Data Communication And Networking

DECAP453

Edited by
Ajay Kumar Bansal

LOVELY
PROFESSIONAL
UNIVERSITY

LOVELY
PROFESSIONAL
UNIVERSITY

Data Communication And Networking

Edited By:
Ajay Kumar Bansal

CONTENT

Unit 01: Introduction To Data Communication and Computer Networks

Dr. Rajni Bhalla, Lovely Professional University

CONTENTS

Objectives

After studying this unit, you will be able to:

- understand the basics of network and analyze the components required for communication.
- identify the different ways of representing the information and describe the main uses of computer networks
- learn different types of networks
- identify how one network differs from another network.
- explain the seven main topologies.

Introduction

The merging of computers and communications has a profound influence on the way systems are organized. The concept of a computer center as a room with a large computer to which the users bring their work for processing is now obsolete. The old model of a single computer servicing all the computational needs of an organization has been replaced by the one in which a large system of separate but interconnected computers does the job. These systems are called computer networks. The two computers are said to be interconnected if they can exchange information. The connection between the computers need not be only via a copper wire or fiber optics or microwaves. A communication satellite can be used for networking the computers.

History of Computer Networks

Following is a brief history of computers, networking, and telecommunication milestones:

1897: CRT (Cathode Ray Tube) credited to Braun

1900–1915: Teletype (telegraph 5 bit) 1915–1020: ARQ (Automatic Repeat request) credited to Van Duuren.

1915–1020: ARQ (Automatic Repeat request) credited to Van Duuren

1930–1940: ENIAC credited to DOD/MIT 1950s:

The 1950s: SAGE (Semi-Automatic Ground Environment) MIT 1950s

1960s: Transistorized Computers–2nd Generation 1961: CTSS (Compatible Time-Sharing System) credited to Cobato/MIT.

1965: Auto Equalization Techniques of Phone lines credited to Lucky et al. 1966: Fiber Glass credited to Kao & Hockman

1967: Integrated Circuits Computers–3rd Generation

1968: Carterfone–FCC Decision in

1969: A group of DoD researchers linked four computers at UCLA, SRI, University of Utah, and the UCSB. They created a network to communicate with one another about government projects. The network was part of the DoD's Advanced Research Project Agency and was dubbed ARPAnet

1972: More than 50 universities and military agencies were linked together on the network. For a short period, it was a top-secret defense project, ensuring that computers could talk to each other in the event of a nuclear attack. The communication system between the sites was called email and was invented by Ray Tomlinson of Bolt, Berank, and Newman.

1973: The defence project links were extended to Norway and England.

1974: Transmission Control Protocol (TCP) was published and the military and educational links diverged. Organizations like NASA began to experiment with computer networks, and the networks began to interconnect and the name Internet was coined.

1976: The Queen sends an email from RSRE Malvern.

1983: TCP/IP become the protocol standard for ARPAnet. Scott Fahlman invents the smiley to convey emotions in email.

1984: In the US, the NSF built high-speed, long-distance lines that connected supercomputer sites across the USA. These eventually replaced the original ARPAnet. In time, NSFnet was joined by other networks at dozens of universities, research laboratories, and high-tech companies. The system for assigning names to computers on the network was introduced — DNS. JANET was launched to connect British Universities.

1986: The NSF established its own faster network NSFnet and Network News Transfer Protocol (NNTP) was introduced making on-line interactive discussion a reality. The backbone speed was 56 Kbps.

1987: 1000 Notes th RFC and 10,000th host.

1988: Robert Tappan Morris releases the first Internet Worm and CERT was set up in response to this. Backbone speed upgraded to 1.544 Mbps. IRC developed.

1989: 100,000th host. Cuckoo's Egg was released by Cliff Stoll telling the true story of East German crackers accessing US installations.

1990: ARPAnet ceased to exist and the Internet effectively took its role.

1991: Gopher, a software program for retrieving information from servers on the Internet was made available by the University of Minnesota. The US Government announced that it no longer intended to restrict activity on the Internet to research. This policy shift was sufficient for 12 companies to co-operate and produce CIX. Phil Zimmerman released PGP. Backbone speed upgraded to 44.736 Mbps.

1992: The World Wide Web became a possibility after CERN, in Switzerland, released hypertext. 1,000,000th Host. The author gets his first dial-up email account with Demon Internet (Nov. 1992).

1993: Mosaic, a software program to browse Web sites written by Marc Andreesen, was released followed by Netscape.

1994: Shopping Malls arrive on the Internet. The UK Treasury goes online and the first cyberbank opens. The first banner adverts appeared for Zima (a drink) and AT&T.

1995: Traditional dialup services (AOL, CompuServe, etc) start to provide dialup services. The Vatican goes online. Several Internet companies go public. Netscape leads the field with the largest ever IPO on NASDAQ. DEC launches AltaVista, which claims to index every HTML page there is. Jeff Bezos launches Amazon.com. eBay is launched.

1996: 9,272 organizations find themselves unlisted after the InterNIC drops their name service as a result of not having paid their domain name fee. Various ISPs suffer extended service outages, bringing into question whether they will be able to handle the growing number of users. AOL (19 hours), Netcom (13 hours), AT&T WorldNet (28 hours - email only). China requires users of the Internet to register with the Police. Saudi Arabia restricts use to universities and hospitals. Domain name tv.com sold to CNET for US$15,000. Backbone speed upgraded to 622 Mbps.

1997: 2000th RFC. 16 Million hosts. 1,000,000th Domain name registered (March 6th for Bonny View Cottage Furniture Company).

1998: 3,000,000th Domain name registered. US Postal authorities allow the purchase of postage stamps online for downloading and printing. Gigabit Ethernet standard ratified. Google is launched.

1999: First full-service bank opens on the Internet (First Internet Bank of Indiana). The first forged web page, looking like Bloomberg, raises the shares of a small company by 31% (7th April). Melissa strikes. 5,000,000th Domain name registered. First Cyberwar starts between Serbia and Kosovo. Shawn Fanning Launches Napster — record labels are furious.

2000: 10,000,000th Domain name registered. French Courts require that 'hate' memorabilia for sale on Yahoo's auction site must be removed. Gnutella is launched. ICANN selects new top-level domains. Backbone is upgraded to IPv6.

2001: Forwarding email becomes illegal in Australia (Digital Agenda Act). Napster was forced to suspend service after legal action. Taliban bans the Internet in Afghanistan. Nimda was released on the Internet.

2002: Distributed denial of Service attack hits 13 DNS root servers, causing national security concerns.

2003: The first official Swiss online election takes place in Anières (7 Jan), SQL Slammer (goes round the world in 10 minutes and takes out 3 of the 13 DNS Servers). Followed by SoBig.F (19 Aug) and Blaster (11 Aug).

2004: Lycos Europe releases a screen saver to help fight spam by keeping spam servers busy with requests (1 Dec). The service is discontinued within a few days after backbone providers block access to the download site and the service causes some servers to crash.

1.1 Data Communication

What is communication?

When two people communicate with one another that is known as communication as shown in Figure 1. When the sender is sending a message, They are sharing their ideas that are called communication.

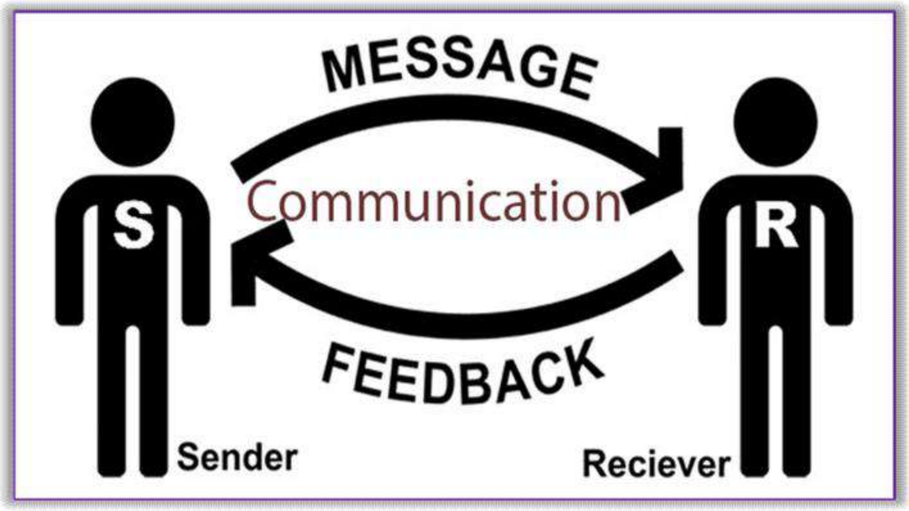


Figure 1 Data Communication

What is Data Communication?

When you are electronically transferring data from one place to another place by using a medium that is called data communication. When you are sharing data with another person that can be in any location or it can be in a remote location also. Communication always happens through some medium. If two people are communicating. We can say that medium is air. Similarly, when the sender wants to send data, the medium will be used to send data. That medium is called a transmission medium as shown in Figure 2. Medium is the physical path through which sender and receiver connect. A medium can be wired or wireless. Examples of the wired medium can be copper wire, twisted pair cable, and fiber optic cable as shown in Figure 3.

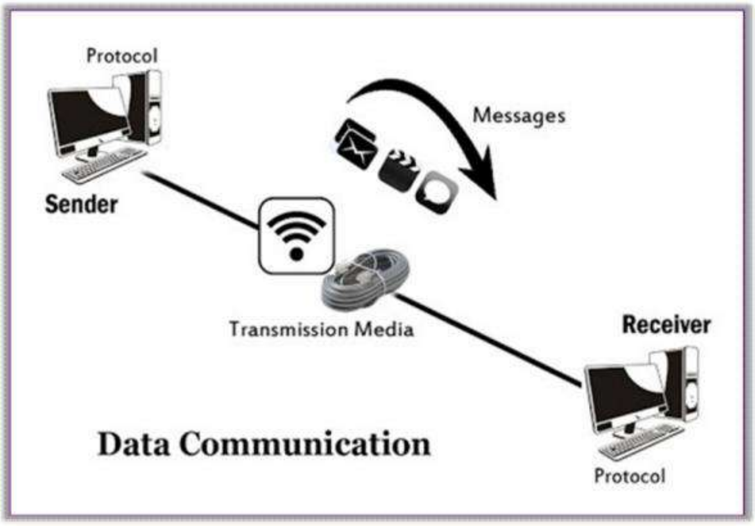


Figure 2 Transmission Media

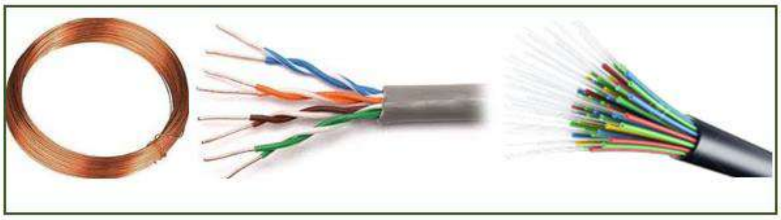


Figure 3 Medium for Data Transfer

1.2 Defining Network

A network consists of two or more computers that are linked to share resources (such as printers and CD-ROMs), exchange files, or allow electronic communications. The computers on a network may be linked through cables, telephone lines, radio waves, satellites, or infrared light beams. The term 'computer network' means an interconnected collection of autonomous computers.

(a) Two computers are said to be interconnected if they can exchange information.

(b) The requirement for computers to be autonomous excludes from our definition systems in which there is a clear master/slave relation.

The key difference between a computer network and a distributed system:

In a distributed system, the existence of multiple autonomous computers is transparent to the user. A distributed system looks like a virtual uni-processor to its users.

With a network, a user must explicitly do the followings:

- log onto one machine (e.g., login),
- submit jobs remotely (e.g., rsh),
- move files around (e.g., rcp, FTP, UUCP),
- and generally handle all the network management personally.

In effect, a distributed system is a special case of a network, one whose software gives it a high degree of cohesiveness and transparency. The network is everywhere. The best examples of the network are railways, radio, water taps, and water supply chain systems in ancient Rome as shown in figure 5.

Figure 5 Best Examples of Network

What type of shared resources can be shared?

When one system sends data to another, it can share resources. Examples of resources are the internet, printer, and file server as shown in Figure 6.

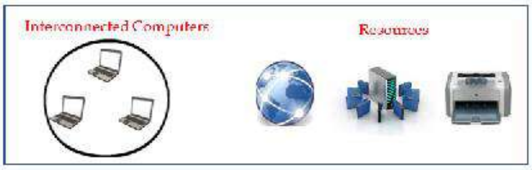


Figure 6 Examples of Shared Resources

Shared resources can be in the form of hardware or they can in the form of software. Resources related to hardware are disks, printers, and scanners. Resources related to software are Files, applications that are also known, and application software as shown in Figure 7. Any type of file can be shared through the network. It can be in form of text, audio, and video.

Figure 4 Best examples of network

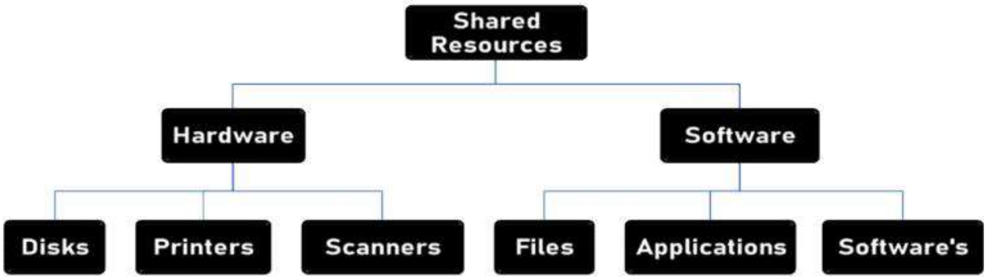


Figure 7 Shared Resources

For an example of hardware sharing, suppose if you only have one printer in your classroom that you want to share with several student computers, you can connect that printer to your computer and then share it so that it is accessible from any computer in the room. One of the major advantages is the utilization of the resource. Proper utilization of resources like the printer is doing here. No need to purchase a separate printer for each system. Maintenance costs also reduce when we share the resource.

1.3 Benefits of computer network

- File-Sharing: We can share a file with a different number of users.
- Hardware Sharing: If you are working in one network and if you share a file Best example is a printer. If 10 faculties are seated in a single room. Instead of purchasing an individual printer for each faculty, all can share a single printer. It reduces maintenance costs and increases resource utilization.
- Application Sharing: By using the internet, we can view and control the application that exists in a remote location. It means we can use applications that exist in some other system.
- User Communication: Two users can communicate with another. Two users can be connected using a router or switch or hub or any other device. They can be connected using peer-to-peer connection also.
- Network Gaming: Nowadays, it is very much important. All students joined in one location to play the game. They are located in different locations. Just to play a game, they are joining in a single location.

1.4 Data Communication and Networking

It is changing the way we do business. Day by day, business trending is changing. Today businesses are not limited to offline. When you want to extend your business, you want to opt online. The sellers are going to the online website so that they sell their products. Seller also wants to join their remote customer who is not in the same location. By this approach, users can expand their business. Due to competition in the market, business decisions have to be made more quickly. For this reason, decision-makers require immediate access to accurate information.

For example:

Why wait a week for that report from Germany to arrive by mail when it could appear instantaneously through computer networks. Business today relies on computer networks and internetwork.

Before we get into details of networking, we need to know: -

1. ***How does the network operate?***
2. ***What type of technologies are available?***
3. ***Which design best fills which set of needs?***

It means what type of network you wanted to build. Suppose if you wanted to connect the system to a small institute. You need to decide either you are going to connect all systems either using a bus topology, ring topology or you want to connect using a hybrid topology. How you are going to connect all systems. To make connectivity between all systems together, basic knowledge is required. If you want all systems able to shared resources and data with one another, it means technology knowledge is very much required. You must know what type of technology is required, which is the best technology to connect systems.

The development of the personal computer brought tremendous changes for:

- Business
- Industry
- Science
- Education

It has changed the way people are looking. We are looking for online education. Students are looking for specialized certificate courses. Similarly, revolution is occurring in data communication and networking. Technology advances are making it possible for communications links to carry more and faster signals. Services are evolving to allow the use of this expanded capacity.

Example: Telephone services extended to have:

- Conference calling
- Call waiting
- Voice mail and
- Caller ID

4. *What is the role they are playing in the research?*

Research in data communications and networking has resulted in new technologies. One goal is to be able to exchange data such as text, audio, and video from all points in the world.

5. *Why would we want to access the internet?*

We want to access the internet to download and upload information quickly and accurately and at any time.

6. ***What** is communication*

When two people connected through a medium so that they can communicate and share knowledge. Sharing can be local or remote.

Local Sharing

Local sharing means two users are directly communication with one another. Example face to face communication as shown in Figure 8.

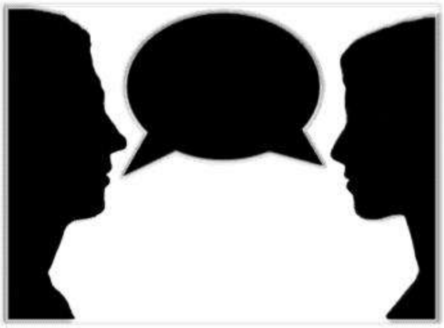


Figure 8 Local Communication

Remote Sharing

Remote sharing means when we talk about distance. User with whom we want to communicate that is not located locally. The user is located at a remote location. It means the user is located at some other place. Example telecommunication means telephony, telegraphy, and television means communication at a distance as shown in Figure 9.

Figure 9 Remote Sharing

What is data?

Data refers to information presented in any form and 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 wire cable. For data communication to occur the communicating devices must be a part of a communicating 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.

- 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. If it's not reaching to correct destination then security is compromised here.
- Accuracy: The system must deliver the data accurately. Data that have been altered in transmission and left uncorrected are unusable. When the sender is sending data, in between no manipulation of data is allowed.
- Timeliness: The system must deliver data promptly. 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**. Especially in case if you are sending audio and video, data must deliver in time. For example, if we play a video and audio is coming after few minutes. So, I cannot enjoy listening to that video. There should not be much delay of data.
- Jitter: Jitter refers to the variation in the packet arrival time. It is an uneven delay in the delivery of audio or video packets. There should not be many variations in the early packet and coming packet as shown in Figure 10. Variations in the packet should less. Jitter in between first and second packet is of 5 minutes and second to third is also 5 minutes and third, to the fourth packet is taking 15 minutes. It is not at all acceptable. Jitter means variation in the packet delay should be consistent as shown in Figure 10.

Components of Data Communication

Five components of data communication play a major role as shown in figure 11:

- Sender
- Receiver
- Message
- Transmission medium
- Protocol

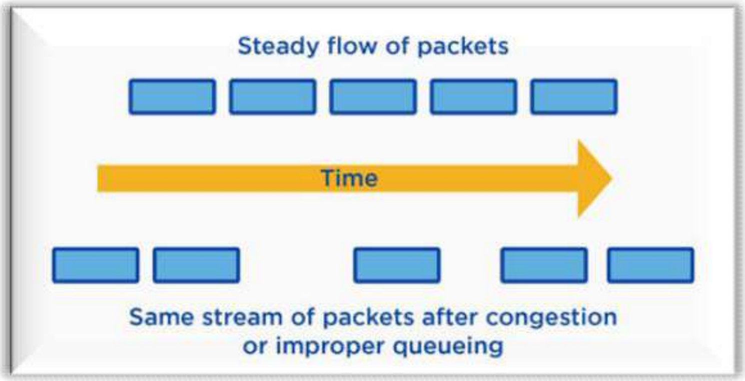

Steady flow of packets
Time
Same stream of packets after congestion or improper queueing

Figure 10 Jitter Variation in Packets

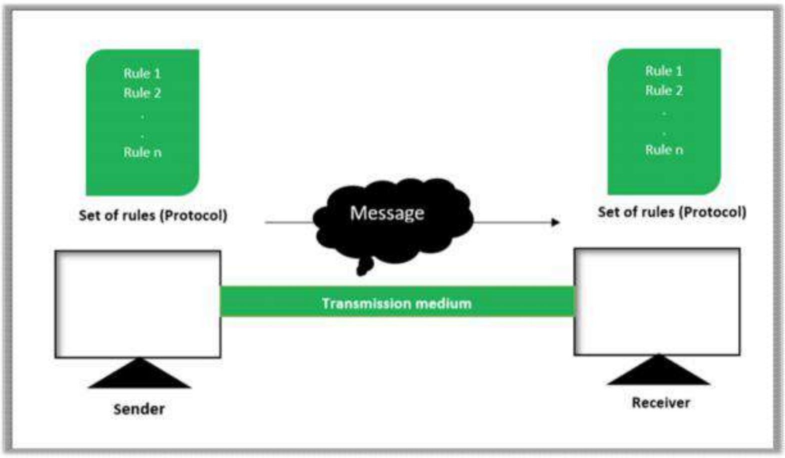


Figure 11 Components of Data Communication

Sender: Sender who sends the data. It can be a workstation from where you are sending the data.

Receiver: Receiver who has the capability of receiving the data. It can also be a workstation that is receiving the data.

Message: The information that users want to send and that can be in the form of text, audio, and video.

Transmission medium: it can be either wired or wireless medium. Wired media are also known as guided media and wireless media means without wire. Examples of wireless media can be microwave, radio wave, Bluetooth, or mobile communication.

Protocol: The sender is going to use some rules and regulations to send a message to the destination. Communication should happen by following some rules. For example: If the sender knows only French and the receiver knows only English. They have to follow some rules to communicate with one another. Otherwise, they won't be able to communicate with one another.

Data Representation

Information can be represented in text, images, numbers, audio, and videos form.

Text is a collection of alphabets. Smaller case and uppercase alphabets are also allowed. These alphabets will be converted into the form of bits. Information can be represented in the form of images also that would be converted into bits. Numbers are in the form of digits. Information can be represented in the form of audio. If you wanted to send recorded sound that can also be sent. When you want o send a video to your friend that can also be in the form of bits.

1.5 Applications Of Computer Networks

The main uses of computer networks are in business, home, mobile and social issues.

Business Applications

Most of the companies have a substantial number of computers. There uses are

- **Resource Sharing:**

The first one is resource sharing that helps to share the resource. Computer networks allow organizations to have units that are placed apart from each other, to share information in a very effective manner. Programs and software in any computer can be accessed by other computers linked to the network. It also allows sharing of hardware equipment, like printers and scanners among varied users as shown in Figure 12.

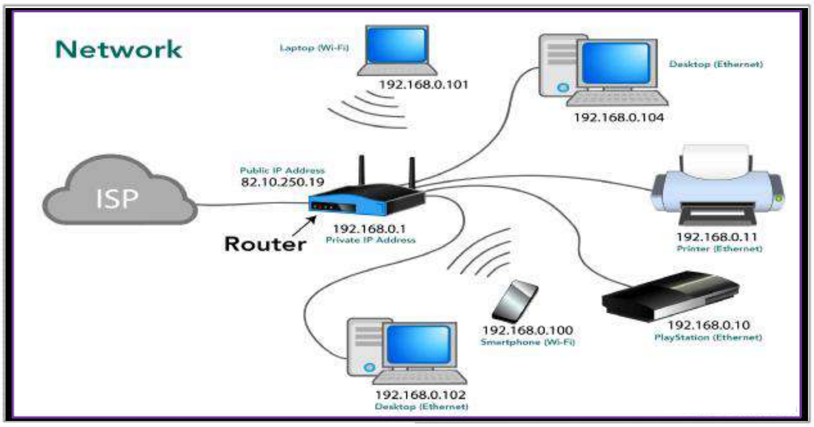


Figure 12 Resource Sharing

- **Server Client Model:**

A client-server model is shown in Figure 13 which consists of two clients and one server and they are connected by a network.

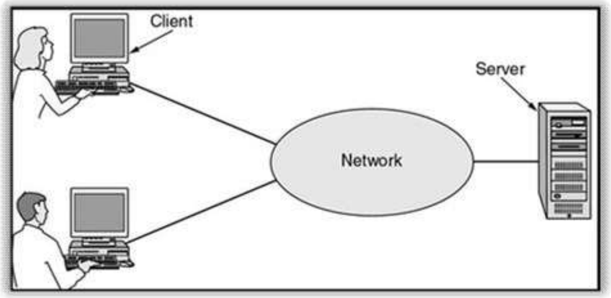


Figure 13 Client/Server Model

The client is sending a request and the server is replying because they are connected through a network. The client-server model describes how a server provides resources and services to one or more clients. Examples of servers include web servers, mail servers, and file servers. Each of these servers provides resources to client devices, such as desktop computers, laptops, tablets, and smartphones.

- **Communication medium**

A computer network can provide a powerful communication medium among employees. Virtually every company that has two or more computers now has e-mail (electronic mail), which employees generally use for a great deal of daily communication as shown in.

Figure 14 Communication Medium

- **E-commerce**

A goal that is starting to become more important in businesses is doing business with consumers over the Internet.

Figure 15 Examples of E-commerce

Airlines, bookstores, and music vendors have discovered that many customers like the convenience of shopping from home as shown in Figure 15. This sector is expected to grow quickly in the future. The most popular forms in E-commerce are listed in

Table 1 Most Popular Forms in E-commerce

The most popular forms are listed	
B2C (Business to Consumer)	When consumer order shoes online
B2B (Business to Business)	Truck manufacturer oredering tires from suppliers
C2C (Consumer to Consumer)	An online auction site(ebay)
G2C (government to Citizen)	reduce the average time for fulfilling citizen’s requests for various government services.
P2P (Peer to Peer)	Buyer and seller transact directly

- **P2P Communication**

Stands for "Peer to Peer." In a P2P network, the "peers" are computer systems that are connected via the Internet. Files can be shared directly between systems on the network without the need for a central server. In other words, each computer on a P2P network becomes a file server as well as a client.

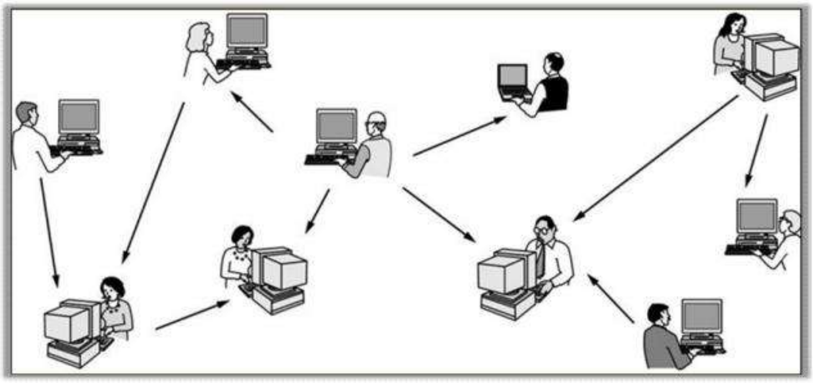


Figure 16 Peer to Peer Communication

Once connected to the network, P2P software allows you to search for files on other people's computers. Meanwhile, other users on the network can search for files on your computer, but typically only within a single folder that you have designated to share. While P2P networking makes file sharing easy and convenient, it also has led to a lot of software piracy and illegal music downloads. Therefore, it is best to be on the safe side and only download software and music from legitimate websites.

Home Applications

Some of the most important uses of the Internet for home users are as follows:

- ***Access to remote information:***
- ***Person-to-person communication***
- ***Interactive entertainment***
- ***Electronic commerce***

Mobile Users

Mobile computers, notebook computers are one of the fastest-growing segments of the entire computer industry. Since having a wired connection is impossible in cars, boats, airplanes, there is a lot of interest in wireless networks. Wired connection is not possible in these areas, so wireless connection comes into the picture. For example, people who usually traveling, they prefer to use portable devices.

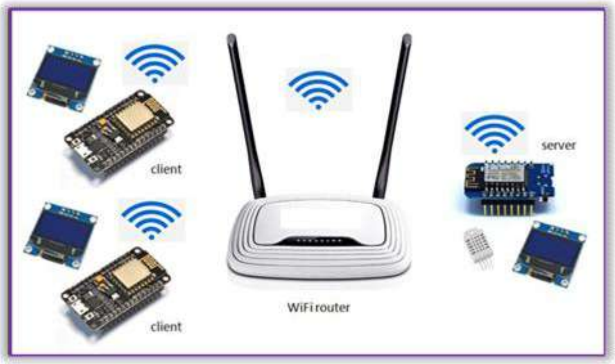


Figure 17 Wireless Hotspots

Wireless hotspots are another kind of wireless network for mobile computers Figure 17. Smartphones such as the popular iPhone, combine aspects of mobile phones and mobile computers. Although wireless networking and mobile computing are often related, they are not identical, as the below Figure 18 shows:

Wireless	Mobile	Applications
No	No	Desktop computer in offices
No	Yes	A notebook computer used in a hotel room
Yes	No	Networks in older, unwired buildings
Yes	Yes	Portable office; PDA for store inventory

Figure 18 Mobile Users

Social Issues

The growth in the availability of affordable computing technology has caused several major shifts in the way that society operates. The majority of these have been for the better, with home computers and the internet providing unlimited access to all of the information ever created and discovered by humanity.

- *Communication Breakdown*

Socializing within a family unit has always been important, as it strengthens the bonds between us and ensures cohesion within the group. But with more and more households owning several computers and numerous portable devices granting access to information and entertainment, some argue that this is leading to a lack of family communication. If each member is engrossed in their laptop, smartphone, or tablet each evening, even communal things like watching television are compromised. Meanwhile, you can see whole families who are out to dinner and still staring into a touchscreen rather than talking to one another. And if you're the one driving to that family dinner and texting while driving, you're a distracted driver, increasing your risk of crashing, and potentially causing death and injury. Increase your digital wellbeing by allowing technology to improve your life and not become a distraction to your life and others. Your life and others are more important than technology.

- *Unauthorized Access*

Unauthorized access is when someone gains access to a website, program, server, service, or other system using someone else's account or other methods. For example, if someone kept guessing a password or username for an account that was not theirs until they gained access, it is considered unauthorized access.

- *Authentication*

Authentication is the process of recognizing a user's identity. It is the mechanism of associating an incoming request with a set of identifying credentials. The credentials provided are compared to those on a file in a database of the authorized user's information on a local operating system or within an authentication server.

- *Ethical Use of Networks*

It is a generally accepted policy to maintain access to the **networks** as open as possible; it is thus unavoidable that certain **ethical** standards be imposed on the activities of individual users. Such demands differ little from those placed on other members of a modern civilized community.

- *dentity Theft*

Identity theft, also known as *identity fraud*, is a crime in which an imposter obtains key pieces of personally identifiable information (PII), such as Social Security or driver's license numbers, to impersonate someone else. The taken information can be used to run up debt purchasing credit, goods, and services in the name of the victim, or to provide the thief with false credentials. In rare cases, an imposter might provide false identification to police, creating a criminal record or leaving outstanding arrest warrants for the person whose identity has been stolen.

- ***Cyberbullying***

Cyberbullying or cyberharassment is a form of bullying or harassment using electronic means. Cyberbullying and cyberharassment are also known as online bullying. It has become increasingly common, especially among teenagers, as the digital sphere has expanded and technology has advanced.

- ***Gaming Addiction***

Video game addiction is the compulsive or uncontrolled use of video games, in a way that causes problems in other areas of the person's life. Often considered a form of a computer addiction or internet addiction, video game addiction has been an increasing concern for parents as video games have become more commonplace and are often targeted at children.

- ***Health & Fitness***

Fitness involves the activity of some sort that stimulates various systems of the body and maintains a certain condition within the body. Health, on the other hand, involves every system of the body and is only achieved through a lifestyle that supports health.

1.6 Types of Networks

- Personal Area Network (PAN)
- Local Area Network (LAN)
- Wide Area Network (WAN)
- Metropolitan Area Network (MAN)

Personal Area Network (PAN)

A personal area network (PAN) is the interconnection of information technology devices within the range of a person, typically within a range of 10 meters. For example, a person traveling with a laptop, a personal digital assistant (PDA), and a portable printer could interconnect them without having to plug anything in, using some form of wireless technology. Typically, this kind of personal area network could also be interconnected without wires to the Internet or other networks as shown in Figure 19.

PAN is organized around a person within a single building. This could be inside a small office or residence. PAN would include one or more computers, telephones, peripheral devices, video game consoles, and other personal entertainment devices. Two or more devices are located on your person and normally connect via Bluetooth.

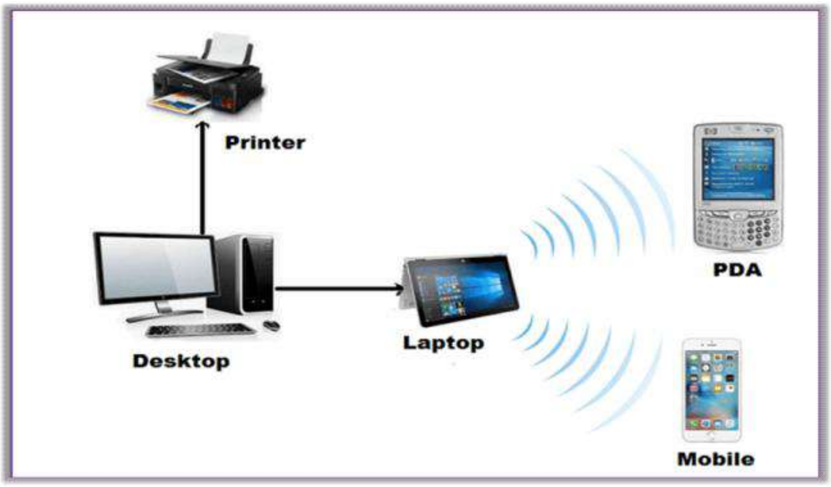


Figure 19 Personal Area Network

Types of PAN

PAN can be of different types. It can be wired PAN or Wireless PAN defined in diagram below:

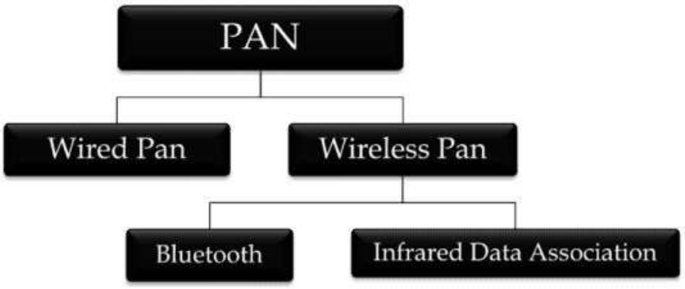


Figure 20 Types of PAN

Wired PAN

- The data cable is an example of the above PAN.
- This is also a Personal Area Network because that connection is for the user's personal use. PAN is used for personal use only.

Wireless Personal Area Network

The two kinds of wireless technologies used for WPAN are Bluetooth and Infrared Data Association. **Bluetooth** uses short-range radio waves over distances up to approximately 10 meters. For example, Bluetooth devices such as a keyboard, pointing devices, audio headsets, printers may connect to personal digital assistants (PDAs), cell phones, or computers wirelessly.

Figure 21 Examples of Bluetooth and Infrared data association

Infrared Data Association (IrDA) uses infrared light, which has a frequency below the human eye's sensitivity. Infrared, in general, is used, for instance, in TV remotes.

Local Area Network

LAN is normally located inside the same building or structure or residence, school, laboratory, or office building. A LAN can be as simple as two computers connected to a single switch. A LAN is very useful for sharing resources, such as data storage and printers. LANs can be built with relatively inexpensive hardware, such as hubs, network adapters, and Ethernet cables. High speed and relatively low cost are the defining characteristics of LANs. It is usually privately owned networks. LANs are typically used for single sites where people need to share resources among themselves but not with the rest of the outside world as shown in Figure 22.

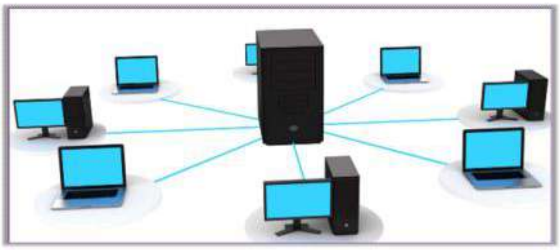


Figure 22 Local Area Network

Two most common transmission technologies in use for local area networks:

- Ethernet over twisted-pair cabling

- Wi-Fi

Ethernet over twisted pair cabling

Wi-Fi(Wireless Fidelity)

Historical technologies include ARCNET, Token Ring, and AppleTalk as shown in Figure 23.

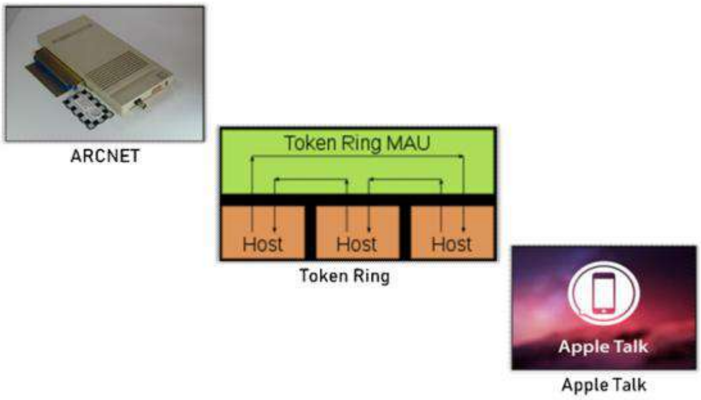


Figure 23 Historical Technologies of LAN

RELATED TO LAN IS WLAN

Another name for WLAN is wi-fi(Wireless Fidelity).

Classification of LAN

Computer networks can logically be classified into 1) peer-to-peer networks and 2) client-server networks as shown in Figure 24.

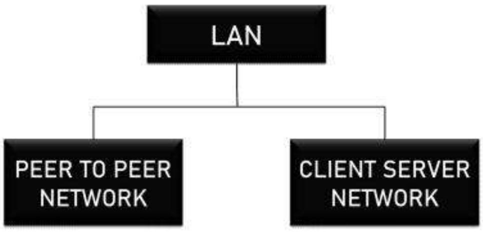


Figure 24 Types of LAN

Peer to Peer Network

Each computer is responsible for making its resources available to other computers on the network. Responsible for setting and maintaining own security for these resources. Each computer is responsible for accessing the required resources from peer-to-peer relationships. Peer to peer network is usually for a small network containing less than 10 computers on a single LAN. Do not have a central control system. There are no servers in peer networks as shown in Figure 25. Peer to Peer networks is amplified into homegroups.

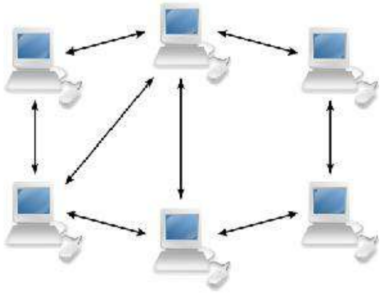


Figure 25 Peer to Peer Network

Advantages and disadvantages of a peer network

Table 2 Advantages and Disadvantages of Peer Network

Advantages	Disadvantages
Use less expensive computer hardware	Not Very Secure
Easy to administer	No central point of storage or file archiving
No NOS required	The additional load on the computer because of resource sharing
More built-in redundancy, easy setup, and low cost	Hard to maintain version control.

Client/Server Network

A computer network in which one centralized, powerful computer (called the server) is a hub to which many less powerful personal computers or workstations (called clients) are connected. The clients run programs and access data that are stored on the server.

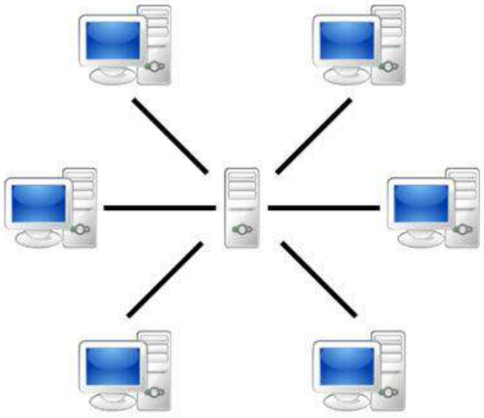


Figure 26 Client/Server Network

Table 3 Advantages and Disadvantages of Client/Server

Advantages	Disadvantages
Very secure	Requires professional administration
Better performance	More hardware intensive
Centralized backup	More software-intensive

Very reliable	Expensive dedicated software

Metropolitan Area Network

A **metropolitan area network** (**MAN**) is a computer network larger than a local area network. MAN knew as *municipal area network.* consists of a computer network across an entire city, college campus, or small region. A MAN is often used to connect several LANs to form a bigger network. A MAN is larger than a LAN, which is typically limited to a single building or site. An example of MAN is shown in Figure 27 below.

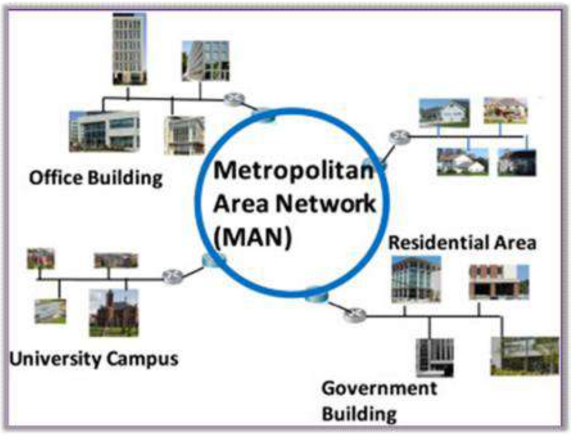


Figure 27 Metropolitan Area Network

Networking technologies used in municipal networks

- Asynchronous Transfer Mode (ATM),
- FDDI,
- and SMDS(Switched Multi-megabit Data Service)

Wide Area Network

A **wide area network**, or **WAN**, occupies a very large area, such as an entire country or the entire world. A WAN can contain multiple smaller networks, such as LANs or MANs. The Internet is the best-known example of a public WAN.

How WAN networks are established?

Computers connected to a wide-area network are often connected through public networks, such as the telephone system. They can also be connected through **leased lines or satellites**. The largest WAN in existence is the Internet.

1.7 Difference Between LAN and WAN

Table 4 Difference Between LAN and WAN

LAN	WAN
Owned by a person=Privately owned	Can be private or public
Operate over small area	Large distance across countries.
Easy to design and maintain	Not Easy
Communication medium is generally coaxiable cables.	Satellite links(because of countries)
Minimum propagation delay(**time** required for a digital signal to travel from the input(s) of a logic gate to the output)	Excessive propagation delay

High data rate(because of small distance)	Low data rate
Broadcasting	Switching techniques

Table 5 Difference Between LAN, MAN, and WAN

Parameters	LAN	WAN	MAN
Ownership of network	Private	Private or Public	Private or Public
Geographical area covered	Small	Very large	Moderate
Design and maintenance	Easy	Not easy	Not easy
Communication medium	Coaxial cable	PSTN or satellite links	Coaxiable cables,PSTN,optical fiber cable,wireless
Bandwidth	Low	High	Moderate
Data Speed	High	Low	Moderate

1.8 Network Topologies

A Network Topology is the arrangement with which computer systems or network devices are connected. Topologies may define both physical and logical aspects of the network. Both logical and physical topologies could be the same or different in the same network.

- ***The physical topology*** is the placement of the various components of a network, including device location and cable installation,
- while ***logical topology*** illustrates how data flows within a network, regardless of its physical design

Point-to-Point

The simplest topology with a dedicated link between two endpoints. The simplest topology is a permanent link between two endpoints. A children's tin can telephone is **one example** of a physical dedicated channel.

Types of Topologies

There are seven main topologies as shown in Figure 28 below:

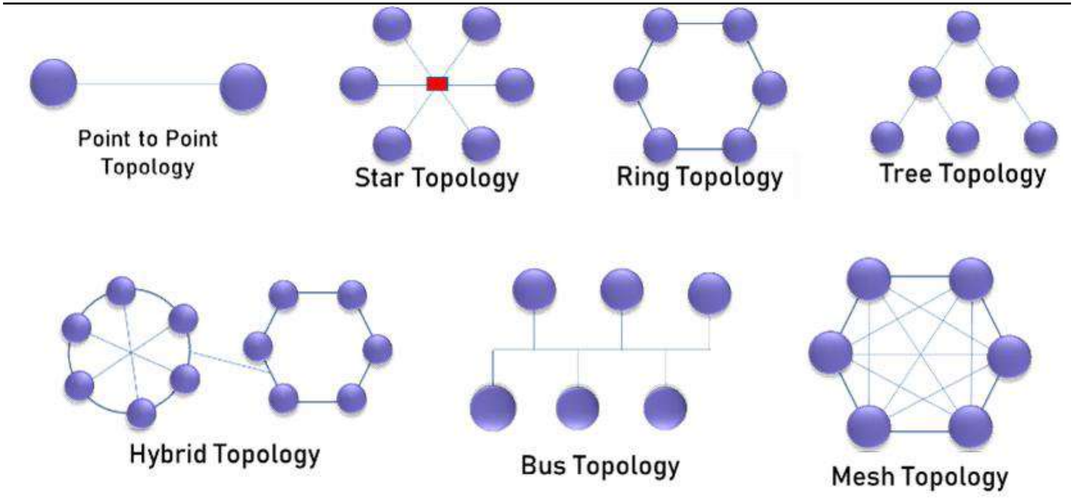
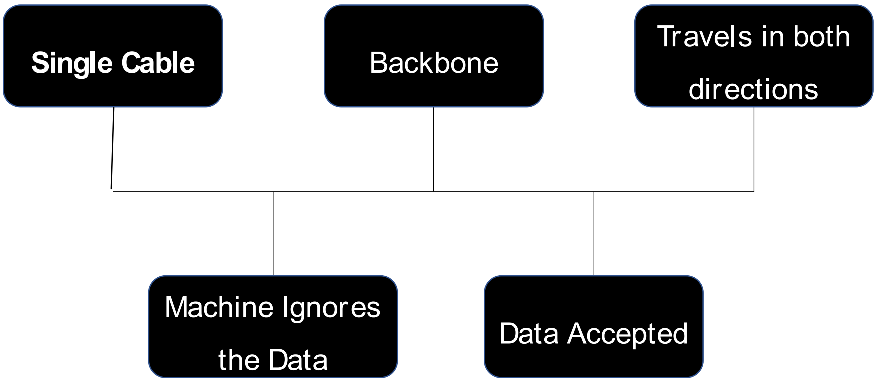


Figure 28 Types of Topologies

Bus Topology

In a bus topology, each node is connected to a single cable, with the help of interface connectors. This central cable is the backbone of the network and is known as the bus. A signal from the source travels in both directions to all machines connected on the bus cable until it finds the intended recipient. If the machine address does not match the intended address for the data, the machine ignores the data. Alternatively, if the data matches the machine address, the data is accepted.



The benefit over other technology

Because the bus topology consists of only one wire, it is rather inexpensive to implement when compared to other topologies. because only one cable is utilized, it can be the single point of failure.

Working of Bus Topology

A signal from the source is broadcasted and it travels to all workstations connected to bus cable. Although the message is broadcasted only the intended recipient, whose MAC address or IP address matches, accepts it. If the MAC /IP address of the machine doesn't match with the intended address, the machine discards the signal. A terminator is added at the ends of the central cable, to prevent the bouncing of signals.

Advantages (benefits) of Linear Bus Topology

1) It is easy to set-up and extends the bus network.
2) The cable length required for this topology is the least compared to other networks.
3) Bus topology costs very little.
4) Linear Bus network is mostly used in small networks. Good for LAN.

Disadvantages (Drawbacks) of Linear Bus Topology

1. There is a limit on central cable length and the number of nodes that can be connected.
2. Dependency on central cable in this topology has its disadvantages. If the main cable (i.e. bus) encounters some problem, the whole network breaks down.
3. Proper termination is required to dump signals. The use of terminators is a must.
4. It is difficult to detect and troubleshoot faults at the individual station.
5. Maintenance costs can get higher with time.
6. The efficiency of Bus network reduces, as the number of devices connected to it increases
7. It is not suitable for networks with heavy traffic.
8. Security is very low because all the computers receive the sent signal from the source.

Star Topology

In local area networks with a star topology, each network host is connected to a central hub with a point-to-point connection. So it can be said that every computer is indirectly connected to every other node with the help of the hub. In Star topology, every node (computer workstation or any other peripheral) is connected to a central node called a hub, router, or switch. The switch is the server and the peripherals are the clients. The network does not necessarily have to resemble a star to be classified as a star network, but all of the nodes on the network must be connected to one central device as shown in Figure 29. All traffic that traverses the network passes through the central hub. The hub acts as a signal_repeater._The star topology is considered the easiest topology to design and implement. Devices typically connect to the hub with Un-shielded Twisted Pair (UTP) Ethernet.

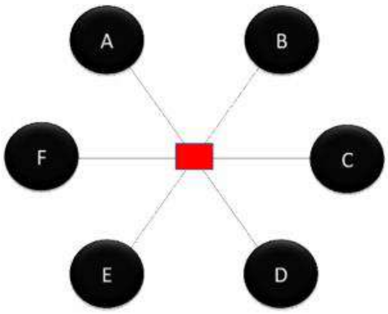


Figure 29 Star Topology

Primary Disadvantage Compared to Other

The primary disadvantage of the star topology is that

1. The hub represents a single point of failure.
2. Compared to the bus topology, a star network generally requires more cable.
3. The use of a hub, a router, or a switch as a central device increases the overall cost of the network.
4. The use of a hub, a router, or a switch as a central device increases the overall cost of the network.

Advantages of Star Topology

- As compared to Bus topology it gives far much better performance, signals don't necessarily get transmitted to all the workstations. A sent signal reaches the intended destination after passing through no more than 3-4 devices and 2-3 links. The performance of the network is dependent on the capacity of the central hub.
- Easy to connect new nodes or devices. In star topology, new nodes can be added easily without affecting the rest of the network. Similarly, components can also be removed easily.
- Failure of one node or link doesn't affect the rest of the network. At the same time, it's easy to detect the failure and troubleshoot it. Centralized management. It helps in monitoring the network.

Ring Topology

A network topology is set up circularly in such a way that they make a closed loop. This way data travels around the ring in one direction and each device on the ring acts as a repeater to keep the signal strong as it travels. Each device incorporates a receiver for the incoming signal and a transmitter to send the data to the next device in the ring. The network is dependent on the ability of the signal to travel around the ring. When a device sends data, it must travel through each device on the ring until it reaches its destination. Every node is a critical link.[4] In a ring topology, there is no server computer present; all nodes work as a server and repeat the signal. Sending and receiving of data takes place with the help of TOKEN.

Token passing

The token contains a piece of information which along with data is sent by the source computer. This token then passes to the next node, which checks if the signal is intended for it. If yes, it receives it and passes the empty to into the network, otherwise passes the token along with the data to the next node. This process continues until the signal reaches its intended destination. The nodes with token are the ones only allowed to send data. Other nodes have to wait for an empty token to reach them. This network is usually found in offices, schools, and small buildings. The structure is shown in Figure 30.

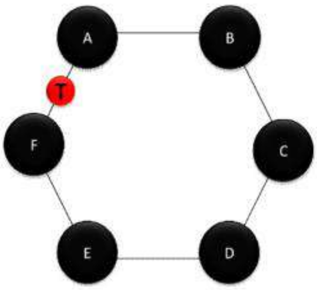


Figure 30 Ring Topology

Advantages of Ring Topology

1. This type of network topology is very organized. Each node gets to send the data when it receives an empty token. This helps to reduce the chances of a collision. Also in a ring topology, all the traffic flows in only one direction at a very high speed.
2. Even when the load on the network increases, its performance is better than that of Bus topology.
3. There is no need for the network server to control the connectivity between workstations.
4. Additional components do not affect the performance of the network.
5. Each computer has equal access to resources.

Disadvantages of Ring Topology

1. The disadvantage of this topology is that if one node stops working, the entire network is affected or stops working.
2. Each packet of data must pass through all the computers between source and destination. This makes it slower than Star topology
3. If one workstation or port goes down, the entire network gets affected.
4. The network is highly dependent on the wire which connects different components.

Mesh Topology

In a mesh topology, each of the network nodes, computer, and other devices, are interconnected with one another as shown in Figure 31. Every node not only sends its signals but also relays data from other nodes. A true mesh topology is the one where every node is connected to every other node in the network. This type of topology is very expensive as there are many redundant connections, thus it is not mostly used in computer networks. It is commonly used in a wireless network. Flooding or routing technique is used in a mesh topology.

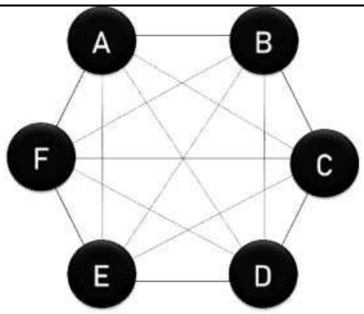


Figure 31 Mesh Topology

Full Mesh Topology

In a full mesh topology, each component is connected to every other component. Even after considering the redundancy factor and cost of this network, its main advantage is that the network traffic can be redirected to other nodes if one of the nodes goes down. Full mesh topology is used only for backbone networks.

Partial Mesh Topology

Here, some of the systems are connected similarly as in mesh topology while the rest of the systems are only connected to 1 or 2 devices. It can be said that in partial mesh, the workstations are 'indirectly' connected to other devices. This one is less costly and also reduces redundancy.

Advantages of Mesh topology

1. Data can be transmitted from different devices simultaneously. This topology can withstand high traffic.
2. Even if one of the components fails there is always an alternative present. So data transfer doesn't get affected.
3. Expansion and modification in topology can be done without disrupting other nodes.

Disadvantages of Mesh Topology

1. There are high chances of redundancy in many of the network connections.
2. The overall cost of this network is way too high as compared to other network topologies.
3. Set-up and maintenance of this topology are very difficult. Even administration of the network is tough.

Tree Topology

Tree Topology integrates the characteristics of Star and Bus Topology. Earlier we saw how in Physical Star network Topology, computers (nodes) are connected by each other through a central hub. And we also saw in Bus Topology, work station devices are connected by the common cable called Bus. In Tree Topology, the number of Star networks are connected using Bus. This main cable seems like the main stem of a tree and other star networks as the branches as shown in Figure 32. It is also called **Expanded Star Topology.** Ethernet protocol is commonly used in this type of topology.

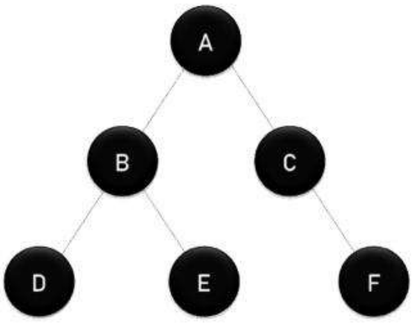


Figure 32 Tree Topology

Advantages of Tree Topology

1. It is an extension of Star and bus Topologies, so in networks where these topologies can't be implemented individually for reasons related to scalability, tree topology is the best alternative.
2. Expansion of Network is possible and easy.
3. Here, we divide the whole network into segments (star networks), which can be easily managed and maintained.
4. Error detection and correction are easy.
5. Each segment is provided with dedicated point-to-point wiring to the central hub.
6. If one segment is damaged, other segments are not affected.

Disadvantages of Tree Topology

1. Because of its basic structure, tree topology, relies heavily on the main bus cable, if it breaks the whole network is crippled.
2. As more and more nodes and segments are added, maintenance becomes difficult.
3. The scalability of the network depends on the type of cable used.

Hybrid Topology

Hybrid, as the name suggests, is a mixture of two different things. Similarly in this type of topology, we integrate two or more different topologies to form a resultant topology that has good points(as well as weaknesses) of all the constituent basic topologies rather than having characteristics of one specific topology. This combination of topologies is done according to the requirements of the organization.

For example:- if there exists a ring topology in one office department while a bus topology in another department, connecting these two will result in Hybrid topology. Remember connecting two similar topologies cannot be termed as Hybrid topology. Star-Ring and Star-Bus networks are the most common examples of the hybrid network as shown in Figure 33.

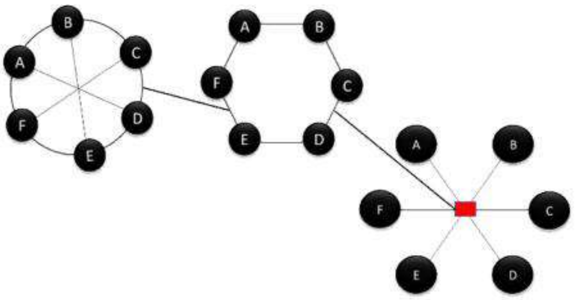


Figure 33 Hybrid Topology

Advantages of Hybrid Network Topology

1. ***Reliable***: Unlike other networks, fault detection and troubleshooting are easy in this type of topology. The part in which fault is detected can be isolated from the rest of the network and required corrective measures can be taken, WITHOUT affecting the functioning of the rest of the network.
2. ***Scalable***: It's easy to increase the size of the network by adding new components, without disturbing existing architecture.
3. ***Flexible***: Hybrid Network can be designed according to the requirements of the organization and by optimizing the available resources. Special care can be given to nodes where traffic is high as well as where chances of fault are high.
4. ***Effective***: Hybrid topology is the combination of two or more topologies, so we can design it in such a way that the strengths of constituent topologies are maximized while their weaknesses are neutralized. For example, we saw Ring Topology has good data

reliability (achieved by the use of tokens) and Star topology has high tolerance capability (as each node is not directly connected to the other but through a central device), so these two can be used effectively in hybrid star-ring topology.

Disadvantages of Hybrid Topology

1. ***The complexity of Design*:** One of the biggest drawbacks of hybrid topology is its design. It's not easy to design this type of architecture and it's a tough job for designers. The configuration and installation process needs to be very efficient.
2. ***Costly Hub*:** The hubs used to connect two distinct networks are very expensive. These hubs are different from usual hubs as they need to be intelligent enough to work with different architectures and should be functional even if a part of the network is down.
3. ***Costly Infrastructure:*** As hybrid architectures are usually larger in scale, they require a lot of cables, cooling systems, sophisticated network devices, etc.

Summary

- A network consists of two or more computers that are linked to share resources (such as printers and CD-ROMs), exchange files, or allow electronic communications. The computers on a network may be linked through cables, telephone lines, radio waves, satellites, or infrared light beams.
- The primary purpose of a computer network is to share resources. The main goal of networking is **Resource sharing.** A second goal is to provide **high reliability** by having alternative sources of supply. Another goal is **saving money.** Another closely related goal is to increase the performance of the system as the workload increases by just adding more processors. With central mainframes, when the system is full, it must be replaced by a larger one, usually at great expense and with even greater disruption to the users. Computer networks provide a powerful communication medium.
- There are two important dimensions for classifying networks — **transmission technology** and **scale**.
- Transmission technology can be classified into two types:
  - Broadcast networks.
  - Point-to-point networks
- **Broadcast networks:** These networks have a single communication channel shared by all the machines on the network.
- **Point-to-point** networks consist of many connections between individual pairs of machines. Multiple routes and intermediate machines may exist between a pair of machines, so routing algorithms play an important role here.
- A collection of interconnected networks is called an **internetwork** or just **Internet**. The **Internet** refers to a specific worldwide Internet that is widely used to connect universities, government offices, companies, and private individuals.
- A network topology is the basic design of a computer network. It details how key network components such as nodes and links are interconnected.
- There are three primary types of network topologies that refer to the physical and logical layout of the network cabling. They are a star, ring, and bus topology.

Keywords

Archive: A computer site advertises and stores a large amount of public domain, shareware software and documentation.

Broadcast Networks: They have a single communication channel, which is shared by all the computers on the network and therefore, any message transmitted by a computer on the network is received by all the computers connected to the channel.

Error Control: The receiving end after the completion of receiving the information must also be capable of dealing with and recognizing corruption.

Local Area Network: A LAN is a form of local (limited distance), shared packet network for computer communications.

Metropolitan Area Network: In MAN, different LANs are connected through a local telephone exchange using one or two cables but not switching elements.

Service Primitives: The primitives enable the service provider to perform some action or report on an action taken by a peer entity.

Wide Area Network: A WAN may be defined as a data communications network that covers a the relatively broad geographic area to connect LANs between different cities with the help of transmission facilities provided by common carriers, such as telephone companies.

Review Questions

Q1: Physical or logical arrangement of a network is

a) Topology
b) Routing
c) Networking
d) Control

Q2: _____________ topology requires a multipoint connection.

a) Star
b) Mesh
c) Ring
d) Bus

Q3: The types of transmission channel or media used for LAN or WAN are

a) Twisted Pair Cables
b) Coaxial Cables
c) Fibre-Optic Cables and Radio Waves
d) All of above

Q4: The _________________ is the physical path by which a message travels from sender to receiver.

a) Protocol
b) Message
c) Transmission medium
d) Sender

Q5: A _____________ is a set of rules governing data communication between two devices

a) Protocol
b) Message
c) Medium
d) Sender

Q6: The ____________ is the device that sends the message

a) Protocol
b) Sender
c) Network
d) Medium

Q7: In a ____________ connection, two and only two devices are connected by a dedicated link.

a) Multipoint
b) Point-to-point
c) (a) and (b)
d) None of the above

Q8: An unauthorized user is a network _____________ issue.

a) Performance
b) Reliability
c) Security
d) All the above

Q9: Which topology requires a centraller controller or a hub?

a) Mesh
b) Star
c) Bus
d) Ring

Q10: Which type of network would use phone lines?

a) Wireless
b) WAN
c) LAN
d) WWAN

Q11: Explain five effective characteristics of data communication.

Q12: Write down applications of computer networks.

Q13: How are computer networks classified? Mention some of the important reasons for the classification of computer networks.

Q14: Diagramtically explain bus topology and ring topology

Q15: Explain the difference between Star, Bus, and Mesh topology.

Q16: Write down the difference between LAN, MAN, and WAN.

Answers

1.	a	2.	d	3.	d	4.	c	5.	a
6.	b	7.	b	8.	c	9.	b	10.	b

Further/Suggested Readings

- Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall.
- Behrouz A. Forouzan and Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies.
- Burton, Bill, *Remote Access for Cisco Networks*, McGraw-Hill Osborne Media, McGraw-Hill Osborne Media.
- Rajneesh Agrawal and Bharat Bhushan Tiwari, *Computer Networks and Internet*, Vikas Publication

- https://www.geeksforgeeks.org/computer-network-tutorials/

Unit 02: Data and Signals

CONTENTS

Objectives

After this lecture, you would be able to:

- understand analog and digital signals
- learn what are the various transmission modes.
- understand what is performance metrics.
- learn the various essential network metrics to monitor.
- Understand the various reasons for transmission impairments
- Understand the concept of protocols.
- Learn the components and functions of protocols

Introduction

One of the major functions of the physical layer is to move data in the form of electromagnetic signals across a transmission medium. Whether you are collecting numericalstatistics from another computer, sending animated pictures from a design workstation,or causing a bell to ring at a distant control center, you are working with the transmissionof**data** across network connections.Generally, the data usable to a person or application are not in a form that can betransmitted over a network. For example, a photograph must first be changed to a formthat transmission media can accept. Transmission media work by conducting energyalong a physical path.

2.1 What are major functions of Physical Layer

The physical layer is responsible for electromagnetic compatibility including electromagnetic spectrum frequency allocation and specification of signal strength, analog bandwidth, etc. The transmission medium may be electrical or optical over optical fiber or a wireless IR communication link.For example, a photograph must first be changed to a form that transmission media can accept.

Analog and Digital Data

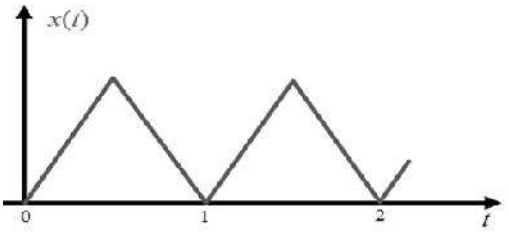
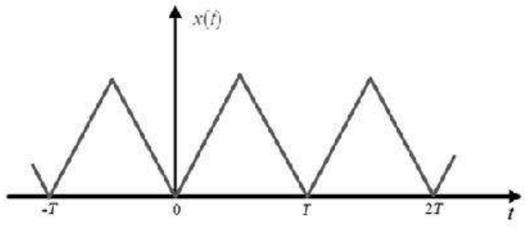
Signals can be classified based on different parameters like the time intervals between which it is being created, or the type of the signal.Both data and the signals that represent them can be either analog or digital in form as shown in Table 1.

Table 1 Analog and Digital Data

Analog Data	Digital Data
The term analog data refers to information that is continuous.	Digital data take on discrete values.

2.2 Signal Classifications

a) Periodic and Non-Periodic Signals

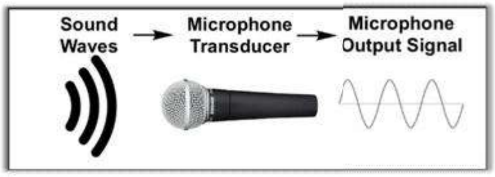
Periodic Signals	Non-Periodic Signals
Completes a pattern and repeats that pattern over subsequent identical periods.	**Changes without exhibiting a pattern**
The completion of one full pattern is called a cycle.	**Any continuous-time signal which is not periodic is called a non-periodic signal.**
	

b) Analog and Digital Data

Examples of Analog data are shown below.

An analog clock that has hour, minute, and second hands gives information in a continuous form.

Sounds made by a human voice take on continuous values.



The example of digital data are shown below:

Digital clock that reports the hours and the minutes will change suddenly from 8:05 to 8:06.

i Analog Signal

- Many levels of intensity
- An infinite number of values.

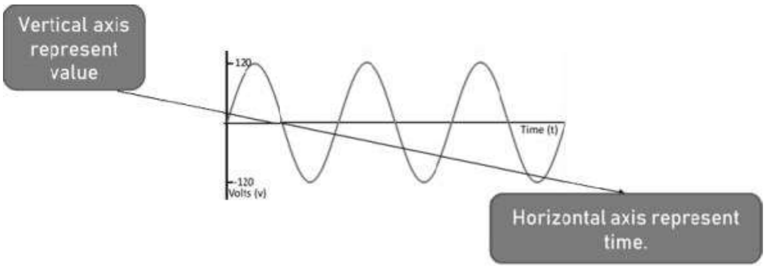

Vertical axis represent value
120
Time (t)
-120
Volts (v)
Horizontal axis represent time.

Figure 1 Analog Signal

ii Digital Signal

- **Limited number of defined values.**
- **Although each value can be any number, it is often as simple as 1 and 0.**

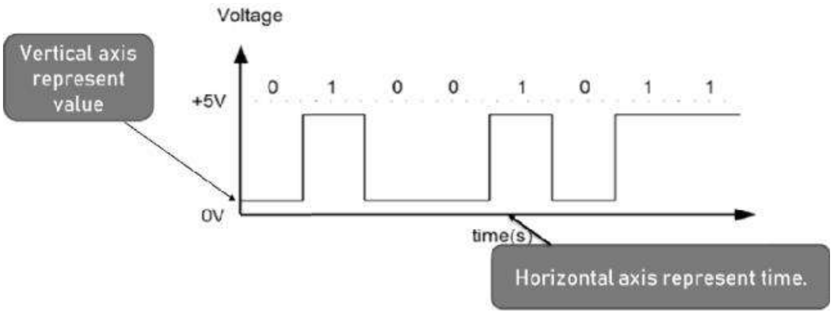

Voltage
Vertical axis represent value
+5V
0 1 0 0 1 0 1 1
0V
time(s)
Horizontal axis represent time.

Figure 2 Digital Signal

Information can also be represented by a digital signal. For example, a 1 can be encoded as a positive voltage and a 0 as zero voltage. A digital signal can have more than two levels. In this case, we can as zero voltage. A digital signal can have more than two levels. In this case, we can send more than 1 bit for each level. The following figure shows two signals, one with two levels and the other with four. In general, if a signal has L levels, each level needs log2L bits.

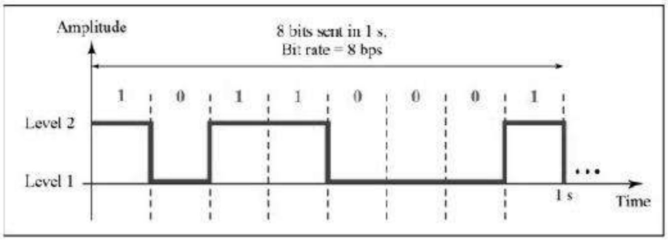


Figure 3 Digital Signal with two levels

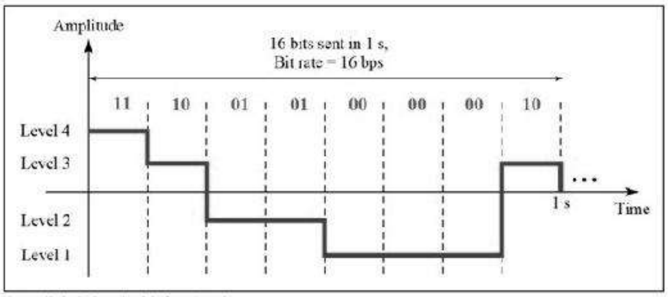


Figure 4 Digital Signal with four levels

Bit Rate:

The bit rate is the number of bits sent in Is, expressed in bits per second (bps).

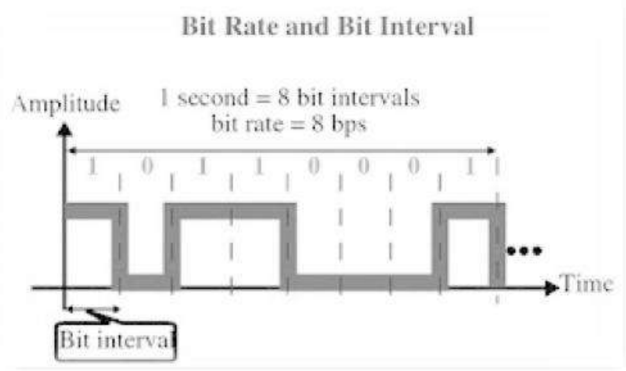


Figure 5 Bit rate and bit interval

The Bit Rate for the above diagram is 8bps and 16bps.

Bit Length:

The bit length is the distance one bit occupies on the transmission medium.

Bit length =propagation speed x bit duration

Baud Rate

Baud Rate is the number of signal unit transmitted per second.

Thus Baud Rate is always less than or equal to bit rate. Baud rate is number of symbols per second.

DifferenceBetweenBit Rate and Baud Rate

Bit Rate	Baud Rate
Bit rate is also defined as per second travel number of bits.	Baud rate is also defined as per second number of changes in signal.
Bit rate emphasized on computer efficiency.	While baud rate emphasized on data transmission.
Bit rate is not used to decide the requirement of bandwidth for transmission of signal.	While baud rate is used to decide the requirement of bandwidth for transmission of signal.

Bit Length

The bit length is the distance one bit occupies on the transmission medium.

Bit length =propagation speed x bit duration

Bit Interval

Data can be represented by a digital signal. For Example a 1 can be encoded as a positivevoltage and a 0 can be encoded as a zero voltage.

2.3 Transmission Mode

Transmission mode refers to the mechanism of transferring of data between two devices connected over a network.It is also called Communication Mode.These modes direct the direction of flow of information.

Modes of Transmission

Transmission mode means transferring of data between two devices. It is also known as communication mode. Buses and networks are designed to allow communication to occur between individual devices that are interconnected. There are three types of transmission mode:-

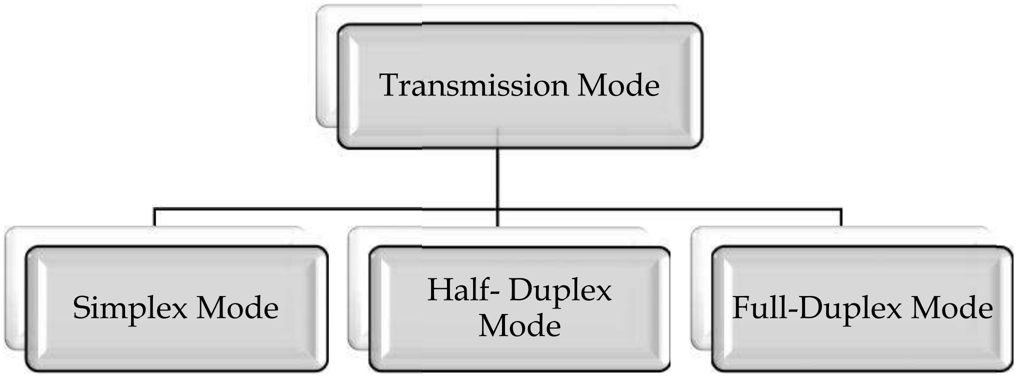


Figure 6 Transmission Modes

These are explained as following below.

a) Simplex Mode –

In Simplex mode, the communication is unidirectional, as on a one-way street. Only one of the two devices on a link can transmit, the other can only receive. The simplex mode can use the entire capacity of the channel to send data in one direction. Example: Keyboard and traditional monitors. The keyboard can only introduce input, the monitor can only give the output.

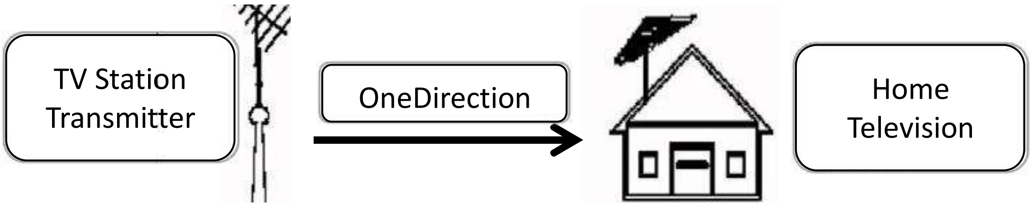


Figure 7 Simplex Mode

b) Half-Duplex Mode –

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

Channel capacity=Bandwidth * Propagation Delay

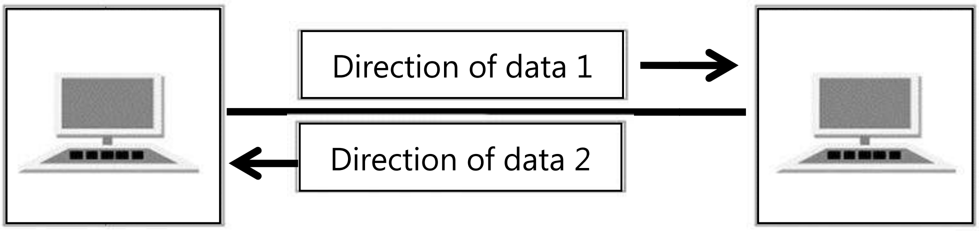


Figure 8 Half Duplex Transmission

c) Full Duplex Mode

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

Either the link must contain two physically separate transmission paths, one for sending and other for receiving.

Or the capacity is divided between signals travelling in both directions.

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

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

Channel Capacity=2* Bandwidth*propagation Delay

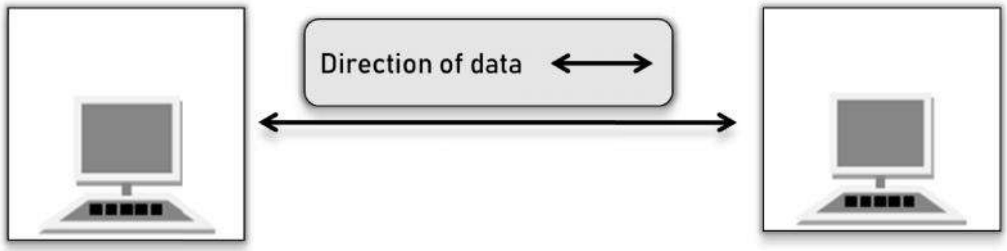


Figure 9 Full Duplex Transmission

Performance Metrics

Network performance can be affected by a number of different factors.It's important for companies to know which network performance metrics are important to examine.However, depending on the specific issues that plague your network, not every metric is going to be important for you to look at.Despite this, there are some metrics that are essential for any businesses to consider.

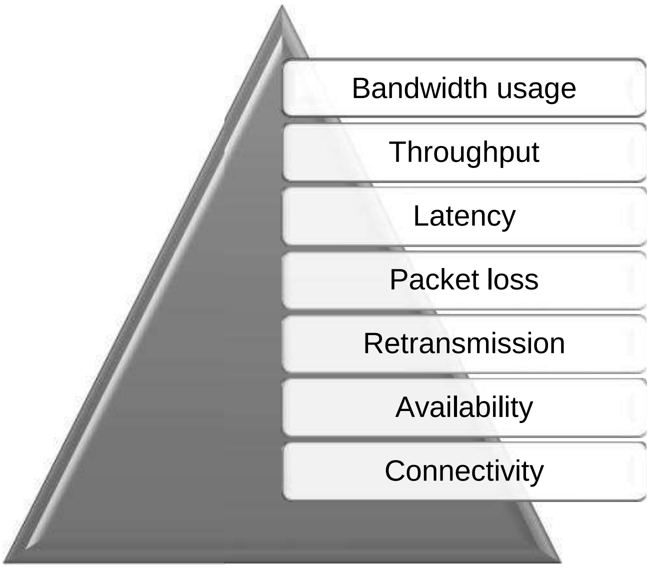


Figure 10: Seven Essential Network Performance Metrics

Bandwidth usage

Bandwidth is the maximum data transmission rate possible on a network. For optimal network operations, you want to get as close to your maximum bandwidth as possible without reaching critical levels. This indicates that your network is sending as much data as it can within a period of time but isn't being overloaded. An NPM can monitor how much bandwidth is currently being used on a network, as well as how much bandwidth is typically used during daily operations. The solution can also alert you when your network is using too much bandwidth.

Throughput

Throughput measures your network's actual data transmission rate, which can vary wildly through different areas of your network. While your network's bandwidth measures the theoretical limit of data transfer, throughput tells you how much data is actually being sent. Specifically, throughput measures the percentage of data packets that are successfully being sent; a low throughput means there are a lot of failed or dropped packets that need to be sent again.

Latency

Latency is the delay that happens between a node or device requesting data and when that data is finished being delivered. This delay can happen for a variety of reasons, but whatever the cause, your NPM solution can track any delays and log them. Consistent delays or odd spikes in delay time indicate a major performance issue; however, because delays can often be undetectable to the human eye, you need a monitoring tool to keep an eye on any delays that happen.

Packet loss

Packet loss examines how many data packets are dropped during data transmissions on your network. The more data packets that are lost, the longer it takes for a data request to be fulfilled. Your IT team should know how many packets are being dropped on average across your infrastructure. A network's Transmission Control Protocol (TCP) interprets when packets are dropped and takes steps to ensure that data packets can still be transmitted; your network team should monitor this system to make sure it's working.

Retransmission

When packets are lost, the network needs to retransmit it to complete a data request. This retransmission rate lets your enterprise know how often packets are being dropped, which is an indication of congestion on your network. You can analyze retransmission delay, or the time it takes for a dropped packet to be retransmitted, to understand how long it takes your network to recover from packet loss.

Availability

Network availability, also known as uptime, simply measures whether or not the network is currently operational. You can never guarantee 100% availability, but you want to be aware of any downtime that happens on your network that you weren't expecting. It's important to be alerted when the network goes down, which network monitoring tools will provide for you. However, you should also be able to discover your actual uptime percentage and how often your network goes down.

Connectivity

Connectivity refers to whether the connections between the nodes on your network are working properly. If there is an improper or malfunctioning connection on your network, it can be a major hurdle for your company. Ideally, every connection should always be operating at peak levels. However, performance issues like malware can target specific nodes or connections to affect performance in that specific area of the network.

2.4 Transmission Impairments

Signals travel through transmission media, which are not perfect. The imperfection causes signal impairment. This means that the signal at the beginning of the medium is not the same as the signal at the end of the medium. What is sent is not what is received.

Reasons for Impairments

Three causes of impairment are attenuation, distortion, and noise.

