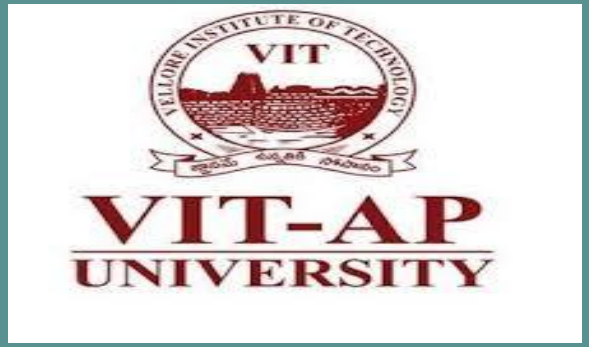
CSE 4004: Web Technologies
MODULE 1: Introduction, HTTP

Faculty: Mrs Daphna Chacko, Assistant Professor, SCOPE, VIT AP



References

1. *"Web Technologies: A computer Science Perspective"*, Jeffrey C Jackson

NB: Images are adapted from Internet

Introduction

- ★ **internet** : Refers to the internetworking of computers to share resources
- ★ **Internet :** globally interconnected system of computers which use the TCP/IP protocols
- ★ **Protocols** : Set of rules followed by communicating devices
- ★ **HTTP** : Protocol used to transfer hypertext
- ★ **FTP** : Protocol used to transfer files
- ★ **TCP** : Transmission control protocol. Deals with port addresses.
    - For example, 80 denotes http, 21 denotes ftp, 443 denotes https
- ★ **IP** : Internet protocol. Used to assign IP addresses to devices
- ★ **IP address** : Numeric id assigned to each device connected to the Internet

What makes the web??

- ★ **The web** : world wide web (www) is a collection of web pages and other web resources linked by hyperlinks and URLs
- ★ **Web pages** : Documents which has hypertext in it
- ★ **Hypertext** : Is a word or picture which is linked to another file from the Internet
- ★ **Markup languages** : Languages that are used to define the parts of the document, using tags. Tags are not displayed but its effects are visible on the text
- ★ **HTML** : is a widely used markup language
- ★ **Web sites** : A collection of related pages
- ★ **Web browsers** : Program that displays the web pages that it retrieves (Eg, Chrome, Firefox)
- ★ **Web server** : A computer that hosts (stores) the webpages

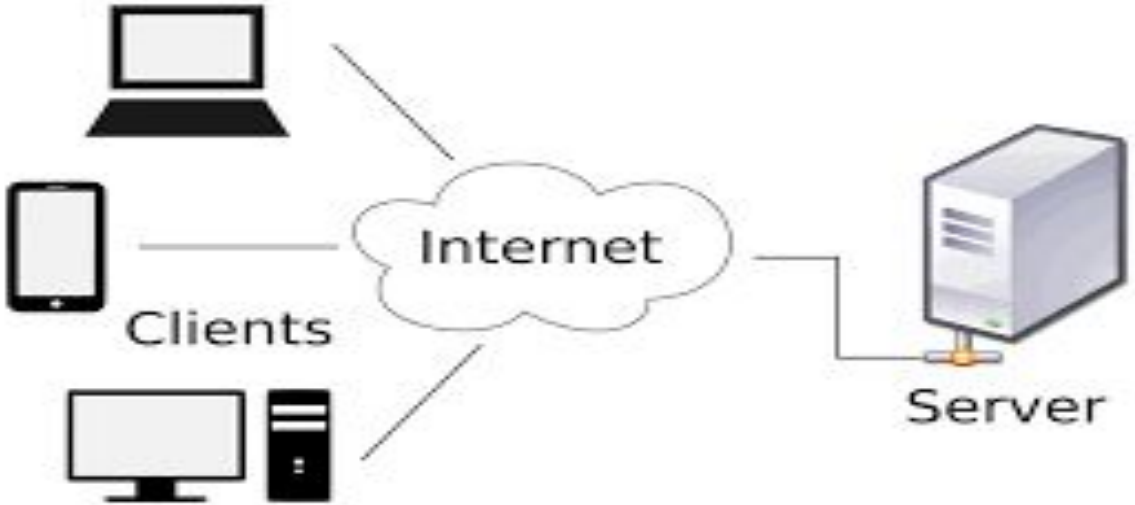
Some more terms related to web

- **Internet Service Provider (ISP)** : A company that provides services related to web
- **Modem** : A hardware device that converts digital data into a form in which it can be transmitted and vice versa.
- **Domain names** : Internet address of the service provider that stores the documents and hosts the sites
- **Domain name system (DNS):** A naming system for computers, services or other devices connected to the Internet or other private networks
- **Uniform Resource Locator (URL)** : Commonly known as web address is a unique address assigned to a web resource.
  - Contains the protocol, domain name and path
  - http://www.vpropel.in/loginn/
- **Internet Engineering Task Force (IETF)**: Develops protocols for Internet, message formats, schemas and languages

