Chapter 22

Network Layer: Delivery, Forwarding, and Routing

Modified by Chao Gao, 2010

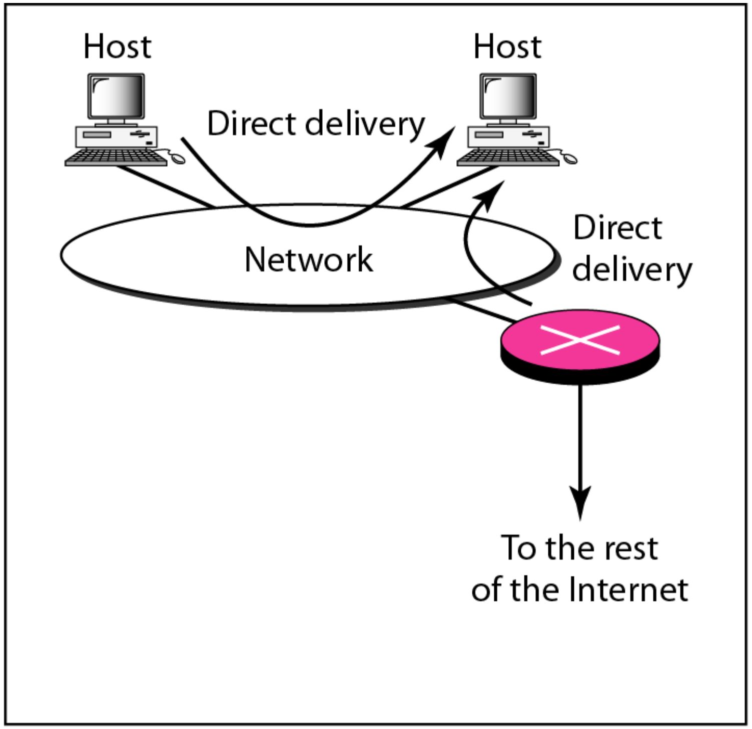
22-1 DELIVERY

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:

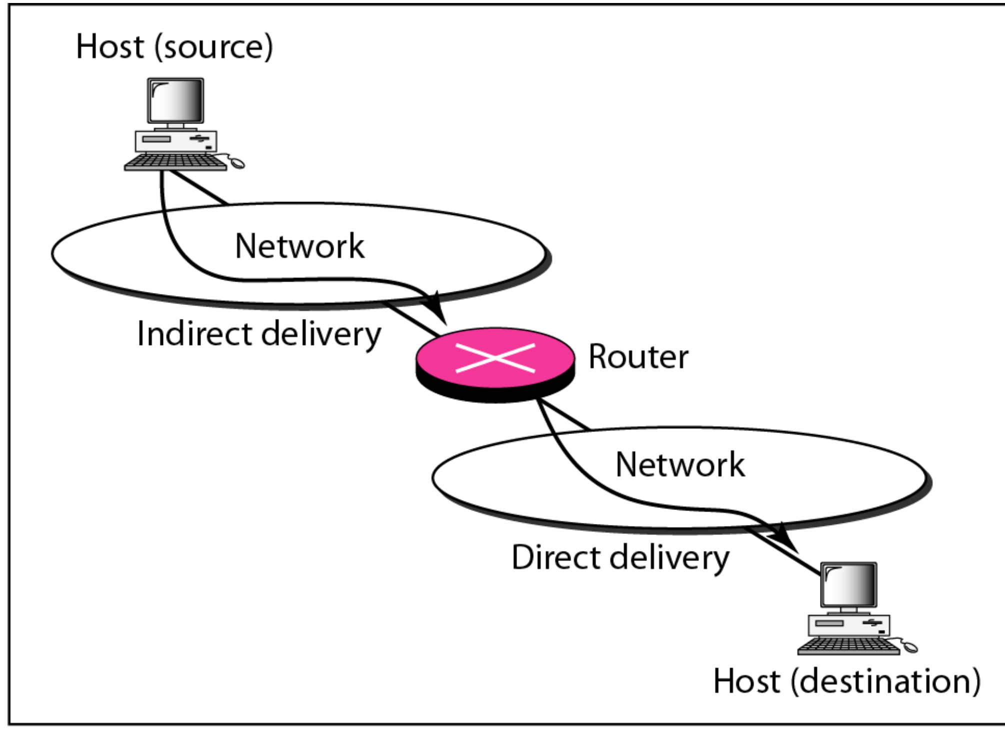
Direct vs Indirect Delivery

Figure 22.1 *Direct and indirect delivery*



a. Direct delivery

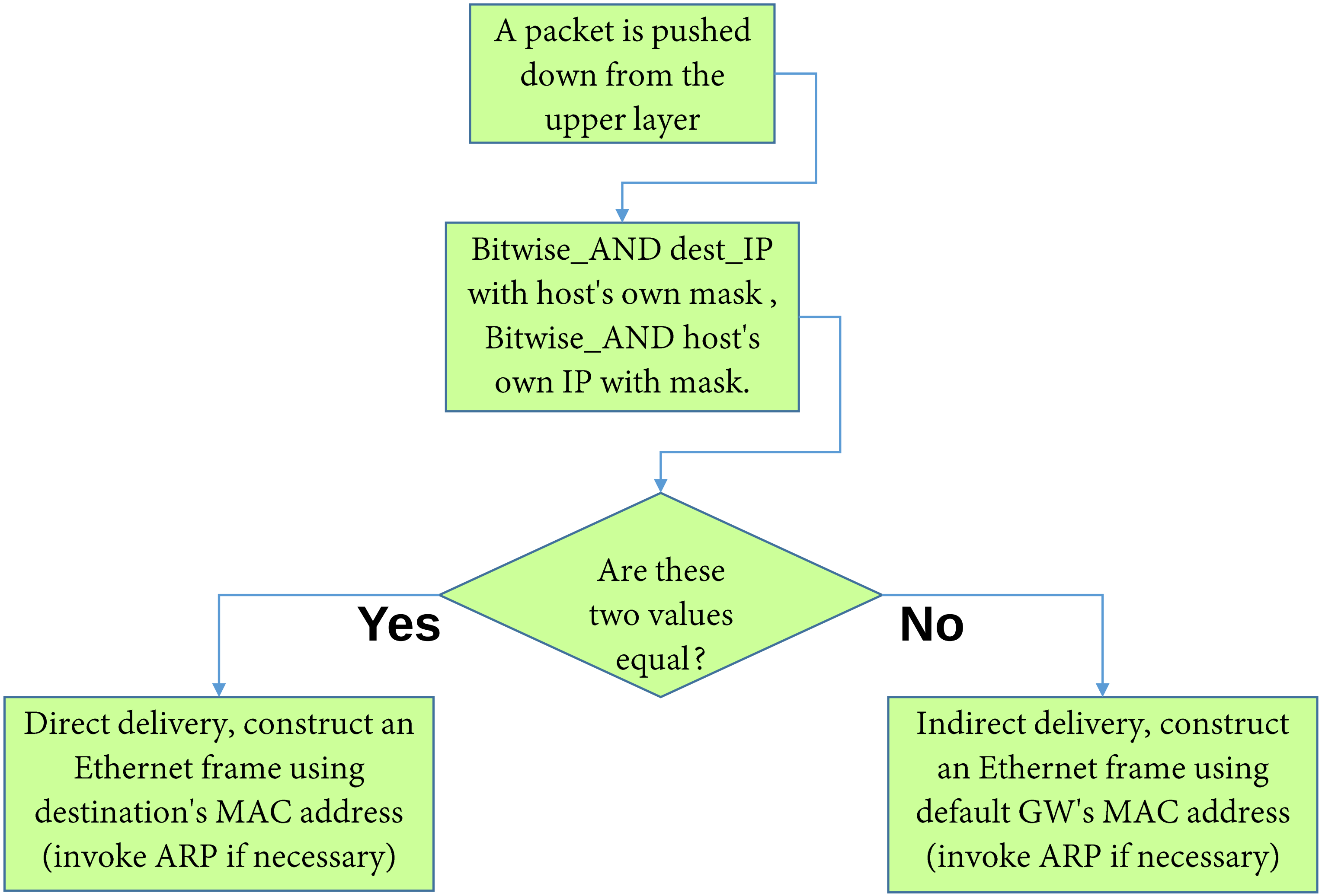
In the same subnetwork.



b. Indirect and direct delivery

Not in the same subnetwork.
Router(s) involved.

A Host Decides Delivery Method by



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

Figure 22.2 *Route based vs next-hop based*

a. Routing tables based on route

b. Routing tables based on next hop

Routing table for host A

Destination	Route
Host B	R1, R2, host B

Destination	Next hop
Host B	R1

Routing table for R1

Destination	Route
Host B	R2, host B

Destination	Next hop
Host B	R2

Routing table for R2

Destination	Route
Host B	Host B

Destination	Next hop
Host B	---

Which is better?

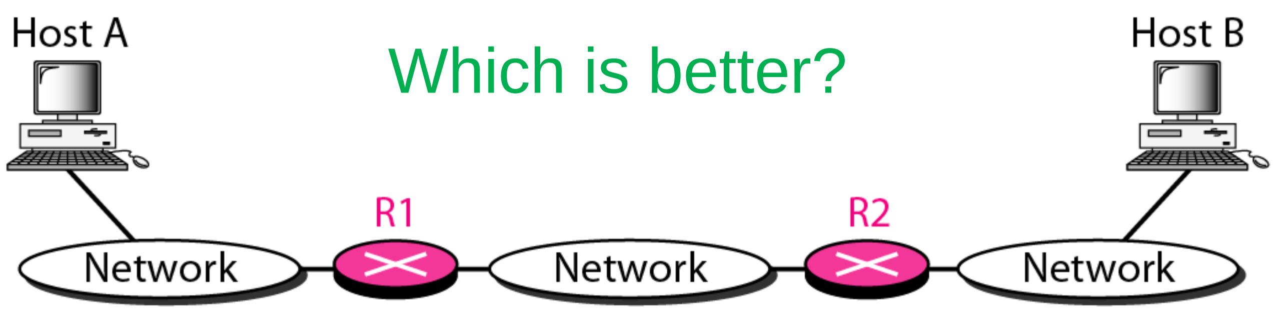


Figure 22.3 *Host-specific vs network-specific method*

Which is better?

