UNIVERSITY of RWANDA

COLLEGE OF SCIENCE AND TECHNOLOGY

SCHOOL OF ICT

Module Title: WEB TECHNOLOGY

Module Code: ITE2163

- Second Year : IS
- By: Mr Frodouard MINANI

Telephone: 0782149182/0739505574 **Email:** frodominani@gmail.com

Indicative Content

UNIT 1:

- **WWW Basics:** History of the Web, Growth of Web in the past decade, Protocols governing the web, Internet Browser, Resource Locators, Search engines, E-mail Services, Internet Software.

UNIT 2:

- **Programming Concepts:** Introduction to HTML, Table Handling in HTML, Creating Forms, and Techniques involved in building a local web page. Client side programming using Java script. Introduction to Client server programming. Server Side Programming concepts.

UNIT 3:

- Introduction to the semantic web
- ***XML***
  - Syntax • DOM
- ***UNIT 4:***
- Semantic web applications

•SW applications •Architectures

- ***Web services***

•Introduction •Conversational vs Functional approaches

•Protocols

Semantic Web Services

•Motivation behind Semantic Web Services

Introduction to the Internet and Web

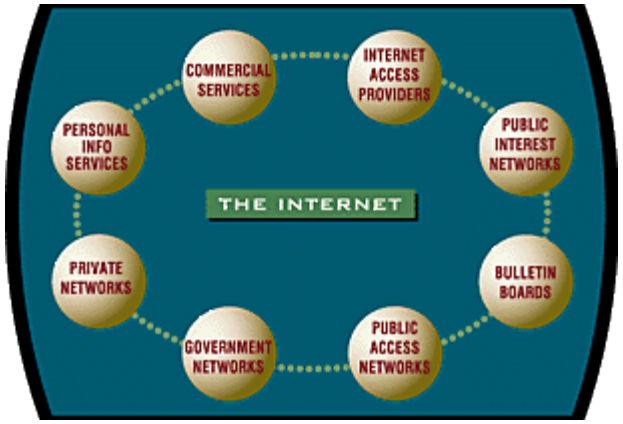
Introduction to the Internet

- The Internet has been called the "**network of networks.**"
- It's a way to link millions of people around the world, each sing a computer connected to a smaller network:

  -the one in their office,

  - to school, or

  - to a commercial network



Structure of the interconnected things (IoT)

➢*The **Internet** has quickly become a global communications system, offering new ways to reach other people and vast new sources of information.*

➢Today, the phone lines and computers are merely the tools of the networked world that make new forms of human communication possible.

➢These new ways to communicate are supported by a variety of simple tools for using the Internet.

➢One aspect of the Internet, **the World Wide Web**, is becoming one of the most popular forms of this new communications medium.

Who is the Owner of the Internet?

- Nobody owns or controls the Internet or the World Wide Web.
- Although millions of individuals and organizations control their own piece of it.

- The Internet does have some governing bodies that propose standards and specifications and help plan for the future.

- Individuals or organizations who just want to browse the Internet, send e-mail or use other basic facilities will need only a computer, modem, communications software and a simple way of accessing the Internet.

What is the Internet?

- The **Internet** is the GSI (Global System of Interconnected) network of computers, which is connected with different types of computers whole over the world.

- In the word, "**INTERNET**" was coming from "Inter-Networking," a mechanism of the Internet to transfer data from a computer to the other computers.

- The **Internet** means an Interconnection of Networks, which are communication protocols for communications between two computers over the network.

- The **Internet** is a huge capability of stored data such as music, video, software and any types of documents.

History of the Internet

- The **Internet** had its roots during the 1960's as a **project of the United States government's Department of Defense**, to create a **non-centralized network**.
- This project was called ARPANET (**Advanced Research Projects Agency Network),** created by the Pentagon's Advanced Research Projects Agency established in 1969 to provide a secure and survivable communications network for organizations engaged in defense-related research.

History of the Internet Cont'

