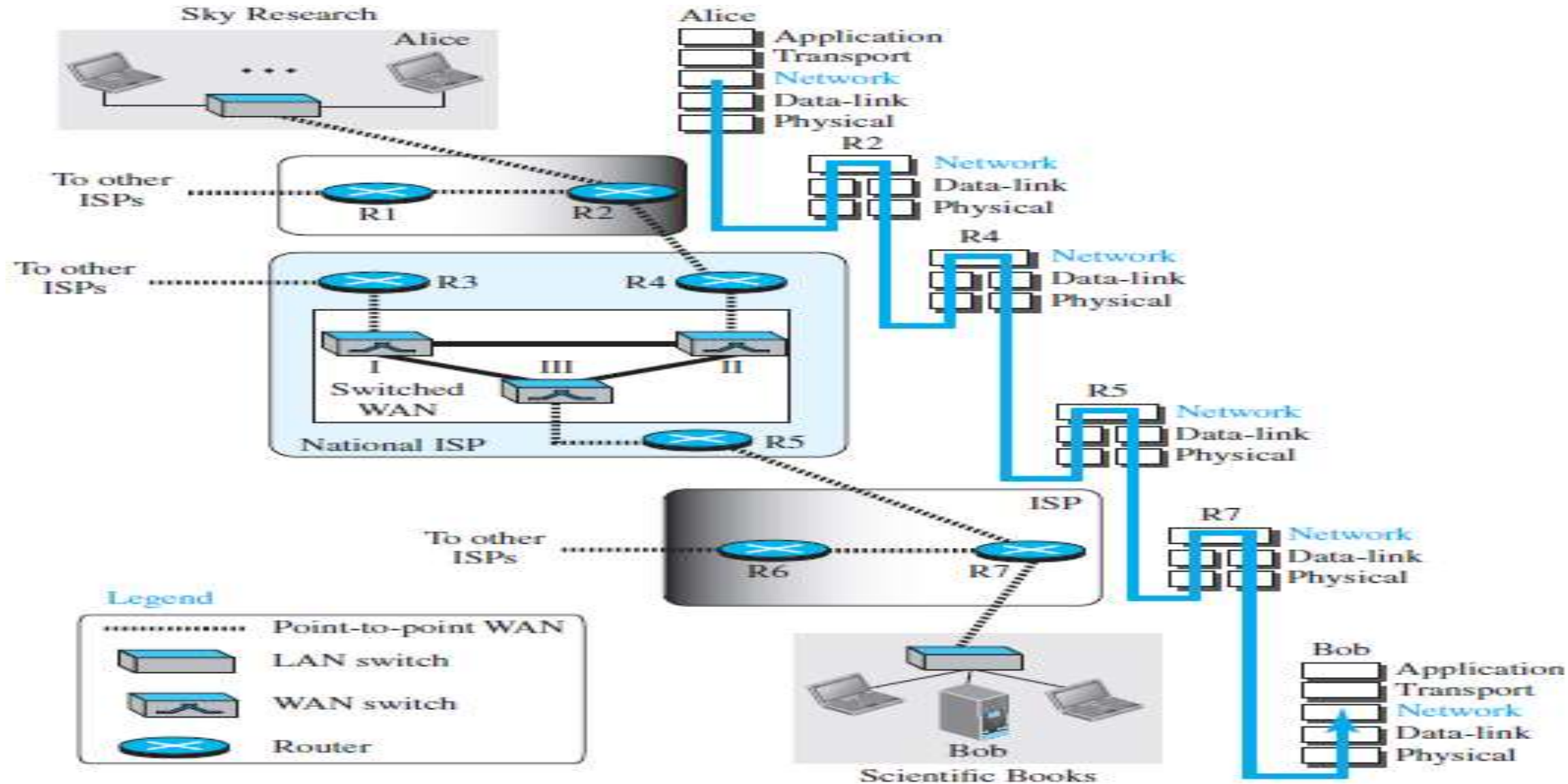
Chapter Three

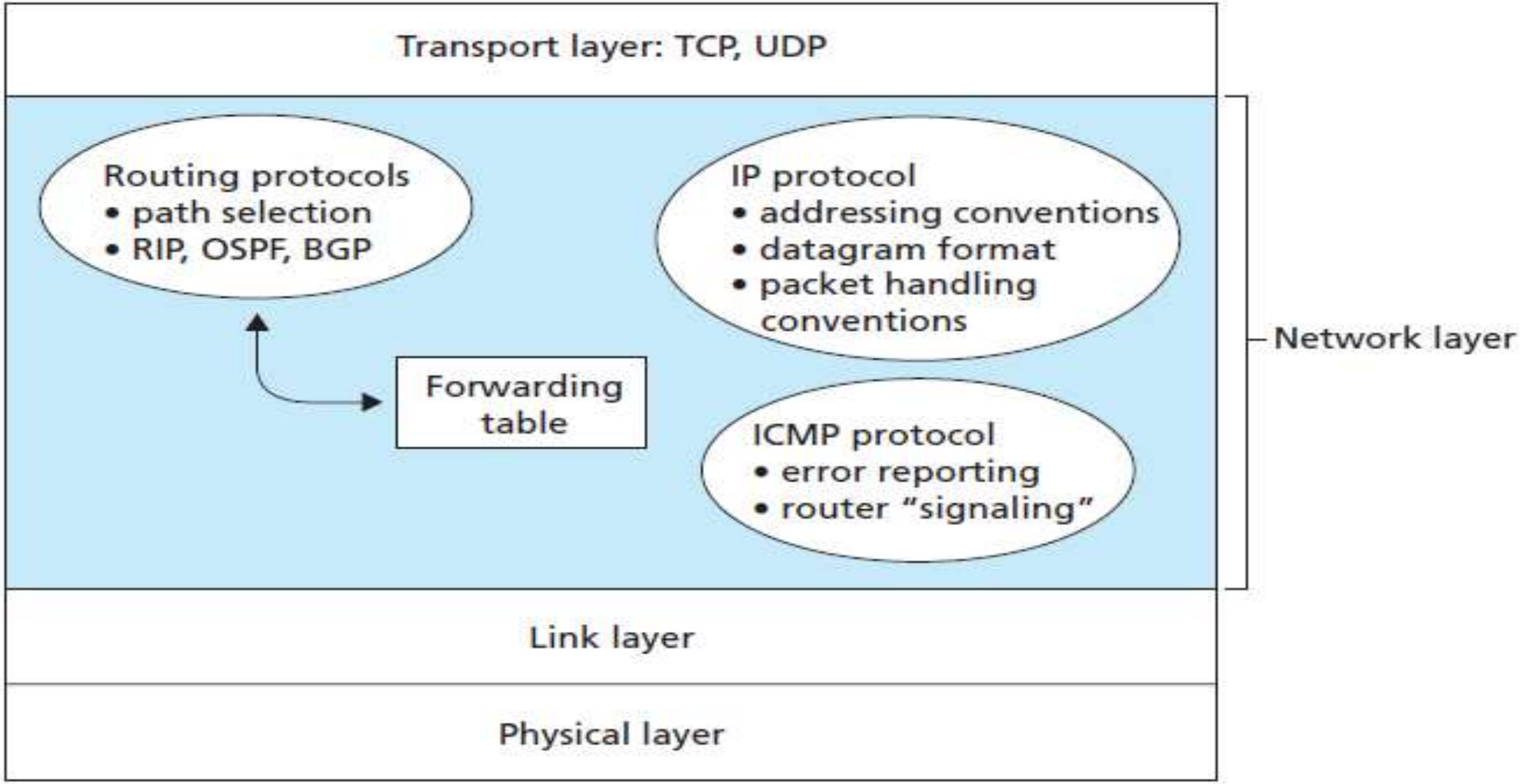
Network Layer And Routing

Introduction

Data Communication and Computer Networks

(CSE3204)


Sky Research
Alice
Alice
Application
Transport
Network
Data-link
Physical
R2
Network
Data-link
Physical
To other ISPs
R1
R2
R4
Network
Data-link
Physical
To other ISPs
R3
R4
I
III
II
Switched WAN
National ISP
R5
R5
Network
Data-link
Physical
ISP
R7
Network
Data-link
Physical
To other ISPs
R6
R7
Legend
Point-to-point WAN
LAN switch
WAN switch
Router
Bob
Application
Transport
Network
Data-link
Physical
Bob
Scientific Books



A look inside the Internet's network layer

Introduction

- It is the lowest layer that deals with delivering of individual packet from source host to destination host.

1. The first duty of the network layer is **packetizing**: encapsulating the payload in a network-layer packet at the source and decapsulating the payload from the network-layer packet at the destination.

2. Forwarding:- involves the transfer of a packet from an incoming link to an outgoing link within a single router.

3. Routing:- involves all of a network's routers, whose collective interactions via routing protocols (algorithms) determine the paths that packets take on their trips from source to destination node.

4. **Logical addressing:** required if a packet passes the network boundary, to distinguish the source and destination systems.

5. **Best path selection:-** Best path is selected by a routing protocol based on the value or metric it uses to determine the distance to reach a network.

Introduction cont….

6. Connection setup:

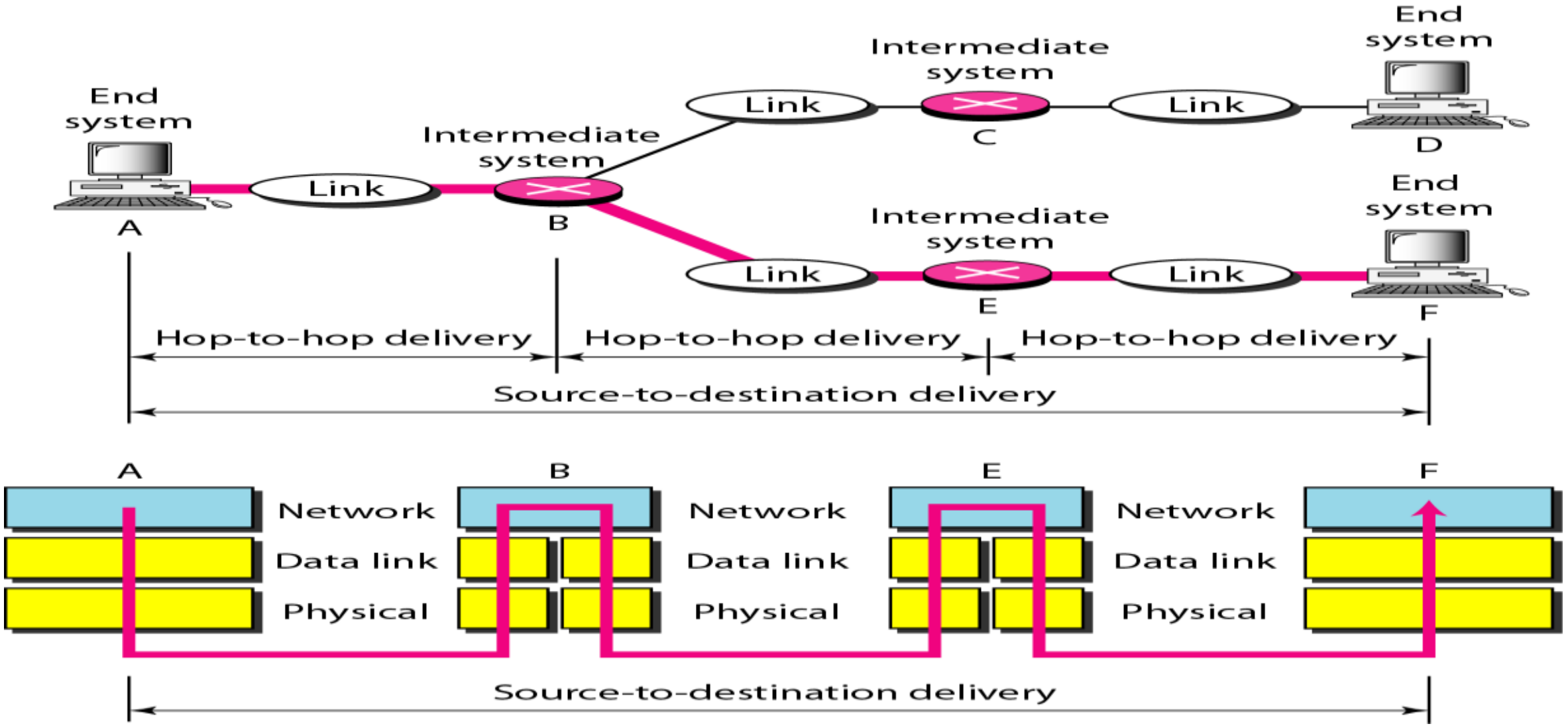
- Connection-oriented communication (virtual circuit network) and
- Connectionless communication (datagram network)

7. Store-and-Forward Packet Switching: A host with a packet to send transmits it to the nearest router. The packet is stored there until it has fully arrived and the link has finished its processing by verifying the checksum.

