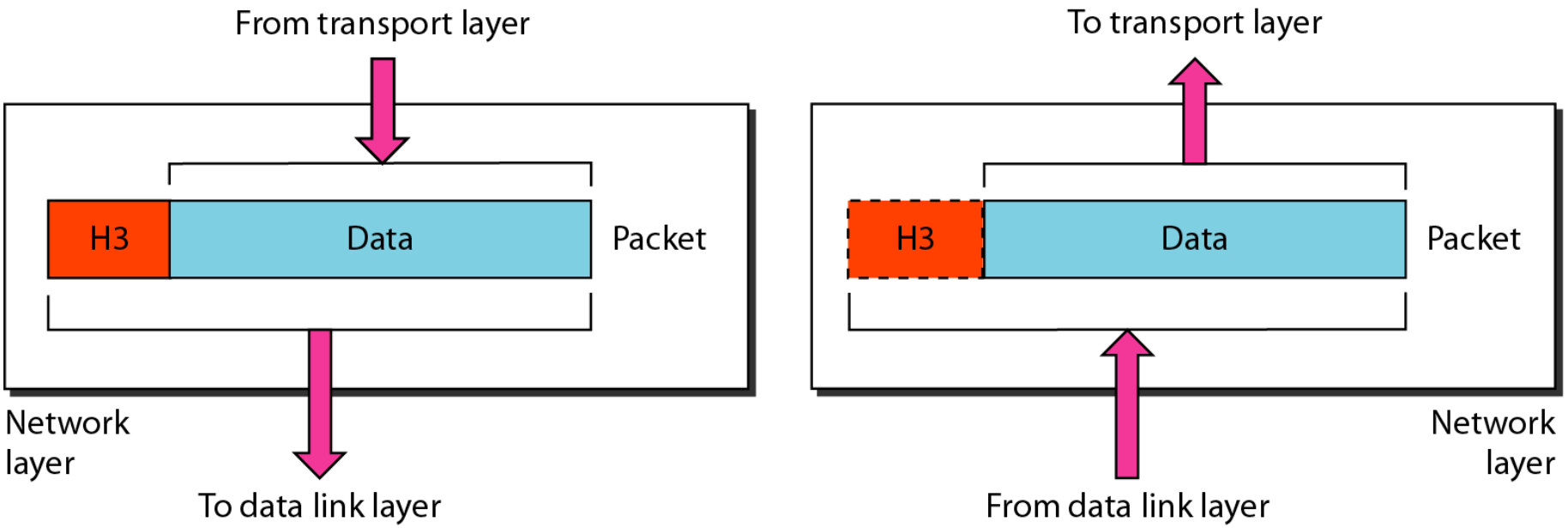
Unit- 4

Network Layer

Mr. Atul Pawar , PCCOE, Pune

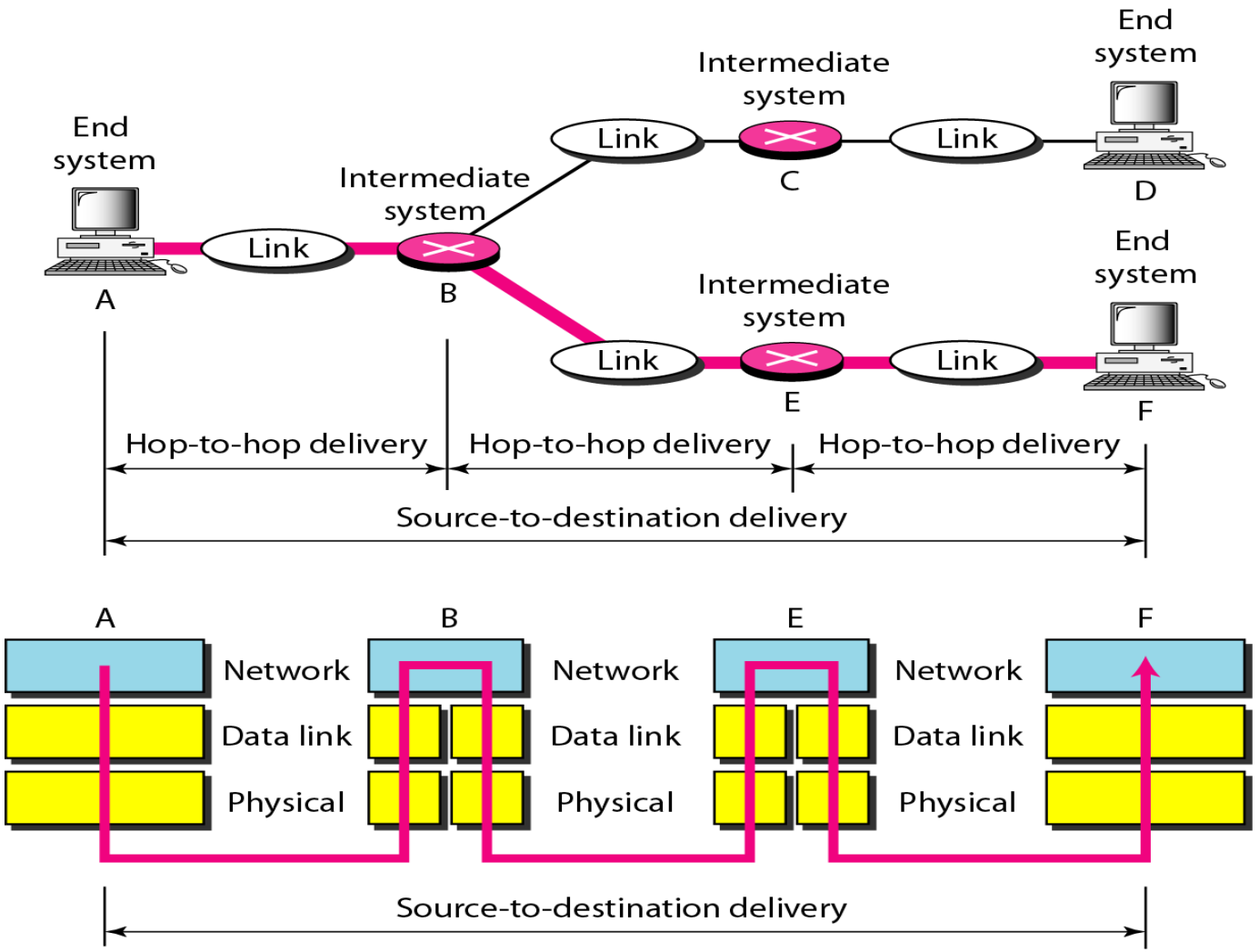
Network layer



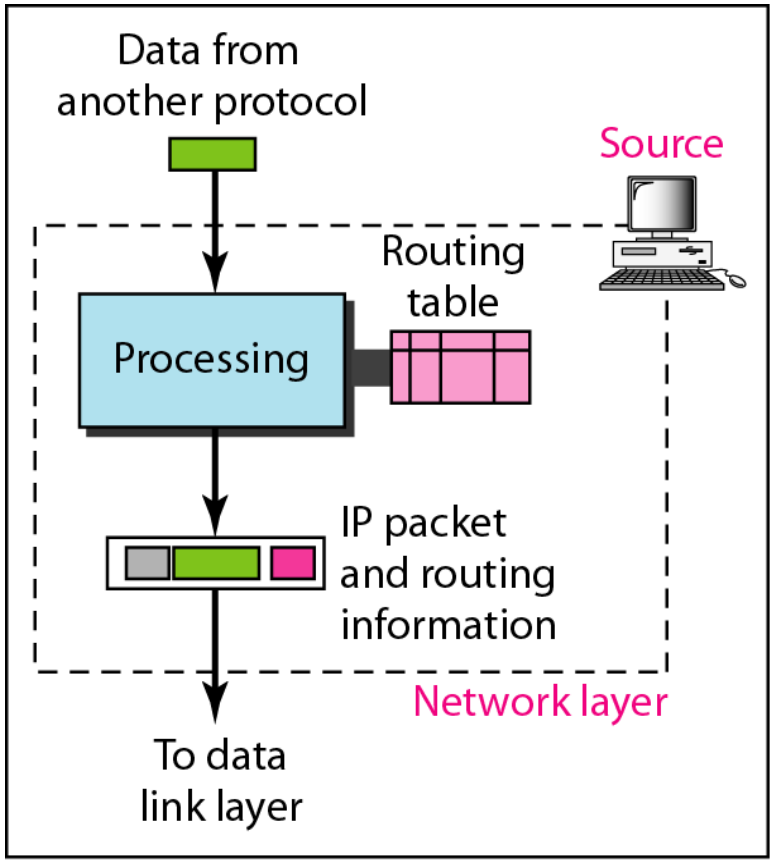
Note

The network layer is responsible for the delivery of individual packets from the source host to the destination host.

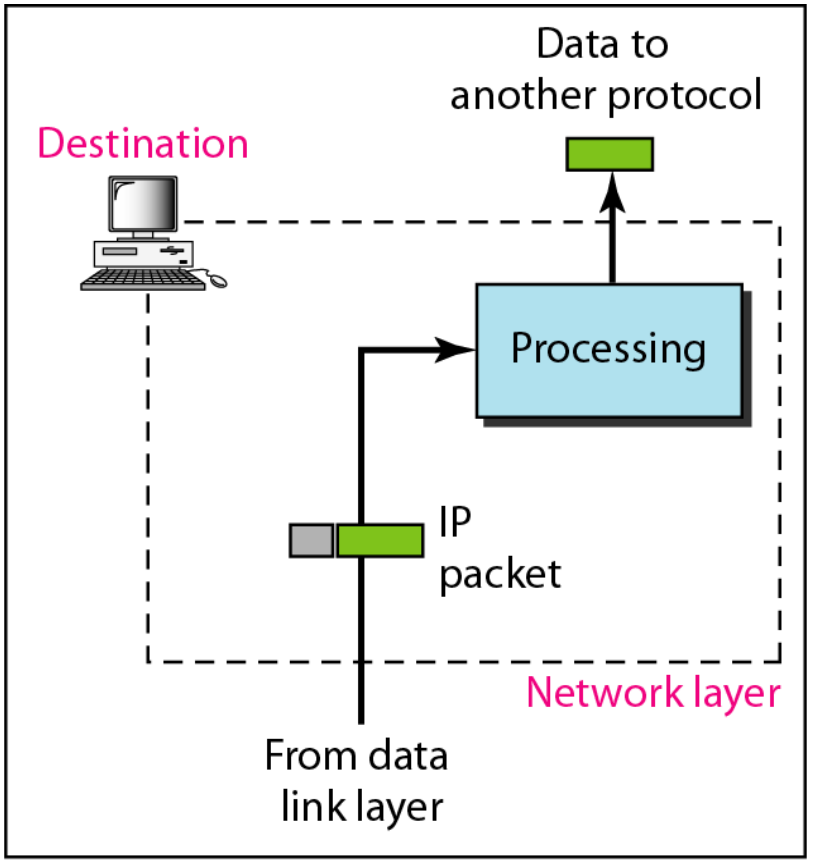
Source to Destination delivery



Network Layer at the Source, router and destination.

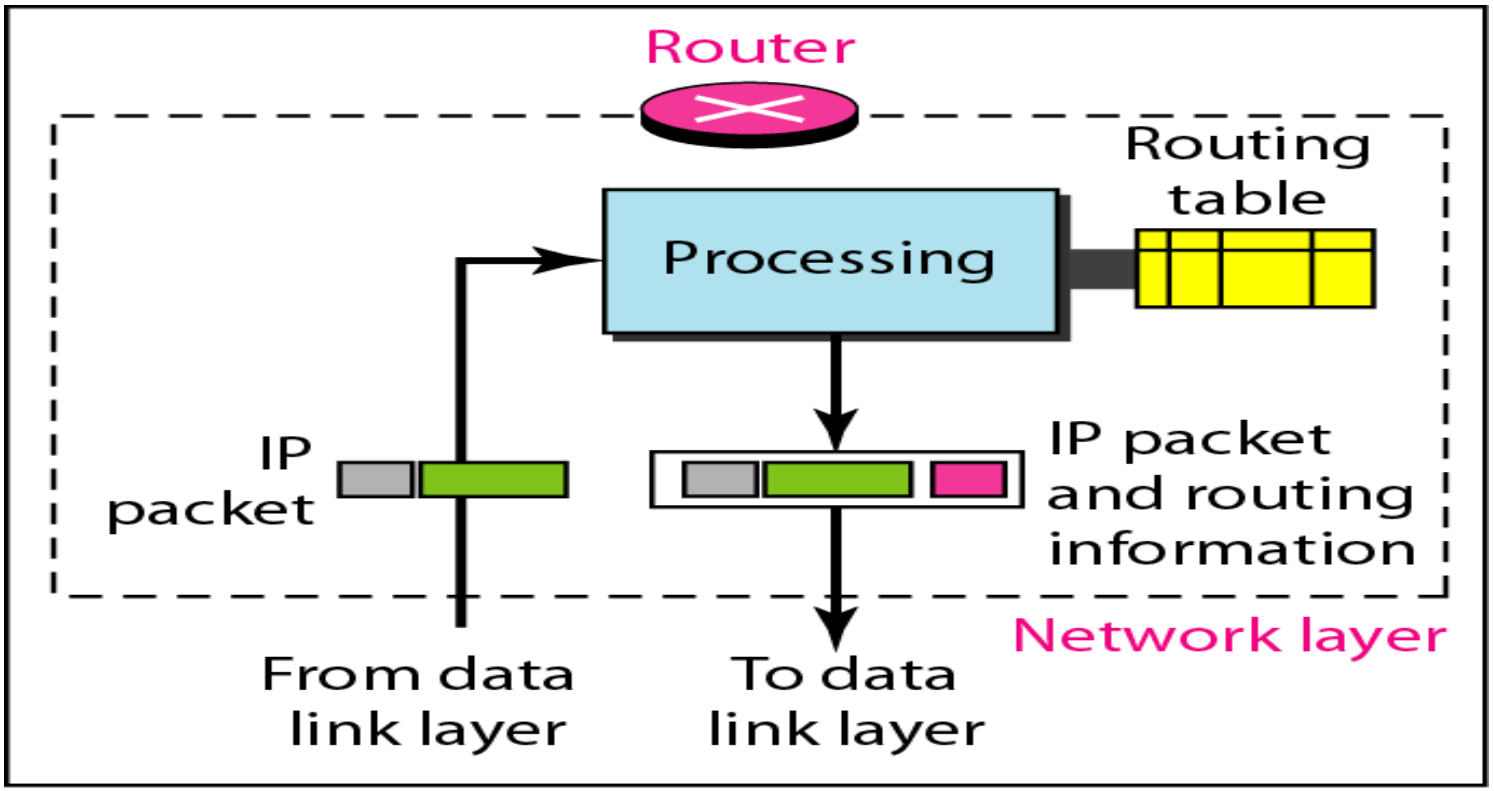


a. Network layer at source



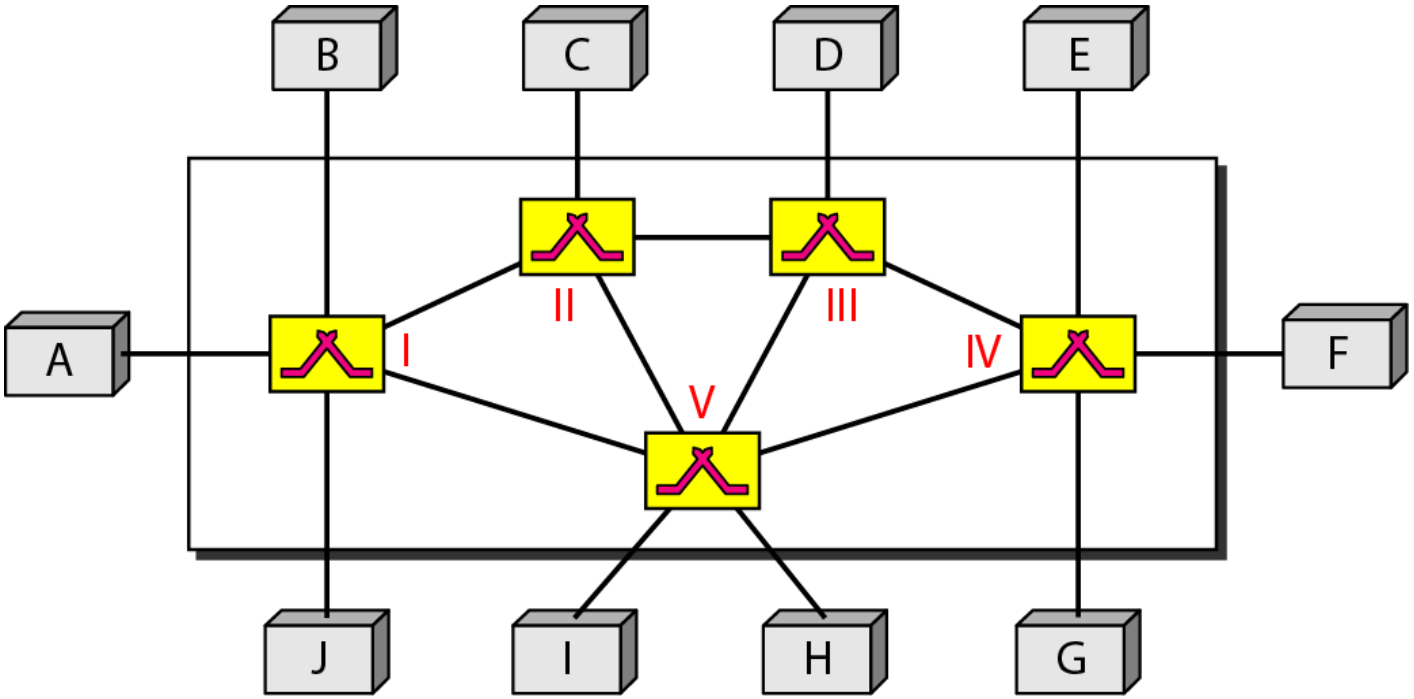
b. Network layer at destination

Network layer at the source, router, and destination (continued)

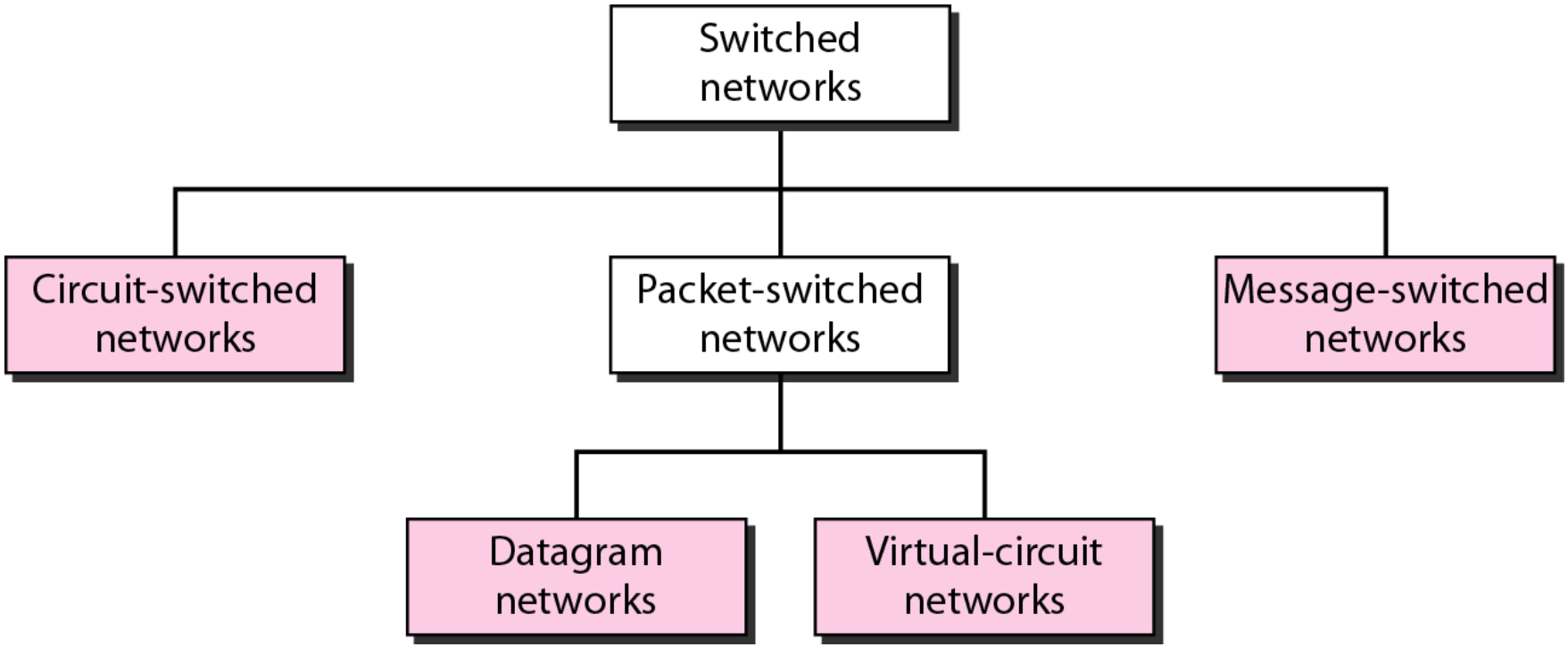


c. Network layer at a router

Switched network



Taxonomy of switched networks



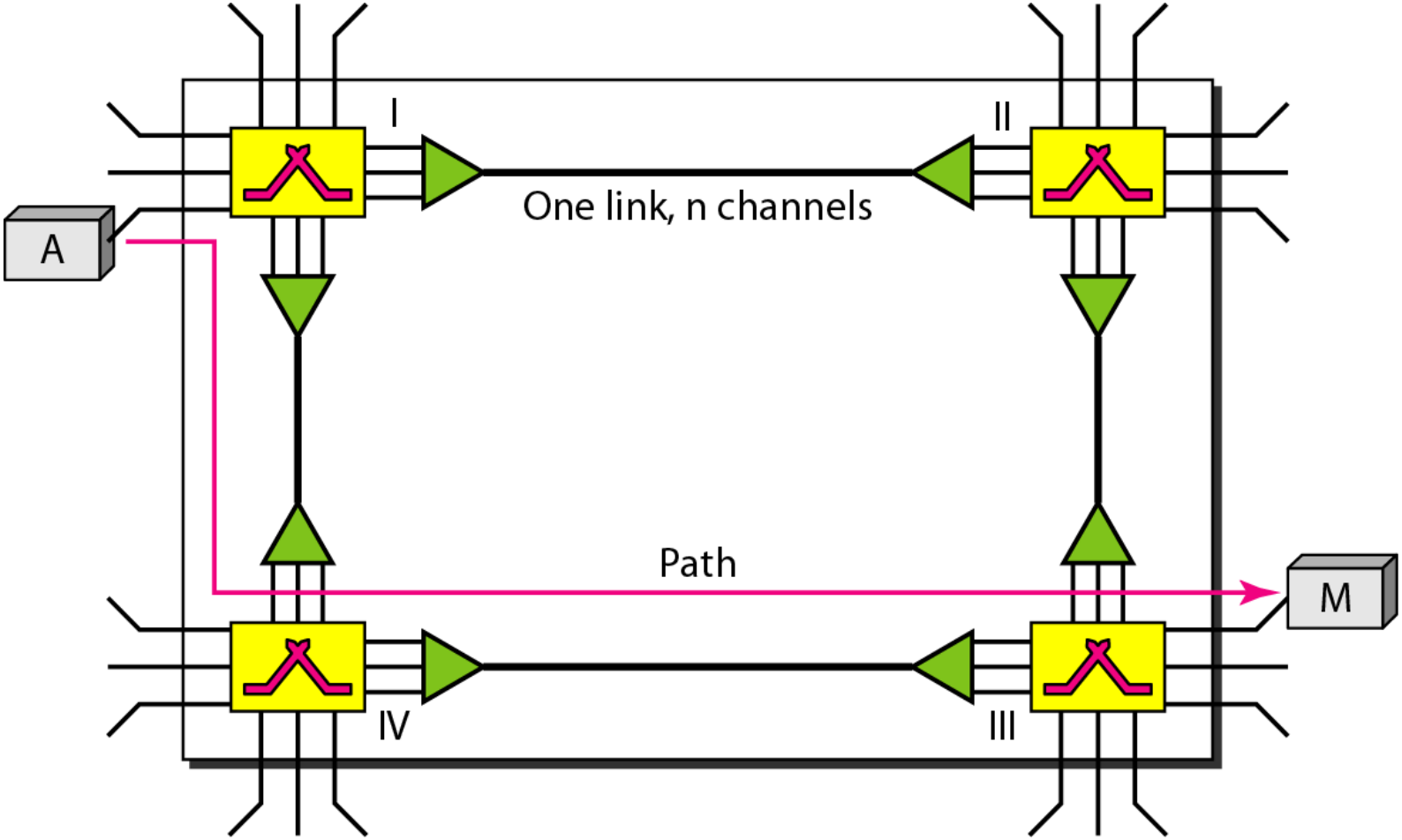
CIRCUIT-SWITCHED NETWORKS

A circuit-switched network consists of a set of switches connected by physical links. A connection between two stations is a dedicated path made of one or more links. However, each connection uses only one dedicated channel on each link. Each link is normally divided into n channels by using FDM or TDM.

Note

A circuit-switched network is made of a set of switches connected by physical links, in which each link is divided into *n* channels.

A trivial circuit-switched network



Circuit switching

1. Takes place at physical layer

2. Stations must make a reservation for the resources to be used during communication.

3.The data are a continuous flow, not packetized.

4. No addressing involved during data transfer.

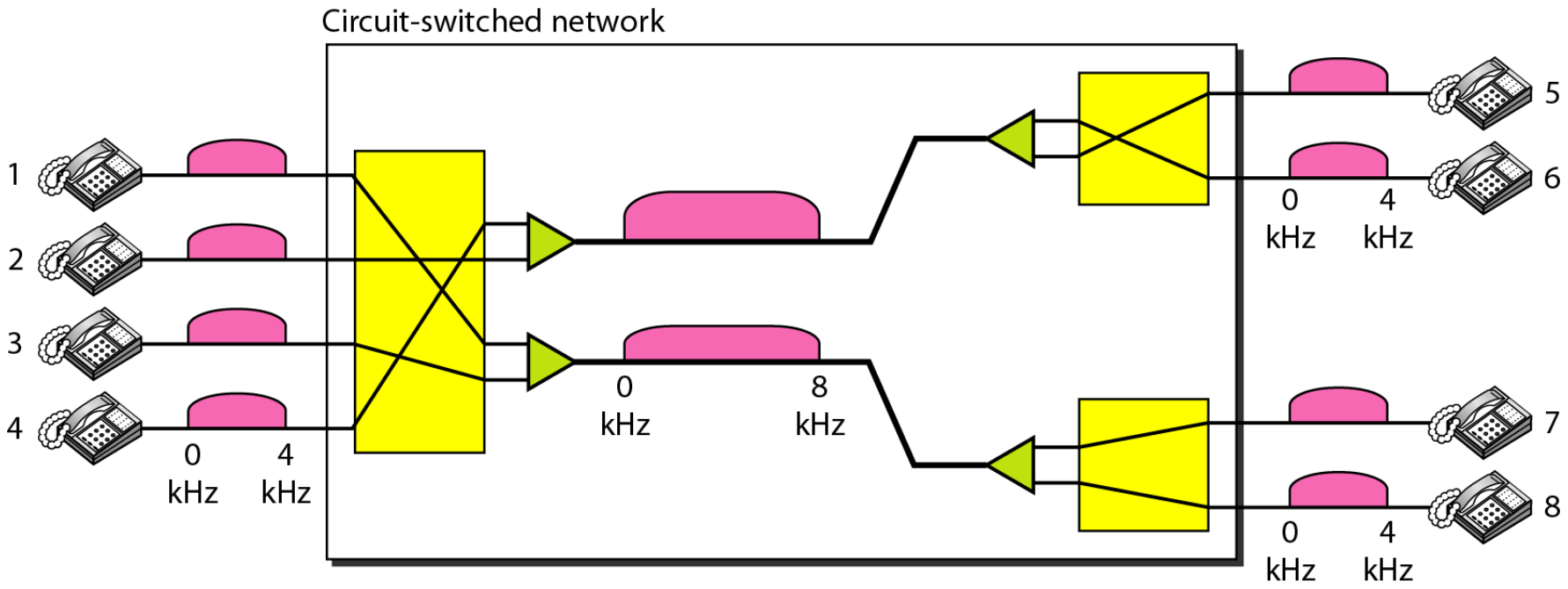
Note

In circuit switching, the resources need to be reserved during the setup phase; the resources remain dedicated for the entire duration of data transfer until the teardown phase.

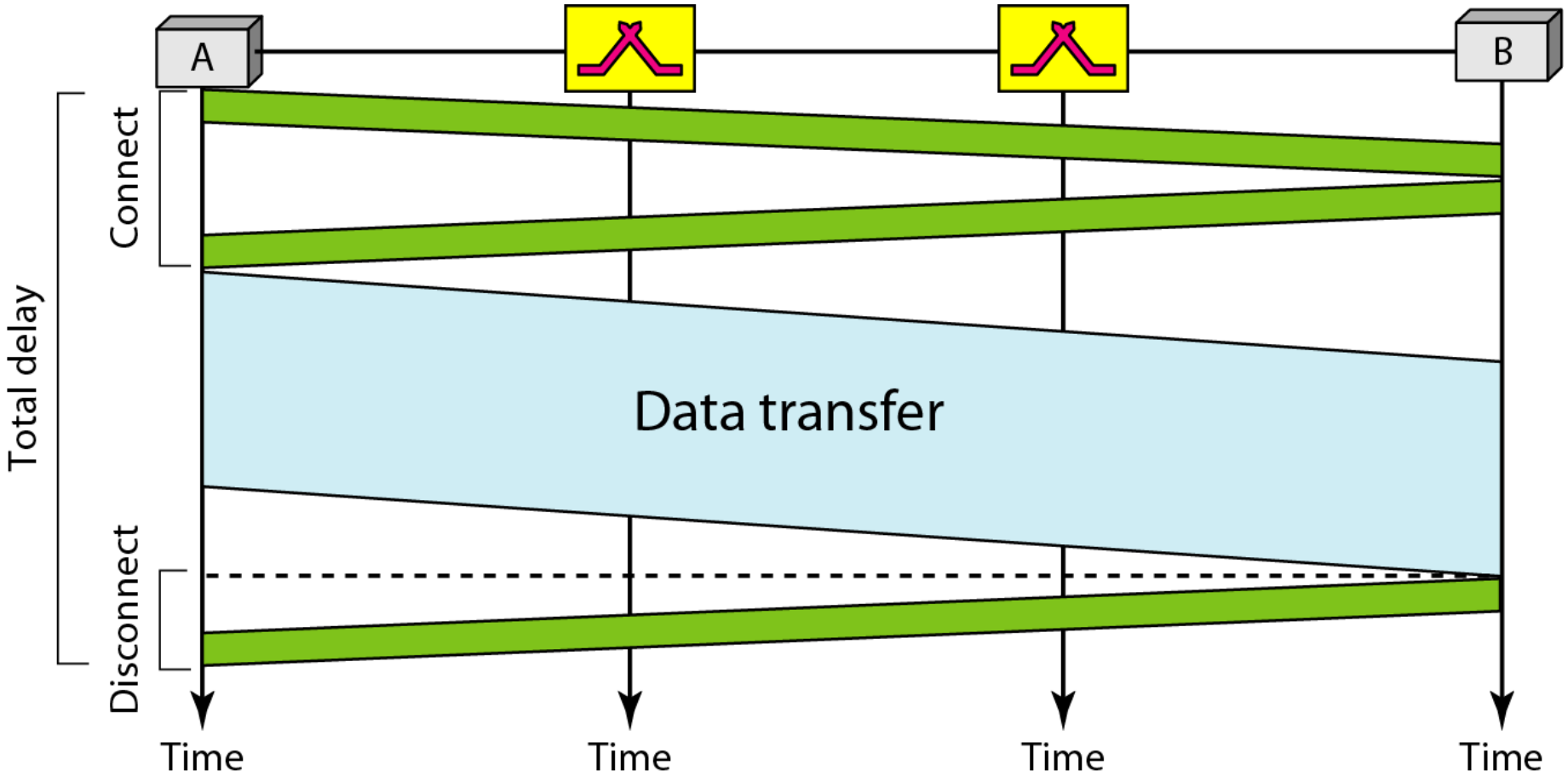
Example

As a trivial example, let us use a circuit-switched network to connect eight telephones in a small area. Communication is through 4-kHz voice channels. We assume that each link uses FDM to connect a maximum of two voice channels. The bandwidth of each link is then 8 kHz. Figure 8.4 shows the situation. Telephone 1 is connected to telephone 7; 2 to 5; 3 to 8; and 4 to 6. Of course the situation may change when new connections are made. The switch controls the connections.

