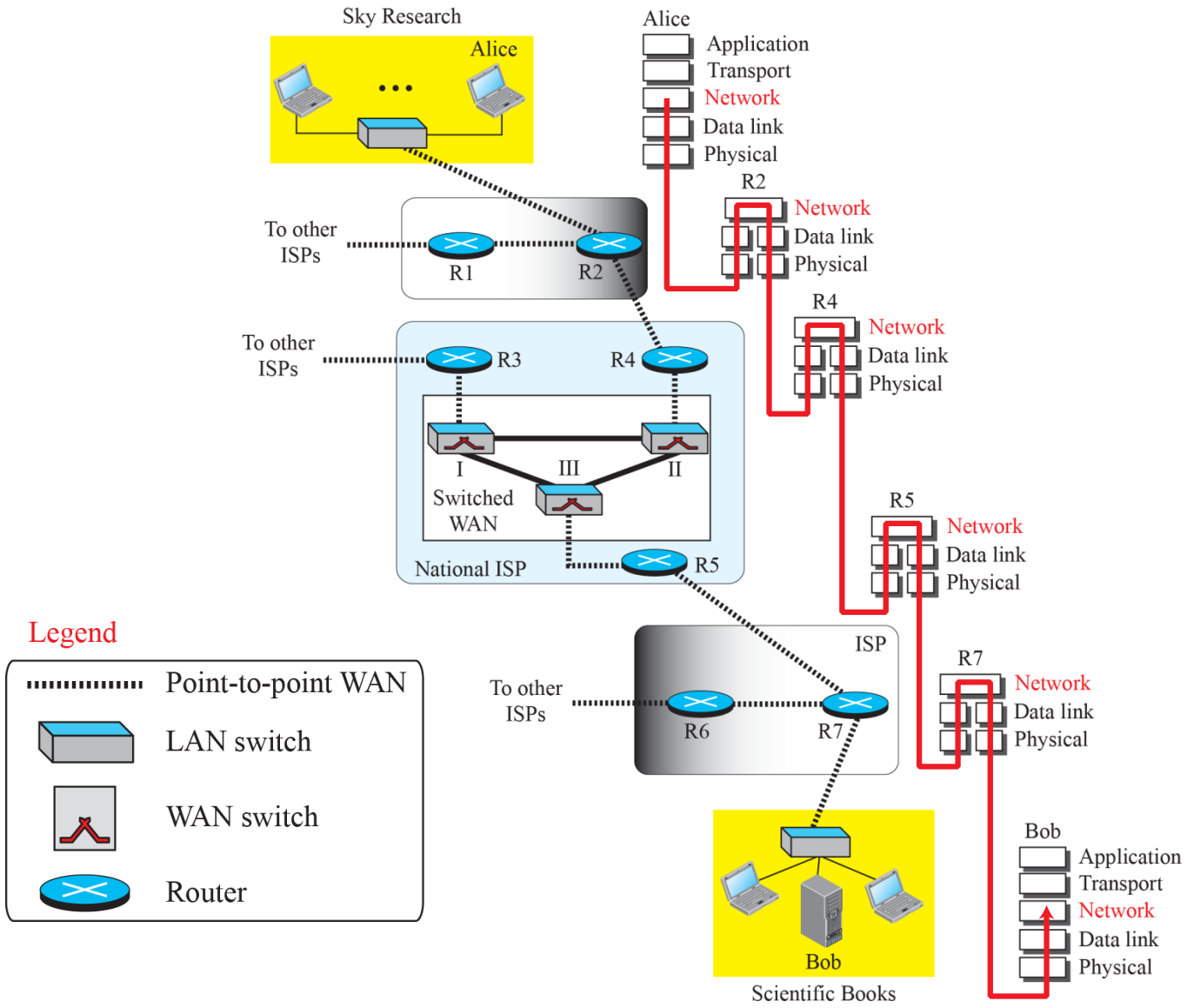
Network Layer

Communication at the network layer



Network-Layer Services

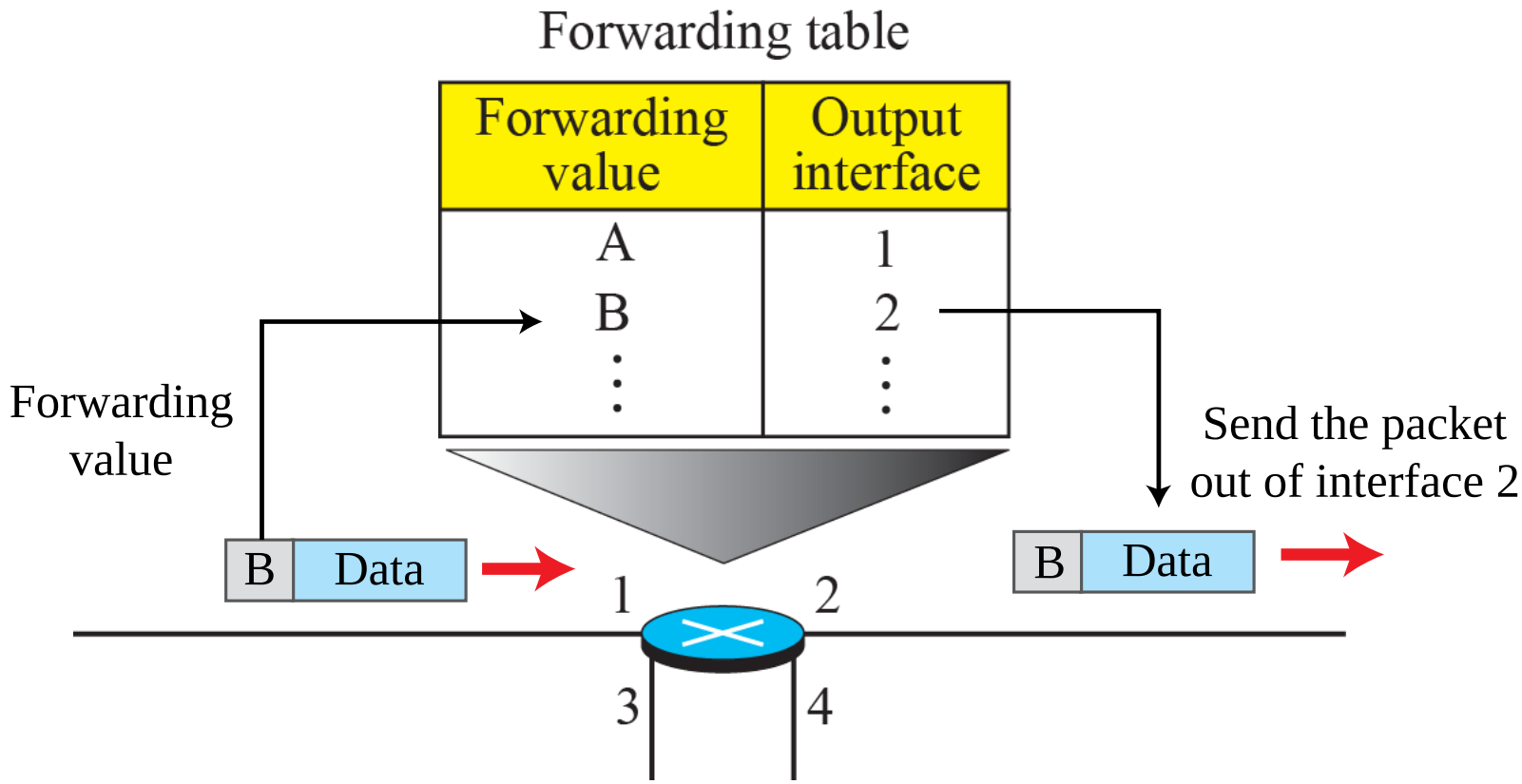
Before discussing the network layer in the Internet today, let's briefly discuss the network-layer services that, in general, are expected from a network-layer protocol.

- Packetizing
- Routing
- Forwarding

(continued)

- Error Control
- Flow Control
- Congestion Control
- Quality of Service
- Security

Forwarding process



Packet Switching

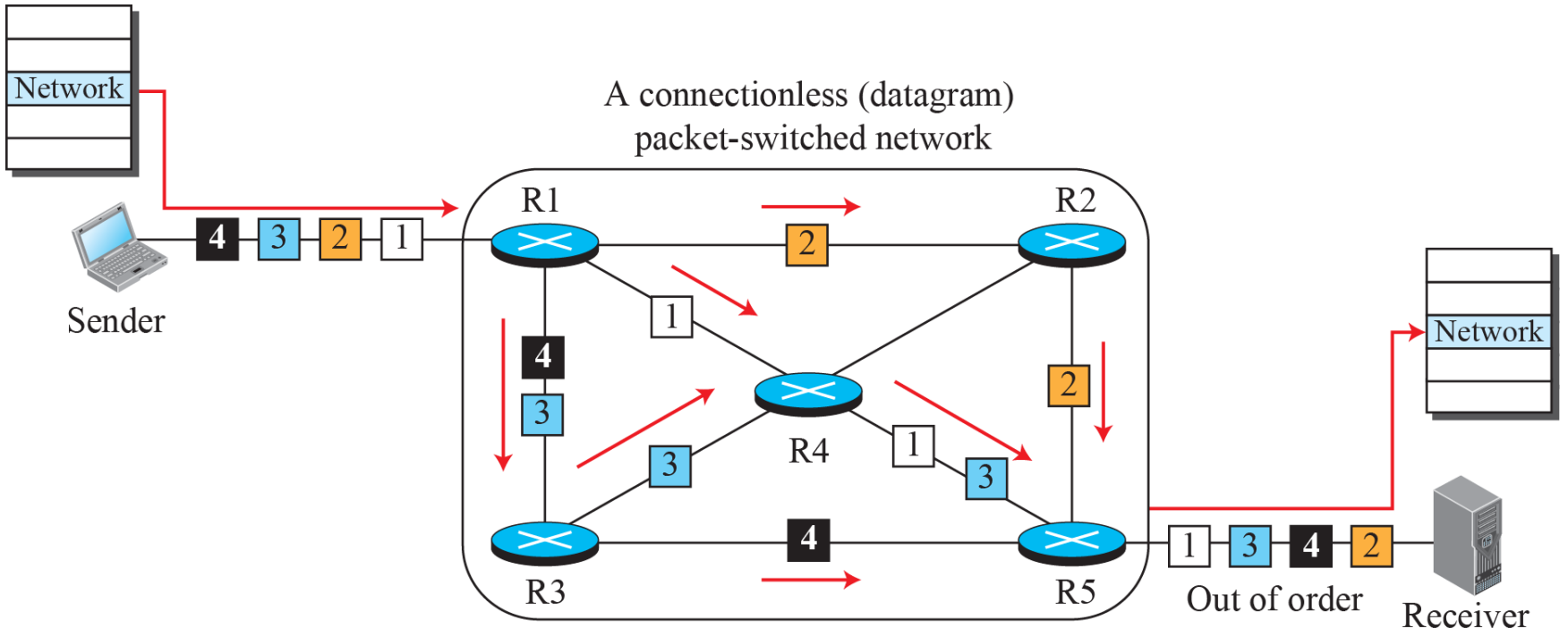
From the discussion of routing and forwarding in the previous section, we infer that a kind of switching occurs at the network layer. A router, in fact, is a switch that creates a connection between an input port and an output port (or a set of output ports), just as an electrical switch connects the input to the output to let electricity flow.

4.1.2 (continued)

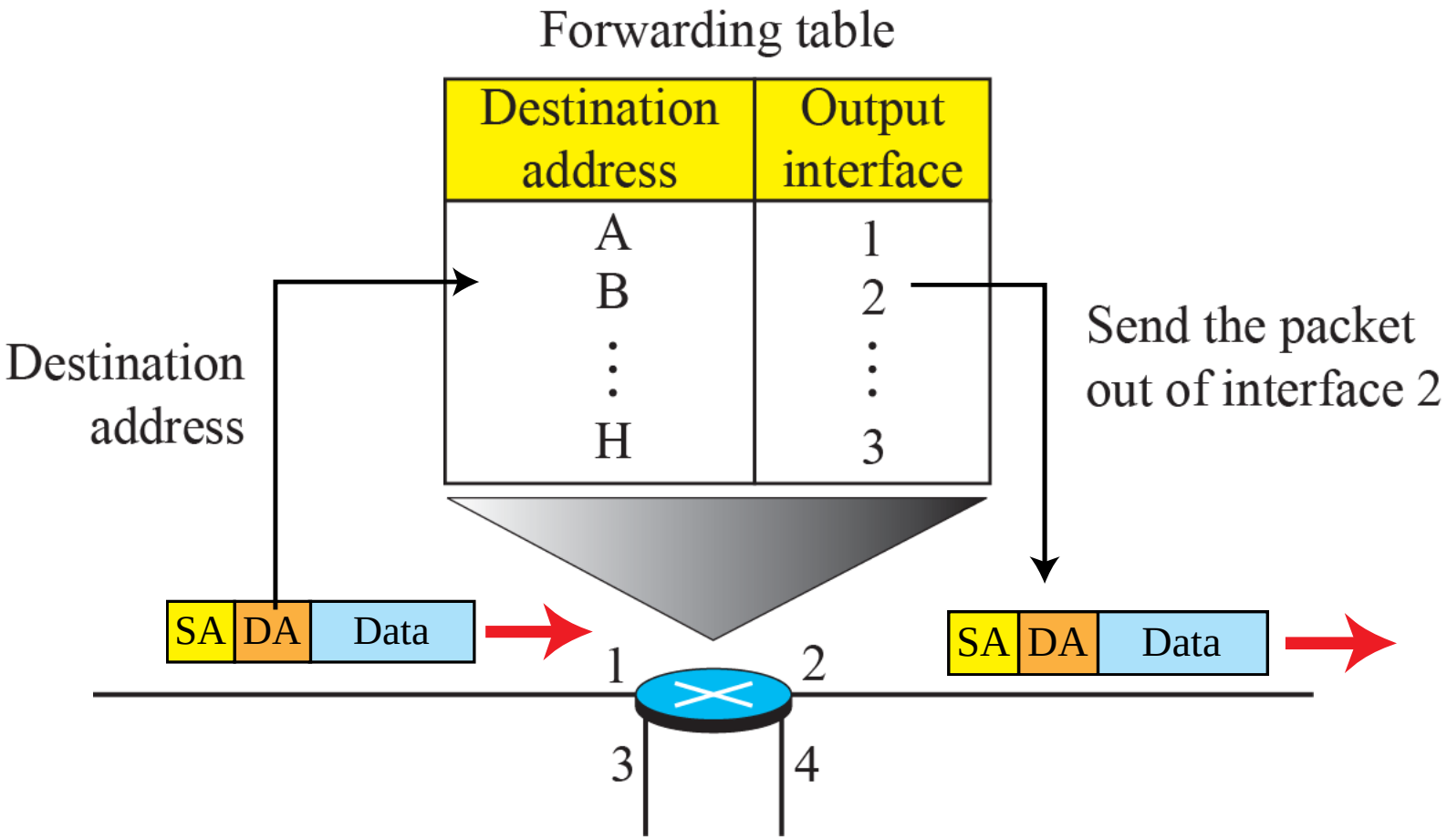
- Datagram Approach

- Virtual-Circuit Approach
  - Setup Phase
  - Data-Transfer Phase
  - Teardown Phase

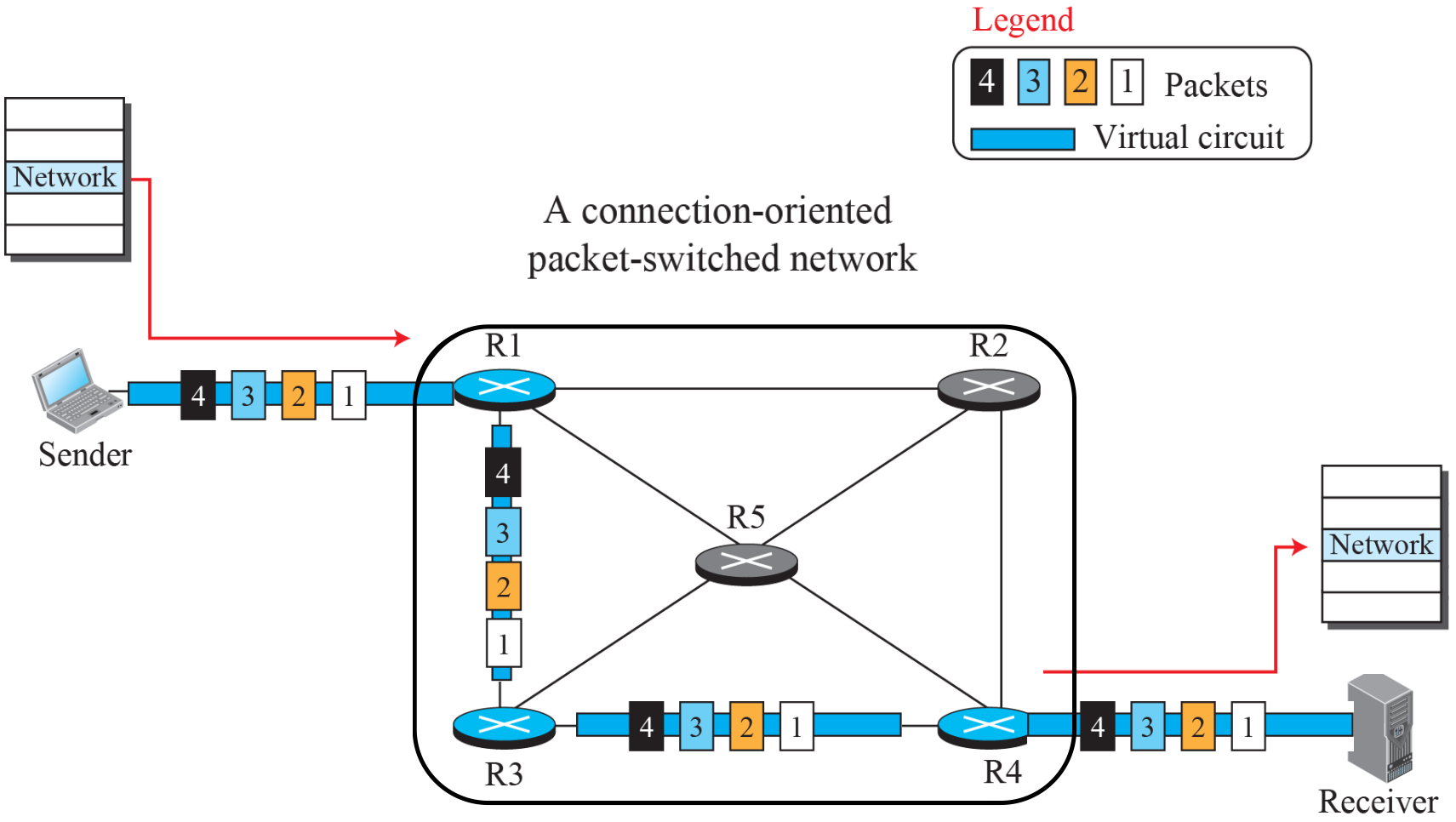
A connectionless packet-switched network



