INTRODUCTION TO NETWORKS

Connecting Digitally

By Umar Khokhar & Binh Tran
Georgia Gwinnett College

Table of Contents

Chapter 1

INTRODUCTION TO NETWORKS

In this chapter, the student will learn the basics of the networks, types of networks, computer components and their functions

- Computer Components
- Network Communication models
- Network components
- Common Network Terminologies
- Network Communication steps

1.1 What is a Network?

A network is a group of computers connected using a transmission medium (e.g. cable, fiber optics or wireless communication technologies). A network is designed to allow the computers to:

- Exchange information
- Share resources

1.2 Fundamentals of Computers

A computer is any digital device which can perform/offer four (4) functions:

- Accepts the input from the users
- Process the users input and execute the instructions
- Stores the data for future usage
- Provides the output in a human understandable format

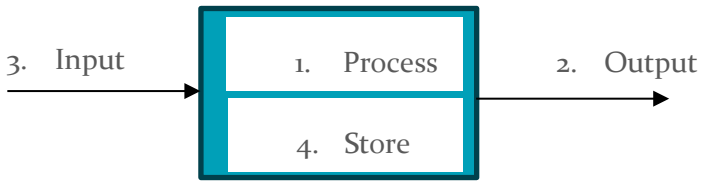


Figure 2.1: Computer Functions

1.2.1 INPUT DEVICES

The input devices allow the users to interact with the computers. The input devices mainly convert the human language into the Machine Language (Binary) e.g. Keyboard, Mouse and Microphone etc. are the examples of the input devices.

1.2.2 OUTPUT DEVICES

The output devices convert the machine language back into the human language (e.g. text, sound, graphics etc.). The monitor, speakers and the printers are the examples of the output devices.

1.2.3 PROCESSING

The component of the computer which executes the instructions of the programs is called CPU (Central Processing Unit). The CPU fetches the instructions from the RAM and executes them.

1.2.4 STORAGE DEVICES

The storage devices hold the data for future processing/usage. The storage devices can be of two types:

- Short Term storage (Main Memory)
- Long term storage (Secondary Memory)

1.2.4.1 Short Term Storage

Volatile devices, which requires constant power supply to retain the contents. Random Access Memory (RAM) is one of the examples of the short-term storage. The RAM is the most active part, or you can say the working memory which holds the currently in use data. If you have a large RAM, it allows you to run heavy applications/software and programs. If the RAM couldn't handle a program or application, then it uses the secondary memory as a backup.

1.2.4.2 Long Term Storage

Non-volatile devices, which retain the content even in the absence of the power supply. The tradition hard drive, Solid State Drive (SSD), USB (Universal Serial Bus), CDs, DVDs and the Blu-ray are the examples of the long-term storage. The traditional hard drive involves platters and moveable read/write head which uses magnetic fields to read/write the data onto it. However, the SSD uses flash memory (registers and blocks) to store the data which makes it extremely faster and expensive.

1.2.5 BUSES

Collection of wires, carrying the data from one component to another component of the computer. The buses connect the various components of the computers to CPU (e.g. CPU to RAM, CPU to I/O devices and CPU to Expansion buses etc.).

1.3 Hierarchy of data representation

The smallest entity a computer can process is called 'bit', which is the acronym for Binary Digits. The bit can be either '0' or '1' where '0' represents the absence of the voltage and '1' represents the presence of the voltage.

To optimize the data storage representation, we use the Bytes, KB and MBs etc. The digital data hierarchy is presented in table 1.1.

1 Byte	8 bits
1 KB (Kilo)	1024 Byte
1 MB (Mega)	1024 KB
1 GB (Giga)	1024 MB
1 TB (Tera)	1024 GB
1 PB (Peta)	1024 TB
1 EB (Exa)	1024 PB
1 ZB (Zeta)	1024 EB
1 YB (Yuta)	1024 ZB

Table 1.1: Data Hierarchy

1.4 Components of the Personal Computer (PC)

The PC involves four (4) major components:

- Motherboard
- RAM
- Storage Device
- Firmware
- Device Controllers

The detailed description of these components is presented as follows:

1.4.1 MOTHERBOARD:

The motherboard is located under the system unit and called the mainboard. All the major computer components e.g. RAM, CPU, ROM, GPU etc. are attached onto it and communicate through the buses. There are many slots and sockets on mainboard onto which the computing components are connected to. The details of some of components are presented as follows:

- **CPU Socket**: Connects the CPU
- **PCI (Peripheral Component Interconnect) Express Expansion bus**: Allows expansion (hardware) of the system.
- **RAM Slot**: Connects the RAM
- **Chipset:** Allows the data transfer between computer components.
- **Main Power Connector:** Power supply to motherboard
- **SATA (Serial Advanced Technology Attachment) Connector:** Connects the SSD or external hard drive.

1.4.2 RAM:

The RAM is also known as main memory which holds the programs and process which are currently being executed. The RAM forwards the instructions to the

CPU and the CPU executes the instructions and forwards the results back to RAM or I/O devices.

1.4.3 STORAGE DEVICES:

See 1.2.4 for details.

1.4.4 FIRMWARE:

The firmware is the computer program installed by the developer of the hardware. The firmware is usually stored onto the ROM (Read Only Memory) and contains information about the bootup instructions. The boot process involves six steps:

- After we power on the computer, it powers on the motherboard.
- Initialize the CPU
- CPU initializes the BIOS routines and runs POST (Power On Self-Test)
- Identifies the devices and looks for OS
- Loads the OS onto RAM
- Starts the OS Services

1.5 Network Components

The components that turn a standalone computer into a networked device are called network components. The detailed description of the network components is presented as follows:

1.5.1 NETWORK INTERFACE CARD (NIC):

The NIC is a physical card (circuit) plugged onto the motherboard which can be wired (Ethernet) or wireless (IEEE-820.11) based interface circuit. The NIC connects the computer to the network devices and with other network resources.

1.5.2 NETWORK MEDIUM:

The network medium is the channel through which the devices over the network communicate with each other. It can be wired (Ethernet, Optical fiber, Co-axial cable etc.) or wireless (WIFI, Radio Frequency or Bluetooth etc.).

1.5.3 INTERCONNECTING DEVICES:

The devices/computers on the network are connected with each other using Switches or hubs which are called interconnecting devices.

1.5.4 NETWORK SOFTWARE:

To communicate with other devices on the network, the devices use either network client software (request for services) or network server software (provides the services).

1.5.5 NETWORK PROTOCOLS:

The computer over the network uses TCP/IP protocol, which is the suite of the protocols and allows the networked computers to perform various network activities e.g. Email exchange, uploading, downloading etc. (SMTP, HTTP, FTP etc.)

1.5.6 NIC DRIVERS:

Receives the data from the network protocols and then forwards to the physical NIC which turns them into signal and sends over the network medium.

1.6 Steps of network communication

The network communication involves mainly four steps, which are presented as follows:

1. The client uses the user application to interact and create a request for service.
2. The client uses the network client software to format the request/message.
3. Then it uses the network protocol and encapsulates the network frame. Firstly, it formulates the request/message into the specified format (e.g. http, https, ftp, smtp etc.) and adds the Source and Destination MAC and IP addresses.
4. The frame will be then forwarded to NIC which converts it into the signals and forwards it to the server (receiver) through network medium. The server will de-encapsulate the request, execute it and forwards the response back to the client in the same way.

Example of network communication using PING Utility:

The ping and trace route are the two ICMP protocol utilities which is a UDP protocol (We will discuss the UDP and TCP protocols in detail in chapter 4). The ping utility is being used to check the network connection with other network devices and resources. Assume that we have multiple devices connected with each other over a LAN using a switch. The figure 2.2 presents the LAN model. The steps of the network communication of Ping are presented as follows:

1. Open the terminal and type “ping 1.1.1.2” (User Application)

2. The ping request (message) will be created (Network Software)
3. The ping message will be encapsulated (Network Protocol), the source and the destination IP and MAC addresses. If the receiver's MAC addresses are not known to the sender, then it broadcast the ARP (Address Resolution Protocol). The ARP request will be received by all the connected nodes and will forward the MAC address of the requested device (If they know the MAC address of the requested device). The response is called ARP response.
4. Then the frame will be forwarded to the NIC. The NIC forwards the frame to the switch which forwards to the correct device.

1.7 Common Network Terminologies

Based on the area/range, the network can be categorized into three types; LAN, WAN and MAN while based on the connectivity, it can either wired or wireless. In this section, we will discuss the common network terminologies which we will use throughout the book. Some of the commonly known terminologies are described as follows:

1.7.1 LOCAL AREA NETWORK (LAN):

Group of computers connected together over a small geographic area (e.g. within a building, home, or organization etc.). The LAN offers faster speed and better security as compared to the other types of networks.

Example of the LAN: Assume that an organization has two office locations (LANs) one in New York and second in Georgia, if they connect both of the locations (networks) using a private cable/optical fiber then despite of distance, the network will still be considered as LAN.

1.7.2 METROPOLITIAN AREA NETWORK (MAN):

Private network which spans over several building within the city. The MAN usually belongs to same organization which has multiple locations across the city. The MAN involves public private (networking technologies) to connect multiple locations.

1.7.3 WIDE AREA NETWORK (WAN):

Connects two or more LANs using public networks (Internet Service Providers).

Example of the WAN: Assume that an organization has two office locations (LANs) one in New York and second in Georgia, if they connect both of the locations (networks) using a public network (ISPs) then, the network will still be considered as WAN. Similarly, if two people connected to same LAN, sitting in the

same room and using WhatsApp for chatting with each other then they are using WAN (since it involves public network/cloud).

1.7.4 INTERNET:

Network of interconnected computer networks is called internet. If we connect all the WANs across the globe, it forms the internet (The largest WAN on the planet).

1.7.5 INTRANET:

The network within the organization which is a private network of the organization that allows the access of the organizational resources to those connected to the internal network.

1.7.6 EXTRANET:

Limits and controls the access of the internal organizational resources to the outsiders. Allows the outsiders to access the organization resources from outside LAN.

1.7.7 PACKET:

When we add the source and destination IP address (Network Layer) to the data then data turns into packet.

1.7.8 FRAME:

Packet with the source and destination (Gateway) MAC addresses (Data link layer) is called Frame.

1.7.9 ENCAPSULATION:

The process of appending the source and the destination IP and MAC addresses is called encapsulation.

1.7.10 SERVER:

A computer that offers services and facilities the other computers on the network.

1.7.11 CLIENT:

A computer which requests for the services from other computers on the network e.g. when we provide an address to google maps then the google maps server executes your request and sends you the result. In this example your computer acts a client machine.

1.8 Network Model

To design a network (LAN or WAN), there are mainly two types of network models: Client-server and the Peer to Peer (P2P). The detailed description of both models is presented as follows:

1.8.1 CLIENT-SERVER:

The most commonly used model to design LAN. In this model, the workstation or PC are considered as clients of the network and if they want to do any activity e.g. resource and internet access then they will request to the server. The server is considered as a powerful machine which handles the requests of all of the clients and also manages the access control.

If we use Microsoft server (software) then it mainly uses two applications: Domain Controllers and the Active Directory. The domain controllers manage the user accounts (credentials) and the active directory controls the access of the resources to the clients.

1.8.2 PEER TO PEER (P2P):

Unlike the client-server, in the P2P all of the computers in the network, can take the role of both client and server simultaneously (based on the need). To better understand the concept of the P2P let us take an example. Assume that we have four computers in a LAN: Computer A, B, C and D. The computer A offers FTP services, Computer B offers SMTP, Computer C offers HTTP and D offers a printer. If computer A wants to take a printout, then it will request to computer D (behaves like client to that computer). Similarly, if computer C wants to transmit an email, then it will forward request to Computer B (behaves like client to that computer). The only problem of the peer-to-peer networking is, one has to memorize/manage numerous credentials (usernames and the passwords) to use the services of available servers. Moreover, there will be no central control of the resources, since all the computers in the network have same level of privilege.

Chapter 2

NETWORKING DEVICES

In this chapter, the student will learn the working, limitation, Pros and Cons of the following networking devices:

- Repeaters
- Hubs
- Switches
- Routers
- Network Interface Cards (NICs)
- Wireless Access Points (WAPs)

2.1 What is a Networking Device?

Device which can establish communication between two or more computers or networks is called networking device.

These networking devices are located on layer 1,2 and 3 of the OSI model (discussed in Chapter 7) based on their functionalities. The hubs and the repeaters operate on layer 1 (Physical Layer), switches and the bridges operate on layer 2 (data link) while layer 3 switches and the routers operate at layer 3 (Network). The detailed description of all these networking devices is presented as follows:

2.2 Repeater

The repeaters connect two Local Area Network (LAN) segments (not the two LANs) and improve the network range by broadcasting and regenerating the original signals. The repeaters operate at layer 1 (Physical) and are not considered intelligent devices, as they do not have access to MAC and IP addresses. To better understand the working of repeaters, let us consider the following example:

Let's say we need to set up a local area network (LAN) that covers an area of more than 200 meters. We'll use a type of Ethernet called 10 Base T, where the "10" means it runs at 10 Mbps and the "T" stands for twisted-pair cable, which can reach up to 100 meters in length. Since our LAN area exceeds 100 meters, we'll need to divide it into two segments. The first segment, up to 100 meters, will have good signal strength. For the next segment, we'll need to regenerate the signal to maintain its strength. To do this, we'll place a device called a repeater at the 100-meter mark. The repeater takes the signal from the first segment, regenerates it to its original strength, and then sends it to the next segment. This way, we can cover the entire 200-meter area while maintaining a strong and reliable signal.

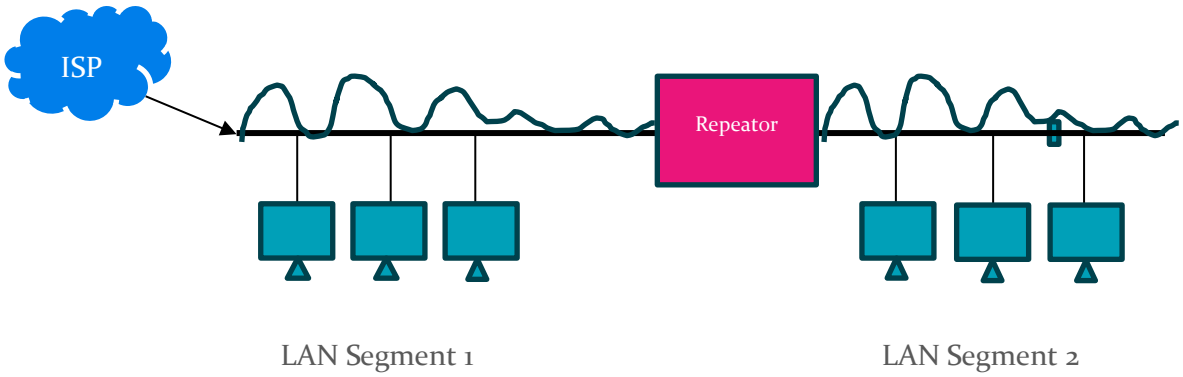


Figure 2.1: Repeater connecting two LAN segments

The repeater has only one collision domain (channel) therefore if computers from different LAN segments start using the same channel simultaneously then there might be a collision between the two signals.

The repeater regenerates/repeats the original signal and doesn't amplify the signal so, it is not considered as amplifier.

2.3 Hubs

Just like repeaters, the hubs also operate on layer 1 of the OSI model. The hub connects the devices within the LAN and transmits the electrical signals (bits). Although, the hubs incorporate the single collision domain, but they use half duplex communication model to avoid collisions. In half duplex model, there is only one communication channel between receiver and the sender (only one entity can send the message at a time). Since, the hubs reside on physical layer and cannot store the MAC addresses of the connected nodes (devices) therefore, when the hub receives any data/information on any one of the ports they broadcast it to all the nodes. To better understand the functionalities of the hub, let us consider the following example.

Assume that we have a hub (placed as centralized device) which connects four (4) workstations (A, B, C and D). If workstation A wants to communicate with workstation D then A will forward the data/frame (having source and destination's IP address and MAC address) to hub. Since, the hub doesn't know the MAC addresses of the connected workstations therefore, it will broadcast the frame to all the workstations. All other workstations (B and C) will discard the frame (since, it is not addressed to them) and only D will open and read the message. However, if the other workstations configure promiscuous mode on their NIC then they can also read and open that frame. This poses not only the real security challenge (confidentiality) but there will be enormous amount of traffic in the network (e.g. A wants to send the data to B then besides the B other connected workstations will also receive a copy of the data). Therefore, the hubs flood all the connected nodes. The figure 2.2 describes the working of the hub (4 ports) in a LAN.

The hubs usually have 2, 4 and 8 ports which means it connects 2, 4 and 8 workstations in LANs. The hubs can be active or passive devices:

The active hubs need a constant power supply to operate, and it can also amplify or regenerate the signals (just like repeaters). The active hubs can be used as an alternative to the repeaters. The passive hubs don't need constant electricity and they can only broadcast the signal/data but cannot regenerate or amplify the signal strength.