Circuit-switched network used in Example



Delay in a circuit-switched network



Note

Switching at the physical layer in the traditional telephone network uses the circuit-switching approach.

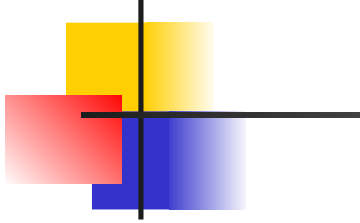
Circuit switching

- ***Advantages:***
  - The communication channel (once established) is dedicated.

- ***Disadvantages:***
  - Possible long wait to establish a connection, (10 seconds, more on long- distance or international calls.) during which no data can be transmitted.
  - More expensive than any other switching techniques, because a dedicated path is required for each connection.
  - Inefficient use of the communication channel, because the channel is not used when the connected systems are not using it.

DATAGRAM NETWORKS

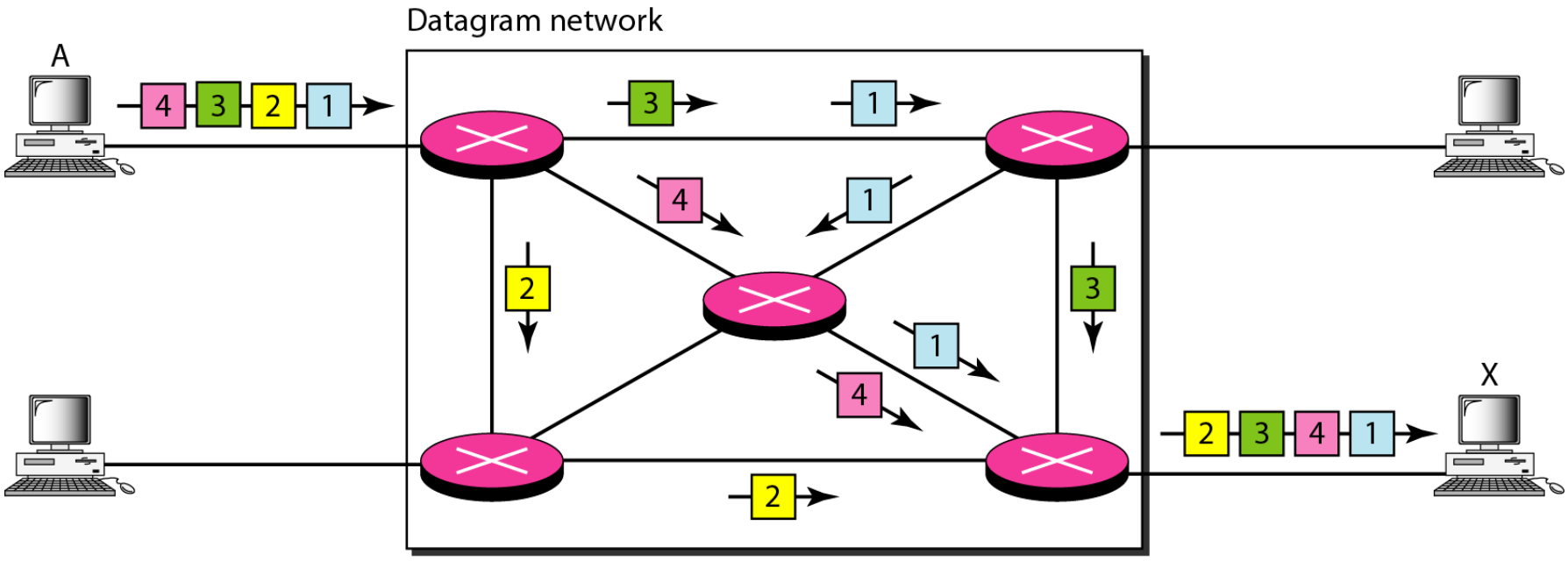
In data communications, we need to send messages from one end system to another. If the message is going to pass through a packet-switched network, it needs to be divided into packets of fixed or variable size. The size of the packet is determined by the network and the governing protocol.



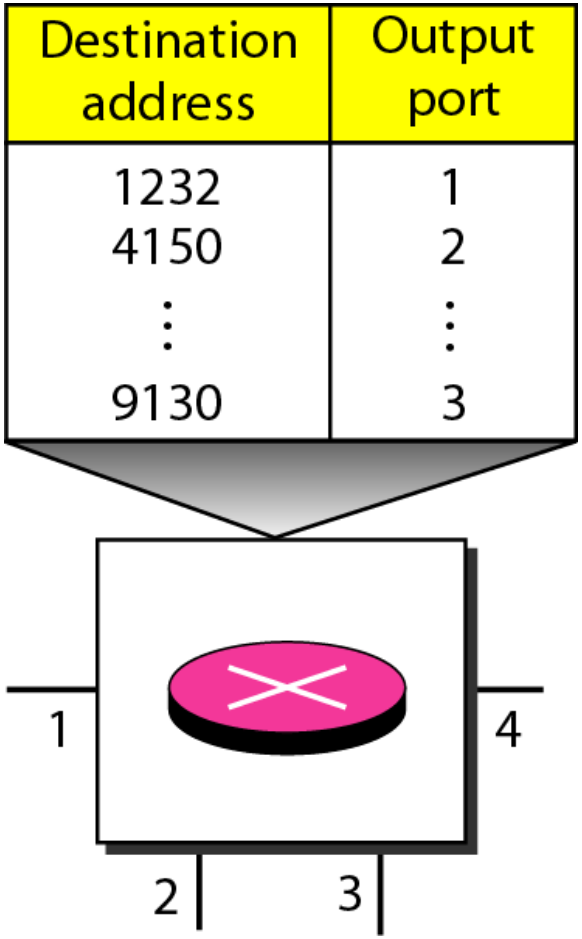
Note

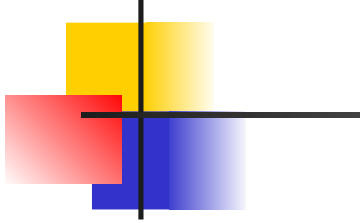
In a packet-switched network, there is no resource reservation; resources are allocated on demand.

A datagram network with four switches (routers)



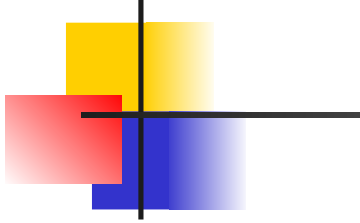
Routing table in a datagram network





Note

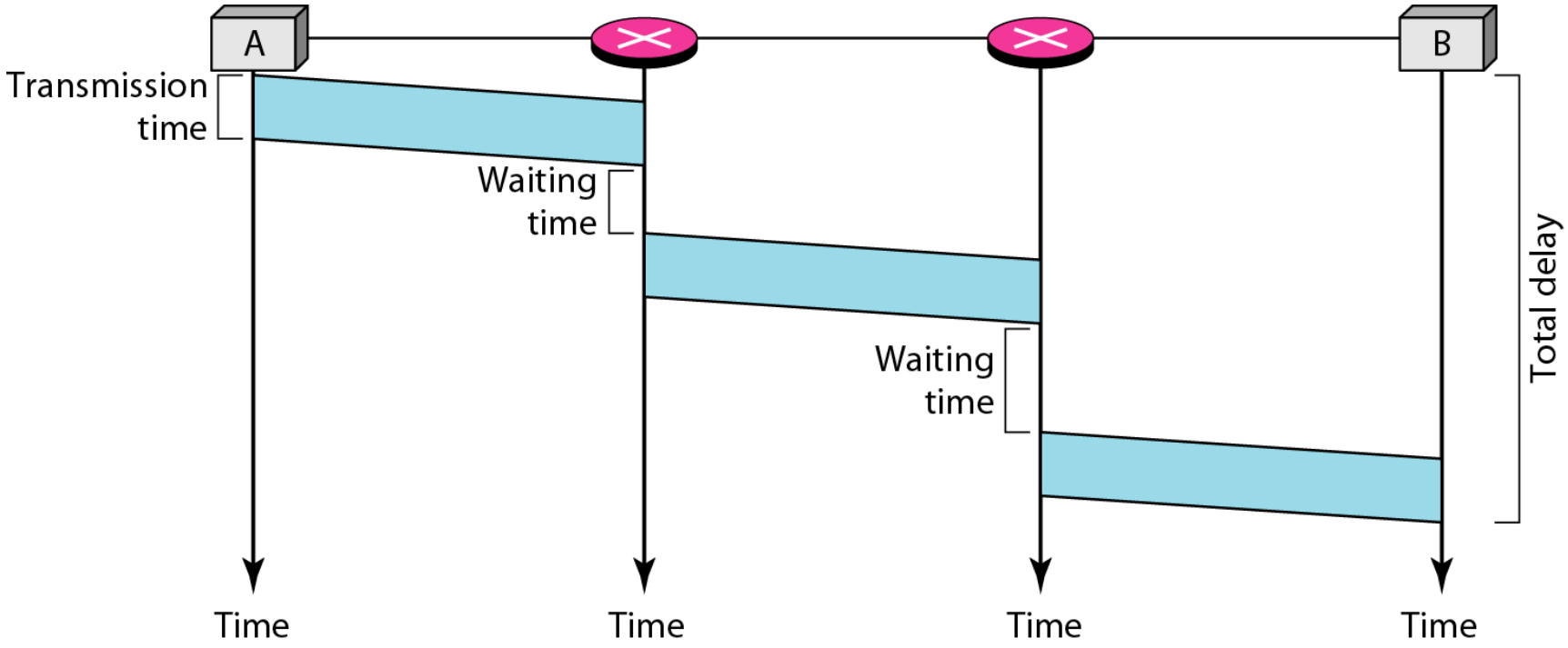
A switch in a datagram network uses a routing table that is based on the destination address.



Note

The destination address in the header of a packet in a datagram network remains the same during the entire journey of the packet.

Delay in a datagram network



Note

Switching in the Internet is done by using the datagram approach to packet switching at the network layer.

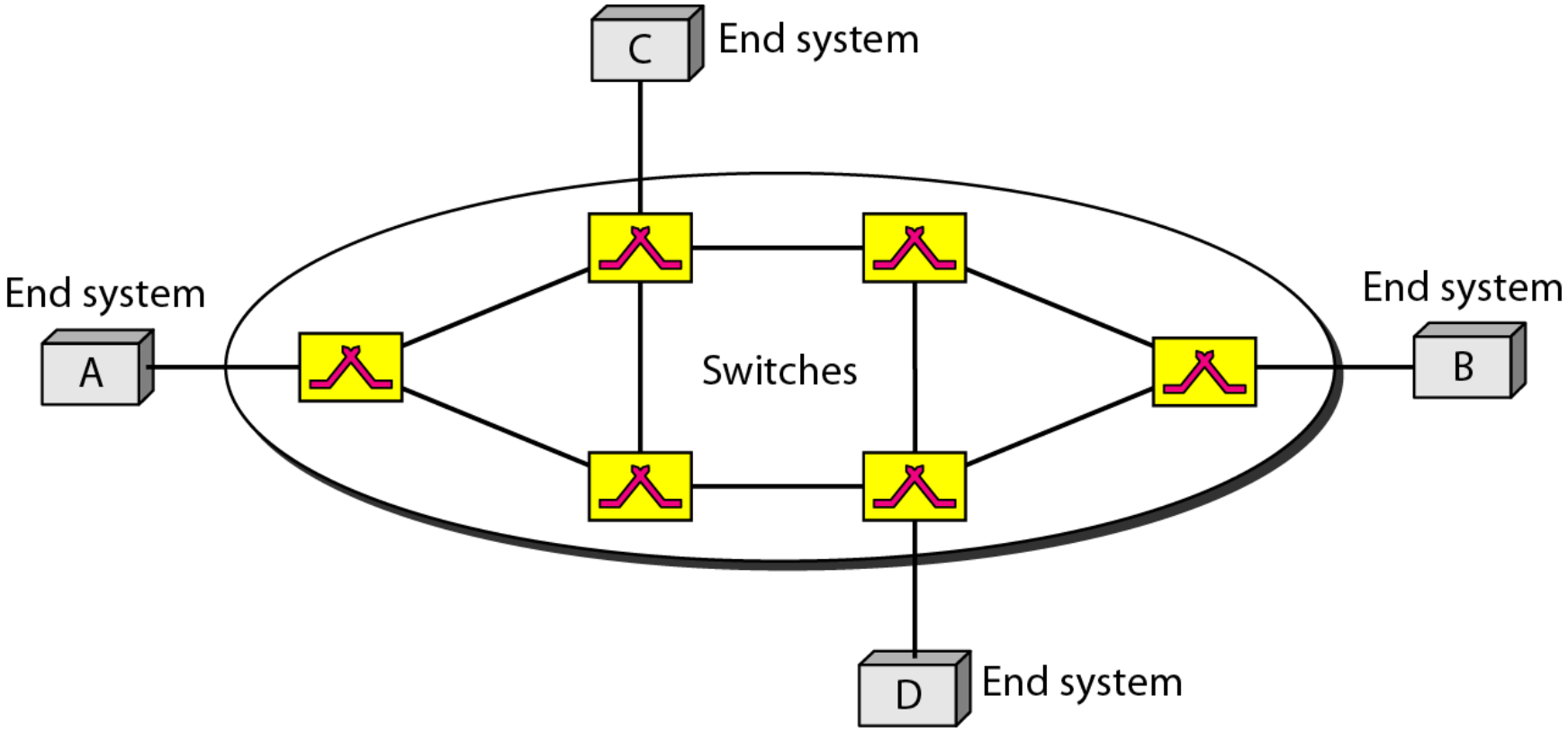
VIRTUAL-CIRCUIT NETWORKS

A virtual-circuit network is a cross between a circuit-switched network and a datagram network. It has some characteristics of both.

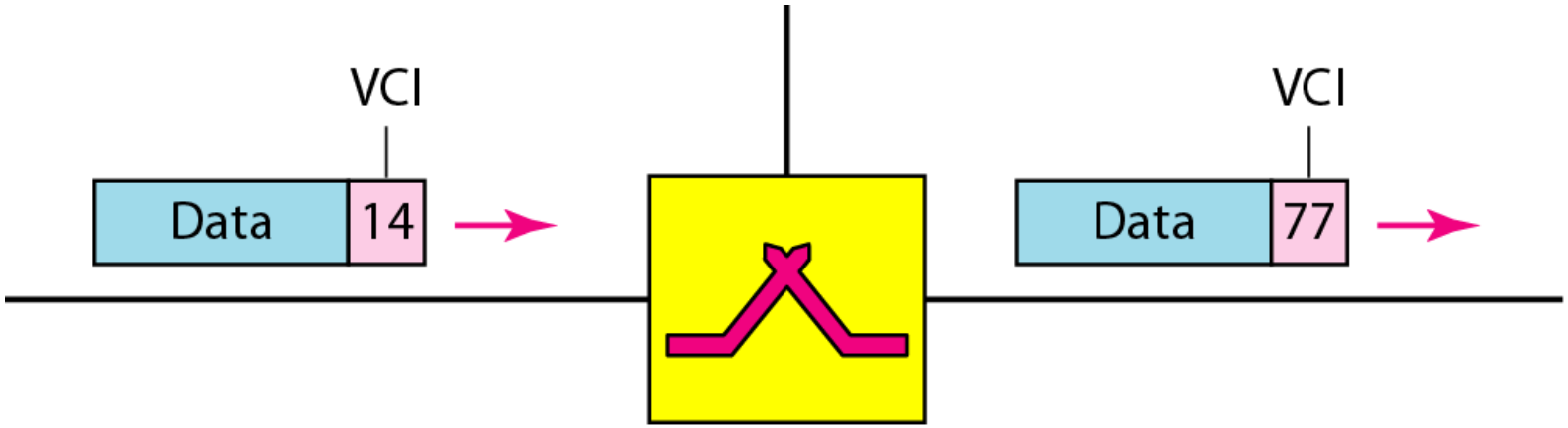
Virtual Circuit Networks

1. As in circuit switched network, there are set up and tear down phases in addition to the data transfer phase.
2. Resources can be allocated during the set up phase or on demand.
3. Data are packetized.
4. All packet follows same path established during the connection.
5. virtual circuit network implemented in data link layer, circuit switched network in physical layer and datagram network in network layer.

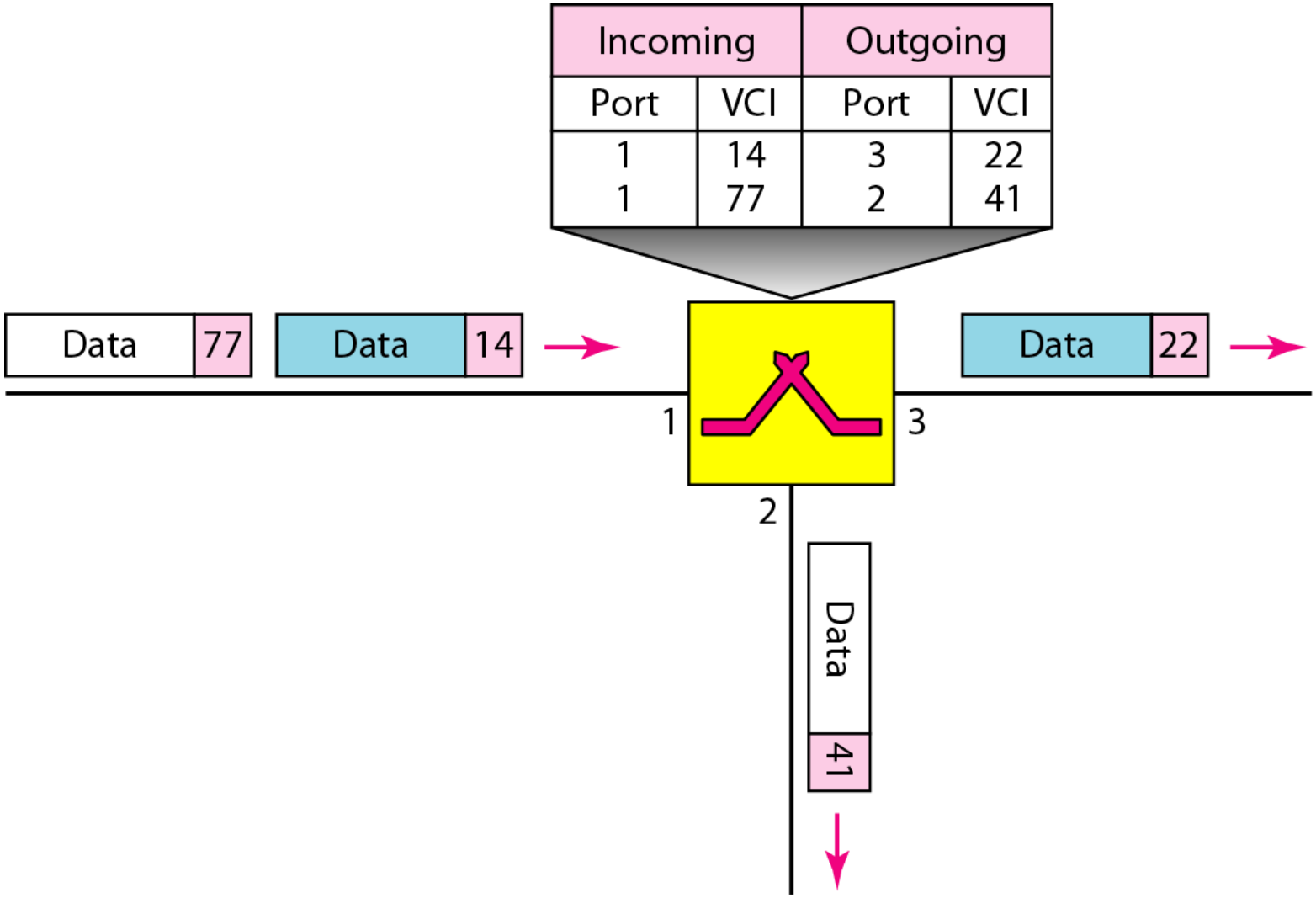
Virtual-circuit network



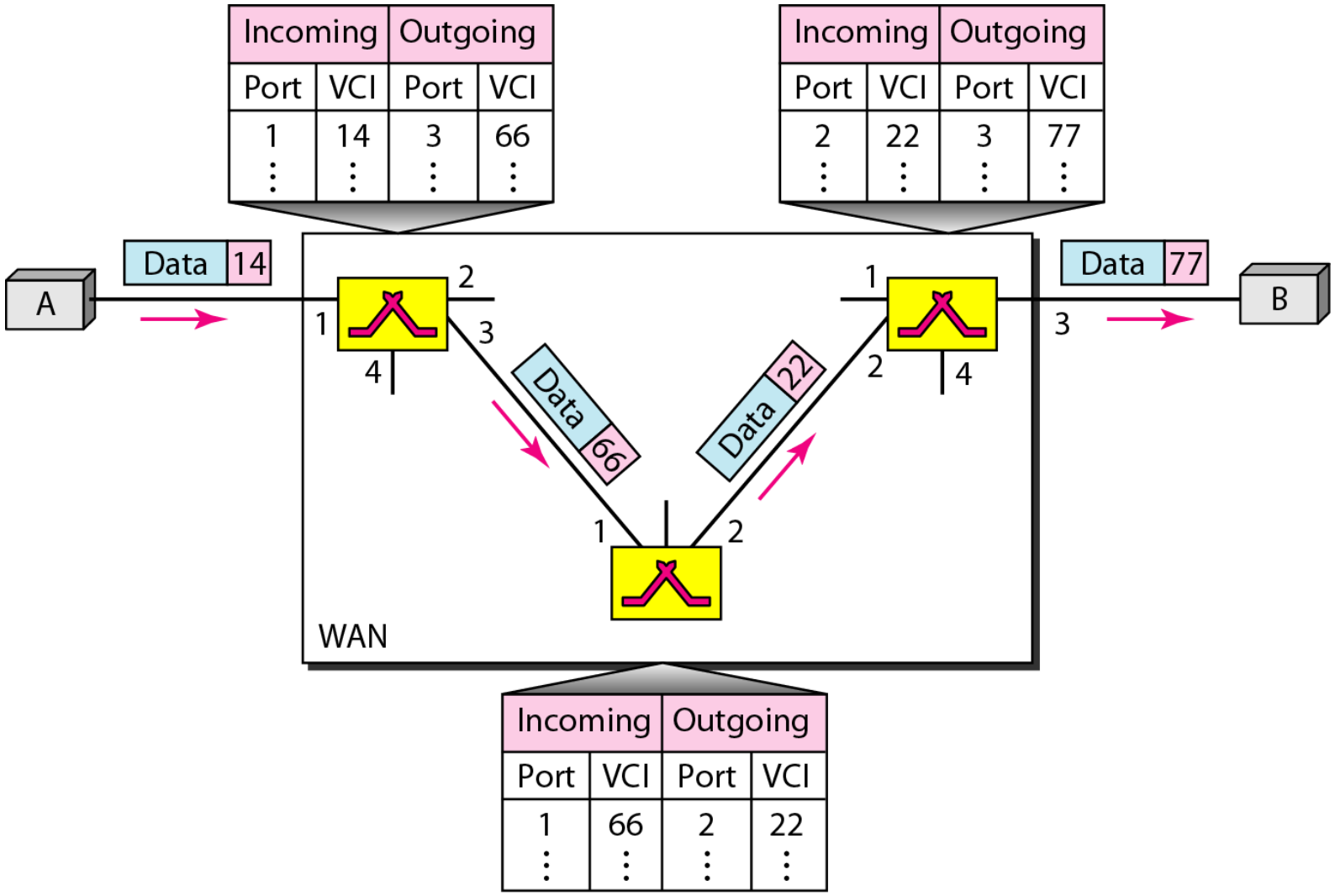
Virtual-circuit identifier



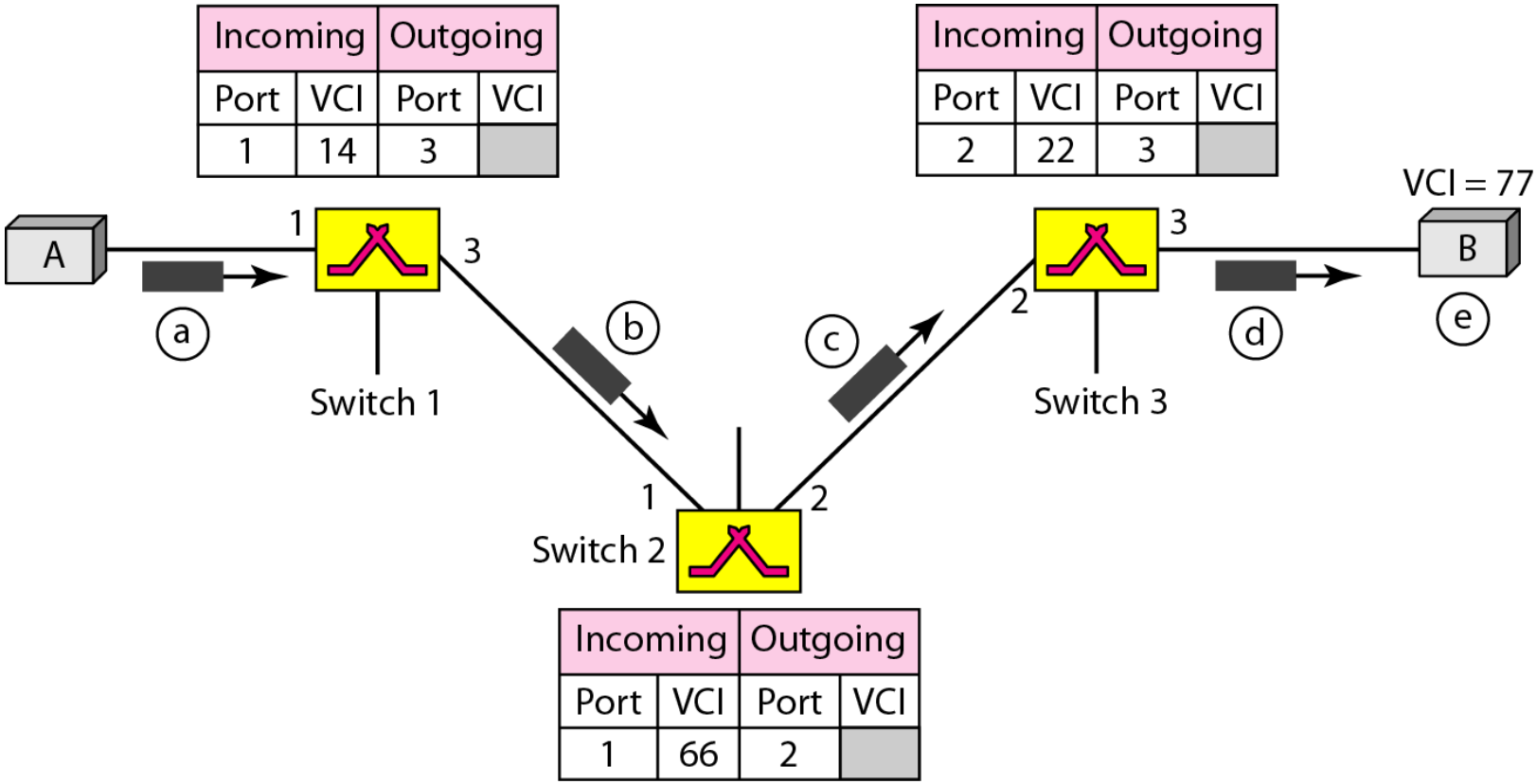
Switch and tables in a virtual-circuit network



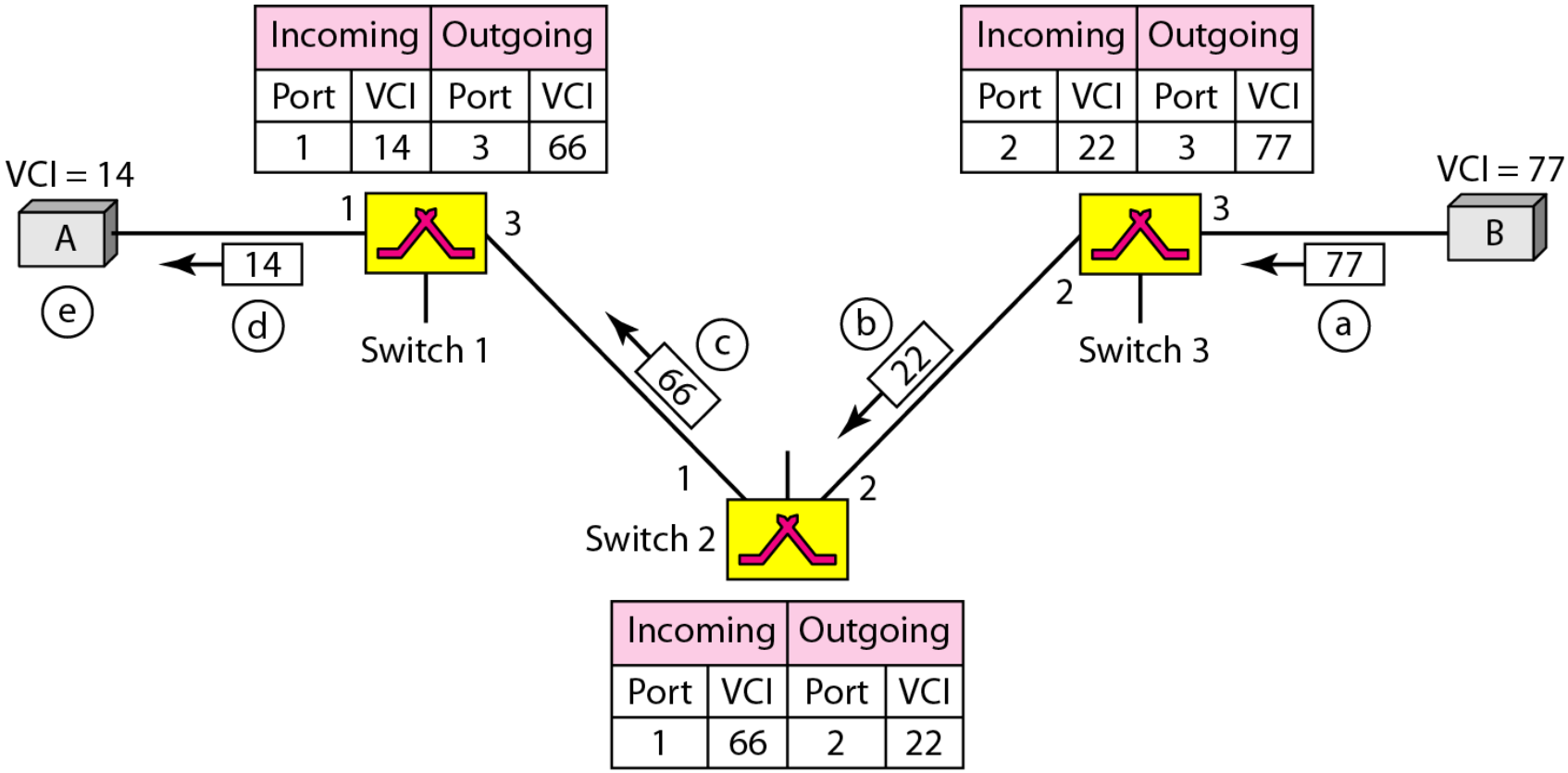
Source-to-destination data transfer in a virtual-circuit network



Setup request in a virtual-circuit network



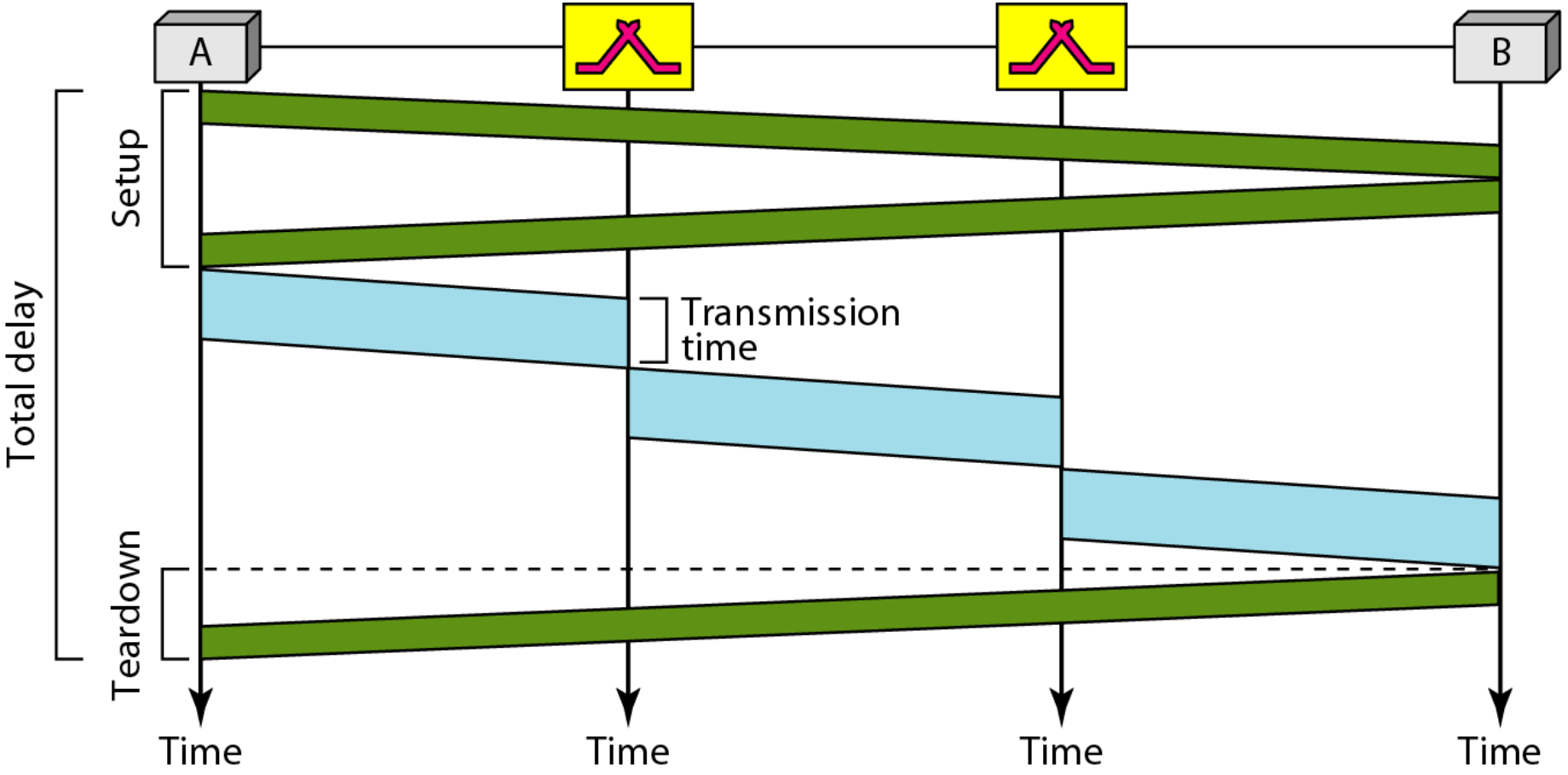
Setup acknowledgment in a virtual-circuit network



Note

In virtual-circuit switching, all packets belonging to the same source and destination travel the same path; but the packets may arrive at the destination with different delays if resource allocation is on demand.

