Unit -1
WEBSITE BASICS, HTML 5, CSS 3, WEB 2.0

1.1 WEB ESSENTIALS

WEBSITES

A website is a set of related web pages typically served from a single web server.

A website is hosted on at least one web server, accessible via a network such as the internet or a private local area network. The pages of a website can usually be accessed from a simple Uniform Resource Locator (URL) otherwise called as web address. The URLs of the pages organize them into a hierarchy.

Terminologies:

Internet: The Internet is a collection of computers around the world connected to each other via high speed series of networks.

World Wide Web (WWW): The World Wide Web – or Web consists of a vast assortment of files and documents that are stored on the computers and written in some form of Hyper Text Markup Language (HTML).

Servers: The computers that store the files are called servers because they can serve requests from many users at the same time.

Browsers: A Web browser is a program that displays Web pages and other documents on the web. Examples: Internet explorer, Firefox, Google Chrome etc.

HTML: HTML, or Hyper Text Markup Language, is the authoring language that describes how a Web page should be displayed by a Web browser. It has two essential features:

- **Hypertext:** When a visitor clicks a link on a Web page, it leads to another web page
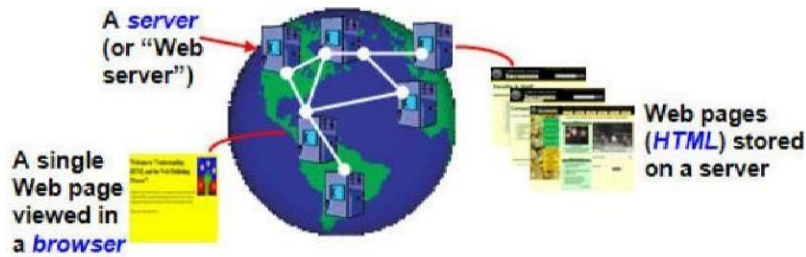
Fig 1.1 Internet and its components

INTERNET

The Internet is a vast, electronic network connecting many millions of computers from every corner of the world.The Internet is a global network of networks.

The Internet is a publicly-accessible network that consists of millions of smaller domestic, academic, business, and government networks. The Internet links are computer networks all over the world so that users can share resources and communicate with each other.People and organizations connect into the Internet so they can access its massive store of shared information.

The internet is a participative medium. Anybody can publish information or create new services.The internet is a **cooperative endeavor** - no organization is in charge of the internet.The following components are essential for an internet connection: Computer, Connection - Phone Line, Cable, DSL, Wireless, Modem, Network Software - TCP/IP, Application Software - Web Browser, Email, etc and Internet Service Provider (ISP).
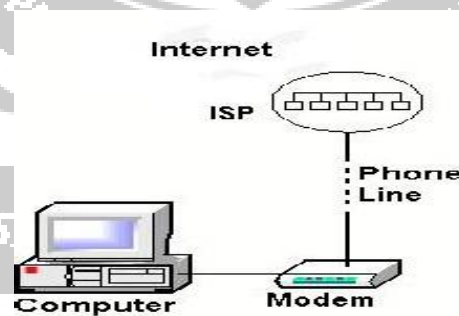

Fig 1.2 Components of internet

Evolution of Internet

The concept of Internet was originated in 1969 and has undergone several technological & infrastructural changes:

- The origin of Internet devised from the concept of Advanced Research Project Agency Network (ARPANET). **ARPANET** was developed by United States Department of Defense.

- Basic purpose of ARPANET was to provide communication among the various bodies of government.

- In 1972, the ARPANET spread over the globe with 23 nodes located at different countries and thus became known as Internet.

- By the time, with invention of new technologies such as TCP/IP protocols, DNS, WWW, browsers, scripting languages etc, Internet provided a medium to publish and access information over the web.

Internet Terminologies

- ➢ **Host:** A computer connected to the Internet is commonly referred to as a host.

- ➢ **Communication services:** The data is passed back and forth between host computers using packets and protocols, such as electronic mail (e-mail) for messaging, file transfer protocol (FTP) for moving files, telnet for accessing information, hypertext transfer protocol (HTTP) for serving up Web sites, custom protocols, etc. They are called communication services.

- ➢ **Internet Service Provider (ISP):** The Internet itself is decentralized-no one is completely responsible or has total control; however, the connection to the Internet is partly controlled by an Internet Service Provider (ISP). Example for ISP: Reliance, Airtel, Idea (IIN) etc.

- ➢ **Online:** When the computer is connected to the internet then it is in online.

- ➢ **Hyperlinks:** Allow a user to quickly move from one web page to another, even if the pages are on different servers in different parts of the world.

- ➢ **Protocols:** They are pre-established means of communication. **Example:** TCP/IP, SMTP.

- ➢ **TCP/IP:** TCP is the protocol that establishes a virtual connection between a destination and a source. TCP guarantees delivery of data and also guarantees that packets will be delivered in the same order in which they were sent. Internet Protocol (IP) is responsible for packaging the little packets of information and delivering them.

- ➢ **Client/ Server model:** TCP/IP uses the client/server model of communication in which a computer user (a client) requests and is provided a service (such as sending a Web page) by another computer (a server) in the network.

