UNIT II NETWORK LAYER PROTOCOLS 9

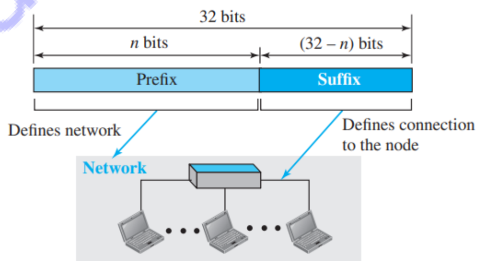
Network Layer – IPv4 Addressing – Network Layer Protocols (IP, ICMP and Mobile IP)- Unicast and Multicast Routing – Intradomain and Interdomain Routing Protocols – IPv6 Addresses – IPv6 – Datagram Format - Transition from IPv4 to IPv6.

Network Layer:

- The network layer is involved at the source host, destination host, and all routers in the path.
- At the source host (Alice), the network layer accepts a packet from a transport layer, encapsulates the packet in a datagram, and delivers the packet to the data-link layer
- At the destination host (Bob), the datagram is decapsulated, and the packet is extracted and delivered to the corresponding transport layer

IPv4 Addressing:

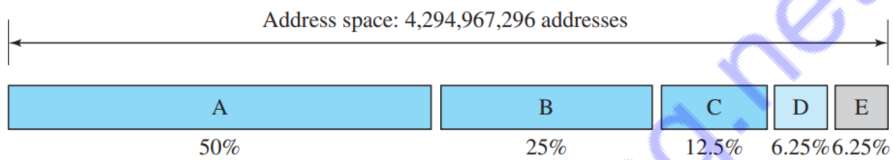
- An IPv4 address is a 32-bit address that uniquely and universally defines the connection of a device (for example, a computer or a router) to the Internet
- IPv4 addresses are unique. They are unique in the sense that each address defines one, and only one, connection to the Internet.
- Two devices on the Internet can never have the same address at the same time
- A 32-bit IPv4 address is also hierarchical, but divided only into two parts. The first part of the address, called the prefix, defines the network; the second part of the address, called the suffix, defines the node (connection of a device to the Internet)
- The prefix length is n bits and the suffix length is (32 − n) bits.
- A prefix can be fixed length or variable length.
- The scheme which uses fixed length prefix is called as classful addressing and the scheme which uses variable-length network prefix is referred to as classless addressing

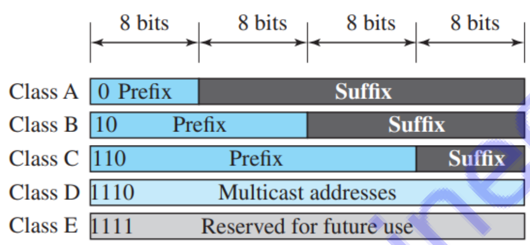


<u>**Classful Addressing**</u>

- When the Internet started, an IPv4 address was designed with a fixed-length prefix, but to accommodate both small and large networks, three fixed-length prefixes were designed instead of one (n = 8, n = 16, and n = 24). The whole address space was divided into five classes (class A, B, C, D, and E).This scheme is referred to as classful addressing
- In class A, the network length is 8 bits, but since the first bit, which is 0, defines the class, we can have only seven bits as the network identifier. This means there are only 2^7 = 128 networks in the world that can have a class A address.

- In class B, the network length is 16 bits, but since the first two bits, which are (10)2, define the class, we can have only 14 bits as the network identifier. This means there are only 2^{14} = 16,384 networks in the world that can have a class B address.
- All addresses that start with (110)2 belong to class C. In class C, the network length is 24 bits, but since three bits define the class, we can have only 21 bits as the network identifier. This means there are 2^{21} = 2,097,152 networks in the world that can have a class C address.
- Class D is not divided into prefix and suffix. It is used for multicast addresses. All addresses that start with 1111 in binary belong to class E.
- As in Class D, Class E is not divided into prefix and suffix and is used as reserve



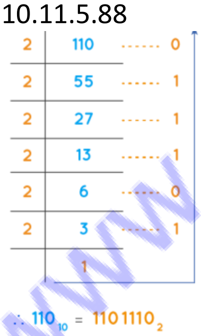


Class	Prefixes	First byte
A	n = 8 bits	0 to 127
B	n = 16 bits	128 to 191
C	n = 24 bits	192 to 223
D	Not applicable	224 to 239
E	Not applicable	240 to 255

Example:

1. Rewrite the following IP addresses using binary notation:

a. 110.11.5.88



Likewise convert everything:

a. 110.11.5.88- 01101110.00001011.00000101.01011000

b. 12.74.16.18

00001100 01001010 00010000 00010010

c. 201.24.44.32

11001001 00011000 00101100 00100000

2. Find the class of the following classful IP addresses:
a.) 130.34.54.12
130 is between 128 and 191 => Class B
b.) 200.34.2.1
200 is between 192 and 223 => Class C
c.) 245.34.2.8
245 is between 240 and 254 => Class E

3.Find the class of each address.
a. 00000001 00001011 00001011 11101111
b. 11000001 10000011 00011011 11111111
c. 14.23.120.8
d. 252.5.15.111

Solution

a. The first bit is O. This is a class A address.
b. The first 2 bits are 1; the third bit is O. This is a class C address.
c. The first byte is 14 (between 0 and 127); the class is A.
d. The first byte is 252 (between 240 and 255); the class is E

Address Depletion:

- In the Internet if the addresses were not distributed properly, the Internet was faced with the problem of the addresses being rapidly used up.
- For an Example consider class A address in this 128 organizations connected and each organization allowed to use with 16,777,216 nodes (2^{32}). Since there may be only a few organizations that are this large, most of the addresses in this class were wasted.
- Class B addresses was designed for midsize organizations, but many of the addresses in this class also remained unused.
- Class C addresses have a completely different flaw in design. The number of addresses that can be used in each network (256) was so small that most companies were not comfortable using a block in this address class. C
- Class E addresses were almost never used, wasting the whole class.

Subnetting and Supernetting:

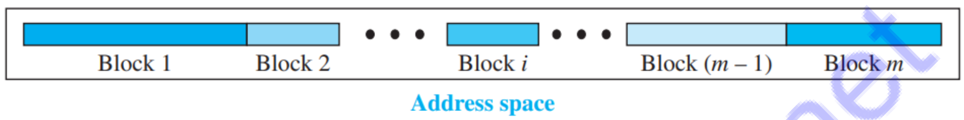
- In subnetting, a class A or class B block is divided into several subnets. Each subnet has a larger prefix length than the original network.
- If a network in class A is divided into four subnets, each subnet has a prefix of nsub = 10. At the same time, if all of the addresses in a network are not used, subnetting allows the addresses to be divided among several organizations.
- This idea did not work because most large organizations were not happy about dividing the block and giving some of the unused addresses to smaller organizations.
- While subnetting was devised to divide a large block into smaller ones, supernetting was devised to combine several class C blocks into a larger block to be attractive

Advantage of Classful Addressing:

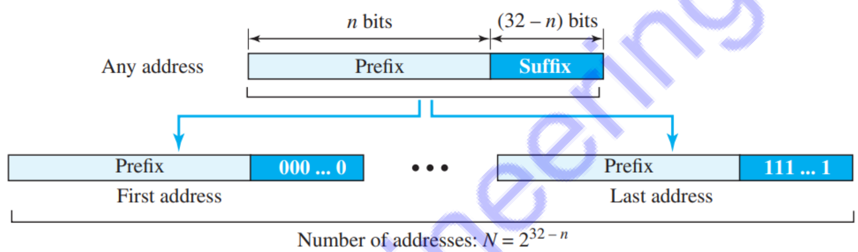
- For the given an address, we can easily find the class of the address and, since the prefix length for each class is fixed, we can find the prefix length immediately.

Classless Addressing:

- Subnetting and supernetting in classful addressing did not really solve the address depletion problem. With the growth of the Internet, it was clear that a larger address space was needed as a long-term solution
- In classless addressing, the whole address space is divided into variable length blocks. The prefix in an address defines the block
- Theoretically, there is a block of 2^0, 2^1 , 2^2 , ..., 2^{32} addresses.
- An organization can be granted one block of addresses



- The prefix length in classless addressing is variable it ranges from 0 to 32
- A small prefix means a larger network; a large prefix means a smaller network.

