Chapter 6
Delivery and Routing of IP Packets

Outline

- Delivery
- Forwarding
- Routing
- Structure of a Router

Delivery v.s. Routing

- Delivery
  - The way a packet is handled by the underlying networks under the control of the network layer
    - Connectionless v.s. connection-oriented service
    - Direct v.s. indirect delivery
- Routing
  - The way routing tables are created to help in forwarding
    - Static and dynamic routing

6.1

Delivery

Delivery

- Delivery of a packet
  - The network layer supervises the handling of the packet by the underlying physical networks

- Two important concepts
  - ***Connection types***
    - *Connection-oriented v.s. connectionless services*
  - ***Direct versus indirect delivery***

Connection Types

- ***Connection-Oriented Services***
  - The network layer protocol first makes a connection before sending a packet
  - There is a relationship between the packets
    - They are sent on the same path in ***sequential order***
  - The decision about the route of a sequence of packets is made only one
    - When the connection is established
- ***Connectionless Services***
  - The network layer treat each packet independently
  - Each packet having no relationship to any other packet
  - IP protocol is a connectionless protocol

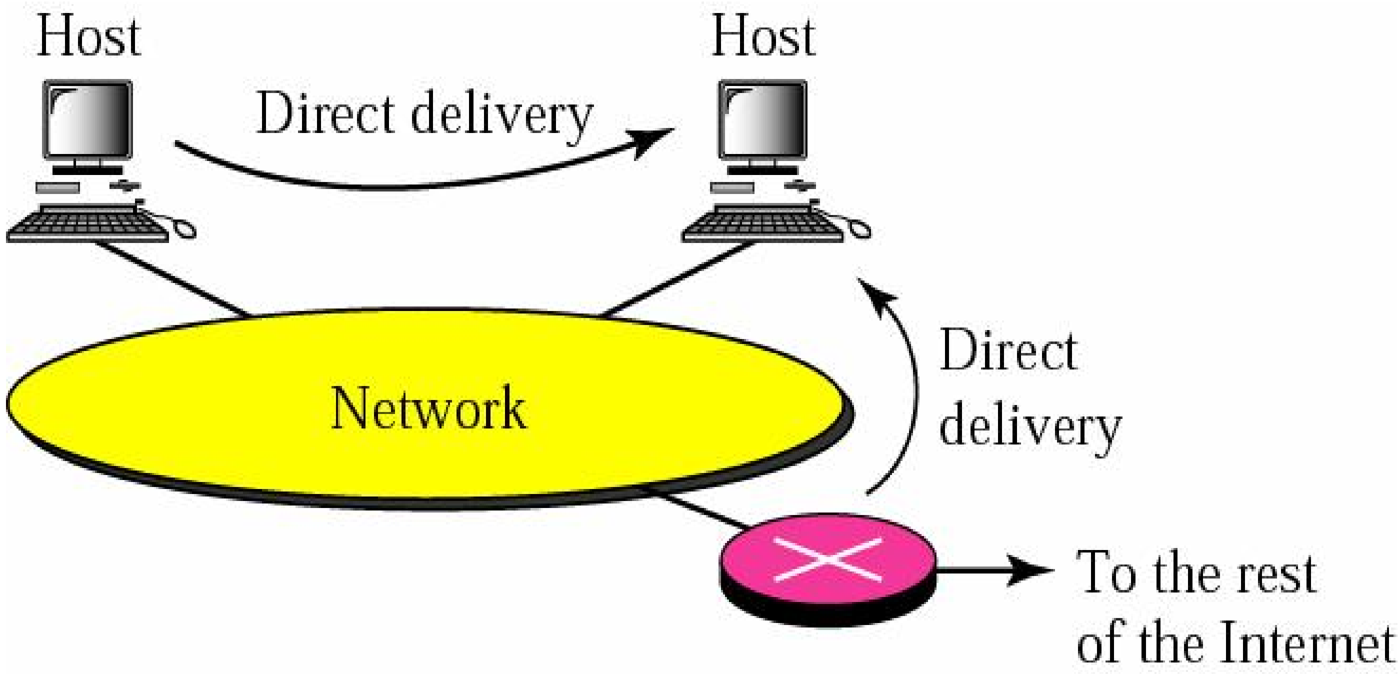
Direct Versus Indirect Delivery

- The delivery of a packet may be direct or indirect

Direct Delivery

- The final destination is a host in the same physical network as the deliverer
  - When the source and destination are located on the same physical network
  - Or the delivery is between the last router and the destination host

Direct Delivery



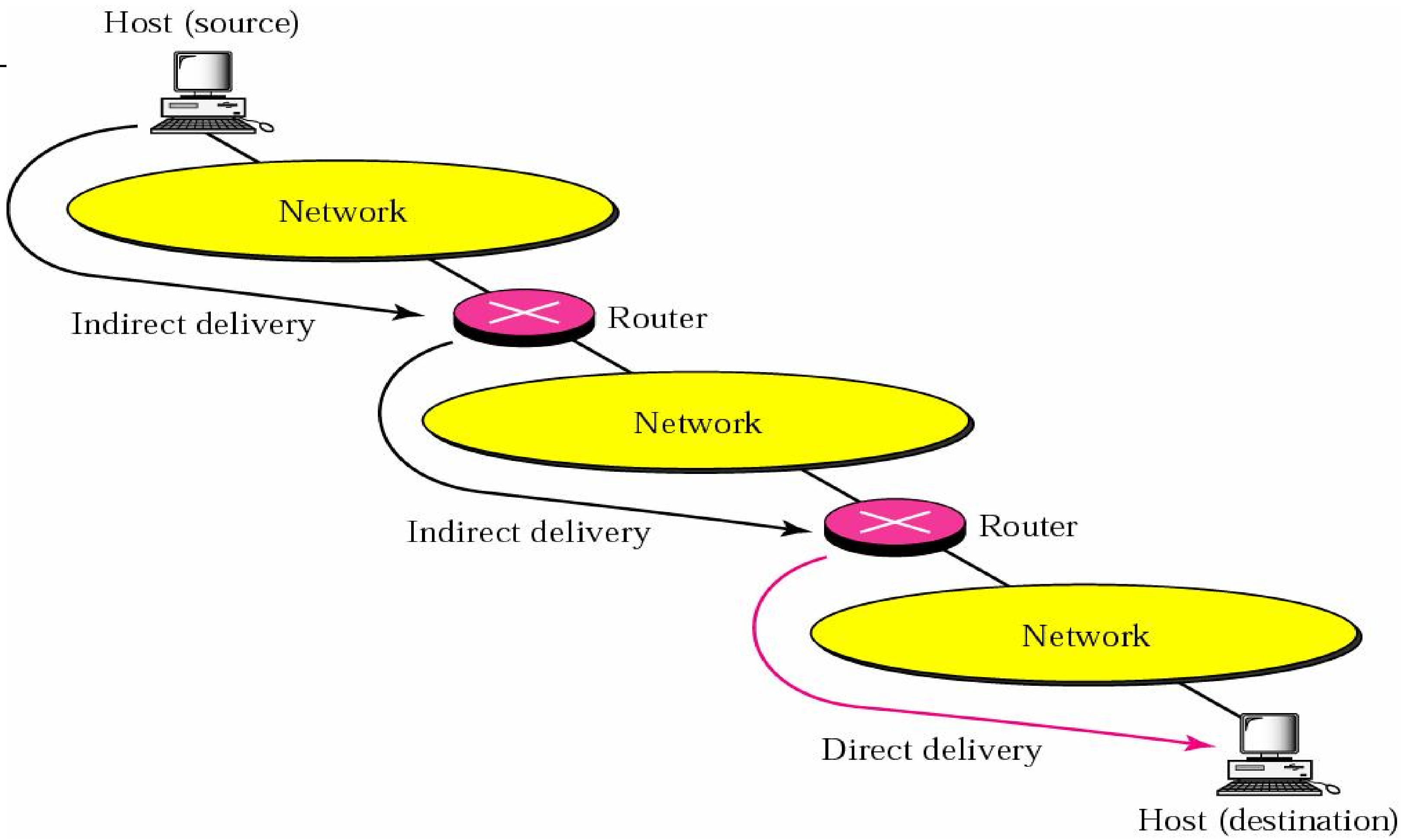
Direct Delivery (Cont.)

- How to determine if the delivery is direct
  - Compare the *network addresses* between *the destination* and *the current network*
- For direct delivery, the sender uses the destination *IP* address to find the destination *physical address*
  - Static method: finding a table
  - Dynamic method: use the address resolution protocol (ARP)

Indirect Delivery

- The destination host and the deliverer are not on the same network
  - Packet goes from routers to routers
- For indirect delivery
  - The sender uses *the destination IP address* and *a routing table* to find *the next router's IP address*
  - Then, the sender uses ARP protocol to find the next router's physical address

Indirect Delivery



Indirect Delivery (Cont.)

