Lecture 10

Network Layer: Routing protocols

0. Overview

In this lecture we will cover the following topics:

   16.1 Unicast routing
   16.2 Multicast routing
   16.3 Summary (part 16)
1. ROUTING PROTOCOLS

- Packets go from source to destination via routers.
- The router consults the routing table.
- Routing table can be:
  - static [does not change automatically] or
  - dynamic [changes automatically]
- Routing protocols are needed to create the routing tables dynamically.
- A routing protocol is a combination of rules and procedures that lets routers in the internet inform one another of changes. It allows routers to share whatever they know about the internet or their neighborhood.

16.1 UNICAST ROUTING

- In unicast routing, there is only one source and only one destination.
- When a router receives a packet, it forwards the packet through only one of its ports (the one belonging to the optimum path) as defined in routing table. It discards the packet, if there is no route.
**Metric of different protocols**

- Metric is the cost assigned for passing through a network.
  - The total metric of a particular router is equal to the sum of the metrics of networks that comprise the route.
  - A router chooses the route with smallest metric.

- **RIP** (Routing Information Protocol): Cost of passing each network is same; it is one hop count.
  - If a packet passes through 10 networks to reach the destination, the total cost is 10 hop counts.

- **OSPF** (Open Shortest Path First): Administrator can assign cost for passing a network based on type of service required.
  - OSPF allows each router to have more than one routing table based on required type of service.
  - Maximum throughput, minimum delay

- **BGP** (Border Gateway Protocol): Criterion is the policy, which is set by the administrator.

---

**Interior and Exterior routing**

- **Autonomous System**: Group of networks and routers under the authority of a single administration.
- Routers inside an autonomous system is referred to as interior routing.
- Routing between autonomous systems is referred to as exterior routing.
Autonomous systems

- Solid lines show the communication between routers that use interior routing protocols.
- Broken lines show the communication between routers that use an exterior routing protocols.

Routing Information Protocol (RIP)

- RIP is based on Distance vector routing.
- Distance vector routing
  - Sharing knowledge about the entire autonomous system: Each router periodically shares its knowledge about the entire autonomous system with its neighbours.
  - Sharing only with neighbours through all its interfaces.
  - Sharing at regular intervals: 30 seconds.
- Routing table
  - Has one entry for each destination network of which the router is aware.
  - Each entry has destination network address, the shortest distance to reach the destination in hop count, and next router to which the packet should be delivered to reach its final destination.
  - Hop count is the number of networks that a packet encounters to reach its final destination.
A distance vector routing table

<table>
<thead>
<tr>
<th>Destination</th>
<th>Hop Count</th>
<th>Next Router</th>
<th>Other information</th>
</tr>
</thead>
<tbody>
<tr>
<td>163.5.0.0</td>
<td>7</td>
<td>172.6.23.4</td>
<td></td>
</tr>
<tr>
<td>197.5.13.0</td>
<td>5</td>
<td>176.3.6.17</td>
<td></td>
</tr>
<tr>
<td>189.45.0.0</td>
<td>4</td>
<td>200.5.1.6</td>
<td></td>
</tr>
<tr>
<td>115.0.0.0</td>
<td>6</td>
<td>131.4.7.19</td>
<td></td>
</tr>
</tbody>
</table>

RIP Updating Algorithm

Receive: a response RIP message
1. Add one hop to the hop count for each advertised destination.
2. Repeat the following steps for each advertised destination:
   1. If (destination not in the routing table)
      1. Add the advertised information to the table.
   2. Else
      1. If (next-hop field is the same)
         1. Replace entry in the table with the advertised one.
      2. Else
         1. If (advertised hop count smaller than one in the table)
            1. Replace entry in the routing table.
   3. Return.
Example of updating a routing table

Initial routing tables in a small autonomous system

- When a router is added to a network, it initializes a routing table for itself, using its configuration file.
- The table consists only the directly attached networks and the hop counts, which are initialized to 1.
- The next-hop field, which identifies the next router, is empty.
Each routing table is updated upon receipt of RIP messages using the RIP updating algorithm.