Client server model

Client : Internet user who has the required software to access various Internet services. Eg web browser

Server : Software/ system that provides service or the information needed. Eg, web servers, mail server, ftp servers



Internet application protocols

- **HTTP** : HyperText Transfer Protocol
- **FTP** : File Transfer Protocol
- **SMTP** : Simple Mail Transfer Protocol
- **TFTP** : Trivial File Transfer Protocol
- **RIP** : Routing Information Protocol
- **IMAP** : Internet Message Access Protocol

HTTP

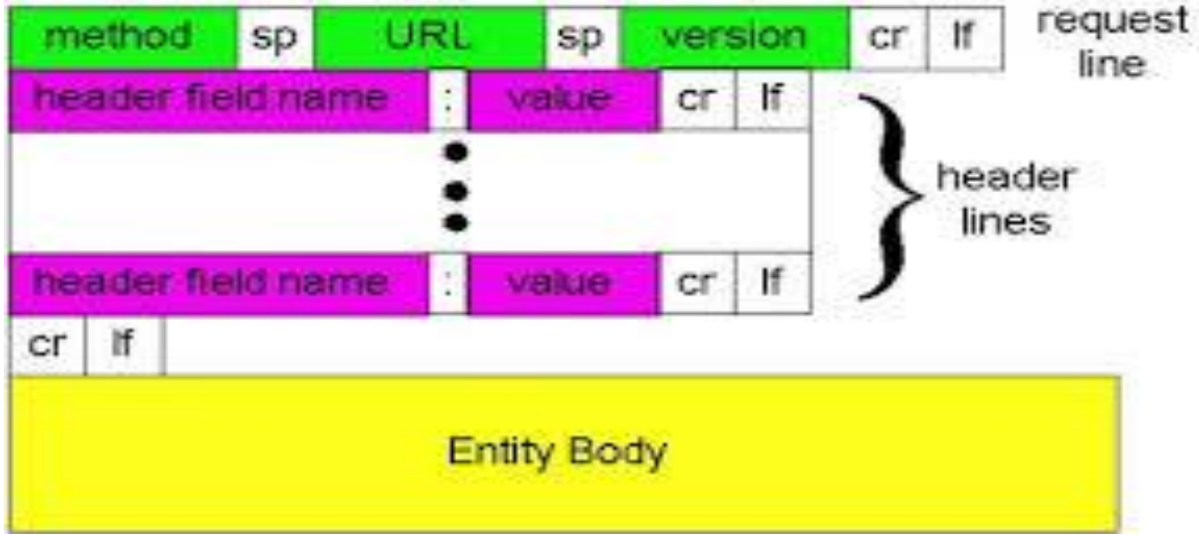
- Application level protocol to transfer hypertext documents on www
- This protocol follows the request-response model based on the client/server based architecture where web browsers, robots and search engines, etc. act like HTTP clients, and the Web server acts as a server.
- A client establishes a TCP connection to a HTTP server
- The client then send a request in the form of an ASCII string
- The server sends a reply as an ASCII string , binary sequence or other formats
- These messages are made up of plain text in a readable form
- HTTP Requests and Replies are treated independently by the client and server. So server does not maintain any specific state about each HTTP transaction.

HTTP request messages

Method	Description
GET	Request for resource from server
POST	Submit data to the server
HEAD	Same as GET but does not return the body
PUT	The data within the request must be stored at the URL supplied, replacing any existing data.
DELETE	Delete a resource
OPTIONS	Return the HTTP methods supported by the server
CONNECT	Client requests the HTTP proxy to forward a TCP connection to some destination. Used to create a TCP/IP tunnel for secure connections using HTTP proxies.

