Module I (REF: Data Communications and Networking, 5th Ed. - by Behrouz A. Forouzan)

Introduction to computer networks – physical structure, topology, types - TCP/IP – architecture, Description of layers, addressing – wired LAN – Ethernet protocol – IEEE project 802 – Standard Ethernet – characteristics, addressing, implementation – wireless LAN – architectural comparison, characteristics, access control – IEEE 802.11 – architecture – LAN connecting devices – hub, switch, router – virtual LAN – architecture, membership, configuration.

NETWORKS

A network is the interconnection of a set of devices capable of communication. In this definition, a device can be a **host** (or an *end system* as it is sometimes called) such as a large computer, desktop, laptop, workstation, cellular phone, or security system. A device in this definition can also be a **connecting device** such as a router, which connects the network to other networks, a switch, which connects devices together, a modem (modulator-demodulator), which changes the form of data, and so on. These devices in a network are connected using wired or wireless transmission **media** such as cable or air.

Network Criteria

A network must be able to meet a certain number of criteria. The most important of these are performance, reliability, and security.

Performance: It can be measured in many ways, including *transit time* and *response time*. Transit time is the amount of time required for a message to travel from one device to another. Response time is the elapsed time between an inquiry and a response. The performance of a network depends on a number of factors, including the number of users, the type of transmission medium, the capabilities of the connected hardware, and the efficiency of the software. Performance is often evaluated by two networking metrics: *throughput* and *delay*. We often need more throughput and less delay.

Reliability: In addition to accuracy of delivery, network reliability is measured by the frequency of failure, the time it takes a link to recover from a failure, and the network's robustness in a catastrophe.

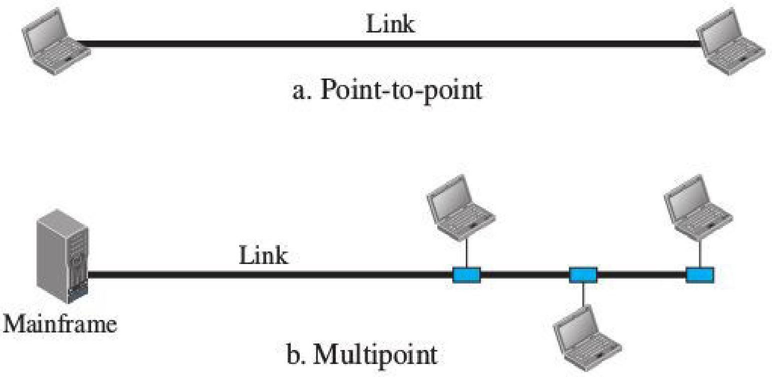
Security: Network security issues include protecting data from unauthorized access, protecting data from damage and development, and implementing policies and procedures for recovery from breaches and data losses.

Physical Structures

A network is two or more devices connected through links. A link is a communications pathway that transfers data from one device to another. For visualization purposes, it is simplest to imagine any link as a line drawn between two points. There are two possible types of connections: point-to-point and multipoint.

Point-to-Point: A point-to-point connection provides a dedicated link between two devices. The entire capacity of the link is reserved for transmission between those two devices. Most point-to-point connections use an actual length of wire or cable to connect the two ends, but other options, such as microwave or satellite links, are also possible. When we change television channels by infrared remote control, we are establishing a point-to-point connection between the remote control and the television’s control system.

Multipoint : A multipoint (also called multidrop) connection is one in which more than two specific devices share a single link. In a multipoint environment, the capacity of the channel is shared, either spatially or temporally. If several devices can use the link simultaneously, it is a *spatially shared* connection. If users must take turns, it is a *timeshared* connection.



a. Point-to-point

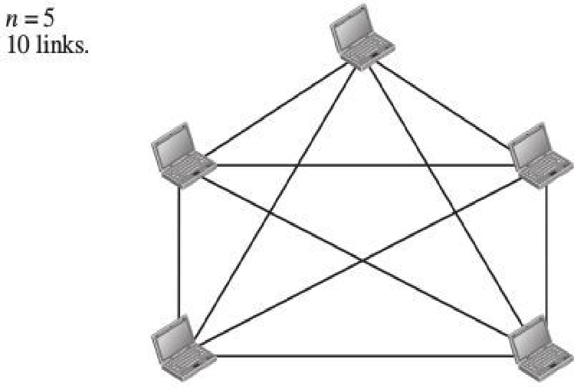
b. Multipoint

Physical Topology

The term physical topology refers to the way in which a network is laid out physically. The topology of a network is the geometric representation of the relationship of all the links and linking devices (usually called nodes) to one another. There are four basic topologies possible: mesh, star, bus, and ring. Combinations of these topologies also exists.

Mesh Topology:

In a mesh topology, every device has a dedicated point-to-point link to every other device. The term dedicated means that the link carries traffic only between the two devices it connects. In a network containing n devices (called *nodes*), each node must be connected to the other n – 1 nodes. We need n(n-1) physical links. However, if each physical link allows communication in both directions (duplex mode), we can divide the number of links by 2. In other words, we can say that in a mesh topology, we need n (n – 1) / 2 duplex-mode links. To accommodate that many links, every device on the network must have n – 1 input/output (I/O) ports to be connected to the other n – 1 stations.



<u>Advantages</u>:

1. Use of dedicated links guarantees that each connection can carry its own data load, thus eliminating the traffic problems that can occur when links must be shared by multiple devices.

2. A mesh topology is robust. If one link becomes unusable, it does not incapacitate the entire system.

3. Privacy or security. When a message travels along a dedicated line, only the intended recipient sees it.

4. Point-to-point links make fault identification and fault isolation easy.

<u>Disadvantages</u>:

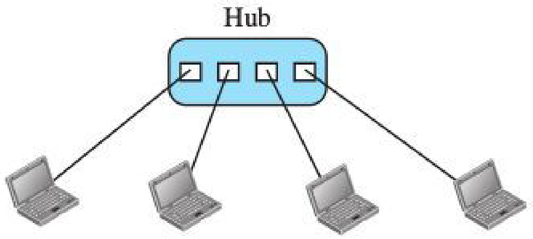
1. Amount of cabling. Every device must be connected to every other device, installation and reconnection are difficult.

2. The bulk wiring can be greater than the available space (in walls, ceilings, or floors) can accommodate.

3. The hardware required to connect each link (I/O ports and cable) can be expensive.

For these reasons a mesh topology is usually implemented in a limited fashion, for example, as a backbone connecting the main computers of a hybrid network that can include several other topologies. One practical example of a mesh topology is the connection of telephone regional offices in which each regional office needs to be connected to every other regional office.

Star Topology:



In a star topology, each device has a dedicated point-to-point link only to a **central controller**, usually called a **hub**. The devices are not directly linked to one another. Unlike a mesh topology, a star topology does not allow direct traffic between devices. The controller acts as an exchange: If one device wants to send data to another, it sends the data to the controller, which then relays the data to the other connected device. The star topology is used in local-area networks (LANs).

<u>Advantages</u>:

1. A star topology is less expensive than a mesh topology.

2. Each device needs only one link and one I/O port to connect it to any number of others. This factor also makes it easy to install and reconfigure.

3. Less cabling is needed.

4. Additions, moves, and deletions involve only one connection: between that device and the hub.

5. Robustness. If one link fails, only that link is affected. All other links remain active, as long as the hub is working.

6. Easy fault identification and fault isolation.

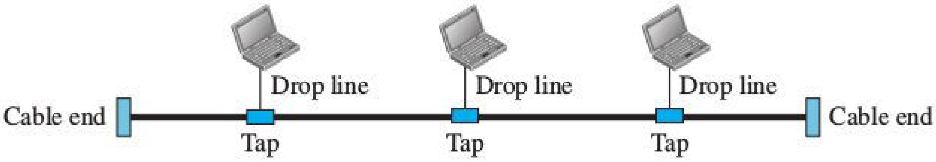
Disadvantages:

1. If the hub goes down, the whole system is dead.

2. More cabling is required in a star than in some other topologies (such as ring or bus).

Bus Topology:

A bus topology is a multipoint connection type. One long cable acts as a **backbone** to link all the devices in a network. Nodes are connected to the bus cable by drop lines and taps. A drop line is a connection running between the device and the main cable. A tap is a connector that either splices into the main cable or punctures the sheathing of a cable to create a contact with the metallic core. As a signal travels along the backbone, some of its energy is transformed into heat. Therefore, it becomes weaker and weaker as it travels farther and farther. For this reason there is a limit on the number of taps a bus can support and on the distance between those taps.



Bus topology was one of the first topologies used in the design of early local-area networks.

Advantages:

1. Ease of installation. Backbone cable can be laid along the most efficient path, then connected to the nodes by drop lines of various lengths.

2. Less cabling than mesh or star topologies.

Disadvantages:

1. Difficult reconnection and fault isolation. A bus is usually designed to be optimally efficient at installation. It can therefore be difficult to add new devices.

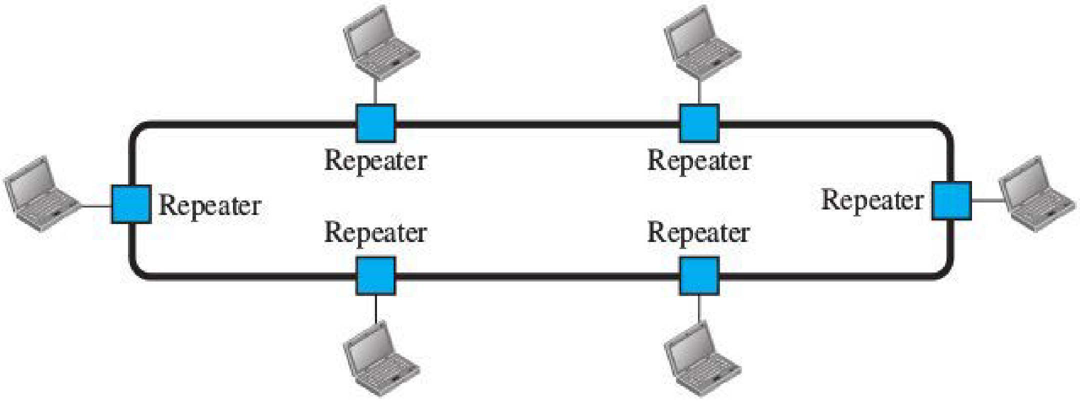
2. Signal reflection at the taps can cause degradation in quality (this degradation can be controlled by limiting the number and spacing of devices connected to a given length of cable).

3. Adding new devices may require modification or replacement of the backbone.

4. A fault or break in the bus cable stops all transmission, even between devices on the same side of the problem. The damaged area reflects signals back in the direction of origin, creating noise in both directions.

Ring Topology:

In a ring topology, each device has a dedicated point-to-point connection with only the two devices on either side of it. A signal is passed along the ring in one direction, from device to device, until it reaches its destination. Each device in the ring incorporates a repeater. When a device receives a signal intended for another device, its repeater regenerates the bits and passes them along. Ring topology was widely used when IBM introduced its Token Ring LAN.



Advantages:

1. Relatively easy to install and reconfigure. To add or delete a device requires changing only two connections.

2. Fault isolation is simplified. Generally, in a ring a signal is circulating at all times. If one device does not receive a signal within a specified period, it can issue an alarm. The alarm alerts the network operator to the problem and its location.

Disadvantages:

1. Unidirectional traffic can be a disadvantage. In a simple ring, a break in the ring (such as a disabled station) can disable the entire network. This weakness can be solved by using a dual ring or a switch capable of closing off the break.

NETWORK TYPES

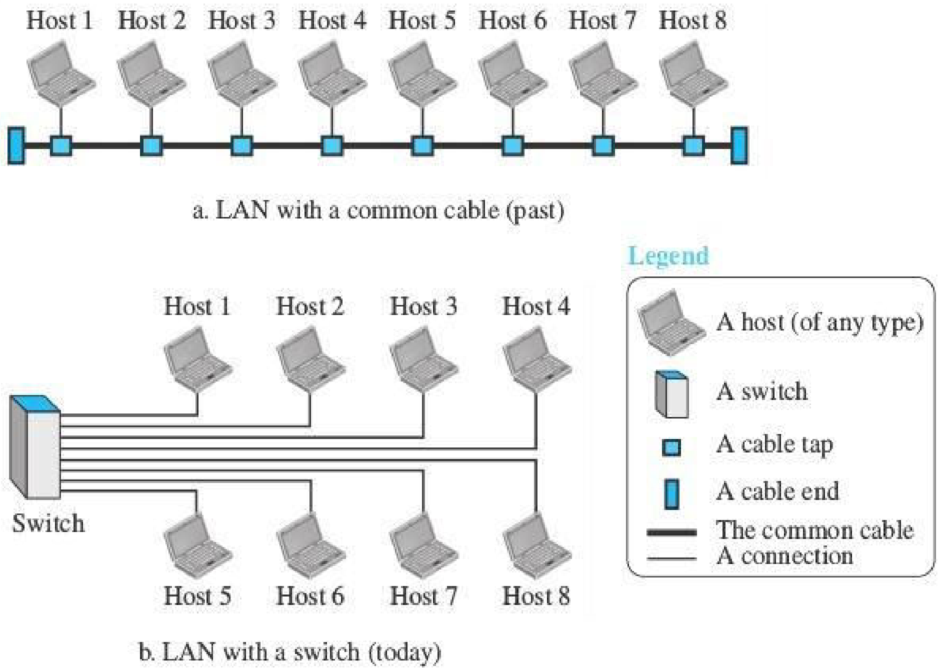
Networks can be divided into different types based on a variety of criteria. Most general classification is based on the size of the network or the geographical area where the network spans. The two types of such networks are Local Area Networks (LANs) and Wide Area Networks (WANs).

LANs:

A **local area network (LAN)** is usually privately owned and connects some hosts in a single office, building, or campus. Depending on the needs of an organization, a LAN can be as simple as two PCs and a printer in someone's home office, or it can extend throughout a company and include audio and video devices. Each host in a LAN has an identifier, an address, that uniquely defines the host in the LAN. A packet sent by a host to another host carries both the source host and the destination host's addresses.

In the past, all hosts in a network were connected through a common cable, which meant that a packet sent from one host to another was received by all hosts. The intended recipient kept the packet; the others dropped the packet. Today, most LANs use a smart connecting switch, which is able to recognize the destination address of the packet and guide the packet to its destination without sending it to all other hosts.

WANs:

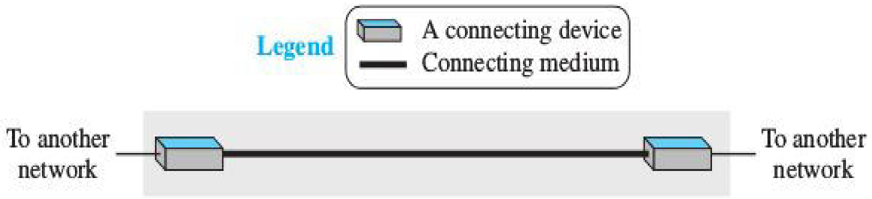


A wide area network (WAN) is an interconnection of devices capable of communication, which are geographically widely separated, spanning a town, a state, a country, or even the world. A LAN interconnects hosts; a WAN interconnects connecting devices such as switches, routers, or modems. A LAN is normally privately owned by the organization that uses it; a WAN is normally created and run by communication companies and leased by an organization that uses it.

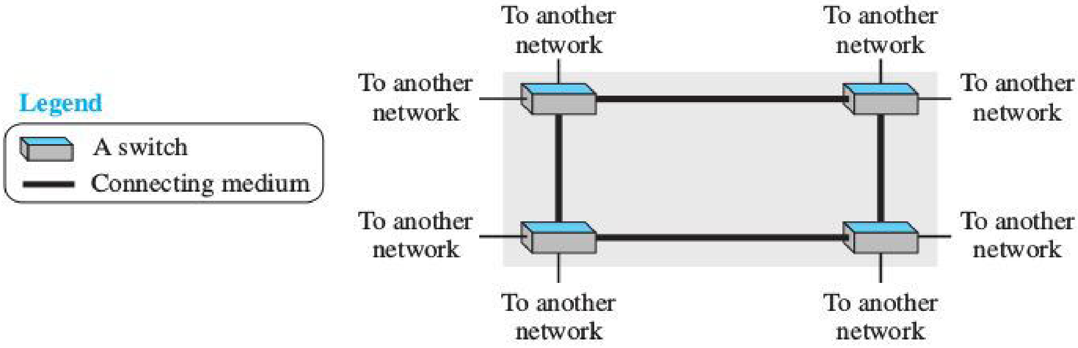
We see two distinct examples of WANs today: point-to-point WANs and switched WANs.

Point-to-Point WAN :

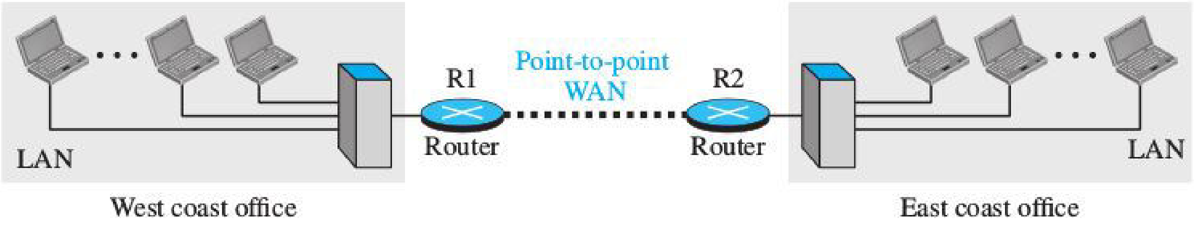
A point-to-point WAN is a network that connects two communicating devices through a transmission media (cable or air).



Switched WAN: A switched WAN is a network with more than two ends. A switched WAN is used in the backbone of global communication today. A switched WAN is a combination of several point-to-point WANs that are connected by switches.



Internetwork: When two or more networks are connected, they make an internetwork, or internet. As an example, assume that an organization has two offices, one on the east coast and the other on the west coast. Each office has a LAN that allows all employees in the office to communicate with each other. To make the communication between employees at different offices possible, the management leases a point-to-point dedicated WAN from a service provider, such as a telephone company, and connects the two LANs. Now the company has an internetwork, or a private internet. When a host in the west coast office sends a message to another host in the same office, the router blocks the message, but the switch directs the message to the destination. On the other hand, when a host on the west coast sends a message to a host on the east coast, router R1 routes the packet to router R2, and the packet reaches the destination.

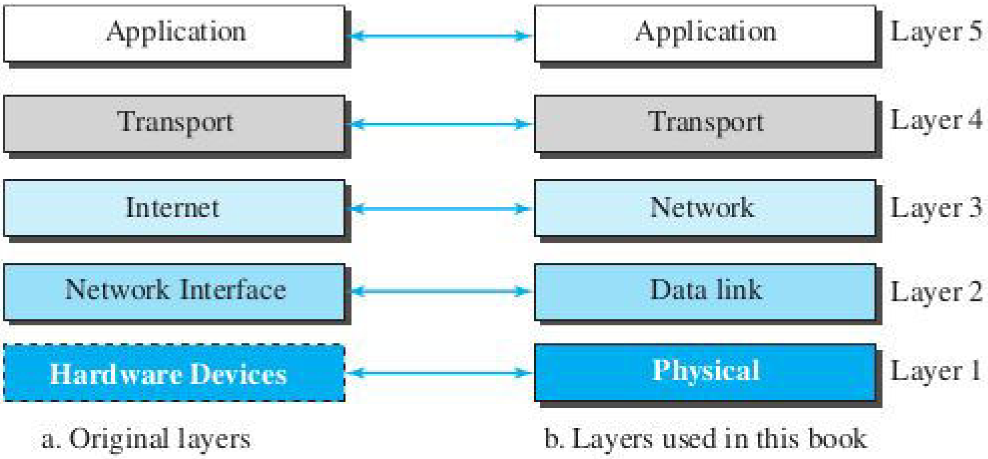


TCP/IP PROTOCOL SUITE

TCP - Transmission Control Protocol

IP - Internet Protocol

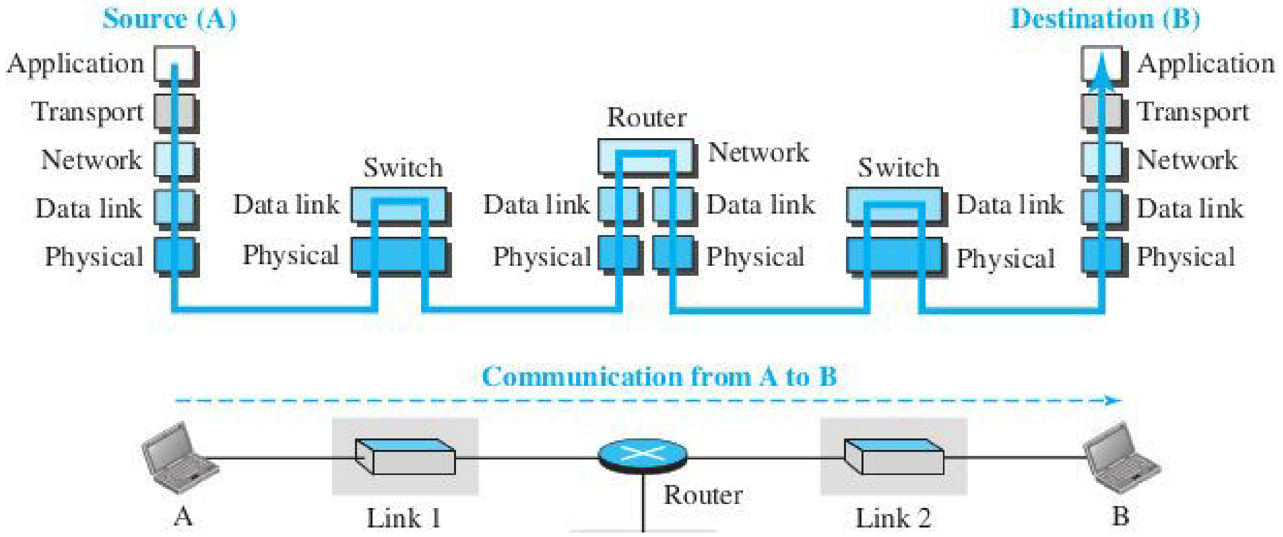
TCP/IP is a protocol suite (a set of protocols organized in different layers) used in the Internet today. It is a hierarchical protocol made up of interactive modules, each of which provides a specific functionality. The term hierarchical means that each upper level protocol is supported by the services provided by one or more lower level protocols. The original TCP/IP protocol suite was defined as four software layers built upon the hardware. Today, however, TCP/IP is thought of as a five-layer model.



Layered Architecture

To know the layered architecture and how the layers act in the network consider a network with three LANs connected to a router through three links (switches). The network diagram and layered architectures are given below.

Let us assume that computer A communicates with computer B. As the figure shows, we have five communicating devices in this communication: source host (computer A), the link-layer switch in link 1, the router, the link-layer switch in link 2, and the destination host (computer B). The layers associated with the devices are based on the function they provide.



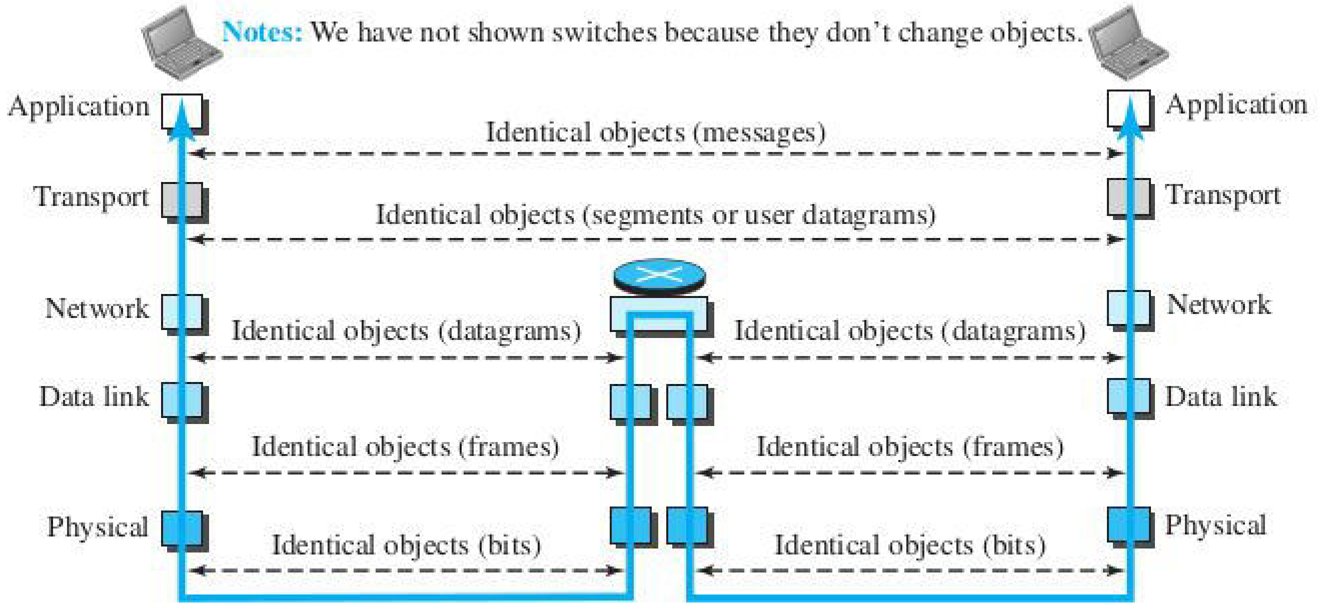
The two hosts are involved in all five layers; the source host needs to create a message in the application layer and send it down the layers so that it is physically sent to the destination host. The destination host needs to receive the communication at the physical layer and then deliver it through the other layers to the application layer.