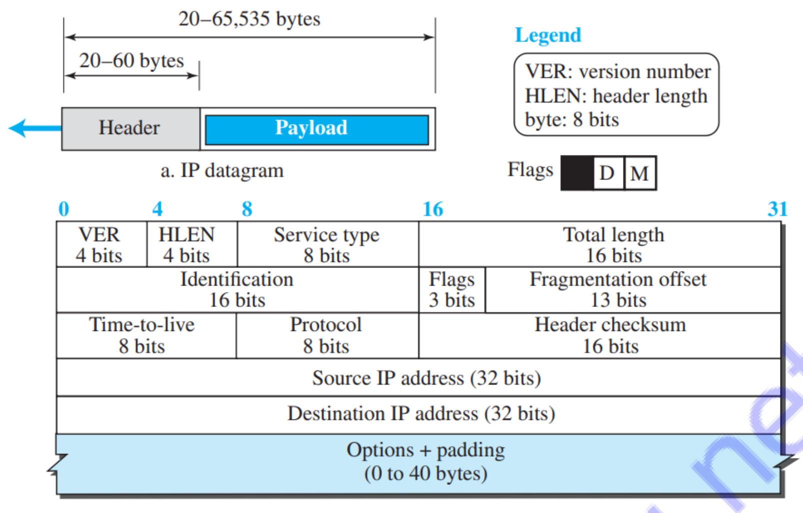


Internet Protocol version 4 (IPv4):

- Internet Protocol version 4 (IPv4), is responsible for packetizing, forwarding, and delivery of a packet at the network layer
- IPv4 is an unreliable datagram protocol—a best-effort delivery service.
- The term best-effort means that IPv4 packets can be corrupted, be lost, arrive out of order, or be delayed, and may create congestion for the network.
- IPv4 is also a connectionless protocol that uses the datagram approach.
- Packets used by the IP are called datagrams.
- Each datagram is handled independently, and each datagram can follow a different route to the destination

Datagram Format:

- IPv4 defines the format of a packet in which the data coming from the upper layer or other protocols are encapsulated
- A datagram is a variable-length packet consisting of two parts: **header and payload (data).** The **header is 20 to 60 bytes** in length and contains information essential to routing and delivery.



Version Number: The 4-bit version number (VER) field defines the version of the IPv4 protocol

Header Length:

- The 4-bit header length (HLEN) field defines the total length of the datagram header in 4-byte words. The IPv4 datagram has a variable-length header.
- When a device receives a datagram, it needs to know when the header stops and the data, which is encapsulated in the packet, starts.
- However, to make the value of the header length (number of bytes) fit in a 4-bit header length, the total length of the header is calculated as 4-byte words.
- The total length is divided by 4 and the value is inserted in the field. The receiver needs to multiply the value of this field by 4 to find the total length.

Service:

- In the original design of the IP header, this field was referred to as type of service (TOS), which defined how the datagram should be handled.

Total Length.

- This 16-bit field defines the total length (header plus data) of the IP datagram in bytes.
- A 16-bit number can define a total length of up to 65,535 (when all bits are 1s). However, the size of the datagram is normally much less than this.
- This field helps the receiving device to know when the packet has completely arrived.
- To find the length of the data coming from the upper layer, subtract the header length from the total length.
- The header length can be found by multiplying the value in the HLEN field by 4.

Length of data = total length − (HLEN) × 4

Identification, Flags, and Fragmentation Offset.

- These three fields are related to the fragmentation of the IP datagram when the size of the datagram is larger than the underlying network can carry.

Time-to-live.

- Due to some malfunctioning of routing protocols a datagram may be circulating in the Internet, visiting some networks over and over without reaching the destination.
- This may create extra traffic in the Internet.

- The time-to-live (TTL) field is used to control the maximum number of hops visited by the datagram.
- When a source host sends the datagram, it stores a number in this field.
- This value is approximately two times the maximum number of routers between any two hosts.
- Each router that processes the datagram decrements this number by one. If this value, after being decremented, is zero, the router discards the datagram

Protocol

- In TCP/IP, the data section of a packet, called the payload, carries the whole packet from another protocol.
- A datagram can also carry a packet from other protocols that directly use the service of the IP, such as some routing protocols or some auxiliary protocols

Header checksum

- IP is not a reliable protocol; it does not check whether the payload carried by a datagram is corrupted during the transmission.
- IP puts the burden of error checking of the payload on the protocol that owns the payload, such as UDP or TCP.
- The datagram header, however, is added by IP, and its error-checking is the responsibility of IP.

Source and Destination Addresses.

- These 32-bit source and destination address fields define the IP address of the source and destination respectively. The source host should know its IP address

Options

- A datagram header can have up to 40 bytes of options. Options can be used for network testing and debugging.
- Although options are not a required part of the IP header, option processing is required of the IP software.

Payload

- Payload, or data, is the main reason for creating a datagram. Payload is the packet coming from other protocols that use the service of IP

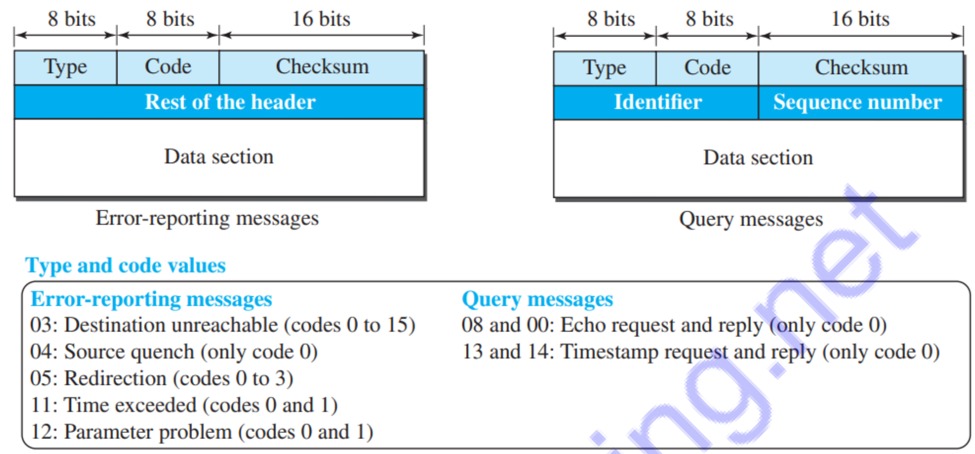
ICMPv4 Internet Control Message Protocol version 4

- The IPv4 has no error-reporting or error-correcting mechanism
- The IP protocol also lacks a mechanism for host and management queries. A host sometimes needs to determine if a router or another host is alive
- The Internet Control Message Protocol version 4 (ICMPv4) has been designed to compensate for the above two deficiencies. It is a companion to the IP protocol.
- It is a companion to the IP protocol. ICMP itself is a network-layer protocol. However, its messages are not passed directly to the data-link layer as would be expected
- When an IP datagram encapsulates an ICMP message, the value of the protocol field in the IP datagram is set to 1 to indicate that the IP payroll is an ICMP message

Messages:

- ICMP messages are divided into two broad categories: **error-reporting messages and query messages.** The error-reporting messages report problems that a router or a host (destination) may encounter when it processes an IP packet.
- The query messages, which occur in pairs, help a host or a network manager get specific information from a router or another host. For example, nodes can discover their neighbors.

- An ICMP message has an 8-byte header and a variable-size data section
- The first field, ICMP type, defines the type of the message. The code field specifies the reason for the particular message type. The last common field is the checksum field. The rest of the header is specific for each message type.