OSPFW

- Open Shortest Path First
- Special routers called autonomous system boundary routers are responsible for dissipating information about other autonomous systems into the current system.
- OSPF divides an autonomous system into areas.
Areas in an Autonomous System

- **Area**: A collection of networks, hosts, and routers all contained within an autonomous system.
- Routers inside an area flood the area with routing information.
- **Area border routers**: Summarize the information about the area and send it to other routers.
- **Backbone area [Primary area]**: All the areas inside an autonomous system must be connected to the backbone. Routers in this area are called as backbone routers. This area identification number is 0.
- If, due to some problem, the connectivity between a backbone and an area is broken, a virtual link between routers must be created by the administration to allow continuity of the functions of the backbone as the primary area.
**OSPF**

- **Metric**
  - Administrator can assign the cost to each route.
  - Based on type of service (minimum delay, maximum throughput, and so on)

- **Link state routing**
  - Sharing knowledge about the neighbourhood: Each router sends the state of its neighbourhood to every other router in the area.
  - Sharing with every other router: By flooding, a process whereby a router sends its information to all its neighbours (through all its output ports). Each neighbour sends the packet to all its neighbours, and so on. Every router that receives the packet sends copies to each of its neighbours. Eventually, every router (without exception) has received a copy of the same information.
  - Sharing when there is a change; Only to its neighbours.
  - Each router should have the exact topology of the internet at every moment.
  - From this topology, a router can calculate the shortest path between itself and each network.

---

**Types of Links (1/5)**

- Point-to-point
- Transient
- Stub
- Virtual
**Point-to-Point Link (2/5)**

- Connects two routers without any other router or host in between.
- Directly connected routers using serial line.
- Only one neighbour.

![Diagram of Point-to-Point Link]

**Transient link (3/5)**

- A network with several routers attached to it.
- Each router has many neighbours.
- Lot of advertisements about their neighbours.
- One of the routers in the network has two duties: true router and designated router because we cannot connect each router to every other router through one single network. Each router has only one neighbour, the designated router (network). On the other hand, the designated router (network) has five neighbours.
- Designated router represents a network. There exists a metric between each node to the designated router but there is no metric from the designated router to any other node.

![Diagram of Transient Link]

![Unrealistic Representation]

![Realistic Representation]
Stub Link (4/5)

- A network that is connected to only one router.
- The data packets enter the network through this single router and leave the network through this same router.

Virtual link (5/5)

- When the link between two routers is broken, the administration may create a virtual link between them, using a longer path that probably goes through several routers.
Example of an internet & Graphical representation

Types of LSAs (1/6)

- To share information about their neighbours, each entity distributes Link State Advertisements (LSAs).

```
<table>
<thead>
<tr>
<th>Link state advertisements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Router link</td>
</tr>
<tr>
<td>Network link</td>
</tr>
<tr>
<td>Summary link to network</td>
</tr>
<tr>
<td>Summary link to AS boundary router</td>
</tr>
<tr>
<td>External link</td>
</tr>
</tbody>
</table>
```
**Router link (2/6)**

- A true router uses this advertisement to announce information about all its links and what is at the other side of the link (neighbours).

![Router diagram](image)

**Network link (3/6)**

- A designated router, on behalf of the transient network, distributes this type of LSA packet.
- The packet announces the existence of all the routers connected to the network.
• A router must also know about the networks outside its area, and the area border routers can provide this information.
• An area border router is active in more than one area.
• It receives router link and network link advertisements and creates a routing table for each area.

• If a router inside an area wants to send a packet outside the autonomous system, it should first know the route to an autonomous boundary router; the summary link to AS boundary router provides this information.
External link (6/6)

- A router inside an autonomous system wants to know which networks are available outside the autonomous system; the external link advertisement provides this information.
- The AS boundary router floods the autonomous system with the cost of each network outside the autonomous system, using a routing table created by a exterior routing protocol.

Dijkstra Algorithm

- Every router in the same area has the same 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.

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.
Shortest-path calculation

- The number next to each node represents the cumulative cost from the root node.
- Note that if a network can be reached through two directions with two cumulative costs, the direction with the smaller cumulative cost is kept, and the other one is deleted.
Link state routing table for router A

<table>
<thead>
<tr>
<th>Network</th>
<th>Cost</th>
<th>Next Router</th>
<th>Other Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>5</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>7</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td>10</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>N4</td>
<td>11</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>N5</td>
<td>15</td>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

BGP

- Border Gateway Protocol
- Inter-autonomous system routing protocol.
- BGP is based on a routing method called path vector routing.
Why D.V. and L.S. are not good enough?