Delay in a virtual-circuit network



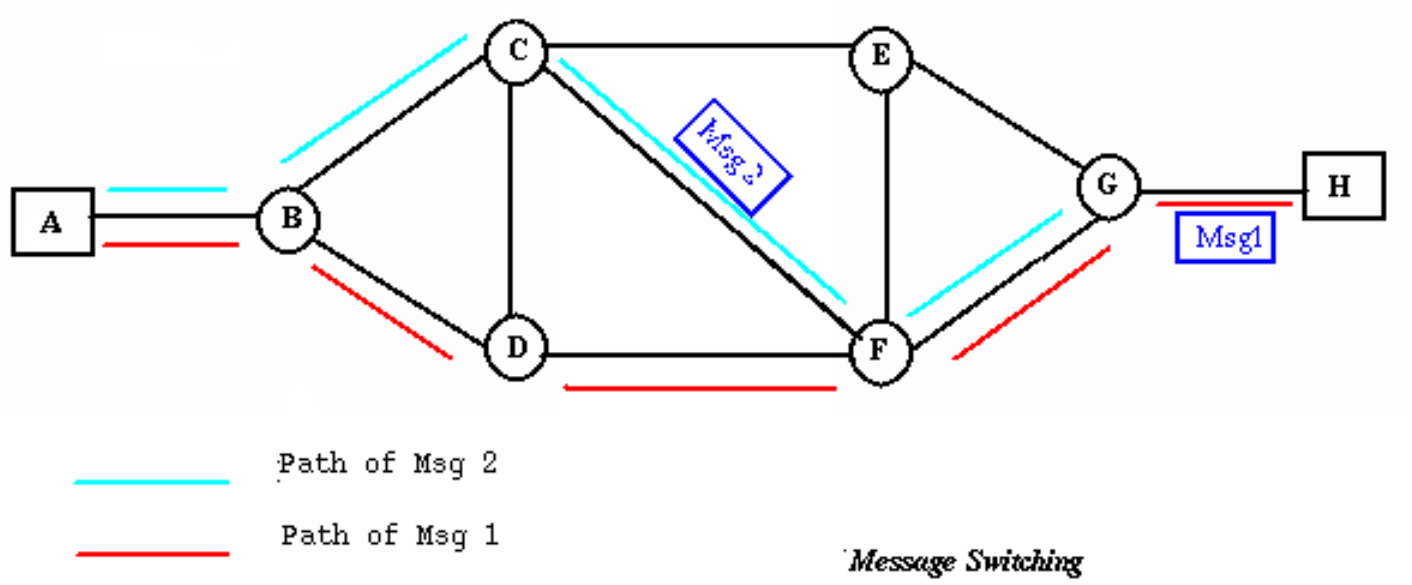
Note

Switching at the data link layer in a switched WAN is normally implemented by using virtual-circuit techniques.

Message Switching

- With message switching there is no need to establish a dedicated path between two stations.
- When a station sends a message, the destination address is appended to the message.
- The message is then transmitted through the network, in its entirety, from node to node.
- Each node receives the entire message, stores it in its entirety on disk, and then transmits the message to the next node.
- This type of network is called a store-and-forward network.

Message Switching



- A message-switching node is typically a general-purpose computer. The device needs sufficient secondary-storage capacity to store the incoming messages, which could be long. A time delay is introduced using this type of scheme due to store- and-forward time, plus the time required to find the next node in the transmission path.

Message Switching

- ***Advantages:***
  - Channel efficiency can be greater compared to circuit-switched systems, because more devices are sharing the channel.
  - Message priorities can be established due to store-and-forward technique.
  - Message broadcasting can be achieved with the use of broadcast address appended in the message.

Message Switching

- ***Disadvantages***
  - Message switching is not compatible with interactive applications.
  - Store-and-forward devices are expensive, because they must have large disks to hold potentially long messages.

Note

Switching at the network layer in the Internet uses the datagram approach to packet switching.

Note

Communication at the network layer in the Internet is connectionless.

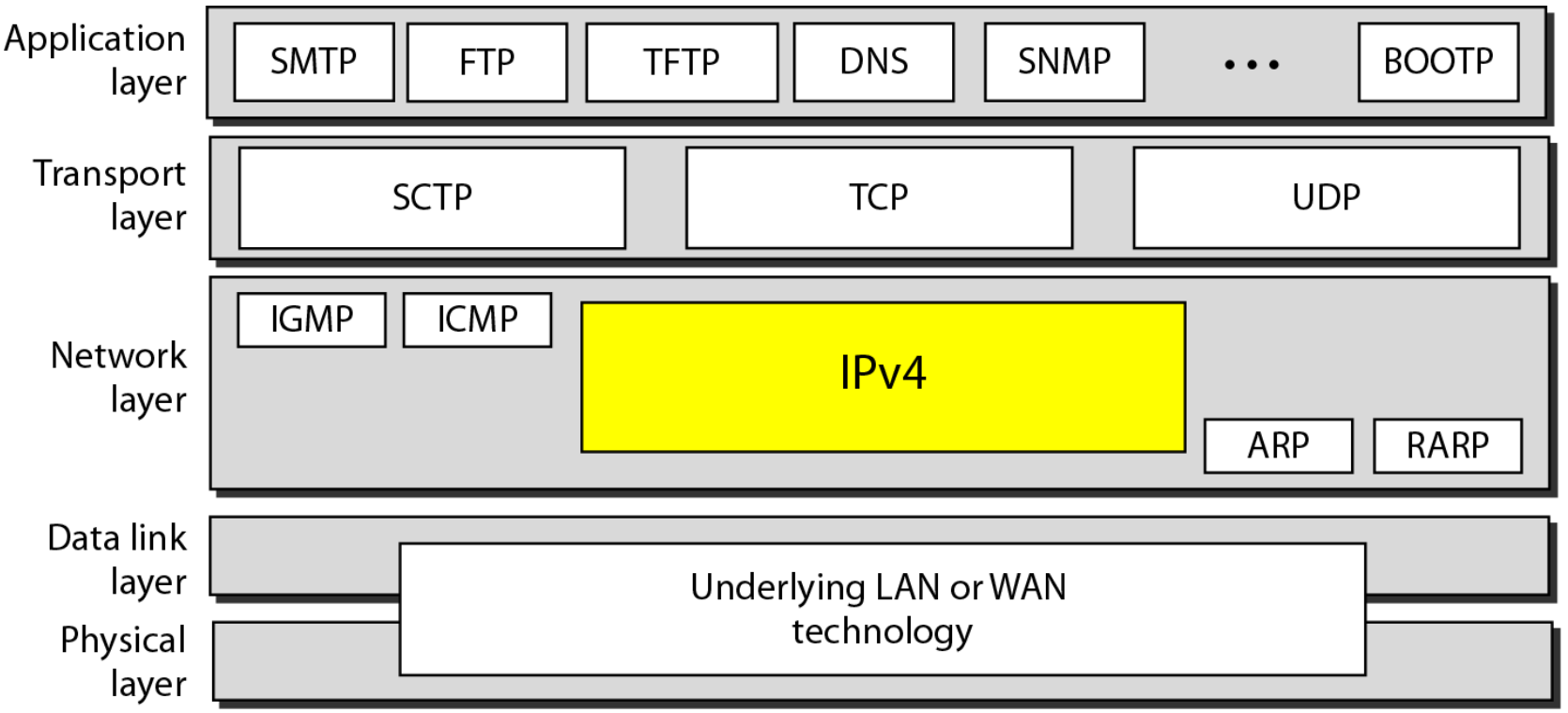
IP (IPv4)

- The Internet Protocol version 4 (IPv4) is the delivery mechanism used by the TCP/IP protocols.
- Unreliable, Connectionless datagram protocol.
- No error control and Flow control.
- Packets in IP layer is called datagram's.

Provides 2^{32} addresses.

- Address is written by dotted-decimal notation i.e. 172.16.10.84

Position of IPv4 in TCP/IP protocol suite



IPv4 datagram format

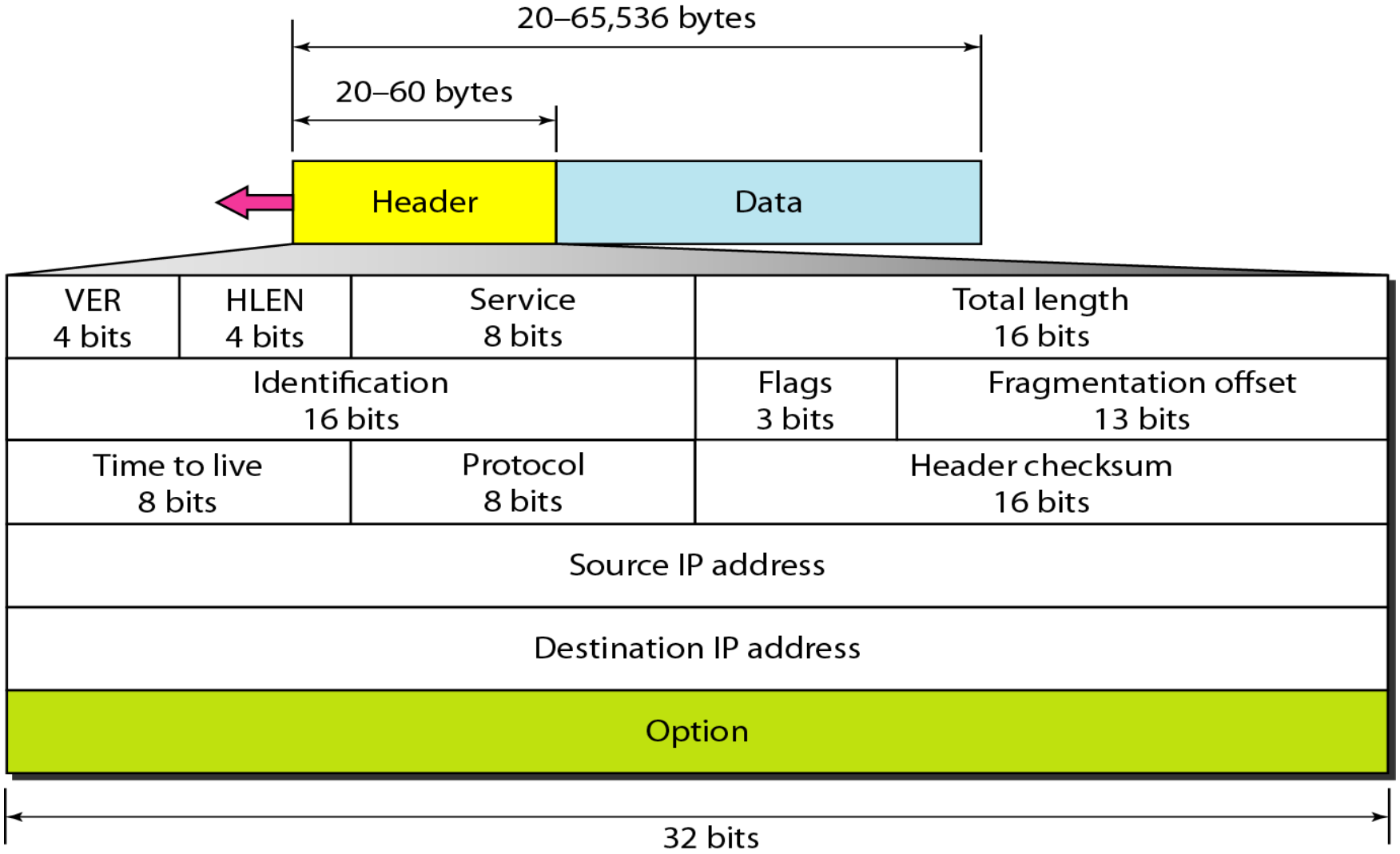
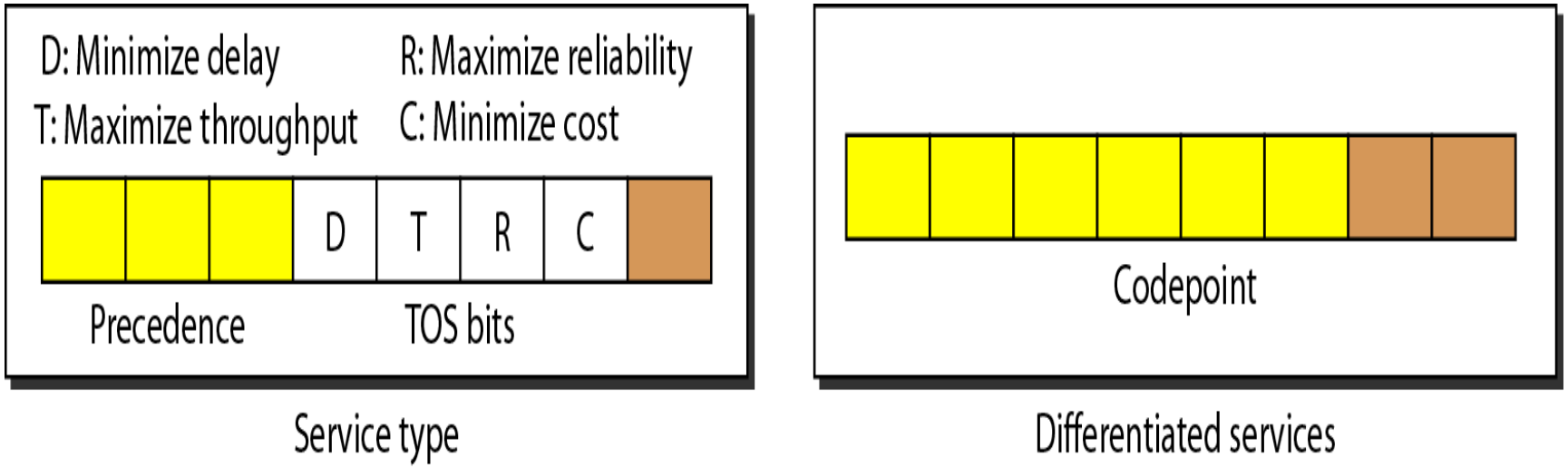


Figure. *Service type or differentiated services*



The precedence subfield was part of version 4, but never used.

Table. *Types of service*

TOS Bits	*Description*
0000	Normal (default)
0001	Minimize cost
0010	Maximize reliability
0100	Maximize throughput
1000	Minimize delay

Table. *Default types of service*

Protocol	*TOS Bits*	*Description*
ICMP	0000	Normal
BOOTP	0000	Normal
NNTP	0001	Minimize cost
IGP	0010	Maximize reliability
SNMP	0010	Maximize reliability
TELNET	1000	Minimize delay
FTP (data)	0100	Maximize throughput
FTP (control)	1000	Minimize delay
TFTP	1000	Minimize delay
SMTP (command)	1000	Minimize delay
SMTP (data)	0100	Maximize throughput
DNS (UDP query)	1000	Minimize delay
DNS (TCP query)	0000	Normal
DNS (zone)	0100	Maximize throughput

DS field

DS field contain 2 subfields:

1. DSCP (Differentiated Services Code Point): 6 bit, defines the Per-Hop Behavior (PHB).
2. CU (currently unused)

Per-Hop Behavior:

PHBs are defined:

3. DE PHB (Default PHB)
4. EF PHB (Expedited Forwarding PHB)
5. AF PHB (Assured Forwarding PHB)

Figure. *Protocol field and encapsulated data*

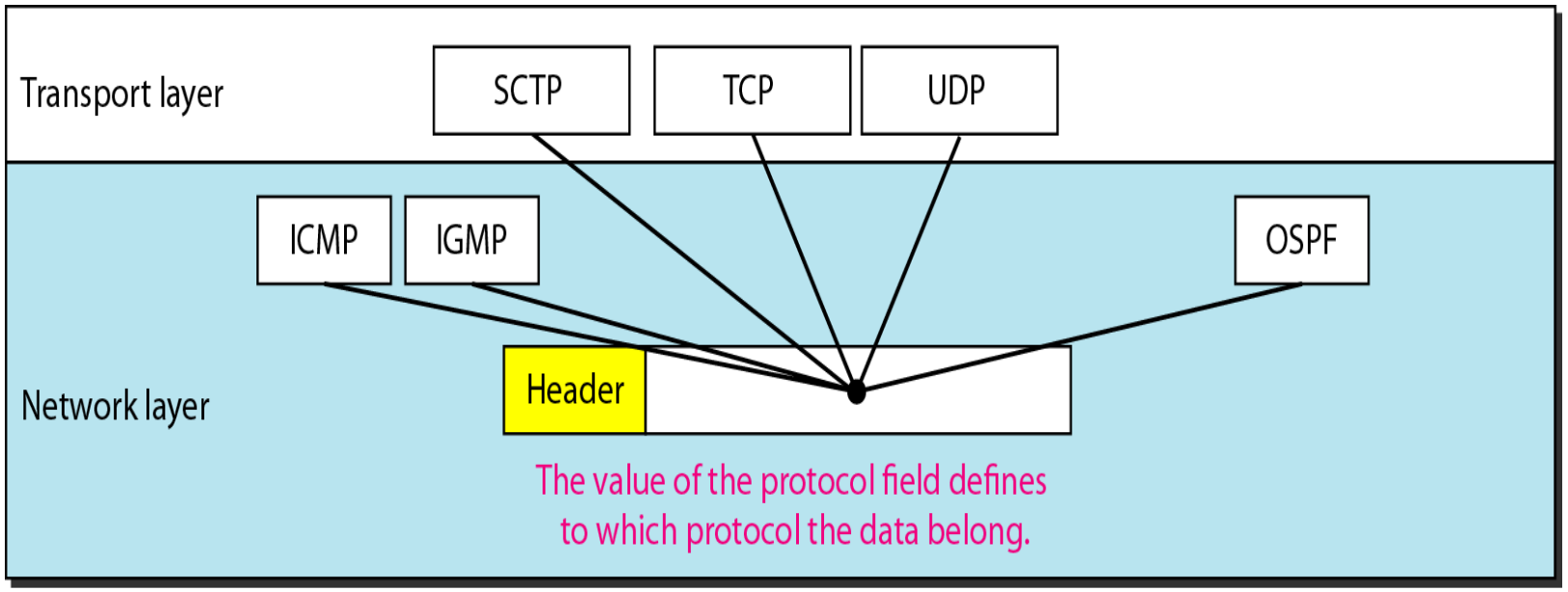
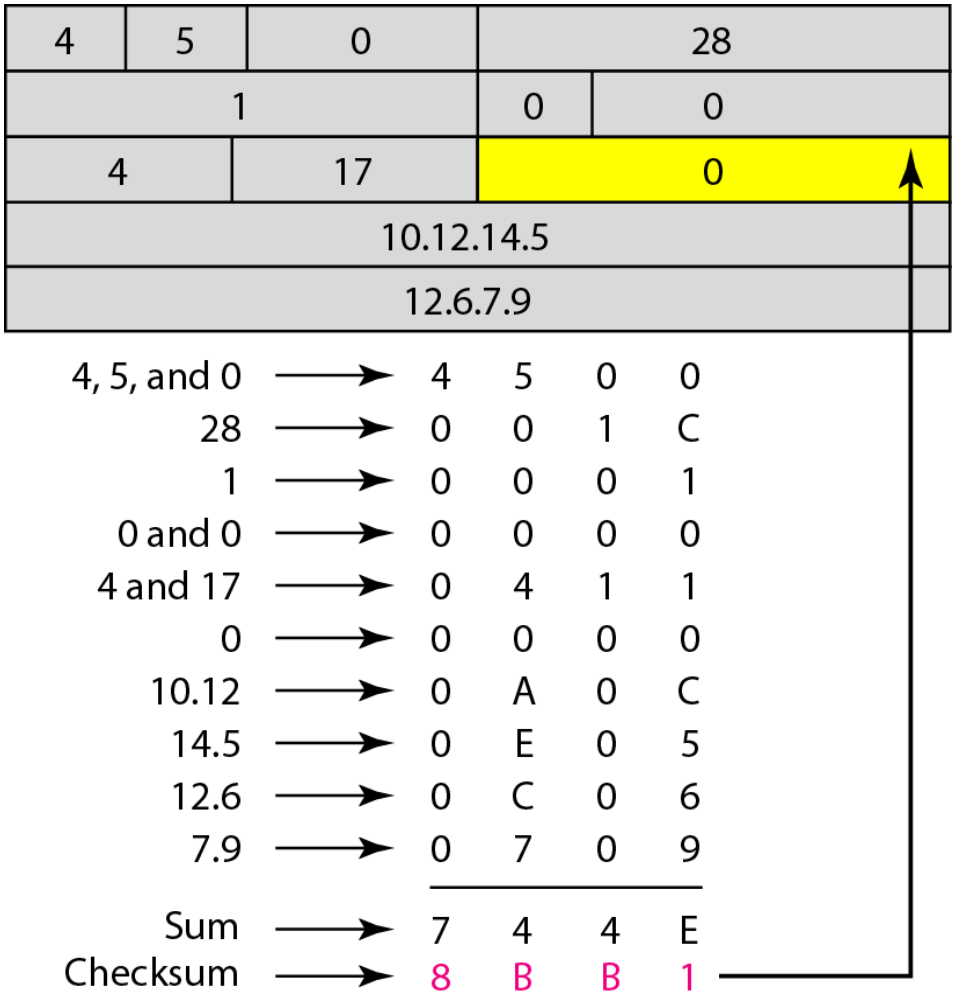


Table. *Protocol values*

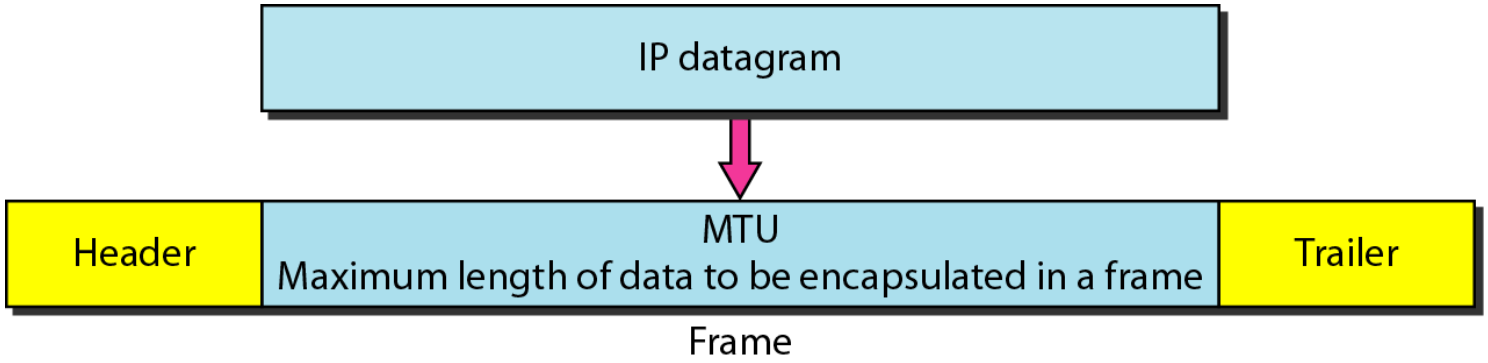
Value	*Protocol*
1	ICMP
2	IGMP
6	TCP
17	UDP
89	OSPF

Example of checksum calculation in IPv4



Fragmentation:

Maximum transfer unit (MTU)



MTUs for some networks

Protocol	*MTU*
Hyperchannel	65,535
Token Ring (16 Mbps)	17,914
Token Ring (4 Mbps)	4,464
FDDI	4,352
Ethernet	1,500
X.25	576
PPP	296

Fields Related to Fragmentation:

1. Identification.

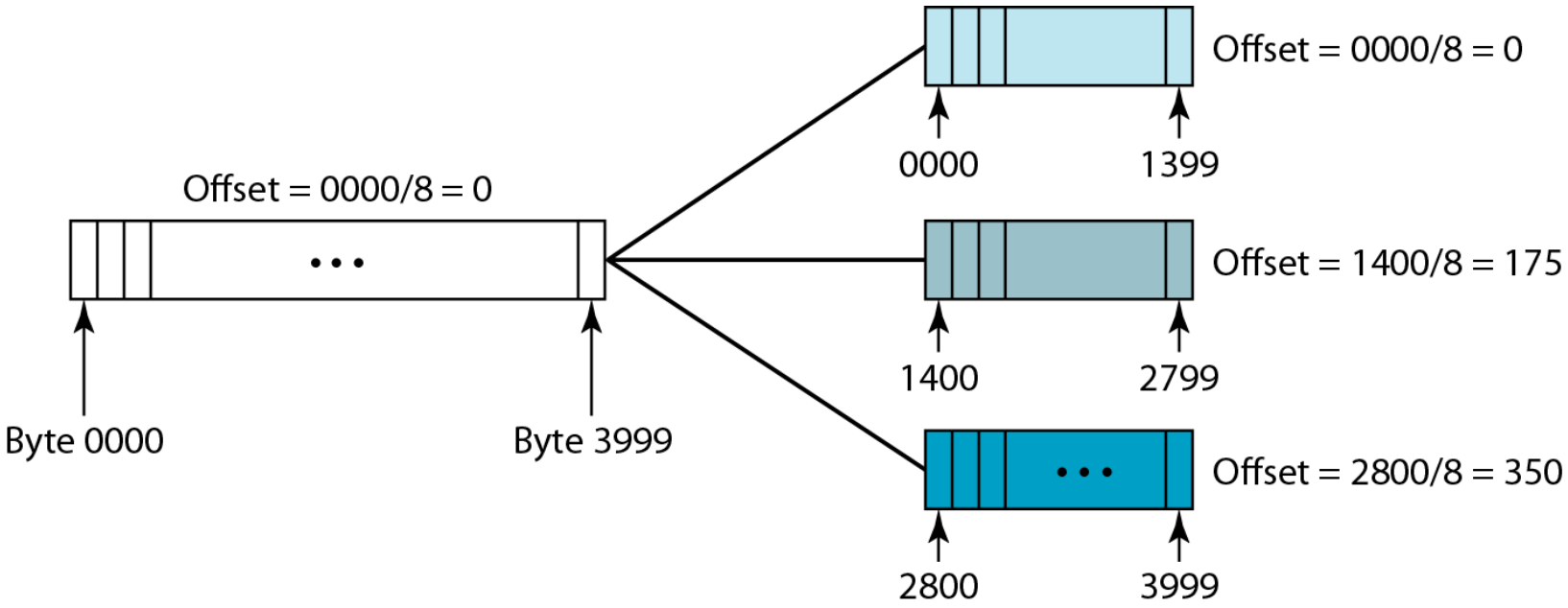
2. Flags used in fragmentation

D: Do not fragment
M: More fragments

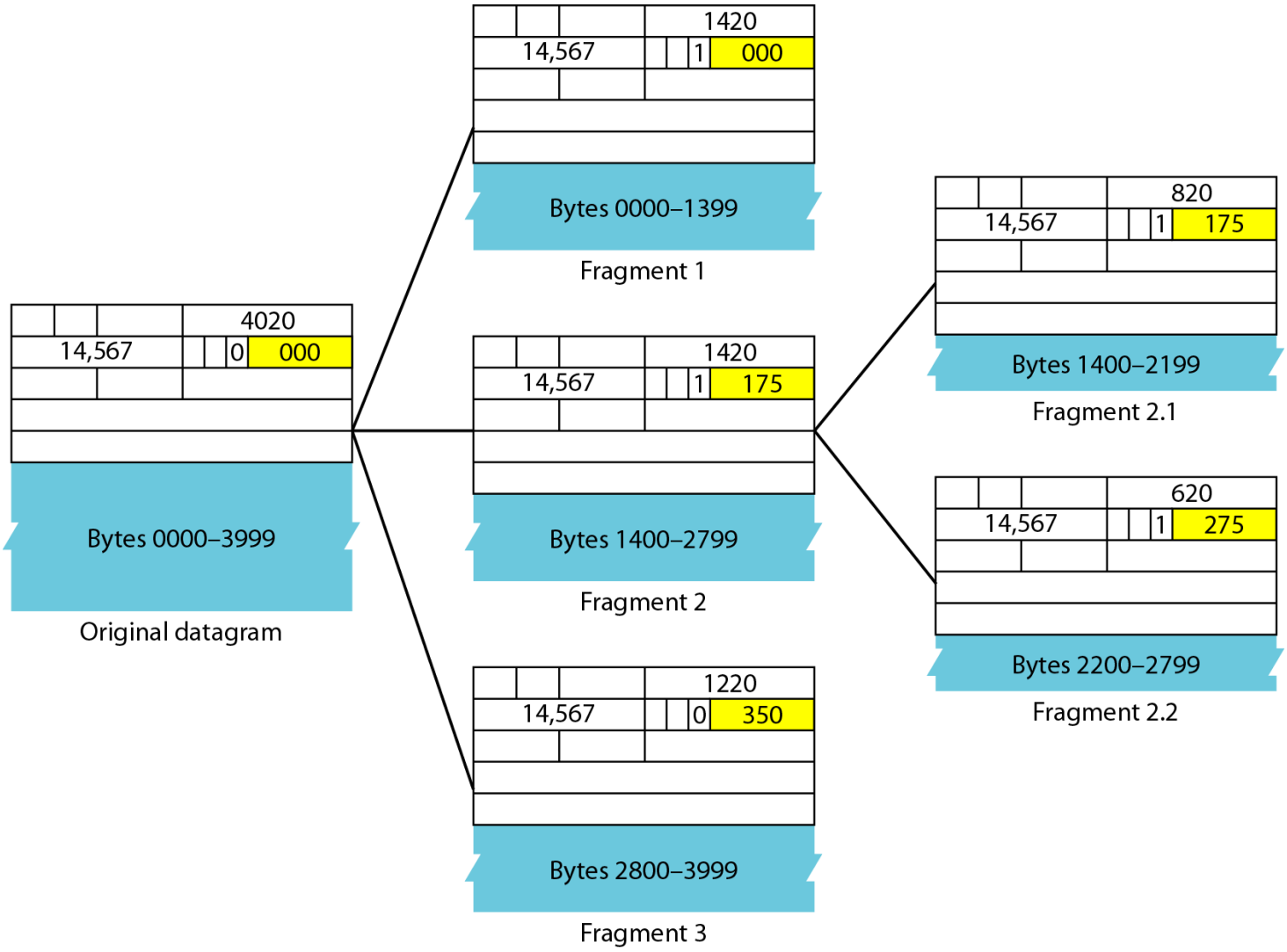
3. Fragmentation Offset.