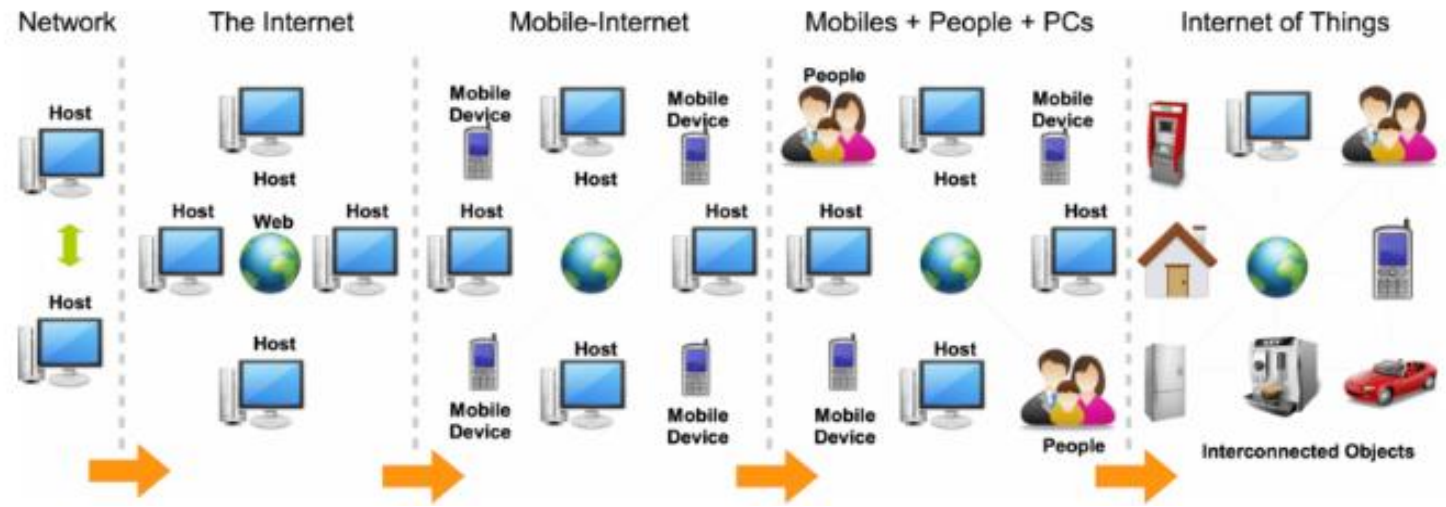
- **ARPANET** was the first WAN which is only **four sites** in 1969.
- The **Internet** evolved from the basic ideas of the **ARPANET** for interconnecting computers for sharing and exchanging information.
- In 1989, the U.S. Government lifted controls on the use of the Internet, and allowed its use for commercial purposes as well.
- Since then, the Internet has grown rapidly to become the world's largest network and now it now interconnects more **than 50,000 networks**, allowing more than 15 millions computers, more than 100 million computer users and more than 155 countries around the world easily communicate with each other.
- The Internet network continues to grow at the rapid pace.

History of the Internet Cont'

- In order to make the network more global a **new sophisticated and standard protocol** was needed.
- They developed **IP (Internet Protocol) technology** which defined how **electronic messages were packaged, addressed, and sent over the network.**
- The **standard protocol** was invented in 1977 and **was called TCP/IP (Transmission Control Protocol/Internet Protocol).**
- TCP/IP allowed users to link various branches of other complex networks directly to the ARPANET, which soon came to be called the Internet.

ARPANET computer

ARPANET Computer



Internet Evolution

Use of the INTERNET

Communication

Research

Education

Financial Transaction

Online Booking

Job Search

Blogging

Shopping

Who Pays for the Internet?

- Because the Internet is not one "thing", it's many things.
- No one central agency exists that charges individual Internet users.
- Rather, individuals and institutions who use the Internet pay a local or regional **Internet service provider** for their share of services.
- And in turn, those smaller Internet service providers might purchase services from an even larger network.
- So basically, everyone who uses the Internet in some way pays for part of it.

What makes the internet work?

- Internet allows many different computers to connect and talk to each other.
- This is possible because of a set of standards, known as **protocols** that govern the transmission of data over the network: **TCP/IP** (Transmission Control Protocol/Internet Protocol).

History and Overview of the Web

Overview of WWW

➢Although many people use the terms World **Wide Web** and **Internet** interchangeably, the World Wide Web actually is a service of the Internet.

➢While the Internet was developed in the late 1960s, the World Wide Web emerged in the early 1990s.

➢The **World Wide Web (WWW),** or **Web**, consists of a worldwide collection of electronic documents and services, distributed across the Internet and linked together by hypertext links.

➢Each electronic document on the Web is called a **Web page**, which can contain text, graphics, animation, audio, and video.

➢Some Web pages are **static (fixed); others are dynamic (changing).**

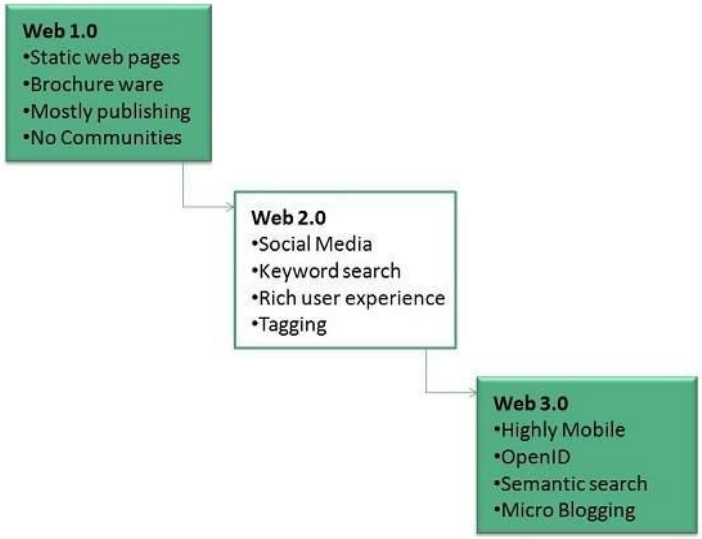
➢**A Web site** is a collection of related Web pages and associated items, such as documents and pictures, stored on a Web server.

