History of the Internet and the Web

CSC 170

The Internet

- The beginnings of the Internet go back to 1969
- It is a worldwide network of interconnected computers and related equipment.
- Used as a "backbone" for other online applications, including email, instant messaging, Skype, streaming video, and the World Wide Web.
- The Web ≠ The Internet!

Early Internet Applications

- Email (1971) - a system of sending electronic mail to accounts on remote machines.
- Telnet (1972) - a tool to let you log in to remote computers
- FTP (1973) - File Transfer Protocol; moves files from one computer to another.
- Gopher (1991) - first menu-driven Internet application used to share documents.
  - Gopher was like the "web" except: no linking (hypertext)
- All of these were (and some still are) command-line, text-based interfaces.

The World Wide Web

The graphical user interface to information stored on computers connected to the Internet.

- A system of interlinked hypertext documents connected by the Internet.
- Proposed by Tim Berners-Lee in 1989, became publically available in late 1991.
- The web's popularity took off in 1993 with the release of the Mosaic web browser. Mosaic's claim to fame is that it was able to display both text *and* images.

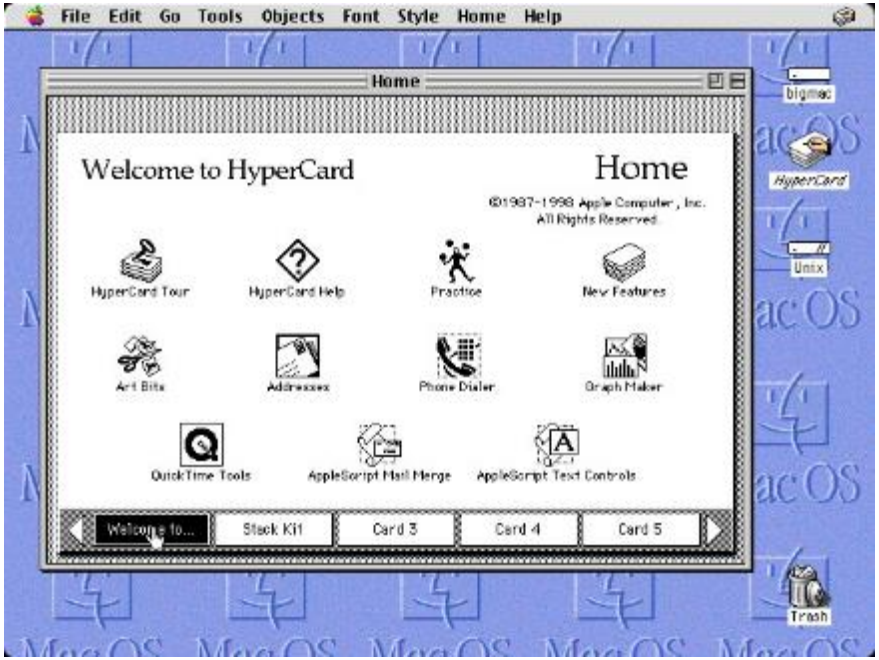
Tim Berners-Lee

- Proposed the concept of the World Wide Web in 1989; introduced it in 1991
- Rejected names included "Information Mesh", "Mine of Information", and "Information Mine"
- Originally intended for distributed document delivery and sharing among physicists. (Academic papers!!!)
- Text-only at first

Tim Berners-Lee:
Three existing technologies

1. Documents on the Internet (servers and connections)
   - Storing and sharing documents on the internet
   - Gopher, other document-sharing systems already existed
2. Markup languages (SGML-based)
   - IBM's GML (1960)
   - Scribe (1980)
   - SGML: Standardized Generalized Markup Language (1986)
3. Hypertext (Apple Hypercard)
   - "Hypermedia" (1965)
   - A hypermedia application for the Apple Macintosh (1987)

Apple HyperCard



1987: free with all new Macs

Withdrawn from sale in March 2004

HyperCard inspired:

- the web
- HTTP
- JavaScript
- ViolaWWW

Marc Andreesen

- Worked for the NCSA as a student; saw need for images on WWW
- In 1993, developed the first GUI browser, Mosaic, that displayed "inline" (embedded) images, along side text
- Timing was perfect because desktop PC Internet access was finally available via SLIP/PPP.
- Left NCSA in 1994 to found Netscape.
- Netscape later went out of business, but lives partly on through mozilla.org, the creator of the Firefox web browser.

