Data Communications and Networking Fourth Edition

Forouzan

Chapter 22

Network Layer: Delivery, Forwarding, and Routing

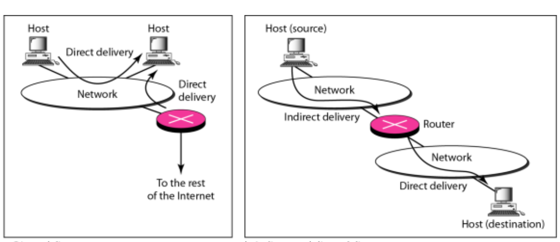
22-1 DELIVERY

The network layer supervises the handling of the packets by the underlying physical networks. We define this handling as the delivery of a packet.

Topics discussed in this section:

Direct Versus Indirect Delivery

Figure 22.1 ***Direct and indirect delivery***



a. Direct delivery b. Indirect and direct delivery

- Direct delivery occurs when the source and destination of the packet are located on the same physical network or when the delivery is between the last router and the destination host.
- If the destination host is not on the same network as the deliverer, the packet is delivered indirectly.
- In an indirect delivery, the packet goes from router to router until it reaches the one connected to the same physical network as its final destination.

22-2 FORWARDING

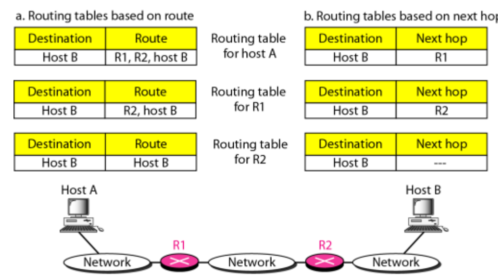
Forwarding means to place the packet in its route to its destination. Forwarding requires a host or a router to have a routing table. When a host has a packet to send or when a router has received a packet to be forwarded, it looks at this table to find the route to the final destination.

Topics discussed in this section:

Forwarding Techniques
Forwarding Process
Routing Table

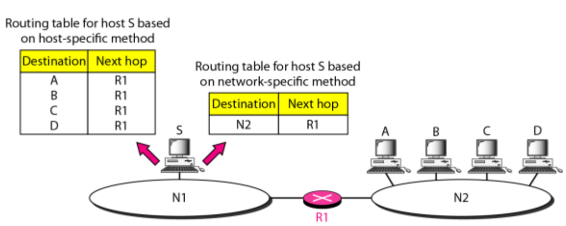
22.5

Figure 22.2 ***Route method versus next-hop method***



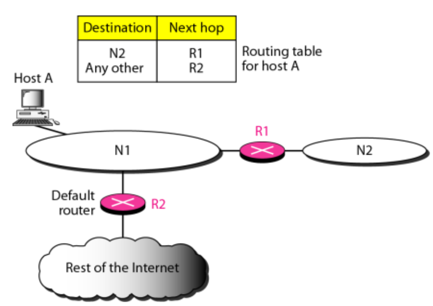
22.6

Figure 22.3 ***Host-specific versus network-specific method***



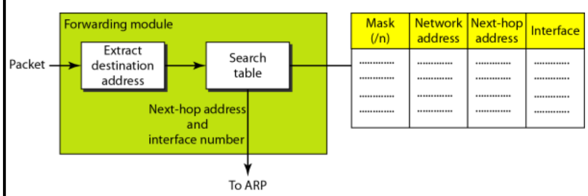
22.7

Figure 22.4 ***Default method***



22.8

Figure 22.5 *Simplified forwarding module in classless address*



22.9

- In classless addressing, the routing table needs to have one row of information for each block involved.
- The table needs to be searched based on the network address (first address in the block). Unfortunately, the destination address in the packet gives no clue about the network address. To solve the problem, we need to include the mask *(/n)* in the table; we need to have an extra column that includes the mask for the corresponding block.

22.10

Note

In classless addressing, we need at least four columns in a routing table.

22.11

Example 22.1

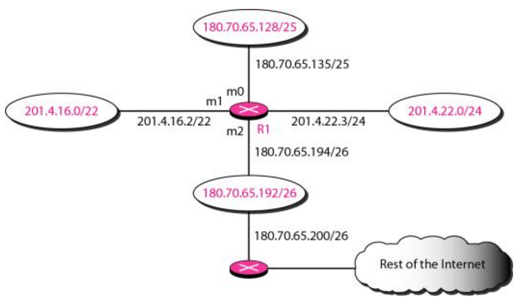
Make a routing table for router R1, using the configuration in Figure 22.6.

Solution

Table 22.1 shows the corresponding table.