- Then forwarded to the next router along the path until it reaches the destination host, where it is delivered. This mechanism is store-and-forward packet switching.

8. Security: network layer provide a security for a connectionless service, for this purpose we need to have another virtual level (IPsec) that changes the connectionless service to a connection-oriented service.

Cont'd...


End system
A
Link
Intermediate system
B
Link
Intermediate system
C
Link
End system
D
Link
Intermediate system
E
Link
End system
F
Hop-to-hop delivery
Hop-to-hop delivery
Hop-to-hop delivery
Source-to-destination delivery
A
B
E
F
Network
Data link
Physical
Network
Data link
Physical
Network
Data link
Physical
Source-to-destination delivery

cont'd….

Services Provided to the Transport Layer:

- The services provided to the transport layer by the network layer can be carefully designed by considering the following transport layer requirements in mind:

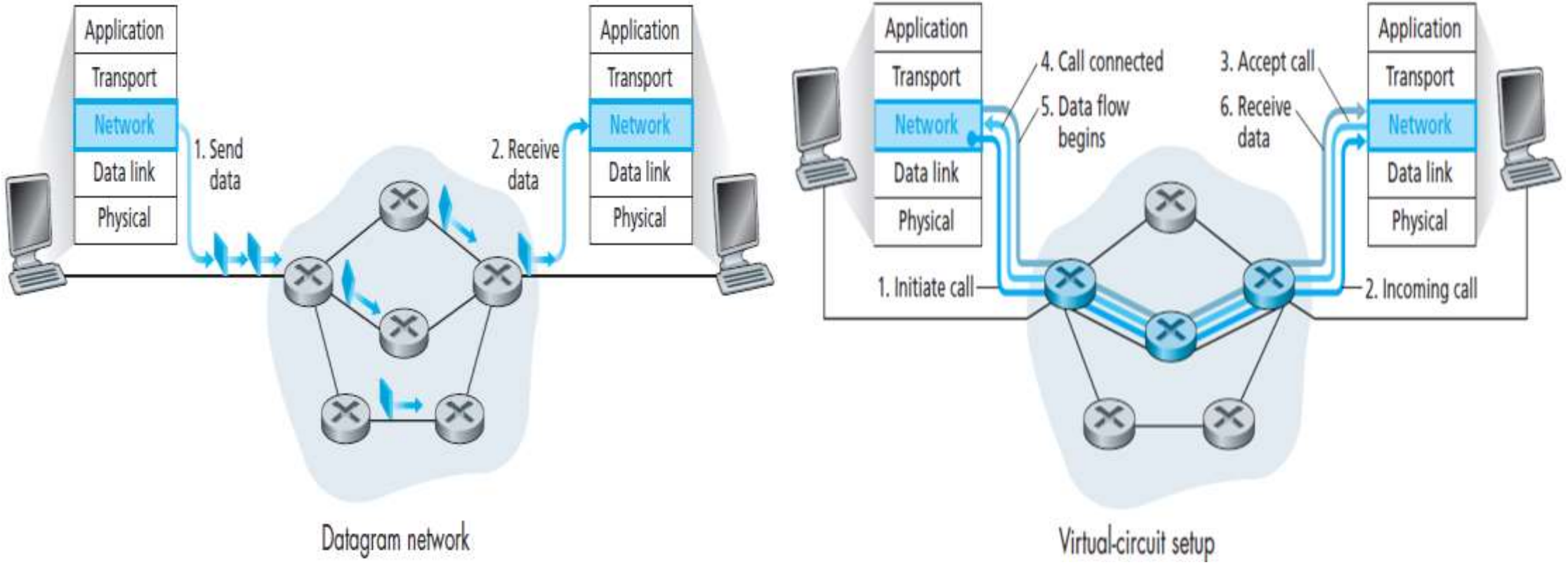
  1.The services should be independent of the router technology.

  2. The transport layer should be shielded from the number, type, and topology of the routers present.

  3. The network addresses made available to the transport layer should use a uniform numbering plan, even across LANs and WANs.

- The network layer should provide connection oriented or connectionless service to transport layer.

Connectionless and Connection-oriented Communication



Comparison of Datagram and Virtual-Circuit Networks

Issue	**Datagram network**	**Virtual-circuit network**
Circuit setup	Not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short VC number
State information	Routers do not hold state information about connections	Each VC requires router table space per connection
Routing	Each packet is routed independently	Route chosen when VC is set up; all packets follow it.
Effect of router failures	None, except for packets lost during the crash	All VCs that passed through the failed router are terminated
Quality of service	Difficult	Easy if enough resources can be allocated in advance for each VC
Congestion control	Difficult	Easy if enough resources can be allocated in advance for each VC

Network-Layer Performance Measures

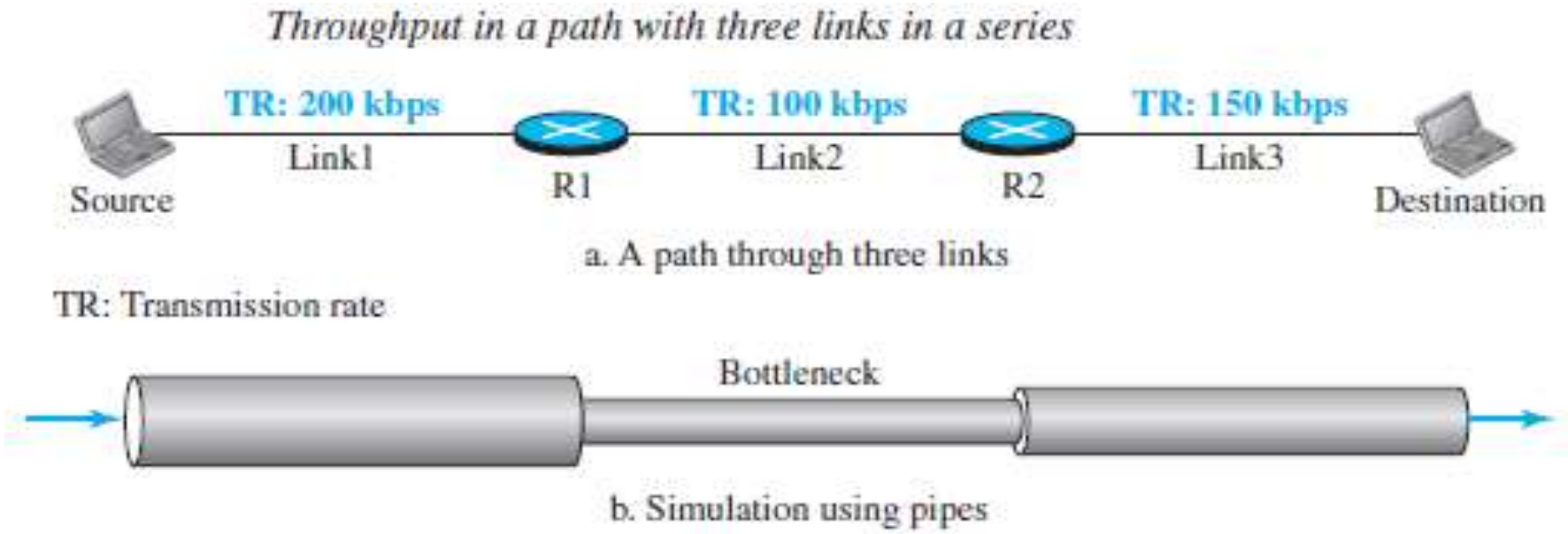
- The performance of a network can be measured in terms of ***delay***, ***throughput***, and ***packet loss (***sometimes ***congestion control*** can also be consider to improve the performance).

❖ The delays in a network can be divided into four types:

- ✓ **Transmission delay**,=(Packet length) / (Transmission rate)
- ✓ **Propagation delay**, =(Distance) / (Propagation speed)
- ✓ **Processing delay**, =(Time required to process a packet in a router or a destination host
- ✓ **Queuing delay**.=The time a packet waits in input and output queues in a router
- ✓ *Total delay = (n + 1) (Delay tr + Delay pg + Delay pr) + (n) (Delay qu)*, by assuming the total number of router in the middle equals to *n*.

Cont'd……

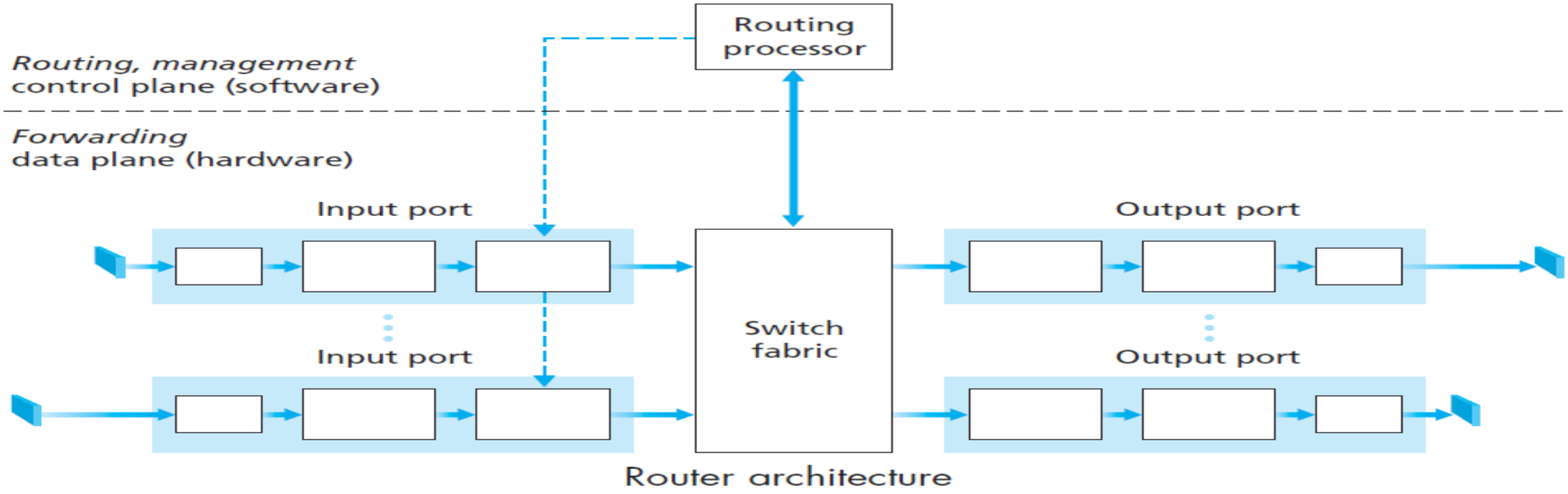
- **Throughput** is the number of bits passing through the point in a second, which is actually the transmission rate of data at that point.
- Another issue that severely affects the performance of communication is the number of packets lost during transmission.
- Throughput = minimum {TR1, TR2, . . . TRn}.



Introduction cont….
Router

Router cont….

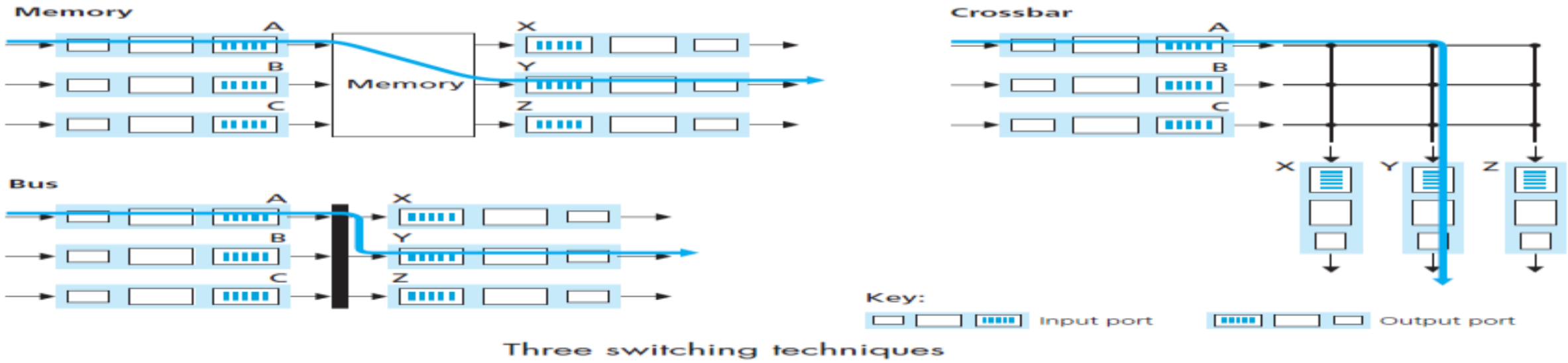
- Two main functions:
  - Run routing algorithms/protocols (e.g., RIP, OSPF, IGRP and others).
  - Forwarding datagrams from incoming to outgoing links



Router architecture

Switching fabric

- It is the heart of the router, it can be used to forward data from incoming port to outgoing port.
- The switching fabric is a crucial component of a router or network switch that facilitates the movement of data packets between input and output ports within the device. It serves as the central switching mechanism that connects the various ports and enables efficient packet forwarding.
- Switching fabric can be implemented in the following ways:
- Switching via *memory*, Switching via *bus and* Switching via *interconnection network (Crossbar).*



Three switching techniques

Output ports

- An output port stores packets received from the switching fabric and transmits these packets on the outgoing link by performing the necessary link-layer and physical-layer functions.
- When a link is bidirectional, an output port will typically be paired with the input port for that link on the same line card (a printed circuit board containing one or more input ports, which is connected to the switching fabric).
- Scheduling discipline chooses among queued datagrams for transmission by considering the issues like:
  - Ordering (First Come First Served), Quality of Service (QoS), fairness (protecting monopoly of resources).

