

Subject Name: Introduction to Web Technology

Module Number: 1

Dr. Amrendra Singh Yadav

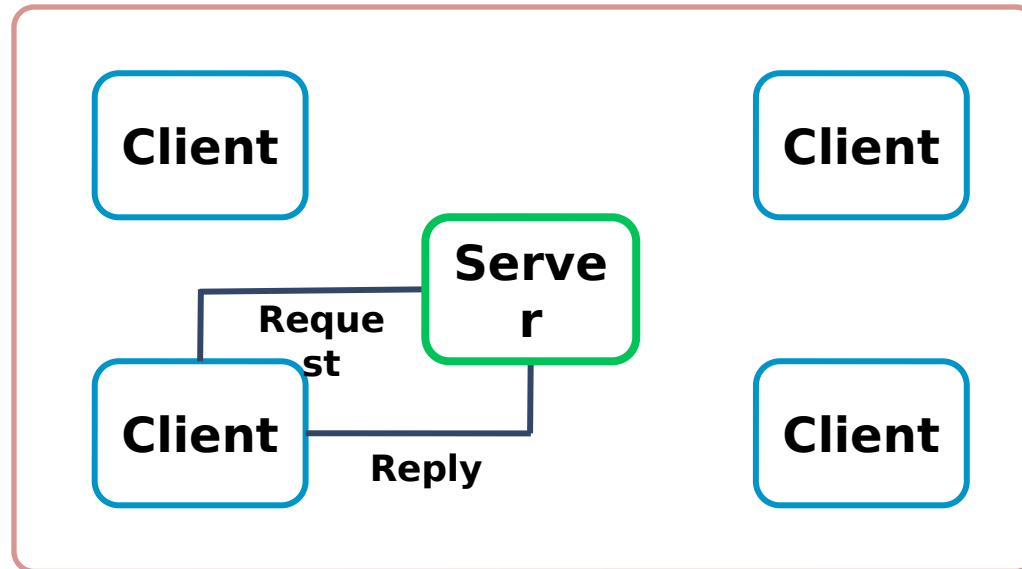
CONTENTS

- Introduction to Internet
- Internet Design Principles
- Internet Protocols
- World Wide Web
- Hyper Text Transfer Protocol
- Web Technologies and Web Services

Client and Server

Client – hardware/software that accesses a service made available by a server.

Server – hardware/software that provides services to other users.

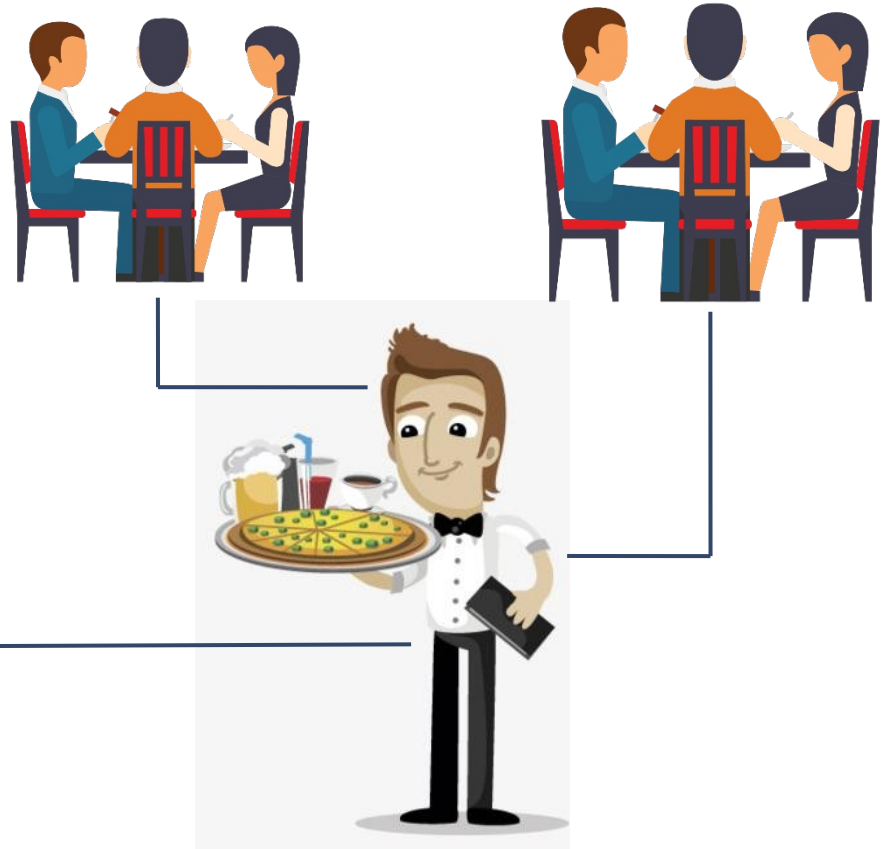


Client Server Analogy

Server is like waiter

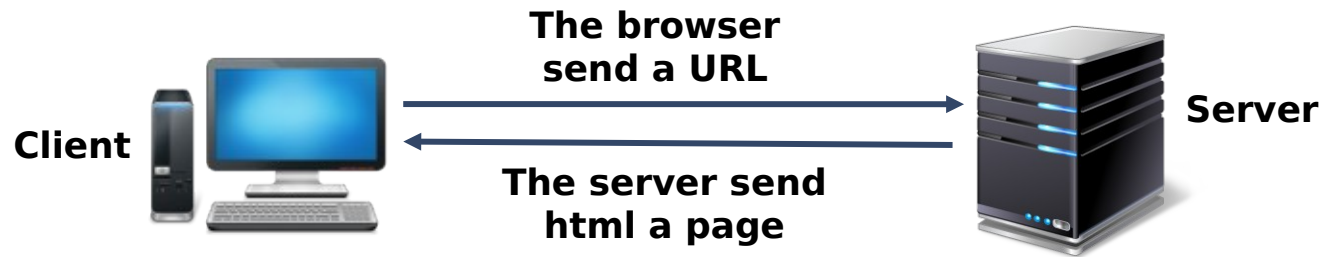
Clients are like customers

- Asks the server for food
- Not allowed in to the kitchen



Client Server Communication

The process of establishing a connection between a client and a server is called as **Client-Server Communication**



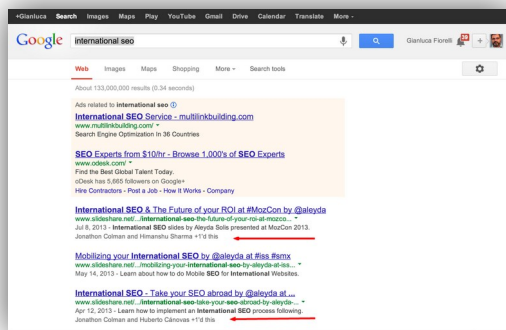
The Internet uses the Client-Server model.



Have you sent an email to a friend living in United States?



Have you connected with friends all over the world on Facebook?



Have you searched something on the internet?

How is it possible for you to communicate with a friend who might be on a different continent?

You and your friend both need to be connected to the internet to be able to do all these things.

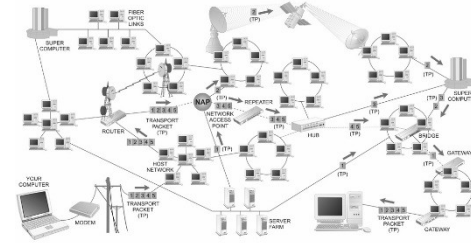
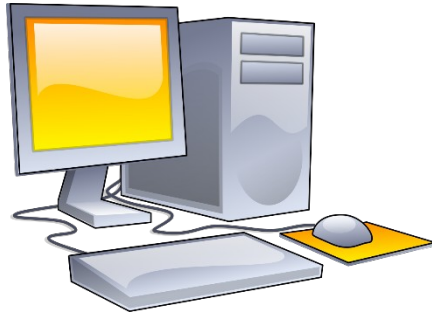


So What is Internet?

Internet is a global network of **interconnected networks and devices** that exchange information.

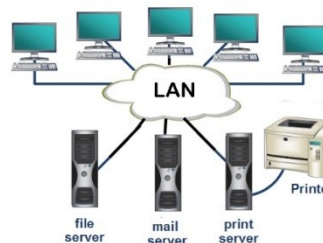


Evolution of Internet



- **Standalone Computers**
 - When computers were first created they were stand alone.
- **Networks**
 - **The need for sharing resources among the computers** led to connecting computers into a network.
- **Inter-network**
 - These networks were then connected together to create what is today called the ‘Internet’.

Need for sharing resources led to the creation of networks and ultimately the ‘Internet’.



Internet – How it came to be?

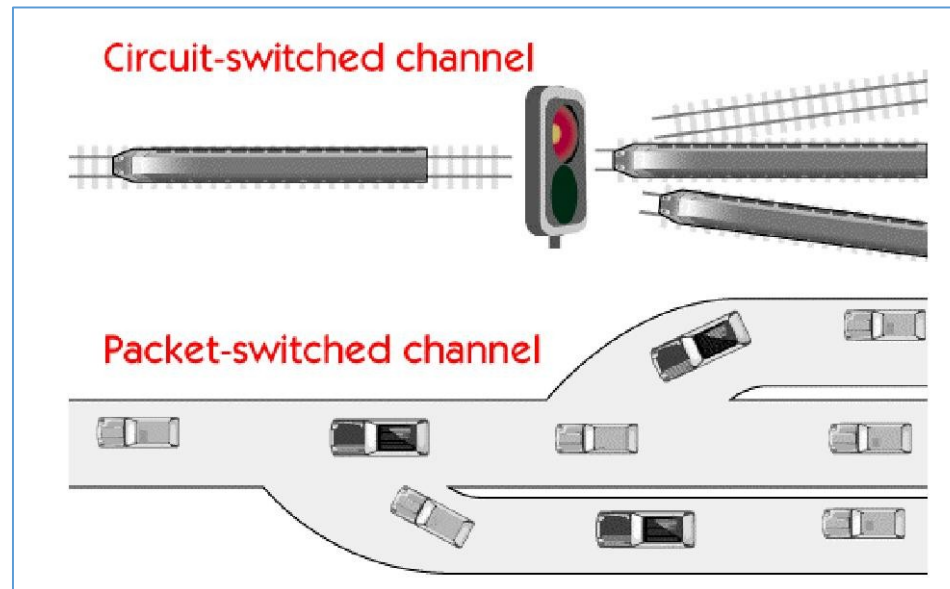
History of Internet

- The Advanced Research Projects Agency Network or ARPANET funded by the United States Department of Defense in **1969**.
- In **1971** Ray Tomlinson developed a system to send electronic mail.
- Dr. Robert M. Metcalfe developed **Ethernet** in **1973**.
- In the **early 80s** the, the current versions of Transmission Control Protocol (TCP) and IP (Internet Protocol) were developed.
- In **1992**, The Computer Emergency Rescue Team (CERT) released the first version of the “World Wide Web”.
- In **1996**, IPV6 – Internet Protocol Version 6 was created.
- In **1999**, a wireless technology called 802.11b or popularly known as “Wi-Fi” was standardized.