- A delivery always involves one direct delivery but zero or more indirect deliveries

- Besides, the last delivery is always a direct delivery

6.2

Forwarding

Forwarding

- ***Forwarding***: place the packet in its route to its destination
- Forwarding requires a host/router to have a *routing table*
- However, with the increase of hosts,
  - The number of entries in the routing table also increase
- Techniques to decrease the table size and handle issues such as security
  - *Next-hop routing*
  - *Network-specific routing*
  - *Host-specific routing*
  - *Default routing*

Next-Hop Routing

- Routing table holds only the address of the next hop
  - Instead of holding information about the complete route

Figure 6.2

Next-Hop Routing

Routing table for host A

Destination	Route
Host B	R1, R2, Host B

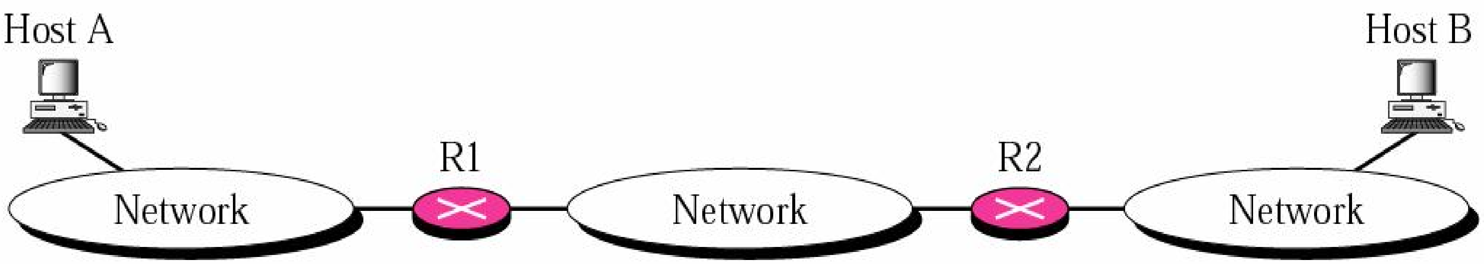
Routing table for R1

Destination	Route
Host B	R2, Host B

Routing table for R2

Destination	Route
Host B	Host B

a. Routing tables based on route



Routing table for host A

Destination	Next Hop
Host B	R1

Routing table for R1

Destination	Next Hop
Host B	R2

Routing table for R2

Destination	Next Hop
Host B	Ñ

b. Routing tables based on next hop

Network-Specific Routing

- Use only one entry to define *the address of the network itself, i.e., network address*
  - Instead of having an entry for every host connected to the same physical network

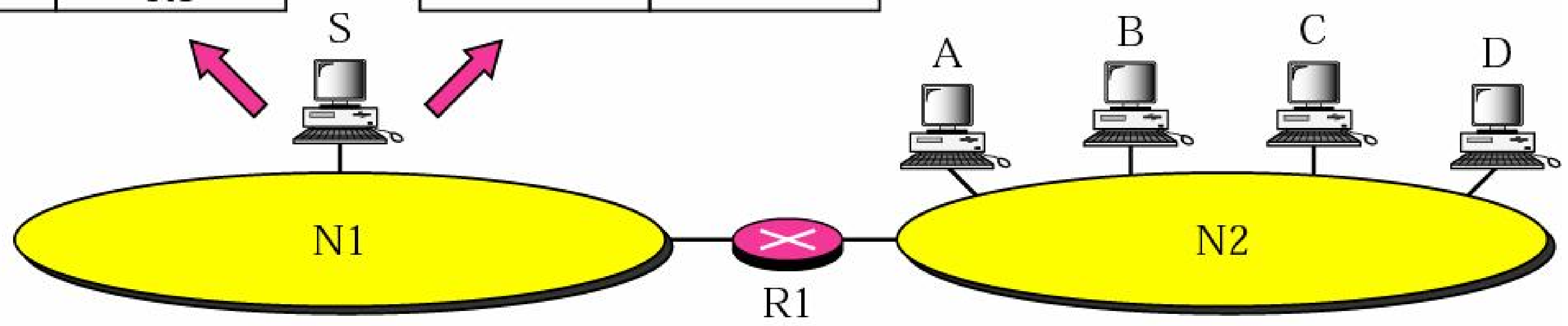
Network-Specific Routing

Routing table for host S based on host-specific routing

Destination	Next Hop
A	R1
B	R1
C	R1
D	R1

Routing table for host S based on network-specific routing

Destination	Next Hop
N2	R1



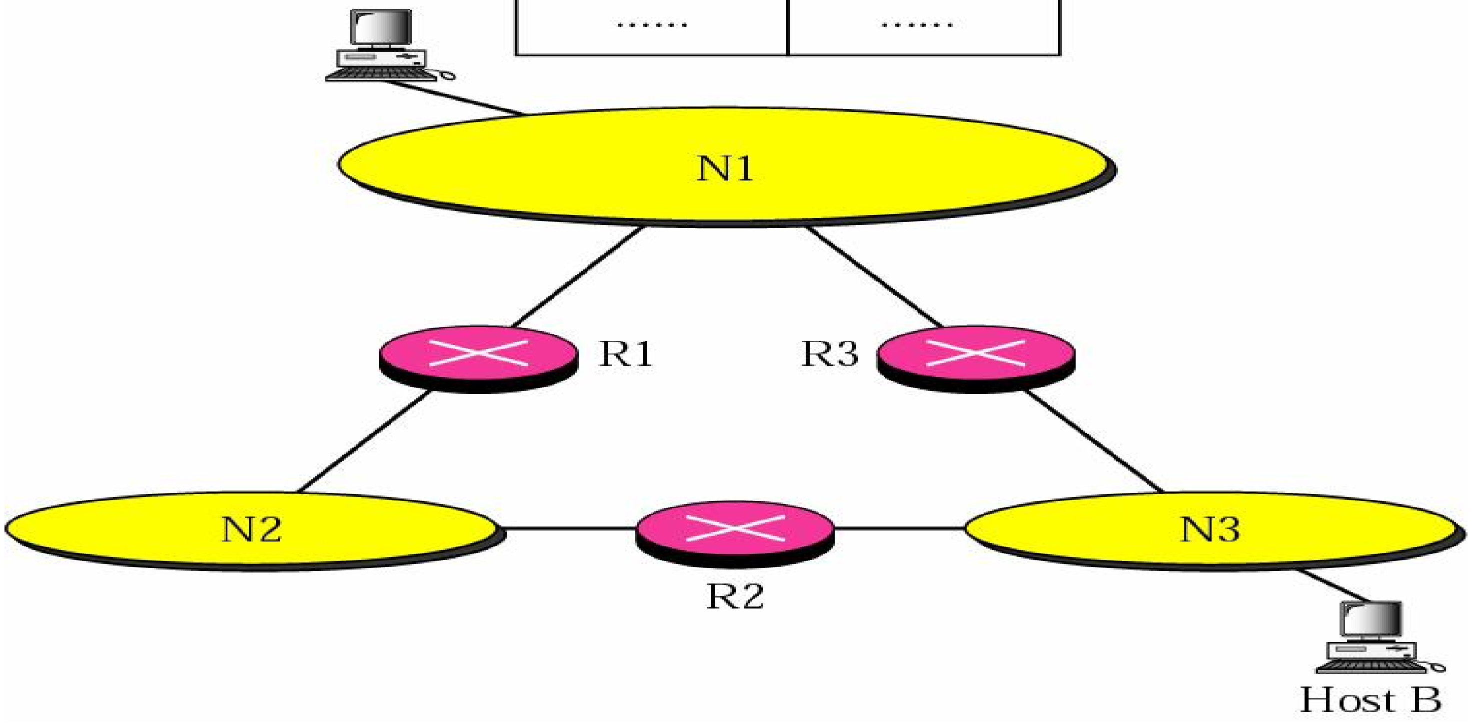
Host-Specific Routing

- The destination host address is given in the routing table
- The inverse of network-specific routing
- Not efficient for performance
  - But, in some occasions, the administrator wants to have more control over routing
    - Checking the route
    - Providing security

Host-Specific Routing

Routing table for host A

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



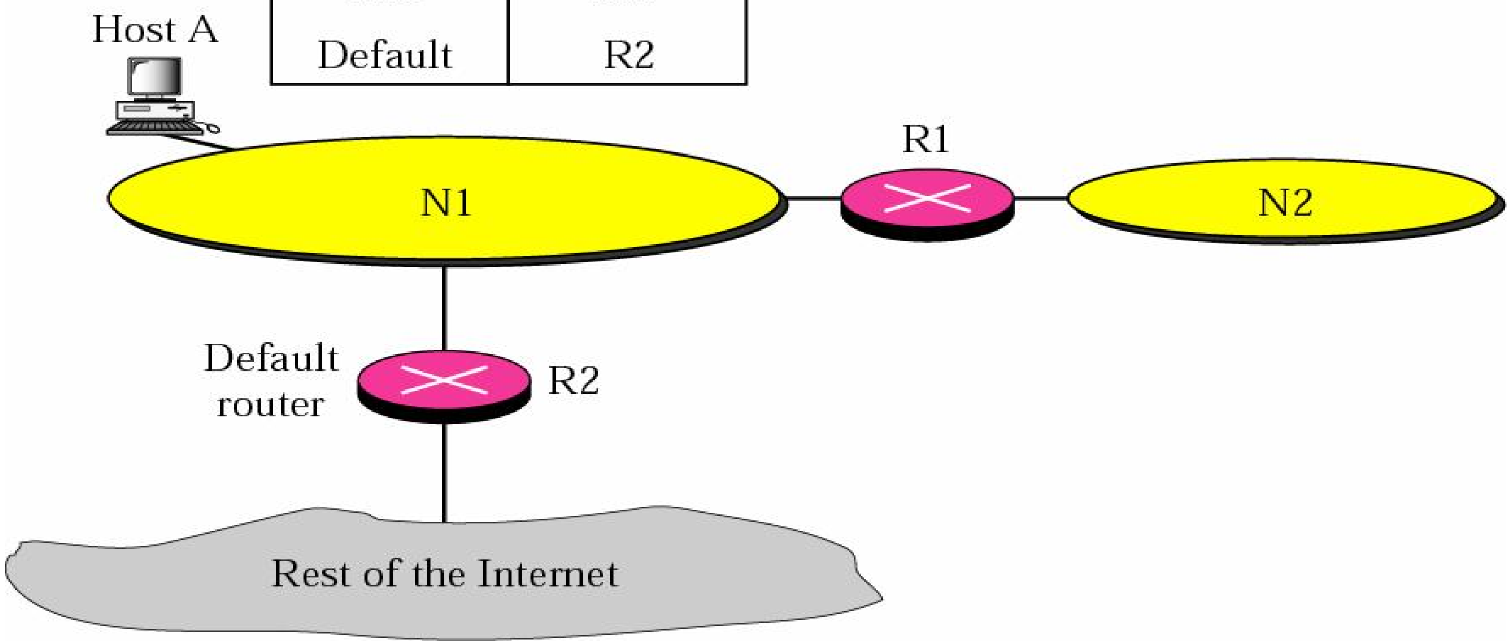
Default Routing

- Instead of listing all networks in the routing table
  - Just use one entry called *default*
  - Network address is 0.0.0.0

Default Routing

Routing table for host A

Destination	Next Hop
N2	R1
......	......
Default	R2



Forwarding with Classful Addressing

- The existence of a default mask in a classful address makes for forwarding process simple