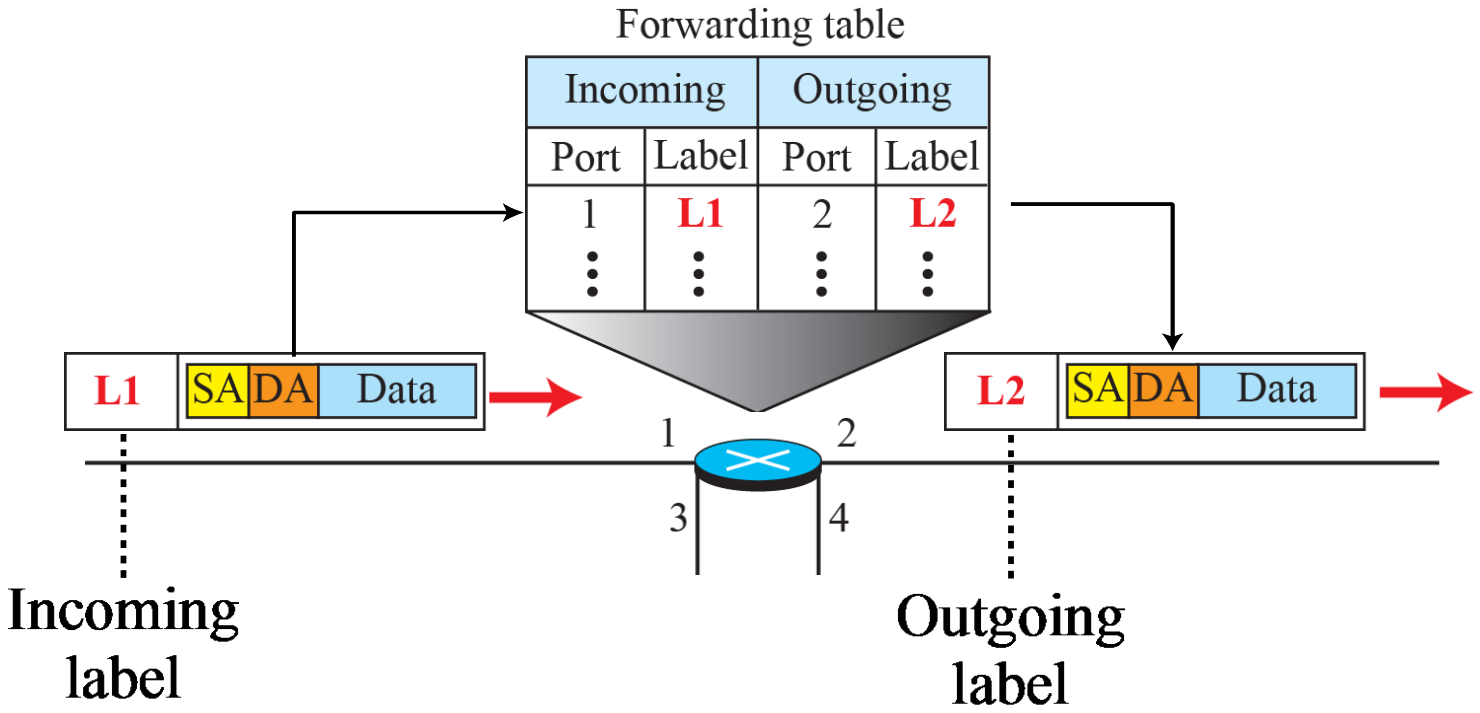
Forwarding process in a router when used in a connectionless network



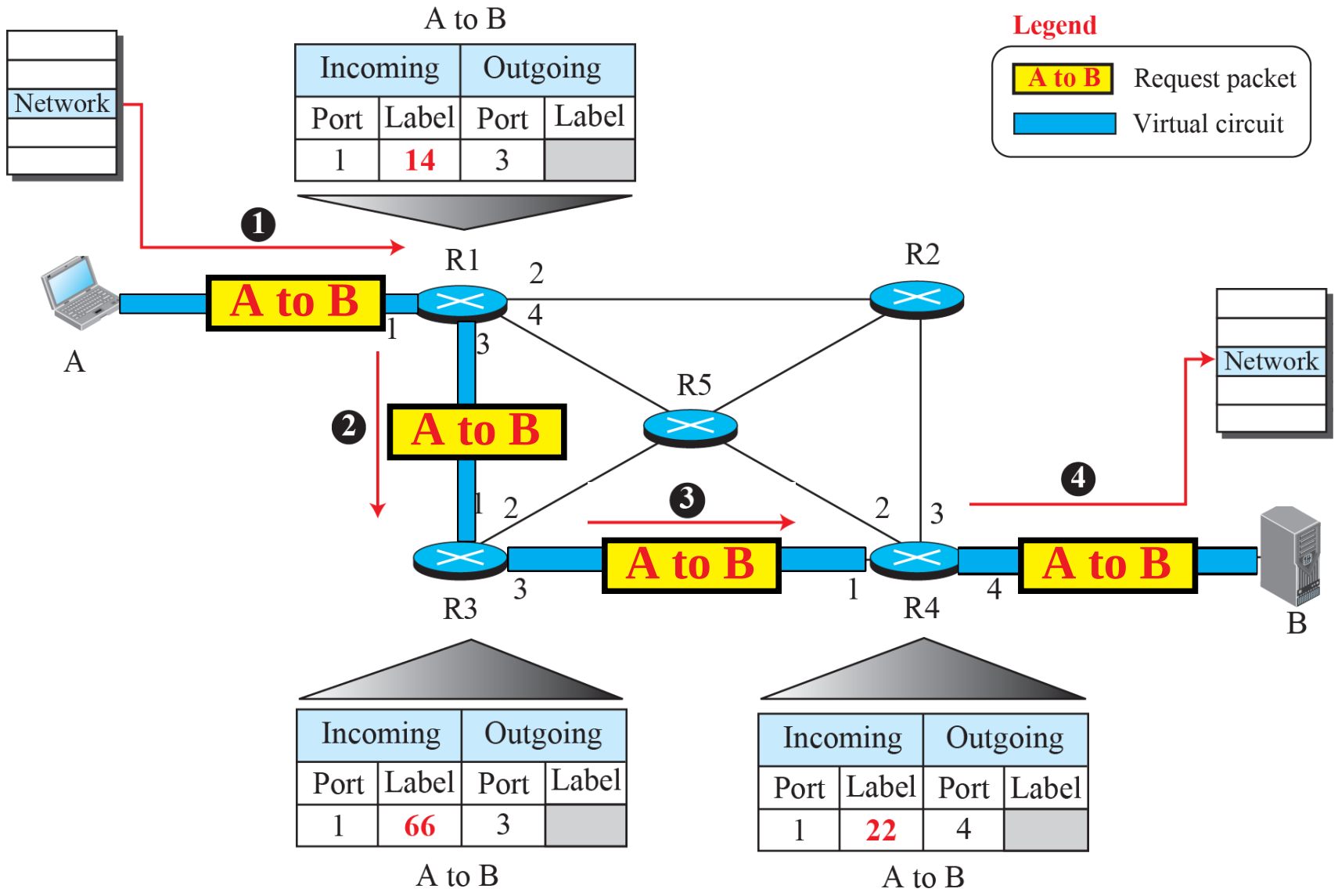
A virtual-circuit packet-switched network



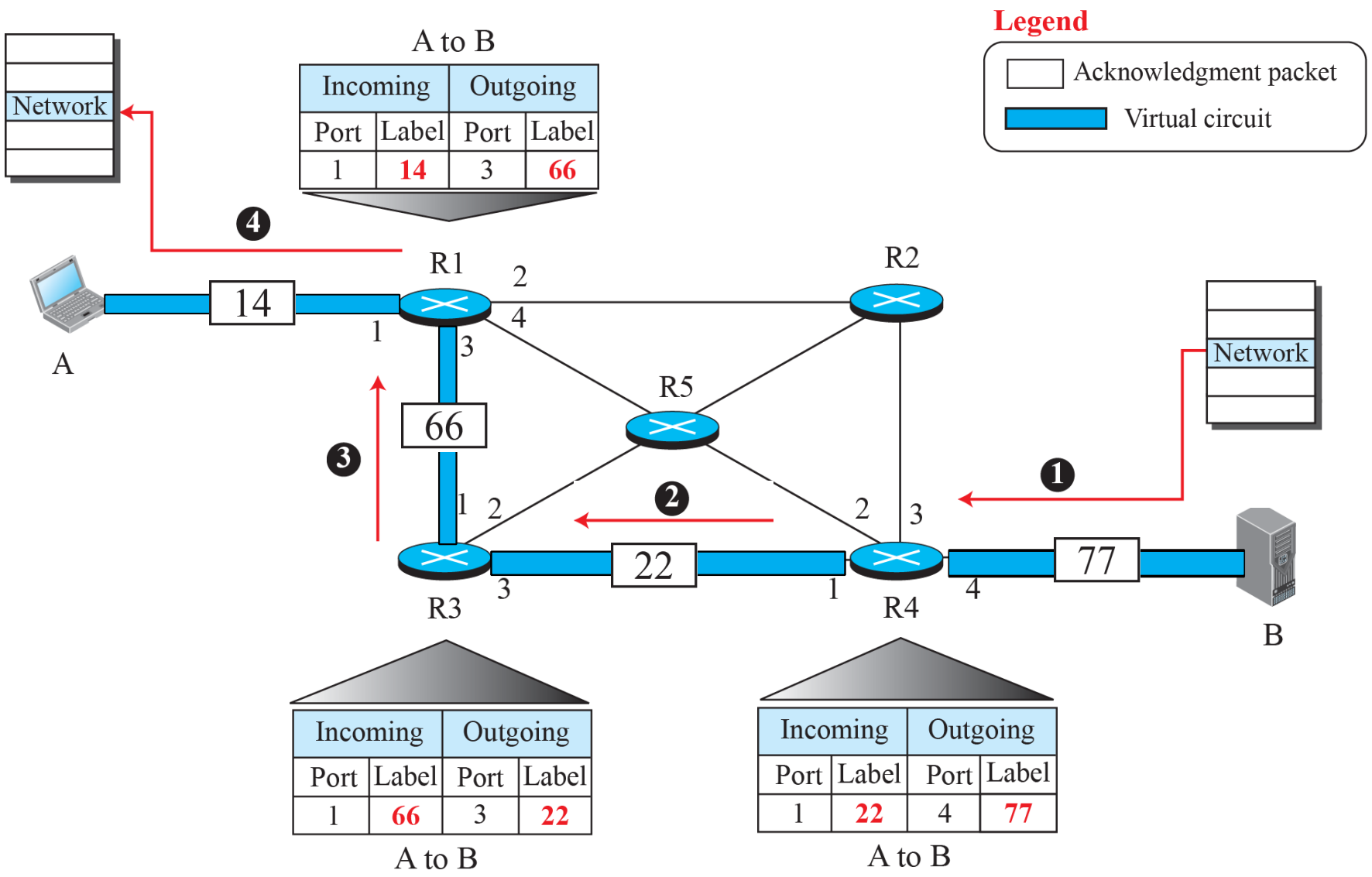
Forwarding process in a router when used in a virtual circuit network



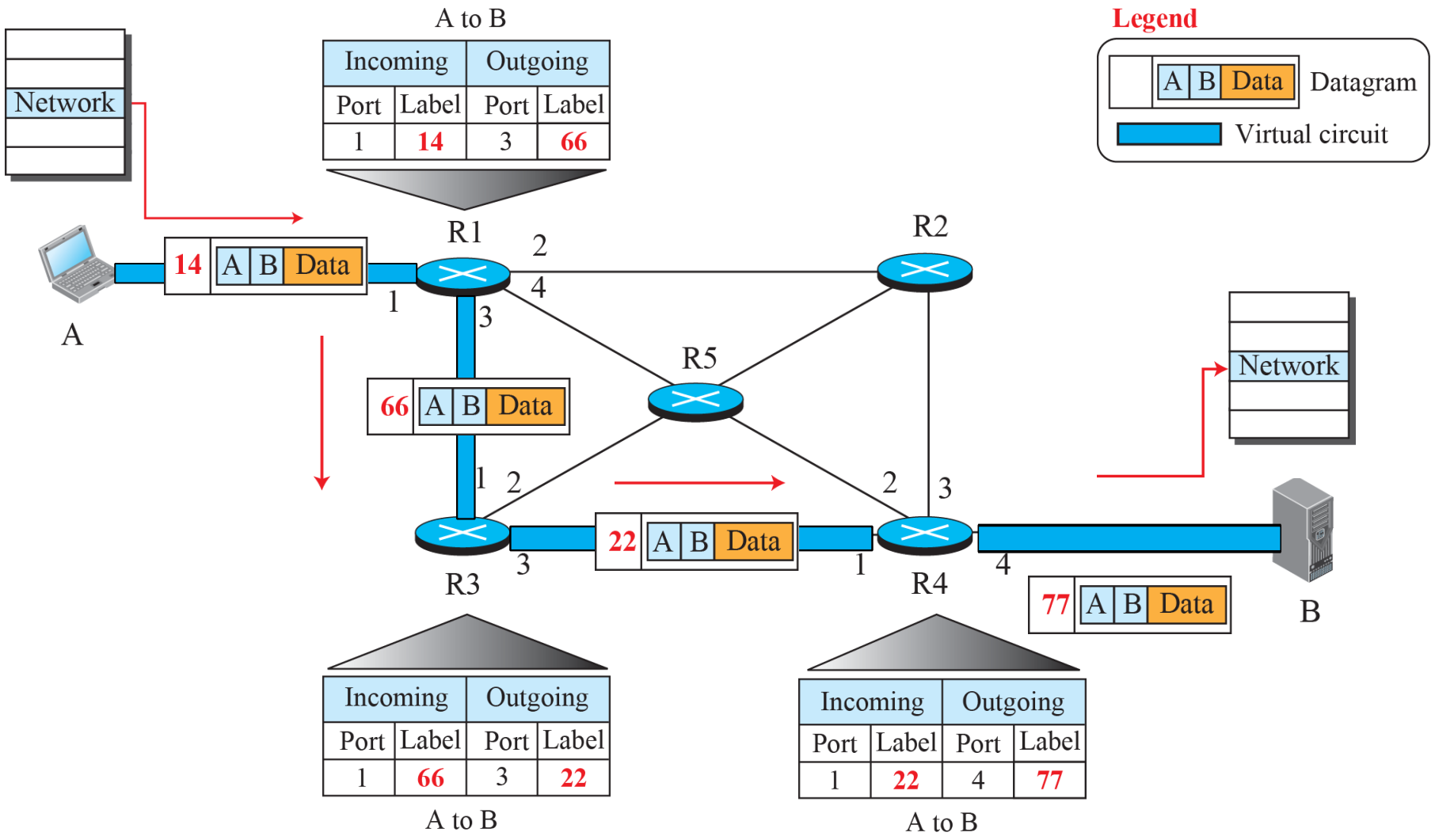
Sending request packet in a virtual-circuit network



Sending acknowledgments in a virtual-circuit network



Sending acknowledgments in a virtual-circuit network



Network-Layer Performance

The upper-layer protocols that use the service of the network layer expect to receive an ideal service, but the network layer is not perfect. The performance of a network can be measured in terms of delay, throughput, and packet loss. We first define these three terms in a packet-switched network before we discuss their effects on performance.

- Delay
- Throughput
- Packet Loss

(continued)

- Delay

- **Transmission Delay**

  $Delay_{tr}$ = (Packet length) / (Transmission rate).

- **Propagation Delay**

  $Delay_{pg}$ = (Distance) / (Propagation speed).

