quickly reacts to topology changes
- All items in the database must be sent to neighbors to form link state packets.
- All neighbors must be trusted in the topology.
- Authentication mechanisms can be used to avoid undesired adjacency and problems.
- No split horizon techniques are possible in the link state routing.

OPEN SHORTEST PATH FIRST (OSPF) ROUTING PROTOCOL

- Open Shortest Path First (OSPF) is a unicast routing protocol developed by working group of the Internet Engineering Task Force (IETF).
  - It is a intradomain routing protocol.
  - It is an open source protocol.
  - It is similar to Routing Information Protocol (RIP)
- OSPF is a classless routing protocol, which means that in its updates, it includes the subnet of each route it knows about, thus, enabling variable-length subnet masks. With variable-length subnet masks, an IP network can be broken into many subnets of various sizes. This provides network administrators with extra network-configuration flexibility. These updates are multicasts at specific addresses (224.0.0.5 and 224.0.0.6).
- OSPF is implemented as a program in the network layer using the services provided by the Internet Protocol
- IP datagram that carries the messages from OSPF sets the value of protocol field to 89
- OSPF is based on the SPF algorithm, which sometimes is referred to as the Dijkstra algorithm

OSPF has two versions – version 1 and version 2. Version 2 is used mostly



Version : OSPF Version

Packet Type : The OSPF packet type

1. Hello
2. Database description
3. Link state request
4. Link state update
5. Link state acknowledgement

Packet length: The length of the protocol packet in bytes

Router ID: IP address of the router that has sent this packet

Area ID: Area ID of the router that has sent this packet

Authentication type: Authorized or not

Authentication Data: Password etc.

OSPF packet format

OSPF Messages – OSPF is a very complex protocol. It uses five different types of messages. These are as follows:

- Hello message (Type 1) – It is used by the routers to introduce itself to the other routers.
- Database description message (Type 2) – It is normally send in response to the Hello message.
- Link-state request message (Type 3) – It is used by the routers that need information about specific Link-State packet.
- Link-state update message (Type 4) – It is the main OSPF message for building Link-State Database.
- Link-state acknowledgement message (Type 5) – It is used to create reliability in the OSPF protocol.

Network Model for OSPF

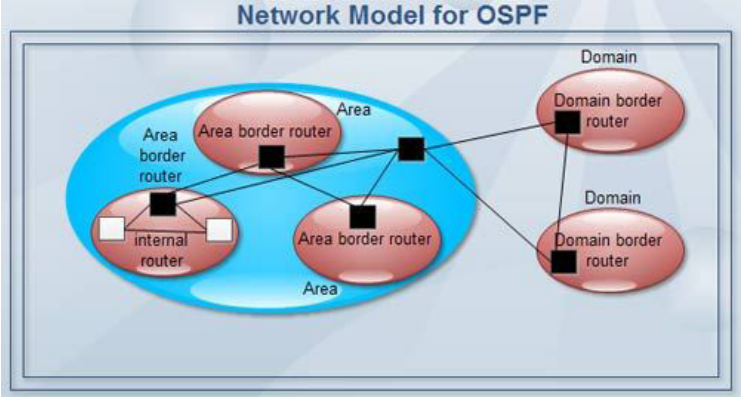
Routers can be classified into 3 types as shown below. One router may play two or more roles. Also, routing information exchanged between these routers is called LSP (Link State Packet).

Domain Border Router This router exchanges route information with routers in other domains. Information thus obtained is included in an OSPF message and transferred to other routers in the same domain (domain to which domain border router belongs). This allows all routers in the same domain to know which domain border router can provide route information to a specific domain.

Internal Router Internal router is a router having its links directly connected to a network within a specific area. That is, internal router does not have any direct links to a network in another area.

Area Border Router This router belongs to two or more areas and notifies the backbone of the outline of its own configuration information so that this outline information can be transferred to other area boundary routers.

- The backbone consists of those networks not contained in any area, their attached routers, and those routers that belong to multiple areas.
- To recapitulate what has been described above, OSPF is an hierarchical routing composed of intra area routing, inter-area routing, inter-domain routing, and so on. This means that if a message needs to be sent from one area to another, this message will sequentially pass.
- Source host --> internal router --> Area border router in the same area --> Domain border router in the same domain -->Destination domain border router --> Destination area border router --> --> --> Destination internal router --> Destination host.



In link-state protocols, there are no restrictions. in number of hope as in distance-vector protocols, and these are aimed at relatively large networks such as Internet backbones. The load on routers will be large however, since processing is complex.

Link State Routing has two phases:

- **Reliable Flooding**

  Initial state: Each node knows the cost of its neighbors.

  Final state: Each node knows the entire graph.

- **Route Calculation**

  Each node uses Dijkstra's algorithm on the graph to calculate the optimal routes to all nodes.

The Link state routing algorithm is also known as Dijkstra's algorithm which is used to find the shortest path from one node to every other node in the network.

The Dijkstra's algorithm is an iterative, and it has the property that after kth iteration of the algorithm, the least cost paths are well known for k destination nodes.

Some Notations:

c(i , j): Link cost from node i to node j. If i and j nodes are not directly linked, then c(i , j) = ∞.

D(v): It defines the cost of the path from source code to destination v that has the least cost currently.

P(v): It defines the previous node (neighbor of v) along with current least cost path from source to v.

