What is a Web Browser?

Introduction

A browser is a software designed to find and display content on the World Wide Web. This content might be **a web page, a pdf document, a picture, a video or any other content**. The richest content is a web page. A web page is usually an HTML document which stands for **Hyper Text Markup Language**. Web browsers are also able to run **CSS (Cascading Style Sheet) and Javascript code**. CSS styles the HTML pages to make them look beautiful. And Javascript gives the page the ability to interact with users, show dynamic content and listen to specific **events**.

Events are fired inside the browser window and tend to be attached to a specific item that resides in it (this might be a single element, set of elements, the HTML document loaded in the current tab or the entire browser window). Different types of events occur on the browser. For example:

* Clicking or hovering the cursor over a button, link, picture, etc.
* Pressing a key.
* Resizing, minimizing, maximizing or closing the browser window.
* Loading a web page.
* Submitting a form.
* Playing, pausing a video.

Browsers have common features like address bar, back and forward buttons, bookmarking options, buttons for refreshing the page and stopping the request and home button to direct the user to the home page.

| **https://docs.google.com/uc?id=1fKwU2GKKP0cqfPRrHqmBYcQU3WDUqi-W** |
| --- |
| *Web browsers* |

Q: What Is Web Browser?  
A: A web browser is a program that you use to view web pages. Some of the most popular web browsers are Microsoft Internet Explorer, Google Chrome, Mozilla Firefox.

Q: What Is Web Browser?  
A: A web browser is a program that you use to view web pages. Some of the most popular web browsers are Microsoft Internet Explorer, Google Chrome, Mozilla Firefox.

 - Interview Q&A

 - Interview Q&A

### Introduction

Web browsers have seven high level components. These are **User Interface, Browser Engine, Rendering Engine, Networking, Javascript Interpreter, UI Backend and Data Storage**.

| **https://docs.google.com/uc?id=1hOJ6ySahv7Y_2DpvXrxTYSAaxaMw2Z0P** |
| --- |
| Web browser structure |

 User Interface :The user interface is where the users interact with the browser.

 Browser Engine :The browser engine is the bridge between the user interface and the rendering engine. It relays the necessary information to the rendering engine for performing a render.

 Rendering Engine :The rendering engine is responsible for rendering the requested content and displaying it on the screen. It may be either an HTML, a CSS or a JS content.

 Networking :The networking component uses the URL given by the user and retrieves the page via a protocol called HTTP or a file via FTP protocol.

 Javascript Interpreter :The Javascript interpreter, as the name suggests, interprets and execute the Javascript code of the web page. Then the results are sent to the rendering engine for displaying.

 UI Backend :The UI Backend is responsible for drawing basic widgets like combo boxes and windows. It uses operating system user interface methods.

 Data Storage :The data storage is persistence layer. The browser may need a small storage area to store, process or show data. So there are some mechanisms for storing like localStorage, sessionStorage, IndexedDB, WebSQL and FileSystem.

## What is HTTP?

### Introduction

**Hypertext Transfer Protocol (HTTP)** is an application layer designed within the framework of internet protocol suite. It is used for transferring **text, image, sound, video or any other type of multimedia**[**files**](https://lms.clarusway.com/mod/lesson/view.php?id=1052). When a web browser tries to reach a specific web address, it wants to get in touch with the computer holding all the necessary [files](https://lms.clarusway.com/mod/lesson/view.php?id=1052). So basically the client web browser sends an **http request** to get an HTML content or any specific data staying inside the server computer. When the server computer gets the request, it returns an **http response** containing the information requested.

The statement written in the address bar is called **Uniform Resource Locator (URL)**. URLs are the keys to specify a computer and its IP address in the network. The browser should use this key to find out the IP address of a particular computer. The system that is responsible to hold the values of these keys is [**Domain Name System (DNS)**](https://lms.clarusway.com/mod/lesson/view.php?id=943) [servers](https://lms.clarusway.com/mod/lesson/view.php?id=1015). If the DNS server knows the address, it answers. Otherwise, the DNS server starts to ask other DNS [servers](https://lms.clarusway.com/mod/lesson/view.php?id=1015) and fetches the information at the end. Therefore DNS gives it back to the browser and the browser can find the way to the correct destination using HTTP and other related protocols.

| **https://docs.google.com/uc?id=1eu-7nSL7a1q2XFiEYcUUtg8BLd5hmqey** |
| --- |
| Client and server model |

HTTP Request and Response

As mentioned before, HTTP (Hypertext Transfer Protocol) is a request-response protocol. A client on one side (web browser) asks or requests something (a resource located on the server) from a server (a computer responsible to respond to that request) and the server on the other side sends a response to that client. But before this request and response activities, a connection should be established. When we open our browser and write down the URL (Uniform Resource Locator), we are making the first step to connect to the resource. The URL is the address of that resource. It could be a website or a web service, a pdf or anything similar. If we pull apart the URL, we will see the structure like the one below.

