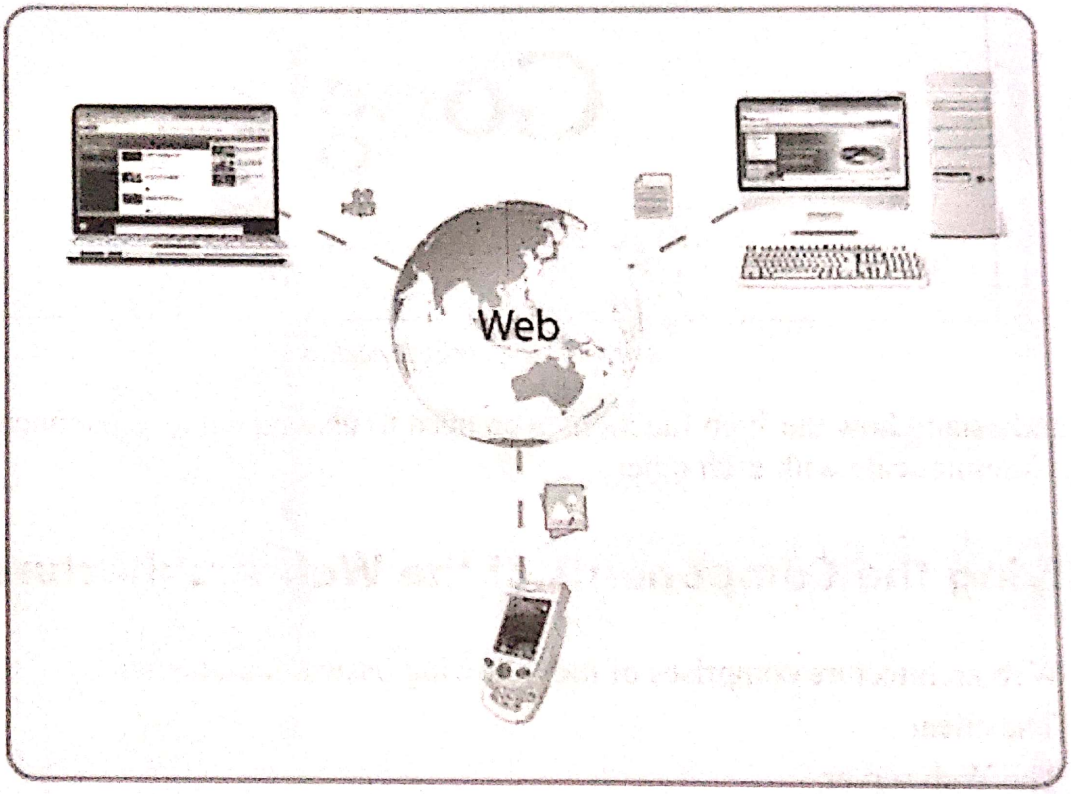
Introduction to Web Architecture

The Internet has been the fastest growing technology. It has changed the way business is conducted. It is a comprehensive system of interconnected networks and is a huge storehouse of information. This is made possible by the means of the World Wide Web (WWW). The following figure depicts the WWW.

