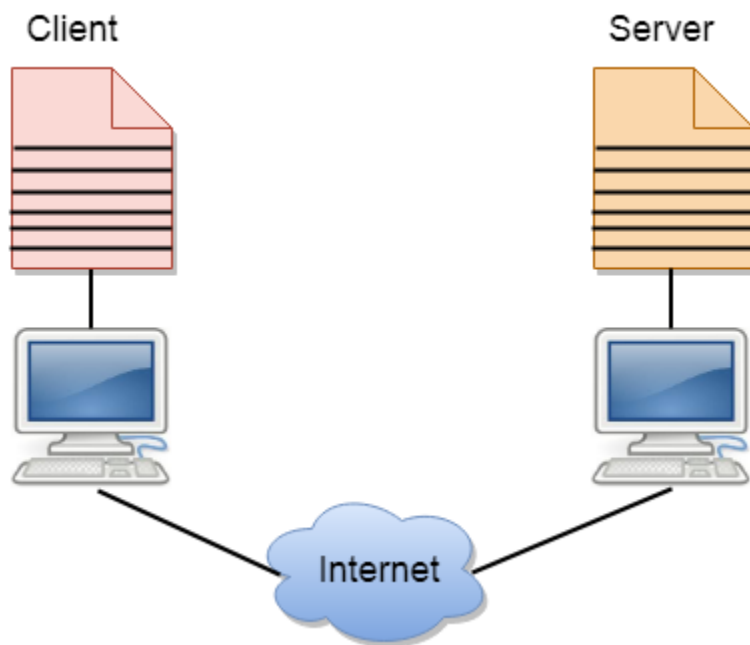


# UNIT 5:

# Application Layer Protocols

## Client and Server model

- A client and server networking model is a model in which computers such as servers provide the network services to the other computers such as clients to perform a user based tasks. This model is known as client-server networking model.
- The application programs using the client-server model should follow the given below strategies:



- An application program is known as a client program, running on the local machine that requests for a service from an application program known as a server program, running on the remote machine.
- A client program runs only when it requests for a service from the server while the server program runs all time as it does not know when its service is required.
- A server provides a service for many clients not just for a single client. Therefore, we can say that client-server follows the many-to-one relationship. Many clients can use the service of one server.
- Services are required frequently, and many users have a specific client-server application program. For example, the client-server application program allows the user to access the files, send e-mail, and so on. If the services are more customized, then we should have one generic application program that allows the user to access the services available on the remote computer.

## Client

A client is a program that runs on the local machine requesting service from the server. A client program is a finite program means that the service started by the user and terminates when the service is completed.

## Server

A server is a program that runs on the remote machine providing services to the clients. When the client requests for a service, then the server opens the door for the incoming requests, but it never initiates the service.

A server program is an infinite program means that when it starts, it runs infinitely unless the problem arises. The server waits for the incoming requests from the clients. When the request arrives at the server, then it responds to the request.

## Advantages of Client-server networks:

- **Centralized:** Centralized back-up is possible in client-server networks, i.e., all the data is stored in a server.
- **Security:** These networks are more secure as all the shared resources are centrally administered.

- **Performance:** The use of the dedicated server increases the speed of sharing resources. This increases the performance of the overall system.
- **Scalability:** We can increase the number of clients and servers separately, i.e., the new element can be added, or we can add a new node in a network at any time.

## Disadvantages of Client-Server network:

- **Traffic Congestion** is a big problem in Client/Server networks. When a large number of clients send requests to the same server may cause the problem of Traffic congestion.
- It does not have a robustness of a network, i.e., when the server is down, then the client requests cannot be met.
- A client/server network is very decisive. Sometimes, regular computer hardware does not serve a certain number of clients. In such situations, specific hardware is required at the server side to complete the work.
- Sometimes the resources exist in the server but may not exist in the client. For example, If the application is web, then we cannot take the print out directly on printers without taking out the print view window on the web.

## NFS

**NFS** is an abbreviation of the **Network File System**. It is a protocol of a distributed file system. This protocol was developed by the [Sun Microsystems](#) in the year of 1984.

It is an architecture of the client/server, which contains a client program, server program, and a protocol that helps for communication between the client and server.

It is that protocol which allows the users to access the data and files remotely over the network. Any user can easily implement the NFS protocol because it is an open standard. Any user can manipulate files as same as if they were on like other protocols. This protocol is also built on the ONC RPC system.

This protocol is mainly implemented on those computing environments where the centralized management of resources and data is critical. It uses the [Transmission Control Protocol \(TCP\)](#) and [User Datagram Protocol \(UDP\)](#) for accessing and delivering the data and files.

Network File System is a protocol that works on all the networks of IP-based. It is implemented in that client/server application in which the server of NFS manages the

authorization, authentication, and clients. This protocol is used with Apple Mac OS, Unix, and Unix-like operating systems such as Solaris, [Linux](#), FreeBSD, AIX.

## Difference Between NFS and CIFS

| NFS                                                                                 | CIFS                                                                         |
|-------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| 1. NFS is an abbreviation of the Network File System.                               | 1. CIFS is an abbreviation of the Common Internet File system.               |
| 2. This protocol is used for sharing the files by Unix and Linux Operating systems. | 2. This protocol is used for sharing the files by Windows Operating systems. |
| 3. It is highly scalable.                                                           | 3. It is low scalable.                                                       |
| 4. The speed of communication is fast.                                              | 4. The speed of communication is medium.                                     |
| 5. The network File system is not a secure protocol.                                | 5. Common Internet File System is more secure than the Network File System.  |
| 6. NFS is not a reliable protocol.                                                  | 6. CIFS is a reliable protocol.                                              |
| 7. This protocol does not provide the session.                                      | 7. This protocol provides the sessions.                                      |
| 8. This protocol is easy to implement and set up.                                   | 8. Its implementation is complex.                                            |
| 9. This protocol uses 111 port for both TCP and UDP.                                | 9. This protocol uses 139 and 445 TCP ports and 137 and 138 UDP ports.       |

# Introduction to Remote Login

**Remote Login** is a process in which user can login into remote site i.e. computer and use services that are available on the remote computer. With the help of remote login a user is able to understand result of transferring and result of processing from the remote computer to the local computer.

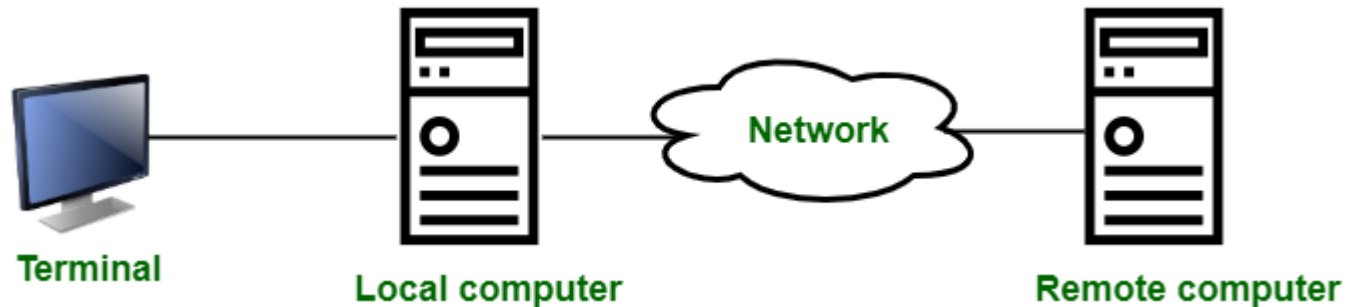
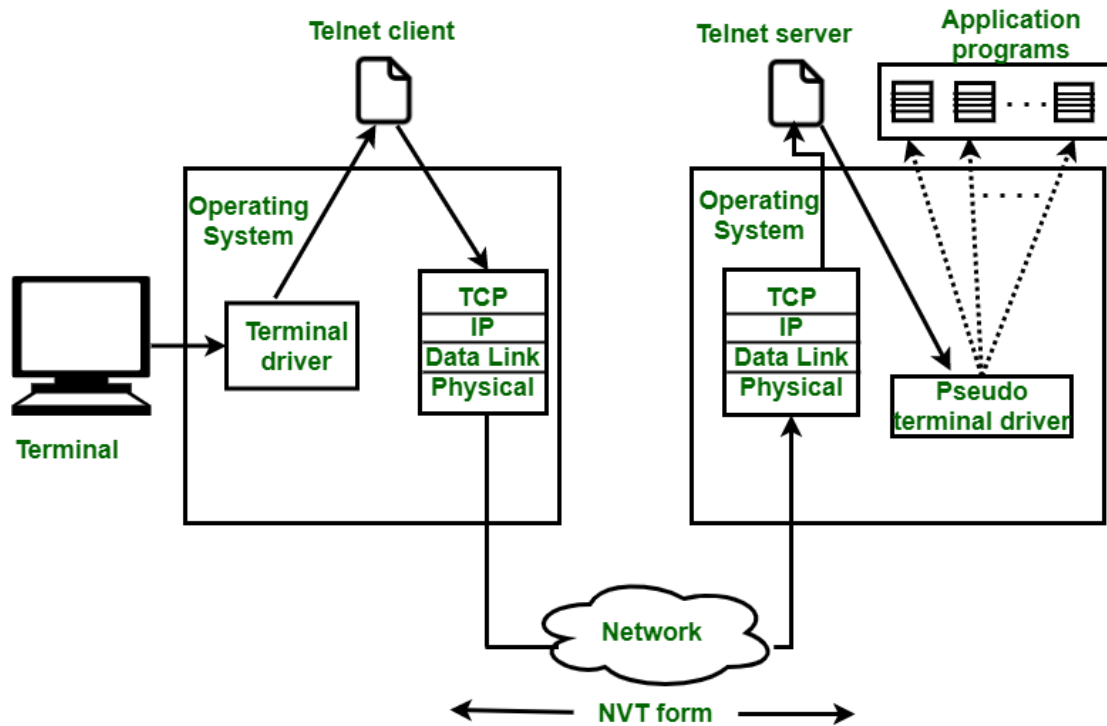


Figure – Remote login

It is implemented using [Telnet](#).

## Procedure of Remote Login :

1. When the user types something on local computer, then local operating system accepts character.
2. Local computer does not interpret the characters, it will send them to TELNET client.
3. TELNET client transforms these characters to a universal character set called **Network Virtual Terminal (NVT) characters** and it will pass them to the local TCP/IP protocol Stack.
4. Commands or text which is in the form of NVT, travel through Internet and it will arrive at the TCP/IP stack at remote computer.
5. Characters are then delivered to operating system and which later on passed to TELNET server.
6. Then TELNET server changes that characters to characters which can be understandable by remote computer.
7. Remote operating system receives character from a **pseudo-terminal driver**, which is a piece of software that pretends that characters are coming from a terminal.
8. Operating system then passes character to the appropriate application program.



