Introduction, Network service model, Datagrams and Virtual circuit service, routing principles: A link state routing algorithm, The Distance vector routing algorithm, Hierarchical routing, The Internet protocol (IP): IPV4 addressing, datagram format, IP datagram fragmentation, Internet Control Message Protocol [ICMP], Network address translator, Routing in the Internet, IPV6, Multicasting routing

Network Layer Functions

Devices of Network Layer mainly focus on routing. Routing may include various tasks aimed to achieve a single goal. These can be:

- Addressing devices and networks.
- Populating routing tables or static routes.
- Queuing incoming and outgoing data and then forwarding them according to quality of service constraints set for those packets.
- Internetworking between two different subnets.
- Delivering packets to destination with best efforts.
- Provides connection oriented and connection less mechanism.

Network Layer Features

With its standard functionalities, Layer 3 can provide various features as:

- Quality of service management
- Load balancing and link management
- Security
- Interrelation of different protocols and subnets with different schema.
- Different logical network design over the physical network design.
- L3 VPN and tunnels can be used to provide end to end dedicated connectivity.
- Helps to communicate end to end devices over the internet. It comes in two flavors.

Network layer protocols

The following protocol operate at the TCP/IP internet layer

- **Internet protocol (IP):** IP provides connectionless, best- effort delivery routing of packet. IP is not concerned with the contents of the packets but looks for a path to the destination. **IPv4 or IPv6**
- **Internet control message protocol (ICMP):** ICMP Provides control and messaging capabilities.
- **Address Resolution Protocol (ARP):** ARP determines the data link layer address or MAC address, for known IP address.
- **Reverse ARP (RARP):** RARP determines the IP address for known MAC address.

Network service model

It means the characteristics of end to end transport of packets between sending and receiving end system. In the sending host, when the transport layer passes a packet to the network layer, specific services that could be provided by the network layer include:

- Guaranteed delivery
- Guaranteed delivery with bounded delay.

Furthermore, the following service could be provided to a flow of packets between a given source and destination:

- In order packet delivery
- Guaranteed minimal bandwidth
- Guaranteed maximum jitter
- Security service.

Virtual Circuit and Datagram Networks

The internet transport layer provides each application a choice between two services UDP (a connectionless service) or TCP (a connection oriented service). In similar manner, a network layer can also provide connectionless service (datagram networks) or connection service (virtual circuit network)

Although these transport layer and network layer service models seem parallel, there are some crucial differences:

i. In transport layer, it is process to process service. But, in network layer, it host to host service.
ii. In all computer network architectures up to now (internet, ATM, frame relay, and soon), the network layer provides either a host to host connection service or host to host connectionless service but not both.

iii. Connection oriented service in transport layer in implemented at the edge of the network in the end systems; however, the network layer connection service is implemented in the network core as well as the end system.

Virtual Circuit (VC) Network

Many network architectures (not internet) including those of ATM and frame relay are VC network and therefore, use connections at the network layer. These network layer connections are called virtual circuits (VCs). Let's now consider how a VC service can be implemented in a computer network.

A VC consists of

1) A path (i.e. a series of links and routers) between the source and destination hosts.
2) VC numbers, one number for each link along the path.
3) Entries in the forwarding table in each router along the path.

A packet belonging to a virtual circuit will carry a VC number in its header. Because a virtual circuit may have a different VC number on each link, each intervening router must replace the VC number of each traversing packet with a new VC number. The new VC number is obtained from the forwarding table.

There are three identifiable phases in a virtual circuit

i) VC setup
ii) Data transfer
iii) VC teardown

Datagram network

Internet is a datagram network in which each time an end system wants to send a packet, it stamps that packet with the address of the destination end system and then pops packet into the network. Routers in a datagram network don't maintain any state information about VCs.

As a packet is transmitted from source to destination, it passes through a series of routers. Each of these routers uses the packets destination address to forward the packet. Specifically, each routers has a forwarding table that maps destination addresses to link interfaces, when a packet arrives at the router, the router uses the packets destination address to look up the appropriate output link interface in the forwarding table. The router then intentionally forwards the packet to that output link interface.

Routing

Once you create an inter network by connecting your WANs and LANs to a router. You'll need to configure local network addresses, such as IP addresses, to all hosts on the internet work so that they can communicate across that internetwork.

The term routing refers to taking a packet from one device and sending it through the network to another device on a different network. Routers don't really care about hosts. They only care about networks and the best path to each network. The logical network address of the destination host is used to get packets to a network through a routed network, and then hardware address of the host is used to deliver the packet from a router to the correct destination host.

Principles: If your network has ho routers, then it is clear that you are not routing. Routers route traffic to the entire network in your internetwork. To be able to route packets, a router must know, at minimum, the following:

→ Destination address
→ Neighbor routers from which it can learn about remote network.
→ Possible routers to all remote networks.
→ The best route to each remote network.
→ How to maintain and verify routing information.

The router learns about remote network from neighboring routers or from an administrator. The router then builds a routing table (a map of the internet work) that describes how to find the remote network. If the network is directly connected, the router already knows how to get to it.

Static Vs Dynamic Routing

If a network is not directly connected to the router the router must use one of two ways to learn how to get to the remote network: static routing or dynamic routing.
Static routing means someone must hand-type all network locations into the routing table. If static routing is used, the administrator is responsible for updating all changes by hand onto all routers.
In dynamic routing, a protocol acts on all neighboring routers. Then the routers update each other about all the networks they know about and place this information into the routing table. If a change occurs in the network, the dynamic routing protocols automatically inform all routers about the event e.g. RIP V1, RIP v2, OSPF, EIGRP.

Routing algorithm: Distance vector vs. link state
There are three classes of routing protocols:

i) Distance vector

The distance-vector protocols are in use today. Find the best path to a remote network by judging distance. For e.g., in the case of RIP routing, each time a packet goes through a router, that's called a hop. The route with the least number of hops to the network is determined to be the best route. The vector indicates the direction to the remote network. E.g.: RIP, IGRP, they periodically send the entire routing table to directly connected neighbors.

ii) Link State

It is also called shortest-path-first protocols in which the routers each create three separate tables. One to keep track of directly attached neighbors, one determines the topology of the entire internet work, and one is used as the routing table. Link-state routers know more about the internet work than any distance-vector routing protocol. E.g. OSPF (Open Shortest Path First). They send updates containing the state of their own links to all other directly connected routers on the network, which is then propagated to their neighbors.

iii) Hybrid

Hybrid protocols use aspects of both distance vector and link state. E.g.: EIGRP.