- **Distance Vector routing**
  - Sometimes we don't want the route with smallest hop count as the preferred route [like, avoiding non-secure routes].
  - D.V routing information provides only the hop count and not the path that leads to that destination.
    - A router that receives a distance vector advertisement packet may be fooled if the shortest path is actually calculated through the receiving router itself.

- **Link State routing**
  - Internet is too big for this routing method
  - To use link state routing for the whole internet would require each router to have a huge link state database.
  - It would also take a long time for each router to calculate its routing table using the Dijkstra algorithm

- **Path Vector routing**
  - Each entry in the routing table contains the destination network, the next router, and the path to reach the destination.
  - The path is usually defined as an ordered list of autonomous systems that a packet should travel through to reach the destination.

---

Path vector routing table

<table>
<thead>
<tr>
<th>Network</th>
<th>Next Router</th>
<th>Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>N01</td>
<td>R01</td>
<td>AS14, AS23, AS67</td>
</tr>
<tr>
<td>N02</td>
<td>R05</td>
<td>AS22, AS67, AS05, AS89</td>
</tr>
<tr>
<td>N03</td>
<td>R06</td>
<td>AS67, AS89, AS09, AS34</td>
</tr>
<tr>
<td>N04</td>
<td>R12</td>
<td>AS62, AS02, AS09</td>
</tr>
</tbody>
</table>
Autonomous boundary routers that participate in path vector routing advertise the reachability of the networks in their own autonomous systems to neighbor autonomous boundary routers.

Concept of neighborhood here is the same as the one described in the RIP or OSPF protocol.

Two autonomous boundary routers connected to the same network are neighbours.

Each router that receives a path vector message verifies that the advertised path is in agreement with its policy (a set of rules imposed by the administrator controlling the routes). If it is, the router updates its routing table and modifies the message before sending it to the next neighbour.

The modification consists of adding its AS number to the path and replacing the next router entry with its own identification.

**Loop prevention**: Path vector avoids this problem by checking the path to see if its own AS is in the list.

**Policy Routing**: Check the AS in the path list against a policy. If it is against the policy, the router can ignore that path and that destination. It does not update its routing table with this path, and it does not send this message to its neighbors. So, routing table entry is not based on metric but on policy.
Path Attributes

- Path is a list of attributes
- Each attribute gives some information about the path
- List of attributes help the receiving router make a better decision when applying its policy.
- Two categories: well-known and optional
  - **Well-known**: Every BGP router should recognize
    - Mandatory
      - ORIGIN: source of routing information [RIP, OSPF, ...]
      - AS_PATH
      - NEXT_HOP
    - Discretionary: Not required to be included in every update message.
  - **Optional**: Need not be recognized by every router
    - Transitive: One that must be passed to the next router by the router that has not implemented this attribute
    - Non-transitive: One that should be discarded if the receiving router has not implemented it.

Types of BGP Messages

- Open: To create a neighborhood relationship
- If the neighbor accepts the neighborhood relationship, it responds with a keep-alive message, which means that a relationship has been established between two routers
- Update message is used by router to withdraw destinations that have been advertised previously, announce a router to a new destination, or do both.
- Keep-alive: Routers exchange this message regularly (before their hold time expires) to tell each other that they are alive.
- Notification: Sent by a router whenever an error condition is detected or a router wants to close the connection.
16.2 MULTICAST ROUTING

- One to many; Source is unicast address, but the destination is a group address (Class D)

- When a router receives a packet, it may forward it through several of its ports

- Router may discard the packet if it is not in the multicast path.

- **Flooding**: A router forwards a packet out of all its port except the one from which the packet came. Flooding provides broadcasting, but it also creates loops. A router will receive the same packet over and over from different ports. Several copies of the same packet are circulated, creating traffic jams.
IGMP

- Internet Group Management Protocol
- Group Management
  - IGMP is not a multicasting routing protocol
  - IGMP is a protocol that manages group membership.
  - In any network, there are one or more multicast routers that distribute multicast packets to hosts or other routers.
  - IGMP helps the multicast router create and update the list of groups in the network for which there is at least one loyal member.

