Unit: 2 Communication in Distributed System

Computer Network and its classification

- A computer network is a system in which multiple computers are connected to each other to share information and resources.
- The physical connection between networked computing devices is established using either cable media or wireless media.
- The best-known computer network is the Internet.

Figure: Computer Network

Local area network (LAN)

- It is privately-owned networks within a single building or campus of up to a few km in size.
- They are widely used to connect personal computers and workstations in company offices and factories to share resources (e.g., printers) and exchange information.
- LANs are easy to design and troubleshoot
- In LAN, all the machines are connected to a single cable.
- Different types of topologies such as Bus, Ring, Star and Tree are used.
- The data transfer rates for LAN is up to 10 Gbps.
- They transfer data at high speeds. High transmission rate are possible in LAN because of the short distance between various computer networks.
- They exist in a limited geographical area.

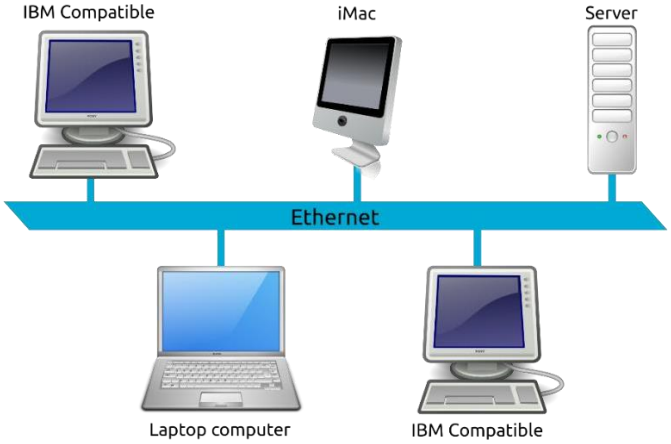

Figure: Local area network (LAN)

Wide area Network (WAN)

- WAN spans a large geographical area, often a country or region.
- WAN links different metropolitan's countries and national boundaries there by enabling easy communication.
- It may be located entirely with in a state or a country or it may be interconnected around the world.
- It contains a collection of machines intended for running user (i.e., application) programs. We will follow traditional usage and call these machines hosts.
- The communication between different users of WAN is established using leased telephone lines or satellite links and similar channels.

Figure: Wide area network (WAN)

Metropolitan area network (MAN)

- MAN is a larger version of LAN which covers an area that is larger than the covered by LAN but smaller than the area covered by WAN.
- A metropolitan area network or MAN covers a city.
- The best-known example of a MAN is the cable television network available in many cities.
- MAN connects two or more LANs.

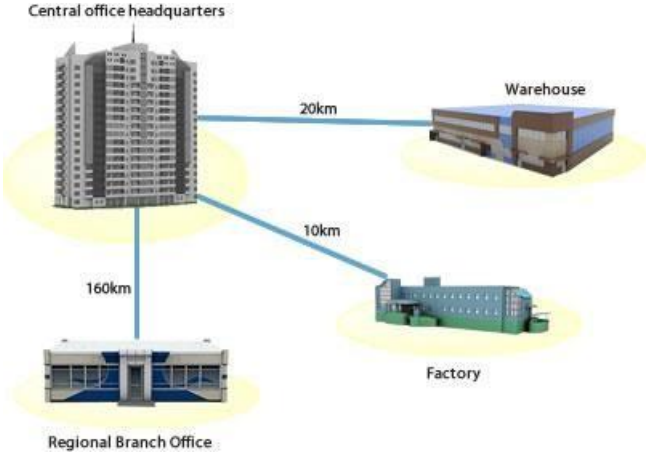

Figure: Metropolitan area network (MAN)

Wireless LAN

- WLAN technology allows computer network devices to have network connectivity without the use of wires, using radio waves.
- These networks use infrared link, Bluetooth or low power radio network operating at a bandwidth of 1-2 Mbps over a distance of 10 m.
- At the top level, wireless networks can be classified as wireless LAN, wireless MAN, and wireless WAN.
- The wireless LANs are further classified as Personal Area networks (e.g. Bluetooth, wireless sensor networks) and Business LAN (e.g. 802.11b).