Hierarchical Routing: intra-AS routing and inter-AS routing
Autonomous system (AS) is a collection of networks under a common administrative domain, which basically means that all routers sharing the same routing table information are in the same AS
According to AS, there are two types of routing protocols:

i. Intra-AS routing/interior Gateway protocol (IGP) e.g.: RIP, OSPF.
ii. Inter-AS routing/exterior Gateway protocol (EGP) e.g.: Border gateway protocol (BGP)

The internet Protocol (IP)

IP is sometimes referred to as an unreliable protocol. This does not mean that IP will not accurately deliver data across a network. IP is unreliable because it does not perform error checking and correction. That function is handled by upper layer protocols from the transport or application layers.
IP performs the following operations:

→ Defines a packet and an addressing scheme
→ Transfers data between the internet layer and network access layer.
→ Routers packets to remote hosts.
→ The main function of IP is forwarding and addressing in the internet.

IPv4 Addressing

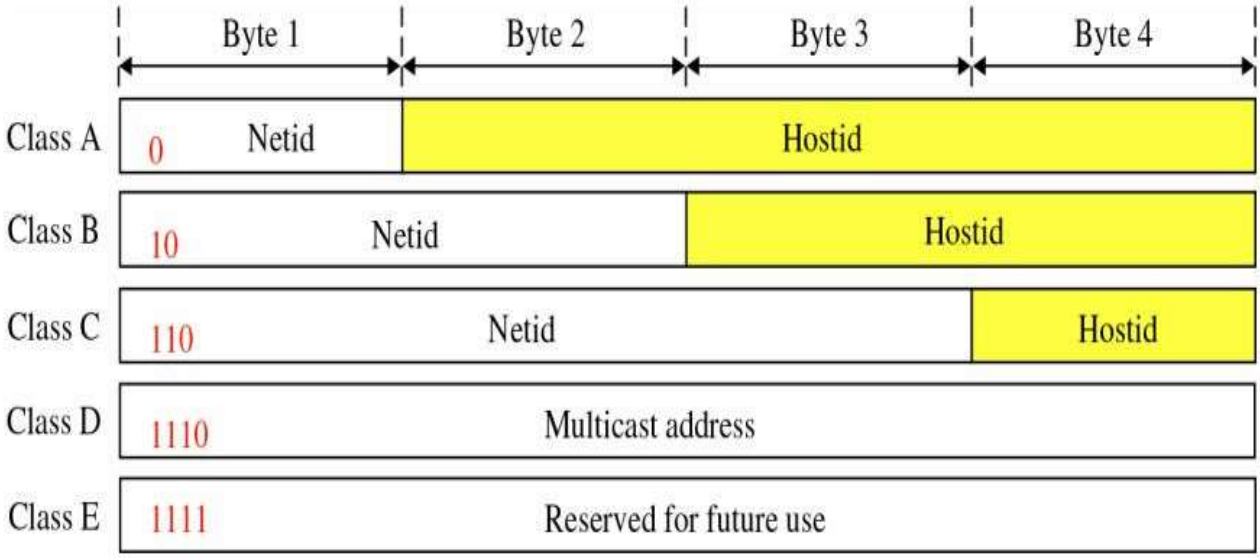
A router's job is to receive a datagram on one link and forward the datagram on some other link, a router necessarily has two or more links to which it is connected. The boundary between the router and any one of its link is called an interface. Because every host and router is capable of sending and receiving IP datagram, IP requires each host and router interface to have its own. IP address thus, an IP address is technically associated with an interface, rather than with the host router containing that interface.
Each IP address is 32 bits long (4 bytes) and thus a total of 2^32 possible IP address. Approximately, there are about 4 billion possible IP addresses. These IP addresses are typically written in so called dotted-decimal notation, in which each byte of the address is written in its decimal from and is separated by a period (dot) from other bytes in the address. For e.g. : consider the IP address 192.168.10.5 the 192 is the decimal equivalent of the first 8 bits of the address, so are the 168, 10 and 5. Thus, the address 192.168.10.5 in binary notation is

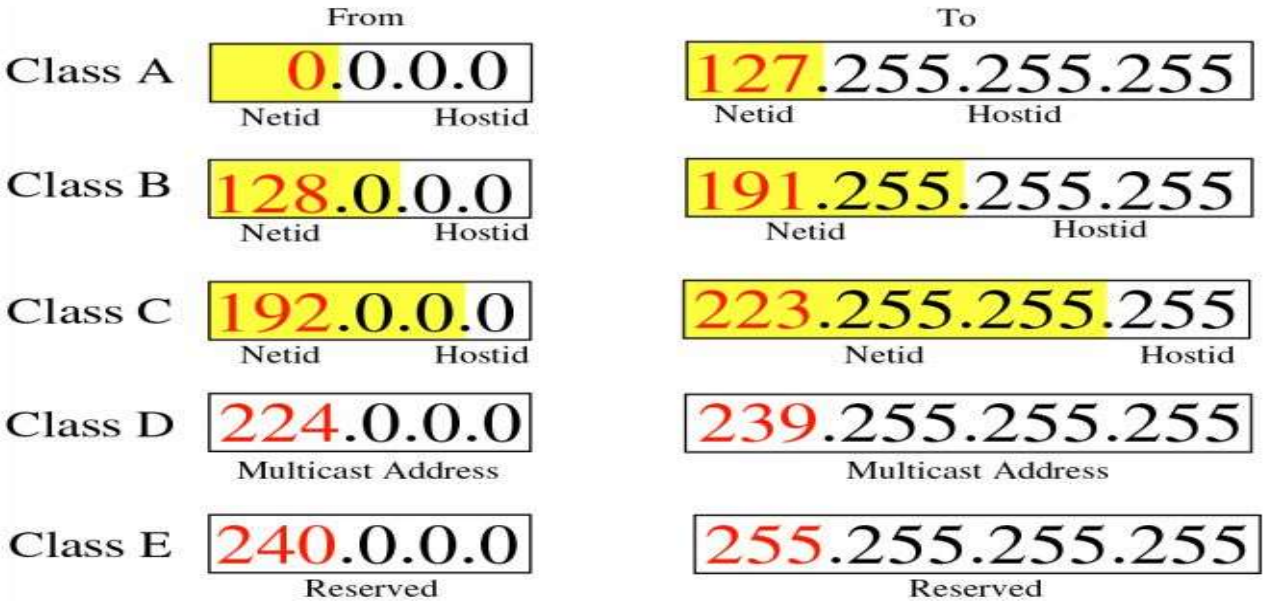
11000000 101010000 00001010 00000101

Each interface an every host and router in the global internet must have an IP address that is globally unique (except for interfaces behind NATs). A portion of an interface's IP address will be determined by the subnet to which it is connected.

Different classes of IPV4 address

An internet address is made of 4 bytes (32 bits) that define a host's connection to a network.
IP address is made up of (netid + hostid)





Class A

- Range: 0 – 127
- So total of 126 (2^{8-1}) Networks are possible and total host = 2^{24} in each Network.
- Default subnet mask is 255.0.0.0

Class B

- Range: 128 – 191
- So total of 2^{16-2} Networks are possible and total host = 2^{16} in each Network.
- Default subnet mask is 255.255.0.0

Class C
- Range: 192 – 223
- So total of 2^{24-3} Networks are possible and total host = 2^8 in each Network.
- Default subnet mask is 255.255.255.0

Class D
- Range: 224 – 239
- Used for Multicasting
- E.g. 224.0.0.1 (group)

Class E
- Range 240-255
- Not used (for future use)

Private Vs Public Address

The people who created the IP addressing scheme also created the IP addressing scheme also created what we call private IP addresses which can be used on a private network, but they are not routable through the Internet. This is designed for the purpose of creating a measure of well-needed security, but it also conveniently saves valuable IP address space.

