UNIT – V

The Application Layer

World Wide Web (WWW)

The World Wide Web (WWW), often called the Web, is a system of interconnected webpages and information that you can access using the Internet. It was created to help people share and find information easily, using links that connect different pages together. The Web allows us to browse websites, watch videos, shop online, and connect with others around the world through our computers and phones.

Key Parts of the Web

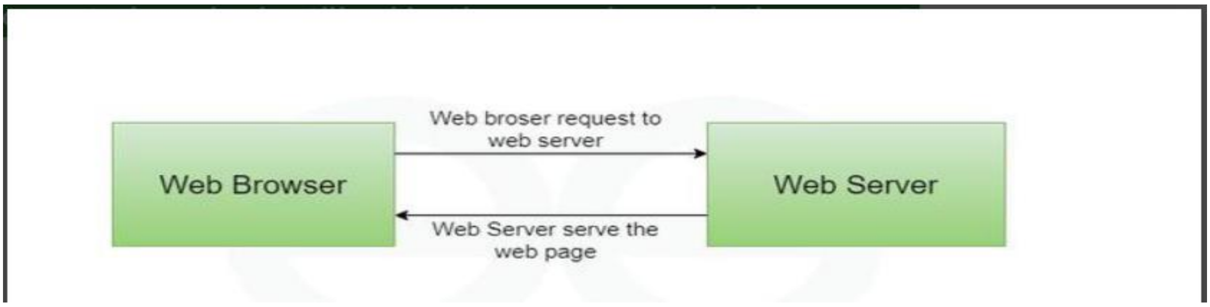
The Web has three main building blocks that make it work:

- **URL (Uniform Resource Locator):** This is the address of a webpage, like https://www.example.com./ It tells your browser exactly where to find the page.
- **HTTP (Hypertext Transfer Protocol):** This is the set of rules that lets your browser and the server talk to each other to send and receive webpages.
- **HTML (Hypertext Markup Language):** This is the code that tells browsers how to display a webpage, including where to put text, pictures, and links.

Working of World Wide Web(WWW)

A Web browser is used to access web pages. Web browsers can be defined as programs which display text, data, pictures, animation and video on the Internet. Hyperlinked resources on the World Wide Web can be accessed using software interfaces provided by Web browsers. Initially, Web browsers were used only for surfing the Web but now they have become more universal.

The below diagram indicates how the Web operates just like client-server architecture of the internet. When users request web pages or other information, then the web browser of your system request to the server for the information and then the web server provide requested services to web browser back and finally the requested service is utilized by the user who made the request.


Web Browser
Web broser request to web server
Web Server serve the web page
Web Server

Challenges of the Web

The Web is amazing, but it has some problems that you should know about:

- **Privacy**: Some websites collect information about you, like what you search for, and might share it without asking.
- **Safety**: Hackers can try to steal your information or send viruses through fake links or ads.
- **False Information**: Not everything on the Web is true, so you need to check if a website is trustworthy.
- **Bullying**: Some people use the Web to be mean or bully others, which can hurt feelings.
- **Too Much Screen Time**: Spending too much time online can make it hard to focus on school or sleep well.
- **Access Issues**: Not everyone has fast Internet, especially in some countries, which makes it harder to use the Web.

History of the WWW

It is a project created, by Tim Berner Lee in 1989, for researchers to work together effectively at CERN. It is an organization, named the World Wide Web Consortium (W3C), which was developed for further development of the web. This organization is directed by Tim Berner's Lee, aka the father of the web. CERN, where Tim Berners worked, is a community of more than 1700 researchers from more than 100 countries. These researchers spend a little time on CERN and the rest of the time they work at their colleges and national research facilities in their home country, so there was a requirement for solid communication so that they can exchange data.

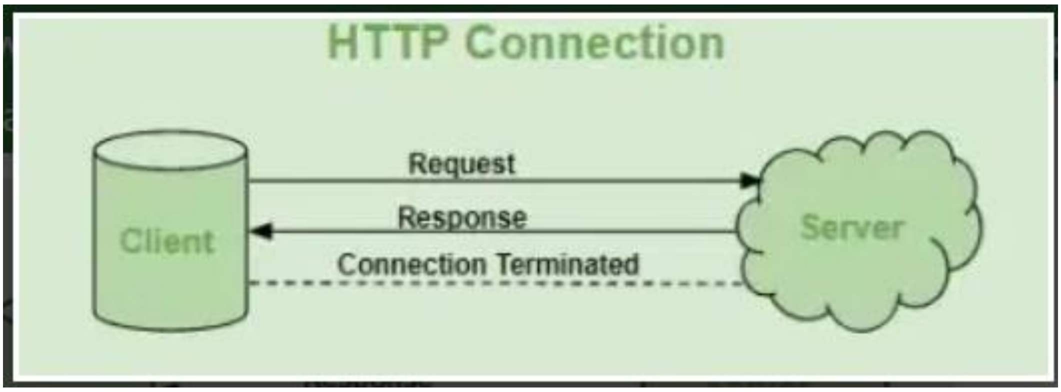
How Web Evolves?