Figure: Wireless LAN

Classification of Network

Basis Of Comparison	LAN	MAN	WAN
Expands to	Local Area Network	Metropolitan Area Network	Wide Area Network
Meaning	A network that connects a group of computers in a small geographical area	It covers relatively large region such as cities, towns	It spans large locality & connects countries together.
Ownership of Network	Private	Private or Public	Private or Public
Design and Maintenance	Easy	Difficult	Difficult
Propagation Delay	Short	Moderate	Long
Speed	High	Moderate	Low
Equipment Needed	Switch, Hub	Modem, Router	Microwave, Radio Transmitters & Receivers
Range(Approx.)	1 to 10 km	In 100 km	Beyond 100 km
Used for	College, School, Hospital	Small towns, City	Country/Continent

OSI Model layered architecture

OSI Layer Architecture

- OSI model is based on a proposal developed by the International Standards Organization (ISO) as a first step toward international standardization of the protocols used in the various layers.
- It was revised in 1995.
- The model is called the OSI (Open Systems Interconnection) Reference Model because it deals with connecting open systems—that is, systems that are open for communication with other systems.
- The OSI model has seven layers.
    - Physical Layer
    - Data Link Layer
    - Network Layer
    - Transport Layer
    - Session Layer
    - Presentation Layer
    - Application Layer