Structure of a request message

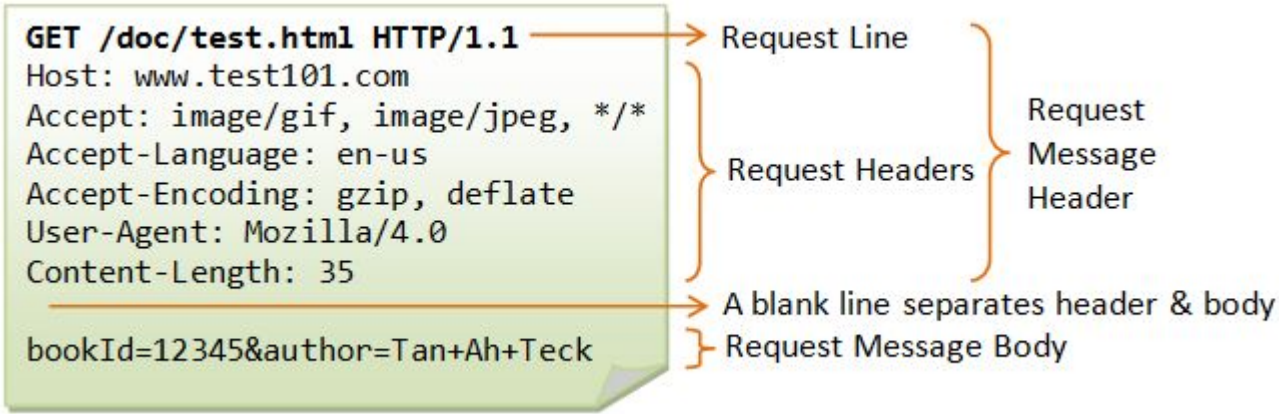
- An HTTP request message consists of a start line followed by a message header and optionally a message body.
- The start line and often the message header consists of printable ascii characters
- A blank line (CR LF) is inserted between the header and message body



Start line

- Every start line or request line consists of 3 parts with a space separating each of them
  - **Request method:** Should be written in uppercase
  - **Request-URI portion of web address** : The URL or URN of a web resource
  - **HTTP version:** At-present the version used is 1.1. The version string must be in uppercase
- An example for a start line :
  - GET /http://www.google.com HTTP/1.1

URN is a unique identifier for a resource rather than specifying the location of the resource.



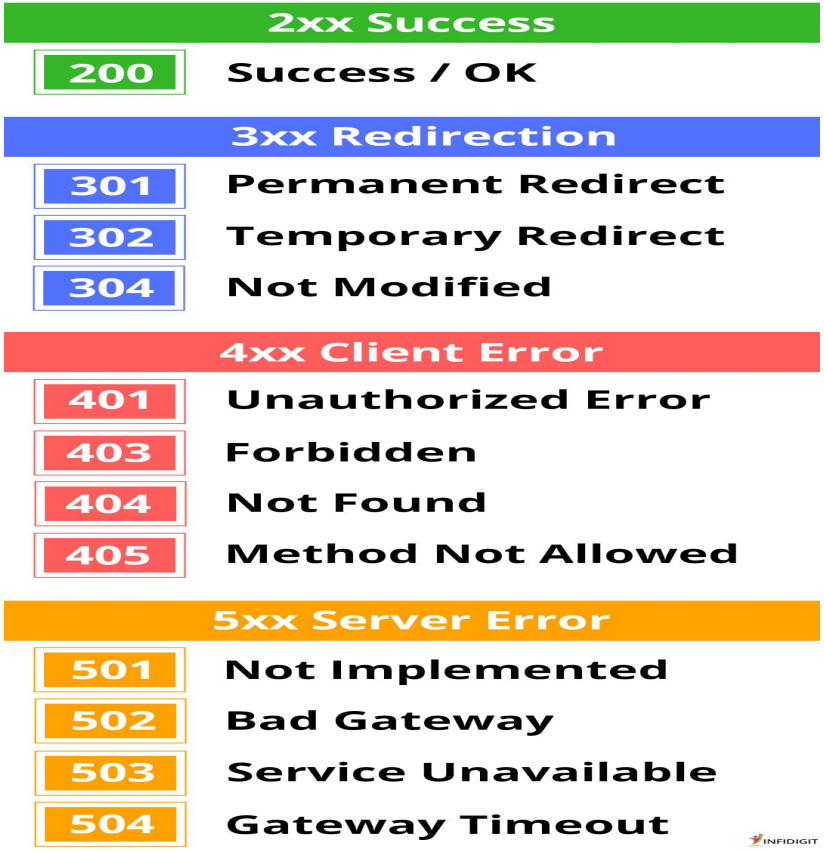
HTTP response

- An HTTP response message consists of a status line followed by a message header and optionally a message body.