N: It is the total number of nodes available in the network.

Algorithm

Initialization

N = {A} // **A is a root node**.

for all nodes v

if v adjacent to A

then D(v) = c(A,v)

else D(v) = infinity

loop

find w not in N such that D(w) is a minimum.

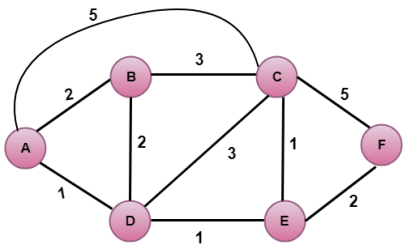
Add w to N

Update D(v) for all v adjacent to w and not in N:

D(v) = min(D(v) , D(w) + c(w,v))

Until all nodes in N

In the above algorithm, an initialization step is followed by the loop. The number of times the loop is executed is equal to the total number of nodes available in the network.



Let's understand through an example:

In the above figure, source vertex is A.

Step 1:

The first step is an initialization step. The currently known least cost path from A to its directly attached neighbors, B, C, D are 2,5,1 respectively. The cost from A to B is set to 2, from A to D is set to 1 and from A to C is set to 5. The cost from A to E and F are set to infinity as they are not directly linked to A.

Step	N	D(B),P(B)	D(C),P(C)	D(D),P(D)	D(E),P(E)	D(F),P(F)
1	A	2,A	5,A	1,A	∞	∞

Step 2:

In the above table, we observe that vertex D contains the least cost path in step 1. Therefore, it is added in N. Now, we need to determine a least-cost path through D vertex.

(a) Calculating shortest path from A to B

v = B, w = D

D(B) = min(D(B) , D(D) + c(D,B))

= min(2, 1+2)>

= min(2, 3)

The minimum value is 2. Therefore, the currently shortest path from A to B is 2.

(b) Calculating shortest path from A to C

v = C, w = D

D(B) = min(D(C) , D(D) + c(D,C))

= min(5, 1+3)

= min(5, 4)

The minimum value is 4. Therefore, the currently shortest path from A to C is 4.

(c) Calculating shortest path from A to E

v = E, w = D

D(B) = min(D(E) , D(D) + c(D,E))

= min(∞, 1+1)

= min(∞, 2)

The minimum value is 2. Therefore, the currently shortest path from A to E is 2.

Step 3:

In the above table, we observe that both E and B have the least cost path in step 2. Let's consider the E vertex. Now, we determine the least cost path of remaining vertices through E.

(a) Calculating the shortest path from A to B.

v = B, w = E

D(B) = min(D(B) , D(E) + c(E,B))

= min(2 , 2+ ∞)

= min(2, ∞)

The minimum value is 2. Therefore, the currently shortest path from A to B is 2.

(b) Calculating the shortest path from A to C.

v = C, w = E

D(B) = min(D(C) , D(E) + c(E,C))

= min(4 , 2+1)

= min(4,3)

The minimum value is 3. Therefore, the currently shortest path from A to C is 3.

(c) Calculating the shortest path from A to F.

v = F, w = E

D(B) = min(D(F) , D(E) + c(E,F))

= min(∞ , 2+2)

= min(∞ ,4)

The minimum value is 4. Therefore, the currently shortest path from A to F is 4.

Step	N	D(B),P(B)	D(C),P(C)	D(D),P(D)	D(E),P(E)	D(F),P(F)
1	A	2,A	5,A	1,A	∞	∞
2	AD	2,A	4,D		2,D	∞
3	ADE	2,A	3,E			4,E

Step 4:

In the above table, we observe that B vertex has the least cost path in step 3. Therefore, it is added in N. Now, we determine the least cost path of remaining vertices through B.

(a) Calculating the shortest path from A to C.

v = C, w = B

D(B) = min(D(C) , D(B) + c(B,C))

= min(3 , 2+3)

= min(3,5)

The minimum value is 3. Therefore, the currently shortest path from A to C is 3.

(b) Calculating the shortest path from A to F.

v = F, w = B

D(B) = min(D(F) , D(B) + c(B,F))

= min(4, ∞)

= min(4, ∞)

The minimum value is 4. Therefore, the currently shortest path from A to F is 4.

Step	N	D(B),P(B)	D(C),P(C)	D(D),P(D)	D(E),P(E)	D(F),P(F)
1	A	2,A	5,A	1,A	∞	∞
2	AD	2,A	4,D		2,D	∞
3	ADE	2,A	3,E			4,E
4	ADEB		3,E			4,E

Step 5:

In the above table, we observe that C vertex has the least cost path in step 4. Therefore, it is added in N. Now, we determine the least cost path of remaining vertices through C.

Calculating the shortest path from A to F.

v = F, w = C

D(B) = min(D(F) , D(C) + c(C,F))

= min(4, 3+5)

= min(4,8)

The minimum value is 4. Therefore, the currently shortest path from A to F is 4.