To accomplish the connection between the ISP and the corporation, the end user, no matter who they are need to use something called Network Address Translation (NAT), which basically takes a private IP address and converts it use on the internet. Many people can use the some real IP address to transmit out onto the internet. Doing things this way saves megatons of address space-good for us all. The reserved private addresses

Class A: 10.0.0.0 through 10.255.255.255

Class B: 172.16.0.0 through 172.16.255.255

Class C: 192.168.0.0 through 192.168.255.255

IP Datagram Format

Different s field used in IP (Version 4) datagram are depicted in fig below:

<table>
<tr><th>Version (4)</th><th>HLEN (4)</th><th>Types of services (8)</th><th colspan="2">Datagram Length (16)</th></tr>
<tr><td colspan="3">Identifier (16)</td><td>Flags (3)</td><td>Fragment Offset (13)</td></tr>
<tr><td colspan="2">TTL (8)</td><td>Protocol (8)</td><td colspan="2">Header Checksum (16)</td></tr>
<tr><td colspan="5">Source IP address (32)</td></tr>
<tr><td colspan="5">Destination IP address (32)</td></tr>
<tr><td colspan="5">Options or Padding not always</td></tr>
<tr><td colspan="5">Data (variable)</td></tr>
</table>

* Number in bracket indicates bits used in that field.

Version: Identifies the version of IP in use. Current version is IPV4.

HLEN: Header length is set to a value to indicate the length of datagram header. Most IP datagram doesn't contain options, so HLEN mostly indicates where the data begins in datagram. Typical IP datagram has 20 bytes header.

Types of services: Identifies different types of services included in IP datagram such as delay, throughput, precedence etc. IP datagram can be real-time or non-real-time as per type of services

Datagram Length: Indicates total length (Data + Header) of the IP datagram. Maximum length if IP datagram is 2^{16}=65535 bytes but in general not more than 1500 bytes.

Identifiers / Flags / Fragment Offset: Identifier (also called Fragment ID) indicates all fragments that belong together. Flags indicate that other fragments to follow. All fragments except last are indicated as 1 and last flag is 0.Fragment offset is used to tell the receiving host how to reassemble the packets.

Time-to-Live (TTL): TTL is used to measure the time a datagram has been in internet. Each Gateway in internet checks this field and discards packet if TTL is 0.

Protocol: this field is used to indicate upper layer protocols (Transport layer) that are to receive the datagram at the destination host. Either TCP or UDP receive the IP datagram at destination.

Header Checksum: Used o detect bit error at the receiving datagram.

Source/Destination address: IP datagram used two 32-bits addresses called source IP address and Destination IP address.

Options: The option field is not used in every datagram. His field is used sometimes for network management and diagnostics.
Data: Data field contains the user data. IP stipulates that the combination of header and Data can't exceed 65535 bytes. Data length varies from protocol to protocol used in network access layer.

IP datagram Fragmentation

Not all network access layer protocols can carry packets of the same size. Some protocols can carry big packets and other protocols can carry small packets. For example, Ethernet packets can carry no more than 1500 bytes of data, where as packets for many wide area network are not more than 576 bytes. The maximum amount of the data that the network access layer (TCP/IP model) protocol can carry is called Maximum Transfer Unit (MTU). Because each IP datagram is encapsulated within the network access layer packet for transport between routers, the MTU of the network access protocol places a hard limit on the size of an IP datagram. The main problems here are that each of the links along the route between sender and receiver can use different network access protocols, and each of these protocols can have different MTUs.

When the size of IP datagram is large than the MTU of Network access layer protocols, this IP datagram need to be fragmented into two or more fragments. These fragments need to be reassembled before they reach to the destination transport layer. Reassembling is done with fragment ID and Fragment Offset. Indeed, both TCP and UDP are expecting to receive complete unfragmented segments from the Internet layer.

The designers of IPV4 felt that the fragmenting, reassembling and possibly again fragmenting and reassembling datagram into the routers would introduce significant complication into the protocol and put a damper on router performance. Fragmentation and reassembly add extra burden at sending routers and receiving hosts. So fragmentation should be minimized as far as possible. This is often done by limiting the TCP /UDP segments to a relatively small size i.e. less than 576 bytes (all network access layer protocols supported by IP are supposed to have MTUs at least 576 bytes. Fragmentation can be entirely eliminated by using an MSS (maximum segment size) of 536 bytes, 20 bytes for TCP header and 20 bytes for IP header.

Fragmentation is supported by only IPV4 not by IPV6.

Features of IP:

- *It is connectionless service:* So without prior call setup, it permits to exchange traffics between two host computers.
- *Datagram could be lost*: As IP is connectionless; it is possible that datagrams could be lost between two end user's stations.
- *IP hides underlying sub network from the end user:* In this context, it creates a virtual network for the end user. This aspect of IP is quite attractive, because it allows different types of networks to attach to an IP gateway. As a reason IP is reasonably simple to install and, because of its connectionless design, it is quite accommodating.
- *IP is unreliable, best effort and datagram type protocol:* It has no reliability mechanisms. It has no error recovery procedures for the underlying sub networks.
- *IP has no flow control mechanisms:* The user datagram may lost , duplicated or even arrive at out of order. It is not the job of IP to deal with most of these problems. It is not the job of IP to deal with most of these problems, as most of the problems are passed to the next upper layer, TCP.
- *IPV4 supports fragmentation:* Fragmentation refers to an operation where in a protocol data unit (PDU) is divided or segmented into smaller units.

Subnetting

→ A subnetwork, or subnet, is a logically visible subdivision of an IP network. The practice of dividing a network into two or more networks is called subnetting.

→ All computers rhat belong to a subnet are addresses with a common, identical, most significant bit group in their IP address. This results in the logical division of an IP address into two fields,

- A network or routing prefix
- The rest field or host identifier

→ The rest field is an identifier for specific host or network interface.

Address class	Bits for subnet mask	Network prefix
A	11111111 00000000 00000000 00000000	/8
B	11111111 11111111 00000000 00000000	/16
C	11111111 11111111 11111111 00000000	/24

Benefits of subnetting
→ Reduced network traffic
→ Simplified management
→ Smaller broadcast domain

Subnet mask
A subnet mask is a 32-bit number that masks an IP address, and divides an IP address into network address and host address. Subnet mask is made by setting the network bits to all 1's and setting host bit to all 0's. Within a given network, two host addresses are reserved for special purpose. The '0' address is assigned a network address and '255' is assigned to a broadcast address, and they cannot be assigned to hosts.