- Discussions
  - Forwarding without subnetting
  - Forwarding with subnetting

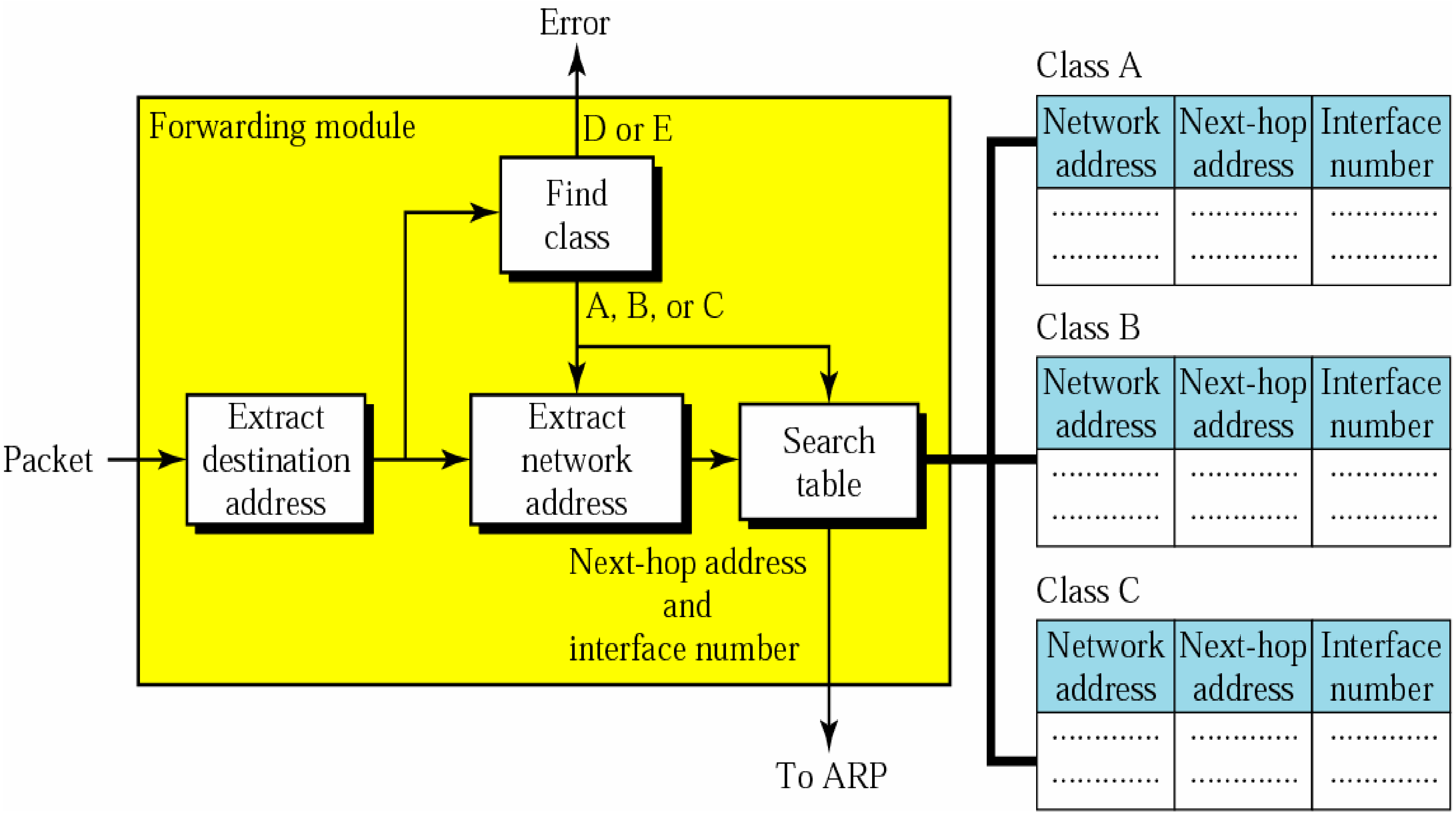
Forwarding without Subnetting

- Most routers in classful addressing are not involved in subnetting
  - Since subnetting happens inside an organization
- A forwarding module would consists of three tables
  - One for each unicast class
  - If support multicast, add one more table to handle class D addresses

Forwarding without Subnetting (Cont.)

- Each routing table has a minimum of three columns
  - ***Network address of the destination network***
    - Tell us where the destination host is located
    - Assume use the network-specific forwarding, not the host-specific forwarding
  - ***Next-hop address***
    - Tell which router the packet must be delivered for an indirect delivery
    - Empty for a direct delivery
  - ***Interface number***
    - Define the outgoing port from which the packet is sent out

Simplified Forwarding Module in Classful Address without Subnetting

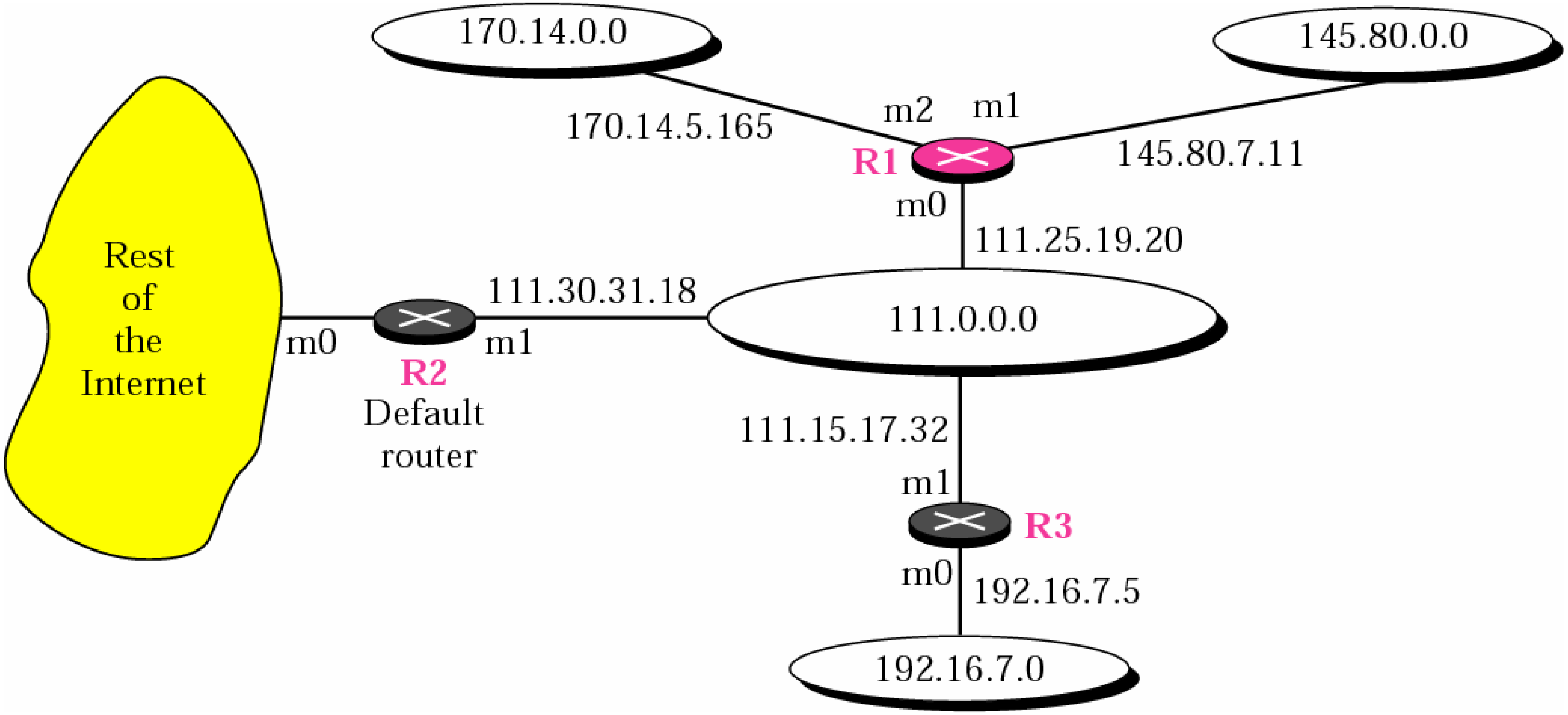


Steps of Forwarding without Subnetting

- Extract the *destination address* of the packet
- *Make a copy of the destination address and find the class of its address*
  - ***Shift*** the copy of the address 28 bits to the right
  - The result is a 4-bit number used to determine the class
- Use the result in steps 2 to extract the *network address*
  - By ***masking off*** the rightmost 8, 16, or 24 bits since we have the mask now
- The *class of the address* and the *network address* are used to find out next-hop information
  - The class determines the table to be searched
  - Then search the table for the network address
    - If found, use the ***next-hop address*** and the ***interface number***; else, use the ***default route***
- Use the ARP module to find the *physical address* of the next router
  - By the next-hop address and the interface number

Example 1

- Show the routing table for router R1



Example 1: Solution

- Following figure shows the three tables used by router R1.
- Some entries in the next-hop address column are empty
  - Because the destination is in the same network to which the router is connected (direct delivery).
  - Thus, the next-hop address used by ARP is simply the destination address of the packet

Table for Example 1

Class A

Network address	Next-hop address	Interface
111.0.0.0	------------	m0

Class B

Network address	Next-hop address	Interface
145.80.0.0	------------	m1
170.14.0.0	------------	m2

Class C

Network address	Next-hop address	Interface
192.16.7.0	111.15.17.32	m0

Default: 111.30.31.18, m0

Example 2

- Router R1 in above figure receives a packet with destination address 192.16.7.14. Show how the packet is forwarded
- Solution
  - The destination address is 11000000 00010000 00000111 00001110.
  - A copy of the address is shifted 28 bits to the right. The result is 00000000 00000000 00000000 00001100 or 12.
    - The destination network is class C.
  - The network address is extracted by masking off the leftmost 24 bits of the destination address; the result is 192.16.7.0.
  - The table for Class C is searched and the network address is found in the first row.
  - The next-hop address 111.15.17.32. and the interface m0 are passed to ARP

