Introduction to Distributed Computing and Basics of Various Networking Protocols.

Prof. Anand D Dave

Department of Information Technology,

Dharmsinh Desai University, Nadiad.

Distributed Computing

- What is Distributed Computing ?
- Early computing was performed on a single processor. Uni-processor computing can be called centralized computing.
- A distributed system is a system whose components are located on different networked computers, which communicate and coordinate their actions by passing messages to one another.
- The components interact with one another in order to achieve a common goal.

Distributed Computing

- Examples of distributed systems vary from SOA-based systems to massively multiplayer online games to peer-to-peer applications.
- A computer program that runs within a distributed system is called a distributed program (and distributed programming is the process of writing such programs).
- There are many different types of implementations for the message passing mechanism, including pure HTTP, RPC-like connectors and message queues.

Distributed Computing

- Distributed computing also refers to the use of distributed systems to solve computational problems. e.g. N-Queen Problem
- In distributed computing, a problem is divided into many tasks, each of which is solved by one or more computers, which communicate with each other via message passing.

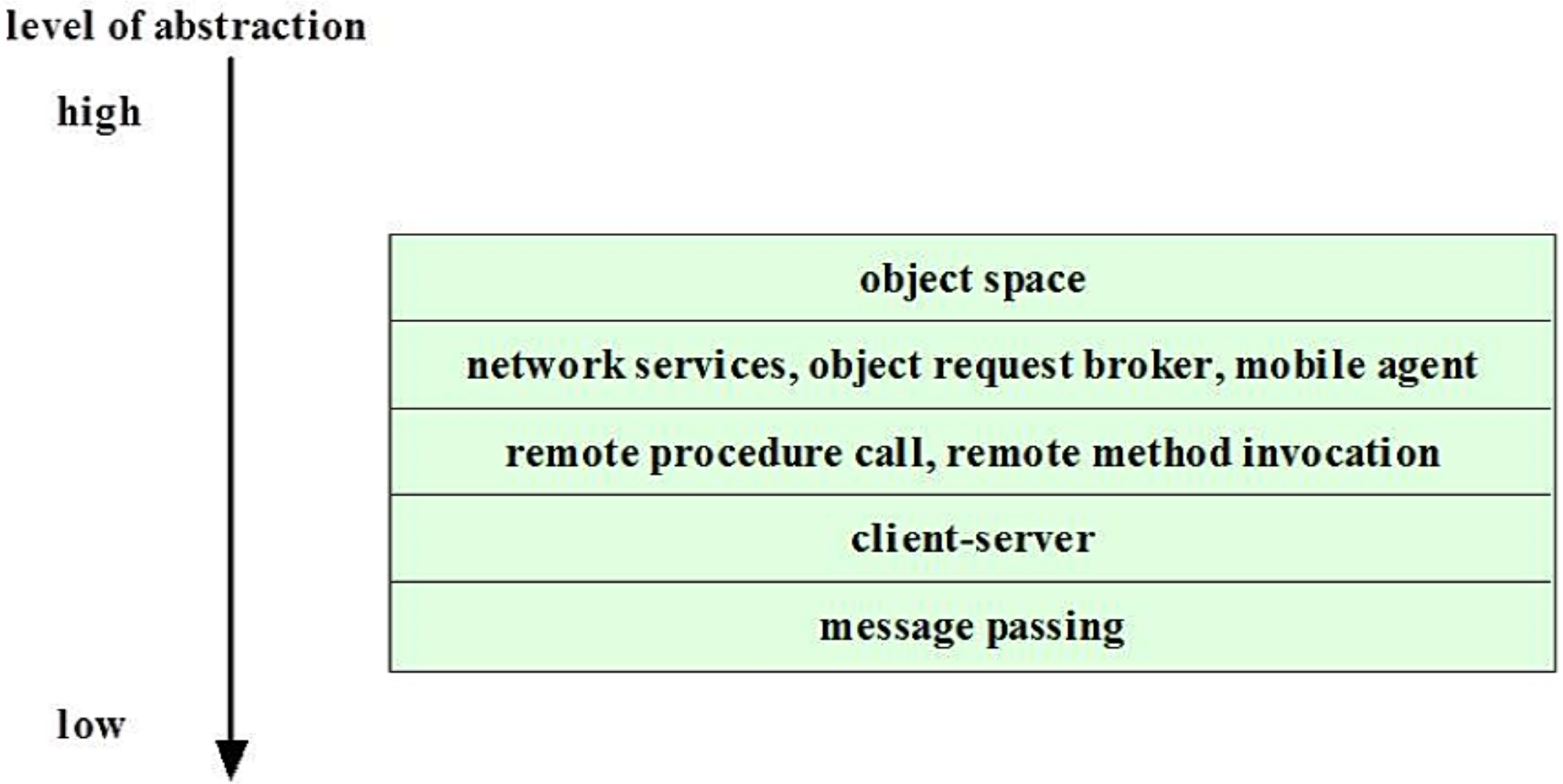
Distributed Computing

- SETI@home ("SETI at home") is an Internet-based public volunteer computing project employing the BOINC software platform created by the Berkeley SETI Research Center and is hosted by the Space Sciences Laboratory, at the University of California, Berkeley.
- Its purpose is to analyze radio signals, searching for signs of extraterrestrial intelligence, and as such is one of many activities undertaken as part of the worldwide SETI effort.

