Department of Computer Science and Engineering
CS6004 Computer Networking
Subject Notes: UNIT-I

Computer Network: Definition

A **computer network** is a set of **computers** connected together for the purpose of sharing resources. The most common resource shared today is connection to the Internet. Other shared resources can include a printer or a file server.

#Goals

- Several machines can share printers, tape drives, etc.
- Reduced cost
- Resource and load sharing
- Programs do not need to run on a single machine
- High reliability
- If a machine goes down, another can take over
- Mail and communication

#Components

A data communications system has five components.

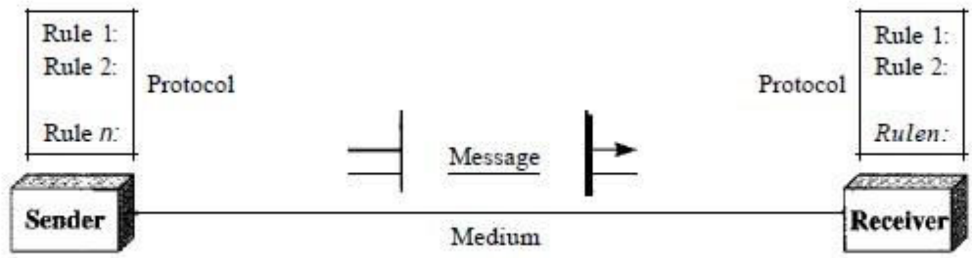

Fig. 1.1 Computer Network: Components

1. Message. The message is the information (data) to be communicated. Popular forms of information include text, numbers, pictures, audio, and video.

2. Sender. The sender is the device that sends the data message. It can be a computer, workstation, telephone handset, video camera, and so on.

3. Receiver. The receiver is the device that receives the message. It can be a computer, workstation, telephone handset, television, and so on.

4. Transmission medium. The transmission medium is the physical path by which a message travels from sender to receiver. Some examples of transmission media include twisted-pair wire, coaxial cable, fiber-optic cable, and radio waves

5. Protocol. A protocol is a set of rules that govern data communications. It represents an agreement between the communicating devices. Without a protocol, two devices may be connected but not communicating.

#Architecture

Network architecture is the design of a communications network. It is a framework for the specification of a network's physical components and their functional organization and configuration.

In telecommunication, the specification of a network architecture may also include a detailed description of products and services delivered via a communications network, as well as detailed rate and billing structures under which services are compensated. The network architecture of the Internet is predominantly expressed by its use of the Internet Protocol Suite, rather than a specific model for interconnecting networks or nodes in the network, or the usage of specific types of hardware link

#Computer Network's: Classifications & Types.

There are three types of network classification

1) LAN (Local area network)
2) MAN (Metropolitan Area network)

3) WAN (Wide area network)

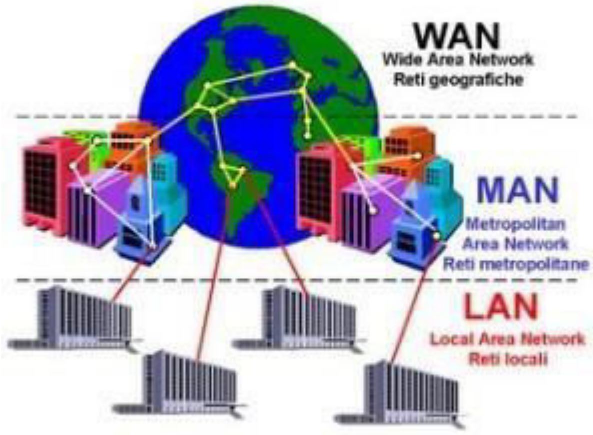

Fig. 1.2 Computer Network: Classifications

1) Local area network (LAN)

LAN is a group of the computers placed in the same room, same floor, or the same building so they are connected to each other to form a single network to share their resources such as disk drives, data, CPU, modem etc. LAN is limited to some geographical area less than 2 km. Most of LAN is used widely is an Ethernet system of the bus topology.

Characteristics of LAN

LAN connects the computer in a single building, block and they are working in any limited area less than 2 km.

Media access control methods in a LAN, the bus-based Ethernet and token ring.

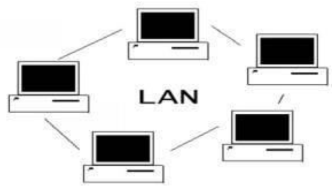

Fig. 1.3 Local area network

2) Metropolitan Area network (MAN)

The metropolitan area network is a large computer network that expands a Metropolitan area or campus. Its geographic area between a WAN and LAN. its expand round 50km devices used are modem and wire/cable.