**Figure – Remote login procedure**

### NVT Character Set :

- With NVT Character set, TELNET client translates characters into NVT form and deliver to network.
- TELNET server translates data and commands from NVT form to the other form that will be understandable by remote computer.
- NVT uses 2 sets of characters, one for data and other for control. Size of both characters is 8-bit bytes.
- For data, NVT is an 8-bit character set in which 7 lowest bits are same as ASCII and highest order bit is 0.
- For control characters, NVT uses an 8-bit character set in which the highest bit is set to 1.



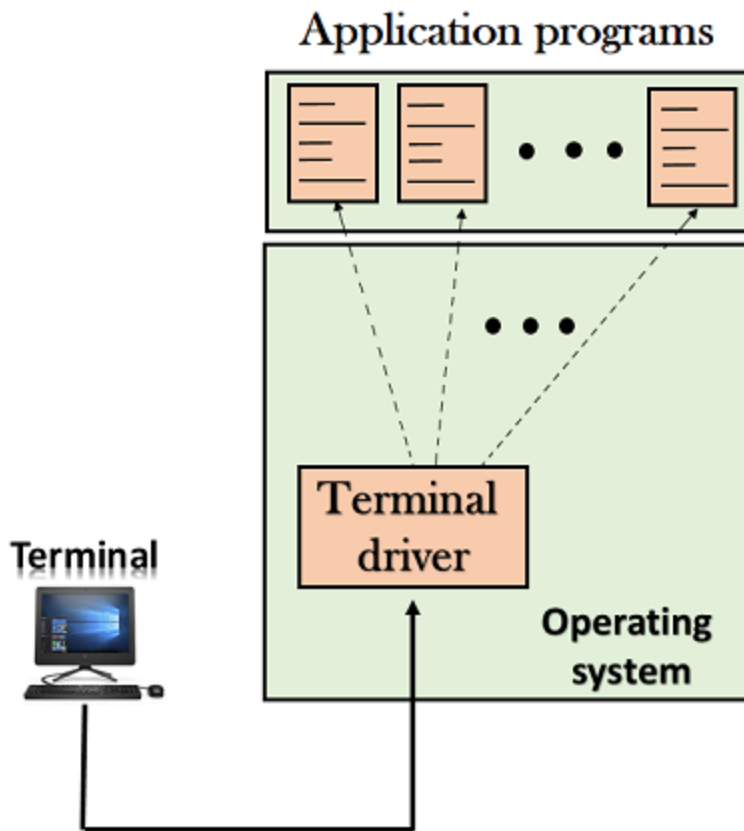
**Figure – NVT Character Set**

# Telnet

- The main task of the internet is to provide services to users. For example, users want to run different application programs at the remote site and transfers a result to the local site. This requires a client-server program such as FTP, SMTP. But this would not allow us to create a specific program for each demand.
- The better solution is to provide a general client-server program that lets the user access any application program on a remote computer. Therefore, a program that allows a user to log on to a remote computer. A popular client-server program Telnet is used to meet such demands. Telnet is an abbreviation for **Terminal Network**.
- Telnet provides a connection to the remote computer in such a way that a local terminal appears to be at the remote side.

## There are two types of login:

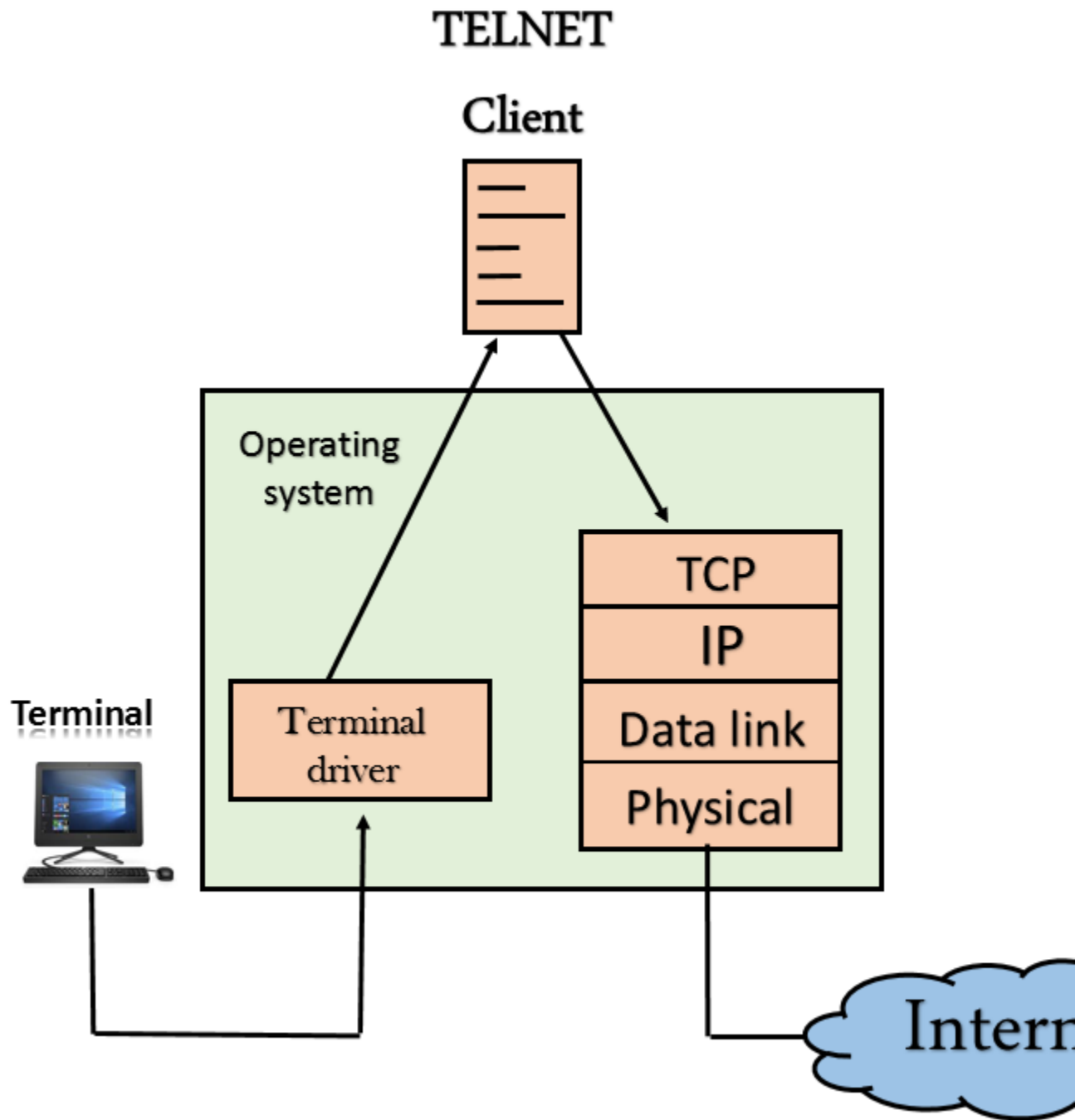
### Local Login



- When a user logs into a local computer, then it is known as local login.
- When the workstation running terminal emulator, the keystrokes entered by the user are accepted by the terminal driver. The terminal driver then passes these characters to the operating system which in turn, invokes the desired application program.
- However, the operating system has special meaning to special characters. For example, in UNIX some combination of characters have special meanings such as control character with "z" means suspend. Such situations do not create any problem as the terminal driver knows the meaning of such characters. But, it can cause the problems in remote login.



## Remote login



- When the user wants to access an application program on a remote computer, then the user must perform remote login.

## **How remote login occurs**

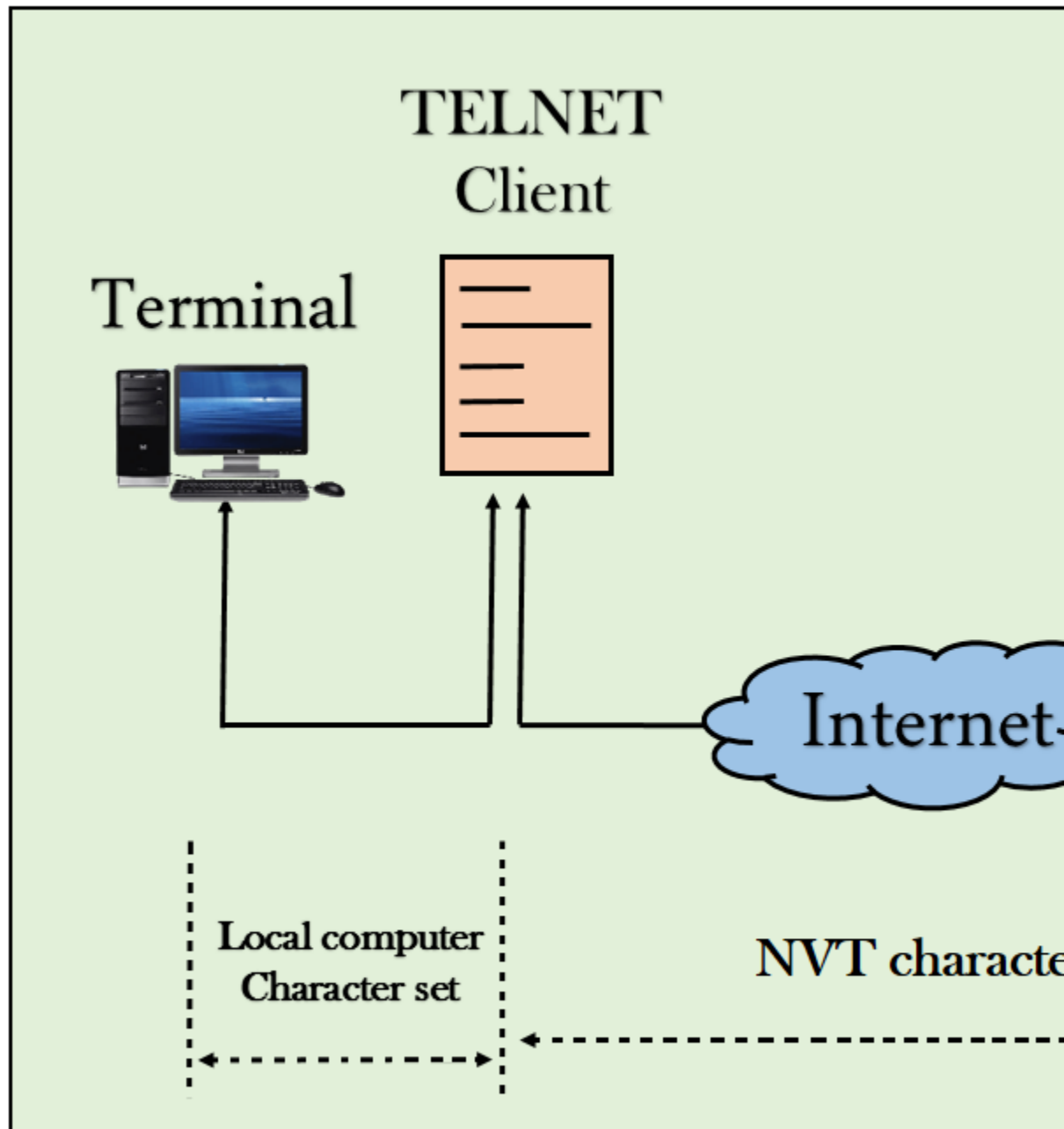
### **At the local site**

The user sends the keystrokes to the terminal driver, the characters are then sent to the TELNET client. The TELNET client which in turn, transforms the characters to a universal character set known as network virtual terminal characters and delivers them to the local TCP/IP stack

### **At the remote site**

The commands in NVT forms are transmitted to the TCP/IP at the remote machine. Here, the characters are delivered to the operating system and then pass to the TELNET server. The TELNET server transforms the characters which can be understandable by a remote computer. However, the characters cannot be directly passed to the operating system as a remote operating system does not receive the characters from the TELNET server. Therefore it requires some piece of software that can accept the characters from the TELNET server. The operating system then passes these characters to the appropriate application program.