Network address – Used to identify the network itself .Data that is sent to any host on that network (198.150.11.1-198.150.11.254) will be seen outside of the local area network as 198.159.11.0. The only time that the host numbers matter is when the data is on the local area network.

Broadcast address – Used for broadcasting packets to all the devices on a network. Data that is sent to the broadcast address will be read by all hosts on that network. The Broadcast Address for above IP addresses is 198.150.12.255.

CIDR (Classless Inter Domain Routing)
CIDR was introduced in 1993 replacing the previous generation of IP address syntax – classful networks. CIDR allowed for more efficient use of IPv4 address space and prefix aggregation, known as route summarization or supernetting.
CIDR allows routers to group routes together to reduce the bulk of routing information carried by core routers. With CIDR, IP addresses and their subnet mask are written as four octets, separated by periods, followed by a forward slash (/) and a two digit number that represents the network mask.
e.g. 10.1.1.0/30
172.16.1.16/28
192.168.1.32/27

ICMP (Internet Control Message Protocol)

The internet protocol is connectionless-mode protocol, and as such, it has no error reporting and error-correcting mechanisms. It relies on a module called the Internet control message protocol (ICMP) to;

a. Reports errors on the processing of a datagram
b. Provide for some administrative and status messages.

ICMP sends messages and reports errors to the source host regarding the delivery of a packet. ICMP notifies the host if a destination is unreachable. ICMP is also responsible for managing and creating a time-exceeded message in the event that the lifetime of the datagram expires. ICMP also performs certain editing functions to determine if the IP header is in error or otherwise unintelligible.
The error and status reporting services of ICMP are summarized as below.

Type	Code	description
0	0	echo reply (ping)
3	0	destination network unreachable
3	1	destination host unreachable
3	2	destination protocol unreachable
3	3	destination port unreachable
3	6	destination network unknown
3	7	destination host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery

11 0 TTL expired
12 0 bad IP header

ICMP packet format

0 7	8 15	16 31
8-bit type	b-bit code	16-bit checksum
Data (contents depend on type and code)		

Type: type of message
Code: Subtype of message
Checksum: 1's complement computed over entire ICMP message (except for the checksum field itself, which is set to zero)
Data: depends on type and code

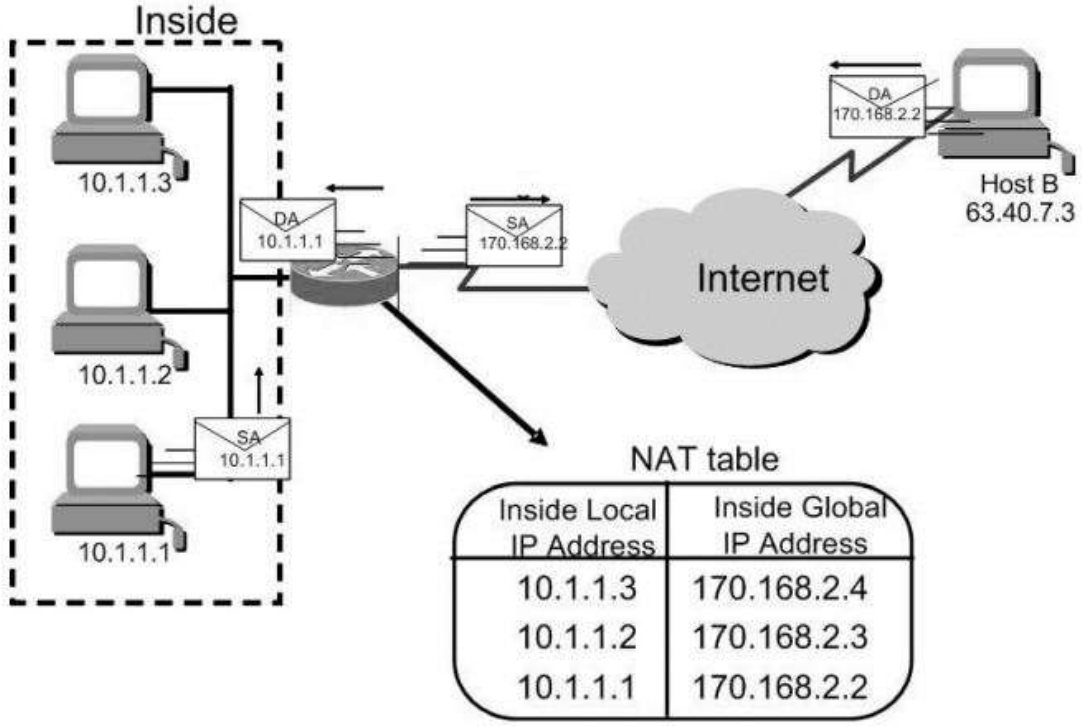
NAT (Network Address Translation)

NAT (Network Address Translation or Network Address Translator) is the translation of an Internet Protocol address (IP address) used within one network to a different IP address known within another network. One network is designated the *inside* network and the other is the *outside*. Typically, a company maps its local inside network addresses to one or more global outside IP addresses and unmaps the global IP addresses on incoming packets back into local IP addresses. This helps ensure security since each outgoing or incoming request must go through a translation process that also offers the opportunity to qualify or authenticate the request or match it to a previous request. NAT also conserves on the number of global IP addresses that a company needs and it lets the company use a single IP address in its communication with the world.