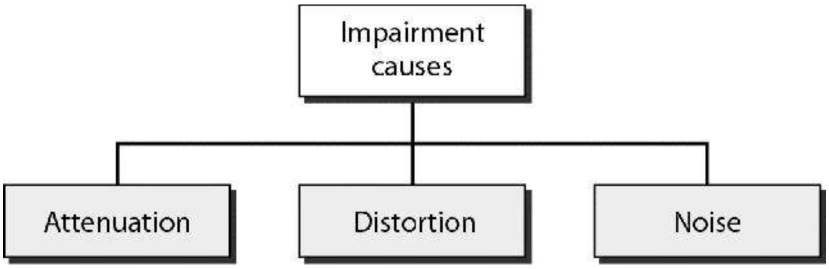


Figure 11 Causes of Impairments

a) Attenuation

It means loss of energy. The strength of signal decreases with increasing distance which causes loss of energy in overcoming resistance of medium. This is also known as attenuated signal. Amplifiers are used to amplify the attenuated signal which gives the original signal back and compensate for this loss.

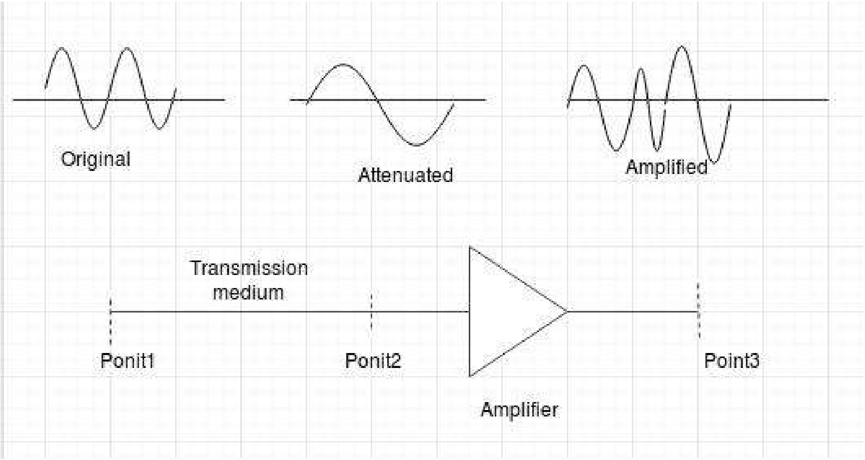


Figure 12 Attenuation

Attenuation is measured in **decibels(dB)**. It measures the relative strengths of two signals or one signal at two different point.

How is Attenuation Measured

To show the loss or gain of energy the unit "decibel" is used.

Attenuation dB = 10log10P2/P1

P1 - input signal

P2 - output signal

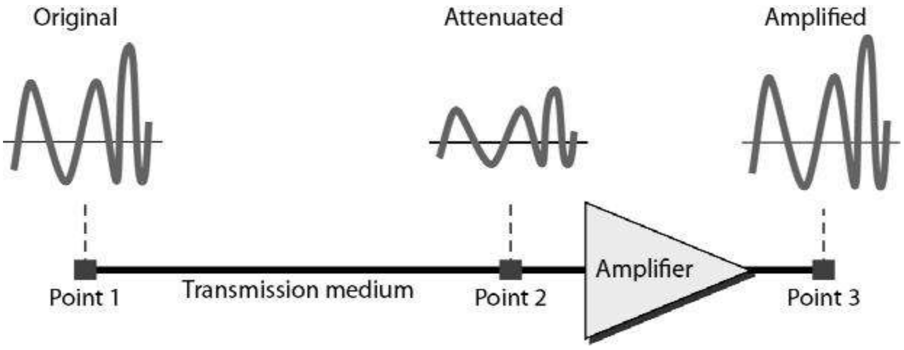


Figure 13 Attenuated Signal

Suppose a signal travels through a transmission medium and its power is reduced to one-half. This means that P2 is (1/2)P1. In this case, the attenuation (loss of power) can be calculated as

$$10\log_{10}\frac{P2/P1}{P1} = 10\log_{10}\frac{0.5\ P1}{P1} = 10\log_{10}0.5 = 10(-0.3) = -3\text{ dB}$$

A loss of 3 dB (–3 dB) is equivalent to losing one-half the power

A signal travels through an amplifier, and its power is increased 10 times. This means that P2 = 10P1. In this case, the amplification (gain of power) can be calculated as

$$10\log_{10}\frac{P2}{P1} = 10\log_{10}\frac{10P1}{P1}$$

$$\mathbf{= 10\log_{10}10 = 10(1) = 10\text{ dB}}$$

One reason that engineers use the decibel to measure the changes in the strength of a signal is that decibel numbers can be added (or subtracted) when we are measuring several points (cascading) instead of just two. In Figure 3.27 a signal travels from point 1 to point 4. In this case, the decibel value can be calculated as

dB = -3 + 7 – 3 = +1

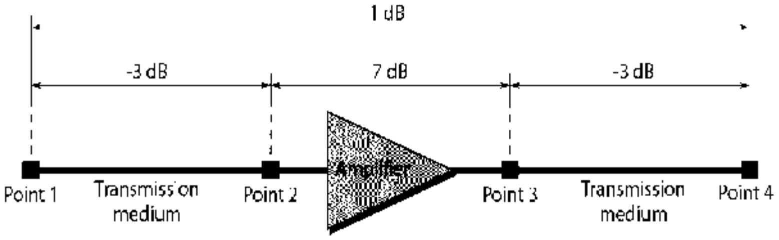


Figure 14 Attenuated Signal over Four Points

b) Distortion

It means changes in the form or shape of the signal. This is generally seen in composite signals made up with different frequencies. Each frequency component has its own propagation speed travelling through a medium. And thats why it delay in arriving at the final destination Every component arrive at different time which leads to distortion. Therefore, they have different phases at receiver end from what they had at senders end.

Example

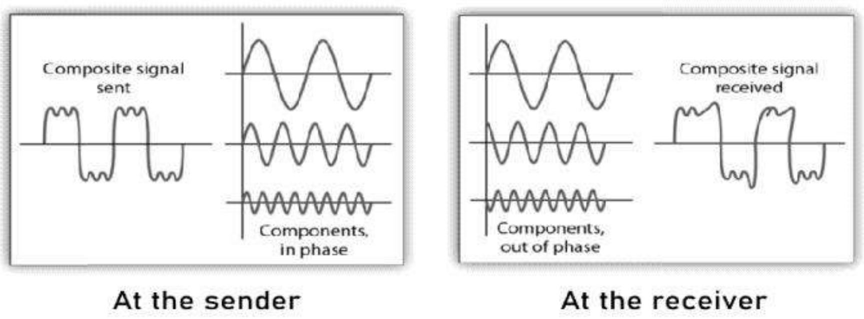


Figure 15 Distortion

c) Noise

The random or unwanted signal that mixes up with the original signal is called noise. There are several types of noise such as induced noise, crosstalk noise, thermal noise and impulse noise which may corrupt the signal.

Induced noise comes from sources such as motors and appliances. These devices act as sending antenna and transmission medium act as receiving antenna. Thermal noise is movement of electrons in wire which creates an extra signal. Crosstalk noise is when one wire affects the other wire. Impulse noise is a signal with high energy that comes from lightning or power lines.

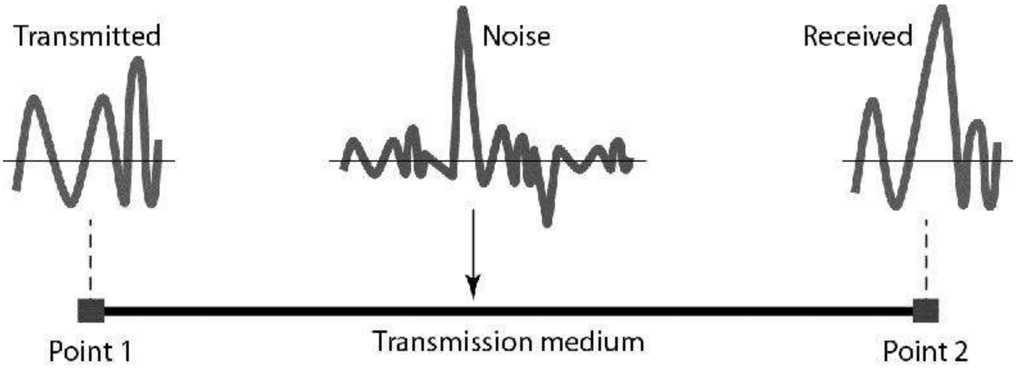


Figure 16 Noise

Signal to Noise Ratio (SNR)

To measure the quality of a system the SNR is often used. It indicates the strength of the signal wrt the noise power in the system.

$$SNR = \frac{Average\ signal\ power}{Average\ noise\ power}$$

It is usually given in dB and referred to as $SNR_{dB.}$

A high SNR means the signal is less corrupted by noise.

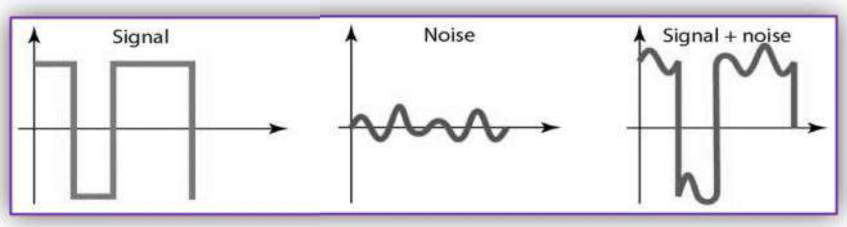


Figure 17 High SNR Signal

- **A low SNR means the signal is more corrupted by noise.**

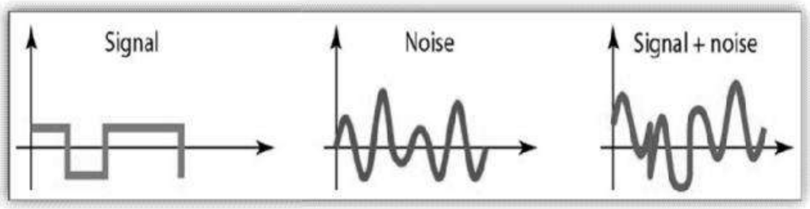


Figure 18 Low SNR Signal

2.5 Protocols

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.Cooperative action is necessary. It should be noted that computer networking is not only to exchange bytes. It is a huge system with several utilities and functions. For exampleerror detection, encryption, routing etc.For proper communication, entities in different systems must speak the same language. There must be mutually acceptable conventions and rules about the content, timing and underlying mechanisms. Those conventions and associated rules are referred as “PROTOCOLS”.

Protocol Architecture

The task of data transfer is broken up into some modules. It is important to understand as to why it is done and how do these modules interact?For example, file transfer could use three modules: File transfer application, communication service module and network access module. Let us see a real-worldexample of the Protocol Architecture.

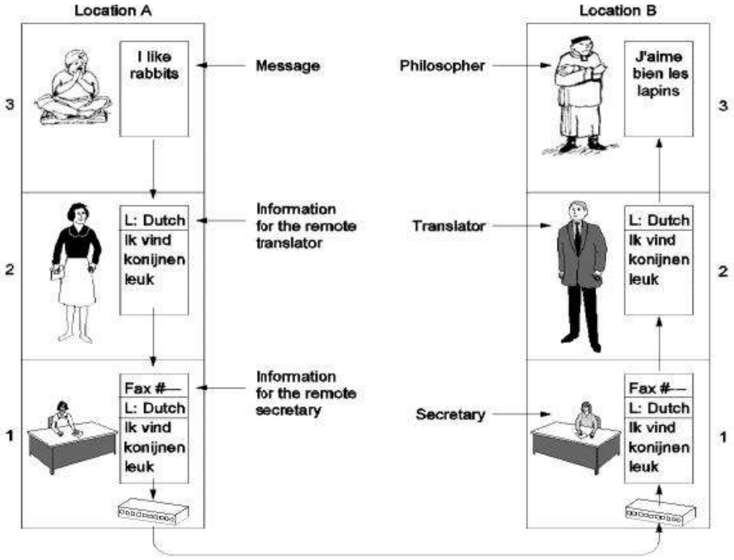


Figure 19Philosopher-Translator-Secretary Architecture

Let us focus on some of the issues: like the peer-to-peer protocols are independent of each otherfor example, secretaries may change the comm. medium to emailor the translators may agree on using another common language. Note that each layer adds a header

2.6 Network Standards

Networking standards define the rules for data communications that are needed for interoperability of networking technologies and processes.Standards help in creating and maintaining open markets and allow different vendors to compete on the basis of the quality of their products while being compatible with existing market products.

Types of Standards

Standards are of two typesi.e.De facto and De jure.

De facto standardsare the standards that are followed without any formal plan or approval by any organization. They have come into existence due to traditions or facts. For example, the HTTP had started as a de facto standard.

De jure standardsare the standards which have been adopted through legislation by any officially recognized standards organization. Most of the communication standards that are used today are de jure standards.

Standards Organizations

Some of the noted standards organizations are

1. International Standards Organization (ISO)
2. International Telecommunication Union (ITU)
3. Institute of Electronics and Electrical Engineers (IEEE)
4. American National Standards Institute (ANSI)
5. Internet Research Task Force (IRTF)
6. Electronic Industries Association (EIA)
7. World Wide Web Consortium (W3C)

1. **International Standards Organization (ISO)-** The International Organization for standardization widely known as ISO, is an international standard-setting body composed of representatives from various national standards organizations. Founded on February 23, 1947, the organization promulgates worldwide proprietary industrial and commercial standards. It has it's headquarters in Geneva, Switzerland. While ISO defines itself as a non-governmental organization, its ability to set standards that often-become law, either through treaties or national standards, makes it more powerful than most non-governmental organizations. In practice, ISO acts as a consortium with strong links to governments. ISO is an international standard-setting body composed of representatives from various national standards organizations the organization promulgates worldwide proprietary industrial and commercial standards.ISO's main products are the International Standards. ISO also publishes Technical Reports, Technical Specifications, Publicly Available Specifications, Technical Corrigenda, and Guides.

2. **International Telecommunication Union (ITU) -** The International Telecommunication Union is the specialized agency of the United Nations which is responsible for information and communication technologies. ITU coordinates the shared global use of the radio spectrum, promotes international cooperation in assigning satellite orbits, works to improve telecommunication infrastructure in the developing world and establishes worldwide standards.ITU also organizes worldwide and regional exhibitions and forums, such as ITU TELECOM WORLD, bringing together representatives of government and the telecommunications and ICT industry to exchange ideas, knowledge and technology.
3. **Institute of Electronics and Electrical Engineers (IEEE) -**IEEE's Constitution defines the purposes of the organization as "scientific and educational, directed toward the advancement

of the theory and practice of Electrical, Electronics, Communications and Computer Engineering, as well as Computer Science, the allied branches of engineering and the related arts and sciences." The IEEE is Not-for-Profit Corporation. It was formed in 1963 by the merger of the Institute of Radio Engineers (IRE, founded 1912) and the American Institute of Electrical Engineers (AIEE, founded 1884). It has more than 400,000 members in more than 160 countries, 45% outside the United States.It is also a leading developer of industrial standards having developed over 900 active industry standards in a broad range of disciplines, including electric power and energy, biomedical technology and health care, information technology, information assurance, telecommunications, consumer electronics, transportation, aerospace, and nanotechnology. IEEE develops and participates in educational activities such as accreditation of electrical engineering programs in institutes of higher learning. IEEE serves as a major publisher of scientific journals and a conference organizer. IEEE is one of the leading standards-making organizations in the world. IEEE performs its standards making and maintaining functions through the IEEE Standards Association (IEEE-SA). One of the more notable IEEE standards is the IEEE 802 LAN/MAN group of standards which includes the IEEE 802.3 Ethernet standard and the IEEE 802.11 Wireless Networking standard.

4. **American National Standards Institute (ANSI) -**Though ANSI itself does not develop standards, the Institute oversees the development and use of standards by accrediting the procedures of standards developing organizations. ANSI accreditation signifies that the procedures used by standards developing organizations meet the Institute's requirements for openness, balance, consensus, and due process.ANSI was originally formed in 1918, when five engineering societies and three government agencies founded the American Engineering Standards Committee (AESC).In 1928, the AESC became the American Standards Association (ASA). In 1966, the ASA was reorganized and became the United States of America Standards Institute (USASI). The present name was adopted in 1969.ANSI also designates specific standards as American National Standards, or ANS, when the Institute determines that the standards were developed in an environment that is equitable, accessible and responsive to the requirements of various stakeholders.The American National Standards process involves:
   - consensus by a group that is open to representatives from all interested parties
   - broad-based public review and comment on draft standards
   - consideration of and response to comments
   - incorporation of submitted changes that meet the same consensus requirements into a draft standard
   - availability of an appeal by any participant alleging that these principles were not respected during the standards-development process.

5. **Internet Research Task Force (IRTF) -**The Internet Research Task Force (IRTF) promotes research of importance to the evolution of the Internet by creating focused, long-term Research Groups working on topics related to Internet protocols, applications, architecture, and technology.The IRTF is a composed of several focused and long-term Research Groups.Research Groups have the stable long-term membership needed to promote the development of research collaboration and teamwork in exploring research issues. Participation is by individual contributors, rather than by representatives of organizations.The IRTF is managed by the IRTF Chair in consultation with the Internet Research Steering Group (IRSG). The IRSG membership includes the IRTF Chair, the chairs of the various Research Groups, and other individuals ("members at large") from the research community selected by the IRTF Chair.The IRTF is managed by the IRTF Chair in consultation with the Internet Research Steering Group (IRSG). The IRSG membership includes the IRTF Chair, the chairs of the various Research Groups, and other individuals ("members at large") from the research community selected by the IRTF Chair.

6. **World Wide Web Consortium (W3C) -**The World Wide Web Consortium (W3C) is the main international standards organization for World Wide Web (abbreviated WWW or W3).Founded and headed by Tim Berners-Lee, the consortium is made up of member organizations which maintain full-time staff for the purpose of working together in the development of standards for the World Wide Web. W3C also engages in education and outreach, develops software and serves as an open forum for discussion about the Web.W3C was created to ensure compatibility and agreement among industry members in the adoption of new standards. Prior to its creation, incompatible versions of HTML were offered by different vendors, increasing the potential for inconsistency between web pages. The

consortium was created to get all those vendors to agree on a set of core principles and components which would be supported by everyone.

Summary

Having discussed the various transmission medias and the different type of signals that they deal with, a proper transmission mode has to be chosen for the type of communication we are undergoing. To ensure the effective and error-free transmission different essential network performance metrics must be considered to make necessary changes in the network. The different transmission impairments need to be understood by an organisation and should be effectively managed to ensure effective data communication

Keywords

Archive: A computer site advertises and stores a large amount of public domain, shareware software and documentation.

Broadcast Networks: They have a single communication channel, which is shared by all the computers on the network and therefore, any message transmitted by a computer on the network is received by all the computers connected to the channel.

Bit Length: It is the distance one bit occupies on the transmission medium.

Baud Rate is the number of signal unit transmitted per second. It is always less than or equal to bit rate.

Distortion: It means changes in the form or shape of the signal. This is generally seen in composite signals made up with different frequencies. Each frequency component has its own propagation speed travelling through a medium. And that's why it delays in arriving at the final destination

Self-Assessment

1. Telegraph Signals are the examples of
   a. Digital Signals
   b. Analog Signals
   c. Impulse Signals
   d. Pulse Train
2. The example of an analog to analog conversion is
   a. Radio
   b. Video
   c. Television
   d. Internet
3. Analog to Analog conversion can be accomplished in
   a. One way
   b. Two ways
   c. Three ways
   d. Four ways
4. In the ____________ transmission mode, communication is unidirectional
   a. Simplex
   b. Half-duplex
   c. Full-duplex
   d. Hybrid
5. The _________ is an example of simplex device.
   a. Repeater
   b. Tap
   c. Walkie-talkie

d. Keyboard

6. In the _________ transmission mode, each station can transmit , but not at the same time

    a. Simplex

    b. Half-duplex

    c. Full-duplex

    d. B and c

7. In the ___________ transmission mode, both stations can transmit and receive at the same time.

    a. Simplex

    b. Half-duplex

    c. Full-duplex

    d. B and c

8. _______ refers to the direction of signal flow between two linked devices

    a. Line configuration

    b. Topology

    c. Transmission mode

    d. Line discipline.

9. The term that refres to loss of strength of a signal is called

    a. Attenuation

    b. Distortion

    c. Noise

    d. Impairments

10. A transmission media can have signal impairment because of

    a. Noise

    b. Attenuation

    c. Distortion

    d. All of above

11. _______ is a type of transmission impairment in which the signal loses strength due to the different propagation speeds of each frequency that makes up the signal.

    a. Attenuation

    b. Noise

    c. Distortion

    d. Decibel

12. ________ is a type of transmission impairment in which an outside source such as crosstalk corrupts a signal.

    a. Attenuation

    b. Noise

    c. Distortion

    d. Decibel

Answers for Self Assessment

1.	A	2.	A	3.	C	4.	A	5.	D
6.	B	7.	B	8.	C	9.	A	10.	D
11.	C	12.	B						

Review Questions

Q1: Difference between bit rate and baud rate.

Q2: Explain difference types of modes of communication.

Q3: Explain difference between full duplex, half duplex and simplex along with examples.

Q4: Write down the standards of organization.

Q5: What do you understand by signals. Explain types of signals also.

Q6: There are always three causes of impairment. Explain them in detail.

Q7: Difference between periodic and non-periodic signals.

Q8: Write note on

a) Bit rate

b) Baud rate

c) Bit length

d) Bit Interval.

Q9: What do you understand by Defacto and DeJure.

Q10: What do you understand by protocol. Explain protocol architecture.

Further Readings

Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall

Behrouz A. Forouzan and Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies

Burton, Bill, *Remote Access for Cisco Networks*, McGraw-Hill Osborne Media McGraw-Hill Osborne Media

Rajneesh Agrawal and Bharat Bhushan Tiwari, *Computer Networks and Internet*, Vikas Publication

UNIT 03: Digital and Analog Transmission

CONTENTS

Objectives

- represent digital data by using digital signals.
- learning schemes to transmit data digitally.
- learn the digital to analog conversions.
- understand the various modulation techniques.

Digital Transmission

A computer network is designed to send information from one point to another. This information needs to be converted to either a digital signal or an analog signal for transmission. In this chapter, we discuss the first choice, conversion to digital signals. First, we discuss digital-to-digital conversion techniques, methods which convert digital data to digital signals. Second, we discuss analog-to-digital conversion techniques, methods which change an analog signal to a digital signaL Finally, we discuss transmission modes

3.1 Digital-to Digital Transmission

We said that data can be either digital or analog. We also said that signals that represent data can also be digital or analog. In this section, we see how we can represent digital data by using digital signals. The conversion involves three techniques:

- Line coding,
- Block coding,
- Scrambling. Line coding is always needed~ block coding and scrambling mayor may not be needed.

Line Coding

Line coding is the process of converting digital data to digital signals. We assume that data, in the form of text, numbers, graphical images, audio, or video, are stored in computer memory as sequences of bits . Line coding converts a sequence of bits to a digital signal. At the sender, digital data are encoded into a digital signal; at the receiver, the digital data are recreated by decoding the digital signal. Figure 1 shows the process.

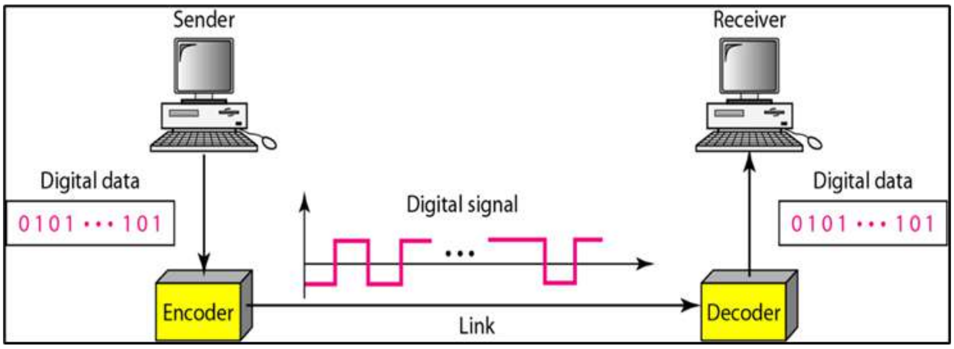


Figure 1 Digital to Digital Transmission

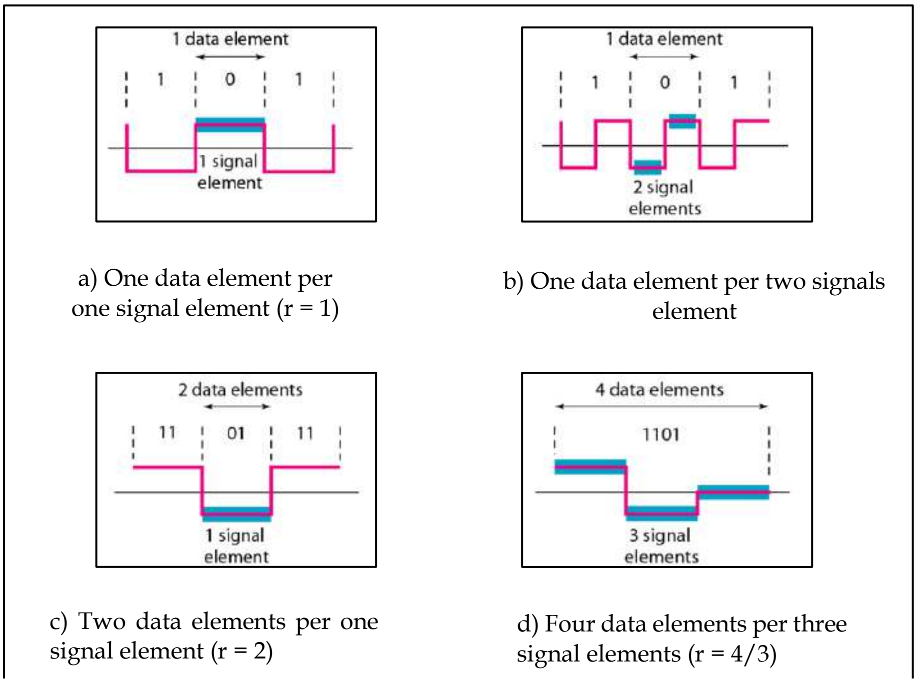
Characteristics

Before discussing different line coding schemes, we address their common characteristics.

Signal Element Versus Data Element

Let us distinguish between a data element and a signal element. In data communications, our goal is to send data elements. A data element is the smallest entity that can represent a piece of information: this is the bit. In digital data communications, a signal element carries data elements. A signal element is the shortest unit (timewise) of a digital signal. In other words, data elements are what we need to send; signal elements are what we can send. Data elements are being carried; signal elements are the carriers.

We define a ratio *r* which is the number of data elements carried by each signal element. Figure 2 shows several situations with different values of *r.*



a) One data element per one signal element (r = 1)

b) One data element per two signals element

c) Two data elements per one signal element (r = 2)

d) Four data elements per three signal elements (r = 4/3)

*Figure SEQ Figure * ARABIC 2 Signal element vs Data element*

In part a of the figure, one data element is carried by one signal element *(r* = 1). In part b of the figure, we need two signal elements (two transitions) to carry each data element *(r* = 1). We will see later that the extra signal element is needed to guarantee synchronization. In part c of the figure, a signal element carries two data elements *(r* = 2). Finally, in part d, a group of 4 bits is being carried by a group of three signal elements *(r* = 4/3). For every line coding scheme we discuss, we will give the value of *r.*

An analogy may help here. Suppose each data element is a person who needs to be carried from one place to another. We can think of a signal element as a vehicle that can carry people. When *r* =1, it means each person is driving a vehicle. When $r > 1$, it means more than one person is travelling in a vehicle (a carpool, for example). We can also have the case where one person is driving a car and a trailer ($r = 1/2$).

Data Rate V/s Signal Rate

The data rate defines the number of data elements (bits) sent in Is. The unit is bits per second (bps). The signal rate is the number of signal elements sent in Is. The unit is the baud. There are several common terminologies used in the literature. The data rate is sometimes called the bit rate; the signal rate is sometimes called the pulse rate, the modulation rate, or the baud rate.

One goal in data communications is to increase the data rate while decreasing the signal rate. Increasing the data rate increases the speed of transmission; decreasing the signal rate decreases the bandwidth requirement. In our vehicle-people analogy, we need to carry more people in fewer vehicles to prevent traffic jams. We have a limited *bandwidth* in our transportation system. The difference between datarate and signal rate is shown in Table 1.

Table 1 Data rate V/s Signal rate

Data rate	Signal rate
The number of data elements (bits) sent in 1s.	The number of signal elements sent in 1s.
The unit is bits per second (bps).	The unit is the baud.
Also known as bit rate.	Also known as pulse rate.

Baseline Wandering

In decoding a digital signal, the receiver calculates a running average of the received signal power. This average is called the *baseline.* The incoming signal power is evaluated against this baseline to determine the value of the data element. A long string of Os or 1s can cause a drift in the baseline (baseline wandering) and make it difficult for the receiver to decode correctly. A good line coding scheme needs to prevent baseline wandering.

DC Components

When the voltage level in a digital signal is constant for a while, the spectrum creates very low frequencies (results of Fourier analysis). These frequencies around zero, called DC (direct-current) *components,* present problems for a system that cannot pass low frequencies or a system that uses electrical coupling (via a transformer). For example, a telephone line cannot pass frequencies below 200 Hz. Also a long-distance link may use one or more transformers to isolate different parts of the line electrically. For these systems, we need a scheme with no DC component.

Self-synchronization

To correctly interpret the signals received from the sender, the receiver's bit intervals must correspond exactly to the sender's bit intervals. If the receiver clock is faster or slower, the bit intervals are not matched and the receiver might misinterpret the signals. A self-synchronizing digital signal includes timing information in the data being transmitted. This can be achieved if there are transitions in the signal that alert the receiver to the beginning, middle, or end of the pulse. If the receiver's clock is out of synchronization, these points can reset the clock.

Built-in Error Detection

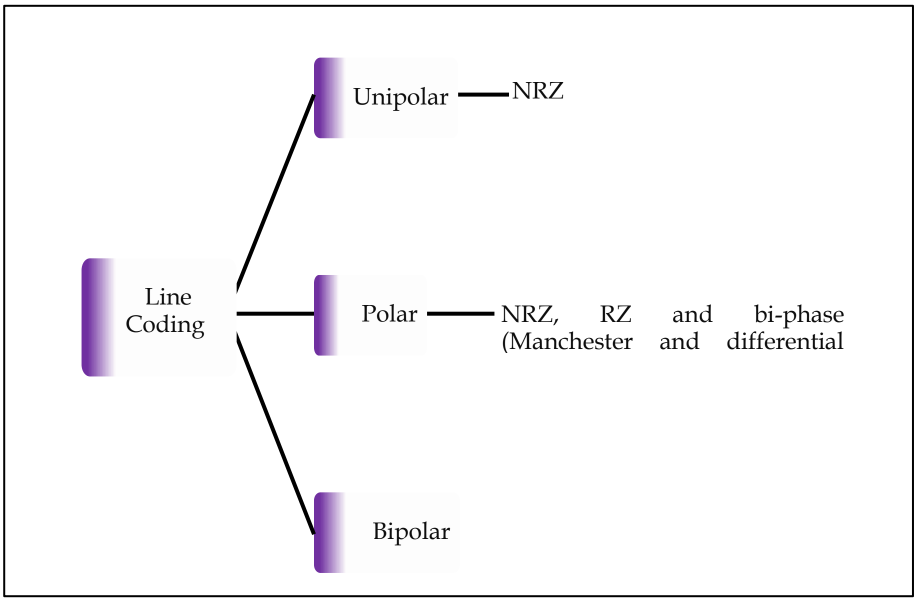
It is desirable to have a built-in error-detecting capability in the generated code to detect some of or all the errors that occurred during transmission. Some encoding schemes that we will discuss have this capability to some extent. Immunity to Noise and Interference Another desirable code characteristic is a code that is immune to noise and other interferences. Some encoding schemes that we will discuss have this capability.

Complexity

A complex scheme is more costly to implement than a simple one. For example, a scheme that uses four signal levels is more difficult to interpret than one that uses only two levels.

3.2 Line Coding Schemes

Line Coding schemes are divided into three categories as shown in Figure 3.



Unipolar

In a unipolar scheme, all the signal levels are on one side of the time axis, either above or below.

NRZ (Non-Return-to-Zero)

Traditionally, a unipolar scheme was designed as a non-return-to-zero (NRZ) scheme in which the positive voltage defines bit I and the zero voltage defines bit O. It is called NRZ because the signal does not return to zero at the middle of the bit. Figure 4 shows a unipolar NRZ scheme.

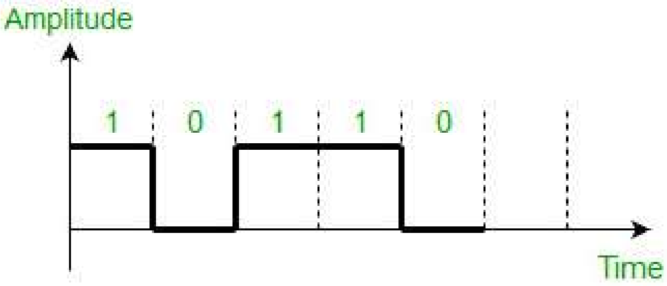


Figure 4 Unipolar NRZ Scheme

NRZ-L and NRZ-I – These are somewhat similar to unipolar NRZ scheme but here we use two levels of amplitude (voltages). For NRZ-L(NRZ-Level), the level of the voltage determines the value of the bit, typically binary 1 maps to logic-level high, and binary 0 maps to logic-level low, and for NRZ-I(NRZ-Invert), two-level signal has a transition at a boundary if the next bit that we are going to transmit is a logical 1, and does not have a transition if the next bit that we are going to transmit is a logical 0.

Note – For NRZ-I we are assuming in the example that previous signal before starting of data set "01001110" was positive. Therefore, there is no transition at the beginning and first bit "0" in current data set "01001110" is starting from +V. Example: Data = 01001110 is shown in Figure 5.

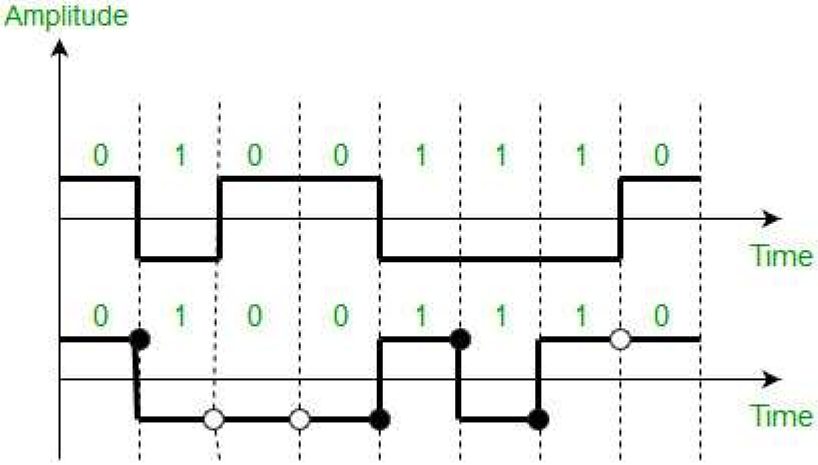


Figure 5 NRZ-L and NRZ-I

Comparison between NRZ-L and NRZ-I: Baseline wandering is a problem for both of them, but for NRZ-L it is twice as bad as compared to NRZ-I. This is because of transition at the boundary for NRZ-I (if the next bit that we are going to transmit is a logical 1). Similarly self-synchronization problem is similar in both for long sequence of 0's, but for long sequence of 1's it is more severe in NRZ-L.

Return to zero (RZ)

One solution to NRZ problem is the RZ scheme, which uses three values positive, negative, and zero. In this scheme signal goes to 0 in the middle of each bit.
Note – The logic we are using here to represent data is that for bit 1 half of the signal is represented by +V and half by zero voltage and for bit 0 half of the signal is represented by -V and half by zero voltage. Example: Data = 01001 as shown in Figure 6.

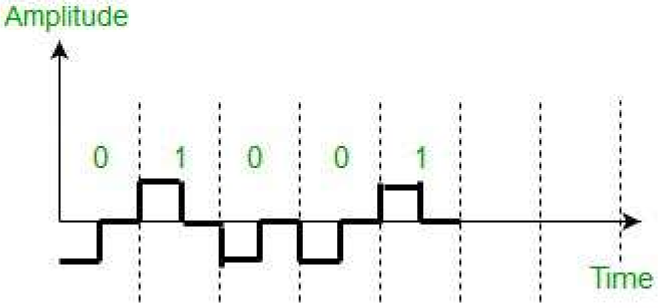


Figure 6 Return to Zero

Main disadvantage of RZ encoding is that it requires greater bandwidth. Another problem is the complexity as it uses three levels of voltage. As a result of all these deficiencies, this scheme is not used today. Instead, it has been replaced by the better-performing Manchester and differential Manchester schemes.

Biphase (Manchester and Differential Manchester)

- **Manchester encoding** is somewhat combination of the RZ (transition at the middle of the bit) and NRZ-L schemes. The duration of the bit is divided into two halves. The voltage remains at one level during the first half and moves to the other level in the second half. The transition at the middle of the bit provides synchronization.
- **Differential Manchester** is somewhat combination of the RZ and NRZ-I schemes. There is always a transition at the middle of the bit but the bit values are determined at the beginning of the bit. If the next bit is 0, there is a transition, if the next bit is 1, there is no transition.

The logic we are using here to represent data using Manchester is that for bit 1 there is transition form -V to +V volts in the middle of the bit and for bit 0 there is transition from +V to -V volts in the middle of the bit. For differential Manchester we are assuming in the example that previous signal before starting of data set "010011" was positive. Therefore there is transition at the beginning and first bit "0" in current data set "010011" is starting from -V. Example: Data = 010011.

The Manchester scheme overcomes several problems associated with NRZ-L, and differential Manchester overcomes several problems associated with NRZ-I as there is no baseline wandering and no DC component because each bit has a positive and negative voltage contribution.

Only limitation is that the minimum bandwidth of Manchester and differential Manchester is twice that of NRZ.

Bipolar schemes

In this scheme there are three voltage levels positive, negative, and zero as shown in *Figure 7*. The voltage level for one data element is at zero, while the voltage level for the other element alternates between positive and negative.

- ***Alternate Mark Inversion (AMI)***

  A neutral zero voltage represents binary 0. Binary 1's are represented by alternating positive and negative voltages.

- ***Pseudoternary***

  Bit 1 is encoded as a zero voltage and the bit 0 is encoded as alternating positive and negative voltages i.e., opposite of AMI scheme. Example: Data = 010010.

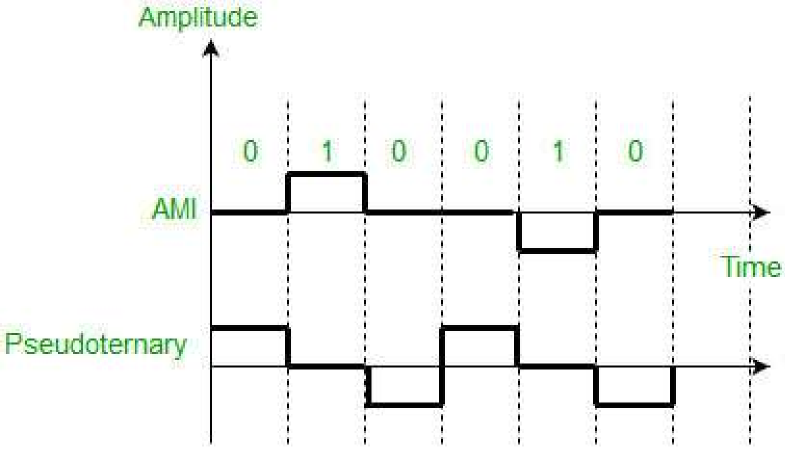


Figure 7: AMI and Pseudoternary

The bipolar scheme is an alternative to NRZ. This scheme has the same signal rate as NRZ, but there is no DC component as one bit is represented by voltage zero and other alternates every time .

Data Communication System

In data communication system, digital and analog communication together plays a very important and integrated role irrespective of many advantages of digital communications over analog. *Figure 8* and *Figure 9* shows the integrated role of digital and analog communication to complete data communication system. *Figure 9* shows that the link between modems is modulated analog signal created by the modem. The communication from PC to modem is consisted of binary signal

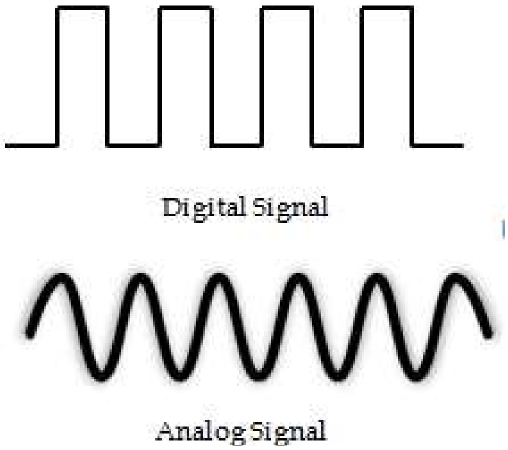


Figure SEQ Figure * ARABIC 8: Digital and Analog Signal

where as the communication between Central Telephone Office (CTO) and modem takes place in modulated analog signal. The communication between CTO to another CTO is by digital signal using time division multiplexers. Thereafter CTO feeds modulated analog signal to modem and modem converts it into binary signal for the PC. We may now say that different types of signals emerging on the communication link and reaching to CTO on their way across a city. These can be multiplexed to share the same communication link for transmitting to destination.

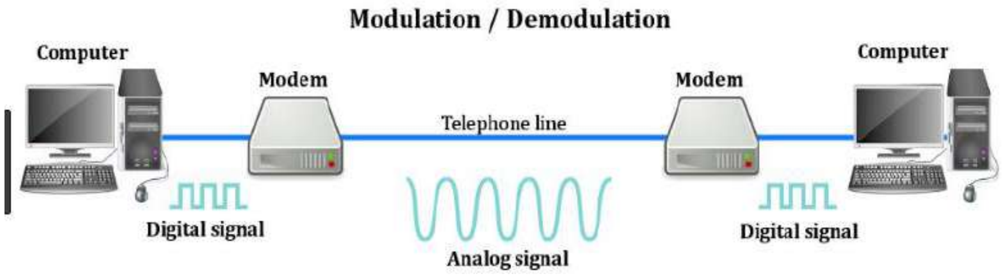


Figure 9: Data Communication System

3.3 Circuits, Channels and Multichanneling

A circuit is a path between two or more points along which an electrical current flows. In data communication, a circuit is considered as a specific path between two or more points along which signals is carried. The signals may be analog, binary or digital. This has been shown in the Figures 9 and 10.

3.4 Modulation of Digital Signal

A digital transmission uses low pass channel with high bandwidth. Similarly, analog transmission is also possible on band pass channels that require converting of binary data or a low pass analog signal into a band pass analog signal. This technique is called modulation. Thus, modulation of binary data or digital to analog modulation is carried over by changing one of the characteristics of the analog signal in accordance with the information in the digital signal. The information in the digital signal is always in the form of 0 and 1. The characteristics of analog signal that are altered are amplitude, frequency and phase of the analog waveform. Based on the change in one of the characteristics, the digital to analog modulation may be of amplitude shift keying (ASK), frequency shift keying (FSK) and phase shift keying (PSK) types. Quadrature amplitude modulation is the fourth category that combines changes in both amplitude and phase to provide better efficiency.

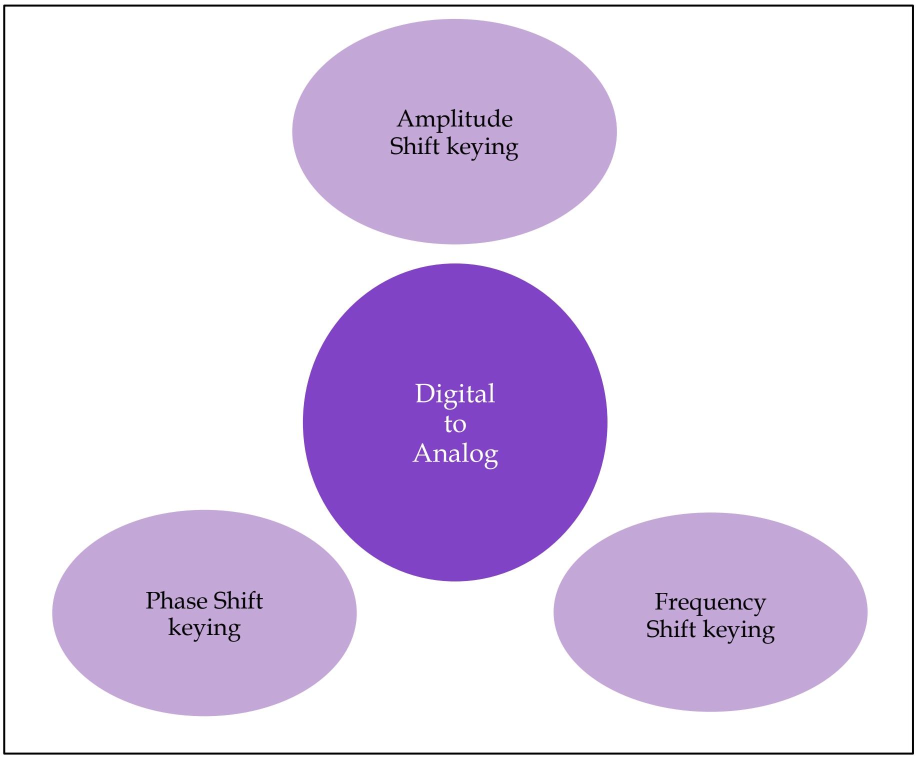


Figure 10: Techniques used for Digital to Analog Conversion

Data Rate

Bit rate is the number of bits (0 or 1) transmitted during 1 second of time. The number of signal changes per unit of time to represent the bits is called the data rate of the modem. That rate is usually expressed in terms of a unit known as a baud. A signal unit may have 1 or more than 1 bits. Therefore, the baud is the number of times per second the line condition can switch from “1” to “0”. Baud rate and bit rate, which are expressed in bits per second, usually are not the same, as several bits may be transmitted through the channel by the modem in each signal change (a few bits can be transmitted as one symbol). The relation between bit rate and baud is expressed that bit rate equals the baud rate times the number of bits represented by each signal unit. Bit rate is always more or equal than baud rate. The reason for baud rate is that it determines the bandwidth required to transmit the signal. The signal may be in the form of pieces or block that may contain bits. A fewer bandwidth required to move these signal unit with large bits for an efficient system. To understand the relation between bit and baud rate, we consider an analogy of car, passengers and highway with signal units, bits and bandwidth respectively. A car has capacity of carrying 5 passengers maximum at a time. Suppose a highway may support only 1000 cars per unit time without congestion. When each car on the highway carries 5 passengers, it is considered that the highway is capable of providing services without congestion. Thus highways services are treated efficient. Consider another case, when all these 5000 passengers wish to go in separate cars, they require 5000 cars and highway can only support 1000 cars at a time. The services offered get deteriorated because highway’s capacity is meant only for 1000 cars. It does not bother as to whether these 1000 cars are carrying 1000 passengers or 5000 passengers or more. To support more cars, the highway needs to widen. Similarly, the number of bauds determines the bandwidth.

Carrier Signal

The carrier signal that is a high frequency signal plays a significant role in the modulation and data transmission. It is the base signal generated by the sending device whose one of the characteristics is altered in accordance with the digital signal to be modulated. The modulating signal or digital signal riding over the carrier signal is transmitted to the receiving device. The receiving device is tuned to the frequency of the carrier signal. Other advantages of the carrier signal are that it provides efficient transmission between sending device and receiving device and needs smaller sizes of antenna because of higher frequency of transmission.

Amplitude Shift Keying (ASK)

ASK describes the technique how the carrier wave is multiplied by the digital signal f (t) so that the strength of the carrier wave is varied to represent binary 0 and 1. In ASK, both the frequency and phase of an analog waveform are kept uniform while amplitude is changed in accordance with the digital signal.

Mathematically, the modulated carrier signal y(t) is:

y(t) = f(t) × sin(2pfct + j) where fc is a carrier frequency and t is instantaneous time. The following Figure 11 represents ASK modulated waveform along with its input

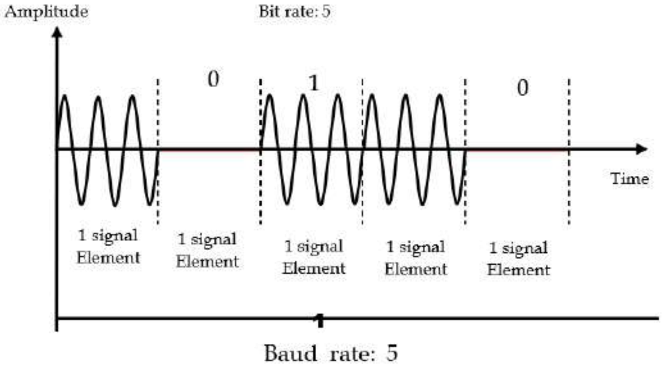


Figure 11 Binary ASK or ON-OFF Keying

The main advantage of ASK is that it is easy to produce and detect. The disadvantages of ASK are that it is highly susceptible to noise interference that changes the amplitude of the signal. A 0 can be changed to 1 and vice versa. Other drawbacks are that the speed of the changing amplitude is limited by the bandwidth of the line and the small amplitude changes suffer from unreliable detection. Telephone lines limit amplitude changes to some 3000 changes per second. The disadvantages of amplitude modulation causes this technique to no longer be used by modems, however, it is used in conjunction with other techniques.

The bandwidth for an ASK signal is mathematically given by:
Bandwidth (BW) = (1 + d) × Nbaud
Where Nbaud is the baud rate and d is the modulating index and may have minimum value as 0.

Frequency Shift Keying

Frequency of analog carrier signal is modified in accordance with the message signal. FSK describes the modulation of a carrier (or two carriers) by using a different frequency for a 1 or 0. In this technique the frequency of the carrier signal is changed according to the data while keeping the amplitude and phase constant. The transmitter sends different frequencies for a 1 than for a 0 as shown in *Figure 12*. The resultant modulated signal may be regarded as the sum of two amplitude modulated signals of different carrier frequency.

Mathematically, the modulated wave y(t) can be shown as y(t) = f1(t) sin(2pfc1t + j) + f2(t) sin(2pfc2t + j) where fc1 and fc2 are different carrier frequencies of two different signals. FSK is classified as wide band if the separation between the two carrier frequencies is larger than the bandwidth of the spectrums. Narrow-band FSK is the term used to describe an FSK signal whose carrier frequencies are separated by less than the width of the spectrum than ASK for the same modulation.

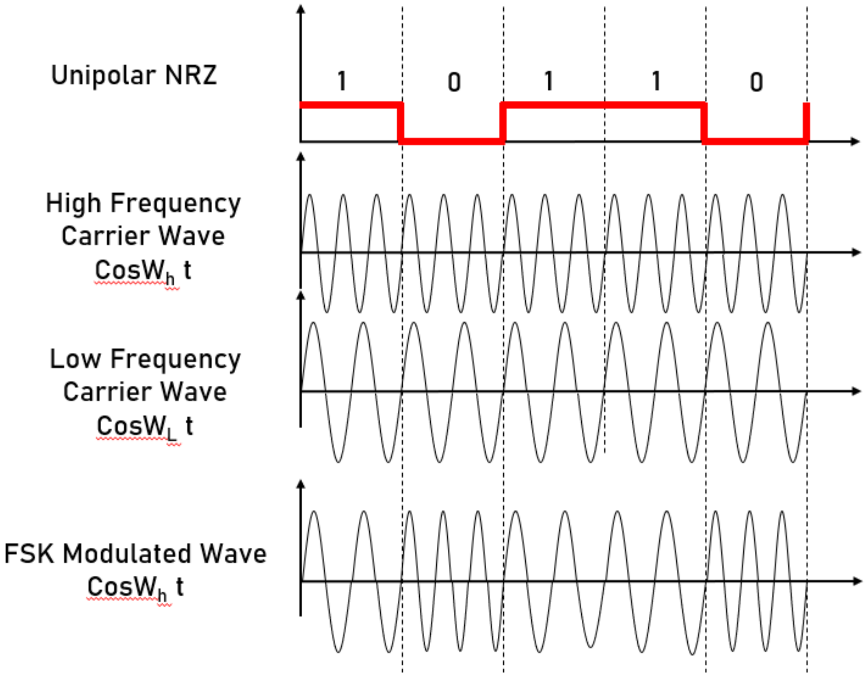

Unipolar NRZ
1
0
1
1
0
High Frequency Carrier Wave CosW$_h$ t
Low Frequency Carrier Wave CosW$_L$ t
FSK Modulated Wave CosW$_h$ t

Figure 12 Frequency Shift Keying

The advantage of FSK is that it provides better immunity from noise because the receiving device looks for specific frequency changes over given number of periods and frequency is almost unaffected from noise. The disadvantages of this technique are that again as it was with amplitude modulation. The rate of frequency changes is limited by the bandwidth of the line, and that distortion caused by the lines makes the detection even harder than amplitude modulation. Today this technique is used in low rate asynchronous modems up to 1200 baud only. The bandwidth for FSK signal is the sum of the baud rate of the signal and the frequency shift. The frequency shift is the difference between the two carrier frequencies.

Phase Shift Keying

In this modulation method a sine wave is transmitted and the phase of the sine wave carries the digital data or the phase of sine wave is varied to represent binary 1 or 0 and both the amplitude and frequency of the analog waveform are kept constant. For a 0, a 0 degrees phase sine wave is transmitted. For a 1, a 180 degrees sine wave is transmitted. As this method involves two states of phase changes, it is called binary PSK or 2-PSK. This technique, in order to detect the phase of each symbol, requires phase synchronization between the receiver's and transmitter's phase. This complicates the receiver's design. The advantages of PSK are that it is immune to noise and is not band limited.

Differential Phase Modulation A sub method of the phase modulation is differential phase modulation. In this method, the modem shifts the phase of each succeeding signal in a certain number of degrees for example, a 0 for 90 degrees and 1 for 270 degrees as illustrated in *Figure 13*.

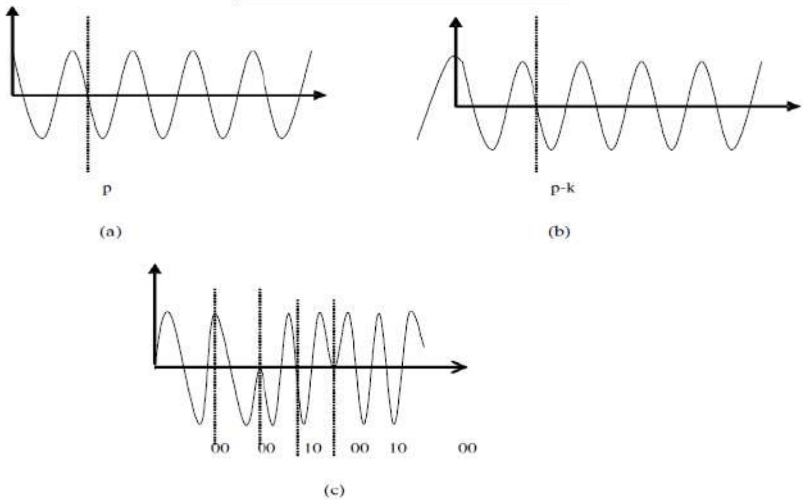


Figure 13: Phase Shift Keying

PSK is a technique, which shifts the period of a wave. The wave shown in Figure 13 (a) has a period of p starting from 0. The wave shown in Figure 13 (b) is the same wave as shown in Figure 13 (c), but its phase has been shifted. Notice that the period starts at the wave's highest point 1 on the vertical axis. It just so happens that we have shifted this wave by one quarter of the wave's full period. We can shift it another quarter, if we want to, so the original wave would be shifted by half of its period. And we could do it one more time, so that it would be shifted three quarters of its original period. This means there exist 4 separate waves and therefore each wave is provided for some binary value. Since there are 4, 2 bits are provided to each wave which is represented below Table 2.

Table 2 Bit Value and Amount of Shift

Bit Value	Amount of Shift
00	None
01	¼
10	½
11	¾

This technique of letting each shift of a wave represent some bit value is phase shift keying. But the real key is to shift each wave relative to the wave that came before it. PSK describes the modulation technique that alters the phase of the carrier. Mathematically, it can be represented as y(t) = f(t) sin(2pfct + j(t)) where jc is phase shift. This method is easier to detect than the previous one. The receiver has to detect the phase shifts between symbols and not the absolute phase.

Binary Phase Shift Keying (BPSK): In the case of two possible phases shift the modulation will be called BPSK - binary PSK. In the case of 4 different phase shifts possibilities for each symbol which means that each symbol represents 2 bits the modulation will be called quadrature PSK (QPSK), and in case of 8 different phase shifts the modulation technique will be called 8-PSK. A single data channel modulates the carrier. A single bit transition, 1 to 0 or 0 to 1, causes a 180-degree phase shift in the carrier. Thus, the carrier is said to be modulated by the data. As this has only two phases, 0 and 1 as shown in *Figure 14* and *Figure 15*. It is therefore a type of ASK with taking the values -1 or 1 and its bandwidth is the same as that of ASK. Phase shift keying offers a simple way of increasing the number of levels in the transmission without increasing the bandwidth by introducing smaller phase shifts. Quadrature phase-shift-keying (QPSK) has four phases such as 0, p/2, p, 3p/2. Consequently, M-ary PSK has M phases given by 2pm/M; m = 0,1…M-1. For a given bit-rate, QPSK requires half the bandwidth of PSK and is widely used for this reason.

The number of times the signal parameter (amplitude, frequency, and phase) is changed per second is called the signaling rate. It is measured in baud. 1 baud = 1 change per second. With binary modulations such as ASK, FSK and BPSK, the signaling rate equals the bit-rate. With QPSK and M-ary PSK, the bit-rate may exceed the baud rate.

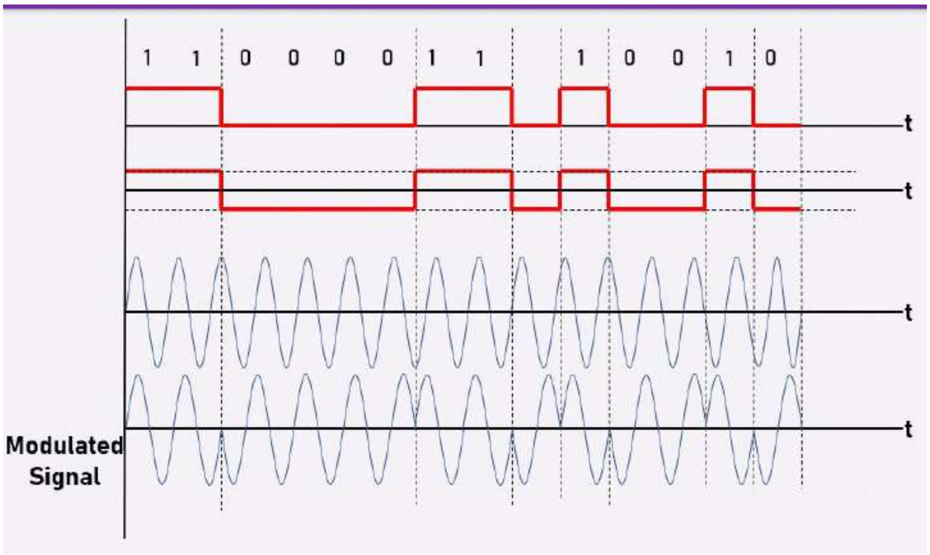


Figure 14 Binary Phase Shift Keying

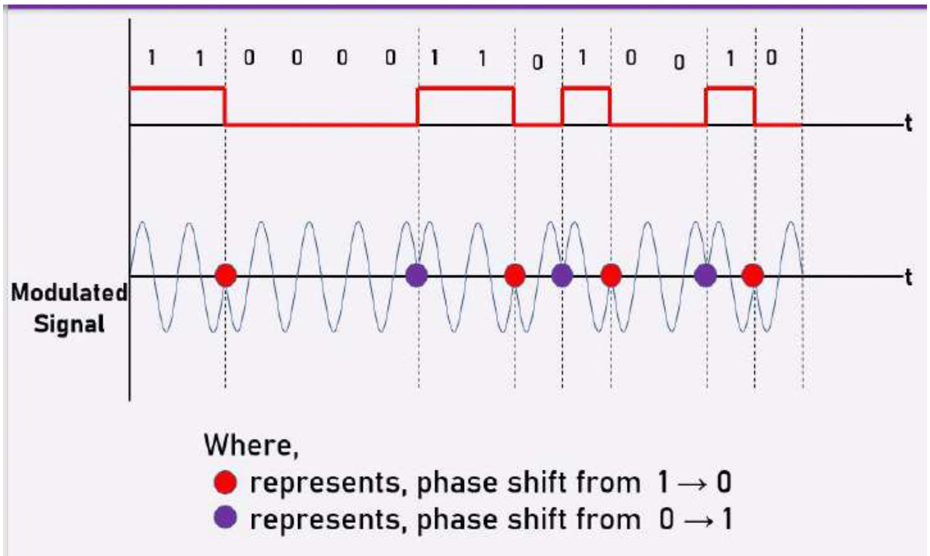

1 1 0 0 0 0 1 1 0 1 0 0 1 0
t
t
Modulated Signal
Where,
● represents, phase shift from 1 → 0
● represents, phase shift from 0 → 1

Figure 15: Output Modulated Signal

Quadrature Phase Shift Keying (QPSK)

Two data channels modulate the carrier. Transitions in the data cause the carrier to shift by either 90 or 180 degrees. This allows transmission of two discrete data streams, identified as I channel (In phase) and Q channel (Quadrature) data. In this method four different phase angles are used.

In QPSK, the four angles are usually out of phase by 90°.

QPSK Modulator

Quadrature Phase Shift Keying (**QPSK**) is a form of Phase Shift Keying in which two bits are **modulated** at once, selecting one of four possible carrier phase shifts (0, 90, 180, or 270 degrees). **QPSK** allows the signal to carry twice as much information as ordinary PSK using the same bandwidth

Analog to Analog Conversion

Analog-to-analog conversion, or analog modulation, is the representation of analog information by an analog signal. An example is radio. Analog to analog can be represented in three ways: Amplitude Modulation, Frequency Modulation and Phase Modulation.

3.5 Amplitude Modulation

This describes the technique in which the carrier wave is multiplied by the digital signal f(t). Mathematically, the modulated carrier signal y(t) is: y(t) = f(t) sin(2p¦ct+j) where fc is a carrier frequency and t is instantaneous time. *Figure 16* shows the technique of amplitude modulation. The main advantage of this technique is that it is easy to produce such signals and also to detect them. This technique has two major disadvantages. The first is that the speed of the changing amplitude is limited by the bandwidth of the line. The second is that the small amplitude changes suffer from unreliable detection. Telephone lines limit amplitude changes to some 3000 changes per second. The disadvantages of amplitude modulation causes this technique to no longer be used by modems, however, it is used in conjunction with other techniques.

QAM (Quadrature Amplitude Modulation)

This technique is based on the basic amplitude modulation. This technique improves the performance of the basic amplitude modulation. In this technique two carrier signals are transmitted simultaneously. The two carrier signals are at the same frequency with a 90 degrees phase shift.

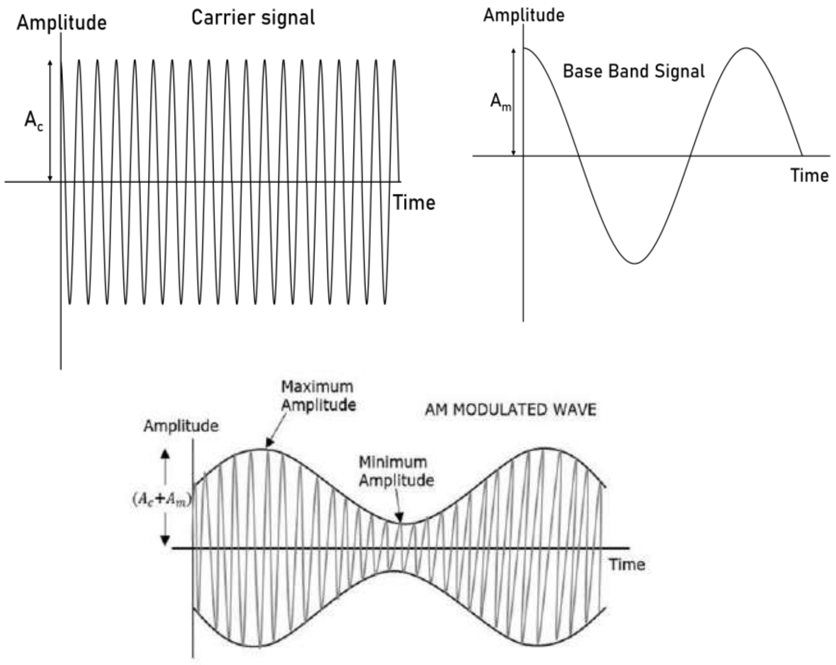


Figure 16: Amplitude Modulation

Advantages and Disadvantages of Amplitude modulation

In order that a radio signal can carry audio or other information for broadcasting or for two-way radio communication, it must be modulated or changed in some way.

Advantages and disadvantages of amplitude modulation are shown below in *Table 3*.

Table 3 Advantages and Disadvantages of Amplitude Modulation

Advantages	Disadvantages
It is simple to implement	It is not efficient in terms of its power usage
It can be demodulated using a circuit consisting of very few components	It is not efficient in terms of its use of bandwidth, requiring a bandwidth equal to twice that of the highest audio frequency.

AM receivers are very cheap as no specialized components are needed.	It is prone to high levels of noise because most noise is amplitude based and obviously AM detectors are sensitive to it.

3.6 Frequency Modulation

Frequency Modulation involves the modulation of the frequency of the analog sine wave where the instantaneous frequency of the carrier is deviated in proportion of the deviation of the modulated carrier with respect to the frequency of the instantaneous amplitude of the modulating signal. It may be said in a simple word that it occurs when the frequency of a carrier is changed based upon the amplitude of input signal. Unlike AM, the amplitude of carrier signal is unchanged. This makes FM modulation more immune to noise than AM and improves the overall signal-to noise ratio of the communications system. Power output is also constant, differing from the varying AM power output. The amount of analog bandwidth necessary to transmit FM signal is greater than the mount necessary for AM, a limiting constraint for some systems. The modulating index for FM is given as below :

$$\beta = fp/fm,$$

where

$\boldsymbol{\beta}$ = Modulation index, f_m = frequency of the modulating signal and f_p = peak frequency deviation.

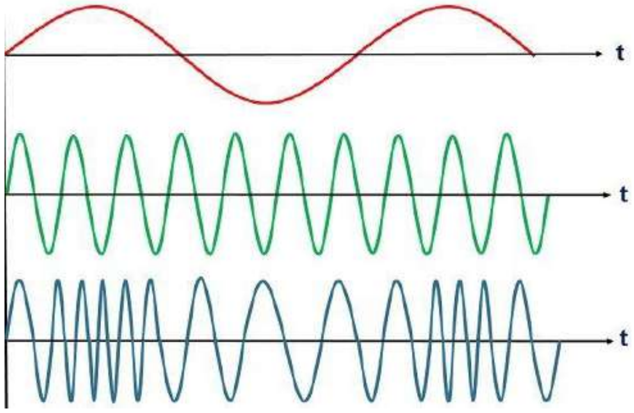


Figure 17: Frequency Modulation

From the *Figure* 17, it is inferred that the amplitude of the modulated signal always remains constant, irrespective of frequency and amplitude of modulating signal. It means that the modulating signal adds no power to the carrier in frequency modulation unlike to amplitude modulation. FM produces an infinite number of side bands spaced by the modulation frequency, fm that is not in case of AM. Therefore, AM considered a linear process whereas FM as a nonlinear process. It is necessary to transmit all side bands to reproduce a distortion free signal. Ideally, the bandwidth of the modulated signal is infinite in this case. In general the determination of the frequency content of an FM waveform is complicated, but when b is small, the bandwidth of the FM signal is 2fm. On the other hand, when b is large, the bandwidth is determined (empirically) as 2 fm (1 + b).

The difference between amplitude modulation and frequency modulation are shown in Table 4.

Table 4. Difference between Amplitude Modulation and Frequency Modulation

AMPLITUDE MODULATION	FREQUENCY MODULATION
In amplitude modulation, the frequency and phase remain the same.	In frequency modulation amplitude and phase remain the same.

Its modulation index varies from 0 to 1.	Its modulation index is always greater than one.
It has only two sidebands.	It has an infinite number of sidebands.
It has simple circuit.	It has complex circuit.
The amplitude of the carrier wave is modified in order to send the data or information.	The frequency of the carrier wave is modified in order to send the data or information.
The amplitude of the carrier wave is modified in order to send the data or information.	The frequency of the carrier wave is modified in order to send the data or information.
It requires low bandwidth in the range of 10 kHz.	It requires high bandwidth in the range of 200 kHz.
It works in a frequency range of 535 to 1705 Kilohertz (KHz).	It works in a frequency range of 88 to 108 Megahertz (MHz).
It operates in the medium frequency (MF) and high frequency (HF).	It operates in the very high frequency.
It has poor sound quality.	It has better sound quality.

Phase Modulation

Phase Modulation (PM) is similar to frequency modulation. Instead of the frequency of the carrier wave changing, the phase of the carrier wave changes. In PM the phase of the carrier is made proportional to the instantaneous amplitude of the modulating signal. Modulating index for PM is given as b = Dj, where Dj is the peak phase deviation in radians. As in the case of angular modulation argument of sinusoidal is varied and therefore we will have the same resultant signal properties for frequency and phase modulation. A distinction in this case can be made only by direct comparison of the signal with the modulating signal wave, as shown in *Figure 18*.

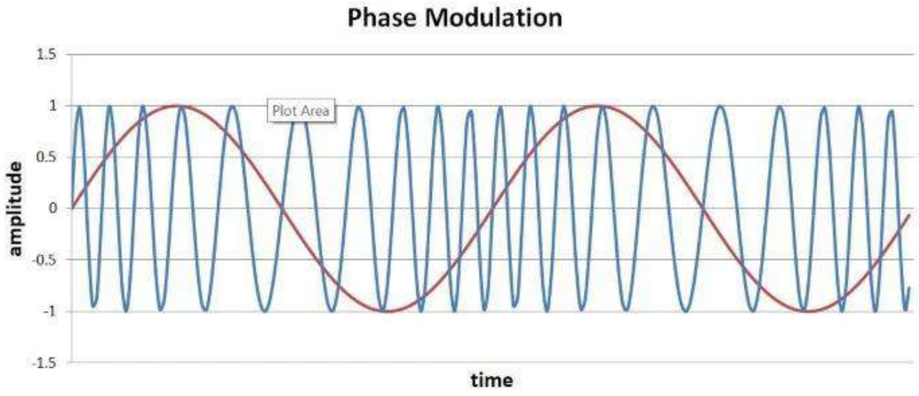


Figure 18: Phase Modulation

Caution Phase modulation and frequency modulation are interchangeable by selecting the frequency response of the modulator so that its output voltage will be proportional to integration of the modulating signal and differentiation of the modulating signal respectively. Bandwidth and power issues are same as that of the frequency modulation.

3.8 Self Assessment

State whether the following statements are true or false:

1. The modulation is the act of translating some high-frequency (base band signal) such as voice, data, etc. to a lower frequency.

2. Amplitude Modulation (AM) involves the modulation of the amplitude of the carrier as analog sine wave.

3. FM considered a linear process whereas AM as a nonlinear process.

4. In PM the phase of the carrier is made proportional to the instantaneous amplitude of the modulating signal.

5. Analog to analog signal conversion involves amplitude modulation and frequency modulation techniques only.

6. The process of changing one of the characteristics of a carrier analog signal based on the information in a digital signal is called ___________ conversion

   a. analog-to-analog
   b. analog-to-digital
   c. digital-to-analog
   d. digital-to-digital

7. In ______________ the frequency of the carrier signal is varied based on the information in a digital signal.

   a. ASK
   b. PSK
   c. FSK
   d. QAM

8. In _____________ the amplitude of the carrier signal is varied based on the information in a digital signal.

   a. ASK
   b. PSK
   c. FSK
   d. QAM

9. In ___________ the phase of the carrier signal is varied based on the information in a digital signal.

   a. ASK
   b. PSK
   c. FSK
   d. QAM

10. Most modern modems use ___________ for digital to analog modulation.

   a. ASK
   b. PSK
   c. FSK
   d. QAM

11. ________ rate is the number of bits per second; ____________ rate is the number of signals unit per second.

   a. Baud; bit

b. Bit; baud
c. Baud; base
d. Base; baud

12. For ___________, the minimum bandwidth required for transmission is equal to the baud rate.

a. ASK
b. PSK
c. FSK
d. a and b

13. In which type of modulation can the bit rate be four times the baud rate ?

a. ASK
b. FSK
c. PSK
d. None of the above

14. In which type of modulation can the bit rate be three times the baud rate?

a. ASK
b. FSK
c. PSK
d. none of the above

15. In which type of modulation can the bit rate be half the baud rate.

a. ASK
b. FSK
c. PSK
d. None of the above.

16. ASK,PSK, FSK and QAM are examples of ______________ conversion.

a. digital-to-digital
b. digital-to-analog
c. analog-to-analog
d. analog-to-digital

17. ___________ conversion is the process of changing one of the characteristics of an analog signal based on the information in the digital data.

a. Digital–to-analog
b. Analog-to-analog
c. Analog-to-digital
d. Digital-to-digital

Answers:

1. False
2. True
3. False
4. True
5. False
6. digital-to-analog
7. FSK
8. PSK
9. PSK
10. QAM
11. Bit; baud
12. a and b
13. PSK
14. None of the above
15. None of the above
16. digital-to-analog
17. Digital-to-analog

Summary

- A circuit is a path between two or more points along which signals is carried. The circuit may be a physical path consisting of wires or it may be wireless. A network, which is wired or wireless involves a number of circuits consisting of a number of intermediate switches.
- A virtual circuit is a logical path selected out of many possible physical paths available between two or more points.
- Multiplexing is the process in which multiple channels are combined for transmission over a common transmission path.
- The digital transmission requires a low pass channel with high bandwidth. The analog transmission can be carried on band pass channels. The different methods that convert binary data or a low pass analog signal into a band pass analog signal is called modulation.
- The digital to analog conversion includes ASK (Amplitude Shift Keying), FSK (Frequency Shift Keying), PSK (Phase Shift Keying), QPSK (Quadrature Phase Shift Keying), QAM (Quadrature Amplitude Modulation) and have been explained under the section Modem Modulation Techniques.
- Analog to analog signal conversion involves amplitude modulation, frequency modulation and phase modulation techniques.

Keywords

Amplitude Modulation: It involves the modulation of the amplitude of the carrier as analog

sine wave.

Amplitude Shift Keying: ASK refers to technique how the carrier wave is multiplied by the

digital signal f(t) so that the strength of the carrier wave is varied to represent binary 0 and 1.

Baud Rate: It is the number of times per second the line condition can switch from "1" to "0".

Binary Phase Shift Keying: BPSK involves two possible phases shift for the modulation.

Carrier Signal: It is the base signal generated by the sending device whose one of the

characteristics is altered in accordance with the digital signal to be modulated.

Differential Phase Shift Keying: In this method, the modem shifts the phase of each succeeding signal in a certain number of degrees.

FDM: In frequency division multiplexing, multiple channels are combined together for

transmission over a single channel.

FDMA: This divides the frequency band into various channels based on the FDM techniques.

Each of these can carry a voice conversation or, with digital service, carry digital data.

Frequency Modulation: Frequency Modulation involves the modulation of the frequency of the analog sine wave.

Frequency Shifted Keying: FSK describes the modulation of a carrier (or two carriers) by using a different frequency for a 1 or 0.

Inter modulation: It involves two (or more) sinusoids effect one another to produce undesired products, that is, unwanted frequencies (noise).

Modems: Refers to the modulator and demodulator that converts analog signal to digital signal and vice versa.

Modulation: It is the act of translating some low-frequency (base band signal) such as voice, data, etc. to a higher frequency.

Multiplexing: Refers to the process in which multiple channels are combined for transmission over a common transmission path.

Phase Modulation: Phase Modulation (PM) is similar to frequency modulation. Instead of the frequency of the carrier wave changing, the phase of the carrier wave changes.

Phase Shift Keying: In this modulation method a sine wave is transmitted and the phase of the sine wave carries the digital data or the phase of sine wave is varied to represent binary 1 or 0 and both the amplitude and frequency of the analog waveform are kept constant.

Quadrature Amplitude Modulation: This technique is based on the amplitude modulation and phase modulation to improve the performance of the amplitude modulation.

Quadrature Phase Shifted Keying: In the case of 4 different phase shifts possibilities for each symbol which means that each symbol represents 2 bits the modulation will be called quadrature PSK (QPSK).

Space Division Switching: Refers to the kind of switch developed for analog environment. Crossbar switch is the simplest possible space division switch where each packet takes a different path through the switch depending on its destination. Time division switching is based on multiplexing for digital transmission.

TDM: Refers to the process to merge data from several sources into a single channel for communication over transmission media like telephone lines, microwave system or satellite system.

TDMA: This is a digital transmission technology that allows a number of channels to access a single radio frequency (RF) channel without interference by allocating unique time slots to each channel.

Virtual circuit: This is a logical path selected out of many possible physical paths available between two or more points.

WDM: This is defined as the fibre-optic transmission technique that employs two or more optical signals having different wavelengths to transmit data simultaneously in the same direction over one fibre, and later on is separated by wavelength at the distant end.

Further/Suggested Readings

- Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall.
- Behrouz A. Forouzan and Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies.
- Burton, Bill, *Remote Access for Cisco Networks*, McGraw-Hill Osborne Media, McGraw-Hill Osborne Media.
- Rajneesh Agrawal and Bharat Bhushan Tiwari, *Computer Networks and Internet*, Vikas Publication

- https://www.geeksforgeeks.org/computer-network-tutorials/

Unit- 04: Network Models

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss concept of process network software and the significance of layering the communication process and related design issues for the layers
- Describe different technologies involved in defining the network hardware
- Explain what are the reference models for computer networks and how they are related with the OSI reference model

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.We can compare the task of networking to the task of solving a mathematics problem with a computer. The fundamental job of solving the problem with a computer is done by computer hardware. However, this is a very tedious task if only hardware is involved. We would need switches for every memory location to store and manipulate data. The task is much easier if software is available. At the highest level, a program can direct the problem-solving process; the details of how this is done by the actual hardware can be left to the layers of software that are called by the higher levels. Compare this to a service provided by a computer network. For example, the task of sending an e-mail from one point in the world to another can be broken into several tasks, each performed by a separate software package. Each software package uses the services of another software package. At the lowest layer, a signal, or a set of signals, is sent from the source computer to the destination computer. In this chapter, we give a general idea of the layers of a network and discuss the functions of each.

4.1 Layered Tasks

We use the concept of layers in our daily life. As an example, let us consider two friends who communicate through postal maiL The process of sending a letter to a friend would be complex if there were no services available from the post office. *Figure* 1shows the steps in this task.

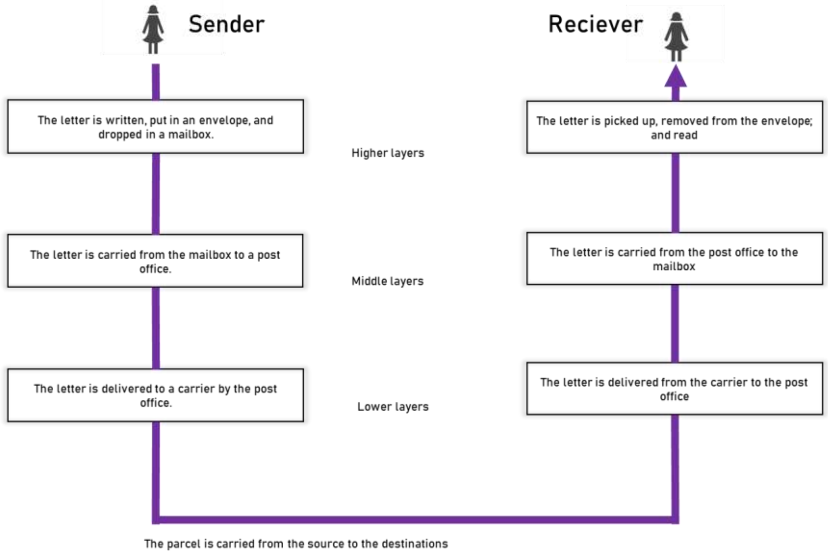

Sender
Reciever
The letter is written, put in an envelope, and dropped in a mailbox.
Higher layers
The letter is picked up, removed from the envelope; and read
The letter is carried from the mailbox to a post office.
Middle layers
The letter is carried from the post office to the mailbox
The letter is delivered to a carrier by the post office.
Lower layers
The letter is delivered from the carrier to the post office
The parcel is carried from the source to the destinations

Figure 1 Tasks Involve in Sending a Letter

Sender, Receiver, and Carrier

In *Figure 1* we have a sender, a receiver, and a carrier that transports the letter. There is a hierarchy of tasks.