## Network Virtual Terminal (NVT)



- The network virtual terminal is an interface that defines how data and commands are sent across the network.
- In today's world, systems are heterogeneous. For example, the operating system accepts a special combination of characters such as end-of-file token running a

DOS operating system *ctrl+z* while the token running a UNIX operating system is *ctrl+d*.

- TELNET solves this issue by defining a universal interface known as network virtual interface.
- The TELNET client translates the characters that come from the local terminal into NVT form and then delivers them to the network. The Telnet server then translates the data from NVT form into a form which can be understandable by a remote computer.

## FTP

- FTP stands for File transfer protocol.
- FTP is a standard internet protocol provided by TCP/IP used for transmitting the files from one host to another.
- It is mainly used for transferring the web page files from their creator to the computer that acts as a server for other computers on the internet.
- It is also used for downloading the files to computer from other servers.

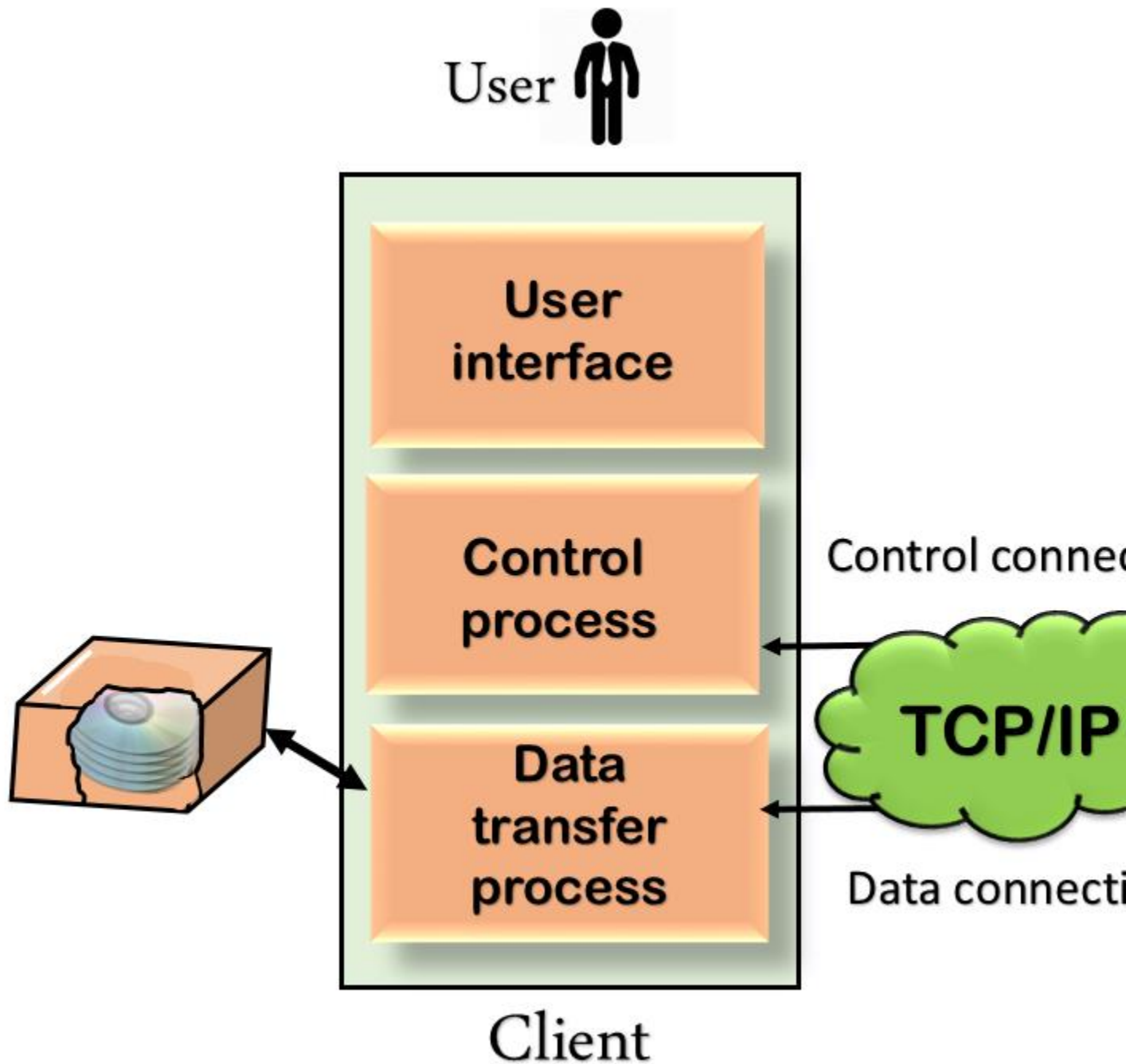
## Objectives of FTP

- It provides the sharing of files.
- It is used to encourage the use of remote computers.
- It transfers the data more reliably and efficiently.

## Why FTP?

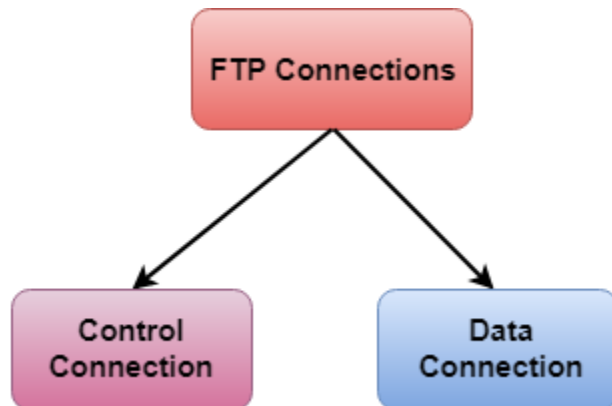
Although transferring files from one system to another is very simple and straightforward, but sometimes it can cause problems. For example, two systems may have different file conventions. Two systems may have different ways to represent text and data. Two systems may have different directory structures. FTP protocol overcomes these problems by establishing two connections between hosts. One connection is used for data transfer, and another connection is used for the control connection.

## Mechanism of FTP



The above figure shows the basic model of the FTP. The FTP client has three components: the user interface, control process, and data transfer process. The server has two components: the server control process and the server data transfer process.

**There are two types of connections in FTP:**



- **Control Connection:** The control connection uses very simple rules for communication. Through control connection, we can transfer a line of command or line of response at a time. The control connection is made between the control processes. The control connection remains connected during the entire interactive FTP session.
- **Data Connection:** The Data Connection uses very complex rules as data types may vary. The data connection is made between data transfer processes. The data connection opens when a command comes for transferring the files and closes when the file is transferred.

## FTP Clients

- FTP client is a program that implements a file transfer protocol which allows you to transfer files between two hosts on the internet.
- It allows a user to connect to a remote host and upload or download the files.
- It has a set of commands that we can use to connect to a host, transfer the files between you and your host and close the connection.
- The FTP program is also available as a built-in component in a Web browser. This GUI based FTP client makes the file transfer very easy and also does not require to remember the FTP commands.

## Advantages of FTP:

- **Speed:** One of the biggest advantages of FTP is speed. The FTP is one of the fastest way to transfer the files from one computer to another computer.
- **Efficient:** It is more efficient as we do not need to complete all the operations to get the entire file.

- **Security:** To access the FTP server, we need to login with the username and password. Therefore, we can say that FTP is more secure.
- **Back & forth movement:** FTP allows us to transfer the files back and forth. Suppose you are a manager of the company, you send some information to all the employees, and they all send information back on the same server.

## Disadvantages of FTP:

- The standard requirement of the industry is that all the FTP transmissions should be encrypted. However, not all the FTP providers are equal and not all the providers offer encryption. So, we will have to look out for the FTP providers that provides encryption.
- FTP serves two operations, i.e., to send and receive large files on a network. However, the size limit of the file is 2GB that can be sent. It also doesn't allow you to run simultaneous transfers to multiple receivers.
- Passwords and file contents are sent in clear text that allows unwanted eavesdropping. So, it is quite possible that attackers can carry out the brute force attack by trying to guess the FTP password.
- It is not compatible with every system.

# Introduction to Electronic Mail

Electronic mail, commonly known as email, is a method of exchanging messages over the internet. Here are the basics of email:

1. An email address: This is a unique identifier for each user, typically in the format of name@domain.com.
2. An email client: This is a software program used to send, receive and manage emails, such as Gmail, Outlook, or Apple Mail.
3. An email server: This is a computer system responsible for storing and forwarding emails to their intended recipients.

To send an email:

1. Compose a new message in your email client.
2. Enter the recipient's email address in the "To" field.
3. Add a subject line to summarize the content of the message.
4. Write the body of the message.