URL = http: // host : port / path ? query

The host is the computer where the resource is stored. The host computers are given a human readable alias and this alias defines the IP address of the hosting computer in the URL. The port is the port number of the hosting computer and the path is the address where the resource file is located in the hosting computer. The query parameter is the value we give to the resource file while we try to establish the connection.

| **https://docs.google.com/uc?id=1mGrnn_dLCJ4IWme3lPPClWEfFa-6kFvM** |
| --- |
| *Request Message* |

If the URL starts with **HTTP** and there is no port value, the default port number is 80. If it starts with **https**, then the default port number is 443. The **path** and the **query** characters are optional. When the connection is established, the medium is available for transferring the request. But what is this request? A request is a text message and it has 4 parts which are request line, headers (optional), a blank line and a message body (optional).

| **https://docs.google.com/uc?id=1BAF_bekv5JkULW_uY5_CAZ2NHN9Hzzsr** |
| --- |
| *Request Message Example* |

The server responds to that request with an HTTP response message. The structure of the response message is similar to the request message. The one thing changed is the status line in place of the request line. A status line consists of the protocol version followed by a numeric status code and its associated meaning. Status code element is a 3-digit integer where the first digit of the status code defines the category of response. There are 5 categories. These categories are in its general state are expressed below.

* 1xx -> Informational
* 2xx -> Success
* 3xx -> Redirection
* 4xx -> Client Error
* 5xx -> Server Error

| **https://docs.google.com/uc?id=1MJO-DS1mavA1ECFKX2vvZ2QOiyfILSZf** |
| --- |
| *Response Message Example* |

The rest is handled by the browser and its engines.

Q: What is status code in HTTP?  
A: It is a standard response code given by web [servers](https://lms.clarusway.com/mod/lesson/view.php?id=1015) on the Internet. It helps to identify the cause of a problem when a web page or other resource does not load properly. There are two major group of HTTP status code error exist:

* 4xx Client Error
* 5xx Server Error

 - Interview Q&A

Q: What is status code in HTTP?  
A: It is a standard response code given by web [servers](https://lms.clarusway.com/mod/lesson/view.php?id=1015) on the Internet. It helps to identify the cause of a problem when a web page or other resource does not load properly. There are two major group of HTTP status code error exist:

* 4xx Client Error
* 5xx Server Error

 - Interview Q&A

Q: What are the header fields in HTTP?  
A: HTTP header fields allow the client and server to pass information with the request and response message. Following are the header fields in HTTP:

* **General header:**It applies for both request and response message.
* **Request header:** It contains information for the request message.
* **Response header:** It is used to contain response header information sent by the web server.
* **Entity header:**It is used to contain more information about the body of the entity.

 - Interview Q&A

Q: What are the header fields in HTTP?  
A: HTTP header fields allow the client and server to pass information with the request and response message. Following are the header fields in HTTP:

* **General header:**It applies for both request and response message.
* **Request header:** It contains information for the request message.
* **Response header:** It is used to contain response header information sent by the web server.
* **Entity header:**It is used to contain more information about the body of the entity.

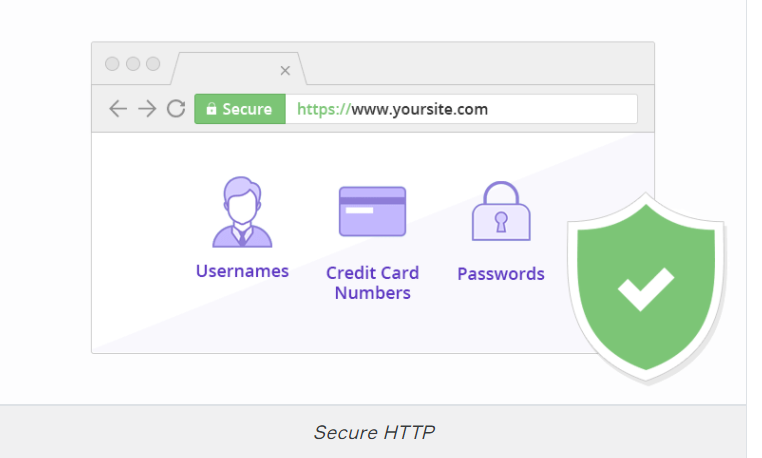
- Interview Q&A

### Certificates and HTTPS

**SSL** stands for Secure Sockets Layer and, in short, it's the standard technology for keeping an internet connection secure and safeguarding any sensitive data that is being sent between two systems, preventing criminals from reading and modifying any information transferred, including potential personal details. The two systems can be a server and a client (for example, a shopping website and browser) or server to server (for example, an application with personal identifiable information or with payroll information).

