

Computer Networks

MODULE 1



What is Computer Network?

We will use the term "computer network" to mean a collection of autonomous computers interconnected by a single technology.

Two computers are said to be interconnected if they are able to exchange information. copper wire; fiber optics, microwaves, infrared, and communication satellites, etc.



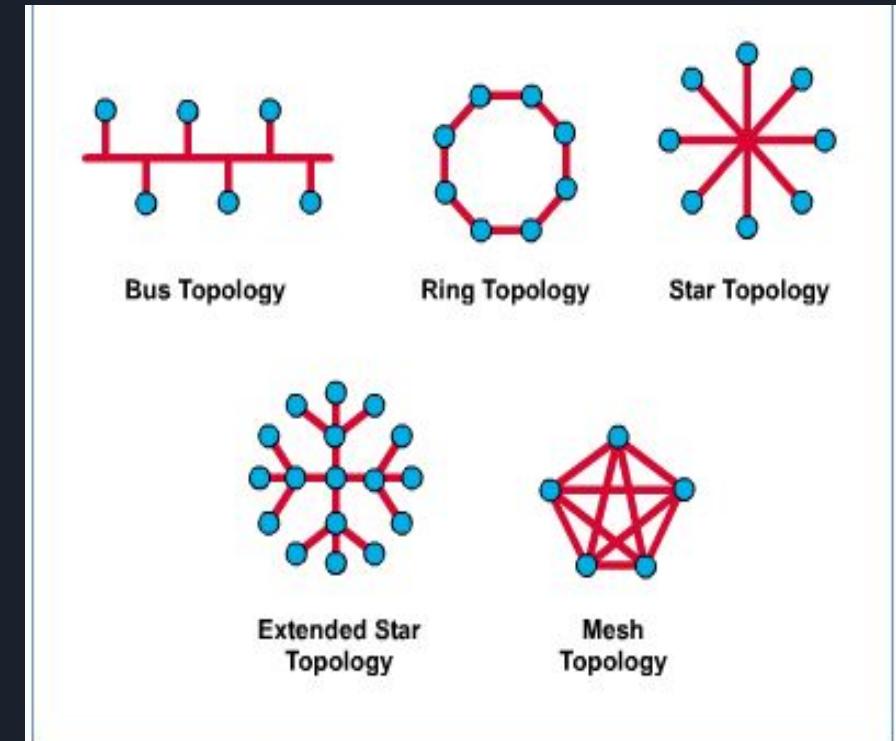
Uses of Computer Networks

1. Business Applications
2. Scientific Applications
3. Home Applications
4. Mobile Users
5. E-Commerce
6. Grid Computing

Tag	Full name	Example
B2C	Business-to-consumer	Ordering books online
B2B	Business-to-business	Car manufacturer ordering tires from supplier
G2C	Government-to-consumer	Government distributing tax forms electronically
C2C	Consumer-to-consumer	Auctioning second-hand products online
P2P	Peer-to-peer	Music sharing

Network Topology

- The network topology defines the way in which computers, printers, and other devices are connected.
- A network topology describes the layout of the wire and devices as well as the paths used by data transmissions.





Network Hardware

(TRANSMISSION TECHNOLOGIES)

1) Broadcast

Broadcast networks have a single communication channel that is shared by all the machines on the network. Short messages, called packets in certain contexts, sent by any machine are received by all the others. (Address Checking required)

2) Point-to-point

In point-to-point networks, there consist of many connections between individual pairs of machines. As a general rule (although there are many exceptions), smaller, geographically localized networks tend to use broadcasting, whereas larger networks usually are point-to-point.

Network Hardware

(SCALE PERSPECTIVE)

1. Personal Area Network (PAN)
2. Local Area Network (LAN)
3. Metropolitan Area Network (MAN)
4. Wide Area Network (WAN)
5. Internet

Interprocessor distance	Processors located in same	Example
1 m	Square meter	Personal area network
10 m	Room	
100 m	Building	
1 km	Campus	Local area network
10 km	City	
100 km	Country	
1000 km	Continent	Metropolitan area network
10,000 km	Planet	
		The Internet



Social Issues

- 1) Network neutrality
- 2) Digital Millennium Copyright Act
- 3) Profiling users
- 4) Phishing



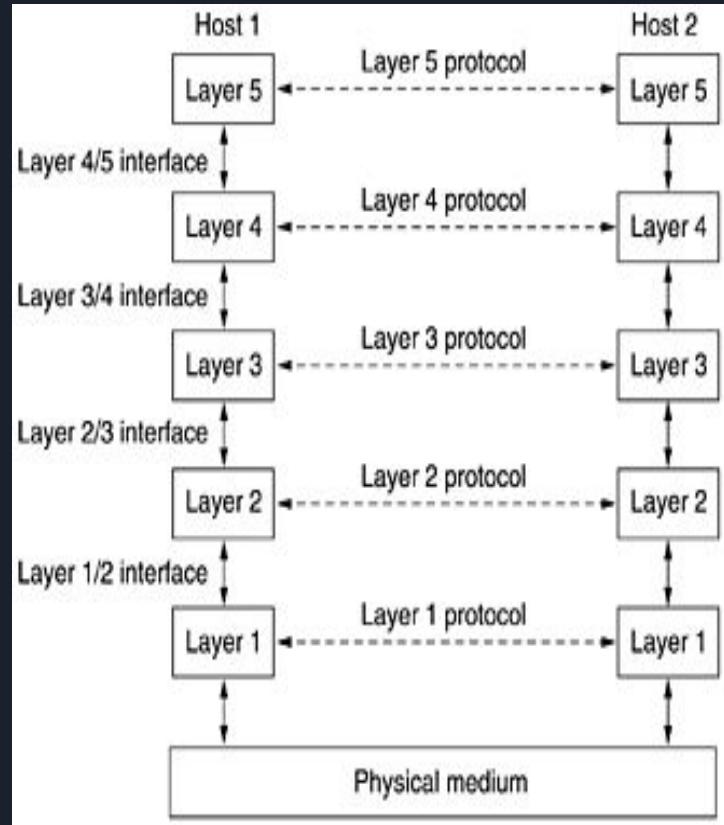
Network Software

- ❑ Protocol Hierarchies
- ❑ Design Issues for the Layers
- ❑ Connection-Oriented and Connectionless Services
- ❑ Service Primitives
- ❑ The Relationship of Services to Protocols

Network Software

Protocol Hierarchies

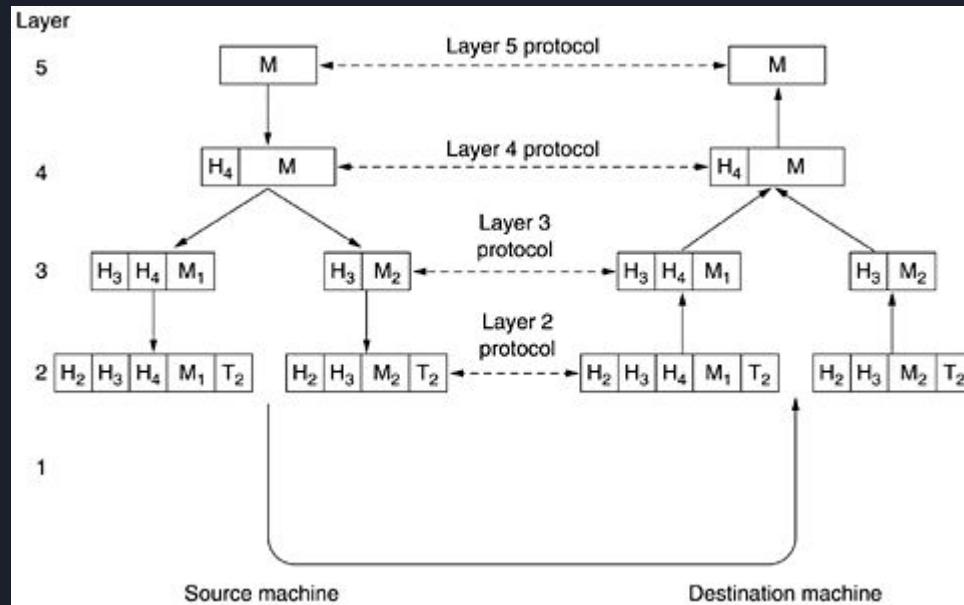
- To reduce their design complexity, most networks are organized as a stack of layers or levels, each one built upon the one below it.
- The purpose of each layer is to offer certain services to the higher layers, shielding those layers from the details of how the offered services are actually implemented.
- A protocol is an agreement between the communicating parties on how communication is to proceed.



Network Software

Protocol Hierarchies

- An example network protocol stack





Network Software

Design Issues for the Layers

- Addressing
- Flow Control
- Error Control
- Multiplexing
- Routing



Network Software

Connection-Oriented and Connectionless Services

- Connection-Oriented Service:
the service user first establishes a connection, uses the connection, and then releases the connection. (e.g., the telephone, tube)
 - Connectionless Service:
Each message carries the full destination address, and each one is routed through the system independent of all the others. (e.g., the postal system)
Usually, connectionless service can not guarantee the order of messages.
-
- ★ In order to enhance the reliability of transmission of connection-oriented service, acknowledge each received message is helpful. For example, the file transfer.
 - ★ However, some applications prefer fast speed than the reliability. For example, the digitized voice traffic, video conference.

Network Software

Connection-Oriented and Connectionless Services

- ❖ 6 Types Of Services

	Service	Example
Connection-oriented	Reliable message stream	Sequence of pages
	Reliable byte stream	Remote login
Connection-less	Unreliable connection	Digitized voice
	Unreliable datagram	Electronic junk mail
	Acknowledged datagram	Registered mail
	Request-reply	Database query



Network Software

Service Primitives

A service is formally specified by a set of primitives (basic operations) available to a user or other entity to access the service.

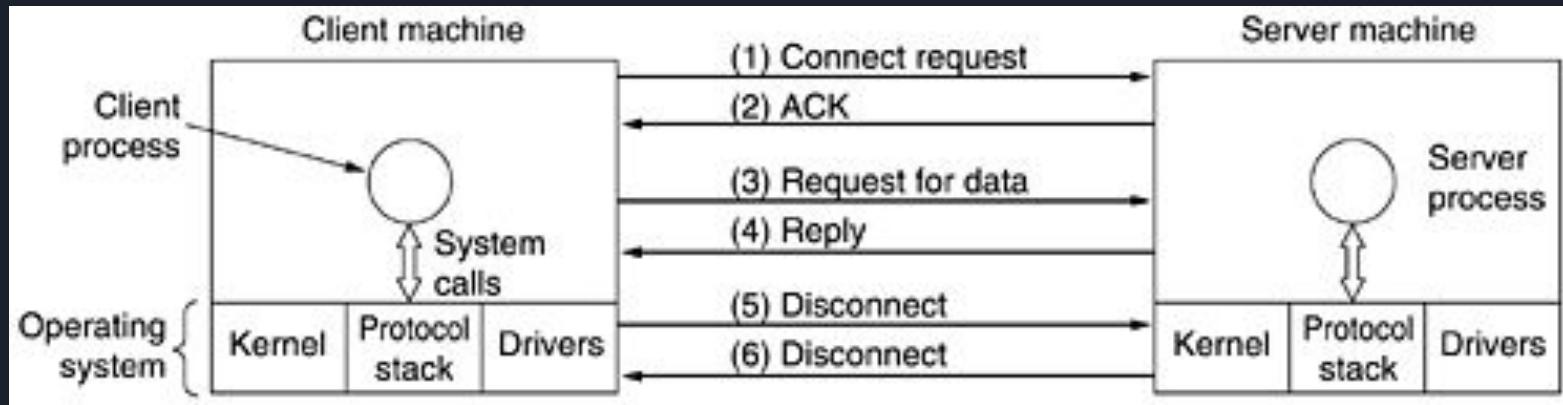
- Five service primitives that provide a simple connection-oriented services

Primitive	Meaning
LISTEN	Block waiting for an incoming connection
CONNECT	Establish a connection with a waiting peer
RECEIVE	Block waiting for an incoming message
SEND	Send a message to the peer
DISCONNECT	Terminate a connection

Network Software

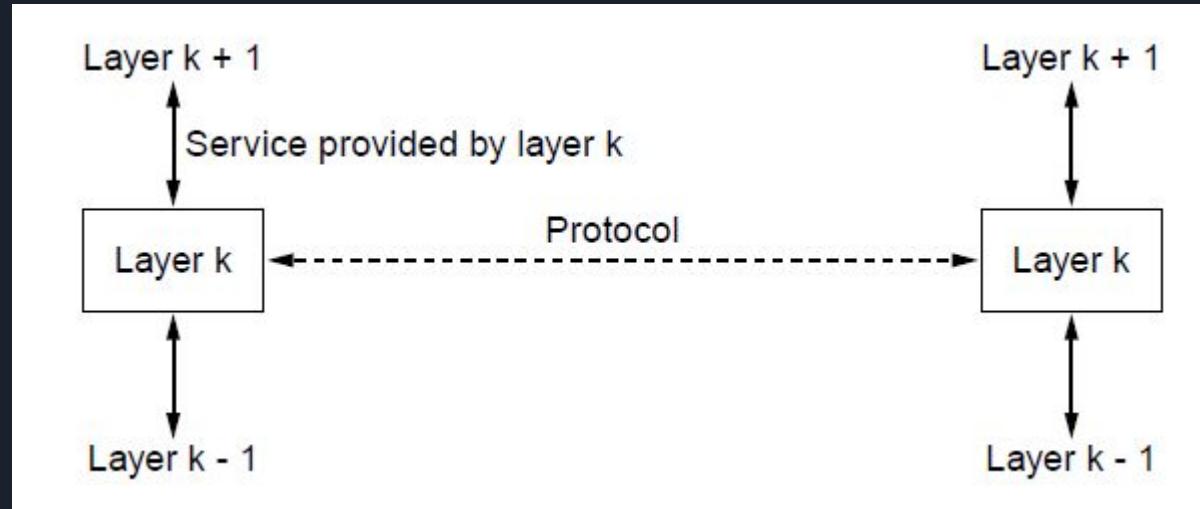
Service Primitives

- Packets sent in a simple client-server interaction on a connection-oriented network



Network Software

Services to Protocols Relationship

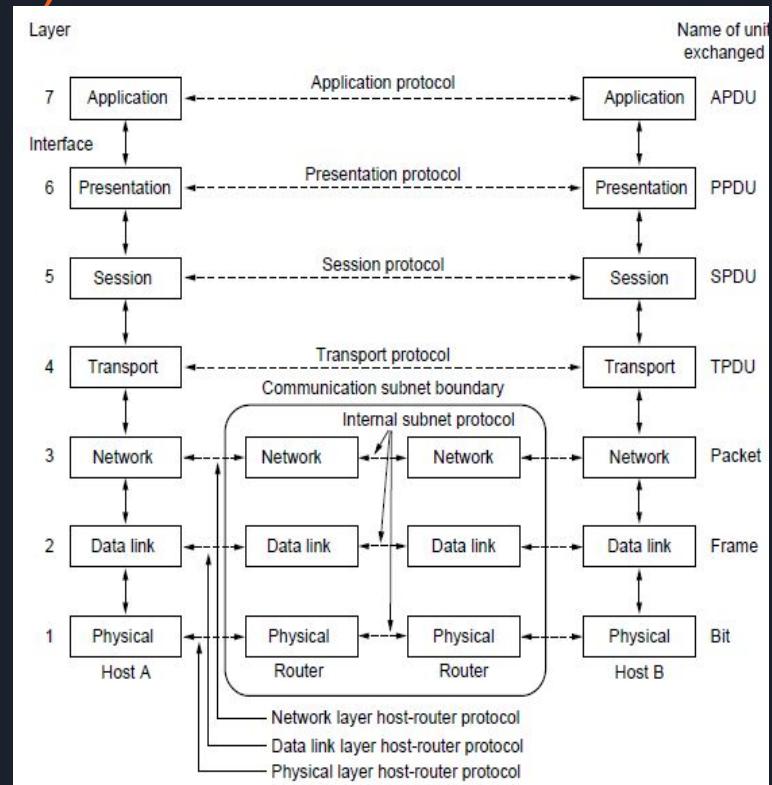


The relationship between a service and a protocol.

Reference Models

OSI (Open Systems Interconnect)

- Developed by the International Standards Organization (ISO)
- OSI means Open Systems Interconnection
- It is rarely used today, while it is actually quite general and still valid, and the features discussed at each layer are still very important.
- OSI model itself is not a network architecture because it does not specify the exact services and protocols to be used in each layer.





Reference Models

OSI (Open Systems Interconnect)

❖ The Physical Layer

- The physical layer is concerned with transmitting raw bits over a communication channel.
- Responsible for electrical signals, light signal, radio signals etc
- Hardware layer of OSI MODEL
- Devices like repeater, hub, cables, ethernet work on this layer and protocols like RS232, ATM, FDDI, Ethernet work on this layer

❖ The Data Link Layer

- The main task of the data link layer is to transform a raw transmission facility into a line that appears free of undetected transmission errors to the network layer.
- Having the sender break up the input data into data frames and transmit the frames sequentially
- It is divided into MAC Layer and LLC Layer.(The MAC sublayer controls how a computer on the network gains access to the data and permission to transmit it.The LLC layer controls frame synchronization, flow control and error checking.)
- Switch works at this level. Protocols are PPP and HDLC.



Reference Models

OSI (Open Systems Interconnect)

❖ The Network Layer

- The network layer controls the operation of the subnet (routing).
- Congestion control, QOS (quality of service)
- Creates logical paths between two hosts across the world wide web called as virtual circuits
- Internetworking, error handling, congestion control and packet sequencing work at this layer. Eg : IGMP , ICMP ,IPX
- Router works at this level.

❖ The transport Layer

- The basic function of the transport layer is to accept data from above, split it up into smaller units if need be, pass these to the network layer, and ensure that the pieces all arrive correctly at the other end.
- The transport layer is a true end-to-end layer, all the way from the source to the destination.
- Responsible for complete data transfer.
- Eg: UDP , RDP ,TCP ,SPX.



Reference Models

OSI (Open Systems Interconnect)

❖ The Session Layer

- Responsible for establishment, management and termination of connections between applications.
- The session layer sets up, coordinates, and terminates conversations, exchanges, and dialogues between the applications at each end. It deals with session and connection coordination.
- Protocols like NFS, NetBios names, RPC, SQL work at this layer.
- It deals with session and connection coordination.

❖ The presentation layer

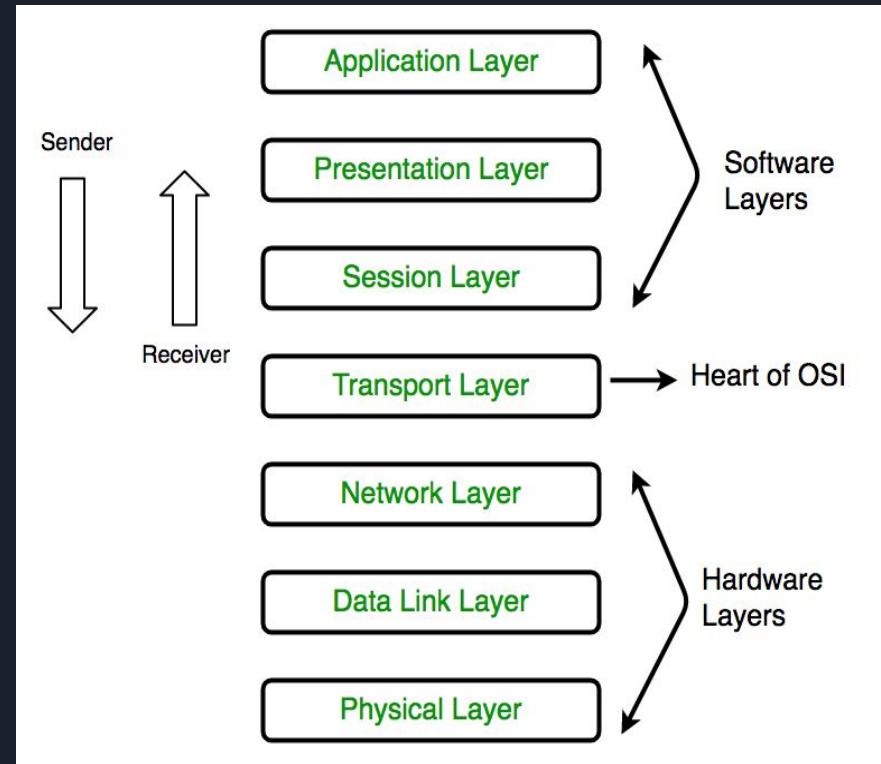
- Responsible for data representation on your screen
- Encryption and decryption of the data
- Data semantics and syntax
- Layer 6 Presentation examples include encryption, ASCII, EBCDIC, TIFF, GIF, PICT, JPEG, MPEG, MIDI.

Reference Models

OSI (Open Systems Interconnect)

❖ The Application Layer

- Application layer supports application, apps, and end-user processes.
- Quality of service
- This layer is responsible for application services for file transfers, e-mail, and other network software services.
- Protocols like Telnet, FTP, HTTP work on this layer.
- This Layer is also called Desktop layer.





Reference Models

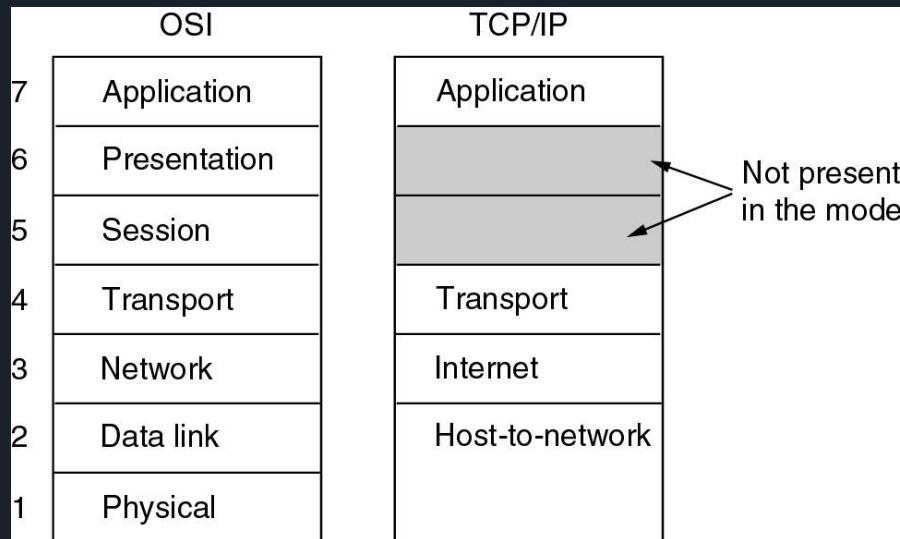
OSI (Open Systems Interconnect)

Critique of the OSI Model and Protocols :

- ❖ Bad timing.
- ❖ Bad technology.
- ❖ Bad implementations.
- ❖ Bad politics

Reference Models

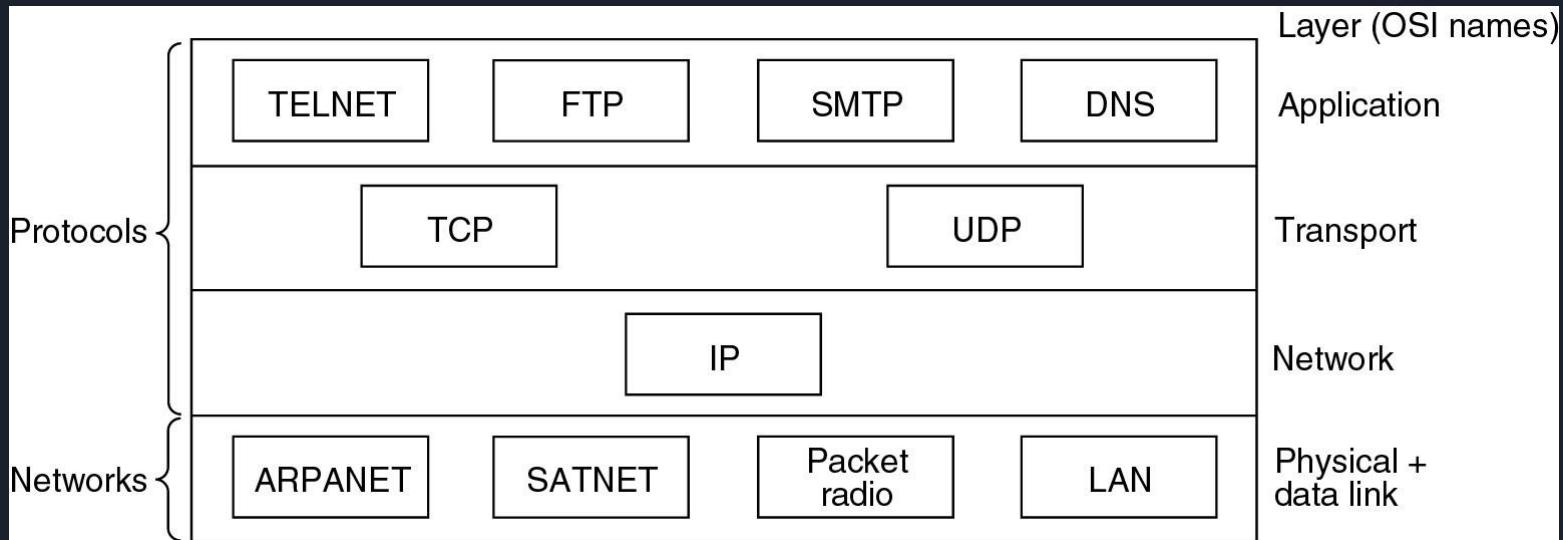
TCP/IP (Transmission Control Protocol / Internet protocol)



Reference Models

TCP/IP (Transmission Control Protocol / Internet protocol)

- ❖ Protocols and networks in the TCP/IP model initially





Reference Models

TCP/IP (Transmission Control Protocol / Internet protocol)

Link Layer

At physical level

- At physical level data consists of stream of bits (bit level synchronization)
- Line configuration (point 2 point / point 2 multipoint)
- Transmission mode - simplex / half - duplex / full duplex
- Topology - star, ring, tree, bus etc

At Logical level

- At logical level data consists of frames.
- Synchronization of frames.
- Flow control, error detection and correction of frames at LLC sublayer.
- Physical addressing, access control at MAC sub layer.



Reference Models

TCP/IP (Transmission Control Protocol / Internet protocol)

Internet Layer

- Packages data received from upper layer into datagram by adding IP header to it, which consist of source and destination address that is used to forward the datagram between nodes and across networks.
- Performs addressing, routing and other functions using following protocols:
 - Internet Protocol (IP) - is a connectionless protocol used for addressing host and routing data-gram's from source to destination host across one or more types of networks.
 - Address Resolution Protocol (ARP) - used to resolve host hardware address or MAC address when only its IP address of internet layer is known.
 - Internet Control Message Protocol (ICMP) - used to report error messages to host machine like destination host unreachable, time exceeded, invalid parameter etc.
 - Internet Group Management Protocol (IGMP) - used by host and routers to establish and manage IP multicast groups
 - Protocols used: IP (IPv4,IPv6), ICMP, ICMPv6, IGMP, ARP, RARP, DHCP etc.



Reference Models

TCP/IP (Transmission Control Protocol / Internet protocol)

Transport Layer

- Responsible for host-to-host message delivery along with flow control, congestion control and error recovery.
- The core protocols of transport layer are Transmission Control Protocol (TCP) and User Datagram Protocol (UDP).
- Either one of them protocol is selected, depending upon the type of communication application software needed.
- In TCP packets are called as segments, while in UDP it is called as datagrams.
- TCP - is a connection-oriented, reliable and stateful protocol.
- UDP - is a connectionless, unreliable and stateless datagram protocol.
- New protocols Datagram Congestion Control Protocol (DCCP), Stream Control Transmission Protocol (SCTP) etc are used.
- Network devices used : Router, Gateway.



Reference Models

TCP/IP (Transmission Control Protocol / Internet protocol)

Application Layer

- Some common application layer protocols used are HTTP , SMTP, TELNET , FTP , DNS , DHCP etc.
- Following are the functions of application layer :
 - Formatting data - Text format may be in ASCII / EBCDIC
 - Representing data - Defines how data is to be represented to the end user.
 - Process-to-process communication - E.g. client and server application software.
 - Creating sessions - E.g. a web browser with one or more tabs, then its application layer responsibility to create and manage separate sessions for each tab.
- Provides services like browsing, e-mail, file transfer, chatting, directory etc.
- Network application software's uses specific protocol and socket or port number for communication like web browser - http protocol - 80 port, file transfer - ftp protocol - 21 port, Logging on remote computer - telnet protocol - 23 port etc.



Reference Models

TCP/IP (Transmission Control Protocol / Internet protocol)

Problems:

- Service, interface, and protocol not distinguished
- Not a general model
- Host-to-network “layer” not really a layer
- No mention of physical and data link layers
- Minor protocols deeply entrenched, hard to replace



Network Standardisation

- ❑ Who's Who in telecommunications
- ❑ Who's Who in international standards
- ❑ Who's Who in internet standards

IEEE 802 Standards

Number	Topic
802.1	Overview and architecture of LANs
802.2 ↓	Logical link control
802.3 *	Ethernet
802.4 ↓	Token bus (was briefly used in manufacturing plants)
802.5	Token ring (IBM's entry into the LAN world)
802.6 ↓	Dual queue dual bus (early metropolitan area network)
802.7 ↓	Technical advisory group on broadband technologies
802.8 †	Technical advisory group on fiber optic technologies
802.9 ↓	Isochronous LANs (for real-time applications)
802.10 ↓	Virtual LANs and security
802.11 *	Wireless LANs
802.12 ↓	Demand priority (Hewlett-Packard's AnyLAN)
802.13	Unlucky number. Nobody wanted it
802.14 ↓	Cable modems (defunct: an industry consortium got there first)
802.15 *	Personal area networks (Bluetooth)
802.16 *	Broadband wireless
802.17	Resilient packet ring

Metric Units

The principal metric prefixes.

MODULE 2

OVERVIEW

- Datalink layer design issues
- flow control ad ARQ techniques.
- Data link protocols-HDLC
- Data link layer in internet.
- MAC sublayer

DATALINK LAYER DESIGN ISSUES

- □ Services provided to the Network Layer
- □ Framing
- □ Error Control
- □ Flow Control

SERVICES PROVIDED TO THE NETWORK LAYER

- Unacknowledged connectionless service
 - No acks, no connection
 - Error recovery up to higher layers
 - For low error-rate links or voice traffic
- Acknowledged connectionless service
 - Acks improve reliability
 - For unreliable channels. E.g.: Wireless System

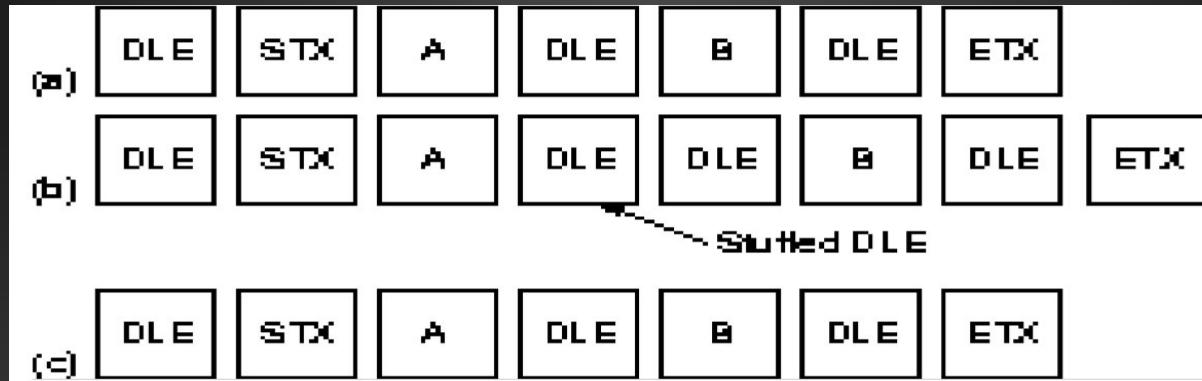
- Acknowledged connection-oriented service
- Equivalent of reliable bit-stream
- Connection establishment
- Packets Delivered In-Order
- Connection Release
- Inter-Router Traffic

FRAMING

- Framing = How to break a bit-stream into frames
- Need for framing: Error Detection/Control work on chunks and not on bit streams of data
- Framing methods:
 - Timing : risky. No network guarantees.
 - Character count: may be garbled by errors
 - Character stuffing: Delimit frame with special characters
 - Bit stuffing: delimit frame with bit pattern
 - Physical layer coding violations

CHARACTER STUFFING

- ❑ Delimit with DLE STX or DLE ETX character flags
- ❑ Insert 'DLE' before accidental 'DLE' in data
- ❑ Remove stuffed character at destination.



BIT STUFFING

- Delimit with special bit pattern (bit flags)
 - Stuff bits if pattern appears in data
 - Remove stuffed bits at destination

(a) 0110111111111111110010

(b) 01101111011111011111010010

Stuffed bits

ERROR CONTROL

- Error Control = Deliver frames without error, in the proper order to network layer
- Error control Mechanisms:
 - Ack/Nack: Provide sender some feedback about other end
 - Time-out: for the case when entire packet or ack is lost
 - Sequence numbers: to distinguish retransmissions from originals

FLOW CONTROL

- Flow Control = Sender does not flood the receiver, but maximizes throughput
- Sender throttled until receiver grants permission

ELEMENTARY DATA LINK PROTOCOLS

Unrestricted Simplex Protocol

Framing only

No error or flow control

No capacity barriers

❖ The actions by the link layer of sender:

fetch the packet from the network layer.

construct the frame.

send the frame.

❖ Receiver side:

Always wait for an event to occur.

Wait for an event frame arrival.

Call the procedure “from_the_physical_layer”.

Process the frame according to the header and trailer.

Pass the frame by calling the procedure “to_network_layer”.

Simplex Stop-and-Wait

- Send one packet
- Wait for Ack before proceeding

Limitations:

for the protocol realization we need a half duplex channel.

sender enter the while loop on receiving an acknowledgement.

Receiver enter the while loop after sending the acknowledgement.

Simplex Protocol for a Noisy Channel

- Automatic Repeat request (ARQ) protocols

The protocols in which the sender wait for a positive acknowledgement before advancing to the next data item are called ARQ(Automatic Repeat Request).

- Positive Ack
- 1-bit sequence number in frames (not in acks) ;sequence number are either 0 or 1.
- Timeout to detect lost frames/acks
- Retransmission
- Can fail under early timeout conditions

Full Duplex Communication

- Piggybacking of ack

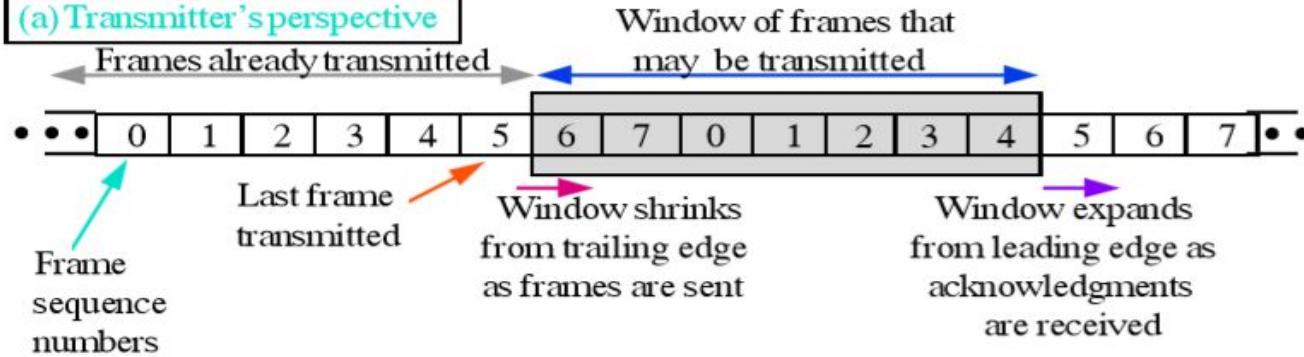
SLIDING WINDOW PROTOCOLS

- Window = Set of sequence numbers to send/receive
- Sender window
 - Sender window increases when ack received
 - Packets in sender window must be buffered at source
 - Sender window may grow in some protocol

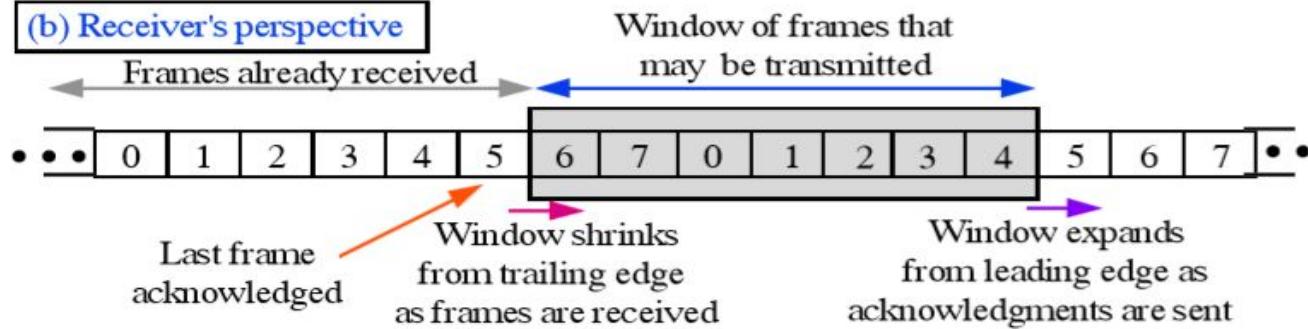
- Receiver window
- Packets outside window discarded
- Window advances when sequence number = low edge of window received
- Receiver window always constant
- Sender transmits W frames before blocking (pipelining)

Sliding Window

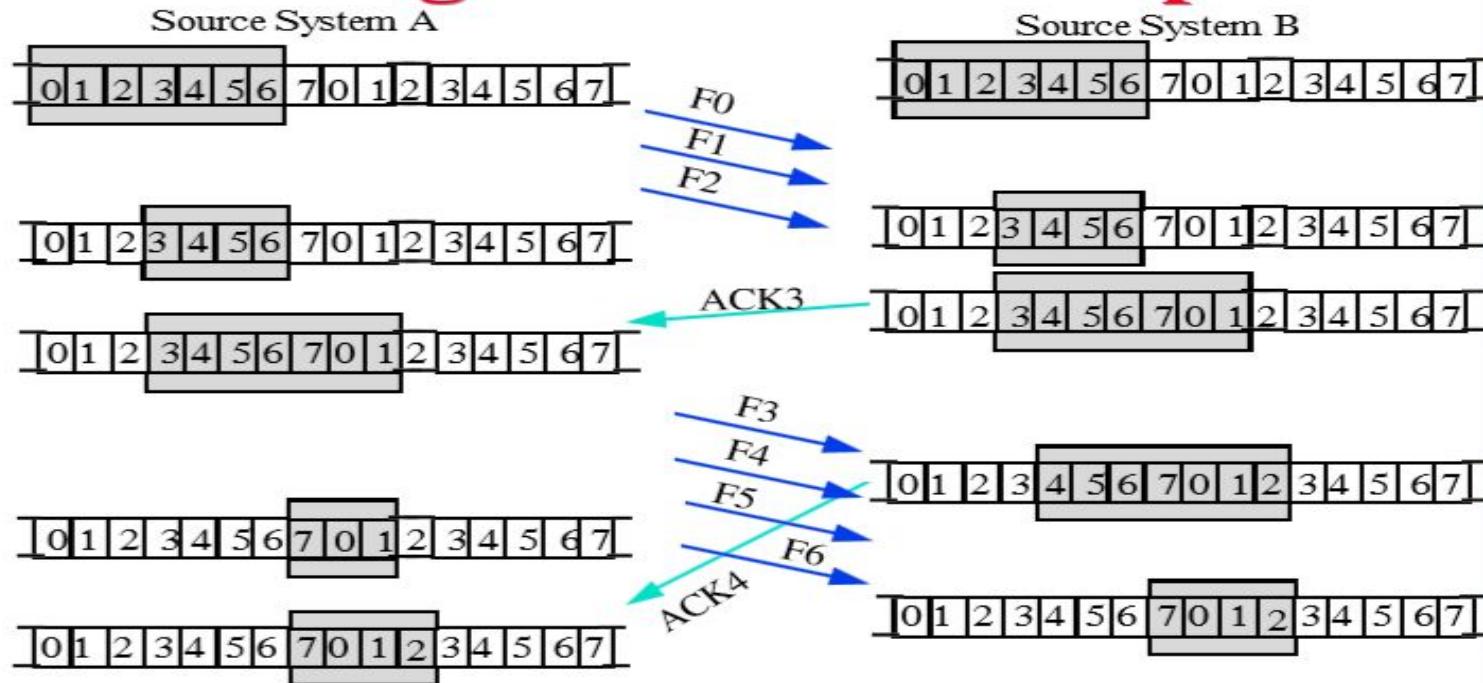
(a) Transmitter's perspective



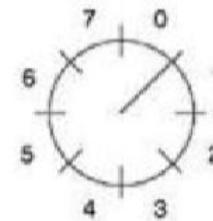
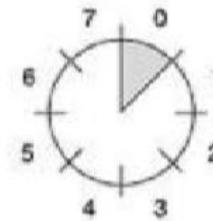
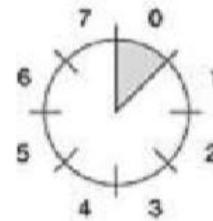
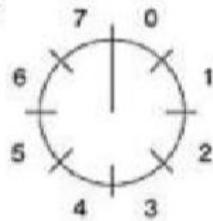
(b) Receiver's perspective



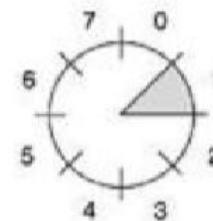
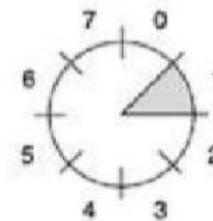
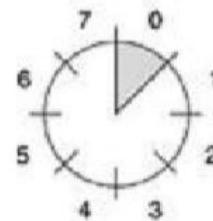
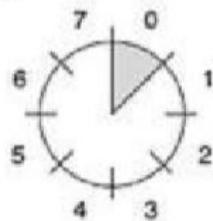
Sliding Window: Example



Sender



Receiver



(a)

(b)

(c)

(d)

Fig.8.A sliding window of size 1, with a 3-bit sequence number (a) Initially (b) After the first frame has been sent (c) After the first frame has been received (d) After the first acknowledgement has been received.

GO BACK N

- Sliding Window + Retransmit all packets starting from earliest unpacked packet
- Acks are cumulative
- Ack the latest packet (N) if all packets (< N) are received.
- Can waste a lot of bandwidth if error rate high.
- If m id=s the number of bits sequence numbers are modulo 2^m
- Maximum seq.no= $(2^m)-1$.
- Send window can slide one or more slots when a valid acknowledgement received.
- In receiver side only one movement is permitted
- Buffering mechanism is used.
- Sometimes the frames can be discarded . This is one disadvantage.

SELECTIVE REPEAT

- If data is received acknowledgement is send, otherwise a Nack is send .
- Out-of-order buffering required at destination
- Acks not cumulative.
- Non-sequential arrival of packets
 - ⇒ ensure no overlap with old window when receiver advances window
 - ⇒ Maximum Window = (Sequence Number Space)/2.
- Receiver window size= $(2^m)-1$.
- In receiver side if Acknowledgement is send , window can move.
- Safe window size is half of 2^m .

- Enhancement to piggybacking
- Auxiliary Ack timer
- Set timer when packet to be acknowledged is received
- If no reverse data, generate Ack when timer goes off
- Send negative ack (Nack) when out-of-order frame received
- Source retransmits Nacknowledged frame
- Naks not cumulative
- Do not send duplicate Nack

HDLC PROTOCOL

- ☐ High-level Data Link Control
- ☐ Bit-oriented, bit-stuffing for transparency
- ☐ Control field for sequence numbers, acks etc
- ☐ Data field arbitrarily long (practical limits: checksum efficiency)
- ☐ CRC using CRC-CCITT generator polynomial.

Frame types:

- Information, Supervisory, Unnumbered frames
- Sliding window with 3-bit sequence number, piggybacked acks
- Connection Management through supervisory frame

- Information
- Supervisory
- unnumbered

Bits	1	3	1	3
(a)	0	Seq	P/F	Next
(b)	1	0	Type	P/F
(c)	1	1	Type	P/F
				Modifier

- **Information Frames:**
- User data
- Piggybacked Acknowledgments: Next frame expected
- Poll/Final = Command/Response
- **Supervisory Frames:**
- Flow and error control
- Go back N and Selective Reject
- Final ⇒ No more data to send
- **Unnumbered Frames:**
- Control
- Mode setting commands and responses
- Information transfer commands and responses
- Recovery commands and responses
- Miscellaneous commands and responses

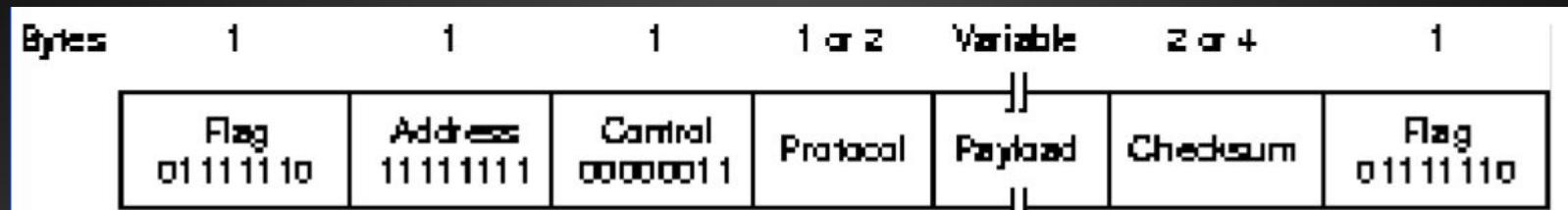
DATA LINK LAYER IN THE INTERNET

- ☐ Point-to-Point link scenarios
- ☐ Router-router leased line (PPP)
- ☐ dial-up host-router connection (SLIP, future PPP)
- ☐ SLIP: Not an approved Internet Standard ⇒ incompatibilities
- ☐ PPP: Official Internet Standard.

