

Chapter One

The Fundamentals

Overview of the internet

- Internet is network of networks (i.e. when two or more devices are connected for sharing data or resources or exchange messages we call it networks)
- No company owns the internet (I.e.no central administration or control for the internet), it is cooperative effort governed by a system of standard and rules.

Origin of the internet

In the 1969 departments of defense (DOD of USA) started a network called ARPANET (advanced research projects administration network) with one computer at California and three at Utah.

Use and service of the internet

E-mail:- allows user to:

- Send a message to just only one user or group of users,
- To read, print, forward, answer (reply) or delete a message
- Attach large documents.


Telnet (Remote login):

- Is away that the users enable to logon to a remote computer and interactively access its resources.
- Uses a special protocol called Network Terminal Protocol
- Protocol is the standard set of rule that governs all communication over a computer network

File Transfer Protocol (FTP):-

- Used to download files from an FTP sites

World Wide Web

- Internet  web
- Internet is a collection of computers and other devices connected by equipment to communicate with each other.

WWW

- Is the most popular application or service of an internet.
- Also called web.
- It consists of a worldwide collection of electronic documents.
 - ❖ An electronic document on the Web is called a Web page, which can contain text, graphics, animation, audio, and video.
 - ❖ Web pages could be static (fixed) or dynamic (changing).
 - ❖ A Web site is a collection of related Web pages stored on the same web server.
 - ❖ A Web Server is a computer that delivers requested Web pages to your computer. The same Web server can store multiple Web sites.

World Wide Web is combination of four ideas:

- Hypertext
- Resource identifiers
- Client-server architecture
- Markup language

Hypertext, that is the ability, in a computer environment, to move from one part of a document to another or from one document to another through internal connections among these documents (called "hyperlinks");

Resource identifiers, that is the ability, on a computer network, to locate a particular resource (computer, document or other resource) on the network through a unique identifier called IP address;

The client-server model, of computing, in which client software or a client computer makes requests of server software or a server computer that provides the client with resources or services, such as data or files;

Markup language, in which characters or codes embedded in text indicate to a computer how to print or display the text, e.g. as in italics or bold type or font.

- The rule making body of the web is called W3C (world- wide web consortium)
- W3C is the non- profit organizations which makes the web standard.
 - The most know web standards are: HTML, CSS, XML, and XHTML.

Origin of the WWW

- The World Wide Web and its associated Hyper Text Transfer Protocol (HTTP) grew out of work done at the European laboratory for particle physics (CERN) in 1990
- Tim Berners-Lee developed HTTP as a network protocol for distributing documents and wrote the first web browser.
- HTTP is simple request/response protocol in which web browser asks for a document and the web server return the document in the form of HTML data stream preceded by a few descriptive headers.

Static & Dynamic web page

Static (fixed) web page

- Is a web page that is written entirely using HTML.
- Each web page is a separate document and there are no databases or external files that are drawn upon.
- The only way to edit this type of website is to go into each page and edit the HTML.
- Difficult to manage and handle the web page
- Hard for Mobility and Modification, also difficult to host up-to-date information
- Visitors to a static Web page all see the same content (unchanged).
- Prebuild document and not sufficient

Dynamic (changing) web page

- Is written using more complex code — such as PHP or ASP — and has a greater degree of functionality.
- Potentially able to make updates without needing any change of HTML document.
- Each page of a dynamic website is generated from information stored in a database or external file.
- Visitors to a static Web page see the changed content.
- Visitors can customize some or all of the viewed content such as weather for a region, or ticket availability for flights.

Why build web pages dynamically?

Build web page dynamically

- The web page is based on data sent by the client.

User can submit two kinds of data Explicit (i.e. HTML form of data) and implicit (i.e. HTTP request header) either kind of input can be used to build the output page.

- The web is derived from data that change frequently.

The web page use information from corporate databases or other server side sources.

Drawbacks of Dynamic Site

- cost more to develop, because they require more complex coding
- You will need to obtain web hosting which supports databases and dynamic languages.
- Needs a skill person

Client-server Architecture

A network architecture in which each computer or process on the network is either a client or a server.

Server (“back end”):

- A database from which a client requests information
- Fulfills a request for information by managing the request or serving the requested information to the client
- Responsible for data storage and management

Client (“front end”):

- Presents an interface to the user
- Gathers information from the user, submits it to a server, then receives, formats, and presents the results returned from the server.

Tip

“Frontend” refers to any aspect of the design process that appears in or relates directly to the browser.