➢**A Web server** is a computer that delivers requested Web pages to your computer.

➢The same Web server can store multiple Web sites.

Evolution of WWW

➢**World Wide Web** was created by **Timothy Berners Lee** in 1989 at **CERN** in **Geneva**



Evolution of WWW

Browsing the Web

➢**A Web browser,** or browser, is application software that allows users to access and view Web pages or access Web 2.0 programs.

➢To browse the Web, you need a computer or mobile device that is connected to the Internet and has a Web browser.

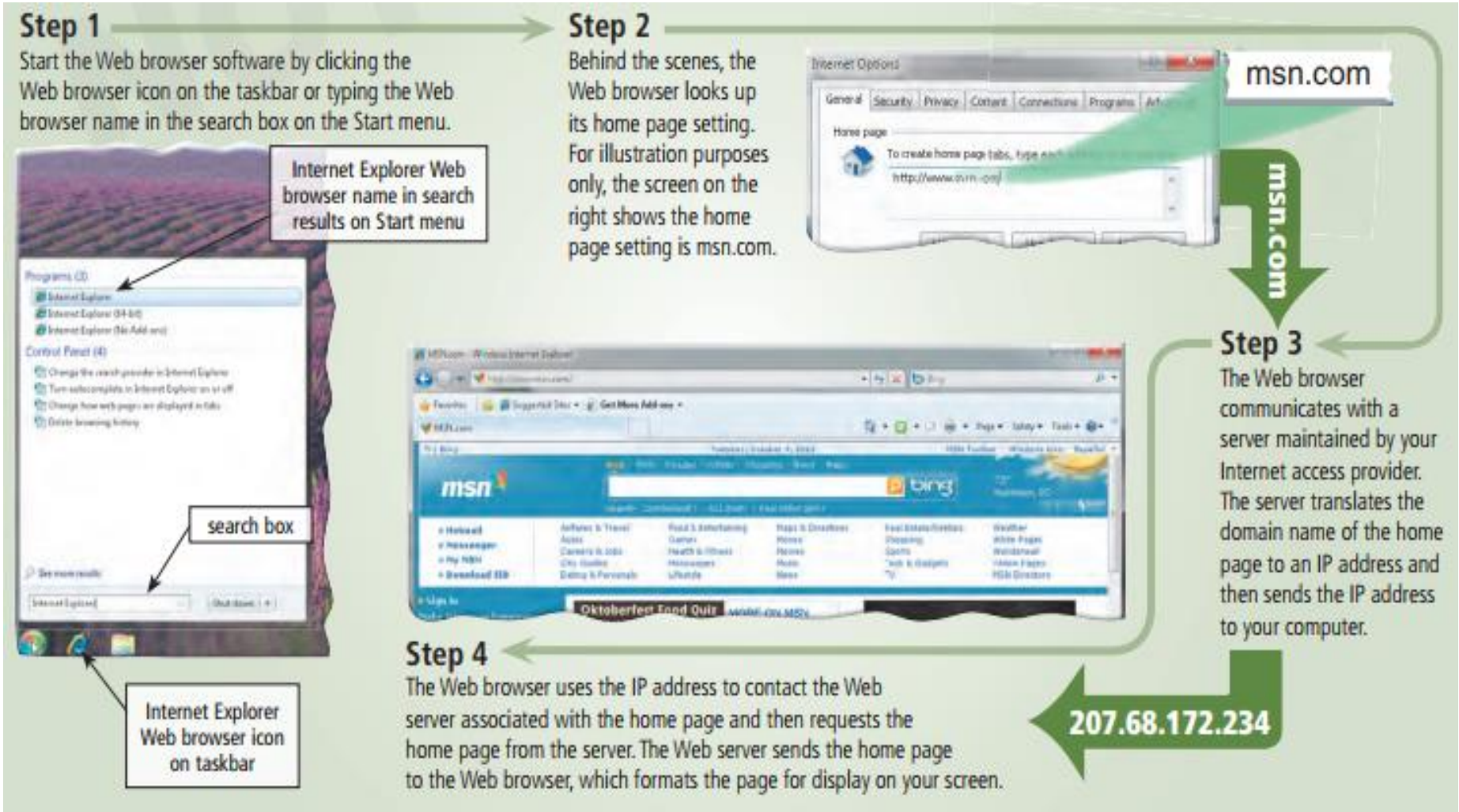
➢The more widely used Web browsers for personal computers are **Internet Explorer, Firefox, Opera, Safari, and Google Chrome.**

➢The browser retrieves and displays a starting Web page, sometimes called the **browser's home page**

✓For a computer or mobile device to display a Web page, the page must be downloaded**.**

✓**Downloading** is the process of a computer or device receiving information, such as a Web page, from a server on the Internet

How a Web Browser Displays a Home Page



How Web browser displays a home page.