The router needs to route the packets from one network to another, hence it does not need transport or application layers.

The switches are used to connect systems in the same network across links, hence they need only upto the data link layer.

LAYERS IN THE TCP/IP PROTOCOL SUITE

To better understand the duties of each layer, we need to think about the logical connections between layers.


Notes: We have not shown switches because they don't change objects.
Application
Transport
Network
Data link
Physical
Identical objects (messages)
Identical objects (segments or user datagrams)
Identical objects (datagrams)
Identical objects (datagrams)
Identical objects (frames)
Identical objects (frames)
Identical objects (bits)
Identical objects (bits)
Application
Transport
Network
Data link
Physical

1. Physical Layer

It is responsible for carrying individual **bits** in a frame across the link. Although the physical layer is the lowest level in the TCP/IP protocol suite, the communication between two devices at the physical layer is a logical communication because under the physical layer the transmission media carries the bits as electrical, optical or electromagnetic signals. Thus the logical unit between two physical layers in two devices is a **bit**.

2. Data-link Layer

The internet is made up of several links connected by routers. It is the duty of the data link layer to move the datagrams across the links, even if the protocols are different. TCP/IP does not define any specific protocol for the data-link layer. It supports all the standard and proprietary protocols. The data-link layer takes a datagram from the network layer and encapsulates it in a packet called a **frame**.

3. Network Layer

The network layer is responsible for creating a connection between the source computer and the destination computer. The communication at the network layer is host-to-host. But in the routers its function is to find and route the best path from source to destination. We can say that the network layer is responsible for host-to-host communication and routing the packet through possible routes.

The major protocol in the network layer is the Internet Protocol (IP). IP defines the format of the packet and structure of addresses at the network layer. These packets are called **datagrams**. IP is also responsible for routing of packets. IP is a connectionless protocol that provides no flow control, no error control, and no congestion control services. If the data transfer needs these services, the higher layers should provide these.

The network layer also includes unicast (one-to-one) and multicast (one-to-many) routing protocols which create forwarding tables for routers to help IP in the routing process.

There are some auxiliary protocols to help IP. The Internet Control Message Protocol (ICMP) helps IP to report some problems when routing a packet. The Internet Group Management Protocol (IGMP) is another protocol that helps IP in multitasking. The Dynamic Host Configuration Protocol (DHCP) helps IP to get the network-layer address for a host. The Address Resolution Protocol (ARP) is a protocol that helps IP to find the link-layer address of a host or a router when its network-layer address is given.

4. Transport Layer

The logical connection at the transport layer is end-to-end. It’s duty is to get the messages from application layer programs and transport them to the transport layer of the destination, where the packets are delivered to corresponding application layer programs. Since there are many types of programs running in the application layer, the transport layer should be independent of the application layer.

The main protocol in the transport layer is a connection oriented protocol called the Transmission Control Protocol (TCP). TCP establishes a connection (logical pipe) between transport layers at two hosts before transferring data. TCP provides flow control, error control and congestion control. The TCP data packets are called **segments**.

Another protocol is the User Datagram Protocol (UDP) which is connection-less. It is a simple protocol and does not provide flow control, error control and congestion control. USP packets are called **user datagrams**. In UDP, each user datagram is an independent entity without being related to the previous or the next one. UDP is used for application that sends short messages and where error control and flow control are not mandatory.

A new protocol called Stream Control Transmission Protocol (SCTP) is used for multimedia applications.

5. Application Layer

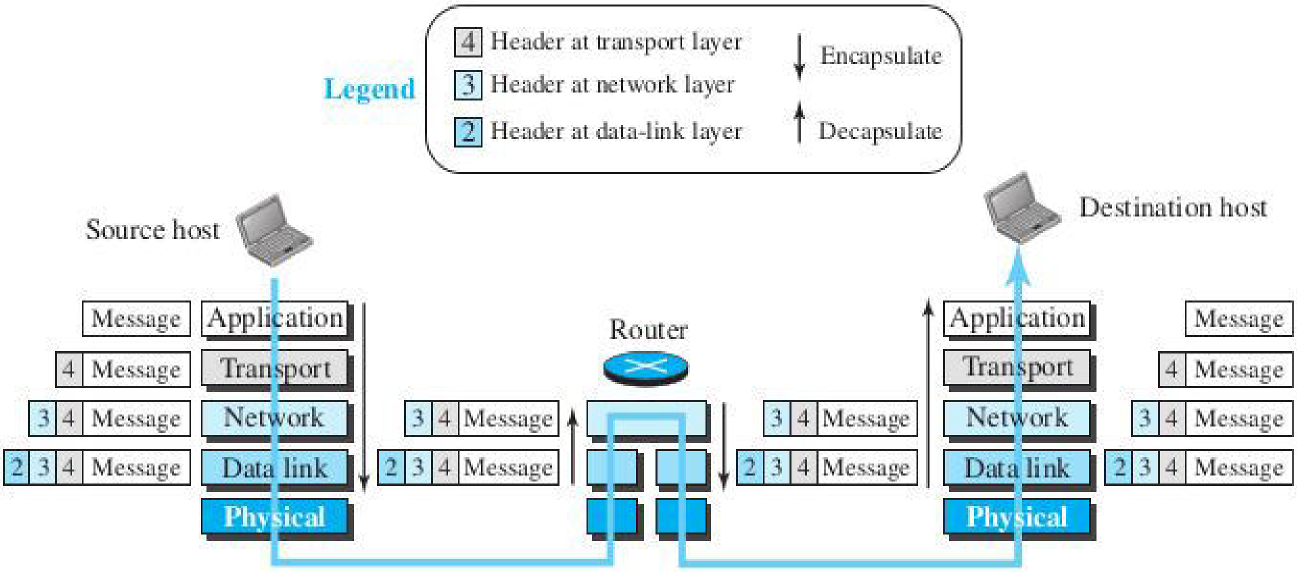
Process-to-process communication is the duty of the application layer. Thus the logical connection between the two application layers is end-to-end. The two application layers exchange **messages** between each other as though there were a bridge between the two layers. Common protocols in this layer are as follows;

- ➢ The Hypertext Transfer Protocol (HTTP) is a vehicle for accessing the World Wide Web (WWW).
- ➢ The Simple Mail Transfer Protocol (SMTP) is the main protocol used in electronic mail (e-mail) service.
- ➢ The File Transfer Protocol (FTP) is used for transferring files from one host to another.
- ➢ The Terminal Network (TELNET) and Secure Shell (SSH) are used for accessing a site remotely.
- ➢ The Simple Network Management Protocol (SNMP) is used by an administrator to manage the Internet at global and local levels.
- ➢ The Domain Name System (DNS) is used by other protocols to find the network-layer address of a computer.
- ➢ The Internet Group Management Protocol (IGMP) is used to collect membership in a group.

> In short, the duty of the application, transport, and network layers is end-to-end. However, the duty of the data-link and physical layers is hop-to-hop, in which a hop is a host or router. In other words, the domain of duty of the top three layers is the internet, and the domain of duty of the two lower layers is the link.

ENCAPSULATION AND DECAPSULATION

Encapsulation happens at the source and decapsulation happens at the destination. Since he routers connects different networks, they have to decapsulate incoming packets and encapsulate them with changed header information. The concept of these in figured as follows with two systems and a router.



There may be intermediate links (switches), but since switches does not change the header information, they do not need encapsulation and decapsulation and hence not included in the figure.

Encapsulation at the Source Host

1. The application layer passes the **message** to the transport layer. These messages may or may not contain header or trailer, but transport layer consider these as messages only.
2. The transport layer consider the massage as payload and adds the transport layer header to it ('4' in the figure). The header contains information necessary for end-to-end delivery of message such as source and destination addresses, flow control, error control, or congestion control information etc. The result is the transport-layer packet, which is called the **segment** (in TCP) and the **user datagram** (in UDP). The transport layer then passes the packet to the network layer.
3. The network layer takes the transport-layer packet as data or payload and adds its own header to the payload. The header contains the addresses of the source and destination hosts (IP address) and some more information used for error checking of the header, fragmentation information, and so on. The result is the network-layer packet, called a **datagram**. The network layer then passes the packet to the data-link layer.
4. The data-link layer takes the network-layer packet as data or payload and adds its own header, which contains the link-layer address of the host or the next hop (the router). The result is the link-layer packet, which is called a **frame**. The frame is passed to the physical layer for transmission.

Decapsulation and Encapsulation at the Router

At the router, both decapsulation and encapsulation happens because the router is connected to two or more links.

1. After the set of bits are delivered to the data-link layer, this layer decapsulates the

datagram from the frame and passes it to the network layer.

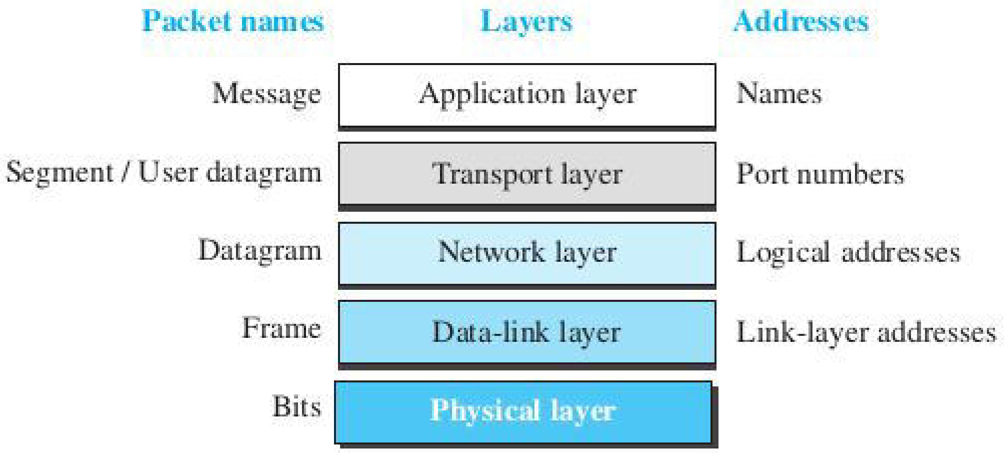
2. The network layer only inspects the source and destination addresses in the *datagram header* and consults its forwarding table to find the next hop to which the datagram is to be delivered. The contents of the datagram should not be changed by the network layer in the router unless there is a need to fragment the datagram if it is too big to be passed through the next link. The datagram is then passed to the data-link layer of the next link.
3. The data-link layer of the next link encapsulates the datagram in a frame and passes it to the physical layer for transmission.

Decapsulation at the Destination Host

At the destination host, each layer only decapsulates the packet received, removes the payload, and delivers the payload to the next-higher layer protocol until the message reaches the application layer. Decapsulation in the host also involves error checking.

Addressing

Any communication that involves two parties needs two addresses: source address and destination address. Even if there are five layers there are only four pair of addresses. Since physical layer just exchanges bits,it does not need any address. The relation between the packet names,layers and addresses are shown in the figure.



Packet names	Layers	Addresses
Message	Application layer	Names
Segment / User datagram	Transport layer	Port numbers
Datagram	Network layer	Logical addresses
Frame	Data-link layer	Link-layer addresses
Bits	Physical layer	

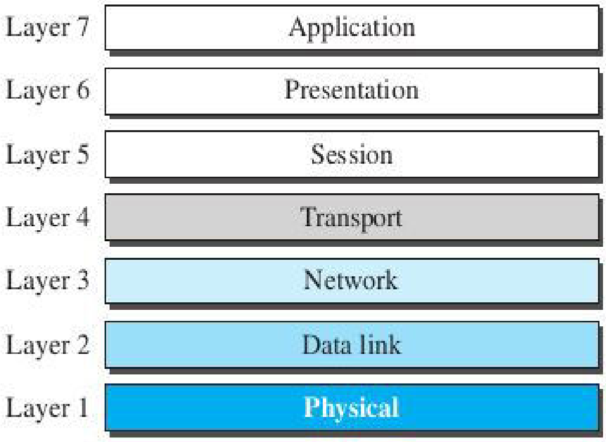
1. The application layer names can be like, example.com, someone@mail.com, etc.
2. The port number define different application level programs running at the same time at the source and destination.
3. Network layer addresses (logical addresses, eg: IP Addresses) uniquely defines the connection of a device in the network.
4. The link-layer addresses, sometimes called MAC addresses or physical addresses, are locally defined addresses, each of which defines a specific host or router in a network (LAN or WAN).

ISO-OSI (International Organization for Standardization - Open Systems Interconnection)

An open system is a set of protocols that allows any two different systems to communicate regardless of their underlying architecture. The purpose of the OSI model is to show how to facilitate communication between different systems without requiring changes to the logic of the underlying hardware and software.

OSI is a model for understanding and designing a network architecture that is flexible, robust, and interoperable.

Layers in OSI model:



OSI vs TCP/IP

1. Number of layers is 7 in ISO-OSI but 4 in TCP/IP.
2. In OSI model, the top three layers are separate, but in TCP/IP these three layers are combined into one application layer.
3. TCP/IP has more than one transport-layer protocol. Some of the functionalities of the session layer are available in some of the transport-layer protocols.
4. Application layer in TCP/IP is a collection of various protocols and different types of applications can be developed at this layer.

Wired LANs: Ethernet

The TCP/IP does not define any protocol for physical layer and data link layer, because it allows any protocol that can provide service to network layer. Data link layer and physical layer are the major layers in LANs and WANs. The major LAN technologies were Ethernet, Token Ring, Token Bu and FDDI (Fiber Distributed Data Interface). Since Ethernet updated itself with time to time, other networks disappeared slowly.

IEEE Project 802

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. The IEEE 802 standard specifies functions of the physical layer and the data-link layer of major LAN protocols.

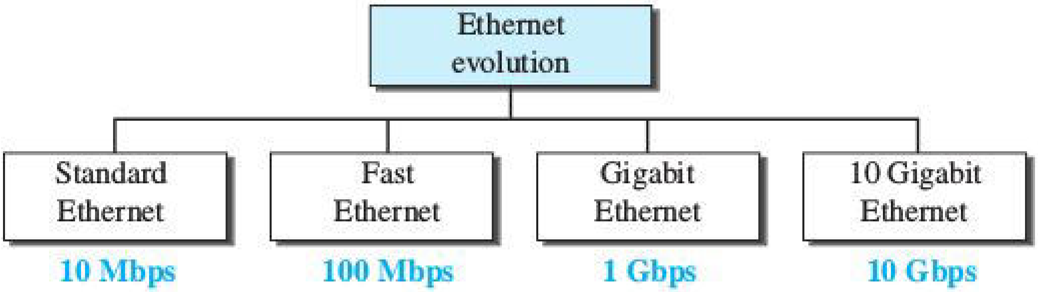
The IEEE has subdivided the data-link layer into two sublayers: **logical link control (LLC)** and **media access control (MAC)**. IEEE has also created several physical-layer standards for different LAN protocols.

Logical Link Control (LLC) sublayer: Flow control, error control and *a part of framing* are done by LLC. The LLC provides a single link control protocol for all IEEE LANs. Hence LLC can interconnect different LANs. It makes the MAC sublayer transparent.

Media Access Control (MAC) sublayer: It defines specific access method for each LAN. For example, it defines CSMA/CD as the media access method for Ethernet LANs and defines the token-passing method for Token Ring and Token Bus LANs. Also a part of framing is done by MAC sublayer.

Ethernet Evolution

From 1970, the Ethernet has gone through different generations. They are as shown in the figure.



STANDARD ETHERNET

It is the Ethernet with 10 Mbps speed. Even though the Ethernet has gone through different generations since the Standard Ethernet, some functionalities are still present in the modern Ethernet.

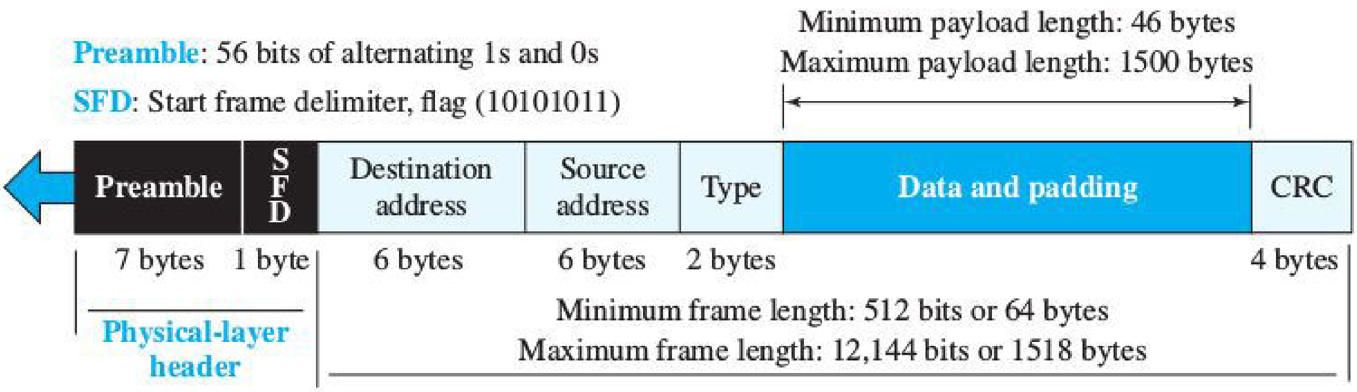
Standard Ethernet Characteristics

- Ethernet provides a *connectionless* service, which means each frame sent is independent of the previous or next frame. Ethernet has no connection establishment or connection termination phases.
- The sender sends a frame whenever it is created; the receiver may or may not be ready for it.
- The receiver may be slow and hence cannot process all the incoming frames, which may result in dropping frames. If a frame drops, the sender will not know about it.
- The network layer protocol is IP and uses the service of Ethernet. IP also is connectionless, and hence it will not know about the drop of frames.
- If the transport layer is also a connectionless protocol, such as UDP, the frame is lost and problem may only come for the application layer.
- However, if the transport layer is TCP, the sender TCP does not receive acknowledgment for its segment and sends it again.
- Ethernet is also *unreliable* like IP and UDP. If a frame is corrupted during transmission and the receiver finds out about the corruption, the receiver drops the frame silently. It is the

duty of high-level protocols to find out about it.

Standard Ethernet Frame Format

The Ethernet contains seven fields, as follows.


Preamble: 56 bits of alternating 1s and 0s
SFD: Start frame delimiter, flag (10101011)
Minimum payload length: 46 bytes
Maximum payload length: 1500 bytes
Preamble
S F D
Destination address
Source address
Type
Data and padding
CRC
7 bytes
1 byte
6 bytes
6 bytes
2 bytes
4 bytes
Physical-layer header
Minimum frame length: 512 bits or 64 bytes
Maximum frame length: 12,144 bits or 1518 bytes

1. Preamble - (7 bytes)

This field contains 7 bytes (56 bits) of alternating 0s and 1s (ie, 101010101…) that tells the receiving system that a frame is coming. This enables the receiver to synchronize its clock if it is out of synchronization. The receiver may skip some bits before the preamble ends, even the end of this pattern gives an alert and a timing pulse to the receiver. *The preamble is added by the physical layer of the sender and it is not part of the frame.*

2. Start Frame Delimiter (SFD) - (1 byte)

This is a one byte flag (101010**11**) to indicate that the next bit after the SFD is the start of the frame. The SFD warns the receiver that this is the last chance for synchronization. *The SFD field also is added by the physical layer.*

3. Destination address (DA) - (6 bytes)

It is the 6 byte (48 bits) link layer address of the destination station *(link layer address is otherwise called as physical address or hardware address or MAC address)*. When the receiver sees its own link-layer address in this field, or a multicast address for a group that the receiver is a member of, or a broadcast address, it decapsulates the data from the frame and passes the data to the upper layer protocol.

4. Source Address (SA) - (6 bytes)

It is the 6 byte (48 bits) link layer address of the sender of the packet.

5. Type - (2 bytes)

This field defines the upper-layer protocol whose packet is encapsulated in the frame. This protocol can be IP, ARP, OSPF, and so on. This field is used for multiplexing and demultiplexing.

6. Data - (46 bytes to 1500 bytes)

This field contains the payload delivered by the upper layer. The length of this data must be a minimum of 46 bytes and a maximum of 1500 bytes. If the data is below 46 bytes, the upper layer should append zeros to the data to make it 46 bytes. The upper layer in the receiver will remove these extra zeros. If the data length is greater than 1500 bytes, it should be defragmented and encapsulated in more than one frame. The receiver will rejoin all fragments.

7. Cyclic Redundancy Check (CRC) - (4 bytes)

This field contains error detection information, in this case a CRC-32. The CRC is calculated over the *addresses, types, and data field*. If the receiver calculates the CRC and finds that it is not zero (corruption in transmission), it discards the frame.

Frame Length

The length of the payload (data field) in a Standard Ethernet frame ranges from 46 to 1500. The other fields in the frame constitutes 18 bytes (6+6+2+4). Hence the minimum length of a Standard Ethernet frame should be 46+18=64 bytes. The maximum length should be 1500+18=1518 bytes.

The maximum length restriction has two reasons:

1. Memory was very expensive when Ethernet was designed; a maximum length restriction helped to reduce the size of the buffer.
2. The maximum length restriction prevents one station from monopolizing the shared medium, blocking other stations that have data to send.

Minimum frame length: 64 bytes **Minimum data length: 46 bytes**
Maximum frame length: 1518 bytes **Maximum data length: 1500 bytes**

Addressing in Standard Ethernet

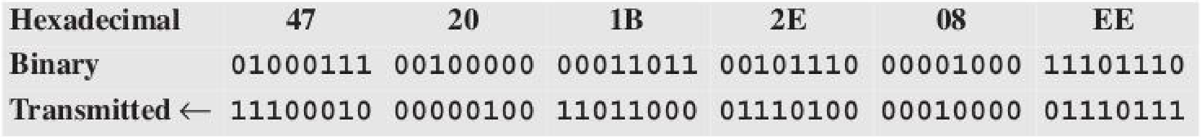
Each station on an Ethernet network (such as a PC, workstation, or printer) has its own **network interface card (NIC)**. The NIC fits inside the station and provides the station with a **unique** link-layer address. The Ethernet address is 6 bytes (48 bits), normally written in **hexadecimal notation**, with a colon between the bytes. For example, the following shows an Ethernet MAC address:

4A:30:10:21:10:1A

<u>Transmission of Address Bits:</u>

The way the addresses are sent out online is different from the way they are written in hexadecimal notation. The transmission is **left to right, byte by byte**; however, **for each byte, the least significant bit is sent first and the most significant bit is sent last.** This means that the bit that defines an address as unicast or multicast arrives first at the receiver. This helps the receiver to immediately know if the packet is unicast or multicast.

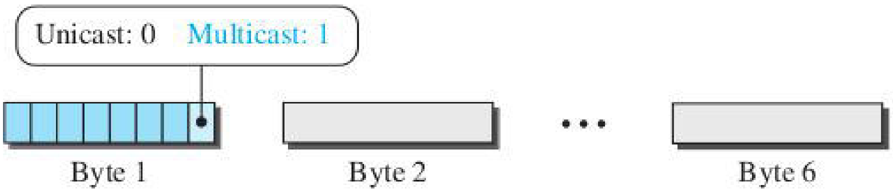
Example: Show how the address 47:20:1B:2E:08:EE is sent out online.



Hexadecimal	47	20	1B	2E	08	EE
Binary	01000111	00100000	00011011	00101110	00001000	11101110
Transmitted ←	11100010	00000100	11011000	01110100	00010000	01110111

To understand Unicast, Multicast, and Broadcast Addresses:

A source address is always a unicast address—the frame comes from only one station. The destination address can be unicast, multicast, or broadcast. If the least significant bit of the first byte in a destination address is 0, the address is unicast; otherwise, it is multicast or broadcast. Hence when the first bit is received, the receiver can understand whether it is unicast or not.