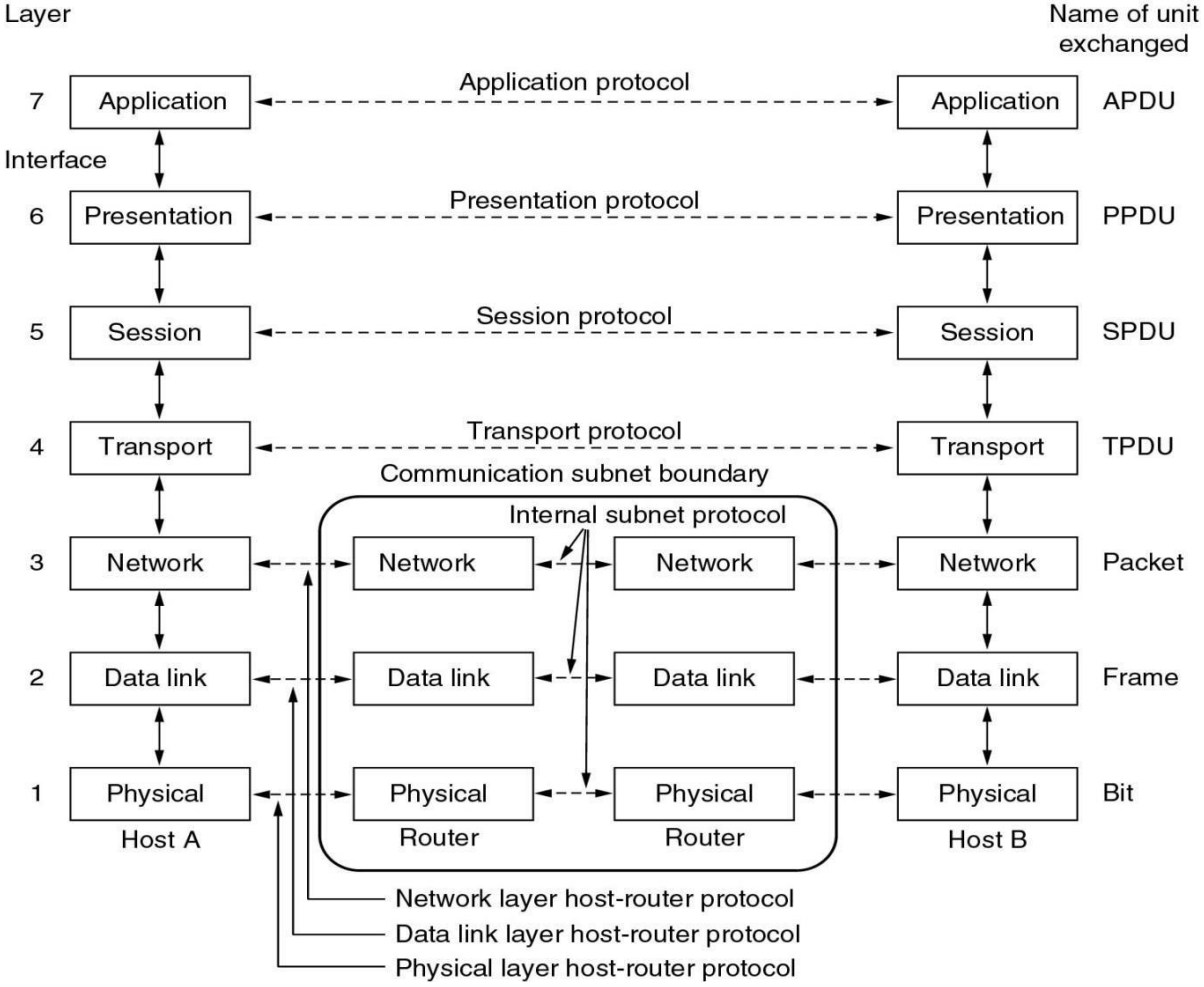

Figure: OSI Reference Model

Physical Layer

- The physical layer, the lowest layer of the OSI model, is concerned with the transmission and

reception of the unstructured raw bit stream over a physical medium.

- It describes the electrical/optical, mechanical, and functional interfaces to the physical medium, and carries the signals for all of the higher layers. It provides:
    (1) Data encoding: modifies the simple digital signal pattern (1s and 0s) used by the PC to better accommodate the characteristics of the physical medium, and to aid in bit and frame synchronization.
    (2) Transmission technique: determines whether the encoded bits will be transmitted by baseband (digital) or broadband (analog) signalling.
    (3) Physical medium transmission: transmits bits as electrical or optical signals appropriate for the physical medium.

Data link Layer

- The data link layer provides error-free transfer of data frames from one node to another over the physical layer, allowing layers above it to assume virtually error-free transmission over the link.
- To do this, the data link layer provides:
    (1) Link establishment and termination: establishes and terminates the logical link between two nodes.
    (2) Frame traffic control: tells the transmitting node to "back-off" (stop) when no frame buffers are available.
    (3) Frame sequencing: transmits/receives frames sequentially.
    (4) Frame acknowledgment: provides/expects frame acknowledgments. Detects and recovers from errors that occur in the physical layer by retransmitting non-acknowledged frames and handling duplicate frame receipt.
    (5) Frame delimiting: creates and recognizes frame boundaries.
    (6) Frame error checking: checks received frames for integrity.
    (7) Media access management: determines when the node "has the right" to use the physical medium.

Network Layer

- The network layer controls the operation of the subnet, deciding which physical path the data should take based on network conditions, priority of service, and other factors.
- To do this, the network layer provides:
    (1) Routing: routes frames among networks.
    (2) Subnet traffic control: routers (network layer intermediate systems) can instruct a sending station to "throttle back" its frame transmission when the router's buffer fills up.
    (3) Frame fragmentation: if it determines that a downstream router's maximum transmission unit (MTU) size is less than the frame size, a router can fragment a frame for transmission and re-assembly at the destination station.
    (4) Logical-physical address mapping: translates logical addresses or names, into physical addresses.

(5) Subnet usage accounting: has accounting functions to keep track of frames forwarded by subnet intermediate systems, to produce billing information.

Transport Layer

- The transport layer ensures that messages are delivered error-free, in sequence, and with no losses or duplications. It relieves (release) the higher layer protocols from any concern with the transfer of data between them and their peers.
- The size and complexity of a transport protocol depends on the type of service it can get from the network layer. For a reliable network layer with virtual circuit capability, a minimal transport layer is required. If the network layer is unreliable and/or only supports datagrams, the transport protocol should include extensive error detection and recovery.
- The transport layer provides:
  (1) Message segmentation: accepts a message from the (session) layer above it, splits the message into smaller units (if not already small enough), and passes the smaller units down to the network layer. The transport layer at the destination station reassembles the message.
  (2) Message acknowledgment: provides reliable end-to-end message delivery with acknowledgments.
  (3) Message traffic control: tells the transmitting station to "back-off" when no message buffers are available.
- Typically, the transport layer can accept relatively large messages, but there are strict message size limits imposed by the network (or lower) layer. Consequently, the transport layer must break up the messages into smaller units, or frames, prepending a header to each frame.
- The transport layer header information must then include control information, such as message start and message end flags, to enable the transport layer on the other end to recognize message boundaries.
- In addition, if the lower layers do not maintain sequence, the transport header must contain sequence information to enable the transport layer on the receiving end to get the pieces back together in the right order before handing the received message up to the layer above.

