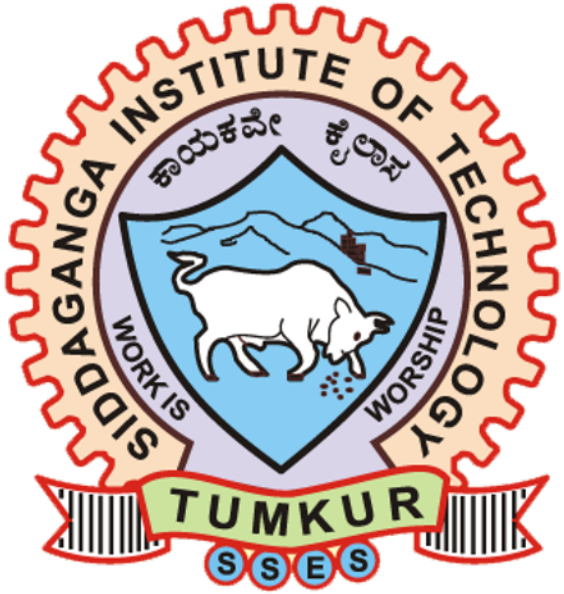
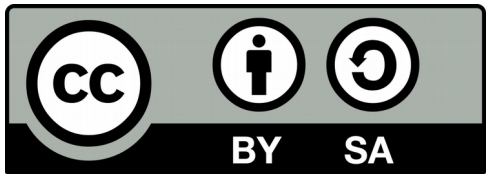
Introduction to Network Layer

SIDDAGANGA INSTITUTE OF TECHNOLOGY

Department of CSE

Prabodh C P





Contents

- IPV4 ADDRESSES
  - Address Space.
  - Hierarchy in addressing
  - Classful Addressing.
  - Classless Addressing.
  - Dynamic Host Configuration Protocol(DHCP).
  - Network Address Resolution (NAT)
- FORWARDING OF IP PACKETS
  - Forwarding Based on Destination Address
  - Forwarding Based on Label
  - Routers as Packet Switches

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.
- IPv4 addresses are unique in the sense that each address defines one, and only one, connection to the Internet.
- If a device has two connections to the Internet, via two networks, it has two IPv4 addresses.

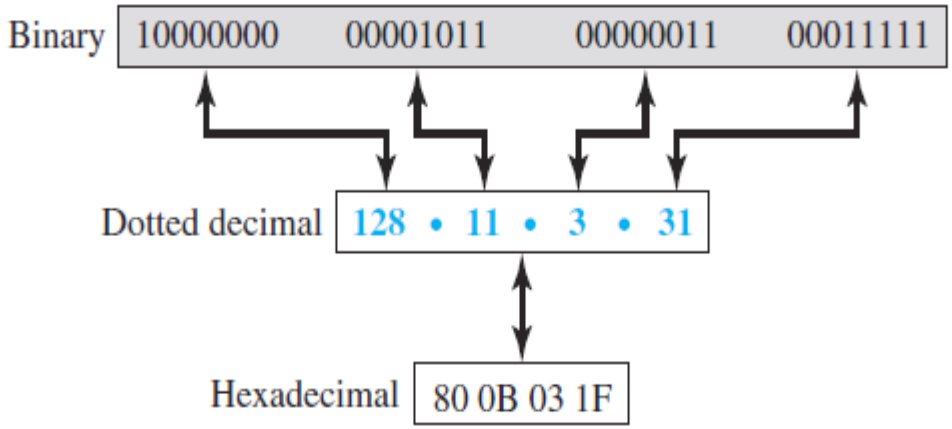
Address Space

- An address space is the total number of addresses used by the protocol.
- An address of **b** bits has an address space of 2^b
- IPv4 uses 32-bit addresses, which means that the address space is 2^{32} or 4,294,967,296. (more than four billion).
- IPV4 addresses are global.

Notation

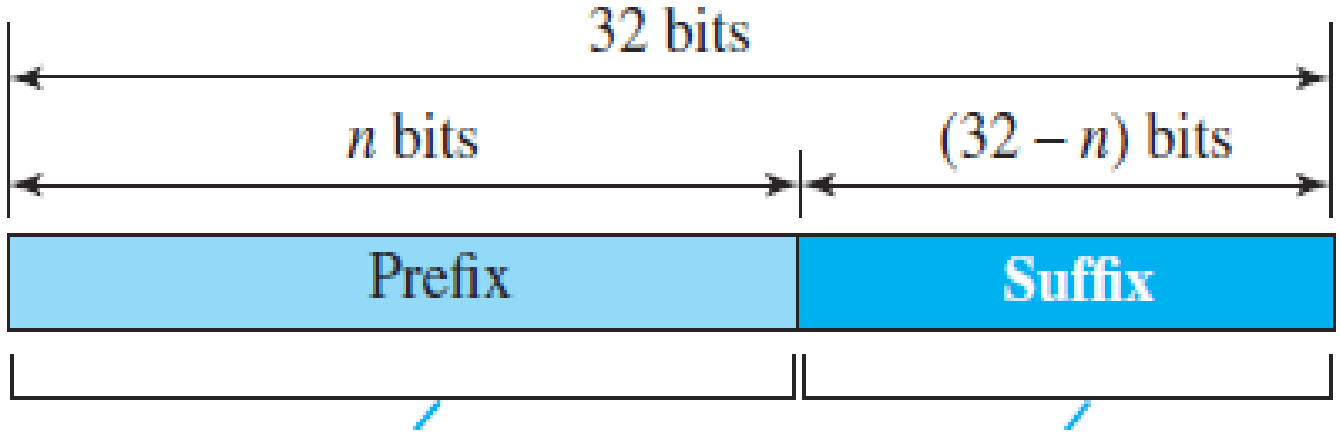
- Three different notations
  - binary notation (base 2),
  - dotted-decimal notation (base 256)
  - hexadecimal notation (base 16).

Figure 18.16 *Three different notations in IPv4 addressing*



Hierarchy in Addressing

- Most communication networks that involves delivery, such as a telephone network or a postal network, the addressing system is hierarchical.
- A 32-bit IPv4 address is also hierarchical, but divided only into two parts.
  - The first part of the address, called the prefix, defines the network;(length : ***n***)
  - the second part of the address, called the suffix, defines the node (connection of a device to the Internet). (length : ***32-n***)



Addressing Schemes

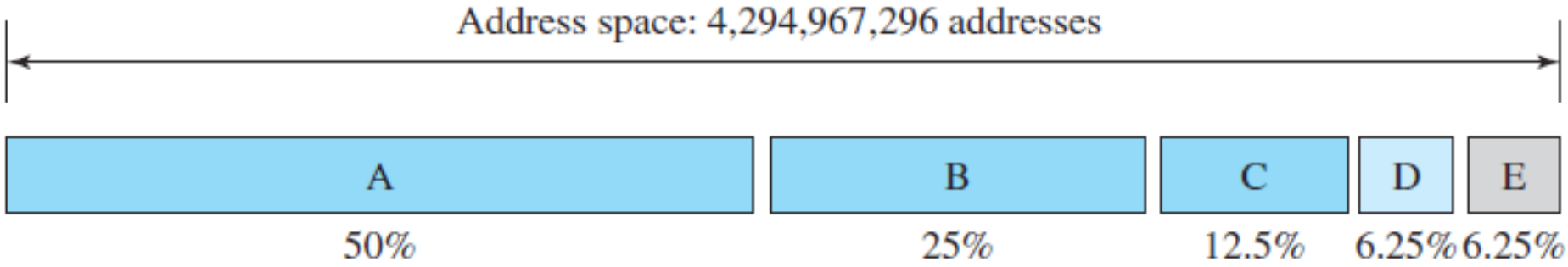
- A prefix can be fixed length or variable length.
- The network identifier in the IPv4 was first designed as a fixed-length prefix. This scheme, which is now obsolete, is referred to as **classful addressing**.
- The new scheme, which is referred to as **classless addressing**, uses a variable-length network prefix.

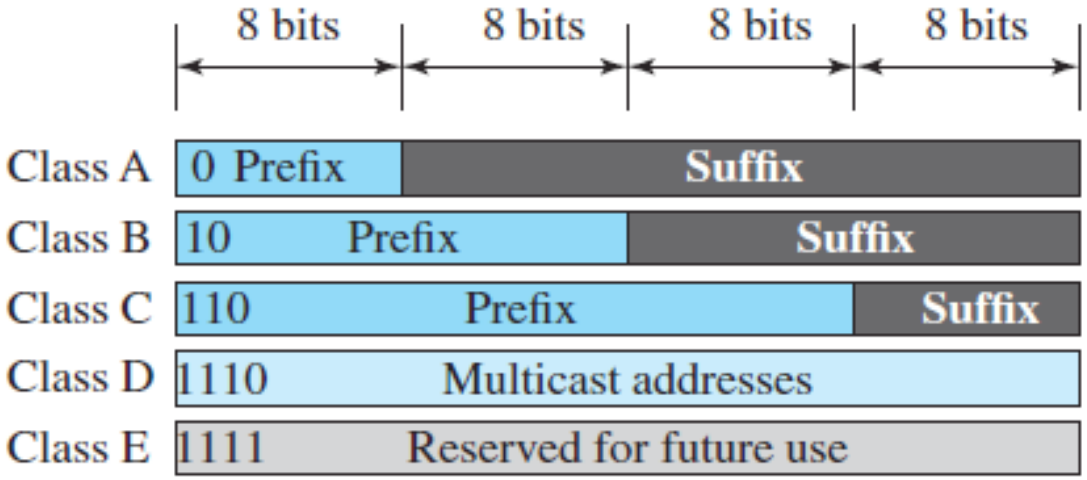
Classful Addressing

- When the Internet started, an IPv4 address was designed with a fixed-length prefix, but to accommodate both small and large networks, three fixed-length prefixes were designed instead of one (n = 8, n = 16, and n = 24).
- The whole address space was divided into five classes (class A, B, C, D, and E)

Classful 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

Classful Addressing

- In class A, the network length is 8 bits, but since the first bit, which is 0, defines the class, we can have only seven bits as the network identifier. There are only 2^7 = 128 networks in the world that can have a class A address.
- class B, network length is 16 bits, first two bits 10, define the class,14 bits as the network identifier. there are only 16,384 networks in the world that can have a class B address.
- In class C, network length is 24 bits , three bits 110 define the class, 21 bits as the network identifier. there are 2,097,152 networks in the world that can have a class C address.
- Class D is not divided into prefix and suffix. It is used for multicast addresses. All addresses that start with 1111 in binary belong to class E. As in Class D, Class E is not divided into prefix and suffix and is used as reserve.

Address Depletion

- Class A:can be assigned to only 128 organizations in the world, but each organization needs to have a single network with 16,777,216 nodes.
- Since there may be only a few organizations that are this large, most of the addresses in this class were wasted (unused).
- Class B addresses were designed for midsize organizations, but many of the addresses in this class also remained unused.
- Class C addresses have a completely different flaw in design, The number of addresses that can be used in each network (256 nodes) was so small that most companies were not comfortable using a block in this address
- Class E addresses were almost never used, wasting the whole class.

Address Depletion

- A network with two nodes uses an entire class C network address, thereby wasting 253 perfectly useful addresses; a class B network with slightly more than 255 hosts wastes over 64,000 addresses.
- The addresses are being rapidly used up, resulting in no more addresses available for organizations and individuals that needed to be connected to the Internet.

Subnetting and Supernetting

- subnetting
  - In subnetting, a class A or class B block is divided into several subnets. Each subnet has a larger prefix length than the original network.
  - subnetting allows the addresses to be divided among several organizations.
  - This idea did not work because most large organizations were not happy about dividing the block and giving some of the unused addresses to smaller organizations.
- Supernetting
  - While subnetting was devised to divide a large block into smaller ones, supernetting was devised to combine several class C blocks into a larger block to be attractive to organizations that need more than the 256 addresses available in a class C block.
  - This idea did not work either because it makes the routing of packets more difficult.

Advantage of Classful Addressing

- Given an address, we can easily find the class of the address and, since the prefix length for each class is fixed, we can find the prefix length immediately.

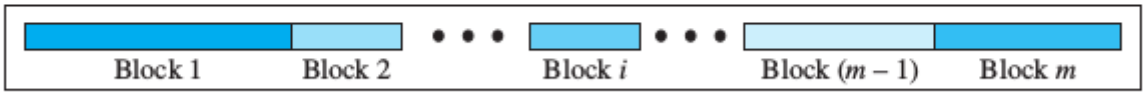
Classless Addressing

- Subnetting and supernetting in classful addressing did not really solve the address depletion problem.
- With the growth of the Internet, it was clear that a larger address space was needed as a long-term solution. The larger address space, however, requires that the length of IP addresses also be increased, which means the format of the IP packets needs to be changed.
- Although the long-range solution has already been devised and is called IPv6 (discussed later),
- A short-term solution was also devised to use the same address space but to change the distribution of addresses to provide a fair share to each organization. The short-term solution still uses IPv4 addresses, but it is called classless addressing

Classless Addressing

- In 1996, the Internet authorities announced a new architecture called classless addressing.
- In classless addressing, the whole address space is divided into variable length blocks.
- The prefix in an address defines the block (network); the suffix defines the node (device). Theoretically, we can have a block of 2^0 , 2^1 , 2^2 , . . . , 2^{32} addresses. One of the restrictions is that the number of addresses in a block needs to be a power of 2.
- the prefix length in classless addressing is variable. We can have a prefix length that ranges from 0 to 32. The size of the network is inversely proportional to the length of the prefix. A small prefix means a larger network; a large prefix means a smaller network.