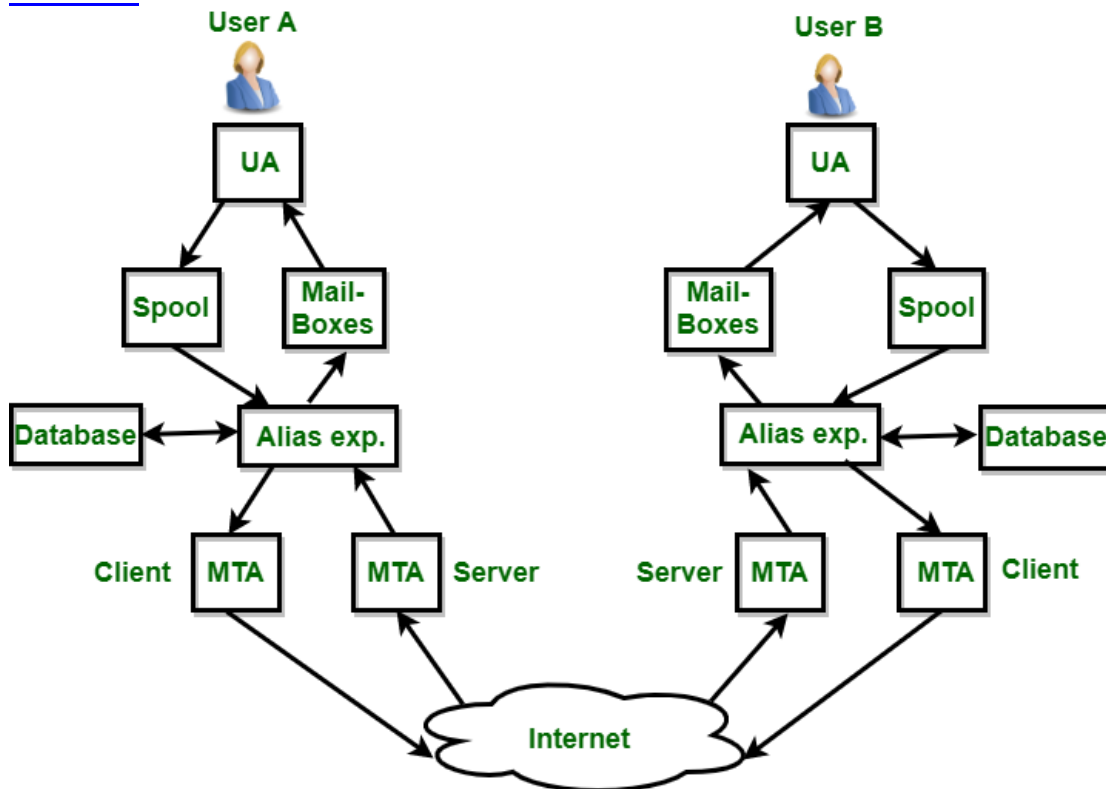
5. Attach any relevant files if needed.
6. Click “Send” to deliver the message to the recipient’s email server.
7. Emails can also include features such as cc (carbon copy) and bcc (blind carbon copy) to send copies of the message to multiple recipients, and reply, reply all, and forward options to manage the conversation.

**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**. It is just like postal mail service. **Components of E-Mail System** : The basic components of an email system are : User Agent (UA), Message Transfer Agent (MTA), Mail Box, and Spool file. These are explained as following below.

1. **User Agent (UA)** : The UA is normally a program which is used to send and receive mail. Sometimes, it is called as mail reader. It accepts variety of commands for composing, receiving and replying to messages as well as for manipulation of the mailboxes.
2. **Message Transfer Agent (MTA)** : MTA is actually responsible for transfer of mail from one system to another. To send a mail, a system must have client MTA and system MTA. It transfer mail to mailboxes of recipients if they are connected in the same machine. It delivers mail to peer MTA if destination mailbox is in another machine. The delivery from one MTA to another MTA is done by [Simple Mail Transfer](#)



## Protocol.



3. **Mailbox** : It is a file on local hard drive to collect mails. Delivered mails are present in this file. The user can read it delete it according to his/her requirement. To use e-mail system each user must have a mailbox . Access to mailbox is only to owner of mailbox.
4. **Spool file** : This file contains mails that are to be sent. User agent appends outgoing mails in this file using SMTP. MTA extracts pending mail from spool file for their delivery. E-mail allows one name, an **alias**, to represent several different e-mail addresses. It is known as **mailing list**, Whenever user have to sent a message, system checks recipient's name against alias database. If mailing list is present for defined alias, separate messages, one for each entry in the list, must be prepared and handed to MTA. If for defined alias, there is no such mailing list is present, name itself becomes naming address and a single message is delivered to mail transfer entity.

### **Services provided by E-mail system :**

- **Composition** – The composition refer to process that creates messages and answers. For composition any kind of text editor can be used.
- **Transfer** – Transfer means sending procedure of mail i.e. from the sender to recipient.
- **Reporting** – Reporting refers to confirmation for delivery of mail. It help user to check whether their mail is delivered, lost or rejected.
- **Displaying** – It refers to present mail in form that is understand by the user.

- **Disposition** – This step concern with recipient that what will recipient do after receiving mail i.e save mail, delete before reading or delete after reading.

Advantages Or Disadvantages:

Advantages of email:

1. Convenient and fast communication with individuals or groups globally.
2. Easy to store and search for past messages.
3. Ability to send and receive attachments such as documents, images, and videos.
4. Cost-effective compared to traditional mail and fax.
5. Available 24/7.

Disadvantages of email:

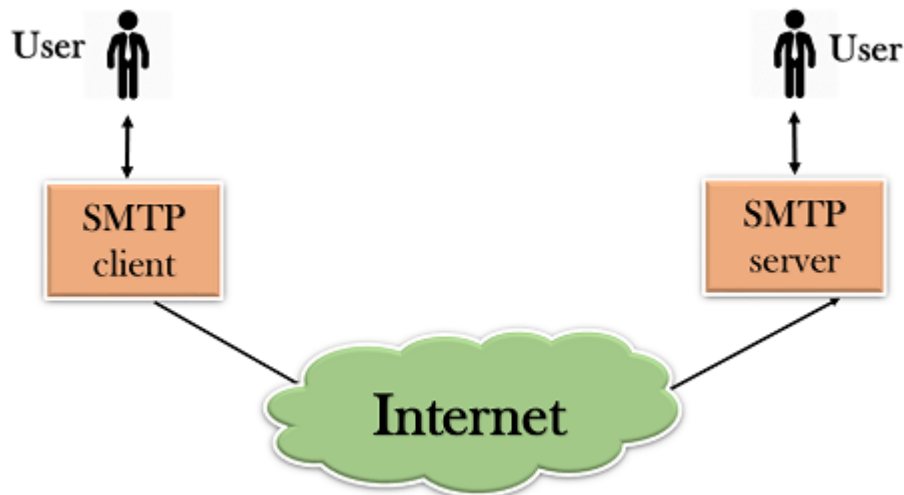
1. Risk of spam and phishing attacks.
2. Overwhelming amount of emails can lead to information overload.
3. Can lead to decreased face-to-face communication and loss of personal touch.
4. Potential for miscommunication due to lack of tone and body language in written messages.
5. Technical issues, such as server outages, can disrupt email service.
6. It is important to use email responsibly and effectively, for example, by keeping the subject line clear and concise, using proper etiquette, and protecting against security threats.

## SMTP

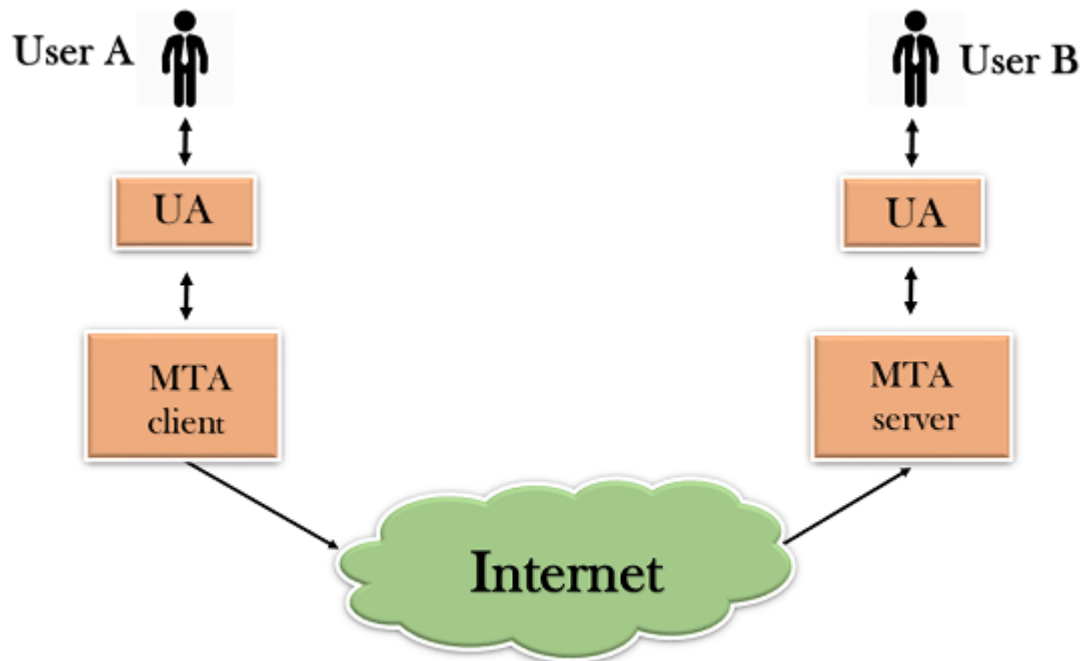
- SMTP stands for Simple Mail Transfer Protocol.
- SMTP is a set of communication guidelines that allow software to transmit an electronic mail over the internet is called **Simple Mail Transfer Protocol**.
- It is a program used for sending messages to other computer users based on e-mail addresses.

- It provides a mail exchange between users on the same or different computers, and it also supports:
  - It can send a single message to one or more recipients.
  - Sending message can include text, voice, video or graphics.
  - It can also send the messages on networks outside the internet.
- The main purpose of SMTP is used to set up communication rules between servers. The servers have a way of identifying themselves and announcing what kind of communication they are trying to perform. They also have a way of handling the errors such as incorrect email address. For example, if the recipient address is wrong, then receiving server reply with an error message of some kind.