- **Processing Delay**

  $Delay_{pr}$ = Time required to process a packet in a router or a destination host

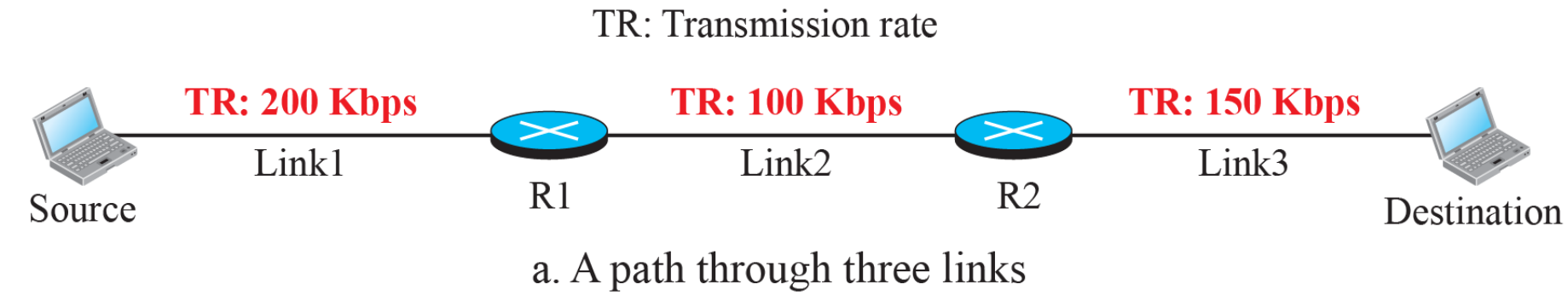
- **Queuing Delay**

  $Delay_{qu}$ = The time a packet waits in input and output queues in a router

- **Total Delay**

  Total delay = $(n + 1)(Delay_{tr} + Delay_{pg} + Delay_{pr}) + (n)(Delay_{qu})$

Throughput in a path with three links in a series

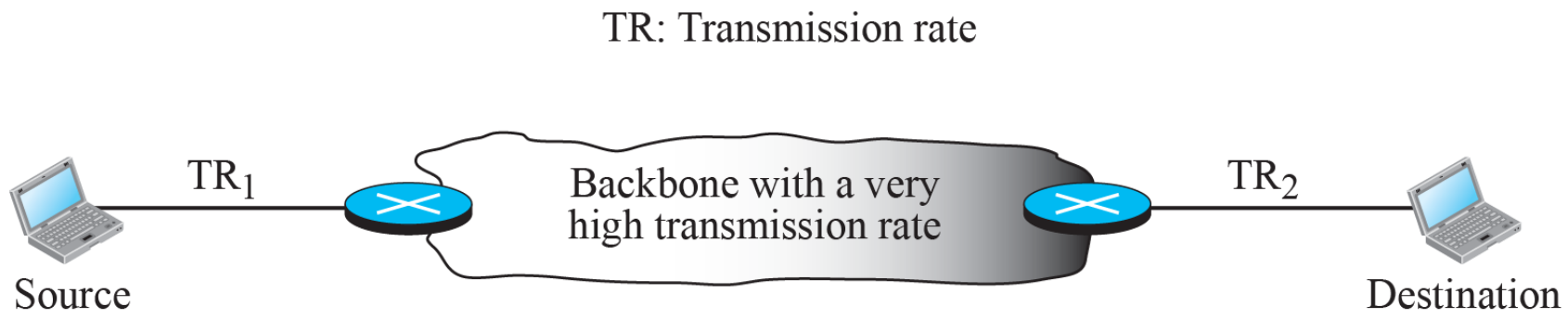


a. A path through three links

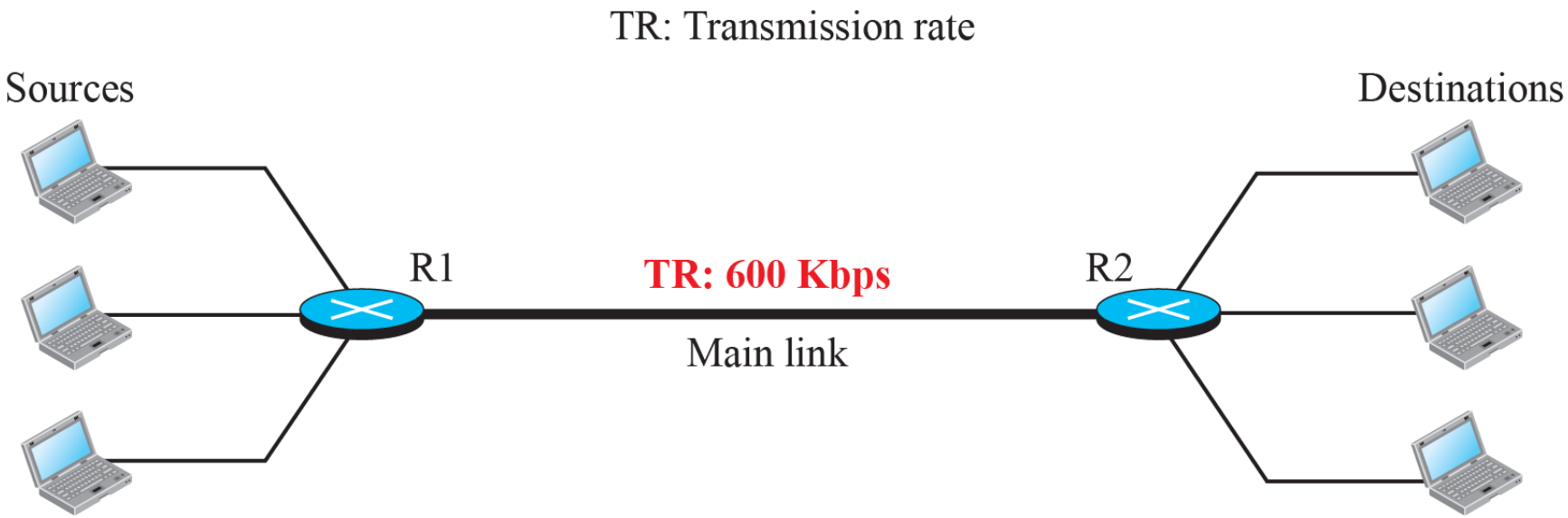
b. Simulation using pipes

Throughput = minimum {TR1, TR2, . . . TRn}.

A path through the Internet backbone



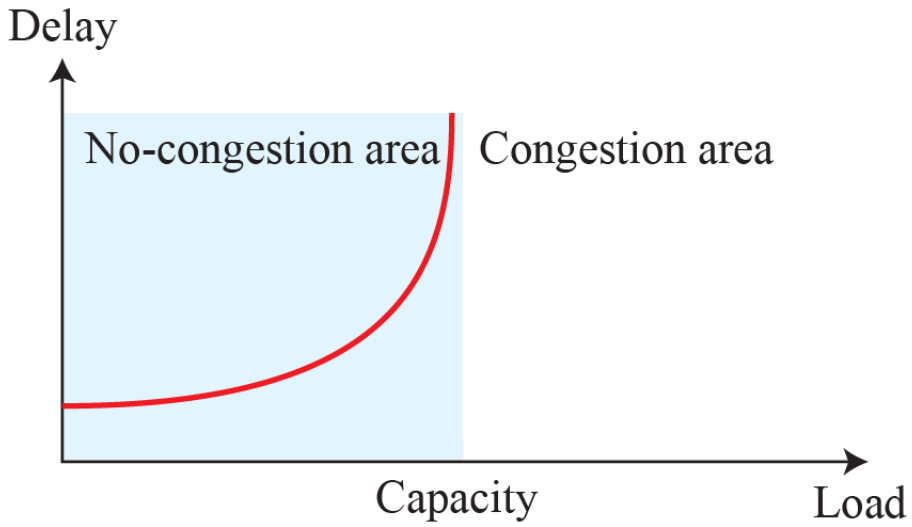
Effect of throughput in shared links



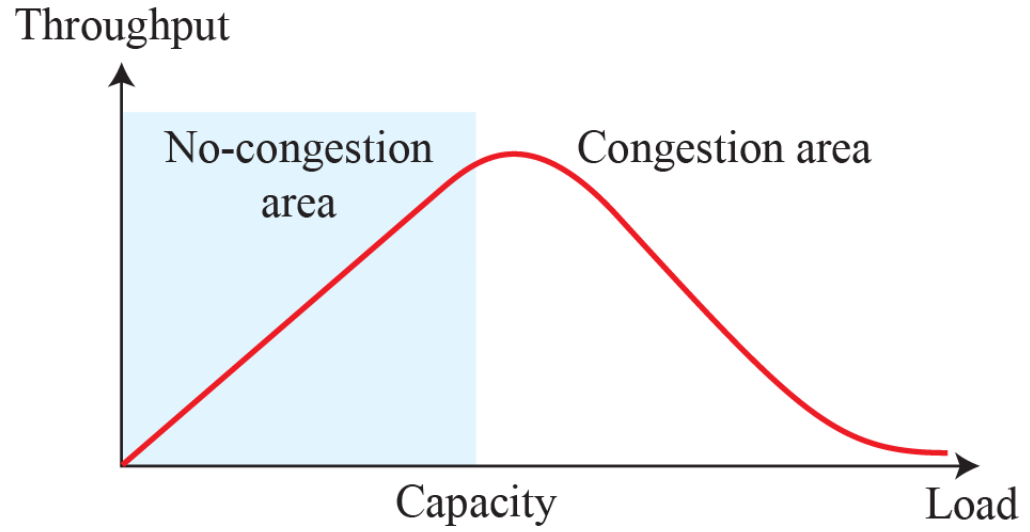
Network-Layer Congestions

Although congestion at the network layer is not explicitly addressed in the Internet model, the study of congestion at this layer may help us to better understand the cause of congestion at the transport layer and find possible remedies to be used at the network layer. Congestion at the network layer is related to two issues, throughput and delay, which we discussed in the previous section.

Packet delay and throughput as functions of load



a. Delay as a function of load



b. Throughput as a function of load

Congestion Control

Open-Loop Congestion Control

- Retransmission Policy
- Window Policy
- Acknowledgment Policy
- Discarding Policy
- Admission Policy

Closed-Loop Congestion Control

- Backpressure
- Choke Packet
- Implicit Signaling
- Explicit Signaling

(continued)

❖ Open-Loop Congestion Control

In open-loop congestion control, policies are applied to prevent congestion before it happens.

- **Retransmission Policy**

The retransmission policy and the retransmission timers must be designed to optimize efficiency and at the same time prevent congestion.

- **Window Policy**

The type of window at the sender may also affect congestion. The Selective Repeat window is better than the Go-Back-N window for congestion control.

- **Acknowledgment Policy**

The acknowledgment policy imposed by the receiver may also affect congestion. If the receiver does not acknowledge every packet it receives, it may slow down the sender and help prevent congestion.

(continued)

❖ Open-Loop Congestion Control

- **Discarding Policy**

A good discarding policy by the routers may prevent congestion and at the same time may not harm the integrity of the transmission.(Wine Policy & Milk Policy)

- **Admission Policy**

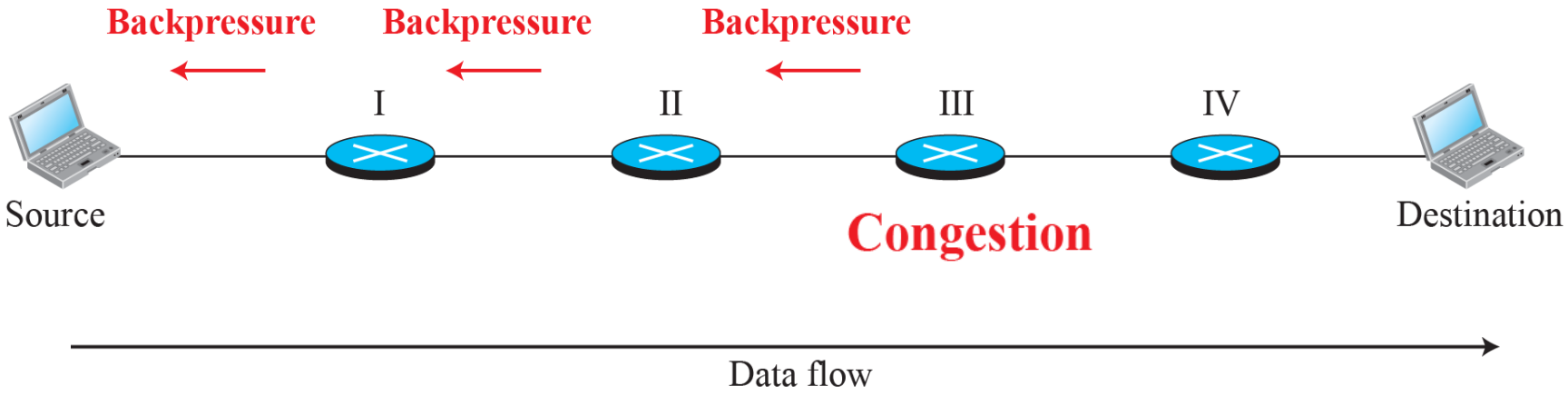
An admission policy, which is a quality-of-service mechanism, can also prevent congestion in virtual-circuit networks. Switches in a flow first check the resource requirement of a flow before admitting it to the network.

(continued)

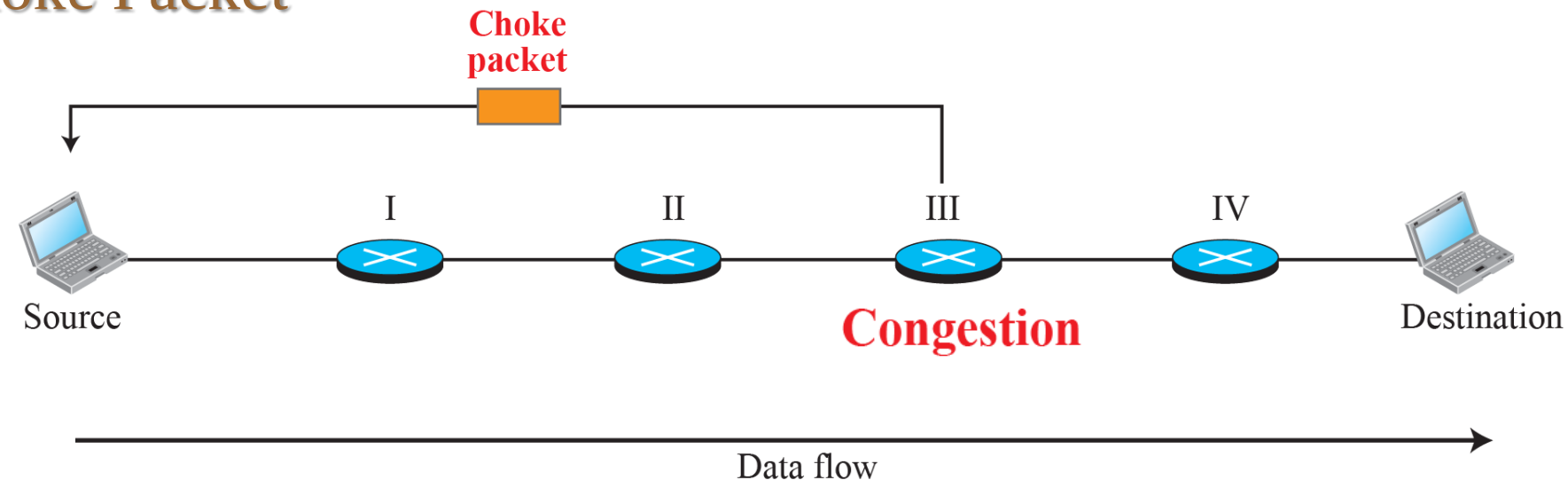
- Congestion Control

Closed-Loop Congestion Control

Backpressure



- Choke Packet



- Implicit Signaling

The source guesses that there is congestion somewhere in the network from other symptoms.

- Explicit Signaling

in the explicit-signaling method, the signal is included in the packets that carry data. Explicit signaling can occur in either the forward or the backward direction.

IPv4 Addresses

The identifier used in the IP layer of the TCP/IP protocol suite to identify the connection of each device to the Internet is called the Internet address or IP address. An IPv4 address is a 32-bit address that uniquely and universally defines the connection of a host or a router to the Internet. The IP address is the address of the connection, not the host or the router, because if the device is moved to another network, the IP address may be changed.

(continued)

- Address Space
- Notation
- Hierarchy in Addressing
- Classful Addressing
  - Address Depletion
  - Disadvantage of Classful Addressing

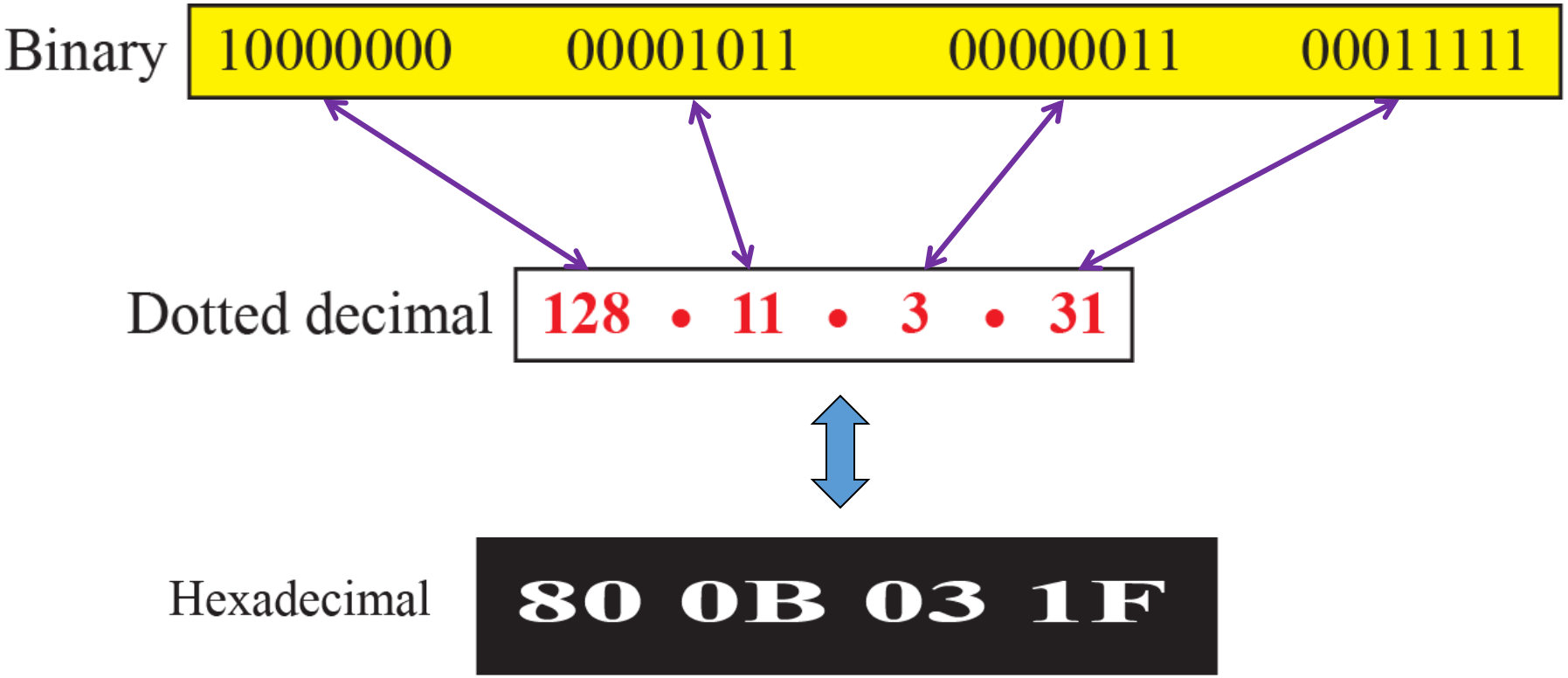
(continued)

- Classless Addressing
  - Prefix Length: Slash Notation
  - Extracting information from an address
  - Address Mask
  - Network Address
  - Block Allocation
  - Subnetting
  - Address Aggregation
  - Special Addresses