Step	**N**	**D(B),P(B)**	**D(C),P(C)**	**D(D),P(D)**	**D(E),P(E)**	**D(F),P(F)**
1	A	2,A	5,A	1,A	∞	∞
2	AD	2,A	4,D		2,D	∞
3	ADE	2,A	3,E			4,E
4	ADEB		3,E			4,E
5	ADEBC					4,E

Final table:

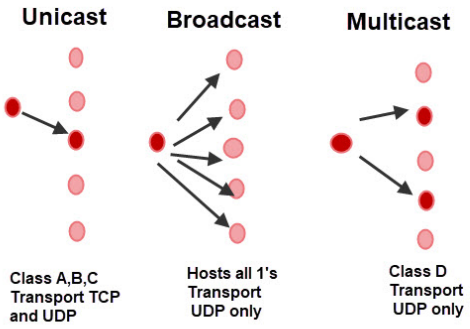
Step	**N**	**D(B),P(B)**	**D(C),P(C)**	**D(D),P(D)**	**D(E),P(E)**	**D(F),P(F)**
1	A	2,A	5,A	1,A	∞	∞
2	AD	2,A	4,D		2,D	∞
3	ADE	2,A	3,E			4,E
4	ADEB		3,E			4,E
5	ADEBC					4,E
6	ADEBCF					

Disadvantage:

Heavy traffic is created in Line state routing due to Flooding. Flooding can cause an infinite looping, this problem can be solved by using Time-to-leave field

MULTICASTING:

- IP multicasting allows a host to send a single packet to thousands of hosts across a routed network i.e. The Internet.



Unicast,Broadcast and Multicast IP Addressing

- It is used mainly for audio (radio) and video distribution.
- In Networking a packet can be sent to:
  - A single host –Unicast = (TCP and UDP)
  - All hosts -Broadcast – (UDP only)
  - A group of hosts – Multicast -(UDP only)

Broadcasts vs Multicasts

- Multicasting is different from IP broadcasting as:
- Broadcasting uses a single IP address. Host bits set to all 1's. There are a range of multicast addresses
- Broadcast messages are not sent through routers but multicast messages are.
- All hosts will receive broadcasts by default
- A host must be configured to receive multicast messages.

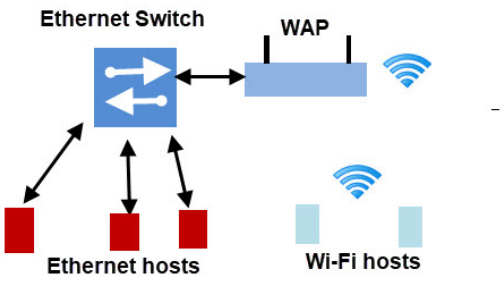
Multicast Addresses

- IPv4 Multicast addresses use the reserved class D address range:
- 224.0.0.0 through 239.255.255.255
- The addresses range between 224.0.0.0 and 224.0.0.255 is reserved for use by routing and maintenance protocols inside a network.
- These addresses aren't forwarded by routers. Many of the multicast addresses are reserved see Muticast Space Registry

Multicasting Works

On a small home or office network any host can send and receive multicast datagra1ms.

Multicasting on a Home Network



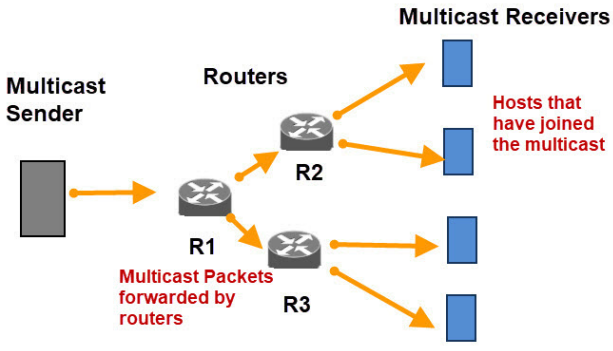
A multicast datagram sent from any host can be received by any host on the network configured to receive messages on that multicast address. A host can send and receive on multiple multicast addresses.

Multicast Groups

- A host that is configured to receive datagrams sent to a multicast address becomes part of a multicast group for that address.
- A group can have 1 to an unlimited number of hosts. Neither hosts or routers maintain a list of individual group members.
- A host can be part of multiple multicast groups and can send to multiple multicast addresses.
- A host can send datagrams to a multicast group address even though there are no members of that group, and a host doesn't need to be a member of a group to send multicast datagrams to that group.

Multicast On the Internet

On the Internet multicast packets need to be forwarded by routers.



Multicasting on the Internet

- A router will determine if any of the hosts on a locally attached network are configured to receive multicast datagrams using IGMP(Internet Group Management Protocol).

- Routers will listen for IGMP messages and periodically send queries on the local subnet. using the multicast group address 224.0.0.1 (Reserved All hosts address).
- Multicast routers do not keep track of which hosts are part of a group, but only need to know if any hosts on that subnet are part of a group.
- If a router receives a multicast datagram from another network and has no members for that group address on any of it's subnets it drops the packet.

Time To Live (TTL)

Each IP Multicast packet uses the time-to-live (TTL) field of the IP header as a scope-limiting parameter. The TTL field controls the number of hops that a IP Multicast packet is allowed to propagate. Each time a router forwards a packet, its TTL is decremented.

A multicast packets whose TTL has expired (is 0) is dropped, without an error notification to the sender. This mechanism prevents messages from needless transmission to regions of the worldwide Internet that lie beyond the subnets containing the multicast group members.

