

#

PS01CMCA52 Computer Networks

1. Introduction and Data Communication Fundamentals

- Introduction to computer networks
- Classification of computer networks
- Transmission media : guided and unguided media.
- Functions of network connecting devices: Amplifier, Repeater, Bridge, Hub, Switch, Router, Gateway, Modems
- Data transmission concepts : transmission modes, multiplexing, switching technologies, asynchronous and synchronous transmission
- Introduction to Local Area Networks (LANs), LAN topologies,
- Gigabit Ethernet

2. Layered Protocols

- Protocols, Protocol hierarchies
- Design issues for the layers
- The OSI reference model and the TCP/IP reference model
- The Internet Protocol (IP), IP addresses, Subnets,
- Introduction to Transmission Control Protocol (TCP), The TCP segment header
- Introduction to User Datagram Protocol (UDP)

A Computer Network : an *interconnected* collection of autonomous *computers*.

- Two computers are said to be interconnected if they are able to exchange information.
- Networks may have various sizes, shapes and forms.

Advantages of Computer Networks

- Resource sharing : users can share programs, equipment and data without regard to the physical location of the resource and users.
- Reliability : improved by providing alternative sources of supply. e.g. selected files may be replicated on multiple machines, so if one of them is unavailable, the other copies could be used. If one machine crashes, the system as a whole can still survive. Presence of multiple CPUs : if one goes down, the others may be used to take over its work.
- Economics : Microprocessors offer a better price/performance ratio than mainframes.
- Incremental growth : Computing power can be added in small increments.
- Communication : Make a human-to-human communication easier. For example, by electronic mail.
- Speed : more total computing power than a mainframe machine.
- Flexibility : Spread the workload over the available machines in the most cost effective way.

Classification of Computer Networks

- **Classification by size**

Personal Area Networks (PANs)
Local Area Networks (LANs)
Metropolitan Area Networks (MANs)
Wide Area Networks (WANs)
The Internet

- **Classification by topology**

Bus
Star
Ring
Tree
Complete
Intersecting Rings
Irregular

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

- **Local Area Networks (LANs)**: Local Area Networks (LANs) are privately-owned networks within a single building or a campus of up to a few kilometers in size. LANs are confined to a small area.
- **Metropolitan Area networks (MANs)** : A Metropolitan Area network (MAN) covers a city. The best known example of a MAN is the cable television network available in many cities. Another example of a MAN is IEEE 802.16. MANs are faster than long-distance WANs, but slower than LANs.

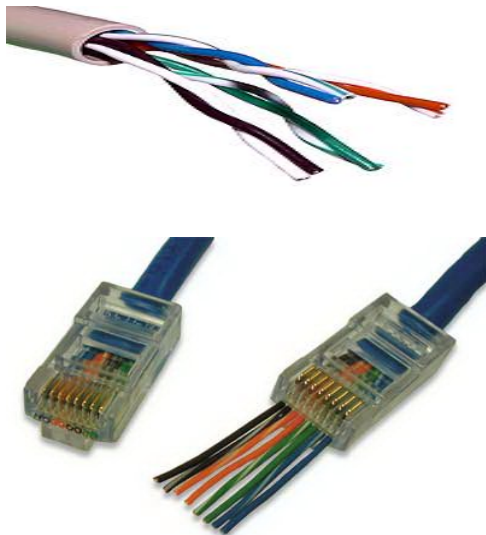
- **Wide Area networks (WANs)** : A Wide Area Network, or WAN, spans a large geographical area, often a country or a continent. It contains a collection of machines intended for running user programs. These machines are called hosts. The hosts in a WAN are connected by a **communication subnet**, or just **subnet** for short. The hosts are owned by the customers, whereas the **communication subnet** is typically owned and operated by a telephone company or Internet service provider. The job of the **subnet** is to carry messages from host to host, just as the telephone system carries words from a speaker to a listener. In most WANs, the **subnet** consists of two distinct components : **transmission lines** and **switching elements**. Transmission lines move bits between machines. Transmission lines can be made of copper wire, optical fiber, or even radio links. **Switching elements** are specialized computers that connect three or more transmission lines. When data arrive on an incoming line, the switching element must choose an outgoing line on which to forward them. These switching computers have been called by various names in the past; the name **router** is now most commonly used.
- **Wireless Networks** : Digital wireless communication is very popular. Wireless networks can be divided into three main categories :
 - * System interconnection
 - * Wireless LANs
 - * Wireless WANs
- System interconnection is all about interconnecting the components of a computer using short-range radio. A short-range wireless network called **Bluetooth** connects various components without wires. Bluetooth also allows digital cameras, headsets, scanners, and other devices to connect to a computer by merely being brought within range. No cables, no driver installation, just put them down, turn them on, and they work.
- **The Internet** : The Internet is a global network of networks.
- **Classification by transmission technology**
 - * **Broadcast networks**
 - have a single communication channel shared by all the machines on the network.
 - short messages (packets) sent by any machine are received by all the others.
 - usage of an address field in the packet : process/ignore the packet.
 - broadcasting : packets received and processed by every machine on the network.
 - * **Point-to-point networks**
 - consist of many connections between individual pairs of machines.
 - Packet journey : may involve visiting one / more intermediate nodes.
 - Possibility of multiple routes : role of routing algorithms

Transmission Media

- Guided Transmission Media
 - Twisted Pair Wires
 - Baseband Coaxial Cable
 - Broadband Coaxial Cable
 - Fiber Optics
- Unguided Transmission Media
 - Wireless transmission through various media

Twisted Pair

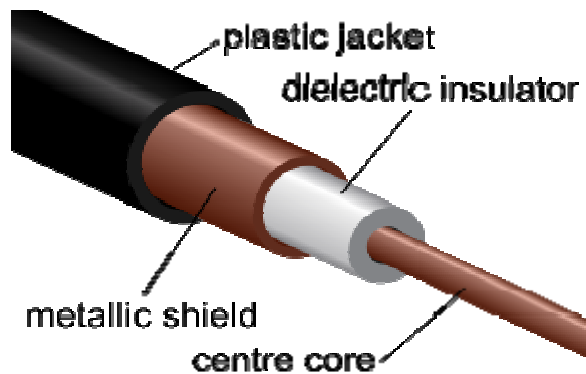
- Consists of two insulated copper wires, typically about 1 mm thick.
- Low cost, adequate performance, widely used, oldest, still most common, supports data transmission speed of several Mbps
- Wires : twisted in a helical form, just like a DNA molecule.
- Twisting : less crosstalk
 - better quality signal over longer distances
- Used in telephone system
- Can run several kms without amplification, but for longer distances, repeaters are used.
- Can be used for transmitting either analog or digital signals.
- Bandwidth = f(thickness of wire, distance traveled)



RJ45 Connectors for Twisted Pair Wires

Coaxial Cable

- Also known as “coax”, better shielding than twisted pairs
- Consists of a stiff copper wire as the core, surrounded by an insulating material. The insulator is encased by a cylindrical conductor, often as a closely-woven braided mesh. The outer conductor is covered with a protective plastic sheath.
- Good combination of high bandwidth, excellent noise immunity.
- Widely used for cable television and MANs.
- Can span longer distances at higher speeds.
- Two kinds of coax cable : (i) 50-ohm cable, commonly used for digital transmission, (ii) 75-ohm cable, used for analog transmission and cable television / Internet over cable.
- Bandwidth = f (cable quality, length, S/N), close to 1 GHz
- BNC connectors are used with coaxial cables.



Fiber Optics

- Can handle much higher bandwidth than copper wires
- Used in high-end networks.
- Low attenuation
 - Repeater are needed every 50 km on long lines versus about every 5 km for copper
- Not affected by power surges, electromagnetic interference, power failures
- Not affected by corrosive chemicals in the air :
Hence, ideal for harsh factory environments
- Optical Fiber : Thin and lightweight
- Fibers do not leak light
- Quite difficult to tap : excellent security against potential wire-tappers.
- Downside of fiber :
 - Less familiar technology requiring skills not all engineers have
 - Can be damaged easily by being bent too much
 - Since optical transmission is inherently unidirectional, two-way communication requires either two fibers or two frequency bands on one fiber.

- Achievable bandwidth with current technology : 50,000 Gbps (50 Tbps)
- Optical transmission system has 3 key components :
 1. Light source
 2. Transmission medium – ultra-thin fiber of glass
 3. Detector – generates an electrical pulse when light falls on it.
- Optical transmission system - unidirectional data transmission system.
- Accepts an *electrical signal*, converts and transmits it by *light pulses*, reconverts the output to an *electrical signal* at the receiving end.
- A *pulse of light* indicates a *binary 1*.
- An *absence of light* indicates a *binary 0*.
- Doesn't leak light
- Works on an interesting principle of physics :
- When a light ray passes from one medium to another (e.g. from fused silica to air), the ray is refracted (bent) at the boundary (e.g., air/silica).
- Light is trapped by *total internal reflection*.
- Amount of refraction – properties of the two media (indices of refraction).
- For angles of incidence above a certain critical value, the light is refracted back into the silica. None of it escapes into air.
- The light ray incident at or above the critical angle is trapped inside the fiber.
- Since any light ray incident on the boundary above the critical angle will be reflected internally, many different rays will be bouncing around at different angles. Each ray is said to have a different mode.
- Since any light ray incident on the boundary above the critical angle will be reflected internally, many different rays will be bouncing around at different angles. Each ray is said to have a different mode. A fiber having this property is called a multimode fiber.
- If the fiber's diameter is reduced to a few wavelengths of light, the fiber acts like a wave guide, and the light can propagate only in a straight line, without bouncing, yielding a single-mode fiber.
- Single-mode fiber – more expensive / widely used.

Wireless Transmission

- Radio Transmission
- Microwave Transmission
- Infrared Transmission
- Lightwave Transmission

Network Connecting Devices

- **Amplifier** : A device used to strengthen weak signals.
- **Repeater** : A repeater is an electronic device that receives a signal, regenerates and retransmits it. It receives the signal before it becomes too weak or corrupted. It regenerates the bits and forwards the refreshed signal. A repeater operates at the physical layer.
- **Bridge** : A data-link layer device that accepts a frame, processes it and forwards it to the next layer.
- Bridges are networking devices that connect networks. Sometimes it is necessary to divide networks into subnets to reduce the amount of traffic on each larger subnet or for security reasons. Once divided, the bridge connects the two subnets and manages the traffic flow between them. Today, network switches have largely replaced bridges.