Routing table for host S based on host-specific method

Destination	Next hop
A	R1
B	R1
C	R1
D	R1

Routing table for host S based on network-specific method

Destination	Next hop
N2	R1

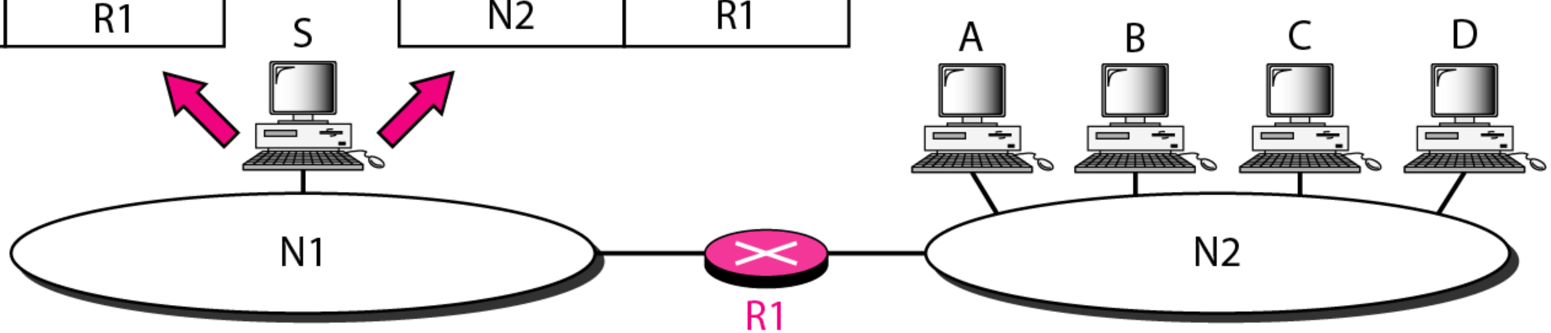
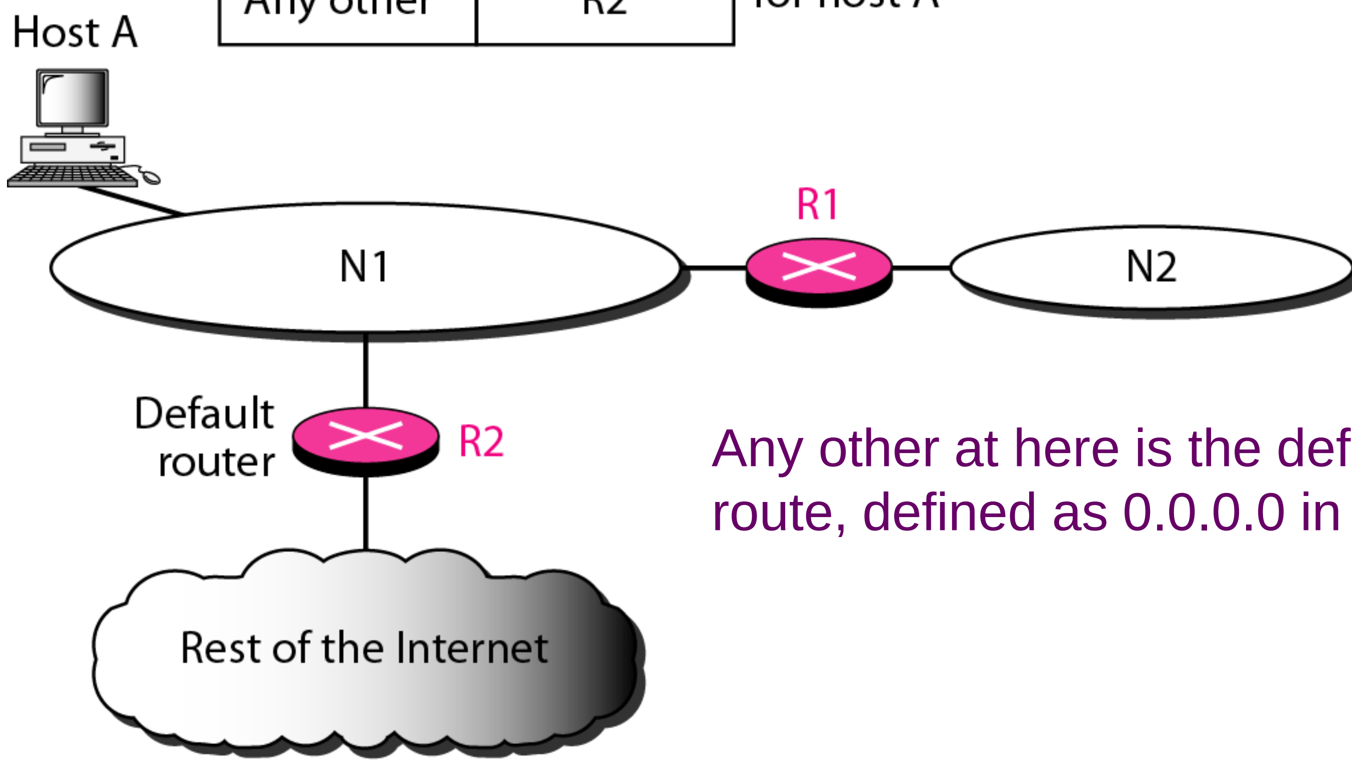


Figure 22.4 *Default route*

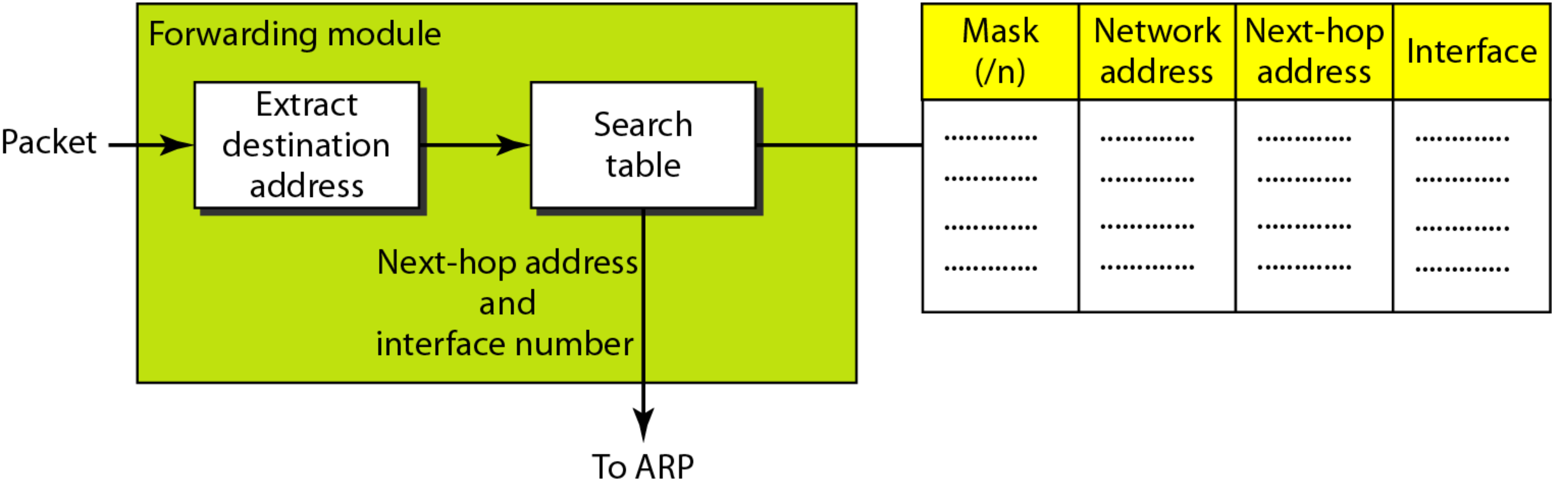
Destination	Next hop
N2	R1
Any other	R2

Routing table for host A



Any other at here is the default route, defined as 0.0.0.0 in IPv4.

Figure 22.5 *Simplified forwarding table in classless address*



Note

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

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.