Distributed Computing Paradigms

- Paradigm means "a pattern, example, or model."
- Characteristics that distinguish distributed applications from conventional applications which run on a single machine. These characteristics are:
- Interprocess communication: A distributed application require the participation of two or more independent entities (processes). To do so, the processes must have the ability to exchange data among themselves.
- Event synchronization: In a distributed application, the sending and receiving of data among the participants of a distributed application must be synchronized

Distributed Application Paradigms



Distributed Computing: Client- Server

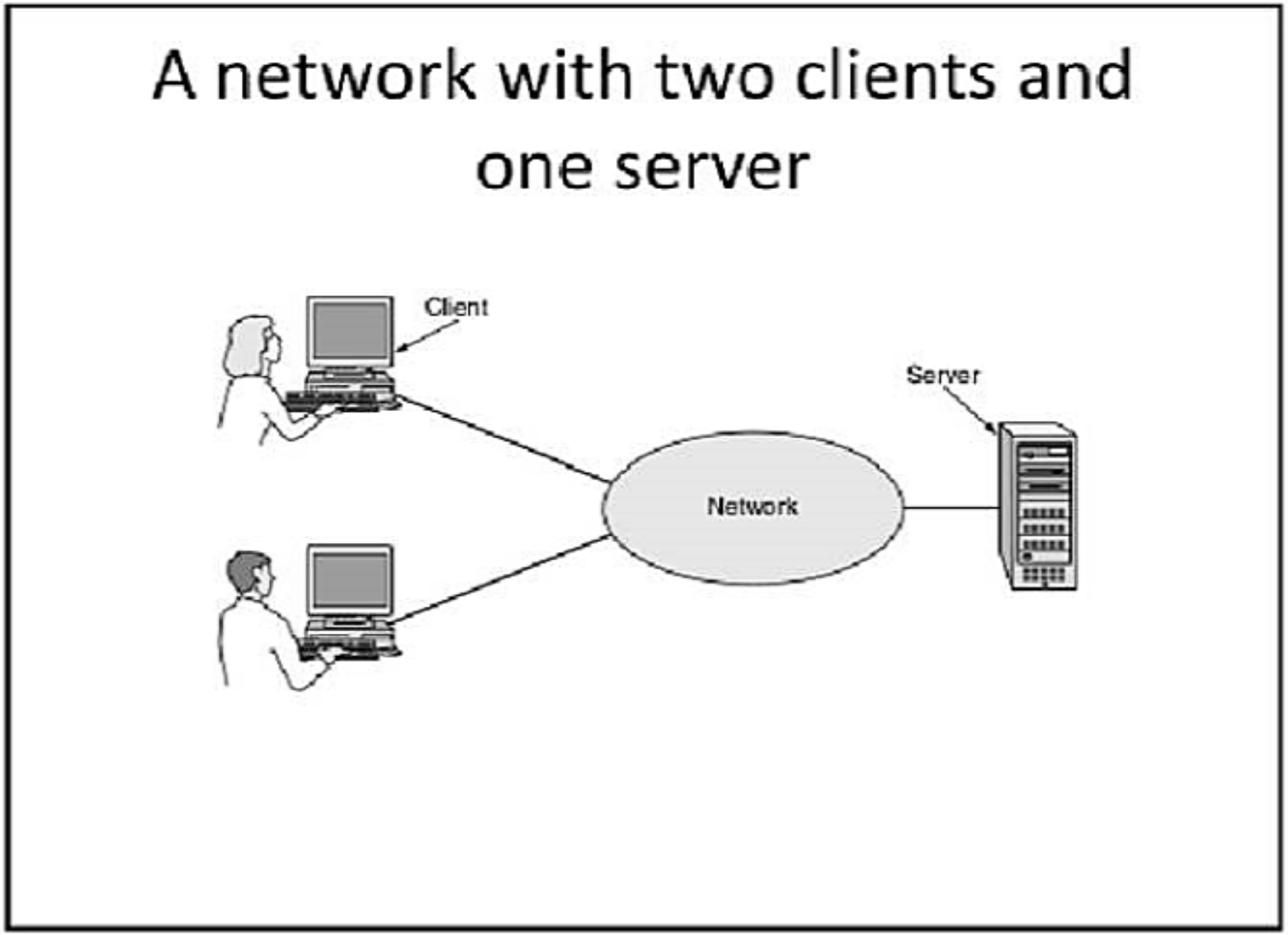
- Network applications have client and server.
- - e.g. browser and web-server
- - e.g. FTP client and FTP server

- More than one client can utilize service of a server.

