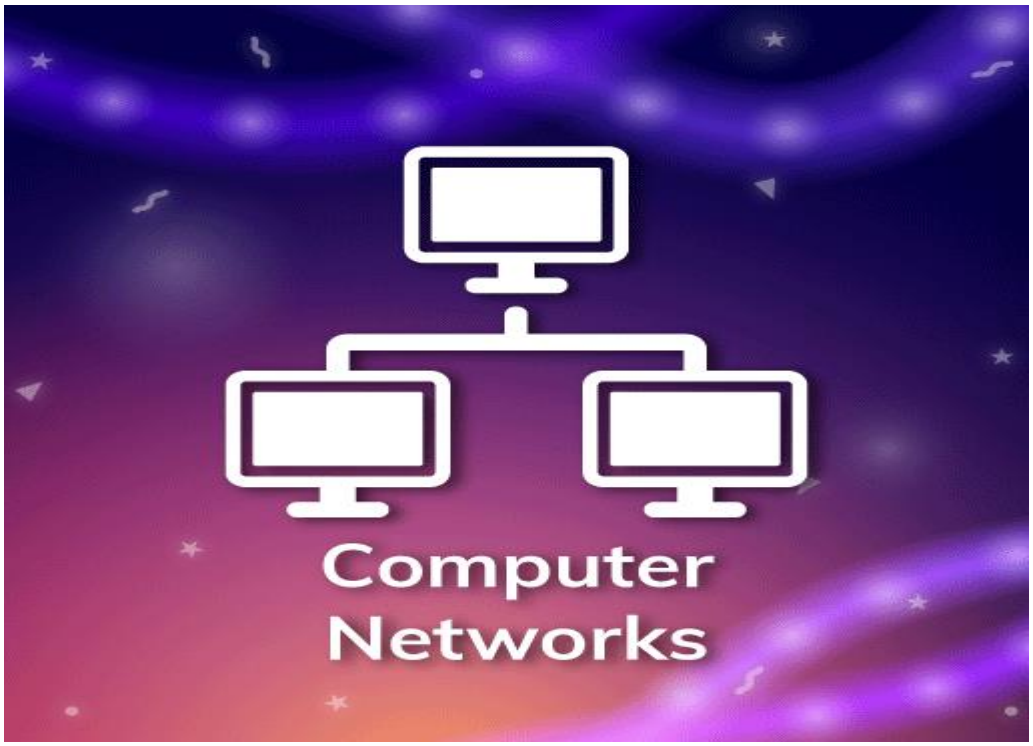
Karnataka State Open University
Mukthagangotri, Mysore-570006

B.C.A Computer science

Second Semester



COMPUTER NETWORKS

COURSE – BCADSC-2.5 **BLOCK –1-4**

With clear, straightforward text and engaging examples, this book brings an active style of learning to the study of computer networking. It can be used as a primary textbook to accompany lecture material or as a companion text to provide hands-on assignments. It is also an ideal guide to self-study for the computing professional. In fact, this book can help any interested readers look under the hood of the network they use every day and become a more informed network consumer.

The book consists of a set of exercises, each of which involves analyzing traces of actual network activity. Important concepts are presented in the context of actual traces of real-world scenarios. Readers learn the details of networking protocols in the best ways possible—by seeing them in action!

thing about the Internet is constant change. Despite these changes, the question we asked in the first edition is just as valid today: What are the underlying concepts and technologies that make the Internet work? The answer is that much of the TCP/IP architecture continues to function just as was envisioned by its creators more than 30 years ago. This isn't to say that the Internet architecture is uninteresting; quite the contrary. Understanding the design principles that underly an architecture that has not only survived but fostered the kind of growth and change that the Internet has seen over the past 3 decades is precisely the right place to start.

CREDIT PAGE

Programme Name: BCA-Computer Science **Year/Semester: II** Semester Block **No:1-4**

Course Name: Computer Networks **Credit:** 4 U**nit No:** 1-16

Course Design Expert Committee

Dr. Sharanappa V Halase Chairman
Vice Chancellor,
Karnataka State Open University,
Mukthagangotri, Mysuru-570006.

Dr. Ashok Kamble Member
Dean (Academic),
Karnataka State Open University,
Mukthagangotri, Mysuru-570 006.

Editorial Committee

Dr. Mahesha D.M., MCA.,PhD Chairman
BOS Chairman,
Assistant Professor & Programme co-ordinator(PG)
DoS&R in Computer Science,
Karnataka State Open University, Mysuru-570 006.

Smt, Suneetha MSc.,(PhD) Member Convener
Dept Chairperson & Programme co-ordinator (UG)
DoS&R in Computer Science,
Karnataka State Open University, Mysuru-570 006.

Dr Bhavya D.N., MTech., PhD Member
Assistant Professor & Programme co-ordinator(UG)
DoS&R in Computer Science,
Karnataka State Open University, Mysuru-570 006.

Dr. Ashoka S B., MSc.PhD Member
External Subject Expert,
Assistant Professor,
DoS&R in Computer Science,
Maharani's Cluster University, palace Road Bangalore-01

Name of Course Writer	No of Blocks	No of Units	Name of Course Editor	No of Units
Dr Ashoka S.B Assistant professor DOS in Computer Science Maharni Cluster University Palce Road Bangalore-560001	Block-1	1-4	**Dr. Sharthkumar Y.H** Professor IS MIT Belawadi Srirangapatna tq Mandya District -571477	1- 4

Dr. Poornima Y Asst, Professor Dos in Computer Science Maharaja's College- 5700021	Block-2	5-8	**Sri. Pradeepa H. G** Asst. Professor DOS in Computer Science JSS women's College Saraswathipuram, Mysore	5-8
Dr.Basappa B Kodada Associate professor Department of Computer ScienceAJ Institute of Engineering and Technology Mangalore	Block-3	9-12	**Sri. Somashekar BM** BE., MTech., (PhD) Assistant Professor, Department of Information Science MIT Naguvanhalli Srirangapatna Taluk 571438	9-12
Smt. Sumanashree.Y.S Chairperson DOS in Computer Science (PG)JSS college of Arts, Science, Commerce OOty Road Mysore	Block-4	13-16	**Smt. Vasanthi** Assistant professor DOS in Computer Science MIT Degree College Vishweshwaranagar Mysore	13-16

Copy Right

Registrar,
Karnataka State Open University,
Mukthagantoghri, Mysore 570 006.

Developed by the Department of Studies and Research in Information Technology, under the guidance of Dean (Academic), KSOU, Mysuru.
Karnataka State Open University, February-2022.
All rights reserved. No part of this work may be reproduced in any form or any other means, without permission in writing from the Karnataka State Open University.
Further information on the Karnataka State Open University Programmes may be obtained from the University's Office at Mukthagangotri, Mysore – 570 006.

Printed and Published on behalf of Karnataka State Open University, Mysore-570 006 by the **Registrar (Administration)-2023**

TABLE OF CONTENTS		
BLOCK 1		**PAGE NO.**
UNIT- 1	Data Communication, Component and Basic Concepts Introduction, Characteristics delivery, Accuracy, Timeliness and Jitter, Components	1-8
UNIT-2	Message, Sender, Receiver, Transmission medium and protocol. Topology – Mesh, Star, Tree, Bus, Ring and Hybrid Topologies.	9-25
UNIT-3	Transmission modes – Simplex, Half Duplex, Full Duplex Categories of networks –LAN, MAN, WAN. DNS, IP address, MAC address,	26-52
UNIT-4	Web browser, ISP, URL, WWW,Broadband Transmissions. Guided Media – TwistedPair Cable, Coaxial Cable, Fiber-Optic Cable.	53-72
BLOCK 1I		
UNIT-5	Unguided Media – Radio Wave Transmission Systems, Microwave Transmission Systems, Infrared Transmission Systems and Satellite Communication System.	73-78
UNIT-6	The OSI Model – Functions of all the Seven Layers. Networking Devices – Functions and Applications of Hub, Switches, Bridges, Repeaters	79-89
UNIT-7	Internetworking Devices – Functions and Applications of Routers and Gateways.IP Addressing – Dynamic IP Addressing, Static IP Addressing,	90-97
UNIT-8	Types of IP Addresses. Protocols – Overview only- TCP, UDP, IP, IPV4, IPV6,	98-114
BLOCK 1II		
UNIT- 9	TCP/IP Suite, SMTP, POP3, SNMP, HTTP, FTP, DNS, ICMP	115-132
UNIT-10	IGMP, ARP, RARP, OSPF, BGP, ALOHA	133-142
UNIT-11	Packet Switching Networks – Network Services and Internal Network Operations,	143-151
UNIT-12	Packet Network Topology, Datagrams and Virtual Circuits, Connectionless Packet Switching,	152-166
BLOCK 1V		
UNIT-13	Virtual Circuit Packet Switching. Routing Concepts – Routing Tables	167-176
UNIT-14	Dijkstar's Shortest Path Routing Algorithm, Congestion Control Algorithms-Leaky Bucket Algorithm.	177-186
UNIT-15	Data Link Issues –Single bit error and Burst Error, concepts of Redundancy, Checksum	187-197
UNIT-16	Single Bit Error correction and Hamming Code correction method.	198-209

UNIT 1: COMPUTER NETWORKS

Structure:

1.0 Objectives
1.1 Introduction
1.2 Data communication Components and basic concepts
1.4 Characteristic delivery, Accuracy, Timelineness and jitter
1.7 Summary
1.8 Keywords
1.9 Questions for self-study
1.10 References

1.0 OBJECTIVES

At the end of this unit you will be able to:

- Discuss data communication
- Explain the component and basic concepts
- State the advantages accuracy

1.1 INTRODUCTION

The Network is a group of computers connected together to share resources such as programs, files, printer, and storage disk. However modern network also includes connections to portable/mobile devices (such as tablet PC, notebook, PDA, digital camera, portable MP3 player, and mobile phone), home entertainment devices (such as TV, video player/recorder, stereo, and radio), home appliances (such as refrigerator and washing machine), and monitoring or sensor devices. Wearable things (such as wristwatch, clothing) and perhaps living objects like human, pet, and tree will be network-able in the not-so-distant future.

Network design is not a hard-to-grasp science, yet it had been only the job of IT professionals or network specialists before because it was only about big company networks with complex wiring and many computers. Nowadays network design comes down to consumer realm because of the growing popularity of wireless ad-hoc networks using IrDA and Bluetooth and the prevailing home networking technologies such as Ethernet, Wi-Fi, HomePNA, and HomePlug. New technologies such as ZigBee and WiMedia (UWB) will give even more connectivity choices for consumer networks.

- **Advantages of computer networks**

  These purposes must be fulfilled by various advantages of networks.

  - **Resource Sharing**

Resource sharing means the goal is to make all programs, data and equipment available to anyone on the network without regard to the physical location of the resource and the user.

Example: Suppose a user happens to be 1000 km away from his data should not prevent him from using the data as though they were local. Also load sharing is another aspect of resource sharing.

  - **High Reliability**

Network provides high reliability by having alternative sources of supply.

Example: Suppose all files could be replicated on two or three machines, so if one of them is unavailable (due to a hardware failure), the other copies could be used. For

military, banking, air traffic control, and other applications, the ability to continue operating the face of hardware problems is of great importance.

- **Low Cost/Saving Money**

Small computers have a much better price/performance ratio than large one. Mainframes are roughly a factor of fourty faster than the fastest single chip microprocessors, but they cost a thousand times more. This imbalance has caused many system designers to build systems consisting of powerful personal computers, as per user, with data kept on one or more shared file server machines.

- **Communications**

Another goal of setting up a computer network has little to do with technology at all. A computer network can provide a powerful communication medium among widely separated people. Using a network, it is easy for two or more people who live far apart to write a report together. i.e. when one author makes a change to the document, which is kept online, the others can see the change immediately, instead of waiting several days for a letter.

- **Uses of Computer Networks**

1. Access to remote programs: A company that has produced a model simulating the world economy may allow its clients to log in over the network and run the program to see how various projected inflation rates, interest rates, and currency fluctuations might affect their business. This approach is often preferable to selling the program outright, especially if the model is constantly being adjusted or requires an extremely large mainframe computer to run.

2. Access to remote data bases: It may soon be easy for the average person sitting at home to make reservations for aero planes, trains, buses, boats, hotels, restaurants, theatres and so on, anywhere in the world with instant confirmation. Home banking and the automated newspaper also fall in this category.

3. Value-added communication facilities: High-quality communication facilities tend to reduce the need for physical proximity. Everyone in the world, have an ability to

send and receive electronic mail. These mails are also be able to contain digitized voice, still pictures and possibly even moving television and video images.

4. using for entertainment purpose.

5. Accessing the information systems like World Wide Web, this contains almost any information.

1.2 DATA COMMUNICATION

Data communications refers to the transmission of this digital data between two or more computers and a computer network or data network is a telecommunications network that allows computers to exchange data. The physical connection between networked computing devices is established using either cable media or wireless media. The best-known computer network is the Internet. This unit should teach you basics of Data Communication and Computer Network (DCN) and will also take you through various advance concepts related to Data Communication and Computer Network.

A system of interconnected computers and computerized peripherals such as printers is called computer network. This interconnection among computers facilitates information sharing among them. Computers may connect to each other by either wired or wireless media.

- **Network Engineering**

Networking engineering is a complicated task, which involves software, firmware, chip level engineering, hardware, and electric pulses. To ease network engineering, the whole networking concept is divided into multiple layers. Each layer is involved in some particular task and is independent of all other layers. But as a whole, almost all networking tasks depend on all of these layers. Layers share data between them and they depend on each other only to take input and send output.

- **Internet**

A network of networks is called an internetwork, or simply the internet. It is the largest network in existence on this planet. The internet hugely connects all WANs and it can have connection to LANs and Home networks. Internet uses TCP/IP protocol suite and

uses IP as its addressing protocol. Present day, Internet is widely implemented using IPv4. Because of shortage of address spaces, it is gradually migrating from IPv4 to IPv6. Internet enables its users to share and access enormous amount of information worldwide. It uses WWW, FTP, email services, audio and video streaming etc. At huge level, internet works on Client-Server model. Internet uses very high speed backbone of fiber optics. To inter-connect various continents, fibers are laid under sea known to us as submarine communication cable.

- **Applications of Communication & Computer Network**

Computer systems and peripherals are connected to form a network. They provide numerous advantages:

- Resource sharing such as printers and storage devices
- Exchange of information by means of e-Mails and FTP
- Information sharing by using Web or Internet
- Interaction with other users using dynamic web pages
- IP phones
- Video conferences
- Parallel computing
- Instant messaging

- **A data communications system has five components**

- **Message:** The message is the information/data to be communicated. Popular forms of information include text, numbers, pictures, audio, and video.
- **Transmitter:** The transmitter is the device that sends the data. It can be a computer, workstation, telephone handset, video camera, and so on.

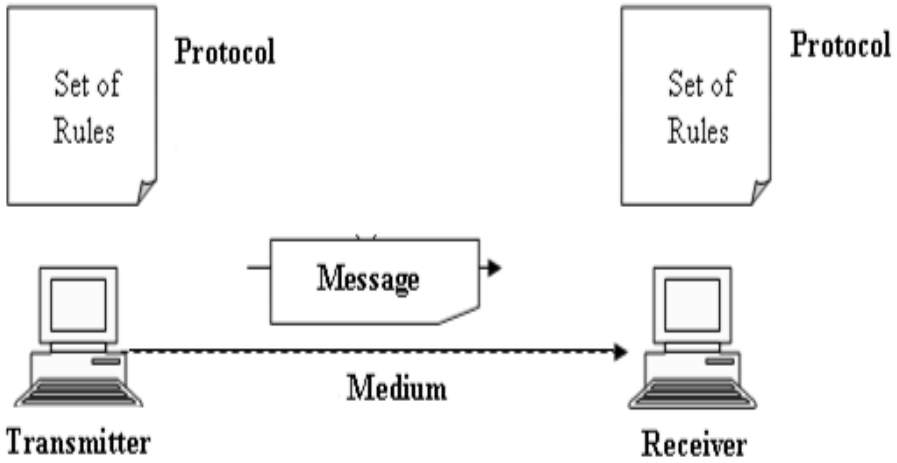


Figure 1.3: Five components of data communication

- **Receiver:** The receiver is the device that receives the data. It can be a computer, workstation, telephone handset, television, and so on.
- **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.
- **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.

- **Data Representation:**

  Nowadays information comes in different forms such as text, numbers, images, audio, and video.

  - **Text**: Text is represented as a bit pattern, a sequence of bits 0s or 1s. A set of bit patterns that is designed to represent text symbols is called a code; the process of representing symbols is called coding. The prevalent coding system called Unicode uses 32 bits to represent a symbol or character. The American Standard Code for Information Interchange (ASCII) now constitutes the first 127 characters in Unicode and is also referred to as Basic Latin.
  - **Numbers**: Numbers are also represented by bit patterns. However, a code such as ASCII is not used to represent numbers; the number is directly converted to a binary number to simplify mathematical operations.
  - **Images**: Images are also represented by bit patterns. In its simplest form, an image is composed of a matrix of pixels (picture elements), where each pixel is a

small dot. For a black and white image, a 1-bit pattern is enough to represent a pixel. In the RGB method to represent color image each color is made of a combination of three primary colors: *red,* green, and blue. Another method is called YCM, in which a color is made of a combination of three other primary colors: yellow, cyan, and magenta.

- **Audio**: Audio refers to the recording or broadcasting of sound or music. Audio is by nature different from text, numbers, or images. It is continuous, not discrete. Even when we use a microphone to change voice or music to an electric signal, we create a continuous signal.
- **Video**: Video refers to the recording or broadcasting of a picture or movie. Video can either be produced as a continuous entity (e.g., by a TV camera), or it can be a combination of images, each a discrete entity, arranged to convey the idea of motion. We can change video to a digital or an analog signal.

1.3 CHRECTERSTIC DELIVARY, ACCURACY, TIMELINENESS, JITTER

The word data refers to information presented in whatever form is agreed upon by the parties creating and using the data. Data communications are the exchange of data between two devices via some form of transmission medium such as a wire cable. For data communications to occur, the communicating devices must be part of a communication system made up of a combination of hardware (physical equipment) and software (programs). The effectiveness of a data communications system depends on four fundamental characteristics: delivery, accuracy, timeliness, and jitter.

I. Delivery- The system must deliver data to the correct destination. Data must be received by the intended device or user and only by that device or user.

2. Accuracy- The system must deliver the data accurately. Data that have been altered in transmission and left uncorrected are unusable.

3. Timeliness- The system must deliver data in a timely manner. Data delivered late are useless. In the case of video and audio, timely delivery means delivering data as they are produced, in the same order that they are produced, and without significant delay. This kind of delivery is called real-time transmission.

4. Jitter- Jitter refers to the variation in the packet arrival time. It is the uneven delay in the delivery of audio or video packets. For example, let us assume that video packets are sent every 3D ms. If some of the packets arrive with 3D-ms delay and others with 4D-ms delay, an uneven quality in the video is the result.

1.4 SUMMARY

This unit introduces internetworking concept and its application in the modern era internet and communication, later it describes the data communication and its component and their basic concepts, finally you can discuss the characteristic delivery and its accuracy and timeliness, jitter.

1.5 KEYWORDS

Simplex
Half-Duplex
Full-Duplex
Transmitter

1.6 QUESTION FOR SELF STUDY

1. What is Computer Networks? Explain in detail.
2. Discuss in detail about data communication network.
3. Explain Component and basic concepts of Networks.
4. Discuss characteristic delivery, accuracy, timelineness, jitter.

1.8 REFERENCE

1. Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.
2. Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.
3. Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT 2: TRANSMISSION MEDIUM

Structure:

2.0 Objectives

2.1 Introduction

2.2 Message, Sender, Receiver,

2.3 Transmission medium and protocol.

2.4 Topology

- 2.4.1 Mesh
- 2.4.2 Star topology
- 2.4.3 Tree topology
- 2.4.4 Bus topology
- 2.4.5 Ring topology
- 2.4.6 Hybrid Topologies.

2.7 Summary

2.8 Keywords

2.9 Questions for self-study

2.10 References

2.0 OBJECTIVES

At the end of this unit you will be able to:

- Understand Message, Sender, Receiver.
- Discuss the Transmission medium and protocol.
- Explain the Mesh, Star, Tree, Bus, Ring and Hybrid Topologies.

2.1 INTRODUCTION

The representation of data can be done through computers as well as other types of telecommunication devices with the help of signals. These are broadcasted from one device to another in the shape of electromagnetic energy. The signals like electromagnetic can travel throughout vacuum, air otherwise other transmission mediums to travel from one sender to another receiver. Electromagnetic energy mainly includes voice, power, radio waves, visible light, UV light, & gamma rays. In the OSI model, the first layer is the physical layer which is dedicated to the transmission media. In data communication, a transmission media is a physical lane between the Tx & the Rx and it is the channel where data can be transmitted from one area to another.

- **What is Transmission Media**

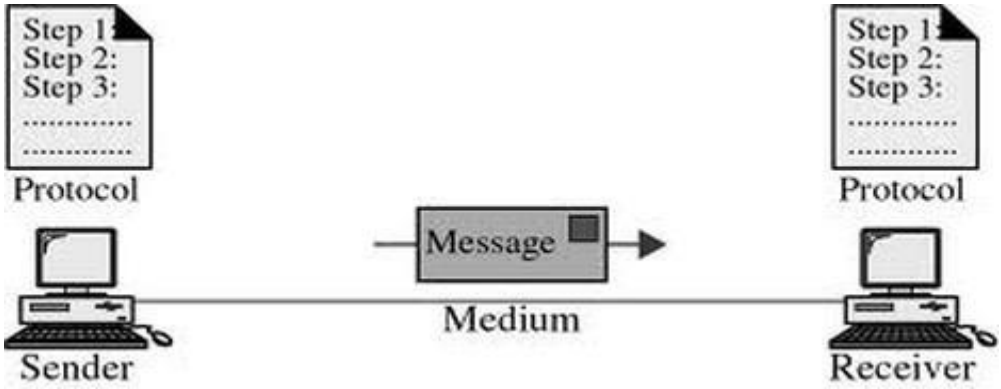
A communication channel that is used to carry the data from the transmitter to the receiver through the electromagnetic signals. The main function of this is to carry the data in the bits form through the Local Area Network (LAN). In data communication, it works like a physical path between the sender & the receiver. For instance, in a copper cable network the bits in the form of electrical signals whereas in a fiber network, the bits are available in the form of light pulses. The quality, as well as characteristics of data transmission, can be determined from the characteristics of medium & signal. The properties of different transmission media are delay, bandwidth, maintenance, cost, and easy installation.

2.2 MESSAGE, SENDER, RECEIVER

Basic Components of a Communication System

The following are the basic requirements for working of a communication system.

1. The sender (source) who creates the message to be transmitted
2. A medium that carries the message
3. The receiver (sink) who receives the message



Data Communication system components

1. **Message:** A message in its most general meaning is an object of communication. It is a vessel which provides information. Yet, it can also be this information. Therefore, its meaning is dependent upon the context in which it is used; the term may apply to both the information and its form.

   2. Sender: The sender will have some kind of meaning she wishes to convey to the receiver. It might not be conscious knowledge; it might be a sub-conscious wish for communication. What is desired to be communicated would be some kind of idea, perception, feeling, or datum. It will be a part of her reality that she wishes to send to somebody else.

   3. Receiver: These messages are delivered to another party. No doubt, you have in mind a desired action or reaction you hope your message prompts from the opposite party. Keep in mind, the other party also enters into the communication process with ideas and feelings that will undoubtedly influence their understanding of your message and their response. To be a successful communicator, you should consider these before delivering your message, then acting appropriately.

   4. Medium: Medium is a means used to exchange / transmit the message. The sender must choose an appropriate medium for transmitting the message else the message might not be conveyed to the desired recipients. The choice of appropriate medium of communication is essential for making the message effective and correctly interpreted by the recipient. This choice of communication medium varies depending upon the features of communication. For instance - Written medium is chosen when a message

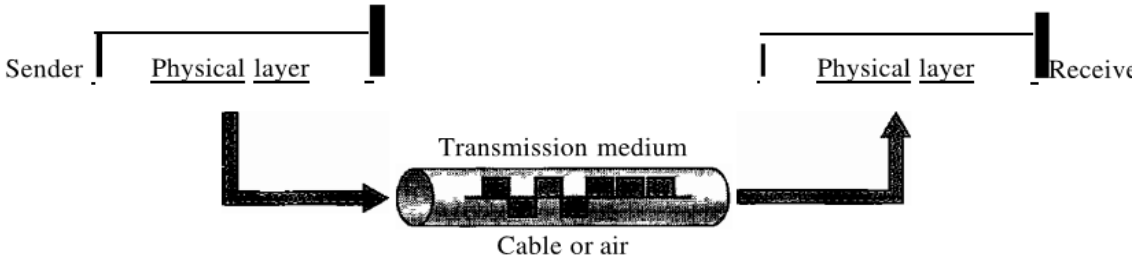
has to be conveyed to a small group of people, while an oral medium is chosen when spontaneous feedback is required from the recipient as misunderstandings are cleared then and there.

5. Protocol: A protocol is a formal description of digital message formats and the rules for exchanging those messages in or between computing systems and in telecommunications. Protocols may include signalling, authentication and error detection and correction syntax, semantics, and synchronization of communication and may be implemented in hardware or software, or both.

6. Feedback: Feedback is the main component of communication process as it permits the sender to analyze the efficacy of the message. It helps the sender in confirming the correct interpretation of message by the decoder. Feedback may be verbal (through words) or non-verbal (in form of smiles, sighs, etc.). It may take written form also in form of memos, reports, etc.

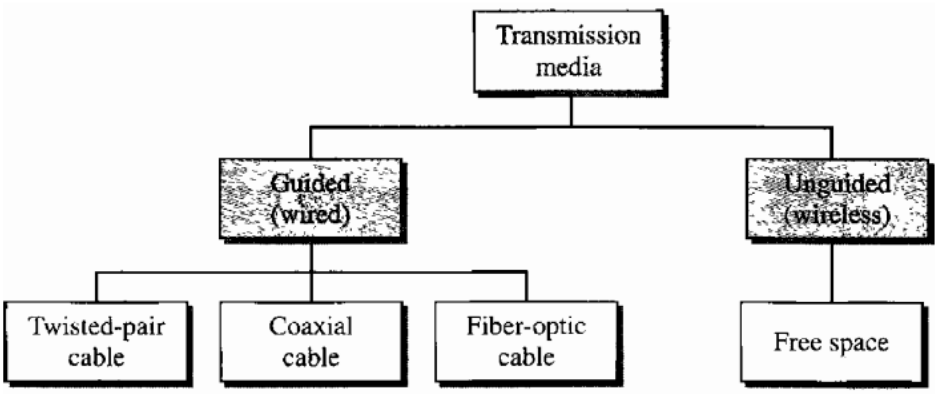
2.3 TRANSMISSION MEDIUM AND PROTOCOL.

A transmission medium can be broadly defined as anything that can carry information from a source to a destination. For example, the transmission medium for two people having a dinner conversation is the air. The air can also be used to convey the message in a smoke signal or semaphore. For a written message, the transmission medium might be a mail carrier, a truck, or an airplane. In data communications the definition of the information and the transmission medium is more specific. The transmission medium is usually free space, metallic cable, or fiber-optic cable. The information is usually a signal that is the result of a conversion of data from another form.



In telecommunications, transmission media can be divided into two broad categories:

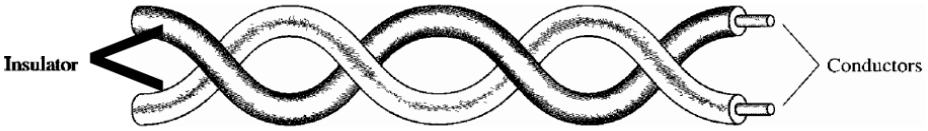
Guided and unguided. Guided media include twisted-pair cable, coaxial cable, and fiber-optic cable. Unguided medium is free space.



Guided media, which are those that provide a conduit from one device to another, include twisted-pair cable, coaxial cable, and fiber-optic cable. A signal travelling along any of these media is directed and contained by the physical limits of the medium. Twisted-pair and coaxial cable use metallic (copper) conductors that accept and transport signals in the form of electric current. Optical fiber is a cable that accepts and transports signals in the form of light.

- **Twisted-Pair Cable**

A twisted pair consists of two conductors (normally copper), each with its own plastic insulation, twisted together.

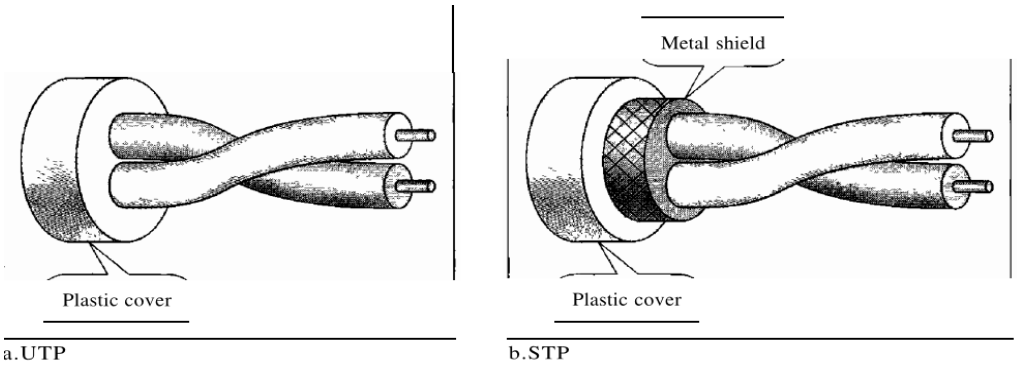


One of the wires is used to carry signals to the receiver, and the other is used only as a ground reference. The receiver uses the difference between the two. In addition to the signal sent by the sender on one of the wires, interference (noise) and crosstalk may affect both wires and create unwanted signals. If the two wires are parallel, the effect of these unwanted signals is not the same in both wires because they are at different locations relative to the noise or crosstalk sources (e,g., one is closer and the other is farther). This results in a difference at the receiver. By twisting the pairs, a balance is maintained. For example, suppose in one twist, one wire is closer to the noise source and

the other is farther; in the next twist, the reverse is true. Twisting makes it probable that both wires are equally affected by external influences (noise or crosstalk). This means that the receiver, which calculates the difference between the two, receives no unwanted signals. The unwanted signals are mostly cancelled out. From the above discussion, it is clear that the number of twists per unit of length (e.g., inch) has some effect on the quality of the cable.

- **Unshielded Versus Shielded Twisted-Pair Cable**

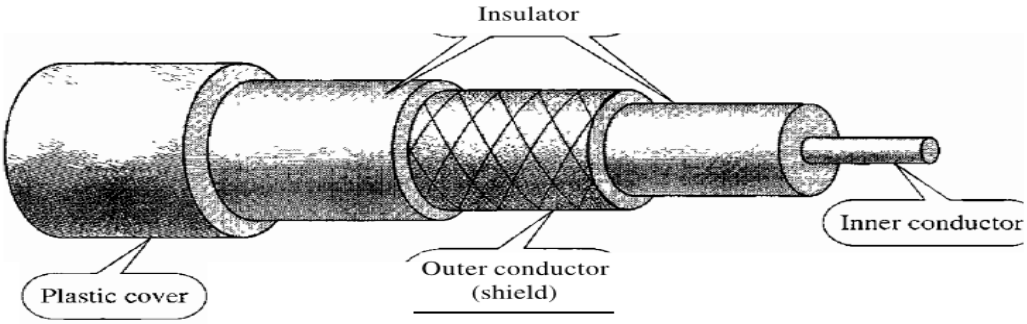
The most common twisted-pair cable used in communications is referred to as unshielded twisted-pair (UTP). IBM has also produced a version of twisted-pair cable for its use called shielded twisted-pair (STP). STP cable has a metal foil or casing improves the quality of cable by preventing the penetration of noise or crosstalk, it is bulkier and more expensive.



a.UTP b.STP

- **Coaxial Cable**

Coaxial cable (or *coax)* carries signals of higher frequency ranges than those in twisted pair cable, in part because the two media are constructed quite differently. Instead of having two wires, coax has a central core conductor of solid or stranded wire (usually copper) enclosed in an insulating sheath, which is, in turn, encased in an outer conductor of metal foil, braid, or a combination of the two. The outer metallic wrapping serves both as a shield against noise and as the second conductor, which completes the circuit. This outer conductor is also enclosed in an insulating sheath, and the whole cable is protected by a plastic cover.



- **Coaxial Cable Standards**

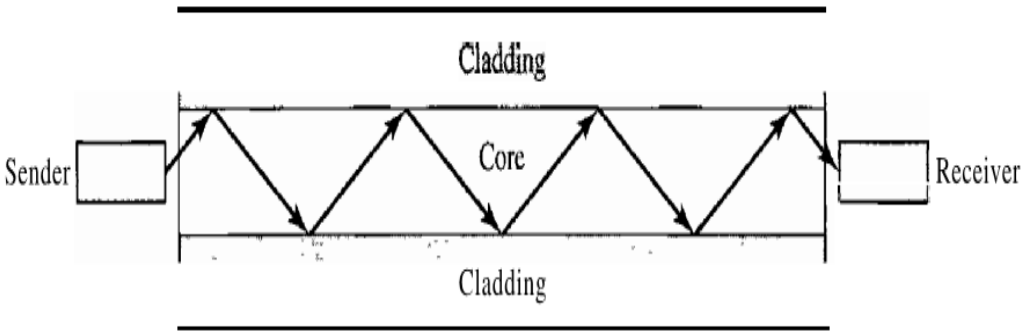
Coaxial cables are categorized by their radio government (RG) ratings. Each RG number denotes a unique set of physical specifications, including the wire gauge of the inner conductor, the thickness and type of the inner insulator, the construction of the shield, and the size and type of the outer casing. Each cable defined by an RG rating is adapted for a specialized function, as shown in Table

Category	*Impedance*	*Use*
RG-59	75 Ω	Cable TV
RG-58	50 Ω	Thin Ethernet
RG-11	50 Ω	Thick Ethernet

- **Fiber-Optic Cable**

A fiber-optic cable is made of glass or plastic and transmits signals in the form of light. To understand optical fiber, we first need to explore several aspects of the nature of light. Light travels in a straight line as long as it is moving through a single uniform substance. If a ray of light traveling through one substance suddenly enters another substance (of a different density), the ray changes direction.

Optical fibres use reflection to guide light through a channel. A glass or plastic core is surrounded by a cladding of less dense glass or plastic. The difference in density of the two materials must be such that a beam of light moving through the core is reflected off the cladding instead of being refracted into it.



- **Protocols:**

A protocol is a set of rules that govern data communications. A protocol defines what is communicated, how it is communicated, and when it is communicated. The key elements of a protocol are syntax, semantics, and timing of Syntax. The term *syntax* refers to the structure or format of the data, meaning the order in which they are presented.

The word *semantics* refers to the meaning of each section of bits. How a particular pattern is to be interpreted, and what action is to be taken based on that interpretation? For example, does an address identify the route to be taken or the final destination of the message?

The term *timing* refers to two characteristics: when data should be sent and how fast they can be sent. For example, if a sender produces data at 100 Mbps but the receiver can process data at only 1 Mbps, the transmission will overload the receiver and some data will be lost.

2.4 TOPOLOGY MESH, STAR, TREE, BUS, RING, HYBRID TOPOLOGIES

The term "TOPOLOGY" refers to the way in which the end points or stations/computer systems, attached to the networks, are interconnected. We have seen that a topology is essentially a stable geometric arrangement of computers in a network. If you want to select a topology for doing networking. You have attention to the following points.

- Application S/W and protocols.
- Types of data communicating devices.
- Geographic scope of the network.
- Cost.
- Reliability.

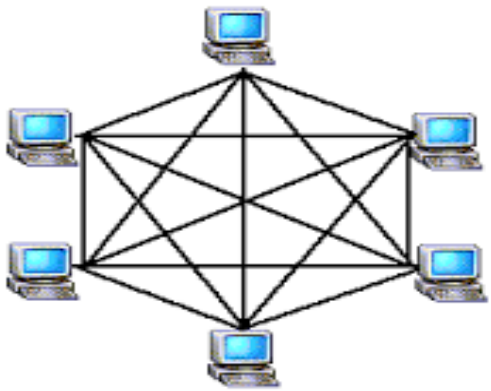
Depending on the requirement there are different Topologies to construct a network.

- **Mesh topology.**
- **Star topology.**
- **Tree topology.**
- **Bus topology.**
- **Ring topology.**

- Ring and mesh topologies are felt convenient for peer to peer transmission.
- Star and tree are more convenient for client server.
- Bus topology is equally convenient for either of them.

2.4.1 Mesh Topology.

In a mesh topology, every device has a dedicated point-to-point link to every other device. The term *dedicated* means that the link carries traffic only between the two devices it connects. In a fully connected mesh network with *n* nodes, the number of physical links is n(n-1). However, if each physical link allows communication in both directions (duplex mode), we need *n (n* -1) /2 duplex-mode links. To accommodate that many links, every device on the network must have *n* – 1 input/output port to be connected to the other *n* - 1 stations. One practical example of a mesh topology is the connection of telephone regional offices in which each regional office needs to be connected to every other regional office.



2.4.1 Mesh **Topology**

- **Advantages of mesh topology:**

A mesh offers several advantages over other network topologies. they are

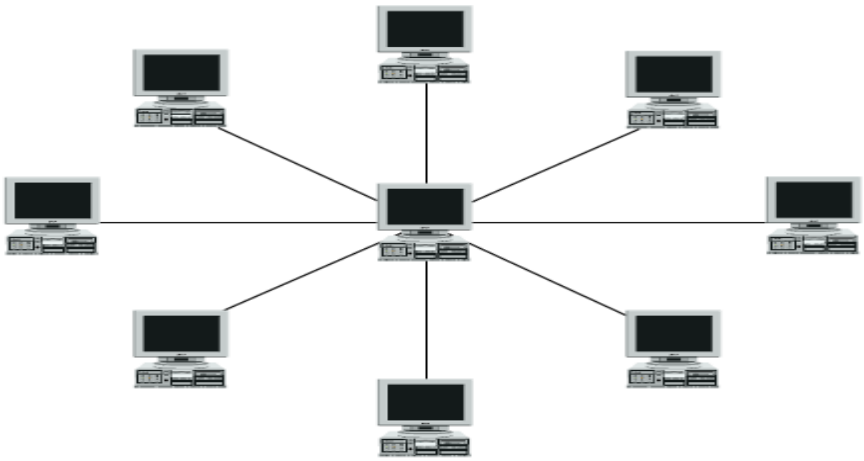
- The network has multiple links, so if one route is blocked then other routes can be used for data communication. Thus the topology is robust.
- The problem of traffic does not arise, as there are dedicated links that guarantees that each connection can carry its own data.
- Point to point links makes fault identification easy. Traffic can be routed to avoid links with suspected problems. This enables to discover the precise location of the fault and aids in finding the cause and solution.
- The topology ensures the advantage of privacy or security. When every message travels along a dedicated line, only the intended recipient sees it. Physical boundaries prevent other users from gaining access to messages.

- **Disadvantages of mesh topology:**
  - The topology is expensive in terms of the cabling cost.
  - Installation is complex as each node is connected to every other node.
  - There is mesh of wiring which can be difficult to manage.

For these reasons a mesh topology is usually implemented in a limited fashion, for example, as a backbone connecting the main computers of a hybrid network that can include several other topologies.

2.4.2 Star Topology:

In a star topology, each node has a dedicated point-to-point link only to a central controller, usually called a hub. The hub manages and controls all functions of the network. The nodes are not directly linked to one another. Unlike a mesh topology, a star topology does not allow direct traffic between nodes. The controller acts as an exchange: If one node wants to send data to another, it sends the data to the controller, which then relays the data to the destination.



2.4.2 Star Topology

The star topology reduces the chance of network failure by connecting all the nodes to a central controller. Failure of a node on a star network is easy to detect and can be removed from the network. However, failure of the central controller will disable communication throughout the whole network.

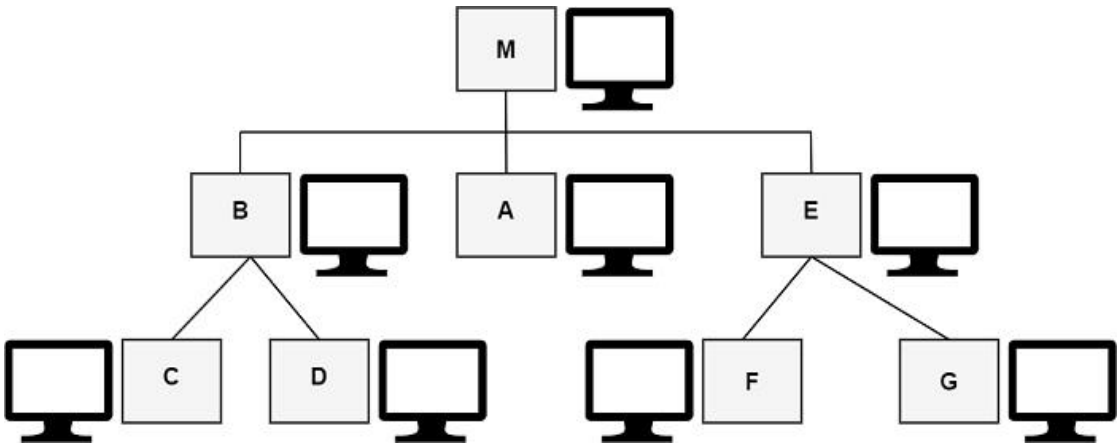
- **Advantages of star topology:**
  - Communication in the network is centrally managed by the hub.
  - The topology is easy to modify and add additional nodes.
  - It is very robust. Failure of one node does not affect the rest of the network.
  - It is easy to detect and isolate faults. Hub can be used to monitor link problems and bypass defective links.
  - It is less expensive than a mesh topology. In a star, each device needs only one link and one I/O port to connect it to any number of others.
  - It easy to install and reconfigure.
  - Far less cabling needs to be housed, and additions, moves, and deletions involve only one connection between that device and the hub.
- **Disadvantages of star topology:**
  - The whole topology depends on the hub or central node. If the hub fails, then the entire network fails.
  - Although a star requires far less cable than a mesh, each node must be linked to a central hub. The topology is relatively expensive in terms of the wiring cost than in some other topologies (such as ring or bus).

The star topology is used in local-area networks (LANs). High-speed LANs often use a star topology with a central hub.

2.4.3 Tree Topology

A tree network topology hierarchically links computers and requires data to circulate through the branches. In this network, various terminals and computers are connected to the main computer in a hierarchical aspect, with each additional device branching from one of the higher level.

It uses simple software to control the network. Tree topology provides a concentration point for control and error resolution. It is the central computer which is at the topmost order of hierarchy. Mostly it is a mainframe computer. This topology is also known as a vertical network or a hierarchical network. The topology is shown in the figure.



2.4.3 Tree Topology

The main computer, M, instructs the traffic flow between various components.

It uses a distributed aspect, and thus every computer is managed by its upper-level computer. In figure C and D, the system is controlled by B, similarly, F and G are controlled by E, and further A, B and E are controlled by M. Thus, indirectly, all the computers are controlled by the main computer M. Still, no doubt the controlling of children system by its parent reduce the workload of main computers. The topology is thus simple from the point of control, but at the same time, it is sure that if the main computer fails, the entire network can fail.

- **Advantages**

The advantages of the tree network topology are as follows

- Controlling and control software is very simple.
- A new node at the lowest levels can be inserted very efficiently.

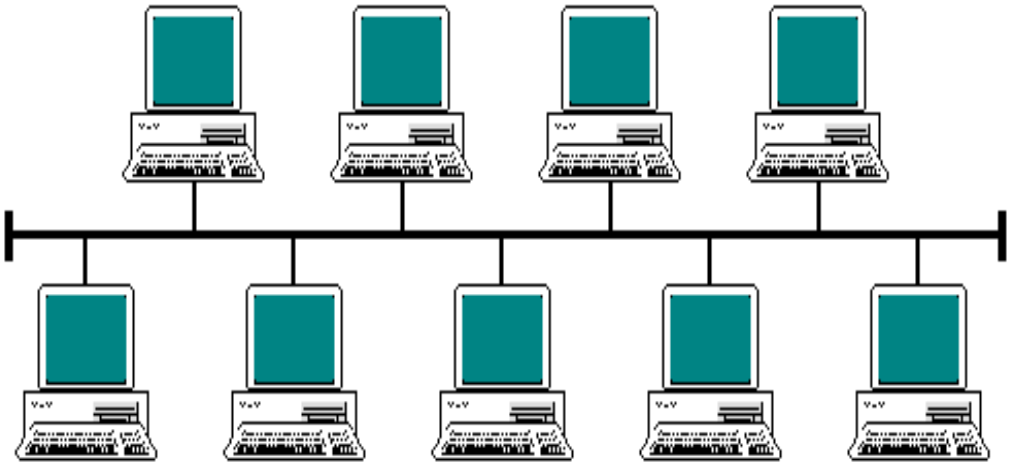
- **Disadvantages**

The disadvantages of the tree network topology are as follows

- If the topmost computer fails, it can lead to the failure of the entire network.
- If an intermediate level computer fails, it can cause the loss of its lower level machines from the network.
- If an original node is to be added at higher levels, it is complicated.
- If the levels boost, it can lead to an intricate network.

2.4.4 BUS Topology

The preceding examples all describe point-to-point connections. A bus topology, on the other hand, is multipoint. One long cable acts as a backbone to link all the devices in a network. On a bus network there is no central point for management and control. These functions are distributed to each node on the bus. Data frames originating at a node are propagated away from the node in both the directions on the bus. Each node on the bus interrogates the data frame destination address. If the destination does not match the node address, the node discards the data frame back on to the bus. If the destination address matches the node address, it accepts the data frame and processing is done.



2.4.4 Bus Topology

Bus topology was the one of the first topologies used in the design of early local area networks. Ethernet LANs can use a bus topology, but they are less popular.

- **Advantages of a bus topology:**
  - The bus topology is cheap, easy to handle and implement.
  - The topology is simple and reliable.
  - It requires less cable than mesh or star topologies.
  - It is easy to extend the network.
- **Disadvantages:**
  - It is difficult to reconnect and to isolates faults.
  - The cable length is limited. This limits the number of stations that can be connected.
  - A break in the bus cable stops all transmission, even between the devices on the same side of the problem. The damaged area reflects signals back in the direction of origin, creating noise in both directions.
  - Adding new devices may require modification or replacement of the backbone. Network performance deteriorates with additional nodes.
  - Problem solving may be more complex.

2.4.5 Ring Topology

In a ring topology, each node has a dedicated point-to-point connection with only the two nodes on either side of it. A signal is passed along the ring in one direction, from node to node, until it reaches its destination. Each node in the ring incorporates a repeater. When a node receives a signal intended for another node, its repeater regenerates the bits and passes them along. Generally, in a ring, a signal is circulating at all times. If one node does not receive a signal within a specified period, it can issue an alarm. The alarm alerts the network operator to the problem and its location.

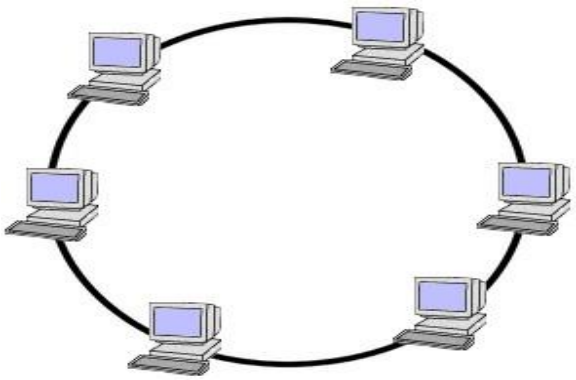


Figure 2.4.5: A Ring topology connecting four stations

- **Advantages:**
  - The topology is relatively easy to install and reconfigure.
  - Every node has equal access to the token and the opportunity to transmit.
  - In a ring, every device is linked to only its immediate neighbors. To add or delete a device requires changing only two connections. Thus the performance of the network is not affected by the addition of nodes.
- **Disadvantages:**
  - In a simple ring, the failure of a single node of the network can cause the entire network to fail. This weakness can be solved by using a dual ring or a switch capable of closing off the break.
  - It is difficult to isolate problems in the network.
  - Unidirectional traffic is a disadvantage.

Ring topology was prevalent when IBM introduced its local-area network Token Ring. Today, the need for higher-speed LANs has made this topology less popular.

2.4.6 Hybrid topologies

Hybrid topology is a network topology that is composed of one or more interconnections of two or more networks that are based upon different physical topologies. It is a mixture of above mentioned topologies. Usually, a central computer is attached with sub-controllers which in turn participate in a variety of topologies.

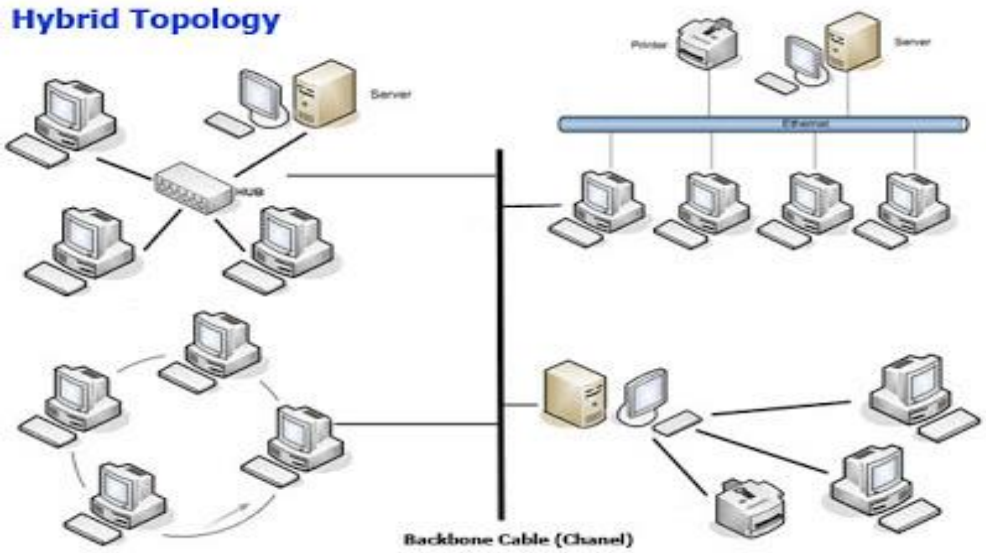


Figure 2.4.6: Hybrid topology

- **Advantages:**
  - The topology is extremely flexible.
  - It is highly reliable.

- **Disadvantages:**
  - The topology is very expensive than other topologies.
  - Hybrid network requires more cabling between its hardware devices.
  - Hybrid network is difficult to set up and troubleshoot.

2.4 SUMMARY

The transfer of data from one device to another through a transmission medium. The communication devices are connected by media links to form a network. The connection is either point-to-point or multipoint. Network topology refers to the physical or logical arrangement of a network. The basic topologies are mesh, star, bus, ring, tree and hybrid technologies.

2.5 KEYWORDS

Traffic
Transmission medium
Components
Twisted-pair

2.6 QUESTION FOR SELF STUDY

1. What is transmission medium? Explain
2. Explain transmission medium and its protocol.
3. Discuss in detail message, sender, receiver.
4. What is mesh topology? Explain
5. Explain Star topology in detail

6. Discuss in detail ring topology.
7. What is BUS topology?
8. Explain RING topology.
9. Discuss in detail hybrid topologies.