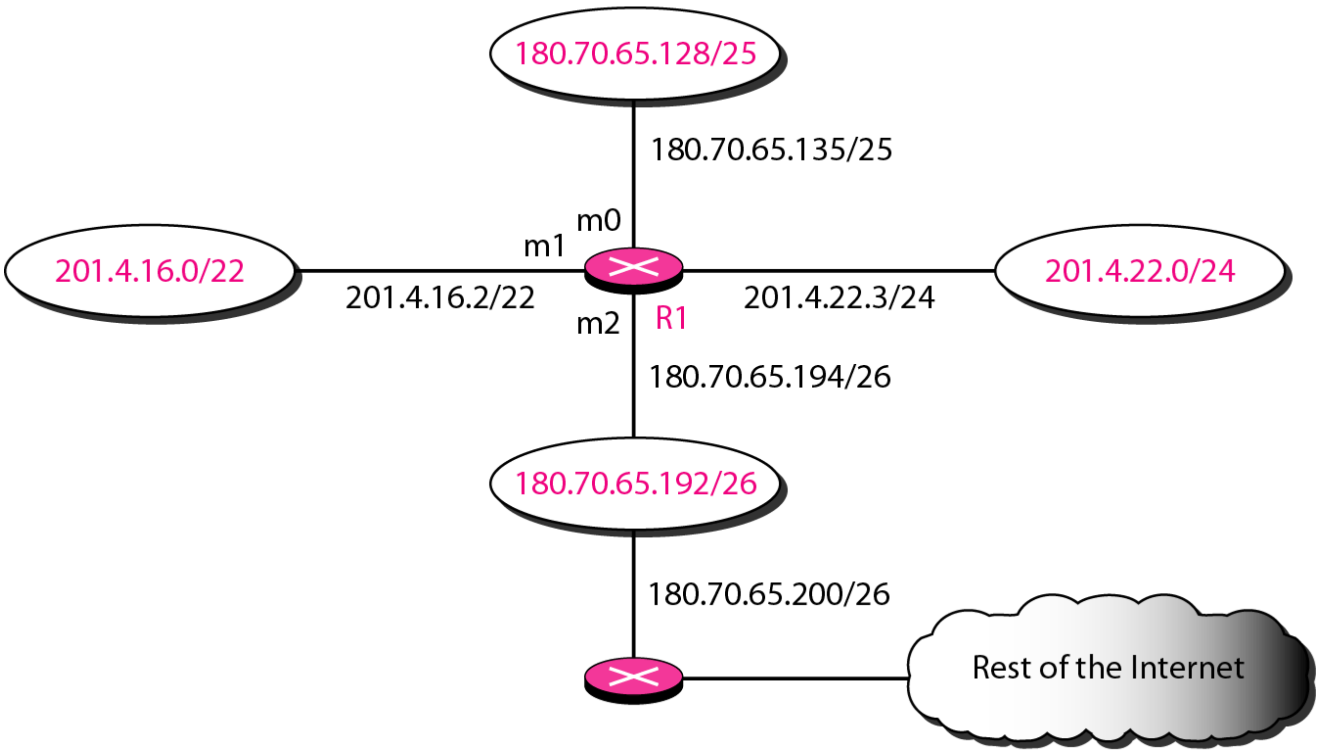


Table 22.1 ***Routing table for router R1 in Figure 22.6***

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

A routing table is always sorted by mask value *n*, so that correct decision will be made.

What is the mask value for "Any"?

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.***

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).***
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.***

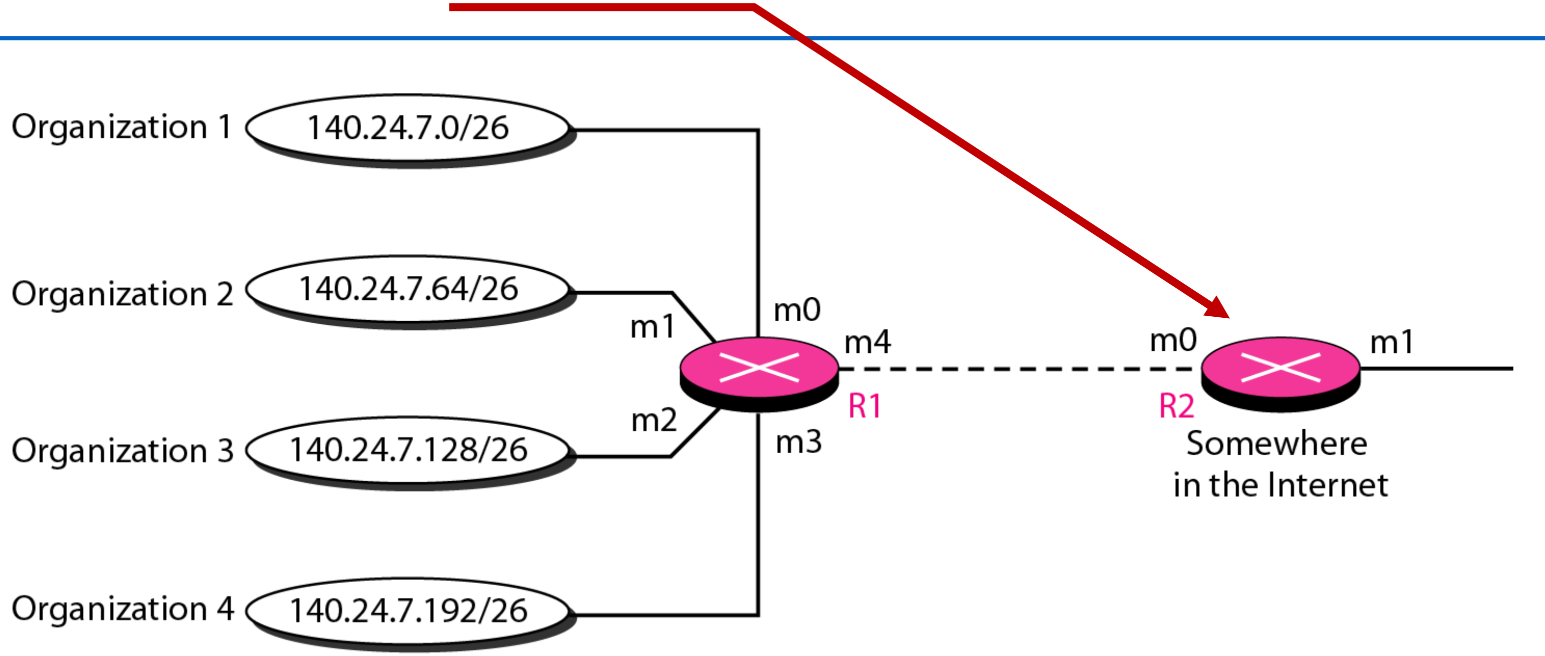
Example 22.4

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 (to simplify the routing table)*



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

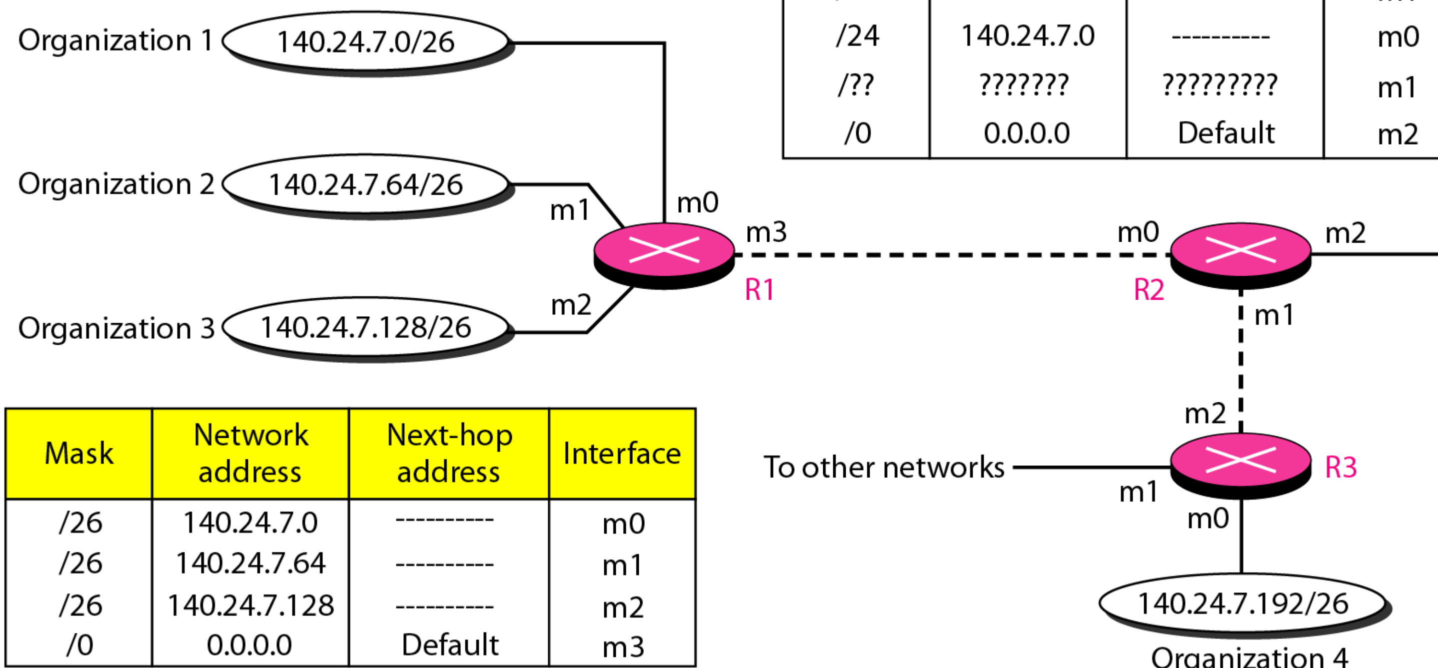
Mask	Network address	Next-hop address	Interface
/24	140.24.7.0	----------	m0
/0	0.0.0.0	Default	m1

Routing table for R2

Figure 22.8 *Longest mask matching (higher priority to match)*

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

This example explains why the routing table is sorted by mask value.

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

Hierachical routing

- In order to reduce the routing table size, a local ISP can be assigned a single, but large block of address with a certain prefix length.
- So that the rest of Internet needs only to know the ISP's netID, and thus the routing table size can be greatly reduced.
- This is the case of Figure 22.7.

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.