A bridge functions by blocking or forwarding data, based on the destination MAC address written into each frame of data. If the bridge believes the destination address is on a network other than that from which the data was received, it can forward the data to the other networks to which it is connected.

- If the address is not on the other side of the bridge, the data is blocked from passing. Bridges “learn” the MAC addresses of devices on connected networks by “listening” to network traffic and recording the network from which the traffic originates.

The advantages of bridges are simple and significant. By preventing unnecessary traffic from crossing onto other network segments, a bridge can dramatically reduce the amount of network traffic on a segment. Bridges also make it possible to isolate a busy network from a not-so-busy one, thereby preventing pollution from busy nodes.

- **Router** : A network layer device that determines the route to be followed by packets.
- **Gateway** : A device that performs protocol conversion.
- **Modem** : Modem stands for MODulator/DEModulator. A modem converts digital signals generated by the computer into analog signals which can be transmitted over a telephone or cable line and transforms incoming analog signals into their digital equivalents.

Data Transmission Concepts

Transmission Modes :

1. Simplex Mode :

In Simplex mode, the communication is unidirectional, as on a one-way street. Only one of the two devices on a link can transmit, the other can only receive. The simplex mode can use the entire capacity of the channel to send data in one direction.

Example: Keyboard and traditional monitors. The keyboard can only introduce input, the monitor can only give the output.

2. Half-Duplex Mode :

In half-duplex mode, each station can both transmit and receive, but not at the same time. When one device is sending, the other can only receive, and vice versa. The half-duplex mode is used in cases where there is no need for communication in both directions at the same time. The entire capacity of the channel can be utilized for each direction.

Example: Walkie-talkie in which message is sent one at a time and messages are sent in both directions.

3. Full-Duplex Mode :

In full-duplex mode, both stations can transmit and receive simultaneously. In full_duplex mode, signals going in one direction share the capacity of the link with signals going in another direction, this sharing can occur in two ways:

- Either the link must contain two physically separate transmission paths, one for sending and the other for receiving.
- Or the capacity is divided between signals travelling in both directions.

Full-duplex mode is used when communication in both directions is required all the time. The capacity of the channel, however, must be divided between the two directions.

Example: Telephone Network in which there is communication between two persons by a telephone line, through which both can talk and listen at the same time.

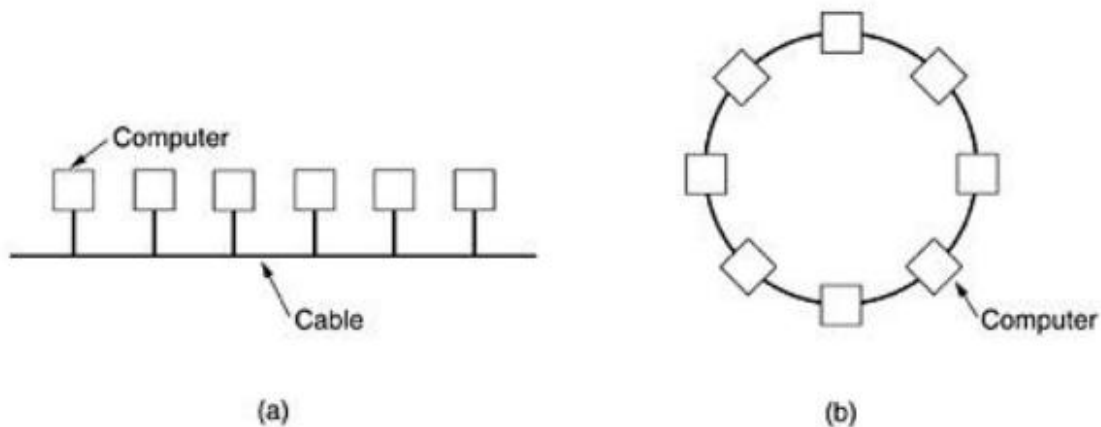
Multiplexing

- A multiplexer has number of I/O lines on one side and a high-speed multiplexed line on the other side.
- FDM : The frequency spectrum is divided among the logical channels, with each user having exclusive possession of his frequency band.
- TDM : In TDM, the users take turns in a round robin fashion, each one periodically getting the entire bandwidth for a little burst of time.

Asynchronous vs Synchronous transmission - (Introduction) : Asynchronous transmission : character-oriented , uses start bit/ stop bits, overhead of extra bits. Synchronous transmission : less overhead, SYN per large-sized block of data.

Introduction to Local Area Networks (LANs)

- Local Area Networks (LANs) are privately-owned networks within a single building or a campus of up to a few kilometers in size.
- Widely used to connect personal computers and workstations in company offices and factories to share resources (e.g., printers) and exchange information.
- LANs are distinguished from other kinds of networks by three characteristics :
 1. their size,
 2. their transmission technology,
 3. their topology.
- LANs are restricted in size.
- Worst-case transmission time is bounded and known in advance. It simplified network management.
- LANs may use a transmission technology consisting of a cable to which all the machines are connected, like a telephone company party lines.
- Traditional LANs run at speeds of 100 Mbps to 1 Gbps.
Newer LANs operate at up to 10 Gbps.
- LANs have low delay (microseconds or nanoseconds) and make very few errors..



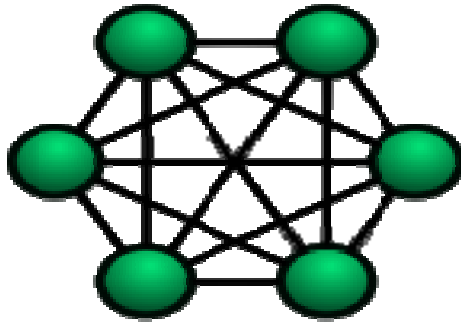
- Various topologies are possible for broadcast LANs.
- The bus and the ring topologies are shown in the figure.
- In a bus (i.e., a linear cable) network, at any instant at most one machine is the master and is allowed to transmit. All other machines are required to refrain from sending.

- An arbitration mechanism is needed to resolve conflicts when two or more machines want to transmit simultaneously. The arbitration mechanism may be centralized or distributed.
- IEEE 802.3, popularly called Ethernet, for example, is a bus-based broadcast network with decentralized control, usually operating at 10 Mbps to 10 Gbps.
- Computers on the Ethernet can transmit whenever they want to. If two or more packets collide, each computer just waits a random time and tries again later.
- A second type of broadcast system is the *ring*.
- In a *ring*, each bit propagates around on its own, not waiting for the rest of the packet to which it belongs.
- Typically, each bit circumnavigates the entire ring in the time it takes to transmit a few bits, often before the complete packet has even been transmitted.
- Some rule is needed for arbitrating simultaneous accesses to the ring.
- Various methods, such as having the machines take turns, are in use.
- IEEE 802.5 (the IBM token ring) is a ring-based LAN operating at 4 and 16 Mbps.
- FDDI is another example of a ring network.
- Broadcast networks can be further divided into static and dynamic, depending on how the channel is allocated. A typical static allocation would be to divide time into discrete intervals and use a round-robin algorithm, allowing each machine to broadcast only when its time slot comes up. Static allocation wastes channel capacity when a machine has nothing to say during its allocated slot, so most systems attempt to allocate the channel dynamically.

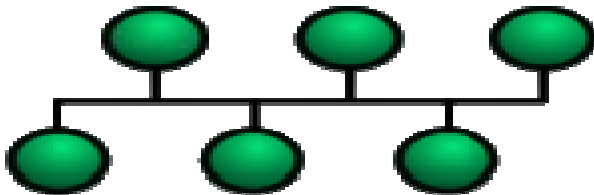
LAN Topologies

- **Network topology** : The specific physical or logical arrangement of the elements of a network.
- **Bus topology** : A network topology in which all nodes, *i.e.*, stations, are connected together by a linear cable.
- **Ring topology** : A network topology in which every node has exactly two branches connected to it.
- **Star topology** : A network topology in which peripheral nodes are connected to a central node.
- **Fully connected topology** : A network topology in which there is a direct path (branch) between any two nodes. *Note*: In a fully connected network with n nodes, there are $n(n-1)/2$ direct paths, *i.e.*, branches. *Synonym* **fully connected mesh network**.
- **Tree topology** : A network topology that, from a purely topologic viewpoint, resembles an interconnection of star networks in that individual peripheral nodes are required to transmit to and receive from one other node only, toward a central node, and are not required to act as repeaters or regenerators.
- **Hybrid topology** : A combination of any two or more network topologies.