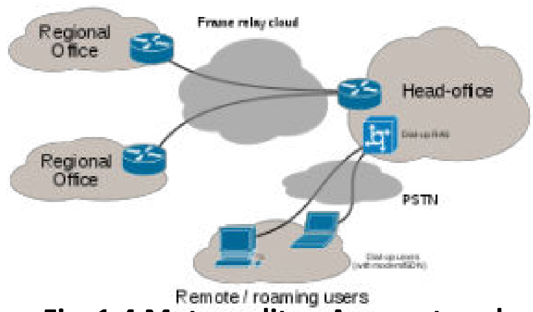

Fig. 1.4 Metropolitan Area network

Characteristics of MAN

1) Its covers the towns and cities (50km)

2) MAN is used by the communication medium for optical fibre cables, it also used for other media.

3) Wide area Network (WAN)

The wide area network is a network which connects the countries, cities or the continents, it is a public communications links. The most popular example of a WAN is the internet. WAN is used to connect LAN so the users and the computer in the one location can communicate with each other.

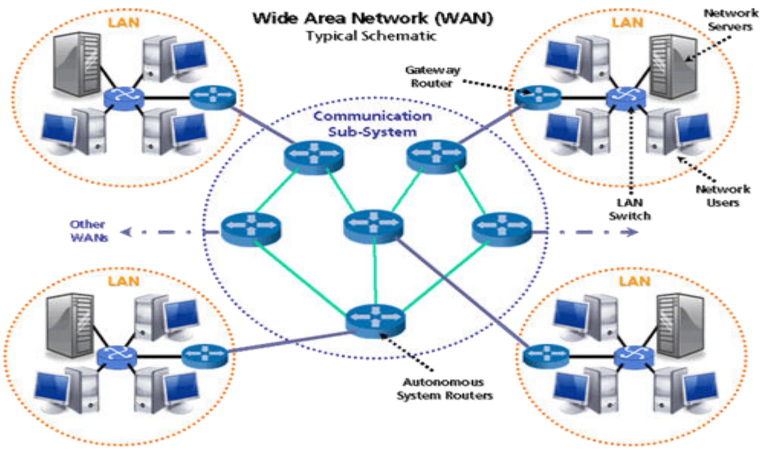

Fig. 1.5 Wide area Network

Characteristics of WAN

1) Its covers the large distances (More than 100 KM).
2) Communication medium used are satellite, telephones which are connected by the routers.

#Layered Architecture:

Protocol hierarchy: - To tackle with the design complexity most of the networks are organize as a set of layers or levels. The fundamental idea of layered architecture is to divide the design into small pieces. The layering provides modularity to the network design. The main duty of each layer is to provide offer services to higher layers, and provide abstraction. The main benefits of layered architecture are modularity and clear interfaces.

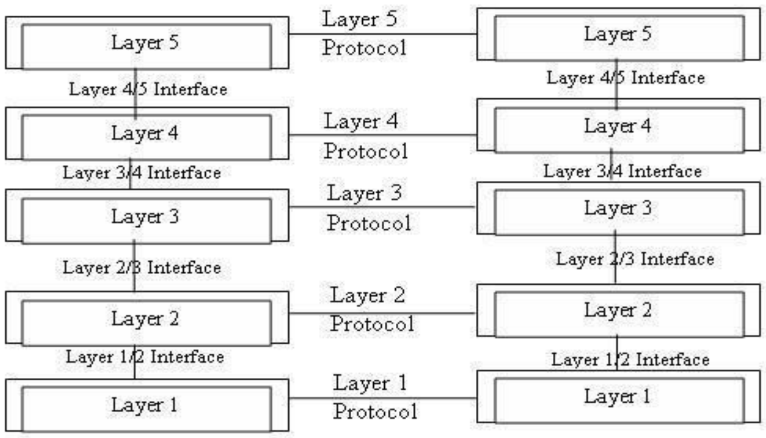

Fig. 1.6 Five Layered Network

#Design Issues:

Layered architecture in computer network design

Layered architectures have several advantages. Some of them are,

- Modularity and clear interface
- Provide flexibility to modify network services
- Ensure independence of layers
- Management of network architecture is easy
- Each layer can be analysed and tested independent of other layers

#Interfaces and Services:

The benefits to layering networking protocol specifications are many including:

Interoperability - Layering promotes greater interoperability between devices from different manufacturers and even between different generations of the same type of device from the same manufacturer.