Internet Addresses

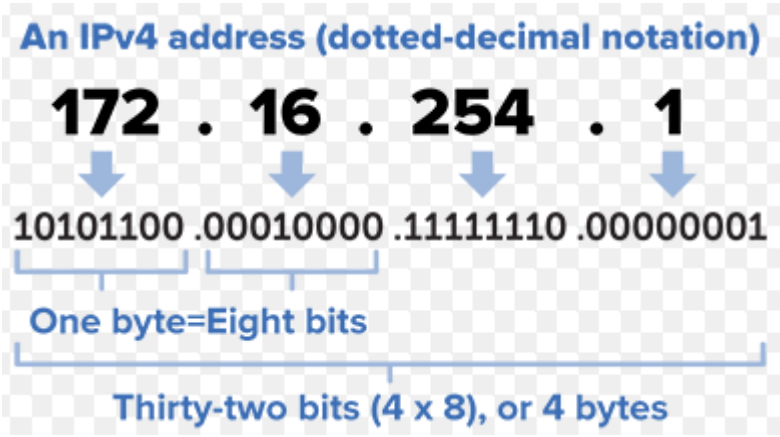
➢**An IP address**, short for Internet Protocol address, is a number that uniquely identifies each computer or device connected to the Internet.

➢The IP address usually consists of four groups of numbers, each separated by a period.

➢The number in each group is between **0** and **255.**

➢ For example, the numbers **72.14.207.9**9 are an IP address.

➢**A domain name** is the text version of an IP address



The IP address and domain name for the Google Web site

- The organization that assigns and controls top-level domains is the **Internet Corporation** for **Assigned Names and Numbers** (ICANN pronounced EYE-can)
- The **domain name system (DNS**) is the method that the Internet uses to store domain names and their corresponding IP addresses.
- When domain name is specified, a DNS server translates the domain name to its associated IP address so that data and information can be routed to the correct computer.
- **A DNS server** is an Internet server that usually is associated with an Internet access provider.

Examples of Generic Top-Level Domains

Generic TLD	Intended Purpose
aero	Aviation community members
biz	Businesses of all sizes
cat	Catalan cultural community
com	Commercial organizations, businesses, and companies
coop	Business cooperatives such as credit unions and rural electric co-ops
edu	Educational institutions
gov	Government agencies
info	Business organizations or individuals providing general information
jobs	Employment or human resource businesses
mil	Military organizations
mobi	Delivery and management of mobile Internet services
museum	Accredited museums
name	Individuals or families
net	Network providers or commercial companies
org	Nonprofit organizations
pro	Certified professionals such as doctors, lawyers, and accountants
tel	Internet communications
travel	Travel industry

Web Addresses

➢A Web page has a unique address, called a **URL (Uniform Resource Locator)** or **Web address**

URL has the following syntax:

protocol*://*hostname*:*port*/*path-and-file-name

- ***Protocol*:** The application-level protocol used by the client and server, e.g., HTTP, FTP, and telnet.
- ***Hostname*:** The DNS domain name (e.g., www.nowhere123.com) or IP address (e.g., 192.128.1.2) of the server.
- ***Port*:** The TCP port number that the server is listening for incoming requests from the clients.
- ***Path-and-file-name*:** The name and location of the requested resource, under the server document base directory.

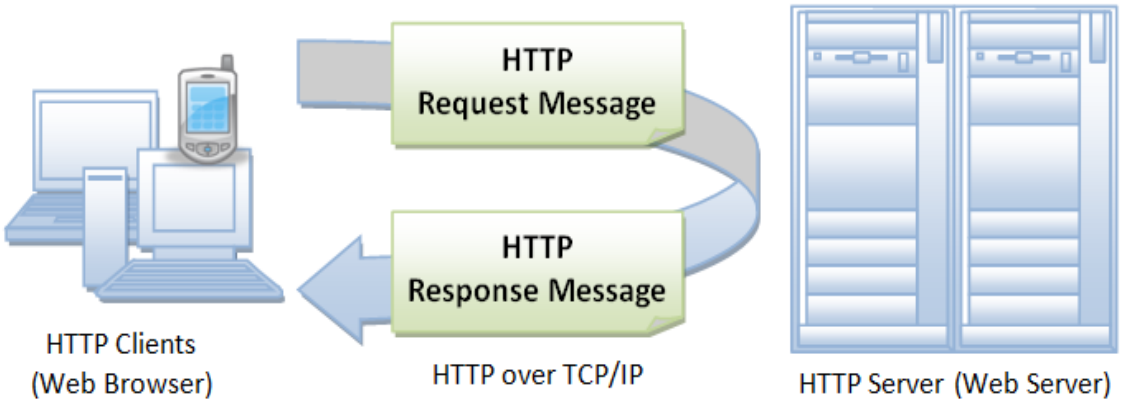
- In the following URL:
- **http://www.nowhere123.com/docs/index.html**, the communication protocol is **HTTP**; the hostname is **www.nowhere123.com**.
- The port number was not specified in the URL, and takes on the default number, which is TCP port **80** for HTTP.
- The path and file name for the resource to be located is **"/docs/index.html**".