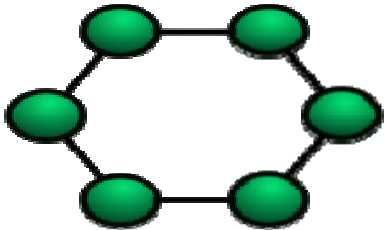
The Fully Connected Topology



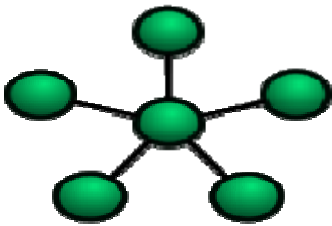
The Bus Topology



The Ring Topology



The Star Topology



Gigabit Ethernet

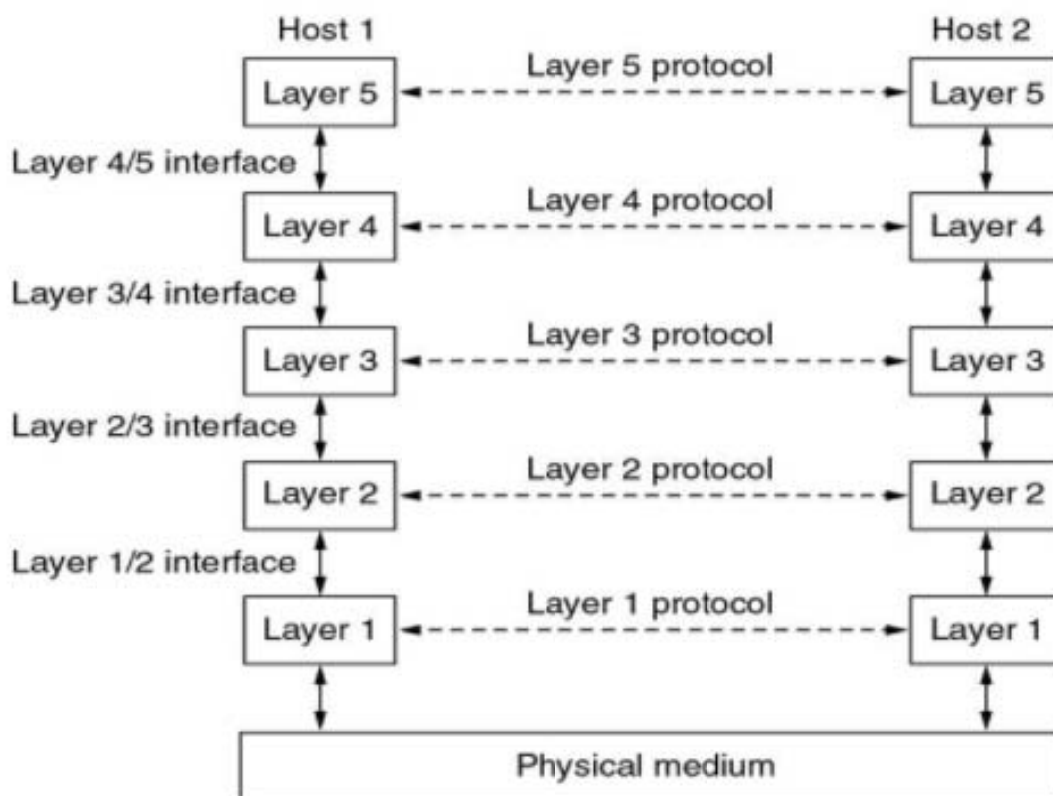
- Ratified by IEEE in 1998 under the name IEEE 802.3 z
- Supports data transmission speed of 1 Gbps.
- Backward compatible with all existing Ethernet standards.
- Supports two different modes of operation : full-duplex mode and half-duplex mode. The full-duplex mode allows traffic in both directions at the same time.
- All configurations of gigabit Ethernet are point-to-point rather than multidrop. In the simplest gigabit Ethernet configuration, two computers are directly connected to each other.

Name	Cable	Max. Segment	Advantages/Features
1000 Base-SX	Fiber optics	550 m	Multimode fiber (50, 62.5 microns) Uses 8B/10B encoding scheme
1000 Base-LX	Fiber optics	5000 m	Single (10 μ) or multimode (50, 62.5 μ) Uses 8B/10B encoding scheme 1.3 μ lasers operating over 10 μ fiber Best choice for campus backbones Expected to be popular
1000 Base-CX	2 Pairs of STP	25 m	Uses Shielded twisted pair Competing with high-performance fiber f cheap UTP from below
1000 Base-T	4 Pairs of UTP	100 m	Standard category 5 UTP Supports clock speed of 125 MHz 4 twisted pairs x 2 bits per twisted pair = 8 data bits per clock cycle

Protocol Hierarchies

- In order to reduce design complexity, most networks are organized as a stack of layers or levels, each one built upon the one below it.
- The number of layers, the name of each layer, the contents of each layer, and the function of each layer differ from network to network.
- The purpose of each layer is to offer certain services to the higher layers, shielding those layers from the details of how the offered services are actually implemented.
- Each layer is a kind of virtual machine, offering certain services to the layer above it.

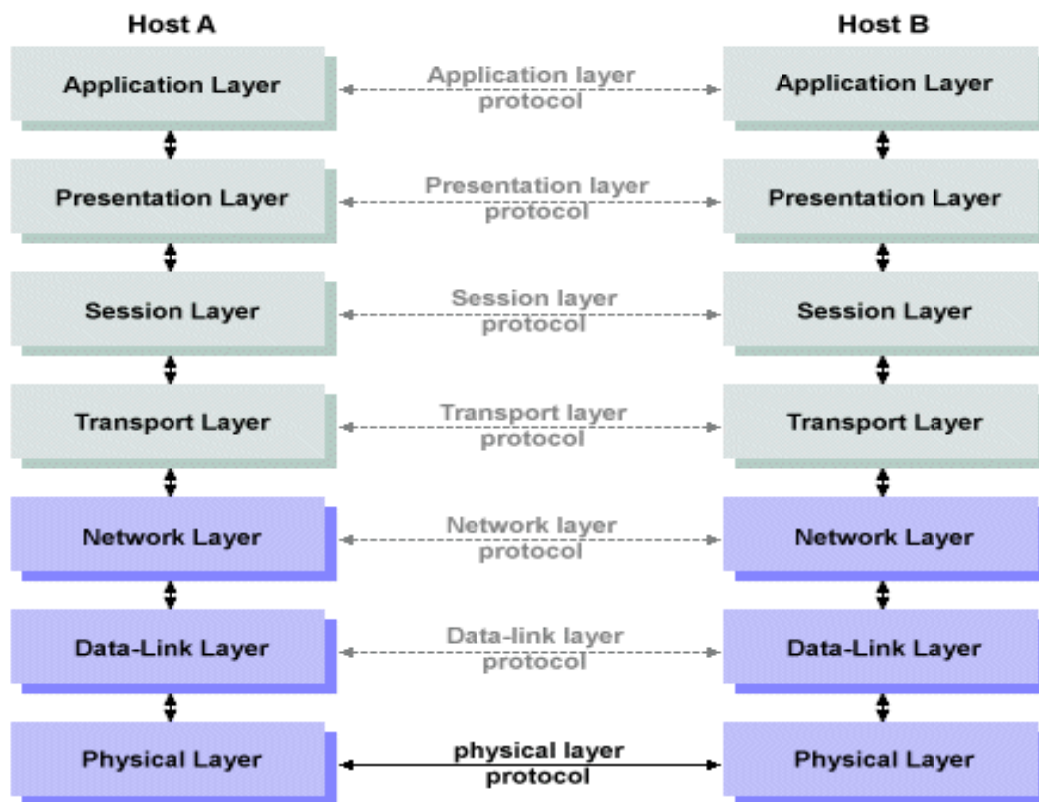
Layers, Protocols and Interfaces



- The fundamental idea is that a particular piece of software (or hardware) provides a service to its users but keeps the details of its internal state and algorithms hidden from them.
- Layer n on one machine carries on a conversation with layer n on another machine. The rules and conventions used in this conversation are collectively known as the layer n protocol.
- Basically, a protocol is an agreement between the communicating parties on how communication is to proceed. Violating a protocol makes communication more difficult.

- A five-layer network is illustrated in the figure.
- The entities comprising the corresponding layers on different machines are called **peers**. The peers may be processes, hardware devices, or even human beings. Peers communicate by using the protocol.

The OSI Reference Model



The OSI Reference Model

1. The Physical Layer

- The physical layer is concerned with transmitting raw bits over a communication channel.
- The design issues at this layer focus on making sure that when one side sends a 1 bit, it is received by the other side as a 1 bit, not as a 0 bit.
- The design issues here largely deal with mechanical, electrical, and timing interfaces, and the physical transmission medium, which lies below the physical layer.

Typical issues at the physical layer :

- How many volts should be used to represent a 1 and how many for a 0 ?

- How many nanoseconds a bit lasts ?
- Whether transmission may proceed simultaneously in both directions ?
- How the initial connection is established and how its is torn down when both sides have finished ?
- How many pins the network connector has and what each pin is used for ?

2. **The Data Link Layer**

- The main task of the data link layer is to transform a raw transmission facility into a line that appears free of undetected transmission errors to the network layer.
- It accomplishes this task by having the sender break up the input data into ***data frames*** (typically a few hundred or a few thousand bytes) and transmit the frames sequentially.
- If the service is reliable, the receiver confirms correct receipt of each frame by sending back an acknowledgement frame.
- Another issue that arises in the data link layer (and most of the higher layers as well) is how to keep a fast transmitter from drowning a slow receiver in data.
- Some traffic regulation mechanism is often needed to let the transmitter know how much buffer space the receiver has at the moment. Frequently, this flow regulation and error handling are integrated.
- Broadband networks have additional issue in the data link layer : how to control access to the shared channel. A special sublayer of the data link layer, the medium access control sublayer deals with this problem.

3. **The Network Layer**

- The network layer controls the operation of the subnet.
- A key design issue at the network layer is **determining how packets are routed from source to destination.**
- **Routes** can be based on static tables that are “wired into” the network and rarely changed. Routes can also be determined at the start of each conversation , for example, a terminal session (e.g., a login to a remote machine). Finally, routes can be highly dynamic, being determined newly for each packet, to reflect the current network load.
- **Congestion control :** If too many packets are present in the subnet at the same time, they will get in one another’s way, forming bottlenecks. The control of such congestion also belongs to the network layer.
- More generally, the quality of service provided (delay, transit time, jitter, etc.) is also a network layer issue.
- **Issues related to internetworking :** When a packet has to travel from one network to another to get to its destination, many problems can arise. The addressing used by the second network may be different from the first one. The second network may not accept the packet at all because it is too large. The protocols may differ, and so on. It is up to the network layer to overcome all these problems to allow heterogeneous networks to be interconnected.

- In broadcast networks, the routing problem is simple, so the network layer is often thin or even non-existent.

4. The Transport Layer

Basic function of the transport layer :

- To accept data from the session layer, split it up into smaller units if required, and pass these to the network layer.
- Ensure that the pieces all arrive correctly at the other end.
- Layer 4 and above are referred to as higher layers.
- Protocols at these levels are **end-to-end** (not concerned with the details of the underlying communication facility).

Purpose of layer 4 :

- To provide a **reliable mechanism** for **exchange of data** between processes in different systems.
- **To ensure that data units are delivered error-free, in sequence, with no losses/duplications.**
- To optimize the use of network services.
- **To provide a requested quality of service to session entities.**

(e.g. session entity might specify acceptable error rates, maximum delay, priority, security, etc.)