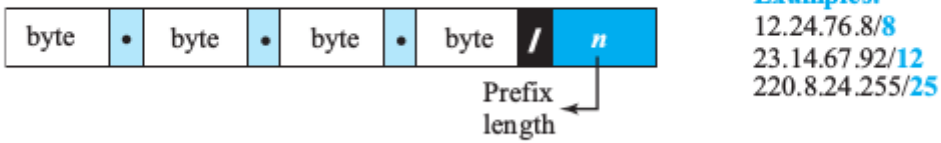
Figure 18.19 *Variable-length blocks in classless addressing*



Prefix Length: Slash Notation

- The first question that we need to answer in classless addressing is how to find the prefix length if an address is given.
- Since the prefix length is not inherent in the address, we need to separately give the length of the prefix.
- In this case, the prefix length, n, is added to the address, separated by a slash.
- The notation is informally referred to as slash notation and formally as classless interdomain routing or CIDR

Figure 18.20 *Slash notation (CIDR)*

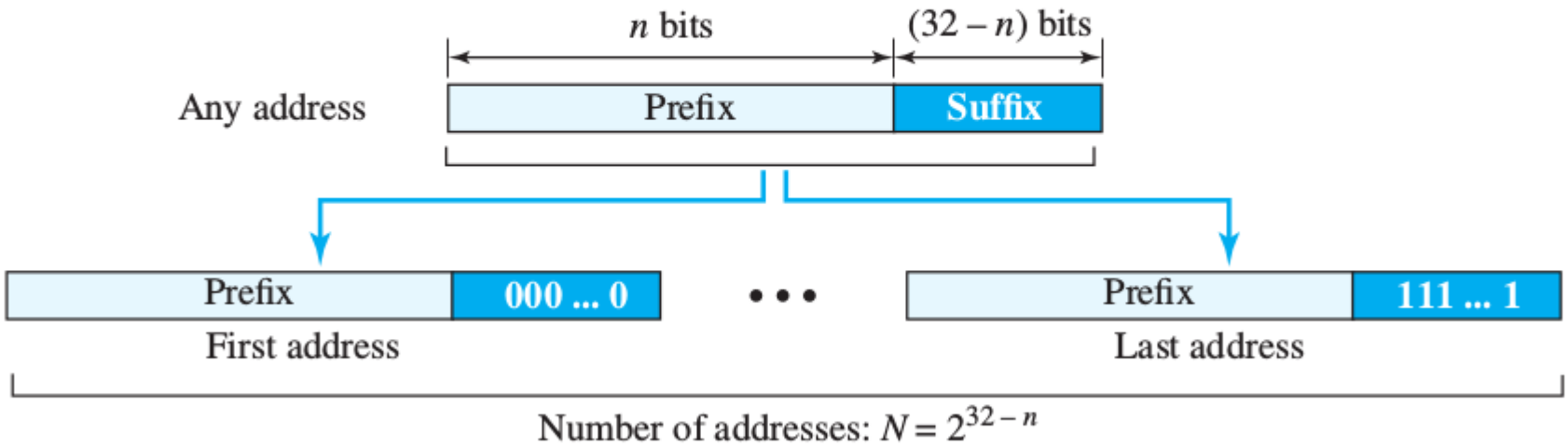


Examples:
12.24.76.8/8
23.14.67.92/12
220.8.24.255/25

Address Information

- The number of addresses in the block is found as N = 2^(32-n);n is prefix.
- To find the first address, we keep the n leftmost bits and set the (32 – n) rightmost bits all to 0s.
- To find the last address, we keep the n leftmost bits and set the (32 – n) rightmost bits all to 1s.

Figure 18.21 *Information extraction in classless addressing*



Example

A classless address is given as **167.199.170.82/ 27** . We can find the above three pieces of information as follows.

The number of addresses in the network is $2^{32-n} = 2^5 = 32$ addresses.

The first address can be found by keeping the first 27 bits and changing the rest of the bits to 0s.

Address: 167.199.170.82/ 27

First address: 167.199.170.64/ 27

10100111 10100111 11000111 11000111

10100111 10100111 11000111 11000000

The last address can be found by keeping the first 27 bits and changing the rest of the bits to 1s.

Address: 167.199.170.82/ 27

Last address: 167.199.170.95/ 27

10100111 10100111 11000111 11000111

10100111 10100111 11000111 11011111

Address Mask

The address mask is a 32-bit number in which

The prefix bits are set to 1s.

The rest of the bits are set to 0s.

A computer can easily find the address mask because it is the complement of $(2^{32-n} - 1)$

1. The number of addresses in the block

   N = NOT (mask) + 1.

2. The first address in the block =(Any address in the block) AND (mask).

3. The last address in the block =(Any address in the block) OR [(NOT (mask)].

Example

A classless address is given as **167.199.170.82/ 27** . We can find the above three pieces of information as follows.

The number of addresses in the network is $2^{32-n} = 2^5 = 32$ addresses.

The first address can be found by keeping the first 27 bits and changing the rest of the bits to 0s.

Address: 167.199.170.82/ 27

First address: 167.199.170.64/ 27

10100111 10100111 11000111 11000111

10100111 10100111 11000111 11000000

The last address can be found by keeping the first 27 bits and changing the rest of the bits to 1s.

Address: 167.199.170.82/ 27

Last address: 167.199.170.95/ 27

10100111 10100111 11000111 11000111

10100111 10100111 11000111 11011111

Example

We repeat previous Example using the mask.

The mask in dotted-decimal notation is 255.255.255.224

Number of addresses in the block:

N = NOT (mask) + 1= 0.0.0.31 + 1 = 32 addresses

First address:

First = (address) AND (mask) = 167.199.170.64

Last address:

Last = (address) OR (NOT mask) = 167.199.170.95

Example

Prefix length:16	Block:	230.8.0.0	230.8.255.255
Prefix length:20	Block:	230.8.16.0	230.8.31.255
Prefix length:26	Block:	230.8.24.0	230.8.24.63
Prefix length:27	Block:	230.8.24.32	230.8.24.63
Prefix length:29	Block:	230.8.24.56	230.8.24.63
Prefix length:31	Block:	230.8.24.56	230.8.24.57

In classless addressing, an address cannot per se define the block the address belongs to.

For example, the address 230.8.24.56 can belong to many blocks

Network Address

Given any address, we can find all information about the block.

The first address, the network address, is particularly important because it is used in routing a packet to its destination network.

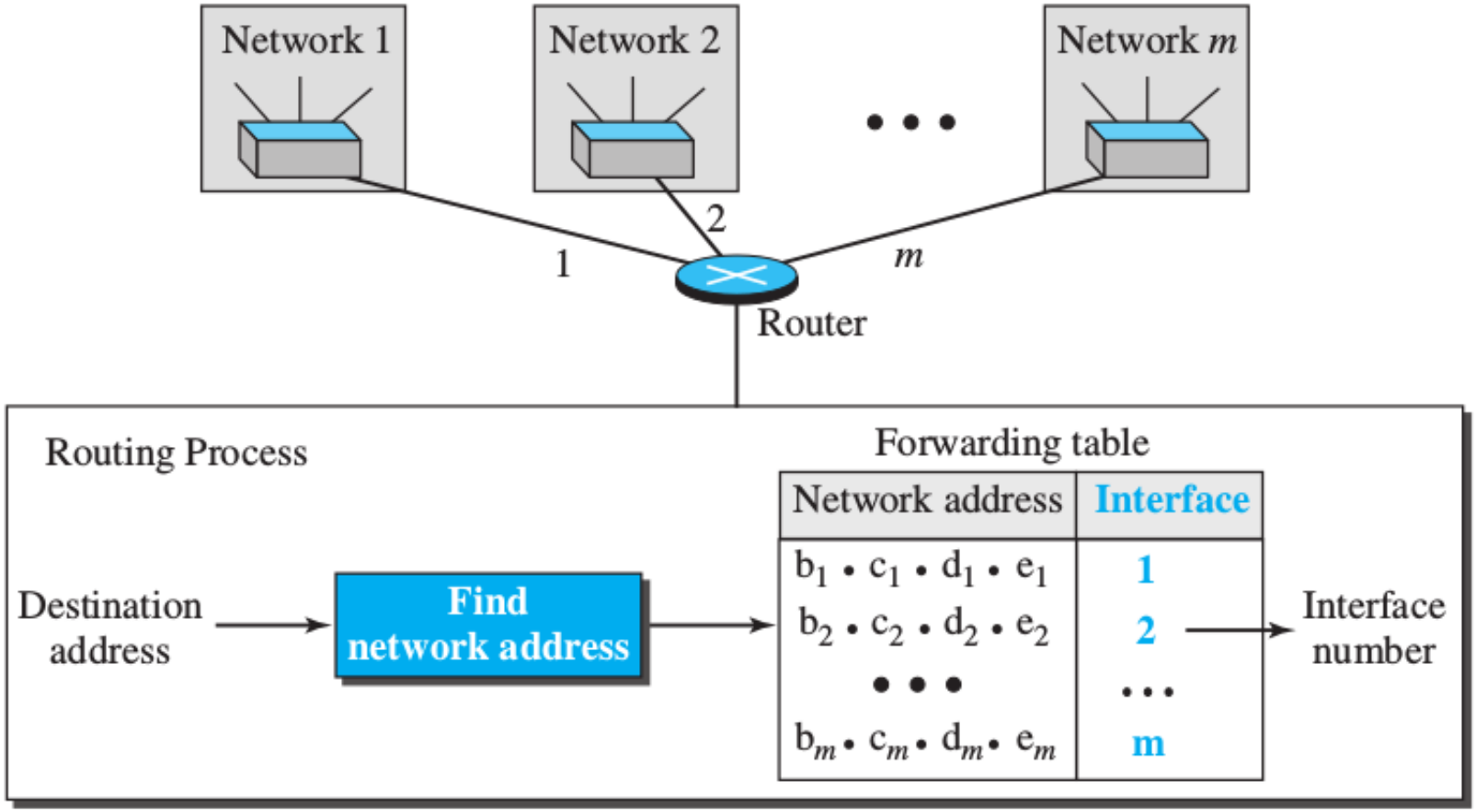
For the moment, let us assume that an internet is made of m networks and a router with m interfaces.

When a packet arrives at the router from any source host, the router needs to know to which network the packet should be sent: from which interface the packet should be sent out.

Once the network address is known, the router consults its forwarding table to find the corresponding interface from which the packet should be sent out.

Each network is identified by its network address

Figure 18.22 *Network address*



Block Allocation

- The ultimate responsibility of block allocation is given to a global authority called the Internet Corporation for Assigned Names and Numbers (ICANN).
- However, ICANN does not normally allocate addresses to individual Internet users.
- It assigns a large block of addresses to an ISP
- For the proper operation of the CIDR, two restrictions need to be applied to the allocated block.

1. The number of requested addresses, N, needs to be a power of 2. The reason is **that**

$N = 2^{32-n}$ or $n = 32 - \log_2 N$. If N is not a power of 2, we cannot have an integer value for n.

2. The requested block needs to be allocated where there is an adequate number of contiguous addresses available in the address space. The first address needs to be divisible by the number of addresses in the block. The reason is that the first address needs to be the prefix followed by (32 – n) number of 0s.

The decimal value of the first address is then

first address = (prefix in decimal) × 2^{32-n} = (prefix in decimal) × N.

Example

An ISP has requested a block of 1000 addresses.

Since 1000 is not a power of 2, 1024 addresses are granted.

The prefix length is calculated as

$n = 32 - \log_2 1024 = 22$.

An available block, 18.14.12.0/ 22 , is granted to the ISP.

The first address in the block is

00010010 00001110 00001100 00000000

It can be seen that the first address in decimal is

302,910,464, which is divisible by 1024.

Subnetting

- Subnetting is used to create more levels of hierarchy
- An organization (or an ISP) that is granted a range of addresses may divide the range into several subranges and assign each subrange to a subnetwork (or subnet).
- A subnetwork can be divided into several sub-subnetworks.
- A sub-subnetwork can be divided into several sub-sub-subnetworks, and so on.

Designing Subnets

The total number of addresses granted to the organization is **N**, the prefix length is **n**,

the assigned number of addresses to each subnetwork is $\mathbf{N_{sub}}$, and prefix length for each subnetwork is $\mathbf{n_{sub}}$.

Steps in designing Subnets

- The number of addresses in each subnetwork should be a power of 2.
- The prefix length for each subnetwork should be found using the following formula:

$$\mathbf{n_{sub} = 32 - \log_2 N_{sub}}$$

- The starting address in each subnetwork should be divisible by the number of addresses in that subnetwork. This can be achieved if we first assign addresses to larger subnetworks.

Example

An organization is granted a block of addresses with the beginning address 14.24.74.0/24. The organization needs to have 3 subblocks of addresses to use in its three subnets: one subblock of 10 addresses, one subblock of 60 addresses, and one subblock of 120 addresses. Design the subblocks.

There are $2^{32-24} = 256$ addresses in this block.

The first address is 14.24.74.0/24

The last address is 14.24.74.255/24.

To satisfy the third requirement, we assign addresses to subblocks, starting with the largest and ending with the smallest one.

The number of addresses in the largest subblock, which requires 120 addresses, is not a power of 2. We allocate 128 addresses. The subnet mask for this subnet can be found as $\mathbf{n_1 = 32 - \log_2 128 = 25.}$ The first address in this block is 14.24.74.0/25; the last address is

14.24.74.127/25

Example contd..

The number of addresses in the second largest subblock, which requires 60 addresses, is not a power of 2. We allocate 64 addresses.

The subnet mask for this subnet can be found as $\mathbf{n_1 = 32 - \log_2 64 = 26.}$

The first address in this block is 14.24.74.128/26; the last address is 14.24.74.191/26

The number of addresses in the smallest subblock, which requires 10 addresses, is not a power of 2. We allocate 16 addresses.