Reasons for Internet Growth in the 1990s

- Removal of the ban on commercial activity
- Development of the World Wide Web by Tim Berners-Lee at CERN
- Development of Mosaic, the first graphics-based web browser at NCSA
- Personal computers were increasingly available and affordable
- Online service providers offered low-cost connections to the Internet

Web Standards and the W3C Consortium

- W3C – World Wide Web Consortium
  - Develops recommendations and prototype technologies related to the Web
  - Produces specifications, called Recommendations, in an effort to standardize web technologies

Internet Standards & Coordination

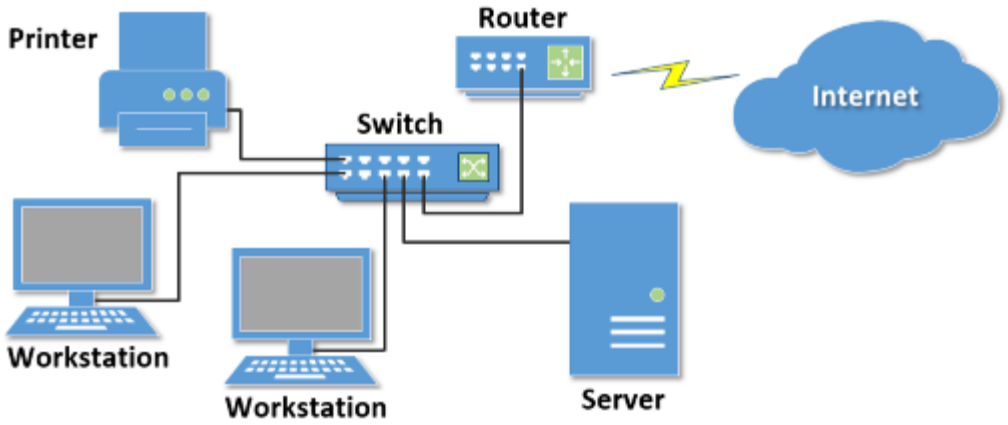
- ICANN - The Internet Corporation for Assigned Numbers & Names
  - Non-profit organization
  - Main function is to coordinate the assignment of:
    - Internet domain names
    - IP address numbers
    - Protocol parameters
    - Protocol port numbers.

Networks

The Internet *is* a Network

Network Overview

- **Network**
  two or more computers connected together for the purpose of communicating and sharing resources



Networks

- LAN – Local Area Network
  - Usually confined to a single building or group of buildings

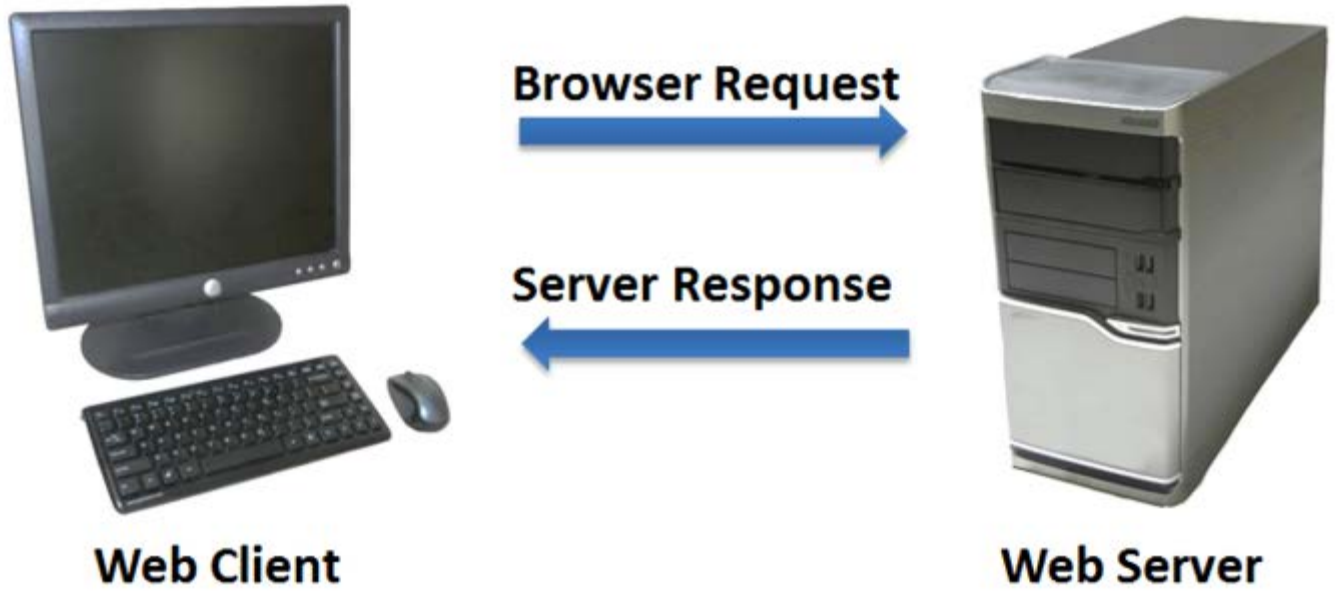
- WAN – Wide Area Network
  - Usually uses some form of public or commercial communications network to connect computers is widely dispersed geographical areas.