- If layer 3 is reliable and/or only supports datagrams, the layer 4 protocol should include extensive error detection and recovery.
- The size and complexity of the transport protocol depends on the type of service received from the layer 3.
- Example of protocol designed for the layer 4 : **TCP**.
- Other possible kinds of transport services are broadcasting of messages to multiple destinations, and transporting isolated messages.
- The transport layer is a true **end-to-end layer**, all the way from the source to the destination. In other words, a program on the source machine carries on a conversation with a similar program on the destination machine, using the message headers and control messages.
- In the lower layers, the protocols are between each machine and its immediate neighbors, and not between the ultimate source and destination machines, which may be separated by many routers.

5. The Session Layer

- The session layer allows users on different machines to **establish sessions** between them.
- Sessions offer various services, including
 - * **dialog control** (keeping track of whose turn it is to transmit)
 - * **token management** (preventing two parties from attempting the same critical operation at the same time)
 - * **synchronization** (checkpointing long transmissions to allow them to continue from where they were after a crash).

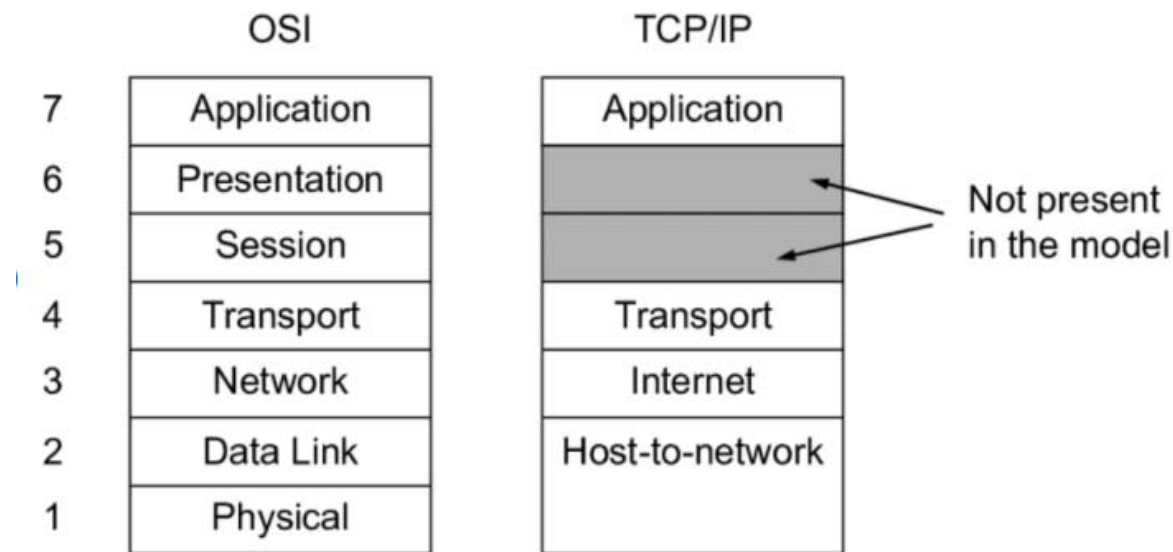
6. The Presentation Layer

- Unlike lower layers, which are mostly concerned with moving bits around, the presentation layer is concerned with the syntax and semantics of the information transmitted.
- In order to make it possible for computers with different data representations to communicate, the data structures to be exchanged can be defined in an abstract way, , along with a standard encoding to be used “on the wire”.
- The presentation layer manages these abstract data structures and allows higher-level data structures (e.g., banking records), to be defined and exchanged.

7. The Application Layer

- The application layer *contains a variety of protocols* that are commonly needed by users.
- One **widely-used application protocol** is **HTTP** (HyperText Transfer Protocol), which is the basis for the World Wide Web.
- When a browser wants a Web page, it sends the name of the page it wants to the server using HTTP.
- The server then sends the page back.
- Other application protocols are used for file transfer, electronic mail, and network news.

The TCP/IP Reference Model



Different Layers of TCP/IP Reference Model

Layer 1: Host-to-network Layer

Layer 2: Internet layer

Layer 3: Transport Layer

Layer 4: Application Layer

The Host-to-network Layer

- Protocol is used to connect to the host, so that the packets can be sent over it.
- Varies from host to host and network to network.

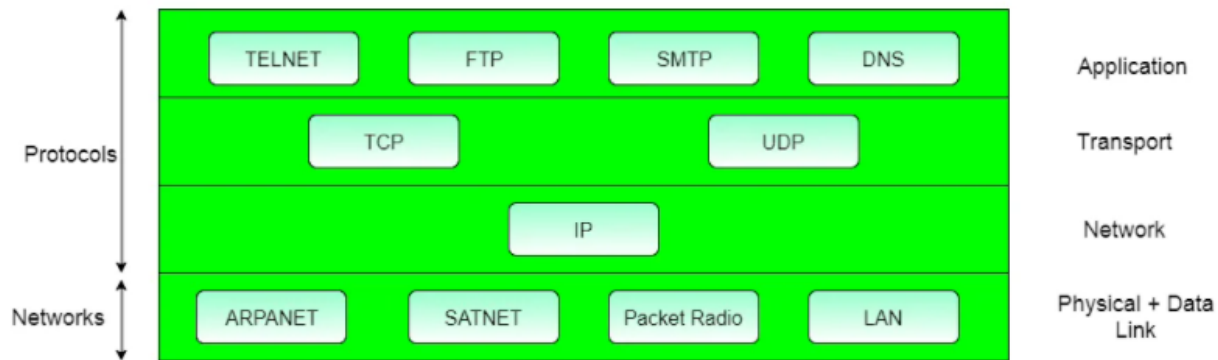
The Internet Layer

- The TCP/IP internet layer is similar in functionality to the OSI network layer.
- Major issues handled : packet routing and avoiding congestion.
- The job of the internet layer is to deliver IP packets where they are supposed to go.
- Its job is to permit hosts to inject packets into any network and have them travel independently to the destination (potentially on a different network). If the packets arrive in a different order than they were sent, then the higher layers have to rearrange them, if in-order delivery is desired.
- The Internet layer defines an official packet format and protocol called IP (Internet Protocol).

The Transport Layer

- The layer above the internet layer in the TCP/IP model is called the transport layer.
- It is designed to allow peer entities on the source and destination hosts to carry on a conversation, just as in the OSI transport layer.
- Two end-to-end transport protocols have been defined here : TCP (Transmission Control Protocol) is a reliable connection-oriented protocol that allows a byte stream originating on one machine to be delivered without error on any other machine in the internet.
- It fragments the incoming byte stream into discrete messages and passes each one on to the internet layer. At the destination, the receiving TCP process reassembles the received messages.
- TCP also handles flow control to make sure a fast sender cannot swamp a slow receiver with more messages than it can handle.
- The second protocol in this layer, UDP (User Datagram Protocol), is an unreliable, connectionless protocol for applications that do not want TCP's sequencing or flow control and wish to provide their own.
- UDP is widely used for one-shot, client-server type request-reply queries and applications in which prompt delivery is more important than accurate delivery, such as transmitting speech or video.

Protocols and networks in the TCP/IP model:



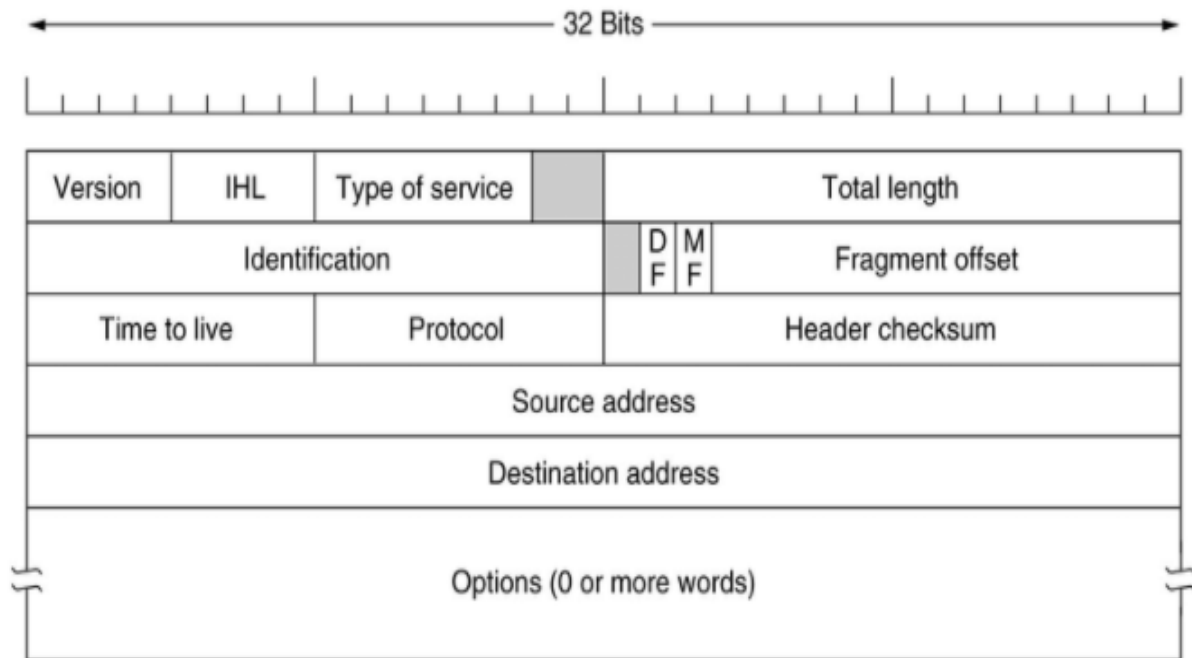
The Application Layer

- The TCP/IP model does not have session or presentation layers.
- The application layer contains all the higher-level protocols.
- Some of them are TELNET, FTP, SMTP, DNS, etc.
- **TELNET** is a two-way communication protocol which allows connecting to a remote machine and run applications on it.
- **FTP**(File Transfer Protocol) is a protocol, that allows File transfer amongst computer users connected over a network. It is reliable, simple and efficient.
- **SMTP**(Simple Mail Transport Protocol) is a protocol, which is used to transport electronic mail between a source and destination, directed via a route.
- **DNS**(Domain Name Server) resolves an IP address into a textual address for Hosts connected over a network.