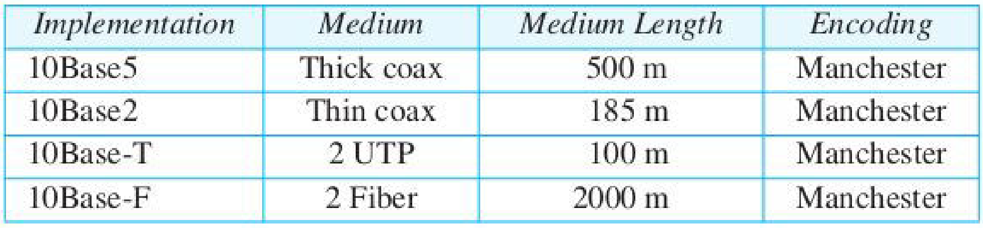
Also the broadcast address means all the forty-eight bits are 1s and is meant to be delivered for all the systems in the LAN.

Example:

1) 4A:30:10:21:10:1A - First byte is 0100 101***0*** - hence uni cast
2) 47:20:1B:2E:08:EE - First byte is 0100 011***1*** - hence multicast
3) FF:FF:FF:FF:FF:FF - All bytes are 1s, hence broadcast

Standard Ethernet Implementation

The most popular implementations are the below four types.

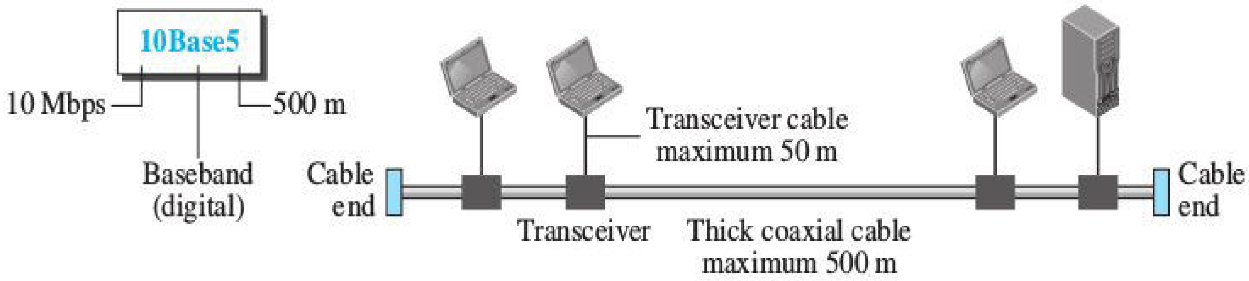


Implementation	*Medium*	*Medium Length*	*Encoding*
10Base5	Thick coax	500 m	Manchester
10Base2	Thin coax	185 m	Manchester
10Base-T	2 UTP	100 m	Manchester
10Base-F	2 Fiber	2000 m	Manchester

In the nomenclature 10BaseX, the number defines the data rate (10 Mbps), the term Base means baseband (digital) signal, and X approximately defines either the maximum size of the cable in 100 meters (for example 5 for 500 or 2 for 185 meters) or the type of cable, T for unshielded twisted pair cable (UTP) and F for fiber-optic.

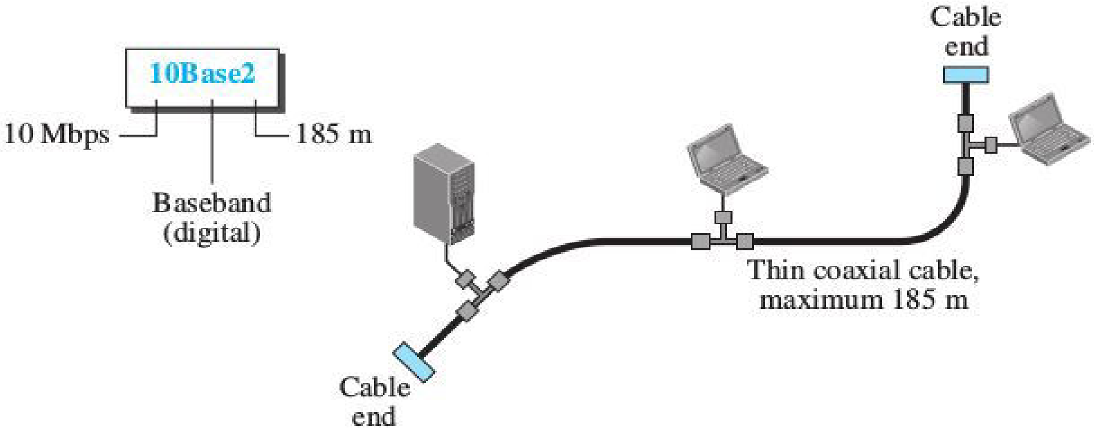
1) 10Base5: Thick Ethernet

It was the first implementation which used thick coaxial cables to connect systems in bus topology. The name thick Ethernet (thicknet) because the coaxial cable was very thick (like garden hose) and very hard to bend. It had a maximum segment length of 500m. If length greater than 500m is needed, repeaters should be used, otherwise cause signal degradation. The transceivers (transmitter/receiver) are connected to the cable via vampire taps. The transceiver is responsible for transmitting, receiving, and detecting collisions. The transceiver is connected to the station via a transceiver cable that provides separate paths for sending and receiving. This means that collision can only happen in the coaxial cable.



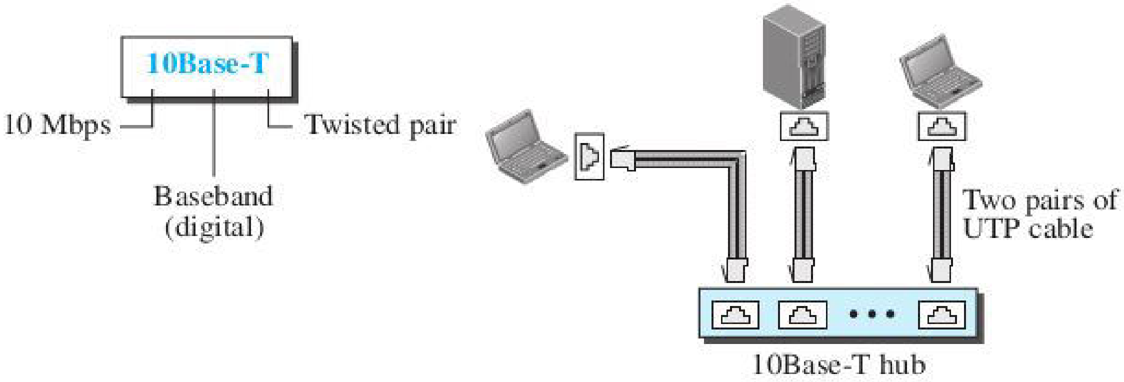
2) 10Base2: Thin Ethernet

It is a bus topology network with thin flexible coaxial cable, hence it is called thin Ethernet or Thinnet. It has a maximum segment length of 185m, more than which needs repeaters. The transceiver is normally part of the network interface card (NIC) inside the system. The collision here occurs in the thin coaxial cable. This implementation is more cost effective than 10Base5 because thin coaxial cable is less expensive than thick coaxial and the *tee* connections are much cheaper than *taps*. Installation is simpler because the thin coaxial cable is very flexible.



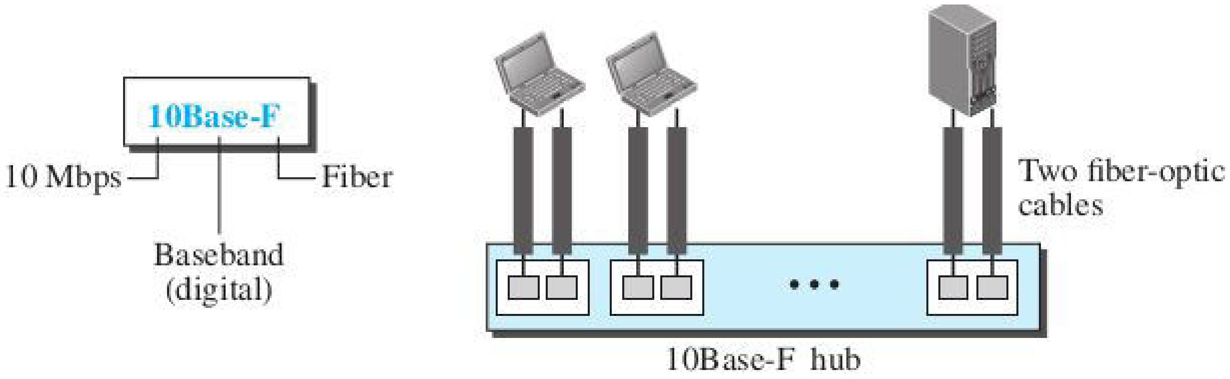
3) 10Base-T (Twisted-Pair Ethernet)

This implementation uses a physical Star topology. All the stations are connected to a central hub via twisted pair cable. The cables have separate pairs of wires for sending and receiving the data. Hence there will be no collision in the cable and the collision happens in the hub. The maximum cable length for a single UTP cable must be 100m to minimize the effect of attenuation.



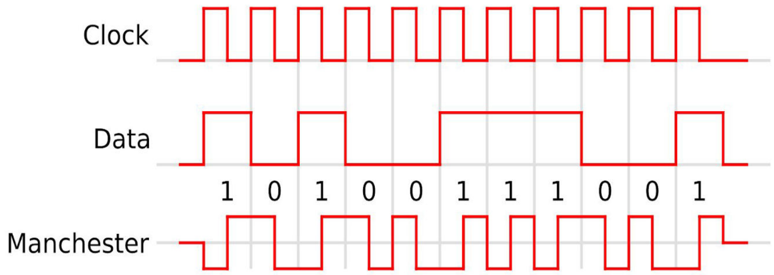
4) 10Base-F (Fiber Ethernet)

It is similar to 10Base-T, but it uses optical fibres as medium. 10Base-F uses a star topology to connect stations to a hub. The stations are connected to the hub using two fiber-optic cables.



Encoding and Decoding

The encoding and decoding uses Manchester type. Manchester code always has a transition at the middle of each bit period which indicates a data bit. It may have a transition at the start of the period also, but they do not carry information. As per IEEE 802.3 (Ethernet) standards, a logic 0 is represented by a high-low signal sequence and a logic 1 is represented by a low-high signal sequence.



Channel Access Method in Ethernet: Carrier Sense Multiple Access/Collision Detect (CSMA/CD)

CSMA/CD is the protocol for channel access in Ethernet networks. On Ethernet, any device can try to send a frame at any time. Each device senses whether the line is idle. If it is, the device begins to transmit its frame. If another device has tried to send at the same time, a collision is said to occur and the frames are discarded. The station then sends a jamming signal to inform others that a collision has happened. Each device then waits a random amount of time and retries until successful in getting its transmission sent. Collision detection in CSMA/CD serves two purposes. If a collision is detected, it means that the frame has not been received and needs to be resent. If a collision is not detected, it is a kind of acknowledgment that the frame was received.

Wireless LAN

Architectural Comparison between Wired and Wireless LANs

1. Medium

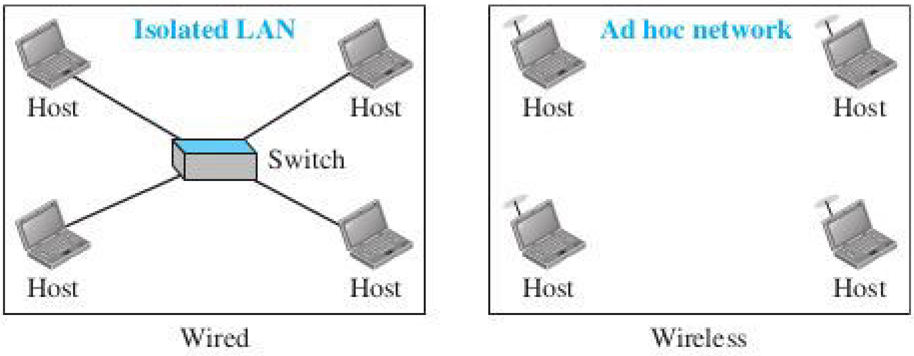
The first difference is that on wired LAN, there are one or more point-to-point full duplex wires between hosts, where as in wireless LAN, the medium is air and signal is generally broadcast. Here all the devices share the same medium.

2. Hosts

In wired LAN, the hosts are physically connected to the network at a point with a fixed link-layer address related to its network interface card (NIC). If it is moved to another network, it has to be physically connected there and then uses the services of that network. In wireless LAN, a host is not physically connected to the network; it can move freely (as we'll see) and can use the services provided by the network.

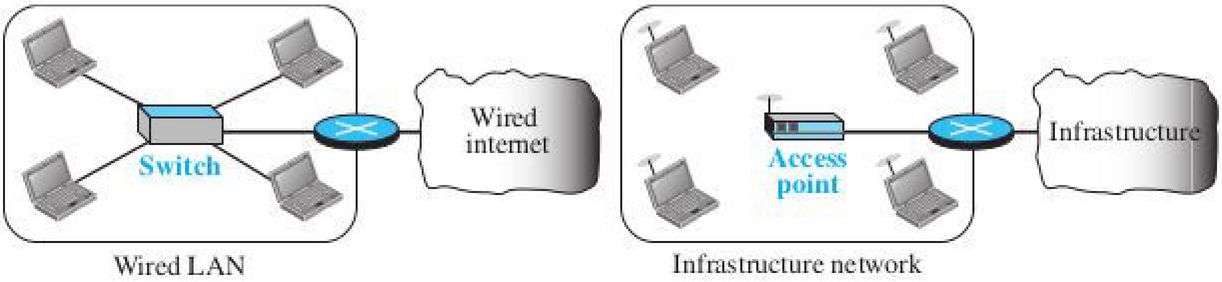
3. Isolated LANs

In wired LAn, the term isolated network means a set of devices connected to a link-layer switch. But in wireless LAN, the isolated network is called ***ad hoc network***, where all the devices communicate freely with each other without the help of a link layer switch.



4. Connection to Other Networks

A wired LAN can be connected to another network or an internetwork such as the Internet using a router. A wireless LAN may be connected to a wired infrastructure network, to a wireless infrastructure network, or to another wireless LAN.



The term Infrastructure network means if the ad hoc network is connected to external network through an *Access Point (AP)*. An access point is gluing two different environments together: one wired and one wireless. Communication between the AP and the wireless host occurs in a wireless environment; communication between the AP and the infrastructure occurs in a wired environment.

5. Moving between wired and wireless environments

If a device need to change from wired to wireless environment, its network interface must be changed. ie, change the network interface cards designed for wired environments to the ones designed for wireless environments and replace the link-layer switch with an access point. In this change, the link-layer addresses will change (because of changing NICs), but the network-layer addresses (IP addresses) will remain the same; we are moving from wired links to wireless links.

Characteristics of Wireless LAN

There are several characteristics of wireless LANs that either do not apply to wired LANs or the existence of which is negligible and can be ignored.

1. Attenuation

   The strength of electromagnetic signals decreases rapidly because the signal disperses in all directions; only a small portion of it reaches the receiver. The situation becomes worse with mobile senders that operate on batteries and normally have small power supplies.

2. Interference

   Another issue is that a receiver may receive signals not only from the intended sender, but also from other senders if they are using the same frequency band.

3. Multipath Propagation

   A receiver may receive more than one signal from the same sender because electromagnetic waves can be reflected back from obstacles such as walls, the ground, or objects. The result is that the receiver receives some signals at different phases (because they travel different paths). This makes the signal less recognizable.

4. Error

   Errors may happen in transmission before reaching the destination. The error is measured in Signal-to-Noise Ratio (SNR). If SNR is high, it means that the signal is stronger than the noise and hence may be able to convert the signal to actual data. If SNR is low, that means noise is stronger and hence very difficult to renvert the actual data.

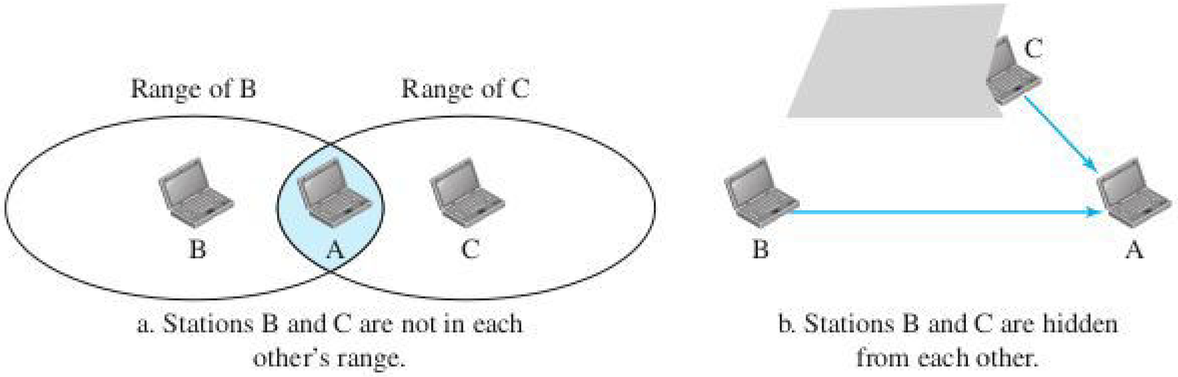
Access Control

Access control in wireless LAN deals with how a wireless host can get access to the shared medium (air). In wired LAN, Ethernet, the CSMA/CD works well because it is easy to check whether a collision has occurred or not in wired medium. But the CSMA/CD algorithm does not work in wireless LANs for three reasons:

1. To detect a collision, a host needs to send and receive at the same time (sending the frame and receiving the collision signal), which means the host needs to work in a duplex mode. Wireless hosts do not have enough power to do so (the power is supplied by batteries). They can only send or receive at one time.

2. Because of the **hidden station problem** (or *hidden terminal problem*).

Consider the situation as seen in the figure below *(figure a)*. There are three wireless devices A, B and C. C in not in the range of B, and hence B does not know the existence of C. Also, B is not in the range of C and hence C does not know B. But A is situated in the ranges of both B and C. This can be represented as in *figure b*.



a. Stations B and C are not in each other's range.

b. Stations B and C are hidden from each other.

Assume that station B is sending data to station A. In the middle of this transmission, station C also has data to send to station A. However, station C is out of B's range and transmissions from B cannot reach C. Therefore C thinks the medium is free. Station C sends its data to A, which results in a collision at A because this station is receiving data from both B and C. In this case, we say that stations B and C are hidden from each other with respect to A. Hidden stations can reduce the capacity of the network because of the possibility of collision.

3. The distance between stations can be great. Signal fading could prevent a station at one end from hearing a collision at the other end.

Due to the above reasons, **CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance)** is used in wireless LANs.

In CSMA/CA, the sender checks whether the medium is busy. If the channel is busy, it will wait for some random amount of time before checking the channel again. If the channel is free, it waits for a short period of time and sends its data. After successful reception of the data the receiver sends back an acknowledgement. If the sender does not receive the acknowledgement, it will send the data again. Some wireless networks use an optional RTS/CTS messages for data transmission. Here if a device wants to send data it will broadcast an RTS (Request To Send) signal first. If the receiver is ready, it will send back a CTS (Clear To Send) signal and then the transmission happens. Since the RTS and CTS are broadcast signals, all other devices in the range will know about the transfer, they will not send any packets and thus avoid collision.

IEEE 802.11 PROJECT

IEEE has defined the specifications for a wireless LAN, called IEEE 802.11, which covers the physical and data-link layers. It is sometimes called wireless Ethernet. Wireless Fidelity (WiFi) is a type of wireless LAN technology certified by WiFi Alliance.

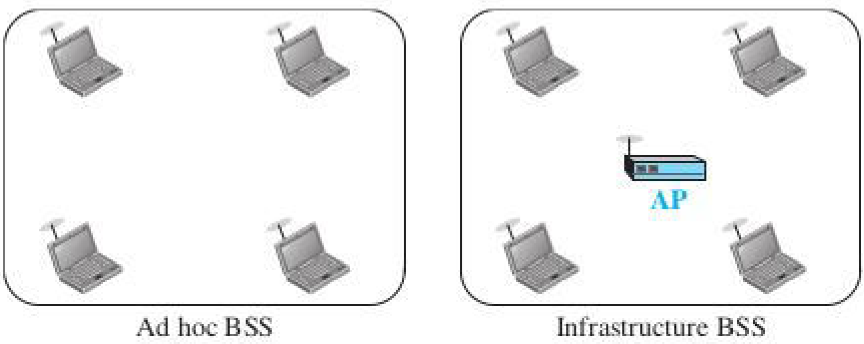
IEEE 802.11 Architecture

The standard defines two kinds of services: the basic service set (BSS) and the extended service set (ESS).

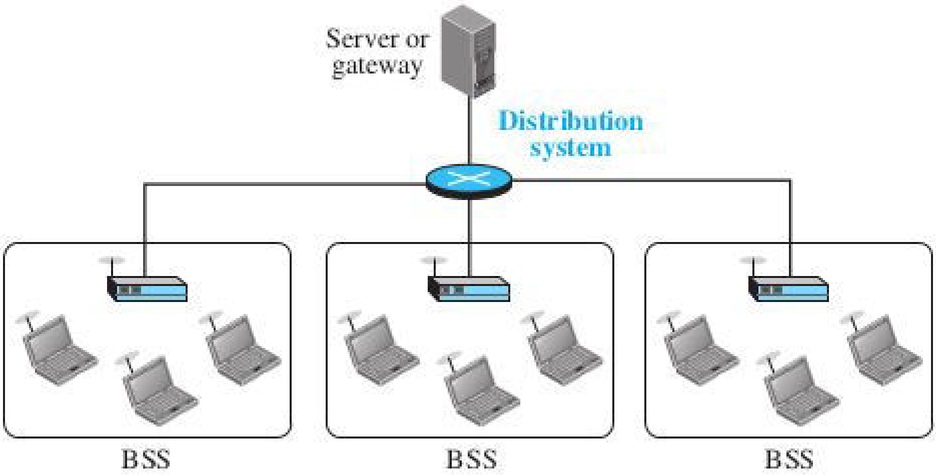
Basic Service Set (BSS)

IEEE 802.11 defines the basic service set (BSS) as the building blocks of a wireless LAN.

A basic service set is made of stationary or mobile wireless stations and an optional central base station, known as the access point (AP). The BSS without AP communicates among each other. They cannot communicate with the stations in another BSS. Such BSS is called Ad hoc BSS. A BSS with an AP is sometimes referred to as an infrastructure BSS. The stations in such BSS communicate with outside networks through this AP.



Extended Service Set (ESS)



An extended service set (ESS) is made up of two or more BSSs with APs. In this case, the BSSs are connected through a *distribution system*, which is a wired or a wireless network. The distribution system connects the APs in the BSSs. The distribution system can be any IEEE LAN such as Ethernet. The extended service set uses two types of stations: mobile and stationary. The mobile stations are normal stations inside a BSS. The stationary stations are AP stations that are part of a wired LAN. The systems in two BSSs communicates through APs.

Station Types:

IEEE 802.11 defines three types of stations based on their mobility in a wireless LAN: *no-transition, BSS-transition, and ESS-transition mobility*.

A station with **no-transition mobility** is either stationary (not moving) or moving only inside a BSS.

A station with **BSS-transition mobility** can move from one BSS to another, but the movement is confined inside one ESS.

A station with **ESS-transition mobility** can move from one ESS to another. However, IEEE 802.11 does not guarantee that communication is continuous during the move.

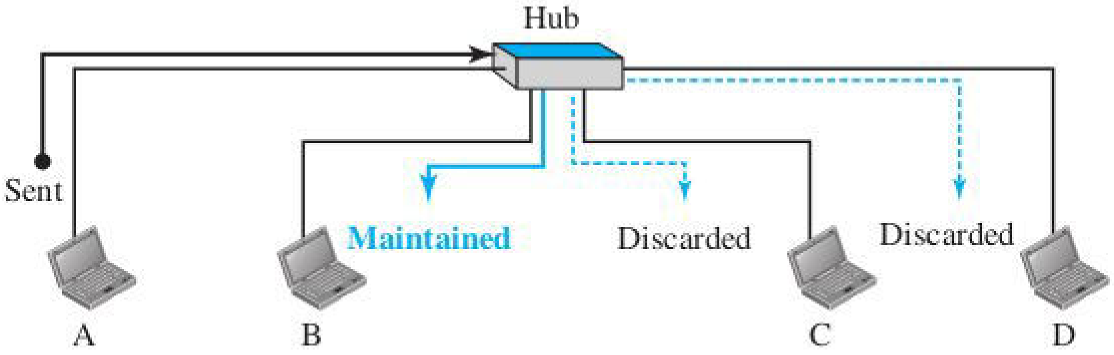
CONNECTING DEVICES

Connecting devices connect hosts together to form networks, or connects networks together. These devices do not need all the layers in TCP/IP protocol suite, but need some layers only based on their duties. The major networking devices are hubs, link-layer switches and routers. Hubs operate in the first layer of the Internet model. Link-layer switches operate in the first two layers. Routers operate in the first three layers.