Routing Processor

- The routing processor:
  - Executes the routing protocols (which we'll study in the next classes and demonstrated in the laboratory sessions),
  - Maintains routing tables (by sharing routing table information among routers) and attached link state information (dependent on the routing algorithm used), and
  - Computes the forwarding table for the router
- It also performs the network management functions.

Internet Protocol (IP) and IP addressing

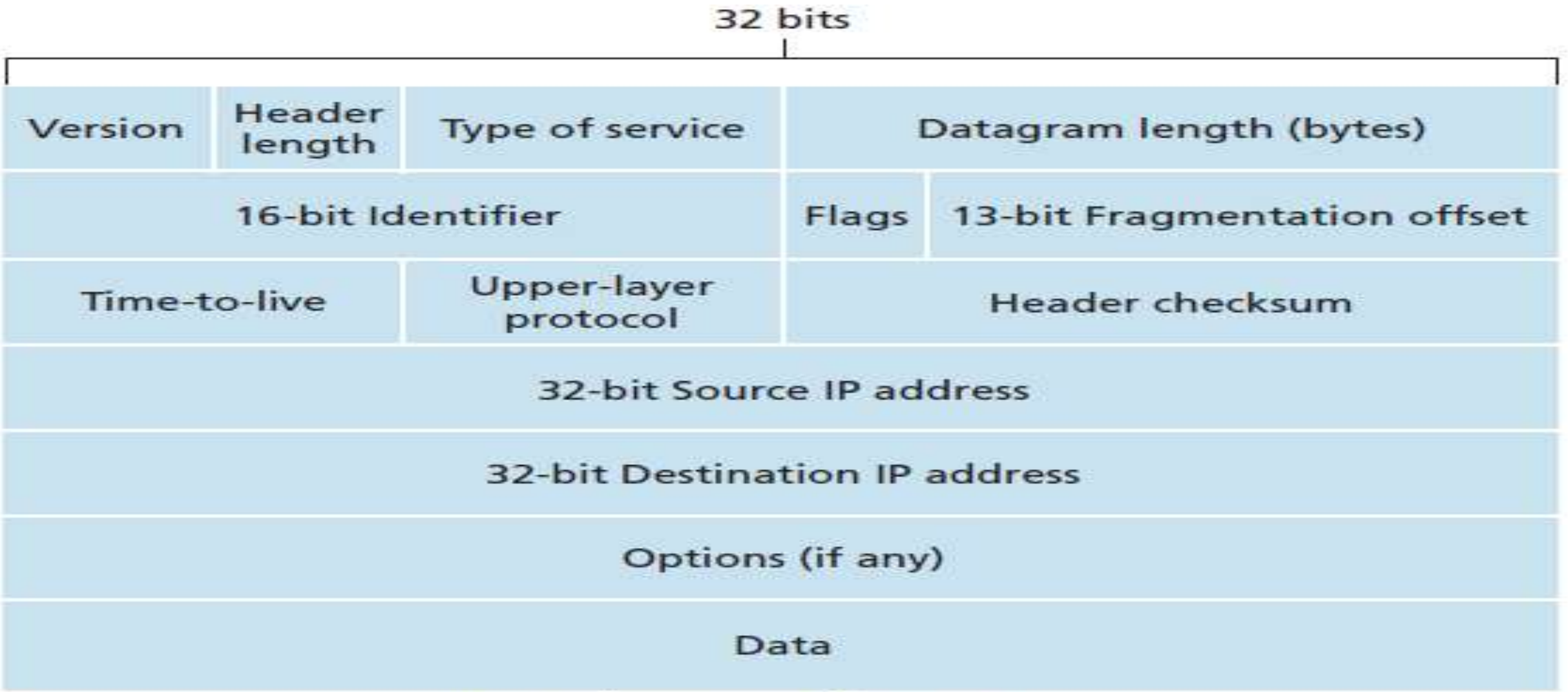
- As with any protocol standard, IP is specified in two parts:
  - The interface with a higher layer (TCP/UDP), specifying the services that IP provides
  - The actual protocol format and mechanisms
- IP basic characteristics:
  - Connectionless - No connection is established before sending data packets.
  - Best Effort (unreliable) - No overhead is used to guarantee packet delivery.
  - Media Independent - Operates independently of the medium carrying the data.

- There are two IP addressing method used in networking:-

  - Internet Protocol Version 4 (IPv4) is the first 32 bits addressing method and

  - Internet Protocol Version 6 (IPv6) 128 bits addressing.

IP Header

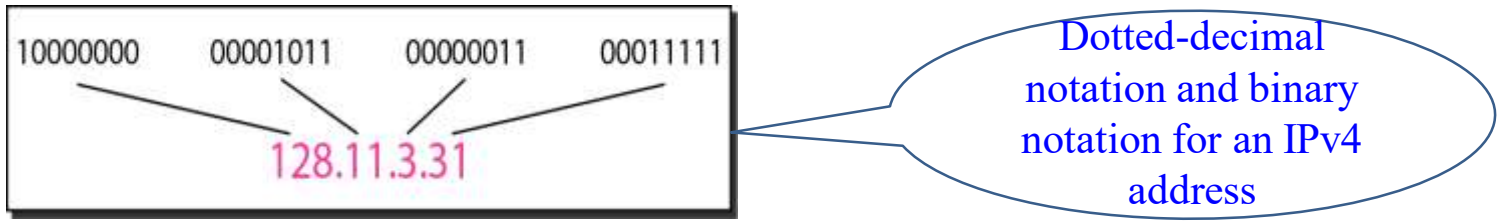


IPv4 datagram format

IPv4 Addressing

- IP address: it is a 32-bit identifier for host and router interfaces (both physical and logical interfaces have an ip address).
- Interface is a boundary between the host (such as a computer or a network device)and the physical link (such as an Ethernet cable or a wireless connection) in computer networking.
- An IP address is technically associated with an interface, rather than with the host or router containing that interface.
- These addresses are typically written in so-called **binary** and **dotted-decimal notation**, in which each byte of the address is written in its decimal form and is separated by a period (dot) from other bytes in the address.

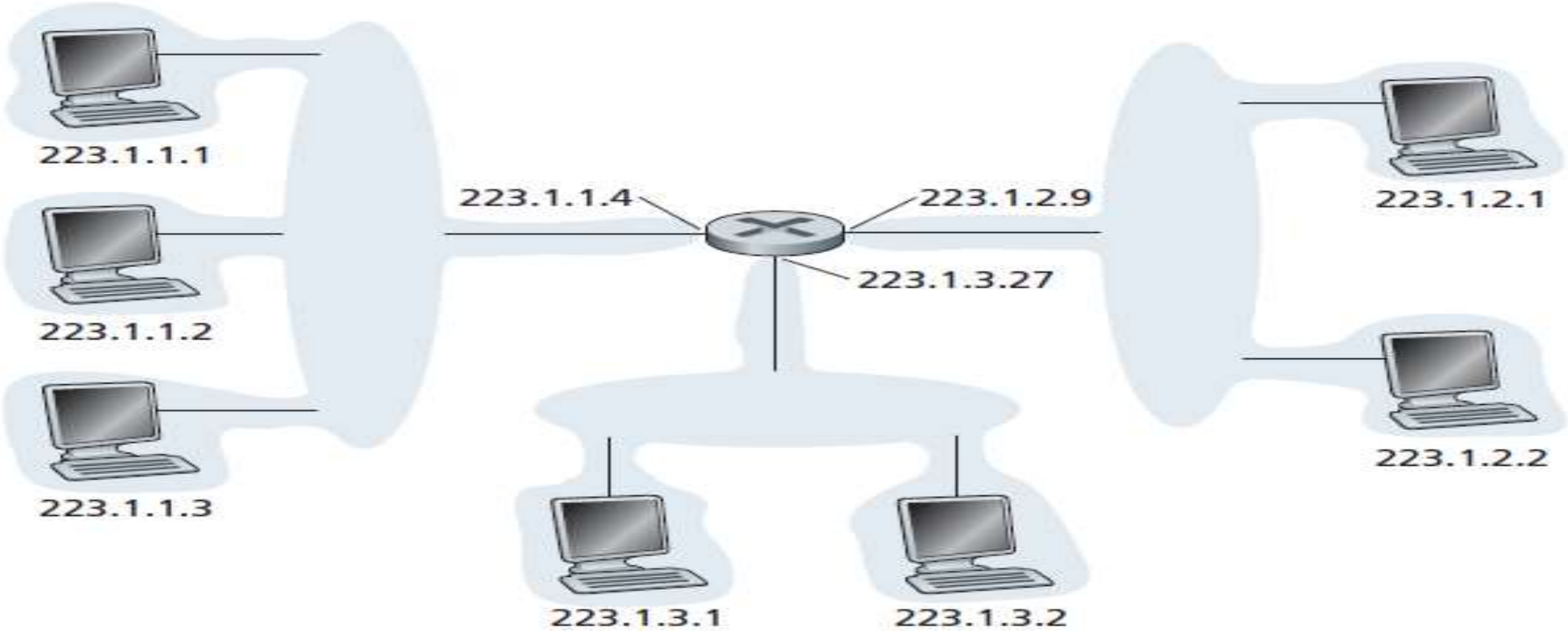
Example:



- These addresses cannot be chosen in a willy-nilly manner, however. A portion of an interface's IP address will be determined by the subnet to which it is connected.

IPv4 uses 32-bit addresses, which means that the address space is 2^{32} or 4,294,967,296 (more than 4 billion). This means that, theoretically, if there were no restrictions, more than 4 billion devices could be connected to the Internet.

IPv4 Addressing cont…



Interface addresses and subnets

Cont'd…

Rules for IPV4 representation

- A mixture of binary notation and dotted-decimal notation is not allowed.
- There can be no more than four numbers.
- Each number needs to be less than or equal to 255.
- There must be no leading zero.

Find the error, if any, in the following IPv4 addresses

a. 111.56.045.78
b. 221.34.7.8.20
c. 75.45.301.14
d. 11100010.23.14.67

Solution

a. There must be no leading zero (045).

b. There can be no more than four numbers.

c. Each number needs to be less than or equal to 255.

d. A mixture of binary notation and dotted-decimal notation is not allowed.