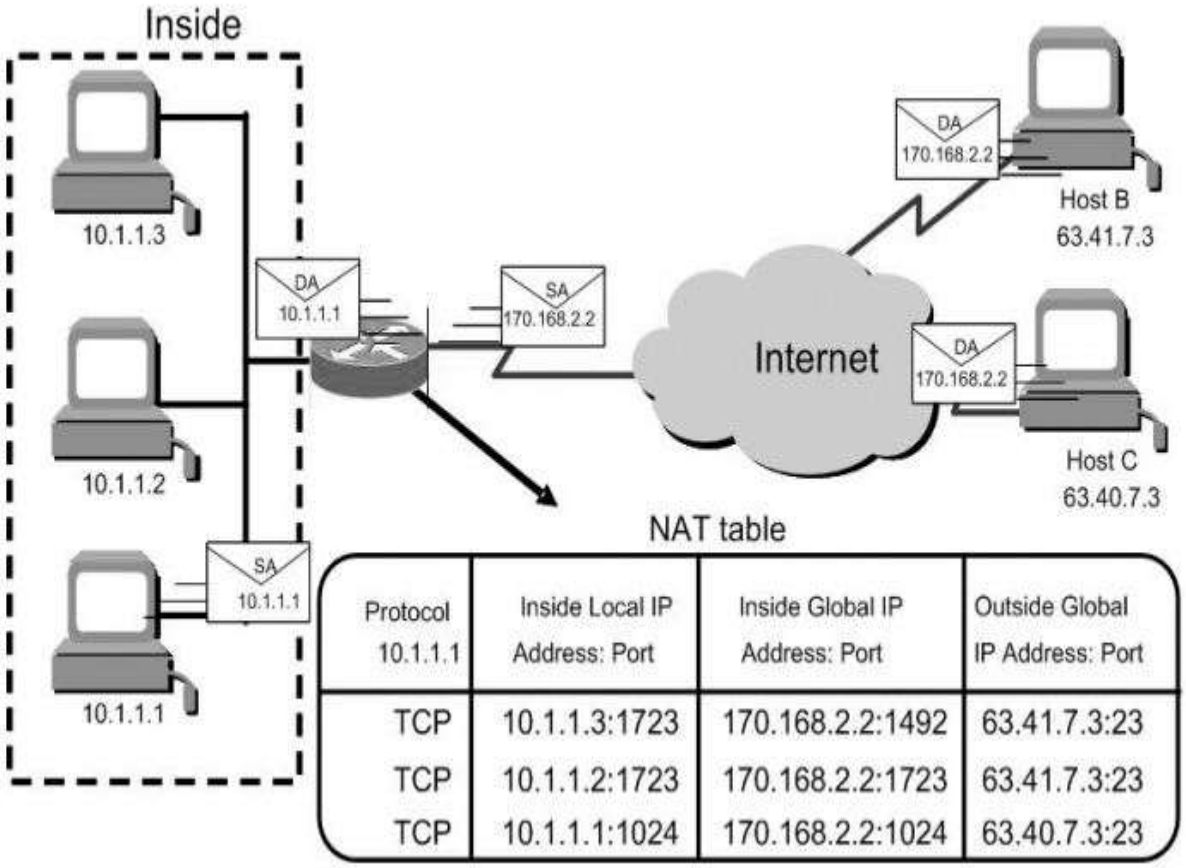
NAT is included as part of a router and is often part of a corporate firewall. Network administrators create a NAT table that does the global-to-local and local-to-global IP address mapping. NAT can also be used in conjunction with *policy routing*. NAT can be statically defined or it can be set up to dynamically translate from and to a pool of IP addresses.

Types of NAT

- **Static NAT**: A local IP address to one global IP address statically



- **Dynamic NAT:** A local IP address to any of a rotating pool of global IP addresses that a company may have



- **NAT Overloading (PAT – Port Address Translation):** A local IP address plus a particular TCP port to a global IP address or one in a pool of them.

NAT Terms

- **Inside local address**—Name of inside source inside translation
- **Outside local address**—Name of destination host before translation
- **Inside global address**—Name of inside host after translation
- **Outside global address**— Name of outside destination host after translation

Need of NAT

i) You need to connect to the internet and your hosts don't have globally unique IP addresses.

ii) You change to a new ISP that requires you to renumber your network.

iii) You need to merge two intranets with duplicate addresses.

Routing in the Internet

1. RIP (Routing Information Protocol)

→ Distance-vector routing protocol.
→ Intra AS routing Protocol.
→ Employs the hop count as a routing metric.
→ RIP prevents routing loops by implementing a limit on the number of hops allowed in a path from the source to a destination.
→ The maximum number of hops allowed for RIP is 15.
→ This hop limit, however, also limits the size of networks that RIP can support.

→ A hop count of 16 is considered an infinite distance and used to deprecate inaccessible, inoperable, or otherwise undesirable routes in the selection process.
→ Periodic updates every 30 seconds, even the topology not changed.
→ In the early deployments, routing tables were small enough that the traffic was not significant. As networks grew in size, however, it became evident there could be a massive traffic burst every 30 seconds, even if the routers had been initialized at random times
→ RIP uses the User Datagram Protocol (UDP) as its transport protocol, and is assigned the reserved port number 520

Types

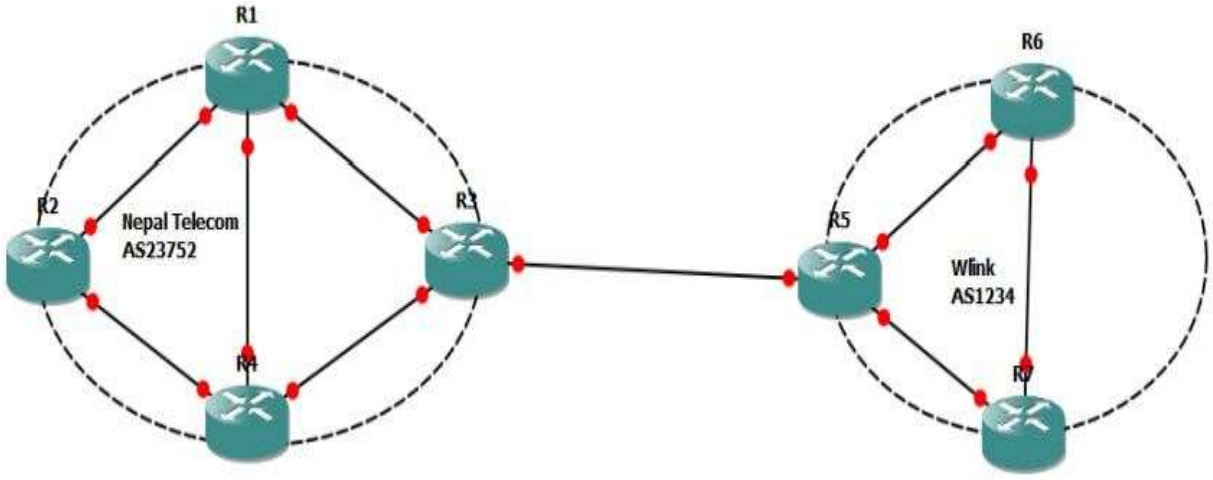
i. RIPv1 (version 1)
   — Uses classful routing.
   — The periodic routing updates do not carry subnet information,
   — Lacking support for variable length subnet masks (VLSM).
   — There is also no support for router authentication, making RIP vulnerable to various attacks.
   — Broadcast is used for database update
ii. RIPv2 (version 2)
   — Includes the ability to carry subnet information, thus supporting Classless Inter-Domain Routing (CIDR).
   — In an effort to avoid unnecessary load on hosts that do not participate in routing,
   — RIPv2 multicasts the entire routing table to all adjacent routers at the address 224.0.0.9, as opposed to RIPv1 which uses broadcast.
   — Support Authentication
iii. RIPng (next generation)
   — Support of IPv6 networking.
   — RIPng sends updates on UDP port 521 using the multicast group FF02::9.