- The field is 13 bits wide, so the offset can be from 0 to 8191.
- Fragments are specified in units of 8 bytes, which is why fragment
length must be a multiple of 8.
*- Uncoincidentally, 8191 * 8 = 65,528, just about the maximum size*
allowed for an IP datagram.

Fragmentation Offset



Detailed fragmentation example



IPv4 ADDRESSES

An IPv4 address is a 32-bit address that uniquely and universally defines the connection of a device (for example, a computer or a router) to the Internet.

Topics discussed in this section:

Address Space
Notations
Classful Addressing
Network Address Translation (NAT)

Note

An IPv4 address is 32 bits long.

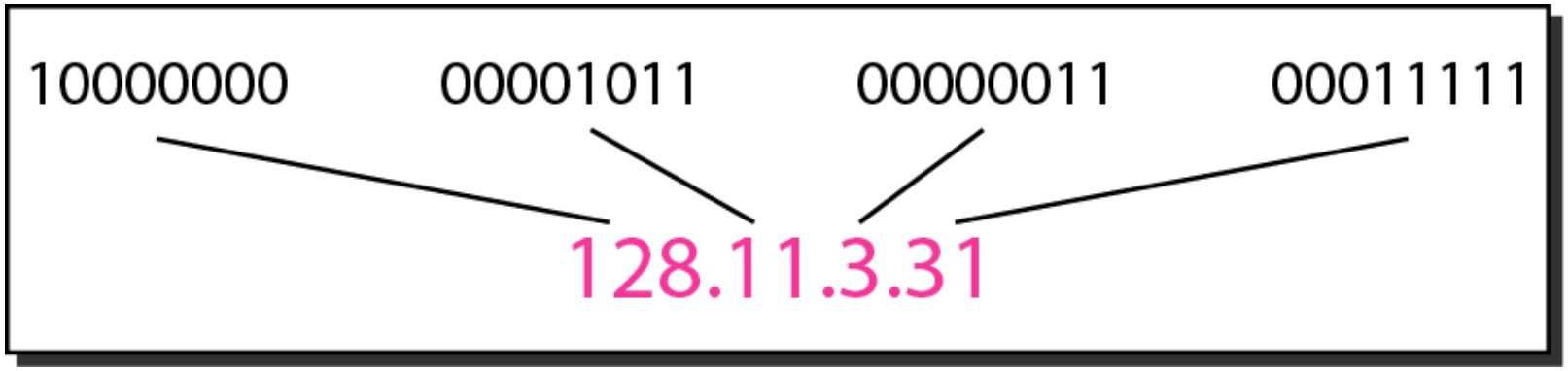
The IPv4 addresses are unique and universal.

1. Address Space

The address space of IPv4 is 2^{32} or 4,294,967,296.

2. Notations

Dotted-decimal notation and binary notation for an IPv4 address.

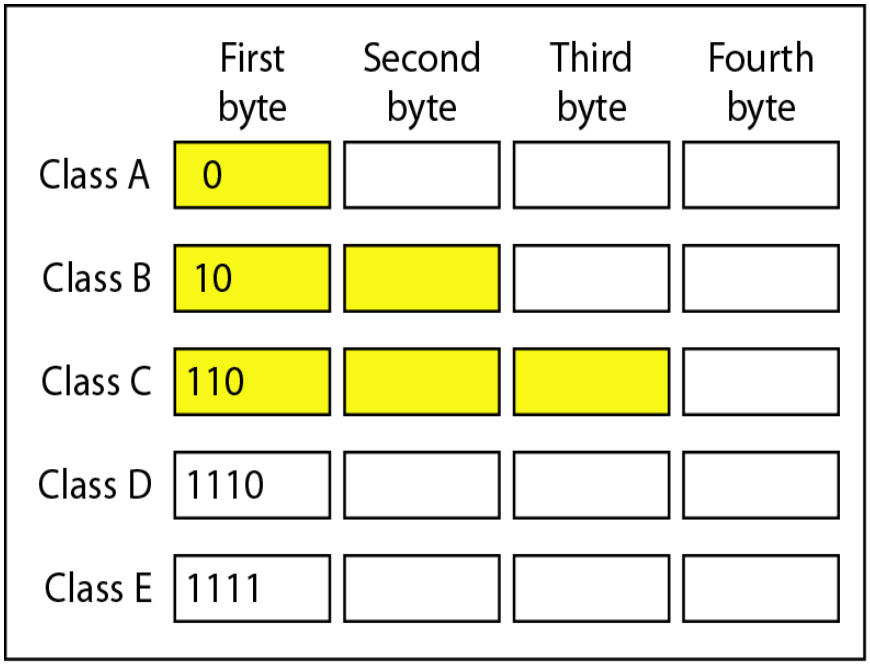


3. Classful addressing.

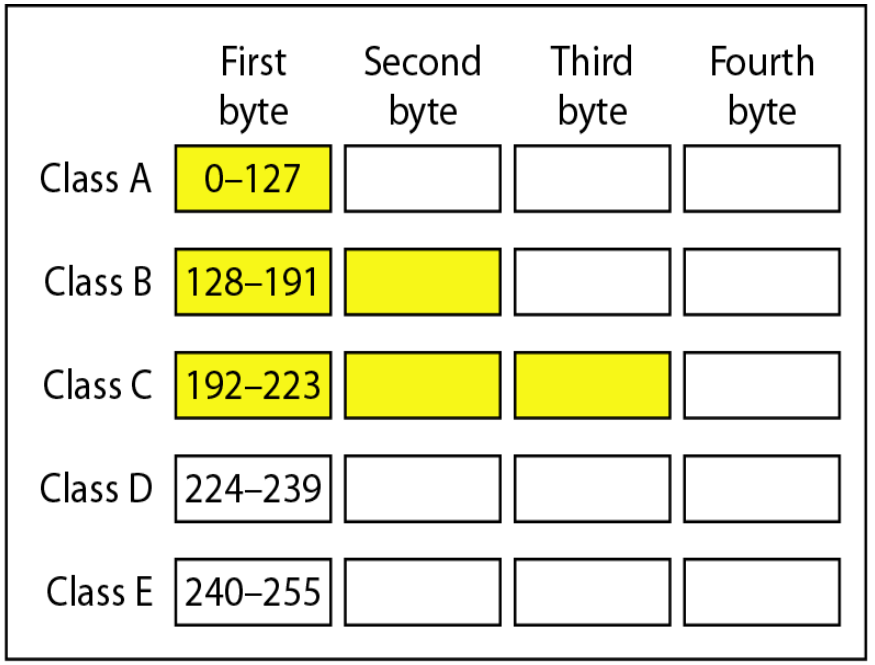
Note

In classful addressing, the address space is divided into five classes:
A, B, C, D, and E.

Finding the classes in binary and dotted-decimal notation



a. Binary notation



b. Dotted-decimal notation

Note

In classful addressing, a large part of the available addresses were wasted.

Note

Classful addressing, which is almost obsolete, is replaced with classless addressing.

Classless Addressing

- To overcome address depletion and give more organizations access to the Internet, classless addressing was designed and implemented.
- In this scheme, there are no classes, but the addresses are still granted in blocks.

Classless Interdomain Routing (CIDR) notation :

Class	*Binary*	*Dotted-Decimal*	*CIDR*
A	**11111111** 00000000 00000000 00000000	**255**.0.0.0	/8
B	**11111111 11111111** 00000000 00000000	**255.255**.0.0	/16
C	**11111111 11111111 11111111** 00000000	**255.255.255**.0	/24

The last column of above Table shows the mask in the form "/n" where n can be 8, 16, or 24 in classful addressing.

This notation is also called slash notation or **Classless Interdomain Routing (CIDR) notation.**

The notation is used in classless addressing.

Address Blocks :

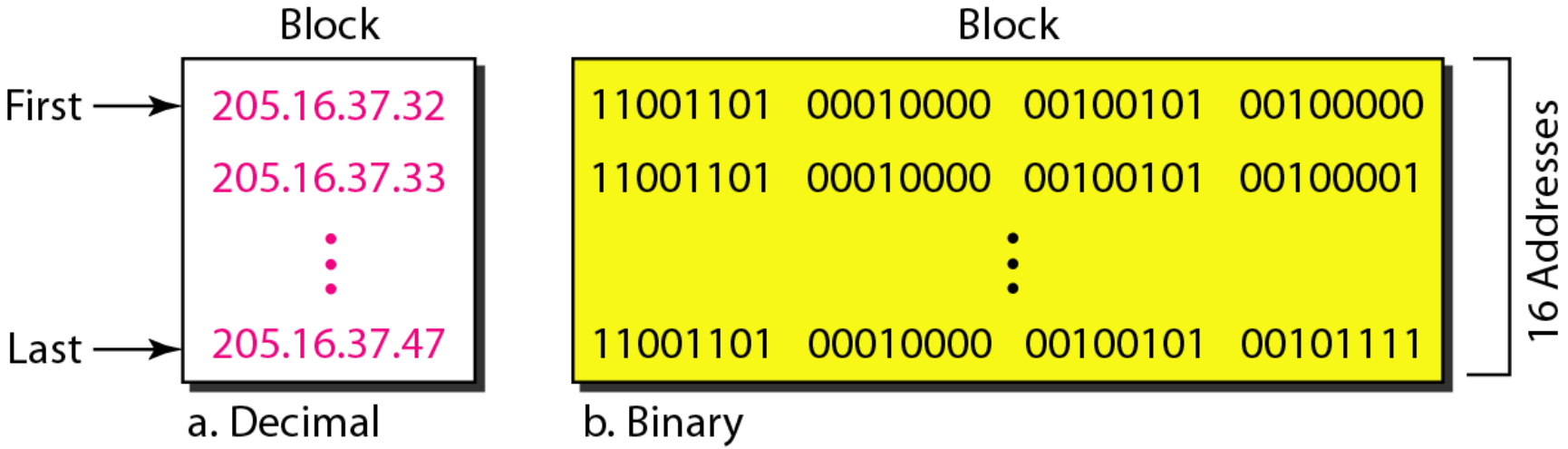
- In classless addressing, when an entity, small or large, needs to be connected to the Internet, it is granted a block (range) of addresses.

- The size of the block (the number of addresses) varies based on the nature and size of the entity.
- For example, a household may be given only two addresses; a large organization may be given thousands of addresses. An ISP, may be given thousands or Hundreds of thousands based on the number of customers it may serve.

- To simplify the handling of addresses, the Internet authorities impose three restrictions on classless address blocks:

1. The addresses in a block must be contiguous, one after another.
2. The number of addresses in a block must be a power of 2 (1, 2, 4, 8, ...).
3. The first address must be evenly divisible by the number of addresses.

Figure 19.3 *A block of 16 addresses granted to a small organization*



Note

In IPv4 addressing, a block of addresses can be defined as x.y.z.t /n in which x.y.z.t defines one of the addresses and the /n defines the mask.

Note

The first address in the block can be found by setting the rightmost $32 - n$ bits to 0s.

Example Continued...

A block of addresses is granted to a small organization. We know that one of the addresses is 205.16.37.39/28. What is the first address in the block?

Solution

The binary representation of the given address is

11001101 00010000 00100101 00100111

If we set 32−28 rightmost bits to 0, we get

11001101 00010000 00100101 00100000

or

205.16.37.32.

This is actually the block shown in Figure 19.3.

Note

The last address in the block can be found by setting the rightmost $32 - n$ bits to 1s.

Example 19.7

Example continued...

Find the last address for the block in previous Example.

Solution

The binary representation of the given address is

11001101 00010000 00100101 00100111

If we set 32 − 28 rightmost bits to 1, we get

11001101 00010000 00100101 00101111

or

205.16.37.47

This is actually the block shown in Figure 19.3.

Note

The number of addresses in the block can be found by using the formula 2^{32-n}.

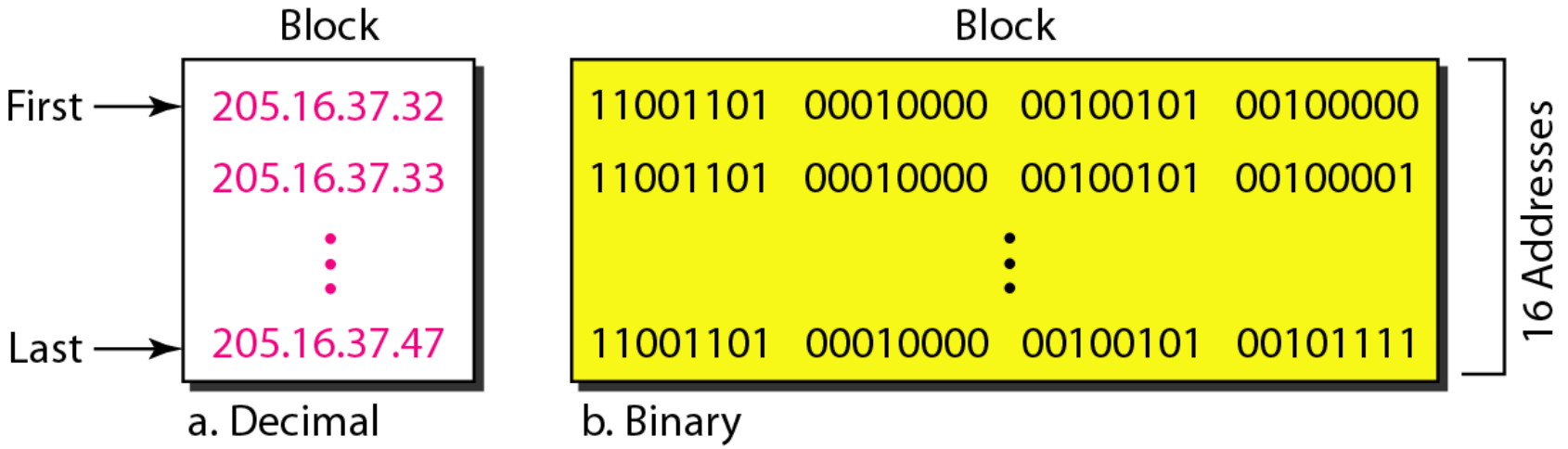
Example Continued...

Find the number of addresses in Example 19.6.

Solution

The value of n is 28, which means that number of addresses is 2^{32-28} or 16.

Figure of Example *A network configuration for the block 205.16.37.32/28*



Example

Another way to find the first address, the last address, and the number of addresses is to represent the mask as a 32-bit binary (or 8-digit hexadecimal) number. This is particularly useful when we are writing a program to find these pieces of information. In Example 19.5 the /28 can be represented as

11111111 11111111 11111111 11110000

(twenty-eight 1s and four 0s).

Find

a. The first address

b. The last address

c. The number of addresses.

Mask :

Another way to find First and Lasst addresses in the block is to use the address mask.

The address mask is a 32 bit number in which the n leftmost bits are set to 1s and the rest of the bits (32-n) are set to 0s.

In classfull addressing, the length of the netid and hostid (in bits) can be find using Mask.

We can use a same mask (also called the default mask), a 32-bit number made of contiguous 1s followed by contiguous 0s.

The masks for classes A, B, and C are shown in below Table.

The concept does not apply to classes D and E.

Class	*Binary*	*Dotted-Decimal*	*CIDR*
A	**11111111** 00000000 00000000 00000000	**255**.0.0.0	/8
B	**11111111 11111111** 00000000 00000000	**255.255**.0.0	/16
C	**11111111 11111111 11111111** 00000000	**255.255.255**.0	/24

Example continued...

Solution

a. The first address can be found by ANDing the given addresses with the mask. ANDing here is done bit by

bit. The result of ANDing 2 bits is 1 if both bits are 1s;

the result is 0 otherwise.

Address:	11001101	00010000	00100101	00100111
Mask:	**11111111**	**11111111**	**11111111**	**11110000**
First address:	11001101	00010000	00100101	00100000

Example continued..

b. The last address can be found by ORing the given addresses with the complement of the mask. ORing here is done bit by bit. The result of ORing 2 bits is 0 if both bits are 0s; the result is 1 otherwise. The complement of a number is found by changing each 1 to 0 and each 0 to 1.

Address:	11001101	00010000	00100101	00100111
Mask complement:	**00000000**	**00000000**	**00000000**	**00001111**
Last address:	11001101	00010000	00100101	00101111

Example continued...

c. The number of addresses can be found by complementing the mask, interpreting it as a decimal number, and adding 1 to it.

Mask complement:	00000000 00000000 00000000 00001111
Number of addresses:	15 + 1 = 16

Note

The first address in a block is normally not assigned to any device;
it is used as the network address that represents the organization to the rest of the world.

1.Limitations of IPv4:

2.- Limited IP address field i.e. 232

3.- No encryption and Authentication is provided.

4.- Minimum delay and resource reservation is not provided for real time audio and video transmission.

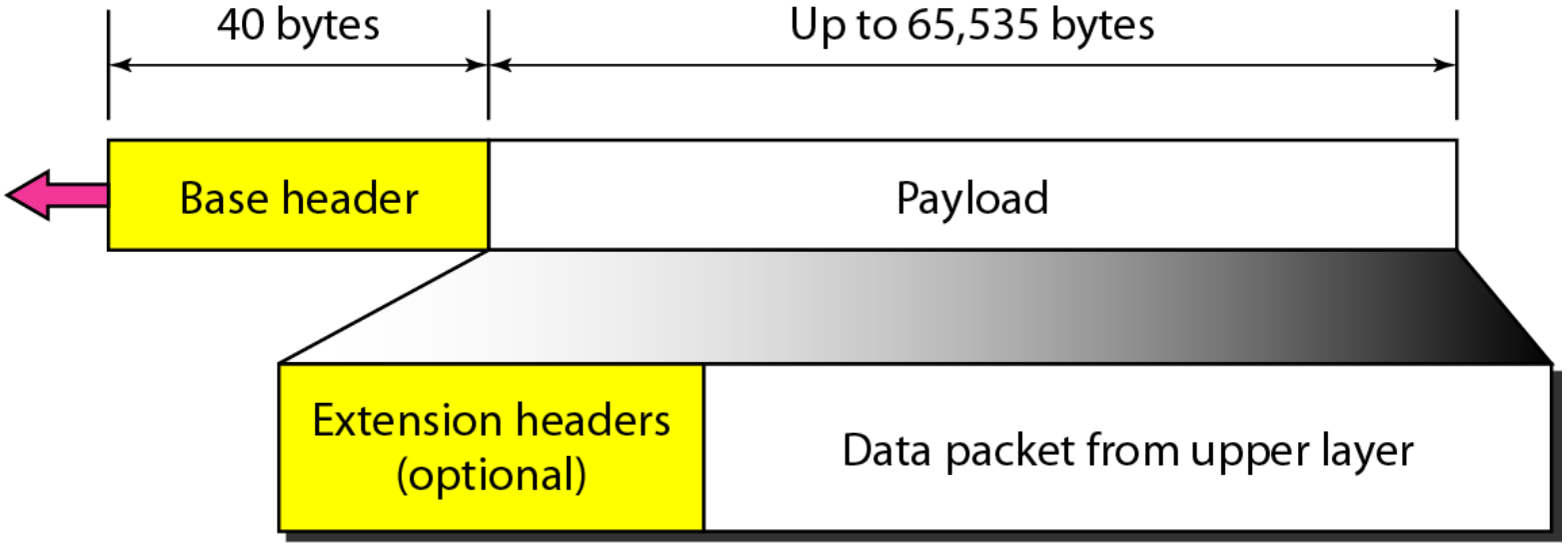
IPv6

The network layer protocol in the TCP/IP protocol suite is currently IPv4. IPv4 has some deficiencies that make it unsuitable for the fast-growing Internet.

Advantages of IPv6:

- **Larger Address Space:** 128 bit address space.
- **Better Header Format:** Options are separated and inserted when needed. It simplifies and speed up routing process.
- **New options:** To allow for additional functionalities.
- **Allowance for extension:** It allow the extension of the protocol if required by new technologies or applications.
- **Support for resource allocation:** Is used to support traffic such as real time audio and video.
- **Support for more security:** Encryption and authentication options in IPv6 provides confidentiality and integrity of the packet.

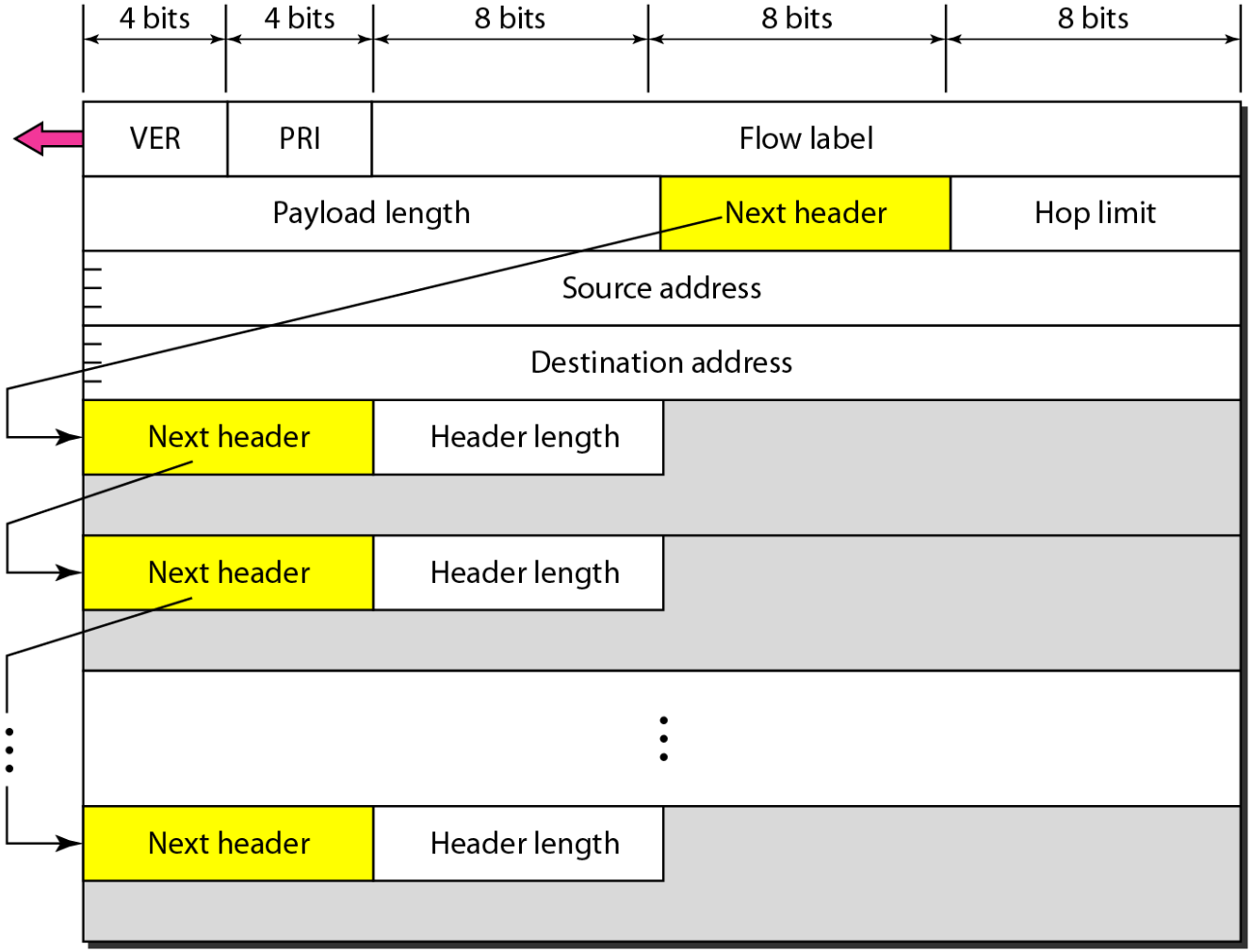
IPv6 datagram header and payload



Note

An IPv6 address is 128 bits long. (16 bytes)

Format of an IPv6 datagram



Version: Specifies the protocol version number. Here 6.
Priority(Traffic class)(4 bits): Defines priority of the packet with respect to traffic congestion.
Flow label(3 bytes): Designed to provide special handling for a particular flow of data. Used by host to label those packets for which it is requesting special handling by routers within a network.
Payload Length (16 bits): Gives total length of IP datagram excluding base header.
Next Header (8 bits): Gives type of next header.
Hop Limit (8 bits): Same as TTL.
Source Address (16 bytes): Gives original address of datagram source.
Destination Address (16 bytes) : Identifies destination address of datagram source.

Table *Priorities for congestion-controlled traffic*

Priority	*Meaning*
0	No specific traffic
1	Background data
2	Unattended data traffic
3	Reserved
4	Attended bulk data traffic
5	Reserved
6	Interactive traffic
7	Control traffic

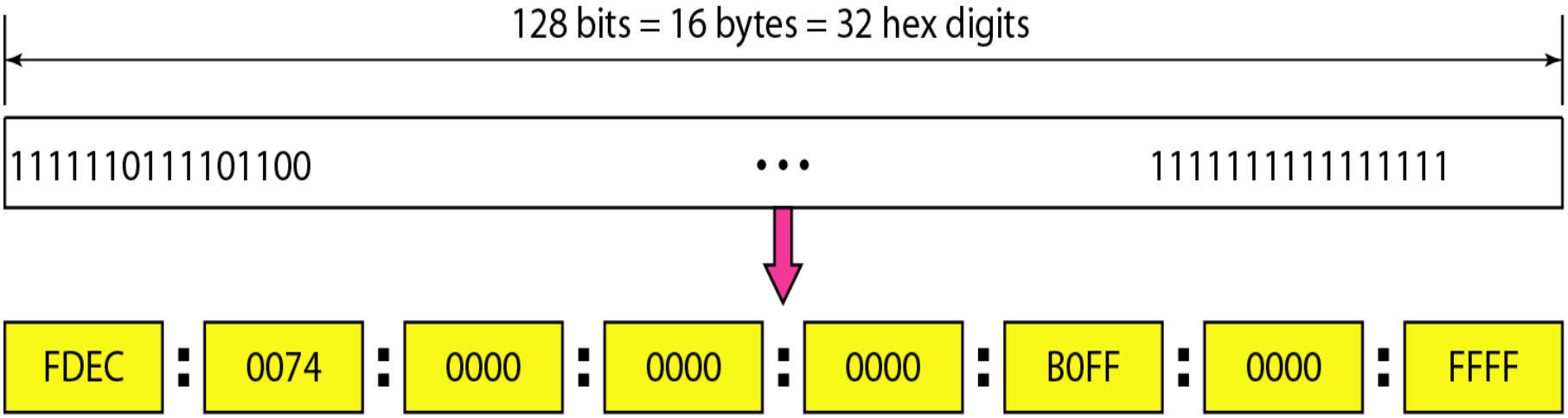
Table -*Priorities for noncongestion-controlled traffic*

Priority	*Meaning*
8	Data with greatest redundancy
. . .	. . .
15	Data with least redundancy

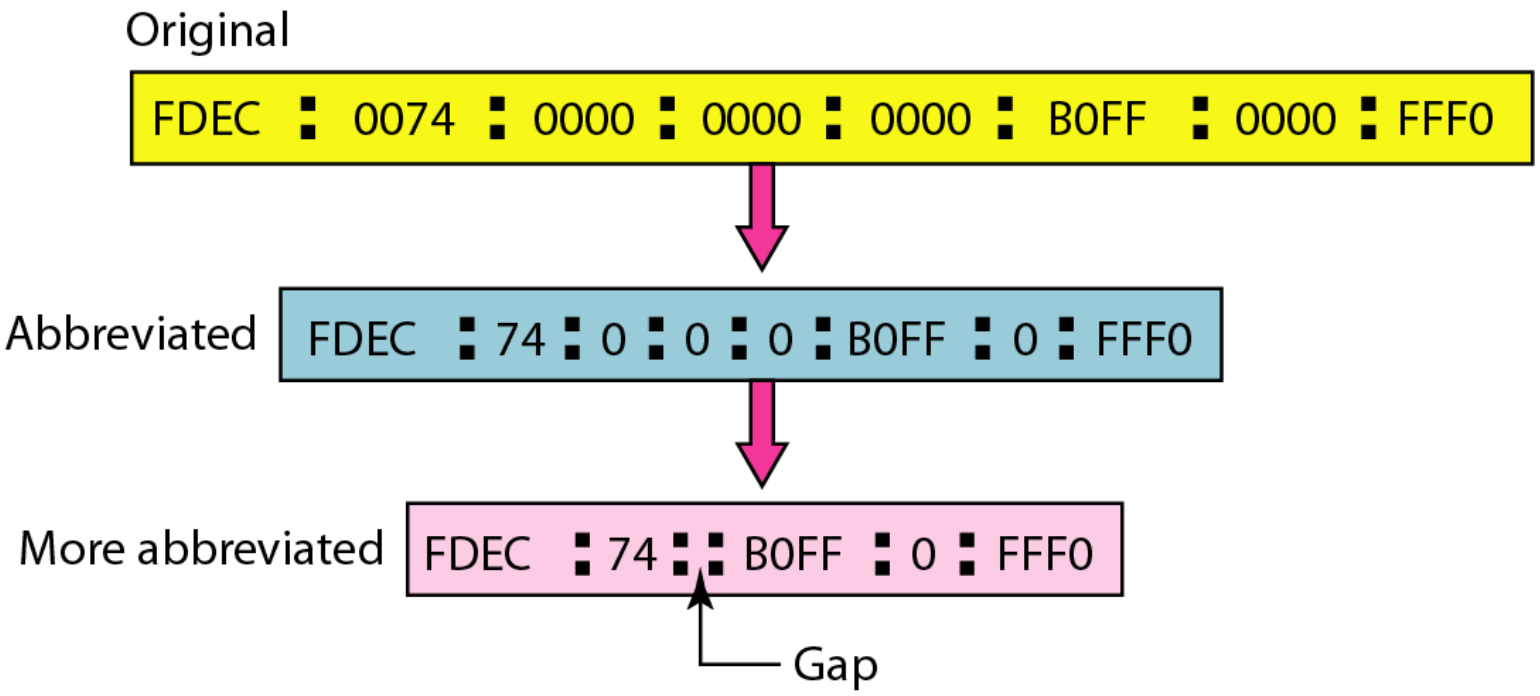
Next header codes for IPv6

Code	*Next Header*
0	Hop-by-hop option
2	ICMP
6	TCP
17	UDP
43	Source routing
44	Fragmentation
50	Encrypted security payload
51	Authentication
59	Null (no next header)
60	Destination option

IPv6 address in binary and hexadecimal colon notation



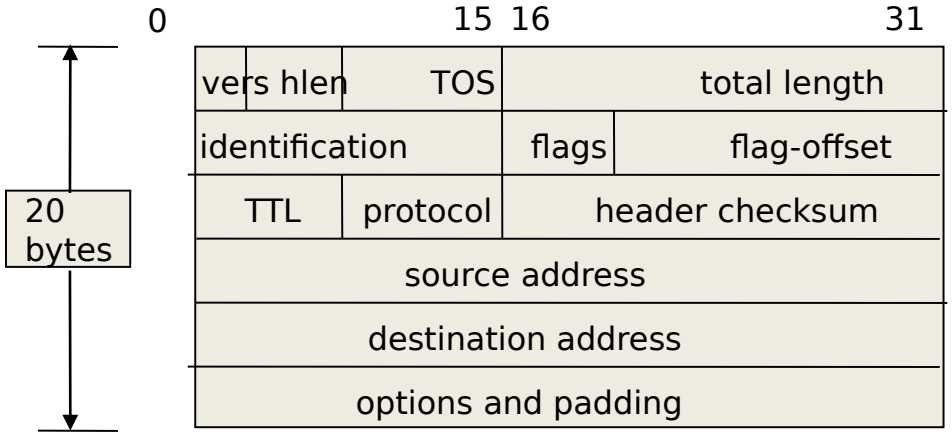
Abbreviated IPv6 addresses



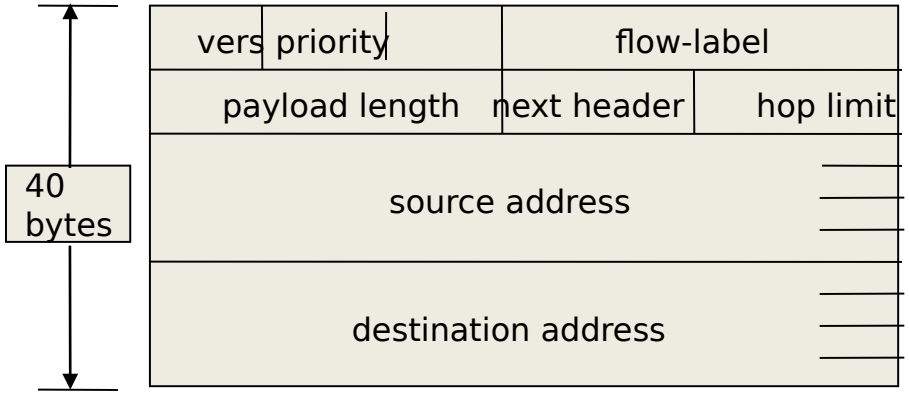
Comparison between IPv4 and IPv6 packet headers

Comparison
1. The header length field is eliminated in IPv6 because the length of the header is fixed in this version.
2. The service type field is eliminated in IPv6. The priority and flow label fields together take over the function of the service type field.
3. The total length field is eliminated in IPv6 and replaced by the payload length field.
4. The identification, flag, and offset fields are eliminated from the base header in IPv6. They are included in the fragmentation extension header.
5. The TTL field is called hop limit in IPv6.
6. The protocol field is replaced by the next header field.
7. The header checksum is eliminated because the checksum is provided by upper-layer protocols; it is therefore not needed at this level.
8. The option fields in IPv4 are implemented as extension headers in IPv6.