Point to Point Protocols

- Reference: RFCs 1661, 1662 1663
- Frame method to unambiguously delineate frames and for error control
- Link Control Protocol(LCP): link management
- Network Control Protocol (NCP): Supports multiple network layers

PPP frame format:



Connection setup:

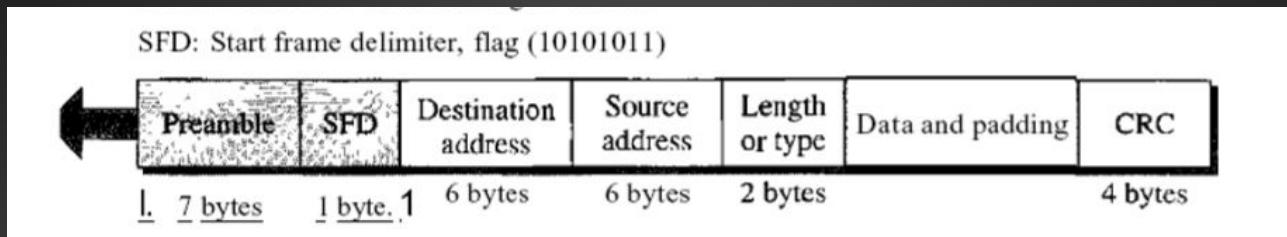
- Home PC Modem calls Internet Provider's router: sets up physical link
- PC sends series of LCP packets as PPP payload + Select PPP (data link) parameters + Authenticate
- PC sends series of NCP packets as PPP payload + Select network parameters + E.g., Get dynamic IP address

Connection Teardown:

- NCP first followed by LCP and physical connection

MAC SUBLAYER

- It frames data received from the upper layer and passes them to the physical layer
- Ethernet does not provide any mechanism for acknowledging received frames, making it what is known as an unreliable medium.
- Acknowledgments must be implemented at the higher layers. The format of the MAC frame is



- D Preamble. The first field of the 802.3 frame contains 7 bytes (56 bits) of alternating 0s and 1s that alerts the receiving system to the coming frame and enables it to synchronize its input timing. The pattern provides only an alert and a timing pulse. The 56-bit pattern allows the stations to miss some bits at the beginning of the frame. The preamble is actually added at the physical layer and is not (formally) part of the frame.
- D Start frame delimiter (SFD). The second field (1 byte: 10101011) signals the beginning of the frame. The SFD warns the station or stations that this is the last chance for synchronization. The last 2 bits is 11 and alerts the receiver that the next field is the destination address.
- Destination address (DA). The DA field is 6 bytes and contains the physical address of the destination station or stations to receive the packet.
- Source address (SA). The SA field is also 6 bytes and contains the physical address of the sender of the packet. We will discuss addressing shortly.
- Length or type. This field is defined as a type field or length field. The original Ethernet used this field as the type field to define the upper-layer protocol using the MAC frame. The IEEE standard used it as the length field to define the number of bytes in the data field. Both uses are common today.
- Data. This field carries data encapsulated from the upper-layer protocols. It is a minimum of 46 and a maximum of 1500 bytes.
- CRC. The last field contains error detection information, in this case a CRC-32

IEEE STANDARDS

IEEE STANDARDS

In 1985, the Computer Society of the IEEE started a project, called Project 802, to set standards to enable intercommunication among equipment from a variety of manufacturers. Project 802 is a way of specifying functions of the physical layer and the data link layer of major LAN protocols.

Topics discussed in this section:

802.3 – Ethernet

802.5 – Token Ring

802.4 – Token Bus

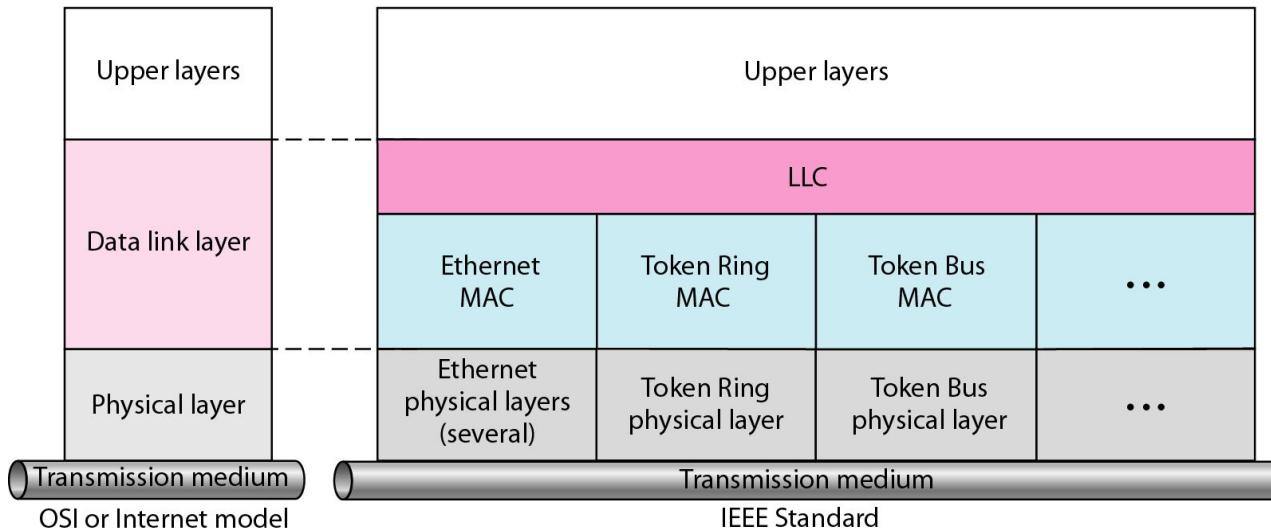
802.11 – BSS

802.15 – Bluetooth

Figure 13.1 IEEE standard for LANs

LLC: Logical link control

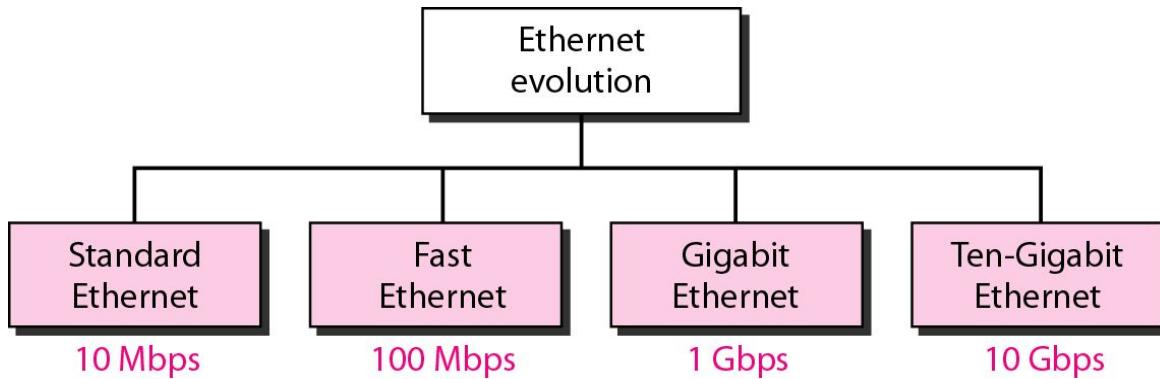
MAC: Media access control



STANDARD ETHERNET

The original Ethernet was created in 1976 at Xerox's Palo Alto Research Center (PARC). Since then, it has gone through four generations. We briefly discuss the **Standard (or traditional) Ethernet** in this section.

Figure 13.3 *Ethernet evolution through four generations*



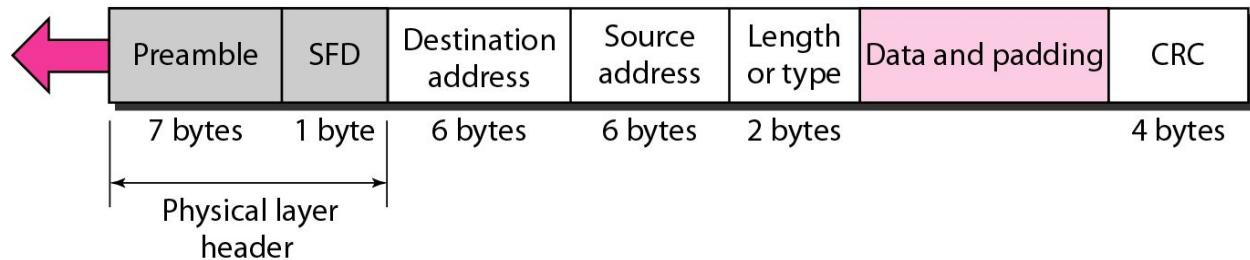
Different Classes of Standard Ethernet

- 10 base 5 – uses thick coaxial cable for a distance of 500m; supports 100 nodes
- 10 base 2 – uses thin coaxial cable for a distance of 185cm; supports 30 nodes
- 10 base T – uses twisted pair cable for a distance of 100m; supports 1024 nodes
- 10 base F – uses optical fibre cable for a distance of 2 km; supports 1024 nodes

Figure 13.4 802.3 MAC frame

Preamble: 56 bits of alternating 1s and 0s.

SFD: Start frame delimiter, flag (10101011)



- ## Preamble

The preamble is responsible for providing the synchronization between the sending and receiving device. It is a series of 56 bits (7 bytes) of alternating 1s and 0s found at the beginning of the frame

- ## Start of Frame Delimiter

The start frame delimiter follows the preamble. As its name implies, it indicates the start of the data frame. The start frame delimiter is 1 byte in length—made up of the following 8-bit sequence—10101011.

- ## Address Fields

Each of the address fields—the destination address and the source address—can be either 2 bytes or 6 bytes in length. If universal addressing is used, the addresses must be 6 bytes each. But if local addressing is used they may be either 2 or 6 bytes long. Both destination and source addresses must be of the same length for all devices on a given network.

- Length Count

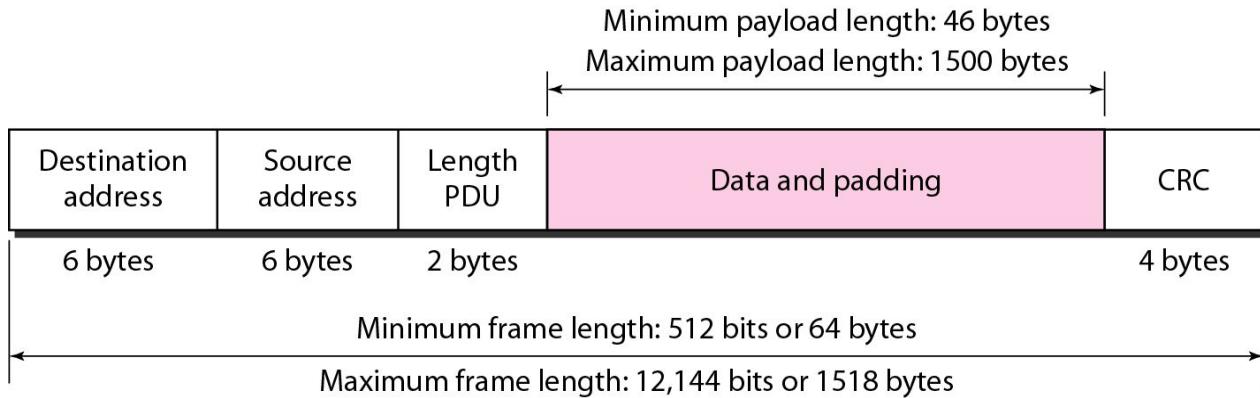
This is a 2-byte field indicating the length of the data field that follows. It is needed to determine the length of the data field in those cases when a pad field is used.
- Information Field

The information field contains the actual data packet to be transmitted. Its length is variable.
- Pad Field

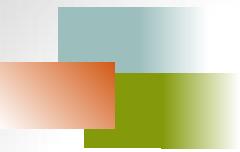
A pad field is used to ensure that the frame meets a minimum length requirement. A frame must contain a minimum number of bytes in order for stations to detect collisions accurately.
- Frame Check Sequence

The frame check field is used as an error-control mechanism. When the transmitting device assembles a frame, it performs a calculation on the bits in the frame. The algorithm used to perform this calculation always results in a 4-byte value. The sending device stores this value in the frame check sequence field.

Figure 13.5 Minimum and maximum lengths



Padding is used when enough data is not there to fill the frame.



Note

Frame length:

Minimum: 64 bytes (512 bits)

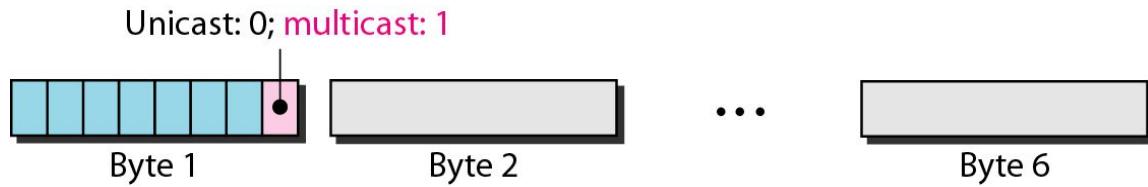
Maximum: 1518 bytes (12,144 bits)

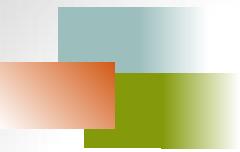
Figure 13.6 *Example of an Ethernet address in hexadecimal notation*

06 : 01 : 02 : 01 : 2C : 4B

6 bytes = 12 hex digits = 48 bits

Figure 13.7 *Unicast and multicast addresses*

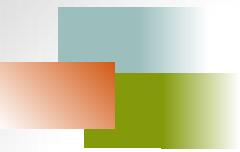




Note

The least significant bit of the first byte
defines the type of address.

If the bit is **0**, the address is unicast;
otherwise, it is multicast.



Note

The broadcast destination address is a special case of the multicast address in which all bits are 1s.

Example 13.1

Define the type of the following destination addresses:

- a. 4A:30:10:21:10:1A
- b. 47:20:1B:2E:08:EE
- c. FF:FF:FF:FF:FF:FF

Solution

To find the type of the address, we need to look at the second hexadecimal digit from the left. If it is even, the address is unicast. If it is odd, the address is multicast. If all digits are F's, the address is broadcast. Therefore, we have the following:

- a. This is a unicast address because A in binary is 1010.
- b. This is a multicast address because 7 in binary is 0111.
- c. This is a broadcast address because all digits are F's.

Example 13.2

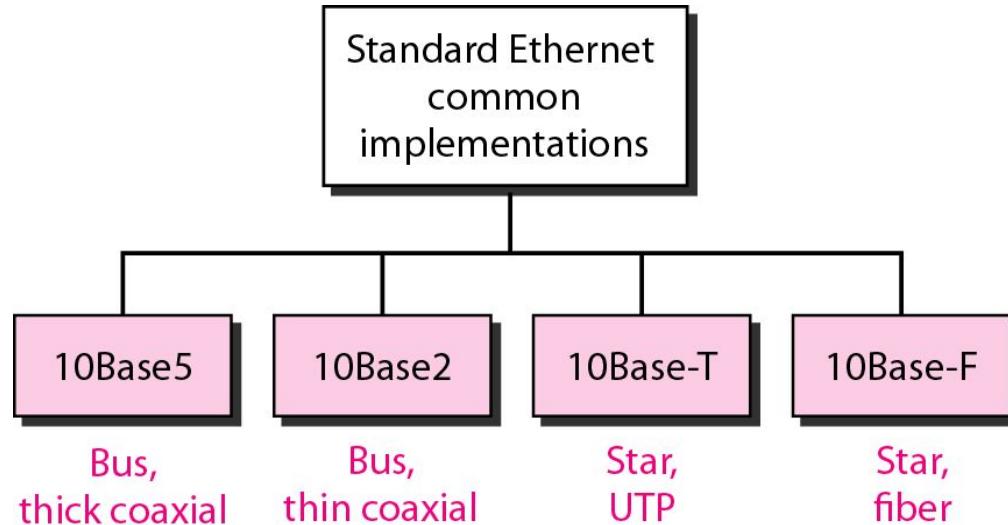
Show how the address **47:20:1B:2E:08:EE** is sent out on line.

Solution

The address is sent left-to-right, byte by byte; for each byte, it is sent right-to-left, bit by bit, as shown below:

← 11100010 00000100 11011000 01110100 00010000 01110111

Figure 13.8 *Categories of Standard Ethernet*



CHANGES IN THE STANDARD

The 10-Mbps Standard Ethernet has gone through several changes before moving to the higher data rates. These changes actually opened the road to the evolution of the Ethernet to become compatible with other high-data-rate LANs.

Topics discussed in this section:

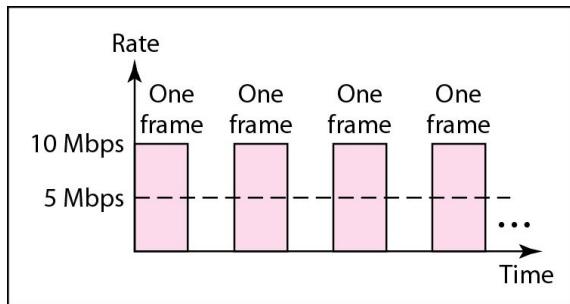
Bridged Ethernet

Switched Ethernet

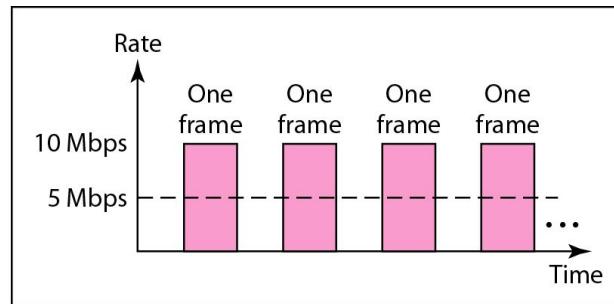
Bridged Ethernet

- Bridges are used to divide LANs
- Helps raise bandwidth
- Helps in forming separate collision domain
- Works in both physical and data link layers
- In physical layer – act as regenerator
- In MAC layer – route the packet based on MAC address
- Collisions are reduced

Figure 13.14 *Sharing bandwidth*

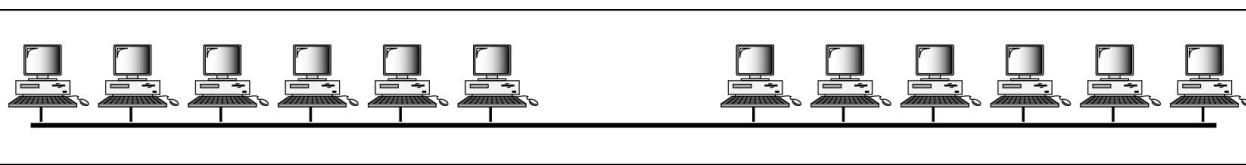


a. First station

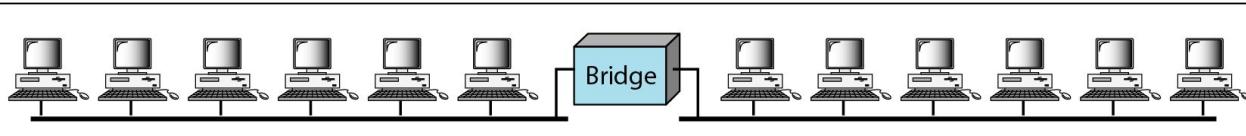


b. Second station

Figure 13.15 *A network with and without a bridge*



a. Without bridging



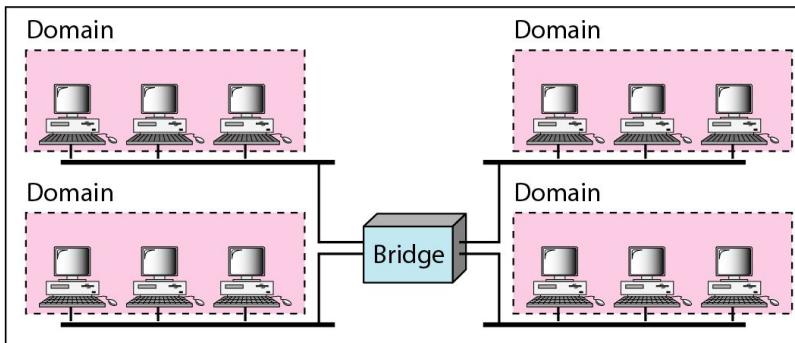
b. With bridging

Figure 13.16 Collision domains in an unbridged network and a bridged network

Domain



a. Without bridging



b. With bridging

Switched Ethernet

- An n-port bridge is called a 2-layer switch
- Each host act as a network
- Collisions are less as all hosts are directly connected to a port

Figure 13.17 *Switched Ethernet*

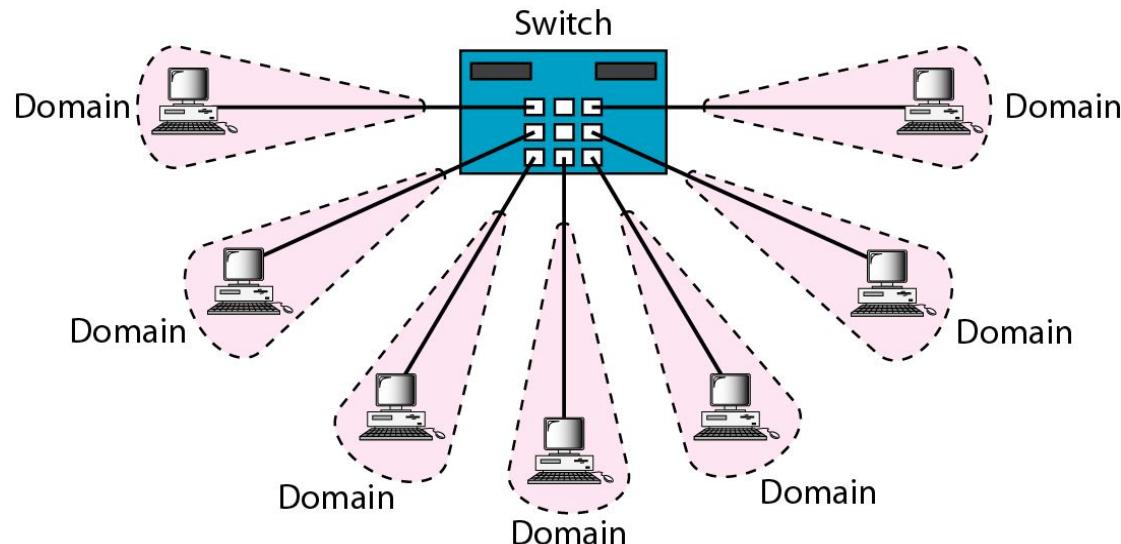
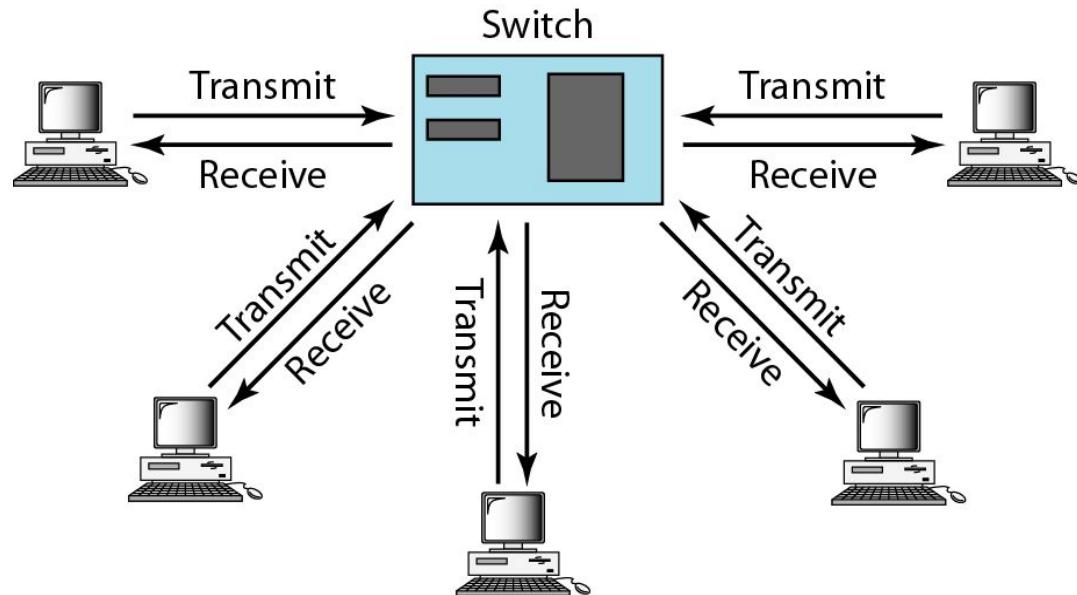


Figure 13.18 Full-duplex switched Ethernet



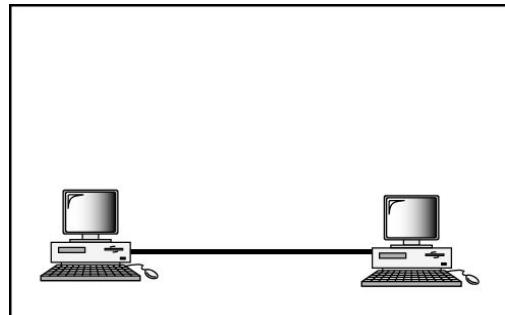
FAST ETHERNET

Fast Ethernet was designed to compete with LAN protocols such as FDDI or Fiber Channel. IEEE created Fast Ethernet under the name 802.3u. Fast Ethernet is backward-compatible with Standard Ethernet, but it can transmit data 10 times faster at a rate of 100 Mbps.

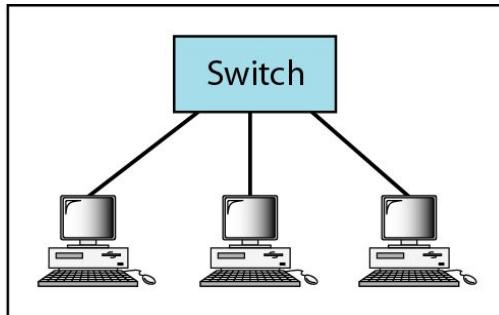
Features

- Compatible with standard Ethernet
- Uses 48 – bit MAC address
- Follows the same frame format of 802.3
- Also having minimum and maximum frame size of 802.3
- Uses only star topology and point-to-point topology
- Uses hub for half-duplex and switch for full-duplex
- Auto-negotiation:
 - Permits connection between incompatible devices
 - Permits a station to check hub capabilities
 - Allows a device to have multiple capabilities

Figure 13.19 *Fast Ethernet topology*

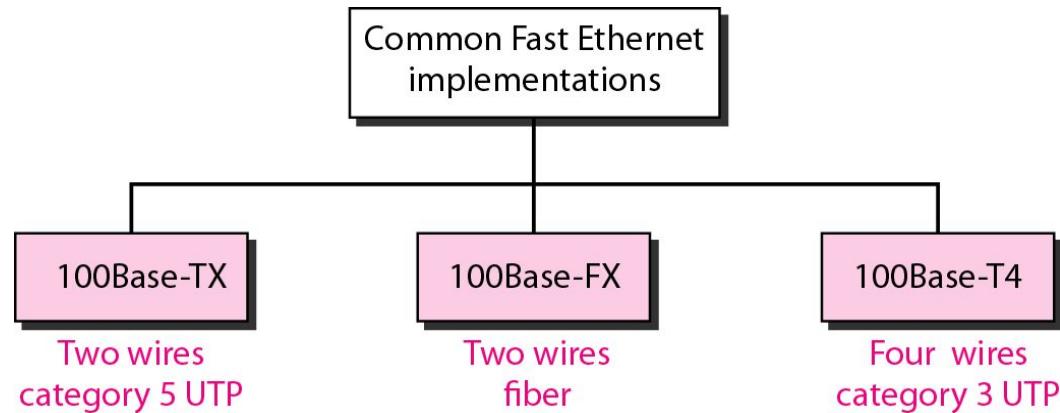


a. Point-to-point



b. Star

Figure 13.20 *Fast Ethernet implementations*



GIGABIT ETHERNET

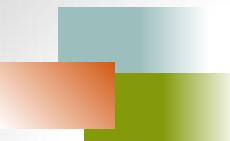
The need for an even higher data rate resulted in the design of the Gigabit Ethernet protocol (1000 Mbps).
The IEEE committee calls the standard 802.3z.

In half-duplex mode, provides three capabilities of services:

- Traditional – 512 bits
- Extended – 512 bytes
- Frame bursting

Frames are not identified separately in frame bursting.

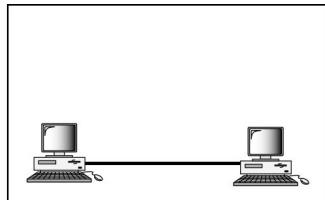
After a frame is sent unwanted bits are send till next frame is ready. This prevent spoofing.



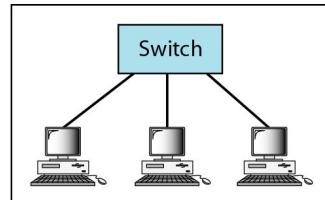
Note

In the full-duplex mode of Gigabit Ethernet, there is no collision; the maximum length of the cable is determined by the signal attenuation in the cable.

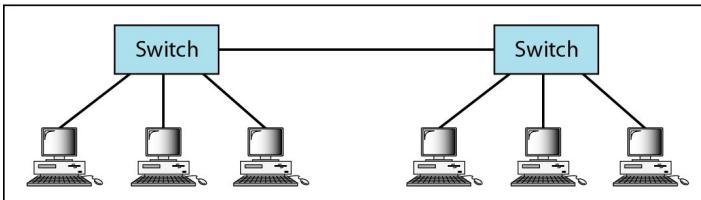
Figure 13.22 *Topologies of Gigabit Ethernet*



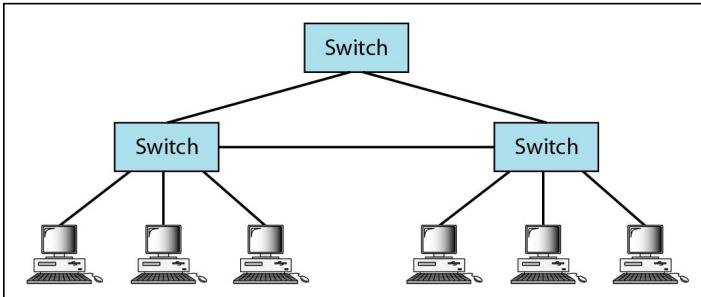
a. Point-to-point



b. Star

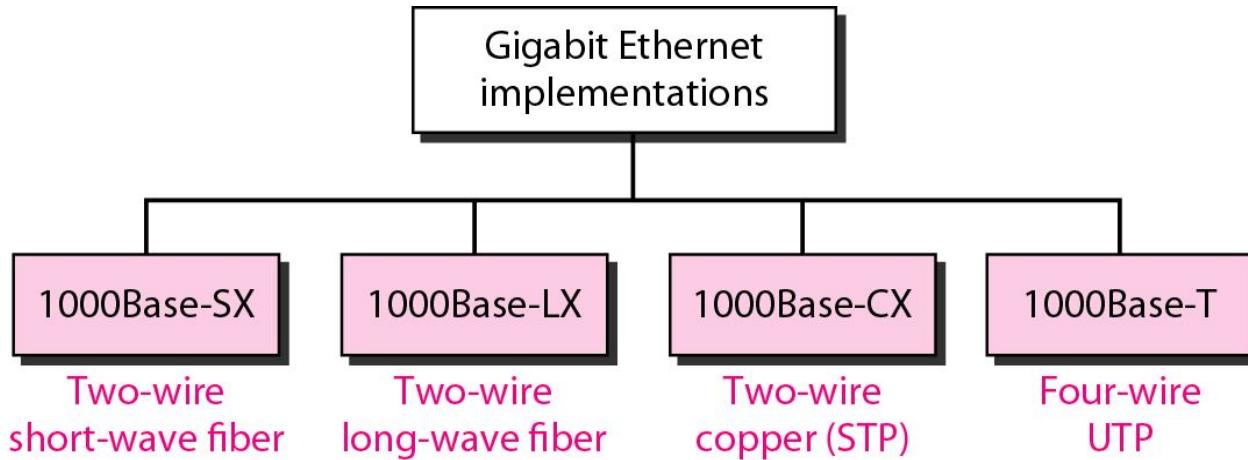


c. Two stars



d. Hierarchy of stars

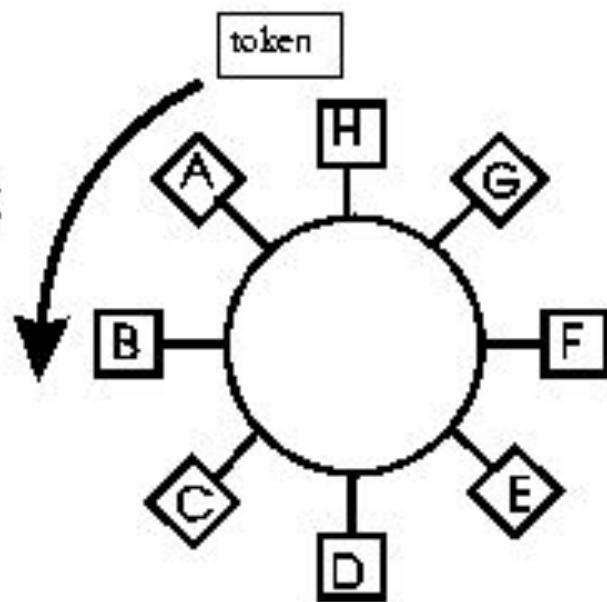
Figure 13.23 *Gigabit Ethernet implementations*



802.5 (TOKEN RING)

- There is a point to point link between stations that form a ring.
- Physical Layer Topology: *Ring*
 - Stations connected in a loop
 - Signals go in only one direction, station-to-station
- In a token ring a special bit format called a token circulated around all the stations.
- Only one station, the one possessing the token is allowed to transmit at any time.

Token
Circulates Ring



Token Ring Operation

- When a station wishes to transmit, it must wait for the token to pass by and seize the token.
 - One approach: change one bit in token which transforms it into a normal data frame and appends frame for transmission.
 - Second approach: station claims token by removing it from the ring.
- The data frame circles the ring and is removed by the transmitting station.
- Each station interrogates passing frame. If destined for station, it copies the frame into local buffer.
- As bits have propagated around the ring & they come back, they are removed from the ring by the sender.

802.5 Frame Format



- Starting delimiter(SF) and ending delimiter(EF) mark the beginning & ending of the frame.
- Access control(AC) consist of token bit, monitor bit, priority bit.
- Destination address(DA) & source address(SA) fields gives the address.
- Checksum(FCS) field is used to detect transmission errors.

- Frame control(FC) has a set of codes which is used to control the frame which is in transmission
- Frame status(FS)

A	C	xx	A	C	xx
---	---	----	---	---	----

A – address recognized bit

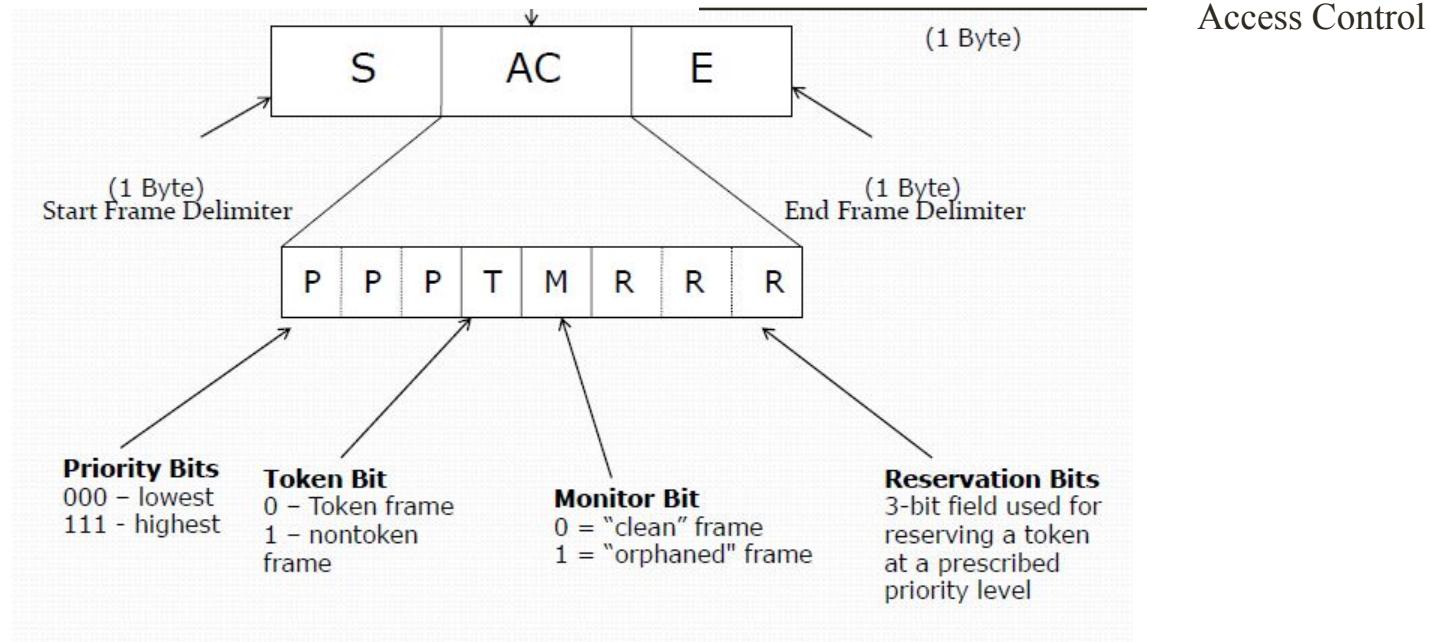
Xx – undefined

C – frame-copied bit

3 combinations:

- A=0 C=0 : Destination not present or not powered.
- A=1 C=0 : destination present but frame not accepted.
- A=1 C=1 : Destination present and frame copied.

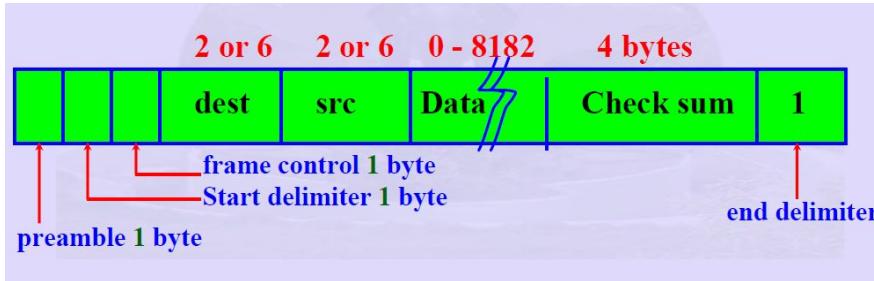
Token Frame Format



802.4(TOKEN BUS)

- Linear tree shaped cable on to which stations are attached.
- Each station knows the address of its left and right neighbors.
- A logical ring is formed.
- Token passing from higher to lower order station address
- Token acquired station transmits for certain amount of time
- Hand over token either at end of time or no frame to transmit
- Prioritise tokens

802.4 Frame Format



Frame control indicates the type of frame it holds. Frame can hold various values including:

- 0 – claim token; doing at the initialisation stage
- 1 – solicit successor one; used when you are adding a node to the ring
- 2 – solicit successor two; again add node to ring
- 3 – who follows; used in case of lost token
- 4 – for resolving contention in case of multiple station addition
- 8 – used to pass the token
- 12 – for setting the successor, mostly used when deletion of node from a ring

Token Holder

- A token holder will issue frame with frame control value(solicit successor one).
- The message is passed to the immediate successor of the token holder.
- If no response is received within the time period, new node will be added as successor of the token handler.
- If more nodes are requested, a contention process will be adding as the successor of the token holder.

802.11(Basic Service Set)

IEEE has defined the specifications for a wireless LAN, called IEEE 802.11, which covers the physical and data link layers.

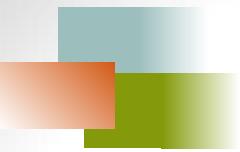
Topics discussed in this section:

DCF

Hidden Station Problem

Exposed Station Problem

PCF



Note

A BSS without an AP is called an **ad hoc** network;
a BSS with an AP is called an **infrastructure** network.

Figure 14.1 Basic service sets (BSSs)

BSS: Basic service set

AP: Access point

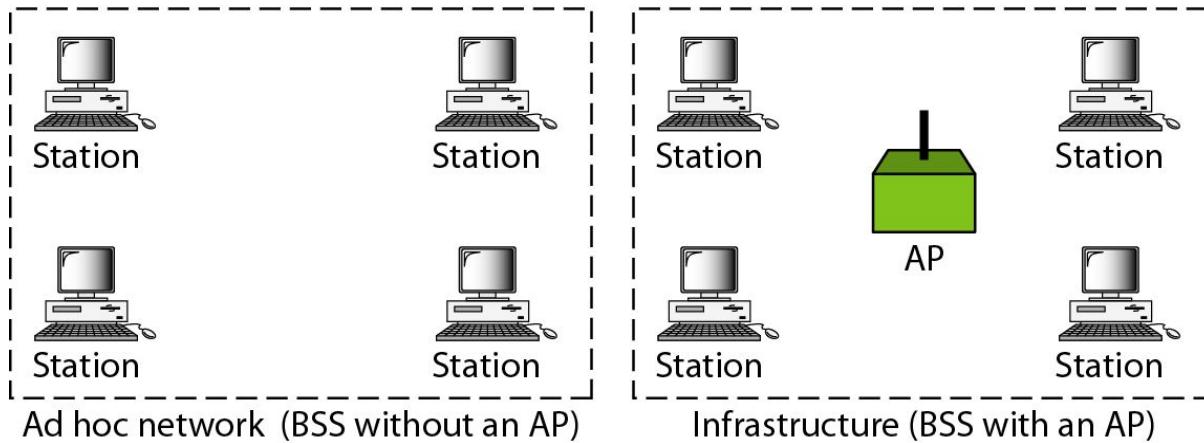
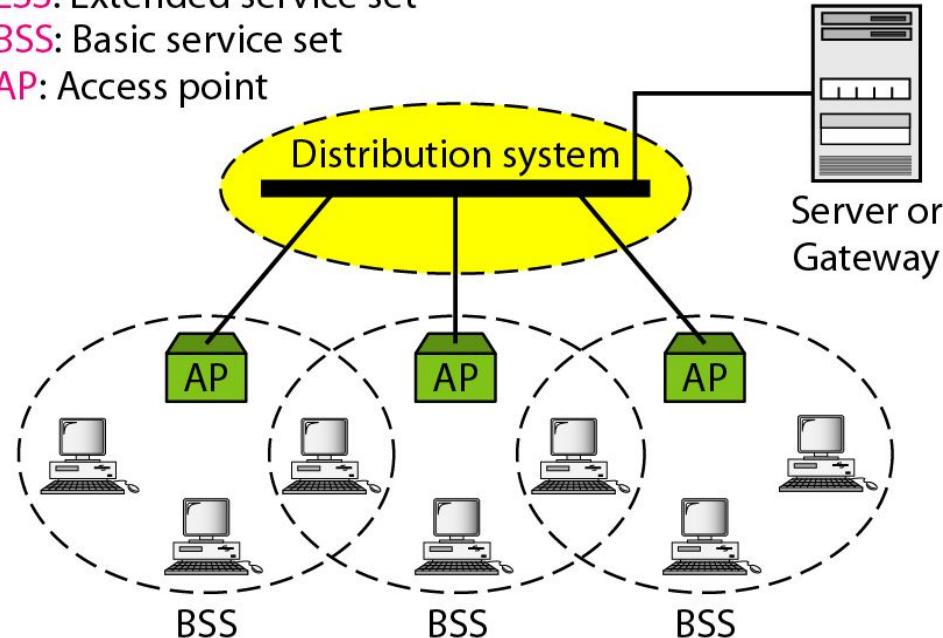


Figure 14.2 *Extended service sets (ESSs)*

ESS: Extended service set

BSS: Basic service set

AP: Access point





Note

Two access mechanisms:

Distributed Co-ordination Function(DCF)

Point Co-ordination Function(PCF)



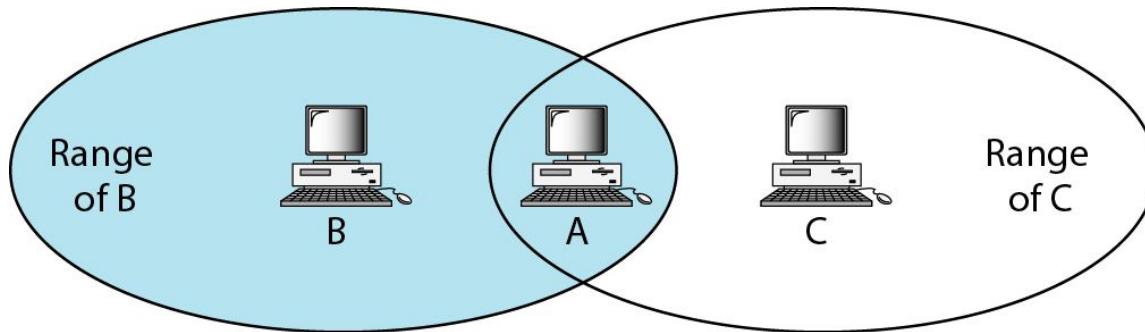
DCF

Note

The problems associated with wireless devices

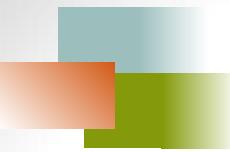
- Hidden station/terminal problem
- Exposed station/terminal problem

Hidden Station Problem



B and C are hidden from each other with respect to A.

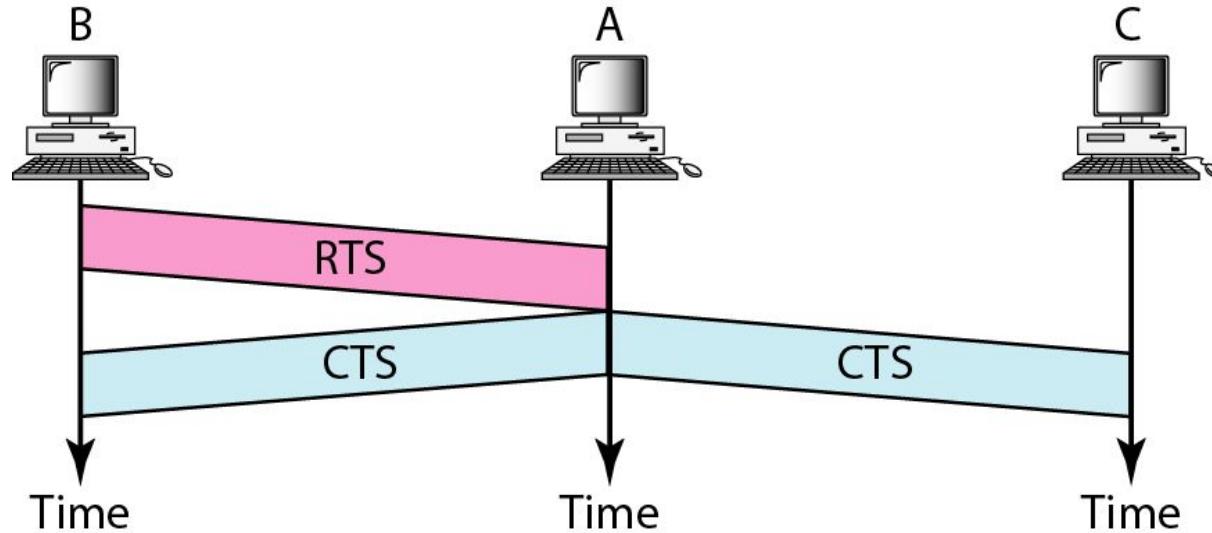
When both B and C send data to A, data collide. This can be removed using CTS and RTS.