- The data section in error messages carries information for finding the original packet that had the error. In query messages, the data section carries extra information based on the type of query

Error Reporting Messages

- Since IP is an unreliable protocol, one of the main responsibilities of ICMP is to report some errors that may occur during the processing of the IP datagram.
- ICMP does not correct errors, it simply reports them.
- Error correction is left to the higher-level protocols.
- Error messages are always sent to the original source because the only information available in the datagram about the route is the source and destination IP addresses.
- ICMP uses the source IP address to send the error message to the source (originator) of the datagram.
- To make the error-reporting process simple, ICMP follows some rules in reporting messages

Destination Unreachable

- The most widely used error message is the destination unreachable (type 3).
- This message uses different codes (0 to 15) to define the type of error message and the reason why a datagram has not reached its final destination

Source Quench

- It informs the sender that the network has encountered congestion and the datagram has been dropped; the source needs to slow down sending more datagrams

Redirection Message

- The redirection message (type 5) is used when the source uses a wrong router to send out its message.
- The router redirects the message to the appropriate router, but informs the source that it needs to change its default router in the future. The IP address of the default router is sent in the message.

Parameter Problem

- A parameter problem message (type 12) can be sent when either there is a problem in the header of a datagram (code 0) or some options are missing or cannot be interpreted (code 1).

Query Messages

- Query messages in ICMP can be used independently without relation to an IP datagram.
- Query messages are used to probe or test the liveliness of hosts or routers in the Internet, find the one-way or the round-trip time for an IP datagram between two devices, or even find out whether the clocks in two devices are synchronized.
- Naturally, query messages come in pairs: request and reply. The echo request (type 8) and the echo reply (type 0) pair of messages are used by a host or a router to test the liveliness of another host or router.
- A host or router sends an echo request message to another host or router; if the latter is alive, it responds with an echo reply message
- The timestamp request (type 13) and the timestamp reply (type 14) pair of messages are used to find the round-trip time between two devices or to check whether the clocks in two devices are synchronized.
- The timestamp request message sends a 32-bit number, which defines the time the message is sent.
- The timestamp reply resends that number, but also includes two new 32-bit numbers representing the time the request was received and the time the response was sent.

MOBILE IP

- Mobile IP, the extension of IP protocol that allows mobile computers to be connected to the Internet at any location where the connection is possible

Addressing-Stationary Hosts:

- The original IP addressing was based on the assumption that a host is stationary, attached to one specific network.
- A router uses an IP address to route an IP datagram. An IP address has two parts: a prefix and a suffix. The prefix associates a host with a network. For example, the IP address 10.3.4.24/8 defines a host attached to the network 10.0.0.0/8.
- The host in the Internet does not have an address that it can carry with itself from one place to another.
- The address is valid only when the host is attached to the network.
- If the network changes, the address is no longer valid.

Mobile Hosts:

- When a host moves from one network to another, the IP addressing structure needs to be modified. Several solutions have been proposed to solve this issue

Changing the Address:

- One simple solution to address the mobile host problem is to change its address as it goes to the new network.

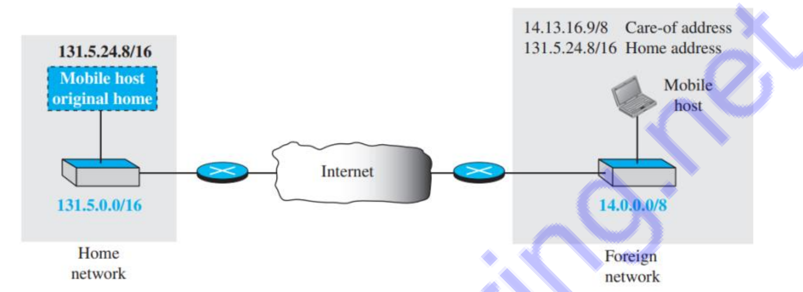
There are many drawbacks with this approach

- The configuration files would need to be changed.
- Each time the computer moves from one network to another, it must be rebooted.
- The DNS tables need to be revised so that every other host in the Internet is aware of the change.

- The host roams from one network to another during a transmission, the data exchange will be interrupted.

Two Addresses

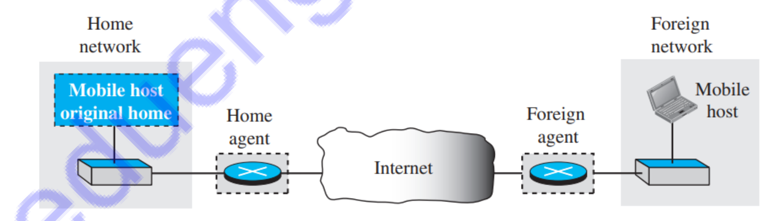
- The approach that is more feasible is the use of two addresses. The host has its original address, called the **home address**, and a temporary address, called the **care-of address**.
- The home address is permanent; it associates the host with its home network, the network that is the permanent home of the host. The care-of address is temporary.
- When a host moves from one network to another, the care-of address changes; it is associated with the foreign network, the network to which the host moves.



- Mobile IP has two addresses for a mobile host: one home address and one care-of address. The home address is permanent; the care-of address changes as the mobile host moves from one network to another

Agents

- To make the change of address transparent to the rest of the Internet requires a home agent and a foreign agent.



Home Agent

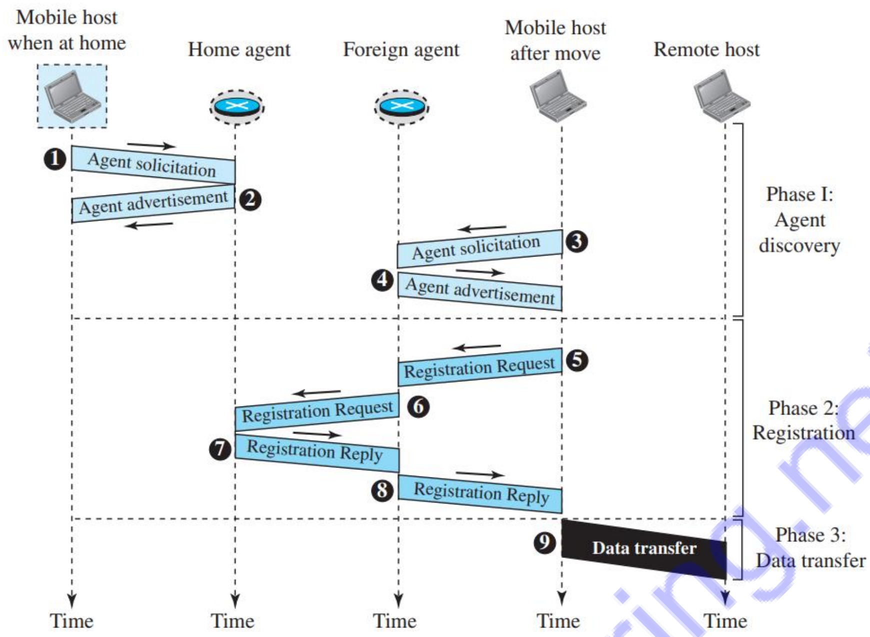
- The home agent is usually a router attached to the home network of the mobile host. The home agent acts on behalf of the mobile host when a remote host sends a packet to the mobile host.
- The home agent receives the packet and sends it to the foreign agent.

Foreign Agent

- The foreign agent is usually a router attached to the foreign network. The foreign agent receives and delivers packets sent by the home agent to the mobile host.
- When the mobile host acts as a foreign agent, the care-of address is called a collocated care-of address.

Three Phases

- To communicate with a remote host, a mobile host goes through three phases: agent discovery, registration, and data transfer



Agent Discovery

- The first phase in mobile communication, agent discovery, consists of two sub phases. A mobile host must discover (learn the address of) a home agent before it leaves its home network. A mobile host must also discover a foreign agent after it has moved to a foreign network

Agent Advertisement

- When a router advertises its presence on a network using an ICMP router advertisement, it can append an agent advertisement to the packet if it acts as an agent

ICMP Advertisement message			
Type	Length	Sequence number	
Lifetime		Code	Reserved
Care-of addresses (foreign agent only)			

Type. The 8-bit type field is set to 16.
Length. The 8-bit length field defines the total length of the extension message (not the length of the ICMP advertisement message).
Sequence number. The 16-bit sequence number field holds the message number. The recipient can use the sequence number to determine if a message is lost.
Lifetime. The lifetime field defines the number of seconds that the agent will accept requests. If the value is a string of 1s, the lifetime is infinite.
Code. The code field is an 8-bit flag in which each bit is set (1) or unset (0).
Care-of Addresses. This field contains a list of addresses available for use as careof addresses

Agent Solicitation

- When a mobile host has moved to a new network and has not received agent advertisements, it can initiate an agent solicitation

Bit	Meaning
0	Registration required. No collocated care-of address.
1	Agent is busy and does not accept registration at this moment.
2	Agent acts as a home agent.
3	Agent acts as a foreign agent.
4	Agent uses minimal encapsulation.
5	Agent uses generic routing encapsulation (GRE).
6	Agent supports header compression.
7	Unused (0).