A local network multicast reaches all immediately-neighboring members of the destination host group (the IP TTL is 1 by default). If a multicast datagram has a TTL greater than 1, the multicast router(s) attached to the local network take responsibility for internetwork forwarding. The datagram is forwarded to other networks that have members of the destination group. On those other member networks that are reachable within the IP time-to-live, an attached multicast router completes delivery by transmitting the datagram as a local multicast.

TTL thresholds in multicast routers prevent datagrams with less than a certain TTL from traversing certain subnets. This can provide a convenient mechanism for confining multicast traffic to within campus or enterprise networks. Several standard settings for TTL are specified for the MBONE: 1 for local net, 15 for site, 63 for region and 127 for world.

Internet Group Management Protocol (IGMP)

Multicast packets from remote sources must be relayed by routers, which should only forward them on to the local network if there is a recipient for the multicast host group on the LAN.

The Internet Group Management Protocol (IGMP) is used by multicast routers to learn the existence of host group members on their directly attached subnets. It does so by sending IGMP queries and having IP hosts report their host group memberships. The basic version of IGMP dates from 1988 and is now a full Internet standard. It is described in RFC 1112.

ion (bits 0-3)	**Type (bits 4-7)**	**Code (bits 8-15)**	**Checksum (bits 16-31)**
Multicast Group Address (Class D)			

IGMP is loosely analogous to ICMP and is implemented over IP. IGMP messages are encapsulated in IP datagrams. IGMP has only two kinds of packets: Host Membership Query and

Host Membership Report, with the same simple fixed format containing some control information in the first word of the payload field and a class D address in the second word

When a process asks its host to join a new multicast host group, the driver creates a hardware multicast address, and an IGMP Host Membership Report with the group address is immediately sent. The host's network interface is expected to map the IP host group addresses to local network addresses as required to update its multicast reception filter. Each host keeps track of its host group memberships, and when the last process on a host leaves a group, that group is no longer reported by the host.

Periodically the local multicast router sends an IGMP Host Membership Query to the "all-hosts" group, to verify current memberships. If all member hosts reported memberships at the same time frequent traffic congestion might result.

This is avoided by having each host delay their report by a random interval if it has not seen a report for the same group from another host. As a result, only one membership report is sent in response for each active group address, although many hosts may have memberships.

IGMP updates are used by multicast routing protocols to communicate host group memberships to neighboring routers, propagating group information through the internetwork. IGMP is used to identify a designated router in the LAN for this purpose. The bandwidth needed to transmit host group information is usually slight compared to the multicast application traffic, so this propagation method is workable. More sophisticated methods enable routers to determine dynamically how to best forward the multicast application traffic, as discussed in the next section.

Advantages of IP Multicast

- Many emerging Internet applications are one-to-many or many-to-many, where one or multiple sources are sending to multiple receivers. Examples are the transmission of corporate messages to employees, communication of stock quotes to brokers, video and audio conferencing for remote meetings and telecommuting, and replicating databases and web site information.
- IP Multicast efficiently supports this type of transmission by enabling sources to send a single copy of a message to multiple recipients who explicitly want to receive the information. This is far more efficient than requiring the source to send an individual copy of a message to each requester (referred to as point-to-point unicast), in which case the number of receivers is limited by the bandwidth available to the sender.
- It is also more efficient than broadcasting one copy of the message to all nodes (broadcast) on the network, since many nodes may not want the message, and because broadcasts are limited to a single subnet.

  Internet Protocol version 6 (IPv6) is the latest revision of the Internet Protocol (IP) and the first version of the protocol to be widely deployed. IPv6 was developed by the Internet Engineering Task Force (IETF) to deal with the long-anticipated problem of IPv4 address exhaustion.

IPv6 – Features:

Larger Address Space

In contrast to IPv4, IPv6 uses 4 times more bits to address a device on the Internet. This much of extra bits can provide approximately 3.4×1038 different combinations of addresses. This address can accumulate the aggressive requirement of address allotment for almost everything in this world. According to an estimate, 1564 addresses can be allocated to every square meter of this earth.

Simplified Header

IPv6's header has been simplified by moving all unnecessary information and options (which are present in IPv4 header) to the end of the IPv6 header. IPv6 header is only twice as bigger than IPv4 provided the fact that IPv6 address is four times longer.

- **End-to-end Connectivity**

  Every system now has unique IP address and can traverse through the Internet without using NAT or other translating components. After IPv6 is fully implemented, every host can directly reach other hosts on the Internet, with some limitations involved like Firewall, organization policies, etc.

- **Auto-configuration**

  IPv6 supports both stateful and stateless auto configuration mode of its host devices. This way, absence of a DHCP server does not put a halt on inter segment communication.

- **Faster Forwarding/Routing**

  Simplified header puts all unnecessary information at the end of the header. The information contained in the first part of the header is adequate for a Router to take routing decisions, thus making routing decision as quickly as looking at the mandatory header.

- **IPSec**

  Initially it was decided that IPv6 must have IPSec security, making it more secure than IPv4. This feature has now been made optional.

- **No Broadcast**

  Though Ethernet/Token Ring are considered as broadcast network because they support Broadcasting, IPv6 does not have any broadcast support any more. It uses multicast to communicate with multiple hosts.

- **Anycast Support**

  This is another characteristic of IPv6. IPv6 has introduced Anycast mode of packet routing. In this mode, multiple interfaces over the Internet are assigned same Anycast IP address. Routers, while routing, send the packet to the nearest destination.