“Backend” refers to the programs and scripts that work on the server behind the scenes to make web pages dynamic and interactive.

Types of Client-server Architecture:-

1. **Two-tire:-**it consist a client and a server.

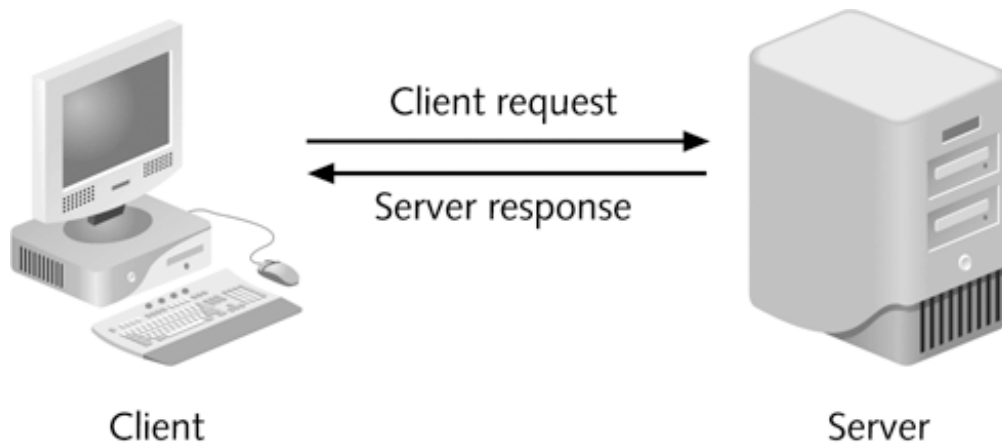


Fig 1:- two-tier client server architecture

2. A **three-tier**:-client/server architecture consists of three distinct pieces:

Client tier, or user interface tier, is the Web browser

Processing tier, or middle tier, handles the interaction between the Web browser clients

And the **data storage tier**

- Performs necessary processing or calculations based on the request from the client tier
- Handles the return of any information to the client tier

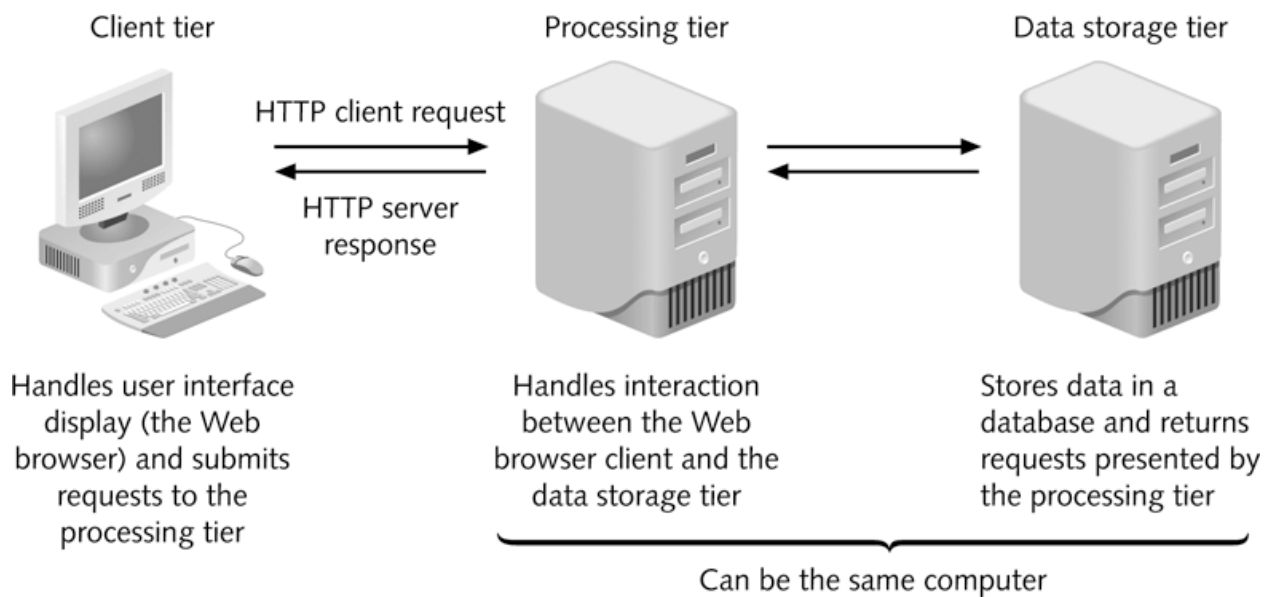


Fig2:- The design of a three-tier client/server architecture

How the Web Work?