22.12

Figure 22.6 *Configuration for Example 22.1*



22.13

Table 22.1 *Routing table for router R1 in Figure 22.6*

Mask	*Network Address*	*Next Hop*	*Interface*
/26	180.70.65.192	—	m2
/25	180.70.65.128	—	m0
/24	201.4.22.0	—	m3
/22	201.4.16.0	....	m1
Any	Any	180.70.65.200	m2

22.14

Example 22.2

Show the forwarding process if a packet arrives at R1 in Figure 22.6 with the destination address 180.70.65.140.

Solution

The router performs the following steps:

1. ***The first mask (/26) is applied to the destination address. The result is 180.70.65.128, which does not match the corresponding network address.***
2. ***The second mask (/25) is applied to the destination address. The result is 180.70.65.128, which matches the corresponding network address. The next-hop address and the interface number m0 are passed to ARP for further processing.***

22.15

Example 22.3

Show the forwarding process if a packet arrives at R1 in Figure 22.6 with the destination address 201.4.22.35.

Solution

The router performs the following steps:

1. ***The first mask (/26) is applied to the destination address. The result is 201.4.22.0, which does not match the corresponding network address.***
2. ***The second mask (/25) is applied to the destination address. The result is 201.4.22.0, which does not match the corresponding network address (row 2).***

22.16

Example 22.3 (continued)

3. The third mask (/24) is applied to the destination address. The result is 201.4.22.0, which matches the corresponding network address. The destination address of the packet and the interface number m3 are passed to ARP.

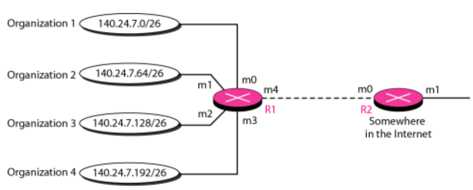
Example 22.4

Show the forwarding process if a packet arrives at R1 in Figure 22.6 with the destination address 18.24.32.78.

Solution

This time all masks are applied, one by one, to the destination address, but no matching network address is found. When it reaches the end of the table, the module gives the next-hop address 180.70.65.200 and interface number m2 to ARP. This is probably an outgoing package that needs to be sent, via the default router, to someplace else in the Internet.

Figure 22.7 *Address aggregation*



Mask	Network address	Next-hop address	Interface
/26	140.24.7.0	----------	m0
/26	140.24.7.64	----------	m1
/26	140.24.7.128	----------	m2
/26	140.24.7.192	----------	m3
/0	0.0.0.0	Default	m4

Routing table for R1

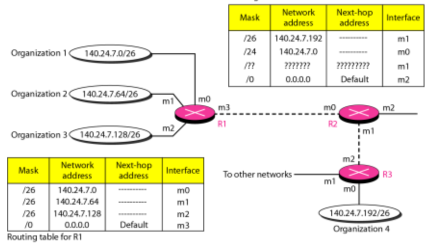
Mask	Network address	Next-hop address	Interface
/24	140.24.7.0	----------	m0
/0	0.0.0.0	Default	m1

Routing table for R2

Figure 22.8 *Longest mask matching*

Routing table for R2

Mask	Network address	Next-hop address	Interface
/26	140.24.7.192	----------	m1
/24	140.24.7.0	----------	m0
/??	???????	?????????	m1
/0	0.0.0.0	Default	m2



Mask	Network address	Next-hop address	Interface
/26	140.24.7.0	----------	m0
/26	140.24.7.64	----------	m1
/26	140.24.7.128	----------	m2
/0	0.0.0.0	Default	m3

Routing table for R1

Mask	Network address	Next-hop address	Interface
/26	140.24.7.192	----------	m0
/??	???????	?????????	m1
/0	0.0.0.0	Default	m2

Routing table for R3

Example 22.5

As an example of hierarchical routing, let us consider Figure 22.9. A regional ISP is granted 16,384 addresses starting from 120.14.64.0. The regional ISP has decided to divide this block into four subblocks, each with 4096 addresses. Three of these subblocks are assigned to three local ISPs; the second subblock is reserved for future use. Note that the mask for each block is /20 because the original block with mask /18 is divided into 4 blocks.

The first local ISP has divided its assigned subblock into 8 smaller blocks and assigned each to a small ISP. Each small ISP provides services to 128 households, each using four addresses.

22.21

Example 22.5 (continued)

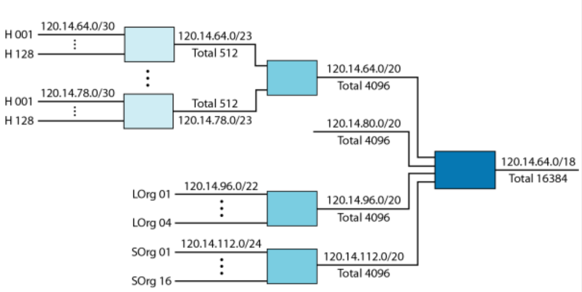
The second local ISP has divided its block into 4 blocks and has assigned the addresses to four large organizations.

