UNIT IV

Congestion Control, Congestion-Avoidance Mechanisms and Quality of Service, Internetworking:

Intra-Domain and Inter-Domain Routings, Unicast Routing Protocols: RIP, OSPF and BGP, Multicast Routing Protocols: DVMRP, PIM-DM, PIM-SM, CBT, MSDP and MOSPF, Spanning Tree Algorithm

4.1 Congestion Control

4.1.1 Congestion-Avoidance Mechanisms And Quality Of Service

Q1. What is congestion control? Explain the principles and techniques of congestion control.

Ans : **(Imp.)**

Meaning

Congestion Control is a type of network layer issue, and it is concerned with what happens when there is more data in the network than can be sent with reasonable packet delays, and there are no lost packets.

Causes

The main cause of congestion is a huge amount of data traffic. But other factors are equally important for making congestion as given below:

1. Sudden arrival of large data (called burst data) from many input lines and trying to access a single output line of a router. In this case, the particular output line is blocked if its bandwidth isn't sufficiently high.
2. Low bandwidth line will produce congestion even if the data rate isn't too high.
3. Mismatch between the speeds of different components of the system may also produce congestion.

Principle

The principles of congestion control are a set of guidelines and techniques that are used to manage network congestion and ensure that the network operates efficiently and effectively. Some of the key principles of congestion control include:

1. **Monitoring network traffic**

   Network congestion can occur when there is more traffic on the network than the network can handle. Therefore, it is essential to monitor network traffic continuously to detect congestion before it becomes a problem.

2. **Feedback-based mechanisms**

   Feedback-based mechanisms are used to control the rate of traffic flow and prevent congestion. These mechanisms involve sending feedback messages to the source of the traffic, indicating the current network conditions and the need to reduce the rate of traffic.

3. **Resource allocation**

   Congestion control involves allocating network resources, such as bandwidth and buffer space, effectively. This ensures that each flow of traffic receives a fair share of network resources and prevents any one flow from monopolizing the network.

4. **Congestion avoidance**

   Congestion avoidance techniques are used to prevent congestion from occurring in the first place. These techniques involve detecting and reacting to early signs of congestion, such as packet loss or delay, and reducing the rate of traffic to prevent congestion from occurring.

5. **Traffic prioritization**

   Congestion control involves prioritizing traffic based on its importance and criticality. This ensures that critical traffic, such as voice or video traffic, is given priority over less critical traffic, such as file downloads.

Techniques

There are several congestion control techniques that can be used to manage network traffic and prevent congestion. Some of the most common techniques include:

1. **Traffic shaping**

   Traffic shaping is a technique that involves regulating the rate of traffic entering the network to ensure that it does not exceed the capacity of the network. This technique can be used to prioritize critical traffic or to limit the amount of non-critical traffic.

2. **Packet dropping**

   Packet dropping is a technique that involves dropping packets when the network becomes congested. This technique can be used to prevent network overload and ensure that critical traffic receives priority.

3. **Resource reservation**

   Resource reservation is a technique that involves reserving network resources, such as bandwidth or buffer space, for specific types of traffic. This technique can be used to ensure that critical traffic receives the necessary resources and priority.

4. **Quality of Service (QoS)**

   QoS is a technique that involves assigning different levels of priority to different types of traffic. This technique can be used to ensure that critical traffic, such as voice or video traffic, receives priority over less critical traffic, such as email or file downloads.

. **Active Queue Management (AQM)**

AQM is a technique that involves monitoring the length of network queues and dropping packets before the queue becomes congested. This technique can be used to prevent buffer overflow and ensure that critical traffic receives priority.

Explicit Congestion Notification (ECN)

ECN is a technique that involves marking packets when congestion is detected. This technique can be used to signal congestion to end hosts and prevent network overload.

Q2. What are Congestion Avoidance Mechanisms ? Explain.

Ans :

Meaning

Congestion avoidance mechanisms are techniques used to prevent or mitigate network congestion, which occurs when the demand for network resources exceeds the available capacity. congestion avoidance mechanisms are critical to the performance and reliability of computer networks. By preventing or mitigating congestion, these techniques ensure that network traffic flows smoothly and efficiently, which is essential for businesses and organizations that rely on their networks for communication and data transfer.

Here are some common congestion avoidance mechanisms used in computer networks:

1. **Traffic Shaping**

   Traffic shaping is a technique that regulates the flow of network traffic by smoothing out bursts of traffic and ensuring that traffic flows within defined limits. It works by controlling the rate at which traffic is sent and received, which helps prevent congestion by preventing traffic from overwhelming network resources. Traffic shaping can be implemented using various techniques such as leaky bucket, token bucket, or rate limiting.

2. **Quality of Service (QoS)**

   QoS is a mechanism that prioritizes network traffic based on its importance or type. By prioritizing traffic, QoS can ensure that important traffic, such as VoIP or video traffic, is given priority over less important traffic, such as email or file transfers. QoS can be implemented using various techniques such as packet classification, marking, and policing.

3. **Random Early Detection (RED)**

   RED is a technique used to prevent congestion by selectively dropping packets before congestion occurs. RED randomly drops packets when the average queue length

exceeds a certain threshold, which helps prevent congestion by allowing congestion to be detected and controlled before it becomes severe. RED is typically used in routers to prevent buffer overflow.

4. **Explicit Congestion Notification (ECN)**

ECN is a technique used to inform network devices of congestion before it occurs. ECN-capable devices can mark packets to indicate congestion, which allows downstream devices to respond appropriately by reducing the amount of traffic they send. ECN is implemented in routers and hosts and is enabled by default in most modern operating systems.

5. **Adaptive TCP**

Adaptive TCP is a technique that dynamically adjusts the transmission rate of TCP connections based on network conditions. By adjusting the transmission rate, adaptive TCP can help prevent congestion by preventing TCP connections from overwhelming network resources. Adaptive TCP uses various techniques such as slow start, congestion avoidance, and fast retransmit to adjust the transmission rate.

4.2 Internet working

4.2.1 Intra-Domain And Inter-domain Routings

Q3. What is Routing? Explain Intra- Domain And Inter-domain Routings.

Ans :

Meaning

Routing is the process of directing network traffic from its source to its destination across a network. It is a fundamental concept in computer networking that enables devices on a network to communicate with each other by forwarding packets of data between them. The routing process involves the use of routing protocols and algorithms that determine the optimal path for data to travel from the source to the destination based on various factors, such as network topology, available bandwidth, and network congestion.

Routing can occur at different layers of the networking stack, including the physical layer, data link layer, network layer, and transport layer. At the network layer, routing is typically performed by devices such as routers, which use routing tables and algorithms to determine the best path for data to travel between networks.

In computer networking, intra-domain routing and inter-domain routing are two distinct routing protocols that are used to direct traffic within and between different autonomous systems (AS).

1. **Intra-domain routing**

Intra-domain routing protocols are used to direct traffic within a single autonomous system (AS). An autonomous system is a collection of networks that are controlled by a single entity or organization. Intra-domain routing protocols are typically used within an organization's network and are designed to ensure that traffic is directed along the most efficient path from the source to the destination within the same AS. Examples of intra-domain routing protocols include OSPF (Open Shortest Path First) and IS-IS (Intermediate System to Intermediate System).

2. **Inter-domain routing**

Inter-domain routing protocols are used to direct traffic between different autonomous systems. Inter-domain routing protocols are designed to ensure that traffic is directed along the most efficient path between different autonomous systems, taking into account factors such as network topology, available bandwidth, and cost. Examples of inter-domain routing protocols include BGP (Border Gateway Protocol) and EGP (Exterior Gateway Protocol).

Q4. Write the differences between Inter domain and Intradomain Routing. **(Imp.)**

Ans :

The following table highlights the major differences between interdomain and intradomain routing protocols.

S.No	Intradomain Routing	Interdomain Routing
1.	Routing algorithm works only within domains.	Routing algorithm works within and between domains.
2.	It need to know only about other routers within their domain.	It need to know only about other routers within and between their domain.
3.	Protocols used in intradomain routing are known as Interior-gate way protocols.	Protocols used in interdomain routing are known as Exterior-gateway protocols.
4.	In this Routing, routing takes place within an autonomous network.	In this Routing, routing takes place between the autonomous networks.
5.	Intradomain routing protocols ignores the internet outside the AS(autonomous system).	Interdomain routing protocol assumes that the internet contains the collection of interconnected AS(autonomous systems).
6.	Some Popular Protocols of this routing are RIP(resource information protocol) and OSPF(open shortest path first).	Popular Protocols of this routing is BGP (Border Gateway Protocol) used to connect two or more AS(autonomous system).

4.3 Unicast Routing Protocols

4.3.1 RIP

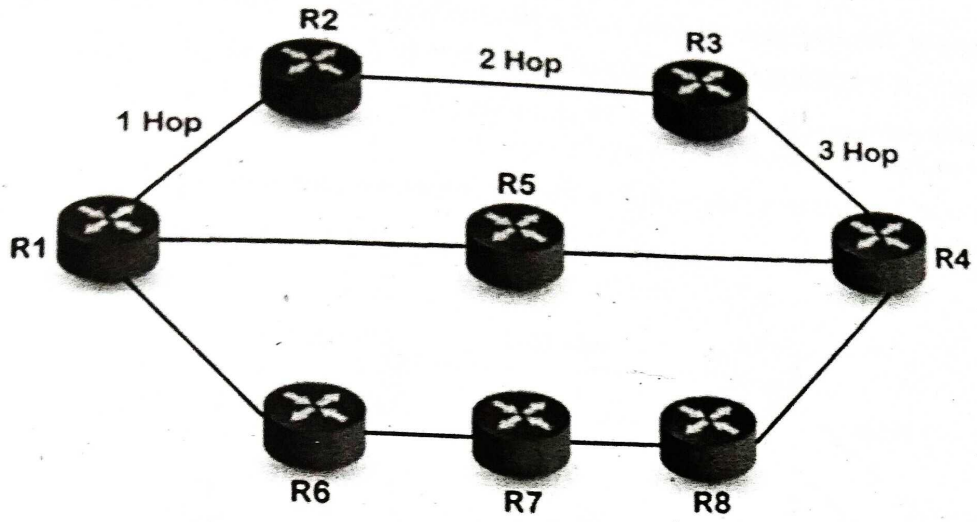
Q5. Explain bout RIP protocol.

Ans :

RIP stands for Routing Information Protocol. RIP is an intra-domain routing protocol used within an autonomous system. Here, intra-domain means routing the packets in a defined domain, for example, web browsing within an institutional area. To understand the RIP protocol, our main focus is to know the structure of the packet, how many fields it contains, and how these fields determine the routing tables.

- RIP is based on the distance vector-based strategy, so we consider the entire structure as a graph where nodes are the routers, and the links are the networks.
- In a routing table, the first column is the destination, or we can say that it is a network address.
- The cost metric is the number of hops to reach the destination. The number of hops available in a network would be the cost. The hop count is the number of networks required to reach the destination.
- In RIP, infinity is defined as 16, which means that the RIP is useful for smaller networks or small autonomous systems. The maximum number of hops that RIP can contain is 15 hops, i.e., it should not have more than 15 hops as 16 is infinity.
- The next column contains the address of the router to which the packet is to be sent to reach the destination.

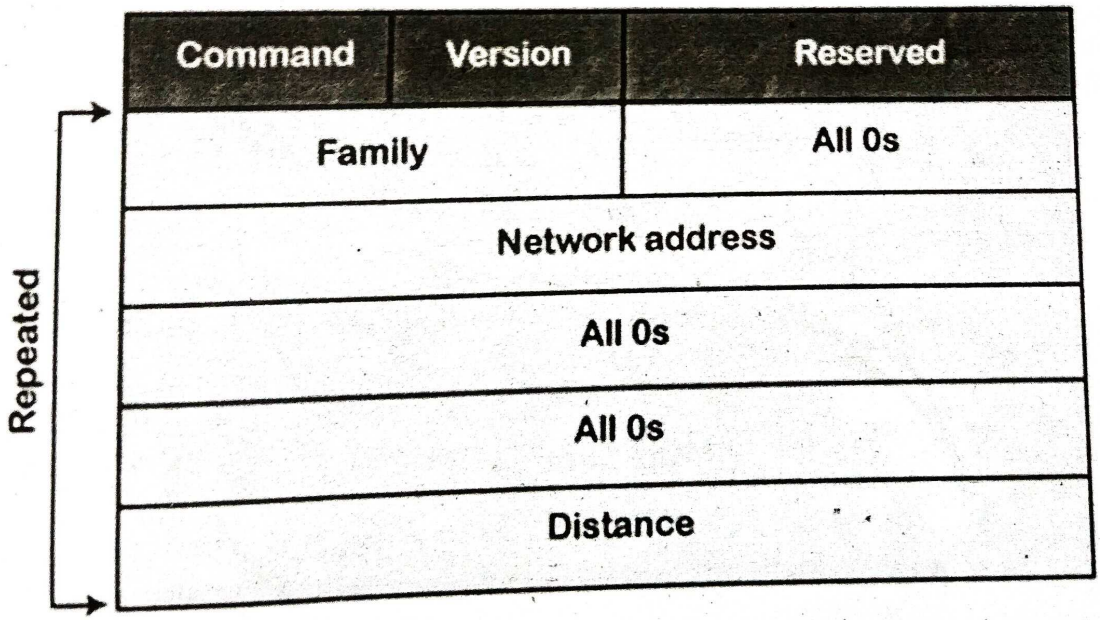
When the router sends the packet to the network segment, then it is counted as a single hop.