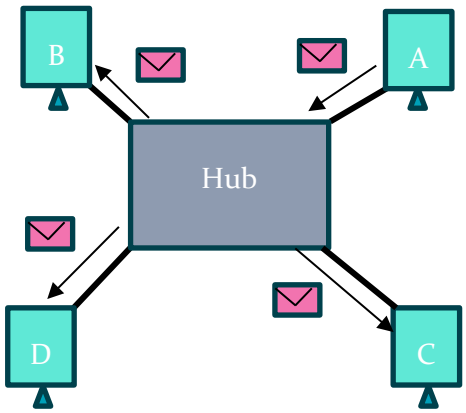


Figure 2.2: LAN using Hub

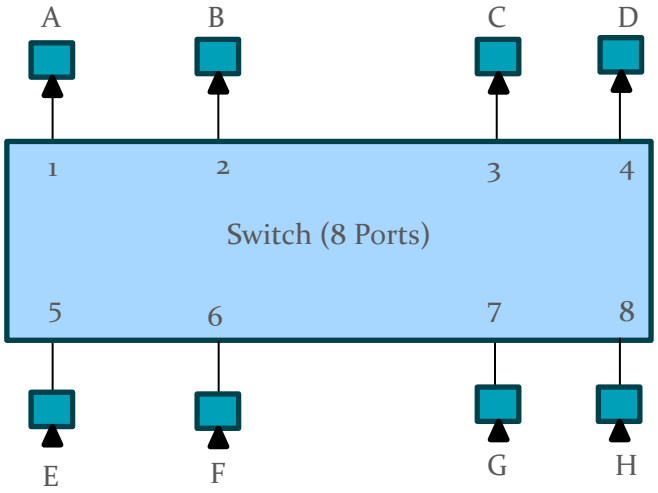
2.4 Switches

Just like hubs, the switches are the networking devices which connect the computers (workstations) in the LAN. Unlike the hubs, the switches operate at layer 2 (data link) of OSI model. The switches use full duplex communication model (supports bidirectional communication model) and maintain a CAM (Content Addressable Memory) where it stores the MAC addresses of the connected devices. If the workstation A wants to send a message to workstation B then A will forward the message to the switch and the switch will forward the message to that specific port with which B (receiver) is connected to. The switches do not broadcast the frames to all connected Usually, the switches contain 8/16/32/48/64 ports, where each port provides a separate collision domain to avoid the collisions. The switches offer only one broadcast domain and offer various data transmission speeds for both wireless and wired transmission mediums. Let us understand the working of the switch using the following example:

Assume that we have 8-port switch and there are eight (8) workstations connected to it. The switch offers multiple collision domains (multiple channels) which means all workstations can communicate with other workstations simultaneously and hence offers a full duplex communication model. Since, switches offer multiple collision domain therefore it can avoid the device collision.

The switch maintains a CAM table to forward the packets to the correct ports with which the requested workstations are connected to. For example, we just have configured a 8-port switch with 8 workstations are connected to it. Initially, the CAM table will be empty, and it will get populated once the workstations will start communicating with each other. Assume that workstation A wants to send a packet to workstation H (where A and H are the MAC addresses of the workstations) then switch will broadcast this packet and add MAC address A next to port 1 in the CAM table (See the figure 2.3 below) so, next time if any workstation wants to communicate with workstation A then switch will not broadcast the

packet and forward it only to Port 1. Upon receiving the packet from workstation, A, if the workstation H wants to respond back to A then switch will forward the packet to Port 1 (since, it knows the MAC address of A) and adds the MAC address H next to port 8 in the CAM table. Similarly, when other workstations forward any packet to each other, the CAM table will get updated and once it will get fully updated then switch will control the traffic flow based on the MAC addresses.



Port No.	MAC
1	A
2	B
3	C
4	D
5	E
6	F
7	G
8	H

Figure 2.3: Switch and CAM table

In conclusion, the switch first broadcast and then after the CAM table gets fully updated then it controls the traffic flow with unicasting and multicasting. There is only one broadcast domain and usually the ethernet switch offers 8/16/32/48/64/128 Ports.

2.4.1 Types of Switches

There are mainly four (4) different types of switches:

a) Store and forward Switch
b) Cut through Switch
c) Fragment Free Switch
d) Adaptive Switching

The detailed description of each type of switch is presented as follows:

2.4.1.1 *Store and forward switch*

The most common type of switch which buffers and verifies each complete frame before transmitting it. This type of switches is usually slow but reliable.

2.4.1.2 *Cut through Switch*

These types of switches don't perform any verification and error checking. They just read the MAC addresses of the frames and forwards them.

2.4.1.3 *Fragment Free Switch*

This type of switch mainly checks the header of the frame (first 64 Bytes) for errors using CRC (Cyclic Redundancy Check). If it finds any errors, then the packet will not be forwarded. This type of switch is also faster than store and forward switch since, it just checks the header of the frame (not the whole frame).

2.4.1.4 *Adaptive Switching*

It gives you option e.g. 1 to make the switch behave like Store and forward, 2 to make it act like cut through switch and 3 to make it like fragment free switch.

2.5 Routers

The routers are the layer 3 (network) devices and use IP addresses for the data transmission. Unlike the switches, which are mainly the network devices (connects the devices/computers within the LAN), the routers are WAN/Internetworking devices which connect the LANs. The routers connect two different LAN (having different Network IDs) and each port of the router has a separate broadcast domain. The router uses the routing tables and route the packets to the correct destination. Let us consider the following example to better understand the difference between a network device (e.g. Switch and the internetworking device (Router).

The networking devices (e.g. Hubs and Switches) connect the computers/devices belong to same network ID within one LAN while the internetworking devices e.g. Routers, connect two or more different LANs (having different Network IDs).

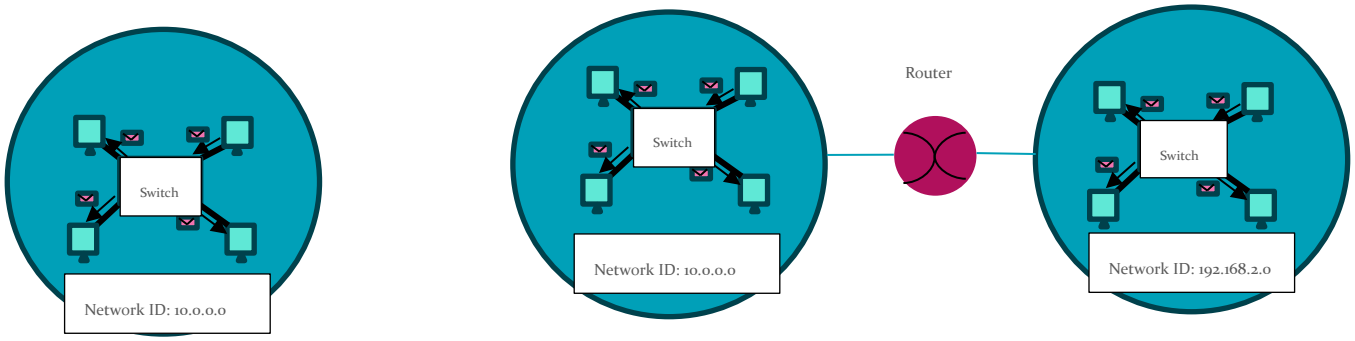


Figure 2.4: Networking Device vs Internetworking Device

The figure 2.4 compares the networking vs internetworking devices. For the LAN (having the network ID 10.0.0.0), all the devices/computers connect through a switch, however for the WAN having two different network (LAN-1' Network ID: 10.0.0.0 and the LAN-2' Network ID: 192.168.2.0), we need a router for connectivity. The router manages the routing table which holds the network IDs of the connected networks, so the router can route the packets to and from the correct network.

Assume that a computer (e.g. 192.168.2.1) from LAN-2 having Network ID: 192.168.2.0 sends a ping packet to a computer (10.0.0.1) on LAN-2 having Network ID: 10.0.0.0. The packet (having source and destination IP address) will go to switch of the LAN, the switch checks the destination IP address (finds the IP address belong to remote network). The switch forwards the packet to the router which check its routing table to find the port with which the network is connected to and forwards the packet to that LAN (switch). Upon receiving the packet from the router, the switch checks its CAM table to find the port number with which the computer with IP address 10.0.0.1 connected to.

Besides the network IDs the routing tables also holds the routing maps of the connected networks to find the shortest and optimal path for the packets to be routed.

2.6 Network Interface Card (NIC)

The NIC is a circuit/card that is installed onto the motherboard and offers an ethernet port which provides an interface between computer the network. The NIC not only connects the computer to the network but also provides an identification (MAC address) to the computer over the network. The NIC is also termed as LAN adapter or network interface controller. So, any device/computer who wants to get connected onto the network, it needs an NIC.

The NIC is of 6 Bytes (48 bits) long e.g. 34:FF:AE:00:AF:EF and stored onto the ROM. The first 3 bytes identifies the Organization Unique Identifier (OUI) issued by IANA and last 3 bytes represent the vendor serial number. The NIC operates on the Physical (layer 1) and the data link layer (layer 2) of the OSI model. The working principle of the NIC (sending and receiving operation) as presented as follows:

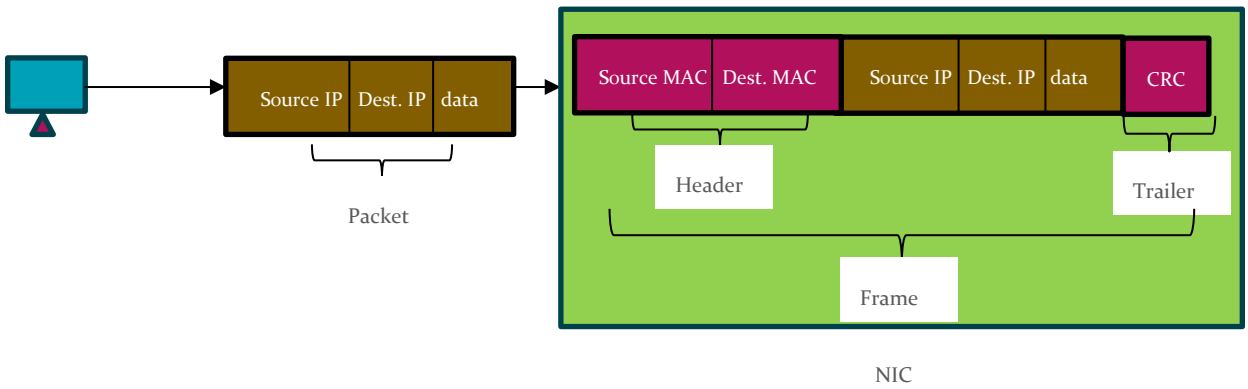


Figure 2.5: NIC Operation (Sending Side)

As shown in the figure 2.5, the NIC receives packet from the protocol and adds the header (source and destination MAC address) and the trailer (Cyclic Redundancy Check for error detection and correction) to the packet and turns the packet into frame. Finally, transmits the signal to the destination over the transmission medium.

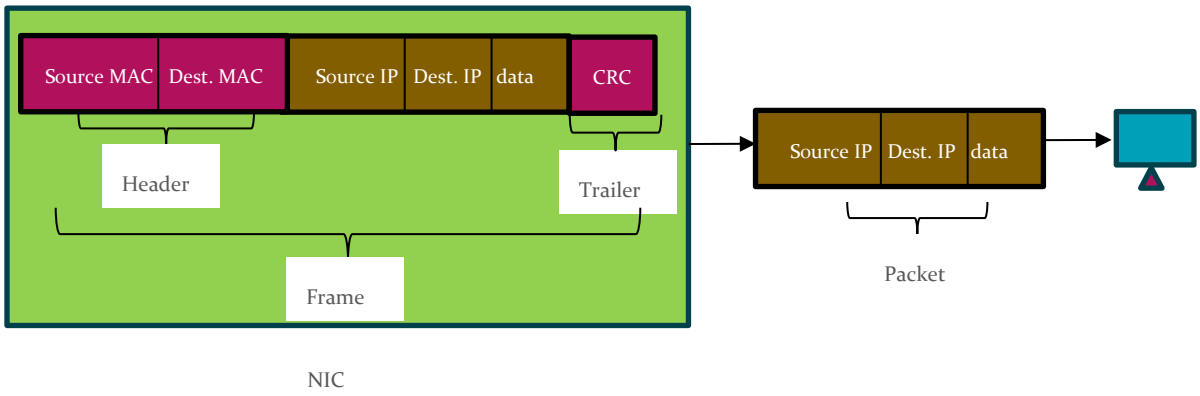


Figure 2.6: NIC Operation (Receiving Side)

As shown in the figure 2.6, upon receiving the signal from the network device, the NIC converts the signal/binary into frame. The NIC further verifies the MAC addresses and checks CRC for errors. Upon successful verification, the NIC removes the header and the trailer and forwards the packet to the protocol.

2.5.1 Modes of NIC

The NIC operates in two modes: Normal and Promiscuous mode. In the normal mode, the NIC captures only those packets which are directed to its MAC address, however in the promiscuous mode, the NIC grabs every packet comes it way, even directed for someone else. For the traffic monitoring and protocol analysis/sniffing, the NIC is configured to promiscuous mode.

2.5.2 Broadcast packet

If a computer wants to broadcast a packet over the network, then the destination MAC address must be set to FF:FF:FF:FF:FF:FF (set binary to all 1).

2.5.3 Wireless NIC

The wireless NIC works like an antenna and uses radio waves to connect the computer to the network (broadcasting wifi signals).

2.7 Wireless Access Points (WAP)

Just like other networking devices (e.g. switches and hubs) connects the LAN devices but without ethernet cables/wires. The major difference between the WAP and the ethernet networking devices is the acknowledgement capability. When the sender sends a frame then receives must have to acknowledge. To avoid the collision (single collision domain), the sender first sends RTS (Request To Send) to WAP and upon receiving CTS (Clear to Send), the sender forwards the frame to the receiver.

Chapter 3

NETWORK TOPLOGIES AND TECHNOLOGIES

In this chapter, the student will learn how to design a network (LAN and WAN) and have a detailed understanding of the LAN technologies

- Network Topologies
- Local Area Network implementation technologies

3.1 Network Topologies

The network topologies depict the design of the network and presents how network physically looks like. The network topologies can be: Physical or logical

- **Physical**: presents how physically the computers/devices are connected to form a network.
- **Logical**: regardless of the physical connectivity, the logical topology presents how the communication between the devices takes place. For example, the computer physically connected in a star format (will be discussed in 3.2) however, the information is travelling in a ring format.

3.2 Types of Network Topologies

There are mainly six (6) types of network topologies, the figure 3.1 presents the types of network topologies.

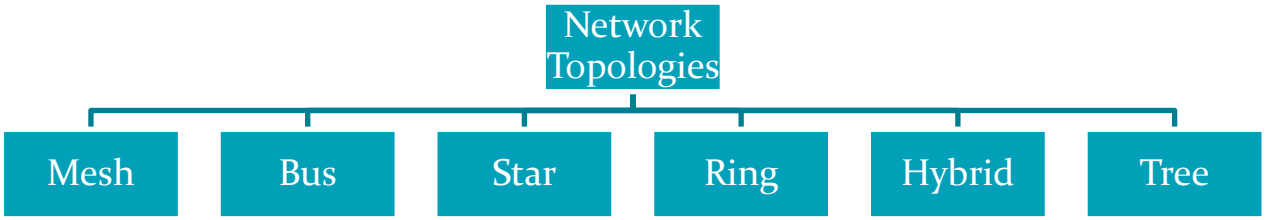


Figure 3.1: Types of Network Topologies

The detailed description of these network topologies is discussed as follows.

3.2.1 Mesh Topology

In the mesh topology, every computer is connected to every other computer in the network (dedicated connection/path). The mesh topology offers the highest level of redundancy and widely used in WAN design. The figure 3.2 presents the mesh topology

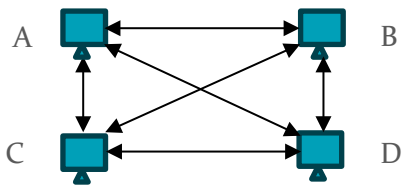


Figure 3.2: Mesh Topology

As shown in the figure 3.2, the mesh topology offers diverse connections between all the computers of the network, which makes it more reliable and faster over other network topologies. If Computer A wants to send a packet to computer B then they have a dedicated connection from A to B which both of the computers can use to exchange information. Similarly, all other computers also have a dedicated (direct) connection with other computers and form a peer-to-peer network model. If any one of the connections gets down, then computers can still communicate using other/alternate connections. The mesh topology is widely used in WAN where routers use mesh topology and connect multiple LANs. The only problem of the mesh topology is the implementation cost, it involves excessive amount of wires/channels which will increase the overall cost of the network. Moreover, the troubleshooting of the network will also be complicated as compared to other network topologies, which makes it less common in LAN design.

3.2.2 Bus Topology

The figure 3.3 presents the bus topology model. The bus topology uses backbone/trunk cable (called Bus) and the devices are connected to the bus directly using T-connector. On both ends of the backbone bus, we use terminators which avoid the reflection of the signal and prevents the signal distortion. The co-axial cable or ethernet (10 Base T etc.) can be used as backbone cable and it doesn't require any networking device (switch or hubs) for implementation of the network. The bus topology requires a smaller number of the cable as compared to mesh topology and also easy to implement. The bus topology mainly broadcast the information and doesn't support unicast and multicast (logically it can) e.g. if computer A wants to send a packet to computer B then it will be broadcasted (every computer connected to the bus will receive a copy of the packet). Because of the broadcasting nature, the bus topology doesn't offer good security and not considered reliable option to implement a LAN. If any segment of the backbone cable gets break, then it can disconnect/disrupt the LAN. Even though, it is easier to add new devices to the backbone cable but each time, we add new devices the overall network goes down.

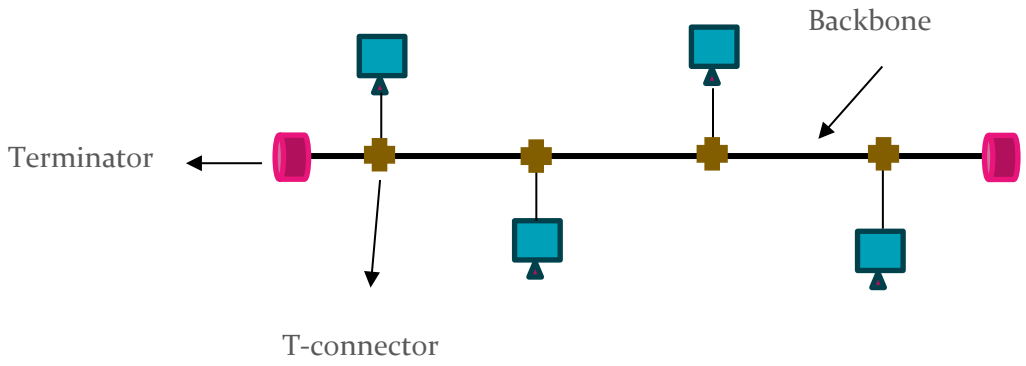


Figure 3.3: Bus Topology

3.2.3 *Star Topology*

The most common topology to implement LAN, where all the computers are connected to a centralized networking device (Switch/hub). Each device needs a single cable (point to point) connection between the switch/hub and computer. In case of cable failure, it can only affect single computer connected to the networking device. The star topology is the easiest and the cost-effective way to design LAN. The star topology offers easy troubleshooting and addition of the devices. The star topology can also be used to extend the LAN range (extended star). The figure 3.4 presents the star topology and the 3.5 depicts the extended star.