- Communication between client and server requires use of protocol. e.g. TCP/IP protocol suite.

- Client adn server are user process, while TCP and IP protocols are normally part of the protocol within the kernel.

Distributed Computing: Client- Server

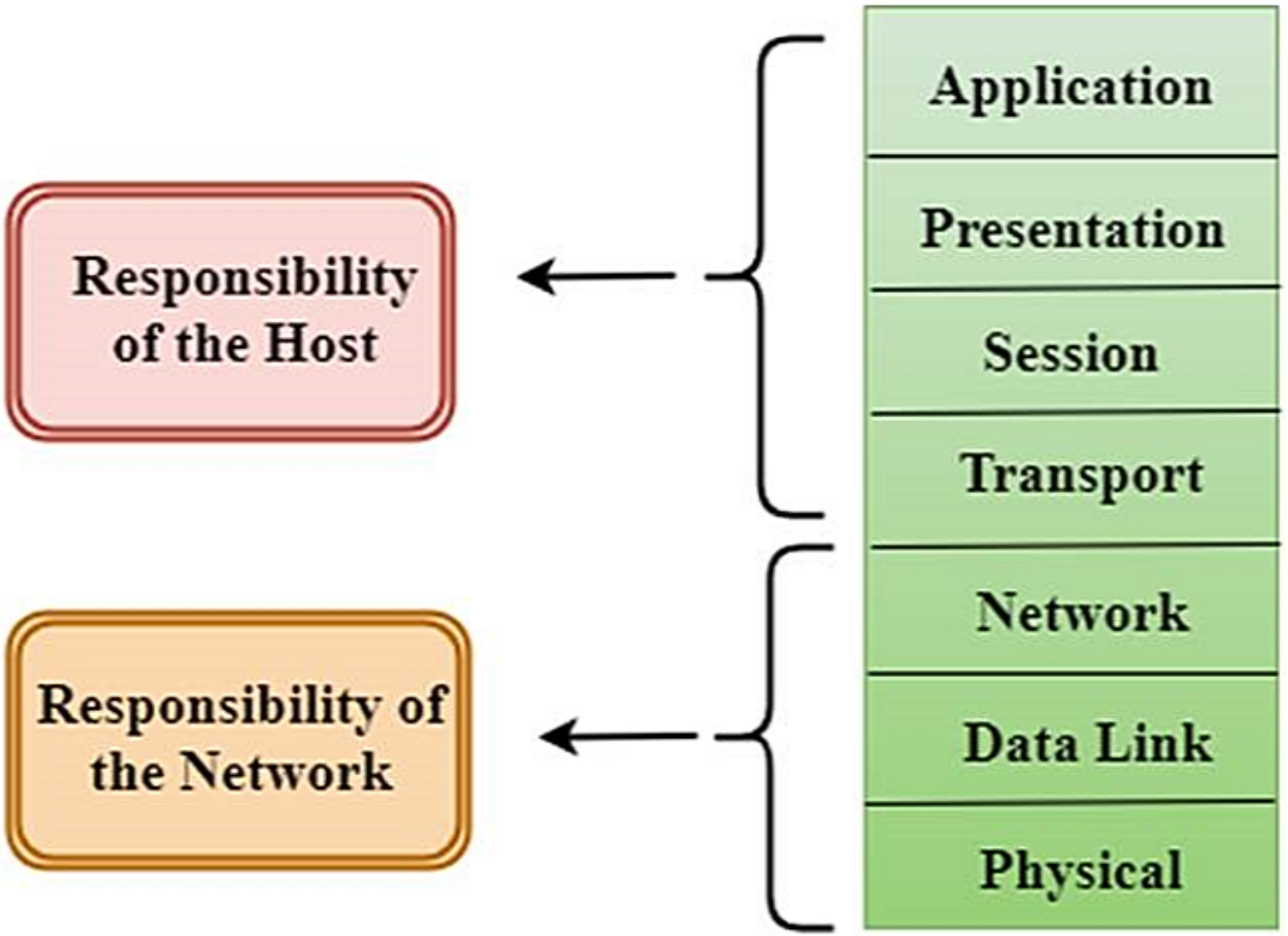


Distributed Computing: Client- Server