World Wide Web(WWW) Evolved so much from web 1.0 to web 4.0 (Future of WWW) as follows:

- Web 1.0 (1990–2000) Introduced static websites,
- while Web 2.0 (2000–2010) brought interactive and social platforms.
- Web 3.0 (2010–2020) focused on the semantic web, enabling machines to understand data and offer personalized experiences.
- Looking ahead, Web 4.0 (2020–2030) is expected to be a fully intelligent web ecosystem powered by AI and integrated web operating systems.

Hypertext Transfer Protocol - HTTP

HTTP (Hypertext Transfer Protocol) is a fundamental protocol of the Internet, enabling the transfer of data between a client and a server. It is the foundation of data communication for the World Wide Web. HTTP provides a standard between a web browser and a web server to establish communication. It is a set of rules for transferring data from one computer to another.



How HTTP Works: Step-by-Step Process

Here’s how HTTP works when you visit a website:

1. Open Web Browser: First, you open your web browser and type a website URL (e.g., www.example.com).
2. DNS Lookup: Your browser asks a Domain Name System (DNS) server to find out the IP address associated with that URL. Think of this as looking up the phone number of the website.
3. Send HTTP Request: Once the browser has the website’s IP address, it sends an HTTP request to the server. The request asks the server for the resources needed to display the page (like text, images, and videos).
4. Server Response: The server processes your request and sends back an HTTP response. This response contains the requested resources (like HTML, CSS, JavaScript) needed to load the page.
5. Rendering the Web Page: Your browser receives the data from the server and displays the webpage on your screen.

Common HTTP Methods

Method	Description
GET	Retrieve data from the server
POST	Send data to the server (like form submissions)
PUT	Update existing data

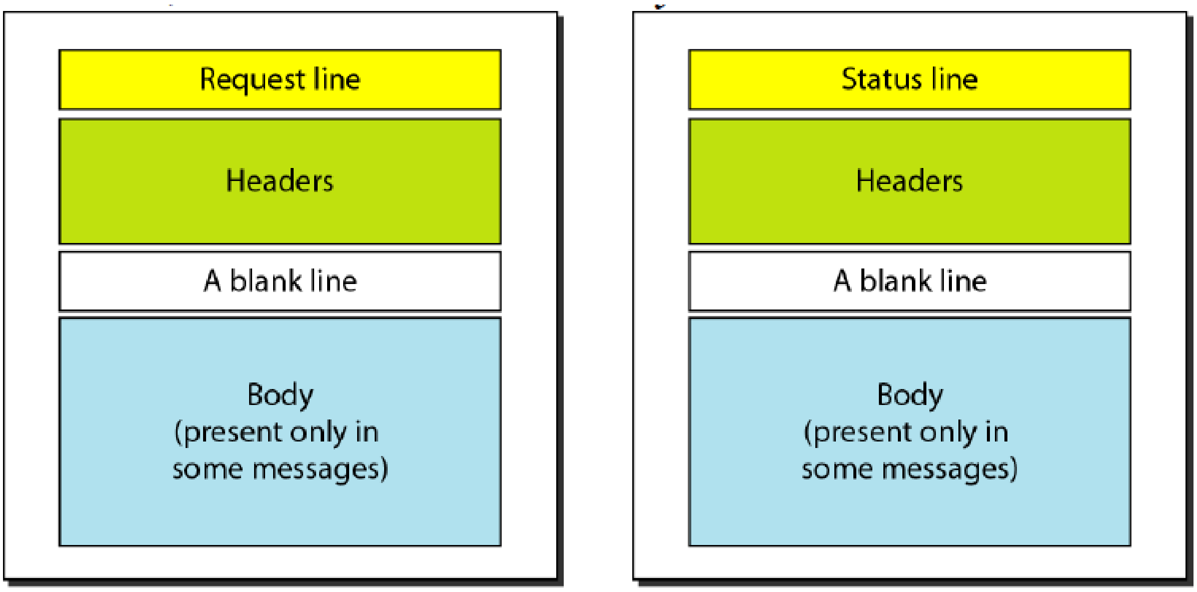
DELETE	Delete data
HEAD	Same as GET but only headers are returned
OPTIONS	Used to retrieve the communication options available for a resource, including supported methods and headers.

Common HTTP Status Codes

Code	Meaning
200 OK	Request successful
301 Moved Permanently	Resource has a new URL
400 Bad Request	Invalid request
401 Unauthorized	Authentication required
403 Forbidden	Access denied
404 Not Found	Page not found
500 Internal Server Error	Server encountered an error

