Internet

The internet is a global information system that is logically linked together by globally unique address. It is able to support communications using time TCP/IP. It can also be described like a network of networks that consists of millions of private, public, academic, business, and government networks, of local to global scope. Vinton Cerf is the father of the internet

* Etymology
  + : Inter – Crossing Boundaries
  + : Network – Connection
* Age: 48, started around 1969. First began with the ARPANET

WWW

Allows documents to be connected to other documents by hypertext link

* Age: 28, started around 1989
* Creator: Tim Berners-Lee
* WAIS: Wide Area Informaation Servers, a specialized database
* GOPHER (protocol): designed for searching, retrieving, and indexing information in services
* USENET: Newsgroup

HTTP stands for Hypertext Transfer Protocol. It is a stateless communications protocol application layer used to access resources (hypertext/hypermedia) on the WWW, invented by Tim Berners-Lee. It was developed by W3C (World Wide Web Consorcium) and IETC (Internet Engineering Task Force)

## HTTP versions

* HTTP 0.9 (1991)
* HTTP 1.0 (RFC 1945, May 1996)
* HTTP 1.1 (RFC 2068, Jan 1997, RFC 2616 Jun 1999) RFC 7230 – Jun 2014
* HTTP 2 RFC 7540 May 2015 (push protocol)

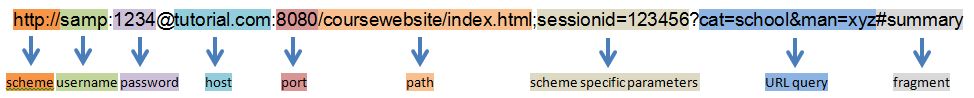
## HTTP Fundamentals

HTTP usually runs on top of TCP/IP using TCP using port 80 by default, or TCP port 443 for HTTPS (HTTP over SSL/TLS (Transport Layer Security)). HTTP resources are identified using URIs (Uniform Resource Identifier) specifically,HTTP URLs

## HTTP Resource Addressing

* Scheme - ex. http: or https:
* Authentication Information Protocol (optional)
* Host - The host name identifies the host that holds the resource. For example, www.example.com. A server provides services in the name of the host, but hosts and servers do not have a one-to-one mapping.
* Port number (optional) - Most servers use the well-known port numbers for HTTP and HTTPS , so most HTTP URLs omit the port number.
* Path - The path identifies the specific resource in the host that the web client wants to access. For example, /software/htp/cics/index.html.
* Scheme (optional)-The scheme identifies the protocol to be used to access the resource on the Internet. It can be HTTP (without SSL) or HTTPS (with SSL).
* URL-encoded query (optional) -The query string is usually a string of name and value pairs; for example, term=bluebird. Name and value pairs are separated from each other by an ampersand (&); for example, term=bluebird&source=browser-search.
* Bookmark/Fragment identifier (optional)

#### Example



HTTP is based on client-server architecture

## Client

Clients, AKA User Agents(UA) host web resources (HTML, CSS, JavaScript)

* Client requests to the server
* User agents – accessing resources
* Examples: Web browsers, web crawlers/spiders, email clients, webtools & app

## Server

The server processes the requests and replies with an HTTP response message.

* Origin server – actual pages or CSS
* 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

### Important notes:

* HTTP 1.1, uses a pull protocol where once a client sends a requests, the server will give what is needed.
* HTTP 2, uses the push protocol where a server can push information.
* Stateless communications protocol ( servers do not keep information about clients in-between requests)
* Cookie - A very small text file
* Cache - Local storage/copy of resource that is fetched from a server

### Acronyms:

* TLS - Transport Layer Security
* Poll - frequent Requests
* MIME - Multipurpose Internet Mail Extensions (text/html)
* DNS - Domain Name System

## HTTP Provides support for other functionalities such as:

### Cache control

* Proxies can be caching servers
* Note: Locality of Memory Principle

### Content media type(MIME) specification

* Tell you what type of content it is. Example: text/html

### Language and character set specification

### Content/transfer coding

* Example: Compression

### Content negotiation

* Example: Mobile and desktop has different content even if it's the same website

### Client server protocol negotiation

* If connections have different protocols. Example: HTTP 1.0 and HTTP 2
* Connections have to negotiate which protocol to use

### Persistent connections

* Introduced in HTTP 1.1
* In HTTP 1, connections where non-persistent so once a request has been fulfilled the connection would cut off. This made sense before since websites back then have little content. However nowadays the non-persistent connection cannot be used anymore. because of overhead. More sites have more requests and for each request there will be a transfer, which called for the need of persistent connections. So when asking for a request, it will inform the server not to close the connection.

### Request pipelining

* Introduced in HTTP 1.1
* There is still overhead in waiting for one request to be finished in a persistent connection. Request pipelining enables you to fetch the next instruction EVEN IF the first one isn't done yet.
* HTTPS
* Status line
  + HTTP protocol version
  + Status code
  + Reason phrase
* Response headers
* Empty line
* Message body

GET

* Most commonly used HTTP method
* Can be combined with conditional and/or range request headers to effect conditional and/or partial resource retrieval
* Must be supported by all general-purpose servers.