The Client/Server Model

- Client/Server can describe a relationship between two computer programs – the "**client**" and the "**server**".
- Client
  - requests some type of service (such as a file or database access) from the server.
- Server
  - fulfills the request and transmits the results to the client over a network

The Internet Client/Server Model

- Client – Web Browser
- Server – Web Server



Web Client

- Connected to the Internet when needed
- Usually runs web browser (client) software *(such as Google Chrome or Firefox)*
- Uses HTTP (Hypertext Transfer Protocol)
- Requests web pages from server
- Receives web pages and files from server

Web Server

- Continually connected to the Internet
- Runs web server software
  such as Apache or IIS (Microsoft's Internet Information Server)
- Uses HTTP (Hypertext Transfer Protocol)
- Receives request for the web page
- Responds to request and transmits status code, web page, and associated files

Internet Protocols

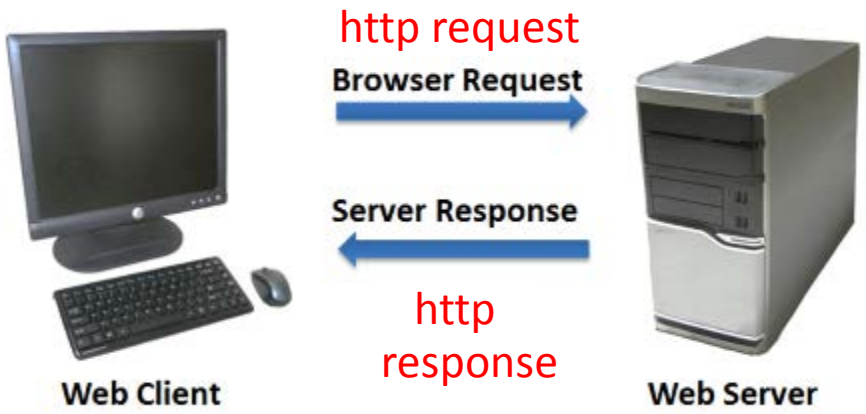
- Protocols
  - Rules that describe the methods used for clients and servers to communicate with each other over a network.
- There is no *single* protocol that makes the Internet and Web work.
- A number of protocols with specific functions are needed.

FTP
File Transfer Protocol

- A set of rules that allow files to be exchanged between computers on the Internet.
- Web developers commonly use FTP to transfer web page files from their computers to web servers.
- FTP is also used to download programs and files from other servers to individual computers.

HTTP - Hypertext Transfer Protocol

- A set of rules for exchanging files such as text, graphic images, sound, video, and other multimedia files on the Web.



- Web browsers send HTTP requests for web pages and their associated files.
- Web servers send HTTP responses back to the web browsers.

TCP/IP
Transmission Control Protocol/ Internet Protocol

- TCP/IP has been adopted as the official communication protocol of the Internet.