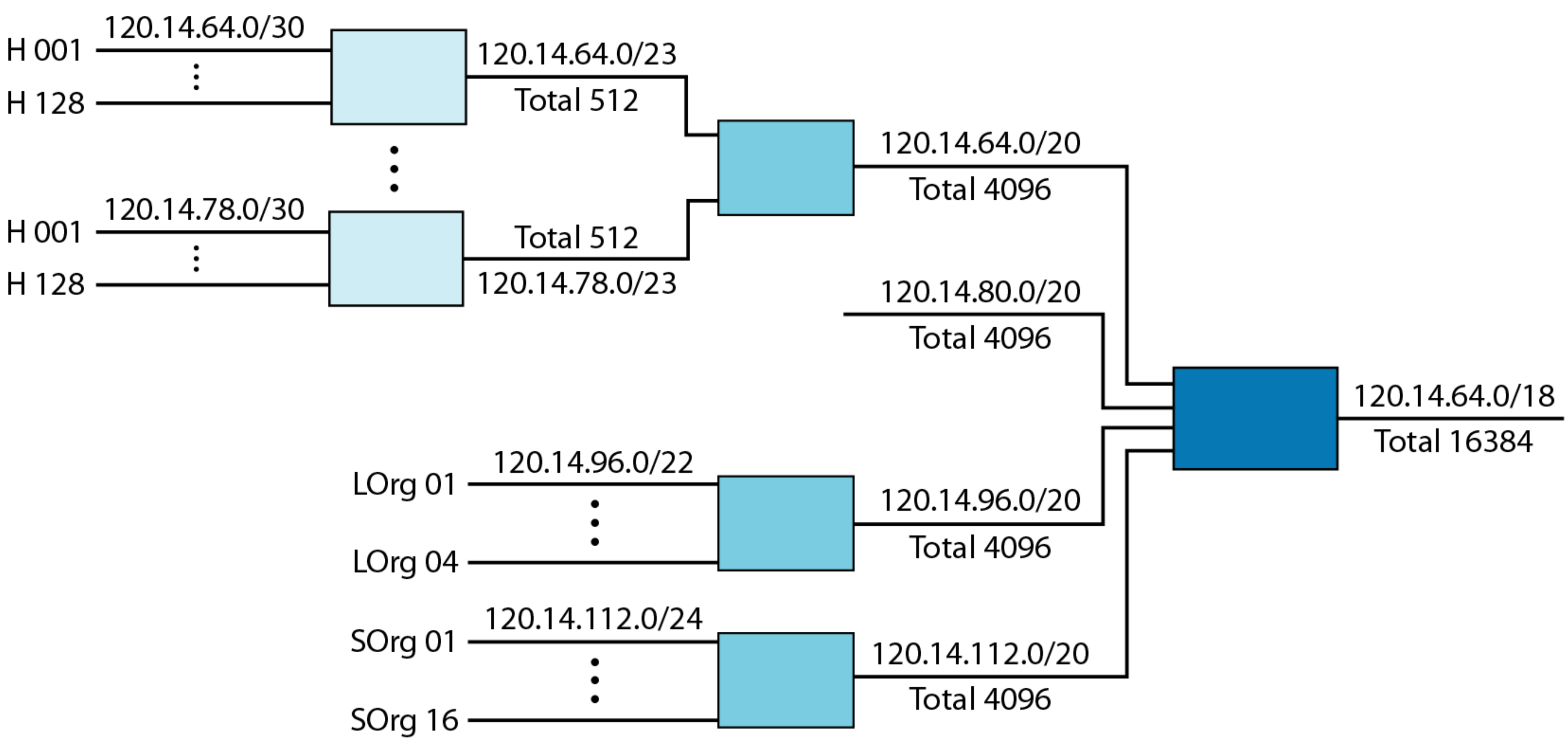
Example 22.5 (continued)

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.

Figure 22.9 *Hierarchical routing with ISPs*



Routing table

- The minimum size of routing table is 4 columns. It is not standardized that how many extra columns can be added.

Figure 22.10 ***Common fields in a routing table***

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

Flag: 5 flags are defined as U(Up), G(Gateway), H(host-specific), D(added-by-redirection), M(Modified-by-redirection)
Reference count: the number of users of this route at the moment.
Use: number of packets transmitted through for the destination.