History of Internet

The Beginning

- In 1964, when **Packet Switching** was invented it opened the doors for computers to connect and communicate.
- Unlike **Circuit Switching**, which required a dedicated channel for the duration of transmission, packet switching divided the data into small parts called ‘packets’ which were then transmitted over the network.



History of Internet

Advanced Research project Agency Network (ARPANET) – ancestor of today's Internet.

- The American Defense Advance Research Projects Agency (DARPA) was ordered by the American government to create a network spanning the United States so that information exchange could be ensured during a potential atomic attack by USSR.
- The network that was developed was called the '*ARPANET*' short for Advanced Research project Agency Network.
- It was a basic network based on packet switching and providing messaging services for sharing information. By 1977, over a 100 nodes were connected through ARPANET.

History of Internet

Electronic Mail – the initial networking application

- In 1972, the initial network application – the electronic mail was introduced.
- March Ray Tomlinson wrote the basic email software motivated by the need of the ARPANET users for an easy and better coordination mechanism.
- Email remained the largest network application for over a decade.

The commonly used email protocols are :

1. **Simple Mail Transfer Protocol (SMTP)** used for sending email messages.
2. **Post Office Protocol (POP)** used for retrieving email messages.

History of Internet

Ethernet – the network technology is born

- Developed by Bob Metcalfe at Xerox PARC in 1973, and is now a dominant network technology in the Internet.
- Ethernet is a relatively low cost, reasonably fast and very popular LAN technology.

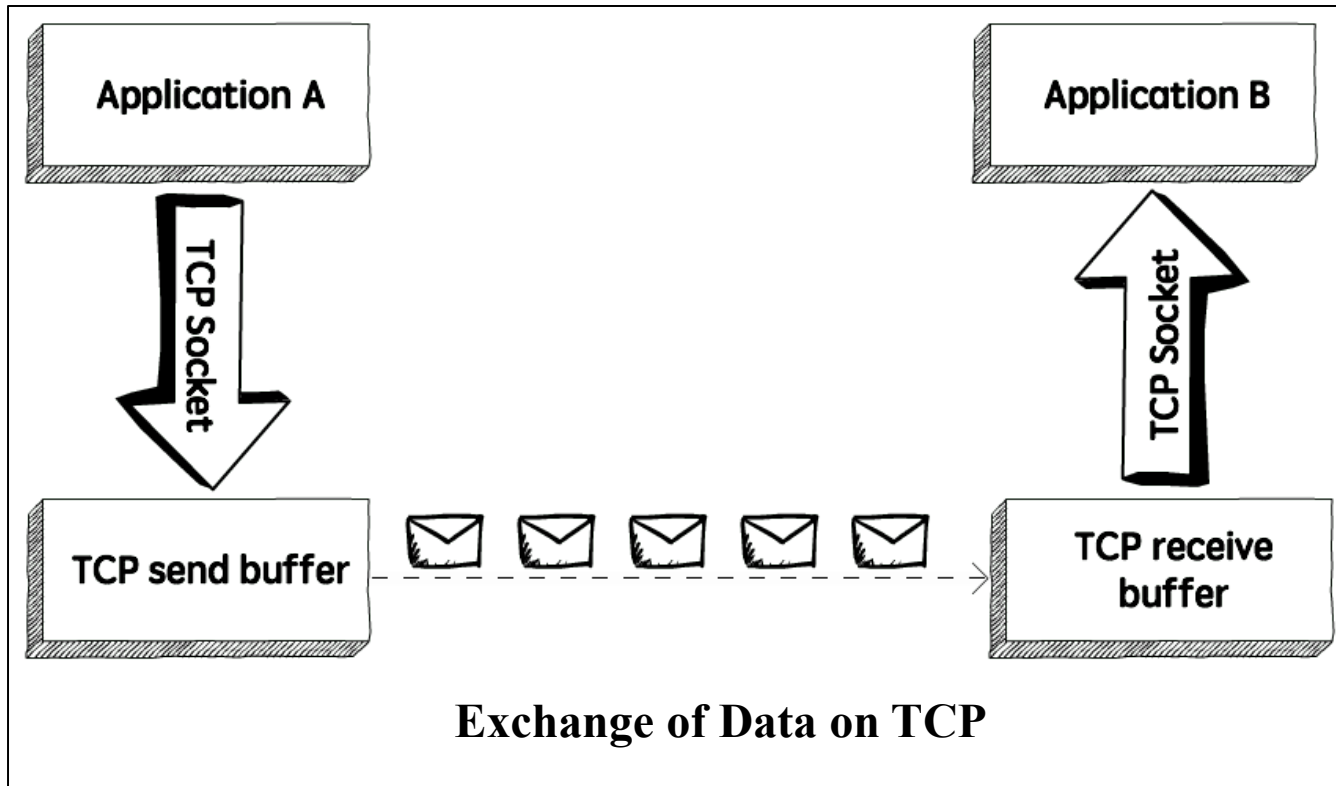


Ethernet Cable

History of Internet

ARPANET adopts Transmission Control Protocol (TCP/IP)

In 1974, ARPANET adopted the TCP/IP protocol created by Robert Kahn and Vint Cerf. By mid 70s.



The Internet emerges!

The original **ARPANET** grew into the ‘internet’.

“The internet design philosophy has changed considerably from the first proposal to the current standards.”

- Dave Clark

Design principles of the internet

1. Multiplexed utilization of existing interconnected networks

Sharing a single communication channel.

2. Survivability in the face of failure

The ability of the internet to keep up the communication services despite loss of networks or gateways.

3. Types of Services

The internet must support multiple types of services.

4. Support for a variety of Network

The internet architecture must accommodate and provide a way for different networks like Local Area Networks, home networks and business networks to connect with each other and exchange information.

Design principles of the internet

5. Distributed management

The critical resources are managed in a distributed fashion by various entities to ensure the fundamental right to freedom is not violated.

6. Cost effectiveness

Making this technology affordable will ensure it is accessible to all and avoid info-exclusion.

7. Host attachment

connecting a device to the internet should be as effortless as possible to ensure less overhead and better experience.

8. Accountability

Accountability is not intrinsic to the current internet architecture, migration towards the accountability framework is ongoing through **Host Identity Protocol (HIP)** and **Accountable Internet Protocol (AIP)**.

What is a Protocol ?

- A protocol can be thought of as an **‘accepted standard’**.
- A protocol is a set of rules that the computers use over a network when they communicate.
- These are rules accepted and followed by the network to successfully transmit data.
- There are a host of protocols which allow users to communicate over the internet. Some of the important ones will be discussed in this chapter.

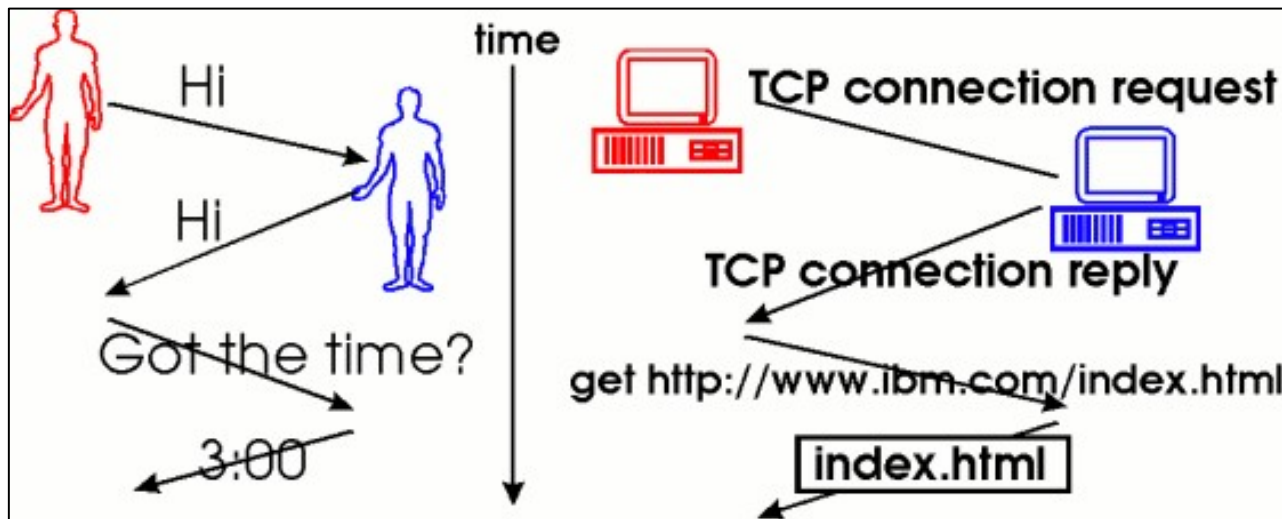


Diagram: Human protocol and a computer network protocol

Transmission Control Protocol (TCP)

The Transmission Control Protocol provides reliable transmission of data in an IP environment.

Among the services it provides are:

- **Stream data transfer**
- **Reliability**
- **Efficient flow control**
- **Full-duplex operation**
- **Multiplexing**

User Datagram Protocol (UDP)

- The User Datagram Protocol is an alternative communication protocol to TCP.
- It is a **connection-less protocol** that does not provide reliability, order or error-checking, but is much faster than TCP.
- UDP is more suitable for real-time applications like video calls.

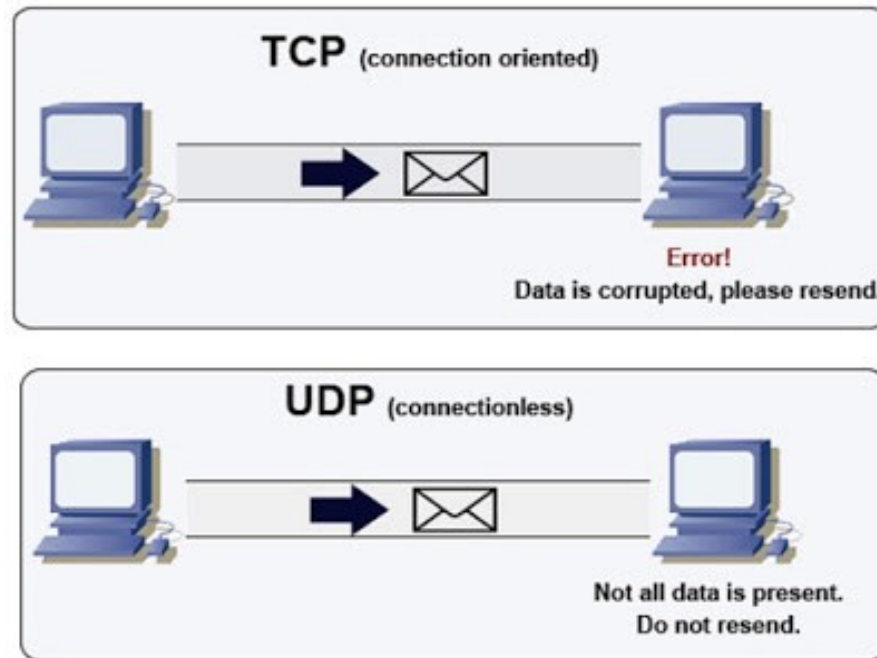


Diagram: TCP vs. UDP

Internet Protocol (IP)

Internet Protocol (IP) has the task of delivering packets from the source host to the destination host solely based on the IP addresses in the packet headers.