At the Sender Site

Let us first describe, in order, the activities that take place at the sender site.

o Higher layer. The sender writes the letter, inserts the letter in an envelope, writes the sender and receiver addresses, and drops the letter in a mailbox.

o Middle layer. The letter is picked up by a letter carrier and delivered to the post office.

o Lower layer. The letter is sorted at the post office; a carrier transports the letter.

On the Way

The letter is then on its way to the recipient. On the way to the recipient's local post office, the letter may actually go through a central office. In addition, it may be transported by truck, train, airplane, boat, or a combination of these.

At the Receiver Site

- Lower layer. The carrier transports the letter to the post office.
- Middle layer. The letter is sorted and delivered to the recipient's mailbox.
- Higher layer. The receiver picks up the letter, opens the envelope, and reads it.

Hierarchy

According to our analysis, there are three different activities at the sender site and another three activities at the receiver site. The task of transporting the letter between the sender and the receiver is done by the carrier. Something that is not obvious immediately is that the tasks must be done in the order given in the hierarchy. At the sender site, the letter must be written and dropped in the mailbox before being picked up by the letter carrier and delivered to the post office. At the receiver site, the letter must be dropped in the recipient mailbox before being picked up and read by the recipient.

Services

Each layer at the sending site uses the services of the layer immediately below it. The sender at the higher layer uses the services of the middle layer. The middle layer uses the services of the lower layer. The lower layer uses the services of the carrier. 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 TCPIIP protocol suite became the dominant commercial architecture because it was used and tested extensively in the Internet; the OSI model was never fully implemented.

4.2 The OSI Model

Established in 1947, the International Standards Organization (ISO) is a multinational body dedicated to worldwide agreement on international standards. An ISO standard that covers all aspects of network communications is the Open Systems Interconnection model. It was first introduced in the late 1970s. An open system is a set of protocols that allows any two different systems to communicate regardless of their underlying architecture. The purpose of the OSI model is to show how to facilitate communication between different systems without requiring changes to the logic of the underlying hardware and software. The OSI model is not a protocol; it is a model for understanding and designing a network architecture that is flexible, robust, and interoperable.

ISO is the organization

OSI is the model

The OSI model is a layered framework for the design of network systems that allows communication between all types of computer systems. It consists of seven separate but related layers, each of which defines a part of the process of moving information across a network (seeFigure 2). An understanding of the fundamentals of the OSI model provides a solid basis for exploring data communications.

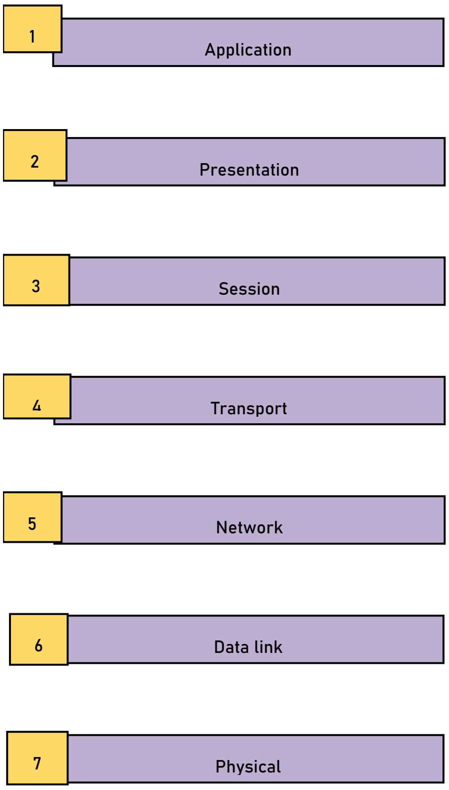


Figure 2 Seven Layers of OSI Model

Layered Architecture

The OSI model is composed of seven ordered layers: physical (layer 1), data link (layer 2), network (layer 3), transport (layer 4), session (layer 5), presentation (layer 6), and application (layer 7). Figure 2.3 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.

In developing the model, the designers distilled the process of transmitting data to its most fundamental elements. They identified which networking functions had related uses and collected those functions into discrete groups that became the layers. Each layer defines a family of functions distinct from those of the other layers. By defining and localizing functionality in this fashion, the designers created an architecture that is both comprehensive and flexible. Most importantly, the OSI model allows complete interoperability between otherwise incompatible systems. Within a single machine, each layer calls upon the services of the layer just below it. Layer 3, for example, uses the services provided by layer 2 and provides services for layer 4. Between machines, layer x on one machine communicates with layer x on another machine. This communication is governed by an agreed-upon series of rules and conventions called protocols. The processes on each machine thatcommunicate ata given layer are called peer-to-peer processes. Communication between machines istherefore a peer-to-peer process using the protocols appropriate to a given layer.

Peer-to-Peer Processes

At the physical layer, communication is direct: In Figure 2.3, 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 thenback up through the layers.

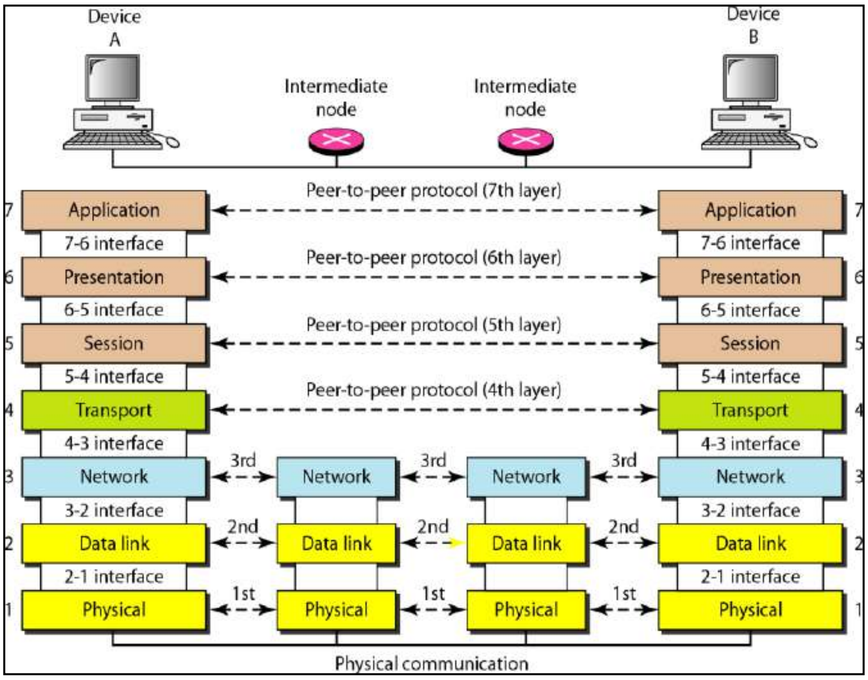


Figure 3 The interaction between layers in OSI Model

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 I 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, 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 I, 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 what 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. In Figure 4, which gives an overall view of the OSI layers, D7 means the data unit at layer 7, D6 means the data unit at layer 6, and so on. The process starts at layer 7 (theapplication 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.

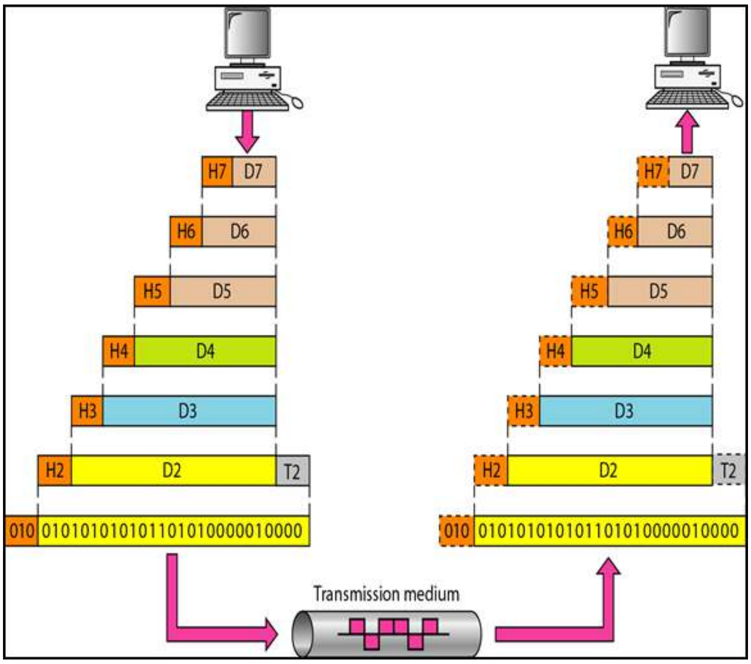


Figure 4 An exchange using the OSI model

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.

Encapsulation

Figure 2.3 reveals another aspect of data communication in the OSI model: encapsulation. A packet (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 (data and header and maybe trailer) from level N. The concept is called encapsulation; 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.

4.3 Layers in the OSI Model

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

Physical Layer

The physical layer coordinates the functions required to carry a bit stream over a physical medium. It deals with the mechanical and electrical specifications of the interface and transmission medium. It also defines the procedures and functions that physical devices and interfaces have to perform for transmission to Occur. Figure 5 shows the position of the physical layer with respect to the transmission medium and the data link layer.

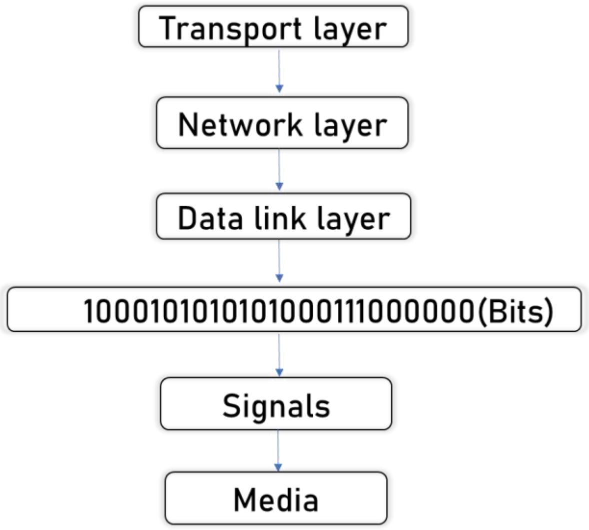


Figure 5 Physical Layer

The physical layer is also concerned with the following:

Physical characteristics of interfaces and medium. The physical layer defines the characteristics of the interface between the devices and the transmission medium. It also defines the type of transmission medium.

Representation of bits. The physical layer data consists of a stream of bits (sequence of Os or 1s) with no interpretation. To be transmitted, bits must be encoded into signals--electrical or optical. The physical layer defines the type of encoding (how Os and I s are changed to signals).

Data rate. The transmission rate-the number of bits sent each second-is also defined by the physical layer. In other words, the physical layer defines the duration of a bit, which is how long it lasts.

Synchronization of bits. The sender and receiver not only must use the same bit rate but also must be synchronized at the bit level. In other words, the sender and the receiver clocks must be synchronized.

Line configuration. The physical layer is concerned with the connection of devices to the media. In a point-to-point configuration, two devices are connected through a dedicated link. In a multipoint configuration, a link is shared among several devices.

Physical topology. The physical topology defines how devices are connected to make a network. Devices can be connected by using a mesh topology (every device is connected to every other device), a star topology (devices are connected through a central device), a ring topology (each device is connected to the next, forming a ring), a bus topology (every device is on a common link), or a hybrid topology (this is a combination of two or more topologies).

Transmission mode. The physical layer also defines the direction of transmission between two devices: simplex, half-duplex, or full-duplex. In simplex mode, only one device can send; the other can only receive. The simplex mode is a one-way communication. In the half-duplex mode, two devices can send and receive, but not at the same time. In a full-duplex (or simply duplex) mode, two devices can send and receive at the same time.

Data Link Layer

The data link layer transforms the physical layer, a raw transmission facility, to a reliable link. It makes the physical layer appear error-free to the upper layer (network layer). Figure 6shows the relationship of the data link layer to the network and physical layers.

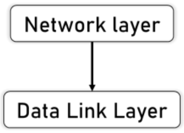


Figure 6 Data Link Layer

Other responsibilities of the data link layer include the following:

- ***Framing***. The data link layer divides the stream of bits received from the network layer into manageable data units called frames.
- ***Physical addressing***. If frames are to be distributed to different systems on the network, the data link layer adds a header to the frame to define the sender and/or receiver of the frame. If the frame is intended for a system outside the sender's network, the receiver address is the address of the device that connects the network to the next one.
- ***Flow control***. If the rate at which the data are absorbed by the receiver is less than the rate at which data are produced in the sender, the data link layer imposes a flow control mechanism to avoid overwhelming the receiver.
- ***Error control***. The data link layer adds reliability to the physical layer by adding mechanisms to detect and retransmit damaged or lost frames. It also uses a mechanism to recognize duplicate frames. Error control is normally achieved through a trailer added to the end of the frame.
- ***Access control***. 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.

Addressing: It can be of two types as shown in .Networks operate in exactly the same way. The physical address of a computer on a LAN, fixed in the network interface card by the manufacturer, works like physical knowledge of where a house is. Messages sent on the same LAN segment, corresponding to a house's neighborhood, can get to a specific physical LAN address. If you want to send a message across the country or around the world, though, you need the equivalent of a country, state, city, and street address. In network terms, the address that you need is called a logical address. The key difference between physical addresses and logical addresses is that, although physical addresses are scattered randomly around the world, logical addresses follow a pattern determined by network administrators and stored in routing tables. Routing tables (used by routers) are the equivalent of street maps, guiding messages to their destination.

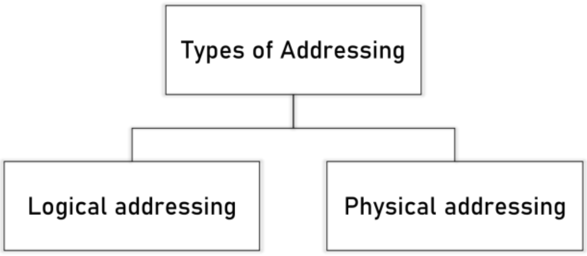


Figure 7 Types of Addressing

Note The data link layer is responsible for moving frames from one hop to the next.

Network Layer

The network layer is responsible for the source-to-destination delivery of a packet, possibly across multiple networks (links). Whereas the data link layer oversees the delivery of the packet between two systems on the same network (links), the network layer ensures that each packet gets from its point of origin to its final destination. If two systems are connected to the same link, there is usually no need for a network layer. However, if the two systems are attached to different networks (links) with connecting devices between the networks (links), there is often a need for the network layer to accomplish source-to-destination delivery. Figure 8 shows the relationship of the network layer to transport layers.

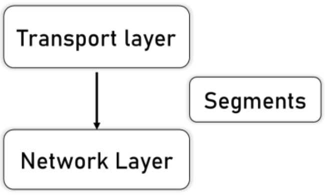


Figure 8 Network Layer

Other responsibilities of the network layer include the following:

Logical addressing. The physical addressing implemented by the data link layer handles the addressing problem locally. If a packet passes the network boundary, we need another addressing system to help distinguish the source and destination systems. The network layer adds a header to the packet coming from the upper layer that, among other things, includes the logical addresses of the sender and receiver.

Routing. When independent networks or links are connected to create *internetworks* (network of networks) or a large network, the connecting devices (called *routers*or *switches)* route or switch the packets to their final destination. One of the functions of the network layer is to provide this mechanism. Figure 9 illustrates end-to-end delivery by the network layer.

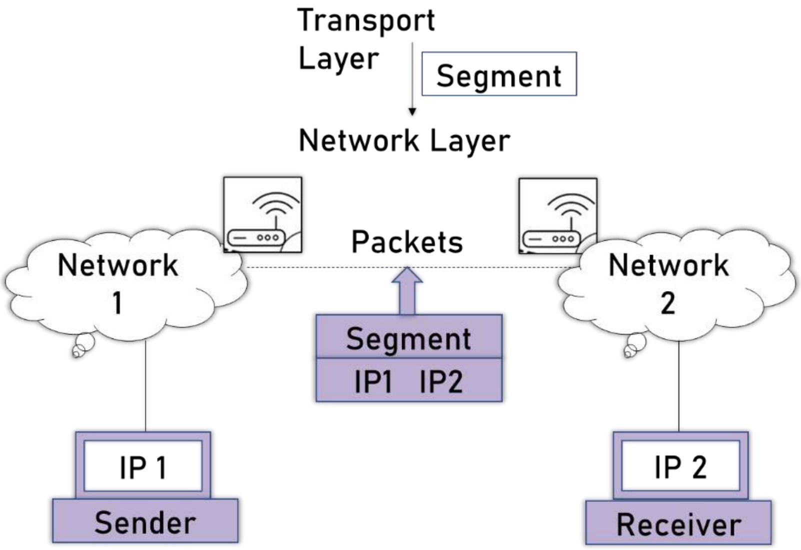


Figure 9 Source to destination delivery

Transport Layer

The transport layer is responsible for process-to-process delivery of the entire message. A process is an application program running on a host. Whereas the network layer oversees source-to-destination delivery of individual packets, it does not recognize any relationship between those packets. Ittreats each one independently, as though each piece belonged to a separate message, whether or not it does. The transport layer, on the other hand, ensures that the whole message arrives intact and in order, overseeing both error control and flow control at the source-to-destination level. Figure 10shows the relationship of the transport layer to the network and session layers.

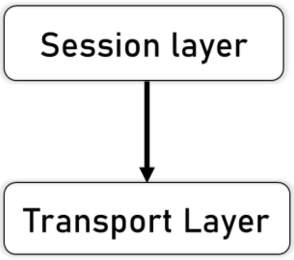


Figure 10 Transport Layer

Other responsibilities of the transport layer include the following:

- ***Service-point addressing***. Computers often run several programs at the same time. For this reason, source-to-destination delivery means delivery not only from one computer to the next but also from a specific process (running program) on one computer to a specific process (running program) on the other. The transport layer header must therefore include 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.
- ***Segmentation and reassembly***. A message is divided into transmittable segments, with each segment containing a sequence number as shown in Figure 11. These numbers enable the transport layer to reassemble the message correctly upon arriving at the destination and to identify and replace packets that were lost in transmission.

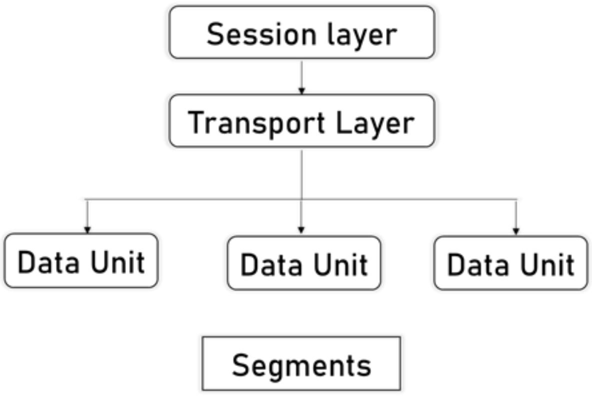


Figure 11 Segmentation

- ***Connection control***. The transport layer can be either connectionless or connection-oriented. A connectionless transport layer treats each segment as an independent packet and delivers it to the transport layer at the destination machine. A connection-oriented transport layer makes a connection with the transport layer at the destination machine first before delivering the packets. After all the data are transferred, the connection is terminated.
- ***Flow control***. Like the data link layer, the transport layer is responsible for flow control. However, flow control at this layer is performed end to end rather than across a single link.
- ***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.

Protocols of Transport Layer

The protocols of transport layer along with difference is explained below *Table 1*.

Table 1 Protocols of Transport Layer

TCP	UDP
Connection Oriented transmission	Connectionless transmission
TCP is slower	UDP is faster
Acknowledgment	No Acknowledgment
Used where full data delivery is must	Does not matter whether we gave received all data of Transport Layer.

Note: The transport layer is responsible for the delivery of a message from one process to another.

Session Layer

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

Specific responsibilities of the session layer include the following:

- 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 (one way at a time) or full-duplex (two ways at a time) mode.
- Authentication and Authorization: Authentication means confirming your own identity, while authorization means granting access to the system. In simple terms, authentication is the process of verifying who you are, while authorization is the process of verifying what you have access to.

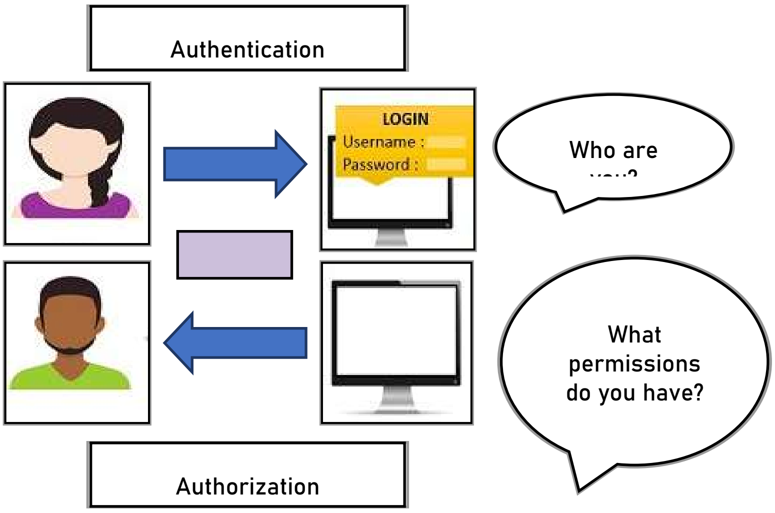


Figure 12 Authentication and Authorization

- Session management
- Synchronization. The session layer allows a process to add checkpoints, or synchronization points, to a stream of data. For example, if a system is sending a file of 2000 pages, it is advisable to insert checkpoints after every 100 pages to ensure that each 100-page unit is received and acknowledged independently. In this case, if a crash happens during the transmission of page 523, the only pages that need to be resent after system recovery are pages 501 to 523. Pages previous to 501 need not be resent. Figure 2.12 illustrates the relationship of the session layer to the transport and presentation layers.

Presentation Layer

The presentation layer performs functions related to the syntax and semantics of the information transmitted that include formatting and displaying of received data by terminals and printers. It is concerned with differences in the data syntax used by communicating applications. This layer is responsible for remedying those differences by resorting to mechanisms that transform the local syntax (specific to the platform in question) to a common one for the purpose of data exchange. For example, it performs encoding of data in a standard agreed upon way to facilitate information exchange among heterogeneous systems using different codes for strings, for example, conversion between ASCII and EBCDIC character codes. It facilitates data compression for reducing the number of bits to be transmitted and encrypts data for privacy and authentication, if necessary.

Translation.The processes (running programs) in two systems are usually exchanging information in the form of character strings, numbers, and so on. The information must be changed to bit streams before being transmitted. Because different computers use different encoding systems, the presentation layer is responsible for interoperability between these different encoding methods. 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. To carry sensitive information, a system must be able to ensure privacy. Encryption means that the sender transforms the original information toanother 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 reduces the number of bits contained in the information. Data compression becomes particularly important in the transmission of multimedia such as text, audio, and video.

Application Layer

The application layer provides support services for user and application tasks. It determineshow the user will use the data network. It allows the user to use the network. For example, itprovides network-based services to the end user.Examples of network services are distributed databases, electronic mail, resource sharing, filetransfers, remote file access and network management. This layer defines the nature of the task to be performed.

4.4 TCP/IP

The OSI Model we just looked at is just a reference/logical model. It was designed to describe the functions of the communication system by dividing the communication procedure into smaller and simpler components. But when we talk about the TCP/IP model, it was designed and developed by Department of Defense (DoD) in 1960s and is based on standard protocols. The layers in the TCP/IP protocol suite do not exactly match those in the OSI model. The original TCP/IP protocol suite was defined as having four layers: host-to- network, internet, transport, and application. However, when TCP/IP is compared to OSI, we can say that the TCP/IP protocol suite is made of five layers: physical, data link, network, transport, and application. TCP/IP Model shown inFigure 13.

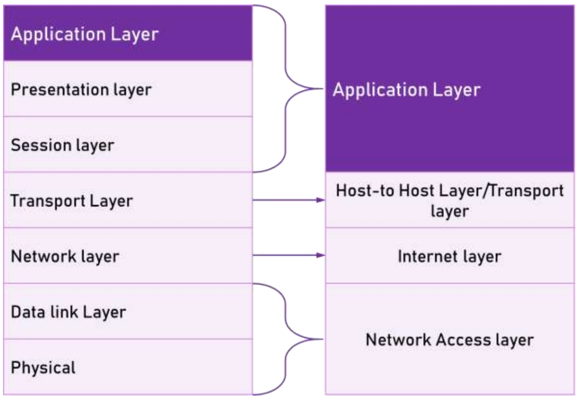


Figure 13 TCP/IP Model (b)

Internet Layer

The packet format and protocol at this layer is called Internet Protocol (IP). IP is a connectionless type service that introduces IP packets into any network. The packets travel independently to the destination. Prior to transmission of data, no logical connection is needed. The TCP/IP Internet layer corresponds to the network layer of the OSI reference model in functionality, as shown in Figure 13.

Transport Layer Notes

The transport layer of TCP/IP model corresponds to the transport layer of the OSI reference model. It is represented by two end-to-end protocols namely, TCP (Transmission Control Protocol) and UDP (User Datagram Protocol). TCP is a reliable connection-oriented protocol and UDP is an unreliable connectionless protocol.

Application Layer

The TCP/IP model was the first of its kind model and therefore did not contain session or presentation layers because of its little use to most of the applications. This layer has all the higher-level protocols, as shown in Figure 13.

Network Access Layer

The layer below the Internet layer is not defined and varies from host and network to network. The TCP/IP model suggests that the host has to connect to the network using some protocol so it can send IP packets over it.

4.5 Similarities and Difference between OSI and TCP/IP

It stands for Transmission Control Protocol/Internet Protocol. The TCP/IP model is a concise version of the OSI model. It contains four layers, unlike seven layers in the OSI model.

Similarities between OSI and TCP/IP

- They share similar architecture.
- They share a common application layer.
- Knowledge by networking professionals.
- Both models assume that packets are switched.

OSI VS TCP/IP

OSI Model	TCP/IP Model
OSI appears to complex because of 7 layers	TCP/IP appears to be a more simpler model and this is mainly due to the fact that it has fewer layers.

Networks are not usually built around the OSI model as it is merely used as a guidance tool.	TCP/IP is considered to be a more credible model- This is mainly due to the fact because TCP/IP protocols are the standards around which the internet was developed therefore it mainly gains creditability due to this reason
The OSI model is bottom to up process of network connection.	TCP/IP is the top to bottom process structure for internet purpose.
OSI defines several more layers of standardized functions	TCP/IP makes no assumptions about what happens above the level of a network session
OSI model is a reference model.	TCP/IP is an implementation of OSI model.
OSI provides layer functioning and also defines functions of all the layers.	TCP/IP model is more based on protocols and protocols are not flexible with other layers.
In OSI model the transport layer guarantees the delivery of packets	In TCP/IP model the transport layer does not guarantees delivery of packets
Follows horizontal approach	Follows vertical approach.
OSI model defines services, interfaces and protocols very clearly and makes clear distinction between them	In TCP/IP it is not clearly separated its services, interfaces and protocols
OSI is a general model	TCP/IP model cannot be used in any other application.
Network layer of OSI model provide both connection oriented and connectionless service.	The Network layer in TCP/IP model provides connectionless service.
It has 7 layers	It has 4 layers
Transport layer guarantees delivery of packets	Transport layer does not guarantees delivery of packets
The protocol are better hidden and can be easily replaced as the technology changes.	It is not easy to replace the protocols
OSI truly is a general model	TCP/IP can not be used for any other

OSI model represents an idea	application TCP/IP network model represents reality in the world

Summary

- To set up a computer network, you'll need three fundamental components: hardware, protocols (software), and applications (useful software). The importance of the idea of layers in networking is also discussed.
- Each two-layer layer serves as an interface for the top layer, allowing each layer to alter with minimal influence on the above levels. This security can be so effective that a programme may be unaware that it is operating on separate hardware.
- There are seven levels in the OSI network paradigm.
- Transmission Control Protocol/Internet Protocol (TCP/IP) is the acronym for Transmission Control Protocol/Internet Protocol. It was created with the goal of defining a set of protocols capable of enabling transparent communications interoperability between computers of any size, independent of the hardware or operating system platforms that support them.
- TCP/IP has grown in popularity over time to become the most widely used protocol today. The public availability of TCP/protocol IP's specifications is one of the reasons for its popularity. TCP/IP may legitimately be called an open system in this regard. TCP/IP is used by most users for file transfers, electronic mail (e-mail), and remote login services.

Keywords

- The Internet protocol suite is a collection of communication protocols that are used on the Internet and other comparable networks.
- Reference Model for Open Systems Interconnection (OSI): The OSI model of data transmission was created by the International Standardization Organization (ISO) in 1984. OSI defines a seven-layer model for defining protocol architectures and functional features, which is utilised by the industry as a frame of reference.
- TCP/IP: Technically, Transmission Control Protocol (TCP) and Internet Protocol (IP) are two separate network protocols. TCP and IP, on the other hand, are so widely used together that TCP/IP has become standard nomenclature to refer to any or both protocols.

Self-Assessment

1. OSI stands for __________

A. open system interconnection
B. operating system interface
C. optical service implementation
D. open service Internet

2. TCP/IP model does not have ______ layer but OSI model have this layer.

A. session layer
B. transport layer
C. application layer

D. network layer

3. Which layer is used to link the network support layers and user support layers?

A. session layer

B. data link layer

C. transport layer

D. network layer

4. TCP/IP model was developed _____ the OSI model.

A. prior to

B. after

C. simultaneous to

D. with no link to

5. Which layer is responsible for process to process delivery in a general network model?

A. network layer

B. transport layer

C. session layer

D. data link layer

6. Which address is used to identify a process on a host by the transport layer?

A. physical address

B. logical address

C. port address

D. specific address

7. Transmission data rate is decided by ____________

A. network layer

B. physical layer

C. data link layer

D. transport layer

8. A layer of the OSI model on one system communicates with the ___ layer of its peer system.

A. above

B. below

C. same

D. None

9. TCP/IP model does not have ______ layer but OSI model have this layer.

A. session layer

B. presentation layer

C. application layer

D. both (a) and (b)

10. Which layer provides the services to user?

A. application layer

B. session layer

C. presentation layer

D. none of the mentioned

11. The number of layers in Internet protocol stack

A. 5

B. 7

C. 6

D. None of the mentioned

12. Transport layer is implemented in

A. End system

B. NIC

C. Ethernet

D. None of the mentioned

Answers for Self Assessment

1.	A	2.	A	3.	C	4.	A	5.	B
6.	C	7.	B	8.	C	9.	D	10.	A
11.	A	12.	A						

Review Questions

1. What are the most significant design considerations for computer-to-computer communication?

2. In the ISO-OSI paradigm, what are the main roles of the network layer? What distinguishes the network layer's packet delivery role from that of the data link layer?

3. In the OSI reference model, what is the objective of layer isolation?

4. Why is the OSI Reference Model so extensively used? What good did it do to establish itself as a data transmission standard?

5. Compare and contrast the OSI reference model with the TCP/IP model.

Further Readings

Andrew S. Tanenbaum, Computer Networks, Prentice Hall

Behrouz A. Forouzan and Sophia Chung Fegan, Data Communications and Networking, McGraw-Hill Companies

Burton, Bill, Remote Access for Cisco Networks, McGraw-Hill Osborne Media McGraw-Hill Osborne Media

Rajneesh Agrawal and Bharat Bhushan Tiwari, Computer Networks and Internet, Vikas

Unit 05 – Physical Layer

CONTENTS

Objectives

- learn what is transmission media.
- differentiate between wired and wireless media.
- learn different types of switching techniques along with their advantages and dis-advantages.
- learn about the various networking devices.

Introduction

As you can see in the Figure 1, whatever medium you use to transfer data from source to destination is important. As a result, this is referred to as a transmission medium. When two people communicate with one another, a medium may be either wired or wireless. In that case, the air serves as a contact or transmitting medium. So, the transmission medium is known as air, but in this case, it is the transmission medium that is transmitting a packet from source to destination via the transmission medium, which can be wired or wireless.

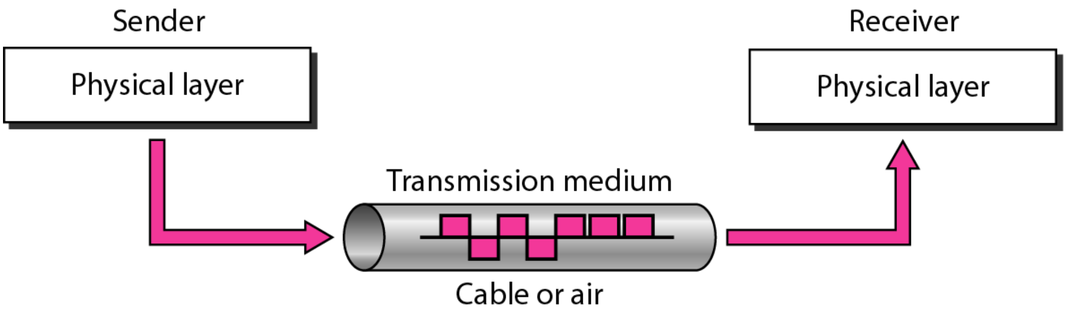


Figure 1 Transmission medium

Notes

As a result, the transmission medium is the means by which data is transmitted from one location to another. As a result, any message can be sent as data that can be translated into binary digits. As a result, the signal is encoded with these digits. The OSI model was discussed in the previous lecture. So, data is sent to a network, the network divides into packets, and the packets are formed by the network. A mail carrier truck or an aero plane could be the delivery medium for a written letter. Similarly, when 2332333

5.1 Classes of transmission media

So, there are two forms of transmission media: guided media and unguided media as shown in Figure 2. Under guided media, we have twisted pair cables, coaxial cables, and fibre optic cables.

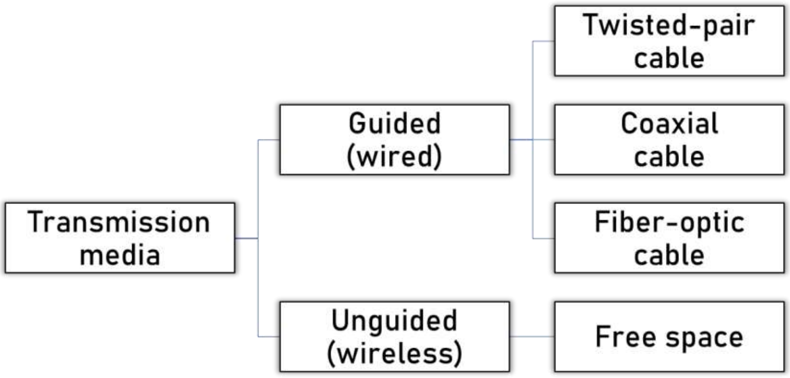


Figure 2 Classes of transmission media

Factors affecting for selecting a transmission media

So, it may be a transmission rate, cost, or ease of installation, which means that all communication media must be installed with care. The installation should be carried out by professional and experienced technicians and supervisors, and it should be resistant to the elements. Distances have an effect on the transmission medium that must be chosen, as well as the form of transmission media that must be chosen.

Guided Media

The guided media refers to a method of directing you through the process of sending data from one location to a nother through a wired link. Wired media is another term for guided media. As a result, those that are tangible or have a physical presence are those. We have a broad range of boundary transmission media. The term "boundary transmission media" refers to the use of cables or wires to link the source and destination.

Bounded transmission media are

a. Twisted-pair cable

b. Coaxial cable

c. Fiber-optic cable

Twisted-pair cable

So, let's look at why we twist the wires. As you can see in the picture, if we don't twist the wires, the two wires are parallel, which means the effect of these unwanted signals isn't the same in both wires. For example, one is here and the other is here, the two parallel wires are here, some noise sources are there, which means my upper wire is closer to the noise source than the lower wire.

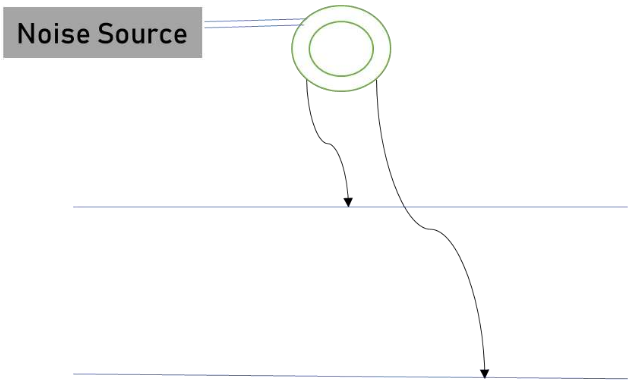


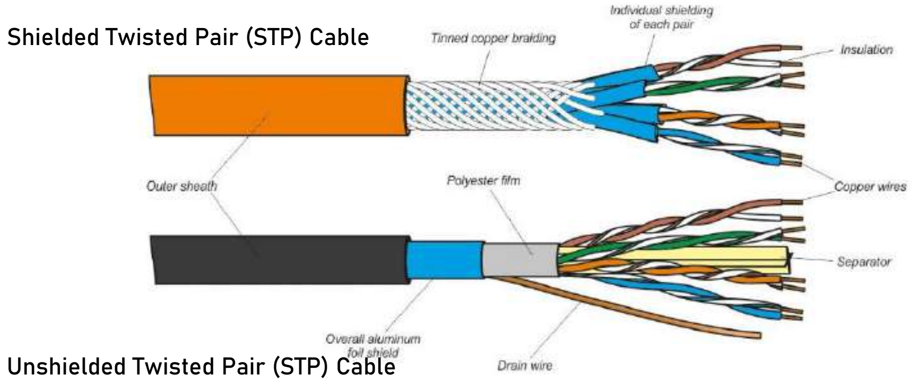
Figure 3 Effect of wire on parallel wires

that if two wires are connected in a parallel manner. As a result, my upper wire would be more affected than my other wires, owing to their different positions in relation to noise or crosstalk. As a consequence, there is a disparity at the receiver.

Similarly, we should twist the wire base to minimise crosstalk or noise because my upper wire will be close to the noise source half of the time, and my lower wire will be close to the noise source half of the time, which is why we twist the wire. So, in this case, one wire is used to transmit data and the other wire is used to receive data. As a result of the twisting, both wires are likely to be similarly influenced by external forces. Since both wires will be affected equally when we twist the wire, since half of the time this wire is closer to the noise source than half of the time this wire is closer to the noise source. As a result, the receiver that determines the difference between the two does not pick up any unnecessary signals. As a result, the undesirable signals are largely suppressed. As a result, it's obvious that the number of twists per unit of length has an effect on the cable's consistency. The higher the twist, the higher the cable's efficiency. As a result, twists often have an effect on the cable's consistency.

UTP and STP cables

Twisted pair cable comes in two varieties. Unshielded twisted pair cable is one kind, while shielded twisted pair cable is another. The most popular form of networking cable is unshielded twisted pair cable. STP is a shielded twisted pair cable that is mostly used in large organisations due to its high cost and bulkiness. It is bulkier because it has an additional metal covering.


Shielded Twisted Pair (STP) Cable
Tinned copper braiding
Individual shielding of each pair
Insulation
Outer sheath
Polyester film
Copper wires
Separator
Overall aluminum foil shield
Drain wire
Unshielded Twisted Pair (STP) Cable

As a result, it is bulkier, and handling this form of cable is difficult. As a result, we mostly use a UTP cable. It's not only for networking; it's also for standard telephone UTP cat 1 connections. There are various types of UTP cables available, but UTP cat five, E is the most common UTP cable, which was developed to replace the old coaxial cable, which was unable to keep up with the ever-increasing demand for faster and more reliable networks. So, we have a variety of different cable categories, and depending on which category you want to achieve, you'll need the required cable. A metal foil or braided mesh covering encases each pair of insulated conductors in shielded twisted pair cable. Both of the metal cases boost the cable's quality, and because the cable's length increases, it has an additional covering. As a result, my cable has a longer length than a UTP cable. It also has the extra covering that prevents noise or crosstalk from penetrating, as well as being bulkier and more costly. As a result, the most common twisted pair cable in communication is known as a UTP cable, which stands for unshielded twisted pair cables.

Table 1 Category of cables

CAT3:	Rarely used today, CAT3 is usually deployed in phone lines. It supports 10 Mbps for up to 100 meters.
CAT4	: Typically used in token ring networks, CAT4 supports 16 Mbps for up to 100 meters
CAT5:	Used in Ethernet-based LANs, CAT5 contains two twisted pairs. It supports 100 Mbps for up to 100 meters.
CAT5e:	Used in Ethernet-based LANs, CAT5e contains four twisted pairs. It supports 1 Gbps for 100 meters.
CAT6:	Used in Ethernet-based LANs and data center networks, CAT6 contains four tightly wound twisted pairs. It supports 1 Gbps for up to 100 meters and 10 Gbps for up to 50 meters.

Category of cables

The EIA (Electronic Industry Association) has created a set of 12 standards that divide UTP cable into seven groups. So there are categories that decide cable quality from one to seven, with one being the lowest and seven being the best, such as cat three cat 4, cat 5, cat 5 e and, cat six as shown in Table 1. So, as I previously said, cat 5e is used in Ethernet-based LANs, and we also have a cat Three that is rarely used today. Cat six is a twisted pair network that uses Ethernet. It supports one gigabit per second, up to 100 metres, and 10 Gbps for up to 50 metres.

UTP Connector

RJ 45 female and RJ 45 mail connector, also known as registered jack, are the most common connectors in the UTP.

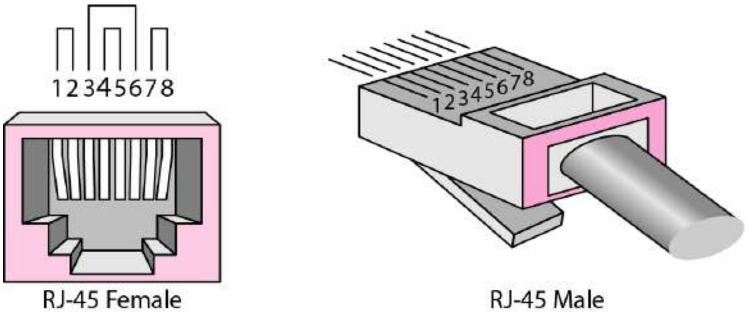


Figure 4 UTP Connector

Advantages of Twisting

It eliminates crosstalk, or the flow of information from one buyer to another. It also protects against external signal interference. It's simple, versatile, and attached, which means that if you want to extend the cable's length, you can do so easily. It's also simple to install and maintain; the steps are straightforward, and there's no need to learn any complicated formulas. It's also lightweight and inexpensive. It is available for purchase by everyone.

Disadvantages of Twisted Pair Cable

A signal with a high attenuation has been lost. Carrying a signal over a long distance without a repeater is a challenge because when the signal loses energy and becomes weaker and weaker, it becomes a problem. So, in that case, I'll need a repeater, and the role of a repeater is to improve the signal. So, low bandwidth, maximum data range from one megabit per second to ten megabits per second, without any additional equipment repeater is also hardware and incurs costs, so this is one of the disadvantages.

5.2.2 Coaxial Cable

Since the two media are built so differently, coaxial cable can carry signals with higher frequency ranges than twisted pair cable. Coax has a solid or regular wire centre core that is surrounded by an insulated layer, as seen in the Figure 5.

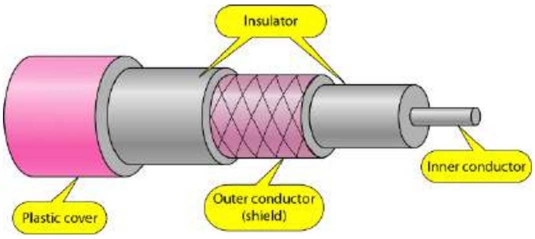


Figure 5 Coaxial Cable

As a result, cable TV connectors companies are the most popular users of coaxial cable. So, there are various types of coaxial cables, and they are classified as coaxial cables based on their radio government ratings.

Connectors

The most popular coaxial cable connectors are Bayonet Neill connectors, which are used to attach coaxial cable to the devices that require coaxial connectors. Now, as you can see in the diagram, the most popular connectors are BNC, BNC T connector, and BNC Terminator as shown in Figure 6.

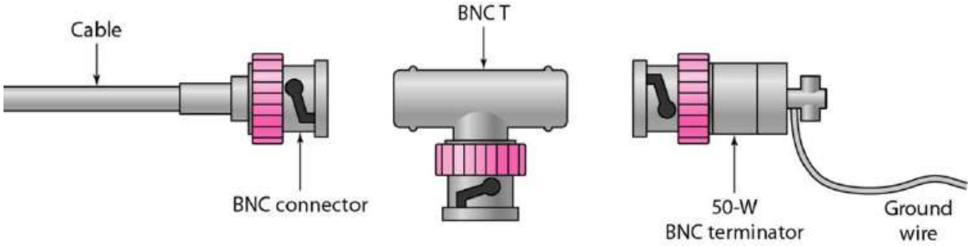


Figure 6 BNC Connector

There are three different types of connectors. As a result, BNC connectors are used to link the end of a cable to a system such as a television set. To avoid signal reflection, a BNC Terminator is used at the cable's end. While coaxial cable has a far higher bandwidth than twisted pair cable, attenuation is far higher in this case, and the signal weakens quickly, necessitating the use of repeaters on a regular basis.

Performance of coaxial cable

While coaxial cable has a far higher bandwidth than twisted pair cable , attenuation is far higher in this case, and the signal weakens quickly, necessitating the use of repeaters on a regular basis. The responsibilities of coaxial cables and how they are used in television signals. It's used to link computers in local area networks, and shorter lengths of this cable are used to connect machines to test equipment including signal generators. It is also used to transmit radio signals.

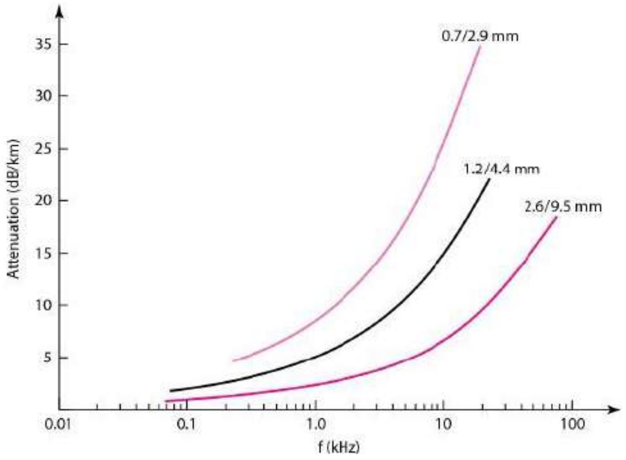
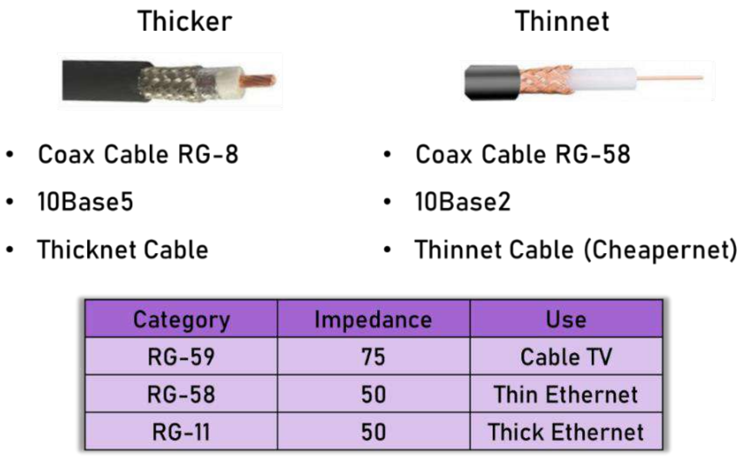


Figure 7 Performance of coaxial cable

Categories of Coaxial Cables

There are two types of coaxial cable available: thin net and thick net Figure *8*. In reality, thick net is still used for backbone wiring, and twisted pair cable is an alternative to thinnet on an Ethernet network. Thicknet was the initial Ethernet wiring, but thinnet is less expensive and easier to install. In the case of thick net, the maximum length is 500 metres; in the case of thin net, the maximum length is 500 metres. It can reach a maximum length of 185 metres. Its names, thicknet and thin net, are self-evident. The term "thick net" refers to a cable that is thicker than the term "thin net," and the term "thin net" refers to a cable that is thinner. Table 1 shows the disparity between thicknet and thinnet.



Thicker

- Coax Cable RG-8
- 10Base5
- Thicknet Cable

Thinnet

- Coax Cable RG-58
- 10Base2
- Thinnet Cable (Cheapernet)

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

Figure 8 Categories of coaxial cable

Table 2 Difference between thicknet and thinnet

Thicknet	Thinnet
Thicker than thinnet	Thinner
Max length 500 meters	Max length 185 meters

Advantages of Coaxial cable

Now consider the benefits of coaxial cable transmission rate over twisted pair cable. It can be used in a mutual cable network and is mostly used for broadband transmission. Both channels are sent to the same wire at the same time. It has a high bandwidth of 400 megabits per second.

Disadvantages of Coaxial Cable

It's pricey, and it's not compatible with twisted pair cable if you want to extend the cable's duration. Since coaxial and twisted pair cables are incompatible, they cannot be linked. As a result, this is one of the most significant disadvantages of a coaxial cable.

Fibre Optic Cables

I believe it is also obvious from the name. And everyone knows that when transmitting data, the speed is extremely high, particularly when compared to twisted and coaxial cables. Fiber optic cables are also used in medical equipment. It is made up of 1000s of strong fibres. As a result, this single strand of fibre is as fine as a human hair. As a result, it transmits data in the form of light shown in Figure 9. This is why it's called fibre optic cable. So, it's made up of an inner glass core surrounded by a glass-like substance with a low refractive index, and the fibre optic cable's core cladding and protective coating.

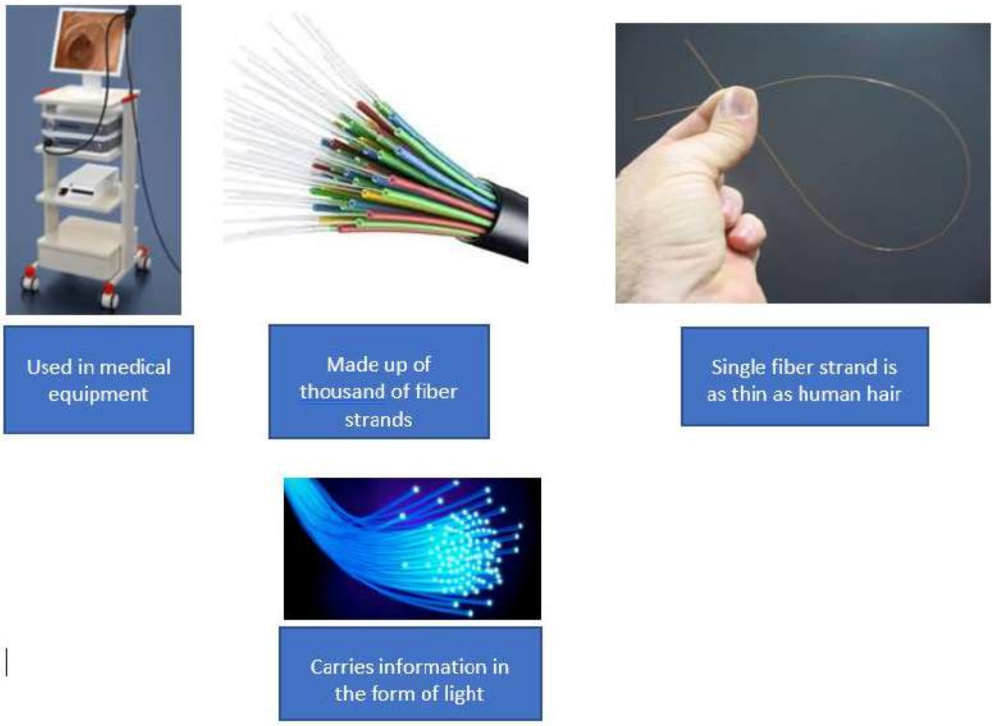

Used in medical equipment
Made up of thousand of fiber strands
Single fiber strand is as thin as human hair
Carries information in the form of light

Figure 9 Fibre Optic Cables

To comprehend the workings of a fibre optic cable, we must first comprehend the fundamental behaviour of light. As light travels through a medium, such as a medium or a container, the speed of the light changes. The refractive index is determined by dividing the speed of light in a vacuum by the speed of light in a medium Figure 10. So while this shift in pace is a fascinating phenomenon, it also brings with it a slew of new words to learn, one of which is refraction.

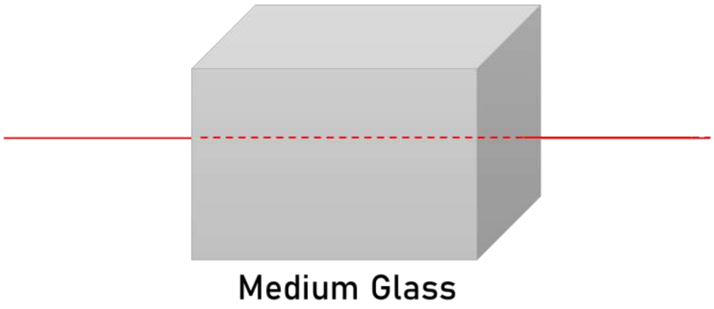


Figure 10 Medium Glass

As light travels through a prism, it bends instead of going straight, as seen in the Figure 11. When light travels from a medium with a high reflective index to one with a lower reflective index, it bends towards the interface, which is known as refraction.

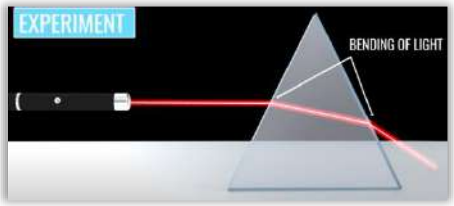


Figure 11 Refraction

Refraction is the explanation why a pencil appears bent in a glass of water, as an example. As a result, the fibre optic cable employs this basic strategy. So, as we raise the refractive index of the glass in real time, the light will bend more and more to the surface; as you increase the refractive index, the light will bend more and more to the surface; as you increase the refractive index, the light will suddenly come to the first medium as a pure reflection; as you increase the refractive index, the light will bend more and more to the surface; as you increase the refractive index, the light will suddenly come to the first medium as a pure reflection. A complete internal reflection is what it's called. So, rather than raising the refractive index, we should increase the incident angle in this situation, resulting in a critical angle.

Figure 12 Pencil appear bent in a glass of water

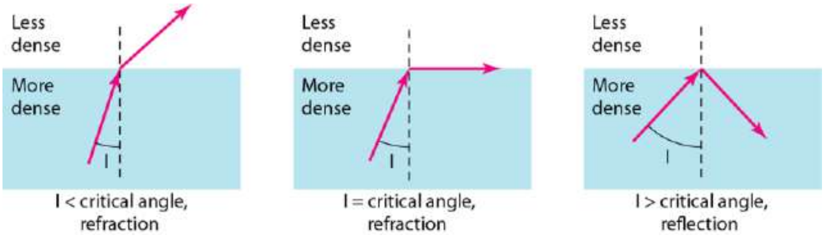


Figure 13 Bending of light ray

When the angle of incidence approaches the critical angle, the ray reflects and moves closer to the surface; as the angle approaches the critical angle, the light bends around the interface. If the angle of incidence is greater than the critical angle, the ray reflects and passes through the denser material again as seen in Figure 13.

Let's look at propagation modes now.

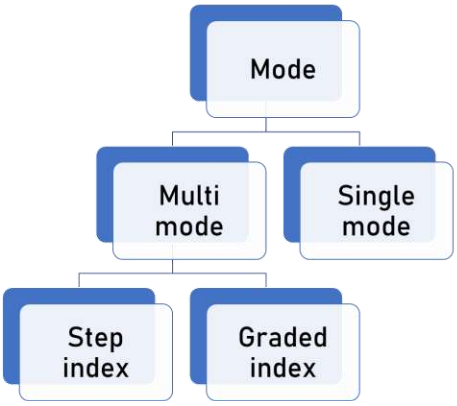


Figure 14 Propagation Modes

So, there are two modes available. One is multimode, while the other is single mode. Multimode is further subdivided into step and graded indexes. So, let's take a look at each one individually. So, what is a multimode name? It refers to the fact that multiple light beams from a single light source pass through the heart in different directions. However, the movement of these beams inside the cable is dependent on the structure; in many cases, the core density remains constant from the centre to the edges. A beam of light travels in a straight line through this constant density until it reaches the interface of the core and the cladding. At the interface, there are unexpected changes due to a lower density. As a result, the angles of a beam motion are alternated. Let's talk about this step index mode now.

In the case of step index, the suddenness of the transition means that there is an abrupt change in the signal, which leads to signal distortion as it travels through the fibre as seen in Figure 15.

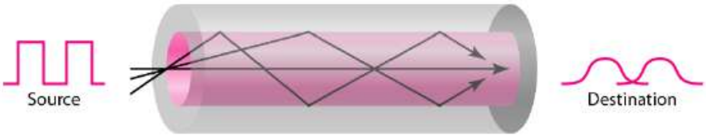


Figure 15 Multimode, Step Index

As you can see in the Figure 16, multimode graded index fibre reduces signal distortion as it travels through the cable. The term index here refers to a reflection index.

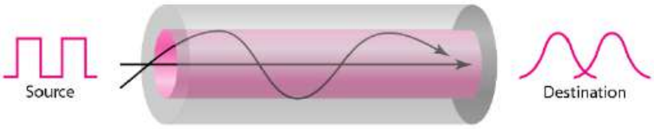


Figure 16 Multimode, Graded Index

Let's talk about single mode now. Single mode now employs a phase index fibre and a highly oriented light source to confine beings to a narrow range of angles, all of which are similar to the horizontal. As you can see Figure 17, since the diameter in single mode is so small, my light is only going in one direction, and there is only one straight line. Since the diameter is so small, it can't go up and down; instead, it can only move in one direction: up and down. As a result, single mode uses a phase index fibre and a highly oriented light source to confine a beam to a narrow range of angles all close to the horizontal.

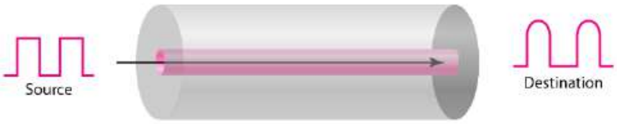


Figure 17 Single Mode

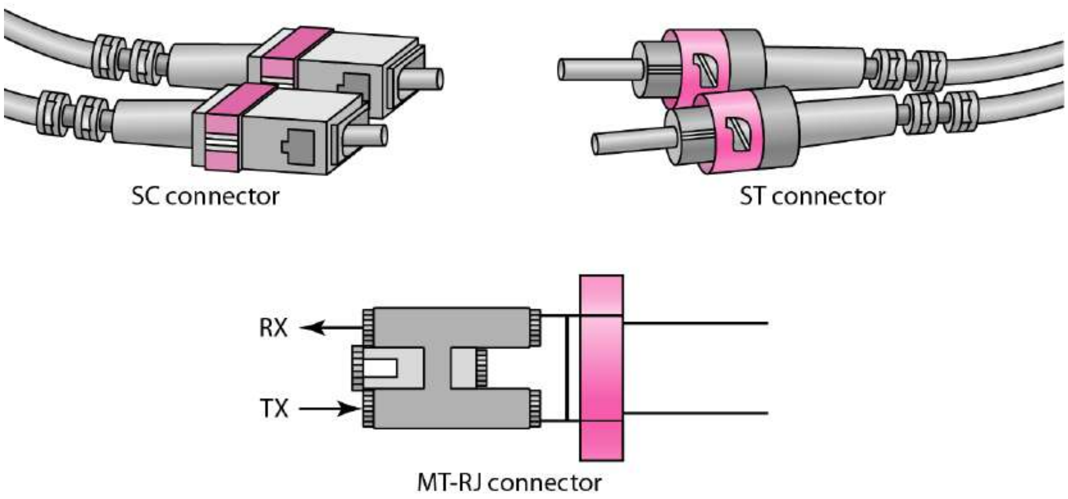


Figure 18 Fibre optic cable connectors

So, there are fibre optic cable connectors. There are SC connectors and ST connectors. These are the names of the connectors used in the five most popular scenarios.

Advantages of Fiber Optic Cable

- It is immune to electrical/mechanical interface.
- Highly suitable for harsh environment
- Secure transmission
- Used for broadband transmission.

Disadvantages

- Installation is difficult(glass is very fragile)
- Connecting two fibers is difficult.
- Connection loss
- Most expensive

Unguided Media: Wireless

Now let's talk about the unguided media that carry electromagnetic waves without the use of any physical conductor. Unguided means that we're not using cables to relay signals; instead, we're using waves that may be electromagnetic waves.

Types of unguided media: -

- Radio Waves
- Microwaves
- Infrared

Unguided media, as seen in the diagram, transport electromagnetic waves without the use of a physical conductor. As a result, signals are typically transmitted over an open space and are thus accessible to anyone with a platform capable of receiving them. As a result, unguided media will move from point A to point B in a variety of ways.

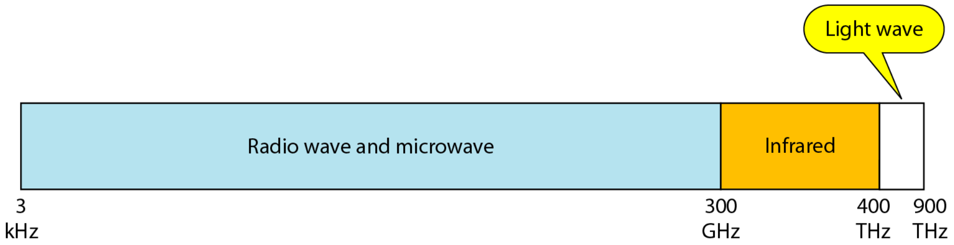


Figure 19 Electromagnetic Spectrum for Wireless Communication

Propagation Methods

Ground, sky, and line of sight propagation are seen in Figure 20 . As a result, in ground propagation, the antenna flies to the lowest part of the atmosphere, hugging the ground. As a result, the lowest frequency signals emitted from the transmitting antenna obey the curvature of the earth in all directions.

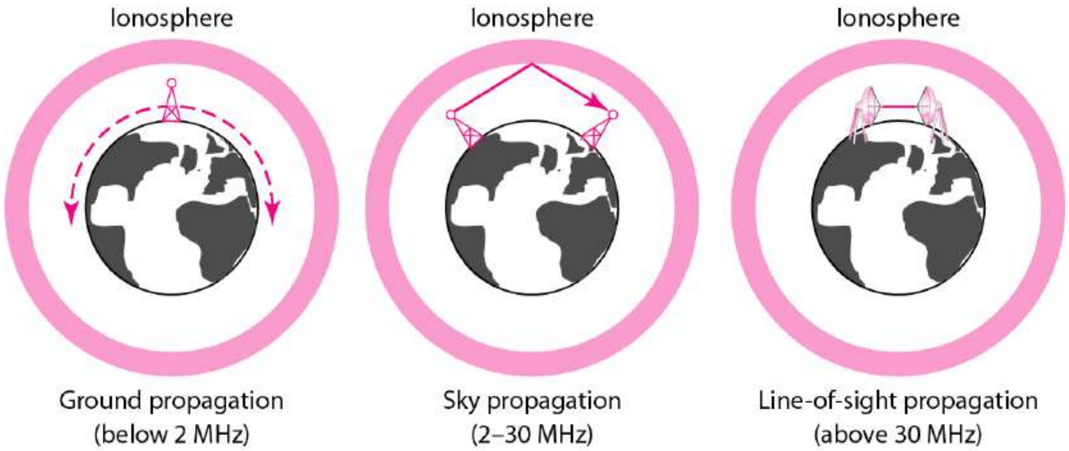


Figure 20 Propagation Methods

The amount of power in the signal determines the distance it can cover; the greater the power, the greater the distance it can cover. What about higher frequency radio waves that propagate upward through the ionosphere? They mirrored back to Earth there. As a result, this method of transmission allows for longer distances while using less fuel. As a result, very high frequency signals are transmitted in straight lines via line of sight propagation. As a result, antenna to antenna is what it means. The signal is transmitted antenna to antenna, whether they are tall apart or wherever they are.

Wireless Transmission Waves

As seen in the Figure 21 below, there are three forms of wireless communication.

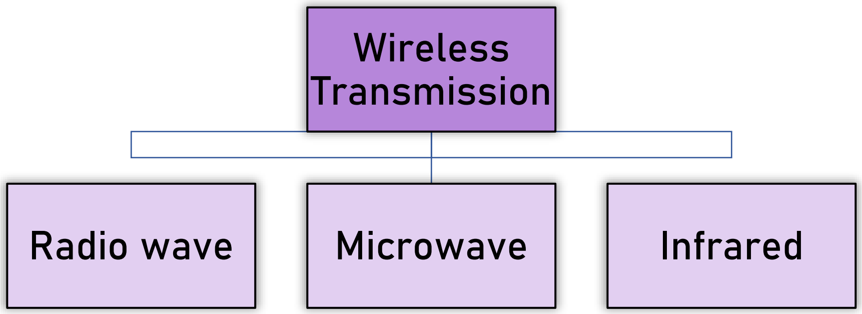


Figure 21 Types of Wireless Transmission

Radio waves

It's a method of data transmission that uses radio waves rather than copper or glass, so energy flows through the air. Radio, television, cellular phones, and other devices use radio communication in some way. Radio waves can pass through walls and even an entire structure. They can travel long or short distances depending on the frequency. Radio waves are used for multicast communications, such as radio and television, and paging systems. They can penetrate through walls. Highly regulated. Omni directional or directional antennas are used to broadcast radio waves depending upon band Figure 22. The transceiver unit, which is consisted of transmitter and receiver along with the antenna, determines the power of RF signal. Other characteristics of radio waves is that in vacuum all electromagnetic waves or radio waves travel at the same speed i.e., at the speed of light which is equal to 3 × 108 meter per seconds. In any medium this speed gets reduced and also becomes frequency dependent. In case of copper the speed of light becomes approximately two thirds of the speed of light.

- The basic features of the radio waves are that:
- They are easy to generate
- They have same velocity in vacuum
- They may traverse long distances
- They are omni directional
- They can penetrate building easily so they find extensive use in communication both indoor and outdoor
- They are frequency dependent. At low frequency they can pass through obstacles well but the power falls off sharply with distance from the source, as power is inversely proportional to cube of the distance from the source. At HF they travel in straight lines and bounce off obstacles.

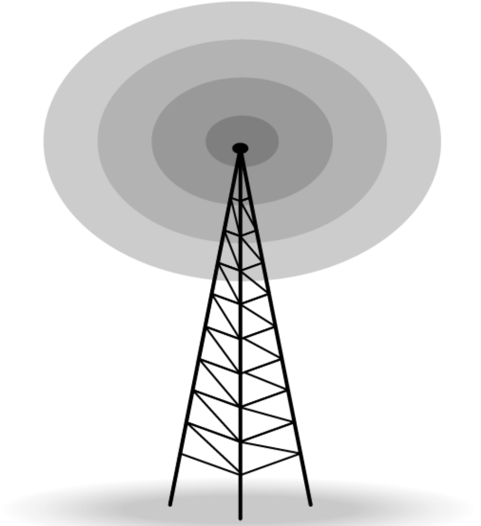


Figure 22 Omnidirectional Antenna

Advantages of Radio Wave

- 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

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

MicroWave

Microwaves are used for unicast communication such as cellular telephones, satellite networks, and wireless LANs. Higher frequency ranges cannot penetrate walls. Use directional antennas - point to point line of sight communications. Microwave radio, a form of radio transmission that uses ultra-high frequencies, developed out of experiments with radar (radio detecting and ranging) during the period preceding World War II. There are several frequency ranges assigned to microwave systems, all of which are in the Giga Hertz (GHz) range and the wavelength in the millimeter range. This very short wavelength gives rise to the term microwave. Such high frequency signals are especially susceptible to attenuation and, therefore must be amplified or repeated after a particular distance. In order to maximize the strength of such a high frequency signal and, therefore, to increase the distance of transmission at acceptable levels, the radio beams are highly focused. The transmit antenna is centered in a concave, reflective metal dish which serves to focus the radio beam with maximum effect on the receiving antenna. The receiving antenna, similarly, is centered in a concave metal dish, which serves to collect the maximum amount of incoming signal.

Unidirectional Antennas

Unidirectional antennas, which send signals in only one direction, are needed for microwaves. For microwave communications, two types of antennas are used: the parabolic dish and the horn antenna, as shown in Figure 23. The geometry of a parabola is used to build a parabolic dish antenna: Any line parallel to the line of symmetry (line of sight) reflects off the curve at different angles, resulting in a point called the focus where all the lines converge. The parabolic dish acts as a funnel, collecting a variety of waves and funnelling them to a single point.

A horn directed at the dish is used to broadcast outgoing transmissions. The microwaves collide with the dish and deflect outward, reversing the receipt route. A horn antenna resembles a massive scoop. Outgoing signals are broadcast up a stem (which resembles a handle) and deflected outward by the curved head in a series of narrow parallel beams. The scooped form of the horn collects received signals, analogous to a parabolic dish, and deflects them down into the stem. Microwaves are very useful when unicast (one-to-one) communication is needed between the sender and the receiver due to their unidirectional properties. They're used in mobile phones.

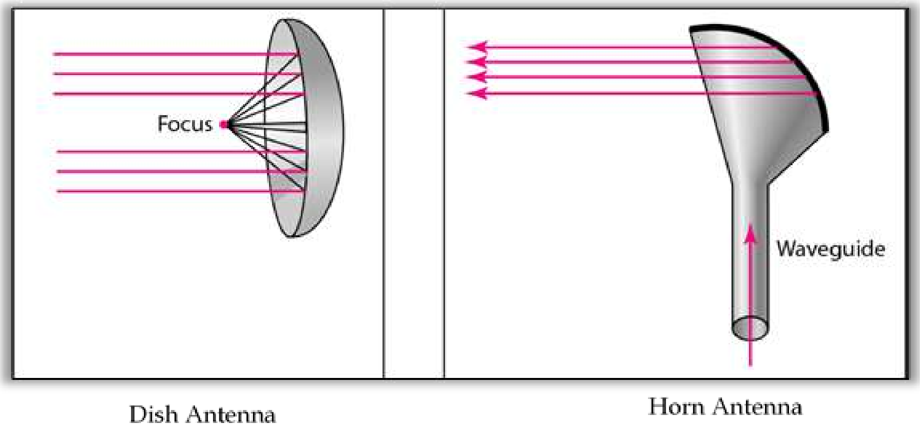


Figure 23 Unidirectional Antennas

Drawbacks

- They cannot pass through microwave buildings.
- Bad weather influences signal transmission.
- They are frequency dependent.

Infrared Signals

For short-range communication, infrared waves with frequencies ranging from 300 GHz to 400 THz (wavelengths ranging from 1 mm to 770 nm) may be used. Because of their high frequency, infrared waves are unable to penetrate walls. This beneficial feature avoids interference between systems; a short-range communication device in one room will not be affected by a system in the next room. When we use our infrared remote control, it does not interfere with our neighbours' use of the remote. This same property, however, renders infrared signals useless for long-range communication. Furthermore, we are unable to use infrared waves outside of a building because the sun's rays produce infrared waves that can disrupt contact.

- Infrared signals can be used for short-range communication in a closed area using line-of-sight propagation**.**
- Infrared communication is unreliable outside the building because the sun rays will interfere with the infrared waves.
- Infrared waves cannot penetrate the walls.

When we use our infrared remote control, it does not interfere with our neighbours' use of the remote. This same property, however, renders infrared signals useless for long-range communication. Furthermore, we are unable to use infrared waves outside of a building because the sun's rays produce infrared waves that can disrupt contact.

Switching techniques

We have several paths to transfer data from source to destination, and the switching technique will choose the best path. So, what is the best path for them to take to send data from source to destination? To make a one-to-one connection, a switching technique is used to link the systems. This is why we're going to look at switching strategies. As we want to transfer my packet from source to destination, which is the best path to take so that I can send my data or packet from source to destination, which is the best path to take so that I can send my data or packet from source to destination, which is the best path to take so that I can send my data or packet from source to destination, which is the best path to take so that I can send my data or packet from source to destination, which is the best path.

Switching Methods

The two broad level switching methods are seen in Figure 24.

So, switching approaches can be link based or connectionless.

- Connection oriented switching means that you must reserve resources before sending data.
- Connectionless switching means that you do not need to reserve resources and can begin sending data as soon as it becomes available.

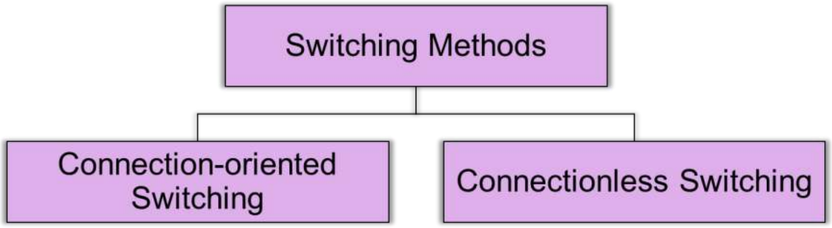


Figure 24 Switching Methods

5.2 Classification of Switching Techniques

So, let's look at the different types of switching. There are three types of switching: circuit switching, packet switching as shown in Figure 25.

Circuit Switching

The first is circuit switching, which is a switching strategy that creates a dedicated path between the sender and the receiver. A dedicated path ensures that no one else can use that resource for that period of time.

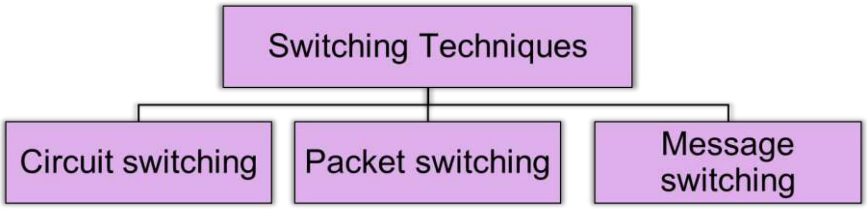


Figure 25 Switching Techniques

When you go to a restaurant and notice that the tables are already reserved, it is written on the wall. What does this mean? It means that no one else can use that table at that time because it has been reserved for a certain person, and only that person can use it. So that is what the word "dedicated" means. If the link is made, a dedicated route will remain in place until the teardown process. Circuit switching in a network functions in the same manner as it does in telephone networks. Circuit switching can be permanent or temporary. When a user tries to transmit data, voice, or video and a request, a signal is transmitted back to the receiver.

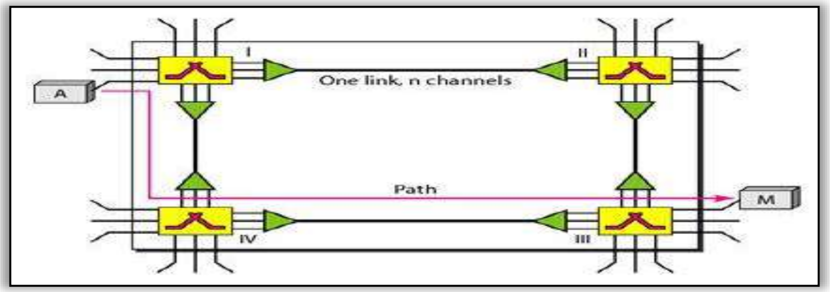


Figure 26 A trivial Circuit Switched network

The receiver then sends back an acknowledgment to ensure that the dedicated route is accessible. In public telephone networks, the data circuit switch is used after receiving the acknowledgment dedicated route transfers. It is the best example of voice communication transfer. As illustrated in the Figure 26. Each link is divided into three channels using the FdM, or TDM, and each link has three connections that operate in three phases.

Applications that use circuit switching may have to go through three phases:-

- Establishing a circuit.
- Transferring the data.
- Disconnecting the circuit

What do the terms "***establishing a circuit"*** and "establishing a process" mean? Creating the circuit entails create a phase Set up a phase means that if you want to submit data, you must first reserve the resources. Channels and turn buffers are examples of resources. switch processing time switch input outputs that must be committed during the data transmission. As you can see in Figure 26, A wishes to give information to m. So, first, A will notify switch four that a dedicated path exists between switch one and switch four, then switch four will create a dedicated path between switch four and switch three, and switch four will inform node m of A's purpose, and node m will agree that it is ready to receive the data.

The ***data transfer*** process will only function after that. Data transfer refers to a continuous flow of data from a source station to a destination station; continuous flow implies that the data will not be split into packets because my resources have been allocated. My data does not need to be divided into packets. As a result, data will flow continuously. And the first is the setup step, in which we reserve resources such as CPU bandwidth, CPU buffer, switch processing time, and switch input output ports, which are the resources I need. Then, once the acknowledgment is received, the source station will begin sending the data, which will be sent in a continuous flow from the source station to the destinations.

Disconnecting the process is the third choice. After the data has been successfully transferred, the circuit will be disconnected.

Important points

- circuit switching, are designed for the voice applications.
- The best suitable example of circuit switching is telephone.
- The connection is established using Dedicated path.
- Circuit switching takes place at the physical layer.

Efficency and Delay

Circuit switching has a lower performance than the other two forms of networks because it requires you to reserve resources. So keep in mind that whenever I mention reserving resources, I'm referring to a cost. To send data, I must first create a link. The total delay is due to the time required to create the link, move the connection, and disconnect the circuit; however, the delay in this type of network is very negligible because there is no waiting time because there is no connection; I have already formed a connection, so there will be no delay; the total delay is due to the time needed to create the connection, transfer the connection, and disconnect the circuit; however, the total delay is due to the time needed to create the connection, transfer the connection, and disconnect the circuit.. Otherwise, there is no delay.

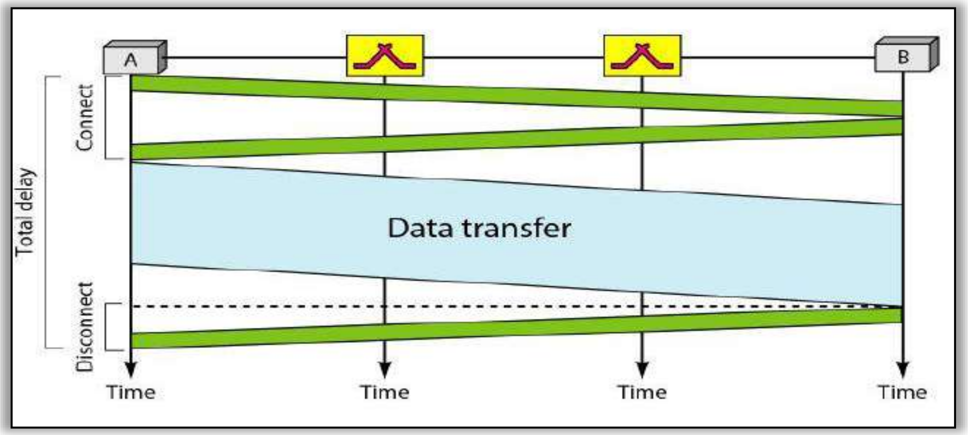


Figure 27 Delay in Circuit Switched Network

And I'm assuming that when you go to the restaurant, you won't have to wait for a table to become available because you've already reserved one. As a result, the time delay would be negligible. Only when linking, transferring, and disconnecting the circuit is there some time for delay.