The subnet mask for this subnet can be found as $\mathbf{n_1 = 32 - \log_2 16 = 27.}$

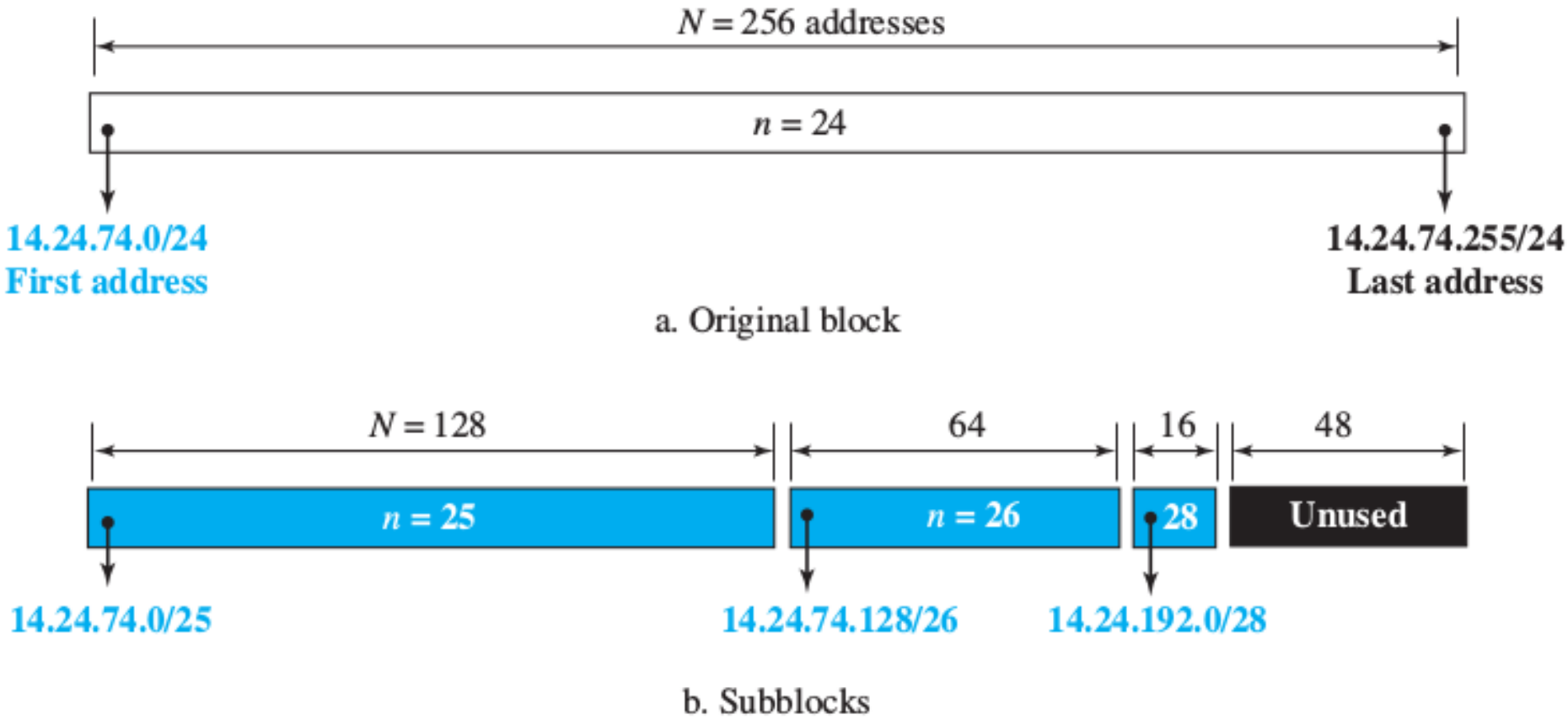
The first address in this block is 14.24.74.192/28; the last address is 14.24.74.207/28

208 address used

48 still available

Subblocks allocation

Figure 18.23 *Solution to Example 18.5*



Address Aggregation

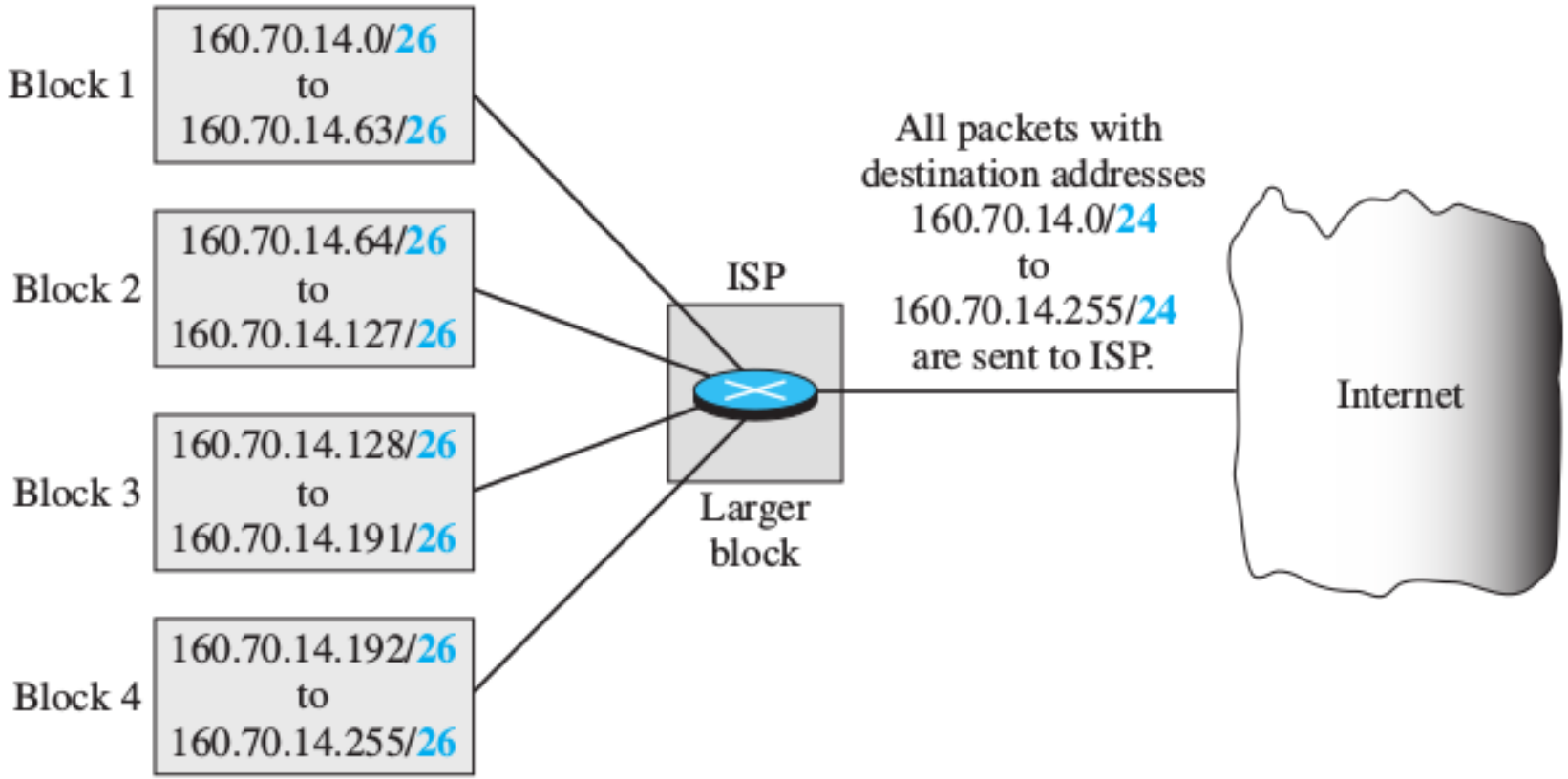
- One of the advantages of the CIDR strategy is address aggregation (sometimes called address summarization or route summarization).
- When blocks of addresses are combined to create a larger block, routing can be done based on the prefix of the larger block.
- ICANN assigns a large block of addresses to an ISP. Each ISP in turn divides its assigned block into smaller subblocks and grants the subblocks to its customers.

Example of address aggregation

- four small blocks of addresses are assigned to four organizations by an ISP.
- The ISP combines these four blocks into one single block and advertises the larger block to the rest of the world.
- Any packet destined for this larger block should be sent to this ISP. It is the responsibility of the ISP to forward the packet to the appropriate organization. T

Address Aggregation

Figure 18.24 *Example of address aggregation*



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. Routers block this datagram

Special Addresses

- 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.
- 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.
- Multicast Addresses
  - The block 224.0.0.0/4 is reserved for multicast addresses.

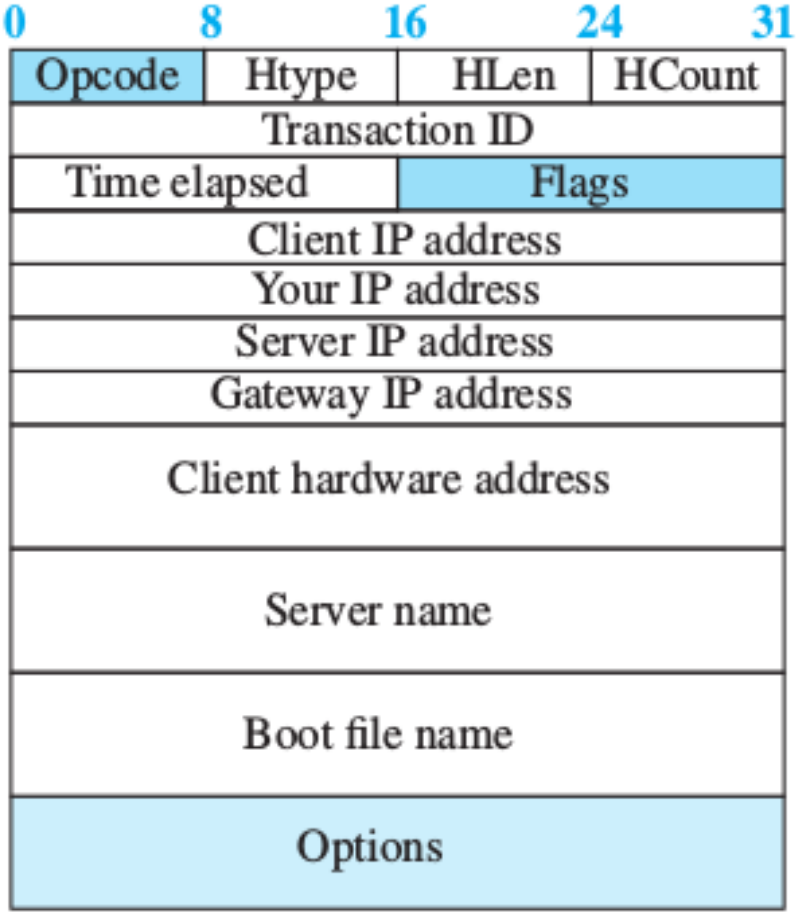
Dynamic Host Configuration Protocol (DHCP)

- ISP can receive a block of addresses directly from ICANN and a small organization can receive a block of addresses from an ISP.
- After a block of addresses are assigned to an organization, the network administrator can manually assign addresses to the individual hosts or routers.
- However, address assignment in an organization can be done automatically using the Dynamic Host Configuration Protocol (DHCP).
- DHCP is an application-layer program, using the client-server paradigm
- DHCP has found such widespread use in the Internet that it is often called a plug-and-play protocol.

DHCP

- Used to assign permanent IP addresses to the host and routers.
- provide a temporary IP address to visitors
- ISP with 1000 granted addresses to service 4000 customers
- Usually four pieces of information are normally needed:
  - the computer address,
  - The prefix,
  - the address of a router, and
  - the IP address of a name server.
- DHCP can be used to provide these pieces of information to the host.

DHCP Message Format

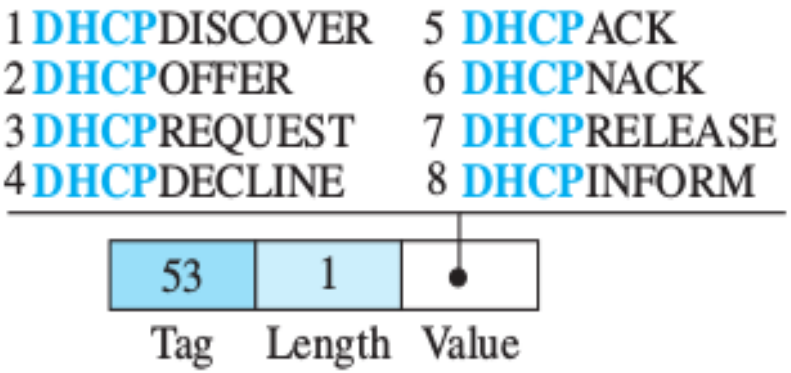


Fields:

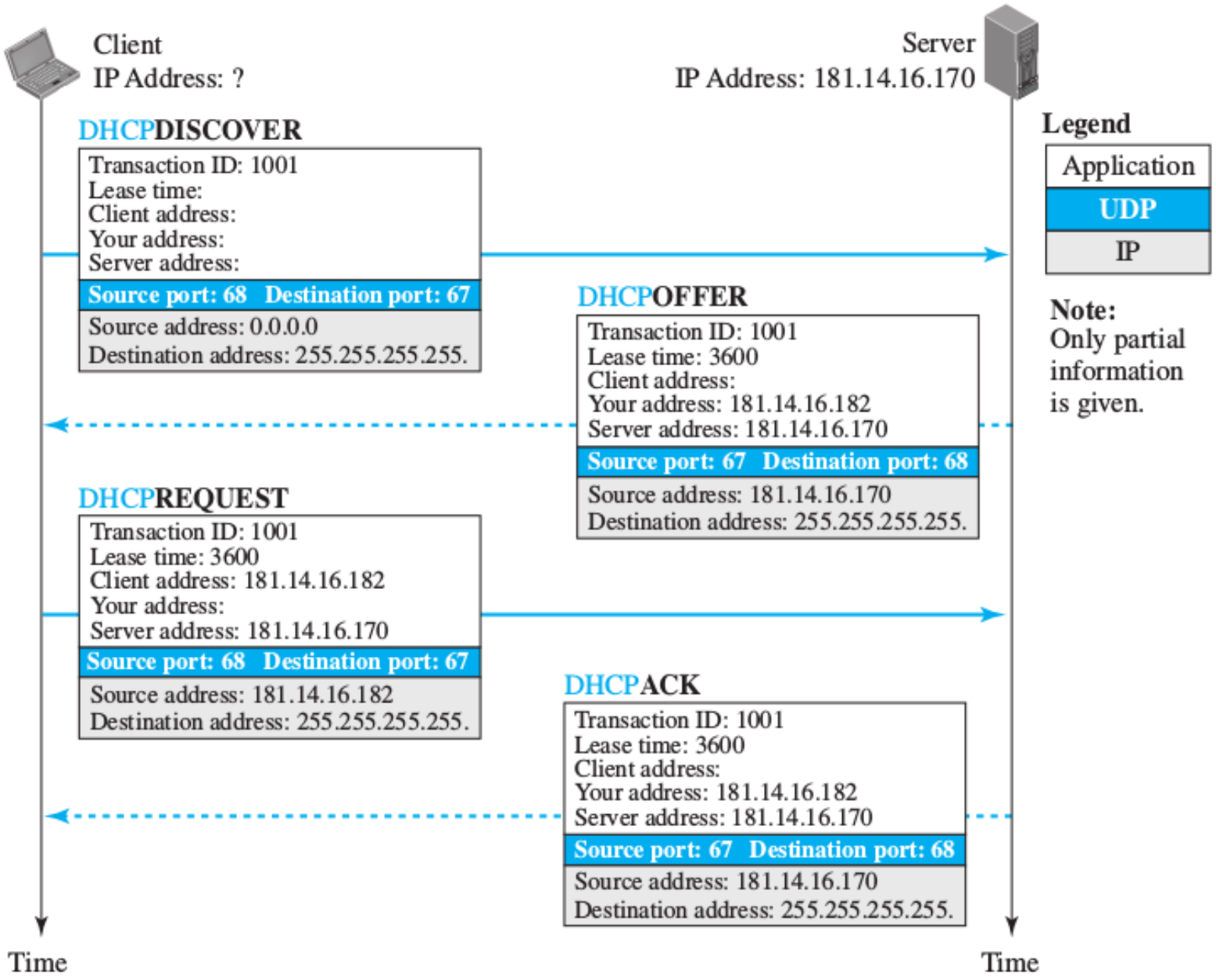
Opcode: Operation code, request (1) or reply (2)
Htype: Hardware type (Ethernet, ...)
HLen: Length of hardware address
HCount: Maximum number of hops the packet can travel
Transaction ID: An integer set by the client and repeated by the server
Time elapsed: The number of seconds since the client started 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

Options Format

- The 64-byte option field has a dual purpose. It can carry either additional information or some specific vendor information.
- The server uses a number, called a magic cookie, in the format of an IP address with the value of 99.130.83.99
- If present, the next 60 bytes are options. An option is composed of three fields: a 1-byte tag field, a 1-byte length field, and a variable-length value field.
- There are several tag fields that are mostly used by vendors.