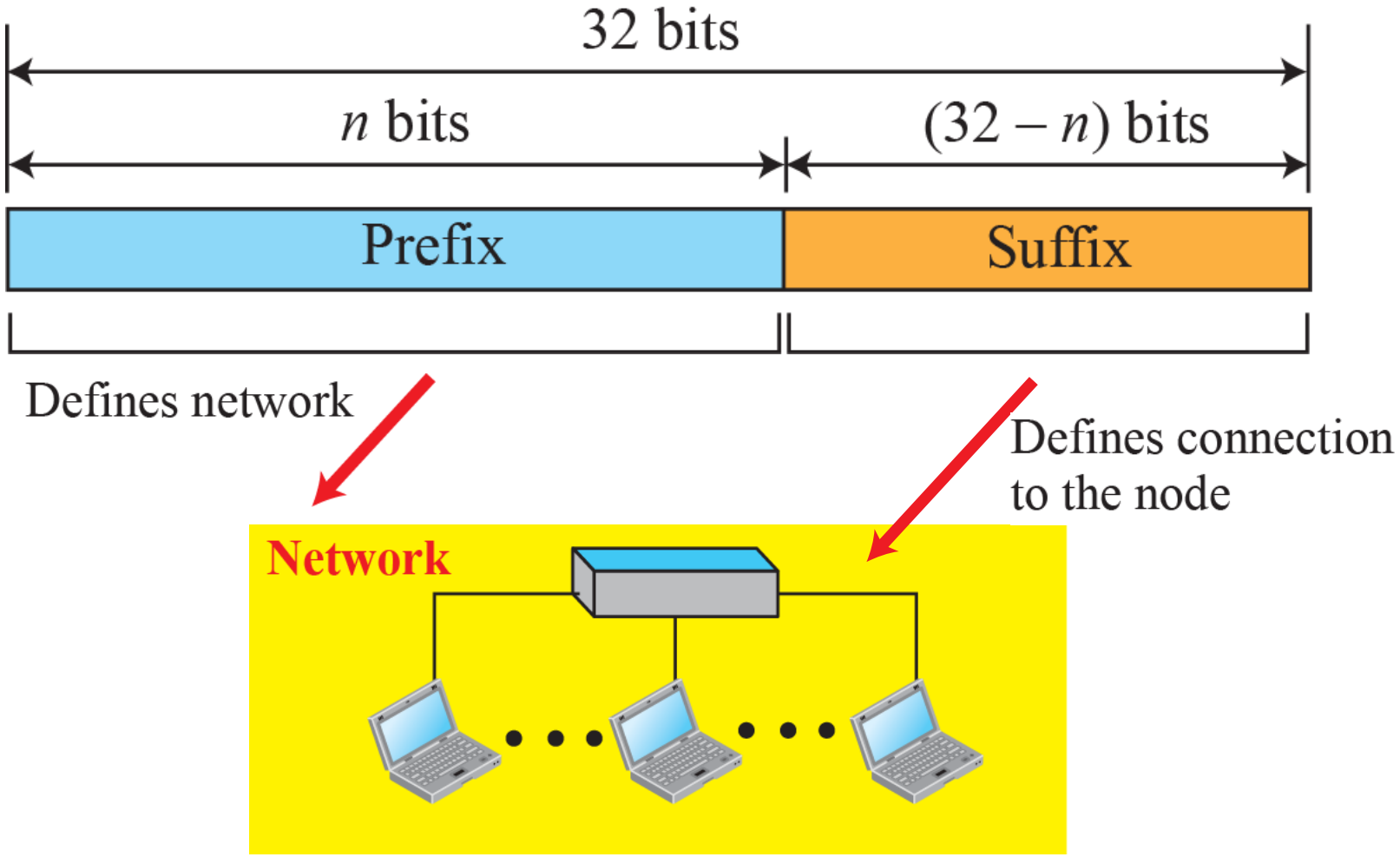
(continued)

- Dynamic Host Configuration Protocol (DHCP)
  - DHCP Message Format
  - DHCP Operation
  - Two Well-Known Ports
  - Using FTP
  - Error Control
  - Transition States
- NAT
  - Address Translation
  - Translation Table

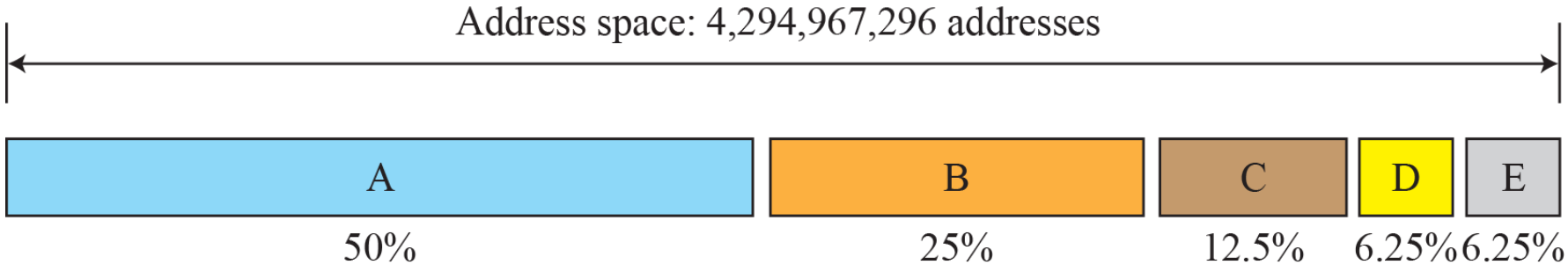
Three different notations in IPv4 addressing

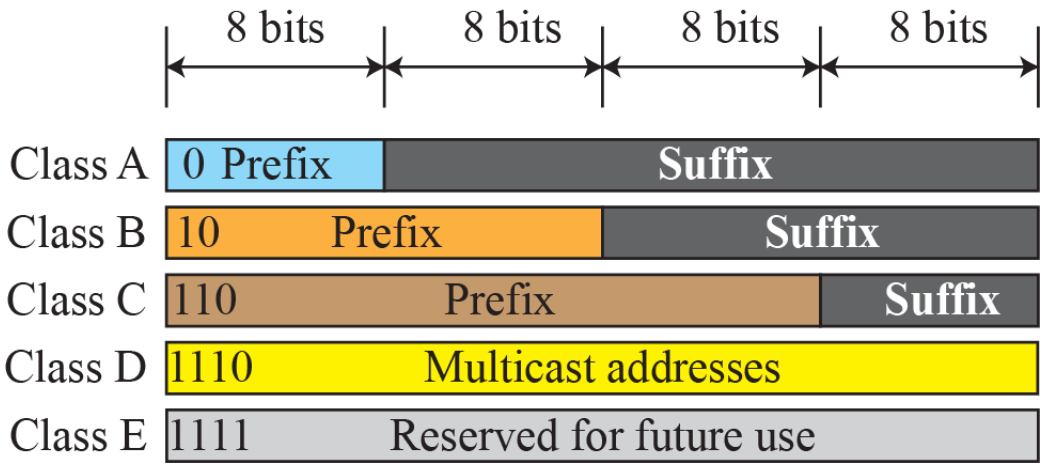


Hierarchy in addressing



Occupation of the address space in classful addressing





Class	Prefixes	First byte
A	n = 8 bits	0 to 127
B	n = 16 bits	128 to 191
C	n = 24 bits	192 to 223
D	Not applicable	224 to 239
E	Not applicable	240 to 255

Class A: 1.0.0.0 to 127.255.255.255	00000001.00000000.00000000.00000000 01111111.11111111.11111111.11111111
Class B: 128.0.0.0 to 191.255.255.255	10000000.00000000.00000000.00000000 10111111.11111111.11111111.11111111
Class C: 192.0.0.0 to 223.255.255.255	11000000.00000000.00000000.00000000 11011111.11111111.11111111.11111111
Class D: 224.0.0.0 to 239.255.255.255	11100000.00000000.00000000.00000000 11101111.11111111.11111111.11111111
Class E: 240.0.0.0 to 255.255.255.255	11110000.00000000.00000000.00000000 11111111.11111111.11111111.11111111

Class A
127 networks (2^7-2)
and 16777214 hosts
(2^{24}-2).

1.0.0.0	00000001.00000000.00000000.00000000
1.0.0.1	00000001.00000000.00000000.00000001
1.0.0.2	00000001.00000000.00000000.00000010
	.
1.255.255.255	00000001 .11111111.11111111.11111111
2.0.0.0	00000010.00000000.00000000.00000000
2.0.0.1	00000010.00000000.00000000.00000001
2.0.0.2	00000010.00000000.00000000.00000010
2.255.255.255	00000010. 11111111.11111111.11111111
3.0.0.0	00000011.00000000.00000000.00000000
3.0.0.1	00000011.00000000.00000000.00000001
3.0.0.2	00000011.00000000.00000000.00000010
3.255.255.255	00000011. 11111111.11111111.11111111

Class A has 127networks (2^7-1) and 1,67,77,214 hosts (2^{24}-2).

Class B has 16384 (2^{14}) Network addresses and 65,534 (2^{16}-2) Host addresses.

Class C gives 2097152 (2^{21}) Network addresses and 254 (2^8-2) Host addresses.

170.50.0.0-class B-network address

170.50.255.255-Broadcast address

170.50. 0.1-First Host address

170.50.255.254-Last Host Address

65,534 hosts

Subnet Mask

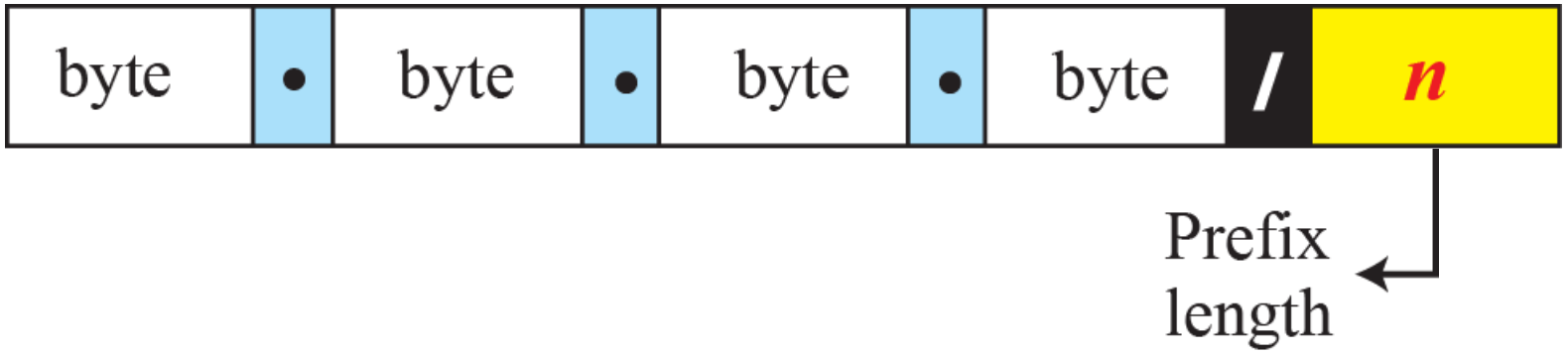
A **subnet mask** is a 32-bit number created by setting host bits to all 0s and setting **network** bits to all 1s. In this way, the **subnet mask** separates the IP address into the **network** and host addresses.

Class A: 255.0.0.0 11111111.00000000.00000000.00000000

Class B: 255.255.0.0 11111111.11111111.00000000.00000000

Class C: 255.255.255.0 11111111.11111111.11111111.00000000

Slash notation (CIDR)



Class A – Default ->/8

Class B – Default ->/16

Class C – Default ->/24

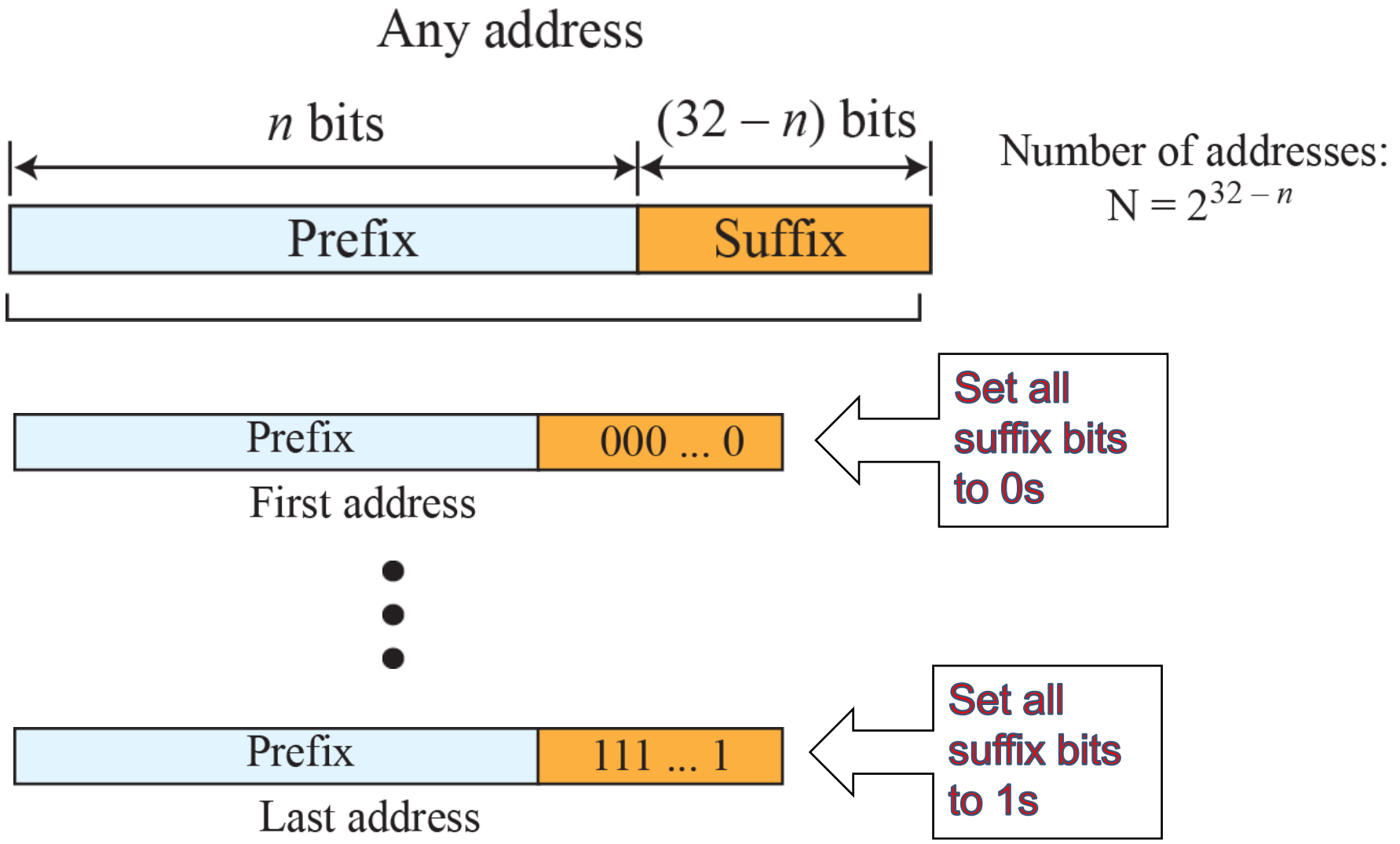
Examples:

12.24.76.8/**8**

23.14.67.92/**12**

220.8.24.255/**25**

Figure 4.34: ***Information extraction in classless addressing***



Network Address	192.168.1.0	11000000.10101000.00000001.00000000
First Host Address	192.168.1.1	11000000.10101000.00000001.00000001
	192.168.1.2	11000000.10101000.00000001.00000010
	192.168.1.3	11000000.10101000.00000001.00000011
	192.168.1.127	11000000.10101000.00000001.01111111
	192.168.1.128	11000000.10101000.00000001.10000000
	192.168.1.129	11000000.10101000.00000001.10000001
	192.168.1.130	11000000.10101000.00000001.10000010
	192.168.1.131	11000000.10101000.00000001.10000011
Last Host Address	192.168.1.254	11000000.10101000.00000001.11111110
Broadcast Address	192.168.1.255	11000000.10101000.00000001.11111111

Network

Host

Network Address	192.168.1.0	11000000.10101000.00000001.00000000
First Host Address	192.168.1.1	11000000.10101000.00000001.00000001
	192.168.1.2	11000000.10101000.00000001.00000010
	192.168.1.3	11000000.10101000.00000001.00000011
Last Host Address		
Broad cast Address	192.168.1.127	11000000.10101000.00000001.01111111
Network address	192.168.1.128	11000000.10101000.00000001.10000000
First Host address	192.168.1.129	11000000.10101000.00000001.10000001
	192.168.1.130	11000000.10101000.00000001.10000010
	192.168.1.131	11000000.10101000.00000001.10000011
Last Host Address	192.168.1.254	11000000.10101000.00000001.11111110
Broadcast Address	192.168.1.255	11000000.10101000.00000001.11111111

Sub Network 1

Sub Network 2

Network

Host

Prefix->/25

IP: 170.50.246.0/18-host

SM: 255.255.192.0

NA: 170.50.192.0

FHA: 170.10.192.1

BA: 170.50.255.255

LHA: 170.50.255.254

10101010.00110010.11110110.00000000

11111111.11111111.11000000.00000000

10101010.00110010.11000000.00000000
00000000.00000000.00111111.11111111

10101010.00110010.11111111.11111111

10.0.0.0	00001010.00000000.00000000.00000000
10.63.255.255	00001010.00111111.11111111.11111111
10.64.0.0	00001010.01000000.00000000.00000000
10.127.255.255	00001010.01111111.11111111.11111111
10.128.0.0	00001010.10000000.00000000.00000000
10.191.255.255	00001010.10111111.11111111.11111111
10.192.0.0	00001010.11000000.00000000.00000000
10.255.255.255	00001010.11111111.11111111.11111111

IP Address (and) Subnet mask=Network address

IP: 130.45.34.36/20 SM: 255.255.240.0

10000010.00101101.00100010.00100100 (ip address)

AND

11111111.11111111.11110000.00000000 (SM) =

10000010.00101101.00100000.00000000 =

130.45.32.0 (the resulting network address)

Network address (OR) Subnet Mask inverse= BA

10000010.00101101.00100000.00000000 (netadress)

OR

00000000.00000000.00001111.11111111 (inverted SM) =

10000010.00101101.00101111.11111111 =

130.45.47.255 (broadcast address)

Example 4.2

Number of addresses in the block:	N = NOT (mask) + 1 = 0.0.0.31 + 1 = 32 addresses
First address:	First = (address) AND (mask) = 167.199.170. 82
Last address:	Last = (address) OR (NOT mask) = 167.199.170. 255