The stream of event that occurs with every web page that appears on screen.

1. You request a web page by either typing its URL(for e.g `http://www.wsu.edu.et`) directly in the browser or by clicking on a link on the page. The URL contains all the information needed to target a specific document on a specific web server on the internet.
2. your browser send an HTTP request on the server named in the URL and asks for a specific file .if the URL specifies a directory(not a file).it is the same as request the default file in that directory.
3. The server looks for the request file and sends an either of the following HTTP response:
 - a) If the page cannot be found. The server return an error message .the message typical says “404” not found.
 - b) If the document is found. The server retrieves the requested file and returns it to the browser.
4. The browser parses the HTML document. If the page contains images(indicated by the HTML `img` tag), the browser contacts the server again to request each image files specified in the markup
5. The browser inserts each image in the document flow where indicated by the `img`. Then, the assembled web page is displayed for your viewing pleasure.

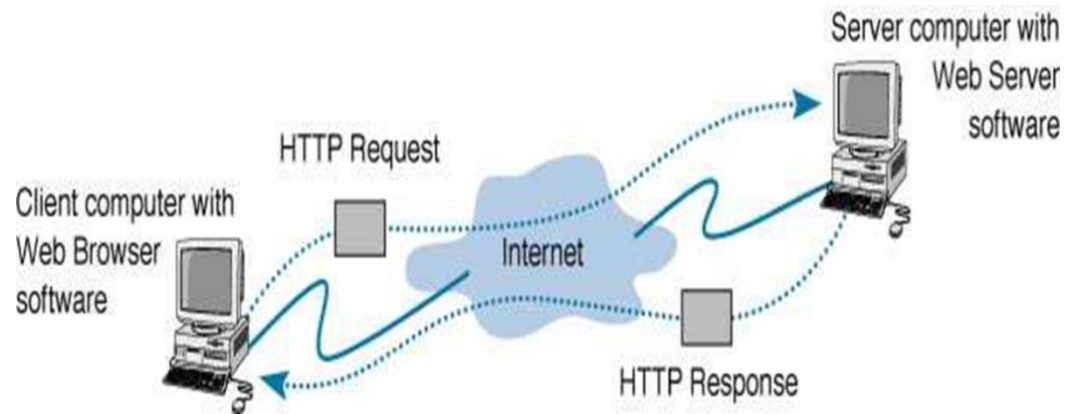


Fig 3:-how the web work

Browser and Web Server

- **Browser**

- It is a program that retrieves information from the Web.
- Its purpose is to retrieve and display information from a Web server by using HTTP protocol.
- It allows any user to access a server easily

There are different web browsers that are available and in use today and they all come with a variety of features. Some of these are:-

- **NCSA Mosaic** :- The first real HTML browser (1993)
- **Microsoft Internet Explorer** :- Most commonly found browsers
- **Opera**:-The fastest browser on Earth
- **Lynx**:-Text based web client
- **Mozilla Firefox Netscape, Safari, Google Chrome** etc.

Web Server

- It is a program that provides documents to clients.

E.g.

- **CERN httpd** (the first webserver)
- **Apache**(most widely used Web server software)
- **IIS**(internet information server)
- **Netscape Web server** ...

Uniform Resource Locators (URLs)

- A URL is the exact address or location of the electronic files/web page to be retrieved on the internet.
- It consists of a protocol, domain name, and sometimes the path to a specific Web page or location on a Web page.
 - Example:

<http://www.Yahoo.com/new/mypage/about.html>

- http:represents the protocol

- WWW.Yahoo.com:- represents the domain name
- New/mypage:- represents the path of the web page
- about.html :- represent the web page name/file name.

Default file

- Many address do not include file(web page) name, but simply point to a directory like this `http://www.Yahoo.com/ new/mypage`
- When a server receive a request for a directory name rather than a specific file, it looks for a default document/file usually named as `index.html` and send it back as a response for display . So when some one types the above URL in their browser what they will actually see is this `http://www.Yahoo.com/ new/mypage/index.html`
- The name of the default file(index file) may vary . And depend on how the server is configured.
 - E.g:
 - Index.html
 - Home.html
 - Default.html ...