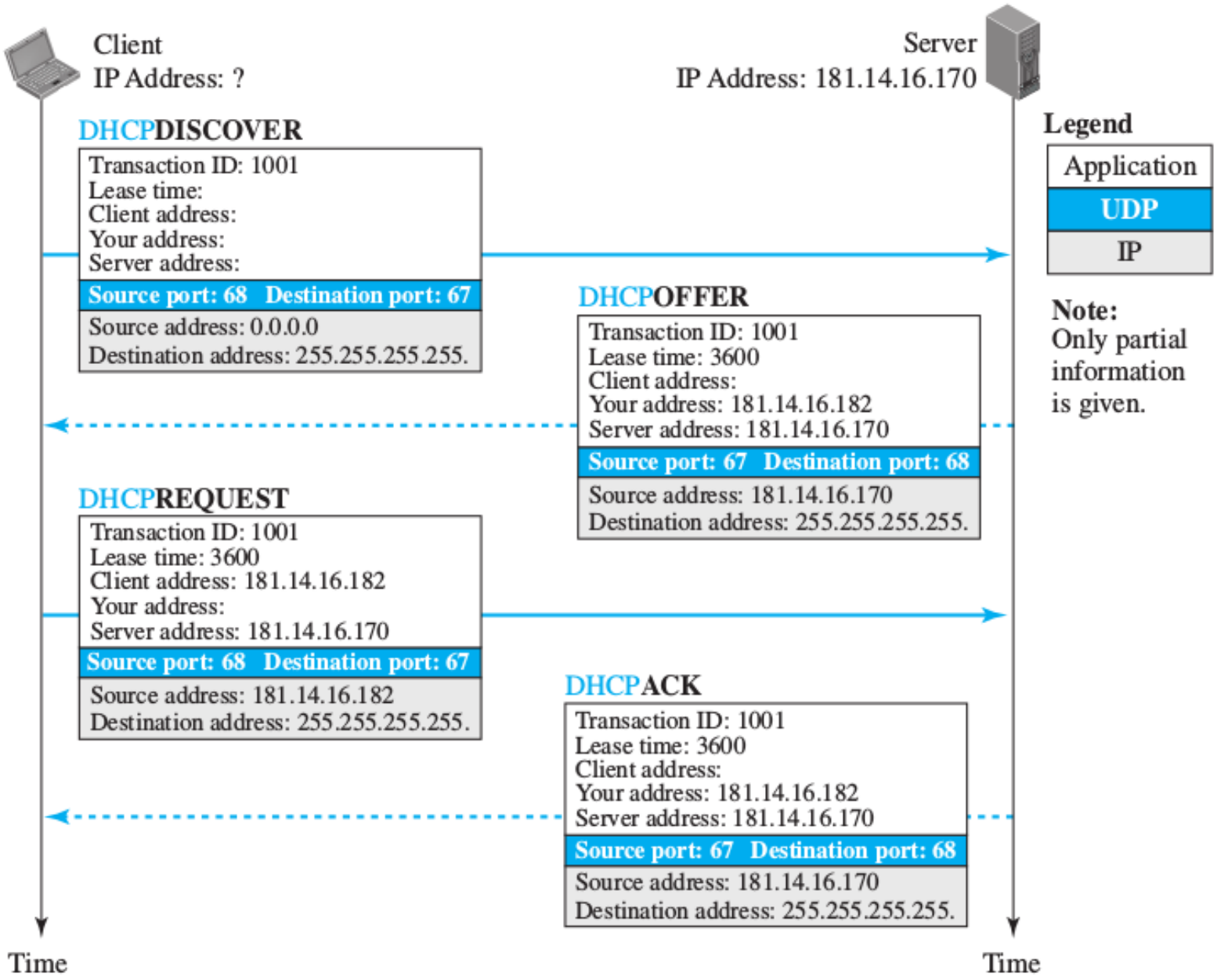
Operation of DHCP



DHCPDISCOVER

- The joining host creates a **DHCPDISCOVER** message in which only the transaction-ID field is set to a random number. No other field can be set because the host has no knowledge with which to do so.
- This message is encapsulated in a UDP user datagram with the source port set to 68 and the destination port set to 67.
- The user datagram is encapsulated in an IP datagram with the source address set to **0.0.0.0** ("this host") and the destination address set to **255.255.255.255** (broadcast address).
- The reason is that the joining host knows neither its own address nor the server address.

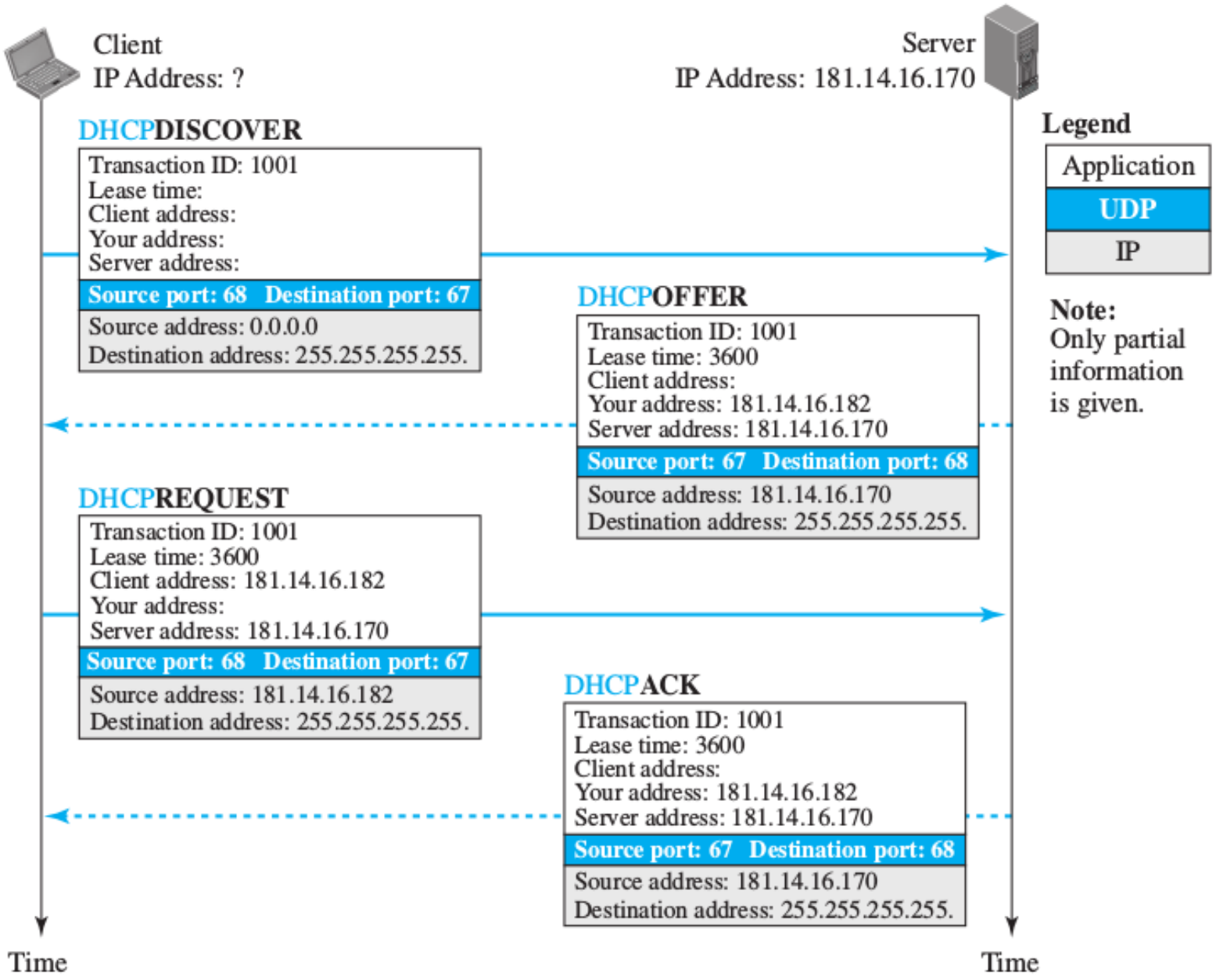
Operation of DHCP



DHCPOFFER

- The DHCP server or servers (if more than one) responds with a **DHCPOFFER** message in which the your address field defines the offered IP address for the joining host and the server address field includes the IP address of the server.
- The message also includes the lease time for which the host can keep the IP address.
- This message is encapsulated in a user datagram with the same port numbers, but in the reverse order.
- The user datagram in turn is encapsulated in a datagram with the server address as the source IP address, but the destination address is a broadcast address, in which the server allows other DHCP servers to receive the offer and give a better offer if they can.

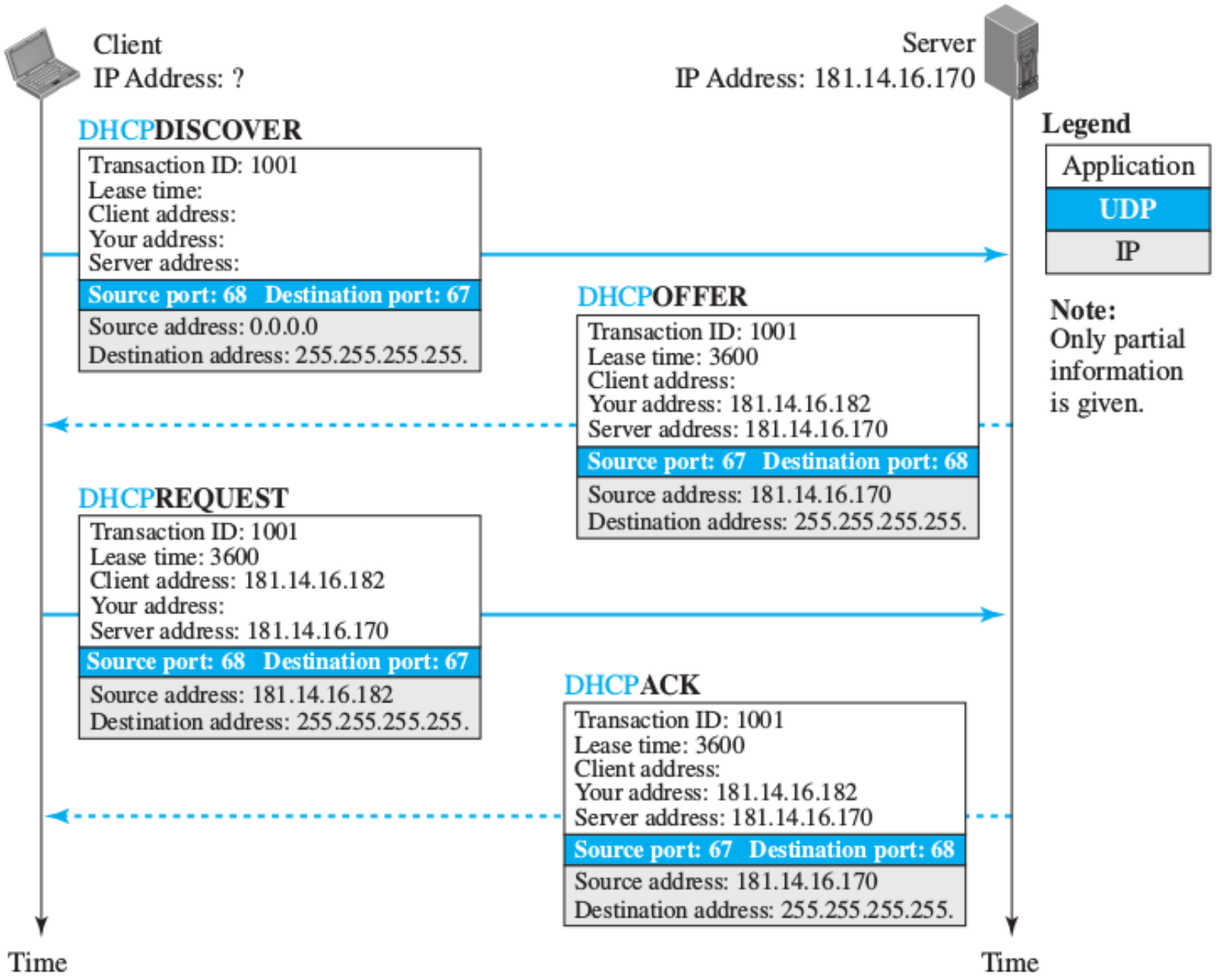
Operation of DHCP



DHCPREQUEST

- The joining host receives one or more offers and selects the best of them.
- The joining host then sends a DHCPREQUEST message to the server that has given the best offer. The fields with known value are set.
- The message is encapsulated in a user datagram with port numbers as the first message.
- The user datagram is encapsulated in an IP datagram with the source address set to the new client address, but the destination address still is set to the broadcast address to let the other servers know that their offer was not accepted.

