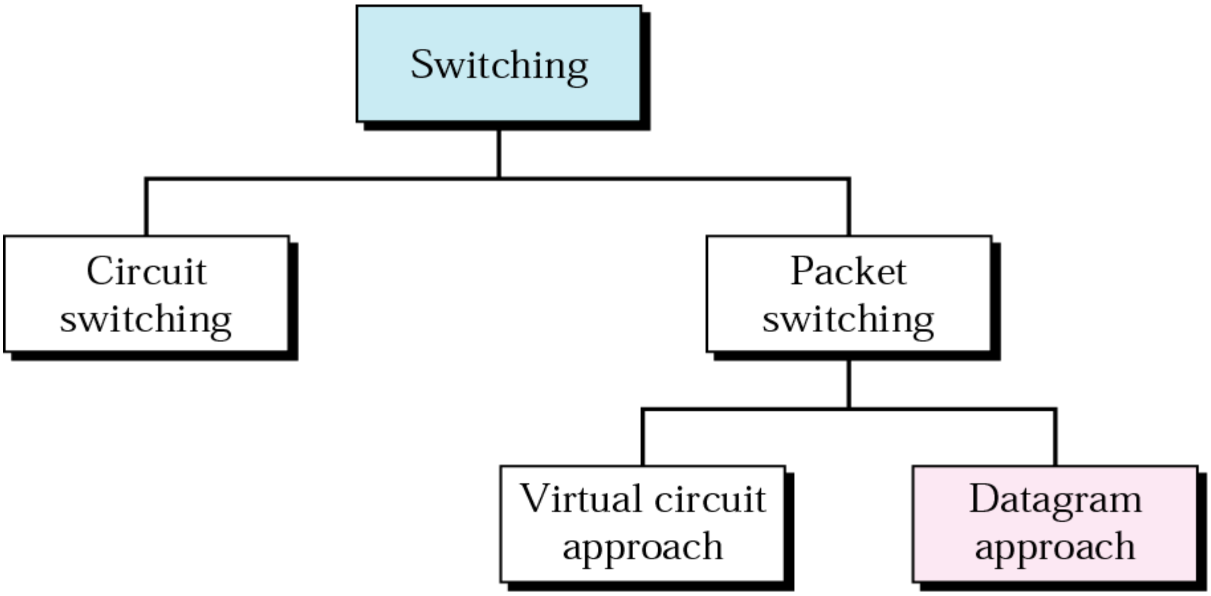
Unit 3

To route the packet, we need any device like switch and router



Switching:

It is used to route the packet to the next router in the internet

It can be divided into two switching

1) Circuit Switching
2) Packet Switching

Circuit Switching:

Consider the following step to establish the connection between sender and receiver

We have the dedicated link between source to destination

The sending data can transmit sequentially one by one which is send by the sender.

And receiver also can receive the packet the same order which is send by the sender.

Packet Switching

It can be divided into two types

I)Virtual circuit
2)Datagram approach

Virtual Circuit:

1) Establish the connection between two nodes
2) Communicate the data
3) Disconnect the link between send to receiver.

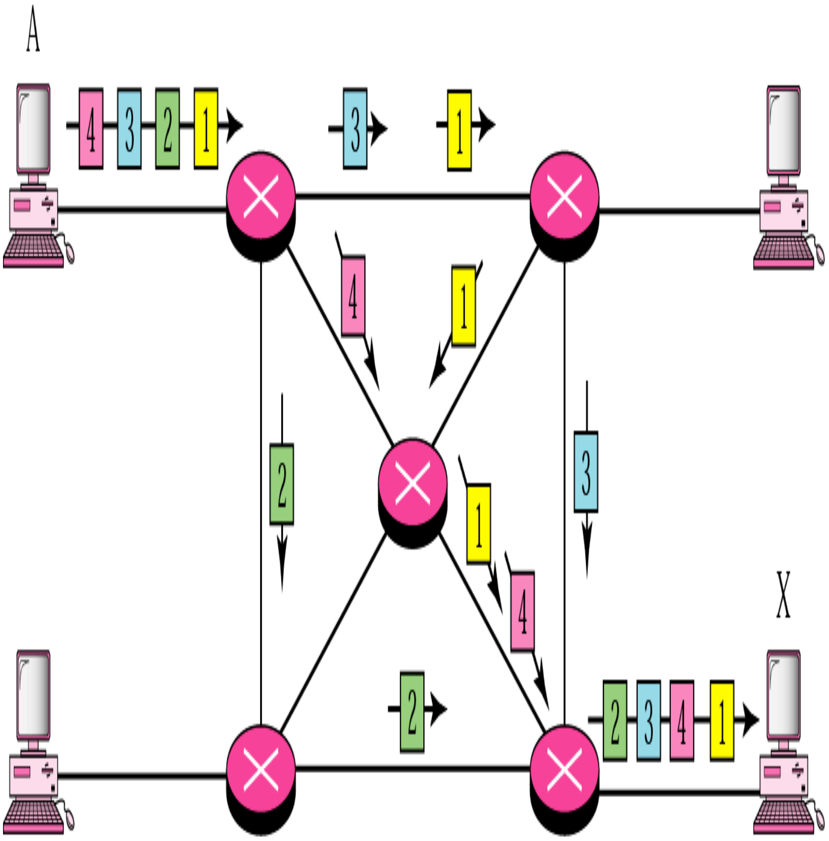
Circuit switching	**Datagram Packet switching**	**VC Packet Switching**

Dynamically create the link by using virtual circuit

Datagram approach:

There is no dedicated link between source to destination, all the packet can transmit the via different route.

At the receiving end , rearrange all the packets in order to get the original data.



• Dedicated transmission path • Continuous transmission • Path stays fixed for entire transmission • Call setup delay • No queueing delay • Busy signal overloaded network • Fixed bandwidth for each circuit • No overhead after call setup (connection overhead setup)	• No dedicated transmission path • Transmission of packets. • Route of each packet is independent • No setup delay • Queueing delays at switches • Delays increase in overloaded networks • Bandwidth is shared by all packets • Overhead in each packet	• No dedicated transmission path • Transmission of packets • Path stays fixed for entire transmission • Call setup delay • Queueing delays at switches • Delays increase in overloaded networks • Bandwidth is shared by all packets • Overhead in each packet

Routing:

The responsibilities of routing is route the packets and also with shortest distance, before route the packet to construct the routing table using RIP and OSPF algorithm based on that routing table only the router has to decide everything.

Example:

Routing table for host A

Destination	Route
Host B	R1, R2, Host B

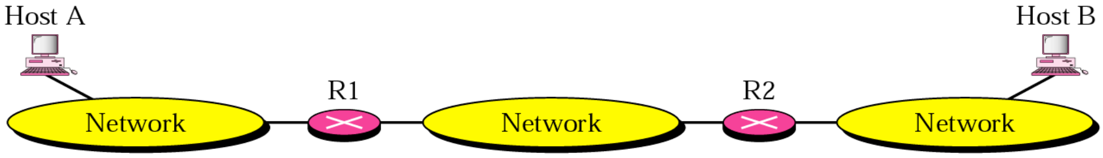
Routing table for R1

Destination	Route
Host B	R2, Host B

Routing table for R2

Destination	Route
Host B	Host B

a. Routing tables based on route



Routing table for host A

Destination	Next Hop
Host B	R1

Routing table for R1

Destination	Next Hop
Host B	R2

Routing table for R2

Destination	Next Hop
Host B	—

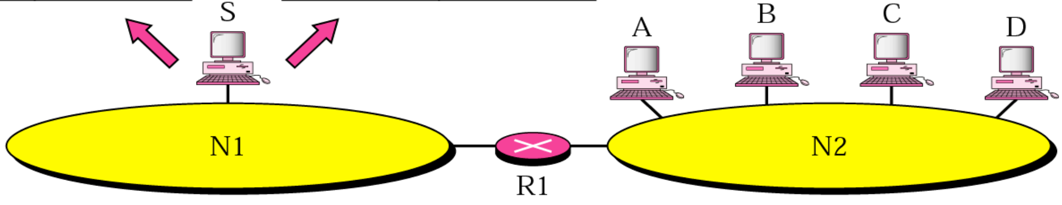
b. Routing tables based on next hop

Routing table for host S based on host-specific routing

Destination	Next Hop
A	R1
B	R1
C	R1
D	R1

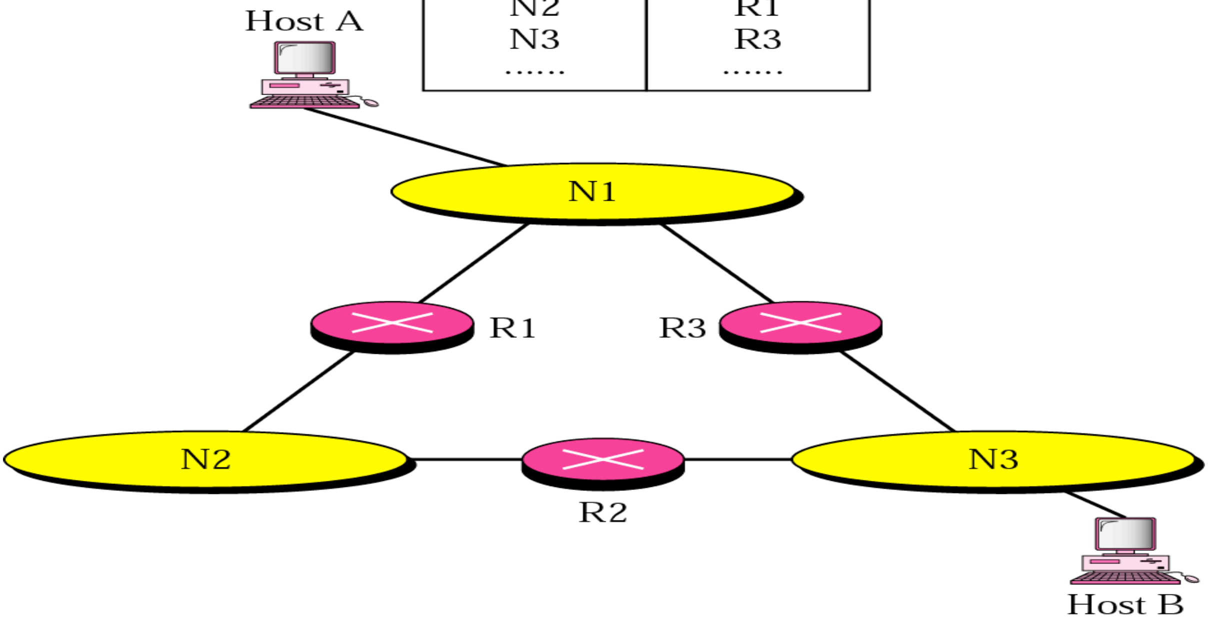
Routing table for host S based on network-specific routing

Destination	Next Hop
N2	R1



Routing table for host A

Destination	Next Hop
Host B	R3
N2	R1
N3	R3
......	......

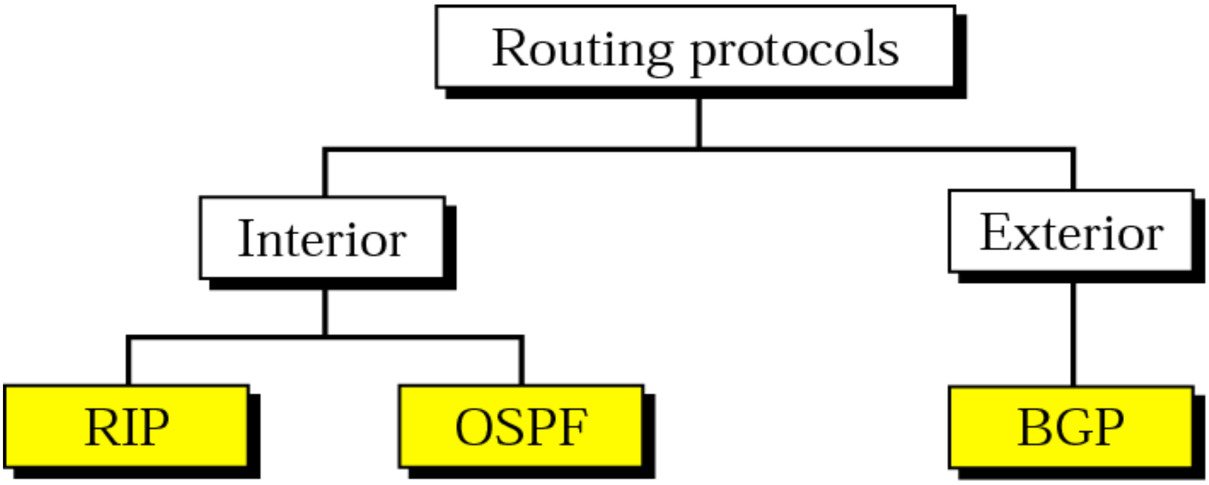


Need for Routing Protocols

- Routing protocols have been created in response to the demand for dynamic routing tables

Metric

- Cost assigned for passing through a network
- Total metric for a particular route is equal the sum of the metrics of networks that comprise the route
- Router chooses the route with shortest / smallest metric
- Metric assigned to each network depends on the type of protocol
  - Routing Information Protocol treat all networks as equals (Cost = one hop count)
  - Open Shortest Path First allows administrator to assign a cost for passing through the network
- RIP (Routing Information Protocol)
- OSPF (Open Shortest Path First)
- BGP (Border Gateway Protocol)
- Internet is so large that one routing protocol cannot handle the task of updating the routing tables of all routers
- Internet is divided into autonomous systems
- Autonomous systems (AS) is a group of networks and routers under the authority of single administration
- Routing inside an autonomous system - interior routing
- Routing between autonomous systems - exterior routing



Routing Information Protocol (RIP)

- Interior routing protocol used inside an autonomous system
- Based on distance vector routing which uses Bellman-Ford algorithm for calculating the routing tables

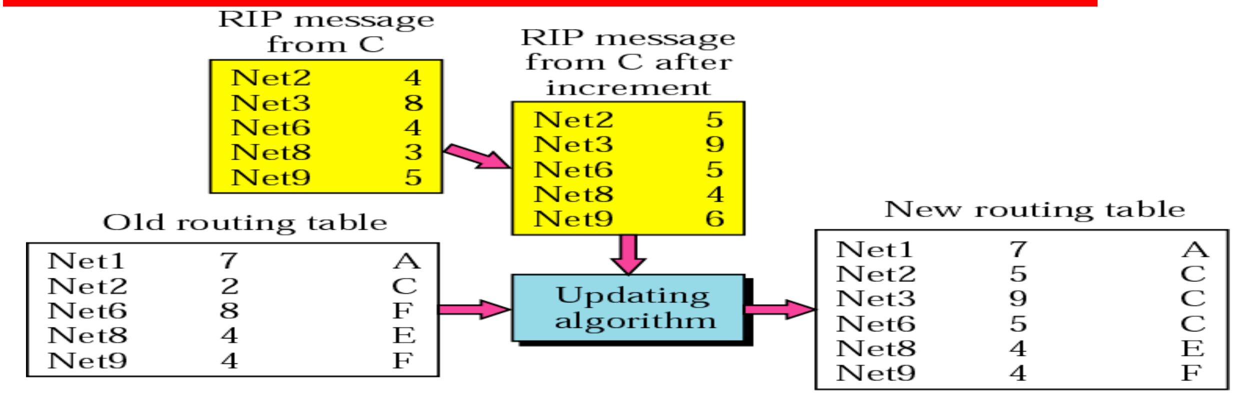
Distance Vector Routing

- Each router periodically shares its knowledge about the entire internet with its neighbours
  - Sharing the knowledge about the entire autonomous system
  - Sharing only with neighbours
  - Sharing at regular intervals