Example 1

1. Change the following IPv4 addresses from binary notation to dotted-decimal notation.

a. 10000001 00001011 00001011 11101111
b. 11000001 10000011 00011011 11111111

Solution

We replace each group of 8 bits with its equivalent decimal number and add dots for separation.

a. 129.11.11.239
b. 193.131.27.255

Example 2: Change the following IPv4 addresses from dotted-decimal notation to binary notation.

a. 111.56.45.78
b. 221.34.7.82

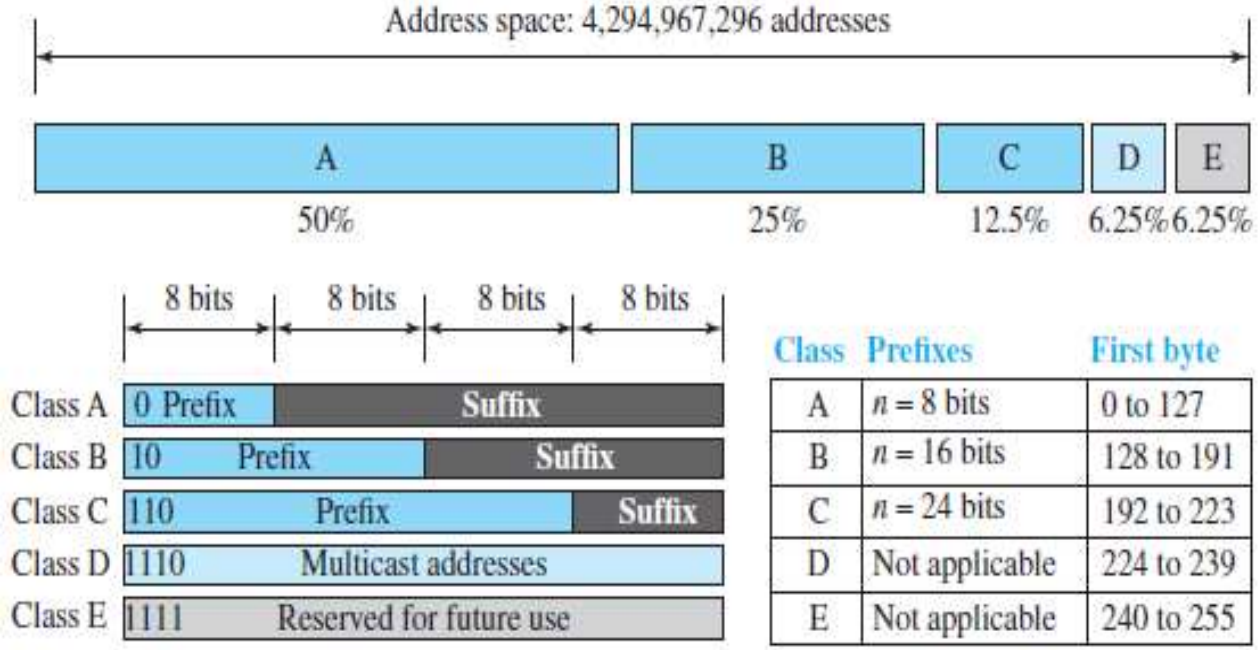
Solution

We replace each decimal number with its binary equivalent.

a. 01101111 00111000 00101101 01001110
b. 11011101 00100010 00000111 01010010

IPv4 Addressing cont'd….
Class full addressing

- Network Classes: the address is coded to allow a variable allocation of bits to specify network and host. This encoding provides flexibility in assigning addresses to hosts and allows a mix of network sizes on an internet.
- In class full addressing, the address space is divided into five classes: A, B, C, D, and E. Each class occupies some part of the address space.
- The three principal network classes are best suited to the following conditions:
  - Class A: Few networks, each with many hosts
  - Class B: Medium number of networks, each with a medium number of hosts
  - Class C: Many networks, each with a few hosts



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

- **Network ID and Host ID**

Example 4

Find class of the following IP addresses?

a. 00000001 00001011 00001011 11101111

b. 11000001 10000011 00011011 11111111

c. 14.23.120.8

d. 252.5.15.111

Solution

a. The first bit is 0. This is a class A address.

b. The first 2 bits are 1; the third bit is 0. This is a class C address.

c. The first byte is 14; the class is A.

d. The first byte is 252; the class is E.

Classes and Blocks

- One problem with classful addressing is that each class is divided into a fixed number of blocks with each block having a fixed size .

Class	*Number of Blocks*	*Block Size*	*Application*
A	128	16,777,216	Unicast
B	16,384	65,536	Unicast
C	2,097,152	256	Unicast
D	1	268,435,456	Multicast
E	1	268,435,456	Reserved

Number of Blocks for class A = 2^7
Number of Blocks for class B = 2^{14}
Number of Blocks for class C = 2^{21}

Block size for class A = 2^{24}
Block size for class B = 2^{16}
Block size for class C = 2^8

Block size=2^n where n is total number of host bit (suffix)

Number of block =2^{M-L} where M is total number of network bit and L is total number of common bit from network ID

IPv4 Addressing cont'd…..(CLA)

- The Internet's address assignment strategy is known as **Classless Inter domain Routing** (CIDR).
- CIDR generalizes the notion of subnet addressing. As with subnet addressing, the 32-bit IP address is divided into two parts and again has the dotted-decimal form *a.b.c.d/x (x is a network prefix)*.
- The *x* most significant bits of an address of the form *a.b.c.d/x* constitute the network portion of the IP address.
- The number of bits used in host portion determines the number of hosts that we can have within the network.
- **Types of Addresses in IPv4 Network**
  - **Network address** - The address by which we refer to the network
  - **Broadcast address** - A special address used to send data to all hosts in the network
  - **Host addresses** - The addresses assigned to the end devices (including router interfaces) in the network
- The size of the network is inversely proportional to the length of the suffix. A small prefix means a larger network; a large prefix means a smaller network.

Subnet and Network Prefixes (Subnet Masking)

- Subnetting is the process of splitting the larger network into a number of smaller networks.
- Sub netting is the process of borrowing bits from the HOST bits, in order to divide the larger network into small subnets.
- Within the sub netted network, the local routers must route on the basis of an extended network number consisting of the network portion of the IP address and the subnet number.
- Because of Subnetting we get the following advantages:
  - Network traffic was reduced
  - Network performance become optimized
  - Network management become simple
  - Network supports large geographic area
- The prefix length (x in the previous slide) is the number of bits in the address that gives us the network portion.

Subnet cont...

Example

(a) Dotted decimal and binary representations of IP address and subnet masks

	Binary Representation	Dotted Decimal
IP address	11000000.11100100.00010001.00111001	192.228.17.57
Subnet mask	11111111.11111111.11111111.11100000	255.255.255.224
Bitwise AND of address and mask (resultant network/subnet number)	11000000.11100100.00010001.00100000	192.228.17.32
Subnet number	11000000.11100100.00010001.001	1
Host number	00000000.00000000.00000000.00011001	25

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

B. default subnet mask

Subnet cont…

How to Create Subnets

- To create a subnet answer the following questions:
  - How many subnets does the chosen subnet mask produced? It can be calculated using 2^{n-dp}, where n is the given network prefix and dp is the default prefix for the given class (i.e. class A, B or C) when IP address is a Classful IP **or** dp is the network prefix given from ISP.
  - How many valid hosts per subnet are available? It can be calculated using $2^{32-n}-2$.
  - What are the valid subnets addresses?
  - What's the broadcast address of each subnets?
  - What are the valid hosts in each subnet?

Subnet cont….

Example 1. consider the following class C Ip-address: 192.168.1.0/25 (255.255.255.128)

192.168.1.0 = Network address

255.255.255.128 = Subnet mask (i.e. 11111111.11111111.11111111.10000000)

- How many subnets? Since 128 is 1 bit exceeded from the default 24 bits on (10000000), the answer would be $2^{25-24}=2^1=2$.
- How many hosts per subnet? We have 7 (32-25) host bits off (10000000), so the equation would be $2^7-2 = 126$ hosts.
- What are the valid subnets? 256-128 = 128. Remember, we'll start at zero and count in our block size, so our subnets are 0, 128.

Subnet cont.....

Example 1. cont……

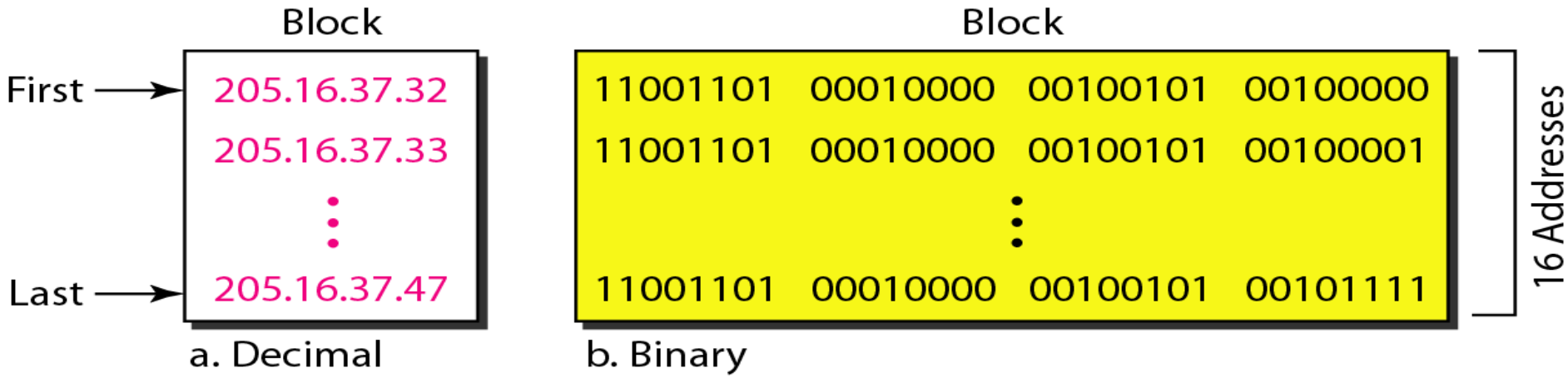
- What is the broadcast address for each subnet? The number right before the next subnet address equals the broadcast address. For the zero subnet 127 and for 128 subnet 255 are a broadcast addresses.
- What are the valid hosts? These are the numbers between the subnet and broadcast address.

	Subnet One	Subnet Two
Subnet	0	128
First host	1	129
Last host	126	254
Broadcast	127	255

Classless addressing

Restriction

- 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.



First Address

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

Example 5

- 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.

Last Address

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

Example 6

Find the last address for the block in Example 5.

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

Number of Addresses

- The number of addresses in the block is the difference between the last and first address. It can easily be found using the formula 2^{32-n}.

Example 7

- Find the number of addresses in Example 5.

Solution

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

Method II

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 the above example 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.

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

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.

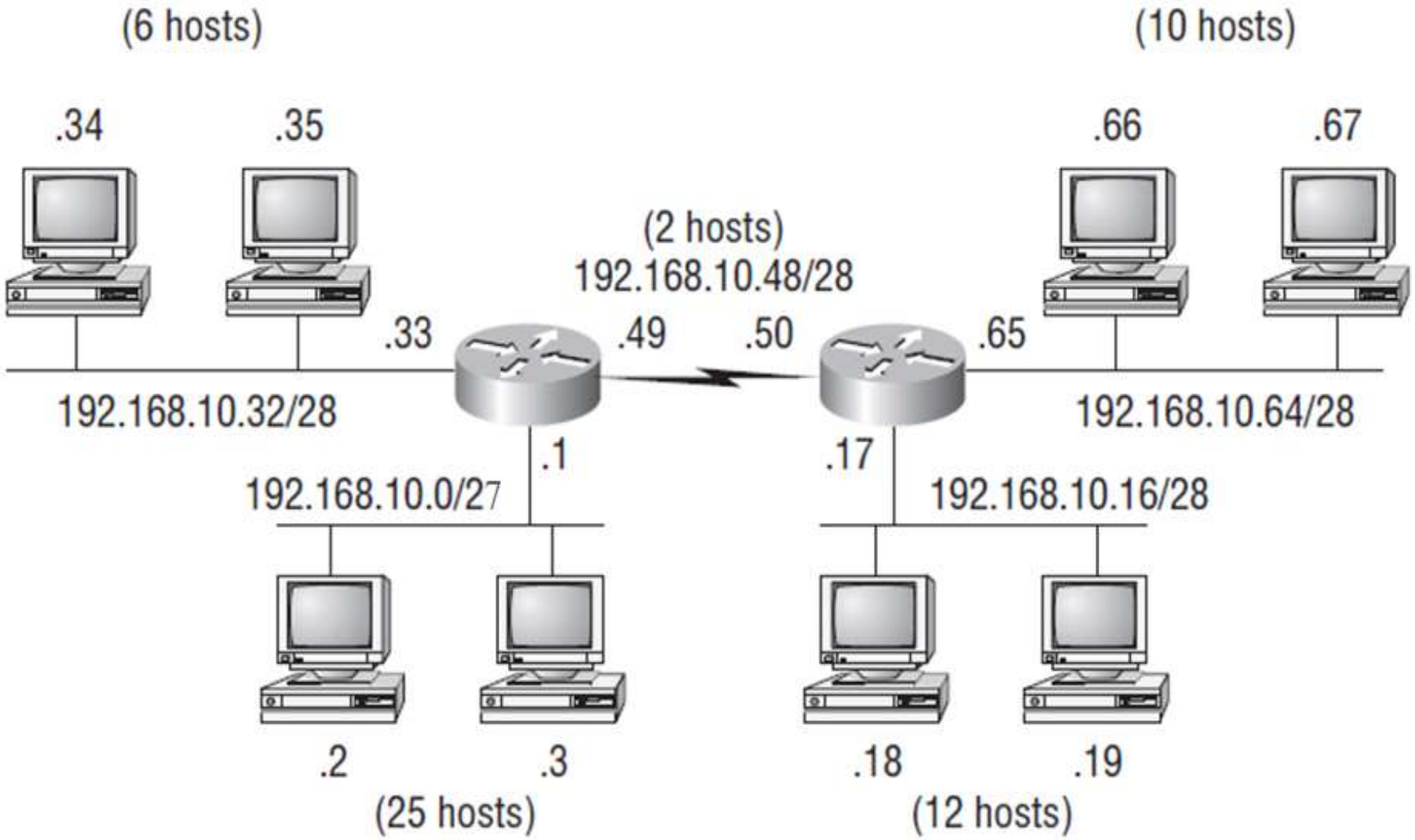
Address:	11001101	00010000	00100101	00100111
Mask complement:	**00000000**	**00000000**	**00000000**	**00001111**
Last address:	11001101	00010000	00100101	00101111

Contd.