IGMP message types
IGMP message format

- Type: 8-bit; Defines the type of message
  - General or special query: 0x11
  - Membership report: 0x16
  - Leave report: 0x17
- Maximum response time
  - 8-bit; Defines the amount of time in which a query must be answered; Value is in tenths of a second
- Checksum: 16-bit field carrying checksum calculated over 8-byte message.
- Group address: 0 for general query message. The value defines the groupid (multicast address of the group) in special query, the membership report and leave report messages.

IGMP type field

<table>
<thead>
<tr>
<th>Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>General or special query</td>
<td>0x11 or 00010001</td>
</tr>
<tr>
<td>Membership report</td>
<td>0x16 or 00010110</td>
</tr>
<tr>
<td>Leave report</td>
<td>0x17 or 00010111</td>
</tr>
</tbody>
</table>
IGMP operation

- IGMP operates locally.
- A multicast router connected to a network has a list of multicast addresses of the groups for which the router distributes packets to groups with at least one loyal member in that network.
- For each group, there is one router which has the duty of distributing the multicast packets destined for that group.
- A host or multicast router can have membership in a group. When a host has membership, it means that one of its processes (an application program) receives multicast packets from some group. When a router has membership, it means that a network connected to one of its other interfaces receives these multicast packets.
- In both cases, the host and the router keep a list of groupids and relay their interest to the distributing router.

Routers R1 & R2 may be distributors for some of the groups [given in router R] in other networks, but not on this network.

Membership report (1/2)

- A host or a router can join a group.
- A host maintains a list of processes that have membership in a group.
- When a process wants to join a new group, it sends its request to the host. The host adds the name of the process and the name of the requested group to its list.
- If this is the first entry for this particular group, the host sends a membership report message. If this is not the first entry, there is no need to send the membership report since the host is already a member of the group; it already receives multicast packets for this group.
Membership report (2/2)

- Router maintains a list of groupids that shows membership for the networks connected to each interface. When there is new interest in a group for any of these interfaces, the router sends out a membership report. This report is sent out of all interfaces except the one from which the new interest comes.

- Membership report is sent twice, one after the other within a few minutes. If the first one is lost or damaged, the second one replaces it.

---

Leave report

- Hosts send a leave report when there is no process interested in a specific group.
- When a multicast router receives a leave report, it cannot immediately purge that group from its list because the report comes from just one host or a router.
- Multicast router generates a specific query with specific groupid to identify whether the group can be purged or not. If no response within the specified response time, the group can be purged from the list.
Monitoring membership

- Multicast router monitors all the hosts or routers in a LAN to see if they want to continue their membership in a group.
- What happens: A case where a only alive host shuts down without sending the leave report.
- Router periodically (by default, every 125sec) sends a general query message. In this message, the group address field is set to 0.0.0.0. This means the query for membership continuation is for all groups in which a host is involved, not just one.
- Query message is sent by only one router (normally called query router) to prevent unnecessary traffic.
- The router expects reply for each group within the maximum response time of 10 sec.
- When a host or router receives the general query message, it responds with a membership report if it is interested in a group. If there is a common interest (two hosts, for example, are interested in the same group), only one response is sent for that group to prevent unnecessary traffic.

General query message

- Host or Router sends a General Query message. The message contains the group address field set to 0.0.0.0.
- Router expects a reply from each host or router within the maximum response time.
- Host or Router responds with a Membership Report if it is interested in the group. If there is a common interest, only one response is sent.
- No Response is indicated if there is no interest in the group.
**Delayed Response**

- When a host or router receives a query message, it does not respond immediately; it delays the response.
- Each host or router uses a random number to create a timer, which expires between 1 and 10 seconds.
- The expiration time can be in steps in 1s or less.
- Each group in the list has its own timer.
- Each host or router waits until its timer has expired before sending a membership report message.
- As the membership report is a broadcast, the waiting host or router receives the report and knows that there is no need for duplication of report message by many hosts.