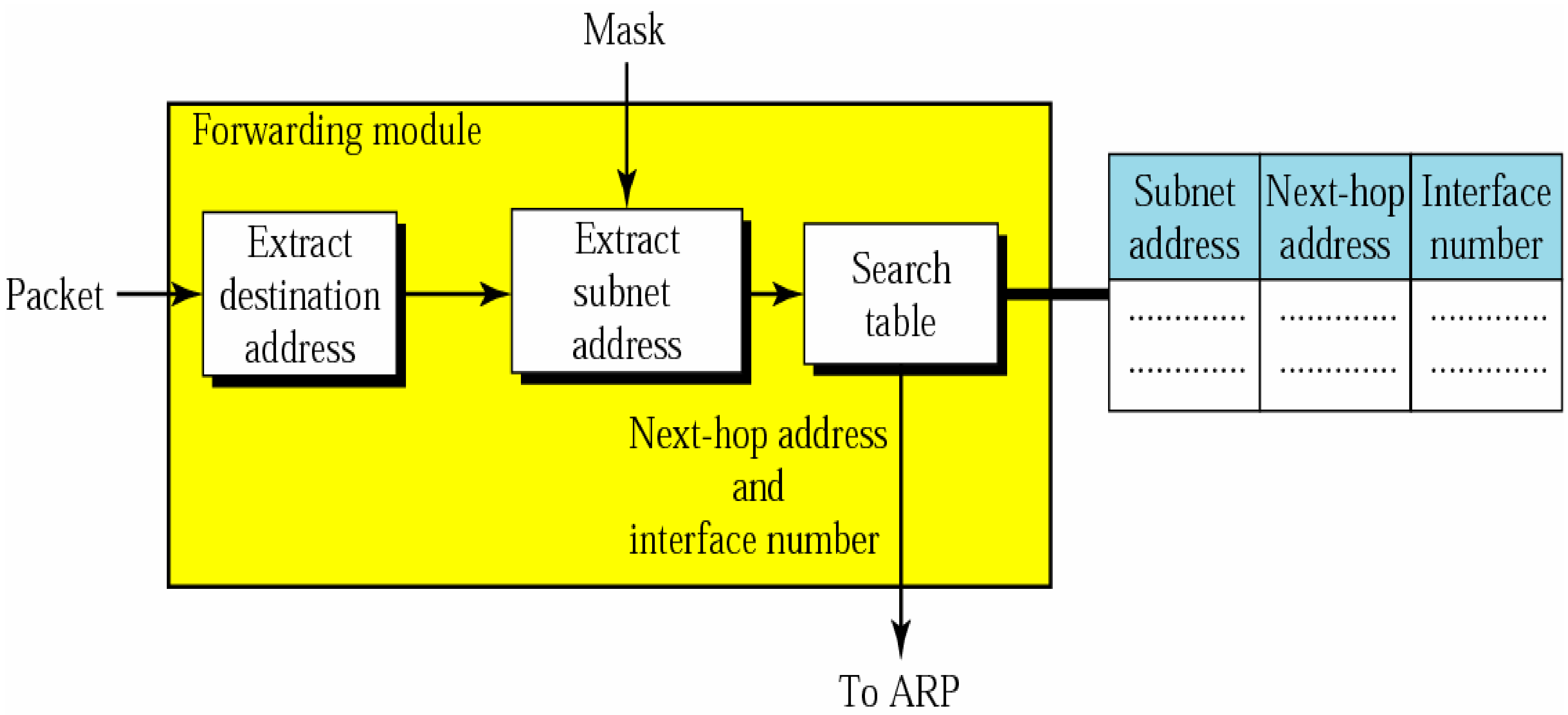
Example 3

- Router R1 in above figure receives a packet with destination address 167.24.160.5. Show how the packet is forwarded.
- Solution
  - The destination address is 10100111 00011000 10100000 00000101.
  - A copy of the address is shifted 28 bits to the right. The result is 00000000 00000000 00000000 00001010 or 10.
    - The class is B.
  - The network address is 167.24.0.0.
  - The table for Class B is searched.
    - No matching network address is found.
    - The packet needs to be forwarded to the default router (the network is somewhere else in the Internet).
  - The next-hop address 111.30.31.18 and the interface number m0 are passed to ARP.

Forwarding with Subnetting

- In classful addressing, subnetting happens inside an organization
- The routers that handle subnetting
  - At the border of the organization site
  - Inside the site boundary
- Number of routing tables
  - If variable-length subnetting is used
    - We need several tables
  - Otherwise, only one is enough

Simplified Forwarding Module in Classful Address with Subnetting



Steps in Forwarding with Subnetting

- Extract the *destination address* of the packet
- Extract the subnet address by the *destination address* and the *mask*
- Search the table by the *subnet address* to find the *next-hop address* and the *interface number*
  - If no mach, use the *default route*
- Pass the *next-hop address* and the *interface number* to the ARP

Example 4

- Following figure 6.11 shows a router connected to four subnets

- Note
  - The router is configured to apply the mask /18 to any destination address

Example 5

- The router in above figure receives a packet with destination address 145.14.32.78. Show how the packet is forwarded
- Solution
  - The mask is /18.
  - After applying the mask, the subnet address is 145.14.0.0.
  - The packet is delivered to ARP with the next-hop address 145.14.32.78 and the outgoing interface m0

Example 6

- A host in network 145.14.0.0 in above figure has a packet to send to the host with address 7.22.67.91. Show how the packet is routed
- Solution
  - The router receives the packet and applies the mask (/18).
    - The network address is 7.22.64.0.
  - The table is searched and the address is not found.
    - The router uses the address of the default router (not shown in figure) and sends the packet to that router

Forwarding with Classless Addressing

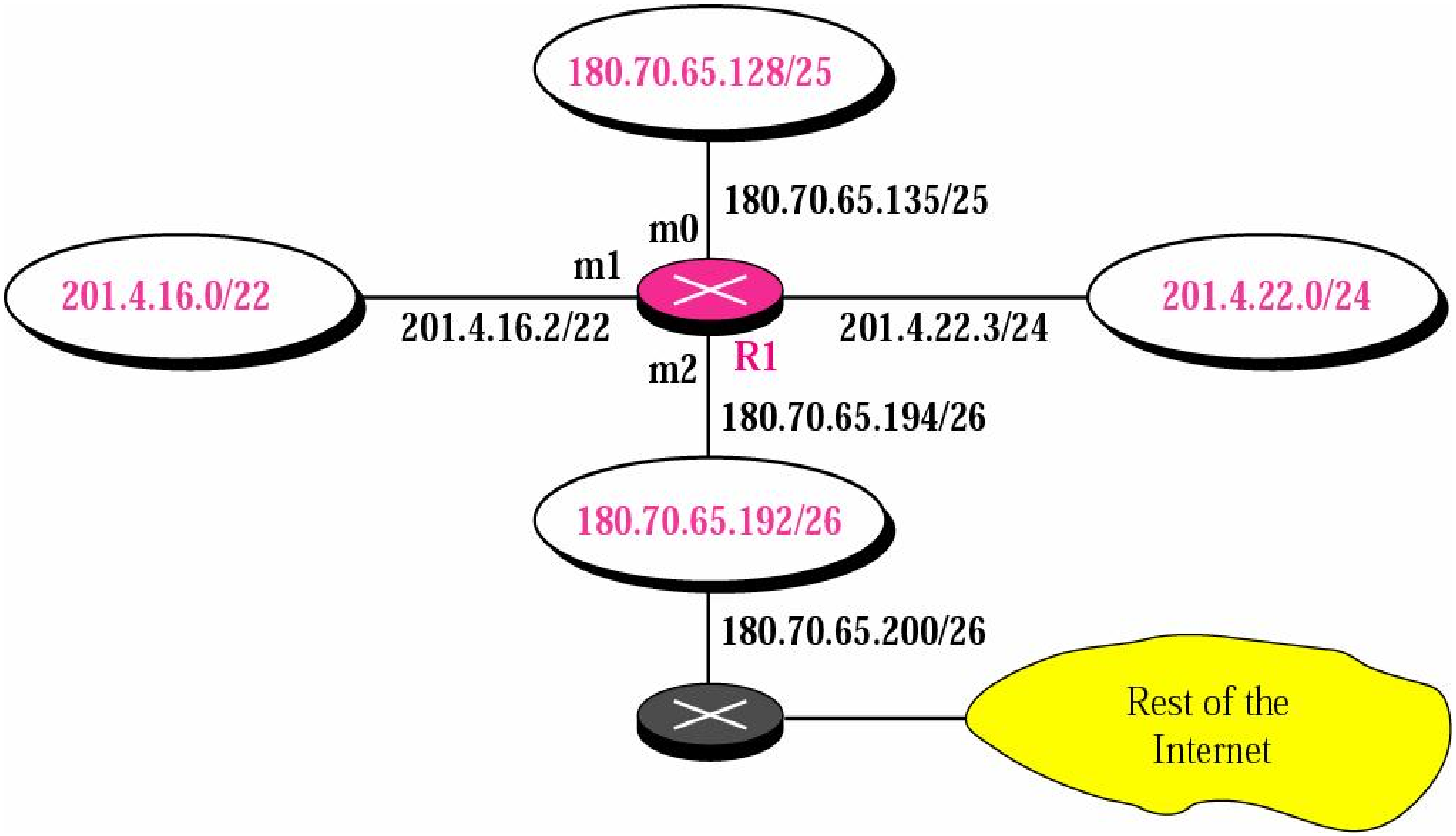
- In classless addressing
  - The whole address space is one entity
  - i.e., there is no classes
  - Thus, forwarding requires one row for each block
- However, cannot derive the network address from the destination address in the packet
- Solution
  - Include the *mask* (/n) of the corresponding block in the table
  - Thus, a classless routing table needs at least four columns

Example 7

- Make a routing table for router R1 using the configuration in Figure 6.13
- *Solution:*

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
Default	Default	180.70.65.200	m2

Configuration for Example 7



Example 8

- Show the forwarding process if a packet arrives at R1 in above figure with the destination address 180.70.65.140
- Solution: the router performs the following steps:
  - 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
  - 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* (the destination address of the packet in this case) and the *interface number m0* are passed to ARP for further processing

Example 9

- Show the forwarding process if a packet arrives at R1 in above figure with the destination address 201.4.22.35
- Solution: the router performs the following steps
  - The first mask (/26) is applied to the destination address.
    - The result is 201.4.22.0 => (X).
  - The second mask (/25) is applied to the destination address.
    - The result is 201.4.22.0 => (X).
  - The third mask (/24) is applied to the destination address.
    - The result is 201.4.22.0 => (0)
    - The destination address of the package and the interface number m3 are passed to ARP

Example 10

- Show the forwarding process if a packet arrives at R1 in above figure with the destination address 18.24.32.78
- Solution
  - This time all masks are applied 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 some place else in the Internet

Example 11

- The routing table for router R1 is given in Table 6.2. Can we draw its topology?