The third local ISP has divided its block into 16 blocks and assigned each block to a small organization. Each small organization has 256 addresses, and the mask is /24.

There is a sense of hierarchy in this configuration. All routers in the Internet send a packet with destination address 120.14.64.0 to 120.14.127.255 to the regional ISP.

22.22

Figure 22.9 *Hierarchical routing with ISPs*



22.23

Routing Table

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

A **static routing table** contains information entered manually. The administrator enters the route for each destination into the table. When a table is created, it cannot update automatically when there is a change in the Internet.

22.24

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

Figure 22.10 *Common fields in a routing table*

Mask	Network address	Next-hop address	Interface	Flags	Reference count	Use
....	....	....	....	....	....	....

- Flags - This field defines up to five flags. Flags are on/off switches that signify either presence or absence.
- The five flags are U (up), G (gateway), H (host-specific), D (added by redirection), and M (modified by redirection).

Example 22.6

One utility that can be used to find the contents of a routing table for a host or router is netstat in UNIX or LINUX. The next slide shows the list of the contents of a default server. We have used two options, r and n. The option r indicates that we are interested in the routing table, and the option n indicates that we are looking for numeric addresses. Note that this is a routing table for a host, not a router. Although we discussed the routing table for a router throughout the chapter, a host also needs a routing table.

Example 22.6 (continued)

$ netstat -rn
Kernel IP routing table

Destination	Gateway	Mask	Flags	Iface
153.18.16.0	0.0.0.0	255.255.240.0	U	eth0
127.0.0.0	0.0.0.0	255.0.0.0	U	lo
0.0.0.0	153.18.31.254	0.0.0.0	UG	eth0

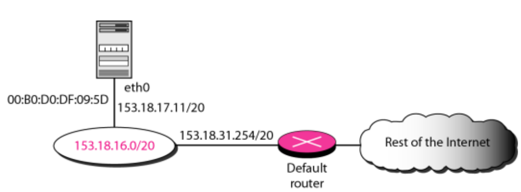
The destination column here defines the network address. The term gateway used by UNIX is synonymous with router. This column actually defines the address of the next hop. The value 0.0.0.0 shows that the delivery is direct. The last entry has a flag of G, which means that the destination can be reached through a router (default router). The Iface defines the interface.

Example 22.6 (continued)

More information about the IP address and physical address of the server can be found by using the ifconfig command on the given interface (eth0).

```
$ ifconfig eth0
eth0   Link encap:Ethernet  HWaddr 00:B0:D0:DF:09:5D
inet addr:153.18.17.11  Bcast:153.18.31.255  Mask:255.255.240.0
. . .
```


Figure 22.11 *Configuration of the server for Example 22.6*



22-3 UNICAST ROUTING PROTOCOLS

A routing table can be either static or dynamic. A static table is one with manual entries. A dynamic table is one that is updated automatically when there is a change somewhere in the Internet. A routing protocol is a combination of rules and procedures that lets routers in the Internet inform each other of changes.

Topics discussed in this section:

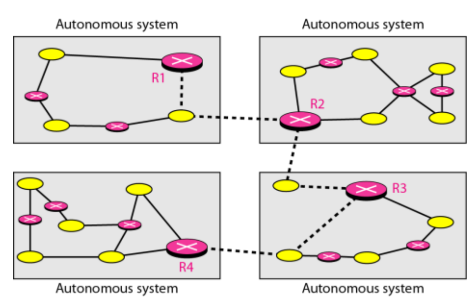
Optimization
Intra- and Interdomain Routing
Distance Vector Routing and RIP
Link State Routing and OSPF
Path Vector Routing and BGP

- Routing protocols have been created in response to the demand for dynamic routing tables.
- A routing protocol is a combination of rules and procedures that lets routers in the internet inform each other of changes. It allows routers to share whatever they know about the internet or their neighbourhood.

- **Optimization**
- A router receives a packet from a network and passes it to another network. A router is usually attached to several networks. When it receives a packet, to which network should it pass the packet?
- The decision is based on optimization: Which of the available pathways is the optimum pathway? What is the definition of the term *optimum?*

- One approach is to assign a cost for passing through a network. We call this cost a metric.
- (RIP), treat all networks as equals. The cost of passing through a network is the same; it is one hop count.
- In Open Shortest Path First (OSPF) , A route through a network can have different costs (metrics).