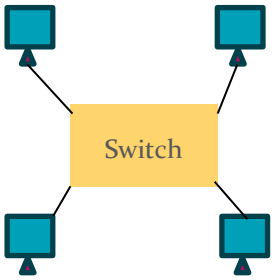


Figure 3.4: Star Topology

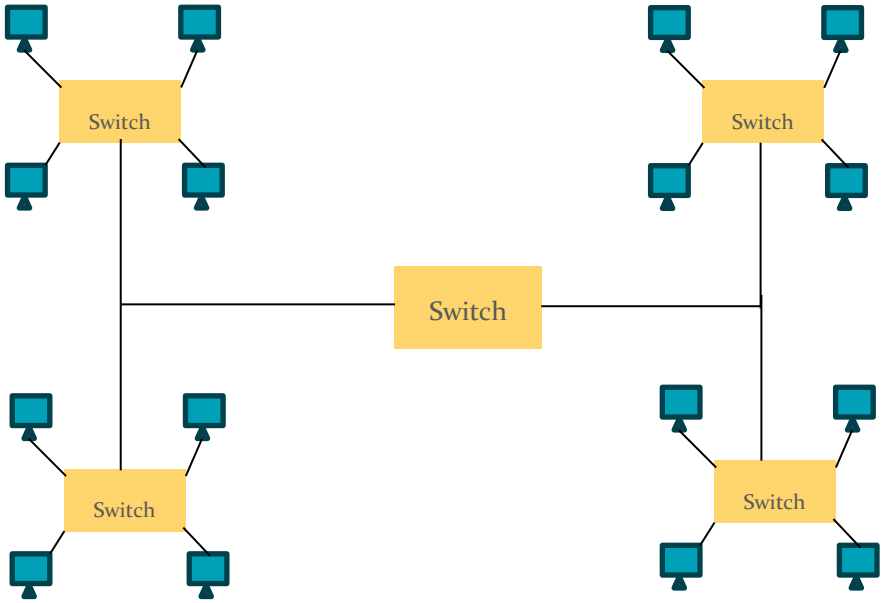


Figure 3.5: Extended Star Topology

3.2.4 *Ring Topology*

The figure 3.6 presents the ring topology, where computers are connected in a ring fashion. The ring topology is also known as token ring since, it uses a token to control the flow of information. In ring topology, the information flows in counterclockwise direction. The information travels from one computer to another in anti-clockwise fashion. If computer A wants to send a packet to B then firstly, it needs to acquire a token and then it will forward the packet to D, D will forward to C then eventually, C will deliver the packet to B.

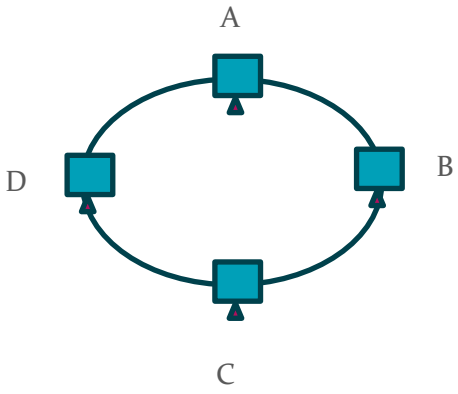


Figure 3.6: Ring Topology

The ring networks are most commonly wired in physical star configuration and uses a device (centralized) Multi Station Access Device (MSAD) to control the traffic using token. Assume that computer A wants to send a packet to B then MSAD issues the token to A and the packet will firstly go to D, then it will be forwarded to C and eventually B will receive it from C. The figure 3.7 depicts the Token ring network using MSAD.

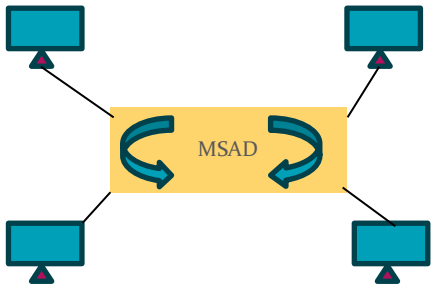


Figure 3.7: Logical Ring (MSAD)Topology

The ring topology can be used to connect various segments of the LAN (star LANs). The following figure 3.8 presents the application of the ring topology.

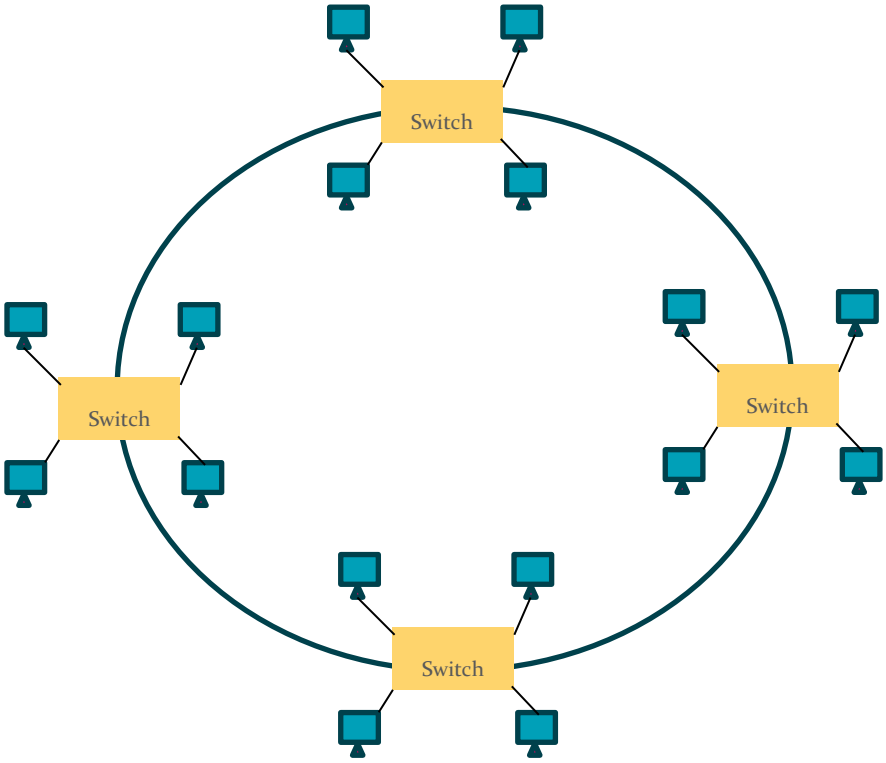


Figure 3.8: Ring-Star Topology

3.2.5 *Hybrid Topology*

The combination of two or more topologies to design a LAN is called hybrid topology. The figure 3.8 (Ring-Star) presents one of Hybrid topology example. To better understand the hybrid topology, let us take the following example. Assume that you have just joined an organization and they have two disconnected segments of LAN, you have been assigned a task to connect both the LAN segments without changing the existing network topologies. Then we can use bus topology to connect both segments without making any significant change in the network design. This will form a hybrid network design topology. The figure 3.9 presents the hybrid topology with the above-mentioned design specifications.

3.2.6 *Tree Topology*

The tree topology is also the hybrid topology which is basically the combination of the bus and star topologies. When we connect different segments of Star LANs using Bus to extend LAN range then this topology is called tree topology.

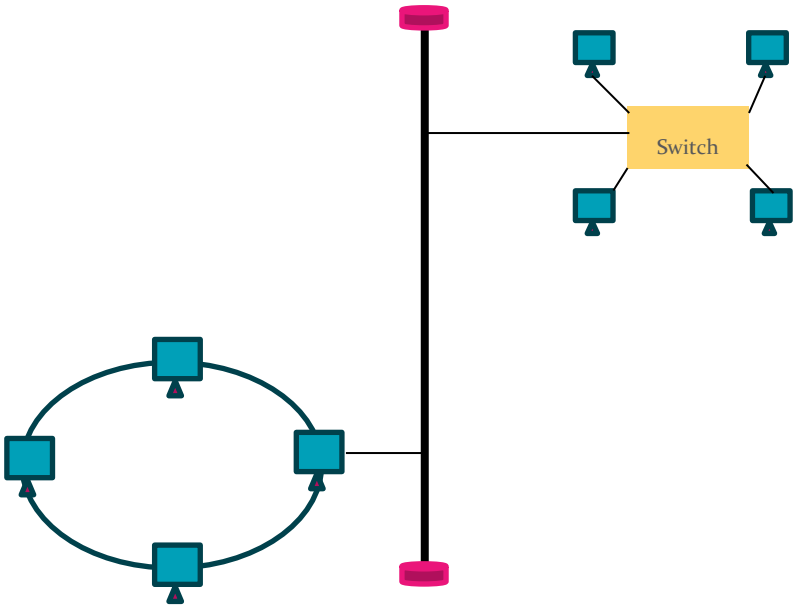


Figure 3.9: Hybrid Topology (Star-Bus-Ring)

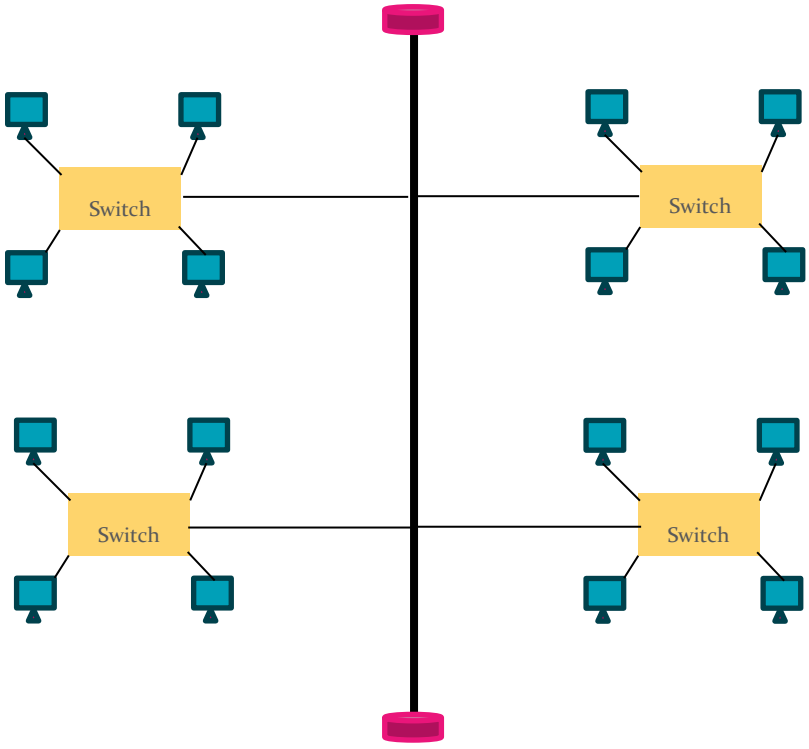


Figure 3.10: Tree Topology (Bus-Star)

3.3 Network Technologies

After designing the network (LAN), we use network technologies which computers use to access the medium (for communication). For LAN, we mainly use two network technologies: Ethernet (Wired) and WIFI (Wireless). These network technologies provide the mechanism for channel accessing, channel specification and frame structure. The detailed description of these network technologies is presented as follows.

3.3.1 Ethernet

In this section, we will discuss:

- Overview of Ethernet
- Addressing
- Frame structure
- 'Media access
- Standards

3.3.1.1 Overview of the Ethernet

The ethernet is introduced by IEEE (Institute of Electrical and Electronics Engineers) in 1983 under the project IEEE-802.3. The ethernet forms wired LAN and one of the most commonly implemented LAN technology. The ethernet mainly uses star and bus topologies to design the LAN and offers network speeds from 1Mbps to 400 Gbps (depends on the channel conditions and type of medium).

3.3.1.2 Ethernet Addressing

The ethernet uses MAC addresses for communication. Every computer in the LAN has an ethernet NIC which provides the unique ID (MAC address) which computers use for LAN information exchange (computers uses ARP protocol to update its ARP cache which holds the MAC addresses of the nearby devices). The ethernet protocol doesn't involve any message acknowledgement and relies on the network protocols to ensure delivery. The ethernet uses Manchester encoding for data transmission over the channel.

3.3.1.3 Ethernet Frame

The figure 3.11 presents the ethernet frame. The ethernet frame involves 7 fields: Preamble, Starting Frame Delimiter (SFD), Destination MAC address, Source MAC address, Length, payload (data) and the Cyclic Redundancy Check (CRC). The preamble and the SFD are added on the physical layer that is why considered as Physical layer header, while remaining five fields are added on Data Link Layer (DLL) of OSI model therefore called DLL header. The detailed description of all fields is discussed as follows:

Preamble: is of 7 Bytes and are used to notify/alert the receiver that the new frame has started (The pattern of the preamble is 101010....).

SFD: is of 1Byte long which differentiates the Physical layer header from DLL header (SFD pattern: 10101011, the last two bits "11" tells the receiver that PL header ends here).

Destination MAC address: 6 Bytes long

Source MAC address: 6 Bytes long

Length: 2 Bytes padding (separates the data/payload from MAC addresses)

Data: The minimum size of the data is 46 Bytes and maximum size can be of 1500 bytes long.

CRC: 4 Bytes for error detection and correction.

The minimum length of the frame can be 64 Bytes (46 Bytes data) and the maximum length of the frame can be 1518 Bytes (1500 Bytes).
(**Note**: For frame size calculation, we consider only the DLL)

Figure 3.11: Ethernet Frame

3.3.1.4 *Media Access*

To access the media, the ethernet uses CSMA/CD (Carrier Sense Multiple Access/Collision Detection). Let us understand CSMA/CD with the following example.

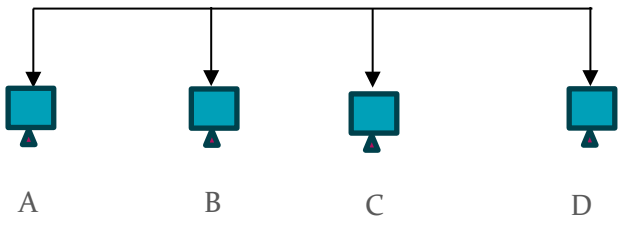


Figure 3.12: Media Access using Bus LAN

Consider the figure 3.12, assume that the computer A sends a packet to computer C and at the same time, the computer D sends a packet to computer B, since they use/share same channel therefore there would be a collision of information. All the connected computers will get a collision notification, however, since there is no acknowledgment in ethernet so, none of them knows whether their information was involved in the collision or not. To improve the media access technique for ethernet, IEEE proposed CSMA, where the computers who want to send the data/packets will firstly check/listen the channel and if there is no traffic then they will send their packet over the channel. However, after a computer checks the channel and sends its information, if another computer broadcasts, then again there will be a collision. To detect the collision (so, the victim computers can retransmit the data), if computer A sends a packet to computer C then it will sense/checks the channel and keep listening to the channel (length of the packet) unless the packet gets to the destination computer. In that way, if any collision occurs (the length of the packet) then the victim computers will retransmit the data. (CSMA/CD)

3.3.1.5 Ethernet Standards

The ethernet channels are usually described as X Base Y (where X represents the data rate (Mbps) and Y represents the distance/description of the cable e.g. 10 Base T (10 Mbps using Unshielded Twisted Pair Cable with 100 meters range) similarly 10 Base 2 (10 Mbps over 200 meters). There are three (3) standards of the ethernet: Standard (10 Mbps), Fast Ethernet (100 Mbps) and Gigabit Ethernet (1 Gbps and up). The table 3.1 summarizes the various ethernet medium standards.

Channel	**Peak Data rate (Mbps)**	**Standard**	**Distance**
Cat 3	*10*	*10 Base T*	*100 meters*
Cat 5	*100*	*100 Base T*	*100 meters*
Cat 5e	*1000*	*1000 Base T*	*100 meters*
Cat 6	*1000*	*1000 Base T*	*100 meters (55 meters for 10 Gbps)*
Cat 7	*10 Gbps*	*10X Base T*	*1000 meters*

Table 3.1: Ethernet Standards

3.3.2 Wifi

In this section we will discuss:

- Overview of the Wifi
- Modes of Operation
- Channels/Frequencies
- Standards of Wifi

- Security Protocols

3.3.2.1 *Overview of the Wifi*

The full form of Wifi is Wireless Fidelity and it was invented by IEEE in 1997 (IEEE-802.11). The main objective of the Wifi is to support mobility over the LAN and offers a Wireless Local Area Network (WLAN).

3.3.2.2 *Modes of Operation*

The Wifi offers two modes of operations: Infrastructure and Adhoc. The infrastructure mode uses centralized device (Access Point) to connect all the devices/computer to form a LAN. In the Adhoc mode, there is no central device and peer to peer (p2p) network gets established using mobile hotspot.

3.3.2.3 *Channels/Frequencies*

The WLAN (IEEE-802.11) routers operates over two (2) frequencies bands: 2.4 GHz and 5 GHz, and called Dual Band routers. The 2.4 GHz offers 14 Channels (3 usable) and 5 GHz offers 42 channels with less electromagnetic interference from household devices. The table 3.2 summarizes the different standards of Wifi.

Standard	**Peak Data rate (Mbps)**	**Band**
IEEE-802.11	*11*	*2.4 GHz*
IEEE-802.11a	*54*	*5 GHz*
IEEE-802.11g	*54*	*2.4 GHz*
IEEE-802.11n	*65-600*	*2.4/5 GHz*
IEEE-802.11ac	*1.3 Gbps*	*2.4/5 GHz*

Table 3.2: WLAN Standards

The IEEE-802.11n uses MIMO (Multi Input and Muti Output) and offers an enhanced internet speed up to 600 Mbps. The IEEE-802.11ac improves the MIMO and incorporates MU-MIMO (Multi-user MIMO) and offers peak speed up to 1.3 Gbps.