Note

The CTS frame in CSMA/CA handshake can prevent collision from a hidden station.

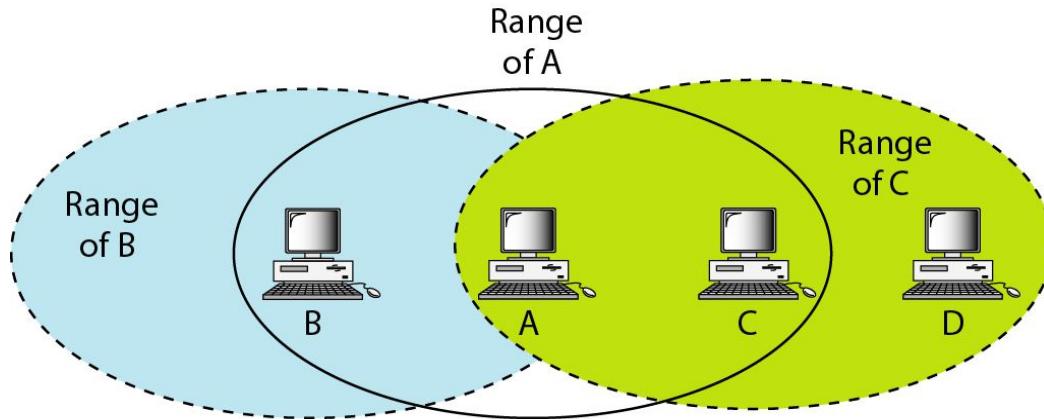
Figure 14.11 *Use of handshaking to prevent hidden station problem*



Preventing collision using CTS

- Suppose B wants to send data to A. Then B sends an RTS to A.
- If A is ready to receive data, it send CTS to all stations in its range. So C now knows that A is going to receive a message from some other station.
- Now B can send data to A without any collision.
- Now if C wants to send, it will first send RTS and only if CTS is received after that, it will send data to A.

Exposed Station Problem



C is exposed to transmission from A to B.

Figure 14.13 *Use of handshaking in exposed station problem*

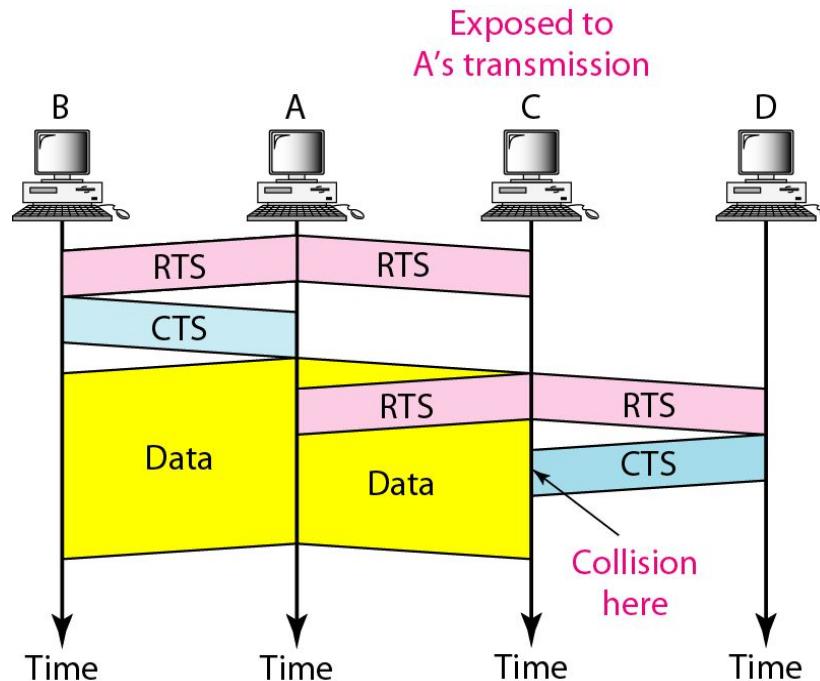


Figure 14.4 CSMA/CA flowchart

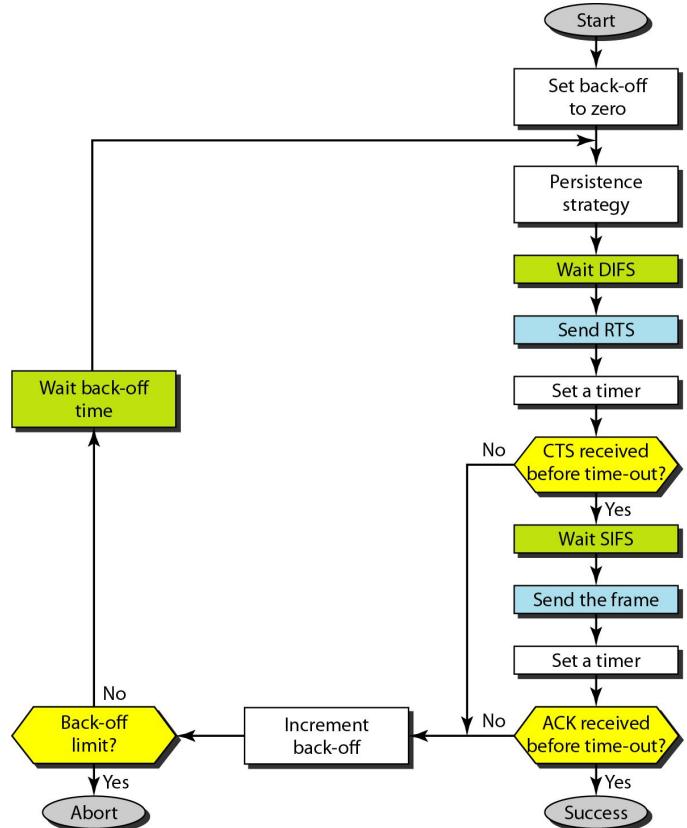
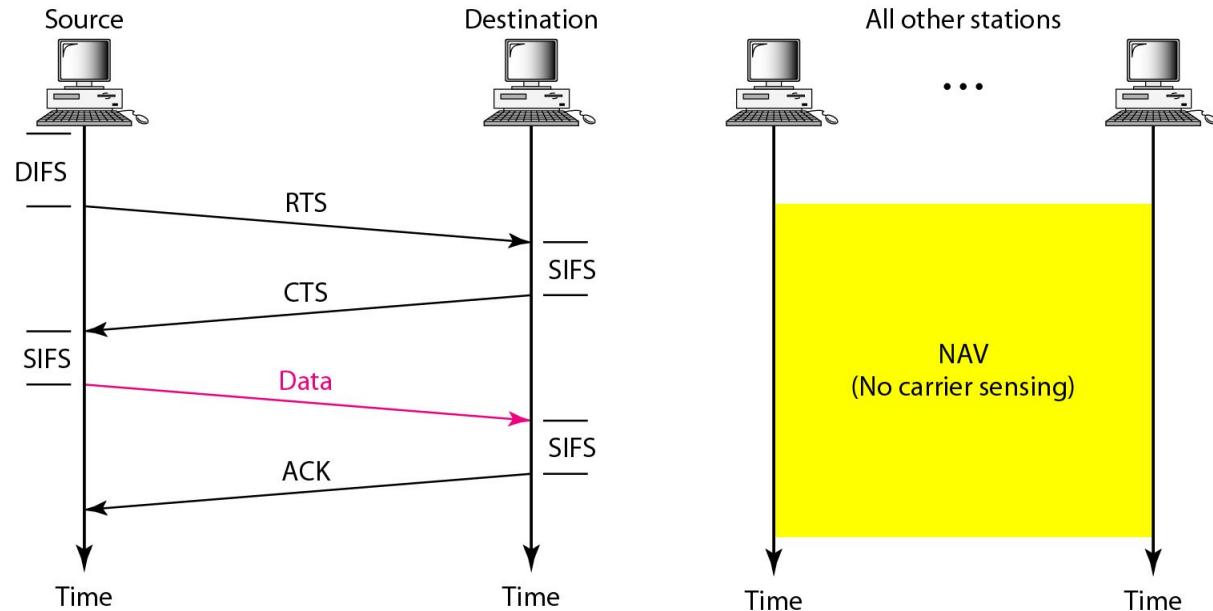
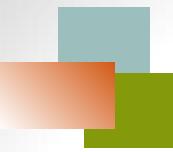


Figure 14.5 CSMA/CA and NAV





PCF

Note

PCF is having two time spans

- PIFS
- SIFS

DIFS > PIFS > SIFS

Figure 14.6 Example of repetition interval

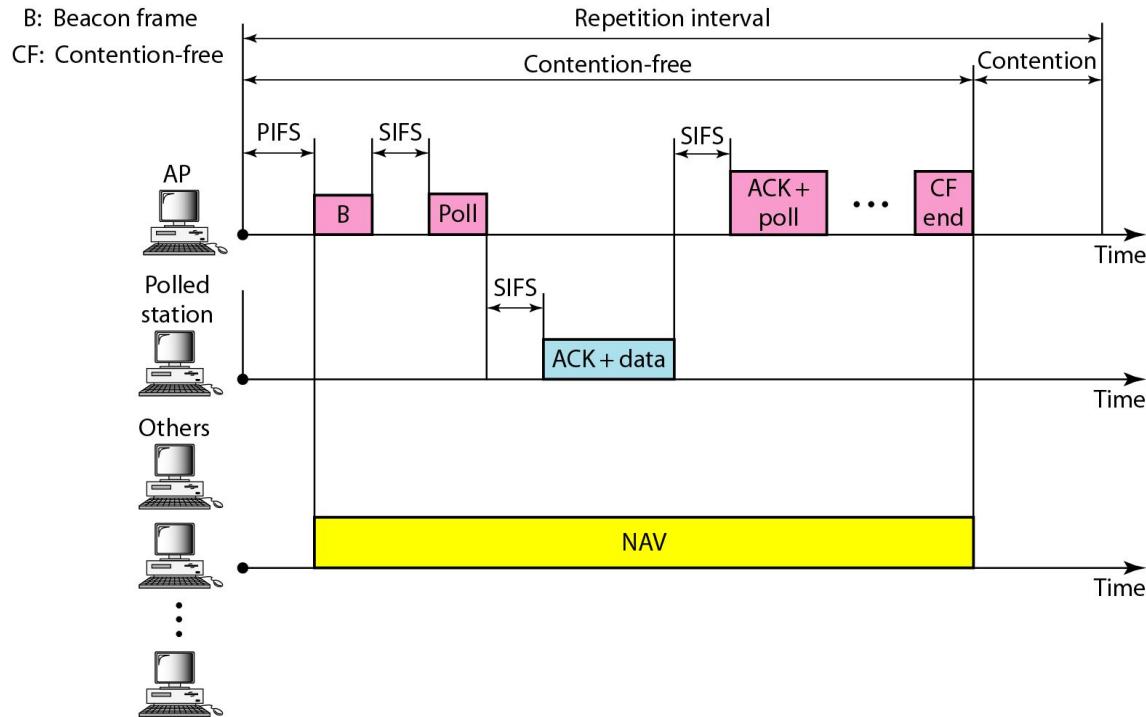
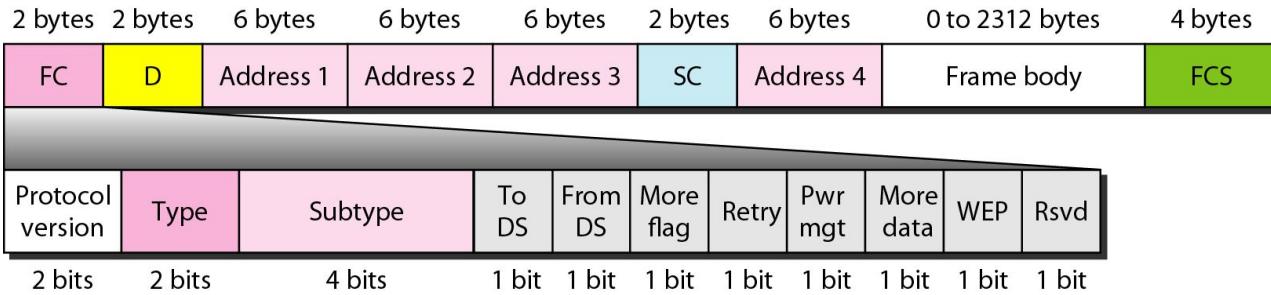


Figure 14.7 Frame format



- Frame control(FC) – tells the type of control frame and other control information.
- D – frame ID
- Sequence Control(SC) – uses stop and wait ARQ
- Frame checksum(FCS) – for error control

Table 14.1 *Subfields in FC field*

Field	Explanation
Version	Current version is 0
Type	Type of information: management (00), control (01), or data (10)
Subtype	Subtype of each type (see Table 14.2)
To DS	Defined later
From DS	Defined later
More flag	When set to 1, means more fragments
Retry	When set to 1, means retransmitted frame
Pwr mgt	When set to 1, means station is in power management mode
More data	When set to 1, means station has more data to send
WEP	Wired equivalent privacy (encryption implemented)
Rsvd	Reserved

Table 14.2 *Values of subfields in control frames*

Subtype	Meaning
1011	Request to Send(CTS)
1100	Clear to Send(CTS)
1101	Acknowledgement(ACK)

Table 14.3 *Addresses*

To DS	From DS	Address 1	Address 2	Address 3	Address 4
0	0	Destination	Source	BSS ID	N/A
0	1	Destination	Sending AP	Source	N/A
1	0	Receiving AP	Source	Destination	N/A
1	1	Receiving AP	Sending AP	Destination	Source

802.15(BLUETOOTH)

Bluetooth is a wireless LAN technology designed to connect devices of different functions such as telephones, notebooks, computers, cameras, printers, coffee makers, and so on. A Bluetooth LAN is an ad hoc network, which means that the network is formed spontaneously.

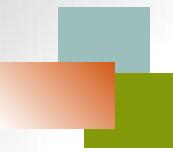
Topics discussed in this section:

Architecture

Bluetooth Layers

Baseband Layer

L2CAP



Note

Piconet is the basic model used in Bluetooth. It consist of one primary station and n secondary stations.

Scatternet is a combination of piconet.

Figure 14.19 Piconet

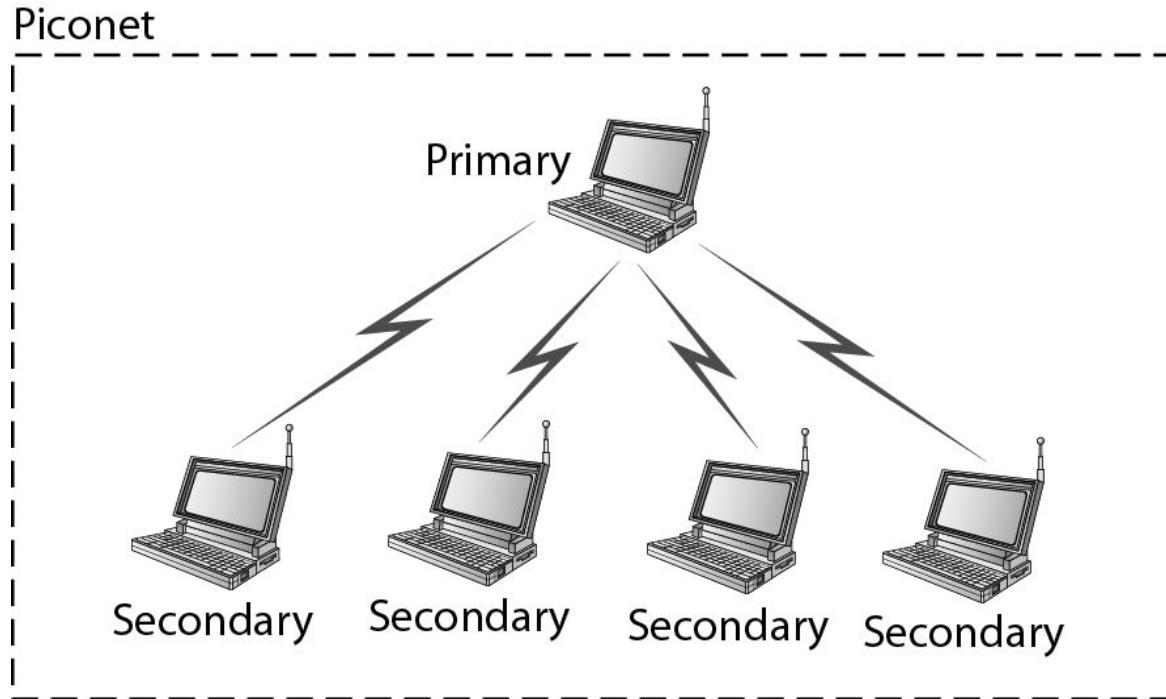
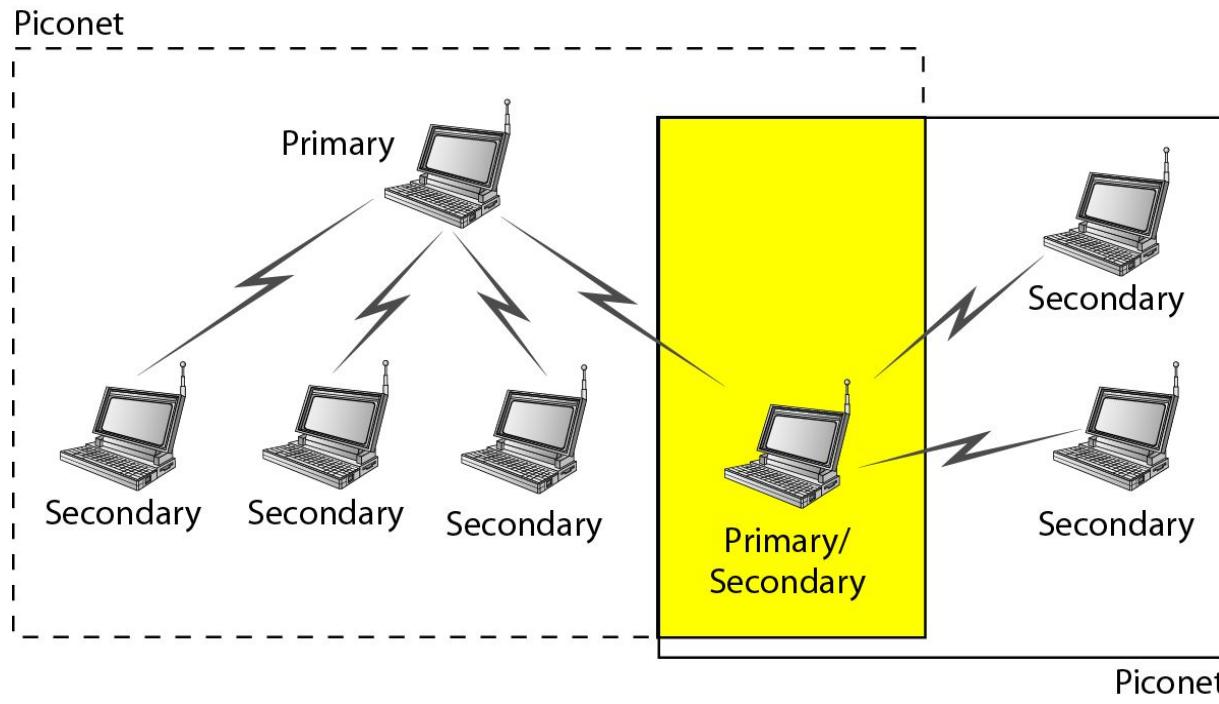


Figure 14.20 Scatternet



Layered Architecture of Bluetooth

1. Radio layer
 - It is similar to physical layer of OSI model
 - Uses 2.4 GHz bandwidth with 79 channels of 1 MHz each
 - Uses FHSS key
 - Slot time is 625 μ s
 - Modulation method – GFSK
2. Baseband layer
 - It is equivalent to MAC layer of data link layer
 - Slot time is 625 μ s
 - Access method is TDMA
 - Support single secondary or multiple secondaries
 - Supports two types of transmission links
 - Synchronous connection oriented link
 - Asynchronous connection oriented link

3. L2CAP(Logical Link Control and Adaptation Protocol)

- Similar to LLC od data link layer
- Duties are:
 - Multiplexing
 - Segmentation and reassembling
 - Quality of Service(QoS)

Figure 14.21 *Bluetooth layers*

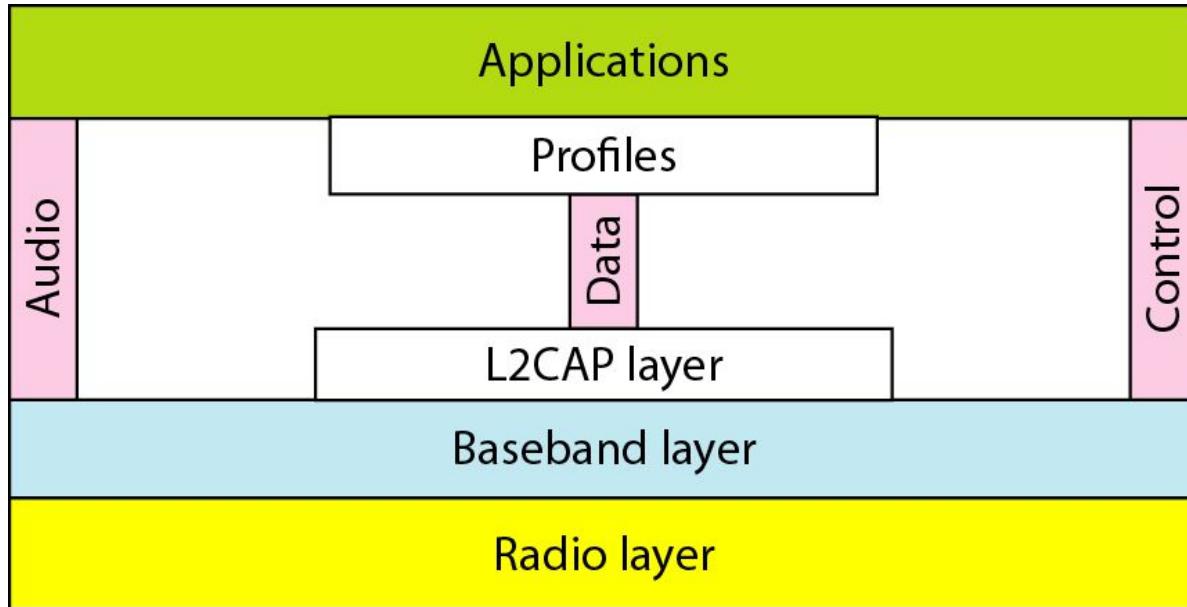
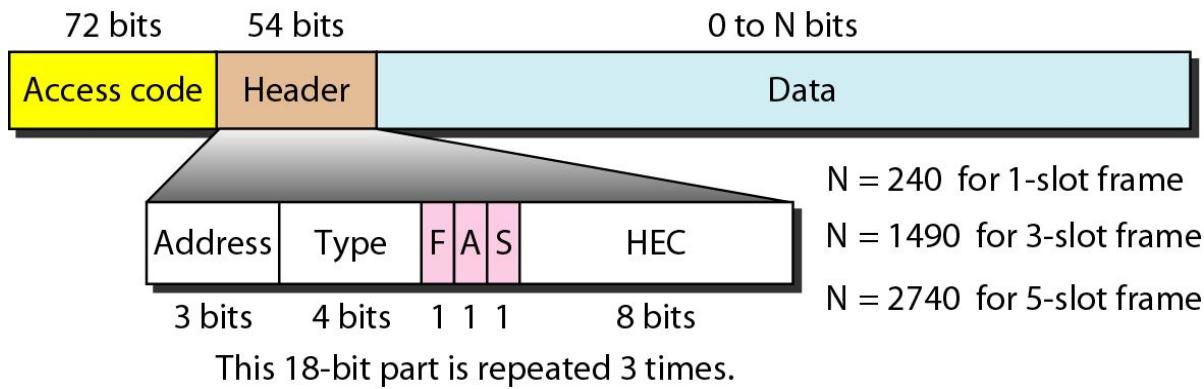


Figure 14.25 L2CAP data packet format



Figure 14.24 Frame format types



F – for flow control

A – stop and wait ARQ

S – sequence number

HEC – header checksum

MODULE 3

Network Layer

Network Layer Services

- **Packetization**
- **Routing and Forwarding**
- **Error control**
- **Flow control**
- **Congestion control**
- **Quality of service**
- **Security**

Routing

- In routing, a packet is routed, hop by hop, from its source to its destination by the help of forwarding tables.
- Routing a packet from its source to its destination means routing the packet from a source router to a destination router.
- There are several routes that a packet can travel from the source to the destination; what must be determined is which route the packet should take.

Least-Cost routing

- The internet can be modeled as a weighted graph, in which each edge is associated with cost.
- One of the ways to interpret the best route from the source router to the destination router is to find the least cost between the two.
- In other words, the source router chooses a route to the destination router in such a way that the total cost for the route is the least cost among all possible routes.
- For this purpose, **least-cost trees** are made, a tree with the source router as the root that spans the whole graph and in which the path between the root and any other node is the shortest.

Routing Algorithms

- Distance Vector Routing
- Link-State Routing
- Flooding

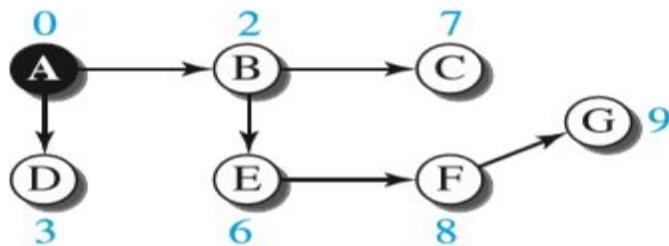
Distance Vector Routing

- In distance-vector routing, the first thing each node creates is its own least-cost tree with the rudimentary information it has about its immediate neighbors.
- The incomplete trees are exchanged between immediate neighbors to make the trees more and more complete and to represent the whole internet.

Distance Vector Routing

- The heart of distance-vector routing is the famous Bellman-Ford equation.
- Distance-vector routing creates a **distance vector**, a one-dimensional array to represent the tree.
- The **name** of the distance vector defines the root, the **indexes** define the destinations, and the **value** of each cell defines the least cost from the root to the destination.

Distance Vector Routing



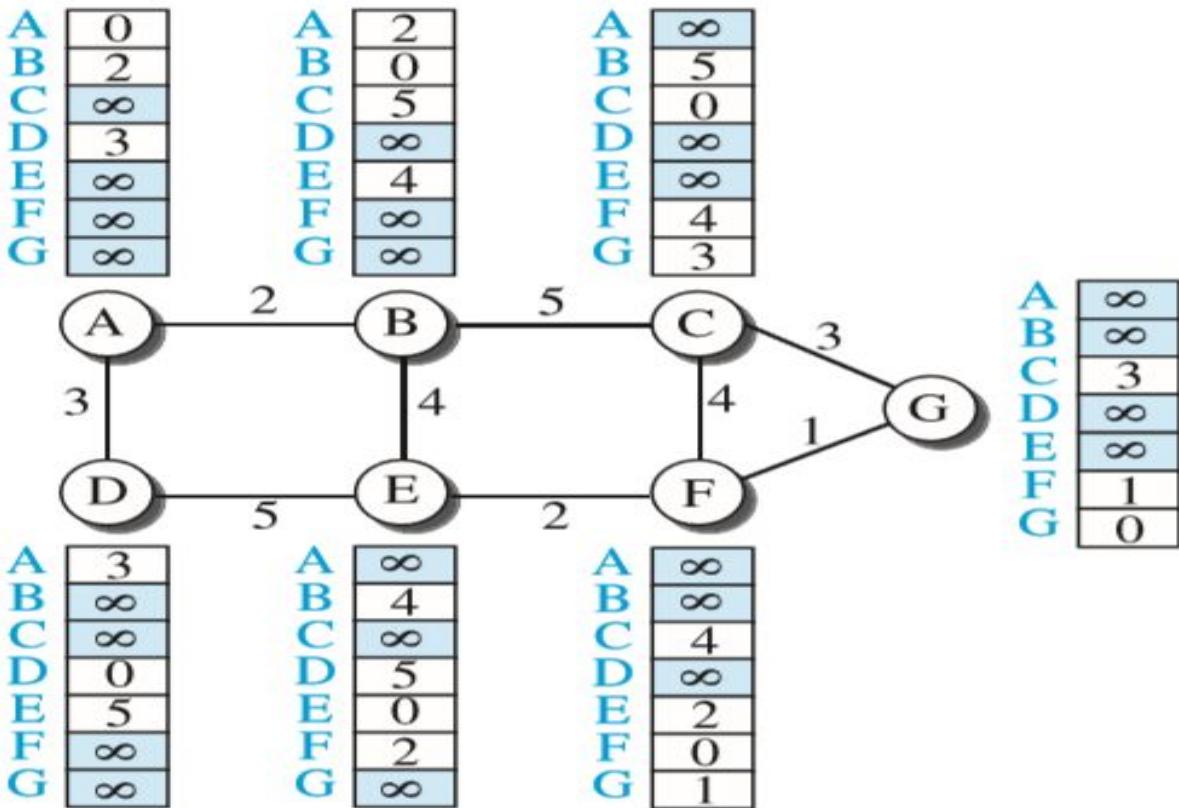
a. Tree for node A

A	0
B	2
C	7
D	3
E	6
F	8
G	9

b. Distance vector for node A

Distance Vector Routing

- Each node in an internet, when it is booted, creates a very rudimentary distance vector with the minimum information the node can obtain from its neighborhood.
- The node sends some greeting messages out of its interfaces and discovers the identity of the immediate neighbors and the distance between itself and each neighbor.
- It then makes a simple distance vector by inserting the discovered distances in the corresponding cells and leaves the value of other cells as infinity.



Distance Vector Routing

- These rudimentary vectors cannot help the internet to effectively forward a packet. For example, node A thinks that it is not connected to node G because the corresponding cell shows the least cost of infinity.
- To improve these vectors, the nodes in the internet need to help each other by exchanging information. After each node has created its vector, it sends a copy of the vector to all its immediate neighbors.
- After a node receives a distance vector from a neighbor, it updates its distance vector using the Bellman-Ford equation.

New B

A	2
B	0
C	5
D	5
E	4
F	∞
G	∞

Old B

A	2
B	0
C	5
D	∞
E	4
F	∞
G	∞

A

A	0
B	2
C	∞
D	3
E	∞
F	∞
G	∞

$$B[] = \min(B[], 2 + A[])$$

New B

A	2
B	0
C	5
D	5
E	4
F	6
G	∞

Old B

A	2
B	0
C	5
D	5
E	4
F	∞
G	∞

E

A	∞
B	4
C	∞
D	5
E	0
F	2
G	∞

$$B[] = \min(B[], 4 + E[])$$

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

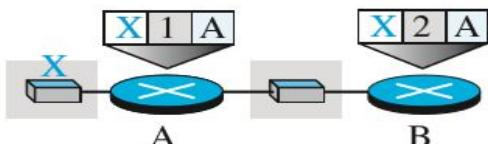
b. Second event: B receives a copy of E's vector.

Distance Vector Routing Problems

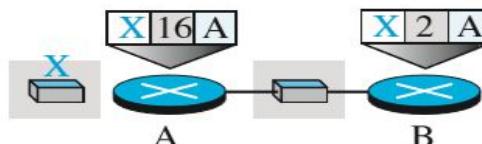
- **Count to Infinity:** A problem with distance-vector routing is that any decrease in cost propagates quickly, but any increase in cost will propagate slowly.
- For a routing protocol to work properly, if a link is broken every other router should be aware of it immediately, but in distance-vector routing, this takes some time. The problem is referred to as count to infinity.
- It sometimes takes several updates before the cost for a broken link is recorded as infinity by all routers.

Distance Vector Routing Problems

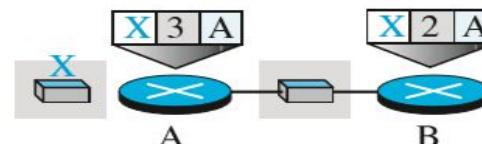
- Two node loop:



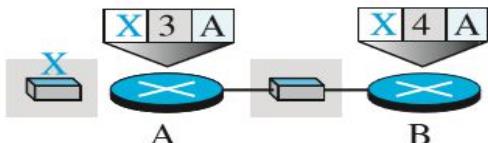
a. Before failure



b. After link failure

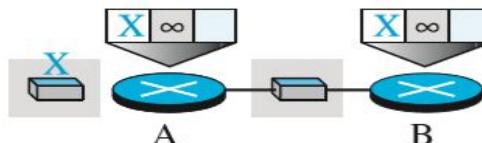


c. After A is updated by B



d. After B is updated by A

...



e. Finally

Distance Vector Routing Problems

- **Split Horizon:** One solution to instability is called split horizon. In this strategy, instead of flooding the table through each interface, each node sends only part of its table through each interface.
- **Poison Reverse:** Replace the infinity value with some warning message.

Link-State Routing

- A routing algorithm that directly uses least cost trees and forwarding tables.
- This method uses the term **link-state** to define the characteristic of a link (an edge) that represents a network in the internet.
- In this algorithm the cost associated with an edge defines the **state** of the link. Links with lower costs are preferred to links with higher costs.

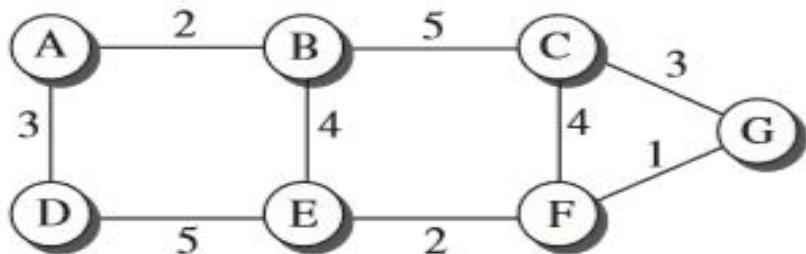
Link-State Routing

- To create a least-cost tree with this method, each node needs to have a complete map of the network, which means it needs to know the state of each link.
- **The collection of states for all links is called the link-state database (LSDB).**
- There is **only one LSDB** for the whole internet.

Link-State Database

- The **Link-State Database** can be represented as a **two-dimensional array**(matrix) in which the value of each cell defines the cost of the corresponding link.

Link-State Routing



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

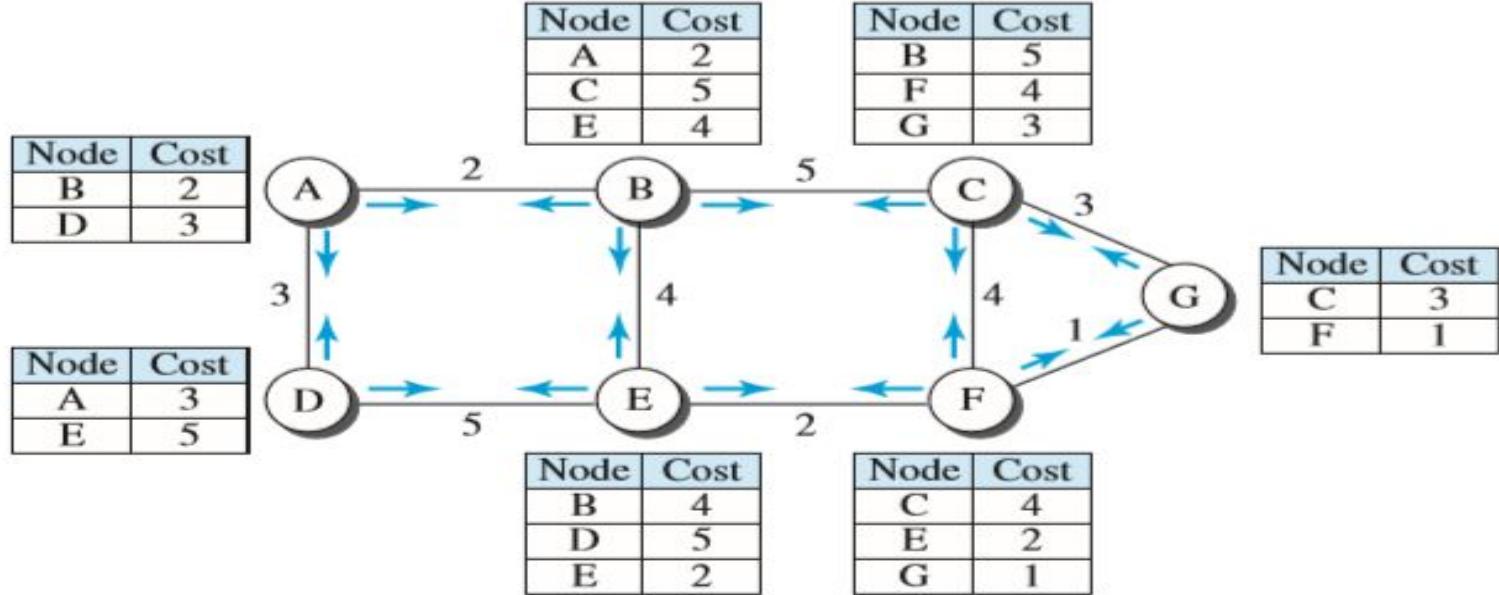
b. Link state database

Link-State Routing

- Creation of LSDB :This can be done by a process called **flooding**.
- **Flooding:** Each node can send some greeting messages to all its immediate neighbors (those nodes to which it is connected directly) to collect two pieces of information for each neighboring node: **the identity of the node** and **the cost of the link**.
- **The combination of these two pieces of information is called the LS packet (LSP);** the LSP is sent out of each interface.

Link-State Routing

- When a node receives an LSP from one of its interfaces, it compares the LSP with the copy it may already have.
- If the newly arrived LSP is older than the one it has (found by checking the sequence number), it discards the LSP.
- If it is newer or the first one received, the node discards the old LSP (if there is one) and keeps the received one.
- It then sends a copy of it out of each interface except the one from which the packet arrived. This guarantees that flooding stops somewhere in the network .



Creation of LSDB

Link-State Routing

- In the **distance-vector routing** algorithm, each router tells its neighbors what it knows about the whole internet; in the **link-state routing** algorithm, each router tells the whole internet what it knows about its neighbors.
- To create a least-cost tree for itself, using the shared LSDB, each node needs to run the famous **Dijkstra Algorithm**.

Routing Protocols

- The protocols used within the same domain are called **Intra-domain** Protocols.
- **RIP and OSPF are Intra-domain Protocols.**

- The protocols used between different domains are called **Interdomain** Protocols.
- **BGP is an Interdomain Protocol.**

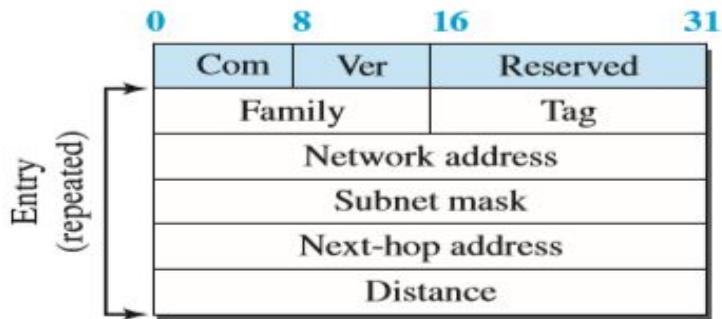
Routing Information Protocol (RIP)

- The Routing Information Protocol (RIP) is one of the most widely used intradomain routing protocols based on the distance-vector routing algorithm .
- The cost is defined as the **number of hops**, which means the number of networks (subnets) a packet needs to travel through from the source router to the final destination host.

Routing Information Protocol (RIP)

- In RIP, the maximum cost of a path can be 15, which means 16 is considered as infinity (no connection). For this reason, RIP can be used only in autonomous systems in which the diameter of the AS is not more than 15 hops.

Routing Information Protocol(RIP)

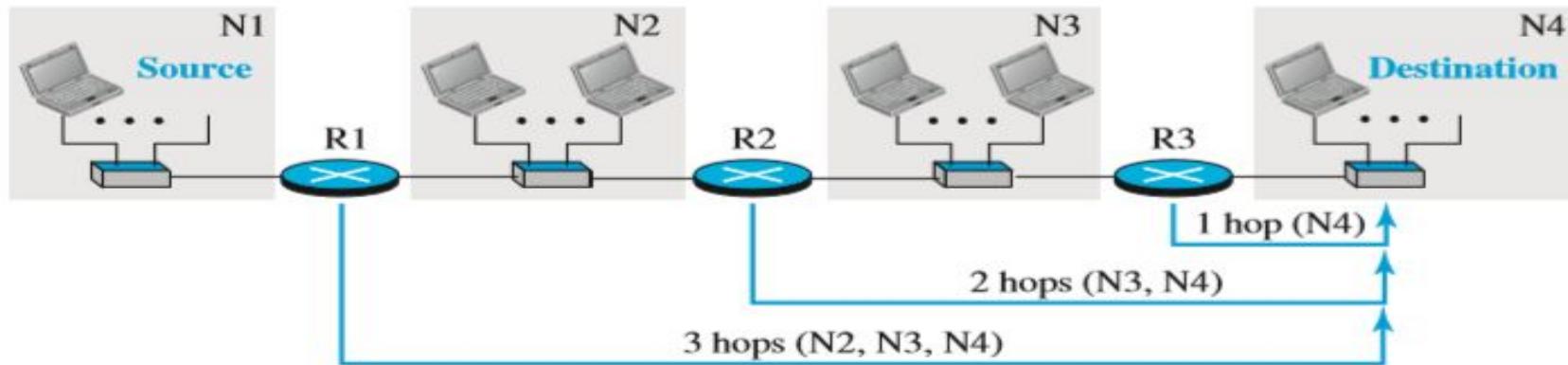


Fields

Com: Command, request (1), response (2)
Ver: Version, current version is 2
Family: Family of protocol, for TCP/IP value is 2
Tag: Information about autonomous system
Network address: Destination address
Subnet mask: Prefix length
Next-hop address: Address length
Distance: Number of hops to the destination

RIP Message format

Routing Information Protocol (RIP)



Routing Information Protocol (RIP)

- This protocol uses **forwarding tables** for routing the packets to required destinations.
- A forwarding table in RIP is a **three-column table**, in which the **first column** is the address of the destination network, the **second column** is the address of the next router to which the packet should be forwarded, and the **third column** is the cost (the number of hops) to reach the destination network.

Routing Information Protocol (RIP)

Forwarding table for R1

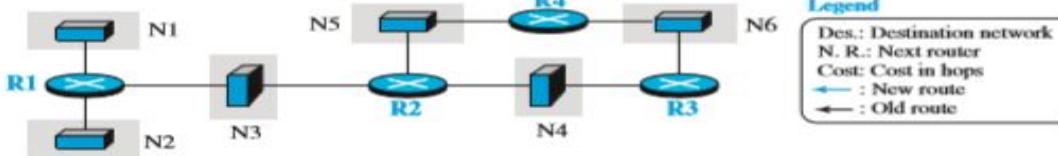
Destination network	Next router	Cost in hops
N1	—	1
N2	—	1
N3	R2	2
N4	R2	3

Forwarding table for R2

Destination network	Next router	Cost in hops
N1	R1	2
N2	—	1
N3	—	1
N4	R3	2

Forwarding table for R3

Destination network	Next router	Cost in hops
N1	R2	3
N2	R2	2
N3	—	1
N4	—	1



R1			R2			R3			R4		
Dest.	N. R.	Cost									
N1	—	1	N3	—	1	N4	—	1	N5	—	1
N2	—	1	N4	—	1	N6	—	1	N6	—	1
N3	—	1	N5	—	1	N6	—	1	N6	—	1

Forwarding tables after all routers booted

New R1			Old R1			R2 Seen by R1			New R3			Old R3			R2 Seen by R3			New R4			Old R4			R2 Seen by R4					
Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost	Dest.	N. R.	Cost
N1	—	1	N1	—	1	N3	R2	2	N3	R2	2	N4	—	1	N3	R2	2	N3	R2	2	N4	—	1	N5	R2	2	N5	R2	2
N2	—	1	N2	—	1	N2	—	1	N4	—	1	N4	—	1	N6	—	1	N4	—	1	N4	—	1	N5	—	1	N5	—	1
N3	—	1	N3	—	1	N5	R2	2	N5	R2	2	N6	—	1	N5	R2	2	N5	R2	2	N6	—	1	N6	—	1	N6	—	1
N4	R2	2	N4	—	1	N6	—	1	N6	—	1	N6	—	1	N6	—	1	N6	—	1	N6	—	1	N6	—	1	N6	—	1
N5	R2	2	N5	—	1	N5	—	1	N5	—	1	N6	—	1	N5	—	1	N5	—	1	N5	—	1	N6	—	1	N6	—	1

Changes in the forwarding tables of R1, R3, and R4 after they receive a copy of R2's table

Final R1			Final R2			Final R3			Final R4		
Dest.	N. R.	Cost									
N1	—	1	N1	R1	2	N1	R2	3	N1	R2	3
N2	—	1	N2	R1	2	N2	R2	3	N2	R2	3
N3	—	1	N3	—	1	N3	R2	2	N3	R2	2
N4	R2	2	N4	—	1	N4	—	1	N4	R2	2
N5	R2	2	N5	—	1	N5	R2	2	N5	—	1
N6	R2	3	N6	R3	2	N6	—	1	N6	—	1

Forwarding tables for all routers after they have been stabilized

Open Shortest Path First (OSPF)

- Open Shortest Path First (OSPF) is also an **intradomain routing** protocol like RIP, but it is based on the link-state routing protocol.
- It also uses **forwarding tables** to route the information to the correct destination.

Open Shortest Path First(OSPF)

- Each OSPF router can create a **forwarding table** after finding the shortest-path tree between itself and the destination using **Dijkstra's algorithm**.