In the above figure, when the router 1 forwards the packet to the router 2 then it will count as 1 hop count. Similarly, when the router 2 forwards the packet to the router 3 then it will count as 2 hop count, and when the router 3 forwards the packet to router 4, it will count as 3 hop count. In the same way, RIP can support maximum upto 15 hops, which means that the 16 routers can be configured in a RIP.

RIP Message Format

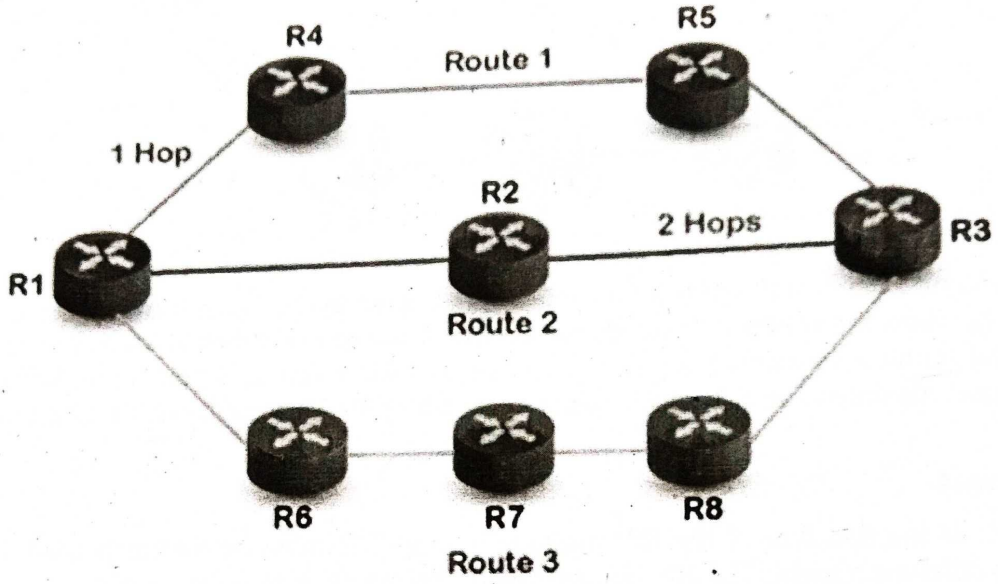
Now, we look at the structure of the RIP message format. The message format is used to share information among different routers. The RIP contains the following fields in a message:



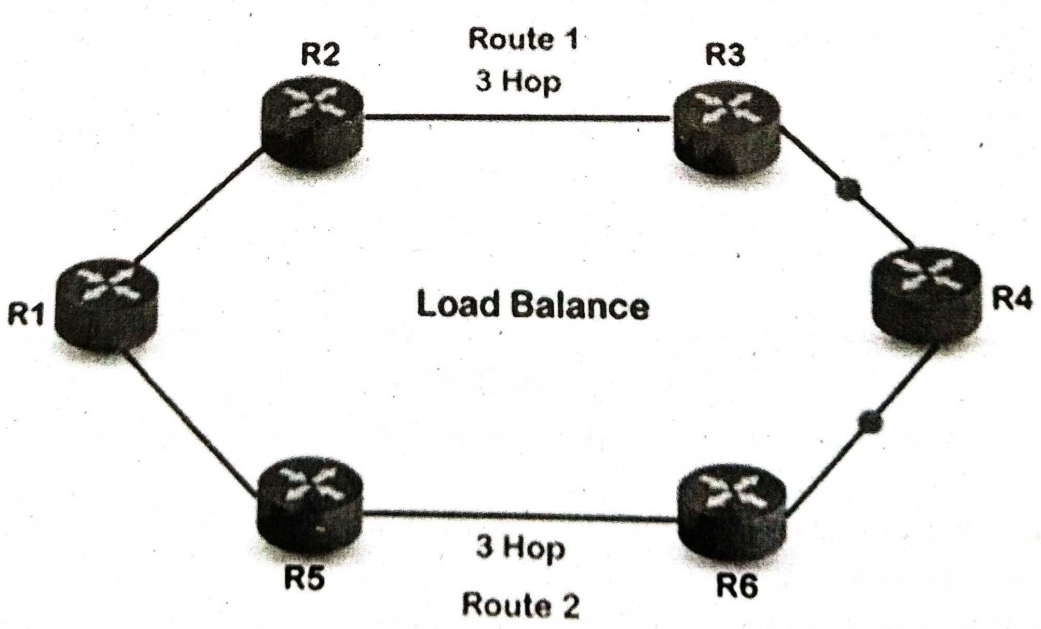
Command: It is an 8-bit field that is used for request or reply. The value of the request is 1, and the value of the reply is 2.

Version: Here, version means that which version of the protocol we are using. Suppose we are using the protocol of version1, then we put the 1 in this field.

- **Reserved:** This is a reserved field, so it is filled with zeroes.
- **Family:** It is a 16-bit field. As we are using the TCP/IP family, so we put 2 value in this field.
- **Network Address:** It is defined as 14 bytes field. If we use the IPv4 version, then we use 4 bytes, and the other 10 bytes are all zeroes.
- **Distance:** The distance field specifies the hop count, i.e., the number of hops used to reach the destination.



If there are 8 routers in a network where Router 1 wants to send the data to Router 3. If the network is configured with RIP, it will choose the route which has the least number of hops. There are three routes in the above network, i.e., Route 1, Route 2, and Route 3. The Route 2 contains the least number of hops, i.e., 2 where Route 1 contains 3 hops, and Route 3 contains 4 hops, so RIP will choose Route 2.



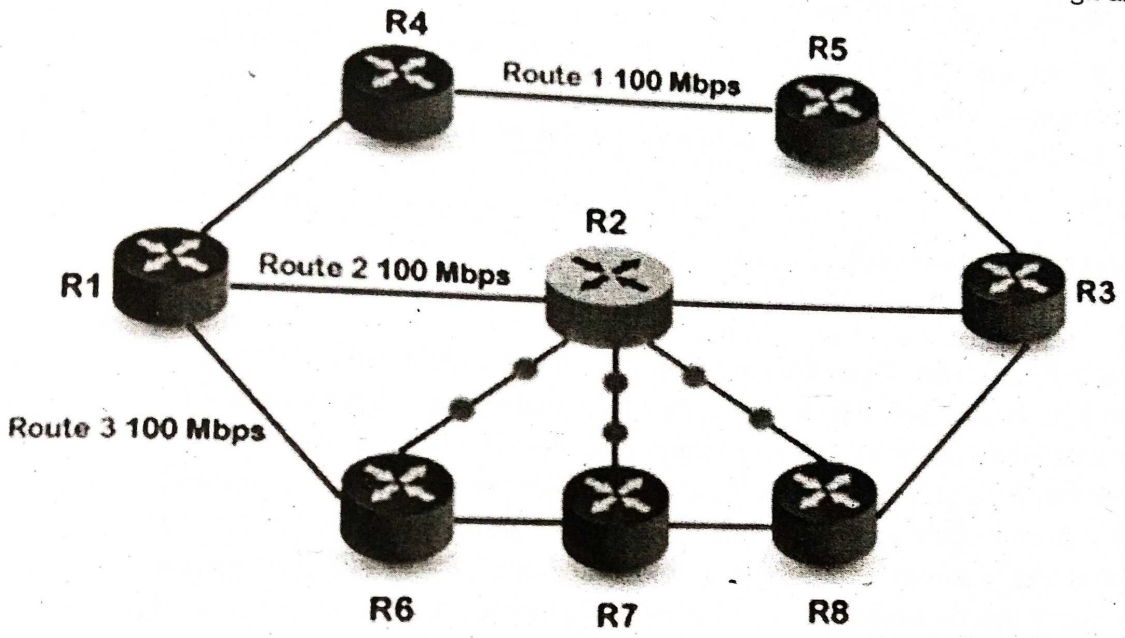
Suppose R1 wants to send the data to R4. There are two possible routes to send data from r1 to r2. As both the routes contain the same number of hops, i.e., 3, so RIP will send the data to both the routes simultaneously. This way, it manages the load balancing, and data reach the destination a bit faster.

Q6. Explain the Disadvantages of RIP.

Ans :

The following are the disadvantages of RIP:

(a) In RIP, the route is chosen based on the hop count metric. If another route of better bandwidth is available, then that route would not be chosen. Let's understand this scenario through an example.



We can observe that Route 2 is chosen in the above figure as it has the least hop count. The Route 1 is free and data can be reached more faster; instead of this, data is sent to the Route 2 that makes the Route 2 slower due to the heavy traffic. This is one of the biggest disadvantages of RIP.

- The RIP is a classful routing protocol, so it does not support the VLSM (Variable Length Subnet Mask). The classful routing protocol is a protocol that does not include the subnet mask information in the routing updates.
- It broadcasts the routing updates to the entire network that creates a lot of traffic. In RIP, the routing table updates every 30 seconds. Whenever the updates occur, it sends the copy of the update to all the neighbors except the one that has caused the update. The sending of updates to all the neighbors creates a lot of traffic. This rule is known as a split-horizon rule.
- It faces a problem of Slow convergence. Whenever the router or link fails, then it often takes minutes to stabilize or take an alternative route; This problem is known as Slow convergence.
- RIP supports maximum 15 hops which means that the maximum 16 hops can be configured in a RIP
- The Administrative distance value is 120 (Ad value). If the Ad value is less, then the protocol is more reliable than the protocol with more Ad value.
- The RIP protocol has the highest Ad value, so it is not as reliable as the other routing protocols.

The following timers are used to update the routing table:

- **RIP update timer : 30 sec**

  The routers configured with RIP send their updates to all the neighboring routers every 30 seconds.

- **RIP Invalid timer : 180 sec**

  The RIP invalid timer is 180 seconds, which means that if the router is disconnected from the network or some link goes down, then the neighbor router will wait for 180 seconds to take the update. If it does not receive the update within 180 seconds, then it will mark the particular route as not reachable.

- **RIP Flush timer : 240 sec**

  The RIP flush timer is 240 second which is almost equal to 4 min means that if the router does not receive the update within 240 seconds then the neighbor route will remove that particular route from the routing table which is a very slow process as 4 minutes is a long time to wait.

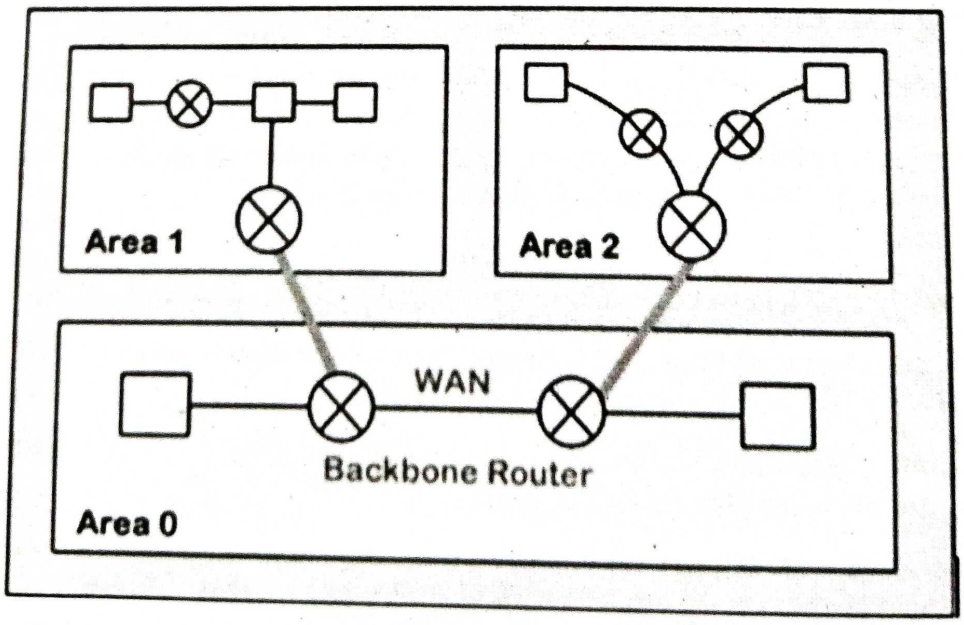
4.3.2 OSPF

Q7. Explain OSPF protocol.

Ans :

The OSPF stands for Open Shortest Path First. It is a widely used and supported routing protocol. It is an intradomain protocol, which means that it is used within an area or a network. It is an interior gateway protocol that has been designed within a single autonomous system. It is based on a link-state routing algorithm in which each router contains the information of every domain, and based on this information, it determines the shortest path. The goal of routing is to learn routes. The OSPF achieves by learning about every router and subnet within the entire network. Every router contains the same information about the network. The way the router learns this information by sending LSA (Link State Advertisements). These LSAs contain information about every router, subnet, and other networking information. Once the LSAs have been flooded, the OSPF stores the information in a link-state database known as LSDB. The main goal is to have the same information about every router in an LSDBs.

OSPF Areas



OSPF divides the autonomous systems into areas where the area is a collection of networks, hosts, and routers. Like internet service providers divide the internet into a different autonomous system for easy management and OSPF further divides the autonomous systems into Areas.