2. OSPF (Open Shortest Path First)

→ Link State Routing Algorithm
→ Cost/Metric = Link Bandwidth
→ Shortest Path Algorithm to calculate best path from source to destination.
→ Open Shortest Path First (OSPF) is an adaptive routing protocol for Internet Protocol (IP) networks.
→ It uses a link state routing algorithm and falls into the group of interior routing protocols, operating within a single autonomous system (AS).
→ OSPF is perhaps the most widely-used interior gateway protocol (IGP) in large enterprise networks
→ It included the ability to carry subnet information, thus supporting Classless Inter-Domain Routing (CIDR).
→ Supports Authentication

3. BGP (Border Gateway Protocol)

→ Exterior Gateway protocol
→ Called Path vector Routing Algorithm.
→ Neighboring BGP routers i.e. BGP peers exchange detailed path information.
→ Used for communicating between two AS

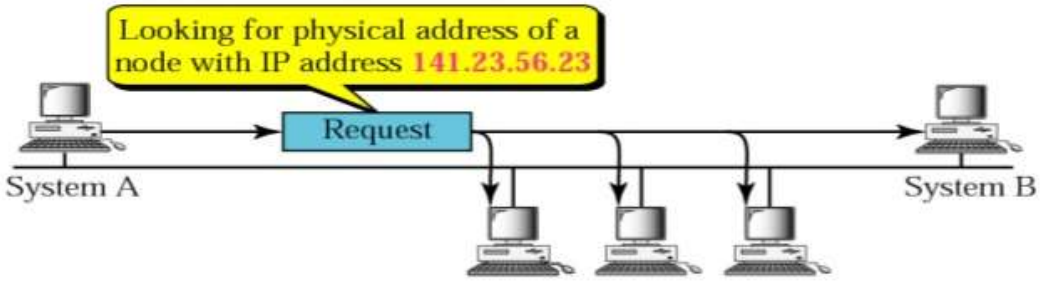


→ Revolves around three activities
- Receiving and filtering route advertisement from directly attached neighbors.
- Route Selection
- Sending route advertisements to neighbors.

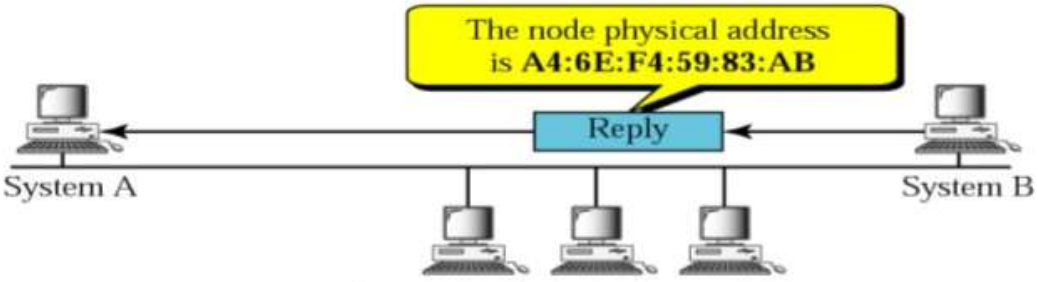
ARP (Address resolution Protocol)

Address Resolution Protocol (ARP) is a telecommunications protocol used for resolution of network layer addresses into link layer addresses

ARP Operation



a. ARP request is broadcast



b. ARP reply is unicast

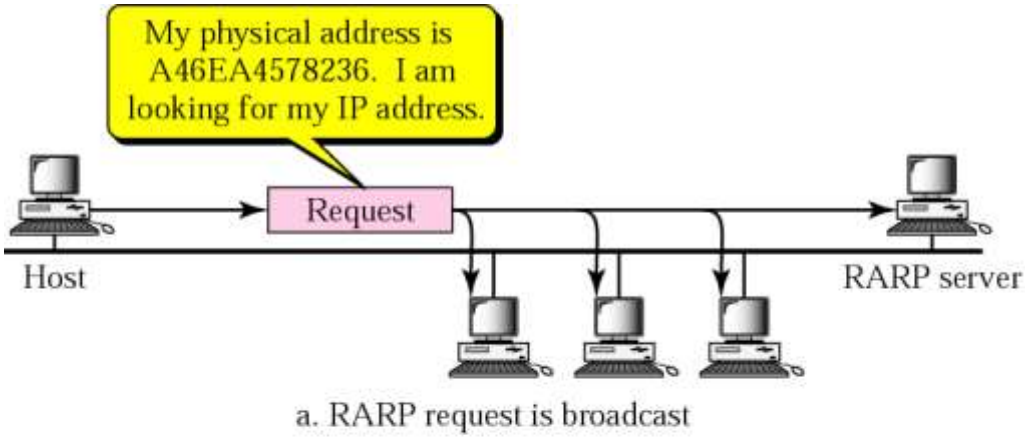
RARP (Reverse ARP)

→ RARP (Reverse Address Resolution Protocol) is a protocol by which a physical machine in a local area network can request to learn its IP address from a gateway server's Address Resolution Protocol (ARP) table or cache.

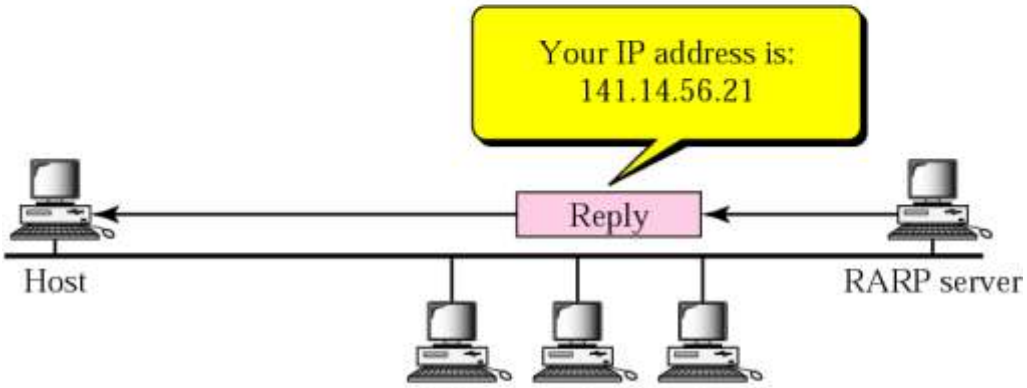
→ A network administrator creates a table in a local area network's gateway router that maps the physical machine (or Media Access Control - MAC address) addresses to corresponding Internet Protocol addresses.

→ When a new machine is set up, its RARP client program requests from the RARP server on the router to be sent its IP address. Assuming that an entry has been set up in the router table, the RARP server will return the IP address to the machine which can store it for future use.

RARP Operation



a. RARP request is broadcast



b. RARP reply is unicast

Introduction to Multicast Routing

→ In computer networking, multicast is the delivery of a message or information to a group of destination computers simultaneously in a single transmission from the source.

→ Copies are automatically created in other network elements, such as routers, but only when the topology of the network requires it.

→ Multicast is most commonly implemented in IP multicast, which is often employed in Internet Protocol (IP) applications of streaming media and Internet television.

→ In IP multicast the implementation of the multicast concept occurs at the IP routing level, where routers create optimal distribution paths for datagrams sent to a multicast destination address.

Internet Group Management Protocol (IGMP)

→ IGMP runs between hosts and the nearest multicast routers.

→ A local host can use it to inform the multicast router that which multicast group it wants to be join, while the multicast routers can use it to poll the LAN periodically, thus determine if known group members are still active.