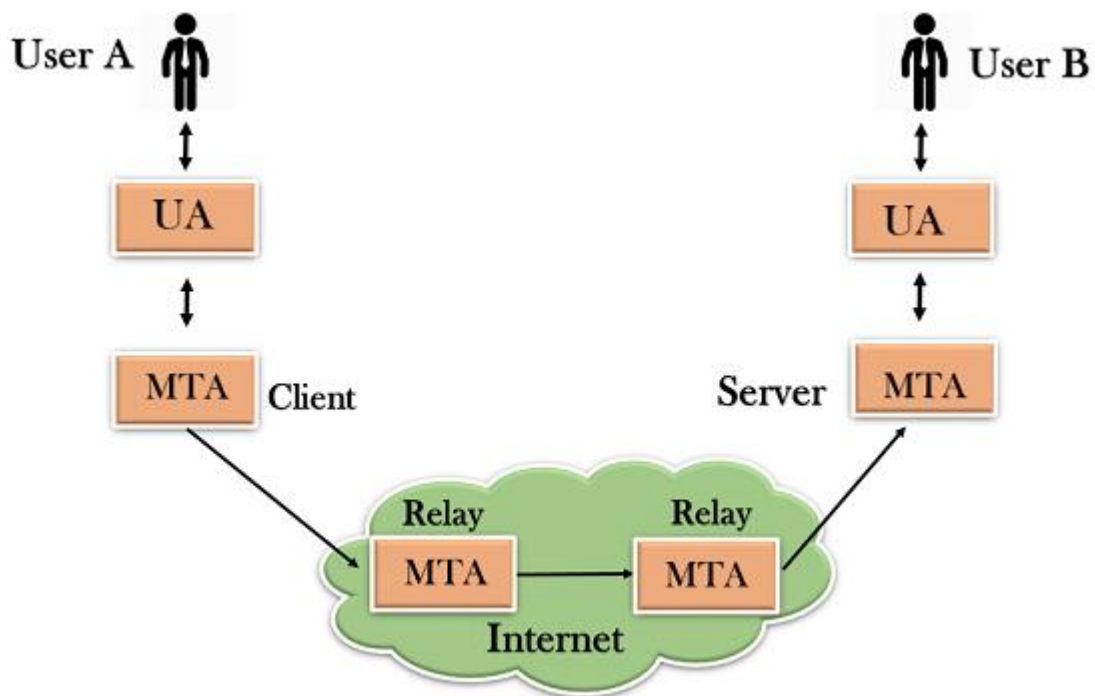
## Components of SMTP



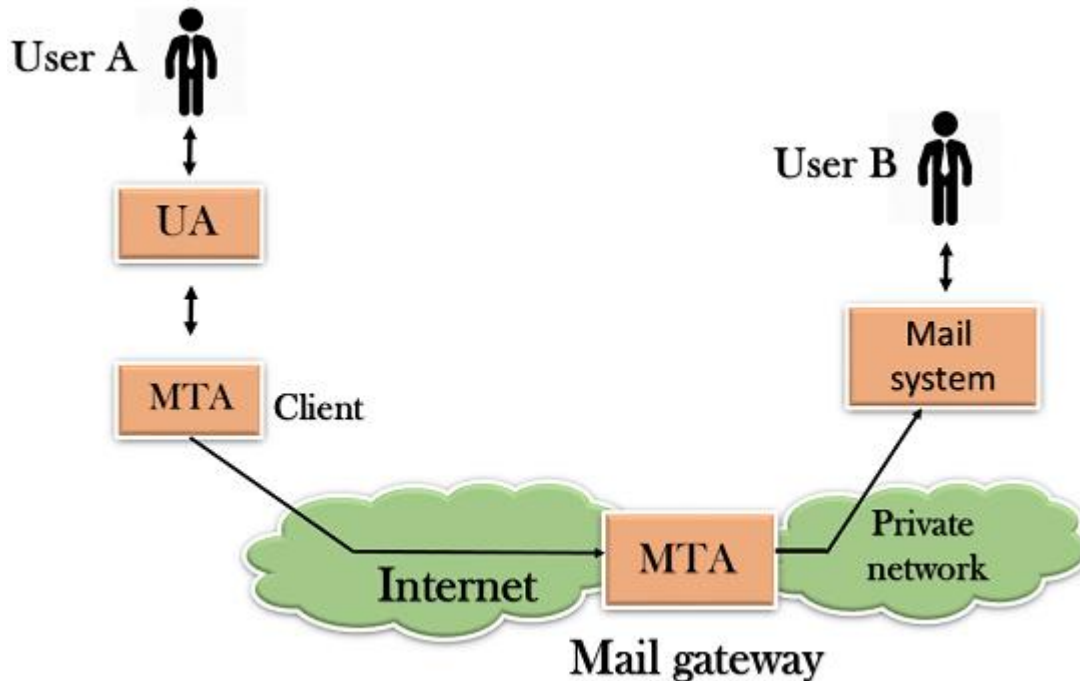
- First, we will break the SMTP client and SMTP server into two components such as user agent (UA) and mail transfer agent (MTA). The user agent (UA) prepares the message, creates the envelope and then puts the message in the envelope. The mail transfer agent (MTA) transfers this mail across the internet.



- SMTP allows a more complex system by adding a relaying system. Instead of just having one MTA at sending side and one at receiving side, more MTAs can be added, acting either as a client or server to relay the email.



- The relaying system without TCP/IP protocol can also be used to send the emails to users, and this is achieved by the use of the mail gateway. The mail gateway is a relay MTA that can be used to receive an email.



## Working of SMTP

1. **Composition of Mail:** A user sends an e-mail by composing an electronic mail message using a Mail User Agent (MUA). Mail User Agent is a program which is used to send and receive mail. The message contains two parts: body and header. The body is the main part of the message while the header includes information such as the sender and recipient address. The header also includes descriptive information such as the subject of the message. In this case, the message body is like a letter and header is like an envelope that contains the recipient's address.
2. **Submission of Mail:** After composing an email, the mail client then submits the completed e-mail to the SMTP server by using SMTP on TCP port 25.
3. **Delivery of Mail:** E-mail addresses contain two parts: username of the recipient and domain name. For example, vivek@gmail.com, where "vivek" is the username of the recipient and "gmail.com" is the domain name. If the domain name of the recipient's email address is different from the sender's domain name, then MSA will send the mail to the Mail Transfer Agent (MTA). To relay the email,

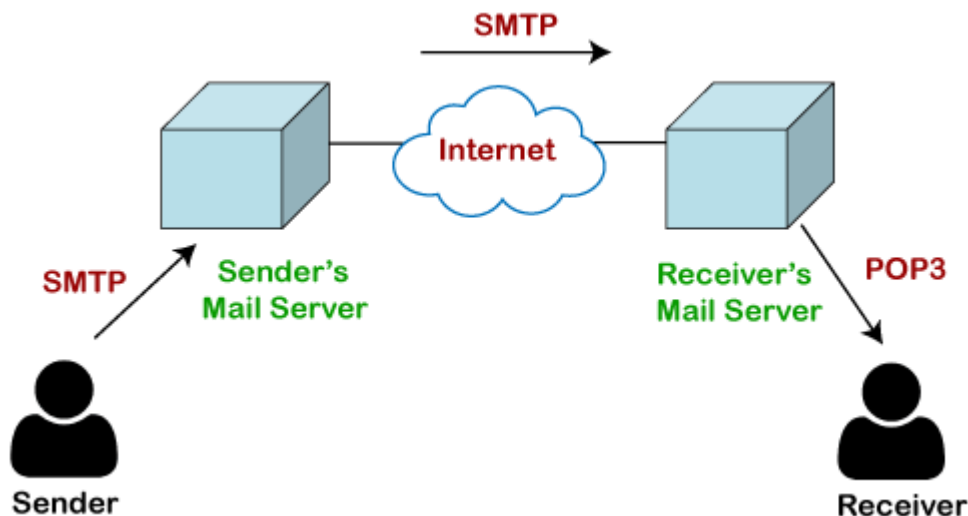
the MTA will find the target domain. It checks the MX record from Domain Name System to obtain the target domain. The MX record contains the domain name and IP address of the recipient's domain. Once the record is located, MTA connects to the exchange server to relay the message.

4. **Receipt and Processing of Mail:** Once the incoming message is received, the exchange server delivers it to the incoming server (Mail Delivery Agent) which stores the e-mail where it waits for the user to retrieve it.
5. **Access and Retrieval of Mail:** The stored email in MDA can be retrieved by using MUA (Mail User Agent). MUA can be accessed by using login and password.

## POP Protocol

The POP protocol stands for Post Office Protocol. As we know that SMTP is used as a message transfer agent. When the message is sent, then SMTP is used to deliver the message from the client to the server and then to the recipient server. But the message is sent from the recipient server to the actual server with the help of the Message Access Agent. The Message Access Agent contains two types of protocols, i.e., POP3 and IMAP.

### How is mail transmitted?



Suppose sender wants to send the mail to receiver. First mail is transmitted to the sender's mail server. Then, the mail is transmitted from the sender's mail server to the receiver's mail server over the internet. On receiving the mail at the receiver's mail server, the mail is then sent to the user. The whole process is done with the help of Email protocols. The transmission of mail from the sender to the sender's mail server and then to the receiver's

mail server is done with the help of the [SMTP protocol](#). At the receiver's mail server, the POP or [IMAP protocol](#) takes the data and transmits to the actual user.

Since SMTP is a push protocol so it pushes the message from the client to the server. As we can observe in the above figure that SMTP pushes the message from the client to the recipient's mail server. The third stage of email communication requires a pull protocol, and POP is a pull protocol. When the mail is transmitted from the recipient mail server to the client which means that the client is pulling the mail from the server.

## What is POP3?

The POP3 is a simple protocol and having very limited functionalities. In the case of the POP3 protocol, the POP3 client is installed on the recipient system while the POP3 server is installed on the recipient's mail server.

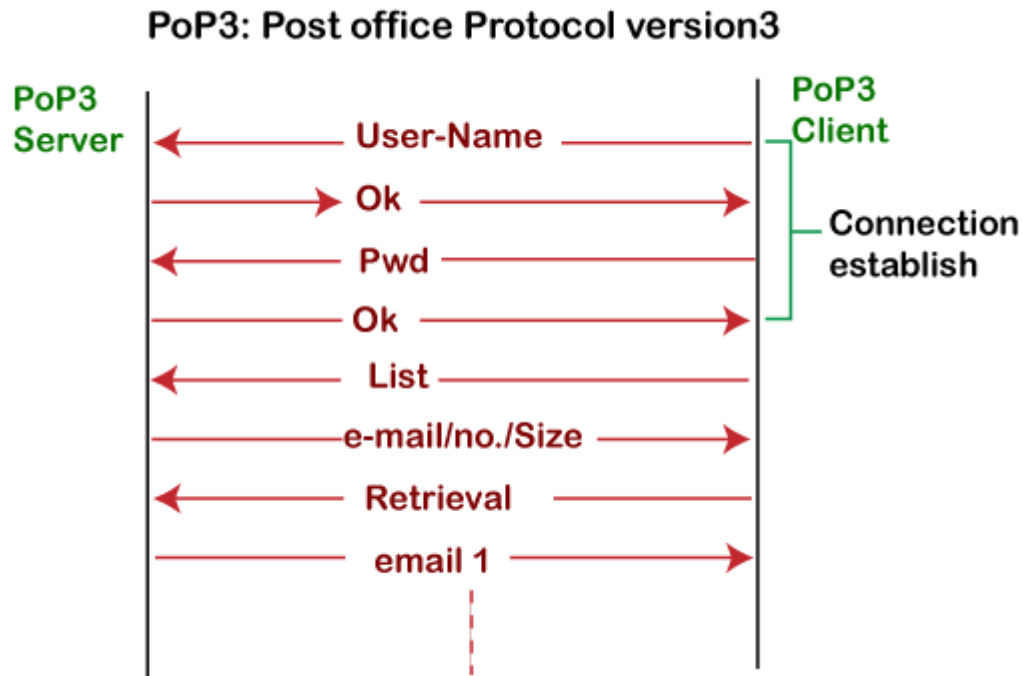
## History of POP3 protocol

The first version of post office protocol was first introduced in 1984 as RFC 918 by the [internet](#) engineering task force. The developers developed a simple and effective email protocol known as the POP3 protocol, which is used for retrieving the emails from the server. This provides the facility for accessing the mails offline rather than accessing the mailbox offline.

