WEB ESSENTIALS

Unit -I

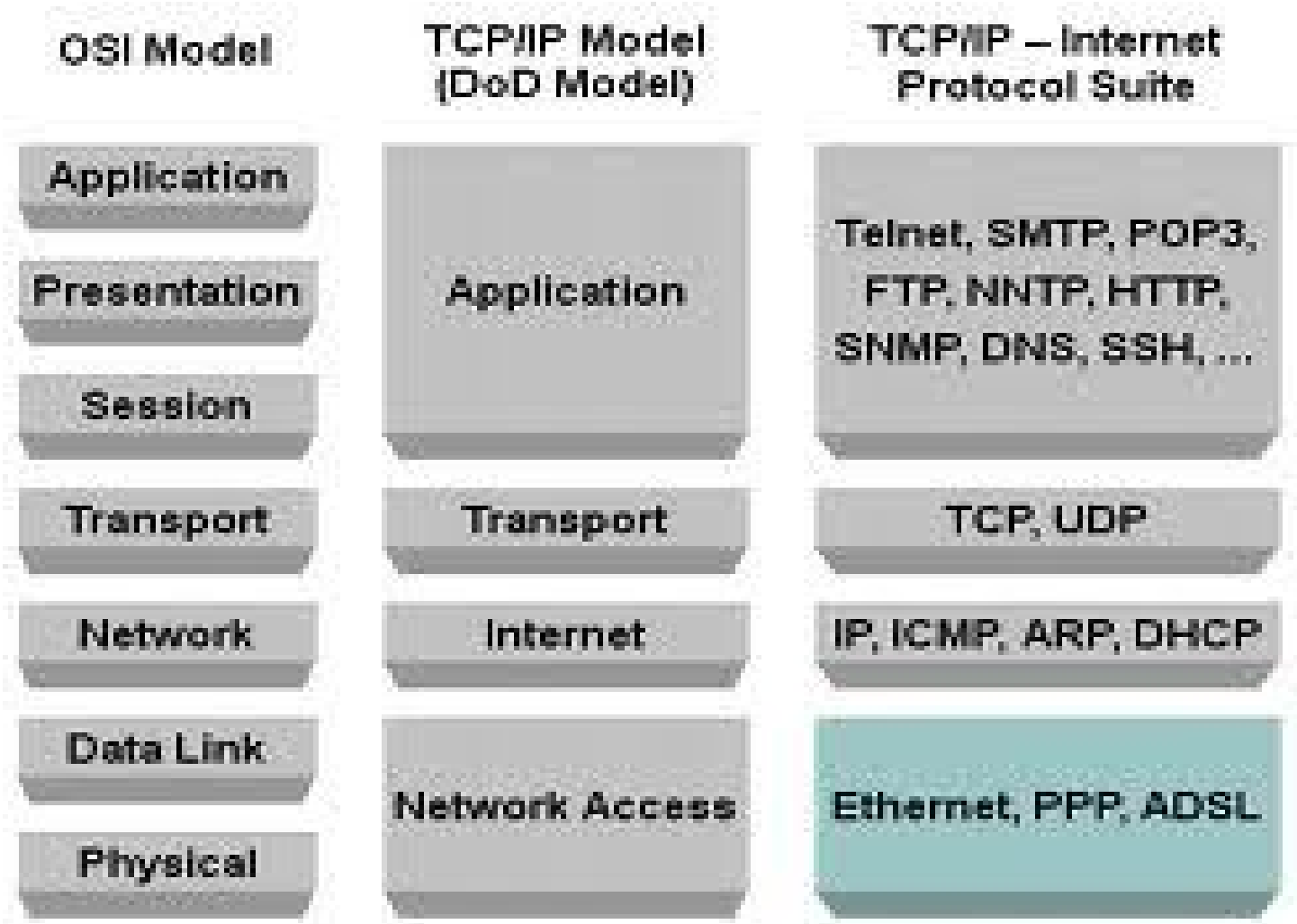
HISTORY OF INTERNET

- ARPA - Advanced Research Projects Agency
  - Root of internet
  - U.S Department of defense research project
  - This ARPANET project was intended for communication between geographically dispersed computers from diff manufacturers running diff OS
  - Emergence of TCP/IP protocols lead to communication between different networks
  - Initially used for military purposes and limited to universities and research institutions
  - Later decided to use internet for commercial purposes

INTERNET PROTOCOLS

- Protocol – agreed formats for communicating between two computers
- TCP/IP
- HTTP
- FTP
- SMTP

OSI - TCP/IP


OSI Model
TCP/IP Model (DoD Model)
TCP/IP – Internet Protocol Suite
Application
Presentation
Session
Transport
Network
Data Link
Physical
Application
Transport
Internet
Network Access
Telnet, SMTP, POP3, FTP, NNTP, HTTP, SNMP, DNS, SSH, ...
TCP, UDP
IP, ICMP, ARP, DHCP
Ethernet, PPP, ADSL

NETWORK LAYER

- IP – Internet Protocol
  - IP Address – 32-bit number
  - Each device on the Internet has one or more IP address
  - Function of IP is to transfer data from one computer to another computer
  - Error detection
  - Corrupted packets are discarded- hence unreliable : packets get lost

TRANSPORT LAYER PROTOCOLS

- Transmission Control Protocol – TCP
- User Datagram Protocol - UDP

TRANSMISSION CONTROL PROTOCOL (TCP)

- Reliable Communication
- Connection Establishment
- Full duplex communication
- Port – Used for communicating with many diff applications on a machine

USER DATAGRAM PROTOCOL(UDP)