2.7 REFERENCES

1. Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.
2. Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.
3. Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT 3: TRANSMISSION MODES, NETWORK HARDWARE, IP ADDRESS

Structure:

3.0 Objectives

3.1 Introduction

3.2 Simplex, Half Duplex, Full Duplex

3.4 LAN, MAN, WAN.

3.5 DNS, IP address, MAC address,

3.6 Summary

3.7 Keywords

3.8 Questions for self-study

3.9 References

3.0 OBJECTIVES

After studying this unit, you will be able to

- Explain the transmission modes
- Understand Simplex, Half Duplex, Full Duplex
- Discuss the LAN, MAN, WAN.
- State the Applications DNS, IP address, MAC address,

3.1 INTRODUCTION

Transmission mode, also known as a communication mode, is the transfer of data between two devices via a communication channel that includes an optical fiber, wireless channels, copper wires, and other storage media. Data is transmitted between two devices in the form of electromagnetic waves. There are numerous data transmission methods in which the message is delivered in the form of a sequence of pulses utilizing digital modulation. The data transmission method was initially used in modems in a computer networking system in the 1940s, then in LANs, WANs, repeaters, and other networking systems. Based on the direction of exchange of information.

The Physical Layer in the Open System Interconnection (OSI) Layer Model is dedicated to data transmission in the network. It primarily determines the direction in which the data must travel to reach the receiving system or node. Data is transmitted between the devices via the communication channel that includes an optical fiber, copper wires, wireless channels, and other storage media.

3.2 SIMPLEX, HALF DUPLEX, FULL DUPLEX

The term transmission mode refers to the transmission of information between two communication devices via an **interaction channel** that indicates the direction of information flow between the devices. There are three primary types of transmission modes based on the direction of the exchange of information. The first is simplex, followed by half duplex, and finally full duplex.

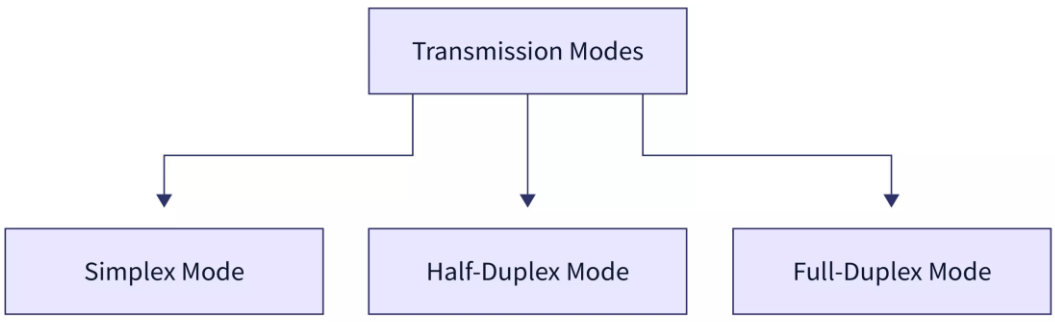


Figure 3.2 Transmission modes

- **Simplex**:

In simplex mode, the communication is unidirectional. Only one of the two devices on a link can transmit; the other can only receive (see Figure 3.2.1). Keyboards and traditional monitors are examples of simplex devices. The keyboard can only introduce input; the monitor can only accept output. The simplex mode can use the entire capacity of the channel to send data in one direction.

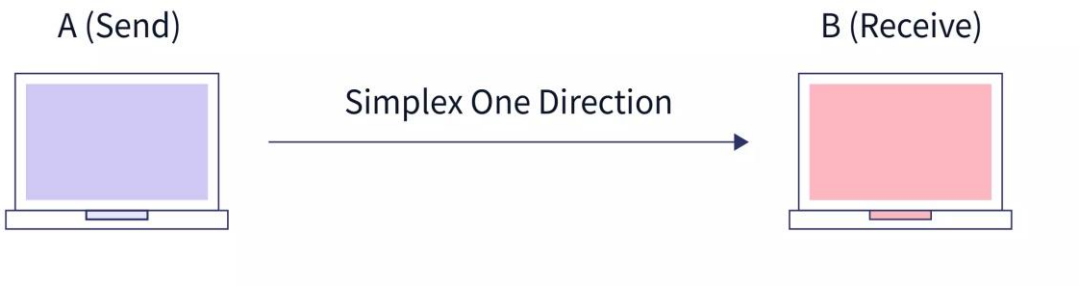


Figure 3.2.1: Simplex communication

Advantages

- The station can broadcast more data at once when operating in simplex mode since it can use the complete bandwidth of the communication channel.
- In a simplex mode of transmission, radio stations can use the complete bandwidth of the communicating channel to send all data in one shot with no data loss.

Disadvantages

- The simplex transmission mode is primarily utilized in corporate environments where rapid response is not necessary because communications primarily involve two-way data exchange.
- Since device communication is unidirectional in simplex mode, there is no intercommunication between them.

- **Half-Duplex**:

In half-duplex mode, both the stations can transmit and receive, but not at the same time. When one device is sending, the other can only receive, and vice versa (see Figure 3.2.2). In a half-duplex transmission, the entire capacity of a channel is taken over by whichever of the two devices is transmitting at the time. Walkie-talkies and CB (citizens band)

radios are both half-duplex systems. 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.

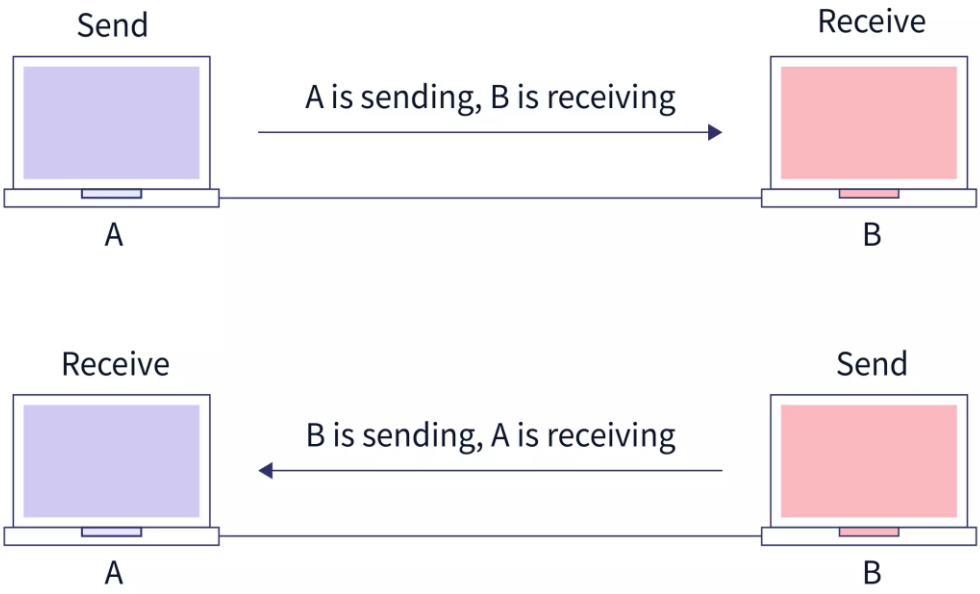


Figure 3.2.2: Half-Duplex communication

Advantages

- In half duplex mode, error detection is available, and if an error occurs, the receiver demands that the sender retransmit the data.
- Since this method of transmission allows for two-way communication, the entire bandwidth of the communication channel is used in only one direction at a time.

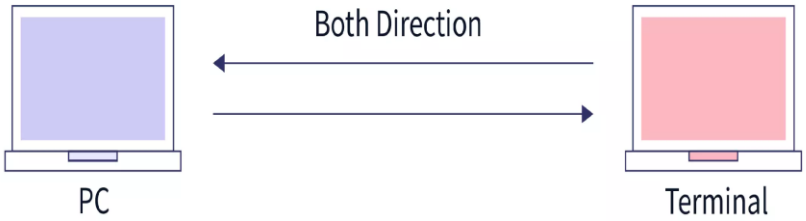
Disadvantages

- In half duplex mode, when one of the devices is sending the data, then another one has to wait. This causes a delay in sending the data at the right time.

- ***Full-Duplex***:

In full-duplex, 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 the other 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 of the channel is divided between signals traveling in both directions.

One common example of full-duplex communication is the telephone network. When two people are communicating by a telephone line, both can talk and listen at the same time. The 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.



Advantages

- The full duplex transmission mode is the fastest because the transmission happens in both ways.
- Both stations can send and receive data simultaneously.

Disadvantages

- If no dedicated path exists between the devices, the communication channel's capacity is divided into two parts.

3.3 LAN, MAN, WAN.

A **Local Area Network** (LAN) is a group of computer and peripheral devices which are connected in a limited area such as school, laboratory, home, and office building. It is a widely useful network for sharing resources like files, printers, games, and other application. The simplest type of LAN network is to connect computers and a printer in someone’s home or office. In general, LAN will be used as one type of transmission medium. It is a network which consists of less than 5000 interconnected devices across several buildings.

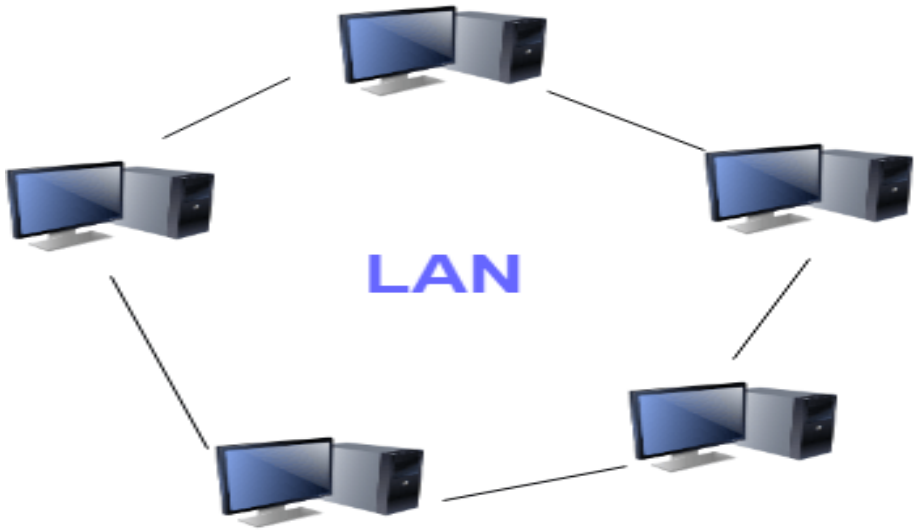


Figure 3.3.1 Local area Network

- **Characteristics of LAN**

Here are the important characteristics of a LAN network:

- It is a private network, so an outside regulatory body never controls it.
- LAN operates at a relatively higher speed compared to other WAN systems.
- There are various kinds of media access control methods like token ring and Ethernet.

- **Advantages of LAN**

Here are the pros/benefits of LAN:

- Computer resources like hard-disks, DVD-ROM, and printers can share local area networks. This significantly reduces the cost of hardware purchases.
- You can use the same software over the network instead of purchasing the licensed software for each client in the network.
- Data of all network users can be stored on a single hard disk of the server computer.
- You can easily transfer data and messages over networked computers.
- It will be easy to manage data at only one place, which makes data more secure.
- Local Area Network offers the facility to share a single internet connection among all the LAN users.

- **Disadvantages of LAN**

Here are the cons/drawbacks of LAN:

- LAN will indeed save cost because of shared computer resources, but the initial cost of installing Local Area Networks is quite high.
- The LAN admin can check personal data files of every LAN user, so it does not offer good privacy.
- Unauthorized users can access critical data of an organization in case LAN admin is not able to secure centralized data repository.
- Local Area Network requires a constant LAN administration as there are issues related to software setup and hardware failures

- **What is WAN (Wide Area Network)?**

WAN (Wide Area Network) is another important computer network that which is spread across a large geographical area. WAN network system could be a connection of a LAN which connects with other LAN's using telephone lines and radio waves. It is mostly limited to an enterprise or an organization.

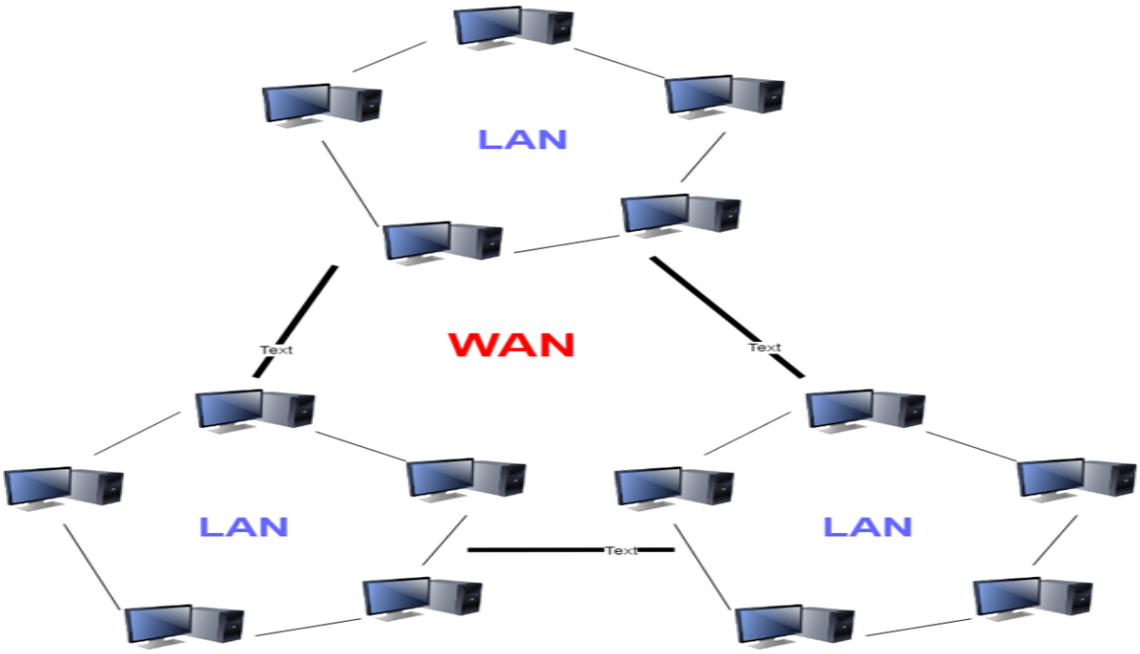


Figure 3.3.2 Wide Area Network

- **Characteristics of WAN**

Below are the characteristics of WAN:

- The software files will be shared among all the users; therefore, all can access to the latest files.
- Any organization can form its global integrated network using WAN.

- **Advantages of WAN**

Here are the benefits/pros of WAN:

- WAN helps you to cover a larger geographical area. Therefore, business offices situated at longer distances can easily communicate.
- Contains devices like mobile phones, laptop, tablet, computers, gaming consoles, etc.
- WLAN connections work using radio transmitters and receivers built into client devices.

- **Disadvantages of WAN**

Here are the drawbacks/cons of WAN network:

- The initial setup cost of investment is very high.
- It is difficult to maintain the WAN network. You need skilled technicians and network administrators.
- There are more errors and issues because of the wide coverage and the use of different technologies.
- It requires more time to resolve issues because of the involvement of multiple wired and wireless technologies.
- Offers lower security compared to other types of network in computer.

- **What is MAN (Metropolitan Area Network)?**

A **Metropolitan Area Network** or MAN is consisting of a computer network across an entire city, college campus, or a small region. This type of network is large than a LAN, which is mostly limited to a single building or site. Depending upon the type of

configuration, this type of network allows you to cover an area from several miles to tens of miles.

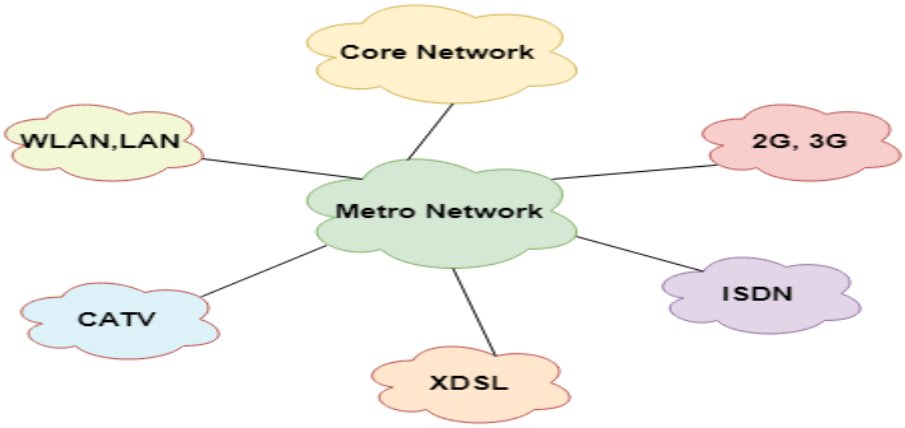


Figure 3.3.3 Metropolitan Area Network (MAN)

- **Characteristics of MAN**

Here are important characteristics of the MAN network:

- It mostly covers towns and cities in a maximum 50 km range
- Mostly used medium is optical fibers, cables
- Data rates adequate for distributed computing applications.

- **Advantages of MAN**

Here are the pros/benefits of MAN network:

- It offers fast communication using high-speed carriers, like fiber optic cables.
- It provides excellent support for an extensive size network and greater access to WANs.
- The dual bus in MAN network provides support to transmit data in both directions concurrently.
- A MAN network mostly includes some areas of a city or an entire city.

- **Disadvantages of MAN**

Here are drawbacks/cons of using the MAN network:

- You need more cable to establish MAN connection from one place to another.
- In MAN network it is tough to make the system secure from hackers

- **Other Types of Computer Networks**

Apart from above mentioned computer networks, here are some other important types of networks:

- WLAN (Wireless Local Area Network)
- Storage Area Network
- System Area Network
- Home Area Network
- POLAN- Passive Optical LAN
- Enterprise private network
- Campus Area Network
- Virtual Area Network
- Let's see all these different types of networks in detail:

1) WLAN

WLAN (Wireless Local Area Network) helps you to link single or multiple devices using wireless communication within a limited area like home, school, or office building. It gives users an ability to move around within a local coverage area which may be connected to the network. Today most modern day's WLAN systems are based on IEEE 802.11 standards.

2) Storage-Area Network (SAN)

A Storage Area Network is a type of network which allows consolidated, block-level data storage. It is mainly used to make storage devices, like disk arrays, optical jukeboxes, and tape libraries.

3) System-Area Network

System Area Network is used for a local network. It offers high-speed connection in server-to-server and processor-to-processor applications. The computers connected on a SAN network operate as a single system at quite high speed.

4) Passive Optical Local Area Network

POLAN is a networking technology which helps you to integrate into structured cabling. It allows you to resolve the issues of supporting Ethernet protocols and network apps. POLAN allows you to use optical splitter which helps you to separate an optical signal from a single-mode optical fiber. It converts this single signal into multiple signals.

5) Home Area Network (HAN):

A Home Area Network is always built using two or more interconnected computers to form a local area network (LAN) within the home. For example, in the United States, about 15 million homes have more than one computer.

These types of network connections help computer owners to interconnect with multiple computers. This network allows sharing files, programs, printers, and other peripherals.

6) Enterprise Private Network:

Enterprise private network (EPN) networks are build and owned by businesses that want to securely connect numerous locations in order to share various computer resources.

7) Campus Area Network (CAN):

A Campus Area Network is made up of an interconnection of LANs within a specific geographical area. For example, a university campus can be linked with a variety of campus buildings to connect all the academic departments.

8) Virtual Private Network:

A VPN is a private network which uses a public network to connect remote sites or users together. The VPN network uses “virtual” connections routed through the internet from the enterprise’s private network or a third-party VPN service to the remote site. It is a free or paid service that keeps your web browsing secure and private over public Wi-Fi hotspots.

3.4 DNS, IP ADDRESS, MAC ADDRESS

The Application Layer: The application layer is the top most layer on the network model where all the applications are found. Application layer consists of application protocols that provide services to user applications. The layers below the application layer provide reliable transport and they do not involve in helping users directly.
However, even in the application layer there is a need for support protocols, to allow the applications to function. Accordingly, we will look at one of these before starting with the applications themselves. The item in question is DNS, which handles naming within the Internet. After that, we will examine real application like electronic mail.

The **Domain Name System** (**DNS**) is a hierarchical naming system built on a distributed database for computers, services, or any resource connected to the Internet or a private network. It associates various information with domain names assigned to each of the participating entities. Most importantly, it translates domain names meaningful to humans into the numerical identifiers associated with networking equipment for the purpose of locating and addressing these devices worldwide.

Although programs theoretically could refer to hosts, mailboxes, and other resources by their network (e.g., IP) addresses, these addresses are hard for people to remember. Also, sending e-mail to *tana@128.111.24.41* means that if Tana's ISP or organization moves the mail server to a different machine with a different IP address, her e-mail address has to change. Consequently, ASCII names were introduced to decouple machine names from machine addresses. In this way, Tana's address might be something like *tana@art.ucsb.edu*. Nevertheless, the network itself understands only numerical addresses, so some mechanism is required to convert the ASCII strings to network addresses. In the following sections we will study how this mapping is accomplished in the Internet. Way back in the ARPANET, there was simply a file, *hosts.txt* that listed all the hosts and their IP addresses. Every night, all the hosts would fetch it from the site at which it was maintained. For a network of a few hundred large timesharing machines, this approach worked reasonably well.

However, when thousands of minicomputers and PCs were connected to the net, everyone realized that this approach could not continue to work forever. For one thing,

the size of the file would become too large. However, even more important, host name conflicts would occur constantly unless names were centrally managed, something unthinkable in a huge international network due to the load and latency. To solve these problems, **DNS** (the **Domain Name System**) was invented. The essence of DNS is the invention of a hierarchical, domain-based naming scheme and a distributed database system for implementing this naming scheme. It is primarily used for mapping host names and e-mail destinations to IP addresses but can also be used for other purposes. DNS is defined in RFCs 1034 and 1035.

- **The DNS Name Space**

Managing a large and constantly changing set of names is a nontrivial problem. In the postal system, name management is done by requiring letters to specify (implicitly or explicitly) the country, state or province, city, and street address of the addressee. By using this kind of hierarchical addressing, there is no confusion between the Marvin Anderson on Main St. in White Plains, N.Y. and the Marvin Anderson on Main St. in Austin, Texas. DNS works the same way.

Conceptually, the Internet is divided into over 200 top-level domains, where each domain covers many hosts. Each domain is partitioned into sub domains, and these are further partitioned, and so on. All these domains can be represented by a tree, as shown in Fig. 3.4.1. The leaves of the tree represent domains that have no sub domains (but do contain machines, of course). A namespace maps each address to a unique name. These can be represented by a tree representing domain and sub domains. The leaves of the tree represent domains that do not have any further sub domain. A leaf domain may contain a single host or group hosts (representing a company or any community).

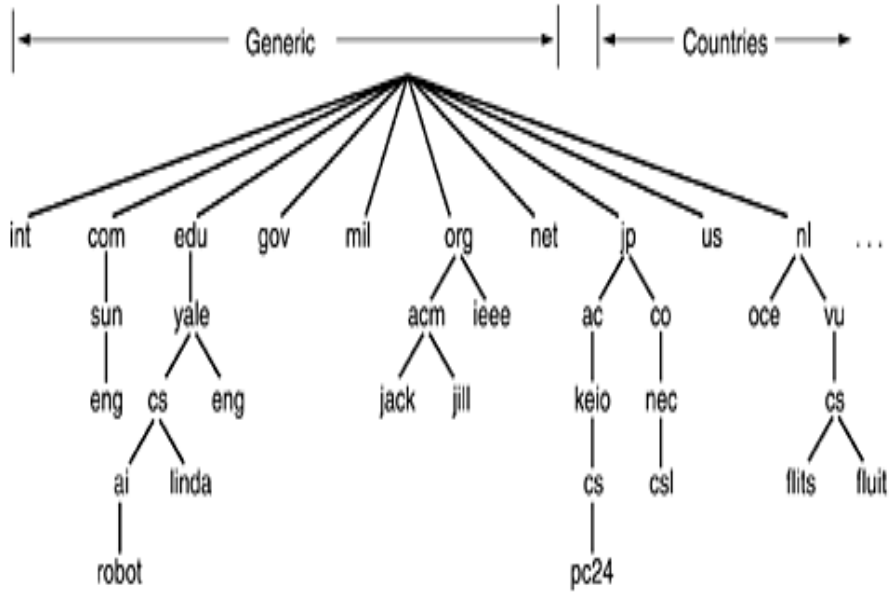


Figure 3.4.1. A portion of the Internet domain name space.

The top-level domains come in two flavors: generic and countries. The original generic domains were *com (commercial), edu* (educational institutions), *gov* (the U.S. Federal Government), *int* (certain international organizations), *mil* (the U.S. armed forces), *net* (network providers), and *org* (nonprofit organizations). The country domains include one entry for every country, as defined in ISO 3166.

In November 2000, ICANN approved four new, general-purpose, top-level domains, namely, *biz* (businesses), *info* (information), *name* (people's names), and *pro* (professions, such as doctors and lawyers). In addition, three more specialized top-level domains were introduced at the request of certain industries. These are *aero* (aerospace industry), *coop* (co-operatives), and *museum* (museums). Other top-level domains will be added in the future.

In general, getting a second-level domain, such as *name-of-company*.com, is easy. It merely requires going to a registrar for the corresponding top-level domain (*com* in this case) to check if the desired name is available and not somebody else's trademark. If there are no problems, the requester pays a small annual fee and gets the name. By now, virtually every common (English) word has been taken in the *com* domain. Try household articles, animals, plants, body parts, etc. Nearly all are taken.

Each domain is named by the path upward from it to the (unnamed) root. The components are separated by periods (pronounced "dot"). Thus, the engineering department at Sun Microsystems might be *eng.sun.com,* rather than a UNIX-style name such as */com/sun/eng.* Notice that this hierarchical naming means that *eng.sun.com.* does not conflict with a potential use of *eng* in *eng.yale.edu.*, which might be used by the Yale English department.

Domain names can be either absolute or relative. An absolute domain name always ends with a period (e.g., *eng.sun.com.*), whereas a relative one does not. Relative names have to be interpreted in some context to uniquely determine their true meaning. In both cases, a named domain refers to a specific node in the tree and all the nodes under it.
Domain names are case insensitive, so *edu*, *Edu*, and *EDU* mean the same thing. Component names can be up to 63 characters long, and full path names must not exceed 255 characters. In principle, domains can be inserted into the tree in two different ways. For example, *cs.yale.edu* could equally well be listed under the US country domain as *cs.yale.ct.us*. In practice, however, most organizations in the United States are under a generic domain, and most outside the United States are under the domain of their country. There is no rule against registering under two top-level domains, but few organizations except multinationals do it (e.g., *sony.com* and *sony.nl*).

Each domain controls how it allocates the domains under it. For example, Japan has domains *ac.jp* and *co.jp* that mirror *edu* and *com*. The Netherlands does not make this distinction and puts all organizations directly under *nl*. Thus, all three of the following are university computer science departments:

1. *cs.yale.edu* (Yale University, in the United States)
2. *cs.vu.nl* (Vrije University, in The Netherlands)
3. *cs. keio.ac.jp* (Keio University, in Japan)

To create a new domain, permission is required of the domain in which it will be included. For example, if a VLSI group is started at Yale and wants to be known as *vlsi.cs.yale.edu*, it has to get permission from whoever manages *cs.yale.edu*. Similarly, if a new university is chartered, say, the University of Northern South Dakota, it must ask the manager of the *edu* domain to assign it *unsd.edu*. In this way, name conflicts are

avoided and each domain can keep track of all its sub domains. Once a new domain has been created and registered, it can create sub domains, such as *cs.unsd.edu*, without getting permission from anybody higher up the tree.

Naming follows organizational boundaries, not physical networks. For example, if the computer science and electrical engineering departments are located in the same building and share the same LAN, they can nevertheless have distinct domains. Similarly, even if computer science is split over Babbage Hall and Turing Hall, the hosts in both buildings will normally belong to the same domain.

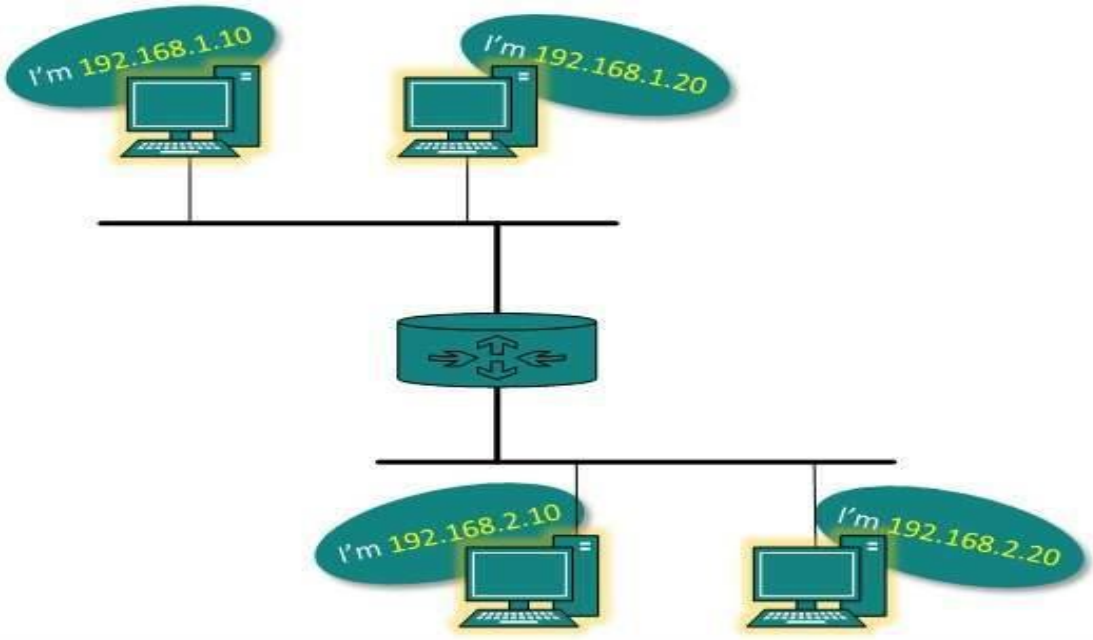
- **IP Address**

A network address always points to host / node / server or it can represent a whole network. Network address is always configured on network interface card and is generally mapped by system with the MAC address (hardware address or layer-2 address) of the machine for Layer-2 communication.

There are different kinds of network addresses in existence:

- IP
- IPX
- AppleTalk

We are discussing IP here as it is the only one we use in practice these days.



IP addressing provides mechanism to differentiate between hosts and network. Because IP addresses are assigned in hierarchical manner, a host always resides under a specific

network. The host which needs to communicate outside its subnet, needs to know destination network address, where the packet/data is to be sent.

Hosts in different subnet need a mechanism to locate each other. This task can be done by DNS. DNS is a server which provides Layer-3 address of remote host mapped with its domain name or FQDN. When a host acquires the Layer-3 Address (IP Address) of the remote host, it forwards all its packet to its gateway. A gateway is a router equipped with all the information which leads to route packets to the destination host.

Routers take help of routing tables, which has the following information:

- **Method to reach the network**

  Routers upon receiving a forwarding request, forwards packet to its next hop (adjacent router) towards the destination.

  The next router on the path follows the same thing and eventually the data packet reaches its destination.

  Network address can be of one of the following:

  - Unicast (destined to one host)
  - Multicast (destined to group)
  - Broadcast (destined to all)
  - Anycast (destined to nearest one)

  A router never forwards broadcast traffic by default. Multicast traffic uses special treatment as it is most a video stream or audio with highest priority. Anycast is just similar to unicast, except that the packets are delivered to the nearest destination when multiple destinations are available.

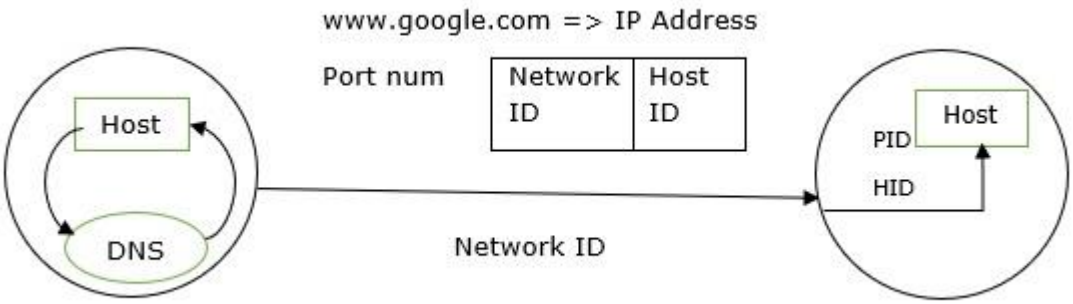
  Computer network is a group of computers or nodes that are linked to each other to share information and resources. Whereas, IP addresses are generally represented by a 32-bit unsigned binary value. It is represented in a dotted decimal format.

  For example, 9.250.7.5 is a valid IP address.

  The IP address consists of a pair of numbers

  IP address = <network number><host number>

  Let see the pictorial representation of two hosts that are communicated with the help of Network and IP address.



- **Explanation**

Step 1 − The service that is used to convert the Domain name to IP address is called Domain Name Service.

Step 2 − The port number is used to Identify a particular host, for well-known services the port numbers are predefined and fixed.

For example

http=>port number=80

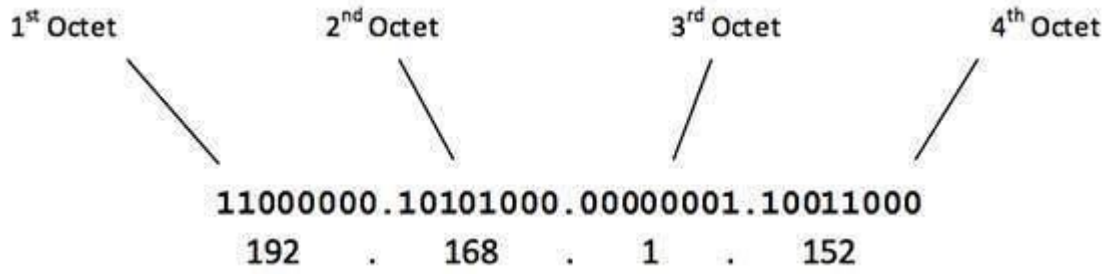
ftp=> port number=21

SMTP=> Port number=25

Step 3 − Even though our intention is to reach google.com, we have to visit DNS first and then getting the IP address of google.com we will get the google home page. Actually this is overhead, also called DNS overhead. This problem can be rectified when we get the IP address of google.com. For the first time, we need to store that in our personal system for some time. Later on, when we try to access the site again we can get the IP address directly from our system. By doing this there is no need to get DNS for every time access which reduces the overheads. If the IP address expires then there is no other alternative, we have to go to DNS.IP addresses are used by the IP protocol helpful to identify a host on the Internet. Generally, speaking, an IP address identifies an interface that is capable of sending and Receiving IP datagrams.

Internet Protocol hierarchy contains several classes of IP Addresses to be used efficiently in various situations as per the requirement of hosts per network. Broadly, the IPv4 Addressing system is divided into five classes of IP Addresses. All the five classes are

identified by the first octet of IP Address.Internet Corporation for Assigned Names and Numbers is responsible for assigning IP addresses.The first octet referred here is the left most of all. The octets numbered as follows depicting dotted decimal notation of IP Address



The number of networks and the number of hosts per class can be derived by this formula

Number of networks = 2^network_bits

Number of Hosts/Network = 2^host_bits – 2

When calculating hosts' IP addresses, 2 IP addresses are decreased because they cannot be assigned to hosts, i.e. the first IP of a network is network number and the last IP is reserved for Broadcast IP.

- **Class A Address**

The first bit of the first octet is always set to 0 (zero). Thus the first octet ranges from 1 – 127, i.e.

00000001 – 01111111
1 – 127

Class A addresses only include IP starting from 1.x.x.x to 126.x.x.x only. The IP range 127.x.x.x is reserved for loopback IP addresses.

The default subnet mask for Class A IP address is 255.0.0.0 which implies that Class A addressing can have 126 networks (2^7-2) and 16777214 hosts (2^{24}-2).

Class A IP address format is thus: **0NNNNNNN**.HHHHHHHH.HHHHHHHH.HHHHHHHH

- **Class B Address**

An IP address which belongs to class B has the first two bits in the first octet set to 10, i.e.

10000000 – 10111111
128 – 191

Class B IP Addresses range from 128.0.x.x to 191.255.x.x. The default subnet mask for Class B is 255.255.x.x.

Class B has 16384 (2^{14}) Network addresses and 65534 (2^{16}-2) Host addresses.

Class B IP address format is: **10NNNNNN.NNNNNNNN**.HHHHHHHH.HHHHHHHH

- **Class C Address**

The first octet of Class C IP address has its first 3 bits set to 110, that is −

11000000 – 11011111
192 – 223

Class C IP addresses range from 192.0.0.x to 223.255.255.x. The default subnet mask for Class C is 255.255.255.x.

Class C gives 2097152 (2^{21}) Network addresses and 254 (2^{8}-2) Host addresses.

Class C IP address format is: **110NNNNN.NNNNNNNN.NNNNNNNN**.HHHHHHHH

- **Class D Address**

Very first four bits of the first octet in Class D IP addresses are set to 1110, giving a range of −

11100000 – 11101111
224 – 239

Class D has IP address range from 224.0.0.0 to 239.255.255.255. Class D is reserved for Multicasting. In multicasting data is not destined for a particular host, that is why there is no need to extract host address from the IP address, and Class D does not have any subnet mask.

- **Class E Address**

This IP Class is reserved for experimental purposes only for R&D or Study. IP addresses in this class ranges from 240.0.0.0 to 255.255.255.254. Like Class D, this class too is not equipped with any subnet mask.

- **MAC Address**

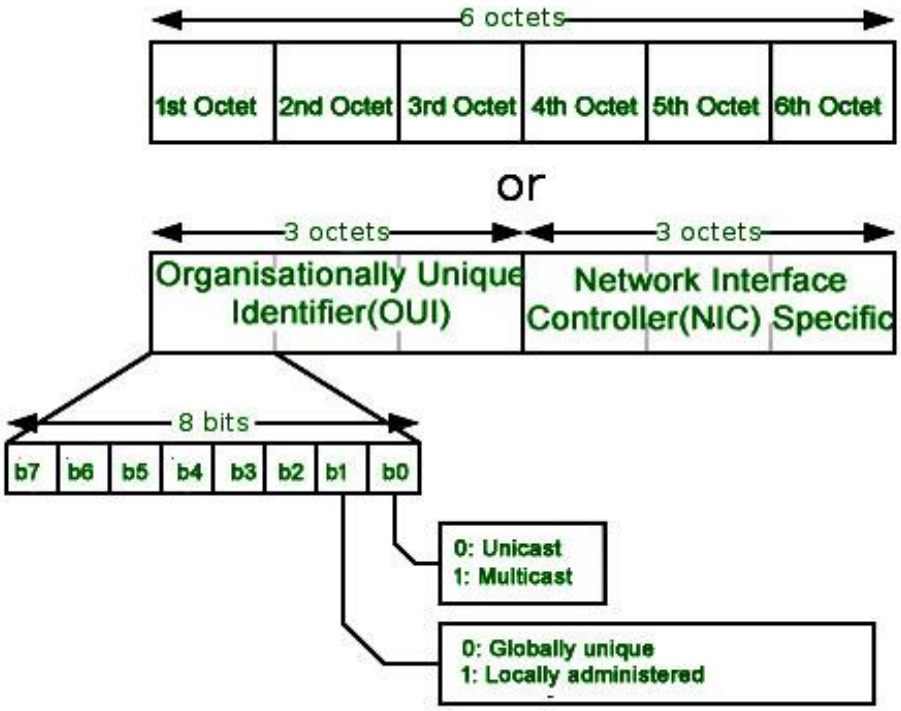
In order to communicate or transfer the data from one computer to another computer, we need some address. In Computer Network various types of addresses are introduced; each works at a different layer. Media Access Control Address is a physical address that works at the Data Link Layer. In this article, we will discuss about addressing DLL, which is MAC Address.

- **Media Access Control (MAC) Address**

MAC Addresses are unique **48-bits** hardware number of a computer, which is embedded into a network card (known as a **Network Interface Card**) during the time of manufacturing. MAC Address is also known as the **Physical Address** of a network device. In IEEE 802 standard, Data Link Layer is divided into two sublayers –

- Logical Link Control(LLC) Sublayer
- Media Access Control(MAC) Sublayer

MAC address is used by the Media Access Control (MAC) sublayer of the Data-Link Layer. MAC Address is worldwide unique since millions of network devices exist and we need to uniquely identify each.



- **Format of MAC Address**

MAC Address is a 12-digit hexadecimal number (6-Byte binary number), which is mostly represented by Colon-Hexadecimal notation. The First 6-digits (say 00:40:96) of MAC Address identifies the manufacturer, called OUI (**Organizational Unique Identifier**). IEEE Registration Authority Committee assigns these MAC prefixes to its registered vendors.

Here are some OUI of well-known manufacturers :

CC:46: D6 - Cisco

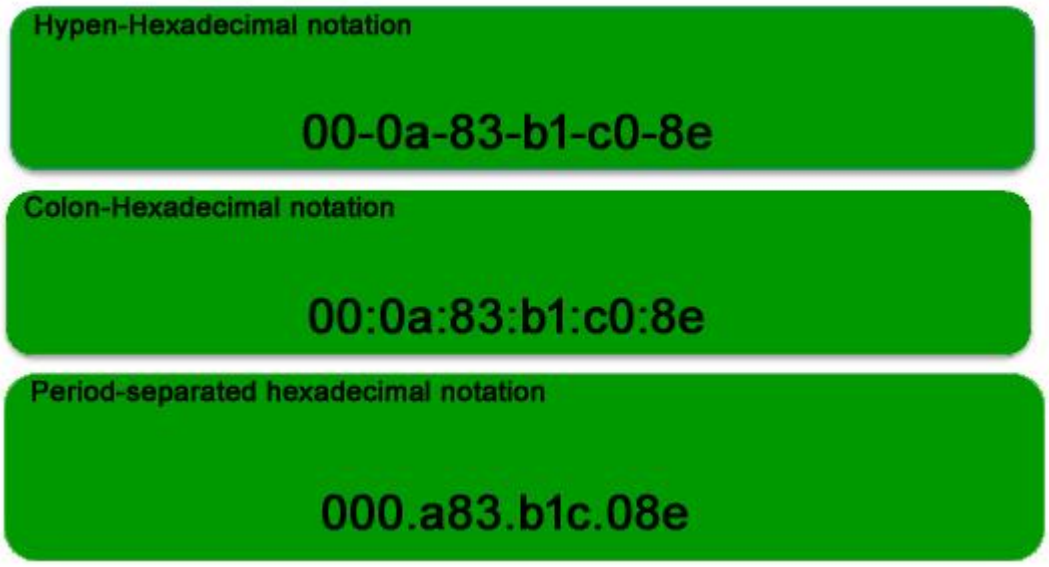
3C:5A: B4 - Google, Inc.

3C:D9:2B - Hewlett Packard

00:9A:CD - HUAWEI TECHNOLOGIES CO., LTD

The rightmost six digits represent **Network Interface Controller**, which is assigned by the manufacturer.

As discussed above, the MAC address is represented by Colon-Hexadecimal notation. But this is just a conversion, not mandatory. MAC address can be represented using any of the following formats:



Note: Colon-Hexadecimal notation is used by *Linux OS* and Period-separated Hexadecimal notation is used by *Cisco Systems*.

- **How to find MAC address?**

Command for UNIX/Linux - *ifconfig -a*

ip link list

ip address show

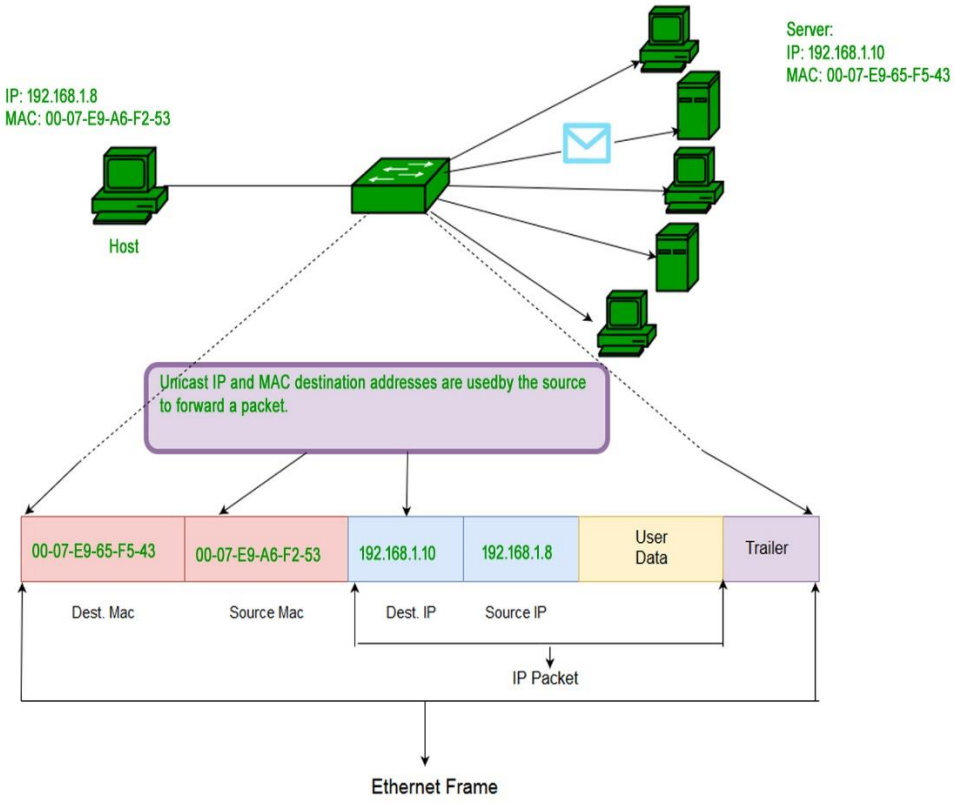
Command forWindows OS - *ipconfig /all*

MacOS - *TCP/IP Control Panel*

Note – LAN technologies like Token Ring, and Ethernet use MAC Addresses as their Physical address but there are some networks (AppleTalk) that do not use MAC addresses.

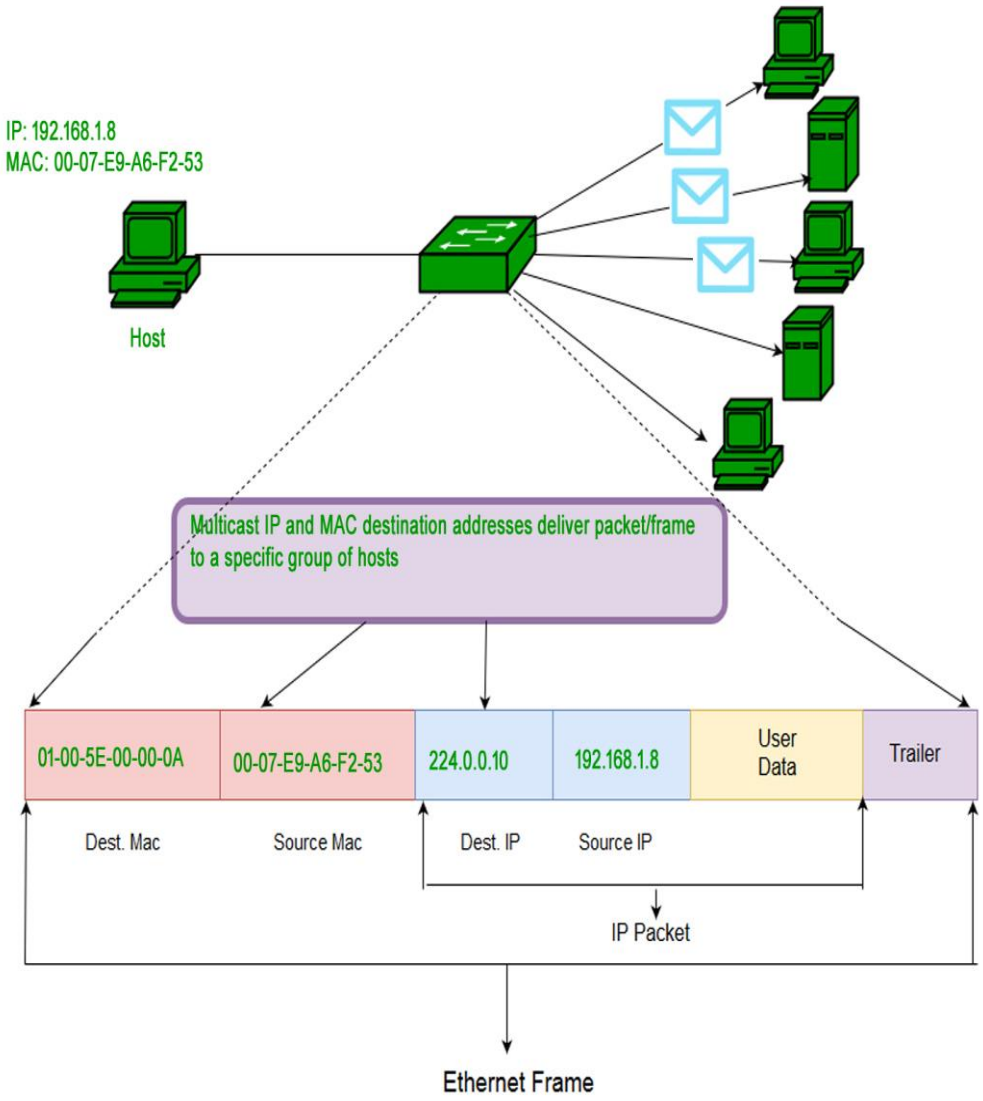
- **Types of MAC Address:**

1. Unicast: A Unicast addressed frame is only sent out to the interface leading to a specific NIC. If the LSB (least significant bit) of the first octet of an address is set to zero, the frame is meant to reach only one receiving NIC. MAC Address of source machine is always Unicast.