---

**Example**

Imagine there are three hosts in a network, as shown in figure below. A query message was received at time 0; the random delay time (in tenths of seconds) for each group is shown next to the group address. Show the sequence of report messages.
Solution

The events occur in this sequence:

1. **Time 12.**
   The timer for 228.42.0.0 in host A expires and a membership report is sent, which is received by the router and every host including host B which cancels its timer for 228.42.0.0.

2. **Time 30.**
   The timer for 225.14.0.0 in host A expires and a membership report is sent, which is received by the router and every host including host C which cancels its timer for 225.14.0.0.

3. **Time 50.**
   The timer for 251.71.0.0 in host B expires and a membership report is sent, which is received by the router and every host.

4. **Time 70.**
   The timer for 230.43.0.0 in host C expires and a membership report is sent, which is received by the router and every host including host A which cancels its timer for 230.43.0.0.

Multicast Trees

- **Objectives of Multicasting are**
  - Each member of the group should receive one, and only one, copy of the multicast packet. Receipt of multiple copies is not allowed.
  - Nonmembers must not receive a copy.
  - There must be no loops in routing; that is, a packet must not visit a router more than once.
  - The path traveled from the source to each destination must be optimal (the shortest path).

- **Source-Based Tree**
  - A single tree is made for each combination of source and group.
    - MOSPF, DVMRP, PIM-DM.

- **Group-Shared Tree**
  - Each group in the system shares the same tree.
  - Tree changes when the group changes but remains the same when the group remains the same.
  - Group determines the tree and not the source.
  - Approaches to find multicast tree: Steiner tree [only theoretical], rendezvous-point tree.
    - CBT, PIM-SP
MBONE

- Only a small fraction of Internet routers are multicast routers.
- A multicast router may not find another multicast router in the neighborhood to forward the multicast packet.
- Tunneling helps to connect the multicast routers logically. Routers enclosed in broken circles are capable of multicasting. To enable multicasting, we make a multicast backbone (MBONE) out of these isolated routers, using the concept of tunneling.

Logical tunnel is established by encapsulating the multicast packet inside a unicast packet.
- The intermediate (nonmulticast) routers forward the packet as unicast routers and deliver the packet from one island to another.
- DVMRP supports both MBONE and tunneling.
Multicast routing protocols

- **DVMRP (Distance Vector Multicast Routing Protocol)**
  - Source-based routing protocol
  - Formation of shortest-path tree
    - No router knows the complete route for a particular destination. Each router knows from which port to send out a unicast packet on the destination address.
    - Optimal tree is determined while the packet travels. When a router receives a packet, the router forwards the packet through some of the ports, based on the source address, and contributes to the formation of the tree; the rest of the tree is made by other down-stream routers.

- **This protocol should accomplish the following:**
  - Must prevent the formation of loops
  - Must prevent duplications; no network receives more than one copy. In addition, the path traveled by a copy is the shortest path from the source to the destination.
  - Must provide for dynamic membership.

- **Reverse Path Forwarding (RPF)**
  - A router forwards the copy that has traveled the shortest path from the source to the router.
  - To find if the packet has traveled the shortest path, RPF uses the unicast routing table of RIP.
  - It pretends that it needs to send a packet to the source and finds if the port given by the routing table is the same from which the packet has arrived.
Reverse Path Forwarding (RPF)

In RPF, the router forwards only the packets that have traveled the shortest path from the source to the router; all other copies are discarded. RPF prevents the formation of loops.

RPF versus RPB

In RPF, looping is avoided but does not guarantee the receipt of only one copy. This is because the packet is forwarded based on source address and not on destination address.

To eliminate duplication, we must define only one parent router for each network. A network can receive a multicast packet from a particular source only through a designated parent router.

For each source, the router sends the packet only out of these ports for which it is the designated parent. This policy is called Reverse Path Broadcasting (RPB). RPB guarantees that the packet reaches every network and that every network receives only one copy.

Select the router with the shortest path to the source as the designated parent router.