Domain Name System (DNS)

- Domain name is a human readable name given to a server for ease of reference by human, i.e. It is an IP address in the textual form rather than numbers.
- The Domain Name System is a system by which it converts human-readable name (domain name) such as “www.wsu.edu.et” to their corresponding IP addresses, 191.57.126.24.
- Without DNS, the Internet would be much less usable.
- Their conversions are done by software systems called domain name servers
- Domain name server is simply a computer that contains a database of domain names and their corresponding IP addresses.

DNS name

- Is constructed hierarchically
- The highest level of the hierarchy being the last component or label of the DNS address
- Label can be up to 63 characters long and are case insensitive

- A maximum length of 255 characters is allowed
- Labels must start with a letter and can only consists of letters, digits and hyphens.
- The final most significant label of a fully qualified name can be:-
 - **Three letter code** :- indicate the type of organization hosting the computer

Code	Meaning
com	Commercial. Now international
edu	Educational
gov	governmental
Int	International origination
mil	military
net	Network related
org	Miscellaneous organization

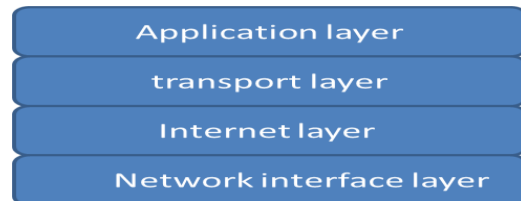
- **Two letter codes**:-indicate the country of origin
et for Ethiopia
kr for KOREA

TCP/ IP Protocols

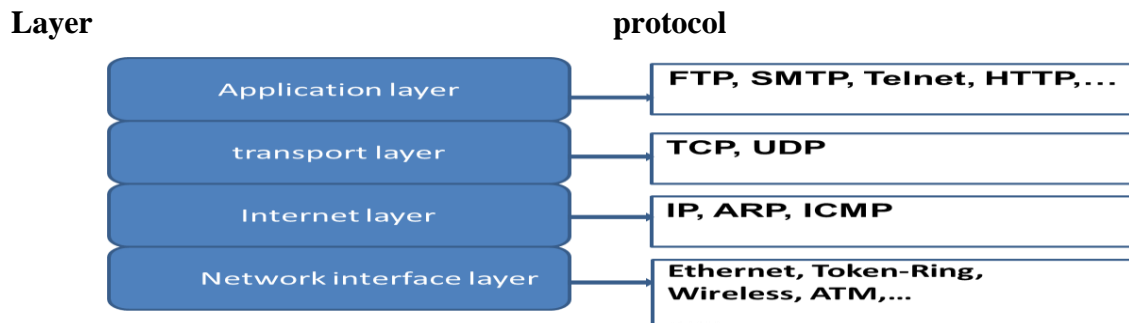
General description of the TCP/IP protocols

- **Protocols** are sets of rules for message formats and procedures that allow machines and application programs to exchange information.
 - *Why we need to build Protocols?*
- The TCP/IP protocol suite is so named for two of its most important protocols:- Transmission Control Protocol (TCP) and Internet Protocol (IP).
- **TCP/IP**
 - is a set of protocols developed to allow cooperating computers to share resources across a network
 - A highly standardized protocol used widely on the Internet
 - The main design goal of TCP/IP was to build an interconnection of networks, referred to as an inter-network, or internet, that provided universal communication services over heterogeneous physical networks.
 - The set of TCP/IP protocols are partitioned into 4-layers:

- Network interface layer
- Internet layer
- Transport layer
- Application layer



- TCP/IP Suite of Protocols



TCP/IP Protocol Suite and services

- **Application layer protocols**
 - provide accurate and efficient data delivery
- **Typical protocols in application layer:**
 - **FTP** – File Transfer Protocol
 - For file transfer
 - **Telnet** – Remote terminal protocol
 - Provides remote login service. It allows a user on one machine to log into another machine on the network.
 - **SMTP** – Simple Mail Transfer Protocol
 - For mail transfer
 - **HTTP** – Hypertext Transfer Protocol
 - For Web browsing
 - **LPD**-Line Printer Daemon

- designed for printer sharing
- **Transport Layer protocols define the rules of**
 - Dividing a large piece of data into segments
 - Reassemble segments into the original piece
- **Typical protocols in transport layer:**
 - TCP – Transmission Control Protocol
 - Provide further functions such as reordering and data resend
 - UDP- User datagram protocol
 - does *not* provide functions such as reordering and data resend