Registration

- The second phase in mobile communication is registration. After a mobile host has moved to a foreign network and discovered the foreign agent, it must register.

Registration Request

- A registration request is sent from the mobile host to the foreign agent to register its care-of address and also to announce its home address and home agent address

Type	Flag	Lifetime
Home address		
Home agent address		
Care-of address		
Identification		
Extensions ...		

Lifetime.

This field defines the number of seconds the registration is valid. If the field is a string of 0s, the request message is asking for deregistration. If the field is a string of 1s, the lifetime is infinite.

Home address. This field contains the permanent (first) address of the mobile host.

Home agent address. This field contains the address of the home agent.

Care-of address. This field is the temporary (second) address of the mobile host.

Identification. This field contains a 64-bit number that is inserted into the request by the mobile host and repeated in the reply message. It matches a request with a reply.

Extensions. Variable length extensions are used for authentication

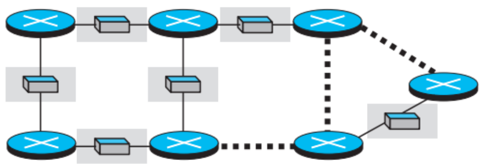
Unicast Routing:

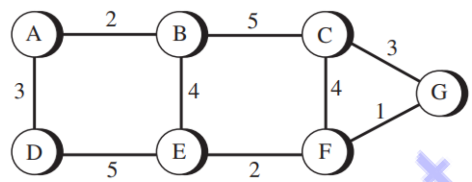
- In unicast routing, a packet is routed, hop by hop, from its source to its destination by the help of forwarding tables.
- The source host needs no forwarding table because it delivers its packet to the default router in its local network.
- The destination host needs no forwarding table because it receives the packet from its default router in its local network.
- Routing a packet from its source to its destination means routing the packet from a source router to a destination router.

Least-Cost Routing

When an internet is modeled as a weighted graph, one of the ways to interpret the best route from the source router to the destination router is to find the least cost between the two.
That is, the source router chooses a route to the destination router in such a way that the total cost for the route is the least cost among all possible routes.

- Consider the figure (below) assume that the best route between A and E is determined. There are two possible routes–one is from A-**D-E with the cost of 8 and the other is A-B-E, with the cost of 6**. Among that the path has **least cost is A-B-E with the cost of 6 has been choosen**.
- This means that each router needs to find the least-cost route between itself and all the other routers to be able to route a packet towards the destination.





Routing Table:

- To route a packet in the network a host or a router has a routing table with an entry for each destination, or a combination of destinations, to route IP packets.
- The routing table can be either static or dynamic.

Static Routing Table

- A static routing table contains information entered manually.
- The administrator enters the route for each destination into the table. When a table is created, it cannot update automatically when there is a change in the Internet.
- The table must be manually altered by the administrator.
- A static routing table can be used in a small internet that does not change very often.

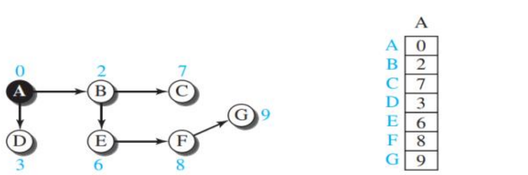
Dynamic Routing Table

- A dynamic routing table is updated periodically by using one of the dynamic routing protocols such as RIP, OSPF, or BGP.
- Whenever there is a change in the Internet, such as a shutdown of a router or breaking of a link, the dynamic routing protocols update all the tables in the routers automatically.

Distance Vector Routing:

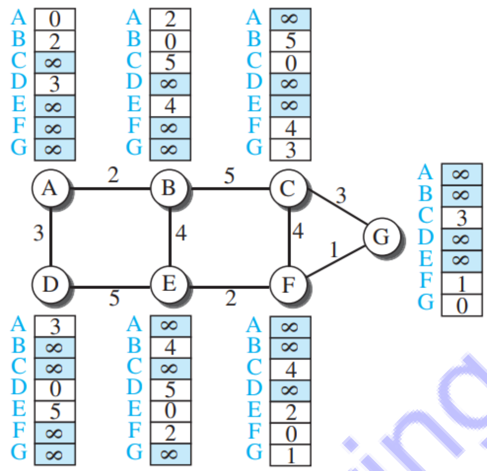
- In distance vector routing, the least-cost route between any two nodes is the route with minimum distance
- In distance-vector routing, the first thing each node creates is its own least-cost tree with the limited information it has about its immediate neighbors.
- The incomplete trees are exchanged between immediate neighbors to make the trees more and more complete and to represent the whole internet.
- In distance-vector routing, a router continuously tells all of its neighbors what it knows about the whole internet
- A least-cost tree is a combination of least-cost paths from the root of the tree to all destinations.
- These paths are graphically fix together to form the tree.
- Distance-vector routing unfixes these paths and creates a distance vector, a one-dimensional array to represent the tree

Example: Consider the graph shown in figure



a. Tree for node A b. Distance vector for node A

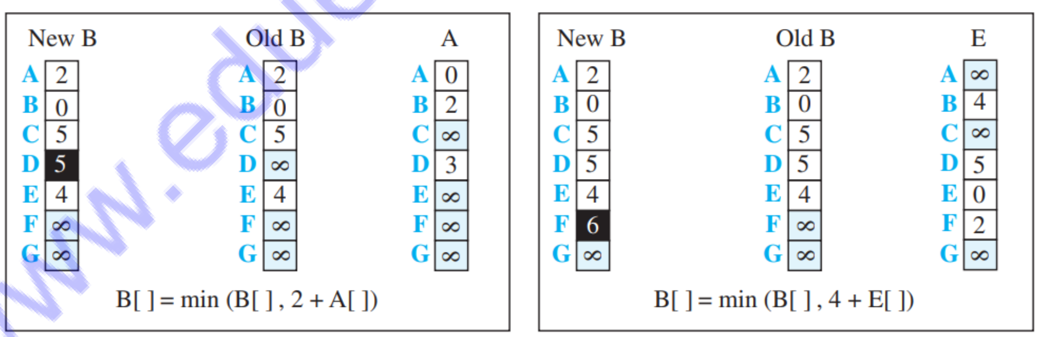
- The node sends some greeting messages to identify the immediate neighbors and the distance between itself and each neighbor.
- It then makes a simple distance vector by inserting the distances in the corresponding cells and leaves the value of other cells not connected as infinity.



- After each node has created its vector, it sends a copy of the vector to all its immediate neighbors.
- After a node receives a distance vector from a neighbor, it updates its distance vector using the Bellman-Ford equation

$$D_{xy} = \min\{D_{xy}, (c_{xz} + D_{zy})\}$$

- Consider the figure, In the first event, node A has sent its vector to node B. Node B updates its vector using the cost $C_{BA} = 2$. In the second event, node E has sent its vector to node B. Node B updates its vector using the cost $C_{EA} = 4$.



a. First event: B receives a copy of A's vector. b. Second event: B receives a copy of E's vector.

- After the first event, node B has one improvement in its vector: its least cost to node D has changed from infinity to 5 (via node A). After the second event, node B has one more improvement in its vector; its least cost to node F has changed from infinity to 6 (via node E).

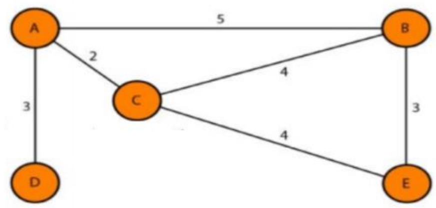
Count to Infinity:

- A problem with distance-vector routing is that any decrease in cost propagates quickly, but any increase in cost will propagate slowly.