Header comparison



IPv4



IPv6

Removed (6)

- ID, flags, flag offset
- TOS, hlen
- header checksum

Changed (3)

- total length => payload
- protocol => next header
- TTL => hop limit

Added (2)

- priority
- flow label

Expanded

- address 32 to 128 bits

Major Improvements of IPv6 Header

- No option field: Replaced by extension header.
- Result in a fixed length, 40-byte IP header.
- No header checksum: Result in fast processing.
- No fragmentation at intermediate nodes: Result in fast IP forwarding.

IPv6 Address Space

- IPv6 address space is 2^{128}
- The designers of IPv6 divided the address into several categories.
- A few leftmost bits, called the *type prefix*, in each address define its category.
- The type prefix is variable in length, but it is designed such that no code is identical to the first part of any other code.

- ***Type prefixes for IPv6 addresses***

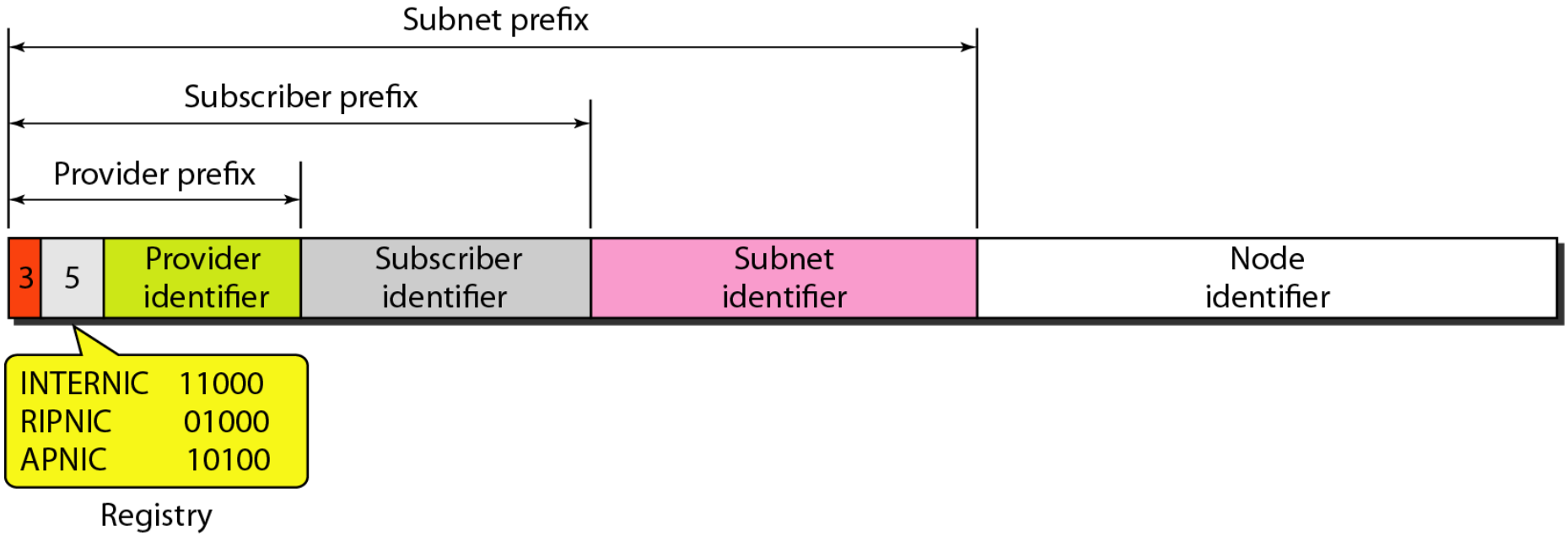
Type Prefix	*Type*	*Fraction*
0000 0000	Reserved	1/256
0000 0001	Unassigned	1/256
0000 001	ISO network addresses	1/128
0000 010	IPX (Novell) network addresses	1/128
0000 011	Unassigned	1/128
0000 1	Unassigned	1/32
0001	Reserved	1/16
001	Reserved	1/8
010	**Provider-based unicast addresses**	**1/8**

- ***Type prefixes for IPv6 addresses (continued)***

Type Prefix	*Type*	*Fraction*
011	Unassigned	1/8
100	Geographic-based unicast addresses	1/8
101	Unassigned	1/8
110	Unassigned	1/8
1110	Unassigned	1/16
1111 0	Unassigned	1/32
1111 10	Unassigned	1/64
1111 110	Unassigned	1/128
1111 1110 0	Unassigned	1/512
1111 1110 10	Link local addresses	1/1024
1111 1110 11	Site local addresses	1/1024
1111 1111	Multicast addresses	1/256

- *Unicast Addresses*
- - A unicast address defines a single computer.
- - IPv6 defines two types of unicast address:
- 1. Geographically based
- 2. Provider – based
- Provider based address is generally used by a normal host as a unicast address.

- ***Prefixes for provider-based unicast address***



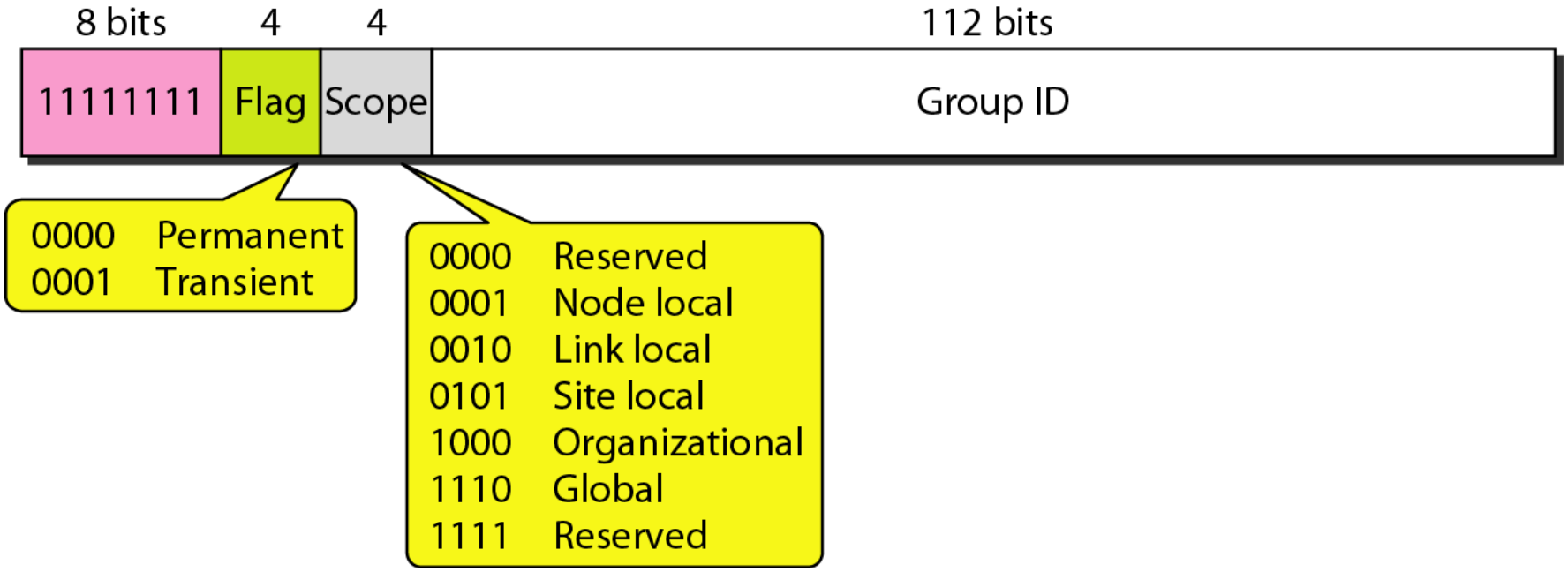
- **Type identifier** : (3 bits) Defines the address as a provider based address.
- **Registry identifier** : (5 bits) Indicates the agency that has registered the address. Currently 3 registry centers have been defined :
  - INTERNIC (11000) North America
  - RIPNIC (01000) European registration
  - APNIC (101000) Asian and Pacific countries

  - **Provider Identifier**: (16 bit) This variable length field identifies the provider for Internet access (such as an ISP)
  - **Subscriber Identifier** : (24 bit) When an organization subscribes to the Internet through a provider, it is assigned a subscriber identification.
  - **Subnet Identifier** : (32 bits) Each subscriber can have many different subnetworks and each subnetwork can have an identifier.
  - **Node Identifier** : (48 bits) Defines the identity of the node connected to a subnet.

Multicast Addresses

- A multicast addresses are used to define a group of hosts instead of just one.
- A packet sent to a multicast address must be delivered to each member of the group.

- ***Multicast address in IPv6***

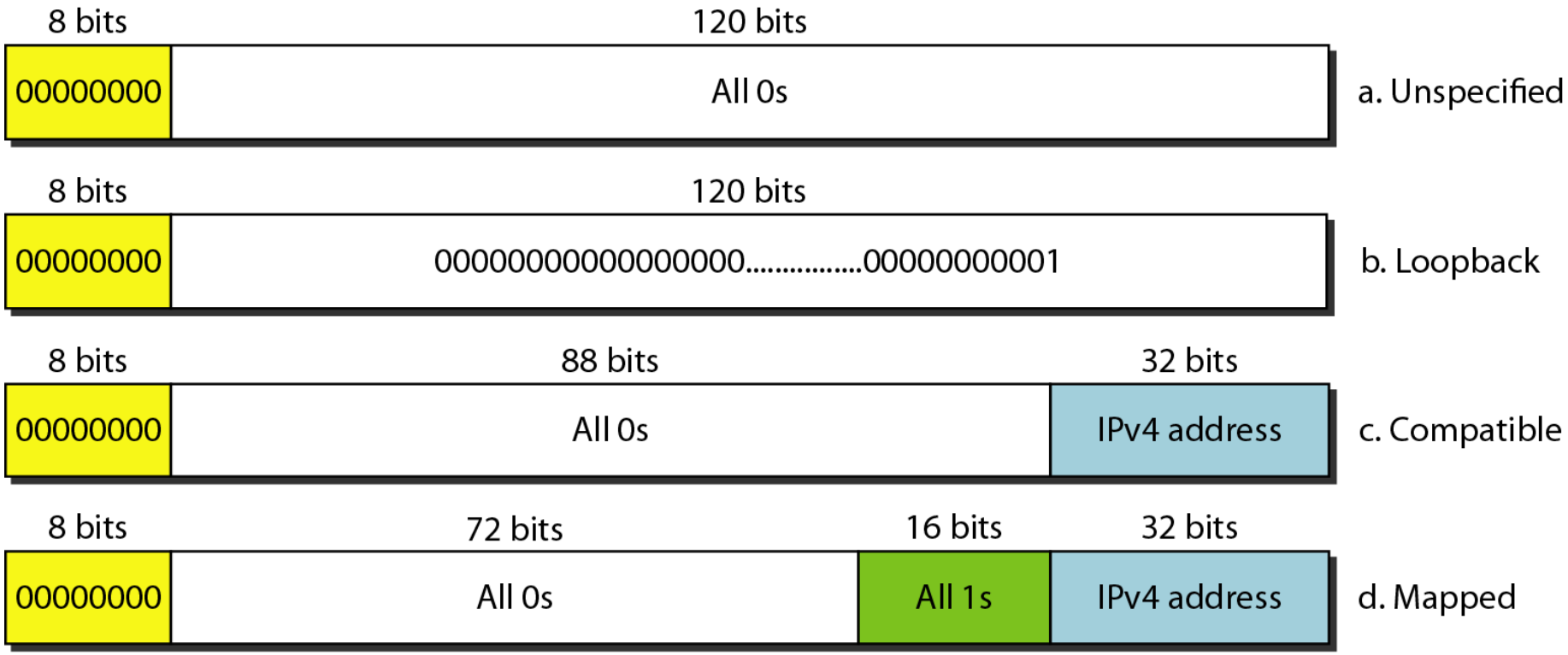


- The second field is a flag that defines the group address as either permenent or transient.
- A **permanent group** address is defined by the Internet authorities and can be accessed at all times.
- A **transient group** address, on the other hand, is used only temporarily.
- The **Scope** field defines the scope of the group address.

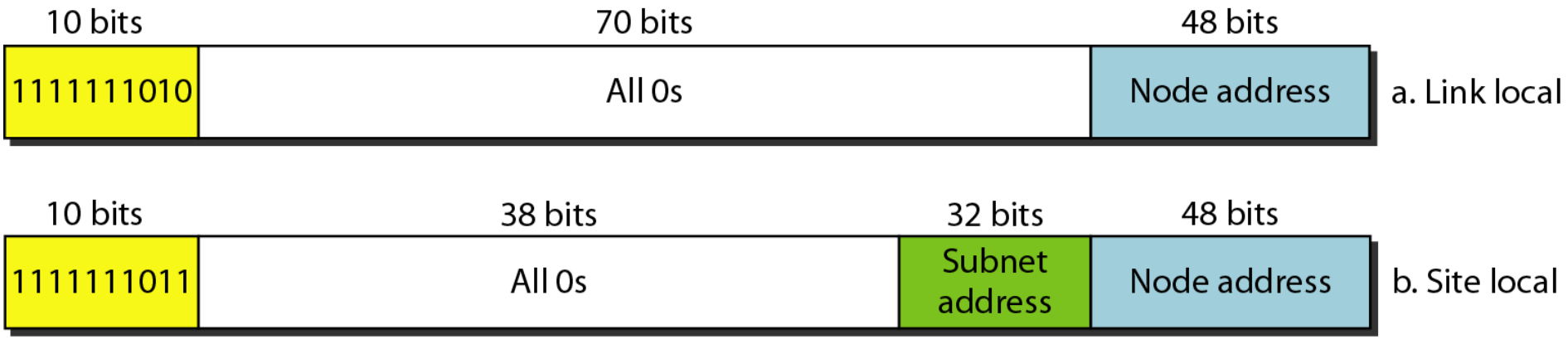
Reserved Addresses:

- **Unspecified address** is used when a host does not know its own address and sends an inquiry to find its address.
- A **Loopback address** is used by a host to test itself without going into the n/w.
- A **Compatible address** is used during the transition from Ipv4 to Ipv6.
- A **mapped address** is also used during transition from Ipv6 to Ipv4..

- ***Reserved addresses in IPv6***



- ***Local addresses in IPv6***



- These addresses are used when an organization wants to use Ipv6 protocol without being connected to the global Internet.
- They provide addressing for private n/w.

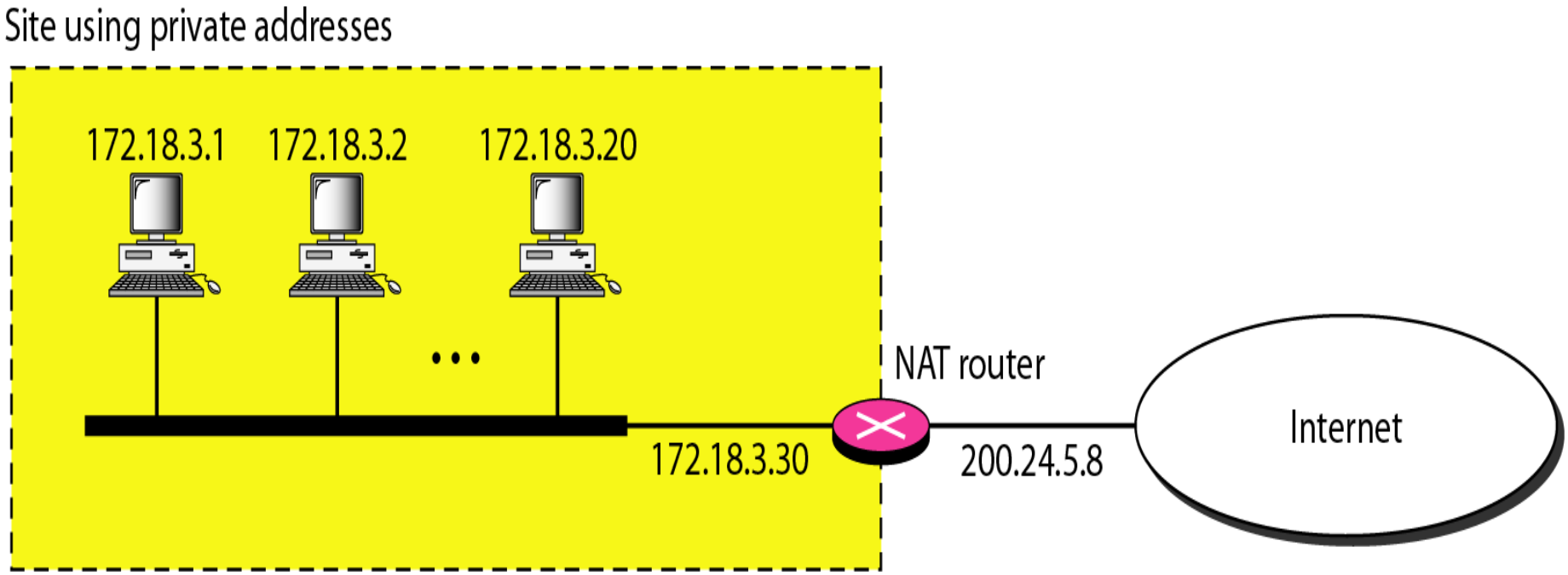
NAT (Network Address Translation)

NAT enables a user to have a large set of addresses internally (private addresses) and one address or a small set of address (global / public address) externally.

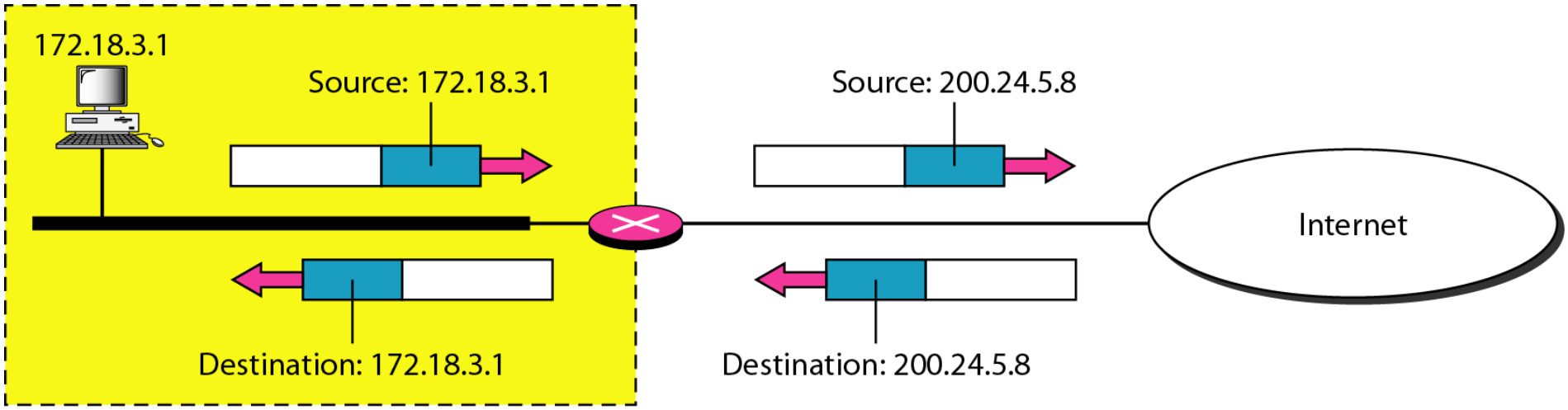
Reserved addresses for private networks

Range			*Total*
10.0.0.0	to	10.255.255.255	2^{24}
172.16.0.0	to	172.31.255.255	2^{20}
192.168.0.0	to	192.168.255.255	2^{16}

A NAT implementation



Addresses in a NAT

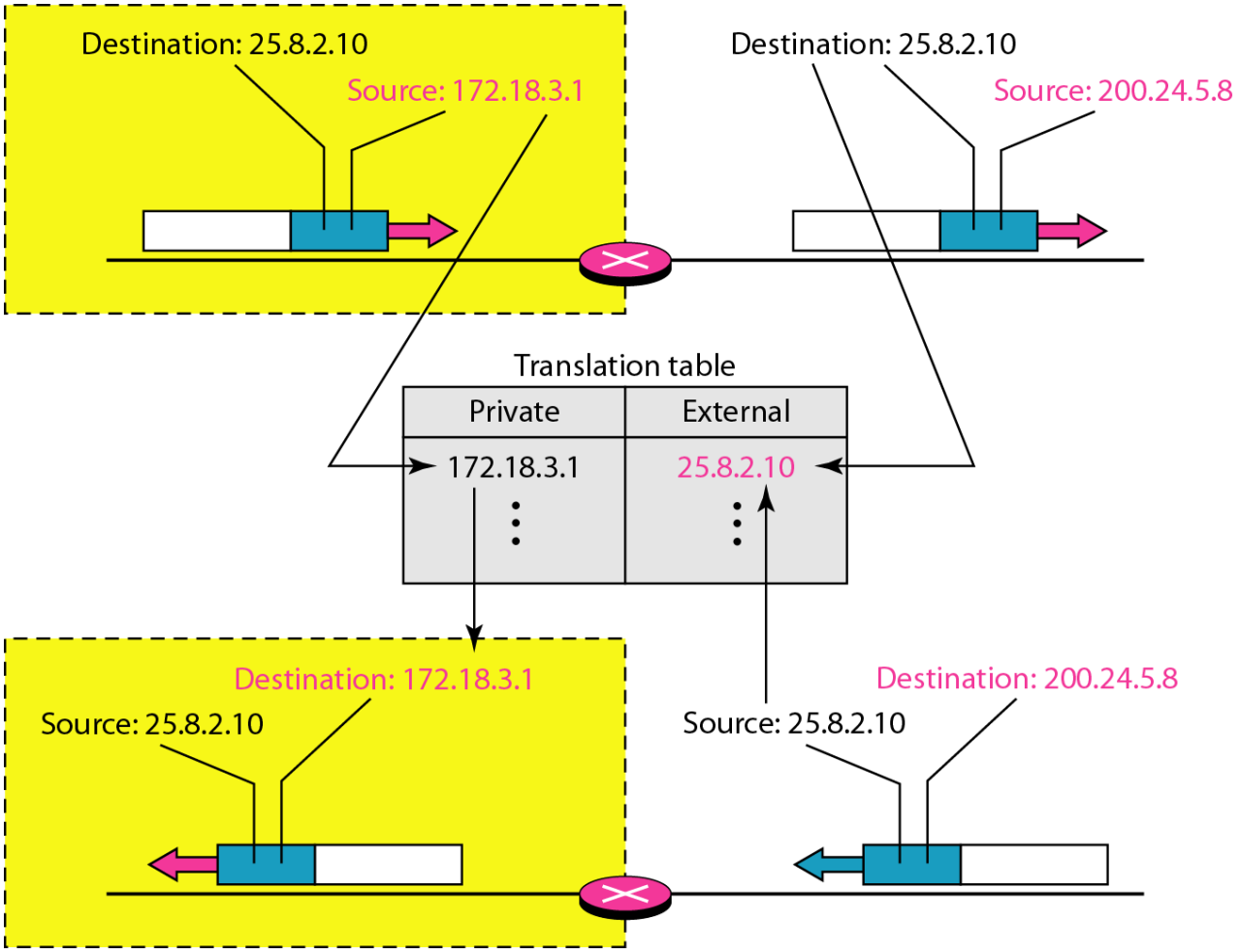


Translation Table:

How does the NAT router know the destination address for a packet coming from the Internet? This problem is solved by using Translation Table.

1. Using One IP Address.
2. Using a Pool of IP Addresses.
3. Using both IP addresses and Port Numbers.

1. Using One IP address NAT address translation:



3. Using Both IP and Port numbers:

Combination of Source address and Destination Port number defines private n/w host.

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
. . .	. . .	. . .	. . .	. . .

ICMP (Internet Control Message Protocol)

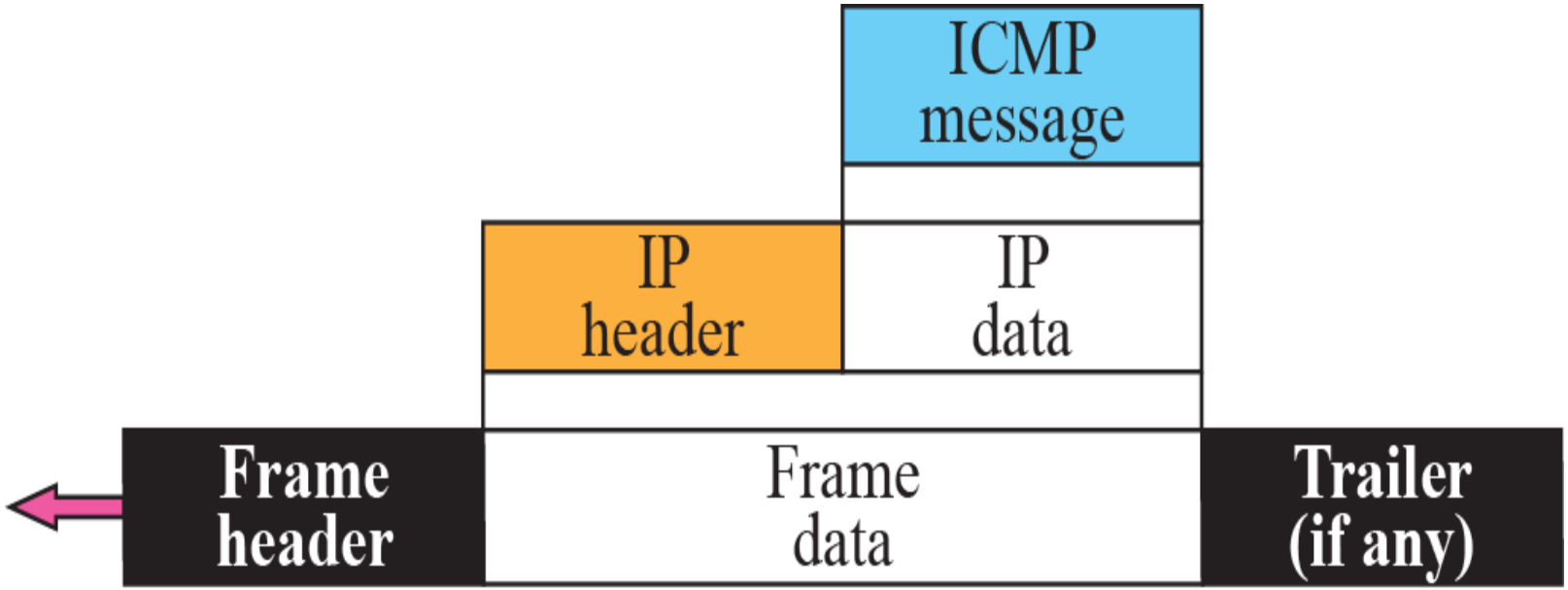
The IP protocol has no error-reporting or error-correcting mechanism. The IP protocol also lacks a mechanism for host and management queries. The Internet Control Message Protocol (ICMP) has been designed to compensate for the above two deficiencies. It is a companion to the IP protocol.

Topics discussed in this section:

Types of Messages
Message Format
Error Reporting and Query
Debugging Tools

- **ICMP allows router to send error or control messages to the host.**
- **It is a communication between the IP software of two machines.**
- **ICMP is a error reporting mechanism not to correct them.**
- **ICMP messages are encapsulated in to IP packet.**
- **ICMP send msg to source only not to intermediate routers.**

ICMP encapsulation



Types Of Messages

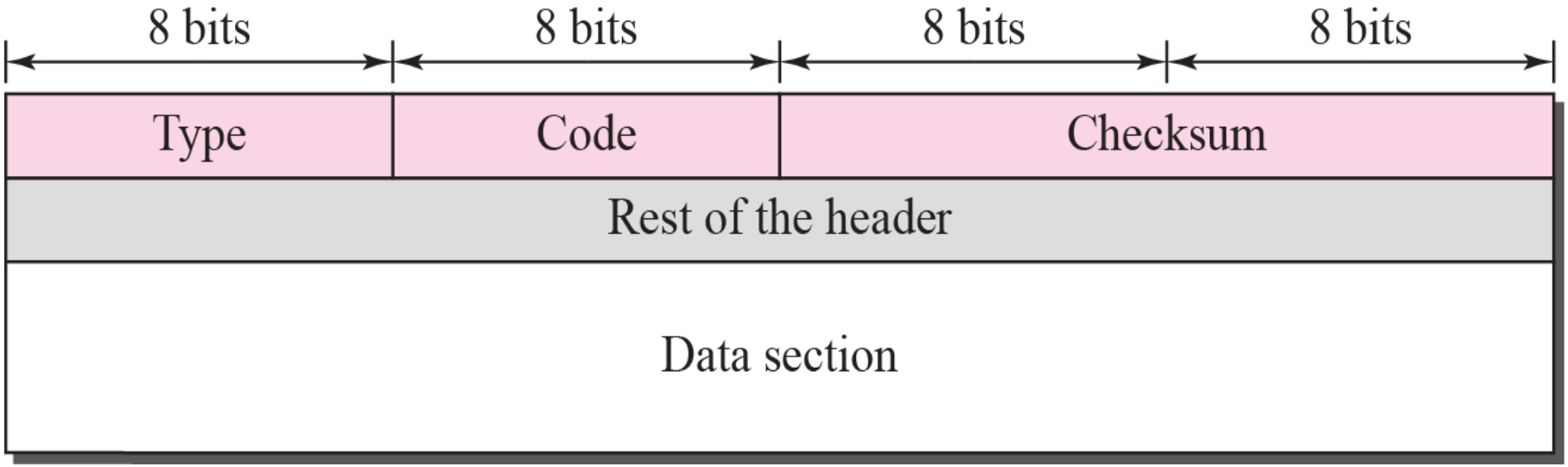
ICMP messages are divided into two broad categories: error-reporting messages and query messages.

The **error-reporting messages** report problems that a router or a host (destination) may encounter when it processes an IP packet.

The **query messages,** which occur in pairs, help a host or a network manager get specific information from a router or another host.