Session Layer

- The session layer allows session establishment between processes running on different stations. It provides:
  (1) Session establishment, maintenance and termination: allows two application processes on different machines to establish, use and terminate a connection, called a session.
  (2) Session support: performs the functions that allow these processes to communicate over the network, performing security, name recognition, logging, and so on.

Presentation Layer

- The presentation layer formats the data to be presented to the application layer. It can be viewed as the translator for the network. This layer may translate data from a format used by the application layer into a common format at the sending station, then translate the common format to a format known to the application layer at the receiving station.
- The presentation layer provides:
  (1) Character code translation: for example, ASCII to EBCDIC.
  (2) Data conversion: bit order, CR-CR/LF, integer-floating point, and so on.

(3) Data compression: reduces the number of bits that need to be transmitted on the network.
(4) Data encryption: encrypt data for security purposes. For example, password encryption.

Application Layer

- The application layer serves as the window for users and application processes to access network services.
- This layer contains a variety of commonly needed functions:
    1. Resource sharing and device redirection
    2. Remote file access
    3. Remote printer access
    4. Inter-process communication
    5. Network management
    6. Directory services
    7. Electronic messaging (such as mail)
    8. Network virtual terminals

TCP/IP Reference model

- Transmission Control Protocol/Internet Protocol (TCP/IP) protocol suite is the engine for the Internet and networks worldwide.
- TCP/IP either combines several OSI layers into a single layer, or does not use certain layers at all.
- TCP/IP is a set of protocols developed to allow cooperating computers to share resources across the network.
- The TCP/IP model has five layers.
    - Application Layer
    - Transport Layer
    - Internet Layer
    - Data Link Layer
    - Physical Network

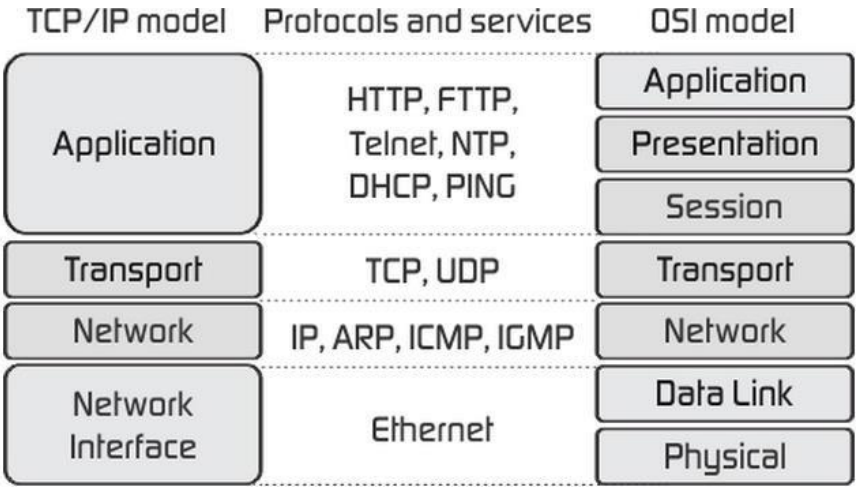

Figure: TCP/IP Reference Model

- As we can see from the above figure, presentation and session layers are not there in TCP/IP model. Also note that the Network Access Layer in TCP/IP model combines the functions of Data Link Layer and Physical Layer.

Application Layer

- Application layer is the top most layer of four layer TCP/IP model.
- Application layer is present on the top of the Transport layer.
- Application layer defines TCP/IP application protocols and how host programs interface with Transport layer services to use the network.
- Application layer includes all the higher-level protocols like DNS (Domain Naming System), HTTP (Hypertext Transfer Protocol), Telnet, SSH, FTP (File Transfer Protocol), TFTP (Trivial File Transfer Protocol), SNMP (Simple Network Management Protocol), SMTP (Simple Mail Transfer Protocol), DHCP (Dynamic Host Configuration Protocol), X Windows, RDP (Remote Desktop Protocol) etc.