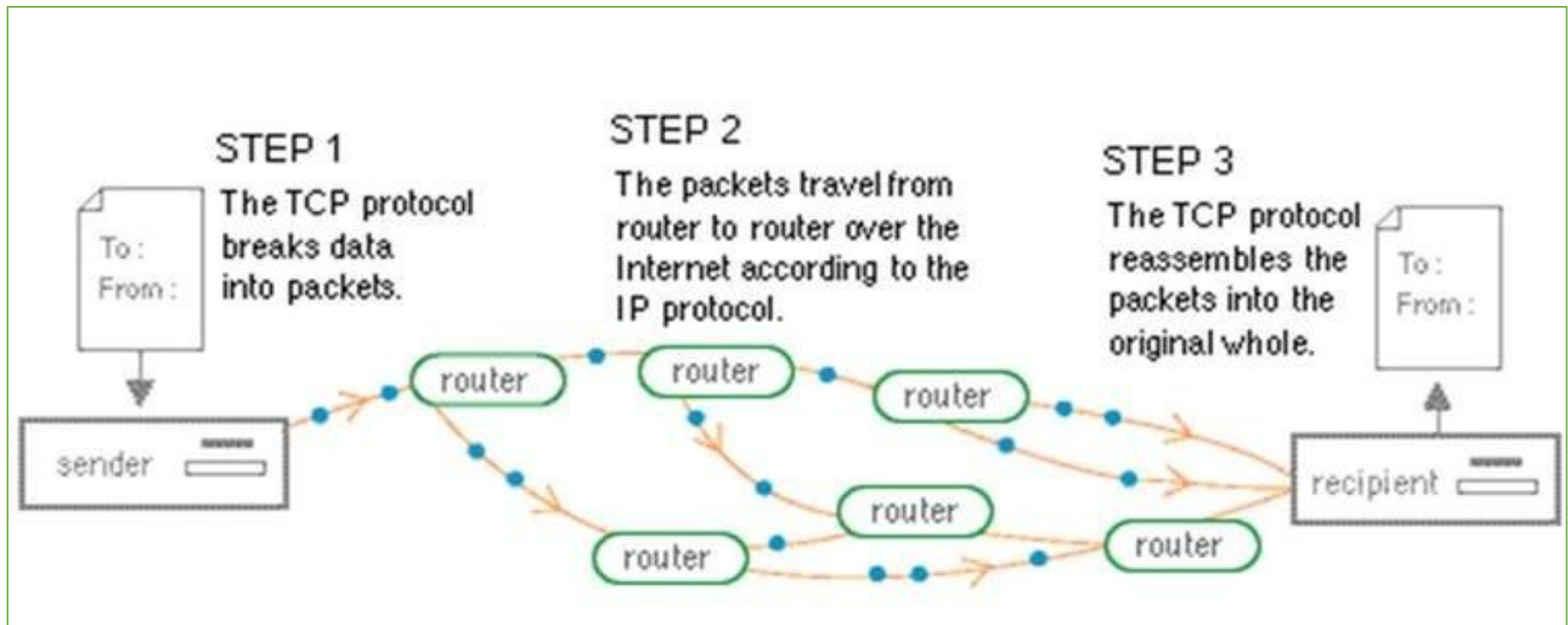


Diagram: The Internet Protocol

IP Address

In the same sense that someone needs your mailing address to send you a letter, a remote computer needs your IP address to communicate with your computer.

There are two standards for IP addresses:

- **IP Version 4 (IPv4)**
- **IP Version 6 (IPv6)**

IPv4

- An IPv4 address is expressed by four numbers separated by dots.
- Each number is the decimal (base-10) representation for an eight-digit binary (base-2) number, also called an **octet**.

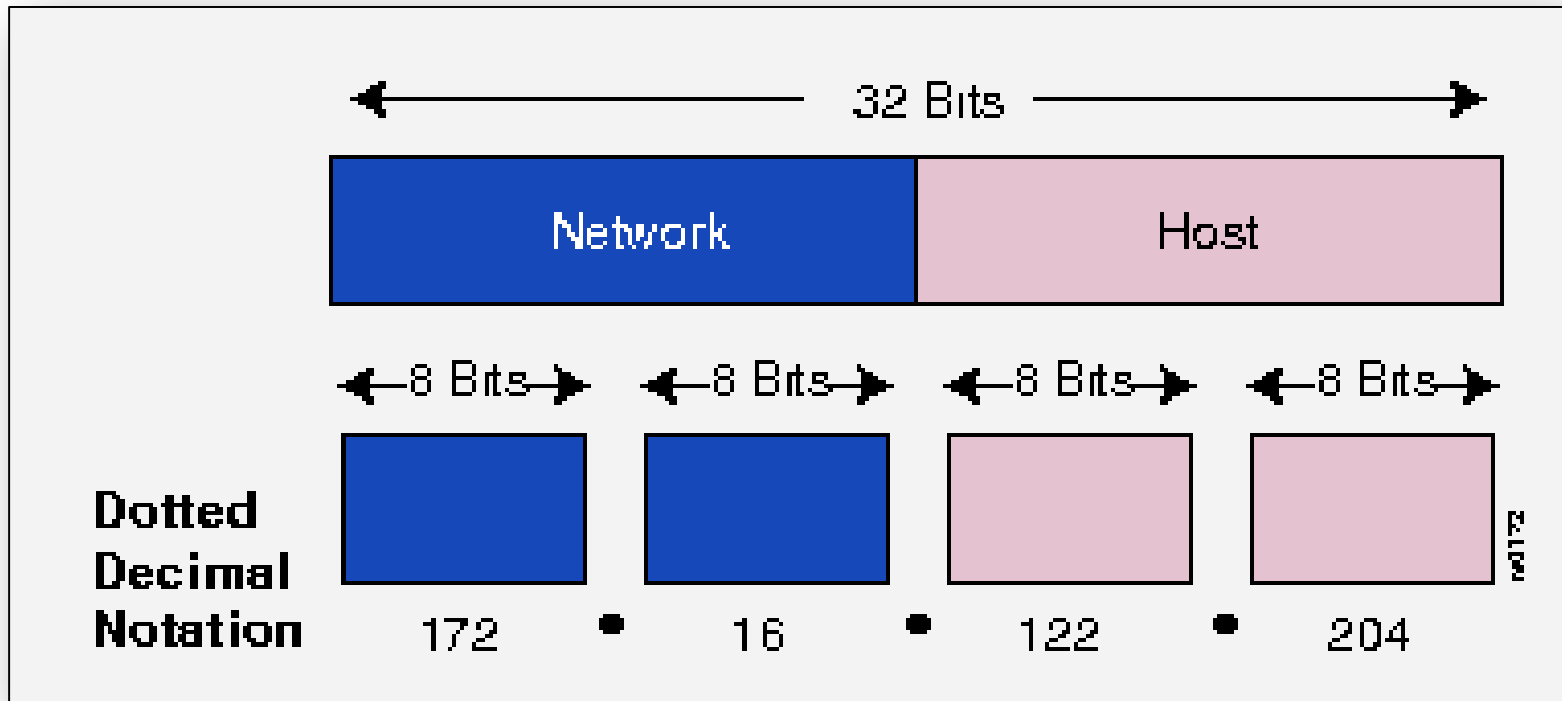


Diagram: IPV4 Address Format

IPV6

- The explosive growth in mobile devices including mobile phones, notebooks and other wireless handheld devices created the need for additional blocks of IP addresses.
- IPv6 uses **128 bit addresses** and offers 2^{128} or approximately 3.4×10^{38} addresses.

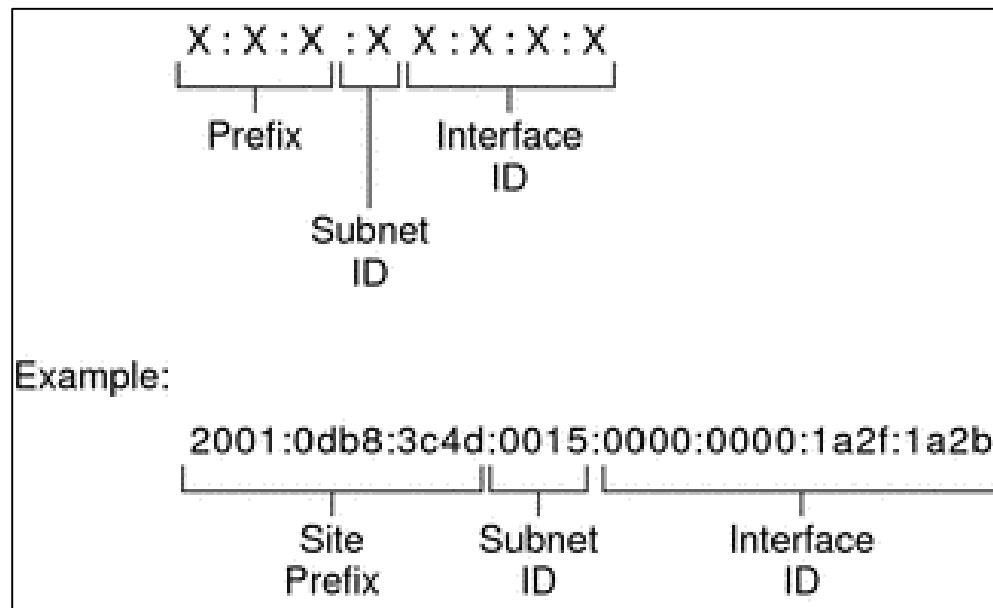


Diagram: IPV6 Address Format

Name or Number!
which is easier to remember?

Domain Name System

- “Domain Name is the name given for any website” (ex: yahoo.com).
- “Domain Name System (DNS) is like a phone directory which stores all the domain names across the web“.

Domain Name Server

Domain Name Server(**DNS**) maps a domain name with its corresponding IP address.