Special Addresses

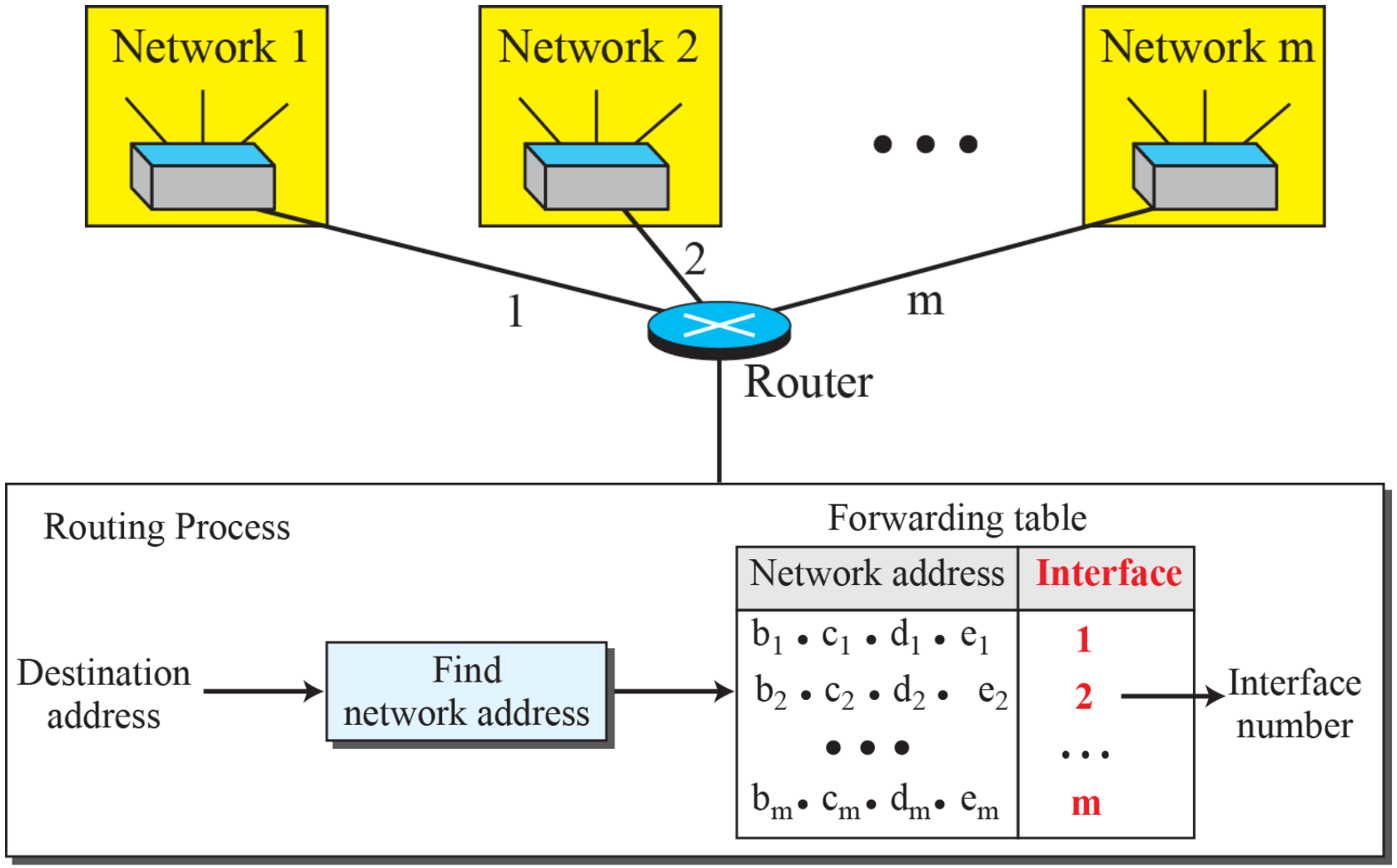
This-host Address: The only address in the block 0.0.0.0/32 is called the this-host address. It is used whenever a host needs to send an IP datagram but it does not know its own address to use as the source address.

Limited-broadcast Address: The only address in the block 255.255.255.255/32 is called the limited-broadcast address. It is used whenever a router or a host needs to send a datagram to all devices in a network.

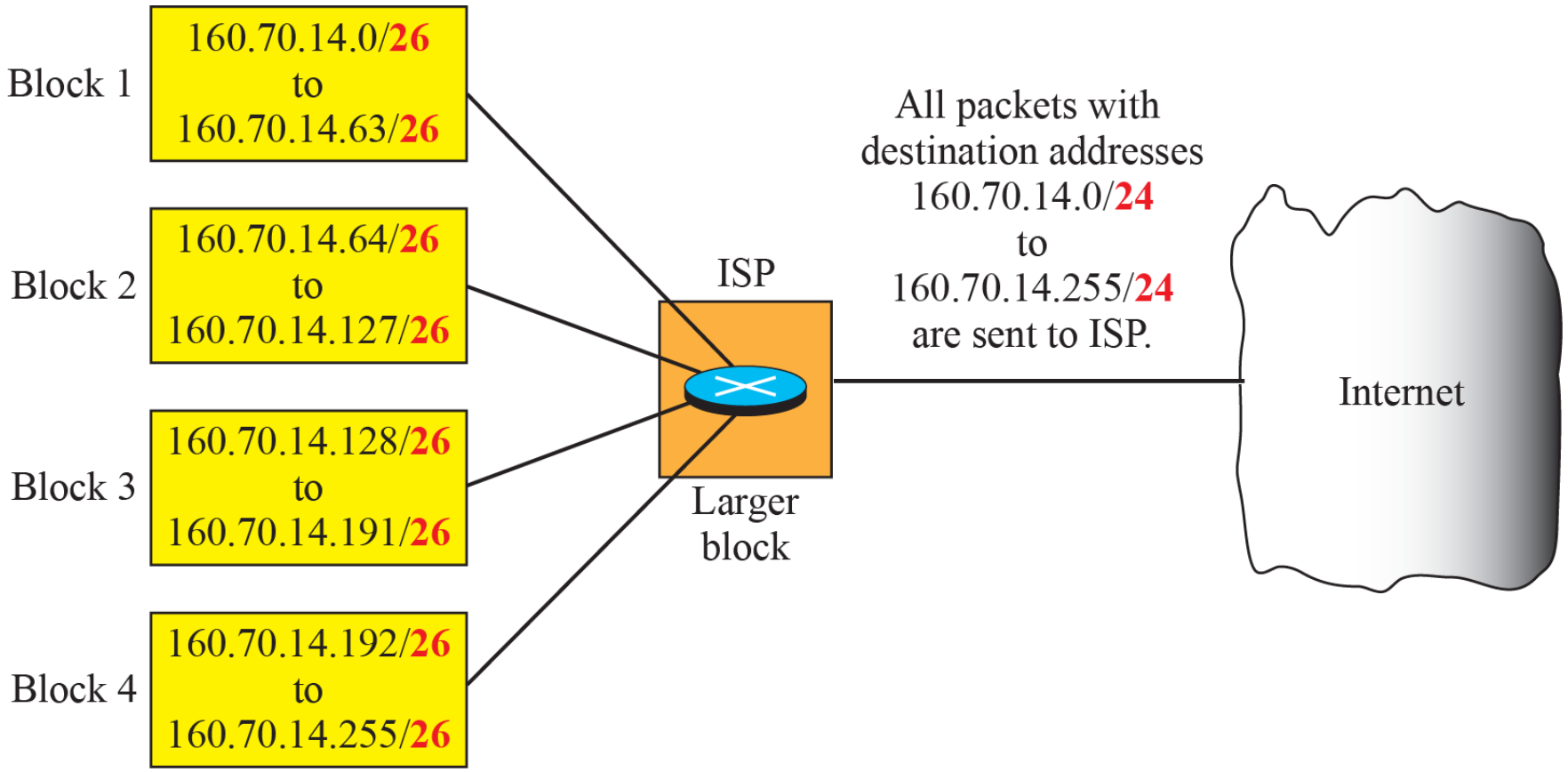
Loopback Address: The block 127.0.0.0/8 is called the loopback address. A packet with one of the addresses in this block as the destination address never leaves the host; it will remain in the host. Any address in the block is used to test a piece of software in the machine. For example, we can write a client and a server program in which one of the addresses in the block is used as the server address.

Private addresses: Four blocks are assigned as private addresses: 10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16, and 169.254.0.0/16.

Network address



Example of address aggregation



DHCP (Dynamic Host configuration Protocol)

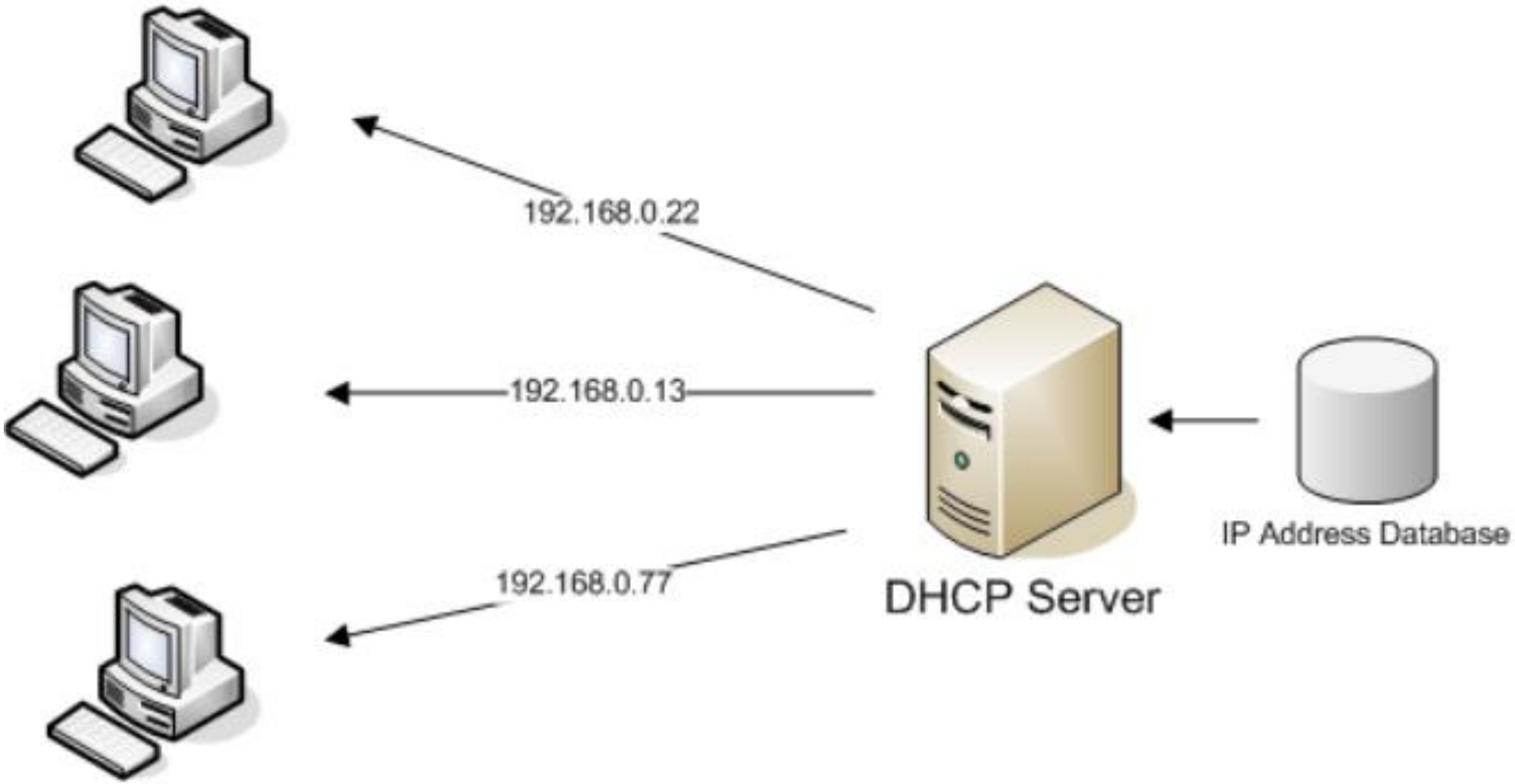
- DHCP is an application-layer program, using the client-server paradigm, that actually helps TCP/IP at the network layer.

- DHCP has found such widespread use in the Internet that it is often called a *plug and-play protocol*.

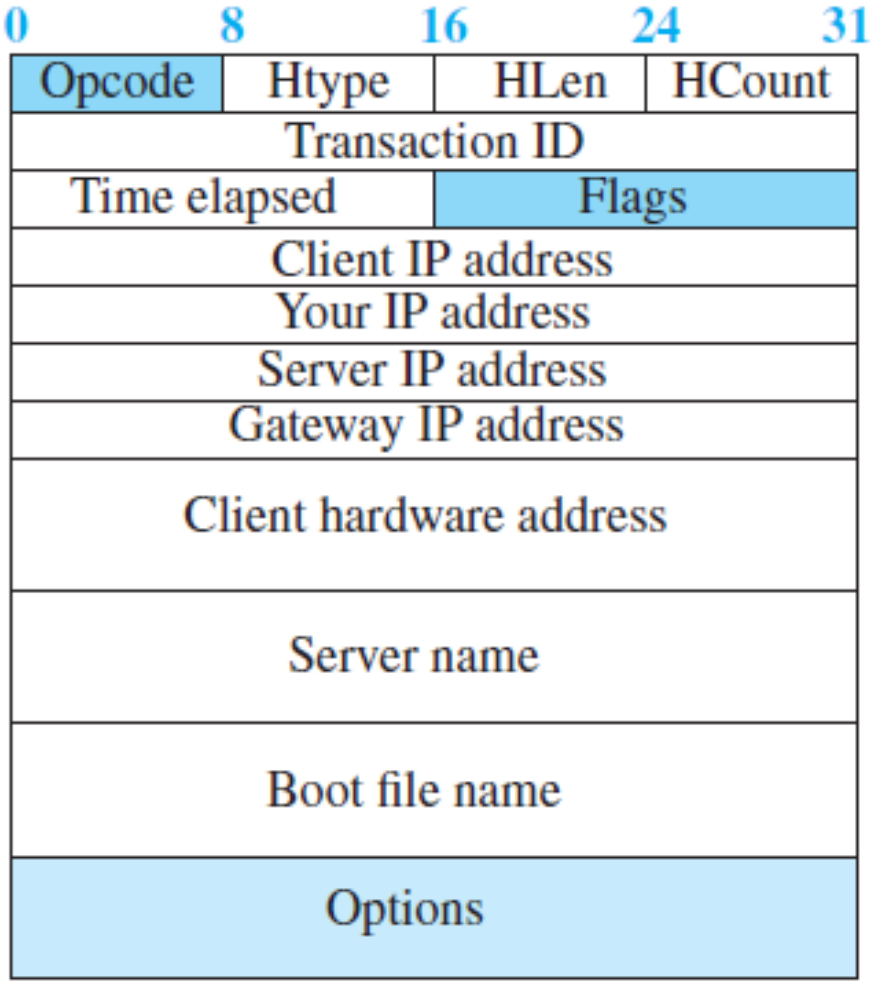
- A network manager can configure DHCP to assign permanent IP addresses to the host and routers. DHCP can also be configured to provide temporary, on demand, IP addresses to hosts.

- DHCP is a client-server protocol in which the client sends a request message and the server returns a response message.

DHCP (Dynamic Host configuration Protocol)



DHCP message format



Fields:

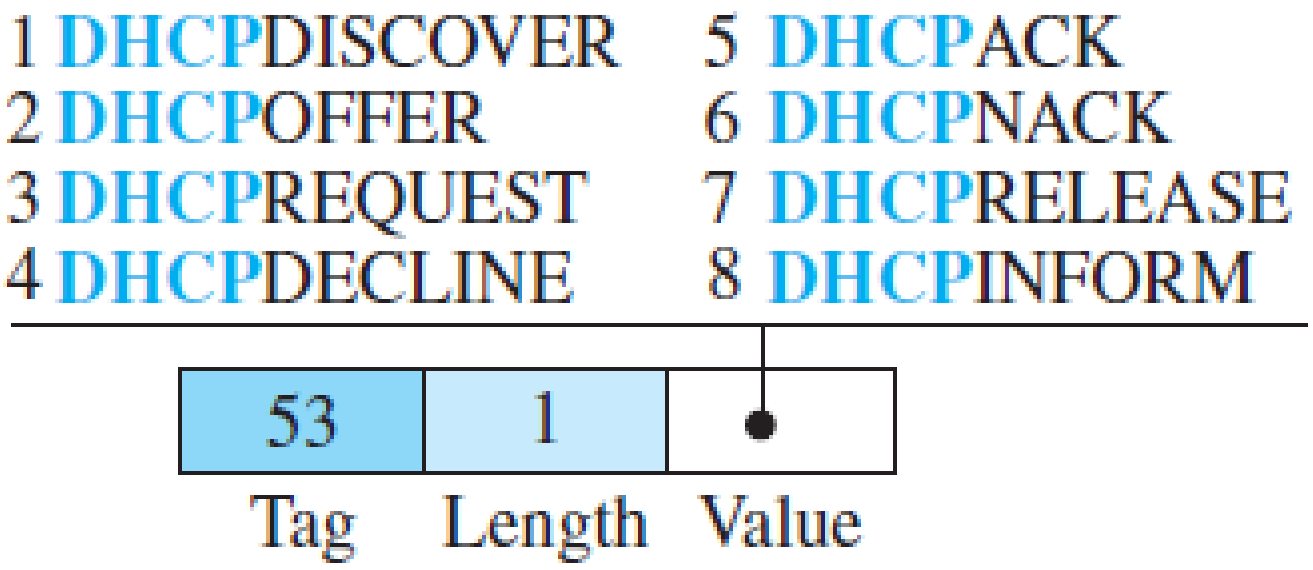
Opcode: Operation code, request (1) or reply (2)
Htype: Hardware type (Ethernet, ...)
HLen: Length of hardware address
HCount: Maximum number of hops the packet can travel
Transaction ID: An integer set by the client and repeated by the server
Time elapsed: The number of seconds since the client started to boot
Flags: First bit defines unicast (0) or multicast (1); other 15 bits not used
Client IP address: Set to 0 if the client does not know it
Your IP address: The client IP address sent by the server
Server IP address: A broadcast IP address if client does not know it
Gateway IP address: The address of default router
Server name: A 64-byte domain name of the server
Boot file name: A 128-byte file name holding extra information
Options: A 64-byte field with dual purpose described in text