Transport Layer

- The purpose of Transport layer is to permit devices on the source and destination hosts to carry on a conversation.
- Transport layer defines the level of service and status of the connection used when transporting data.
- The transport layer provides the end-to-end data transfer by delivering data from an application to its remote peer.
- The most-used transport layer protocol is the Transmission Control Protocol (TCP), which provides:
    (1) Reliable delivery data
    (2) Duplicate data suppression
    (3) Congestion control
    (4) Flow control
- Another transport layer protocol is the User Datagram Protocol (UDP), which provides:
    (1) Connectionless
    (2) Unreliable
    (3) Best-effort service
- UDP is used by applications that need a fast transport mechanism and can tolerate the loss of some data.

Network Layer (Internet Layer)

- The internet layer also called the network layer.
- Internet layer pack data into data packets known as IP datagrams, which contain source and destination address (logical address or IP address) information that is used to forward the datagrams between hosts and across networks.
- The Internet layer is also responsible for routing of IP datagrams.
- Internet Protocol (IP) is the most important protocol in this layer.
- It is a connectionless protocol that does not assume reliability from lower layers. IP does not provide reliability, flow control or error recovery.

- IP provides a routing function that attempts to deliver transmitted messages to their destination.
- These message units in an IP network are called an IP datagram.
- Example: IP, ICMP, IGMP, ARP, and RARP.

Network Interface Layer (Network Access Layer)

- Network Access Layer defines details of how data is physically sent through the network, including how bits are electrically or optically signalled by hardware devices that interface directly with a network medium, such as coaxial cable, optical fiber, or twisted pair copper wire.
- The protocols included in Network Access Layer are Ethernet, Token Ring, FDDI, X.25, Frame Relay etc.

OSI(Open System Interconnection)	**TCP/IP (Transmission Control Protocol/ Internet Protocol)**
OSI provides layer functioning and also defines functions of all the layers.	TCP/IP model is more based on protocols and protocols are not flexible with other layers.
In OSI model the transport layer guarantees the delivery of packets	In TCP/IP model the transport layer does not guarantees delivery of packets.
Follows horizontal approach	Follows vertical approach.
OSI model has a separate presentation layer	TCP/IP doesn't have a separate presentation layer
OSI is a general model.	TCP/IP model cannot be used in any other application.
Network layer of OSI model provide both connection oriented and connectionless service.	The Network layer in TCP/IP model provides connectionless service.
OSI model has a problem of fitting the protocols in the model	TCP/IP model does not fit any protocol
Protocols are hidden in OSI model and are easily replaced as the technology changes.	In TCP/IP replacing protocol is not easy.
OSI model defines services, interfaces and protocols very clearly and makes clear distinction between them.	In TCP/IP it is not clearly separated its services, interfaces and protocols.
It has 7 layers	It has 4 layers

Inter process communication (IPC) in message passing system

- Inter process communication (IPC) basically requires information sharing among two or more processes.

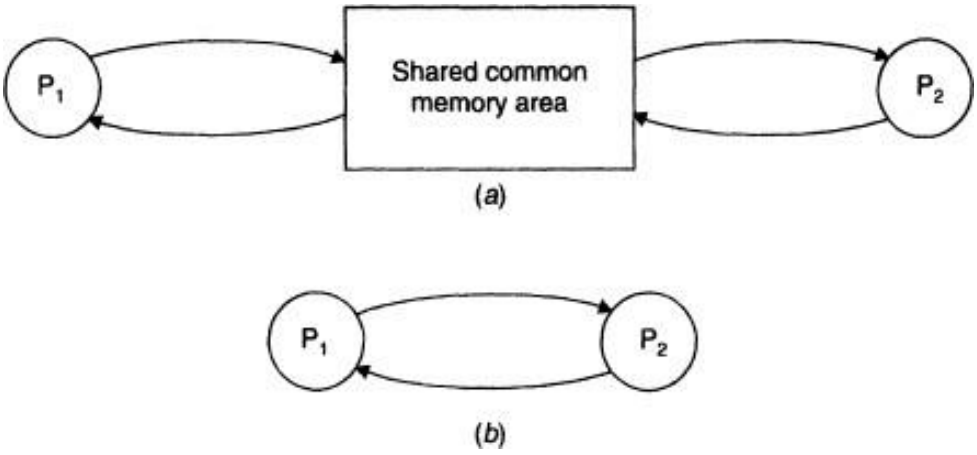

Figure: Inter process communication (a) Shared data approach (b) Message passing approach