Hubs

A repeater is a device that accepts weak signals from one channel on a port, regenerates and retimes the original bit pattern and sends the signal to another port. Thus the repeaters are used to overcome length restrictions. In star topology, there are many ports. Such multiport repeaters are called **hubs** (or **active hubs**). There are passive hubs which act as signal junction among many systems but do not regenerate the signal.

Hubs work at the *physical layer* and does not know any addresses or contents in the data packets.



When a packet from station A to station B arrives at the hub, the signal representing the frame is regenerated to remove any possible corrupting noise, but the hub forwards the packet from all outgoing ports except the one from which the signal was received. In other words, the frame is broadcast. All stations in the LAN receive the frame, but only station B keeps it. The rest of the stations discard it.

> A **hub** or a **repeater** is a physical-layer device. They just regenerate the corrupted bits and send them out to every other port. They do not have a link-layer address and they do not check the link-layer address of the received frame.

Link-Layer Switches

A link-layer switch (or switch) operates in both the physical and the data-link layers. As a physical-layer device, it regenerates the signal it receives. As a link-layer device, the link-layer switch can check the MAC addresses (source and destination) contained in the frame.

Filtering ability of Switch: The switch can check the destination address of a frame and can decide to which outgoing port the frame should be sent. The switch keeps a forwarding table consisting of its port numbers and the MAC addresses of the devices connected to that port. This table helps in filtering.

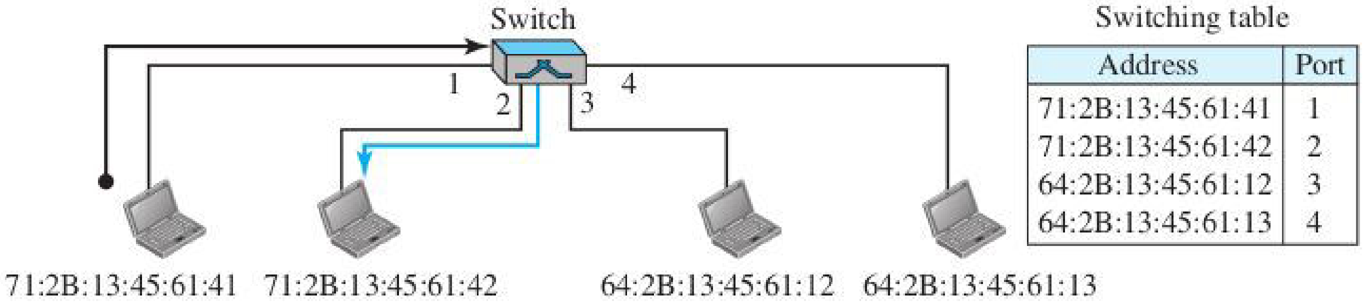
Transparent Switches: A transparent switch is a switch in which the stations are completely unaware of the switch's existence. If a switch is added or deleted from the system, there is no need of reconfiguration of the stations. According to the IEEE 802.1d specification, a system equipped with transparent switches must meet three criteria:

1. Frames must be forwarded from one station to another.
2. The forwarding table is automatically made by learning frame movements in the network.
3. Loops in the system must be prevented.

Working of a Switch

A switch does two functions; learning and forwarding. Learning is the process of knowing the physical addresses of the systems connected to the ports. Forwarding is the process of transferring packets to proper systems.

Learning: Consider the figure below.



Switching table

Address	Port
71:2B:13:45:61:41	1
71:2B:13:45:61:42	2
64:2B:13:45:61:12	3
64:2B:13:45:61:13	4

When a switch is just switched on, its switching table (or forwarding table) does not have any entries. ie, it cannot deliver packets to correct destination. When a packet arrives at a port it makes an entry in table containing the port number and sender MAC address of the incoming packet. Since it does not have a table entry for the destination MAC address, it floods the packet to every other port. Thus initially, the switch acts just like a hub. After some time, its table will be filled and then the switch can forward the packet to exact destination.

An example scenario is as follows;

1. When station A sends a frame to station D, the switch does not have an entry for either D or A. The frame goes out from all three ports; the frame floods the network. However, by looking at the source address, the switch learns that station A must be connected to port 1. This means that frames destined for A, in the future, must be sent out through port 1. The switch adds this entry to its table. The table has its first entry now.

2. When station D sends a frame to station B, the switch has no entry for B, so it floods the

network again. However, it adds one more entry to the table related to station D.

3. The learning process continues until the table has information about every port. However, note that the learning process may take a long time. For example, if a station does not send out a frame (a rare situation), the station will never have an entry in the table.

Advantages of Switches

Collision Elimination: There is no need for carrier sensing and collision detection; each host can transmit at any time.

Connecting Heterogeneous Devices: A link-layer switch can connect devices that use different protocols at the physical layer (data rates) and different transmission media. As long as the format of the frame at the data-link layer does not change, a switch can receive a frame from a device that uses twisted-pair cable and sends data at 10 Mbps and deliver the frame to another device that uses fiber-optic cable and can receive data at 100 Mbps.

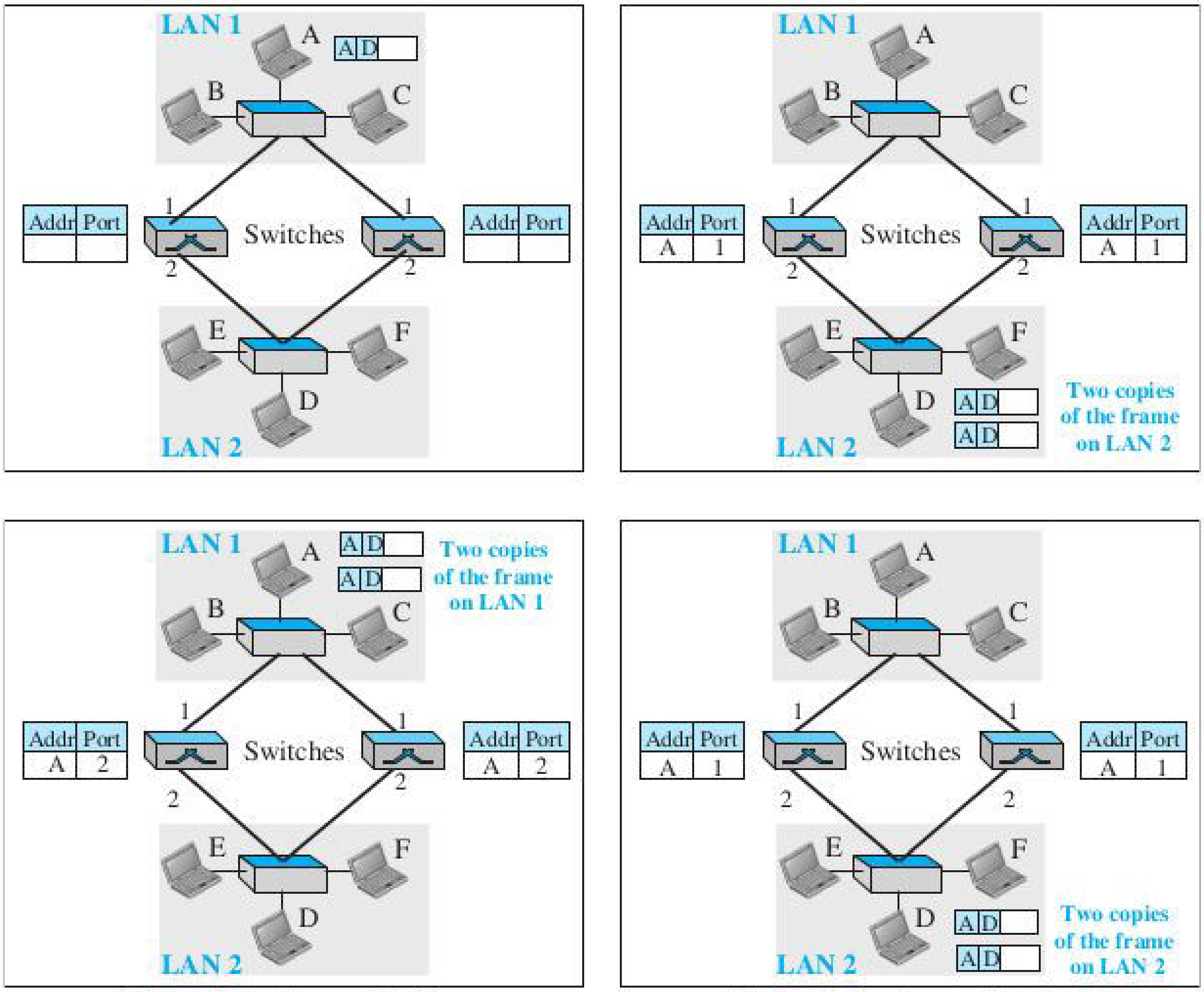
Differences between Hub and Switch

Feature	Hub	Switch
Layer	Physical layer (ie, layer 1 device)	Physical layer and Data link layer (ie, layer 2 device)
Function	Acts as an electrical junction. (multi port repeater)	Allows one-to-one communication. (multiport bridge)
Type of transmission	Works by broadcasting	Initially broadcasting, then unicasting or multicasting as required.
Transfer units	Bits	Frames
Ports	LImited number of ports (4-12). Efficiency decreases with increase in number of ports.	Many ports. (upto 48 is common).
Software	May not have software	Controlled by software/firmware
Transmission mode	Half duplex	Half/full duplex
Collision domain	Only one collision domain, the hub itself.	Different ports have separate collision domain.
Filtering	No filtering capacity	Filtering based on MAC addresses

Loop Problem in switches

Systems administrators like to have redundant switches (more than one switch between a pair of LANs) to make the system more reliable. If a switch fails, another switch takes over until the failed one is repaired or replaced. Redundancy can create loops in the system, which is very undesirable. Loops can be created only when two or more broadcasting LANs (those using hubs, for example) are connected by more than one switch.

Consider such a situation given below. There are two LANs connected by two switches.


LAN 1
A
B
C
A D
Addr Port
Switches
1
2
E
F
D
LAN 2
Two copies of the frame on LAN 2
Two copies of the frame on LAN 1
A 1
A 2

a. Station A sends a frame to station D

b. Both switches forward the frame

c. Both switches forward the frame

c. Both switches forward the frame

1. Station A sends a frame to station D. The tables of both switches are empty. Both forward the frame and update their tables based on the source address A.

2. Now there are two copies of the frame on LAN2. The copy sent out by the left switch is received by the right switch, which does not have any information about the destination address D; it forwards the frame. The copy sent out by the right switch is received by the left switch and is sent out for lack of information about D.The tables of both switches are updated, but still there is no information for destination D.

3. Now there are two copies of the frame on LAN 1. Step 2 is repeated, and both copies are sent to LAN2.

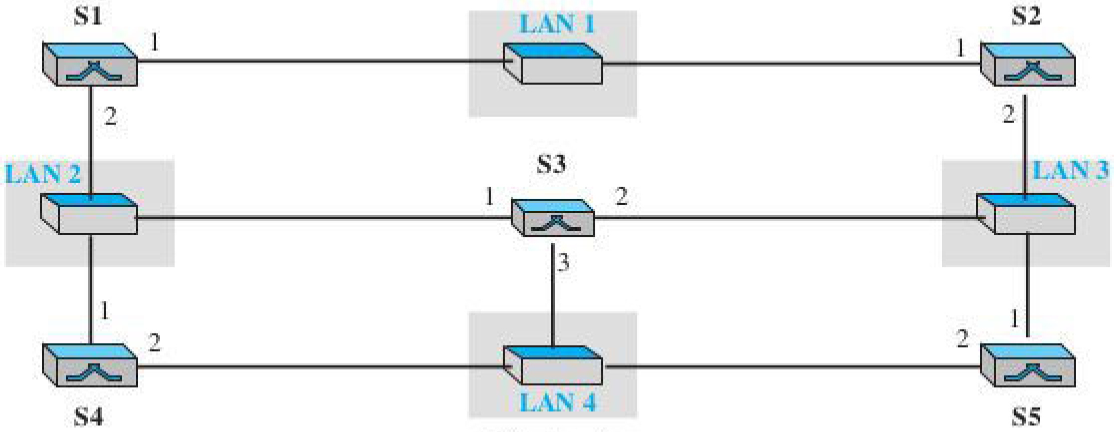
4. The process continues on and on. Note that switches are also repeaters and regenerate frames. So in each iteration, there are newly generated fresh copies of the frames.

Solving the Loop problem: the Spanning Tree Algorithm

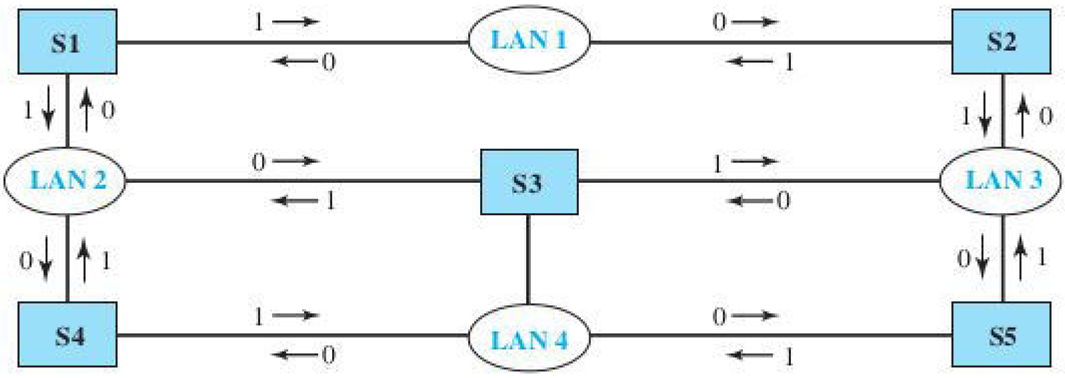
IEEE specifies that switches have to create spanning trees to avoid loop problem. A **spanning tree** is a tree in which there is no loop, but every node can be reached from every other node. And there is only one path between every two nodes.

Without changing the physical topology, the switches make a logical topology based on some spanning tree algorithms so that there is no loop.

Consider a network with five switches and four LANs. These are represented as nodes in the diagram below. To find the spanning tree, we need to assign a cost (metric) to each link, which is generally the minimum hop count. The hop count is normally 1 from a switch to the LAN and 0 in the reverse direction.



a. Actual system



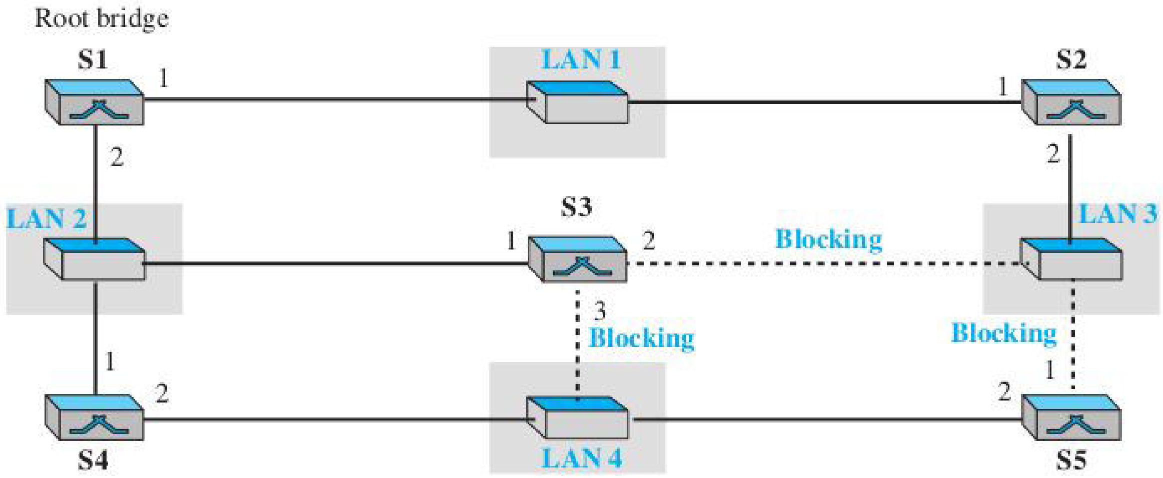
b. Graph representation with cost assigned to each arc

The process for finding the spanning tree involves the following steps (**Algorithm)**:

1. Each switch broadcasts its own unique ID so that all switches know which one has the smallest ID.
2. The switch with the smallest ID is selected as the root switch of the tree (here switch S1).
3. The algorithm tries to find the shortest path (a path with the shortest cost) from the root switch to every other switch or LAN. Shortest path can be found using algorithms such as Dijkstra algorithm.
4. The combination of the shortest paths creates the shortest tree.
5. Based on the spanning tree, we mark the ports that are part of it, the forwarding ports, which forward a frame that the switch receives. We also mark those ports that are not part

of the spanning tree, the blocking ports, which block the frames received by the switch.

The final logical topology is given below.



Routers

A router is a device used to connect different networks running different protocols. *In other words, a router is an internetworking device; it connects independent networks to form an internetwork.*

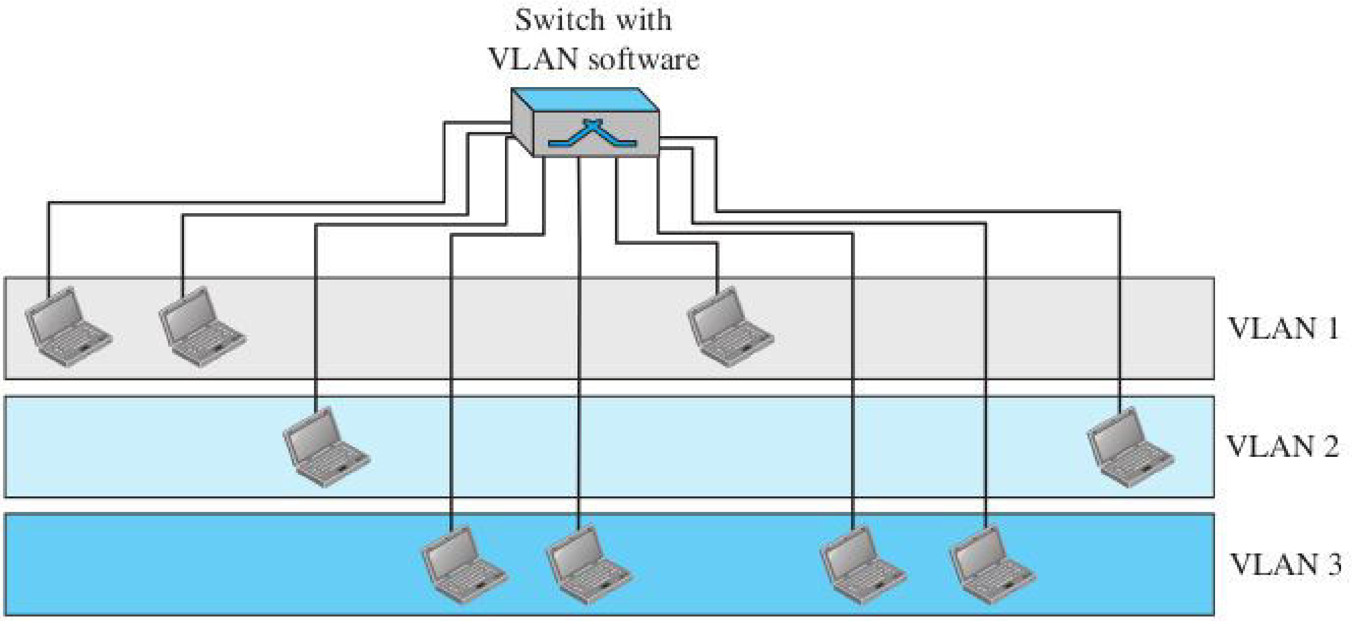
A router is a three-layer device; it operates in the physical, data-link, and network layers. As a physical-layer device, it regenerates the signal it receives. As a link-layer device, the router checks the physical addresses (source and destination) contained in the packet. As a network-layer device, the router checks the network-layer addresses.

There are three major differences between a router and a repeater or a switch.

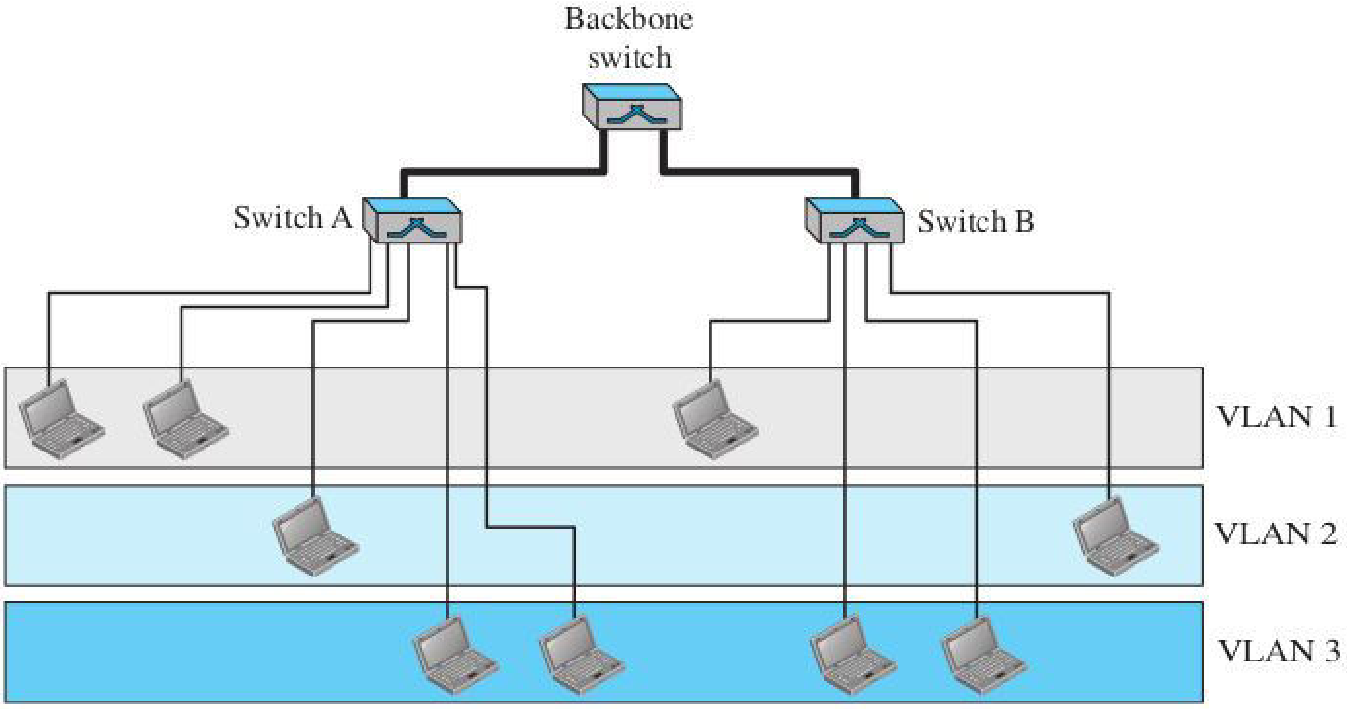
1. A router has a physical and logical (IP) address for each of its interfaces.
2. A router acts only on those packets in which the link-layer destination address matches the address of the interface at which the packet arrives.
3. A router changes the link-layer address of the packet (both source and destination) when it forwards the packet.

Virtual LAN

A station is considered part of a LAN if it physically belongs to that LAN. The criterion of membership is geographic. But it is possible to divide some systems in a network virtually into different groups so that the groups can not communicate each other. This division can be done by software in the switches. Such settings are available in manageable switches. Such virtually created LANs are called Virtual LANs (VLANs). This is done without any change in physical topology of the network.



VLAN technology even allows the grouping of stations connected to different switches in a VLAN. Thus physically the systems may be seen in different segments, they may be under the same VLANs. The figure below shows a local area network with two switches and three VLANs. Stations from switches A and B belong to each VLAN.



VLAN Membership

The memberships to a particular VLAN can be done based on different criteria such as MAC addresses, IP Addresses, physical port numbers (interface numbers) or a combination of all these.