Figure 22.12 ***Autonomous systems***



Intra- and Interdomain Routing

- An autonomous system (AS) is a group of networks and routers under the authority of a single administration.
- Routing inside an autonomous system is referred to as intradomain routing.

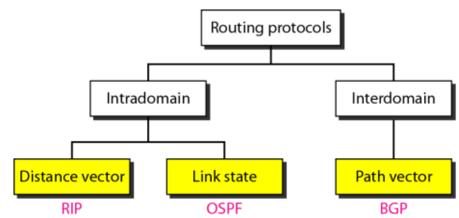
Routing between autonomous systems is referred to as interdomain routing.

Each autonomous system can choose one or more intradomain routing protocols to handle routing inside the autonomous system.

However, only one interdomain routing protocol handles routing between autonomous systems

22.37

Figure 22.13 *Popular routing protocols*



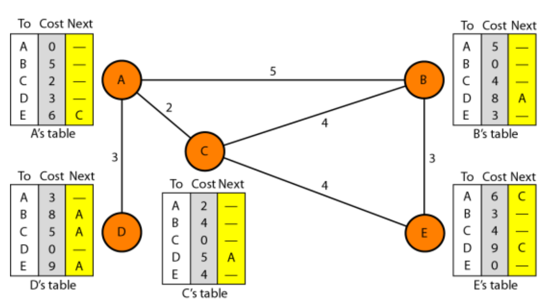
22.38

Distance Vector Algorithm

- In distance vector routing, the least-cost route between any two nodes is the route with minimum distance. In this protocol, as the name implies, each node maintains a vector (table) of minimum distances to every node.
- The table at each node also guides the packets to the desired node by showing the next stop in the route (next-hop routing).

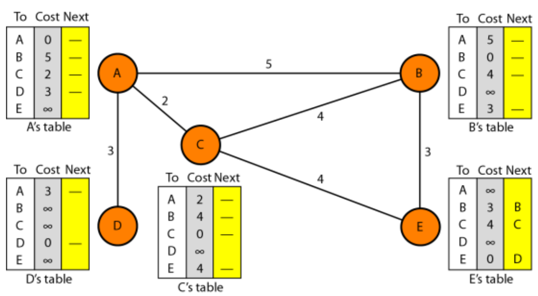
22.39

Figure 22.14 *Distance vector routing tables*



22.40

Figure 22.15 *Initialization of tables in distance vector routing*



22.41

- *Initialization*
- Each node can know only the distance between itself and its immediate neighbors, those directly connected to it.
- So for the moment, we assume that each node can send a message to the immediate neighbors and find the distance between itself and these neighbors.

22.42

- *Sharing*
- Although node A does not know about node E, node C does. So if node C shares its routing table with A, node A can also know how to reach node E. On the other hand, node C does not know how to reach node D, but node A does. If node A shares its routing table with node C, node C also knows how to reach node D. In other words, nodes A and C, as immediate neighbors, can improve their routing tables if they help each other.

22.43

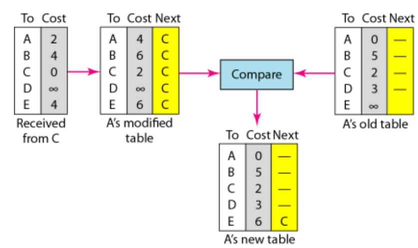
- A node therefore can send only the first two columns of its table to any neighbour.
- In other words, sharing here means sharing only the first two columns.

22.44

Note

In distance vector routing, each node shares its routing table with its immediate neighbors periodically and when there is a change.

Figure 22.16 *Updating in distance vector routing*



- *Updating*
- When a node receives a two-column table from a neighbor, it needs to update its routing table. Updating takes three steps:
- 1. The receiving node needs to add the cost between itself and the sending node to each value in the second column. The logic is clear. If node C claims that its
- distance to a destination is x mi, and the distance between A and C is y mi, then the distance between A and that destination, via C, is $x + y$ mi.

- 2. The receiving node needs to add the name of the sending node to each row as the third column if the receiving node uses information from any row.
- The sending node is the next node in the route.
- 3. The receiving node needs to compare each row of its old table with the corresponding row of the modified version of the received table.

- a. If the next-node entry is different, the receiving node chooses the row with the smaller cost. If there is a tie, the old one is kept.
- b. If the next-node entry is the same, the receiving node chooses the new row. For example, suppose node C has previously advertised a route to node X with distance 3. Suppose that now there is no path between C and X; node C now advertises this route with a distance of infinity. Node A must not ignore this value even though its old entry is smaller. The old route does not exist any more. The new route has a distance of infinity.