The Internet Protocol (IP)

- The format of the IP datagrams : An IP datagram consists of a header part and a text part.
- The header has a 20-byte fixed part and a variable length optional part.
- It is transmitted in big-endian order : from left to right, with the high-order bit of the Version field going first.
- The Version field keeps track of which version of the protocol the datagram belongs to.
- Currently a transition between IPv4 and IPv6 is going on, has already taken years, and by no means close to being finished.
- IPv5 was an experimental real-time stream protocol that was never widely used.

The IPv4 (Internet Protocol) header



- Since the header length is not constant, a field in the header, IHL, is provided to tell how long the header is, in 32-bit words. The minimum value is 5, which applies when no options are present. The maximum value of this 4-bit field is 15, which limits the header to 60 bytes, and thus the Options field to 40 bytes. For some options, such as one that records the route a packet has taken, 40 bytes is too small, making that option useless.
- The *Type of service* field is intended to distinguish between different classes of service. Various combinations of reliability and speed are possible. For digitized voice, fast delivery beats accurate delivery. For file transfer, error-free transmission is more important than fast transmission.
- Originally, a 6-bit field contained (from left to right), a three-bit Precedence field and three flags, D, T, and R. The Precedence field was a priority, from 0 (normal) to 7 (network control packet). The three flag bits allowed the host to specify what it cared most about from the set {Delay, Throughput, Reliability}.
- In theory, these fields allow routers to make choices between, for example, a satellite link with high throughput and high delay or a leased line with low throughput and low delay. In practice, current routers often ignore the *Type of Service* field altogether.
- The *Total length* includes everything in the datagram – both header and data. The maximum length is 65,535 bytes.
- The ***Identification*** field is needed to allow the destination host to determine which datagram a newly arrived fragment belongs to. All the fragments of a datagram contain the same Identification value.

- Next comes an unused bit and then two 1-bit fields.
- DF stands for Don't Fragment. It is an order to the routers not to fragment the datagram because the destination is incapable of putting the pieces back together again. For example, when a computer boots, its ROM might ask for a memory image to be sent to it as a single datagram. By marking the datagram with the DF bit, the sender knows it will arrive in one piece, even if this means that the datagram must avoid a small-packet network on the best path and take a suboptimal route. All machines are required to accept fragments of 576 bytes or less.
- MF stands for More Fragments.
- All fragments except the last one have this bit set. It is needed to know when all fragments of a datagram have arrived.
- The Fragment offset field tells where in the current datagram this fragment belongs. All fragments except the last one in a datagram must be a multiple of 8 bytes, the elementary fragment unit. Since 13 bits are provided, there is a maximum of 8192 fragments per datagram, giving a maximum datagram length of 65,536 bytes, one more than the *Total length* field.
- The Time to live field is a counter used to limit packet lifetimes. It is supposed to count time in seconds, allowing maximum lifetime of 255 sec. It must be decremented on each hop and is supposed to be decremented multiple times when queued for a long time in a router.
- In practice, it just counts hops. When it hits zero, the packet is discarded and a warning packet is sent back to the source host. This feature prevents datagrams from wandering around forever.
- When the network layer has assembled a complete datagram, it needs to know what to do with it. The Protocol field tells it which transport process to give it to. TCP is one possibility, but so are UDP and some others. The numbering of protocols is global across the entire Internet. Protocols and other assigned numbers are contained in an online database located at www.iana.org.
- The Header checksum field verifies the header only. Such a checksum is useful for detecting errors generated by bad memory words inside a router.
- The algorithm is to add up all the 16-bit halfwords as they arrive, using one's complement arithmetic and then take the one's complement of the result.
- For purposes of this algorithm, the *Header checksum* is assumed to be zero upon arrival.
- The *Header checksum* must be recomputed at each hop because at least one field always changes (the *Time to live* field).
- The Source address and Destination address indicate the network number and host number.
- The Options field is provided to allow subsequent versions of the protocol to include information not present in the original design, to permit experimenters to try out new ideas, and to avoid allocating header bits to information that is rarely needed.
- The options are variable length. Each option begins with a 1-byte code identifying the option. Some options are followed by a 1-byte option length field, and then one or more data bytes.

- The **Options** field is padded out to a multiple of four bytes.
- Originally, five options were defined, but since then some new ones have been added.
- Current complete list of options is now maintained online at www.iana.org/assignments/ip-parameters.

Some of the IP options.

Option	Description
Security	Specifies how secret the datagram is
Strict source routing	Gives the complete path to be followed
Loose source routing	Gives a list of routers not to be missed
Record route	Makes each router append its IP address
Timestamp	Makes each router append its address and timestamp

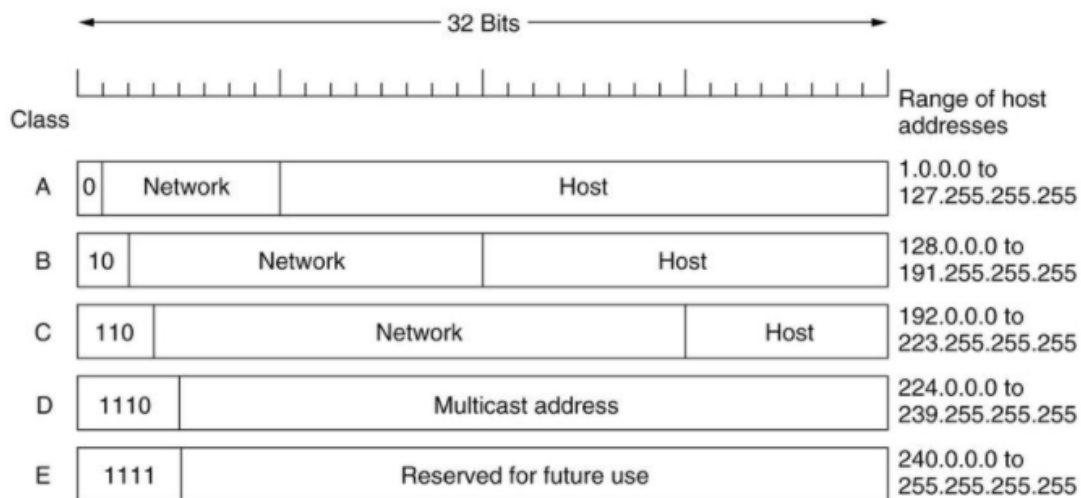
- The **Security** option tells how secret the information is. In theory, a military router might use this field to specify not to route through certain countries the military considers to be “bad guys”. In practice, all routers ignore it.
- The **Strict source routing** option gives the complete path from source to destination as a sequence of IP addresses. The datagram is required to follow that exact route. It is most useful for system managers to send emergency packets when the routing tables are corrupted, or for making timing measurements.
- The **Loose source routing** option requires the packet to traverse the list of routers specified, and in the order specified, but it is allowed to pass through other routers on the way. Allows a few routes to force a particular path.
- The **Record route** option tells the routers along the path to append their IP address to the option field. This allows system managers to track down bugs in the routing algorithms (“Why are packets from Ahmedabad to New Delhi visiting Mumbai first?”). When ARPANET was first set up, no packet ever passed through more than nine routers, so 40 bytes of option was ample. Now it is too small.
- The **Timestamp** option is like the *Record route* option, except that in addition to recording its 32-bit IP address, each router also records a 32-bit timestamp. This option is also used for debugging routing algorithms.

IP Addresses

- Every host and router on the Internet has an IP address, which encodes its network number and host number. The combination is unique.
- No two machines on the Internet have the same IP address.
- All IPv4 addresses are 32 bits long.
- IP addresses are used in the Source address and Destination address fields of IP packets.
- An IP address does not actually refer to a host. It really refers to a network interface, so if a host is on two networks, it must have two IP addresses.
- For several decades, IP addresses were divided into five categories. This allocation is called **classful addressing**. It is no longer used, but references to it in the literature are still common.

The class A, B, C, and D formats allow for up to

- 128 networks with 16 million hosts each,
- 16,384 networks with up to 64K hosts, and
- 2 million networks (e.g., LANs) with up to 256 hosts each.
- Also supported is multicast, in which a datagram is directed to multiple hosts.
- Addresses beginning with 1111 are reserved for future use
- Network numbers are managed by a non-profit corporation called **ICANN (Internet Corporation for Assigned Names and Numbers)** to avoid conflicts.



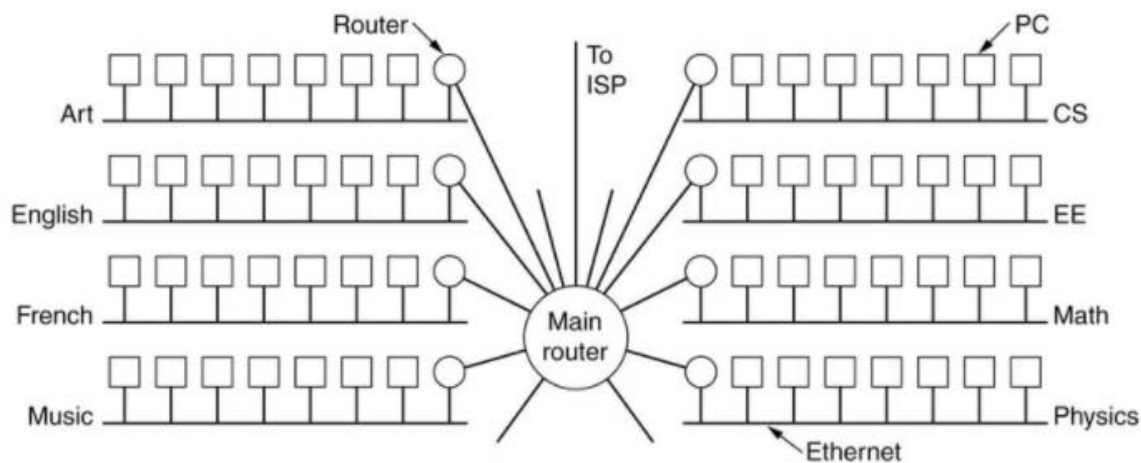
Special IP addresses.