1. Interface Numbers

   Some VLAN vendors use switch interface numbers as a membership characteristic. For example, the administrator can define that stations connecting to ports 1, 2, 3, and 7 belong to VLAN 1, stations connecting to ports 4, 10, and 12 belong to VLAN 2, and so on.

2. MAC Addresses

   Some VLAN vendors use the 48-bit MAC address as a membership characteristic. For example, the administrator can stipulate that stations having MAC addresses E2:13:42:A1:23:34 and F2:A1:23:BC:D3:41 belong to VLAN 1.

3. IP Addresses

   Some VLAN vendors use the 32-bit IP address as a membership characteristic. For example, the administrator can stipulate that stations having IP addresses 181.34.23.67, 181.34.23.72, 181.34.23.98, and 181.34.23.112 belong to VLAN 1.

4. Multicast IP Addresses

   Some VLAN vendors use the multicast IP address as a membership characteristic. Multicasting at the IP layer is now translated to multicasting at the data link layer.

5. Combination

   The administrator can choose one or more characteristics when installing the software.

Configuration

VLANs and stations can be configured in three ways; manually, semi automatically, and automatically.

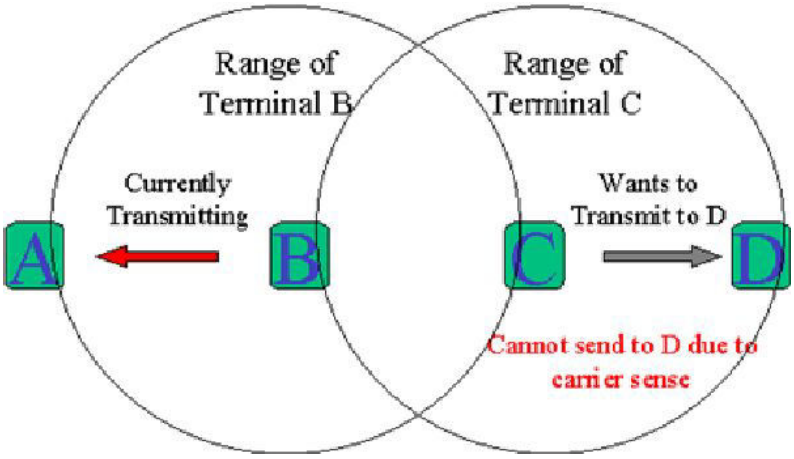
Manual Configuration: The administrator manually configure the VLAN based on one of the different criteria such as port numbers, MAC addresses, etc.

Automatic Configuration: In an automatic configuration, the stations are automatically connected or disconnected from a VLAN using criteria defined by the administrator. For example, the administrator can define the project number as criterion for being a member of a group. When a user changes projects, he or she automatically migrates to a new VLAN.

Semi Automatic Configuration: A semiautomatic configuration is somewhere between a manual configuration and an automatic configuration. Usually, the initializing is done manually, with migrations done automatically.

Exposed Terminal Problem

Exposed terminal occurs in wireless LANs. Consider a network containing 4 nodes, say A, B, C and D as in figure.



Here A and C are in the range of B. B and D are in the range of C. Suppose B is sending data to A. The problem occurs when C tries to send data to D at the same time. Even though both transfers are independent, C cannot send data. This is because when C checks the channel if it is free, it will sense that the channel is not free as B is sending data. C can send data only after B finishes sending the data. This problem is called Exposed Terminal problem.

Module II – Network Layer

Network layer services – Packetizing, routing and forwarding, other services – Performance – delay, throughput, packet loss, congestion control – IPV4 address – address space, classful addressing, classless addressing, subneting – DHCP – Internet protocol (IP) – datagram format, fragmentation – IPV4 datagram security – Routing algorithms – Distance-vector, Link-state, path vector – unicasting, multicasting, broadcasting

Services

- Responsible for the routing of packets from source host to destination host.
- Use IP address for routing.
- Found on end devices and some connecting devices such as router.
- Major Services are given below.

1. Packetizing

- Main duty of network layer
- Packetizing: encapsulating the payload (data received from upper layer) in a network-layer packet at the source and decapsulating the payload from the network-layer packet at the destination.
- Network layer carries the upper layer payload from source to destination *without changing the payload content*.
- The source network layer adds its header that contains the source and destination addresses and some other information that is required by the network-layer protocol.
- If the payload is too large, it is fragmented at the network layer of the source device/intermediary routers.
- In such case, the *destination* network layer keeps the packet until all the fragments come, then reassembles the fragments and delivers them to the upper layer protocol.

2. Routing and Forwarding

- Routing
  - o It is the process of leading the packets from source to destination through different intermediate LANs, WANs and routers.
  - o Network layer in routers finds the best route among all routes using static/dynamic routing protocols and routing tables.
- Forwarding:
  - o When a packet arrives at one of the router interfaces, the router forwards the packet to another interface based on the routing tables.
  - o The forwarding can be to another attached network (in unicast routing) or to some attached networks (in multicast routing).

3. Other Services

- Error Control:
  - o Since the datagram may be fragmented at each router, error detection in network layer is inefficient.
  - o But there is a checksum field for the header to control any corruption in the header.

  - The Internet uses ICMP which provides some kind of error control if the datagram is discarded or has some unknown information in the header.
- Flow Control:
  - *It regulates the rate of packet flow from the sender if the receiver is slower to accept the packets.*
  - *This occurs when the receiver sends feedback to the sender to regulate the flow.*
  - The network layer <u>does not</u> directly provides flow control because;
    1. The lack of flow control makes this layer simpler.
    2. The upper layers can use buffers to store packets if they are slow.
    3. The upper layers provides flow control, hence it is not needed at the network layer.
- Congestion Control:
  - It is not directly provided at the network layer as it is done by upper layers.
- Quality of Service (QoS):
  - Since the Internet is used for a variety of services, QoS is important.
  - But to keep the network layer simpler, it is provided at the upper layers.
- Security:
  - Since the Internet is used by a huge number of people, security is a big concern.
  - The connectionless network layer can be made into connection oriented by using IPSec.

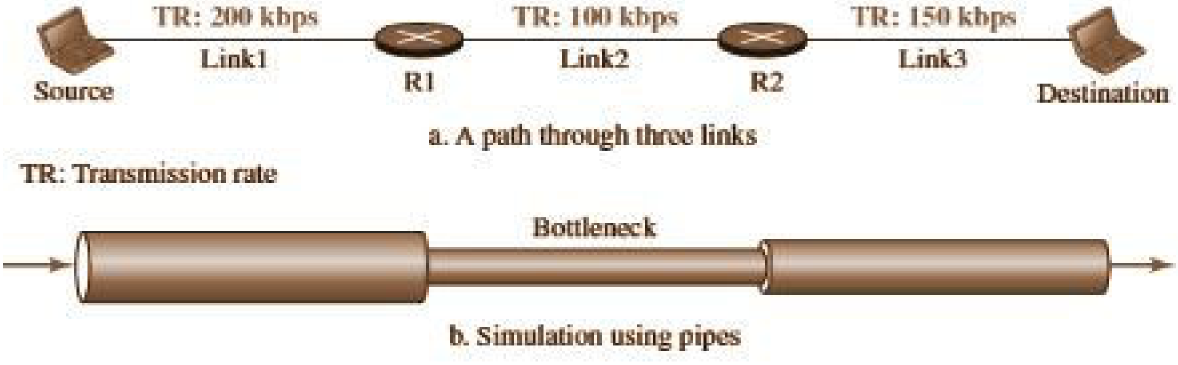
Network Layer Performance

- It can be measured in terms of *delay*, *throughput*, *packet loss* and *Congestion control*.
- **Delay**: Four types; transmission delay, propagation delay, processing delay, and queuing delay.
  - **Transmission Delay**:
    - Delay in sending the packet (first bit to last bit) by the source host or router.
    - If first bit is placed on the line at time t1 and last bit at t2, then transmission delay is t2-t1.
    - Transmission delay is longer for a longer packet and shorter if the sender can transmit faster.
    - Delay_{tr} = (Packet length) / (Transmission rate).
    - For a Fast Ethernet LAN with 100 Million bits per second and with a packet length of 10000 bits, the transmission delay is 10000/100000000=100 microsecond.
  - **Propagation Delay**
    - The time taken for a bit to travel from point A to point B in the transmission media.
    - Depends on the propagation speed of the media, which is 3×10^8 meters/second in a vacuum and normally much less in a wired medium.
    - It also depends on the distance of the link.
    - Delay_{pg} = (Distance) / (Propagation speed).
    - If the distance of a cable link in a point-to-point WAN is 2000 meters and the propagation speed of the bits in the cable is 2×10^8 meters/second, then the propagation delay is 10 microseconds.

- **Processing Delay**
  - The time required for a router or a destination host to receive a packet from its input port, remove the header, perform an error detection procedure, and deliver the packet to the output port (in the case of a router) or deliver the packet to the upper-layer protocol (in the case of the destination host).
  - The processing delay may be different for each packet.
  - Delay_{pr} = Time required to process a packet in a router or a destination host

- **Queuing Delay**
  - Happens in router
  - A router has an input queue connected to each of its input ports to store packets waiting to be processed
  - The router also has an output queue connected to each of its output ports to store packets waiting to be transmitted.
  - The queuing delay (Delay_{qu}) is the total time for which a packet stays in these queues.
- **Total Delay**
  - For *n* routers;
    Total delay = (n+1) (Delay_{tr} + Delay_{pg} + Delay_{pr}) + (n) (Delay_{qu})

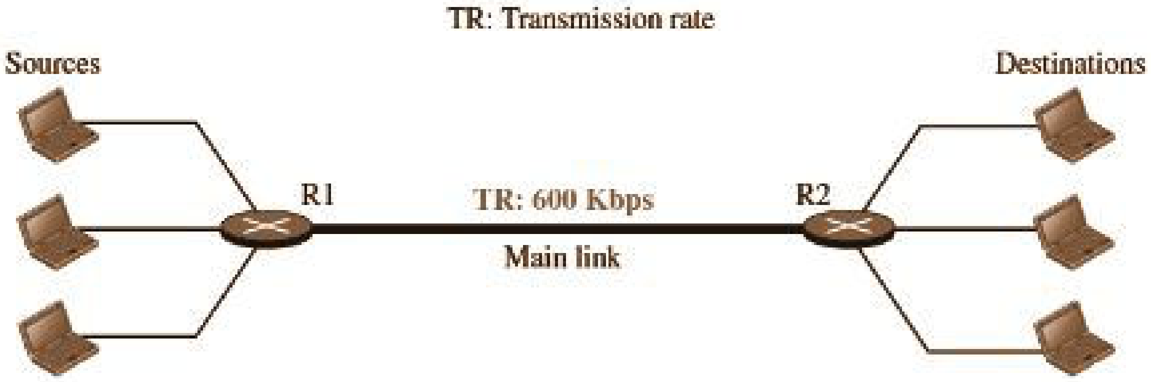
- **Throughput**
  - Throughput at a point in a network is the number of bits passing through that point in a second.
  - Actually it is the transmission rate at that point.
  - Since a source to destination path contains a number of links, there may be different rates.



a. A path through three links

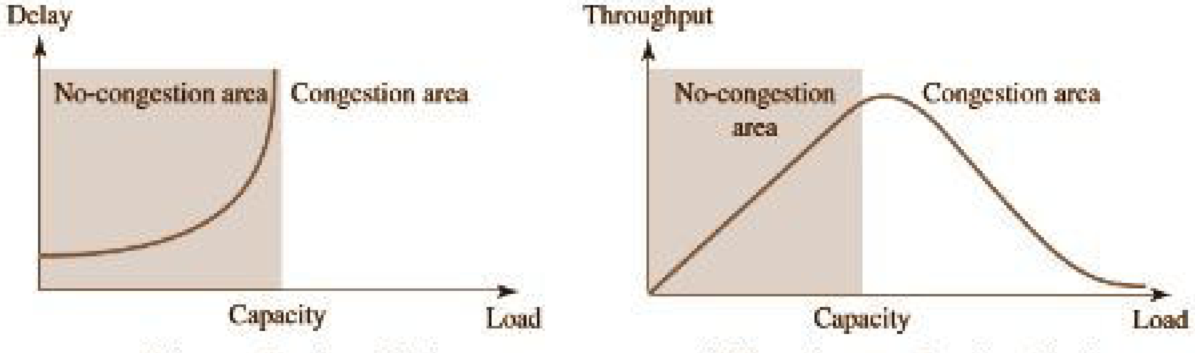
b. Simulation using pipes

  - Throughput = minimum of all transmission rates
  - In shared medium *shown below*, the transmission rates are shared among all the three routers (in the sender), hence only 200kbps can be calculated for throughput. This is same with receiver.



- **Packet Loss**
  - The packets may be lost during transmission.
  - Also, when the router buffers become full, the following packets may be lost.
  - These packets need to be resent, but which in turn cause overflow and packet lose.

- **Congestion Control**
  - Congestion at the network layer is related to two issues, throughput and delay.
  - When the load is less than the capacity, the delay is minimum.
  - As the load increases to the network capacity, the delay increases sharply that it becomes infinity.
  - The throughput increases to the maximum as the network reaches its capacity, but after that the throughput decreases.
  - Congestion occurs at this area.



a. Delay as a function of load

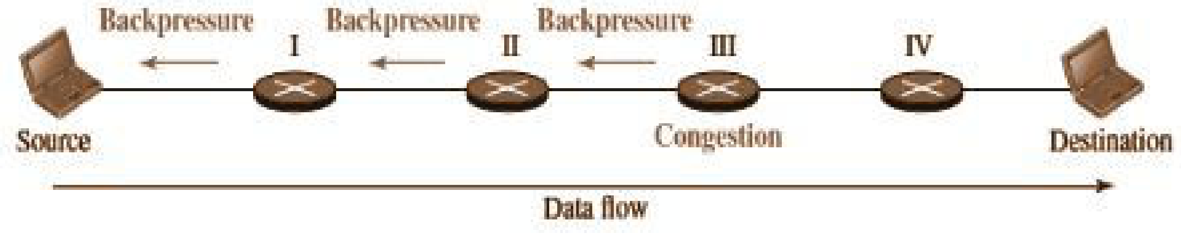
b. Throughput as a function of load

  - Congestion control refers to techniques and mechanisms that can either prevent congestion before it happens or remove congestion after it has happened.
  - Congestion control has two broad categories: open-loop congestion control (prevention) and closed-loop congestion control (removal).

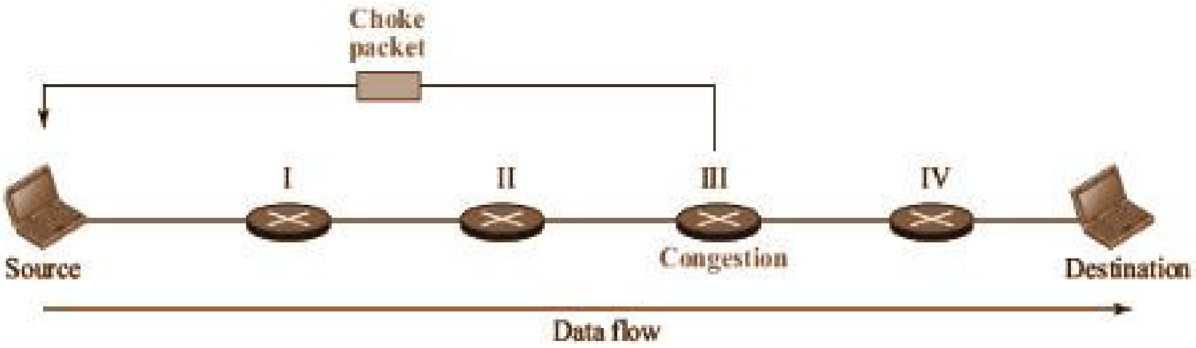
  - **<u>Open-Loop Congestion Control (congestion prevention policies)</u>**
    - **Retransmission Policy**
      - If the sender feels that the packet is corrupted or lost, the packet needs to be retransmitted.
      - A good retransmission policy and optimized timers can prevent congestion.
    - **Window Policy:**

        - Type of window affects congestion.
        - Selective Repeat window is better than the Go-Back-*N* window for congestion control.
    - **Acknowledgement Policy**
        - Several approaches are used here;
        - If the receiver does not send acknowledgment for all packets, it may slow down the sender.
        - A receiver may send an acknowledgment only if it has a packet to be sent or a special timer expires.
        - A receiver may decide to acknowledge only *N* packets at a time.
    - **Discarding Policy**
        - A good discarding policy that does not interrupt the integrity of data can prevent the congestion.
        - For example, in audio transmission, if the policy is to discard less sensitive packets when congestion is likely to happen, the quality of sound is still preserved and congestion is prevented or alleviated.
    - **Admission Policy**
        - Seen in Virtual-circuit network as a Quality of Service mechanism.
        - The router can deny virtual circuit creation if there is congestion present or likely to happen in the network.

- **<u>Closed Loop Congestion Control (Congestion Removal Policies)</u>**
    - Try to alleviate congestion after it happens.
    - Some types are;
    - **Back pressure**
        - Here a 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, and so on.
        - It is a node-to-node congestion control that starts with a node and propagates, in the opposite direction of data flow, to the source.
        - Only used in virtual-circuit network.



    - **Choke Packet**
        - A choke packet is a packet sent by a node directly to the source to inform it of congestion.
        - The intermediate nodes through which the packet has traveled are not warned.
        - In ICMP, the warning message is a 'source quench' message.



- **Implicit Signaling**
  - In this, there is no communication between the congested node or nodes and the source.
  - The source guesses that there is congestion somewhere in the network from other symptoms (no ack for a long time, delayed ack, etc).
  - Thus the source slow down.
- **Explicit Signaling**
  - The node that experiences congestion can explicitly send a signal to the source or destination.
  - Apart from choke-packet method, here the signal is included in the packets that carry data

IPv4 address

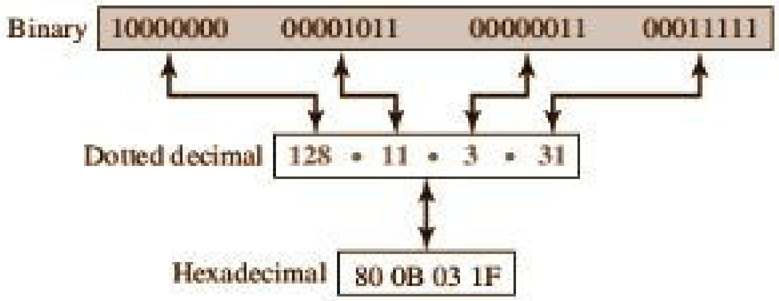
- IP address is a network layer logical address defined to identify each host.
- An IPv4 address is a 32-bit address that uniquely and universally defines the connection of a host or a router to the *Internet.*
- The IP address is the address of the connection, not the host or the router, because if the device is moved to another network, the IP address may be changed.

Address Space

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

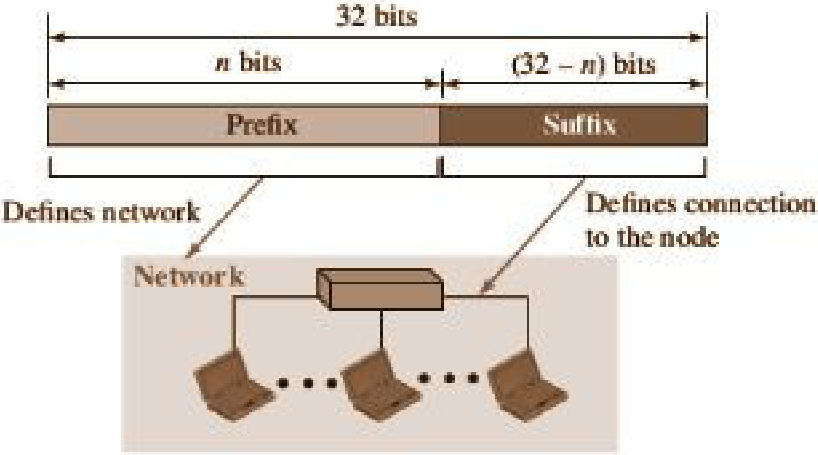
IPv4 notation

- Three notation: binary notation (base 2), hexadecimal notation (base 16) and dotted-decimal notation.
- Binary notation: IPv4 address is displayed as 32 bits; each byte is separated by space.
- Hexadecimal notation: Four bits form a hexadecimal number. Often used in network programming.
- Dotted decimal notation: Most common form and easily readable. Each byte is represented as decimal and each byte is separated by dot.



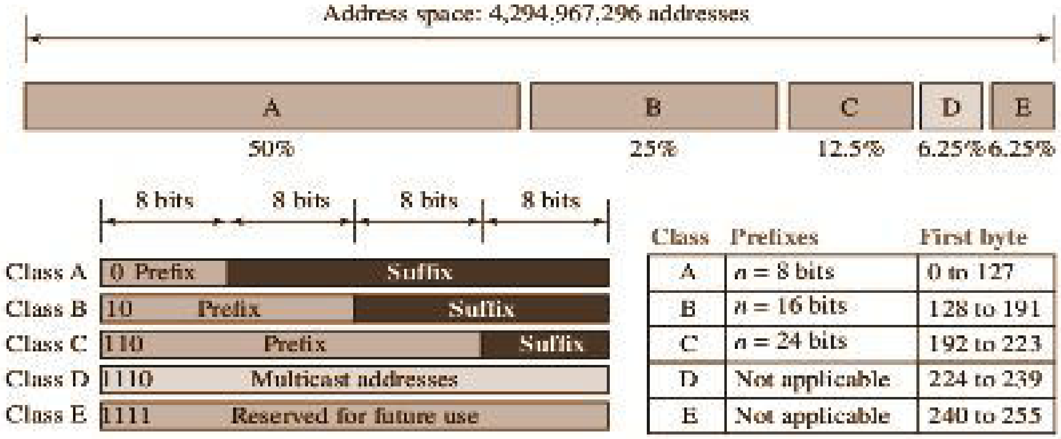
Hierarchy in Addressing

A 32-bit IPv4 address is divided only into two parts. The first part of the address, called the prefix, defines the network (say *n* bits); the second part of the address, called the suffix, defines the node or host (*32-n* bits).



Classful Addressing

- To accommodate both small and large networks, three fixed-length prefixes were designed (n = 8, n = 16, and n = 24).
- The whole address space was divided into five classes (class A, B, C, D, and E) .
- This scheme is referred to as classful addressing.



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

Subnetting and Supernetting

In subnetting, a class A or class B block is divided into several subnets. Each subnet has a larger prefix length than the original network. For example, if a network in class A is divided into four subnets, each subnet has a prefix of n_{sub} = 10. This idea did not work because most large organizations were not happy about dividing the block and giving some of the unused addresses to smaller organizations.

Supernetting was devised to combine several class C blocks into a larger block to be attractive to organizations that need more than the 256 addresses available in a class C block. This idea did not work either because it makes the routing of packets more difficult.

Advantage of Classful Addressing

If we see an address, we can easily find the class of the address and the prefix length.

Classless Addressing

- In classful addressing, the number of networks and hosts had limitations due to the fixed length prefix part. This can be overcome by using classless addressing where variable length prefix is used for each network.
- Here we can have blocks of 1 address, 2 addresses, 4 addresses, 8 addresses and so on.
- It is possible for the ISP to assign network addresses for organizations or individuals as per their requirement.
- The number of bits on the network part is given along with the address by using a *slash* (/) notation.
- This representation is called Classless Inter Domain Routing or CIDR .

  Eg: 12.24.76.8 /8
  23.14.67.92 /12
  220.8.24.255 /25

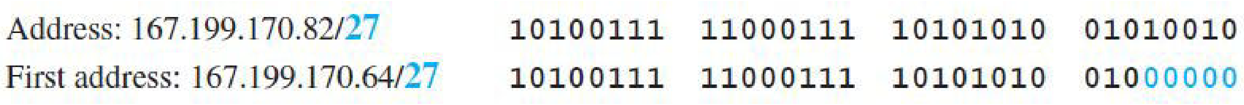
Extracting Information from an Address

If there are *n* bits for the prefix field (ie ***/n*** notation), we can directly interpret three information.

1. The number of addresses in the block (ie, block size) is found as N = $2^{32} - n$.
2. To find the first address, keep the n leftmost bits and set the (32 − n) rightmost bits all to 0s.
   This address is called the network address. This cannot be given to a host.
   The next address will be the first host address.
3. To find the last address,keep the n leftmost bits and set the (32 − n) rightmost bits all to 1s.
   This address is called the broadcast address of that network. This can't be given to a host.
   The previous address will be the last host address.