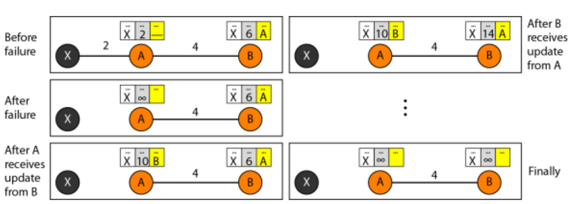
22.49

When to Share

- Periodic Update A node sends its routing table, normally every 30 s, in a periodic update. The period depends on the protocol that is using distance vector routing.
- Triggered Update A node sends its two-column routing table to its neighbors anytime there is a change in its routing table. This is called a triggered update.

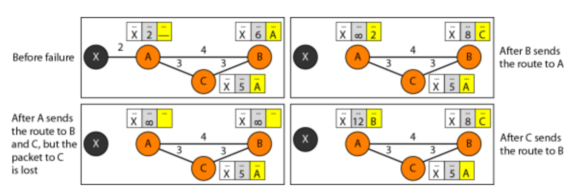
22.50

Figure 22.17 *Two-node instability*



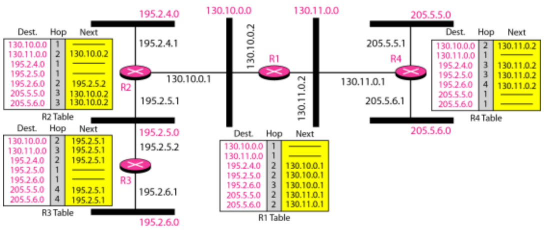
22.51

Figure 22.18 *Three-node instability*



22.52

Figure 22.19 ***Example of a domain using RIP***



R2 Table

Dest.	Hop	Next
130.10.0.0	1	———
130.11.0.0	2	130.10.0.2
195.2.4.0	1	———
195.2.5.0	1	———
195.2.6.0	2	195.2.5.2
205.5.5.0	3	130.10.0.2
205.5.6.0	3	130.10.0.2

R3 Table

Dest.	Hop	Next
130.10.0.0	2	195.2.5.1
130.11.0.0	3	195.2.5.1
195.2.4.0	2	195.2.5.1
195.2.5.0	1	———
195.2.6.0	1	———
205.5.5.0	4	195.2.5.1
205.5.6.0	4	195.2.5.1

R1 Table

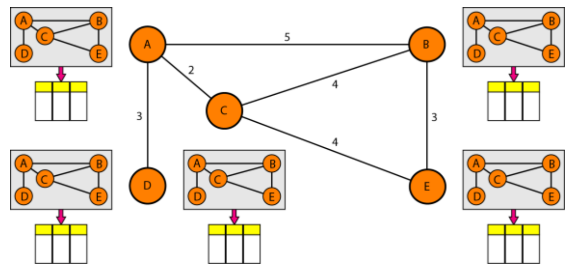
Dest.	Hop	Next
130.10.0.0	1	———
130.11.0.0	1	———
195.2.4.0	2	130.10.0.1
195.2.5.0	2	130.10.0.1
195.2.6.0	3	130.10.0.1
205.5.5.0	2	130.11.0.1
205.5.6.0	2	130.11.0.1

R4 Table

Dest.	Hop	Next
130.10.0.0	2	130.11.0.2
130.11.0.0	1	———
195.2.4.0	3	130.11.0.2
195.2.5.0	3	130.11.0.2
195.2.6.0	4	130.11.0.2
205.5.5.0	1	———
205.5.6.0	1	———