Routing Table

Every router keeps a routing table that has one entry for each destination network of which the router is aware

Destination	Hop Count	Next Router	Other information
163.5.0.0	7	172.6.23.4	
197.5.13.0	5	176.3.6.17	
189.45.0.0	4	200.5.1.6	
115.0.0.0	6	131.4.7.19	



RIP message from C

Net2	4
Net3	8
Net6	4
Net8	3
Net9	5

RIP message from C after increment

Net2	5
Net3	9
Net6	5
Net8	4
Net9	6

Old routing table

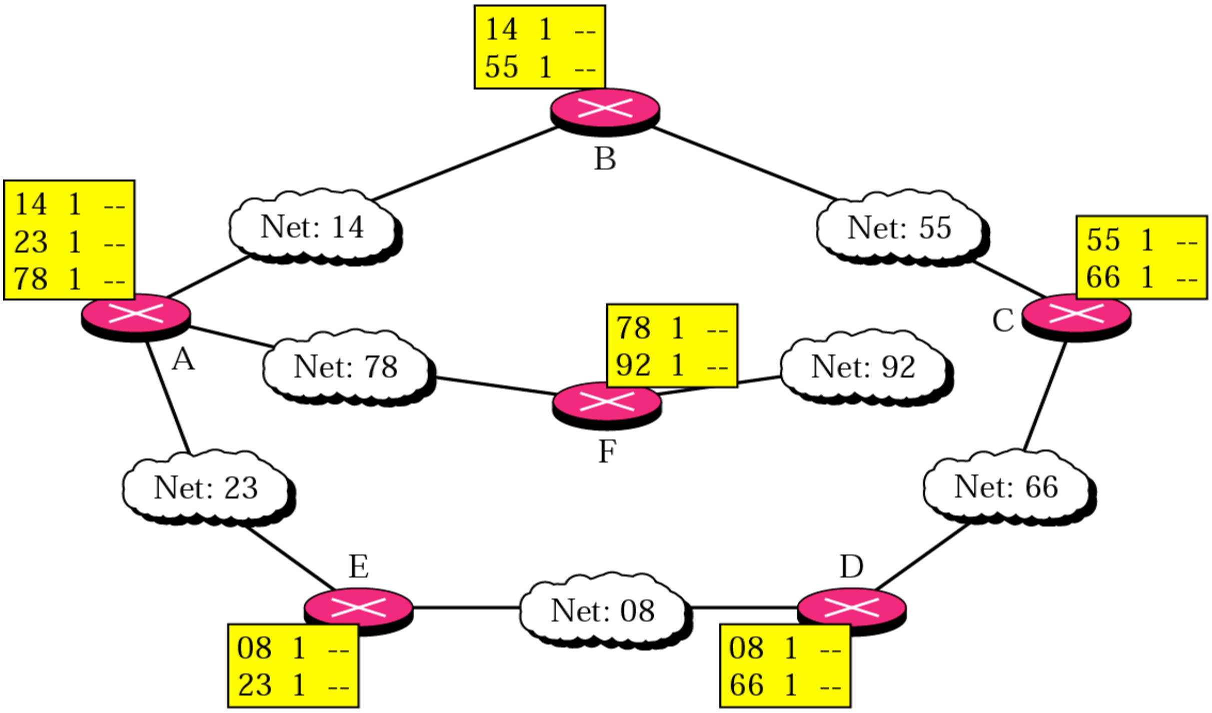
Net1	7	A
Net2	2	C
Net6	8	F
Net8	4	E
Net9	4	F

New routing table

Net1	7	A
Net2	5	C
Net3	9	C
Net6	5	C
Net8	4	E
Net9	4	F

Net1: No news, do not change
Net2: Same next hop, replace
Net3: A new router, add
Net6: Different next hop, new hop count smaller, replace
Net8: Different next hop, new hop count the same, do not change
Net9: Different next hop, new hop count larger, do not change

Initially table contains only directly attached networks and the hop counts- Next hop field which identifies the next router is empty



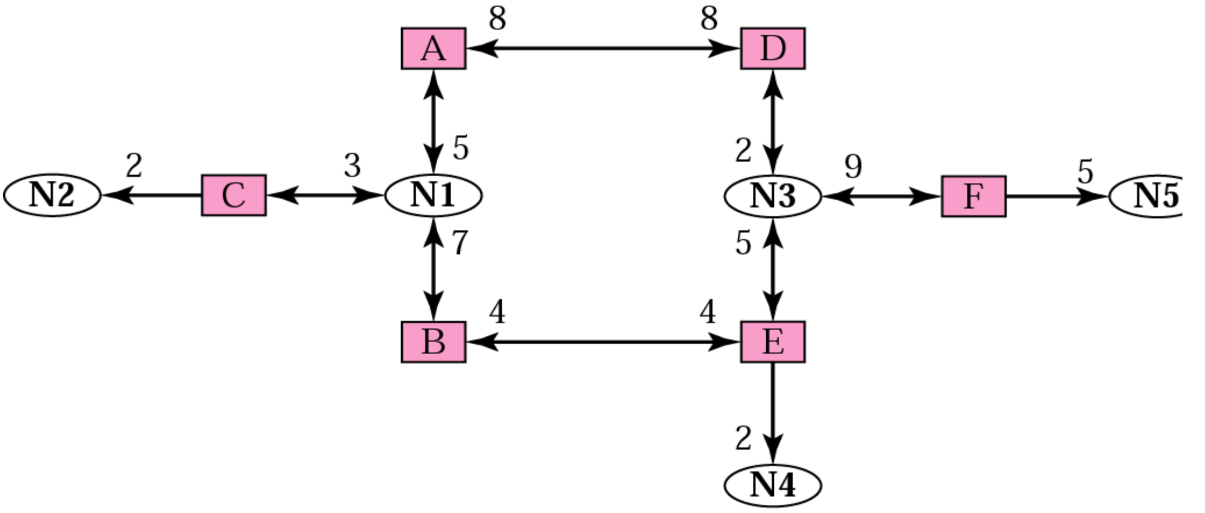
Open Shortest Path First (OSPF)

- To handle routing efficiently and in a timely manner, OSPF divides an autonomous system into areas
- Area:
  - Collection of networks, hosts, and routers all contained within an autonomous system
  - All networks inside an area must be connected

Link State Routing

- Process by which each router shares its knowledge about its neighbourhood with every router in the area
  - Sharing knowledge about the neighbourhood
  - Sharing with every other router – flooding
  - Sharing when there is a change – results in lower internet traffic than that required by distance vector

routing



Link State Database

- Every router in an area receives the router link and network link LSAs from every other router and forms a link state database
- Every router in the same area has the same link state database
- Link state database – tabular representation of the topology of the internet inside an area – shows relationship between each router and its neighbours including the metrics