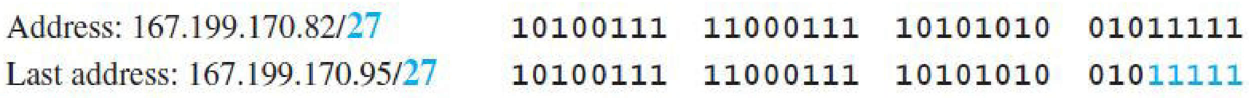
Eg: Consider the address 167.199.170.82/**27**
Here the network bits are 27 and host bits are 32-27=5

1. Number of total addresses in this network = 2^5 = 32
2. The first address can be found by keeping the first 27 bits and changing the rest of the bits to 0s.



Address: 167.199.170.82/27	10100111	11000111	10101010	01010010
First address: 167.199.170.64/27	10100111	11000111	10101010	01000000

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



Address: 167.199.170.82/27 10100111 11000111 10101010 01011111

Last address: 167.199.170.95/27 10100111 11000111 10101010 01011111

Network address: 167.199.170.64 /27
First host: 167.199.170.65 /27
Last host: 167.199.170.94 /27
Broadcast address: 167.199.170.95 /27

Subnet mask

To identify the network address, the system use an address called the subnet mask. This address is obtained by setting all network bits to 1 and all host bits to 0. The system makes an AND operation with the given IP address and the result will be the network address of that IP address.

Subnet mask in the above example: 11111111 11111111 11111111 111***00000***
255. 255. 255. 224

Ie, subnet mask is **255.255.255.224 /27**

Example:

An ISP has requested a block of 1000 addresses. Since 1000 is not a power of 2, 1024 addresses are granted. To have 1024 hosts, we need 10 bits at the host part (2^{10} = 1024). Hence we need 22 bits at the network part. Thus an address x.y.z.p **/22** is granted to the ISP.

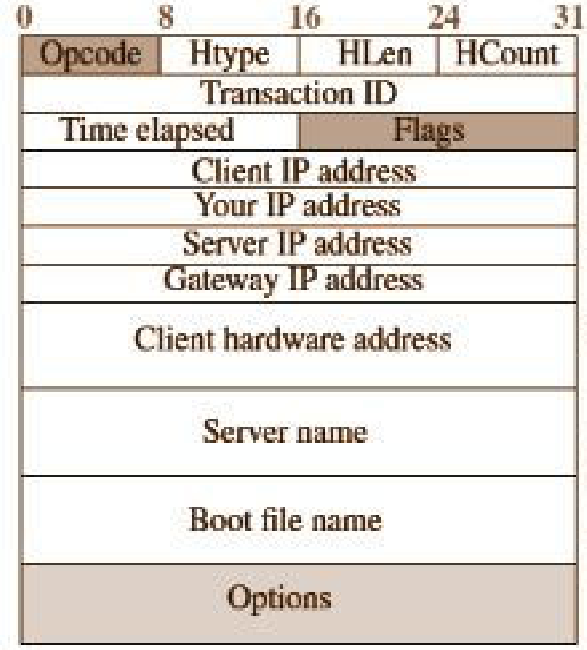
Special Addresses

- This-host address:
  - 0.0.0.0/32 is called this-host address.
  - This will be the address when the system is booted up.
  - Whenever a host needs to send an IP datagram but it does not know its own address to use as the source address, it use this address.
- Limited-broadcast address:
  - 255.255.255.255/32
  - Used whenever a router or a host needs to send a datagram to all devices in a network.
  - But the routers block this type of addresses.
- Loopback address:
  - 127.0.0.0/8
  - The packet with this destination network address never leaves the host.
  - Used to test whether the NIC is working properly.
  - Also used to test client-server programs.
- Private Addresses
  - Used for private use, ie for internal networks.
  - Class A: 10.0.0.0/8
  - Class B: 172.16.0.0/12
  - Class C: 192.168.0.0/16
  - 169.254.0.0/16 used by the client systems when DHCP fails.
- Multicast Addresses
  - The block 224.0.0.0/4 is reserved for multicast addresses.

Dynamic Host Configuration Protocol (DHCP)

- It is an application layer client-server protocol which uses UDP and IP.
- The DHCP is used to assign IP addresses to the client computers automatically.
- A DHCP server is provided with a pool of IP addresses from which it assigns IP addresses to the requesting DHCP clients.
- It is a modification of an older protocol called BOOTP.
- Generally an organization or an ISP gets a block of IP addresses from ICANN (Internet Corporation for Assigned Names and Numbers). From these addresses, the administrator can either assign addresses manually to the client systems or install DHCP server to assign addresses automatically.
- DHCP can be configured to assign IP addresses permanently or temporarily.
- In addition to IP address, DHCP can assign other parameters like subnet masks, gateway address, DNS server address, etc to the client systems.
- DHCP is common in all type of routers, hotspots and access points.

DHCP Message Format



Fields:

Opcode: Operation code, request (1) or reply (2)

Htype: Hardware type (Ethernet, ...)

HLen: Length of hardware address

HCount: Maximum number of hops the packet can travel

Transaction ID: An integer set by the client and repeated by the server

Time elapsed: The number of seconds since the client started to boot

Flags: First bit defines unicast (0) or multicast (1); other 15 bits not used

Client IP address: Set to 0 if the client does not know it

Your IP address: The client IP address sent by the server

Server IP address: A broadcast IP address if client does not know it

Gateway IP address: The address of default router

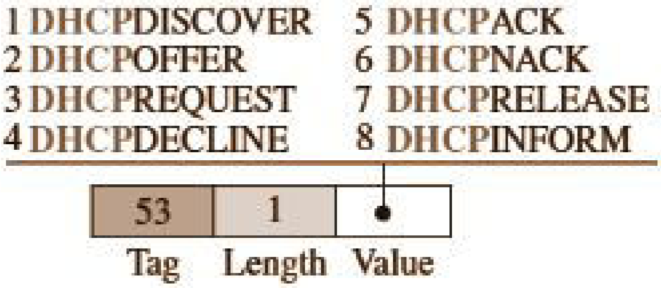
Server name: A 64-byte domain name of the server

Boot file name: A 128-byte file name holding extra information

Options: A 64-byte field with dual purpose described in text

Common format for request and reply is given below.

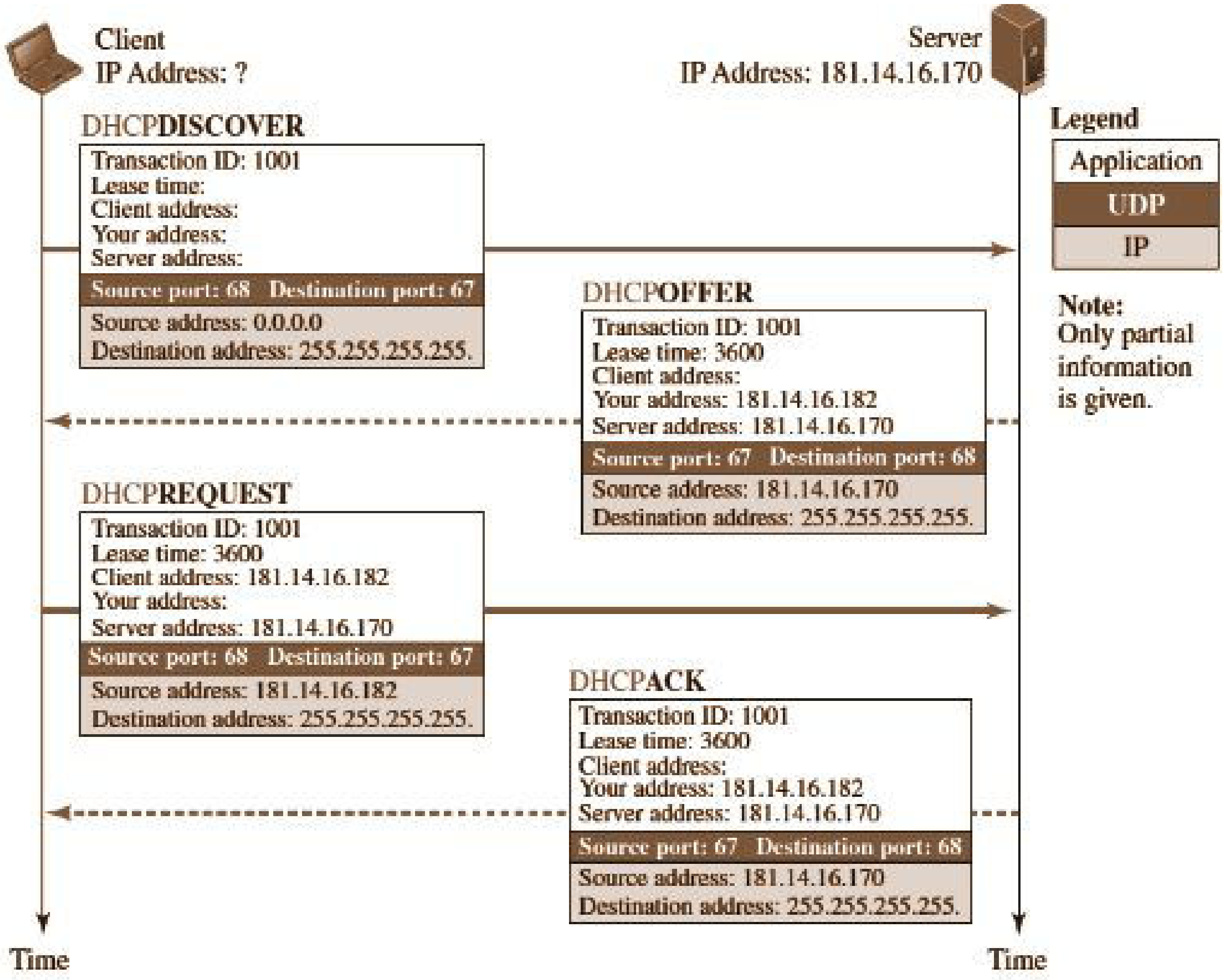
- The 64-byte option field has a dual purpose. It can carry either additional information or some specific vendor information.
- The server uses a number, called a magic cookie, in the format of an IP address with the value of **99.130.83.99**. When the client finishes reading the message, it looks for this magic cookie. If present, the next 60 bytes are options.
- An option is composed of three fields: a 1-byte tag field, a 1-byte length field, and a variable-length value field.
- If the tag field is 53, the value field defines one of the 8 message types given below.



DHCP Operation

The DHCP employs a connectionless service model, using the User Datagram Protocol (UDP). It is implemented with two UDP port numbers for its operations which are the same as for the BOOTP protocol. UDP port number 67 is the destination port of a server, and UDP port number 68 is used by the client.

DHCP operations fall into four phases: server discovery, IP lease offer, IP lease request, and IP lease acknowledgement. These stages are often abbreviated as DORA for Discovery, Offer, Request, and Acknowledgement.



DHCPDISCOVER

Transaction ID: 1001
Lease time:
Client address:
Your address:
Server address:
Source port: 68 Destination port: 67
Source address: 0.0.0.0
Destination address: 255.255.255.255.

DHCPOFFER

Transaction ID: 1001
Lease time: 3600
Client address:
Your address: 181.14.16.182
Server address: 181.14.16.170
Source port: 67 Destination port: 68
Source address: 181.14.16.170
Destination address: 255.255.255.255.

DHCPREQUEST

Transaction ID: 1001
Lease time: 3600
Client address: 181.14.16.182
Your address:
Server address: 181.14.16.170
Source port: 68 Destination port: 67
Source address: 181.14.16.182
Destination address: 255.255.255.255.

DHCPACK

Transaction ID: 1001
Lease time: 3600
Client address:
Your address: 181.14.16.182
Server address: 181.14.16.170
Source port: 67 Destination port: 68
Source address: 181.14.16.170
Destination address: 255.255.255.255.

1. The joining host (client) creates a DHCPDISCOVER message in which only the transaction-ID field is set to a random number. No other field can be set because the host has no knowledge with which to do so. This message is encapsulated in a UDP user datagram with the source port set to

68 and the destination port set to 67. The user datagram is encapsulated in an IP datagram with the source address set to 0.0.0.0 (“this host”) and the destination address set to 255.255.255.255 (broadcast address). The reason is that the joining host knows neither its own address nor the server address.

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

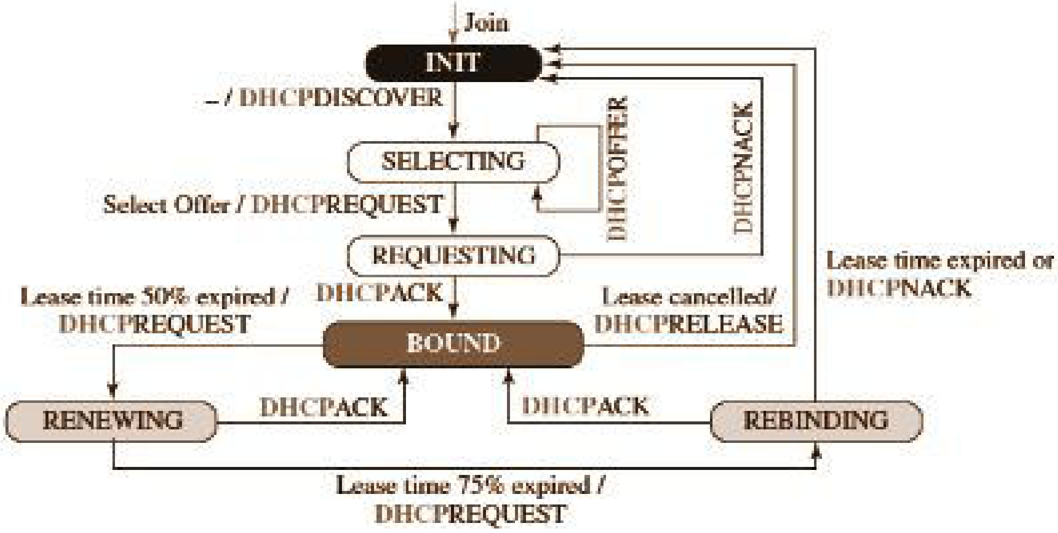
3. The joining host receives one or more offers and selects the best of them. The joining host then sends a DHCPREQUEST message to the server that has given the best offer. The fields with known value are set. The message is encapsulated in a user datagram with port numbers as the first message. The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

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

In the DHCPACK message, the server defines the *pathname of a file* in which the client can find complete information of the server. The client can then use a file transfer protocol (FTP) to obtain the rest of the needed information.

Transition States

The following is the state transition diagram of a DHCP client, if it is considered as a Finite State Machine (FSM).



Join
INIT
– / DHCPDISCOVER
SELECTING
DHCPOFFER
DHCPNACK
Select Offer / DHCPREQUEST
REQUESTING
Lease time expired or DHCPNACK
Lease time 50% expired / DHCPREQUEST
DHCPACK
Lease cancelled/ DHCPRELEASE
BOUND
RENEWING
DHCPACK
DHCPACK
REBINDING
Lease time 75% expired / DHCPREQUEST

INTERNET PROTOCOL (IP)

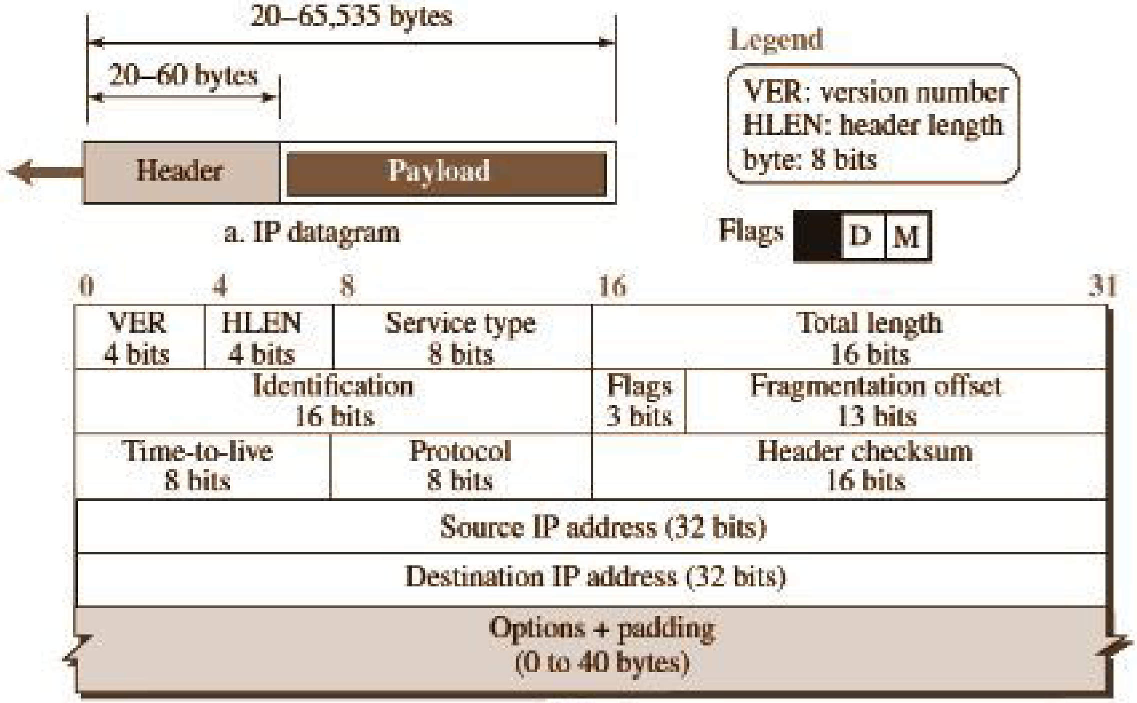
Datagram Format

Packets used by the IP are called datagrams.

A datagram is a variable-length packet consisting of two parts: header and payload (data).

The header is 20 to 60 bytes in length and contains information essential to routing and delivery.

Version Number: The 4-bit version number (VER) field defines the version of the IPv4 protocol, which, obviously, has the value of 4.



b. Header

Header Length: The 4-bit header length (HLEN) field defines the total length of the datagram header in *4-byte words*. ie, a length of 5 means 5 x 4 bytes = 20 bytes. The header size ranges between 5 to 15. (ie, 20 bytes to 60 bytes).

Service Type: In the original design of the IP header, this field was referred to as type of service (TOS), which defined how the datagram should be handled. Examples are minimize delay, maximize throughput, maximize reliability, etc. It was never widely used as originally defined, and its meaning has been subsequently re-defined for use by a technique called *Differentiated Services (DS or DiffServ).*

Total Length: This 16-bit field defines the total length (header plus data) of the IP datagram in bytes. A 16-bit number can define a total length of up to 65,535 (when all bits are 1s). However, the size of the datagram is normally much less than this.

Length of data = total length − (HLEN) × 4

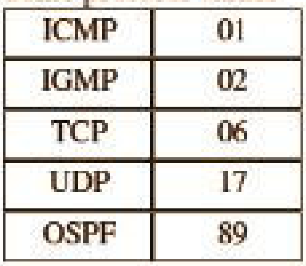
One use of this field is that, minimum payload length of an Ethernet frame is 46 bytes. If the size of the datagram is less than 46, zeros are padded to make it 46. When a receiver decapsulates the datagram, it needs to check the total length field to determine how much is really data and how much is padding.

Identification, Flags, and Fragmentation Offset: These three fields are related to the fragmentation of the IP datagram when the size of the datagram is larger than the underlying network can carry.

Time-to-live (TTL): Due to some malfunctioning of routing protocols (discussed later) a datagram may be circulating in the Internet, visiting some networks over and over without reaching the destination. This may create extra traffic in the Internet. The time-to-live (TTL) field is used to control the maximum number of hops (routers) visited by the datagram. When a source host sends the datagram, it stores a number in this field. This value is approximately two times the maximum number of routers between any two hosts. Each router that processes the datagram decrements this number by one. If this value, after being decremented, is zero, the router discards the datagram.

Protocol: Shows the protocol used by the payload. This field provides multiplexing at the source and demultiplexing at the destination. Some protocols and values are given below.

Some protocol values



ICMP	01
IGMP	02
TCP	06
UDP	17
OSPF	89

Header checksum: IP does not have an error checking mechanism for the payload. But it has a checksum field to ensure that the **header** information (source and destination IP addresses, fragmentation info, protocol field etc) are received at the destination without any alteration or corruption. Since the value of some fields, such as TTL, which are related to fragmentation and options, may change from router to router, the checksum needs to be recalculated at each router.

Source and Destination Addresses: 32 bit IPv4 addresses of the source and destination machines. The value of these fields must remain unchanged during the time the IP datagram travels from the source host to the destination host.

Options: A datagram header can have up to 40 bytes of options. Options can be used for network testing and debugging.

Payload: Payload is the packet coming from other protocols that use the service of IP.

Fragmentation

There will be a number of protocols in the network that are connected through routers.

Each router decapsulates the IP datagram from the frame it receives, processes it, and then encapsulates it in another frame.

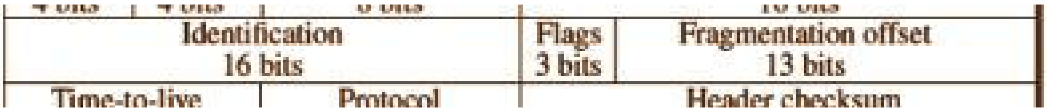
The frame length in each protocol will be different.

The maximum payload length that can be encapsulated in a frame is called the Maximum Transfer Unit (MTU). MTU in an Ethernet Frame is 1500 bytes.

When a datagram is to be propagated from a link with larger MTU to a link with smaller MTU, the router that connects these two links divides the datagram accordingly. This process is called fragmentation. A datagram may be fragmented many times in many routers before it reaches the destination. All the fragments are reassembled only at the destination device. In each fragmentation, most of the header information are copied to all fragments except the flags, fragmentation offset, and total length.The checksum also is recalculated.

Fields Related to Fragmentation

There are three fields in an IP datagram that are related to fragmentation: identification, flags and fragmentation offset.

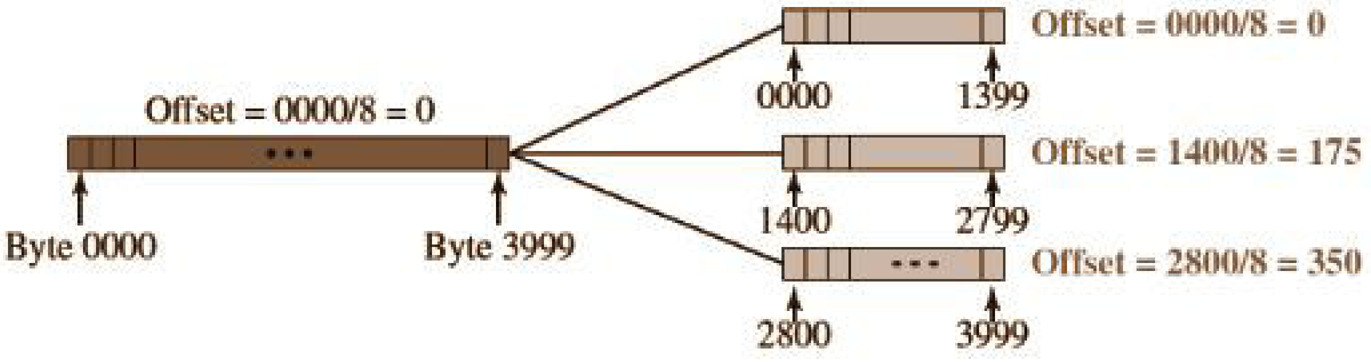


Identification 16 bits		Flags 3 bits	Fragmentation offset 13 bits
Time-to-live	Protocol	Header checksum	

The 16-bit identification field in combination with the source IP address identifies a datagram originating from the source host. The sender keeps a counter whose value is used as the identification number. This value is copied to all fragments of the same datagram and used to identify all the fragments that are parts of the same datagram.