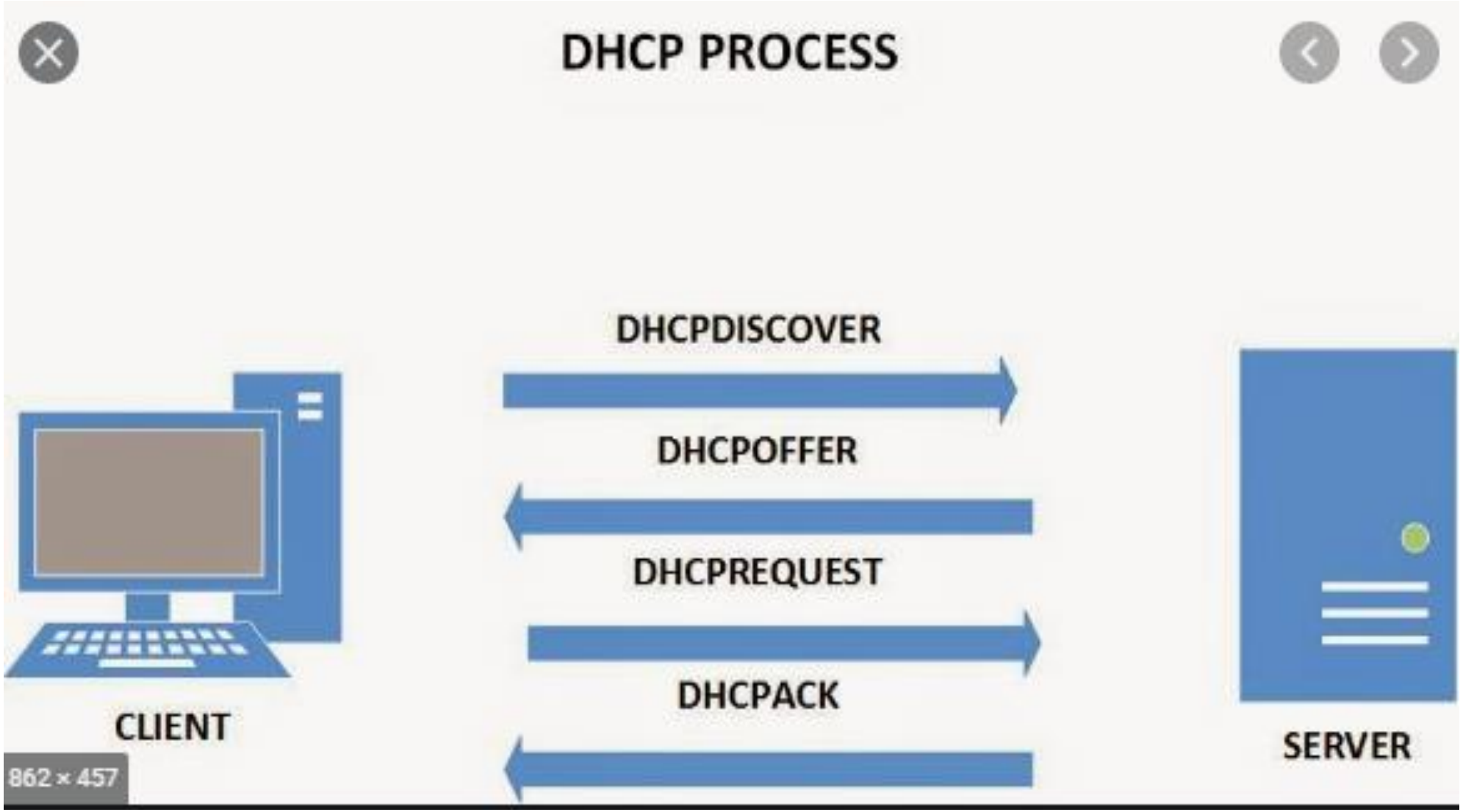
An option is composed of three fields:

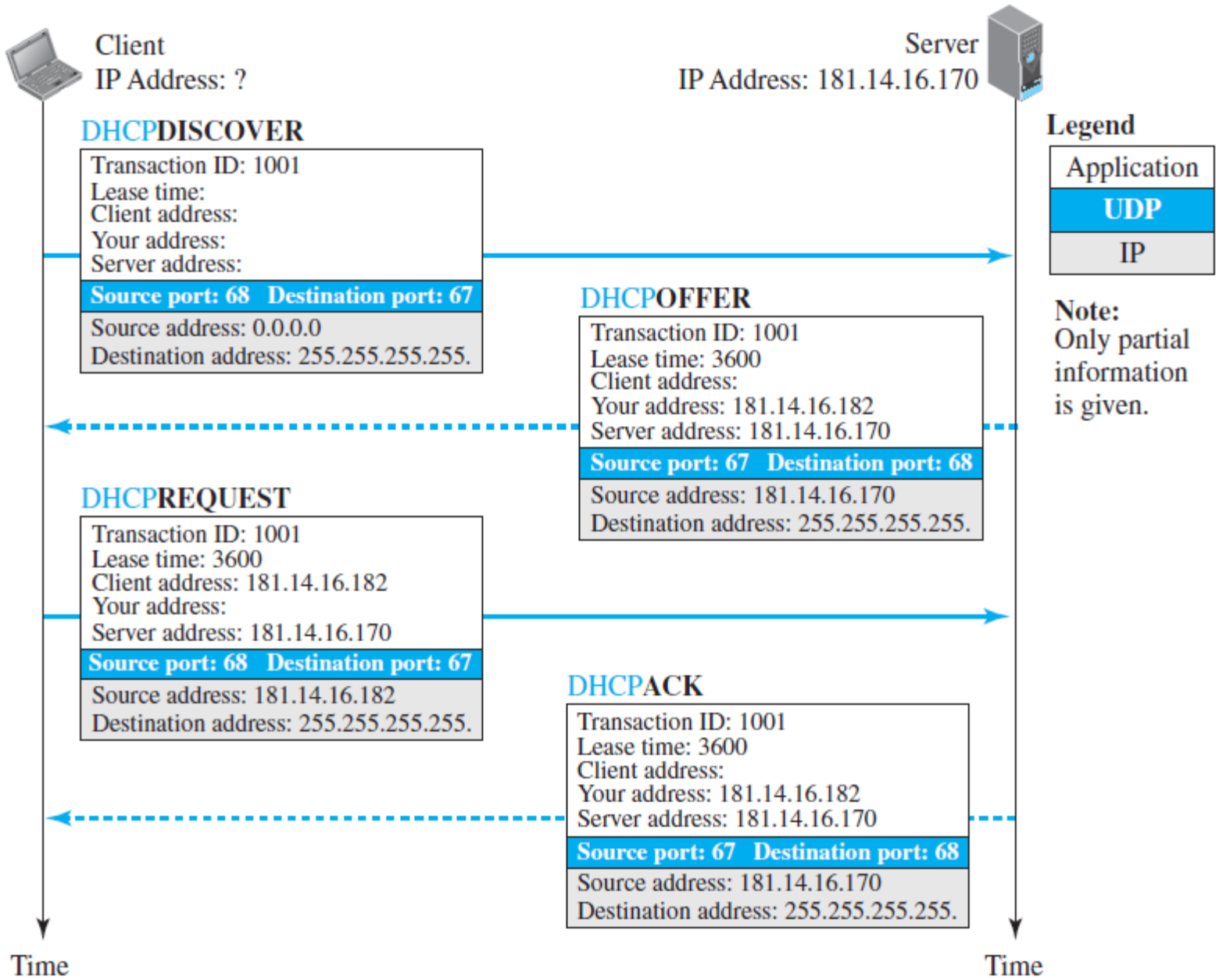
a 1-byte tag field,

a 1-byte length field, and

a variable-length value field.

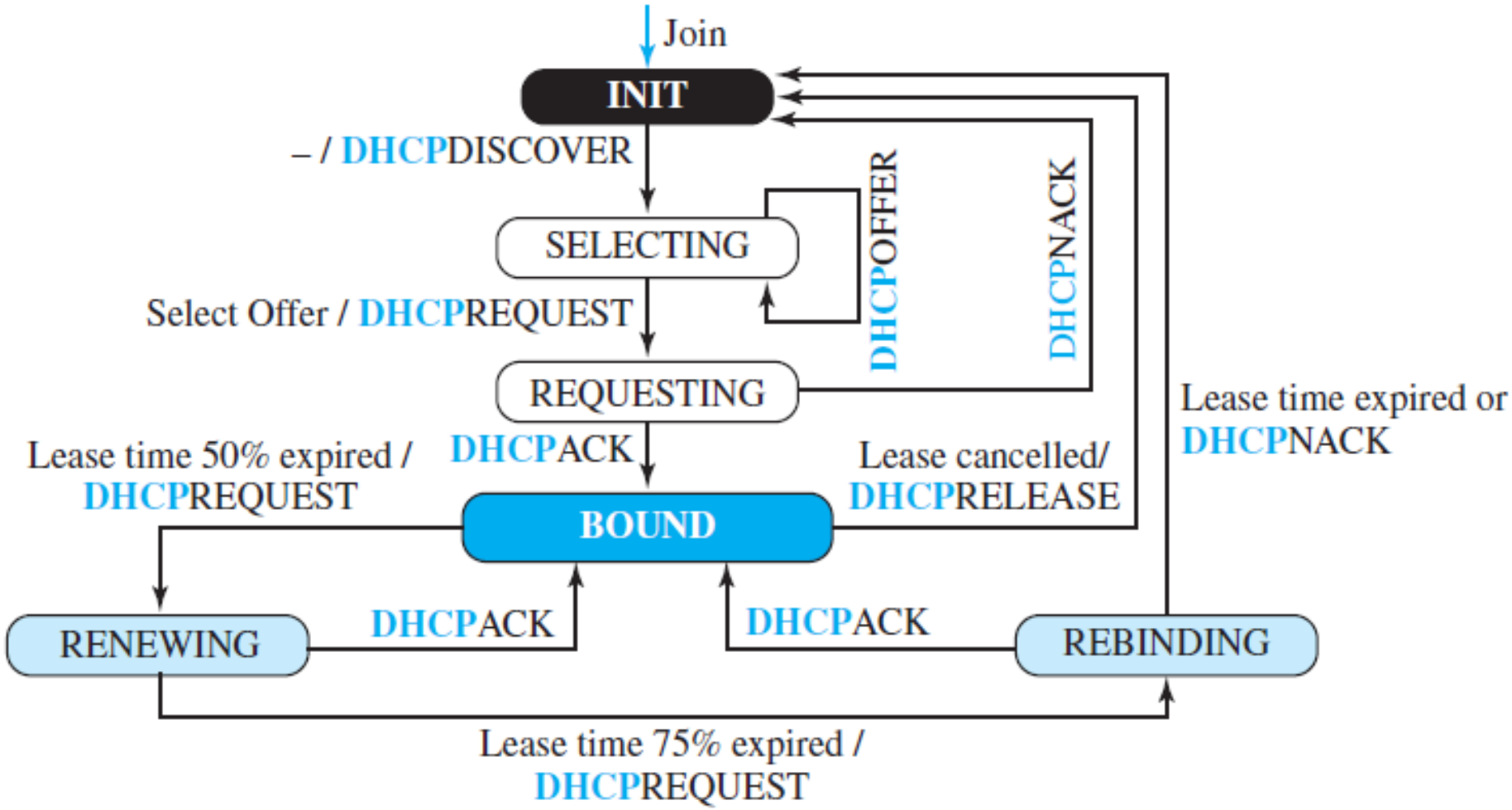



DHCP PROCESS
DHCPDISCOVER
DHCPOFFER
DHCPREQUEST
DHCPACK
CLIENT
SERVER
862 × 457

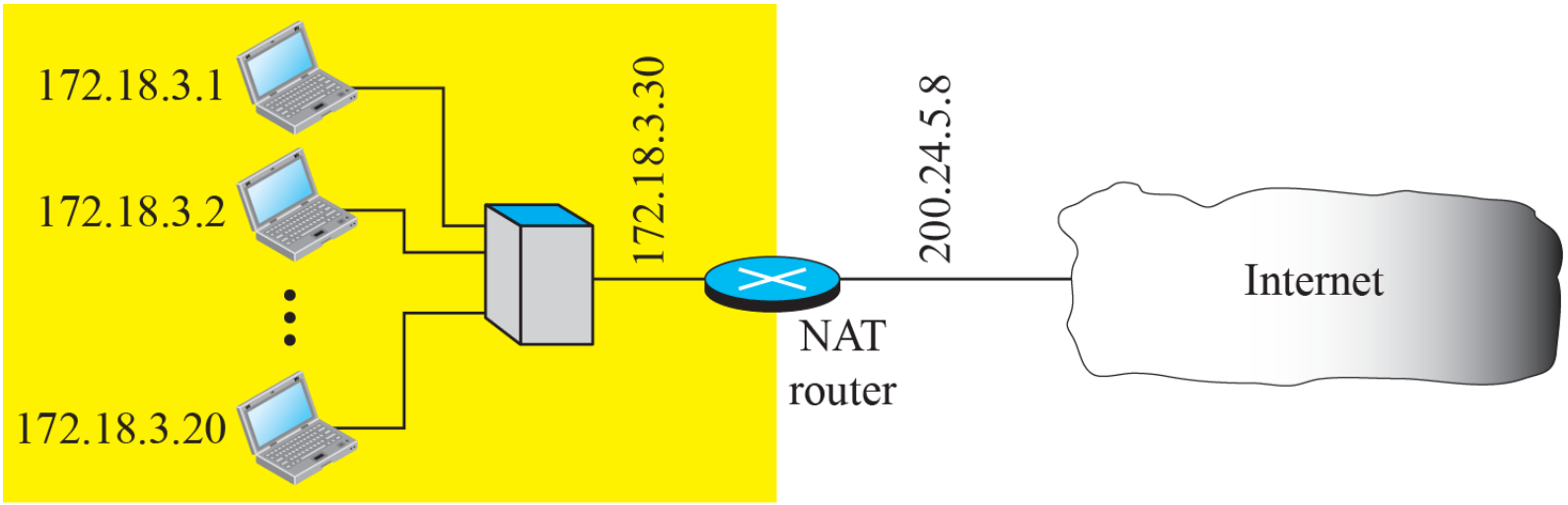


DHCP uses two well-known ports (68 and 67)

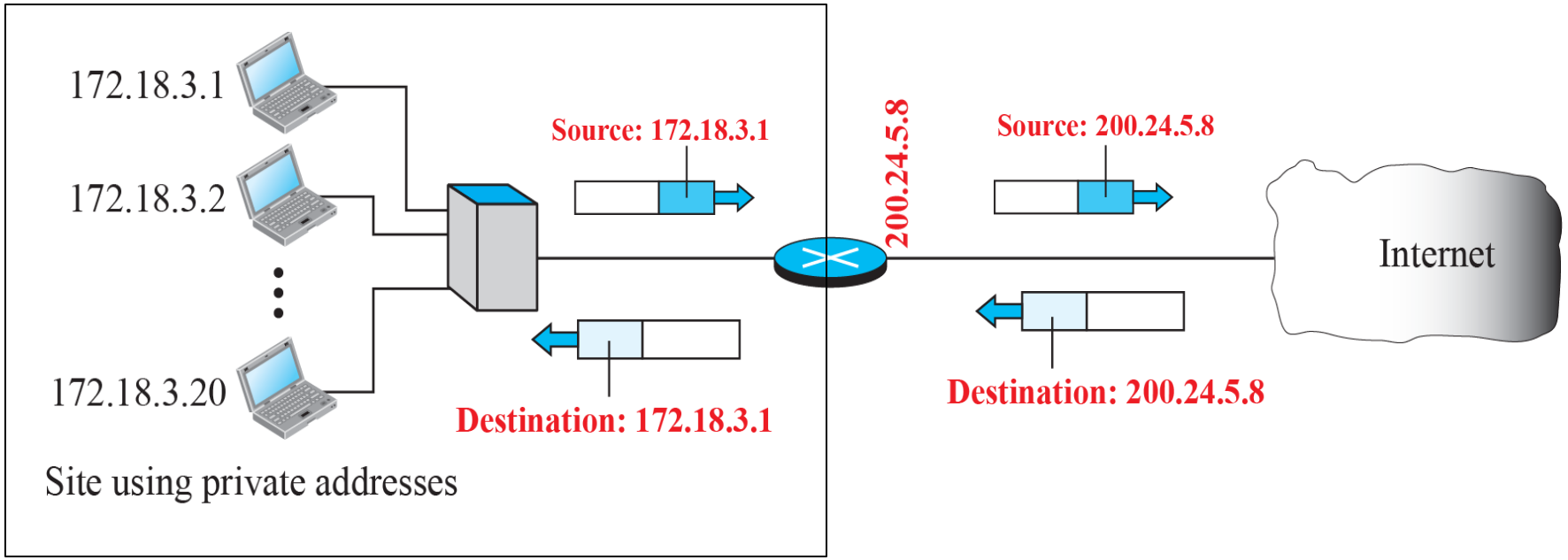
FSM for the DHCP client



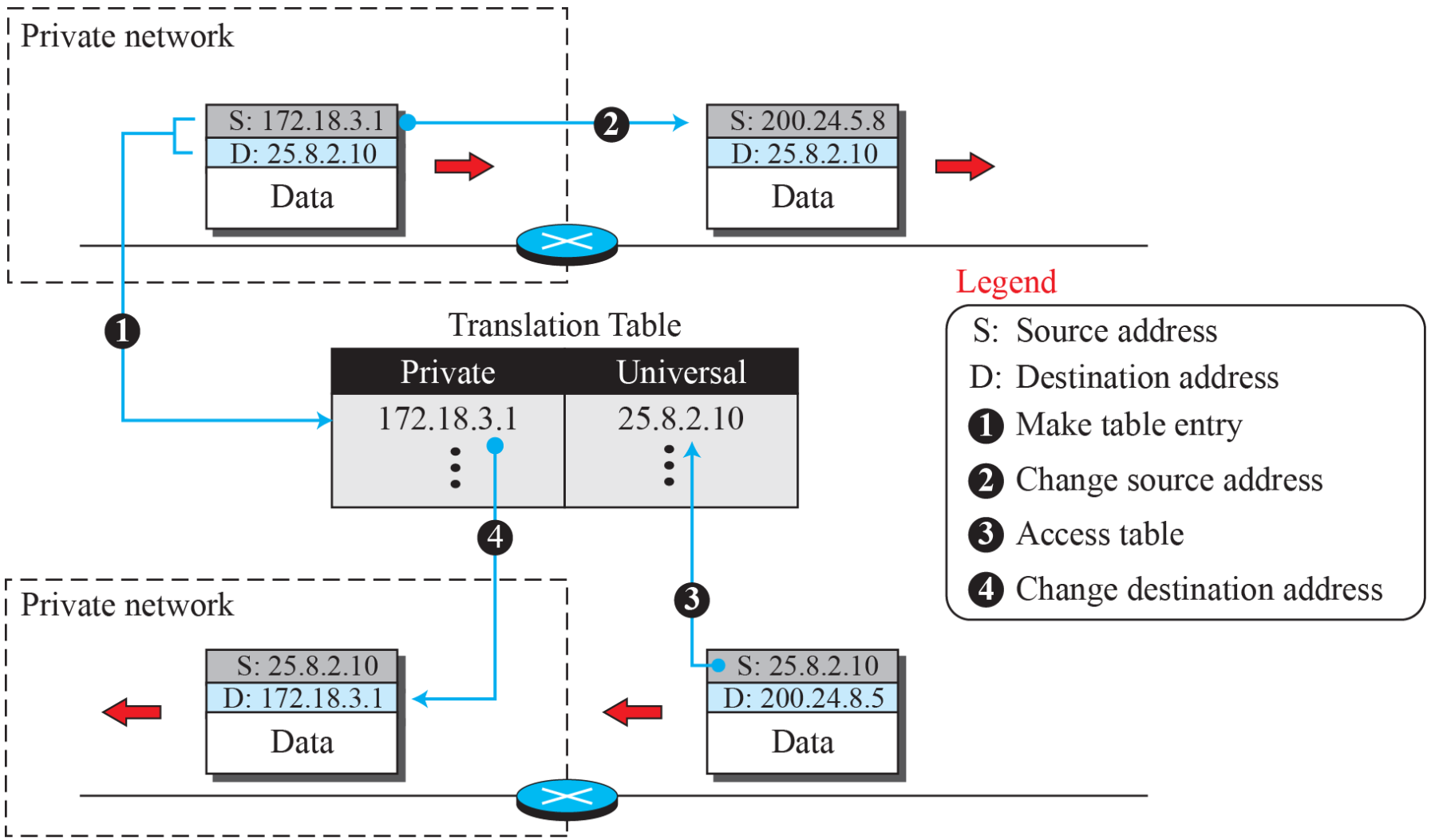
NAT(Network address translation)