Greater Compatibility - One of the greatest of all the benefits of using a hierarchal or layered approach to networking and communications protocols is the greater compatibility between devices, systems and networks that this delivers.

Better Flexibility - Layering and the greater compatibility that it delivers goes a long way to improving the flexibility. Particularly in terms of options and choices.
Increased Life Expectancy - Increased product working life expectancies as backwards compatibility is made considerably easier. Devices from different technology generations can co-exist thus the older units do not get discarded immediately newer technologies are adopted.
Scalability - Experience has shown that a layered or hierarchal approach to networking protocol design and implementation scales better than the horizontal approach.
Mobility - Greater mobility is more readily delivered whenever we adopt the layered and segmented strategies into our architectural design
Cost Effective Quality - The layered approach has proven time again and again to be the most economical way of developing and implementing any system be they are small, simple, large or complex makes no difference. This ease of development and implementation translates to greater efficiency and effectiveness which in turn translates into greater economic rationalization and cheaper products while not compromising quality.
Standardization and Certification - The layered approach to networking protocol specifications facilitates a more streamlined and simplified standardization and certification process; particularly from an "industry" point of view. This is due to the clearer and more distinct definition and demarcation of what functions occur at each layer when the layered approach is taken.
Rapid Application Development (RAD) - Workloads can be evenly distributed which means that multiple activities can be conducted in parallel thereby reducing the time taken to develop, debug, optimize and package new technologies ready for production implementation.

#Connection Oriented & Connectionless Services, Service primitives, Design issues & its functionality

• Connection-oriented

There is a sequence of operation to be followed by the users of connection-oriented service. They are:

1. Connection is established
2. Information is sent
3. Connection is released

In connection-oriented service we must establish a connection before starting the communication. When connection is established we send the message or the information. Then we release the connection.
Connection oriented service is more reliable than connectionless service. Example of connection oriented is TCP (Transmission Control Protocol) protocol.

• Connectionless

It is similar to postal services, as it carries the full address where the message (letter) is to be carried. Each message is routed independently from source to destination. The order of message sent can be different from the order received.
In connectionless the data is transferred in one direction from source to destination without checking that destination is still there or not or if it prepared to accept the message. Authentication is not needed in this. Example of Connectionless service is UDP (User Datagram Protocol) protocol.

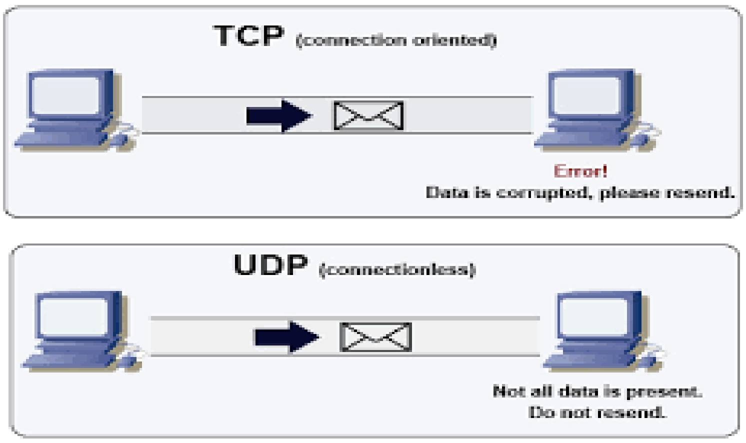

Fig. 1.7 Connection Oriented & Connectionless Services

#Service Primitives

Connection Oriented Service Primitives

LISTEN	Block waiting for an incoming connection
CONNECTION	Establish a connection with a waiting peer
RECEIVE	Block waiting for an incoming message
SEND	Sending a message to the peer
DISCONNECT	Terminate a connection

Connectionless Service Primitives

UNIDATA	This primitive sends a packet of data
FACILITY, REPORT	Primitive for enquiring about the performance of the network, like delivery statistics.