Applications of Multicast

→ Video/audio conference

→ IP TV, Video on Demand

→ Advertisement, Stock, Distance learning

→ Distributed interactive gaming or simulations

→ Voice-over-IP

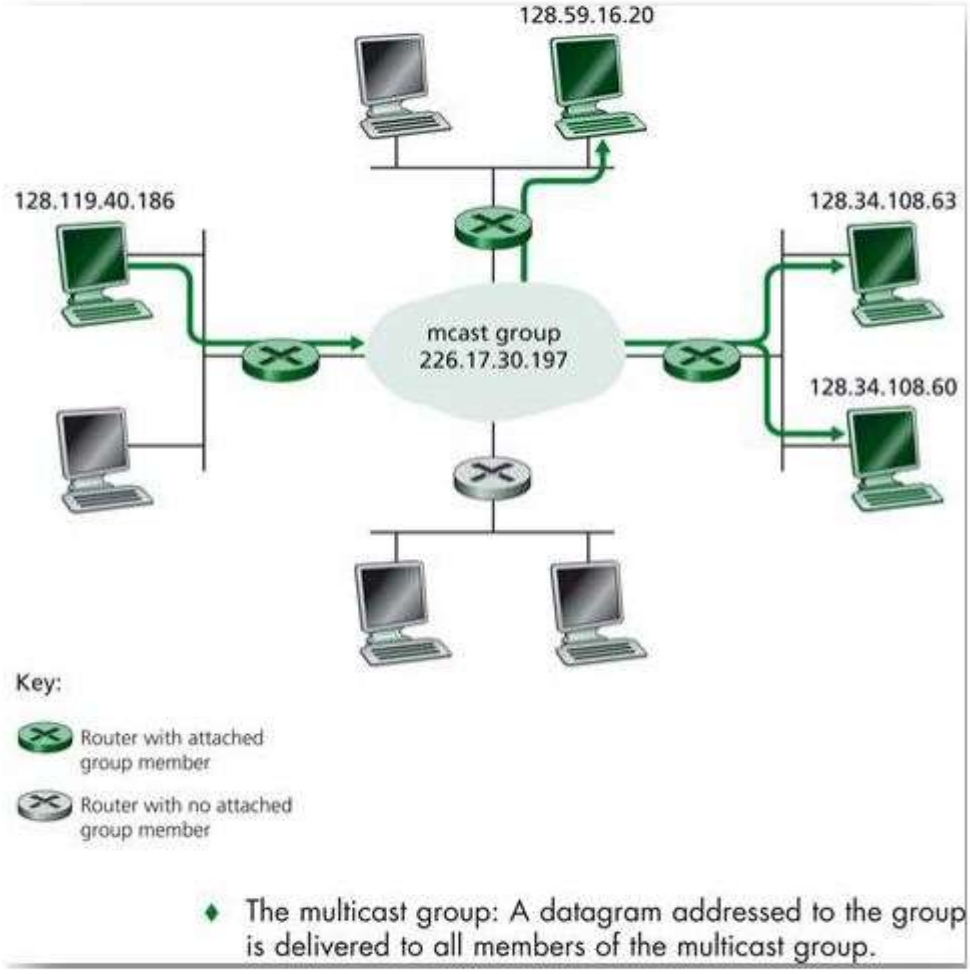
→ Synchronizing of distributed database, websites

How multicast?

→ Using Class D in IP v 4 (224-239) or addresses that begin with 1111 1111 (FF) in IP v 6

→ e.g. 224.0.0.1, FF5B:2D9D:DC28:0000:0000:FC57:D4C8:1FFF

→ Rather than sending a separate copy of the data for each recipient, the source sends the data only once using the multicast group, and routers along the way to the destinations make copies as needed.



IPv6

→ This huge growth in Internet use has not only led to increased demand for better, faster technology, but has also increased the demand for addresses from which to send and receive information.

→ 128 bits adresses

→ 2^{128} IP addresses developed

→ Every grain of sand on the planet can be IP-addressable

Limitations of IPv4

→ Address Space

→ Various unnecessary and Variable header fields

→ Fragmentation in Router

→ Addressing Model

→ NAT

→ Broadcast Versus Multicast

→ Quality of Service

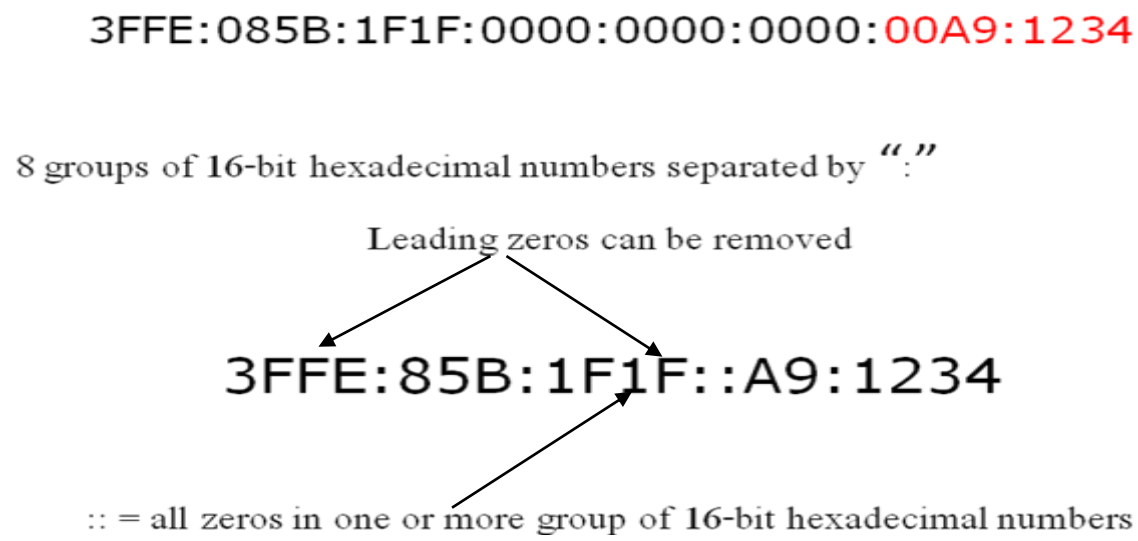
Most important changes introduced in IPv6

→ Expanded addressing capabilities
- Size increases from 32 bits to 128 bits. This ensures that the IP address wouldn't run out of IP addresses.
- In addition to unicast and multicast addresses, it introduced anycast address, which allows a datagram to be delivered to any one of a group of hosts.

