

UNIT 1



INTRODUCTION TO THE WEB

WEB ARCHITECTURE



- WWW is a two tier architecture consisting of
 - Web servers – produce and deliver information
 - Web clients – retrieve and display information
- Three primary concepts are involved in this process,
 - HTML
 - URL
 - HTTP

HTML



- HTML is a primary language used to encode documents containing hyperlinks.
- HTML 1st version was released in 1991 by Tim Berners-Lee(inventor of WWW)
- Final modified version in may 2001, known as HTML4.01
- HTML documents consist of case-sensitive tags.
- Documents contents are embedded in starting and ending tags. Ex) `<p>web technology</p>`



- Starting tag may have optional attributes that control the behaviour of the tag.
- Documents are linked by special tags called anchor tags. These tags are also called hyperlinks.
- HTML allows to embed all kinds of data (text,images,audio,video,etc)
- HTML documents are viewed by a software called 'browser'.
- When a user selects a hyperlink, the web browser retrieves the document and displays it on the screen.

URL



- WWW is a repository of information called resources.
- URLs are used to address these resources.
- URL is a general addressing protocol, Uniform Resource Identifier.
- URLs are location independent addressing mechanism used on the WWW.

HTTP



- HTTP is an application layer protocol.
- HTTP is stateless.
- HTTP is a request-response protocol.

- It consists of the following steps ,
 - HTTP client establishes a TCP connection to the HTTP server.
 - Client sends an HTTP request to the server specifying the resource it wants to access.
 - Server sends an HTTP response containing the desired information.

MAJOR ISSUES IN WEB SOLUTION DEVELOPMENT



- HTML Performance
- Database issues
- Network bandwidth issues

HTML Performance



- Avoid using too many images on web
- Use nested tables but do not use too many levels of nesting
- Frames are another source bottlenecks; use tables as far as possible.
- Avoid redundant tags
- Avoid using too many comments
- Use relative path names
- Avoid AJAX(Asynchronous Java script And XML) technology – increase the server load.

DATABASE ISSUES



- Many web applications use backend database servers. A good database can increase the performance.
- The following points should be considered while working with databases.
 - Avoid dynamic queries.
 - Use stored procedures.
 - Inspect the join type used in SQL query. A join may increase or decrease the performance of the query.
 - Replicate the database if load on the server is high.

NETWORK BANDWIDTH ISSUES



- In many cases, network performance may degrade overall performance.
- It is advisable to run the web server and the database server on different machines.