- ➤ **IP address:** It is the address of the machine. It is a four byte unique number that identifies a system on the Internet.

- ➤ **Domain Name Services (DNS):** They link text to our numeric IP addresses, allowing users to use the DNS as a proxy for the IP address. The IP addresses are often provided by the ISP. Each site must register the name for a cost through a DNS hosting service. DNS host servers then are used to convert our text DNS address to its digital IP address equivalent.

  - • **Universal Resource Locators:** URL's are a way of identifying information on a server. A URL gives the protocol, the domain, the directory, and even the file. A URL consists of the following parts:protocol (such as http:// or ftp://), host name (the Web server's IP address or domain name), directory (i.e. folder) and file name

- ➤ **World Wide Web:** The World Wide Web consists of all the Web sites and pages served on the Internet via HTTP. It is a hypermedia-based system for browsing Internet sites. It is named the web because it is made of many sites linked together; users can travel from one site to another by clicking on hyperlinks. Text, graphics, sound, and video can all be accessed. **Tim Berners-Lee** invented the World Wide Web in 1989 while working at CERN, the European Particle Physics Laboratory.

BASIC INTERNET PROTOCOLS

> **Protocol is a set of mutually accepted and implemented rules at both ends of the communications channel for the proper exchange of information.**

TCP / IP

Transmission Control Protocol/Internet Protocol (TCP /IP) is a suite of communication protocols used to interconnect network devices on the internet. TCP/IP can also be used as a communications protocol in a private network.TCP/IP specifies how data is exchanged over the internet by providing end-to-end communications that identify how it should be broken into packets, addressed, transmitted, routed and received at the destination.

TCP/IP requires little central management, and it is designed to make networks reliable, with the ability to recover automatically from the failure of any device on the network.TCP defines how applications can create channels of communication across a network. It also manages how a message is assembled into smaller packets before they are then transmitted over the internet and reassembled in the right order at the destination address.

IP defines how to address and route each packet to make sure it reaches the right destination. Each gateway computer on the network checks this IP address to determine where to forward the message.A key element of IP is the IP address, which is simply a 32 -bit number. IP addresses are normally written as a sequence of four decimal numbers separated by periodsas in 192.0.34.166. When an application on the source computer wants to send

information to a destination, the application calls IP software on the source machine and provides it with data to be transferred along with an IP address for each of the source and destination computers.

The IP software running on the source creates a packet, which is a sequence of bits representing the data to be transferred along with the source and destination IP addresses and some other header information, such as the length of the data.If the destination computer is on the same local network as the source, then the IP softwarewill send the packet to the destination directly via this network.

If the destination is onanother network, the IP software will send the packet to a gateway, which is a device that is connected to the source computer's network as well as to at least one other network.The gateway will select a computer on one of the other networks to which it is attached and send the packet on to that computer. This process will continue, until the packet reaches the destination computer.

IP software on that computer will receive the packet and pass its data up to an application that is waiting for the data.TCP, the Transmission Control Protocol, is a higher-level protocol that extends IP to provide additional functionality, including reliable communication based on the concept of a connection.
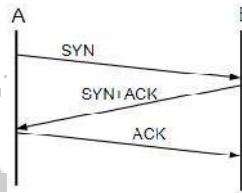

Fig 1.5: Three way handshaking of TCP

A connection is established between TCP software running on two machines by one of the machines sending a connection-request message via IP to the other. If the connection is accepted by B, then B returns a message to A requesting a connection in the other direction. If A responds affirmatively, thenthe connection is established. Notice that this means that A and B can both send messages. Once a connection has been established, TCP provides reliable data transmission by demanding an acknowledgment for each packet it sends via IP. TCP has contains port to communicate with many different applications on amachine.

UDP, DNS and Domain Names

UDP is connectionless and unreliable protocol. It doesn't require making a connection with the host to exchange data. Since UDP is unreliable protocol, there is no mechanism for ensuring that data sent is received. UDP transmits the data in form of a datagram.

UDP provides protocol port used i.e. UDP message contains both source and destination port number, that makes it possible for UDP software at the destination to deliver the message to correct application program.

Source port	Destination port
Length	UDP checksum
Data	

Fig 1.3: UDP Datagram

Differences between TCP and UDP

TCP	UDP
TCP is a connection-oriented protocol.	UDP is a connectionless protocol.
As a message makes its way across the internet from one computer to another. This is connection based.	UDP is also a protocol used in message transport or transfer. This is not connection based which means that one program can send a load of packets to another and that would be the end of the relationship.
TCP is suited for applications that require high reliability, and transmission time is relatively less critical	UDP is suitable for applications that need fast, efficient transmission, such as games. UDP's stateless nature is also useful for servers that answer small queries from huge numbers of clients.
TCP is used by HTTP, HTTPs, FTP, SMTP, Telnet	UDP is used by DNS, DHCP, TFTP, SNMP, RIP, VOIP.
TCP rearranges data packets in the order specified.	UDP has no inherent order as all packets are independent of each other. If ordering is required, it has to be managed by the application layer.
The speed for TCP is slower than UDP.	UDP is faster because there is no error-checking for packets.
There is absolute guarantee that the data transferred remains intact and arrives in the same order in which it was sent.	There is no guarantee that the messages or packets sent would reach at all.
TCP header size is 20 bytes	UDP Header size is 8 bytes.
Data is read as a byte stream, no distinguishing indications are transmitted to signal message (segment) boundaries	Packets are sent individually and are checked for integrity only if they arrive. Packets have definite boundaries which are honored upon receipt, meaning a read operation at the receiver socket will yield an entire message as it was originally sent