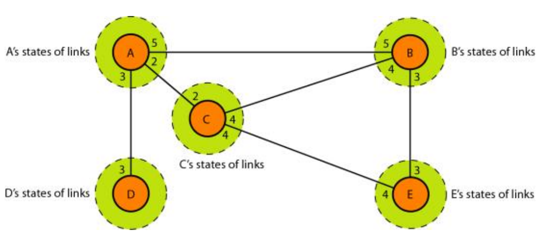
22.53

Figure 22.20 ***Concept of link state routing***



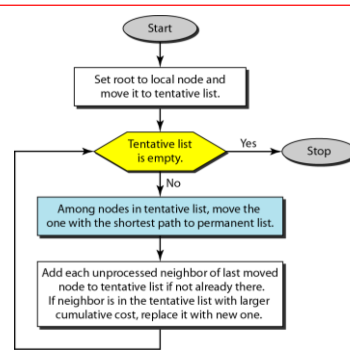
22.54

Figure 22.21 ***Link state knowledge***



22.55

Figure 22.22 ***Dijkstra algorithm***



22.56

Figure 22.23 *Example of formation of shortest path tree*

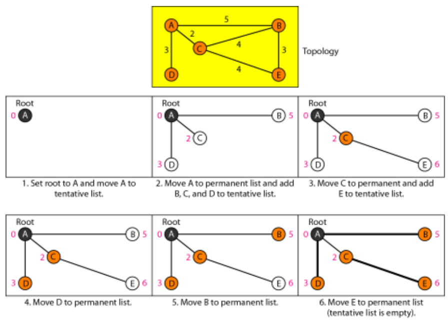


Table 22.2 *Routing table for node A*

Node	*Cost*	*Next Router*
A	0	—
B	5	—
C	2	—
D	3	—
E	6	C

Figure 22.24 *Areas in an autonomous system*

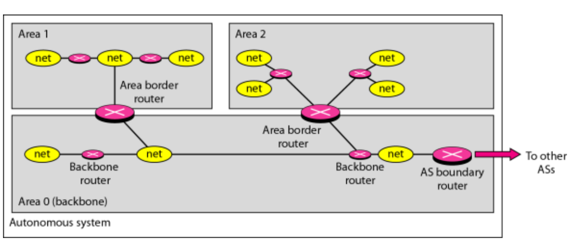


Figure 22.25 *Types of links*

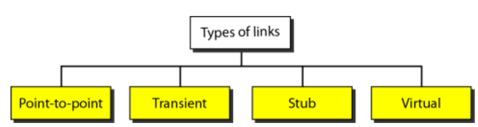


Figure 22.26 *Point-to-point link*

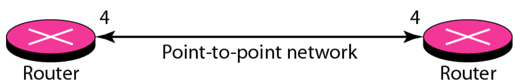


Figure 22.27 *Transient link*

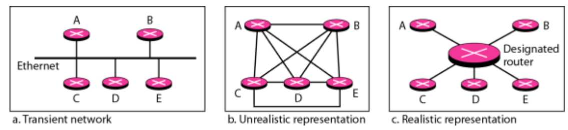


Figure 22.28 *Stub link*

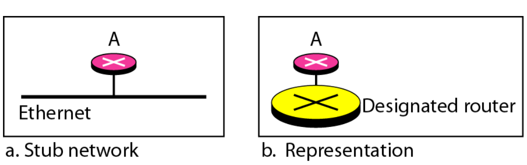


Figure 22.29 *Example of an AS and its graphical representation in OSPF*

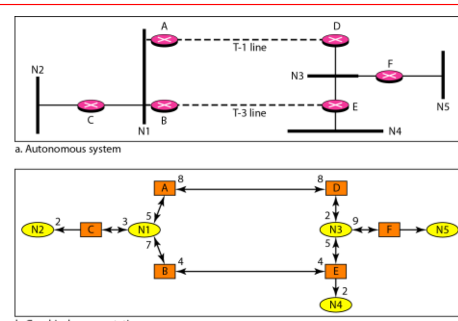
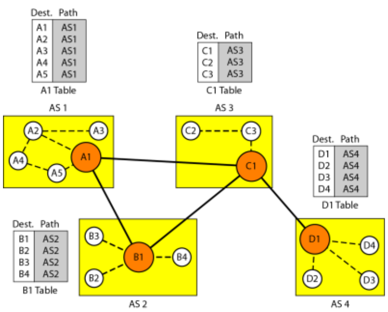


Figure 22.30 *Initial routing tables in path vector routing*



22.65

Figure 22.31 *Stabilized tables for three autonomous systems*

Dest.	Path
A1 ... A5	AS1 ... AS1
B1 ... B4	AS1-AS2 ... AS1-AS2
C1 ... C3	AS1-AS3 ... AS1-AS3
D1 ... D4	AS1-AS2-AS4 ... AS1-AS2-AS4

A1 Table

Dest.	Path
A1 ... A5	AS2-AS1 ... AS2-AS1
B1 ... B4	AS2 ... AS2
C1 ... C3	AS2-AS3 ... AS2-AS3
D1 ... D4	AS2-AS3-AS4 ... AS2-AS3-AS4

B1 Table

Dest.	Path
A1 ... A5	AS3-AS1 ... AS3-AS1
B1 ... B4	AS3-AS2 ... AS3-AS2
C1 ... C3	AS3 ... AS3
D1 ... D4	AS3-AS4 ... AS3-AS4

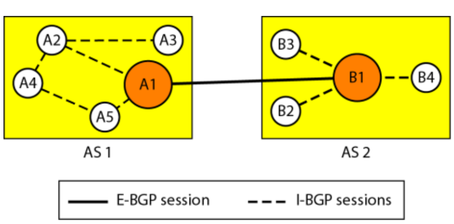
C1 Table

Dest.	Path
A1 ... A5	AS4-AS3-AS1 ... AS4-AS3-AS1
B1 ... B4	AS4-AS3-AS2 ... AS4-AS3-AS2
C1 ... C3	AS4-AS3 ... AS4-AS3
D1 ... D4	AS4 ... AS4

D1 Table

22.66

Figure 22.32 *Internal and external BGP sessions*



22.67

22-4 MULTICAST ROUTING PROTOCOLS

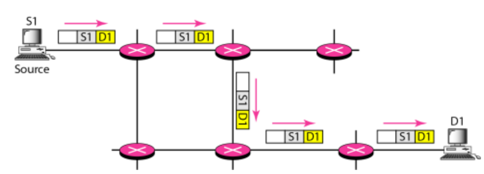
In this section, we discuss multicasting and multicast routing protocols.