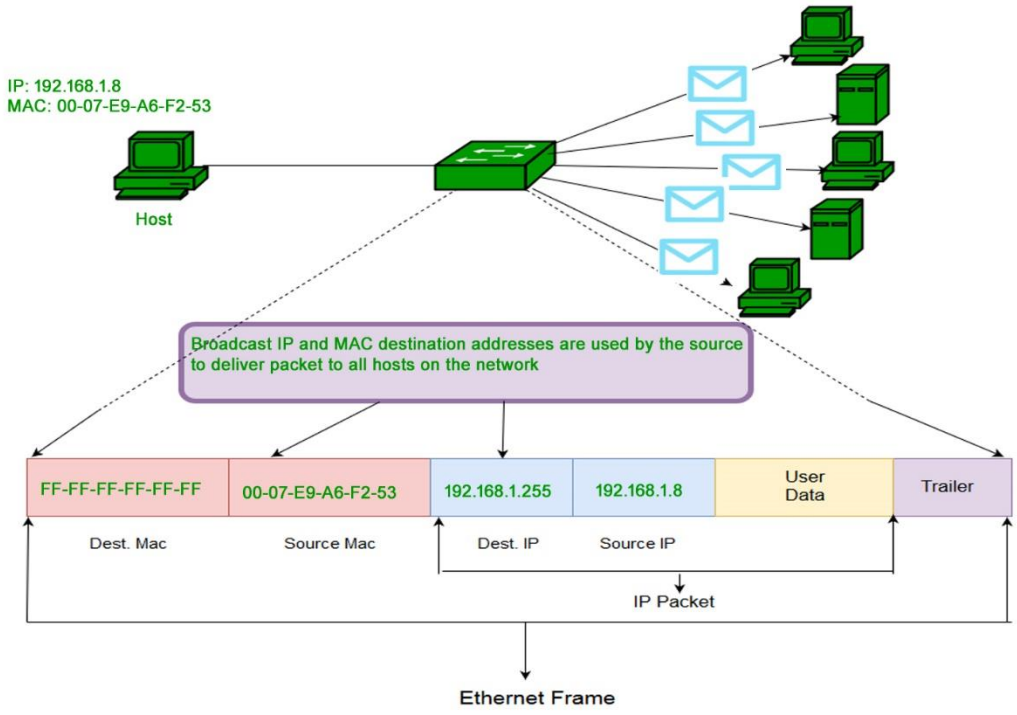


2. Multicast: The multicast address allows the source to send a frame to a group of devices. In Layer-2 (Ethernet) Multicast address, LSB (least significant bit) of the first

octet of an address is set to one. IEEE has allocated the address block 01-80-C2-xx-xx-xx (01-80-C2-00-00-00 to 01-80-C2-FF-FF-FF) for group addresses for use by standard protocols.



3. Broadcast: Similar to Network Layer, Broadcast is also possible on the underlying layer (Data Link Layer). Ethernet frames with ones in all bits of the destination address (FF-FF-FF-FF-FF-FF) are referred to as the broadcast addresses. Frames that are destined with MAC address FF-FF-FF-FF-FF-FF will reach every computer belonging to that LAN segment.



- **What is MAC Cloning:**

Some ISPs use MAC addresses in order to assign an IP address to the gateway device. When a device connects to the ISP, the DHCP server records the MAC address and then assigns an IP address. Now the system will be identified through the MAC address. When the device gets disconnected, it loses the IP address. If the user wants to reconnect, the DHCP server checks if the device is connected before. If so, then the server tries to assign the same IP address (in case the lease period has not expired). In case user changed the router, the user has to inform the ISP about new MAC address because the new MAC address is unknown to ISP, so the connection cannot be established.

Or the other option is **Cloning**, user can simply clone the registered MAC address with ISP. Now router keeps reporting the old MAC addresses to ISP and there will be no connection issue.

- **Characteristics of MAC address:**

Media Access Control address (MAC address) is a unique identifier assigned to most network adapters or network interface cards (NICs) by the manufacturer for identification and used in the Media Access Control protocol sub-layer.

An Ethernet MAC address is a 48-bit binary value expressed as 12 hexadecimal digits (4 bits per hexadecimal digit). MAC addresses are in a flat structure and thus they are not routable on the Internet. Serial interfaces do not use MAC addresses. It does NOT contain a network and host portion with the address. It is used to deliver the frame to the destination device.

3.6 SUMMARY

Transmission mode, also known as a communication mode, is the transfer of data between two devices via a communication channel that includes an optical fiber, wireless channels, copper wires, and other storage media. The connection is either simplex, half duplex, Full duplex The term transmission mode refers to the transmission of information between two communication devices via an **interaction channel** that indicates the direction of information flow between the devices. So the LAN stands for local area network. MAN stands for metropolitan area network. WAN stands for wide area network. Operates in small areas such as the same building or campus. While IP addresses are used to uniquely identifies the connection of the network with that device takes part in a network. MAC Address is used to ensure the physical address of the computer. It uniquely identifies the devices on a network.

3.7 KEYWORDS

LAN

WAN

MAN

MAC

3.8 QUESTION FOR SELF STUDY

1. What is Simplex mode? Explain
2. Explain Half-Duplex.
3. What is *Full-Duplex*? Explain
4. Explain LAN.

5. What is WAN? Explain
6. Explain MAN
7. What is DNS? Explain
8. Explain IP address.
9. What is MAC address.

3.9 REFERENCES

1 Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.

2 Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.

3 Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT 4: WEB BROWSER

Structure:

4.0 Objectives

4.1 Introduction

4.2 ISP

4.3 URL

4.4 WWW

4.5 Broadband Transmissions

4.6 Guided Media – Twisted Pair Cable, Coaxial Cable, Fiber-Optic Cable

4.7 Summary

4.8 Keywords

4.9 Questions for self-study

4.10 References

4.0 OBJECTIVES

After studying this unit, you will be able to

- Explain the web browser
- Understand ISP, URL, WWW.Broad band transmission
- Discuss the Guided Media – Twisted Pair Cable
- State the Applications Coaxial Cable, Fiber-Optic Cable

4.1 INTRODUCTION

Today web browsers are easily accessible and can be used on devices like computer, laptops, mobile phones, etc. but this evolution of making browsers available for easy use took many years.

Given below are some salient points which one must know with regard to the history of web browsers:

- **"Worldwide Web"** was the first web browser created by Tim Berners Lee in 1990. This is completely different from the World Wide Web we use today
- In 1993, the **"Mosaic"** web browser was released. It had the feature of adding images and an innovative graphical interface. It was the "the world's first popular browser"
- After this, in 1994, Marc Andreessen (leader of Mosaic Team) started working on a new web browser, which was released and was named **"Netscape Navigator"**
- In 1995, **"Internet Explorer"** was launched by Microsoft. It soon overtook as the most popular web browser
- In 2002, **"Mozilla Firefox"** was introduced which was equally as competent as Internet Explorer
- Apple too launched a web browser in the year 2003 and named it **"Safari"**. This browser is commonly used in Apple devices only and not popular with other devices
- Finally, in the year 2008, Google released "Chrome" and within a time span of 3 years it took over all the other existing browsers and is one of the most commonly used web browsers across the world

- **Functions of Web Browser**

Our dependency on the Internet has massively increased. Stated below are functions of web browsers and how are they useful:

- The main function is to retrieve information from the World Wide Web and making it available for users

- Visiting any website can be done using a web browser. When a URL is entered in a browser, the web server takes us to that website
- To run Java applets and flash content, plugins are available on the web browser
- It makes Internet surfing easy as once we reach a website we can easily check the hyperlinks and get more and more useful data online
- Browsers user internal cache which gets stored and the user can open the same webpage time and again without losing extra data
- Multiple webpages can be opened at the same time on a web browser
- Options like back, forward, reload, stop reload, home, etc. are available on these web browsers, which make using them easy and convenient

- **Types of Web Browser**

The functions of all web browsers are the same. Thus, more than the different types there are different web browsers which have been used over the years.

Discussed below are different web browser examples and their specific features:

- **Worldwide Web**

  - The first web browser ever
  - Launched in 1990
  - It was later named "Nexus" to avoid any confusion with the World Wide Web
  - Had the very basic features and less interactive in terms of graphical interface
  - Did not have the feature of bookmark

- **Mosaic**

  - It was launched in 1993
  - The second web browser which was launched
  - Had a better graphical interface. Images, text and graphics could all be integrated
  - It was developed at the National Center for Supercomputing Applications
  - The team which was responsible for creating Mosaic was lead by Marc Andreessen
  - It was named "the world's first popular browser"

- **Netscape Navigator**

  - It was released in 1994
  - In the 1990s, it was the dominant browser in terms of usage share

- More versions of this browser were launched by Netscape
- It had an advanced licensing scheme and allowed free usage for non-commercial purposes

- **Internet Explorer**

  - It was launched in 1995 by Microsoft
  - By 2003, it has attained almost 95% of usage share and had become the most popular browsers of all
  - Close to 10 versions of Internet Explorer were released by Microsoft and were updated gradually
  - It was included in the Microsoft Windows operating system
  - In 2015, it was replaced with “Microsoft Edge”, as it became the default browser on Windows 10

- **Firefox**

  - It was introduced in 2002 and was developed by Mozilla Foundation
  - Firefox overtook the usage share from Internet Explorer and became the dominant browser during 2003-04
  - Location-aware browsing was made available with Firefox
  - This browser was also made available for mobile phones, tablets, etc.

- **Google Chrome**

  - It was launched in 2008 by Google
  - It is a cross-platform web browser
  - Multiple features from old browsers were amalgamated to form better and newer features
  - To save computers from malware, Google developed the ad-blocking feature to keep the user data safe and secure
  - Incognito mode is provided where private searching is available where no cookies or history is saved
  - Till date, it has the best user interface
  - Apart from these, Opera Mini web browser was introduced in 2005 which was specially designed for mobile users. Before the mobile version, the computer

version “Opera” was also released in 1995. It supported a decent user interface and was developed by Opera Software.

4.2 ISP (Internet Service Provider)

ISP stands for Internet Service Provider which is a term used to refer to a company that provides internet access to people who pay the company or subscribe to the company for the same. For their services, the customers have to pay the internet service provider a nominal fee which varies according to the amount of data they actually use or the data plan which they purchase. An Internet Service Provider is also known as an Internet Access Provider or an online service provider. An Internet Service Provider is a must if one wants to connect to the internet. The first Internet Service Provider was Telenet. Telenet was the commercialized version of the ARPANET – a precursor to the internet, of sorts. Telenet was introduced in 1974. Since then, many Internet Service Providers have entered the scene and this was partly because of the proliferation of the internet as a commodity that fuelled the consumerist attitude of the people. Pretty soon, an Internet Service Provider called “The World” came to be in vogue and ever since it started serving its customers today in 1989 has cemented itself as the first archetypal Internet Service Provider. Examples of major Internet Service Providers include Google Fiber, Verizon, Jio, AT&T etc.

- **Characteristics**
  - **E-mail Account:** Many Internet Service Providers offer an e-mail address to their consumers.
  - **User Support:** Professionals and an increasing number of lay users prefer an ISP that can provide them with customer support so that they have someone they can refer to if things go awry.
  - **Access to high-speed internet:** Probably the most obvious item on this list as this feature of an Internet Service Provider lies literally in its name. Furthermore, the higher the speed an Internet Service Provider can offer one, the better it’s standing in the market and the more customers it can attract.
  - **Spam Blocker:** An Internet Service Provider that hinders its customers’ productivity by way of not blocking spam and displaying frequent ads is not

something that is generally favoured in the market today. Therefore, many of the Internet Service Providers offer spam blocking features to their customers.

- **Web Hosting:** Some of the ISPs offer web hosting services to their clientele as well.

- **Different types of ISP connections**
  - DSL
  - Wi-Fi broadband
  - mobile broadband
  - fibre optic broadband
  - cable broadband

- **List of ISP**
  - Reliance Jio
  - Vodafone Idea
  - Airtel
  - BSNL
  - Hathway
- **Advantages**
  - The customer need not then bother with either the technicalities or finances of investing and inventing a web browser to work with. An ISP can readily do all of this for its customers.
  - Many ISPs, being professional companies, provide its clientele with high-speed internet and that is not possible if one decides to sidesteps these companies.
  - ISPs offer a very high degree of reliability and availability
  - The ISPs are secure – they offer a tremendous deal of protection against viruses and use only the latest software patches whilst operating and thereby, maintaining the integrity of the browser.
  - User do not need to invest in user's own web server.
  - ISP's should give the best uptime guarantee.

- **Disadvantages**
  - Because of the range of options available in the market and due to cut-throat competition, some of the ISPs have been accused of violating the customers' trust by way of inflated pricing, data losses, etc. It is true that using an ISP makes the customer entirely dependent on it.
  - If an Internet Service Provider is stretched thin because of hosting too many sites on a shared server, it can compromise the quality of the customers' data by way of slow download rates and poor performance of websites.
  - User need to trust user's ISP for uptime and security.
  - ISP can directly affect user if the it gets blacklisted.

4.3 URL (Uniform Resource Locator)

A URL (Uniform Resource Locator) is a unique identifier used to locate a resource on the Internet. It is also referred to as a web address. URLs consist of multiple parts -- including a protocol and domain name -- that tell a web browser how and where to retrieve a resource. End users use URLs by typing them directly into the address bar of a browser or by clicking a hyperlink found on a webpage, bookmark list, in an email or from another application.

- **How is a URL structured?**

The URL contains the name of the protocol needed to access a resource, as well as a resource name. The first part of a URL identifies what protocol to use as the primary access medium. The second part identifies the IP address or domain name -- and possibly subdomain -- where the resource is located.URL protocols include HTTP (Hypertext Transfer Protocol) and HTTPS (HTTP Secure) for web resources, mail to for email addresses, FTP for files on a File Transfer Protocol (FTP) server, and telnet for a session to access remote computers. Most URL protocols are followed by a colon and two forward slashes; "mail to" is followed only by a colon.

Optionally, after the domain, a URL can also specify:

- a path to a specific page or file within a domain;

- a network port to use to make the connection;
- a specific reference point within a file, such as a named anchor in an HTML file; and
- a query or search parameters used -- commonly found in URLs for search results.

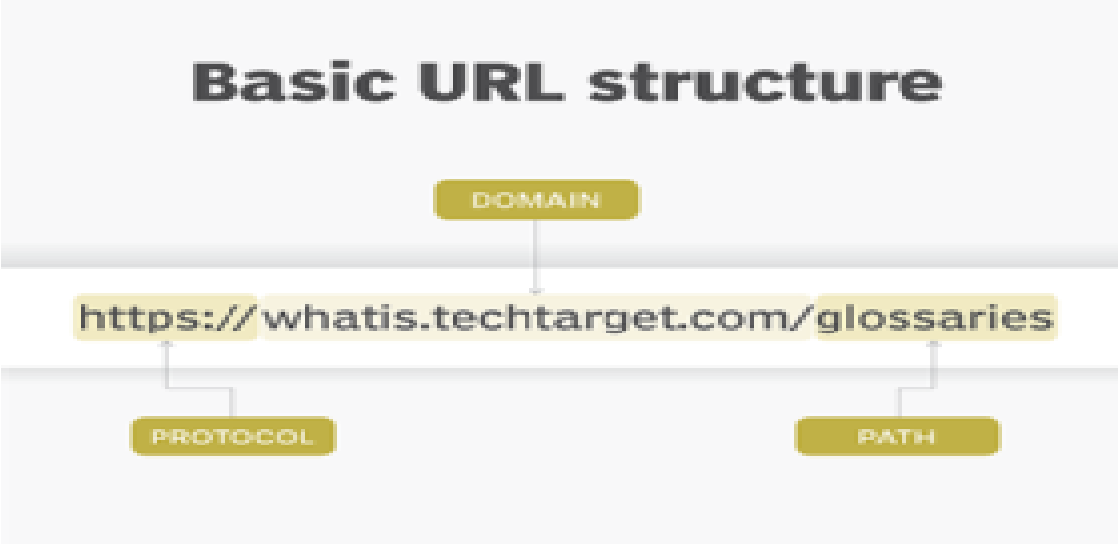
- **Importance of a URL design**

URLs can only be sent over the Internet using the ASCII character-set. Because URLs often contain non-ASCII characters, the URL must be converted into a valid ASCII format. URL encoding replaces unsafe ASCII characters with a "%" followed by two hexadecimal digits. URLs cannot contain spaces.

- **URL examples**

When designing URLs, there are different theories about how to make the syntax most usable for readers and archivists. For example, in the URL's path, dates, authors, and topics can be included in a section referred to as the "slug." Consider, for example, the URL for this definition:

https://www.techtarget.com/searchnetworking/definition/URL



Look past the protocol (identified as HTTPS) and the permalink (www.techtarget.com) and we see the file path includes two paths (search networking and definition) and the title of the definition (URL).

Additionally, some URL designers choose to put the date of the post, typically, as (YYYY/MM/DD).

- **Parts of a URL**

Using the URL **https://www.techtarget.com/whatis/search/query?q=URL** as an example, components of a URL can include:

- **The protocol or scheme.** Used to access a resource on the internet. Protocols include http, https, ftps, mailto and file. The resource is reached through the domain name system (DNS) name. In this example, the protocol is https.
- **Host name or domain name.** The unique reference the represents a webpage. For this example, whatis.techtarget.com.
- **Port name.** Usually not visible in URLs, but necessary. Always following a colon, port 80 is the default port for web servers, but there are other options. For example, port80.
- **Path.** A path refers to a file or location on the web server. For this example, search/query.
- **Query.** Found in the URL of dynamic pages. The query consists of a question mark, followed by parameters. For this example, ?.
- **Parameters.** Pieces of information in a query string of a URL. Multiple parameters can be separated by ampersands (&). For this example, q=URL.
- **Fragment.** This is an internal page reference, which refers to a section within the webpage. It appears at the end of a URL and begins with a hashtag (#). Although not in the example above, an example could be #history in the URL https://en.wikipedia.org/wiki/Internet#History.

Other examples of parts of a URL can include:

- The URL **mailto: president@whitehouse.gov** initiates a new email addressed to the mailbox president in the domain whitehouse.gov.

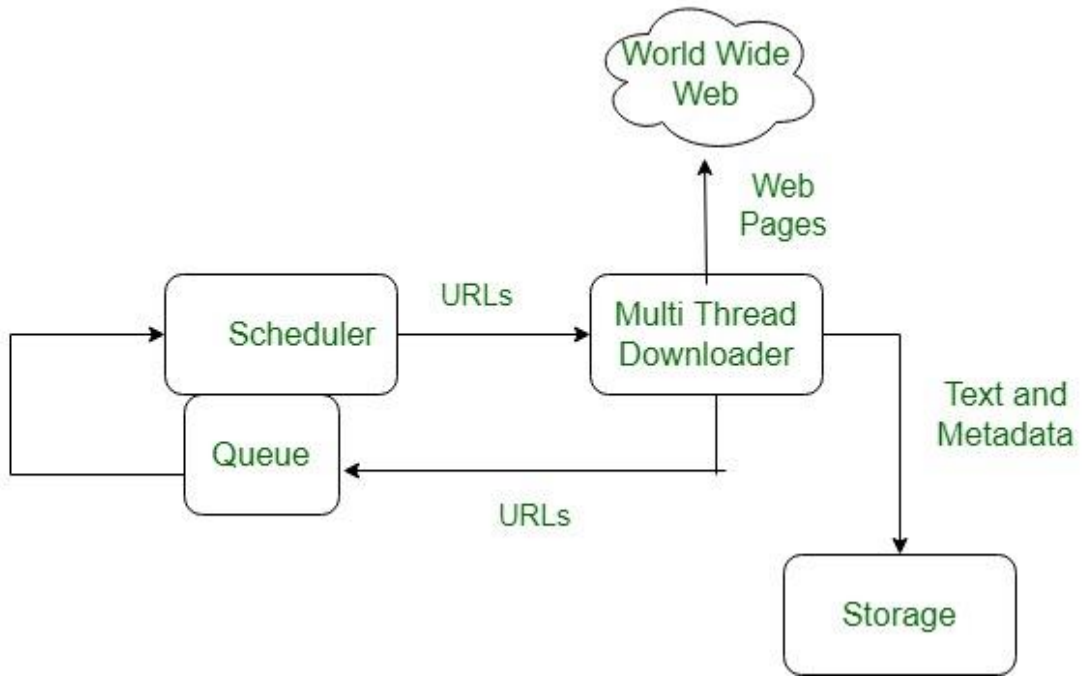
- The URL **ftp://www.companyname.com/whitepapers/widgets.ps** specifies the use of the FTP protocol to download a file.

4.4 WWW (World wide web)

The **World Wide Web** is abbreviated as WWW and is commonly known as the web. The WWW was initiated by CERN (European library for Nuclear Research) in 1989. WWW can be defined as the collection of different websites around the world, containing different information shared via local servers (or computers). It is a project created, by Timothy Berner Lee in 1989, for researchers to work together effectively at CERN. is an organization, named the World Wide Web Consortium (W3C), which was developed for further development of the web. This organization is directed by Tim Berner's Lee, aka the father of the web.

- **System Architecture:**

From the user's point of view, the web consists of a vast, worldwide connection of documents or web pages. Each page may contain links to other pages anywhere in the world. The pages can be retrieved and viewed by using browsers of which internet explorer, Netscape Navigator, Google Chrome, are the popular ones. The browser fetches the page requested interprets the text and formatting commands on it, and displays the page, properly formatted, on the screen. The basic model of how the web works are shown in the figure below. Here the browser is displaying a web page on the client machine. When the user clicks on a line of text that is linked to a page on the abd.com server, the browser follows the hyperlink by sending a message to the abd.com server asking it for the page.



Here the browser displays a web page on the client machine when the user clicks on a line of text that is linked to a page on abd.com, the browser follows the hyperlink by sending a message to the abd.com server asking for the page.

- **Working of WWW:**

The World Wide Web is based on several different technologies: Web browsers, Hypertext Markup Language (HTML) and Hypertext Transfer Protocol (HTTP).
A Web browser is used to access web pages. Web browsers can be defined as programs which display text, data, pictures, animation and video on the Internet. Hyperlinked resources on the World Wide Web can be accessed using software interfaces provided by Web browsers. Initially, Web browsers were used only for surfing the Web but now they have become more universal. Web browsers can be used for several tasks including conducting searches, mailing, transferring files, and much more. Some of the commonly used browsers are Internet Explorer, Opera Mini, and Google Chrome.

- **Features of WWW:**

  - Hypertext Information System
  - Cross-Platform
  - Distributed
  - Open Standards and Open Source

- Uses Web Browsers to provide a single interface for many services
- Dynamic, Interactive and Evolving.
- "Web2.0"

- **Components of the Web:**

There are 3 components of the web:

1. **Uniform Resource Locator (URL):** serves as a system for resources on the web.
2. **Hypertext Transfer Protocol (HTTP):** specifies communication of browser and server.
3. **Hyper Text Markup Language (HTML):** defines the structure, organization and content of a webpage.

4.5 Broadband Transmission

Broadband Transmission is a signalling system that transfers electromagnetic energy-carrying received signals across a wide frequency range. A broadband service's capacity often allows for constant transmission of several data transfers. Broadband offers high-speed internet service through various techniques, including fiber optic cables, connectivity modes, cables, DSL, and satellites.

Broadband describes several high-capacity communication standards that can send data, phone calls, and videos over distant locations and at incredibly fast speeds. It may use multiple communication channels simultaneously. Each data channel is represented by modulation on a particular frequency band for which sending or receiving equipment must be tuned. Dial-up is no longer necessary thanks to broadband, which is always accessible. Its impact is wide-ranging since it enables videoconferencing, data transfer, high-quality and fast accessibility of information, and other things in various contexts like healthcare, learning, and technical advancement.

- **How the Transmission Works**

An Analog signal is how data is transmitted over broadband. Several communications may co-occur because each transmission is given a specific network bandwidth. Since

the broadband transmission is linear, two channels are required to transmit and receive information.

This might be performed by using two wires, one for transmitting and one for providing or by designating frequencies for transmitting and another for receiving very much along the line. Frequency-division combining is possible with broadband.

- **Transmission Techniques**

Various transmission techniques are part of broadband, including:

- Digital Subscriber Line (DSL)
- Cable Modem
- Fiber
- Wireless
- Satellite
- Broadband over Power lines (BPL)

Various elements would influence your decision about broadband technology. Some factors include your location (urban or rural), broadband access to the network is bundled with other things (such as voice calling and entertainment systems), the cost, and usage level.

- **Example**

An undergraduate student, needs a broadband connection because she has to continuously join online classes, research for her assignments, and watch educational videos. The internet connection is much needed for her to have a smooth operation.

To have a Wi-Fi network connection, broadband is a must because through this, data can be transferred to PC, Mobile phones and other devices. It is a single system (wire) that can transport many signals at the same time.

- In broadband, high-speed data transfer, a wired connection can simultaneously transport massive amounts of data.
- Multiple signals are sent on different frequencies using broadband activation.

4.6 Guided Media – Twisted Pair Cable, Coaxial Cable, Fiber-Optic Cable

Networking to be effective, raw stream of data is to be transported from one device to other over some medium. Various transmission media can be used for transfer of data.

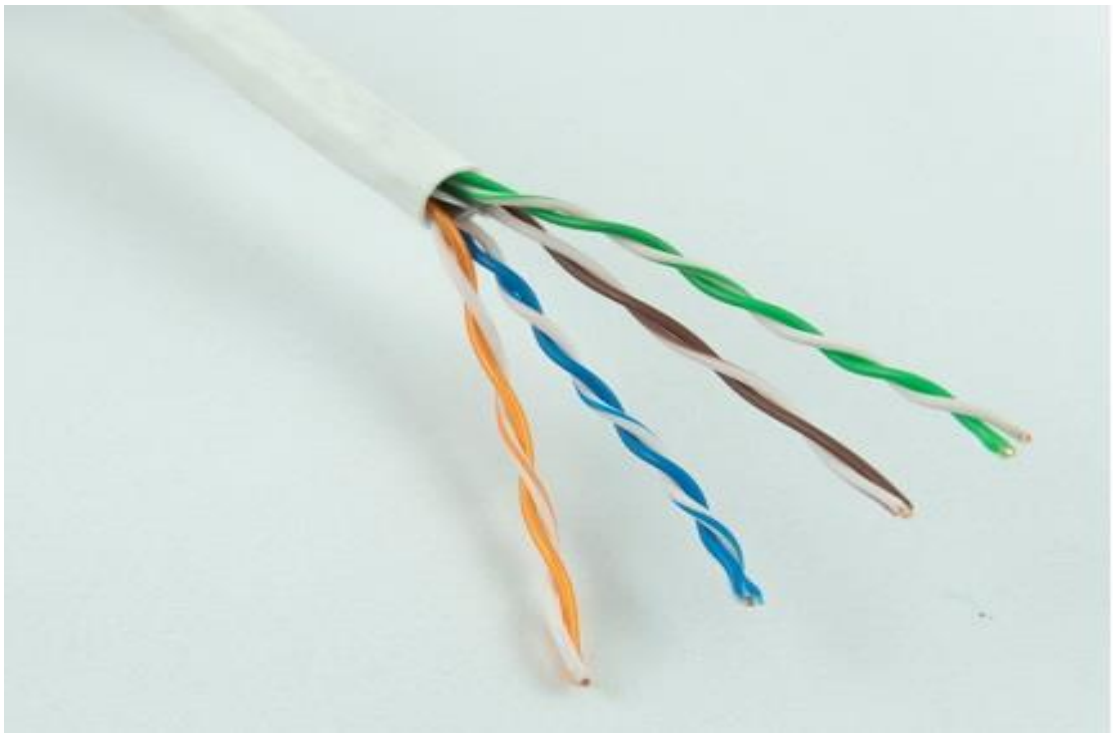
These transmission media may be of two types

- **Guided** − In guided media, transmitted data travels through cabling system that has a fixed path. For example, copper wires, fibre optic wires, etc.
- **Unguided** − In unguided media, transmitted data travels through free space in form of electromagnetic signal. For example, radio waves, lasers, etc.

Each transmission media has its own advantages and disadvantages in terms of bandwidth, speed, delay, cost per bit, ease of installation and maintenance, etc. Let's discuss some of the most commonly used media in detail.

- **Twisted Pair Cable**

Copper wires are the most common wires used for transmitting signals because of good performance at low costs. They are most commonly used in telephone lines. However, if two or more wires are lying together, they can interfere with each other's signals. To reduce this electromagnetic interference, pair of copper wires are twisted together in helical shape like a DNA molecule. Such twisted copper wires are called **twisted pair**. To reduce interference between nearby twisted pairs, the twist rates are different for each pair.



Up to 25 twisted pair are put together in a protective covering to form twisted pair cables that are the backbone of telephone systems and Ethernet networks.

Advantages of twisted pair cable

Twisted pair cable are the oldest and most popular cables all over the world. This is due to the many advantages that they offer

- Trained personnel easily available due to shallow learning curve
- Can be used for both analog and digital transmissions
- Least expensive for short distances
- Entire network does not go down if a part of network is damaged

- **Disadvantages of twisted pair cable**

With its many advantages, twisted pair cables offer some disadvantages too

- Signal cannot travel long distances without repeaters
- High error rate for distances greater than 100m
- Very thin and hence breaks easily
- Not suitable for broadband connections

- **Shielding twisted pair cable**

To counter the tendency of twisted pair cables to pick up noise signals, wires are shielded in the following three ways

- Each twisted pair is shielded.
- Set of multiple twisted pairs in the cable is shielded.
- Each twisted pair and then all the pairs are shielded.

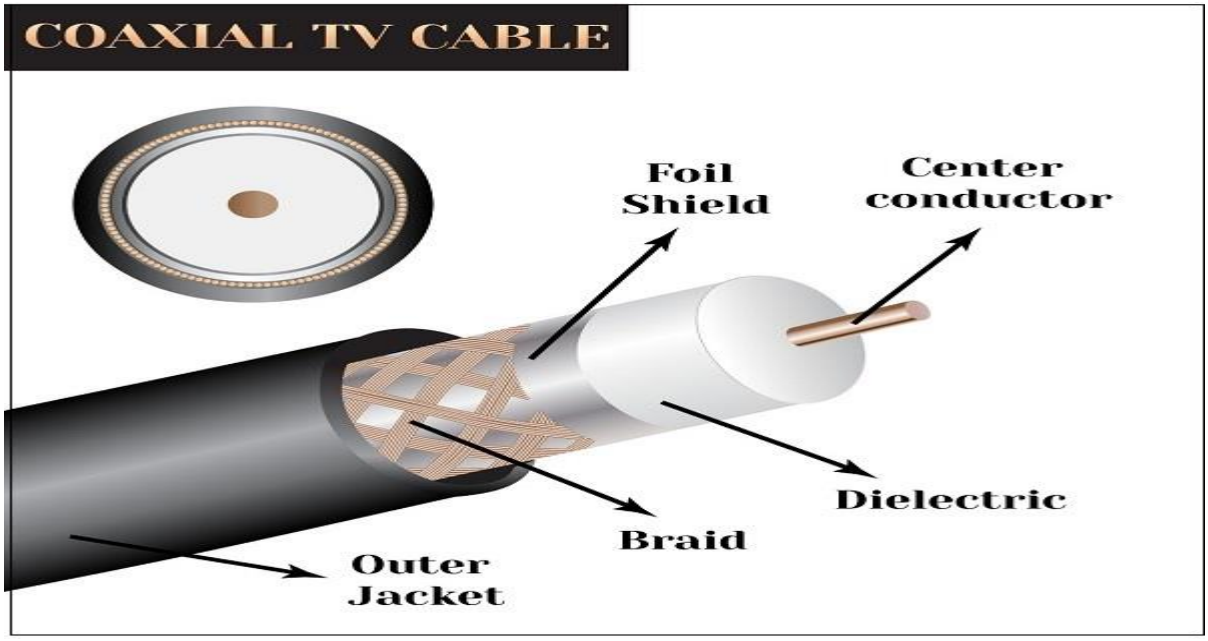
Such twisted pairs are called shielded twisted pair (STP) cables. The wires that are not shielded but simply bundled together in a protective sheath are called unshielded twisted pair (UTP) cables. These cables can have maximum length of 100 metres.Shielding makes the cable bulky, so UTP are more popular than STP. UTP cables are used as the last mile network connection in homes and offices.

- **Coaxial Cable**

Coaxial cables are copper cables with better shielding than twisted pair cables, so that transmitted signals may travel longer distances at higher speeds. A coaxial cable consists of these layers, starting from the innermost

- Stiff copper wire as core
- Insulating material surrounding the core
- Closely woven braided mesh of conducting material surrounding the insulator
- Protective plastic sheath encasing the wire

Coaxial cables are widely used for cable TV connections and LANs.



- **Advantages of Coaxial Cables**

These are the advantages of coaxial cables

- Excellent noise immunity
- Signals can travel longer distances at higher speeds, e.g. 1 to 2 Gbps for 1 Km cable
- Can be used for both analog and digital signals
- Inexpensive as compared to fibre optic cables
- Easy to install and maintain

- **Disadvantages of Coaxial Cables**

These are some of the disadvantages of coaxial cables

- Expensive as compared to twisted pair cables
- Not compatible with twisted pair cables

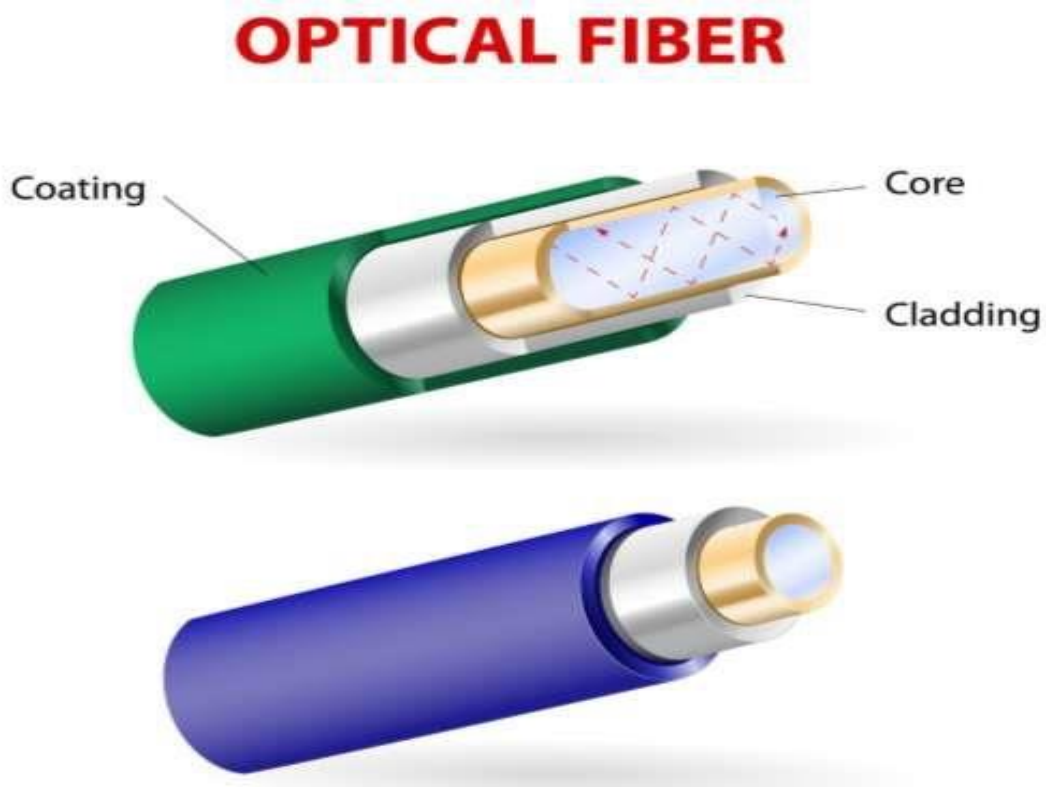
- **Optical Fibre**

Thin glass or plastic threads used to transmit data using light waves are called optical fibre. Light Emitting Diodes (LEDs) or Laser Diodes (LDs) emit light waves at the source, which is read by a detector at the other end. Optical fibre cable has a bundle of such threads or fibres bundled together in a protective covering. Each fibre is made up of these three layers, starting with the innermost layer

- Core made of high quality silica glass or plastic

- Cladding made of high quality silica glass or plastic, with a lower refractive index than the core
- Protective outer covering called buffer

Note that both core and cladding are made of similar material. However, as refractive index of the cladding is lower, any stray light wave trying to escape the core is reflected back due to total internal reflection.



Optical fibre is rapidly replacing copper wires in telephone lines, internet communication and even cable TV connections because transmitted data can travel very long distances without weakening. **Single node** fibre optic cable can have maximum segment length of 2 kms and bandwidth of up to 100 Mbps. **Multi-node** fibre optic cable can have maximum segment length of 100 kms and bandwidth up to 2 Gbps.

- **Advantages of Optical Fibre**

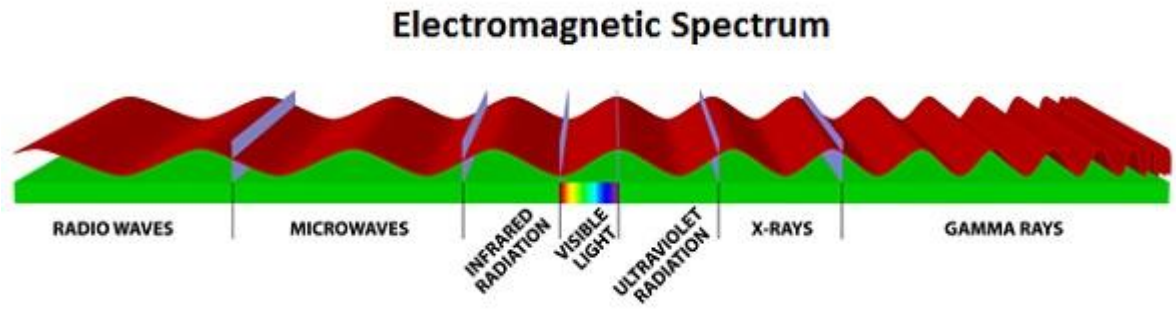
  Optical fibre is fast replacing copper wires because of these advantages that it offers
  - High bandwidth
  - Immune to electromagnetic interference
  - Suitable for industrial and noisy areas
  - Signals carrying data can travel long distances without weakening Disadvantages of Optical Fibre

Despite long segment lengths and high bandwidth, using optical fibre may not be a viable option for every one due to these disadvantages

- Optical fibre cables are expensive
- Sophisticated technology required for manufacturing, installing and maintaining optical fibre cables
- Light waves are unidirectional, so two frequencies are required for full duplex transmission

- **Infrared**

Low frequency infrared waves are used for very short distance communication like TV remote, wireless speakers, automatic doors, hand held devices etc. Infrared signals can propagate within a room but cannot penetrate walls. However, due to such short range, it is considered to be one of the most secure transmission modes.



- **Radio Wave**

Transmission of data using radio frequencies is called radio-wave transmission. We all are familiar with radio channels that broadcast entertainment programs. Radio stations transmit radio waves using transmitters, which are received by the receiver installed in our devices.

Both transmitters and receivers use antennas to radiate or capture radio signals. These radio frequencies can also be used for direct voice communication within the allocated range. This range is usually 10 miles.

- **Advantages of Radio Wave**

These are some of the advantages of radio wave transmissions

- Inexpensive mode of information exchange
- No land needs to be acquired for laying cables
- Installation and maintenance of devices is cheap

- **Disadvantages of Radio Wave**

These are some of the disadvantages of radio wave transmissions

- Insecure communication medium
- Prone to weather changes like rain, thunderstorms, etc.

4.7 SUMMARY

Web Browser A software application used to access information on the World Wide Web is called a Web Browser. When a user requests some information, the web browser fetches the data from a web server and then displays the webpage on the user's screen. ISP stands for Internet Service Provider which is a term used to refer to a company that provides internet access to people who pay the company or subscribe to the company for the same. Networking to be effective, raw stream of data is to be transported from one device to other over some medium. Various transmission media can be used for transfer of data through guided and unguided media.

4.8 KEYWORDS

Web Browser

ISP

WWW

URL

Guided

Unguided

4.9 QUESTION FOR SELF STUDY

1. What is web browser? explain its applications with an example.
2. What is ISP? Explain its advantages and disadvantages.
3. Explain WWW with an example.
4. What is URL and its Applications? explain
5. Explain Broad band transmission.
6. Explain with an example Guided Media – Twisted Pair Cable, Coaxial Cable, Fiber-Optic Cable.

4.10 REFERENCES

1. Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.
2. Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.
3. Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT 5: UNGUIDED TRANSMISSION MEDIA

Structure

5.0 Objectives

5.1 Introduction

5.2 Unguided Media-Wireless, radio waves, microwaves, infrared and Satellite Communication System

5.3 Summary

5.4 Keywords

5.5 Questions for Self Study

5.6 Suggested readings and References

5.0 OBJECTIVES

At the end of this unit you will be able to

- Analyze Transmission Media-Unguided Media-Wireless, radio waves, microwaves, infrared and Satellite Communication System.

5.1 INTRODUCTION

- In this unit we study about the Transmission Media- Unguided Media-Wireless, radio waves, microwaves, infrared and Satellite Communication System.

.

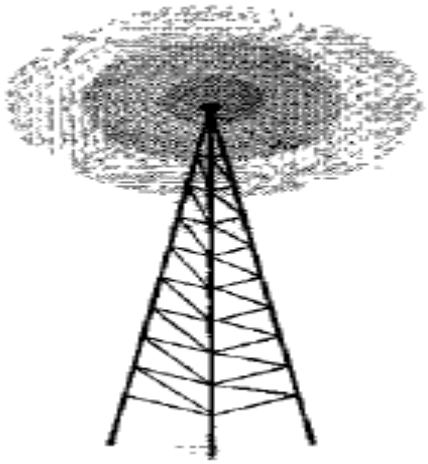
5.2 UNGUIDED MEDIA-WIRELESS, RADIO WAVES, MICROWAVES, INFRARED

Unguided media transport electromagnetic waves without using a physical conductor. This type of communication is often referred to as wireless communication. Signals are normally broadcast through free space and thus are available to anyone who has a device capable of receiving them. Unguided signals can travel from the source to destination in several ways: ground propagation, sky propagation, and line-of-sight propagation.In ground propagation, radio waves travel through the lowest portion of the atmosphere, hugging the earth. These low-frequency signals emanate in all directions

from the transmitting antenna and follow the curvature of the planet. Distance depends on the amount of power in the signal: The greater the power, the greater the distance.In sky propagation, higher-frequency radio waves radiate upward into the ionosphere (the layer of atmosphere where particles exist as ions) where they are reflected back to earth. This type of transmission allows for greater distances with lower output power. In line-or-sight propagation, very high-frequency signals are transmitted in straight lines directly from antenna to antenna. Antennas must be directional, facing each other, and either tall enough or close enough together not to be affected by the curvature of the earth. Line-of-sight propagation is tricky because radio transmissions cannot be completely focused. The section of the electromagnetic spectrum defined as radio waves and microwaves is divided into eight ranges, called *bands,* each regulated by government authorities. These bands are rated from *very low frequency* (VLF) to *extremely high frequency* (EHF).We can divide wireless transmission into three broad groups: radio waves, micro- waves, and infrared waves.

- **Radio Waves**

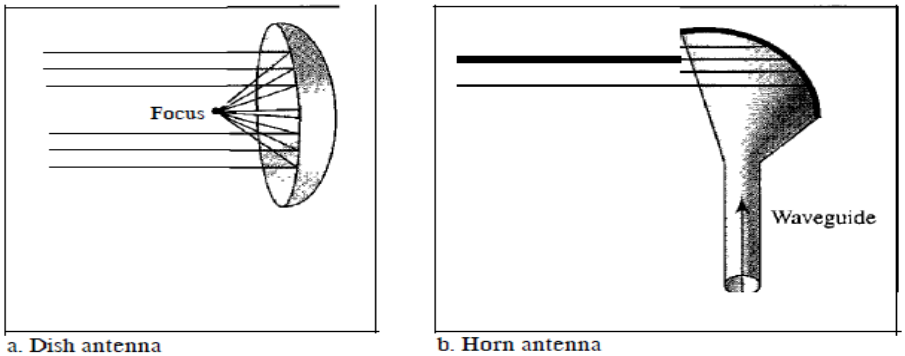
Although there is no clear-cut demarcation between radio waves and microwaves, electromagnetic waves ranging in frequencies between 3 kHz and 1 GHz are normally called radio waves; waves ranging in frequencies between 1 and 300 GHz are called microwaves. Radio waves are omnidirectional. When an antenna transmits radio waves, they are propagated in all directions. This means that the sending and receiving antennas do not have to be aligned. A sending antenna sends waves that can be received by any receiving antenna. The omnidirectional property has a disadvantage, too. The radio waves transmitted by one antenna are susceptible to interference by another antenna that may send signals using the same frequency or band. Radio waves, particularly those waves that propagate in the sky mode, can travel long distances. This makes radio waves a good candidate for long-distance broadcasting such as AM radio.Radio waves, particularly those of low and medium frequencies, can penetrate walls. This characteristic can be both an advantage and a disadvantage. It is an advantage because, for example, an AM radio can receive signals inside a building. It is a disadvantage because we cannot isolate a communication to just inside or outside a building. The radio wave band is relatively narrow, just under 1 GHz, compared to the microwave band.



- **Microwaves**

Electromagnetic waves having frequencies between I and 300 GHz are called microwaves. Microwaves are unidirectional. When an antenna transmits microwave waves, they can be narrowly focused. This means that the sending and receiving antennas need to be aligned. The unidirectional property has an obvious advantage. A pair of antennas can be aligned without interfering with another pair of aligned antennas. The following describes some characteristics of microwave propagation:

- Microwave propagation is line-of-sight. Since the towers with the mounted antennas need to be in direct sight of each other, towers that are far apart need to be very tall.
- Very high-frequency microwaves cannot penetrate walls. This characteristic can be a disadvantage if receivers are inside buildings.
- The microwave band is relatively wide, almost 299 GHz. Therefore, wider subbands can be assigned, and a high data rate is possible
- Use of certain portions of the band requires permission from authorities.



a. Dish antenna b. Horn antenna

- **Infrared**

Infrared waves, with frequencies from 300 GHz to 400 THz (wavelengths from 1 mm to 770 nm), can be used for short-range communication. Infrared waves, having high frequencies, cannot penetrate walls. This advantageous characteristic prevents interference between one system and another; a short-range communication system in one room cannot be affected by another system in the next room. When we use our infrared remote control, we do not interfere with the use of the remote by our neighbors. However, this same characteristic makes infrared signals useless for long-range communication. In addition, we cannot use infrared waves outside a building because the sun's rays contain infrared waves that can interfere with the communication.

- **Satellite communications are comprised of 2 main components:**

The satellite itself is also known as the space segment, and is composed of three separate units, namely the fuel system, the satellite and telemetry controls, and the transponder. The transponder includes the receiving antenna to pick-up signals from the ground station, a broad band receiver, an input multiplexer, and a frequency converter which is used to reroute the received signals through a high powered amplifier for downlink. The primary role of a satellite is to reflect electronic signals. In the case of a telecom satellite, the primary task is to receive signals from a ground station and send them down to another ground station located a considerable distance away from the first. This relay action can be two-way, as in the case of a long distance phone call. Another use of the satellite is when, as is the case with television broadcasts, the ground station's uplink is then down linked over a wide region, so that it may be received by many different customers possessing compatible equipment. Still another use for satellites is observation, wherein the satellite is equipped with cameras or various sensors, and it merely downlinks any information it picks up from its vantage point.

- **The Ground Station.**

This is the earth segment. The ground station's job is two-fold. In the case of an uplink, or transmitting station, terrestrial data in the form of baseband signals, is passed through a baseband processor, an up converter, a high powered amplifier, and through a parabolic dish antenna up to an orbiting satellite. In the case of a downlink,

or receiving station, works in the reverse fashion as the uplink, ultimately converting signals received through the parabolic antenna to base band signal. Communications satellite is station in the space that receives microwaves signals from an earth-based station, amplifies the signals, and broadcasts the signals back over a wide area to many earth-based stations. Communications satellites are usually placed about 22,300 miles above the Earth's equator and moves at the same rate as the Earth. Applications of communications satellite include television and radio broadcasts, videoconferencing, paging, and global positioning systems.

- **Advantages of satellites**

Lots of data can be sent simultaneously.

Allow high quality broadband communication across continents.

- **Disadvantages of satellites**

The fee to launch a satellite is extremely expensively.

The infrastructure needed to access satellite communications is also expensive.

5.3 SUMMARY

In this unit, we discussed about the Analyze Transmission Media-guided media, twisted-pair cable, coaxial cable, fiber optic cable. Also studied Unguided Media-Wireless, radio waves, microwaves, infrared.

5.4 KEYWORDS

Guided, Unguided, Infrared, Radio waves, micro waves

5.5 QUESTIONS FOR SELF STUDY

1. What are the different wireless transmissions medium?
2. Which transmission media would have greater start-up cost, hardwiring or wireless transmission media?
3. Explain different types of wireless transmission media? in detail and also compare them
4. Write a note on Unguided Media-Wireless, radio waves, microwaves, infrared and Satellite.

5.6 Reference

1. Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.

2. Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.
3. Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT -6: NETWORK MODELS

Structure

6.0 Objectives

6.1 Introduction

6.2 The OSI model

6.3 Layers in the OSI model

6.4 Check your progress

6.5 Summary

6.6 Keywords

6.7 Questions for self-study

6.8 References

6.0 OBJECTIVES

After studying this unit, you will be able to

- ✓ Explain the layered architecture of the OSI model.
- ✓ Describe the peer to peer processes.
- ✓ Learn the concept of encapsulation
- ✓ Differentiate the roles and functions of the layers in the OSI model.

6.1 INTRODUCTION

A network is a combination of hardware and software that sends data from one location to another. The hardware consists of the physical equipment that carries signals from one point of the network to another. The software consists of instruction sets that make possible the services that we expect from a network.

The layered model that dominated data communications and networking literature before 1990 was the Open Systems Interconnection (OSI) model. Everyone believed that the OSI model would become the ultimate standard for data communications, but this did not happen. The TCP/IP protocol suite became the dominant commercial architecture because it was used and tested extensively in the Internet; the OSI model was never fully implemented. In this unit, we give the overview of the OSI model and the layers in it; discuss the functions of each layer.

6.2 THE OSI MODEL

The 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. The model is called the ISO **OSI** (**Open Systems Interconnection**) Reference Model because it deals with connecting open systems. An open system is a set of protocols that allows any two different systems to communicate regardless of their underlying architecture. The OSI model is not a protocol; it is a model for understanding and designing a network architecture that is flexible, robust, and interoperable.

6.2.1 Layered Architecture:

The OSI model is a layered framework for the design of network systems that allow communication between all types of computer systems. It consists of seven separate but related layers; they are - Physical, Data link, Network, Transport, Session, Presentation, and Application. The layers are ordered as shown in the figure.

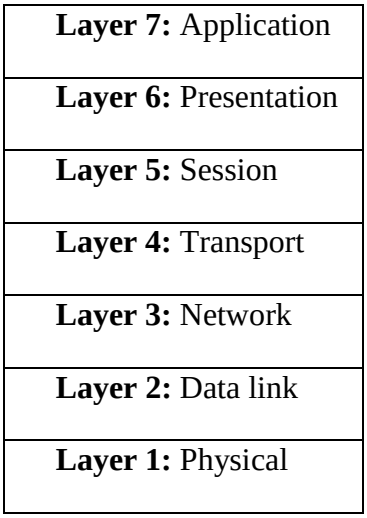


Figure 6.1: Layers of the OSI model

Figure 6.1 shows the layers involved when a message is sent from device A to device B. As the message travels from A to B, it may pass through many intermediate nodes. These intermediate nodes usually involve only the first three layers of the OSI model.