Routers that exist inside the area flood the area with routing information

In Area, the special router also exists. The special routers are those that are present at the border of an area, and these special routers are known as Area Border Routers. This router summarizes the information about an area and shares the information with other areas.

All the areas inside an autonomous system are connected to the backbone routers, and these backbone routers are part of a primary area. The role of a primary area is to provide communication between different areas.

Working

There are three steps that can explain the working of OSPF:

Step 1:

The first step is to become OSPF neighbors. The two connecting routers running OSPF on the same link creates a neighbor relationship.

Step 2:

The second step is to exchange database information. After becoming the neighbors, the two routers exchange the LSDB information with each other.

Step 3:

The third step is to choose the best route. Once the LSDB information has been exchanged with each other, the router chooses the best route to be added to a routing table based on the calculation of SPF.

Router forms a neighbor relationship

The first thing is happened before the relationship is formed is that each router chooses the router ID.

Router ID (RID): The router ID is a number that uniquely identifies each router on a network. The router ID is in the format of the IPv4 address. There are few ways to set the router ID, the first way is to set the router ID manually and the other way is to let the router decides itself.

The following is the logic that the router chooses to set the router ID:

- **Manually assigned**

The router checks whether the router ID is manually set or not. If it manually set, then it is a router ID. If it is not manually set, then it will choose the highest 'up' status loopback interface IP address. If there are no loopback interfaces, then it will choose the highest 'up' status non-loopback interface IP address.

Two routers connected to each other through point to point or multiple routers are connected can communicate with each other through an OSPF protocol. The two routers are adjacent only when both the routers send the HELLO packet to each other. When both the routers receive the acknowledgment of the HELLO packet, then they come in a two-way state. As OSPF is a link state routing protocol, so it allows to create the neighbor relationship between the routers. The two routers can be neighbors only when they belong to the same subnet, share the same area id, subnet mask, timers, and authentication. The OSPF relationship is a relationship formed between the routers so that they can know each other. The two routers can be neighbors if atleast one of them is designated router or backup designated router in a network, or connected through a point-to-point link.

Types

A link is basically a connection, so the connection between two routers is known as a link.

There are four types of links in OSPF:

1. **Point-to-point link**

The point-to-point link directly connects the two routers without any host or router in between.

2. **Transient link**

When several routers are attached in a network, they are known as a transient link.

The transient link has two different implementations:

Unrealistic topology

When all the routers are connected to each other, it is known as an unrealistic topology.

Realistic topology

When some designated router exists in a network then it is known as a realistic topology. Here designated router is a router to which all the routers are connected. All the packets sent by the routers will be passed through the designated router.

3. **Stub link**

It is a network that is connected to the single router. Data enters to the network through the single router and leaves the network through the same router.

4. **Virtual link**

If the link between the two routers is broken, the administration creates the virtual path between the routers, and that path could be a long one also.

OSPF Message Format

The following are the fields in an OSPF message format:

Version(8)	Type(8)	Message (16)
Source IP address		
Area Identification		
Chcek sum	Auth.Type	
Authentication (32)		

Version

It is an 8-bit field that specifies the OSPF protocol version.

Type

It is an 8-bit field. It specifies the type of the OSPF packet.

- **Message**

It is a 16-bit field that defines the total length of the message, including the header. Therefore, the total length is equal to the sum of the length of the message and header.

- **Source IP address**

It defines the address from which the packets are sent. It is a sending routing IP address.

- **Area identification**

It defines the area within which the routing takes place.

- **Checksum**

It is used for error correction and error detection.

- **Authentication type**

There are two types of authentication, i.e., 0 and 1. Here, 0 means for none that specifies no authentication is available and 1 means for pwd that specifies the password-based authentication.

- **Authentication**

It is a 32-bit field that contains the actual value of the authentication data.

Q8. Explain about OSPF Packets.

Ans :

There are five different types of packets in OSPF:

1. **Hello packet**

The Hello packet is used to create a neighborhood relationship and check the neighbor's reachability. Therefore, the Hello packet is used when the connection between the routers need to be established.

2. **Database Description**

After establishing a connection, if the neighbor router is communicating with the system first time, it sends the database information about the network topology to the system so that the system can update or modify accordingly.

3. **Link state request**

The link-state request is sent by the router to obtain the information of a specified route.

Suppose there are two routers, i.e., router 1 and router 2, and router 1 wants to know the information about the router 2, so router 1 sends the link state request to the router 2. When router 2 receives the link state request, then it sends the link-state information to router 1.

Link state update

The link-state update is used by the router to advertise the state of its links. If any router wants to broadcast the state of its links, it uses the link-state update.

Link state acknowledgment

The link-state acknowledgment makes the routing more reliable by forcing each router to send the acknowledgment on each link state update. For example, router A sends the link state update to the router B and router C, then in return, the router B and C sends the link- state acknowledgment to the router A, so that the router A gets to know that both the routers have received the link-state update.

Q9. Explain about OSPF States.

Ans :

The device running the OSPF protocol undergoes the following states:

- **Down**

If the device is in a down state, it has not received the HELLO packet. Here, down does not mean that the device is physically down; it means that the OSPF process has not been started yet.

- **Init**

If the device comes in an init state, it means that the device has received the HELLO packet from the other router.

- **2WAY**

If the device is in a 2WAY state, which means that both the routers have received the HELLO packet from the other router, and the connection gets established between the routers.

- **Exstart**

Once the exchange between the routers get started, both the routers move to the Exstart state. In this state, master and slave are selected based on the router's id. The master controls the sequence of numbers, and starts the exchange process.

- **Exchange**

In the exchange state, both the routers send a list of LSAs to each other that contain a database description.

- **Loading**

On the loading state, the LSR, LSU, and LSA are exchanged.

- **Full**

Once the exchange of the LSAs is completed, the routers move to the full state.

Q10. Explain about Router attributes.

Ans :

Before going to the Extract state, OSPF chooses one router as a Designated router and another router as a backup designated router. These routers are not the type, but they are the attributes of a router. In the case of broadcast networks, the router selects one router as a designated router and another router as a backup designated router. The election of designated and the backup designated router is done to avoid the flooding in a network and to minimize the number of adjacencies. They serve as a central point for exchanging the routing information among all the routers. Since point-to-point links are directly connected, so DR and BDR are not elected.

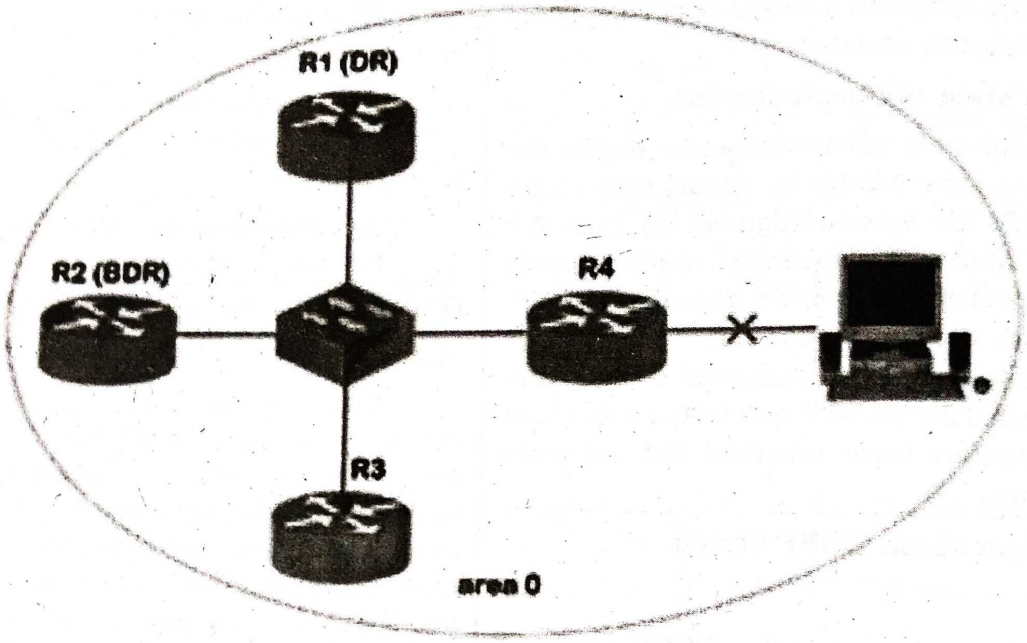
If DR and BDR are not elected, the router will send the update to all the adjacent neighbors, leading to the flooding in a network. To avoid this problem, DR and BDR are elected. Each non-DR and non-BDR send the update only to the DR and BDR instead of exchanging it with other routers in a network segment. DR then distributes the network topology information to other routers in the same area whereas the BDR serves a substitute for the DR. The BDR also receives the routing information from all the router but it does not distribute the information. It distributes the information only when the DR fails.

The multicast address 224.0.0.6 is used by the non-DR and non-BDR to send the routing information to the DR and BDR. The DR and BDR send the routing information to the multicast address 224.0.0.5.

Based on the following rules, the DR and BDR are elected:

- The router with the highest OSPF priority is chosen as the DR. By default, the highest priority is set as 1.
- If there is no highest priority, then the router with the highest router Id is chosen as the DR, and the router with the second-highest priority is chosen as the BDR.

Let's understand this scenario through an example.



In the above figure, R1 is chosen as the DR, while R2 is chosen as the BDR as R1 has the highest router ID, whereas the R2 has the second-highest router ID. If the link fails between R4 and the system, then R4 updates only R1 and R4 about its link failure. Then, DR updates all the non-DR and non-BDR about the change, and in this case, except R4, only R3 is available as a non-DR and non-BDR.

4.3.3 BGP

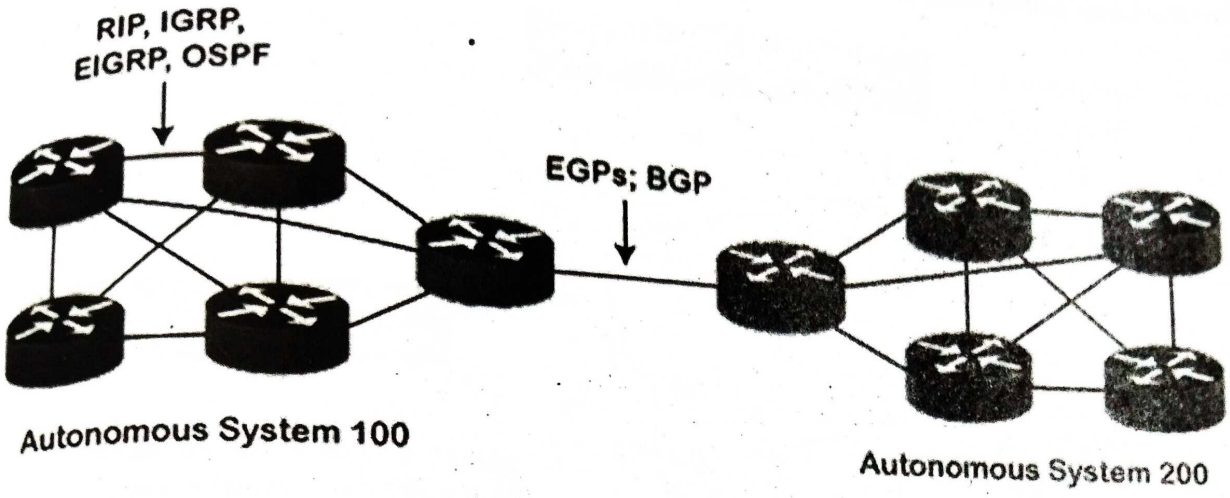
Q11. Explain BGP protocol.

Ans :

It is an interdomain routing protocol, and it uses the path-vector routing. It is a gateway protocol that is used to exchange routing information among the autonomous system on the internet.

There are many versions of BGP, such as:

- BGP version 1: This version was released in 1989 and is defined in RFC 1105.
- BGP version 2: It was defined in RFC 1163.
- BGP version 3: It was defined in RFC 1267.
- BGP version 4: It is the current version of BGP defined in RFC 1771.



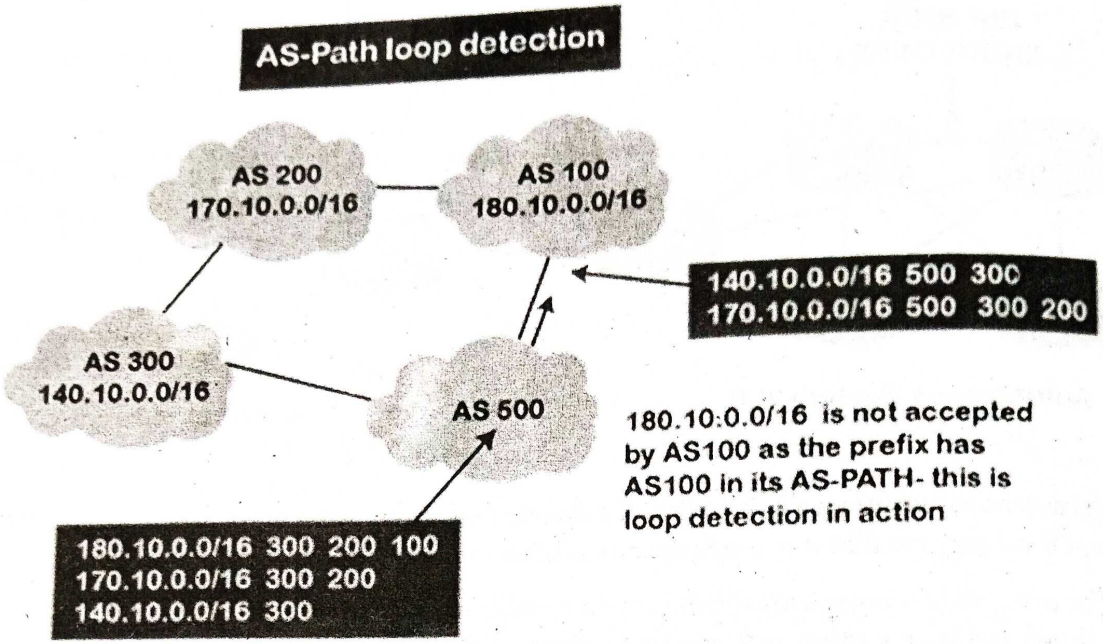
An autonomous system is a collection of networks that comes under the single common administrative domain. Or we can say that it is a collection of routers under the single administrative domain.