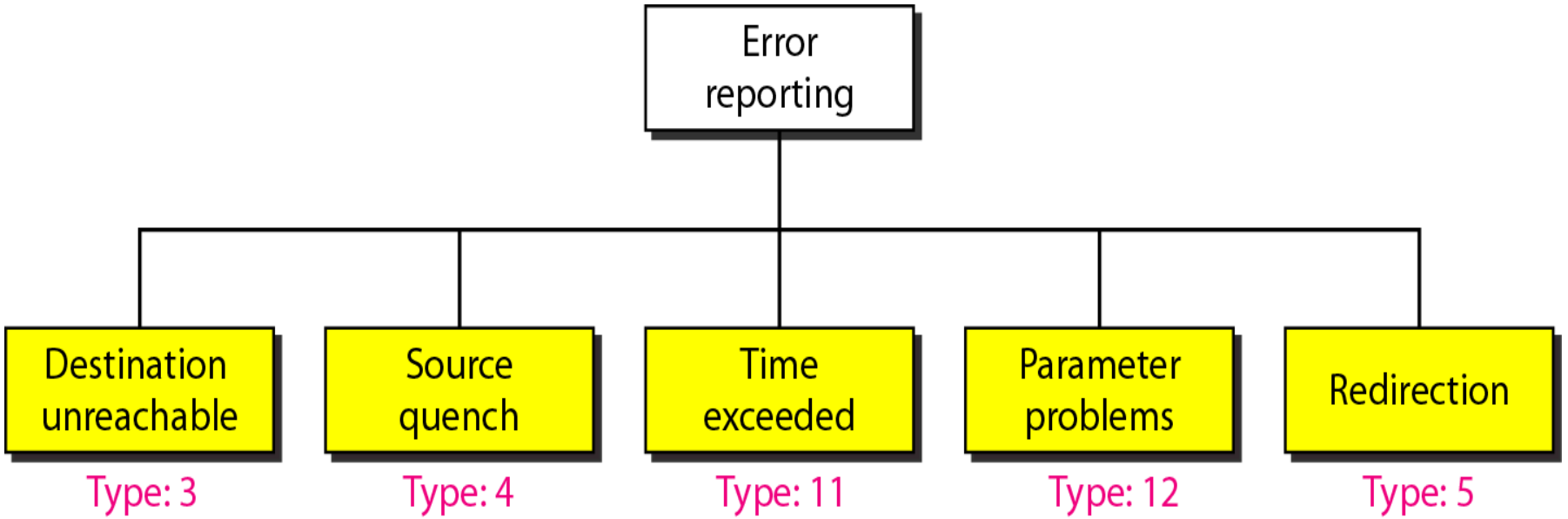
Message Format

General format of ICMP messages



- **Type (8 bits):** Used to identify type of ICMP message i.e. Whether the msg is error reporting or query msg.
- **Code (8 bits):** Provides information or parameters of msg type.
- **Checksum (16 bits):** Gives checksum of the ICMP msg.
- **Data section (64 bits):** IP header and original datagram.
- **Rest of the Header:** Unused

Types of Messages



Note

ICMP always reports error messages to the original source.

1. Destination-unreachable

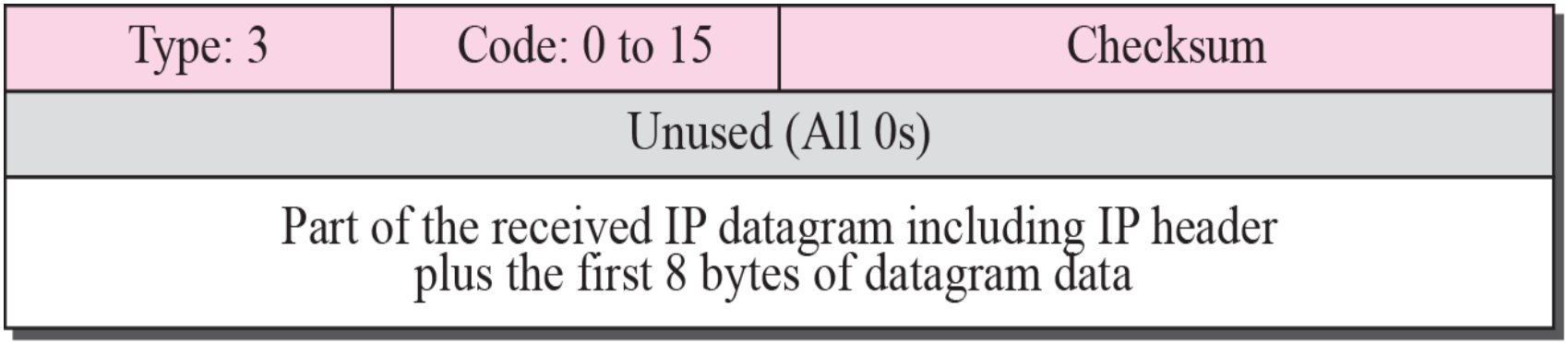
- Packet is not forwarded to destination due to some problems.

Destination-unreachable messages with codes 2 or 3 can be created by either a router or the destination host.

- The router or the host sends a destination- unreachable message back to the source host only.

When a router cannot forward or deliver an IP datagram, it sends a destination unreachable message back to the original source.

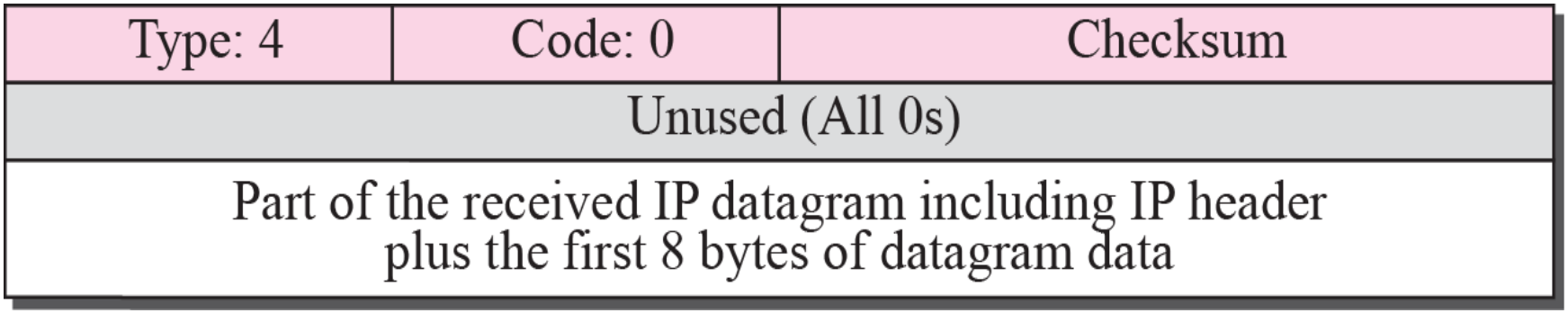
The CODE field specifies details
0: network unreachable
1: host unreachable
2: protocol unreachable
3: port unreachable
4: fragmentation needed and DF (don't fragment) set
5: source route failed
Etc.



2. Source-quench

A source-quench message informs the source that a datagram has been discarded due to congestion in a router or the destination host.

The source must slow down the sending of datagrams until the congestion is relieved.

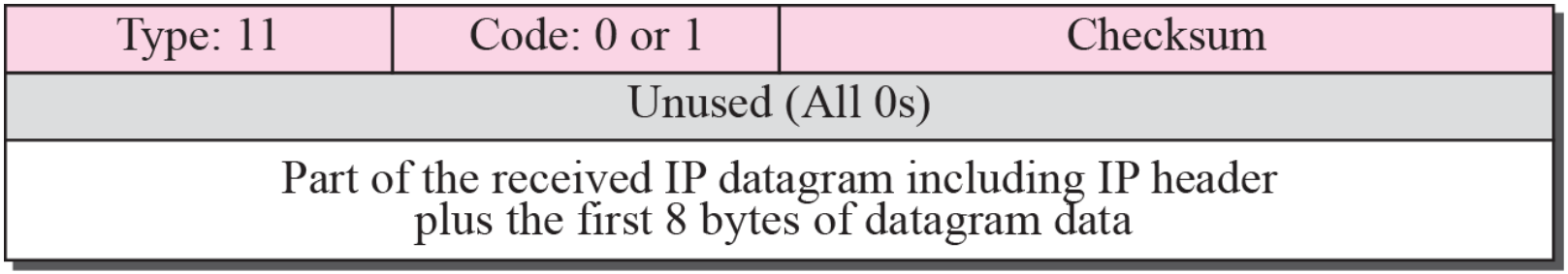


One source-quench message is sent for each datagram that is discarded due to congestion.

3. Time-exceeded

Whenever a router decrements a datagram with a time-to-live value to zero, it discards the datagram and sends a time-exceeded message to the original source.

When the final destination does not receive all of the fragments in a set time, it discards the received fragments and sends a time-exceeded message to the original source.



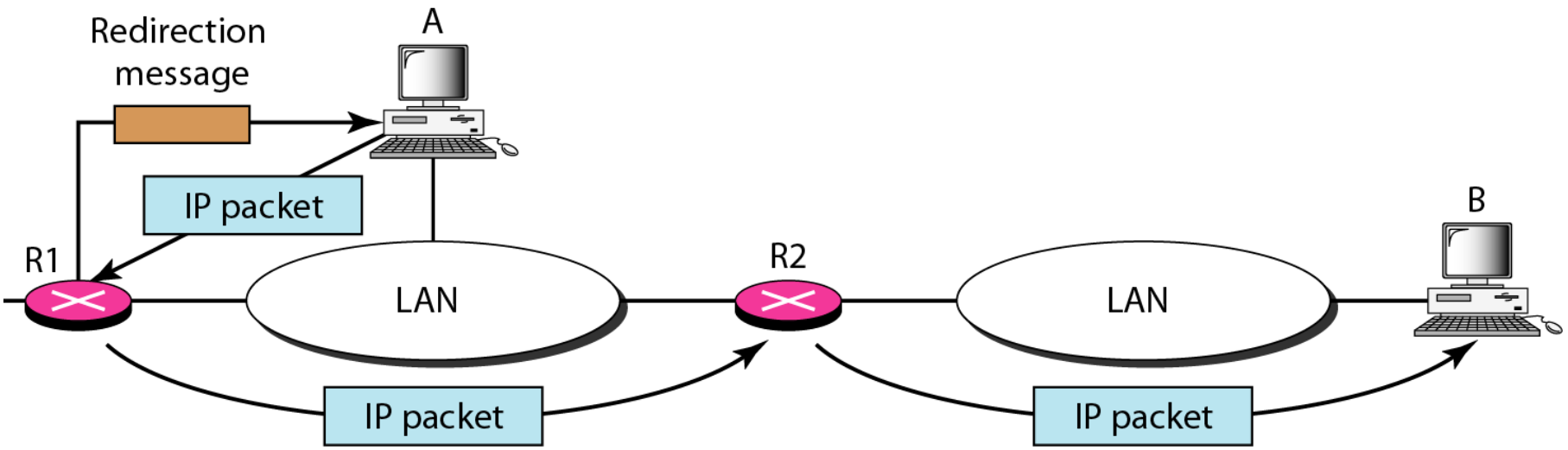
In a time-exceeded message, code 0 is used only by routers to show that the value of the time-to-live field is zero.

Code 1 is used only by the destination host to show that not all of the fragments have arrived within a set time.

4. Parameter-problem

When there is a problem in the header part of the datagram then a parameter-problem message can be created by a router or the destination host.

5. Redirection concept

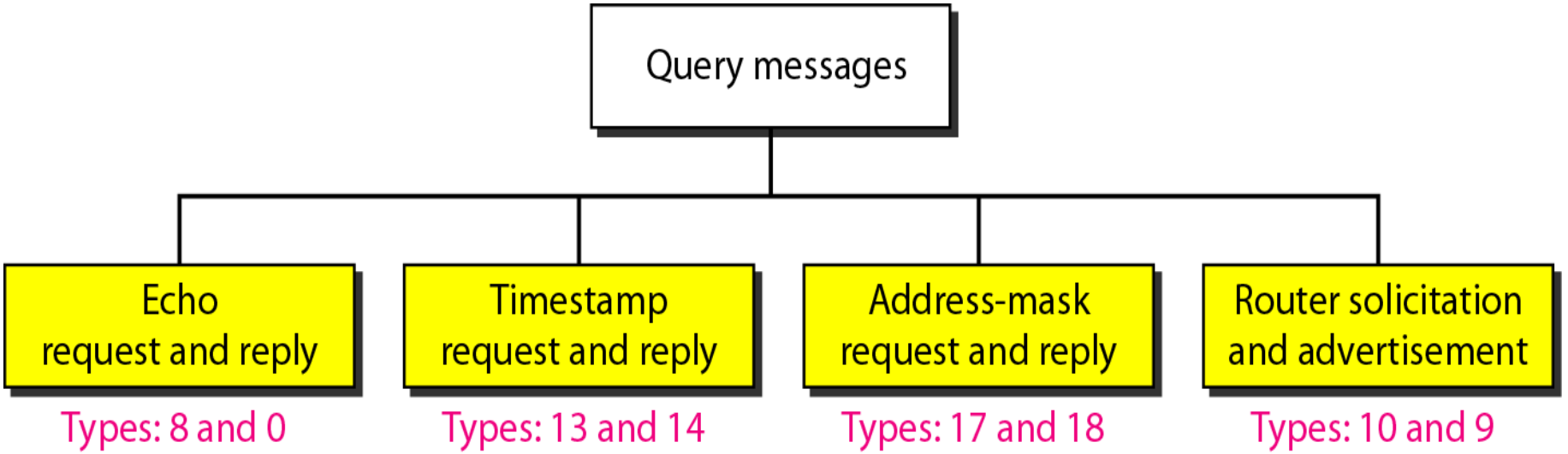


A host usually starts with a small routing table that is gradually updated.

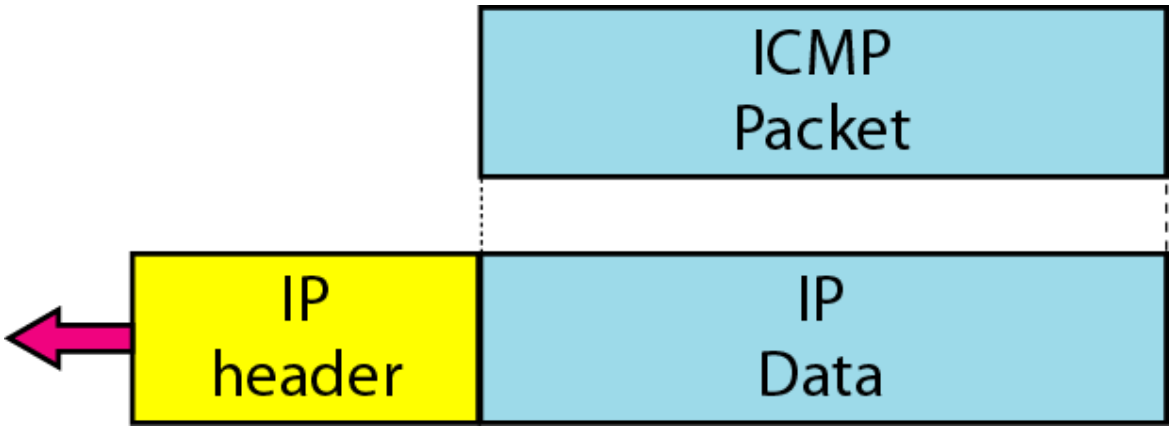
One of the tools to accomplish this is the redirection message.

A redirection message is sent from a router to a host on the same local network.

Query messages



Encapsulation of ICMP query messages:



1. Echo-request and reply message

Echo-request and echo-reply messages can test the reachability of a host.

An echo-request message can be sent by a host or router.

An echo-reply message is sent by the host or router that receives an echo-request message.

This is usually done by invoking the ping command.

2. Timestamp-request and timestamp-reply message

Type 13: request
Type 14: reply

Type: 13 or 14	Code: 0	Checksum
Identifier		Sequence number
Original timestamp		
Receive timestamp		
Transmit timestamp		

Timestamp-request and timestamp-reply messages can be used to calculate the round-trip time between a source and a destination machine. It can also used to synchronized the clocks in two machines.

3. Address mask request and reply

When source knows his IP address but dont know its Subnet mask that time source send Address mask request to router.

Router receive request and send subnet mask address to source.

At the time of request Address mask field is zeros and the time of reply its contain Address mask.

4. Router Solitation and Advertisement

A host want to send data to another network host for that he also know the Address of the routers connected to its network.

For that the host broadcast a router solicitation message.

The router receiving this message send the routing information using router Advertisement message.

A router can also periodically sends router advertisement messages even if no host has solicited.

UNICAST ROUTING PROTOCOLS

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

Intra- and Interdomain Routing :

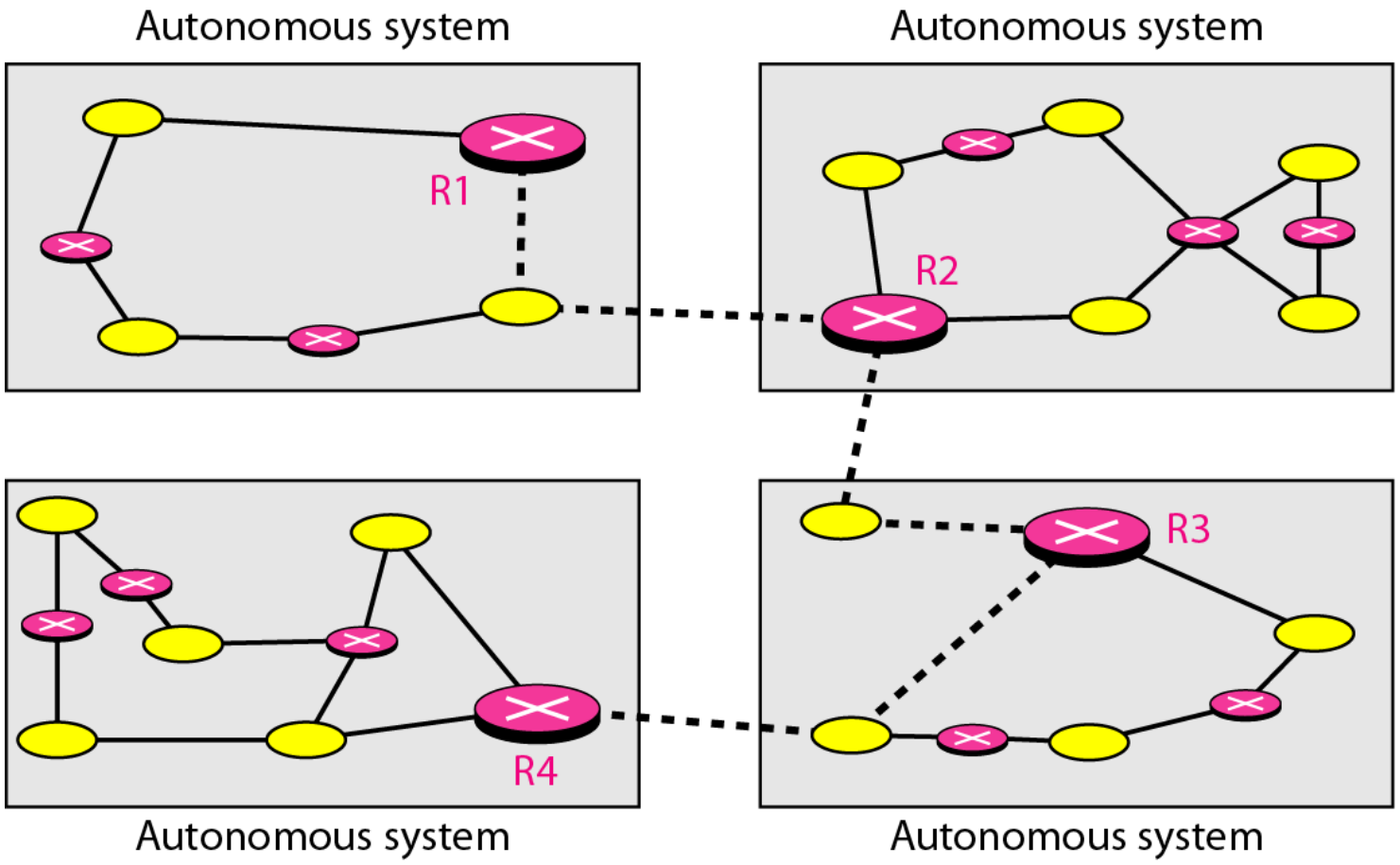
Today, an internet can be so large that one routing protocol cannot handle the task of updating the routing tables of all routers.

For this reason, an internet is divided into autonomous systems.

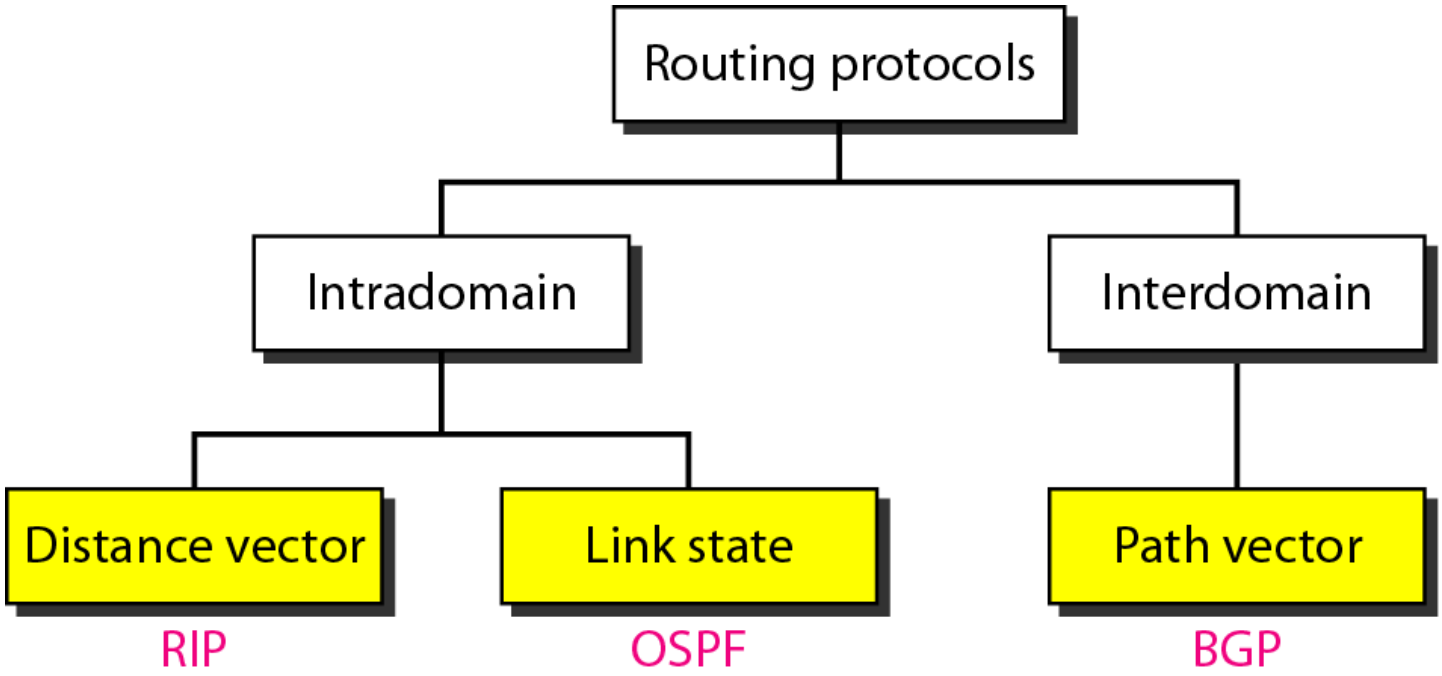
An **autonomous system (AS)** is a group of networks and routers under the authority of a single administration.

- Routing inside an autonomous system is referred to as **intradomain routing.**
- Routing between autonomous systems is referred to as **interdomain routing.**

- ***Autonomous systems***



- ***Popular routing protocols***



Dynamic Routing Algorithms: (Adaptive)

- **Distance Vector Routing Algorithm.**
- **Link State Routing Algorithm.**

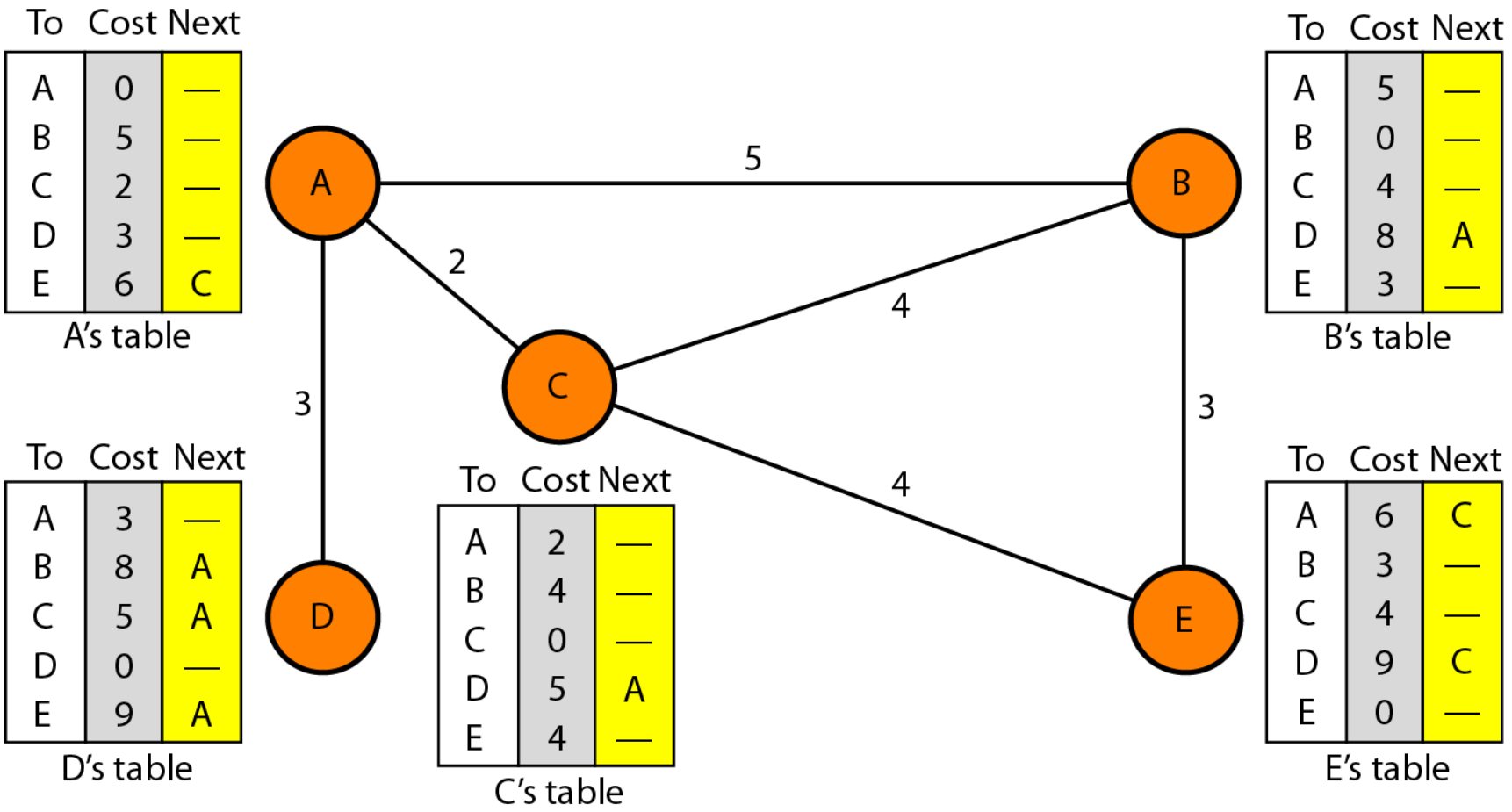
1.Distance Vector Routing:

In it, a node tells its neighbors about its distance to every other node in the network.
In it, the least cost route between any two nodes is the route with minimum distance.
In it, each node maintains a table (vector table) of minimum distance to every node.

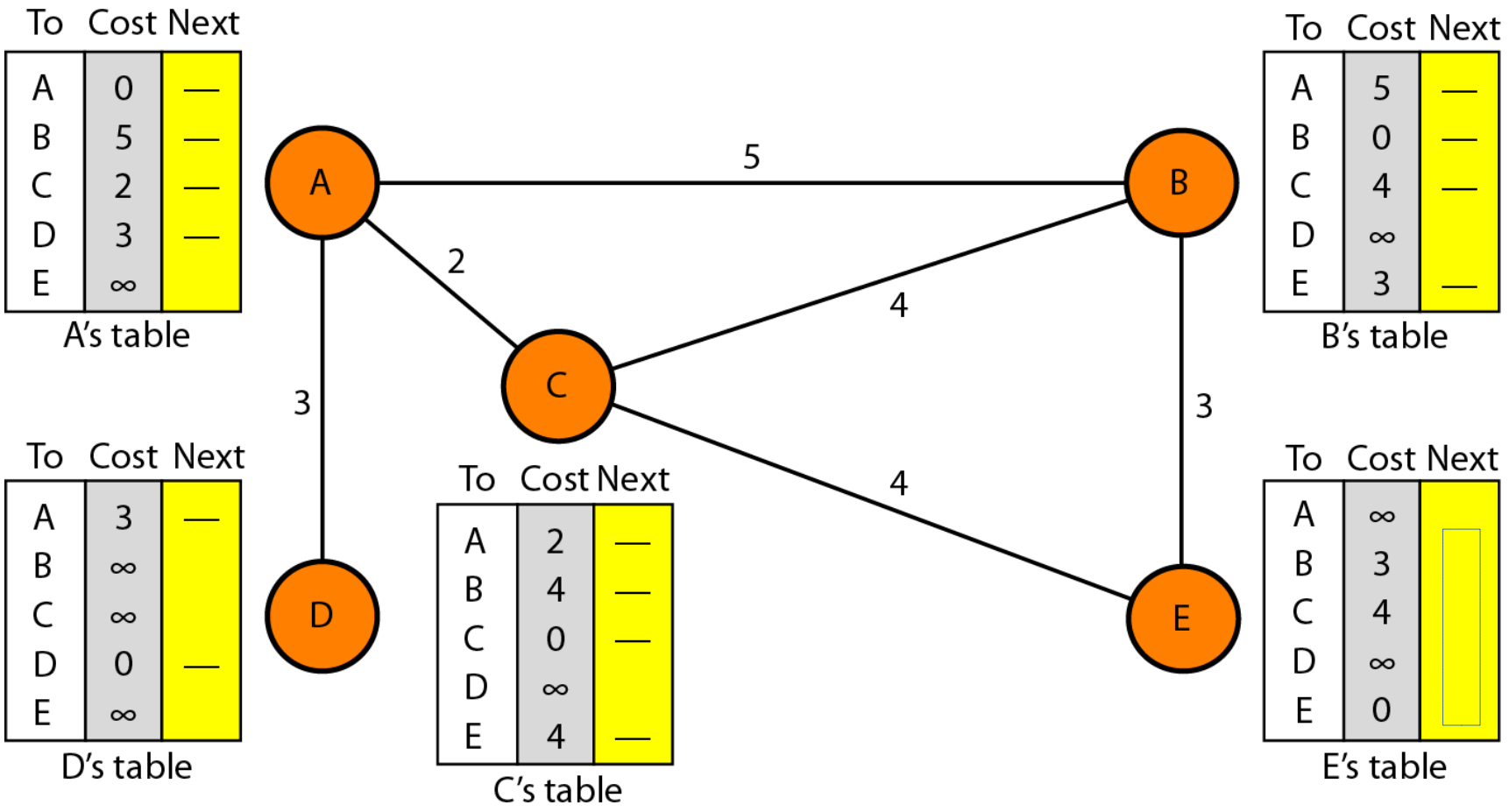
This algorithm follow 3 steps:

2.1. Initialization.
3.2. Sharing.
4.3. Updating.

Distance vector routing tables



Initialization of tables in distance vector routing



Sharing

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

It shared whole table with neighbors but some time shared only first two columns of the vector table.

Updating

Updating takes 3 steps:

1.The receiving node needs to add the cost between itself and the sending node to each value in the second column i.e.