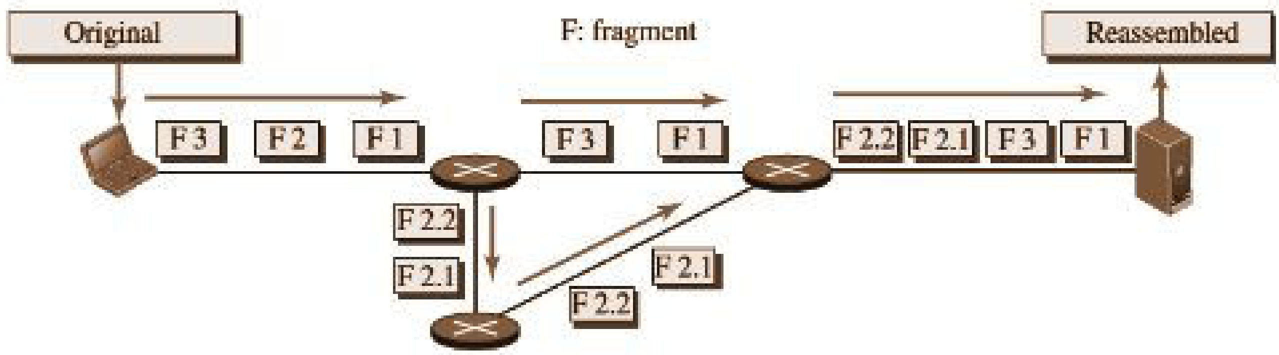
The 3-bit flags field defines three flags. The leftmost bit is reserved (not used). The second bit (D bit) is called the **do not fragment** bit. If its value is 1, the machine must not fragment the datagram. If it cannot pass the datagram through any available physical network, it discards the datagram and sends an ICMP error message to the source host. If its value is 0, the datagram can be fragmented if necessary. The third bit (M bit) is called the **more fragment** bit. If its value is 1, it means the datagram is not the last fragment; there are more fragments after this one. If its value is 0, it means this is the last or only fragment.

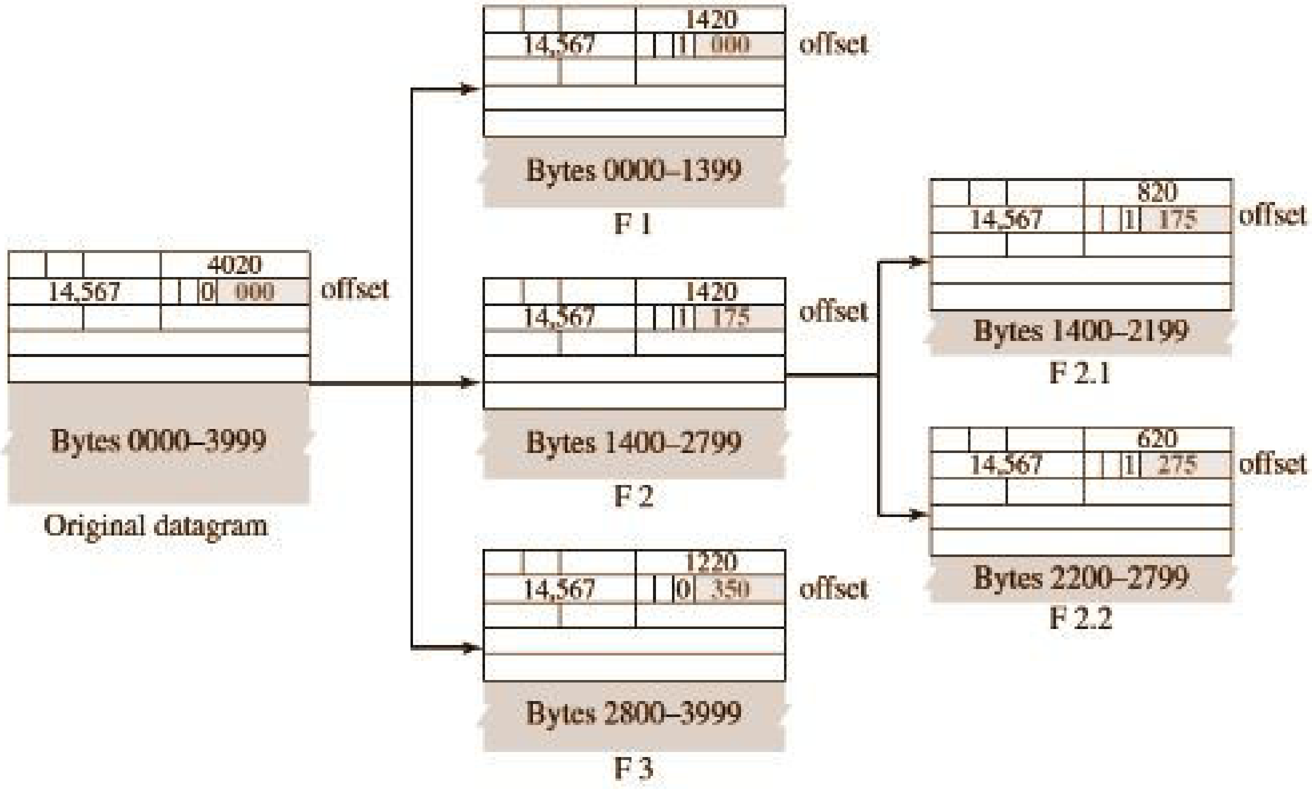
The 13-bit fragmentation offset field shows the relative position of this fragment with respect to the whole datagram. It is the offset of the data in the original datagram measured in units of 8 bytes. Figure shows a datagram with a data size of 4000 bytes fragmented into three fragments. The bytes in the original datagram are numbered 0 to 3999. The first fragment carries bytes 0 to 1399. The offset for this datagram is 0/8 = 0. The second fragment carries bytes 1400 to 2799; the offset value for this fragment is 1400/8 = 175. Finally, the third fragment carries bytes 2800 to 3999. The offset value for this fragment is 2800/8 = 350.



Example of fragmentation:

Consider source with MTU 4000 Bytes. The first link has an MTU 1400 Bytes and the downward link has an MTU 800 Bytes. In the below diagram, the second datagram is again fragmented with an MTU 800 Bytes.





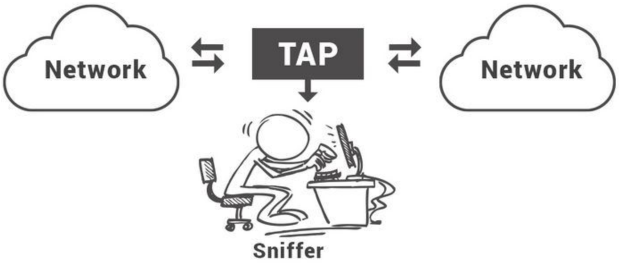
Security of IPv4 Datagrams

No security was provided for the IPv4 protocol.

There are three security issues that are particularly applicable to the IP protocol: packet sniffing, packet modification, and IP spoofing.

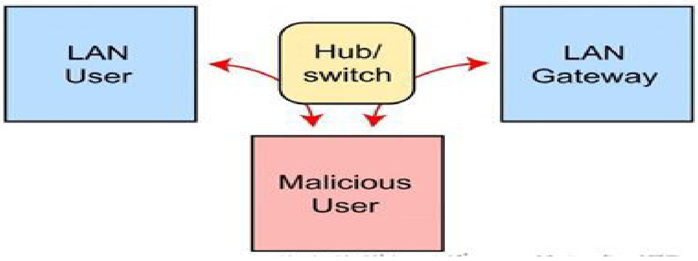
Packet Sniffing:

An intruder may intercept an IP packet and make a copy of it. Packet sniffing is a passive attack, in which the attacker does not change the contents of the packet. This type of attack is very difficult to detect because the sender and the receiver may never know that the packet has been copied. Although packet sniffing cannot be stopped, encryption of the packet can make the attacker's effort useless. The attacker may still sniff the packet, but the content is not detectable.



Packet Modification:

The second type of attack is to modify the packet. The attacker intercepts the packet, changes its contents, and sends the new packet to the receiver. The receiver believes that the packet is coming from the original sender. This type of attack can be detected using a data integrity mechanism. The receiver, before opening and using the contents of the message, can use this mechanism to make sure that the packet has not been changed during the transmission.



IP Spoofing:

An attacker can masquerade as somebody else and create an IP packet that carries the source address of another computer. An attacker can send an IP packet to a bank pretending that it is coming from one of the customers. This type of attack can be prevented using an origin authentication mechanism.

IPSec

The IP packets today can be protected from the previously mentioned attacks using a protocol called IPSec (IP Security). This protocol, which is used in conjunction with the IP protocol, creates a connection-oriented service between two entities in which they can exchange IP packets without worrying about the three attacks discussed above.

Defining Algorithms and Keys: Before making secure channel, the two devices agree on some available algorithms and keys to be used for security purposes.

Packet Encryption: The packets exchanged between two parties can be encrypted for privacy using one of the encryption algorithms and a shared key agreed upon in the first step.

Data Integrity: Data integrity guarantees that the packet is not modified during the transmission. If the received packet does not pass the data integrity test, it is discarded.

Origin Authentication: IPSec can authenticate the origin of the packet to be sure that the packet is not created by an imposter.

ROUTING ALGORITHMS

Several algorithms have been designed to find the best route a packet can travel.

The difference between these methods are in the way they interpret the least cost and the way they create the least-cost tree for each node.

Major routing algorithms are Distance-Vector routing, Link-State routing and Path-Vector routing.

Distance-Vector Routing (DV Routing)

Outline: 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. A router continuously tells all of its neighbors what it knows about the whole internet.

DV routing uses **Bellman-Ford Equation** to find the least cost between nodes. This equation is the heart of DV routing.

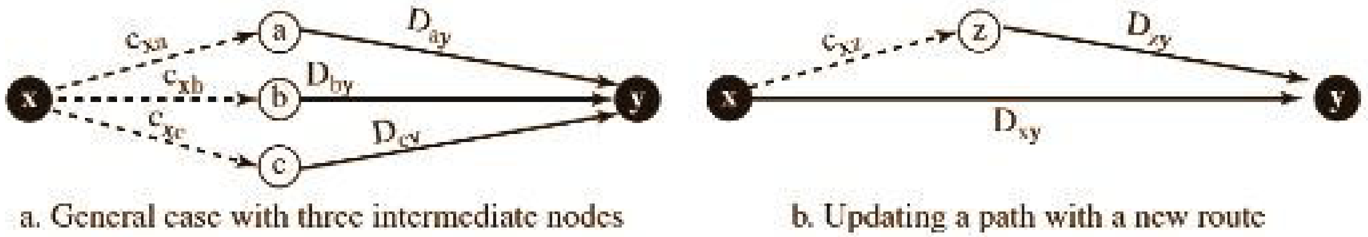
Bellman-Ford Equation: This equation is used to find the least cost (shortest distance) between a source node, x, and a destination node, y, through some intermediary nodes (a, b, c, . . .) when the costs between the source and the intermediary nodes and the least costs between the intermediary nodes and the destination are given.

The following shows the general case in which $\mathbf{D_{ij}}$ is the shortest distance and $\mathbf{c_{ij}}$ is the cost between nodes **i** and **j**.

$$D_{xy} = \min\{(c_{xa} + D_{ay}), (c_{xb} + D_{by}), (c_{xc} + D_{cy}), \ldots\}$$

To update an existing least cost with a least cost through an intermediary node, such as **z**, if the latter is shorter;

$$D_{xy} = \min\{D_{xy}, (c_{xz} + D_{zy})\}$$



a. General case with three intermediate nodes

b. Updating a path with a new route

Thus this equation helps to find new least cost paths from existing least cost paths.

Distance Vectors:

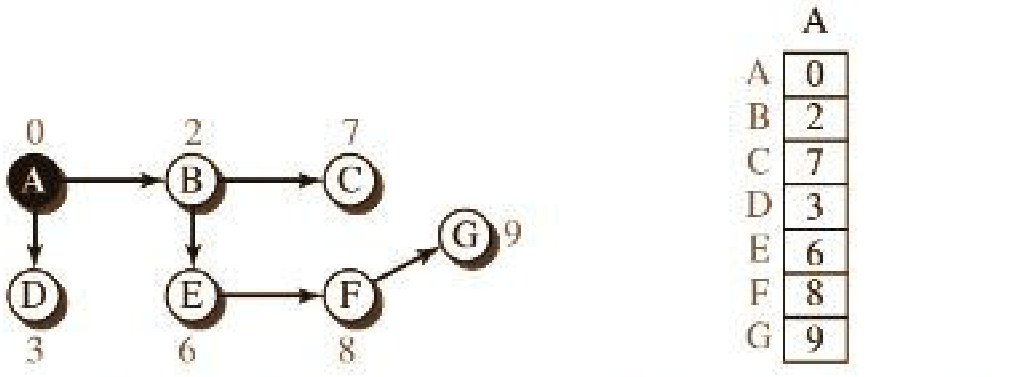
A least-cost tree is a combination of least-cost paths from the root of the tree to all destinations.

A distance vector in a node is a one-dimensional array to represent the tree with the node as the root.

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*.

A distance vector does not give the path to the destinations; it gives only the least costs to the destinations.

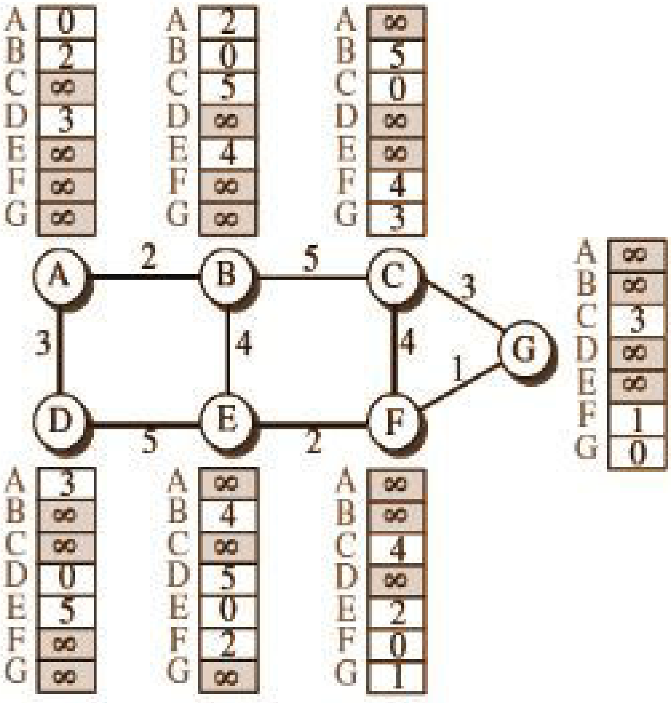
Example: See the figure;



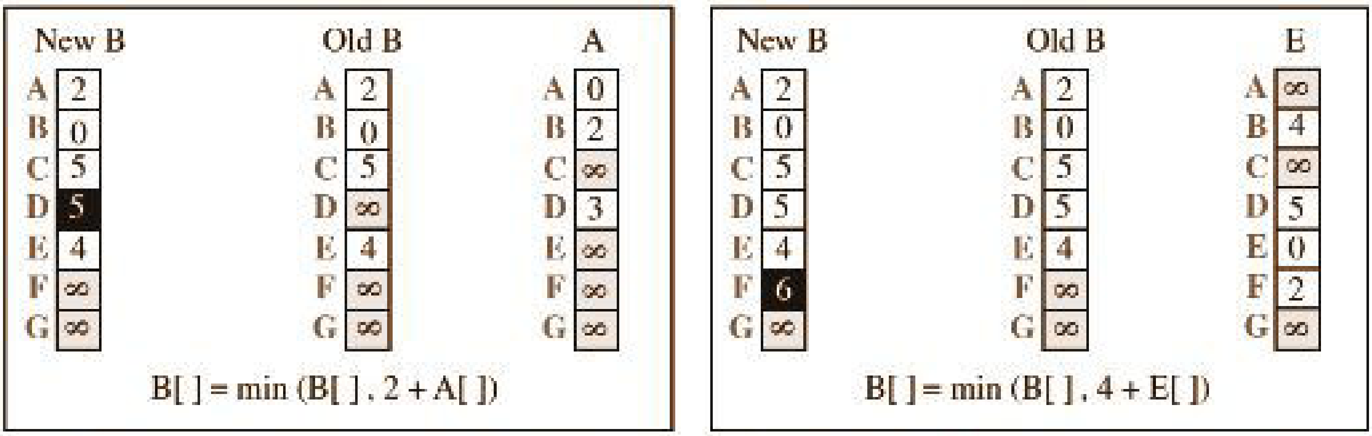
a. Tree for node A b. Distance vector for node A

Distance Vector creation: Each node in an internet, when it is booted, creates a very rudimentary distance vector. 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.

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 (second case). It also updates the number if nodes in the internet.


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

Initial vector creation



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

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

	New B		Old B		E
A	2	A	2	A	∞
B	0	B	0	B	4
C	5	C	5	C	∞
D	5	D	5	D	5
E	4	E	4	E	0
F	6	F	∞	F	2
G	∞	G	∞	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.

Exchanging vectors eventually stabilizes the system and allows all nodes to find the ultimate least cost between themselves and any other node. Then it uses the DV algorithm to find the least path.

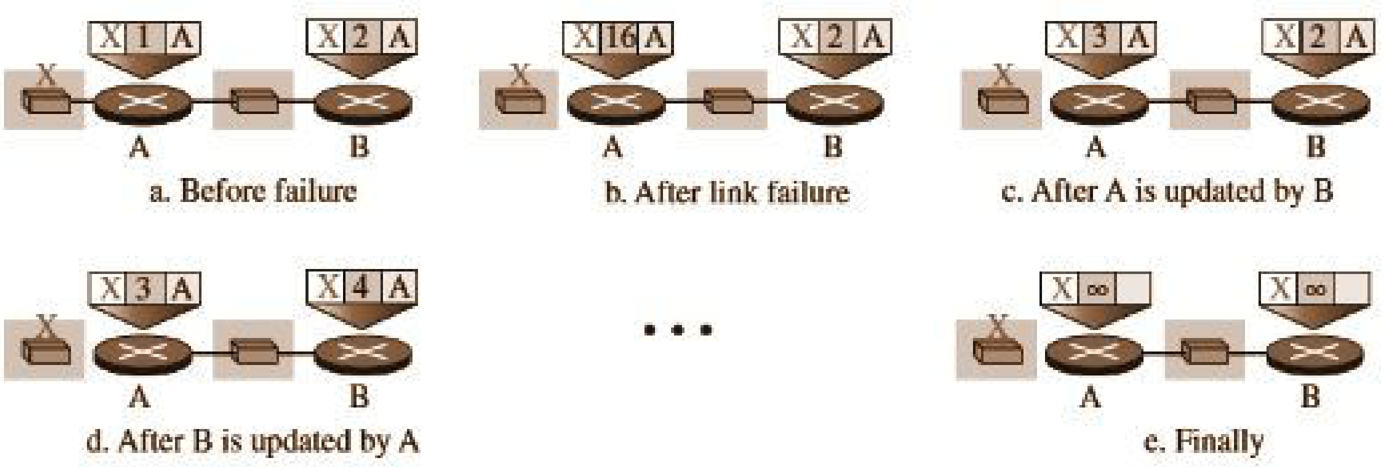
Distance-Vector Routing Algorithm

```
Distance_Vector_Routing ( )
{
        // Initialize (create initial vectors for the node)
        D[myself ] = 0
        for (y = 1 to N)
           {
                if (y is a neighbor)
                        D[y] = c[myself ][y]
                else
                        D[y] = ∞
           }
        Send vector {D[1], D[2], …, D[N]} to all neighbors
         // Update (improve the vector with the vector received from a neighbor)
         repeat (forever)
           {
                 wait (for a vector Dw from a neighbor w or any change in the link)
                 for (y = 1 to N)
                    {
                         D[y] = min [D[y], (c[myself ][w] + Dw[y ])] // Bellman-Ford equation
                    }
                if (any change in the vector)
```

```
            send vector {D[1], D[2], …, D[N]} to all neighbors
        }
} // End of Distance Vector
```

Count to Infinity problem

A problem with distance-vector routing is that any decrease in cost (good news) propagates quickly, but any increase in cost (bad news) will propagate slowly. For a routing protocol to work properly, if a link is broken (cost becomes infinity), 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.



a. Before failure

b. After link failure

c. After A is updated by B

d. After B is updated by A

e. Finally

Count to infinity problem

Split Horizon solution

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. If, according to its table, node B thinks that the optimum route to reach X is via A, it does not need to advertise this piece of information to A; the information has come from A (A already knows). In this case, when X goes down, node A keeps the value of infinity as the distance to X. Later, when node A sends its forwarding table to B, node B also corrects its forwarding table. The system becomes stable after the first update: both node A and node B know that X is not reachable.

Poison Reverse strategy

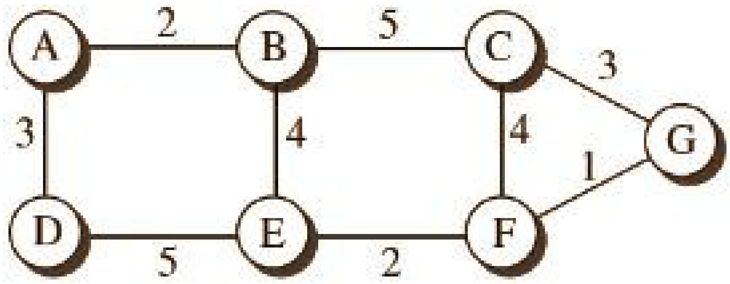
Using the split-horizon strategy has one drawback. Normally, the corresponding protocol uses a timer, and if there is no news about a route, the node deletes the route from its table. When node B in the previous scenario eliminates the route to X from its advertisement to A, node A cannot guess whether this is due to the split-horizon strategy (the source of information was A) or because B has not received any news about X recently. In the poison reverse strategy B can still advertise the value for X, but if the source of information is A, it can replace the distance with infinity as a warning.

Link-State Routing (LS Routing)

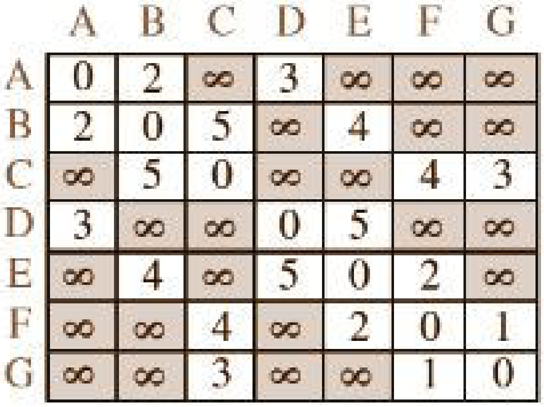
- The cost associated with an edge (link) defines the state of the link. Hence the term.
- Links with lower costs are preferred.
- If the cost of a link is infinity, it means that the link does not exist or has been broken.

Link-State Database (LSDB): The term link-state defines the characteristic of a link (an edge) that represents a network in the internet. To create a least-cost tree with this method, each node needs to have a complete map of the network. The collection of states for all links is called the link-state database (LSDB). There is only one LSDB for the whole internet; each node needs to have a duplicate of it to be able to create the least-cost tree. The LSDB can be represented as a two-dimensional array(matrix) in which the value of each cell defines the cost of the corresponding link.

Example of Link-State DataBase (LSDB):



a. The weighted graph

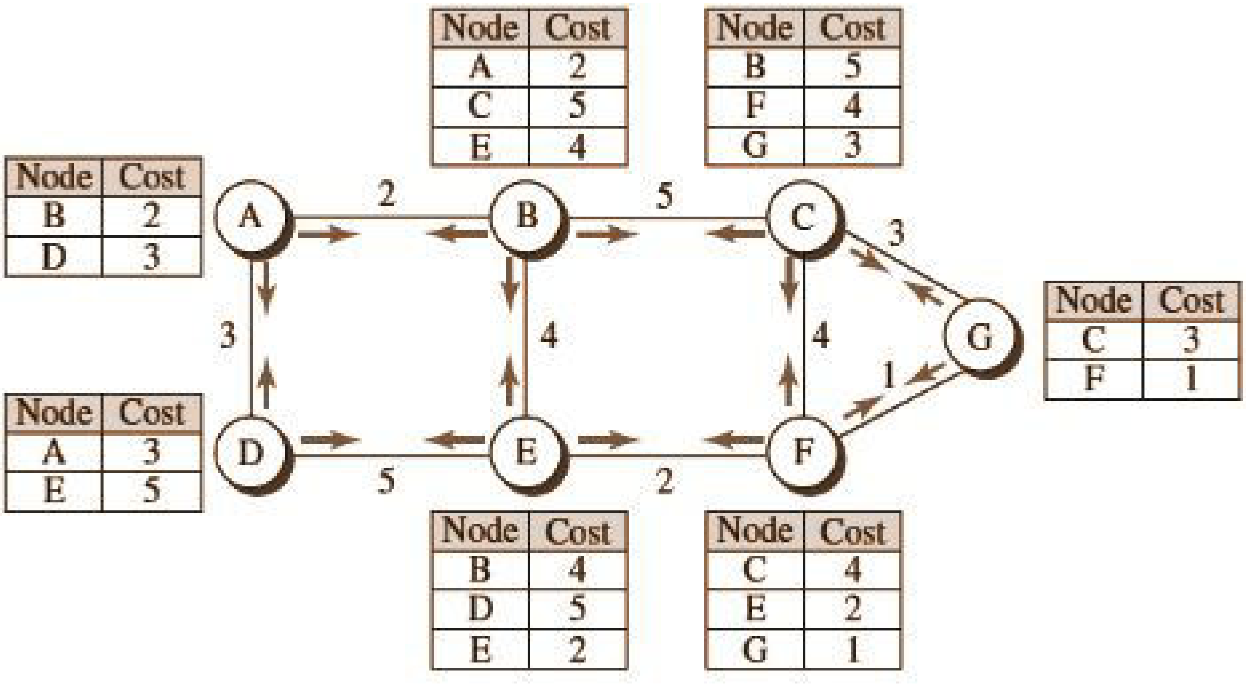


	A	B	C	D	E	F	G
A	0	2	∞	3	∞	∞	∞
B	2	0	5	∞	4	∞	∞
C	∞	5	0	∞	∞	4	3
D	3	∞	∞	0	5	∞	∞
E	∞	4	∞	5	0	2	∞
F	∞	∞	4	∞	2	0	1
G	∞	∞	3	∞	∞	1	0

b. Link state database

Each node creates LSDB of the whole network by a process called **flooding**. Each node sends some greeting messages to all its immediate neighbors 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.