- An alternative to TCP which is also built over IP
- Unreliable communication.
- Two way communication is not possible.
- Advantage – Speed over TCP
- Generally used for sending short messages which expect short responses (eg. DNS)

APPLICATION LAYER PROTOCOLS

- SMTP – Simple Mail Transfer Protocol
  - Transfer of email between different email servers
- FTP – File Transfer Protocol
  - Transfer files between machines
- Telnet - Execute commands types into one computer on a remote computer
- Http – Communicating between web servers and clients

WORLD WIDE WEB

- Internet - Sharing of information
- WWW is one of the technologies used for managing the information over internet
- Two types of software
  - Server – that provides information
  - Client - that access information
- This server and client communicate over Internet using a protocol built on top of TCP/IP – which is HTTP
- Client requests a document and Server returns the requested document

WORLD WIDE WEB

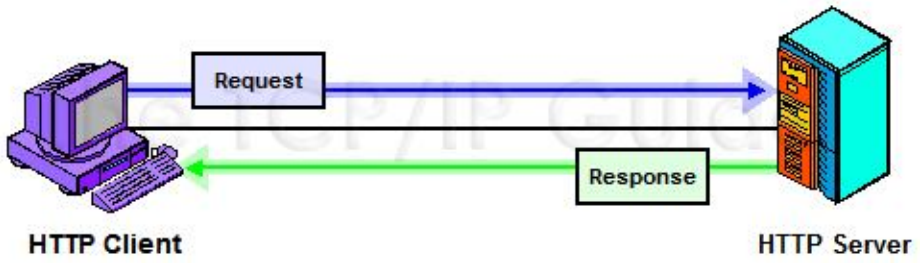
- Documents are generally web pages which are linked to other documents on the Web
- Internet – collection of machines connected over IP
- World Wide Web – collection of machines on the internet that provide information via HTTP

WEB

- Web 1.0 – Information Web
- Web 2.0 – Interactive or Social web
- Web 3.0 – Intelligent or Semantic Web

HYPERTEXT TRANSFER PROTOCOL (HTTP)

- It is an application layer protocol
- It specifies how web clients and server communicate
- Follows a *request-response model*



- HTTP communication consists of an HTTP Request message sent by a client to a server, which replies with an HTTP Response

HTTP REQUEST

- The HTTP client sends a request message formatted according to the rules of the HTTP standard—an *HTTP Request*.
- This message specifies the resource that the client wishes to retrieve, or includes information to be provided to the server.
- Message structure
  - Start line
  - Header field(s)
  - Message body

HTTP REQUEST (START LINE)

- GET /path/servlet.jsp HTTP/1.1
  - Req method — Req-URI — Protocol/Version

- Request Method
  - GET – return the response as specified by Req-URI
  - POST – return the response based on the data sent through request
  - HEAD –same as GET ,but without message body in response

HTTP REQUEST (START LINE)

- OPTIONS – return list of methods that can be used to access the resource
- PUT – store the information(msg body of the req) in the server
- DELETE – delete the resource identifed by the Req-URI
- TRACE – return a copy of the complete http req msg.(testing purposes)

HTTP REQUEST (START LINE)

- REQUEST-URI
  - http:// + value of Host header field + Request-URI = Uniform Resource Identifier(URI)
  - An URI is an identifier that is intended to be associated with a particular resource on WWW.
- HTTP Protocol/Version – HTTP/1.1
  - Initial version – 0.9
  - Second version – 1.0
  - Current version - 1.1

HTTP REQUEST (HEADER FIELD)

- This section of an HTTP Request contains the request headers, which are used to communicate information about the client environment.
- Few of these headers are: Content-Type, User-Agent, Accept-Encoding, Content-Length, Accept-Language, Host, etc. Field Name : Field value

HTTP REQUEST (HEADER FIELD)

- GET /path/index.html HTTP/1.1
- Proxy-Connection: Keep-Alive
- User-Agent: Mozilla/5.0 [en] (X11; I; Linux 2.2.3 i686)
- Host: www.vit.ac.in
- Accept: image/gif, image/x-xbitmap, image/jpeg, image/pjpeg, image/png
- Accept-Encoding: gzip
- Accept-Language: en
- Accept-Charset: iso-8859-1, *, utf-8

HTTP RESPONSE

- The server reads and interprets the request. It takes action relevant to the request and creates an *HTTP Response* message, which it sends back to the client.
- The response message indicates whether the request was successful, and may also contain the content of the resource that the client requested, if appropriate.
- Response structure
  - Status line
  - Header field(s)
  - Message body

HTTP RESPONSE (STATUS LINE)

- HTTP/1.1 200 OK

  - Version
  - Response Code
  - Reason Phrase

- Response Code
  - 1xx- Informational
  - 2xx –Success
  - 3xx- Redirect
  - 4xx –Client error
  - 5xx – Server error

HTTP RESPONSE (STATUS LINE)

- Some of the status codes with reason phrase
  - 200 OK
  - 301 Moved Permanently
  - 307 Temporary Redirect
  - 401 Unauthorized
  - 404 Not found
  - 500 Internal Server Error

HTTP RESPONSE (STATUS LINE)

- HTTP/1.0 200 OK
- Date: Fri, 13 Nov 2012 06:57:43 GMT
- Content-Location: http://www.vit.ac.in/site.asp
- Etag: "07db14afa76be1:1074"
- Last-Modified: Thu, 05 Nov 2009 20:01:38 GMT
- Content-Length: 7931
- Content-Type: text/html
- Server: Microsoft-IIS/4.0
- Age: 922
- Proxy-Connection: close

A SIMPLE TRANSACTION