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

Subnet cont'd…..
Variable length subnet masking(VLSM)

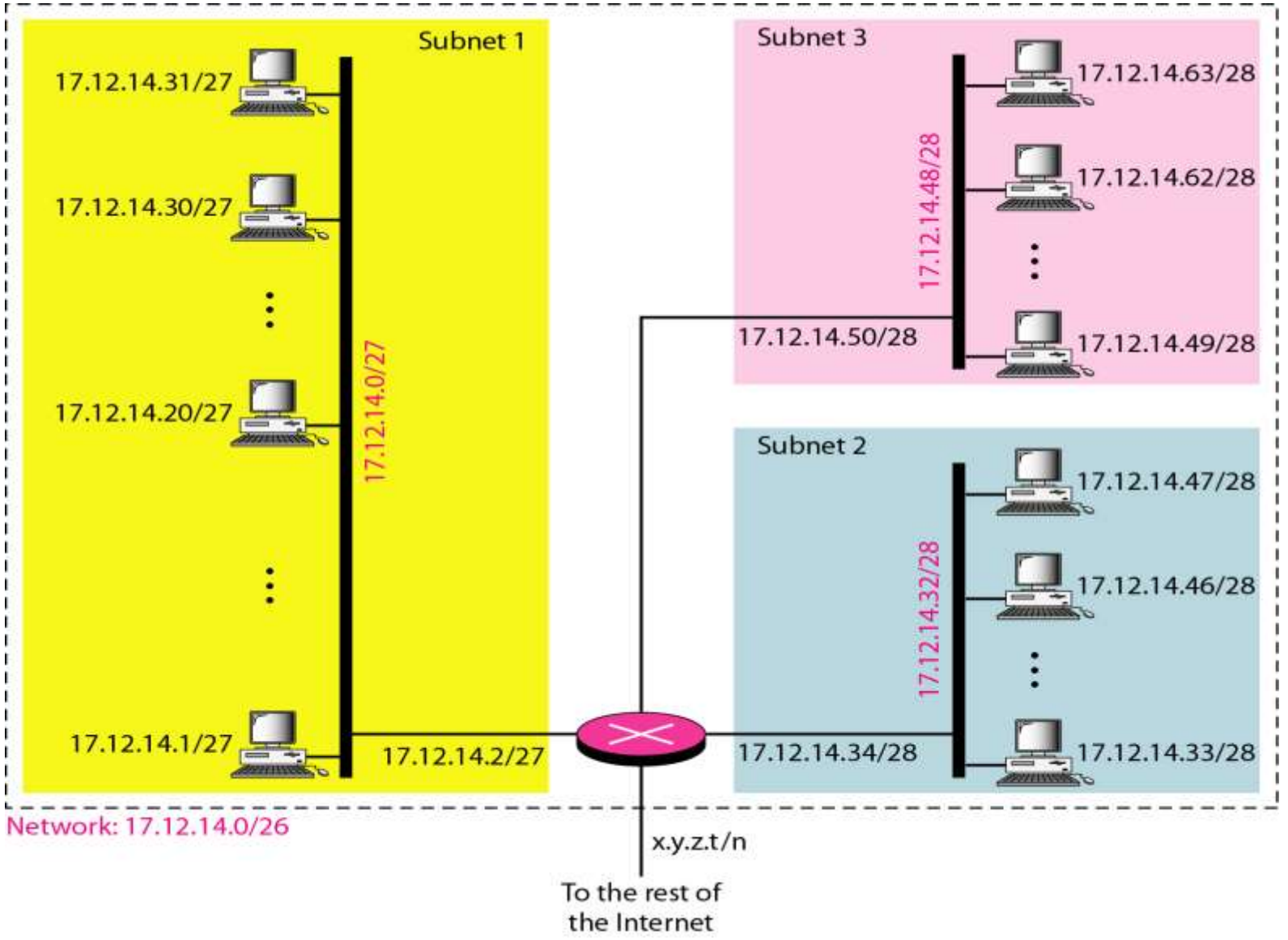


Example 7

- Suppose an organization is given the block 17.12.14.0/26, which contains 64 addresses. The organization has three offices and needs to divide the addresses into three sub blocks of **32**, **16**, and **16** addresses.
- We can find the new masks by using the following arguments:
  1. Suppose the mask for the first subnet is n1, then 2^{32-n1} must be 32, which means that **n1 = 27**.
  2. Suppose the mask for the second subnet is n2, then 2^{32-n2} must be 16, which means that **n2 = 28**.
  3. Suppose the mask for the third subnet is n3, then 2^{32-n3} must be 16, which means that **n3 = 28**.
- This means that we have the masks 27, 28, 28 with the organization mask being 26.

Contd..

Read reserved IPV4 address range

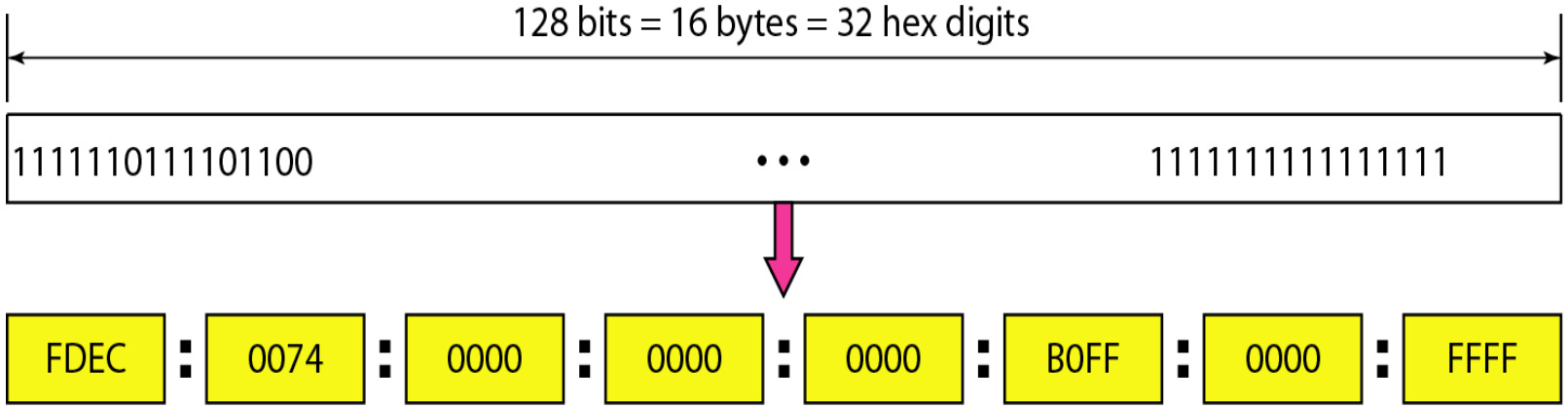


IPv6: Structure and Address Space

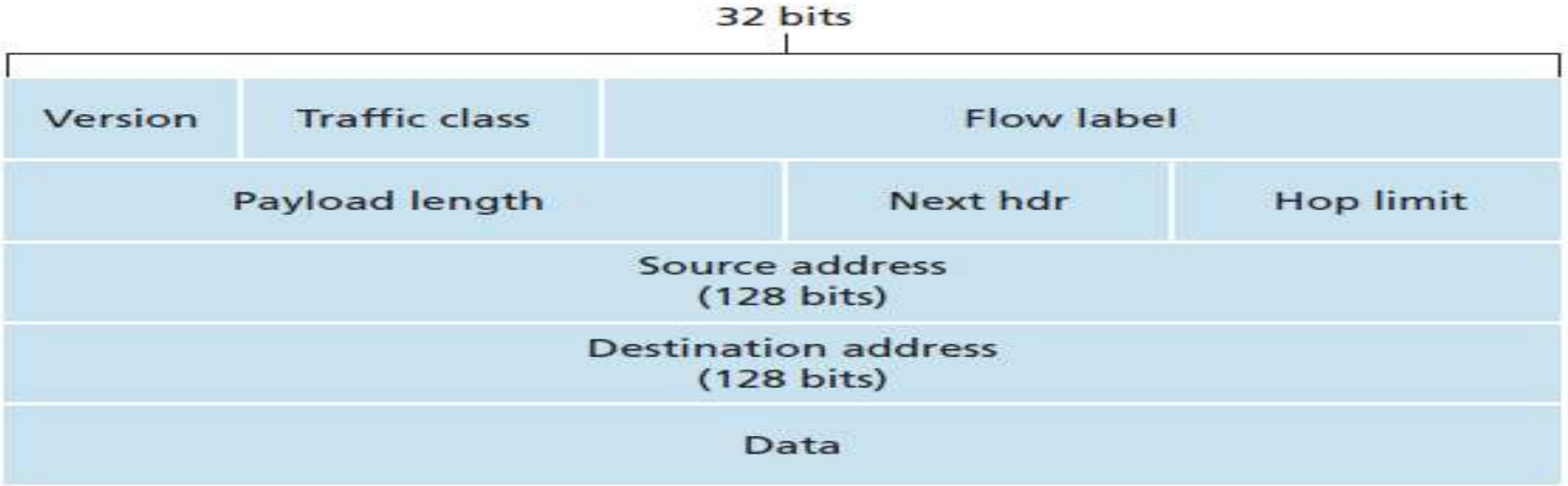
- The IP (IPv4) has been the foundation of the Internet and virtually all multivendor private internetworks.
- This protocol is reaching the end of its useful life and a new protocol, known as IPv6 (IP version 6), has been defined to ultimately replace IP.
- The driving motivation for the adoption of a new version of IP was the limitation imposed by the 32-bit address field in IPv4.
- With a 32-bit address field, it is possible in principle to assign 2^{32}different addresses, which is over 4 billion possible addresses.
- Reasons for the inadequacy of 32-bit addresses include the following:
  - The is convenient but wasteful of the address space. two-level structure of the IP address
  - Networks are multiplying rapidly.
  - Growth of TCP/IP usage into new areas will result in a rapid growth in the demand for unique IP addresses.
  - Multiple IP addresses are required for a single host.

IPv6 cont.....

- IPv6 includes the following enhancements over IPv4:
  - **Expanded address space**: IPv6 uses 128-bit addresses.
  - **Improved option mechanism**: IPv6 options are placed in separate optional headers that are located between the IPv6 header and the transport-layer header. It also makes it easier to add additional options.
  - **Increased addressing flexibility**: IPv6 includes the concept of an anycast address in addition to the existing casting mechanisms in IPv4 (i.e. Unicast, Multicast and Broadcast).
  - **Support for resource allocation**: IPv6 enables the labeling of packets belonging to a particular traffic flow for which the sender requests special handling. This aids in the support of specialized traffic such as real-time video.



IPv6 Header



- The following fields are defined in IPv6:
  - **Version**: this 4-bit field identifies the IP version number. IPv6 carries a value of 6 in this field.
  - **Traffic class**: this 8-bit field is similar in spirit to the TOS field we saw in IPv4.
  - **Flow label:** this 20-bit field is used to identify a flow (the datagram that needs special treatments due to different reasons) of datagrams.

IPv6 Addresses

- IPv6 addresses are 128 bits in length. Addresses are assigned to individual interfaces on nodes, not to the nodes themselves.
- IPv6 allows three types of addresses:
  - **Unicast:** An identifier for a single interface.
  - **Anycast:** An identifier for a set of interfaces. A packet sent to an anycast address is delivered to one of the interfaces identified by that address (the "nearest" one, according to the routing protocols' measure of distance).
  - **Multicast:** An identifier for a set of interfaces.

General unicast address format (routing prefix size varies)

bits	48 (or more)	16 (or fewer)	64
field	*routing prefix*	*subnet id*	*interface identifier*

General multicast address format

bits	8	4	4	112
field	*prefix*	*flg*	*sc*	*group ID*

bits	10	54	64
field	*prefix*	*zeroes*	*interface identifier*

IPv6 Addresses cont.….