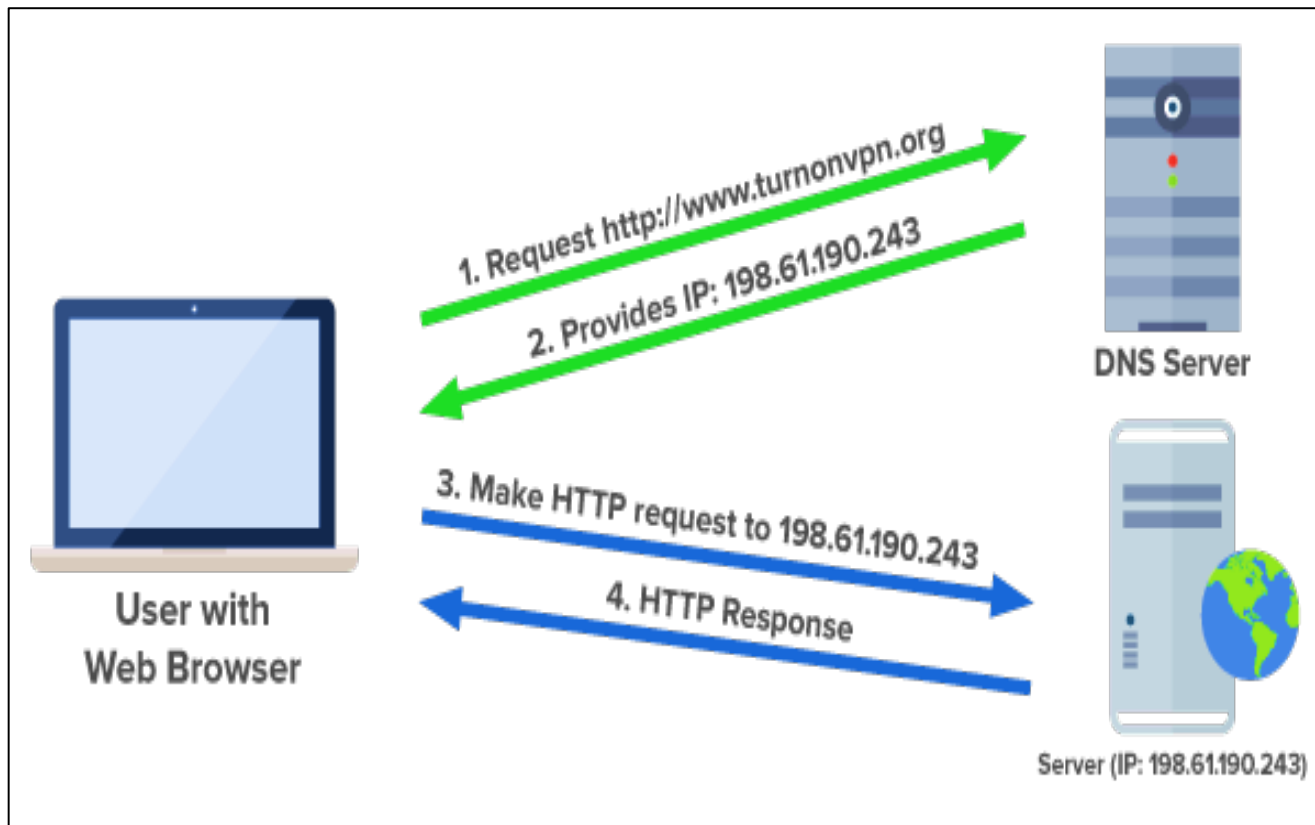


Diagram: Domain Name Server

File Transfer Protocol (FTP)

- The File Transfer Protocol is used to transfer files from one host computer to the other over a TCP/IP network. It creates two different connections – a **control connection** and a **data connection**.
- The control connection is used for sending control information and the data connection is used for sending the actual file.
- Port 21 is used for control connection whereas data connection uses port 20.

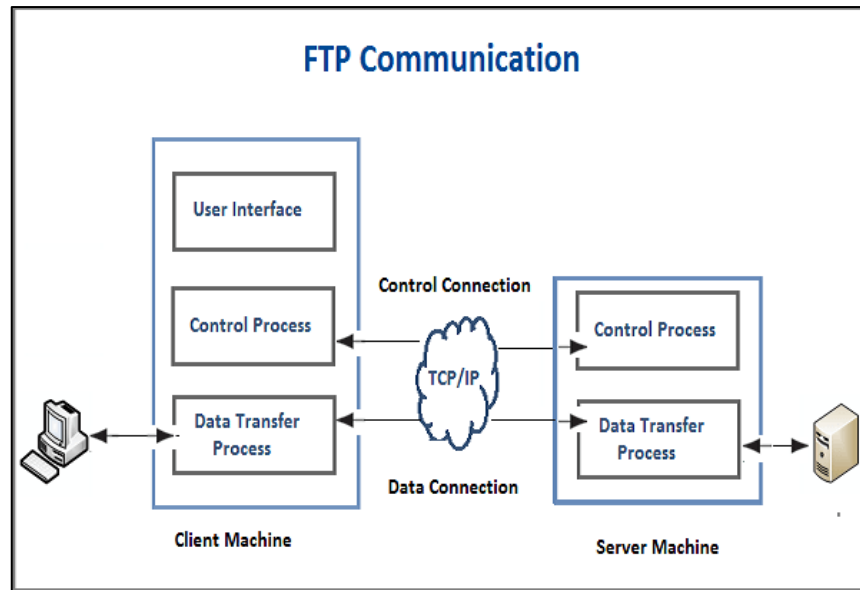


Diagram: File Transfer protocol

SMTP (Simple Mail Transfer Protocol)

- The Simple Mail Transfer Protocol is used for sending e-mail messages between servers.
- It uses the port 25.
- SMTP because it is limited in its ability to queue messages at the receiving end, it is used with one of two other protocols – POP (Post Office Protocol) or IMAP (Internet Message Access Protocol) which allow the user to save the messages in a mailbox on the server and retrieve them when they want.

POP3 (Post Office Protocol)

- This protocol is used by local e-mail clients to **retrieve e-mail** from the server.
- It uses TCP port 110 to establish connection with the server and then it sends a request for a particular mailbox using user name and password.
- The user can then retrieve the list of mail messages one by one. POP3 has two modes - Delete and keep. In the delete mode, mail is deleted from mailbox and in the keep mode, mail resides in the mailbox.

SMTP and POP3 Protocols

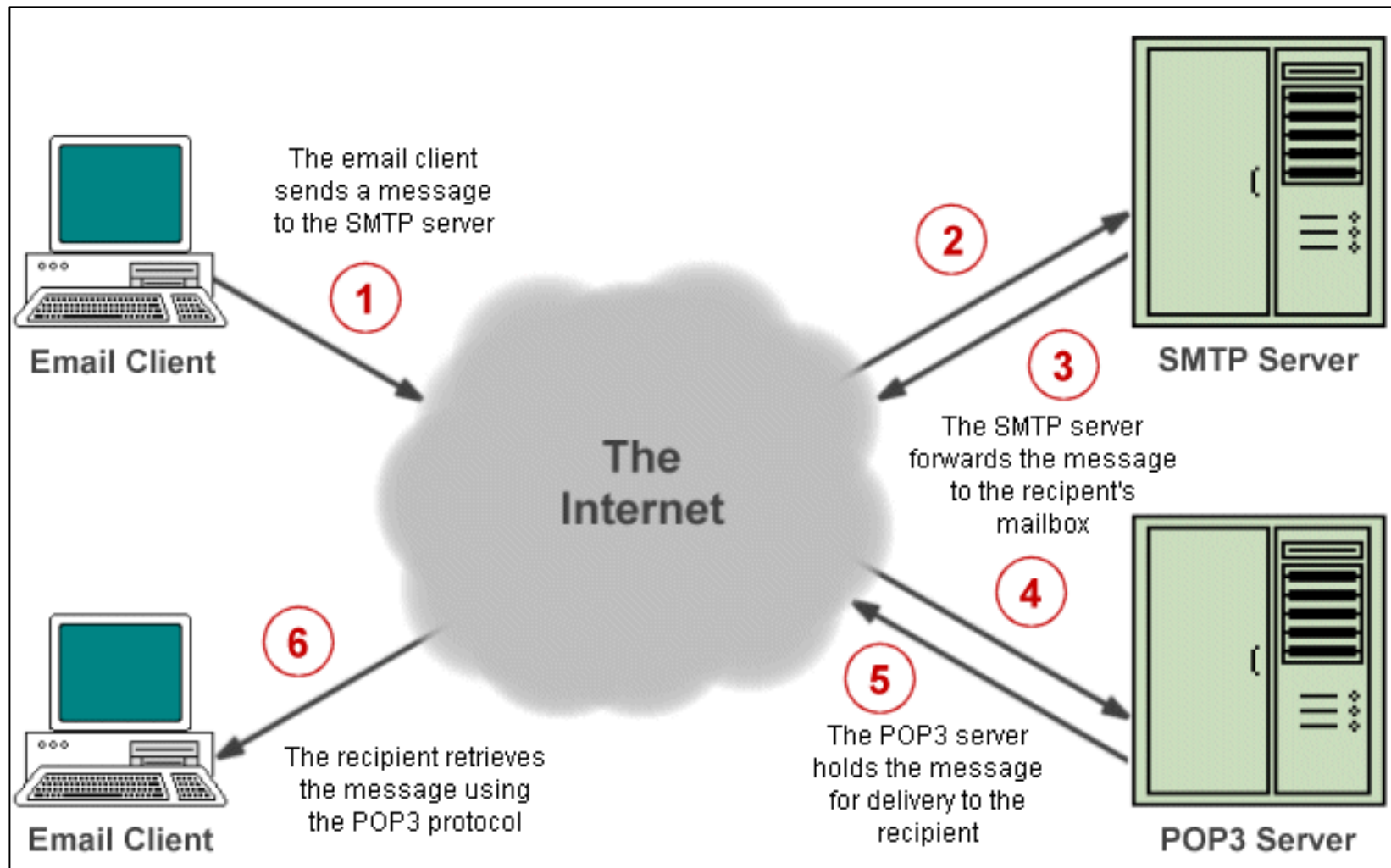


Diagram: SMTP and POP3 protocols

Interactive Mail Access Protocol (IMAP)

- By default, messages that have been downloaded to a recipient's computer are deleted from the mail server.
- The **Interactive Mail Access Protocol (IMAP)** is a more sophisticated mail protocol that stores all incoming and outgoing mail on the server so that mail clients with mailboxes on the server can access their e-mail from anywhere.
- Mail is not downloaded to the user's PC, and is only deleted from the client's mailbox if the client specifies that it is to be deleted.

TELNET (Telecommunications Network)

- Telnet is a remote login protocol.
- The basic purpose of Telnet is to provide a means by which keyboard commands typed by a remote user can cross the network and become input for a different computer. Screen output related to the session then crosses the network from that different computer (the server) to the client system.
- However, because of security issues its use over public internet should be avoided.