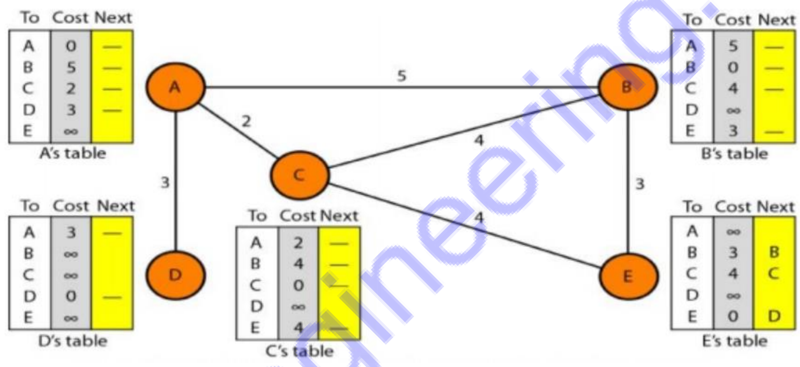
- For a routing protocol to work properly, if a link is broken, every other router should be aware of it immediately, but in distance-vector routing, this takes some time. The problem is referred to as count to infinity.

Example:

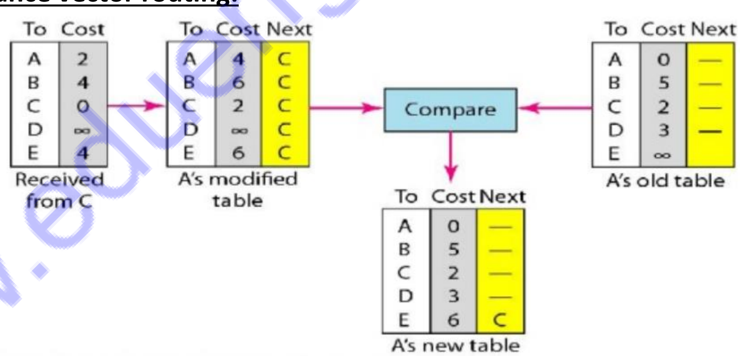
Consider the graph shown in Figure



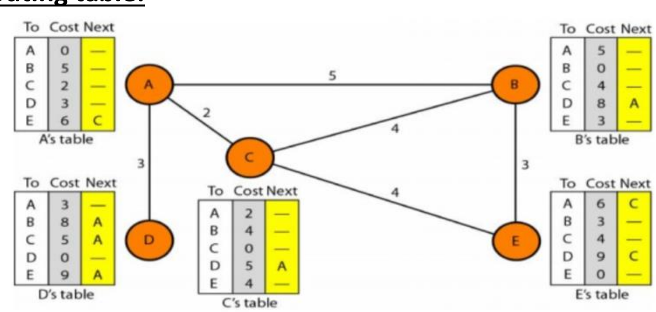
<u>Initialization Table:</u>



<u>Updating Distance vector routing:</u>



<u>Distance vector routing table:</u>

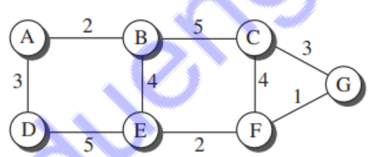


Link state Routing:

- A routing algorithm that directly creates least-cost trees and forwarding tables is link-state (LS) routing.
- This method uses the term link-state to define the characteristic of a link (an edge) that represents a network in the internet.
- In this algorithm the cost associated with an edge defines the state of the link.
- Links with lower costs are preferred to links with higher costs; if the cost of a link is infinity, it means that the link does not exist or has been broken.

Link-State Database (LSDB):

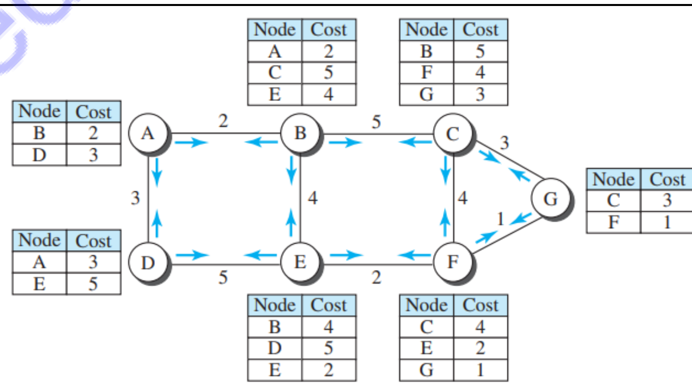
- To create a least-cost tree with this method, each node needs to have a complete map of the network, which means it needs to know the state of each link.
- The collection of states for all links is called the link-state database (LSDB). There is only one LSDB for the whole internet This process of creating LSDB is called flooding.
- Each node can send some greeting messages to all its immediate neighbors to collect two pieces of information for each neighboring node: the identity of the node and the cost of the link. The combination of these two pieces of information is called the LS packet (LSP)
- When a node receives an LSP from one of its interfaces, it compares the LSP with the copy it may already have. If the newly arrived LSP is older than the one it has (by checking the sequence number), it discards the LSP. If it is newer or the first one received, the node discards the old LSP (if there is one) and keeps the received one.
- It then sends a copy of it out of each interface except the one from which the packet arrived



a. The weighted graph

	A	B	C	D	E	F	G
A	0	2	∞	3	∞	∞	∞
B	2	0	5	∞	4	∞	∞
C	∞	5	0	∞	∞	4	3
D	3	∞	∞	0	5	∞	∞
E	∞	4	∞	5	0	2	∞
F	∞	∞	4	∞	2	0	1
G	∞	∞	3	∞	∞	1	0

b. Link state database

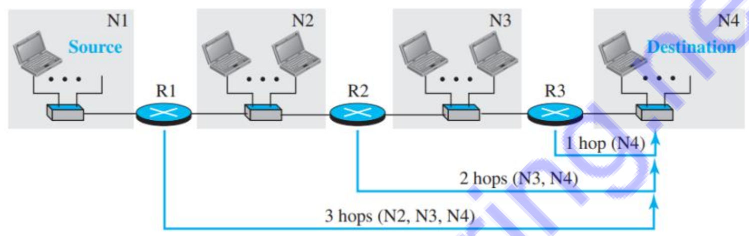


- In the distance-vector routing algorithm, each router tells its neighbors what it knows about the whole internet. In the link-state routing algorithm, each router tells the whole internet what it knows about its neighbors.

Intradomain Protocols:

Routing Information Protocol (RIP):

- The Routing Information Protocol (RIP) is one of the most widely used intradomain routing protocols based on the distance-vector routing algorithm
- RIP was started as part of the Xerox Network System (XNS), but it was the Berkeley Software Distribution (BSD) version of UNIX.
- A router in this protocol basically implements the distance-vector routing algorithm
- First, since a router in an AS needs to know how to forward a packet to different networks (subnets) in an AS, RIP routers advertise the cost of reaching different networks instead of reaching other nodes in a theoretical graph.
- Second, to make the implementation of the cost simpler, the cost is defined as the number of hops, which means the number of networks (subnets) a packet needs to travel through from the source router to the final destination host



Forwarding Tables

- A forwarding table in RIP is a three-column table in which the first column is the address of the destination network, the second column is the address of the next router to which the packet should be forwarded, and the third column is the cost (the number of hops) to reach the destination network.

Forwarding table for R1

Destination network	Next router	Cost in hops
N1	——	1
N2	——	1
N3	R2	2
N4	R2	3

Forwarding table for R2

Destination network	Next router	Cost in hops
N1	R1	2
N2	——	1
N3	——	1
N4	R3	2

Forwarding table for R3

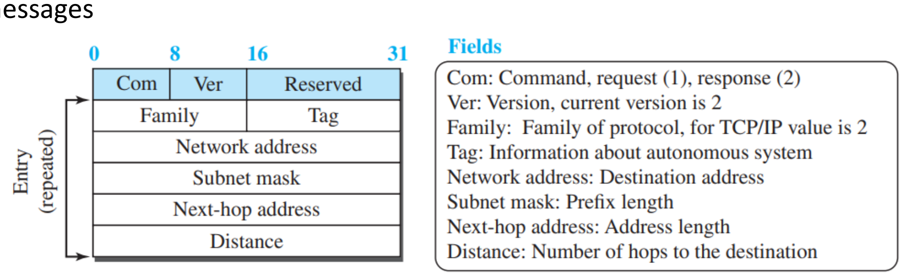
Destination network	Next router	Cost in hops
N1	R2	3
N2	R2	2
N3	——	1
N4	——	1

RIP Implementation:

- RIP is implemented as a process that uses the service of UDP on the well-known port number 520.
- RIP has gone through two versions: RIP-1 and RIP-2. The second version is backward compatible with the first section; it allows the use of more information in the RIP messages that were set to 0 in the first version.