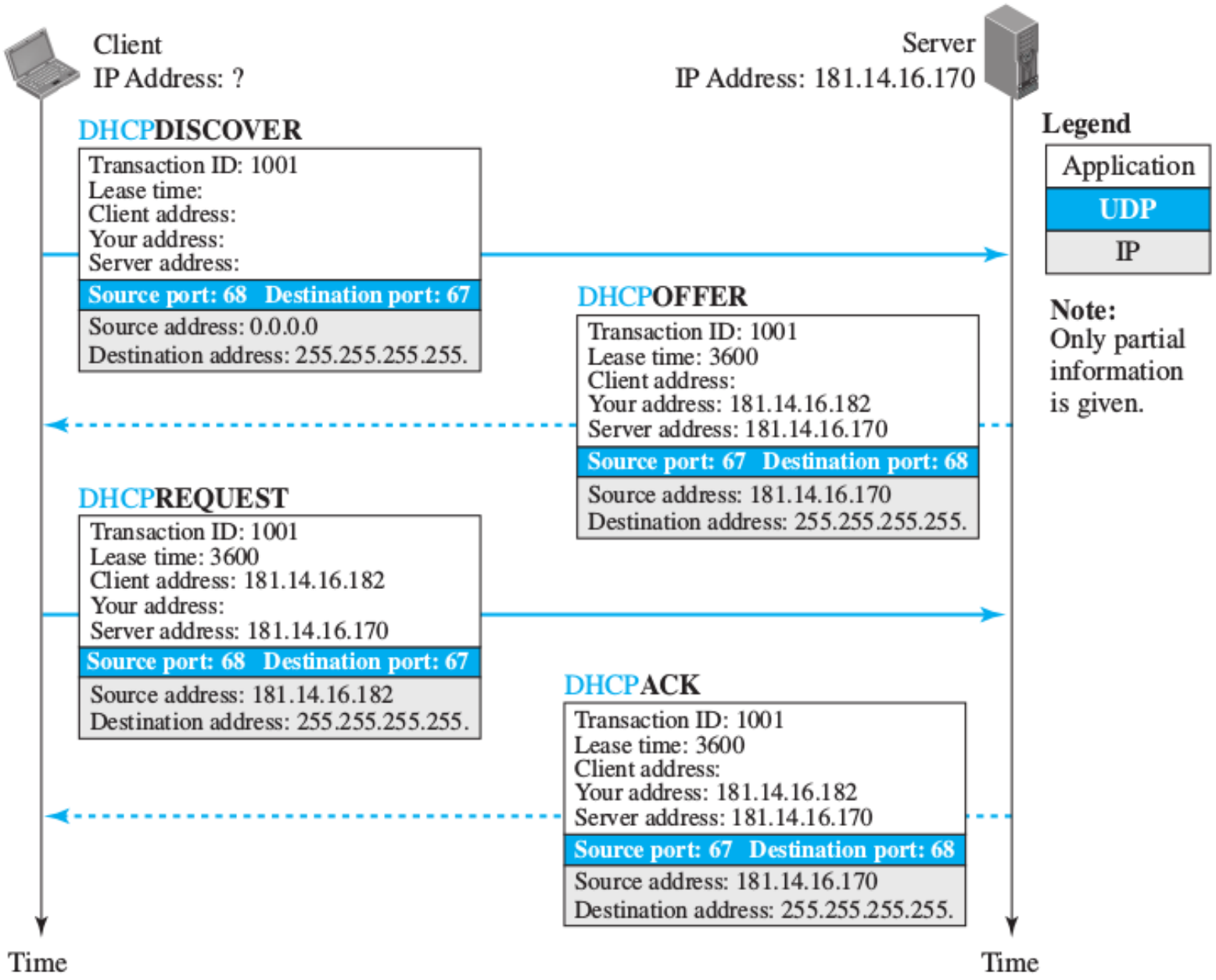
Operation of DHCP



DHCPACK

- Finally, the selected server responds with a DHCPACK message to the client if the offered IP address is valid.
- If the server cannot keep its offer (for example, if the address is offered to another host in between), the server sends a DHCPNACK message and the client needs to repeat the process.
- This message is also broadcast to let other servers know that the request is accepted or rejected.

Operation of DHCP



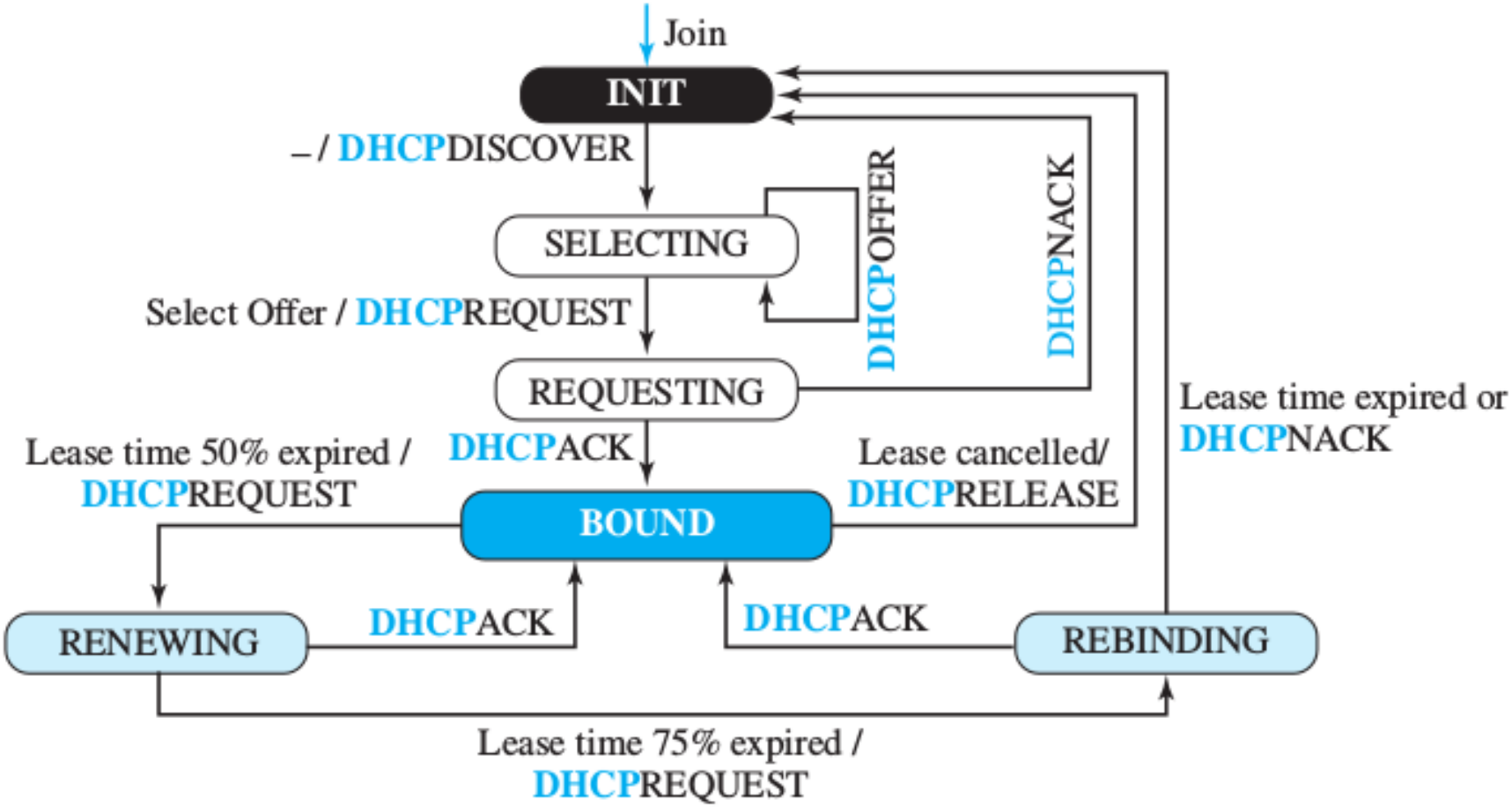
Using FTP

- The server does not send all of the information that a client may need for joining the network.
- In the DHCPACK message, the server defines the pathname of a file in which the client can find complete information such as the address of the DNS server.
- The client can then use a file transfer protocol to obtain the rest of the needed information.

Error Control

- DHCP uses the service of UDP, which is not reliable. To provide error control, DHCP uses two strategies.
- First, DHCP requires that UDP use the checksum.
- Second, the DHCP client uses timers and a retransmission policy if it does not receive the DHCP reply to a request.

FSM for the DHCP client

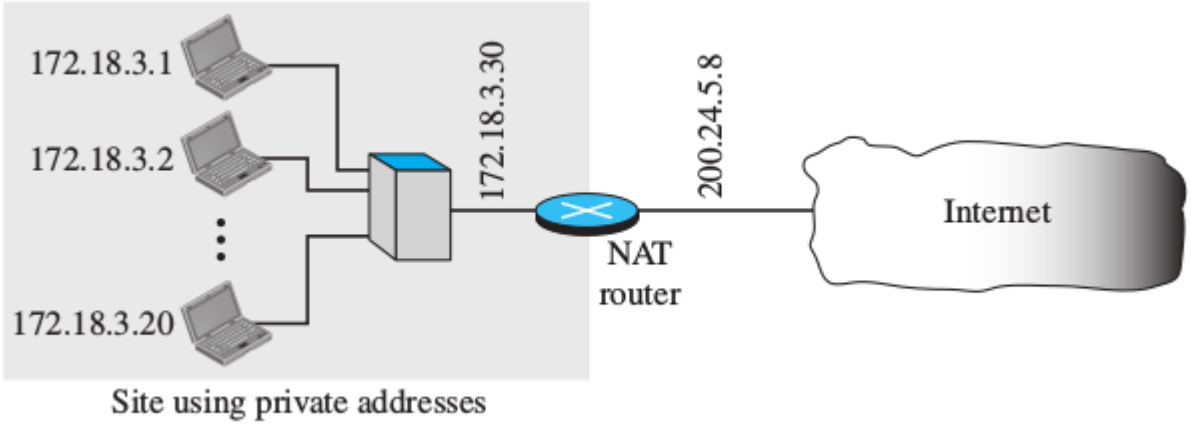


Network Address Resolution (NAT)

- Suppose that an ISP has granted a small range of addresses to a small business
- If the business grows and the ISP may not be able to grant the demand because the addresses before and after the range may have already been allocated to other networks.
- But usually only a portion of computers in a small network need access to the Internet simultaneously
- The number of allocated addresses does not have to match the number of computers in the network.
- So the systems can use private addresses locally and few global addresses

Network Address Resolution (NAT)

- A technology that can provide the mapping between the private and universal addresses is Network Address Translation (NAT).
- The technology allows a site to use a set of private addresses for internal communication and a set of global Internet addresses (at least one) for communication with the rest of the world.
- The site must have only one connection to the global Internet through a NAT-capable router that runs NAT software.

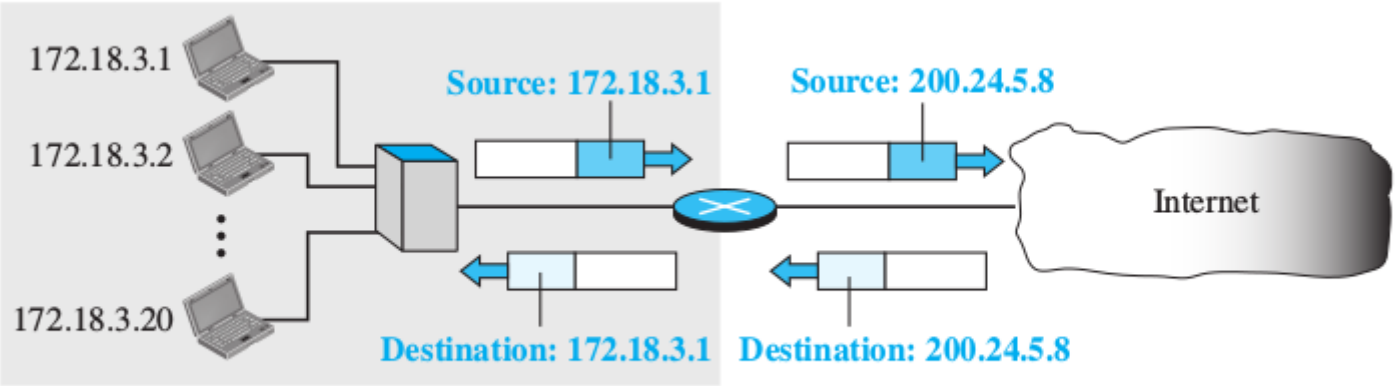


Network Address Resolution (NAT)

- the private network uses private addresses.
- The router that connects the network to the global address uses one private address and one global address.
- The private network is invisible to the rest of the Internet; the rest of the Internet sees only the NAT router with the address 200.24.5.8.