Design issues & its functionality

- **Justifying a Network**: - Some applications may be best satisfied by individual point to point connections to handle very specific communication requirements.
- **Scope**: - The scope of the network is viewed as bounded on one side by the offerings of the common carriers who provide communication facilities from which the network is built and on the other side by the application on which it is interconnected.
- Manageability:-
- **Network Architecture**: - While designing the network architecture, network may be a single homogeneous mesh comprised of a single type of node and a single type of link. Network architecture might be hierarchical network with one type link riding on another.
- **Switch Mode**: - For data transmission, different types of switching methods are possible. These are packet switching, circuit switching and hybrid switching.
- **Node Placement and sizing**: - A fundamental problem in the topological optimization of a network is the selection of the network node sites and where to place multiplexers, hubs and switch.
- **Link Topology and sizing**: - It involves selecting the specific links interconnecting nodes. At the highest level, that is where the architecture of the network is derived. Thus a hierarchy that include a backbone as well as LAN'S may be defined. It is possible to permit the backbone to be a mesh while LAN is constrained to be trees.
- **Routing**: - It involves selecting paths for each requirements. At higher level, this involves selecting the routing procedure itself.

#ISO-OSI Reference Model

#Principles of OSI Reference Model

The OSI reference model has 7 layers. The principles that were applied to arrive at the seven layers can be briefly summarized as follows:

1. A layer should be created where a different abstraction is needed.
2. Each layer should perform a well-defined function.
3. The function of each layer should be chosen with an eye toward defining internationally standardized protocols.
4. The layer boundaries should be chosen to minimize the information flow across the interfaces.
5. The number of layers should be large enough that distinct functions need not be thrown together in the same layer out of necessity and small enough that architecture does not become unwieldly.

Feature of OSI Model:

1. Big picture of communication over network is understandable through this OSI model.
2. We see how hardware and software work together.
3. We can understand new technologies as they are developed.
4. Troubleshooting is easier by separate networks.
5. Can be used to compare basic functional relationships on different networks.