RIP Messages:

- Two RIP processes, a client and a server, like any other processes, need to exchange messages



- RIP has two types of messages: request and response. A request message is sent by a router that has just come up or by a router that has some time-out entries. A request message can ask about specific entries or all entries.
- A response (or update) message can be either solicited or unsolicited. A solicited response message is sent only in answer to a request message. It contains information about the destination specified in the corresponding request message

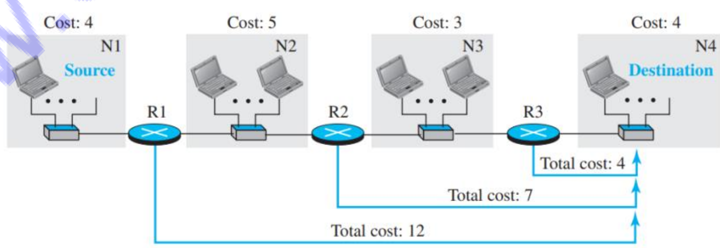
RIP Algorithm:

RIP implements the same algorithm as the distance-vector routing algorithm. There are some changes added in that algorithm

- Instead of sending only distance vectors, a router needs to send the whole contents of its forwarding table in a response message.
- The receiver adds one hop to each cost and changes the next router field to the address of the sending router. We call each route in the modified forwarding table the received route and each route in the old forwarding table the old route.
- The received router selects the old routes as the new ones except in the following three cases:
  1. If the received route does not exist in the old forwarding table, it should be added to the route.
  2. If the cost of the received route is lower than the cost of the old one, the received route should be selected as the new one.
  3. If the cost of the received route is higher than the cost of the old one, but the value of the next router is the same in both routes, the received route should be selected as the new one.
  4. The new forwarding table needs to be sorted according to the destination route

Open Shortest Path First (OSPF):

- OSPF is an open protocol, which means that the specification is a public document.
- Like RIP, the cost of reaching a destination from the host is calculated from the source router to the destination network.
- However, each link (network) can be assigned a weight based on the throughput, round-trip time, reliability, and so on. An administration can also decide to use the hop count as the cost.



Forwarding Tables

- Each OSPF router can create a forwarding table after finding the shortest-path tree between itself and the destination using Dijkstra's algorithm
- Compared with RIP, which is normally used in small ASs, OSPF was designed to be able to handle routing in a small or large autonomous system.
- However, the formation of shortest-path trees in OSPF requires that all routers flood the whole AS with their LSPs to create the global LSDB.

- Although this may not create a problem in a small AS, it may have created a huge volume of traffic in a large AS.
- To prevent this, the AS needs to be divided into small sections called areas. Each area acts as a small independent domain for flooding LSPs. OSPF uses two level of hierarchy in routing: the first level is the autonomous system, the second is the area.

Forwarding table for R1

Destination network	Next router	Cost
N1	——	4
N2	——	5
N3	R2	8
N4	R2	12

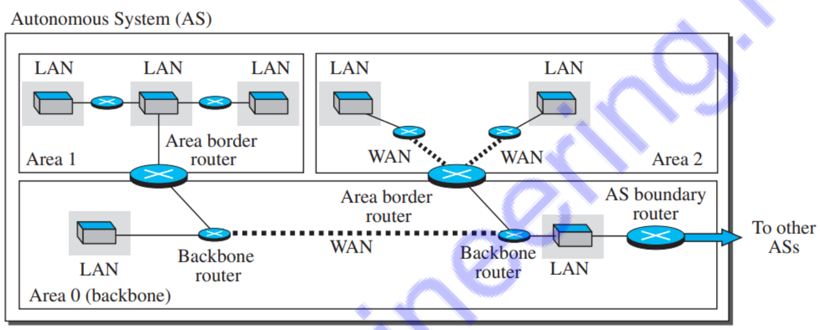
Forwarding table for R2

Destination network	Next router	Cost
N1	R1	9
N2	——	5
N3	——	3
N4	R3	7

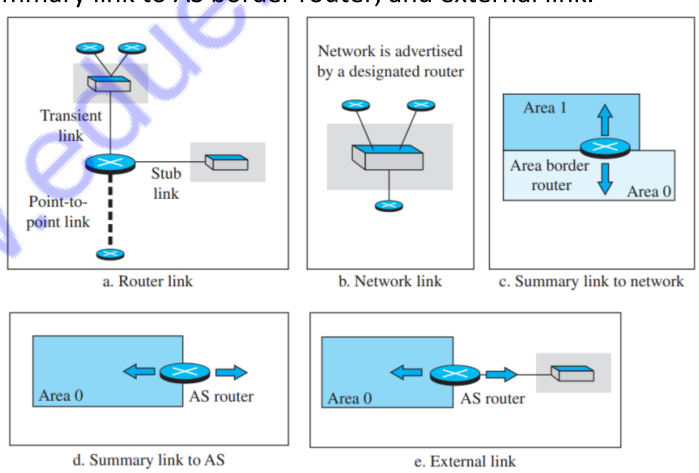
Forwarding table for R3

Destination network	Next router	Cost
N1	R2	12
N2	R2	8
N3	——	3
N4	——	4

- The routers in the backbone area are responsible for passing the information collected by each area to all other areas



- OSPF is based on the link-state routing algorithm, which requires that a router advertise the state of each link to all neighbors for the formation of the LSDB.
- There are five types of link-state advertisements are there: router link, network link, summary link to network, summary link to AS border router, and external link.

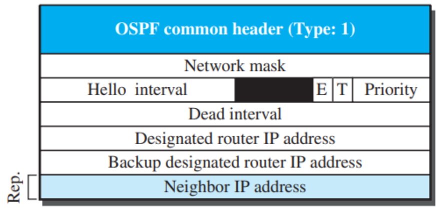


OSPF Messages:

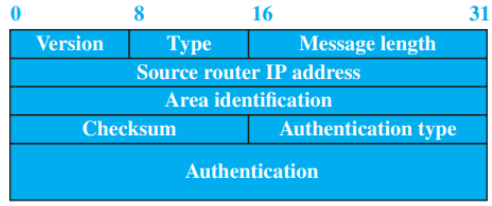
OSPF is a very complex protocol; it uses five different types of messages.

- The hello message (type 1) is used by a router to introduce itself to the neighbors and announce all neighbors that it already knows.
- The database description message (type 2) is normally sent in response to the hello message to allow a newly joined router to acquire the full LSDB.
- The linkstate request message (type 3) is sent by a router that needs information about a specific LS.

- The link-state update message (type 4) is the main OSPF message used for building the LSDB.



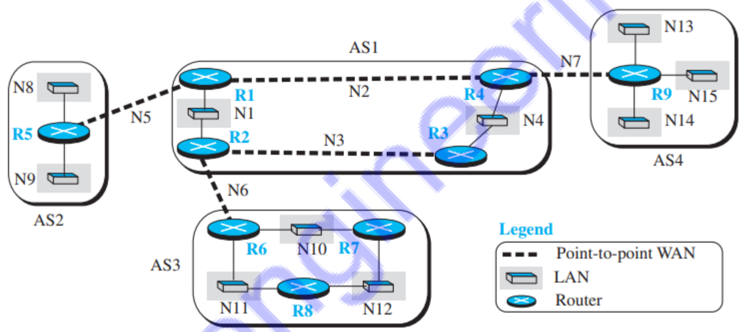
Hello message



OSPF common header

Border Gateway Protocol Version 4 (BGP4):

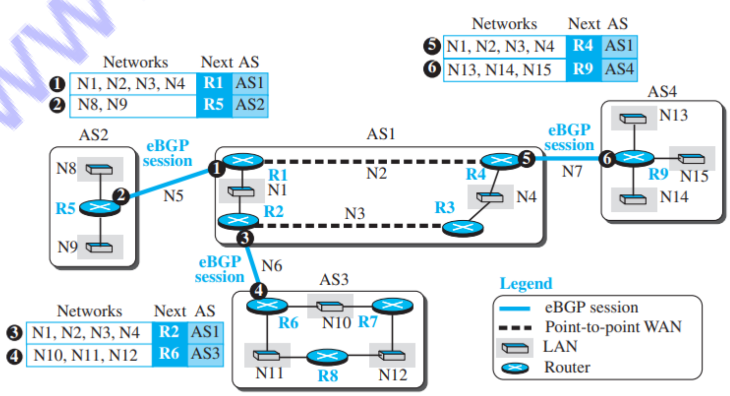
- The Border Gateway Protocol version 4 (BGP4) is the only interdomain routing protocol used in the Internet today. BGP4 is based on the path-vector algorithm.
- BGP, and in particular BGP4, is a complex protocol. AS2, AS3, and AS4 are stub autonomous systems; AS1 is a transient one.
- Each router in each AS knows how to reach a network that is in its own AS, but it does not know how to reach a network in another AS.