0 0	This host
0 0 . . . 0 0 Host	A host on this network
1 1	Broadcast on the local network
Network 1 1 1 1 . . . 1 1 1 1	Broadcast on a distant network
127 (Anything)	Loopback

- Network addresses, which are 32-bit numbers, are usually written in dotted decimal notation. In this format, each of the 4 bytes is written in decimal, from 0 to 255.
- The lowest IP address is 0.0.0.0 and the highest is 255.255.255.255.
- The values 0 and -1 (all 1s) have special meanings.
- The value 0 means this network or this host.
- The value of -1 is used as a broadcast address to mean all hosts on the indicated network.
- The IP address 0.0.0.0 is used by hosts when they are being booted.
- IP addresses with 0 as network number refer to the current network. These addresses allow machines to refer to their own network without knowing its number (but they have to know its class to know how many 0s to include).
- The address consisting of all 1s allows broadcasting on the local network, typically a LAN.
- The addresses with a proper network number and all 1s in the host field allow machines to send broadcast packets to distant LANs anywhere in the Internet.
- All addresses of the form 127.xx.yy.zz are reserved for loopback testing. Packets sent to that address are not put out onto the wire; they are processed locally and treated as incoming packets. This allows packets to be sent to the local network without the sender knowing its number.

Subnets

- All the hosts in a network must have the same network number.
- This property of IP addressing can cause problems as networks grow.
- For example, consider a university that started out with one class B network used by the Computer Science Department for the computers on its Ethernet. A year later, the Mathematics Department wanted to get on the Internet, so they bought a repeater to extend the CS Ethernet to their building. As time went on, many other departments acquired computers and the limit of 4 repeaters per Ethernet was quickly reached. A different organization was required.
- Getting a second network address would be hard to do since network addresses are scarce and the university already had enough addresses for over 60,000 hosts.

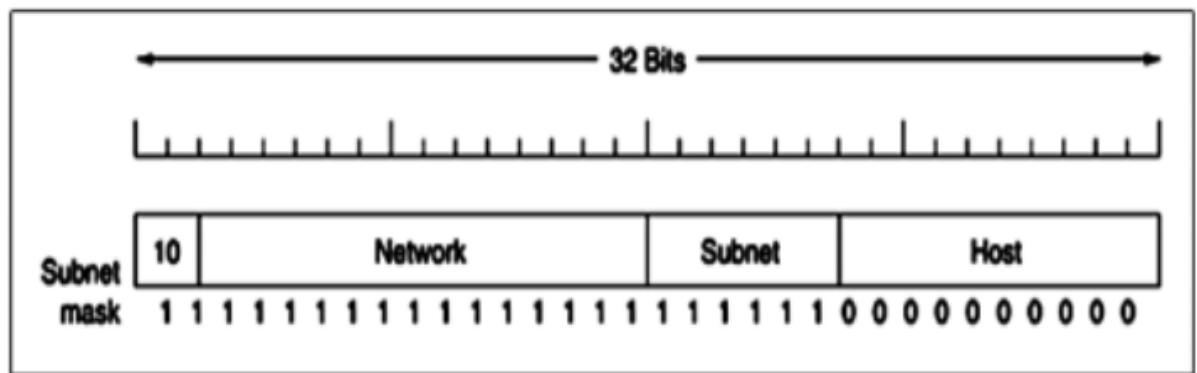


A campus network consisting of LANs for various departments.

-
- The solution is to allow a network to be split into several parts for internal use but still act like a single network to the outside world.
 - A typical campus network nowadays might look like the one shown in the figure, with a main router connected to an ISP or regional network and numerous Ethernets spread around campus in different departments. Each of the Ethernets has its own router connected to the main router.
 - In the Internet literature, the parts of a network are called subnets. For example, in our case, Ethernets. This usage conflicts “subnet” to mean the set of all routers and communication lines in a network. It will be clear from the context which meaning is intended.

- When a packet comes into the main router, how does it know which subnet (Ethernet) to give it to ?
- One possibility : to have a table with 65,536 entries in the main router telling which router to use for each host on campus. However, this idea would require a very large table in the main router and a lot of manual maintenance as hosts are added, moved, or taken out of service.
- Instead, a different scheme is invented.
- Basically, instead of having a single class B address with 14 bits for the network number and 16 bits for the host number, some bits are taken away from the host number to create a subnet number.
- For example, if the university has 35 departments, it could use a 6-bit subnet number and a 10-bit host number, allowing for up to 64 Ethernets, each with a maximum of 1022 hosts (0 and -1 are not available, as mentioned earlier). This split could be changed later if it turns out to be the wrong one.
- To implement subnetting, the main router needs a subnet mask that indicates the split between network + subnet number and host.
- Subnet masks are also written in dotted decimal notation, with the addition of a slash followed by the number of bits in the network + subnet part. For the example of the given Figure, the subnet mask can be written as 255.255.252.0.
- An alternative notation is /22 to indicate that the subnet mask is 22 bits long.

A class B network subnetted into 64 subnets



In this example, the first subnet might use IP addresses starting at 130.50.4.1.

The second subnet might start at 130.50.8.1.

The third subnet might start at 130.50.12.1, and so on.

Subnet 1 : 10000010 00110010 000001|00 00000001

Subnet 2 : 10000010 00110010 000010|00 00000001

Subnet 3 : 10000010 00110010 000011|00 00000001

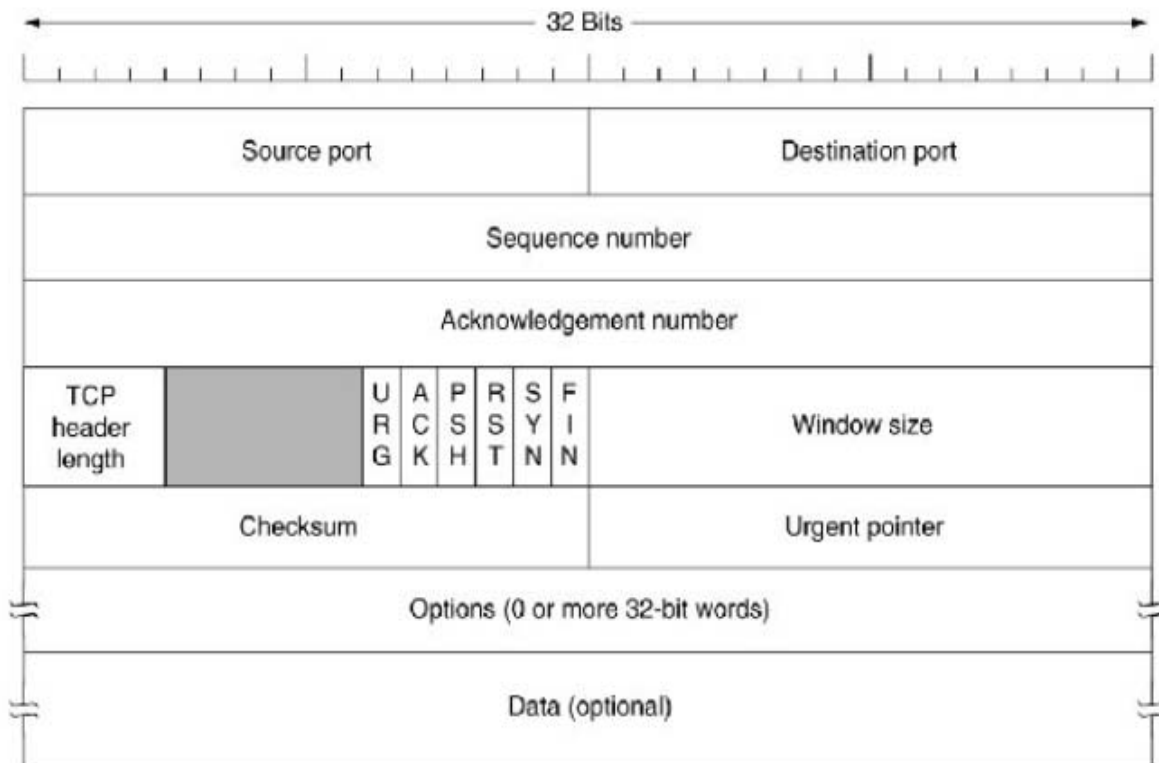
- Outside the network, subnetting is not visible, so allocating a new subnet does not require contacting ICANN or changing any external databases.
- To see how subnets work, it is necessary to explain how IP packets are processed at a router.

- Each router has a table listing some number of (network,0) IP addresses and some number of (this-network, host) IP addresses. The first kind tells how to get to distant networks. The second kind tells how to get to local hosts.
- Associated with each table is the network interface to use to reach the destination, and certain other information.
- When an IP packet arrives, its destination address is looked up in the routing table.
- If the packet is for a distant network, it is forwarded to the next router on the interface given in the table.
- If it is a local host (e.g., on the router's LAN), it is sent directly to the destination.
- If the network is not present, the packet is forwarded to a default router with more extensive tables.
- This algorithm means that each router only has to keep track of other networks and local hosts, not (network, host) pairs, greatly reducing the size of the routing table.
- When subnetting is introduced, the routing tables are changed, adding entries of the form (this-network, subnet, 0) and (this-network, this-subnet, host).
- Thus, a router on subnet k knows how to get to all the other subnets and also how to get to all the hosts on subnet k.
- It does not have to know the details about hosts on other subnets.
- In fact, all that needs to be changed is to have each router do a Boolean AND with the network's subnet mask to get rid of the host number and look up the resulting address in its tables (after determining which network class it is).
- For example, a packet addressed to 130.50.15.6 and arriving at the main router is ANDed with the subnet mask 255.255.252.0/22 to give the address 130.50.12.0.
- This address is looked up in the routing tables to find out which output line to use to get to the router for subnet 3.
- Subnetting thus reduces router table space by creating a three-level hierarchy consisting of network, subnet, and host.

TCP

- Every byte on a TCP connection has its own 32-bit sequence number.
- Separate 32-bit sequence numbers are used for acknowledgement and for the window mechanism.
- The sending and receiving TCP entities exchange data in the form of segments.
- A TCP segment consists of a fixed 20-byte header (plus an optional part) followed by zero or more data bytes.
- The TCP software decides how big segments should be. It can accumulate data from several writes into one segment or can split data from one write over multiple segments.
- Two limits restrict the segment size.
- First, each segment, including the TCP header, must fit in the 65,515-byte IP payload.
- Second, each network has a maximum transfer unit (MTU) and each segment must fit in the MTU.
- In practice, the MTU is generally 1500 bytes (the Ethernet payload size) and thus defines the upper bound on segment size.
- The basic protocol used by TCP entities is the sliding window protocol.
- When a sender transmits a segment, it also starts a timer.
- When the segment arrives at the destination, the receiving TCP entity sends back a segment (with/without data) bearing an acknowledgement number equal to the next sequence number it expects to receive.
- If the sender's timer goes off before the acknowledgement is received, the sender transmits the segment again.
- Segments can arrive out of order, so bytes 3072-4095 can arrive but cannot be acknowledged because bytes 2048-3071 have not turned up yet.
- Segments can also be delayed so long in transit that the sender times out and retransmits them.
- The retransmissions may include different byte ranges than the original transmission, requiring a careful administration to keep track of which bytes have been correctly received so far.
- However, since each byte in the stream has its own unique offset, it can be done.