HTTP Messages

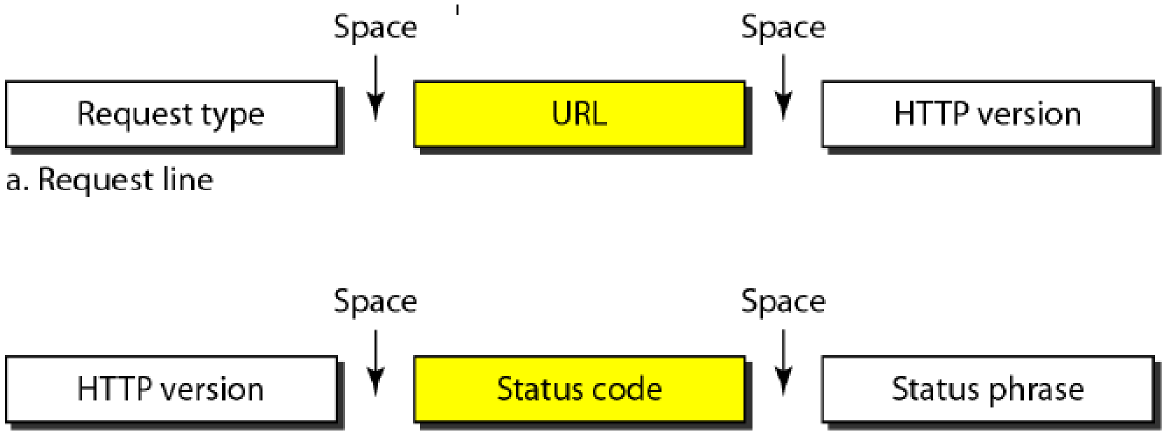
- The formats of the request and response messages are similar; both are shown in below figure. A request message consists of a request line, a header, and sometimes a body. A response message consists of a status line, a header, and sometimes a body.



Request message

Response message

Request and Status Lines- The first line in a request message is called a request line; the first line in the response message is called the status line.



a. Request line

b. Status line

Request type. This field is used in the request message. In version 1.1 of HTTP, several request types are defined. The request type is categorized into methods as defined in below table.

- **URL-**The uniform resource locator (URL) is a standard for specifying any kind of information on the Internet..
- **Version-**The most current version of HTTP is 1.1.
- **Status code-** This field is used in the response message. The status code field is similar to those in the FTP and the SMTP protocols. It consists of three digits. Whereas the codes in the 100 range are only informational, the codes in the 200 range indicate a successful request.
- **Status phrase-** This field is used in the response message. It explains the status code in text form.
- **Header** -The header exchanges additional information between the client and the server. For example, the client can request that the document be sent in a special format, or the server can send extra information about the document. The header can consist of one or more header lines. Each header line has a header name, a colon, a space, and a header value.
- A header line belongs to one of four categories:
    - general header
    - request header
    - response header
    - and entity header.
    - **General header -**The general header gives general information about the message and can be present in both a request and a response. Below table lists some general headers with their descriptions.

- **Request header -** The request header can be present only in a request message. It specifies the client's configuration and the client's preferred document format. See below Table for a list of some request headers and their descriptions.
- **Response header -** The response header can be present only in a response message. It specifies the server's configuration and special information about the request. See below Table for a list of some response headers with their descriptions.
- **Entity header-** The entity header gives information about the body of the document. Although it is mostly present in response messages, some request messages, such as POST or PUT methods, that contain a body also use this type of header. See below Table for a list of some entity headers and their descriptions.

The Domain Name System:

- An application layer protocol defines how the application processes running on different systems, pass the messages to each other.
- DNS stands for Domain Name System.
- DNS is a directory service that provides a mapping between the name of a host on the network and its numerical address.
- DNS is required for the functioning of the internet.
- DNS is a service that translates the domain name into IP addresses. This allows the users of networks to utilize user-friendly names when looking for other hosts instead of remembering the IP addresses.
- To map a name onto an IP address, an application program calls a library procedure called Resolver, passing it the name as a parameter. The resolver sends a UDP packet to a local DNS server, which then looks up the name and returns the IP address to the resolver, which then returns it to the caller.
- Armed with the IP address, the program can then establish a TCP connection with the destination, or send it UDP packets.
- For example, suppose the FTP site at gmail had an IP address of 132.147.165.50, most people would reach this site by specifying

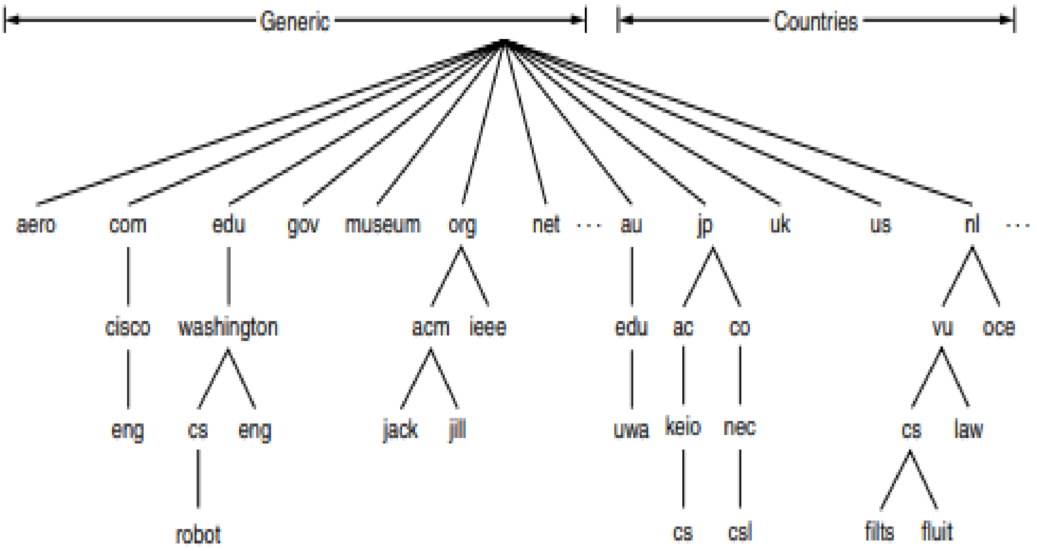
ftp.EduSoft.com. Therefore, the domain name is more reliable than IP address.