For example, an organization can contain multiple routers having different locations, but the single autonomous number system will recognize them. Within the same autonomous system or same organization, we generally use IGP (Interior Gateway Protocol) protocols like RIP, IGRP, EIGRP, OSPF. Suppose we want to communicate between two autonomous systems. In that case, we use EGP (Exterior Gateway Protocols). The protocol that is running on the internet or used to communicate between two different autonomous number systems is known as BGP (Border Gateway Protocol). The BGP is the only protocol that is running on the internet backbone or used to exchange the routes between two different autonomous number systems. Internet service providers use the BGP protocol to control all the routing information.

Features

- **Open standard-** It is a standard protocol which can run on any window device.
- **Exterior Gateway Protocol** -It is an exterior gateway protocol that is used to exchange the routing information between two or more autonomous system numbers.
- **Inter AS-domain routing** - It is specially designed for inter-domain routing, where inter AS-domain routing means exchanging the routing information between two or more autonomous number system.
- **Supports internet** - It is the only protocol that operates on the internet backbone.
- **Classless** - It is a classless protocol.
- **Incremental and trigger updates** - Like IGP, BGP also supports incremental and trigger updates.
- **Path vector protocol** - The BGP is a path vector protocol. Here, path vector is a method of sending the routes along with routing information.
- **Configure neighborhood relationship** - It sends updates to configure the neighborhood relationship manually. Application layer protocol

BGP's Loop prevention mechanism

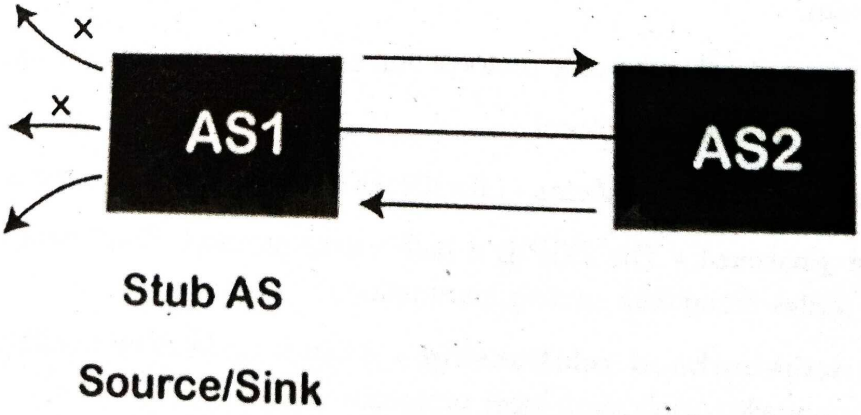


There is a possibility that when you are connecting to the internet, then you may be advertising route 10.0.0.0 to some autonomous system, then it is advertised to some other autonomous system. Then there is a possibility that the same route is coming back again. This creates a loop. But, in BGP, there is a rule that when the router sees its own AS number for example, as shown in the above figure, the network 180.10.0.0/16 is originating from the AS 100, and when it sends to the AS 200, it is going to carry its path information, i.e., 180.10.0.0/16 and AS 100. When AS 200 sends to the AS 300, AS 200 will send its path information 180.10.0.0/16 and AS path is 100 and then 200, which means that the route originates from AS 100, then reaches 200 and finally reaches to 300. When AS 300 sends to the AS 500, it will send the network information 180.10.0.0/16, and AS path is 100, 200, and then 300. If AS 500 sends to the AS 100, and AS 100 sees its own autonomous number inside the update, it will not accept it. In this way, BGP prevents the loop creation.

Types of Autonomous systems

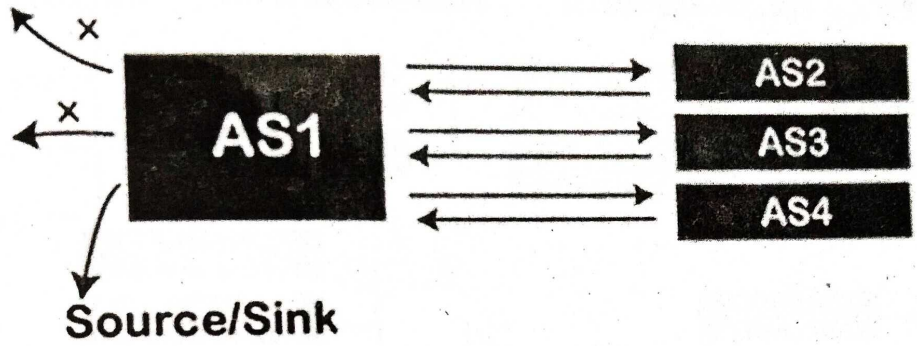
The following are the types of autonomous systems:

- Stub autonomous system



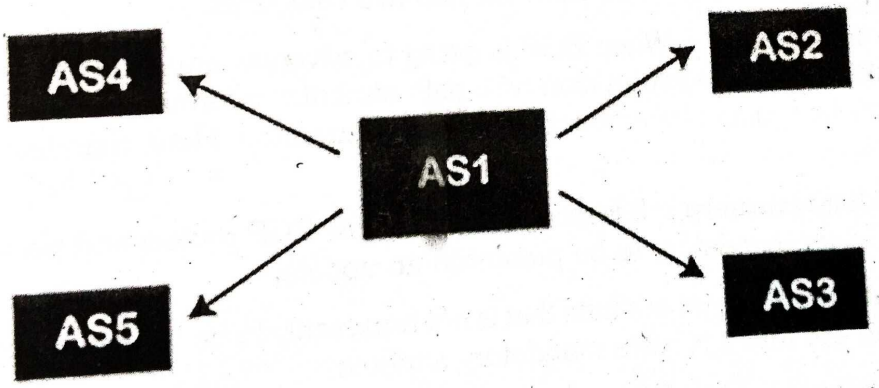
It is a system that contains only one connection from one autonomous system to another autonomous system. The data traffic cannot be passed through the stub autonomous system. The Stub AS can be either a source or a sink. If we have one autonomous system, i.e., AS1, then it will have a single connection to another autonomous system, AS2. The AS1 can act either as a source or a sink. If it acts as a source, then the data moves from AS1 to AS2. If AS1 acts as a sink, means that the data gets consumed in AS1 which is coming from AS2, but the data will not move forward from AS1.

- Multihomed autonomous system



It is an autonomous system that can have more than one connection to another autonomous system, but it can still be either a source or a sink for data traffic. There is no transient data traffic flow, which means that the data can be passed from one autonomous system.

- Transient Autonomous System

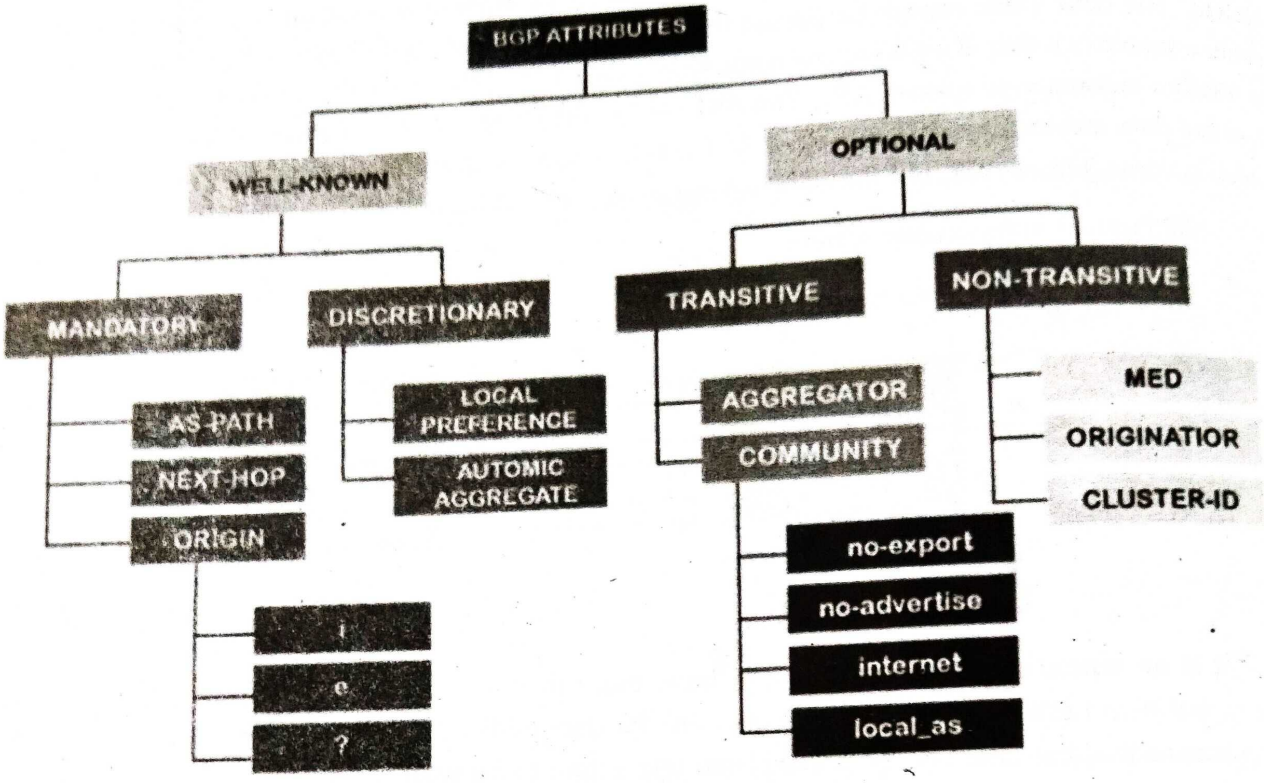


The transient autonomous system is a multihomed autonomous system, but it also provides transient traffic flow.

Path attributes

The BGP chooses the best route based on the attributes of the path.

As we know that path-vector routing is used in the border gateway routing protocol, which contains the routing table that shows the path information. The path attributes provide the path information. The attributes that show or store the path information are known as path attributes. This list of attributes helps the receiving router to make a better decision while applying any policy. Let's see the different types of attributes. The path attribute is broadly classified into two categories:



1. **Well-known attribute :** It is an attribute that should be recognized by every BGP router.

   The well-known attribute is further classified into two categories:

- **Well-known mandatory :** When BGP is going to advertise any network, but it also advertises extra information, and that information with path attributes information. The information includes AS path information, origin information, next-hop information. Here, mandatory means that it has to be present in all the BGP routing updates.
- **Well-known discretionary :** It is recognized by all the BGP routers and passed on to other BGP routers, but it is not mandatory to be present in an update.

2. **Optional attribute :** It is an attribute that is not necessarily to be recognized by every BGP router. In short, we can say that it is not a mandatory attribute.

   The optional attribute is further classified into two categories:

- **Optional transitive:** BGP may or may not recognize this attribute, but it is passed on to the other BGP neighbors. Here, transitive means that if the attribute is not recognized, then it is marked as a partial.
- **Optional non-transitive:** If the BGP cannot recognize the attribute, it ignores the update and does not advertise to another BGP router.