Introduction to TCP/IP protocol suite

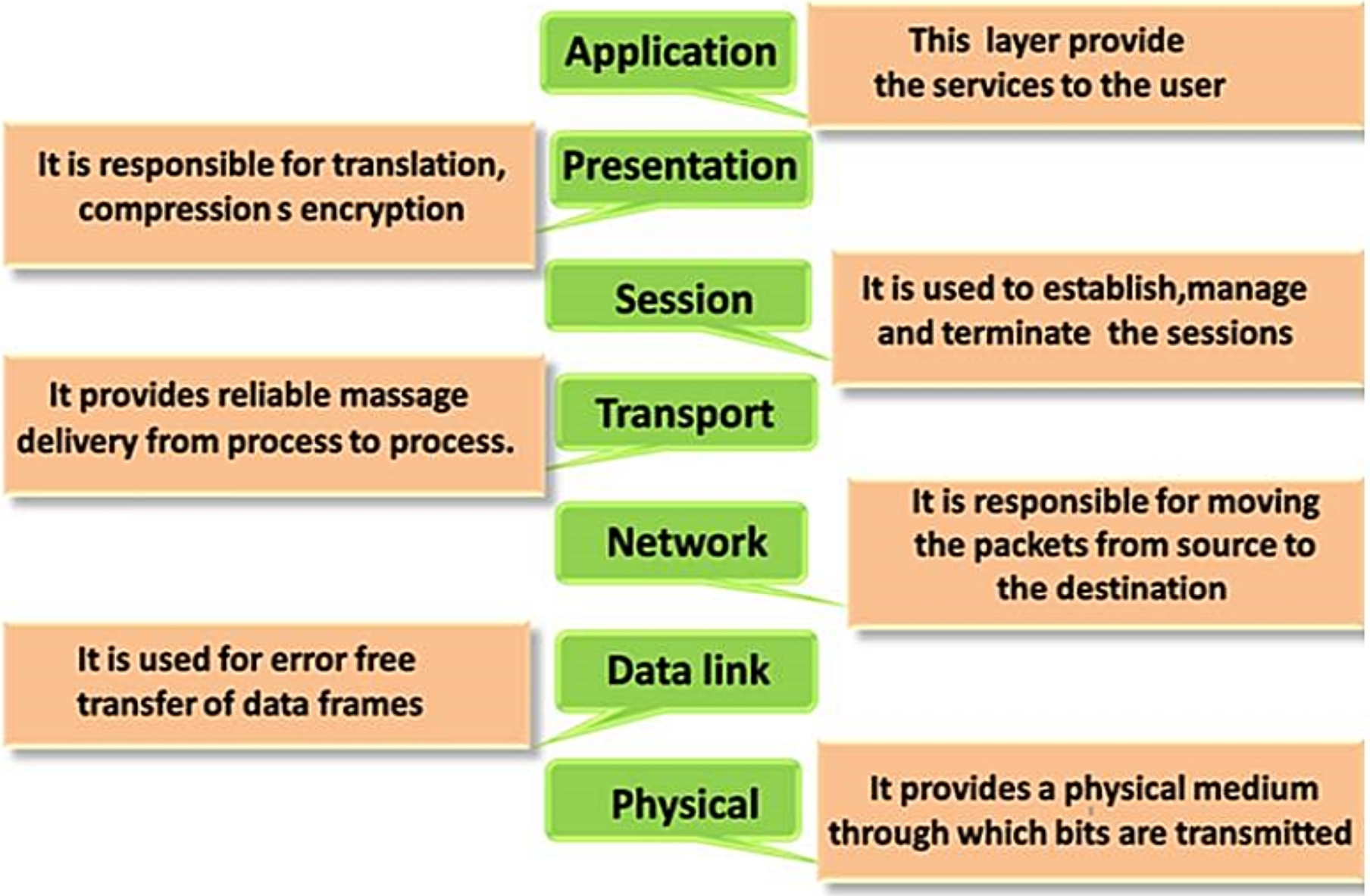
- General design Issue of protocol layers
  - Addressing
  - Error Control
  - Flow Control
  - Multiplexing
  - Routing
- TCP/IP protocol has four Layers
  -Application
  -Transport
  -Network
  -Host to Network (Data-link and Physical)