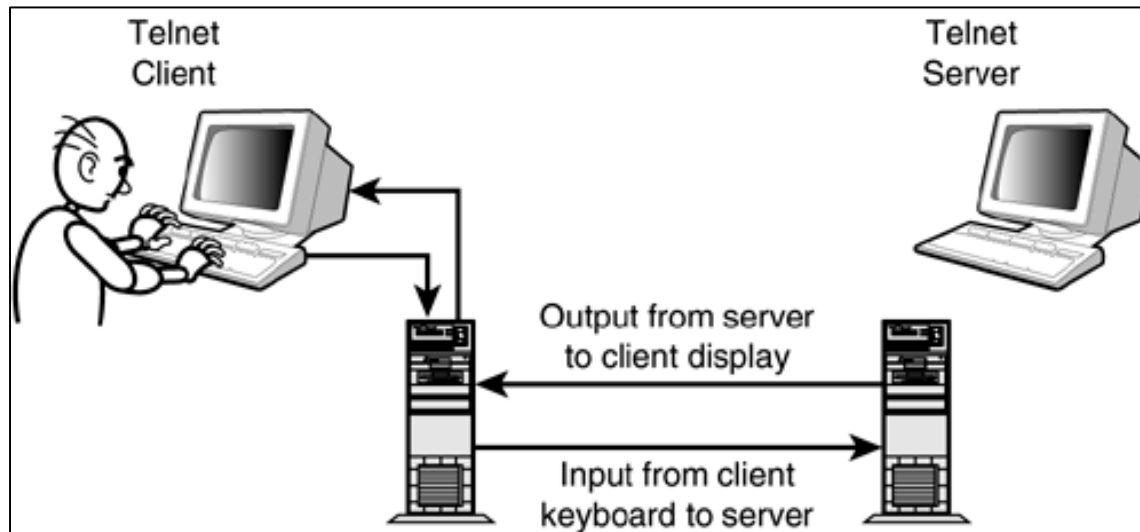
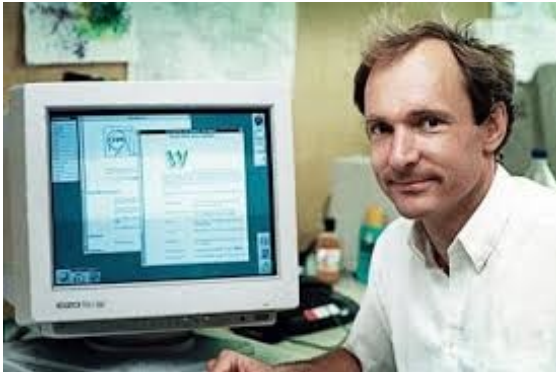


Diagram: Telnet Protocol

Is the Internet same as the World Wide Web?

WHY?

The web was born at **CERN** in Geneva in 1990, to allow its researchers who were geographically dispersed to share documents using a hypertext system.



Tim Berners Lee
Father of World Wide
Web

The web is not the same as the internet, it is a service provided by the internet.

**The World Wide Web is a
system of interlinked hypertext
documents accessed via the
internet.**

History of the Web

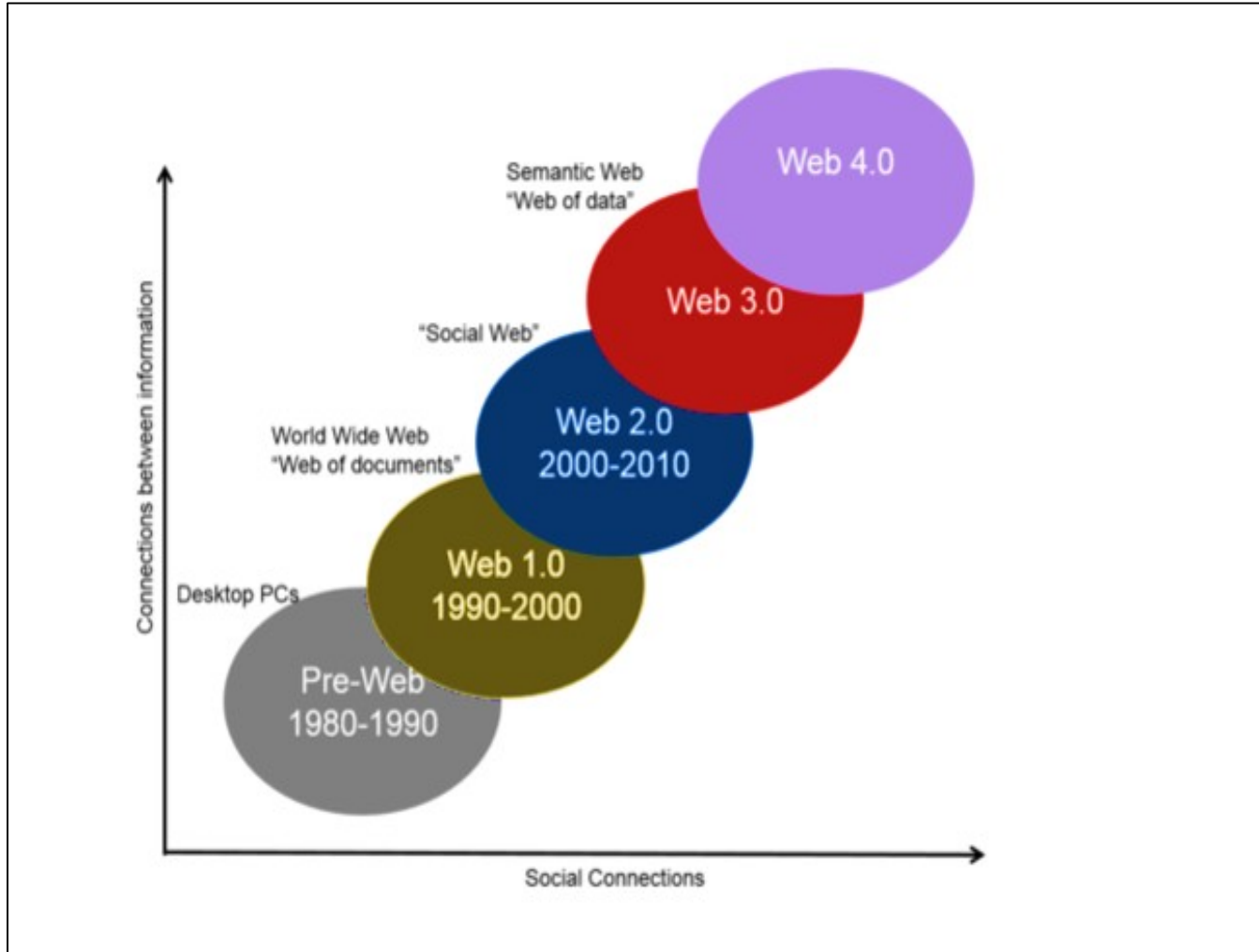


Diagram: World Wide Web Timeline

Web 1.0

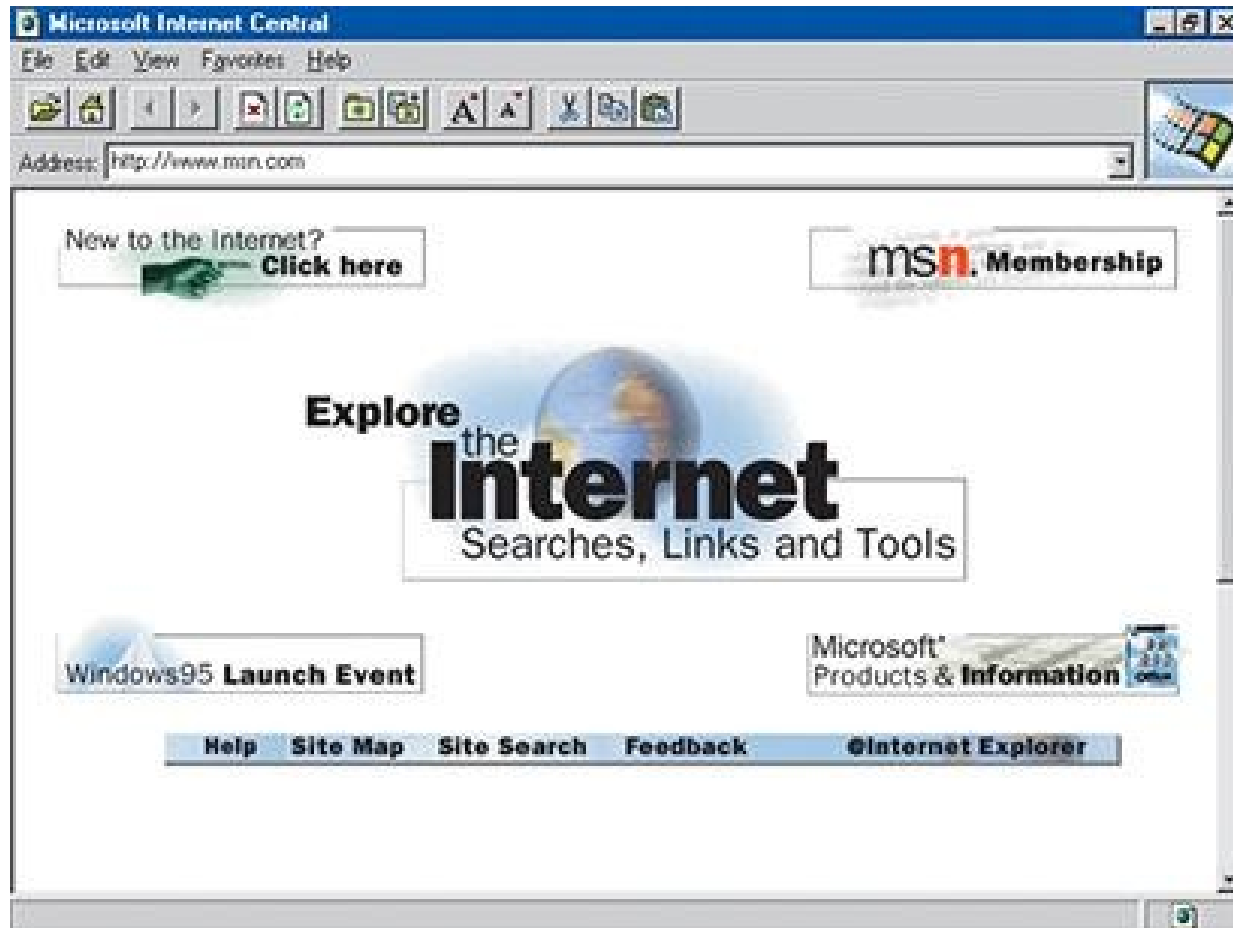


Diagram: Internet Explorer web page from 1995

Web 2.0



Web 3.0

SEMANTIC WEB

- ✓ Semantic Web concepts bring meaning to the Web.
- ✓ Data is represented in a way such that, machines can understand it.
- ✓ Data to Smart Data.