When a node receives an LSP from one of its interfaces, it compares the LSP with the copy it may already have. When a node receives an LSP, it compares that LSP with its own copy of LSP. Based on the sequence number, it retains the latest version of the LSP. It then sends a copy of it out of each interface except the one from which the packet arrived (this is called *flooding*). This guarantees that flooding stops somewhere in the network (where a node has only one interface). After receiving all new LSPs, each node creates the comprehensive LSDB. This LSDB is the same for each node. Then it starts creating least-cost tree using **Dijkstra Algorithm.**



Node	Cost
B	2
D	3

Node	Cost
A	2
C	5
E	4

Node	Cost
B	5
F	4
G	3

Node	Cost
C	3
F	1

Node	Cost
A	3
E	5

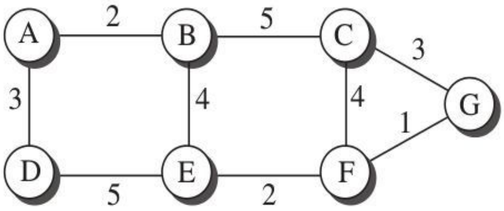
Node	Cost
B	4
D	5
E	2

Node	Cost
C	4
E	2
G	1

Initial stage of LSP formation

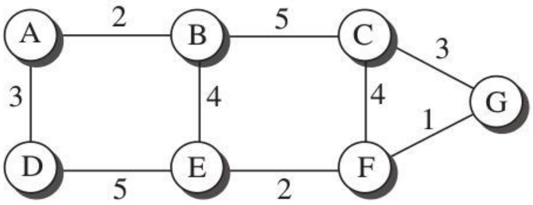
This algorithm has the following steps:

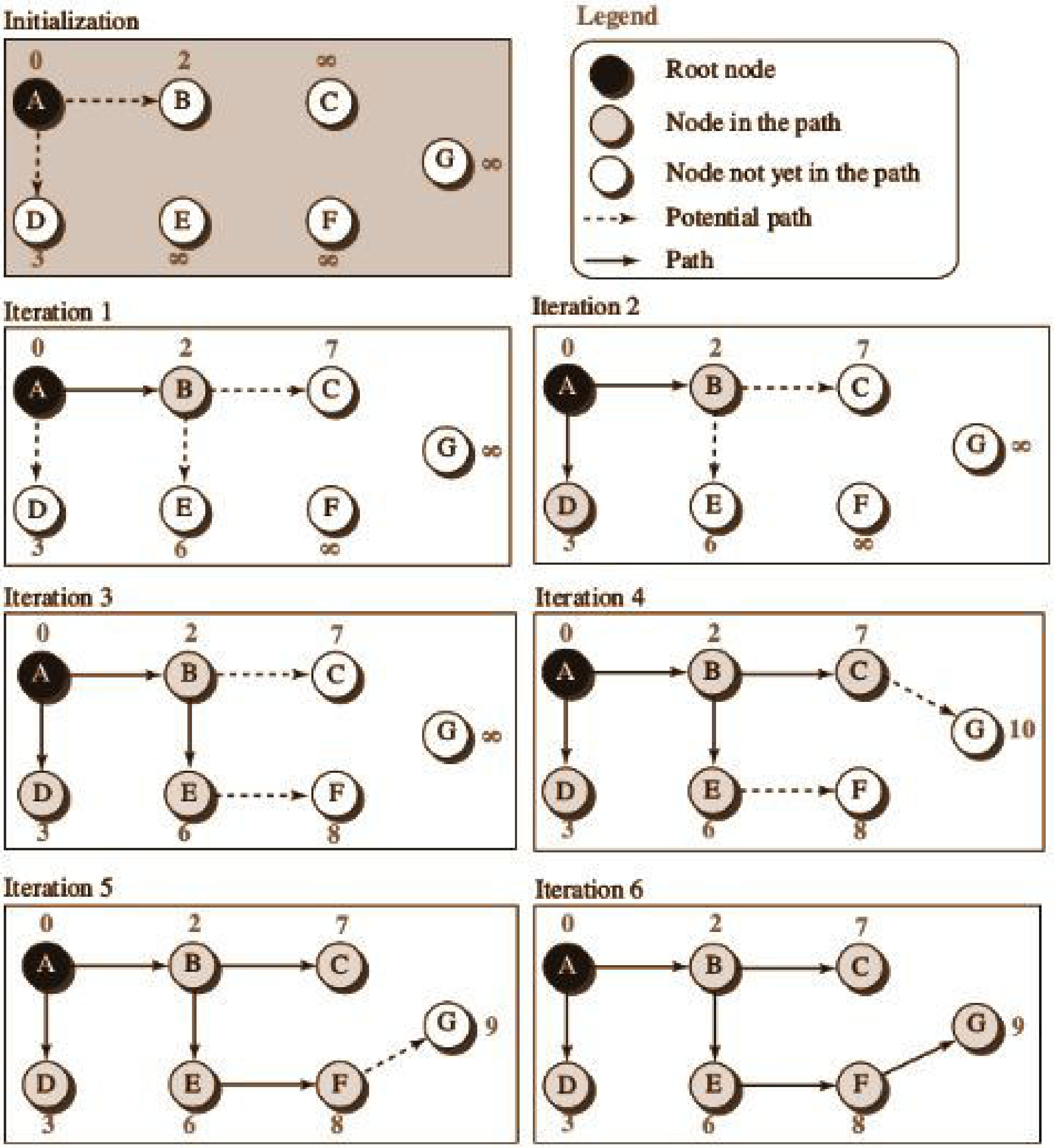
1. The node chooses itself as the root of the tree, creating a tree with a single node, and sets the total cost of each node based on the information in the LSDB.
2. The node selects one node, among all nodes not in the tree, which is closest to the root, and adds this to the tree. After this node is added to the tree, the cost of all other nodes not in the tree needs to be updated because the paths may have been changed.
3. The node repeats step 2 until all nodes are added to the tree.



```
Dijkstra's Algorithm ( )
{
        // Initialization
        Tree = {root}          // Tree is made only of the root
        for (y = 1 to N)       // N is the number of nodes
        {
                if (y is the root)
                        D[y] = 0              // D[y] is shortest distance from root to node y
                else if (y is a neighbor)
                        D[y] = c[root][y]     // c[x][y] is cost between nodes x and y in LSDB
                else
                        D[y] = ∞
        }
// Calculation
        repeat
        {
                find a node w, with D[w] minimum among all nodes not in the Tree
                Tree = Tree ∪ {w} // Add w to tree
                        // Update distances for all neighbors of w
                for (every node x, which is a neighbor of w and not in the Tree)
                {
                        D[x] = min{D[x], (D[w] + c[w][x])}
                }
        } until (all nodes included in the Tree)
} // End of Dijkstra
```

Figure for the steps in LS Routing:





Initialization

Legend

Iteration 1

Iteration 2

Iteration 3

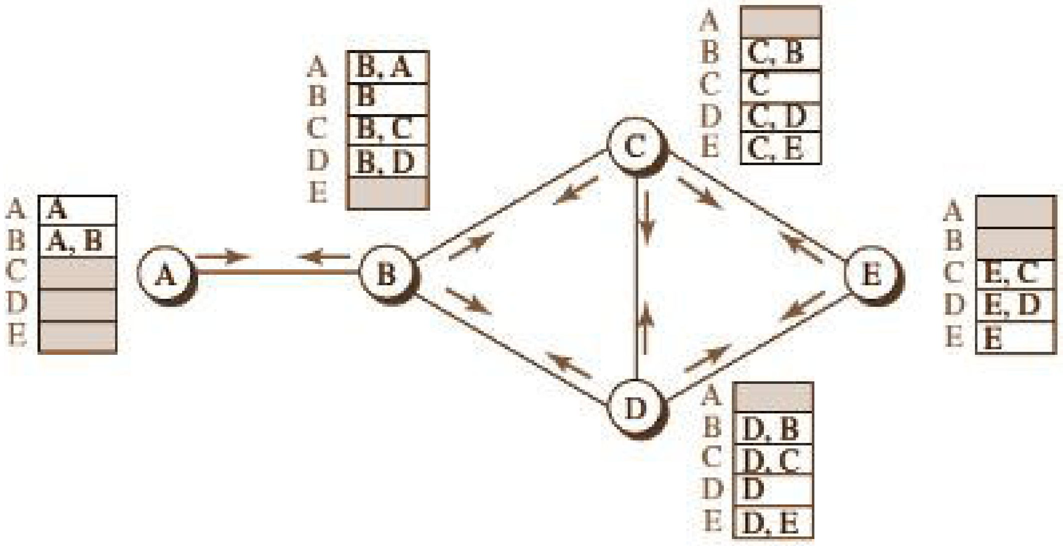
Iteration 4

Iteration 5

Iteration 6

Path-Vector Routing (PV Routing)

- Not based on least-cost path, but best-cost path.
- Does not have the drawbacks of DV and LS routing.
- Mainly used by Internet Service Providers (ISP).
- Used when certain policies are to be imposed on routers; such as to avoid some path.
- The best route is determined by the source using the policy it imposes on the route. In other words, the source can control the path.
- Thus each router has its own spanning tree.
- A source may apply several policies at the same time.
- When a node is booted up, it creates a path vector based on the information from its neighbors; then passes these path vectors to its neighbors.

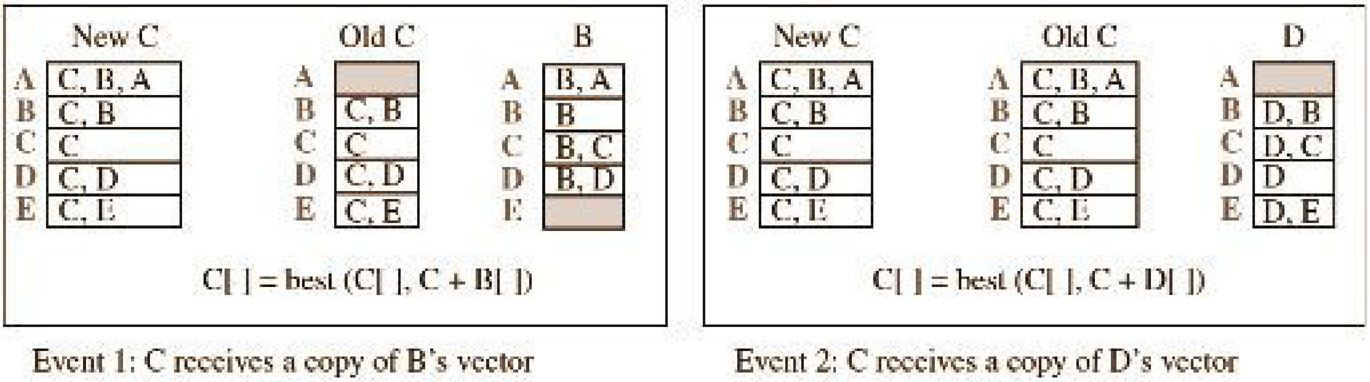


Path vectors made at booting time

- Then the neighbors update their vectors by using these informations and so on.
- The updation occurs based on the equation;

  Path(x, y) = best {Path(x, y), [(x + Path(v , y)]} for all v 's in the internet.

  The operator (+) means to add x to the beginning of the path.
- If Path (v, y) includes x, that path is discarded to avoid a loop in the path.



	New C	Old C	B
A	C, B, A		B, A
B	C, B	C, B	B
C	C	C	B, C
D	C, D	C, D	B, D
E	C, E	C, E	

C[ ] = best (C[ ], C + B[ ])

Event 1: C receives a copy of B's vector

	New C	Old C	D
A	C, B, A	C, B, A	
B	C, B	C, B	D, B
C	C	C	D, C
D	C, D	C, D	D
E	C, E	C, E	D, E

C[ ] = best (C[ ], C + D[ ])

Event 2: C receives a copy of D's vector

Two instances of C's path vector updation

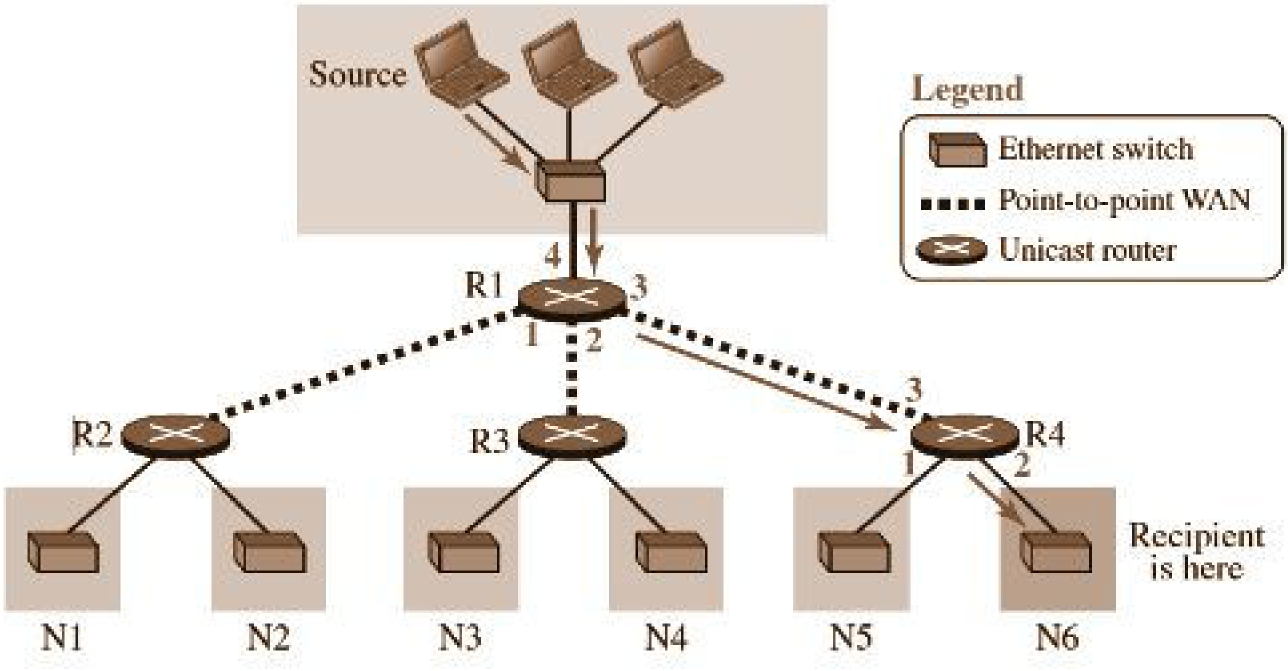
<u>Path-Vector Algorithm (for a node)</u>

```
Path_Vector_Routing ( )
{
        // Initialization
        for (y = 1 to N)
        {
                if (y is myself)
                        Path[y] = myself
                else if (y is a neighbor)
                        Path[y] = myself + neighbor node
                else
                        Path[y] = empty
        }
```

```
    Send vector {Path[1], Path[2], …, Path[y]} to all neighbors
    // Update
    repeat (forever)
    {
        wait (for a vector $Path_w$ from a neighbor w)
        for (y = 1 to N)
        {
            if ($Path_w$ includes myself)
                discard the path // Avoid any loop
            else
                Path[y] = best {Path[y], (myself + $Path_w$[y])}
        }
        If (there is a change in the vector)
            Send vector {Path[1], Path[2], …, Path[y]} to all neighbors
    }
} // End of Path Vector
```

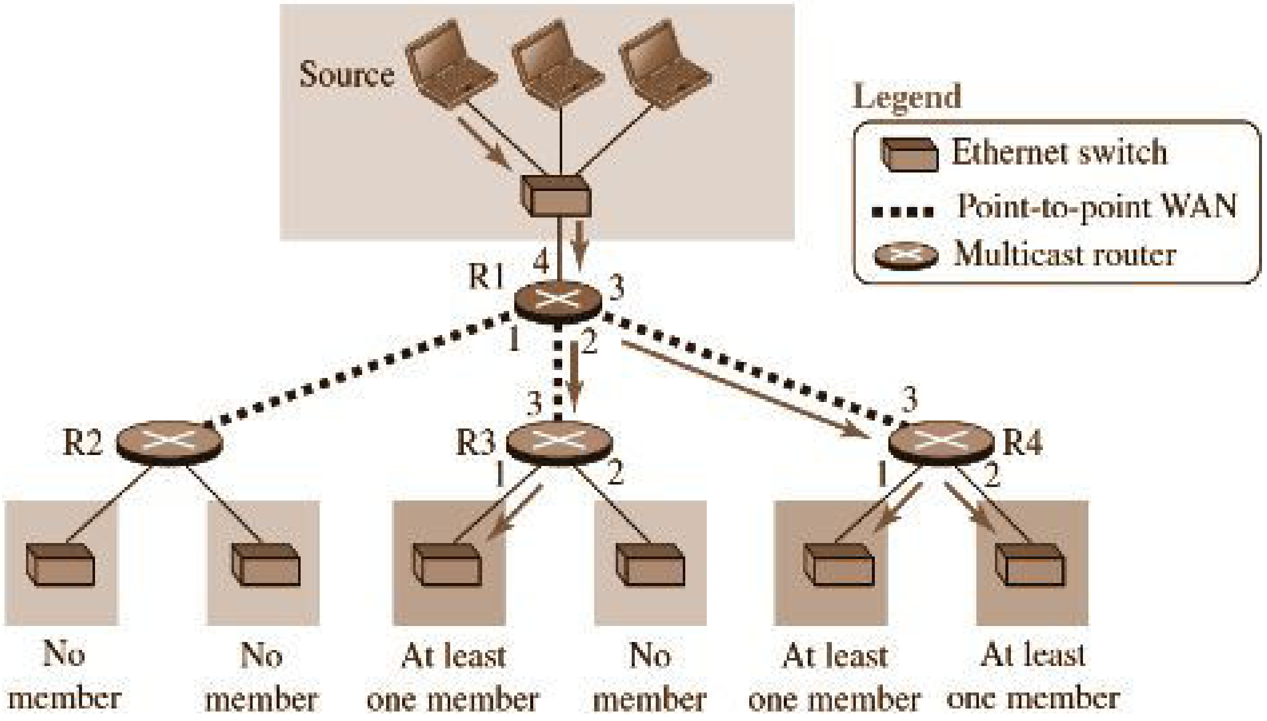
Unicasting

- One source and one destination network.
- One to one relationship between source and destination.
- Each router in the path of the datagram forward the packet to one and only one of its interfaces.



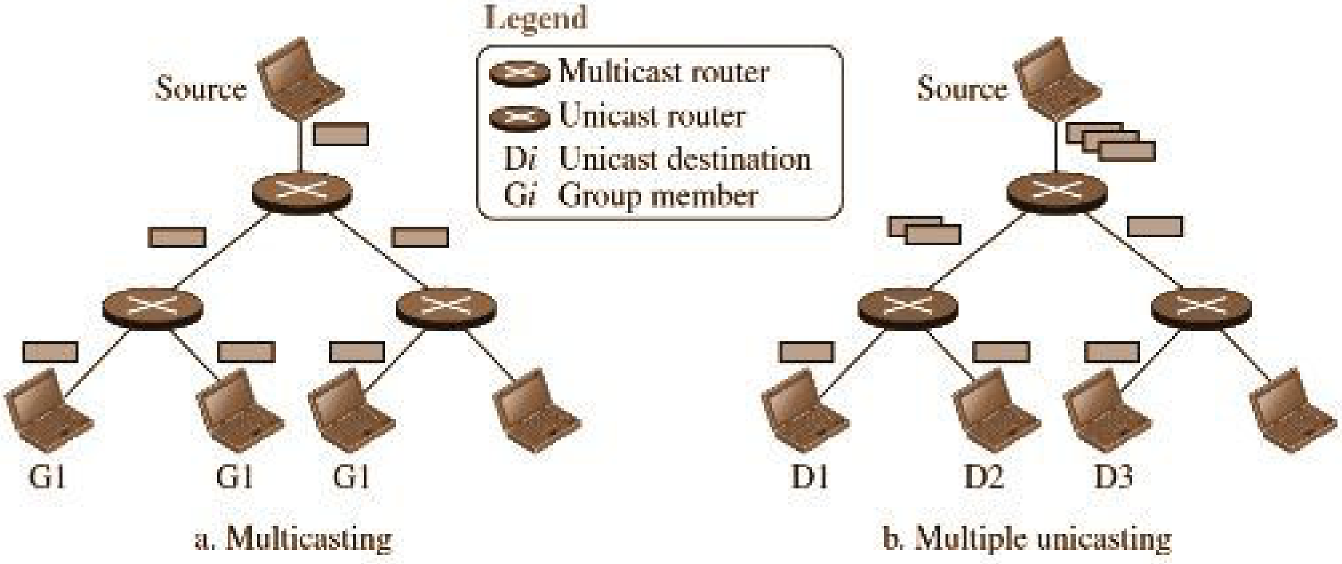
Multicasting

- One source and a group of destinations.
- The relationship is one to many.
- The source address is a unicast address, but the destination address is a group address.



Multicasting versus Multiple Unicasting

Multicasting	Multiple unicasting
➢ Single packet from the source that is duplicated by the routers. ➢ Same destination address on all packets. ➢ Only a single copy of the packet travels between any two routers. ➢ Needs less bandwidth. ➢ Only a single message, hence there is no delay. ➢ Eg: Group messaging.	➢ Several packets from the source. ➢ Packets have different destination addresses. ➢ There may be multiple copies traveling between two routers. ➢ Needs more bandwidth. ➢ Delay in packet creation if the recipients are more. ➢ Eg: e-mail message to a number of people.

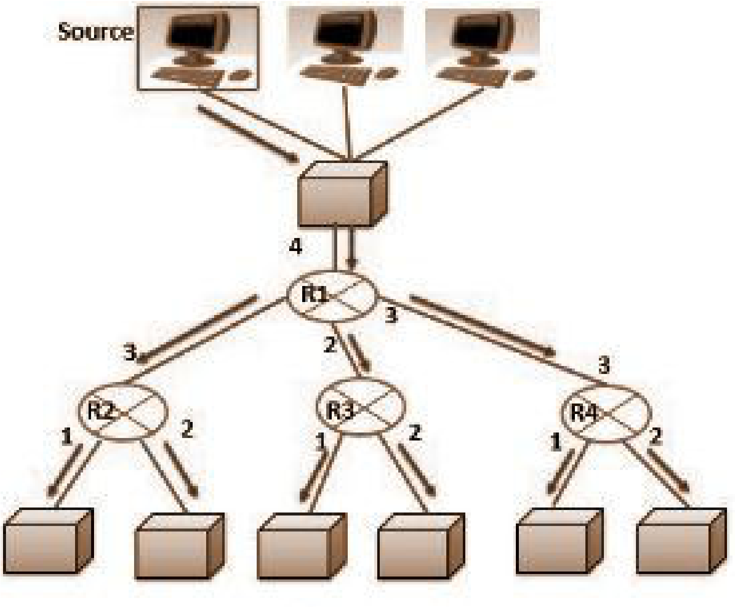


Multicast Applications

- ❏ Access to Distributed Databases
- ❏ Information Dissemination: Different types of business groups
- ❏ Teleconferencing.
- ❏ Distance Learning.

Broadcasting

- One-to-all communication.
- A host sends a packet to all hosts in an internet.
- Not generally used as it may create a huge volume of traffic and use a huge amount of bandwidth.
- Partial broadcasting is done in the Internet.
- Controlled broadcasting may also be done in a domain mostly as a step to achieve multicasting.



Broadcast