As you can see in the Figure 27, the overall delay in the circuit switch network is due to the time required to establish the link, transfer data, and disconnect the circuit, so the delay in this time is very small.

Advantages of Circuit Switching

The benefits of circuit switching include the creation of a dedicated channel that no one else can use. And if I talk about dedicated, that means my packet loss will be very low, and if there is a dedicated channel, overhead will be very low because everyone is on the same path. If I want to get to a specific station and everyone is taking the same path. I'm also following the same paths because I don't have to think about which way to go, whether left or right, because I'm simply following the same direction. Improves data transfer rate significantly because there is no waiting period and no delay. As a result, the data transfer rate is strong here, which reduces data loss. If you want to go to a wedding but don't know where to go. You're just following path one after another, and some dedicated link is created for you, and you're just following the route, one after another, so data loss will be extremely rare here, and my data will not be lost. It improves delay in the data flow.

Disadvantages of circuit switching

- The waiting time lasts long, and there is no data transfer.
- Each connection has a dedicated path, and this gets expensive.
- When connected systems do not use the channel, it is kept idle.

Message Switching

As message switching is used, the entire message is treated as a data unit, and each node receives the entire message and buffers it before forwarding it to the next node. A message switching shown in Figure 28.

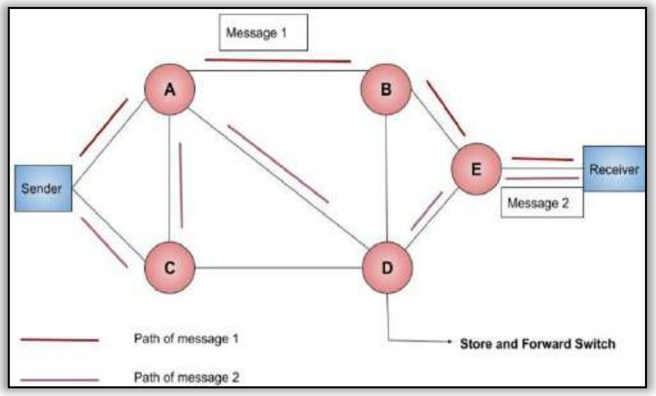


Figure 28 Message Switching

This form of network is known as a store and forward network since no dedicated route is established between the sender and the receiver. We are not allocating any resources between the sender and the recipient, implying that there is no dedicated route. It's a combination of dropping store and forward strategies.

Advantages of Message Switching

- Stores the message when the next node is not available
- Reduces traffic congestion.
- Data channels are shared by network devices.
- Manages traffic efficiently by assigning priorities.

Drawbacks of Message Switching

- Transit path needs enough storage.
- Message switching is very slow.
- Message switching was not a solution for streaming media and real-time applications.

Packet Switching

Datagram and virtual circuit packet switching are the two methods of packet switching. Let's look at what packet switching is and why it's called that because the data is split into packets. The packet switching technique is a switching method that sends the message all at once. However, it is broken down into smaller bits and sent separately. Messages are split into smaller pieces known as packets, with each packet having its own sequence number, so that we can place them in order when we receive them. The length of the packet may be set or variable; the switching information is contained in the header of each packet; the type of information header contains as shown in Figure 29. The source and destination nodes, as well as the intermediate number, are listed in the header. The sequence number is the intermediate number since, in this case, each packet is taking a different path to reach its destination. As a result, when the packets arrive, we must sort them into orders. So, who's going to assist in putting the packet together? That is your sequence number, and it is with the aid of that sequence number that we are placing the packets in the correct order. The message will be sent again if the packet is lost or corrupted.

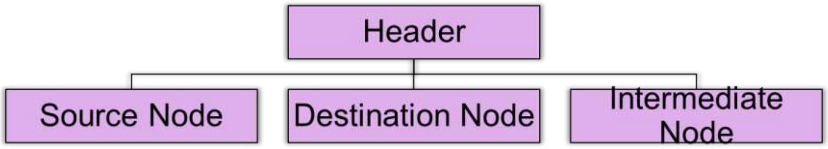


Figure 29 Details of Header

Modes of Packet Switching

So, what are the different types of packet switching? Packet switching may be connection-oriented or connectionless. In a connectionless network, packets can pass through the network using the shortest route possible, with all packets falling on separate paths to the destination being reassembled at the receiving end.

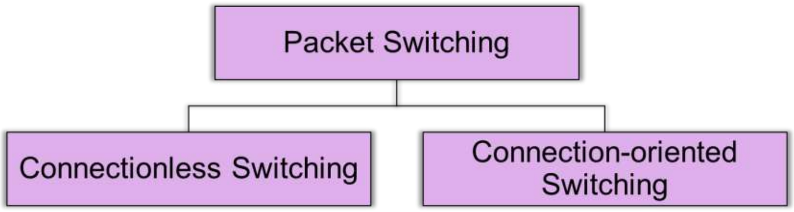


Figure 30 Modes of Packet Switching

Since this is the source and this is the destination, if I try to send data from source to destination, it will divide the data into little small packets. So that's why we'll reassemble the packet only if you're segmenting it. After you've divided the package, it's time to put it together. We'll reassemble the packets and reassemble the message when we receive them. If the correct order is met, an acknowledgment message indicating that the packet was successfully received is also sent.

Two Basic Forms of Packet Switching

There are two ways to treat packets. It may be a virtual circuit or a datagram switching as shown in Figure 31.

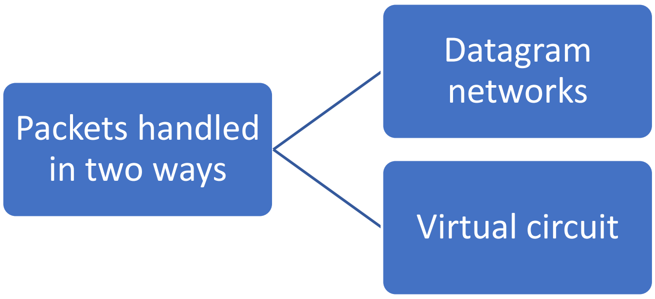


Figure 31 Forms of Packet Switching

But, first and foremost, let's talk about datagram networks.

Datagram network

It is a packet switching technology in which a packet is referred to as a datagram, hence the name datagram switching. Now, each packet contains destination information, which the switch uses to forward the packet to the next destination. Packets can now follow any practical path. So, in a datagram switching technique, the path is not set, i.e., it is not committed to the fact that you will only use this path to get to your destination. On the cloud, a variety of paths would be available. As I previously said, the modes of packet switching are connectionless and connection based, respectively, and datagram switching is a connection less switching. As a result, datagram switching, also known as connectionless switching, involves intermediate nodes making routing decisions in order to forward packets. As a result, datagram networks are also known as connectionless networks. What does connection less mean? It means that the switch does not store information about the connection state. As seen in the Figure 32 below. A is sending a packet of data, which is divided into packets, each of which has a sequence number and contains information about the source, destination, and sequence number.

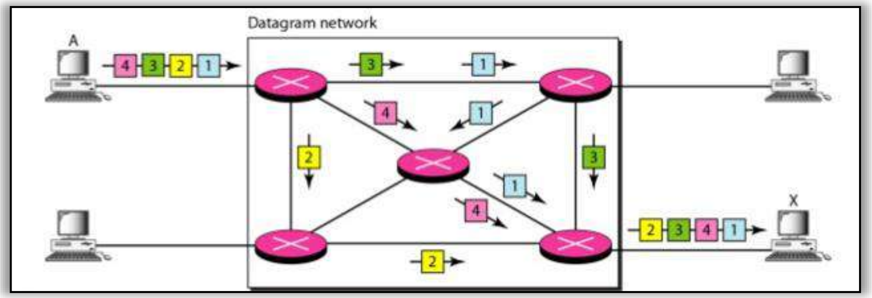


Figure 32 Datagram Network

So 1234, and as you can see here, each packet takes a different path to reach its destination, which is x. As a result, when they arrive at their destination, they are not in order 1432. This is the order, so we can place our previous message in our order by looking at their sequence number, which we can organise in the order routing table. Each switch has a routing table that is based on the destination address. Routing tables are dynamic and are modified on a regular basis if no new nodes are added to the network. Routing tables are dynamic and are updated on a regular basis if no new nodes are added to the network. As a result, the destination address now assists in getting packets to their intended destination by forwarding them to the next destination. So, in a datagram network, the destination address in the header of a packet stays the same during the entire journey of the packet performance of datagram networks is unquestionably better than a circuit switch network.

And I want to submit a packet every time. I don't need to set up a link or a phase, and I don't need to reserve resources because it costs a lot of money; resources are only allocated when there are packets to be transferred, which means resources are only reserved on demand. If you have a packet to send during datagram switching. At that time, if only you will submit the packet, which ensures that only resources will be allocated at that time. In the case of datagram switching, if you want to send a packet, resources will be allocated only on request. Since resources are not allocated in advance, there is a greater delay in a datagram network than in a virtual circuit network if you have a packet to send delay.

Virtual Circuit Networks

Since it is a combination of datagram and circuit switching, it is link driven. It is expensive to implement because if you want to transmit data, a global packet is transmitted to the network. Resources are allocated for time periods, and it is a highly efficient medium of transmission. It will also set aside all resources so that the packets can be sent to their destinations. Only if you have a package to send will the global packet be sent, which will reserve the resources as if you needed them. If you want to send a packet, you'll also need CPU bandwidth and a buffer to hold the packet for a period of time. As a result, in the case of a virtual circuit, a pre-planned route is formed prior to the message being sent. As seen in the Figure 33 below, call request and a call Accept packets are used to create the sender-receiver link. Paths are reserved in this case for the length of the logical relation. If I want to end the link, a simple request will be sent to end the connection request. So call request call accept indicates that a link has been established; if you have data to send, it will send the data transfer, and an acknowledgment will be sent indicating that my data has been received successfully; and finally, we can terminate the connection by sending a simple request.

Since a global packet is sent in this case as well, packets are sent to the recipient in the same order as they were sent by the sender. But that's how the money are set aside. As a result, every packet takes the same route to its destination. A virtual circuit is a secure network link. In each packet, there is no need for overhead.

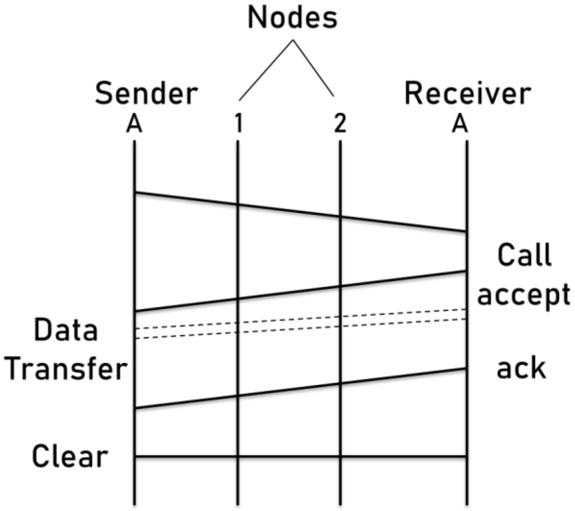


Figure 33 Working of virtual circuit

Advantages of Virtual circuit

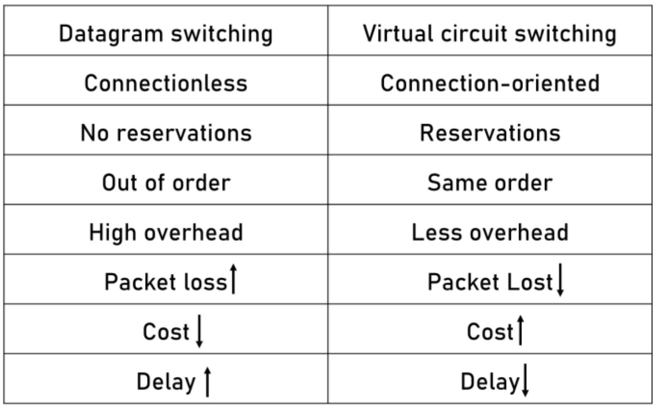
- Packets are delivered to the receiver in the same order sent by the sender.
- Virtual circuit is a reliable network circuit.
- There is no need for overhead in each packet.
- Single global packet overhead is used in virtual circuit.

Disadvantages of Virtual Circuit

- Virtual circuit is costly to implement.
- It provides only connection-oriented service.
- Always a new connection set up is required for transmission.

5.3 Difference between Datagram switching and virtual circuit switching

The following Figure 34 illustrates the distinction between datagram switching and virtual circuit switching.



Datagram switching	Virtual circuit switching
Connectionless	Connection-oriented
No reservations	Reservations
Out of order	Same order
High overhead	Less overhead
Packet loss↑	Packet Lost↓
Cost↓	Cost↑
Delay↑	Delay↓

Figure 34 Datagram switching vs Virtual circuit switching

Networking Devices

As hardware devices that are used to link computers, printers, fax machines, and other electronic devices, you will be able to learn about various networking devices and understand the functioning and functionality of various networking devices. These devices send data over the same or a different network in a fast, safe, and accurate manner. As a result, we must investigate all of the devices that assist in the transfer of data from source to destination. Those are what we refer to as networking devices.

Use of Networking Devices

Figure 35 depicts the number of networking devices in operation.

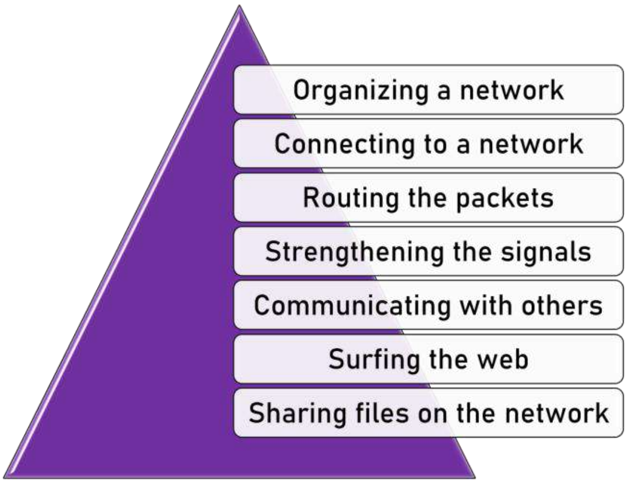

Organizing a network
Connecting to a network
Routing the packets
Strengthening the signals
Communicating with others
Surfing the web
Sharing files on the network

Figure 35 Use of Networking Devices

So, networking devices are the devices that are used for organizing a network, linking our network, routing the package, and improving the signals because my signal loses strength when it travels from source to destination.

5.4 Types of Networking Devices

Types of networking devices are shown in Figure 36.

1. Modem
2. Hub
3. Switch
4. Router
5. NIC
6. Repeater
7. Bridge
8. Gateway

Figure 36 Types of Networking Devices

Modem

Modem is short for "Modulator-Demodulator. An analog signal is converted to a digital signal, and digital signals are converted to analog signals. The modulator is used to transform digital signals to analog signals. The demodulator modem is used to communicate over telephone lines when an analogue signal is converted to a digital signal as seen in Figure 37.

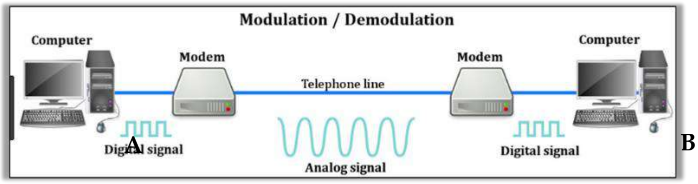


Figure 37 MODEM

How fast the modem can transfer and receive the data?

- At Slow Speeds, modems are measured in baud rate.
- At Higher Speeds, modems are measured in bits per second.
- Higher the speed, the faster you can send and receive data over the network.

Working of MODEM

So, as you can see in the diagram, if A wants to send data to B, here's an example. However, we must first use the modem before transmitting the signals to the telephone lines. So, a modem can transform digital data from a device to an analogue signal so that it can be sent over a telephone line. Then there's the fact that those cables can only accept analogue signals. Finally, we want to give a signal to machine B, which will only accept digital data. So, once again, we must use a modem to transform this analogue signal into a digital signal, which is known as a demodulator.

Which Layer Modem Works in OSI Model?

Modem works at Physical Layer of the OSI model.

HUB

The hub is a networking system that connects different types of cables to centralized network traffic via a single connecting point hub, which also serves as a central controller for all connected devices as a master controller. Hub serves as a master controller. It actually goes to whatever communication is taking place, and it will go through the hub, as seen in this example. The hub now serves as a central controller for all devices that want to connect.

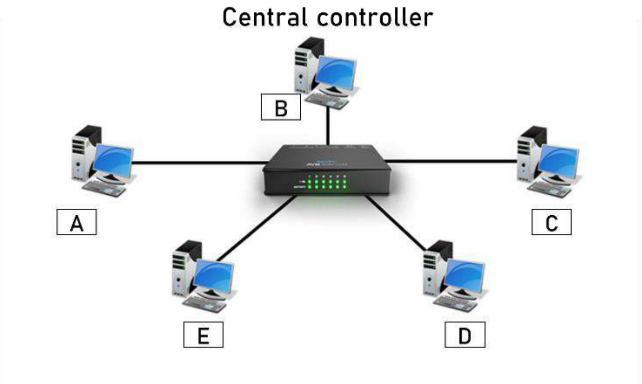


Figure 38 Hub

If A wants to send data to C, instead of sending the data directly to C, A can send the data to hub. As a result, the hub serves as a central controller, and no devices are attached directly to one another. Hub will still broadcast data, which means it will send data to C, E, and D. As a result, the hub will still broadcast data, which means it will send information to all connected devices. Following that, even the C's acknowledgment will be sent to all of the units. To begin, A sends data to C, which will arrive at the hub, which will broadcast the data, and C will accept it. Other nodes B, E, and D will search the header information and determine that the packet is not for them, so they will discard it. Then C will send the acknowledgment, and since Cs cannot send acknowledgements directly to A, C will send to the hub again, and the hub will broadcast the data again because this is a hub property. So it will give to B, E, and D, as well as A, with A accepting and the rest of the devices discarding the packet. To link topology segments of land and track network traffic, a hub with multiple ports is used. It organises and monitors all data sent and received between computers. Now, in order to prevent data collisions, we use the CSMA CD protocols, so if A sends data and c sends data at the same time, the data will be lost. When hub was broadcasting and C was sending the message. So, when A has already sent data to C and C has begun sending data, when the hub broadcasts the data, there will be a collision. So, to prevent these collisions, we use a CSMA CD technique, which you will understand and which will give you all of the details of this lecture when it comes up in the future lectures.

Types of HUB

Active hubs, passive hubs, and intelligent hubs are the three types of hubs as show in Figure 39. In the case of an active hub, it performs the same functions as a passive hub, but before forwarding the data signal, it amplifies it. As a result of this added functionality, active hub is also known as a repeater, as it amplifies the signal.

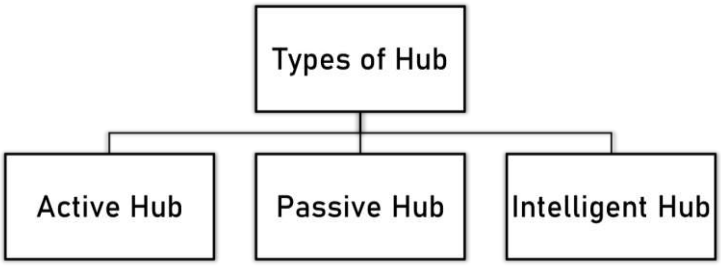
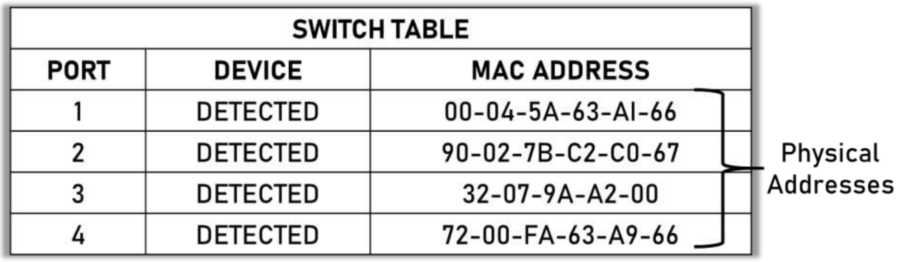


Figure 39 Types of HUB

The passive hub comes next. It transmits data signals in the same format in which they are received. It has no impact on a signal in any way. It operates in the same way as an active hub, but it also has remote control capabilities. They also offer networking devices the ability to use different data rates. It also allows an administrator to track traffic passing through the hub and configure each port in the hub, as well as determine which layer hub is active. As a result, the OSI physical layer and the TCP IP model are used by the hub.

Switch

The switch is a multiple LAN connecting system that receives data packets from multiple input ports and forwards them to a particular output port. When opposed to a hub, a switch is a more intelligent system. This is what a hub does: it takes data and sends it to all of the ports that are connected to the hub, while a switch, which is an intelligent system, receives data from one of the devices but only sends it to the specific intended recipient to whom it belongs. Let's say A wants to give some information to B. As a result, A will send data to switch, and switch will send data to C exclusively. It will not broadcast the data, but the way they are transmitting it, a switch will actually learn the physical addresses of the devices that are connected to it and store these physical addresses in a table called a MAC address.



SWITCH TABLE		
PORT	DEVICE	MAC ADDRESS
1	DETECTED	00-04-5A-63-AI-66
2	DETECTED	90-02-7B-C2-C0-67
3	DETECTED	32-07-9A-A2-00
4	DETECTED	72-00-FA-63-A9-66

Physical
Addresses

Figure 40 Switch Table

We're using a switch, which has a MAC address, which is the system's physical address. That is why we call switches intelligent devices because of how they send data to a specific recipient to which it belongs, using a MAC address to send data only to the computer to which your data belongs. But these are your MAC addresses, aren't they? 00,04,5A,63,AI,66. This is your MAC address, which is 12 digits long.

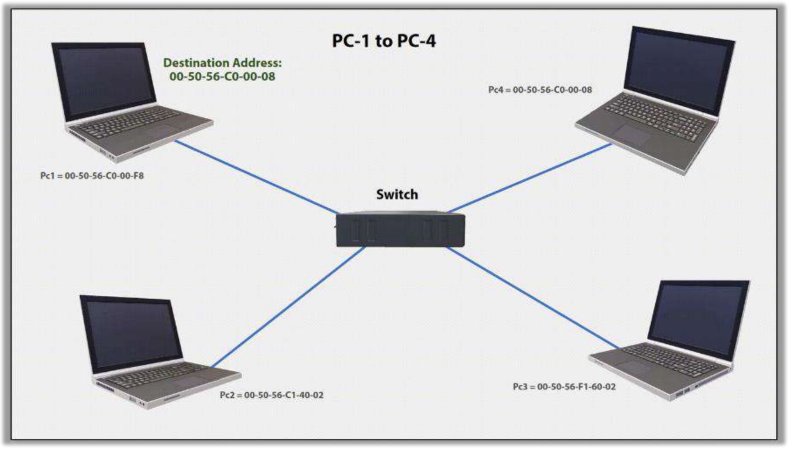


Figure 41 Switch Example

As you can see Figure 41, PC1 needs to transmit data to PC4, so PC1 sends the data to the switch, but the switch does not broadcast the data like the hub did; instead, the switch can only send the data to the intended receiver, which is PC4, since the MAC address is written on the header. As you can see, PC1 needs to transmit data to PC4, so PC1 sends the data to the switch, but the switch does not broadcast the data like the hub did; instead, the switch can only send the data to the intended receiver, which is PC4, since the MAC address is written on the header.

As a result, the MAC address is included in the post, and the data is sent to PC 4 via the MAC address switch. Only the intended recipient will receive it. As a result, the switch is aware of the MAC addresses of all systems connected to it. Now, more than ever, switch are favored over hubs because they eliminate unwanted traffic on the network.

Various Switching Methods

In the Figure 42 Different methods of switching below, various switching strategies are depicted.The first is the cut-through technique. The next choice is to use the store and forward process.

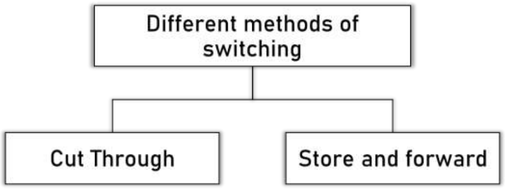


Figure 42 Different methods of switching

Cut Through

If I use the cut through process, the data will be split into small packets whenever I send it. In the case of the cut through process, the data will now be forwarded by the switch. The turn forwards the packet as soon as it arrives. As a result, each packet will take a different route to reach its destination. It does not wait for all of the data to arrive in a cut through. It simply forwards the packet as soon as it arrives. As a result, packet loss is more likely in the case of cut through and packet loss.

Store and forward

First and foremost, the switch waits for all of the data to arrive at the destination successfully before forwarding the packet to the intended recipient, then it will wait for all of the data to arrive before forwarding the packet to the destination to the intended recipient, so it will wait for all of the data to arrive before sending the packet to the destination. Packet failure is extremely unlikely in the case of store and forward strategies because it waits for all of the data to arrive before moving on to the next.

Difference between Hub and switch

Hub	Switch
Only detects that a device is physically connected to it.	Can detect specific devices that are connected to it because it keeps a record of MAC ADDRESS of those devices.

Similarity Between Hub and Switch

- Both are used to share data within a local area network, such as a home or business network.
- Both are not accustomed to exchanging data with each other outside of their own network, such as over the internet.
- Since a computer must be able to read IP addresses in order to exchange data or route data outside their own network or to another network, such as the internet.
- switch and hub Both may not read the IP address; we need a router to read the IP address. As a result, data is routed from one network to another based on IP addresses.

Router

When data packets are sent, the router inspects the IP address to decide whether the packet was intended for its own network or for another network. The router is particularly the gateway offer network. Example of router is shown in Figure 43.

Figure 43 Router

It will allocate to the same receiver if it fits within the same network. If it isn't intended for the same network, don't use it. It is intended for use in another network, and it will transmit data to that network using IP addresses, as routers do. Routers bind two or more topologically related or dissimilar networks, such as LANs and WLANs. It will share the available bandwidth through a network of computers. As a hardware firewall, it offers better security against hacking. Routers are smart enough to figure out the shortest and quickest path from source to destination, which is the easiest way to send data from one location to another.

So, if you want to send data from source to destination, routers can assist you in finding the shortest path. There are a variety of routes to choose from. So, which path is the fastest, and which one is the shortest? Such that my packet arrives at its destination in the shortest time possible. As an example, the shortest route from node 11, N11, to node N3 is found using the router as shown in Figure 44. When a data packet is forwarded to its final destination, the linked routers are registered.

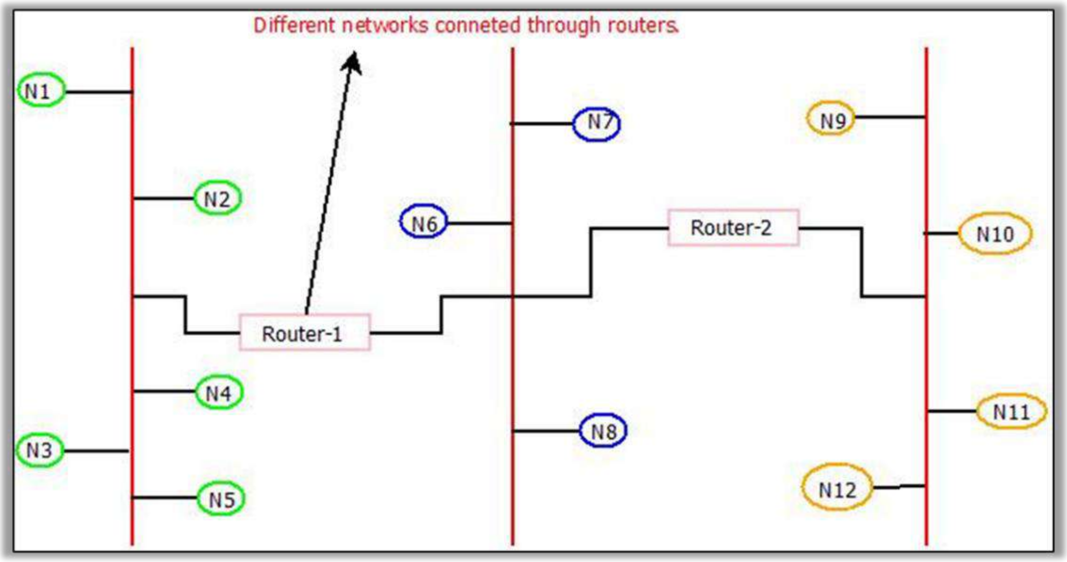


Figure 44 Description router

As a result, these records are held in a database table called the routing table. So, a routing table can be designed statistically or dynamically, so the router knows which route to take to send the data since it is updated on a regular basis. If any new networks access the network router, the destination addresses are still modified. As a result, routers are aware of each system's IP address. The IP address is the unique identifier for each device. So, routers know what the next hop is, where data will be transferred, and how long it will take to get data from source to destination. One of the functions of the router is to link many small networks into a larger network. By linking small networks together, we can create a larger network. So, one of the most popular routers, and one that is very familiar nowadays, is a wireless router, which allows a user to connect easily without installing any cables, and they also allow a user to connect easily without installing any cables, so if we are talking about wireless implies that we are not using any cables and can connect wirelessly.

Important points

- ***Which Layer Router Operates?***

  Operates at **network layer** of OSI model.

- When it comes to the network layer, two things must come to mind: the IP address and the router.
- the router read the system's IP address.
- The router can use this information to determine the best path for your packets or data to travel from source to destination in the shortest time possible.

Network interface card.

It is a circuit board or a card that allows computers to communicate over a network through cables or wirelessly through a small antenna network Figure 45. It connects to a network via cables cat5, coaxial cable, fibre optics, etc. When someone tells you the other term for a network interface card, it's a LAN adapter, network adapter, or network card. You may also use the terms LAN adapter, network adapter, or network card to describe your device.

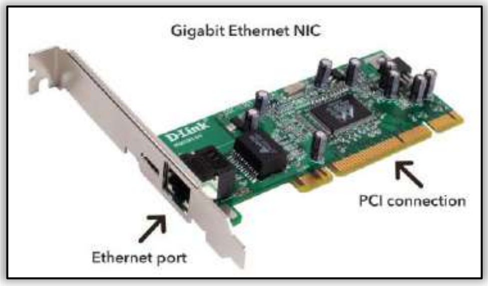


Figure 45 Network Interface Card

Types of NIC

The wired network interface card and the wireless network interface card are the two types of network interface cards.

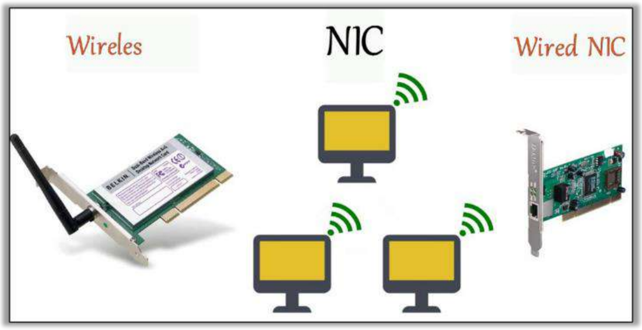


Figure 46 Types of NIC

Wireless NIC

Wireless NIC means in laptop, we are using a NIC card, it can be connected through cable. Also, but if it is a wireless NIC it can connect through the wireless modem. It works on the radio frequency.

Wired NIC

A wired network is one of the most popular wired configurations. Ethernet cables are used to transmit data between connected PCs in most wired networks. A single router will link all the computers in a small wired network. Multiple routers or switches connected to each other are common in larger networks. A cable modem, T1 line, or other form of Internet connection is usually connected to one of these devices. Wired may also apply to peripheral devices. Since several keyboards and mice are now wireless, the term "wired" is frequently used to refer to input devices that attach to a USB port as seen in Figure 47. Monitors and external hard drives use cables as well, although they are rarely referred to as wired devices because wireless alternatives are rarely available.

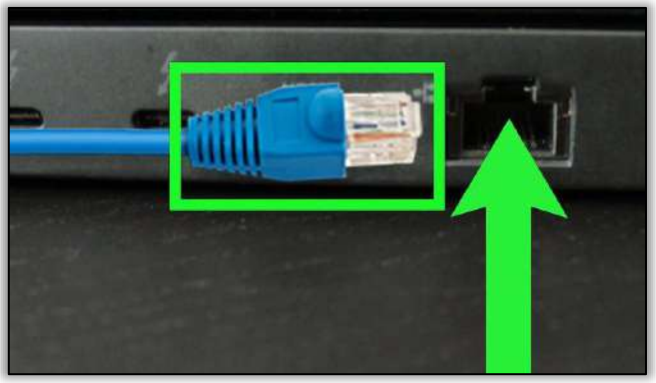


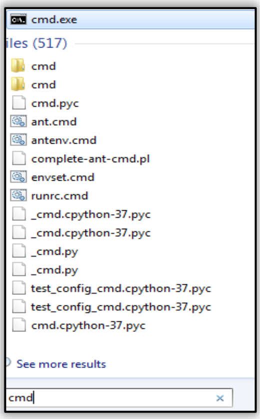
Figure 47 Wired NIC

- On desktop computer, we use wired NIC. If NIC is wired then connect through RJ45 port. By connection cable with RJ45 you communicate with internet.

How to check whether NIC card is working or not?

It's a straightforward procedure.

The first move is to go to the start menu, select Run, and type CMD into the command prompt that appears in Figure 48.



cmd.exe

iles (517)

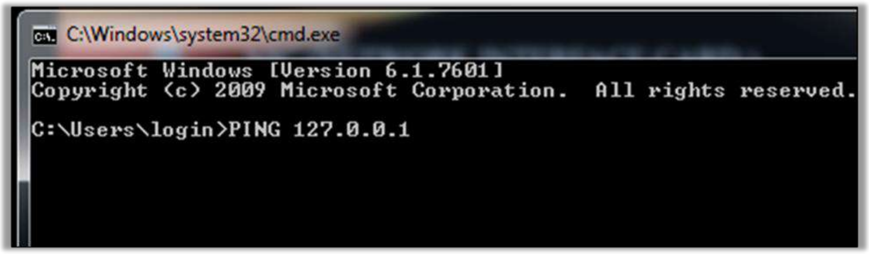
cmd
cmd
cmd.pyc
ant.cmd
antenv.cmd
complete-ant-cmd.pl
envset.cmd
runrc.cmd
_cmd.cpython-37.pyc
_cmd.cpython-37.pyc
_cmd.py
_cmd.py
test_config_cmd.cpython-37.pyc
test_config_cmd.cpython-37.pyc
cmd.cpython-37.pyc

See more results

cmd

Figure 48 Open command prompt

In that case, when the company manufactures the NIC, it assigns the NIC, which is a network interface card, a loopback address in phase two. I'd like to point out that we're writing PING space 127.0.0.1 as seen in Figure 49.

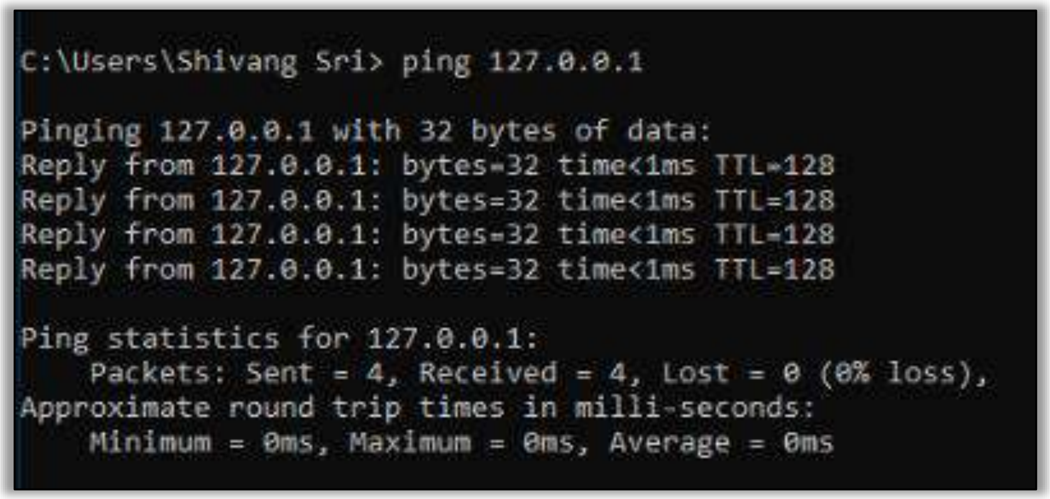


C:\Windows\system32\cmd.exe

```
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation.  All rights reserved.

C:\Users\login>PING 127.0.0.1
```

Figure 49 To check loopback address



```
C:\Users\Shivang Sri> ping 127.0.0.1

Pinging 127.0.0.1 with 32 bytes of data:
Reply from 127.0.0.1: bytes=32 time<1ms TTL=128
Reply from 127.0.0.1: bytes=32 time<1ms TTL=128
Reply from 127.0.0.1: bytes=32 time<1ms TTL=128
Reply from 127.0.0.1: bytes=32 time<1ms TTL=128

Ping statistics for 127.0.0.1:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 0ms, Maximum = 0ms, Average = 0ms
```

Figure 50 Reply of loopback address

So if you get a response, it means everything is fine as seen in Figure 50.

LAN Standards

Different LAN standards are:

- ISO
- ANSI
- EIA/TIA
- IEEE

But there are several LAN standards that are developing the standards that we are discussing on the NIC that we are publishing, the first of which is the ISO. So this form of organisation describes flow management, promotes flow management, and checks and promotes the consistency of the norm. The ANSI American National Standard international organisation is the next LAN standard. This is the organisation that governs all cards relevant to NIC network equipments, regulations, and quality control. The next one is your EIA, TIA, which is used when we do wired communication via cable and link one device to another via cable. Then there are two colour norms that we use. So EIA and TIA agreed on colour coding, for example, if I wanted to connect two identical devices, such as one PC to the other PC, we'd use a crossover cable. If I want to communicate, I'll use a PC with a router, which means those are two different devices. In that case, the straight cable patch straight cable would be used. So, when do you use a straight cable and when should you use a crossover cable? These are the criteria that EIA TIA decides on. The IEEE is the next organisation, which develops all communication standards. The 802.11 standard was the first land standard. 802.1, 802.2, 802.3 45678, and so on. It is essentially a correspondence that is decided by 12 bits. So latest one is 802.17 is there.

So now, a MAC address for the network interface card that we're discussing is divided into two bits, one of which is known as the manufacturing ID and the other as the serial number as seen in Figure 51.

00-04-AC	F3-1C—D4
Manafacture ID	Serial number

Figure 51 Parts of MAC Address

When a hacker tries to steal data from a target PC, the MAC address is punched on the network interface card, just like students. ISP stands for internet service provider, or security department. Cybercrime can be used to track down the intruder, and the serial number can be used to pinpoint their location. At this time, the individual is present at which site. As a result, we are able to track down the hacker using his MAC address.

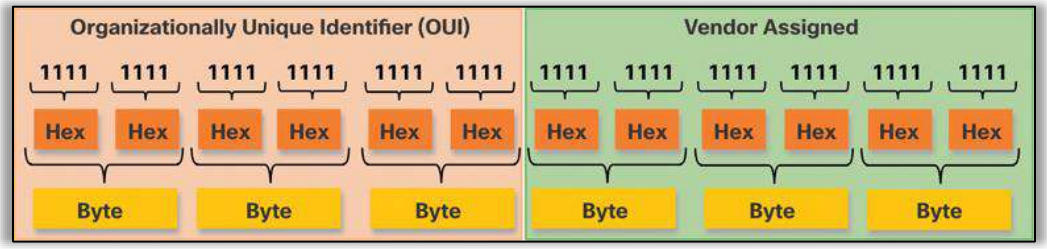


Figure 52 MAC Address format

A MAC address is a hexadecimal number of 12 digits. As a result, when you convert this to binary. As a result, it has been reduced to 48 bits. As a result, the MAC address is a long-term address. A physical address is another term for it. It is difficult to change this permanent address in practice; however, you can change the MAC address virtually with the aid of a variety of applications. Since a single hexadecimal digit represents four binary bits, an Ethernet MAC address consists of a 48-bit binary value. Hexadecimal is used to label an Ethernet address. As seen in this Figure 52 As a result, 12 hexadecimal values can be used to represent a 48-bit Ethernet MAC address.

Which Layer NIC Operates?

Operates on physical and data link layer of OSI model.

Repeater

The term "repeater" refers to a device that increases the signal's intensity as it passes from source to destination. So it's likely that your signals begin to lose energy; in fact, they begin to lose energy. As a result, we'll need an amplifier or other repeaters to increase the signal's intensity or improve it.

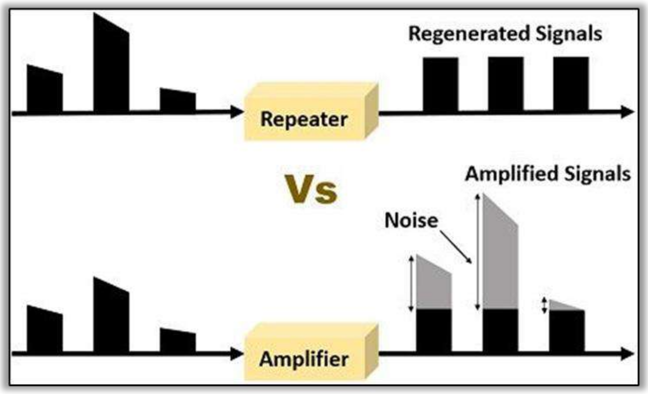


Figure 53 Repeater

It operates at the physical layer of the OSI model, and it's ideal for long-distance networks. The bus topology's main benefit is that it removes unnecessary noise from incoming signals, which is why my signal was losing energy. If you wish to receive the exact signal that is being sent, you must also give the exact signal. As a result, we'll have to use the repeater repeater.

Bridge

A bridge is a networking system that links two or more LANs together, treating them as a single network as seen in Figure 54. So, this is what distinguishes it from a router; in a router, we link two different networks together and they act as a separate network.

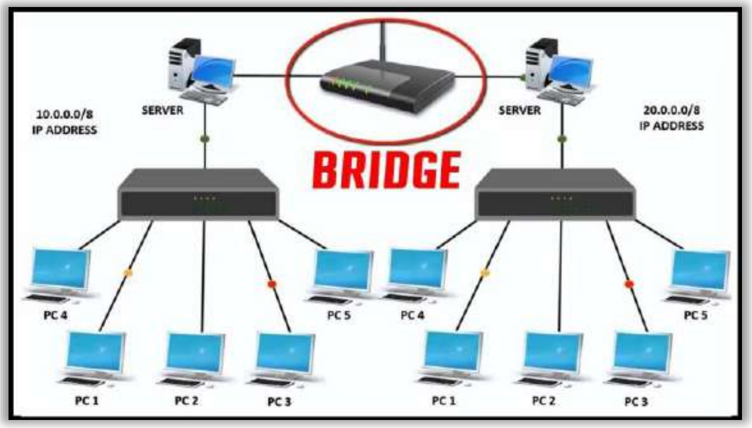


Figure 54 Bridge

A bridge is a networking system that links two or more LANs together, treating them as a single network. So this is what distinguishes it from a router; in a router, we link two different networks together and they act as a separate network. Multiple networks are linked together in this situation, but they are considered as a single network. As a result, bridges are used when the number of LANs grows and network traffic becomes too much to handle. So we use the bridge to divide the LAN network traffic into segments and thus reduce the network traffic. As a result, a bridge will send data over two separate protocols, such as Ethernet and Token Bus. It examines the frame's MAC address. If they fit, the data is forwarded; if they don't, the frame is discarded.

At which layer bridge operates?

The OSI model's data link layer is where bridge operates.

Gateway

A gateway is a computer network node (router) that serves as a critical stopover point for data on its way to or from other networks. We can connect and send data back and forth thanks to gateways. Without gateways, the Internet will be useless to us (as well as a lot of other hardware and software). The gateway is the device in an office that routes traffic from a workstation to the outside network that serves Web pages. The Internet Service Provider, which gives you access to the entire Internet, is the hub for basic Internet connections at home.

A node is essentially a physical location where data is halted for transport or reading/use. (A node is a device or modem; a computer cable is not.) A few node notes are as follows:

- A gateway or a host node may serve as a stopping point on the Internet.
- A node is a machine that manages the traffic that your Internet Service Provider (ISP) receives.

Your gateway is the modem (or modem-router combo) your ISP offers so you can connect to their network if you have a wireless network at home that allows your entire family access to the Internet. On the other hand, the device that your Internet Service Provider (ISP) uses to monitor all of the data traffic it receives and sends out is a node. When a computer server serves as a portal, it also acts as a firewall and proxy server. Unwanted traffic and outsiders are kept out of a private network by a firewall. A proxy server is software that sits between the applications you use on your computer (such as a Web browser) and the computer server that supports your network. The proxy server's job is to ensure that the real server can handle the request.

Keywords

Bandwidth: Refers to the range of frequencies assigned to a channel.

Bounded Media: Refers to the wired transmission systems that employ physical media, which are tangible.

Coaxial Cable: It is a very robust shielded copper wire two-conductor cable in which a solid center conductor runs concentrically (coaxial) inside a solid outer circular conductor.

Frequency Spectrum: Refers to the range of frequencies being supported by a particular transmission medium.

Gauge: Gauge is a measure of the thickness of the conductor.

Graded Index Multimode Fiber: In the case of a graded index multimode fiber, the index of refraction across the core is gradually changed from a maximum at the center to a minimum near the edges, hence the name graded index.

Monomode/Singlemode fiber: This has a thinner inner core. In this case, the core diameter of about 9 µm is much closer in size to the wavelength of light being propagated, about 1.3 µm. This limits the light transmission to a single ray or mode of light to propagate down the core of the fiber.

Multimode Fiber: The core diameter is relatively large compared to a wavelength of light.

Optical Fiber: Optical fiber carries the transmitted information in the form of a fluctuating beam of light in a glass fiber rather than as an electrical signal on a wire.

Cable Modem: It works on the principle of modems and provides access to data signal sent through the cable television infrastructure.

Cell Site: A circular geographical area that handles cellular phones within its defined physical boundary.

Microwave Radio: It is a form of radio transmission that uses ultra-high frequencies.

Radio: A technique where data is transmitted using radio waves and therefore energy travels through the air rather than copper or glass.

Review Questions

1. Describe how satellite communication is different from radio broadcast?
2. Explain different types of networking devices along with advantages and disadvantages.
3. What are the different transmission mediums over which data communication devices can provide service?
4. What are the major limitations of twisted pair wire?
5. Describe how satellite communication is different from radio broadcast?

Self Assessment

1. Which is the smallest unit amongst the following with reference to the ATM-

   a) transmission path

   b) virtual path

   c) virtual circuit

   d) all are of the same size

2. A device that provides a central connection point for cables is –

   a) Switch

   b) Hub

   c) Gateway

   d) Proxy Server

3. A device that helps prevent congestion and data collisions –

   a) Switch

   b) Hub

   c) Gateway

   d) Proxy Server

4. Transmission media directly controlled by

   a) Physical layer

   b) Data link Layer

   c) Network Layer

   d) Session Layer

5) guided media provides a conduit from one device to another, includes

   a) twisted pair cable

   b) fiber optic cable

   c) coaxial cable

   d) All of above

6) RG-59 is used in

   a) radio

   b) thick ethernet

   c) thin ethernet

   d) cable tv

7) A local telephone network is an example of a _______ network.

a) Packet switched

b) Circuit switched

c) Bit switched

d) Line switched

8) Most packet switches use this principle ____________

a) Stop and wait

b) Store and forward

c) Store and wait

d) Stop and forward

9) In _________ systems, resources are allocated on demand.

a) packet switching

b) circuit switching

c) line switching

d) frequency switching

Answers

1 (c)	2 (c)	3 (a)
4 (a)	5 (d)	6 (d)
7 (b)	8 (b)	9 (a)

Further Readings

Andrew S. Tanenbaum, Computer Networks, Prentice Hall

Behrouz A. Forouzan and Sophia Chung Fegan, Data Communications and Networking, McGraw-Hill Companies

Burton, Bill, Remote Access for Cisco Networks, McGraw-Hill Osborne Media

McGraw-Hill Osborne Media

Rajneesh Agrawal and Bharat Bhushan Tiwari, Computer Networks and Internet, Vikas Publication

Dr. Rajni Bhalla, Lovely Professional University

Unit 06: Data link layer – error control and flow control

CONTENTS

Objectives

- what is block coding
- learn different types of error detection mechanism
- Different types of framing techniques

Introduction

When transmitting data from source to destination, it is likely that the data will get compromised during transmission. Your bits shift from one to zero or zero to one between transmissions as seen in Figure 1.

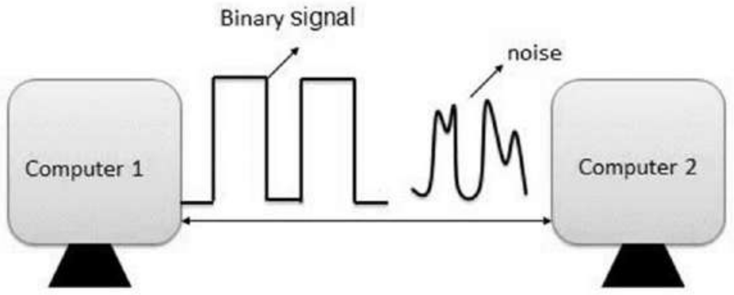


Figure 1 Error

So, how do we prove that my pieces aren't tampered with? or How would the receiver be able to tell whether the bits produce an error? As a result, we must learn the pathways for detecting errors. As a result, data can be manipulated when being transmitted. This is the reason we need to learn error detection mechanism.

6.1 What is the block coding?

If I want my receiver to know whether the bits are corrupted or not, I use block coding, which divides our messages into a block called a data term. So, along with the data terms, we'll add some redundant bits. When we combine the redundant terms with the data universe, we get our code name. If I write 101, I'm using data terms. This is a list of my data terms. Now, if we add r redundant bits to this data word, it becomes the code word; if we add r redundant bits to this data word, it becomes the code word.

This is referred to as a code term. As a result, the sender sends these bits to the recipient. As a result, we'll use these redundant bits to determine if my data is properly corrected and the same data that the sender sent, or if my bits are incorrect. So we can build a combination of two raise to power k data words using k bit. With n bits, we can make a dataword that combines two raise to power n datawords.

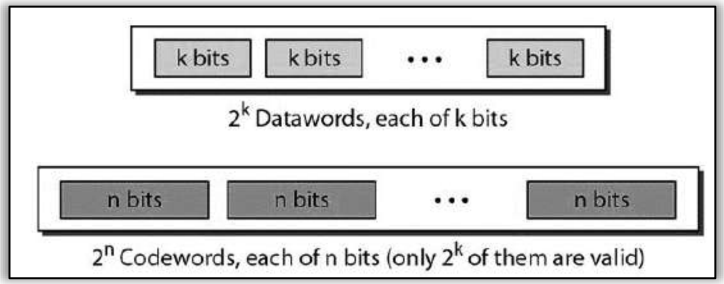

k bits
k bits
• • •
k bits
2^k Datawords, each of k bits
n bits
n bits
• • •
n bits
2^n Codewords, each of n bits (only 2^k of them are valid)

Figure 2 Block Coding

Ques:- How can errors be detected by using block coding?

Ans:- So, we have two options for determining if my data is right or not: my potential recipient already has a valid codeword, and that valid code will determine if my data is correct or not. If they don't align, the bits are corrupted, and the original code word has been replaced by an invalid one as seen in Figure 3.

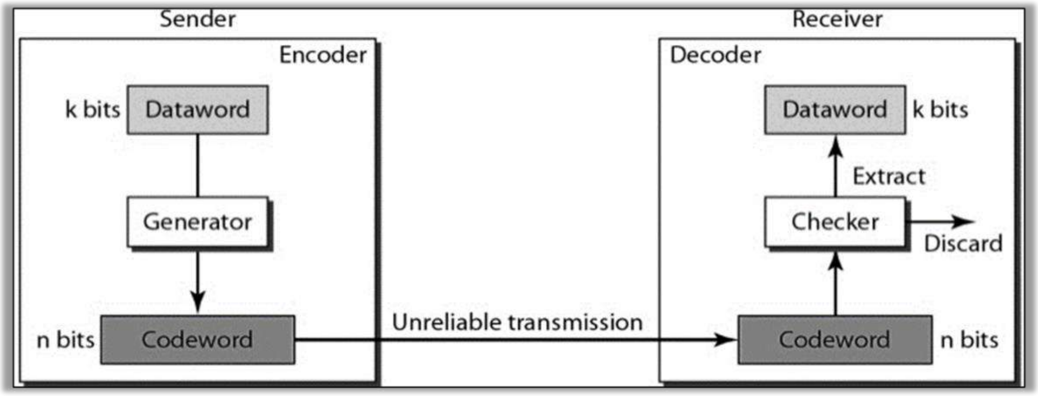


Figure 3 Error Detection

So, let's start with a few topics that are directly or indirectly relevant to error detection and correction mechanisms. First and foremost, we use only a small number of integers while we are using modular arithmetic. As a result, we define a modulus, which is a spectrum of zero to n minus one. If n is 12, the range would be zero to 12 minus one, which is zero to 11. As a result, we only use the integers 0 to 11 inclusive. Only integers in the range of zero to n minus one are used in modulo-N arithmetic. In modulus arithmetic, there is no carry and if you do a subtraction there is no carry even if you do an addition there is no carry, I hope you people have seen this, in case of addition and subtraction the outcome is almost the same, the result is almost the same zero plus 0 and zero both are the same here, in case of addition and subtraction the result is almost the same, the result is almost the same zero plus 0 . Whenever we use the error correction schemes, we'll use the exclusive OR estate. So you must understand what an exclusive is. If the bits are the same, the result will be zero; if the bits are different, the result will be one. Similarly, these are the results of the exclusive OR, and this is what we'll use to assist you in error detecting mechanisms, such as when calculating the cyclic redundancy search. Even so, in addition to performing, we use this XOR property.

Types of Error

There are bits that can be changed during transmissions whenever a bit flow from one place to another. So, single bit error, multi bit error, and burst error are the three categories of errors as seen in Figure 4.

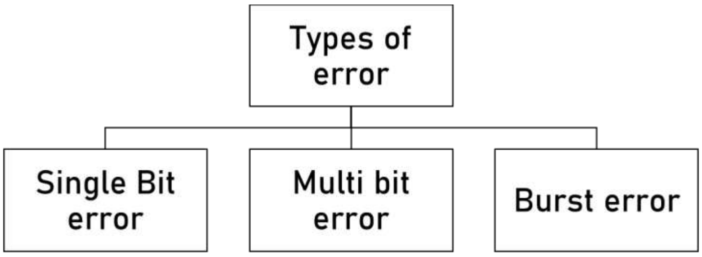


Figure 4 Types of Error

Single Bit Error

If only a single bit is corrupted, you can see that only a single bit from zero to one is corrupted. As a result, this is a well-known example of a single bit as seen in Figure 5.

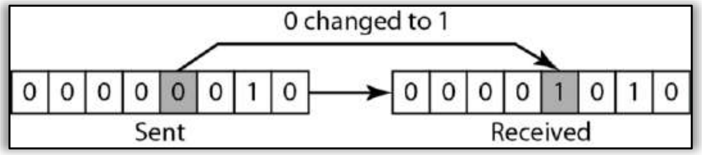


Figure 5 Single bit Error

Multi-bit Error

If two bits in two separate places are corrupted. As a result, this is referred to as a multi bit malfunction.

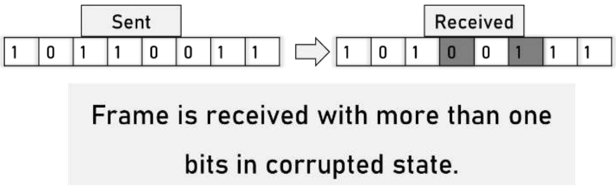

Sent
1 0 1 1 0 0 1 1
Received
1 0 1 0 0 1 1 1
Frame is received with more than one bits in corrupted state.

Figure 6 Multi-bit Error

Burst Error

A third type of error is a burst error, which occurs when several bits are scrambled at the same time. The fact that we have 100 instead of 011 indicates that three consecutive bits are compromised. As a result, this is referred to as a burst malfunction.

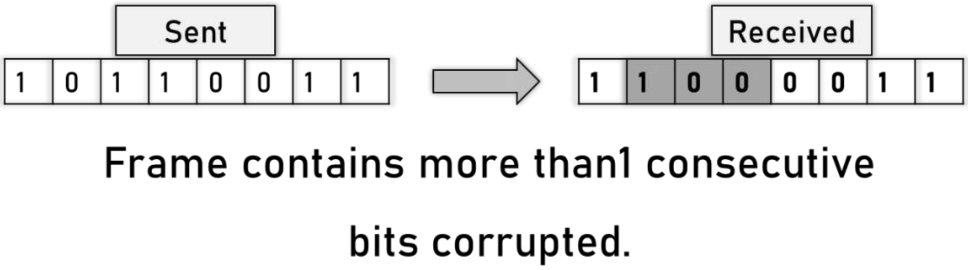


Figure 7 Burst Error

So, single bit error, multi bit error, and burst error are the three categories of errors. For several bits, just a single bit is compromised. When more than one bit is corrupted in multiple places, this is referred to as a multi bit error. When consecutive bits are corrupted, this is referred to as a burst error.

Redundancy

We add the redundant bits together with the data if we have a data expression. As a result, these redundant bits assist the receiver in determining if the bits are corrupted or false. So these are the pieces that assist the receiver in determining the outcome. We have a list of data words. This is referred to as data terms, and it isn't real data. As a result, it can be made up of 8 bits or any other number of bits. We then combine the redundant bits with the data. As a result, this amount becomes my code term. So we're sending duplicate bits, which means we're sending extra bits.

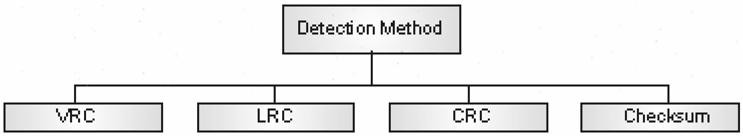


Figure 8 Four Types of Redundancy Check

Vertical Redundancy Check: The vertical redundancy check (VCR), also known as a parity check, is the most effective and least costly method for error detection. Any data unit is appended with a redundant bit called a parity bit in this procedure, bringing the total number of bits in the unit (including the parity bit) to an even number.

Longitudinal Redundancy Check: In longitudinal redundancy check (LRC), a block of bits is organized in a table (rows and columns). For example, instead of sending a block of 32 bits, we organize them in a table made of four rows and eight columns, as shown in Figure 9. We then calculate the parity bit for each column and create a new row of eight bits, which are the parity bits for the whole block.

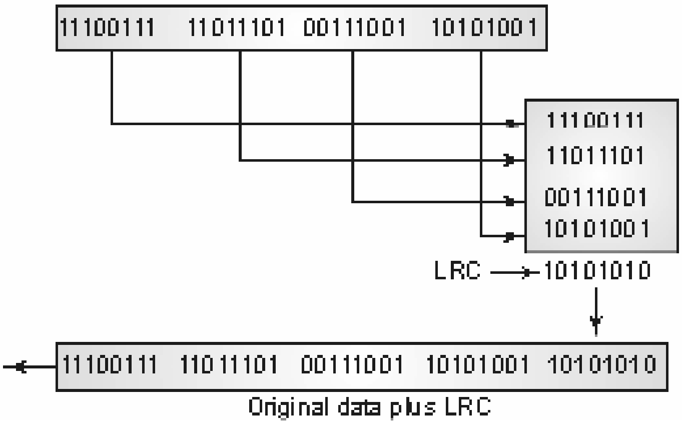


Figure 9 LRC

CRC and checksum will discuss later in this chapter.

6.2 Error Control Mechanisms May Work in Two Different Ways

We have two options for detecting and correcting a mistake. So we have two mechanisms: one is called error detecting, and the other is called error correction as seen in Figure 10. To begin, there is a distinction to be made between error detection and error correction. We only consider whether the error has occurred or not in the case of error detection; we do not consider which bit has an error, the magnitude of the error, whether it is a single bit error or a burst error in the case of error detection, but we do in the case of error correction.

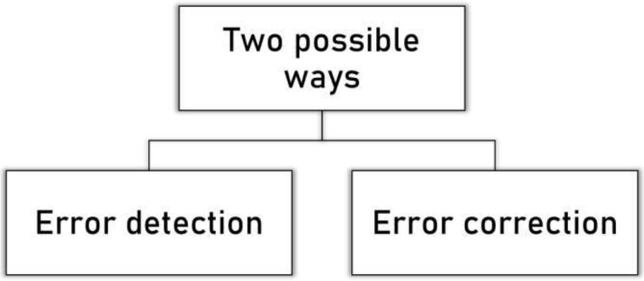


Figure 10 Error Control Mechanism

In the case of error correction, the receiver needs to see how many bits and how big the error is. As a result, in order to correct an error, we must know the precise number of bits that contain the error. We also need to know where the error is located, because the number of errors and their magnitude are critical factors in error correction.

Detection vs Correction

Table 1 Detection vs Correction

Error Detection	Error Correction
we are looking only to see if any error has occurred.	we need to know the exact number of bits that are corrupted and more importantly, their location in the message
Not even interested in the number of errors	The number of the errors and the size of the message are important factors
A single-bit error is the same for us as a burst error	If we need to correct one single error in an 8-bit data unit, we need to consider eight possible error locations;

Error Detection

Errors in the received frames are detected by means of

- Simple Parity Check
- Two-Dimensional parity Check
- Checksum
- and Cyclic Redundancy Check (CRC).

Simply Parity Check

Let's go over each one one by one. One is a quick parity check; it may be even or odd parity; in this case, we'll count the number of ones. If I'm counting, because I want equal parity, I'll calculate the amount of one and add to make it even. Let me demonstrate with an example how we can add the redundant bit as seen in Figure 11.

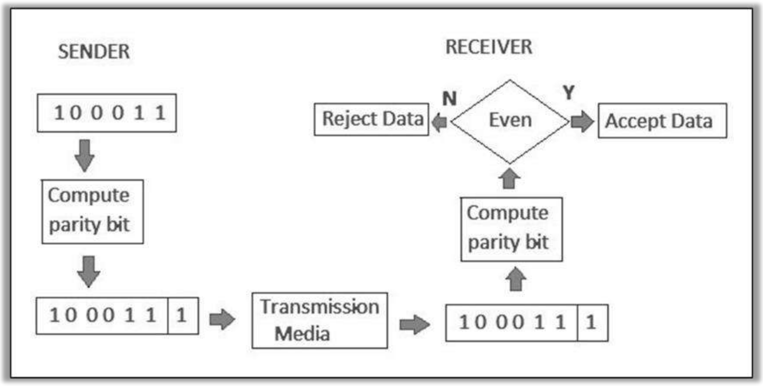


Figure 11 Simple Parity Check

Single Parity Bit

Single parity screening is an easy and inexpensive mechanism for detecting errors. A redundant bit, also known as a parity bit, is appended to the end of the data unit in this technique to make the number of 1s equal. As a result, the cumulative number of bits sent will be nine. If the number of 1s bits is odd, parity bit 1 is appended at the end of the data unit; if the number of 1s bits is even, parity bit 0 is appended. The parity bit is determined from the received data bits and added to the received parity bit at the receiving end. Even-parity checking is a procedure that produces a cumulative number of 1s that is even.

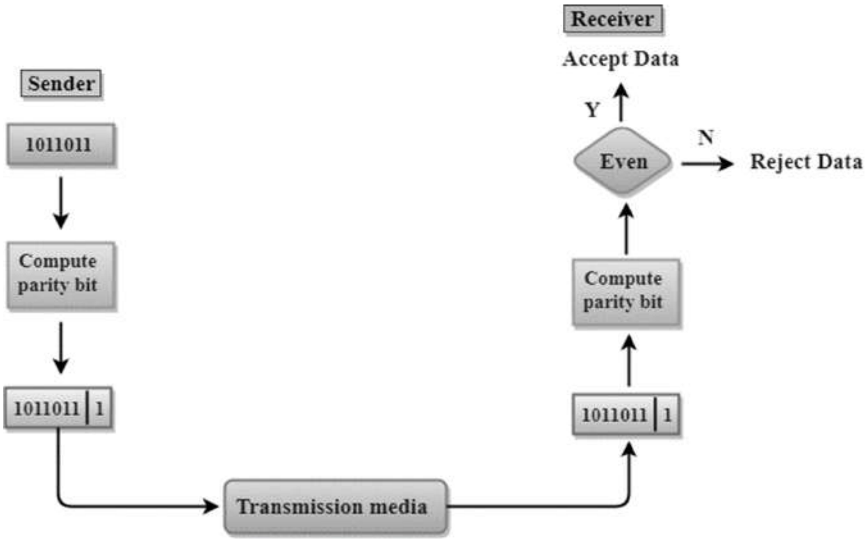


Figure 12 Single Parity Bit

Drawbacks of Single Parity Checking

It can only detect single-bit errors which are very rare.

If two bits are interchanged, then it cannot detect the errors.

Two-Dimensional Parity Check

Two-Dimensional Parity Check, which organizes data in the form of a table, will increase performance.

For each row, parity check bits are computed, which is analogous to a single-parity check.

A block of bits is separated into rows in a Two-Dimensional Parity scan, and the redundant row of bits is applied to the whole block.

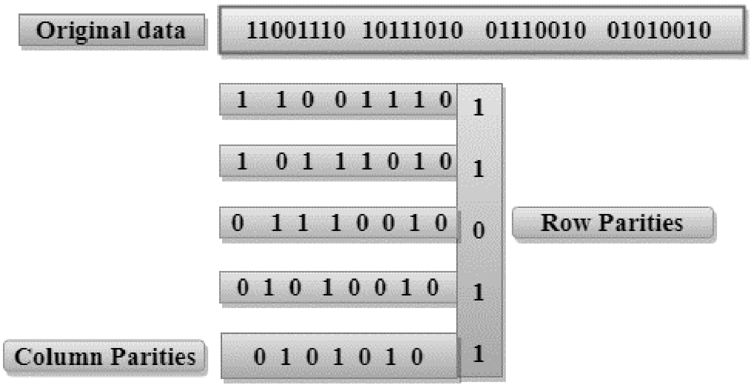


Figure 13 Two-Dimensional Parity Check

Drawbacks Of 2D Parity Check

The 2D Parity checker would not be able to spot the error if two bits in one data unit are corrupted and two bits in the same place in another data unit are not corrupted.

In certain instances, this technique may not be able to spot 4-bit or more errors.

Checksum

A checksum is a method for detecting errors that is based on the principle of redundancy.

It is divided into two parts:

Checksum Generator

A Checksum is generated at the sending side. Checksum generator subdivides the data into equal segments of n bits each, and all these segments are added together by using one's complement arithmetic. The sum is complemented and appended to the original data, known as checksum field. The extended data is transmitted across the network.

Suppose L is the total sum of the data segments, then the checksum would be ?L

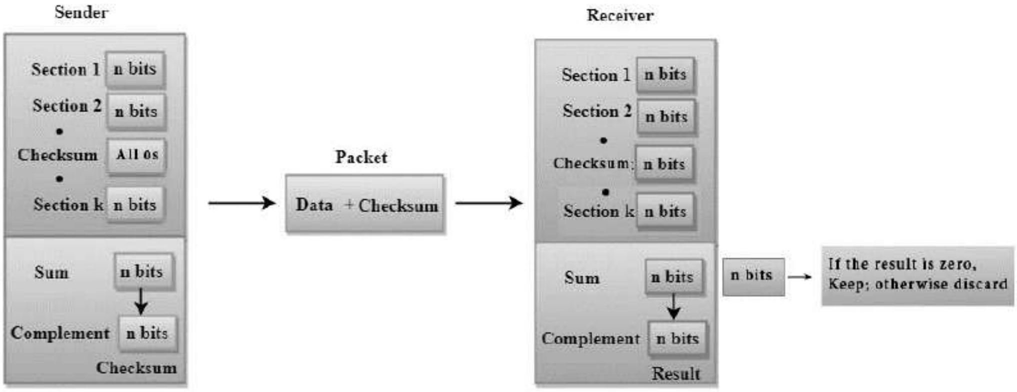


Figure 14 Checksum

The Sender follows the given steps as shown in Figure 14:

The block unit is divided into k sections, and each of n bits.

All the k sections are added together by using one's complement to get the sum.

The sum is complemented and it becomes the checksum field.

The original data and checksum field are sent across the network.

Checksum Checker

A Checksum is verified at the receiving side. The receiver subdivides the incoming data into equal segments of n bits each, and all these segments are added together, and then this sum is complemented. If the complement of the sum is zero, then the data is accepted otherwise data is rejected.

The Receiver follows the given steps:

The block unit is divided into k sections and each of n bits.

All the k sections are added together by using one's complement algorithm to get the sum.

The sum is complemented.

If the result of the sum is zero, then the data is accepted otherwise the data is discarded.

Cyclic Redundancy Check (CRC)

CRC is a redundancy error technique used to determine the error.

Following are the steps used in CRC for error detection:

In CRC technique, a string of n 0s is appended to the data unit, and this n number is less than the number of bits in a predetermined number, known as division which is n+1 bits.

Secondly, the newly extended data is divided by a divisor using a process is known as binary division. The remainder generated from this division is known as CRC remainder.

Thirdly, the CRC remainder replaces the appended 0s at the end of the original data. This newly generated unit is sent to the receiver.

The receiver receives the data followed by the CRC remainder. The receiver will treat this whole unit as a single unit, and it is divided by the same divisor that was used to find the CRC remainder.

If the resultant of this division is zero which means that it has no error, and the data is accepted.

If the resultant of this division is not zero which means that the data consists of an error. Therefore, the data is discarded.

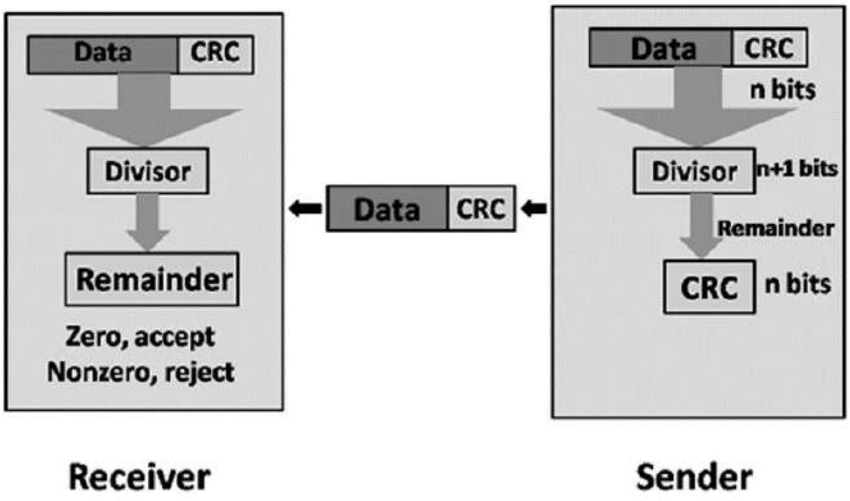


Figure 15 Cyclic Redundancy Check

Let's look at an example to better explain this concept:

Assume the divisor is 1001 and the initial data is 11100.

CRC Generator

- A CRC generator uses a modulo-2 division. Firstly, three zeroes are appended at the end of the data as the length of the divisor is 4 and we know that the length of the string 0s to be appended is always one less than the length of the divisor.
- Now, the string becomes 11100000, and the resultant string is divided by the divisor 1001.
- The remainder generated from the binary division is known as CRC remainder. The generated value of the CRC remainder is 111.

- CRC remainder replaces the appended string of 0s at the end of the data unit, and the final string would be 11100111 which is sent across the network as seen in Figure 16.

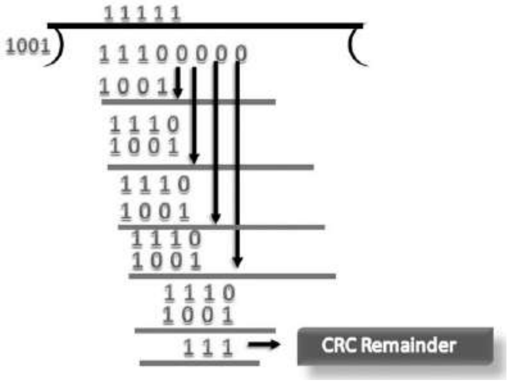


Figure 16 CRC Generator

CRC Checker

- The functionality of the CRC checker is similar to the CRC generator.
- When the string 11100111 is received at the receiving end, then CRC checker performs the modulo-2 division.
- A string is divided by the same divisor, i.e., 1001.
- In this case, CRC checker generates the remainder of zero. Therefore, the data is accepted as seen in Figure 17.

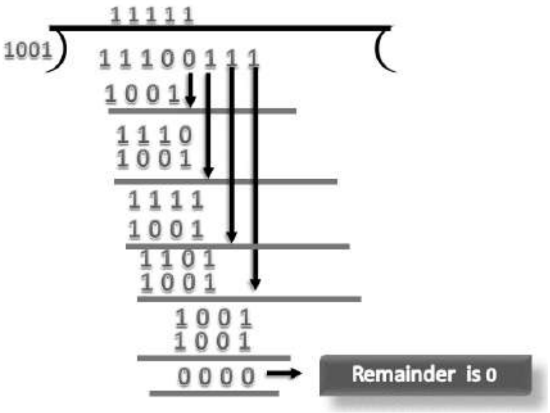


Figure 17 CRC Checker

6.3 Error Correction

Error correction is much more complex than error prediction, as we previously said. The receiver just needs to realize that the sent codeword is invalid in error detection; in error correction, the receiver must locate (or guess) the original codeword sent. We may assume that error correction requires more redundant bits than error detection.

Errors and Errors correcting code

In error correction, the encoder and decoder have different structures.

Bits can be corrupted when being transferred over a data network due to congestion and network issues. Errors are caused by distorted bits, which cause the recipient to receive erroneous data. Under the limits of the algorithm, error correcting codes determine the precise number of bits that have been corrupted and the origin of the corrupted bits.Error-correcting codes (ECC) are a series of numbers developed by algorithms for detecting and correcting errors in data transmitted over a noisy medium.

Error-correcting codes can be divided into two categories.

Block codes-

The message is broken down into fixed-size chunks of bits, with redundant bits inserted for error detection and correction.

Convolutional codes use data streams of arbitrary length to create parity symbols, which are created by sliding a Boolean function over the data stream.

Hamming Distance

So, the first is our hamming distance, which is the number of places where the related symbols vary from two strings of equal length. But, using hamming distance, we'll figure out how many bits are distinct.

Example

Find the Hamming Distance Between Two Pairs of Words

Ques:- The Hamming distance d(OOO, 011) is

Ans:- $\frac{000}{011}$ = 2 (two 1s)

$\frac{10101}{11110}$ = 3 (three 1s)

Minimum Hamming Distance

The shortest Hamming distance for all possible pairs of strings in a series of equal length strings is called the minimum Hamming distance.

Example

Suppose there are four strings 010, 011, 101 and 111.

010 ⊕ 011 = 001, d(010, 011) = 1.

010 ⊕ 101 = 111, d(010, 101) = 3.

010 ⊕ 111 = 101, d(010, 111) = 2.

011 ⊕ 101 = 110, d(011, 101) = 2.

011 ⊕ 111 = 100, d(011, 111) = 1.

101 ⊕ 111 = 010, d(011, 111) = 1.

Hence, the Minimum Hamming Distance, d_{min} = 1.

Hamming Code

Hamming code is a block code that can identify and correct single-bit errors while detecting up to two simultaneous bit errors. It was created by R.W. Hamming for the purpose of error correction. The source encodes the message using this coding process by adding redundant bits into the message. Extra bits are produced and placed at unique locations in the message to allow error detection and correction. When the destination receives this message, it performs recalculations in order to locate errors and determine which bit location is incorrect.

Hamming Code is a method of encoding a letter.

The procedure used by the sender to encode the message encompasses the following steps −

Step 1 − Calculation of the number of redundant bits.

Step 2 − Positioning the redundant bits.

Step 3 − Calculating the values of each redundant bit.

Once the redundant bits are embedded within the message, this is sent to the user.

Decoding a Hamming Code message

When a message is sent, the recipient performs recalculations to find and correct errors. The recalculation procedure is as follows:

Step 1: Determine how many redundant bits there are.

Step 2: Putting the obsolete pieces in their proper places.

Step 3: Testing for parity.

Step 4: Identifying and correcting errors

Step 1 − Calculation of the number of redundant bits

Using the same formula as in encoding, the number of redundant bits are ascertained.

$2^r \geq m + r + 1$ where *m* is the number of data bits and *r* is the number of redundant bits.

Step 2 − Positioning the redundant bits

The *r* redundant bits placed at bit positions of powers of 2, i.e. 1, 2, 4, 8, 16 etc.

Step 3 − Parity checking

Parity bits are calculated based upon the data bits and the redundant bits using the same rule as during generation of c_1,c_2 ,c_3 ,c_4 etc. Thus

c_1 = parity(1, 3, 5, 7, 9, 11 and so on)

c_2 = parity(2, 3, 6, 7, 10, 11 and so on)

c_3 = parity(4-7, 12-15, 20-23 and so on)

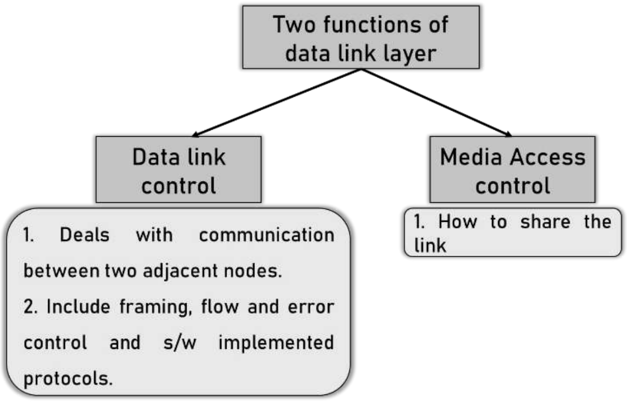
Step 4 − Error detection and correction

The decimal equivalent of the parity bits binary values is calculated. If it is 0, there is no error. Otherwise, the decimal value gives the bit position which has error. For example, if $c_1c_2c_3c_4$ = 1001, it implies that the data bit at position 9, decimal equivalent of 1001, has error. The bit is flipped to get the correct message.

6.4 Data Link Control

Data link management and media access control are the two primary features of the data link layer as shown in Figure 18. The first, data link control, is concerned with the architecture and procedures for node-to-node connectivity between two neighbouring nodes. This feature is covered in this chapter. The data link layer's second feature is media access control, or how to share the link.

Framing, flow and error management, and software-implemented protocols are all data link control functions that ensure seamless and stable frame transfer between nodes. The first topic in this chapter is framing, or how to arrange the pieces held by the physical layer. The flow and error management are then discussed.


Two functions of data link layer

Data link control

1. Deals with communication between two adjacent nodes.

2. Include framing, flow and error control and s/w implemented protocols.

Media Access control

1. How to share the link

Figure 18 Data Link control

FRAMING

Moving bits in the form of a signal from the source to the destination is what data transfer in the physical layer entails. Bit synchronisation is provided by the physical layer, which ensures that the sender and recipient use the same bit durations and timing.

The data link layer, on the other hand, must cram bits into frames in such a way that each frame can be distinguished from the previous one. Framing is a technique used by our mail department. The basic act of sealing a letter in an envelope divides one piece of data from another; the envelope acts as a delimiter. By inserting a sender address and a recipient address, framing in the data link layer divides a message from one source to a destination, or from other messages to other destinations. The sender address aids the receiver in acknowledging the arrival of the packet; the destination address determines where the packet will go.

While the whole message could fit in one frame, this is rarely achieved. One explanation for this is that frames can be very broad, making flow and error control inefficient. When a message is sent in a single big frame, then a single bit error requires the whole message to be resent. A single-bit error impacts only the small frame when a message is split into bigger frames.

Two Types of Framing

A. Fixed sized framing

B. Variable sized framing

Fixed-Size Framing

The size of the frames may be constant or variable. There is no need to define the limits of the frames in fixed-size framing; the size will serve as a delimiter. The ATM wide-area network, which uses fixed-size frames called cells, is an example of this method of framing.