- IPv6 addresses are represented by treating the 128-bit address as a sequence of 8 16-bit numbers, and representing this in the form of eight hexadecimal numbers divided by colons, for example:

  2001:0DB8:0055:0000:CD23:0000:0000:0205

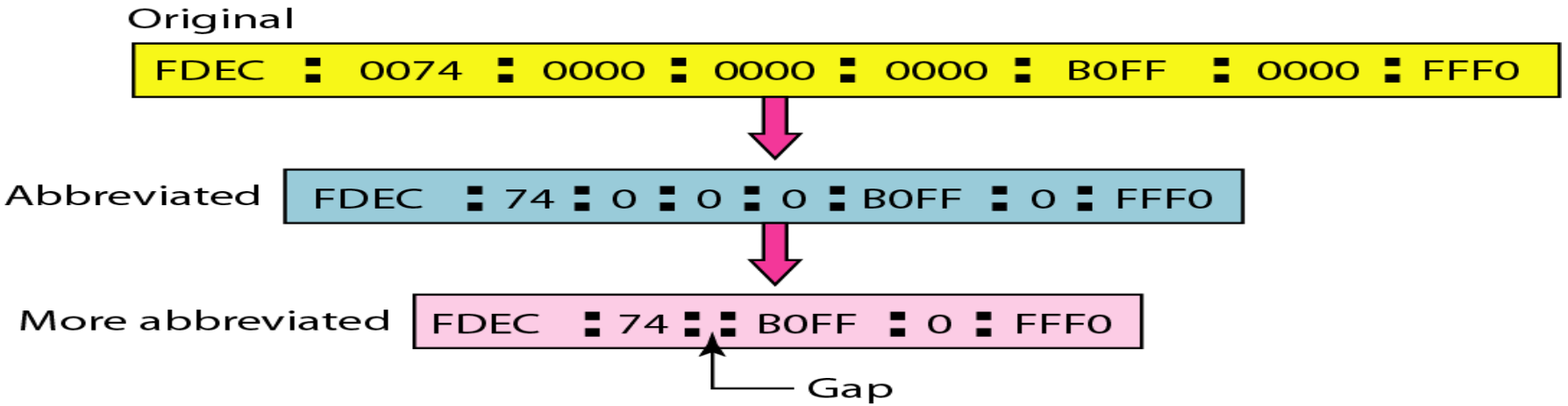
- One to three zeroes that appear as the leading digits in any colon-delimited hexadecimal grouping may be dropped (2001:0DB8:55:0:CD23:0:0:0205)
- A group of all zeroes, or consecutive groups of all zeroes, can be substituted by a double colon, but this may only be done once in an address. **Example:** 2001:DB8:55:0:CD23:0:0:0205**or** 2001:DB8:55:0:CD23::205

Structure of IPv6

ipv6-address/prefix-length

- prefix-length is a decimal value specifying how many of the leftmost contiguous bits of the address comprise the prefix. Example: 2001:0DB8:55:0:CD23::0205/48 **or** ::/64 **or** ABC::B:0:1234/80

Example 2



- Expand the address 0:15::1:12:1213 to its original.

Solution

- We first need to align the left side of the double colon to the left of the original pattern and the right side of the double colon to the right of the original pattern to find how many 0s we need to replace the double colon.

```
xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx
 0:   15:                  :    1:   12:1213
```

This means the original address is:

```
0000:0015:0000:0000:0000:0001:0012:1213
```

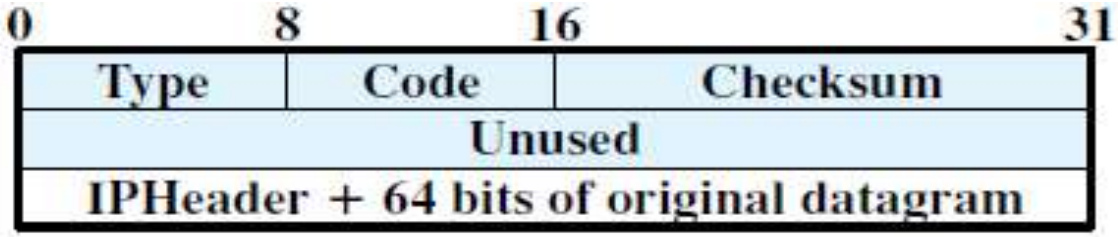
Internet Control Message Protocol (ICMP)

- ICMP provides a means for transferring messages from routers and hosts each other.
- In essence, ICMP provides feedback about problems in the communication environment.

Examples: when a datagram cannot reach its destination, or when the router does not have the buffering capacity to forward a datagram.

- An ICMP message is sent in response to a datagram, either by a router along the datagram's path or by the intended destination host.
- Although ICMP is, in effect, at the same level as IP in the TCP/IP architecture, it is a user of IP.
- An ICMP message is constructed and then passed down to IP, which encapsulates the message with an IP header and then transmits the resulting datagram in the usual fashion.

ICMP cont….



Destination unreachable; time exceeded; source quench

- An ICMP message starts with a 64-bit header consisting of the following:
  - **Type (8 bits):** Specifies the type of ICMP message.
  - **Code (8 bits):** Used to specify parameters of the message that can be encoded in one or a few bits.
  - **Checksum (16 bits):** Checksum of the entire ICMP message.
  - **Parameters (32 bits):** Used to specify more lengthy parameters.

```
Router#ping 192.168.2.25

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.2.25, timeout is 2
seconds:
.....
Success rate is 0 percent (0/5)

Router#
```

ICMP cont....

- ICMP error report kinds:
  - The **destination unreachable** message covers a number of contingencies.
  - A router will return a **time exceeded** message if the lifetime of the datagram expires.
  - A syntactic or semantic error in an IP header will cause a **parameter problem** message.
  - The **echo and echo reply messages (pinging and replaying the ping message)**provide a mechanism for testing that communication is possible between entities.

```
C:\>ping 200.1.2.120

Pinging 200.1.2.120 with 32 bytes of data:

Request timed out.
Request timed out.
Request timed out.
Request timed out.

Ping statistics for 200.1.2.120:
    Packets: Sent = 4, Received = 0, Lost = 4 (100% loss),

C:\>ping 200.1.2.255

Pinging 200.1.2.255 with 32 bytes of data:

Reply from 200.1.2.1: Destination host unreachable.
Reply from 200.1.2.1: Destination host unreachable.
Reply from 200.1.2.1: Destination host unreachable.
Reply from 200.1.2.1: Destination host unreachable.

Ping statistics for 200.1.2.255:
    Packets: Sent = 4, Received = 0, Lost = 4 (100% loss),
```

```
C:\>ping 200.1.2.1

Pinging 200.1.2.1 with 32 bytes of data:

Reply from 200.1.2.1: bytes=32 time<1ms TTL=255
Reply from 200.1.2.1: bytes=32 time<1ms TTL=255
Reply from 200.1.2.1: bytes=32 time=1ms TTL=255
Reply from 200.1.2.1: bytes=32 time<1ms TTL=255

Ping statistics for 200.1.2.1:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 0ms, Maximum = 1ms, Average = 0ms
```

ICMP cont……

- The **source quench** message provides a rudimentary form of flow control.
- Either a router or a destination host may send this message to a source host, requesting that it reduce the rate at which it is sending traffic to the internet destination.
- A router sends a **redirect message** to a host on a directly connected router to advise the host of a better route to a particular destination.
- The **timestamp and timestamp** reply messages provide a mechanism for sampling the delay characteristics of the internet.
- The **address mask request and address mask** reply messages are useful in an environment that includes subnets. The address mask request and reply messages allow a host to learn the address mask for the LAN to which it connects.

ICMPv6

- ICMPv6 is used by IPv6 nodes to report errors encountered in processing packets, and to perform other internet-layer functions, such as diagnostics (ICMPv6 "ping"). ICMPv6 is an integral part of IPv6 and MUST be fully implemented by every IPv6 node.
- ICMPv6 messages are grouped into two classes: error messages and informational messages.

ICMPv6 packet

Bit offset	0–7	8–15	16–31
0	Type	Code	Checksum
32	Message body		

ICMPv6 error messages

1 Destination Unreachable

2 Packet Too Big

3 Time Exceeded

4 Parameter Problem

ICMPv6 informational messages:

128 Echo Request

129 Echo Reply

Address Mapping

- **Logical address (IP address):** it is a network address uniquely identify networks in the wide area network.
- **Physical address:** it is a MAC (Medium Access Control) address, which provides a physical address for a host port attached to the LAN.
- To deliver an IP datagram to a destination host, a mapping must be made from the IP address to the subnetwork (MAC) address for that last hop.
- If a datagram traverses one or more routers between source and destination hosts, then the mapping must be done in the final router, which is attached to the same subnetwork as the destination host.
- If a datagram is sent from one host to another on the same subnetwork, then the source host must do the mapping.

Address Mapping cont….

- For this purpose, a number of approaches are possible, it include:
  - Each system can maintain a local table of IP addresses and matching subnetwork addresses for possible correspondents.
  - A centralized directory can be maintained on each subnetwork that contains the IP-subnet address mappings.
  - An address resolution protocol can be used.

```
Router>sh arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  200.1.1.9              -   0060.70AE.C003  ARPA   GigabitEthernet0/2
Internet  200.1.1.10            28   0090.0CC4.D202  ARPA   GigabitEthernet0/2
Internet  200.1.1.129           -   0060.70AE.C002  ARPA   GigabitEthernet0/1
Internet  200.1.1.130           28   0030.A365.0699  ARPA   GigabitEthernet0/1
Internet  200.1.1.131           28   0001.965E.ADB1  ARPA   GigabitEthernet0/1
Internet  200.1.1.132           28   00E0.B001.EDA9  ARPA   GigabitEthernet0/1
Internet  200.1.1.133           28   0090.0CEA.23A9  ARPA   GigabitEthernet0/1
Router>
```

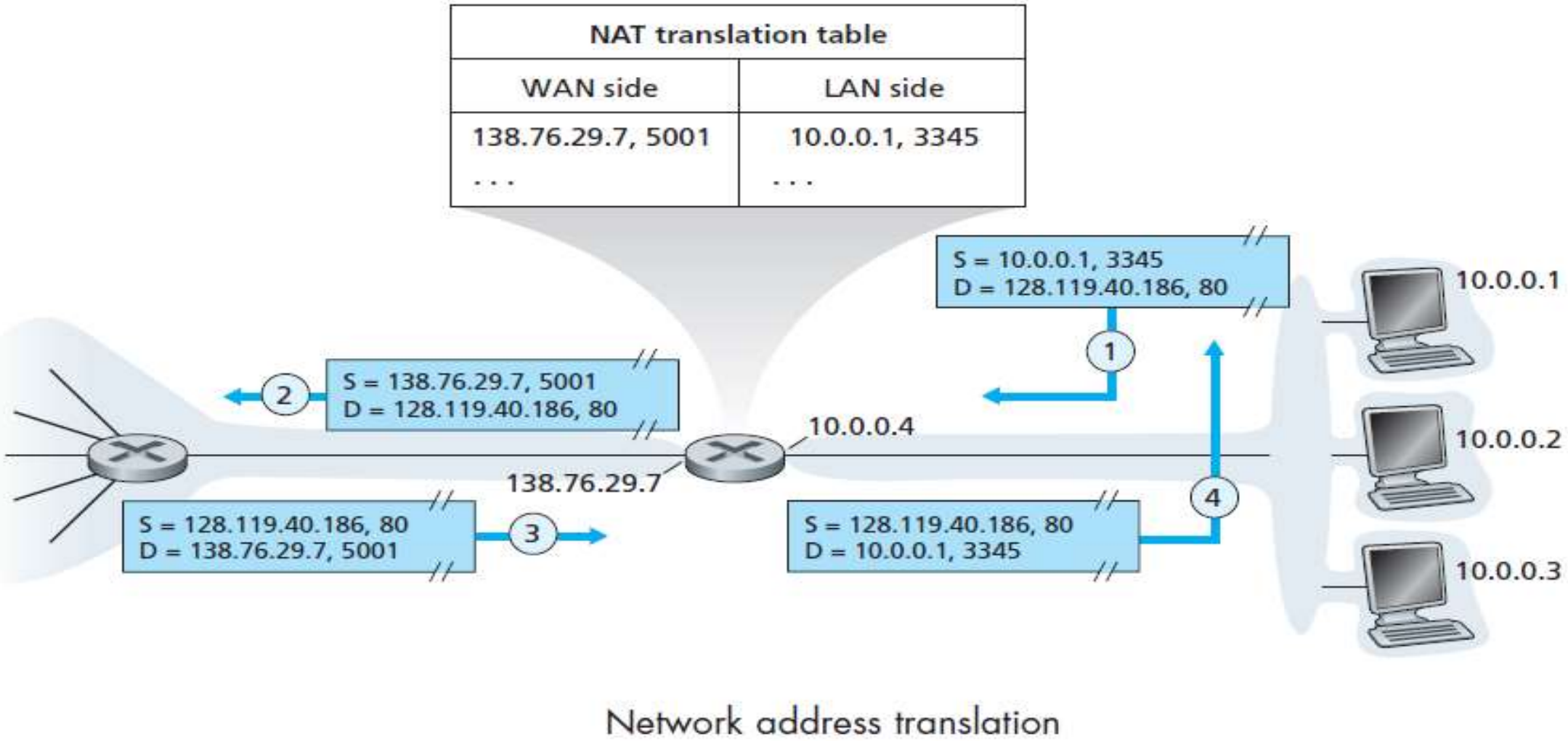
```
C:\>arp -a
  Internet Address      Physical Address      Type
  192.168.3.1           00e0.b0b3.c201        dynamic
  192.168.3.2           0040.0b07.28e8        dynamic
```

Address Resolution Protocol (ARP)

- ARP allows dynamic distribution of the information needed to build tables to translate an IP address A into a 48-bit Ethernet address; the protocol can be used for any broadcast network. ARP exploits the broadcast property of a LAN.
- ARP works as follows:

1. Each system on the LAN maintains a table of known IP-subnetwork address mappings.
2. When a subnetwork address is needed for an IP address, and the mapping is not found in the system's table, the system uses ARP directly on top of the LAN protocol (e.g., IEEE 802) to broadcast a request. The broadcast message contains the IP address for which a subnetwork address is needed.
3. Other hosts on the subnetwork listen for ARP messages and reply when a match occurs. The reply includes both the IP and subnetwork addresses of the requested and replying host.
4. The original request includes the requesting host's IP address and subnetwork address. Any interested host can copy this information into its local table, avoiding the need for later ARP messages.
5. The ARP message can also be used simply to broadcast a host's IP address and subnetwork address, for the benefit of others on the subnetwork.

