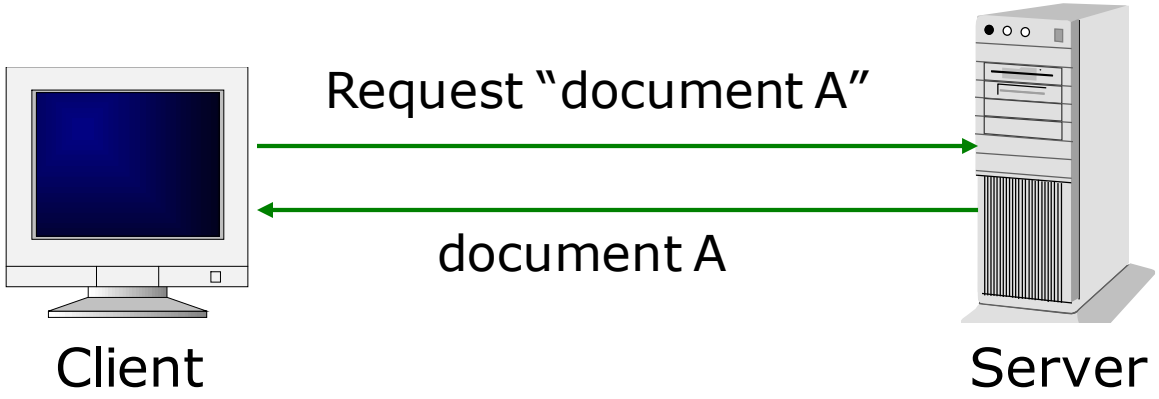
Web Essentials

Clients, Servers, and Communication

Web Essentials

- **Client**: web browsers, used to surf the Web
- **Server** systems: used to supply information to these browsers
- Computer **networks**: used to support the browser-server communication



Internet Vs Web

- **The Internet**: a inter-connected computer networks, linked by wires, cables, wireless connections, etc.
- **Web**: a collection of interconnected documents and other resources.
- The world wide web (**WWW**) is accessible via the Internet, as are many other services including email, file sharing, etc.

How does the Internet Work?

- Through communication protocols
- A **communication protocol** is a specification of how communication between two computers will be carried out
  - **IP** (Internet Protocol): defines the packets that carry blocks of data from one node to another
  - **TCP** (Transmission Control Protocol) and **UDP** (User Datagram Protocol): the protocols by which one host sends data to another.
  - Other application protocols: **DNS** (Domain Name Service), **SMTP** (Simple Mail Transmission Protocol), and **FTP** (File Transmission Protocol)

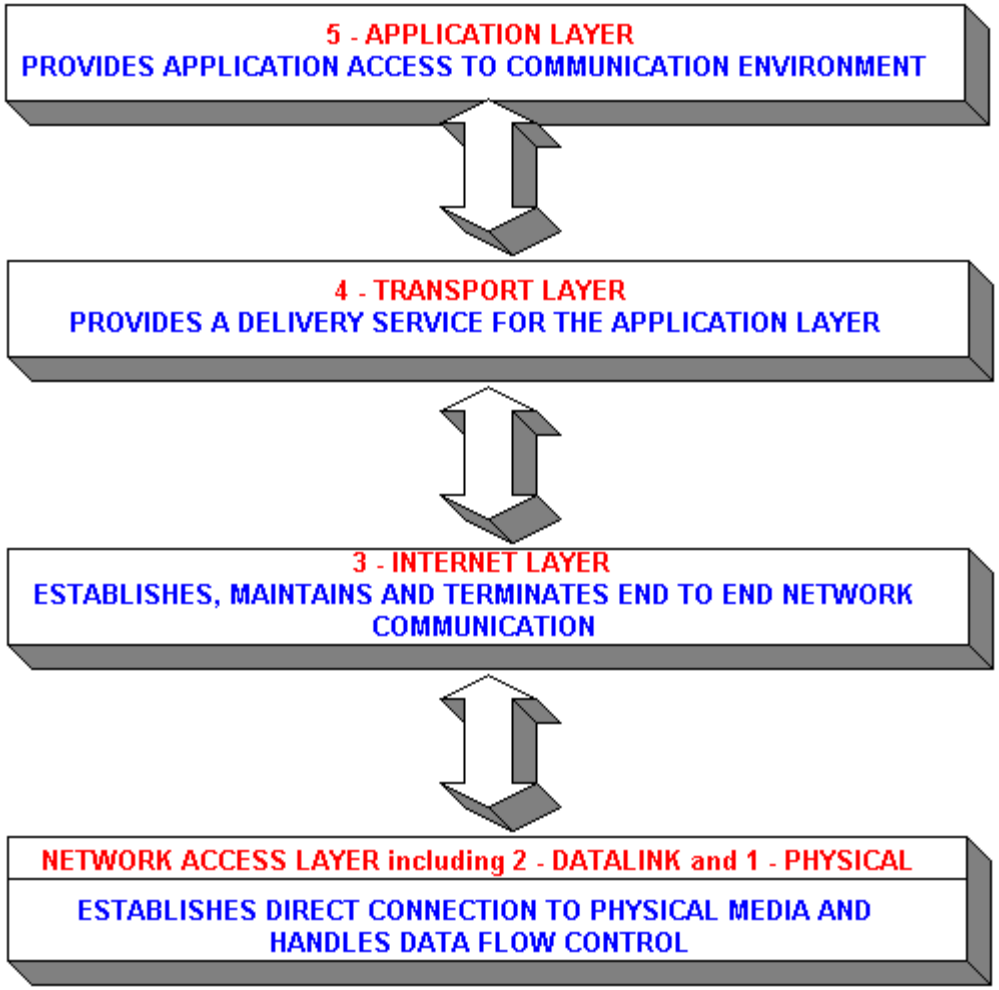
The Internet Protocol (IP)