RPB creates a shortest-path broadcast tree from the source to each destination. It guarantees that each destination receives one and only copy of the packet.
Reverse Path Multicasting (RPM)

- RPB does not multicast the packet, it broadcasts it.
- To be efficient, the multicast packet must reach only those networks that have active members for that particular group.
- In DVMRP, the first packet is broadcast to every network. The remaining packets is based on pruning and grafting. This is called as RPM.
- **Pruning**: Procedure that stops the sending of messages from an interface.
- **Grafting**: Procedure that resumes the sending of multicast messages from an interface.
- Pruning and Grafting are done by IGMP.
- RPM adds pruning and grafting to RPB to create a multicast shortest-path tree that supports dynamic membership changes.

---

RPF, RPB, and RPM

- a. RPF
- b. RPB
- c. RPM (after pruning)
- d. RPM (after grafting)
MOSPF

- **Multicast Open Shortest Path First**
  - Uses multicast link state routing to create source-based trees.
  - First, the tree is a least-cost tree (using a metric) instead of a shortest-path tree.
  - Second, the tree is made all at once instead of gradually (the tree is said to be premade, prepruned, and ready to be used).
- Least-Cost Trees
  - Each router knows the entire topology of the network.
  - Each router uses Dijkstra’s algorithm to create a least-cost tree that has the router as the root and the rest of the routers as nodes of the tree.
  - Least-cost trees in MOSPF is different for each router.

Unicast tree and multicast tree

- In multicast routing, we need one tree for each source-group pair, and the root must be the source.
- This is done using the database by asking the router to use Dijkstra’s algorithm to create a tree with the source as the root.
- Three problems exist:
  - Algorithm uses unicast addresses but the tree we need requires group addresses.
  - Membership can change frequently.
  - Applying Dijkstra algorithm for each multicast packet is very expensive timewise.
**Solution to the problems**

- Add a new link state update packet to associate the unicast address of a host with the group address or addresses the host is sponsoring. It is called a group membership LSA.
  - We make a tree that contains all the hosts belonging to a group, but we use the unicast address of the host in the calculation.

- Link state packets can also solve the second problem if they are sent whenever there is a change in the membership.

- The router can calculate the least-cost trees on demand (when it receives the first multicast packet). In addition, the tree can be saved in the cache memory for future use by the same source-group pair. MOSPF is a data-driven protocol.

---

**CBT**

- **Core-Based Tree**
  - Group-shared protocol that uses a core as the root of the tree.
  - Autonomous system is divided into regions, and a core (center router or rendezvous router) is chosen for each region.
  - Formation of tree
    - After rendezvous router is selected, every router is informed of the unicast address of the selected router.
    - All routers send a unicast join message that passes through all routers that are located between sender and rendezvous router.
    - Each intermediate router extracts the necessary information from the message, such as the unicast address of the sender and the port through which the packet has arrived, and forwards the message to the next router in the path.
    - When the rendezvous router has received all join messages from every member of the group, the tree is formed. Now every router knows its upstream and downstream router.
Shared-group tree with rendezvous router

- If a router wants to leave the group, it sends a leave message to its upstream router. The upstream router removes the link to that router from the tree and forwards the message to the upstream router, and so on.

DVMRP, MOSPF & CBT

<table>
<thead>
<tr>
<th>DVMRP and MOSPF</th>
<th>CBT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree is made from the root</td>
<td>Tree is made from the leaves</td>
</tr>
</tbody>
</table>

| | |
| The tree is first made (broadcasting) and then pruned. | There is no tree in the beginning; the joining (grafting) gradually makes the tree. |
Sending a multicast packet to the rendezvous router

- A multicast packet is send from source to rendezvous router and it forwards the message to all members of the group.
- Packet from source to members of group as below:
  - Source [may be or may be part of the tree] encapsulates the multicast packet inside a unicast packet with the unicast destination address of the core and sends it to the core. This part of delivery is done using a unicast address; the only recipient is the core router.
  - Core decapsulates the unicast packet and forwards it to all “interested” ports, which is part of the tree and is not pruned by IGMP.
  - Each router that receives the multicast packet, in turn, forwards it to all interested ports.