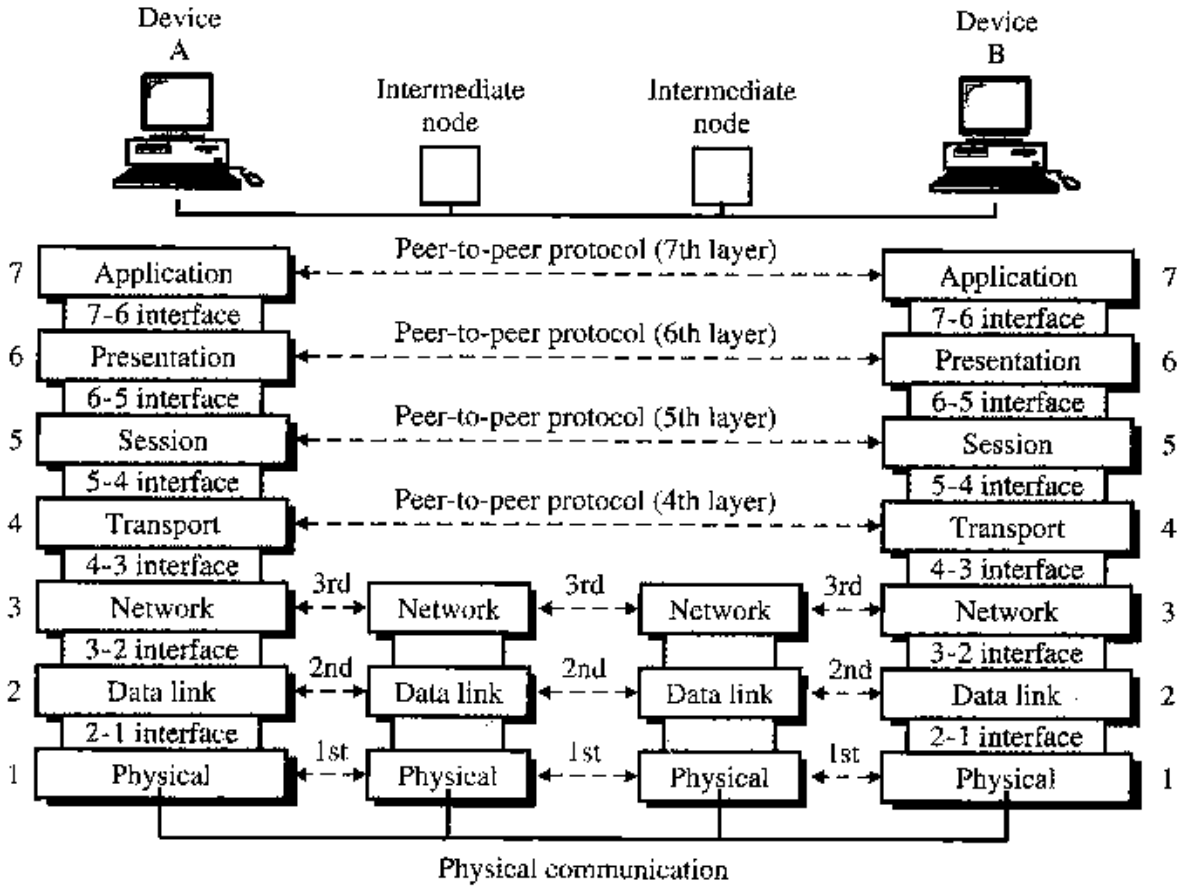


Figure 6.2: Interaction between layers in OSI model

In developing the model, the related functions were collected into discrete groups that became the layers. Each layer defines a family of functions distinct from those of the other layers. Thus the layered architecture is both comprehensive and flexible. Within a single machine, each layer calls upon the services of the layer just below it. Between machines, layer *x* on one machine communicates with layer *x* on another machine. This communication is governed by an agreed-upon set of rules and conventions called protocols. The processes on each machine that communicate at a given layer are called peer-to-peer processes. Communication between machines is therefore a peer-to-peer process using the protocols appropriate to a given layer.

6.2.2 Peer-to-Peer Processes:

At the physical layer, communication is direct: In Figure 6.2.2, device A sends a stream of bits to device B (through intermediate nodes). At the higher layers, however, communication must move down through the layers on device A, over to device B, and then back up through the layers. Each layer in the sending device adds its own information to the message it receives from the layer just above it and passes the whole package to the layer just below it.

At layer 1 the entire package is converted to a form that can be transmitted to the receiving device. At the receiving machine, the message is unwrapped layer by layer, with each process receiving and removing the data meant for it. For example, layer 2 removes the data meant for it, and then passes the rest to layer 3. Layer 3 then removes the data meant for it and passes the rest to layer 4, and so on.

- **Interfaces between Layers:**

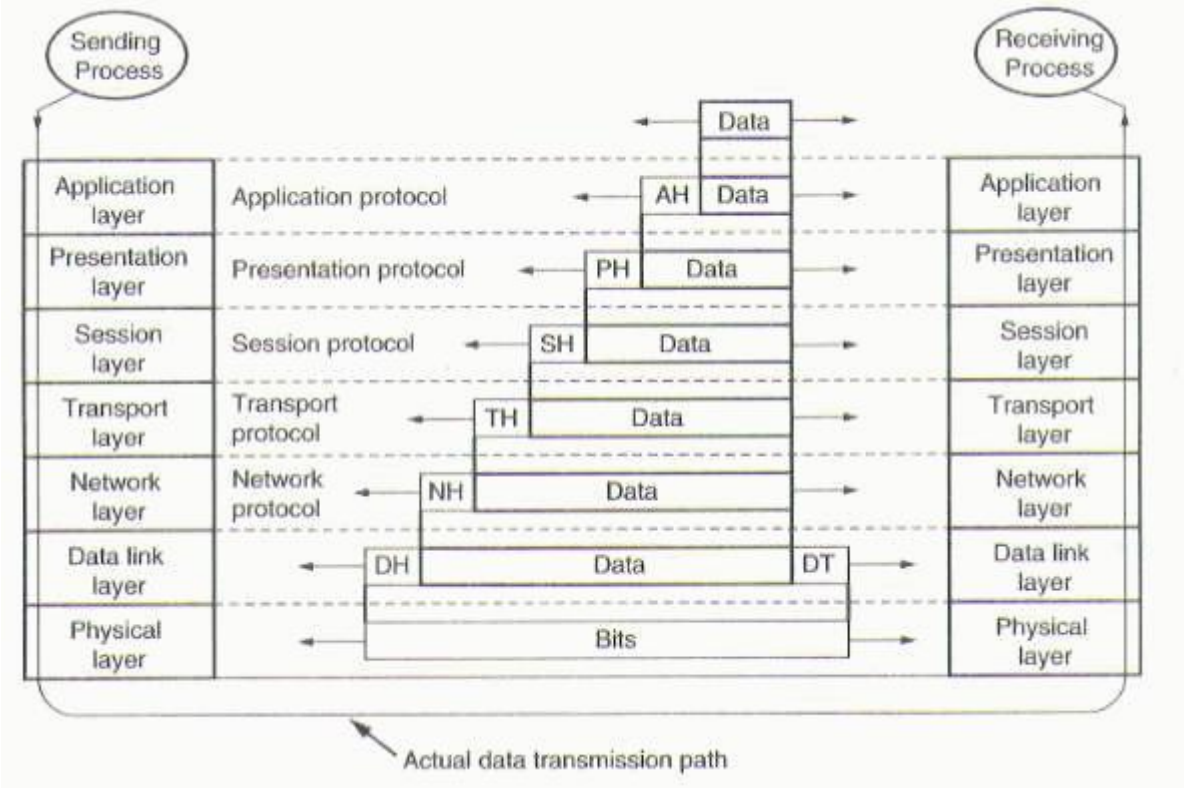
The passing of the data and network information down through the layers of the sending device and back up through the layers of the receiving device is made possible by an interface between each pair of adjacent layers. Each interface defines the information and services, a layer must provide for the layer above it. Well-defined interfaces and layer functions provide modularity to a network. As long as a layer provides the expected services to the layer above it, the specific implementation of its functions can be modified or replaced without requiring changes to the surrounding layers.

- **Organization of the Layers:**

The seven layers can be thought of as belonging to three subgroups. Layers 1, 2, and 3-physical, data link, and network-are the network support layers; they deal with the physical aspects of moving data from one device to another such as electrical specifications, physical connections, physical addressing, and transport timing and reliability. Layers 5, 6, and 7-session, presentation, and application-can be thought of as the user support layers; they allow interoperability among unrelated software systems. Layer 4, the transport layer, links the two subgroups and ensures that the data the lower layers have transmitted is in a form that the upper layers can use.

The upper OSI layers are almost always implemented in software; lower layers are a combination of hardware and software, except for the physical layer, which is mostly hardware.

Figure 6.3, gives an overall view of the OSI layers. The process starts at layer 7 (the application layer), then moves from layer to layer in descending, sequential order. At each layer, a **header,** or possibly a **trailer,** can be added to the data unit. Commonly, the trailer is added only at layer 2. When the formatted data unit passes through the physical layer (layer 1), it is changed into an electromagnetic signal and transported along a physical link.



.

Figure 6.3: Exchange using the OSI model

Upon reaching its destination, the signal passes into layer 1 and is transformed back into digital form. The data units then move back up through the OSI layers. As each block of data reaches the next higher layer, the headers and trailers attached to it at the corresponding sending layer are removed, and actions appropriate to that layer are taken. By the time it reaches layer 7, the message is again in a form appropriate to the application and is made available to the recipient.

6.2.3 Encapsulation:

Encapsulation is another aspect of data communications in the OSI model as revealed in Figure 6.3. A packet containing header and data at level 7 is encapsulated in a packet at level 6. The whole packet at level 6 is encapsulated in a packet at level 5, and so on. In other words, the data portion of a packet at level *N* - 1 carries the whole packet containing data, header and (maybe) trailer from level *N*. The concept is called *encapsulation.* The level *N* - 1 is not aware of which part of the encapsulated packet is data and which part is the header or trailer. For level *N* - 1, the whole packet coming from level *N* is treated as one integral unit.

6.3 LAYERS IN THE OSI MODEL

In this section we briefly describe the functions of each layer in the OSI model.

6.3.1 Physical Layer:

The physical layer is the lowest layer of the OSI model, 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.

The physical layer is concerned with the following:

1. Defining the characteristics of the interface between the devices and the transmission medium. It also defines the type of transmission medium.
2. To transmit a stream of bits which is a sequence of 0s and 1s, must be encoded into signals--electrical or optical. The physical layer defines the type of encoding.
3. The physical layer defines the duration of a bit, which is how long it lasts.
4. The sender and receiver not only must use the same bit rate but also must be synchronized at the bit level.
5. The physical layer is concerned with the connection of devices to the media.
6. The various physical topologies that devices are connected to make a network.
7. Defining the direction of transmission between two devices: simplex, half-duplex, or full-duplex.

6.3.2 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.

The responsibilities of the data link layer include the following:

1. Divides the stream of bits received from the network layer into manageable data units called frames.
2. Adding a header to the frame to define the sender and receiver of the frame, if the frame is to be distributed to different systems on the network.
3. To impose a flow control mechanism to avoid overwhelming the receiver, if the data receiving rate at the receiver is less than the data production rate at the sender.
4. It is responsible for the detection and retransmission of damaged or lost frames, to recognize duplicate frames. Thus controls the error in transmission.

5. When two or more devices are connected to the same link, data link layer protocols are necessary to determine which device has control over the link at any given time.

6.3.3 Network Layer:

The network layer is responsible for the delivery of a packet from source-to-destination, possibly across multiple networks. A need for the network layer to accomplish source-to-destination delivery arises whenever the two systems are attached to different networks.

The responsibilities of the network layer include the following:

1. The physical addressing is implemented by the data link layer that addresses the problem locally. To distinguish the source and destination systems, the network layer adds the logical addresses of the sender and receiver.
2. Responsible for the mechanism of routing the packets to their final destination.
3. Has accounting functions to keep track of frames forwarded by subnet intermediate systems, to produce billing information.
4. Routers can instruct a sending station to "throttle back" its frame transmission when the router's buffer fills up.
5. 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-assemble at the destination station.

6.3.4 Transport Layer:

The transport layer ensures that messages are delivered error-free, in sequence, and with no losses or duplications. It relieves 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 datagram, the transport protocol should include extensive error detection and recovery. The transport layer provides:

1. Service point addressing: Source-to-destination delivery means delivery from a specific process (running program) on one computer to a specific process (running program) on the other. So the transport layer header includes a type

of address called a *service-point address* (or port address). The network layer gets each packet to the correct computer; the transport layer gets the entire message to the correct process on that computer.

2. Message segmentation: Accepts a message from the session layer, splits the message into smaller units, and passes the smaller units down to the network layer. The transport layer at the destination station reassembles the message.
3. Message acknowledgment: Provides reliable end-to-end message delivery with acknowledgments.
4. Flow control: The transport layer is responsible for end to end flow control.
5. Error control. Like the data link layer, the transport layer is responsible for error control. However, error control at this layer is performed process-to process rather than across a single link. The sending transport layer makes sure that the entire message arrives at the receiving transport layer without error (damage, loss, or duplication). Error correction is usually achieved through retransmission.

6.3.5 Session Layer:

The services provided by the first three layers are not sufficient for some processes. The session layer is the network *dialog controller*. It establishes, maintains, and synchronizes the interaction among different communicating systems.

Specific responsibilities of the session layer include the following:

1. Dialog control. The session layer allows two systems to enter into a dialog. It allows the communication between two processes to take place in either half duplex or full-duplex mode.
2. Synchronization. The session layer allows a process to add checkpoints, or synchronization points, to a stream of data.

6.3.6 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 responsibilities of the presentation layer include:

Translation: The presentation layer at the sender changes the information from its sender-dependent format into a common format. The presentation layer at the receiving machine changes the common format into its receiver-dependent format.

Encryption: The layer transforms the original message to another form and sends the resulting message out over the network. Decryption reverses the original process to transform the message back to its original form.Compression: Data compression becomes particularly important in the transmission of multimedia such as text, audio, and video. The layer compresses the data before transmission.

6.3.7 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 such as electronic mail, remote file access and transfer, shared database management, and other types of distributed information services.

Specific services provided by the application layer include the following:

1. Network virtual terminal. A network virtual terminal is a software version of a physical terminal, and it allows a user to log on to a remote host. The remote host believes it is communicating with one of its own terminals and allows the user to log on.
2. File transfer, access, and management. This application allows a user to access files in a remote host, to retrieve files from a remote computer for use in the local computer, and to manage or control files in a remote computer locally.
3. Mail services. This application provides the basis for e-mail forwarding and storage.
4. Directory services. This application provides distributed database sources and access for global information about various objects and services.

6.4 CHECK YOUR PROGRESS

1. How many layers are there in the OSI model?
2. **Which layer is responsible for converting data into electrical signals?**
3. The layer which establishes, maintains, and terminates communications between applications located on different devices is _____

4. _____**layer is responsible for the** routing of the data packets from the source to destination.

5. **Which layer defines how data is formatted, presented, encoded, and converted for use on the network?**
6. The _____ layer links the network support layers and the user support layers.
7. The application layer is a _____ support layer.
8. The data link layer is a _____ support layer.
9. Interface defines the _____and _____ a layer must provide for the layer above it.

Answers:

1. Seven
2. Physical layer
3. Session
4. Network
5. Presentation layer
6. Transport
7. User
8. Network
9. Information, services

6.5 SUMMARY

The OSI model is based on a proposal developed by the ISO for the standardization of the protocols. The OSI model provides guidelines for the development of universally compatible networking protocols. The model has seven layers. The upper layers (5-7) of the OSI model deal with application issues and generally are implemented only in software. The highest layer, the application layer, is closest to the end-user. Both users and application layer processes interact with software applications that contain a communications component. The lower layers (1-4) of the OSI model handle data transport issues. The physical layer and the data link layer are implemented in hardware and software. The lowest layer, the physical layer, is closest to the physical network medium and is responsible for actually placing information on the medium.

6.6 KEYWORDS

Layer, protocol, interface, OSI model - physical, data link, network, transport, session, presentation, and application layers.

6.7 QUESTIONS FOR SELF STUDY

1. Which layers in the OSI model are the network support layers?
2. Which layer in the OSI model is the user support layer?
3. What is the difference between network layer delivery and transport layer delivery?
4. What is a peer-to-peer process?
5. How does information get passed from one layer to the next in the OSI model?
6. What are headers and trailers, and how do they get added and removed?
6. What are the concerns of the physical layer in the OSI model?
7. What are the responsibilities of the data link layer in the OSI model?
8. What are the responsibilities of the network layer in the OSI model?
9. What are the responsibilities of the transport layer in the OSI model?
10. What is the difference between a port address, a logical address, and a physical address?
11. Name some services provided by the application layer in the OSI model.

6.8 REFERENCES

1. Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.
2. Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.
3. Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT 7: CONNECTING DEVICES (Networking Devices)

Structure:

7.0 Objectives

7.1 Introduction

7.2 Connecting devices

7.3 Passive Hubs and active hubs

7.4 Bridges

7.5 Two layer switches

7.6 Gateway

7.7 Summary

7.8 Keywords

7.9 Questions

7.10 References

7.0 OBJECTIVES

After studying this unit, you will be able to

- Name the connecting devices used in a network
- Explain router
- Describe the functions of bridges, gateway

7.1 INTRODUCTION

LANs do not normally operate in isolation. They are connected to one another or to the Internet. To connect LANs, or segments of LANs, we use connecting devices. Connecting devices can operate in different layers of the Internet model. Here, we discuss only those that operate in the physical and data link layers.

7.2 CONNECTING DEVICES

we divide **connecting devices** into five different categories based on the layer in which they operate in a network

The five categories contain devices which can be defined as

1. Those which operate below the physical layer such as a passive hub.

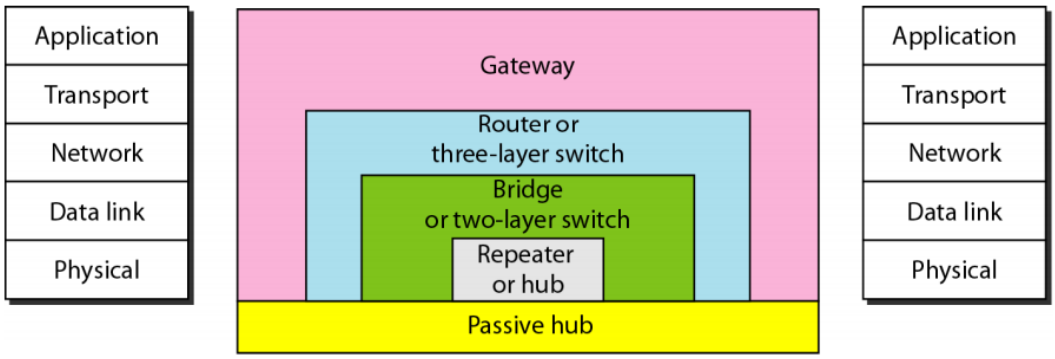
2. Those which operate at the physical layer (a repeater or an active hub).

3. Those which operate at the physical and data link layers (a bridge or a two-layer switch).

4. Those which operate at the physical, data link, and network layers (a router or a

three-layer switch).

5. Those which can operate at all five layers (a gateway).

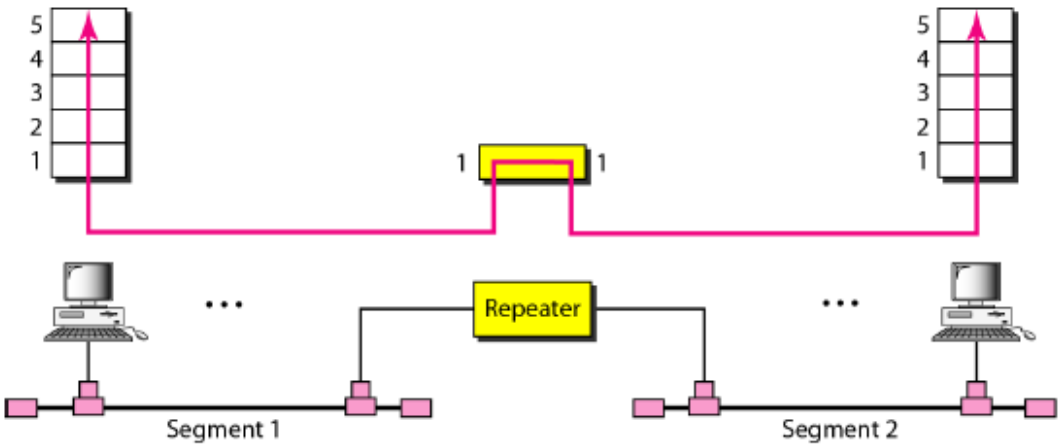


7.3 PASSIVE HUBS

A passive hub is just a connector. It connects the wires coming from different branches. In a star-topology Ethernet LAN, a passive hub is just a point where the signals coming from different stations collide; the hub is the collision point. This type of a hub is part of the media; its location in the Internet model is below the physical layer.

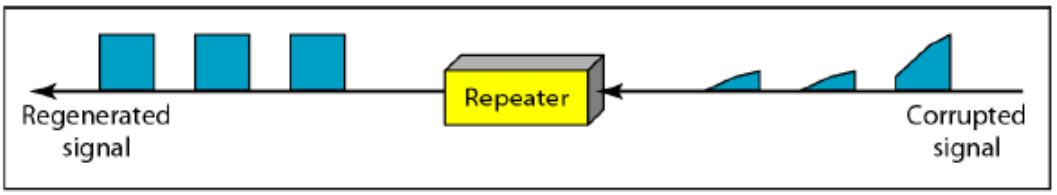
- **Repeaters**

A repeater is a device that operates only in the physical layer. Signals that carry information within a network can travel a fixed distance before attenuation endangers the integrity of the data. A repeater receives a signal and, before it becomes too weak or corrupted, regenerates the original bit pattern. The repeater then sends the refreshed signal. A repeater can extend the physical length of a LAN, as shown in Figure .

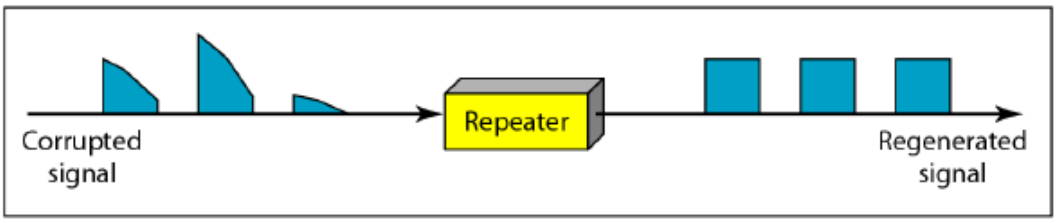


A repeater does not actually connect two LANs; it connects two segments of the same LAN. The segments connected are still part of one single LAN. A repeater is not a device that can connect two LANs of different protocols. The location of a repeater on a link is vital. A repeater must be placed so that a signal reaches it before any noise changes the meaning of any of its bits. A little noise can alter the precision of a bit's voltage without destroying its identity (see Figure 15.3). If the corrupted bit travels much farther, however, accumulated noise can change its meaning completely. At that point, the original voltage is not recoverable, and the error needs to be corrected. A repeater placed on the line before the legibility of the signal becomes lost can still read the signal well enough to determine the intended voltages and replicate them in their original form.

.



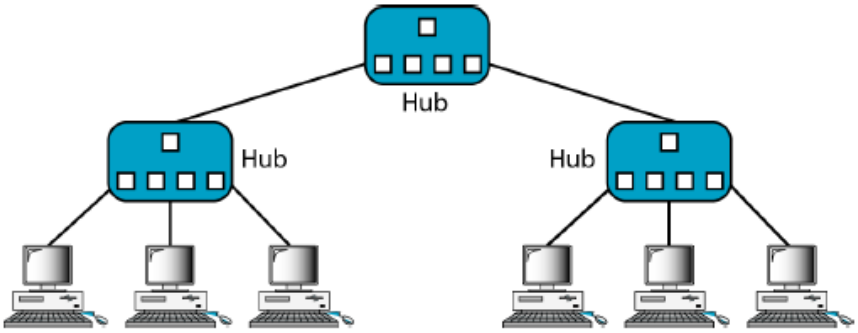
a. Right-to-left transmission.



b. Left-to-right transmission.

- **ACTIVE HUBS**

An active hub is actually a multipart repeater. It is normally used to create connections between stations in a physical star topology. We have seen examples of hubs in some Ethernet implementations (lOBase-T, for example). However, hubs can also be used to create multiple levels of hierarchy, as shown in Figure 15.4. The hierarchical use of hubs removes the length limitation of 10Base-T (100 m).

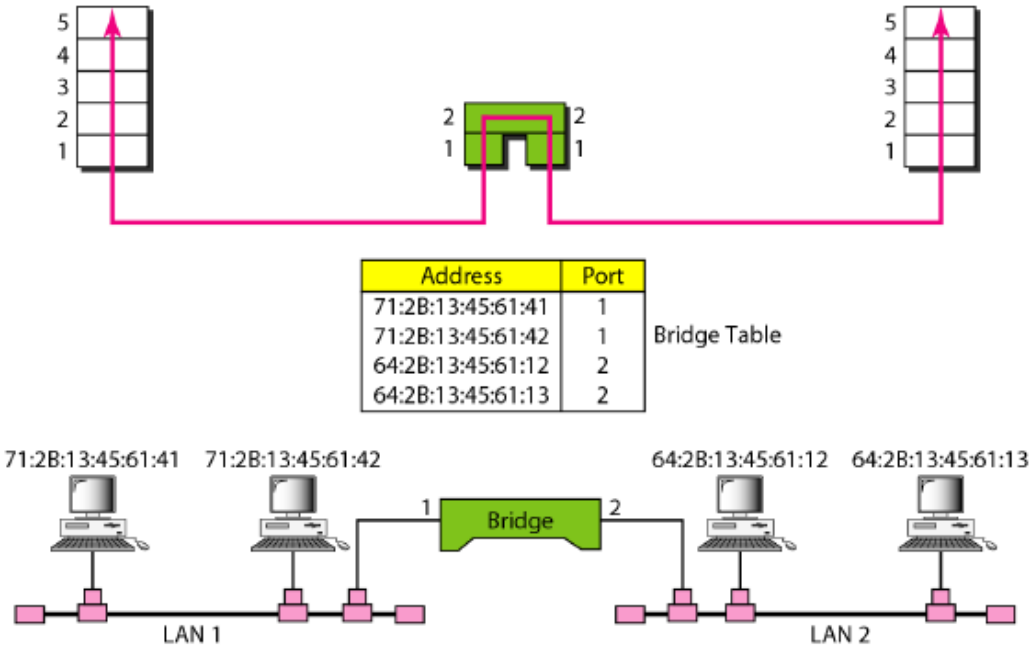


7.4 BRIDGES

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

Filtering

One may ask, What is the difference in functionality between a bridge and a repeater? A bridge has filtering capability. It can check the destination address of a frame and decide if the frame should be forwarded or dropped. If the frame is to be forwarded, the decision must specify the port. A bridge has a table that maps addresses to ports. A bridge has a table used in filtering decisions. Let us give an example. In Figure, two LANs are connected by a bridge. If a frame destined for station 712B13456142 arrives at port 1, the bridge consults its table to find the departing port. According to its table, frames for 7l2B13456142 leave through port 1; therefore, there is no need for forwarding, and the frame is dropped. On the other hand, if a frame for 712B13456141 arrives at port 2, the departing port is port 1



and the frame is forwarded. In the first case, LAN 2 remains free of traffic; in the second case, both LANs have traffic. In our example, we show a two-port bridge; in reality a bridge usually has more ports.

Note also that a bridge does not change the physical addresses contained in the frame.

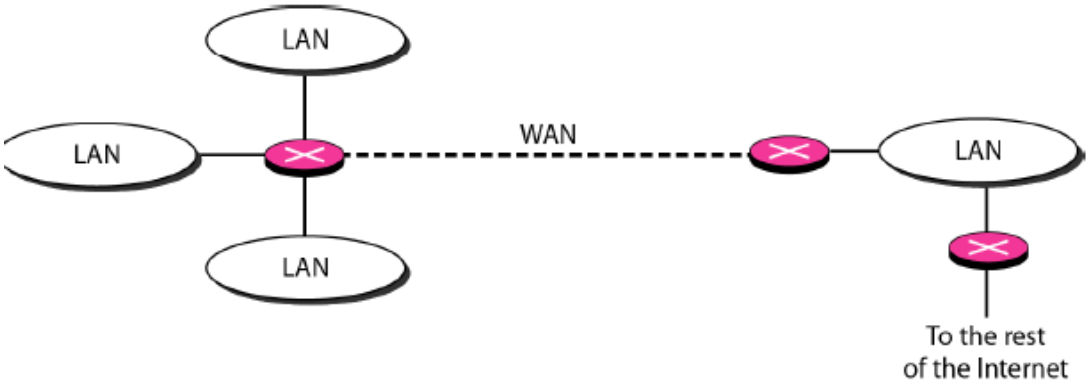
7.5 TWO-LAYER SWITCHES

When we use the term *switch*, we must be careful because a switch can mean two different things. We must clarify the term by adding the level at which the device operates.We can have a two-layer switch or a three-layer switch. A **three-layer switch** is used at the network layer; it is a kind of router. The **two-layer switch** performs at the physical and data link layers. A two-layer switch is a bridge, a bridge with many ports and a design that allows better (faster) performance. A bridge with a few ports can connect a few LANs together. A bridge with many ports may be able to allocate a unique port to each station, with each station on its own independent entity. This means no competing traffic (no collision, as we saw in Ethernet). A two-layer switch, as a bridge does, makes a filtering decision based on the MAC address of the frame it received. However, a two-layer switch can be more sophisticated. It can have a buffer to hold the frames for processing. It can have a switching factor that forwards the frames faster. Some new two-layer switches, called *cut-through* switches, have

been designed to forward the frame as soon as they check the MAC addresses in the header of the frame.

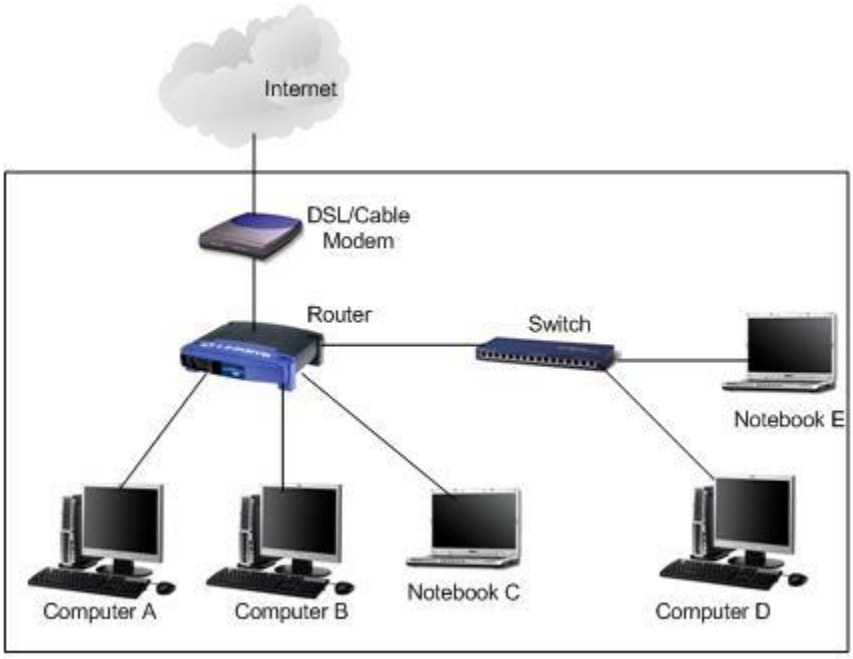
- **ROUTERS**

A router is a three-layer device that routes packets based on their logical addresses (host-to-host addressing). A router normally connects LANs and WANs in the Internet and has a routing table that is used for making decisions about the route. The routing tables are normally dynamic and are updated using routing protocols. We discuss routers and routing in greater detail in Chapters 19 and 21. Figure 15.11 shows a part of the Internet that uses routers to connect LANs and WANs.



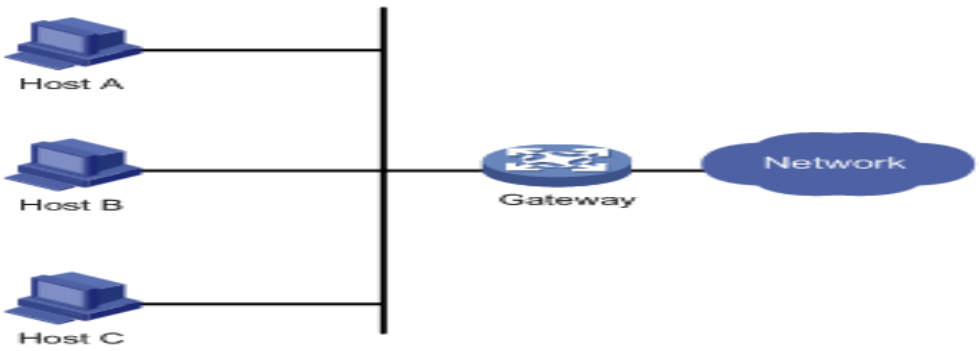
- **THREE-LAYER SWITCHES**

A three-layer switch is a router, but a faster and more sophisticated. The switching fabric in a three-layer switch allows faster table lookup and forwarding. In this book, we use the terms *router* and *three-layer switch* interchangeably.



7.6 GATEWAY

Although some textbooks use the terms *gateway* and *router* interchangeably, most of the literature distinguishes between the two. A gateway is normally a computer that operates in all five layers of the Internet or seven layers of OSI model. A gateway takes an application message, reads it, and interprets it. This means that it can be used as a connecting device between two internetworks that use different models. For example, a network designed to use the OSI model can be connected to another network using the Internet model. The gateway connecting the two systems can take a frame as it arrives from the first system, move it up to the OSI application layer, and remove the message.



7.7 SUMMARY

In this unit, we study about the connecting devices used in the network to connect devices within the network, between the networks of same type and between the

networks of different type.

7.8 KEYWORDS

Router, Bridge, Switch, Hub, Gateway

7.9 QUESTIONS

1. Name the connecting devices used in the network.
2. Distinguish between active and passive hub.
3. What are the functions of repeater?
4. Define router.
5. What is a gateway?
6. Write the functions of bridge.

7.10 REFERENCES

1. Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.
2. Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.
3. Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT 8: IP ADDRESSING AND PROTOCOLS

Structure

8.0 Objectives
8.1 Introduction
8.2 TCP, IP, UDP,IPV4,IPV6
8.3 Summary/Let us Sum up
8.4 Keywords
8.5 Questions for Self Study
8.6 Suggested readings and References

8.0 OBJECTIVES

At the end of this unit you will be able to understand:

- TCP, IP, UDP, IPV4, IPV6.

8.1 INTRODUCTION

In this unit, we will study about the TCP, IP, UDP, IPV4, IPV6.

8.2 TCP, IP, UDP, IPV4, IPV6.

IP: Internet Protocol:

IP addresses are an essential part of computer networking. They play an important role in sending and receiving information on the internet. Every device that connects to an internet network has an IP address, which means there are billions of IP addresses that exist. In this unit, we're going to discuss IP addresses, how they work, static IP vs dynamic IP, and more.

- **What is an IP address?**

An Internet Protocol address, or IP address, is a **unique identifier** assigned to every device on a TCP/IP network. The Internet Protocol is the set of rules that outlines how data should be transported across the internet or local networks. IP addresses help **identify devices and allow them to communicate** with each other. Internally, IP addresses are stored as numbers. The Domain Name System (DNS) allows us to use words to identify different servers on the internet, such as an application, server, or website. When we type a URL into our search bar, DNS looks up that domain's IP address and returns it to our network device. There are two main versions of IP addresses: IPv4 and IPv6.

IP addresses are strings of four numbers separated by characters. For example, an IP address could look like this: 152.132.4.23. They're produced by a division of the Internet Corporation for Assigned Names and Numbers to help make the internet more secure and accessible. We can think of IP addresses as physical home addresses. We can exchange addresses with friends and family. Those addresses give us a destination that allows us to communicate with friends and family through different communication methods like birthday cards, letters, and more.

- **How does an IP address work?**

All devices communicate with one another using the Internet Protocol (IP). Here's how it works:Before our devices connect to the internet, they **connect to a network** that's connected to the internet. This network gives us access to the internet. For example, the network that we use at home will most likely be our Internet Service Provider (ISP). After this, our ISP will assign an IP address to our device. Our internet activity will go through the ISP and then be routed back to us using our assigned IP address.

If we leave our home and take our personal device with us, our IP address from our home network doesn't come with us. Let's say we go to a hotel. When we want to use the internet at the hotel, we'd probably connect to their Wi-Fi network. Since we're

using a new network, we're temporarily assigned a **new IP address**. This temporary IP address is assigned to us by the hotel's ISP. **Can our IP addresses change?** Yes, they can! If we turn our modem or router off, our IP address may change. We can also contact our ISP to change our IP address. This is one of the reasons why DNS is so important. Instead of directly informing others when our IP address changes, we can directly inform our DNS server. This means that any other device that contacts the DNS server will get the updated information for the new IP address.

- **Static IP addresses:**

A static IP address is explicitly allocated to a device rather than one that a DHCP server has assigned. Because it does not change, it is called static.

Static IP addresses can be configured on routers, phones, tablets, desktops, laptops, and any other device that can use an IP address. This can be done either by the device itself handing out IP addresses or by manually typing the IP address into the device.

If you want to host a website from your home, have a file server on your network, utilize networked printers, forward ports to a specific device, run a print server, or use a remote access application, you'll need a static IP address. DNS servers are an example of a static IP address at work.

- **Static IP addresses**

Pros

- **Remote access**: Static IP addresses make it easy for us to work remotely using a Virtual Private Network (VPN).
- **Server hosting**: Static IP addresses make it easy for people to find us using DNS.
- **DNS support**: With static IP, it's easier to manage DNS servers.
- **Geolocation services**: With static IP addresses, our geolocation services are more accurate. This is because our services will match the IP address to its physical location.
- **Reliable connection**: A static IP address is fixed, which typically results in a more reliable connection.

- **Easy to find**: A static IP address can make it easier to find specific devices on a network.

Cons

- **Security concerns**: With a static IP address, anyone with the proper tools can find where our devices are located. VPNs can help with this.
- **Cost**: Static IP addresses are not as cost-effective as dynamic IP addresses. Typically, ISPs charge more for them.

- **Dynamic IP addresses**

A dynamic IP address is an IP address that can **regularly change**. An ISP will buy a large number of dynamic IP addresses and assign them to their customer's devices. Dynamic IP addresses are often reassigned. Reassigning IP addresses helps internet providers save money and ensure a higher level of **security**. It also means that they don't need to take the time to reestablish any network connections if we go on a vacation or move to a new location.Dynamic IP addresses are more common for **consumer equipment** and personal use. A dynamic IP address is assigned to a device by our ISP's Dynamic Host Configuration Protocol (DHCP) servers. The DHCP server typically uses network routers to assign addresses to devices.

- **Dynamic IP addresses**

Pros

- **Easy configuration**: DHCP servers automatically assign IP addresses to our devices, so we don't need to worry about setting it up ourselves.
- **Cost**: Dynamic IP addresses are usually cheaper than static IP addresses.
- **Unlimited IP addresses**: Dynamic IP addresses can be reused. Whenever our devices need a new dynamic IP address, our network or router can automatically configure them for us.
- **Security**: Dynamic IP addresses make it more difficult for potential attackers to locate our networked devices. This is because dynamic IP addresses can change frequently, so it's harder to track a device. This helps with physical and online security. We can also increase our security measures by using a VPN.

Cons

- **DNS compatibility**: If we wanted to host an email server, for example, it may be difficult to use a dynamic IP address because DNS doesn't work well with dynamic IP addresses. We could use a dynamic DNS service, but those tend to be expensive.
- **Remote connectivity**: If we don't have the proper remote access software, it'll be difficult to connect using a dynamic IP address. A VPN can help with this.
- **Increased downtime**: Sometimes, our ISP can't assign us a dynamic IP address. This can slow down our internet connection.
- **Inaccurate geolocation**: Dynamic IP addresses may affect our geolocation services because our IP address may not reflect our physical location.

- **At the transport layer**,

TCP/IP defines three protocols: Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and Stream Control Transmission Protocol (SCTP). At the network layer, the main protocol defined by TCP/IP is the Internetworking Protocol (IP); there are also some other protocols that support data movement in this layer.

- **Physical and Data Link Layers**

At the physical and data link layers, TCP/IP does not define any specific protocol. It supports all the standard and proprietary protocols. A network in a TCP/IP internetwork can be a local-area network or a wide-area network.

- **Network Layer**

At the network layer (or, more accurately, the internetwork layer), TCP/IP supports the Internetworking Protocol. IP, in turn, uses four supporting protocols: ARP, RARP, ICMP, and IGMP. Each of these protocols is described in greater detail in later units.

- **Internetworking Protocol (IP)**

The Internetworking Protocol (IP) is the transmission mechanism used by the TCP/IP protocols. It is an unreliable and connectionless protocol--a best-effort delivery service. The term best effort means that IP provides no error checking or tracking. IP

assumes the unreliability of the underlying layers and does its best to get a transmission through to its destination, but with no guarantees. IP transports data in packets called datagrams, each of which is transported separately. Datagrams can travel along different routes and can arrive out of sequence or be duplicated. IP does not keep track of the routes and has no facility for reordering data- grams once they arrive at their destination. The limited functionality of IP should not be considered a weakness, however. IP provides bare-bones transmission functions that free the user to add only those facilities necessary for a given application and thereby allows for maximum efficiency.

- **Address Resolution Protocol (ARP)**

The Address Resolution Protocol (ARP) is used to associate a logical address with a physical address. On a typical physical network, such as a LAN, each device on a link is identified by a physical or station address, usually imprinted on the network interface card (NIC). ARP is used to find the physical address of the node when its Internet address is known. ARP .

- **Reverse Address Resolution Protocol (RARP)**

The Reverse Address Resolution Protocol (RARP) allows a host to discover its Inter- net address when it knows only its physical address. It is used when a computer is con nected to a network for the first time or when a diskless computer is booted.

The Internet Control Message Protocol (ICMP) is a mechanism used by hosts and gateways to send notification of datagram problems back to the sender. ICMP sends query and error reporting messages.

- **Internet Group Message Protocol**

The Internet Group Message Protocol (IGMP) is used to facilitate the simultaneous transmission of a message to a group of recipients.

- **Transport Layer**

Traditionally the transport layer was represented in TCP/IP by two protocols: TCP and UDP. IP is a host-to-host protocol, meaning that it can deliver a packet from one

physical device to another. UDP and TCP are transport level protocols responsible for delivery of a message from a process (running program) to another process. A new transport layer protocol, SCTP, has been devised to meet the needs of some newer applications.

- **User Datagram Protocol**

The User Datagram Protocol (UDP) is the simpler of the two standard TCP/IP transport protocols. It is a process-to-process protocol that adds only port addresses, checksum error control, and length information to the data from the upper layer. UDP

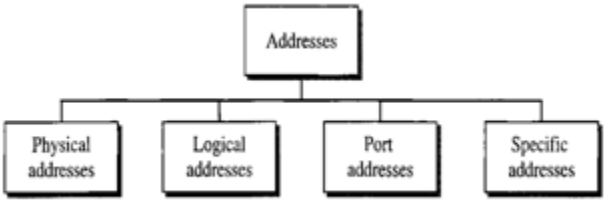
- **Transmission Control Protocol**

The Transmission Control Protocol (TCP) provides full transport-layer services to applications. TCP is a reliable stream transport protocol. The term stream, in this context, means connection-oriented: A connection must be established between both ends of a transmission before either can transmit data. At the sending end of each transmission, TCP divides a stream of data into smaller units called segments. Each segment includes a sequence number for reordering after receipt, together with an acknowledgment number for the segments received. Segments are carried across the internet inside of IP datagrams. At the receiving end, TCP collects each datagram as it comes in and reorders the transmission based on sequence numbers. **Stream Control Transmission Protocol** The Stream Control Transmission Protocol (SCTP) provides support for newer applications such as voice over the Internet. It is a transport layer protocol that combines the best features of UDP and TCP.

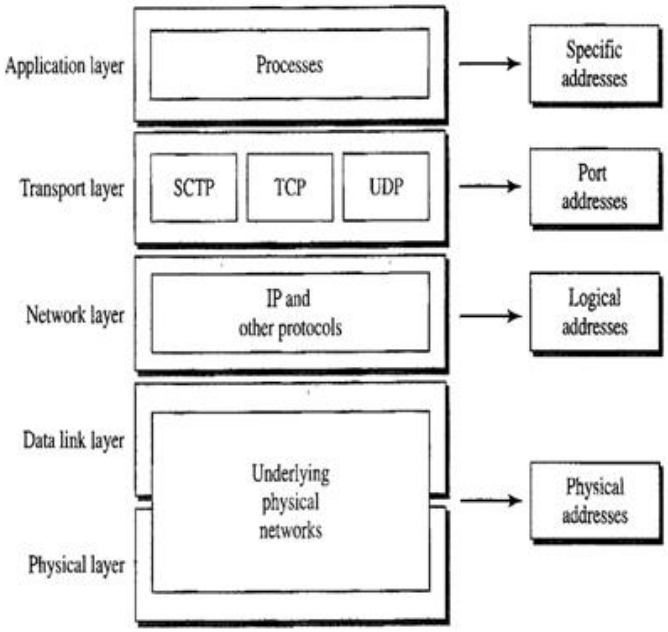
3.4 Addressing-Physical, Logical, Port and Specific addresses.

- **ADDRESSING**

Four levels of addresses are used in an internet employing the TCP/IP protocols: physical (link) addresses, logical (IP) addresses, port addresses, and specific addresses .



Each address is related to a specific layer in the TCP/IP architecture.

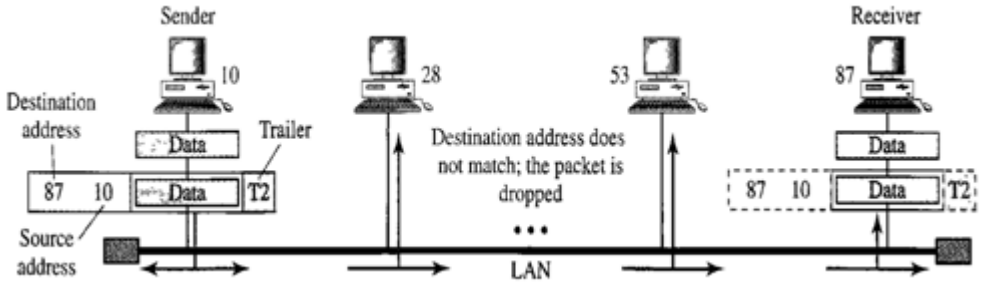


- **Physical Addresses**

The physical address, also known as the link address, is the address of a node as defined by its LAN or WAN. It is included in the frame used by the data link layer. It is the lowest-level address. The size and format of these addresses vary depending on the network. For example, Ethernet uses a 6-byte (48-bit) physical address that is imprinted on the network interface card (NIC). In Figure ,a node with physical address 10 sends a frame to a node with physical address 87. The two nodes are connected by a link (bus topology LAN). At the data link layer, this frame contains physical (link) addresses in the header. These are the only addresses needed. The rest of the header contains other information needed at this level. The trailer usually contains extra bits needed for error detection. As the figure shows, the computer with physical address 10 is the sender, and the computer with physical address 87 is the receiver. The data link layer at the sender receives data from an upper layer. It encapsulates the data in a frame, adding a header and a trailer. The header, among

other pieces of information, carries the receiver and the sender physical (link) addresses. Note that in most data link protocols, the destination address, 87 in this case, comes before the source address (10 in this case).

The intended destination computer, however, finds a match between the destination address in the frame and its own physical address. The frame is checked, the header and trailer are dropped, and the data part is decapsulated and delivered to the upper layer.



Most local-area networks use a 48-bit (6-byte) physical address written as 12 hexadecimal digits; every byte (2 hexadecimal digits) is separated by a colon, say 07:0t:02:01:2C:4B,A 6-byte (12 hexadecimal digits) physical address

- **Logical Addresses**

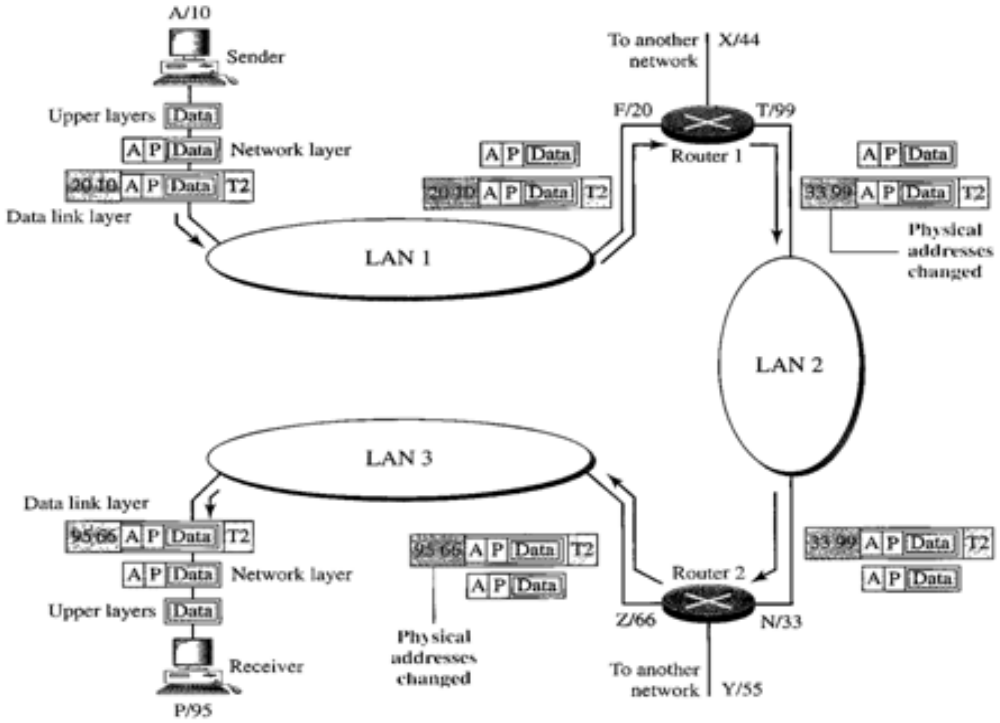
Logical addresses are necessary for universal communications that are independent of underlying physical networks. Physical addresses are not adequate in an internetwork environment where different networks can have different address formats. A universal addressing system is needed in which each host can be identified uniquely, regardless of the underlying physical network.

The logical addresses are designed for this purpose. A logical address in the Internet is currently a 32-bit address that can uniquely define a host connected to the Internet. No two publicly addressed and visible hosts on the Internet can have the same IP address.

- **Example**

Figure shows a part of an internet with two routers connecting three LANs. Each device (computer or router) has a pair of addresses (logical and physical) for each connection. In this case, each computer is connected to only one link and therefore has only one pair of addresses. Each router, however, is connected to three networks (only two are shown in the figure). So each router has three pairs of addresses, one for

each connection. Although it may obvious that each router must have a separate physical address for each connection, it may not be obvious why it needs a logical address for each connection.



The computer with logical address A and physical address 10 needs to send a packet to the computer with logical address P and physical address 95. We use letters to show the logical addresses and numbers for physical addresses. The sender encapsulates its data in a packet at the network layer and adds two logical addresses (A and P). Note that in most protocols, the logical source address comes before the logical destination address (contrary to the order of physical addresses). The network layer, however, needs to find the physical address of the next hop before the packet can be delivered. The network layer consults its routing table and finds the logical address of the next hop (router 1) to be F. The ARP discussed previously finds the physical address of router 1 that corresponds to the logical address of 20. Now the network layer passes this address to the data link layer, which in turn, encapsulates the packet with physical destination address 20 and physical source address 10.The frame is received by every device on LAN 1, but is discarded by all except router 1, which finds that the destination physical address in the frame matches with its own physical address. The router decapsulates the packet from the frame to read the logical destination address P. Since the logical destination address does not match the router's

logical address, the router knows that the packet needs to be forwarded. The router consults its routing table and ARP to find the physical destination address of the next hop (router 2), creates a new frame, encapsulates the packet, and sends it to router 2.