In 1985, the post office protocol version 2 was introduced in RFC 937, but it was replaced with the post office protocol version 3 in 1988 with the publication of RFC 1081. Then, POP3 was revised for the next 10 years before it was published. Once it was refined completely, it got published on 1996.

Although the POP3 protocol has undergone various enhancements, the developers maintained a basic principle that it follows a three-stage process at the time of mail retrieval between the client and the server. They tried to make this protocol very simple, and this simplicity makes this protocol very popular today.

**Let's understand the working of the POP3 protocol.**



To establish the connection between the POP3 server and the POP3 client, the POP3 server asks for the user name to the POP3 client. If the username is found in the POP3 server, then it sends the ok message. It then asks for the password from the POP3 client; then the POP3 client sends the password to the POP3 server. If the password is matched, then the POP3 server sends the OK message, and the connection gets established. After the establishment of a connection, the client can see the list of mails on the POP3 mail server. In the list of mails, the user will get the email numbers and sizes from the server. Out of this list, the user can start the retrieval of mail.

Once the client retrieves all the emails from the server, all the emails from the server are deleted. Therefore, we can say that the emails are restricted to a particular machine, so it would not be possible to access the same mails on another machine. This situation can be overcome by configuring the email settings to leave a copy of mail on the mail server.

## Advantages of POP3 protocol

**The following are the advantages of a POP3 protocol:**

- It allows the users to read the email offline. It requires an internet connection only at the time of downloading emails from the server. Once the mails are downloaded from the server, then all the downloaded mails reside on our PC or hard disk of our computer, which



can be accessed without the internet. Therefore, we can say that the POP3 protocol does not require permanent internet connectivity.

- It provides easy and fast access to the emails as they are already stored on our PC.
- There is no limit on the size of the email which we receive or send.
- It requires less server storage space as all the mails are stored on the local machine.
- There is maximum size on the mailbox, but it is limited by the size of the hard disk.
- It is a simple protocol so it is one of the most popular protocols used today.
- It is easy to configure and use.

## Disadvantages of POP3 protocol

**The following are the advantages of a POP3 protocol:**

- If the emails are downloaded from the server, then all the mails are deleted from the server by default. So, mails cannot be accessed from other machines unless they are configured to leave a copy of the mail on the server.
- Transferring the mail folder from the local machine to another machine can be difficult.
- Since all the attachments are stored on your local machine, there is a high risk of a virus attack if the virus scanner does not scan them. The virus attack can harm the computer.
- The email folder which is downloaded from the mail server can also become corrupted.
- The mails are stored on the local machine, so anyone who sits on your machine can access the email folder.

## What is World Wide Web?

World Wide Web, which is also known as a Web, is a collection of websites or web pages stored in web servers and connected to local computers through the internet. These websites contain text pages, digital images, audios, videos, etc. Users can access the content of these sites from any part of the world over the internet using their devices such as computers, laptops, cell phones, etc. The WWW, along with internet, enables the retrieval and display of text and media to your device.



The building blocks of the Web are web pages which are formatted in HTML and connected by links called "hypertext" or hyperlinks and accessed by HTTP. These links are electronic connections that link related pieces of information so that users can access the desired information quickly. Hypertext offers the advantage to select a word or phrase from text and thus to access other pages that provide additional information related to that word or phrase.

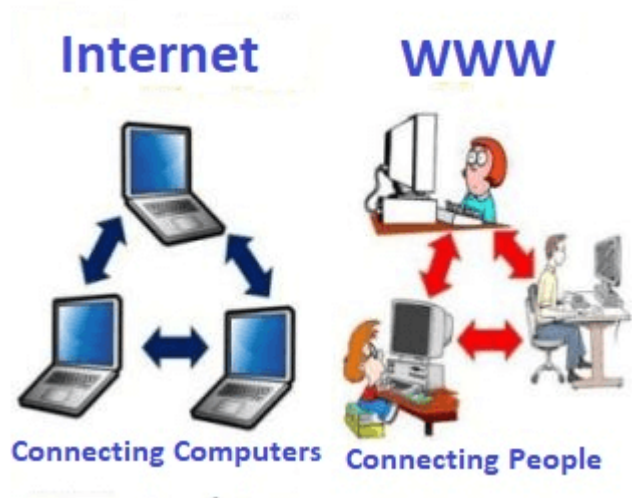
A web page is given an online address called a Uniform Resource Locator (URL). A particular collection of web pages that belong to a specific URL is called a website, e.g., *www.facebook.com*, *www.google.com*, etc. So, the World Wide Web is like a huge electronic book whose pages are stored on multiple servers across the world.

Small websites store all of their WebPages on a single server, but big websites or organizations place their WebPages on different servers in different countries so that when users of a country search their site they could get the information quickly from the nearest server.

So, the web provides a communication platform for users to retrieve and exchange information over the internet. Unlike a book, where we move from one page to another in a sequence, on World Wide Web we follow a web of hypertext links to visit a web page and from that web page to move to other web pages. You need a browser, which is installed on your computer, to access the Web.

## Difference between World Wide Web and Internet:

Some people use the terms 'internet' and 'World Wide Web' interchangeably. They think they are the same thing, but it is not so. Internet is entirely different from WWW. It is a worldwide network of devices like computers, laptops, tablets, etc. It enables users to send emails to other users and chat with them online. For example, when you send an email or chatting with someone online, you are using the internet.



But, when you have opened a website like google.com for information, you are using the World Wide Web; a network of servers over the internet. You request a webpage from your computer using a browser, and the server renders that page to your browser. Your computer is called a client who runs a program (web browser), and asks the other computer (server) for the information it needs.

## History of the World Wide Web:



The World Wide Web was invented by a British scientist, Tim Berners-Lee in 1989. He was working at CERN at that time. Originally, it was developed by him to fulfill the need of automated information sharing between scientists across the world, so that they could easily share the data and results of their experiments and studies with each other.

CERN, where Tim Berners worked, is a community of more than 1700 scientists from more than 100 countries. These scientists spend some time on CERN site, and rest of the time they work at their universities and national laboratories in their home countries, so there was a need for reliable communication tools so that they can exchange information.

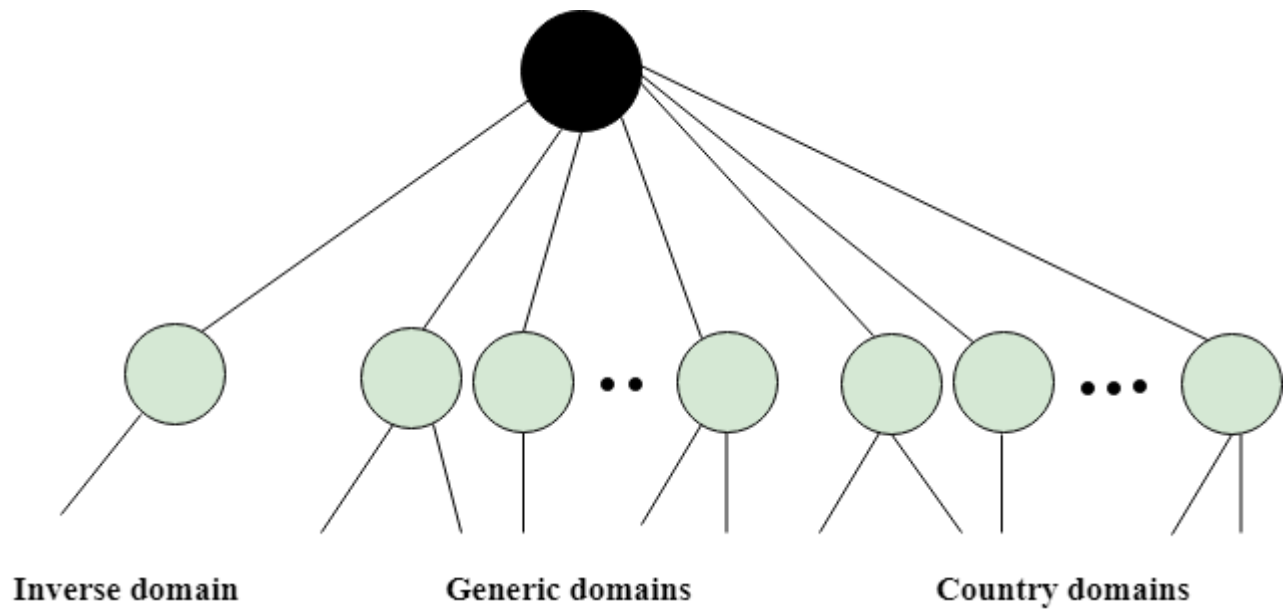
Internet and Hypertext were available at this time, but no one thought how to use the internet to link or share one document to another. Tim focused on three main technologies that could make computers understand each other, HTML, URL, and HTTP. So, the objective behind the invention of WWW was to combine recent computer technologies, data networks, and hypertext into a user-friendly and effective global information system.

## DNS

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.
- Each node in a tree has a domain name, and a full domain name is a sequence of symbols specified by dots.
- 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.
- For example, suppose the FTP site at EduSoft 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.

DNS is a TCP/IP protocol used on different platforms. The domain name space is divided into three different sections: generic domains, country domains, and inverse domain.