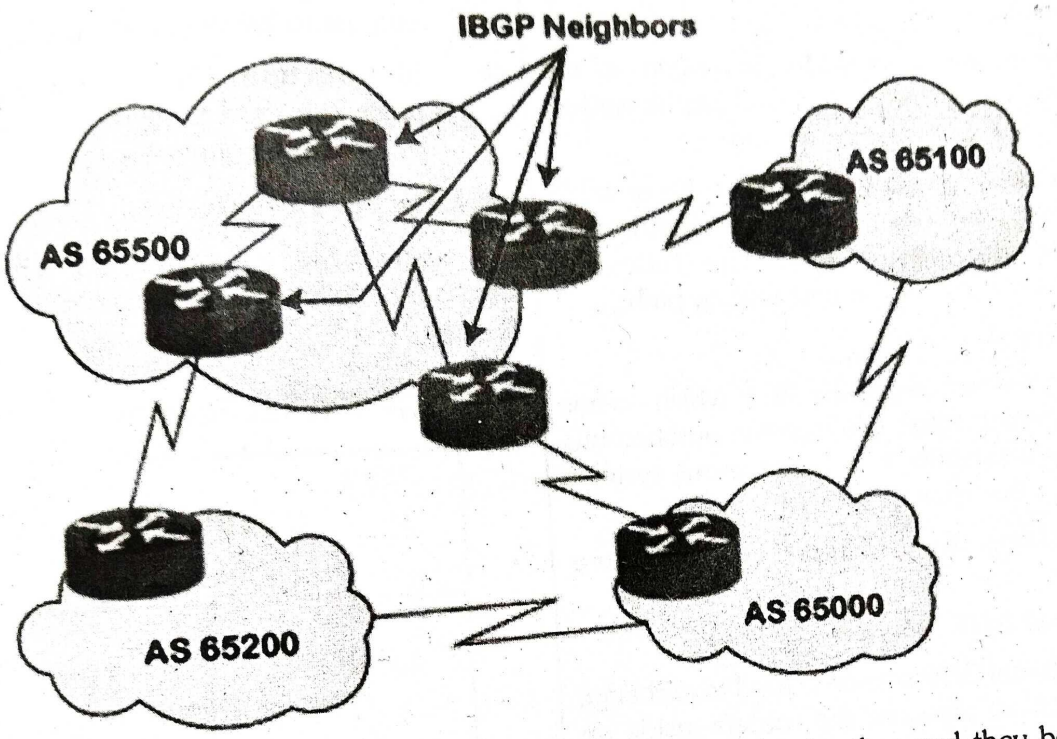
BGP Neighbors

BGP neighborship is similar to the OSPF neighborship, but there are few differences. BGP forms the neighboring relationship with the help of the TCP connection on port number 179 and then exchanges the BGP updates. They exchange the updates after forming the neighbor relationship. In BGP, the neighbor relationship is configured manually. BGP neighbors are also known as BGP peers or BGP speakers.

There are two types of neighbor relationship:

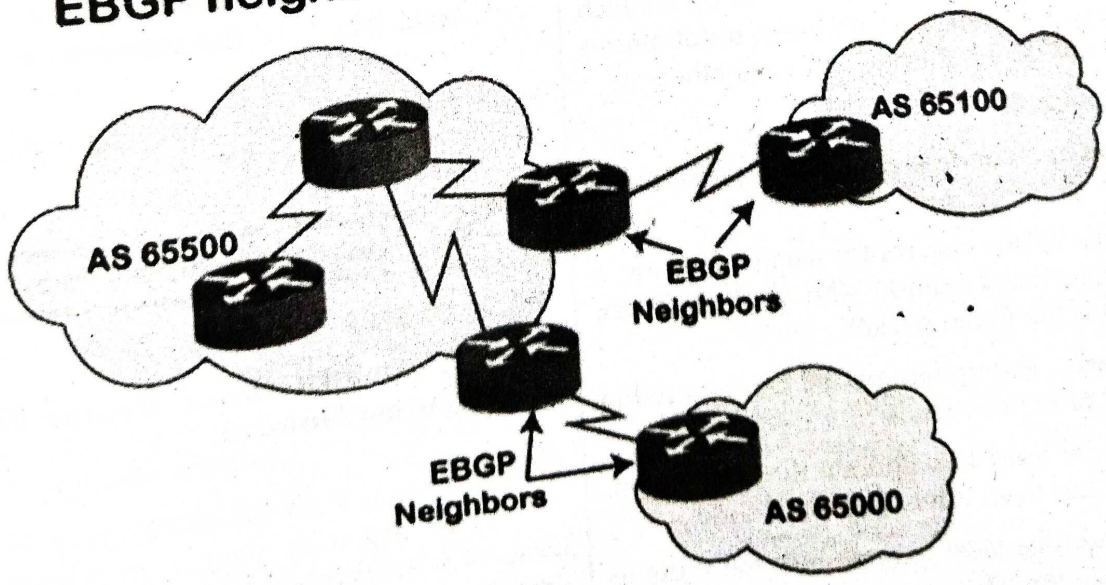
IBGP (Internal BGP): If all the routers are neighbors of each other and belong to the same autonomous number system, the routers are referred to as an IBGP.

IBGP neighbors



EBGP (External BGP): If all the routers are neighbors of each other and they belong to the different autonomous number systems, then the routers are referred to as an EBGP.

EBGP neighbors



BGP Tables

There are three types of BGP tables:

- **Neighbor table:** It contains the neighbors who are configured by the administrator manually. The neighbor relationship has to be manually configured by using the neighbor command.
- **BGP forwarding table:** It contains all the routes advertised in BGP and can be verified using the following command:
- **IP routing table:** The IP routing table contains the best path routes required to reach the destination. The following command shows the best routing path:

BGP Sessions

When we talk about the BGP, which means that the communication between the autonomous systems. Let's consider two autonomous systems having five nodes each.

BGP sessions are classified into two categories:

1. **Internal BGP session**

   The internal BGP session is used to exchange information between the routers inside an autonomous system. In short, we can say that the routing information is exchanged between the routers of the same autonomous system.

2. **External BGP session**

   The external BGP session is a session in which nodes or routers of different autonomous systems communicate with each other.

Types of packets

There are four different types of packets exist in BGP:

- **Open:** When the router wants to create a neighborhood relation with another router, it sends the Open packet.
- **Update:** The update packet can be used in either of the two cases:

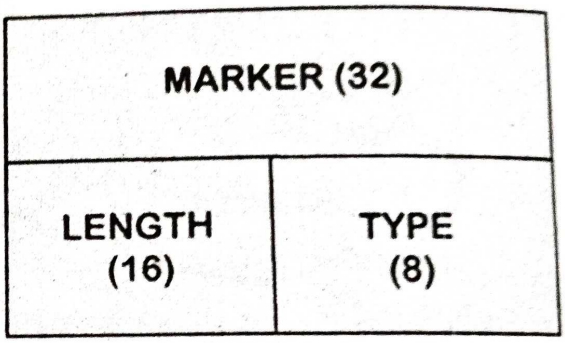
1. It can be used to withdraw the destination, which has been advertised previously.
2. It can also be used to announce the route to the new destination.

- **Keep Alive:** The keep alive packet is exchanged regularly to tell other routers whether they are alive or not. For example, there are two routers, i.e., R1 and R2. The R1 sends the keep alive packet to R2 while R2 sends the keep alive packet to R1 so that R1 can get to know that R2 is alive, and R2 can get to know that R1 is alive.
- **Notification:** The notification packet is sent when the router detects the error condition or close the connection.

BGP Packet Format

Now we will see the format in which the packet travels. The following are the fields in a BGP packet format:

BGP Packet Format



1. **Marker:** It is a 32-bit field which is used for the authentication purpose.
2. **Length:** It is a 16-bit field that defines the total length of the message, including the header.
3. **Type:** It is an 8-bit field that defines the type of the packet.

4.4 Multicast Routing Protocols

4.4.1 DVMRP

Q12. Explain Distance Vector Multicast Routing Protocol.

Ans :

The distance vector multicast routing protocol is multicast routing protocol that takes the routing decision based upon the source address of the packet.

- This algorithm constructs the routing tree for a network.
- Whenever a router receives a packet, it forwards it to some of its ports based on the source address of packet.

2. It must prevent the formation of duplicate packets.
3. It must ensure that the path traveled by a packet is the shortest from its source to the router.
4. It should provide dynamic membership.

To accomplish this, the DVMR algorithm uses a process based on following decision making strategies:

1. **Reverse Path Forwarding (RPF)**

- In this strategy, the router only forwards those packets that have traveled the shortest path from source to destination.
- To achieve this, the router pretends that it has a packet to send to the source from where the packet has arrived.
- In this way, the shortest path to the sender of the packet is computed.
- If the same route is followed by the received packet, it is forwarded to the next router and it is discarded otherwise.
- The reverse path forwarding ensures that the network receives a copy of the packet without formation of loops. A loop occurs when a packet that has left the router may come back again from another interface or the same interface and be forwarded again.
- RPF does not guarantee that there would be no duplicate packets in the network i.e. the network may receive two or more copies.
- The reason for this is that the routing is based on the source address and not on the destination address.

2. **Reverse Path Broadcasting (RPB)**

- In order to solve the problem, RPB is used.
- In this method, one parent router is defined for each network.
- The network could accept the multicast packets from this parent router only.
- This router sends packets to those ports for which it is designated as parent.
- Thus, RPB principle' allows a router to broadcast the packet in the network.

This creates duplicate packets on the network and reduces the network efficiency.

3. **Reverse Path Multicasting (RPM)**

- To overcome the problem of broadcasting in RPB, Reverse Path Multicasting in used.
- In this the desired multicast network tree is created by using two different methods: Pruning and grafting.
- A router can send a prune message to its upstream router whenever it finds that its network is not interested in a multicast packet. In this way a router prunes (cuts) its network from multicasting.
- If a router receives prune message from all the downstream routers, it in turn, sends a prune message to its upstream router.
- A router can also send a graft message to its upstream router if it finds that its network is again interested in receiving the multicast packet. In this way, graft message forces the upstream router to resume sending the multicast message. The network is again grafted (joined).

4. **Multicast Open Shortest Path First (MOSPF)**

- Multicast open shortest path first is the multicast version of open shortest path first protocol.
- It is an extension of OSPF that uses multicast link state routing method to create source based trees.
- The method used by MOSPF is different from DVMRP.
- The first difference is· that in this method, the tree is least cost tree instead of shortest path tree.
- The second difference is that the tree is not made gradually. It is made immediately it is prepared and ready to use.

4.4.2 PIM-DM

Q13. What is PIM?

Ans :

Meaning

PIM (Protocol Independent Multicast) is a family of multicast routing protocols that are designed to efficiently distribute multicast traffic across a network. PIM can be used in different modes depending on the network topology and the type of multicast traffic being distributed.

Modes

1. **Dense Mode (PIM-DM)**

   Dense mode is used in networks with high-density multicast traffic where most of the network nodes are interested in the multicast traffic. In PIM-DM, multicast traffic is initially flooded to all network nodes, and then pruned back based on the network topology and multicast group membership.

2. **Sparse Mode (PIM-SM)**

   Sparse mode is used in networks with sparse multicast traffic where only a small subset of nodes are interested in the multicast traffic. In PIM-SM, multicast traffic is forwarded only to those networks that have receivers for that multicast group.

   PIM also has a Bidirectional mode (PIM-BIDIR) that is used in networks where multicast traffic is bidirectional, such as video conferencing.

   PIM is a protocol-independent multicast routing protocol, which means that it can work with different unicast routing protocols, such as OSPF, BGP, or IS-IS. PIM uses different mechanisms to build multicast distribution trees and to forward multicast traffic, such as Reverse Path Forwarding (RPF) and Shared Trees.

Component

The main components of the PIM (Protocol Independent Multicast) protocol are as follows:

1. **Multicast group membership management**

   PIM includes mechanisms to manage multicast group membership and to distribute multicast traffic to interested receivers. PIM can work with different multicast group membership protocols, such as IGMP or MLD.

2. **Multicast routing protocols**

   PIM can work with different unicast routing protocols, such as OSPF or BGP, to exchange routing information and to build multicast distribution trees.

3. **Multicast forwarding mechanisms**

   PIM includes different mechanisms to forward multicast traffic, such as Reverse Path Forwarding (RPF) and Shared Trees. RPF ensures that multicast traffic is forwarded in the direction of the root of the multicast tree and avoids loops. Shared Trees allow multiple multicast groups to share the same distribution tree, reducing the overhead of building separate trees for each group.

4. **Multicast traffic control mechanisms**

   PIM includes mechanisms to control multicast traffic, such as pruning, which removes branches of the multicast tree where there are no receivers. Pruning reduces unnecessary multicast traffic and conserves network resources.

5. **Multicast rendezvous points**

   In PIM-SM, a multicast rendezvous point (RP) is used to bootstrap multicast traffic distribution and to build multicast trees. The RP is a centralized point where multicast traffic is initially sent before being distributed to interested receivers.

Q14. Explain PIM –DM protocol.

Ans :

PIM Dense-Mode

PIM-DM (Protocol Independent Multicast - Dense Mode) is a multicast routing protocol that is used to efficiently distribute multicast traffic in a dense network with high bandwidth connectivity. It is a flood-and-prune protocol, which means that multicast traffic is initially flooded to all connected networks and then pruned back based on the network topology and multicast group membership.

In PIM-DM, multicast traffic is forwarded to all directly connected networks until it reaches a

network where there are receivers interested in the traffic. Once a receiver is found, the multicast traffic is forwarded only to those networks that have receivers for that multicast group. PIM-DM builds a multicast distribution tree that is rooted at the source and extends to all receivers.

PIM-DM uses a reverse path forwarding (RPF) algorithm to prevent loops and to ensure that multicast traffic is forwarded in the direction of the root of the multicast tree. PIM-DM also uses a prune mechanism to remove branches of the multicast tree where there are no receivers.

PIM-DM is suitable for networks with high-density multicast traffic and where bandwidth is not an issue. It is less suitable for networks with sparse multicast traffic and limited bandwidth.

Being Built Back to the Sender.