Forwarding table for R1

Destination network	Next router	Cost
N1	—	4
N2	—	5
N3	R2	8
N4	R2	12

Forwarding table for R2

Destination network	Next router	Cost
N1	R1	9
N2	—	5
N3	—	3
N4	R3	7

Forwarding table for R3

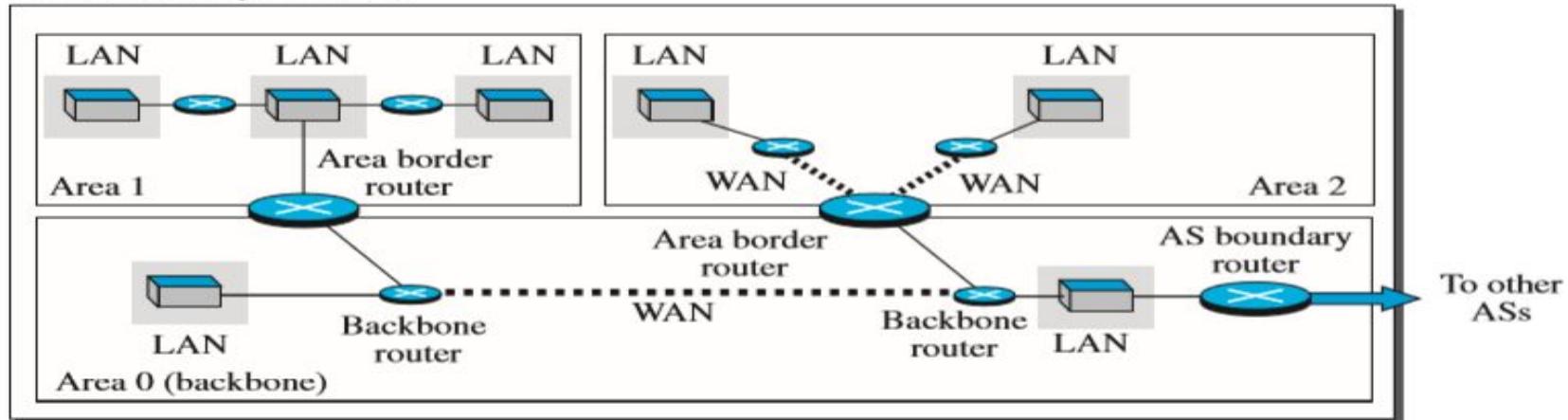
Destination network	Next router	Cost
N1	R2	12
N2	R2	8
N3	—	3
N4	—	4

Open Shortest Path First(OSPF)

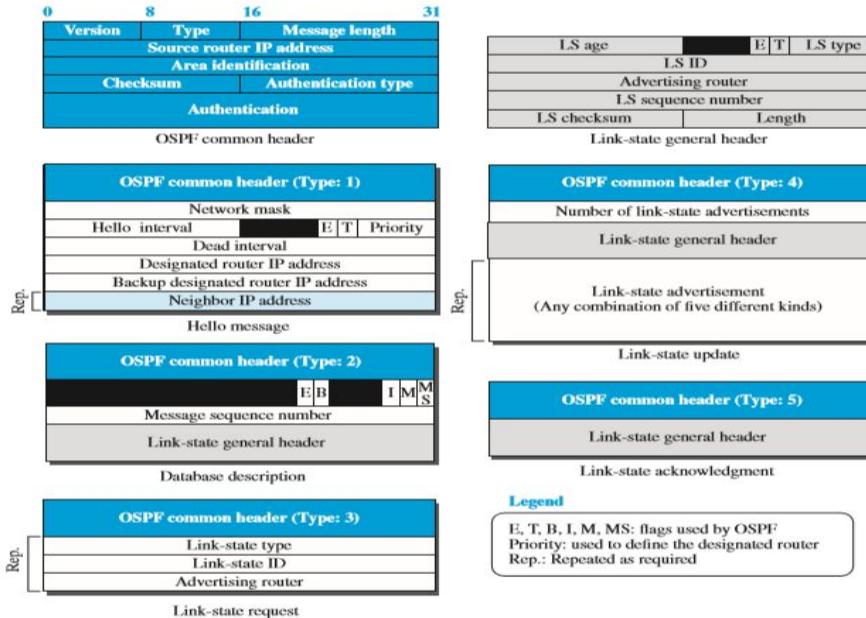
- OSPF was designed for autonomous systems.
- The formation of shortest-path trees in OSPF requires that all routers flood the whole AS with their LSPs to create the global LSDB.
- The AS needs to be divided into small sections called **areas**. Each area acts as a small independent domain for flooding LSPs.
- OSPF uses another level of hierarchy in routing: **the first level is the autonomous system, the second is the area.**

Open Shortest Path First(OSPF)

Autonomous System (AS)



Open Shortest Path First (OSPF)

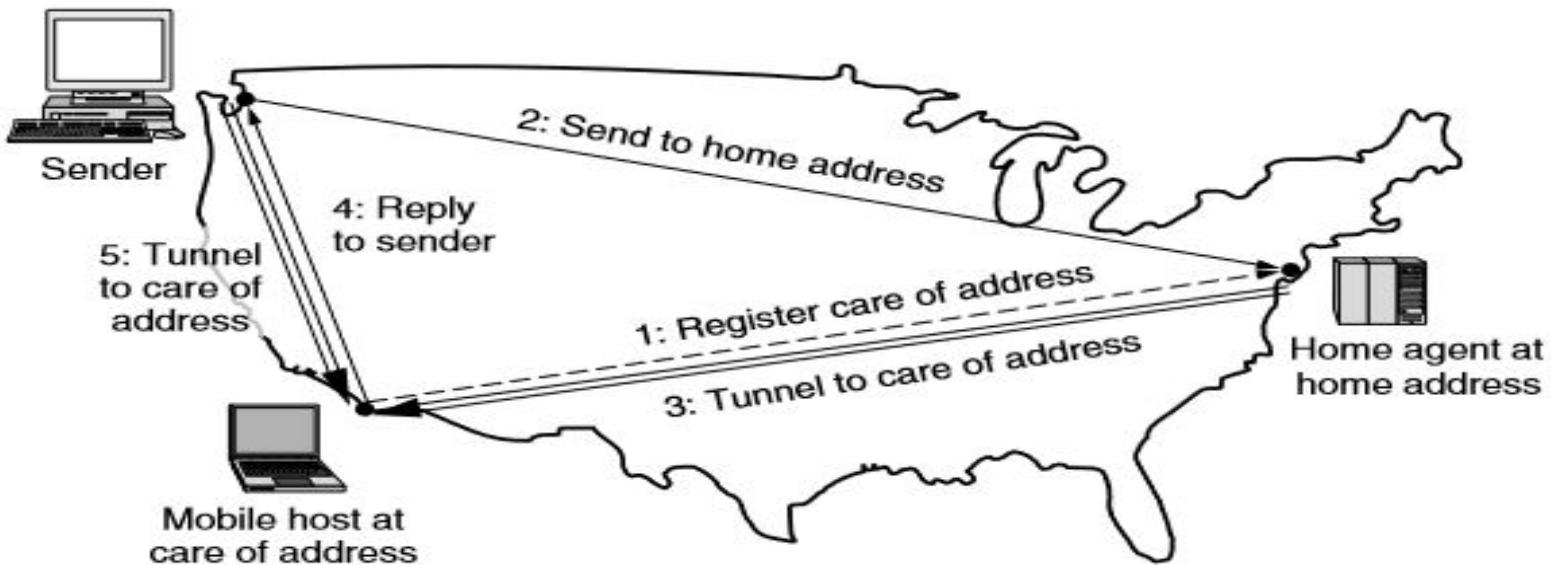


OSPF
Message
Format

Routing in mobile hosts

- The basic idea used for mobile routing in the Internet and cellular networks is for the mobile host to tell a host at the home location where it is now.
- This host, which acts on behalf of the mobile host, is called the **home agent**. Once it knows where the mobile host is currently located, it can forward packets so that they are delivered.
- The local network address acquired by a mobile host is known as **Care of address**.

Routing in mobile hosts



Routing in mobile hosts

- Step 1: Once the mobile host has care of address, it can tell its home agent where it is now. It does this by sending a registration message to the home agent with the care of address.
- Step 2: Next, the sender sends a data packet to the mobile host using its permanent address .

Routing in mobile hosts

- Step 3: It then wraps or encapsulates the packet with a new header and sends this bundle to the care of address. This mechanism is called **tunneling**.
- Step 4: The mobile host then sends its reply packet directly to the sender. The overall route is called **triangle routing** because it may be circuitous if the remote location is far from the home location.

Routing in mobile hosts

- Step 5: Subsequent packets can be routed directly to the mobile host by tunneling them to the care of address, bypassing the home location entirely.
- If connectivity is lost for any reason as the mobile moves, the home address can always be used to reach the mobile.

MODULE 4

Index

- Congestion Control Algorithms – QoS
- Internetworking- Network layer in internet.

What is congestion?

A state occurring in network layer when the message traffic is so heavy that it slows down network response time.

Effects of Congestion:

- As delay increases, performance decreases.
- If delay increases, retransmission occurs, making situation worse.

Causes of Congestion:

The various causes of congestion in a subnet are:

- The input traffic rate exceeds the capacity of the output lines. If suddenly, a stream of packet start arriving on three or four input lines and all need the same output line. In this case, a queue will be built up. If there is insufficient memory to hold all the packets, the packet will be lost.
- Increasing the memory to unlimited size does not solve the problem. This is because, by the time packets reach front of the queue, they have already timed out (as they waited the queue).
- When timer goes off source transmits duplicate packet that are also added to the queue. Thus same packets are added again and again, increasing the load all the way to the destination.

- Congestion is also caused by slow links. This problem will be solved when high speed links are used. But it is not always the case.
- Sometimes increase in link bandwidth can further deteriorate the congestion problem as higher speed links may make the network more unbalanced.

- Congestion in a subnet can occur if the processors are slow. Slow speed CPU at routers will perform the routine tasks such as queuing buffers, updating table etc slowly. As a result of this, queues are built up even though there is excess line capacity.
- The routers are too slow to perform bookkeeping tasks (queuing buffers, updating tables, etc.).
- The routers' buffer is too limited.

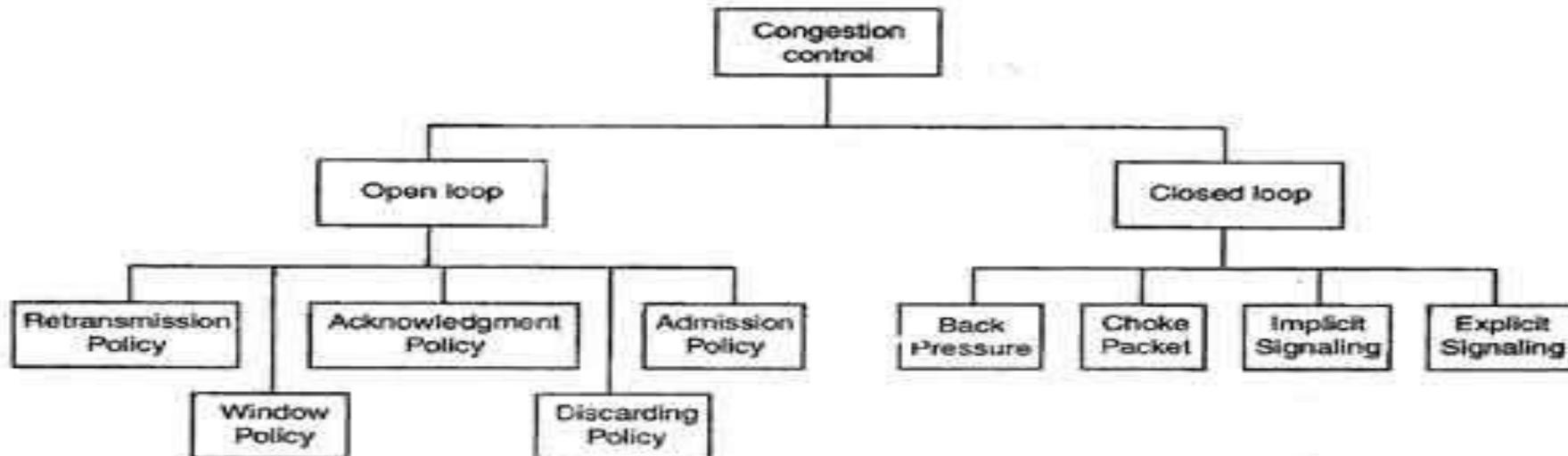
How to correct the Congestion Problem:

Congestion Control refers to techniques and mechanisms that can either prevent congestion, before it happens, or remove congestion, after it has happened.

Congestion control mechanisms are divided into two categories, one category prevents the congestion from happening and the other category removes congestion after it has taken place.

Two categories are:

Open loop And Closed loop



Types of Congestion Control Methods

Open Loop Congestion Control

In this method, policies are used to prevent the congestion before it happens. Congestion control is handled either by the source or by the destination.

The various methods used for open loop congestion control are:

- Retransmission Policy
- Window Policy
- Admission Policy
- Acknowledgement Policy
- Discarding Policy

Retransmission Policy

- The sender retransmits a packet, if it feels that the packet it has sent is lost or corrupted.
- However retransmission in general may increase the congestion in the network. But we need to implement good retransmission policy to prevent congestion.
- The retransmission policy and the retransmission timers need to be designed to optimize efficiency and at the same time prevent the congestion.

Window Policy

- To implement window policy, selective reject window method is used for congestion control.
- Selective Reject method is preferred over Go-back-n window as in Go-back-n method, when timer for a packet times out, several packets are resent, although some may have arrived safely at the receiver. Thus, this duplication may make congestion worse.
- Selective reject method sends only the specific lost or damaged packets.

Acknowledgement Policy

The acknowledgement policy imposed by the receiver may also affect congestion.

- If the receiver does not acknowledge every packet it receives it may slow down the sender and help prevent congestion.
- Acknowledgments also add to the traffic load on the network. Thus, by sending fewer acknowledgements we can reduce load on the network.
- To implement it, several approaches can be used:
 1. A receiver may send an acknowledgement only if it has a packet to be sent.
 2. A receiver may send an acknowledgement when a timer expires.
 3. A receiver may also decide to acknowledge only N packets at a time.

Discarding Policy

- A router may discard less sensitive packets when congestion is likely to happen.
- Such a discarding policy may prevent congestion and at the same time may not harm the integrity of the transmission.

Admission Policy

- An admission policy, which is a quality-of-service mechanism, can also prevent congestion in virtual circuit networks.
- Switches in a flow first check the resource requirement of a flow before admitting it to the network.
- A router can deny establishing a virtual circuit connection if there is congestion in the "network or if there is a possibility of future congestion.

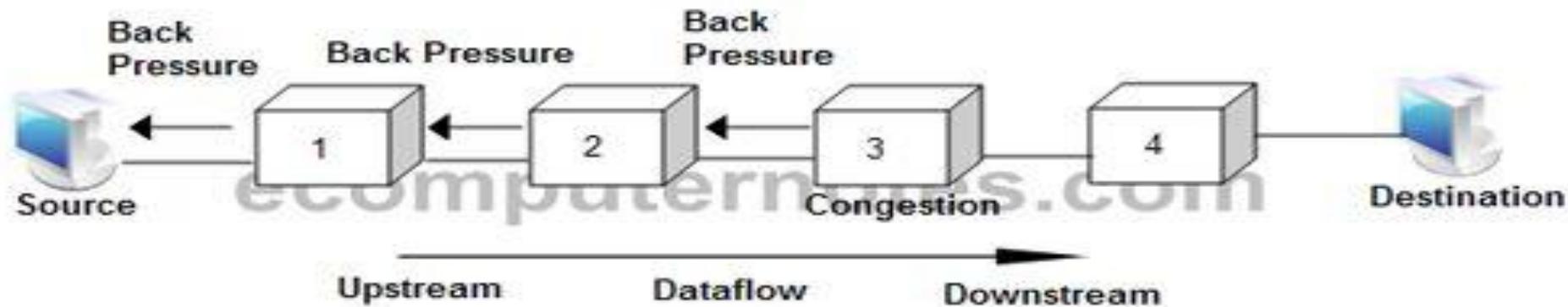
Closed Loop Congestion Control

Closed loop congestion control mechanisms try to remove the congestion after it happens.

The various methods used for closed loop congestion control are:

1. Back Pressure
2. Choke Packet
3. Implicit Signaling
4. Explicit Signaling

Backpressure

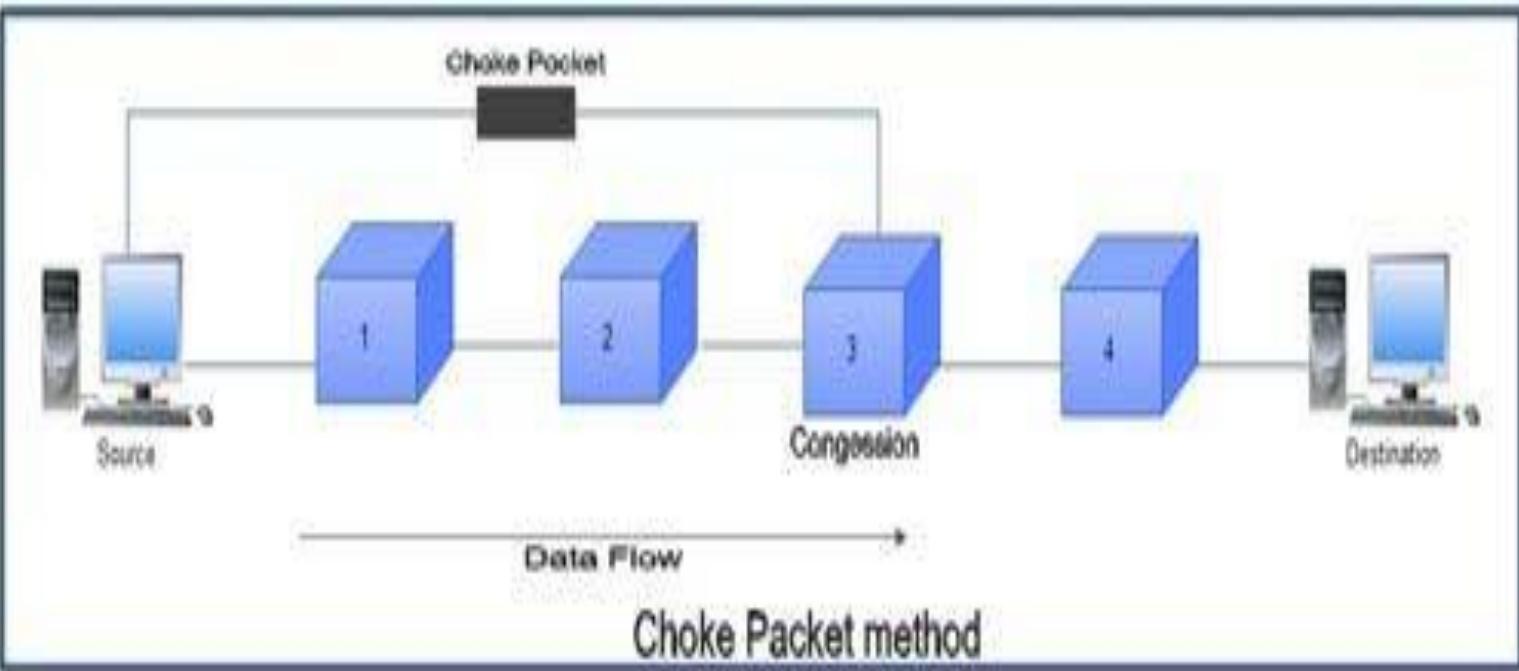


Backpressure Method

- The backpressure technique can be applied only to virtual circuit networks. In such virtual circuit each node knows the upstream node from which a data flow is coming.
- In this method of congestion control, the congested node stops receiving data from the immediate upstream node or nodes.
- This may cause the upstream node or nodes to become congested, and they, in turn, reject data from their upstream node or nodes.
- As shown in fig node 3 is congested and it stops receiving packets and informs its upstream node 2 to slow down. Node 2 in turns may be congested and informs node 1 to slow down. Now node 1 may create congestion and informs the source node to slow down. In this way the congestion is alleviated. Thus, the pressure on node 3 is moved backward to the source to remove the congestion.

Choke Packet

- In this method of congestion control, congested router or node sends a special type of packet called choke packet to the source to inform it about the congestion.
- Here, congested node does not inform its upstream node about the congestion as in backpressure method.
- In choke packet method, congested node sends a warning directly to the source station *i.e.* the intermediate nodes through which the packet has traveled are not warned.



Implicit Signaling

- In implicit signaling, there is no communication between the congested node or nodes and the source.
- The source guesses that there is congestion somewhere in the network when it does not receive any acknowledgment. Therefore the delay in receiving an acknowledgment is interpreted as congestion in the network.
- On sensing this congestion, the source slows down.
- This type of congestion control policy is used by TCP.

Explicit Signaling

- In this method, the congested nodes explicitly send a signal to the source or destination to inform about the congestion.
- Explicit signaling is different from the choke packet method. In choke packed method, a separate packet is used for this purpose whereas in explicit signaling method, the signal is included in the packets that carry data .
- Explicit signaling can occur in either the forward direction or the backward direction .
- In backward signaling, a bit is set in a packet moving in the direction opposite to the congestion. This bit warns the source about the congestion and informs the source to slow down.
- In forward signaling, a bit is set in a packet moving in the direction of congestion. This bit warns the destination about the congestion. The receiver in this case uses policies such as slowing down the acknowledgements to remove the congestion.

Quality of Service Networking

Quality of Service (QoS) refers to the capability of a network to provide better service to selected network traffic over various technologies, including Frame Relay, Asynchronous Transfer Mode (ATM), Ethernet and 802.1 networks, SONET, and IP-routed networks that may use any or all of these underlying technologies.

Techniques to improve QoS:

- Proper Schedule(using the proper queue concept at the intermediate nodes)
- Traffic Shaping
- Token Bucket Algorithm

Traffic Shaping

- Shaping is used to create a traffic flow that limits the full bandwidth potential of the flow(s). This is used many times to prevent the overflow problem mentioned in the introduction. For instance, many network topologies use Frame Relay in a hub-and-spoke design.
- In this case, the central site normally has a high-bandwidth link (say, T1), while remote sites have a low-bandwidth link in comparison (say, 384 Kbps). In this case, it is possible for traffic from the central site to overflow the low bandwidth link at the other end.
- Shaping is a perfect way to pace traffic closer to 384 Kbps to avoid the overflow of the remote link. Traffic above the

Leaky Bucket Algorithm(just for a reference to help understanding token bucket algorithm)

- It is a traffic shaping mechanism that controls the amount and the rate of the traffic sent to the network.
- A leaky bucket algorithm shapes bursty traffic into fixed rate traffic by averaging the data rate.
- Imagine a bucket with a small hole at the bottom.
- The rate at which the water is poured into the bucket is not fixed and can vary but it leaks from the bucket at a constant rate. Thus (as long as water is present in bucket), the rate at which the water leaks does not depend on the rate at which the water is input to the bucket.

- Also, when the bucket is full, any additional water that enters into the bucket spills over the sides and is lost.
- The same concept can be applied to packets in the network. Consider that data is coming from the source at variable speeds. Suppose that a source sends data at 12 Mbps for 4 seconds. Then there is no data for 3 seconds. The source again transmits data at a rate of 10 Mbps for 2 seconds. Thus, in a time span of 9 seconds, 68 Mb data has been transmitted.
- If a leaky bucket algorithm is used, the data flow will be 8 Mbps for 9 seconds. Thus constant flow is maintained.

Token Bucket Algorithm

- The leaky bucket algorithm allows only an average (constant) rate of data flow. Its major problem is that it cannot deal with bursty data.
- A leaky bucket algorithm does not consider the idle time of the host. For example, if the host was idle for 10 seconds and now it is willing to sent data at a very high speed for another 10 seconds, the total data transmission will be divided into 20 seconds and average data rate will be maintained. The host is having no advantage of sitting idle for 10 seconds.
- To overcome this problem, a token bucket algorithm is used. A token bucket algorithm allows bursty data transfers.
- A token bucket algorithm is a modification of leaky bucket in which leaky bucket contains tokens.

- In this algorithm, a token(s) are generated at every clock tick. For a packet to be transmitted, system must remove token(s) from the bucket.
- Thus, a token bucket algorithm allows idle hosts to accumulate credit for the future in form of tokens.
- For example, if a system generates 100 tokens in one clock tick and the host is idle for 100 ticks. The bucket will contain 10,000 tokens.
- Now, if the host wants to send bursty data, it can consume all 10,000 tokens at once for sending 10,000 cells or bytes.
- Thus a host can send bursty data as long as bucket is not empty.

Internetworking

Internetworking is the practice of connecting a computer network with other networks through the use of gateways that provide a common method of routing information packets between the networks.

The resulting system of interconnected networks are called an *internetwork*, or simply an *internet*. Internetworking is a combination of the words *inter*("between") and networking; not *internet-working* or *international-network*.

The Network Layer

- The transport layer enables the applications to efficiently and reliably exchange data. Transport layer entities expect to be able to send segment to any destination without having to understand anything about the underlying subnetwork technologies.
- Many subnetwork technologies exist. Most of them differ in subtle details (frame size, addressing, ...). The network layer is the glue between these subnetworks and the transport layer. It hides to the transport layer all the complexity of the underlying subnetworks and ensures that information can be exchanged between hosts connected to different types of subnetworks.

Static routing

The simplest solution is to pre-compute all the routing tables of all routers and to install them on each router. Several algorithms can be used to compute these tables.

A simple solution is to use shortest path routing and to minimise the number of intermediate routers to reach each destination. More complex algorithms can take into account the expected load on the links to ensure that congestion does not occur for a given traffic demand. These algorithms must all ensure that :

- All routers are configured with a route to reach each destination
- none of the paths composed with the entries found in the routing tables contain a cycle. Such a cycle would lead to a forwarding loop.

Distance Vector Routing

Distance vector routing is a simple distributed routing protocol. It allows routers to automatically discover the destinations reachable inside the network as well as the shortest path to reach each of these destinations. The shortest path is computed based on *metrics* or *costs* that are associated to each link. We use $l.\text{cost}$ to represent the metric that has been configured for link l on a router.

Each router maintains a routing table. The routing table R can be modelled as a data structure that stores, for each known destination address d , the following attributes :

- $R[d].link$ is the outgoing link that the router uses to forward packets towards destination d
- $R[d].cost$ is the sum of the metrics of the links that compose the shortest path to reach destination d
- $R[d].time$ is the timestamp of the last distance vector containing destination d

Link State Routing

- Link state routing is the second family of routing protocols. While distance vector routers use a distributed algorithm to compute their routing tables, link-state routers exchange messages to allow each router to learn the entire network topology.
- Based on this learned topology, each router is then able to compute its routing table by using a shortest path computation [Dijkstra1959].

For link-state routing, a network is modelled as a *directed weighted graph*. Each router is a node, and the links between routers are the edges in the graph. A positive weight is associated to each directed edge and routers use the shortest path to reach each destination. In practice, different types of weight can be associated to each directed edge :

- Unit weight. If all links have a unit weight, shortest path routing prefers the paths with the least number of intermediate routers.
- Weight proportional to the propagation delay on the link. If all link weights are configured this way, shortest path routing uses the paths with the smallest propagation delay.
- Where C is a constant larger than the highest link bandwidth in the network. If all link weights are configured this way, shortest path routing prefers higher bandwidth paths over lower bandwidth paths

IPv4

- ▶ Internet Protocol version 4 (IPv4) is the fourth version in the development of the Internet Protocol (IP) and the first version of the protocol to be widely deployed. IPv4 is described in IETF publication RFC 791 (September 1981), replacing an earlier definition (RFC 760, January 1980).

Network Basics

- ▶ **Hosts** - Hosts are said to be situated at ultimate end of the network, i.e. a host is a source of information and another host will be the destination. Information flows end to end between hosts. A host can be a user's PC, an internet Server, a database server etc.
- ▶ **Media** - If wired, then it can be copper cable, fiber optic cable, and coaxial cable. If wireless, it can be free-to-air radio frequency or some special wireless band. Wireless frequencies can be used to interconnect remote sites too.
- ▶ **Hub** - A hub is a multiport repeater and it is used to connect hosts in a LAN segment. Because of low throughputs hubs are now rarely used. Hub works on Layer-1 (Physical Layer) of OSI Model.

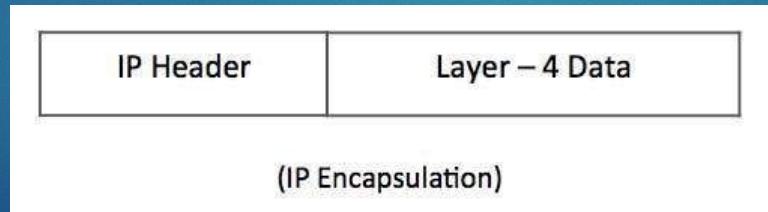
- ▶ **Switch** - A Switch is a multiport bridge and is used to connect hosts in a LAN segment. Switches are much faster than Hubs and operate on wire speed. Switch works on Layer-2 (Data Link Layer), but Layer-3 (Network Layer) switches are also available.
- ▶ **Router** - A router is Layer-3 (Network Layer) device which makes routing decisions for the data/information sent for some remote destination. Routers make the core of any interconnected network and the Internet.
- ▶ **Gateways** - A software or combination of software and hardware put together, works for exchanging data among networks which are using different protocols for sharing data.
- ▶ **Firewall** - Software or combination of software and hardware, used to protect users data from unintended recipients on the network/internet.

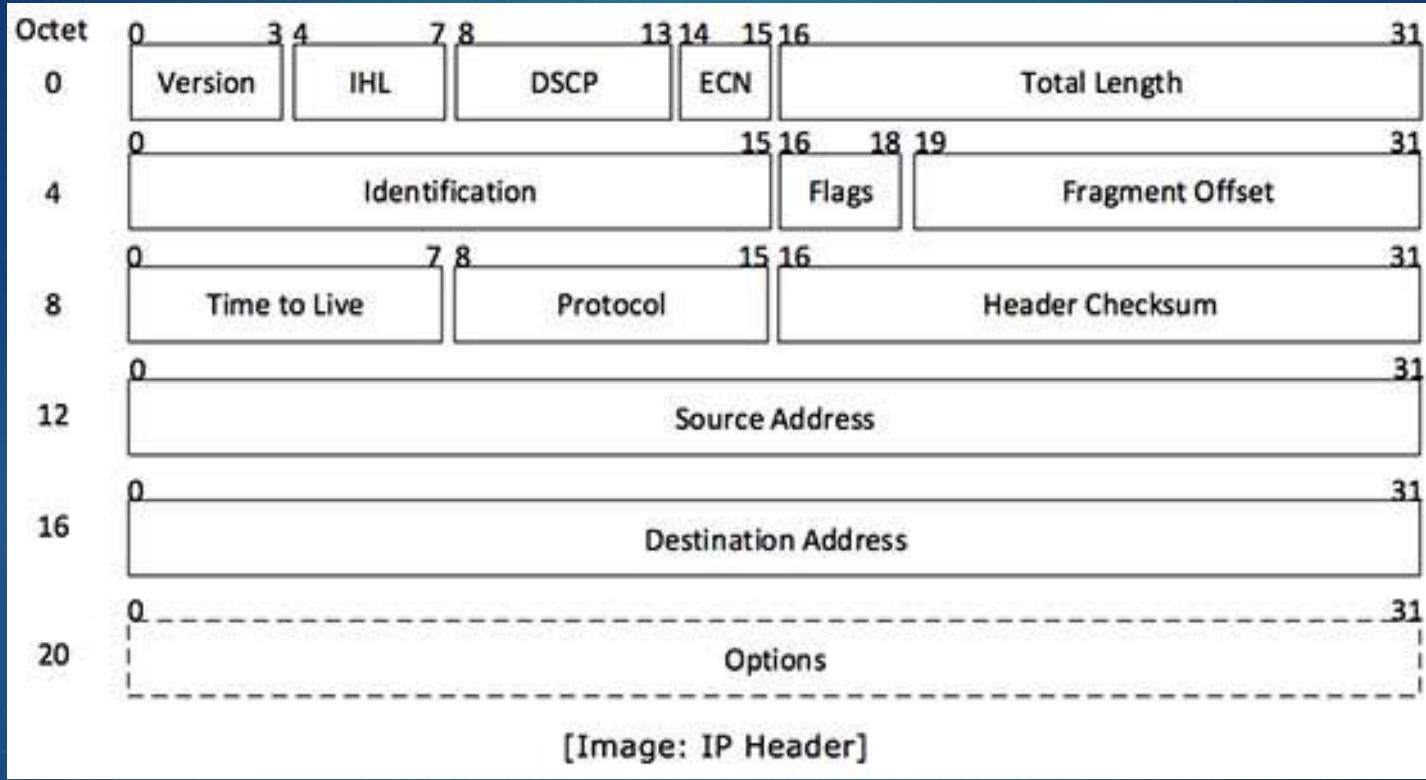
Host Addressing

- ▶ Communication between hosts can happen only if they can identify each other on the network. In a single collision domain (where every packet sent on the segment by one host is heard by every other host) hosts can communicate directly via MAC address.
- ▶ MAC address is a factory coded 48-bits hardware address which can also uniquely identify a host. But if a host wants to communicate with a remote host, i.e. not in the same segment or logically not connected, then some means of addressing is required to identify the remote host uniquely. A logical address is given to all hosts connected to Internet and this logical address is called **Internet Protocol Address**.

IPv4 - Packet Structure

- ▶ Internet Protocol being a layer-3 protocol (OSI) takes data Segments from layer-4 (Transport) and divides it into packets. IP packet encapsulates data unit received from above layer and add to its own header information.
- ▶ The encapsulated data is referred to as IP Payload. IP header contains all the necessary information to deliver the packet at the other end.





IP header includes many relevant information including Version Number, which, in this context, is 4. Other details are as follows:

- ▶ **Version:** Version no. of Internet Protocol used (e.g. IPv4).
- ▶ **IHL:** Internet Header Length; Length of entire IP header.
- ▶ **DSCP:** Differentiated Services Code Point; this is Type of Service.
- ▶ **ECN:** Explicit Congestion Notification; It carries information about the congestion seen in the route.
- ▶ **Total Length:** Length of entire IP Packet (including IP header and IP Payload).
- ▶ **Identification:** If IP packet is fragmented during the transmission, all the fragments contain same identification number. to identify original IP packet they belong to.
- ▶ **Flags:** As required by the network resources, if IP Packet is too large to handle, these 'flags' tells if they can be fragmented or not. In this 3-bit flag, the MSB is always set to '0'.
- ▶ **Fragment Offset:** This offset tells the exact position of the fragment in the original IP Packet.

- ▶ **Time to Live:** To avoid looping in the network, every packet is sent with some TTL value set, which tells the network how many routers (hops) this packet can cross. At each hop, its value is decremented by one and when the value reaches zero, the packet is discarded.
- ▶ **Protocol:** Tells the Network layer at the destination host, to which Protocol this packet belongs to, i.e. the next level Protocol. For example protocol number of ICMP is 1, TCP is 6 and UDP is 17.
- ▶ **Header Checksum:** This field is used to keep checksum value of entire header which is then used to check if the packet is received error-free.
- ▶ **Source Address:** 32-bit address of the Sender (or source) of the packet.
- ▶ **Destination Address:** 32-bit address of the Receiver (or destination) of the packet.
- ▶ **Options:** This is optional field, which is used if the value of IHL is greater than 5. These options may contain values for options such as Security, Record Route, Time Stamp, etc.

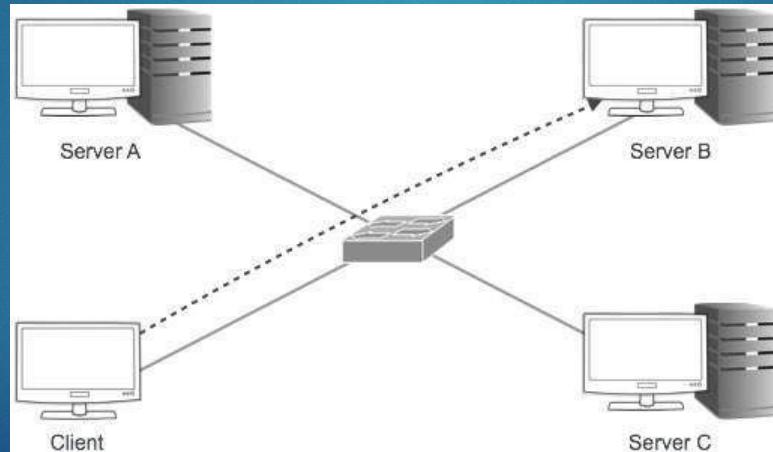
IPv4 - Addressing

IPv4 supports three different types of addressing modes :

- ▶ Unicast Addressing Mode
- ▶ Broadcast Addressing Mode
- ▶ Multicast Addressing Mode

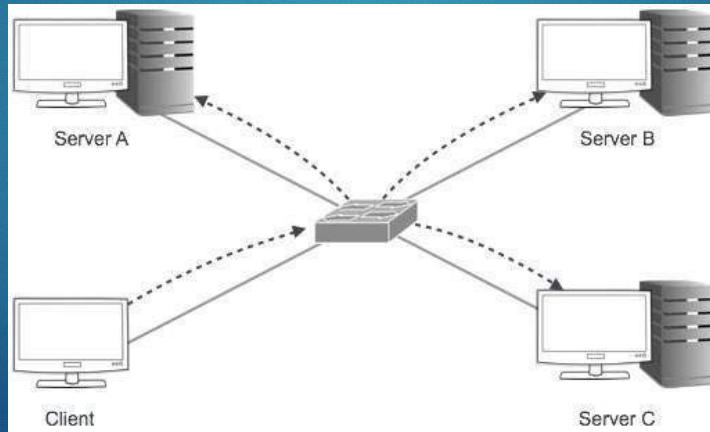
Unicast Addressing Mode:

- In this mode, data is sent only to one destined host. The Destination Address field contains 32-bit IP address of the destination host. Here the client sends data to the targeted server:



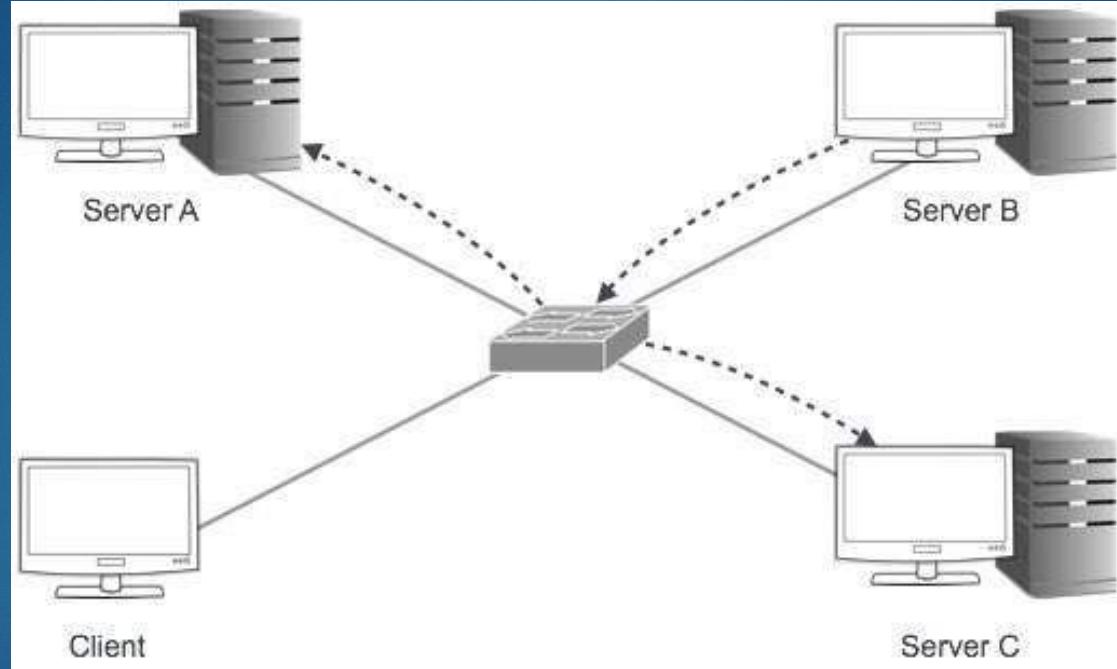
Broadcast Addressing Mode:

- In this mode, the packet is addressed to all the hosts in a network segment. The Destination Address field contains a special broadcast address, i.e. **255.255.255.255**. When a host sees this packet on the network, it is bound to process it. Here the client sends a packet, which is entertained by all the Servers:



Multicast Addressing Mode:

- ▶ This mode is a mix of the previous two modes, i.e. the packet sent is neither destined to a single host nor all the hosts on the segment. In this packet, the Destination Address contains a special address which starts with 224.x.x.x and can be entertained by more than one host.
- ▶ Here a server sends packets which are entertained by more than one servers. Every network has one IP address reserved for the Network Number which represents the network and one IP address reserved for the Broadcast Address, which represents all the hosts in that network.



Hierarchical Addressing Scheme

- ▶ IPv4 uses hierarchical addressing scheme. An IP address, which is 32-bits in length, is divided into two or three parts as depicted:

8 bits	8 bits	8 bits	8 bits
Network	Network	Sub-Network	Host

- ▶ A single IP address can contain information about the network and its sub-network and ultimately the host. This scheme enables the IP Address to be hierarchical where a network can have many sub-networks which in turn can have many hosts.

Subnet Mask

- ▶ The 32-bit IP address contains information about the host and its network. It is very necessary to distinguish both. For this, routers use Subnet Mask, which is as long as the size of the network address in the IP address. Subnet Mask is also 32 bits long.
- ▶ If the IP address in binary is ANDed with its Subnet Mask, the result yields the Network address. For example, say the IP Address is 192.168.1.152 and the Subnet Mask is 255.255.255.0 then:

IP	192.168.1.152	11000000	10101000	00000001	10011000	ANDed
Mask	255.255.255.0	11111111	11111111	11111111	00000000	
Network	192.168.1.0	11000000	10101000	00000001	00000000	Result

- ▶ This way the Subnet Mask helps extract the Network ID and the Host from an IP Address. It can be identified now that 192.168.1.0 is the Network number and 192.168.1.152 is the host on that network.

Binary Representation

- ▶ The positional value method is the simplest form of converting binary from decimal value. IP address is 32 bit value which is divided into 4 octets. A binary octet contains 8 bits and the value of each bit can be determined by the position of bit value '1' in the octet.

MSB	8 th	7 th	6 th	5 th	4 th	3 rd	2 nd	1 st	LSB
	1	1	1	1	1	1	1	1	
Positional Value	128	64	32	16	8	4	2	1	

- ▶ Positional value of bits is determined by 2 raised to power (position – 1), that is the value of a bit 1 at position 6 is $2^{(6-1)}$ that is 2^5 that is 32. The total value of the octet is determined by adding up the positional value of bits. The value of 11000000 is $128+64 = 192$.

IPv4 - Address Classes

- ▶ Internet Protocol hierarchy contains several classes of IP Addresses to be used efficiently in various situations as per the requirement of hosts per network.
- ▶ Broadly, the IPv4 Addressing system is divided into five classes of IP Addresses. All the five classes are identified by the first octet of IP Address
- ▶ The number of networks and the number of hosts per class can be derived by this formula:

$$\text{number of networks} = 2^{\text{network_bits}}$$

$$\text{number of hosts per networks} = 2^{\text{host_bits}} - 2$$

- ▶ When calculating hosts' IP addresses, 2 IP addresses are decreased because they cannot be assigned to hosts, i.e. the first IP of a network is network number and the last IP is reserved for Broadcast IP.

Class A Address

- ▶ The first bit of the first octet is always set to 0 (zero). Thus the first octet ranges from 1 – 127.
- ▶ Class A addresses only include IP starting from 1.x.x.x to 126.x.x.x only. The IP range 127.x.x.x is reserved for loopback IP addresses.
- ▶ The default subnet mask for Class A IP address is 255.0.0.0 which implies that Class A addressing can have 126 networks (2^7 -2) and 16777214 hosts (2^{24} -2).
- ▶ Class A IP address format is thus: **0NNNNNNN.HHHHHHHH.HHHHHHHH.HHHHHHHH**

Class B Address

- ▶ An IP address which belongs to class B has the first two bits in the first octet set to 10.
- ▶ Class B IP Addresses range from 128.0.x.x to 191.255.x.x. The default subnet mask for Class B is 255.255.x.x.
- ▶ Class B has 16384 (2^{14}) Network addresses and 65534 ($2^{16}-2$) Host addresses.
- ▶ Class B IP address format
is: **10NNNNNN.NNNNNNNN.HHHHHHHH.HHHHHHHH**

Class C Address

- ▶ The first octet of Class C IP address has its first 3 bits set to 110.
- ▶ Class C IP addresses range from 192.0.0.x to 223.255.255.x. The default subnet mask for Class C is 255.255.255.x.
- ▶ Class C gives $2097152 (2^{21})$ Network addresses and $254 (2^8 - 2)$ Host addresses.
- ▶ Class C IP address format
is: **110NNNN.NNNNNNNN.NNNNNNNN.HHHHHHHH**

Class D Address

- ▶ Very first four bits of the first octet in Class D IP addresses are set to 1110.
- ▶ Class D has IP address range from 224.0.0.0 to 239.255.255.255. Class D is reserved for Multicasting. In multicasting data is not destined for a particular host, that is why there is no need to extract host address from the IP address, and Class D does not have any subnet mask.

Class E Address

- ▶ This IP Class is reserved for experimental purposes only for R&D or Study. IP addresses in this class ranges from 240.0.0.0 to 255.255.255.254. Like Class D, this class too is not equipped with any subnet mask.

IPv4 - Subnetting

- ▶ Each IP class is equipped with its own default subnet mask which bounds that IP class to have prefixed number of Networks and prefixed number of Hosts per network. Classful IP addressing does not provide any flexibility of having less number of Hosts per Network or more Networks per IP Class.
- ▶ CIDR or **Classless Inter Domain Routing** provides the flexibility of borrowing bits of Host part of the IP address and using them as Network in Network, called Subnet. By using subnetting, one single Class A IP address can be used to have smaller sub-networks which provides better network management capabilities.

Class A Subnets

- ▶ In Class A, only the first octet is used as Network identifier and rest of three octets are used to be assigned to Hosts (i.e. 16777214 Hosts per Network). To make more subnet in Class A, bits from Host part are borrowed and the subnet mask is changed accordingly.
- ▶ For example, if one MSB (Most Significant Bit) is borrowed from host bits of second octet and added to Network address, it creates two Subnets ($2^1=2$) with ($2^{23}-2$) 8388606 Hosts per Subnet.
- ▶ The Subnet mask is changed accordingly to reflect subnetting.
- ▶ In case of subnetting too, the very first and last IP address of every subnet is used for Subnet Number and Subnet Broadcast IP address respectively. Because these two IP addresses cannot be assigned to hosts, sub-netting cannot be implemented by using more than 30 bits as Network Bits, which provides less than two hosts per subnet.

Class B Subnets

- ▶ By default, using Classful Networking, 14 bits are used as Network bits providing (2^{14}) 16384 Networks and ($2^{16}-2$) 65534 Hosts.
- ▶ Class B IP Addresses can be subnetted the same way as Class A addresses, by borrowing bits from Host bits.