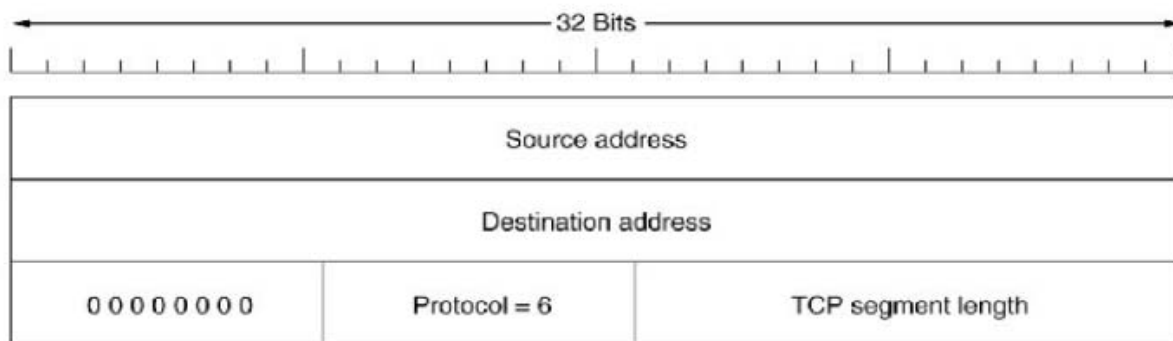
The TCP Segment Header



- Figure shows the layout of a TCP segment.
- Every segment begins with a fixed-format, 20-byte header.
- The fixed header may be followed by header options.
- After the options, if any, up to $65,535 - 20 - 20 = 65,495$ data bytes may follow, where the first 20 refer to the IP header and the second to the TCP header.
- Segments without any data are legal and are commonly used for acknowledgements and control messages.
- Let us dissect the TCP header field by field.
- The **Source port** and **Destination port** fields identify the local end points of the connection.
- The well-known ports are defined at www.iana.org but each host can allocate the others as it wishes.
- A port plus its host's IP address forms a 48-bit unique end point.
- The source and destination end points together identify the connection.
- The **Sequence number** and **Acknowledgement number** fields perform their usual functions. Both the fields are 32-bit long because every byte of data is numbered in a TCP stream.

- The **Acknowledgement number** field specifies the next byte expected, not the last byte correctly received.
- The **TCP header length** field tells how many 32-bit words are contained in the TCP header. This information is needed because the Options field is of variable length, so the header is, too. Technically, this field really indicates the start of the data within the segment, measured in 32-bit words, but that number is just the header length in words, so the effect is the same.
- Next comes a **6-bit field** that is **not used**.
- Next, there are **six 1-bit flags**.
- **URG** is set to 1 if the Urgent pointer is in use. The Urgent pointer is used to indicate a byte offset from the current sequence number at which urgent data are to be found. This facility is in lieu of interrupt messages.
- The **ACK** bit is set to 1 to indicate that the Acknowledgement number is valid. If ACK=0, the segment does not contain an acknowledgement so the Acknowledgement number field is ignored.
- The **PSH** bit indicates PUSHed data. The receiver is hereby kindly requested to deliver the data to the application upon arrival and not buffer it until a full buffer has been received (which it might otherwise do for efficiency).
- The **RST** bit is used to reset a connection that has become confused due to a host crash or some other reason. It is also used to reject an invalid segment or refuse an attempt to open a connection. In general, if you get a segment with the RST bit on, you have a problem on your hands.
- The **SYN** bit is used to establish connections.
- The connection request has SYN=1 and ACK=0 to indicate that the piggyback acknowledgement field is not in use.
- The connection reply does bear an acknowledgement, so it has SYN=1 and ACK=1.
- In essence the SYN bit is used to denote CONNECTION REQUEST and CONNECTION ACCEPTED, with the ACK bit used to distinguish between those two possibilities.
- The **FIN** bit is used to release a connection. It specifies that the sender has no more data to transmit.
- Flow control in TCP is handled using a variable-sized sliding window.
- The **Window size** field tells how many bytes may be sent starting at the byte acknowledged.
- A Window size field of 0 is legal and says that the bytes up to and including Acknowledgement number – 1 have been received, but that the receiver is currently badly in need of a rest and would like no more data for the moment, thank you. The receiver can later grant permission to send by transmitting a segment with the same Acknowledgement number and a nonzero Window size field.
- In TCP, acknowledgements and permission to send additional data are completely decoupled. In effect, a receiver can say : I have received bytes up through k but I do not want any more just now. This decoupling gives additional flexibility.

- A **Checksum** is also provided for extra reliability. It checksums the header, the data, and the conceptual pseudoheader. When performing this computation, the TCP Checksum field is set to zero and the data field is padded out with an additional zero byte if its length is an odd number.
- The checksum algorithm is simply to add up all the 16-bit words in 1's complement and then to take the 1's complement of the sum.
- As a consequence, when the receiver performs the calculation on the entire segment, including the Checksum field, the result should be 0.
- The pseudoheader contains the 32-bit IP addresses of the source and destination machines, the protocol number for TCP (6), and the byte count for the TCP segment (including the header).
- Including the pseudoheader in the TCP checksum computation helps detect misdelivered packets. UDP uses the same pseudoheader for its checksum.



The pseudoheader included in the TCP checksum.

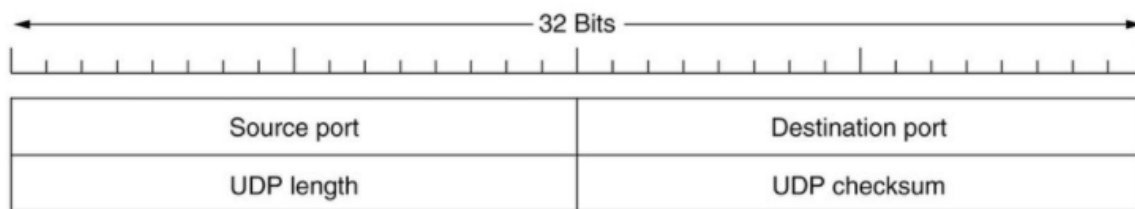
- The *Options* field provides a way to add extra facilities not covered by the regular header.
- The most important option is the one that allows each host to specify the maximum TCP payload it is willing to accept.
- Using large segments is more efficient than using small ones because the 20-byte header can then be amortized over more data, but small hosts may not be able to handle big segments.
- During connection setup, each side can announce its maximum and see its partner's/
- If a host does not use this option, it defaults to a 536-byte payload. All Internet hosts are required to accept TCP segments of $536 + 20 = 556$ bytes. The maximum segment size in the two directions need not be the same.

- Another option : Window scale option allowing a sender and a receiver to negotiate a window scale factor.

Introduction to UDP

- **UDP** stands for *User Datagram Protocol*.
- UDP is a *connectionless transport protocol*.
- UDP provides a way for applications to send encapsulated IP datagrams and send them without having to establish a connection.
- UDP transmits *segments* consisting of an *8-byte header followed by the payload*.
- The two ports serve to identify the end points within the source and destination machines.
- When a UDP packet arrives, its payload is handed to the process attached to the destination port.
- The *source port* is primarily need when a reply must be sent back to the source.

The UDP Header



- The *UDP length* field includes the 8-byte header and the data.
- The *UDP checksum* is optional and stored as 0 if not computed (a true computed 0 is stored as all 1s). Turning it off is foolish unless the quality of the data does not matter. (e.g., digitized speech).
- UDP does not do flow control, error control, or retransmission upon receipt of a bad segment. All of that is up to the user processes.
- What it does do is provide an interface to the IP protocol with the added feature of demultiplexing multiple processes using the ports.
- One area where *UDP* is especially *useful* is in *client-server situations*.
- Client-server situation in which UDP is useful :

Often, the client sends a short request to the server and expects a short reply back. If either the request or reply is lost, the client can just time out and try again. Not only is the code simple, but fewer messages are required (one in each direction) than with a protocol requiring an initial setup.

- Domain Name System (DNS) is an application that uses UDP this way. A program that needs to look up the IP address of some host name, for example, www.cs.berkeley.edu, can send a UDP packet containing the host name to a DNS server. The server replies with

a UDP packet containing the host's IP address. No setup is needed in advance and no release is needed afterward. Just two messages go over the network.

Routing Algorithms

- The **main function of the network layer** is **routing packets from the source machine to the destination machine.**
- A **routing algorithm** is that part of the network layer software responsible for deciding which output line an incoming packet should be transmitted on. If the subnet uses datagrams internally, this decision must be made anew for every arriving data packet since the best route may have changed since last time. If the subnet uses virtual circuits internally, routing decisions are made only when a new virtual circuit is being set up.
- Routing algorithms : generally grouped into two major classes : non-adaptive and adaptive.
- **Non-adaptive routing algorithms** do not base their routing decisions on measurements or estimates of the current traffic and topology. Instead, the choice of the **route** to use to get from I to J (for all I and J) is **computed in advance, off-line, and downloaded to the routers when the network is booted.** This procedure is sometimes called **static routing.**
- **Adaptive routing algorithms,** in contrast, change their routing decisions to reflect changes in the **topology**, and usually the **traffic** as well. Adaptive algorithms differ in where they get their information (e.g., locally, from adjacent routers, or from all routers), when they change the routes (e.g., every T sec, when the load changes or when the topology changes), and what metric is used for optimization (e.g., distance, number of hops, or estimated transit time).