Address Translation

- All of the outgoing packets go through the NAT router, which replaces the source address in the packet with the global NAT address.
- All incoming packets also pass through the NAT router, which replaces the destination address in the packet



Site using private addresses

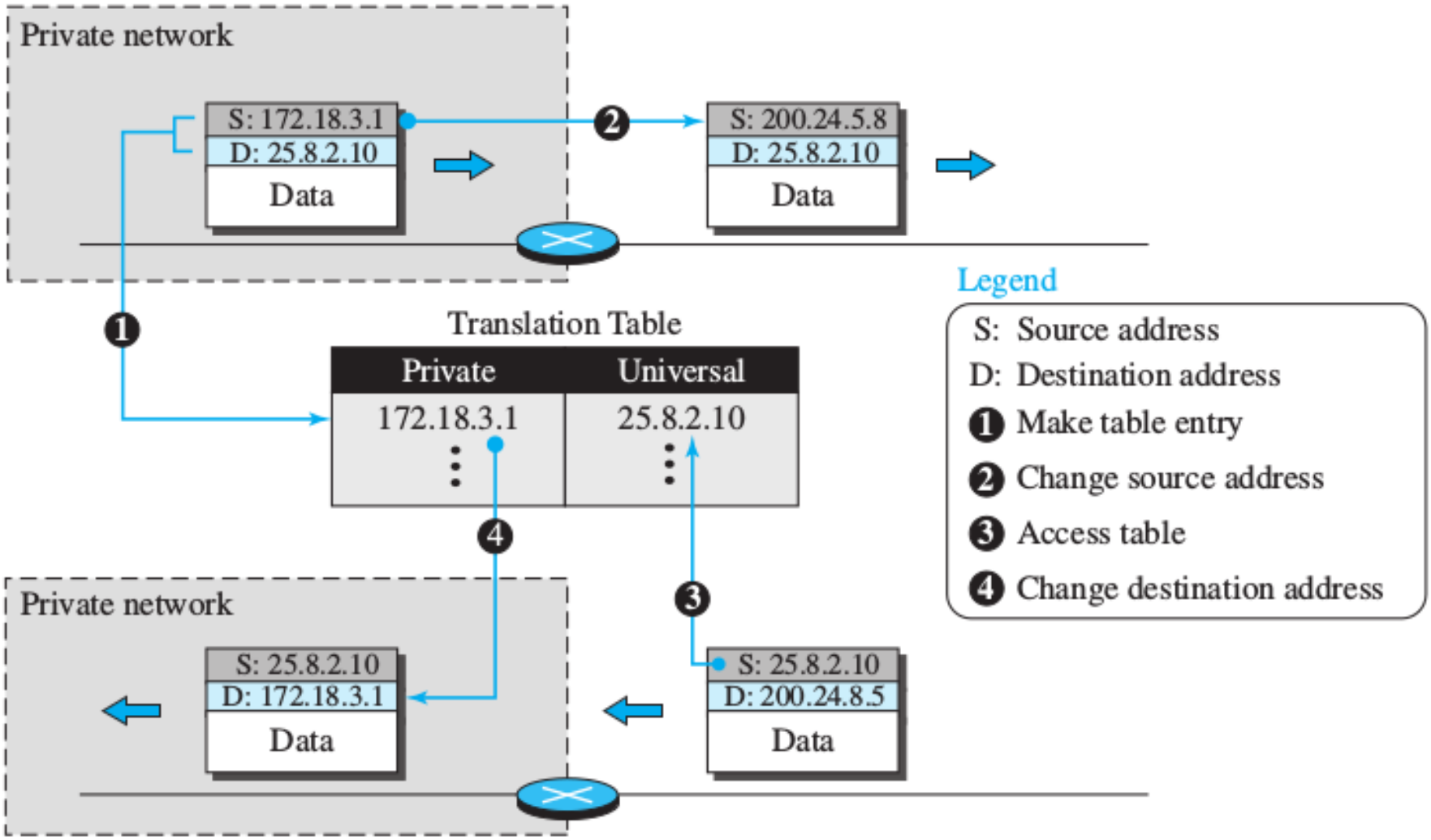
Translation Table

- Translating the source addresses for an outgoing packet is straightforward.
- What is the destination address for a packet coming from the Internet?
- Which private address should it be mapped?
- This is solved if the NAT router has a translation table.

Using One IP Address

- Here a translation table has only two columns: the private address and the external address (destination address of the packet).
- When the router translates the source address of the outgoing packet, it also makes note of the destination address - where the packet is going.
- When the response comes back from the destination, the router uses the source address of the packet (as the external address) to find the private address of the packet.

Translation



Using One IP Address

- In this strategy, communication must always be initiated by the private network.
- The NAT mechanism described requires that the private network start the communication.
- NAT is used mostly by ISPs that assign a single address to a customer.
- The customer, however, may be a member of a private network that has many private addresses.
- In this case, communication with the Internet is always initiated from the customer site, using a client program such as HTTP, TELNET, or FTP to access the corresponding server program.
- when e-mail that originates from outside the network site is received by the ISP e-mail server, it is stored in the mailbox of the customer until retrieved with a protocol such as POP.

Using a Pool of IP Addresses

- The use of only one global address by the NAT router allows only one private-network host to access a given external host.
- To remove this restriction, the NAT router can use a pool of global addresses.
- For example, instead of using only one global address (200.24.5.8), the NAT router can use four addresses (200.24.5.8, 200.24.5.9, 200.24.5.10, and 200.24.5.11)
- four private-network hosts can communicate with the same external host at the same time
- No more than four connections can be made to the same destination.

Using Both IP Addresses and Port Addresses

- To allow a many-to-many relationship between private-network hosts and external server programs, we need more information in the translation table.
- If the translation table has five columns, instead of two, that include the source and destination port addresses and the transport-layer protocol, the ambiguity is eliminated
- the combination of source address (25.8.3.2) and destination port address (1401) defines the private network host to which the response should be directed.
- Private ports should be unique

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
⋮	⋮	⋮	⋮	⋮

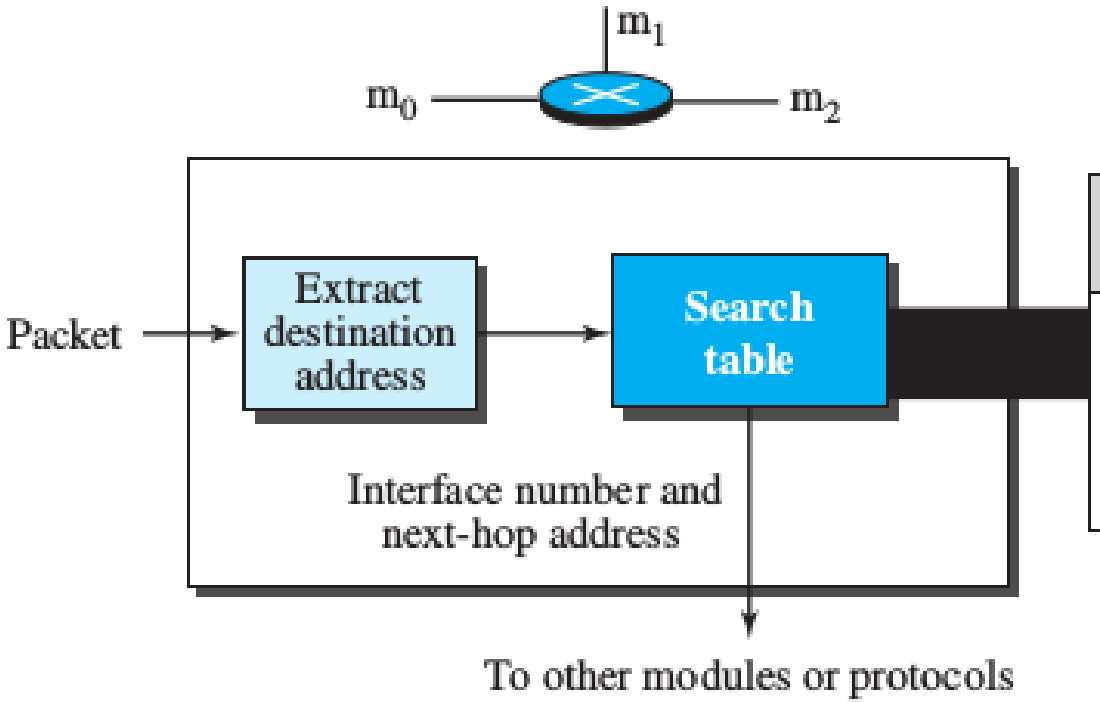
FORWARDING OF IP PACKETS

- Forwarding means to place the packet in its route to its destination
- Internet today is made of a combination of links (networks), forwarding means to deliver the packet to the next hop.
- IP protocol was originally designed as a connectionless protocol, today the tendency is to change it to a connection-oriented protocol.
- connectionless protocol
  - forwarding is based on the destination address of the IP datagram.
- connection-oriented protocol
  - forwarding is based on the label attached to an IP datagram.

Forwarding Based on Destination Address

- Forwarding requires a host or a router to have a forwarding table.
- During Forwarding it looks at this table to find the next hop to deliver the packet to.
- A classless forwarding table needs to include four pieces of information: the mask, the network address, the interface number, and the IP address of the next router.
- The job of the forwarding module is to search the table, row by row.

Simplified forwarding module in classless address

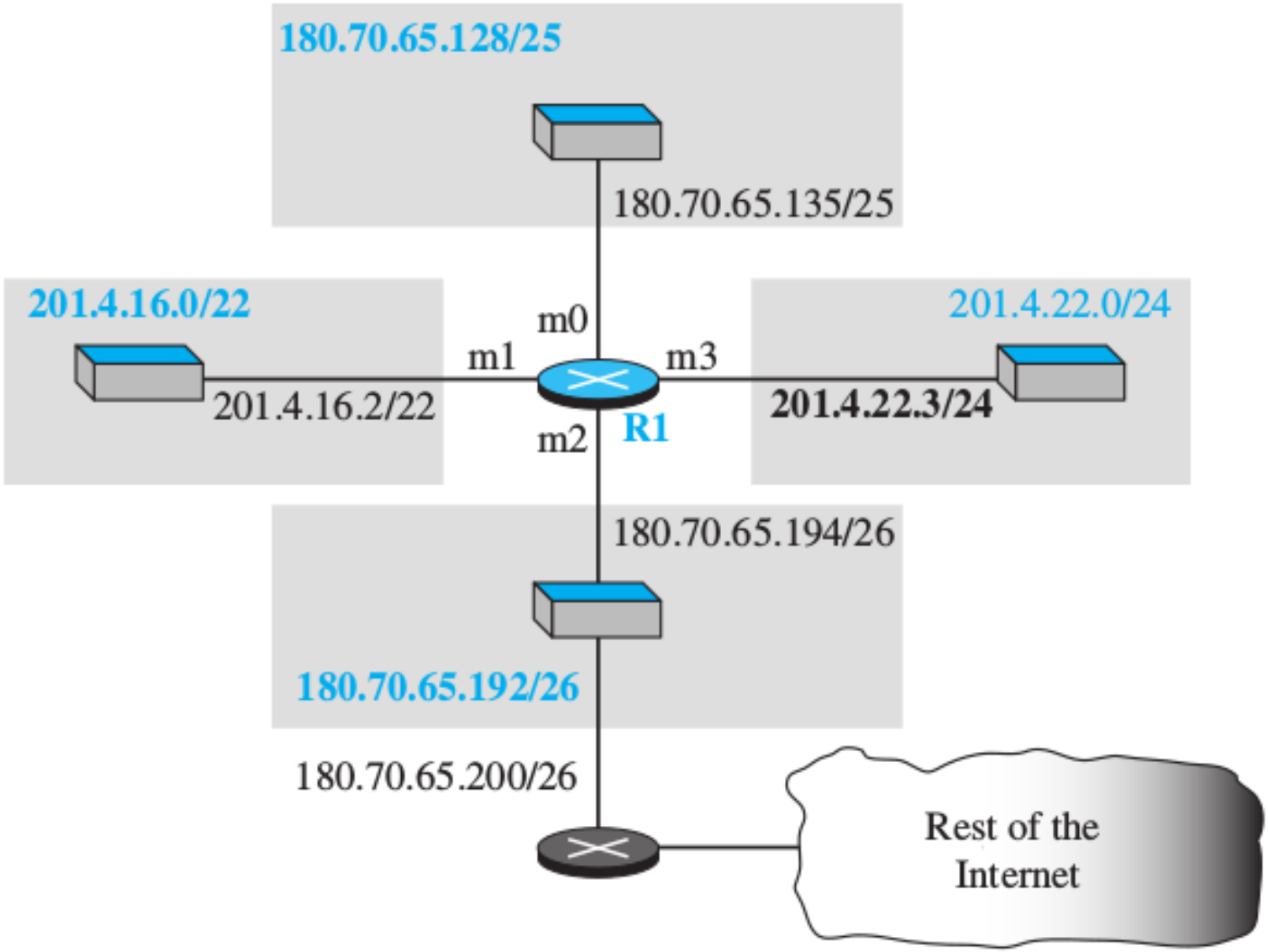


Forwarding table

Network address including mask	Next-hop IP address	Interface
$x_0.y_0.z_0.t_0/n_0$	.....	m_0
$x_1.y_1.z_1.t_1/n_1$	.....	m_1
$x_2.y_2.z_2.t_2/n_1$	.....	m_2

- In each row, the **n** leftmost bits of the destination address (prefix) are kept and the rest of the bits (suffix) are set to 0s.
- If the resulting address (which we call the network address), matches with the address in the first column, the information in the next two columns is extracted; otherwise the search continues.
- Normally, the last row has a default value in the first column (not shown in the figure), which indicates all destination addresses that did not match the previous rows.

Example


180.70.65.128/25
180.70.65.135/25
201.4.16.0/22
m0
m1
m3
201.4.22.0/24
201.4.16.2/22
201.4.22.3/24
m2
R1
180.70.65.194/26
180.70.65.192/26
180.70.65.200/26
Rest of the
Internet

Forwarding table for router R1 using prefix bits

Table 18.2 *Forwarding table for router R1 in Figure 18.33*

Network address/mask	*Next hop*	*Interface*
180.70.65.192/26	—	m2
180.70.65.128/25	—	m0
201.4.22.0/24	—	m3
201.4.16.0/22	—	m1
Default	180.70.65.200	m2

Leftmost bits in the destination address	*Next hop*	*Interface*
`10110100 01000110 01000001 11`	—	m2
`10110100 01000110 01000001 1`	—	m0
`11001001 00000100 00011100`	—	m3
`11001001 00000100 000100`	—	m1
Default	180.70.65.200	m2

Exercise

Show the forwarding process if a packet arrives at R1 in Figure 18.33 with the destination address 180.70.65.140.

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.