3.3.2.4 *Wifi Security Protocols*

- *WEP:* In 1998, to improve the security of the WLAN channels and avoid sniffing, IEEE proposed a new protocol WEP (Wired Equivalent Protocol) which encrypts the data before transmission over the insecure (air) channel.
- *WPA and WPA2:* The WEP receives numerous attacks right after its introduction, therefore in 2003, IEEE proposed a new protocol WPA (Wifi Protected Access) to combat against security attacks. Besides the

encryption, WPA also involves access code (key) to avoid the piggy backers from joining the network. However, the encryption algorithm proves to be extremely weak and reported to be vulnerable against many disclosure attacks which leads towards development of WPA2. The WPA2 keeps the same access control mechanism while replaces the encryption schemes (uses AES as encryption algorithm). The incorporation of AES significantly improves the security of the WPA2 and avoids all of the existing attacks. The WPA2 enterprise uses an authentication server (RADIUS, TACACS etc.), offers the captive portal and provides the centralized access control and maintains the logs.

Chapter 4

TCP/IP PROTOCOL

In this chapter, we will discuss TCP/IP protocol, its working and layered architecture. After this chapter, the students will be able to:

- Understand the fundamentals of network protocol
- Analyze the applications of TCP/IP protocol
- TCP vs UDP protocols

4.1 TCP/IP Protocol

The protocol is the set of rules for data communication between two or more computers. The OSI model was proposed in 1976 and have never been implemented and used as a reference model for network communication (The OSI will be discussed in Chapter 6). The TCP/IP protocol is a suite of protocol that allows two or more computers to communicate over the network. The TCP/IP protocol comprises of four layers (Application, Transport, Internet, and Network access). The figure 4.1 presents the TCP/IP protocol with the TCP frame formation.

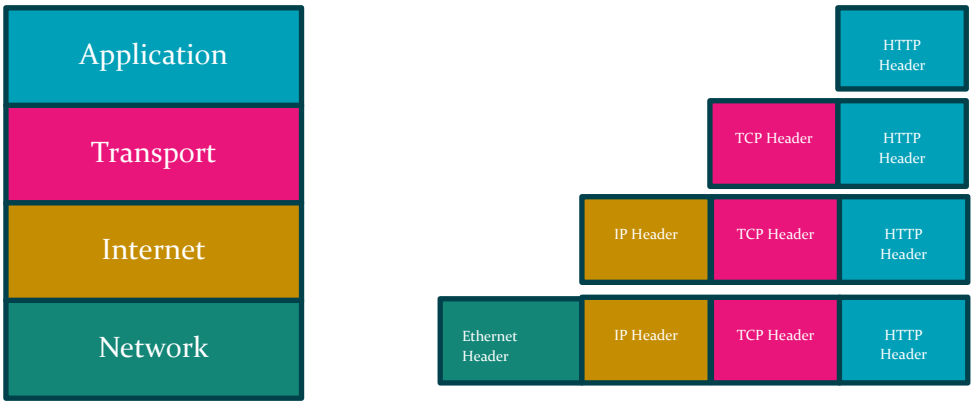


Figure 4.1: TCP/IP Protocol layered architecture

The end user applications run on the application layer of the TCP/IP protocol e.g. Browsers and web applications etc., and it provides the user interaction. Let us understand the roles of all of the four (4) layers with the TCP frame formation. Assume that user wants to access a webpage, then on the application layer, application will create a **HTTP header** (request for the webpage). If user wants to download/upload then, it will create the FTP header so, it mainly depends on the user application and request. The HTTP Header (comprises of the request) will go to transport layer, where TCP (Transmission Control Protocol) and UDP (User Datagram Protocol) run. If the request comprises of TCP protocol (which it does, since we are using HTTP), then transport layer firstly does the segmentation (turns the data into small packets) then adds the source and destination port numbers. The segment (HTTP header and TCP Header) with port numbers will go to internet layer. The Internet layer adds the **IP Header** which includes the routing information and IP addresses of the source and destination. The packet (HTTP Header, TCP Header and IP Header) finally gets to the network access layer, which adds the source and gateway MAC addresses. We add the gateway MAC address instead of destination MAC address since for LAN communication (frame delivery), we need to add the MAC address of the gateway/router.

Consider the following example to better understand the network communication between LAN device and a remote server.

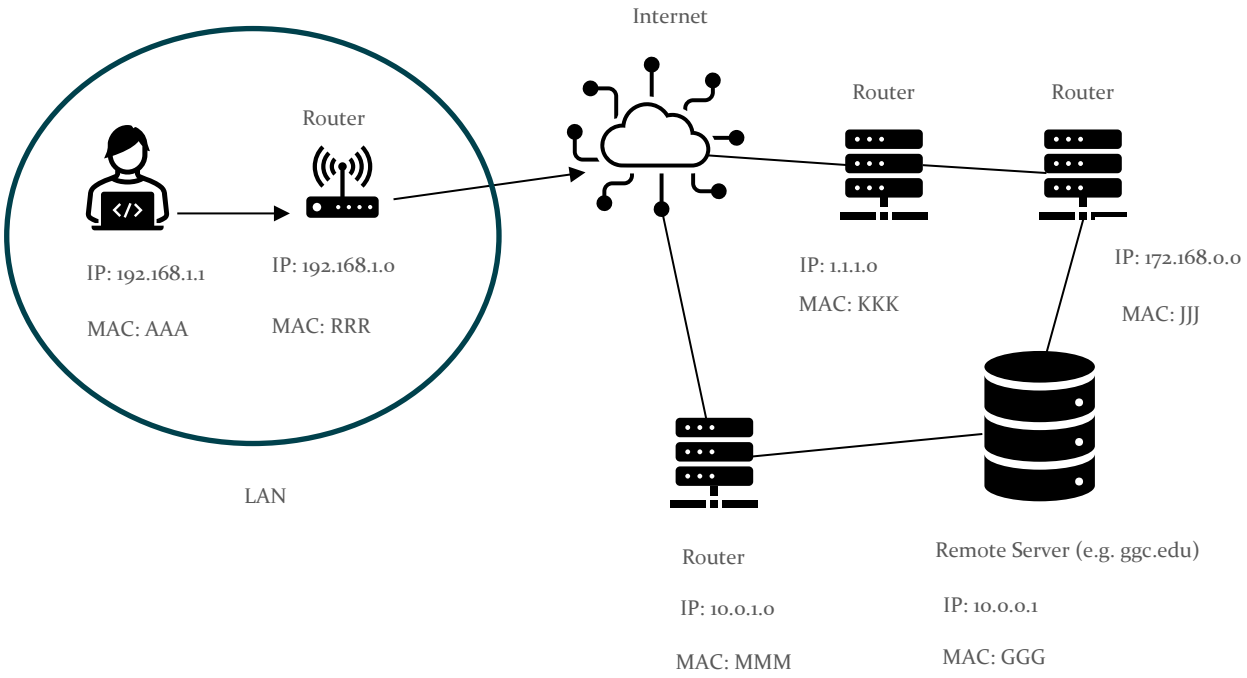


Figure 4.2: Remote Server communication

As shown in figure 4.2, assume that a user having IP address 192.168.1.1 and MAC address AAA wants to download some contents (files) from remote server having IP address 10.0.0.1 (ggc.edu). Then firstly, the user uses the application (web browser) and creates a request. As per TCP/IP protocol, the application layer will create a FTP header (FTP to download and upload the files) and forwards that to transport layer. The transport layer performs segmentation and adds the source and destination port numbers (which can be 20 or 21). Then intermediate packet will be forwarded to internet layer, which adds the source and destination IP addresses (Source IP: 192.168.1.1 and Destination IP:10.0.0.1) and other routing information. The packet will then arrive at network layer (performs datalink layer and physical layer responsibility), the network layer adds the MAC addresses of the source (AAA) and the gateway (RRR) and the finally the frame will be turned into signal for transmission. The figure 4.3 shows the complete frame. When the frames reach at the router (RRR), the router opens the frame and finds the destination IP address. The router calculates the route and finds the best/shortest path and re-encapsulates the frame and replaces the gateway MAC addresses with the optimal router (MMM) (best path), rest of the fields will remain same. The router (MMM) will finally deliver the frame to remote server (ggc.edu) which repeats the same process to send the response.

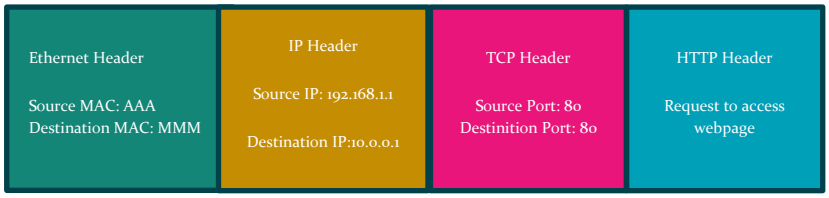


Figure 4.3: TCP Frame

4.2 TCP/IP Layered Architecture

In this section, we will discuss the details of four (4) layers of TCP/IP protocol.

4.2.1 Application layer

The application layer includes all applications and protocols that provide end user interaction e.g. Browsers, web and mobile applications etc. The detailed description of the protocols run on application layer is presented as follows:

4.2.1.1 HTTP

HTTP stands for Hyper Text Transfer Protocol, which allows the user to request from the remote web servers. The http protocol uses numerous methods to perform desired actions e.g.

Get: requests data from specified resource.

Post: submits the data to be processed by the remote server e.g. credentials.

Patch: Updates the stored data

Delete: removes the data on the remote server.

The http protocol uses Port 80 to operate.

4.2.1.2 HTTPs

The https allows the users to forward the requests in a secure manner by encrypting the user input and the response from the server. The https protocol incorporates TLS (Transport layer Security)/SSL (Secure Socket Layer) and uses asymmetric encryption to ensure the security of the user data. The https avoids the sniffing and interception and runs on port 443.

4.2.1.3 FTP

The FTP (File Transfer Protocol) is a standardized protocol for uploading and the downloading. The FTP protocol transfers the large files across the remote computers securely. The protocol runs on the port 20, 21.

4.2.1.4 DNS

The DNS (Domain Name System) protocol connects the users to the websites using their domain name instead of IP addresses. The DNS

protocol resolves the domain names into IP addresses and runs on port 53.

4.2.1.5 SMTP

The SMTP (Simple Mail Transfer Protocol) transfers the email over the internet and controls its flow. The SMTP protocol runs on the port 25.

4.2.1.6 POP3

The POP3 (Post Office Protocol 3) downloads the emails from remote email server to the local computers and its runs on port 110.

4.2.1.7 IMAP

The IMAP (Internet Message Access Protocol) provides the synchronization of the emails between different devices (e.g. outlook and the web-based emails) and IMAP runs on port 143.

4.2.1.8 Telnet

Allows the users to remotely access the devices (remote) using command terminal. Because of the insecure connectivity, the telnet protocol has been obsolete. The protocol runs on port 23 and usually closed by the OS developers to avoid the possible security threats.

4.2.1.9 SSH

The SSH (Secure Shell) also offers the remote accessing and has widely replaced the telnet protocol because of better security advantage over telnet. The SSH protocol runs on port 22.

4.2.1.10 DHCP

The DHCP (Dynamic Host Configuration Protocol) runs on port 67/68 and assigns the automatic IP addressing.

4.2.2 *Transport layer*

The transport layer performs three main actions:

- Encapsulate the data into small segments
- Adds source and destination port numbers
- Adds the ISN (Initial Sequence Number), which allows the receiver to reassemble the frame.

Both of the protocols TCP and UDP run on the transport layer. The TCP protocols (HTTP, HTTPs, FTP and SMTP etc.) are connection oriented where the sender uses SYN protocol to establish the connection with receiver. However, the UDP protocols (e.g. ICMP, ARP etc.) are connectionless, where sender sends the data without making any connection with the receiver. The UDP protocol is faster than TCP protocol, however, proves to be unreliable. The detailed description of the SYN protocol is presented as follows:

SYN Protocol:

The TCP protocols use SYN protocol to establish a connection between the sender and the receiver. The SYN protocol involves three (3) way handshake process:

- The sender sends TCP Packet with SYN flag.
- Upon receiving, the TCP packet with SYN flag, if the receiver is up (running that specific service e.g. http, ftp, smtp etc. which sender wants to access/request), the receiver responds with TCP packet with SYN and ACK flags.
- Then sender sends the TCP packet with ACK flag

After this 3-way handshake, the connection is established, and sender sends the request to the receiver. The figure 4.4 presents the SYN protocol.

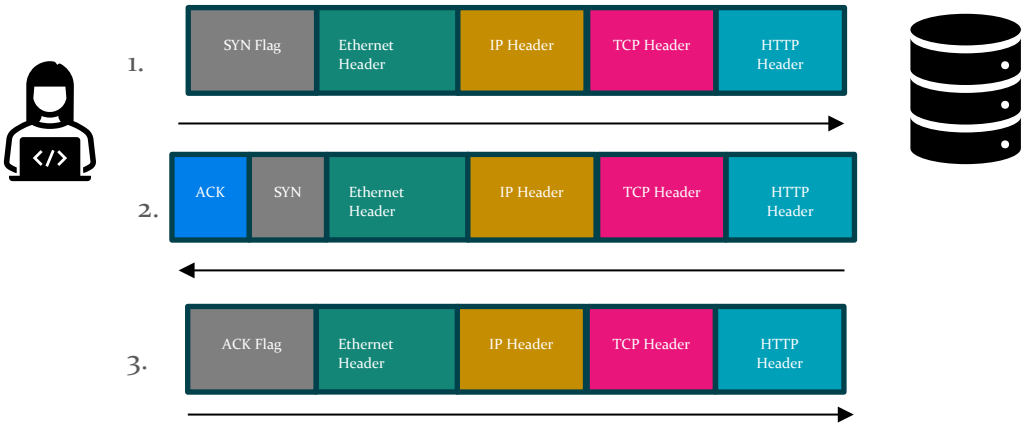


Figure 4.4: SYN Protocol

4.2.3 *Internet layer*

The Internet layer adds the source, and the destination IP addresses and only UDP protocol runs on this layer e.g. ICMP (Internet Control Message Protocol) which offers Ping and Trace route utilities, Address Resolution Protocol (ARP) and other routing protocols (RIP, OSPF etc.). The routers and layer-3 switches run on this layer which uses IP addresses and routes the packet without establishing any connection with receiver.

4.2.4 *Network layer*

The network layer behaves like physical layer (OSI) and adds the source and gateway MAC addresses to the TCP/UDP packets. The protocols such as Ethernet, Wifi etc. run on this layer.

Chapter 5

IP ADDRESSING

In this chapter, we will discuss IPv4 the IPv6 its working and major differences. After this chapter, the students will be able to:

- Understand the basics of IPv4 and IPv4 and their structure
- Compare NAT vs PAT
- Implementation of LAN/WAN using IPv4 and IPv6

5.1 IP Addressing

The IP (Internet Protocol) is the name of the computer over internet. The computers use two addresses for communication over the network: Physical Address (MAC) and the IP address. The MAC address is used for communication over the LAN while the IP address allows the computers to communicate with remote devices over WAN. In this chapter, we will discuss the IPV4 and IPV6 structure and addressing methods. The detailed description of the IP addressing is presented as follows:

5.2 IPV4

In this section, we will discuss the structure, classes and addressing methods.

5.2.1 *Structure of IPV4*

The IPV4 address is of 32-bits long which comprises of four (4) segments (each segment is 8-bits long). Since, each segment is 8-bits long therefore the value of each segment ranges from 0 to 255. The IP address depicts two IDs: Network ID and the Host ID. The network ID is fixed and represents the organization while the host ID is dynamic and assigned to any host, requesting for the IP address.

$$w.x.y.z$$

Where w,x,y,z represents the four segments/octets (contains 8-bits) e.g. if one has an IP address 192.168.10.1 then computer interpret it as: (Binary)

192.168.10.1

11000000.10101000.00001010.00000001 (Binary Equivalent)

32-bits

5.2.2 *Classes of IPV4*

The IPV4 addresses are categorized into five (5) classes, which defines the size of the network. The first octet tells the class of the IP address, the ranges of the first octet (classification of the IP address) is presented in table 5.1.

Class of IP address	**Range (first Octet)**	**Application/Usage**	**Possible Host IP Addresses**
A	1 →126	Large Network (Corporate)	$2^{24} \approx 16\,M$
B	128→191	Medium Sized Organization	$2^{16} \approx 65K$
C	192→223	Small networks	$2^8 = 255$
D	224→239	Multicasting/Research	N/A
E	240→255	Research	N/A

Table 5.1: Classes of IPV4

As we can see from the table 5.1:

- **Class A IP address**: $network.host.host.host$
  So, the first (1) octet is reserved for the network ID and remaining three (3) octets are used for hosts which allows $\approx 16\ M$ IP addresses and can form a large network.
- **Class B IP address**: $network.network.host.host$
  The first two (2) octets are reserved for the network ID and remaining two (2) octets are used for hosts which allows $\approx 65\ K$ IP addresses and can form a medium sized network.
- **Class C IP address**: $network.network.network.host$
  The first three (3) octets are reserved for the network ID and one (1) octet is used for hosts which allows $= 255$ IP addresses and can form a small network (household)

The class D and E are not used for IP addressing and are reserved for research purposes.

5.2.3 *Loopback IPV4 Address*

In IPV4, the range 127.0.0.0 – 127.255.255.255 is reserved for the loop back address (local host). Usually, we use 127.0.0.1 for the local host troubleshooting, testing and development.

5.2.4 *Network ID*

The network ID is used to identify the specific network (organization's identification) to which computers and devices belong to. The network ID cannot be assigned to the host and reserved for the network identification only. The network part is represented as '1' which represents the saturation of the network part and '0' is used to represent the host part. So, to calculate the network ID, the network part (based on the class of IP Address) remains intact, and the host part will be represented as '0'. The network ID calculation for Class A, B and C with examples are presented as follows:

Class A: If we have an IP address 15.0.0.2, then as we know it belongs to Class A and first octet is reserved for network part and remaining three (3) are assigned for host. Therefore, the network ID with which this IP address belongs to will be 15.0.0.0 (starting IP address).

Class B: If we have an IP address 173.28.1.9, then as we know it belongs to Class B and first two octets are reserved for network part and remaining two (2) are assigned for host. Therefore, the network ID with which this IP address belongs to will be 173.28.0.0 (starting IP address).

Class C: If we have an IP address 193.28.1.10, then as we know it belongs to Class C and first three octets are reserved for network part and last one is assigned for host. Therefore, the network ID with which this IP address belongs to will be 193.28.1.0 (starting IP address).