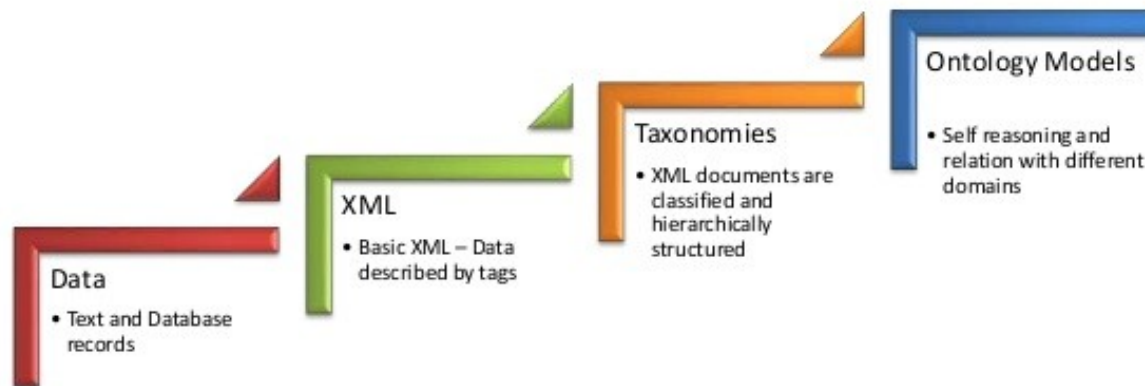


Diagram: The Semantic Web


Web 1.0 / 2.0 / 3.0 Summary

Crawl	Walk	Run
Web 1.0	Web 2.0	Web 3.0
Mostly Read-Only	Wildly Read-Write	Portable & Personal
Company Focus	Community Focus	Individual Focus
Home Pages	Blogs / Wikis	Lifestreams / Waves
Owning Content	Sharing Content	Consolidating Content
Web Forms	Web Applications	Smart Applications
Directories	Tagging	User Behavior
Page Views	Cost Per Click	User Engagement
Banner Advertising	Interactive Advertising	Behavioral Advertising
Britannica Online	Wikipedia	The Semantic Web
HTML / Portals	XML / RSS	RDF / RDFS / OWL

The World Wide Web Consortium

In 1994, Tim Berners-Lee with the help of Massachusetts Institute of Technology founded the World Wide Web Consortium (W3C), an international community which is devoted to developing ‘Open web standards’.

The W3C describes itself as,



“The World Wide Web Consortium exists to realize the full potential of the web”



Web Architecture

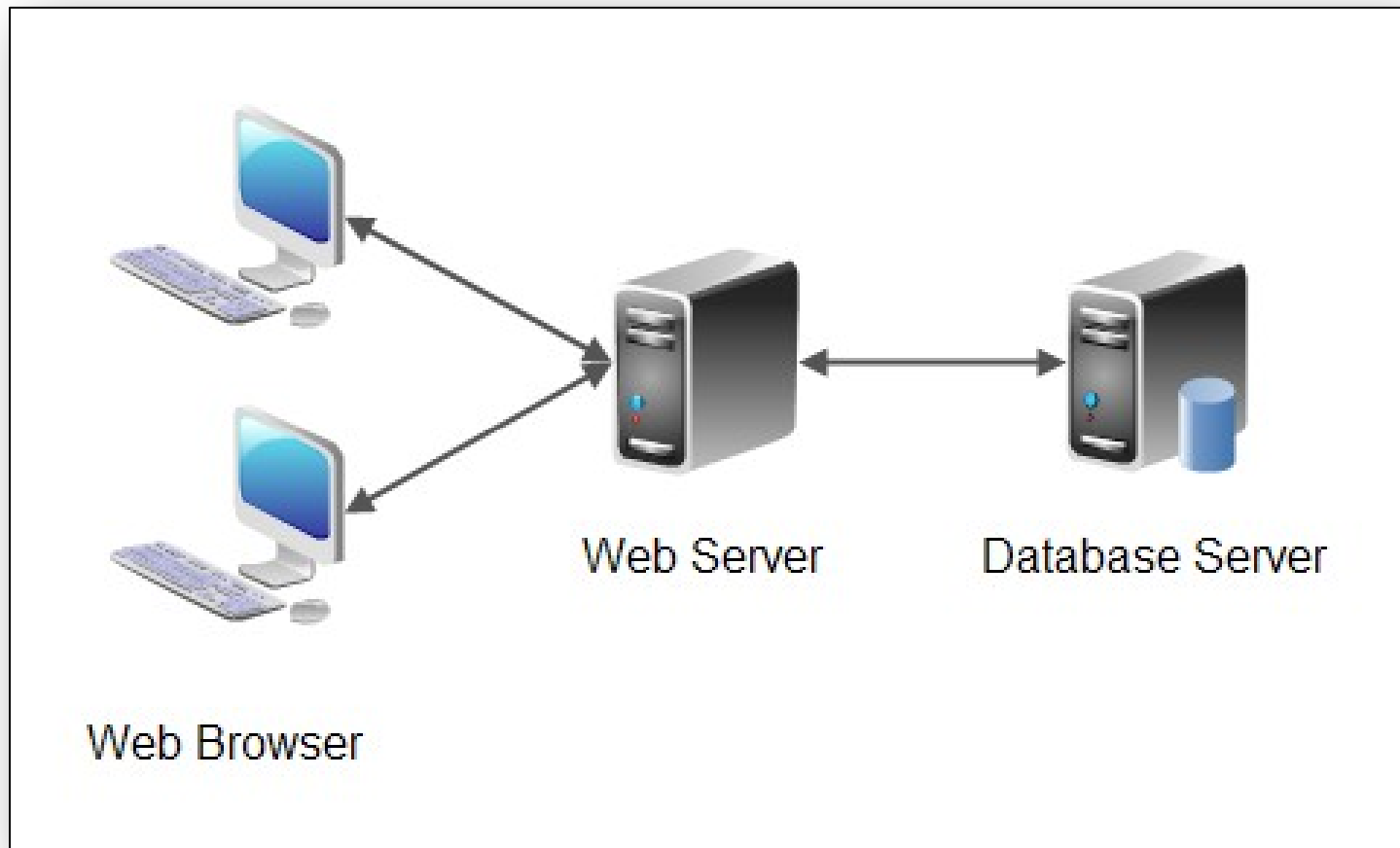


Diagram: The Web Architecture

Web Server



- A web server can be both software and hardware.
- It is a program that responds to the requests made by the client.
- A computer where the websites are hosted or stored can also be called as a web server.

Web Clients

- A web client can be both software and hardware.
- A web client is an application that communicates with a web server, using Hypertext Transfer Protocol (HTTP).
- A Web client contains two parts: dynamic Web pages and the Web browser.
- Dynamic Web pages are produced by components that run in the Web tier, and a Web browser delivers Web pages received from the server.



Unified Resource Identifier (URI)

- The web servers **host web resources**.
- A web resource is a content source, it can be anything from a static file, a live image from a camera or dynamic content resources.
- Each web server resource has a name to identify the resource uniquely called the ‘**Uniform Resource Identifier**’ or **URI**.
- For example:

Below is an example of an image resource URI.

<http://www.example.com/specials/polar.gif>

The Uniform Resource Identifier has two forms:

1. Uniform Resource Locator
2. Uniform Resource Name

Unified Resource Locator (URL)

Uniform Resource Locator is the most common form of resource identifier.

A URL defines the specific location of the resource on a particular server.

Example: <http://www.yahoo.com/images/logo.gif> is the URL for the Yahoo! Website's logo.

http://

Scheme

www.yahoo.com

Address

/images/logo.gif

**Resource
location**

Unified Resource Name (URN)

- **Uniform Resource Name** serves as a unique name to a resource irrespective of the location where it resides.
- Being location-independent these resources can move freely and can be accessed by multiple network access protocols.
- Uniform Resource Names are still in experimental stage and not yet widely adopted.
- For example,

The Uniform Resource Name *urn:ietf:rfc:2141* can be used to name the internet standards document “*RFC 2141*”.

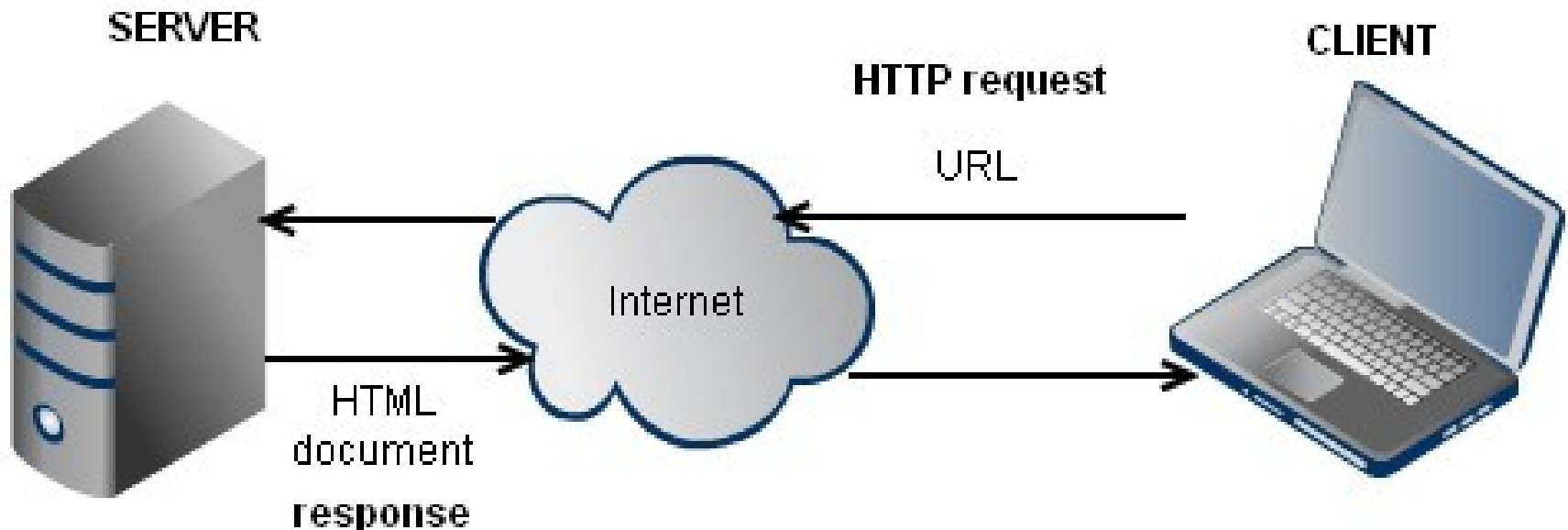
Web Browser

- A web browser is an application program used to request, view and traverse the web pages.
- It can display information in the form of text, images, multimedia and more.
- We can upload or download files using the browser.



Hyper Text Transfer Protocol

Hyper Text Transfer Protocol or HTTP is the protocol that is used to enable a web browser and a web server to communicate.