Topics discussed in this section:

Unicast, Multicast, and Broadcast
Applications
Multicast Routing
Routing Protocols

22.68

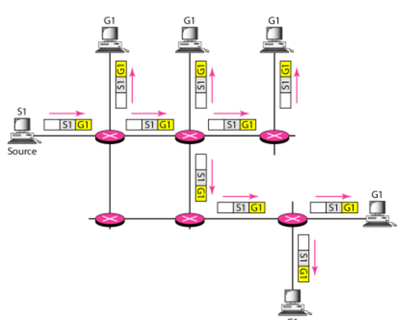
Figure 22.33 *Unicasting*



Note

In unicasting, the router forwards the received packet through only one of its interfaces.

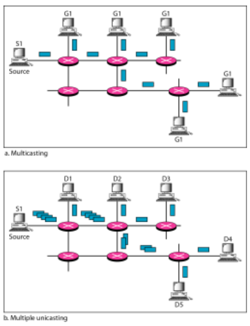
Figure 22.34 *Multicasting*



Note

In multicasting, the router may forward the received packet through several of its interfaces.

Figure 22.35 *Multicasting versus multiple unicasting*



22.73

Note

Emulation of multicasting through multiple unicasting is not efficient and may create long delays, particularly with a large group.

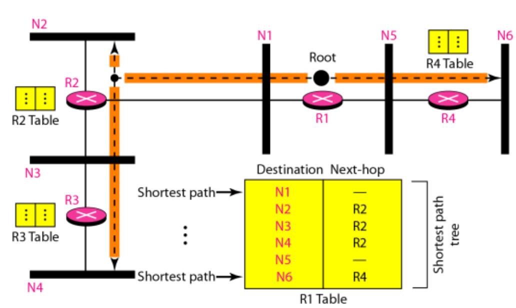
22.74

Note

In unicast routing, each router in the domain has a table that defines a shortest path tree to possible destinations.

22.75

Figure 22.36 *Shortest path tree in unicast routing*



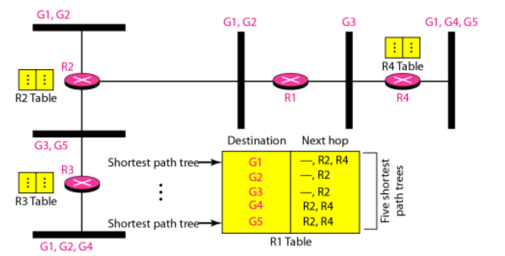
22.76

Note

In multicast routing, each involved router needs to construct a shortest path tree for each group.

22.77

Figure 22.37 ***Source-based tree approach***



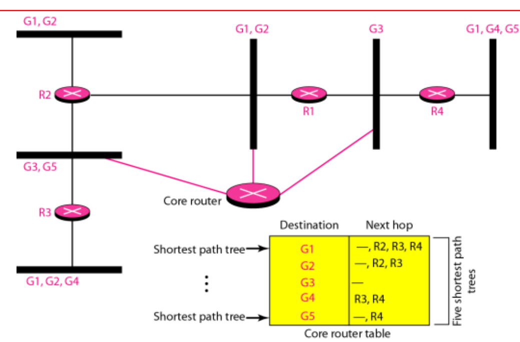
22.78

Note

In the source-based tree approach, each router needs to have one shortest path tree for each group.

22.79

Figure 22.38 ***Group-shared tree approach***

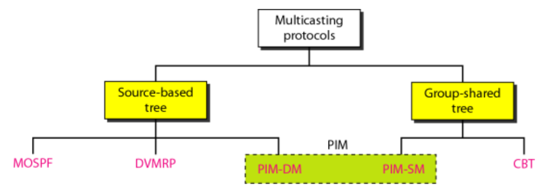


22.80

Note

In the group-shared tree approach, only the core router, which has a shortest path tree for each group, is involved in multicasting.

Figure 22.39 *Taxonomy of common multicast protocols*



Note

Multicast link state routing uses the source-based tree approach.

Note

Flooding broadcasts packets, but creates loops in the systems.