THE DNS NAME SPACE:

- The Internet is divided into several hundred top level domains, where each domain covers many hosts.
- Each domain is partitioned into sub domains, and these are further partitioned as so on. All these domains can be represented by a tree, in which the leaves represent domains that have no sub domains.
- A leaf domain may contain a single host, or it may represent a company and contains thousands of hosts. Each domain is named by the path upward from it to the root. The components are separated by periods (pronounced “dot”)

The Top domain comes in 2 flavors:-

- **Generic:**com(commercial), edu(educational instructions), mil(the U.S armed forces, government), int (certain international organizations), net(network providers), org (non profit organizations).
- **Country:** include 1 entry for every country. Domain names can be either absolute (ends with a period e.g. eng.sum.com) or relative (doesn’t end with a period). Domain names are case sensitive and the component names can be up to 63 characters long and full path names must not exceed 255 characters



RESOURCE RECORDS:

- Every domain can have a set of resource records associated with it. For a single host, the most common resource record is just its IP address.
- When a resolver gives a domain name to DNS, it gets both the resource records associated with that name i.e., the real function of DNS is to map domain names into resource records.
- A resource record is a 5-tuple and its format is as follows:

  1.Domain Name

  2.Time to live

  3.Type

  4.Class

  5. Value

  Domain _name : Tells the domain to which this record applies.

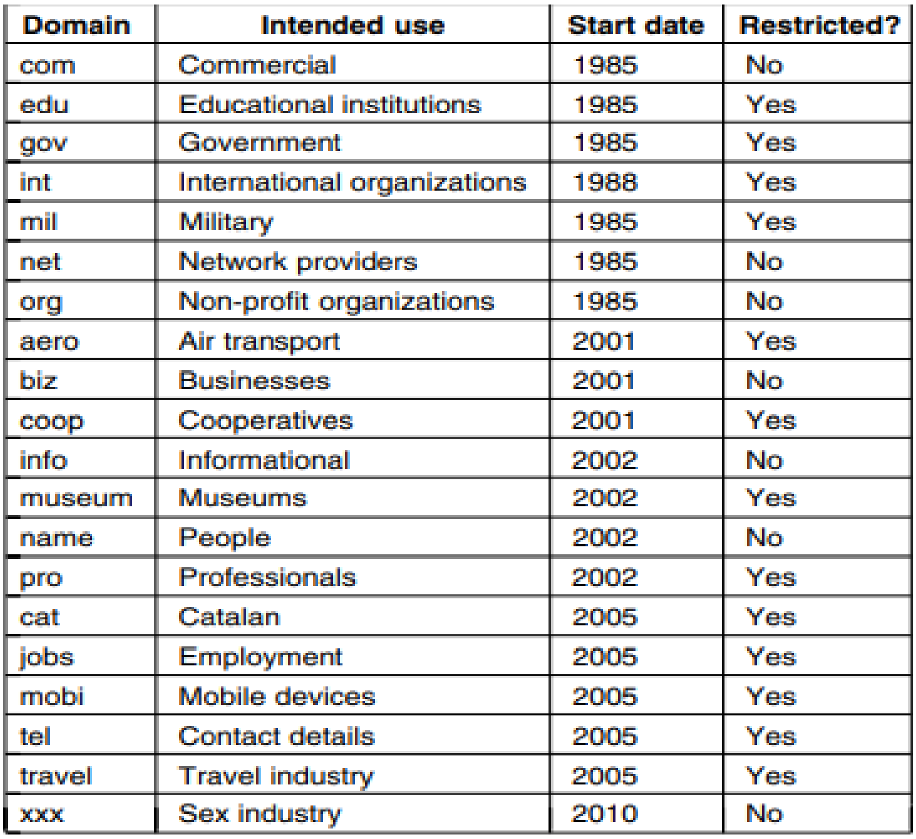
  Time- to- live : Gives an identification of how stable the record is (High Stable = 86400 i.e. no. of seconds /day) (High Volatile = 1 min)

  Type: Tells what kind of record this is.

  Class: It is IN for the internet information and codes for non internetinformation ,other code can be used.

  Value: This field can be a number a domain name or an ASCII string

  Generic top-level domains



Domain	Intended use	Start date	Restricted?
com	Commercial	1985	No
edu	Educational institutions	1985	Yes
gov	Government	1985	Yes
int	International organizations	1988	Yes
mil	Military	1985	Yes
net	Network providers	1985	No
org	Non-profit organizations	1985	No
aero	Air transport	2001	Yes
biz	Businesses	2001	No
coop	Cooperatives	2001	Yes
info	Informational	2002	No
museum	Museums	2002	Yes
name	People	2002	No
pro	Professionals	2002	Yes
cat	Catalan	2005	Yes
jobs	Employment	2005	Yes
mobi	Mobile devices	2005	Yes
tel	Contact details	2005	Yes
travel	Travel industry	2005	Yes
xxx	Sex industry	2010	No

Email (Electronic Mail)

- **Electronic Mail** (e-mail) is one of most widely used services of Internet. This service allows an Internet user to send a **message in formatted manner (mail)** to the other Internet user in any part of world.
- Message in mail not only contain text, but it also contains images, audio and videos data. The person who is sending mail is called **sender** and person who receives mail is called **recipient**.