Variable-Size Framing

In variable-size framing, which is common in local area networks. We need a way to identify the end of one frame and the start of the next in variable-size framing. A character-oriented approach and a bit-oriented approach have also been used in the past for this reason.

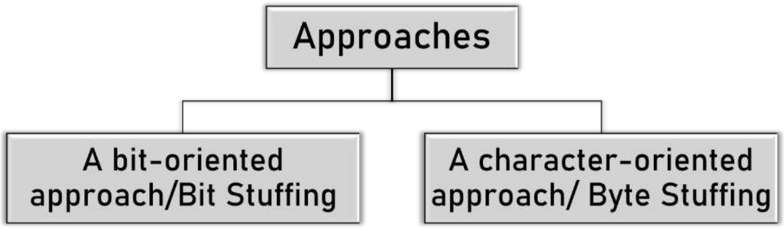

Approaches
A bit-oriented approach/Bit Stuffing
A character-oriented approach/ Byte Stuffing

Figure 19 Variable Size Framing

Character Oriented Protocols

The data to be transported in a character-oriented protocol is 8-bit characters from an encoding scheme such as ASCII (see Appendix A). The header, which typically contains the source and destination addresses as well as other control information, and the trailer, which contains redundant bits for error detection and correction, are both multiples of 8 bits. An 8-bit (I-byte) flag is applied to the beginning and end of each frame to distinguish it from the next. The start or end of a frame is signaled by the flag, which is made up of protocol-dependent special characters. The format of a frame in a character-oriented protocol is seen in Figure 20.

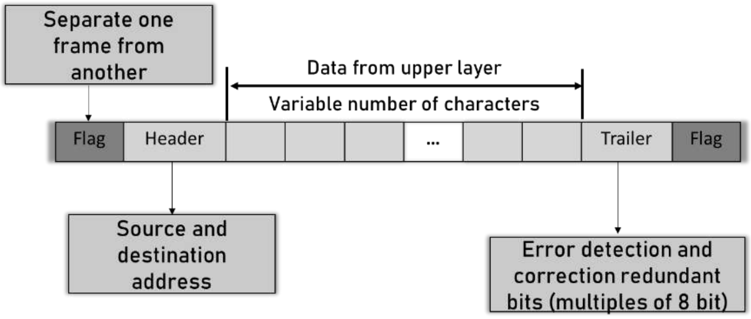

Separate one frame from another
Data from upper layer
Variable number of characters
Flag
Header
...
Trailer
Flag
Source and destination address
Error detection and correction redundant bits (multiples of 8 bit)

Figure 20 A frame in a character-oriented protocols

When the data link layers only exchanged text, character-oriented framing was common. Every character that isn't used in text chat may be chosen as the flag. Other types of content, such as graphs, audio, and video, are now sent. Any pattern used on the flag may also be included in the data. If this occurs, the receiver may believe it has reached the end of the frame when it finds this pattern in the centre of the data. A byte-stuffing technique was applied to character-oriented framing to solve this issue. When a character has the same pattern as the flag, byte stuffing (or character stuffing) adds a separate byte to the data portion of the frame. An extra byte has been stuffed into the data row. This byte, which has a predefined bit pattern, is known as the escape character (ESC). The receiver extracts the ESC character from the data segment and considers the next character as data rather than a delimiting flag whenever it sees it.

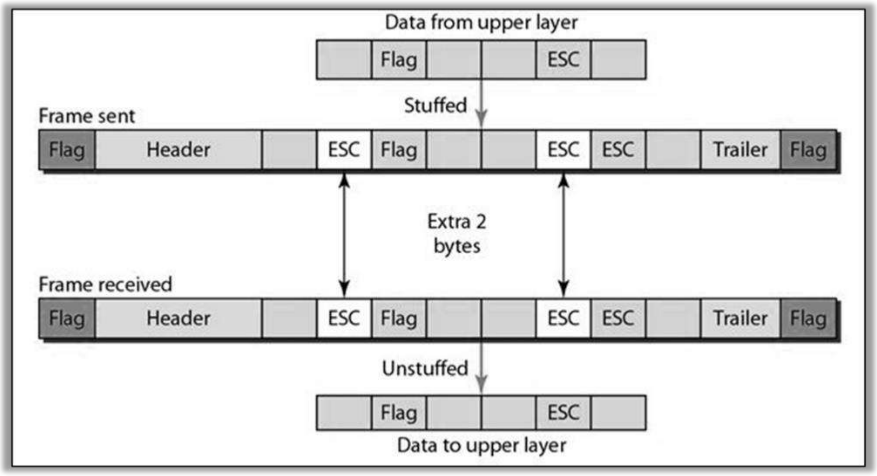


Figure 21 Byte stuffing and unstuffing

The inclusion of the flag in the data section of the frame is allowed by byte stuffing by the escape character, but it introduces a new challenge. What happens if one or more escape characters are accompanied by a flag in the text? The receiver discards the escape character but holds the flag, which is mistakenly read as the frame's end. The escape characters that are part of the text must also be labelled by another escape character to solve this dilemma. To put it another way, if the escape character is already present in the code, an additional one is inserted to indicate that the second one is still present. The scenario is depicted in Figure 21. Another issue with data communications is character-oriented protocols. Unicode and other universal coding schemes use 16-bit and 32-bit characters, which clash with 8-bit characters.

Bit-Oriented Protocols

The data segment of a frame in a bit-oriented protocol is a series of bits that the upper layer interprets as text, image, audio, video, and so on. We do need a delimiter to distinguish one frame from the next, in addition to headers (and perhaps trailers). To specify the beginning and end of the frame, most protocols use the special 8-bit pattern flag 01111110 as the delimiter as shown in Figure 22.

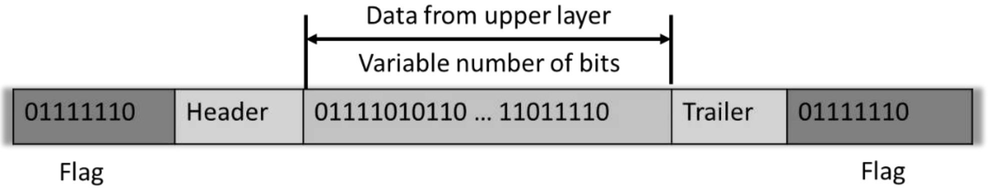


Figure 22 Bit-Oriented Protocol

This flag has the same potential for causing problems as the byte-oriented protocols. That is, if the flag pattern exists in the details, we must notify the receiver that the frame is not yet complete. To avoid the pattern appearing like a flag, we stuff one single bit (rather than one byte).

The receiver finally removes the extra stuffed bit from the results. It's worth noting that the extra bit is inserted after one 0 and five 1s, regardless of the next bit's value. This ensures that the flag field sequence does not appear in the frame by accident.

Bit stuffing at the sender and bit elimination at the receiver are shown in Figure 23. It's worth noting that even though we end up with a 0 after five 1s, we still stuff an O. The receiver will delete the zero.

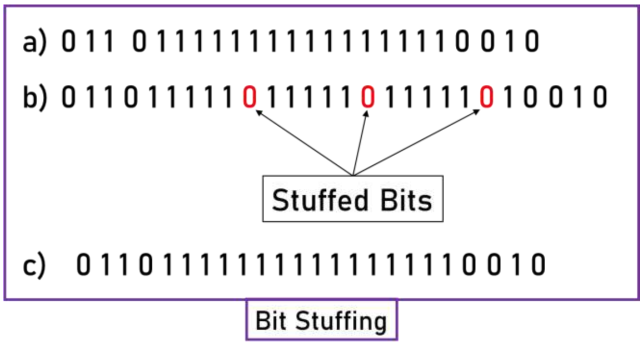

a) 0 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 1 0
b) 0 1 1 0 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 0 1 0 0 1 0
Stuffed Bits
c) 0 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 1 0
Bit Stuffing

Figure 23 Bit Stuffing and Unstuffing

6.5 Flow Control and Error Control

Data transmission necessitates the cooperation of at least two computers, one of which sends data and the other of which receives it. Even such a simple arrangement necessitates a great deal of planning in order for an understandable transaction to take place. Flow control and error control are the two most critical functions of the data link network. These functions are collectively referred to as data link management. The basic functions of data link layer are framing, Error control and flow control as seen in Figure 24.

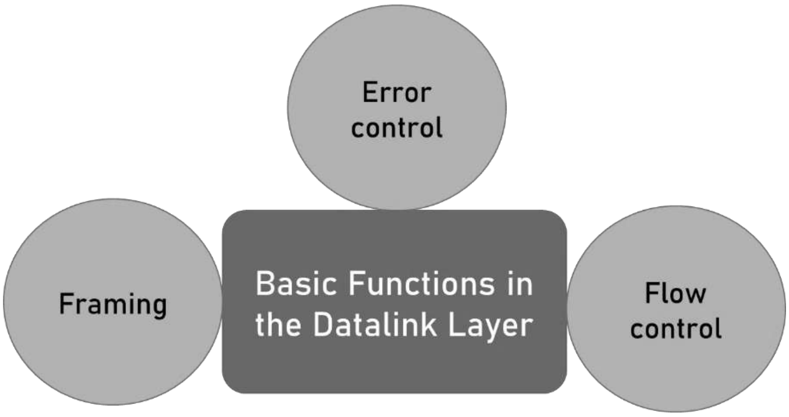


Figure 24 Basic Functions of Data Link Layer

6.6 Protocols

Let's take a look at how the data link layer can use framing, flow management, and error control to deliver data from one node to another. In most cases, the protocols are implemented in software using one of the standard programming languages. To keep our discussions language-free, we wrote a version of each protocol in pseudocode that focuses mostly on the process rather than delving into the nuances of language rules.

Types of Data Link Protocols

The protocols are divided into those that can be used for noiseless (error-free) channels and those that can be used for noisy (error-producing) channels. The protocols in the first group can't be used in real life, so they're useful for learning about noisy channel protocols. The classifications as seen in Figure 25.

There is a distinction between the protocols we address here and the protocols used in actual networks. The data frames pass from one node, referred to as the sender, to another node, referred to as the receiver, in all of the protocols we address.

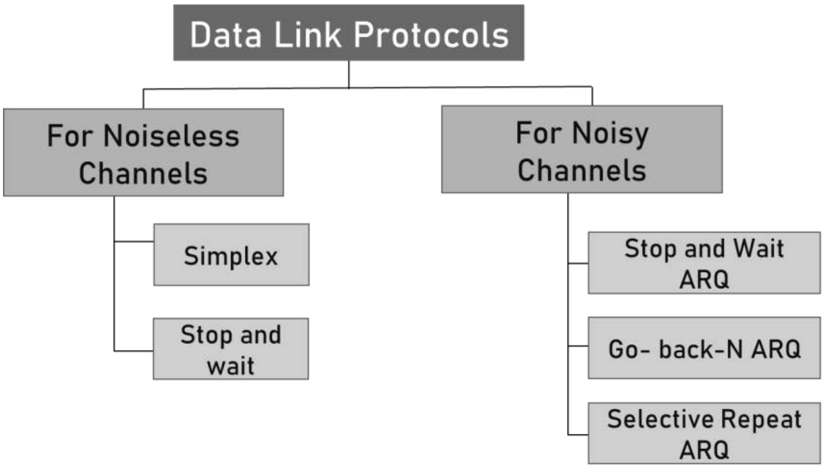


Figure 25 Classifications of protocols

Data flows in just one direction, despite the fact that special frames called acknowledgment (ACK) and negative acknowledgment (NAK) will flow in the opposite direction for flow and error monitoring. In a real-world network, data link protocols are bidirectional, allowing data to flow in all directions. Piggybacking is a method used in these protocols to provide flow and error control information such as ACKs and NAKs in the data frames. We picked the latter for our topic because bidirectional protocols are more complicated than unidirectional protocols. They can be generalised to bidirectional protocols if they are understood.

6.7 NOISELESS CHANNELS

Let's pretend we have an ideal channel with no missing, duplicated, or compromised frames. For this sort of channel, we add two protocols. The first is a protocol that does not use flow control, while the second does. Of necessity, neither has error management since we assumed the channel to be fully noiseless.

Simplest Protocol

For lack of a better term, we call our first protocol the Simplest Protocol. It has no flow or error management. It is a one-way protocol, meaning data frames only pass in one direction, from sender to receiver. We assume that the receiver can accommodate any frame it receives right away, with a loading time that is negligible. The receiver's data link layer separates the header from the frame right away and hands the data packet to its network layer, which will accept it right away. To put it another way, the receiver should never be overloaded with pictures.

Design

Flow control is not needed in this scheme. The sender site's data link layer receives data from its network layer, creates a frame out of it, and transfers it. The receiver site's data link layer collects a frame from its physical layer, removes data from it, and sends it to its network layer. The sender and receiver's data link layers provide transmission facilities to their network layers. For the actual exchange of bits, the data link layers depend on the services offered by their physical layers (such as signalling, multiplexing, and so on). A design as seen in Figure 26.

We need to go into the protocol that all data link layers use. The sender location would not be able to transmit a frame until it receives a data packet from its network layer. Until a frame arrives, the recipient location cannot send a data packet to the network layer. If the protocol is applied as a process, the concept of events must be introduced. The procedure at the sender site is still running; no action is taken before the network layer makes a request. The process at the receiver site is still running, but no action is taken before the physical layer sends a warning. Since they don't know when the associated incidents will occur, both operations are still going.

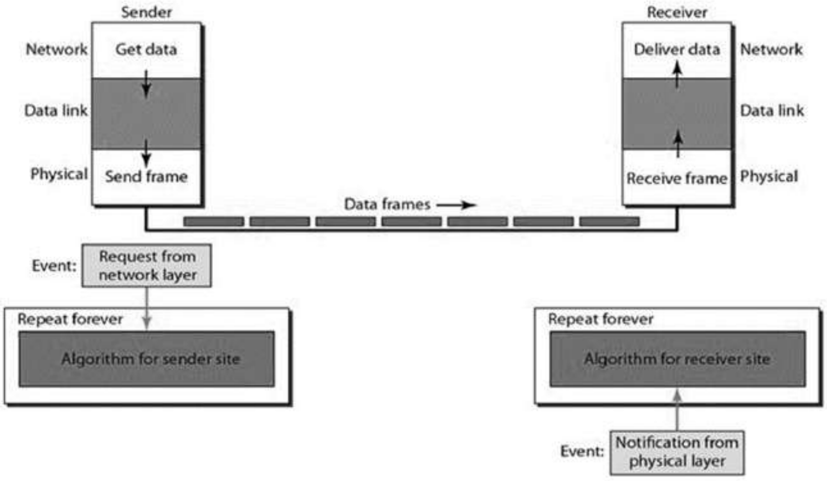
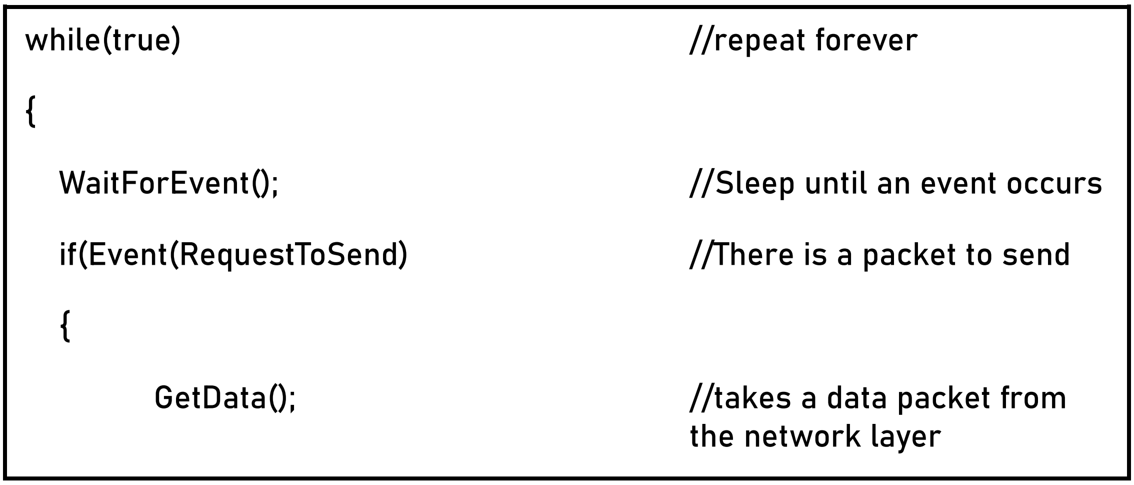


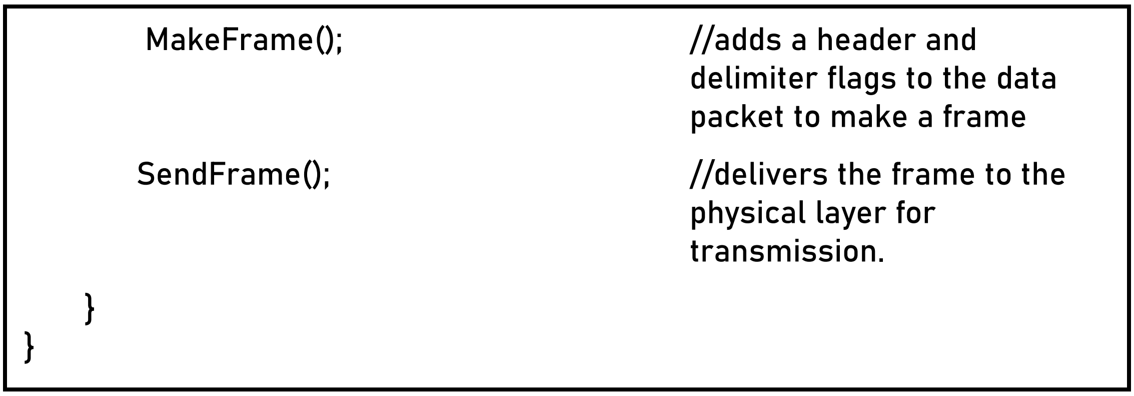
Figure 26 The design of the simplest protocol with no flow or error control

Algorithms

The protocol at the sender site is depicted in Algorithm 1.

Algorithm 1 The simplest protocol has a sender-site algorithm.


```
while(true)                              //repeat forever
{
   WaitForEvent();                       //Sleep until an event occurs
   if(Event(RequestToSend)               //There is a packet to send
   {
         GetData();                      //takes a data packet from
                                         the network layer
```

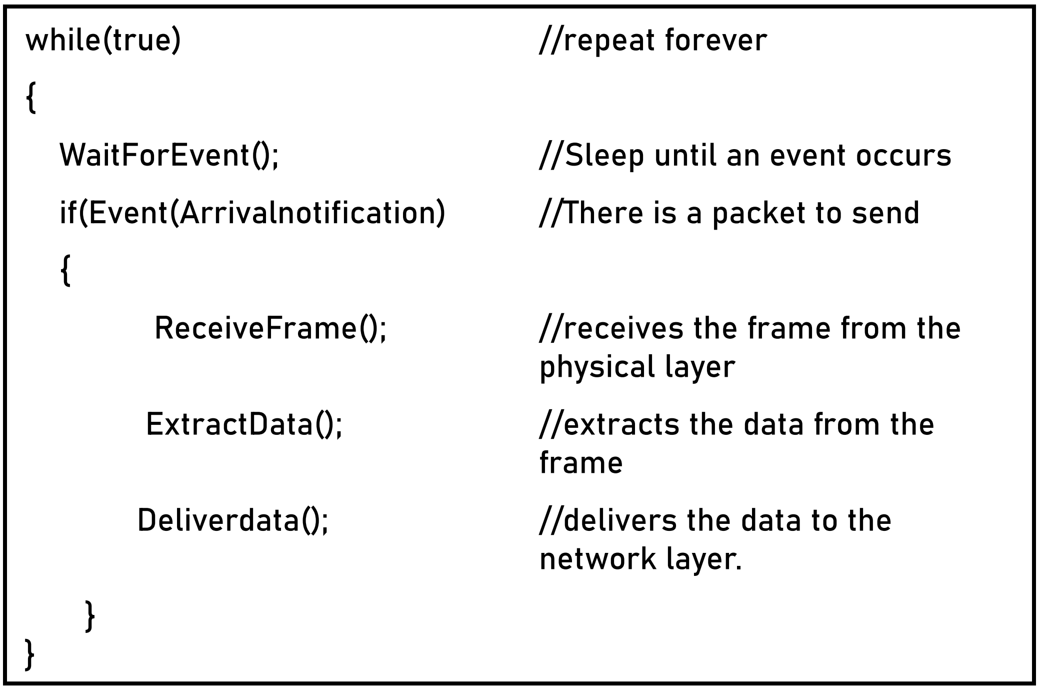


```
        MakeFrame();                    //adds a header and
                                        delimiter flags to the data
                                        packet to make a frame

        SendFrame();                    //delivers the frame to the
                                        physical layer for
                                        transmission.

    }
}
```

Algorithm 2 The simplest protocol has a receiver-site algorithm.



```
while(true)                             //repeat forever
{
  WaitForEvent();                       //Sleep until an event occurs
  if(Event(Arrivalnotification)         //There is a packet to send
  {
         ReceiveFrame();                //receives the frame from the
                                        physical layer
        ExtractData();                  //extracts the data from the
                                        frame
       Deliverdata();                   //delivers the data to the
                                        network layer.
    }
}
```

Stop-and-Wait Protocol

When data frames arrive at the receiver site quicker than they can be processed, they must be stored before they are needed. The recipient, in most cases, does not have enough storage capacity, particularly if it is processing data from multiple sources. Frames may be discarded or service may be denied as a result of this. To keep the recipient from becoming overburdened for pictures, we need to advise the sender to slow down in some way. There has to be more. The sender receives input from the recipient. The Stop-and-Wait Protocol gets its name from the fact that the sender sends one frame, waits for clarification from the recipient (all right, go ahead), and then sends the next frame. Data frames remain unidirectional, but auxiliary ACK frames (simple tokens of acknowledgment) are sent in the opposite direction.

Example

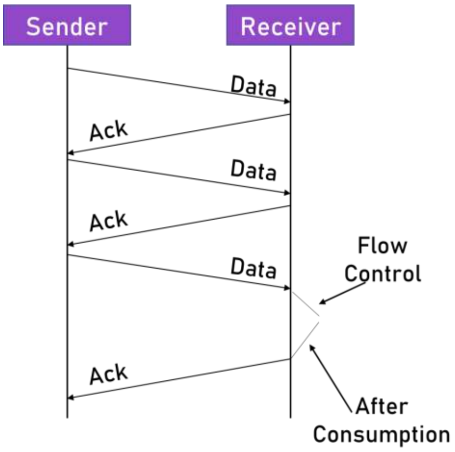


Figure 27 Stop and Wait Protocol

Figure 27 plays a collaboration sample using this protocol. It's all quite straightforward. The sender sends one frame and waits for the receiver's answer. When the ACK comes up, the next frame is sent by the sender. It's worth noting that sending two frames in the protocol puts the sender in danger. The sender is involved in four cases, while the recipient is involved in two.

6.8 NOISY CHANNELS

Since the Stop-and-Wait Protocol demonstrates how to incorporate flow control into it, Noiseless channels, unlike their predecessors, do not exist. We may choose to dismiss the mistake (as we always do). We either need to apply error management to our protocols (which we don't), or we need to add error control to our protocols.

Stop-and-Wait Automatic Repeat Request

The Stop-and-Wait Automatic Repeat Request (Stop-and-Wait ARQ) is our first protocol, which applies a basic error control feature to the Stop-and-Wait Protocol. Let's take a look at how this protocol does error detection and correction. We need to apply redundancy bits to our data frame to find and repair corrupted frames. Frames that have been lost are more difficult to deal with than those that have been corrupted. In previous protocols, we used There was no way to tell what frame it was. It's possible that the obtained frame is the right one. or a copy, or an out-of-order frame To solve the problem, number the frames. When it comes to the When a recipient gets an out-of-order data frame, it means the frames were either sent in the wrong order or were sent in the wrong order. duplicated or misplaced.

In this protocol, the completed and lost frames must be resent. If the recipient agrees, If the sender does not react when an error occurs, how can the sender know which frames to resend? To To avoid this, the sender saves a copy of the sent frame. It begins at the same time. Since an ACK frame can be corrupted and lost, it needs redundancy bits as well. as well as a series number This protocol's ACK frame includes a sequence number region. In this protocol, the sender actually discards or rejects a compromised ACK frame. One that is out of order.

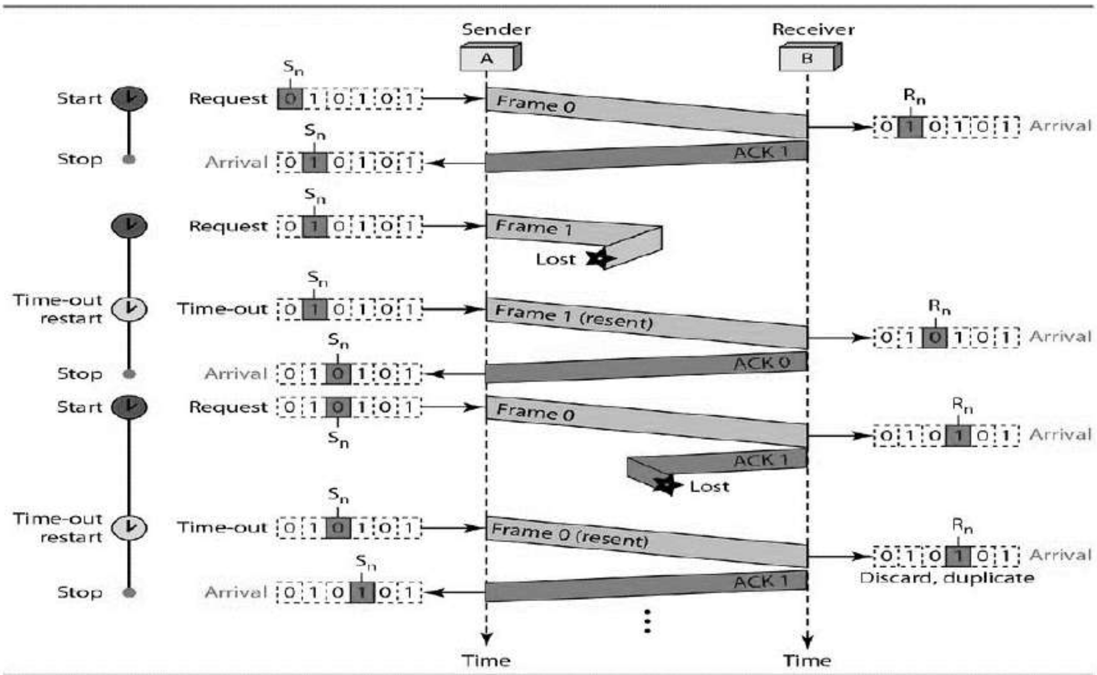


Figure 28 Stop and Wait ARQ

In Figure 28, frame an is submitted and accepted as an indication of Stop-and-Wait ARQ. After the time-out, frame 1 is missed and resent. The timer is stopped after the received frame 1 is acknowledged. The frame has been submitted and accepted, but the acknowledgement has been misplaced. Since the sender has no way of knowing if the frame or the acknowledgment has been lost, it resends frame 0, which is accepted, after the time-out.

Go-Back-N Automatic Repeat Request

Many frames must be in transition when waiting for acknowledgement to increase transmission performance (filling the pipe). To put it another way, we need to keep the channel busy while the sender waits for acknowledgement by allowing several frames to be outstanding. Go-Back-N Automatic Repeat Request is the only . We will send several frames in this protocol before collecting acknowledgements, and we hold a backup of these frames before the acknowledgements arrive.

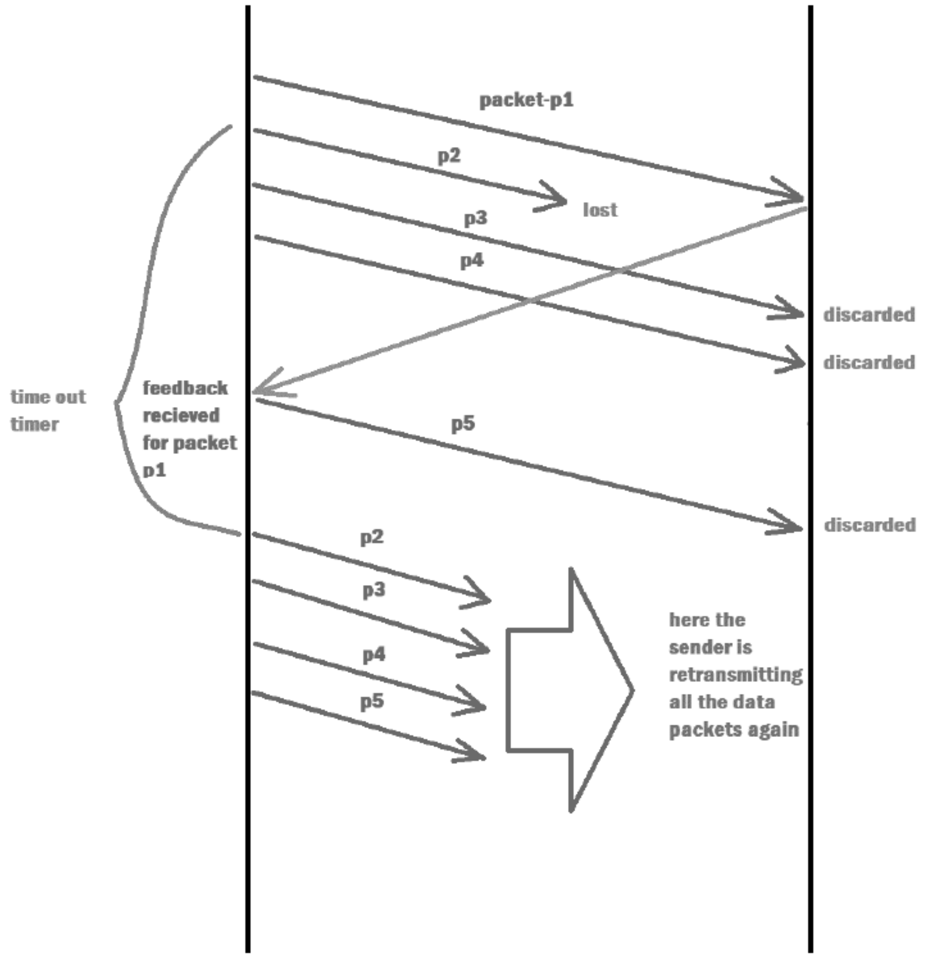

packet-p1
p2
lost
p3
p4
discarded
discarded
time out timer
feedback recieved for packet p1
p5
discarded
p2
p3
p4
p5
here the sender is retransmitting all the data packets again

Figure 29 Go-Back-N ARQ

Selective-Repeat ARQ

The go-back-n protocol works well when errors are low, but it loses a lot of bandwidth on retransmitted frames when the line is poor. The selective repeat protocol is an alternative technique that allows the receiver to accept and buffer frames after a broken or missing one.

discerning Retransmit only those packets that are currently missing (due to errors) a second time:

The receiver must be able to accept out-of-order packets.

Since packets must be released in order to the higher layer, the recipient must be able to buffer any packets.

Retransmission requests:

Implicit – The receiver accepts any positive packet; packets that do not receive an ACK before the timeout are considered to be missing or incorrect. It's worth noting that this method must be used to ensure that any packet is obtained at some stage.

Explicit NAK (selective reject) – An explicit NAK (selective reject) can only request the retransmission of one packet. This method might speed up the retransmission, but it isn't strictly necessary. In reality, one or both methods are used.

Selective Repeat Protocol

Except that buffers are used and the recipient and sender each retain a window of space, this protocol (SRP) is almost identical to the GBN protocol. Where the connection is extremely unreliable, SRP works well.

Sender's Windows (Ws) = Receiver's Windows (Wr).

- In the SR protocol, the window size should be smaller than or equal to half the sequence number. This is to prevent packets from being misidentified. If the windows size is greater than half the sequence number space, the sender can send new packets that the receiver interprets as retransmissions if an ACK is lost.
- The sender will send new packets as long as the total number of unacknowledged packets is less than W.
- After a timeout – or a NAK if NAK is used – the sender retransmits un-ACKed packets.
- The right packets are acknowledged by the receiver.
- The receiver saves the right packets before they can be sent to the higher layer in time.
- The sender and receiver windows in Selective Repeat ARQ must be no more than one-half of a metre wide.

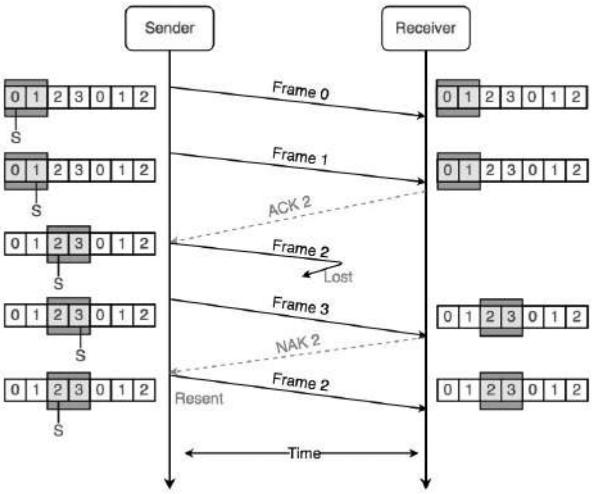


Figure 30 Selective Repeat Protocol

Figure 30: Only frames for which a NAK is sent are retransmitted by the sender.

The Selective Repeat Protocol (SRP) has the same utility as the GO-Back-N protocol.

Efficiency = N/(1+2a)

Where a = Propagation delay / Transmission delay

Buffers = N + N

Sequence number = N (sender side) + N (Receiver Side)

Piggybacking

Data frames flow in only one direction in the three protocols we discussed in this section, though control information such as ACK and NAK frames will pass in both directions. Data frames flow in both ways in real life: from node A to node B and from node B to node A. This necessitates the flow of control knowledge in both directions. Piggybacking is a method for increasing the reliability of bidirectional protocols. When a frame carries data from point A to point B, it can also carry control information about arrived (or lost) frames from point B; when a frame carries data from point B to point A, it can also carry control information about arrived (or lost) frames from point A.

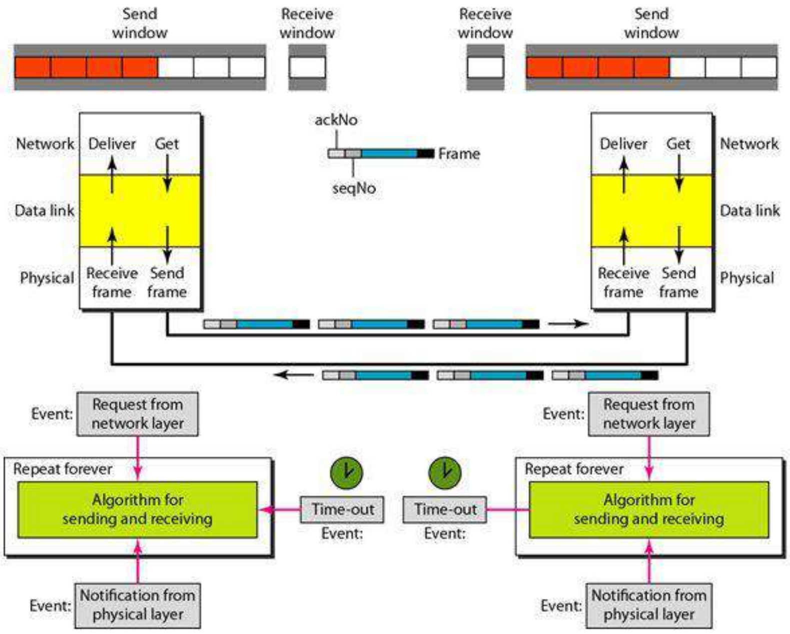


Figure 31 Design of piggybacking in Go-Back-NARQ

Figure 31 shows a specification for a Go-Back-N ARQ that uses piggybacking. Each node now has two windows: one for sending and one for receiving data. Both include the use of a timer. Request, entry, and time-out are three forms of incidents in which all are concerned. The arrival case, on the other hand, is complicated; when a frame arrives, the site must manage all control details and the frame itself.

All of these issues must be addressed in a single situation: the arrival event. At each location, the request event only uses the send window; the arrival event requires both windows. The fact that all pages would use the same algorithm is a vital aspect of piggybacking. Since it would merge two arrival events into one, this algorithm is difficult.

Self-Assessment

1. Which can be used as an intermediate device in between transmitter entity and receiver entity?

a) IP router

b) Microwave router

c) Telephone switch

d) All of the mentioned

2. Which is more efficient?

a) Parity check

b) Cyclic redundancy check

c) Parity & Cyclic redundancy check

d) None of the mentioned

3. CRC uses

a) Multiplication

b) Binary division

c) Multiplication & Binary division

d) None of the mentioned

4. In ___________ coding, we divide our message into blocks, each of k bits,called _____________.

a) block; blockwords

b) linear; datawords

c) block; datawords

d) none of the above

5. Checksums use _________ arithmetic.

a) one's complement arithmetic
b) two's complement arithmetic
c) either (a) or (b)
d) none of the above

6. The checksum of 1111 and 1111 is _________.

a) 0000
b) 1111
c) 1110
d) 0111

7. In cyclic redundancy checking, the divisor is _______ the CRC.

a) one bit less than
b) one bit more than
c) The same size as
d) none of the above

8. In cyclic redundancy checking, what is the CRC?

a) The quotient
b) The dividend
c) The divisor
d) The remainder

9. Stop and Wait Automatic Repeat request is a special case of

a) simplest protocol
b) Go Back-N Automatic Repeat request
c) Selective Repeat Automatic Repeat Request
d) stop and wait

10. The send window in the Go-Back-N Protocol is an abstract conceot defining an imaginary box with

a) one variable
b) two variables
c) three variables
d) four variables

11. The stop and wait uses the link of

a) modulation
b) full duplex
c) half duplex
d) de modulation

12. In block coding, we divide our message into blocks, is called

a) code blocks

b) packet blocks

c) code words

d) data words

Answers: Self-Assessment

1.	d	2.	b	3.	c	4.	c
5.	a	6.	a	7.	b	8.	d
9.	b	10.	c	11.	c	12.	d

Summary

- The strategies for gaining access to a shared communication channel and ensuring secure data transfer are described in the data link layer. Framing, checksums, error detection and correction, acknowledgment, flow control, and encapsulating packets from the network layer to frames are some of its key responsibilities.
- Unacknowledged connectionless service, acknowledged connectionless service, and acknowledged connection-oriented service are all provided by the data link layer.
- The simplest type of error detection is parity check, which requires the receiver to count only the number of 1s in the obtained data stream with an additional parity bit.
- A checksum is a basic redundancy scan that is used to find data errors.
- Cyclic Redundancy Check is a method for adding a data string to packets of information that can be used to detect errors in the data packets that is commonly used in computer networks.
- On an error-free communication channel, the Stop and Wait protocol is the simplest to adopt and proves to be the most effective. An error-free contact channel, on the other hand, is virtually impossible.
- The Go Back N protocol necessitates buffer management, making it difficult to maintain source and destination devices in line. It's actually the least effective since it retransmits all subsequent frames in the event of a frame failure, wasting a lot of bandwidth.
- Selective Repeat is an enhancement of the Go Back N protocol that attempts to make more usage of bandwidth by reducing the amount of retransmissions by retransmitting just one frame rather than the whole sequence. As a result, Selective Repeat is a safer option.

Keywords

The Point-to-Point Protocol (PPP) is a data link layer protocol that links two communicating link-level peers at either end of a point-to-point link.

Selective Repeat: Provides buffers at the source and destination hosts, allowing the source node to provide multiple remaining frames at the same time and the destination node to accept out-of-order frames and store them in its window.

Simplex is a term that refers to Stop and Wait: After propagation, the source node waits for the destination node to accept it. Following receipt of the acknowledgment, the loop is repeated again

Checksum: An algorithm for calculating the binary values in a packet or other block of data and storing the results with the data to correlate with a new checksum at the other end.

Cyclic Redundancy Check: A method for adding a data string to information packets that can be used to find errors in the data packets.

Error management entails sequencing frames and submitting confirmation control frames.

Controlling the rate of data transfer between two source and destination hosts is referred to as flow control.

Framing: The data link layer splits the bit stream into frames to provide a secure transmission of bit streams to the network layer.

Go Back N: Using buffers, the Go Back N protocol allows the source computer to have multiple outstanding frames at the same time.

Even parity and odd parity methods are used in parity checks. The receiver's procedure is straightforward since it just has to count the number of 1s in the obtained data stream with the inclusion of a parity bit.

Acknowledged Connectivity-oriented Operation: Before any data transfer, the data link layer offers this service to the network layer by creating a connection between the source and destination hosts.

Acknowledged Connectionless Service: When each frame transmitted between two hosts is sent correctly, it is referred to as acknowledged connectionless service.

Review Questions

1. What is the data link protocol?

2. What advantages does Selective Repeat sliding window protocol offer over Go Back N

protocol?

3. What is the purpose of flow control?

4. Describe how does finite state machine model carry out protocol verification?

5. What are different data link protocols available? Why does PPP have become popular?

6. Explain error detection techniques.

7. Explain hamming code with example.

Further/Suggested Readings

- Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall.
- Behrouz A. Forouzan and Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies.
- Burton, Bill, *Remote Access for Cisco Networks*, McGraw-Hill Osborne Media, McGraw-Hill Osborne Media.
- Rajneesh Agrawal and Bharat Bhushan Tiwari, *Computer Networks and Internet*, Vikas Publication
- https://www.geeksforgeeks.org/computer-network-tutorials/

Dr. Rajni Bhalla, Lovely Professional University

Unit 07: Data Link Layer-medium access protocol

CONTENTS

Objectives

- understand bit-oriented protocol.
- know the frame format and types of HDLC and PPP frame.
- understand multiple access protocol.

Introduction

The frame is clearly viewed as a set of bits with no semantics or context in a bit based approach. Regardless of the frame contents, a bit oriented protocol can switch data frames. HDLC is a bit-oriented protocol.

HDLC

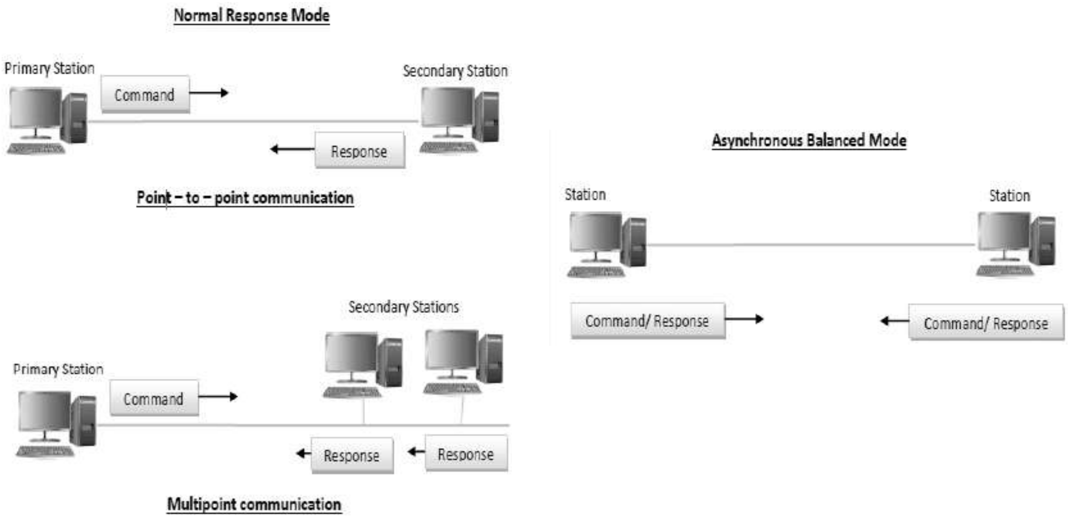
HDLC stands for high-level Data Link Control protocol, which ensures that data frames are delivered reliably over a network or communication link. HDLC also offers framing, data transparency, error detection and correction, and even flow management, among other things.

7.1 History of HDLC

Let's look at the history of HDLC. The synchronous Data Link Control protocol, developed by IBM, is an example of a bit based protocol. The SDLC is currently standardised as HDLC synchronous Data Link Control, who standardises as an HDLC. The SDLC was standardised by the ISO as the high level Data Link Control protocol, and SDLC is really important for the data link layer. It is not commonly used, but it is critical for many data link layer protocols. It's a bit-oriented protocol that can be used for point-to-point and multi-point communication.

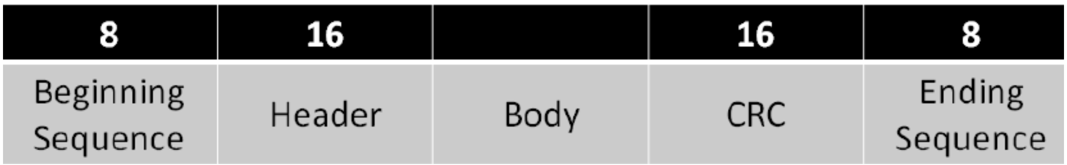
Transfer mode of HDLC

As you can see in the Figure 1, HDLC has two transfer modes: normal response mode and asynchronous balanced mode. There are two kinds of stations in normal response mode. A primary station sends orders, and a secondary station responds to the commands sent. In point-to-point and multi-point correspondence, the normal response mode is used. Asynchronous Balanced mode is when the setup is balanced, which ensures that each station can send and receive commands. It is only used for one-to-one correspondence.



HDLC Frame format

The first is the beginning sequence, which is 8 bits long and shows you how many bits will be used to indicate the start and end of the frame as seen in . This is also known as the flag bits, and it consists of eight bits, meaning that this is the start of the series. After that, there's a 16-bit header. Your body sections, which are variable in volume, are the third one. We can't define the scale since the body is so big. This body one would receive whatever data it receives from the network layer. The CRC stands for error detection with cyclic redundancy scan, and it has a 16-bit duration. Then comes the ending sequence, which is eight bits long and indicates the end of the frame; the starting and ending sequences are both 06 times one and none. This sequence is also sent while the connection is idle such that the sender and receiver's clocks remain in sync. The starting and ending sequences are 0, 6 times one, then zero.



8	16		16	8
Beginning Sequence	Header	Body	CRC	Ending Sequence

Figure 2 HDLC Frame format

During some times when the connection is idle, this sequence is also transmitted. Why? So that the sender and receiver can keep their clocks synchronized. The next field is the header, which I've already showed you, followed by the address and control fields, which hold the address and control foot. And there's the body, which is a variable-size payload of whatever data we're getting from the satellite. The next section is the shell, which is a variable-size payload information that any data we get from the network layer will be inserted into this body one CRC cyclic redundancy search error detection mechanism. We've already seen that any data is collected from the network layer in the datalink layer is added to the header and the trailer trailer, which are the error detection bits and CRC. So, header refers to the header, and trailer header refers to the trailer header, which has a 16-bit CRC. A trailer is what we call it in this country. Error correction parts will be used in the teaser.

Types of HDLC Frame

There are two fields in the HDLC frame format header: one is known as the address field, and the other is known as the control field. So, what is the purpose of a control field? The purpose is to define the type of HDLC frame.

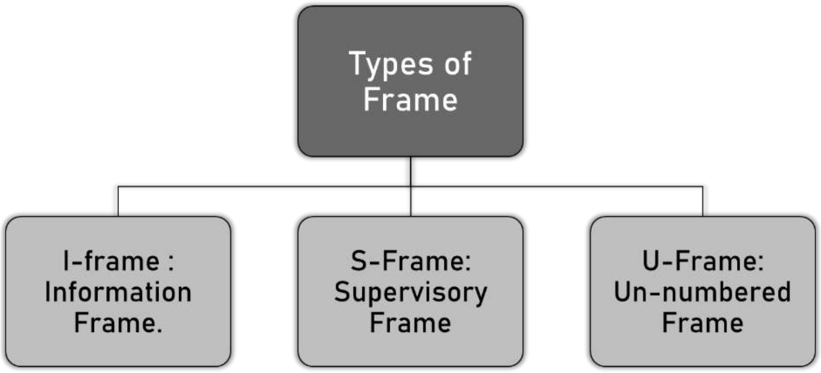

Types of Frame
I-frame : Information Frame.
S-Frame: Supervisory Frame
U-Frame: Un-numbered Frame

Figure 3 Types of Frames

It's as if you're trying to figure out what kind of HDLC frame you have. The control field determines the frame class, and the control field is, in turn, a part of the header. So there are three different kinds of HDLC frames: I-frame, S-frame, and U-frame as shown in Figure 4. The knowledge frame is also known as the I-frame. A supervisor frame is also known as a S frame. Unnumbered frames are also known as U-frames.

I-Frame	1st bit is 0
S-Frame	1st two bits is 10
U-Frame	1st two bits is 11

Figure 4 Types of HDLC Frames

The control field determines the frame forms. As you can see in the table, we have an I-frame, an S-frame, and a U-frame. An I-frame is one in which the first bit in the control field is zero; an S-frame is one in which the first bit in the control field is one zero; and a U-frame is one in which the first two bits or one zero is in the control field. It's a S frame in a U frame, which stands for unnumbered frame. A unnumbered frame is one in which the first bit is one.If the first bit in the control field is zero, the frame is an Iframe, which means it carries information. If the first two bits in the control field are one, zero, the frame is a supervisor frame.So, the supervisor frame plays a part in error management and flow control systems, and the third one is the unnumbered frame. It's used for a variety of tasks, including connection management.

Byte Oriented Protocol

Consider the frame as a string of bytes or a character. That is why, as we discussed in the previous lecture, it is often regarded as a character-oriented approach.

In general, we have three protocols.

BISYNC -> PPP -> Binary Synchronous Communication Protocol

PPP -> Point-to-Point Protocol.

DDCMP -> Message Protocol for Digital Data Communication

7.2 Point-to-Point Protocol

It is a data link layer protocol that is available. The Point to Point Protocol (PPP) is a van protocol that is widely used over internet connections. When we say "internet connections," we're referring to two routers. The Point to Point Protocol (P2P) is a commonly used protocol. It's also known as the van protocol. Since there are two major applications for the Point to Point Protocol.the two most popular applications It's commonly used in internet networks that have a lot of traffic and a lot of speed. Obviously, the internet is one of the stuff that has a high load and a fast pace at the same time. It's often used to send data from two multipoint-to-point computers, or two computers that are directly connected.

Components of PPP

- Encapsulation Component
- Link Control Protocol (LCP)
- Authentication Protocols (AP)
- Network Control Protocols (NCPs)

Encapsulation Component

It encapsulates data grams in order for them to be transferred over this physical layer.

Link Control Protocol (LCP)

It is in charge of creating, configuring, checking, maintaining, and terminating connections, as well as negotiating ties. It also makes use of functionality across two end point connections.

Authentication Protocols (AP)

Password authentication protocol and challenge Handshake Authentication Protocol, also known as PAP protocol and CHAP protocol, are two authenticated protocols of Point to Point Protocol that authenticate endpoints for usage of services. The acronyms PAP and CHAP stand for Password Authentication Protocol and Challenge Handshake Authentication Protocol, respectively.

Network Control Protocols (NCPs)

The conditions and services for the networks are negotiated using this protocol.

Frame format of PPP

PPP frames are typically used to encapsulate packets of data or information that only contain configuration data or data. PPP is based on the same fundamental format as HDLC. PPP usually has one more field, the protocol field. After the control field and before the input or data field, this protocol field appears.

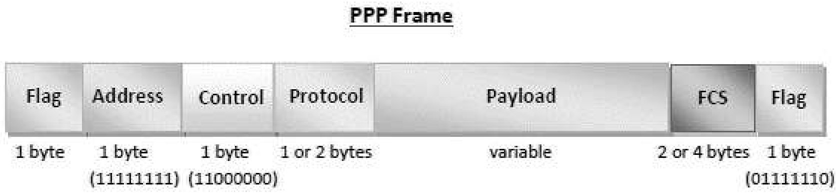


Figure 5 Frame format of PPP

Flag field –

PPP frames are identical to HDLC frames in that they both begin and end with the regular HDLC flag. It always has a 1-byte value, which is 01111110 in binary.

Address filed –

The address field essentially serves as a transmitted address. All 1's basically means that all of the stations are able to embrace frame in this case. It has a 1-byte value, or 11111111 binary value. Specific station addresses are not given or assigned by PPP, on the other hand.

Control filed –

In HDLC, this area primarily uses the U-frame (Unnumbered frame) format. The control field is needed for various purposes in HDLC, but in PPP, it is set to 1 bit, or the binary value 00000011. This 1 byte is used for a data link that does not need a connection.

Protocol field –

This area essentially defines the datagram's network protocol. It usually determines the type of packet in the data field, or what is being transported in the data field. This field is 1 or 2 bytes long and is used to identify the PDU (Protocol Data Unit) encapsulated by the PPP frame.

Data filed –

It normally includes the datagram from the upper layer. For standard PPP data frames, the network layer datagram is especially encapsulated in this region. The length of this field varies rather than being continuous.

FCS filed –

This area normally includes a checksum for the purpose of detecting errors. It can be 16 bits or 32 bits in length. It's even done on the address, access, protocol, and also information fields. Characters are applied to the frame for monitoring and error management.

Byte Stuffing in PPP frame

When the flag sequence appears in the packet, byte stuffing is used in the PPP payload field such that the recipient does not believe it to be the end of the frame. Each byte that comprises the same byte as the flag byte or the escape byte has the escape byte, 01111101, packed before it. Before transferring the message into the network layer, the receiver extracts the escape byte.

7.3 Multiple Access Protocols in Computer Network

The Data Link Layer is responsible for transmission of data between two nodes. Its main functions are-

- Data Link Control
- Multiple Access Control

Data Link control –

The data link control is responsible for reliable transmission of message over transmission channel by using techniques like framing, error control and flow control. For Data link control refer to – Stop and Wait ARQ

Multiple Access Control –

If the sender and receiver have a dedicated connection, the data link control layer is sufficient; but, if the sender and receiver do not have a dedicated link, several stations will reach the channel at the same time. To reduce collisions and prevent crosstalk, multiple access protocols are needed.When an instructor asks a question in a classroom full of students, and all of the students (or stations) begin responding at the same time (send data at the same time), a lot of confusion is generated (data duplication or data loss). It is then the teacher's task (multiple access protocols) to control the students and get them answer one at a time.For data sharing on non-dedicated networks, protocols are thus necessary. Multiple control protocols can be further broken down into the following subcategories:

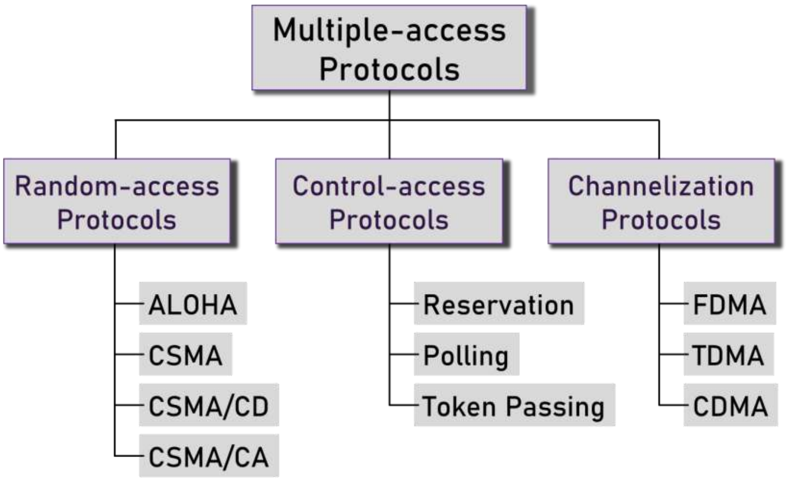


Figure 6 Multiple Access protocols

7.4 Random-access protocols

It manages station connection to the transmission line. It sends out a connection that necessitates the use of an access control system. As a result, random access protocols are used. We use a variety of methods. So, Aloha CSMA, CSMA CD, and CSMA CA, respectively.

(a) ALOHA

It was created for a wireless LAN, but it can also be used for a shared medium. Multiple stations will relay data at the same time in this scenario. This is why, in this situation, we are confronted with a crash and data that is jumbled. If you can see in the Figure 7, if two cars begin transmitting data at the same time or begin using the station at the same time, there is a risk of a crash.

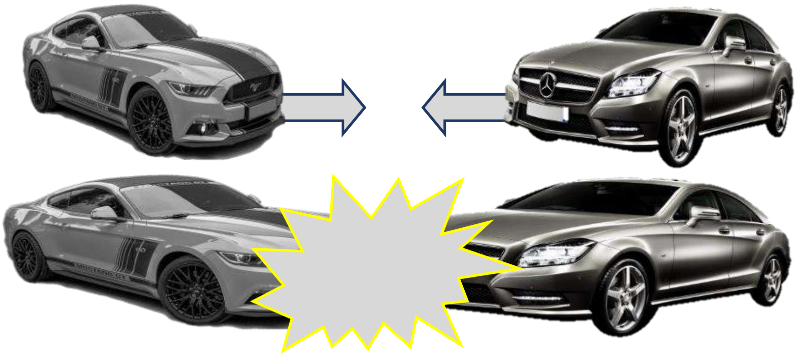


Figure 7 Collision

If more than one station begins using the channel at the same time, a collision can occur. There are two separate interpretations of aloha. One is pure Aloha. The second is slotted ALOHA as shown in Figure 8.

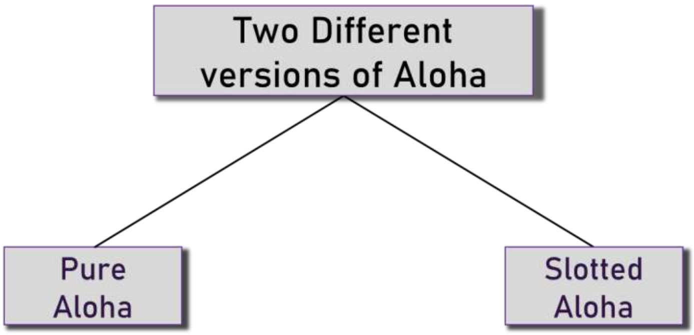


Figure 8 Different versions of ALOHA

Pure Aloha

When a station transmits data, it waits for a response. If the acknowledgment does not arrive within the allotted time, the station waits a random period of time (Tb) before re-sending the results. Since multiple stations take varying amounts of time to wait, the chances of another collision are reduced.

Vulnerable Time = 2* Frame transmission time

Throughput = G exp{-2*G}

Maximum throughput = 0.184 for G=0.5

Slotted Aloha

It's equivalent to pure aloha, except that we split time into slots and data can only be sent at the start of each slot. If a station runs out of time, it must wait for the next available slot. This lowers the chances of a crash.

Vulnerable Time = Frame transmission time

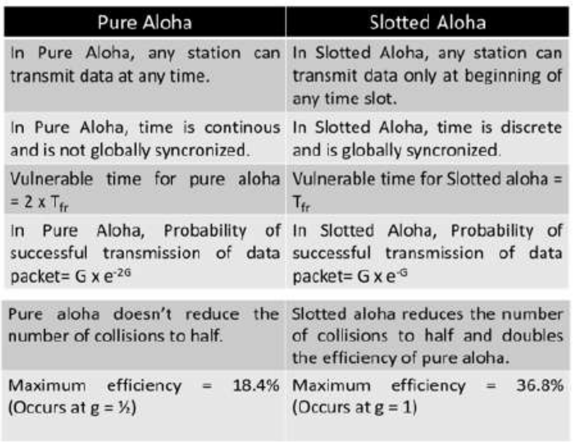
Throughput = G exp{-*G}

Maximum throughput = 0.368 for G=1

Difference between Pure Aloha and Slotted Aloha

The difference between pure aloha and slotted aloha is shown in Table 1.

Table 1 Pure Aloha vs Slotted Aloha



Pure Aloha	Slotted Aloha
In Pure Aloha, any station can transmit data at any time.	In Slotted Aloha, any station can transmit data only at beginning of any time slot.
In Pure Aloha, time is continous and is not globally syncronized.	In Slotted Aloha, time is discrete and is globally syncronized.
Vulnerable time for pure aloha = 2 x T_{fr}	Vulnerable time for Slotted aloha = T_{fr}
In Pure Aloha, Probability of successful transmission of data packet= G x e^{-2G}	In Slotted Aloha, Probability of successful transmission of data packet= G x e^{-G}
Pure aloha doesn't reduce the number of collisions to half.	Slotted aloha reduces the number of collisions to half and doubles the efficiency of pure aloha.
Maximum efficiency = 18.4% (Occurs at g = ½)	Maximum efficiency = 36.8% (Occurs at g = 1)

(b) CSMA

Since the station must first sense the medium (for idle or busy) before transmitting data, Carrier Sense Multiple Access means less collisions. It sends data if the channel is idle; otherwise, it waits for the channel to become idle. However, because of the propagation delay, there is also a risk of a collision in CSMA.

For example: Station A will first feel the medium before sending results. It will begin transmitting data if the channel is found to be idle. If station B asks to send data and senses the medium, it will also find it idle and send data by the time the first bit of data is sent (delayed due to propagation delay) from station A. As a consequence, data from stations A and B will collide.

CSMA access methods: -

1-persistent: The node detects the channel and sends the data if it is idle; otherwise, it continuously checks the medium for idleness and transmits unconditionally(with 1 probability) when the channel becomes idle.

Non-persistent: If the channel is idle, the node sends the data; if not, it tests the medium after a random period of time (not continuously) and transmits when it is found idle.

P-persistent: The node detects the medium and, if it is idle, sends data with a probability of p. If the data isn't sent ((1-p) chance), it waits a while and then scans the medium again; if it's already idle, it sends with p probability. This process is repeated before the frame is submitted. Wifi and packet radio networks both use it.

O-persistent: The superiority of nodes is determined ahead of time, and transmission takes place in that order. Node waits for its time slot to transmit data if the medium is idle.

(c) CSMA/CD

Carrier detects multiple entry points and detects collisions. If a collision is observed, stations may stop data transmission.

(d) CSMA/CA

Multiple access is detected by the carrier, and collisions are avoided. The sender receives recognition signals as part of the collision detection process. The data is successfully transmitted if there is only one signal (its own), so if there are two signals (its own and the one with which it collided), a collision has occurred.

7.5 Controlled access

In monitored access, the stations exchange data to determine which station has the authority to transmit. To stop message collisions on shared medium, it only requires one node to send at a time.

The three controlled-access methods are:

- Reservation
- Polling
- Token Passing

Reservation

Until sending info, a station must make a reservation using the reservation process. There are two types of periods on the timeline:

1. A fixed-length reservation interval

2. Variable frame data transfer time

Polling

The polling procedure is similar to a roll call in class. A handler, like the coach, sends a message to each node in turn.

One serves as the main station (controller), while the others serve as secondary stations. Both data must be exchanged via the controller. The address of the node being chosen for access is included in the message received by the controller.

While all nodes receive the message, only the one to which it is sent responds and sends data, if any. If there is no evidence, a “poll reject” (NAK) message is normally returned as shown in Figure 10 and Figure 9.The polling messages have a high overhead, and the controller's reliability is highly dependent.

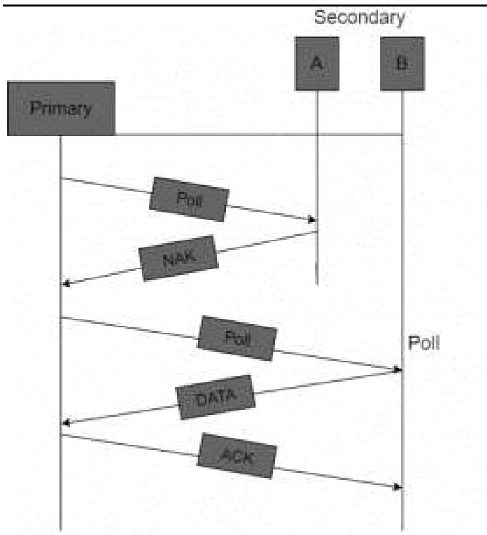


Figure 10 Polling (a)

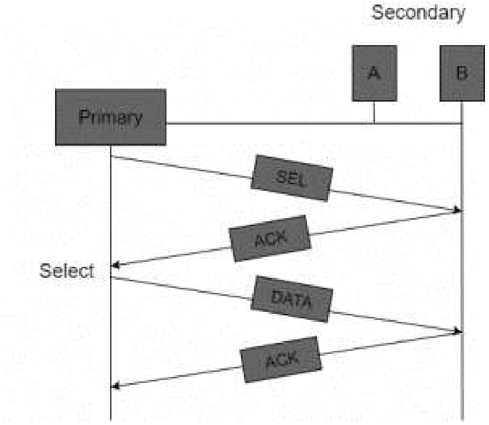


Figure 9 Polling (b)

Token passing

The stations in a token passing scheme are theoretically bound to one another in the form of a loop, and station access is controlled by tokens.

A token is a small message or a special bit pattern that circulates from one station to the next in a predetermined order.

Tokens are exchanged from one station to the next in the ring in the case of Token ring, while in the case of Token bus, each station uses the bus to transfer tokens to the next station in a predetermined order.

In all instances, the token denotes the ability to submit. When a station receives the token and has a frame queued for transmission, it will transfer the frame before passing the token to the next station. If there is no queued loop, it merely transfers the token as seen in Figure 11.

Following the transmission of a frame, each station must wait for all N stations (including itself) to send the token to their neighbors, as well as the other N – 1 station to send a frame if they have one.

There are issues such as token duplication or loss, insertion of a new station, and replacement of a station that must be addressed in order for this scheme to operate correctly and reliably.

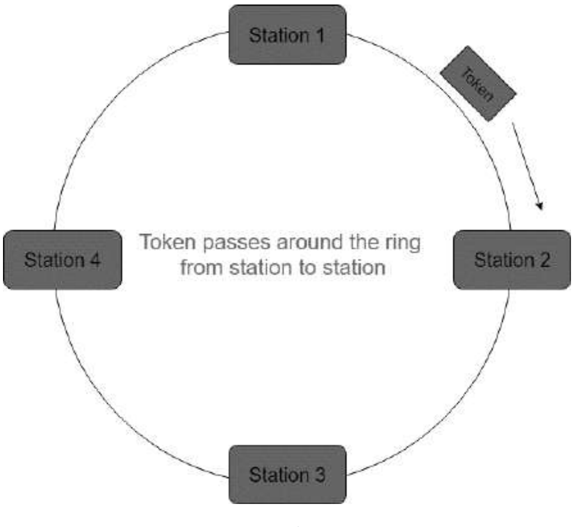


Figure 11 Token Passing

7.6 Channelization Protocols

The usable bandwidth of the link is shared in time, frequency, and code among multiple stations to allow them to access the channel at the same time.

FDMA (Frequency Division Multiple Access)

The usable spectrum is split into equivalent bands, allowing each station to have its own band. To stop crosstalk and disruption, guard bands have been added to ensure that no two bands overlap.

TDMA (Time Division Multiple Access)

The bandwidth is used by a number of different stations. To prevent collisions, time is split into slots, and stations are assigned to relay data during these slots. However, there is a synchronisation overhead since each station must know the time slot. Add synchronisation bits to each slot solves this issue. Another problem with TDMA is propagation delay, which can be overcome by using guard bands.

CDMA (Code Division Multiple Access)

Both signals are carried on a single channel at the same time. There is no such thing as bandwidth or time division. If there are several people in a room speaking at the same time, for example, flawless data reception is also possible if only two people talk the same language. Similarly, data from several stations can be broadcast in various code languages at the same time.

Summary

- The data link layer is divided into two sub layers. The upper sub layer is in charge of data link management, while the lower sub layer is in charge of addressing mutual media connectivity.
- The Finite State Machine model is a technique for verifying the protocol's correctness. The data link protocols PPP and HDLC are commonly used.
- Bluetooth is a proprietary open wireless technology protocol for transmitting data over short distances (using short wavelength radio communications in the ISM band from 2400 to 2480 MHz) between fixed and mobile devices, allowing for the development of highly secure personal area networks (PANs).
- In CSMA/CA, the sender collects the acknowledgement to identify potential collisions, and if there is just one acknowledgement present (its own), the data-frame has been transmitted successfully.
- Until submitting info, a station must make a reservation using the reservation access system. Intervals are used to divide time. A reservation frame precedes the data frames sent in that interval in each interval.

Keywords

Point-to-Point (PPP): It's a data link layer protocol that links two connecting link-level peers at either end of the link over a point-to-point link.

Aloha: Multiple access (MA) to the shared medium is possible with ALOHA. This arrangement has the potential for collisions. When one station sends data, another can try to send data at the same time. The data from the two stations clash, resulting in a jumbled mess.

CSMA: Carrier sense multiple access with collision detection (CSMA/CD) adds collision detection to the CSMA algorithm. After sending a picture, a station tracks the medium to see if the transmission was accurate. If that's the case, the station is over. If there is a collision, though, the frame is sent again.

Controlled access: In controlled access, the stations confer to determine who has the authority to deliver. It is not possible for a station to submit until it has received permission from other stations.

Channelization: Channelization is a multiple-access system in which a link's usable bandwidth is exchanged between separate stations in terms of time, distance, or code.

FDMA: The usable spectrum is split into frequency bands in frequency-division multiple access (FDMA). A band is assigned to each station for data transmission. To put it another way, each band is assigned to a particular station and remains with that station at all times.

CDMA: To gain multiple access in code-division multiple access (CDMA), the stations use separate codes. CDMA is based on coding theory and employs number sequences known as chips.

Self Assessment

1. .... describes the techniques to access a shared communication channel and reliable transmission of data frame in computer communication environment.
2. .... does not include any connection setup or release and does not deal with frame recovery due to channel noise.
3. .... refers to a reliable transfer of bit streams to the network layer for which the data link layer breaks the bit stream into frames.
4. .... controls mismatch between the source and destination hosts data sending and receiving speed and therefore dropping of packets at the receiver end.
5. We have categorized access methods into _________ groups
   a. Two
   b. Three
   c. Four
   d. Five
6. In _______, the stations share the bandwidth of the channel in time.
   a. FDMA
   b. CDMA
   c. TDMA
   d. none of the above
7. In the _______ method, the stations in a network are organized in a logical ring.
   a. Polling
   b. token passing
   c. reservation
   d. none of the above
8. _________ augments the CSMA algorithm to detect collision.
   a. CSMA/CD
   b. CSMA/CA
   c. either (a) or (b)
   d. both (a) and (b)
9. In the _________ method, after the station finds the line idle, it sends its frame immediately. If the line is not idle, it continuously senses the line until it finds it idle.
   a. p-persistent
   b. non persistent
   c. 1-persistent
   d. none of the above
10. In the _____ method, time is divided into intervals. In each interval, a reservation frame precedes the data frames sent in that interval.

a. token passing
b. Reservation
c. Polling
d. none of the above

11. In the _______ method, each station has a predecessor and a successor.

a. token passing

b. polling

c. reservation

d. none of the above

12. The vulnerable time for CSMA is the ________propagation time.

a. three times

b. two times

c. the same as

d. none of the above

Answers for Self Assessment

1.	Data link layer	2.	Unacknowledged connectionless sevice	3.	framing	4.	Rate of data transmission	5.	A
6.	C	7.	B	8.	A	9.	C	10.	B
11.	A	12.	C						

Review Questions

1. What is the data link protocol?

2. List three categories of multiple access protocols.

3. How can a collision be avoided in CSMA/CD network?

4. Compare and contrast CSMA/CD and token passing access methods.

5. Is Slotted Aloha always better than Aloha? Explain your answer with justification.

6. How does PPP transmit data grams over serial point-to-point links?

7. How does PPP establish link for authenticated transfer of file?

8. What are different data link protocols available? Why does PPP have become popular?

9. How does the data link layer accomplish the transmission of data from the source network layer to the destination network layer?

10. Compare and contrast a random-access protocol with a channelizing protocol.

11. Do we need a multiple access protocol when we use the local loop of the telephone company to access the Internet? Why?

12. Define channelization and list three protocols in this category

Further Readings

Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall.

Behrouz A. Forouzan and Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies.

J. D. Spragins, *Telecommunications Protocols and Design*, Addison-Wesley.

Unit 08: Network layer - logical addressing

CONTENTS

Objectives

- what is an IP address?
- types of IP address.
- how to find the public IP address
- classification of IP addresses
- comparison between ipv4 and ipv6
- understand Classful Addressing and classless addressing
- understand NAT, ICMP, ARP, and RARP

Introduction

How can you locate your computer's or smartphone's IP address? Computers on the internet connection with one another through underground or underwater cables, as well as wirelessly. If I want to copy a file from the internet, I think my machine has to have a URL so that other computers on the internet can identify and locate it. In Internet terms, a computer's address is known as an IP address; an IP address is essentially a device's identity; without an IP address, no system can connect with another on the internet. Let's look at another case to better explain what an IP address is. Assume if anyone wishes to deliver mail here and needs a home address. You must have a home address to deliver mail to others. Similarly, whether you want to send or communicate with the machine's programs, or if you want to copy something from any internet-connected device to your machine, the computer requires an address as seen in Figure 1.

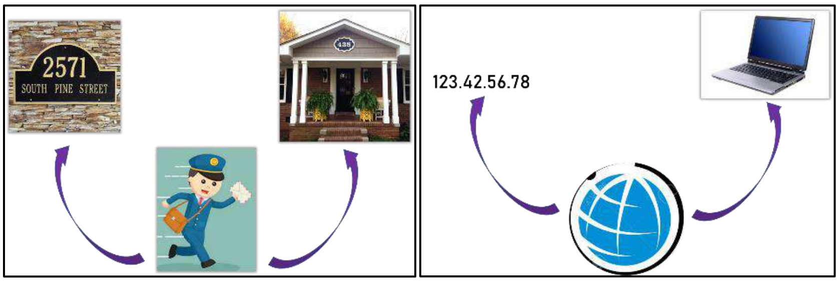


Figure 1 Comparison of a home address with IP address

Such that machines on the internet give you a file that you wish to download, that address is called an IP address in internet terms. An IP address is nothing more than a set of numbers written in a certain format. Now, this is your home address, as determined by your local postman, and it is to this address that you receive all mail. And your home's identity is based on its address. Someone would like to deliver a message. So, the postman will send the letter to you because they know this is the address and we have to bring this letter to this same address. Similarly, the machine has an IP address that is used to identify it on the internet, and you have access to all downloadable files. This computer's address is: As a result, this is the system's IP address.

8.1 IP address and its full form

The Internet Protocol Address is the entire version of the IP address. What exactly is the concept of Internet Protocol? The Internet Protocol (IP) is a series of laws that govern how the Internet functions.

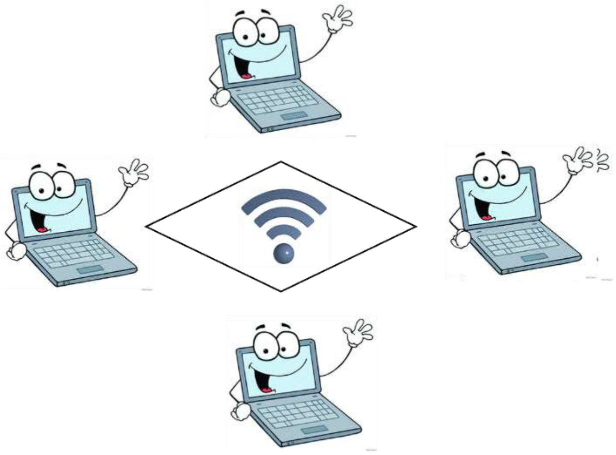


Figure 2 IP Address

Two Types of IP Addresses

So now we have two kinds of IP addresses, one for IP version four and the other for IP version six as seen in Figure 3.

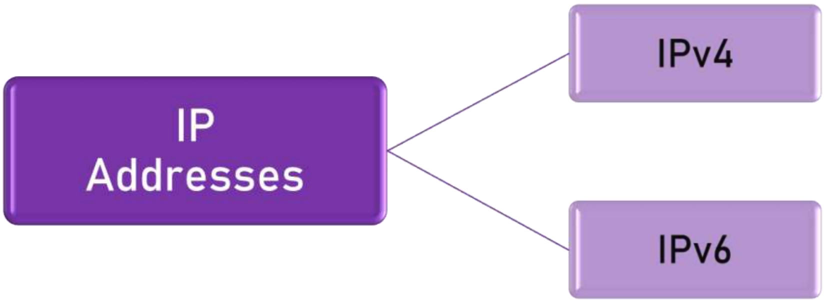


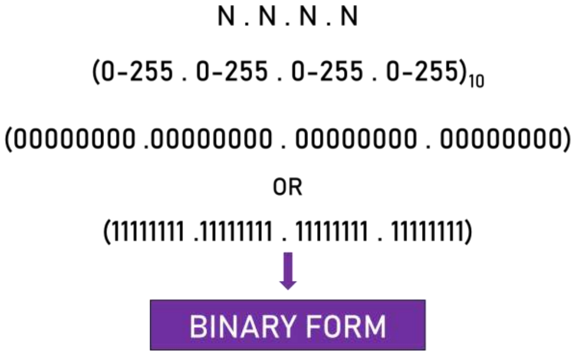
Figure 3 Two types of IP Address

What exactly is IP version 4? Why have we gone on to IP version six while we have IP version four? But, first and foremost, let's define IP version four. The full form of IP version four is internet protocol version four, which consists of four numbers separated by a dot and dot n dot n dot n, which means there are four numbers separated by a dot operator as seen below.

N . N . N . N

0-255 . 0-255 . 0-255 . 0-255

The IP address is made up of numbers in the range of zero 0 to 255. On 192.168.0.1, I can write. However, if I send a message to 256.168.0.1, this is an invalid IP address since the range should be between 0 and 255. However, machines are incapable of comprehending decimal numbers. Since we're written in decimal numbers, then every range we're writing in between 0 and 255 is in decimal numbers. Computers, on the other hand, do not grasp decimal numbers so they are perplexed by them. So, what is the answer to this? As a result, we must translate these decimal numbers into binary numbers that the machine can comprehend. As a result, the binary spectrum can be written as shown in Figure 4.



N . N . N . N

$(0\text{-}255 . 0\text{-}255 . 0\text{-}255 . 0\text{-}255)_{10}$

(00000000 .00000000 . 00000000 . 00000000)

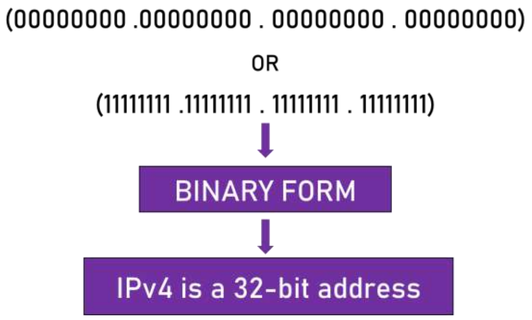
OR

(11111111 .11111111 . 11111111 . 11111111)

BINARY FORM

Figure 4 IP in Binary Form

So IP version four is a 32-bit address, which is a special sequence of ones or zeros allocated to each computer; we call it unique since no two devices will use the same IP address as seen in Figure 5.



(00000000 .00000000 . 00000000 . 00000000)

OR

(11111111 .11111111 . 11111111 . 11111111)

BINARY FORM

IPv4 is a 32-bit address

Figure 5 IP is a 32-bit address

That is why it is referred to as a specific address. Any device connecting to the internet will have a different IP address. They have a specific IP address any time they connect to the internet. So, the question now is how many devices they can address. As a result, a total of 2power32, or roughly 4 billion computers, can be addressed and wired to the internet. Version four of IP. But it's only capable of addressing 4 billion devices. As a result, we can only bind 4 billion computers as shown below.

2^{32} = 4,294,967,296 Devices

ISSUE with 32-bit IPv4

Nowadays, every machine, every user, has a laptop, tabs, and several devices; they have a desktop, their desktop, and they use smartphones as well. As a result, they like the IP address of any device. Also in today's world, not every machine is present in every household. Everyone has their scheme. Everyone has their smart screen much of the time. They want to connect to the internet, and to do so, they would need IP addresses. As a result, IP version four will not be able to supply IP addresses to all of them because it can only communicate with 4 billion addresses. As a result, it can only have 4 billion addresses. So far, we've surpassed the 4 billion mark, and we're on our way to IP version six.