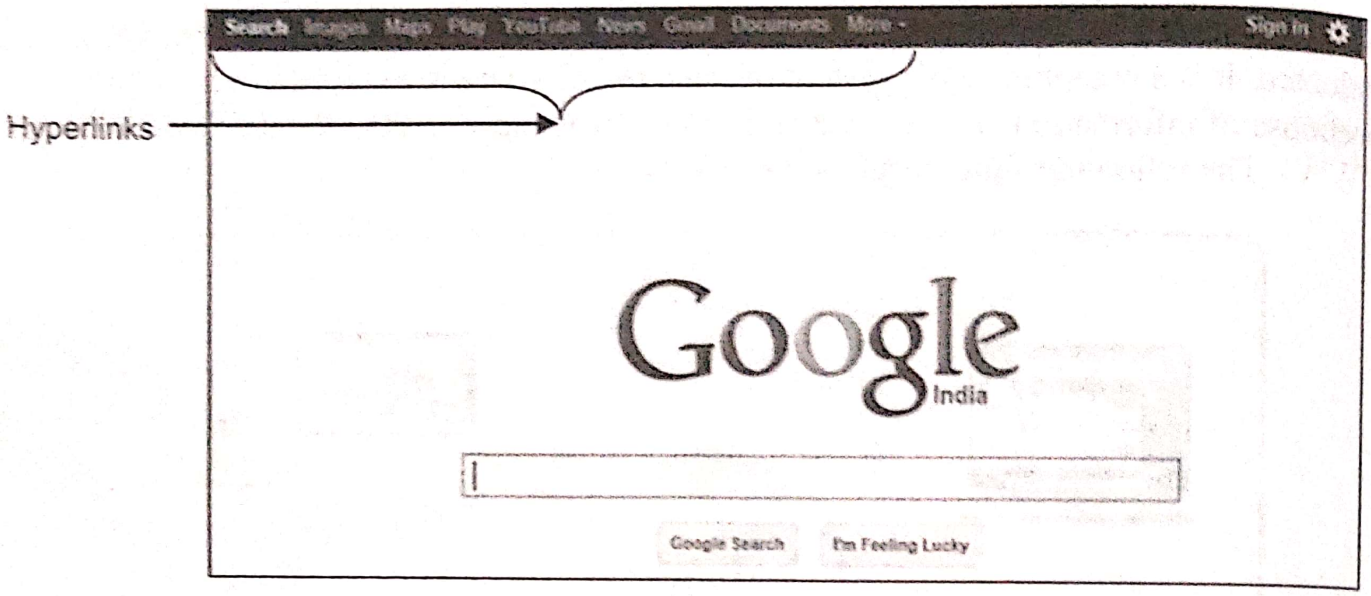


The World Wide Web

The WWW, popularly known as the Web, is a collection of several Web pages. A Web page is an electronic document containing text, images, audio, or videos. These are interlinked with each other. This interconnection among Web pages is achieved by using hypertexts. A hypertext is the highlighted or underlined text on a Web page. A hypertext connects the content on one Web page to the content on another Web page. For this reason, the hypertexts are also known as hypertext links or hyperlinks. Clicking a hyperlink opens the Web page that the hyperlink is linked to.

A set of interconnected Web pages displaying related information on a particular subject is called a website. Each website has a unique address on the Internet. This address is known as the Uniform Resource Locator (URL). Websites are hosted on a Web server and are accessed by using client applications, such as a Web browser. A Web browser lets you specify the URL of a website and opens the home page of the website. The home page of a website is the first page of the website that contains links to all other pages.

The following figure shows a Web page with hyperlinks.



A Web Page with Hyperlinks

To understand how the Web functions, you need to understand its components and how they communicate with each other.

Identifying the Components of the Web Architecture

The Web architecture comprises of the following basic components:

- The client
- The Web server
- Protocol
- URL

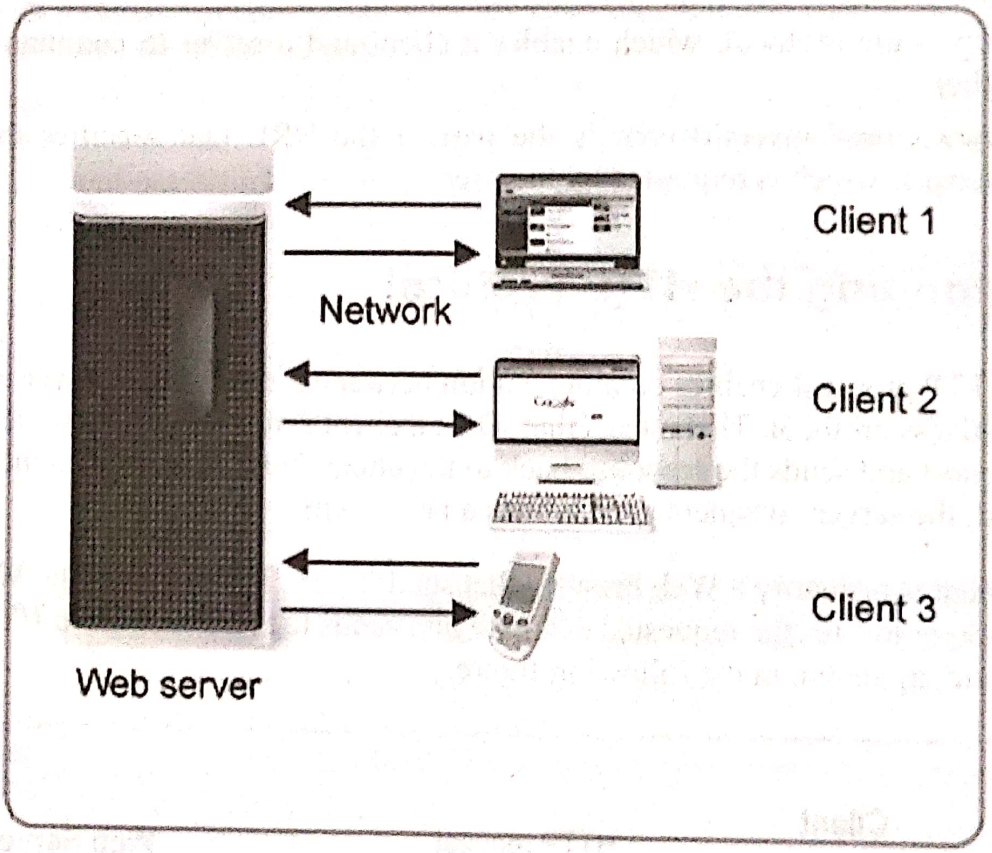
The Client

The client is any application that enables you to access the contents on the Web. These applications are popularly known as Web browsers. Internet Explorer, Chrome, and Mozilla Firefox are some of the popular Web browsers. By using any of these Web browsers, the client sends a request for the resources, such as a Web page, an image, an audio file, a video file, or any other document stored on the Web server.

The Web Server

A server that delivers Web pages, images, or other resources is known as a Web server. It is either a computer or software running on a computer that is responsible for handling the client requests. When a Web server receives a request from the client, it interprets the request and responds with the requested resource.

The following figure depicts the client-server interaction.



The Client-server Interaction

The preceding figure displays three devices: a laptop, a desktop, and a mobile phone accessing and sending a request to the Web server. The Web server processes these requests and sends the response to the respective clients. This communication between the client and the Web server is done using a protocol.

Protocol

A protocol is a set of rules that defines how computers must communicate with each other over a network. These rules define common guidelines, such as the data format and the speed of data transfer. However, there are different types of protocols, which can be classified based on their use, such as browsing the Internet, and sending and receiving e-mails. For example, the Hyper Text Transfer Protocol (HTTP) protocol enables exchange of information across the Internet.

URL

A URL is a string that refers to a Web resource. A client initiates a request for a resource by typing the URL into the Web browser. For example, if you want to access the website of a university named, GrantUniversity, you can initiate the request by typing the following URL in the Web browser:

- http://www.GrantUniversity.com

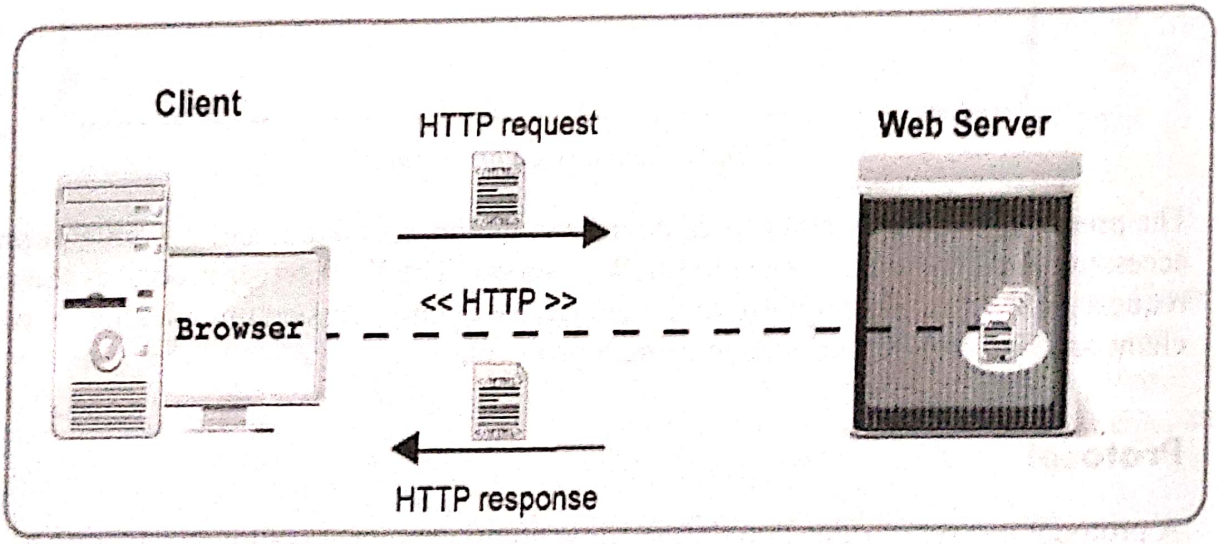
In the preceding example:

- http is the protocol, which enables a client and a server to communicate with each other.
- www.GrantUniversity.com is the part of the URL that specifies the name of the website, which is requested by the user.

Understanding the HTTP Protocol

The HTTP protocol enables communication between a client and a server over the Web. It is a stateless protocol. This means that when a client sends a request, the server processes the request and sends the response back to the client. Now, when the client sends another request, the server considers the client as a new client.

The client is primarily a Web browser that sends an HTTP request to the Web server. The Web server locates the requested resource and sends the response as an HTTP response to the client, as shown in the following figure.