Distributed Computing: Client- Server
OSI Reference Model



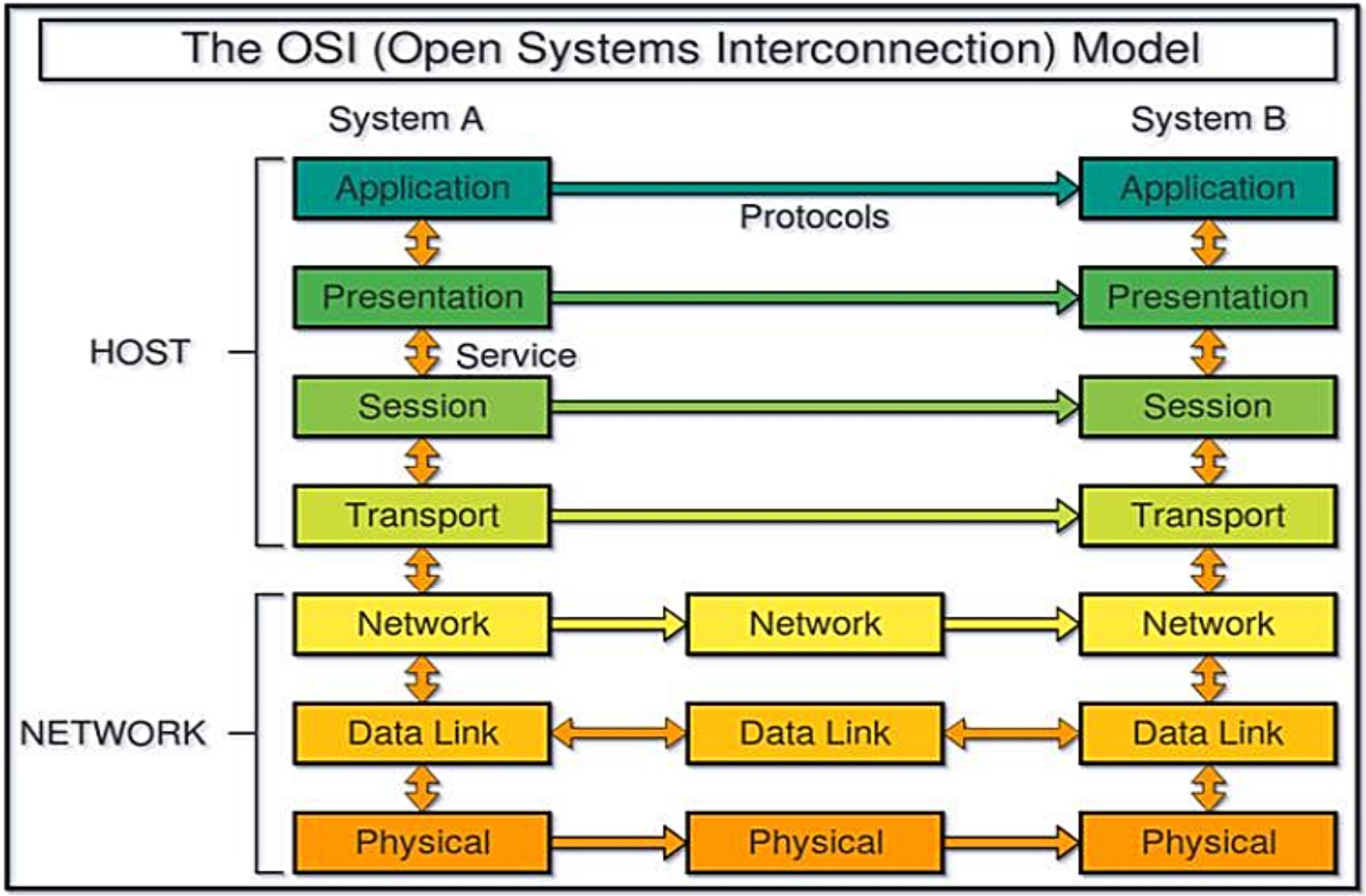
Distributed Computing: Client- Server

Functions of the OSI Lavers

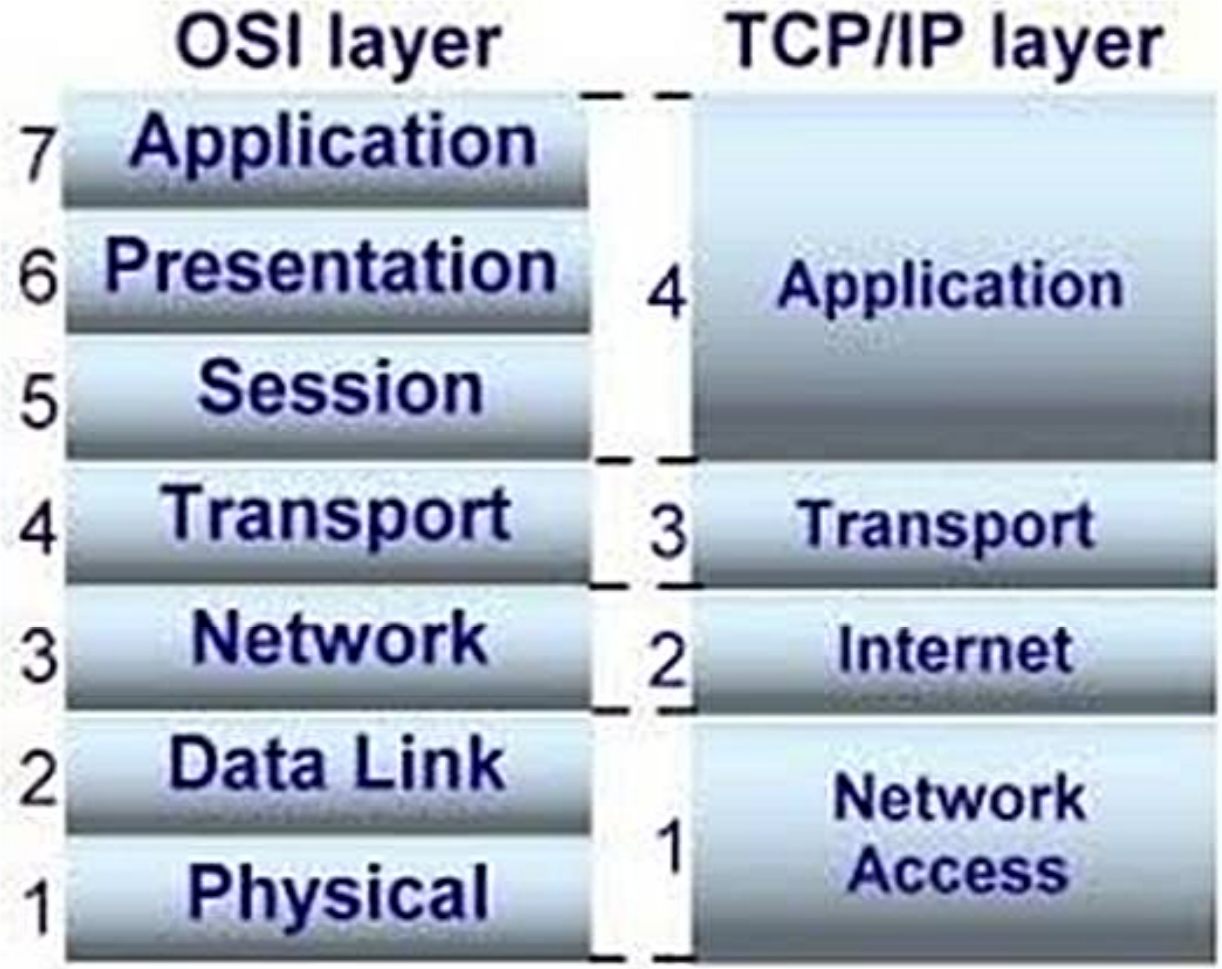


Distributed Computing: Client- Server

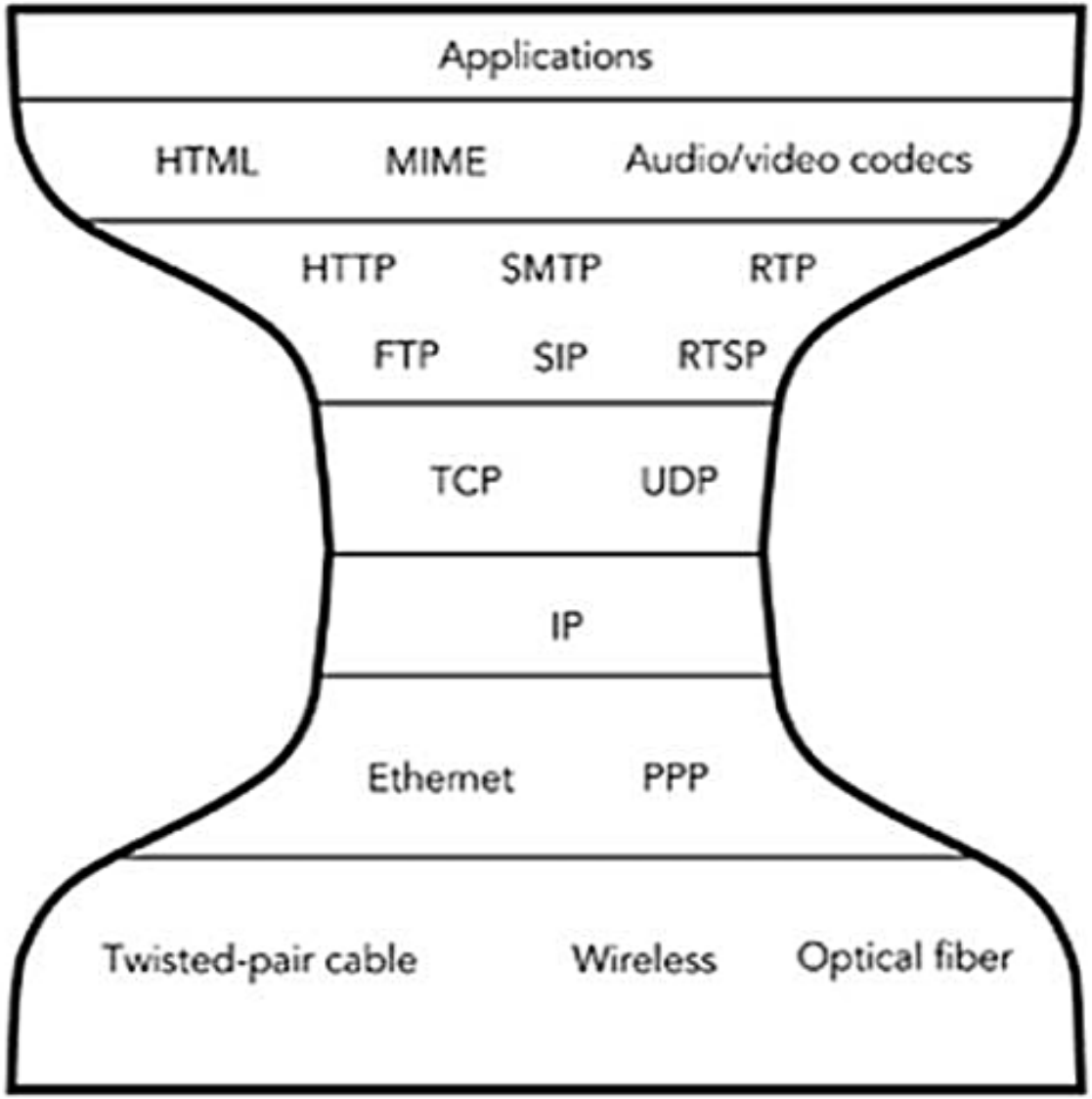
Working of the OSI Layers



Distributed Computing: Client- Server
OSI Model Vs TCP/IP



Distributed Computing: Client- Server TCP/IP Hourglass Model



Distributed Computing: Client- Server (Communication Aspects)

- There are two addresses for the communication :
- Hardware address:
  - NIC card comes with Physical address from manufacturer
  - It is 6 byte MAC address

    This address identifies a machine in a local network

- IP address :
  - It is logical host address
  - Used to perform internetwork routing.
  - 4 byte in IPv4

    Identifies a machine in a global network.

Distributed Computing: Client- Server (Communication Aspects)