Network Address Translation (NAT)



Network address translation

NAT cont....

- The basic idea behind NAT is for the ISP to assign each home or business a single (few) public IP address(es) for Internet traffic.
- Within the customer network, every computer gets a unique IP address, which is used for routing in the intranet traffic. However, just before a packet exits the customer network and goes to the ISP, an address translation from the unique internal IP address to the shared public IP address takes place.
- This translation makes use of three ranges of IP addresses that have been declared as private. The only rule is that no packets containing these addresses may appear on the Internet itself.
- If all datagrams arriving at the NAT router from the WAN have the same destination IP address, then how does the router know the internal host to which it should forward a given datagram?
- The router uses a **NAT translation table**, and to include port numbers as well as IP addresses in the table entries.

Protocol 10.1.1.1	Inside Local IP Address: Port	Inside Global IP Address: Port	Outside Global IP Address: Port
TCP	10.1.1.3:1723	170.168.2.2:1492	63.41.7.3:23
TCP	10.1.1.2:1723	170.168.2.2:1723	63.41.7.3:23
TCP	10.1.1.1:1024	170.168.2.2:1024	63.40.7.3:23

NAT cont….

- NAT has enjoyed widespread deployment in recent years. But it has some limitations:
- **First** port numbers are meant to be used for addressing processes, not for addressing hosts.
- **Second** routers are supposed to process packets only up to layer 3.
- **Third** the NAT protocol violates the so-called end-to-end argument; that is, hosts should be talking directly with each other, without interfering nodes modifying IP addresses and port numbers.
- **Fourth** it interferes with P2P applications, including P2P file-sharing applications and P2P Voice-over-IP applications.

Routing Protocols and their classifications..

- **Autonomous system (AS)** is essentially a collection of IP networks and routers under the same administration that share a common routing strategy.

  Example ISP

- IGPs provide information on reachable interior destinations to the outside work
- Classified to distance vector routing and link state routing protocols
- EGPs provide information on reachable exterior destinations to the interior routers.
- EGPs are typically used between ISPs.
- Another difference between them is the **route determination**.
- IGPs choose the best path based on **distance, delay, or bandwidth**.
- Therefore, how the metrics are used is an important design issue for IGPs.
- EGPs choose routes commonly according to a **routing policy**.

IGP:- Within a single autonomous system

- ☞Single network administration

Routing Protocols and their classifications

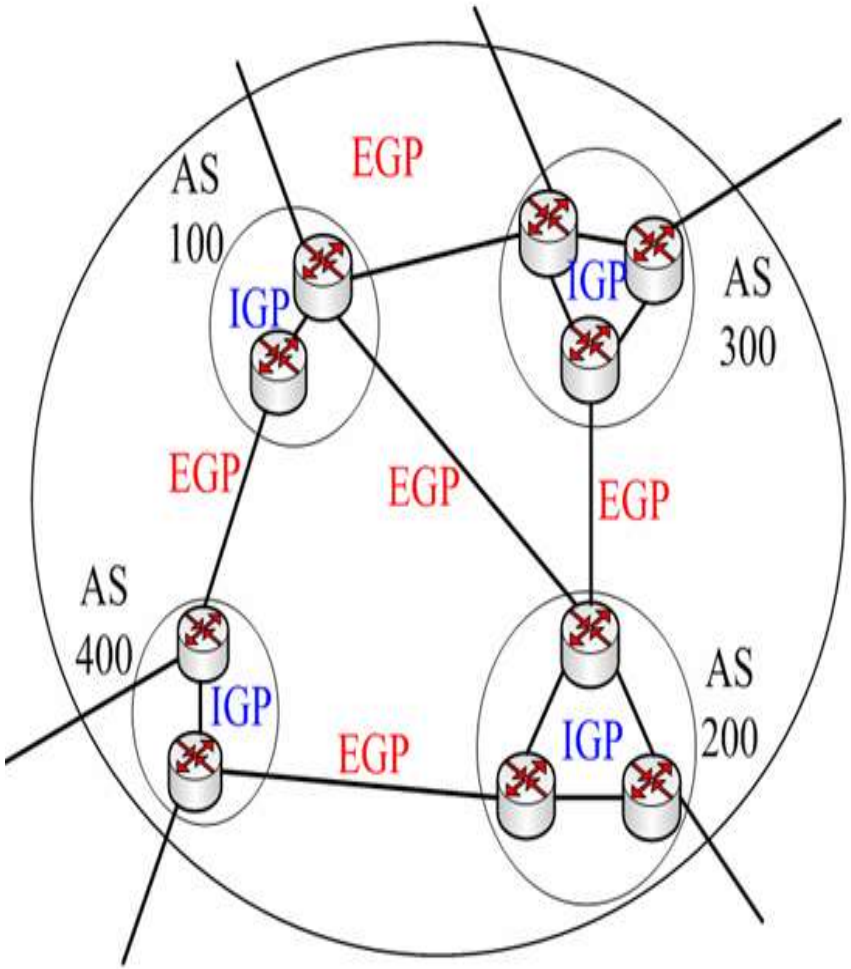
- Different routing protocols use different algorithms and approaches to gather and disseminate routing information about their area.
- Routing protocols are classified into the following categories:

	Interior Gateway Protocols				Exterior Gateway Protocols
	Distance Vector		Link-State		Path Vector
IPv4	RIPv2	EIGRP	OSPFv2	IS-IS	BGP-4
IPv6	RIPng	EIGRP for IPv6	OSPFv3	IS-IS for IPv6	BGP-MP

Interior Vs. Exterior Routing Protocols

- The routing protocols used within an autonomous system (AS) are called **interior gateway protocols (IGPs)**, and the ones used between ASs are called **exterior gateway protocols (EGPs)**.

Contd.



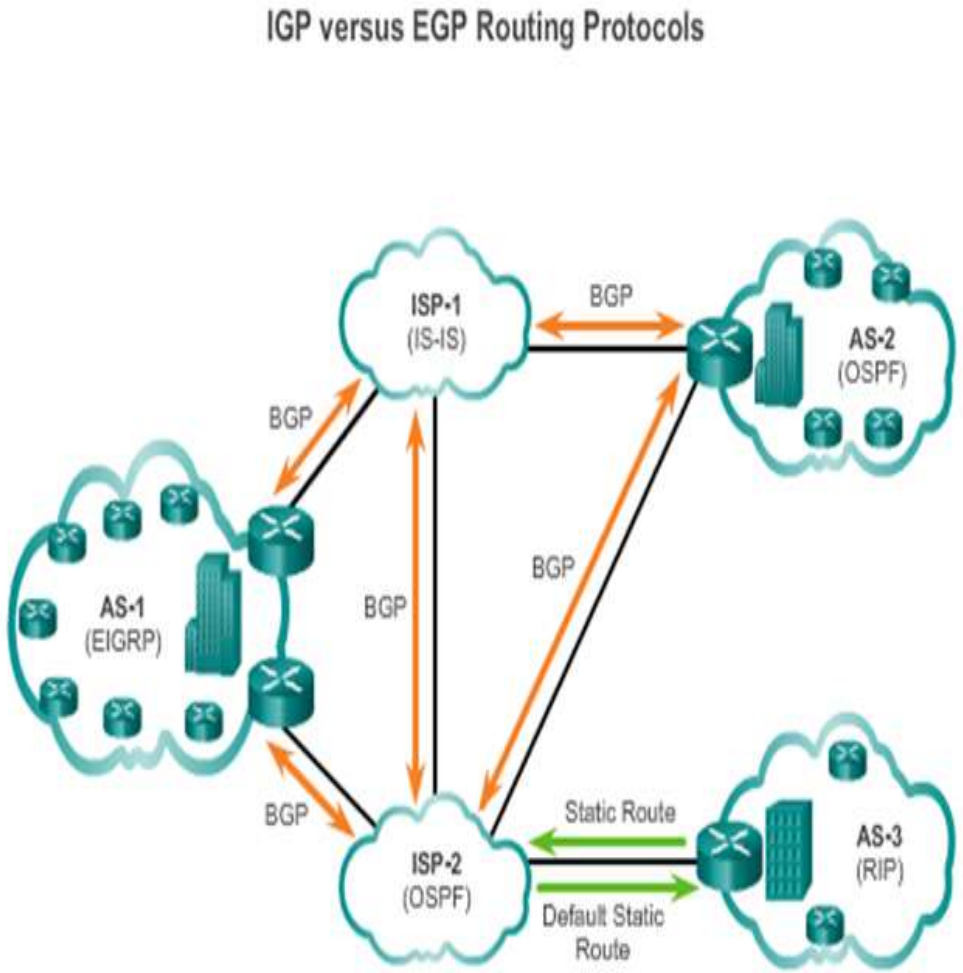


Fig. 4.2.1 The Internet comprises a large number of autonomous systems

Distance Vector Vs. Link State Routing protocols

Distance Vector

- Updates frequently
- Each router is "aware" of only its immediate neighbors
- Slow convergence
- Prone to routing loops
- Easy to configure
- Fewer router resources required
- Updates require more bandwidth
- Does not "understand" the topology of the network

Link State

- Updates are event triggered
- Each router is "aware" of all other routers in the "area"
- Fast convergence
- Less subject to routing loops
- More difficult to configure
- More router resource intensive
- Updates require less bandwidth
- Has detailed knowledge of distant networks and routers

Routing Algorithm

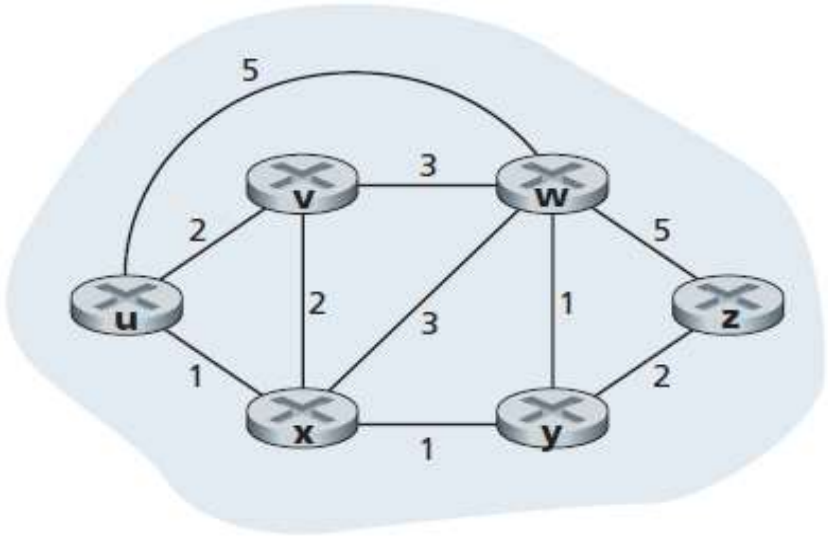
- A routing algorithm is a set of rules and procedures used by routers in a network to determine the best path for forwarding network traffic from a source to a destination.
- Routing algorithms consider various factors such as network topology, link metrics, and routing policies to make routing decisions.
- It is that part of the network layer software responsible for deciding which output line an incoming packet should be transmitted on.
- If the network uses datagrams internally, this decision must be made a new path for every arriving data packet.
- If the network uses virtual circuits internally, routing decisions are made only when a new virtual circuit (sometimes called session) is being set up.
- A host is attached directly to one router, the default router for the host.
- The purpose of a routing algorithm is then simple: given a set of routers, with links connecting the routers, a routing algorithm finds a “good” path from source router to destination router.
- Typically, a good path is one that has the least cost.