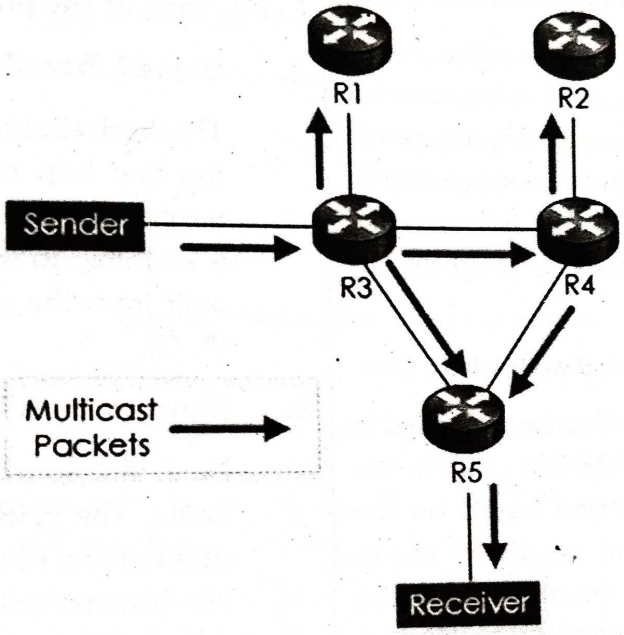


Fig : . PIM-DM Flooding ; (S, G)

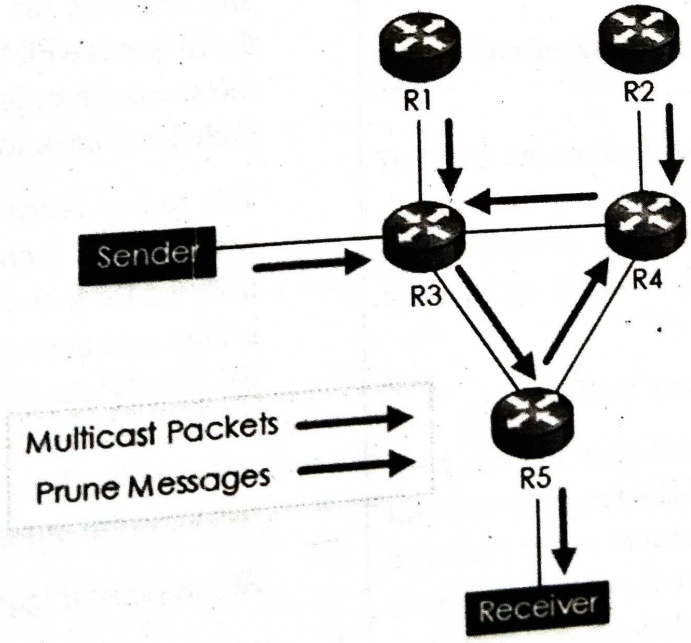


Fig:. PIM-DM Pruning

Working of PIM_DM

Here are the main steps involved in the working of PIM-DM:

1. **Discovery of multicast sources**

   The first step in PIM-DM is the discovery of multicast sources. Multicast sources are identified using multicast group addresses and their associated network interface addresses.

2. **Building the multicast distribution tree**

   PIM-DM builds a multicast distribution tree by flooding multicast traffic to all network nodes. This flood occurs in an expanding ring search manner, which means that the multicast traffic is initially flooded to the directly connected network nodes, and then to progressively more distant nodes.

3. **Pruning the multicast distribution tree**

   After the multicast traffic has been flooded to all network nodes, PIM-DM prunes the multicast distribution tree based on the network topology and multicast group membership. Pruning involves removing branches of the multicast tree where there are no receivers, reducing the amount of multicast traffic on the network.

4. **Maintaining the multicast distribution tree**

   PIM-DM continuously monitors the multicast distribution tree to detect changes in network topology or multicast group membership. If a change is detected, PIM-DM updates the tree accordingly.

5. **Forwarding multicast traffic**

   Once the multicast distribution tree has been built and pruned, PIM-DM forwards multicast traffic to only those network nodes that have requested it. This reduces unnecessary multicast traffic on the network and conserves network resources.

4.4.3 PIM-SM

Q15. Explain PIM-SM protocol.

Ans :

Meaning

PIM-SM works in the opposite manner to PIM-DM; with PIM sparse mode no multicast traffic is forwarded unless some requests it. PIM-SM works via the use of an RP (Rendezvous Point).

Let us look at the process that PIM-SM takes:

1. **Shared Tree (RP to Receiver)**

   The receiver sends an IGMP Join message to the first hop router (FHR), i.e its direct neighboring router. This router will then send a PIM Join to the RP. A shared tree is then built from the receiver to RP based upon (*, G).

2. **Shortest Path Tree (Source to RP)**

   Next, the source starts to send multicast traffic. The FHR encapsulates the multicast packet into a PIM register message and sends via unicast to the RP router. The RP decapsulates the packet and checks the multicast group to see if it has any state for any receivers for the multicast group. If so the RP sends a PIM Join message back towards the source, in order to build an SPT (Shortest Path Tree) back to the source.

   The source sends another multicast packet. however now there is a SPT (aka source tree) from the RP to the FHR. Therefore the packet is now sent across the distribution tree to the RP. The RP receives the packet nativity (with S, G), and in turn, sends a register-stop message back to the sources FHR to stop receiving the register messages (via unicast).

   At this point (Figure 5), we now have a

   - source tree from RP to the source,
   - shared tree from RP to the receiver.

3.

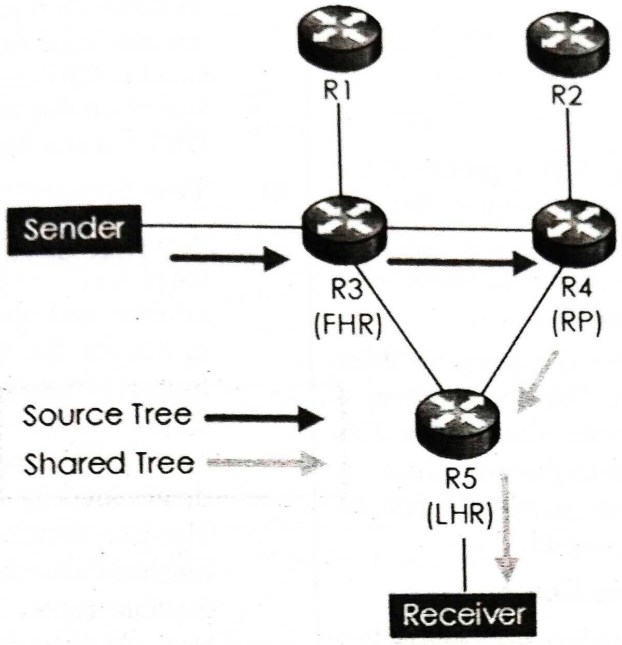


Fig:. PIM Sparse, SPT and Shared Tree

3. Shortest Path Tree Switchover

Although having the RP join back towards the source removes the encapsulation overhead, it does not completely optimize the forwarding paths. As for our receivers, the path via the RP (shared tree) still may be suboptimal. To overcome this, the process to migrate the shared tree to source tree, as known as the SPT switch over begins.

First, the LHR sees the source address of the multicast stream. Now that the LHR knows the source address, it sends an (S, G) Join to the source of the multicast stream. This builds an SPT from LHR to FHR. Finally, a Prune message is sent from the LHR to the RP in order to remove the previously used (RP to LHR) shared tree (Figure 6).

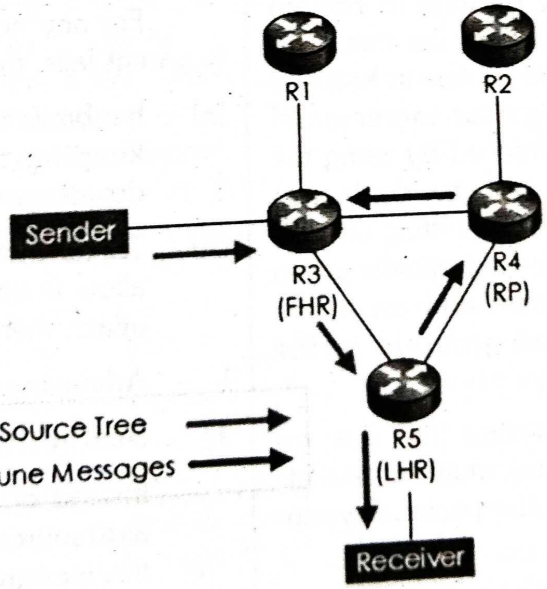


Fig:. SPT Switchover

4.4.4 CBT

Q16. What is CBT ? Explain it.

Ans :

Meaning

The Core-Based Tree (CBT) protocol is a group-shared protocol that uses a core as the root of the tree. The autonomous system is divided into regions, and a core (center router or rendezvous router) is chosen for each region.

There is a core address and a group identifier associated with every group. The core address is the normal unicast address of the core router. This address is used to get packets to the tree. Once on the tree, the packet is multicast based on a globally unique group identifier or group-id.

1) Identification of Core Router

The placement of a group's core should reflect that group's characteristics since the core placement assists in optimizing the routes between any sender and group members on the tree. A router could become a core when a host on one of its attached subnets wishes to initiate a group. Or in case of a single sender, the router nearest to it could become a core. The topic of core placement is open for research.

2) Data Forwarding

Unicast routing is used to route multicast packets to a multicast tree, allowing multicast groups and multicast packets to remain "invisible" to routers not on the tree. This allows for CBT unaware routers in between and is a good strategy for incremental deployment. This is achieved by using the unicast address of the core in the destination field of multicast packets originating off-tree. Data packets destined for a particular group tree carry the group core address in the "destination field" and group-id in the "option" field of IP packet's header.

Once on the corresponding tree (i.e. on arrival at an on-tree router), multicast packets span the tree based on the packet's group-identifier, or group-id.

Core-address in the "destination field" is discarded and group-id in the "option" field is placed in the "destination field". This leads to faster on-tree switching since it is faster to process fixed length header than an extended header. CBT routers forward arriving packets based on the information contained in their CBT Forwarding Information Base.

3) Tree Formation

When a receiver joins a multicast group, its local CBT router looks up the multicast address and obtains the address of the Core router for the group. It then sends a Join-Request message for the group towards the Core. The Join-Request is forwarded to the next-hop router on the path to the core as determined by the unicast forwarding table. The join continues its journey until it either reaches the addressed core, or reaches a CBT capable router that is already part of the tree. At this point, the join's journey is terminated by the receiving router and a Join-Ack is sent. At each CBT router traversed by the Join-Ack, forwarding state is instantiated. In this way, a multicast tree is built.

When a receiver wants to quit a multicast group, same procedure is followed (Quit-Req and Quit-Ack).

Path (or) Node Failure

Link failure is recognizable as a result of "Keep-Alive" mechanism operating between adjacent on-tree routers.

For any non-core router, if a parent or path to parent fails, there are two options

(a) It submits a new Join-Request message, hence keeping the failure transparent to the downstream branch. OR

(b) Tell downstream routers about the failure and allow them to independently attempt to re-attach themselves to the tree.

I. Advantages of Core Based Trees

1) Scalability

Instead of one tree per (source,group) pair as in source based trees, there is one multicast tree per group. So the amount of state that needs to be stored at each router on the tree is O(number of groups) i.e. link information

per tree. Moreover, all routers need not support or implement this protocol for this protocol to work i.e. routers which have no members wrt a particular group need not maintain any information as to the existence of that group.

2) **Tree Creation is receiver based**

Only a router interested in becoming a part of the group (or is on the path between a potential member and the tree) is involved in becoming a part of the tree for that particular group. So tree building overhead is restricted to these routers.

3) It is independent of underlying unicast routing algorithm, resulting in a much simplified multicast tree formation across domain boundaries.

II. **Disdvantages of Core Based Trees**

1) **Core Placement**

Core based trees may not provide the most optimal paths between members of a group.

2) **The Core as a Single Point of Failure**

This problem can be solved by having multiple cores associated with each tree, at the cost of increased complexity.

Two choices with multiple core nodes

(a) **Single Core CBT Trees**

We have multiple "backup" cores to increase the probability that every network node can reach at least one of the cores of a CBT tree. There are multiple cores, which join each other at group initiation time. The primary core is considered the "central-hub" of a tree, with additional nodes simply providing an element of robustness to the design. If the primary core should fail, the recovery scenario is same as that in case of Path or Node Failure. * This has the dynamic Join Overhead.

(b) **Multiple Core CBT Trees**

In this subsets of tree attached to each of core routers. It may lead to optimization of routes between those members. There must be an explicit protocol operating amongst the "backup cores" to handle failure, unlike the earlier case.

4.4.5 MSDP AND MOSPF

Q17. What is MSDP protocol? Explain how to configure it.

Ans : **(Imp.)**

Meaning

MSDP (Multicast Source Discovery Protocol) is a protocol used in multicast networks to distribute information about active multicast sources across multiple Autonomous Systems (ASes). It is used in combination with PIM (Protocol Independent Multicast) to enable multicast traffic to flow between multiple PIM domains.

Components

Here are the main components and features of MSDP:

1. **MSDP Speakers**

MSDP Speakers are routers that participate in MSDP and exchange information about multicast sources with other MSDP Speakers in different PIM domains. MSDP Speakers are usually located at the border of two different PIM domains.

2. **MSDP Peers**

MSDP Peers are MSDP Speakers that have established a peering relationship with each other. They exchange information about active multicast sources using the MSDP protocol.

3. **Multicast Source Active Advertisement (SA) messages**

MSDP uses SA messages to advertise the existence of active multicast sources to other MSDP Peers. SA messages are sent by MSDP Speakers and contain information about the multicast group address, the source address, and the RP (Rendezvous Point) associated with the multicast group.