- In a given network a machine can communicate with other machine by the socket:
- Socket is nothing but: (Ip-addresss : port Number).

- The IP Address can be of different classes (Class A, B ,C and D).
- It is used to identify the Machine in global network.
- Port number is of 2 Bytes and used to identify the process
- Well-Known port numbers

  0 - 1023 for standard network services and defined by the Internet Assigned Numbers Authority(IANA)

Distributed Computing: Client- Server

TCP vs UDP

TABLE 1: COMPARISON OF UDP AND TCP

	TCP	UDP
Connection	Application processes make a connection before messages can be exchanged.	Application processes exchange messages without creating a connection.
Usage	Suitable for applications that require high reliability, and transmission time is relatively less critical.	Suitable for applications that need fast, efficient transmission, and reliability is less critical.
Use by application layer protocols	File transfer (FTP), e-mail (SMTP, POP and IMAP) and Web (HTTP).	Multimedia applications (VoIP, video, online multiplayer games) and DNS (client-server communication).
Reliability	Guarantees delivery of application messages without error and in proper order.	No guarantee that messages will reach the receiving application. Furthermore, messages may arrive out of order.
Ordering of data segments	Rearrange data segments in the order specified.	Has no inherent order as all segments are independent of each other.
Acknowledgement	Segments are acknowledged when received	No acknowledgment
Flow control	Congestion-control mechanism that regulates the transport-layer sender when one or more links between the source and destination hosts become excessively congested.	UDP does not have an option for flow control.
Error checking	Erroneous segments are retransmitted from the sender to the receiver.	Erroneous segments are discarded. Error recovery is not attempted.