**TLS** (Transport Layer Security) is just an updated, more secure, version of SSL. We still refer to our security certificates as SSL because it is a more commonly used term.

**HTTPS** (Hyper Text Transfer Protocol Secure) appears in the URL when a website is secured by an SSL certificate. The details of the certificate, including the issuing authority and the corporate name of the website owner, can be viewed by clicking on the lock symbol on the browser bar.



Q: What are SSL certificates?  
A: SSL is a standard security protocol which ensures confidentiality and integrity of data while in transit. It encrypts the data flow between the web browser and web server, hence ensures confidentiality. Also, web server and browser exchanges key to decrypt the data, which ensures the integrity of data.

Q: What are SSL certificates?  
A: SSL is a standard security protocol which ensures confidentiality and integrity of data while in transit. It encrypts the data flow between the web browser and web server, hence ensures confidentiality. Also, web server and browser exchanges key to decrypt the data, which ensures the integrity of data.

 - Interview Q&A

Q: What are the benefits of HTTPS certificate?  
A: The major benefits of HTTPS certificate are:

* Customer information like credit card number and ATM pin is encrypted and cannot be easily tracked.
* Customers trust and prefer to purchase from the sites that use HTTPS protocol.
* This protocol shows authenticate register domain as secure connection.

 - Interview Q&A

Q: What are the benefits of HTTPS certificate?  
A: The major benefits of HTTPS certificate are:

* Customer information like credit card number and ATM pin is encrypted and cannot be easily tracked.
* Customers trust and prefer to purchase from the sites that use HTTPS protocol.
* This protocol shows authenticate register domain as secure connection.

 - Interview Q&A

## Cookies

### Cookies

A **cookie** is a tiny little file that’s stored on your computer. It contains the address of the web site and codes that your browser sends back to the web site each time you visit a page there. Cookies don’t usually contain personal information or anything dangerous; they’re usually innocuous and useful.

When you browse the web, the webserver needs to know who you are if you want to do things that require logging in or putting items in a virtual shopping cart or completing any other process that requires web site to remember information about you as you move from page to page. The most commonly used trick that allows web sites to keep track of what you’re doing is called setting cookies.

The **advantages** of cookies are:  
\* The cookies are simple to use & implement,  
\* They do not require any server resources,  
\* They are stored on the user’s computer, so, no extra burden on the server,  
\* They are light in size so they occupy less memory and you do not need to send back the data to the server.

The **disadvantages** of cookies are:  
\* The cookies are not secure as they are stored in a clear text, hence no sensitive information should be stored in cookies,  
\* They may pose to a possible security risk because anyone can open and tamper with the cookies.

Q: What is Cookie?  
A: Cookie provides a simple way to identify session among a group of HTTP/HTML requests. The cookie value is often an index into a table stored in the memory of a Web server that points to an in-memory object holding the user's records. This has many potential problems: If the user's request is routed to a different server in a subsequent request, the session information is unknown to the server.  
If the user is routed to a different server and the server is part of an application cluster, then all the [servers](https://lms.clarusway.com/mod/lesson/view.php?id=1015) that could receive the user's request must have a way to synchronize the session data. Storing cookies and synchronizing sessions among clusters of server usually requires configuration, storage space, and memory.

Q: What is Cookie?  
A: Cookie provides a simple way to identify session among a group of HTTP/HTML requests. The cookie value is often an index into a table stored in the memory of a Web server that points to an in-memory object holding the user's records. This has many potential problems: If the user's request is routed to a different server in a subsequent request, the session information is unknown to the server.  
If the user is routed to a different server and the server is part of an application cluster, then all the [servers](https://lms.clarusway.com/mod/lesson/view.php?id=1015) that could receive the user's request must have a way to synchronize the session data. Storing cookies and synchronizing sessions among clusters of server usually requires configuration, storage space, and memory.

 - Interview Q&A

 - Interview Q&A