Mask	*Network Address*	*Next-Hop Address*	*Interface Number*
/26	140.6.12.64	180.14.2.5	m2
/24	130.4.8.0	190.17.6.2.0	m1
/16	110.70.0.0	-----------------	m0
/16	180.14.0.0	-----------------	m2
/16	190.17.0.0	-----------------	m1
Default	Default	110.70.4.6	m0

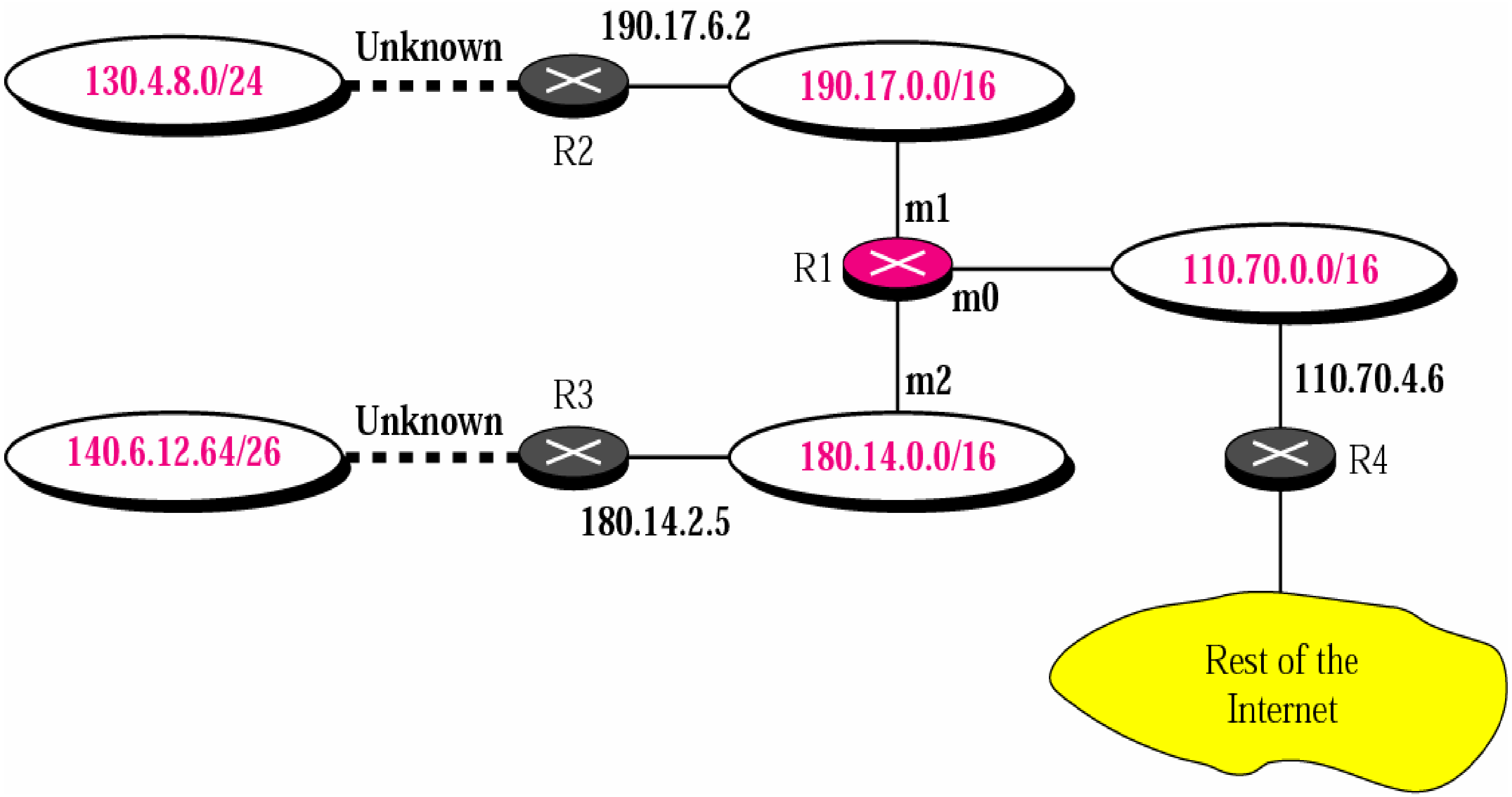
Example 11: Solution

- We know some facts but we don’t have all for a definite topology.
- Router R1 has three interfaces: m0, m1, and m2.
- There are three networks directly connected to router
  - 110.70.0.0/16, 180.14.0.0/16, 190.17.0.0/16
- There are two networks indirectly connected to R1
  - 140.6.12.64/26 and 130.4.8.0/24
- There must be at least three other routers involved
  - From the next-hop column

Example 11: Solution (Cont.)

- One router, the default router, is connected to the rest of the Internet.
- However,
  - We do not know if network 130.4.8.0 and 140.6.12.64 is directly connected to router R2 and R3 or through a point-to-point network (WAN) and another router.
- Figure 6.14 shows our guessed topology

Guessed topology for Example 6



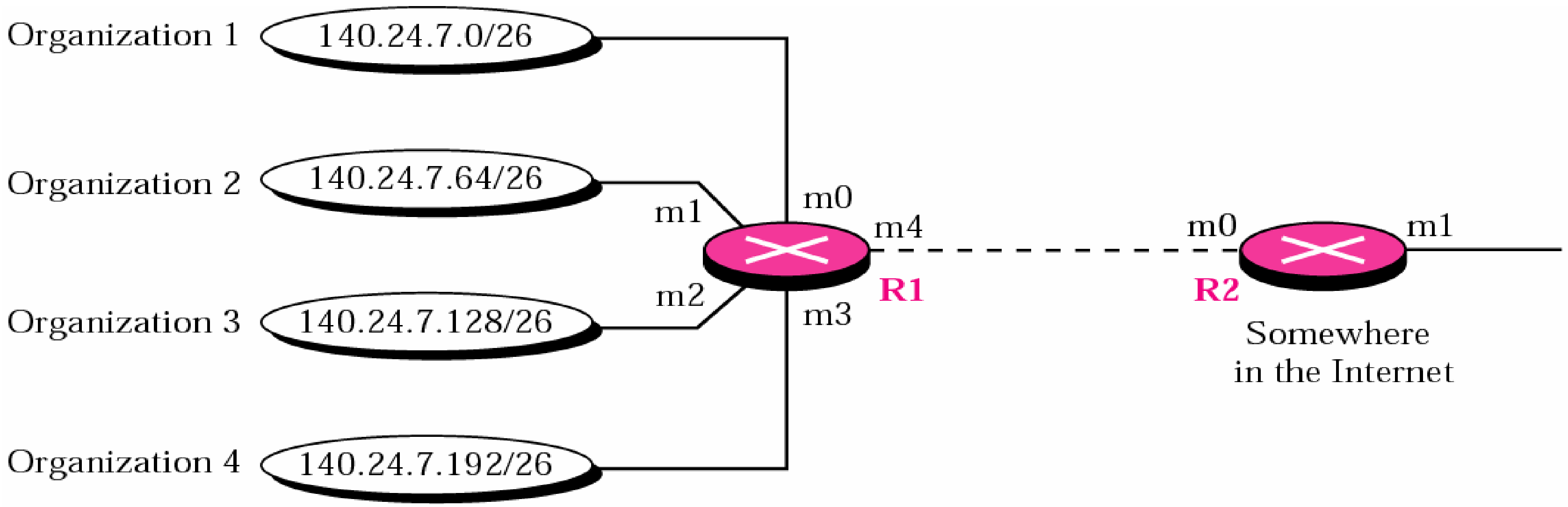
Address Aggregation

- In classful address
  - There is only one entry for each site
  - Even this site is subnetted
- In classless address
  - The number of entries will increase
    - Since classless addressing is to divide up the whole address space into manageable blocks
- Problem:
  - Vast routing table and increased overhead to search table

Address Aggregation (Cont.)

- Solutions: ***address aggregation***
- In the following figure
  - R1 has a longer routing table
    - Connect four organization
  - R2 has a small routing table
    - Any packet with destination 140.24.7.0~140.24.7.255 is sent out from interface m0
      - Regardless of the organization number
    - Called ***address aggregation***
      - The blocks of addresses for four organization are aggregated into one large block
      - We do not have to specify each organization with an entry

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 router	m4

Routing table for R1

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

Routing table for R2

Longest Mask Matching

- However, if organization 4 is not geographically close to the other three routers
  - Can we still use the address aggregation and assign 140.24.7.192/26 to organization 4?
- Ans: Yes
  - Since routing in classless addressing uses ***longest mask matching***
- ***Longest mask matching***
  - Routing table is sorted from the longest mask to the shortest mask
  - If there are three masks, /27, /26, /24
    - /27 is the first entry and /24 is the last one