C to Destination = x and C to A = y then

A to destination (via C) = x+y

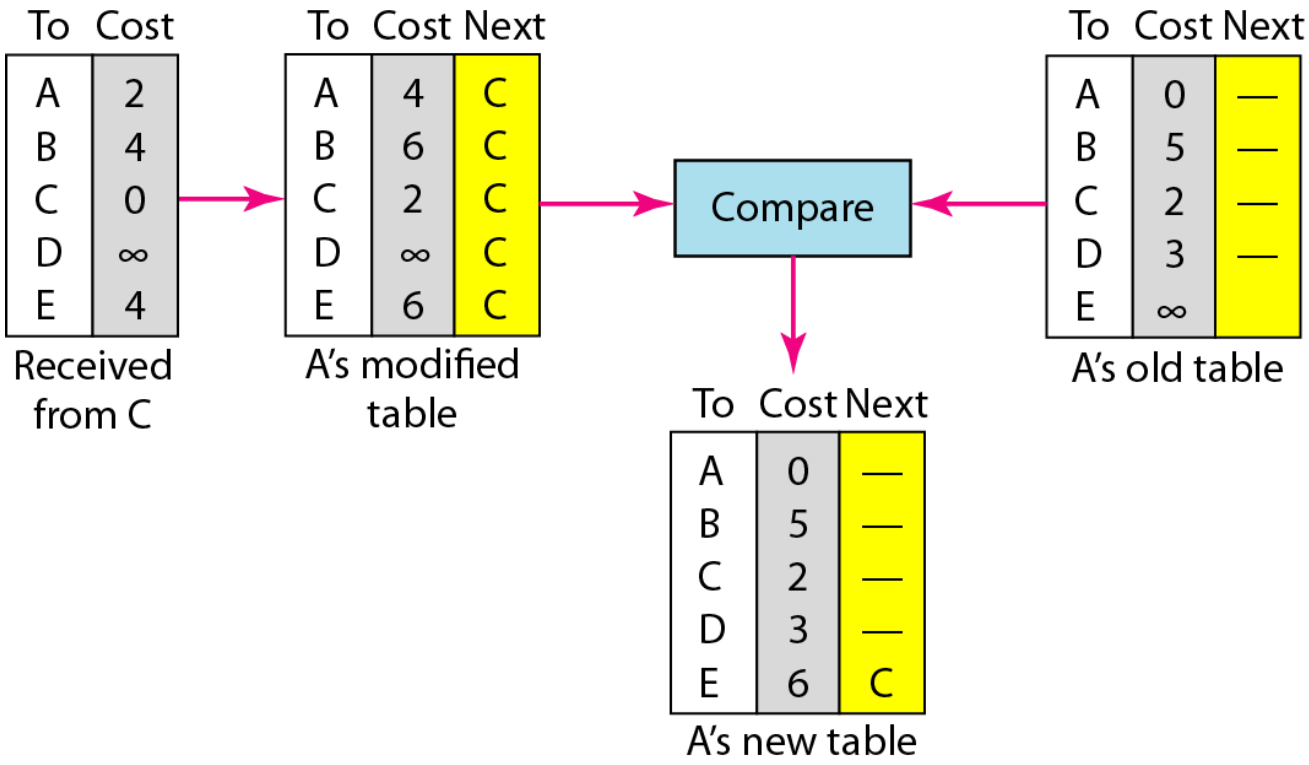
2. The receiving node needs to add the name of the sending node to each row as the third column.

3. Receiving node compare each row of old table with modified version of the received table.

a. If the next node entry is different then chooses the smaller cost row.

b. If the next node entry is same then chooses the new row.

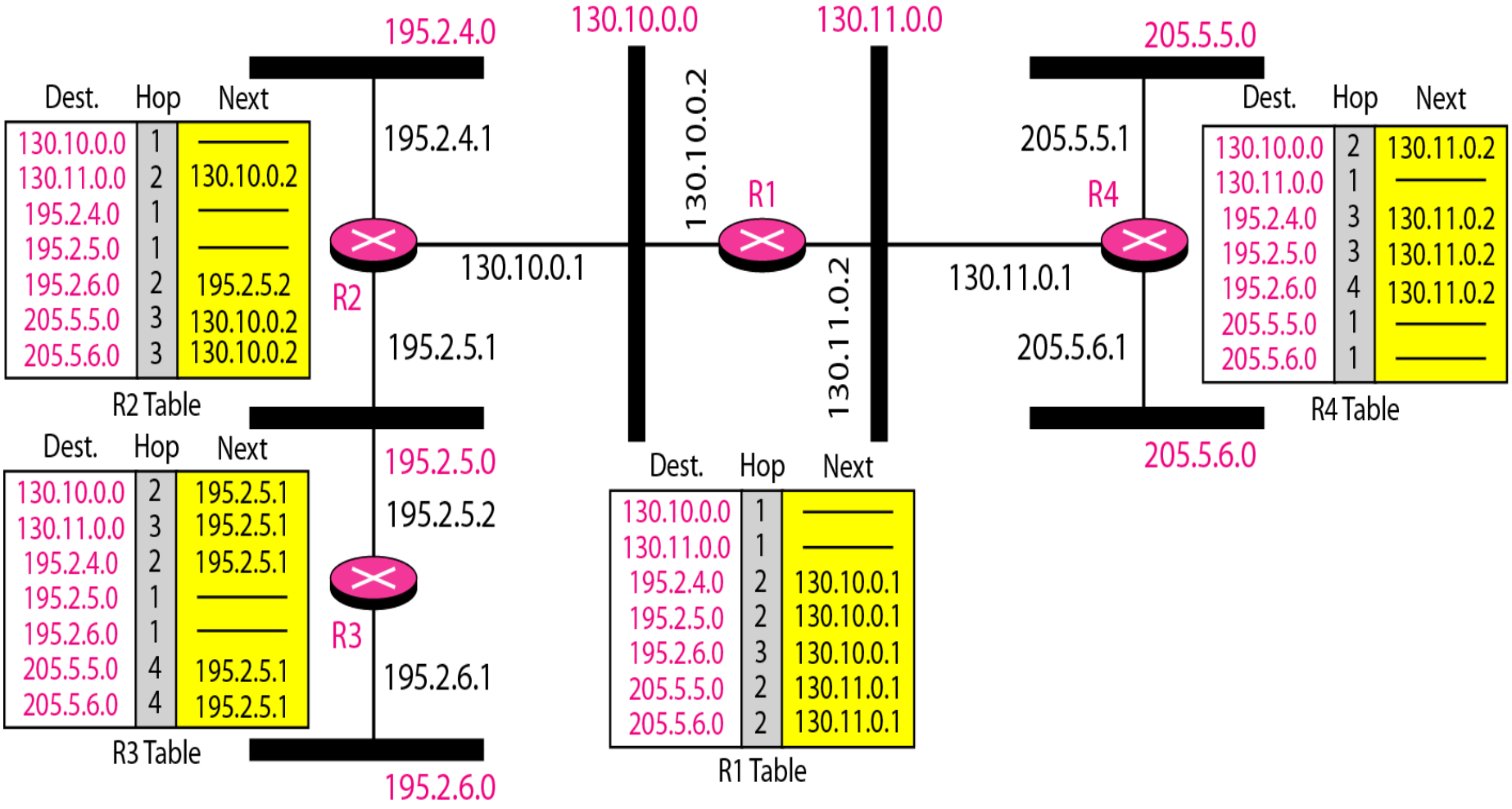
Updating in distance vector routing



Routing Information Protocol

- **RIP is a Intradomain routing protocol.**
- **Used in Autonomous system.**
- **RIP implements distance vector routing directly with some considerations:**
  - *Routers have routing tables, n/w do not.*
  - *Destination in a routing table is a n/w.*
  - *Distance is defined as the number of links(n/w's) to reach the destination (Hop Count).*
  - *Any route in autonomous sys. using RIP cannot have more than 15 hops.*
  - *Next node columns defines the address of next router.*

Example of a domain using RIP



Packet Format

Command (8 bit): Specify types of msg i.e.

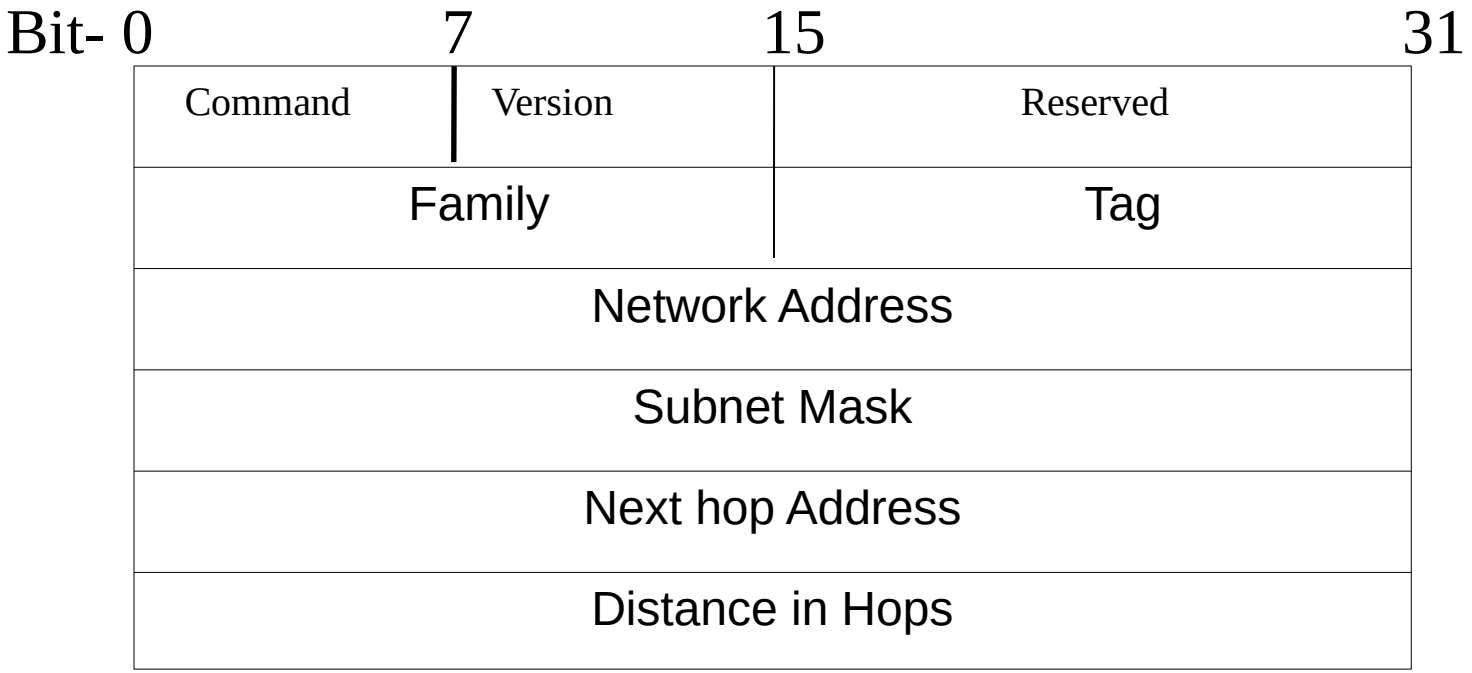
1= Request 2= Response

Version (8 bit): Defines version i.e. RIPv1 or RIPv2, current ver2.

Family (16 bit): Gives family of the protocol.Eg. TCP/IP = 2.

Tag : Information about Autonomous System

IP address(32 bit): Gives Destination address.



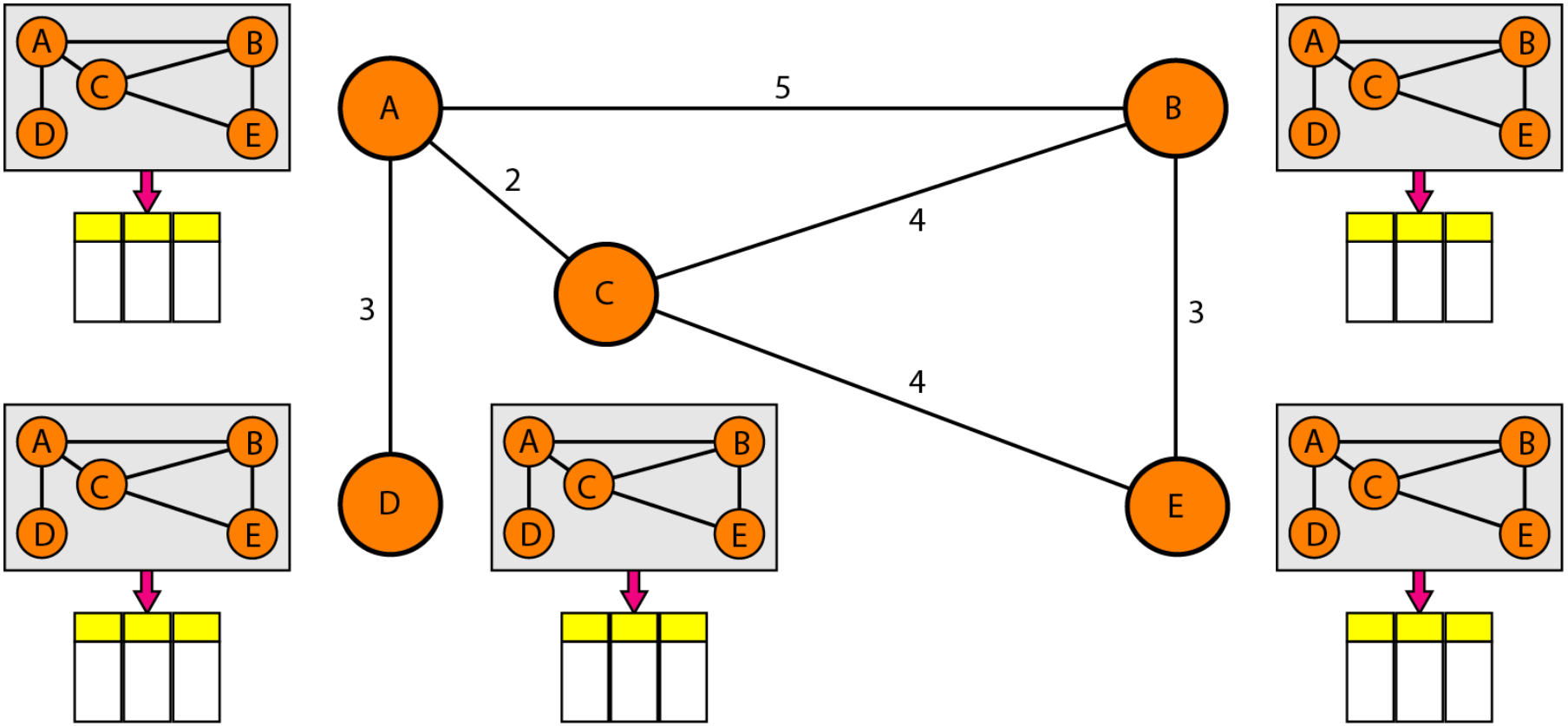
Link State Routing

Link state routing is based on the assumptions that, although the global knowledge about the topology is not clear, each node has partial knowledge: it knows the state (type, conditions and cost) of its links. It create Least cost tree using least cost table and Dijakstra Algo.

5 Operations:

1. Learning about the neighbors.
2. Measuring Line Cost.
3. Building Link State Packets (LSP).
4. Distributing the link state packets.
5. Computing the new routes.

Concept of link state routing



Link state knowledge

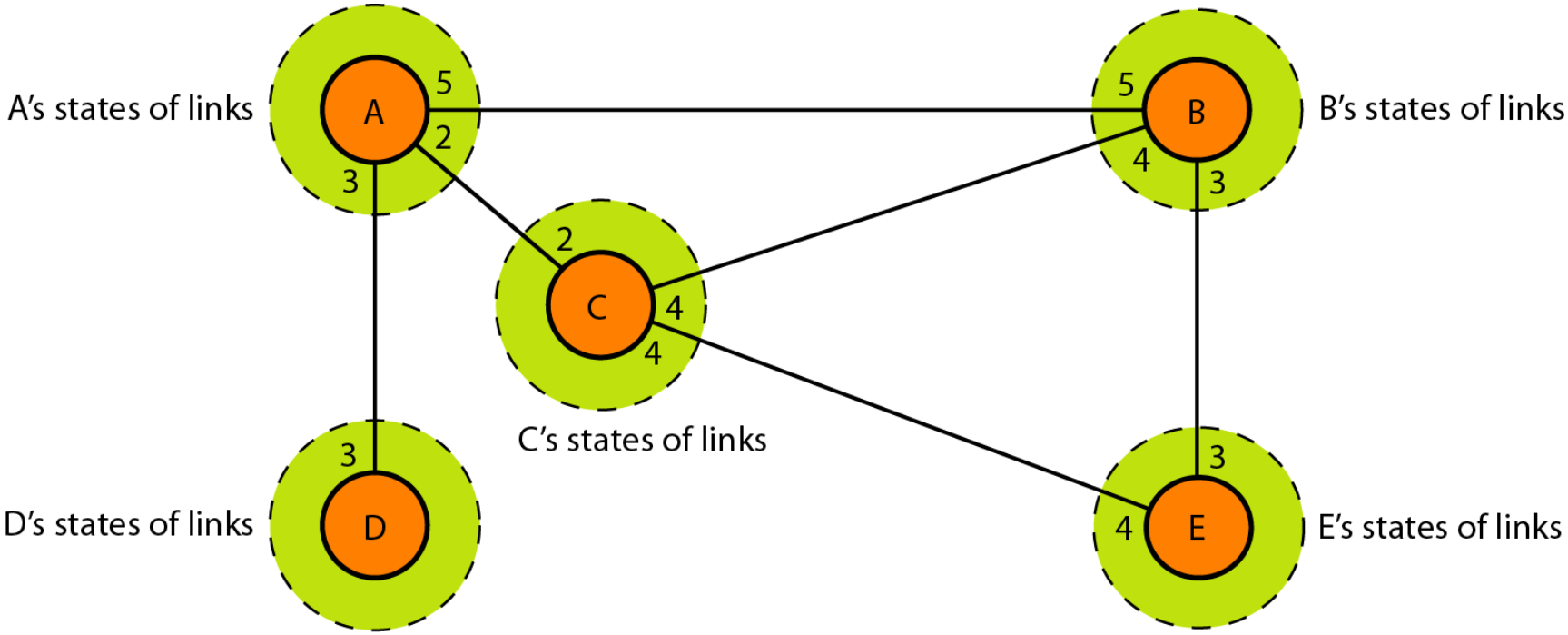
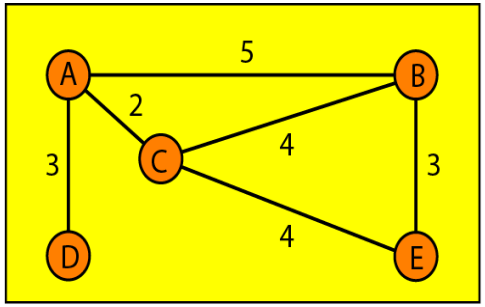
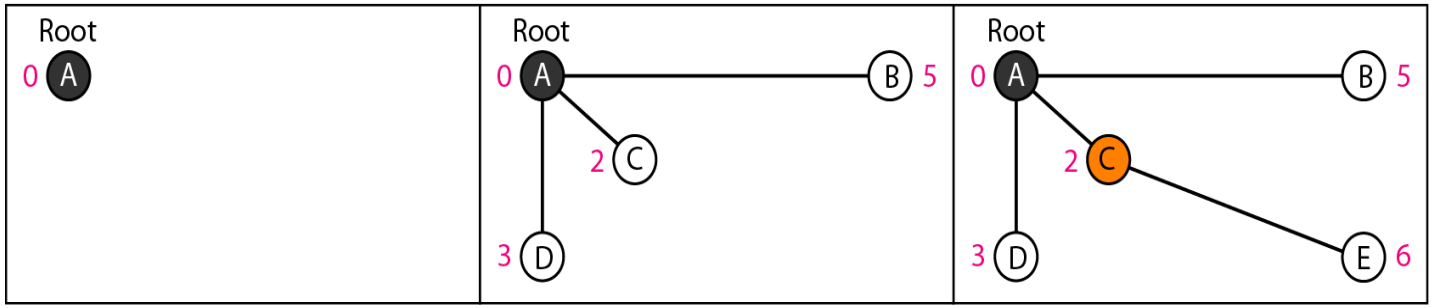


Fig. *Example of formation of shortest path tree*



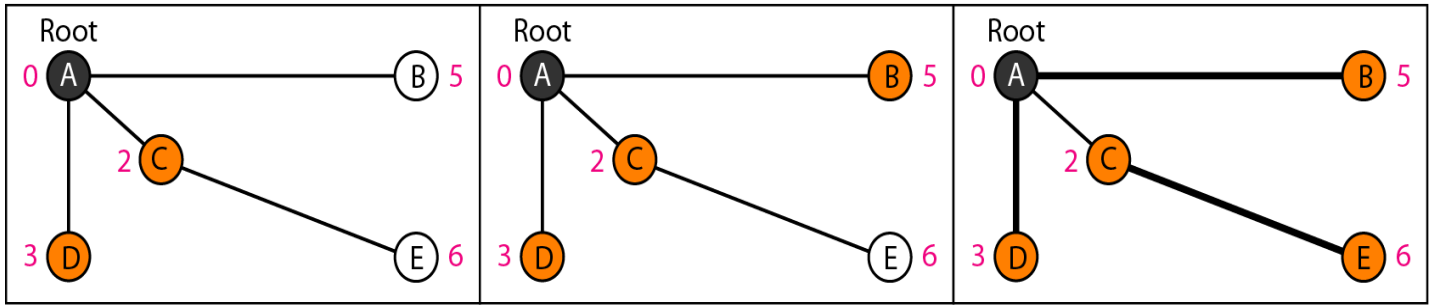
Topology



1. Set root to A and move A to tentative list.

2. Move A to permanent list and add B, C, and D to tentative list.

3. Move C to permanent and add E to tentative list.



4. Move D to permanent list.

5. Move B to permanent list.

6. Move E to permanent list (tentative list is empty).

OSPF (Open Shortest Path First)

- **OSPF is a Intradomain routing protocol.**

- **Based on Least Cost Tree.**
- **Uses Link state routing algo. i.e. Dijkstra's algo.**

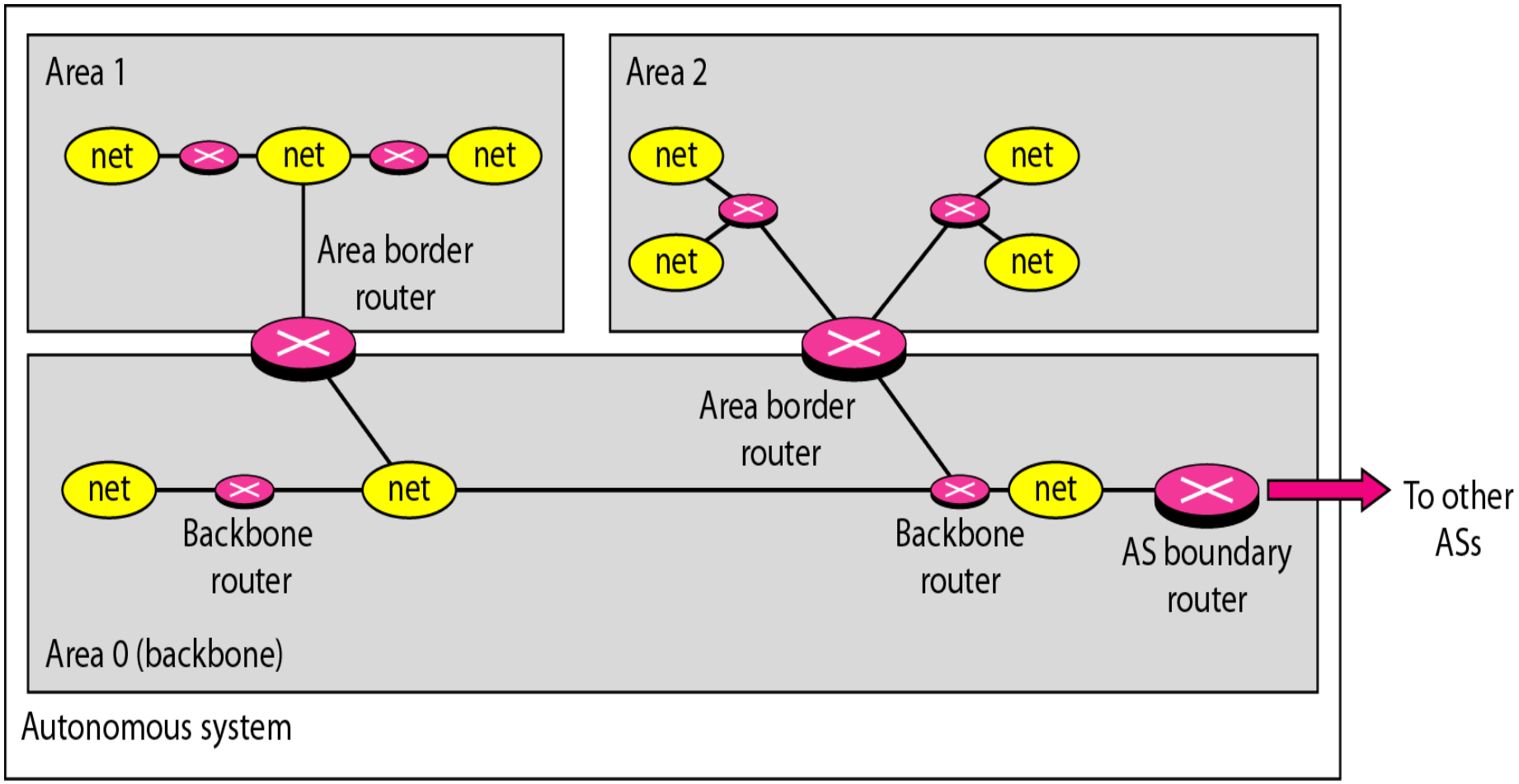
- **Used in Autonomous system.**

  - **Area's.**

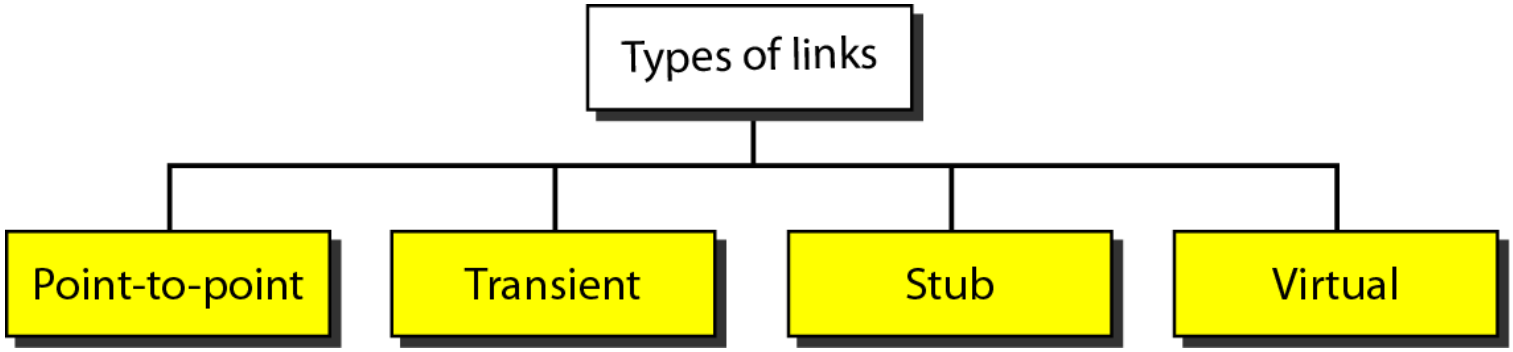
  - **Backbone: All the areas inside the AS must be connected to the backbone.**

  - **Backbone routers.**

Areas in an autonomous system



Types of links



Features of OSPF:

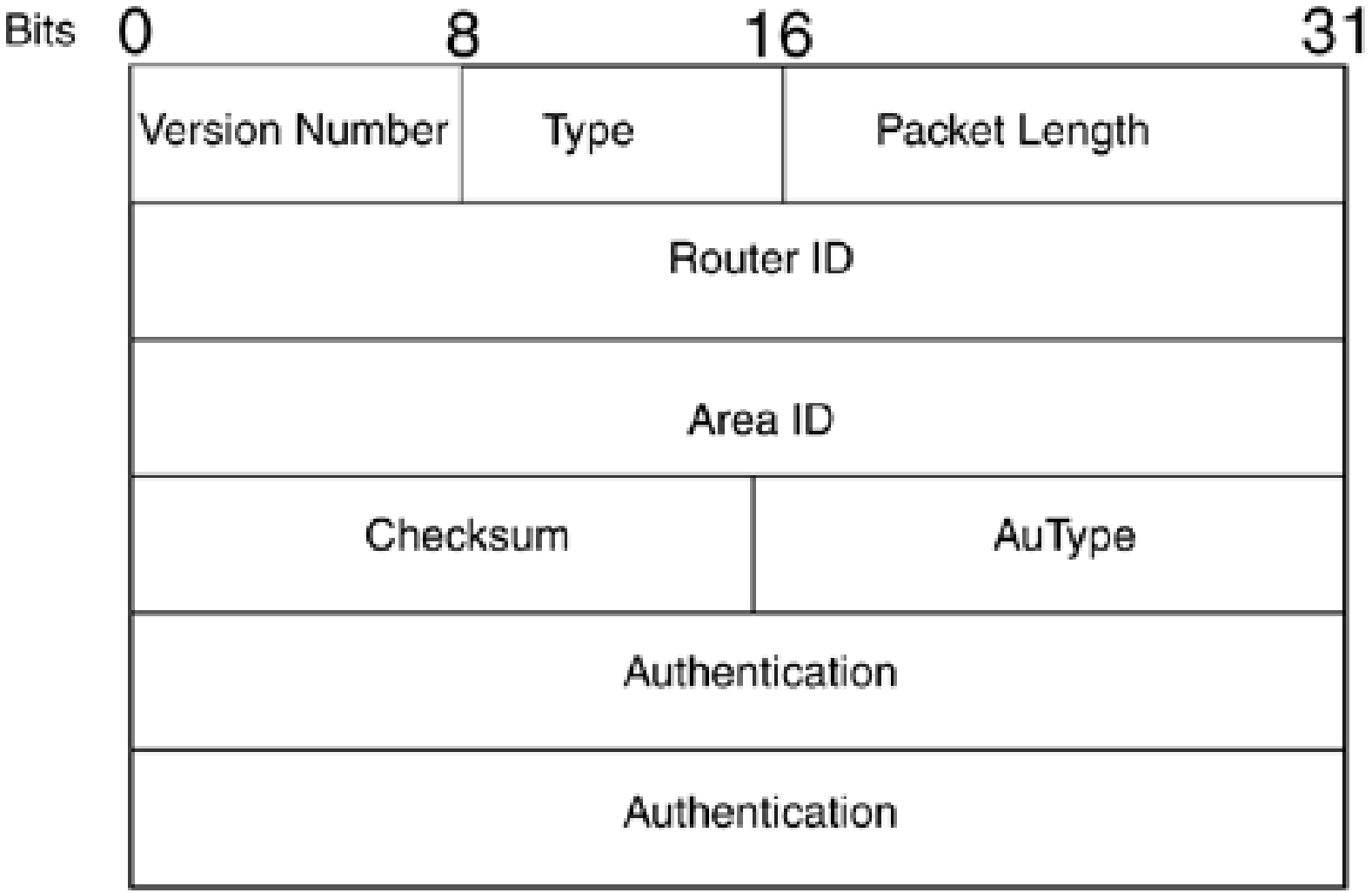
- **Type of service:**
  It provide different services to different routes.
  Provide services according to the nature of applications.

2. Load balancing:
Balance traffic load through multiple routes.

3. AS is divided into Areas:

4. Security:
Info. between routers are authenticated.

OSPF Packet header:



- **Version Number**— This field represents the current version number of OSPF. The latest version is 2. Version 1 is not compatible with Version 2.

- **Type—** This field indicates which of the five types of OSPF packets is appended at the end of this header.

- **Packet Length**— This field contains the length of the entire OSPF packet, including the OSPF header.

- **Source Router ID**— This field contains the 4-byte IP address. The source router ID is used to uniquely identify the router throughout the autonomous system.

- **Area ID**— This field designates the area number to which this packet belongs. This is also a 4-byte number. There are two ways to write this: Area 1 or Area 0.0.0.1. There is no difference between the two.

- **Checksum**— This field includes the checksum for the entire OSPF packet, excluding the authentication for data corruption.

- **AuType**— This field contains the type code for the authentication:

- *0* means that there is a null authentication.
  - *1* means that the authentication type is plain text.