4. **Rendezvous Point (RP)**

An RP is a designated router that acts as a root for a particular multicast group. MSDP uses RPs to distribute information about active multicast sources to other PIM domains. Each PIM domain can have one or more RPs, and each RP must be configured with a list of multicast groups that it is responsible for.

5. **Shared Tree**

MSDP allows for the creation of shared trees that connect multiple PIM domains. Shared trees are used to forward multicast traffic between PIM domains, and they are built by joining the multicast group at the RP.

Overall, MSDP is an important protocol for enabling multicast traffic to flow between different PIM domains. It provides a way for MSDP Peers to discover active multicast sources in other domains and to forward multicast traffic between domains using shared trees. MSDP helps to reduce unnecessary multicast traffic by allowing for the creation of shared trees and by limiting the scope of multicast traffic to only those PIM domains that have active receivers.

Configuring MSDP

Configuring MSDP involves the following steps:

1. Enable PIM on the routers that will participate in MSDP.
2. Configure the Rendezvous Points (RPs) in each PIM domain. The RP is the root of the shared tree for a particular multicast group.
3. Configure the MSDP peer relationships between MSDP Speakers in different PIM domains. Each MSDP Speaker must be configured with the IP address of its MSDP peers.
4. Configure the MSDP SA filters. SA filters determine which SA messages are forwarded by an MSDP Speaker. You can use SA filters to limit the scope of SA messages and reduce unnecessary multicast traffic.

Here's an example configuration for MSDP:

```
Router(config)# ip multicast-routing
Router(config)# interface Ethernet0/0
Router(config-if)# ippim sparse-mode
Router(config)# access-list 10 permit 239.1.1.0 0.0.0.255
Router(config)# access-list 10 permit 239.1.2.0 0.0.0.255
Router(config)# ippimrp-address 10.1.1.1 group-list 10
Router(config)# msdp peer 192.168.1.1 connect-source Loopback0
Router(config)# msdp originator-id Loopback0
Router(config)# msdpsa-filter 10 deny 239.1.2.0/24
```

In this configuration, we enable multicast routing on the router and enable PIM sparse-mode on the Ethernet0/0 interface. We then configure an access-list to permit traffic for the multicast groups 239.1.1.0/24 and 239.1.2.0/24.

Next, we configure the RP address for the multicast group 239.1.1.0/24 as 10.1.1.1 and apply the group-list 10 to limit the scope of the RP.

We then configure an MSDP peer relationship with the IP address 192.168.1.1 and specify the source interface as Loopback0. We also configure the MSDP originator ID as Loopback0.

Finally, we configure an SA filter to deny SA messages for the multicast group 239.1.2.0/24. This will prevent SA messages for this group from being forwarded by the MSDP Speaker.

Q18. Explain about MOSPF protocol.

Ans :

Meaning

Multicast Open Shortest Path First (MOSPF) is the extension of the Open Shortest Path First (OSPF) protocol, which is used in unicastrouting.

It also uses the sourcebased tree approach to multicasting. If the internet is running a unicast link-state routing algorithm, the idea can be extended to provide a multicast link-state routing algorithm.

Each router in the internet has a link-state database (LSDB) that can be used to create a shortest-path tree. To extend unicasting to multicasting, each router needs to have another database, as with the case of unicast distance-vector routing, to show which interface has an active member in a particular group.

Now a router goes through the following steps to forward a multicast packet received from source S and to be sent to destination G,

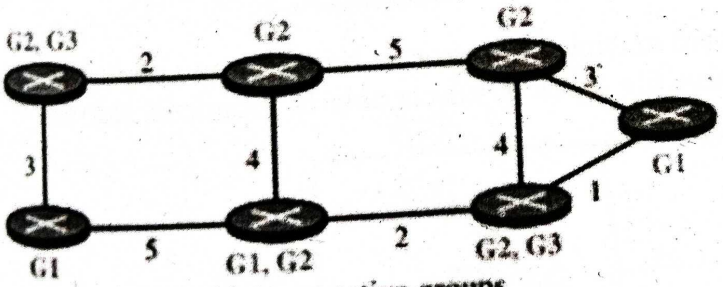
1. The router uses the Dijkstra algorithm to create a shortest-path tree with S (Source) as the root and all destinations in the internet as the leaves. Note that this shortest-path tree is different from the one the router normally uses for unicast forwarding, in which the root of the tree is the router itself. In this case, the root of the tree is the source of the packet defined in the source address of the packet. The router is capable of creating this tree because it has the LSDB, the whole topology of the internet; the Dijkstra algorithm can be used to create a tree with any root, no matter which router is using it. The point we need to remember is that the shortest-path tree created this way depends on the specific source. For each source we need to create a different tree.

2. The router finds itself in the shortest-path tree created in the first step. In other words, the router creates a shortest-path subtree with itself as the root of the subtree.

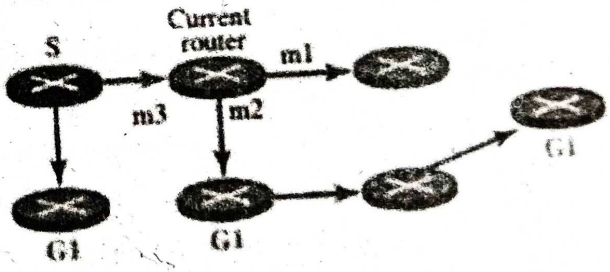
3. The shortest-path subtree is actually a broadcast subtree with the router as the root and all networks as the leaves. The router now uses a strategy similar to the one we describe in the case of DVMRP to prune the broadcast tree and to change it to a multicast tree. The IGMP protocol is used to find the information at the leaf level. MOSPF has added a new type of link state update packet that floods the membership to all routers. The router can use the information it receives in this way and prune the broadcast tree to make the multicast tree.

4. The router can now forward the received packet out of only those interfaces that correspond to the branches of the multicast tree. We need to make certain that a copy of the multicast packet reaches all networks that have active members of the group and that it does not reach those networks that do not.

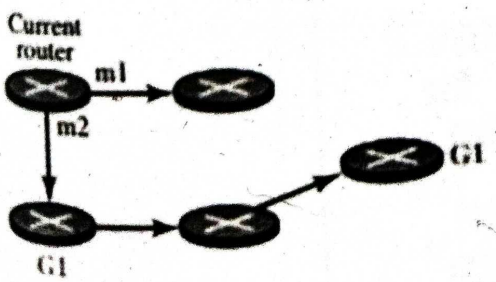
Below figure shows an example of using the steps to change a graph to a multicast tree. For simplicity, we have not shown the network, but we added the groups to each router. The figure shows how a source-based tree is made with the source as the root and changed to a multicast subtree with the root at the current router.



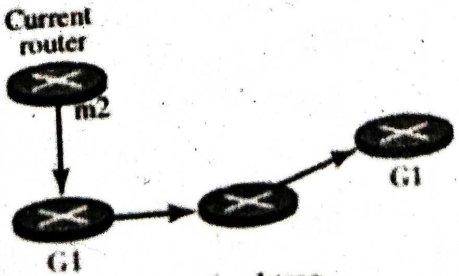
a. An internet with some active groups



b. S-G1 shortest-path tree



c. S-G1 subtree seen by current router



d. S-G1 pruned subtree

Forwarding table for current router

Group-Source	Interface
S, G1	m2
...	...

Example of tree formation in MOSPF

4.4.6 Spanning Tree Algorithm

Q19. Explain Spanning Tree algorithm working principle. **(Imp.)**

Ans :

Meaning

Spanning Tree Protocol (STP) is a communication protocol operating at data link layer the OSI model to prevent bridge loops and the resulting broadcast storms. It creates a loop " free topology for Ethernet networks.

Working Principle

A bridge loop is created when there are more than one paths between two nodes in a given network. When a message is sent, particularly when a broadcast is done, the bridges repeatedly rebroadcast the same message flooding the network. Since a data link layer frame does not have a time-to-live field in the header, the broadcast frame may loop forever, thus swamping the channels.

Spanning tree protocol creates a spanning tree by disabling all links that form a loop or cycle in the network. This leaves exactly one active path between any two nodes of the network. So when a message is broadcast, there is no way that the same message can be received from an alternate path. The bridges that participate in spanning tree protocol are often called spanning tree bridges.

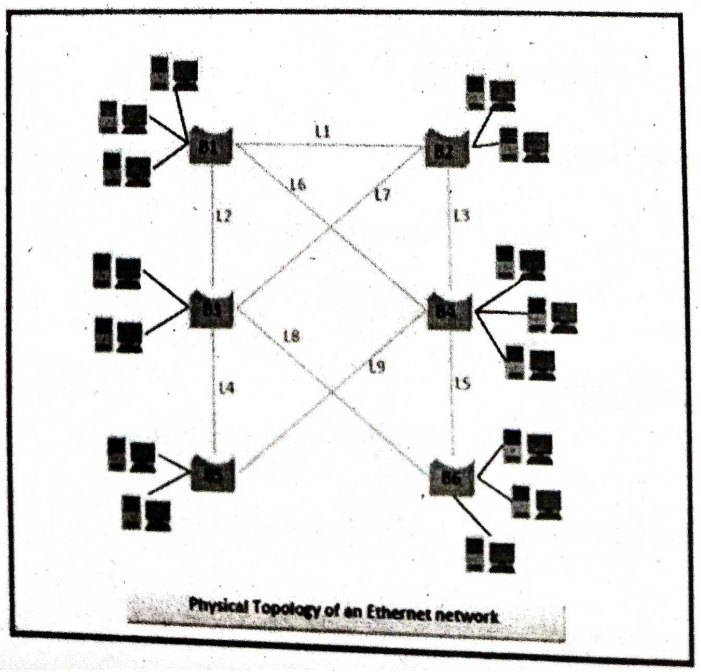
To construct a spanning tree, the bridges broadcast their configuration routes. Then they execute a distributed algorithm for finding out the minimal spanning tree in the network, i.e. the spanning tree with minimal cost. The links not included in this tree are disabled but not removed.

In case a particular active link fails, the algorithm is executed again to find the minimal spanning tree without the failed link. The communication continues through the newly formed spanning tree. When a failed link is restored, the algorithm is re-run including the newly restored link.

Example

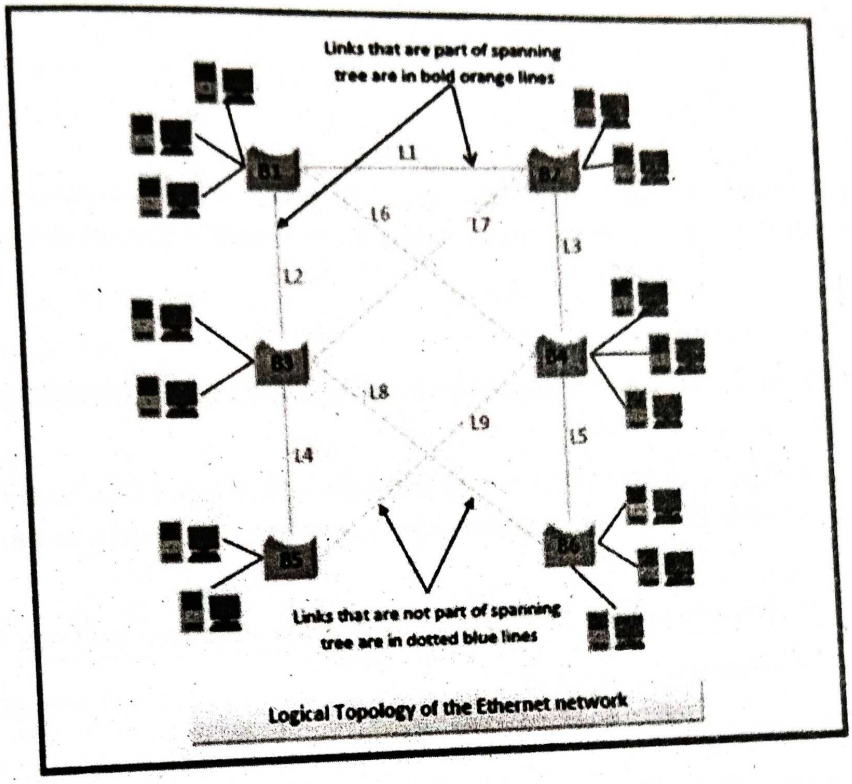
Let us consider a physical topology, as shown in the diagram, for an Ethernet network that comprises of six interconnected bridges. The bridges are named {B1, B2, B3, B4, B5, B6} and several nodes are connected to each bridge. The links between two bridges are named {L1, L2, L3, L4, L5, L6, L7, L8, L9}, where L1 connects B1 and B2, L2 connects B1 and B3 and so on. It is assumed that all links are of uniform costs.

From the diagram we can see that there are multiple paths from a bridge to any other bridge in the network, forming several bridge loops that makes the topology susceptible to broadcast storms.