Distributed Computing: Client- Server Protocols Involved in communication

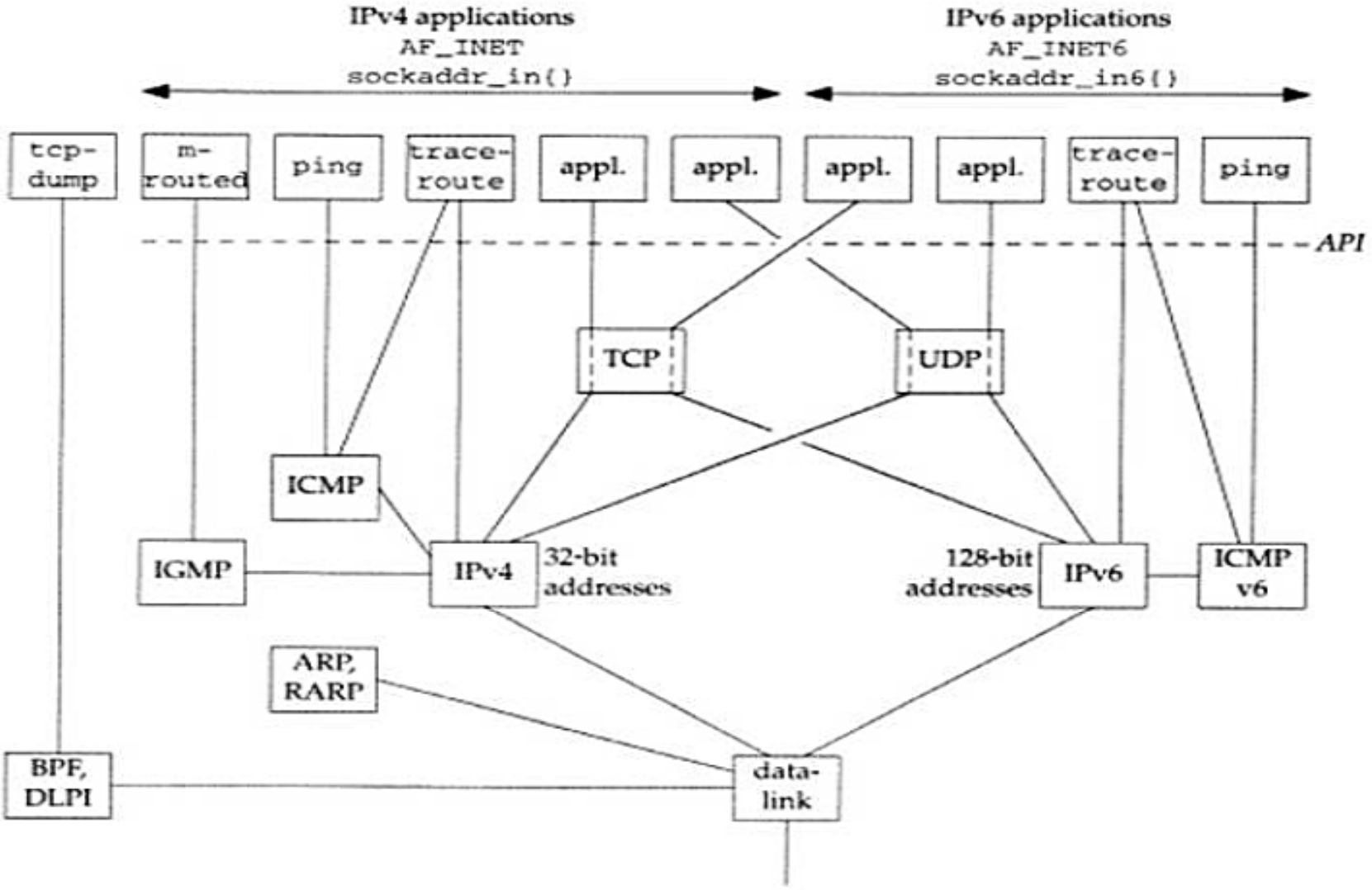


Figure 2.1 Overview of TCP/IP protocols.

Distributed Computing: Client- Server Important Protocols

- IPV4

  - Internet Protocol version 4

  - 32 bit addresses

  - It provides packet delivery service for TCP,UDP,ICMP and IGMP

- ICMP (Internet Control Message Protocol)

  - It handles error and control information between routers and hosts

  - These messages are normally generated and processed by the TCP/IP networking software itself, not by user processes.

Distributed Computing: Client- Server Important Protocols

- IGMP(Internet Group Management Protocol)

  - IGMP is used with multicasting, which is optional with IPv4.

- ARP(Address Resolution Protocol)

  - APR maps an IPv4 address into hardware address.

- RARP(Reverse Address Resolution Protocol)

  - RARP maps a hardware address into an IPv4 address.

References

- Sources :
- https://www.javatpoint.com/osi-model
- https://www.tutorialspoint.com/software_architecture_design/distributed_architecture.htm
- https://sites.google.com/site/prajapatiharshadb/class-notes-for-students
- UNIX Network Programming by W. Richard Stevens, Prentice Hall Publication
- Distributed Computing: Concepts &Applications: by M. L. Liu Addison Wiselly