Format of E-mail :

An e-mail consists of three parts that are as follows :

1. Envelope

2. Header

3. Body

1. Envelope :

The envelope part encapsulates the message. It contains all information that is required for sending any e-mail such as destination address, priority and security level. The envelope is used by MTAs for routing message.

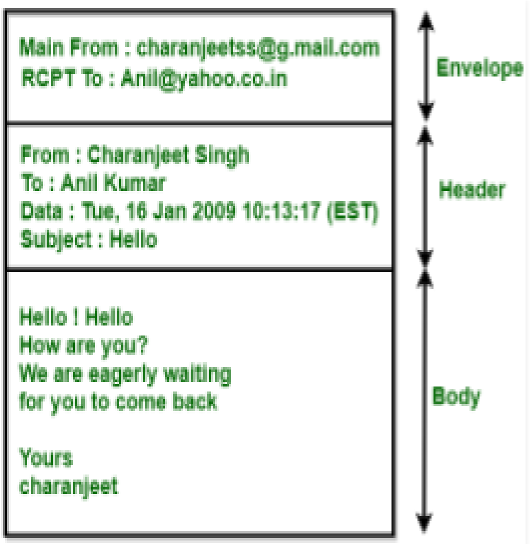
2. Header :
The header consists of a series of lines. Each header field consists of a single line of ASCII text specifying field name, colon and value. The main header fields related to message transport are :

- **To:** It specifies the DNS address of the primary recipient(s).
- **Cc :** It refers to carbon copy. It specifies address of secondary recipient(s).
- **BCC:** It refers to blind carbon copy. It is very similar to Cc. The only difference between Cc and Bcc is that it allow user to send copy to the third party without primary and secondary recipient knowing about this.
- **From :** It specifies name of person who wrote message.
- **Sender :** It specifies e-mail address of person who has sent message.
- **Received :** It refers to identity of sender's, data and also time message was received. It also contains the information which is used to find bugs in routing system.
- **Return-Path:** It is added by the message transfer agent. This part is used to specify how to get back to the sender.

3. Body:-

The body of a message contains text that is the actual content/message that needs to be sent, such as “Employees who are eligible for the new health care program should contact their supervisors by next Friday if they want to switch.” The message body also may include signatures or automatically generated text that is inserted by the sender’s email system.

- The above-discussed field is represented in tabular form as follows :



Simple Mail Transfer Protocol(SMTP)

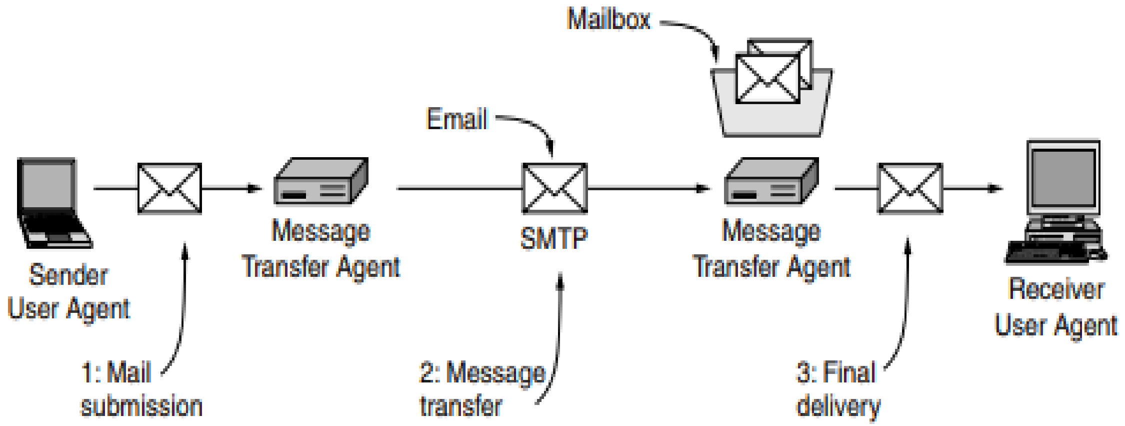
SMTP mainly stands for Simple Mail Transfer Protocol. Basically, the actual transfer of mail is done through the message transfer agents(MTA). Thus in order to send the mail, the system must have the **client MTA** and in order to receive the mail, the system must have a server MTA.

- In order to define the MTA client and server on the Internet, there is a formal way and it is known as Simple Mail Transfer Protocol(SMTP).
- SMTP also makes the use of TCP/IP for sending and receiving e-mail.
- SMTP is based on the client/server model.
- The original standard port for SMTP is Port 25.
- Using this protocol, the client who wants to send the e-mail first opens a TCP connection to the SMTP server and then sends the e-mail across the TCP connection. It is important to note that the SMTP server is always in listening mode. As soon as it listens for the TCP connection from any client then the connection is Initiated on port 25 and after the successful connection, the client sends the e-mail/message immediately.