IPv6

Let's take a look at IP version six now. IP version 6 is a 128-bit address written as a series of eight hexadecimal digits separated by the column in human-readable mode. IP version six can be written as, as you can see here, in a human-readable format. However, computers do not understand hexadecimal, so we must convert it back to binary. It'll be divided into 128 ones and zeros. Since each device linked to the internet is given this sequence of zeros and ones, since no two machines should be allocated the same IP address, this is a special series number. In this example, IP version four is the same, so both would be unique. With IP version six, 2^{128} computers to the internet.

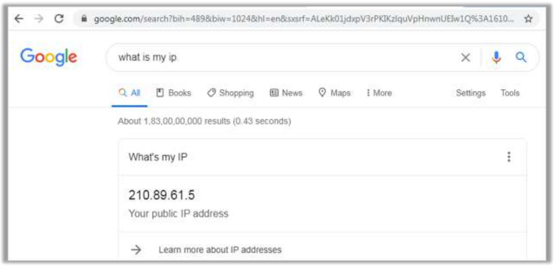
This is the value that has been seen here: 2 power of 128 equals as seen in Equation 1. So, with this many addresses, it will address this many devices. This importance indicates that it can handle this many computers, which is much more than sufficient for future generations.

$$2^{128} = 340{,}282{,}366{,}920{,}938{,}463{,}463{,}374{,}607{,}431{,}768{,}211{,}456$$

Equation 1

How to Find Your Computer Public IP Address?

So, first and foremost, go to Google and type in "what is my IP." As you can see, we've opened Google and are typing in "what is my IP." As a result, it will show your public IP address. Even now, the only concern is how you'll search your smartphone's public IP address.



google.com/search?bih=489&biw=1024&hl=en&sxsrf=ALeKk01jdxpV3rPKlKzlquVpHnwnUEIw1Q%3A1610...

what is my ip

About 1,83,00,00,000 results (0.43 seconds)

What's my IP

210.89.61.5

Your public IP address

Learn more about IP addresses

Figure 6 What is my IP

So, in Google, write down my IP address. As a result, Google will inform you of your smartphone's public IP address. One very important thing to remember is that we won't be able to connect to the internet without IP versions four and six; an IP address is needed for an internet connection. You believe that if you don't have an IP address, you won't be able to log in to the internet, which means you won't be able to download or upload something, and you won't be able to check something. But you'll need an IP address if you want to connect to the internet.

Classification of IP Addresses

IP addresses are divided into two categories: dynamic IP addresses and static IP addresses.

Dynamic IP Address

When you bind your device or smartphone to the internet, this is what happens. Then there's Internet Service Provider ISP, which is a company that provides an IP address from a pool of valid IP addresses. You have an IP address, then you can connect to the internet. Once connected, you can transmit and receive data to and from other devices on the internet. So, the next time you want to connect to the Internet, do so. ISP will assign you a new IP address with the same availability set, which is why it is called dynamic. Dynamic means that it changes all of the time. DHCP assigns a unique IP address to every device.

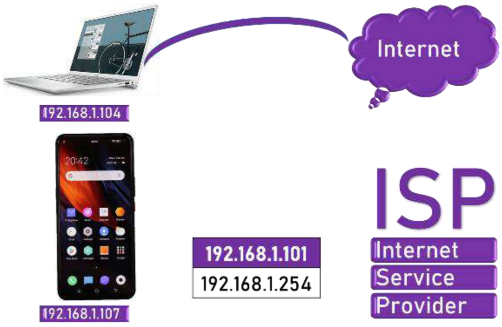


Figure 7 Dynamic IP Address

Static IP Address

It is the only other classification that remains constant. Who is using this static IP address and who are the domain name servers that are using this IP address? What is the DNS server actually, these are the computers that help you to open a website on your computer?

A static IP address provides information, such as the device is located in which continent. In which country, which city, and which is the internet service provider, that is providing the internet connection to that device.

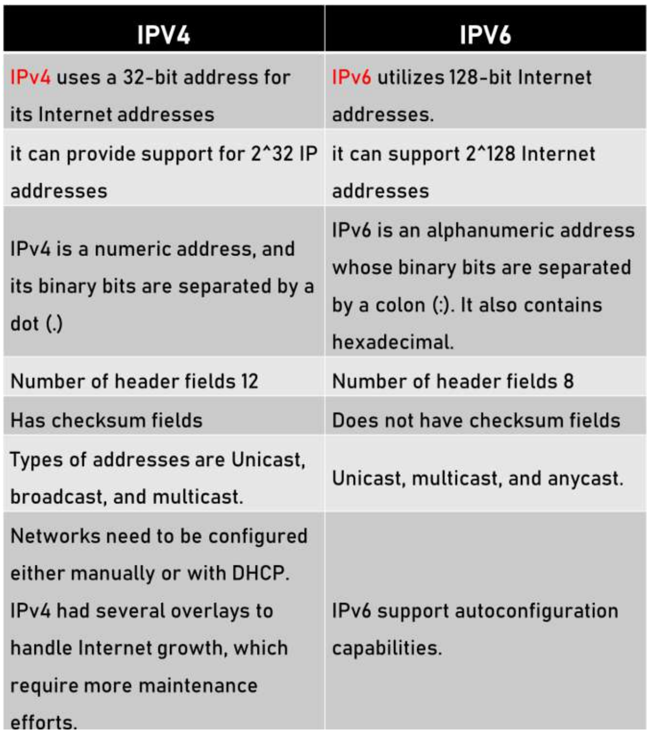
Figure 8 Example of Static IP Address

It is an internet service provider that assigns IP addresses to computers so that they can connect to the Internet. We can monitor the location of a computer connecting to the internet until we know the ISP. As a result, IP addresses allow billions of devices to be identified. What is the disadvantage of using a static IP address? One of the most significant disadvantages is that it is less safe so it is easy to trace. As a result, we can classify it as less stable.

8.2 Comparison between IPv4 and IPV6

Table 1 shows the differences between IPv4 and IPv6.

Table 1 Comparison between IPv4 and IPv6



IPV4	IPV6
IPv4 uses a 32-bit address for its Internet addresses	IPv6 utilizes 128-bit Internet addresses.
it can provide support for 2^32 IP addresses	it can support 2^128 Internet addresses
IPv4 is a numeric address, and its binary bits are separated by a dot (.)	IPv6 is an alphanumeric address whose binary bits are separated by a colon (:). It also contains hexadecimal.
Number of header fields 12	Number of header fields 8
Has checksum fields	Does not have checksum fields
Types of addresses are Unicast, broadcast, and multicast.	Unicast, multicast, and anycast.
Networks need to be configured either manually or with DHCP. IPv4 had several overlays to handle Internet growth, which require more maintenance efforts.	IPv6 support autoconfiguration capabilities.

8.3 Classful Addressing

From 1981 before Classless Inter-Domain Routing was implemented in 1993, classful addressing was a network addressing the Internet's architecture. Centered on four address bits, this addressing system splits the IP address into five distinct groups. Group A, B, C, D, and E are the five categories in which classful addressing is divided. An IP address is a 32-bit unique address with a 2power32 address space.

In general, IP addresses are written in one of two ways: dotted decimal notation or hexadecimal notation as seen in Figure 9. It means classful addressing can be represented in binary also, and it can be represented as a decimal also.

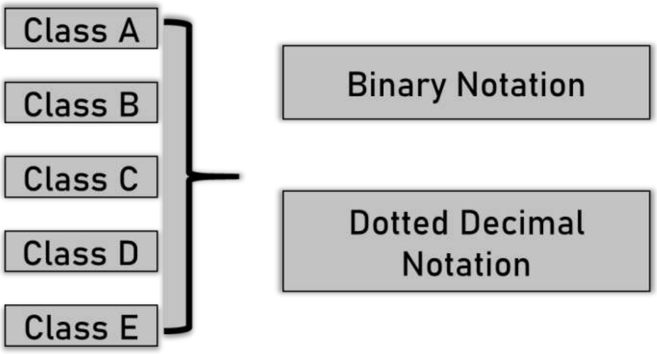


Figure 9 Classful Addressing

Binary Notation

In binary notation, the IP version four address is displayed as 32-bits each octet is often referred to as a byte. So, it is common to hear an IP version four address refers to as a 32-bit address or a four-byte address.

The following is an example of an IPv4 address in binary notation:

01110101 10010101 00011101 00000010

As a result, the IP version four address has been compressed and made easier to read. Internet addresses are normally written with the decimal point dividing the bytes in decimal form.

Decimal Notation

So, we have two tools, one of which is binary and the other of which is decimal, and the same binary is translated into decimal notation. So, there are two types of notations: dotted decimal and binary.

Dotted-decimal notation and binary notation for an IPv4 address

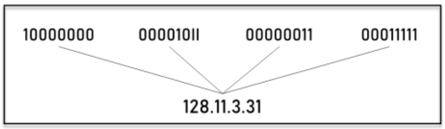


Figure 10 Notation for IPv4

As a result, the IP version four address has been compressed and made easier to read. Internet addresses are normally written with the decimal point dividing the bytes in decimal form as seen in Figure 10.

Important points to remember

1. There must be no leading zero (045).
2. There can be no more than four numbers in an IPv4 address.
3. Each number needs to be less than or equal to 255 (301 is outside this range).

255.272.255.255

4. A mixture of binary notation and dotted-decimal notation is not allowed

172. 00101000. 254. 11001100

Find the error, if any, in the following IPv4 addresses

a. 111.56.045.78

b. 221.34.7.8.20

c. 75.45.301.14

d. 11100010.23.14.67

- When writing in decimal notation, it cannot be begun from zero like it is in the first one.
- It is divided into five octets in the B, so five octets are not allowed.
- The range for the third one should be between 0 and 255. The 301 series is not accurate.
- The fourth one uses a combination of binary and dotted decimal notation.

How would you be able to know to which class it belongs

When your IP address is written in binary notation and the first bit is 0, it means it's a class A address. If the first two bits of your IP address are one zero, your IP address is classified as Class B. If the first three bits of the binary notation are 110, It is classified as Class C. It belongs to Class D if the first four bits are triple one and zero. It belongs to Class E if the first four bits are 11110 as seen in Figure 11.

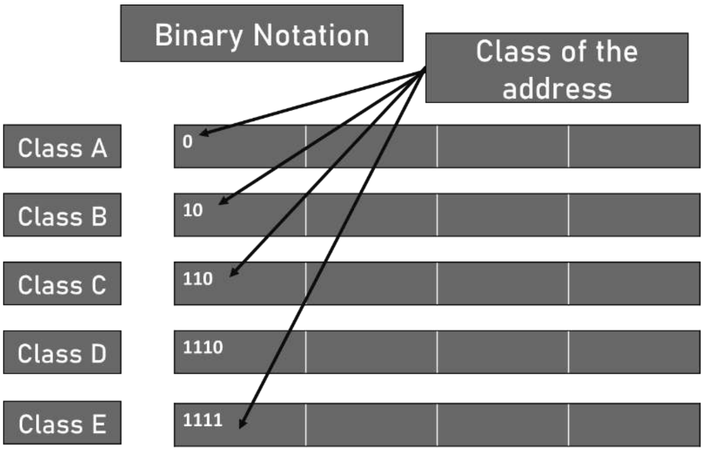


Figure 11 Class of the address in Classful Addressing

Example of Binary Notation

So, since the first bit is zero, it is classified as Class A. Let's look at another case. So far, only the first three bits have been discovered. The first two bits are one, and the third bit is zero, indicating that this is a Class C address as seen below in Figure 12.

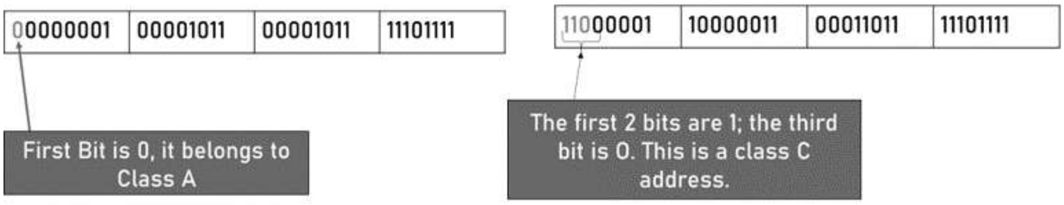

00000001
00001011
00001011
11101111
First Bit is 0, it belongs to Class A
11000001
10000011
00011011
11101111
The first 2 bits are 1; the third bit is 0. This is a class C address.

Figure 12 Example of Binary Notation

Example of Dotted Decimal Notation

If my IP address is expressed in decimal form. So, how can I determine the class it belongs to? So, there's a range, and it's divided into classes, and your IP addresses are divided into those classes. The size of Class A is 0 to 127. Class B has a range of 128 to 191, Class C has a range of 192 to 223 and Class D has a range of 224 to 239, and Class E has a range of 242 to 255.

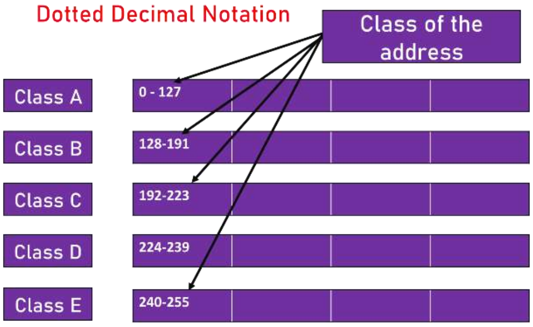
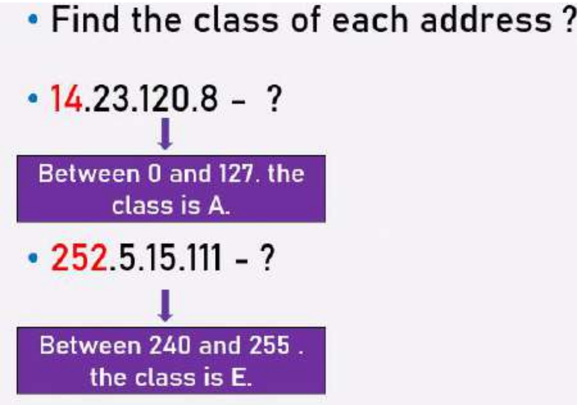


Figure 13 Dotted Decimal Notation

So, if the first byte is 14, we now have a total of 14 bytes. It indicates that they fall within the range of 0 to 127, which is class A. 252, and that they fall within the range of 240 to 255, which is class E.



- Find the class of each address ?
- 14.23.120.8 – ?

Between 0 and 127. the class is A.

- 252.5.15.111 - ?

Between 240 and 255 . the class is E.

Figure 14 Example of decimal notation

Division of Classes

How can we know which part is a network ID and which part is a host ID if groups are split into two parts? The bits used for network ID and host ID, as well as the total number of networks and hosts possible in that class, are determined by the IP address class. Each computer connecting to a network is assigned an IP address by the ISP or network administrator. As a result, an IP address in class A, B, or C is split into Net ID and host ID in a classful addressing scheme. The length of these pieces varies depending on the address's class. So keep in mind that Class D and E are reserved addresses.

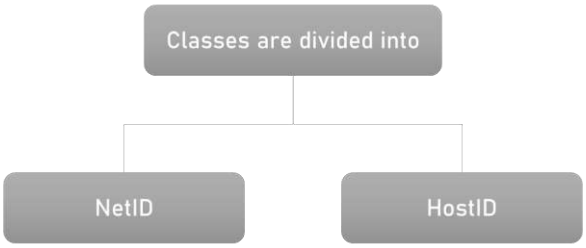
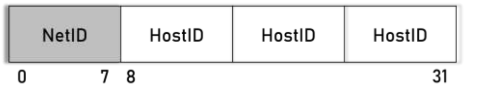


Figure 15 Division of Classes

Division in Class A

So, in class A, the first eight bits represent the net ID and the remaining bits represent the host ID; the remaining bits representing host ID, so the host ID is represented by 24 bits as seen in Figure 16.



0.0.0.0 to 0.255.255.255
127.0.0.0 to 127.255.255.255

We never assigned these to any hosts…these representing default network.

0 – default network

127---Loopback address…check whether NIC is working properly or not

Figure 16 Division in Class A

Important Points about Class A

- IP address begins with 0

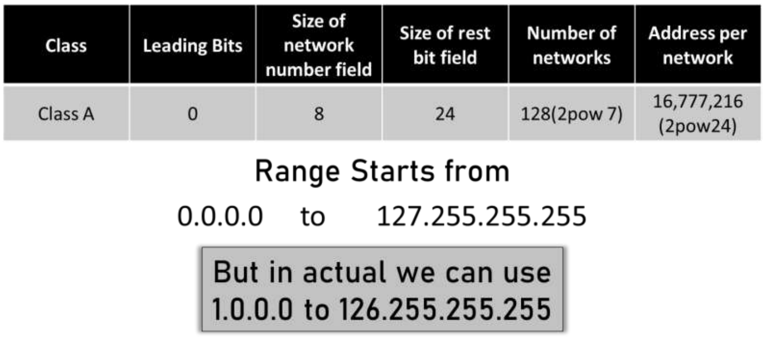
- 7 remaining bits in the network part
  - Only 128 possible class A networks.
- 24 bits in the local part
  - Over 16 million hosts per class A network.
- All class A network parts are assigned or reserved

As you can see in Figure 16, zero to seven show the network element, so how many bits are there, total eight bits, and how many bits are there for the host, 24 bits, so each block has eight bits and each byte has eight bits, so eight plus eight plus eight equals 24 bits. But, as you can see, we're going to represent it in binary notation, as you can see here.

Exceptions in the class A IP addresses

The term "exception" refers to the fact that there are certain networks that we are unable to access, such as network 0, which is the default network. 127 is the loopback address, which is used to determine if the network interface card is functioning correctly or not. It is also known as the loopback address. Table 2 shows a summary of Class A IP Addresses.

Table 2 Summary of Class A IP Address



Class	Leading Bits	Size of network number field	Size of rest bit field	Number of networks	Address per network
Class A	0	8	24	128(2pow 7)	16,777,216 (2pow24)

Range Starts from

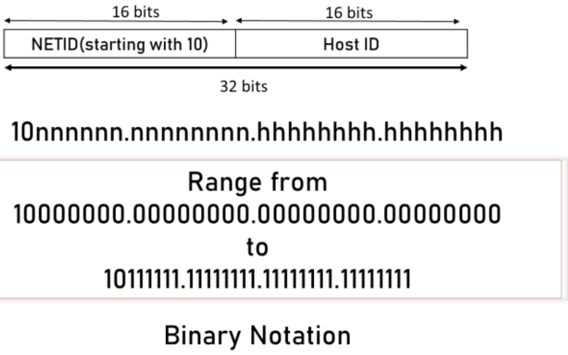
0.0.0.0 to 127.255.255.255

But in actual we can use
1.0.0.0 to 126.255.255.255

When a leading bit is 0, it indicates that the first bit should also be zero. The network number area is eight characters long. The remainder of the bit is 24 network area, which means that it is first in the network field, so all of the other bits represent the host, so 24 plus eight equals 32 number of networks 128 address per network two raised to the power of 24. But now we're using one to 126, which is the range of classes zero to 127, except zero, is reserved for the default network and 127 is the loopback address. But we might assume that we're using a scale of one to 126.

Division in Class B

The network portion is represented by the first 16 bits. The host part is represented by the next 16 pieces as seen in Figure 17. A balance between the number of networks and the number of hosts is achieved by using 16 bits for the network portion and 16 bits for the host part.



10nnnnnn.nnnnnnnn.hhhhhhhh.hhhhhhhh

Range from
10000000.00000000.00000000.00000000
to
10111111.11111111.11111111.11111111

Binary Notation

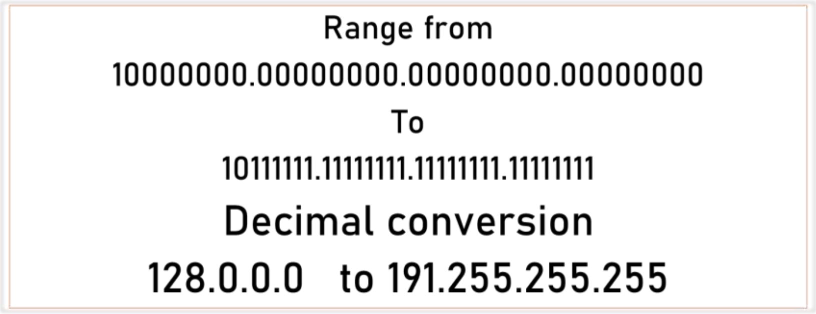
Figure 17 Division in Class B

Important Points in Class B

- 16 bits for network part 16 bits for the host part

- A balance between the number of networks and hosts.
- IP address begins with 10(1st zero in the second position)
- 14 remaining bits in the network part
  - Over 16,000 possible class B network
- 16 bits in the local part
  - Over 65,000 possible hosts

The network portion is represented by the first 16 bits. The host part is represented by the next 16 pieces. A balance between the number of networks and the number of hosts is achieved by using 16 bits for the network portion and 16 bits for the most part. So, in decimal notation, we'll write down more, such as several zero to 127 for class A, and 128 to 191 for class B. We've translated the binary to decimal, so 128 to 191 is the result as seen in Figure 18.



Range from

10000000.00000000.00000000.00000000

To

10111111.11111111.11111111.11111111

Decimal conversion

128.0.0.0 to 191.255.255.255

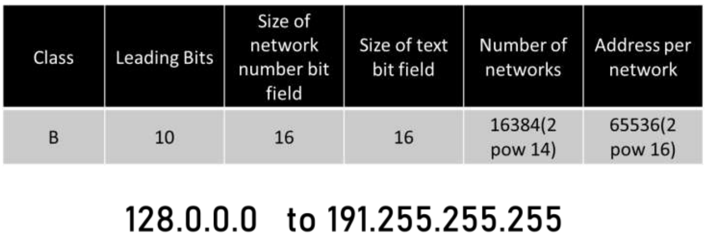
Figure 18 Conversion to Decimal Notation

Exceptions in Class B Addresses

- Exceptions in Class B IP Address
  - 169.254.x.x
- We never assigned this IP address to any host because this range reserve for APIPA.
  - 169.254.x.x is a private IP Addressing space reserved by Microsoft which it assigns automatically to your Network Adapter if it cannot get an IP Address from the DHCP Server

Table 3 has a list of Class B. The first two bits should be one zero, as we remember. Only then may we assume it is a Class B address, with a network number bit field of 16 bits and a text bit field of 16 bits. The number of networks can be represented by 2 power 14 and the number of addresses per network by 2 power 16. As a result, the scale is 128 to 191.

Table 3 Summary of Class B IP Address

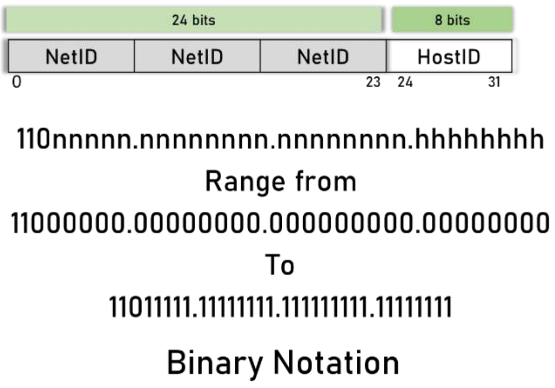


Class	Leading Bits	Size of network number bit field	Size of text bit field	Number of networks	Address per network
B	10	16	16	16384(2 pow 14)	65536(2 pow 16)

128.0.0.0 to 191.255.255.255

Division in Class C

When it comes to class C, the first 24 bits represent the network, while the last eight bits represent the host as seen in Figure 19. It is a common and widely used IP address that starts with the 110 or the first zero in the third position in the case of a Class C address. There are over 2 million potential class C networks with 21 more bits in the network part and eight bits in the local part.



110nnnnn.nnnnnnnn.nnnnnnnn.hhhhhhhh

Range from

11000000.00000000.000000000.00000000

To

11011111.11111111.111111111.11111111

Binary Notation

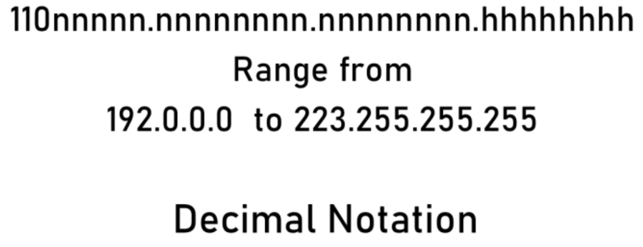
Figure 19 Division in Class C

So, as you can see in the binary notation, the first three bits are set, which is 110, and the next eight bits are representing the network element, and the last eight bits are representing the host part, which we represented with the 111 and the 000, as in the binary notation here.

Important points in Class C

- 24 bits for the network address and 8 bits for host
- Popular and commonly used
- IP address begins with 110(1st zero in 3d position)
- 21 more bits in the network part
    - Over 2 million possible class C networks
- 8 bits in the local part
    - Only 256 possible hosts per class C network.

When we translate this to decimal notation, it becomes 192 to 223 as seen in **Error! Reference source not found.**.



110nnnnn.nnnnnnnn.nnnnnnnn.hhhhhhhh

Range from

192.0.0.0 to 223.255.255.255

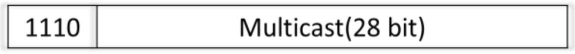
Decimal Notation

So, let's look at this overview of a Class C leading bits: 110 since it begins with 110, the size of the network number bit field is 24, the text bit field is eight, and the number of networks is two lift to power 21 because 110 is reserved out of 24 network bits. The number of addresses per network has been increased from 2 power 8 which equals to 256. The summary of Class C is shown in Table4.**Error! Reference source not found.**

Class D

The IP address range now begins with 224 and ends with 239, which is the range of IP addresses. The first four bits are always 1110, which is used for multicasting IP addresses and always starts with zero. As a result, it delivers a message to a network of hosts rather than just one. There are no network or host bits specified. The first four bits are always triple one, zero, as you can see. When we are expressed in binary form as seen in Table 4.

Table 4 Range of Class D

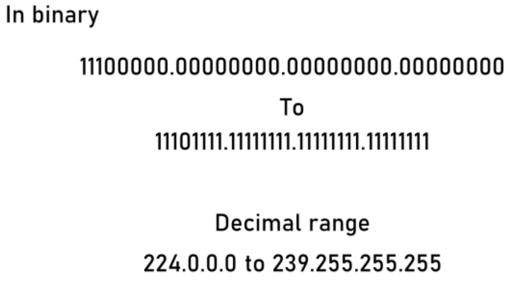


1110	Multicast(28 bit)

Important Points in Class D

- Starts from 224.0.0.0 to 239.255.255.255
- There is no subnet mask for Class D.
- Used for multicasting (one to many types)
- IP address begins with 1110
- Used for multicasting, not defining networks
    - Sending a message to a group of hosts
    - Not just to one
    - Do not define network or host bits
    - Different from broadcasting
    - Used for videoconferencing.

As a result, when we translate this to decimal notation. As a result, the ranges are 224 to 239 as shown in Figure 20.



In binary

11100000.00000000.00000000.00000000

To

11101111.11111111.11111111.11111111

Decimal range

224.0.0.0 to 239.255.255.255

Figure 20 Decimal Conversion in Class D

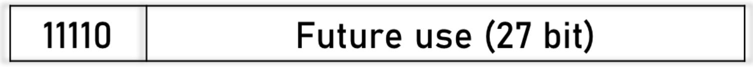
Popular Class D IP Addresses

- 224.0.0.0: Base address reserved
- 224.0.0.1: Used for all multicasting host groups
- 224.0.0.2: Used for all subnet routers
- 224.0.0.5 and 224.0.0.6: Used by Open Shortest Path First, an interior gateway protocol for all network segment routing information

Class E

Class E IP addresses are only used for experimental and testing purposes. Class E IP addresses range from 240.0.0.0 to 255.255.255.254. There is no sub-net mask for this class. The first octet of class E's higher-order bits is always set to 11110 as shown in Table 5.

Table 5 Range of Class E



11110	Future use (27 bit)

Important points in Class E

- Class E IP addresses starts from 240.0.0.0. to 255.255.255.255
- Reserved for experimental purposes.
- IP address begins with 1111.

In Binary Form

11110000.00000000.00000000. 00000000

To

11111111.11111111.11111111. 11111111

In decimal form

240.0.0.0

To

255.255.255.255

Range of Special IP addresses

169.254.0.0 – 169.254.0.16	Link-local addresses
127.0.0.0 – 127.0.0.8	Loop-back addresses
0.0.0.0 – 0.0.0.8	used to communicate within the current network.

Issues with Classful Addressing

There are many issues with classful addressing, the most serious of which is the lack of a network class capable of supporting a medium-sized domain efficiently. A Class C network with 254 hosts is typically too thin, whereas a Class B network with 65,534 hosts is much too big. A request for a network address block from a medium domain is typically addressed by assigning a Class B network address rather than several Class C network addresses, losing thousands of possible host addresses, and causing a detrimental effect on the Internet routing tables.

Design of Classful Addressing

- Class A addresses were designed for large organizations with a large number of attached hosts or routers.
- Class B addresses were designed for midsize organizations with tens of thousands of attached hosts or routers.
- Class C addresses were designed for small organizations with a small number of attached hosts or routers

8.4 Classless Addressing

Subnetting is a technique for reducing IP address waste in a block. We use the host id bits of a classf ul IP address as the net id bits. We specify the IP address and the number of bits for the mask (usua lly followed by a '/' symbol), for example, 192.168.1.1/28. The subnet mask is identified by setting the specified number of bits out of 32 to 1, for example, in the given address, we need to set 28 out of 32 bits to 1 and the remainder to 0, resulting in 255.255.255.240 as the subnet mask.

Important points in Classless Addressing

- IP Addressing with Classless Addressing is a more advanced scheme.
- It improves the efficiency of IP address allocation.
- It takes the place of the older class-based addressing scheme.
- Classless Inter-Domain Routing is another name for it (CIDR).

CIDR Block

CIDR automatically assigns a block of IP Addresses depending on such rules when a user requests a certain number of IP Addresses. This block includes the number of IP addresses that the user has requested. The CIDR block is a set of IP addresses.

Rules for Creating CIDR Block-

The following three rules are used to build a CIDR block:

Rule -01:

The CIDR block's IP addresses must all be contiguous.

Rule-02:

The block's dimension must be presentable as a power of two.

The total number of IP addresses in the block is the block's size. Any CIDR block will still have a size of 2pow1, 2pow2, 2pow3, 2pow4, 2pow5, and so on.

Rule-03:

The block's first IP address must be divisible by the block's height.

CIDR Notation

CIDR IP Addresses look like-

a.b.c.d / n

- They finish with a dash and a number known as an IP network prefix.
- The IP network prefix specifies the number of bits used to identify a network.
- The remainder of the bits was used to identify hosts in the network.

Example

An example of a CIDR IP Address is-

182.0.1.2 / 28

It implies-

- The detection of the network is done with 28 bits.
- The remaining four bits are used to identify hosts in the network.

8.5 Subnetting

Subnetting is the practice of breaking up a big network into smaller networks. A single-wide network is similar to a city with no sectors or street addresses. A postman in such a town could take 3 to 4 days to locate a single address. He will conveniently locate any address in less than an hour if the town is separated into sectors and streets.

Let's look into another scenario. There will be a planned power outage due to repairs. If the town is split into sectors, the electric department will make a local announcement for the affected sector instead of a general announcement. Aside from these two examples, there are several other real-life examples of massive structures being broken into smaller parts. The same principle applies to computer networks. Subnetting is a computer networking technique that divides a wide IP network into smaller IP networks called subnets. The default class A, B, and C networks each have 16777214, 65534, and 254 hosts. When there are a large number of hosts on a single network, problems such as broadcast, collision, and congestion arise.

Let's look at a case. There are four divisions of a company: manufacturing, distribution, growth, and management. There are 50 users in each department. A private class C IP network was used by the company. Both computers can run in a single big network if there is no subnetting.

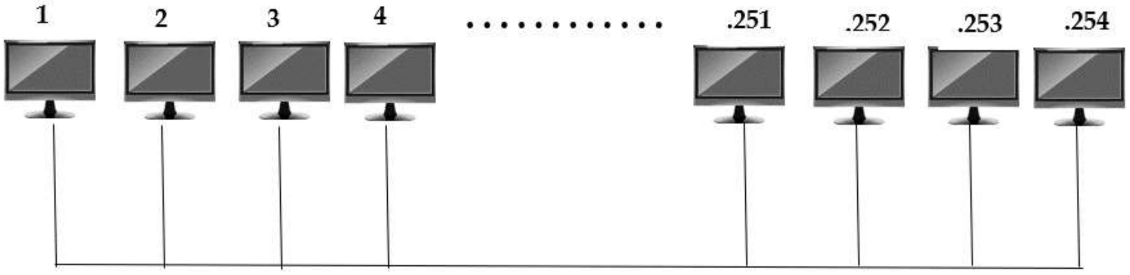


Figure 21 A single large Class C IP Network

To enter and supply information in a network, computers use transmitted signals. In a computer network, a transmitted message is an announcement message that is heard from all hosts in the network. Since all machines are connected to the same network, they can receive all broadcast signals, whether or not the broadcast messages are important to them. This network can be split into subnets in the same way as a city is divided into sectors. When a network is split into subnets, machines can only transmit broadcasts that pertain to them. Since the corporation has four departments, the network can be divided into four subnets.

Description	Network1	Network2	Network3	Network4
Network address	192.168.1.0	192.168.1.64	192.168.1.128	192.168.1.192
Valid hosts	192.168.1.1 to 192.168.1.62	192.168.1.65 to 192.168.1.126	192.168.1.129 to 192.168.1.190	192.168.1.193 to 192.168.1.254
Broadcast address	192.168.1.63	192.168.1.127	192.168.1.191	192.168.1.255

Advantages of Subnetting

- Subnetting is a technique for dividing a big network into smaller networks. Small networks are simple to administer.
- By having only the transmitted traffic that is important to the subnet, subnetting eliminates network traffic.
- Subnetting increases the network's overall efficiency by reducing unwanted traffic.
- Subnetting improves network stability by blocking a subnet's traffic inside a subnet.
- Subnetting decreases the number of IP addresses needed.

Disadvantage of Subnetting

- To connect, different subnets need an intermediary system known as a router.
- Getting more subnets means losing more IP addresses so each subnet has its network and broadcast addresses.
- Subnetting increases the network's complexity. The subnetted network must be managed by an experienced network administrator.

8.6 Network Address Translation

It's a method of converting private IP addresses to public IP addresses. It was first mentioned in RFC 1631.

What is the reason for its implementation?

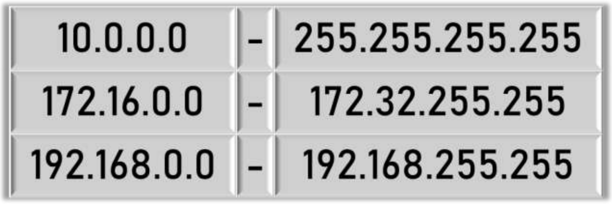
NAT was implemented due to the scarcity of IP addresses. There was a time when we believed that the Internet's expansion would be halted due to the scarcity of IP addresses. If we have five networks, we want internet access to each of them. since the need for IP addresses has risen As a result, we assumed that the internet's growth would eventually slow.

The solution to the implementation of IP Addresses

However, a short-term solution to the issue of IP address depletion or exhaustion has been given. One is IP version six, which is a long-term workaround that we have already implemented. One is CIDR classless inter-domain routing, which is a short-term solution, and the other is NAT network address translation, which is a long-term solution.

NAT is a method for converting private IP addresses to public IP addresses. It's a method of preserving IP addresses. There are a variety of private addresses available as seen in Table 6.

Table 6 Range of Addresses

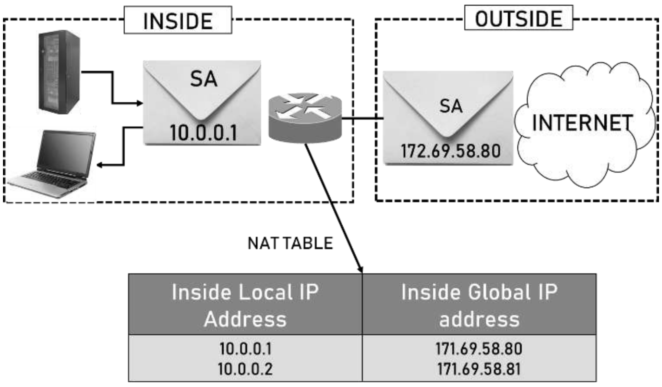


10.0.0.0	-	255.255.255.255
172.16.0.0	-	172.32.255.255
192.168.0.0	-	192.168.255.255

A variety of private addresses are available. These IP addresses have been allocated to private networks for you to connect to the Internet.

Role of NAT

Network address conversion can translate these IP addresses into public IP addresses, allowing them to access the internet. For example, we would use NAT to allocate a single public IP address to all of the systems that have private IP addresses, allowing them to connect to the Internet.



Inside Local IP Address	Inside Global IP address
10.0.0.1 10.0.0.2	171.69.58.80 171.69.58.81

We have a setup with one server with a private IP address of 10.0.0.1. We also have a router with NAT installed, and every router maintains a routing table with a public IP address that they use to map which private address to which public IP address, which is known as a global IP address. And that device would have a global IP address, making it visible on the internet. It will now map 10.0.0.1 to 172.69.58.80 with the aid of the router. It says that the device will not be recognized by the 10.0.0.1 server. That is a non-public IP address. 172.69.58.80 will include this information. The public IP address is that.

NAT Working

A border router has one interface in the local (inside) network and one interface in the global (outside) network and is designed for NAT. NAT transfers a local (private) IP address to a global (public) IP address as a packet travels outside the local (inside) network.

The global (public) IP address of a packet is translated to a local (private) IP address as it reaches the local network.

The packets will be lost and an Internet Control Message Protocol (ICMP) host unreachable packet to the destination will be transmitted if NAT runs out of addresses, i.e. no addresses are left in the pool setup.

What is the purpose of masking port numbers?

Assume two hosts A and B are linked in a network. Now, on the host side, they both request the same destination, on the same port number, say 1000, at the same time. If NAT just performs IP address conversion, when their packets arrive at the NAT, all of their IP addresses are disguised by the network's public IP address and sent to the destination. Destination will respond using the router's public IP address. As a result, when NAT receives a response, it would be unknown which reply belongs to which host (because source port numbers for both A and B are the same). As a result, NAT masks the source port number and creates an entry in the NAT table to prevent such a problem.

Network Address Translation (NAT) Types

There are 3 ways to configure NAT:

Static NAT – In this configuration, a single unregistered (Private) IP address is mapped to a legally registered (Public) IP address, resulting in a one-to-one mapping of local and global addresses. This is a popular term for web hosting. This is not used in businesses since many devices need Internet connectivity, and providing Internet access necessitates the use of a public IP address.

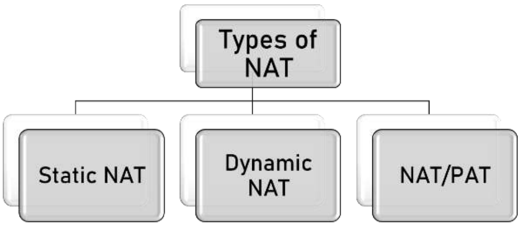


Figure 22 Types of NAT

As 3000 devices need Internet access, the organization would have to purchase 3000 public addresses, which would be very expensive.

Dynamic NAT: An unregistered IP address is converted into a registered (Public) IP address from a pool of public IP addresses in this form of NAT. The packet would be lost if the pool's IP address is not open, since only a limited number of private IP addresses can be converted to public addresses. If a pool of two public IP addresses is available, only two private IP addresses can be interpreted at any given time. Since several private IP addresses are mapped to a pool of public IP addresses, if a third private IP address tries to access the Internet, the packet will be lost. When the number of people who wish to use the Internet is set, NAT is used. This is also very expensive since the company would purchase a large number of global IP addresses to create a pool.

PAT: NAT overload is also known as Port Address Translation (PAT). Many local (private) IP addresses can be mapped to a single registered IP address using this method. The traffic is differentiated by port numbers, which identity which traffic belongs to which IP address. This is the most common method. After all, it is cost-effective because thousands of users can be connected to the Internet with only one single global (public) IP address.

Advantages of NAT

- NAT protects lawfully assigned IP addresses.
- It ensures anonymity by masking the device's IP address when sending and receiving traffic.
- When a network changes, address renumbering is no longer necessary.

Disadvantage of NAT

- Switching route delays occur as a result of translation.
- Although NAT is enabled, certain applications will not work.
- Tunneling protocols like IPsec get more complicated.
- Also, as a network layer system, the router should not tamper with port numbers (transport layer), but it does due to NAT.

ICMP (Internet control message protocol)

Which layer does ICMP Internet Control Message Protocol run in? It runs in the network layer, so which packet has the greatest chance of being discarded? What does that mean when we're sending a packet from source to destination and there's noise, or there's a congestion problem, or there's a buffer problem?

Any packets are discarded so that traffic can flow and congestion can be relieved.

Which packet has the greatest likelihood of being discarded?

The ICMP stands for Internet Managed Messaging Protocol. The first one is the ICMP then IGMP then UDP, and then TCP. Routers search for ICMP packet to discard so that they can make a space for the highest priority protocol

ICMP<IGMP<UDP<TCP

Which is the highest priority protocol?

TCP.

So, there are two situations in which we use the ICMP:

- error management and request and respond as shown in Figure 23.

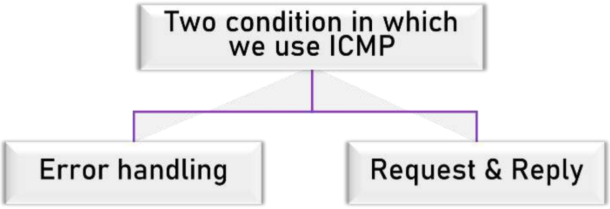


Figure 23 Two situations to use ICMP

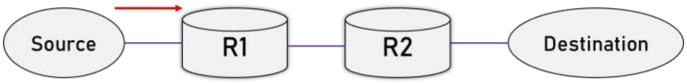


Figure 24 ICMP(a)

If a source sends a packet, and the source is going to transmit an IP packet, the source must send the IP packet to the destination, and it must send it via Routers R1 and R2 as shown in Figure 24. So, my packet arrived at R1 first, then R2. However, when my packet reaches R2, it is rejected. My packet was suddenly discarded due to congestion or a buffer problem; there was so much traffic, and the packet was discarded. As a result, the source's IP packet has been discarded. The router, R2, is responsible for informing the source that your packet has been discarded. As a result, the packet must be resent. So, in this case, the ICMP packet (Internet Control Message Protocol) is used to provide clarification that a packet has been lost. As a result, an ICMP packet containing the source and destination addresses will be sent to the source to notify them that their packet has been discarded as seen in Figure 25. How we deal with the mistake in this situation.

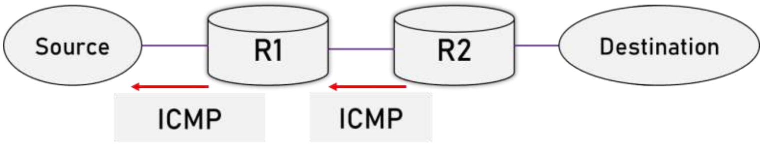


Figure 25 ICMP(b)

Request and response are the second.

Using ICMP packets, the sender will make a direct request to the destination. When studying ICMP, there are certain fundamental laws that you must adhere to. What are the basic rules? We will use the ICMP packet as input if the IP packet is discarded.

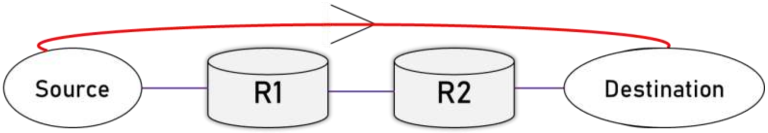


Figure 26 ICMP Request and Reply

Some Basic Rules

If IP packet is discarded

- We can use the ICMP packet as feedback.

If the ICMP packet is discarded

- No ICMP packet will be used as feedback.

However, let's use the ICMP packet as an example to see what the issue is. So, what sort of issue might it be? We've got a starting point and an endpoint. The IP packet will be sent to routers r1 and r2, and it is assumed that R2 will discard the packet due to a congestion problem. Now R2 is going to use the feedback, and which feedback packet it uses, it uses the ICMP packet as feedback. (R2, S) is written in brackets in this case. S is the destination to whom it wishes to warn that your packet has been lost, and R2 is the source one and second. When it arrived at r1, however, r1 still discarded the packet due to a congestion problem or buffer. R1 now has to notify R2 that the packet you sent for the source has been discarded. In this case, the source is R1 and the destination is R2. As a result of the ICMP comments, the packet will be sent to R2 once more. Still, owing to some congestion, R2 has a lot of congestion and a lot of other protocols, and R2's highest protocols are TCP and if ICMP packets are present, it will discard the lowest priority one, which is the ICMP packet.

So R2 will discard the ICMP packet; in this situation, R2 must remind r1 that whatever packet you sent me, whatever information it contained, has been lost. But, in this case, R2 sends the information to r1, where r2 is the source and r1 is the destination. It's sending to r1, so r1 can discard the packet once more. Since ICMP is the packet with the lowest priority. If there is congestion, they can dump the ICMP envelope, which has the lowest priority. Now, using an ICMP packet, r1 can send a feedback type or feedback request to R2. R2 will discard once more, and this will continue to happen over and over. So, endless loop, what do you think is going on here? As a result of the law violation, you will be trapped in an endless loop as shown in Figure 27.

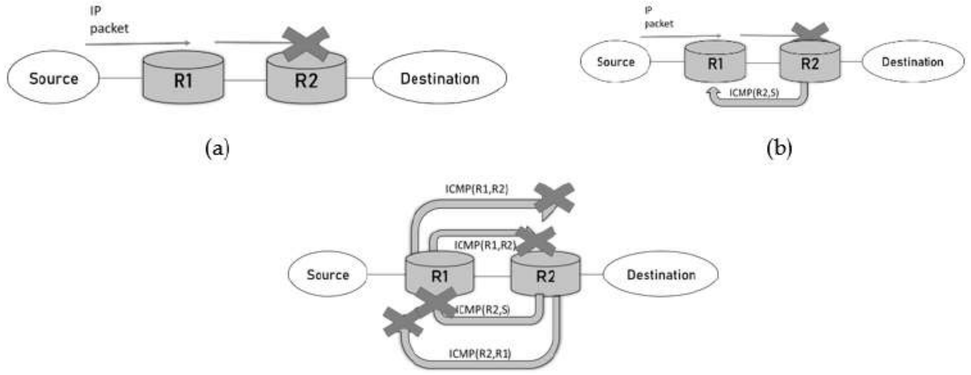


Figure 27 Issue in using ICMP packet

Important point

Always note that if you are sending an ICMP packet and it is misplaced, you cannot use an ICMP packet as feedback; you must use another packet. If an IP packet is refused, it is thrown out. The ICMP packet can be used as input. However, if an ICMP packet is destroyed, it cannot be used as feedback.

8.7 Address Resolution protocol

To send/receive messages, most computer programs/applications use a logical address (IP address), but the direct correspondence takes place over a physical address (MAC address), which is layer 2 of the OSI model. As a result, our goal is to obtain the destination MAC address, which is needed for communication with other devices. This is where ARP comes into play; it aims to convert IP addresses to physical addresses as seen in Figure 28.

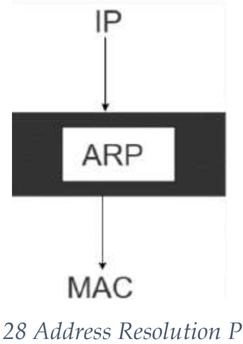


Figure 28 Address Resolution Protocol

ARP stands for Address Resolution Protocol and is one of the most relevant protocols in the OSI model's Network layer. From a host's established IP address, ARP determines the hardware address, also known as the Media Access Control (MAC) address.

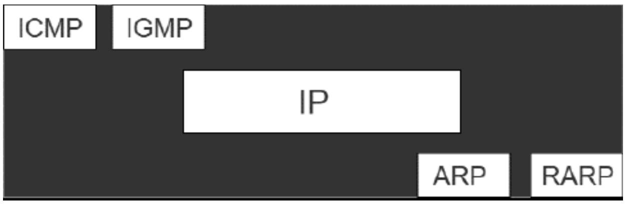


Figure 29 Network Layer

ARP Working

Consider a computer that wishes to connect with another device over the internet. What is the purpose of ARP? Is it broadcasting a packet to all of the originating network's devices? The network devices peel the data link layer header from the protocol data unit (PDU) called frame and send the packet to the network layer (OSI layer 3), where the packet's network ID is validated against the destination IP's network ID, and if they are identical, it responds to the source with the destination's MAC address. The above procedure is repeated until the second last network computer in the path to the destination is validated, at which point ARP responds with the destination MAC address.

8.8 Reverse Address Resolution Protocol

Reverse Address Resolution Protocol (RARP) is an acronym for Reverse Address Resolution Protocol, which is a computer networking protocol that allows a client computer to obtain its IP address from a gateway server's Address Resolution Protocol table or cache. In the gateway-router, the network administrator builds a table that maps the MAC address to the corresponding IP address. This protocol is used to transfer data between two server points. The client does not need to know who the server is capable of serving its request in advance. An administrator must configure each server's Medium Access Control (MAC) addresses individually. RARP is only capable of supporting IP addresses. When a replacement computer is set up, it can or may not have a connected disc that can hold the IP Address indefinitely, so the RARP client software must request the IP Address from the router's RARP server. Under the assumption that an entry has been created in the router table, the RARP server will return the IP address to the machine.

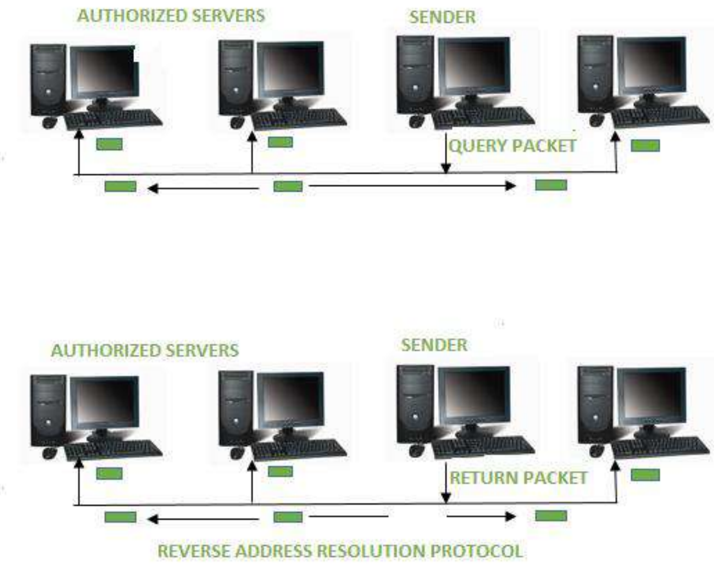


Figure 30 Reverse Address Resolution Protocol

RARP Working

The RARP is a Network Access Layer protocol that allows data to be sent between two points in a network.

Each network member has two distinct addresses: an IP (logical) address and a MAC (physical) address (the physical address). The IP address is allocated by program, and the MAC address is built into the hardware after that.

Any ordinary device on the network will act as a RARP server and respond to RARP requests. It must, however, store the data of all MAC addresses as well as their allocated IP addresses. If the network receives a RARP order, only these RARP servers can respond. The data packet must be transmitted over very low-cost network layers. This means that the package is sent to each of the participants at the same time. For an Ethernet broadcast address and its physical address, the client sends out a RARP order. The server then informs the client of its IP address.

Summary

- A global identification scheme that uniquely identifies every host and router is required at the network layer for packet transmission from host to host.
- An IPv4 address is 32 bits long and identifies a host or router on the Internet specially and universally.
- The netid is the part of the IP address that defines the network in classful addressing.
- The part of the IP address that specifies the host or router on the network is known as the host in classful addressing.
- The relation of a computer to a network is defined by its IP address.
- IPv4 addresses are divided into five categories. The number of hosts per network permitted in Classes A, B, and C varies. Multicasting is in Class D, and Class E is reserved.
- The first byte of an address can conveniently be used to evaluate its class.
- For unicast correspondence, addresses in classes A, B, or C are often used.
- Multicast correspondence is carried out using addresses in class D.
- Subnetting splits a wide network into many smaller ones, introducing an intermediate level of IP addressing hierarchy.
- We can partition the address space into variable-length blocks with classless addressing.
- IPv6 addresses are divided into three categories: unicast, anycast, and multicast.
- Reverse Address Resolution Protocol (RARP) is an acronym for Reverse Address Resolution Protocol, which is a computer networking protocol that allows a client computer to obtain its IP address from a gateway server's Address Resolution Protocol table or cache.

Keywords

***IP Address*:** An Internet Protocol address (IP address) is a numerical mark assigned to each interface (e.g., computer, printer) in a computer network that communicates using the Internet Protocol.

IP Protocol: The Internet Protocol (IP) is the principal communications protocol used for relaying datagrams (packets) across an internetwork using the Internet Protocol Suite.

NAT: Network address translation (NAT) is a way of converting one IP address space to another by changing network address information in packets' IP headers as they are in transit via a traffic routing system.

ARP: ARP stands for Address Resolution Protocol, and is one of the most relevant protocols in the OSI model's Network layer.

RARP: Where only the ethernet address is identified and the IP address is needed, RARP offers the opposite service to ARP.

PAT: Different devices on a LAN can be mapped to a single public IP address using Port Address Translation (PAT), which is an extension of Network Address Translation (NAT).

Classful Addressing: The address space is divided into five classes of Classful addressing: A, B, C, D, and E. Each of these classes has an IP address set that is valid.

Classless Addressing: Unlike classful addressing, classless IPv4 addressing does not partition the address space into classes.

Subnetting: Subnetting is the process of dividing a larger network into smaller networks to preserve security.

Review Questions

1. Explain the IP Protocol. What makes it different from the TCP protocol?
2. What are IP addresses, and what do they mean? Describe how an IP address is formatted.
3. Distinguish between IPV4 and IPV6 addressing, as well as their grouping.
4. Explain what subnetting is and how it works.
5. Difference between ARP and RARP.

Self Assessment

1. What is the format of the IP address?
   a) 34 bit
   b) 64 bit
   c) 16 bit
   d) 32 bit
2. Version 6 of the IP address has how many bits.
   a) 64 bits
   b) 128 bits
   c) 32 bits
   d) 256 bits
3. How many versions/s of IP's are there?
   a) 4 versions
   b) 3 versions
   c) 2 versions
   d) 1 version
4. Which technology allows a large number of private IP addresses to be represented by a smaller number of public IP addresses?
   a) NAT
   b) NTP
   c) RFC 1631
   d) RFC 1918
5. What is the effect of the overload keyword in a static NAT translation configuration?
   a) It enables port address translation.
   b) It enables the use of a secondary pool of IP addresses when the first pool is depleted.
   c) It enables the inside interface to receive traffic.
   d) It enables the outside interface to forward traffic.
6. What is the first step in the NAT configuration process?

a) Define inside and outside interfaces.

b) Define public and private IP addresses.

c) Define IP address pools.

d) Define global and local interfaces.

7. Which of the following is the ethernet broadcast address used in ARP and RARP requests?

   a) 255.255.255.255

   b) 08:00:20:11:aa:01

   c) ff:ff:ff:ff:ff:ff

   d) 224.0.0.0

8. Which of the following describes the function of ARP?

   a) It is used to map a 32-bit IP address to a 48-bit ethernet address.

   b) It is used to map a 48-bit ethernet address to a 32-bit IP address.

   c) It is used to map a 32-bit ethernet address to a 48-bit IP address.

   d) It is used to map a 48-bit IP address to a 32-bit ethernet address.

9. In classful addressing, the address space is divided into

   a. Five classes

   b. Ten classes

   c. Fifteen classes

   d. Three classes

10. In classless addressing, there are no classes but the addresses are still granted in

   a. Sections

   b. Blocks

   c. Codes

   d. All of the above.

11. In IPv4 Addresses, classful addressing is replaced with

   a) Classless Addressing

   b) Classful Addressing new version

   c) Classful Advertising

   d) Classless Advertising

12.The first address in a block is used as the network address that represents the

   a) Class Network

   b) Entity

   c) Organization

   d) DataCodes

Answers

1.	a	2.	b	3.	c	4.	a
5.	a	6.	a	7	c	8.	a
9.	a	10.	b	11.	a	12.	c

Further Readings

Achyut S Godbole and Atul Kahate, *Web Technologies,* Tata McGraw Hill.

Andrew S. Tanenbaum, *Computer Networks,* Prentice Hall.

Behrouz A. Forouzan, Sophia Chung Fegan, *Data Communications and Networking,*

McGraw-Hill Companies.

Douglas Comer, Computer Networks and Internets with Internet Applications, 4th

Edition, Prentice-Hall.

Ferguson P., Huston G., John Wiley & Sons, Inc., 1998. *Quality of Service: Delivering*

QoS on the Internet and in Corporate Networks.

J. D. Spragins, Telecommunications Protocols and Design, Addison Wesley.

McDysan, David E. and Darren L. Spohn, *ATM Theory and Applications*, McGraw-

Hill Osborne Media, 1998.

Nassar, Daniel J., *Ethernet and Token Ring Optimization*, iUniverse.com, 2000.

Spurgeon, Charles E. Ethernet, *The Definitive Guide.* O'Reilly & Associates, 2000.

William A Shay, Understanding Communication and Networks, 3rd Edition, Thomson Press.

Unit 08: Network layer – routing

CONTENTS

Objectives

- what is Unicast Routing.
- three major protocols for unicast routing.
- classification of Routing Algorithms.

Introduction

At the network layer, the Internet Protocol (IP) serves as the primary protocol. IP is a best-effort distribution protocol, but it lacks certain functionality including flow control and error control. It's a logical addressing-based host-to-host protocol. Other protocols are needed to make IP more open to certain specifications of today's internetworking. To construct a mapping between physical and logical addresses, protocols are used. Logical addresses are used in IP packets. However, these packets must be encapsulated in a frame, which requires physical addresses (node-to-node). We'll see the ARP, or Address Resolution Protocol, is designed specifically for this purpose. We sometimes need to reverse map a physical address to a logical address. Since the Internet Protocol lacks flow and error management, another protocol, ICMP, was created to provide warnings. It records network or destination host congestion and other forms of errors.

9.1 Unicast

Unicast routing refers to propagation from a single source to a single receiver, often known as point-to-point correspondence between sender and receiver. TCP and HTTP are two examples of unicast protocols. TCP stands for transmission control protocol, which is the most widely used unicast protocol. It is a connection-oriented protocol that relies on the recipient site's recognition. The hypertext transmission protocol (HTTP) stands for hypertext transfer protocol. It is a collaboration protocol that is based on object-oriented principles. Intradomain routing protocol, intranetwork routing protocol, and intranet routing protocol are two types of routing protocols as shown in Figure 1. Intradomain routing protocols are used to route packets within a given area, such as within an institutional network for e-mail or Web browsing. An interdomain routing protocol, on the other hand, is a method for routing packets through domain networks.

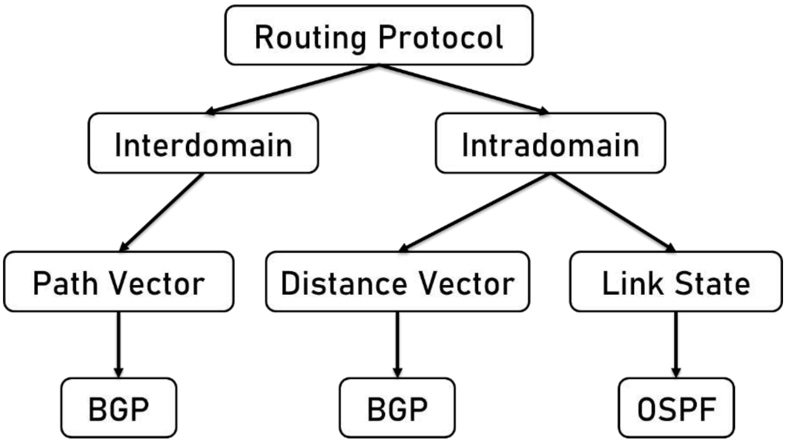


Figure 1 Types of Routing Protocol

For unicast routing, we have three main protocols: distance vector routing, connection state routing, and route vector routing as shown in Figure 2.

9.2 Distance Vector Routing

In distance vector routing, the least cost for any two nodes is the path with the shortest distance, and each node keeps a table of the shortest distances to each other. The term "distance vector" refers to how routes are marketed using two characteristics.

Distance: This metric determines how long it is to the destination network and is dependent on hop count, cost, bandwidth, distance, and other factors.

Vector: The vector specifies the next-hop router's or exit interface's path in order to reach the destination.

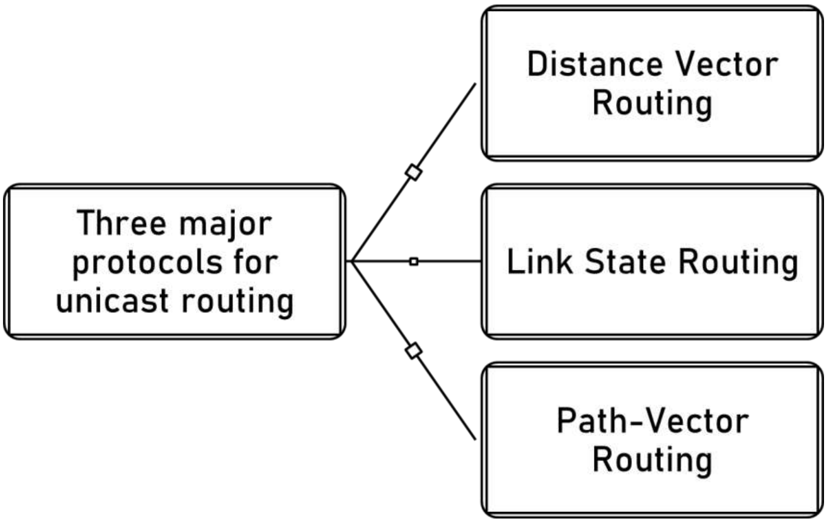


Figure 2 Three major protocols for unicast routing

Example: We present a system of five nodes and the tables that go with them in Figure 3.

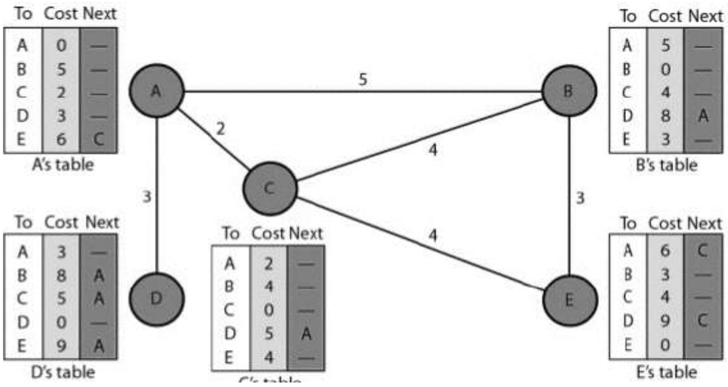


Figure 3 Distance Vector Routing Tables

The table for node A illustrates how to get from this node to every other node.

Our lowest cost to enter node E, for example, is 6. C is a stop along the way.

Intilaization

Figure 3 shows that the tables are stable; each node knows how to access every other node and how much it will cost. However, this is not the case in the beginning. Only the distance between each node and its immediate neighbours, all that are closely bound to it, can be known. So, for the time being, we'll presume that each node will send a message to its immediate neighbours and calculate the distance between them.

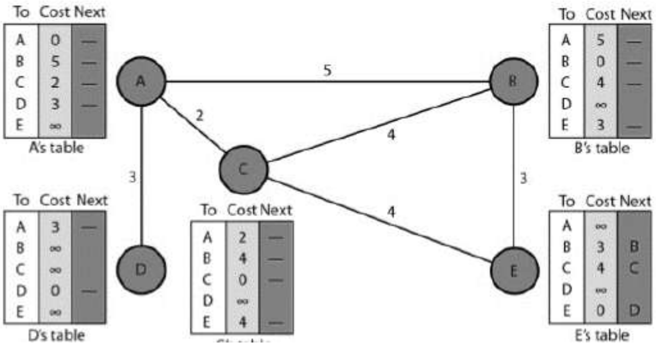


Figure 4 Initialization of Tables in Distance Vector Routing Protocols

The original tables for each node are seen in Figure 4. Any entry that isn't a neighbour is given an infinite distance (unreachable).

Sharing

The exchange of information between neighbors is at the heart of distance vector routing. While node A is unaware of node E, node C is. As a result, if node C shares its routing table with node A, node A will still figure out how to get to node E. Node C, on the other hand, has no idea how to get to node D, but node A does. If node A and node C share a routing table, node C will also know how to access node D. In other words, as immediate neighbors, nodes A and C will boost their routing tables by assisting one another. There's just one problem. How much of the table does each neighbour share? The table of a neighbour is unknown to a node. The best answer for each node is to give its entire table to its neighbour and let him determine the parts to keep and which to delete.

The neighbour, on the other hand, does not benefit from the third column of a table (next stop). This column must be replaced with the sender's name when the neighbour collects a table. The sender of the table is the next node if all of the rows can be included. As a result, a node can only give the first two columns of its table to all of its neighbours. To put it another way, sharing here just refers to the first two columns.

Updating

A node must change its routing table when it receives a two-column table from a neighbour. There are three approaches to updating:

1. For each value in the second column, the receiving node must apply the cost between itself and the transmitting node.
2. If the receiving node uses information from another row, it must apply the name of the transmitting node as the third column of each row. The next node in the path is the sending node.
3. Each row of the receiving node's old table must be compared to the corresponding row of the updated version of the received table.
    a. Where the next-node entry differs, the receiving node selects the row with the lowest rate. The old tie is maintained if there is one.
    b. The receiving node selects the new row if the next-node entry is the same. Assume node C previously advertised a distance-based path to node X. Assume there is no longer a path between C and X; node C now advertises this route with an infinite radius. About the fact that its old entry is smaller, Node A must not disregard this value. The old road is no longer in use. The current path is infinite in duration.

Node A changes its routing table after obtaining the partial table from node C, as seen in Figure 5. When Do You Share?

When does a node transmit its partial routing table (which only has two columns) to all of its immediate neighbours? The table is submitted both on a regular basis and when the table changes.

Periodic Update: Every 30 seconds, a node sends its routing table in a periodic update. The length of time is determined by the protocol that uses distance vector routing.

Triggered Update: When a node's two-column routing table changes, it sends a triggered update to its neighbours. This is referred to as an activated upgrade.

The following factors can influence the outcome.

1. A node receives a table from a peer and updates it, resulting in modifications in its own table.

2. A node senses a defect in one of the adjacent connections, resulting in a drop in distance to infinity.

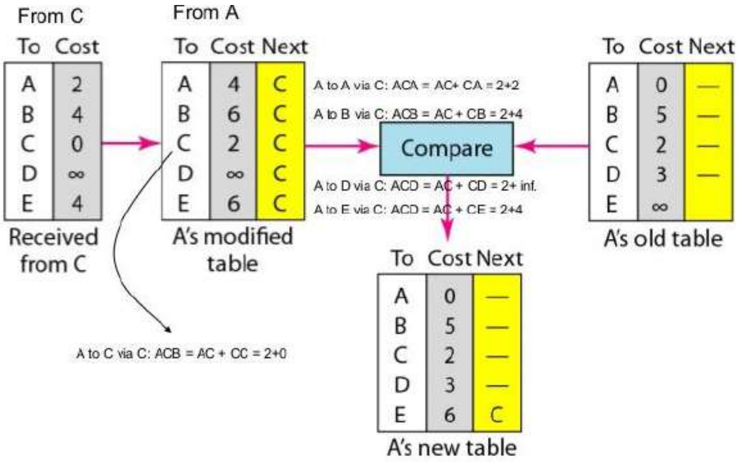


Figure 5 Updating in Distance Vector Routing Table

9.3 Link State Routing

The second family of routing protocols is connection state routing. Link-state routing uses link-state routers to exchange messages that allow each router to learn the entire network topology, while distance vector routers use a distributed algorithm to compute their routing tables. Each router will then compute its routing table using a shortest path calculation based on the learned topology.

Features of link state routing protocols

Link State Packet: A small packet that includes routing information is known as a connection state packet.

Link State Database – A database containing data collected from link state packets.

Shortest path first algorithm (Dijkstra algorithm) – A database estimation produces the shortest path.

Routing Table: A list of known routes and interfaces is called a routing table.

Calculation of shortest path

To find shortest path, each node need to run the famous **Dijkstra algorithm**. This famous algorithm uses the following steps:

Step-1: The node is selected as the tree's root node, resulting in a tree with just one node. Set the cumulative cost of each node to a value depending on the information in the Connection State Database.

Step-2: The node now chooses the node that is closest to the root of all the nodes not in the tree-like structure and connects it to the tree.

The tree's outline is transformed.

Step 3: After this node has been connected to the stack, all non-tree nodes' costs must be modified so the paths could have shifted.

Step-4: The node repeats Steps 2 and 3 until all of the nodes in the tree have been inserted.

In contrast to Distance Vector protocols, Link State protocols have:

1. It necessitates a significant amount of memory.
2. Shortest route computations necessitate a large number of CPU circles.
3. If a network uses a small amount of bandwidth, it responds easily to topology changes.
4. To shape connection state packets, all objects in the database must be sent to neighbours.
5. In the topology, all neighbours must be trusted.
6. To prevent unwanted adjacency and complications, authentication methods may be used.
7. In link state routing, split horizon strategies are not feasible.

Open shortest path first (OSPF) routing protocol –

1. OSPF (Open Shortest Path First) is a unicast routing protocol developed by an Internet Engineering Task Force working group (IETF).
2. It's a protocol for routing within a domain.
3. It's a free and transparent protocol.
4. It's akin to the Routing Information Protocol (RIP) (RIP).
5. OSPF is a classless routing protocol, meaning it uses the subnet of each path it recognises in its notifications, allowing for variable-length subnet masks. An IP network can be divided into several subnets of different sizes using variable-length subnet masks. This gives network operators more consistency in network setup.
6. OSPF is implemented as a network layer application that utilises the facilities offered by the Internet Protocol IP datagram that transports OSPF messages. OSPF sets the protocol field value to 89.
7. The OSPF algorithm, also known as the Dijkstra algorithm, is based on the SPF algorithm.
8. Version 1 and version 2 of OSPF are available. Version 2 is the most widely used.

OSPF Messages – The OSPF protocol is a complicated one. It employs five distinct message forms. The following are some of them:

1. Hello messages (Form 1)
2. Message from the database (Type 2)
3. Message with a link-state request (Type 3)
4. Message about a link-state update (Type 4)

5. Message of Link-State Recognition (Type 5)

9.4 Path Vector routing

The cost of reaching a given destination is not used by a route vector protocol to decide whether or not each path open is loop free. Instead, route vector protocols analyse the path to the destination to determine if it is loop-free.

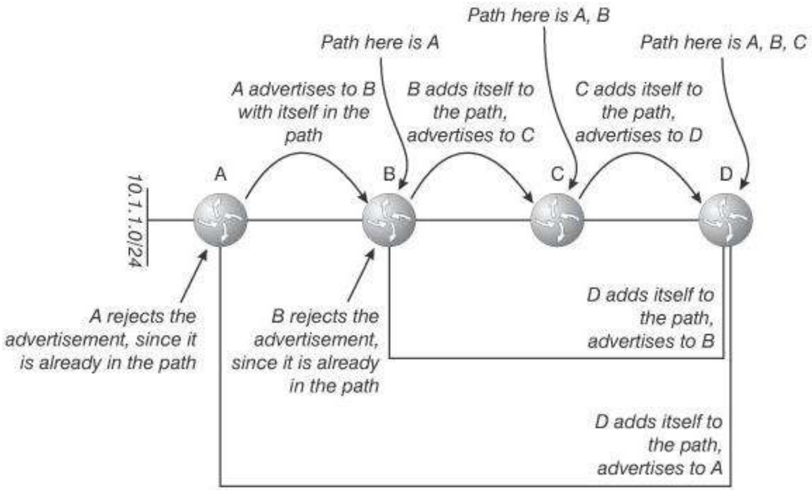

Path here is A, B
Path here is A
Path here is A, B, C
A advertises to B with itself in the path
B adds itself to the path, advertises to C
C adds itself to the path, advertises to D
A
B
C
D
10.1.1.0/24
A rejects the advertisement, since it is already in the path
B rejects the advertisement, since it is already in the path
D adds itself to the path, advertises to B
D adds itself to the path, advertises to A

Figure 6 Path Vector Routing

By logging each hop the routing advertising traverses through the network, a route vector protocol ensures loop-free routes through the network. In this case, router A informs router B that the 10.1.1.0/24 network is reachable. When router B receives this data, it joins the path and broadcasts it to router C. Router C joins the road and informs router D that the 10.1.1.0/24 network can be reached in this direction. As Router D receives the route advertising, it joins the road as well. When router D tries to advertise to router A that it will hit 10.1.1.0/24, router A rejects the advertising because the corresponding direction vector in the advertisement suggests that router A is already in the path. When a router encounters a message in which it is already a part of the path, it rejects it because accepting the request will result in a routing information loop.

9.5 Border Gateway Protocol(BGP)

Rather than considering a single router as a single point in the path to any given destination, BGP applies the path vector principle on a broader scale. Each autonomous system is treated as a single point on the road to any given destination by BGP as shown in Figure 7.

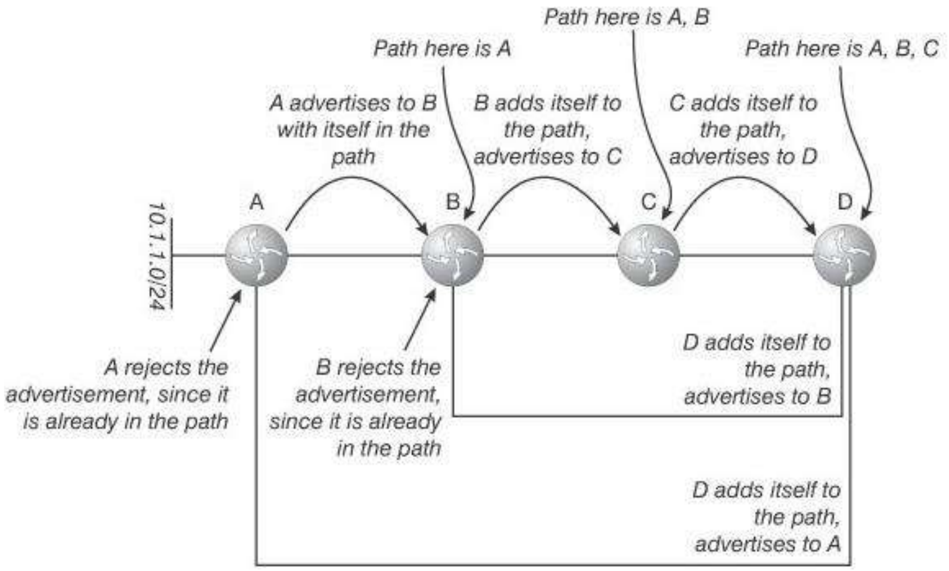

Path here is A, B
Path here is A
Path here is A, B, C
A advertises to B with itself in the path
B adds itself to the path, advertises to C
C adds itself to the path, advertises to D
10.1.1.0/24
A
B
C
D
A rejects the advertisement, since it is already in the path
B rejects the advertisement, since it is already in the path
D adds itself to the path, advertises to B
D adds itself to the path, advertises to A

Figure 7 Path vector over a set of autonomous systems

This case is similar to Figure 6, except that instead of a single router, each autonomous device is treated as a point along the road. The network 10.1.1.0/24, also known as a prefix, is marketed along with a list of autonomous systems over which the upgrade has passed; this list of autonomous systems is referred to as the AS Route. AS 65100 creates the prefix 10.1.1.0/24 and advertises it to AS 65200, connecting itself to the AS Route. AS 65200 is added to the AS Route, and the prefix is advertised to AS 65300.

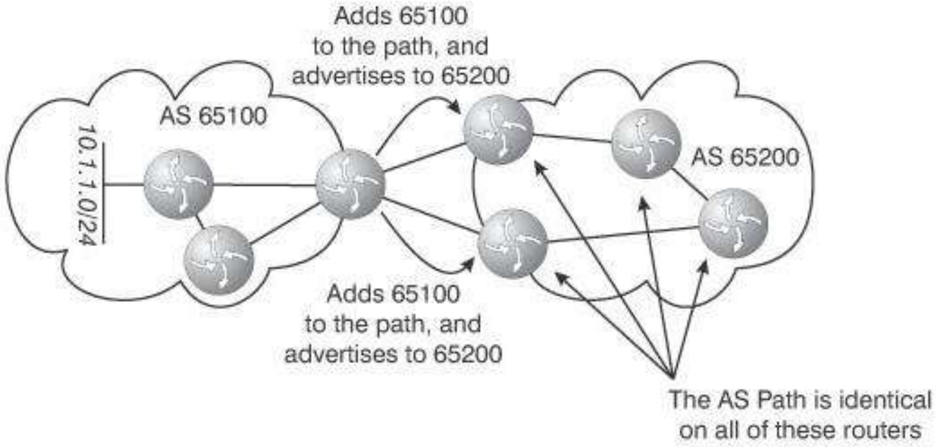

Adds 65100 to the path, and advertises to 65200
AS 65100
10.1.1.0/24
AS 65200
Adds 65100 to the path, and advertises to 65200
The AS Path is identical on all of these routers

Figure 8 BGP routing with an AS

When AS 65300 advertises the prefix 10.1.1.0/24 to AS 65100, it is refused because the 65100 sees that its local AS is already in the AS Path, and accepting the route will result in a routing information loop as shown in Figure 8. The main reason BGP considers an entire autonomous device as a single hop in the AS Path is to mask the AS's topological information. AS 65200, for example, has no way of knowing what the route through AS 65100 looks like; all it knows is that the destination is reachable through AS 65100. One interesting side effect of considering each autonomous system as a single entity for which the autonomous system route vector is connected is that BGP can only detect loops between autonomous systems without additional knowledge or laws, and it cannot guarantee loop-free paths within an AS as shown in Figure 8.

9.6 Types of Casting in a computer network

The cast term here signifies some data(stream of packets) is being transmitted to the recipient(s) from client(s) side over the communication channel that helps them to communicate. Let's see some of the "cast" concepts that are prevailing in the computer networks field.

1. Unicast

When only one sender and one receiver are involved, this method of data transfer is advantageous. So, in a nutshell, it is a one-to-one transmission. When a device with an IP address of 10.1.2.0 in one network needs to transfer traffic (data packets) to a device with an IP address of 20.12.4.2 in another network, unicast is used. Unicast example is shown in Figure 9.

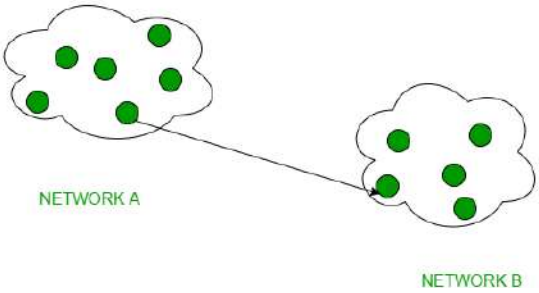


Figure 9 Unicast Example

Broadcast

There are two methods of broadcasting conversion (one-to-all) techniques:

Limited Broadcasting

If you need to transmit a stream of packets to all of the computers on your network, broadcasting w ill come in handy as shown in Figure 10. To do this, it will append 255.255.255 (all 32 bits of the IP address set to 1) to the destination address of the datagram (packet) header, which is reserved for information transmission to all recipients from a single client (sender) across the network.