PIM (Protocol Independent Multicast)

- PIM-DM & PIM-SM are two independent multicast routing protocols, which are unicast-protocol-dependent.
- PIM-DM (Dense Mode)
  - Unicast protocol dependent
  - Used when there is a possibility that each router is involved in multicasting
  - Use of broadcast is justified because almost all routers are involved in the process.
  - Source-based routing protocol that uses RPF and pruning/grafting strategies for multicasting
  - Operation is like DVMRP but unicast protocol independent.
  - It assumes that the autonomous system is using a unicast protocol and each router has a table that can find the outgoing port that has an optimal path to a destination. This unicast protocol can be a distance vector protocol (RIP) or link state protocol (OSPF).
**PIM-SM (Sparse mode)**

- Used when there is a slight possibility that each router is involved in multicasting.
- Use of protocol that broadcasts is not justified.
- Protocol like CBT that uses a group-shared tree is more appropriate.
- A group-shared routing protocol that has a rendezvous point (RP) as the source of the tree.
- Like CBT but does not require acknowledgement from a join message. In addition, it creates a backup set of RPs for each region to cover RP failures.
- PIM-SM can switch from group-shared tree to source-based tree strategy if necessary. This can happen if there is a dense area of activity far from the RP.

**Multicasting is applied in distributed databases, information dissemination, distance learning, and particularly multimedia communications.**

---

**16.3 SUMMARY (part 16) [1/3]**

- A metric is the cost assigned for passage of a packet through a network.
- A router consults its routing table to determine the best path for a packet.
- An autonomous system (AS) is a group of networks and routers under the authority of a single administration.
- RIP and OSPF are popular interior routing protocols used to update routing tables in an AIS.
- RIP is based on distance vector routing, in which each router shares, at regular intervals, its knowledge about the entire AS with its neighbor.
- A RIP routing table entry consists of a destination network address, the hop count to that destination, and the IP address of the next router.
- OSPF divides an AS into areas, defined as collections of networks, hosts, and routers.
- OSPF is based on link state routing, in which each router sends the state of its neighborhood to every other router in the area. A packet is sent only if there is a change in the neighborhood.
- OSPF defines four types of links (networks): point-to-point, transient, stub, and virtual.
16.3 SUMMARY (part 16) [2/3]

- Five types of link state advertisements (LSAs) disperse information in OSPF: router link, network link, summary link to network, summary link to AS boundary router, and external link.
- A router compiles all the information from the LSAs it receives into a link state database. This database is common to all routers in an area.
- An LSA is a multfield entry in a link state update packet.
- BGP is an interautonomous system routing protocol used to upadate routing tables.
- BGP is based on a routing method called path vector routing. In this method, the ASs through which a packet must pass are explicitly listed.
- There are four types of BGP messages: open, update, keep-alive, and notification.
- The Internet Group Management Protocol (IGMP) helps multicast routers create and update a list of loyal members related to a router interface.
- The three IGMP message types are the query message, the membership report, and the leave report.
- A host or router can have membership in a group.
- A host maintains a list of processes that have membership in a group.
- A router maintains a list of group-ids that shows group membership for each interface.

16.3 SUMMARY (part 16) [3/3]

- Multicasting applications include distributed databases, information dissemination, teleconferencing, and distance learning.
- For efficient multicasting we use a shortest-path spanning tree to represent the Communication path.
- In a source-based tree approach to multicast routing, the source-group combinations determines the tree.
- In a group-based tree approach to multicast routing, the group determines the tree.
- DVRMP is a multicast routing protocol that uses the distance routing protocol to create a source-based tree.
- In reverse path forwarding (RPF), the router forwards only the packets that have traveled the shortest path from the source to the router.
- Reverse path broadcasting (RPB) creates a shortest-path broadcast tree from the source to each destination. It guarantees that each destination receives one and only one copy of the packet.
- Reverse path multicasting (RPM) adds pruning and grafting to RPB to create a multicast shortest-path tree that supports dynamic membership changes.
- MOSPF is a multicast protocol that uses multicast link state routing to create a source-based least-cost tree.
- The Core-Based Tree (CBT) protocol is a multicast routing protocol that uses a core as the root of the tree.
- PIM-DM is a source-based routing protocol that uses RPF and pruning and grafting strategies to handle multicasting.
- PIM-SM is a group-shared routing protocol that is similar to CBT and uses a rendezvous point as the source of the tree.
- For multicasting between two noncontiguous multicast routers, we make a multicast backbone (MBONE) to enable tunneling.
References