Status line

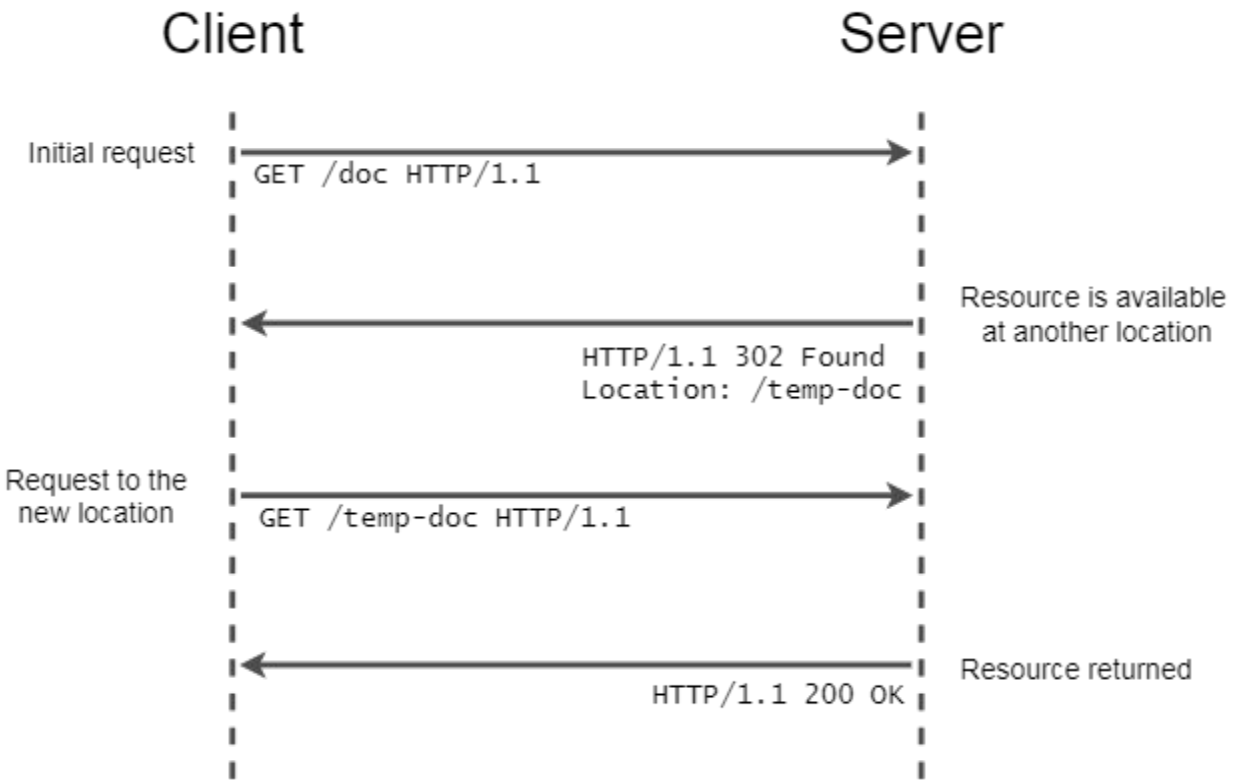
- The status line consists of the following fields
  - **HTTP version:** Version used by the server when formatting the response
  - **Status code :** Indicates the type of response as a 3 digit decimal numbers
  - **Reason phrase :** Text string that presents the information represented by the numeric status code
- An example
  - HTTP/1.1 200 OK
- The body of a response having 200 as status code should contain the resource requested by the client.