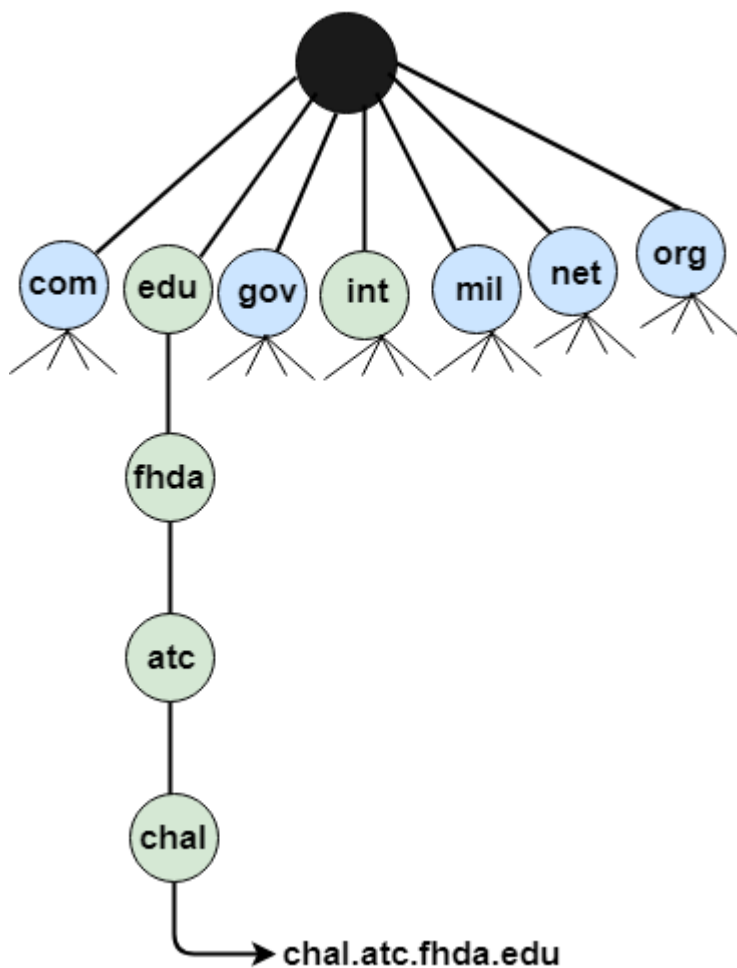
## Generic Domains

- It defines the registered hosts according to their generic behavior.
- Each node in a tree defines the domain name, which is an index to the DNS database.
- It uses three-character labels, and these labels describe the organization type.

| Label | Description                        |
|-------|------------------------------------|
| aero  | Airlines and aerospace companies   |
| biz   | Businesses or firms                |
| com   | Commercial Organizations           |
| coop  | Cooperative business Organizations |
| edu   | Educational institutions           |
| gov   | Government institutions            |
| info  | Information service providers      |

|        |                                        |
|--------|----------------------------------------|
| int    | International Organizations            |
| mil    | Military groups                        |
| museum | Museum & other nonprofit organizations |
| name   | Personal names                         |
| net    | Network Support centers                |
| org    | Nonprofit Organizations                |
| pro    | Professional individual Organizations  |

### Root level



## Country Domain

The format of country domain is same as a generic domain, but it uses two-character country abbreviations (e.g., us for the United States) in place of three character organizational abbreviations.

## Inverse Domain

The inverse domain is used for mapping an address to a name. When the server has received a request from the client, and the server contains the files of only authorized clients. To determine whether the client is on the authorized list or not, it sends a query to the DNS server and ask for mapping an address to the name.

## Working of DNS

- DNS is a client/server network communication protocol. DNS clients send requests to the server while DNS servers send responses to the client.
- Client requests contain a name which is converted into an IP address known as a forward DNS lookups while requests containing an IP address which is converted into a name known as reverse DNS lookups.
- DNS implements a distributed database to store the name of all the hosts available on the internet.
- If a client like a web browser sends a request containing a hostname, then a piece of software such as **DNS resolver** sends a request to the DNS server to obtain the IP address of a hostname. If DNS server does not contain the IP address associated with a hostname, then it forwards the request to another DNS server. If IP address has arrived at the resolver, which in turn completes the request over the internet protocol.

## HTTP

- HTTP stands for **HyperText Transfer Protocol**.
- It is a protocol used to access the data on the World Wide Web (www).
- The HTTP protocol can be used to transfer the data in the form of plain text, hypertext, audio, video, and so on.

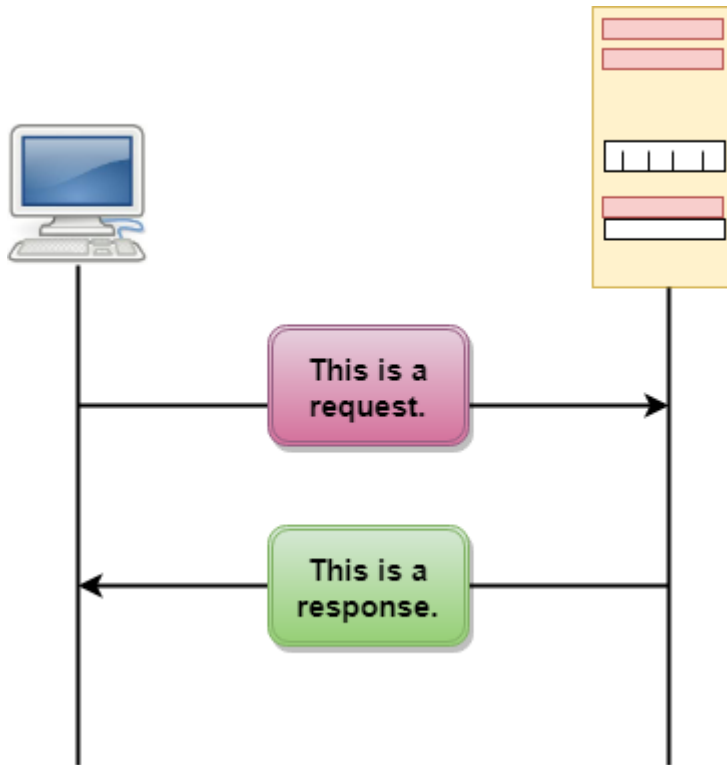
- This protocol is known as HyperText Transfer Protocol because of its efficiency that allows us to use in a hypertext environment where there are rapid jumps from one document to another document.
- HTTP is similar to the FTP as it also transfers the files from one host to another host. But, HTTP is simpler than FTP as HTTP uses only one connection, i.e., no control connection to transfer the files.
- HTTP is used to carry the data in the form of MIME-like format.
- HTTP is similar to SMTP as the data is transferred between client and server. The HTTP differs from the SMTP in the way the messages are sent from the client to the server and from server to the client. SMTP messages are stored and forwarded while HTTP messages are delivered immediately.

## Features of HTTP:

- **Connectionless protocol:** HTTP is a connectionless protocol. HTTP client initiates a request and waits for a response from the server. When the server receives the request, the server processes the request and sends back the response to the HTTP client after which the client disconnects the connection. The connection between client and server exist only during the current request and response time only.
- **Media independent:** HTTP protocol is a media independent as data can be sent as long as both the client and server know how to handle the data content. It is required for both the client and server to specify the content type in MIME-type header.
- **Stateless:** HTTP is a stateless protocol as both the client and server know each other only during the current request. Due to this nature of the protocol, both the client and server do not retain the information between various requests of the web pages.



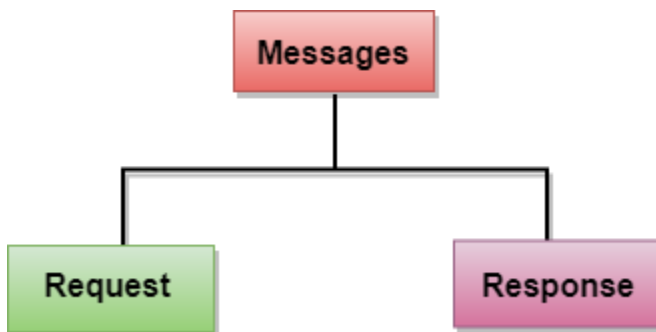
# HTTP Transactions



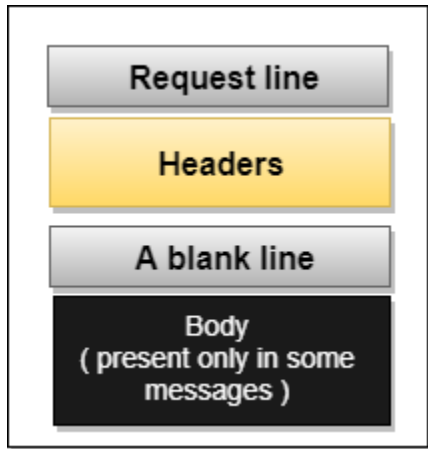
The above figure shows the HTTP transaction between client and server. The client initiates a transaction by sending a request message to the server. The server replies to the request message by sending a response message.

## Messages

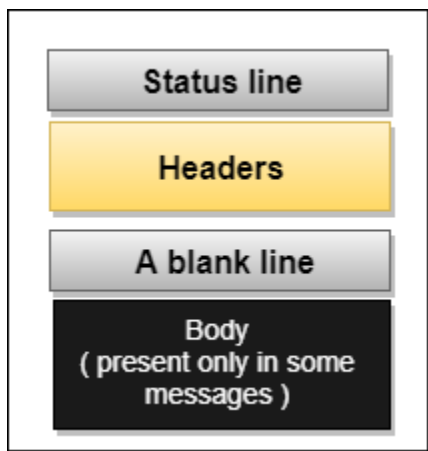
HTTP messages are of two types: request and response. Both the message types follow the same message format.



**Request Message:** The request message is sent by the client that consists of a request line, headers, and sometimes a body.



**Response Message:** The response message is sent by the server to the client that consists of a status line, headers, and sometimes a body.



## Uniform Resource Locator (URL)

- A client that wants to access the document in an internet needs an address and to facilitate the access of documents, the HTTP uses the concept of Uniform Resource Locator (URL).
- The Uniform Resource Locator (URL) is a standard way of specifying any kind of information on the internet.
- The URL defines four parts: method, host computer, port, and path.



- **Method:** The method is the protocol used to retrieve the document from a server. For example, HTTP.
- **Host:** The host is the computer where the information is stored, and the computer is given an alias name. Web pages are mainly stored in the computers and the computers are given an alias name that begins with the characters "www". This field is not mandatory.
- **Port:** The URL can also contain the port number of the server, but it's an optional field. If the port number is included, then it must come between the host and path and it should be separated from the host by a colon.
- **Path:** Path is the pathname of the file where the information is stored. The path itself contain slashes that separate the directories from the subdirectories and files.

## HTML - Hypertext Markup Language



Today, everyone is well aware of the usefulness of the website and the impact it is creating in the lives of people. Even you are able to access this content using a website. Whether it be a social networking, music service provider website, or educational content website, every website is created using **Hypertext Markup Language (HTML)**. We could find the significance of HTML by having a close look at the website building process. It is based on HTML code written in an editor, and it provides the necessary structure for the components of the page. The components are header, footer, main content, and inline elements. Hence, HTML is a useful tool that facilitates webpage creation with the help of HTML elements.

