**DEFINITION**

Hypertext Transfer Protocol (HTTP) is an application-level **protocol** which is used to access **resources** (eg. hypertext, hypermedia) on the **World Wide Web***(redirect to WWW definition)*through the **Internet**. This protocol is a set of rules that defines the format of data that is being exchanged within or between computers (MDN web docs). HTTP is the language used by applications like browsers, server, proxies to communicate with one another.

**HTTP Communication**

In order for you(the client) to access websites like Youtube, Facebook, Twitter or CNN, you need a web browser where you can enter the URL of the site you want to visit. What the browser does is to connect to the server of the website and send a request message (using HTTP) after a connection with the server has been established. When the server receives the request message, it will try to interpret the message and send a response message (using HTTP) containing the HTML file of the site you want to visit or an HTTP 404 error message which means that the resource being requested is not found in the server. I the resource is found, the browser will then interpret the response message and try to display the html file in your browser. When the request has been serviced, the connection between the client and server will be terminated. The response message may contain links to other files which needs to be accessed in order for the website to properly render, the connection must be maintained even after the first request has been serviced. (Persistent connection will be discussed here in detail)

**Intermediaries**

In a request or response message, there are intermediaries that play their role to help in the exchange of information. The three common forms of intermediaries are the following:

* Proxy

A proxy is a forwarding agent, receiving requests for a URI in its absolute form, rewriting all or part of the message, and forwarding the reformatted request toward the server identified by the URI. (rfc2616)

* Gateway

A gateway is a receiving agent, acting as a layer above some other server(s) and, if necessary, translating the requests to the underlying server's protocol. (rfc2616)

* Tunnel

A tunnel acts as a relay point between two connections without changing the messages; tunnels are used when the communication needs to pass through an intermediary (such as a firewall) even when the intermediary cannot understand the contents of the messages.(rfc2616)

**Some Characteristics**

**HTTP is a client-server protocol.** Requests are sent by the client which is usually a Web browser, to the server which provides a response. Clients or user-agents is not necessarily a web browser, it could also be a robot.

* Origin server - this is the actual machine where the resource resides
* Proxy servers - this are servers that serve the request of a client on behalf of the origin server which is in the case of cache.

**HTTP is a pull protocol.** The pull of resources by the user-agent will not happen unless the user-agent sends a request

**HTTP serves and forgets.** After serving the requests sent by user-agents, the Web server forgets the request and serve another request from another user-agent or the same user-agent.

Components of HTTP-based systems