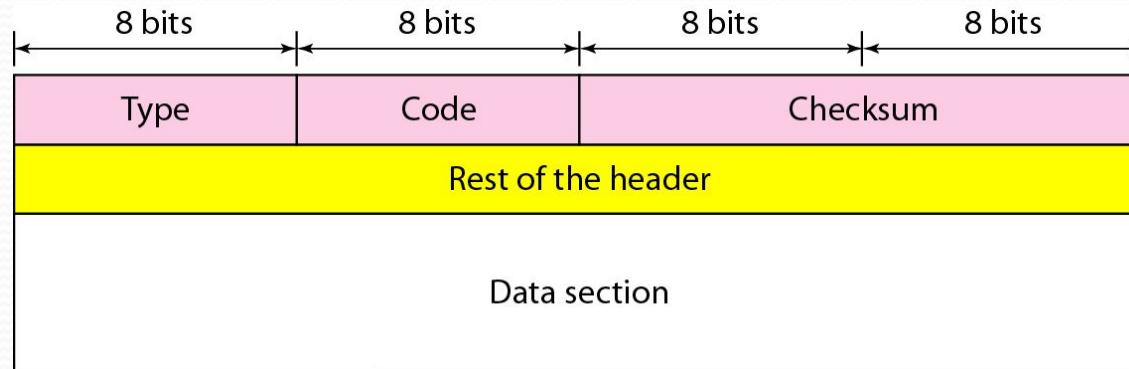
Class C Subnets

- ▶ Class C IP addresses are normally assigned to a very small size network because it can only have 254 hosts in a network. Given below is a list of all possible combination of subnetted Class B IP address:

MODULE -5

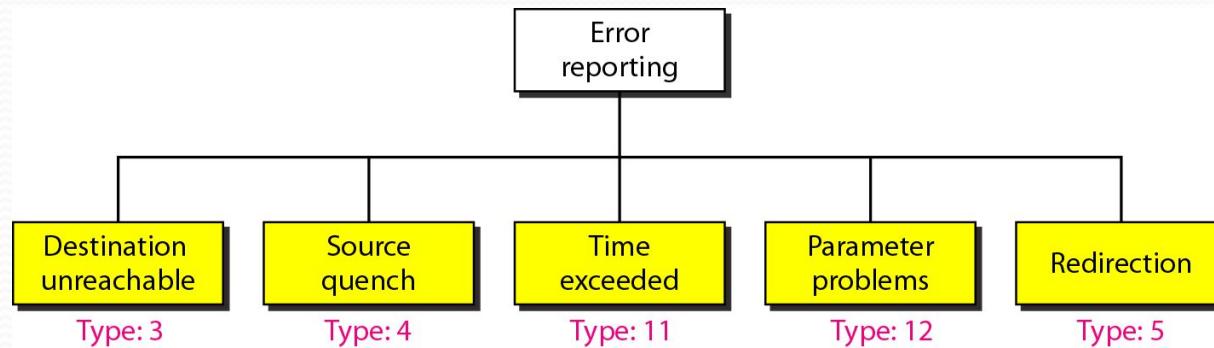
INTERNET CONTROL PROTOCOLS

- ICMP
- 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.
- General format of ICMP messages

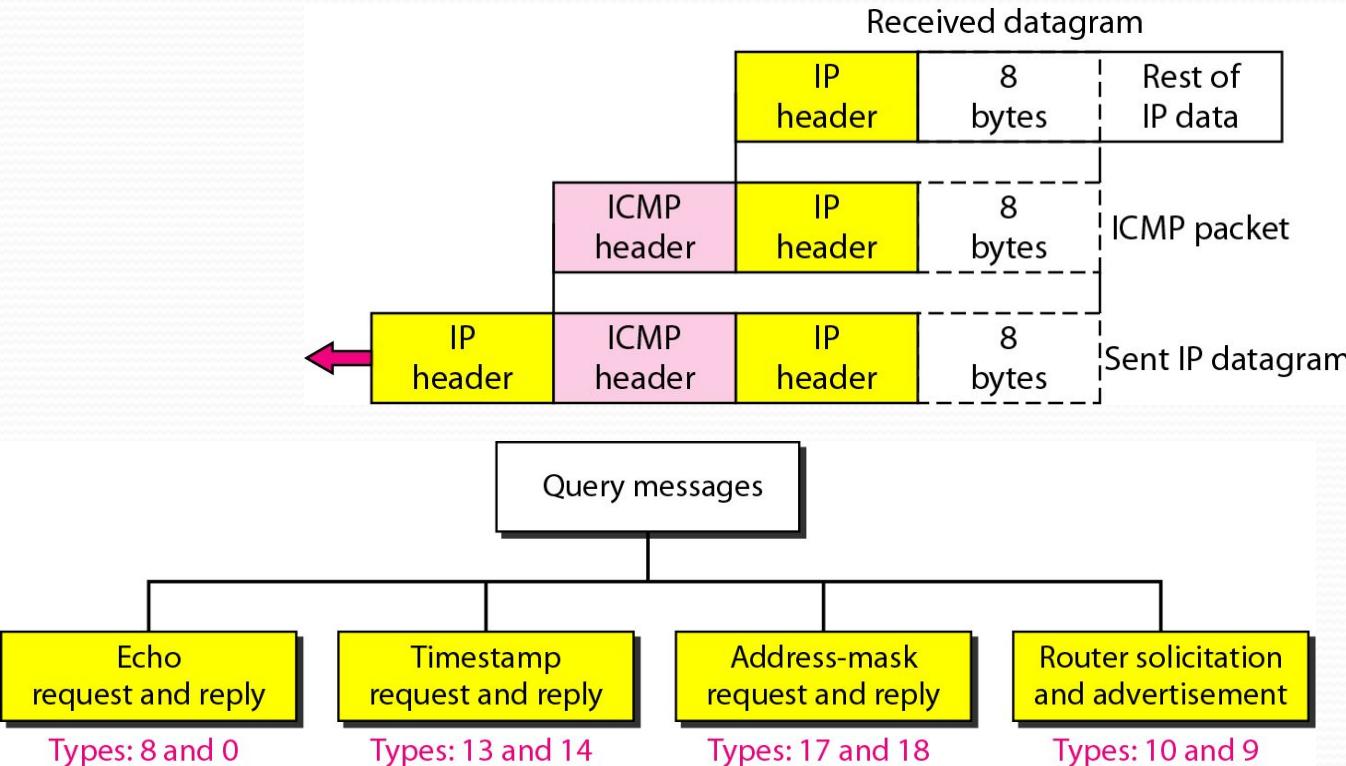


Error-reporting messages

- Important points about ICMP error messages:
- No ICMP error message will be generated in response to a datagram carrying an ICMP error message.
- No ICMP error message will be generated for a fragmented datagram that is not the first fragment.
- No ICMP error message will be generated for a datagram having a multicast address.
- No ICMP error message will be generated for a datagram having a special address such as 27.0.0.0 or 0.0.0.0.



• Contents of data field for the error messages

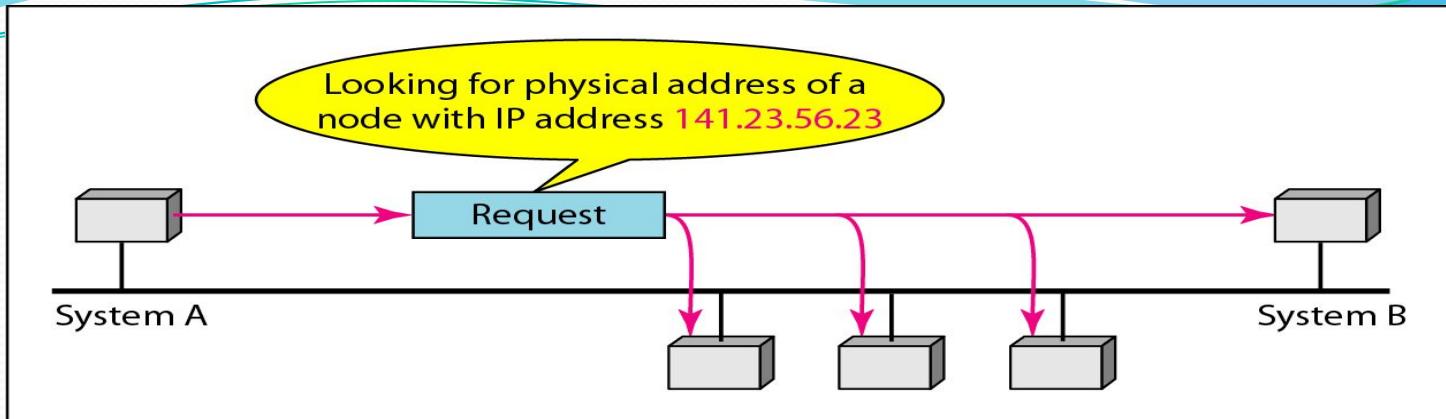


- Encapsulation of ICMP query messages

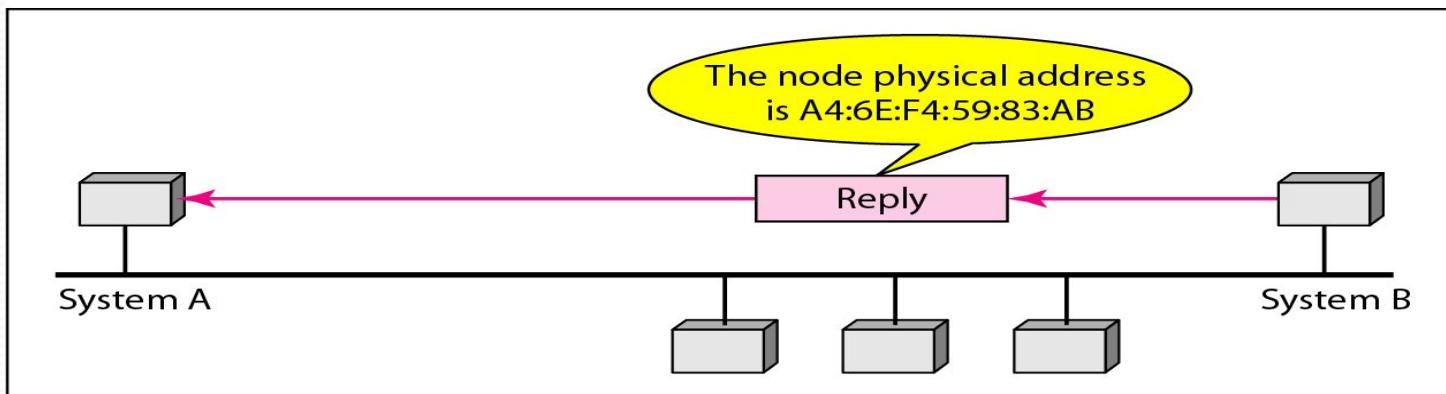


ARP

- The Address Resolution Protocol (ARP) is a communication protocol used for discovering the link layer address, such as a MAC address, associated with a given network layer address, typically an IPv4 address. This mapping is a critical function in the Internet protocol suite.
- ARP can be useful if the ARP reply is cached (kept in cache memory for a while).
- An ARP request is broadcast whereas an ARP reply is unicast

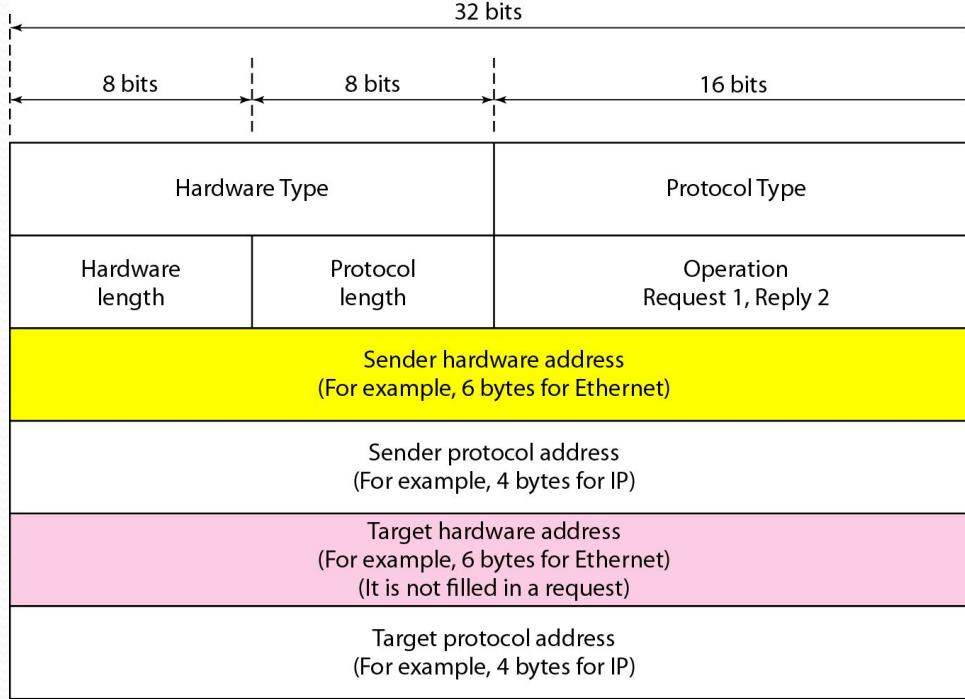


a. ARP request is broadcast

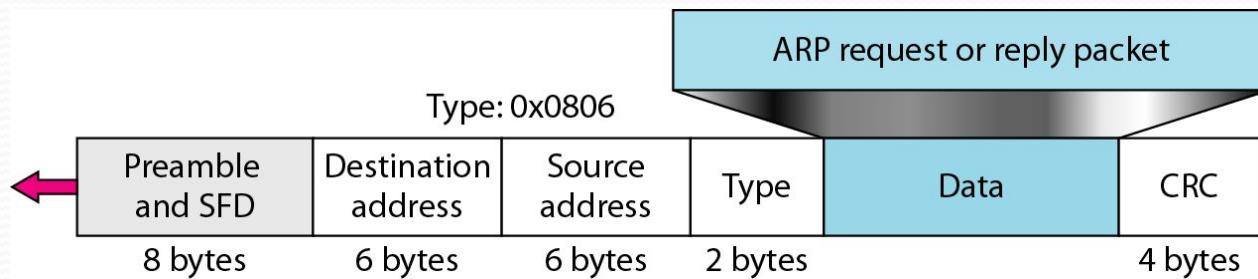


b. ARP reply is unicast

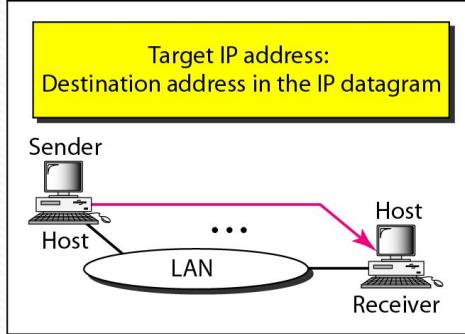
ARP packet



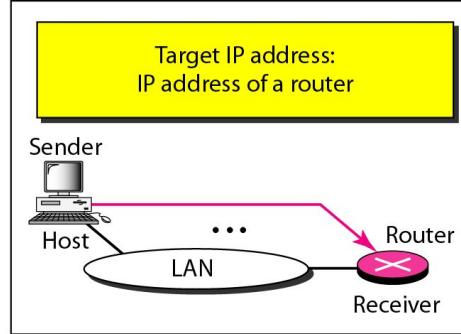
● Encapsulation of ARP packet



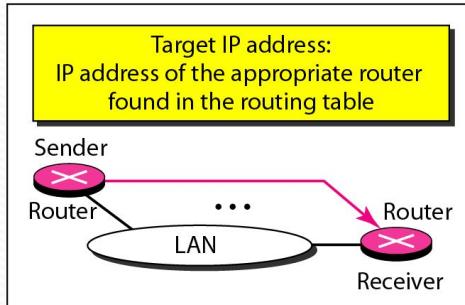
Four cases using ARP



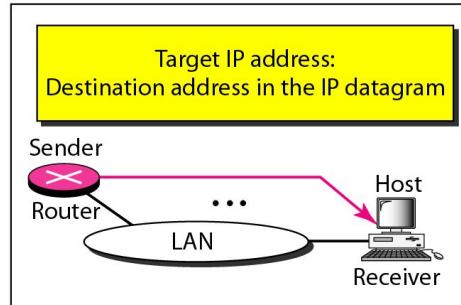
Case 1. A host has a packet to send to another host on the same network.



Case 2. A host wants to send a packet to another host on another network. It must first be delivered to a router.



Case 3. A router receives a packet to be sent to a host on another network. It must first be delivered to the appropriate router.



Case 4. A router receives a packet to be sent to a host on the same network.

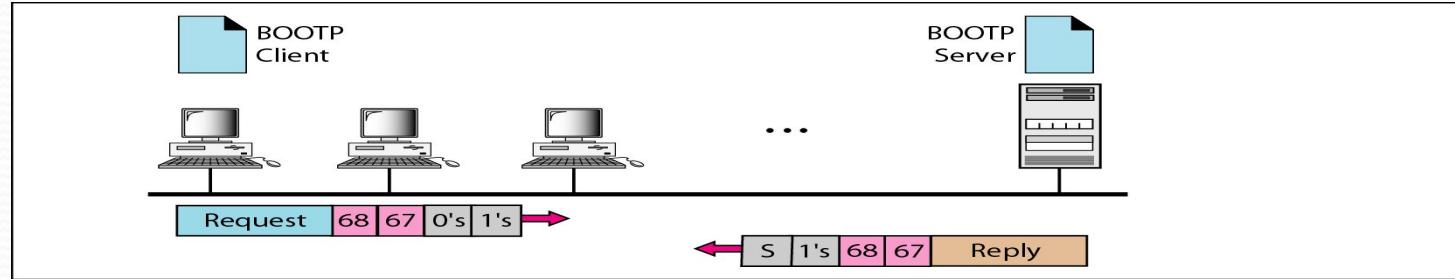
RARP

- Given LAN address, gives the IP address
- A RARP message is created and broadcast on the local network.
- The machine on the local network that knows the logical address will respond with a RARP reply.
- Broadcasting is done at data link layer.
- Broadcast requests does not pass the boundaries of a network.
- Usually for booting diskless workstation.
 - Gets the OS image from remote file server.
 - Same image for all machines.
 - Machine broadcasts its LAN address.
 - Remote RARP server responds with machine's IP address.

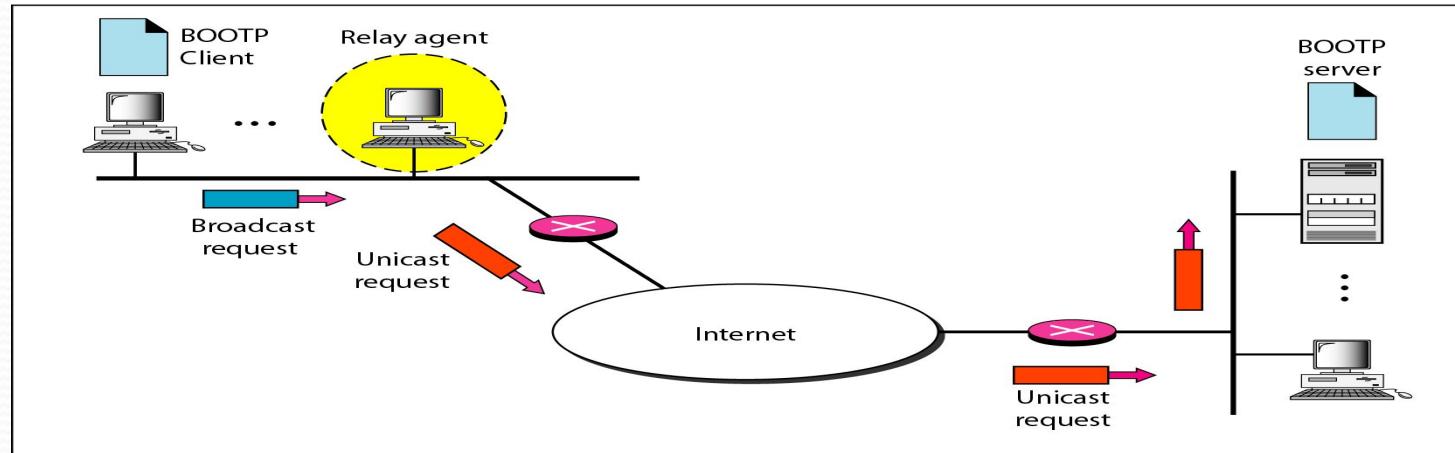
BOOTP

- RARP broadcasts are not forwarded by routers.
- Need RARP server on every network.
- BOOTP uses UDP messages that are forwarded by routers.
 - Also provides additional information such as IP address of file server holding OS image, subnet mask, etc.

BOOTP client and server on the same and different networks



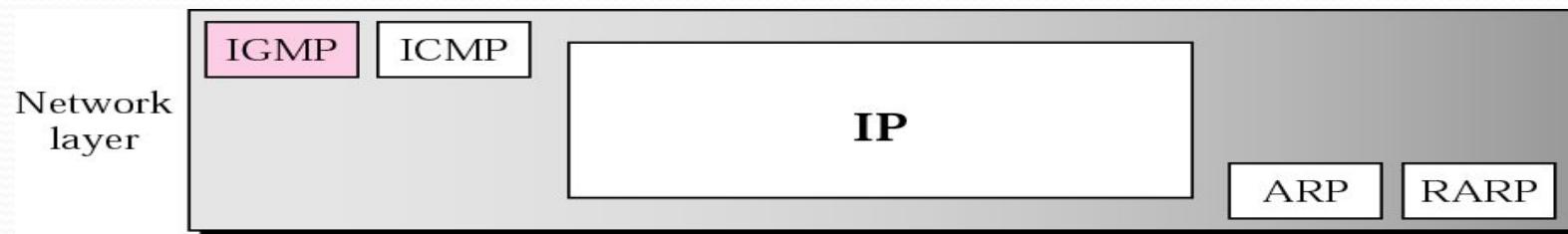
a. Client and server on the same network



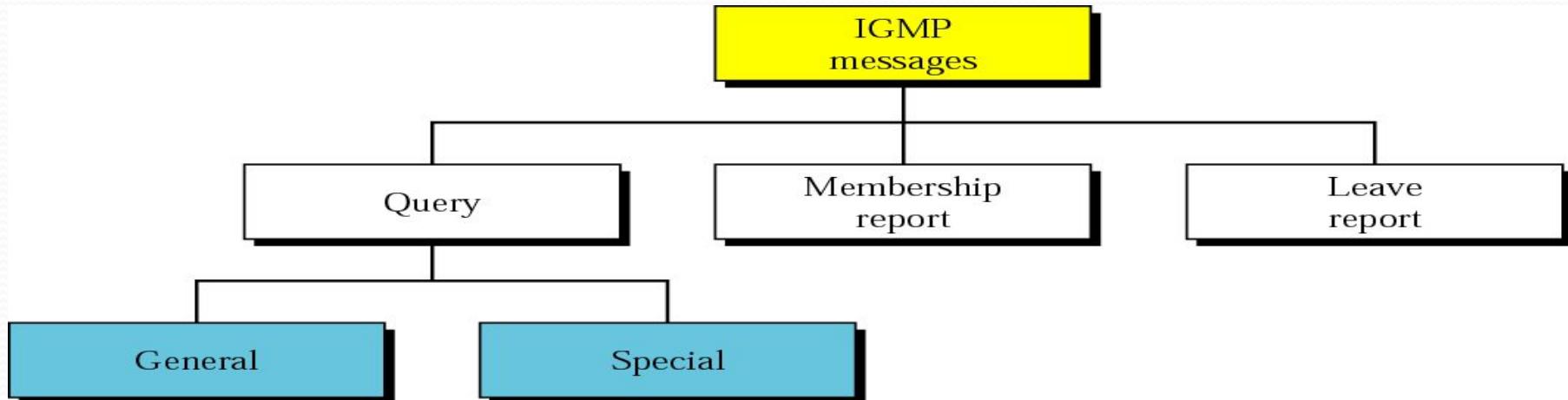
b. Client and server on different networks

Internet Group Management Protocol

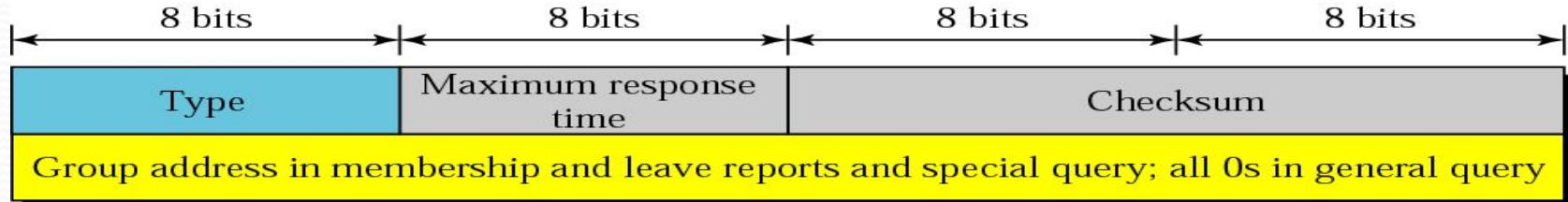
- IGMP is a protocol that manages group membership. The IGMP protocol gives the multicast routers information about the membership status of hosts (routers) connected to the network. .



- IGMP has three types of messages: the query, the membership report, and the leave report. There are two types of query messages, general and special



IGMP message format

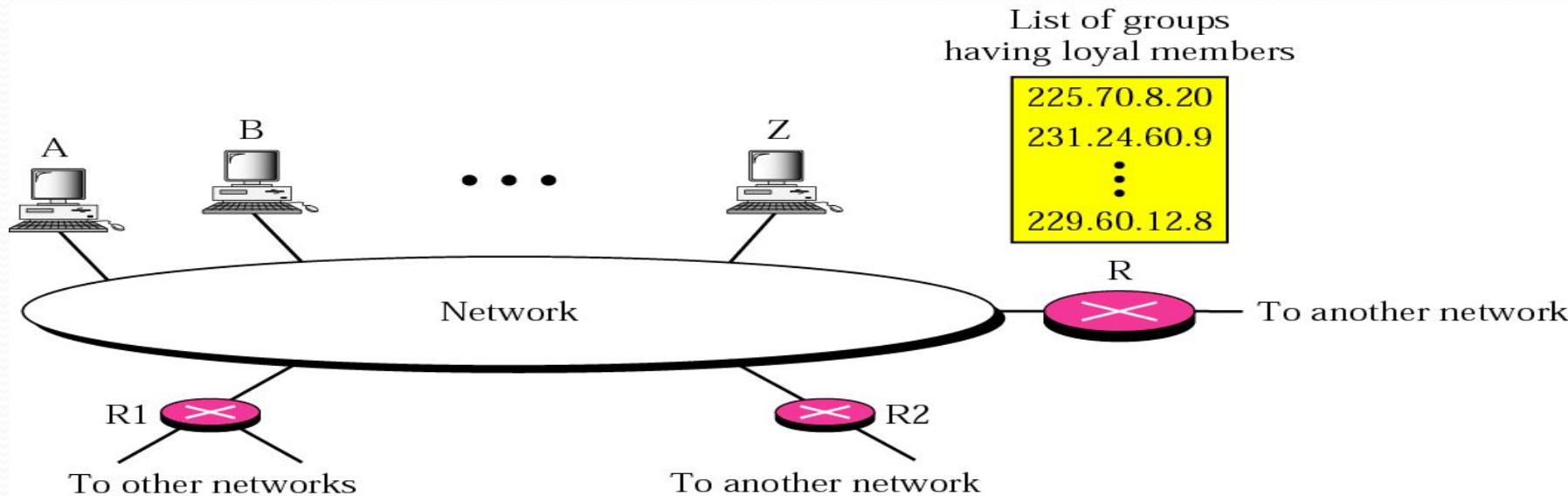


IGMP type field

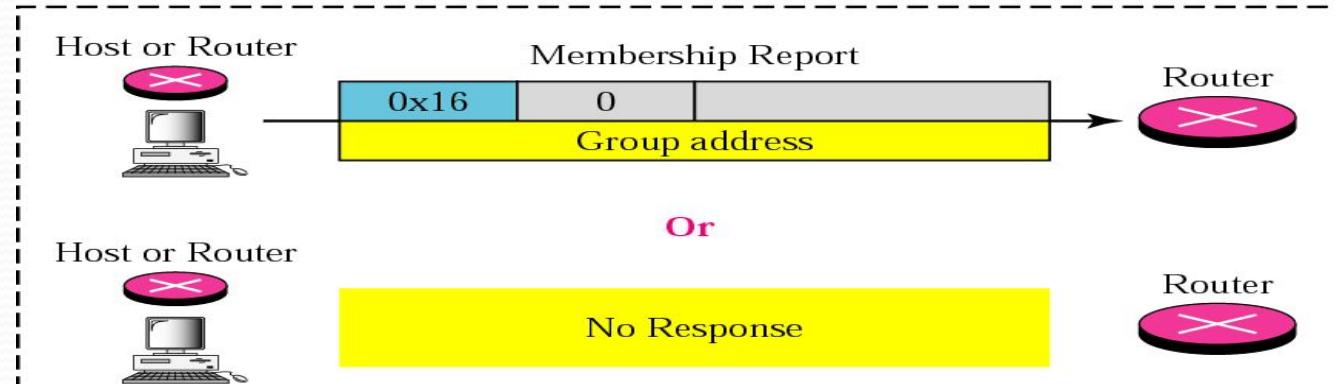
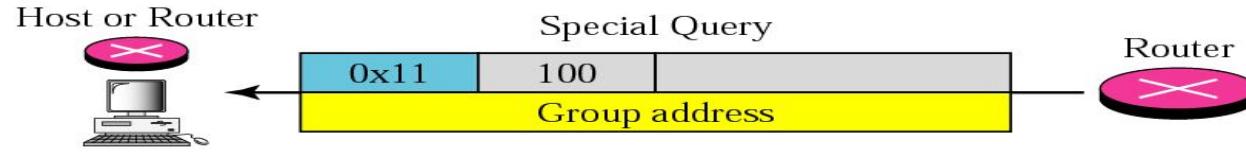
Type	Value
General or Special Query	0x11 or 00010001
Membership Report	0x16 or 00010110
Leave Report	0x17 or 00010111

IGMP OPERATION

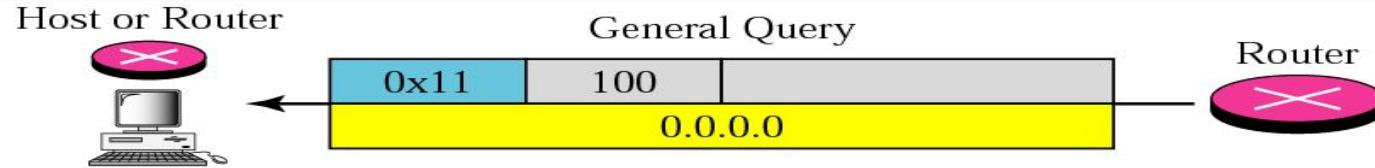
- A multicast router connected to a network has a list of multicast addresses of the groups with at least one loyal member in that network. For each group, there is one router that has the duty of distributing the multicast packets destined for that group.



- In IGMP, a membership report is sent twice, one after the other



- The general query message does not define a particular group

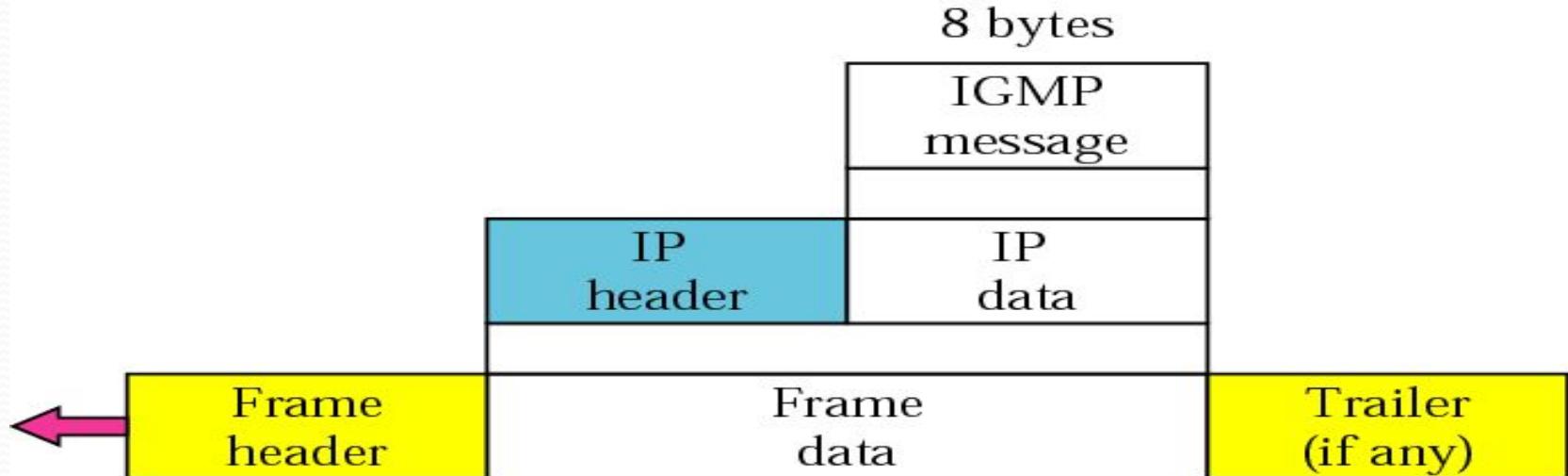


Or



ENCAPSULATION

- The IGMP message is encapsulated in an IP datagram, which is itself encapsulated in a frame.



EXTERIOR ROUTING PROTOCOLS

- EXTERIOR ROUTING PROTOCOLS ARE USED TO EXCHANGE ROUTING INFORMATION BETWEEN AUTONOMOUS SYSTEMS. THE ROUTING INFORMATION PASSED BETWEEN AUTONOMOUS SYSTEMS IS CALLED REACHABILITY INFORMATION. REACHABILITY INFORMATION IS SIMPLY INFORMATION ABOUT WHICH NETWORKS CAN BE REACHED THROUGH A SPECIFIC AUTONOMOUS SYSTEM.
- RFC 1771 DEFINES BORDER GATEWAY PROTOCOL (BGP), THE LEADING EXTERIOR ROUTING PROTOCOL, AND PROVIDES THE FOLLOWING DESCRIPTION OF THE ROUTING FUNCTION OF AN AUTONOMOUS SYSTEM:

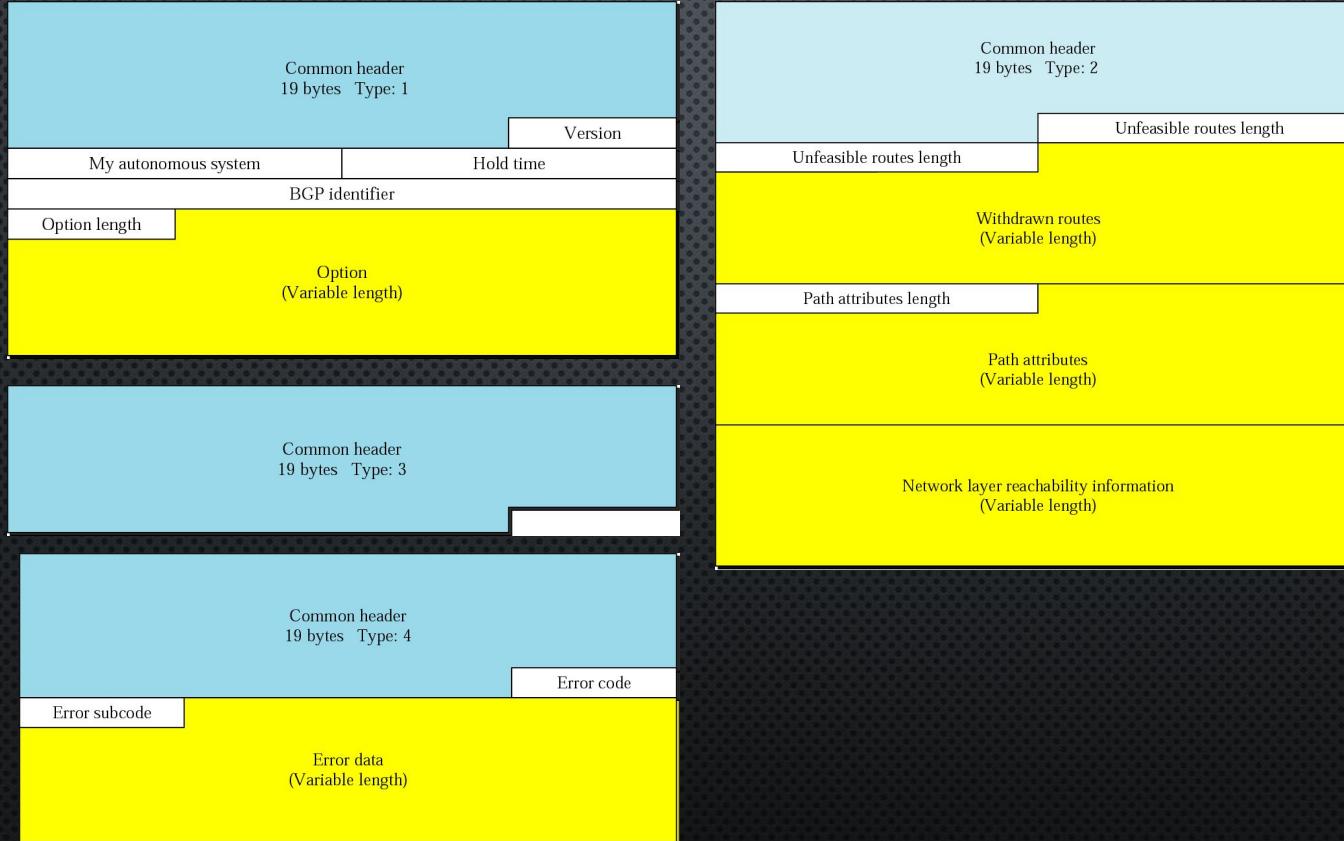
BORDER GATEWAY PROTOCOL (RFC 1771)

- BASED ON THE PATH VECTOR ROUTING.
- DISTANCE-VECTOR PROTOCOL NOT PREFERRED FOR INTER-AS ROUTING (EXTERIOR ROUTING PROTOCOL)
 - ASSUMES ALL ROUTERS HAVE A COMMON DISTANCE METRICS TO JUDGE ROUTE PREFERENCES.
 - IF ROUTERS HAVE DIFFERENT MEANINGS OF A METRIC, IT MAY NOT BE POSSIBLE TO CREATE STABLE, LOOP FREE ROUTES.
 - A GIVEN AS MAY HAVE DIFFERENT PRIORITIES FROM ANOTHER AS.
 - GIVES NO INFORMATION ABOUT THE ASs THAT WILL BE VISITED.
- LINK-STATE ROUTING PROTOCOL
 - DIFFERENT METRICS.
 - FLOODING IS NOT REALISTIC.

- LINK-STATE ROUTING PROTOCOL
 - DIFFERENT METRICS.
 - FLOODING IS NOT REALISTIC.
- PATH VECTOR ROUTING
 - NO METRICS,
 - INFORMATION ABOUT WHICH NETWORKS CAN BE REACHED BY A GIVEN ROUTER AND ASSES TO BE CROSSED.
- DIFFERS FROM DVA
 - PATH VECTOR APPROACH DOES NOT INCLUDE A DISTANCE OR COST ESTIMATE
 - LISTS ALL OF THE ASSES VISITED TO REACH DESTINATION NETWORK

- LOOP PREVENTION IN BGP:
 - CHECKS THE PATH BEFORE UPDATING ITS DATABASE. (IF ITS AS IS IN THE PATH IGNORE THE MESSAGE)
- POLICY ROUTING:
 - IF A PATH CONSIST OF AN AS AGAINST THE POLICY OF THE CURRENT AS, MESSAGE DISCARDED.

BGP MESSAGE FORMAT



IPV6 - HEADERS

- THE WONDER OF IPv6 LIES IN ITS HEADER. AN IPv6 ADDRESS IS 4 TIMES LARGER THAN IPv4, BUT SURPRISINGLY, THE HEADER OF AN IPv6 ADDRESS IS ONLY 2 TIMES LARGER THAN THAT OF IPv4. IPv6 HEADERS HAVE ONE FIXED HEADER AND ZERO OR MORE OPTIONAL (EXTENSION) HEADERS.
- ALL THE NECESSARY INFORMATION THAT IS ESSENTIAL FOR A ROUTER IS KEPT IN THE FIXED HEADER. THE EXTENSION HEADER CONTAINS OPTIONAL INFORMATION THAT HELPS ROUTERS TO UNDERSTAND HOW TO HANDLE A PACKET/FLOW.

FIXED HEADER

	4-11	12-31	
0-3	Version	Traffic Class	Flow Label
32-47	Payload Length	Next Header	Hop Limit
64-191	Source Address		
192-288	Destination Address		

- **VERSION** (4-BITS): IT REPRESENTS THE VERSION OF INTERNET PROTOCOL, I.E. 0110.
- **TRAFFIC CLASS** (8-BITS): THESE 8 BITS ARE DIVIDED INTO TWO PARTS. THE MOST SIGNIFICANT 6 BITS ARE USED FOR TYPE OF SERVICE TO LET THE ROUTER KNOWN WHAT SERVICES SHOULD BE PROVIDED TO THIS PACKET. THE LEAST SIGNIFICANT 2 BITS ARE USED FOR EXPLICIT CONGESTION NOTIFICATION (ECN).
- **FLOW LABEL** (20-BITS): THIS LABEL IS USED TO MAINTAIN THE SEQUENTIAL FLOW OF THE PACKETS BELONGING TO A COMMUNICATION. THE SOURCE LABELS THE SEQUENCE TO HELP THE ROUTER IDENTIFY THAT A PARTICULAR PACKET BELONGS TO A SPECIFIC FLOW OF INFORMATION. THIS FIELD HELPS AVOID RE-ORDERING OF DATA PACKETS. IT IS DESIGNED FOR STREAMING/REAL-TIME MEDIA.
- **PAYLOAD LENGTH** (16-BITS): THIS FIELD IS USED TO TELL THE ROUTERS HOW MUCH INFORMATION A PARTICULAR PACKET CONTAINS IN ITS PAYLOAD. PAYLOAD IS COMPOSED OF EXTENSION HEADERS AND UPPER LAYER DATA. WITH 16 BITS, UP TO 65535 BYTES CAN BE INDICATED; BUT IF THE EXTENSION HEADERS CONTAIN HOP-BY-HOP EXTENSION HEADER, THEN THE PAYLOAD MAY EXCEED 65535 BYTES AND THIS FIELD IS SET TO 0.

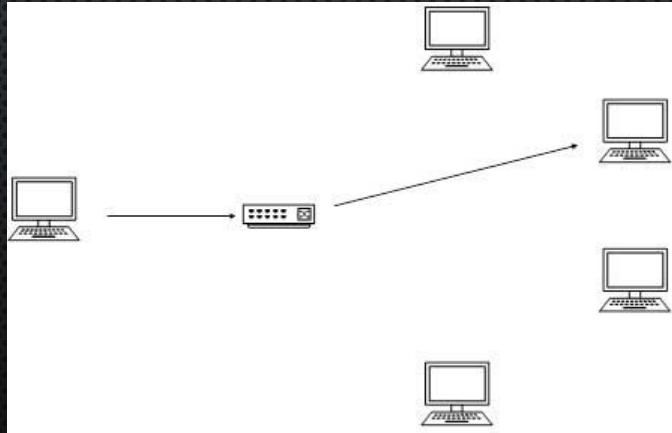
- **NEXT HEADER** (8-BITS): THIS FIELD IS USED TO INDICATE EITHER THE TYPE OF EXTENSION HEADER, OR IF THE EXTENSION HEADER IS NOT PRESENT THEN IT INDICATES THE UPPER LAYER PDU. THE VALUES FOR THE TYPE OF UPPER LAYER PDU ARE SAME AS IPv4's.
- **HOP LIMIT** (8-BITS): THIS FIELD IS USED TO STOP PACKET TO LOOP IN THE NETWORK INFINITELY. THIS IS SAME AS TTL IN IPv4. THE VALUE OF HOP LIMIT FIELD IS DECREMENTED BY 1 AS IT PASSES A LINK (ROUTER/HOP). WHEN THE FIELD REACHES 0 THE PACKET IS DISCARDED.
- **SOURCE ADDRESS** (128-BITS): THIS FIELD INDICATES THE ADDRESS OF ORIGINATOR OF THE PACKET.
- **DESTINATION ADDRESS** (128-BITS): THIS FIELD PROVIDES THE ADDRESS OF INTENDED RECIPIENT OF THE PACKET.

IPV6 - ADDRESSING MODES

- IN COMPUTER NETWORKING, ADDRESSING MODE REFERS TO THE MECHANISM OF HOSTING AN ADDRESS ON THE NETWORK. IPv6 OFFERS SEVERAL TYPES OF MODES BY WHICH A SINGLE HOST CAN BE ADDRESSED. MORE THAN ONE HOST CAN BE ADDRESSED AT ONCE OR THE HOST AT THE CLOSEST DISTANCE CAN BE ADDRESSED.

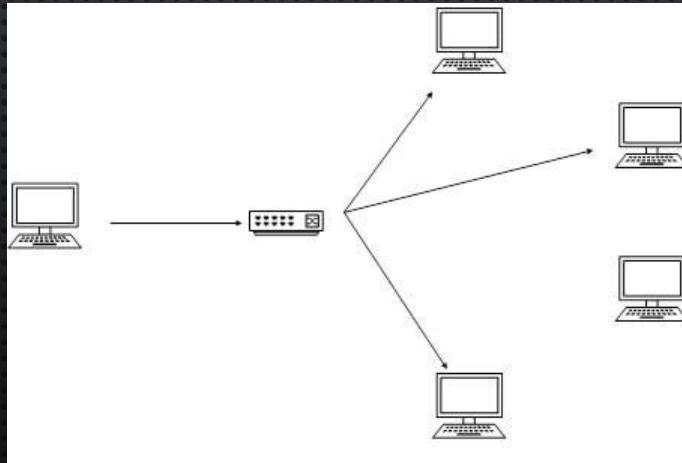
UNICAST

- IN UNICAST MODE OF ADDRESSING, AN IPv6 INTERFACE (HOST) IS UNIQUELY IDENTIFIED IN A NETWORK SEGMENT. THE IPv6 PACKET CONTAINS BOTH SOURCE AND DESTINATION IP ADDRESSES. A HOST INTERFACE IS EQUIPPED WITH AN IP ADDRESS WHICH IS UNIQUE IN THAT NETWORK SEGMENT.
- WHEN A NETWORK SWITCH OR A ROUTER RECEIVES A UNICAST IP PACKET, DESTINED TO A SINGLE HOST, IT SENDS OUT ONE OF ITS OUTGOING INTERFACE WHICH CONNECTS TO THAT PARTICULAR HOST.