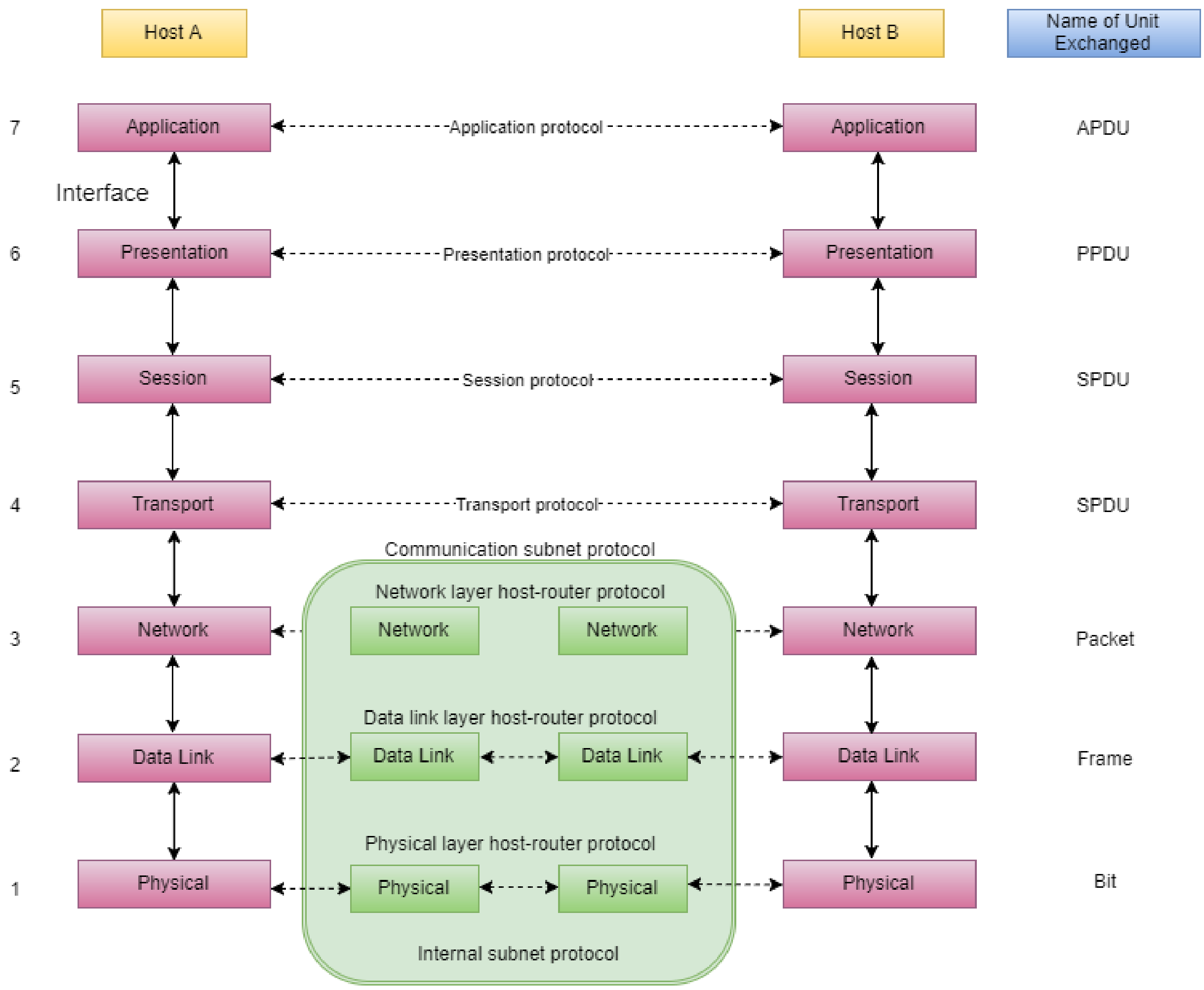

Fig. 1.8 OSI Reference Model

#Description of Different Layers:

Layer 1: The Physical Layer:

1. It is the lowest layer of the OSI Model.
2. It activates, maintains and deactivates the physical connection.
3. It is responsible for transmission and reception of the unstructured raw data over network.
4. Voltages and data rates needed for transmission is defined in the physical layer.
5. It converts the digital/ analog bits into electrical signal or optical signals.
6. Data encoding is also done in this layer.

Layer 2: Data Link Layer:

1. Data link layer synchronizes the information which is to be transmitted over the physical layer.
2. The main function of this layer is to make sure data transfer is error free from one node to another, over the physical layer.
3. Transmitting and receiving data frames sequentially is managed by this layer.
4. This layer sends and expects acknowledgements for frames received and sent respectively. Resending of non-acknowledgement received frames is also handled by this layer.
5. This layer establishes a logical layer between two nodes and also manages the Frame traffic control over the network. It signals the transmitting node to stop, when the frame buffers are full.

Layer 3: The Network Layer:

1. It routes the signal through different channels from one node to other.
2. It acts as a network controller. It manages the Subnet traffic.
3. It decides by which route data should take.
4. It divides the outgoing messages into packets and assembles the incoming packets into messages for higher levels.

Layer 4: Transport Layer:

1. It decides if data transmission should be on parallel path or single path.
2. Functions such as Multiplexing, Segmenting or Splitting on the data are done by this layer
3. It receives messages from the Session layer above it, converts the message into smaller units and passes it on to the Network layer.

4. Transport layer can be very complex, depending upon the network requirements.

Layer 5: The Session Layer:

1. Session layer manages and synchronize the conversation between two different applications.
2. Transfer of data from source to destination session layer streams of data are marked and are resynchronized properly, so that the ends of the messages are not cut prematurely, and data loss is avoided.

Layer 6: The Presentation Layer:

1. Presentation layer takes care that the data is sent in such a way that the receiver will understand the information (data) and will be able to use the data.
2. While receiving the data, presentation layer transforms the data to be ready for the application layer.
3. Languages (syntax) can be different of the two communicating systems. Under this condition presentation layer plays a role of translator.
4. It performs Data compression, Data encryption, Data conversion etc.

Layer 7: Application Layer:

1. It is the topmost layer.
2. Transferring of files disturbing the results to the user is also done in this layer. Mail services, directory services, network resource etc are services provided by application layer.
3. This layer mainly holds application programs to act upon the received and to be sent data.

Merits of OSI reference model:

1. OSI model distinguishes well between the services, interfaces and protocols.
2. Protocols of OSI model are very well hidden.
3. Protocols can be replaced by new protocols as technology changes.
4. Supports connection-oriented services as well as connectionless service.

Demerits of OSI reference model:

1. Model was devised before the invention of protocols.
2. Fitting of protocols is tedious task.
3. It is just used as a reference model.

#TCP/IP Reference Model:

The TCP/IP reference model was developed prior to OSI model. The major design goals of this model were,

1. To connect multiple networks together so that they appear as a single network.
2. To survive after partial subnet hardware failures.
3. To provide a flexible architecture.

Unlike OSI reference model, TCP/IP reference model has only 4 layers. They are,

1. Host-to-Network Layer
2. Internet Layer
3. Transport Layer
4. Application Layer

Layer 1: Host-to-network Layer

1. Lowest layer of the all.
2. Protocol is used to connect to the host, so that the packets can be sent over it.
3. Varies from host to host and network to network.