Dijkstra Algorithm

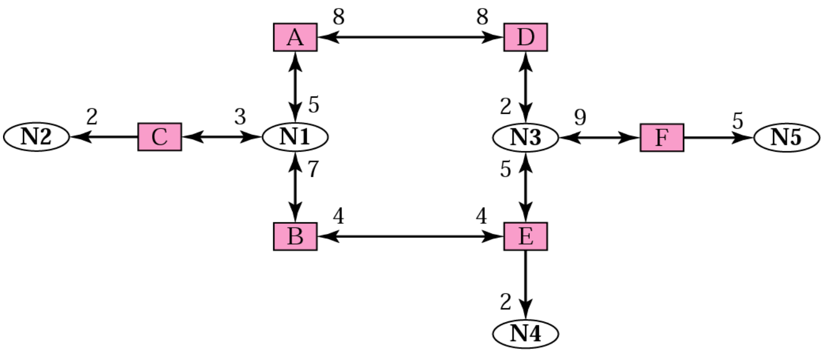
- To calculate routing table each router applies Dijkstra algorithm to its link state database
- Dijkstra algorithm calculates the shortest path between two points on a network using a graph made up of nodes and edges
- Algorithm divides the nodes into two sets: tentative and permanent. It chooses nodes, makes them tentative, examines them and if they pass the criteria, makes them permanent

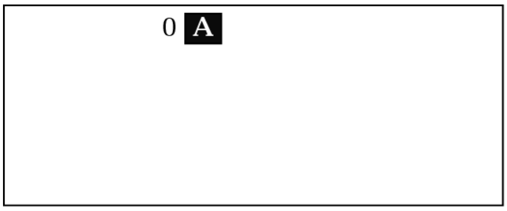
Routing Table

- Each router uses the shortest-path tree method to construct its routing table
- Routing table shows cost of reaching each network in the area
- To find the cost of reaching networks outside of the area, routers use the summary link to network, the summary link to boundary router and the external link advertisements

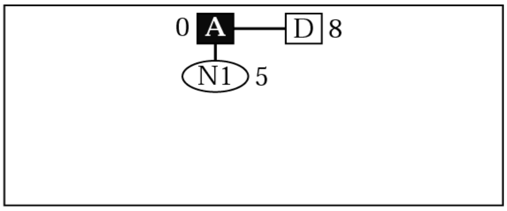
Dijkstra Algorithm

1. Start with the local node (router): the root of the tree.
2. Assign a cost of 0 to this node and make it the first permanent node.
3. Examine each neighbor node of the node that was the last permanent node.
4. Assign a cumulative cost to each node and make it tentative.
5. Among the list of tentative nodes
    1. Find the node with the smallest cumulative cost and make it permanent.
    2. If a node can be reached from more than one direction
        1. Select the direction with the shortest cumulative cost.
6. Repeat steps 3 to 5 until every node becomes permanent.

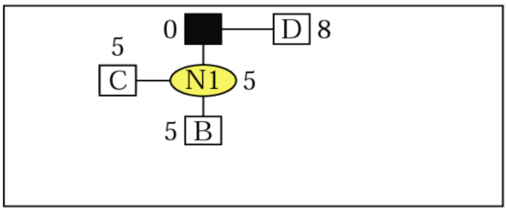




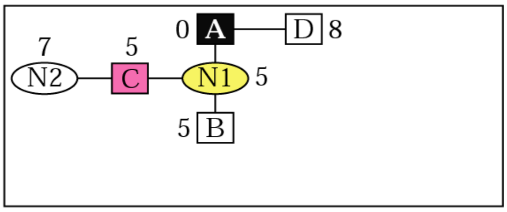
Start with A



Make A permanent, add its neighbors

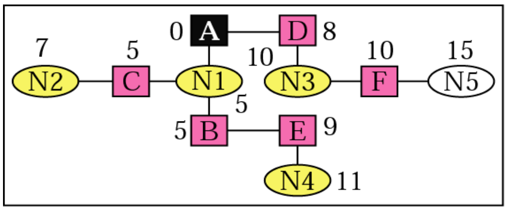


Make N1 permanent, add its neighbors

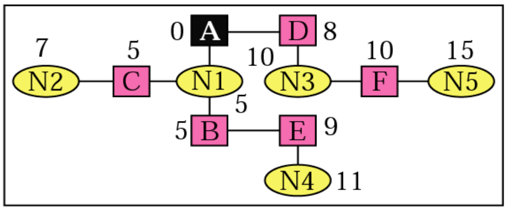


Make C permanent, add its neighbors

⋮



Make N4 permanent



Make N5 permanent

Routing Table

- Each router uses the shortest-path tree method to construct its routing table
- Routing table shows cost of reaching each network in the area
- To find the cost of reaching networks outside of the area, routers use the summary link to network, the summary link to boundary router and the external link advertisements

Routing Table for Router A

Network	Cost	Next Router	Other Information
N1	5	C	
N2	7	D	
N3	10	B	
N4	11	D	
N5	15	C	

BGP(Border Gateway Protocol):

Two major inter domain routing protocols in the recent history of the Internet.

Exterior Gateway Protocol (EGP):

EGP had a number of limitations: it has topological constraints i.e., it adapts to tree topology, where Autonomous systems are connected only as parent and child and not as peers.

The replacement for EGP: **Border Gateway Protocol.** BGP assumes the Internet as an arbitrarily interconnected set of ASs.

Unlike the simple tree-structured Internet, today's Internet consists of an **interconnection** of multiple backbone networks and sites that are connected to each in arbitrary ways.

Some large corporations connect directly to 1/more backbones, while others connect to smaller, non-backbone providers.

Many service providers exist mainly to provide service to consumers. These providers arrange to interconnect with each other at a single "peering point".

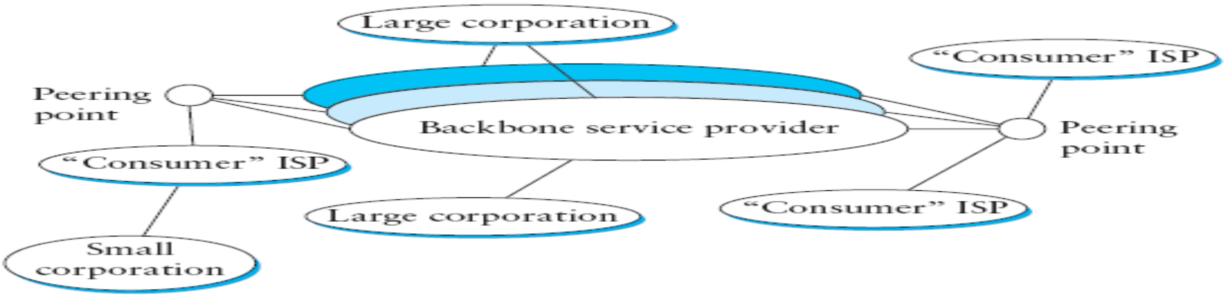
We can classify ASs into **three** types:

- ■ ***Stub AS***: an AS that has only a single connection to one other AS; such an AS will only carry local traffic with in that AS. The small corporation in figure is an eg., of a stub AS.

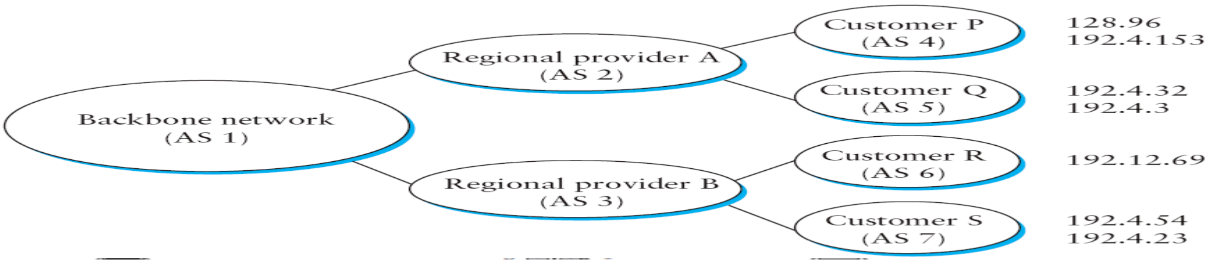
- ■ **Multihomed AS**: an AS that has connections to more than one other AS but that refuses to carry transit traffic; for example, the large corporation at the top
- ■ ***Transit AS***: an AS that has connections to more than one other AS and that is designed to carry both transit and local traffic, such as the backbone providers.