Note the physical addresses in the frame. The source physical address changes from 10 to 99. The destination physical address changes from 20 (router 1 physical address) to 33 (router 2 physical address). The logical source and destination addresses must remain the same; otherwise the packet will be lost. At router 2 we have a similar scenario. The physical addresses are changed, and a new frame is sent to the destination computer. When the frame reaches the destination, the packet is decapsulated. The destination logical address P matches the logical address of the computer. The data are decapsulated from the packet and delivered to the upper layer. Note that although physical addresses will change from hop to hop, logical addresses remain the same from the source to destination.

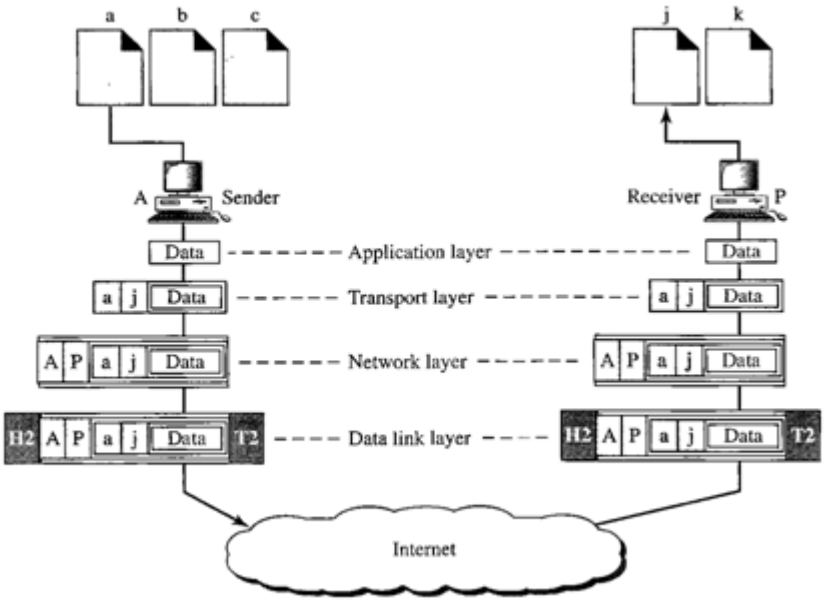
- **Port Addresses**

The IP address and the physical address are necessary for a quantity of data to travel from a source to the destination host. However, arrival at the destination host is not the final objective of data communications on the Internet. A system that sends nothing but data from one computer to another is not complete. Today, computers are devices that can run multiple processes at the same time. The end objective of Internet communication is a process communicating with another process. For example, computer A can communicate with computer C by using TELNET. At the same time, computer A communicates with computer B by using the File Transfer Protocol (FTP). For these processes to receive data simultaneously, we need a method to label the different processes. In other words, they need addresses. In the TCP/IP architecture, the label assigned to a process is called a port address. A port address in TCP/IP is 16 bits in length.

- **Example**

Suppose two computers are communicating via the Intermit. The sending computer is running three processes at this time with port addresses a, b, and c. The receiving computer is running two processes at this time with port addresses j and k. Process a in the sending computer needs to communicate with process j in the receiving computer. Note that although both computers are using the same application, FTP, for example, the port addresses are different because one is a client program and the other

is a server program. To show that data from process a need to be delivered to process j, and not k, the transport layer encapsulates data from the application layer in a packet and adds two port addresses (a and j), source and destination. The packet from the transport layer is then encapsulated in another packet at the network layer with logical source and destination addresses (A and P). Finally, this packet is encapsulated in a frame with the physical source and destination addresses of the next hop. We have not shown the physical addresses because they change from hop to hop inside the cloud designated as the Internet. A port address is a 16-bit address represented by one decimal number.



- **Specific Addresses**

Some applications have user-friendly addresses that are designed for that specific address. Examples include the e-mail address (for example, sumti@fhda.edu) and the Universal Resource Locator (URL) (for example, www. mhhe.com). The first defines the recipient of an e-mail, the second is used to find a document on the World Wide Web. These addresses, however, get changed to the corresponding port and logical addresses by the sending computer.

- **IPv4 ADDRESSES**

An IPv4 address is a 32-bit address that uniquely and universally defines the connection of a device (for example, a computer or a router) to the Internet. **An IPv4 address is 32 bits long.**IPv4 addresses are unique. They are unique in the sense that

each address defines one, and only one, connection to the Internet. Two devices on the Internet can never have the same address at the same time. We will see later that, by using some strategies, an address may be assigned to a device for a time period and then taken away and assigned to another device. On the other hand, if a device operating at the network layer has m connections to the Internet, it needs to have m addresses. We will see later that a router is such a device. The IPv4 addresses are universal in the sense that the addressing system must be accepted by any host that wants to be connected to the Internet.

- **The IPv4 addresses are unique and universal.**

Address Space

A protocol such as IPv4 that defines addresses has an address space. An address space is the total number of addresses used by the protocol. If a protocol uses N bits to define an address, the address space is 2^N because each bit can have two different values (0 or 1) and N bits can have 2^N values. IPv4 uses 32-bit addresses, which means that the address space is 2^{32} or 4,294,967,296 (more than 4 billion). This means that, theoretically, if there were no restrictions, more than 4 billion devices could be connected to the Internet. We will see shortly that the actual number is much less because of the restrictions imposed on the addresses.

- **The address space of IPv4 is 2^{32} or 4,294,967,296.**

Notations There are two prevalent notations to show an IPv4 address: binary notation and dotted decimal notation.

- **Binary Notation**

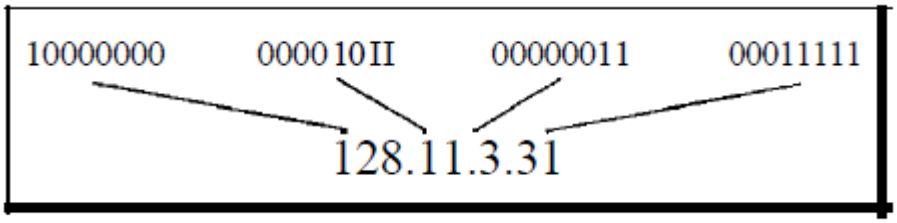
In binary notation, the IPv4 address is displayed as 32 bits. Each octet is often referred to as a byte. So it is common to hear an IPv4 address referred to as a 32-bit address or a 4-byte address. The following is an example of an IPv4 address in binary notation: 01110101 10010101 00011101 00000010

- **Dotted-Decimal Notation**

To make the IPv4 address more compact and easier to read, Internet addresses are

usually written in decimal form with a decimal point (dot) separating the bytes. The following is the dotted decimal notation of the above address:
117.149.29.2 Figure 19.1 shows an IPv4 address in both binary and dotted-decimal notation. Note that because each byte (octet) is 8 bits, each number in dotted-decimal notation is a value ranging from 0 to 255.

Figure 3.4 Dotted-decimal notation and binary notation for an IPv4 address



- # IPv6 ADDRESSES:

Structure

An IPv6 address consists of 16 bytes (octets); it is 128 bits long.

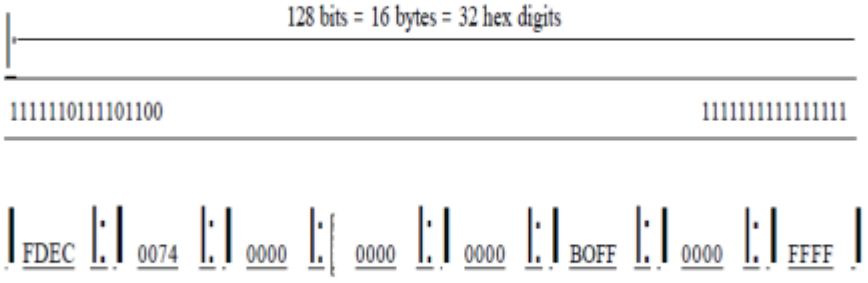
An IPv6 address is 128 bits long.

Hexadecimal Colon Notation

To make addresses more readable, IPv6 specifies hexadecimal colon notation. In this notation, 128 bits is divided into eight sections, each 2 bytes in length. Two bytes in hexadecimal notation requires four hexadecimal digits. Therefore, the address consists of 32 hexadecimal digits, with every four digits separated by a colon, as shown in Figure 3.4.1

Figure 3.4.1 ***IPv6 address in binary and hexadecimal colon notation***

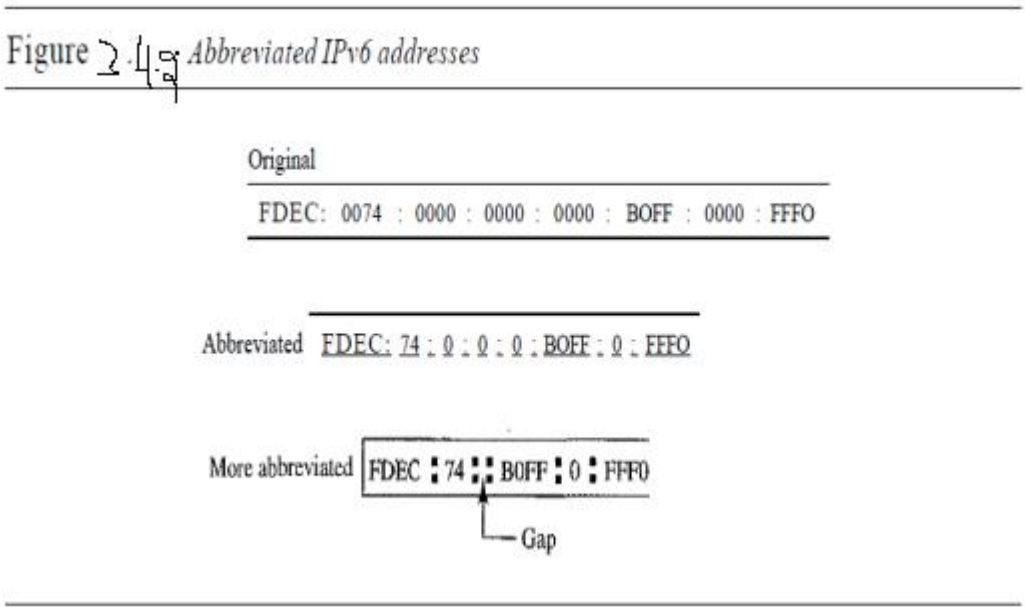
Figure 3.4.1 *IPv6 address in binary and hexadecimal colon notation*



Abbreviation

Although the IP address, even in hexadecimal format, is very long, many of the digits are zeros. In this case, we can abbreviate the address. The leading zeros of a section (four digits between two colons) can be omitted. Only the leading zeros can be dropped, not the trailing zeros (see Figure 19.15).

Figure 3.4.2 Abbreviated IPv6 addresses



Using this form of abbreviation, 0074 can be written as 74, OOOF as F, and 0000 as O. Note that 3210 cannot be abbreviated. Further abbreviations are possible if there

are consecutive sections consisting of zeros only. We can remove the zeros altogether and replace them with a double semicolon. Note that this type of abbreviation is allowed only once per address. If there are two runs of zero sections, only one of them can be abbreviated. Reexpansion of the abbreviated address is very simple: Align the unabbreviated portions and insert zeros to get the original expanded address.

- **Address Space**

IPv6 has a much larger address space; 2128 addresses are available. The designers of IPv6 divided the address into several categories. A few leftmost bits, called the type prefix, in each address define its category. The type prefix is variable in length, but it is designed such that no code is identical to the first part of any other code. In this way, there is no ambiguity; when an address is given, the type prefix can easily be determined.

8.3 SUMMARY

In this unit, we learnt about the TCP/IP Protocol Suite, Physical, Data Link, network, Transport and an Application Layers. Addressing-Physical, Logical, Port and Specific addresses.

8.4 KEYWORDS

TCP/IP, UDP, ARP, URL

8.5 QUESTIONS FOR SELF STUDY

1. Explain TCP/IP Protocol Suite
2. Discuss briefly Physical, Data Link, Network, Transport and an Application Layers.
3. Elucidate Addressing-Physical, Logical, Port and Specific addresses.

8.6 REFERENCE

1. Data Communications and Networking – Behrouz A. Forouzan, 4th Edition, Tata McGraw-Hill, 2006.
2. Communication Networks: Fundamental Concepts and Key Architectures - Alberto Leon, Garcia and Indra Widjaja, 3rd Edition, Tata McGraw- Hill, 2004.

3. Data and Computer Communication, William Stallings, 8th Edition, Pearson Education, 2007.

UNIT – 9: PROTOCOLS

Structure

9.0 Objectives
9.1 Introduction
9.2 TCP/IP suite
9.3 SMTP
9.4 POP3
9.5 SNMP
9.6 HTTP
9.7 DNS
9.8 FTP
9.9 Summary
9.10 Keywords
9.11 Questions
9.12 Reference

9.0 OBJECTIVES

After going through this lesson you will be able to

- Describe TCP/IP protocol suite
- Elucidate SMTP protocol
- Understand POP3 protocol
- Analyse SNMP protocol
- Discuss HTTP protocol
- Describe DNS protocol
- Discuss FTP protocol

9.1 INTRODUCTION

A network protocol is an established set of rules that determine how data is transmitted between different devices in the same network. Essentially, it allows connected devices to communicate with each other, regardless of any differences in their internal processes, structure or design. Network protocols are the reason you can easily communicate with people all over the world, and thus play a critical role in modern digital communications.

9.2 TCP/IP PROTOCOL SUITE

The protocol stack used on the Internet is the Internet Protocol Suite. It is usually called TCP/IP after two of its most prominent protocols, but there are other proto-cols as well. The *TCP/IP model* is based on a five-layer model for networking. From bottom (the link) to top (the user application), these are the physical, data link, net- work, transport, and application layers. Not all layers are completely defined by the model, so these layers are "filled in" by external standards and protocols. The layers have names but no numbers, and although sometimes people speak of "Layer 2" or "Layer 3," these are not TCP/IP terms. Terms like these are actually from the OSI Reference Model.

The TCP/IP stack is *open*, which means that there are no "secrets" as to how it works. (There are "open systems" too, but with TCP/IP, the systems do not have to be "open" and often are not.) Two compatible end-system applications can communicate regardless of their underlying architectures, although the connections between layers are not defined.

The TCP/IP or Internet model is not the only standard way to build a protocol suite or stack. The Open Standard Interconnection (OSI) reference model is a seven- layer model that loosely maps into the five layers of TCP/IP. Until the Web became widely popular in the 1990s, the OSI reference model, with distinctive names and numbers for its layers, was proposed as the standard model for all communication networks. Today, the OSI reference model (OSI-RM) is often used as a learning tool to introduce the functions of TCP/IP.

The TCP/IP stack is comprised of modules. Each module provides a specific function, but the modules are fairly independent. The TCP/IP layers contain relatively independent protocols that can be used depending on the needs of the system to provide whatever function is desired. In TCP/IP, each higher layer protocol is sup- ported by lower layer protocols. The whole collection of protocols forms a type of hourglass shape, with IP in the middle, and more and more protocols up or down from there.

Suite, Stack, and Model

The term "protocol stack" is often used synonymously with "protocol suite" as an implementation of a reference model. However, the term "protocol suite" properly refers to a collection of all the protocols that can make up a layer in the reference model. The Internet protocol suite is an example of the Internet or TCP/IP refer- ence model protocols, and a TCP/IP protocol stack implements one or more of these protocols at each layer.

The TCP/IP Layers

The TCP/IP protocol stack models a series of protocol layers for networks and systems that allows communications between any types of devices. The model consists of five separate but related layers, as shown in Figure 1.9. The Internet protocol suite is based on these five layers. TCP/IP says most about the network and transport layers, and a lot about the application layer. TCP/IP also defines how to interface the network layer with the data link and physical layers, but is not directly concerned with these two layers themselves.

The Internet protocol suite assumes that a layer is there and available, so TCP/IP does not define the layers themselves. The stack consist of protocols, not implementa- tions, so describing a layer or protocols says almost nothing about how these things should actually be built.

Not all systems on a network need to implement all five layers of TCP/IP. Devices using the TCP/IP protocol stack fall into two general categories: a *host* or *end system* (ES) and an *intermediate node* (often a router) or an *intermediate system* (IS). The

User Application Programs

Application Layer
Transport Layer
Network Layer

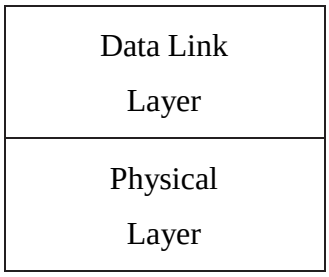


FIGURE : The five layers of TCP/IP. Older models often show only four layers, combining the physical anddata link layers.

intermediate nodes usually only involve the first three layers of TCP/IP (although many of them still have all five layers for other reasons, as we have seen).

In TCP/IP, as with most layered protocols, the most fundamental elements of the process of sending and receiving data are collected into the groups that become the layers. Each layer's major functions are distinct from all the others, but layers canbe combined for performance reasons. Each implemented layer has an *interface* with the layers above and below it (except for the application and physical layers, of course)and provides its defined service to the layer above and obtains services from the layer below. In other words, there is a *service interface* between each layer, but these are notstandardized and vary widely by operating system.

TCP/IP is designed to be comprehensive and flexible. It can be extended to meet new requirements, and has been. Individual layers can be combined for implementation purposes, as long as the service interfaces to the layers remain intact. Layers can even be split when necessary, and new service interfaces defined. Services are provided to the layer above after the higher layer provides the lower layer with the command, data, and necessary parameters for the lower layer to carry out the task.

Layers on the same system provide and obtain services to and from adjacent layers. However, a *peer-to-peer protocol process* allows the same layers on different systems to communicate. The term *peer* means every implementation of some layer is essentially equal to all others. There is no "master" system at the protocol level. Communications between peer layers on different systems use the defined protocols appropriate to the given layer.

In other words, *services* refer to communications between layers within the same process, and *protocols* refer to communications between processes. This can be con- fusing, so more information about these points is a good idea.

Protocols and Interfaces

It is important to note that when the layers of TCP/IP are on different systems, they are *only* connected at the physical layer. Direct peer-to-peer communication between all other layers is impossible. This means that all data from an application have to flow "down" through all five layers at the sender, and "up" all five layers at the receiver to reach the correct process on the other system. These data are sometimes called a *service data unit* (SDU).

Each layer on the sending system adds information to the data it receives from the layer above and passes it all to the layer below (except for the physical layer, which has no lower layers to rely on in the model and actually has to send the bits in a form appropriate for the communications link used).

Likewise, each layer on the receiving system unwraps the received message, often called a *protocol data unit* (PDU), with each layer examining, using, and stripping off the information it needs to complete its task, and passing the remainder up to the next layer (except for the application layer, which passes what's left off to the application program itself). For example, the data link layer removes the wrapper meant for it, usesit to decide what it should do with this data unit, and then passes the remainder up to the network layer.

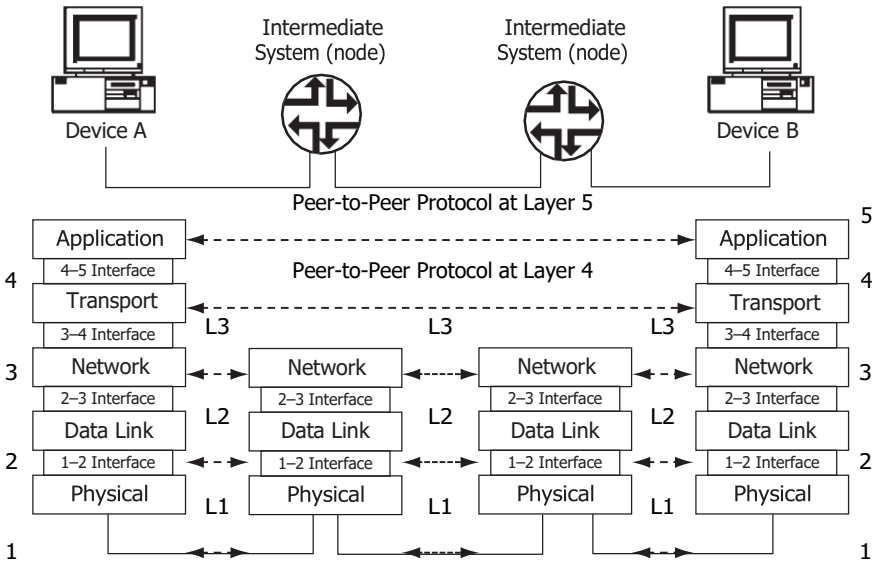


FIGURE 1.10 Physical Communication Links

Protocols and interfaces, showing how devices are only physically connected at the lowest layer (Layer 1). Note that functionally, intermediate nodes only require the bottom three layers of the model.

The whole interface and protocol process is shown in Figure 1.10. Although TCP/IP layers only have names, layer numbers are also used in the figure, but only for illustra- tion. (The numbers come from the ISO-RM.)

As shown in the figure, there is a natural grouping of the five-layer protocol stack at the network layer and the transport layer. The lower three layers of TCP/IP, some- times called the network support layers, must be present and functional on all systems, regardless of the end system or intermediate node role. The transport layer links the upper and lower layers together. This layer can be used to make sure that what was sent was received, and what was sent is useful to the receiver (and not, for example, a stray PDU misdirected to the host or unreasonably delayed).

The process of encapsulation makes the whole architecture workable. Encapsu- lation of one layer's information inside another layer is a key part of how TCP/IP works.

THE LAYERS OF TCP/IP

TCP/IP is mature and stable, and is the only protocol stack used on the Internet. This book is all about networking with TCP/IP, but it is easy to get lost in the particulars of TCP/IP if some discussion of the general tasks that TCP/IP is intended to accomplish is not included. This section takes a closer look at the TCP/IP layers, but only as a general guide to how the layers work.

- **Physical Layer:** Contains all the functions needed to carry the bit stream over a physical medium to another system.
- **Data Link Layer:** Organizes the bit stream into a data unit called a "frame" and delivers the frame to an adjacent system.
- **Network Layer:** Delivers data in the form of a packet from source to destina-tion, across as many links as necessary, to non-adjacent systems.
- **Transport Layer:** Concerned with process-to-process delivery of information.

- **Application Layer:** Concerned with differences in internal representation,user interfaces, and anything else that the user requires.

9.3 SMTP (SIMPLE MAIL TRANSFER PROTOCOL)

SMTP is the standardization for transmission of electronic mails on the Internet. It is used by the e-mail server for sending and receiving messages, but the client host-based application only uses it for sending messages to the mail server. For receiving purposes, they use POP3 or IMAP. It is a TCP/IP application layer protocol and the TCP port used by the mail servers is 25 while the mail clients use the port 587 or 465 for communication.

The outlook mail system of Microsoft system, Gmail and Yahoo mail, deploy SMTP for sending and retrieving emails from the exterior world whereas for interior mail exchange between their respective

How SMTP works

SMTP operates on a client/server model as a three-stage operation. Initially, an electronic mail server uses SMTP to deliver an e-mail to the e-mail server from the Outlook or Gmail clients. Second, the e-mail server sends an e-mail to the e-mail server via SMTP as a relay tool. Third, to upload inbound mails via IMAP and position them on the recipient's inbox, the receiving server uses an email client.

Components of SMTP

First, we split down the SMTP client and the SMTP server into two elements such as User Agent (UA) and Message Teansfer Agent (MTA). The User Agent (UA) plans the message and produces the box. The message is then inserted in the envelope. Thismail is transferred over the internet through the mail transfer agent (MTA).

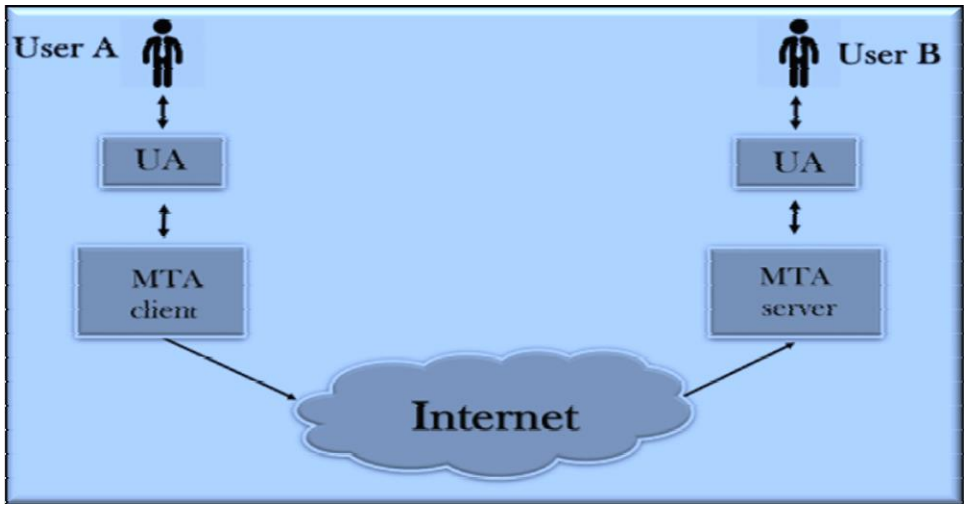


Figure 4A.1:User Agent (UA) and Message Teansfer Agent (MTA) of SMTP

By incorporating a relay system, SMTP permits a more complex system. More MTAs may be included, either as a client or server in order to transmit the e-mail, ratherthan just an MTA at sending and an MTA at the receiving end.

9.4 POP3

It is an easy protocol to access the e-mail boxes remotely. TheRFC 1225 is the protocol. Version 3 (POP3) of the Post Office Protocol is the protocol that helps a recipient to accept an email from the remote mail server.

SMTP requires the host that receives the mail to be stillonline; otherwise it would be impossible to bind to TCP. On behalf of the customers, the POP server collects the mail.

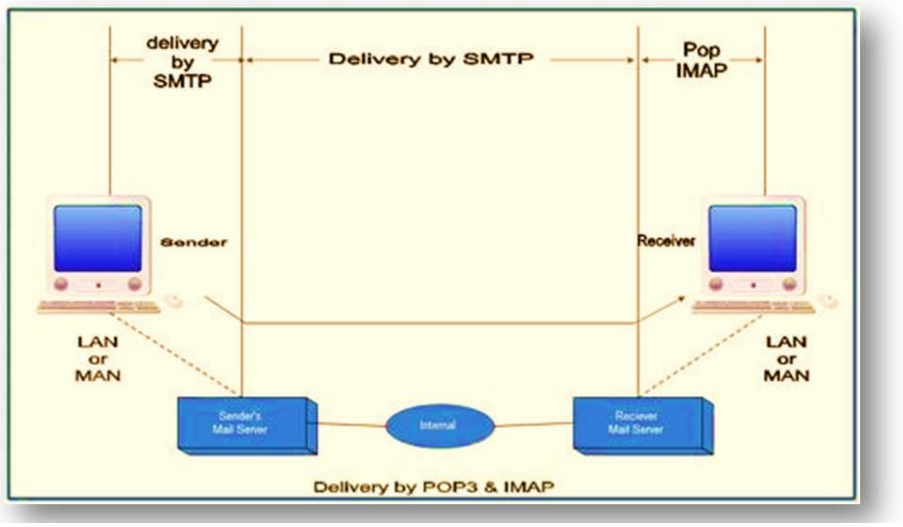


Figure 4A.5: Delivery by POP3

A POP3 server can hold messages for each user before a POP3 client such as Microsoft Outlook 98, Microsoft Outlook Express and Microsoft Mail and News links to the downloads and reads. A POPS client sits in a Transmission Control Protocol session (TPT) using a TCP port 110, logs in to the internet and then issues a set of POP3 commands to receive a POP3 message from a POP3 server.

stat: It asks the server for the number of messages waiting to beretrieved.
list: It determines the size of each message to be retrieved.
retr: It retrieves individual messages d.
Quit: Ends the POP3session.

Though POP3 has experienced a number of changes, the developers maintained that a three-stage process was followed between the customer and the server during mail retrieval. They triedto

make this protocol very easy and today it is very popular with thissimplification.

Let's understand the working of the POP3 protocol.

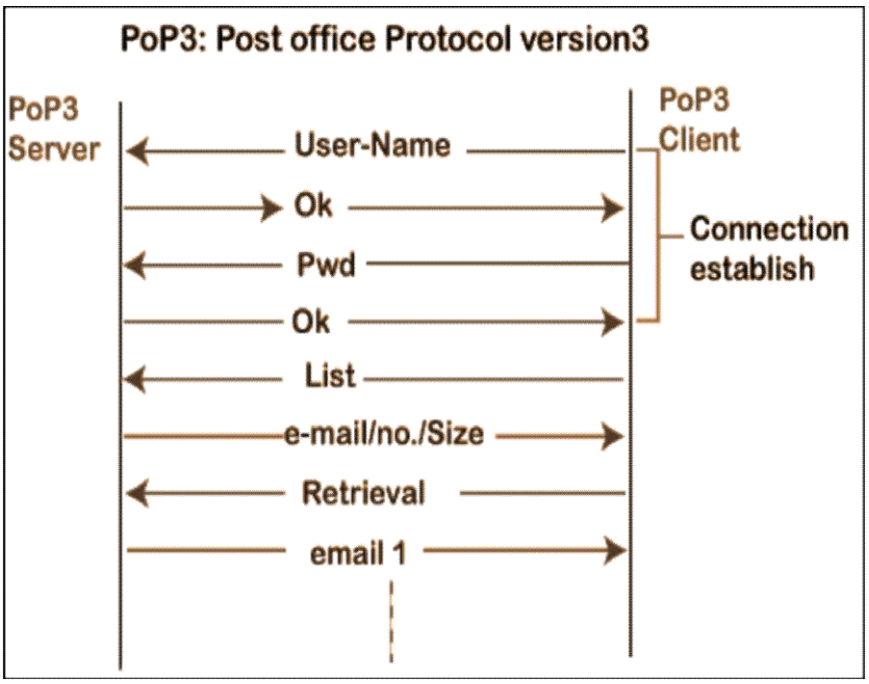


Figure: Working of POP3

The POP3 server asks for the POP3 client user name in order to create a connection between the server POP3 and the POP3 client.If the POP3 server identifies the user name, the ok message is sent. Then it asks the POP3 client's password and then transfers the POP3 client to the POP3 server. The POP3 server sends an OK message if the password fits, and the link is established. The client will see the POP3 mail server mail list after establishing a connection. In themail list, the customer gets the server email numbers and sizes. The user will launch mail recovery from this collection. Once the client receives all of the user addresses, all the server emails will be removed. We may also assume that emails are only contained on a certain computer, so that the same mails cannot be read on another machine. The emailed settings can be modified to leave an email copy on the email server to resolve this condition.

Advantages of POP3 protocol

The following are the advantages of a POP3 protocol:

- Users can read an electronic offline email. Only when uploading e-mails from the cloud

would an internet connection entail. Until the emails are downloaded from the cloud, all of the mails downloaded can be obtained without internet from our PC or hard disk of your phone. There is, therefore, no permanentInternet connectivity in the POP3 protocol.

- Secure and fast access to e-mails, because they are saved on ourPC already.
- The size of the email we receive or send does not restrict.
- Less space for server store is needed when all the mail on thelocal computer is stored.
- The mailbox has the full mailbox capacity, but the hard disk sizeis limited.
- The protocol is simple, so it is one of today's most commonprotocols. Configuring and using is simple.

Disadvantages of POP3 protocol

- **The following are the advantages of a POP3 protocol:**
- When e-mails from a server are downloaded, all mails are defaultremoved from the server. Therefore, mails from other computers cannot be read without a backup of the email on the server being configured.
- It can be difficult to move the mail folder from the local computer to another machine.
- Since all attachments are placed on your local computer, if the
- virus scanner does not search them, then the possibility of a virusattack would be high. The assault of the virus could destroy the machine.
- The downloaded email directory of the mail server will corrupt it as well.
- The mail would be stored at the local machine, allowing anyone on the computer to enter the folder of the email.

9.5 SNMP

SNMP stands for the for simple network management protocol. SNMP is an internet device control system. It offers a series of internet surveillance and management activities. SNMP hasthe boss and agent in two components. The director is a host who manages and tracks a variety of organizations like routers. It is a protocol on the application layer where a few manager stations can accommodate a number of operators. On the level of implementation, the protocol can monitor devices produced by various manufacturers and deployed on various physical networks. This network consists of heterogeneous LANs and WANs linked by routers or gateway.

Managers & Agents

A controller is a host operating SNMP client, while the agent is a router running the SNMP server software. Internet management is accomplished by simple dialog between a boss and an agent.

The agent is used to maintain records in a database while the manager is used for data access. For example, a router can store suitable variables such as certain received and forwarded packets, while the manager can compare those variables to decide whether the router is loaded or not. Agents are also able to assist with the management process. When something goes wrong, an agent sends an alert to the server software on the agent

Management with SNMP has three basic ideas:
A manager tests the agent by asking the data which representsthe agent's behaviour.

A manager also forces a certain role to be exercised by resetting the value in the agent database.

An Agent also contributes by advising the manager of an unusual situation to the management process.

SNMP

SNMP defines five types of messages: Get Request, Get NextRequest, Set Request, Get Response, and Trap.

Get Request: The Get Request message is sent to the agent (server) from a manager (client) to get the variable value.

Get Next Request: The Message Get Next Request is sent from the administrator to the agent to get the variable value. The value of the entries in a table is derived from this sort of post. If the managerdoes not know the entry indices, he cannot locate the values. The Message Get Next Request is used to describe an entity in such contexts.

Get Response: In addition to the Get Request and Get Next Requestquery, the Get Reponse message is submitted by the agent to the boss. This message includes the manager's requested value for an attribute.

Set Request: The message Set Request is sent by the manager to theagent to specify a value in a variable.

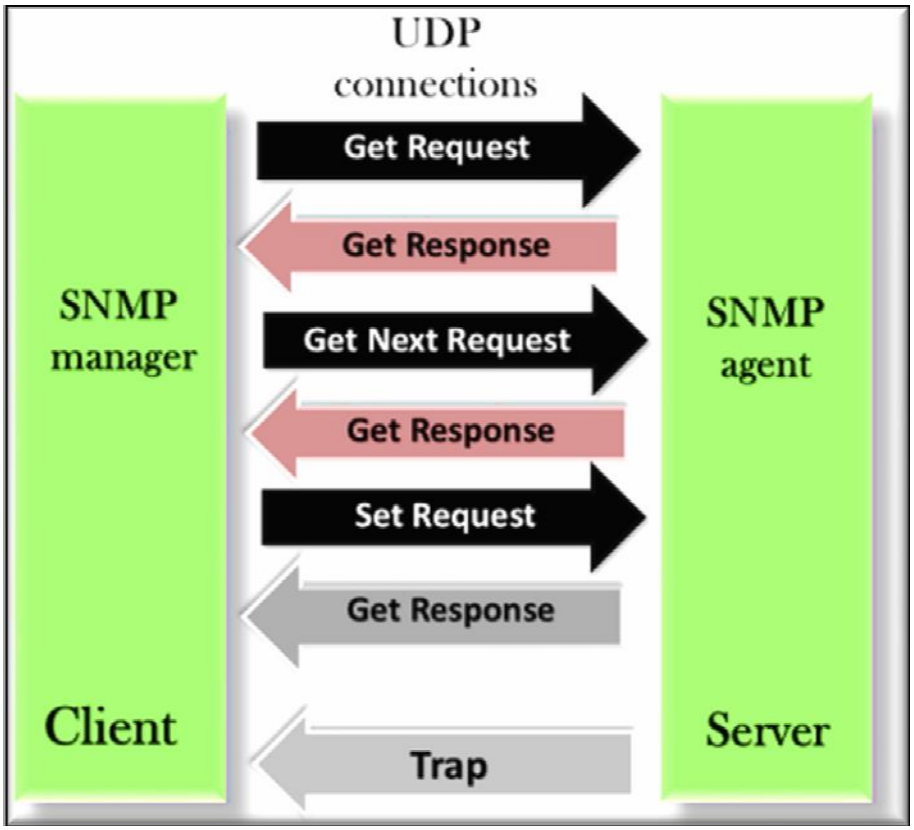


Figure: Types of SNMP Messages

Trap: Trap message is sent by an employee to an event planner. For instance, if the agent is restarted, it notifies the manager and sends the reboot time.

9.6 HTTP

A protocol at the application level for distributed, shared, hyper mediated information systems is the Hypertext Transfer Protocol (HTTP). Since 1990, this has become the basis of theWorld Wide Web data exchange (i.e., internet). HTTP is a generic and stateless protocol, that can also be used with extensions to its request methods, error codes and headers for other purposes. The HTTP is primarily a TCP/IP networking protocol used for delivering World Wide Web data (HTML files, picture files, query results, and so on). TCP 80 is the main port, although other portscan also be used. Computers interact with each other in a structured manner. The HTTP defines how the request data of clients areconstructed, transmitted and responded to the server by the servers.

Features of HTTP are:

There are three basic features that make HTTP a simple but powerful protocol:

HTTP is connectionless: The HTTP client, i.e., a browser starts an HTTP request and the client waits for the answer after a request is made. The server handles the request and sends a reply to the client, then client disconnects the link. Thus, during current request and response, the client and server know each other. New connections are made with new requests.

HTTP is media independent: HTTP is a media-independent protocol that means every dataform can be transmitted via HTTP, as long as both the server andthe recipient know how to manage the data contents. The type of content must be specified in the MIME type header for both the client and the server.

HTTP is stateless: HTTP is a stateless protocol as both the client and server know each other only during the current request. Due to this nature of the protocol, both the client and server do not retain the information between various requests of the web pages.

Basic Architecture

The following diagram shows a very basic architecture of a web application and depicts where HTTP sits:

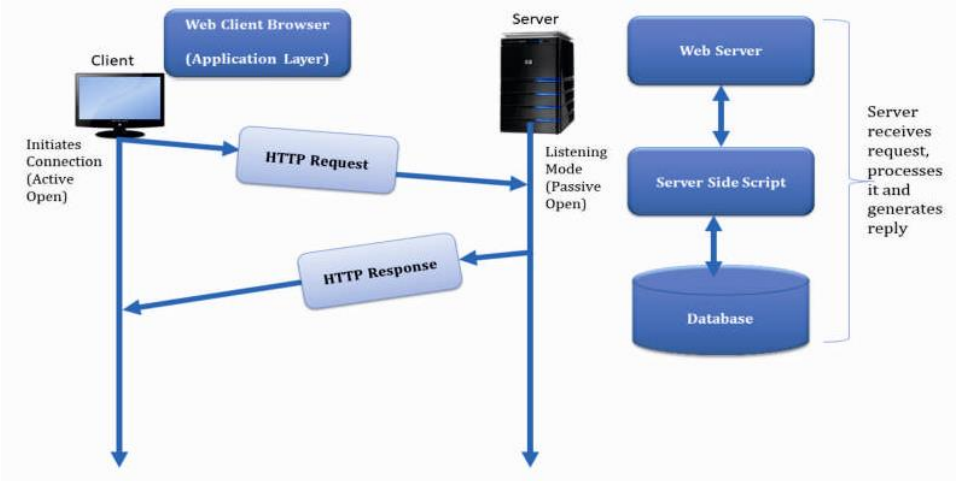


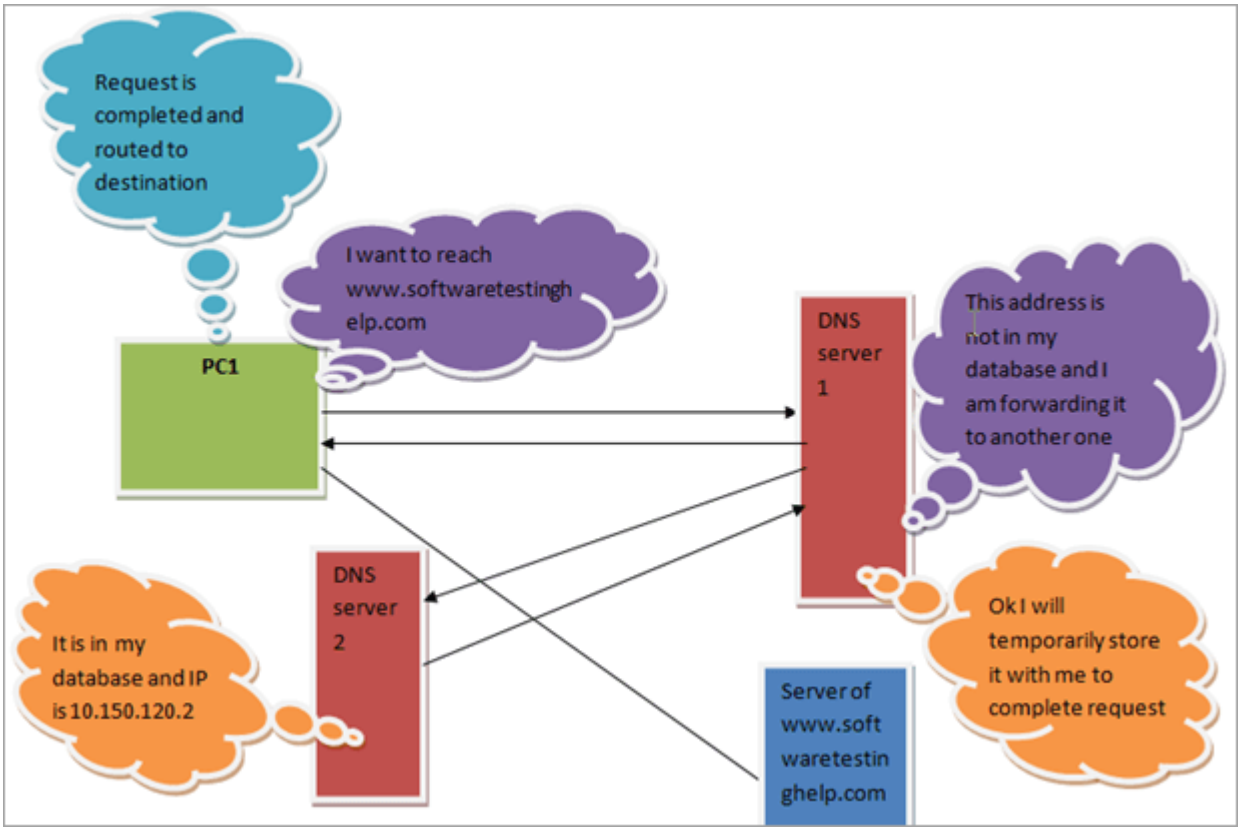
Figure 4.1: HTTP Protocol Architecture

The HTTP protocol is an architectural client/serverrequest/response protocol where web browsers, robots and search engines, etc. function as HTTP clients, and the Web server functions as a server. The HTTP uses the principle of Uniform ResourceLocator (URL)for clients who wish to access a database over the internet and who need address and promote access to documents. The figure above shows the HTTP transaction from client to server. Through sending the request message to the server, the recipient initiates a transaction. A response message is sent from the server to the request.

9.7 DNS (Domain Name Server)

If any user from the personal computer, laptop or tablet uses the Internet and tries to login into some website then the user is using DNS for sure. Thus it is very important to understand the working on a domain name server.

PC's, laptop or tablets don't understand the language of a web address, which means the domain name like Google.com to make them understand for which site we are looking for. Thus DNS came into the role and provides the host with the mapped IP address in respect to the domain name of the website.



As shown in the above figure, when we request for a web page from our PC on the Internet like PC1 is requesting for www.softwaretestinghelp.com, then resolving the domain name query and providing the respective IP address in return is the part of work of the DNS server.

DNS server stores the database of all the relevant IP addresses mapped with their respective domain names.

The DNS query for requesting the IP address in respect to the domain name goes to the DNS server 1 from PC1. The server checks within itself, if it has the IP address regarding the query, and it returns a DNS response with the resolution.

Otherwise, it forwards it to another DNS server 2 requesting for information. This time it gets the resolution from the DNS 2 and it gets mapped with the IP address i.e. 10.150.120.2 corresponding to the Domain name in response and sends it back to PC1.

The PC1 now have the destination IP address and it can communicate further with the known IP address as per the routing.

Now the question arises, as of how the PC will come to know which DNS should be used to get the IP address.

The answer to this is when we connect our system to the ISP, the network devices like a router or switch which assigns the routing information and other configurations as well send which or how many DNS server the PC should connect with to get the address translation.

9.8 FTP (File Transfer Protocol)

It is one of the widely used application layer protocol of the TCP/IP protocol suite. FTP is basically used to exchange data between two host devices over the Internet or Intranet securely.

It is referred to as one of the safest modes of file sharing among systems, and thus it is deployed by large industries, universities, and offices.

It works in the client-server model and thus the user needs an FTP client program to run FTP on its system. The common types of FTP client program include Filezilla and Dreamweaver etc.

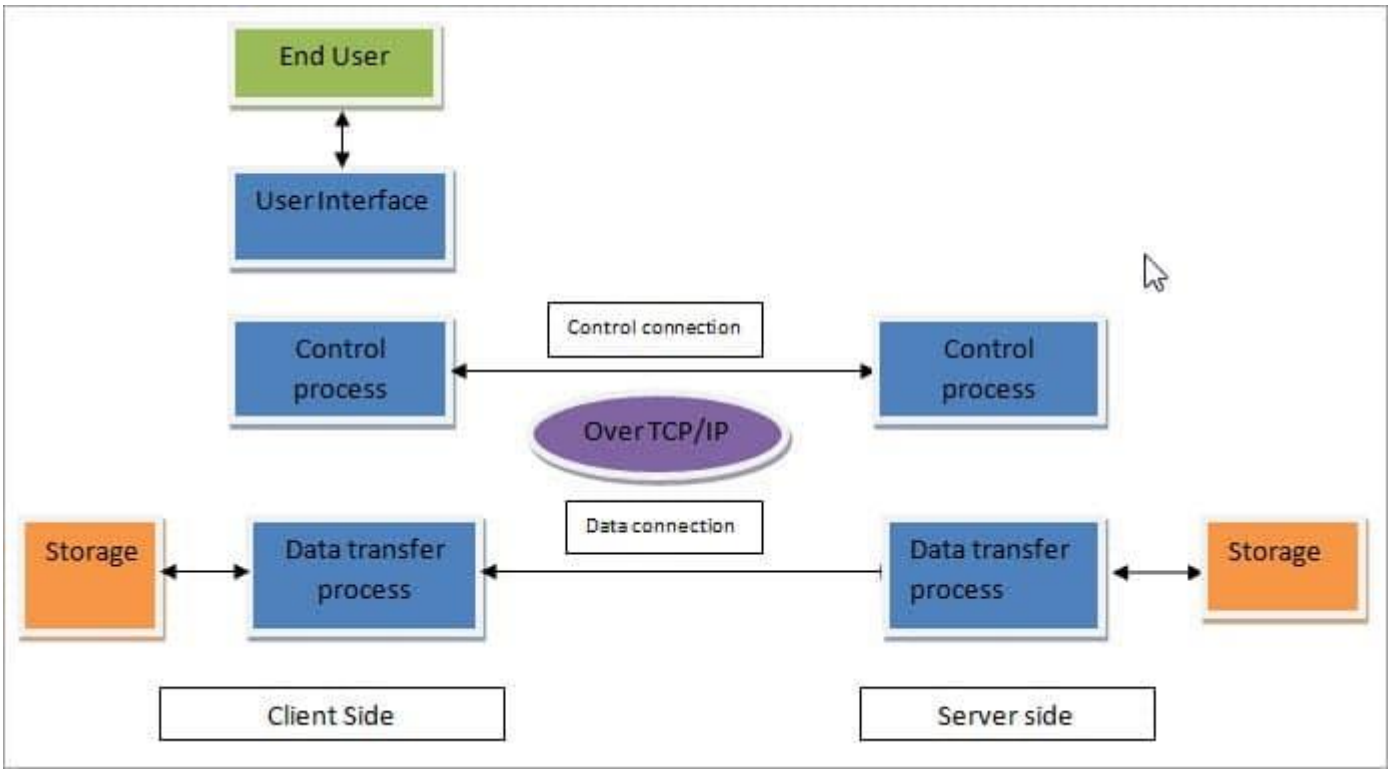
The data transfer takes place only in one direction at a time. The FTP protocol carry out many duties apart from file transfer like creation and deletion of data files, listing, renaming, etc.

The FTP Model

In this model, one host behaves as the client and another host as a server. The one who requests for file-sharing or data is the client host and one which in response completes the request is the server host.

Firstly, the FTP connection is established between the client and server computer and data exchange take place after that. Two channels come into the picture of FTP connection i.e. control channel and data channel.

The control channel establishes the connection between the client and server and remains open for the overall session. The control channel port number is 21 in TCP/IP. While the data channel opens when the client request for a file sharing and get closed after the completion of the request by the server.



Two processes naming data transfer process (DTP) and protocol interpreter (PI) are used in managing the communication between the client and the server. The DTP establishes and manages the connection for the data channel, while PI manages the DTP by applying commands given by the control channel.

The server host end PI is accountable for analyzing the commands received from the client host end via the control channel, connection establishment, and in running the DTP. The client PI is accountable for forwarding the FTP commands, receiving the response from the server and establishment of the connection with the FTP server.

After the establishment of a connection between the FTP client and the FTP server, the client builds up the connection and sends the FTP commands to the server. The server analyzes them and in response completes the request.

Now the server end PI sends the port detail on which the files will be forwarded to the client DTP. The client DTP then waits for the data to arrive at the decided port from the server.

The FTP Response

To make out a secure and reliable file transfer between the client and server, it is important that the server and client should remain in synchronization with each other.

Thus for each command executed by the client, a user is acknowledged by the response and the action is performed by the server host in order. The response consists of a 3 digit code plus a text (a character string is separated from digit by a space) denoting the processing of the commands.

9.9 SUMMARY

In this unit we have learnt the concept of TCP/IP suite. We also discussed SMTP and POP3 protocols. At the end of this unit reader are able to figure out the concepts of SNMP, HTTP, DNS and FTP.

9.10 KEYWORDS

TCP/IP, SMTP, POP3, SNMP, HTTP, DNS and FTP

9.11 QUESTIONS

1. Explain TCP/IP protocol suite.
2. Discuss SMTP protocol.
3. With neat diagram explain POP3.
4. Describe HTTP protocol.
5. Discuss DNS protocol.
6. Briefly explain FTP protocol.