Layer 2: Internet layer

1. Selection of a packet switching network which is based on a connectionless internetwork layer is called a internet layer.
2. It is the layer which holds the whole architecture together.
3. It helps the packet to travel independently to the destination.
4. Order in which packets are received is different from the way they are sent.
5. IP (Internet Protocol) is used in this layer.
6. The various functions performed by the Internet Layer are:
    - Delivering IP packets
    - Performing routing
    - Avoiding congestion

Layer 3: Transport Layer

1. It decides if data transmission should be on parallel path or single path.
2. Functions such as multiplexing, segmenting or splitting on the data is done by transport layer.
3. The applications can read and write to the transport layer.
4. Transport layer adds header information to the data.
5. Transport layer breaks the message (data) into small units so that they are handled more efficiently by the network layer.
6. Transport layer also arrange the packets to be sent, in sequence.

Layer 4: Application Layer

The TCP/IP specifications described a lot of applications that were at the top of the protocol stack. Some of them were TELNET, FTP, SMTP, DNS etc.

1. TELNET is a two-way communication protocol which allows connecting to a remote machine and run applications on it.
2. FTP (File Transfer Protocol) is a protocol, that allows File transfer amongst computer users connected over a network. It is reliable, simple and efficient.
3. SMTP (Simple Mail Transport Protocol) is a protocol, which is used to transport electronic mail between a source and destination, directed via a route.
4. DNS (Domain Name Server) resolves an IP address into a textual address for Hosts connected over a network.
5. It allows peer entities to carry conversation.
6. It defines two end-to-end protocols: TCP and UDP
   - **TCP (Transmission Control Protocol):** It is a reliable connection-oriented protocol which handles byte-stream from source to destination without error and flow control.
   - **UDP (User-Datagram Protocol):** It is an unreliable connection-less protocol that does not want TCPs, sequencing and flow control. Example: One-shot request-reply kind of service.

Merits of TCP/IP model

1. It operated independently.
2. It is scalable.
3. Client/server architecture.
4. Supports number of routing protocols.
5. Can be used to establish a connection between two computers.

Demerits of TCP/IP

1. In this, the transport layer does not guarantee delivery of packets.
2. The model cannot be used in any other application.
3. Replacing protocol is not easy.
4. It has not clearly separated its services, interfaces and protocols.

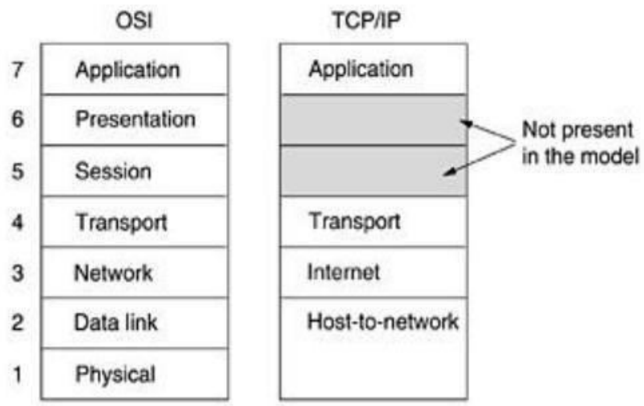

Fig 1.9 The TCP/IP reference model.

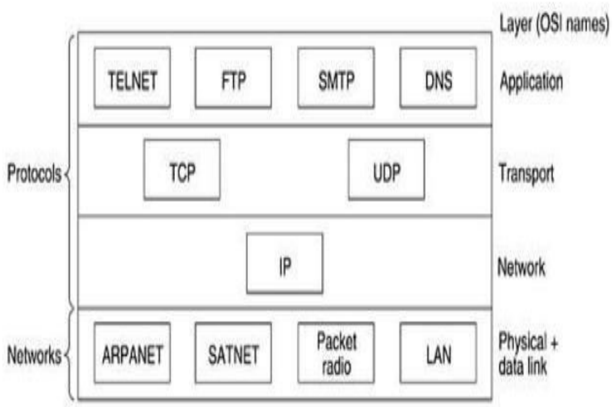

Fig 1.10 Protocols in the TCP/IP model initially.