HEAD

* Identical to GET, except the message body is not included in the response

POST

* Request that the origin server accept the entity enclosed in the request as a new subordinate of the resource identified by the Request-URI in the Request-Line.

OPTIONS

* Request for information about the communication options available on the request/response chain identified by the Request-URI. This method allows the client to determine the options and/or requirements associated with a resource, or the capabilities of a server, without implying a resource action or initiating resource retrieval.

TRACE

* Request the server to “echo” back to the client the received request
* Typically used for testing/diagnostics of the request chain

PUT

* Request the server to store the enclosed entity in the message under the specified request URI

DELETE

* Request the server to delete the resource identified by the request URI

CONNECT

* Reserved for use of tunneling proxy servers

Idempotent Methods

* The methods GET, HEAD, PUT and DELETE share this property.
* D
  + DARPA -Defense Advanced Research Projects Agency
    - - an Agency from the US defence task to invent technology to be used by the military.
    - - working on packet switched vs circuit switch
    - - store and forward
  + Discussion System- Collection of newsgroup where users post messages and distributed via usenet servers
  + DNS- Domain Name System
* E
  + Email - Electronic Mail
* H
  + HTTP (Hypertext Transfer Protocol) - It is a stateless communications protocol application layer used to access resources (hypertext/hypermedia) on the WWW, invented by Tim- Berners Lee.
* I
  + Internet – global information system that is logically linked together by globally unique address
  + IOT- Internet of things
* L
  + LAN - Local Area Network
* M
  + MIME- Multipurpose Internet Mail Extensions (text/html)
* P
  + Pocketswitch- circuit switch – calling somebody
  + Poll- frequent Requests
* T
  + TCP – Transmission Control Protocol
  + TLS- Transport Layer Security
  + Transceiver- has both transmitter and a receiver
* U
  + URL – Uniform Resource Locator, addresses a specific document
  + Usenet - Worldwide distributed discussion system available on computer

## What is HTML?

HTML is a markup language devised to allow website creation. HTML is used to create electronic documents (called pages) that are displayed on the World Wide Web. Each page contains a series of connections to other pages called hyperlinks. Every web page you see on the Internet is written using one version of HTML code or another.

### HyperText Markup Language

* HyperText is the method by which you move around on the web — by clicking on special text called hyperlinks which bring you to the next page. The fact that it is hyper just means it is not linear — i.e. you can go to any place on the Internet whenever you want by clicking on links — there is no set order to do things in.
* Markup is what HTML tags do to the text inside them. They mark it as a certain type of text (italicised text, for example).
* HTML is a Language, as it has code-words and syntax like any other language.
* The <!DOCTYPE html> declaration defines this document to be HTML5.
* The <html> element is the root element of an HTML page.
* The <head> element contains meta information about the document.
* The <title> element specifies a title for the document.li>
* The <body> element contains the visible page content.
* The <h1> element defines a large heading.
* The <p> element defines a paragraph.

### HTML Versions

* HTML 1.0 - 1991
* HTML 2.0 - November 24, 1995; published as RFC 1866.
* HTML 3.0 - HTML 3.0 builds upon HTML 2.0 and provides full backwards compatibility.
* HTML 3.2 - January 14, 1997; Published as a W3C Recommendation.
* HTML 4.0 - December 18, 1997; Published as a W3C Recommendation; supports more multimedia options, scripting languages, style sheets, better printing facilities, and documents that are more accessible to users with disabilities.
* HTML 4.01 - December 24, 1999; This document obsoletes previous versions of HTML 4.0. This enhances the functionality and interoperability of the Web.
  + Strict – deprecated
  + Transitional – still make use the deprecated
  + Frameset – using frames
* HTML 5 (working draft)- October 28, 2014; In this version, new features are introduced to help Web application authors, new elements are introduced based on research into prevailing authoring practices, and special attention has been given to defining clear conformance criteria for user agents in an effort to improve interoperability.
* HTML 5.1 - November 1, 2016; The W3C started working on the draft of HTML 5.1 with the goal of fixing some of the issues that were left open in HTML5.
* HTML 5.2 - February 28, 2017; A working draft; The W3C aims to produce an HTML 5.2 Recommendation in late 2017 that would obsolete HTML 5.1.

### Important Notes:

* World Wide Web Consortium (W3C) - An international community that oversees the development of HTML and open standards to ensure the long-term growth of the Web
* W3C Recommendation – documents specify the HTML standards.
* Tim Berners-Lee –invented HTML.

## Document Object Model (DOM)

It is a programming interface for HTML and XML documents. It provides a structured representation of the document and it defines a way that the structure can be accessed from programs so that they can change the document structure, style and content.