- TCP and IP have different functions that work together to ensure reliable communication over the Internet.
  - TCP = "the envelope"
  - IP = "the address on the envelope"

TCP
Transmission Control Protocol

- Purpose is to ensure the integrity of communication
- Breaks files and messages into individual units called packets

IP
Internet Protocol

- A set of rules that controls how data is sent between computers on the Internet.
- IP routes a packet to the correct destination address.
- The packet gets successively forwarded to the next closest router (a hardware device designed to move network traffic) until it reaches its destination.

IP Address

- Each device connected to the Internet has a unique numeric IP address.

- These addresses consist of a set of four groups of numbers, called octets.

  74.125.225.78 will get you Google!

- An IP address may correspond to a domain name.

The Domain Name System (DNS)

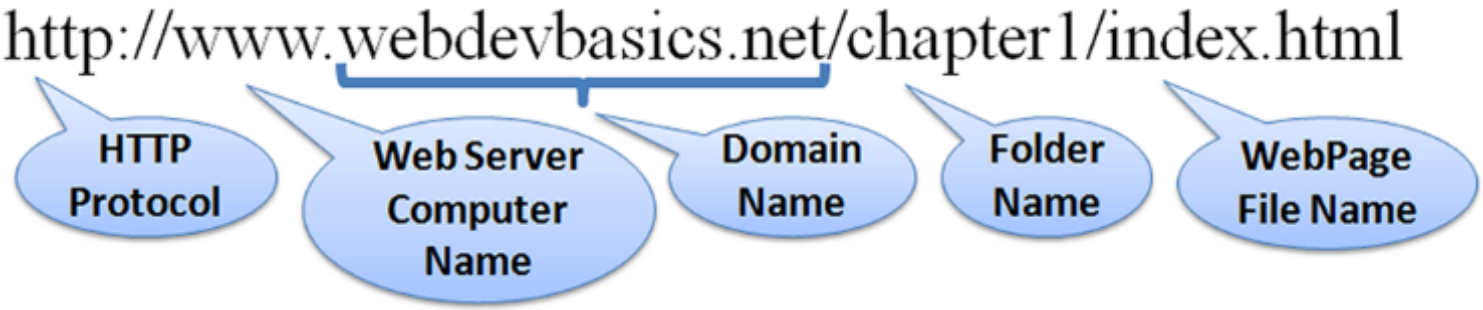
Domain Name

- Locates an organization or other entity on the Internet
- Domain Name System
  - Divides the Internet into logical groups and understandable names
  - Associates unique computer IP Addresses with the text-based domain names you type into a web browser
  - Browser: http://google.com
  - IP Address: 74.125.225.78

Uniform Resource Identifier

- **URI – Uniform Resource Identifier**
  - identifies a resource on the Internet

- **URL – Uniform Resource Locator**
  - a type of URI which represents the network location of a resource such as a web page, a graphic file, or an MP3 file.



TLD
Top-Level Domain Name

- A top-level domain (TLD) identifies the right-most part of the domain name.

- Current generic TLDs:
  .com, .org, .net, .mil, .gov, .edu, .int, .aero, .asia, .cat, .jobs, .name, .biz, .mobi, .museum, .info, .coop, .post, .pro, .tel, .travel, .xxx

