*Internet*

* A global system of interconnected computer networks that use the standard Internet Protocol Suite (TCP/IP) to serve billions of users worldwide.
* It is a *network of networks* that consists of millions of private, public, academic, business, and government networks, of local to global scope, that are linked by a broad array of electronic, wireless and optical networking technologies.

*Vinton Cerf*

* Father of the internet

*World Wide Web*

* A global information medium which users can read and write via computers connected to the Internet. The term is often mistakenly used as a synonym for the Internet itself, but the Web is a service that operates over the Internet, as e-mail does.
* In September 1994, Berners-Lee founded the World Wide Web Consortium (W3C) at the Massachusetts Institute of Technology with support from the Defense Advanced Research Projects Agency(DARPA) and the European Commission. It comprised various companies that were willing to create standards and recommendations to improve the quality of the Web.

*Tim-Berners Lee*

* Father of the web

HTTP (hypermedia)

* Application layer used primarily to retrieve hypertext (on hypermedia) documents and resources on the World Wide Web
* Jointly developed by the W3C and the IETF

Protocol

* Set of rules need to be followed.

Fundamentals

* HTTP typically runs on top of TCP/IP, using TCP port 80 by default (TCP port 443 for HTTPS).
* HTTP resources are identified using URIs (specifically, HTTP URLs)
  + Scheme (*http:* or *https:*)
  + (optional) authentication information
  + Host and (optional) port number
  + Path (resolved to the document root on the server) to the resource
  + (optional) scheme-specific parameters
  + (optional) URL-encoded query
  + (optional) bookmark (or fragment identifier)

<http://jo:secret@myserver.com:8080/products/price.jsp;sessionid=123456?cat=school&man=xyz#summary>

scheme username password host port path scheme specific parameters start of query URL query fragment

* HTTP is based on client-server architecture
* Clients, aka user agents (UA):
  + Web browsers, web crawlers, email clients, other end user tools and applications
* Servers:
  + Origin servers, proxy servers, gateways, tunnels
* HTTP uses a request-response standard protocol
  + The client sends an HTTP request message to the server
  + The server processes the request and replies with an HTTP response message
* HTTP is a stateless communications protocol
  + Servers do not keep information about clients in between requests
  + Web applications effect session tracking using mechanism such as cookies on URL-encoded session information to keep track of related client requests
* HTTP provides support for other functionalities such as cache control, content media type (MIME) specification, language and character set specification, content/transfer coding, client-server protocol negotiations, persistent connections, request pipelining, etc.