- Two basic methods for information sharing are as follows:
  - Original sharing, or shared-data approach.
  - Copy sharing, or message-passing approach.
- In the shared-data approach, the information to be shared is placed in a common memory area that is accessible to all processes involved in an IPC.
- In the message-passing approach, the information to be shared is physically copied from the sender process's space to the address space of all the receiver processes.
- A message-passing system is a subsystem of distributed operating system that provides a set of message-based IPC protocols, and does so by shielding the details of complex network protocols and multiple heterogeneous platforms from programmers.

It enables processes to communicate by exchanging messages and allows programs to be written by using simple communication primitives, such as send and receive.

Desirable features of a good message passing system

Simplicity

- A message passing system should be simple and easy to use.
- It must be straight forward to construct new applications and to communicate with existing one by using the primitives provided by message passing system.

Uniform Semantics

- In a distributed system, a message-passing system may be used for the following two types of inter process communication:
- Local communication, in which the communicating processes are on the same node.
- Remote communication, in which the communicating processes are on different nodes.
- Semantics of remote communication should be as close as possible to those of local communications.

Efficiency

- An IPC protocol of a message-passing system can be made efficient by reducing the number of message exchanges, as far as practicable, during the communication process.
- Some optimizations normally adopted for efficiency include the following:
- Avoiding the costs of establishing and terminating connections between the same pair of processes for each and every message exchange between them.
- Minimizing the costs of maintaining the connections;
- Piggybacking of acknowledgement of previous messages with the next message during a connection between a sender and a receiver that involves several message exchanges.

Reliability

- A good IPC protocol can cope with failure problems and guarantees the delivery of a message.
- Handling of lost messages involves acknowledgement and retransmission on the basis of timeouts.
- A reliable IPC protocol should also be capable of detecting and handling duplicate messages.
- Correctness
- Correctness is a feature related to IPC protocols for group communication.
- Issues related to correctness are as follows:
    (1) Atomicity: It ensures that every message sent to a group of receivers will be delivered to either all of them or none of them.
    (2) Ordered delivery: It ensures that messages arrive to all receivers in an order acceptable to the application.
    (3) Survivability: It guarantees that messages will be correctly delivered despite partial failures of processes, machines, or communication links.

Flexibility

- Not all applications require the same degree of reliability and correctness of the IPC protocol.
- The IPC protocols of a message passing system must be flexible enough to cater to the various needs of different applications.
- The IPC primitives should be such that user have the flexibility to choose and specify types and levels of reliability and correctness requirement of their applications.
- IPC primitives must also have the flexibility to permit any kind of control flow between the cooperating processes, including synchronous and asynchronous send/receive.

Security

- A good message passing system must also be capable of providing a secure end to end communication.
- Steps necessary for secure communication include the following:
- Authentication of the receiver of a message by the sender.
- Authentication of the sender of a message by its receivers.
- Encryption of a message before sending it over the network.

Portability

- The message passing system should itself be portable.
- It should be possible to easily construct a new IPC facility on another system by reusing the basic design of the existing message passing system.

- The application written by using primitives of IPC protocols of the message passing system should be portable.
- This may require use of an external data representation format for the communication taking place between two or more processes running on computers of different architectures.

Issues in IPC by Message Passing

- A message is a block of information formatted by a sending process in such a manner that it is meaningful to the receiving process.
- It consists of a fixed-length header and a variable-size collection of typed data objects.