TCP is heavy-weight. TCP requires three packets to set up a socket connection, before any user data can be sent. TCP handles reliability and congestion control.	UDP is lightweight. There is no ordering of messages, no tracking connections, etc. It is a small transport layer designed on top of IP.
TCP does Flow Control. TCP requires three packets to set up a socket connection, before any user data can be sent. TCP handles reliability and congestion control.	UDP does not have an option for flow control
TCP does error checking	UDP does error checking, but no recovery options.
Acknowledge segments	No Acknowledgment
Handshaking is done	No handshake (connectionless protocol)

File Transfer Protocol (FTP)

FTP is used to copy files from one host to another. FTP offers the mechanism for the same in following manner:

➢ FTP creates two processes such as **Control Process and Data Transfer Process** at both ends i.e. at client as well as at server.

➢ FTP establishes two different connections: one is for data transfer and other is for control information.

➢ Control connection is made between control processes while Data Connection is made between data transfer process.

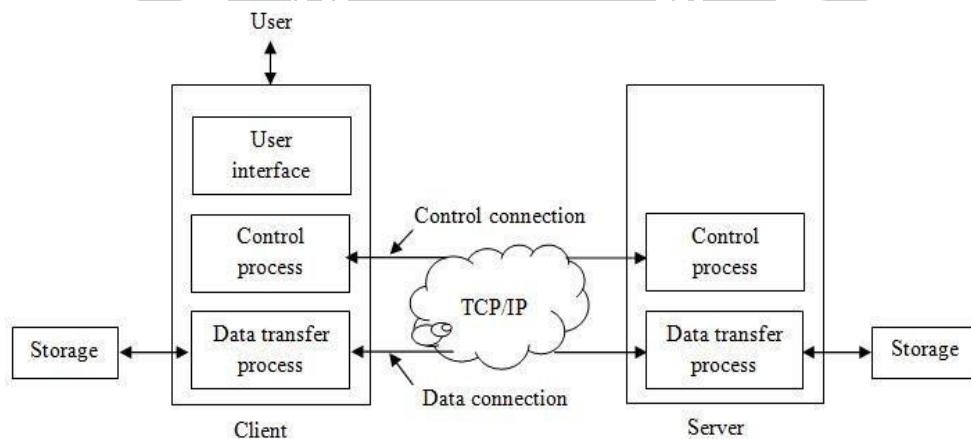➢ FTP uses port 21 for the control connection and Port 20 for the data connection.

Fig 1.4 FTP

Trivial File Transfer Protocol (TFTP)

Trivial File Transfer Protocol is also used to transfer the files but it transfers the files **without authentication**. Unlike FTP, TFTP does not separate control and data information. Since there is no authentication exists, TFTP lacks in security features therefore it is not recommended to use TFTP.

TFTP makes use of UDP for data transport. Each TFTP message is carried in separate UDP datagram.The first two bytes of a TFTP message specify the type of message. The TFTP session is initiated when a TFTP client sends a request to upload or download a file.The request is sent from an ephemeral UDP port to the UDP port 69 of a TFTP server.

Differences between FTP and TFTP

FTP	TFTP
Authentication is done before transferring files.	No authentication is done before transferring files
The underlying protocol employed is TCP.	The underlying protocol employed is UDP.
Port 20 is used as control port and 21 for data transfers.	Port numbers: 3214, 69, 4012 are used.
Reliable data transfer is provided.	Unreliable data transfer

Telnet

Telnet is a protocol used to log in to remote computer on the internet. There are a number of Telnet clients having user friendly user interface. The following diagram shows a person is logged in to computer A, and from there, he remotely logged into another computer B.
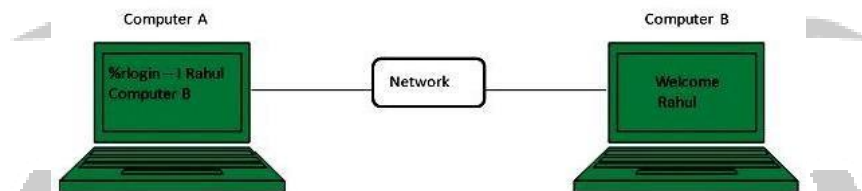

Fig 1.5 Telnet

Hyper Text Transfer Protocol (HTTP)

HTTP is a communication protocol. It defines mechanism for communication between browser and the web server. It is also called request and response protocol because the communication between browser and server takes place in request and response pairs. It is a**stateless** protocol (i.e.) the history of the communication between server and client is notstored in any form.