- **Mobility**

IPv6 was designed keeping mobility in mind. This feature enables hosts (such as mobile phone) to roam around in different geographical area and remain connected with the same IP address. The mobility feature of IPv6 takes advantage of auto IP configuration and Extension headers.

- **Enhanced Priority Support**

  IPv4 used 6 bits DSCP (Differential Service Code Point) and 2 bits ECN (Explicit Congestion Notification) to provide Quality of Service but it could only be used if the end-to-end devices support it, that is, the source and destination device and underlying network must support it.

  In IPv6, Traffic class and Flow label are used to tell the underlying routers how to efficiently process the packet and route it.

- **Smooth Transition**

  Large IP address scheme in IPv6 enables to allocate devices with globally unique IP addresses. This mechanism saves IP addresses and NAT is not required. So devices can send/receive data among each other, for example, VoIP and/or any streaming media can be used much efficiently.

  Other fact is, the header is less loaded, so routers can take forwarding decisions and forward them as quickly as they arrive.

- **Extensibility**

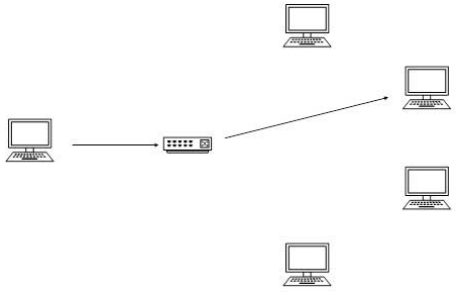
  One of the major advantages of IPv6 header is that it is extensible to add more information in the option part. IPv4 provides only 40-bytes for options, whereas options in IPv6 can be as much as the size of IPv6 packet itself.

IPV6 - Addressing Modes

In computer networking, addressing mode refers to the mechanism of hosting an address on the network. IPv6 offers several types of modes by which a single host can be addressed. More than one host can be addressed at once or the host at the closest distance can be addressed.

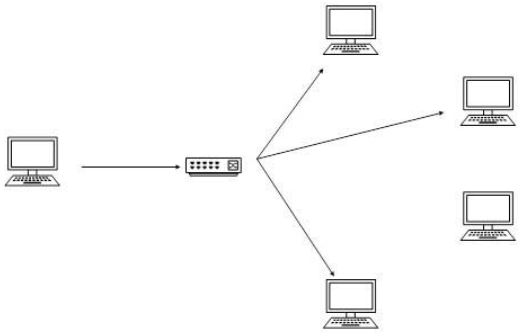
Unicast

In unicast mode of addressing, an IPv6 interface (host) is uniquely identified in a network segment. The IPv6 packet contains both source and destination IP addresses. A host interface is equipped with an IP address which is unique in that network segment.When a network switch or a router receives a unicast IP packet, destined to a single host, it sends out one of its outgoing interface which connects to that particular host.



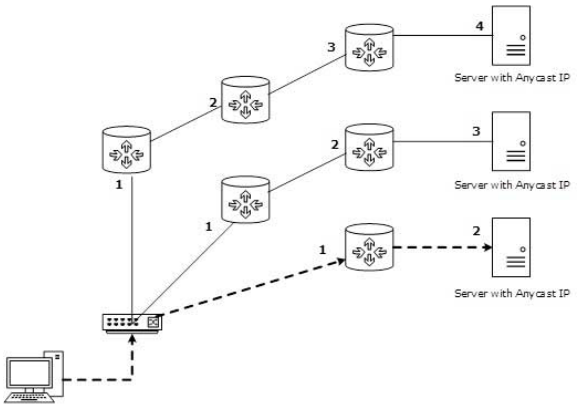
Multicast

The IPv6 multicast mode is same as that of IPv4. The packet destined to multiple hosts is sent on a special multicast address. All the hosts interested in that multicast information, need to join that multicast group first. All the interfaces that joined the group receive the multicast packet and process it, while other hosts not interested in multicast packets ignore the multicast information.



Anycast

IPv6 has introduced a new type of addressing, which is called Anycast addressing. In this addressing mode, multiple interfaces (hosts) are assigned same Anycast IP address. When a host wishes to communicate with a host equipped with an Anycast IP address, it sends a Unicast message. With the help of complex routing mechanism, that Unicast message is delivered to the host closest to the Sender in terms of Routing cost.



Address Structure

An IPv6 address is made of 128 bits divided into eight 16-bits blocks. Each block is then converted into 4-digit Hexadecimal numbers separated by colon symbols.

For example, given below is a 128 bit IPv6 address represented in binary format and divided into eight 16-bits blocks:

0010000000000001 0000000000000000 0011001000111000 1101111111100001

0000000001100011 0000000000000000 0000000000000000 1111111011111011

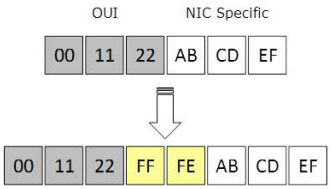
Each block is then converted into Hexadecimal and separated by ':' symbol:

2001:0000:3238:DFE1:0063:0000:0000:FEFB

Interface ID

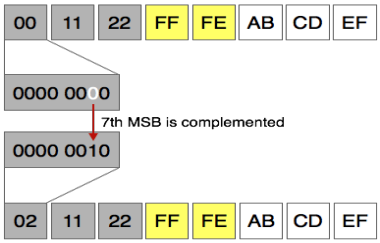
IPv6 has three different types of Unicast Address scheme. The second half of the address (last 64 bits) is always used for Interface ID. The MAC address of a system is composed of 48-bits and represented in Hexadecimal. MAC addresses are considered to be uniquely assigned worldwide. Interface ID takes advantage of this uniqueness of MAC addresses.