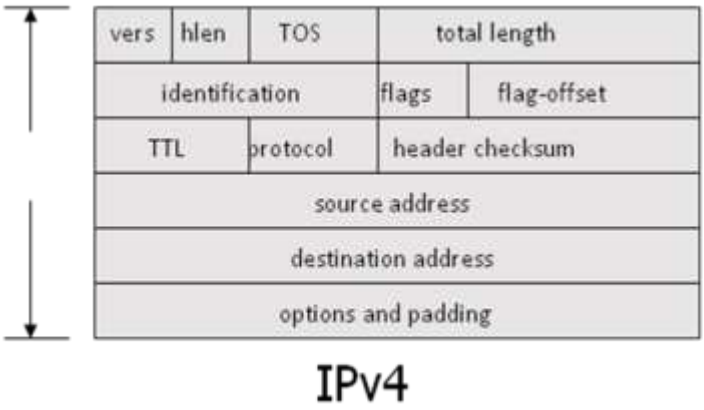
→ A streamlined 40 bytes header
- Allows for faster processing of the IP datagram

→ Flow labeling and priority
- Has an elusive definition of flow.(according to quality of service or real time service e.g. audio and video transfer)

128-bit IPv6 Address



Header comparison

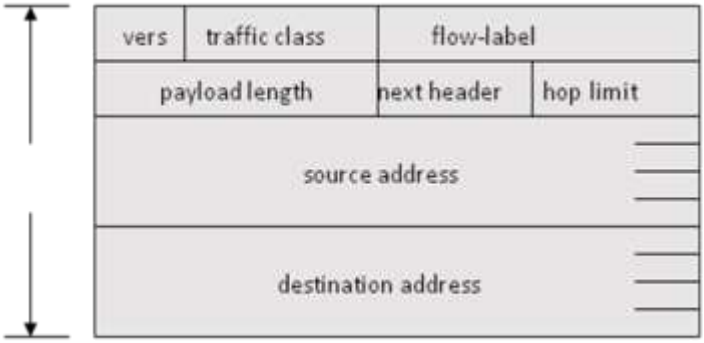


IPv4

Removed (6)

- ID, flags, flag offset
- TOS, hlen
- header checksum

Changed (3)



IPv6

Added (2)

- traffic class
- flow label

Expanded

- address 32 to 128 bits

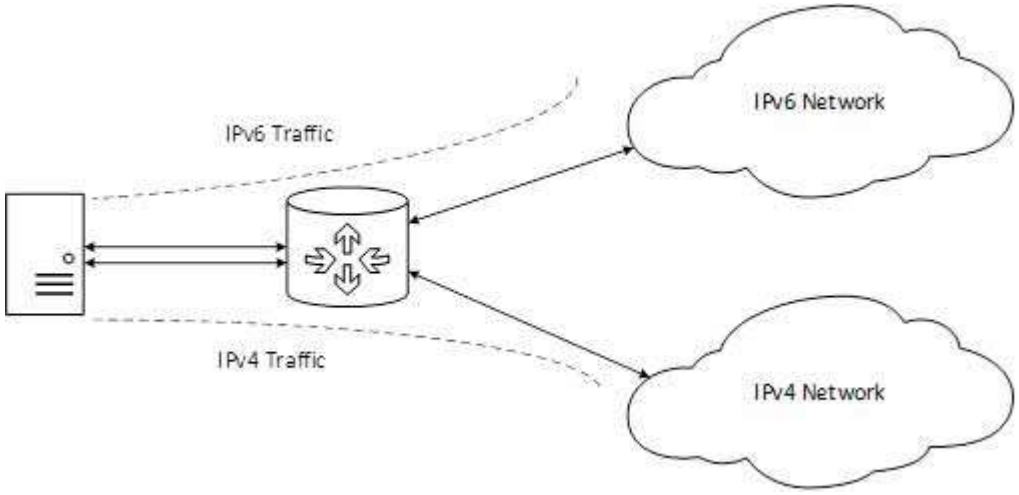
No longer present in IPv6

→ Fragmentation/Reassembly
  — Result in fast IP forwarding
→ Header checksum:
  — Result in fast processing.
→ Option field:
  — Replaced by extension header. Result in a fixed length, 40-byte IP header.

Transition from IPv4 to IPv6

1) Dual Stack

A router can be installed with both IPv4 and IPv6 addresses configured on its interfaces pointing to the network of relevant IP scheme.

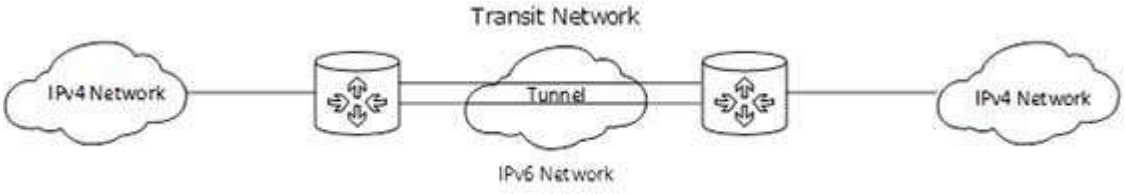


[*Image: Dual Stack Router*]

In the above diagram, a server having IPv4 as well as IPv6 address configured for it can now speak with all the hosts on both the IPv4 as well as the IPv6 networks with the help of a Dual Stack Router. The Dual Stack Router, can communicate with both the networks. It provides a medium for the hosts to access a server without changing their respective IP versions.

2) Tunneling

In a scenario where different IP versions exist on intermediate path or transit networks, tunneling provides a better solution where user's data can pass through a non-supported IP version.

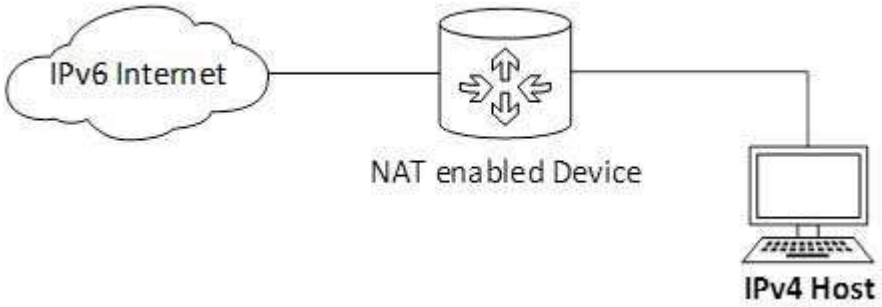


[*Image: Tunneling*]

The above diagram depicts how two remote IPv4 networks can communicate via a Tunnel, where the transit network was on IPv6. Vice versa is also possible where the transit network is on IPv6 and the remote sites that intend to communicate are on IPv4.

3) NAT Protocol Translation

This is another important method of transition to IPv6 by means of a NAT-PT (Network Address Translation – Protocol Translation) enabled device. With the help of a NAT-PT device, actual can take place happens between IPv4 and IPv6 packets and vice versa. See the diagram below:



[*Image: NAT - Protocol Translation*]

A host with IPv4 address sends a request to an IPv6 enabled server on Internet that does not understand IPv4 address. In this scenario, the NAT-PT device can help them communicate. When the IPv4 host sends a request packet to the IPv6 server, the NAT-PT device/router strips down the IPv4 packet, removes IPv4 header, and adds IPv6 header and passes it through the Internet. When a response from the IPv6 server comes for the IPv4 host, the router does vice versa.