9.12 REFERENCES

- Introduction to Data Communications and Networking by Behrouz Forouzan.
- Computer Networks by Andrew S Tanenbaum.
- Networking Essentials – Third Edition – Jeffrey S. Beasley, Piyasat Nilkaew

UNIT – 10: PROTOCOLS

Structure

10.0 Objectives

10.1 Introduction

10.2 ARP
10.3 RARP
10.4 ICMP
10.5 IGMP
10.6 OSPF
10.7 BGP
10.8 Summary
10.9 Keywords
10.10 Questions
10.11 Reference

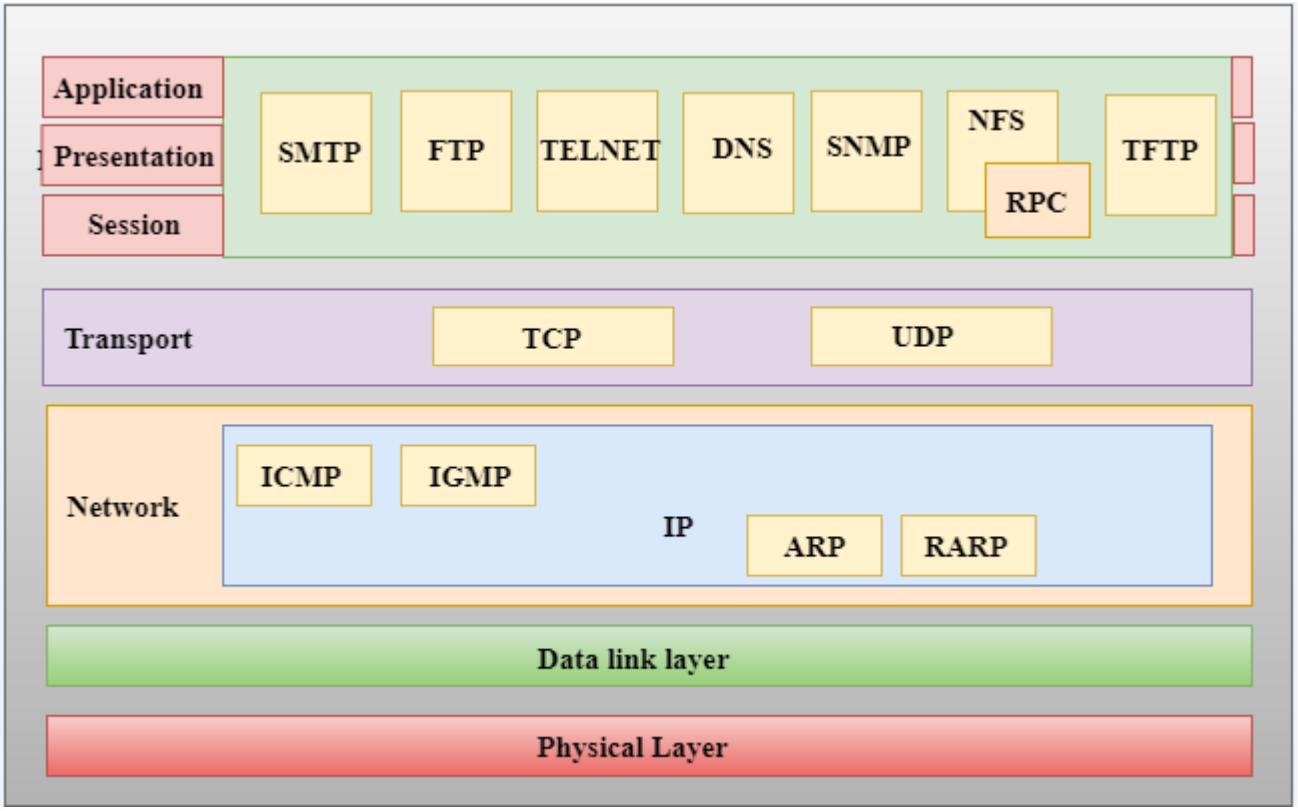
10.0 OBJECTIVES

At the end of this unit you will be able to

- Understand the concept of ARP protocol
- Discuss RARP protocol
- Elucidate ICMP and IGMP
- Describe OSPF protocol
- Explain BGP protocol

10.1 INTRODUCTION

The OSI model, and any other network communication model, provide only a conceptual framework for communication between computers, but the model itself does not provide specific methods of communication. Actual communication is defined by various communication protocols. In the context of data communication, a protocol is a formal set of rules, conventions and data structure that governs how computers and other network devices exchange information over a network. In other words, a protocol is a standard procedure and format that two data communication devices must understand, accept and use to be able to talk to each other. Protocols of different layer as shown in the below diagram.



10.2 ADDRESS RESOLUTION PROTOCOL (ARP)

The Address Resolution Protocol (ARP) is used to associate a logical addresswith a physical address. On a typical physical network, such as a LAN, eachdevice on a link is identified by a physical or station address, usually imprinted on the network interface card (NIC).

ARP is used to find the physical address of a node when its Internet address(IP address) is known. If a machine talks to another machine in the same network, it requires its physical or MAC address. But, since the application has given the destination's IP address it requires some mechanism to bind the IP address with its MAC address. This is done through Address Resolution protocol (ARP).IP address of the destination node is broadcast and the destination node informs the source of its MAC address.

- Assume broadcast nature of LAN
- Broadcast IP address of the destination
- Destination replies it with its MAC address.
- Source maintains a cache of IP and MAC address bindings

But this means that every time machine A wants to send packets to machineB, A has to send an ARP packet to resolve the MAC address of B and hencethis will increase the traffic load too much, so to reduce the communicationcost computers that use ARP maintains a cache of recently acquired IP_to_MAC address bindings, i.e. they do not have to use ARP repeatedly. Several

refinements of ARP are possible: When machine A wants to send packets to macine B, it is possible that machine B is going to send packets tomachine A in the near future. So to avoid ARP for machine B, A should putits IP_to_MAC address binding in the special packet while requesting for theMAC address of B. Since A broadcasts its initial request for the MAC address of B, every machine on the network should extract and store in its cache the IP_to_MAC address binding of A When a new machine appears on the network (e.g. when an operating system reboots) it can broadcast its IP_to_MAC address binding so that all other machines can store it in their caches. This will eliminate a lot of ARP packets by all other machines, whenthey want to communicate with this new machine.

10.3 REVERSE ADDRESS RESOLUTION PROTOCOL

The Reverse Address Resolution Protocol (RARP) allows a host to discover its Internet address when it knows only its physical address. It is used when a computer is connected to a network for the first time or whena diskless computer is booted.

RARP is a protocol by which a physical machine in a local area network canrequest to learn its IP address from a gateway server's Address Resolution Protocol table or cache. This is needed since the machine may not have permanently attached disk where it can store its IP address permanently. A network administrator creates a table in a local area network's gateway routerthat maps the physical machine (or Medium Access Control - MAC) addresses to corresponding Internet Protocol addresses. When a new machine is set up, its RARP client program requests from the RARP server on the router to be sent its IP address. Assuming that an entry has been set up in the router table, the RARP server will return the IP address to the machine which can store it for future use.

Drawbacks of RARP

- Since it operates at low level, it requires direct addresses to the network which makes it difficult for an application programmer tobuild a server.
- It doesn't fully utilizes the capability of a network like Ethernet which is enforced to send a minimum packet size since the replyfrom the server contains only one small piece of information, the32-bit internet address.

10.4 INTERNET CONTROL MESSAGE PROTOCOL (ICMP)

The Internet Control Message Protocol (ICMP) is a mechanism used by hosts and gateways to send notification of datagram problems back to thesender. ICMP sends query and error reporting messages.

This protocol discusses a mechanism that gateways and hosts use to communicate control or error information. The Internet protocol provides unreliable, connectionless datagram service, and that a datagram travels fromgateway to gateway until it reaches one that can deliver it directly to its

finaldestination. If a gateway cannot route or deliver a datagram, or if the gatewaydetects an unusual condition, like network congestion, that affects its abilityto forward the datagram, it needs to instruct the original source to take actionto avoid or correct the problem. The Internet Control Message Protocol allows gateways to send error or control messages to other gateways or hosts; ICMP provides communication between the Internet Protocol software on one machine and the Internet Protocol software on another. This is a special purpose message mechanism added by the designers to the TCP/IP protocols.This is to allow gateways in an internet to report errors or provide information about unexpected circumstances. The IP protocol itself containsnothing to help the sender test connectivity or learn about failures.

Error reporting vs. Error Correction

ICMP only reports error conditions to the original source; the source must relate errors to individual application programs and take action to correct problems. It provides a way for gateway to report the error It does not fully specify the action to be taken for each possible error. ICMP is restricted to communicate with the original source but not intermediate sources. **ICMP Message Delivery**

ICMP messages travel across the internet in the data portion of an IP datagram, which itself travels across the internet in the data portion of an IPdatagram, which itself travels across each physical network in the data portion of a frame. Datagram carrying ICMP messages are routed exactly like datagram carrying information for users; there is no additional reliabilityor priority. An exception is made to the error handling procedures if an IP datagram carrying an ICMP messages are not generated for errors that resultfrom datagram carrying ICMP error messages.

ICMP Message Format

It has three fields; an 8-bit integer message TYPE field that identifies the message, an 8-bit CODE field that provides further information about the message type, and a 16-bit CHECKSUM field (ICMP uses the same additivechecksum algorithm as IP, but the ICMP checksum only covers the ICMP message).In addition, ICMP messages that report errors always include the header and first 64 data bits of the datagram causing the problem. The ICMPTYPE field defines the meaning of the message as well as its format.

The Types include:

TYPE FIELD	ICMP MESSAGE TYPE
0	ECHO REPLY
3	DESTINATION UNREACHABLE
4	SOURCE QUENCH

5	REDIRECT(CHANGE A ROUTE)
8	ECHO REQUEST
11	TIME EXCEEDED FOR A DATAGRAM
12	PARAMETER PROBLEM ON A DATAGRAM
13	IMESTAMP REQUEST
14	TIMESTAMP REPLY
15	INFORMATION REQUEST(OBSOLETE)
16	INFORMATION REPLY(OBSOLETE)
17	ADDRESS MASK REQUEST
18	ADDRESS MASK REPLY TESTING DESTINATION

10.5 INTERNET GROUP MESSAGE PROTOCOL (IGMP)

The Internet Group Message Protocol (IGMP) is used to facilitate thesimultaneous transmission of a message to a group of recipients.

It is a communications protocol used by hosts and adjacent routers on IPv4 networks to establish multicast group memberships. IGMP is an integral partof IP multicast.

IGMP can be used for one-to-many networking applications such as online streaming video and gaming, and allows more efficient use of resources when supporting these types of applications.

IGMP is used on IPv4 networks. Multicast management on IPv6 networks ishandled by Multicast Listener Discovery (MLD) whichuses ICMPv6 messaging in contrast to IGMP's bare IP encapsulation.

Types of messages in IGMP

IGMP has three types of messages: the query, the membership report, andthe leave report. There are two types of query messages as shown in the figure below: general and special.

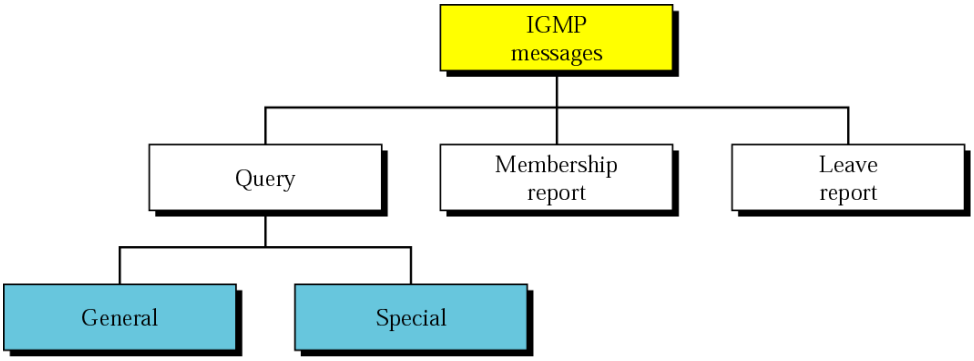


Fig: Types of messages in IGMP

10.6 OPEN SHORTEST PATH FIRST (OSPF)

The Open Shortest Path First OSPF protocol is an intradomain routing protocol, its based on the concept of link state routing protocol. Its domain is also called an autonomous system. OSPF protocol divides an autonomous system into areas and subsections. An area is a group of networks, hosts, and routers all contained within an autonomous system. An autonomous system canbe split into many different areas. All networks inside an area must be connected to each other through a link. Routers inside an area flood the area with the help of routing information. At the border or boundary of an area, special routers called area border routers. the areas inside an autonomous system is a special area called the backbone AS. All the areas inside an AS must be connected to the backbone of the system. The area identification of the backbone is zero. Figure 8.8 shows an autonomous system and its areas.

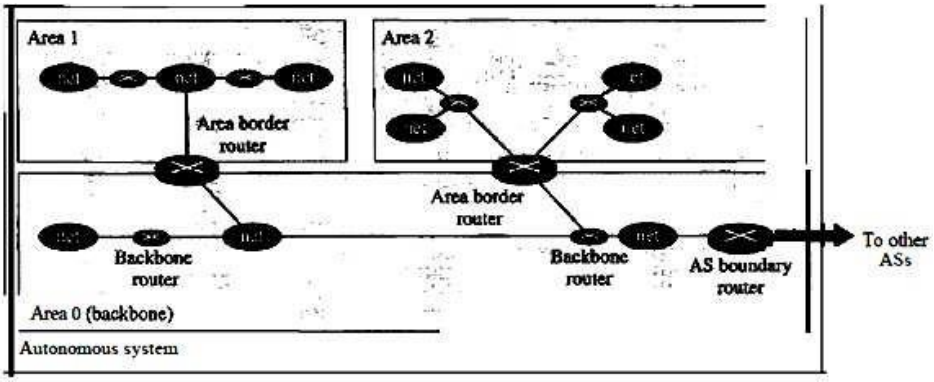
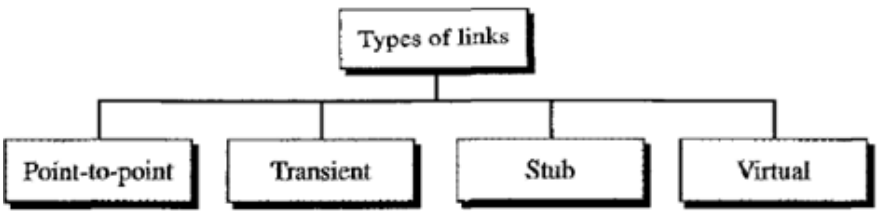


Fig. Areas in an autonomous system

Types of Links in OSPF:

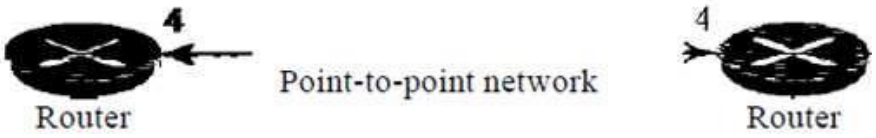
A connection is called a link. Four types of links have beendefined:

- point-to-point,
- transient,
- stub, and
- virtual



- A point-to-point link connects two routers without any help of any other host or router in

between. An example of this type of linkis two routers connected to each other by a telephone line or a T line.There is no requirement to assign a network address to this type of link.



- **Fig. Point-to-point link**

- A transient link is a network link with several routers attached with each other. The data can come into the network through any of the routers and leave through any router. For example, consider theEthernet in Figure 18. Router A has routers B, C, D, and E asneighbor's nodes. Router B has routers A, C, D, and E as neighbor's nodes.

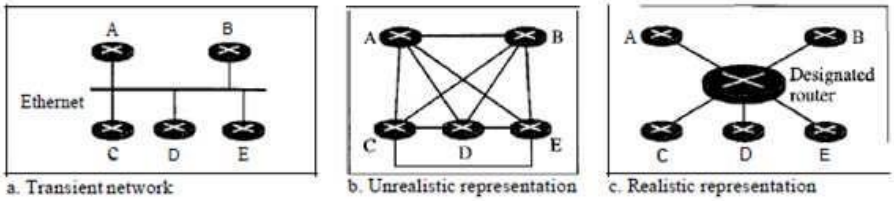


Fig. Transient link

- A stub link is a network that is connected to only one router inthe network. The data packets come into the network through this single router and leave the network through this same router.

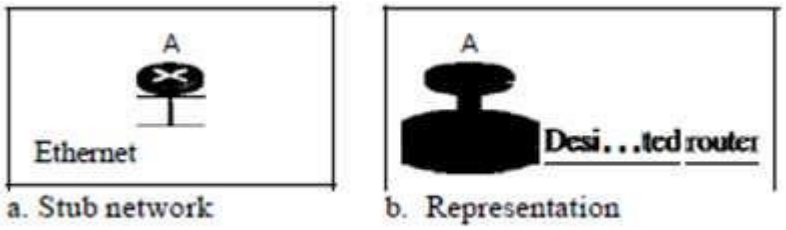


Fig. Stub link

When the link between two routers is broken, the management may create a virtual link between two routers, using a longer path that probably goes through several routers.

10.7 BORDER GATEWAY PROTOCOL (BGP)

Border Gateway Protocol is an interdomain routing protocol,

It uses path vector routing. For example, a large business that manages its own network and has full control over it is an autonomous system. A local ISP that provides services to local customers is called an autonomous system. This divides autonomoussystems into three categories: stub, multihomed, and transit.

Stub AS. A stub AS has only one connection to another AS. The interdomain data traffic in a stub AS can be either created or terminated in the AS.

Multihomed AS. A multihomed AS has more than one connection to other ASs, but it is still only a source or sink for data traffic.

Transit AS. A transit AS is a multihomed AS that also allowstemporary traffic. Good examples of transit ASs are national and international ISPs (Internet backbones).

External and Internal BGP

BGP has two types of sessions: external BGP (E-BGP) and internal BGP sessions. The E-BGP session is used to interchange information between two speaker nodes. It belongs to two different AS. The I-BGP session is used to exchange routing informationbetween two routers inside an AS. Figure 21 shows the details

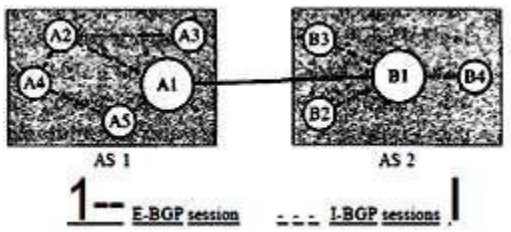


Fig. Internal and external BGP sessions

BGP squint advertises routes to each other on the network. Two of the more important attributes are AS-PATH and NEXT- HOP:

AS-PATH:

This attribute contains the ASs through which the advertisement for the prefix has passed. When a prefix is passed intoan AS, the AS adds its ASN to the ASPATH attribute.

NEXT-HOP:

Providing the critical link between the inter-AS and intra-AS routing protocols, the NEXT-HOP

attribute has a subtle but important use. The NEXT-HOP is the router interface that begins the AS-PATH.

BGP Route Selection

AS BGP uses eBGP and iBGP to distribute routes to all the routers. From this issue, a router may learn about more than one route to any one prefix, in which case the router must select one of the possible routes. The input into the route selection process is the set of all routes that have been acquired and accepted by the router.If there are more than two routes to the same prefix, then BGP sequentially invokes the following elimination rules until one route remains:

- Routes are assigned a local priority value as one of their attributes. The local priority of a route could have been set by therouter. This is a policy decision that is boosted up to the AS's network administrator. The routes with the giant local preference values are selected.
- The route with the shortest path (AS-PATH) is selected. If this rule applies the only rule for route selection, then BGP would be using a DV algorithm for path determination, where the distance metric uses the number of AS hops preferably than the number ofrouter hops.
- From the endure routes the route with the closest NEXT-HOP router is selected. Here, closest means cost of the least-cost path, set on by the intra-AS algorithm, is the smallest
- If more than one route is still leftover, the router uses BGPidentifiers to select the route.

10.8 SUMMARY

In this unit we have discussed in detail ARP protocol and also covered RARP and also discussed about ICMP and IGMP protocols. At the end of this unit also covered in detail about OSPF and BGP protocols.

10.9 KEYWORDS

ARP, RARP, ICMP, IGMP, OSPF and BGP

10.10 QUESTIONS

1. Describe ARP protocol.
2. Explain RARP protocol.

3. Discuss ICMP protocol.
4. Explain Types of messages in IGMP
5. Briefly explain types of link in OSPF.
6. Elucidate BGP protocol.

10.11 REFERENCE

- Introduction to Data Communications and Networking by Behrouz Forouzan.
- Computer Networks by Andrew S Tanenbaum.
- Networking Essentials – Third Edition – Jeffrey S. Beasley, Piyasat Nilkaew

UNIT – 11: PACKET SWITCHING NETWORKS

Structure

11.0 Objectives

11.1 Introduction

11.2 Packet switching networks

11.3 Network services and internal network operations

11.4 Summary

11.5 Keywords

11.6 Questions

11.7 Reference

11.0 OBJECTIVES

At the end of this unit you will be able to

- Elucidate Packet switching networks
- Discuss Network services and network operations

11.1 INTRODUCTION

Packet switching is the transfer of small pieces of data across various networks. These data chunks or “packets” allow for faster, more efficient data transfer. Often, when a user sends a file across a network, it gets transferred in smaller data packets, not in one piece. For example, a 3MB file will be divided into packets, each with a packet header that includes the origin IP address, the destination IP address, the number of packets in the entire data file, and the sequence number.

There are two major types of packet switching:

Connectionless Packet Switching. This classic type of packet switching includes multiple packets, each individually routed. This means each packet contains complete routing information—but it also means different paths of transmission and out-of-order delivery are possible, depending on the fluctuating loads on the network’s nodes (adapters, switches and routers) at the moment. This kind of packet switching is sometimes called datagram switching.

Each packet in connectionless packet switching includes the following information in its header section:

- Source address
- Destination address
- Total number of packets
- Sequence number (Seq#) for reassembly

Once the packets reach their destination via various routes, the receiving devices rearrange them to form the original message.

Connection-Oriented Packet Switching. In connection-oriented packet switching, also called virtual circuit switching or circuit switching, data packets are first assembled and then numbered. They then travel across a predefined route, sequentially. Address information is not needed in circuit switching, because all packets are sent in sequence.

11.2 PACKET SWITCHING NETWORKS

Traditional telephone networks operate on the basis of circuit switching. A call setup process reserves resources (time slots) along a path so that the stream of voice samples can be transmitted with very low delay across the network. The resources allocated to a call cannot be used by other users for the duration of the call. This approach is inefficient when the amount of information transferred is small or if information is produced in bursts, as is the case in many computer

applications. In this chapter we examine networks that transfer blocks of information called packets. Packet-switching networks are better matched to computerapplications and can also be designed to support real-time applications such as telephony.

We can view packet networks from two perspectives. One perspective involves an external view of the network and is concerned with the services that the network provides to the transport layer that operates above it at the end systems. Here we are concerned with whether the network service requires the setting up of a connection and whether the transfer of user data is providedwith quality-of-service guarantees. Ideally the definition of the network servicesis independent of the underlying network and transmission technologies. Thisapproach allows the transport layer and the applications that operate above it to be designed so that they can function over any network that provides the givenservices.

A second perspective on packet networks is concerned with the internal operation of a network. Here we look at the physical topology of a network, the interconnection of links, switches, and routers. We are concerned with theapproach that is used to direct information across the network: datagrams, orvirtual circuits. We are also concerned with addressing and routing procedures, as well as with dealing with congestion inside the network. We must also managetraffic flows so that the network can deliver information with the quality of service it has committed to.

It is useful to compare these two perspectives in the case of broadcast net- works and LANs from the previous chapter and the switched packet networksconsidered here. The first perspective, involving the services provided to the layerabove, does not differ in a fundamental way between broadcast and switchedpacket networks. The second perspective, however, is substantially different. In the case of LANs, the network is small, addressing is simple, and the frame istransferred in one hop so no routing is required. In the case of packet-switching networks, addressing must accommodate extremely large-scale networks and must work in concert with appropriate routing algorithms. These two challenges, addressing and routing, are the essence of the network layer.

In this chapter we deal with the general issues regarding packet-switching networks. Later chapters deal with specific architectures, namely, Internet Protocol (IP) packet networks and asynchronous transfer mode (ATM) packetnetworks. The chapter is organized as follows:

- Network services and internal network operation. We elaborate on the two perspectives on networks, and we discuss the functions of the network layer, including internetworking.
- Physical view of networks. We examine typical configurations of packet- switching networks. This section defines the role of multiplexers, LANs, switches, and routers in network and internetwork operation.
- Datagrams and virtual circuits. We introduce the two basic approaches to operating a packet network, and we use IP and ATM as examples of theseapproaches.

Packet switching is a method of transferring the data to a network in form of packets. In order to transfer the file fast and efficiently manner over the network and minimize the transmission latency, the data is broken into small pieces of variable length, called **Packet**. At the destination, all these small parts (packets) have to be reassembled, belonging to the same file. A packet composes of payload and various control information. No pre-setup or reservation of resources is needed.

Packet Switching uses **Store and Forward** technique while switching the packets; while forwarding the packet each hop first stores that packet then forward. This technique is very beneficial because packets may get discarded at any hop due to some reason. More than one path is possible between a pair of sources and destinations. Each packet contains Source and destination address using which they independently travel through the network. In other words, packets belonging to the same file may or may not travel through the same path. If there is congestion at some path, packets are allowed to choose different paths possible over an existing network.

Packet-Switched networks were designed to overcome the *weaknesses* of Circuit-Switched networks since circuit-switched networks were not very effective for small messages.

Advantage of Packet Switching over Circuit Switching:

- More efficient in terms of bandwidth, since the concept of reserving circuit is not there.
- Minimal transmission latency.
- More reliable as a destination can detect the missing packet.
- More fault tolerant because packets may follow a different path in case any link is down, Unlike Circuit Switching.
- Cost-effective and comparatively cheaper to implement.

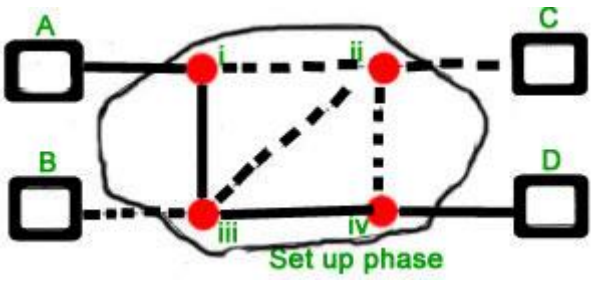
The disadvantage of Packet Switching over Circuit Switching :

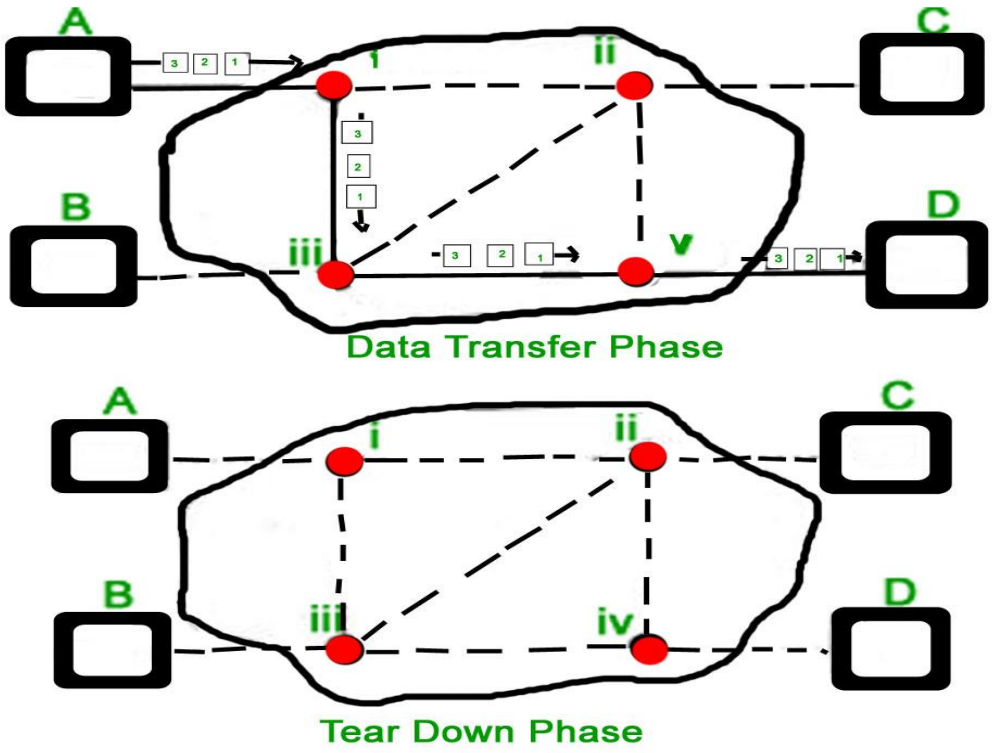
- Packet Switching doesn't give packets in order, whereas Circuit Switching provides ordered delivery of packets because all the packets follow the same path.
- Since the packets are unordered, we need to provide sequence numbers for each packet.
- Complexity is more at each node because of the facility to follow multiple paths.
- Transmission delay is more because of rerouting.
- Packet Switching is beneficial only for small messages, but for bursty data (large messages) Circuit Switching is better.

Modes of Packet Switching :

1. Connection-oriented Packet Switching (Virtual Circuit) :

Before starting the transmission, it establishes a logical path or virtual connection using signaling protocol, between sender and receiver and all packets belongs to this flow will follow this predefined route. Virtual Circuit ID is provided by switches/routers to uniquely identify this virtual connection. Data is divided into small units and all these small units are appended with help of sequence numbers. Overall, three phases take place here- The setup, data transfer and tear down phase.





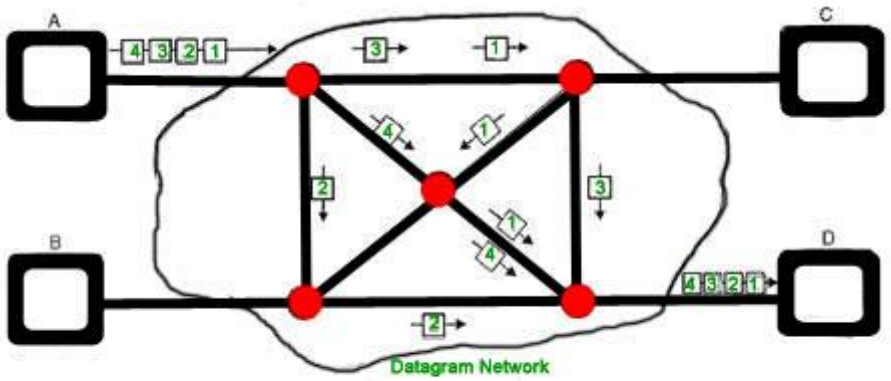
Phases in virtual circuit packet switching

All address information is only transferred during the setup phase. Once the route to a destination is discovered, entry is added to the switching table of each intermediate node. During data transfer, packet header (local header) may contain information such as length, timestamp, sequence number, etc.

Connection-oriented switching is very useful in switched WAN. Some popular protocols which use the Virtual Circuit Switching approach are X.25, Frame-Relay, ATM, and MPLS(Multi-Protocol Label Switching).

Connectionless Packet Switching (Datagram):

Unlike Connection-oriented packet switching, In Connectionless Packet Switching each packet contains all necessary addressing information such as source address, destination address and port numbers, etc. In Datagram Packet Switching, each packet is treated independently. Packets belonging to one flow may take different routes because routing decisions are made dynamically, so the packets arrived at the destination might be out of order. It has no connection setup and teardown phase, like Virtual Circuits.
Packet delivery is not guaranteed in connectionless packet switching, so reliable delivery must be provided by end systems using additional protocols.



Datagram Packet Switching

A---R1---R2---B

A is the sender (start)

R1, R2 are two routers that store and forward data

B is receiver(destination)

Delays in Packet switching :

1. Transmission Delay
2. Propagation Delay

3. Queuing Delay
4. Processing Delay

11.3 NETWORK SERVICES AND INTERNAL NETWORK OPERATION

The essential function of a network is to transfer information among the usersthat are attached to the network or internetwork. In Figure 7.1 we show that thistransfer may involve a single block of information or a sequence of blocks thatare temporally related. In the case of a single block of information, we are interested in having the block delivered correctly to the destination, and we may also be interested in the delay experienced in traversing the network. In the case of a sequence of blocks, we may be interested not only in receiving the blocks correctly and in the right sequence but also in delivering a relatively unimpaired temporal relation.

Figure 7.2 shows a transport protocol that operates end to end across a network. The transport layer peer processes at the end systems accept messagesfrom their higher layer and transfer these messages by exchanging segments endto end across the network. The figure shows the interface at which the networkservice is visible to the transport layer. The network service is all that matters tothe transport layer, and the manner in which the network operates to provide the service is irrelevant.

The network service can be connection-oriented or connectionless. A connectionless service is very simple, with only two basic interactions between the transport layer and the network layer: a request to the network that it send a packet and an indication from the network that a packet has arrived. The usercan request transmission of a packet at any time, and does not need to inform the network layer that the user intends to transmit information ahead of time. A connectionless service puts total responsibility for error control, sequencing, and flow control on the end-system transport layer.

The network service can be connection-oriented. In this case the transportlayer cannot request transmission of information until a connection has been set up. The essential points here are that the network layer must be informed about the new flow that is about to be applied to the network and that the network layer maintains state information about the flows it is handling. During call setup, parameters related to usage and quality of service may be negotiated and network resources may be allocated to ensure that the user flow can be handled as required. A connection-release procedure may also be required to terminate the connection. It is clear that providing connection-oriented serviceentails greater complexity than connectionless service in the network layer.

It is also possible for a network layer to provide a choice of services to theuser of the network. For example, the network layer could offer: (1) best-effortconnectionless service; (2) low-delay connectionless service; (3) connection-oriented reliable stream service; and (4) connection-oriented transfer of packets with delay and bandwidth guarantees. It is easy to come up with examples of applications that can make use of each of these services. However, it does not follow that all the

services should be offered by the network layer. Two inter- related reasons can be given for keeping the set of network services to a mini-mum: the end-to-end argument and the need for network scalability.

When applied to the issue of choice of network services, the end-to-end argument suggests that functions should be placed as close to the applicationas possible, since it is the application that is in the best position to determine whether a function is being carried out completely and correctly. This argumentsuggests that as much functionality as possible should be located in the transportlayer or higher and that the network services should provide the minimum functionality required to meet application performance.

Up to this point we have considered only the services offered by the network layer. Let us now consider the internal operation of the network. Figure 7.3 shows the relation between the service offered by the network and the internaloperation. We say that the internal operation of a network is connectionless ifpackets are transferred within the network as datagrams. Thus in the figure each packet is routed independently. Consequently packets may follow different paths from α to β and so may arrive out of order. We say that the internal operation ofa network is connection-oriented if packets follow virtual circuits that have been

established from a source to a destination. Thus to provide communications between α and β, routing to set up a virtual circuit is done once, and thereafterpackets are simply forwarded along the established path. If resources are reserved during connection setup, then bandwidth, delay, and loss guarantees can be provided.

The fact that a network offers connection-oriented service, connectionless service, or both does not dictate how the network must operate internally. In discussing TCP and IP, we have already seen that a connectionless packet net- work (e.g., IP) can support connectionless service (UDP) as well as connection- oriented service (TCP). We will also see that a connection-oriented network (e.g.,ATM) can provide connectionless service as well as connection-oriented service. We discuss virtual-circuit and datagram network operation in more detail in a later section. However, it is worthwhile to compare the two at this point at a high level.

The approach suggested by the end-to-end argument keeps the network service (and the network layer that provides the service) as simple as possible while adding complexity at the edge only as required. This strategy fits very well with the need to grow networks to very large scale. We have seen that the value of a network grows with the community of users that can be reached and withthe range of applications that can be supported. Heaping the core of the network simple and adding the necessary complexity at the edge enhances the scalabilityof the network to larger size and scope.

This reasoning suggests a preference for a connectionless network, which hasmuch lower complexity than a connection-oriented network. The reasoning does allow the possibility for some degree of ‘‘connection orientation‘‘ as a means toensure that applications can receive the proper level of

performance. Indeed current research and standardization efforts (discussed in Chapter 10) can be viewed as an attempt in this direction to determine an appropriate set of network services and an appropriate mode of internal network operation.

We have concentrated on high-level arguments up to this point. What do these arguments imply about the functions that should be in the network layer?Clearly, functions that need to be carried out at every node in the network mustbe in the network layer. Thus functions that route and forward packets need tobe done in the network layer. Priority and scheduling functions that direct how packets are forwarded so that quality of service is provided also need to be in thenetwork layer. Functions that belong in the edge should, if possible, be implemented in the transport layer or higher. A third category of functions can be implemented either at the edge or inside the network. For example, while congestion takes place inside the network, the remedy involves reducing input flows at the edge of the network. We will see that congestion control has been implemented in the transport layer and in the network layer.

Another set of functions is concerned with making the network service independent of the underlying transmission systems. For example, different

transmissions systems (e.g., optical versus wireless) may have different limits on the frame size they can handle. The network layer may therefore be called upon to carry out segmentation inside the network and reasssembly at the edge. Alternatively, the network could send error messages to the sending edge, requesting that the packet size be reduced. A more challenging set of functions arises when the ‘‘network‘‘ itself may actually be an internetwork. In this case the network layer must also be concerned not only about differences in the size of theunits that the component networks can transfer but also about differences in addressing and in the services that the component networks provide.

11.6 SUMMARY

At the end of this unit we have learnt packet switching networks in that we gone through advantages and disadvantages of packet switching networks and also modes of packet switching. In the last section of this unit we have discussed network services and internal network operations.

11.7 KEYWORDS

Packet switching, Packet, Congestion, virtual circuit and datagram

11.8 QUESTIONS

1. Describe packet switching networks.
2. With neat diagram explain connection-oriented packet switching.
3. Write the advantages and disadvantages of packet switching.
4. Briefly Explain network services.

5. Discuss internal network operations.

11.9 REFERENCE

- Introduction to Data Communications and Networking by Behrouz Forouzan.
- Computer Networks by Andrew S Tanenbaum.
- Networking Essentials – Third Edition – Jeffrey S. Beasley, Piyasat Nilkaew

UNIT – 12: PACKET NETWORK TOPOLOGY

Structure

12.0 Objectives

12.1 Introduction

12.2 Packet network topology

12.3 Datagrams and Virtual circuits

12.4 Connectionless packet switching

12.5 Summary

12.6 Keywords

12.7 Questions

12.8 Reference

12.0 OBJECTIVES

At the end of this unit you will be able to

- Describe Packet network topology
- Elucidate Datagrams and Virtual circuits
- Discuss Connectionless packet switching

12.1 INTRODUCTION

This topology comprises multiple domains consisting of routers interconnected by point-to-point data links, LANs, and wide area networks such as ATM. The principal task of a packet-switching network is to provide connectivity among users.

12. 2 PACKET NETWORK TOPOLOGY

This section considers existing packet-switching networks. We present an end-to-end view of existing networks from a personal computer, workstation, or serverthrough LANs and the Internet and back.

First let us consider the way in which users access packet networks. Figure12.1 shows an access multiplexer where the packets from a number of users share a transmission line. This system arises for example, in X.25, frame relay, and ATM networks, where a single transmission line is shared in the access to a wide area packet-switching network. The multiplexer combines the typically burst flows of the individual computers into aggregated flows that make efficient use of thetransmission line. Note that different applications within a single computer cangenerate multiple simultaneous flows to different destinations. From a logical point of view, the link can be viewed as carrying either a single aggregated flowor a number of separate packet flows. The network access node forwards packets into a backbone packet network.

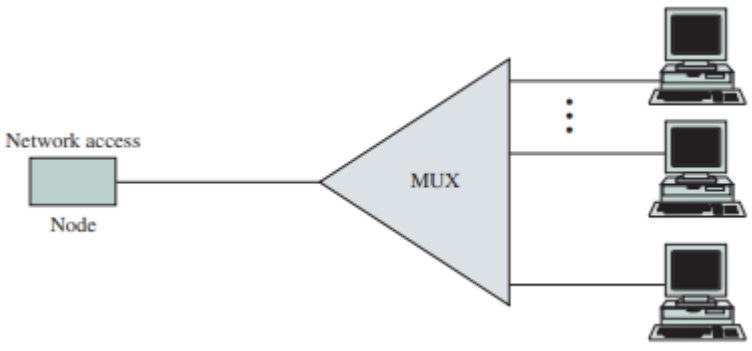


Fig12.1 : Access Multiplexer

Local area networks (LANs) provide the access to packet-switching networksin many environments. As shown in Figure 12.2a, computers are connected to ashared transmission medium. Transmissions are broadcast to all computers in the network. Each computer is identified by a unique physical address, and soeach station listens for its address to receive transmissions. Broadcast and multi-cast transmissions are easily provided in this environment.

LANs allow the sharing of resources such as printers, databases, and soft- ware among a small community of users. LANs can be extended through the use of bridges or LAN switches, as shown in Figure 12.2b. Here the LAN switch forwards inter-LAN traffic based on the physical address of the frames. Traffic local to each LAN stays local, and broadcast transmissions are forwarded to the other attached LANs. Switches can interconnect more than two LANs.

Multiple LANs in an organization, in turn, are interconnected into campus networks with a structure such as that shown in Figure 7.6. LANs for a large group of users such as a department are interconnected in an extended LAN through the use of LAN switches, identified by lowercase s in the figure. Resources such as servers and databases that are primarily of use to this department are kept within the subnetwork. This approach reduces delays in accessing the resources and contains the level of traffic that leaves the subnetwork. Each subnetwork has access to the rest of the organization through a router R that accesses the campus backbone network. A subnetwork also uses the campus backbone to reach the ‘‘outside world‘‘ such as the Internet or other sites belonging to the organization through a gateway router. Depending on the type of organization, the gateway may implement firewall functions to control the traffic that is allowed into and out of the campus network.

Servers containing critical resources that are required by the entire organization are usually located in a data center where they can be easily maintained and where security can be enforced. As shown in Figure 12.3, the critical servers may be provided with redundant paths to the campus backbone network. These servers are usually placed near the backbone network to minimize the numberof hops required to access them from the rest of the organization.

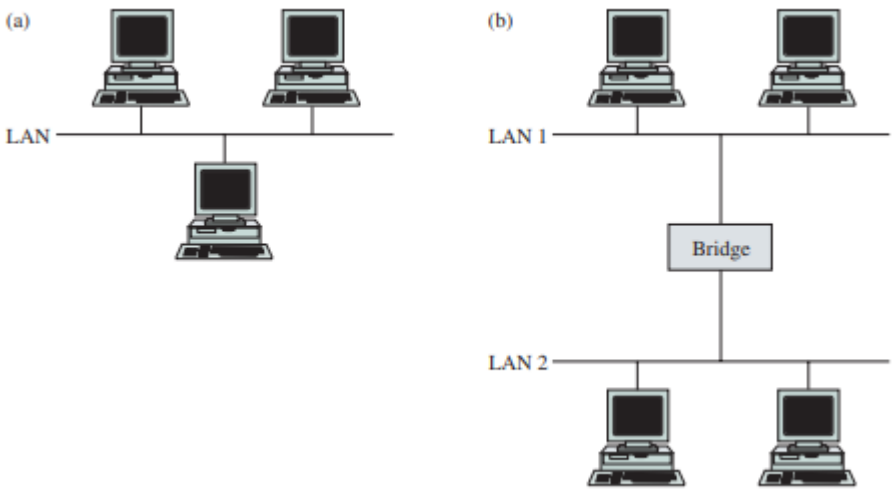


Fig:12.2 Local area networks

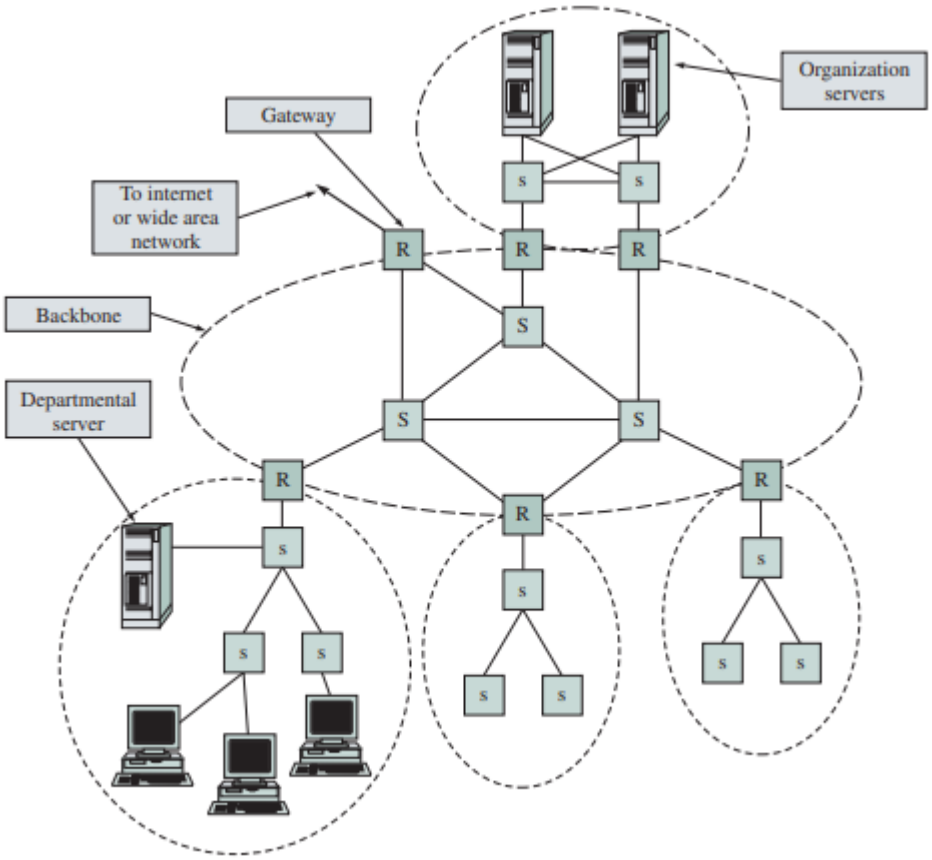


Fig:12.3 Campus networks

The traffic within an extended LAN is delivered based on the physical LANaddresses. However, applications in host computers operate on the basis of logical IP addresses. Therefore, the physical address corresponding to an IP address needs to be determined every time an IP packet is to be transmitted over a LAN. This address resolution problem can be solved by using IP address to physical address translation tables. In the next chapter we discuss the AddressResolution Protocol that IP uses to solve this problem.

The routers in the campus network are interconnected to form the campus backbone network, depicted by the mesh of switches, designated S, in Figure 12.3.Typically, for large organizations such as universities these routers are interconnected by using very high speed LANs, for example, Gigabit Ethernet or an ATM network. The routers use the Internet Protocol (IP), which enables them to operate over various data link and network technologies. The routers exchange information about the state of their links to dynamically calculate routing tablesthat direct packets across the campus network. This approach allows the net-work to adapt to changes in traffic pattern as well as changes in topology due to faults in equipment.

The routers in the campus network form a domain or autonomous system. The term domain indicates that the routers run the same routing protocol. Theterm autonomous system is used for one or more domains under a single administration. All routing decisions inside the autonomous system are

independent of any other network.

Organizations with multiple sites may have their various campus networksinterconnected through routers interconnected by leased digital transmission lines or frame relay connections. In this case access to the wide area network may use an access multiplexer such as the one shown in Figure 7.4. In addition the campus network may be connected to an Internet service provider through one or more border routers as shown in Figure 12.4. To communicate with othernetworks, the autonomous system must provide information about its networkroutes in the border routers. The border router communicates on an interdomain level, whereas other routers in a campus network operate at the intradomain level.

A national ISP provides points of presence (POPs) in various cities wherecustomers can connect to their network. The ISP has its own national networkfor interconnecting its POPs. This network could be based on ATM; it might use IP over SONET; or it might use some other network technology. The ISPs in turn exchange traffic as network access points (NAPs), as shown in Figure 12.5

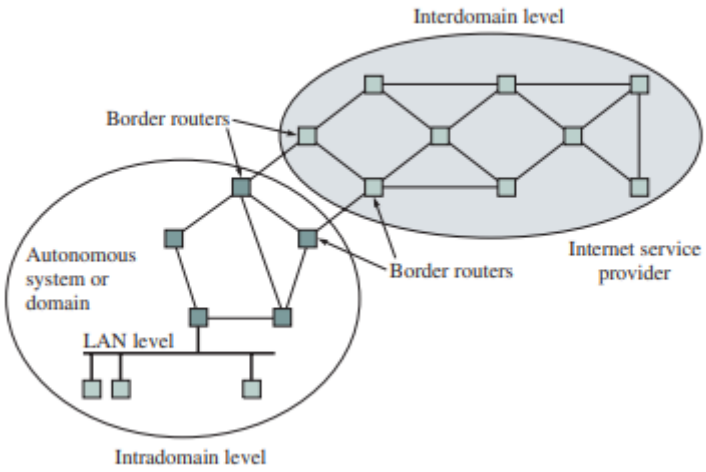


Fig12.4: Intradomain and interdomain levels

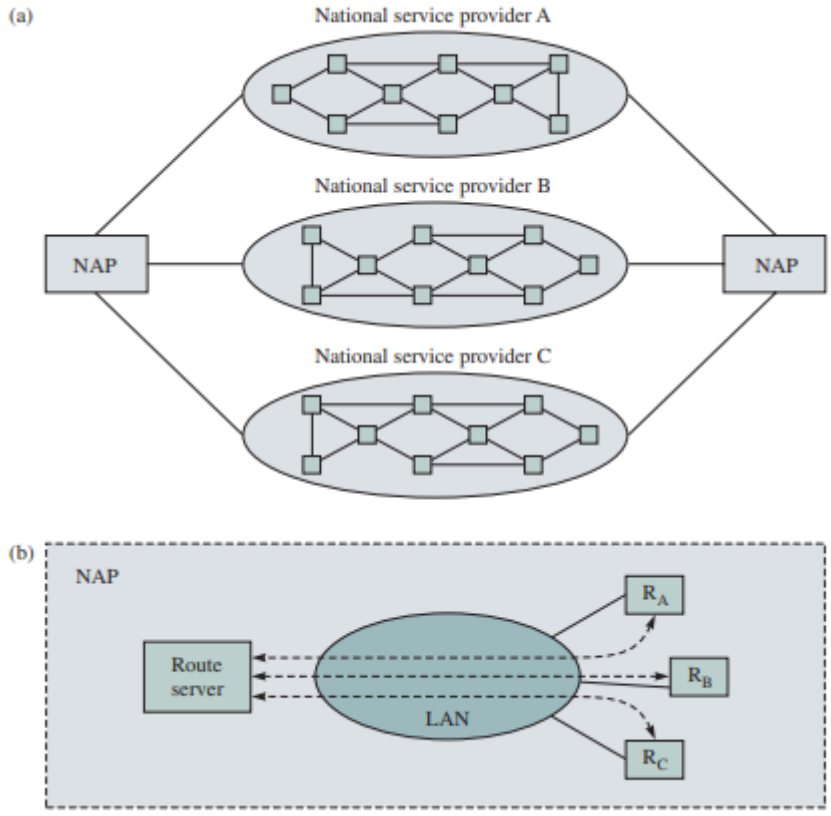


Fig12.5: National ISPs exchange traffic at NAPs; Routing information is exchanged through route servers

A NAP is a high-speed LAN or switch at which the routers from different ISPs can exchange traffic, and as such NAPs are crucial to the interconnectivityprovided by the Internet. (As discussed in Chapter 1, four NAPs were originallyset up by the National Science Foundation). The ISPs interconnected to a NAPneed to exchange routing information. If there are n such ISPs, then n n 1 pairwise route exchanges are required. This problem is solved by introducing aroute server as shown in Figure 7.8b. Each ISP sends routing information to theroute server, which knows the policies of every ISP. The route server in turndelivers the processed routing information to the ISPs. Note that a national service provider also has the capability of interconnecting a customer‘s various sites by using its own IP network, so the customer‘s sitesappear as a single private network. This configuration is an example of a virtualprivate network (VPN). Small office and home (SOHO) users obtain packet access through ISPs. Theaccess is typically through modem dial-up, but it could be through ADSL, ISDN, or cable modem. When a customer connects to an ISP, the customer is assigned an IP address for the duration of the connection.[1] Addresses are sharedin this way because the ISP has only a limited number of addresses. If the ISP is only a local provider, then it must connect to a regional or national provider andeventually to a NAP. Thus we see that a multilevel hierarchical network topology arises for the Internet which is much more decentralized than traditional telephone networks.This topology comprises multiple domains consisting of routers interconnected

by point-to-point data links, LANs, and wide area networks such as ATM.