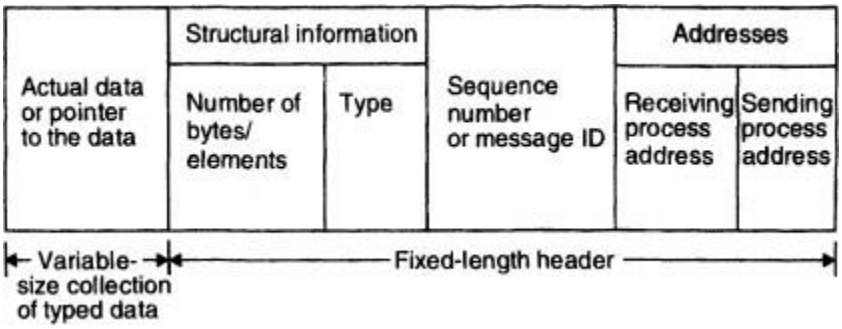

Figure: A typical message structure.

- The header usually consists of the following elements:
- Address:
  - It contains characters that uniquely identify the sending and receiving processes in the network.
- Sequence number.
  - This is the message identifier (ID), which is very useful for identifying lost messages and duplicate messages in case of system failures.
- Structural information.
  - This element also has two parts.
  - The type part specifies whether the data to be passed on to the receiver is included within the message or the message only contains a pointer to the data, which is stored somewhere outside the contiguous portion of the message.
  - The second part of this element specifies the length of the variable-size message data.
- In the design of the IPC protocol for message passing system, the following important issues need to be considered:
  - Who is the sender?
  - Who is the receiver?
  - Is there one receiver or many receivers?
  - Is the message guaranteed to have been accepted by receivers?
  - Does the sender need to wait for the reply?
  - What should be done if the catastrophic event such as node crash or a communication link failure occurs during the course of communication?
  - What should be done if the receiver is not ready to accept the message: will the message be discarded or stored in a buffer? In the case of buffering what would be done if the buffer is full?
  - If there are outstanding messages for a receiver, can it choose the order in which to service the outstanding messages?

Failure Handling in IPC

- During inter process communication partial failures such as a node crash or communication link failure may lead to the following problems:

Loss of request message

- This may happen either due to the failure of communication link betwe en the sender and receiver or because the receiver's node is down at the time the request message reaches there.

Loss of response message

- This may happen either due to the failure of communication link between the sender and receiver or because the sender's node is down at the time the response message reaches there.

Basic concept of client server model

- The structure of the operation system is like a group of cooperating processes called the servers, which offer services to the users called the clients.
- Both run on the same microkernel as the user processes.
- A machine can run as single/multiple clients or single/multiple servers, or a mix of both.
- This model depicted in Figure2.23 uses the connectionless Request Reply protocol thus reducing the overhead of the connection oriented TCP/IP or OSI protocol.

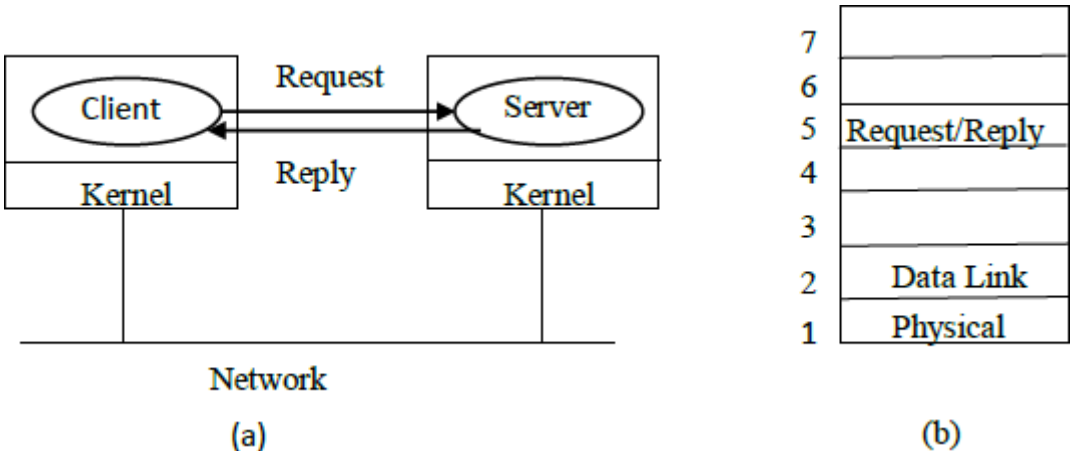

Figure: Client Server model

- The Request Reply protocol works as follows:
  - The client requesting the service sends a request message to the server.
  - The server completes the task and returns the result indicating that it has performed the task or returns an error message indicating that it has not performed the task.