The Hyper Text Transfer Protocol

HTTP Kitchen Analogy

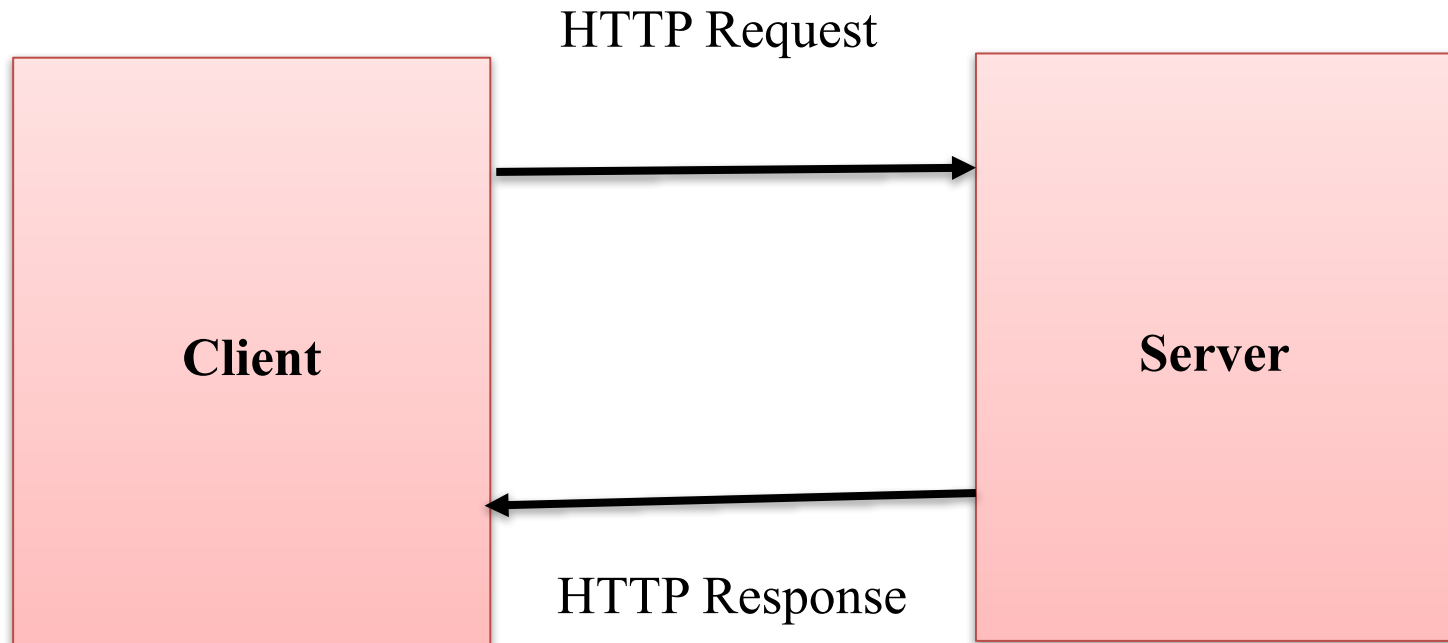
Hyper Text Transfer Protocol

1. Suppose I want to access a web page like *www.example.com/myvideos/vid1*.
2. The HTTP protocol uses the URL to identify the server, it makes a connection and sends the request to the server and terminates the connection.
3. The server receives the request and checks if the requested resource is available, if so it reestablishes the connection with the client and sends the response.
4. The browser can make multiple requests and each request is processed independent of the previous requests.

Eating in a Restaurant

1. The first thing you need to do is find the restaurant you want to eat from.
2. Once in the restaurant you would send an order to the kitchen for whatever you like to eat.
3. The kitchen checks to see if the ordered item is available, if yes it is delivered to you on your table.
4. You can make multiple orders while in the restaurant and the kitchen while processing a given order does not keep in mind what your previous orders were.

HTTP Client-Server Model



HTTP MIME(Multipurpose Internet Mail Extensions) Types

Document	MIME type
HTML formatted text document	<i>text/html</i>
Plain ASCII text document	<i>text/plain</i>
JPEG version	<i>image/jpeg</i>
GIF format	<i>image/gif</i>
Apple quick time movie	<i>video/quicktime</i>
Microsoft power point presentation	<i>application/vnd.ms-powerpoint</i>

HTTP Transaction

<i>Method URI HTTP version</i>	<i>HTTP version Status code</i>	Start line
<i>Name1: value1</i> <i>Name2: value2</i> <i>Name3: value3</i>	<i>Name1: value1</i> <i>Name2: value2</i> <i>Name3: value3</i>	Headers
	<i>Requested Resource</i>	Body

Request Message

Response Message

HTTP Message Formats

Using the format let us write request and response messages for our index.html example.

Request Message

```
GET    /index.html  HTTP 1.0  
Host: www.example.com  
Accept: text/html  
Accept-language: en-us
```

Response Message

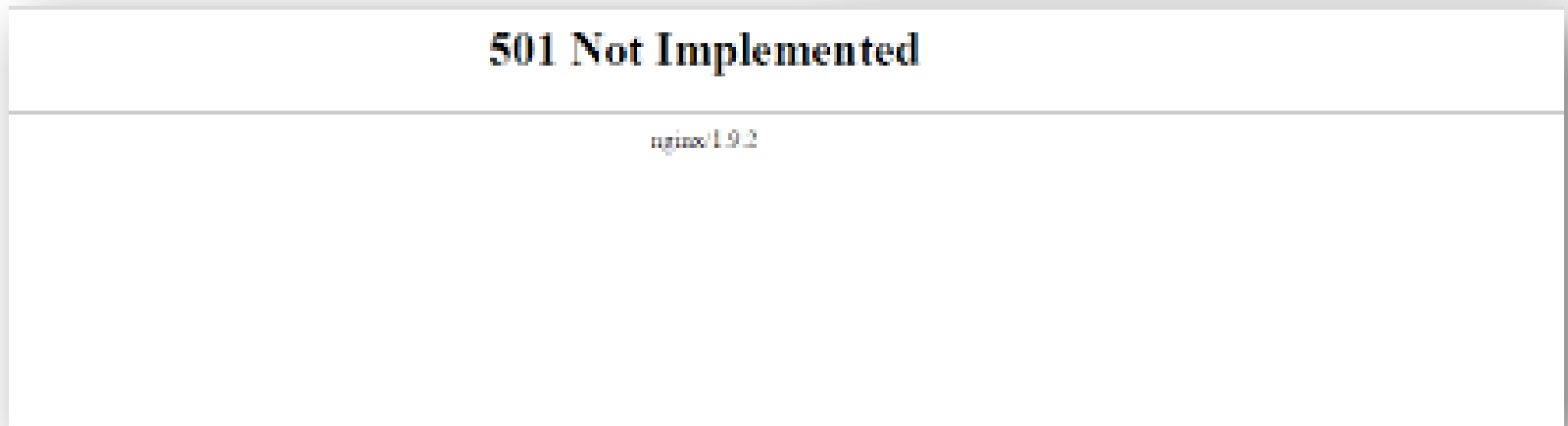
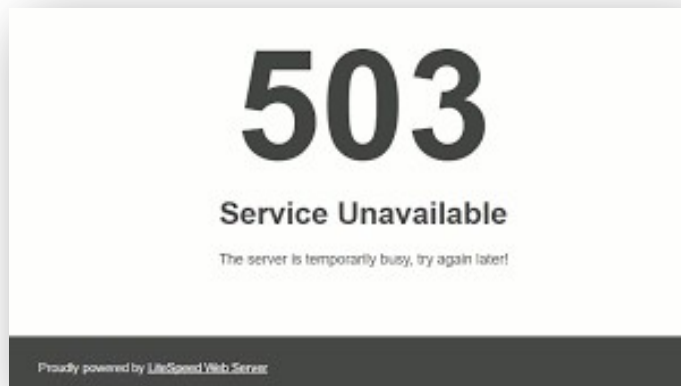
```
HTTP 1.0    200 : OK  
Host: www.example.com  
Accept: text/html  
Accept-language: en-us  
index.html
```

Request and Response Messages

HTTP Methods

Method	Description
GET	Used to get a named resource from the server using URI.
PUT	Replace all representations of the resource with the new upload.
POST	Used to send data to the server like customer information using HTML form
DELETE	Deletes the resource from the server
HEAD	Similar to GET, but transfers the status line and header section of the message
CONNECT	Establishes a connection with the server identified by the URI

The HTTP Status Codes



Web Application

A **Web application** (Web app) is an **application** program that is stored on a remote server and delivered over the Internet through a browser interface.

It can be:

Static Web Application

- A collection of related web pages that may contain text, links, audio, images and videos.
- It requires only client-side processing.

Dynamic Web Application

- A dynamic web application is interactive.
- They can automatically update different sections based on information from other applications or databases. Users see different inputs based on the input given.
- Dynamic web applications require back-end processing in addition to client-side processing.
- Facebook app is an example of dynamic web application.

Static and Dynamic Web Applications

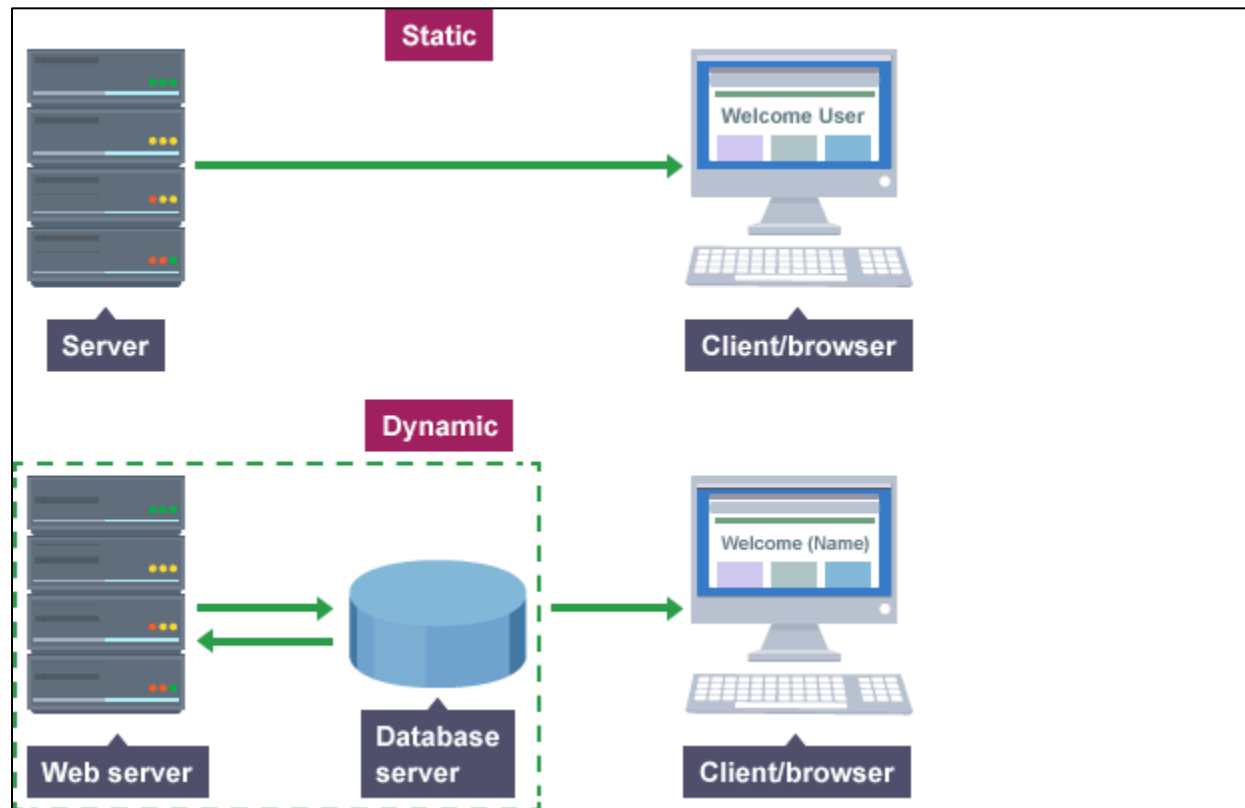


Diagram: Static vs. Dynamic Web Applications

Front-end and Back-end Web Development

The **web development functions** can be categorized into five areas:

1. Preparing page structure.
2. Organizing and managing content hierarchy.
3. Serving content to the user.
4. Capturing user input.
5. Performing back-end processing and integration.