Architecture and Services of SMTP

It will provide an overview of how email systems are organized and what they can do. The architecture of the email system is consists of two kinds of subsystems:

1. The user agents- A user agent is normally a program (sometimes called a mail reader) that accepts a variety of commands for composing, receiving, and replying to messages, as well as for manipulating mailboxes
2. Message transfer agents- Which move the messages from the source to the destination. We will also refer to message transfer agents informally as mail servers.



Architecture of the email system.

3. Sending E-mail

- To send an e-mail message, a user must provide the message, the destination address, and possibly some other parameters.
- The message can be produced with a free-standing text editor, a word processing program, or possibly with a specialized text editor built into the user agent.
- The destination address must be in a format that the user agent can deal with. Many user agents expect addresses of the form user@dns-address.

4. Reading E-mail

- When a user agent is started up, it looks at the user's mailbox for incoming e-mail before displaying anything on the screen.
- Then it may announce the number of messages in the mailbox or display a one-line summary of each one and wait for a command

5. Message Transfer

- The message transfer system is concerned with relaying messages from the originator to the recipient. The simplest way to do this is to establish a transport connection from the source machine to the destination machine and then just transfer the message.

6. Message Formats
   RFC 822 Messages consist of a primitive envelope (described in RFC 821), some number of header fields, a blank line, and then the message body.

   Each header field (logically) consists of a single line of ASCII text containing the field name, a colon, and, for most fields, a value.

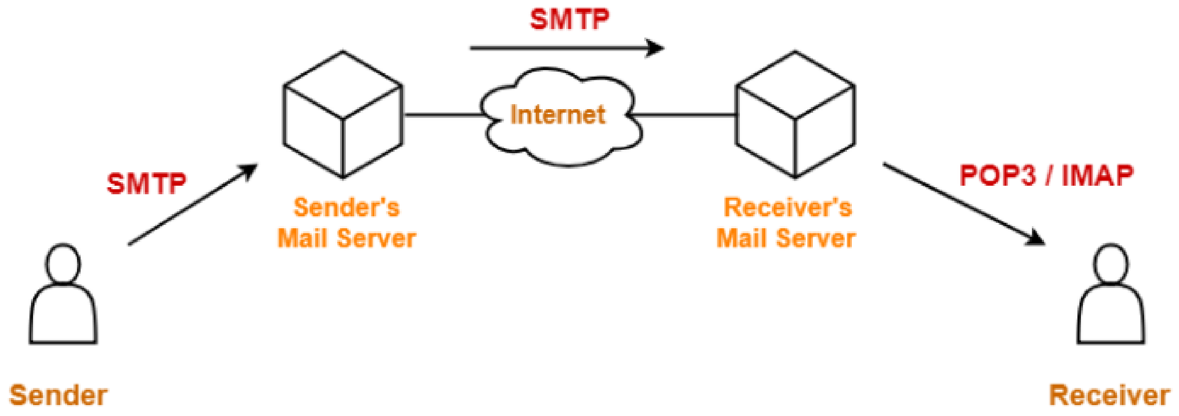
Post Office Protocol (POP)

- POP3, which is an abbreviation for Post Office Protocol 3, is again a protocol used for receiving email. Much like the physical version of a post office clerk, POP3 receives and holds email for an individual until they pick it up.
- POP3 is being designed for receiving the mails.
- POP3 is being implemented on port number 110.
- POP3 is a MAA (Message Access Agent) for accessing the messages from mail box.
- POP3 allows retrieving and organizing mails from mailbox on receiver mail server to receiver's computer.

Working of POP

All the incoming messages are stored on the POP server until the user login by using an email client and downloads the message to their computer. After the message is downloaded by the user it gets deleted from the server.

As we know that the SMTP is used to transfer the email message from the server to the server, basically POP is used to collect the email with an email client from the server and it does not include means to send messages.



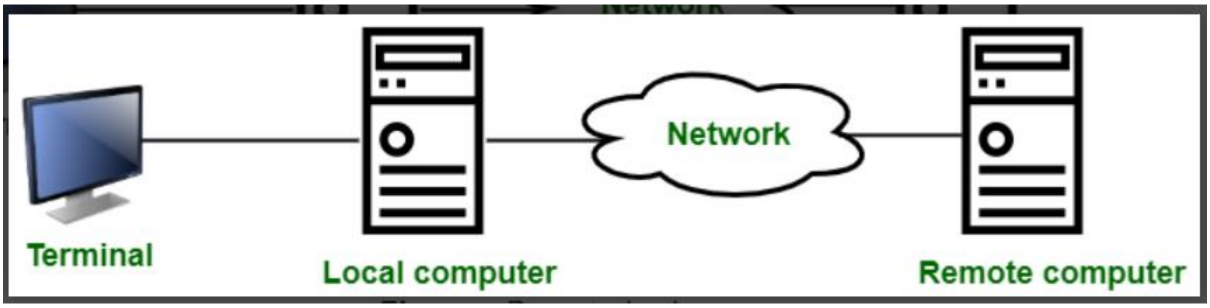
If any user tries to check all the recent emails then they will establish a connection with the **POP3** at the server-side. The user sends the username and password to the server machine for getting the proper authentication. After getting the connection, users can receive all text-based emails and store them on their local terminal (machine), then finally discard all server copies and then breaks the connection from the server machine.