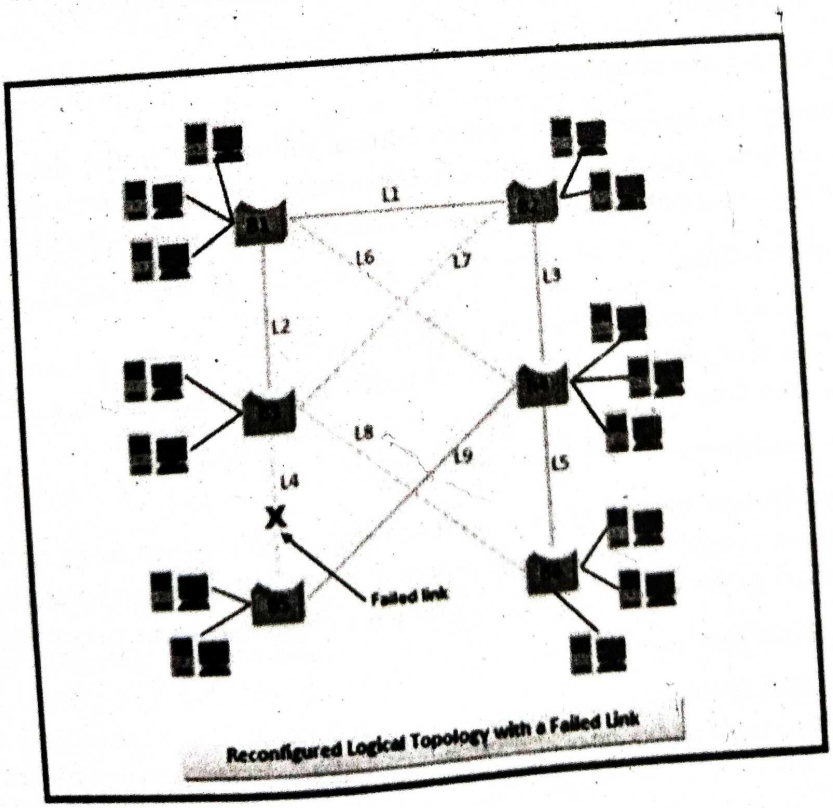
Physical Topology of an Ethernet network

According to spanning tree protocol, links that form a cycle are disabled. Thus, we get a logical topology so that there is exactly one route between any two bridges. One possible logical topology is shown in the following diagram below containing links {L1, L2, L3, L4, L5}-



Logical Topology of the Ethernet network

In the above logical configuration, if a situation arises such that link L4 fails. Then, the spanning tree is reconstituted leaving L4. A possible logical reconfiguration containing links {L1, L2, L3, L5, L9} is as follows -



Reconfigured Logical Topology with a Failed Link

Short Question & Answers

1. What is congestion control?

Ans :

Meaning

Congestion Control is a type of network layer issue, and it is concerned with what happens when there is more data in the network than can be sent with reasonable packet delays, and there are no lost packets.

Causes

The main cause of congestion is a huge amount of data traffic. But other factors are equally important for making congestion as given below:

1. Sudden arrival of large data (called burst data) from many input lines and trying to access a single output line of a router. In this case, the particular output line is blocked if its bandwidth isn't sufficiently high.
2. Low bandwidth line will produce congestion even if the data rate isn't too high.

2. Principles of Congestion Control.

Ans :

1. Monitoring network traffic

Network congestion can occur when there is more traffic on the network than the network can handle. Therefore, it is essential to monitor network traffic continuously to detect congestion before it becomes a problem.

2. Feedback-based mechanisms

Feedback-based mechanisms are used to control the rate of traffic flow and prevent congestion. These mechanisms involve sending feedback messages to the source of the traffic, indicating the current network conditions and the need to reduce the rate of traffic.

3. Resource allocation

Congestion control involves allocating network resources, such as bandwidth and buffer space, effectively. This ensures that each flow of traffic receives a fair share of network resources and prevents any one flow from monopolizing the network.

4. Congestion avoidance

Congestion avoidance techniques are used to prevent congestion from occurring in the first place. These techniques involve detecting and reacting to early signs of congestion, such as packet loss or delay, and reducing the rate of traffic to prevent congestion from occurring.

5. Traffic prioritization

Congestion control involves prioritizing traffic based on its importance and criticality. This ensures that critical traffic, such as voice or video traffic, is given priority over less critical traffic, such as file downloads.

3. **What is Routing?**

Ans :

Meaning

Routing is the process of directing network traffic from its source to its destination across a network. It is a fundamental concept in computer networking that enables devices on a network to communicate with each other by forwarding packets of data between them. The routing process involves the use of routing protocols and algorithms that determine the optimal path for data to travel from the source to the destination based on various factors, such as network topology, available bandwidth, and network congestion.

Routing can occur at different layers of the networking stack, including the physical layer, data link layer, network layer, and transport layer. At the network layer, routing is typically performed by devices such as routers, which use routing tables and algorithms to determine the best path for data to travel between networks.

4. **Write the differences between Inter domain and Intradomain Routing.**

Ans :

The following table highlights the major differences between interdomain and intradomain routing protocols.

S.No	Intradomain Routing	Interdomain Routing
1.	Routing algorithm works only within domains.	Routing algorithm works within and between domains.
2.	It need to know only about other routers within their domain.	It need to know only about other routers within and between their domain.
3.	Protocols used in intradomain routing are known as Interior-gate way protocols.	Protocols used in interdomain routing are known as Exterior-gateway protocols.
4.	In this Routing, routing takes place within an autonomous network.	In this Routing, routing takes place between the autonomous networks.
5.	Intradomain routing protocols ignores the internet outside the AS(autonomous system).	Interdomain routing protocol assumes that the internet contains the collection of interconnected AS(autonomous systems).

5. **RIP.**

Ans :

RIP stands for Routing Information Protocol. RIP is an intra-domain routing protocol used within an autonomous system. Here, intra-domain means routing the packets in a defined domain, for example, web browsing within an institutional area. To understand the RIP protocol, our main focus is to know the structure of the packet, how many fields it contains, and how these fields determine the routing tables.

- RIP is based on the distance vector-based strategy, so we consider the entire structure as a graph where nodes are the routers, and the links are the networks.

- In a routing table, the first column is the destination, or we can say that it is a network address.
- The cost metric is the number of hops to reach the destination. The number of hops available in a network would be the cost. The hop count is the number of networks required to reach the destination.

6. Disadvantages of RIP.

Ans :

- The RIP is a classful routing protocol, so it does not support the VLSM (Variable Length Subnet Mask). The classful routing protocol is a protocol that does not include the subnet mask information in the routing updates.
- It broadcasts the routing updates to the entire network that creates a lot of traffic. In RIP, the routing table updates every 30 seconds. Whenever the updates occur, it sends the copy of the update to all the neighbors except the one that has caused the update. The sending of updates to all the neighbors creates a lot of traffic. This rule is known as a split-horizon rule.
- It faces a problem of Slow convergence. Whenever the router or link fails, then it often takes minutes to stabilize or take an alternative route; This problem is known as Slow convergence.
- RIP supports maximum 15 hops which means that the maximum 16 hops can be configured in a RIP
- The Administrative distance value is 120 (Ad value). If the Ad value is less, then the protocol is more reliable than the protocol with more Ad value.

7. OSPF protocol.

Ans :

The OSPF stands for Open Shortest Path First. It is a widely used and supported routing protocol. It is an intradomain protocol, which means that it is used within an area or a network. It is an interior gateway protocol that has been designed within a single autonomous system. It is based on a link-state routing algorithm in which each router contains the information of every domain, and based on this information, it determines the shortest path. The goal of routing is to learn routes. The OSPF achieves by learning about every router and subnet within the entire network. Every router contains the same information about the network. The way the router learns this information by sending LSA (Link State Advertisements). These LSAs contain information about every router, subnet, and other networking information. Once the LSAs have been flooded, the OSPF stores the information in a link-state database known as LSDB.

8. Explain BGP protocol.

Ans :

It is an interdomain routing protocol, and it uses the path-vector routing. It is a gateway protocol that is used to exchange routing information among the autonomous system on the internet.

There are many versions of BGP, such as:

- BGP version 1: This version was released in 1989 and is defined in RFC 1105.
- BGP version 2: It was defined in RFC 1163.
- BGP version 3: It was defined in RFC 1267.
- BGP version 4: It is the current version of BGP defined in RFC 1771.

9. Distance Vector Multicast Routing Protocol.

Ans :

The distance vector multicast routing protocol is multicast routing protocol that takes the routing decision based upon the source address of the packet.

- This algorithm constructs the routing tree for a network.
- Whenever a router receives a packet, it forwards it to some of its ports based on the source address of packet.

2. It must prevent the formation of duplicate packets.
3. It must ensure that the path traveled by a packet is the shortest from its source to the router.
4. It should provide dynamic membership.

10. Reverse Path Forwarding.

Ans :

- In this strategy, the router only forwards those packets that have traveled the shortest path from source to destination.
- To achieve this, the router pretends that it has a packet to send to the source from where the packet has arrived.
- In this way, the shortest path to the sender of the packet is computed.
- If the same route is followed by the received packet, it is forwarded to the next router and it is discarded otherwise.
- The reverse path forwarding ensures that the network receives a copy of the packet without formation of loops. A loop occurs when a packet that has left the router may come back again from another interface or the same interface and be forwarded again.

11. Reverse Path Broadcasting.

Ans :

- In order to solve the problem, RPB is used.
- In this method, one parent router is defined for each network.
- The network could accept the multicast packets from this parent router only.
- This router sends packets to those ports for which it is designated as parent.
- Thus, RPB principle' allows a router to broadcast the packet in the network.

This creates duplicate packets on the network and reduces the network efficiency.

12. Multicast Open Shortest Path First.

Ans :

- Multicast open shortest path first is the multicast version of open shortest path first protocol.
- It is an extension of OSPF that uses multicast link state routing method to create source based trees.
- The method used by MOSPF is different from DVMRP.
- The first difference is that in this method, the tree is least cost tree instead of shortest path tree.
- The second difference is that the tree is not made gradually. It is made immediately it is prepared and ready to use.

13. PIM –DM protocol.

Ans :

PIM-DM (Protocol Independent Multicast - Dense Mode) is a multicast routing protocol that is used to efficiently distribute multicast traffic in a dense network with high bandwidth connectivity. It is a flood-and-prune protocol, which means that multicast traffic is initially flooded to all connected networks and then pruned back based on the network topology and multicast group membership.

In PIM-DM, multicast traffic is forwarded to all directly connected networks until it reaches a network where there are receivers interested in the traffic. Once a receiver is found, the multicast traffic is forwarded only to those networks that have receivers for that multicast group. PIM-DM builds a multicast distribution tree that is rooted at the source and extends to all receivers.

PIM-DM uses a reverse path forwarding (RPF) algorithm to prevent loops and to ensure that multicast traffic is forwarded in the direction of the root of the multicast tree. PIM-DM also uses a prune mechanism to remove branches of the multicast tree where there are no receivers.

Choose the Correct Answers

1. In _____ congestion control, policies are applied to prevent congestion before it happens. [a]
   (a) Open-loop (b) Closed-loop
   (c) Either (a) or (b) (d) Neither (a) nor (b)

2. In _____ we try to avoid traffic congestion. [a]
   (a) Congestion control (b) Quality of service
   (c) Either (a) or (b) (d) Both (a) and (b)

3. Which type of routing protocol is typically used within a single organization or company? [a]
   (a) Intra-domain routing protocol (b) Inter-domain routing protocol
   (c) Both a and b (d) None of the above

4. Among the following which is inter-domain protocol [c]
   (a) RIP (b) OSPF
   (c) BGP (d) All the above

5. Among the following which is multicast routing protocol [d]
   (a) RIP (b) OSPF
   (c) BGP (d) DVMRP

6. Amongth following which protocol uses path-vector routing [c]
   (a) RIP (b) OSPF
   (c) BGP (d) DVMRP

7. Which PIM mode is typically used in networks with a high density of multicast traffic? [a]
   (a) Dense mode (b) Sparse mode
   (c) Source-specific multicast (SSM) mode (d) Any-source multicast (ASM) mode

8. Which PIM mode is typically used in networks with a low density of multicast traffic? [b]
   (a) Dense mode (b) Sparse mode
   (c) Source-specific multicast (SSM) mode (d) Any-source multicast (ASM) mode

9. Which of the following is true about MOSPF routers? [c]
   (a) They are used to discover multicast sources in a single network domain
   (b) They are used to discover multicast sources across multiple network domains
   (c) They are used to establish the best path for multicast traffic to reach its destination
   (d) None of the above

10. What is the purpose of the Root Bridge in Spanning Tree Protocol? [c]
   (a) To serve as the central switch in the network
   (b) To ensure all switches are synchronized
   (c) To determine the shortest path to all switches in the network
   (d) None of the above

Fill in the blanks

1. _____is a technique that regulates the flow of network traffic by smoothing out bursts of traffic and ensuring that traffic flows within defined limits.
2. The packet sent by a node to the source to inform it of congestion is called _____
3. _____ is a technique used to prevent congestion by selectively dropping packets before congestion occurs
4. _____ is the process of directing network traffic from its source to its destination across a network.
5. When several routers are attached in a network, they are known as a _____ link.
6. Full form of MOSPF_____
7. The _____ protocol is a group-shared protocol that uses a core as the root of the tree.
8. _____ protocol for managing multicast traffic across multiple network domains
9. _____ is a communication protocol operating at data link layer the OSI model to prevent bridge loops and the resulting broadcast storms.
10. _____ protocol is used to discover multicast sources across multiple network domains

Answers

1. Traffic shaping
2. Choke
3. Random Early Detection
4. Routing
5. Transient
6. Multicast Open Shortest Path First
7. Core-Based Tree (CBT)
8. PIM
9. Spanning Tree Protocol (STP)
10. MSDP