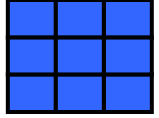


Address translation



Translation





Five-column translation table

Private address	*Private port*	*External address*	*External port*	*Transport protocol*
172.18.3.1	1400	25.8.3.2	80	TCP
172.18.3.2	1401	25.8.3.2	80	TCP
⋮	⋮	⋮	⋮	⋮

NEXT GENERATION IP

The address depletion of IPv4 and other shortcomings of this protocol prompted a new version of IP protocol in the early 1990s. The new version, which is called Internet Protocol version 6 (IPv6) or IP new generation (IPng) was a proposal to augment the address space of IPv4 and at the same time redesign the format of the IP packet and revise some auxiliary protocols such as ICMP.

IPv6 Changes

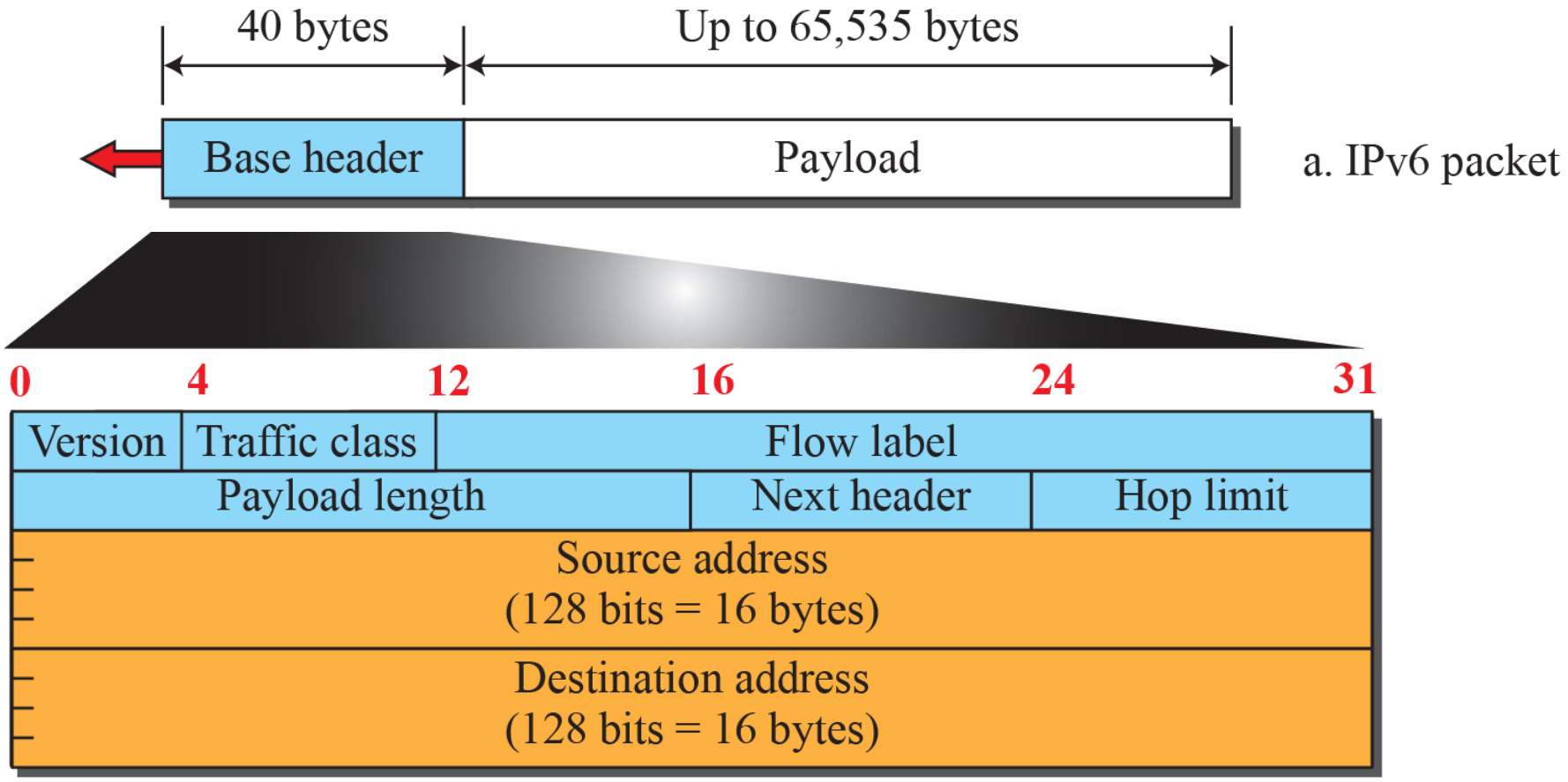
- Larger address space. An IPv6 address is 128 bits long. Compared with the 32-bit address of IPv4, this is a huge (296 times) increase in the address space.
- Better header format. IPv6 uses a new header format in which options are separated from the base header and inserted, when needed, between the base header and the data. This simplifies and speeds up the routing process because most of the options do not need to be checked by routers.
- New options. IPv6 has new options to allow for additional functionalities.
- Allowance for extension. IPv6 is designed to allow the extension of the protocol if required by new technologies or applications.
- Support for resource allocation. In IPv6, the type-of-service field has been removed, but two new fields, traffic class and flow label, have been added to enable the source to request special handling of the packet. This mechanism can be used to support traffic such as real-time audio and video.
- Support for more security. The encryption and authentication options in IPv6 provide confidentiality and integrity of the packet.

Packet Format

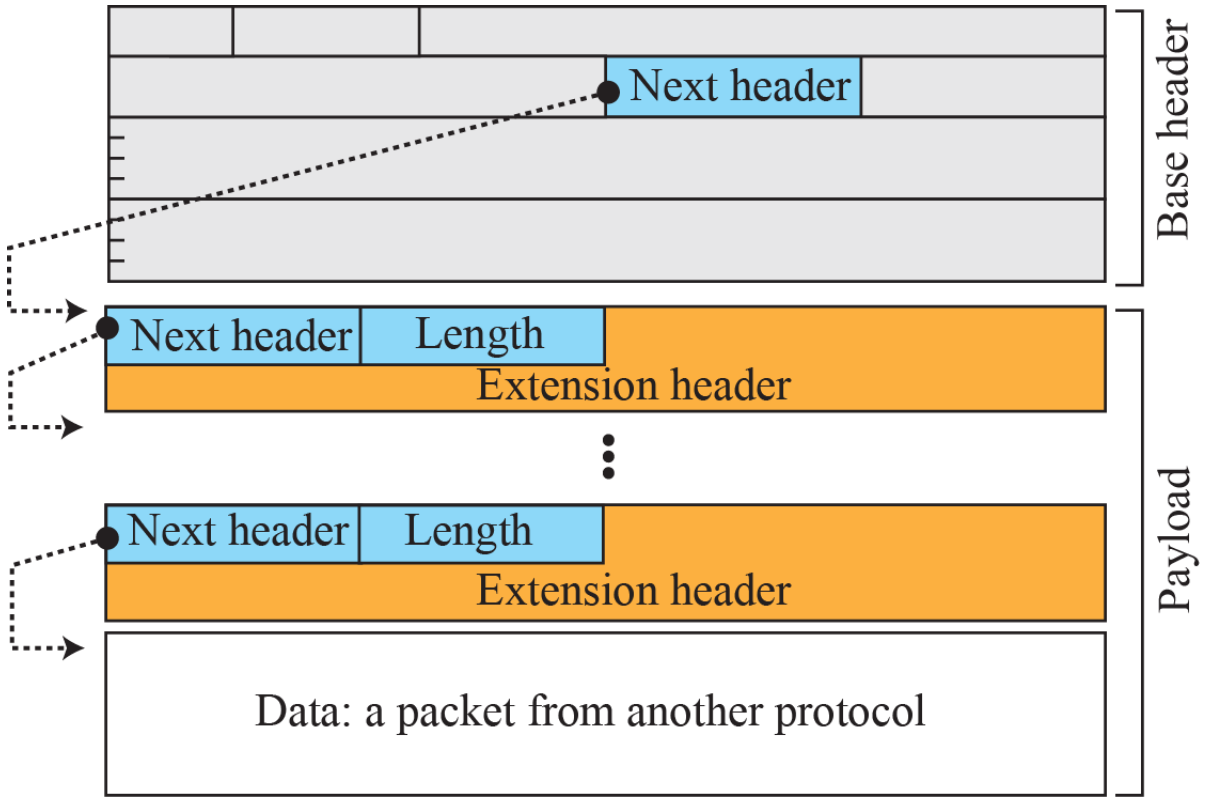
The IPv6 packet is shown in Figure. Each packet is composed of a base header followed by the payload. The base header occupies 40 bytes, whereas payload can be up to 65,535 bytes of information.

- Concept of Flow and Priority in IPv6
- Fragmentation and Reassembly
- Extension Headers

IPv6 datagram

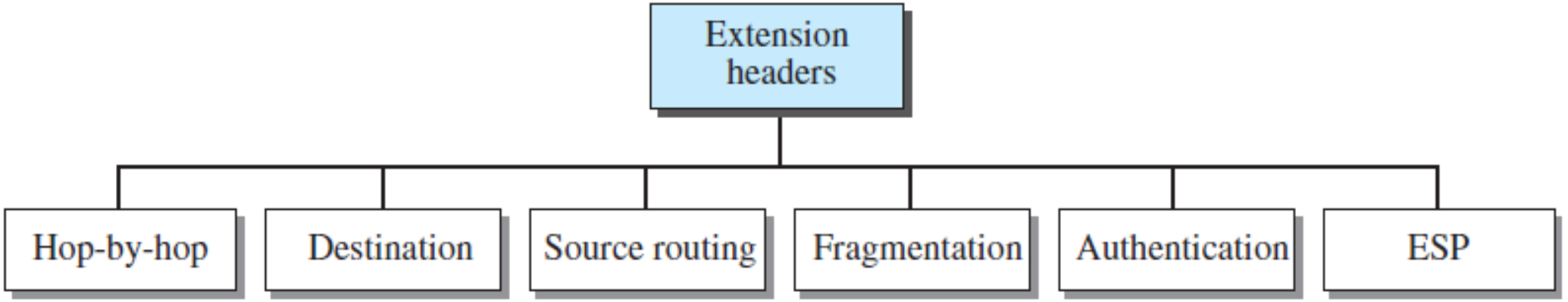


Payload in an IPv6 datagram



Some next-header codes

00: Hop-by-hop option
02: ICMPv6
06: TCP
17: UDP
43: Source-routing option
44: Fragmentation option
50: Encrypted security payload
51:Authentication header
59: Null (no next header)
60: Destination option



- ***Hop-by-Hop Option -*** The *hop-by-hop option* is used when the source needs to pass information to all routers visited by the datagram.
  - ***Pad1.*** This option is 1 byte long and is designed for alignment purposes. Some options need to start at a specific bit of the 32-bit word. If an option falls short of this requirement by exactly one byte, Pad1 is added.
  - ***PadN.*** PadN is similar in concept to Pad1. The difference is that PadN is used when 2 or more bytes are needed for alignment.
  - ***Jumbo payload.*** *T*he length of the payload in the IP datagram can be a maximum of 65,535 bytes. However, if for any reason a longer payload is required, we can use the jumbo payload option to define this longer length.

- ***Destination Option -*** The **destination option** is used when the source needs to pass information to the destination only. Intermediate routers are not permitted access to this information.
- ***Source Routing -*** The source routing extension header combines the concepts of the strict source route and the loose source route options of IPv4.
- ***Fragmentation -*** The concept of **fragmentation** in IPv6 is the same as that in IPv4. However, the place where fragmentation occurs differs. In IPv4, the source or a router is required to fragment if the size of the datagram is larger than the MTU of the network over which the datagram travels. In IPv6, only the original source can fragment. A source must use a **Path MTU Discovery technique** to find the smallest MTU supported by any network on the path. The source then fragments using this knowledge.
- ***Authentication -*** The **authentication** extension header has a dual purpose: it validates the message sender and ensures the integrity of data.
- ***Encrypted Security Payload -*** The **encrypted security payload (ESP)** is an extension that provides confidentiality and guards against eavesdropping.

Comparison of Options between IPv4 and IPv6

- The no-operation and end-of-option options in IPv4 are replaced by Pad1 and PadN options in IPv6.
- The record route option is not implemented in IPv6 because it was not used.
- The timestamp option is not implemented because it was not used.
- The source route option is called the *source route extension header* in IPv6.
- The fragmentation fields in the base header section of IPv4 have moved to the fragmentation extension header in IPv6.
- The authentication extension header is new in IPv6.
- The encrypted security payload extension header is new in IPv6.

IPv6 Addressing

The main reason for migration from IPv4 to IPv6 is the small size of the address space in IPv4. In this section, we show how the huge address space of IPv6 prevents address depletion in the future. We also discuss how the new addressing responds to some problems in the IPv4 addressing mechanism. An IPv6 address is 128 bits or 16 bytes (octets) long, four times the address length in IPv4.

(continued)

- Address Space
- Three Address Types
- Address Space Allocation
  - Global Unicast Addresses
  - Special Addresses
  - Other Assigned Blocks

Colon Hexadecimal

1111111011110110 ... 1111111100000000

FEF6:BA98:7654:3210:ADEF:BBFF:2922:FF00

- The leading zeros of a section can be omitted. Using this form of abbreviation, 0074 can be written as 74.
- Further abbreviation, often called zero compression, can be applied to colon hex notation if there are consecutive sections consisting of zeros only. We can remove all the zeros altogether and replace them with a double semicolon. Note that this type of abbreviation is allowed only once per address. If there is more than one run of zero sections, only one of them can be compressed.