County Code TLDs

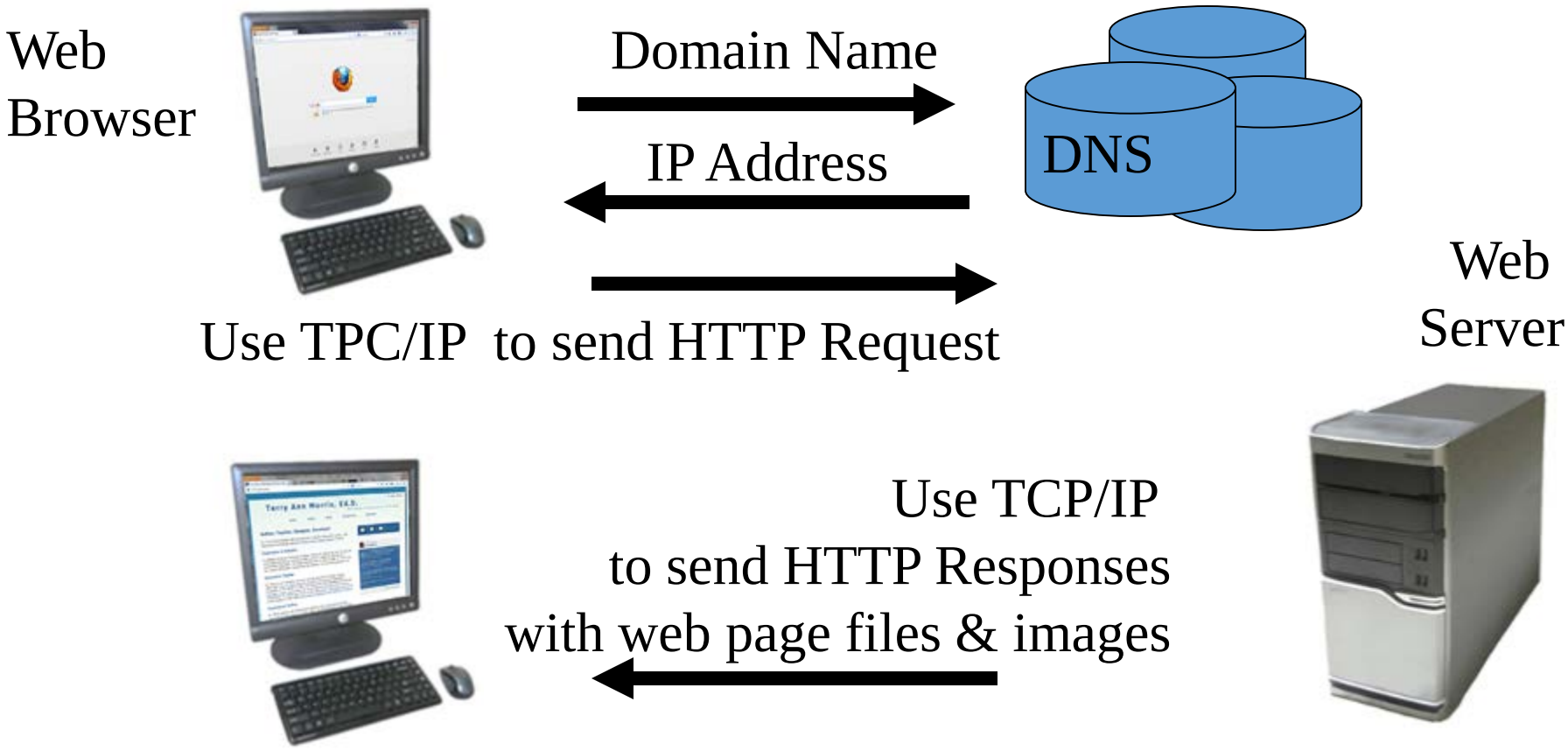
- Two character codes originally intended to indicate the geographical location (country) of the web site.

- In practice, it is fairly easy to obtain a domain name with a country code TLD that is not local to the registrant.

- Examples:
  - .tv, .ws, .au, .jp, .uk
  - See http://www.iana.org/cctld/cctld-whois.htm

Domain Name System

- The Domain Name System (DNS) associates Domain Names with IP addresses.



Markup Languages of the web

HTML

Markup Languages

- SGML – Standard Generalized Markup Language
  - A standard for specifying a markup language or tag set
- HTML – Hypertext Markup Language
  - The set of markup symbols or codes placed in a file intended for display on a web browser.

•XML – eXtensible Markup Language

- A text-based language designed to describe, deliver, and exchange structured information.

- It is not intended to replace HTML –
  it is intended to extend the power of HTML by separating data from presentation.

Markup Languages (3)

- XHTML – eXtensible Hypertext Markup Language
  - Developed by the W3C as the reformulation of HTML 4.0 as an application of XML.
  - It combines the formatting strengths of HTML 4.0 and the data structure and extensibility strengths of XML.

HTML 5

- The next version of HTML4 and XHTML
- http://www.w3.org/html/