HTTP STATUS CODES

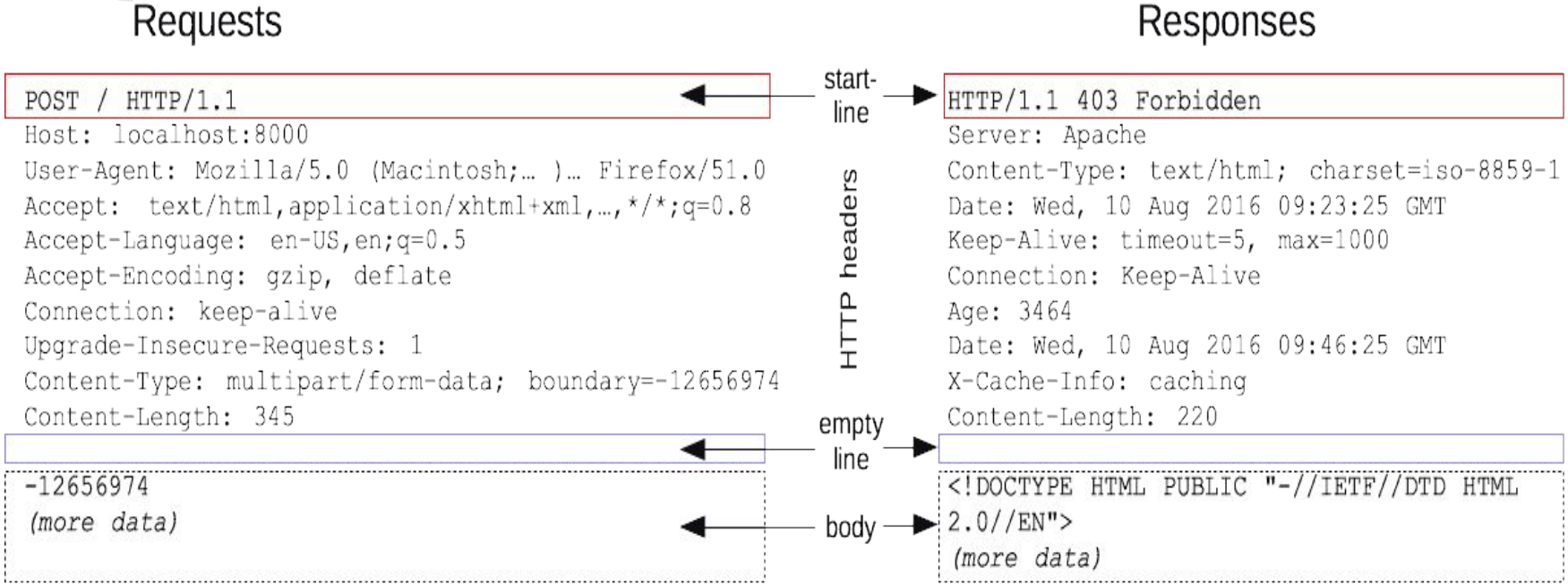


1xx Informational codes

Exchange of messages between client and server

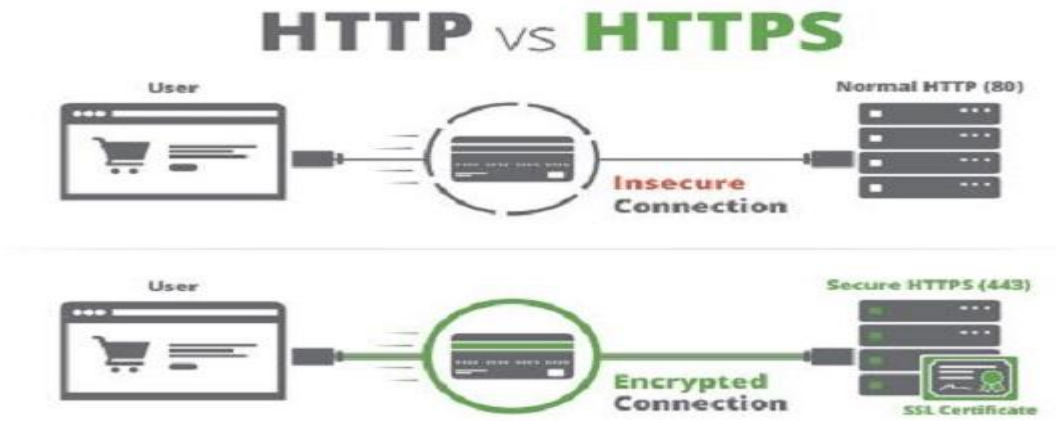


Request from a client and a server's response



HTTP over SSL or HTTPS

- HTTPS = HTTP + SSL i.e It is the result of layering the HTTP on top of SSL/TLS (Secure Socket Layer/Transport Layer Security).
- HTTPS (Hypertext Transfer Protocol over Secure Socket Layer) or HTTP over SSL was developed by Netscape.
- Used to exchange sensitive information which needs to be secured and to prevent unauthorized access.



Basic working of HTTPS

❑ difference between HTTP and HTTPs

S.NO	HTTP	HTTPs
1.	It is hypertext transfer protocol.	It is hypertext transfer protocol with secure.
2.	It is not secure & unreliable.	It is secure & reliable.
3.	HTTP URLs begin with http://.	HTTPs URLs begin with https://.
4.	It uses port 80 by default .	It was use port 443 by default.
5.	It is subject to man-in-the-middle & eavesdropping attacks.	It is designed to withstand such attacks & is considered secure against such attacks.

GENIUS

References

1. *"Web Technology: A Developer's Perspective"*, 2nd Edition, N P Gopalan, J Akilandeswari
2. *"Web Technologies: A Computer Science Perspective"*, Jeffrey C Jackson

NB: Images are adapted from Internet