5.2.5 *Broadcast ID*

The broadcast ID is used to broadcast the packet to all the nodes connected to specific LAN. Unlike network ID, to calculate the broadcast ID, we saturate the host part of the specific IP address. The broadcast ID calculation for Class A, B and C with examples are presented as follows:

Class A: If we have an IP address 15.0.0.2, then as we know it belongs to Class A and first octet is reserved for network part and remaining three (3) are assigned for host. Therefore, the broadcast ID will be 15.255.255.255 (last IP address).

Class B: If we have an IP address 173.28.1.9, then as we know it belongs to Class B and first two octets are reserved for network part and remaining two (2) are assigned for host. Therefore, the broadcast ID will be 173.28.255.255.

Class C: If we have an IP address 193.28.1.10, then as we know it belongs to Class C and first three octets are reserved for network part and last one is assigned for host. Therefore, the broadcast ID will be 193.28.255.255.

5.2.6 *Subnet Mask*

The subnet mask differentiates the network part from the host part of a specific IP address. The subnet mask shows the saturation of the IP address where the network part of the IP address is represented as all eight bits as '1' and '0' is used to represent the host part. For class A, the subnet mask for any specific IP address is 255.0.0.0, for class B 255.255.0.0 and for class C 255.255.255.0.

Example: Calculate the network ID, Broadcast ID and the subnet Mask for 192.168.1.30?

Network ID: 192.168.1.0
Broadcast ID:192.168.1.255
Subnet Mask: 255.255.255.0

5.2.7 *Private IP addressing*

As we have discussed, devices need to have a specific IP address to communicate over the internet. However, we can generate $2^{32} \approx 4.2\ Billion$ IP addresses using IPV4 (32-bits) which do not satisfy the need of modern computational environment (IoTs etc.). Therefore, we use private IP addressing to design the LANs, these private IP addresses are used only for the LAN and cannot communicate outside the LAN. The ranges of private IP addresses of Class A, B and C are as follows:
Class A→ 10.0.0.0 – 10.255.255.255
Class B →172.16.0.0 – 172.31.255.255
Class C→192.168.0.0 – 192.168.255.255

5.2.8 *NAT*

The NAT (Network Address Translation) translates the public IP address into private IP address and otherwise. The NAT allows multiple private IP addresses to use a single public IP address and avoid the need of extensive IP addresses. To understand the NAT, let us take an example.

Assume that we have three hotels: A, B, and C. Hotel A is located in Georgia, Hotel B is in Alabama, and Hotel C is in Florida. Each hotel has ten (10) rooms. If we assign room numbers 101, 102, …, 110 to Hotel A, we can use the same numbering for the other two hotels. If someone from Hotel B wants to send a letter to someone in room 101 at Hotel A, they only need to know the address of Hotel A. NAT uses a similar concept. It creates a LAN using private IP addresses (reserved for LAN) and uses a single public IP address to connect its LAN to the internet. Figure 5.1 illustrates the NAT concept.

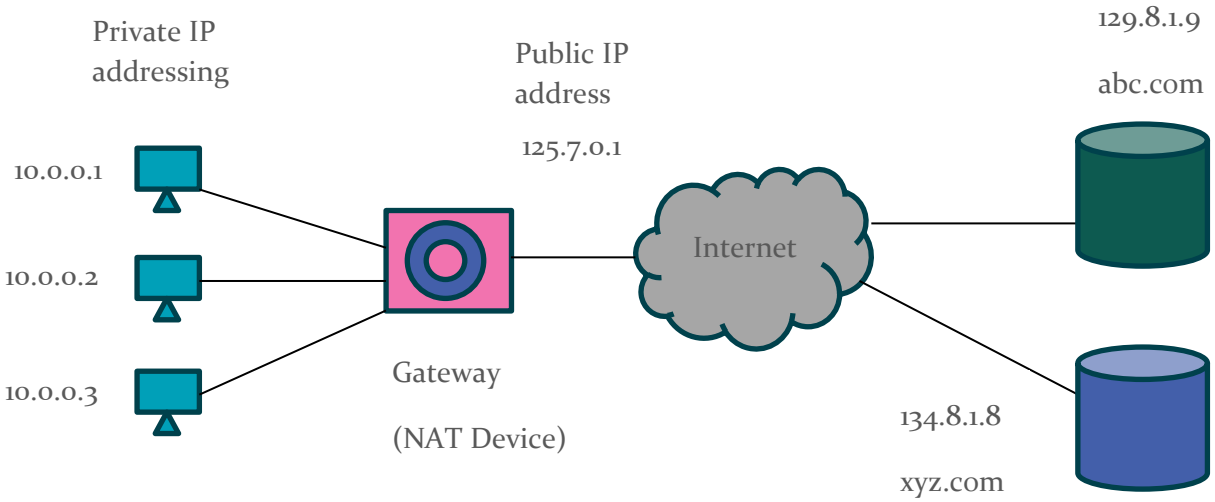


Figure 5.1: NAT Network Model

As shown in the figure 5.1, we have designed a LAN using Class A private IP addressing (10.0.0.0). Assume that we have three computer assigned private IP address (10.0.0.1 – 10.0.0.3) and computer A (10.0.0.1) wants to access abc.com then this request will go to gateway (NAT device, having Public IP address 125.7.0.1) which will update NAT table and sends the request to abc.com using it public IP address (acts a proxy). Upon receiving the request, the abc.com (assumes that the request came from 125.7.0.1) resolves the request and sends the response back to NAT device. The NAT device uses the NAT table and forward the response back to the specific computer (10.0.0.1) who originates the request. The NAT table for the specific example is presented as below:

Private IP address	Public IP address
10.0.0.1	129.8.1.9 (abc.com)

Table 5.2: NAT Table

5.2.9 *PAT*

The PAT (Port Address Translation) is the specific type of the NAT. In PAT, besides the public and private IP address, the port numbers of private and public devices are also stored onto NAT table. This avoids the problem of broadcasting and allows the NAT device to distinguish between the traffic of various devices of the LAN. For example, if two computers (10.0.0.1 and 10.0.0.2) requests to xyz.com simultaneously then to distinguish their traffic, the NAT device will use different ports (specific port for each device) e.g. port 5532 to forward 10.0.0.1 and port 4567 for 10.0.0.2. This will allow the NAT device to forward the response to the correct destination. The table 5.3 presents the PAT for the above-mentioned scenario.

Port Number (Gateway)	Private IP address	Public IP address	Port Public Device (xyz.com)
5532	10.0.0.1	134.8.1.8	8080
4567	10.0.0.2	134.8.1.8	4444

Table 5.2: NAT (using PAT) Table

The PAT will also allow the private devices to use multiple services from the remote servers, which can be differentiated with private device port numbers.

5.3 CIDR

The CIDR (Classless Inter Domain Routing) offers way more flexibility than the class-oriented IP addressing (e.g. Class A, B and C etc.). For example, we want to design a small LAN with 20 computers then we need only 20 IP addresses. If we use the class-oriented IP addressing e.g. Class C private IP address 192.168.1.0 then we will have 254 IP addresses for the hosts (192.168.1.0 is reserved for network ID and the 192.168.1.255 is the Broadcast ID). Since, we just need 20 IP addresses, therefore 234 IP addresses will be wasted (Unused IP addresses). The CIDR offers more flexibility and avoids unused IP addresses. The CIDR:

- Offers no classes of the IP addresses (no restriction)
- Uses Blocks (32 bits) which have two parts (Network ID/Block ID and Host ID). The figure 5.1 presents the CIDR block

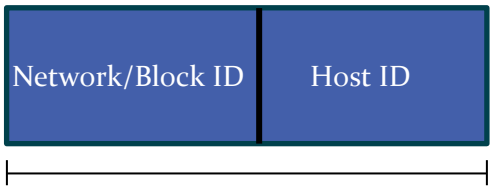


Figure 5.2: CIDR Block

- CIDR Notation: $w.x.y.z/n$ where n indicates the number of network bits (which are represented as 1).

Let us understand CIDR IP addressing using the following example.

Assume that we have an IP address 200.20.10.40/28 then:

a) Find the subnet Mask?
b) Calculate the Network ID?
c) Calculate the Broadcast ID?
d) How many usable (host) IP address can be generated?

a) Find the subnet Mask?

To find the subnet mask for IP address 200.20.10.40/28, we know that there are 28 bits for network part and only 4 bits are left for host. We will write 28 bits (all 1's) and use 0 for the host, as shown follows:

11111111.11111111.11111111.11110000

The decimal value for the IP address will be:

255.255.255.240

b) Calculate the network ID?

To calculate the network ID, first three octets of the IP address will remain same, and we will write the binary for the last octet:

200.20.10.00101000

Now, as we know the last four bits 0010 are reserved for network (to make it 28 bits network) and since the first four bits of last octet are reserved for host so, we will make first four bits to 0000 to find the network ID.

200.20.10.00100000

The decimal value will be:

200.20.10.32

c) Calculate the Broadcast ID?

To calculate the broadcast ID, we will make all the host bits to 1:

200.20.10.00101111

The decimal value will be:

200.20.10.47

d) Number of usable IP addresses?

Since, we have only four (4) bits for host: $2^4 = 16$, where two (2) IP addresses 200.20.10.32 and 200.20.10.47 are reserved for network ID and broadcast ID respectively. Therefore, we will have 14 usable IP addresses.

5.4 IPV6

In this section, we will discuss the IPV6 structure, features and transition methods. The IPV6 offers the enormous amount of IP addresses which completely avoids the shortage of IP addressing problem (IPV4). You may have a question, why don't we have IPV5? The answer is, we do have IPV5 but that is internet streaming protocol and have not designed for IP addressing. The basic features/properties of the IPV6 is described as follows:

- The length of IPV6 is 128 bits and it has divided into eight (8) segments where each segment is of 16 bits (2 Bytes).
- IPV6 doesn't use any classes (Classless).
- No Need for NAT and PAT
- No broadcasting (offers multicasting)
- Expressed in Hexadecimal
- Auto Configuration (if DHCP server is not working)
- No Backward compatibility with IPV4

The hexadecimal numbers are composed of four (4) bits. The table 5.1 presents the hexadecimals numbers with decimal and binary values:

5.4.1 Hexadecimal to binary conversion

To understand the hexadecimal to binary conversion, let us understand the following example:

Hexadecimal number: AE2F

Binary: 1010111000101111 (A→1010, E→1110, 2→0010, F→1111)

5.4.2 Binary to hexadecimal conversion

To understand the hexadecimal to binary conversion, let us understand the following example:

Binary number: 1010111000101111

Firstly, we will split it into chunks of four (4) bits and then find their hexadecimal value using the table 5.1.

Hexadecimal: 1010 1110 0010 1111 (A→1010, E→1110, 2→0010, F→1111) = AE2F

Decimal	Hexadecimal	Binary
0	0	0000
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
10	A	1010
11	B	1011
12	C	1100
13	D	1101
14	E	1110
15	F	1111

Table 5.4: Hexadecimal number

5.4.3 *Acronyms/Notations of IPV6*

- Assume that we have a IPV6 address: **(Scenario 1)**

$$FFFF:0890:CDAE:0000:0000:AE02:FFFF:A123$$

If we have two or more consecutive segments of IP address having all zeros then we can use"::" notation to represent the consecutive zero segments:

$$FFFF:0A890:CDAE::AE02:FFFF:A123$$

- **Scenario 2**

$$FFFF:0890:0DAE:0000:0000:AE02:FFFF:A123$$

If we zero on the left side, then it can be ignored in IP address notation, so the above IP address can be written as:

$FFFF:890:DAE::AE02:FFFF:A123$

- **Scenario 3**

  0000:0000:0000:0000:0000:0000:0000:A123

  The above IP addresses can be shortened to:

  ::A123

5.4.4 *IPV6 Loopback address*

In IPV6, we use the following address:

0000:0000:0000:0000:0000:0000:0000:0001

It can be shortened as:

0:0:0:0:0:0:0:1 *or* ::1 *or* ::1/128 *(where 128 represents all the network bits)*

5.4.5 *IPV6 Mapping IP*

To map (convert) the IPV4 into IPV6, we use following IPV6 address:

::FFFF:0:0/96

Where 96 bits are reserved for network part and 32 bits are reserved for the IPV4 address.

5.4.6 *Link Local Address*

As mentioned above, IPv6 performs autoconfiguration (dynamic) in a LAN. If the DHCP server is down or unavailable, IPv6 uses the address range FE80::/10 for link-local addressing and performs Automatic Private IP Addressing (APIPA).

5.4.7 *Mapping of IPV4 into IPV6*

To map IPv4 onto IPv6, we use ::FFFF:0:0/96, where the last two segments (16 bits) are used for IPv4. To better understand, let us consider the following example:

Example: Calculate the IPV6 equivalent of (IPV4) 201.192.32.40?

The mapping of IPV6 to IPV4 involves four steps:

a) IPV6 address for mapping

   0000:0000:0000:0000:0000:FFFF:0000:0000/96

b) Calculate the Binary of the IPV4 (to be mapped):

   201.192.32.40

   201→11001001

   192→11000000

32→00100000
40→00101000

c) Split each octet into 2 segments (4 bits) and calculate hexadecimal value:

201→11001001→1100 1001 →C9
92→11000000→1100 0000→C0
32→00100000→0010 0000→20
40→00101000→0010 1000→28

Hexadecimal

Split (2 segments)

d) Add the IPV4 hexadecimal values into last two segments of the mapping equation:

$$0000{:}\,0000{:}\,0000{:}\,0000{:}\,0000{:}\,FFFF{:}\,C9C0{:}\,2028$$

The above IP address can also be written as:

$$::FFFF{:}\,C9C0{:}\,2028/96$$

5.4.8 Transition Methods

To support IPV6 and IPV4 (smooth transitioning), there are two methods: Dual Stack Router and Tunnelling methods.

- Dual Stack Routers
  The dual stack routers support both IPV6 and IPV4 infrastructures and allows the devices to communicate using either of the protocols without using any transition method.
- Teredo Tunnelling
  The teredo tunnelling connects IPV4 device to IPV6 infrastructure by encapsulating the IPV4 address into 128 bits IPV6 address. The figure 5.2 presents the IPV4 conversion into IPV6. The teredo tunnelling is used to establish a direct connection between IPV4 device to remote IPV6 device/infrastructure.

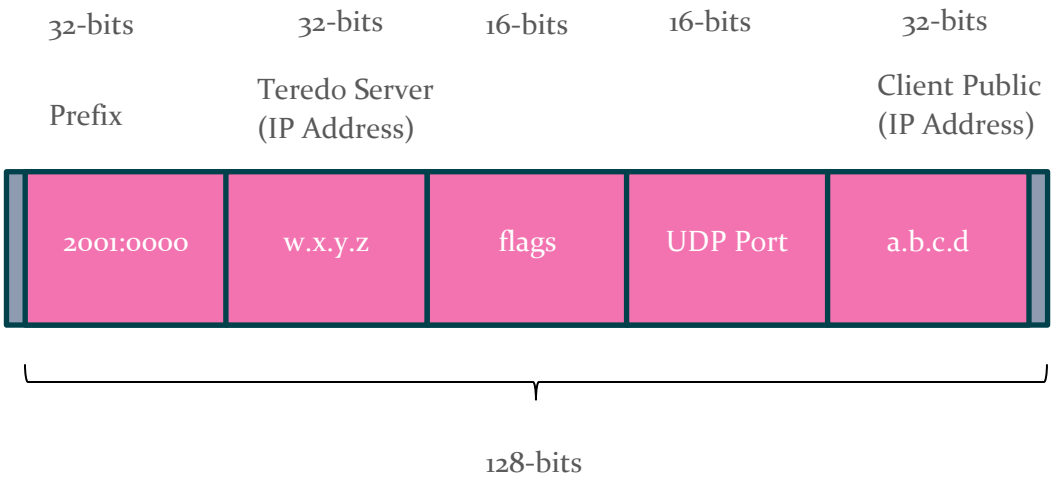


Figure 5.3: Teredo Tunnelling

Chapter 6

OSI MODEL

In this chapter, we will discuss OSI model and its functionalities. After this chapter, the students will be able to:

- Understand the working of OSI model
- Differentiate between various layers of OSI model functionalities
- Configuration and understand the DHCP and DNS protocols

6.1 OSI Model

The OSI (Open System Interconnection) model was introduced by ISO (International Organization for Standardization) in 1976 which allows two or more devices to communicate regardless of their architecture. The OSI model serves as reference model which has never been implemented, it just provides guidance (steps) for communication between the computers over the network. As discussed in Chapter 4, we use TCP/IP (proposed in 1986) protocol for communication (which compliance to OSI model guidance). The OSI model comprises of seven layers (7) and each layer has a well-defined network functionality.

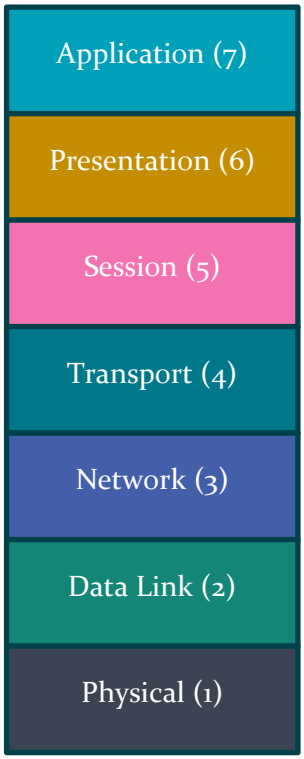


Figure 6.1: OSI Model

To better understand the OSI model and each layer functionality, let us consider the following example.