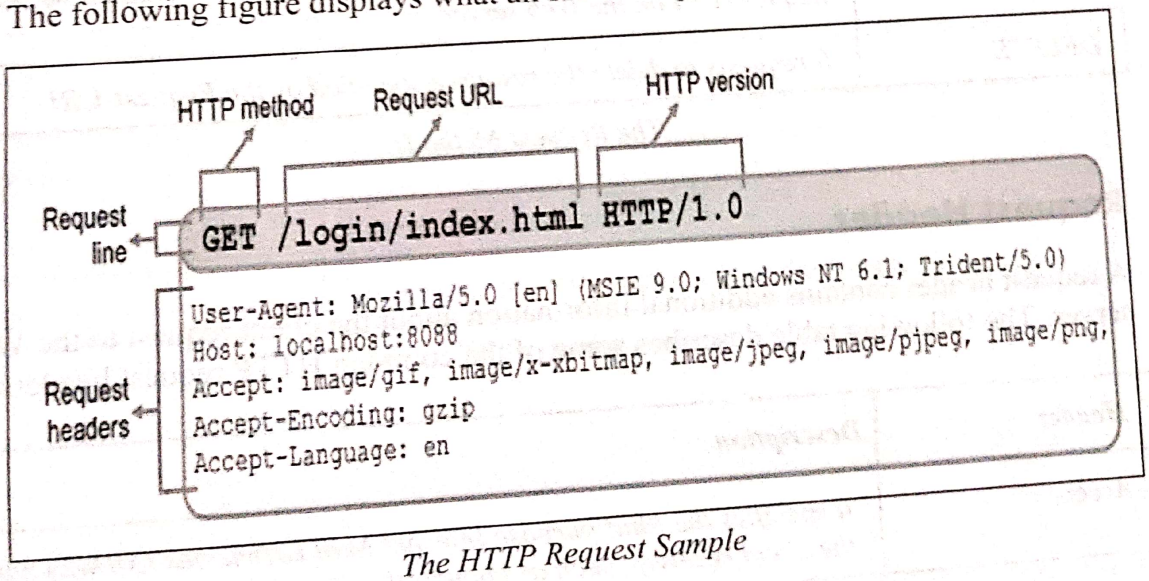


The HTTP Request and Response Cycle

The HTTP request consists of:

- **A request line**: It consists of:
  - **Request method**: It defines the method that is used by the client to request for a resource, such as GET, POST, and HEAD.
  - **Request URL**: It defines the location of the requested resource in the Web server.
  - **Version**: It defines the version of the HTTP protocol used for communication.
- **Request header**: Contains additional information about the client, such as the name and version of the Web browser.
- **Message body**: It is an optional part of the HTTP request. It contains the data entered by a user that needs to be sent to the Web server for processing. For example, the username and password information entered by a user on a login page are sent as a part of the message body to the Web server for authentication, when the submit button is clicked by the user. However, if the user requests for a resource by clicking a hyperlink or hypertext, the message body will be empty.

The following figure displays what an HTTP request consists of.

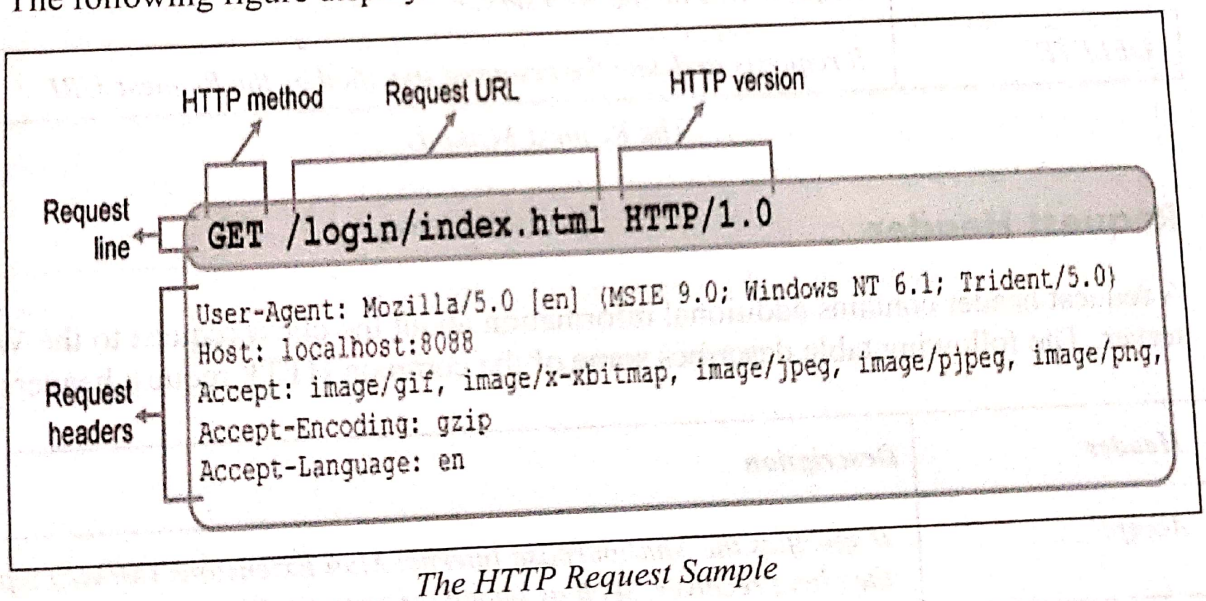


The HTTP Request Sample

The HTTP request consists of:

- **A request line**: It consists of:
  - **Request method**: It defines the method that is used by the client to request for a resource, such as GET, POST, and HEAD.
  - **Request URL**: It defines the location of the requested resource in the Web server.
  - **Version**: It defines the version of the HTTP protocol used for communication.
- **Request header**: Contains additional information about the client, such as the name and version of the Web browser.
- **Message body**: It is an optional part of the HTTP request. It contains the data entered by a user that needs to be sent to the Web server for processing. For example, the username and password information entered by a user on a login page are sent as a part of the message body to the Web server for authentication, when the submit button is clicked by the user. However, if the user requests for a resource by clicking a hyperlink or hypertext, the message body will be empty.

The following figure displays what an HTTP request consists of.



The HTTP Request Sample

Request Methods

HTTP provides various methods to request a resource stored in a Web server. The following table lists some of the request methods and their corresponding operations performed by the Web server.

Request Method	*Operation*
GET	*It retrieves the information specified in the request URL. In addition, it submits the user information encoded as a part of the request URL.*
POST	*It sends a block of data to the Web server for processing. The data is sent as a part of the message body.*
HEAD	*It retrieves only the header information of the response and not the information of the message body.*
PUT	*It uploads the information provided in the request identified by the Request-URI on the Web server.*
DELETE	*It requests to delete the resource specified by the Request-URI.*

The Request Methods

Request Header

A request header contains additional information about the client request to the Web server. The following table describes some of the common HTTP request headers.

Header	*Description*
Accept	*It specifies the Multipurpose Internet Mail Extensions (MIME) types the client receives, such as image/gif or image/jpeg.*
Host	*It specifies the Internet host and the port number of the requested resource.*
User-Agent	*It specifies the information about the client initiating the request, such as the name and version of the Web browser.*
Accept-Language	*It specifies the set of languages that are accepted by the client, such as en is used for English.*
Referer	*It specifies the URI/address of the client from which the Request-URI is received.*

The HTTP Request Header Fields

The HTTP response consists of:

- **Status line**: It consists of the following information:
  - **Status code and reason phrase**: The status code is a three-digit integer value and the reason phrase is a text message that describes the status code.
  - **Version**: It defines the version of the HTTP protocol.
- **Response header**: It contains additional information about the Web server and the response sent to the client, such as age and location.
- **Message body**: It consists of the response in the form of Hyper Text Markup Language (HTML) format.