Google: Linux routing table

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
. . .
```

22-3 UNICAST ROUTING PROTOCOLS

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

Topics discussed in this section:

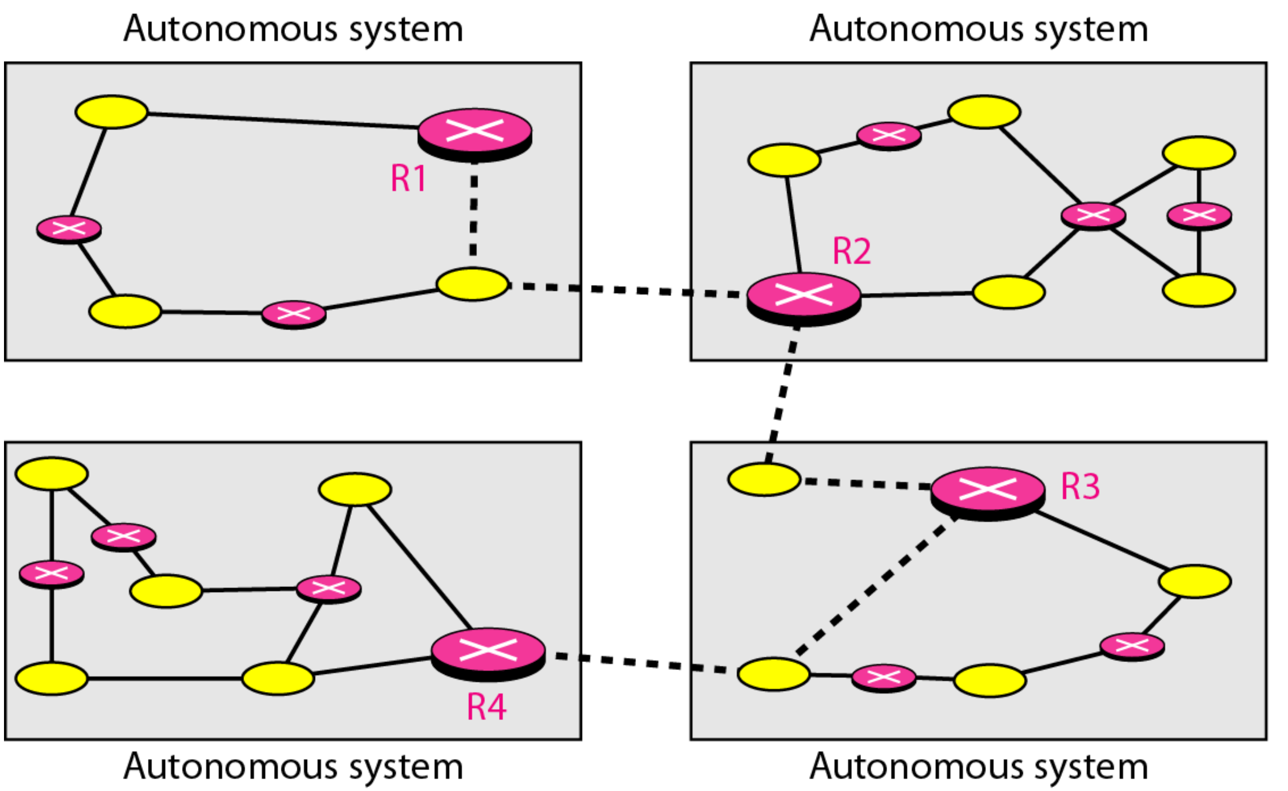
Optimization
Distance Vector Routing and RIP
Link State Routing and OSPF

How to optimize?

- Optimization means to route the package from the source to destiniation via the **most efficient path**.
- usually a "metric" is used to value the cost of a link. The higher the metric value, the higher the cost.

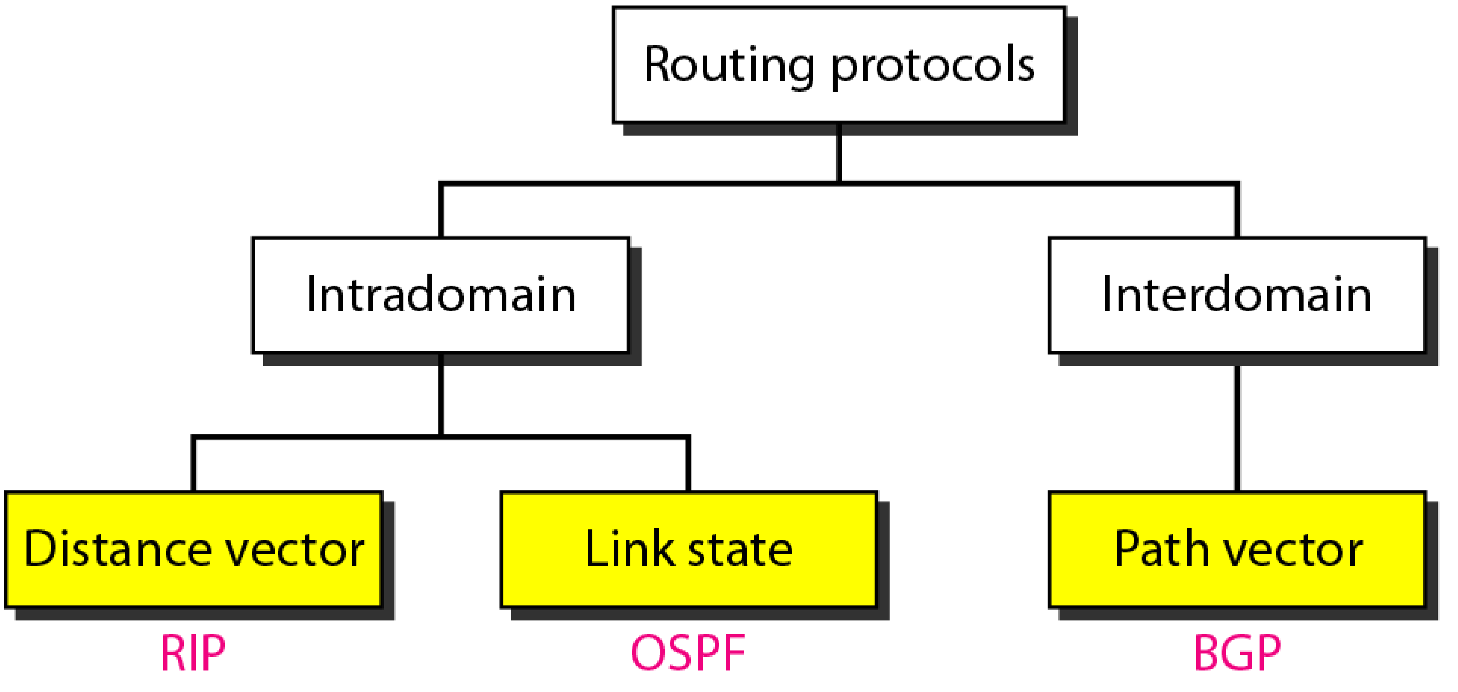
- RIP uses hop-count as metric.
- OSPF uses link bandwidth as metric.

Figure 22.12 *Autonomous systems*



Both RIP and OSPF are used inside an AS.

Figure 22.13 *Popular routing protocols*



Note

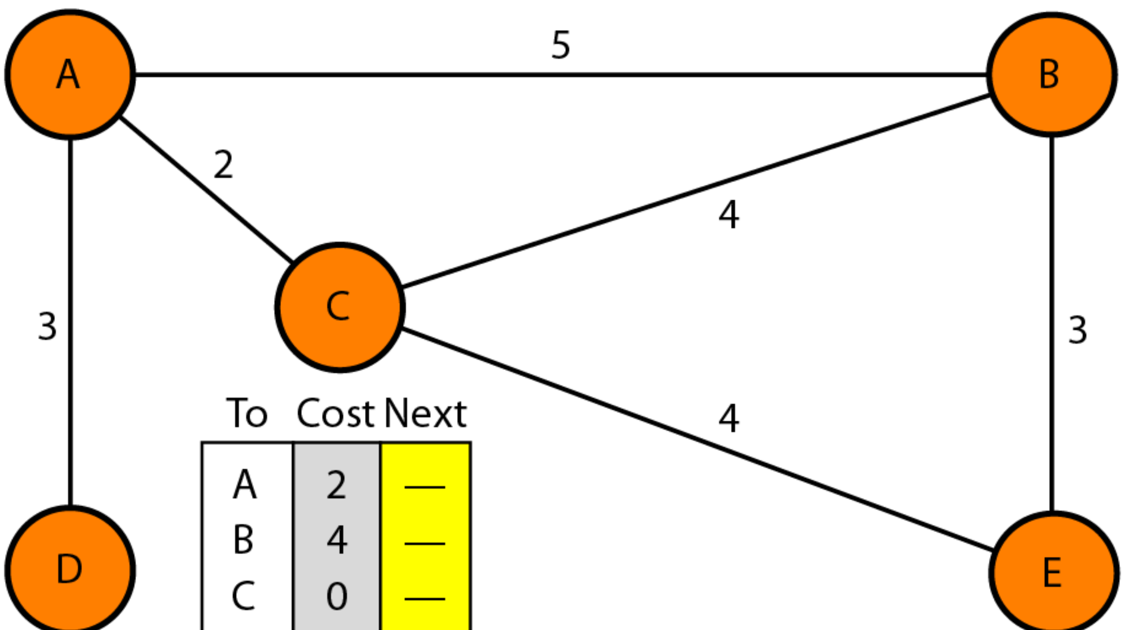
In distance vector routing, each node shares its routing table with its immediate neighbors periodically.

This algorithm is also called as Bellman-Ford algorithm.

Bellman-Ford Algorithm

1. Each router starts with an initial table, which only knows its direct-connected networks.
2. Each router periodically broadcasts its routing table to all its neighbor routers.
3. Upon receiving a routing table, the receiver:
   a) Increases hop count by one of each entry
   b) Compares the incoming table with its own table
   c) If the incoming entry is new, add it to own table
   d) If the incoming entry exists, but hop count is smaller, replace the existing entry by marking the next hop to be the sender.
   e) Otherwise do nothing.

Figure 22.15 *Initialization of tables in distance vector routing*



To	Cost	Next
A	0	—
B	5	—
C	2	—
D	3	—
E	∞	

A's table

To	Cost	Next
A	5	—
B	0	—
C	4	—