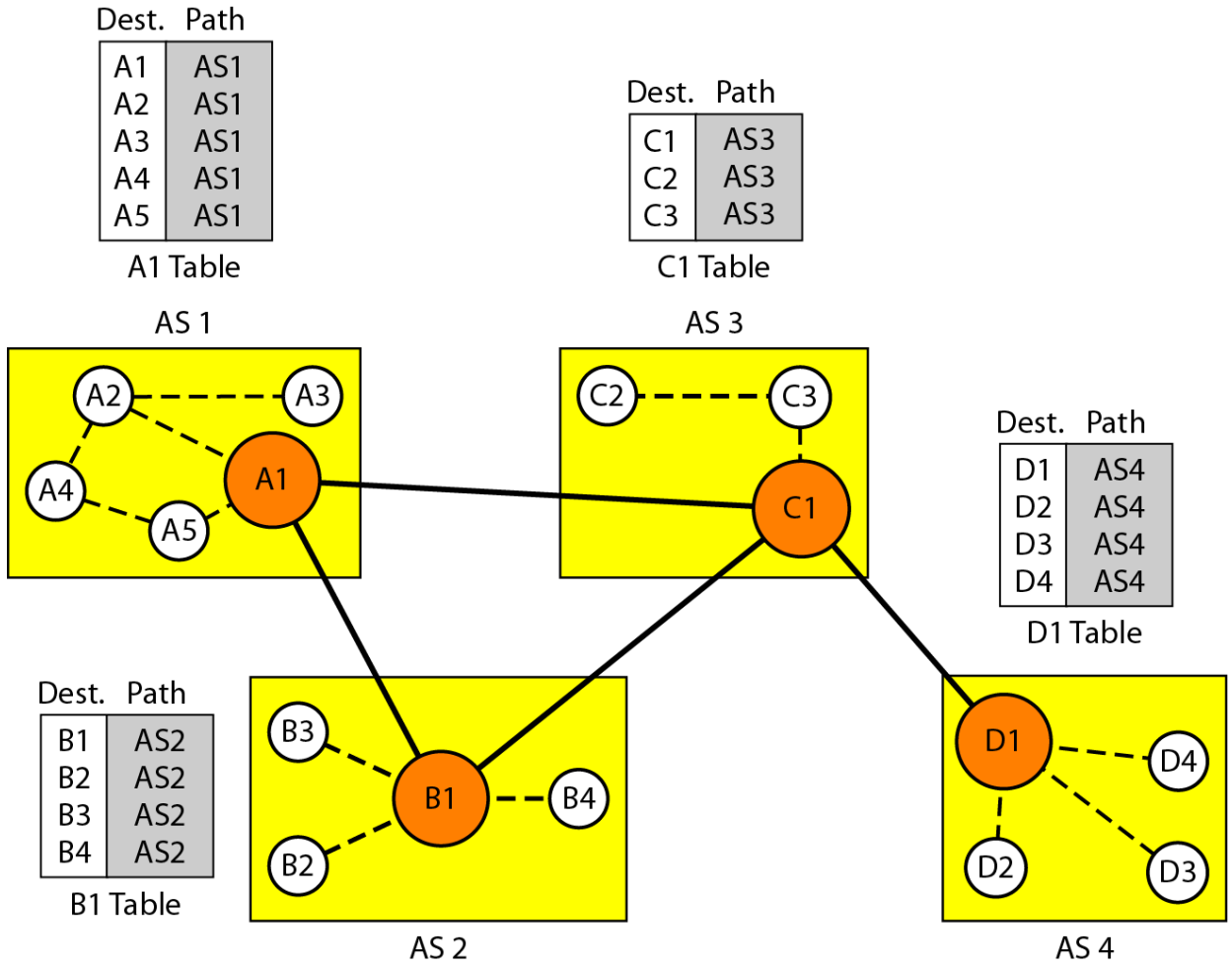
- **Authentication**— This 64-bit field contains the authentication key in case of plain-text authentication.

Path Vector Routing

- *It is Interdomain routing.*
- *Based on best spanning Tree.*
- *Similar to Distance vector routing.*
- *Speaker node in AS creates a routing table & advertise it to speaker nodes in the neighboring AS's.*

  - *Initialization*
  - *Sharing*
  - *Updating*

Figure. *Initial routing tables in path vector routing*



Stabilized tables for three autonomous systems

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

A1 Table

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

B1 Table

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

C1 Table

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

D1 Table

Border Gateway Protocol (BGP)

- **Interdomain routing protocol.**
- **Exchange routing information between 2 different AS.**
- **Based on "Path Vector routing"**
- **Exchange n/w reachability information among BGP routers.**

BGP perform 3 functional procedures:

1. Neighbor acquisition :

- **-** It is used to exchange routing info. Between routers of 2 diff. AS.
- **-** It uses Open msg & keep alive msg.

2. Neighbor reachability :

- Maintain a procedure by Keep alive msg.

3. Network reachability :

- Update routing info. Using Update msg.

Types of BGP Messages

OPEN

To negotiate and establish peering

UPDATE

To exchange routing information

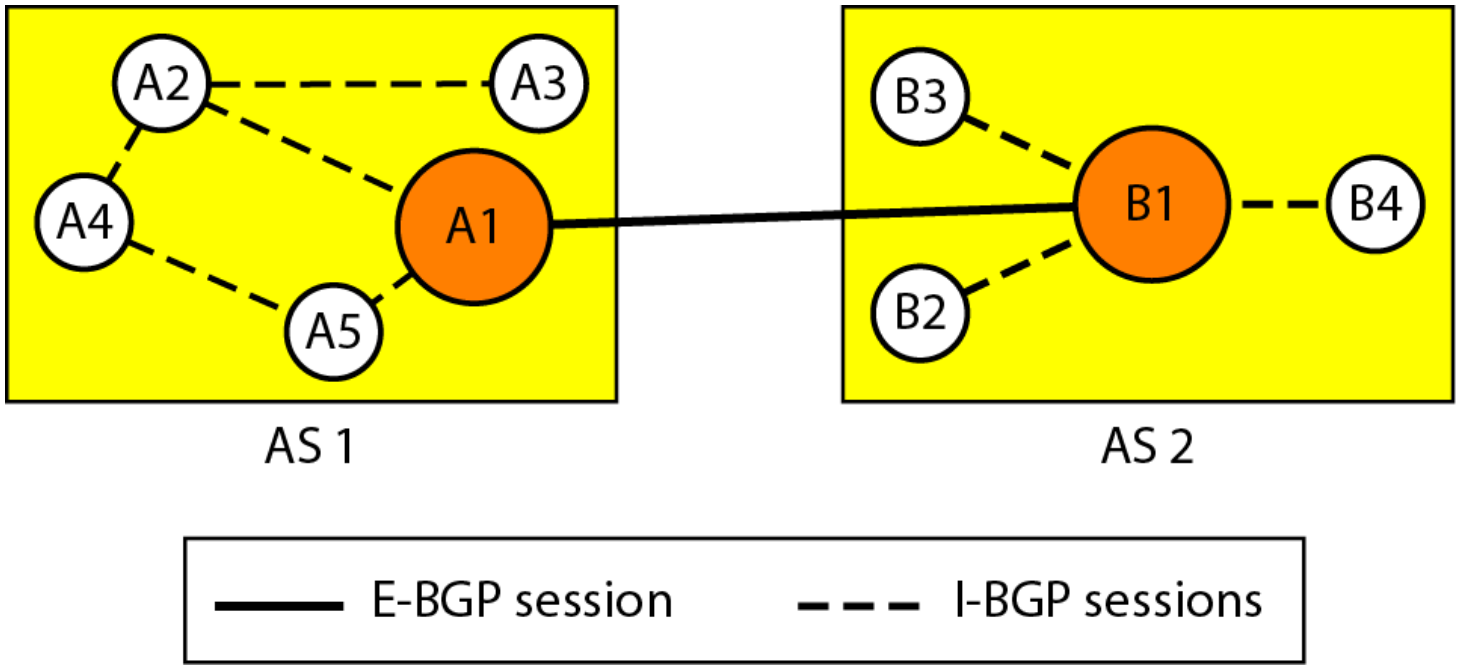
KEEPALIVE

To maintain peering session

NOTIFICATION

To report errors.

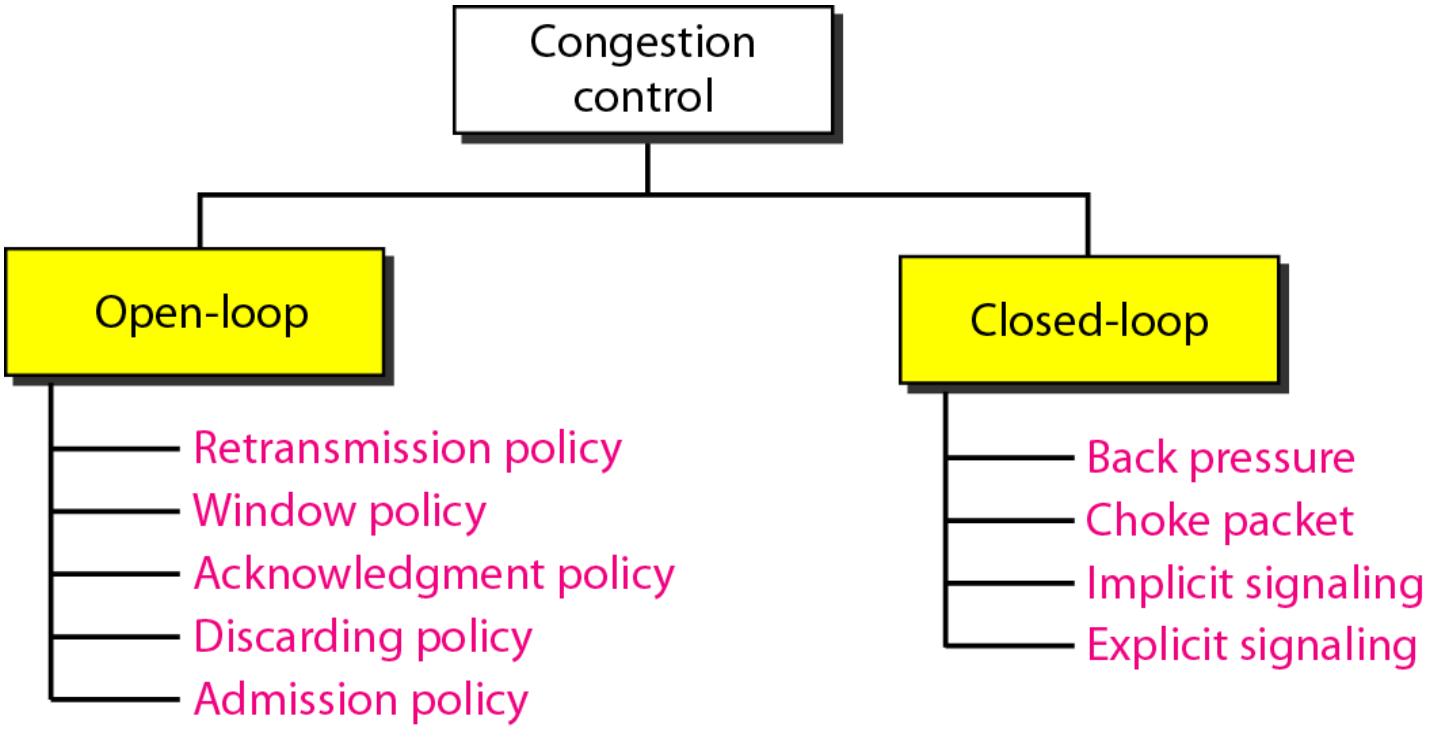
- ***Internal and external BGP sessions***



CONGESTION CONTROL

- Too many packets present in (a part of) the network causes packet delay and loss that degrades performance. This situation is called congestion.
- We would like to design networks that avoid congestion where possible and do not suffer from congestion collapse if they do become congested.
- If routers have an infinite amount of memory, congestion gets worse, not Better.
- Congestion control has to do with making sure the network is able to carry the offered traffic.
- Flow control, in contrast, relates to the traffic between a particular sender and a particular receiver.

Congestion control categories



Open – Loop Congestion Control:

1.Retransmission Policy:

Good Retransmission policy can prevent congestion.

2. Window Policy:

a. Selective Repeat window.

b. Go-Back-N window.

2.Acknowledgment Policy:

Minimum acknowledgment means minimum load on network.

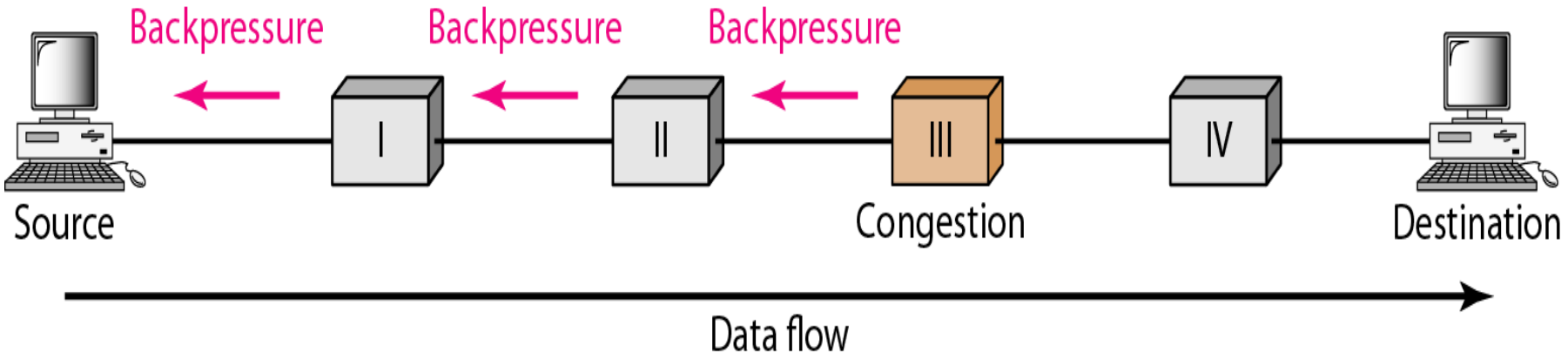
3.Discarding Policy:

4.Admission Policy:

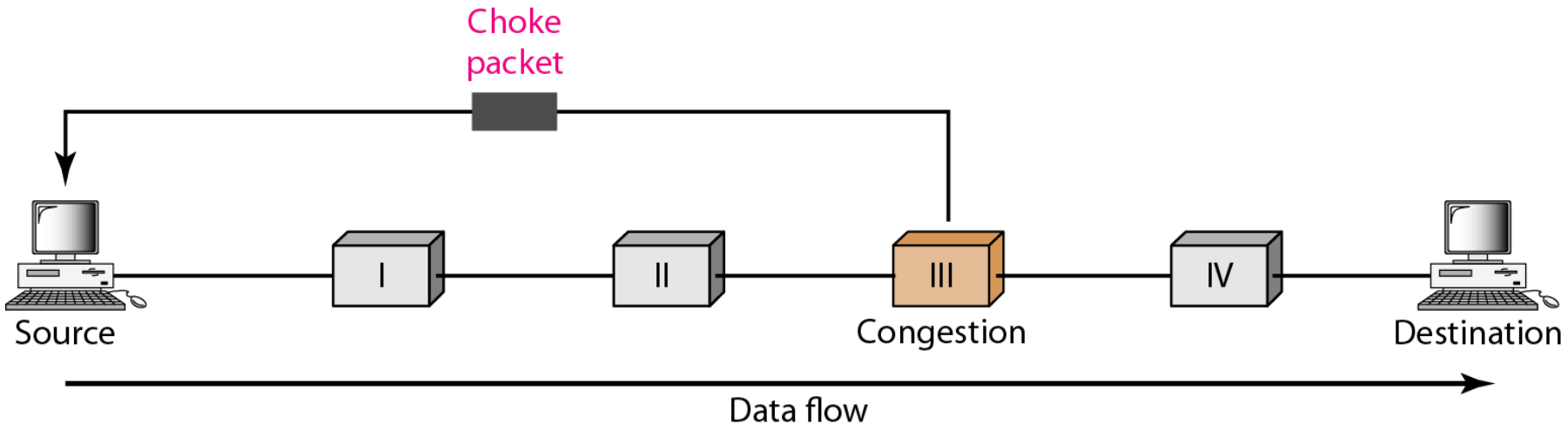
It is a QoS mechanism.

Closed – Loop Congestion Control

1. Backpressure method for alleviating congestion



2. Choke packet



3. Implicit Signaling

4. Explicit Signaling.

- *Backward Signaling.*
- *Forward Signaling.*

Multi-Protocol Label Switch (MPLS)

- MPLS, is a networking technology that routes traffic using the shortest path based on “labels,” rather than network addresses, to handle forwarding over private wide area networks.
- MPLS can reduce latency (the delay in sending/receiving data).
- It also reduces congestion on the paths that have just been avoided as a result of traffic.
- In an MPLS network, incoming packets are assigned a "label" by a "label edge router (LER)".

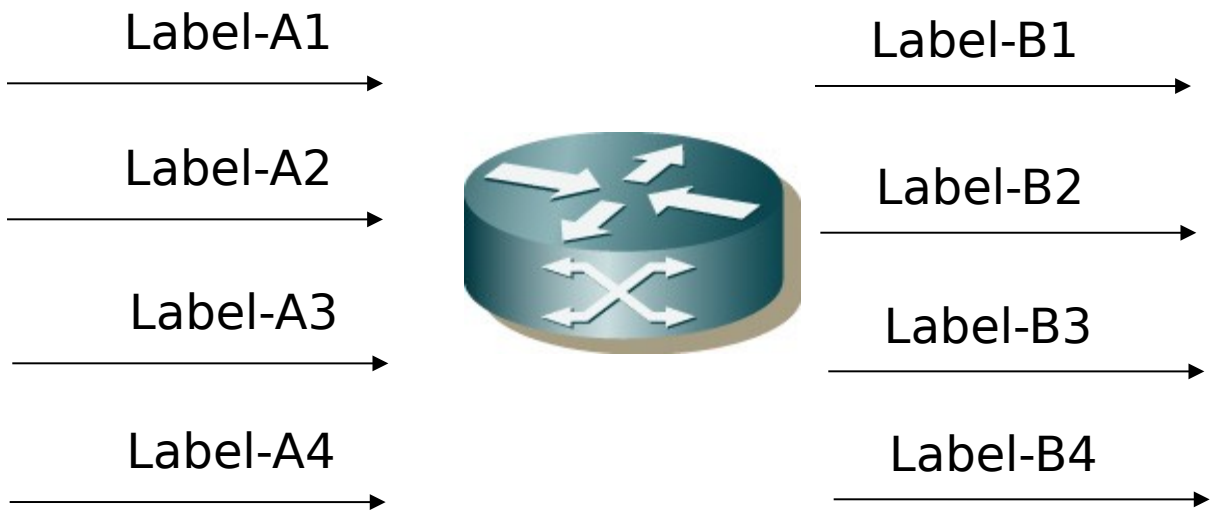
Multi-Protocol Label Switch (MPLS)

- Packets are forwarded along a "label switch path (LSP)" where each "label switch router (LSR)" makes forwarding decisions based solely on the contents of the label.
- At each hop, the LSR strips off the existing label and applies a new label which tells the next hop how to forward the packet.
- Label Switch Paths (LSPs) are established by network operators for a variety of purposes, such as to guarantee a certain level of performance, to route around network congestion, or to create IP tunnels for network-based virtual private networks.

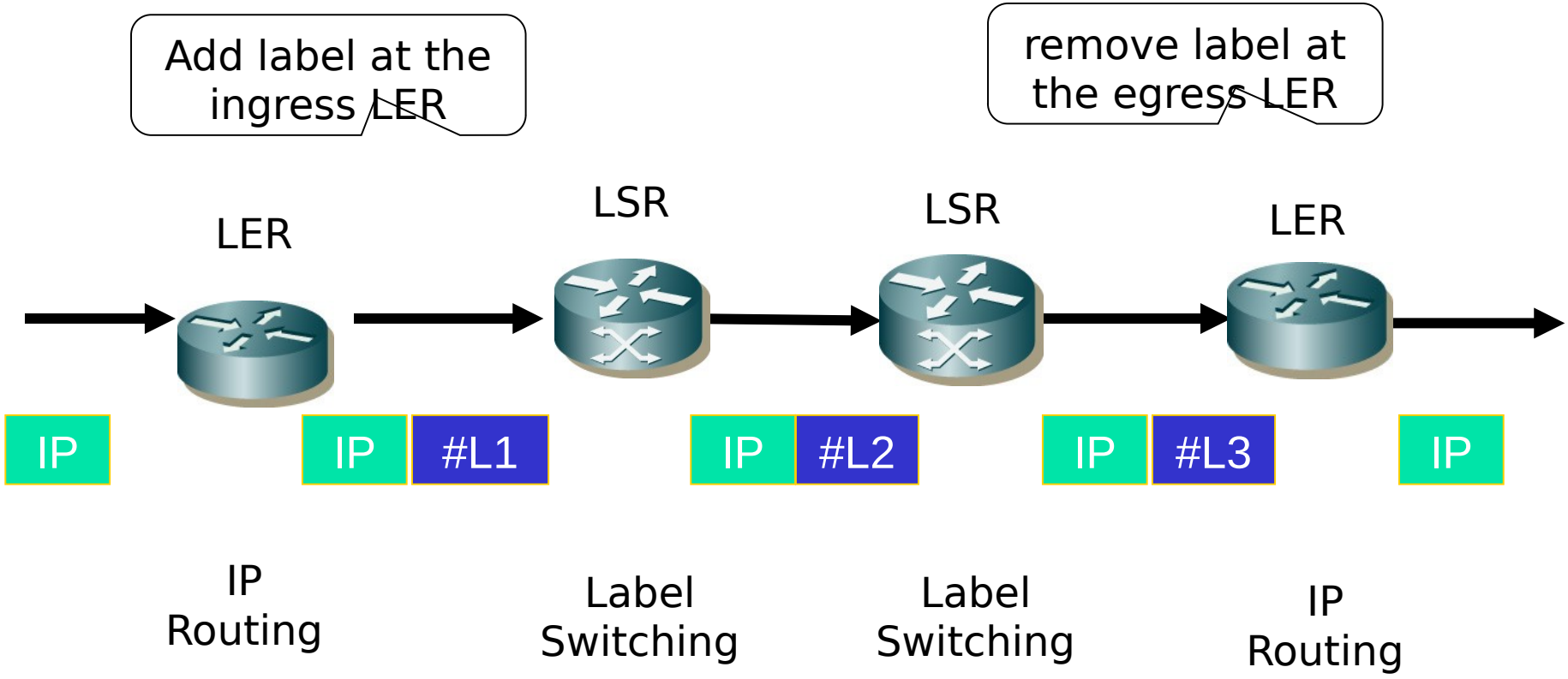
MPLS

- A protocol to establish an end-to-end path from source to the destination
- A *hop-by-hop* forwarding mechanism
- Use labels to set up the path
  - Require a *protocol* to set up the labels along the path
- It builds a *connection-oriented* service on the IP network

Label Substitution (swapping)



How does it work?



Mobile IP

- **Developed as a means for transparently dealing with problems of mobile users**
- Enables hosts to stay connected to the Internet regardless of their location and without changing IP addresses
- Requires no changes to software of non-mobile hosts/routers
- **Requires addition of some infrastructure**
- Has no geographical limitations
- Requires no modifications to IP addresses
- Supports security

Mobile IP Entities

- Mobile Node (MN)
  - The entity that moves from network to network
  - Assigned a permanent IP called its *home address* to which other hosts send packets regardless of MN's location
- Home Agent (HA)
  - Router with additional functionality
  - Located on home network of MN
  - Mobility binding of MN's IP with its *Care of Address* (COA)
  - Forwards packets to appropriate network when MN is away – uses encapsulation

Mobile IP Entities contd.

- Foreign Agent (FA)
  - Another router with enhanced functionality
  - Used to send/receive data between MN and HA
  - Advertises itself periodically
- Care-of-address (COA)
  - Address which identifies MN's current location
  - Sent by FA to HA when MN attaches
  - Usually the IP address of the FA
- Correspondent Node (CN)
  - End host to which MN is corresponding (eg. a web server)

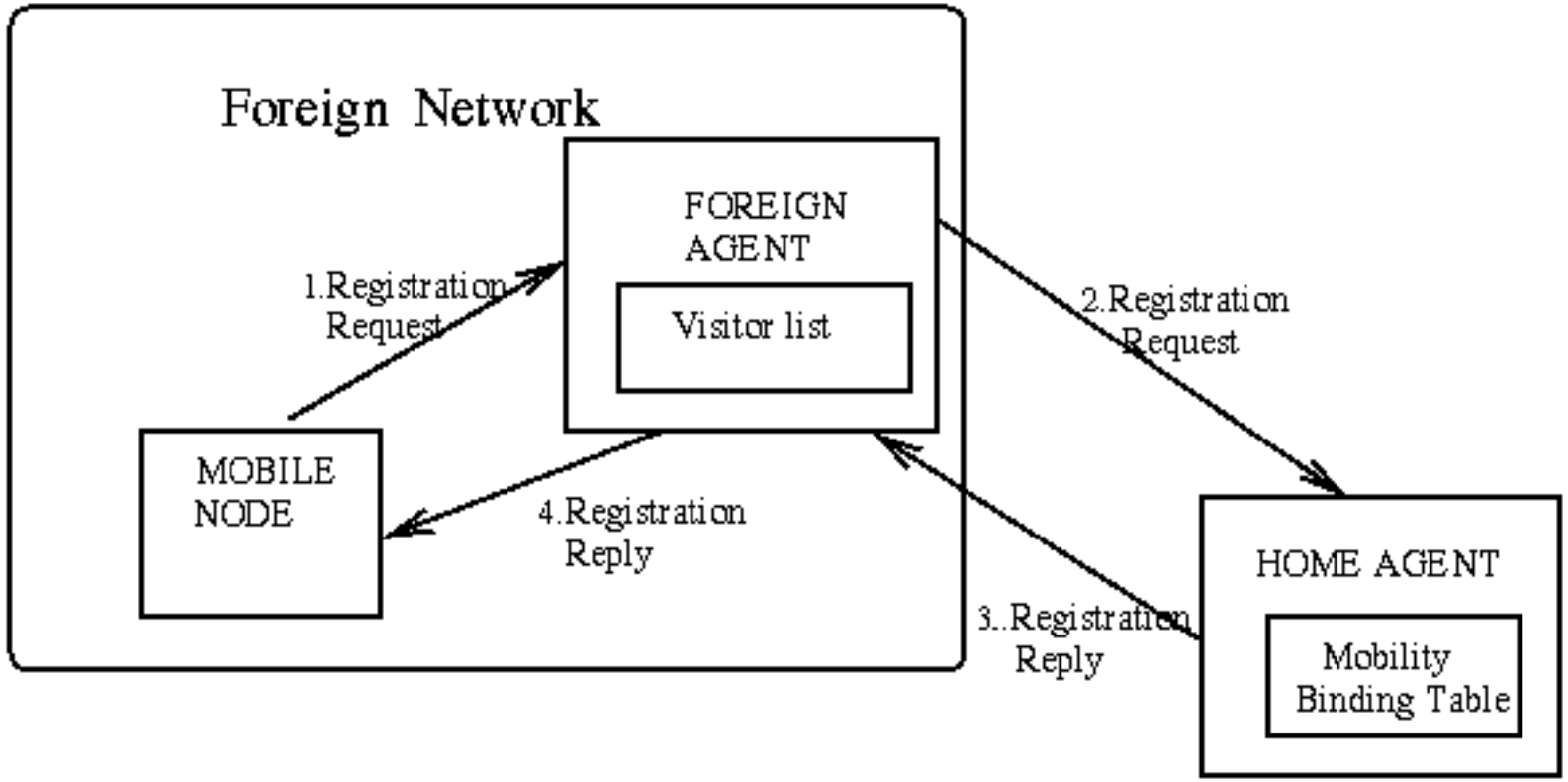
Mobile IP Support Services

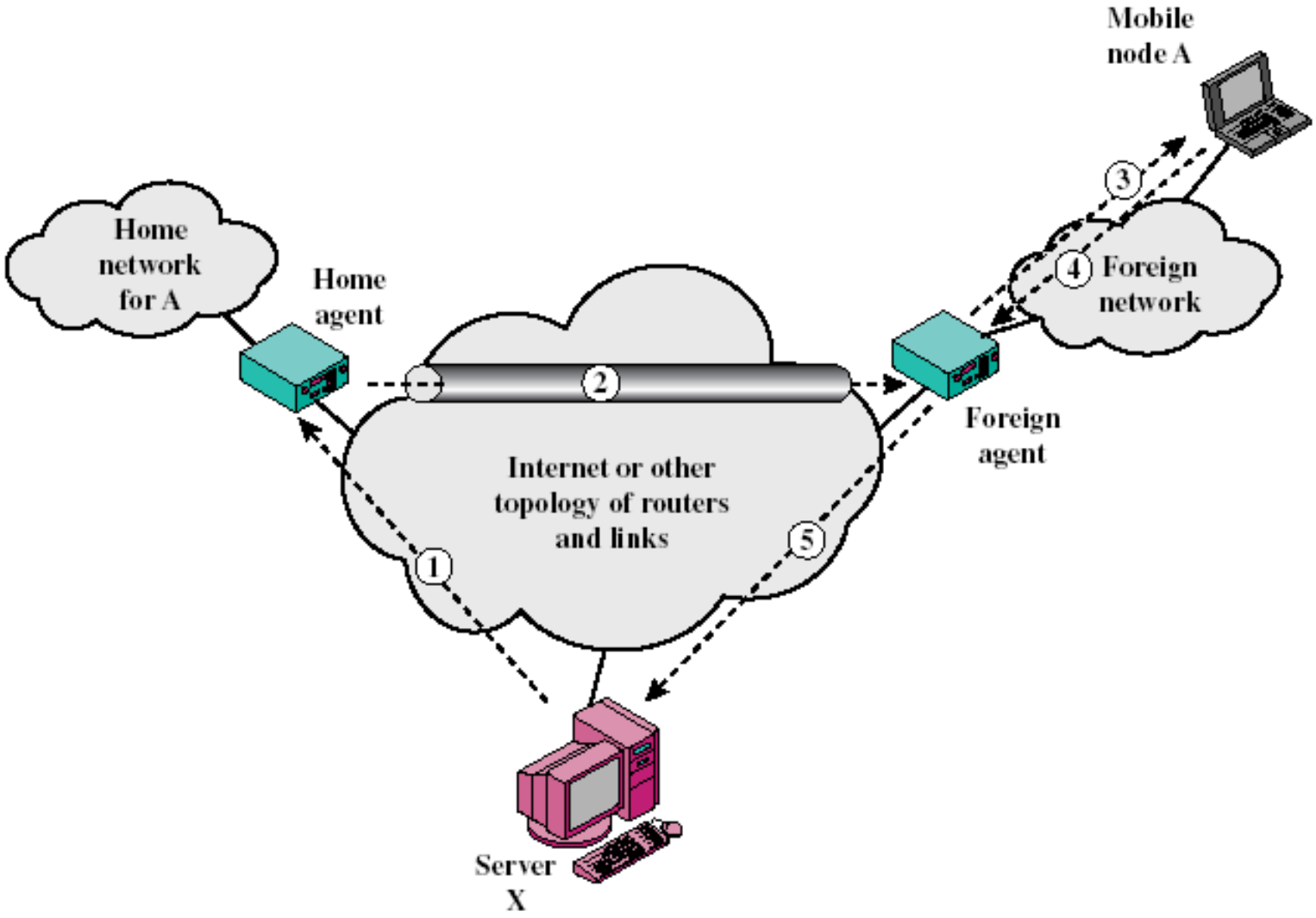
- Agent Discovery
  - HA's and FA's broadcast their presence on each network to which they are attached
  - MN's listen for advertisement and then initiate registration
- Registration
  - When MN is away, it registers its COA with its HA, via FA
  - Registration control messages sent via UDP to well known port
- Encapsulation/decapsulation – just like standard IP only with COA

Mobile IP Operation

- A MN listens for agent advertisement and then initiates registration
  - If responding agent is the HA, then mobile IP is not necessary
- After receiving the registration request from a MN, the HA acknowledges and registration is complete
  - Registration happens as often as MN changes networks
- HA intercepts all packets destined for MN
  - This is simple unless sending application is on or near the same network as the MN
  - HA masquerades as MN
  - There is a specific lifetime for service before a MN must re-register
  - There is also a de-registration process with HA if an MN returns home

Registration Process




Mobile
node A
Home
network
for A
Home
agent
3
4
Foreign
network
2
Foreign
agent
Internet or other
topology of routers
and links
1
5
Server
X