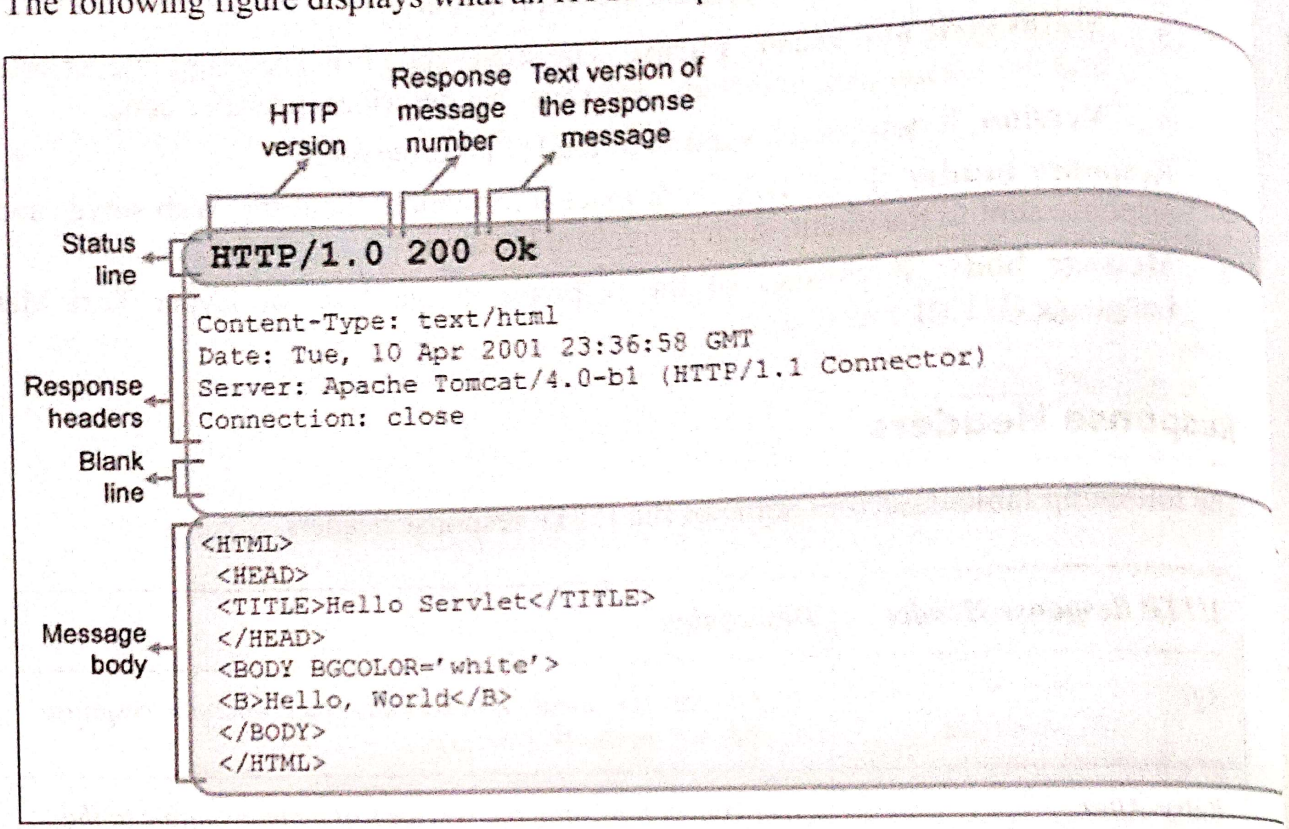
Response Headers

The following table describes some of the HTTP response headers.

HTTP Response Header	*Description*
Age	*It specifies the duration in seconds since when the response is created on the Web server.*
Retry-After	*It specifies how long the service will be unavailable to the client. Used with a response with status code 503 (Service Unavailable).*
Server	*It specifies the information about the server.*

The HTTP Response Headers

The following figure displays what an HTTP response consists of.



The HTTP Response Sample

Status Codes

The following table lists some of the HTTP status codes and their corresponding reason phrases.

Code	*Reason Phrase*
200	*OK*
201	*Created*
202	*Accepted*
301	*Moved Permanently*
304	*Not Modified*
400	*Bad request*
401	*Unauthorized*

Code	*Reason Phrase*
402	*Payment required*
404	*Not found*
500	*Internal Server Error*
503	*Service Unavailable*
505	*HTTP Version Not Supported*

The HTTP Status Codes

The components of the Web architecture, such as the client, URL, the HTTP protocol, and the Web server, together make the Web a powerful information sharing system. However, these days, the Web is not only used for information sharing, but also for creating and modifying information. This has led to the development of the dynamic Web pages.

A dynamic Web page is a Web page that responds to user actions or is dynamically created by a Web program on the basis of a user's request. Consider the example of a dynamic Web page of a banking website that allows you to enter your account number and credentials. After validating and authenticating the details entered by you, your account details are fetched from a database and displayed on the Web page. Such dynamic Web pages can be created by developing a Web application.

Just a minute:

Which one of the following components of the Web architecture delivers Web pages, images, or other resources over the Web?

1. *Client*
2. *Server*
3. *URL*
4. *Protocol*

Answer:

2. *Server*