The figure 6.2 presents the OSI layers functionalities (sending and receiving side). Assume that we have two companies: ABC Inc. and XYZ Inc. The ABC Inc. is Germany based and XYZ Inc. is an American company. Both companies have seven (7) employees, the list of employees (substituted with OSI layers) is as follows:

OSI layer	Substituted Employee (ABC Inc.)	Substituted Employee (XYZ Inc.)
Application	Andy	Alex
Presentation	Pam	Parker
Session	Sam	Sky
Transport	Tom	Tim
Network	Nancy	Nadi
Data Link	David	Danny
Physical	Pat	Pablo

Table 6.1: OSI layer vs Employee (Example)

Let say, Andy (CEO of ABC Inc.) wrote a report of 400 pages and wants to send the report to Alex (CEO of XYZ Inc.). It involves the following steps:

1. Andy forwards the report to Pam and ask him to send it to Alex.
2. Pam finds that the report is in German language, but Alex understands only English. So, Pam translates the report into English and forwards it to Sam.
3. Upon receiving 400 pages report, Sam picks up the phone and calls sky (his counterpart at XYZ Inc.). Sam informs Sky that we will be sending a 400 pages report to your company, please let me know once you receive the report. Sam further forwards the report to Tom.
4. To ensure the delivery of the report, Tom divided the report into four (4) segments (100 pages each) and put it inside four (4) envelops. Tom sends the report to Nancy for further action on the report.
5. Nancy performs three jobs:
    a. Adds the sender (Andy's name) and the Receiver (Alex's name) (To-From) information.
    b. Calculates the best route (shortest) to destination.
    c. Forwards four (4) envelops and the shortest route to David.
6. David adds:
    a. Physical addresses
    b. CRC (Cyclic Redundancy Check) for error detection and correction.
    c. Forwards the route and the envelops to Pat
7. Pat delivers the envelops to Pablo (XYZ Inc.).
8. Pablo forwards it to Danny, who checks and makes sure all the envelops are sealed (not tampered).
9. Danny forwards it to Nadi, who ensures the addresses of the envelops to Tim.
10. Tim combines all the pages again into one (1) 400 pages report and forwards to Sky.

11. Sky informs his counterpart of ABC Inc. Sam and lets him know that we have received your report.
12. Sky forwards the report to Parker, who ensure its language (English, which Alex can understand).
13. Parker forwards the report to Alex, who can take necessary actions on the report.

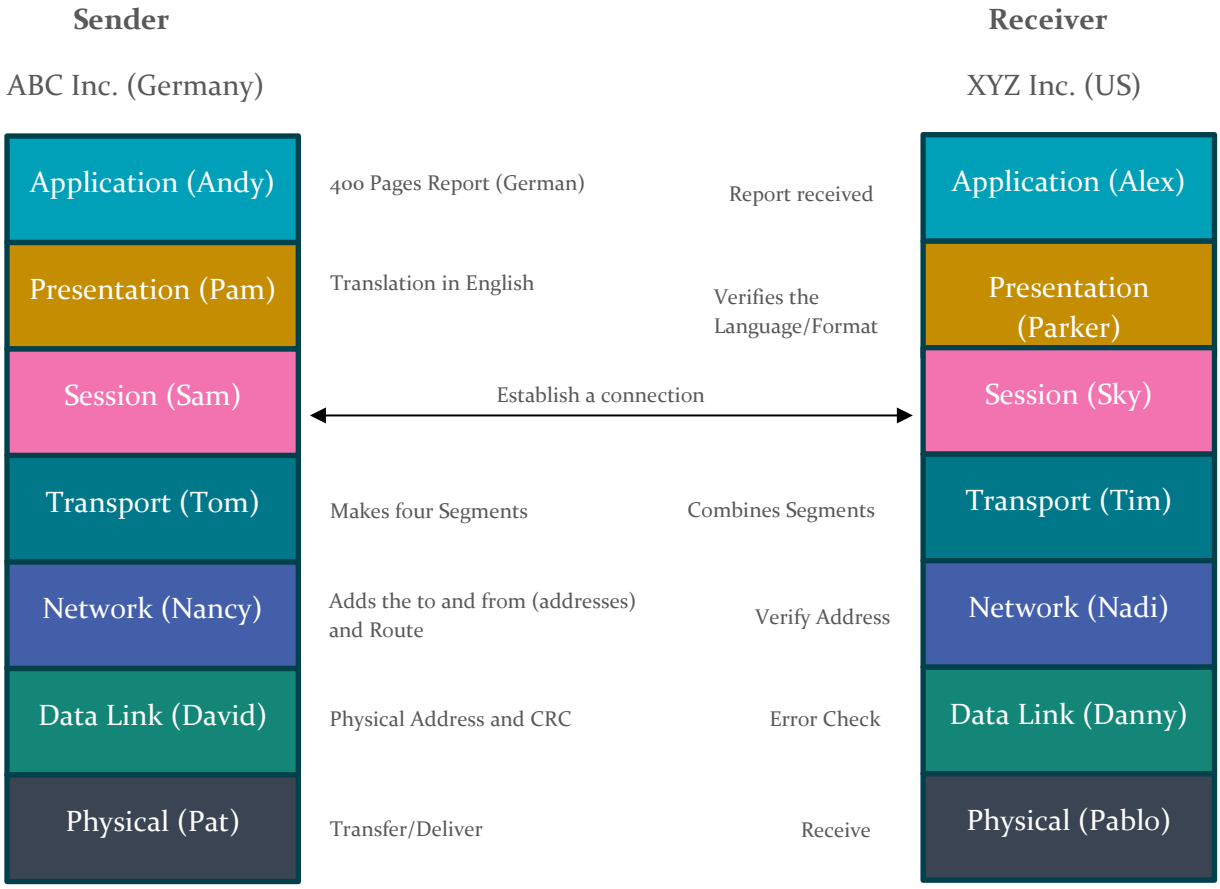


Figure 6.2: OSI Model (Communication Example)

Now, let us interpret the above steps with real-world (network) terms.

1. The application layer provides the end user interaction (runs software and applications) which allows the user to make requests. The application layer runs protocols such as HTTP, HTTPs, SMTP, FTP and DNS etc.
2. The presentation layer performs encryption, decryption, coding, decoding, compression and decompression.
3. The session layer manages the connections (sessions) e.g. login, negotiations and session cookies and tokens.
4. The transport layer performs segmentations (breaks the data into small segments) and it supports both TCP and UDP protocols.

5. The network layer performs the IP addressing, subnetting and manages the network routes. The protocols such as ICMP, ARP, IPX, RIP and OSPF run on this layer. The routers and layer 3 switches also run on network layer.
6. The data link adds the MAC addresses of the source and the gateway. We also use CRC (Cyclic Redundancy Check) for error detection and correction. The switches, Wireless Access Point (WAP) and Bridges run on data link layer and also support ethernet protocol.
7. The lowest layer (physical) includes the communication channels, hubs, repeaters and RJ 45 connectors etc. and converts the data into signals.

After the data reaches at transport layer, it is converted into segments, the segments are further converted into packet when it reaches at network layer. When we add source and destination IP addresses to the segment, then it becomes packet. This packet further goes to data link layer, where we add source and destination (gateway) MAC addresses and packets becomes frame. Finally, the frame will be converted into signals (bits) at physical layer.

6.2 DHCP

The DHCP (Dynamic Host Configuration Protocol) assigns automatic IP addresses for a specific leased time created by scope. The DHCP operates on application layer and works in client-server networking model.

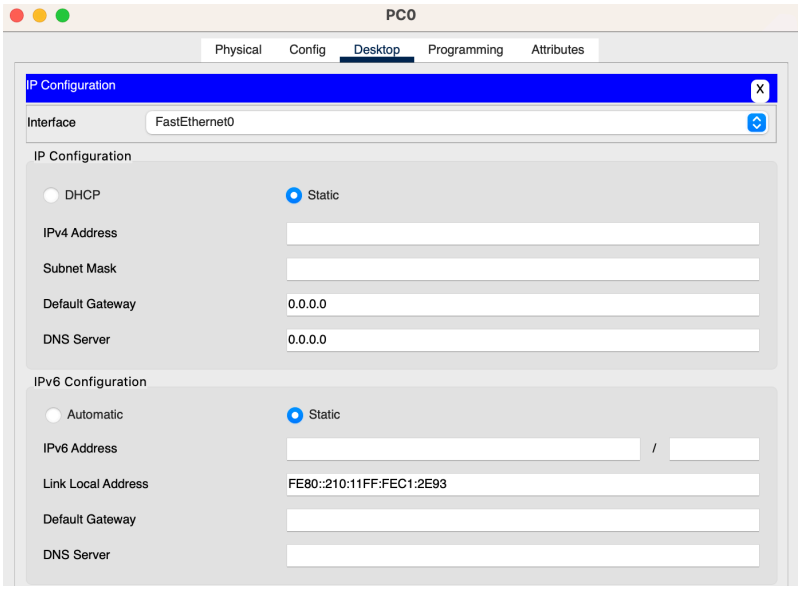


Figure 6.3: IP Configuration (DHCP)

For IP addresses configuration, the network administrators can either use static or dynamic IP configuration. In Dynamic configurations, the network administrators use A

DHCP server manages the automatic assignment of IP addresses. To better understand DHCP configuration, let's consider the following example:

Assume that the network ID of our LAN is 192.168.1.0. Since 192.168.1.0 belongs to Class C (subnet mask: 255.255.255.0), it can create up to 254 usable IP addresses. In DHCP configuration, the DHCP server randomly assigns IP addresses to host computers (requesting an IP address) for a specific lease duration. To support exclusions for printers and other LAN resources, DHCP allows reservations for static IP addresses.

If the DHCP server is unavailable or runs out of IP addresses, computers use APIPA (Automatic Private IP Addressing) with addresses in the 169.254.x.x range to communicate with other devices over the LAN. If the computers are on different subnets, a DHCP relay agent is used to facilitate communication.

6.3 DNS

For communication over the WAN, devices use IP addresses. To access any website (web server), we need to know its IP address. Since there are billions of websites on the internet, it is difficult to use IP addresses to access websites directly. Therefore, we use domain names to access websites. The DNS (Domain Name System) resolves the FQDN (Fully Qualified Domain Name) into an IP address. Figure 6.4 illustrates the working of a DNS server.

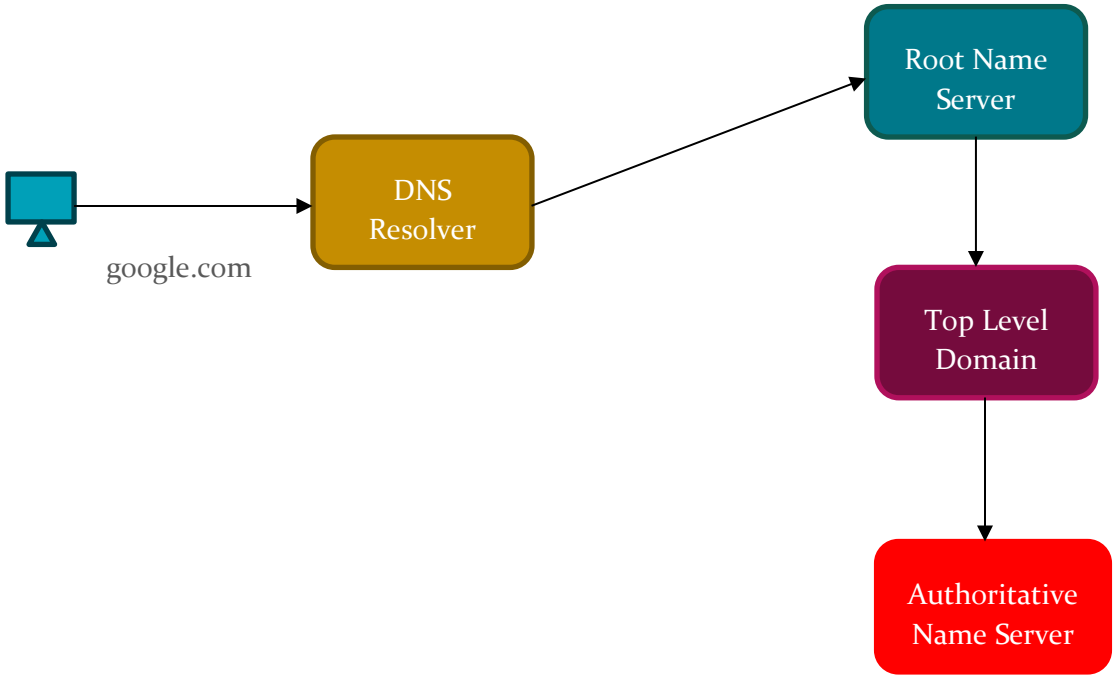


Figure 6.4: DNS Server

As shown in Figure 6.4, if a user wants to access google.com (or any other website), the request will go to the DNS resolver (ISP). If the DNS resolver has the IP address of the requested domain, it will translate the domain name into the IP address. Otherwise, the DNS resolver sends the domain name to a Root Name Server. The Root Name Server forwards the domain to the concerned Top-Level Domain (TLD), which uses the Authoritative Name Server (ANS) to find the IP address of the domain. The ANS holds the DNS records: hostname, MX, CNAME, and PTR. Figure 6.5 depicts the hierarchy of Root Servers and TLDs.

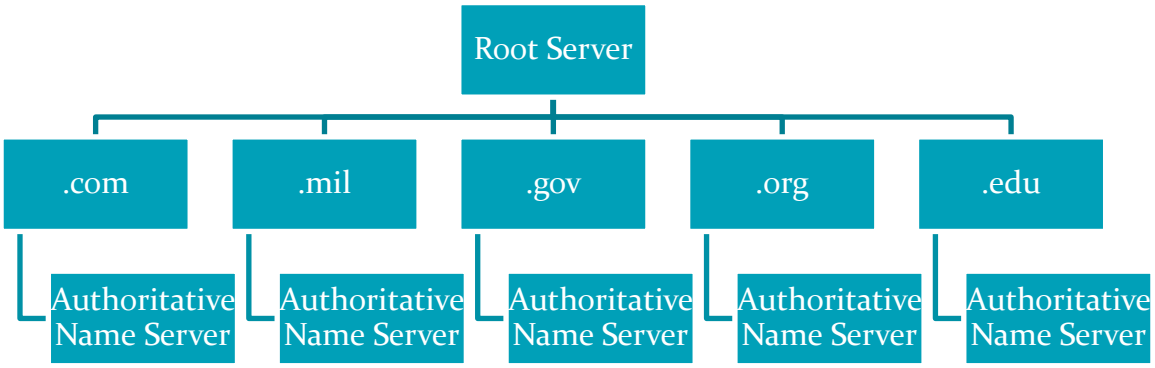


Figure 6.5: Root Server Hierarchy

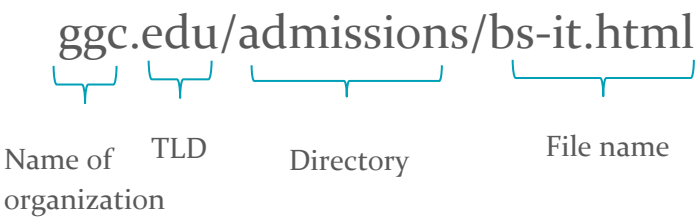


Figure 6.6: Components of domain name

Chapter 7

SUBNETTING

In this chapter, we will discuss subnetting and after this chapter, the students will be able to:

- Understand the why do we need subnetting
- Design various types of subnetting

7.1 What is Subnetting?

Dividing a large network into smaller subnets (small networks) is called subnetting. Subnetting allows network administrators to implement better security and network manageability. For example, a company with four divisions—HR, Marketing, Sales, and Production—can use subnetting to create separate broadcast domains and differentiate the internal systems of each division. This creates four different subnets, each with its own subnet and broadcast IDs, and VLANs can be established to separate these networks.

Subnetting makes IP addresses classless since it involves borrowing bits from the host part to design the subnet. By increasing the number of bits in the network part, subnetting reduces the number of host bits available.

7.2 Types of Subnetting

There are mainly two (2) types of subnetting:

- Fixed Length (Subnet)
- Variable Length (Subnet)

In the fixed length subnetting, we divide the large network into equal sized subnets, offering the same number of hosts. However, the variable length subnetting allows the different sized subnets (based on the requirement). The detailed description of both of subnetting types are discussed as follows:

7.2.1 Fixed Length Subnetting

As mentioned earlier, in fixed sized subnetting, we create same size subnets which offer equal number of IP addresses. To better understand the fixed length subnetting, let us consider the following example.

- **Example with two equal sized Subnets:**

Assume that we have a class C IP address 193.168.1.0 and we need to create two (2) subnets then calculate the Subnet ID and Broadcast ID of each subnet.
As we know, the 193.168.1.0 belongs to Class C so, therefore the subnet mask of the IP address is 255.255.255.0. So, 24 bits are reserved for the network part and 8 bits are available for the host part. Since, we need to divide this network into two subnets therefore, we will borrow only 1 bit from the host part and make it the part of the network. This will make the IP addresses, a classless IP addresses, since we have increased the network part and reduced the host part. The figure 7.1 presents the subnetting of 193.168.1.0 into two subnets.