Protocols governing the web

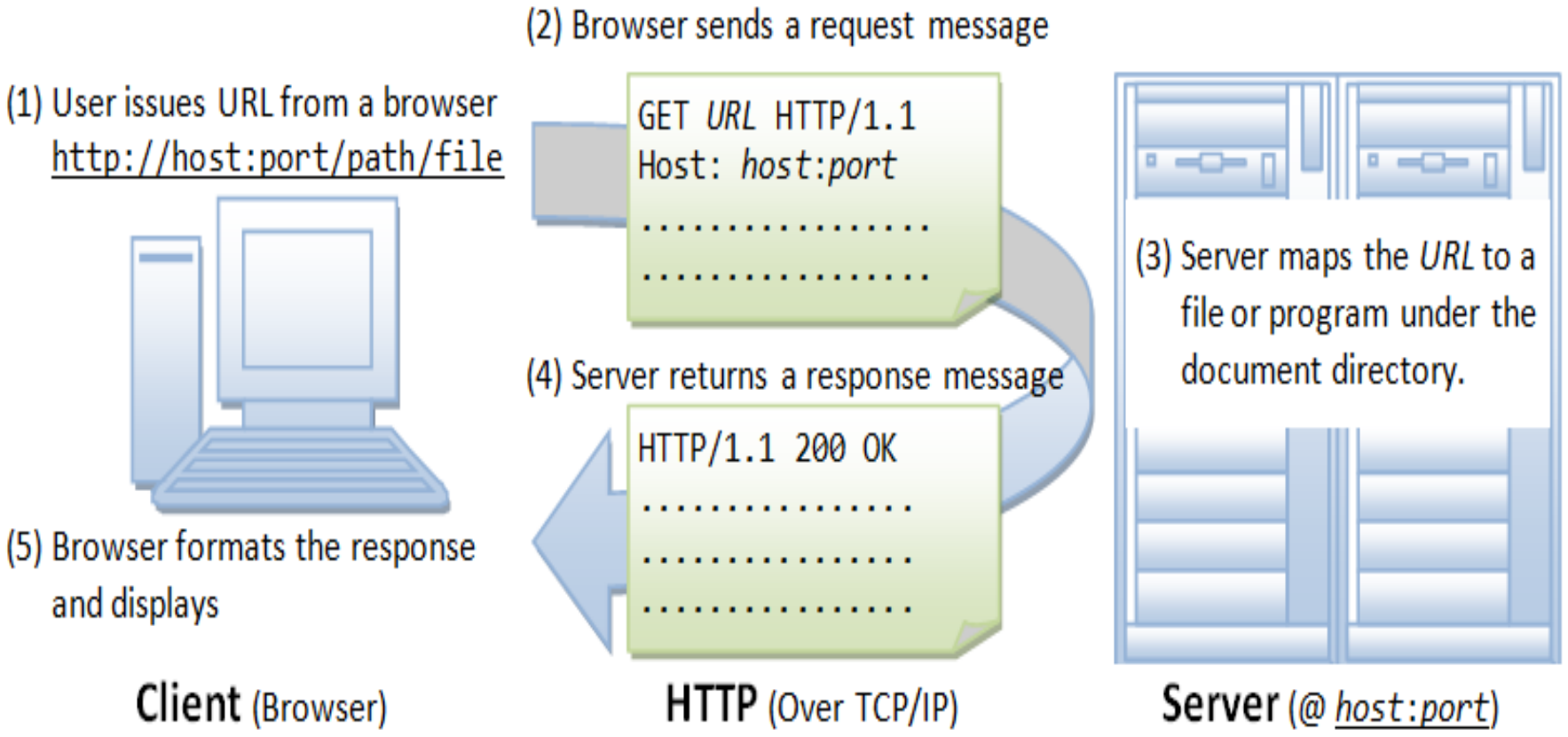
- A **Protocol** is a set of rules.
- **Protocols** allow two computers to communicate over media such as wireless or hardwired technologies.
- When computers communicate with each other, there needs to be a common set of rules and instruction that each computer follows.
- **Protocols** are the language of computers.

Some types of Protocols

Hyper Text Transfer Protocols (HTTP)