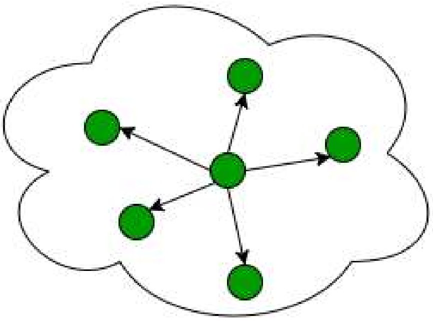


Figure 10 Network Cluster

Direct Broadcasting

If a node on one network needs to send a packet stream to all of the devices on the other network, this is helpful. This is accomplished by converting all of the destination address's Host ID portion bits to 1, which is referred to as Direct Broadcast Address in the datagram header for information transfer.

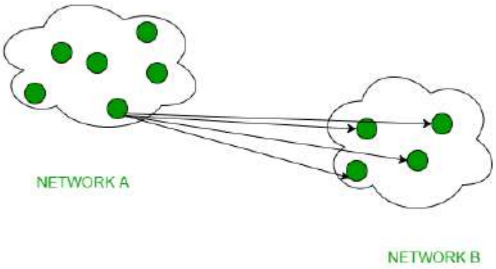


Figure 11 Direct Broadcasting

For video and audio delivery, television networks primarily use this mode. In computer networks, one of the most important protocols in this class is Address Resolution Protocol (ARP), which is used to resolve IP addresses into physical addresses, which is needed for underlying communication.

Multicast

In multicasting, data transfer traffic is shared between one or more senders and one or more receivers. The traffic in this system is split into unicast (one-to-one) and broadcast (multicast) (one-to-all). Multicast allows servers to send single copies of data streams to hosts that require them.

For IP multicast to operate, it needs the assistance of other protocols such as IGMP (Internet Group Management Protocol) and Multicast routing. Class D is also reserved for multicast classes in Classful IP addressing.

9.7 Routing in Adhoc Network

Challenges in Adhoc network

Dynamic topology, unreliable connections, restricted resources, lower connection bandwidth security, and no default router are all problems in an ad hoc network. There could be one other major problem in routing: there are no physical connections.

What does physical connection mean?

It means that when nodes travel from one location to another, wireless connections are formed and broken. As a result of the regular disconnections and partitions, no physical relations exist. As a result, routing protocols are classified as proactive, reactive, or hybrid as seen in Figure 12.

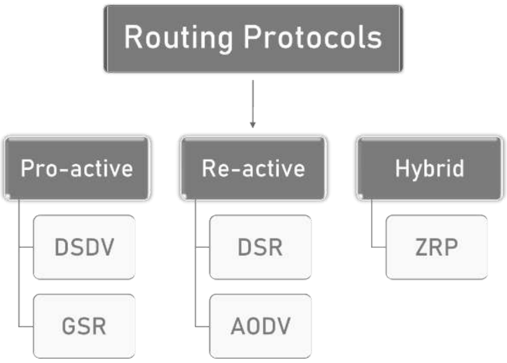


Figure 12 Routing Protocols

Proactive Routing protocols

First and foremost, in a constructive routing, proactive. Until packets are sent, the routing tables are established. The following is an example of a connection state distance factor. Each node in the network is aware of the routes to all other nodes. Each mobile node has its own routing table, which stores information about routes to all potential destination mobile nodes. So, in the case of constructive, because the topology in the mobile ad hoc network is complex, as I previously said, these routing tables are changed regularly, either every 30 seconds or 35 seconds, depending on how quickly the network topology varies. As a result, constructive routing protocol has a drawback. As a result, it does not operate well with large networks so the routing table entries become too large as they must retain and route information to all available nodes.

1. ***DSDV(Destination Sequenced Distance Vector Routing Protocol)***

It is a proactive table driven routing protocol that extends the distance vector routing protocol and is built on the Bellman Ford routing algorithm. DSDV is a proactive table driven routing protocol that falls under the proactive. Due to the count to infinity issue, the distance vector routing protocol was not suitable for a mobile ad hoc network. So, we have a solution for this in the form of the destination sequence distance vector routing protocol. Any routing entry in the routing table that is managed by each node now has a destination sequence number associated with it. Only if the entry consists of a new modified path to the destination with a higher sequence number will a node contain the new upgrade in the table. As a result, it is proactive which is why it is referred to as a table guided routing protocol.

2. *GSR(Global State Routing)*

The next protocol is the global state routing protocol, which extends the wired network's connection state routing and is built on Dijkstra's routing algorithm. Since each node floods the link state routing information directly through the whole network, the link state routing protocol is not suitable for mobile ad hoc networks. As a result, global flooding can cause a network's control packets to become congested. In the last slides of the previous lecture, I explained distance relation state route vector routing. So, if you're not sure what's causing the rain, you should attend the session. What is connection state routing and how does it work? Since you can get all of the details from the floods, what is distance vector routing? As a result, the global state routing protocol was introduced as a workaround. Really, link state routing packets are not flooded into the network by global state routing.

Reactive

This are also known as on-demand routing protocols in the case of reactive routing. This path is only found when it is needed or required. Path exploration is accomplished by flooding route request packets across the mobile network. It is divided into two phases: path exploration and route management. The path is only discovered when it is required.

1. *DSR(Dynamic Source Routing protocol)*

The DSR dynamic source routing protocol is the next step. It's a reactive or on-demand routing protocol, which means the path is only found when it's required. Path exploration is accomplished by flooding route request packets across the mobile network. There are two stages of it: path exploration and route management.

Route Discovery

This process decides the most efficient route for data packet transfer between the source and destination mobile nodes.

Route Maintenance

Since the topology of a mobile ad-hoc network is complex, this step performs route maintenance. As a result, there are several instances of connection breakage resulting in network collapse between mobile nodes.

2. AODV (*Ad-Hoc On Demand Vector Routing protocol)*

It's an on-demand/reactive routing protocol. It is a dynamic source routing protocol (DSR) extension that aids in the elimination of the protocol's disadvantage. When the source mobile node sends the data packet to the destination mobile node after route discovery in DSR, it also includes the full path in its header. As a result, as the network grows in complexity, the length of the full route grows, as does the size of the data packet header, slowing down the entire network. As a result, the Ad-Hoc On Demand Vector Routing protocol was developed as a solution. The key distinction is about how the route is stored: AODV stores it in the routing table, while DSR stores it in the data packet's header. It works in the same way, with two phases: route exploration and route maintenance.

Hybrid routing Protocol

It essentially incorporates the benefits of both reactive and proactive routing protocols in one package. These protocols are flexible in design, adapting to the source and destination mobile nodes' zone and location. One of the most popular hybrid routing protocol is Zone Routing Protocol (ZRP).

Summary

- Whether the deliverer (host or router) and the destination are on the same network, the packet delivery is called direct; if the deliverer (host or router) and the destination are on separate networks, the packet delivery is called indirect.
- Instead of a full list of the stops the packet must make, the next-hop approach only lists the location of the next hop in the routing table; in the network specific method, all hosts on a network share one routing table entry.
- The full IP address of a host is given in the routing table in the host-specific method.

- To avoid massive routing tables, classless addressing necessitates hierarchical and regional routing.
- The entries in a static routing table are modified manually by an administrator, while the entries in a dynamic routing table are updated automatically by a routing protocol.
- RIP is based on distance vector routing, in which each router shares its knowledge of the entire AS with its neighbours at regular intervals.
- Dijkstra's algorithm is used to determine OSPF routing tables.
- BGP is a routing protocol for interautonomous systems that is used to update routing tables.
- Only packets that have travelled the shortest distance from the source to the router are forwarded in reverse path forwarding (RPF).
- Link State Routing tries to find its neighbours and learn their network addresses so that the router can choose the shortest path. Many classes are used to relay packets in hierarchical routing. Broadcast and multicast routings are used to send a single packet to multiple recipients based on whether they are members of a broadcast or multicast network.
- The shortest path to - destination in the network is identified by traversing the tree, and the Dijkstra algorithm is the most commonly used shortest path first algorithm.
- Based on the advertised information about the direction and distance for each destination, which are stored in a local database, the distance vector algorithms are used to decide which route is the safest path to each destination.

Keywords

Adaptive Algorithms: They can change their routing decisions in response to changes in the topology and traffic, and they can automatically adjust routing details as the network configuration changes.

Distance Vector Routing: It maintains a routing table and exchanges its routing table with each of its neighbors so that their routing tables get updated.

Flow-based Routing: It considers both the topology and the load.

Hierarchical routing: Intra-domain and inter-domain routing are used in hierarchical routing.

Link State Routing allows each router in the network to learn the network topology and build a routing table based on it.

Multicast is a term that refers to one or more network interfaces that are spread across several subnets. One-to-many connectivity is possible.

Multicast Routing: This refers to transmitting data to well-defined groups with a large number of participants but a limited number in comparison to the whole network.

Review Questions

1. Describe briefly how hierarchal algorithm works.
2. What is the main purpose of using router in a network?
3. Differentiate between:
    a. Connectionless and connection-oriented service
    b. Interior and Exterior Routing
    c. Link state and distance vector routing
4. Difference between proactive and reactive routing protocols.
5. What is the purpose of BGP?
6. Why do OSPF messages propagate faster than RIP messages?

Self-Assessment

1. Multiple access schemes are used to allow ________ mobile users to share simultaneously a finite amount of radio spectrum.
   a. Many
   b. One
   c. Two
   d. Ten-Fifteen
2. Multiple access protocols is divided into
   a. One
   b. Two
   c. Three
   d. Four
3. The use of hierarchy in routing tables can ____________ the size of routing tables.
   a. Reduce
   b. Increase
   c. Both a and b
   d. None of the above.
4. If there is only one routing sequence for each source destination pair, the scheme is known as …..
   a. static routing
   b. fixed alternative routing
   c. standard routing
   d. dynamic routing
5. Count-to-Infinity problem occurs in ....
   a. distance vector routing
   b. short path first
   c. link state routing
   d. hierarchical routing
6. Link state packets are built in ....
   a. short path first
   b. distance vector routing
   c. link state routing
   d. hierarchical routing

7. In which routing method do all the routers have a common database?
   a. Distance Vector
   b. Link Vector
   c. Shortest path
   d. Link State
8. In distance vector routing algorithm, the routing tables are updated ....
   a. by exchanging information with the neighbours
   b. automatically
   c. using the backup database
   d. by the server
9. Distance vector routing algorithm is implemented in Internet as ....

a. OSPF

b. RIP

c. ARP

d. APR

10. Which of the following routing algorithm takes into account the current network load.

a. broadcast

b. shortest path

c. flooding

d. distance vector routing

11. In distance vector routing the delay metric is ....

a. number of hops

b. geographical distance

c. number of neighbours

d. queue length

12. A well -defined groups that are numerically large in size but small compared to the network as a whole are used in ....

a. Unicast routing

b. Multicast routing

c. Broadcast routing

d. Telecast routing

Answers

1.	a	2.	c	3.	a	4.	b
5.	a	6.	b	7.	d	8.	a
9.	b	10.	d	11.	d	12.	b

Further Readings

Achyut S Godbole and Atul Kahate, *Web Technologies,* Tata McGraw Hill.

Andrew S. Tanenbaum, *Computer Networks,* Prentice Hall.

Behrouz A. Forouzan, Sophia Chung Fegan, *Data Communications and Networking,*

McGraw-Hill Companies.

Douglas Comer, *Computer Networks and Internets with Internet Applications,* 4th Edition, Prentice Hall.

Ferguson P., Huston G., John Wiley & Sons, Inc., 1998. *Quality of Service: Delivering QoS on the Internet and in Corporate Networks*.

J. D. Spragins, *Telecommunications Protocols and Design*, Addison Wesley.

McDysan, David E. and Darren L. Spohn, *ATM Theory and Applications*, McGraw-Hill Osborne Media, 1998.

Nassar, Daniel J., *Ethernet and Token Ring Optimization*, iUniverse.com, 2000.

Spurgeon, Charles E. Ethernet, *The Definitive Guide.* O'Reilly & Associates, 2000.

William A Shay, *Understanding Communication and Networks,* 3rd Edition, Thomson Press.

Unit 10: Transport layer-protocols

CONTENTS

Objectives

- understanding services of transport layer
- difference between connection-oriented and connection-less service.
- Reliable vs Unreliable Delivery
- TCP 3-Way Handshake Process

Introduction

The fourth layer from the top is the transport layer. The transport layer's primary function is to provide direct connectivity services to application processes operating on various hosts. The transport layer allows application processes operating on separate hosts to communicate logically. About the fact that application processes on separate hosts are not physically connected, application processes use the transport layer's logical connectivity to transmit messages to one another. End systems are equipped with transport layer protocols, but network routers are not. The network programmes on a computer network may use more than one protocol. TCP and UDP, for example, are two transport layer protocols that supply the network layer with separate networks. Multiplexing/demultiplexing is supported by all transport layer protocols. It also offers other features such as secure data transmission, guaranteed bandwidth, and guaranteed latency. Each programme in the application layer is capable of sending a message over TCP or UDP as seen in Figure 1. These two protocols are used by the programme to communicate. In the internet layer, both TCP and UDP can connect with the internet protocol. The programmes have access to the transport layer and can read and write to it. As a result, contact can be described as a two-way mechanism.

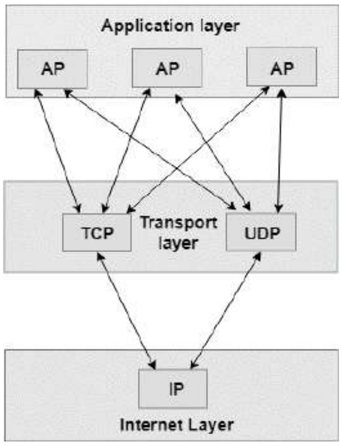


Figure 1 Transport Layer

10.1 Services provided by the Transport Layer

The transport layer provides facilities that are identical to those offered by the data link layer. The data link layer offers services within a single network, while the transport layer provides services through several networks in an internetwork. The physical layer is controlled by the data link layer, while the lower layers are controlled by the transport layer.

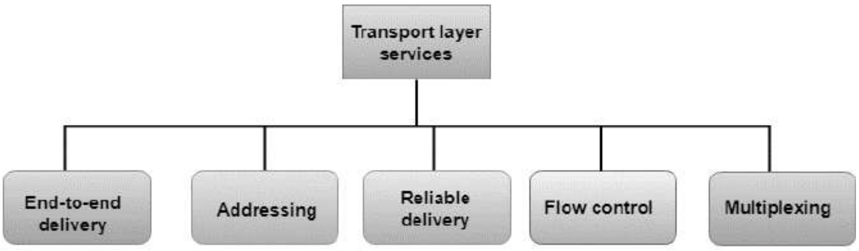


Figure 2 Services Provided by Transport Layer

The five types of services offered by transport layer protocols are as follows:

- **End-to-end delivery**
- **Addressing**
- **Reliable delivery**
- **Flow control**
- **Multiplexing**

End to End delivery

The transport layer sends the entire message to its intended recipient. As a result, it means that a whole message is sent from source to destination.

Reliable delivery

Through retransmitting missing and corrupted packets, the transport layer delivers redundancy services. The reliable delivery has four aspects:

- ***Error control***
- ***Sequence control***
- ***Loss control***
- ***Duplication control***

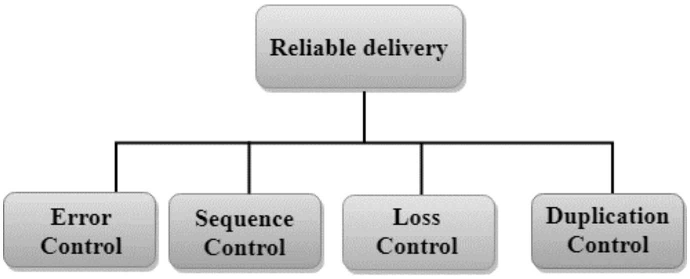


Figure 3 Reliable Delivery

Error Control

- The primary role of reliability is **Error Control**. In reality, no transmission will be 100 percent error-free delivery. Therefore, transport layer protocols are designed to provide error-free transmission.
- The data link layer also has an error management feature, but it only guarantees error-free transmission from node to node. End-to-end reliability is not guaranteed by node-to-node reliability.
- The data link layer examines each network for errors. If an error occurs within one of the routers, the data link layer would not be able to detect it. It only detects errors that occurred between the beginning and the end of the connection. As a result, the transport layer checks for errors from beginning to end to guarantee that the packet arrives intact.

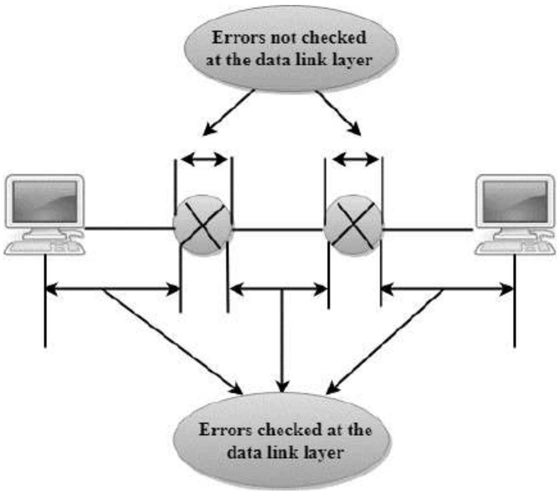

Errors not checked at the data link layer
Errors checked at the data link layer

Figure 4 Error at Data Link Layer

Sequence Control

Sequence monitoring, which is applied at the transport layer, is the second component of reliability.

The transport layer is in charge of ensuring that packets obtained from the upper layers can be used by the lower layers on the receiving end. It means that the different segments of a transmission can be properly reassembled on the receiving end.

Loss Control

The third factor of durability is loss control. The transport layer means that all of a transmission's fragments, not just any of them, arrive at their destination. A transport layer assigns sequence numbers to all transmission fragments on the sending end. The receiver's transport layer will use these sequence numbers to locate the missing portion.

Duplicate Control

The fourth factor of durability is duplication control. The transport layer ensures that no redundant data reaches its intended location. Sequence numbers are used to locate missing packets, as well as to identify and discard duplicate fragments by the recipient.

Flow Control

Flow management is used to keep the sender from sending too many data to the recipient. When a receiver is overwhelmed with files, it discards packets and requests that they be retransmitted. As a result, network interference increases, lowering system performance. Flow regulation is handled by the transport layer.

Multiplexing

Multiplexing is used by the transport layer to increase transmission reliability.

- *Multiplexing can occur in two ways:*
  - ***Upward multiplexing*:** Upward multiplexing means multiple transport layer connections use the same network connection. To make more cost-effective, the transport layer sends several transmissions bound for the same destination along the same path; this is achieved through upward multiplexing as seen in .

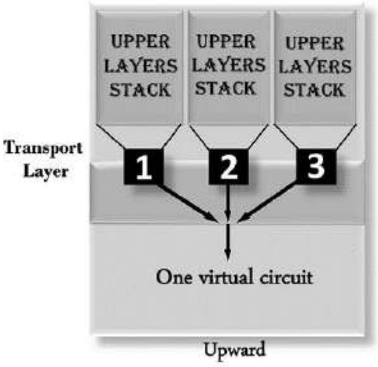


Figure 5 Upward Multiplexing

Downward multiplexing refers to the use of multiple network links through one transport layer link. Downward multiplexing requires the transport layer to divide a link into multiple directions in order to increase throughput. When networks have a low or sluggish bandwidth, this form of multiplexing is used as seen in Figure 6.

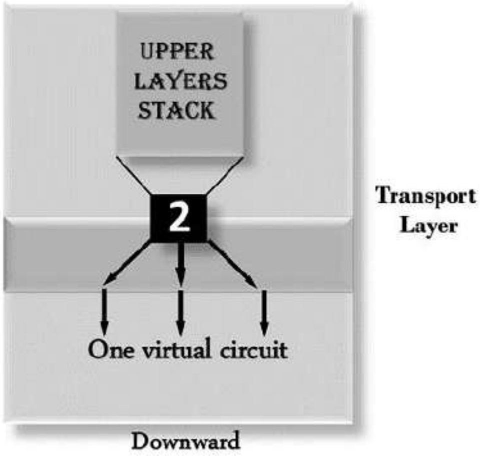


Figure 6 Downward Multiplexing

10.2 Difference between Connection-Oriented and Connectionless Service

To build communications between two or more computers, Connection-oriented and Connectionless Services are used. The establishment and termination of a link for sending data between two or more devices is part of a connection-oriented service. Connectionless operation, on the other hand, does not necessitate the establishment of any connection or termination mechanism in order to transmit data over a network.

Connection-oriented service

A network infrastructure that was planned and built since the telephone system is known as a connection-oriented service. Until sending data over the same or separate networks, a link-oriented service is used to establish an end-to-end connection between the sender and the receiver. Packets are delivered to the recipient in the same order as they were sent to the sender in connection-oriented operation. It employs a handshake procedure to establish a link between the recipient and the sender in order to transfer data over the network. As a result, it is often referred to as a dependable network infrastructure. Assume a sender wishes to transmit data to a recipient. The sender then sends a request packet in the form of a SYN packet to the recipient. After that, the receiver sends a (SYN-ACK) signal/packets in response to the sender's message.

That signifies the recipient's acceptance of the sender's request to begin contact with the receiver. The letter or data will now be sent to the recipient by the sender.

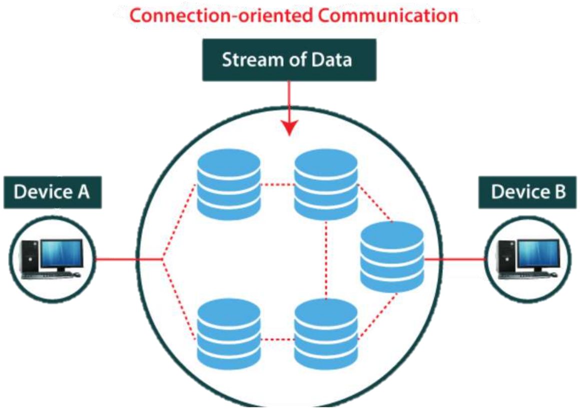


Figure 7 Connection-Oriented Communication

Similarly, a recipient may reply or send data in the form of packets to the sender. A sender may terminate a link by sending a signal to the receiver after successfully exchanging or transmitting data. As a result, we can conclude that it is a dependable network operation.

Example of connection-oriented is TCP.

TCP

TCP (Transmission Control Protocol) is a connection-oriented protocol that establishes links in the same or separate networks to facilitate communication between two or more computer devices. That is the most important protocol for transferring data from one end to the other using the internet protocol. TCP/IP is the abbreviation for Transmission Control Protocol/Internet Protocol.

10.3 Connectionless Service

A link is similar to the postal system in that each letter follows a separate path from the source to the destination address. In a network system, a connectionless service is used to transmit data from one end to the other without establishing a link. As a result, no link must be established before transmitting data from the sender to the recipient. It is not a dependable network service since it does not guarantee the delivery of data packets to the receiver, and data packets will arrive at the receiver in any order. As a result, the data packet does not follow a predetermined direction. The transmitted data packet is not received by the recipient in a connectionless operation due to network interference, and the data may be lost.

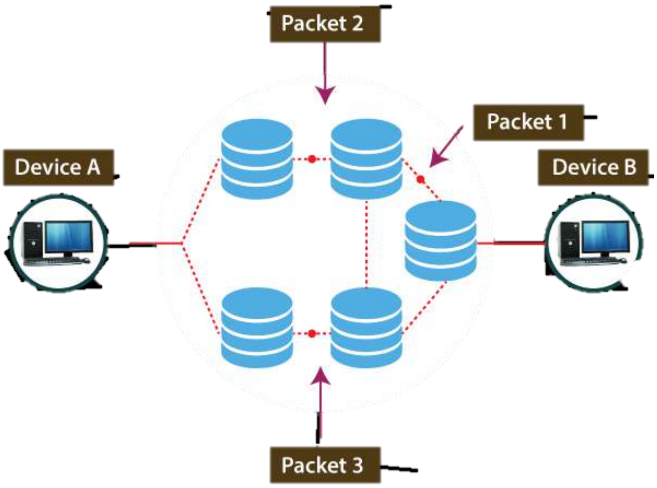


Figure 8 Connectionless Communication

Since it is a connectionless service, a sender can transmit some data to the receiver without first forming a link. The sender's data would be included in the packet or data sources that hold the receiver's address. Data can be sent and received in any order in a connectionless operation. However, it does not promise that the packets will be sent to the correct location.

Example of connectionless service is : UDP

The UDP (User Datagram Protocol) is a connectionless protocol that allows two or more devices to communicate without having to create a link. A sender sends data packets to the receiver with the destination address in this protocol. A UDP does not guarantee that data packets are sent to the right destination or that the sender's data is acknowledged.

The difference between connection-oriented and connectionless is shown in Table 1.

Table 1 Connection-oriented vs Connectionless service

Connection-oriented service	Connectionless Service
It is built on the telephone system in terms of architecture and development.	It is a postal system-based operation.
Until sending data over the same or a separate network, it is used to provide an end-to-end link between the senders and the receiver.	It is used to transmit data packets from senders to receivers without establishing a link.
It establishes a simulated connection between the sender and the recipient.	Between the sender and the recipient, no virtual link or route is established.
Before sending data packets to the recipient, it needs authentication.	Before transmitting data packets, it does not require authentication.
The sender's data packets are returned in the same order that they were delivered.	The order in which data packets are received differs from the order in which they are sent by the sender.
The data packets must be sent over a larger bandwidth.	The data packets must be sent over a low-bandwidth connection.
As it ensures data packets pass from one end to the other with a link, it is a more secure connection service.	Since it does not guarantee the transmission of data packets from one end to the other in order to create a link, it is not a secure connection service.

Since it provides an end-to-end link between the sender and receiver through data transmission, there is no interference.	Owing to the lack of an end-to-end link between the source and receiver for data packet transmission, there could be congestion.
A connection-oriented service is the Transmission Control Protocol (TCP).	Connectionless services include the User Datagram Protocol (UDP), Internet Protocol (IP), and Internet Control Message Protocol (ICMP).

10.4 Reliable vs Unreliable Delivery

- If the application layer program needs reliability, we use a reliable transport layer protocol by implementing flow and error control at the transport layer.
- This means a slower and more complex service.
- In the Internet, there are three common different transport layer protocols. UDP is connectionless and unreliable.
- TCP and SCTP are connection oriented and reliable.

10.5 TCP/IP protocol suite

The original TCP IP protocol suite defines two transport layer protocols. One is UDP, and the other is TCP, which we discussed when I taught you about the OSI model.

Position of UDP in the TCP/IP protocol suite

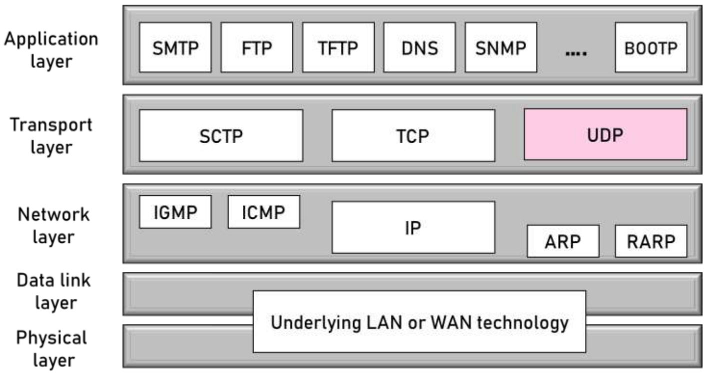


Figure 9 Position of UDP in the TCP/IP protocol suite

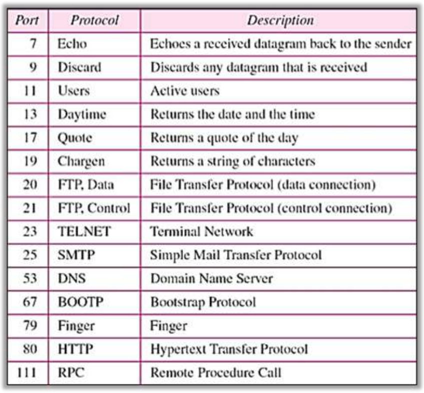
In this case, we'll start with a UDP, which is the easier of the two. Before we get into the TCP. We've also developed a new transport protocol called SCTP.

UDP

So, first and foremost, we'll talk about the UDP protocol, which stands for user datagram protocol in its full form. A connectionless and insecure transport protocol is what it's called. It adds little to the IP's services other than providing process-to-process communication rather than HOST-to-HOST communication. It also does only rudimentary error checking. Why does a mechanism want to use UDP if it is too powerless? As a result of UDP. If there is a downside. It comes with a slew of benefits. The main benefits are that UDP is a very basic protocol that uses very little overhead.

If a process needs to send a small message and isn't concerned about acknowledgement or reliability, it should use UDP. Sending a small message with UDP requires much less contact between the sender and the recipient than sending a small message with TCP or SCTP. We have well-known ports for UDP, as seen in the Figure 10.

Few ports may be used for both UDP and TCP. For example, FTP can use port 21 for either UDP or TCP. SNMP uses two port numbers, 161 and 162, with different purposes.



Port	Protocol	Description
7	Echo	Echoes a received datagram back to the sender
9	Discard	Discards any datagram that is received
11	Users	Active users
13	Daytime	Returns the date and the time
17	Quote	Returns a quote of the day
19	Chargen	Returns a string of characters
20	FTP, Data	File Transfer Protocol (data connection)
21	FTP, Control	File Transfer Protocol (control connection)
23	TELNET	Terminal Network
25	SMTP	Simple Mail Transfer Protocol
53	DNS	Domain Name Server
67	BOOTP	Bootstrap Protocol
79	Finger	Finger
80	HTTP	Hypertext Transfer Protocol
111	RPC	Remote Procedure Call

Figure 10 Well Known Ports for UDP

UDP Header

The UDP header is an 8-byte defined and basic header, while the TCP header will range from 20 to 60 bytes. The first 8 bytes contain all required header information, while the remainder is text. Since each UDP port number field is 16 bits long, the range for port numbers is 0 to 65535; port number 0 is reserved. Port numbers are used to differentiate between various user queries or processes.

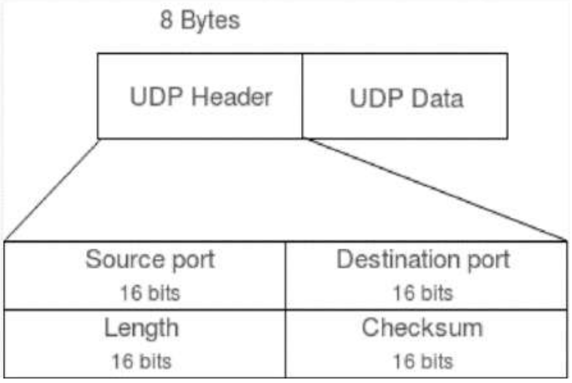


Figure 11 UDP Datagram Protocol

Source Port : Source Port is a two-byte field that identifies the source's port number.

Destination Port: This is a two-byte field that identifies the destined packet's path.

Length: The length of UDP, including the header and data, is measured in bytes. It's a region of 16 bits.

Note: Checksum calculation is not needed in UDP, unlike TCP. UDP does not have error management or flow control. As a result, UDP relies on IP and ICMP to record errors.

10.6 Applications of UDP

- When the size of the data is small, there is less concern for flow and error management, because it is used for basic request-response communication.
- Since UDP supports packet swapping, it is a good protocol for multicasting.
- Few routing upgrade protocols, such as RIP, use UDP (Routing Information Protocol).
- Often used in real-time systems where uneven delays between portions of a sent message cannot be tolerated.
- Following implementations uses UDP as a transport layer protocol:

- NTP (Network Time Protocol)
- DNS (Domain Name Service)
- BOOTP, DHCP.
- NNP (Network News Protocol)
- Quote of the day protocol
- TFTP, RTSP, RIP.

Some activities may be performed by the application layer using UDP, such as route tracing, route recording, and route time stamping.

UDP receives datagrams from the Network Layer, attaches their headers, and sends them to the customer. As a result, it works quickly.

If you delete the checksum field from UDP, it becomes a null protocol.

1. Reduce the amount of computing resources used.
2. When transferring via Multicast or Broadcast.
3. Real-time packet transfer, primarily in multimedia applications

10.7 TCP 3-Way Handshake Process

TCP stands for Transmission Control Protocol, which means it does everything to ensure the data is transmitted in a secure manner.

The latest TCP/IP suite paradigm governs the mechanism of communication between devices over the internet (stripped out version of OSI reference model).

The Application layer is the top layer of the TCP/IP architecture, from which client-side network applications such as web browsers create connections with the server. The information is moved from the application layer to the transport layer, which is where our subject appears.

TCP and UDP (User Datagram Protocol) are two important protocols in this layer, with TCP being the most common (since it provides reliability for the connection established). However, UDP can be used to test the DNS registry in order to get the binary equivalent of the website's domain name.

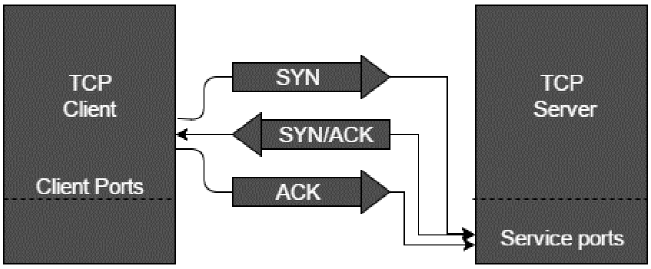


Figure 12 TCP Handshake

Positive Acknowledgement with Re-transmission is a feature of TCP that ensures secure communication (PAR). The transport layer's Protocol Data Unit (PDU) is known as segment. Still, once it gets an acknowledgment, a system using PAR can resend the data unit. If the data unit received at the receiver's end is corrupted (it scans the data with the transport layer's checksum feature for Error Detection), the section is discarded. As a result, the sender must resend the data device about which there is no positive acknowledgment.

As you can see from the above mechanism, three segments are exchanged between the sender (client) and receiver (server) in order to create a secure TCP link as seen in Figure 13.

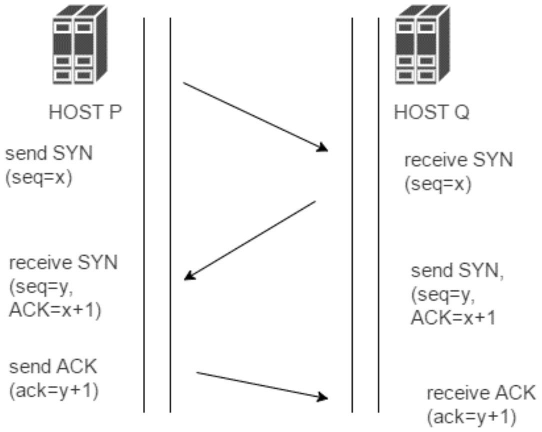


Figure 13 Working

Stage 1 (SYN): In the first step, the client needs to create a link with the server, so it sends a segment with SYN (Synchronize Sequence Number), which tells the server that the client is likely to begin contact and with what sequence number it will begin segments.

Step 2 (SYN + ACK): When the server receives a client message, it sets the SYN-ACK signal bits. Acknowledgement (ACK) denotes the answer of the segment it got, while SYN denotes the sequence number from which it is likely to begin the segments.

Step 3 (ACK): In the final step, the client accepts the server's response and the two of them create a secure link to begin the data transfer. The relation parameter (sequence number) for one direction is established and acknowledged in steps 1 and 2. The relation parameter (sequence number) for the other direction is established and acknowledged in steps 2 and 3. A full-duplex connectivity is developed with these.

Note

When forming relations between the client and the server, the initial sequence numbers are chosen at random.

Summary

- The transport layer of the OSI reference model allows direct data transmission between source and destination machines by using network layer services such as IP to pass PDUs of data between the two communicating machines.
- The transport layer is an end-to-end or source-to-destination layer. To ensure full data sharing, the OSI Transport layer protocol (ISO-TP) manages end-to-end monitoring and error checking. It allows “peer to peer” contact with the destination machine's transport entity (remote peer).
- The transport layer adds a secure layer on top of the network layer's insecure networks. Option negotiation among various quality of service criteria provides consumer applications with efficient, dependable, and cost-effective transportation services.
- Flow control regulates data transfer between devices, ensuring that the transmitting device sends no more data than the receiving device can handle. Data from multiple applications can be sent over a single physical connection using multiplexing.
- On top of the network layer, transport primitives are an efficient means of transmitting data. The transport layer facilities tend to be identical to those delivered at the data link layer. However, they vary in several respects, with the data link layer using physical channels to bind two routers and the transport layer using subnets.

- UDP is a connectionless, insecure protocol that helps processes run faster by reducing the burden on their CPUs. The performance challenges, which lack a scientific paradigm to back them up, are backed up by personal perceptions and examples. They try to solve problems in computer networks by testing network efficiency, designing systems for improved performance, handling TPDUs quickly, and developing protocols for future high-performance networks.

Keywords

Addressing: Addressing or tagging a frame is handled by the Transport Layer.

Connection Establishment Delay: That's the length of time it takes for the destination system to accept that a connection has been demanded. Obviously, the shorter the wait time, the better the service.

Connection Establishment Failure Probability: Because of network congestion, a lack of table space, or other internal issues, the link does not develop within the defined delay.

Connection Establishment/Release: A naming function is used in the transport layer for forming and releasing links across the network, so that a process on one computer may indicate with whom it wants to interact.

Demultiplexing: Where multiple links are multiplexed, demultiplexing is needed at the receiving end.

Error Control: To prevent errors caused by missing or overlapping segments, the transport layer assigns specific segment sequence numbers to individual message packets, forming virtual circuits with only one virtual circuit per session.

Flow Control: The fundamental principle of flow control is to keep a quick and slow mechanism in synergy. The transport layer makes it possible for a fast process to keep up with a sluggish one.

Fragmentation: When the transport layer receives a large message from the session layer, it divides the message into smaller units as required.

Multiplexing: The transport layer provides several network links to increase throughput.

Throughput: In a given time period, it specifies the amount of bytes of user data transmitted per second. It is calculated separately for each contact source.

Transmission Control Protocol: It allows a secure data distribution service with error detection and correction from beginning to end.

User Datagram Protocol(UDP): It is an insecure connectionless datagram protocol in which the transmitting terminal does not validate if data has been received by the receiving terminal.

Review Questions

1. When the facilities delivered at both layers are almost identical, how is the transport layer different from the data link layer?
2. Why transport layer is required when both the network and transport layers provide connectionless and connection oriented services?
3. What are the different quality of services parameters at the transport layer?
4. Why UDP is used when it provides unreliable connectionless service to the transport layer?
5. What is the purpose of flow control?
6. Describe the TCP and its major advantages over UDP.

Self-Assessment

1. UDP known as
    a. User Datagram Protocol
    b. Unity Data Packet

   c. User datagram Packet
   d. None of Above

2. TCP is a
   a. Telnet Control Protocol
   b. Transmission Cent Protocol
   c. Transmission Control Protocol
   d. None of above
3. FTP uses port
   a. 22
   b. 23
   c. 24
   d. 21
4. Which of the following is false with respect to UDP?
   a. Connection-oriented
   b. Unreliable
   c. Transport layer protocol
   d. Low overhead
5. What is the main advantage of UDP?
   a. More overload
   b. Reliable
   c. Low overhead
   d. Fast
6. The ______ field is used to detect errors over the entire user datagram.
   a. udp header
   b. checksum
   c. source port
   d. destination port
7. Which is the correct expression for the length of UDP datagram?
   a. UDP length = IP length – IP header's length
   b. UDP length = UDP length – UDP header's length
   c. UDP length = IP length + IP header's length
   d. UDP length = UDP length + UDP header's length
8. Port number used by Network Time Protocol (NTP) with UDP is ________
   a. 161
   b. 123
   c. 162
   d. 124
9. What is the header size of a UDP packet?
   a. 8 bytes
   b. 8 bits
   c. 16 bytes
   d. 124 bytes
10. Beyond IP, UDP provides additional services such as _______
   a. Routing and switching
   b. Sending and receiving of packets
   c. Multiplexing and demultiplexing

d. Demultiplexing and error checking

11. Beyond IP, UDP provides additional services such as _______
    a. Routing and switching
    b. Sending and receiving of packets
    c. Multiplexing and demultiplexing
    d. Demultiplexing and error checking

12. IP Security operates in which layer of the OSI model?
    a. Network
    b. Transport
    c. Application
    d. Physical

Self-Assessment Answers

1.	a	2.	c	3.	d	4.	a
5.	c	6.	b	7.	a	8.	b
9.	a	10.	d	11	d	12.	a

Further Readings

Achyut S Godbole and Atul Kahate published, *Web Technologies,* Tata McGraw Hill.

Andrew S. Tanenbaum, *Computer Networks,* Prentice Hall.

Behrouz A. Forouzan, Sophia Chung Fegan, *Data Communications and Networking,* McGraw-Hill Companies.

Douglas Comer, Computer Networks and Internets with Internet Applications, 4th Edition, Prentice Hall.

Ferguson P., Huston G., Quality of Service: Delivering QoS on the Internet and in *Corporate Networks*, John Wiley & Sons, Inc., 1998.

J. D. Spragins, Telecommunications Protocols and Design, Addison-Wesley.

McDysan, David E. and Darren L. Spohn. ATM Theory and Applications, McGraw-Hill Osborne Media, 1998.

Nassar, Daniel J., *Ethernet and Token Ring Optimization*. iUniverse.com, 2000.

Spurgeon, Charles E. Ethernet, *The Definitive Guide*. O'Reilly & Associates, 2000.

William A Shay, Understanding Communication and Networks, 3rd Edition, Thomson Press.

Dr. Rajni Bhalla, Lovely Professional University

Unit 11: Transport layer - congestion control and QoS

CONTENTS

Objectives

- understand traffic descriptor and traffic profiles.
- understanding congestion control categories.
- learning flow characteristics.
- techniques to improve quality of service.
- learning traffic shaping techniques.

Introduction

So, first and foremost, congestion management and quality of service are two problems that are so closely linked that improving one automatically improves the other. You're missing one strategy if you're ignoring the other as well. As a result, most strategies for preventing or eliminating congestion in a network often increase network quality of operation. A major target of congestion management and quality of service is data traffic. In congestion control, we aim to reduce traffic congestion, and in quality of service, we strive to build an optimal atmosphere for traffic.

11.1 Data Traffic

But, before we get into congestion management and service efficiency, let's take a look at data flow. So, under the heading of data flow, we'll talk about two terms:

- traffic descriptors and
- traffic profiles.

Traffic Descriptor

You can see the data rate, absolute burst size peak rate, and average data rate in the Figure 1 below. What does the information imply? The average data rate is the number of bits transmitted over time divided by the number of seconds in that time, indicating how much data you will transmit in a given second. The average data rate is a very helpful traffic characteristic because it shows how much bandwidth the traffic requires on average.

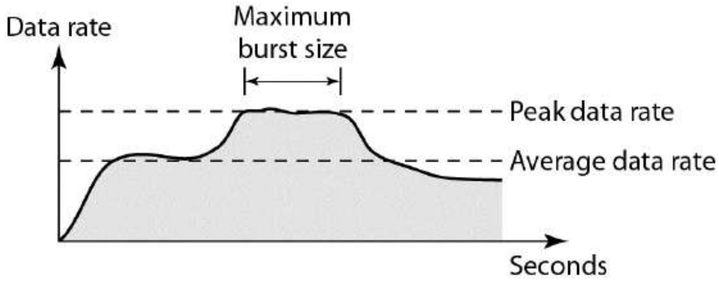


Figure 1 Traffic Descriptors

The maximum y axis value, the peak data rate, is a very significant metric, as you can see in the graph. Peak data rate indicates that there is a sudden increase in traffic, which is why the bar is rising. Peak data rate indicates that there was a lot of traffic at the time, and burst duration indicates how long the traffic was at its peak. As a result, peak data rate is a critical metric since it shows the network's maximum bandwidth, which is needed for traffic to flow without interruption. So, although the peak data rate is certainly a key advantage for the network, the actual burst size should normally be overlooked. if the peak value's period is very short. Let's look at a road case. Suddenly, there is a lot of traffic because of an accident, so whether that traffic is just for a few seconds, or only a few minutes, it is not going to affect any traffic, then we can tolerate this sort, but if it is for a long period of time, it would certainly affect. So at the end, always remember that effective bandwidth is the bandwidth that the network needs to allocate for the flow of traffic, the effective bandwidth is a function of three values

Traffic Profiles

A data flow should provide one of the following traffic profiles for our purposes. As you can see in the Figure 2, we have three figures: one for steady bitrate, another for volatile bitrate, and the third for burst bitrate. I believe it is obvious from the diagram and the expression constant bitrate that it refers to a fixed rate at which time the traffic is unchanged. There has been no improvement in traffic. The traffic is moving at a really fast pace. As a result, the average data rate and peak data rate are always the same in this form of traffic.

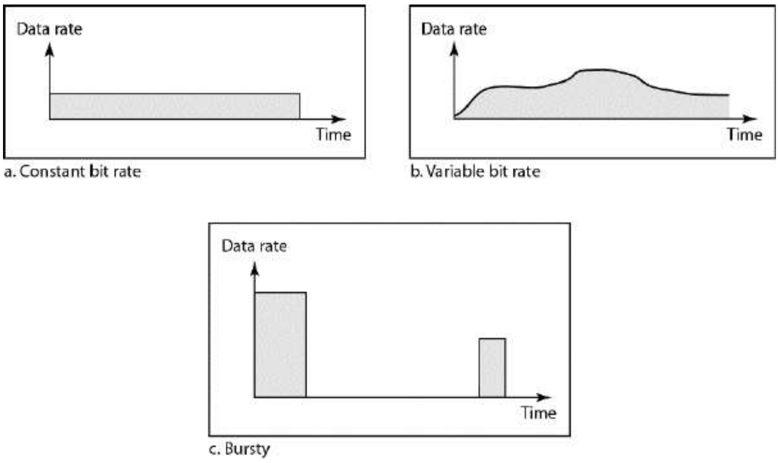


Figure 2 Three Traffic Profiles

The predictable network knows how much bandwidth to set aside for this kind of traffic ahead of time. Variable bitrate is the next choice. In this class of segment, the rate of data flow varies over time, implying that the changes are gradual rather than abrupt. There are days where there is more traffic than there is when there is less traffic. We know what we do if the timing is nine to five if we are operating with numbers, such as whether we are learning at a university or in a college. Right, but there will undoubtedly be a lot of traffic at 5 p.m. But we'll certainly have to deal with traffic, and there will be some variation: there will be a high, and then there will be a low.

As a result, rather than abrupt and sharp shifts, the data flow changes over time with gradual transitions. The average data rate and the peak data rate are both different in this form of flow. Typically, the maximal burst size is a small number. Since it does not need to be reshaped, this kind of traffic is much more difficult to manage than continuous bitrate traffic. The next one is the bursty data rate, which gets its name from the fact that there was no traffic for a long time but then a massive influx of traffic.

As a result, the data rate varies abruptly in a brief amount of time, implying that it is jumping from zero. As an example, one Mbps can be converted to microseconds and vice versa. It's possible that it'll stay at this level for a bit. Because of the abrupt shift in traffic, average bitrate and peak data rate, or both, are different values in this situation. If you can see in the diagram, there was no data flow when there was no traffic, but suddenly there is a lot of traffic. Bursty traffic congestion in the network may occur if the demand on the network is greater than the network's capability. For instance, there is excessive traffic on the road. The road's power was smaller than the load it could handle. Congestion management refers to the mechanisms and procedures used to control congestion and maintain load below availability. Congestion is a significant problem in a packet switch network.

Network Performance

So, let's talk about network performance and how it affects it. As you can see in the diagram, there are input queues and output queues. Since routers and switches have queues buffers that carry packets before and after processing, network or inter network congestion exists.

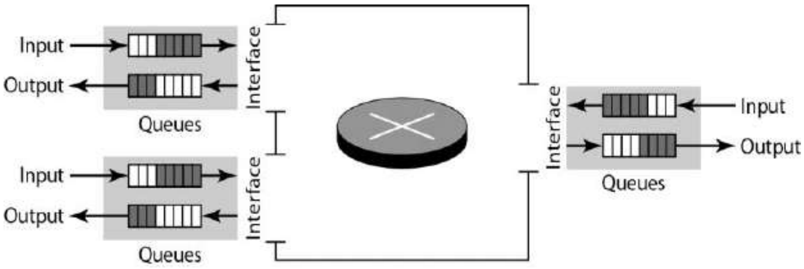


Figure 3 Queues in a Router

Each interface on a router, for example, has its own input and output queue as seen in Figure 3. When a packet is sent at the inbound terminal. There are three stages in the process. The packet is placed at the bottom of the input queue and will be reviewed later. The packet is removed from the input queue by the router's processing machine. The packet is placed in the proper output queue until it enters the front of the queue and uses its routing table and destination address to locate the source. And now it's your turn to submit. As a result, teachers, we must be mindful of two problems. If the rate of packet arrival exceeds the rate of packet processing, so indeed. The input queues are becoming increasingly long. The performance queue becomes longer and longer if the packet departure rate is lower than the packet processing rate. As a function of load throughput, packet delay and throughput indicate processor performance. So, congestion management requires two metrics that assess a network's performance: throughput and delay throughput. Throughput refers to the network's efficiency. If the power is less than the load, the throughput would inevitably be lower. When the power exceeds the load. As a result, we will almost certainly achieve high network reliability, allowing us to avoid congestion before it occurs or to alleviate congestion after it occurs as seen in Figure 4.

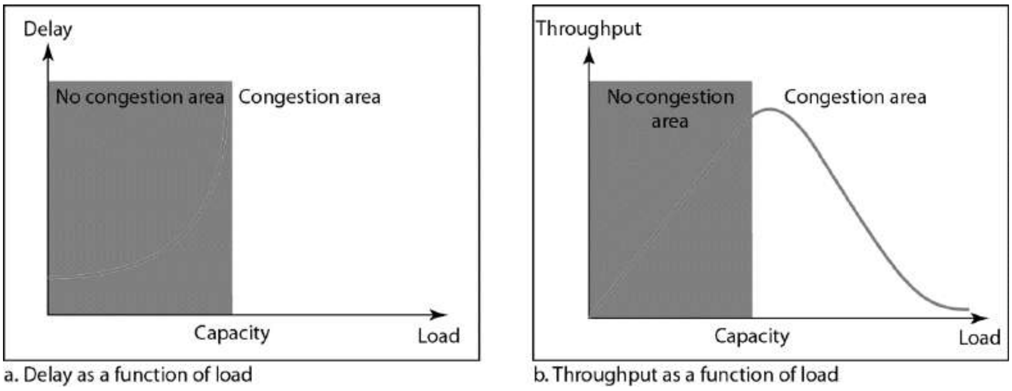


Figure 4 Packet Delay and throughput as a function of Load

11.2 Congestion Control

As a result, we can split these congestion management methods into two categories: open loop and closed loop. Open loop refers to the avoidance of congestion before it occurs. Closed loop refers to the steps that can be taken before congestion occurs.

- open-loop congestion control (prevention) and
- closed-loop congestion control (removal).

Now that we know the congestion has already happened, the next step is to figure out how to get rid of it. I understand that there is still so much traffic on the lane, but I'm not sure how to alleviate it.

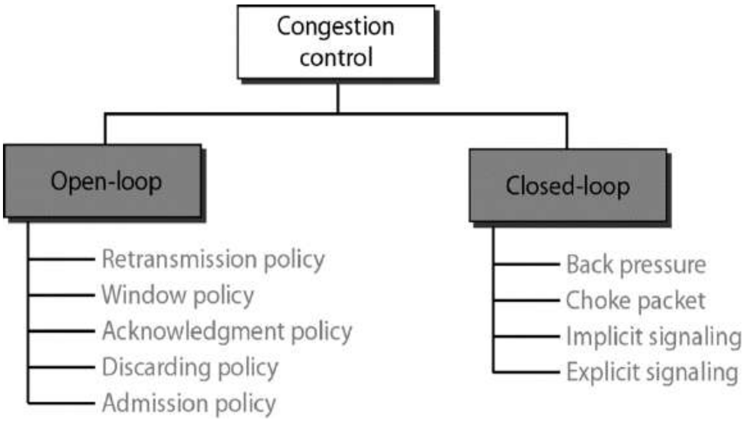


Figure 5 Categories of Congestion Control

Open Loop

To begin with, open loop, we are going to take several steps to avoid congestion and ensure that it does not exist. So, we're going to talk about open loop congestion management strategies first, and then we'll talk about other tactics. As a result, they are used to avoid congestion until it occurs. Either the source or the destination is in charge of congestion management. Retransmission policy, window policy, acknowledgment policy, discarding policy, and admission policy are some of the strategies used in open loop.

Retransmission Policy

The first is a collection of retransmission policies. Retransmission is often inevitable if the sender suspects that a packet has been destroyed or compromised, in which case the packets must be resent. If you're retransmitting the packet, this is the case. It is certain that it will increase congestion. It is likely that your packet will not meet its destination due to congestion because there is still so much traffic, and your packet will get trapped in the traffic. But, if you retransmit it, it will increase traffic, potentially increasing network congestion.

A successful retransmission strategy, on the other hand, will avoid congestion. What we should do here is plan the retransmission protocol and retransmission timers to maximise reliability while still preventing congestion. TCP's retransmission strategy, for example, is intended to avoid or reduce congestion.

Window policy

The second is the window policy. We're talking about a selective repeat window, as you can probably tell from the label. As a result, congestion can be affected in this kind of window at the sender. For congestion management, the Selective Repeat Window is preferable to the Go Back. We've also covered these two strategies in previous videos: selective repeat and go back N.

Congestion can also be affected by the type of window on the sender side. Although several packets may be successfully received at the receiver side, some packets in the Go-back-n window are resent. This duplication has the potential to worsen the network's congestion. As a result, the Selective Repeat Window can be used because it sends the same packet that was missed.

Discarding Policy

A positive discarding strategy implemented by routers is that it allows them to avoid congestion while still partly discarding compromised or less sensitive packages while maintaining message accuracy. When transmitting audio files, routers will discard less sensitive packets to avoid congestion while maintaining the audio file's consistency.

Acknowledgment Policy

Since recognition is a component of network load, the acknowledgment policy enforced by the recipient may have an impact on congestion. Congestion caused by acknowledgment can be avoided using a variety of methods. Rather than sending acknowledgement for a single packet, the receiver should send acknowledgement for N packets. And when a packet must be sent or a timer expires should the recipient send an acknowledgment.

Admission Policy

A system should be used in admission policies to avoid congestion. Until transmitting a network flow further, switches in a flow can review the resource requirements. To avoid more congestion, the router should refuse building a virtual network link if there is a risk of congestion or if the network is still congested.

Closed Loop

In the closed loop strategy, we attempt to alleviate congestion after it has occurred. Different protocols will utilize a variety of mechanisms to accomplish their goals. So, the first is back pressure.

Back Pressure

As seen in the diagram below. Back pressure refers to a congestion control mechanism in which a congestion node blocks data from the upstream node or nodes it is connected to. Back pressure is a node to node congestion control that starts with the node and propagates in the opposite direction of a data flow to the source, back pressure is a node to node congestion control that starts with the node and propagates in the opposite direction of a data flow to the source, only a virtual circuit network in which each node knows the upstream node for whom a data flow is arriving may use the backpressure approach. Backpressure, as seen in the picture, indicates that they are moving in the other direction. Assume that node three in the Figure 6 receives more data than it can process. So, node three node three, we have congestion indicated on the diagram, which indicates it has more input data, which means capacity is smaller but demand is greater. Congestion has already happened at this time.

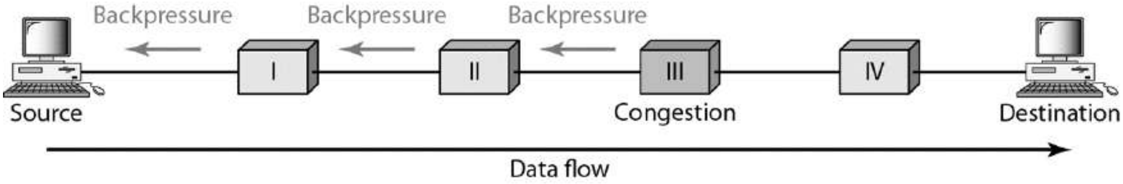


Figure 6 Backpressure

So, what it will do is drop some packets in their input buffer, and it will surely tell the node to upstream mode, which means it will notify the node instantly that it is connected to this. As a result, it will tell the node to slow down. Please don't send me any more packets, since node two will undoubtedly alert node one. If node two is congested, it notifies node one to slow down, which may cause congestion. If this is the case, node one notifies the source offer data that it should slow down. This, in turn, relieves congestion over time. It's worth noting that the pressure on node three is directed backwards to the source in order to relieve congestion; they aren't immediately informing the source. They're alerting the upstream node; node three is telling node two to shut down; node two will check to see if it's already full, and then it'll tell node one to shut down; node one will tell the source to slow down; and node one will tell the source not to transmit any more packets. Because there is already congestion on the network, they are not immediately telling the source in the event of backpressure.

Choke Packet

Choke packets are packets delivered by a node to the source warning them of congestion. They are directly telling the source of the congestion. The backpressure approach differs from the choke packet approach in that the warning is sent from one node to its upstream node. Although the warning may finally reach the source station in the choke packet, it is from the router, which has met congestion straight to the source station. The intermediary nodes, such as node second or node first, through which the packet has passed, are not being alerted here.

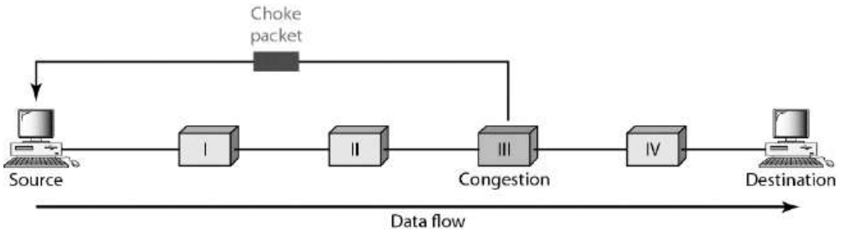


Figure 7 Choke Packet

The Figure 7 is drawn here, as you can see in the illustration. This form of control may be seen in the ICMP Internet Control Message Protocol, as you can see. When an internet router becomes overburdened with IP datagrams. It's possible that some of them will be discarded. So, it sends a source quench ICMP message (Internet Control Message Protocol) to the source code, and the warning message is sent straight to the source station, bypassing the intermediate routers. I have informed you that the ICMP packet is the lowest priority packet when I taught you about the ICMP Internet Control Message Protocol.

Implicit Signaling

If there is no communication between the crowded node and the source, who will inform about the congestion? In this instance, the source will assume that the network is crowded, especially if it has been for a long time. Because the source is not receiving any type of acknowledgement or attention, it will presume that the network is congested. When a sender sends several packets and no acknowledgement is received over an extended period of time, the source will assume that there is a congestion and that packets are not arriving.

Explicit Signaling

If a node encounters congestion, it will send a packet to the source or destination to advise them of the problem. As a result, congestion can occur as a result of the packet or the acknowledgement. As a result, it is also a responsibility to tell both the source and the destination about the congestion. As a result, explicit communication can take place in either a forward or a backward manner. When forward signalling is used, the signal is transmitted towards the direction of the congestion. The destination is alerted to the traffic. There is a source node destination node in the case of forward destination means the destination. In the case of a forward signalling receiver, use the following strategies to avoid further congestion. In the instance of backward signalling, informed sources advised that it was necessary to slow down and stop transmitting extra packets due to congestion.

11.3 Quality of Service

Quality of service (QoS) is an internetworking topic that has gotten more attention than it has been defined. Informally, we might describe quality of service as something that a flow strives towards.

Flow Characteristics

A flow has traditionally been assigned four characteristics: reliability, delay, jitter, and bandwidth as shown in Figure 8.

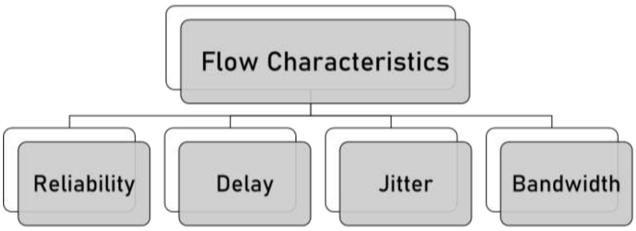


Figure 8 Flow Characteristics

Reliability

A flow requires the quality of reliability. Losing a packet or acknowledgement due to a lack of dependability necessitates retransmission. The sensitivity of application programmes to dependability, on the other hand, is not the same. Electronic mail, file transfer, and Internet connection, for example, are more crucial than phone or audio conferencing in terms of reliability.

Delay

Another flow feature is source-to-destination latency. Again, various applications may tolerate delays to varying degrees. Telephony, audio conferencing, video conferencing, and remote log-in all require minimal latency in this situation, whereas file transmission and e-mail are less critical.

Jitter

Jitter is the difference in delay between packets in the same flow. For example, if four packets leave at 0:01:02:03 and arrive at 20:21:22:23, they all have the same 20-unit delay. If the aforementioned four packets arrive at 21, 23, 21, and 28, on the other hand, they will have distinct delays: 21,22, 19, and 24.

The first scenario is perfectly suitable for audio and video applications; the second situation is not. It makes no difference whether the packets arrive with a small or lengthy delay for these applications as long as the delay is consistent across all packets. The second example is not acceptable for this application. The variance in packet delay is referred to as jitter. High jitter denotes a considerable variance in delays; low jitter denotes a modest variance.

Bandwidth

Various applications necessitate various bandwidths. To refresh a colour screen in video conferencing, we need to transfer millions of bits every second, although the total amount of bits in an e-mail may not even approach a million.

11.4 Techniques to improve quality of service

We go through various approaches that may be utilized to improve service quality. We'll go through four typical ways briefly: scheduling, traffic shaping, admission control, and resource reservation.

Scheduling

Different flows of packets arrive at a switch or router for processing. A smart scheduling strategy balances and appropriately distributes the various flows. Several scheduling approaches have been developed to improve service quality. Three of them are discussed in this article: ***FIFO queuing, priority queuing, and weighted fair queuing.***

FIFO Queuing

Packets wait in a buffer (queue) until the node (router or switch) is ready to handle them in first-in, first-out (FIFO) queuing. The queue will fill up if the average arrival rate is higher than the average processing rate, and incoming packets will be deleted. Those who have had to wait for a bus at a bus stop are familiar with a FIFO queue. Figure 9 is a conceptual representation of a FIFO queue?

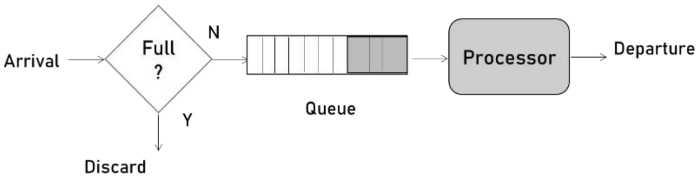


Figure 9 FIFO Queuing

Priority Queuing

Packets are initially assigned to a priority class in priority queuing. There is a separate queue for each priority class. The highest-priority queue packets are handled first. The lowest-priority queue packets are handled last. It's worth noting that the system doesn't cease servicing a queue until it's completely empty. demonstrates priority queuing with two degrees of priority. Figure 10 demonstrates priority queuing with two degrees of priority.

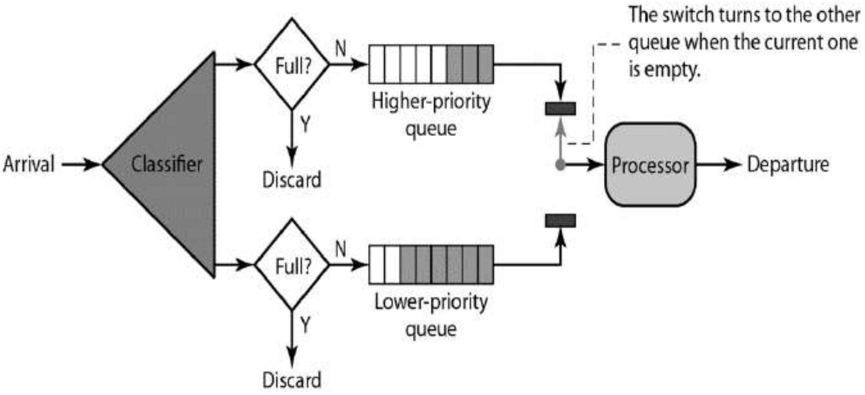


Figure 10 Priority Queuing

Because higher-priority traffic, such as multimedia, may reach its destination with less latency, a priority queue can provide better QoS than a FIFO queue. There is, however, a possible disadvantage. If a high-priority queue has a continuous flow, packets in lower-priority queues will never be processed. This is referred to as starving.

Weighted Fair Queuing

Weighted fair queuing is a superior scheduling strategy. The packets are still allocated to distinct classes and accepted to distinct queues using this method. The queues, on the other hand, are weighted according to their priority; a greater priority indicates a larger weight.

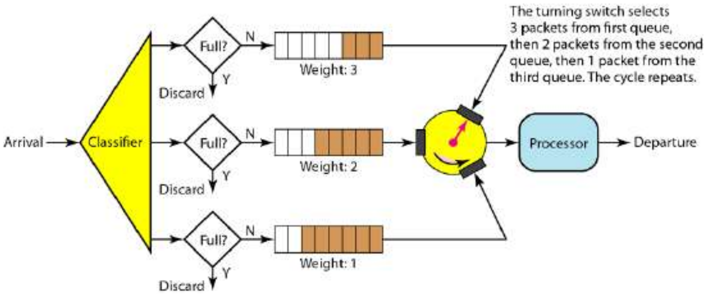

The turning switch selects 3 packets from first queue, then 2 packets from the second queue, then 1 packet from the third queue. The cycle repeats.
Arrival
Classifier
Full?
N
Y
Discard
Weight: 3
Full?
N
Y
Discard
Weight: 2
Full?
N
Y
Discard
Weight: 1
Processor
Departure

Figure 11 Weighted Fair Queuing

The system processes packets in each queue round-robin, with the number of packets chosen from each queue dependent on their weight. For example, if the weights are 3, 2, and 1, the first queue processes three packets, the second queue two, and the third queue one.

All weights can be equal if the system does not assign priority to the classes. As a result, we have fair and prioritised queuing. The methodology is shown in Figure 11 with three classes.

11.5 Traffic Shaping

Traffic shaping is a method of controlling the volume and pace of data transmitted over a network. Leaky bucket and token bucket are two strategies that may be used to shape traffic.

Leaky Bucket Algorithm

Water leaks from a bucket with a tiny hole at the bottom at a consistent pace as long as there is water in the bucket. Unless the bucket is empty, the pace at which the water leaks is unrelated to the pace at which the water is added to it. The input rate can change, but the output rate is always the same. In networking, a method known as leaky bucket helps smooth out spikes in traffic. Bursty pieces are collected in a bucket and distributed at a steady rate. Figure 12 depicts the impacts of a leaking bucket.

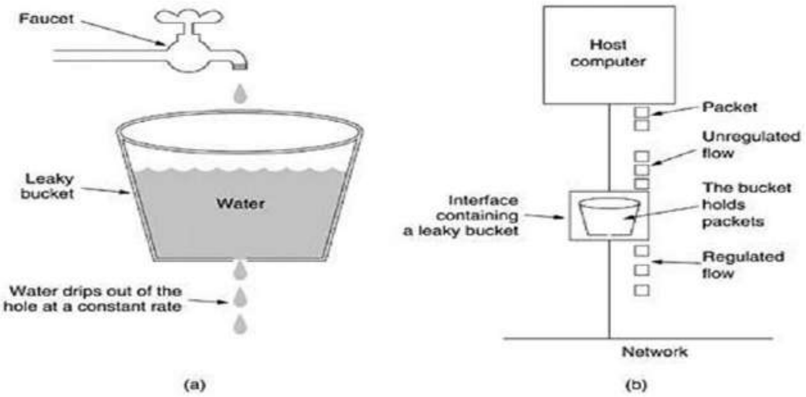


Figure 12 Leaky Bucket Algorithm

In the diagram, we suppose that the network has allocated 3 Mbps of bandwidth to a host. The leaky bucket is used to shape the input traffic so that it adheres to the commitment. Figure 12 shows a total of 24 Mbits of data, the host delivers a burst of data at a rate of 12 Mbps for 2 seconds. After 5 seconds of silence, the host delivers data at a rate of 2 Mbps for 3 seconds, totaling 6 Mbits of data. The host has transmitted 30 Mbits of data in lOs in total. The leaky bucket smooths traffic by transmitting data at a 3 Mbps rate over the same 10 seconds. The first burst may have harmed the network if it hadn't been for the leaky bucket, since it would have consumed more bandwidth than was allocated to this server. We can also understand how the leaking bucket may help to avoid congestion. Consider the motorway at rush hour as an example (bursty traffic). Congestion on our highways may be eliminated if commuters simply stagger their working hours.

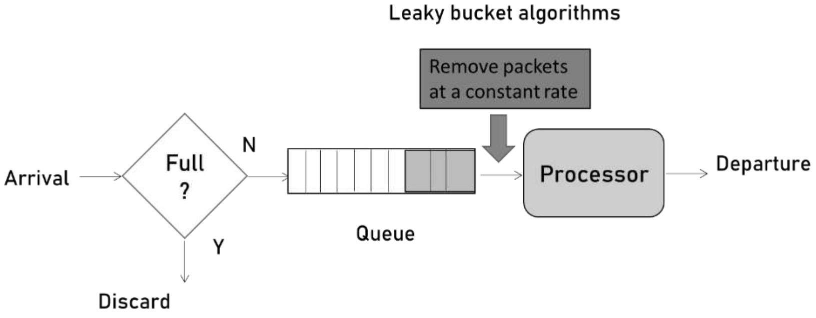


Figure 13 Leaky Bucket Implementation

Figure 13 depicts a simple leaky bucket implementation. The packets are stored in a FIFO queue. If the traffic is made up entirely of fixed-size packets. At each tick of the clock, the operation eliminates a set number of packets from the queue. If the traffic is made up of packets of varying lengths, the fixed output rate must be determined by the amount of bytes or bits.

Token Bucket

Having a leaking bucket is really restricting. An idle host is not credited. For example, if a host does not send for an extended period of time, its bucket will become empty. The leaky bucket now only supports an average rate if the host has bursty data. The duration of the host's inactivity is not taken into account. The token bucket method, on the other hand, allows idle hosts to save credit for the future in the form of tokens. The system transfers n tokens to the bucket for each tick of the clock. For each cell (or byte) of data supplied, the system eliminates one token. The bucket accumulates 10,000 tokens if n is 100 and the host is idle for 100 ticks. The host can now eat all of these tokens in a single tick with 10,000 cells, or 1000 ticks with 10 cells each. In other words, as long as the bucket isn't empty, the host can deliver bursty data. Figure 14 demonstrates the concept.

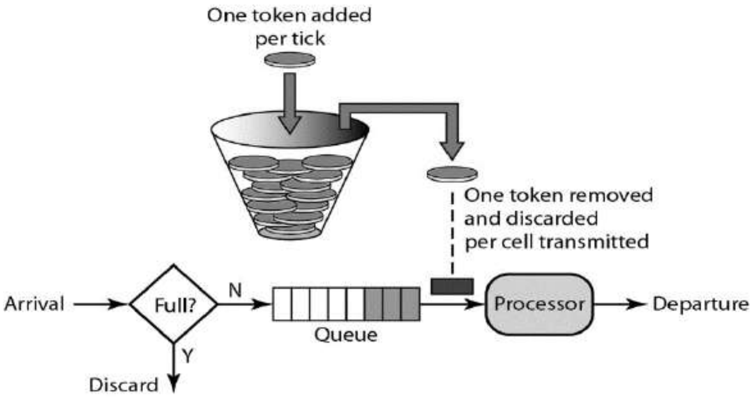


Figure 14 Token Bucket Algorithm

A counter may simply be used to implement the token bucket. The token's value is set to zero. The counter is increased by one each time a token is inserted. The counter is decremented by one each time a unit of data is transmitted. The host is unable to transfer data when the counter is 0.

Combining token bucket and leaky bucket

Both strategies can be used together to credit an idle host while also regulating traffic. After the token bucket, the leaky bucket is used; the rate of the leaky bucket must be greater than the rate of tokens dropped in the bucket.

Resource Reservation

A data flow necessitates the use of resources such as a buffer, bandwidth, and CPU time, among others. When these resources are booked ahead of time, the quality of service is increased. In this part, we'll look at one QoS model called Integrated Services, which relies significantly on resource reservation to improve service quality.

Admission Control

Admission control is the process by which a router or switch accepts or rejects a flow based on established parameters known as flow requirements. Before accepting a flow for processing, a router examines the flow specifications to verify if the router's capacity (bandwidth, buffer size, CPU speed, and so on) and past commitments to other flows are sufficient to handle the incoming flow.

Summary

- When there are too many packets in one region of the network, the subnet's performance suffers. As a result, a network's communication channel is considered crowded if packets transiting the path incur delays that are much greater than the way's propagation delay. When packets never reach their destination, the delay approaches infinite, this is referred to as extremely congested.
- The terms "congestion control" and "flow control" are not interchangeable. Flow control is concerned with point-to-point traffic between a particular source host and a specific destination host, whereas congestion is a global phenomena affecting all hosts, all routers, the store-and-forward processing within the routers, and so on.
- The computer network, which is also a system, is separated into two categories according to control theory. There are two types of solutions: open loop and closed loop.
- The traffic management capability helps you to make the most of your network's resources while also ensuring that resources that haven't been expressly assigned are used efficiently. The majority of traffic management will be determined by transmit priority and bandwidth availability. Delay-sensitive traffic is given a higher transmit priority in the transmit priority.

- The leaky bucket technique is used in network traffic shaping and rate limiting applications. The technique allows you to manage the rate at which data is injected into a network, allowing you to handle data rate burstiness.

Keywords

Congestion: When packets crossing a network's communication channel incur delays that are much greater than the path's propagation delay, the channel is said to be congested.

IP Address: An Internet Protocol address (IP address) is a numerical label assigned to each device (e.g., computer, printer) participating in a computer network that uses the Internet Protocol for communication.

IP Protocol: The Internet Protocol Suite's datagrams (packets) are relayed across an internetwork using the Internet Protocol Suite's Internet Protocol (IP).

Traffic Shaping: Attempts to normalise the average data transfer rate.

Review Questions

1. Explain the general principles of congestion.
2. What do you understand by QoS? Describe the basic QoS structure.
3. Discuss the following two algorithms:
    a. Leaky Bucket
    b. Token Bucket
4. What are two types of congestion control? Where is congestion control implemented in each case?
5. Explain all traffic shaping techniques.
6. Write down techniques to improve quality of service.
7. Difference between token bucket and leaky bucket algorithm.

Self-Assessment

1. The technique in which a congested node stops receiving data from the immediate upstream node or nodes is called as
    a. Explicit signalling
    b. Back pressure
    c. Implicit signalling
    d. Redundant signals
2. A leaky bucket algorithm shapes bursty traffic into fixed-rate traffic by averaging the
    a. Data rate
    b. Average rate
    c. Traffic rate
    d. Traffic shaping
3. Two classes of services have been defined for
    a. Integrated services
    b. Quality data services
    c. Technical services
    d. Protocol services
4. In open-loop congestion control, policies are applied to
    a. Prevent congestion
    b. Discard congestion
    c. Maximize congestion

 d. Eliminate congestion

5. A mechanisim to control the amount and the rate of the traffic sent to the network is called
 a. Traffic congestion
 b. Traffic flow
 c. Traffic control
 d. Traffic shaping
6. Scheduling is done by
 a. Weighted fair queuing
 b. FIFO
 c. Random
 d. LIFO
7. Which of the following is a congested control algorithm.
 a. The leaky bucket
 b. Token bucket
 c. Resource reservation
 d. All of above
8. In Congestion Control, the packet is put at the end of the input queue while waiting to be
 a. checked
 b. entered
 c. read
 d. interpret
9. Integrated Services is based on flow based Quality OF Service model designed for
 a. CPU
 b. Data Node
 c. IP
 d. Traffic Shaping
10. The token bucket can easily be implemented with a counter, initialized by
 a. 0
 b. 1
 c. -1
 d. -2
11. In Congestion, to define the maximum data rate of the traffic we use
 a. Average Data Packet
 b. Peak Data Rate
 c. Packet Data Rate
 d. Average Data Rate
12. In Congestion, the maximum burst size normally refers to the maximum length of time the traffic is generated at the
 a. Average Rate
 b. Packet Rate
 c. Protocol Rate
 d. Peak Rate

Self-Assessment Answer

1.	b	2.	a	3.	a	4.	a
5.	d	6.	d	7.	d	8.	a
9.	c	10.	a	11.	b	12.	d

Further Readings

Achyut S Godbole and Atul Kahate, *Web Technologies,* Tata McGraw Hill.

Andrew S. Tanenbaum, *Computer Networks,* Prentice Hall.

Behrouz A. Forouzan, Sophia Chung Fegan, *Data Communications and Networking,* McGraw-Hill Companies.

Douglas Comer, Computer Networks and Internets with Internet Applications, 4th Edition, Prentice Hall.

Ferguson P., Huston G., John Wiley & Sons, Inc., 1998. *Quality of Service: Delivering* QoS on the Internet and in Corporate Networks.

J. D. Spragins, Telecommunications Protocols and Design, Addison Wesley.

McDysan, David E. and Darren L. Spohn, *ATM Theory and Applications*, McGraw-Hill Osborne Media, 1998.

Nassar, Daniel J., *Ethernet and Token Ring Optimization*, iUniverse.com, 2000.

Spurgeon, Charles E. Ethernet, *The Definitive Guide.* O'Reilly & Associates, 2000.

William A Shay, Understanding Communication and Networks, 3rd Edition, Thomson Press.

Dr. Rajni Bhalla, Lovely Professional University

Unit 12: Application layer – services and protocols

CONTENTS

Objectives

- understand Telnet, FTP
- understand E-mail, POP, IMAP
- understand domain name system
- understand WWW and HTTP

Introduction

Terminal NETwork is the abbreviation for Terminal Network. It's a protocol that allows one device to communicate with another on the same network. It is a basic TCP/IP protocol for virtual terminal services that is provided by ISO. The local machine is the one that initiates the communication. The remote computer is the computer to which the connection is made, i.e. the computer that recognises the connection. When the link between the local and remote computers is created. During a telnet session, whatever is going on on the remote computer is viewed on the local computer. Telnet works on a client/server model. The telnet client software is used on the local computer, while the telnet server programme is used on the remote computer.

12.1 Telnet Commands

A prefix character, Interpret As Command (IAC), with the code 255, is used to identify telnet commands. After IAC, there are command and choice codes. The command's basic format is seen in the Table 1 below.

Table 1Telnet Command Format

IAC	Command Code	Option Code

The Table 2 are some of the most commonly used TELNET commands:

Table 2 Telnet Commands

Character	Decimal	Binary	Meaning
WILL	251	11111011	1. Offering to enable. 2. Accepting a request to enable.
WON'T	252	11111100	1. Rejecting a request to enable. 2. Offering to disable. 3. Accepting a request to disable.
DO	253	11111101	1. Approving a request to enable. 2. Requesting to enable.
DON'T	254	11111110	1. Disapproving a request to enable. 2. Approving an offer to disable. 3. Requesting to disable.

The Table 3 are some commonly used telnet options:

Table 3 Telnet Options

Code	Option	Meaning
0	Binary	This translates to an 8-bit binary transmission.
1	Echo	It will repeat the information obtained from one hand to the other.
3	Suppress go ahead	After info, it will switch off the go-ahead signal.
5	Status	It will inquire about TELNET's position.
6	Timing mark	It establishes the timing markers.
8	Line Width	It determines the width of the rows.
9	Page Size	It determines how many lines are on a list.
24	Terminal type	It was used to specify the terminal type.

32	Terminal Speed	It regulated the terminal's pace.
34	Line Mode	It will switch to line mode.

Operational Modes:

Most telnet implementation operates in one of the following three modes :

- Default mode
- Character mode
- Line mode

Default Mode

- This mode is used if none of the other modes are invoked.
- In this mode, the client performs the echoing.
- In this mode, the user types a character, and the client enchoes it on the screen, but it does not submit it until the whole line is over.