Routing Algorithms

Some of the examples of various routing algorithms are :

- Shortest Path Routing
- Flooding
- Distance Vector Routing
- Link State Routing
- Hierarchical Routing
- Broadcast Routing
- Multicast Routing
- Routing for Mobile Hosts
- Routing in Ad Hoc Networks

Flooding

- Static routing algorithm
- Every incoming packet is sent out on every outgoing line except the one it arrived on.
- Flooding generates vast numbers of duplicate packets.
- Some measures must be taken to damp the process.

e.g. Have a hop counter contained in the header of each packet, which is decremented at each hop. The packet may be discarded when the counter reaches zero.

Ideally, the hop counter should be initialized to the length of the path from source to destination.

- Another technique for damming the flood :

Keep track of which packets have been flooded, to avoid sending them out a second time. The source router can put a sequence number in each packet it receives from its hosts. Each router then needs a list per source router telling which sequence numbers originating at that source have already been seen. If an incoming packet is on the list, it is not flooded. To prevent the list from growing without bound, each list should be augmented by a counter, k, meaning that all sequence numbers through k have been seen. If a packet is a duplicate, it is discarded. The full list below k is not needed, since k effectively summarizes it.

Selective Flooding

- In selective flooding algorithm, the routers do not send every incoming packet out on every line, but only on those lines that are going approximately in the right direction. There is usually little point in sending a westbound packet on an eastbound line.

Uses of Flooding

- In military applications, where large number of routers may be blown to bits at any instant, the tremendous robustness of flooding is highly desirable.
- Distributed database applications – to update all the databases concurrently.
- Wireless networks – All messages transmitted by a station can be received by all other stations within its radio range, which is in fact flooding.
- As a metric against which other routing algorithms can be compared

Link State Routing

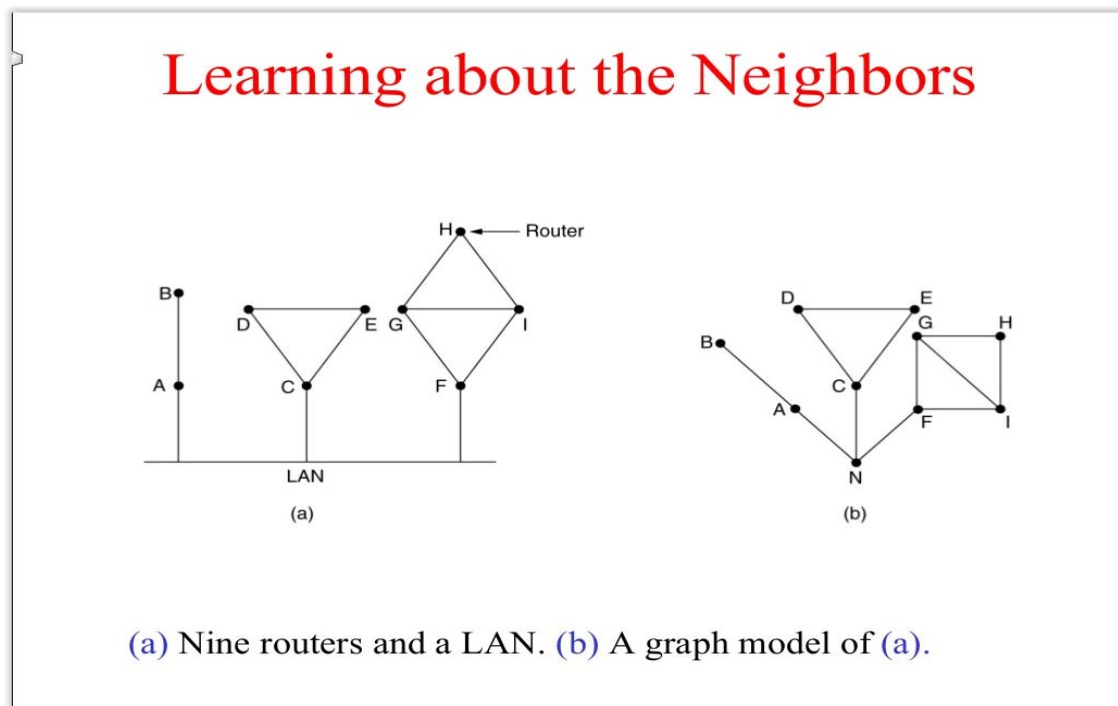
- Distance vector routing was used in the ARPANET until 1979, when it was replaced by link state routing.
- Two primary problems of distance vector routing.
 1. Line bandwidth was not considered while selecting the route.
 2. The algorithm often took too long to converge (the count-to-infinity problem).
- For these reasons, the distance vector routing algorithm was replaced by an entirely new algorithm, now called link state routing algorithm.

- The link state routing algorithm is a dynamic routing algorithm. (Adaptive routing algorithm).
- Each router must do the following :
 1. Discover its neighbors and learn their network addresses.
 2. Measure the delay or cost to each of its neighbors.
 3. Construct a packet telling all it has just learned.
 4. Send this packet to all other routers.
 5. Compute the shortest path to every other router.

Step 1. Each router must discover its neighbors and learn their network addresses.

Learning about the Neighbors

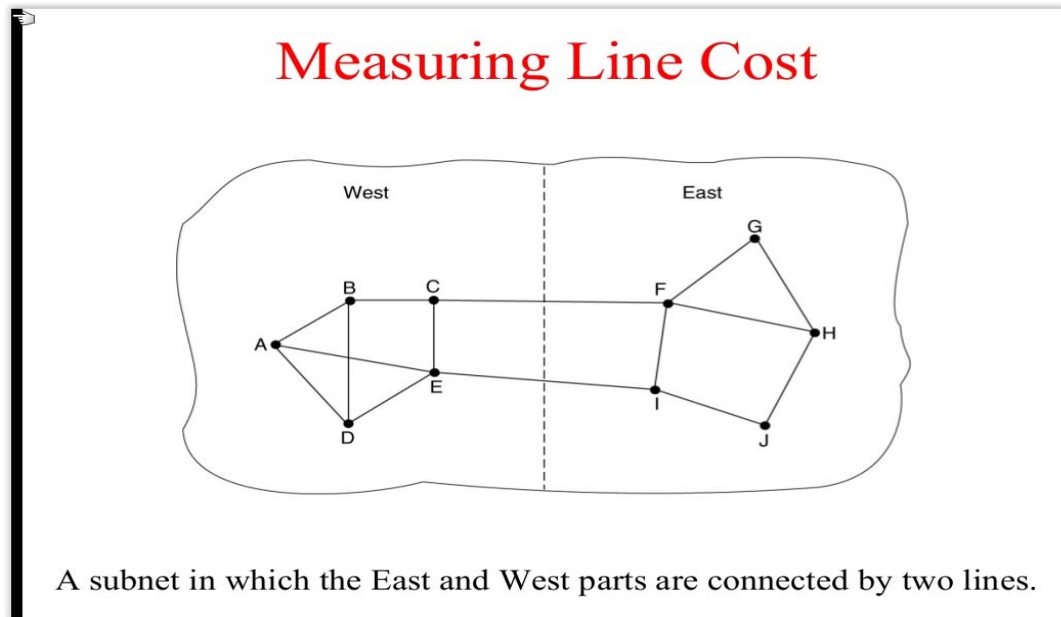
- The router sends a special HELLO packet on each point-to-point line.
- The router on the other end is expected to send back a reply telling who it is.
- When two or more routers are connected by a LAN, the situation is slightly more complicated.
- Consider the figure which illustrates a LAN to which three routers, A, C, and F, are directly connected. Each of these routers is connected to one or more additional routers.
- One way to model the LAN is to consider it as a node itself.
- We can introduce a new, artificial node N, to which A, C, and F are connected.



Step 2 Each router must measure the delay or cost to each of its neighbors.

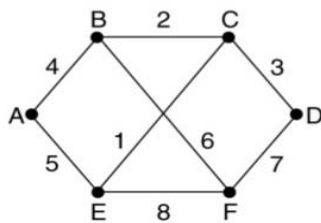
- Measuring Line Cost
- Each router must have a reasonable estimate of the delay to each of its neighbors.

- Send over the line a special ECHO packet that the other side is required to send back immediately.
- By measuring the round-trip time and dividing it by two, the sending router can get a reasonable estimate of the delay.
- For even better results, the test can be conducted several times, and the average can be used.
- Interesting issue : Whether to take the load into account when measuring the delay ?
- To factor the load in, the round-trip timer must be started when the ECHO packet is queued.
- To ignore the load, the timer should be started when the ECHO packet reaches the front of the queue.



Step 3. Construct a packet telling all it has just learned.
 This step requires constructing the Link State Packets.

Building Link State Packets



(a)

Link		State		Packets	
A	B	C	D	E	F
Seq.	Seq.	Seq.	Seq.	Seq.	Seq.
Age	Age	Age	Age	Age	Age
B 4	A 4	B 2	C 3	A 5	B 6
E 5	C 2	D 3	F 7	C 1	D 7
	F 6	E 1		F 8	E 8

(b)


(a) A subnet. (b) The link state packets for this subnet.

Step 4 Each router must send the link state packet

to all other routers.

- Distributing Link State Packets
- The basic distribution algorithm : The fundamental idea is to use flooding to distribute the link state packets.
- To keep the flood in check, each packet contains a sequence number that is incremented for each new packet sent.
- Routers keep track of all the (source router, sequence#) pairs they see.
- When a new link state packet comes in, it is checked against the list of packets already seen.
- If it is new, it is forwarded on all lines except the one it arrived on.
- If it is a duplicate packet, it is discarded.
- If a packet with a seq# lower than the highest one seen so far ever arrives, it is rejected as being obsolete since the router has more recent data.
- Issues :
- If the sequence numbers wrap around, confusion will reign.
- If a router ever crashes, it will lose track of its seq number.
- If a seq# is ever corrupted, and 65,540 is received instead of 4 (a 1-bit error) packets 5 to 65,540 will be rejected as obsolete.
- Solution : Use the age field.

Step 4. Distributing the Link State Packets : The packet buffer for router B in the previous figure :



Distributing the Link State Packets

Source	Seq.	Age	Send flags			ACK flags			Data
			A	C	F	A	C	F	
A	21	60	0	1	1	1	0	0	
F	21	60	1	1	0	0	0	1	
E	21	59	0	1	0	1	0	1	
C	20	60	1	0	1	0	1	0	
D	21	59	1	0	0	0	1	1	

Step 5. Compute the shortest path to every other router.