The principal task of a packet-switching network is to provide connectivityamong users. The preceding description of the existing packet-switching network infrastructure reveals the magnitude of this task. Routers exchange information among themselves and use routing protocols to build a consistent set of routing tables that can be used in the routes to direct the traffic flows in these networks. The routing protocols must adapt to changes in network topology due to the introduction of new nodes and links or to failures in equipment. Different routing algorithms are used within a domain and between domains. A key concernhere is that the routing tables result in stable traffic flows that make efficient useof network resources. Another concern is to keep the size of routing tables manageable even as the size of the network continues to grow at a rapid pace.In this chapter we show how hierarchical addressing structures can help address this problem. A third concern is to deal with congestion that inevitably occurs inthe network. It makes no sense to accept packets into the network when they are likely to be discarded. Thus when congestion occurs inside the network, that is,buffers begin filling up as a result of a surge in traffic or a fault in equipment, thenetwork should react by applying congestion control to limit access to the net-work only to traffic that is likely to be delivered. A final concern involves providing the capability to offer Quality-of-Service guarantees to some packet flows. We deal with these topics also in the remainder of the chapter

12.3 DATAGRAMS AND VIRTUAL CIRCUITS

A network is usually represented as a cloud with multiple input sources and output destinations as shown in Figure 12.6. A network can be viewed as a generalization of a physical cable in the sense of providing connectivity between multiple users.

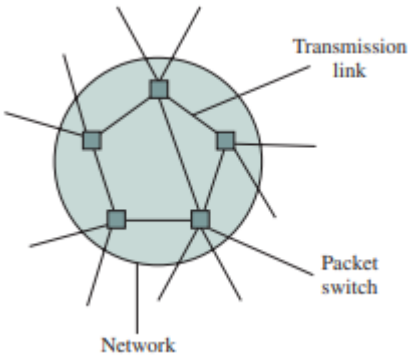


Fig 12.6: Switched networks

Unlike a cable, a switched network is geographically distributed and consists of a graph of transmission lines (i.e., links) interconnected by switches (nodes). These transmission and switching resources are configured to enable the flow of information among users. Networks provide for the interconnection of sources to destinations on a dynamic basis. Resources are typically allocated to an information flow only when needed. In this manner the resources are shared among the

communityof users resulting in efficiency and lower costs. There are two fundamental approaches to transferring information over a packet-switched network. The first approach, called connection-oriented, involves setting up a connection acrossthe network before information can be transferred. The setup procedure typically involves the exchange of signaling messages and the allocation of resources along the path from the input to the output for the duration of the connection. The second approach is connectionless and does not involve a prior allocation of resources. Instead a packet of information is routed independently from switch to switch until the packet arrives at its destination. Both approaches involve theuse of switches or routers to direct packets across the network.

Structure of Switch†Router

Figure shows a generic switch consisting of input ports, output ports, an interconnection fabric, and a switch controller/processor. Input ports and output ports are usually paired. A line card typically handles several input/output ports. The line card implements physical and data link layer functions. Thus the card is concerned with symbol timing and line coding. It is also concerned with framing, physical layer addressing, and error checking. For widely deployed standards,the line card also implements medium access control and data link protocols inhardware with a special-purpose chip set. The line card also contains some buffering to handle the speed mismatch between the transmission line and the interconnection fabric. The controller/processor can carry out a number of functions depending on the type of packet switching. The function of the interconnection fabric is to transfer packets between the line cards. Note that Figure 12.7 shows an ‘‘unfolded‘‘ version of the switch in which the line cards appear twice,once with input ports and again with output ports. In the actual implementation the transmit and receive functions take place in a single line card. However, thefunction of various types of switch architectures is easier to visualize this way.

We elaborate on the operation of the switches as we develop the two approachesto packet switching. A simple switch can be built by using a personal computer or a workstationand inserting several network interface cards (NICs) in the expansion slots as shown in Figure 12.8.

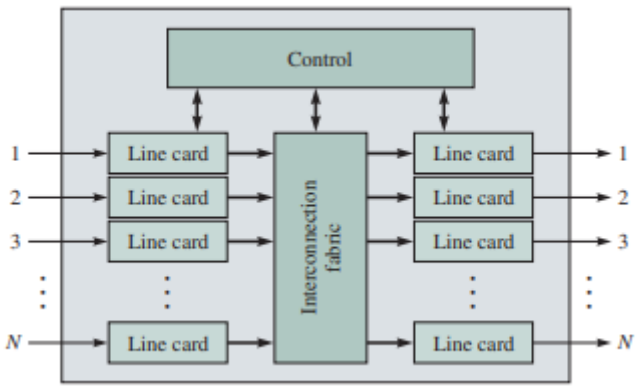


Fig12.7: Components of generic switch/router

The frames that arrive at the NICs are de-encapsulated,and the packets are transferred by using the I/O bus from the NIC to main memory. The processor performs the required routing and protocol processing, formats the packet header, and then forwards the packet by transferring it frommain memory to the appropriate NIC.

The simple setup in Figure 12.9 reveals the three basic resources and potential bottlenecks in switches: processing, memory, and bus (interconnection) bandwidth. Processing is required to implement the protocols, and hence the processing capacity places a limit on the maximum rate at which the switch can operate. Memory is required to store packets, and hence the amount of memory available determines the rate at which packets are lost, thus placing another limit on the load at which the switch can be operated.

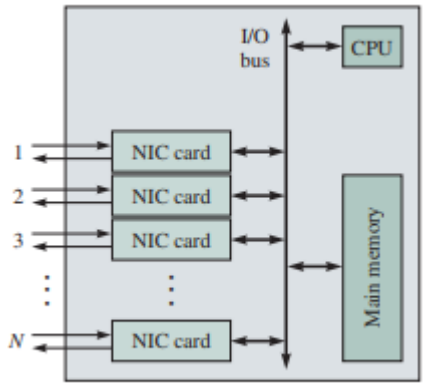


Fig12.8: Building a switch from a general purpose computer

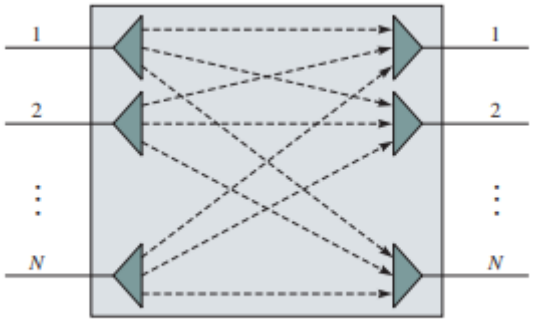


Figure12.9: shows that the flows that enter the switch in effectare demultiplexed at the input port

In this approach the memory bandwidth, which is the rate at which information can be read in and out of RAM, also places a limit on the aggregate rate of the switch. Finally,the I/O bus bandwidth places a limit on the total rate at which information canbe transferred between ports. Different switch architectures configure these basicresources so that target aggregate switch capacities are met in a cost-effective manner.

Each input and output port in a switch/router typically contains multiplexedstreams of packets. Figure shows that the flows that enter the switch in effectare demultiplexed at the input port. The switch or router then directs the packets to output ports. Each output port can be viewed as a multiplexer that precedesthe outgoing transmission line. Thus we see that switches and routers play a key role in controlling where the packet flows are placed in a network. By controlling packet flows, the network bandwidth can be used efficiently and the performacne can be optimized. We return to this discussion when we discuss Quality-of- Service mechanisms later in the chapter.

Computer networks that provide connection-oriented services are called Virtual Circuits while those providing connection-less services are called Datagram networks. For prior knowledge, the Internet which we use is actually based on a Datagram network (connection-less) at the network level as all packets from a source to a destination do not follow the same path. Let us see what are the highlighting differences between these two hot debated topics here:

Virtual Circuits:

1. It is connection-oriented, meaning that there is a reservation of resources like buffers, CPU, bandwidth, etc. for the time in which the newly setup VC is going to be used by a data transfer session.
2. The first sent packet reserves resources at each server along the path. Subsequent packets will follow the same path as the first sent packet for the connection time.
3. Since all the packets are going to follow the same path, a global header is required. Only the first packet of the connection requires a global header, the remaining packets generally don't require global headers.
4. Since all packets follow a specific path, packets are received in order at the destination.
5. Virtual Circuit Switching ensures that all packets successfully reach the Destination. No packet will be discarded due to the unavailability of resources.
6. From the above points, it can be concluded that Virtual Circuits are a highly reliable method of data transfer.
7. The issue with virtual circuits is that each time a new connection is set up, resources and extra information have to be reserved at every router along the path, which becomes problematic if many clients are trying to reserve a router's resources simultaneously.
8. It is used by the ATM (Asynchronous Transfer Mode) Network, specifically for Telephone calls.

Datagram:

1. It is a connection-less service. There is no need for reservation of resources as there is no dedicated path for a connection session.

2. All packets are free to use any available path. As a result, intermediate routers calculate routes on the go due to dynamically changing routing tables on routers.
3. Since every packet is free to choose any path, all packets must be associated with a header with proper information about the source and the upper layer data.
4. The connection-less property makes data packets reach the destination in any order, which means that they can potentially be received out of order at the receiver's end.
5. Datagram networks are not as reliable as Virtual Circuits.
6. The major drawback of Datagram Packet switching is that a packet can only be forwarded if resources such as the buffer, CPU, and bandwidth are available. Otherwise, the packet will be discarded.
7. But it is always easy and cost-efficient to implement datagram networks as there is no extra headache of reserving resources and making a dedicated each time an application has to communicate.
8. It is generally used by the IP network, which is used for Data services like the Internet.

12.4 CONNECTIONLESS PACKET SWITCHING

Packet switching has its origin in message switching, where a message is relayedfrom one station to another until the message arrives at its destination. At the source each message has a header attached to it to provide source and destina- tion addresses. CRC checkbits are attached to detect errors. As shown in Figure 12.10, the message is transmitted in a store-and-forward fashion. The message is transmitted in its entirety from one switch to the next switch. Each switch per-forms an error check, and if no errors are found, the switch examines the header to determine the next hop in the path to the destination. If errors are detected, aretransmission may be requested. After the next hop is determined, the message waits for transmission over the corresponding transmission link. Because the transmission links are shared, the message may have to wait until previously queued messages are transmitted. Message switching does not involve a call setup. Message switching can achieve a high utilization of the transmission line. This increased utilization is achieved at the expense of queueing delays. Loss of messages may occur when a switch has insufficient buffering to storethe arriving message. End-to-end mechanisms are required to recover from these losses.

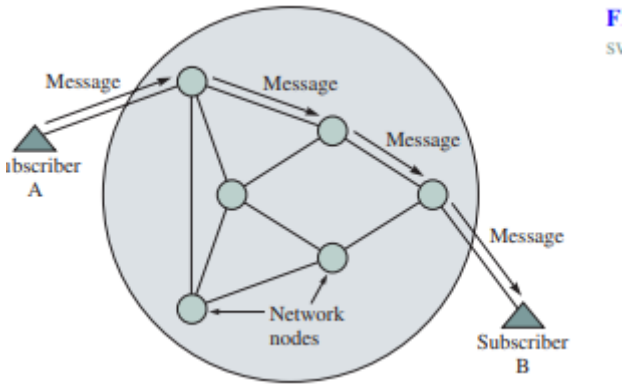


Fig12.10: Message Switching

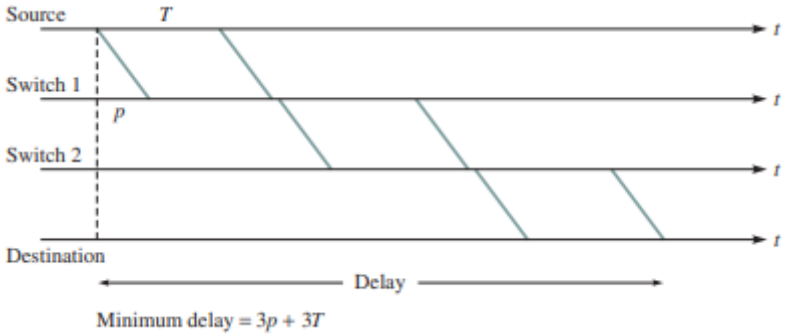


Fig12.11: Delays in message switching

Figure 12.11 shows the total delay that is incurred when a message is trans- mitted over a path that involves two intermediate switches. The message must first traverse the link that connects the source to the first switch. We assume that this link has a propagation delay of p seconds.[3] We also assume that the messagehas a transmission time of T seconds. The message must next traverse the link connecting the two switches, and then it must traverse the link connectingthe second switch and the destination. For simplicity we assume that the propagation delay and the bit rate of the transmission lines are the same. It then follows that the minimum end-to-end messge delay is 3p 3T. Note that this delay does not take into account any queueing delays that may be incurred in thevarious links waiting for prior messages to be transmitted. It also does not takeinto account the times required to perform the error checks or any associatedretransmissions.

Example—Long Messages rersus Packets

Suppose that we wish to transmit a large message (L 10^6 bits) over two hops.Suppose that the transmission line in each hop has an error rate of p 10^{-6} and that each hop does error checking and retransmission. How many bits need to betransmitted using message switching?

If we transmit the message in its entirety, the probability that the messagearrives correctly after the first hop is

$$P_c = (1 - p)^L = (1 - 10^{-6})^{1000000} \approx e^{-Lp} = e^{-1} \approx 1/3$$

Therefore, on the average it will take three tries to get the message over thefirst hop. Similarly, the second hop will require another three full message trans- missions on the average. Thus б Mbits will need to be transmitted to get the 1Mbit message across.

Now suppose that the message is broken up into ten 10^5-bit packets. The probability that a packet arrives correctly after the first hop is

$$P^r_c = (1 - 10^{-б})^{100000} \approx e^{-1/10} \approx 0.90$$

Thus each packet needs to be transmitted 1/0.90 = 1.1 times on the average.The message gets transmitted over each hop by transmitting an average of

1.1 Mbit. The total number of bits transmitted over the two hops is then

2.2 Mbits.

The preceding example reiterates our observation on ARQ protocols that the probability of error in a transmitted block increases with the length of the block.Thus very long messages are not desirable if the transmission lines are noisy because they lead to a larger rate of message retransmissions. This situation is one reason that it is desirable to place a limit on the maximum size of the blocks that can be transmitted by the network. Thus long messages should be broken into smaller blocks of information, or packets.

Message switching is also not suitable for interactive applications because it allows the transmission of very long messages that can impose very long waiting delays on other messages. By placing a maximum length on the size of the blocksthat are transmitted, packet switching limits the maximum delay that can be imposed by a single packet on other packets. Thus packet switching is more suitable than message switching for interactive applications.

In the datagram, or connectionless packet-switching approach, each packet isrouted independently through the network. Each packet has an attached header that provides all of the information required to route the packet to its destination. When a packet arrives at a packet switch, the destination address (and possibly other fields) in the header are examined to determine the next hop inthe path to the destination. The packet is then placed in a queue to wait until the given transmission line becomes available. By sharing the transmission line among multiple packets, packet switching can achieve high utilization at the expense of packet queueing delays. We note that routers in the Internet are packet switches that operate in datagram mode.

Because each packet is routed independently, packets from the same sourceto the same destination may traverse different paths through the network as shown in Figure 7.15. For example, the routes may change in response to a network fault. Thus packets may arrive out of order, and resequencing may berequired at the destination.

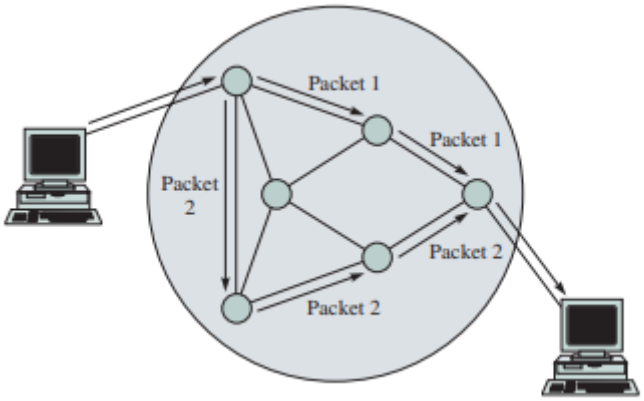


Fig12.12: Datagram packet switching

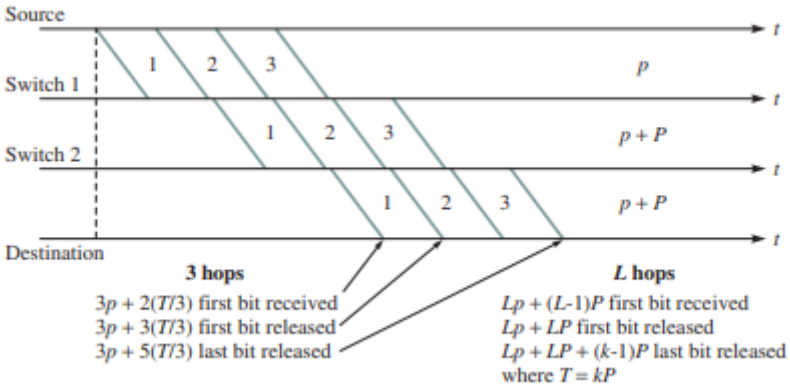


Fig12.13: Delays in packet switching

Figure 12.13 shows the delay that is incurred by transmitting a message that isbroken into three separate packets. Here we assume that the three packets follow the same path and are transmitted in succession. We neglect the overhead due toheaders and suppose that each packet requires P T/3 seconds to transmit. Thethree packets are transmitted successively from the source to the first packet switch.

The first packet in Figure 12.13 arrives at the first switch after p P seconds. Assuming that the packet arrives correctly, it can begin transmission over thenext hop after a brief processing time. The first packet is received at the second packet switch at time 2p 2P. Again we assume that the packet begins transmis-sion over the final hop after a brief processing time. The first packet then arrives at the final link at time 3p 3P. As the first packet traverses the network, the subsequent packets follow immediately, as shown in the figure. In the absence oftransmission errors, the final packet will arrive at the destination at time3p 3P 2P 3p 5P 3p T 2P, which is less than the delay incurredin the message switching example in Figure 7.14. In general, if the path followed by a sequence of packets consists of L hops with identical propagation delaysand transmission speeds, then the total delay incurred by a message that consistsof k packets is given by

$$Lp + LP + (k — 1)P$$

In contrast, the delay incurred using message switching is

$$Lp + LT = Lp + L(kP)$$

Thus message switching involves an additional delay of L 1 k 1 P. (We note that the above delays neglect the queueing and processing times at the various hops in the network.

Figure 12.14 shows a routing table that contains an entry for each possible destination for a small network. This entry specifies the next hop that is to be taken by packets with the given destination. When a packet arrives, the destina- tion address in the header is used to perform a table lookup. The result of the lookup is the number of the output port to which the packet must be forwarded.

When the size of the network becomes very large, this simple table lookup is not feasible, and the switch/router processor needs to execute a route lookup algo- rithm for each arriving packet.

In datagram packet switching, the packet switches have no knowledge of a ‘‘connection‘‘ even when a source and destination exchange a sequence of pack-ets. This feature makes datagram packet switching robust with respect to faults in the network. If a link or packet switch fails, the neighboring packet switches react by routing packets along different links and by sharing the fault informa-tion with other switches. This process results in the setting up of a new set of routing tables. Because no connections are set up, the sources and destinations need not be aware of the occurrence of a failure in the network. The processors in the switch/routers execute a distributed algorithm for sharing network state information and for synthesizing routing tables.

Destination address	Output port
0785	7
1345	12
1566	6
2458	12

Fig12.14: Routing table

The design of the routing table is a key issue in the proper operation of a packet-switching network. This design requires knowledge about the topology of the network as well as of the levels of traffic in various parts of the network. Another issue is that the size of the tables can become very large as the size of the network increases. We discuss these issues further later in the chapter.

IP Internetworks

The Internet Protocol provides for the connectionless transfer of packets across an interconnected set of networks called an internet. In general the componentnetworks may use different protocols so the objective of IP is to provide com- munications across these dissimilar networks. Each device that is attached to an IP internet has a two-part address: a network part and a host part. To transmitan IP packet, a device sends an IP packet encapsulated using its local network protocol to the nearest router. The routers are packet switches that act as gate-ways between the component networks. The router performs a route lookup algorithm on the network part of the destination address of the packet to determine whether the destination is in an immediately accessible network or, if not, to determine the next router in the path to the destination. The router thenforwards the IP packet across the given network by encapsulating the IP packet using the format and protocol of the given network. In other words, IP treats thecomponent networks as data link layers whose role is to transfer the packet tothe next router or to the destination. IP packets are routed in connectionless fashion from router to router until the destination is reached.

12.5 SUMMARY

In this unit we have discussed about packet network topology in detail. Here we also explained in detail datagrams and virtual circuits. At the end of this unit in detail discussed about connectionless packet switching.

12.6 KEYWORDS

LAN, Virtual circuit, Datagram, Internet protocol.

12.7 QUESTIONS

1. Briefly explain packet network topology.
2. Write a short note on datagram.
3. Differentiate virtual circuit and datagram.
4. With neat diagram explain connectionless packet switching
5. Discuss IP Internetworks

12.8 REFERENCES

- Introduction to Data Communications and Networking by Behrouz Forouzan.
- Computer Networks by Andrew S Tanenbaum.
- Networking Essentials – Third Edition – Jeffrey S. Beasley, Piyasat Nilkaew

UNIT-13

VIRTUAL CIRCUIT PACKET SWITCHING AND ROUTING CONCEPTS

Structure:

13.0 Objectives

13.1 Introduction

13.2 Virtual circuit packet switching

13.3 Routing concepts

13.4 Routing tables

13.5 Summary

13.6 Keywords

13.7 Questions for self study

13.8 References

13.0 OBJECTIVES

After studying this unit, you will be able to

- Explain virtual circuit packet switching
- Describe routing concepts
- Elucidate routing tables

13.1 INTRODUCTION

Packet Switched service transfers the data from source to destination. Data is transferred on a type of network in which small units of data called packets are transferred. Each packet contains a destination address within it, where it has to be received.

This type of communication between receiver & sender is connectionless. The internet is also a connectionless network. Most of the traffic over the internet uses packet switched service.

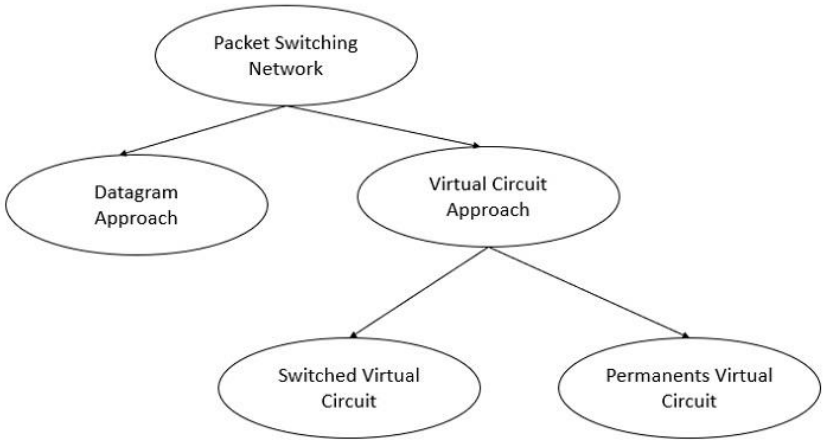
Voice call over the internet is a type of packet switching, where each end of the conversation is divided into small packets & is reassembled later into a complete message. Packets are made up of a header and a payload, header directs it to the destination & data in the payload is the actual data which is to be transferred.

13.2 VIRTUAL CIRCUIT PACKET SWITCHING

The packet switching network is of two types −

- Datagram packet switching
- Virtual circuit packet switching

The packet switching network is diagrammatically represented as follows −



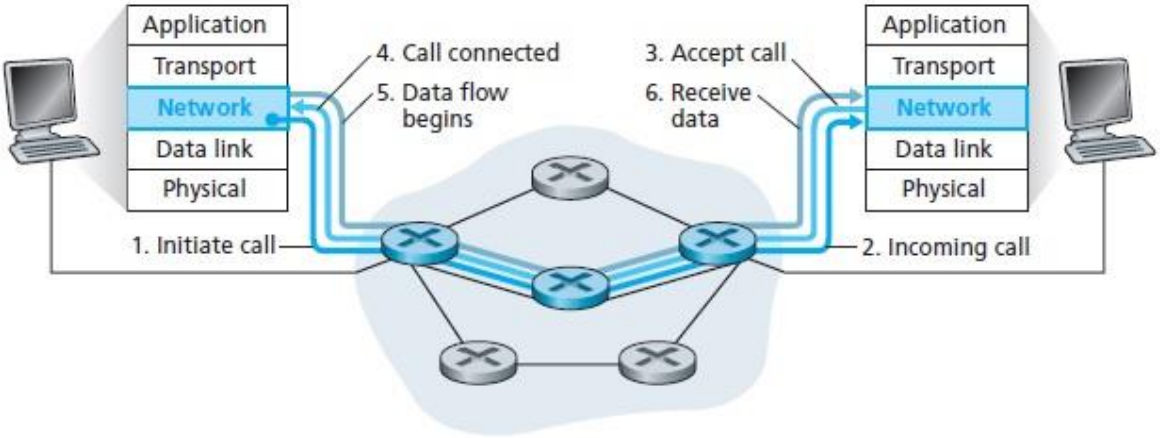
Virtual circuit switching

It is a network where a virtual connection is established between source and the destination. Through this network, packets will be transferred during any call. The path established between two points appears as a dedicated physical circuit. Therefore, it is called a virtual circuit. It is a type of packet switching.

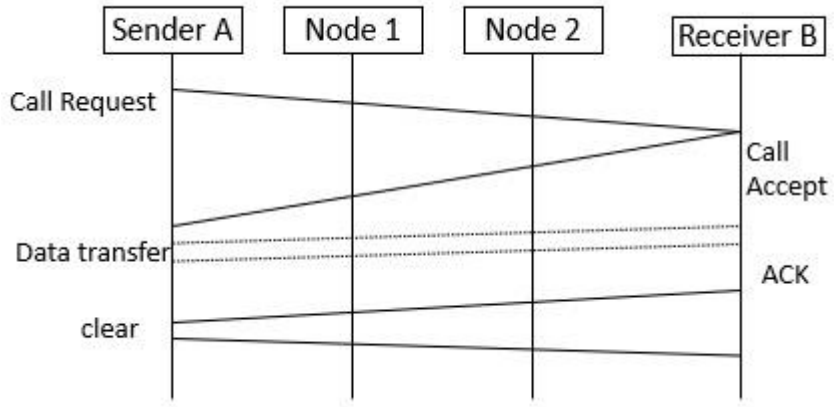
It is a connection-oriented service, where the first packet goes and reserves the resources for the subsequent packets. For examples − X.25 and frame relay.

Virtual Circuit packet switching involves the establishment of a fixed path, often called a virtual circuit or a connection. There are setup and teardown phases in addition to the data transfer phase. Resources can be allocated during the setup phase. Data are packetized and each packet carries an address in the header. However, the address in the header has local jurisdiction (it defines what should be the next switch and the channel on which the packet is being canied), not end-to-end jurisdiction. The reader may ask how the intermediate switches know where to send the packet if there is no final destination address carried by a packet. all packets follow the same path established during the connection. A virtual-circuit network is normally implemented in the data link layer.

The pictorial representation of virtual circuit connection over a telephone call is as follows −



Let's try to understand the concept with the help of data flow diagram as shown below −



Explanation:

Step 1 − Sender A establishes a call request connection to connect with the receiver.

Step 2 − Receiver B establishes a call accepting connection to connect with sender.

Step 3 − Data will be transferred whenever the router is established.

Step 4 − Node 1 and Node 2 are intermediate nodes between sender and receiver, the data will be transferred by connecting two nodes virtually.

Step 5 − after transmitting the data, an ACK will be sent by the receiver by saying a message is received.

Step 6 − A clear signal will be sent if the user wants to terminate the connection.

Advantages:

The advantages of virtual circuit are as follows −

- Packets are delivered in the same order as they all follow the same route between the source & the destination.
- The overhead is smaller as full address is not required on each packet as they all follow the same established path.
- The connection is more reliable as it is one to one connection.
- Less chances of data loss.

Disadvantages:

The disadvantages of virtual circuit are as follows −

- The switching equipment should be powerful.
- Re-establishment of the network is difficult as if there is any failure. All calls need to be re-established.

13.3 ROUTING CONCEPTS

Routing

Routing is the process of path selection in any network. A computer network is made of many machines, called *nodes*, and paths or links that connect those nodes. Communication between two nodes in an interconnected network can take place through many different paths. Routing is the process of selecting the best path using some predetermined rules.

- A Router is a process of selecting path along which the data can be transferred from source to the destination. Routing is performed by a special device known as a router.
- A Router works at the network layer in the OSI model and internet layer in TCP/IP model
- A router is a networking device that forwards the packet based on the information available in the packet header and forwarding table.
- The routing algorithms are used for routing the packets. The routing algorithm is nothing but a software responsible for deciding the optimal path through which packet can be transmitted.
- The routing protocols use the metric to determine the best path for the packet delivery. The metric is the standard of measurement such as hop count, bandwidth, delay, current load on the path, etc. used by the routing algorithm to determine the optimal path to the destination.
- The routing algorithm initializes and maintains the routing table for the process of path determination.

Working of Router:

Data moves along any network in the form of data packets. Each data packet has a header that contains information about the packet's intended destination. As a packet travels to its destination, several routers might route it multiple times. Routers perform this process millions of times each second with millions of packets.

When a data packet arrives, the router first looks up its address in a routing table. This is similar to a passenger consulting a bus timetable to find the best bus route to their destination. Then the router forwards or moves the packet onward to the next point in the network.

Types of Routing:

There are two different types of routing, which are based on how the router creates its routing tables:

Static routing

In static routing, a network administrator uses static tables to manually configure and select network routes. Static routing is helpful in situations where the network design or parameters are expected to remain constant. Static routing is a process in which we have to manually add routes to the routing table.

The static nature of this routing technique comes with expected drawbacks, such as network congestion. While administrators can configure fallback paths in case a link fails, static routing generally decreases the adaptability and flexibility of networks, resulting in limited network performance.

Advantages:

- No routing overhead for router CPU which means a cheaper router can be used to do routing.
- It adds security because an only administrator can allow routing to particular networks only.
- No bandwidth usage between routers.

Disadvantage –

- For a large network, it is a hectic task for administrators to manually add each route for the network in the routing table on each router.

- The administrator should have good knowledge of the topology. If a new administrator comes, then he has to manually add each route so he should have very good knowledge of the routes of the topology.

Dynamic routing:

In dynamic routing, routers create and update routing tables at runtime based on actual network conditions. They attempt to find the fastest path from the source to the destination by using a dynamic routing protocol, which is a set of rules that create, maintain, and update the dynamic routing table.

The biggest advantage of dynamic routing is that it adapts to changing network conditions, including traffic volume, bandwidth, and network failure.

Dynamic routing makes automatic adjustments of the routes according to the current state of the route in the routing table. Dynamic routing uses protocols to discover network destinations and the routes to reach them. **RIP** and **OSPF** are the best examples of dynamic routing protocols. Automatic adjustments will be made to reach the network destination if one route goes down.

A dynamic protocol has the following features:

1. The routers should have the same dynamic protocol running in order to exchange routes.
2. When a router finds a change in the topology then the router advertises it to all other routers.

Advantages –

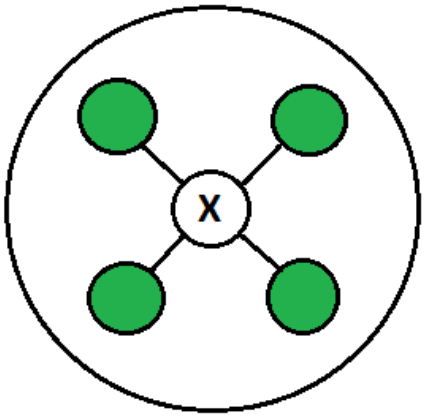
- Easy to configure.
- More effective at selecting the best route to a destination remote network and also for discovering remote network.

Disadvantage –

- Consumes more bandwidth for communicating with other neighbors.
- Less secure than static routing.

13.4 ROUTING TABLES

A Router is a networking device that forwards data packets between computer networks. This device is usually connected to two or more different networks. When a data packet comes to a router port, the router reads address information in packet to determine out which port the packet will be sent. For example, a router provides you with the internet access by connecting your LAN with the Internet.



When a packet arrives at a Router, it examines destination IP address of a received packet and makes routing decisions accordingly. Routers use *Routing Tables* to determine out which interface the packet will be sent. A routing table lists all networks for which routes are known. Each router's routing table is unique and stored in the RAM of the device.

Routing Table:

A routing table is a set of rules, often viewed in table format, which is used to determine where data packets traveling over an Internet Protocol (IP) network will be directed. All IP-enabled devices, including routers and switches, use routing tables. See below a Routing Table:

Destination	**Subnet mask**	**Interface**
128.75.43.0	255.255.255.0	Eth0
128.75.43.0	255.255.255.128	Eth1
192.12.17.5	255.255.255.255	Eth3
default		Eth2

The entry corresponding to the *default* gateway configuration is a network destination of 0.0.0.0 with a network mask (netmask) of 0.0.0.0. The Subnet Mask of default route is always 0.0.0.0 .

Entries of an IP Routing Table:

A routing table contains the information necessary to forward a packet along the best path toward its destination. Each packet contains information about its origin and destination. Routing Table provides the device with instructions for sending the packet to the next hop on its route across the network.

Each entry in the routing table consists of the following entries:

1. **Network ID:**

   The network ID or destination corresponding to the route.
2. **Subnet Mask:**

   The mask that is used to match a destination IP address to the network ID.
3. **Next Hop:**

   The IP address to which the packet is forwarded
4. **Outgoing Interface:**

   Outgoing interface the packet should go out to reach the destination network.
5. **Metric:**

   A common use of the metric is to indicate the *minimum number of hops* (routers crossed) to the network ID.

Routing table entries can be used to store the following types of routes:

- Directly Attached Network IDs
- Remote Network IDs
- Host Routes
- Default Route
- Destination

When a router receives a packet, it examines the destination IP address, and looks up into its **Routing Table** to figure out which interface packet will be sent out.

A host or a router has a routing table with an entry for each destination, or a combination of destinations, to route IP packets. The routing table can be either static or dynamic.

Static Routing Table:

A **static routing table** contains information entered manually. The administrator enters the route for each destination into the table. When a table is created, it cannot update automatically when there is a change in the Internet. The table must be manually altered by the administrator.

A static routing table can be used in a small internet that does not change very often, or in an experimental internet for troubleshooting. It is poor strategy to use a static routing table in a big internet such as the Internet

Dynamic Routing Table:

A dynamic routing table is updated periodically by using one of the dynamic routing protocols such as RIP, OSPF, or BGP. Whenever there is a change in the Internet, such as a

shutdown of a router or breaking of a link, the dynamic routing protocols update all the tables in the routers (and eventually in the host) automatically.

The routers in a big internet such as the Internet need to be updated dynamically for efficient delivery of the IP packets

13.5 SUMMARY

In this unit, we have studies about router, routing table and virtual circuit switching. Virtual circuit switching is a packet switching methodology whereby a path is established between the source and the final destination through which all the packets will be routed during a call. This path is called a virtual circuit because to the user, the connection appears to be a dedicated physical circuit.

A router is a device that connects two or more packet-switched networks or subnetworks. It serves two primary functions: managing traffic between these networks by forwarding data packets to their intended IP addresses, and allowing multiple devices to use the same Internet connection.

There are several types of routers, but most routers pass data between LANs (local area networks) and WANs (wide area networks). A LAN is a group of connected devices restricted to a specific geographic area. A LAN usually requires a single router.

A WAN, by contrast, is a large network spread out over a vast geographic area. Large organizations and companies that operate in multiple locations across the country, for instance, will need separate LANs for each location, which then connect to the other LANs to form a WAN. Because a WAN is distributed over a large area, it often necessitates multiple routers and switches.

Think of a router as an air traffic controller and data packets as aircraft headed to different airports (or networks). Just as each plane has a unique destination and follows a unique route, each packet needs to be guided to its destination as efficiently as possible. In the same way that an air traffic controller ensures that planes reach their destinations without getting lost or suffering a major disruption along the way, a router helps direct data packets to their destination IP address.

In order to direct packets effectively, a router uses an internal routing table — a list of paths to various network destinations. The router reads a packet's header to determine where it is going, then consults the routing table to figure out the most efficient path to that destination. It then forwards the packet to the next network in the path.

13.6 KEYWORDS

Virtual circuit switching, router, routing table, static routing, dynamic routing

13.7 QUESTIONS FOR SELF STUDY

1. Explain Virutal circuit switching. Write the pictorial representation of virtual circuit connection over a telephone call.
2. Define router. Explain routing concepts.
3. Explain static routing and dynamic routing.
4. Write a note on routing table.

13.8 REFERENCES

1. https://www.tutorialspoint.com/
2. https://www.javatpoint.com/
3. https://aws.amazon.com/
4. https://www.geeksforgeeks.org/
5. https://faculty.ksu.edu.sa/sites/default/files/2-chapter2-routing-algorithms.pdf
6. https://ecomputernotes.com/
7. Data communication and networking fourth edition by Behrouz A. Forouzan
8. Computer networks fifth edition by Tanenbaum ,Pearson Education India

UNIT-14

ROUTING ALGORITHMS

Structure:

14.0 Objectives

14.1 Introduction

14.2 Adoptive and Non-adoptive routing algorithms

14.3 Dijkstra's shortest path routing algorithm

14.4 congestion control algorithms

14.4.1 Leaky bucket algorithm

14.4.2 Token bucket algorithm

14.5 Summary

14.6 Keywords

14.7 Questions for self study

14.8 References

14.0 OBJECTIVES

After studying this unit, you will be able to

- Explain Dijkstra's routing algorithm
- Describe shortest path routing algorithm
- Explain Congestion control algorithm
- Elucidate Leaky bucket algorithm

14.1 INTRODUCTION

Routing algorithm

In order to transfer the packets from source to the destination, the network layer must determine the best route through which packets can be transmitted. Whether the network layer provides datagram service or virtual circuit service, the main job of the network layer is to provide the best route. The routing protocol provides this job. The routing protocol is a routing algorithm that provides the best path from the source to the destination. The best path is the path that has the "least-cost path" from source to the destination. Routing is the process of forwarding the packets from source to the destination but the best route to send the packets is determined by the routing algorithm.

14.2 ADOPTIVE AND NON-ADOPTIVE ROUTING ALGORITHMS

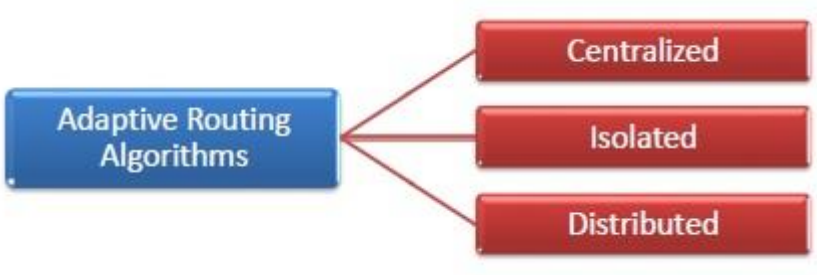
The Routing algorithm is divided into two categories:

- Adaptive Routing algorithm
- Non-adaptive Routing algorithm

Adaptive routing algorithms, also known as dynamic routing algorithms, makes routing decisions dynamically while transferring data packets from the source to the destination. These algorithms construct routing tables depending on the network conditions like network traffic and topology. They try to compute computes the best path, i.e. "least – cost path", depending upon the hop count, transit time and distance.

Types of Adaptive Routing Algorithms

The three popular types of adaptive routing algorithms are shown in the following diagram −



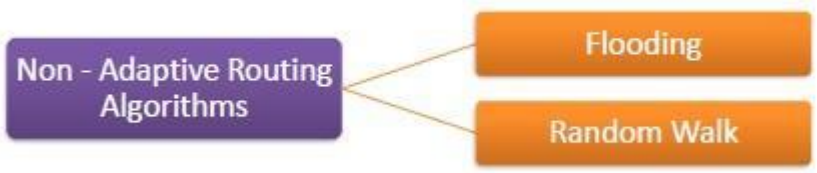
- **Centralized algorithm** − In centralized routing, one centralized node has the total network information and takes the routing decisions. It finds the least-cost path between source and destination nodes by using global knowledge about the network. So, it is also known as global routing algorithm. The advantage of this routing is that only the central node is required to store network information and so the resource requirement of the other nodes may be less. However, routing performance is too much dependent upon the central node. An example of centralized routing is link state routing algorithm.
- **Isolated algorithm** − In this algorithm, the nodes make the routing decisions based upon local information available to them instead of gathering information from other nodes. They do not have information regarding the link status. While this helps in fast decision making, the nodes may transmit data packets along congested network resulting in delay. The examples of isolated routing are hot potato routing and backward learning.

- **Distributed algorithm** − This is a decentralized algorithm where each node receives information from its neighboring nodes and takes the decision based upon the received information. The least-cost path between source and destination is computed iteratively in a distributed manner. An advantage is that each node can dynamically change routing decisions based upon the changes in the network. However, on the flip side, delays may be introduced due to time required to gather information. Example of distributed algorithm is distance vector routing algorithm.

Non-adaptive routing algorithms, also known as static routing algorithms, do not change the selected routing decisions for transferring data packets from the source to the destination. They construct a static routing table in advance to determine the path through which packets are to be sent.

The static routing table is constructed based upon the routing information stored in the routers when the network is booted up. Once the static paths are available to all the routers, they transmit the data packets along these paths. The changing network topology and traffic conditions do not affect the routing decisions.

Types of Non − adaptive Routing Algorithms



- **Flooding** − In flooding, when a data packet arrives at a router, it is sent to all the outgoing links except the one it has arrived on. Flooding may be of three types−
    - **Uncontrolled flooding** − Here, each router unconditionally transmits the incoming data packets to all its neighbours.
    - **Controlled flooding** − They use some methods to control the transmission of packets to the neighbouring nodes. The two popular algorithms for controlled flooding are Sequence Number Controlled Flooding (SNCF) and Reverse Path Forwarding (RPF).
    - **Selective flooding** − Here, the routers don't transmit the incoming packets only along those paths which are heading towards approximately in the right direction, instead of every available paths.
- **Random walks (RW)** − This is a probabilistic algorithm where a data packet is sent by a router to any one of its neighbours randomly. The transmission path thereby

formed is a random walk. RW can explore the alternative routes very efficiently. RW is very simple to implement, requires small memory footprints, does not topology information of the network and has inherent load balancing property. RW is suitable for very small devices and for dynamic networks.

14.3 DIJKSTRA'S SHORTEST PATH ROUTING ALGORITHM

The goal of shortest path routing is to find a path between two nodes that has the lowest total cost, where the total cost of a path is the sum of arc costs in that path.

For example, Dijikstra uses the nodes labelling with its distance from the source node along the better-known route.

Shortest path can be calculated only for the weighted graphs. The edges connecting two vertices can be assigned a nonnegative real number, called the weight of the edge. A graph with such weighted edges is called a weighted graph.

Let G be a weighted graph. Let u and v be two vertices in G, and let P be a path in G from u to v. The weight of the path P is the sum of the weights of all the edges on the path P, Which is also called the weight of v from u via P.

Let G be a weighted graph representing a highway structure. Suppose that the weight of an edge represents the travel time. For example, to plan monthly business trips, a salesperson wants to find the shortest path (that is, the path with the smallest weight) from her or his city to every other city in the graph. Many such problems exist in which we want to find the shortest path from a given vertex, called the source, to every other vertex in the graph. This section describes the shortest path algorithm, also called the greedy algorithm, developed by Dijkstra.

Dijkstra's Shortest Path Algorithm

Given a vertex, say vertex (that is, a source), this section describes the shortest path algorithm.

The general algorithm is:

1. Initialize the array smallestWeight so that smallestWeight[u] = weights[vertex, u].

2. Set smallestWeight[vertex] = 0.

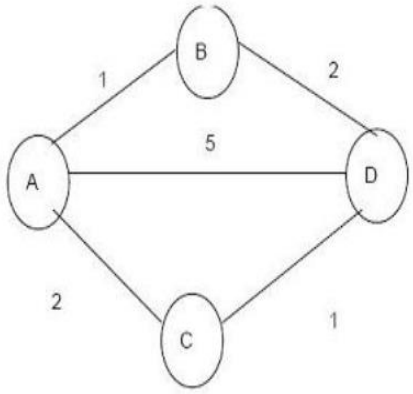
3. Find the vertex, v, that is closest to vertex for which the shortest path has not been determined.

4. Mark v as the (next) vertex for which the smallest weight is found.

5. For each vertex w in G, such that the shortest path from vertex to w has not been determined and an edge (v, w) exists, if the weight of the path to w via v is smaller than its current weight, update the weight of w to the weight of v + the weight of the edge (v, w).

Because there are n vertices, repeat Steps 3 through 5, n – 1 times.

Example : Shortest Path



SOURCE : A

Edge	Cost	Path
B	1	A-B
C	2	A-C
D	5	A-D

Direct Cost
Select A-B

Edge	Cost	Path
B	1	A-B
C	2	A-C
D	3	A-B-D

Therefore A-B-D (3) < A-D (5)
Adjusted from B
Select A-C

Edge	Cost	Path
B	1	A-B
C	2	A-C
D	3	A-B-D

Therefore A-B-D (3) < A-D(5)

14.4 CONGESTION CONTROL ALGORITHMS

What is **congestion**?

A state occurring in network layer when the message traffic is so heavy that it slows down network response time.

Effects of Congestion

- As delay increases, performance decreases.
- If delay increases, retransmission occurs, making situation worse.

Congestion causes choking of the communication medium. When too many packets are displayed in a method of the subnet, the subnet's performance degrades. Hence, a network's communication channel is called congested if packets are traversing the path and experience delays mainly over the path's propagation delay.

Congestion control algorithms

- Congestion Control is a mechanism that controls the entry of data packets into the network, enabling a better use of a shared network infrastructure and avoiding congestive collapse.
- Congestive-Avoidance Algorithms (CAA) are implemented at the TCP layer as the mechanism to avoid congestive collapse in a network.
- There are two congestion control algorithm which are as follows:

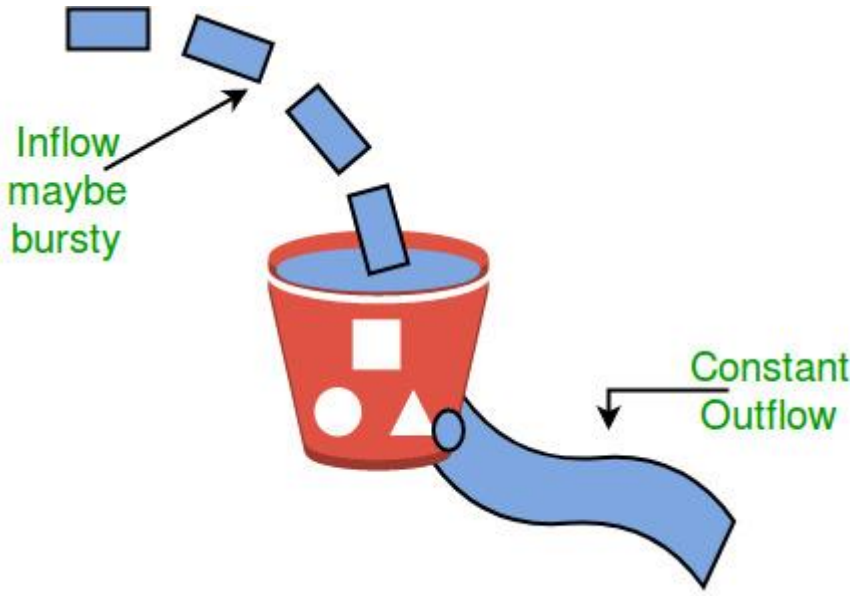
Leaky bucket and token bucket algorithm

14.4.1 Leaky Bucket Algorithm

- The leaky bucket algorithm discovers its use in the context of network traffic shaping or rate-limiting.
- A leaky bucket execution and a token bucket execution are predominantly used for traffic shaping algorithms.
- This algorithm is used to control the rate at which traffic is sent to the network and shape the burst traffic to a steady traffic stream.
- The disadvantages compared with the leaky-bucket algorithm are the inefficient use of available network resources.
- The large area of network resources such as bandwidth is not being used effectively.

Let us consider an example to understand this algorithm.

Imagine a bucket with a small hole in the bottom.No matter at what rate water enters the bucket, the outflow is at constant rate.When the bucket is full with water additional water entering spills over the sides and is lost.



Similarly, each network interface contains a leaky bucket and the following **steps** are involved in leaky bucket algorithm:

1. When host wants to send packet, packet is thrown into the bucket.
2. The bucket leaks at a constant rate, meaning the network interface transmits packets at a constant rate.
3. Bursty traffic is converted to a uniform traffic by the leaky bucket.
4. In practice the bucket is a finite queue that outputs at a finite rate.

14.4.2 Token bucket Algorithm

- The leaky bucket algorithm has a rigid output design at an average rate independent of the bursty traffic.
- In some applications, when large bursts arrive, the output is allowed to speed up. This calls for a more flexible algorithm, preferably one that never loses information. Therefore, a token bucket algorithm finds its uses in network traffic shaping or rate-limiting.
- It is a control algorithm that indicates when traffic should be sent. This order comes based on the display of tokens in the bucket.
- The bucket contains tokens. Each of the tokens defines a packet of predetermined size. Tokens in the bucket are deleted for the ability to share a packet.

- When tokens are shown, a flow to transmit traffic appears in the display of tokens.
- No token means no flow sends its packets. Hence, a flow transfers traffic up to its peak burst rate in good tokens in the bucket.

Need of token bucket Algorithm:-

The leaky bucket algorithm enforces output pattern at the average rate, no matter how bursty the traffic is. So in order to deal with the bursty traffic we need a flexible algorithm so that the data is not lost. One such algorithm is token bucket algorithm.

Steps of this algorithm is as follows:

In regular intervals tokens are thrown into the bucket.

1. The bucket has a maximum capacity.
2. If there is a ready packet, a token is removed from the bucket, and the packet is sent.
3. If there is no token in the bucket, the packet cannot be sent.