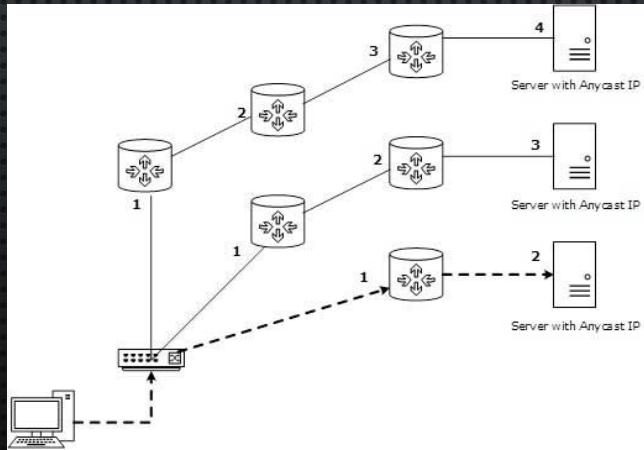
MULTICAST

- THE IPv6 MULTICAST MODE IS SAME AS THAT OF IPv4. THE PACKET DESTINED TO MULTIPLE HOSTS IS SENT ON A SPECIAL MULTICAST ADDRESS. ALL THE HOSTS INTERESTED IN THAT MULTICAST INFORMATION, NEED TO JOIN THAT MULTICAST GROUP FIRST.
- ALL THE INTERFACES THAT JOINED THE GROUP RECEIVE THE MULTICAST PACKET AND PROCESS IT, WHILE OTHER HOSTS NOT INTERESTED IN MULTICAST PACKETS IGNORE THE MULTICAST INFORMATION.



ANYCAST

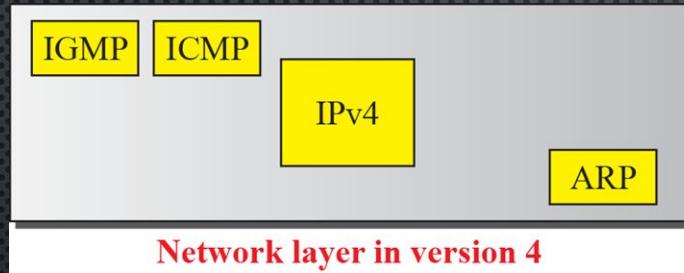
- IPv6 HAS INTRODUCED A NEW TYPE OF ADDRESSING, WHICH IS CALLED ANYCAST ADDRESSING. IN THIS ADDRESSING MODE, MULTIPLE INTERFACES (HOSTS) ARE ASSIGNED SAME ANYCAST IP ADDRESS. WHEN A HOST WISHES TO COMMUNICATE WITH A HOST EQUIPPED WITH AN ANYCAST IP ADDRESS, IT SENDS A UNICAST MESSAGE.
- WITH THE HELP OF COMPLEX ROUTING MECHANISM, THAT UNICAST MESSAGE IS DELIVERED TO THE HOST CLOSEST TO THE SENDER IN TERMS OF ROUTING COST.



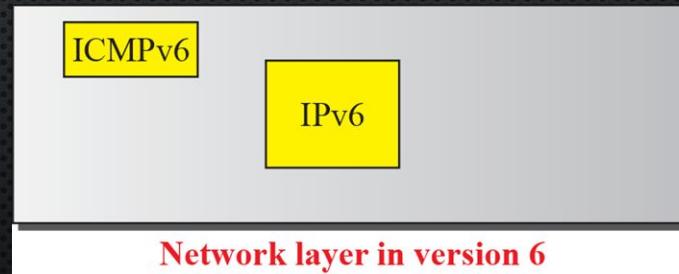
ICMP VERSION 6

- ANOTHER PROTOCOL THAT HAS BEEN MODIFIED IN VERSION 6 OF THE TCP/IP PROTOCOL SUITE IS ICMP. THIS NEW VERSION, INTERNET CONTROL MESSAGE PROTOCOL VERSION 6 (ICMPv6), FOLLOWS THE SAME STRATEGY AND PURPOSES OF VERSION 4. ICMPv6, HOWEVER, IS MORE COMPLICATED THAN ICMPv4: SOME PROTOCOLS THAT WERE INDEPENDENT IN VERSION 4 ARE NOW PART OF ICMPv6 AND SOME NEW MESSAGES HAVE BEEN ADDED TO MAKE IT MORE USEFUL.

COMPARISON OF NETWORK LAYERS IN VERSION 4 AND VERSION 6

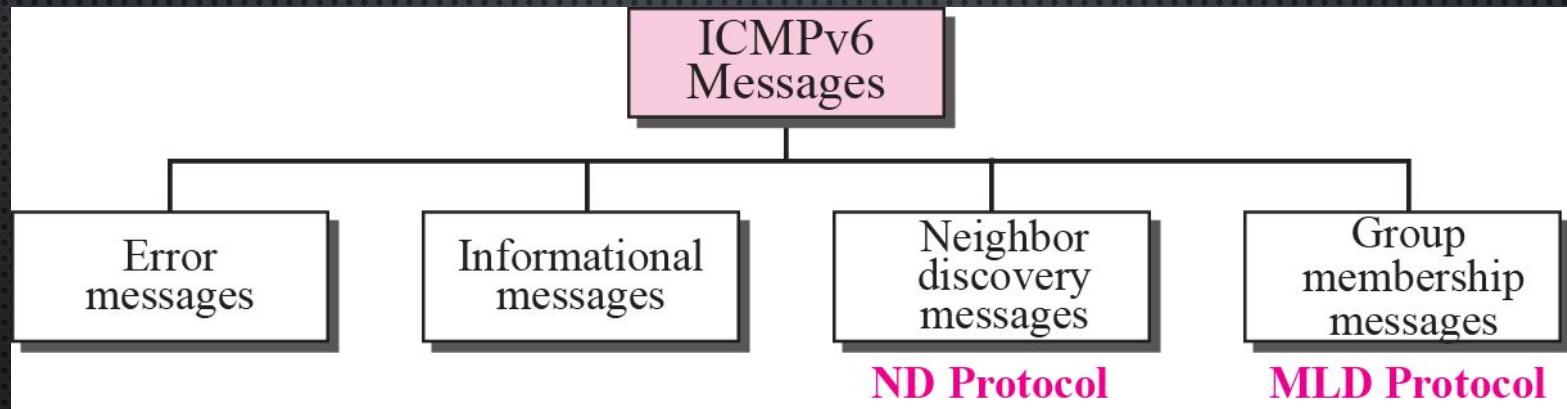


Network layer in version 4



Network layer in version 6

TAXONOMY OF ICMPV6 MESSAGES



28-1 INTRODUCTION

Another protocol that has been modified in version 6 of the TCP/IP protocol suite is ICMP. This new version, Internet Control Message Protocol version 6 (ICMPv6), follows the same strategy and purposes of version 4. ICMPv6, however, is more complicated than ICMPv4: some protocols that were independent in version 4 are now part of ICMPv6 and some new messages have been added to make it more useful.

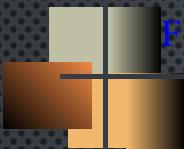
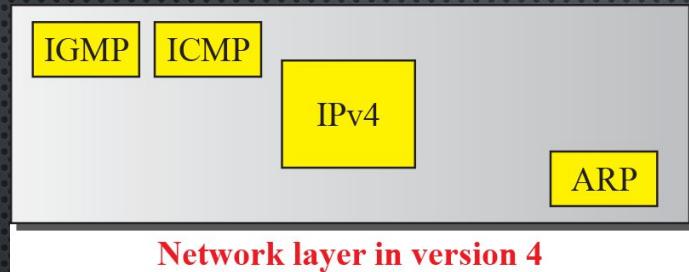
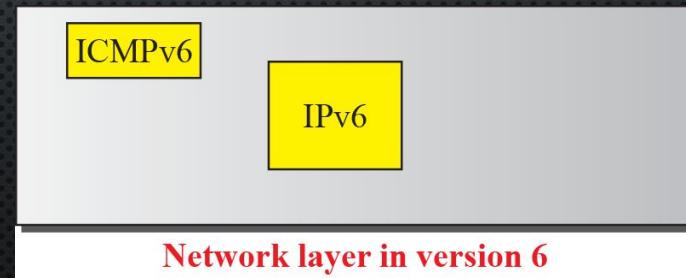


Figure 28.1 Comparison of network layers in version 4 and version 6



Network layer in version 4



Network layer in version 6

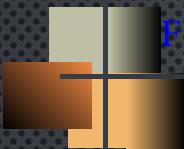
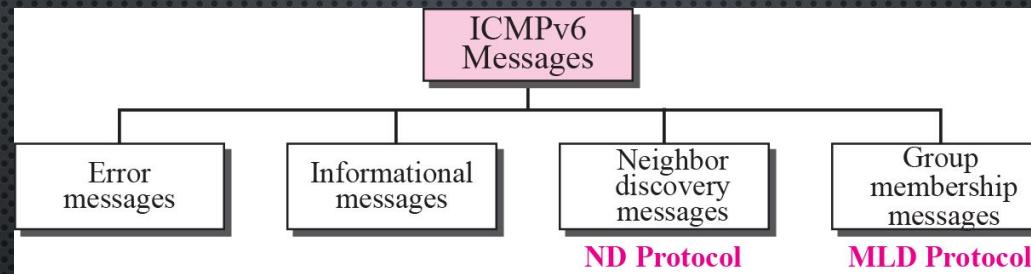


Figure 28.2 *Taxonomy of ICMPv6 messages*



28-2 ERROR MESSAGES

As we saw in our discussion of version 4, one of the main responsibilities of ICMP is to report errors. Four types of errors are handled: destination unreachable, packet too big, time exceeded, and parameter problems (see Figure 28.3).

Topics Discussed in the Section

- ✓ Destination-Unreachable Message
- ✓ Packet-Too-Big Message
- ✓ Time-Exceeded Message
- ✓ Parameter-Problem Message

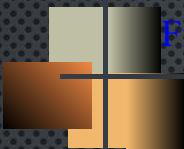
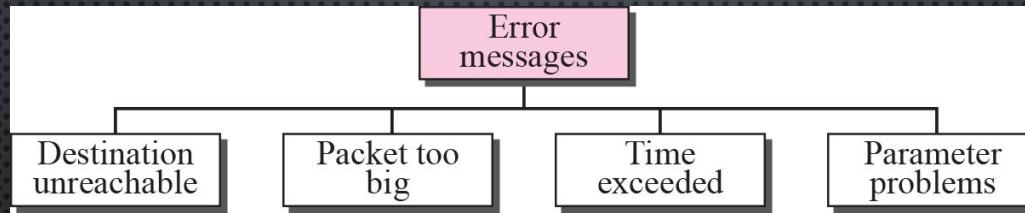


Figure 28.3 *Error-reporting messages*



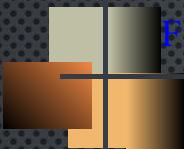
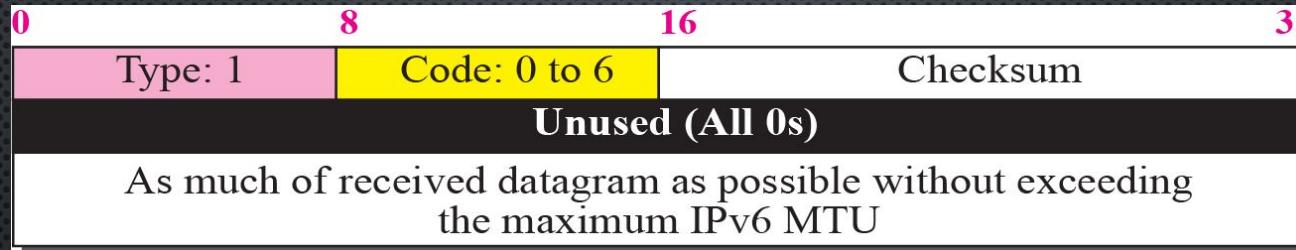


Figure 28.4 Destination unreachable message



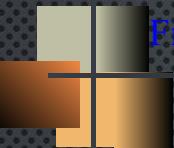
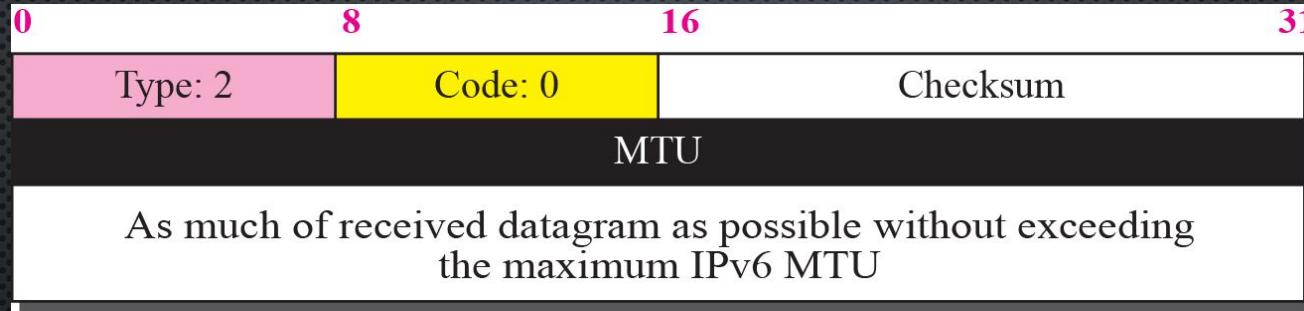


Figure 28.5 *Packet-too-bit message*



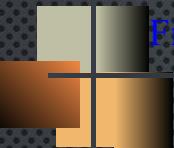
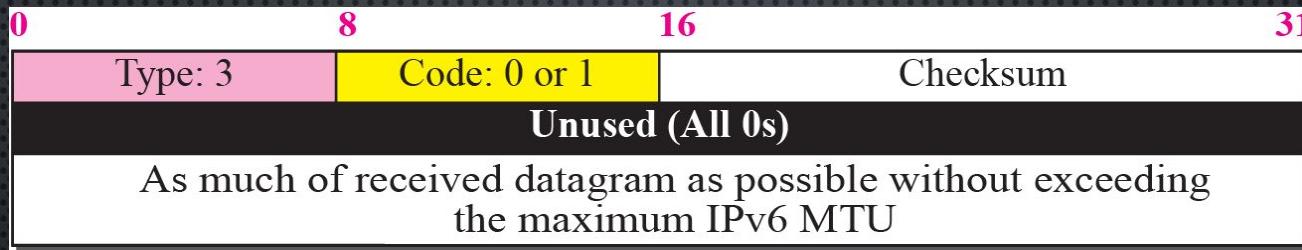


Figure 28.6 Time-exceeded message



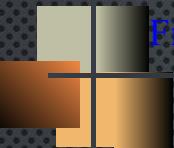
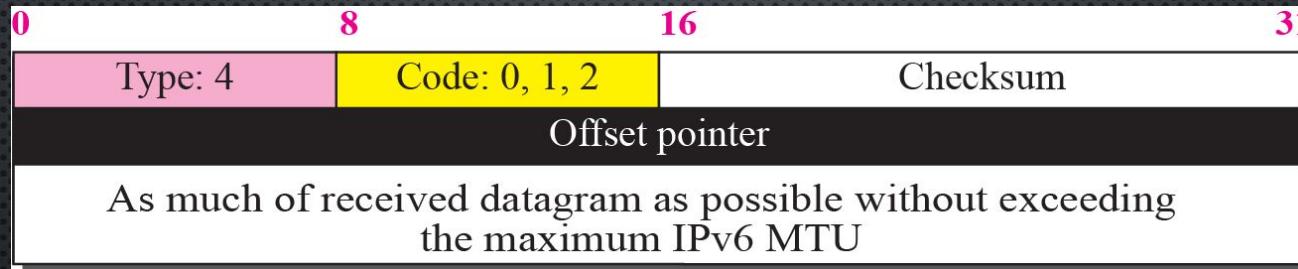


Figure 28.7 Parameter-problem message



28-3 INFORMATIONAL MESSAGES

Two of the ICMPv6 messages can be categorized as informational messages: echo request and echo reply messages. As discussed in Chapter 9, the echo request and echo response messages are designed to check if two devices in the Internet can communicate with each other. A host or router can send an echo request message to another host; the receiving computer or router can reply using the echo response message.

Topics Discussed in the Section

- ✓ Echo-Request Message
- ✓ Echo-Reply Message

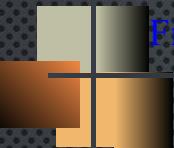
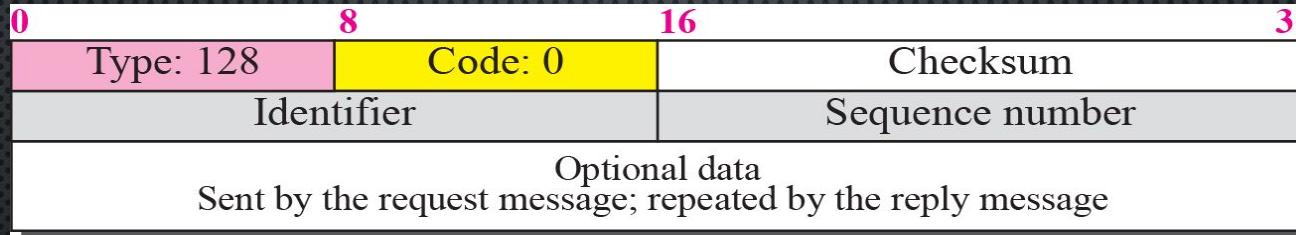


Figure 28.8 Echo-request message



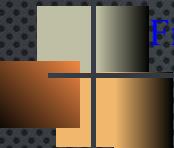


Figure 28.9 Echo-reply message

0	8	16	31			
Type: 129	Code: 0	Checksum				
Identifier	Sequence number					
Optional data Sent by the request message; repeated by the reply message						

28-4 NEIGHBOR-DISCOVERY MESSAGES

Several messages in the ICMPv6 have been redefined in ICMPv6 to handle the issue of neighbor discovery. Some new messages have also been added to provide extension. The most important issue is the definition of two new protocols that clearly define the functionality of these group messages: the Neighbor-Discovery (ND) protocol and the Inverse-Neighbor-Discovery (IND) protocol. These two protocols are used by nodes (hosts or routers) on the same link (network).

Topics Discussed in the Section

- ✓ Router-Solicitation Message
- ✓ Router-Advertisement Message
- ✓ Neighbor-Solicitation Message
- ✓ Neighbor-Advertisement Message
- ✓ Redirection Message
- ✓ Inverse-Neighbor-Solicitation Message
- ✓ Inverse-Neighbor-Advertisement Message

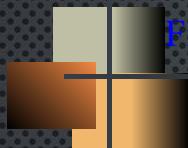
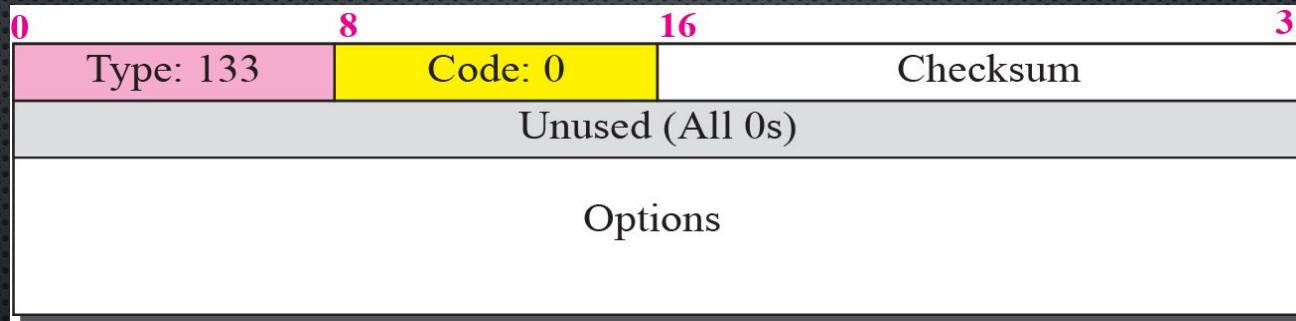


Figure 28.10 Router-solicitation message



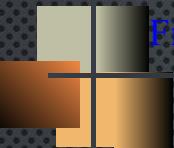
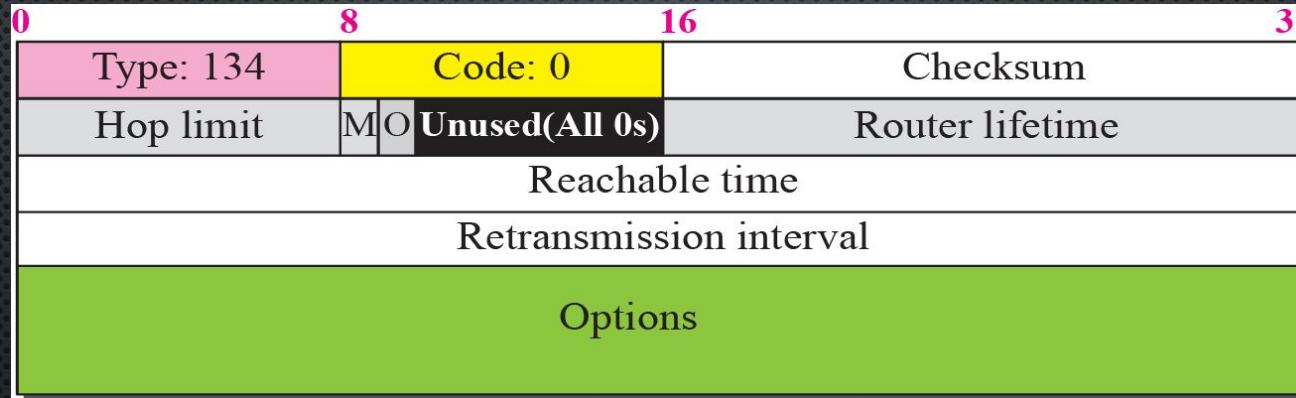


Figure 28.11 Router-advertisement message



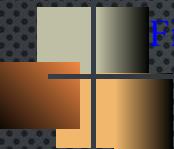
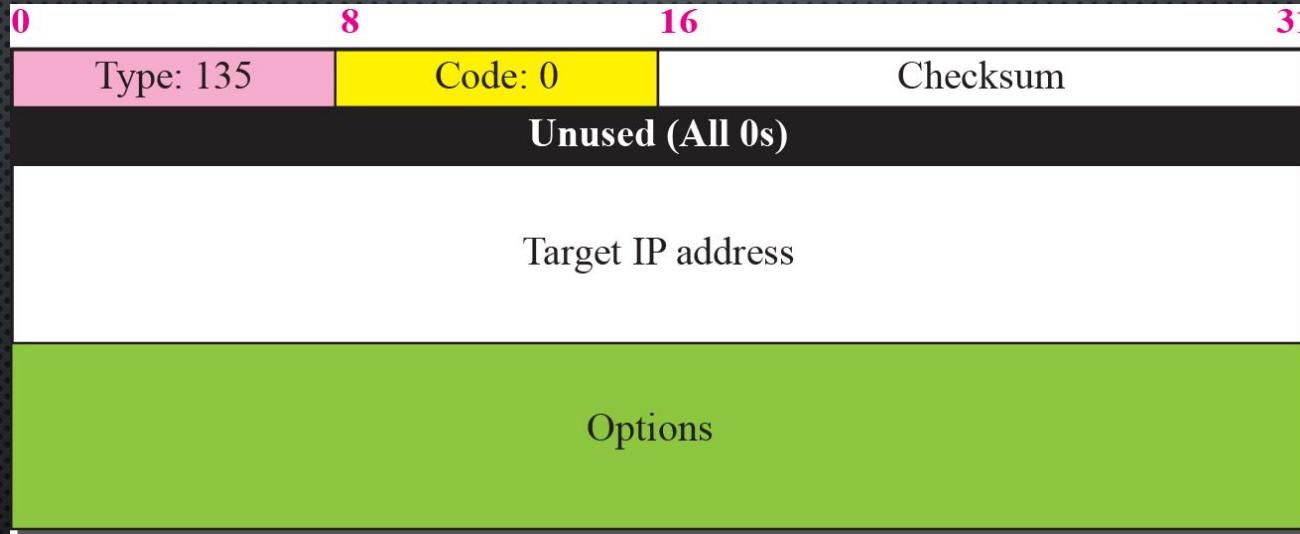


Figure 28.12 *Neighbor-solicitation message*



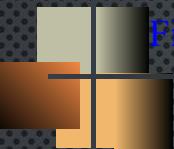


Figure 28.13 Neighbor advertisement message

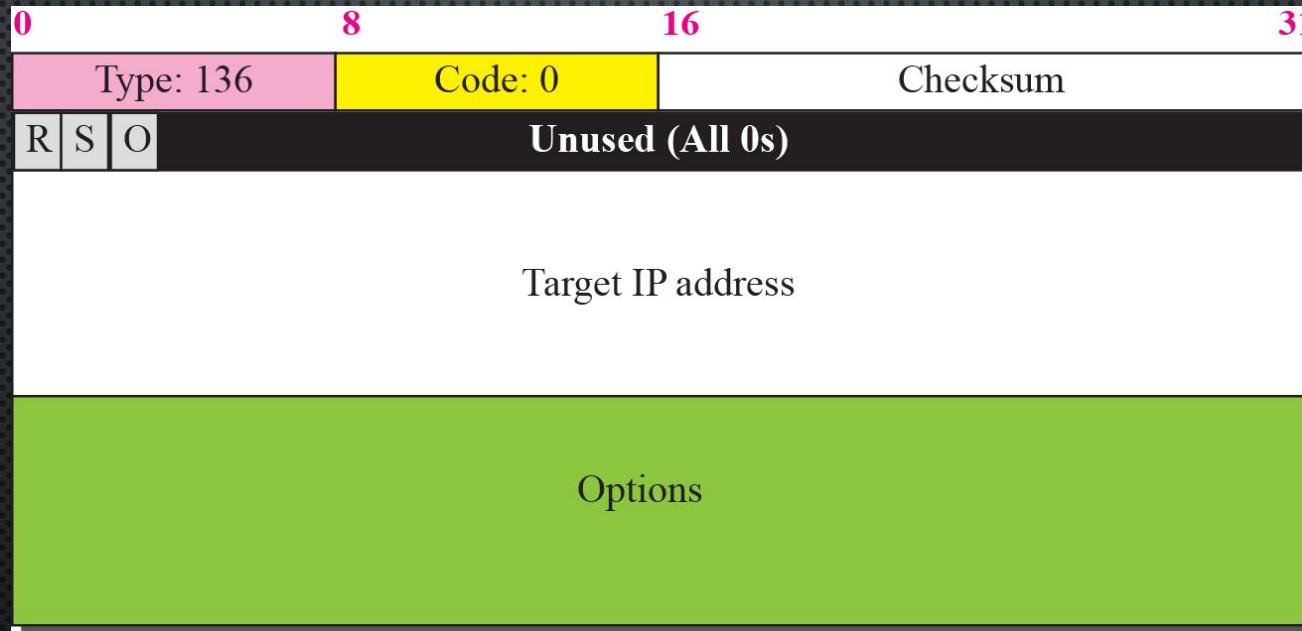
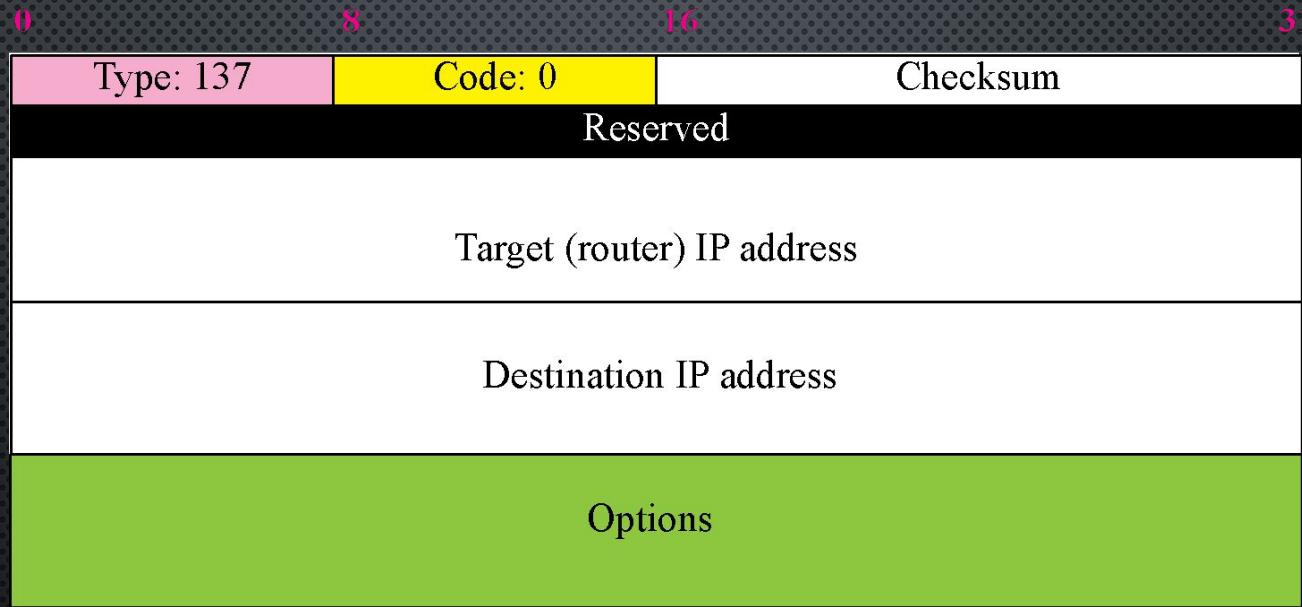


Figure 28.14 *Redirection message*



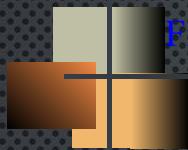
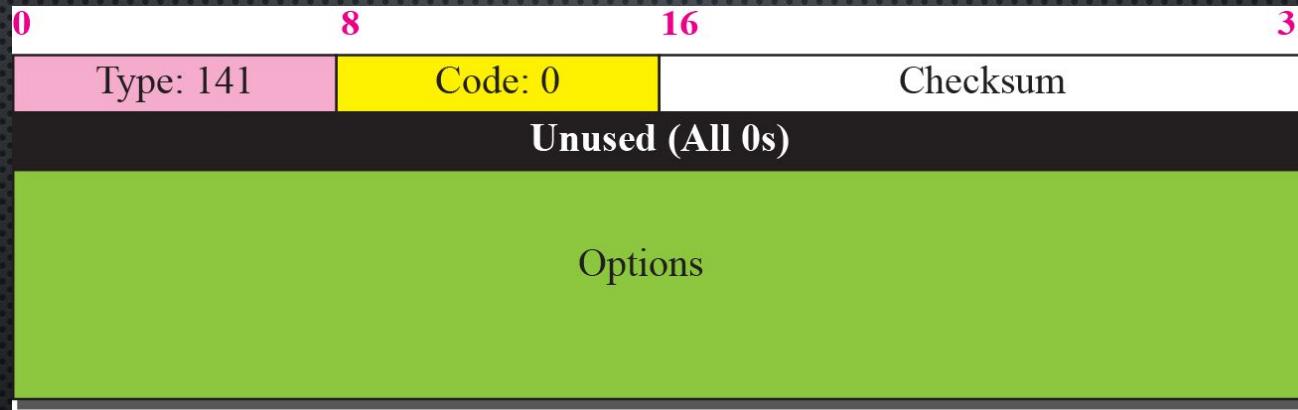


Figure 28.15 *Inverse-neighbor-solicitation message*



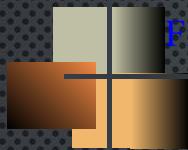


Figure 28.16 *Inverse-neighbor-advertisement message*



Module 6

Transport Layer

The transport layer is responsible for process-to-process delivery—the delivery of a packet, part of a message, from one process to another.

23-2 USER DATAGRAM PROTOCOL (UDP)

The User Datagram Protocol (UDP) is called a connectionless, unreliable transport protocol. It does not add anything to the services of IP except to provide process-to-process communication instead of host-to-host communication.

Table 23.1 *Well-known ports used with UDP*

<i>Port</i>	<i>Protocol</i>	<i>Description</i>
7	Echo	Echoes a received datagram back to the sender
9	Discard	Discards any datagram that is received
11	Users	Active users
13	Daytime	Returns the date and the time
17	Quote	Returns a quote of the day
19	Chargen	Returns a string of characters
53	Nameserver	Domain Name Service
67	BOOTPs	Server port to download bootstrap information
68	BOOTPc	Client port to download bootstrap information
69	TFTP	Trivial File Transfer Protocol
111	RPC	Remote Procedure Call
123	NTP	Network Time Protocol
161	SNMP	Simple Network Management Protocol
162	SNMP	Simple Network Management Protocol (trap)

Figure 23.9 *User datagram format*

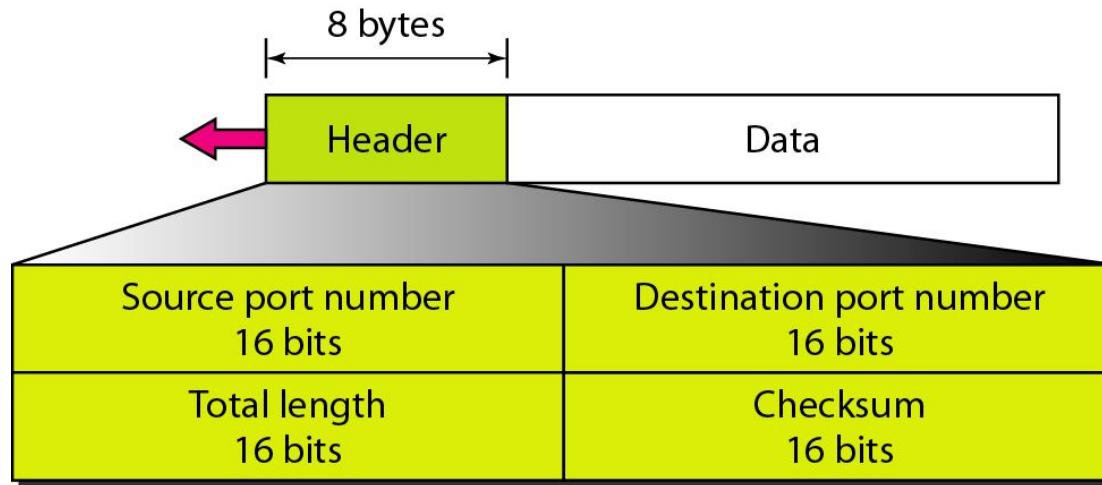


Figure 23.10 Pseudoheader for checksum calculation

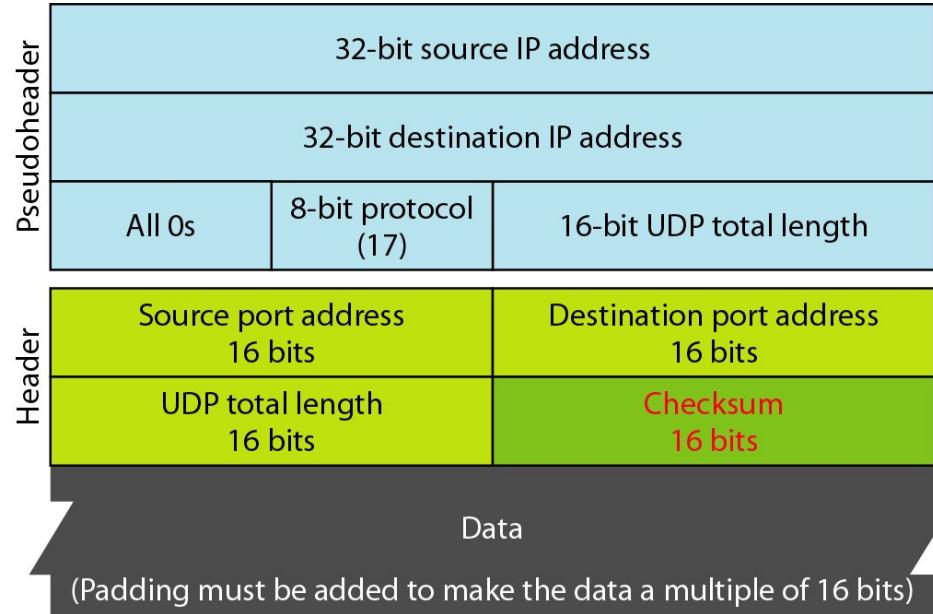
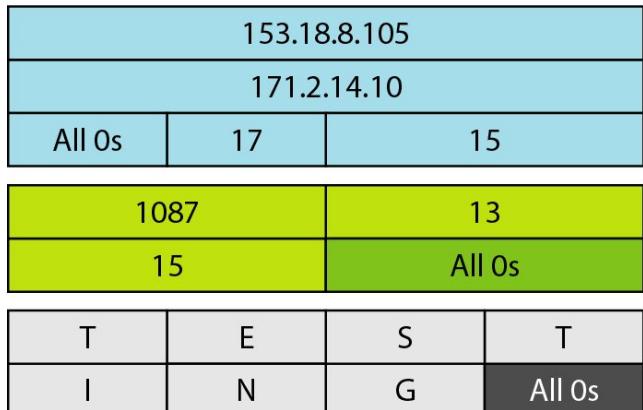




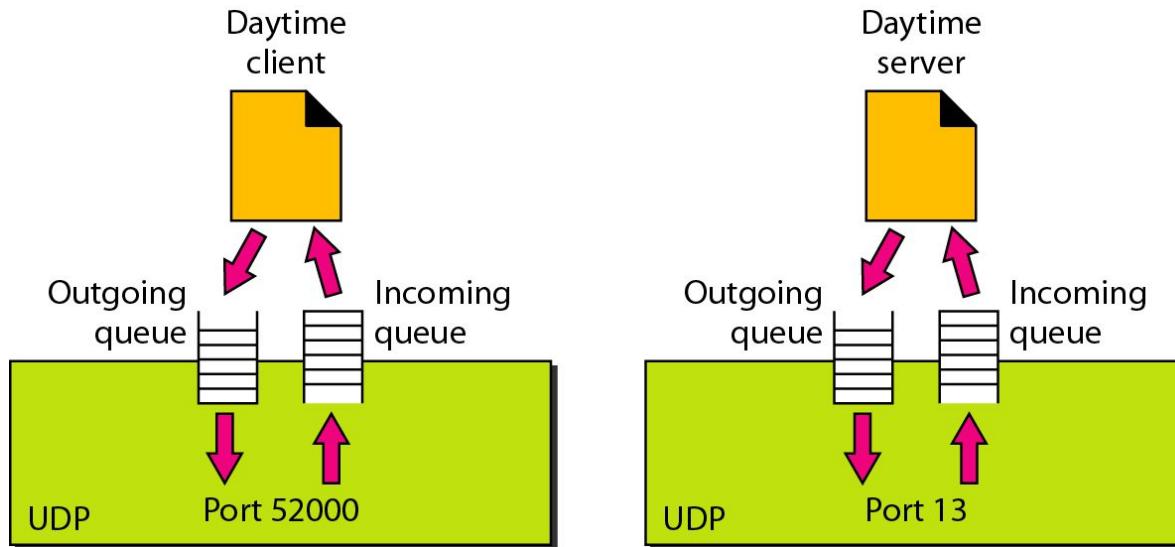
Figure 23.11 shows the checksum calculation for a very small user datagram with only 7 bytes of data. Because the number of bytes of data is odd, padding is added for checksum calculation. The pseudoheader as well as the padding will be dropped when the user datagram is delivered to IP.

Figure 23.11 Checksum calculation of a simple UDP user datagram



10011001 00010010	→ 153.18	
00001000 01101001	→ 8.105	
10101011 00000010	→ 171.2	
00001110 00001010	→ 14.10	
00000000 00010001	→ 0 and 17	
00000000 00001111	→ 15	
00000100 00111111	→ 1087	
00000000 00001101	→ 13	
00000000 00001111	→ 15	
00000000 00000000	→ 0 (checksum)	
01010100 01000101	→ T and E	
01010011 01010100	→ S and T	
01001001 01001110	→ I and N	
01000111 00000000	→ G and 0 (padding)	
10010110 11101011		→ Sum
01101001 00010100		→ Checksum

Figure 23.12 *Queues in UDP*



23-3 TCP

TCP is a connection-oriented protocol; it creates a virtual connection between two TCPS to send data. In addition, TCP uses flow and error control mechanisms at the transport level.

Table 23.2 Well-known ports used by TCP

<i>Port</i>	<i>Protocol</i>	<i>Description</i>
7	Echo	Echoes a received datagram back to the sender
9	Discard	Discards any datagram that is received
11	Users	Active users
13	Daytime	Returns the date and the time
17	Quote	Returns a quote of the day
19	Chargen	Returns a string of characters
20	FTP, Data	File Transfer Protocol (data connection)
21	FTP, Control	File Transfer Protocol (control connection)
23	TELNET	Terminal Network
25	SMTP	Simple Mail Transfer Protocol
53	DNS	Domain Name Server
67	BOOTP	Bootstrap Protocol
79	Finger	Finger
80	HTTP	Hypertext Transfer Protocol
111	RPC	Remote Procedure Call

Figure 23.13 *Stream delivery*

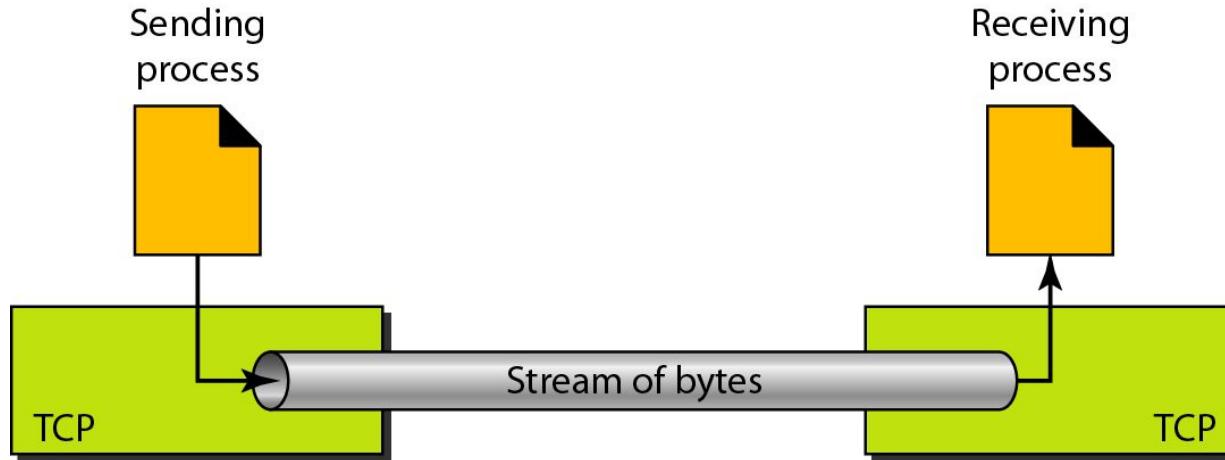


Figure 23.14 *Sending and receiving buffers*

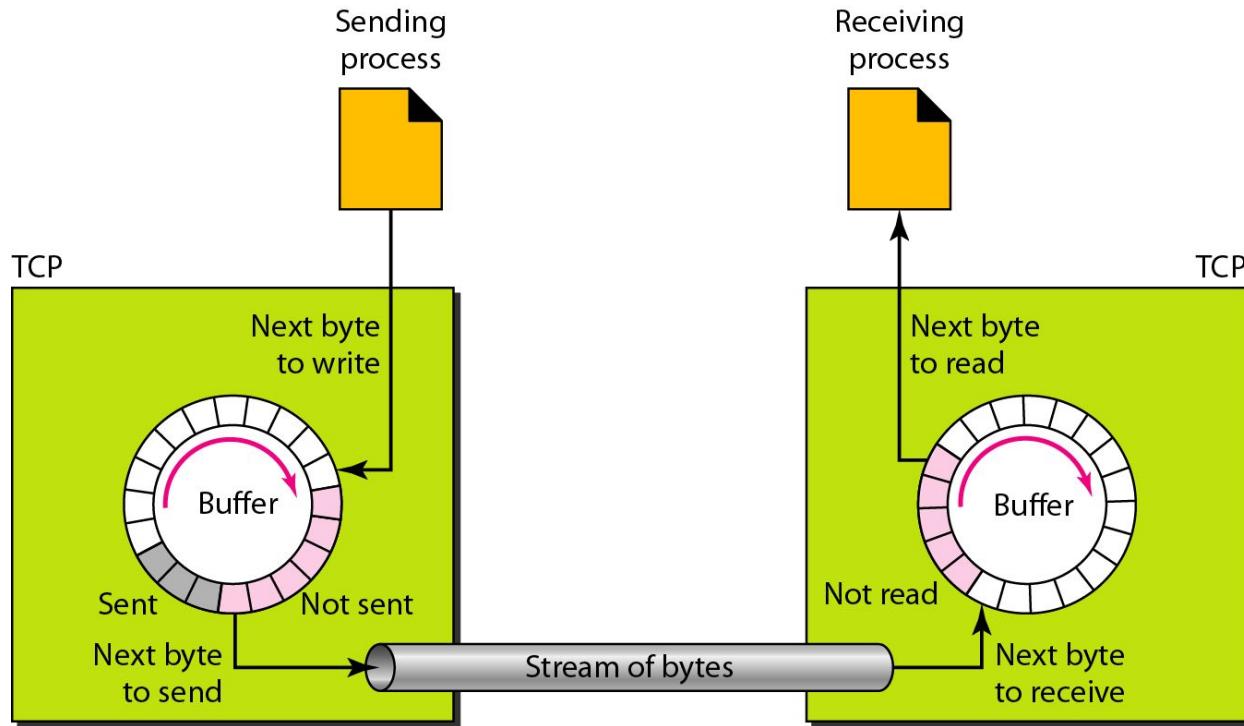
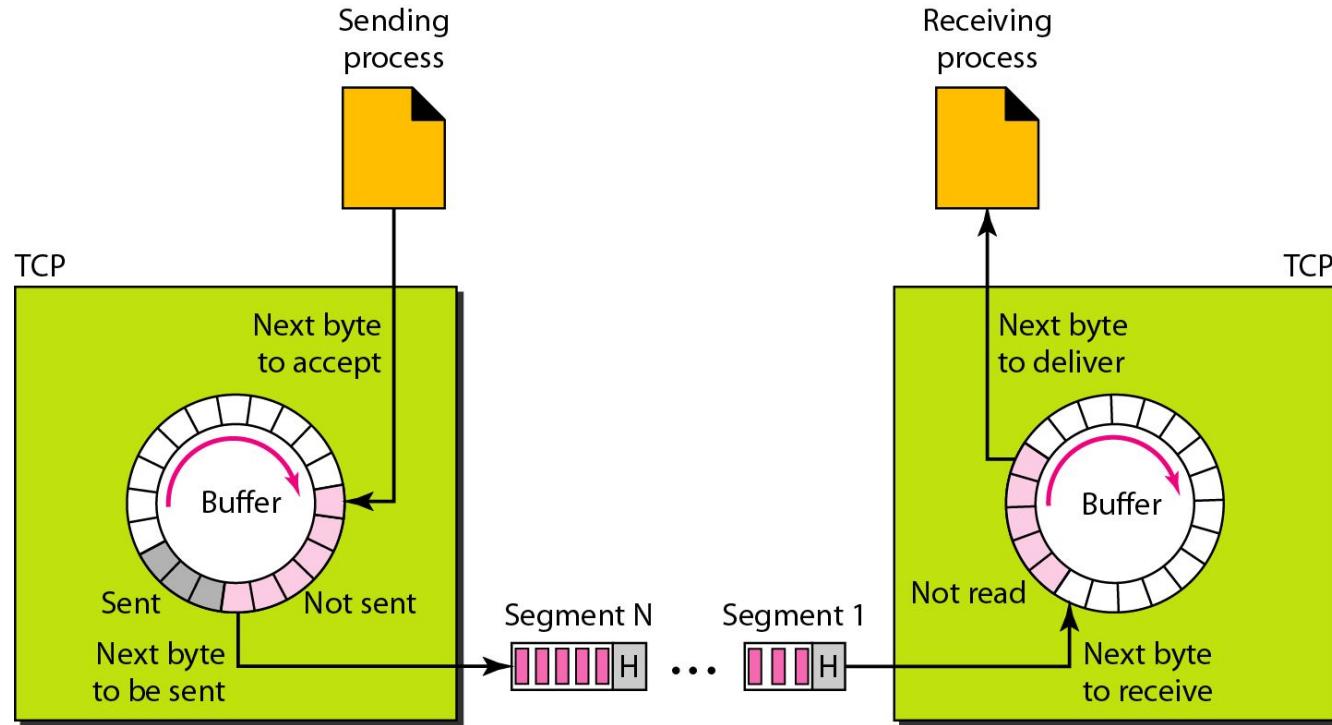
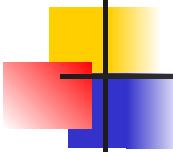


Figure 23.15 *TCP segments*



Note



The bytes of data being transferred in each connection are numbered by TCP.