In order to retrieve a message from the server following steps are taken;

- Firstly a TCP connection is established by the client using port 110.
- The client identifies itself to the server.
- After that client issues a series of POP3 commands.

Remote Loging

Remote login is the process of connecting to a computer or network from a different physical location to access its resources, files, and applications as if you were sitting in front of it. This is achieved through network connections and specific protocols, with common tools like SSH for secure connections and the older, less secure Telnet. It allows users to work, manage files, or provide technical support from anywhere, provided there is an internet or network connection.



How it works

- A user initiates a connection from a local computer to a remote computer using a remote login tool.
- The user authenticates themselves, typically with a username and password, to verify their identity.
- Commands typed on the local machine are sent to the remote machine, where they are processed.
- The results are sent back to the local machine to be displayed, making the remote computer's resources available on the local one.

TELNET

Telnet is a client-server application protocol in computer networks that enables a user to establish a remote, text-based session on another machine. An admin can use Telnet to log in and issue commands to a remote server or network device as if they were physically sitting in front of it.

The protocol is one of the oldest in use today, dating back to 1969. Due to its significant security flaws, it has largely been replaced by more modern and secure alternatives like SSH (Secure Shell).

How Telnet works

The Telnet protocol functions over a single TCP/IP connection, typically using port 23.

1. **Connection:** A user starts a Telnet client on their local machine to connect to a Telnet server running on a remote host.
2. **Network Virtual Terminal (NVT):** Since different systems use different command-line standards, Telnet uses a Network Virtual Terminal (NVT) to act as a universal, intermediary format. The client translates keystrokes and commands from the local system into NVT format before sending them across the network.
3. **Client-server communication:** When the Telnet server on the remote machine receives the NVT-formatted characters, it translates them into a format that

the remote operating system can understand. The remote OS processes the commands and sends the output back to the server.

4. **Display:** The Telnet server returns the output to the client in NVT format, which is then translated back and displayed on the local user's screen.

Advantages

- Simplicity and Ease of Use: Telnet is a straightforward, text-based protocol that is simple to learn and requires minimal configuration.
- Lightweight and Fast: The absence of encryption or a graphical interface makes Telnet a fast and low-overhead tool for quick network checks.
- Universal: The protocol is platform-independent and can be used on almost any operating system.
- Network Service Testing: Telnet is still useful for testing network services and diagnosing connectivity issues by checking if a port on a remote server is open and responding

Disadvantages and security risks

- No encryption: Telnet does not encrypt the data transmitted over a network. All information, including usernames and passwords, is sent in plain text, making it highly vulnerable to interception by attackers.
- Vulnerability to attacks: The lack of encryption and authentication makes Telnet susceptible to eavesdropping, man-in-the-middle attacks, and session hijacking.
- Ineffective authentication: The standard Telnet authentication process is weak, which has allowed improperly configured devices to be easily exploited by malware.
- No support for modern features: Telnet is limited to text-only communication and does not support graphical interfaces or advanced features like file transfer.

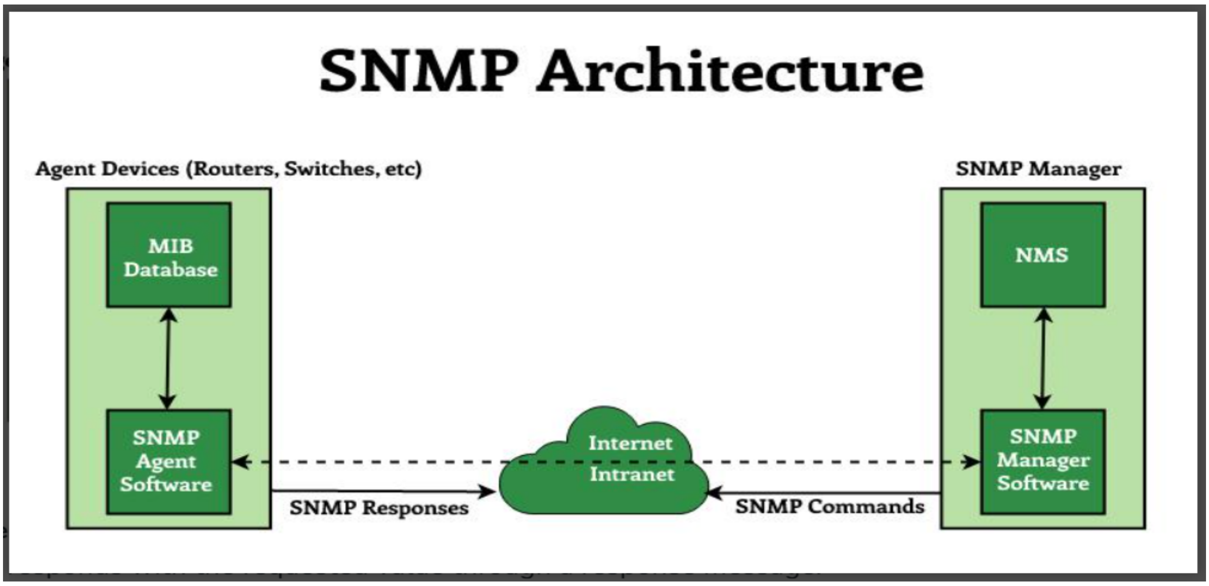
Simple Network Management Protocol (SNMP)

Simple Network Management Protocol (SNMP) is a widely used protocol for network management that provides a standardized framework for monitoring and managing network devices such as routers, switches, servers, printers, firewalls, and load balancer. It operates within the application layer of the Internet protocol suite and allows network administrators to manage network performance, find and solve network problems, and plan for network growth.