The **goal** is to find **any** path to the intended destination that is loop-free. That is, we are more concerned with reachability than optimality.

- The paths must be compliant with the policies of various ASs along the path
- There are a few reasons why interdomain routing is hard. **The first** is simply a matter of scale. An Internet backbone router must be able to forward any packet destined anywhere in the Internet. Maintaining info of all routes at the backbone router is difficult.



- The second challenge in interdomain routing arises from the autonomous nature of the domains. Note that each domain may run its own interior routing protocols and use any scheme it chooses to assign metrics to paths.
- The third challenge involves the issue of trust. Provider A might be unwilling to believe certain advertisements from provider B for fear that provider B will advertise erroneous routing information.
- When configuring BGP, the administrator of each AS picks at least one node to be a **"BGP speaker"** , which is essentially a spokesperson for the entire AS. That BGP speaker establishes **BGP sessions** to other BGP speakers in other ASs
- In addition to the BGP speakers, the AS has one or more border **"gateways",** which need not be the same as the speakers. **The border gateways** are the routers through which packets enter and leave the AS.
- BGP differs from Distance Vector and Link State routings as it advertises **complete paths** as an **enumerated list of ASs** to reach a particular network.



- In the above fig, providers are transit n/w and customer network are stubs.
- A BGP speaker for the AS of provider A (AS2) advertises its reachability info to customers P and Q i.e., it advertises n/ws 128.96, 192.4.153, 192.4.32 & 192.4.3 can be reached from it (AS2).
- The backbone n/w on receiving this advertise, advertises the networks reachable thru AS2 can be reached along the path (AS1, AS2). Similar reasoning applies for networks R and S i.e., their reachability is (AS1, AS3).
- An important job of BGP is to prevent the establishment of looping paths.

- Consider 3 interconnected AS1, AS2 & AS3. AS1 learns that it can reach network 10.0.1 through AS2, so it advertises this fact to AS3, who in turn advertises it back to AS2.
- AS2 could now decide that AS3 was the place to send packets destined for 10.0.1; AS3 sends them to AS1; AS1 sends them back to AS2; and they would loop forever.
- Such loops are prevented in BGP by carrying the complete AS path in the routing messages.
- Here ad received by AS2 from AS3 contains an AS of path AS3, AS1, AS2. AS1 finds itself in this path and concludes that it is not a useful path to use.
- For this to work, the numbering used in ASs should be unique, to prevent looping.
- Also an AS will only advertise routes that it considers good enough for itself. If a BGP speaker has a choice of several different routes to a destination, it will choose the best one according to its own local policies, and then that will be the route it advertises.
- BGP speakers need to be able to cancel previously advertised paths if a critical link or node on a path goes down. This is done with a form of negative advertisement known as a withdrawn route**.** Both positive and negative info is carried in a BGO update message.

IPV6(internet Protocol Version 6)

Internet Protocol version 6 (IPv6) is the latest revision of the Internet Protocol (IP) and the first version of the protocol to be widely deployed. IPv6 was developed by the Internet Engineering Task Force (IETF) to deal with the long-anticipated problem of IPv4 address exhaustion. This tutorial will help you in understanding IPv6 and its associated terminologies along with appropriate references and examples.

Feature:

The successor of IPv4 is not designed to be backward compatible. Trying to keep the basic functionalities of IP addressing, IPv6 is redesigned entirely. It offers the following features:

- **Larger Address Space**

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

- **Simplified Header**

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

- **End-to-end Connectivity**

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

- **Auto-configuration**

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

- **Faster Forwarding/Routing**

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

- **IPSec**

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

- **No Broadcast**

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

- **Anycast Support**

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

- **Mobility**

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

- **Enhanced Priority Support**

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

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

- **Smooth Transition**

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

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

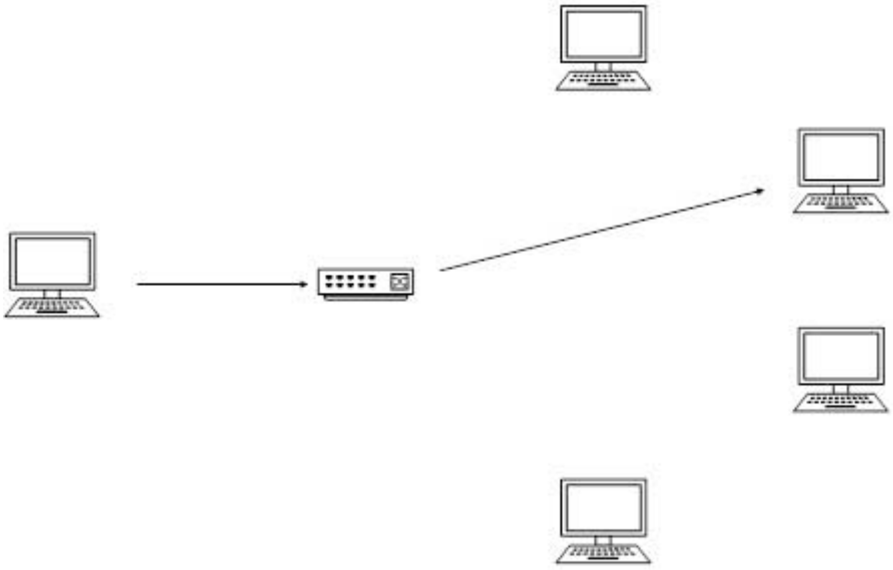
- **Extensibility**

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

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

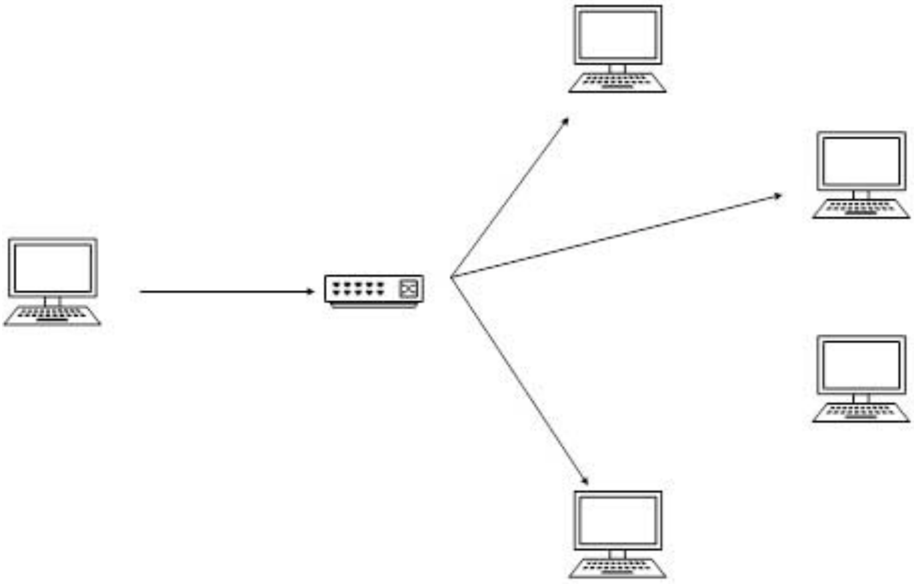
Unicast

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



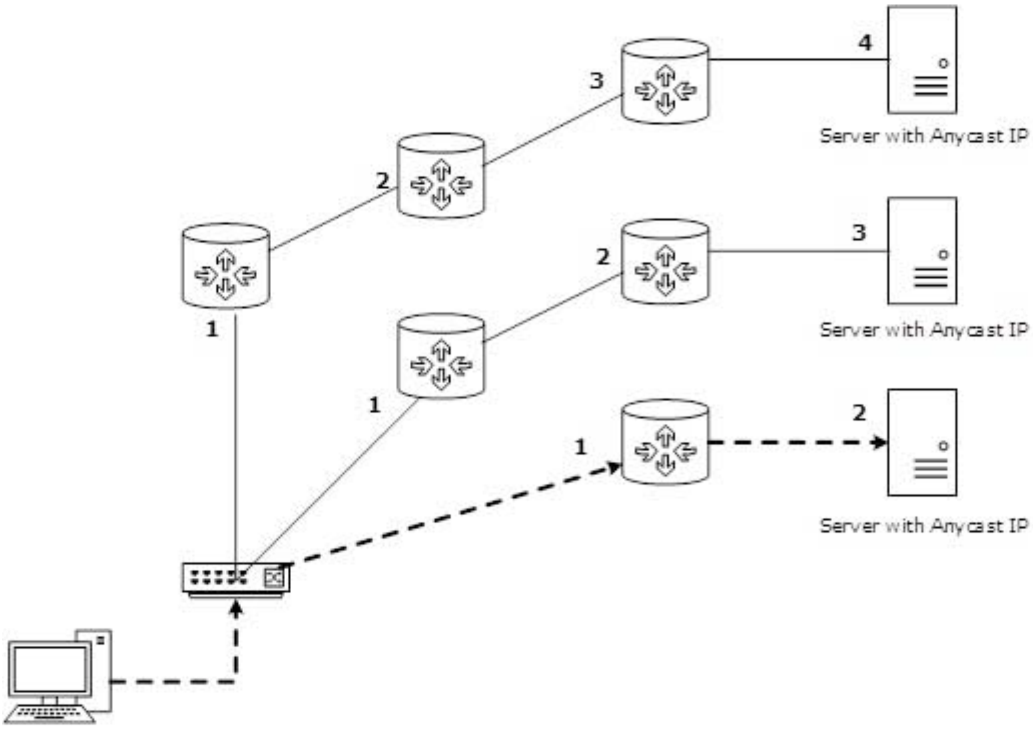
Multicast

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



Anycast

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

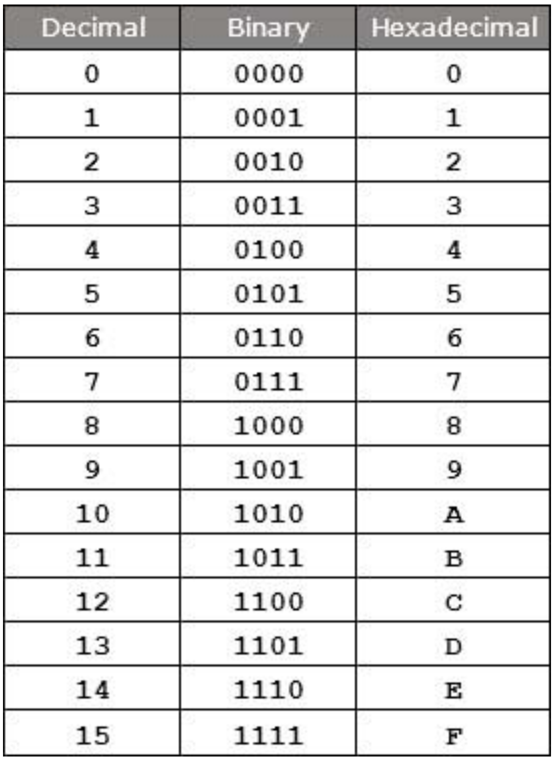


Let’s take an example of TutorialPoints.com Web Servers, located in all continents. Assume that all the Web Servers are assigned a single IPv6 Anycast IP Address. Now when a user from Europe wants to reach TutorialsPoint.com the DNS points to the server that is physically located in Europe itself. If a user from India tries to reach Tutorialspoint.com, the DNS will then point to the Web Server physically located in Asia. Nearest or Closest terms are used in terms of Routing Cost.

In the above picture, when a client computer tries to reach a server, the request is forwarded to the server with the lowest Routing Cost.

Hexadecimal Number System

Before introducing IPv6 Address format, we shall look into Hexadecimal Number System. Hexadecimal is a positional number system that uses radix (base) of 16. To represent the values in readable format, this system uses 0-9 symbols to represent values from zero to nine and A-F to represent values from ten to fifteen. Every digit in Hexadecimal can represent values from 0 to 15.



Decimal	Binary	Hexadecimal
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F

Address Structure

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

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

```
0010000000000001 0000000000000000 0011001000111000 1101111111100001
0000000001100011 0000000000000000 0000000000000000 1111111011111011
```

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

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

Even after converting into Hexadecimal format, IPv6 address remains long. IPv6 provides some rules to shorten the address. The rules are as follows:

Rule.1: Discard leading Zero(es):

In Block 5, 0063, the leading two 0s can be omitted, such as (5th block):

```
2001:0000:3238:DFE1:63:0000:0000:FEFB
```

Rule.2: If two of more blocks contain consecutive zeroes, omit them all and replace with double colon sign ::, such as (6th and 7th block):

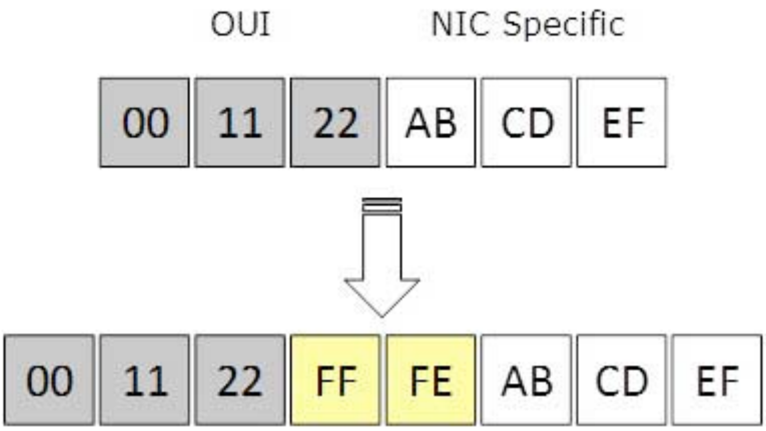
```
2001:0000:3238:DFE1:63::FEFB
```

Consecutive blocks of zeroes can be replaced only once by :: so if there are still blocks of zeroes in the address, they can be shrunk down to a single zero, such as (2nd block):

```
2001:0:3238:DFE1:63::FEFB
```

Interface ID

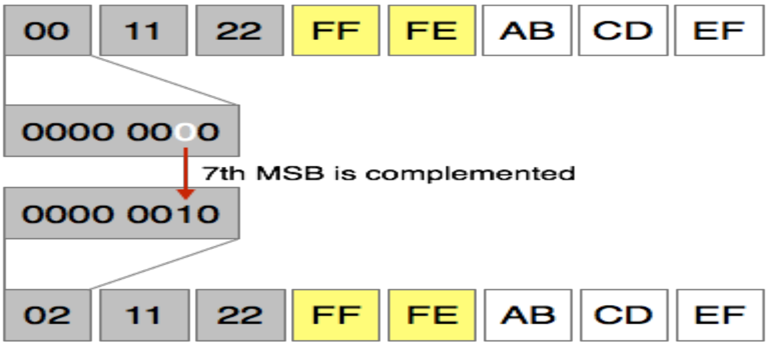
IPv6 has three different types of Unicast Address scheme. The second half of the address (last 64 bits) is always used for Interface ID. The MAC address of a system is composed of 48-bits and represented in Hexadecimal. MAC addresses are considered to be uniquely assigned worldwide. Interface ID takes advantage of this uniqueness of MAC addresses. A host can auto-configure its Interface ID by using IEEE's Extended Unique Identifier (EUI-64) format. First, a host divides its own MAC address into two 24-bits halves. Then 16-bit Hex value 0xFFFE is sandwiched into those two halves of MAC address, resulting in EUI-64 Interface ID.