The numbering starts with a randomly generated number.

The value in the sequence number field of a segment defines the number of the first data byte contained in that segment.

The value of the acknowledgment field in a segment defines the number of the next byte a party expects to receive.

The acknowledgment number is cumulative.

Figure 23.16 *TCP segment format*

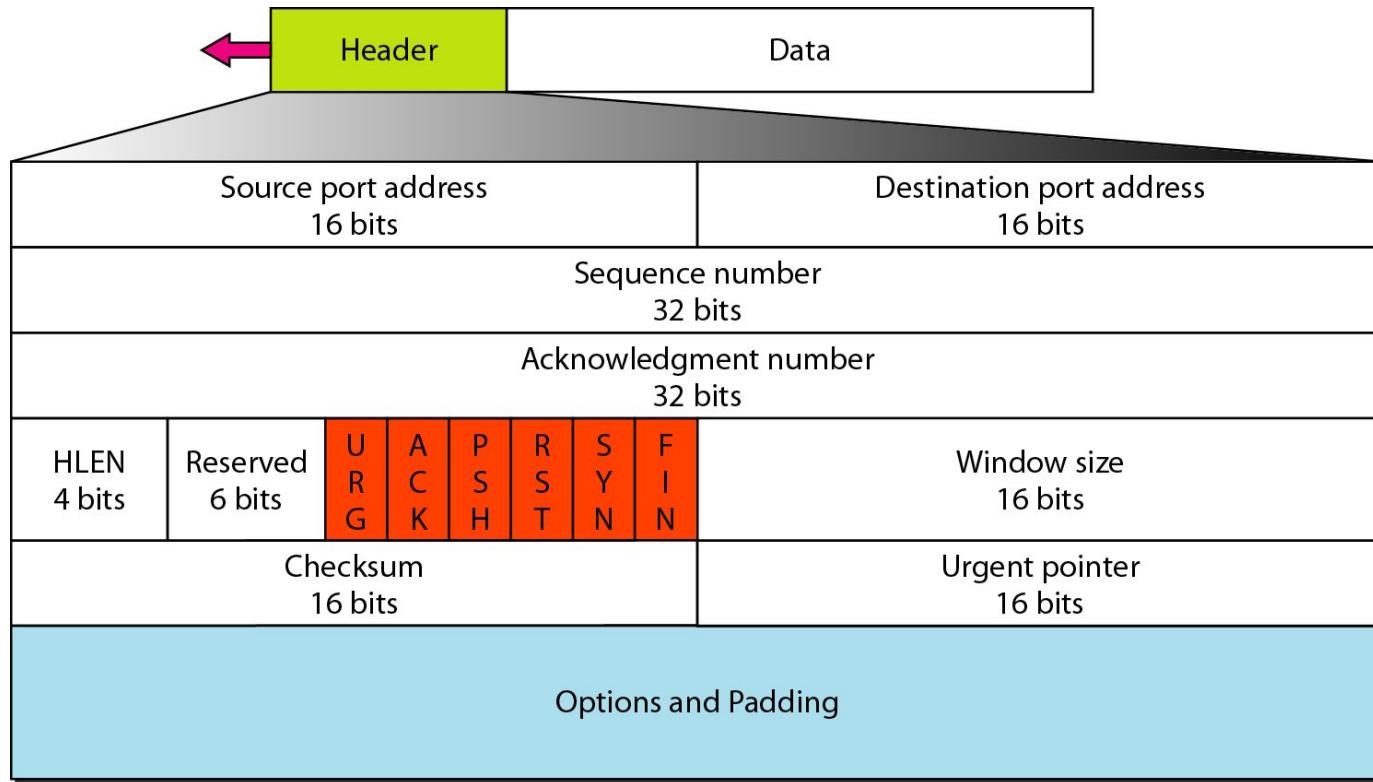


Figure 23.17 *Control field*

URG: Urgent pointer is valid

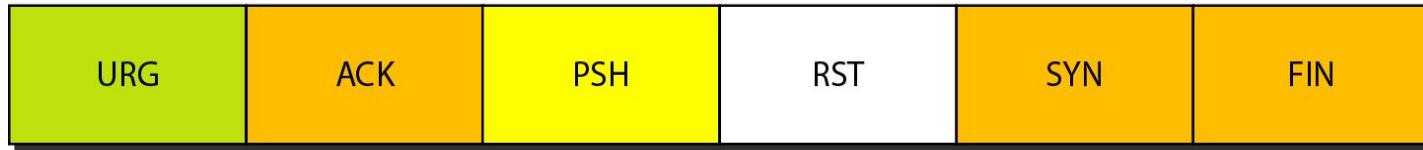
ACK: Acknowledgment is valid

PSH: Request for push

RST: Reset the connection

SYN: Synchronize sequence numbers

FIN: Terminate the connection



<i>Flag</i>	<i>Description</i>
URG	The value of the urgent pointer field is valid.
ACK	The value of the acknowledgment field is valid.
PSH	Push the data.
RST	Reset the connection.
SYN	Synchronize sequence numbers during connection.
FIN	Terminate the connection.

Figure 23.18 Connection establishment using three-way handshaking

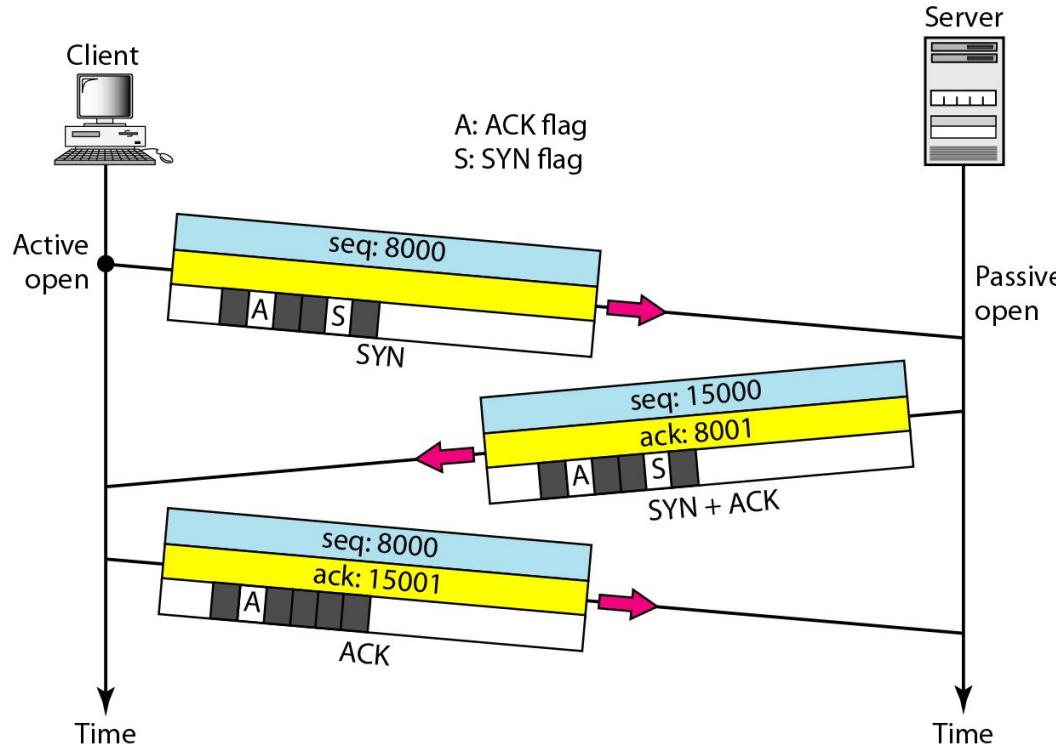


Figure 23.19 Data transfer

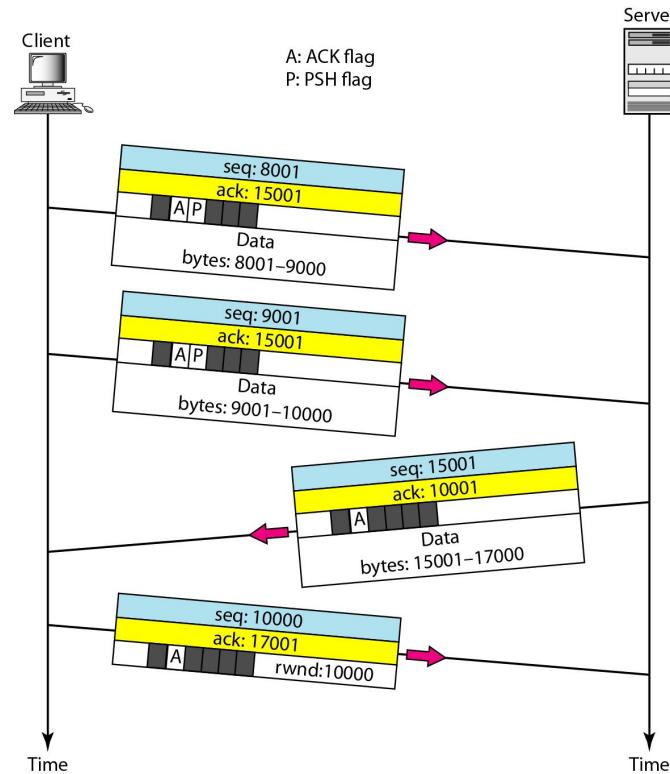
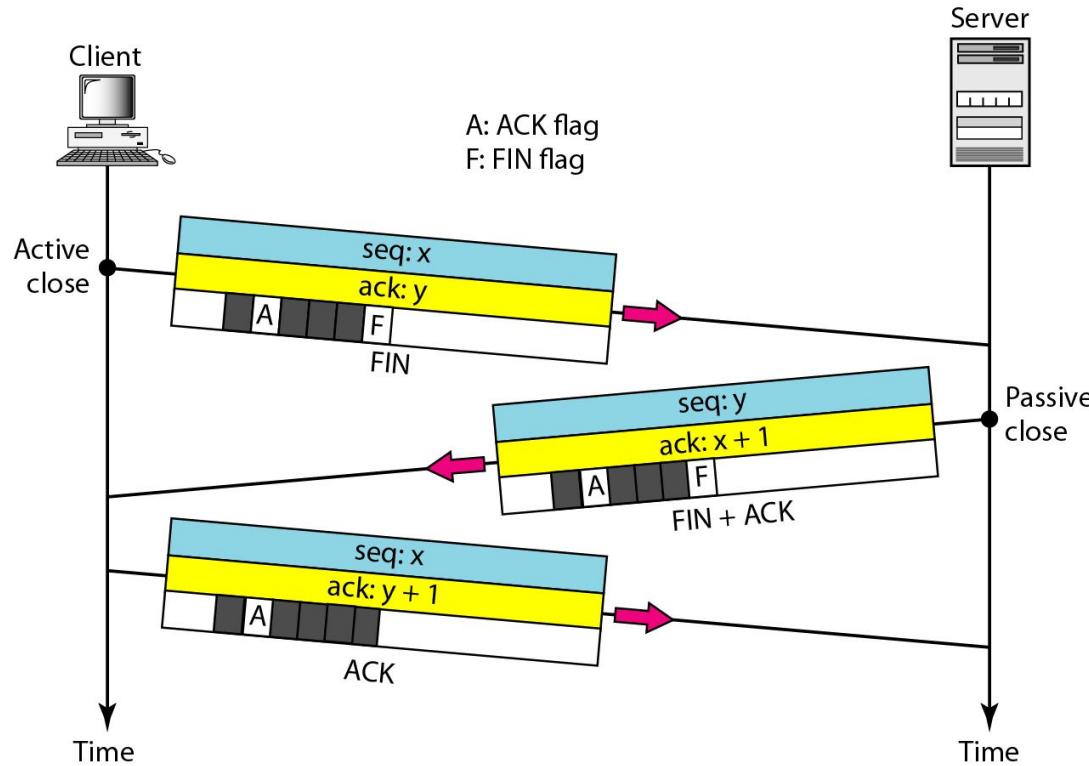
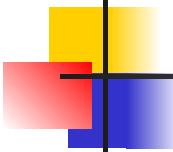


Figure 23.20 Connection termination using three-way handshaking





Note

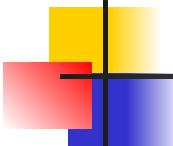
A SYN segment cannot carry data, but it consumes one sequence number.

A SYN + ACK segment cannot carry data, but does consume one sequence number.

An ACK segment, if carrying no data, consumes no sequence number.

The FIN segment consumes one sequence number if it does not carry data.

The FIN + ACK segment consumes one sequence number if it does not carry data.



Note

A sliding window is used to make transmission more efficient as well as to control the flow of data so that the destination does not become overwhelmed with data.

TCP sliding windows are byte-oriented.

In modern implementations, a retransmission occurs if the retransmission timer expires or three duplicate ACK segments have arrived.

No retransmission timer is set for an ACK segment.

Data may arrive out of order and be temporarily stored by the receiving TCP,

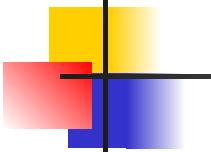
but TCP guarantees that no out-of-order segment is delivered to the process.

Application Layer

The application layer enables the user, whether human or software, to access the network. It provides user interfaces and support for services such as electronic mail, file access and transfer, access to system resources, surfing the world wide web, and network management.

26-3 FILE TRANSFER

Transferring files from one computer to another is one of the most common tasks expected from a networking or internetworking environment. As a matter of fact, the greatest volume of data exchange in the Internet today is due to file transfer.



Note

FTP uses the services of TCP. It needs two TCP connections.

The well-known port 21 is used for the control connection and the well-known port 20 for the data connection.

Figure 26.21 *FTP*

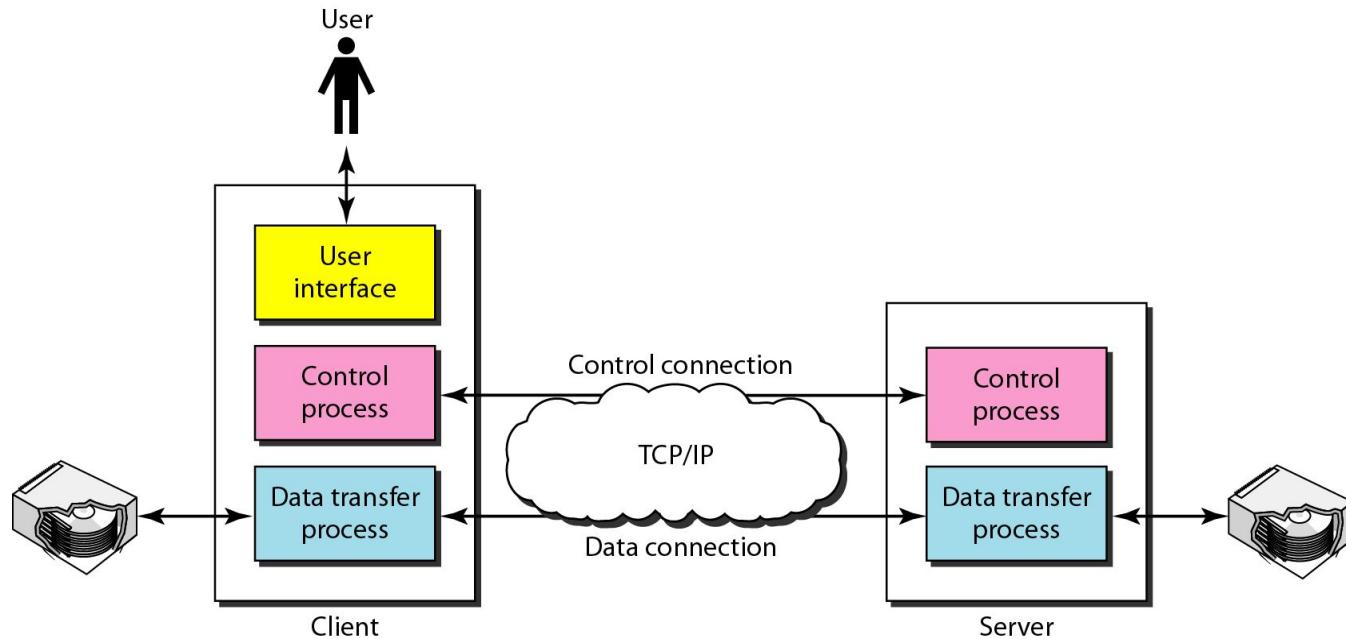


Figure 26.22 *Using the control connection*

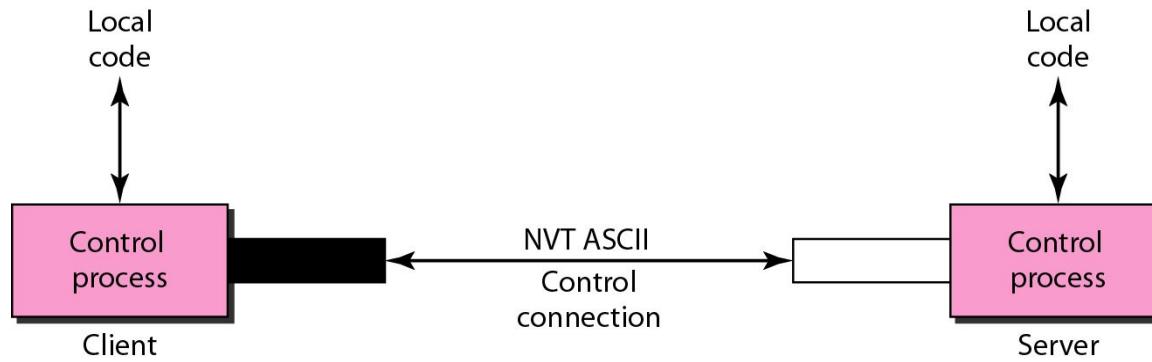
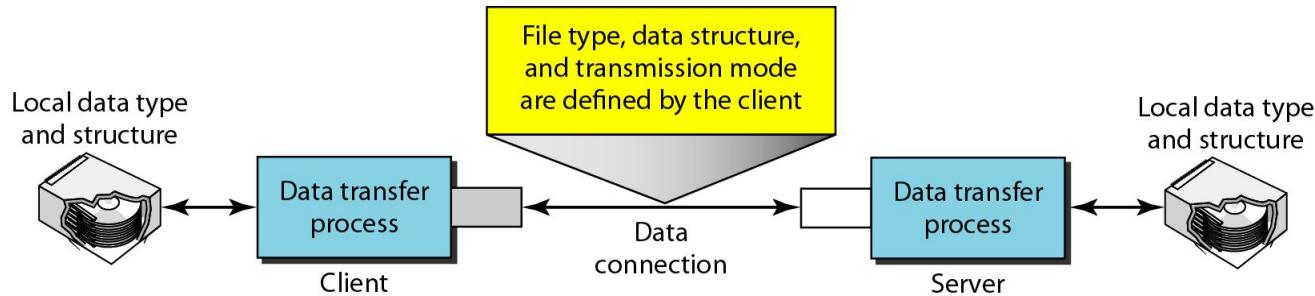


Figure 26.23 *Using the data connection*



25-2 DOMAIN NAME SPACE

To have a hierarchical name space, a domain name space was designed. In this design the names are defined in an inverted-tree structure with the root at the top. The tree can have only 128 levels: level 0 (root) to level 127.

Figure 25.2 Domain name space

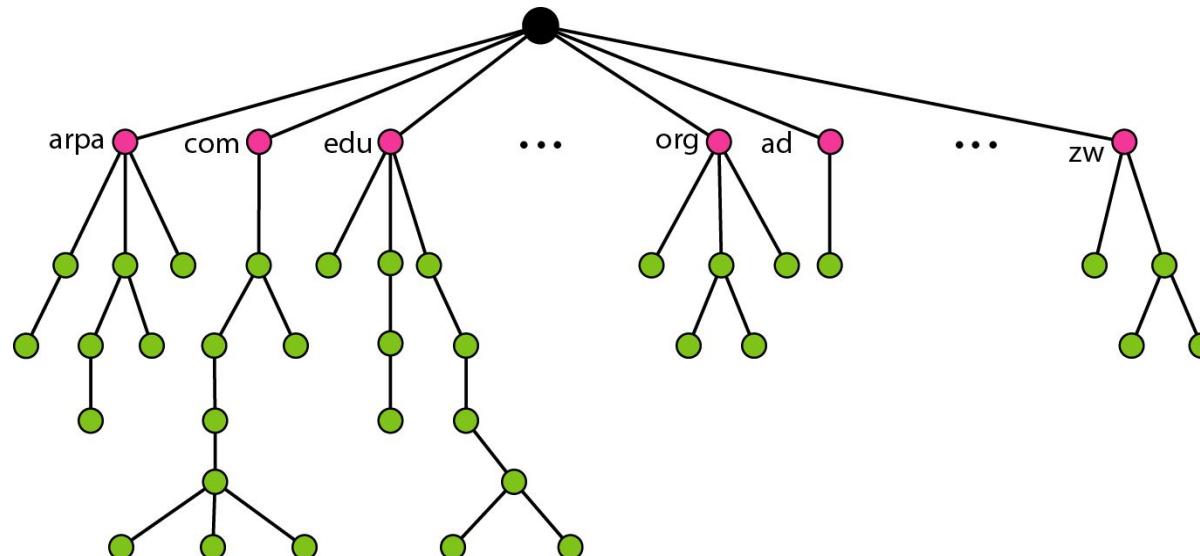
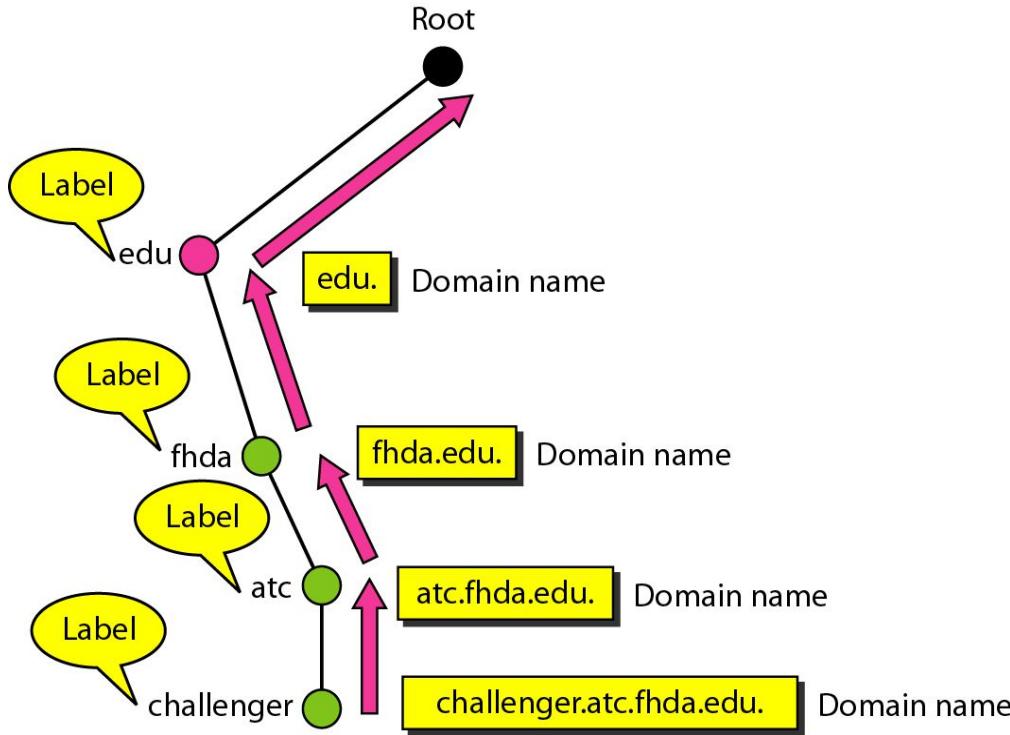


Figure 25.3 Domain names and labels



25-4 DNS IN THE INTERNET

DNS is a protocol that can be used in different platforms. In the Internet, the domain name space (tree) is divided into three different sections: generic domains, country domains, and the inverse domain.

Generic Domains

Country Domains

Inverse Domain

Figure 25.9 *Generic domains*

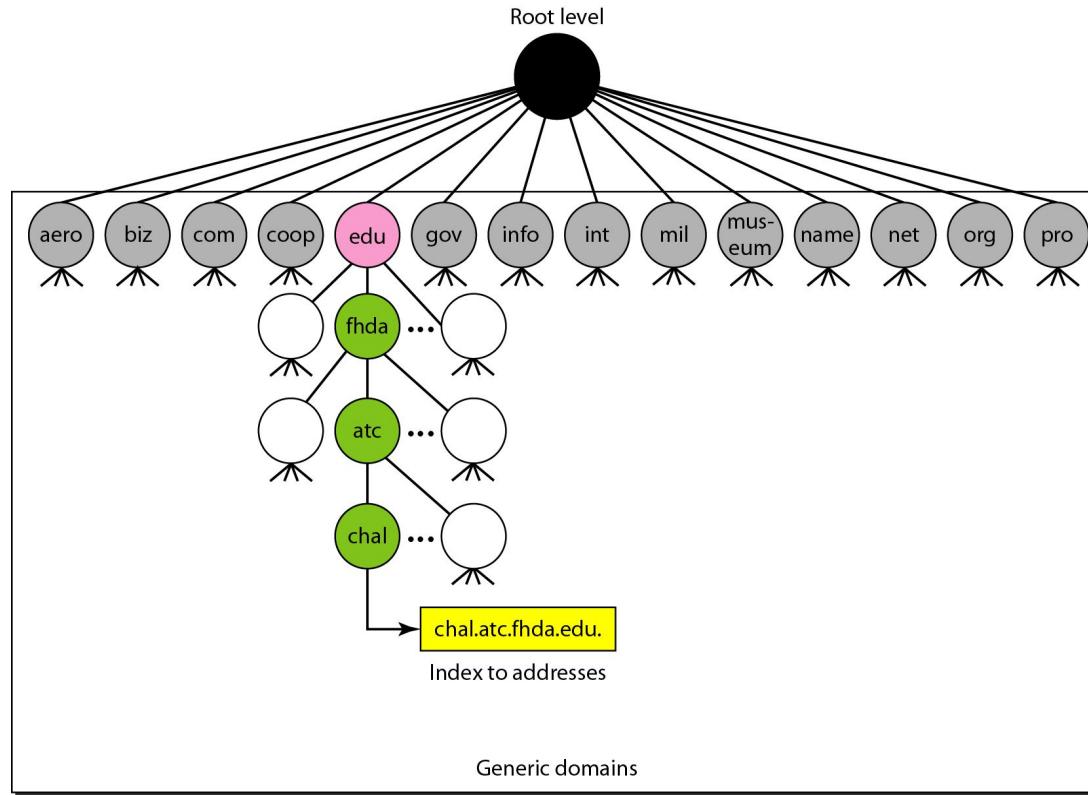


Figure 25.10 *Country domains*

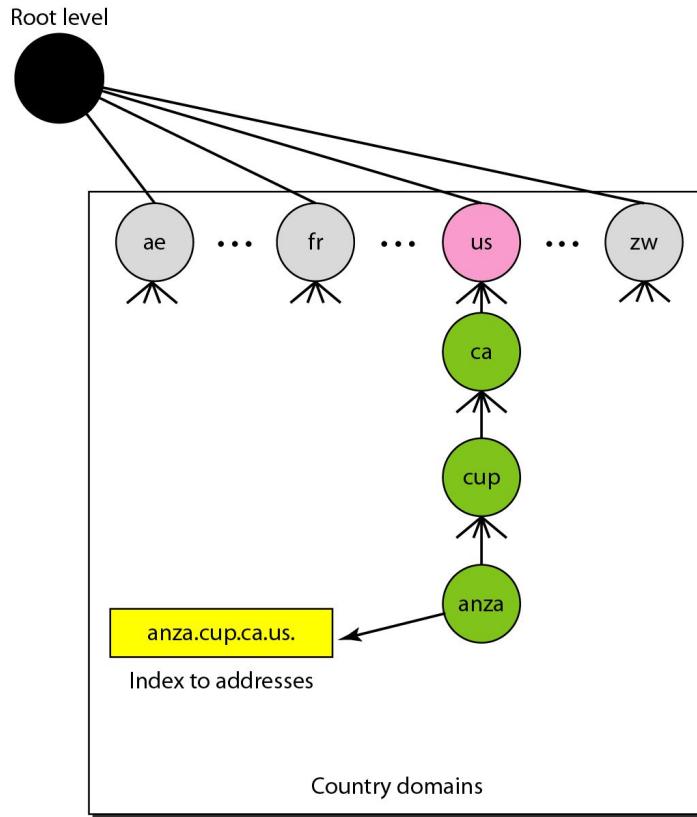
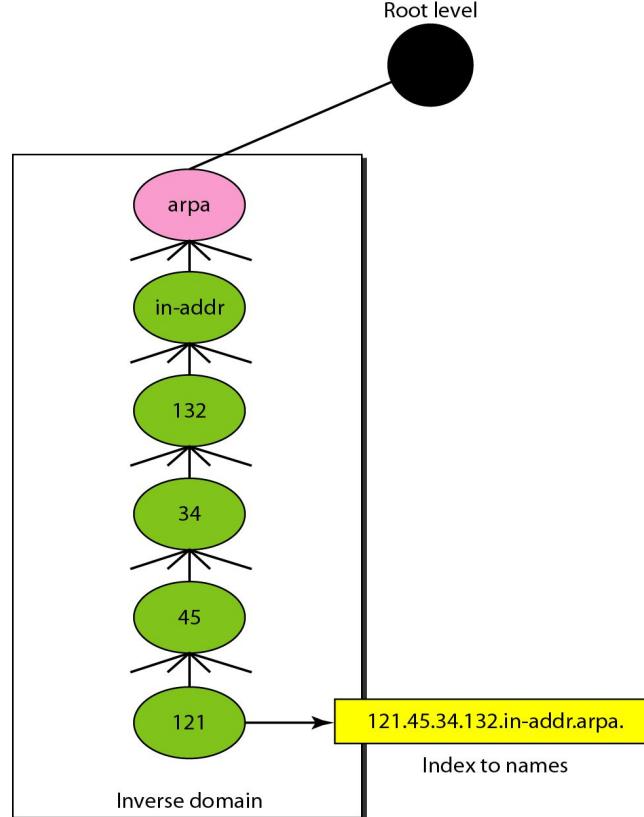


Figure 25.11 Inverse domain



25-5 RESOLUTION

Mapping a name to an address or an address to a name is called name-address resolution.

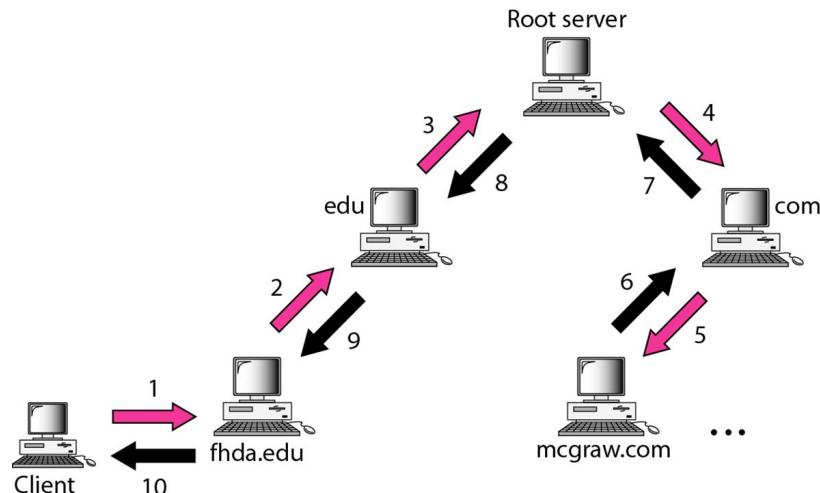


Figure 25.12 *Recursive resolution*

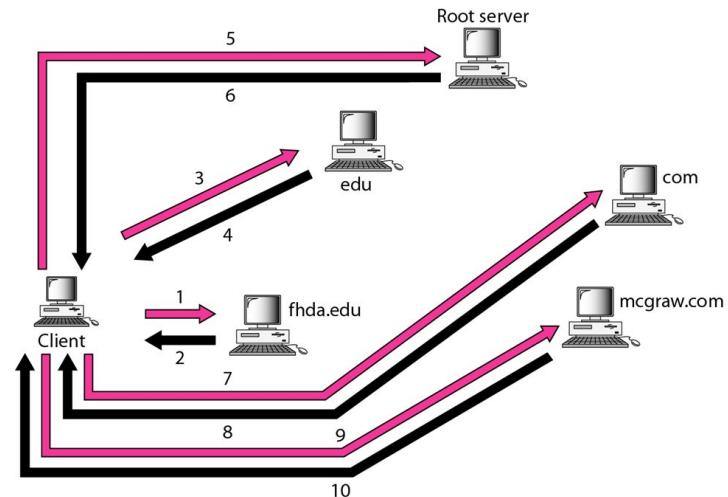
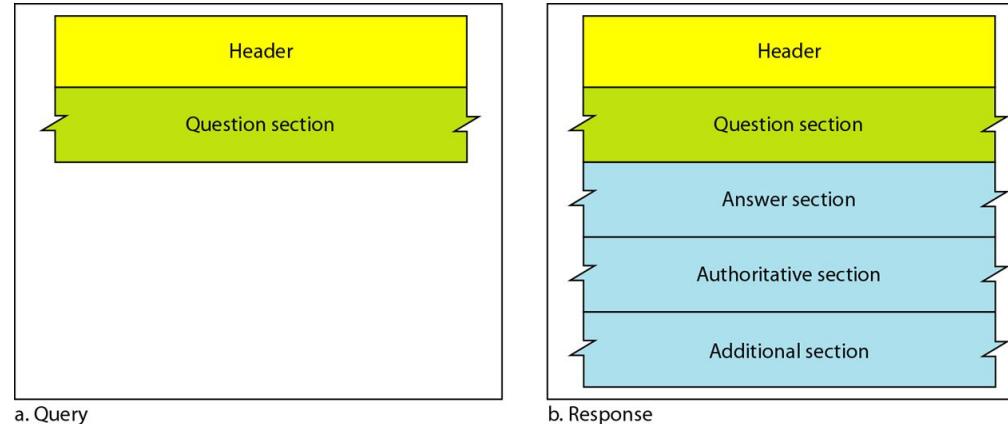


Figure 25.13 *Iterative resolution*

25-6 DNS MESSAGES

DNS has two types of messages: query and response. Both types have the same format. The query message consists of a header and question records; the response message consists of a header, question records, answer records, authoritative records, and additional records.



25-7 TYPES OF RECORDS

As we saw in Section 25.6, two types of records are used in DNS. The question records are used in the question section of the query and response messages. The resource records are used in the answer, authoritative, and additional information sections of the response message.

25-8 REGISTRARS

How are new domains added to DNS? This is done through a registrar, a commercial entity accredited by ICANN. A registrar first verifies that the requested domain name is unique and then enters it into the DNS database. A fee is charged.

25-9 DYNAMIC DOMAIN NAME SYSTEM (DDNS)

The DNS master file must be updated dynamically. The Dynamic Domain Name System (DDNS) therefore was devised to respond to this need. In DDNS, when a binding between a name and an address is determined, the information is sent, usually by DHCP to a primary DNS server. The primary server updates the zone. The secondary servers are notified either actively or passively.

25-10 ENCAPSULATION

DNS can use either UDP or TCP. In both cases the well-known port used by the server is port 53. UDP is used when the size of the response message is less than 512 bytes because most UDP packages have a 512-byte packet size limit. If the size of the response message is more than 512 bytes, a TCP connection is used.

26-2 ELECTRONIC MAIL

One of the most popular Internet services is electronic mail (e-mail). The designers of the Internet probably never imagined the popularity of this application program. Its architecture consists of several components that we discuss in this chapter.

Figure 26.6 First scenario : When the sender and the receiver of an e-mail are on the same system, we need only two user agents.

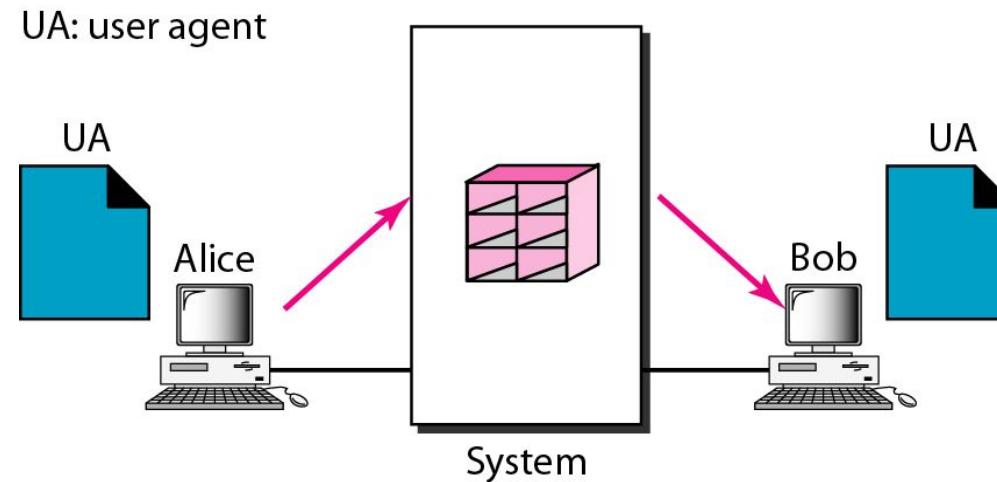


Figure 26.7 Second scenario: When the sender and the receiver of an e-mail are on different systems, we need two UAs and a pair of MTAs (client and server).

UA: user agent

MTA: message transfer agent

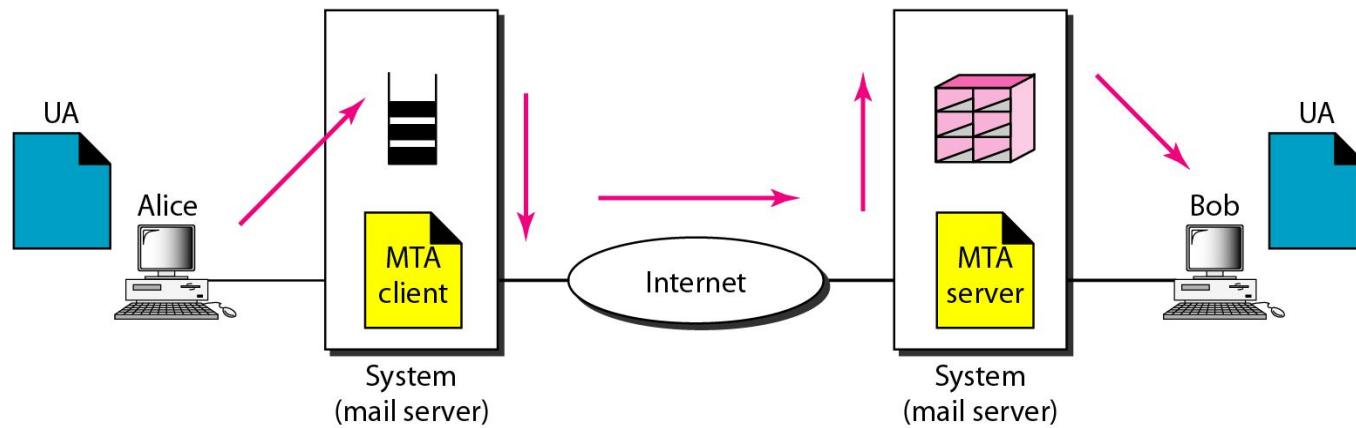


Figure 26.8 Third scenario: When the sender is connected to the mail server via a LAN or a WAN, we need two UAs and two pairs

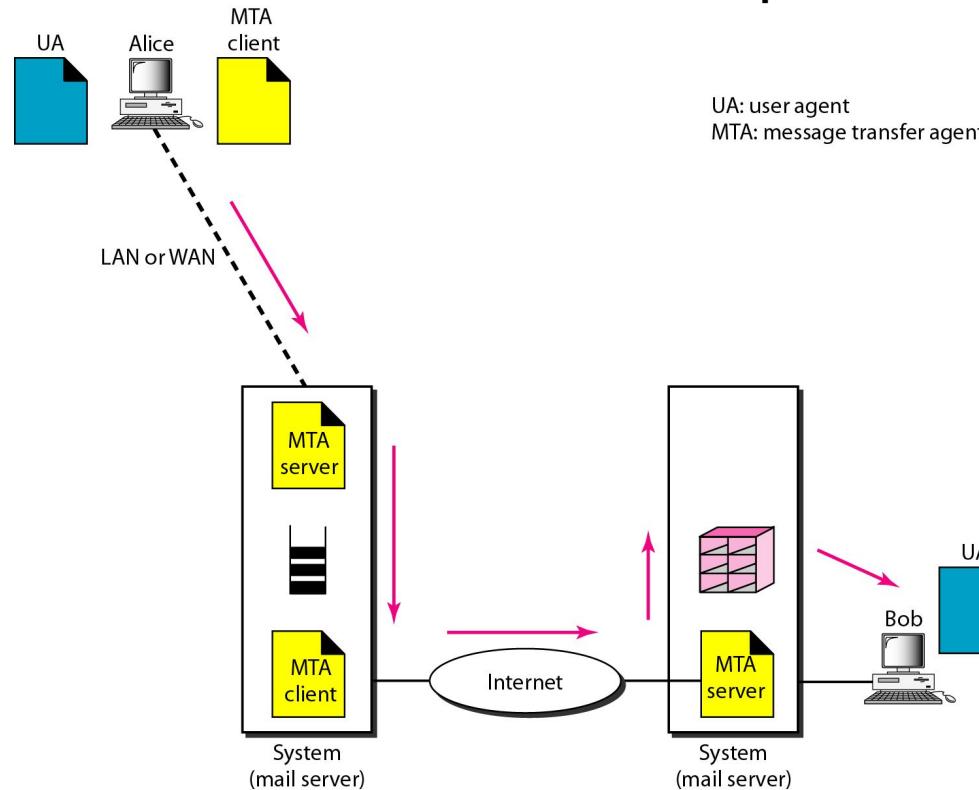


Figure 26.9 *Fourth scenario: When both sender and receiver are connected to the mail server via a LAN or a WAN, we need two UAs, two pairs of MTAs and a pair of MAAs.*