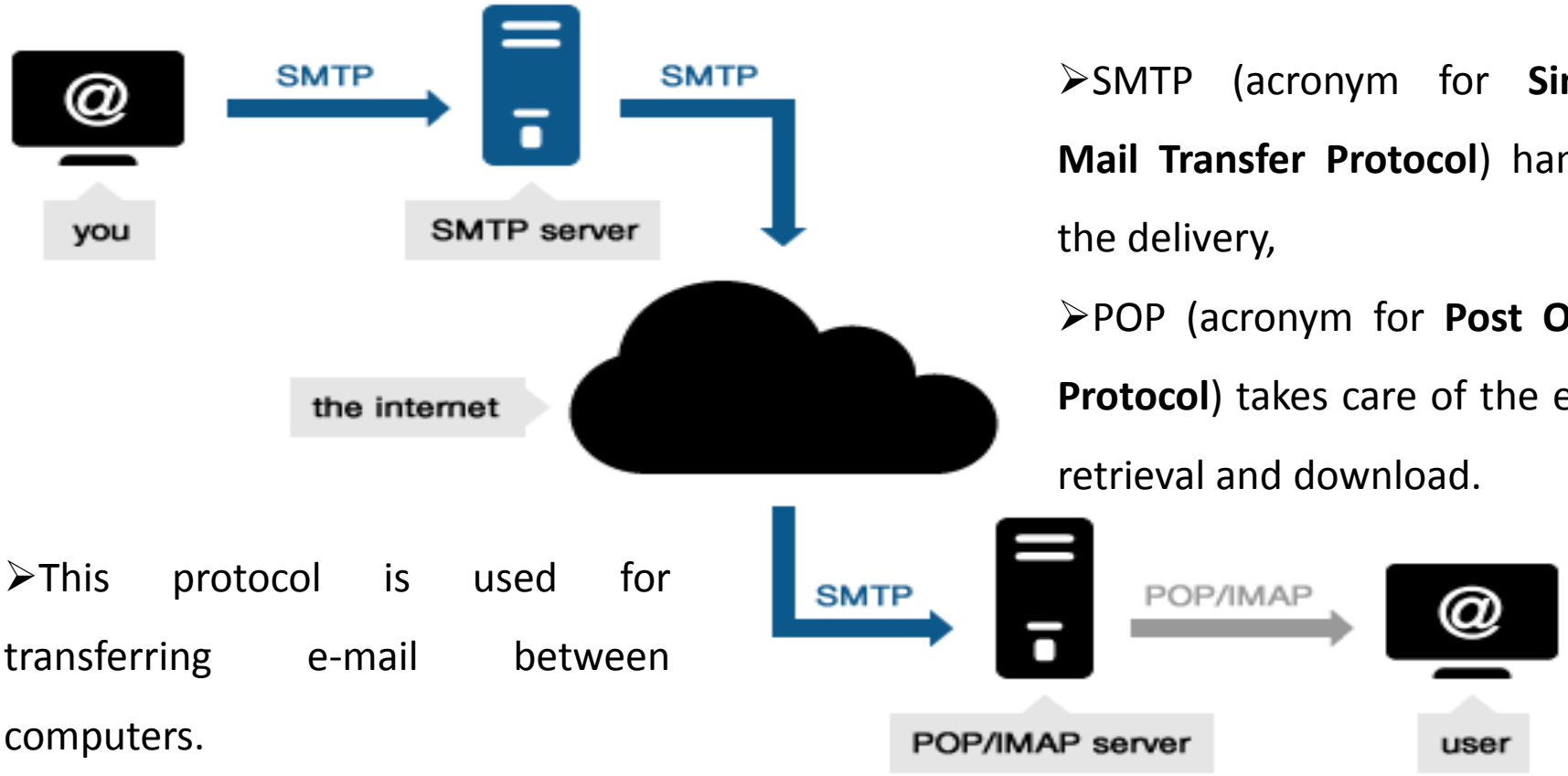


- HTTP is an *asymmetric request-response client-server* protocol as illustrated.
- An HTTP client sends a request message to an HTTP server.
- The server, in turn, returns a response message. In other words, HTTP is a ***pull protocol***, the client ***pulls*** information from the server (instead of server *pushes* information down to the client).



How HTPP works

Simple Mail Transfer Protocol (SMTP)



➢SMTP (acronym for **Simple Mail Transfer Protocol**) handles the delivery,

➢POP (acronym for **Post Office Protocol**) takes care of the email retrieval and download.

➢This protocol is used for transferring e-mail between computers.

➢**Simple Mail Transfer Protocol**, and defines the method that handles the process of **email exchange and delivery** across IPs.

IMAP: Internet Message Access Protocol which allows the client program to manipulate the e-mail message on the server without downloading them on the local computer.

File Transfer Protocol (FTP)

➢FTP is used to upload files on server and download files from server.

FTP: REQUEST

- This request is initiated by the client
- Allowing the File Transfer Protocol to begin
- Sub-commands
  - store, delete, lookup, open and close
  - Allows client to interact with the server before they download or upload a file