#Comparison of the OSI and TCP/IP Reference Models:

OSI (Open System Interconnection)	TCP/IP (Transmission Control Protocol / Internet Protocol)
1. OSI is a generic, protocol independent standard, acting as a communication gateway	1. TCP/IP model is based on standard protocols around which the Internet has developed. It is a communication

between the network and end user.	protocol, which allows connection of hosts over a network.
2. In OSI model the transport layer guarantees the delivery of packets.	2. In TCP/IP model the transport layer does not guarantees delivery of packets. Still the TCP/IP model is more reliable.
3. Follows vertical approach.	3. Follows horizontal approach.
4. OSI model has a separate Presentation layer and Session layer.	4. TCP/IP does not have a separate Presentation layer or Session layer.
5. Transport Layer is Connection Oriented.	5. Transport Layer is both Connection Oriented and Connection less.
6. Network Layer is both Connection Oriented and Connection less.	6. Network Layer is Connection less.
7. OSI is a reference model around which the networks are built. Generally, it is used as a guidance tool.	7. TCP/IP model is, in a way implementation of the OSI model.
8. Network layer of OSI model provides both connection oriented and connectionless service.	8. The Network layer in TCP/IP model provides connectionless service.
9. OSI model has a problem of fitting the protocols into the model.	9. TCP/IP model does not fit any protocol
10. Protocols are hidden in OSI model and are easily replaced as the technology changes.	10. In TCP/IP replacing protocol is not easy.
11. OSI model defines services, interfaces and protocols very clearly and makes clear distinction between them. It is protocol independent.	11. In TCP/IP, services, interfaces and protocols are not clearly separated. It is also protocol dependent.
12. It has 7 layers	12. It has 4 layers

#Network Standardization

International Organization for Standardization One of the most important standards-making bodies is the International Organization for Standardization (ISO),2 which makes technical recommendations about data communication interfaces (see **www.iso.org).** ISO is based in Geneva, Switzerland. The membership is composed of the national standards organizations of each ISO member country.

International Telecommunications Union—Telecommunications Group the Telecommunications Group (ITU-T) is the technical standards-setting organization of the United Nations International Telecommunications Union, which is also based in Geneva (see **www.itu.int**). ITU is composed of representatives from about 200-member countries. Membership was originally focused on just the public telephone companies in each country, but a major reorganization in 1993 changed this, and ITU now seeks members among public- and private-sector organizations who operate computer or communications networks (e.g., RBOCs) or build software and equipment for them (e.g., AT&T).

American National Standards Institute: The American National Standards Institute (ANSI) is the coordinating organization for the U.S. national system of standards for both technology and nontechnology (see **www.ansi.org**). ANSI has about 1,000 members from both public and private organizations in the United States. ANSI is a standardization organization, not a standards-making body, in that it accepts standards developed by other organizations and publishes them as American standards. Its role is to coordinate the development of voluntary national standards and to interact with ISO to develop national standards that comply with ISO's international recommendations. ANSI is a voting participant in the ISO.

#Queueing Models: Little's Theorem, Queueing System:

Queueing theory is the mathematical study of waiting lines or queues. In queueing theory a model is constructed so that queue lengths and waiting times can be predicted. Queueing theory is generally considered a branch of operations research because the results are often used when making business decisions about the resources needed to provide a service.

#Little's Theorem

We begin our analysis of queueing systems by understanding Little's Theorem. Little's theorem states that:
The average number of customers (N) can be determined from the following equation: **N = λT**
Here lambda is the average customer arrival rate and T is the average service time for a customer.
Consider the example of a restaurant where the customer arrival rate (lambda) doubles but the customers still spend the same amount of time in the restaurant (T). These will double the number of customers in the restaurant (N). By the same logic if the customer arrival rate remains the same but the customers service time doubles, this will also double the total number of customers in the restaurant.

Queueing System Classification

With Little's Theorem, we have developed some basic understanding of a queueing system. To further our understanding, we will have to dig deeper into characteristics of a queueing system that impact its performance. For example, queueing requirements of a restaurant will depend upon factors like:

- How do customers arrive in the restaurant? Are customer arrivals more during lunch and dinner time (a regular restaurant)? Or is the customer traffic more uniformly distributed (a cafe)?
- How much time do customers spend in the restaurant? Do customers typically leave the restaurant in a fixed amount of time? Does the customer service time vary with the type of customer?
- How many tables does the restaurant have for servicing customers?

The above three points correspond to the most important characteristics of a queueing system.
They are explained below:

Arrival Process

- The probability density distribution that determines the customer arrivals in the system.
- In a messaging system, this refers to the message arrival probability distribution.

Service Process

- The probability density distribution that determines the customer service times in the system.
- In a messaging system, this refers to the message transmission time distribution. Since message transmission is directly proportional to the length of the message, this parameter indirectly refers to the message length distribution.