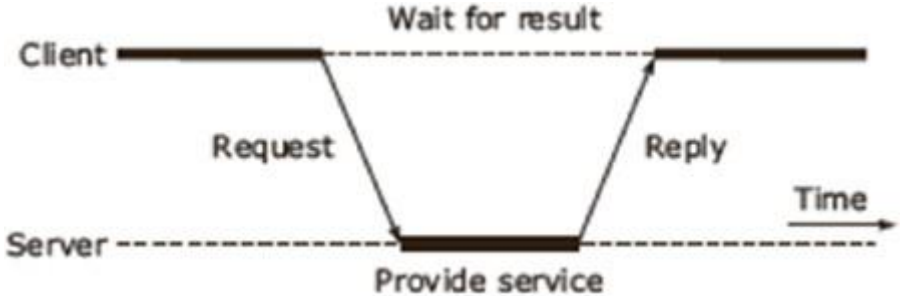

Figure: The Client server interaction

- Figure shows the client server interaction.
- Here reply serves as acknowledgement to the request.

- As seen in the above figure (b) of client server model, only three layers of the protocol are used.
- The physical and the data link layer are responsible for transferring packets between the client and the server through hardware.
- There is no need for routing and establishment of connection.
- The Request Reply protocol defines the set of request and replies to the corresponding requests. The session management and other higher layers are not required.

Remote procedure call (RPC)

- Remote Procedure Call (RPC) is a protocol that one program can use to request a service from a program located in another computer on a network without having to understand the network's details.
- A procedure call is also sometimes known as a function call or a subroutine call.
- RPC uses the client-server model.

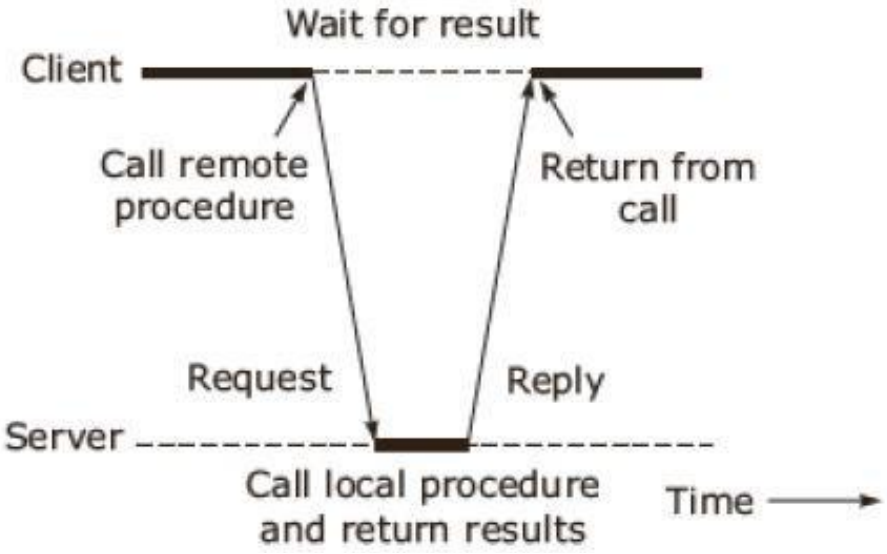

Figure: RPC Model

- It includes mainly five elements:
  1. The Client
  2. The Client stub (stub: Piece of code used for converting parameters)
  3. The RPC Runtime (RPC Communication Package)
  4. The Server stub
  5. The Server

The Client

- It is user process which initiates a remote procedure call.
- The client makes a perfectly normal call that invokes a corresponding procedure in the client stub.

The Client stub

- On receipt of a request it packs a requirement into a message and asks to RPCRuntime to send.
- On receipt of a result it unpacks the result and passes it to client.

RPC Runtime

- It handles transmission of messages between client and server.

The Server stub

- It unpacks a call request and make a perfectly normal call to invoke the appropriate procedure in the server.
- On receipt of a result of procedure execution it packs the result and asks to RPC Runtime to send.

The Server

- It executes an appropriate procedure and returns the result from a server stub.