FTP: RESPONSE

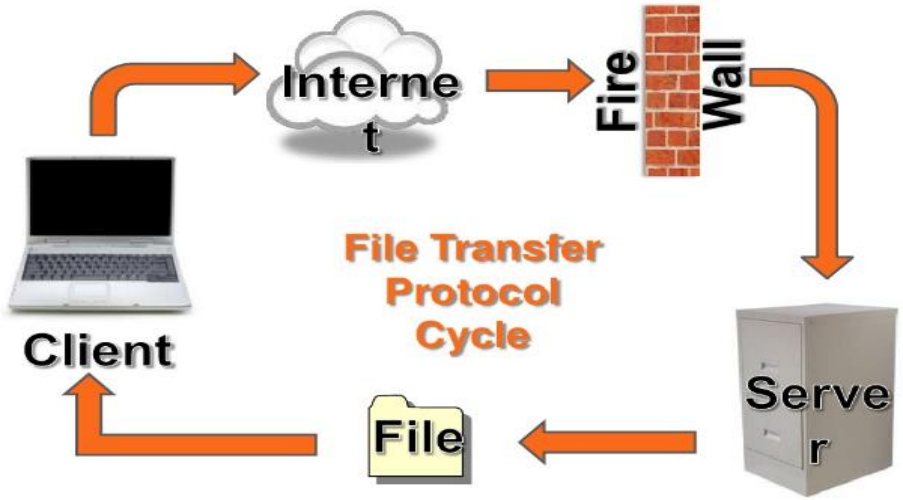
- Server will inform the client that it is ready to either receive or send a file
- The sub-commands used in this transaction are “ready to send” and “ready to receive”

FTP: TRANSFER

- The actual transaction will occur in the transfer step

FTP: TERMINATE

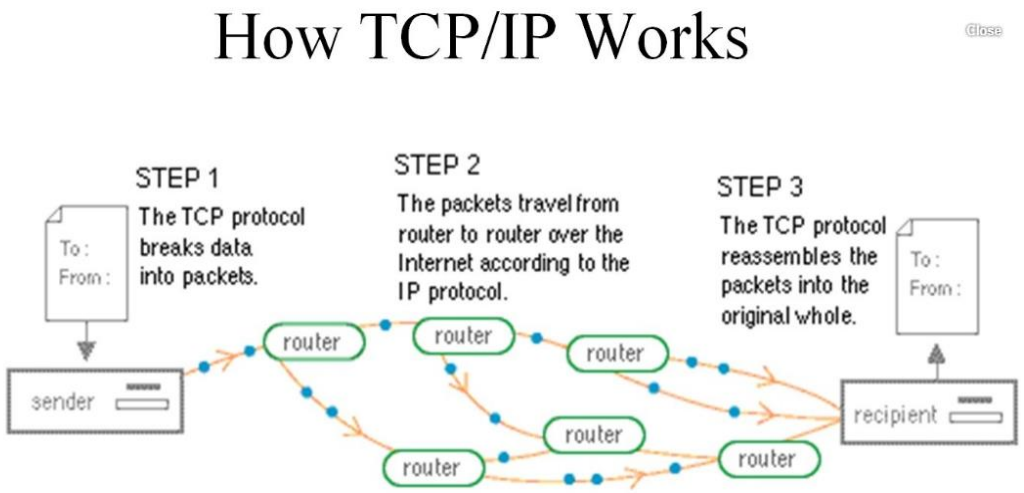
- Two types of terminate transactions
  - successful terminate
  - unsuccessful terminate



FTP Protocol

Transmission Control Protocol/Internet Protocol (TCP/IP)

- TCP/IP stands for **Transmission Control Protocol/Internet Protocol.**
- It is the communication language or protocol used to connect hosts on the internet.
- TCP/IP works on client-server communication model; here the computer user as a client sends requests by another computer or server in the network.
- In TCP/IP, each communications start from one point (host computer) in the network to another point (another host computer).



Search Engines

- A search engine is a program that finds Web sites, Web pages, images, videos, news, maps, and other information related to a specific topic.

- A **search engine** is helpful in locating information for which you do not know an exact Web address or are not seeking a particular Web site

- Search engines require that you enter a word or phrase, called **search text or search query.**

Widely Used Search Tools

Search Tool	Web Address	Search Engine	Subject Directory
A9	a9.com	X	
AlltheWeb	alltheweb.com	X	
AltaVista	altavista.com	X	
AOL Search	search.aol.com	X	
Ask	ask.com	X	
Bing	bing.com	X	
Cuil (pronounced cool)	cuil.com	X	
Dogpile	dogpile.com	X	
Excite	excite.com	X	X
Gigablast	gigablast.com	X	X
Google	google.com	X	X
Lycos	lycos.com	X	
MSN	msn.com	X	X
Open Directory Project	dmoz.org	X	X
WebCrawler	webcrawler.com	X	
Yahoo!	yahoo.com	X	X

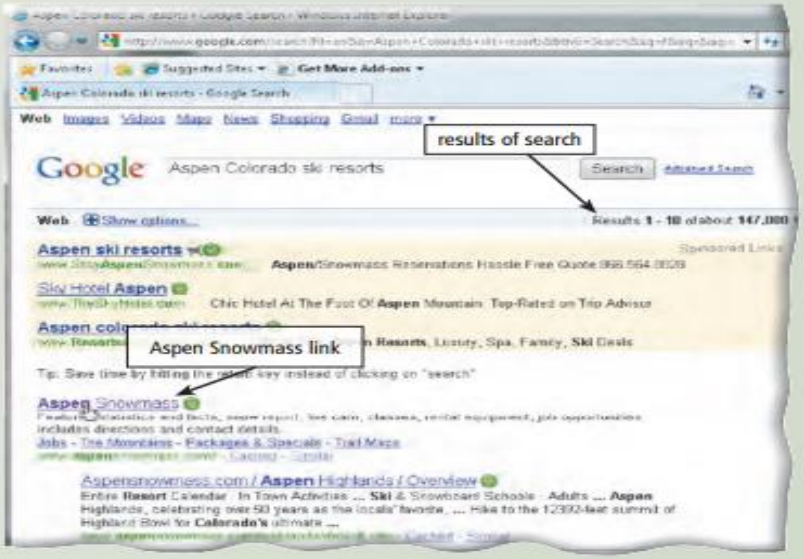
Popular search engines and subject directories.