Routing Algorithm cont….

- The design of routing algorithm may have the following properties into consideration: correctness, simplicity, robustness, stability, fairness, and efficiency.
- Routing algorithms can be grouped into two major classes: non-adaptive and adaptive.
  - **Non-adaptive (static) algorithms:** use fixed routes that are preconfigured or predetermined before network operation
  - **Adaptive (dynamic) algorithms:** continuously monitor the network, react to changes in network conditions, and dynamically adjust routing decisions to optimize the flow of traffic.
    - These dynamic routing algorithms differ in where they get their information, when they change the routes, *and* what metric is used for optimization.
- Others classify routing algorithms into: global or decentralized, link-state or distance-vector and load sensitive or load-insensitive.

Shortest Path Algorithm

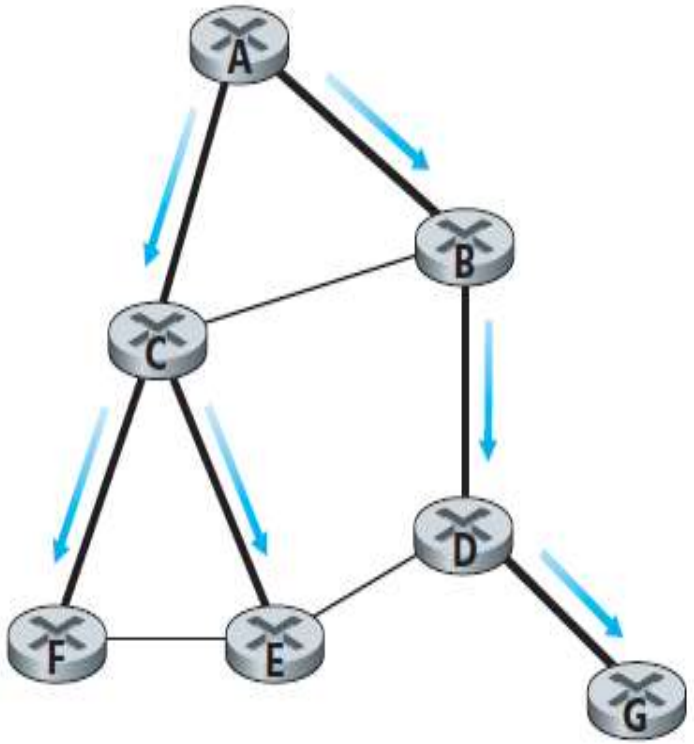
- The idea is to build a graph of the network, with each node of the graph representing a router and each edge of the graph representing a communication line, or link.
- To choose a route between a given pair of routers, the algorithm just finds the shortest path between them on the graph.
- The concept of a shortest path deserves some explanation.
  - One way of measuring path length is the number of hops.
  - Another metric is the geographic distance in kilometers.



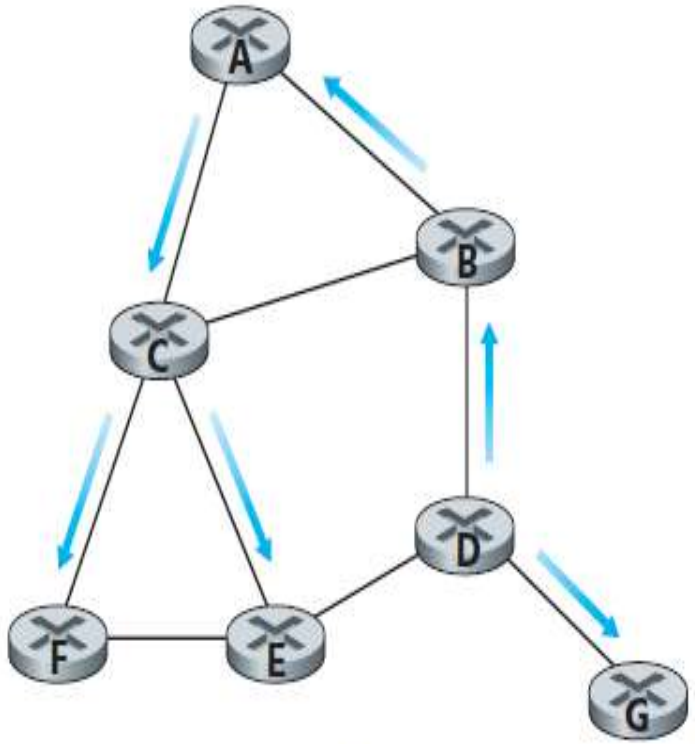
Flooding

- In this algorithm every incoming packet is sent out on every outgoing line except the one it arrived on.
- Flooding obviously generates vast numbers of duplicate packets, in fact, an infinite number unless some measures are taken to damp the process.
  - One such measure is to have a hop counter contained in the header of each packet that is decremented at each hop, with the packet being discarded when the counter reaches zero.
  - Another technique for damming the flood is to have routers keep track of which packets have been flooded, to avoid sending them out a second time.
- It is effective for broadcasting information, tremendously robust and requires little in the way of setup.

Flooding cont….



a. Broadcast initiated at A



b. Broadcast initiated at D

Hierarchical Routing

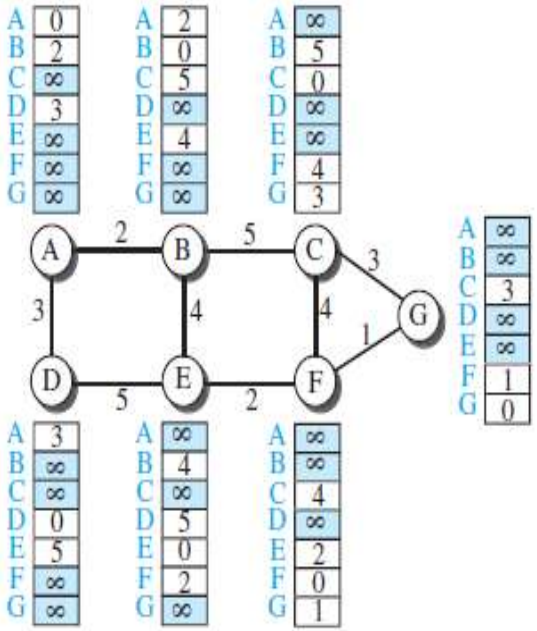
- As networks grow in size, the router routing tables grow proportionally. It affects router memory, CPU time and bandwidth to send status reports to others.
- At a certain point, it is no longer feasible for every router to have an entry for every other router, so the routing will have to be done hierarchically.
- When hierarchical routing is used, the routers are divided **regions.**
- Each router knows all the details about how to route packets to destinations within its own region but knows nothing about other regions.
- When a single network becomes very large, an interesting question is "how many levels should the hierarchy have?"

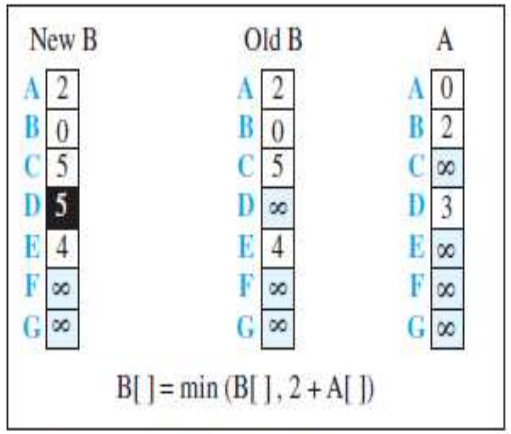
Distance Vector Routing

- A distance vector routing algorithm operates by having each router maintain a table (i.e., a vector) giving the best known distance to each destination and which link to use to get there.
- These tables are updated by exchanging information with the neighbors.
- In distance vector routing, each router maintains a routing table indexed by, and containing one entry for each router in the network.
- This entry has two parts:
  - The preferred outgoing line to use for that destination and
  - An estimate of the distance to that destination.
- The commonly used distance vector routing protocols are:
  - RIP (Routing Information Protocol),
  - BGP (Border Gateway Protocol),
  - IGRP (Interior Gateway Routing Protocol) and the original ARPAnet.

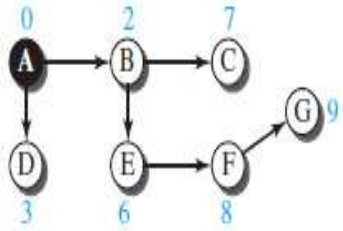
Distance Vector Routing cont….

The first distance vector for an internet

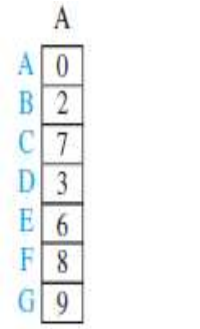




First event: B receives a copy of A's vector.



Tree for node A



Distance vector for node A

Link State Routing

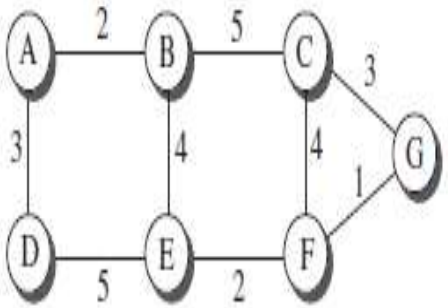
- The idea behind link state routing can be stated as five parts. Each router must do the following things to make it work:

1. Discover its neighbors and learn their network addresses.
2. Set the distance or cost metric to each of its neighbors.
3. Construct a packet telling all it has just learned.
4. Send this packet to and receive packets from all other routers.
5. Compute the shortest path to every other router.

- In effect, the complete topology is distributed to every router.

Link state cont....

Example of a link-state database

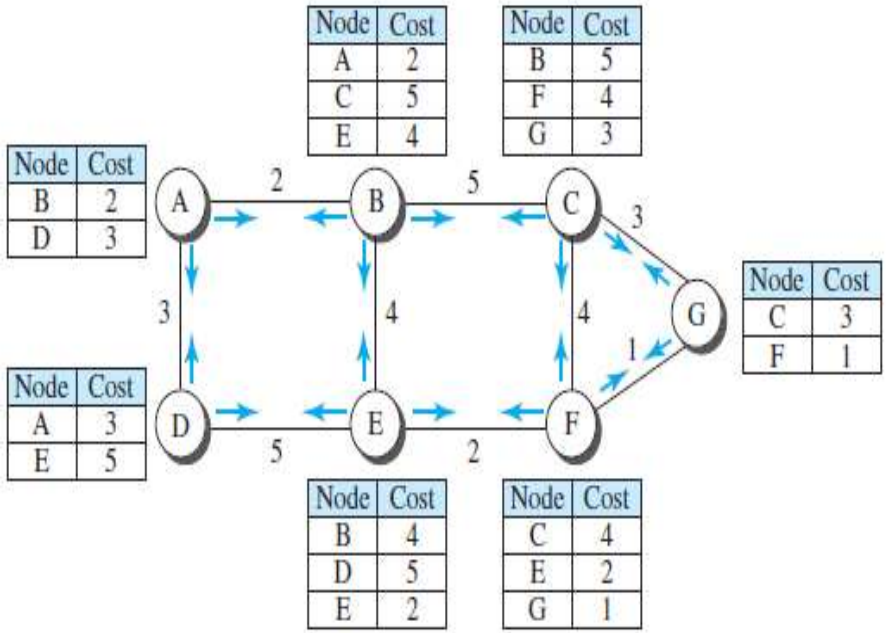


a. The weighted graph

	A	B	C	D	E	F	G
A	0	2	∞	3	∞	∞	∞
B	2	0	5	∞	4	∞	∞
C	∞	5	0	∞	∞	4	3
D	3	∞	∞	0	5	∞	∞
E	∞	4	∞	5	0	2	∞
F	∞	∞	4	∞	2	0	1
G	∞	∞	3	∞	∞	1	0

b. Link state database

LSPs created and sent out by each node to build LSDB



Link State Routing cont…..

- Compared to distance vector routing, link state routing requires more memory and computation.
- For a network with n routers, each of which has k neighbors, the memory required to store the input data is proportional to **kn**, which is at least as large as a routing table listing all the destinations.
- Also, the computation time grows faster than **kn**, even with the most efficient data structures, an issue in large networks.
- Nevertheless, in many practical situations, link state routing works well because it does not suffer from slow convergence problems.
- The two commonly used link state routing protocols **IS-IS** (Intermediate System-Intermediate System) and **OSPF** (Open Shortest Path First).

Next Chapter IV

Transport Layer