A host can auto-configure its Interface ID by using IEEE's Extended Unique Identifier (EUI-64) format. First, a host divides its own MAC address into two 24-bits halves. Then 16-bit Hex value 0xFFFE is sandwiched into those two halves of MAC address, resulting in EUI-64 Interface ID.



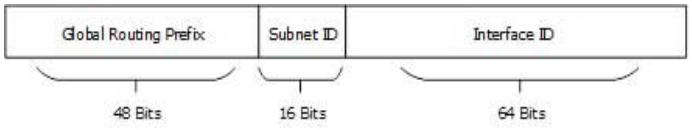
Conversion of EUI-64 ID into IPv6 Interface Identifier

To convert EUI-64 ID into IPv6 Interface Identifier, the most significant 7th bit of EUI-64 ID is complemented. For example:



Global Unicast Address

This address type is equivalent to IPv4's public address. Global Unicast addresses in IPv6 are globally identifiable and uniquely addressable.



Global Routing Prefix: The most significant 48-bits are designated as Global Routing Prefix which is assigned to specific autonomous system. The three most significant bits of Global Routing Prefix is always set to 001.

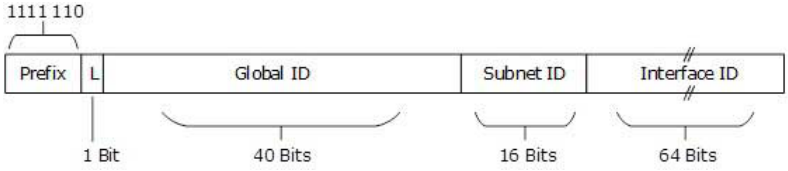
Link-Local Address :

Auto-configured IPv6 address is known as Link-Local address. This address always starts with FE80. The first 16 bits of link-local address is always set to 1111 1110 1000 0000 (FE80). The next 48-bits are set to 0, thus:

Link-local addresses are used for communication among IPv6 hosts on a link (broadcast segment) only. These addresses are not routable, so a Router never forwards these addresses outside the link.

Unique-Local Address :

This type of IPv6 address is globally unique, but it should be used in local communication. The second half of this address contain Interface ID and the first half is divided among Prefix, Local Bit, Global ID and Subnet ID.



Prefix is always set to 1111 110. L bit, is set to 1 if the address is locally assigned. So far, the meaning of L bit to 0 is not defined. Therefore, Unique Local IPv6 address always starts with 'FD'.

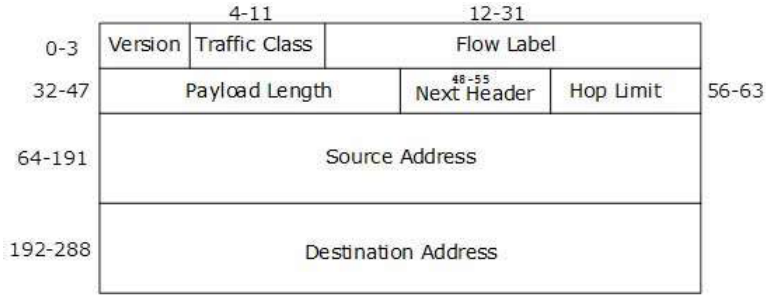
The scope of Link-local address is limited to the segment. Unique Local Address are locally global, but are not routed over the Internet, limiting their scope to an organization's boundary. Global Unicast addresses are globally unique and recognizable. They shall make the essence of Internet v2 addressing.

IPV6 HEADERS

- An IPv6 address is 4 times larger than IPv4, but surprisingly, the header of an IPv6 address is only 2 times larger than that of IPv4.

- IPv6 headers have one Fixed Header and zero or more Optional (Extension) Headers.
- All the necessary information that is essential for a router is kept in the Fixed Header.
- The Extension Header contains optional information that helps routers to understand how to handle a packet/flow.

Fixed Header



IPv6 fixed header is 40 bytes long and contains the following information.

S.N.	Field & Description
1	**Version** (4-bits): It represents the version of Internet Protocol, i.e. 0110.
2	**Traffic Class** (8-bits): These 8 bits are divided into two parts. The most significant 6 bits are used for Type of Service to let the Router Known what services should be provided to this packet. The least significant 2 bits are used for Explicit Congestion Notification (ECN).
3	**Flow Label** (20-bits): This label is used to maintain the sequential flow of the packets belonging to a communication. The source labels the sequence to help the router identify that a particular packet belongs to a specific flow of information. This field helps avoid re-ordering of data packets. It is designed for streaming/real-time media.
4	**Payload Length** (16-bits): This field is used to tell the routers how much information a particular packet contains in its payload. Payload is composed of Extension Headers and Upper Layer data. With 16 bits, up to 65535 bytes can be indicated; but if the Extension Headers contain Hop-by-Hop Extension Header, then the payload may exceed 65535 bytes and this field is set to 0.
5	**Next Header** (8-bits): This field is used to indicate either the type of Extension Header, or if the Extension Header is not present then it indicates the Upper Layer PDU. The values for the type of Upper Layer PDU are same as IPv4's.
6	**Hop Limit** (8-bits): This field is used to stop packet to loop in the network infinitely. This is same as TTL in IPv4. The value of Hop Limit field is decremented by 1 as it passes a link (router/hop). When the field reaches 0 the packet is discarded.
7	**Source Address** (128-bits): This field indicates the address of originator of the packet.
8	**Destination Address** (128-bits): This field provides the address of intended recipient of the packet.