[Image: EUI-64 Interface ID]

Conversion of EUI-64 ID into IPv6 Interface Identifier

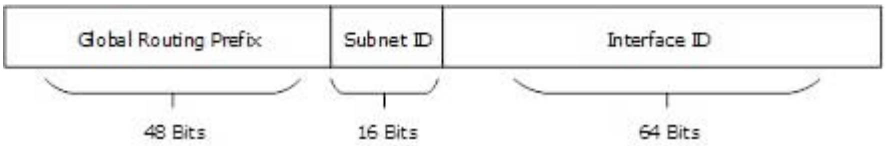
To convert EUI-64 ID into IPv6 Interface Identifier, the most significant 7th bit of EUI-64 ID is complemented. For example:



[Image: IPV6 Interface ID]

Global Unicast Address

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

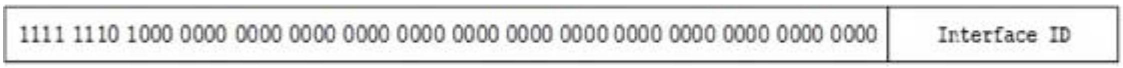


[*Image: Global Unicast Address*]

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

Link-Local Address

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

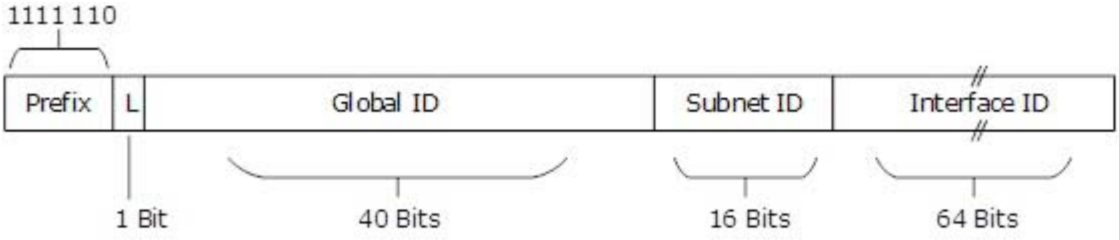


[*Image: Link-Local Address*]

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

Unique-Local Address

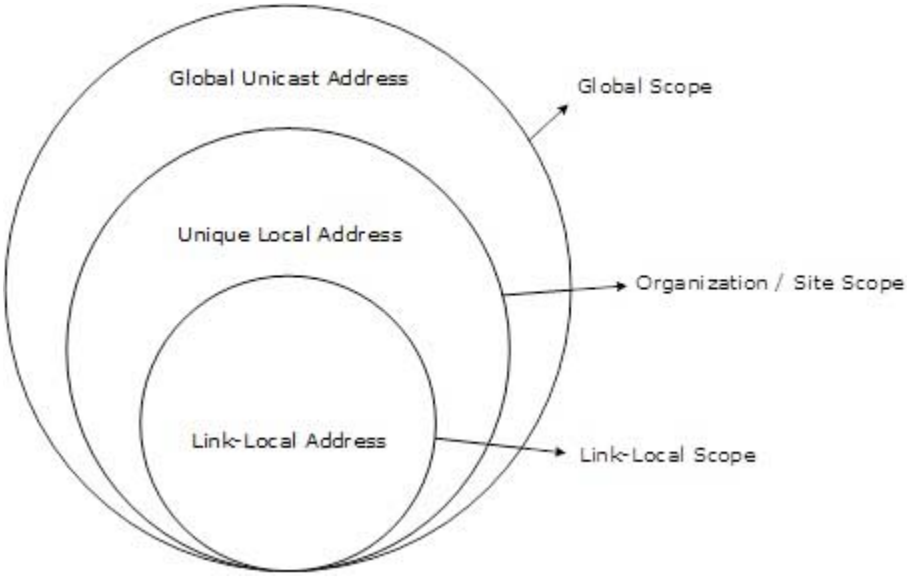
This type of IPv6 address is globally unique, but it should be used in local communication. The second half of this address contain Interface ID and the first half is divided among Prefix, Local Bit, Global ID and Subnet ID.



[*Image: Unique-Local Address*]

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

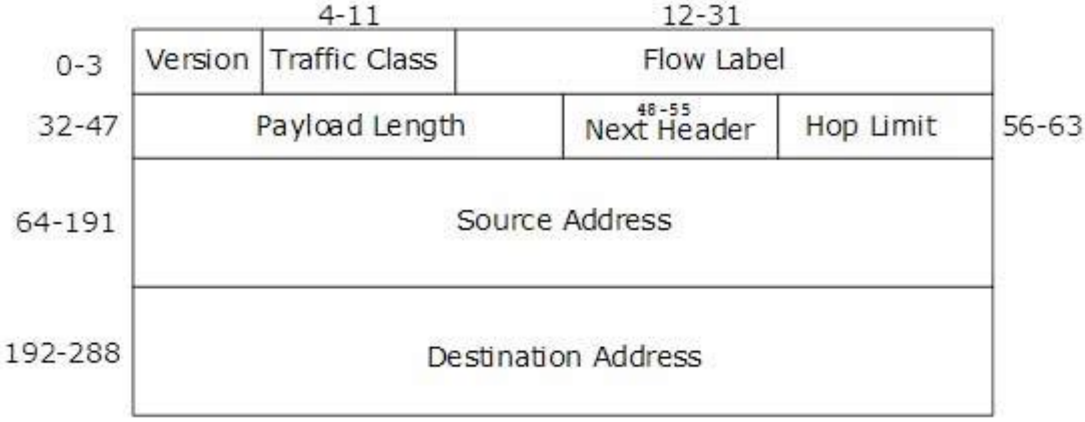
Scope of IPv6 Unicast Addresses:



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

The wonder of IPv6 lies in its header. An IPv6 address is 4 times larger than IPv4, but surprisingly, the header of an IPv6 address is only 2 times larger than that of IPv4. IPv6 headers have one Fixed Header and zero or more Optional (Extension) Headers. All the necessary information that is essential for a router is kept in the Fixed Header. The Extension Header contains optional information that helps routers to understand how to handle a packet/flow.

Fixed Header



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

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

	recipient of the packet.

Multicast:

Number of different algorithms that may potentially be employed by multicast routing protocols:

- Flooding

- Spanning Trees

- Reverse Path Broadcasting (RPB)

- Truncated Reverse Path Broadcasting (TRPB)

- Reverse Path Multicasting (RPM)

- Core-Based Trees

Multicast Routing algorithm

- Distance Vector Multicast Routing Protocol (DVMRP)

- Multicast OSPF (MOSPF)

- Protocol-Independent Multicast (PIM)

Introduction

There are three fundamental types of IPv4 addresses: unicast, broadcast, and multicast. A unicast address is designed to transmit a packet to a single destination. A broadcast address is used to send a datagram to an entire subnetwork. A multicast address is designed to enable the delivery of datagrams to a set of hosts that have been configured as members of a multicast group in various scattered subnetworks. Multicasting is not connection oriented.