TCP-Transmission Control Protocol

- Takes large blocks of information from an application and breaks them into segments. It numbers and sequences each segment so that the destination's TCP protocol can put the segments back into the order the application intended. After these segments are sent, TCP (on the transmitting host) waits for an acknowledgment of the receiving end's TCP virtual circuit session, retransmitting those that aren't acknowledged.
- Before a transmitting, host starts to send segments down the model, the sender's TCP protocol contacts the destination's TCP protocol to establish a connection. What is created is known as a *virtual circuit*. *This type of communication is called connection-oriented.*
- **UDP – User Datagram Service**
 - UDP does *not sequence the segments and does not care in which order the segments arrive* at the destination.
 - But after that, UDP sends the segments off and forgets about them. It doesn't follow through, check up on them, or even allow for an acknowledgment of safe arrival—complete abandonment. Because of this, it's referred to as an unreliable protocol. This does not mean that UDP is ineffective, only that it doesn't handle issues of reliability.

- Further, UDP doesn't create a virtual circuit, nor does it contact the destination before delivering information to it. Because of this, it's also considered a connectionless protocol.
- TCP for reliability and UDP for faster transfers.

Key Concepts of Transport Layer Protocols

TCP	UDP
Sequenced	Unsequenced
Reliable	Unreliable
Connection-oriented	Connectionless
Virtual circuit	Low overhead

- **Internet layer protocols**
 - define the rules of how to find the routes for a packet to the destination
 - It only gives best effort delivery. Packets can be delayed, corrupted, lost, duplicated, out-of-order
- Typical protocols in internet layer:
 - **IP** – Internet Protocol
 - Provide packet delivery
 - **ARP** – Address Resolution Protocol
 - Define the procedures of network address / MAC address translation
 - **ICMP** – Internet Control Message Protocol
 - Define the procedures of error message transfer

- **Network Interface layer:-**

- Formats IP datagrams at the Network layer into packets that specific network technologies can understand and transmit.
- Responsible for sending and receiving TCP/IP packets on the network medium (physical/Data Link)
- Applicable LAN technologies
 - Ethernet, Token Ring, FDDI etc.
- Applicable WAN technologies
 - X.25 (old), Frame Relay, ATM etc.
- Note that some technologies such as ATM and FDDI may be used at both the WAN and the LAN levels

Hypertext Transfer Protocol (HTTP)

- Hyper Text Markup Language (HTML):- is the language used to describe the inside of web document.
- Hypertext Transfer Protocol(HTTP):-
 - Is the protocol used described how this documents are sent over the internet.
 - Is designed to enable communications between clients and servers.
 - Is a protocol for requesting Document over the network.
- Prescribes the rules by which browser (a web client application) make request and servers supply response.
- HTTP consists of a set of command written as lines of ordinary ASCII text. when you have a web browser, you don't enter HTTP command directly .instead when you type a URL or click hyperlink the browser translate your action in to HTTP command that request the document from the server specified in the URL the web server finds the web document and send it back to the browser, where it's send back to the browser, where it's displayed, along with its associated graphics and other hyperlinks.
- HTTP protocol include ways to:-
 - Ask for a document by name

- Determine who the user is
- Decide how to handle outdate resources
- Indicate the results of a request
- And other use full function

HTTP Methods

- Two commonly used methods for a request-response between a client and server are: GET and POST.
 - **GET** - Requests data from a specified resource.
 - **POST** - Submits data to be processed to a specified resource.
- **The GET Method**
 - Note that query strings (name/value pairs) is sent in the URL of a GET request:
`/test/demo_form.asp?name1=value1&name2=value2`
- **GET requests**
 - remain in the browser history
 - can be bookmarked
 - should never be used when dealing with sensitive data
 - have length restrictions
 - should be used only to retrieve data
- **The POST Method**
 - **Note that query strings (name/value pairs) is sent in the HTTP message body of a POST request:**
 - POST /test/demo_form.asp HTTP/1.1
Host: w3schools.com
name1=value1&name2=value2
- **POST requests:**
 - are never cached
 - do not remain in the browser history
 - cannot be bookmarked
 - have no restrictions on data length

Other HTTP Request Methods

- **PUT**:-a request for the server to store the data in the request and the new contents on the specified URI.
- **DELETE**: - a request for the server to delete the resource named in the URI.
- **OPTIONS**:-a request for the information about the request methods the server support.
- **TRACK**:- a request for the web server to echo the HTTP request and its headers.
- **CONNECT**:-Converts the request connection to a transparent TCP/IP channel