In IPv4, a host that wants to communicate with another host on the network needs to have an IP address acquired either by means of DHCP or by manual configuration. As soon as a host is equipped with some valid IP address, it can speak to any host on the subnet.

To communicate on layer-3, a host must also know the IP address of the other host. Communication on a link, is established by means of hardware embedded MAC Addresses. To know the MAC address of a host whose IP address is known, a host sends ARP broadcast and in return, the intended host sends back its MAC address.

In IPv6, there are no broadcast mechanisms. It is not a must for an IPv6 enabled host to obtain an IP address from DHCP or manually configured, but it can auto-configure its own IP.

ARP has been replaced by ICMPv6 Neighbor Discovery Protocol.

- **Neighbor Discovery Protocol**

  A host in IPv6 network is capable of auto-configuring itself with a unique link-local address. As soon as host gets an IPv6 address, it joins a number of multicast groups. All communications related to that segment take place on those multicast addresses only. A host goes through a series of states in IPv6:

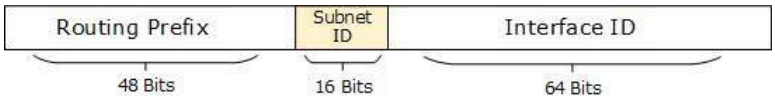
- **Neighbor Solicitation**: After configuring all IPv6's either manually, or by DHCP Server or by auto-configuration, the host sends a Neighbor Solicitation message out to FF02::1/16 multicast address for all its IPv6 addresses in order to know that no one else occupies the same addresses.
- **DAD (Duplicate Address Detection):** When the host does not listen from anything from the segment regarding its Neighbor Solicitation message, it assumes that no duplicate address exists on the segment.
- **Neighbor Advertisement**: After assigning the addresses to its interfaces and making them up and running, the host once again sends out a Neighbor Advertisement message telling all other hosts on the segment, that it has assigned those IPv6 addresses to its interfaces.

Once a host is done with the configuration of its IPv6 addresses, it does the following things:

- **Router Solicitation**: A host sends a Router Solicitation multicast packet (FF02::2/16) out on its segment to know the presence of any router on this segment. It helps the host to configure the router as its default gateway. If its default gateway router goes down, the host can shift to a new router and makes it the default gateway.
- **Router Advertisement**: When a router receives a Router Solicitation message, it response back to the host, advertising its presence on that link.
- **Redirect:** This may be the situation where a Router receives a Router Solicitation request but it knows that it is not the best gateway for the host. In this situation, the router sends back a Redirect message telling the host that there is a better 'next-hop' router available. Next-hop is where the host will send its data destined to a host which does not belong to the same segment.

IPV6 Subnetting:

IPv6 addresses use 128 bits to represent an address which includes bits to be used for subnetting. The second half of the address (least significant 64 bits) is always used for hosts only. Therefore, there is no compromise if we subnet the network.



- 16 bits of subnet is equivalent to IPv4's Class B Network. Using these subnet bits, an organization can have another 65 thousand of subnets which is by far, more than enough.
- Thus, routing prefix is /64 and host portion is 64 bits. We can further subnet the network beyond 16 bits of Subnet ID, by borrowing host bits; but it is recommended that 64 bits should always be used for hosts addresses because auto-configuration requires 64 bits.
- IPv6 subnetting works on the same concept as Variable Length Subnet Masking in IPv4.
- 48 prefixes can be allocated to an organization providing it the benefit of having up to /64 subnet prefixes, which is 65535 sub-networks, each having 264 hosts. A /64 prefix can be assigned to a point-to-point connection where there are only two hosts (or IPv6 enabled devices) on a link.

Routing

Routing concepts remain same in case of IPv6 but almost all routing protocol have been redefined accordingly. We have seen in Communication in IPv6 segment, how a host speaks to its gateway.

Routing is a process to forward routable data choosing best route among several available routes or path to the destination. A router is a device which forwards data which is not explicitly destined to it.

There exist two forms of routing protocols

- **Distance Vector Routing Protocol:** A router running distance vector protocol advertises its connected routes and learns new routes from its neighbors. The routing cost to reach a destination is calculated by means of hops between the source and destination. A Router generally relies on its neighbor for best path selection, also known as "routing-by-rumors". RIP and BGP are Distance Vector Protocols.
- **Link-State Routing Protocol**: This protocol acknowledges the state of a Link and advertises to its neighbors. Information about new links is learnt from peer routers. After all the routing information has been converged, Link-State Routing Protocol uses its own algorithm to calculate best path to all available links. OSPF and IS-IS are link
- Routing protocols and both uses Dijkstra's Shortest Path First algorithm.