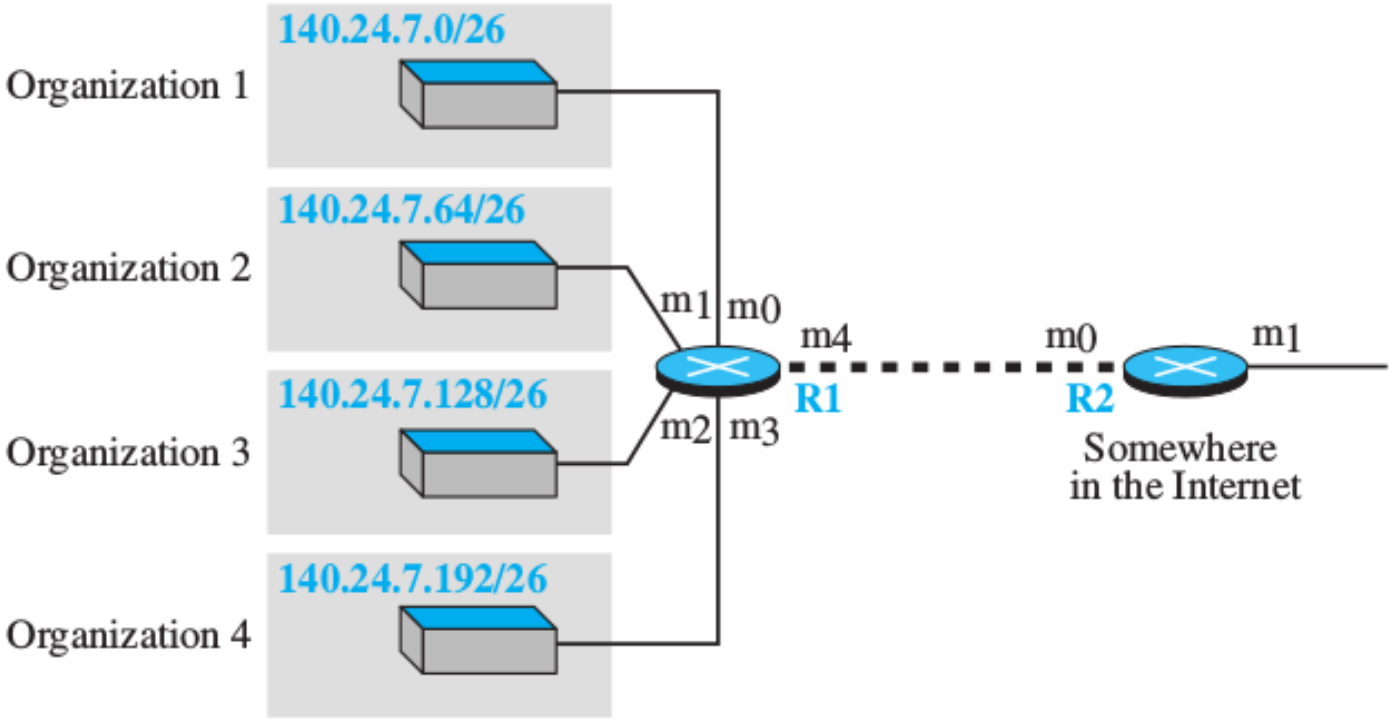
2. The second mask (/25) is applied to the destination address. The result is 180.70.65.128, which matches the corresponding network address.

3.The next-hop address and the interface number m0 are extracted for forwarding the packet.

Address Aggregation

- When we use classful addressing, there is only one entry in the forwarding table for each site outside the organization.
- When we use classless addressing, it is likely that the number of forwarding table entries will increase.
- This is because the intent of classless addressing is to divide up the whole address space into manageable blocks.
- The increased size of the table results in an increase in the amount of time needed to search the table.
- To alleviate the problem, the idea of address aggregation was designed.

Address Aggregation



Forwarding table for R1

Network address/mask	Next-hop address	Interface
140.24.7.0/26	----------	m0
140.24.7.64/26	----------	m1
140.24.7.128/26	----------	m2
140.24.7.192/26	----------	m3
0.0.0.0/0	address of R2	m4

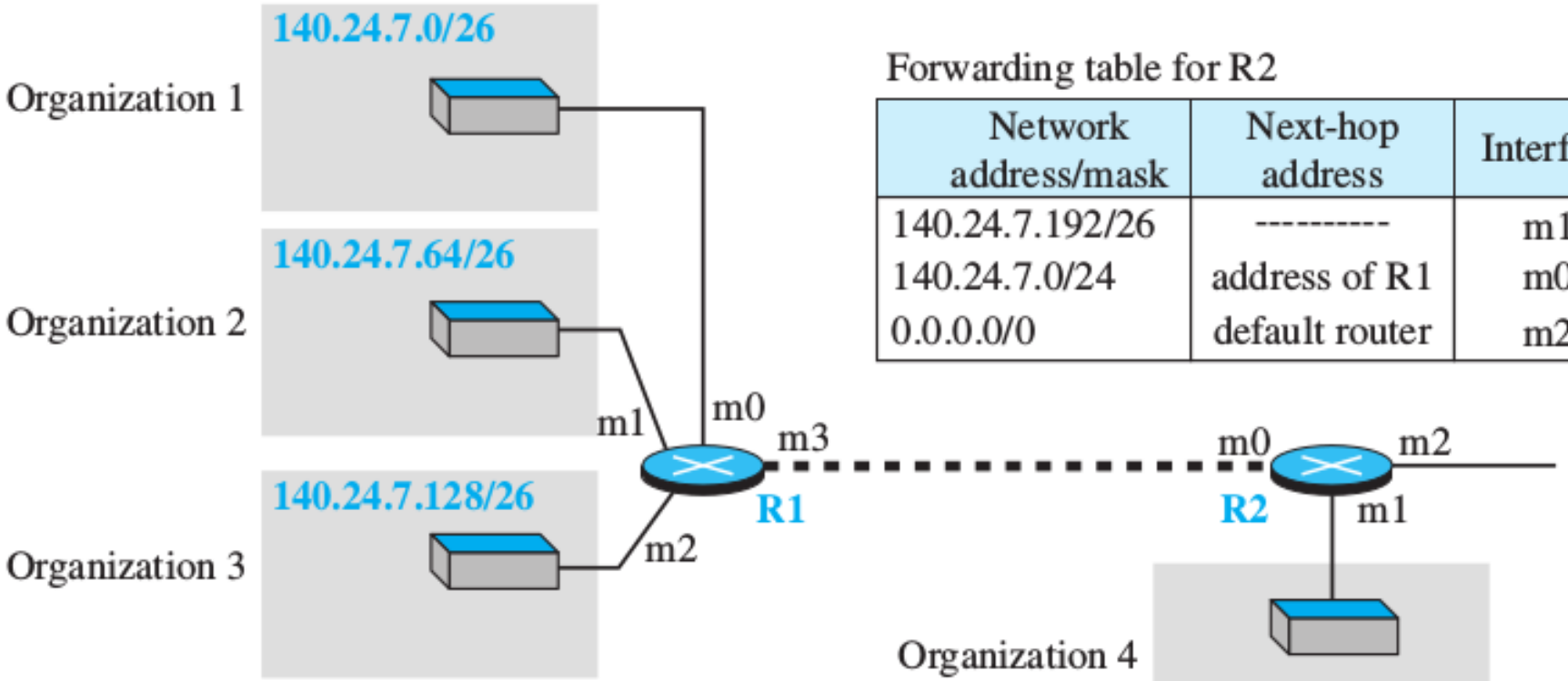
Forwarding table for R2

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

Address Aggregation

- R1 is connected to networks of four organizations that each use 64 addresses. R2 is somewhere far from R1.
- R1 has a longer forwarding table because each packet must be correctly routed to the appropriate organization.
- R2, on the other hand, can have a very small forwarding table.
- For R2, any packet with destination 140.24.7.0 to 140.24.7.255 is sent out from interface m0 regardless of the organization number.
- This is called address aggregation because the blocks of addresses for four organizations are aggregated into one larger block.
- R2 would have a longer forwarding table if each organization had addresses that could not be aggregated into one block.

Longest Mask Matching



Forwarding table for R2

Network address/mask	Next-hop address	Interface
140.24.7.192/26	----------	m1
140.24.7.0/24	address of R1	m0
0.0.0.0/0	default router	m2

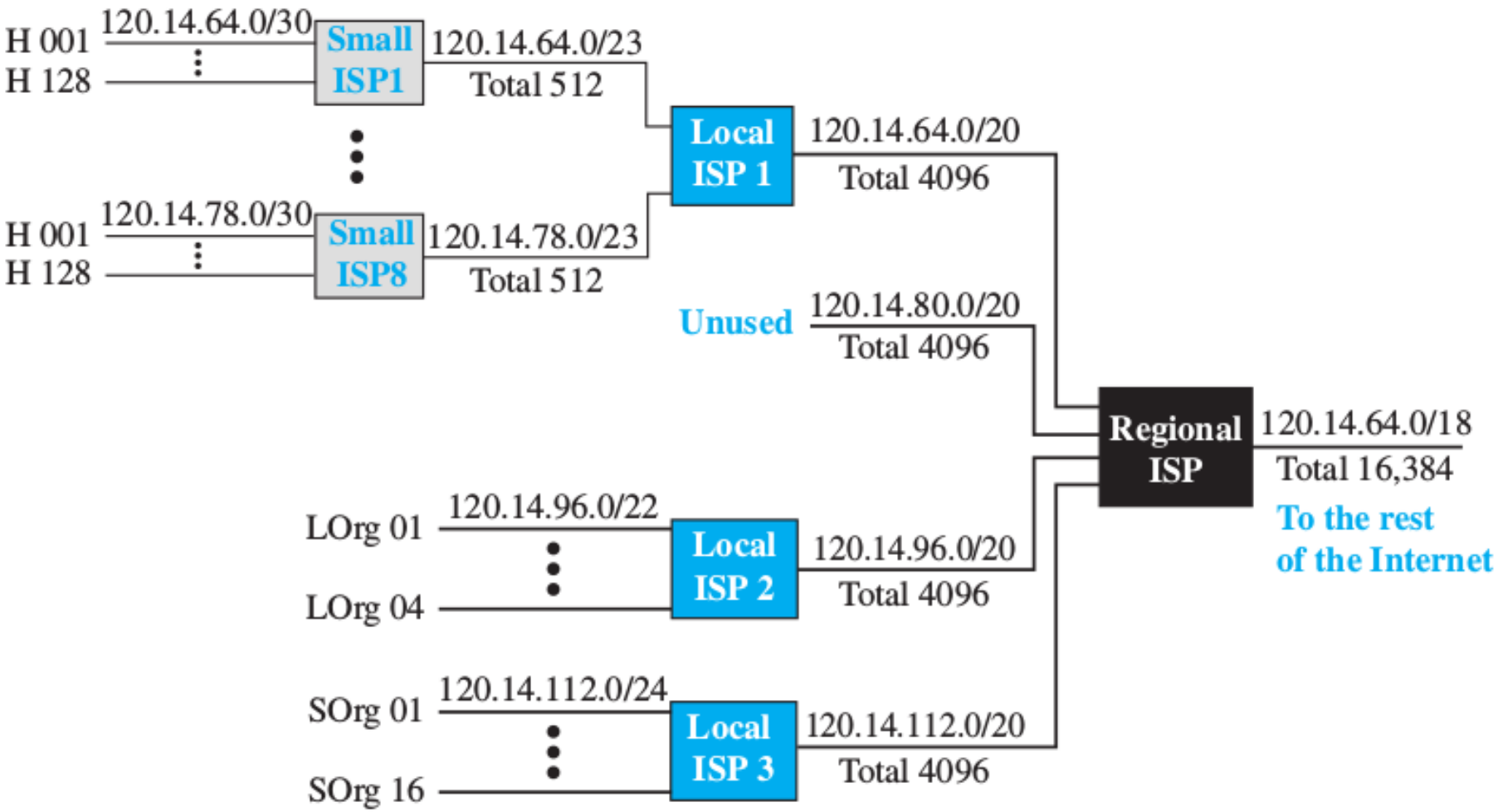
Forwarding table for R1

Network address/mask	Next-hop address	Interface
140.24.7.0/26	----------	m0
140.24.7.64/26	----------	m1
140.24.7.128/26	----------	m2
0.0.0.0/0	default router	m3

Longest Mask Matching

- This principle states that the forwarding table is sorted from the longest mask to the shortest mask.
- Suppose a packet arrives at router R2 for organization 4 with destination address 140.24.7.200.
- The first mask at router R2 is applied, which gives the network address 140.24.7.192.
- The packet is routed correctly from interface m1 and reaches organization 4.
- This would not have happened if the forwarding table is not sorted.