- $Network\ ID\ (Subnet\ 1) =\ 193.168.1.0$
- $Broadcast\ ID\ (Subnet\ 1) =\ 193.168.1.127$

- *Network ID (Subnet 2) =* 193.168.1.128
- *Broadcast ID (Subnet 2) =* 193.168.1.255

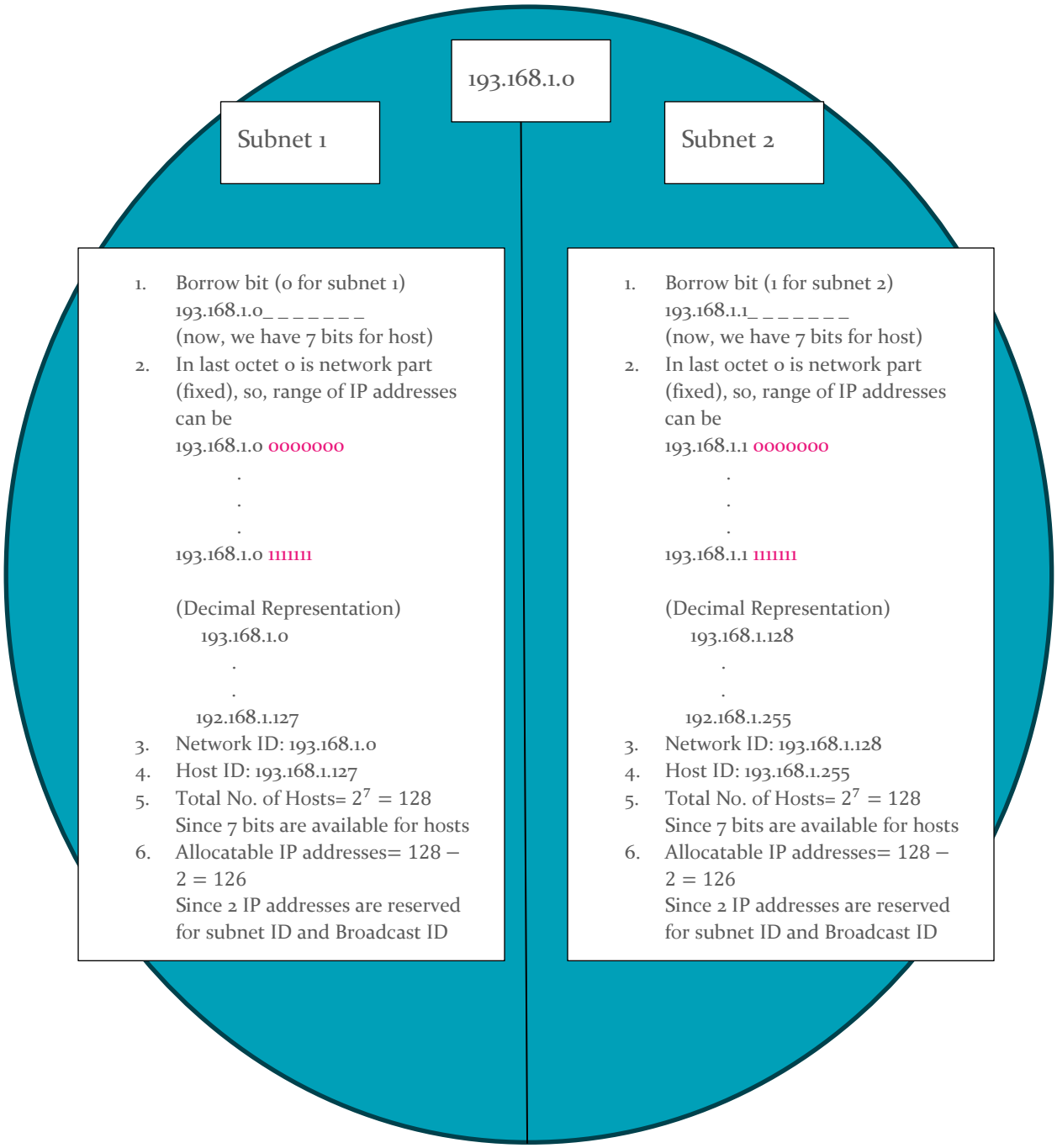


Figure 7.1: Fixed Length subnetting for two subnets

- **Example with four equal sized Subnets:**

Assume that we have a class C IP address 193.168.1.0 and we need to create four (4) subnets then calculate the Subnet ID and Broadcast ID of each subnet.

As we know, the 193.168.1.0 belongs to Class C so, therefore the subnet mask of the IP address is 255.255.255.0. So, 24 bits are reserved for the network part and 8 bits are available for the host part. Since, we need to divide this network into four subnets therefore, we will borrow 2 bits from the host part and make it the part of the network.

$$00 \rightarrow Subnet\ 1$$
$$01 \rightarrow Subnet\ 2$$
$$10 \rightarrow Subnet\ 3$$
$$11 \rightarrow Subnet\ 4$$

As mentioned earlier, the IP addresses will become classless, since we have increased the network part and reduced the host part. The figure 7.2 presents the subnetting of 193.168.1.0 into four (4) subnets.

Subnet 1

- $Network\ ID\ (Subnet\ 1) =$ 193.168.1.0
- $Broadcast\ ID\ (Subnet\ 1) =$ 193.168.1.63

Subnet 2

- $Network\ ID\ (Subnet\ 2) =$ 193.168.1.64
- $Broadcast\ ID\ (Subnet\ 1) =$ 193.168.1.127

Subnet 3

- $Network\ ID\ (Subnet\ 2) =$ 193.168.1.128
- $Broadcast\ ID\ (Subnet\ 1) =$ 193.168.1.191

Subnet 4

- $Network\ ID\ (Subnet\ 2) =$ 193.168.1.192
- $Broadcast\ ID\ (Subnet\ 1) =$ 193.168.1.255

Each subnet will create $2^6 = 64,$ IP address, where two (2) IP addresses are reserved for subnet ID and the broadcast ID.

$$Total\ No.of\ Usable\ IP\ addresses = 2^8 - 8 = 248$$

The detailed working of the subnetting example is presented as below:

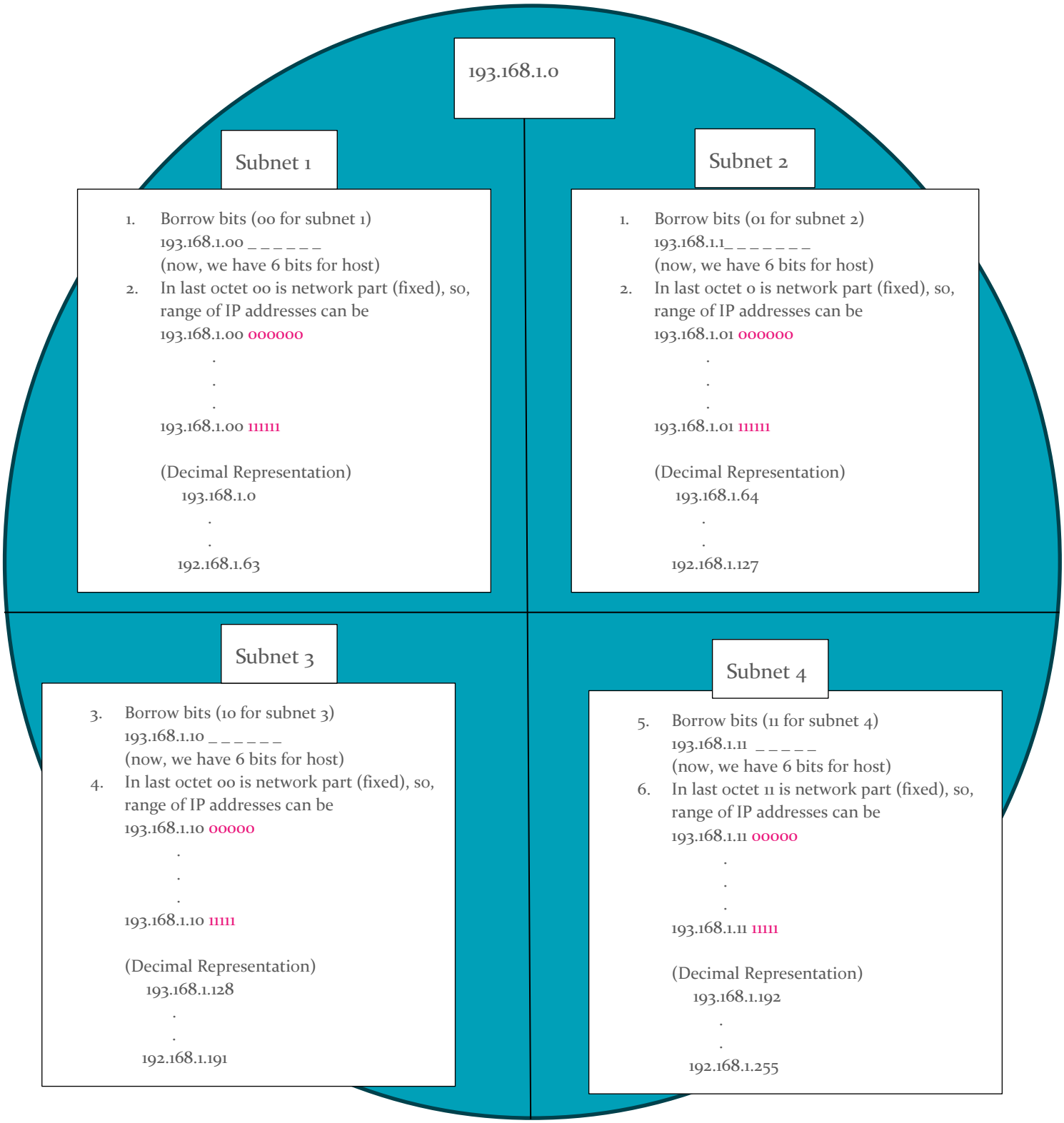


Figure 7.2: Fixed Length subnetting for four (4) subnets

7.2.2 *Variable Length Subnetting*

In the variable length subnetting, we can create subnets of different sizes offering different host IP addresses. In variable length, firstly, we divide the segments/subnets into even number subnet then further split the subnets based unevenly based on the requirement. For example, if you want to create three (3) subnets then firstly, we will divide the network into two subnets (0 for subnet 1 and 1 for subnet 2), then further divide the subnet 2 into two subnets (1 already borrowed for network ID, we will borrow another bit to differentiate further subnets); 10 and 11 for subnet 2 and subnet 3. The detailed description of the VLS (Variable Length Subnetting).

- **Example with variable length three Subnets:**

Assume that we have a class C IP address 193.168.1.0 and we need to create three (3) subnets, where the subnet 1 should offer 128 IP addresses and other two subnet should offer 64 IP addresses. Then calculate the Subnet ID and Broadcast ID of each subnet.

As we know, the 193.168.1.0 belongs to Class C so, therefore the subnet mask of the IP address is 255.255.255.0. So, 24 bits are reserved for the network part and 8 bits are available for the host part. We will firstly divide into two parts (0 and 1) then the second subnet will be further divided into two subnets (10 and 11), where each subnet offers 6 bits for host.

$$0 \rightarrow Subnet\ 1\ (7\ bits\ for\ host)$$
$$10 \rightarrow Subnet\ 2\ (6\ bits\ for\ host)$$
$$11 \rightarrow Subnet\ 3\ (6\ bits\ for\ host)$$

As mentioned earlier, the IP addresses will become classless, since we have increased the network part and reduced the host part. The figure 7.3 presents the subnetting of 193.168.1.0 into three (3) subnets.

Subnet 1

- $Network\ ID\ (Subnet\ 1) =$ 193.168.1.0
- $Broadcast\ ID\ (Subnet\ 1) =$ 193.168.1.127

Subnet 2

- $Network\ ID\ (Subnet\ 2) =$ 193.168.1.128
- $Broadcast\ ID\ (Subnet\ 1) =$ 193.168.1.191

Subnet 3

- $Network\ ID\ (Subnet\ 2) =$ 193.168.1.192
- $Broadcast\ ID\ (Subnet\ 1) =$ 193.168.1.255

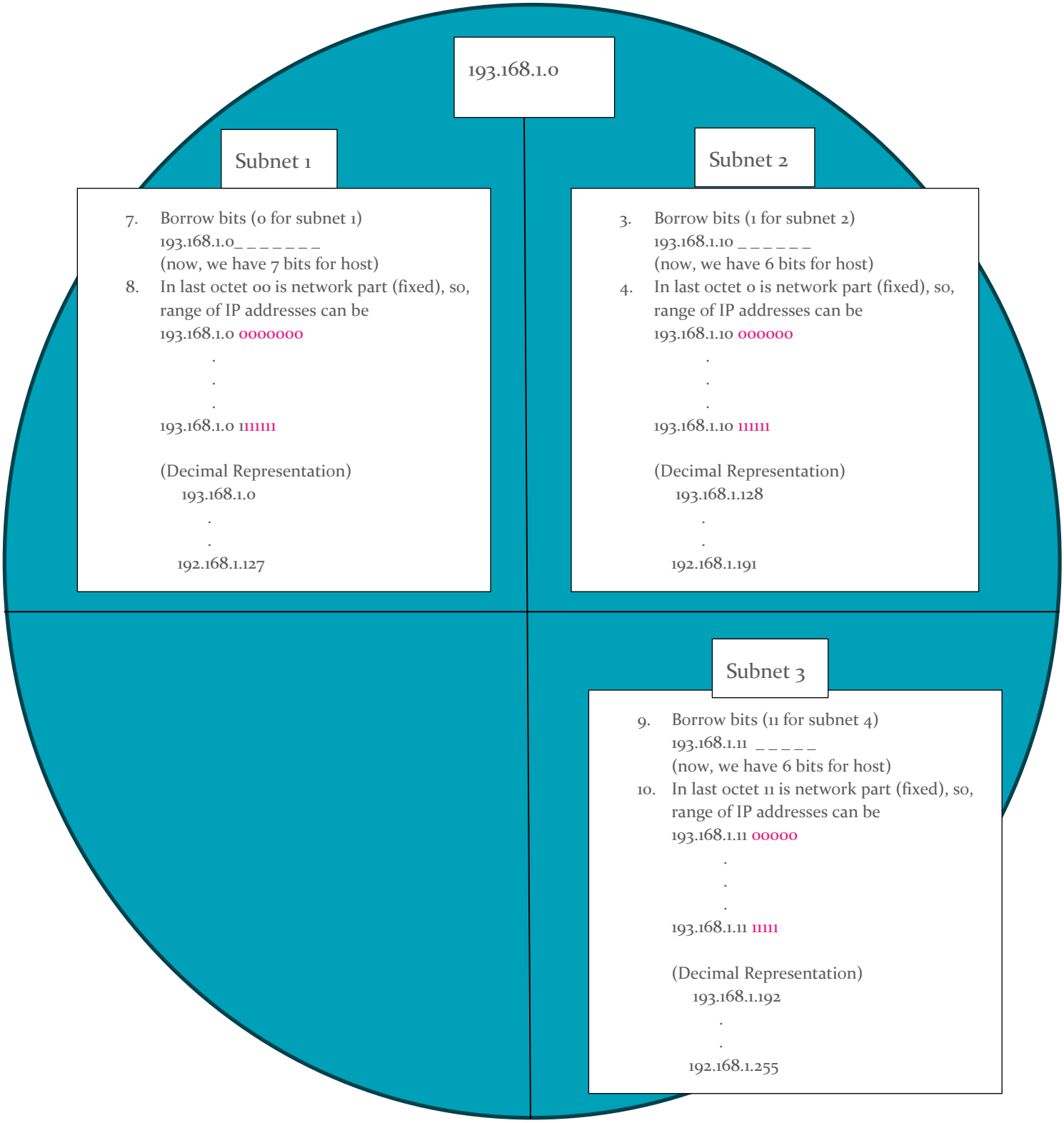


Figure 7.2: Variable Length subnetting for three (3) subnets

Chapter 8

NETWORK OPERATING SYSTEMS

In this chapter, we will discuss Network Operating Systems and its functionalities. After this chapter, the students will be able to:

- Explain the main components of an OS, the kernel, the file system and the processes
- Differentiate between client versus server operating systems

8.1Network Operating Systems Overview

The Operating System (OS) of a computer is what a user interacts with when requesting resources of a network. A Network Operating System (NOS) is one that is designed specifically to be connected to the network as opposed to a non-networked operating system. This chapter discusses the major components common to almost all operating systems. Then continues to differentiate between client and server operating systems as well as introduce the concept of virtualization.

The OS of a computer allows for convenient access for users and applications to send instructions to the computer hardware components. It controls access to the Central Processing Unit (CPU), memory (RAM), storage devices, internal and external devices such as cameras, printers, scanners etc.) and consists of 3 major components.

8.2 Components of an Operating System

The major components of any operating system include:

- Kernel
- File system
- Processes and services

8.2.1 THE KERNEL

The Kernel is effectively the manager of the Operating system. The manager schedules everyone and everything as it relates to resources that the operating system has. The kernel determines which services or processes to run, allowing for higher priority processes to run first as well as balances the request for resources with the available physical resources. The Kernel manages random access memory (RAM) and input/output devices (I/O) so that they are only accessed by one process or service at a time. It is the Kernel that performs these crucial tasks to allow the overall operating system to run smoothly. The Task Manager is an applet that one can use to see the numerous processes that the Kernel is managing as well as their resource usage.

The Kernel is commonly depicted in the Hardware Abstraction Layer (HAL) just above the hardware and includes the task scheduler, memory manager, file system manager, graphical user interface manager.

Figure 8-1 is a simplified illustration of the Windows OS structure, with the Kernel above the hardware layer.

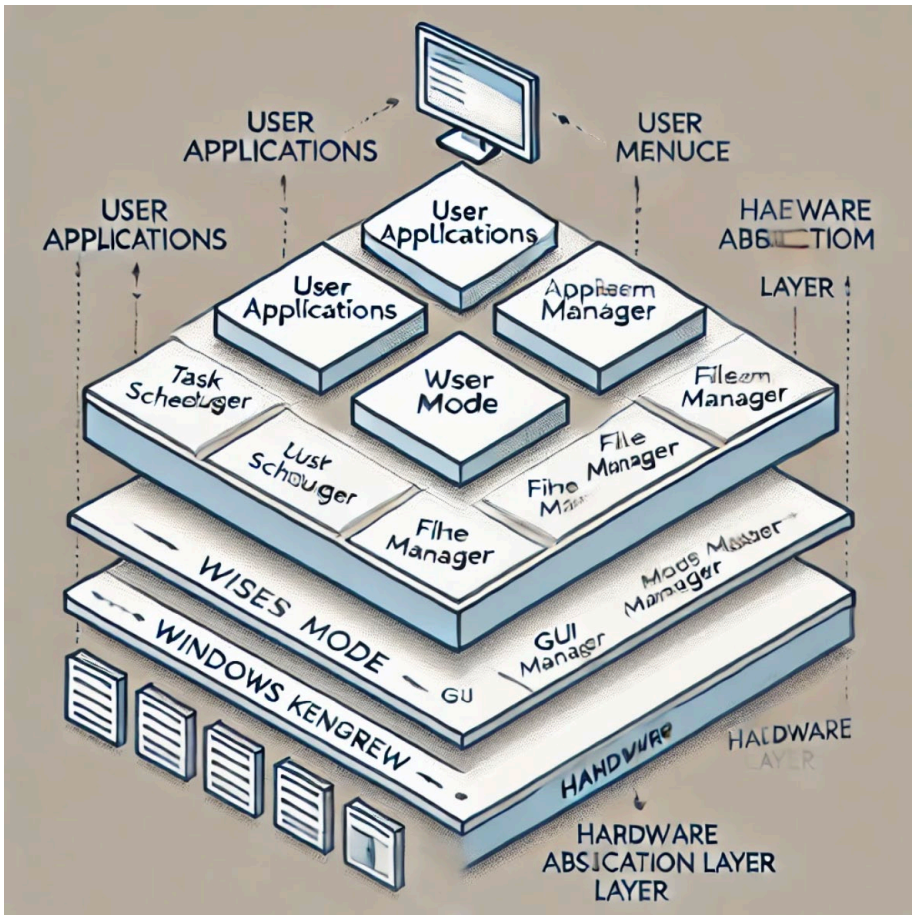


Figure 8-1 The Windows OS structure

8.2.2 FILE SYSTEM

A file system is the way that an operating system organizes and manages files that are stored whether they be on a local hard drive, networked hard drive or cloud hard drive. There are many different types of file systems, and each system allocates space for files differently. However, they all share a similar purpose which includes the following objectives:

- Provide a way for users and applications to open and save files.
- Provide an efficient method to organize space on a storage device
- Provide a structured method for filing and storing files.
- Provide an indexing system for fast retrieval of files.
- Provide a secure way to access the files for authorized users.