Routing protocols can be divided in two categories:

- **Interior Routing Protocol**: Protocols in this categories are used within an Autonomous System or organization to distribute routes among all routers inside its boundary. Examples: RIP, OSPF.

- **Exterior Routing Protocol**: Whereas an Exterior Routing Protocol distributes routing information between two different Autonomous Systems or organization. ***Examples:*** BGP.

Routing protocols

- **RIPng**

  RIPng stands for Routing Information Protocol Next Generation. This is an Interior Routing Protocol and is a Distance Vector Protocol. RIPng has been upgraded to support IPv6.

- **OSPFv3**

  Open Shortest Path First version 3 is an Interior Routing Protocol which is modified to support IPv6. This is a Link-State Protocol and uses Djikrasta's Shortest Path First algorithm to calculate best path to all destinations.

- **BGPv4**

  BGP stands for Border Gateway Protocol. It is the only open standard Exterior Gateway Protocol available. BGP is a Distance Vector protocol which takes Autonomous System as calculation metric, instead of number of routers as Hop. BGPv4 is an upgrade of BGP to support IPv6 routing.

Benefits of IPv6

While increasing the pool of addresses is one of the most often-talked about benefit of IPv6, there are other important technological changes in IPv6 that will improve the IP protocol:

- No more NAT (Network Address Translation)
- Auto-configuration
- No more private address collisions
- Better multicast routing
- Simpler header format
- Simplified, more efficient routing
- True quality of service (QoS), also called "flow labeling"
- Built-in authentication and privacy support
- Flexible options and extensions
- Easier administration (say good-bye to DHCP)

Difference between IPv4 vs. IPv6

Basis for differences	IPv4	IPv6
Size of IP address	IPv4 is a 32-Bit IP Address.	IPv6 is 128 Bit IP Address.
Addressing method	IPv4 is a numeric address, and its binary bits are separated by a dot (.)	IPv6 is an alphanumeric address whose binary bits are separated by a colon (:). It also contains hexadecimal.

Basis for differences	IPv4	IPv6
Number of header fields	12	8
Length of header filed	20	40
Checksum	Has checksum fields	Does not have checksum fields
Example	12.244.233.165	2001:0db8:0000:0000:0000:ff00:0042:7879
Type of Addresses	Unicast, broadcast, and multicast.	Unicast, multicast, and anycast.
Number of classes	IPv4 offers five different classes of IP Address. Class A to E.	lPv6 allows storing an unlimited number of IP Address.
Configuration	You have to configure a newly installed system before it can communicate with other systems.	In IPv6, the configuration is optional, depending upon on functions needed.
VLSM support	IPv4 support VLSM (Virtual Length Subnet Mask).	IPv6 does not offer support for VLSM.
Fragmentation	Fragmentation is done by sending and forwarding routes.	Fragmentation is done by the sender.
Routing Information Protocol (RIP)	RIP is a routing protocol supported by the routed daemon.	RIP does not support IPv6. It uses static routes.
Network Configuration	Networks need to be configured either manually or with DHCP. IPv4 had several overlays to handle Internet growth, which require more maintenance efforts.	IPv6 support autoconfiguration capabilities.
Best feature	Widespread use of NAT (Network address translation) devices which allows single NAT address can mask thousands of non-routable addresses, making end-to-end integrity achievable.	It allows direct addressing because of vast address Space.
Address Mask	Use for the designated network from host portion.	Not used.
SNMP	SNMP is a protocol used for system management.	SNMP does not support IPv6.
Mobility & Interoperability	Relatively constrained network topologies to which move restrict mobility and interoperability capabilities.	IPv6 provides interoperability and mobility capabilities which are embedded in network devices.

Basis for differences	IPv4	IPv6
Security	Security is dependent on applications - IPv4 was not designed with security in mind.	IPSec(Internet Protocol Security) is built into the IPv6 protocol, usable with a proper key infrastructure.
Packet size	Packet size 576 bytes required, fragmentation optional	1208 bytes required without fragmentation
Packet fragmentation	Allows from routers and sending host	Sending hosts only
Packet header	Does not identify packet flow for QoS handling which includes checksum options.	Packet head contains Flow Label field that specifies packet flow for QoS handling
DNS records	Address (A) records, maps hostnames	Address (AAAA) records, maps hostnames
Address configuration	Manual or via DHCP	Stateless address autoconfiguration using Internet Control Message Protocol version 6 (ICMPv6) or DHCPv6
IP to MAC resolution	Broadcast ARP	Multicast Neighbour Solicitation
Local subnet Group management	Internet Group Management Protocol GMP)	Multicast Listener Discovery (MLD)
Optional Fields	Has Optional Fields	Does not have optional fields. But Extension headers are available.
IPSec	Internet Protocol Security (IPSec) concerning network security is optional	Internet Protocol Security (IPSec) Concerning network security is mandatory
Dynamic host configuration Server	Clients have approach DHCS (Dynamic Host Configuration server) whenever they want to connect to a network.	A Client does not have to approach any such server as they are given permanent addresses.
Mapping	Uses ARP(Address Resolution Protocol) to map to MAC address	Uses NDP(Neighbour Discovery Protocol) to map to MAC address
Combability with mobile devices	IPv4 address uses the dot-decimal notation. That's why it is not suitable for mobile networks.	IPv6 address is represented in hexadecimal, colon- separated notation. IPv6 is better suited to mobile networks.