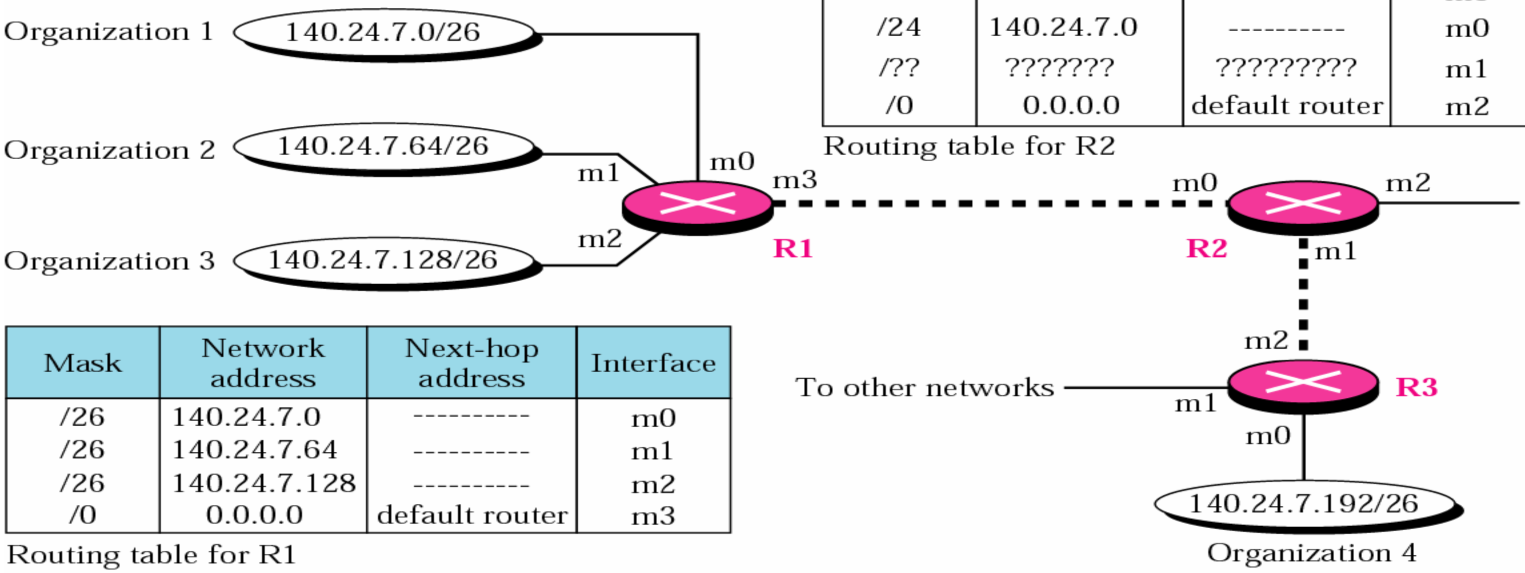
Longest Mask Matching (Cont.)

- Example, in the following figure
  - Organization 4 is moved
  - If an packet with destination address 140.24.7.200 is arrived at R2
    - The first mask at router R2 is applied => match
    - Route from interface m1 and reach organization 4
  - If not use ***longest mask matching***
    - Apply the /24 mask would result in the incorrect routing
- Thus
  - We can still aggregate organization 1~3 into a large block while still assign 140.24.7.192/26 to organization 4

Longest Mask Matching (Cont.)

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 router	m2

Routing table for R2



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 router	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 router	m2

Routing table for R3

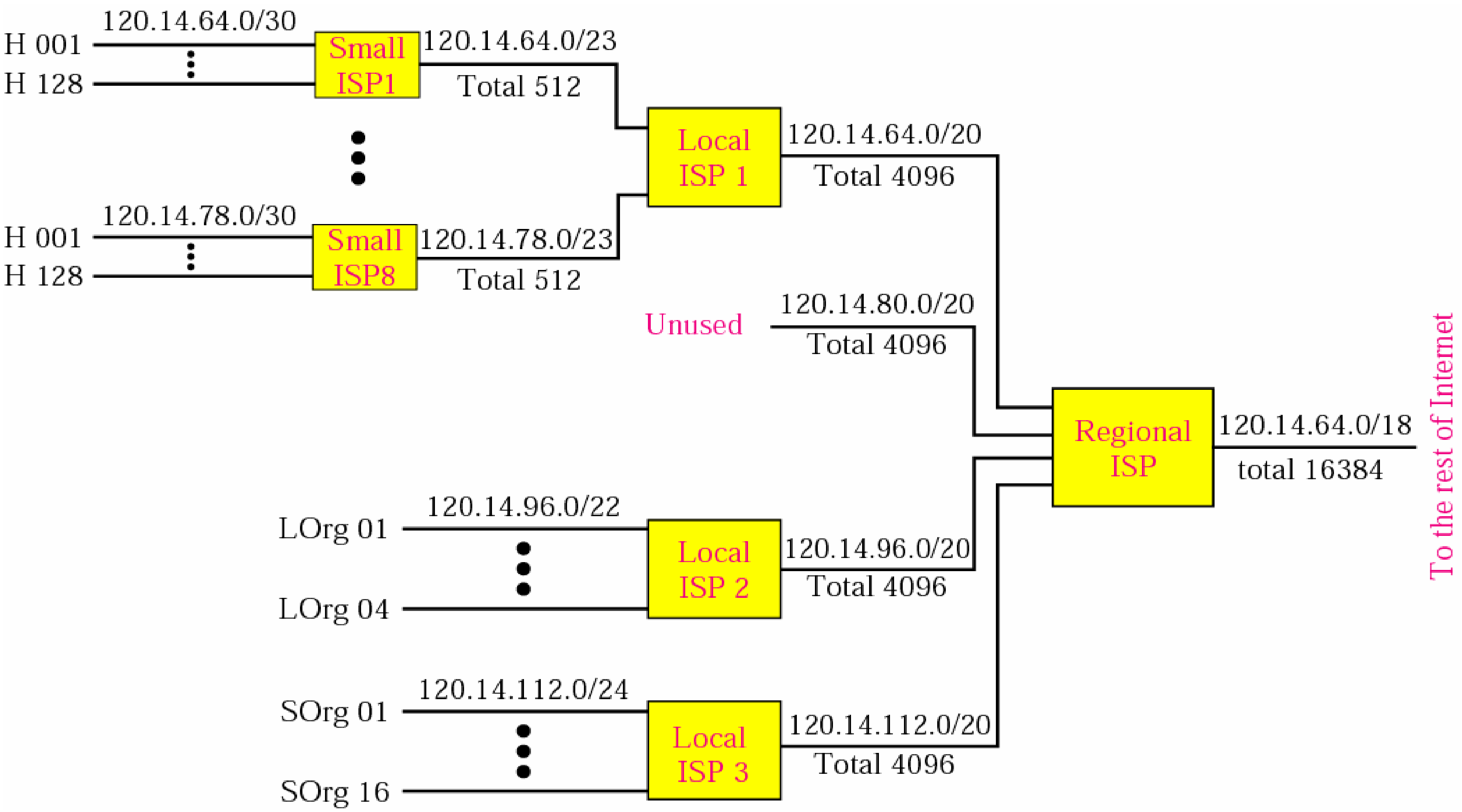
Hierarchical Routing

- To further reduce the size of routing table
  - Create a sense of hierarchy in the Internet architecture and routing tables
- Internet is divided into international and national ISP
  - National ISP are divided into regional ISPs
  - Regional ISP are divided into local ISPs
- If a local ISP divide its block (a.b.c.d/n) into smaller blocks
  - However, to the rest of the Internet, routers do not know such a division
  - Thus, all customers of the local ISP are still defined as a.b.c.d/n to the rest of Internet

Example 12

- Consider the following figure
  - A regional ISP is granted 16384 addresses starting from 120.14.64.0/18
  - 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

Hierarchical Routing with ISPs



Example 12 (Cont.)

- The first local ISP has divided its assigned subblock into 8 smaller blocks
  - Assigned each to a small ISP.
  - Each small ISP provides services to 128 households (H001 to H128), each using four addresses.
- The second local ISP has divided its block into 4 blocks
  - Assigned the addresses to 4 large organizations (LOrg01 to LOrg04).
- The third local ISP has divided its block into 16 blocks
  - Assigned each block to a small organization (SOrg01 to SOrg15).

Example 12 (Cont.)

- 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.
  - The regional ISP sends every packet with destination address 120.14.64.0 to 120.14.79.255 to Local ISP1.
  - Local ISP1 sends every packet with destination address 120.14.64.0 to 120.14.64.3 to H001.

Geographical Routing

- To decrease the size of the routing table even further
  - Extend hierarchical routing to include geographical routing
- Divide the entire address space into a few large blocks
  - A block to North America
  - A block to Europe
  - A block to Asia
  - A block to Africa
- The routers of ISP outside of Europe have only one entry for packets to Europe in their routing table

Routing Table Search Algorithm

- The algorithms used in *search* routing tables in classful addressing must be changed for classless addressing

- The algorithm used for *updating* routing tables in classful addressing must be changed for classless addressing
  - Mentioned in Chapter 14

Searching in Classful Addressing

- The routing table in classful addressing is organized as a list
- However, to make search easier, the routing table can be divided into three buckets (areas)
  - One for each class
- When a packet arrives, applies the *default mask* to find the corresponding bucket (class A, B, or C)
  - Notably, from a address, we can derive which class it belongs to

Searching in Classless Addressing

- In classless routing, we can also use buckets
  - However, 32 buckets are used instead of three
    - Each buckets corresponding to each prefix length
  - When a packet arrives, try the longest prefix (/32), then the next prefix (/31) and so on until matched
    - *Longest match method*
- However, this search method would also take quite a long time
  - Use other data structures such as tree or binary tree

Combination

- Modern routers are all based on classless addressing
  - All include mask in the routing table
  - The classful addressing is treated as a special case of classless addressing
    - Use the mask /24, /16, /8

6.3

ROUTING

Static Versus Dynamic Routing Table

- Static Routing Table
  - The entries are entered *manually*
  - Cannot be updated unless manually altered by administrator
- Dynamic Routing Table
  - Update periodically using dynamic routing protocol
    - RIP, OSPF, or BGP
  - If a router shutdown or a link is broken
    - Update the tables accordingly

Routing Table

- As mentioned before, routing has a minimum of three columns
  - However, some of today's routers have even more columns
  - Vendor dependent
- Common fields in modern routing table
  - Mask, network address, next-hop address, interface, flags, reference count, use
- **Mask**: define the mask applied for the entry
- **Network address**:
  - If host-specific routing, define the address of the destination host

Routing Table (Cont.)

- **Next-hop address**
  - The address of the next-hop router
- **Interface**
  - The name of interface
- **Reference count**
  - The number of *users* that are using this route at any moment
- **Use**
  - The number of *packets* transmitted through this router for the corresponding destination

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

Routing Table (Cont.)

- Flags
  - U (Up): the router is up and running.
    - If not present, cannot forward packet to this router
  - G (Gateway): destination is in another network and use *indirect delivery*
    - If not present, use direct delivery
  - H (Host-specific): the entry in the destination field is host-specific address
    - If not present, destination field is network-specific address

Routing Table (Cont.)

- Flags
  - D (Added by redirection): routing information for this destination has been *added* by a *redirection message* from ICMP.
  - M (Modified by redirection): routing information for this destination has been *modified* by a *redirection message* from ICMP.
    - Discuss in Chapter 9

Figure 6-8

Routing Table

Mask	Destination address	Next-hop address	Flags	Reference count	Use	Interface
255.0.0.0	124.0.0.0	145.6.7.23	UG	4	20	m2
....	....		....	....	....	....
....	....	....	....	....	....	....