D	∞	
E	3	—

B's table

To	Cost	Next
A	3	—
B	∞	
C	∞	
D	0	—
E	∞	

D's table

To	Cost	Next
A	2	—
B	4	—
C	0	—
D	∞	
E	4	—

C's table

To	Cost	Next
A	∞	
B	3	B
C	4	C
D	∞	
E	0	D

E's table

Figure 22.16 *Updating in distance vector routing*

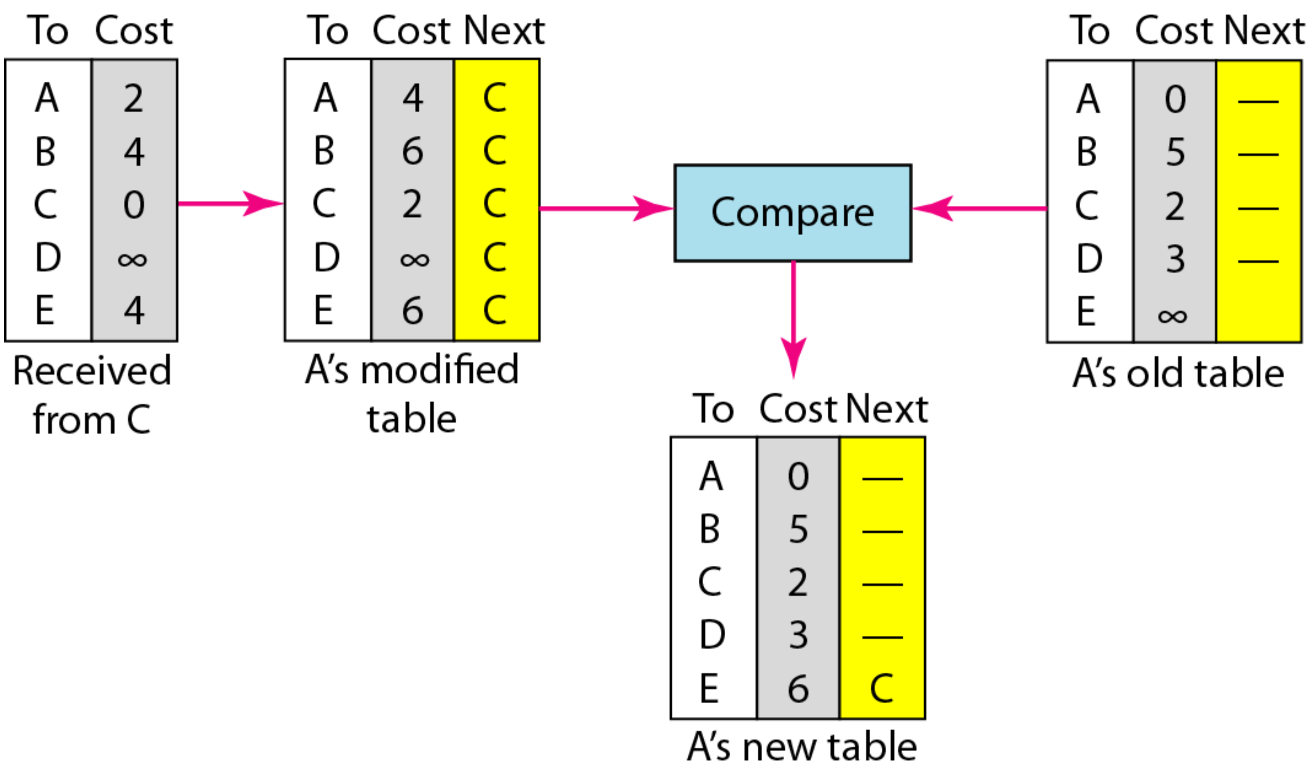


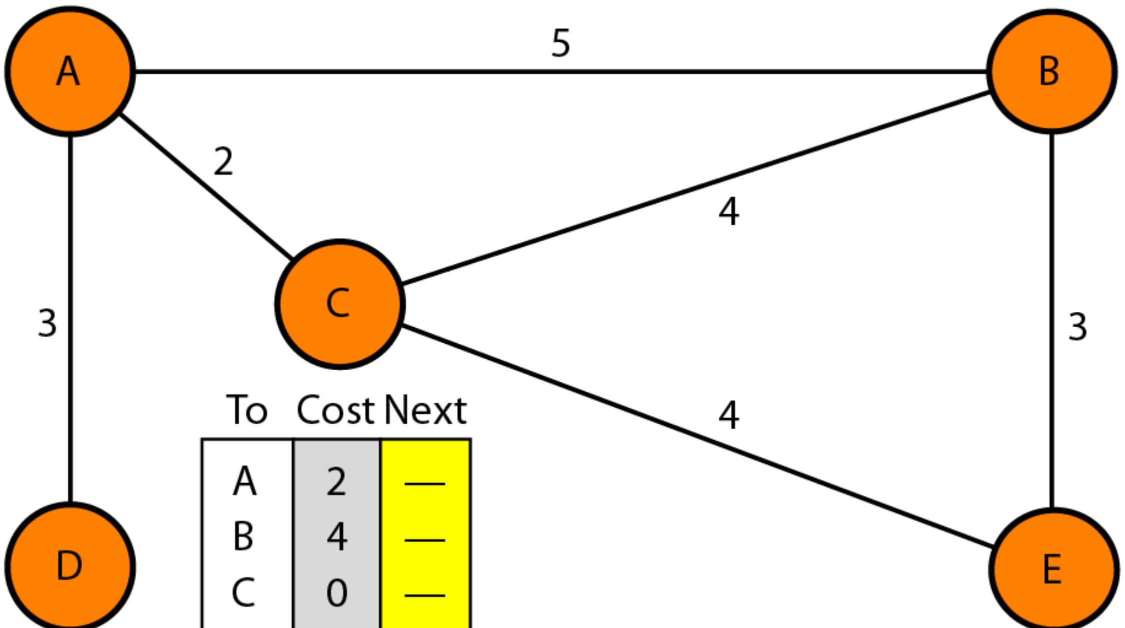
Figure 22.14 *Distance vector routing tables*

To	Cost	Next
A	0	—
B	5	—
C	2	—
D	3	—
E	6	C

A's table

To	Cost	Next
A	5	—
B	0	—
C	4	—
D	8	A
E	3	—

B's table



To	Cost	Next
A	3	—
B	8	A
C	5	A
D	0	—
E	9	A

D's table

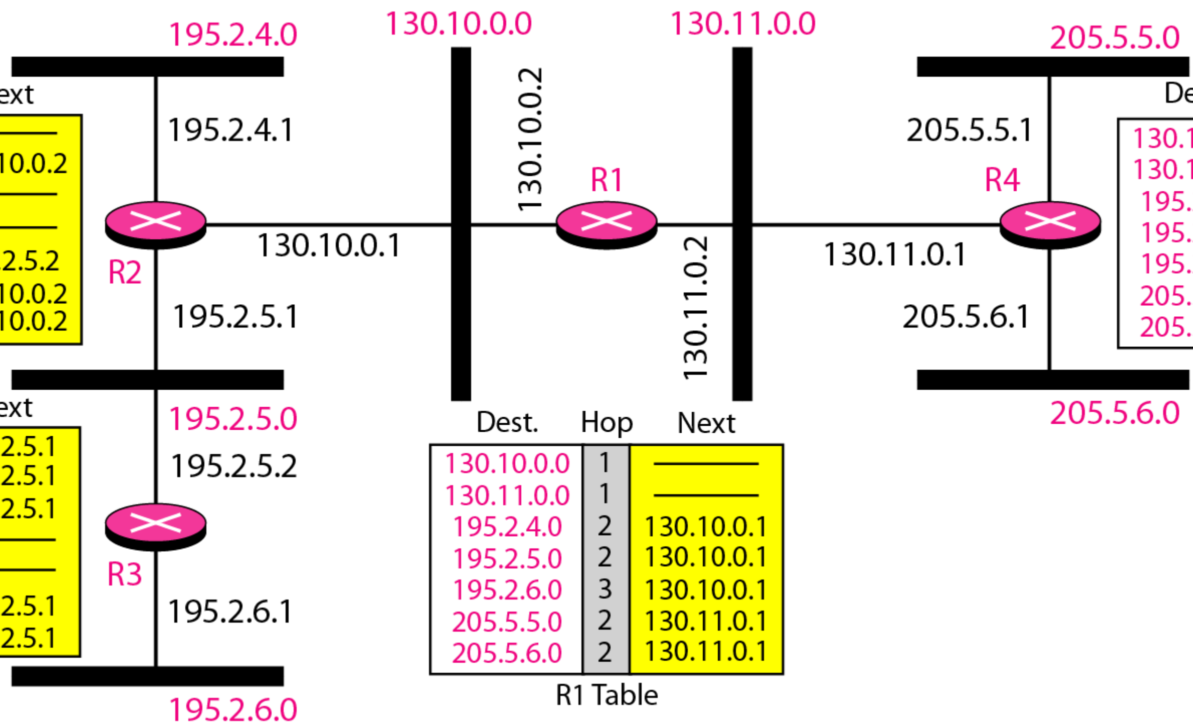
To	Cost	Next
A	2	—
B	4	—
C	0	—
D	5	A
E	4	—

C's table

To	Cost	Next
A	6	C
B	3	—
C	4	—
D	9	C
E	0	—

E's table

Figure 22.19 *Example of a domain using RIP (Routing Information Protocol)*



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

R2 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

R3 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

R1 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	—

R4 Table

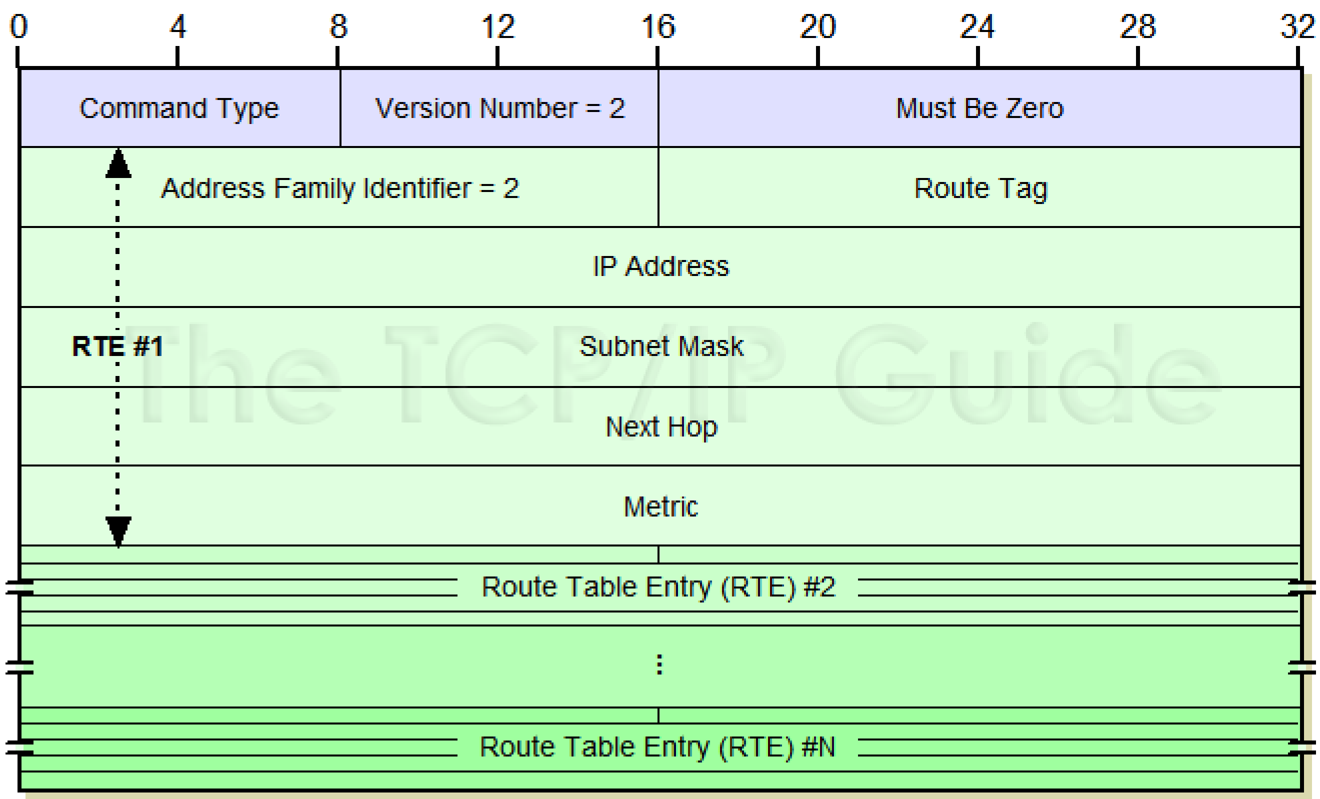
RIP in Practice

- Two versions of RIP, RIPv1 and RIPv2.
- Both use hop count as distance vector. But,
  - RIPv1 is classful. RIP2 is classless (supports Variable Length Subnet Mask VLSM).
  - RIPv1 broadcasts (dst:255.255.255.255) routing table, RIPv2 multicasts (224.0.0.9).

- Because table is broadcast/multicast periodically, the slow neighbor negotiation mechanism limits the use of RIP in a large network (with diameter > 15 hops).

RIPv2 Message Format

- More reading: http://ericleahy.com/index.php/implement-ipv4-rip-version-2-ripv2/.



http://www.tcpipguide.com/free/t_RIPVersion2RIP2MessageFormatandFeatures-3.htm

Link-State Routing (OSPF)