- The BGP may be external BGP (eBGP) or internal BGP (iBGP)

<u>**Operation of External BGP (eBGP)**</u>

- When the software is installed on two routers, they try to create a TCP connection using the well-known port 179.
- A pair of client and server processes continuously communicates with each other to exchange messages.
- The two routers that run the BGP processes are called BGP peers or BGP speakers.

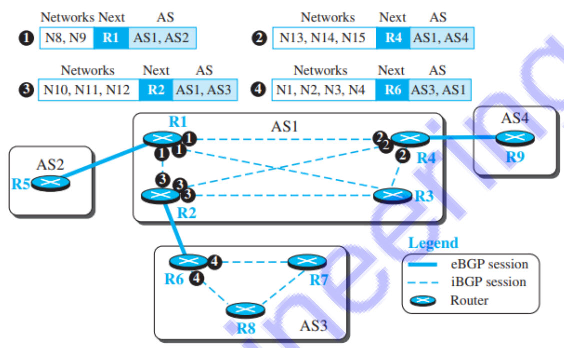


- The eBGP variation of BGP allows two physically connected border routers in two different ASs to form pairs of eBGP speakers and exchange messages.

- However, there is a need for a logical TCP connection to be created over the physical connection to make the exchange of information possible.
- Each logical connection in BGP parlance is referred to as a session.

<u>Operation of Internal BGP (iBGP):</u>

- The iBGP protocol is similar to the eBGP protocol in that it uses the service of TCP on the well-known port 179, but it creates a session between any possible pair of routers inside an autonomous system. However, some points should be made clear.
- First, if an AS has only one router, there cannot be an iBGP session. For example, we cannot create an iBGP session inside AS2 or AS4 in our internet.
- Second, if there are n routers in an autonomous system, there should be [n × (n − 1) / 2] iBGP sessions in that autonomous system to prevent loops in the system.
- Each router needs to advertise its own reachability to the peer in the session instead of flooding what it receives from another peer in another session



- The first message (numbered 1) is sent by R1 announcing that networks N8 and N9 are reachable through the path AS1-AS2, but the next router is R1.
- This message is sent, through separate sessions, to R2, R3, and R4. Routers R2, R4, and R6 do the same thing but send different messages to different destinations.
- The interesting point is that, at this stage, R3, R7, and R8 create sessions with their peers, but they actually have no message to send.
- After R1 receives the update message from R2, it combines the reachability information about AS3 with the reachability information it already knows about AS1 and sends a new update message to R5.
- Now R5 knows how to reach networks in AS1 and AS3. The process continues when R1 receives the update message from R4

Networks	Next	Path
N8, N9	R5	AS1, AS2
N10, N11, N12	R2	AS1, AS3
N13, N14, N15	R4	AS1, AS4

Path table for R1

Networks	Next	Path
N8, N9	R1	AS1, AS2
N10, N11, N12	R6	AS1, AS3
N13, N14, N15	R1	AS1, AS4

Path table for R2

Networks	Next	Path
N8, N9	R2	AS1, AS2
N10, N11, N12	R2	AS1, AS3
N13, N14, N15	R4	AS1, AS4

Path table for R3

Networks	Next	Path
N8, N9	R1	AS1, AS2
N10, N11, N12	R1	AS1, AS3
N13, N14, N15	R9	AS1, AS4

Path table for R4

Networks	Next	Path
N1, N2, N3, N4	R1	AS2, AS1
N10, N11, N12	R1	AS2, AS1, AS3
N13, N14, N15	R1	AS2, AS1, AS4

Path table for R5

Networks	Next	Path
N1, N2, N3, N4	R2	AS3, AS1
N8, N9	R2	AS3, AS1, AS2
N13, N14, N15	R2	AS3, AS1, AS4

Path table for R6

Networks	Next	Path
N1, N2, N3, N4	R6	AS3, AS1
N8, N9	R6	AS3, AS1, AS2
N13, N14, N15	R6	AS3, AS1, AS4

Path table for R7

Networks	Next	Path
N1, N2, N3, N4	R6	AS3, AS1
N8, N9	R6	AS3, AS1, AS2
N13, N14, N15	R6	AS3, AS1, AS4

Path table for R8

Networks	Next	Path
N1, N2, N3, N4	R4	AS4, AS1
N8, N9	R4	AS4, AS1, AS2
N10, N11, N12	R4	AS4, AS1, AS3

Path table for R9

IPv6 ADDRESSING:

- The main reason for migration from IPv4 to IPv6 is the small size of the address space in IPv4.
- An IPv6 address is 128 bits or 16 bytes (octets) long, four times the address length in IPv4
- A computer normally stores the address in binary, but it is clear that 128 bits cannot easily be handled by humans
- Binary notation is used when the addresses are stored in a computer.
- The colon hexadecimal notation (or colon hex for short) divides the address into eight sections, each made of four hexadecimal digits separated by colons

Abbreviation

- Although an IPv6 address, even in hexadecimal format, is very long, many of the digits are zeros.
- The leading zeros of a section can be omitted. Using this form of abbreviation, 0074 can be written as 74, 000F as F, and 0000 as 0. Note that 3210 cannot be abbreviated.
- Further abbreviation, often called zero compression, can be applied to colon hex notation if there are consecutive sections consisting of zeros only. We can remove all the zeros and replace them with a double semicolon.

FDEC:0:0:0:0:BBFF:0:FFFF ⟶ FDEC::BBFF:0:FFFF

- IPv6 uses hierarchical addressing. The address space of IPv6 contains 2^{128} addresses. This address space is 296 times the IPv4 address—definitely no address depletion

<u>Address Types</u>

In IPv6, a destination address can belong to one of three categories: unicast, anycast, and multicast.

<u>Unicast Address</u>

A unicast address defines a single interface (computer or router).

The packet sent to a unicast address will be routed to the intended recipient.

<u>Anycast Address</u>

- An anycast address defines a group of computers that all share a single address.
- A packet with an anycast address is delivered to only one member of the group, the most reachable one. An anycast communication is used, for example, when there are several servers that can respond to an inquiry.
- The request is sent to the one that is most reachable. The hardware and software generate only one copy of the request; the copy reaches only one of the servers.
- IPv6 does not designate a block for anycasting; the addresses are assigned from the unicast block.

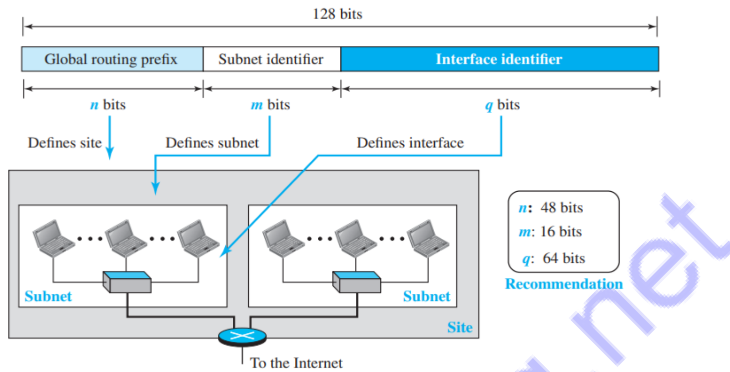
<u>Multicast Address</u>

- A multicast address also defines a group of computers. However, there is a difference between anycasting and multicasting.
- In anycasting, only one copy of the packet is sent to one of the members of the group; in multicasting each member of the group receives a copy.