Flags

U The router is up and running.
G The destination is in another network.
H Host-specific address.
D Added by redirection.
M Modified by redirection.

Example 13

- In UNIX or LINUX, ***netstat*** is used to find the contents of a routing table for a host or router
- The following shows the listing of the contents of the default server.
  - Note that this is a routing table for a host, not a router.
  - A host also needs a routing table

Example 13 (Cont.)

$ 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.**

- **Gateway** = the address of next hop
  - 0.0.0.0 = direct delivery
- **G** in Flags means the destination is reached through a router
- Interface ***lo*** is a virtual loopback interface

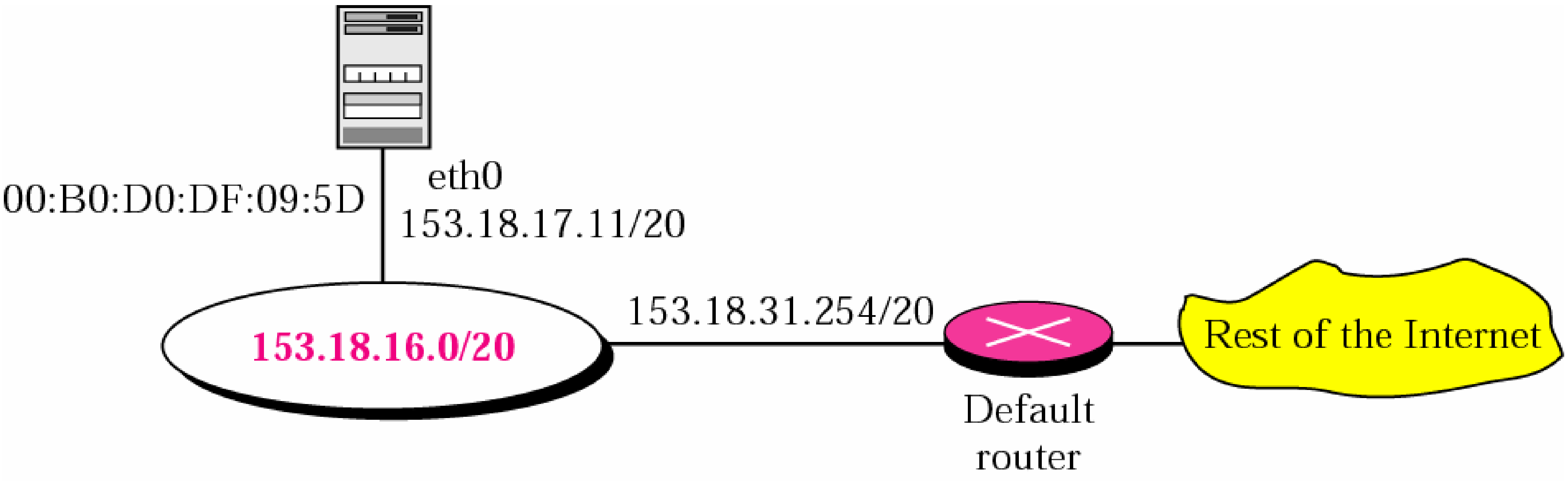
Example 13 (Cont.)