- A key element of IP is IP address, a 32-bit number
- The Internet authorities assign ranges of numbers to different organizations
- IP is responsible for moving packet of data from node to node
- A packet contains information such as the data to be transferred, the source and destination IP addresses, etc.
- Packets are sent through different local network through gateways
- A checksum is created to ensure the correctness of the data; corrupted packets are discarded
- IP-based communication is unreliable

The Transmission Control Protocol (TCP)

- TCP is a higher-level protocol that extends IP to provide additional functionality: reliable communication
- TCP adds support to detect errors or lost data and to trigger retransmission until the data is correctly and completely received
- Connection
- Acknowledgment

TCP/IP Protocol Suites



HTTP, FTP, Telnet, DNS, SMTP, etc.

TCP, UDP

IP (IPv4, IPv6)

The World Wide Web (WWW)

- **WWW** is a system of interlinked, hypertext documents that runs over the Internet
- Two types of software:
  - **Client**: a system that wishes to access the information provided by servers must run client software (e.g., web browser)
  - **Server**: an internet-connected computer that wishes to provide information to others must run server software
  - Client and server applications communicate over the Internet by following a protocol built on top of TCP/IP – HyperText Transport Protocol (**HTTP**)

Basics of the WWW

- **Hypertext**: a format of information which allows one to move from one part of a document to another or from one document to another through **hyperlinks**
- Uniform Resource Locator (**URL**): unique identifiers used to locate a particular resource on the network
- **Markup language**: defines the structure and content of hypertext documents

Web Client: Browser

- Makes HTTP requests on behalf of the user
  - Reformat the URL entered as a valid HTTP request
  - Use DNS to convert server's host name to appropriate IP address
  - Establish a TCP connection using the IP address
  - Send HTTP request over the connection and wait for server's response
  - Display the document contained in the response
    - If the document is not a plain-text document but instead is written in HTML, this involves rendering the document (positioning text, graphics, creating table borders, using appropriate fonts, etc.)

Web Servers

- Main functionalities:
  - Server waits for connect requests
  - When a connection request is received, the server creates a new process to handle this connection
  - The new process establishes the TCP connection and waits for HTTP requests
  - The new process invokes software that maps the requested URL to a resource on the server
  - If the resource is a file, creates an HTTP response that contains the file in the body of the response message
  - If the resource is a program, runs the program, and returns the output

Static Web: HTML/XHTML, CSS

- **HTML** stands for **H**yper**T**ext **M**arkup **L**anguage
  - It is a text file containing small markup tags that tell the Web browser how to display the page
- **XHTML** stands for e**X**tensible **H**yper**T**ext **M**arkup **L**anguage
  - It is identical to HTML 4.01
  - It is a stricter and cleaner version of HTML
- **CSS** stands for **C**ascading **S**tyle **S**heets
  - It defines how to display HTML elements

The End