## What is HTML?

HTML is the abbreviation for Hypertext Markup Language. This standard markup language is used for the creation of web pages. HTML consists of the elements that guide the browser about how to display the content, and a simple HTML document consists of the tags.

HTML is the standard language that is used for designing the content on the website. Cascading Style Sheets (CSS) and Javascript (JS) are the technologies that help the HTML perform better while designing a website. Hence these technologies can be used along with the HTML.

HTML is responsible for rendering the HTML documents from the local memory locations and web servers to the multimedia pages. HTML includes cues of the actual document. Additionally, it also describes the structure of the web page.



## History and Development of HTML

- Elements in HTML act as the building blocks for HTML pages. Interactive forms can also be added to the rendered page, including constructs, images, and other objects. Structured documents can be created with the help of HTML, and it is possible with the help of structural semantics like headings, lists, paragraphs, quotes, links, and other items. Angle bracket is used for the tags that form the HTML elements. Some tags can directly input content to the page, for example tags like `<img/>` and `<input/>`. However, many tags like `<p>` surround the content and provide the information of a document. Such tags may include sub-elements. The task of the browser is not to display the tags but interpret the page's content and display it to the user.
- Programs written in other scripting languages, such as JavaScript, can be embedded in the HTML. The look and layout of the content can also be changed by the inclusion of CSS.

The use of CSS is encouraged by the **World Wide Consortium (W3C)**, which currently maintains CSS standards and is formerly responsible for the HTML standards.

- **Tim Berners-Lee** was a physicist and a contractor at CERN. He proposed a system for using and sharing information; the prototype was known as ENQUIRE. Berners went on to propose an internet-based system which was a hypertext system in **1989**. In **1990**, Berners-Lee wrote browser and server software along with specifying HTML. Berners Lee and Robert Cailliau collaborated for funding; however, CERN didn't adopt their project formally. **Robert Cailliau** was a data system engineer at CERN. He also listed the areas in which hypertext can be used, and the encyclopedia was first. It was the list stated in his notes.
- "HTML Tags" was the first description of the HTML document that was available publicly, and it was mentioned by Tim Berners-Lee for the first time in **1991**. There were about 18 elements that consisted of initial and simple HTML design.
- Some of the elements still exist in HTML 4, and these are about 11. These were influenced by the Standard Generalized Markup Language (SGML)-based documentation. SGML is in house documentation format at CERN.
- HTML is used to compose text, images, etc. into web pages; it is used to interpret the content as a markup language.

## Tags in HTML

There are many tags in HTML which facilitate the process of describing web page to the server. These tags are arranged in a specific manner to get the desired output. Some of the important tags of HTML are discussed below -

1. **<!Doctype>** - It describes the type of document
2. **<html>** - describes that the document is an html document
3. **<title>** - describes title of the document
4. **<body>** - describes the body of the document
5. **<h1> to <h6>** - describes the headings in html
6. **<p>** - describes a paragraph
7. **<br>** - produces a line break of single line
8. **<hr>** - describes a thematic content change

9. `<!--?-->` - Describes a comment

Tags in html are useful entities that allow a developer to manipulate the look and functionalities of a web page. Additionally, some of the tags have closing tags that include a forward slash, indicating a closing tag. Examples of the closing tags are `body`, `title`, and `html`. However, some tags do not have closing tags; examples of such tags are `<br>` and `<img>`. Such elements are often known as empty or void elements in HTML.

Other useful tags in html are -

1. `<b>` - It is used to bold the written text.
2. `<i>` - It is used to make the written text italic.
3. `<u>` - It is used to underline the content written in the tag.
4. `<li>` - It is used to form a list in the content.
5. `<ol>` - It is used to form an ordered list in the content.
6. `<ul>` - It is used to form an unordered list in the content.
7. `<marquee>` - Adds scrolling effect to the text or image in the content.

## Elements in HTML

HTML elements are enclosed by a starting tag and a closing tag and contain other content in them. The closing tags are closed using a forward slash. However, some elements do not require closing tags, and these are known as **void elements**.

## Difference between HTML Tags and HTML Elements

| HTML Tags                                                                      | HTML Elements                                                            |
|--------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Tags in HTML hold the HTML element.                                            | Elements in HTML hold the content.                                       |
| Tags in HTML start with '<' and end with '>'.                                  | HTML elements are described as anything that is written within the tags. |
| Tags in HTML are similar to the keywords, and each tag has a specific meaning. | General content is specified by the HTML element.                        |

# Uses of HTML

HTML acts as the markup language which is used for the creation of web pages and web applications. It is excessively utilized in web development when combined with JavaScript and CSS.

Some of the most important applications of HTML are listed below

1. **Web Development:**

Creating web pages is one of the most used applications of HTML. All pages contain some tags, and these tags define the pages and hyperlinks are used to connect the pages. Every page is written with the help of HTML code.

2. **Web Document Creation:**

Another use of HTML is in document creation. Document creation is done using the basic tags and Document Object Model (DOM). There are three sections of the web document: title, head, and body. The information that identifies the document, such as the title and important keywords, is included in the head. The title is visible in the browser, whereas the body section contains the main content viewed by the consumer. These sections are created using HTML tags, and it is done using specific tags for a purpose. Hence the title, head, and body are kept in a loop.

3. **Internet Navigation:**

Internet navigation is possible through hypertext functionality. It is one of the most widely used functions of HTML. It is the text that refers the user to another webpage, facilitating the user to navigate easily. It makes navigation within the web pages and easy, even if these are on different servers.

4. **Cutting Edge Feature:**

HTML5 is the game-changer in the website development business. HTML5 has the latest set of standards and APIs that can be perfectly implemented using Google Chrome.

5. **Responsive images on web pages:**

Responsive images can be set up using queries. The developer can fully control how the user can render the image, and it can be done using the srcset attribute of the img element by combining it with picture elements. Hence images with size variations can be uploaded.

6. **Use in Game Development:**

The advent of HTML5 and CSS3 and a light-fast JavaScript engine provide a rich

experience to the user. For game development, most of the API features can be utilized while others can be eliminated.

#### 7. **Data Entry Support with HTML:**

Data entry work can be easily carried out using HTML5. Using the HTML5 standards and APIs, the fields such as text, data, etc., can be added. It is useful for data entry, and end-user will have better data entry.



## Web Development Using HTML

Heading, paragraphs, images, data, etc., are the components of a webpage. HTML has a heavy application in web development, as developers use HTML for frontend development. The task of the HTML element is to specify the type of information on each webpage. For example, a paragraph is indicated by 'p'. HTML codes are written to relate different items on one page.

### Working of HTML in a Webpage

The HTML is crucial for any webpage that you wish to design. The HTML code is used for designing the basic structure of the website, and it can be used to manipulate elements with the help of CSS and JavaScript.

For example, HTML can be used for specifying headings on the web page. Afterwards, it can be edited using the appropriate size, color, and size that fits the organization's branding.

HTML is used to indicate the function of structural elements to the search engine and makes the summarizing of content easier.

For example, Mark captions an image as a "figcaption" element and enclose the image and caption in the 'figure' element. In this case, it is easy for the search engine to understand that the two pieces are related to each other, and the caption is the description of the stored image.



## What is HTML5?

[HTML5](#) is the latest HTML version that is utilized in structuring content on the website. It is designed such that it can present every content that can be consumed on the World Wide Web.

## Difference between HTML and HTML5

| HTML                                                                                      | HTML5                                                                                   |
|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| It requires flash player support to support audio and video; otherwise, it won't support. | It can easily support audio and video with the help of <audio> and <video> tags.        |
| Cookies are used to store temporary data.                                                 | Temporary data is stored using SQL databases and application cache.                     |
| Running JavaScript in a browser is not allowed.                                           | HTML5 has JS Web worker API that makes it possible to run JavaScript in the background. |
| VML, Silver-light, flash, etc., are the technologies that make vector graphics possible.  | Vector graphic is integral to HTML5.                                                    |
| Drag and drop effects are not allowed.                                                    | Drag and drop effects are allowed.                                                      |
| Shapes such as circles, rectangles, triangles are not possible in HTML.                   | Shapes such as circles, rectangles, triangles are allowed in HTML5.                     |
| Old browsers support HTML.                                                                | New browsers such as Mozilla, Chrome, etc., support HTML5.                              |
| Long and complicated Doctype declaration.                                                 | Simple and easy Doctype declaration.                                                    |
| HTML is less mobile-friendly.                                                             | HTML5 is more mobile-friendly.                                                          |
| Absence of elements such as nav, header.                                                  | New elements such as nav, header are present.                                           |

|                                                                        |                                                                       |
|------------------------------------------------------------------------|-----------------------------------------------------------------------|
| Long and complex character encoding.                                   | Easy character encoding as compared to HTML.                          |
| Geolocation tracking using the web browser is almost impossible.       | Geolocation tracking is possible with the help of JS Geolocation API. |
| HTML cannot handle the inaccurate syntax.                              | HTML5 can handle the inaccurate syntax.                               |
| Some attributes, such as charset, ping, etc., are not present in HTML. | HTML5 includes attributes such as charset, ping, etc.                 |

## Advantages and Disadvantages of HTML

### Advantages

- It is a widely used language.
- It is supported by all browsers.
- It can be easily learnt and used.
- Light-weighted, fast to load content.
- There is no need for extra software for its operation as it is built by default in every window.
- Ease of use.
- Loose Syntax.
- Writing HTML is easier compared to other languages.
- New programmers can easily code in HTML and gain expertise.
- The web page designing is further simplified with the use of templates.
- Has many applications and hence is useful for web designing beginners.
- Acts as support for all websites.
- Built and operated on almost all websites.
- The use of HTML for data storage has increased significantly.
- No need to buy any software as HTML is free of cost.
- Present by default and hence saves money for the developer.
- Code writing is short and convenient as there are several tags and attributes which make the task easier.

## Disadvantages

- HTML is a static language and, therefore, cannot produce dynamic results alone.
- Many times the structure of HTML becomes tough to understand.
- Errors can cost you money.
- It takes time to create tables, lists, forms and maintain colour schemes. Hence it is time-consuming.
- HTML is not useful for dynamic pages.
- Simple web page creation requires writing a long code.
- Checking the deprecated tags is required because if HTML is used with another language, it may overwrite the code and change the tags. Therefore, it would require knowledge of the other language as well.
- HTML has limited security features.
- Complexity is produced if we wish to write long code in HTML.
- HTML only creates plain and static pages.
- Easy web pages would also require rigorous code writing.