User Interface

The user interface is how a user requests something from the file system. When a user clicks on a file to open it, the user interface calls the file system to request to open the file. If the entity is an application, then the application will load into memory along with all associated files and dynamic libraries. If the file is a document, then the application assigned to open the file gets loaded into memory so it can open the file. When the file system does not know what type of file the user is requesting then it gives the user an option to select the application to open the file with such as the Internet browser.

When users modify a file and save it then the new file is overwritten to the old file and saved to the storage device. The File Explorer in Windows is an example of a File manager program. As you develop your knowledge and understanding you can learn about more complex file systems to allow you to set up and troubleshoot file system related issues.

Disk Drive Space Organization

The space required to store entities on a disk drive is divided into smaller chunks called "sectors". Usually, most disk drives have a size of 512 bytes per sector. Sectors are then grouped into clusters also known as a "block". A "block" is the smallest size that can be stored by a file on the disk. For example, if you have a file system that groups eight sectors to make a block, then each block is 4096 (4K) bytes. If a file that you store is only 2500 bytes then the remaining 1596 bytes is not used so effectively wasted space. This is the reason why depending on the size of our file and file types you should use correct size per sector to maximize disk usage space.

Structured Hierarchical File System

File systems typically organize files in a structured method which is usually done via folders or directories. One system uses a hierarchy for storing folders and files. The "root" of the system generally describes the top of the hierarchy or the top of the system. Under that root you will find folders or directories which then can also contain more folders and directories and ultimately where the files are placed. This structured filing method can represent a disk drive or any other type of storage independent of whether the device is local or networked.

In Windows double-clicking allows opening the folders and sub-folders to open and view their contents. The hierarchical system makes it easier for the user as well as the system to organize and maximize efficiency when attempting to find the 100s or 1000s of files that the file system usually stores. This also allows for faster indexing and searching of files to speed up retrieval of files as needed.

Figure 8-2 shows a logical diagram of a typical Windows file system.

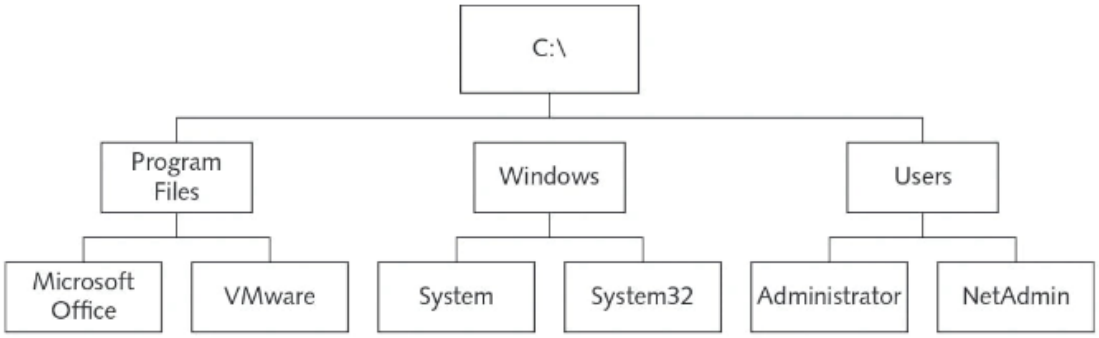


Figure 8-2 A hierarchical file system

8.2.3 PROCESSES AND SERVICES

A process is a program that is loaded into memory and run by the CPU. Normally a full application is made up of many processes and it is these processes together that help run the program such as a Web browser or a Productivity application such as Word, Excel or Powerpoint.

When a user interacts with an application, it is essentially these processes that the user is requesting processing from. If there is no interaction with the user then these entities are normally called services and runs in the background. DHCP, DNS, Windows File Sharing, Client for Microsoft Networks and many other network application layer protocols also run as services that are part of the overall operating system.

These network services are critical because they allow a computer to have features that use network resources to share with the user such as the Web browser displaying a website from the Web Server. Another example if the Domain Name Service (DNS) client which allows users to just type in the website domain name such as www.intel.com rather than must know the Internet Protocol (IP) address of the web server that holds the page.

The Task Manager is an applet that one can see all the processes and services running on the computer. It has tabs so you can view processes, performance, services, CPU usage and amount, Memory usage and amount and allows one to stop and start processes. An operating system can run 10s and 100s of processes simultaneously allowing for high levels of efficiency and multitasking.

The ability for an Operating system to run more than one process or application at once is called multi-tasking and this is why computers have become so efficient allowing one to listen to music, while working on a foreground application while the system is running various background applications all at once. Figure 8-3 shows a view of the Task Manager showing the numerous processes currently running on this system.

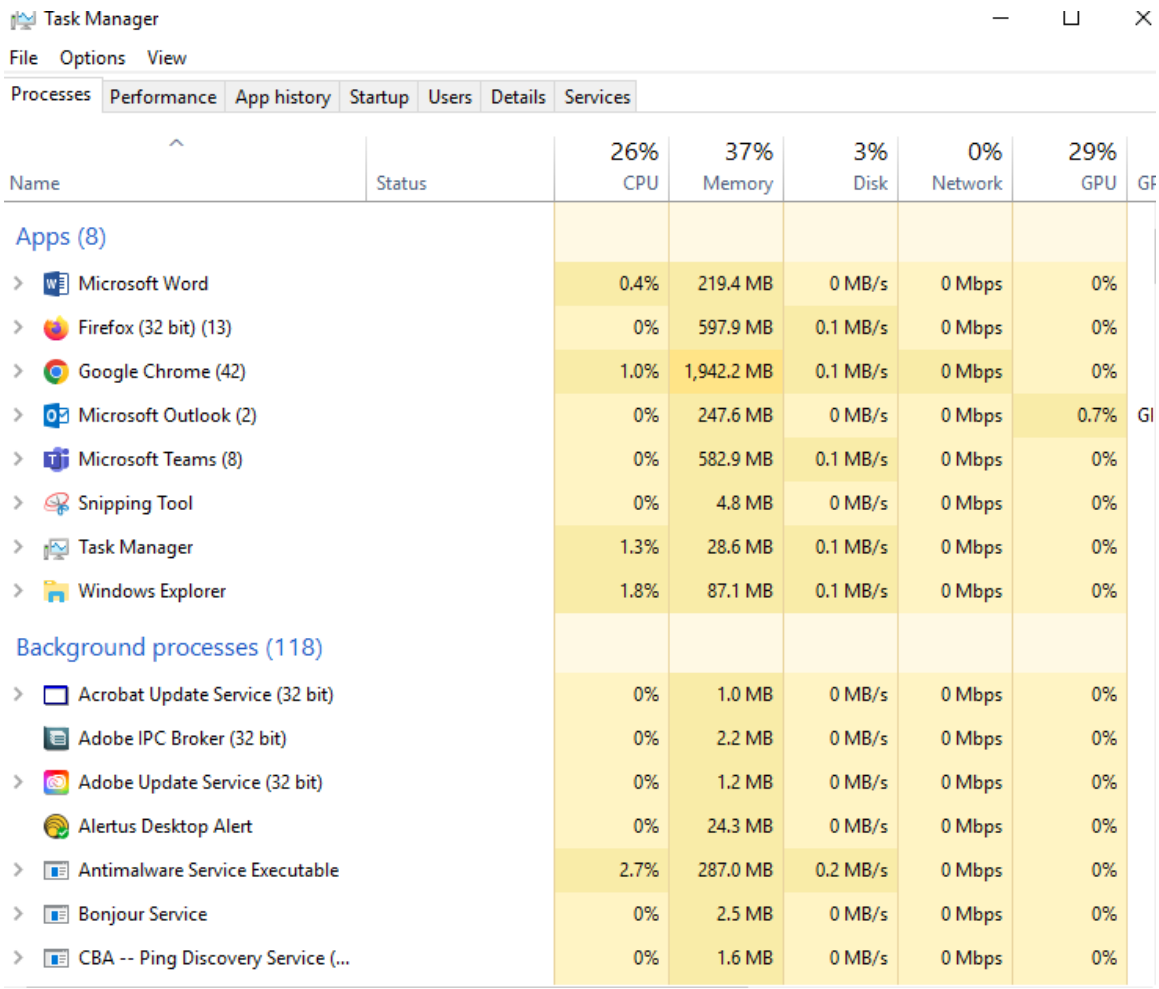


Figure 8-3 The Task Manager processes view

8.3 Client versus Server Operating Systems

8.3.1 CLIENT OPERATING SYSTEMS

A client operating system is one that is normally run on a client machine such as a tablet, laptop, desktop or stand-alone machine. Windows 10, 11 or Mac OS X are examples of client operating systems, but these now include many featured such as file and print sharing and file system security which historically are only found on server operating systems.

These days there is overlap between the client operating system and server operation system; however, the determining factor on whether one needs a client OS, or a server OS is what is the purpose of the system in the overall network.

If the purpose of the system is to “serve” other systems, then one would be better suited to use a server operating system discussed next. If the purpose is more independent and used to “connect to” another system, then the client operating system is recommended.

The client operating system is by no means limited and these days can not only run applications and processes but also allow network access through various network client features such as:

- HTTP client
- File-sharing client
- E-mail client
- DHCP client
- DNS client

Additional client software could be installed on a client operating system allowing for a myriad of applications and features but that is beyond the scope of this chapter.

8.3.2 SERVER OPERATING SYSTEMS

Traditionally a Server operating system installed on Servers in a network provided network services and functions which the client computers were not able to run. Contemporarily, the operating system installed on the client is very similar to that on the server so that distinction is now blurred. The main difference is the number and type of network services that are made available and how the resources on a Server are to be used.

For example, Windows Server 2022 had already been configured with Client for Microsoft Networks as well as DHCP and DNS client services as well as Active Directory service which are unavailable in Windows client operating systems. In Linux or Mac OS some Server operating systems are designed only for server systems.

The CPU, RAM and disk usage in Server operating systems are typically optimized to run as background network services focusing on responses to numerous client requests. They also have more security and fault-tolerant features including:

- Centralized user account and computer management
- Centralized storage
- Name resolution and address assignment
- Server and network fault tolerance

The Windows Server 2022 includes a centralized account management, Active Directory Service which includes authentication and authorization. Active Directory is a directory service that manages how and when users can log on to the network with their username and password and access resources that are allowed for them. The type of Server that runs

Active Directory is known as a domain controller, and users and computers with accounts in Active Directory are referred to as domain members. Figure 8-4 shows a screenshot of the Active Directory Users and Computers management console. The right pane shows the names, security type and description of the groups and accounts whereas the left pane shows the folders that are used to organize the accounts and resources.

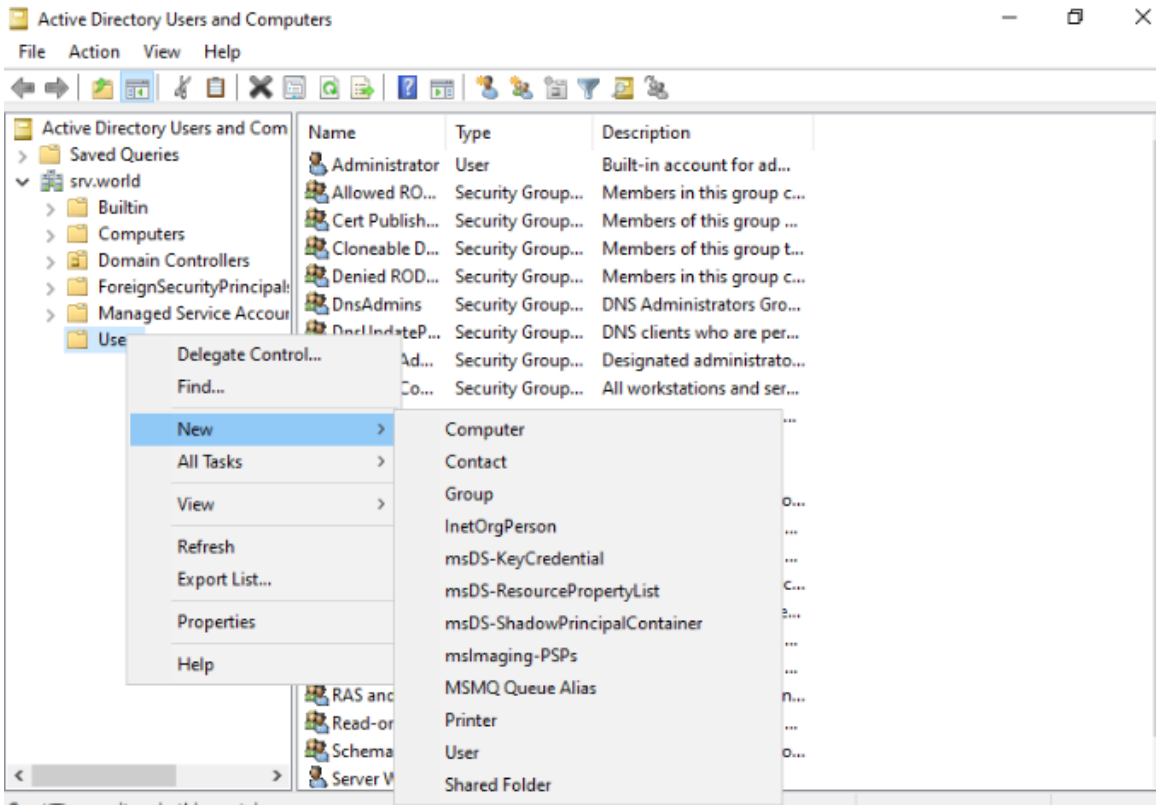


Figure 8-4 Windows Server 2022 Active Directory users and computers management console

Centralized Storage

With the massive size of files being stored and processed, network administrators are always tasked with maximizing and maintaining storage resources. Network storage includes file sharing, which means users can store documents over the network on a network server and allow others to access it. Other storage purposes include storage of email, private and public files, application databases, data backups, streaming data storage.

Initially servers used disk drives that were locally attached to run the installed Server operating system along with the applications as well as user data files. Over the years the size of these files for small to medium size business networks grew so large that managing

storage between a constant burden. Many network and server administrators made use of other new more efficient technologies such as:

- Network-attaches storage devices
- Storage area networks
- Cloud storage

Network-Attached Storage

Network-attached storage (NAS) are usually devices that include built-in networking features primarily for the purpose of sharing storage to all users on the network. A NAS could be a full server with specialized software installed along with many physical storage disk drives or it could be a network appliance without a graphical user interface and solely configured through a web browser type of interface. Many NAS devices have full integration with Windows Active Directory or other file authentication and authorization mechanisms.

Essentially NAS devices are themselves another device on the network, but with security features to allow and deny access from other network client devices such as tablets, laptops, desktops and workstations. This alleviates the need for those client devices to have big individualized storage devices attached to them and allows for centralized monitoring and maintenance.

Storage Area Network

You should already be familiar with the Local Area Network (LAN) and the Wide Area Network (WAN) as well as other entities such as the Metropolitan Area Network (MAN), Campus Area Network (CAN), Home Area Network (HAN) and Personal Area Network (PAN). Another such network is known as the Storage Area Network (SAN). The SAN is a high-speed network storage solution designed to effectively replace direct attached disk drives on servers. SAN technology allows all servers in the area with the correct permissions to connect to the SAN storage device and access resources from there. Servers could use the SAN storage device to boot their Server Operating System or Network Operating Systems.

This high-speed centralized solution offers many benefits such as better reliability and fault-tolerance than traditional storage options and essentially is the beginning of Cloud storage. Power requirements are lower, and maintenance is more centralized, lowering the probability of failures as compared to Servers with locally attached storage. Examples of this SAN technologies include Fibre-Channel and iSCSI. These network technologies allow for large arrays of disk drives to be shared and fully accessible to all Servers which in turn will service the clients that connect to them making requests. In most cases the Client computers are not even aware of the location of their data resources, not knowing if

they are on a SAN storage device or Server local storage depending in the network speed and throughput.

Cloud Storage

Many larger enterprise types of companies need more storage capabilities and even those that a SAN can offer. When this happens, the next step is to use Cloud storage. Due to physical limitations whether this is through personnel or space limitations the Cloud can be the solution.

Cloud storage is storage devices that are hosted usually by their party known as the Cloud Service Provider (CSP). The most common today are Amazon Web Services (AWS) and Microsoft Azure but there are many more cloud service providers in the market. The company essentially becomes a customer and migrates their data to the Cloud Service Provider storage devices.

There are numerous benefits to this method because the company will not need to manage and maintain which includes securing and backing up of their local storage. These tasks and no offloaded to the Cloud Service Provider. The company only must pay for the storage it uses so there is no worry about "running out of space" or having too much unused storage resulting in waste of resources. This migration allows companies to focus on their business objectives rather than focus on the Information Technology tasks of storage and services.

The customer has full control of assigning permissions and allowing user access and allocation without having to maintain the physical hardware. The budget model for the organization thus migrates from a Capital Expenditure (CapEx) to an Operating Expenditure (OpEx) model allowing for smaller organizations to compete with more established larger organizations.

Despite all these benefits there are some security concerns because of the sensitivity of the company data; however, using the shared responsibility model has businesses choosing this very popular option.

Using Cloud storage still requires the need for the applications and services performed by Servers thus Server Operating Systems and Network Operating Systems continue to grow at a rapid pace with no end in sight.