Number of Servers

- Number of servers available to service the customers.
- In a messaging system, this refers to the number of links between the source and destination nodes.

Based on the above characteristics, queueing systems can be classified by the following convention:

A/S/n

Where A is the arrival process, S is the service process and n is the number of servers. A and S are can be any of the following:

M (Markov)	Exponential probability density
D (Deterministic)	All customers have the same value
G (General)	Any arbitrary probability distribution

Examples of queueing systems that can be defined with this convention are:

- **M/M/1:** This is the simplest queueing system to analyse. Here the arrival and service time are negative exponentially distributed (poison process). The system consists of only one server. This queueing system can be applied to a wide variety of problems as any system with a very large number of independent customers can be approximated as a Poisson process. Using a Poisson process for service time however is not applicable in many applications and is only a crude approximation. Refer to M/M/1 Queueing System for details.

• **M/D/n:** Here the arrival process is poison and the service time distribution is deterministic. The system has n servers. (e.g. a ticket booking counter with n cashiers.) Here the service time can be assumed to be same for all customers)

• **G/G/n:** This is the most general queueing system where the arrival and service time processes are both arbitrary. The system has n servers. No analytical solution is known for this queueing system.

Queueing Models: Little's Theorem, Queueing System -3

M/M/c queue

In queueing theory, a discipline within the mathematical theory of probability, the M/M/c queue (or Erlang–C model is a multi-server queueing model. In Kendall's notation it describes a system where arrivals form a single queue and are governed by a Poisson process, there are c servers and job service times are exponentially distributed. It is a generalisation of the M/M/1 queue which considers only a single server. The model with infinitely many servers is the M/M/∞ queue.

M/M/1 queue

In queueing theory, a discipline within the mathematical theory of probability, an M/M/1 queue represents the queue length in a system having a single server, where arrivals are determined by a Poisson process and job service times have an exponential distribution. The model name is written in Kendall's notation. The model is the most elementary of queueing models and an attractive object of study as closed-form expressions can be obtained for many metrics of interest in this model. An extension of this model with more than one server is the M/M/c queue.

M/M/∞

In queueing theory, a discipline within the mathematical theory of probability, the **M/M/∞ queue** is a multi-server queueing model where every arrival experiences immediate service and does not wait In Kendall's notation it describes a system where arrivals are governed by a Poisson process, there are infinitely many servers, so jobs do not need to wait for a server. Each job has an exponentially distributed service time. It is a limit of the M/M/c queue model where the number of servers *c* becomes very large. The model can be used to model bound lazy deletion performance.

M/G/1

In queueing theory, a discipline within the mathematical theory of probability, an **M/G/1 queue** is a queue model where arrivals are **M**arkovian (modulated by a Poisson process), service times have a **G**eneral distribution and there is a single server. The model name is written in Kendall's notation, and is an extension of the M/M/1 queue, where service times must be exponentially distributed. The classic application of the M/G/1 queue is to model performance of a fixed head hard disk.

Model definition

A queue represented by a M/G/1 queue is a stochastic process whose state space is the set {0,1,2,3...}, where the value corresponds to the number of customers in the queue, including any being served. Transitions from state i to $i + 1$ represent the arrival of a new customer: the times between such arrivals have an exponential distribution with parameter λ. Transitions from state i to $i - 1$ represent a customer who has been served, finishing being served and departing: the length of time required for serving an individual customer has a general distribution function. The lengths of times between arrivals and of service periods are random variables which are assumed to be statistically independent.

Scheduling policies

Customers are typically served on a first-come, first-served basis, other popular scheduling policies include

- processor sharing where all jobs in the queue share the service capacity between them equally
- last-come, first served without pre-emption where a job in service cannot be interrupted
- last-come, first served with pre-emption where a job in service is interrupted by later arrivals, but work is conserved
- generalized foreground-background (FB) scheduling also known as least-attained service where the jobs which have received least processing time so far are served first and jobs which have received equal service time share service capacity using processor sharing
- shortest job first without pre-emption (SJF) where the job with the smallest size receives service and cannot be interrupted until service completes
- pre-emptive shortest job first where at any moment in time the job with the smallest original size is served

- shortest remaining processing time (SRPT) where the next job to serve is that with the smallest remaining processing requirement

Service policies are often evaluated by comparing mean sojourn times in the queue. If service times that jobs require are known on arrival then the optimal scheduling policy is SRPT. Policies can also be evaluated using a measure of fairness.