Hierarchical Routing & Geographical Routing



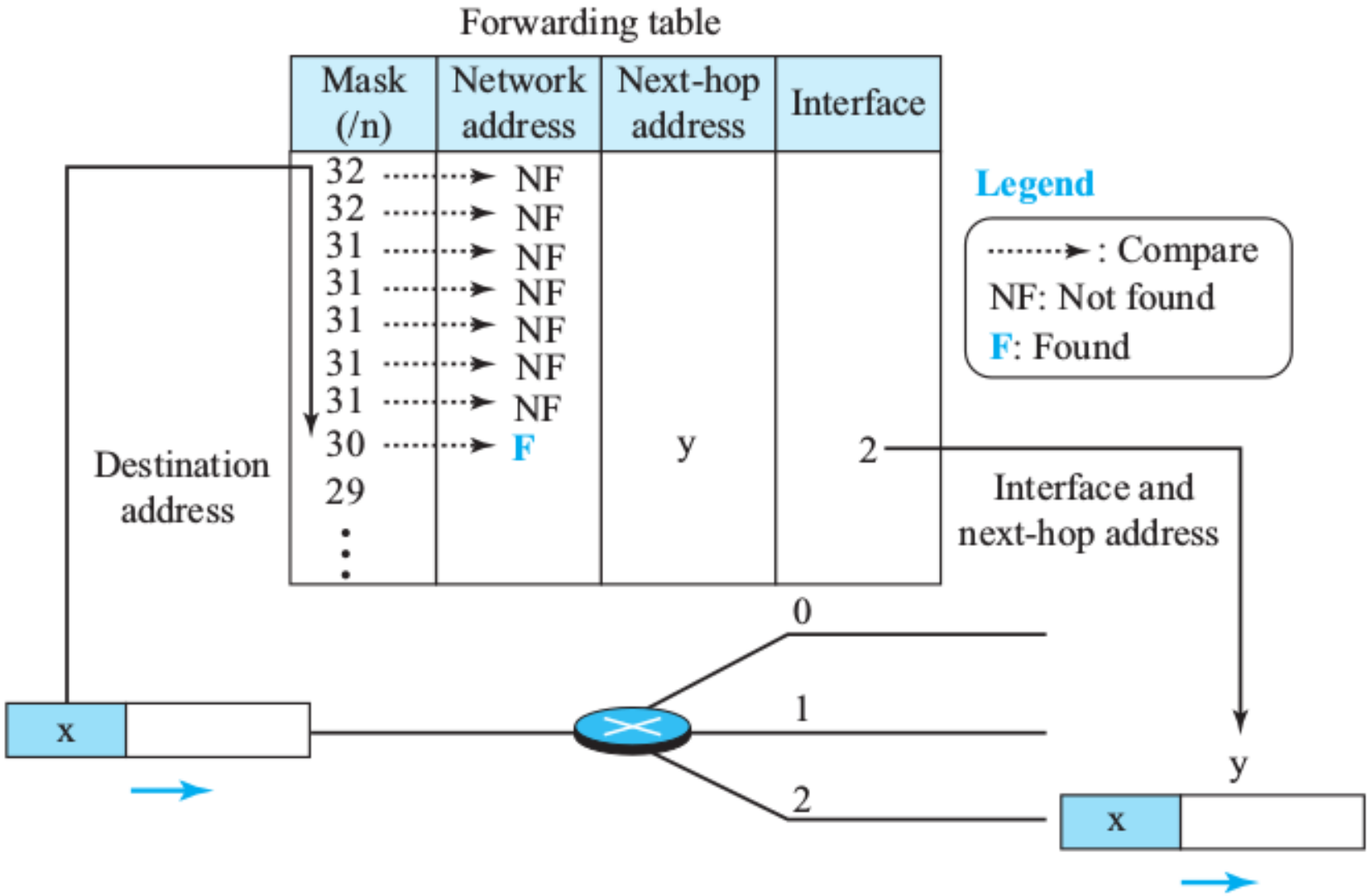
Forwarding Table Search Algorithms

- Search by the longest prefix match
- Takes time
- Choose a different data structure tree, binary tree, trie etc..

Forwarding Based on Label

- Change IP to behave like a connection-oriented protocol in which the routing is replaced by switching
- In a connection-less network (datagram approach), a router forwards a packet based on the destination address in the header of the packet.
- In a connection-oriented network (virtual-circuit approach), a switch (not a router) forwards a packet based on the label attached to the packet.
- Routing is normally based on searching the contents of a table; switching can be done by accessing a table using an index.
- In other words, routing involves searching; switching involves accessing.

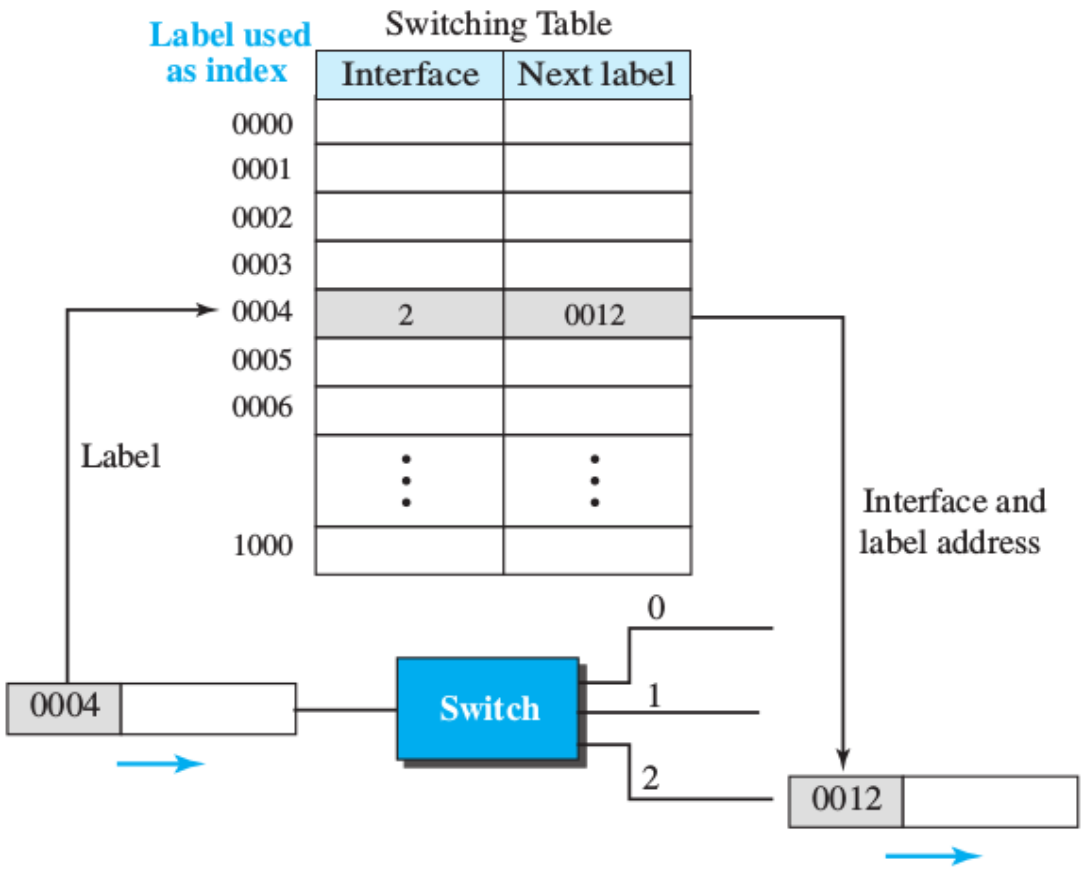
Forwarding based on destination address



Forwarding based on destination address

- When the forwarding algorithm gets the destination address of the packet, it needs to delve into the mask column.
- For each entry, it needs to apply the mask to find the destination network address.
- It then needs to check the network addresses in the table until it finds the match.
- The router then extracts the next-hop address and the interface number to be delivered to the data-link layer.

Forwarding based on label



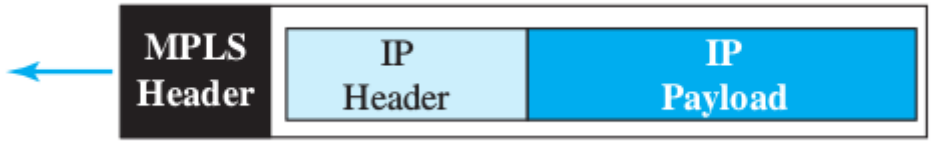
- This is a simple example of using a label to access a switching table.
- Since the labels are used as the index to the table, finding the information in the table is immediate.

Multi-Protocol Label Switching (MPLS)

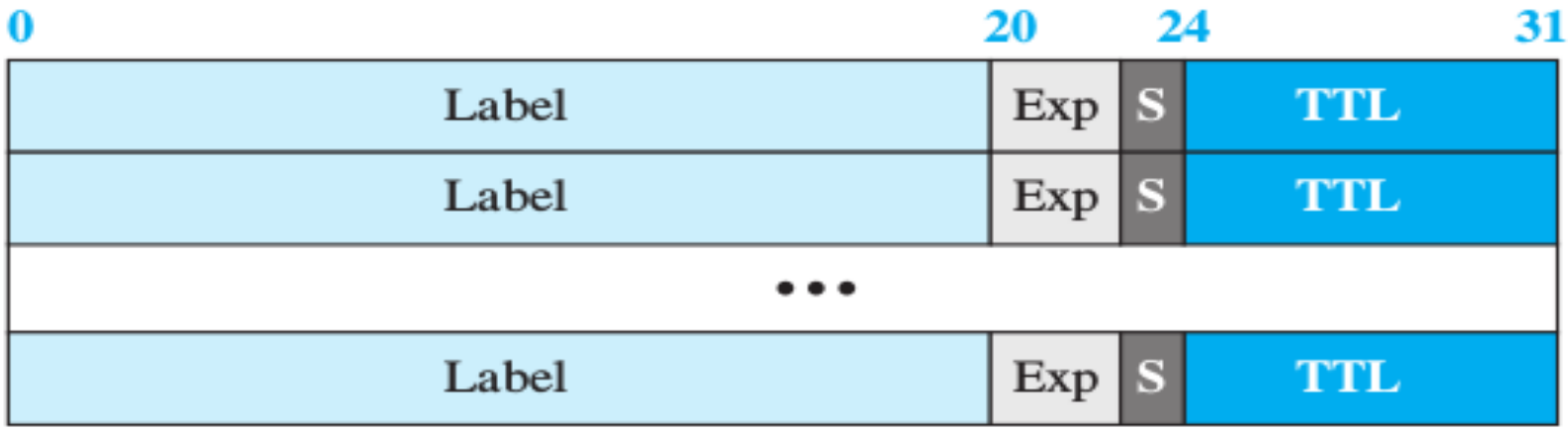
- During the 1980s, several vendors created routers that implement switching technology.
- Later IETF approved a standard that is called Multi-Protocol Label Switching.
- In this standard, some conventional routers in the Internet can be replaced by MPLS routers, which can behave like a router and a switch.
- When behaving like a router, MPLS can forward the packet based on the destination address; when behaving like a switch, it can forward a packet based on the label.

A New Header

- To simulate connection-oriented switching using a protocol like IP, the first thing that is needed is to add a field to the packet that carries the label.
- The IPv4 packet format does not allow this extension
- The solution is to encapsulate the IPv4 packet in an MPLS packet
- The whole IP packet is encapsulated as the payload in an MPLS packet and an MPLS header is added.



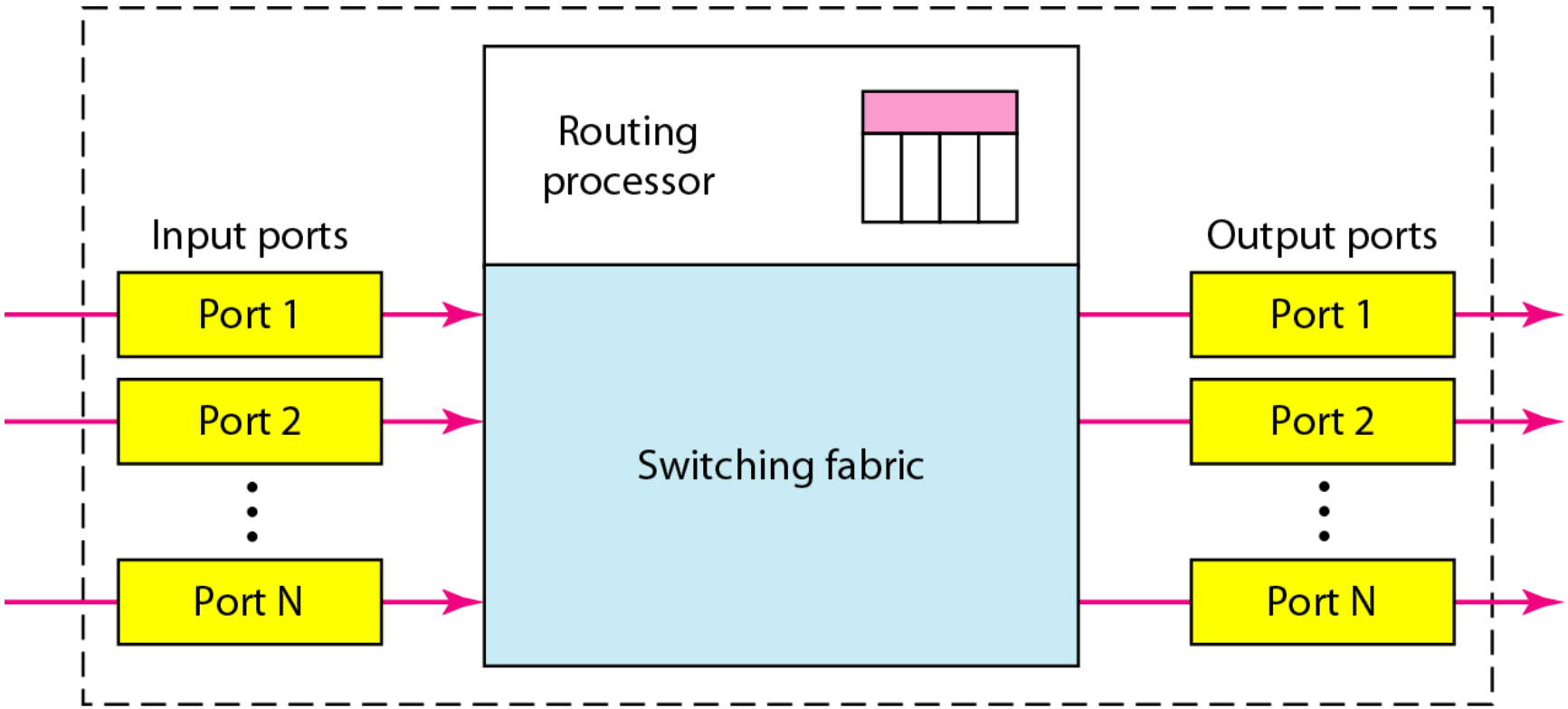
MPLS header made of a stack of labels



- **Label**. This 20-bit field defines the label that is used to index the forwarding table in the router.
- **Exp**. This 3-bit field is reserved for experimental purposes.
- **S**. The one-bit stack field defines the situation of the subheader in the stack. When the bit is 1, it means that the header is the last one in the stack.
- **TTL**. This 8-bit field is similar to the TTL field in the IP datagram. Each visited router decrements the value of this field. When it reaches zero, the packet is discarded to prevent looping.

Routers as Packet Switches

- As we may have guessed by now, the packet switches that are used in the network layer are called routers.
- Routers can be configured to act as either a datagram switch or a virtual-circuit switch.
- The structure of a Packet-Switch, explained earlier, is shown below



END