<u>Global Unicast Addresses</u>

  - The block in the address space that is used for unicast (one-to-one) communication between two hosts in the Internet is called the global unicast address block.
  - CIDR for the block is 2000::/3, which means that the three leftmost bits are the same for all addresses in this block (001).

- The size of this block is 2125 bits, which is more than enough for Internet expansion for many years to come.
- An address in this block is divided into three parts: global routing prefix (n bits), subnet identifier (m bits), and interface identifier (q bits)



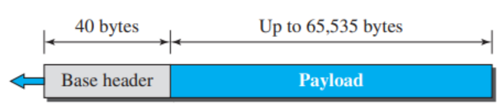
IPv6 PROTOCOL:

The change of the IPv6 address size requires the change in the IPv4 packet format.

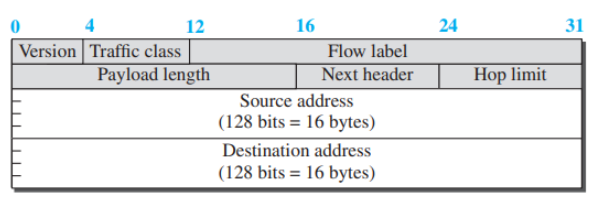
- **Better header format**. IPv6 uses a new header format in which options are separated from the base header and inserted, when needed, between the base header and the data.
- **New options.** IPv6 has new options to allow for additional functionalities.
- **Allowance for extension.** IPv6 is designed to allow the extension of the protocol if required by new technologies or applications.
- **Support for resource allocation.** In IPv6, the type-of-service field has been removed, but two new fields, traffic class and flow label, have been added to enable the source to request special handling of the packet. This mechanism can be used to support traffic such as real-time audio and video.
- **Support for more security.** The encryption and authentication options in IPv6 provide confidentiality and integrity of the packet.

Packet Format:

Each packet is composed of a base header followed by the payload. The base header occupies 40 bytes, whereas payload can be up to 65,535 bytes of information.



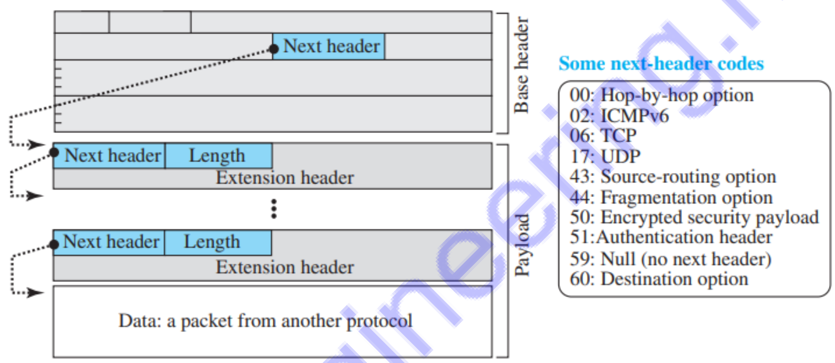
a. IPv6 packet



b. Base header

- **Version.** The 4-bit version field defines the version number of the IP. For IPv6, the value is 6.
- **Traffic class.** The 8-bit traffic class field is used to distinguish different payloads with different delivery requirements. It replaces the type-of-service field in IPv4.

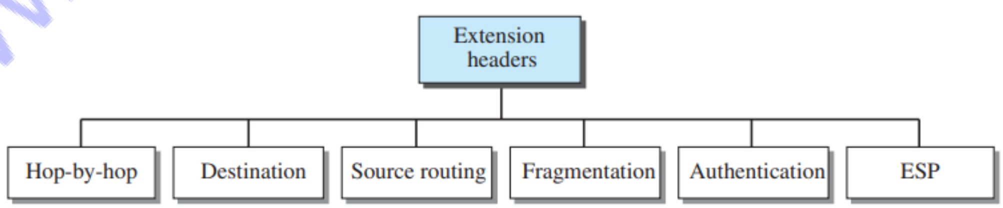
- **Flow label**. The flow label is a 20-bit field that is designed to provide special handling for a particular flow of data.
- **Payload length.** The 2-byte payload length field defines the length of the IP datagram excluding the header.
- **Next header**. The next header is an 8-bit field defining the type of the first extension header or the type of the data that follows the base header in the datagram.
- **Hop limit**. The 8-bit hop limit field serves the same purpose as the TTL field in IPv4.
- **Source and destination addresses.** The source address field is a 16-byte (128-bit) Internet address that identifies the original source of the datagram. The destination address field is a 16-byte (128-bit) Internet address that identifies the destination of the datagram.
- **Payload.** The payload field in IPv6 differ from IPv4



The payload in IPv6 means a combination of zero or more extension headers (options) followed by the data from other protocols.

Extension Header

- An IPv6 packet is made of a base header and some extension headers.
- The length of the base header is fixed at 40 bytes. However, to give more functionality to the IP datagram, the base header can be followed by up to six extension headers.
- These are hop-by-hop option, source routing, fragmentation, authentication, encrypted security payload, and destination option.



Hop-by-Hop - The hop-by-hop option is used when the source needs to pass information to all routers visited by the datagram

Destination Option The destination option is used when the source needs to pass information to the destination only

Source Routing The source routing extension header combines the concepts of the strict source route and the loose source route options of IPv4.

Authentication The authentication extension header has a dual purpose: it validates the message sender and ensures the integrity of data

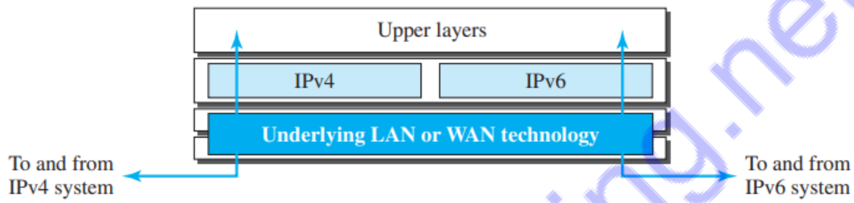
Encrypted Security Payload The encrypted security payload (ESP) is an extension that provides confidentiality and guards against eavesdropping.

TRANSITION FROM IPv4 TO IPv6:

There are three strategies have been devised for transition: dual stack, tunneling, and header translation

Dual Stack:

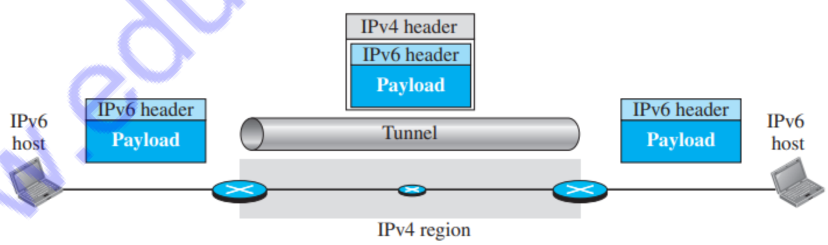
- It is recommended that all hosts, before migrating completely to version 6, have a dual stack of protocols during the transition. In other words, a station must run IPv4 and IPv6 simultaneously until all the Internet uses IPv6



- To determine which version to use when sending a packet to a destination, the source host queries the DNS. If the DNS returns an IPv4 address, the source host sends an IPv4 packet. If the DNS returns an IPv6 address, the source host sends an IPv6 packet.

Tunneling

- Tunneling is a strategy used when two computers using IPv6 want to communicate with each other and the packet must pass through a region that uses IPv4.
- To pass through this region, the packet must have an IPv4 address. So the IPv6 packet is encapsulated in an IPv4 packet when it enters the region, and it leaves its capsule when it exits the region. It seems as if the IPv6 packet enters a tunnel at one end and emerges at the other end.



Header

- Translation Header translation is necessary when the majority of the Internet has moved to IPv6 but some systems still use IPv4.
- The sender wants to use IPv6, but the receiver does not understand IPv6. Tunneling does not work in this situation because the packet must be in the IPv4 format to be understood by the receiver.