Architecture of SNMP

There are mainly three main components in SNMP architecture:

- **SNMP Manager:** It is a centralized system used to monitor the network. It is also known as a Network Management Station (NMS). A router that runs the SNMP server program is called an agent, while a host that runs the SNMP client program is called a manager.
- **SNMP agent:** It is a software management software module installed on a managed device. The manager accesses the values stored in the database, whereas the agent maintains the information in the database. To ascertain if the router is congested or not, for instance, a manager can examine the relevant variables that a router stores, such as the quantity of packets received and transmitted.
- **Management Information Base:**MIB consists of information on resources that are to be managed. This information is organized hierarchically. It consists of objects instances which are essentially variables. A MIB, or collection of all the objects under management by the manager, is unique to each agent. System, interface, address translation, IP, UDP , and EGP , ICMP , TCP are the eight categories that make up MIB. The MIB object is home to these groups.



SNMP Messages

- **GetRequest** : It is simply used to retrieve data from SNMP agents. In response to this, the SNMP agent responds with the requested value through a response message.
- **GetNextRequest :** To get the value of a variable, the manager sends the agent the GetNextRequest message. The values of the entries in a table are retrieved using this kind of communication. The manager won't be able to access the values if it doesn't know the entries' indices. The GetNextRequest message is used to define an object in certain circumstances.

- **SetRequest :** It is used by the SNMP manager to set the value of an object instance on the SNMP agent.
- **Response :** When sent in response to the Set message, it will contain the newly set value as confirmation that the value has been set.
- **Trap :** These are the message sent by the agent without being requested by the manager. It is sent when a fault has occurred.
- **InformRequest :** It was added to SNMPv2c and is used to determine if the manager has received the trap message or not. It is the same as a trap but adds an acknowledgement that the trap doesn't provide.

Characteristics of SNMP

- SNMP is used to monitor network.
- It detects any network faults.
- It can also be used to configure remote devices.
- It allows a standardized way of collecting information about all kinds of devices from various manufacturers among the networking industry.

Advantages of SNMP

- It is easy to implement.
- Agents are widely implemented.
- Agent level overhead is minimal.
- It is robust and extensible.
- Polling approach is good forLAN based managed object.
- It offers the best direct manager agent interface.

Limitation of SNMP

- It does not scale well.
- There is no object orietned data view.
- It has no standard control definition.
- It has many implementation specific (private MIB) extensions.
- It has high communication overhead due to polling

Internet Message Access Protocol (IMAP)

- **Internet Message Access Protocol (IMAP)** is an application-layer protocol that operates as a contract for receiving emails from the mail server. It was designed by Mark Crispin in 1986 as a remote access mailbox protocol; the current version of IMAP is IMAP4.
- It is used as the most commonly used protocol for retrieving emails. This term is also known as Internet mail access protocol, Interactive mail access protocol, and Interim mail access protocol.
- It retrieves messages from your email provider to your email client (such as on your Mac or PC). Importantly, it does not delete messages from the email service after you download them until the user explicitly deletes them.
- It allows you to view and manage your emails across multiple devices because it synchronizes the emails between your email client and the server. This

synchronization ensures that changes such as read status, deletions, and folder organization are reflected consistently across all devices connected to the same email account.

Working of IMAP

IMAP follows Client-server Architecture and is the most commonly used email protocol. It is a combination of client and server process running on other computers that are connected through a network.

This protocol resides over the TCP/IP protocol for communication. Once the communication is set up the server listens on port 143 by default which is non-encrypted. For the secure encrypted communication port, 993 is used.

The following steps are taken for the working of the IMAP :

- Email client Gmail establishes a connection with Gmail's SMTP server.
- By approving the sender's and recipient's email addresses, the SMTP server verifies (authenticates) that the email can be sent.
- The email is sent to the Outlook SMTP server by Gmail's SMTP server.
- The recipient's email address is authenticated by the Outlook SMTP server.
- IMAP or POP3 is used by the Outlook SMTP server to deliver the email to the Outlook email client.

Architecture of IMAP

The Internet Message Access Protocol (IMAP) protocol is a client-server model that allows users to access and view email messages stored on remote servers Here is a summary of the events:

- **IMAP clients:** An IMAP client is an email application or software that users use to communicate with their email accounts. Examples include Microsoft Outlook, Mozilla Thunderbird, Apple Mail, and mobile email applications. The client communicates with the server to receive, manage, and send email messages.
- **IMAP Server:** The IMAP server manages email messages and manages user mailboxes. It responds to requests from IMAP clients, and provides access to email folders and messages.

  The server stores emails in a structured format, usually organized in user-defined folders or mailboxes. Common IMAP server software includes Dovecot, Courier IMAP, Cyrus IMAP, and Microsoft Exchange Server.
- **Network Protocols:** It works over TCP/IP (Transmission Control Protocol/Internet Protocol) networks, and allows an client to connect to a server over the Internet or local area networks.