Note

RPF eliminates the loop in the flooding process.

Figure 22.40 *Reverse path forwarding (RPF)*

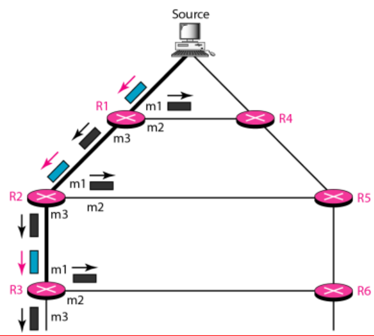


Figure 22.41 *Problem with RPF*

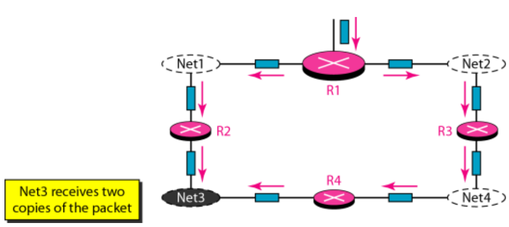
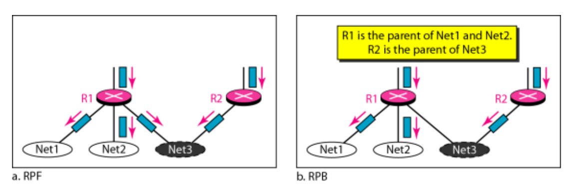


Figure 22.42 *RPF Versus RPB*

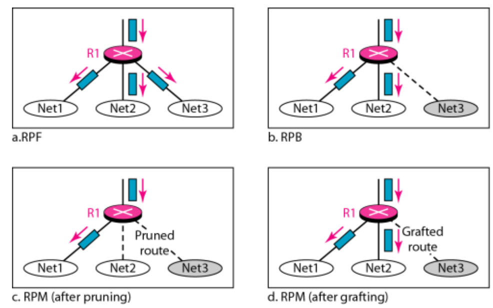


Note

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.

22.89

Figure 22.43 *RPF, RPB, and RPM*



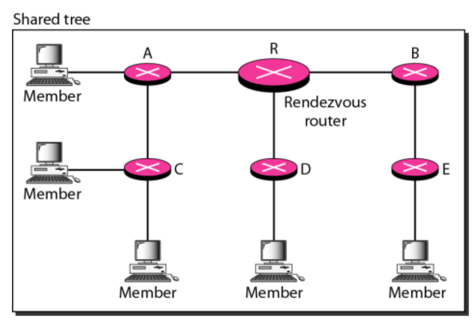
22.90

Note

RPM adds pruning and grafting to RPB to create a multicast shortest path tree that supports dynamic membership changes.

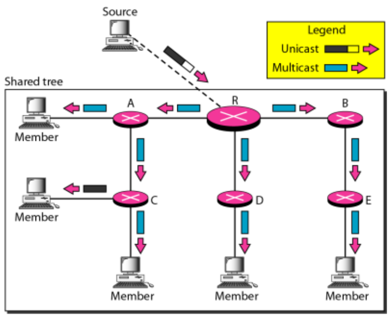
22.91

Figure 22.44 *Group-shared tree with rendezvous router*



22.92

Figure 22.45 *Sending a multicast packet to the rendezvous router*



22.93

Note

In CBT, the source sends the multicast packet (encapsulated in a unicast packet) to the core router. The core router decapsulates the packet and forwards it to all interested interfaces.

22.94

Note

PIM-DM is used in a dense multicast environment, such as a LAN.

22.95

Note

PIM-DM uses RPF and pruning and grafting strategies to handle multicasting.
However, it is independent of the underlying unicast protocol.

22.96

Note

PIM-SM is used in a sparse multicast environment such as a WAN.

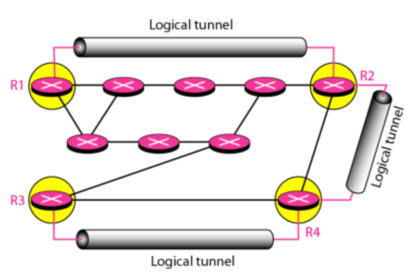
22.97

Note

PIM-SM is similar to CBT but uses a simpler procedure.

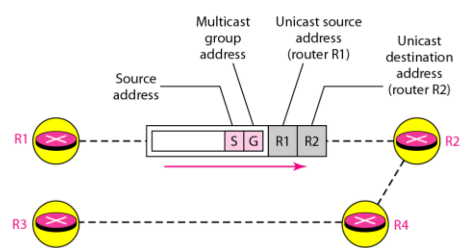
22.98

Figure 22.46 ***Logical tunneling***



22.99

Figure 22.47 ***MBONE***



22.100