Character Mode

- In this mode, each character typed by the client is sent to the server.
- In this mode, the server usually enchoes the character to be viewed on the client's computer.

Line Mode

- Line editing like echoing, character erasing etc is done from the client side.
- Client will send the whole line to the server.

12.2 File Transfer Mode(FTP)

FTP (File Transfer Protocol) is an application layer protocol that allows you to transfer files between local and remote file systems. It, like HTTP, runs on top of TCP. FTP uses two TCP links in parallel to pass a file: a control link and a data connection as seen in Figure 1.

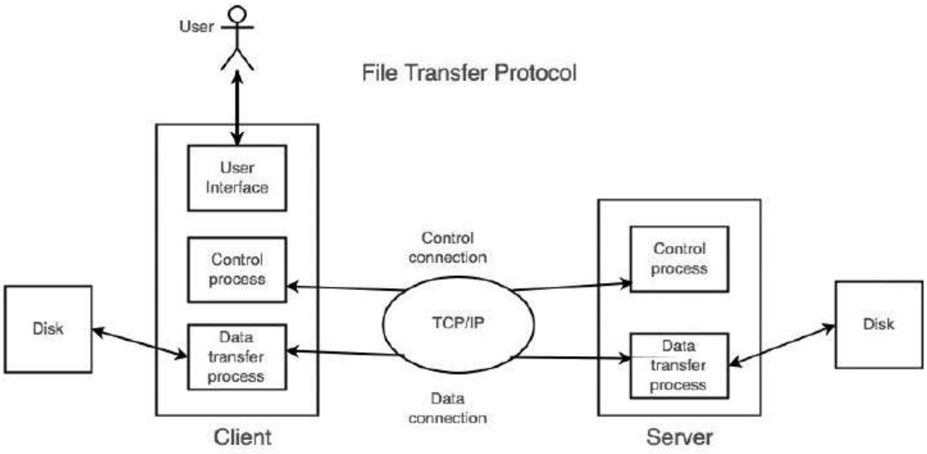


Figure 1 File Transfer Protocol

What is control connection?

FTP uses a control link to send control information such as user identity, passwords, commands to modify the remote directory, commands to retrieve and store files, and so on. On port number 21, the control link is created.

What is data connection?

FTP allows use of a data link to transfer the entire request. On port number 20, a data link is created.

Since FTP uses a different control link, the control information is sent out-of-band. Some protocols use the same TCP link to transmit their request and response header lines, as well as the data. They are said to give their control details in-band as a result of this. HTTP and SMTP are two examples of this.

FTP Session

When a client and server have an FTP session, the client establishes a control TCP connection with the server. Over this, the client sends control detail. When the server receives this information, it establishes a data link with the client. Over a single data link, only one file can be sent. The control link, on the other hand, is kept active during the user session. HTTP is stateless, which means it does not need to keep track of any user state. However, FTP must keep track of the user's status during the session.

Data Structures

FTP supports three different kinds of data structures:

File Structure – There is no internal structure of file-structure, and the file is treated as a single series of data bytes.

Record Structure – A file with a record structure is made up of consecutive entries.

Page Structure – A file with a page structure is made up of separate indexed files.

FTP Commands

Some FTP commands are as follows:

USER: The user identifier is sent to the server with this instruction.

PASS: The user password is sent to the server with this instruction.

CWD: This command helps the user to store or retrieve files in a separate directory or dataset without changing his username or accounting records.

RMD: The directory listed in the path-name is deleted as a directory with this instruction.

MKD: This command creates a directory in the directory defined by the pathname.

PWD: The name of the current working directory is returned in the response from this order.

RETR: This command instructs the remote host to establish a data link and transfer the requested file over it.

STOR: This order triggers a file to be saved in the remote host's current directory.

LIST: Sends a document for a list of all the files in the directory to be shown.

ABOR: This order instructs the server to terminate the previous FTP service command and all data transfers associated with it.

QUIT: This order terminates a USER, and the server closes the control connection if no file transfer is in progress.

FTP Responses – Some of the FTP responses are as follows:

200 Command okay.

530 Not logged in.

331 User name okay, need a password.

225 Data connection open; no transfer in progress.

221 Service closing control connection.

551 Requested action aborted: page type unknown.

502 Command not implemented.

503 Bad sequence of commands.

504 Command not implemented for that parameter.

Anonymous FTP

On certain pages with publicly accessible files, anonymous FTP is allowed. These files can be accessed without the use of a username or password. Instead, by contrast, the username is anonymous and the password is visitor. User access is severely restricted here. The user can be able to copy files but not move through folders, for example.

12.3 E-mail

Electronic mail (e-mail) is one of the most used Internet services. This service enables an Internet user to send a formatted message (mail) to another Internet user located anywhere in the world. Messages in the mail contain not only text, but also photographs, audio, and video files. The person who sends mail is known as the sender, and the person who collects mail is known as the receiver. It's similar to the postal mail service.

Components of E-Mail System:

The following are the main components of an email system: User Agent (UA), Message Transfer Agent (MTA), Mail Box, and Spool file. This are clarified in the following paragraphs:

User Agent(UA): The UA is usually a service that sends and receives mail. It is also known as a mail reader. It supports a wide range of commands for sending, receiving, and replying to messages, as well as manipulating mailboxes.

Message Transfer Agent(MTA):

MTA is in charge of transferring mail from one device to another. A device must have both a client MTA and a system MTA in order to send email. If the receivers are wired to the same machine, it sends mail to their mailboxes. If the destination mailbox is on another node, it sends mail to a peer MTA. Simple Mail Transfer Protocol is used to deliver messages from one MTA to another as seen in Figure 2.

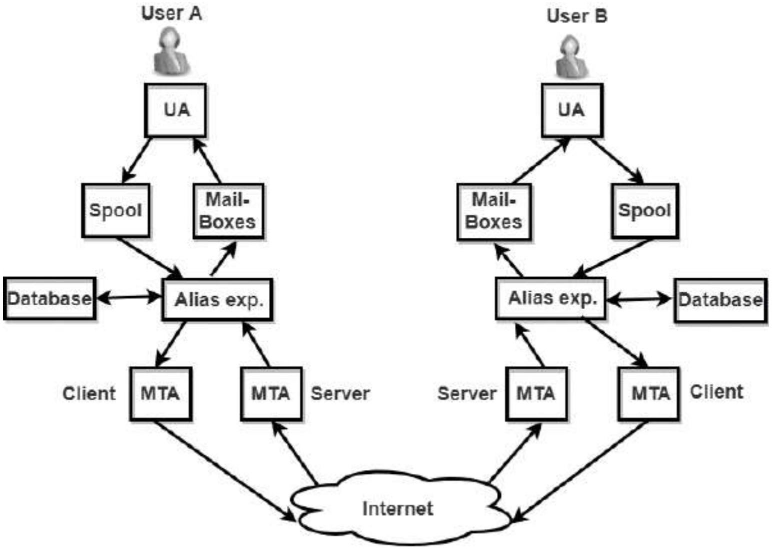


Figure 2 Message Transfer Agent

Mailbox

It's a local hard drive disc that collects e-mails. This file contains delivered emails. The user can read it or delete it depending on his or her needs. Each customer must have a mailbox in order to use the e-mail system. Just the mailbox user has access to the mailbox.

Spool File

This file holds all of the emails that need to be sent. SMTP is used by the user agent to append outgoing mails to this register. For distribution, MTA removes pending mail from the spool register. In e-mail, a single name, known as an alias, may be used to describe several e-mail addresses.

Whenever a user has to send a post, the machine checks the name of the receiver against the alias database. If a mailing list is present for a given alias, separate messages must be prepared and handed to MTA, one for each entry in the list. If no mailing list exists for the given alias, the name becomes the identifying address, and a single letter is sent to the mail transfer individual.

Services provided by E-mail system:

Composition:

The term "composition" refers to the method of creating messages and responses. Every kind of text editor may be used to compose.

Transfer

The sending of mail from the sender to the receiver is referred to as a transfer.

Reporting

The term "reporting" applies to the announcement of postal delivery. It allows users to see if their mail has been delivered, lost, or rejected.

Displaying

It applies to current mail in a format that the user can comprehend.

Disposition

This move is concerned with the recipient's behaviour after receiving mail, such as saving it, deleting it before reading it, or deleting it after reading it.

12.4 POP and IMAP3

Both POP3 (Post Office Protocol 3) and IMAP (Internet Message Access Protocol) are MAA (Message Accessing Agent) protocols that are used to retrieve messages from the mail server and deliver them to the recipient's device. Spam and virus filters are aware of all of these protocols. POP3 is more rigid and complex than IMAP.

Difference Between POP3 and IMAP: Table 4 Difference between POP3 and IMAP.

Table 4 Difference between POP3 and IMAP

Post Office Protocol	Internet message access protocol (IMAP)
POP is a straightforward protocol that allows you to download messages from your Inbox to your machine.	IMAP is a more sophisticated protocol that helps you to see all of the folders on the mail server.
On port 110, the POP server listens, and on port 995, the POP with SSL safe (POP3DS) server listens.	On port 143, the IMAP server listens, and on port 993, the IMAP with SSL secure(IMAPDS) server listens.
POP3 allows you to view your email from just one computer at a time.	Messages are accessible from a variety of platforms.
It is necessary to download the mail on the local machine in order to read it.	Before downloading the message, you can read it in half.
The recipient cannot organise his or her emails in the mail server's mailbox.	The emails can be organised directly on the mail server by the recipient.
On the mail server, the user cannot make, erase, or rename emails.	On the mail server, the user will build, erase, and rename emails.
Before uploading mail to the local system, the user cannot browse the content.	Until uploading, a user can scan the content of an email for a certain string.

There are two modes available: delete and keep. After retrieval, the mail is removed from the mailbox in delete mode. The mail is kept in the mail box after retrieval in keep mode.	Multiple backup backups of the letter are stored on the mail system, so that even if a local server's message is lost, the mail will still be recovered.
Local email tools can be used to make changes to the mail.	Online interface or email programme changes are synchronised with the server.
Many of the messages are downloaded at the same time.	Prior to uploading, the message header can be accessed.

12.5 Domain Name System(DNS)

DNS is a service that converts a host's name to an IP address. The Domain Name System (DNS) is a hierarchical network that is applied as a hierarchy of name servers. It's an application layer protocol that allows clients and servers to send and receive messages.

Requirement

A host is known by its IP address, but people have a hard time recalling numbers, and IP addresses are not static, so a mapping is needed to convert a domain name to an IP address. As a result, DNS is used to transform a website's domain name to a numerical IP address.

Domain

There are many types of DOMAIN:

1. ***Generic Domain***: All of these are common domains:.com (commercial),.edu (educational),.mil (military),.org (non-profit organisation), and.net (similar to commercial).
2. ***Country domain*** .in (india) .us .uk
3. ***Inverse domain*** if we want to know what is the domain name of the website. Ip to domain name mapping. So, DNS can provide both the mapping for example to find the ip addresses of geeksforgeeks.org then we have to type nslookup www.geeksforgeeks.org.

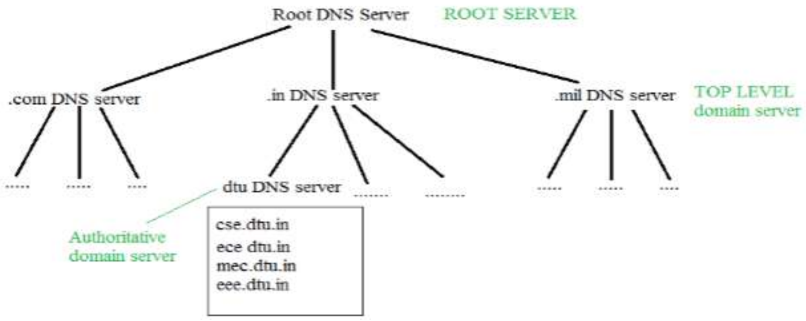


Figure 3 Organization of Domain

It is very difficult to determine the IP address associated with a website and there are millions of them. With all of those websites, we should be able to produce the IP address almost instantly; there should be no significant delay. Database organisation is critical.

DNS Record

What is the validity of a domain name and an IP address? What time is it to live? as well as other details pertaining to the domain name These documents are organised in a tree-like format.

Namespace

A list of titles that can be either flat or hierarchical. A naming scheme retains a list of name-to-value bindings – given a name, a resolution function returns the value that corresponds to it.

Name Server

It's a resolution process that's been put into action. DNS (Domain Name System) is an Internet name service. A zone is an administrative entity, and a domain is a subtree.

Name to Address Resolution

The host requests that the domain name be resolved by the DNS name server. The name server then returns to the host the IP address that corresponds to that domain name, allowing the host to bind to that IP address in the future.

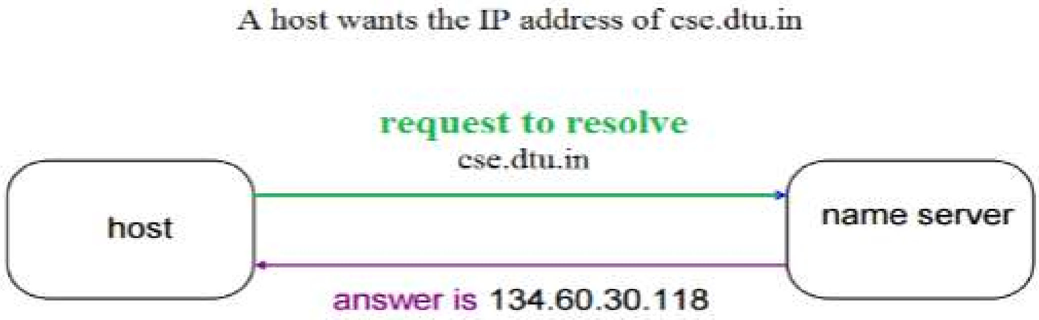


Figure 4 Name Address Resolution

Hierarchy of Name Servers

Root name servers: Name servers that are unable to address the name touch it. If the name mapping is unknown, it contacts the authoritative name server. The mapping is then obtained, and the IP address is returned to the host.

Top level server: It's in charge of com, org, and edu, as well as all top-level country domains like uk, fr, ca, and in. They have information on authoritative domain servers and know the names and IP addresses of each second-level domain's authoritative name server.

Authoritative name servers: This is the DNS registry for the enterprise, and it provides authoritative hostname to IP routing for the organization's servers. It can be kept up to date by a company or a service provider. To get to cse.dtu.in, we must first query the root DNS server, which will then direct us to the top-level domain server, and finally to the authoritative domain name server, which contains the IP address. As a result, the associative ip address will be returned by the authoritative domain server.

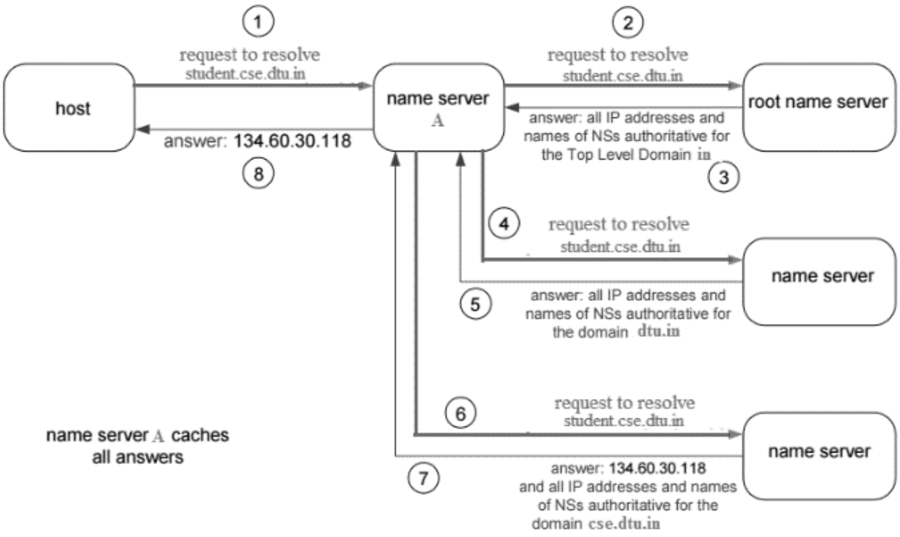

1
request to resolve
student.cse.dtu.in
host
answer: 134.60.30.118
8
name server
A
2
request to resolve
student.cse.dtu.in
root name server
answer: all IP addresses and names of NSs authoritative for the Top Level Domain in
3
4
request to resolve
student.cse.dtu.in
name server
5
answer: all IP addresses and names of NSs authoritative for the domain dtu.in
6
request to resolve
student.cse.dtu.in
name server
7
answer: 134.60.30.118 and all IP addresses and names of NSs authoritative for the domain cse.dtu.in
name server A caches all answers

Figure 5 Domain name server

The client computer sends a request to the local name server, which, if root cannot locate the address in its database, sends a request to the root name server, which then routes the request to an intermediate or authoritative name server. Any hostName to IP address mappings can also be found on the root name server. The definitive name server is still known to the intermediate name server. Finally, the IP address is returned to the local name server, which then forwards it to the host.

12.6 World Wide Web (WWW)

The World Wide Web, also known as the web, is abbreviated as WWW. In 1989, CERN (European Library for Nuclear Research) launched the World Wide Web.

History

It's a project started in 1989 by Timothy Berner's Lee to help CERN researchers collaborate more efficiently. The World Wide Web Consortium (W3C) is a non-profit organisation dedicated to furthering web growth. Tim Berners-Lee, dubbed the "Father of the Internet," is in charge of this organisation.

System Architecture

From the user's perspective, the internet is a large, global network of documentation or web sites. Each page can provide links to other sites on the internet. The pages can be retrieved and accessed using a variety of browsers, including Internet Explorer, Netscape Navigator, Google Chrome, and others. The browser retrieves the requested page, interprets the text and formatting commands on it, and shows the page on the screen, correctly formatted.

The fundamental model of how the internet operates is shown in the Figure 6 below. On the client computer, the browser is viewing a web address. When a user clicks on a line of text that links to a page on the abd.com site, the browser follows the hyperlink by requesting the page from the abd.com server.

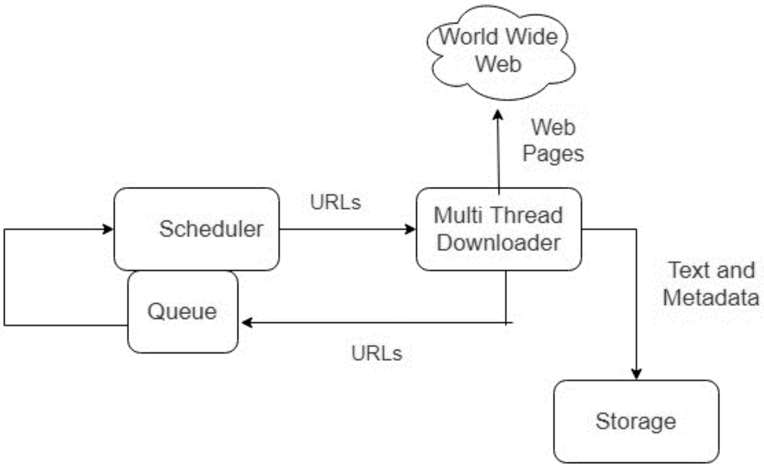


Figure 6 Basic model of web

The browser is now viewing a web page from the client computer. When a user clicks on a line of text that links to an abd.com website, the vbrowser supports the hyperlink by submitting a request to the abd.com server for the page.

Working of WWW

Web servers, Hypertext Markup Language (HTML), and Hypertext Transfer Protocol (HTTP) are among the technologies that make up the World Wide Web (HTTP).

To view webpages, you'll need a Web server. Web browsers are applications that use the Internet to view text, documents, images, animation, and video.

Web browsers provide a software interface for accessing hyperlinked resources on the World Wide Web. Initially, Web browsers were mainly used for browsing the Internet, but they have since been more widely used. Web browsers can be used for a variety of functions, including searching, mailing, and uploading files, among others. Internet Explorer, Opera Mini, and Google chrome are some of the most popular browsers.

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.
- “Web 2.0”

Components of Web

The site is made up of three parts:

Uniform Resource Locator (URL): serves as a web-based resource management system.

Hypertext Transfer Protocol (HTTP): specifies communication of browser and server.

Hyper Text Markup Language (HTML): defines structure, organization and content of webpage.

12.7 Hypertext transfer protocol

The HyperText Transfer Protocol (HTTP) stands for HyperText Transfer Protocol. Tim Berner is the one who came up with the idea. HyperText is a particular category of text that is coded using a common coding language known as HyperText Markup Language (HTML). HTTP/2 is the most recent version of HTTP, released in May 2015. HyperText Transmission Protocol is a collection of protocols for transferring hypertext between two computers.

HTTP establishes a networking protocol between a web browser and a web server. It's a series of guidelines for moving data from one device to another. On the World Wide Web, data such as text, photographs, and other multimedia files are exchanged. When a computer user opens their tab, they are inadvertently using HTTP. It's an implementation protocol for hypermedia knowledge systems that are distributed and interactive.

How it works

To begin, if we want to access a website, we must first open a web browser and then enter the website's URL (e.g., www.facebook.com). This URL has now been forwarded to the Domain Name Server (DNS). The DNS server will then search their cache for records for this URL, and then return an IP address to the web browser that corresponds to this URL. The browser will now send requests to the actual server as seen in Figure 7.

The link will be closed until the server has sent data to the device. If we want something different from the server, we must re-establish the relationship between the client and the server.

The connection will be ended once the server has sent data to the client. If we want anything different from the server, we must re-establish the connection between the client and the server.

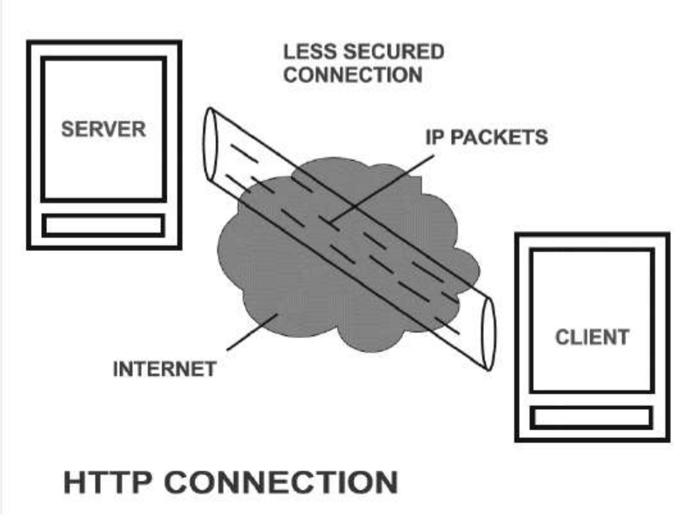


Figure 7 HTTP Connection

History

Tim Berners-Lee and his CERN team are credited with creating HTTP and related technologies.

1. HTTP version 0.9 –

This was the original version of HTTP, released in 1991.

2. HTTP version 1.0 –

RFC 1945 (Request For Comments) was included to HTTP version 1.0 in 1996.

3. HTTP version 1.1 –

RFC 2068 was introduced in HTTP version 1.1 in January 1997. In June 1999, RFC 2616 was published, which included improvements and modifications to the HTTP version 1.1 standard.

HTTP version 2.0 –

On May 14, 2015, the HTTP version 2.0 standard was published as RFC 7540.

HTTP version 3.0 –

Version 3.0 of HTTP is based on an earlier RFC draught. It has been renamed HyperText Transfer Protocol QUIC, which is a Google-developed transport layer network protocol.

Characteristics of HTTP:

- HTTP is an IP-based communication protocol for transferring data from a server to a client and vice versa:
- The server handles a client request, and the server and client are only aware of each other for the present request and response time.
- Any sort of data may be transmitted as long as the server and client are both compliant.
- Once data has been shared, the servers and clients are no longer linked.
- It's a client-server-based request-and-response protocol.
- It is a connection-less protocol because the server does not remember anything about the client when the connection is ended, and the client does not remember anything about the server.
- It's a stateless protocol since neither the client nor the server expects anything from the other, yet they can still interact.

Advantages

- Because there are fewer simultaneous connections, memory and CPU utilisation are minimal.
- Because there are fewer TCP connections, network congestion is reduced.
- Since handshaking is done at the beginning of the connection, latency is minimised because future requests do not require handshaking.
- Reports without disconnecting the connection might be the cause of the problem.
- HTTP provides request or response pipelining.

Disadvantages

- To establish communication and transport data, HTTP demands a lot of electricity.
- HTTP is less secure, because it does not employs any encryption mechanism like https utilises TLS to encrypt typical http requests and answer.
- HTTP is not suited for mobile devices and is excessively chatty.
- HTTP does not allow for true data exchange since it is insecure.
- Clients do not stop connections until they have received all of the data from the server; as a result, the server must wait for the data to be completed and is unavailable to new clients during this period.

Summary

- The Domain Name System (DNS) allows for rapid translation of IP address text from a directory of billions of addresses in a fraction of a second. Domain ideas, which employ hierarchical structures of text addresses translation, might make this possible. The servers that keep track of addresses are dispersed throughout the globe.
- To send data, HTTP employs the TCP transport service via sockets. The HTTP client establishes a TCP connection with the HTTP server by utilising sockets on port 80. The server responds to client queries with HTML pages and objects after accepting the connection from the client. HTML pages and other objects are therefore sent back and forth between the client browser and the web server.
- The World Wide Online, or Web, is an information system in which documents and other web resources are identified by Uniform Resource Locators (URLs), which may be connected together via hyperlinks, and are accessible over the Internet.

- Electronic mail is one of the most widely used network services, and it employs a user agent and a message transfer agent to transmit messages from a user's inbox to remote mailboxes. Websites have been given a new lease on life thanks to multimedia apps, which have made them more dynamic. The amalgamation of many media such as text, images, video, and sound into a single medium has made a significant contribution.
- The File Transfer Protocol (FTP) is a standard communication protocol for transferring computer files over a computer network from a server to a client. FTP is based on a client–server architecture, with the client and server having independent control and data connections.
- Electronic mail (often known as e-mail) is a technique of sending and receiving messages ("mail") between persons who use electronic equipment.

Keywords

Browser: A browser is a piece of software that your computer uses to access the Internet and view WWW content.

Domain Name System (DNS): It is responsible for defining the protocol that allows clients and servers to interact with one another. DNS allows a system to employ a resolver, which converts the host name to an IP address that the server can comprehend.

Electronic mail: It refers to the electronic form of postal mail that employs a user agent and a message transfer agent to deliver the message to the appropriate mailbox. Multimedia

HTTP: HTTP (Hypertext Transfer Protocol) is a network protocol that is used to visit any website.

World wide web (WWW): The World Wide Web (WWWW) is a system for displaying text, graphics, and audio that has been downloaded from the internet. The content and hyperlinks in a hypertext page are written in HyperText Markup Language (HTML), and the page is given an Internet address called a Uniform Resource Locator (URL) (URL).

E-mail: E-mail, in full electronic mail, messages transmitted and received by digital computers through a network.

FTP: Files are either uploaded or downloaded to the FTP server when you provide them over FTP. The files are transmitted from a personal computer to the server when you upload them. The files are transmitted from the server to your own computer when you download them.

POP: For transmitting messages from an e-mail server to an e-mail client, the post office protocol (POP) is the most widely used message request protocol on the Internet. POP3 is an e-mail protocol in which the client requests new messages from the server, and the server "pops" all new messages to the client.

IMAP: The Internet Message Access Protocol (IMAP) is an Internet standard protocol for retrieving email messages from a mail server via a TCP/IP connection by email clients. RFC 3501 is the standard that defines IMAP.

Review Questions

1. What is FTTP?
2. Write advantages and disadvantages of HTTP.
3. Write down components and features of WWW.
4. Write down most commonly used telnet commands.
5. Explain E-mail. Write down components of E-mail.
6. What is domain name system?
7. Write down hierarchy of DNS.
8. Write down difference between POP3 and IMAP.

Self-Assessment

1. The default connection type used by HTTP is _________
   a. Persistent
   b. Non-persistent
   c. Can be either persistent or non-persistent depending on connection request
   d. None of the mentioned
2. The HTTP request message is sent in _________ part of three-way handshake.
   a. First
   b. Second
   c. Third
   d. Fourth
3. The first line of HTTP request message is called _____________
   a. Request line
   b. Header line
   c. Status line
   d. Entity line
4. Dynamic web page ______
   a. is same every time whenever it displays
   b. generates on demand by a program or a request from browser
   c. both is same every time whenever it displays and generates on demand by a program or a request from browser
   d. is different always in a predefined order
5. What is a web browser?
   a. a program that can display a web page
   b. a program used to view html documents
   c. it enables user to access the resources of internet
   d. all of the mentioned
6. Expansion of FTP is __________
   a. Fine Transfer Protocol
   b. File Transfer Protocol
   c. First Transfer Protocol
   d. Fast Transfer Protocol
7. The data transfer mode of FTP, in which all the fragmenting has to be done by TCP is ________
   a. Stream mode
   b. Block mode
   c. Compressed mode
   d. Message mode
8. Which of these is not a medium for e-mail?
   a. Intranet
   b. Internet
   c. Extranet
   d. Paper
9. If the sender wants an option enabled by the receiver, it sends a ______ command.
   a) WILL
   b) DO

c) WONT

d) None of above

10. FTP uses the srvices of

a) UDP

b) IP

c) TCP

d) None of above

11. Which of these do not provide free E-mail?

a. Hotmail

b. Rediff

c. WhatsApp

d. Yahoo

12. Telnet protocol is used to establish a connection to __________

a. TCP port number 21

b. TCP port number 22

c. TCP port number 23

d. TCP port number 25

Answers

1.	a	2.	c	3.	a	4.	b
5.	d	6.	b	7.	a	8.	d
9.	b	10.	c	11.	c	12.	c

Further Readings

Andrew S. Tanenbaum, *Computer Networks,* Prentice Hall.

Behrouz A. Forouzan, Sophia Chung Fegan, *Data Communications and Networking,* McGraw-Hill Companies.

Burton, Bill, *Remote Access for Cisco Networks,* McGraw-Hill Osborne Media.

Dr. Rajni Bhalla, Lovely Professional University

Unit 13 : Internet and WWW

CONTENTS

Objectives

- understand IP security (IPsec)
- understanding Email Security
- understand VPN
- VPN Privacy
- difference between Paid VPN and Free VPN
- understand digital signature.
- understand digital certificate.

Introduction

13.1 IP security (IPSec)

IP security (IPSec) is an Internet Engineering Task Force (IETF) standard suite of protocols that offer data authentication, integrity, and secrecy between two communication points over an IP network. It also specifies how packets are encrypted, decrypted, and authenticated. It specifies the protocols for safe key exchange and key management.

Uses of IP Security:

The following are some of the things IPsec may be used for.

- To encrypt data at the application layer.
- To ensure that routers delivering routing data across the public internet are secure.
- To give authentication without encryption, such as confirming that data comes from a recognized sender.

- To safeguard network data by establishing circuits utilizing IPsec tunnelling, in which all data transported between the two endpoints is encrypted, similar to a VPN connection.

Components of IP Security

It consists of the following elements:

1. ***Encapsulating Security Payload (ESP) –***

   Data integrity, encryption, authentication, and anti-replay are all included. It also supports payload authentication.

2. ***Authentication Header (AH) –***

   It also has data integrity, authentication, and anti-replay capabilities, but no encryption. Anti-replay protection guards against unwanted packet transfer. It does not ensure the privacy of data.

IP HDR	AH	TCP	DATA

3. ***Internet Key Exchange –***

   It's a network security protocol that allows two devices to dynamically exchange encryption keys and discover a means to communicate across a Security Association (SA). To facilitate secure communication, the Security Association (SA) creates common security properties between two network entities. The Internet Security Association (ISAKMP) and the Key Management Protocol (ISAKMP) offer a framework for authentication and key exchange. ISAKMP describes how Security Associations (SAs) are created and how direct connections between two IPsec-enabled hosts are established.

   IKE (Internet Key Exchange) protects the content of messages and serves as an open framework for implementing common algorithms like SHA and MD5. Each packet is assigned a unique identification by the algorithm's IP sec users. The device may then determine if a packet is valid or not using this identification. Unauthorized packets are deleted and not delivered to the intended recipient.

IP HDR	TCP	DATA

Original Packet

IP HDR	ESR HDR	TCP	Data	ESP Trailer	ESP Authentication

←--------- Encryption ----→

←--------- Authentication ---------→

Working of IP Security

- The host determines whether or not the packet should be sent via IPsec. These packets activate the security policy on their own. When the system transmitting the packet uses adequate encryption, this is accomplished. The host additionally verifies whether or not the incoming packets are correctly encrypted.
- Then there was IKE. Phase 1 begins with the two hosts (using IPsec) authenticating themselves to begin a secure communication. It has two different modes.

- The Main mode, which provides additional security, and the Aggressive option, which allows the host to quickly construct an IPsec circuit.
- Now, IKE Phase 2 takes place via a secure channel, with the two hosts negotiating the sort of cryptographic algorithms to employ and agreeing on secret keying material to use with those methods.
- The data is then sent via the newly established IPsec encrypted tunnel. The hosts use IPsec SAs to encrypt and decode these packets.
- When the communication between the hosts is finished or the session expires, the IPsec tunnel is ended by both hosts discarding the keys.

13.2 Email-Security

Email hacking may be accomplished in a variety of ways. Whether it's a spam virus or phishing. What exactly is spam? Sending unwanted mass emails to those who have not requested for them is known as email spamming.

Spam

Spam is junk mail delivered by commercial businesses as an advertising for their products and services. For such mails, we don't have an address. As a result, they are infected with the spam virus. Some emails may contain files containing harmful script, which, when performed on your computer, will mostly damage your vital data.

Phishing email

Phishing is when someone sends an email to a user that claims to be authentic but isn't. Its main goal is to steal sensitive data such as usernames, passwords, and credit card numbers.

Search emails contain a link to a malware-infected website that directs the user to input information on a phoney website that looks and feels real. So, kids, keep in mind one very crucial point:

Spams can cause a slew of issues:

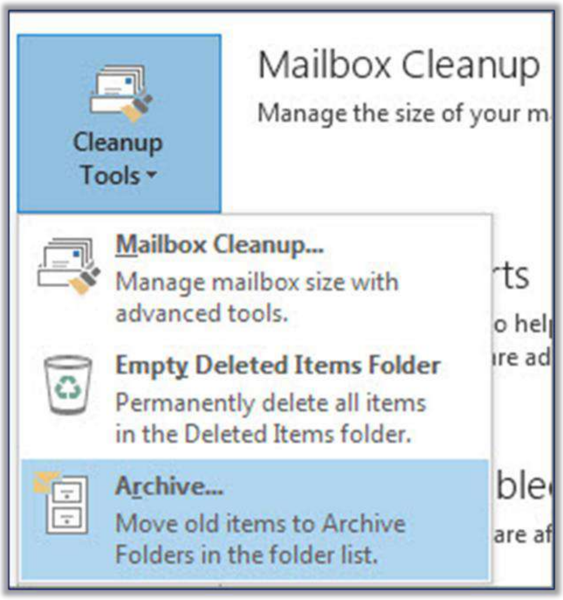
- It floods your email account with unwanted emails, which may result in the loss of information emails; it consumes bandwidth, slowing the speed at which mails are delivered; some unsolicited emails may include viruses, which may harm your computer; it costs time and energy in reviewing and deleting junk emails or spam; it consumes bandwidth, slowing the speed at which mails are delivered; and it spends time and energy in reviewing and deleting junk emails or spam; and it consumes bandwidth, slowing the speed at which mail.

Blocking Spams

Use a different email address for sending messages to a newsgroup or mailing list than the one you use for personal correspondence. Your email address should not be posted on the internet since it might be easily spammed. Avoid responding to emails from unknown senders, and never buy anything in response to spam that promises a product, since these are all variables that might lead to identity theft.

How to clear up and archive your emails

You may go to the File tab and find the cleaning tools there, and if you click on the drop-down arrow, you'll see the choice of archives. Choose the option to archive this folder and all subfolders. Then, select the folder you wish to archive by clicking on it as seen in Figure 1. Choose an older date from the archived items. a list Now, select Browse to name and save a new dot PST file, and then click OK. As a result, you will be able to clean up and archive in this manner.



Cleanup Tools ▾

Mailbox Cleanup

Manage the size of your m

Mailbox Cleanup...
Manage mailbox size with advanced tools.

Empty Deleted Items Folder
Permanently delete all items in the Deleted Items folder.

Archive...
Move old items to Archive Folders in the folder list.

Figure 1 Email Cleanup

13.3 Virtual Private Network (VPN)

A virtual private network is the full version of the term VPN. What is a home network? You have an internet connection that you will use either to connect to the internet or to connect to the internet using a model or a router. Let's look at how we are actually connecting to the internet.

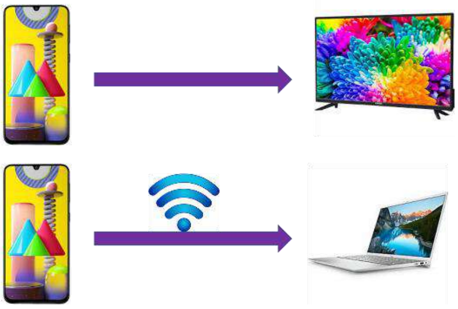


Figure 2 Internal Network

Assume we have a laptop, a television, and a cell phone. These three gadgets are present. Assume you want to send anything to your television, or you're working on your laptop. You're communicating by using a wireless connection that's connected to your phone and laptop as shown in Figure 2.

In a nutshell, these two types of networks are referred to as internal networks. This is your internal network if you're working on a laptop with a phone. Assume that any video you're seeing, such as the one you're seeing right now, is stored in a data centre. And that data centre may be in the United States, Canada, France, Germany, or somewhere else, so you could watch this movie from there. So any location implies you may access that movie; for example, if you are in India and want to view a film that is in Canada, Germany, or another country, you may do so. Let's look at Google as an example. Consider how they really request going to Google.com. Now you'll submit a request via a router, modem, or hub. Everyone has been given an address and will be working from that place.

If you wanted to open a google.com, you would send a request to the server, which would then be routed through the router. So the router has an address, and you can connect to the server, and the server responds to you, and you can sink the google.com domain.

What are the benefits of this sort of network in which IP addresses are routed to the server?

You will be able to connect to the Internet using this address.

As a result, the downsides include that your ISP service provider can track which websites you visit and what you track. Because your IP address is known by internet service providers. It's possible that your government is keeping an eye on you as well. When you're receiving anything from the server or uploading anything to the server, your data might be stolen. As a result, it's conceivable that your information will be stolen. We're relieved that we were able to connect to the open Wi-Fi networks. Your data may be encrypted at times. Hackers can access your data at any moment. A lot of stuff on the internet is confined to a specific region or nation. Assume you're sitting in India and you notice a video that's available in the United States or Canada, and you want to watch it. And such videos are banned on the internet, so you won't be able to see them if you're in a different nation.

In China, for example, you cannot access the Facebook website. You won't be able to register a Facebook account since China has already blocked it as seen in Figure 3. What is the greatest solution if you wish to utilise the Facebook site? As a result, we have a solution in the form of a VPN.

Figure 3 Banned in China

You may access any other site or programme that is on separate places on your machine by utilising a virtual private network. This is where you'll find the VPN box, and anything you're sending will be secured as well. Neither your internet service provider nor the government have any idea what you're talking about. If you make a request, the server assumes that you are both from the same country. It signifies you are receiving the response in accordance with the request. If someone tried to sniff your packet, it would be impossible for them to read it since it is encrypted. You can utilise content from another region as well as any website that is written in your area. You may also view using a virtual private network. It means you'll be able to access and utilise all information safely.

Definition of VPN

A virtual private network (VPN) is a sort of private network that communicates via public telecommunications, such as the Internet, rather than leased lines. It grew in popularity as more people worked in remote places and needed to learn how VPNs function.

Let's understand VPN by an example:

Consider a scenario in which a bank's headquarters are in Washington, D.C. This workplace has a local network with about 100 machines on it. Assume the bank has additional branches in Mumbai, India, and Tokyo, Japan. The old technique of creating a secure connection between the head office and the branch was to use a leased line between the branches and the head office, which was both expensive and inconvenient. We were able to effectively solve this problem thanks to VPN.

Some fantastic attributes are highlighted here:

- VPN also guarantees security by establishing a secure connection between the client and the VPN server.
- Many banned websites may be accessed with a VPN.

- VPN allows you to browse anonymously by masking your IP address.
- Also, the best Search engine optimization (SEO) is done by evaluating data from VPN providers that give country-by-country statistics on how people are using a given product. Many online marketing managers utilise this strategy of SEO to develop fresh ideas.

VPN and its legality

In most nations, using a VPN is legal. The legality of utilising a VPN service is determined by the country in question, as well as its geopolitical connections with other countries. A trustworthy and secure VPN is always lawful if you're not planning on using it for any unlawful purposes, such as online fraud, cyber theft, or downloading copyrighted information in some countries.

According to Bloomberg, China has determined to shut all VPNs (virtual private networks) starting next year. VPNs are used by many Chinese Internet users to access websites that are prohibited by China's so-called "great firewall." This is done to prevent information from being leaked to competing governments and to improve information security.

13.4 Difference between Free VPN and Paid VPN

The difference between Free VPN and Paid VPN is shown in Table 1.

Table 1 Difference between Paid VPN and Free VPN

	Free VPN	**Paid VPN**
Cost	Free	From USD 7 per month
Number of VPN Servers	Generally, up to 5-7	Generally, more than 40
Advertising	Yes	No
Connection speed	Low	High
Monthly Traffic Volume	Limited	Not Limited
Technical Support	No support or it is too low	Quick support, 24/7
Security gurantees	No	Yes
Personal servers	Never	Some developers provide
No logs policy	No guarantees	Yes

13.5 Digital Signature

Encryption: Process of transforming electronic data into cypher text, which is incomprehensible to everyone but authorised individuals. This ensures the safety of the data.

Decryption: The procedure for converting code into data.

- The message is encrypted at the sender's end using various encryption techniques, and then decoded at the receiver's end using decryption methods.
- Encryption and decryption techniques are employed to ensure data security when certain messages, such as usernames and passwords, must be kept secure.

Types of Encryption

1. ***Symmetric Encryption***: A key is used to encrypt data, and the same key is used to decode it.

2. ***Asymmetric Encryption***: Public key cryptography is another name for asymmetric cryptography. It encrypts and decrypts data using public and private keys. The public key is the one of the pair of keys that may be shared with anybody. The private key is the other key in the pair that is kept secret and only known by the owner. A message can be encrypted using one of the keys; The decryption key is the polar opposite of the one used to encrypt the communication.

Public Key

Everyone is aware of the key. The ex-public key of A is 7, and this fact is well known.

Private Key

Only the person whose private key it is is aware of it.

Authentication

Any procedure by which a system confirms the identity of a person who seeks to access it is known as authentication.

Non-repudiation

Non-repudiation refers to the process of ensuring that a transmitted communication was transmitted and received by the parties claiming to have transmitted and received it. Non-repudiation is a method of ensuring that the sender of a communication cannot later dispute sending it, and that the recipient cannot later dispute receiving it.

Integrity

to be sure the message wasn't tampered with during transmission

Message Digest

The representation of text as a single string of numbers, achieved with the use of a formula known as a one-way hash function. A digital signature is an electronic mechanism of authentication that is created by encrypting a message digest with a private key.

13.6 Digital Signature

A digital signature is a mathematical approach for verifying the integrity and validity of a communication, software, or digital document.

Key Generation Algorithms

Digital signatures are digital signatures that guarantee that a communication was transmitted by a certain sender. Authenticity and integrity should be ensured while conducting digital transactions; otherwise, data might be tampered with, or someone might pretend to be the sender and expect a response.

Signing Algorithms

Signing algorithms, such as email programmes, establish a one-way hash of the electronic data to be signed to establish a digital signature. The hash value is subsequently encrypted using the private key by the signing algorithm (signature key). The digital signature consists of this encrypted hash, as well as other information such as the hashing algorithm.

The data is appended with the digital signature and delivered to the verifier. Because a hash function turns any arbitrary input into a much shorter fixed length result, it's better to encrypt the hash rather than the full message or document. This saves time since a shorter hash value must now be signed instead of a big message, and hashing is significantly quicker than signing.

Signature Verification Algorithms

Along with the data, the Verifier obtains a Digital Signature. It then processes the digital signature and the public key (verification key) using the Verification algorithm and generates a value. It also creates a hash value by applying the same hash algorithm to the incoming data. The hash value and the verification algorithm result are then compared. If they're both equal, the digital signature is genuine; otherwise, it's not.

The steps for establishing a digital signature are as follows:

- Message digest is computed by applying hash function on the message and then message digest is encrypted using private key of sender to form the digital signature. (digital signature = encryption (private key of sender, message digest) and message digest = message digest algorithm(message)).
- The message is then sent with a digital signature. (A message with a digital signature is sent).
- The digital signature is decrypted by the receiver using the sender's public key. (As only the sender possesses his private key, only the sender may encrypt using his private key, which can then be decrypted by the sender's public key, ensuring authenticity.)
- The message digest has now been delivered to the recipient.
- The message digest can be computed by the receiver from the message (actual message is sent with the digital signature).
- For integrity, the message digest calculated by the receiver and the message digest (obtained by decrypting a digital signature) must be the same.
- The message digest is calculated using a one-way hash function, which is a hash function in which computing the hash value of a message is simple but computing the message from the hash value is complex.

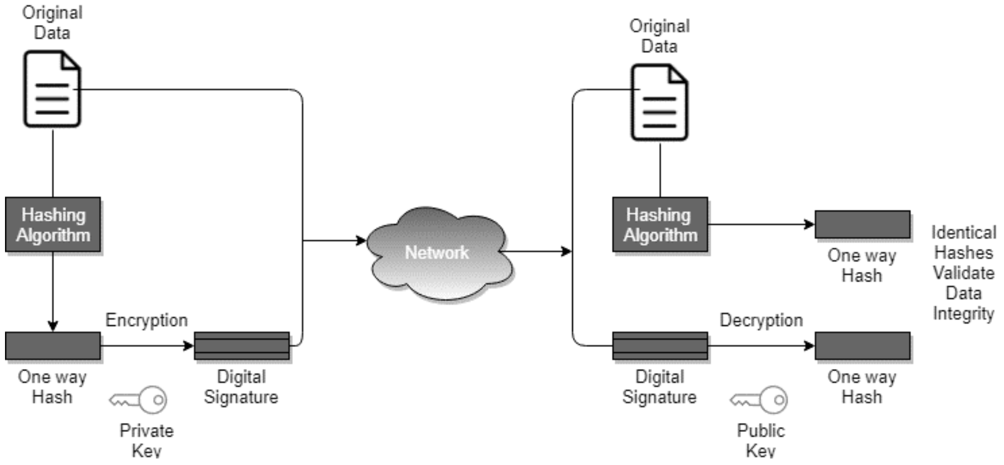


Figure 4 Digital Certificate

A trustworthy third party issues a digital certificate that validates the sender's identity to the receiver and the recipient's identity to the sender.

A certificate issued by a Certificate Authority (CA) to validate the identity of the certificate holder is known as a digital certificate. The CA creates an encrypted digital certificate that includes the applicant's public key and other identifying information. A digital certificate is used to associate a public key with a specific person or business.

Digital Certificate contains

1. Name of certificate holder.
2. Serial number which is used to uniquely identify a certificate, the individual or the entity identified by the certificate
3. Expiration dates.
4. Copy of certificate holder's public key.(used for decrypting messages and digital signatures)
5. Digital Signature of the certificate issuing authority.

Digital certificate is also sent with the digital signature and the message.

13.7 Digital certificate vs Digital signature

The difference between digital signature and digital certificate is shown in Table 2.

Table 2 Digital signature vs Digital certificate

Digital signature	Digital certificate
A digital signature is a string of decimals that is affixed to a file to assist with identifying the signer and ensuring its integrity	A digital certificate is itself a file that is used to assert identity and to facilitate encrypted connections.
Digital signature must ensure that the data or information remains secure from the moment it is sent	Digital certificate is that the certificate binds the digital signature to the object
Digital signatures are used to validate the sent data.	Digital certificates are used to validate the identity of the sender,
It follows Digital Signature Standard (DSS).	It follows X.509 Standard Format

Summary

- Internet Protocol Security (IPsec) is a secure network protocol suite that authenticates and encrypts data packets to offer safe encrypted communication between two computers over an IP network. It's a protocol that's utilised in virtual private networks.
- Any technique that secures email content and accounts from unwanted access is referred to as email security. Email security procedures are in place at email service providers to protect customer accounts and information from hackers.
- A virtual private network (VPN) creates a secure connection between you and the internet. All of your data traffic is routed over an encrypted virtual tunnel via the VPN. When you access the internet, this masks your IP address, making its location opaque to everyone. External assaults are also protected by a VPN connection.
- A digital signature is a mathematical approach for verifying the integrity and validity of a communication, software, or digital document. It's the digital counterpart of a handwritten signature or a stamped seal, but it has a lot more security built in.
- A public key certificate is an electronic document used to confirm the ownership of a public key in cryptography. It is also known as a digital certificate or an identity certificate.
- Encryption is the process of encoding data in cryptography. This technique transforms plaintext into ciphertext, which is an alternate representation of the information.
- Decryption is the process of restoring encrypted data to its original state. In most cases, it's a reversal of the encryption process. Because decryption requires a secret key or password, it decodes the encrypted information so that only an authorised user may decrypt the data.
- Spamming is the practise of sending many unsolicited messages to a large number of people for the purpose of commercial advertising, non-commercial preaching, or any other unlawful purpose, or just sending the same message to the same user over and over again.

Keywords

IP security: In IPv4 and IPv6 network packets, the IP security architecture (IPsec) offers cryptographic protection for IP datagrams.

internet key exchange: The mechanism used to establish up a security association in the IPsec protocol suite is known as Internet Key Exchange (IKE).

Email security: Any technique that secures email content and accounts from unwanted access is referred to as email security. Email security procedures are in place at email service providers to protect customer accounts and information from hackers.

Phishing: Phishing is a type of cybercrime in which a person acting as a genuine organisation contacts a target or targets by email, phone, or text message to persuade them to provide sensitive data such as personally identifying information, banking and credit card information, and passwords.

VPN: The term "virtual private network" refers to the ability to create a secure network connection while using public networks.

Digital signature: A digital signature is exactly what it sounds like: an electronic version of the traditional paper and pen signature.

Encryption: Encryption is the process of converting data into a secret code that hides the real meaning of the data.

Decryption: Decryption is the process of turning encoded/encrypted data into a form that a person or a machine can read and understand.

Public key: A public key is a huge numerical number used to encrypt data in cryptography. A software programme can produce the key, but it's more common for it to be issued by a trusted, recognised authority and made public through a publicly accessible repository or directory.

Private key: The owners keep their private keys hidden. Public keys are disseminated and used to validate credentials and authenticate nodes.

Authentication: The process of ascertaining if someone or something is who or what it claims to be is known as authentication.

Review Questions

1. Difference between Free VPN and paid VPN.
2. What is IP Security?
3. What is E-mail security?
4. Explain components of IP Security.
5. Write down steps for establishing a digital signature.
6. Write down difference between digital signature and digital certificate.
7. Explain types of signature.
8. What are the benefits of this sort of network in which IP addresses are routed to the server?
9. Explain working of IP Security
10. Difference between private key and public key.

Self- Assessment

1. IPSec is designed to provide security at the _________
   a) Transport layer
   b) Network layer
   c) Application layer
   d) Session layer
2. In tunnel mode, IPSec protects the ______
   a) Entire IP packet

b) IP header

c) IP payload

d) IP trailer

3. Which component is included in IP security?

a) Authentication Header (AH)

b) Encapsulating Security Payload (ESP)

c) Internet key Exchange (IKE)

d) All of the mentioned

4. Which of them is not a major way of stealing email information?

a) Stealing cookies

b) Reverse Engineering

c) Password Phishing

d) Social Engineering

5. The process of transforming plain text into unreadable text.

a) Decryption

b) Encryption

c) Network Security

d) Information Hiding

6. A process of making the encrypted text readable again.

a) Decryption

b) Encryption

c) Network Security

d) Information Hiding

7. A digital signature is a mathematical technique which validates?

a) authenticity

b) integrity

c) Non-repudiation

d) All of the above

8. Which algorithm algorithm provides the private key and its corresponding public key?

a) Key generation algorithm

b) Signature verifying algorithm

c) Signing algorithm

d) None of the above

9. The field that covers a variety of computer networks, both public and private, that are used in everyday jobs.

a) Artificial Intelligence

b) ML

c) Network Security

d) IT

10. The process of verifying the identity of a user.

a) Authentication

b) Identification

c) Validation

d) Verification

11. An algorithm in encryption is called _____________

a) Algorithm
b) Procedure
c) Cipher
d) Module

12. The information that gets transformed in encryption is ____________

a) Plain text
b) Parallel text
c) Encrypted text
d) Decrypted text

Answers

1(b)	2(a)	3(d)	4(b)
5(b)	6(a)	7(d)	8(a)
9(c)	10(a)	11(c)	12(a)

Further Readings

Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall

Behrouz A. Forouzan and Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies

Burton, Bill, *Remote Access for Cisco Networks*, McGraw-Hill, Osborne Media

Dale Tesch/Greg Abelar, Security Threat Mitigation and Response: Understanding *CS-MARS*, Cisco Press, Sep. 26, 2006.

Gary Halleen/Greg Kellogg, Security Monitoring with Cisco Security MARS, Cisco Press, Jul. 6, 2007.

Unit 14: Network Security

CONTENTS

Objectives

- understand basic of network security.
- understand network security issues
- learn security goals
- understand security services
- approaches of network security

Introduction

To begin, let's define some fundamental concepts in network security, such as plaintext, ciphertext, encryption, decryption, cryptography, and cryptanalyst as seen in Figure 1. And there's the key. So, what do all of them imply? I'll explain you what they imply in a very easy way so you can comprehend what plaintext ciphertext encryption decryption means.

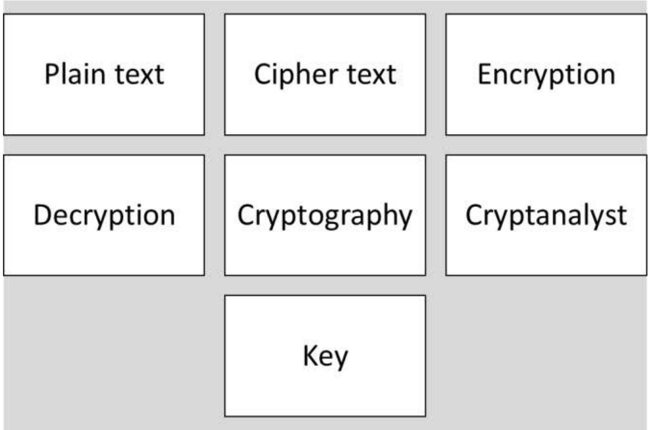


Figure 1 Basics of Network Security

14.1 Basics of Network security terms

A mono alphabetic substitution is a type of keyword encryption. The letter matchings of the cypher alphabet to the plain alphabet are determined by a keyword that serves as the key.The cypher alphabet is formed with the keyword matching to A, B, C, and so on until the keyword is exhausted, at which point the remaining cipher text letters are utilised in alphabetical sequence, omitting those previously used in the key.

Encryption

The first line of input includes the keyword you want to use. The string you must encrypt is on the second line of input.

Plain text : A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Encrypted : K R Y P T O S A B C D E F G H I J L M N Q U V W X Z

All As become Ks, all Bs become Rs, and so on when the keyword is KRYPTOS. Using the keyword "kryptos" to encrypt the message "knowledge is power."

Encrypting the message: Knowledge is Power
Encoded message: IlmWjbaEbgqNmWbp

Few points to be noted in this method:

- All of the messages are written in capital letters.
- Although you can use whitespace, special characters, and digits in your keyword, they will not be considered.
- Whitespace, special characters, and integers are unaffected by the encryption process.

Decryption

To decode the message, compare the provided message's location in the encrypted text to the plain text.

Plain text : A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Encrypted : K R Y P T O S A B C D E F G H I J L M N Q U V W X Z

Message: PTYBIATLEP
Deciphered Text : DECIPHERED

Now, how we generate the deciphered string?
We search for ‘P’ in Encrypted Text and compare its position with plain text letter and generate that letter. So ‘P’ becomes ‘D’, ‘T’ becomes ‘E’, ‘Y’ becomes ‘C’ and so on.

Cryptography

Cryptography is a method of safeguarding information and communications by encoding it in a way that only the people who need to know can interpret and interpret it. As a result, unwanted access to information is prevented. The word "crypt" stands for "hidden," and the suffix "graphy" stands for "writing."The procedures used to safeguard information in cryptography are derived from mathematical principles and a set of rule-based calculations known as algorithms that change signals in ways that make them difficult to decipher.These algorithms are used to generate cryptographic keys, digitally sign documents, verify data privacy, browse the internet, and secure sensitive transactions like as credit card and debit card transactions.

Techniques used For Cryptography:

In today's computer era, cryptography is generally connected with the process of converting plain text to cypher text, which is text encoded in such a way that only the intended receiver of the information can decode it, a process known as encryption. Decryption refers to the process of converting encrypted text to plain text.

Features of Cryptography are as follows:

Confidentiality: Only the individual for whom the information is meant has access to it, and no one else has access to it.

Integrity: Information cannot be changed in storage or in transit between the sender and the intended receiver without being noticed.

Non-repudiation: The information creator/sender cannot dispute that he or she intends to convey information at a later time.

Authentication: The sender's and receiver's identities have been verified. The information's destination/origin is also confirmed.

Types of cryptography

There are three forms of cryptography in general:

1. *Symmetric Key Cryptography*

It is an encryption scheme in which the sender and receiver of a message encrypt and decode messages using a single shared key. Symmetric Key Systems are quicker and easier to use, but they have the drawback of requiring the sender and receiver to exchange keys in a safe way. Data Encryption Technology is the most widely used symmetric key encryption system (DES).

2. *Hash functions*

This algorithm does not make use of any keys. A hash value with a defined length is computed based on the plain text, making it difficult to reconstruct the plain text's contents. Hash functions are used by several operating systems to encrypt passwords.

3. *Asymmetric key cryptography*

A pair of keys is used to encrypt and decode data in this system. Encryption is done using a public key, while decryption is done with a private key. The terms "public key" and "private key" are not interchangeable. Even though everyone knows the public key, the intended receiver can only decode it since he is the only one who knows the private key.

Cryptanalyst

The decoding and examination of codes, cyphers, or encrypted text is known as cryptanalysis. Cryptanalysis is a method of searching for algorithm flaws and breaking into cryptography or information security systems using mathematical formulae.

The following are examples of cryptanalysis attacks:

Known-Plaintext Analysis (KPA): With knowing partial plaintext, the attacker decrypts ciphertexts.

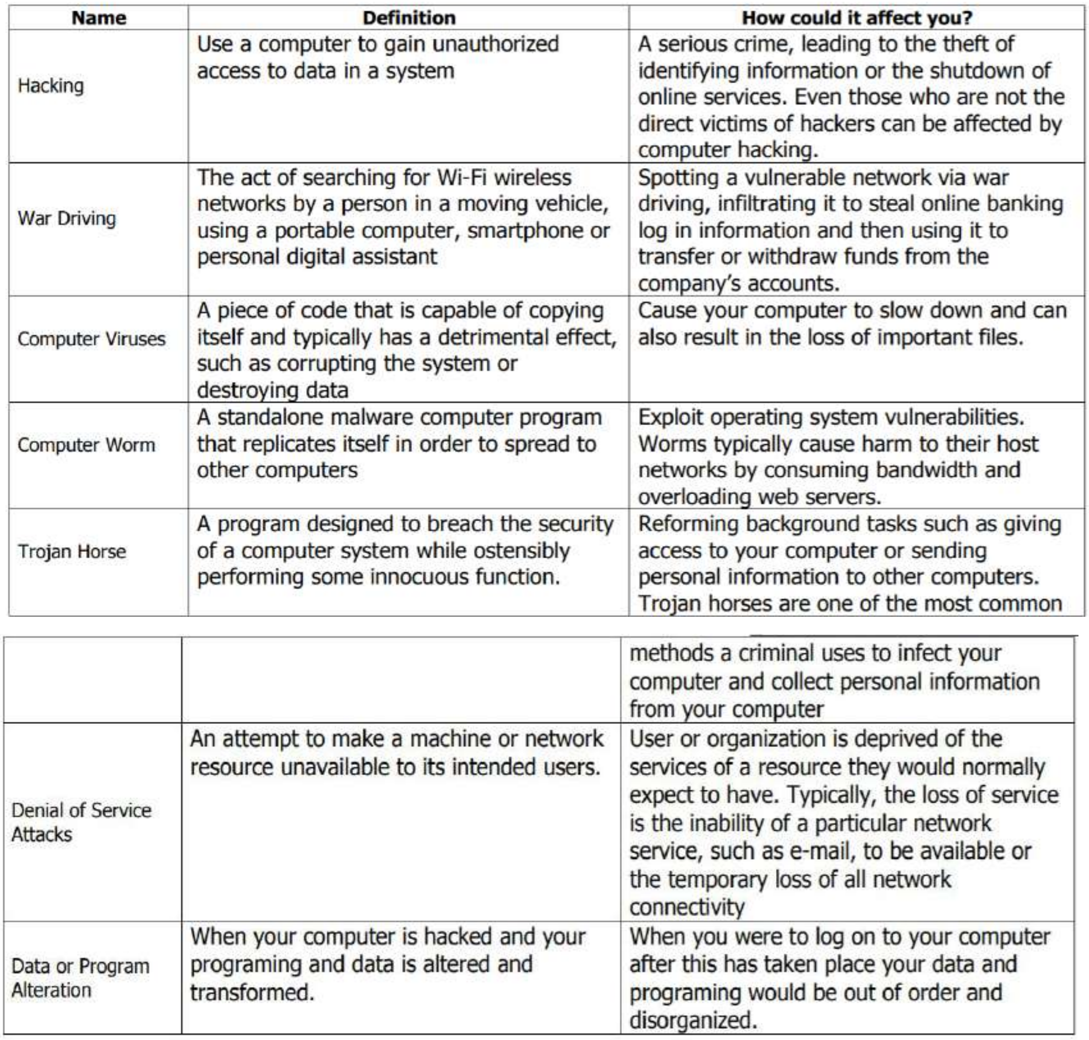
Chosen-Plaintext Analysis (CPA):Using the same mathematical methodology, the attacker creates ciphertext that matches any picked plaintext.

Ciphertext-Only Analysis (COA): The attacker employs ciphertext collections that are well-known.

Man-in-the-Middle (MITM) Attack: When two parties utilise message or key sharing to communicate through a channel that appears to be secure but is really compromised, an attack occurs. This technique is used by the attacker to intercept messages as they move across the communications channel. MITM attacks are avoided by using hash functions.

Adaptive Chosen-Plaintext Attack (ACPA):This attack employs selected plaintext and ciphertext depending on data learnt from previous encryptions, similar to a CPA.

14.2 Network Security Issues



Name	Definition	How could it affect you?
Hacking	Use a computer to gain unauthorized access to data in a system	A serious crime, leading to the theft of identifying information or the shutdown of online services. Even those who are not the direct victims of hackers can be affected by computer hacking.
War Driving	The act of searching for Wi-Fi wireless networks by a person in a moving vehicle, using a portable computer, smartphone or personal digital assistant	Spotting a vulnerable network via war driving, infiltrating it to steal online banking log in information and then using it to transfer or withdraw funds from the company's accounts.
Computer Viruses	A piece of code that is capable of copying itself and typically has a detrimental effect, such as corrupting the system or destroying data	Cause your computer to slow down and can also result in the loss of important files.
Computer Worm	A standalone malware computer program that replicates itself in order to spread to other computers	Exploit operating system vulnerabilities. Worms typically cause harm to their host networks by consuming bandwidth and overloading web servers.
Trojan Horse	A program designed to breach the security of a computer system while ostensibly performing some innocuous function.	Reforming background tasks such as giving access to your computer or sending personal information to other computers. Trojan horses are one of the most common methods a criminal uses to infect your computer and collect personal information from your computer
Denial of Service Attacks	An attempt to make a machine or network resource unavailable to its intended users.	User or organization is deprived of the services of a resource they would normally expect to have. Typically, the loss of service is the inability of a particular network service, such as e-mail, to be available or the temporary loss of all network connectivity
Data or Program Alteration	When your computer is hacked and your programing and data is altered and transformed.	When you were to log on to your computer after this has taken place your data and programing would be out of order and disorganized.

14.3 Security Goals

Data is extremely vulnerable to assaults during transmission. To fulfil his malicious goals, an attacker can target the communication channel, collect the data, and read or re-insert a bogus message.Network security is concerned not only with the security of the computers at each end of the communication chain, but also with the security of the entire network.

▪ C.I.A
Integrity
Confidentiality
Availability

Figure 2 Security Goals

Protecting the usefulness, dependability, integrity, and safety of a network and its data is what network security is all about. Effective network security prevents a wide range of dangers from infiltrating and propagating throughout a network.

Confidentiality, Integrity, and Availability are the three main goals of network security as seen in Figure 2. The CIA triangle is widely used to depict these three pillars of network security as seen in Figure 3.

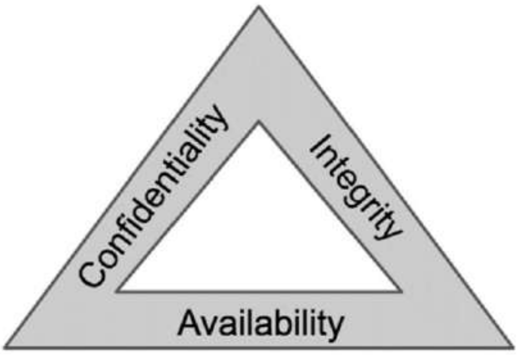


Figure 3 CIA Triangle

Confidentiality: The purpose of secrecy is to keep sensitive company information safe from prying eyes. The confidentiality component of network security ensures that data is only accessible to those who are authorised to see it.

Integrity:This aim entails ensuring and maintaining data correctness and consistency. The purpose of integrity is to ensure that data is accurate and not tampered with by unauthorised individuals.

Availability: The purpose of availability in Network Security is to ensure that data, network resources, and services are always available to authorised users when they need them.

14.4 Security Services

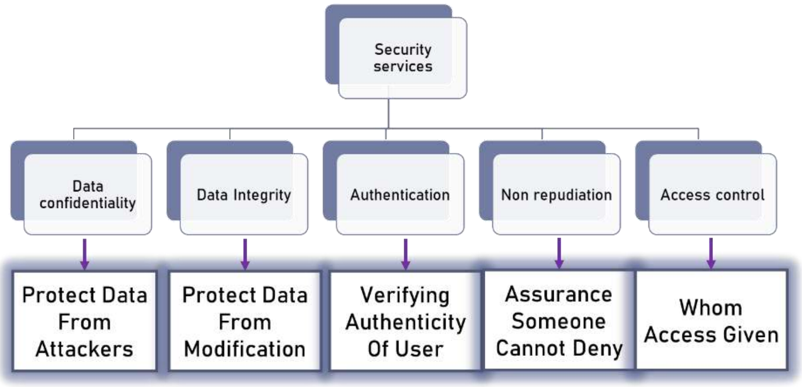

Security services
Data confidentiality
Data Integrity
Authentication
Non repudiation
Access control
Protect Data From Attackers
Protect Data From Modification
Verifying Authenticity Of User
Assurance Someone Cannot Deny
Whom Access Given

Figure 4 Security Services

Data confidentiality

The safeguarding of data from unlawful disclosure.

Connection Confidentiality: The protection of all user data on a connection. Connectionless **Confidentiality:** The protection of all user data in a single data block.

Selective-Field Confidentiality: The confidentiality of selected fields within the user Data on a connection or in a single data block.

Traffic Flow Confidentiality: The protection of the information that might be Derived from observation of traffic flows.

Data Integrity

The assurance that data received is identical to what was sent by an entity.authorised entity (i.e., no changes, additions, or deletions)or a rerun).

Authentication

The act of verifying a claim, such as a computer system user's identity, is known as authentication. Authentication is the process of validating a person's or thing's identity, as opposed to identification, which is the act of indicating that person's or thing's identity.

Non-repudiation

The term "non-repudiation" refers to a circumstance in which the author of a statement is unable to effectively contest the statement's authorship or the validity of an associated contract. When the legitimacy of a signature is being questioned in court, the word is frequently used. The genuineness is "repudiated" in this situation.

Access control

Individuals are authenticated and authorised to access the information they are permitted to see and use through access restrictions.

14.5 Approaches to network security

The Castillo of San Felipe de Barajas in Cartagena is the Spanish colonies' largest and most powerful stronghold. The fort, which was built between 1639 and 1657 and subsequently expanded significantly in 1762, defended the "portal to the new world" against several assault and was never captured.A fleet of Spanish ships patrolling in concentric "rings of defence" would distract or distract any opponent approaching the stronghold, preventing any threat to the citadel. If the adversary were to battle through these rings, the opponent would be protected from cannon fire and ground attack by the high, thick walls, while a sequence of batteries and parapets placed to cover each other would eliminate attack.

The fortress's vast maze of tunnels deep within its core is perhaps the most fascinating aspect. If the fort was ever taken, the defenders would escape into this maze, reorganise, and retake the citadel from within.These multiple levels of defence, though not a perfect parallel, highlight three unique approaches to security. Preventing a danger from forming in the first place, particularly through addressing its underlying causes, is one security strategy.

When a hazard cannot be avoided, security as protection seeks to mitigate, if not completely eradicate, the hazard. However, if we are unable to totally shield ourselves from the danger, security as resilience analyses our capacity to "bounce back" and change the manner in which it impacts our social systems - our capacity to adapt to genuine dangers.

What they target: the fundamental causes of threats, the threats themselves, or the qualities of the referent that is endangered is the essential component that distinguishes the three methods. However, as shown in the table below, the three techniques can be separated in a variety of additional ways.

Table 1 Approaches to network security

Security Approach:	Security as Prevention	Security as Control (Protection)	Security as Resilience
Definition	This strategy aims to prevent dangers from forming in the first place by addressing the root issues that lead to their emergence.	This strategy aims to manage, defend against, or remove a clear threat.	Security as resilience focuses on the ability of social systems to "bounce back" and recover from shocks when dangers cannot be managed or removed.It is about a society's flexibility and adaptation, as well as its rigidities

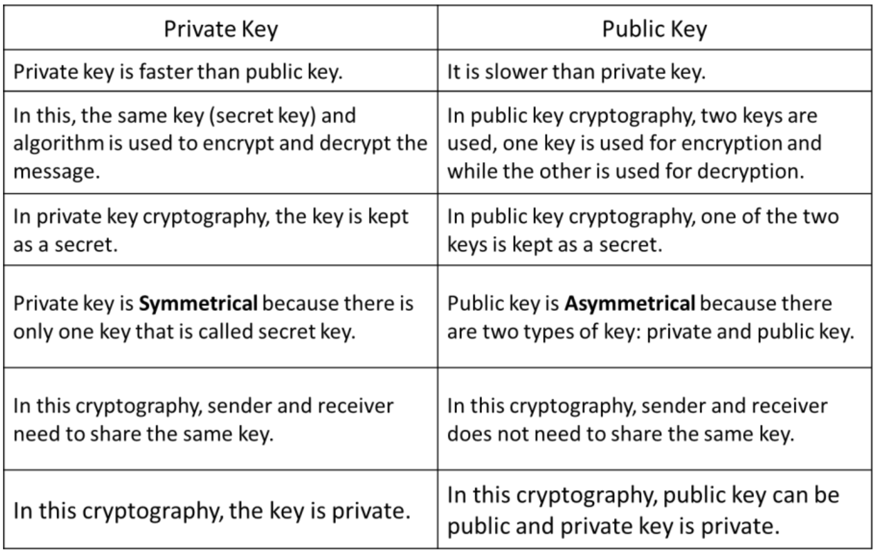
			and how to lessen its vulnerability to disturbance and collapse.
Focus/Target of the Approach	Threats and their underlying causes	The dangers they pose	The endangered referent's capacity to recover or adapt in the face of threats, in particular.
Ontology	The combination of broad underlying structural reasons and proximate factors that mobilise these foundations into tangible threats produces threats.	A Newtonian-mechanistic cosmos with a reasonable number of variables and reasonably straightforward and predictable causality. Problems may be broken down and isolated into their constituent elements.	A complicated cosmos with too many variables to count, emergent features, non-linear causality, phase shifts, limited predictability, and unexpected shocks. Threats originate from the interplay of several variables and are not reducible to individual components.
Threat Type	Economic disparity, relative deprivation, poor governance, and political exclusion as underlying causes of civil war are best suited.	Best suited to threats that can be reduced to specific persons, organisations, and events and then defended against or removed, such as a rival state developing a nuclear programme or a traditional military threat.	Complex adaptive systems that can self-organize (not reducible to individual players) and adapt to defensive measures while eluding elimination, such as the drug trade, which has confounded four decades of drug prohibition, are most adapted to threats.
Example & Approach: The Drug Trade	Reduced demand.	The drug war targets prominent drug lords and organisations with military strikes, with the ultimate objective of eradicating the drug trade by force; crop eradication; and drug interdiction.	Harm reduction measures that lessen the impacts of drug use; decriminalization of some drugs to ease the enforcement burden; and enforcement measures that target only the most violent traffickers.
Example & Approach: Terrorism	Addresses both the frustrations that drive terrorist acts and the societal factors that attract recruits.	Drone strikes and other counter-terrorism tactics are used to locate, apprehend, or murder terrorist leaders, while intelligence and enforcement efforts	Decoupling, contingency planning, and redundancy ensure that our social systems (trade, transport, communications, energy, and so on)

		disrupt current terrorist plots.	can bounce back after an assault; heightened security measures do not overwhelm social systems with complexity, interruptions, resource needs, and rigidities.

14.6 Difference between private key and public key

The difference between private key and public key are shown in Table 2.

Table 2 Private key vs Public key



Private Key	Public Key
Private key is faster than public key.	It is slower than private key.
In this, the same key (secret key) and algorithm is used to encrypt and decrypt the message.	In public key cryptography, two keys are used, one key is used for encryption and while the other is used for decryption.
In private key cryptography, the key is kept as a secret.	In public key cryptography, one of the two keys is kept as a secret.
Private key is **Symmetrical** because there is only one key that is called secret key.	Public key is **Asymmetrical** because there are two types of key: private and public key.
In this cryptography, sender and receiver need to share the same key.	In this cryptography, sender and receiver does not need to share the same key.
In this cryptography, the key is private.	In this cryptography, public key can be public and private key is private.

Summary

- Data on the network is not confidential, thus it must be kept safe from unauthorized users who may be lurking behind networked workstations.
- The malevolent goals might include knocking down network servers, utilizing people's sensitive information such as credit card details for illegal activities, and disrupting large corporations' websites. As a result, it aims to protect data and prevent eavesdroppers from listening in and stealing it.User data on a computer is also safeguarded by granting password-protected access to data and resources, allowing only authorised users to access them. Identifying criminals and foiling their attempts to inflict damage to the network and other resources are also important security considerations.
- Authentication entails validating the antecedents of the person who has requested remote machine services or access, either physically or over e-mail, before permitting him or her to do so. The process of authenticating a person's identification to a distant system is known as authentication.
- The authenticity of a message received by a remote machine is referred to as integrity. In other words, the message transmitted by the source machine is exactly the same as it was

before. The cyclic redundancy coding approach will not suffice in this scenario, as attackers in the system or communication channel may modify the message on purpose.

- Confidentiality: It assures that no one can read the message while it is in transit. This demands the use of encryption techniques in the future.
- With the use of encryption and decryption techniques, the communication is encrypted at the sender end and decoded at the receiving end to ensure anonymity. The secret key and public key procedures are the two options, each with its own set of benefits and drawbacks.
- Transposition and substitution In traditional cryptography, there are two types of cyphers. The encryption mechanism handles parts of the message differently in substitution and transposition.

Keywords

Cipher text: The encrypted message formed by applying the method to the plaintext message using the secret key is known as ciphertext.

IP-spoofing: IP spoofing, like honey pots, entails a computer successfully impersonating a trusted server/resource and intercepting data packets.

Maliciously: Maliciously programmed websites generate chartable webpages that allow users to make donations while also collecting important personal information.

Packet Sniffers: Packet sniffers are devices that intercept data streams over a network in order to gather sensitive information such as usernames, passwords, and credit card numbers.

Password Attacks: A 'Password Attack' includes a number of techniques used by hackers tosteal passwords.

Phishing: Emails with titles such as, "URGENT: Update Account Status" are all attempts by a spammer to "phish" the account details.

Plain text: It is the text message to be transmitted on which an algorithm is applied.

Private key: The key of a key pair, which is used to create a digital signature. It can be used to sign a message that only the corresponding public key can verify

Public key:It is the key of a key pair that is used to verify a digital signature. Key pair consists of private and public key.

Secret Key: They constitute a part of algorithm for encryption and decryption of the message.

Self Assessment

1. In cryptography, what is cipher?

A. algorithm for performing encryption and decryption
B. encrypted message
C. both algorithm for performing encryption and decryption and encrypted message
D. decrypted message

2. In asymmetric key cryptography, the private key is kept by __________

A. sender
B. receiver
C. sender and receiver
D. all the connected devices to the network

3. Cryptanalysis is used __________

A. to find some insecurity in a cryptographic scheme
B. to increase the speed
C. to encrypt the data
D. to make new ciphers

4. The process of transforming plain text into unreadable text.
A. Decryption
B. Encryption
C. Network Security
D. Information Hiding

5.A process of making the encrypted text readable again.

A. Decryption
B. Encryption
C. Network Security
D. Information Hiding

6. Public-key cryptography is also known as ?

A. asymmetric cryptography
B. symmetric cryptography
C. Both A and B
D. None of the above

7. Which of the following keys are known only to the owner?

A. public key
B. protected key
C. private key
D. unique key

8. PKI stands for?

A. public key infrastructure
B. private key infrastructure
C. public key instance
D. private key instance

9. In which of the following, a person is constantly followed/chased by another person or group of several peoples?

A. Phishing
B. Bulling
C. Stalking
D. Identity theft

10.Which of the following refers to the violation of the principle if a computer is no more accessible?

A. Access control

B. Confidentiality

C. Availability

D. All of the above

11.The term "TCP/IP" stands for_____

A. Transmission Contribution protocol/ internet protocol

B. Transmission Control Protocol/ internet protocol

C. Transaction Control protocol/ internet protocol

D. Transmission Control Protocol/ internet protocol

12. In the computer networks, the encryption techniques are primarily used for improving the ________

A. Security

B. Performance

C. Reliability

D. Longevity

Answers for Self Assessment

1.	A	2.	B	3.	A	4.	B	5.	A
6.	A	7.	C	8.	A	9.	C	10.	C
11.	B	12.	A					.	

Review Questions

1. What are different criterions to keep information private when it is sent over a public network?
2. How does the encryption affect performance of network?
3. There are certain information bases on the Internet that need to be prevented by undesirable person to get. How can undesirable person be kept from accessing this?
4. How do we keep our own and other people's computers safe from hackers? Explain with the help of a hypothetical situation.
5. What is a Cipher? Why are cipher used for large messages?
6. Describe briefly two kinds of security attacks, which can be directed against an Internet connected computer system
7. What is the difference between secret key and public key encryption?
8. What is cryptography? What are the benefits of using this technique?
9. What do you mean by substitution and transposition ciphers? Differentiate between the two.

Further Readings

Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall

Behrouz A. Forouzan and Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies

Burton, Bill, *Remote Access for Cisco Networks*, McGraw-Hill, Osborne Media

Dale Tesch/Greg Abelar, Security Threat Mitigation and Response: Understanding *CS-MARS*, Cisco Press, Sep. 26, 2006.

Gary Halleen/Greg Kellogg, Security Monitoring with Cisco Security MARS, Cisco Press, Jul. 6, 2007.

LOVELY PROFESSIONAL UNIVERSITY
Jalandhar-Delhi G.T. Road (NH-1)
Phagwara, Punjab (India)-144411
For Enquiry: +91-1824-521360
Fax.: +91-1824-506111
Email: odl@lpu.co.in