How to Use a Search Engine

Step 1
Type the search engine's Web address (in this case, google.com) in the Address bar in the Web browser.

Step 2
Press the ENTER key. When the Google home page is displayed, type `Aspen Colorado ski resorts` as the search text and then point to the Google Search button.

Step 3
Click the Google Search button. When the results of the search are displayed, scroll through the links and read the descriptions. Point to the Aspen Snowmass link.



Step 4
Click the Aspen Snowmass link to display a Web page with a description and links to skiing in Aspen.

E-mail Services

- Electronic mail, or e-mail, is the most frequently used service on the Internet.
- E-mail Protocols are set of rules that help the client to properly transmit the information to or from the mail server.

Why use Email?

- You can send a message any time, anywhere.
- You can send the same message to several people at the same time.
- You can forward information to co-workers without retyping it.
- You can save time. E-mail is fast, usually taking no more than a few minutes to be received.
- You can e-mail electronic documents and the recipients can then edit and return revised versions.
- You can send messages around the world as easily as to co-workers in the next office.

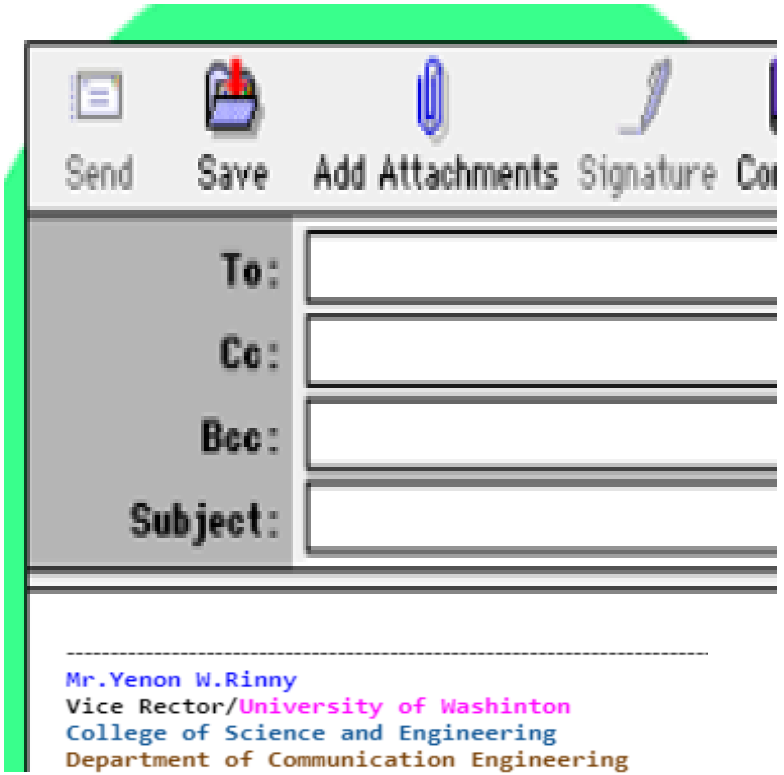
How Email works on the Internet?

- Like a postal address, an e-mail address specifies the destination of an electronic message.
- An Internet e-mail address looks like this: **user_name@domain_name**.
- The user name is a unique name that identifies the recipient.
- The domain name is the address. Many people can share the same domain name.

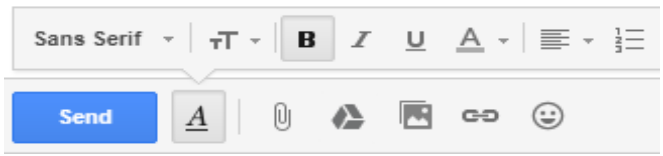
- E-mail is sent and received through electronic "post offices" known as mail servers.
- To read your e-mail, you must retrieve it from the mail server.
- Once you enter the address of the recipient, compose your message, and click Send, your e-mail software handles the delivery.

Anatomy of an Email message

- **To** contains the e-mail addresses of the **recipients.** This is a mandatory entry.
- **CC**, short for **Carbon Copy**.
- **BCC:** short for **Blind Carbon Copy** contains the e-mail addresses of other recipients who receive copies, but their names and addresses are hidden from the other recipients.
- **Subject** contains the **main topic** of the message.
- **Attachment** contains the **names of files** that you may be sending, for example, a word-processing document or a spreadsheet.



Signatures



QUIZ1 / 5

- Explain why HTTP is a pull protocol

- Give the difference between “www” and webpage.