A multicast datagram is delivered to destination group members with the same “best-effort” reliability as a standard unicast IP datagram. This means that a multicast datagram is not guaranteed to reach all members of the group, or arrive in the same order relative to the transmission of other packets. The only difference between a multicast IP packet and a unicast IP packet is the presence of a “group address” in the Destination Address field of the

IP header. Instead of a Class A, B, or C IP address, multicasting employs a Class D destination address format (224.0.0.0- 239.255.255.255).

Multicast Groups

Individual hosts are free to join or leave a multicast group at any time. There are no restrictions on the physical location or the number of members in a multicast group. A host may be a member of more than one multicast group at any given time and does not have to belong to a group to send messages to members of a group.

Group Membership Protocol

A group membership protocol is employed by routers to learn about the presence of group members on their directly attached subnetworks. When a host joins a multicast group, it transmits a group membership protocol message for the group(s) that it wishes to receive, and sets its IP process and network interface card to receive frames addressed to the multicast group. This receiver-initiated join process has excellent scaling properties since, as the multicast group increases in size, it becomes ever more likely that a new group member will be able to locate a nearby branch of the multicast distribution tree.

Multicast Addressing

A multicast address is assigned to a set of receivers defining a multicast group. Sendersuse the multicast address as the destination IP address of a packet that is to be transmitted to all group members.

Class D Addresses

An IP multicast group is identified by a Class D address. Class D addresses have their

high-order four bits set to “1110” followed by a 28-bit multicast group ID. Expressed in standard “dotted-decimal” notation, multicast group addresses range from 224.0.0.0 to 239.255.255.255. Figure 3 shows the format of a 32- bit Class D address. 0 1 2 3 31 1 1 1 0 Multicast Group ID 28 bits

The Internet Assigned Numbers Authority (IANA) maintains a list of registered IP multicast groups. The base address 224.0.0.0 is reserved and cannot be assigned to any group. The block of multicast addresses ranging from 224.0.0.1 to 224.0.0.255 is reserved for the use of routing protocols and other low-level topology discovery or maintenance protocols.

Multicast routers should not forward a multicast datagram with a destination address in this range, regardless of its TTL.

The remaining groups ranging from 224.0.1.0 to 239.255.255.255 are assigned to various multicast applications or remain unassigned. From this range, 239.0.0.0 to 239.255.255 .255 are to be reserved for site-local "administratively scoped" applications, not Internet-wide applications. Some of the well-known groups include: "all systems on this subnet" 24.0.0.1), "all routers on this subnet"(224.0.0.2), "all DVMRP routers" (224.0.0.4), "all OSPF routers" (224.0.0.5), "IETF-1-Audio" (224.0.1.11),

Mapping a Class D Address to an Ethernet Address

The IANA has been allocated a reserved portion of the IEEE-802 MAC-layer multicast address space. All of the addresses in IANA's reserved block begin with 01-00-5E (hex). A simple procedure was developed to map Class D addresses to this reserved address block. This allows IP multicasting to take advantage of the hardware-level multicasting supported by network interface cards.

Transmission and Delivery of Multicast Datagrams

When the sender and receivers are members of the same (LAN) subnetwork, the transmission and reception of multicast frames are relatively simple processes. The source station simply addresses the IP packet to the multicast group, the network interface card maps the Class D address to the corresponding IEEE-802 multicast address, and the frame is sent. Receivers that wish to capture the frame notify their IP layer that they want to receive datagrams addressed to the group.

Things become much more complicated when the sender is attached to one subnetwork and receivers reside on different subnetworks. In this case, the routers are required to implement a multicast routing protocol that permits the construction of multicast delivery trees and supports multicast data packet forwarding. In addition, each router needs to implement a group membership protocol that allows it to learn about the existence of group members on its directly attached subnetworks.

Internet Group Management Protocol (IGMP)

The Internet Group Management Protocol (IGMP) runs between hosts and their immediately neighboring multicast routers. The mechanisms of the protocol allow a host to inform its local router that it wishes to receive transmissions addressed to a specific multicast group. Also, routers periodically query the LAN to determine if known group members are still active. If there is more than one router on the LAN performing IP multicasting, one of the routers is elected "querier" and assumes the responsibility of querying the LAN for group members.

Based on the group membership information learned from the IGMP, a router is able to determine which (if any) multicast traffic needs to be forwarded to each of its "leaf" subnetworks. Multicast routers use this information, in conjunction with a multicast routing protocol, to support IP multicasting across the Internet.

Multicast Forwarding Algorithms

IGMP provides the final step in a multicast packet delivery service since it is only concerned with the forwarding of multicast traffic from the local router to group members on directly attached subnetworks. IGMP is not concerned with the delivery of multicast packets between neighboring routers or across an internetwork.

To provide an Internet-wide delivery service, it is necessary to define multicast routing protocols. A multicast routing protocol is responsible for the construction of multicast packet delivery trees and performing multicast packet forwarding. This section explores a number of different algorithms that may potentially be employed by multicast routing protocols:

Multicast service model

Host interface

Host-router interactions (IGMP)

Multicast Routing

Distance Vector

Link State

Shared tree

Motivation

Efficient delivery to multiple destinations (e.g. video broadcast)

Network-layer support for one-to-many addressing :

* Publish/subscribe communications model

*Don’t need to know destinations

IP Multicast service model Communications based on groups

Special IP addresses (Class D in IPv4) represent “multicast groups”. Anyone can join group to receive packets . Anyone can send to group, Sender need not be part of group , Dynamic group membership – can join and leave at will , Unreliable datagram service, Extension to unicast IP Group membership not visible to hosts, No synchronization.

Elements of IP Multicast

(1) Host interface
Application visible multicast API
Multicast addressing
Link-layer mapping
(2) Host-Router interface
IGMP
(3) Router-Router interface υ Multicast routing protocols

Host interface:

Senders (not much new) : Set TTL on multicast packets to limit “scope”
Scope can be administratively limited on per-group basis, Send packets to multicast address, represents a group , Unreliable transport (no acknowledgements)
Receivers (two new interfaces) : Join multicast group (group address) ,Leave multicast group (group address) . Typically implemented as a socket option in most networking API
Multicast addressing:
Special address range: Class D (3 MSBs set to 1) 224.0.01- 239.255.255.255, Reserved by IANA for multicast
Which address to use for a new group? No standard, Global random selection, Per-domain addressing.
Which address to use to join an existing group?: No standard, Separate address distribution protocol (may use multicast)
Internet Group Management Protocol (IGMP) : Host-router interface
Goal: Communicate group membership between hosts and routers
Soft-state protocol: Hosts explicitly inform their router about membership , Must periodically refresh membership report , Routers implicitly timeout groups that aren’t refreshed , Why isn’t explicit “leave group” message sufficient?
How IGMP works:

Router broadcasts membership query to 224.0.01 (all-systems group) with TTL=1

- Hosts start random timer (0-10 sec) for each group they have joined
- When a host's timer expires for group G, send membership report to group G, with TTL=1
- When a member of G hears a report, they reset their timer for G
- Router times out groups that are not "refreshed" by some host's report

Multicast routing

Router-router interface: Goal: Build distribution tree for multicast packets

Efficient tree (ideally, shortest path) & Low join/leave latency

Several approaches:

Distance Vector/Link State

Leverage existing unicast routing protocols.

Shared tree:

Unicast/multicast hybrids.

Multicast routing taxonomy:

Source-based tree :Separate shortest path tree for each source

Flood and prune (DVMRP, PIM-DM) i) Send multicast traffic everywhere ii) Prune edges that are not actively subscribed to group, Link-state (MOSPF): Routers flood groups they would like to receive i) Compute shortest-path trees on demand

Shared tree (PIM-SM) : Single distributed tree shared among all sources , Specify rendezvous point (RP) for group, Senders send packets to RP, receivers join at RP, RP multicasts to receivers; Fix-up tree for optimization Source-based