- When powered on, a router has initialized routing table of itself (i.e., the links it's directly using).
- All router will generate a "link-state advertisement" (LSA) according to their initial routing table. The LSAs are flooded throughout the whole network (AS).
- Once there is an update in the routing table (by receiving a LSA), a router uses **Dijkstra algorithm** to find shortest path (to all the known routers).

OSPF Link Cost

- Each link cost is calculated in this way (cited from Cisco):
  - Cost = 10^8/bandwidth in bps
  - Example: a 10base-T Ethernet link has cost = $\frac{10^8}{10^7}$ = 10.
  - In Cisco routers, use "ip ospf cost <value>" can manually set the cost at a given interface.

A fast Ethernet link has cost of 1.

Figure 22.20 *Concept of link state routing*

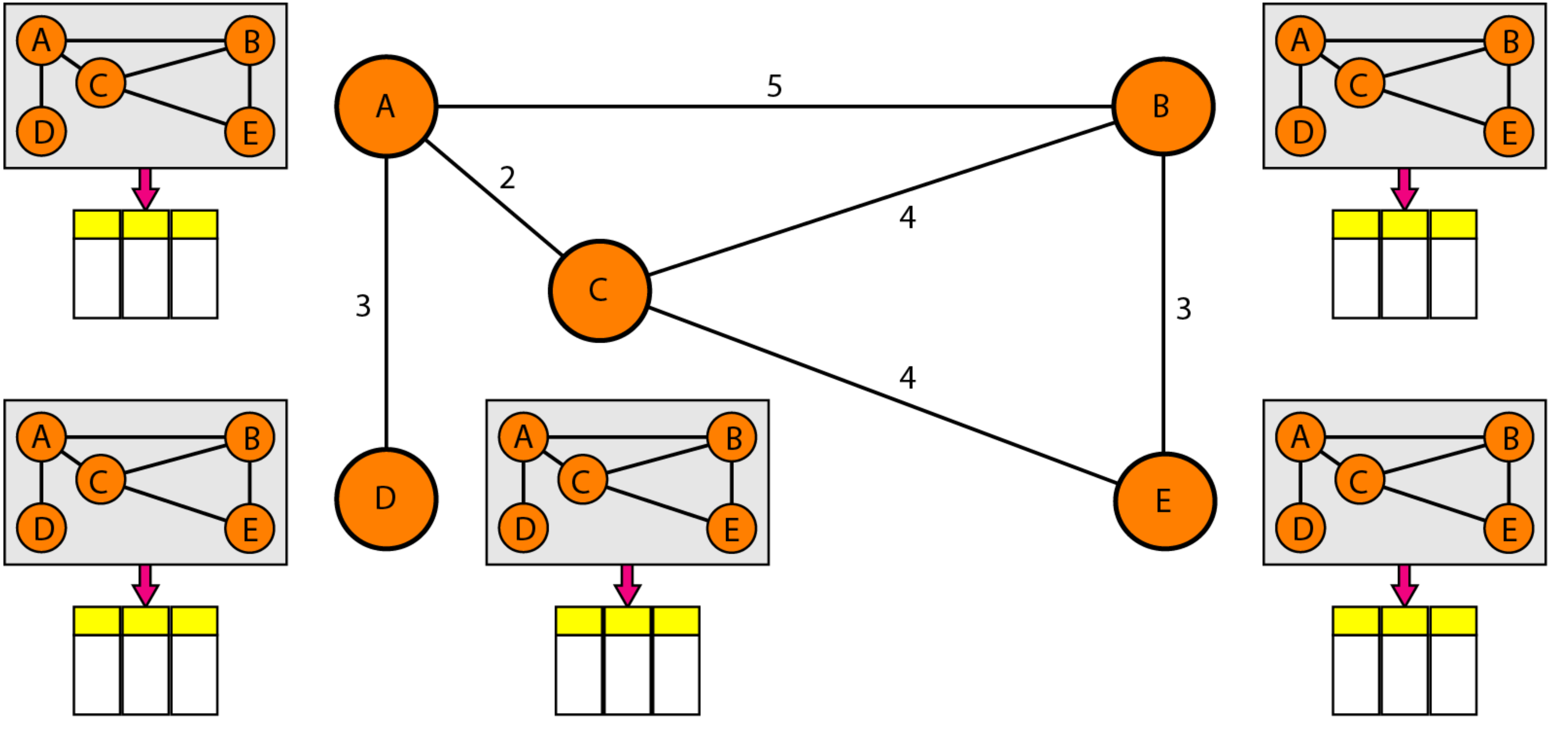


Figure 22.21 *Link state knowledge (initial state)*

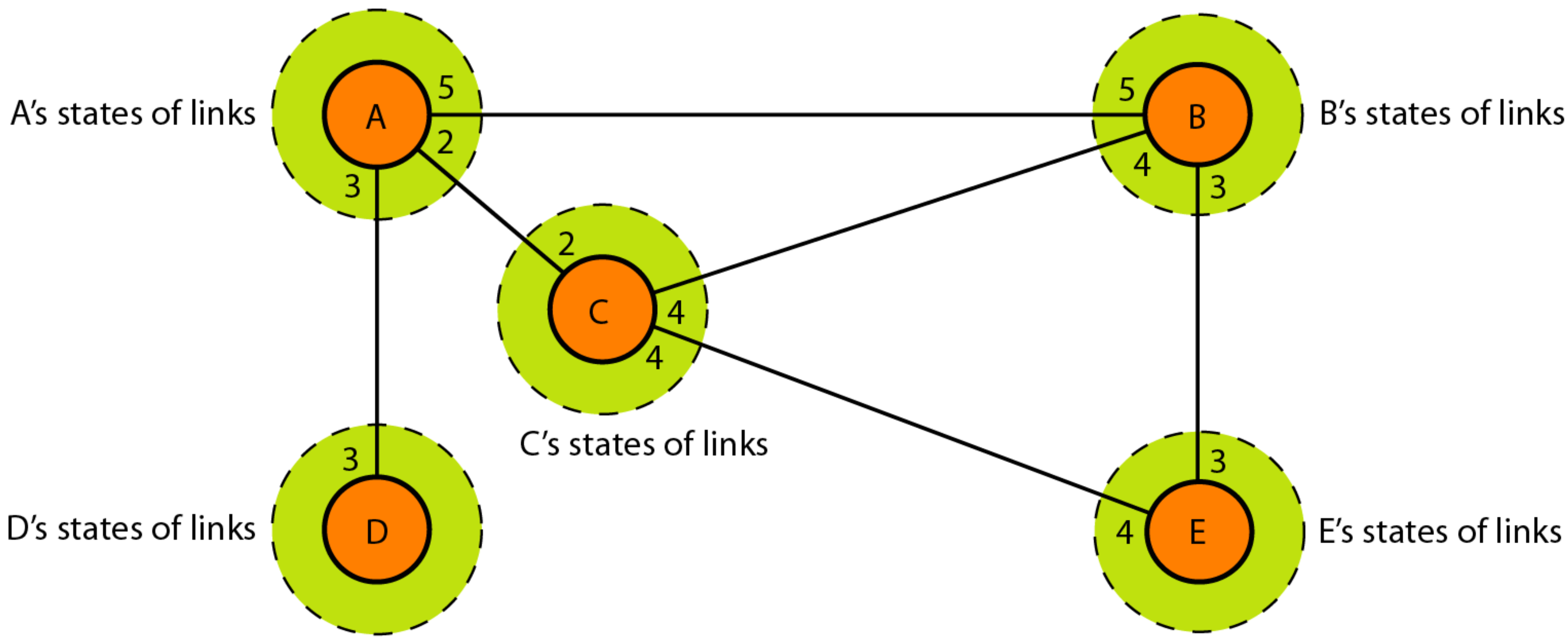


Figure 22.22 *Dijkstra algorithm*

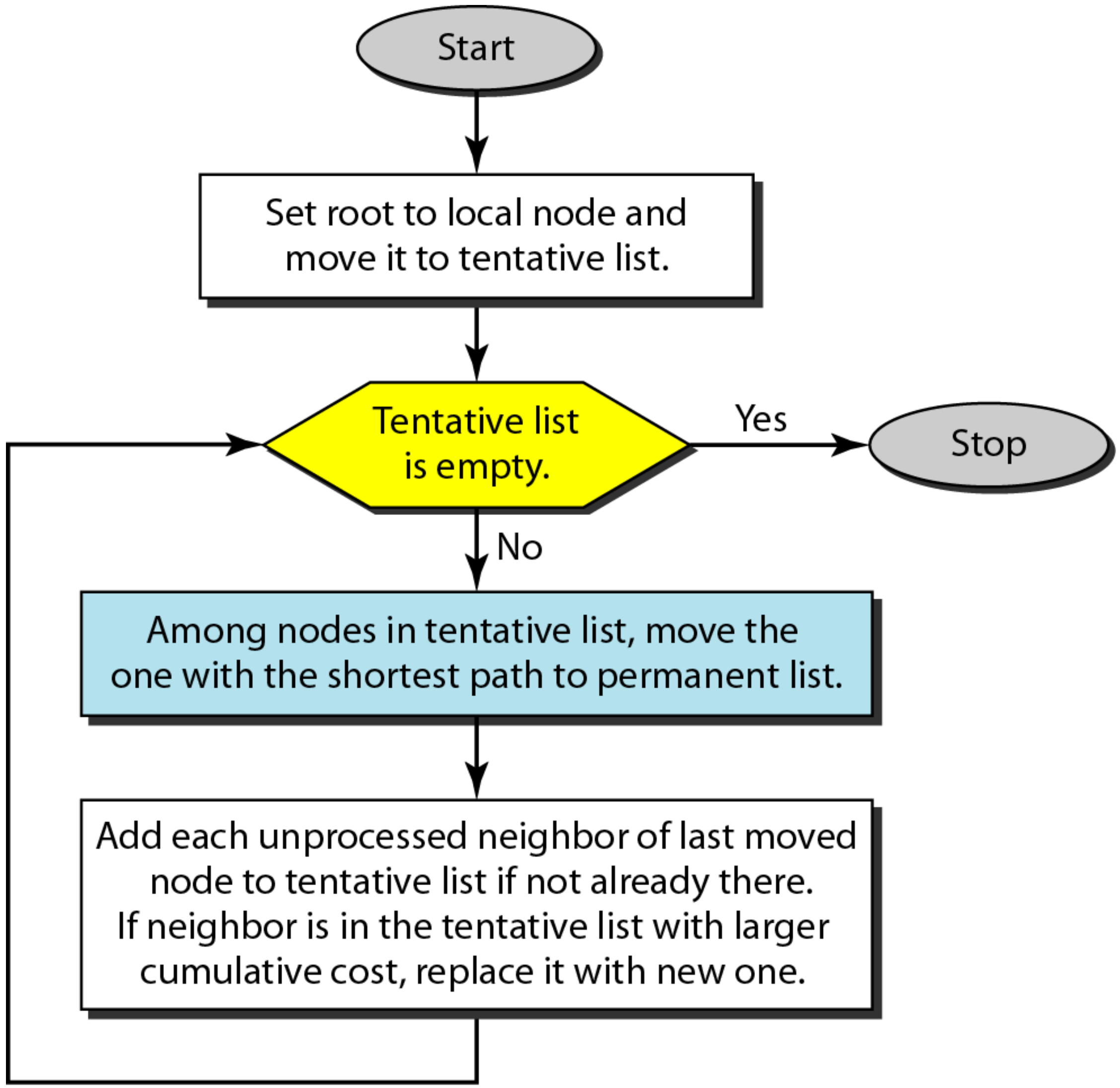


Figure 22.23 *Example of formation of shortest path tree*

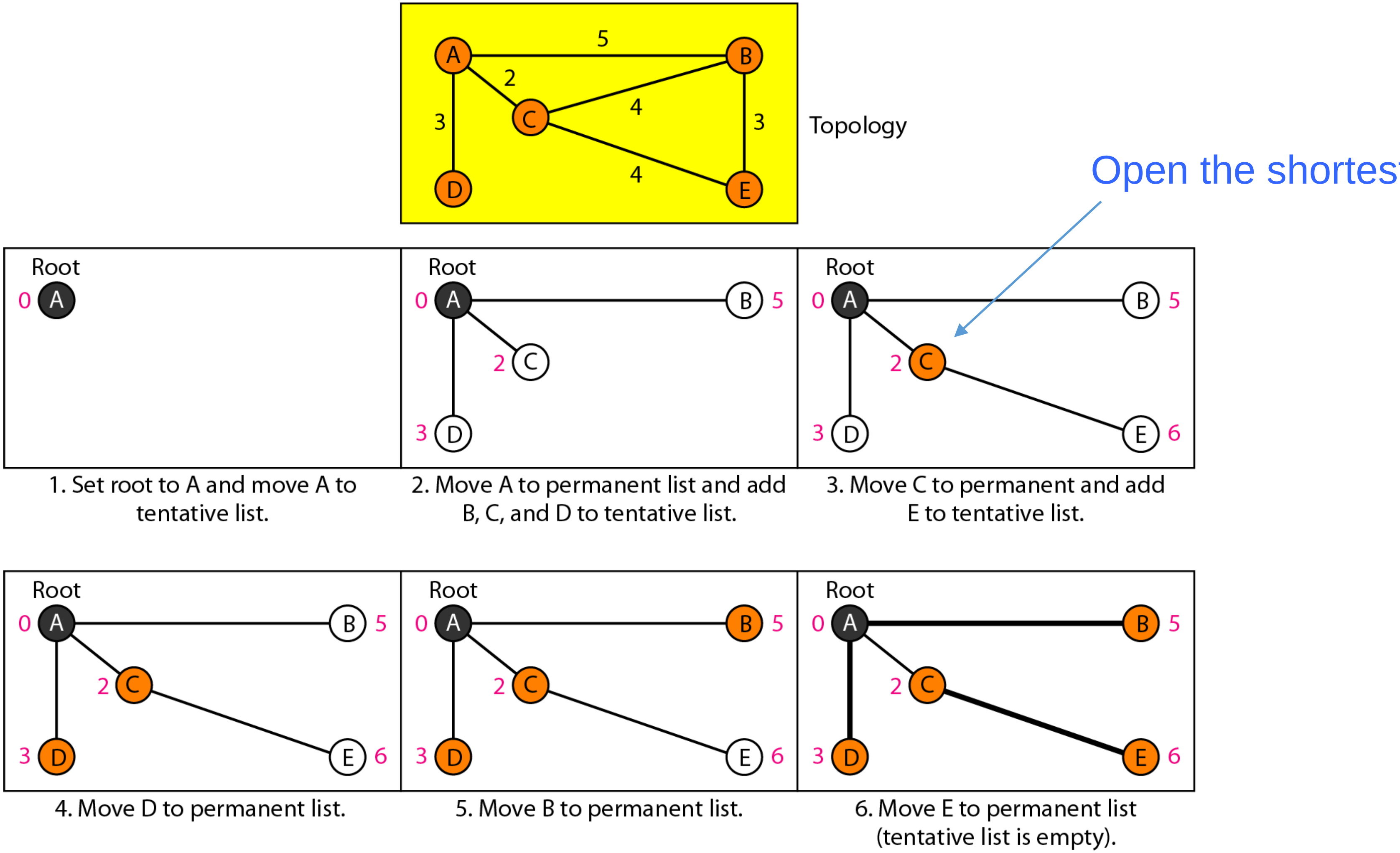


Table 22.2 ***Routing table for node A***

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

Case Study: Cisco Router Configuration

Cisco routers

- We will focus on Cisco's **branch** routers.
  - Cisco 3900, 3800, 3200, 2900, 2800, 1900, 1800, 800 series
  - There are other types of routers, such as mobile internet routers, service provider edge routers, etc.

Connect to a router

- When first use, the router has to be connected via Serial Port.
  - Use Windows **Hyperterminal** program.
  - Set baudrate 9600, 8-bit, no parity, 1 stopbit.
  - Once the connection is established, hit **Enter** key a prompt will show on the terminal window as:

```
cisco>_
```

Enter into privilege mode

- When first connect, it is in EXEC mode. The user can use such **unprivileged** commands as show, ping, telnet, and rlogin, etc, but cannot change configuration.
- To enter privilege mode, use command “enable”. You need a password here.

```
cisco>enable
cisco#_
```

Sub-modes

- From the privilege mode you can enter into a number of sub-modes. Once entered in a sub-mode, you will see the mode name in the prompt:

- Type command "exit" or hit "ctrl_z" will return to a parent mode.

```
cisco(sub-mode-name)#_
```

Display configuration

- Use “show” command with different parameters:

```
cisco#show interfaces
cisco#show ip protocols
cisco#show ip route
cisco#show ip arp
```

Configure ports

- A router has a number of physical ports, each is attached to specific subnet. So each port must have proper IP address.

```
cisco#config
cisco (config)#interface serial 1/1
cisco (config-if)#ip address 192.168.0.3 255.255.255.0
ciscoconfig-if)#no shutdown
cisco (config-if)#ctrl-Z
cisco#
```

Unconfigure

- In the Cisco IOS, the way to reverse or delete the results of any command is to simply put “no” in front of it., if you want to unassign the IP address you had assigned:

```
cisco(config)#interface serail 1/1
cisco(config-if)#no ip address 192.168.155.2 255.255.255.0
cisco(config-if)#ctrl-Z
cisco#show interface serial 1/1
```

Static routing (manual entry)

- First of all, enable the routing function:

```
cisco(config)#ip routing
cisco(config)#ctrl_Z
```

- Static routing can be configured:

```
cisco#config
cisco (config)#ip route  192.168.3.0 255.255.255.0  192.168.150.1
cisco (config)#ctrl-Z
cisco#show ip route
```

Destination network

Destination network mask

Next hop address

Dynamic routing

- Use RIP in an intranet.

```
cisco#config
cisco(config)#router rip
cisco(config-router)#network 192.168.100.0
cisco(config-router)#network 192.168.101.0
cisco(config-router)#ctrl-Z
cisco#show ip protocols
```

These two networks must have been configured at certain interfaces.

Finish the configuration

- Type command

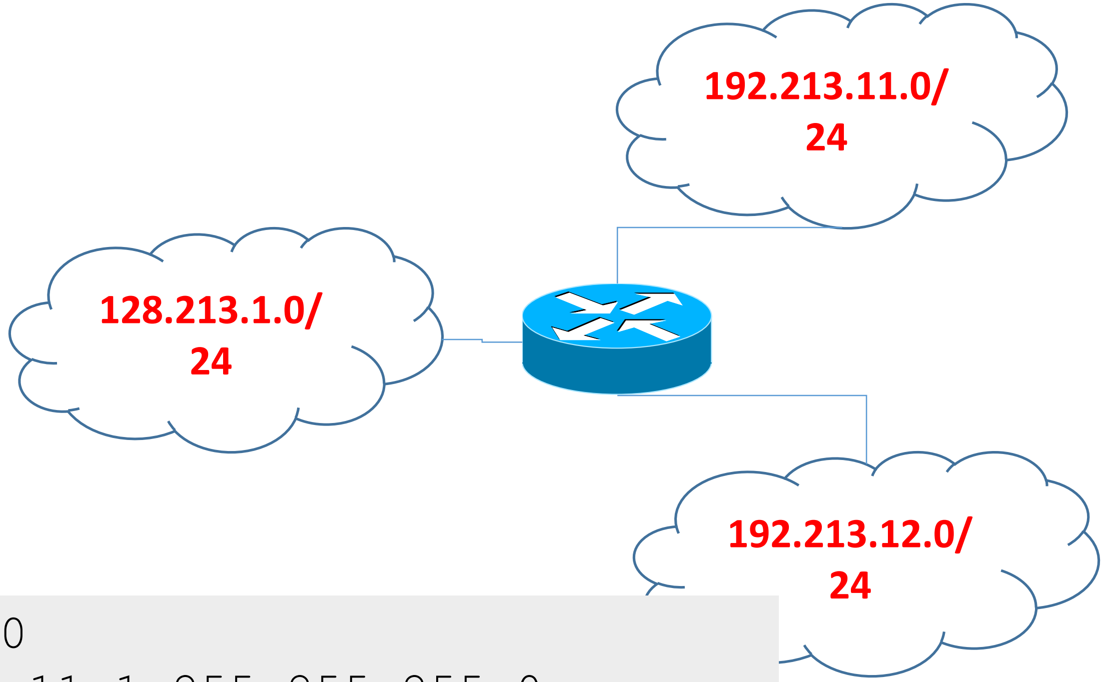
```
cisco#copy running-config startup-config
```

CISCO OSPF Configuration

- Enabling OSPF involves two steps in config mode:
  - Enable OSPF process using "***router ospf <process-id>***" command
  - Assign areas to the interfaces using "***network <network-addr> <mask> <area-id>***" command.

- Process-id can be given any value. A router may run multiple OSPF processes.

Example




```
Cisco# interface Ethernet0
Cisco# ip address 192.213.11.1 255.255.255.0
Cisco# interface Ethernet1
Cisco# ip address 192.213.12.2 255.255.255.0
Cisco# interface Ethernet2
Cisco# ip address 128.213.1.1 255.255.255.0

Cisco# router ospf 100
Cisco# network 192.213.0.0 0.0.255.255 area 0.0.0.0
Cisco# network 128.213.1.1 0.0.0.0 area 23
```