FDEC:0:0:0:0:BBFF:0:FFFF → FDEC::BBFF:0:FFFF

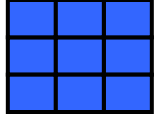
Address Space

The address space of IPv6 contains 2^{128} addresses. This address space is 2^{96} times the IPv4 address—definitely no address

340282366920938463374607431768211456 addresses

Three Address Types

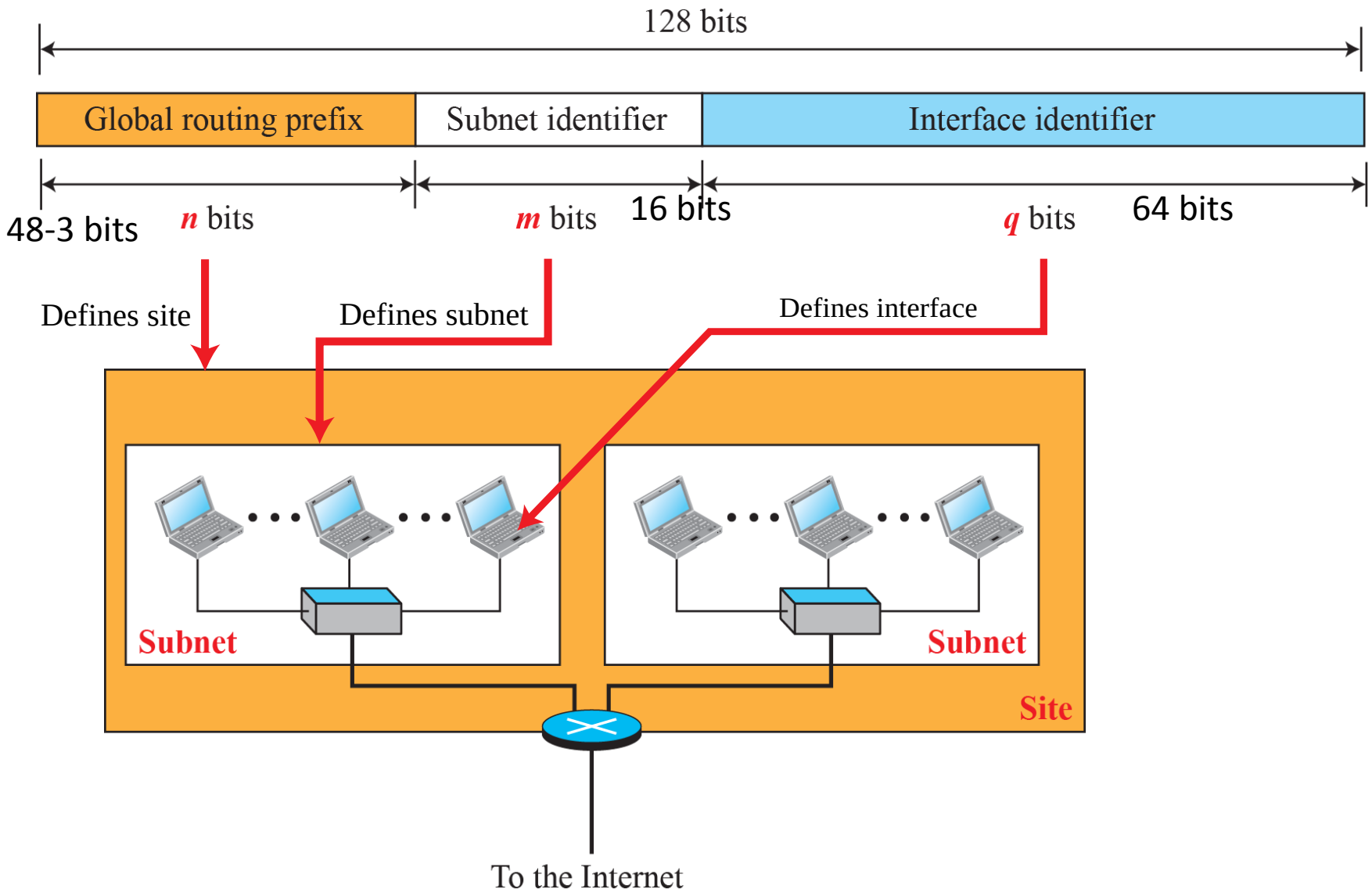
- A unicast address defines a single interface (computer or router). The packet sent to a unicast address will be routed to the intended recipient.
- An anycast address defines a group of computers that all share a single address. A packet with an anycast address is delivered to only one member of the group, the most reachable one.
- A multicast address also defines a group of computers. IPv6 does not define broadcasting



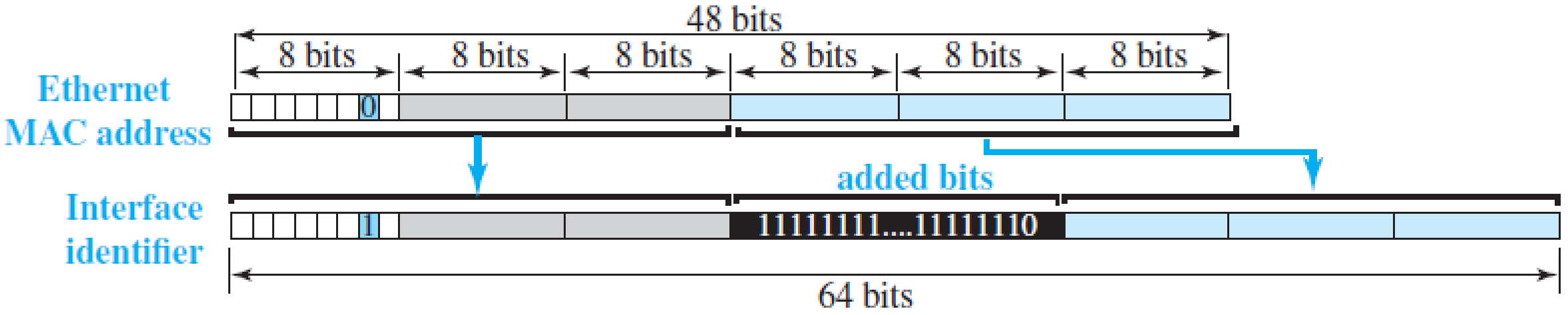
Prefixes for assigned IPv6 addresses

Block prefix	*CIDR*	*Block assignment*	*Fraction*
0000 0000	0000::/8	Special addresses	1/256
001	**2000::/3**	**Global unicast**	**1/8**
1111 110	FC00::/7	Unique local unicast	1/128
1111 1110 10	FE80::/10	Link local addresses	1/1024
1111 1111	FF00::/8	Multicast addresses	1/256

Global unicast address



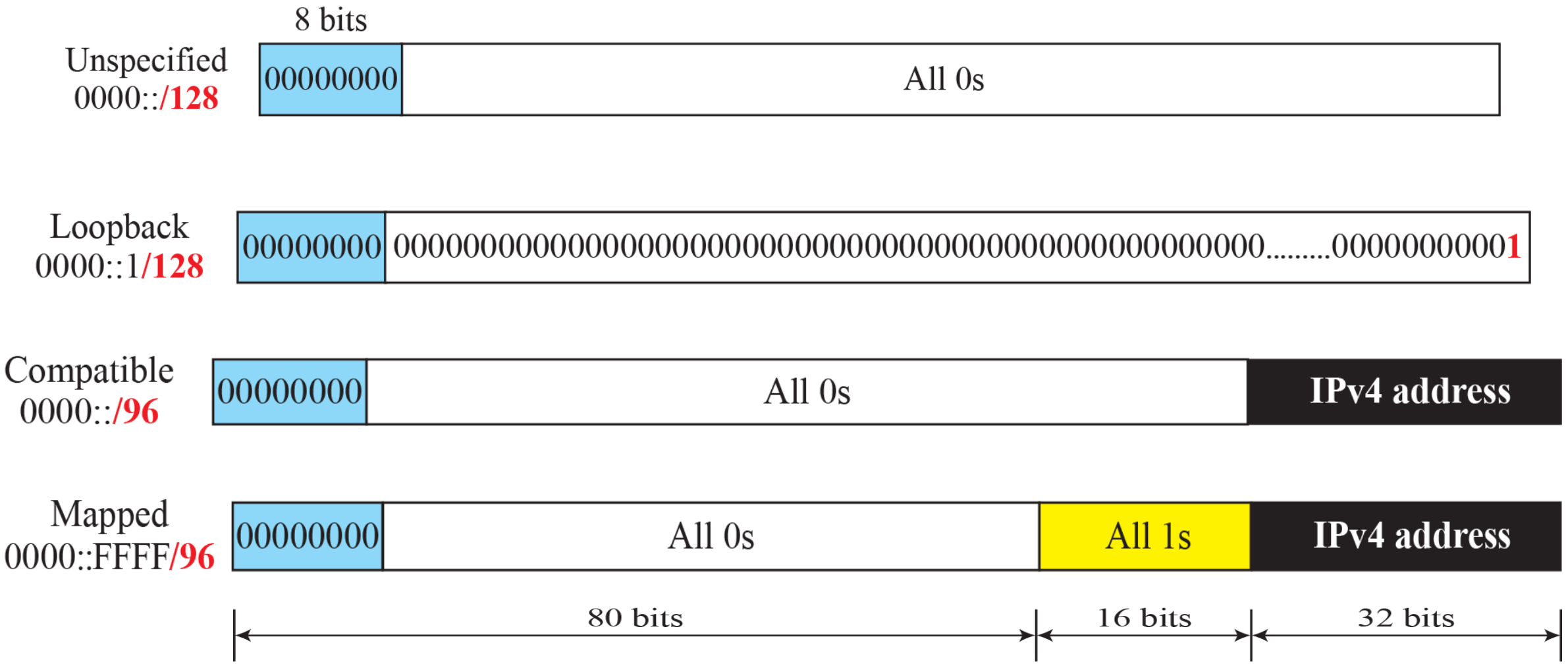
Mapping Ethernet MAC Address



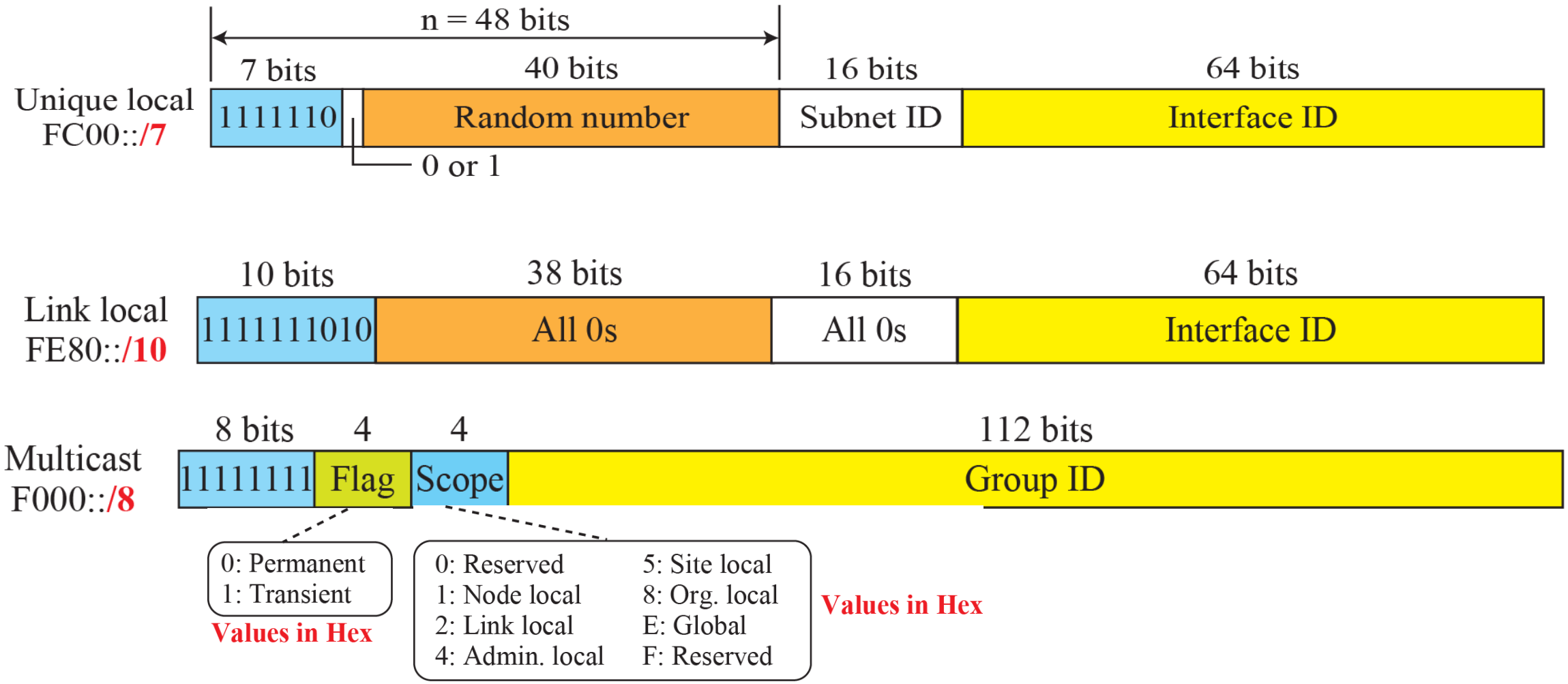
An organization is assigned the block 2000:1456:2474/48. What is the IPv6 address of an interface in the third subnet if the IEEE physical address of the computer is (**F5-A9-23-14-7A-D2**)$_{16}$?

2000:1456:2474:0003:F7A9:23FF:FE14:7AD2/128

Special addresses



Unique local unicast block

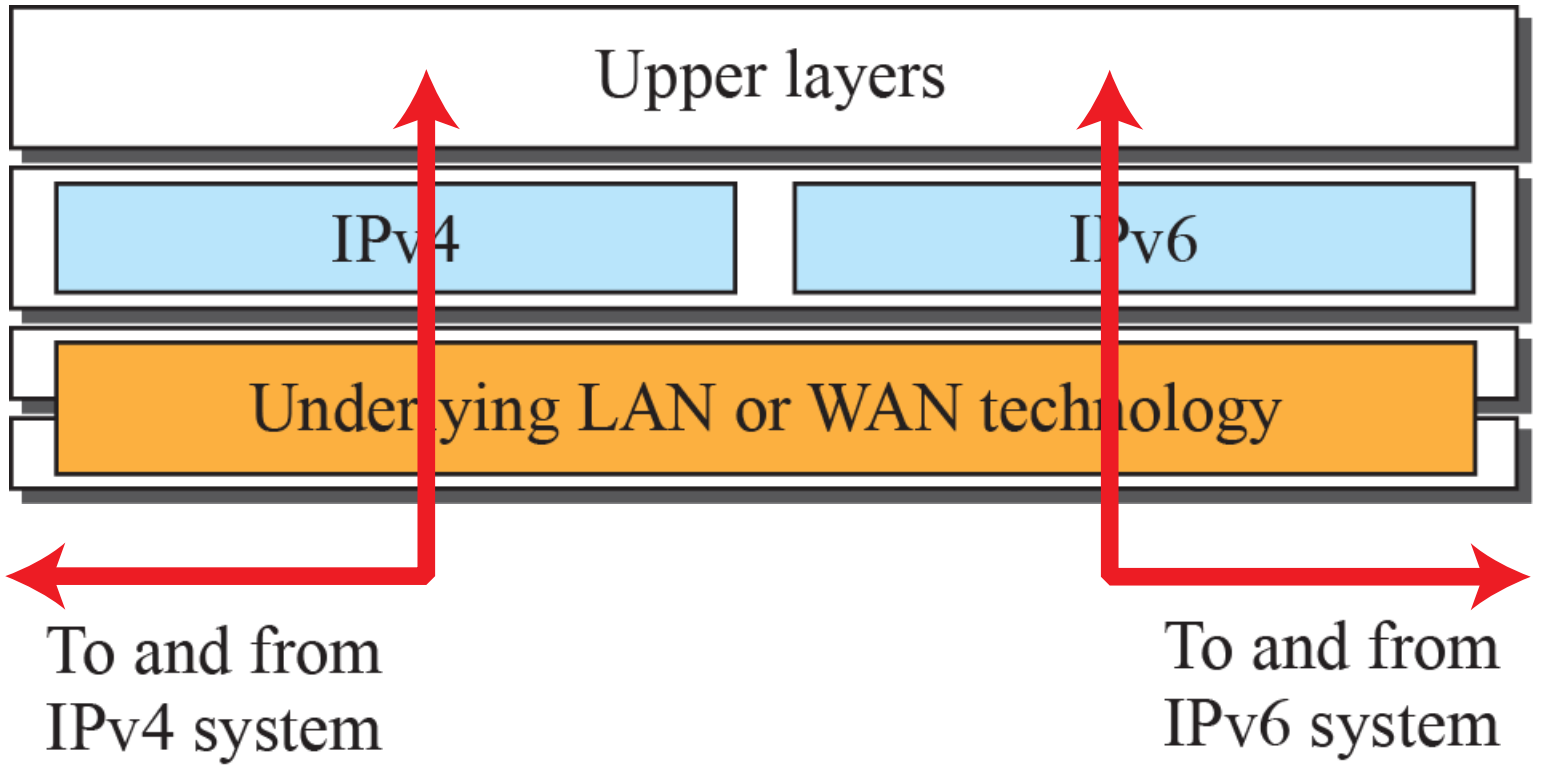


Transition from IPv4 to IPv6

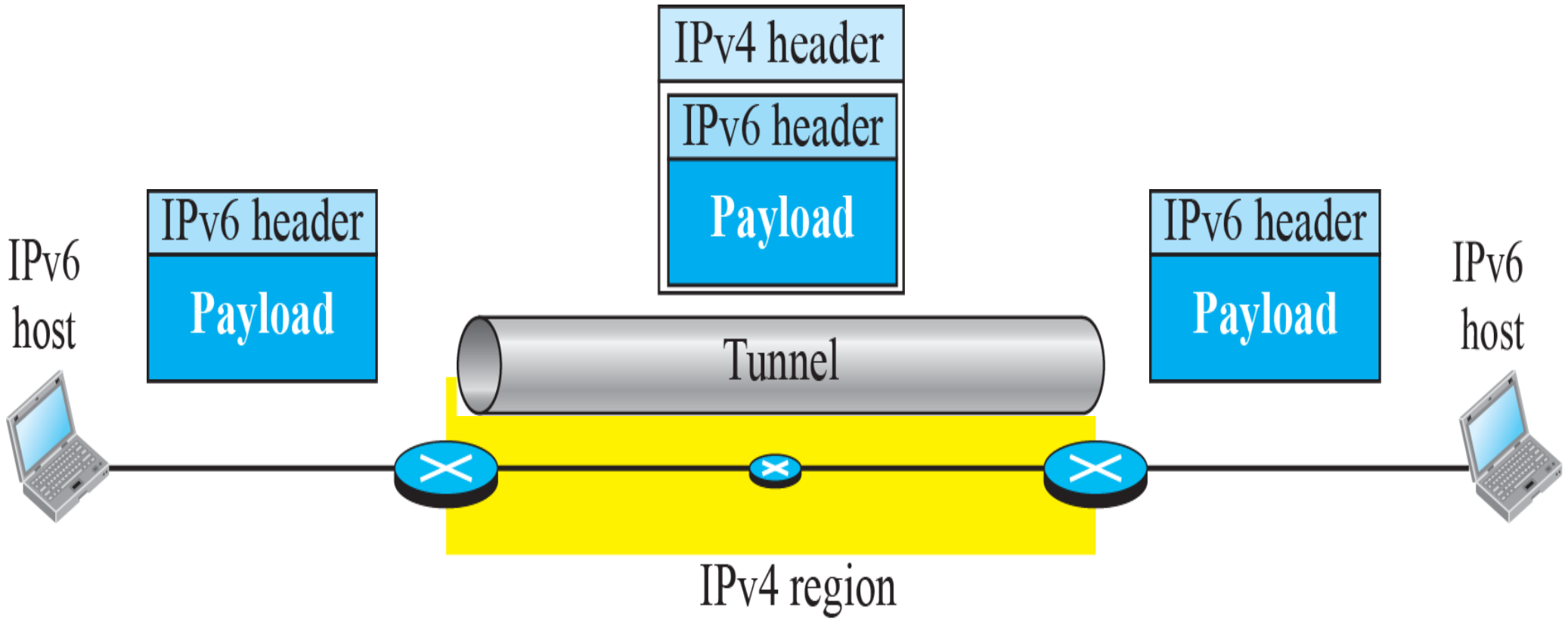
Although we have a new version of the IP protocol, how can we make the transition to stop using IPv4 and start using IPv6? The first solution that comes to mind is to define a transition day on which every host or router should stop using the old version and start using the new version.

- Dual Stack
- Tunneling
- Header Translation

Dual stack


Upper layers
IPv4
IPv6
Underlying LAN or WAN technology
To and from
IPv4 system
To and from
IPv6 system

Tunneling strategy



Header translation strategy