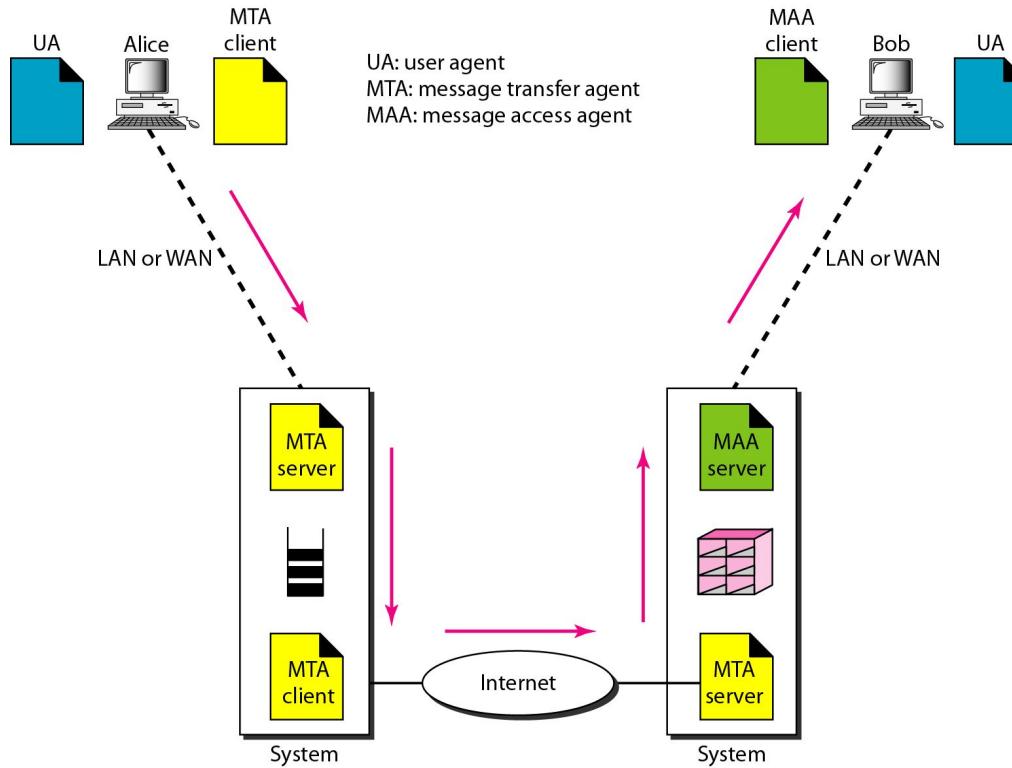


Figure 26.10 *Push versus pull in electronic email*

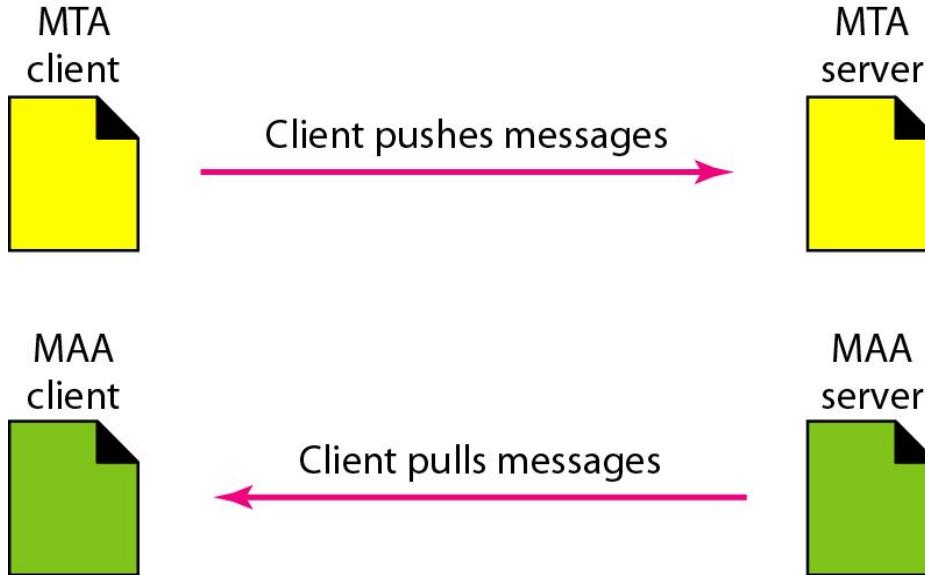


Figure 26.11 *Services of user agent*

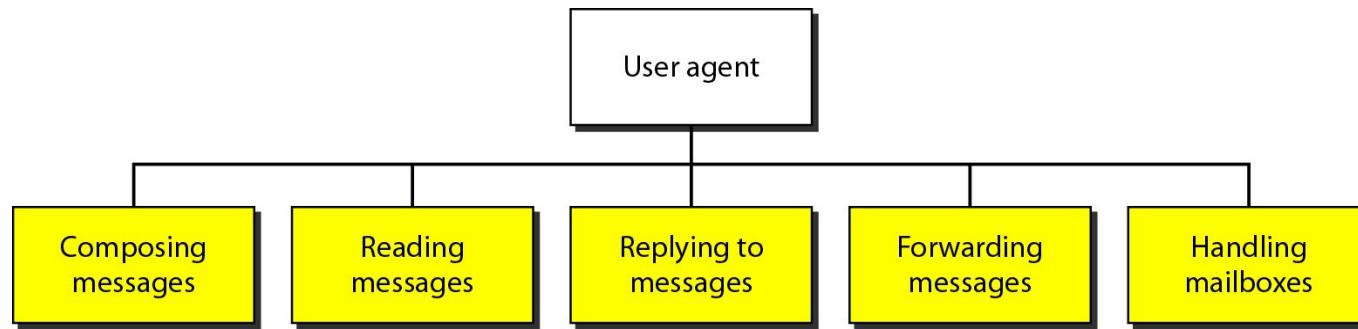


Figure 26.12 Format of an e-mail

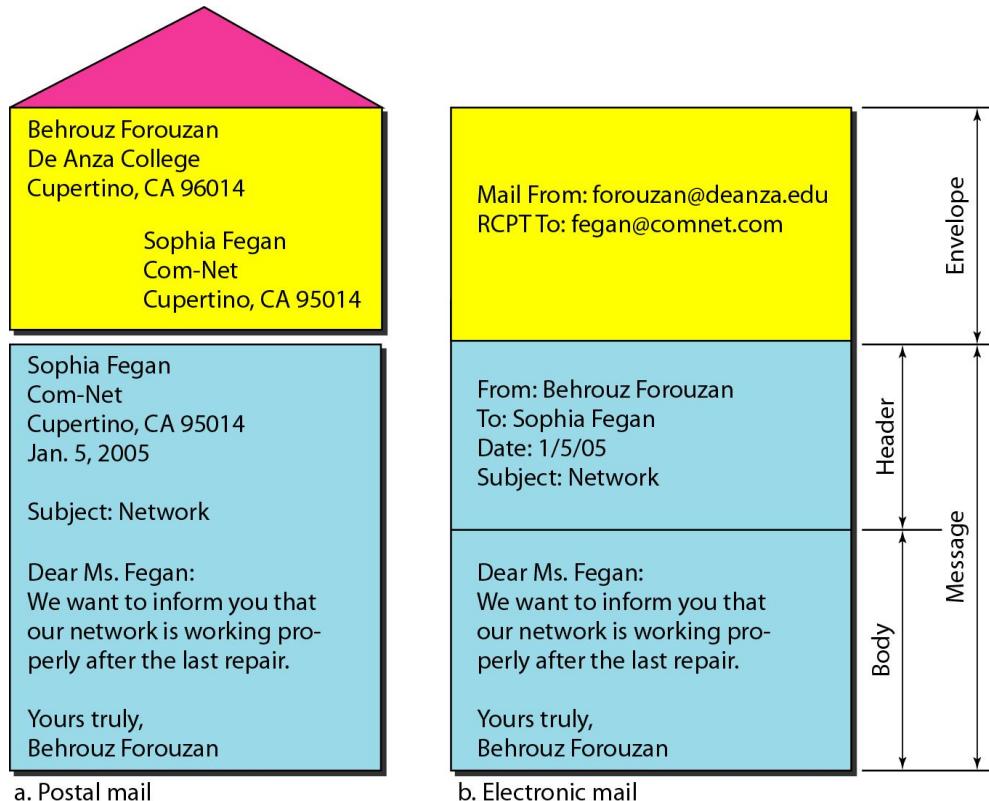


Figure 26.13 *E-mail address*

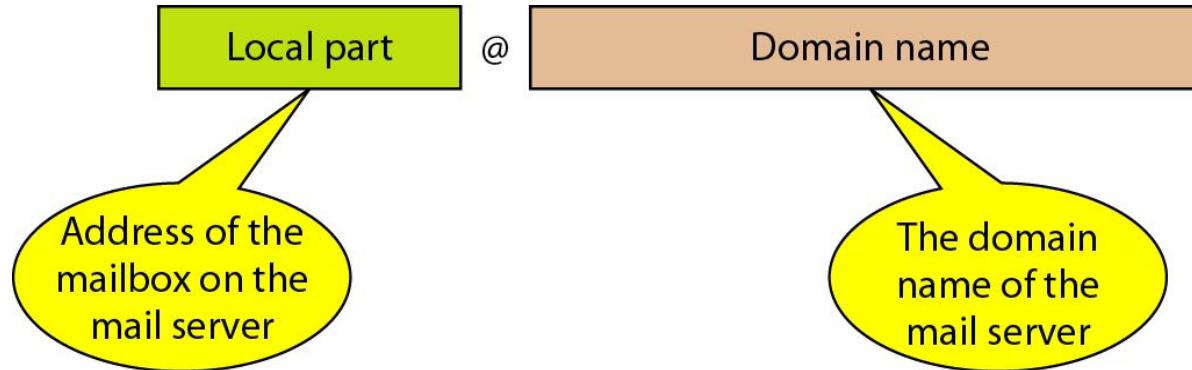


Figure 26.16 SMTP range

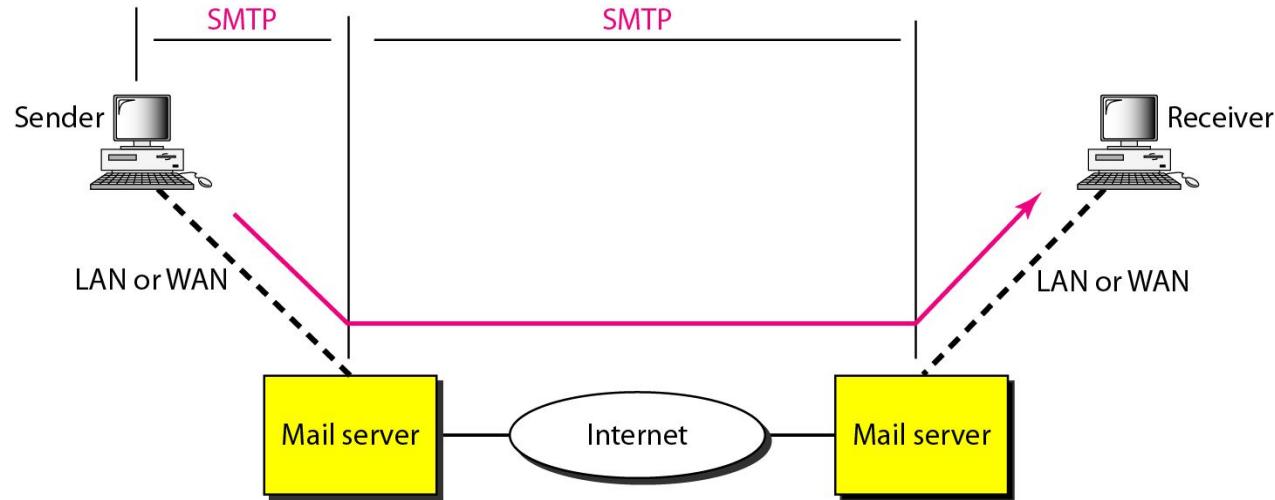


Figure 26.17 Commands and responses

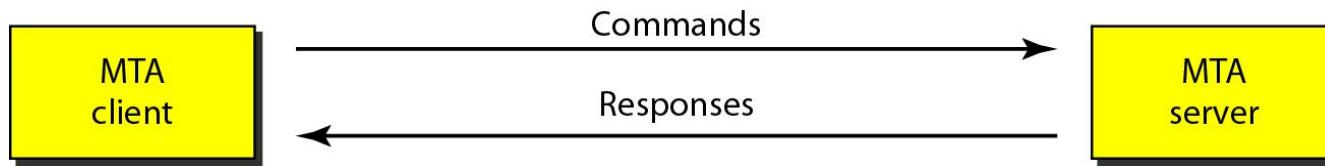


Table 26.7 *Commands*

<i>Keyword</i>	<i>Argument(s)</i>
HELO	Sender's host name
MAIL FROM	Sender of the message
RCPT TO	Intended recipient of the message
DATA	Body of the mail
QUIT	
RSET	
VRFY	Name of recipient to be verified
NOOP	
TURN	
EXPN	Mailing list to be expanded
HELP	Command name
SEND FROM	Intended recipient of the message
SMOL FROM	Intended recipient of the message
SMAL FROM	Intended recipient of the message

Table 26.8 Responses

<i>Code</i>	<i>Description</i>
Positive Completion Reply	
211	System status or help reply
214	Help message
220	Service ready
221	Service closing transmission channel
250	Request command completed
251	User not local; the message will be forwarded
Positive Intermediate Reply	
354	Start mail input
Transient Negative Completion Reply	
421	Service not available
450	Mailbox not available
451	Command aborted: local error
452	Command aborted: insufficient storage

<i>Code</i>	<i>Description</i>
Permanent Negative Completion Reply	
500	Syntax error; unrecognized command
501	Syntax error in parameters or arguments
502	Command not implemented
503	Bad sequence of commands
504	Command temporarily not implemented
550	Command is not executed; mailbox unavailable
551	User not local
552	Requested action aborted; exceeded storage location
553	Requested action not taken; mailbox name not allowed
554	Transaction failed

Figure 26.19 POP3 and IMAP4

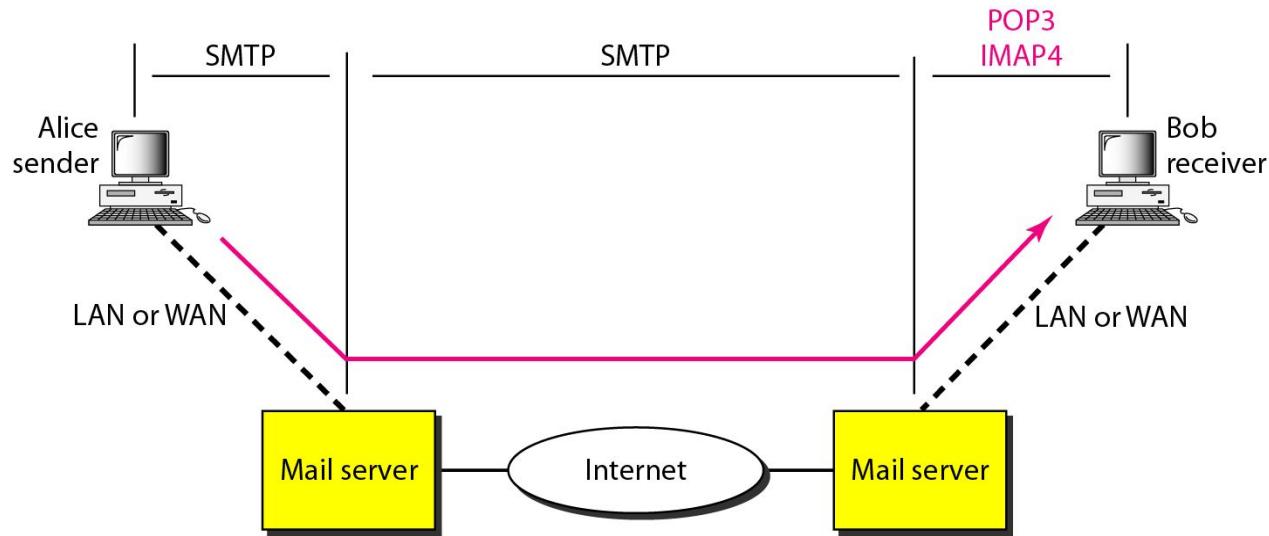


Figure 26.20 *The exchange of commands and responses in POP3*

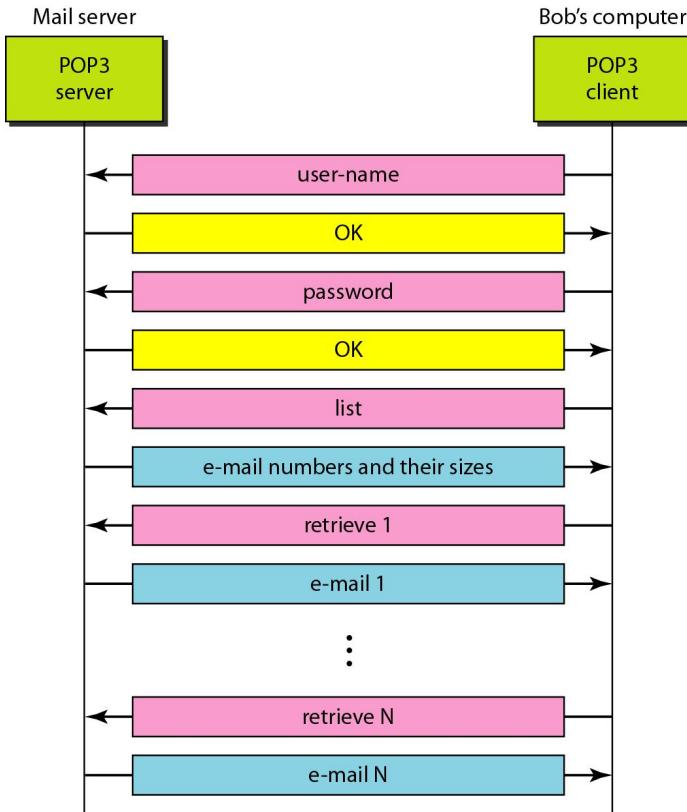


Figure 26.14 *MIME*

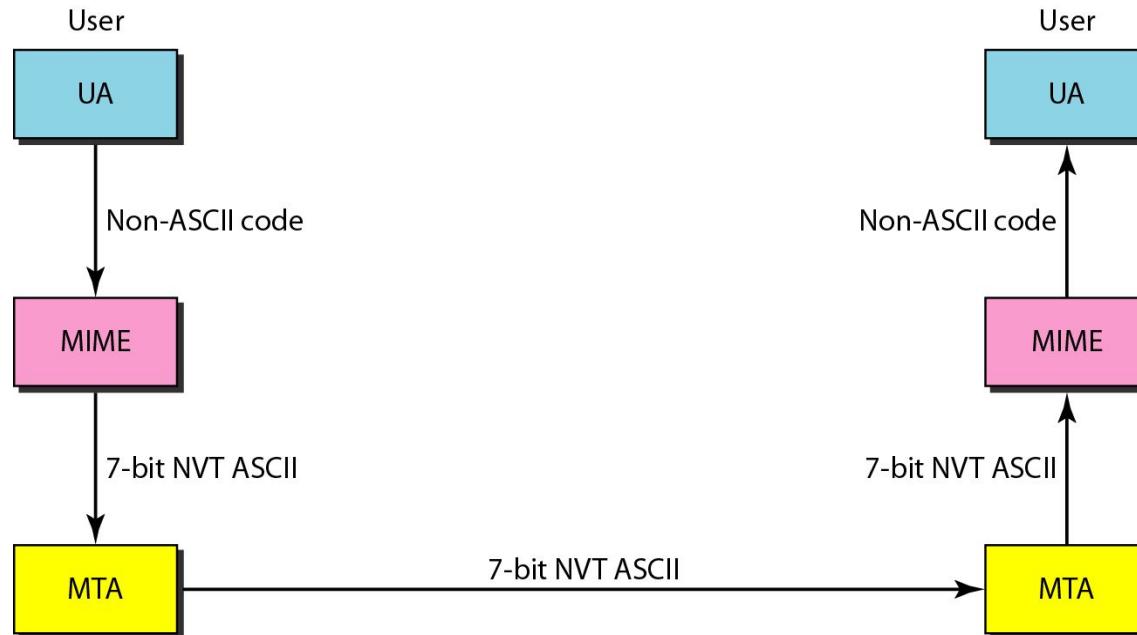


Figure 26.15 *MIME header*

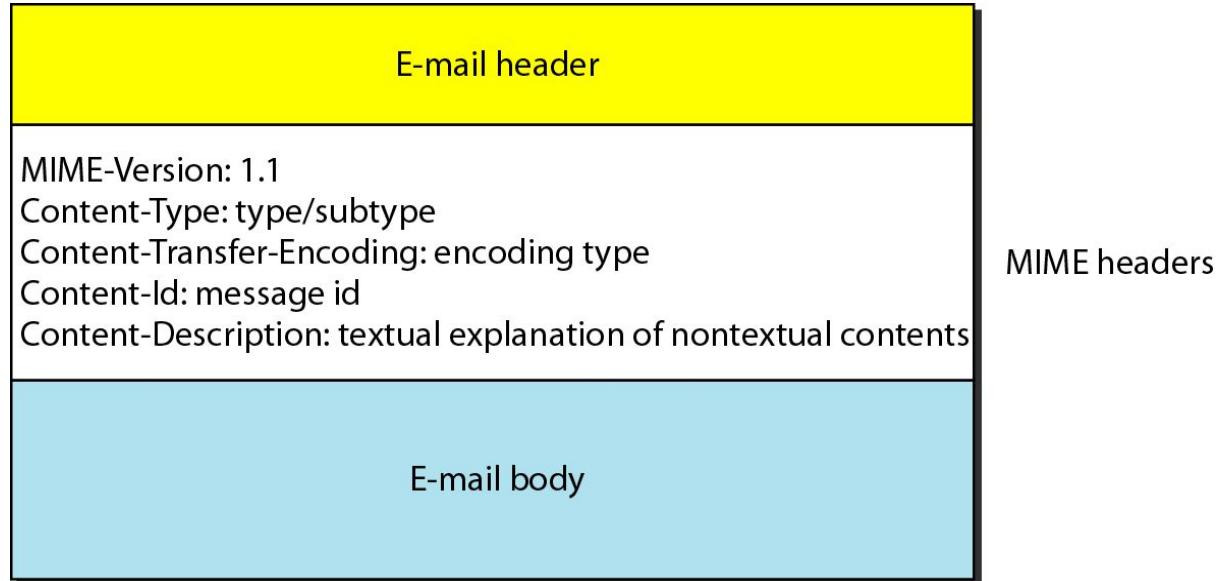


Table 26.5 *Data types and subtypes in MIME*

Type	Subtype	Description
Text	Plain	Unformatted
	HTML	HTML format (see Chapter 27)
Multipart	Mixed	Body contains ordered parts of different data types
	Parallel	Same as above, but no order
	Digest	Similar to mixed subtypes, but the default is message/RFC822
	Alternative	Parts are different versions of the same message
Message	RFC822	Body is an encapsulated message
	Partial	Body is a fragment of a bigger message
	External-Body	Body is a reference to another message
Image	JPEG	Image is in JPEG format
	GIF	Image is in GIF format
Video	MPEG	Video is in MPEG format
Audio	Basic	Single-channel encoding of voice at 8 kHz
Application	PostScript	Adobe PostScript
	Octet-stream	General binary data (8-bit bytes)

Table 26.6 *Content-transfer-encoding*

Type	Description
7-bit	NVT ASCII characters and short lines
8-bit	Non-ASCII characters and short lines
Binary	Non-ASCII characters with unlimited-length lines
Base-64	6-bit blocks of data encoded into 8-bit ASCII characters
Quoted-printable	Non-ASCII characters encoded as an equals sign followed by an ASCII code

28-1 NETWORK MANAGEMENT SYSTEM

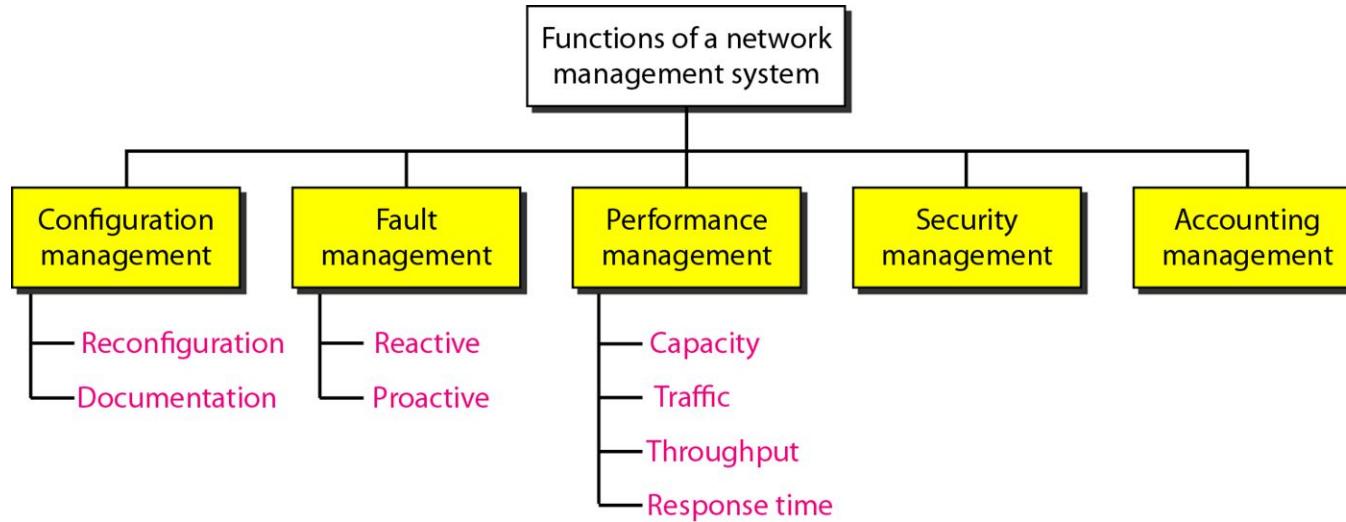
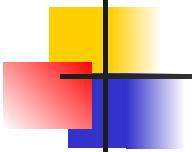


Figure 28.1 *Functions of a network management system*

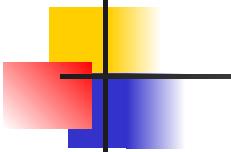
28-2 SIMPLE NETWORK MANAGEMENT PROTOCOL (SNMP)

The Simple Network Management Protocol (SNMP) is a framework for managing devices in an internet using the TCP/IP protocol suite. It provides a set of fundamental operations for monitoring and maintaining an internet.



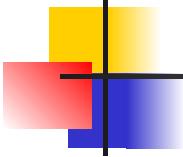
Note

SNMP defines the format of packets exchanged between a manager and an agent. It reads and changes the status (values) of objects (variables) in SNMP packets.



Note

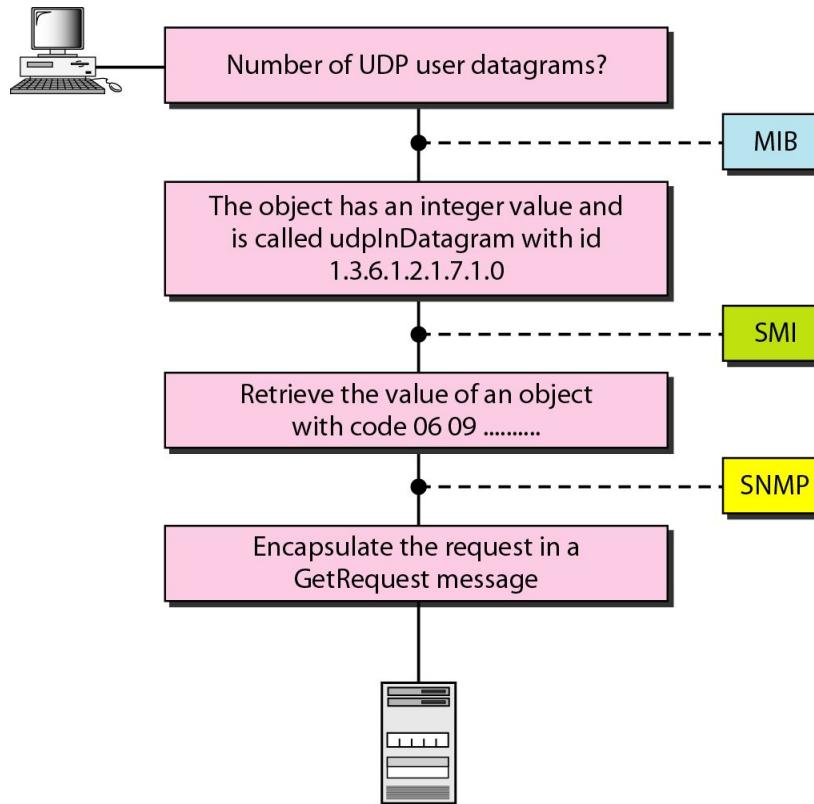
SMI defines the general rules for naming objects, defining object types (including range and length), and showing how to encode objects and values. SMI does not define the number of objects an entity should manage or name the objects to be managed or define the association between the objects and their values.



Note

MIB creates a collection of named objects, their types, and their relationships to each other in an entity to be managed.

Figure 28.4 Management overview



Note

All objects managed by SNMP are given an object identifier.

**The object identifier always starts with
1.3.6.1.2.1.**

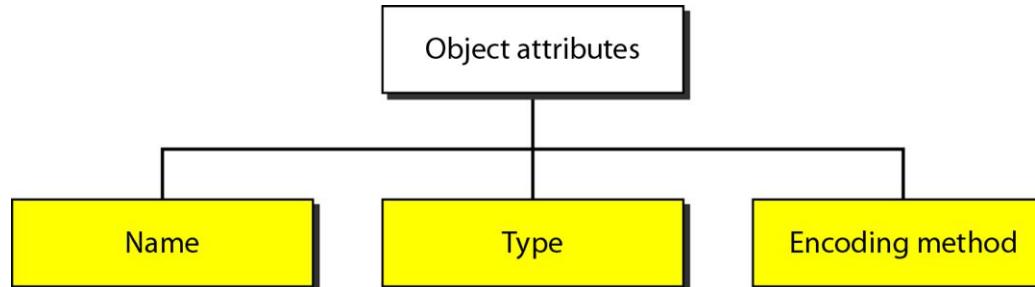


Figure 28.9 *Encoding format*

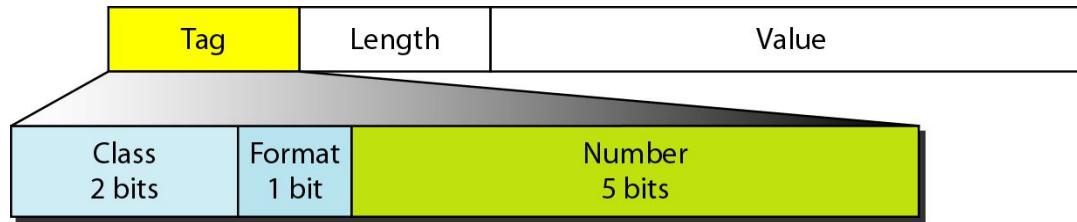
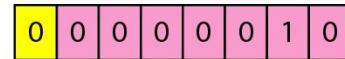


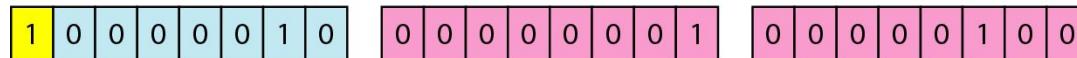
Table 28.2 *Codes for data types*

<i>Data Type</i>	<i>Class</i>	<i>Format</i>	<i>Number</i>	<i>Tag (Binary)</i>	<i>Tag (Hex)</i>
INTEGER	00	0	00010	00000010	02
OCTET STRING	00	0	00100	00000100	04
OBJECT IDENTIFIER	00	0	00110	00000110	06
NULL	00	0	00101	00000101	05
Sequence, sequence of	00	1	10000	00110000	30
IPAddress	01	0	00000	01000000	40
Counter	01	0	00001	01000001	41
Gauge	01	0	00010	01000010	42
TimeTicks	01	0	00011	01000011	43
Opaque	01	0	00100	01000100	44

Figure 28.10 *Length format*



a. The colored part defines the length (2).



b. The shaded part defines the length of the length (2 bytes);
the colored bytes define the length (260 bytes).

Figure 28.11 Example 28.1, INTEGER 14

02	04	00	00	00	0E
00000010	00000100	00000000	00000000	00000000	00001110
Tag (integer)	Length (4 bytes)	Value (14)			

Figure 28.12 Example 28.2, OCTET STRING “HI”

04	02	48	49
00000100	00000010	01001000	01001001
Tag (String)	Length (2 bytes)	Value (H)	Value (I)

Figure 28.13 Example 28.3, ObjectIdentifier 1.3.6.1

06	04	01	03	06	01
00000110	00000100	00000001	00000011	00000110	00000001
Tag (Objectld)	Length (4 bytes)	Value (1)	Value (3)	Value (6)	Value (1)

1.3.6.1 (iso.org.dod.internet)

Figure 28.20 SNMP PDUs

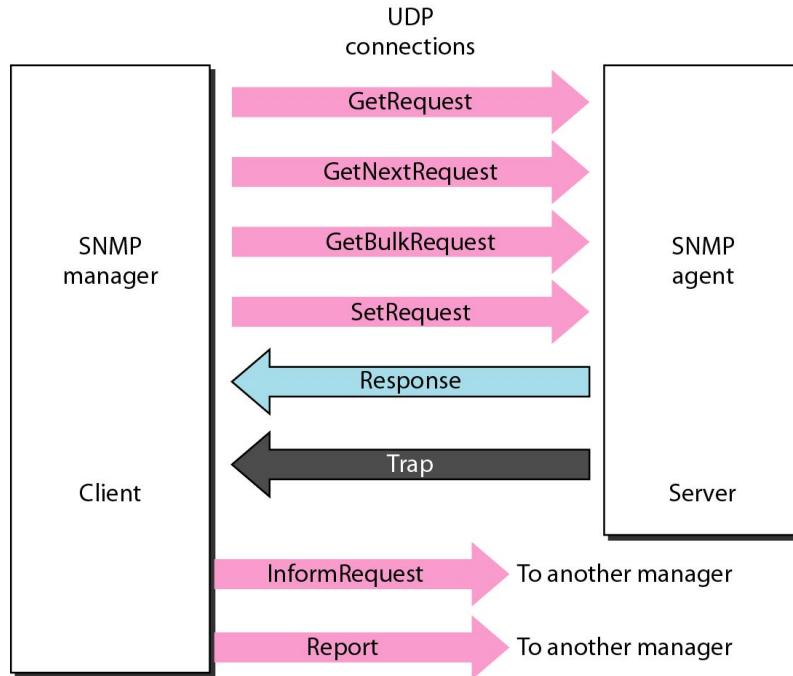
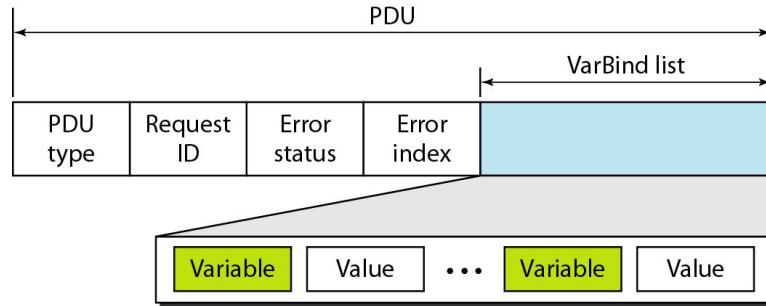


Figure 28.21 SNMP PDU format



Differences:

1. Error status and error index values are zeros for all request messages except GetBulkRequest.
2. Error status field is replaced by nonrepeater field and error index field is replaced by max-repetitions field in GetBulkRequest.

Table 28.3 *Types of errors*

<i>Status</i>	<i>Name</i>	<i>Meaning</i>
0	noError	No error
1	tooBig	Response too big to fit in one message
2	noSuchName	Variable does not exist
3	badValue	The value to be stored is invalid
4	readOnly	The value cannot be modified
5	genErr	Other errors

Figure 28.22 *SNMP message*

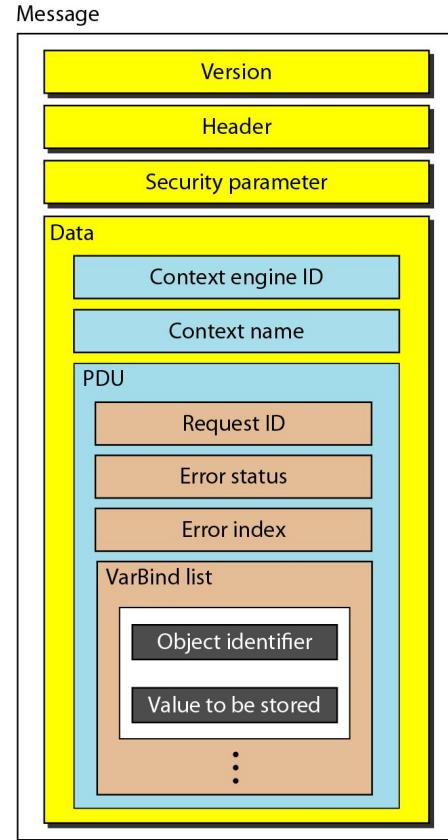


Table 28.4 *Codes for SNMP messages*

<i>Data</i>	<i>Class</i>	<i>Format</i>	<i>Number</i>	<i>Whole Tag (Binary)</i>	<i>Whole Tag (Hex)</i>
GetRequest	10	1	00000	10100000	A0
GetNextRequest	10	1	00001	10100001	A1
Response	10	1	00010	10100010	A2
SetRequest	10	1	00011	10100011	A3
GetBulkRequest	10	1	00101	10100101	A5
InformRequest	10	1	00110	10100110	A6
Trap (SNMPv2)	10	1	00111	10100111	A7
Report	10	1	01000	10101000	A8

Figure 28.23 Example 28.5

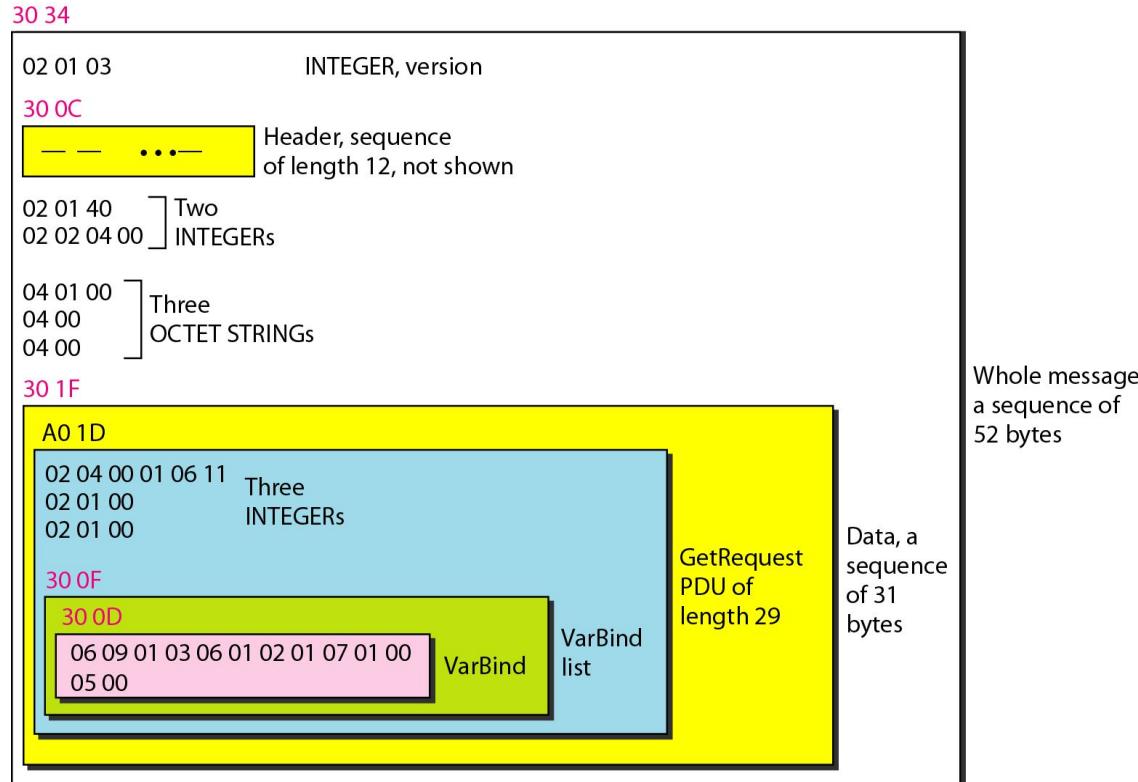
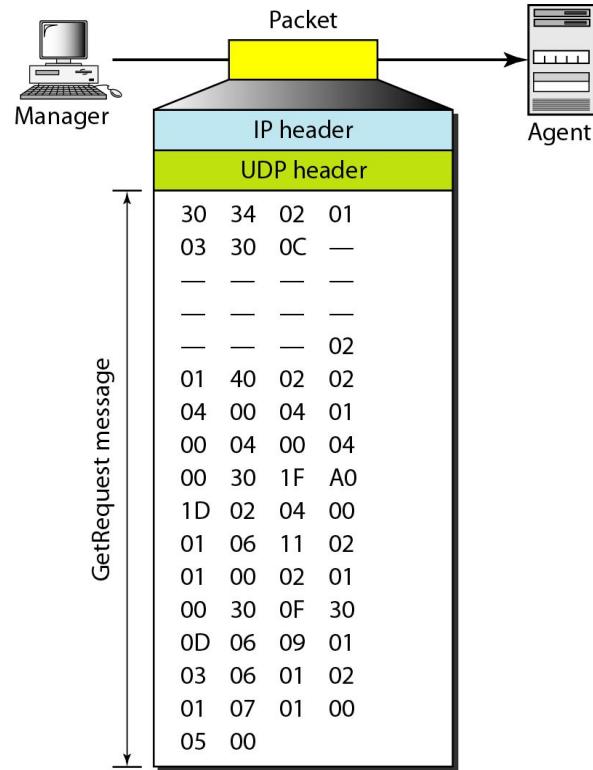


Figure 28.24 *GetRequest message*



WWW

The WWW today is a distributed client/server service, in which a client using a browser can access a service using a server. However, the service provided is distributed over many locations called sites.

Figure 27.1 *Architecture of WWW*

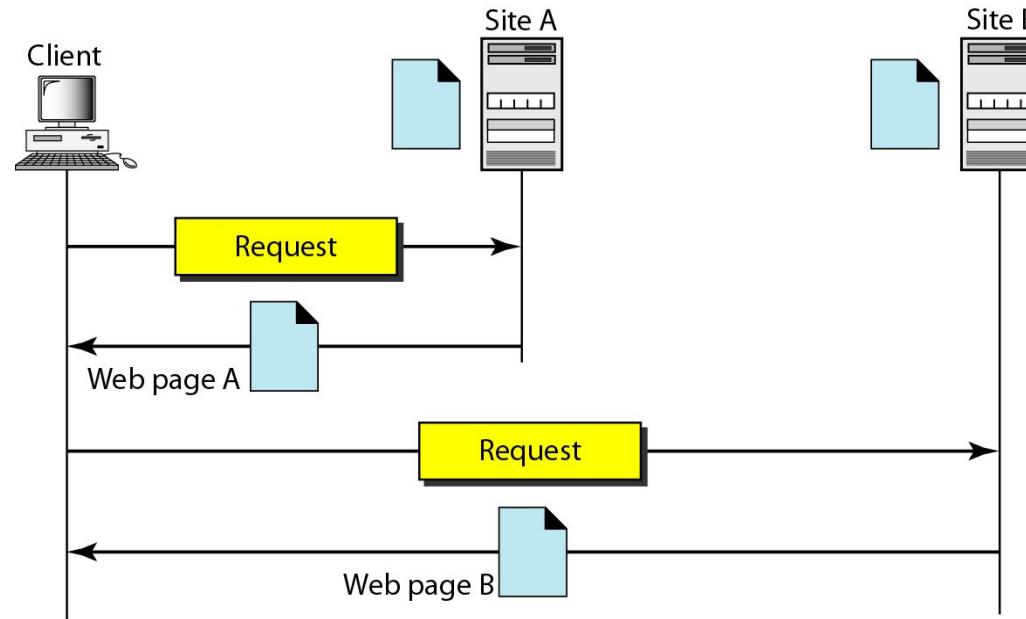
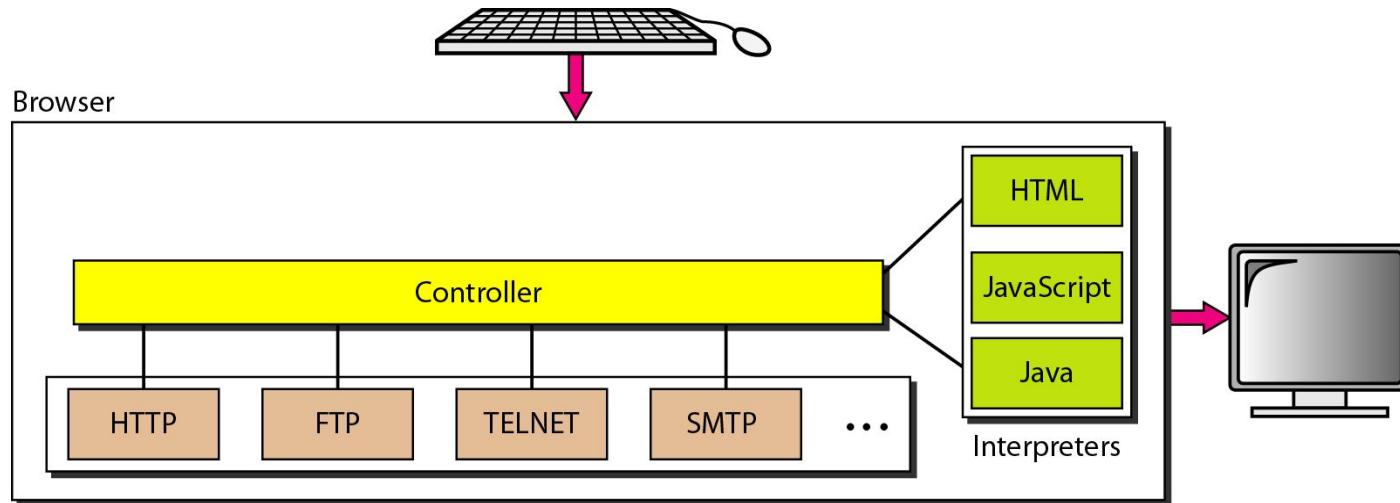


Figure 27.2 Browser



27-2 WEB DOCUMENTS

*The documents in the WWW can be grouped into three broad categories: **static**, **dynamic**, and **active**. The category is based on the time at which the contents of the document are determined.*

Figure 27.4 *Static document*

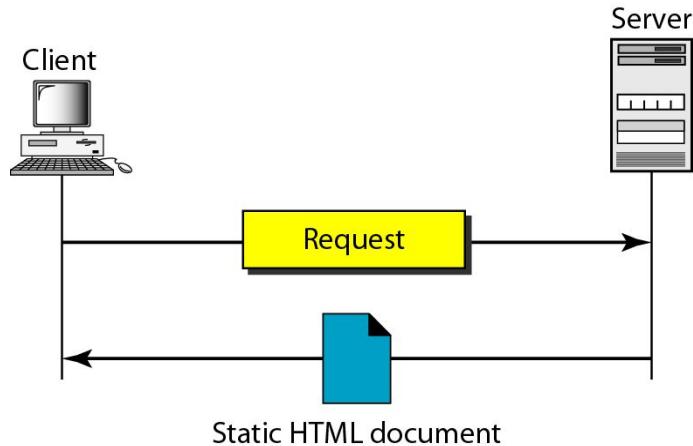


Figure 27.9 Dynamic document using server-site script

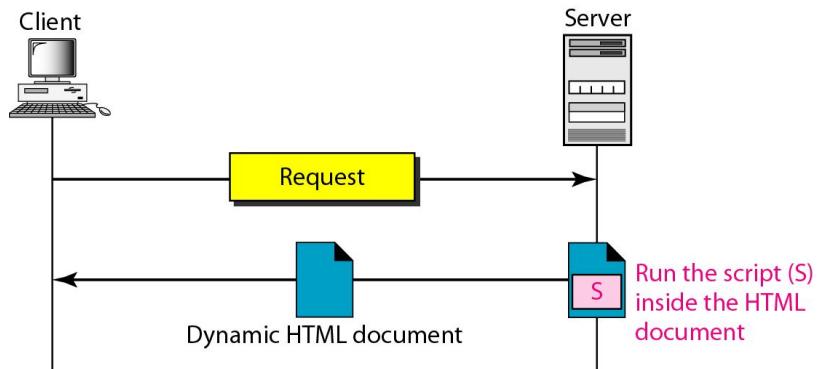


Figure 27.11 Active document using client-site script