- More information about the IP address and physical address of the server
  - Can be found 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
....
```

Example 13 (Cont.)

- From the above information, we can deduce the configuration of the server as shown below

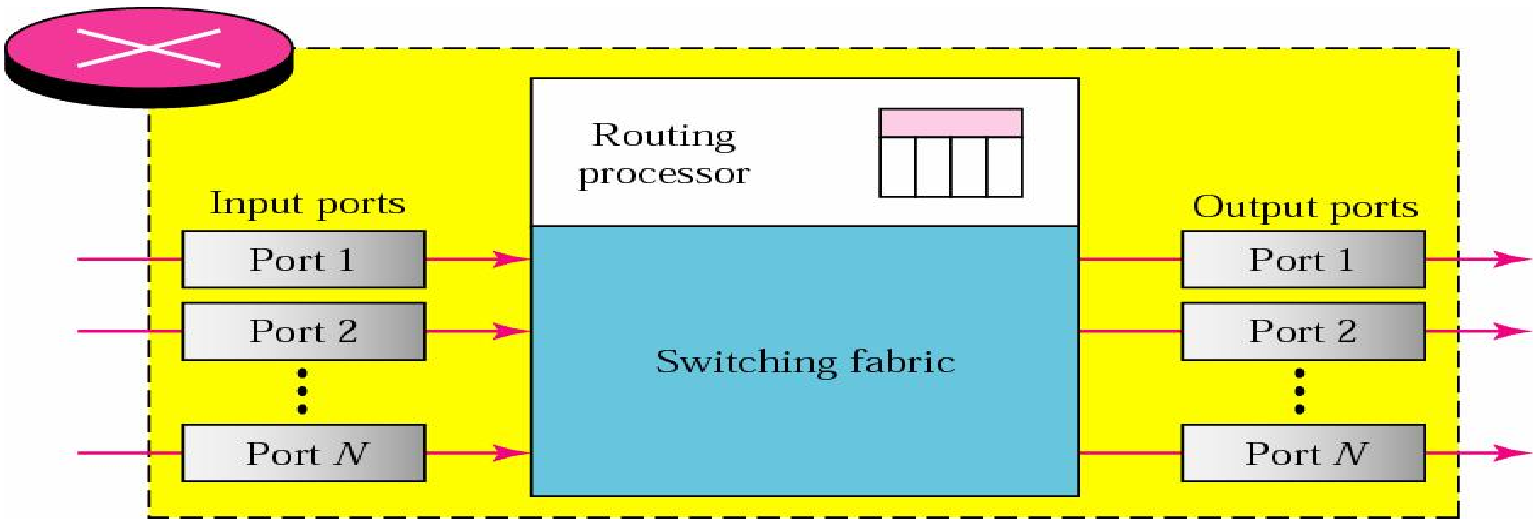


6.4

STRUCTURE OF A ROUTER

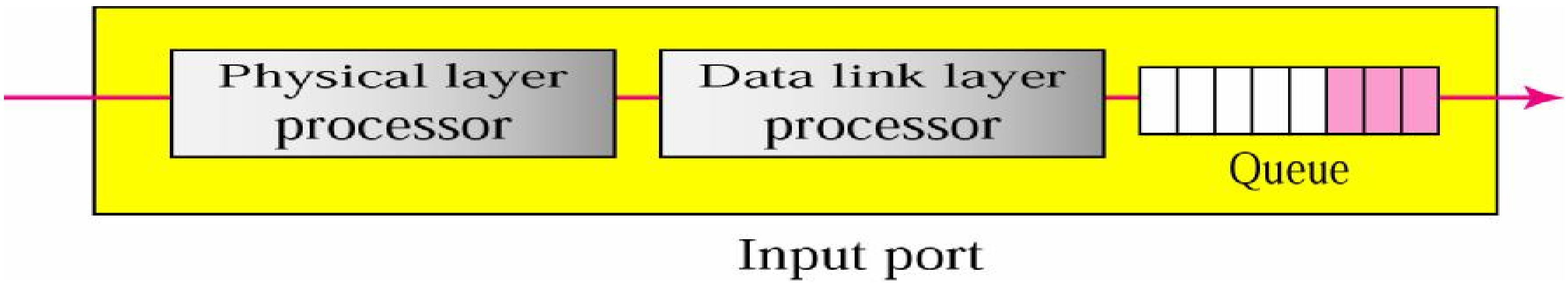
Router Components

- A router has four components
  - Input port
  - Output port
  - Routing processor
  - Switching fabric



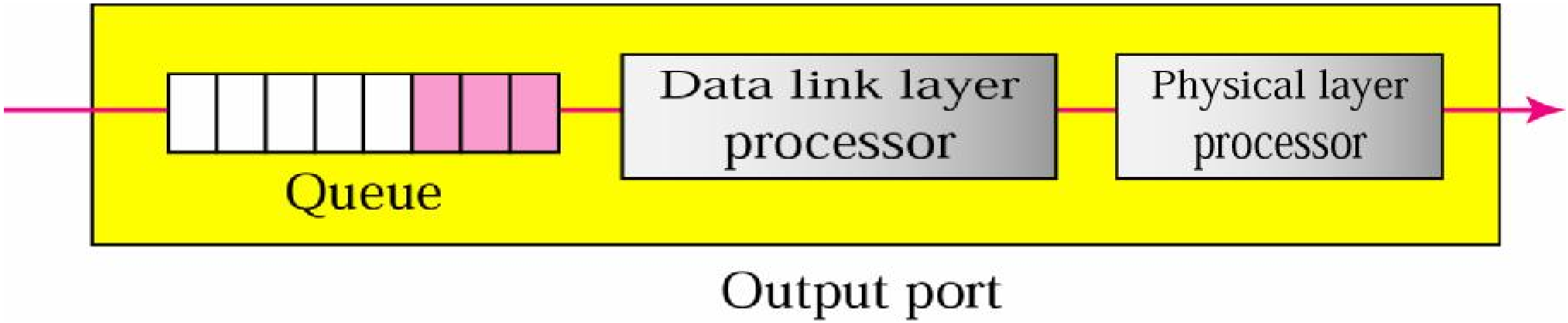
Input Port

- Perform the physical and data link layer functions
  - Bits are constructed from the received signal
  - Packet is decapsulated from the frame
  - Errors are detected and corrected if possible
  - Buffered in queues before packets are directed to switching fabric



Output Ports

- Perform the same functions as the input port, but in the reverse order
  - First, the outgoing packets are queued
  - Then the packet is encapsulated in a frame
  - Finally, the physical and MAC layer functions are applied to send the frame



Routing Processor

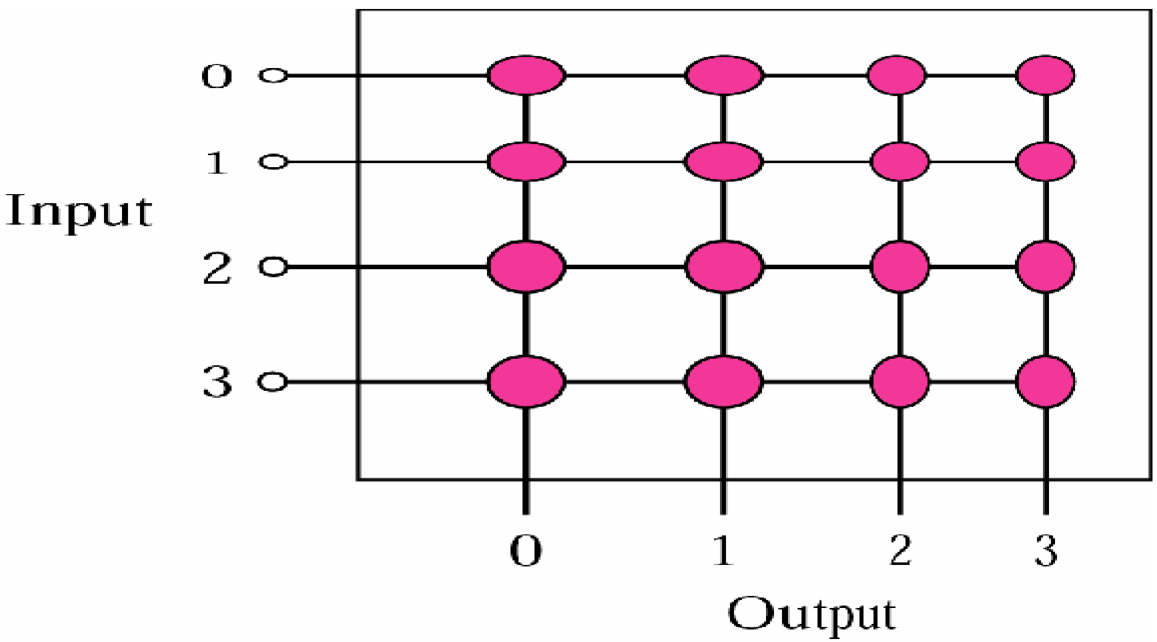
- Perform the functions of the network layer
- Perform *table lookup*
  - Search routing table to find the address of the next hop and the output port number

Switching Fabrics

- Most difficult task in a router
  - Move the packet from the input queue to the output queue
  - The speed affects
    - The size of the input/output queue
    - The overall delay in packet delivery
- Solution: switch fabrics
- Some of the fabrics
  - Crossbar Switch
  - Banyan Switch
  - Batcher-Banyan Switch

Crossbar Switch

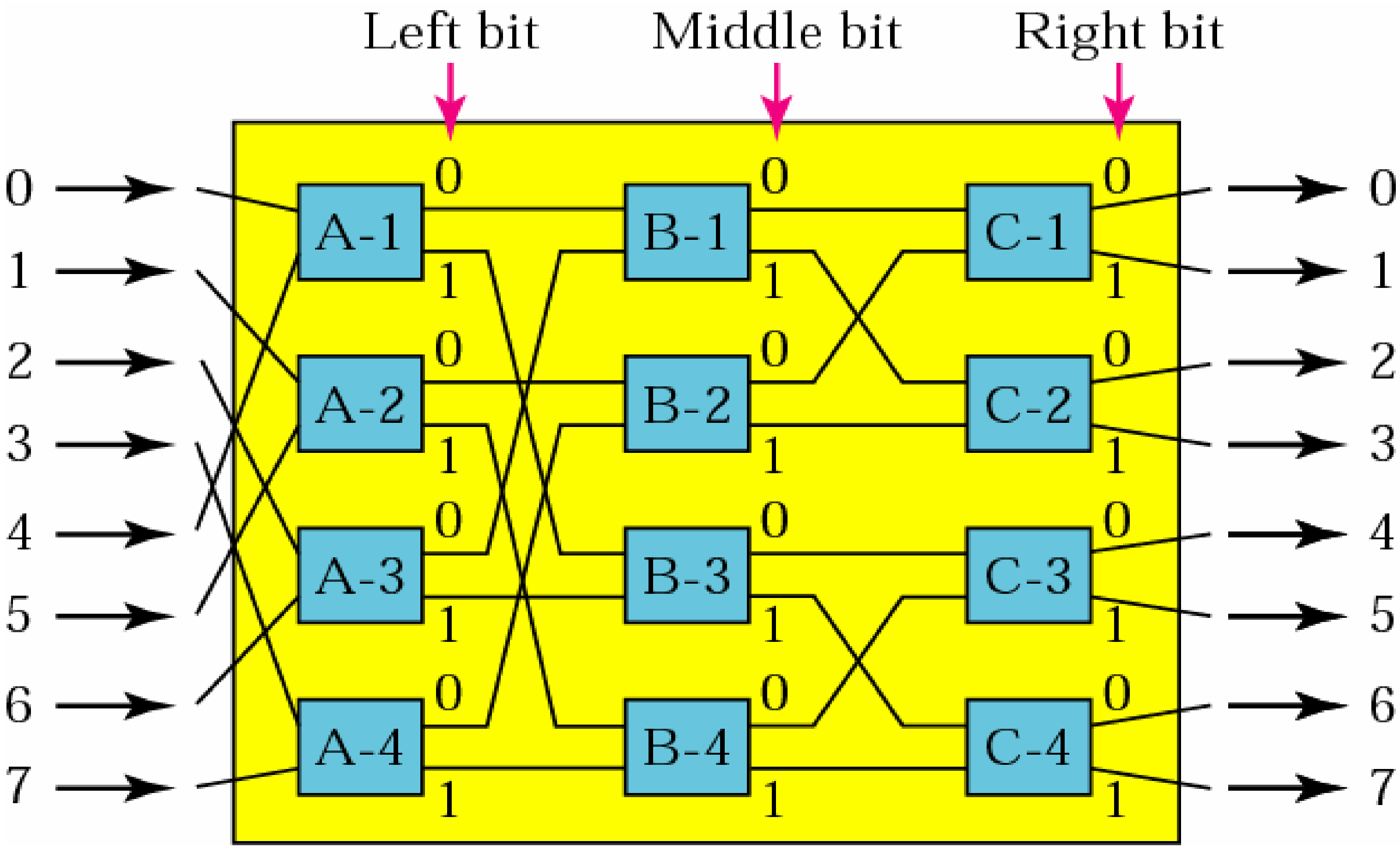
- Connect *n* inputs to *n* output in a grid
- Each ***crosspoint*** has a electronic microswtich



Banyan Switch

- A ***multistage*** switch
  - A *microswitches* at each stage that route packets based on the output port represented as a ***binary string***
  - For *n* input and *n* output, we have $log_2(n)$ stages
    - At each stage, we need *n/2* microswitches
    - See the following figure
      - The number of stages is $\log_2(8) = 3$
  - The first stage routes packets based on the highest bit of the binary string, and so on
  - Elegant design that every input port can connect to any one output port

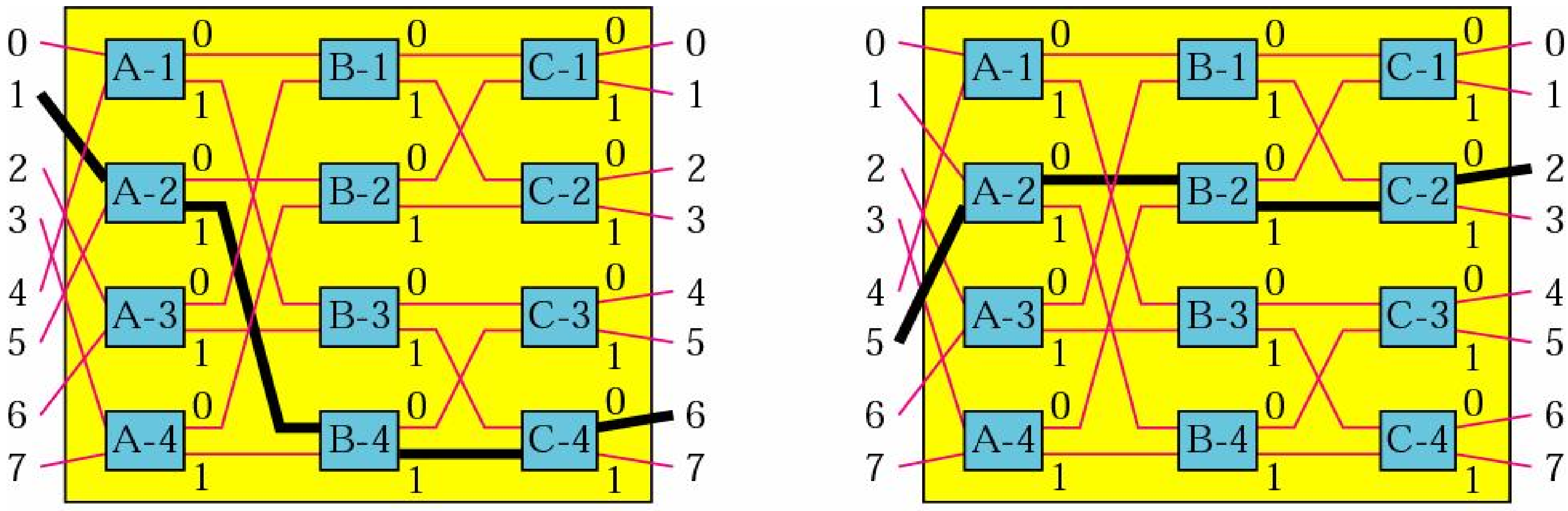
A Banyan Switch



Banyan Switch (Cont.)

- In the next slide
  - A packet has arrived at input port 1 and go to output port 6 (110 in binary)
  - First microswtich (A-2) routes the packet based on the first bit (1)
  - Second microswtich (B-2) routes the packet based on the second bit (1)
  - Third microswtich (C-2) routes the packet based on the third bit (0)

Examples of Routing in a Banyan Switch



a. Input 1 sending to output 6 (110)

b. Input 5 sending to output 2 (010)

Batcher-Banyan Switch

- Problem with the banyan switch
  - The possibility of internal collisions even when two packets are not heading for the same output port
- One solutions: Batcher Switch
  - Sort the incoming packets according to their final destination

Batcher-Banyan Switch (Cont.)

- Batch-Banyan Switch
  - A combination of batch switch and banyan switch
  - A trap is added between batch switch and banyan switch
    - Prevent packets with the same output destination from passing the banyan switch simultaneously
    - Only one packet for each destination is allowed at each tick

Batcher-Banyan Switch