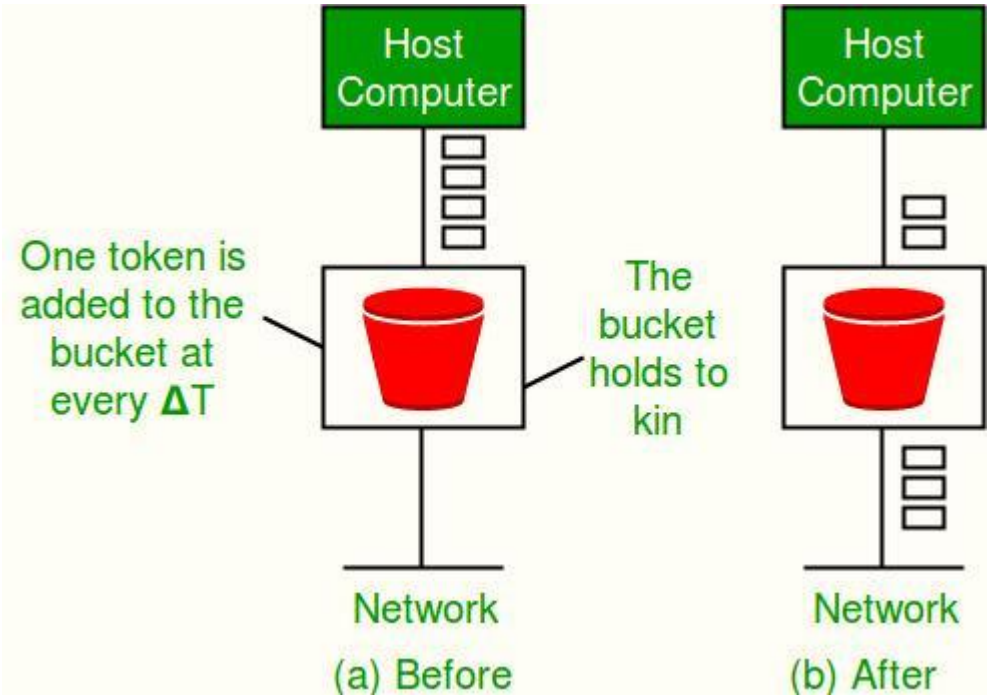
Let's understand with an example,

In figure (A) we see a bucket holding three tokens, with five packets waiting to be transmitted. For a packet to be transmitted, it must capture and destroy one token. In figure (B) We see that three of the five packets have gotten through, but the other two are stuck waiting for more tokens to be generated.

Ways in which token bucket is superior to leaky bucket: The leaky bucket algorithm controls the rate at which the packets are introduced in the network, but it is very conservative in nature. Some flexibility is introduced in the token bucket algorithm. In the token bucket, algorithm tokens are generated at each tick (up to a certain limit). For an incoming packet to be transmitted, it must capture a token and the transmission takes place at the same rate. Hence some of the busty packets are transmitted at the same rate if tokens are available and thus introduces some amount of flexibility in the system.

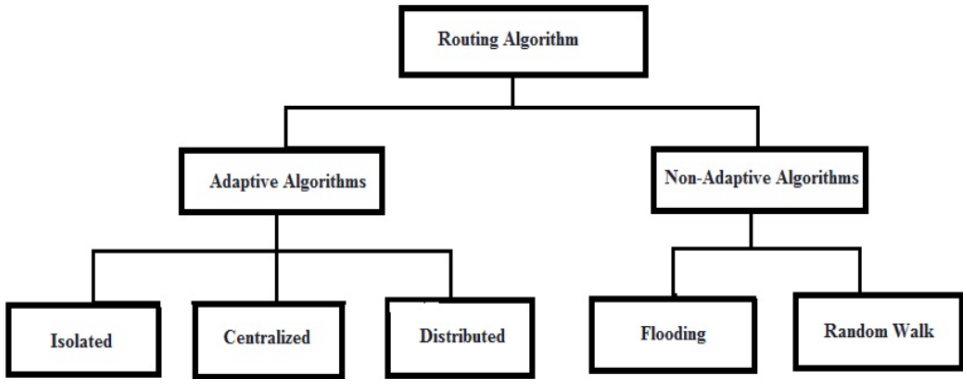
Formula: $M * s = C + \rho * s$ where S – is time taken M – Maximum output rate ρ – Token arrival rate C – Capacity of the token bucket in byte

Let's understand with an example,



14.5 SUMMARY

A routing algorithm is a procedure that lays down the route or path to transfer data packets from source to the destination. They help in directing Internet traffic efficiently. After a data packet leaves its source, it can choose among the many different paths to reach its destination.



Routing v/s Flooding:

Routing	Flooding
Rouitng Table is required	No routing table is required
May give shortest path	Always gives shortest path
Less reliable	More reliable
Traffic is less	Traffic is high
No duplicate packets	Duplicate packets are present

14.6 KEYWORDS

Leaky bucket algorithm, token bucket algorithm, Dijkstra's Shortest Path Algorithm, adoptive routing algorithms, Non-adoptive routing algorithms,flodding, random walk, isolated, distributed, centralized

14.7 QUESTIONS FOR SELF STUDY

1. Explain adoptive routing algorithms.
2. Explain Non-adoptive routing algorithms.
3. Describe Dijkstra's shortest path algorithm with example.
4. In detail, explain leaky bucket algorithm.
5. Explain token bucket algorithm.
6. Give reasons, why token bucket algorithm is superior over leaky bucket algorithm.

14.8 REFERENCES

1. https://www.tutorialspoint.com/
2. https://www.javatpoint.com/
3. https://aws.amazon.com/
4. https://www.geeksforgeeks.org/
5. https://faculty.ksu.edu.sa/sites/default/files/2-chapter2-routing-algorithms.pdf
6. https://ecomputernotes.com/

UNIT-15

DATA LINK ISSUES and ERROR DETECTING METHODS

Structure:

15.0 Objectives

15.1 Introduction

15.2 Data link layer design issues

15.3 Error detection methods

15.3.1 Simple parity check

15.3.2 Two-dimensional Parity check

15.3.2 Check sum

15.3.3 Cyclic Redundancy Check

15.4 Summary

15.5 Keywords

15.6 Questions for self study

15.7 References

15.0 OBJECTIVES

After studying this unit, you will be able to

- Understand the data link layer design issues
- Explain error detecting methods
- Distinguish between single parity error and burst error
- Explain check sum and cyclic redundancy check

15.1 INTRODUCTION

The data link layer in the OSI (Open System Interconnections) Model is in between the physical layer and the network layer. This layer converts the raw transmission facility provided by the physical layer to a reliable and error-free link.

The main functions and the design issues of this layer are

- Providing services to the network layer
- Framing
- Error Control
- Flow Control

The primary function of this layer is to provide a well-defined service interface to network layer above it.

Error is a condition when the receiver's information does not match with the sender's information. During transmission, digital signals suffer from noise that can introduce errors in the binary bits travelling from sender to receiver. That means a 0 bit may change to 1 or a 1 bit may change to 0.To avoid this, we use error-detecting codes which are additional data added to a given digital message to help us detect if any error has occurred during transmission of the message.

Basic approach used for error detection is the use of redundancy bits, where additional bits are added to facilitate detection of errors.

Some popular techniques for error detection are:

1. Simple Parity check
2. Two-dimensional Parity check
3. Checksum
4. Cyclic redundancy check

15.2 DATA LINK LAYER DESIGN ISSUES

Data-link layer is the second layer after the physical layer. The data link layer is responsible for maintaining the data link between two hosts or nodes.

Before going through the design issues in the data link layer. Some of its sub-layers and their functions are as following below.

The data link layer is divided into two sub-layers :

1. **Logical Link Control Sub-layer (LLC) –**

   Provides the logic for the data link, Thus it controls the synchronization, flow control, and error checking functions of the data link layer. Functions are –

   (i) Error Recovery.

   (ii) It performs the flow control operations.

   (iii) User addressing.

2. **Media Access Control Sub-layer (MAC) –**

   It is the second sub-layer of data-link layer. It controls the flow and multiplexing for transmission medium. Transmission of data packets is controlled by this layer. This layer is responsible for sending the data over the network interface card. Functions are –

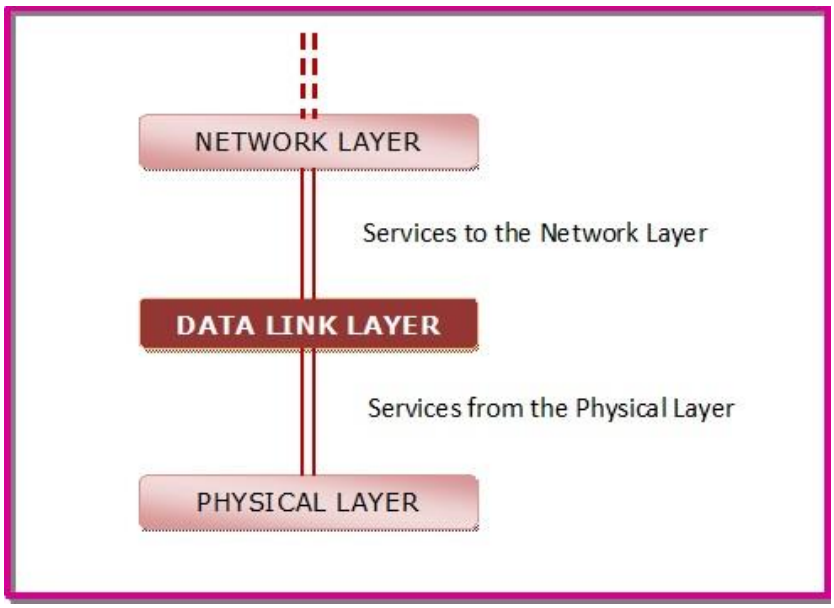
   - **(i)** To perform the control of access to media.
   - **(ii)** It performs the unique addressing to stations directly connected to LAN.
   - **(iii)** Detection of errors.

Design issues with data link layer are:

1. **Services provided to the network layer –**

   The data link layer act as a service interface to the network layer. The principle service is transferring data from network layer on sending machine to the network layer on destination machine. This transfer also takes place via DLL (Data link-layer).

In the OSI Model, each layer uses the services of the layer below it and provides services to the layer above it. The data link layer uses the services offered by the physical layer. The primary function of this layer is to provide a well-defined service interface to network layer above it.



The types of services provided can be of three types −

- Unacknowledged connectionless service
- Acknowledged connectionless service
- Acknowledged connection - oriented service

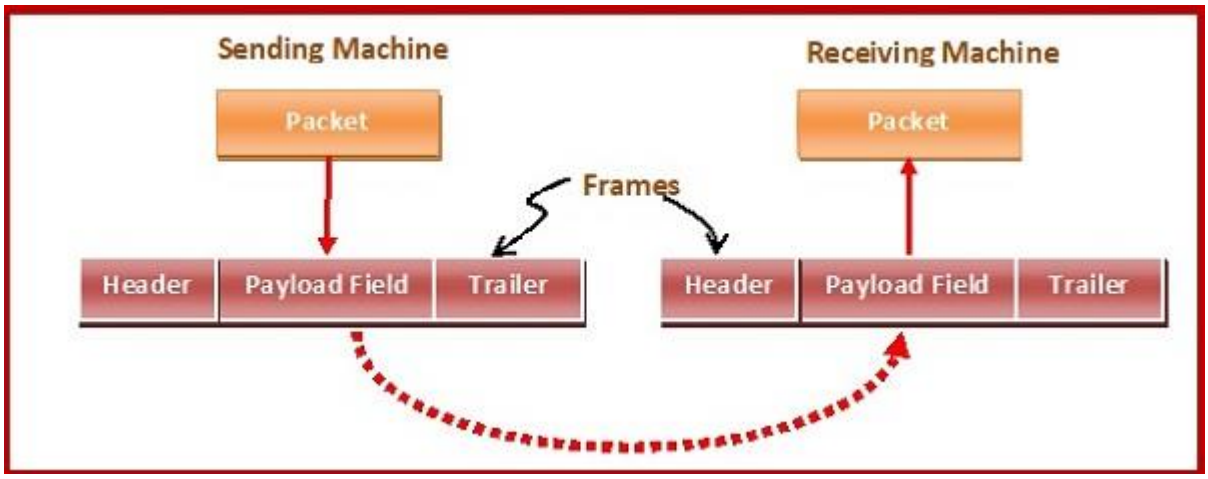
2. **Frame synchronization –**

   The source machine sends data in the form of blocks called frames to the destination machine. The starting and ending of each frame should be identified so that the frame can be recognized by the destination machine.

   The data link layer encapsulates each data packet from the network layer into frames that are then transmitted.

   A frame has three parts, namely −

- Frame Header
- Payload field that contains the data packet from network layer
- Trailer



3. **Flow control –**

Flow control is done to prevent the flow of data frame at the receiver end. The source machine must not send data frames at a rate faster than the capacity of destination machine to accept them.

The data link layer regulates flow control so that a fast sender does not drown a slow receiver. When the sender sends frames at very high speeds, a slow receiver may not be able to handle it. There will be frame losses even if the transmission is error-free. The two common approaches for flow control are −

- Feedback based flow control
- Rate based flow control

There are many reasons such as noise, cross-talk etc., which may help data to get corrupted during transmission. The upper layers work on some generalized view of network architecture and are not aware of actual hardware data processing. Hence, the upper layers expect error-free transmission between the systems. Most of the applications would not function expectedly if they receive erroneous data. Applications such as voice and video may not be that affected and with some errors they may still function well.

4. **Error control –**

Error control is done to prevent duplication of frames. The errors introduced during transmission from source to destination machines must be detected and corrected at the destination machine.

The data link layer ensures error free link for data transmission. The issues it caters to with respect to error control are −

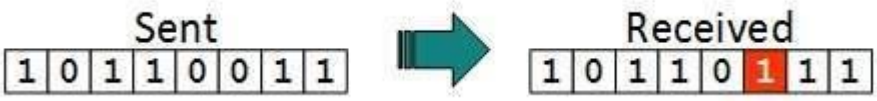
- Dealing with transmission errors
- Sending acknowledgement frames in reliable connections
- Retransmitting lost frames
- Identifying duplicate frames and deleting them
- Controlling access to shared channels in case of broadcasting

Data-link layer uses some error control mechanism to ensure that frames (data bit streams) are transmitted with certain level of accuracy. But to understand how errors is controlled, it is essential to know what types of errors may occur.

Types of Errors

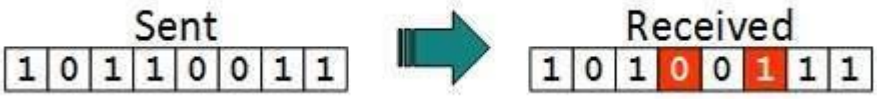
There may be three types of errors:

- **Single bit error**



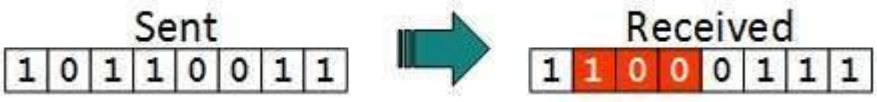
In a frame, there is only one bit, anywhere though, which is corrupt.

- **Multiple bits error**



Frame is received with more than one bit in corrupted state.

- **Burst error**



Frame contains more than1 consecutive bits corrupted.

Error control mechanism may involve two possible ways:

- Error detection
- Error correction

15.3 ERROR DETECTION METHODS

Error Detection

Errors in the received frames are detected by means of Parity Check and Cyclic Redundancy Check (CRC). In both cases, few extra bits are sent along with actual data to confirm that bits received at other end are same as they were sent. If the counter-check at receiver' end fails, the bits are considered corrupted.

Whenever a message is transmitted, it may get scrambled by noise or data may get corrupted. To avoid this, we use error-detecting codes which are additional data added to a given digital message to help us detect if any error has occurred during transmission of the message.

Basic approach used for error detection is the use of redundancy bits, where additional bits are added to facilitate detection of errors.

Some popular techniques for error detection are:

1. Simple Parity check
2. Two-dimensional Parity check
3. Checksum
4. Cyclic redundancy check

These methods are implemented either at Data link layer or Transport Layer of OSI Model.

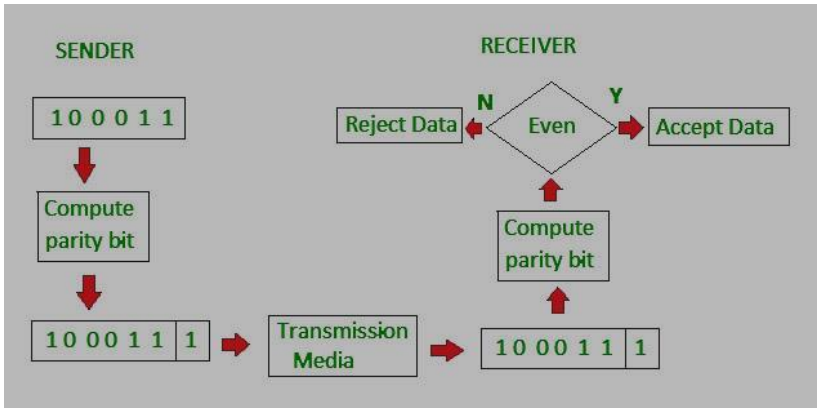
Error Detecting Methods

15.3.1. Simple Parity check

Blocks of data from the source are subjected to a check bit or parity bit generator form, where a parity of :

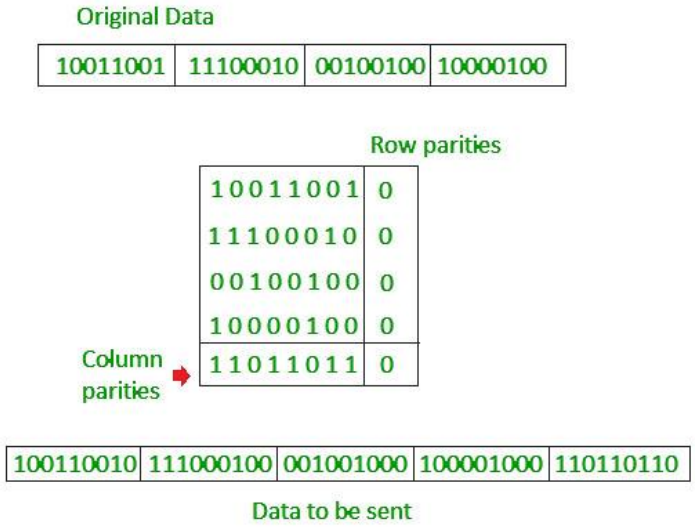
- 1 is added to the block if it contains odd number of 1's, and
- 0 is added if it contains even number of 1's

This scheme makes the total number of 1's even, that is why it is called even parity Checking.



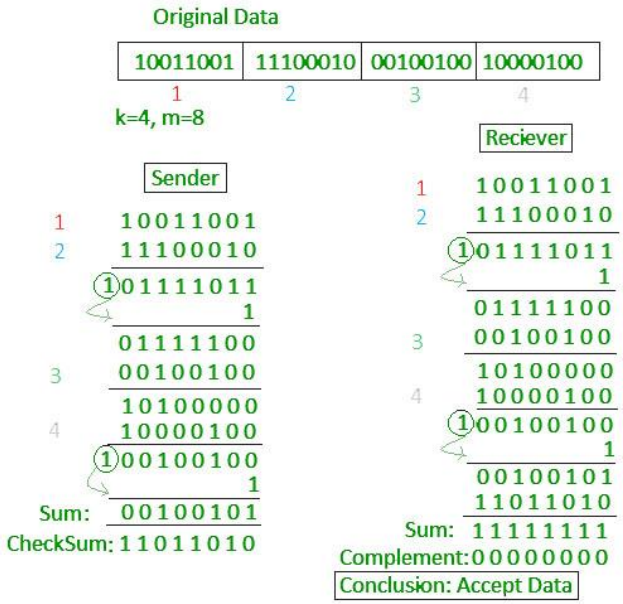
15.3.2 Two-dimensional Parity check

Parity check bits are calculated for each row, which is equivalent to a simple parity check bit. Parity check bits are also calculated for all columns, then both are sent along with the data. At the receiving end these are compared with the parity bits calculated on the received data.



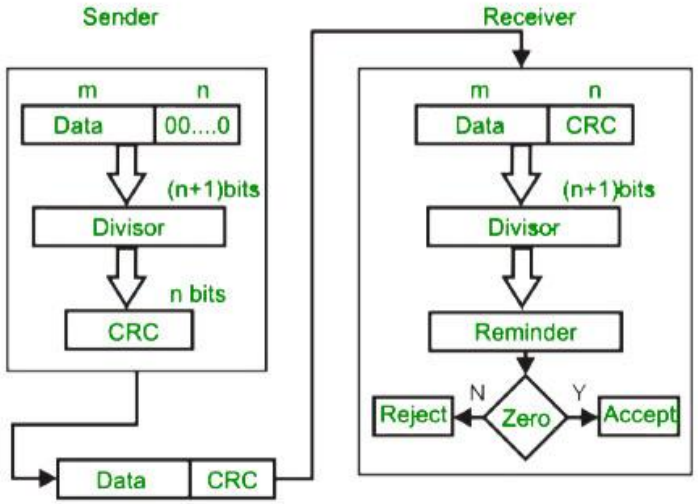
15.3.3. Checksum

- In checksum error detection scheme, the data is divided into k segments each of m bits.
- In the sender's end the segments are added using 1's complement arithmetic to get the sum. The sum is complemented to get the checksum.
- The checksum segment is sent along with the data segments.
- At the receiver's end, all received segments are added using 1's complement arithmetic to get the sum. The sum is complemented.
- If the result is zero, the received data is accepted; otherwise discarded.

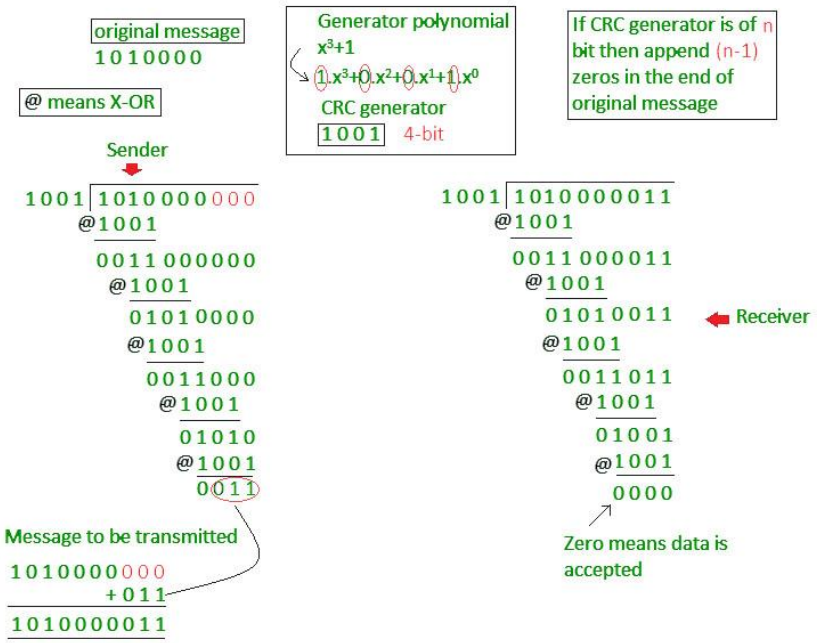


15.3.4. Cyclic redundancy check (CRC)

- Unlike checksum scheme, which is based on addition, CRC is based on binary division.
- In CRC, a sequence of redundant bits, called cyclic redundancy check bits, are appended to the end of data unit so that the resulting data unit becomes exactly divisible by a second, predetermined binary number.
- At the destination, the incoming data unit is divided by the same number. If at this step there is no remainder, the data unit is assumed to be correct and is therefore accepted.
- A remainder indicates that the data unit has been damaged in transit and therefore must be rejected.



Example :



15.4 SUMMARY

Error

A condition when the receiver's information does not match with the sender's information. During transmission, digital signals suffer from noise that can introduce errors in the binary bits travelling from sender to receiver. That means a 0 bit may change to 1 or a 1 bit may change to 0.

Parity Check

One extra bit is sent along with the original bits to make number of 1s either even in case of even parity, or odd in case of odd parity.

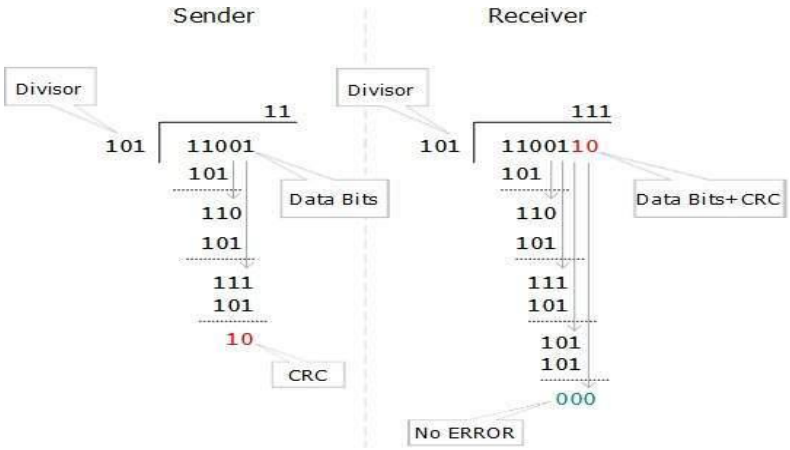
The sender while creating a frame counts the number of 1s in it. For example, if even parity is used and number of 1s is even then one bit with value 0 is added. This way number of 1s remains even. If the number of 1s is odd, to make it even a bit with value 1 is added.

The receiver simply counts the number of 1s in a frame. If the count of 1s is even and even parity is used, the frame is considered to be not-corrupted and is accepted. If the count of 1s is odd and odd parity is used, the frame is still not corrupted.

If a single bit flips in transit, the receiver can detect it by counting the number of 1s. But when more than one bit is erroneous, then it is very hard for the receiver to detect the error.

Cyclic Redundancy Check (CRC)

CRC is a different approach to detect if the received frame contains valid data. This technique involves binary division of the data bits being sent. The divisor is generated using polynomials. The sender performs a division operation on the bits being sent and calculates the remainder. Before sending the actual bits, the sender adds the remainder at the end of the actual bits. Actual data bits plus the remainder is called a code word. The sender transmits data bits as codewords.



At the other end, the receiver performs division operation on codewords using the same CRC divisor. If the remainder contains all zeros the data bits are accepted, otherwise it is considered as there some data corruption occurred in transit.

15.5 KEYWORDS

Error detecting methods, error correction methods, CRC, single parity error, frame control, error control, flow control, check sum, simple parity check, multi-dimensional parity check

15.6 QUESTIONS FOR SELF STUDY

1. Explain design issues with data link layer**.**
2. List and explain the sublayers of data link layer.
3. Explain error detecting methods with example.
4. Describe flow control and error control.

15.7 REFERENCES

1. https://www.tutorialspoint.com/

2. https://www.javatpoint.com/
3. https://aws.amazon.com/
4. https://www.geeksforgeeks.org/
5. https://faculty.ksu.edu.sa/sites/default/files/2-chapter2-routing-algorithms.pdf
6. https://ecomputernotes.com/

UNIT-16

ERROR CORRECTION METHODS

Structure:

16.0 Objectives

16.1 Introduction

16.2 error correction methods

16.2.1 Hamming Code

16.2.2 Single bit error correction code

16.3 Summary

16.4 Keywords

16.5 Questions for self study

16.6 References

16.0 OBJECTIVES

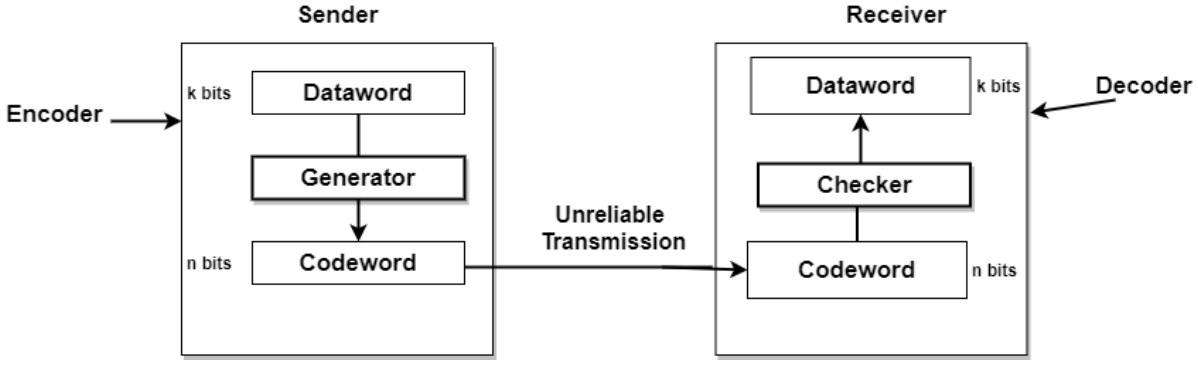
After studying this unit, you will be able to

- Explain error correction methods
- Describe hamming code method
- Explain single bit error correction method

16.1 INTRODUCTION

Error Correction

In Error Detection, the receiver only needs to know that the received code word is invalid; But in Error Correction the receiver needs to guess the Original code word that is sent. In this way, error Correction is much more difficult than Error Detection. The need for redundant bits is more during error correction rather than for error detection.



Structure of encoder and decoder in the error correction

In order to detect or correct the errors, there is a need to send some extra bits along with the data. These extra bits are commonly known as Redundant bits.

As we had learned in the previous tutorial that original data is divided into segments of k bits; it is referred to as **dataword**. When we add **r** redundant bits to each block in order to make the length;**n=k+r** then it is referred to as **Codeword**.

There are two ways to handle the error correction:

1. Whenever an error discovered, the receiver can have the sender in order to retransmit the entire data unit. This technique is known as the **Backward Error correction technique**. This technique is simple and inexpensive in the case of wired transmission like fiber optics; there is no expense in retransmitting the data. In the case of wireless transmission, retransmission costs too much thus forward error correction technique is used then.
2. The receiver can use an error-correcting code that automatically contains certain errors. This technique is known as the **Forward Error Correction technique**.

In order to correct the errors, one has to know the exact position of the error. For example, In case if we want to calculate a single-bit error, the error correction code then mainly determines which one of seven bits is in the error.

In order to achieve this, we have to add some additional redundant bits.

Suppose r (as the redundant bits) and d indicates the total number of data bits. In order to calculate the redundant bits(r), the given formula is used;

$2r= d+r+1$

Error correction is mainly done with the help of the Hamming code.

16.2 ERROR CORRECTION METHODS

Detection versus Correction

The correction of errors is more difficult than the detection. In error detection, we are looking only to see if any error has occurred. The answer is a simple yes or no. We are not even interested in the number of errors. A single-bit error is the same for us as a burst error.

In error correction, we need to know the exact number of bits that are corrupted and more importantly, their location in the message. The number of the errors and the size of the message are important factors. If we need to correct one single error in an 8-bit data unit, we need to consider eight possible error locations; if we need to correct two errors in a data unit of the same size, we need to consider 28 possibilities. You can imagine the receiver's difficulty in finding 10 errors in a data unit of 1000 bits.

In the digital world, error correction can be done in two ways:

- **Backward Error Correction** When the receiver detects an error in the data received; it requests back the sender to retransmit the data unit.

- **Forward Error Correction** When the receiver detects some error in the data received, it executes error-correcting code, which helps it to auto-recover and to correct some kinds of errors.

The first one, Backward Error Correction, is simple and can only be efficiently used where retransmitting is not expensive. For example, fiber optics. But in case of wireless transmission retransmitting may cost too much. In the latter case, Forward Error Correction is used.

To correct the error in data frame, the receiver must know exactly which bit in the frame is corrupted. To locate the bit in error, redundant bits are used as parity bits for error detection. For example, we take ASCII words (7 bits data), then there could be 8 kind of information we need: first seven bits to tell us which bit is error and one more bit to tell that there is no error.

16.2.1 Hamming Code

It is a technique developed by R.W. hamming. This can be applied to data units of any length. This code mainly uses the relationship between data and redundancy bits.

The hamming code technique, which is an **error-detection and error-correction technique**, was proposed by **R.W. Hamming**. Whenever a data packet is transmitted over a network, there are possibilities that the data bits may get lost or damaged during transmission.

Let's understand the Hamming code concept with an example:

Let's say you have received a **7-bit Hamming code** which is **1011011.**

First, let us talk about the redundant bits.

The **redundant bits** are some extra binary bits that are not part of the original data, but they are generated & added to the original data bit. All this is done to ensure that the data bits don't get damaged and if they do, we can recover them.

Now the question arises, how do we determine the number of redundant bits to be added?

We use the formula, $2^r >= m+r+1$; where **r = redundant bit** & **m = data bit**.

From the formula we can make out that there are **4 data bits** and **3 redundancy bits**, referring to the received **7-bit hamming code**.

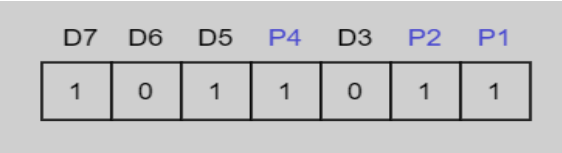
Parity Bit:

It is a bit appended to the data bits which ensures that the total number of 1's are even (even parity) or odd (odd parity).

While checking the parity, if the total number of 1's are odd then write the value of parity bit **P1**(or **P2** etc.) as **1** (which means the error is there) and if it is even then the value of parity bit is **0** (which means no error).

Hamming Code in Error Detection

As we go through the example, the first step is to identify the bit position of the data & all the bit positions which are powers of 2 are marked as parity bits (e.g. 1, 2, 4, 8, etc.). The following image will help in visualizing the received hamming code of 7 bits.

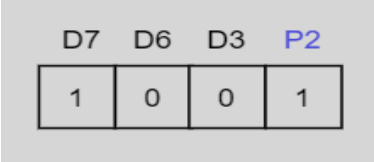


First, we need to detect whether there are any errors in this received hamming code.

Step 1: For checking parity bit **P1**, use **check one and skip one** method, which means, starting from P1 and then skip P2, take D3 then skip P4 then take D5, and then skip D6 and take D7, this way we will have the following bits,

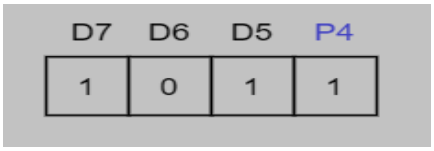
As we can observe the total number of bits is odd so we will write the value of parity bit as **P1 = 1**. This means the **error is there**.

Step 2: Check for P2 but while checking for P2, we will use the **check two and skip two** methods, which will give us the following data bits. But remember since we are checking for P2, so we have to start our count from P2 (P1 should not be considered).



As we can observe that the number of 1's are even, then we will write the value of **P2 = 0**. This means **there is no error**.

Step 3: Check for P4 but while checking for P4, we will use the **check four and skip four** methods, which will give us the following data bits. But remember since we are checking for P4, so we have started our count from P4(P1 & P2 should not be considered).



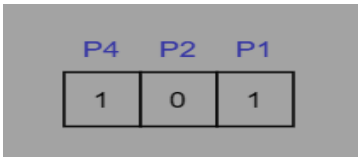
As we can observe that the number of 1's is odd, then we will write the value of **P4 = 1**. This means the error is there.

So, from the above parity analysis, P1 & P4 are not equal to 0, so we can clearly say that the received hamming code has errors.

Hamming Code: Error Correction

Since we found that the received code has an error, so now we must correct them. To correct the errors, use the following steps:

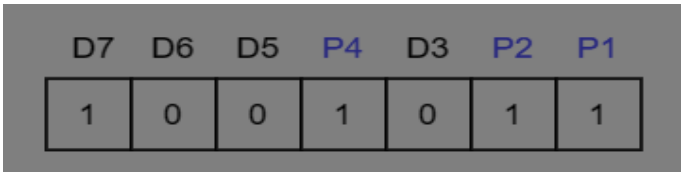
Now the error word E will be:



Now we have to determine the decimal value of this error word **101** which is **5 (22 *1 + 21 * 0 + 20 *1 = 5)**.

We get **E = 5**, which states that the error is in the fifth data bit. To correct it, just invert the fifth data bit.

So the correct data will be:



General Algorithm of Hamming code: Hamming Code is simply the use of extra parity bits to allow the identification of an error.

1. Write the bit positions starting from 1 in binary form (1, 10, 11, 100, etc).
2. All the bit positions that are a power of 2 are marked as parity bits (1, 2, 4, 8, etc).
3. All the other bit positions are marked as data bits.
4. Each data bit is included in a unique set of parity bits, as determined its bit position in binary form. **a.** Parity bit 1 covers all the bits positions whose binary representation includes a 1 in the least significant position (1, 3, 5, 7, 9, 11, etc). **b.** Parity bit 2 covers all the bits positions whose binary representation includes a 1 in the second position from the least significant bit (2, 3, 6, 7, 10, 11, etc). **c.** Parity bit 4 covers all the bits positions whose binary representation includes a 1 in the third position from the least significant bit (4–7, 12–15, 20–23, etc). **d.** Parity bit 8 covers all the bits positions whose binary representation includes a 1 in the fourth position from the least significant

bit bits (8–15, 24–31, 40–47, etc). **e.** In general, each parity bit covers all bits where the bitwise AND of the parity position and the bit position is non-zero.

5. Since we check for even parity set a parity bit to 1 if the total number of ones in the positions it checks is odd.
6. Set a parity bit to 0 if the total number of ones in the positions it checks is even.

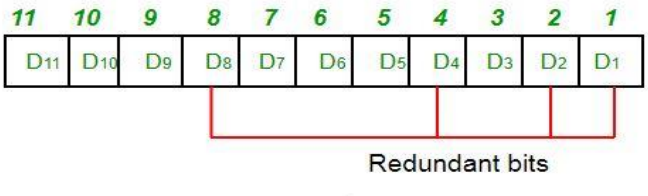
Position	R8	R4	R2	R1
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1

R1 -> 1,3,5,7,9,11
R2 -> 2,3,6,7,10,11
R3 -> 4,5,6,7
R4 -> 8,9,10,11

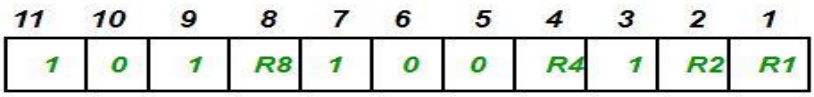
Determining the position of redundant bits – These redundancy bits are placed at positions that correspond to the power of 2.

As in the above example:

- The number of data bits = 7
- The number of redundant bits = 4
- The total number of bits = 11
- The redundant bits are placed at positions corresponding to power of 2- 1, 2, 4, and 8

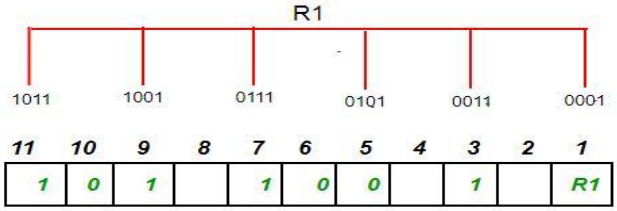


- Suppose the data to be transmitted is 1011001, the bits will be placed as follows:

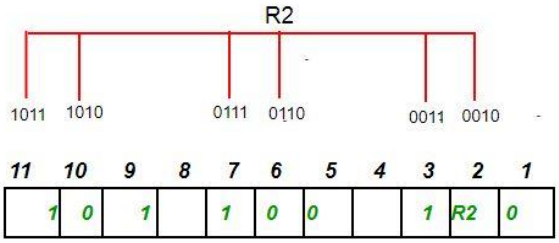


Determining the Parity bits:

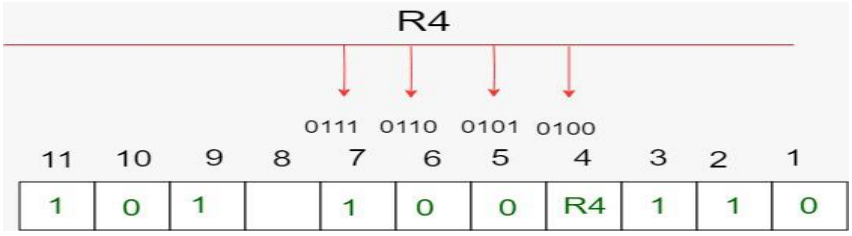
- R1 bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the least significant position. R1: bits 1, 3, 5, 7, 9, 11



- To find the redundant bit R1, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R1 is an even number the value of R1 (parity bit's value) = 0
- R2 bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the second position from the least significant bit. R2: bits 2,3,6,7,10,11

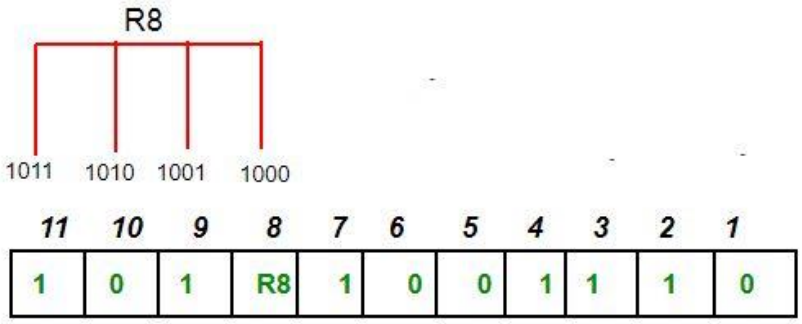


- To find the redundant bit R2, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R2 is odd the value of R2(parity bit's value)=1
- R4 bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the third position from the least significant bit. R4: bits 4, 5, 6, 7



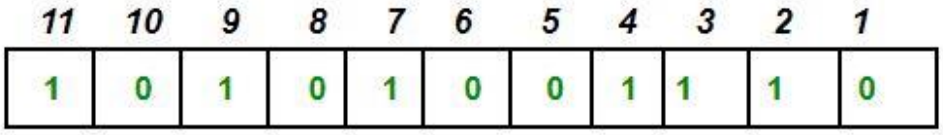
To find the redundant bit R4, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R4 is odd the value of R4(parity bit's value) = 1

1. R8 bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the fourth position from the least significant bit. R8: bit

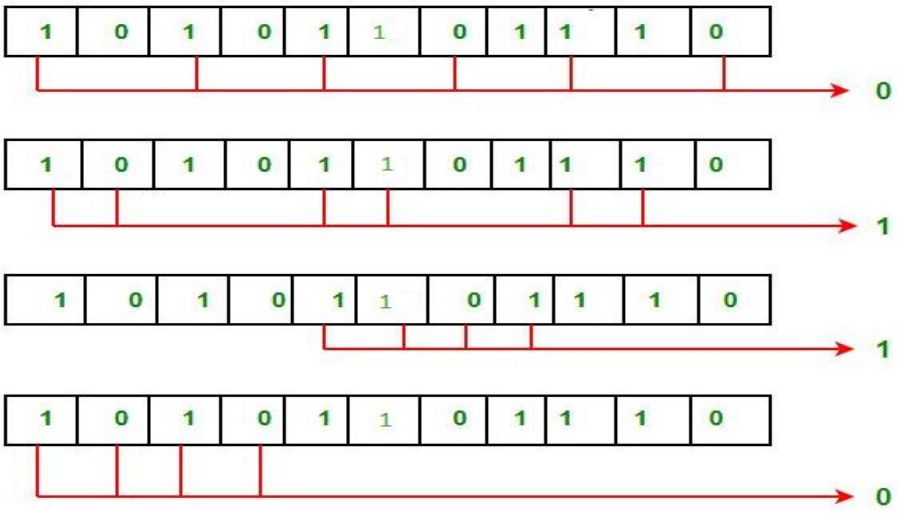


8,9,10,11

- To find the redundant bit R8, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R8 is an even number the value of R8(parity bit's value)=0. Thus, the data transferred is:



Error detection and correction: Suppose in the above example the 6th bit is changed from 0 to 1 during data transmission, then it gives new parity values in the binary number:



The bits give the binary number 0110 whose decimal representation is 6. Thus, bit 6 contains an error. To correct the error the 6th bit is changed from 1 to 0.

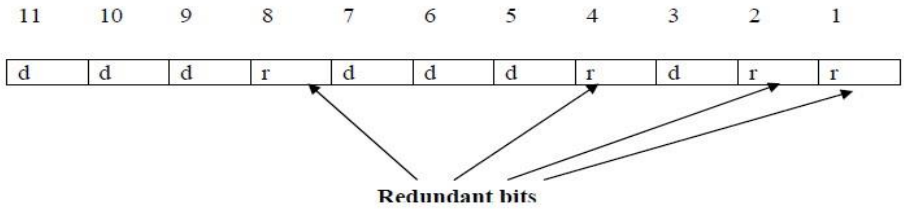
16.2.2 Single bit error correction code:

To calculate the numbers of redundant bits (r) required to correct d data bits, let us find out the relationship between the two. So we have (d+r) as the total number of bits, which are to be transmitted; then r must be able to indicate at least d+r+1 different values. Of these, one value means no error, and remaining d+r values indicate error location of error in each of d+r locations. So, d+r+1 states must be distinguishable by r bits, and r bits can indicates 2r states. Hence, 2r must be greater than d+r+1.

$2^r >= d + r + 1$

The value of r must be determined by putting in the value of d in the relation. For example, if d is 7, then the smallest value of r that satisfies the above relation is 4. So the total bits, which are to be transmitted is 11 bits (d + r = 7 + 4 = 11).

Now let us examine how we can manipulate these bits to discover which bit is in error. A technique developed by R.W.Hamming provides a practical solution. The solution or coding scheme he developed is commonly known as Hamming Code. Hamming code can be applied to data units of any length and uses the relationship between the data bits and redundant bits as discussed.



Positions of redundancy bits in hamming code

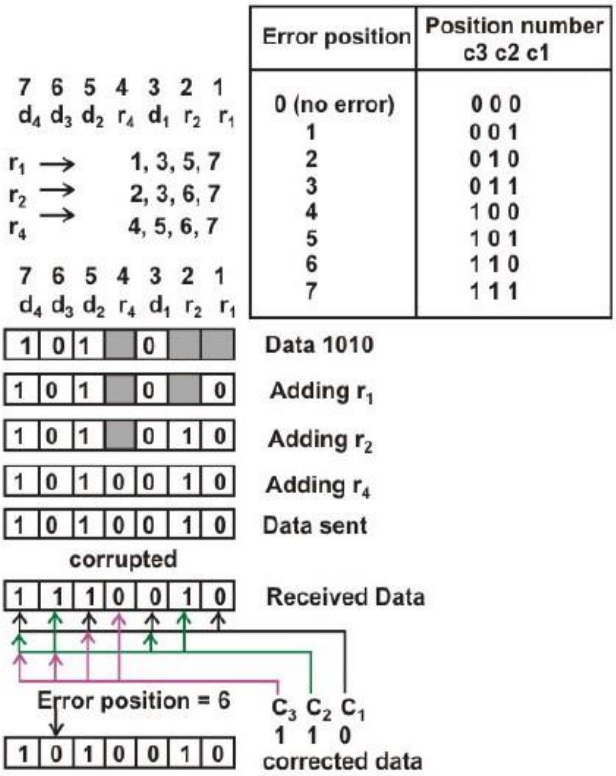
Basic approach for error detection by using Hamming code is as follows:

» To each group of m information bits k parity bits are added to form (m+k) bit code

» Location of each of the (m+k) digits is assigned a decimal value.

» The k parity bits are placed in positions 1, 2, …, 2k-1 positions.–K parity checks are performed on selected digits of each codeword.

» At the receiving end the parity bits are recalculated. The decimal value of the k parity bits provides the bit-position in error, if any.

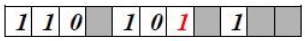


Use of hamming code for error correction for a 4-bit data

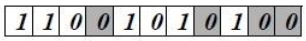
Hamming code is used for correction for 4-bit numbers ($d_4d_3d_2d_1$) with the help of three redundant bits ($r_3r_2r_1$). For the example data 1010, first r_1 (0) is calculated considering the parity of the bit positions, 1, 3, 5 and 7. Then the parity bits r_2 is calculated considering bit positions 2, 3, 6 and 7. Finally, the parity bits r_4 is calculated considering bit positions 4, 5, 6 and 7 as shown. If any corruption occurs in any of the transmitted code 1010010, the bit position in error can be found out by calculating $r_3r_2r_1$ at the receiving end. For example, if the received code word is 1110010, the recalculated value of $r_3r_2r_1$ is 110, which indicates that bit position in error is 6, the decimal value of 110.

Example:

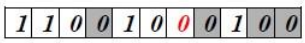
Let us consider an example for 5-bit data. Here 4 parity bits are required. Assume that during transmission bit 5 has been changed from 1 to 0 . The receiver receives the code word and recalculates the four new parity bits using the same set of bits used by the sender plus the relevant parity (r) bit for each set . Then it assembles the new parity values into a binary number in order of r positions (r_8, r_4, r_2, r_1).



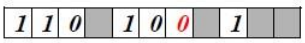
Data to be send



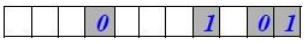
Data to be send along with redundant bits



Data Received



Data Received Minus Parity Bits



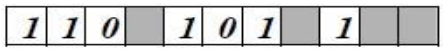
Parity bits recalculated

Calculations:

Parity recalculated (r_8, r_4, r_2, r_1) = 01012 = 510.

Hence, bit 5^{th} is in error i.e. d_5 is in error.

So, correct code-word which was transmitted is :



16.3 SUMMARY

Data can be corrupted during transmission. Some applications require that errors be detected and corrected. In a single-bit error, only one bit in the data unit has changed. A burst error means that two or more bits in the data unit have changed. To detect or correct errors, we need to send extra (redundant) bits with data. There are two main methods of error correction: forward error correction and correction by retransmission.

The correction of errors is more difficult than the detection. In error detection, we are looking only to see if any error has occurred. The answer is a simple yes or no. We are not even interested in the number of errors. A single-bit error is the same for us as a burst error.

In error correction, we need to know the exact number of bits that are corrupted and more importantly, their location in the message. The number of the errors and the size of the message are important factors.

The Hamming distance between two words is the number of differences between corresponding bits. The minimum Hamming distance is the smallest Hamming distance between all possible pairs in a set of words.

Hamming code is a set of error-correction codes that can be used to **detect and correct the errors** that can occur when the data is moved or stored from the sender to the receiver. It is a **technique developed by R.W. Hamming for error correction**. **Redundant bits –** Redundant bits are extra binary bits that are generated and added to the information-carrying bits of data transfer to ensure that no bits were lost during the data transfer. The number of redundant bits can be calculated using the following formula:

$2^r \geq m + r + 1$

Where, r = redundant bit, m = data bit

Suppose the number of data bits is 7, then the number of redundant bits can be calculated using: = 2^4 ≥ 7 + 4 + 1 Thus, the number of redundant bits= 4 **Parity bits.** A parity bit is a bit appended to a data of binary bits to ensure that the total number of 1's in the data is even or odd. Parity bits are used for error detection. There are two types of parity bits:

1. **Even parity bit:** In the case of even parity, for a given set of bits, the number of 1's are counted. If that count is odd, the parity bit value is set to 1, making the total count of occurrences of 1's an even number. If the total number of 1's in a given set of bits is already even, the parity bit's value is 0.
2. **Odd Parity bit –** In the case of odd parity, for a given set of bits, the number of 1's are counted. If that count is even, the parity bit value is set to 1, making the total count of occurrences of 1's an odd number. If the total number of 1's in a given set of bits is already odd, the parity bit's value is 0.

16.4 KEYWORDS

Error detection, error correction, even parity, odd parity, redundancy, hamming code, hamming distance, single bit error correction

16.5 QUESTIONS FOR SELF STUDY

1. Distinguish between error detection and error correction.
2. Explain odd parity and even parity with example.
3. Write the algorithm for hamming code.
4. Explain Hamming code for error detection.
5. Describe Hamming code for error correction.
6. How do you determine the position of redundant bits? Explain.
7. How do you determine the parity bits? Explain.
8. Explain single bit error correction with example.

16.6 REFERENCES

1. https://www.studytonight.com/computer-networks
2. https://www.tutorialspoint.com/
3. https://www.javatpoint.com/
4. https://aws.amazon.com/
5. https://www.geeksforgeeks.org/
6. https://faculty.ksu.edu.sa/sites/default/files/2-chapter2-routing-algorithms.pdf
7. https://ecomputernotes.com