Different web technologies are required to accomplish these functions. Web development can be divided into **two categories**:

6. Front-end Web Development

7. Back-end Web Development

Front-end Development is also called the client-side development

- It is used for developing what the user sees when they open a website or load a web application.
- It is used to develop the structure and design of websites.

Back-end development - Most of the code that makes a web application work is on the back-end.

- The back-end code runs on the server therefore, a back-end developer must have a good understanding of the programming languages, database and server architecture.
- The server-side programming can be divided into four main components:
 - ❖ **The Servers**
 - ❖ **Databases**
 - ❖ **Middleware**
 - ❖ **Programming languages and framework.**

Web Technologies

Web technologies is a general term referring to the many languages and multimedia packages that are used in conjunction with one another, to produce dynamic web sites. Some of the important technologies are:

- **Hyper Text Markup Language (HTML)** - HTML is the **standard markup language for creating web pages** and web applications.
- **Cascading Style Sheets (CSS)** - the Cascading Styling Sheets is used to style an HTML document.
- **JavaScript** - It is a dynamic programming language which when applied to a HTML document can make the websites interactive with the user.
- **Extensible Markup Language (XML)** - it is an extensible language which can be used to create specific mark-up languages such as chemical mark-up.
- **ASP.NET** - ASP .NET is a web framework for building dynamic web sites, web applications and web services.

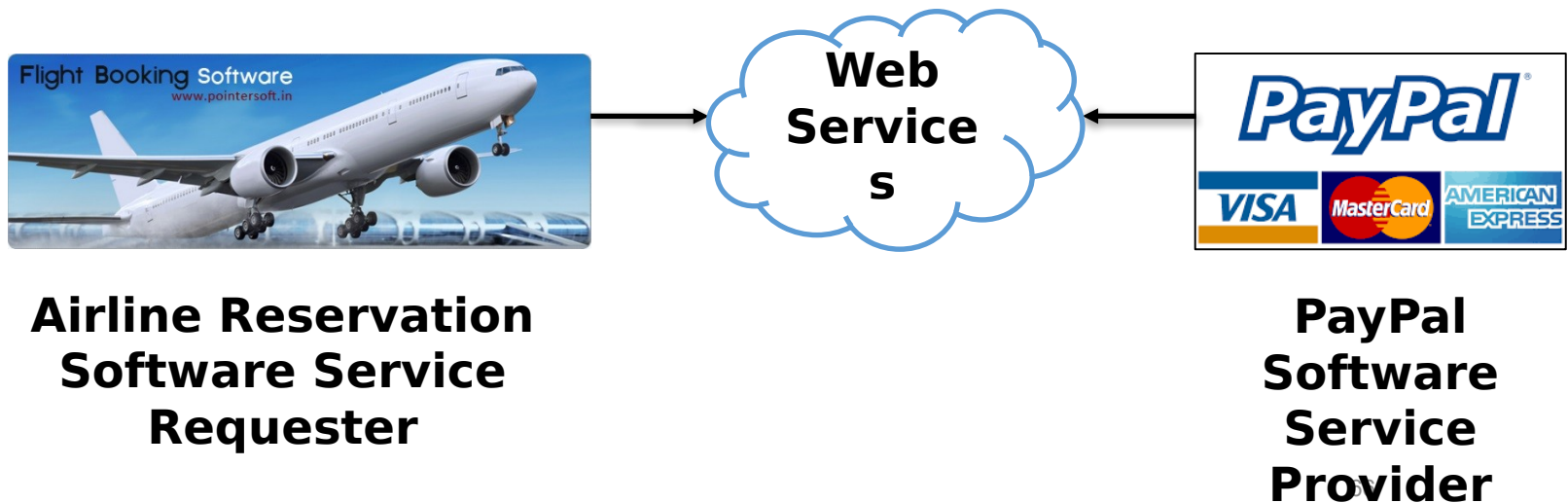
Web Technologies

- **Servlets** – It is a server-side programming language. An efficient and powerful solution for creating dynamic web content.
- **Java Server Pages (JSP)** - The Java Server Pages is a server-side technology that is used to create dynamic web pages based on HTML, XML or other document types. Java code is inserted into HTML using JSP tag `<% ----Java code----%>`.
- **Hypertext Preprocessor (PHP)** - PHP is an open source general purpose scripting language for server-side development. It can be embedded in HTML.
- **Asynchronous JavaScript And XML (AJAX)** - AJAX is not a programming language in itself. It uses the XMLHttpRequest object to request data from the server and java script and HTML DOM to display or use the data. The most appealing characteristic of the AJAX is that it can communicate with the server, exchange data and update a page without even refreshing the page.

Web Services analogy

Consider the **aircraft reservation software**,

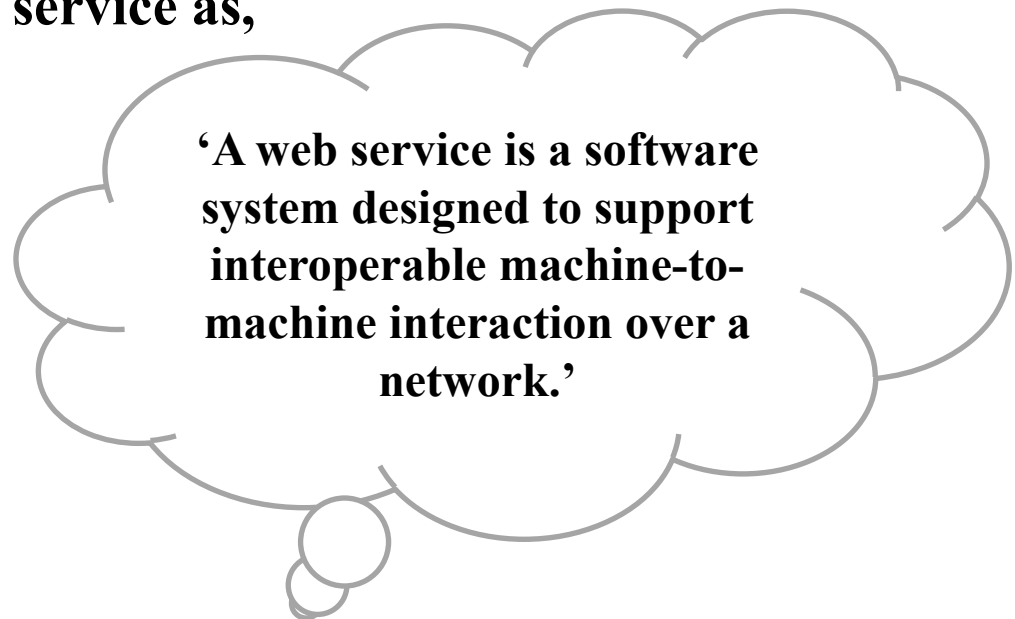
- In addition to doing reservations, the software also requires to communicate with e-commerce sites like **PayPal for payments** made online.
- But we know that the reservation system and the PayPal software are separate systems written in different languages.
- Communication between the two **heterogeneous** software systems happens through a special kind of web application known as a **web service**.



Web Services

- Web service is a service offered by one electronic device to another via the World Wide Web.

The W3C defines a web service as,

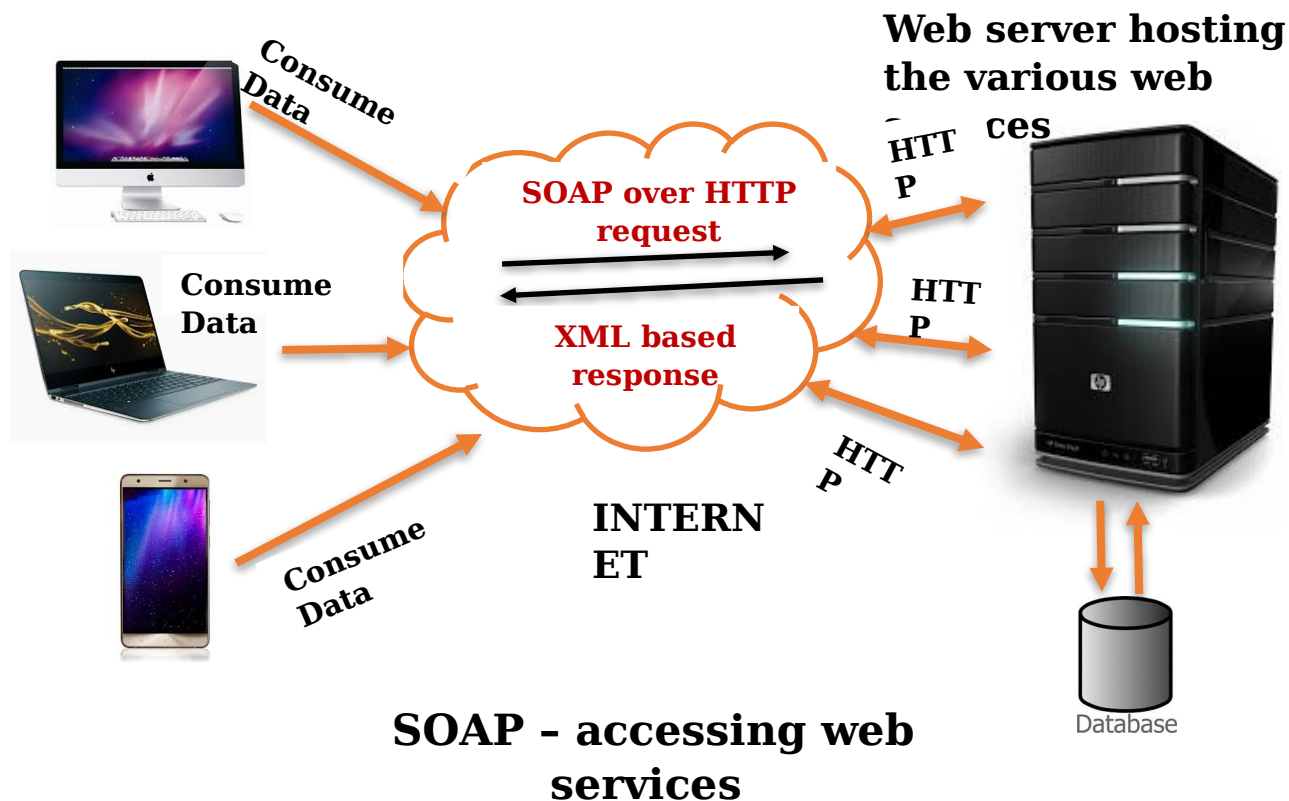


Web Service Components

- **SOAP** – Simple Object Access Protocol
- **REST** – Representational State Transfer
- **WSDL** – Web Services Description Language
- **UDDI** – Universal Description, Discovery and Integration

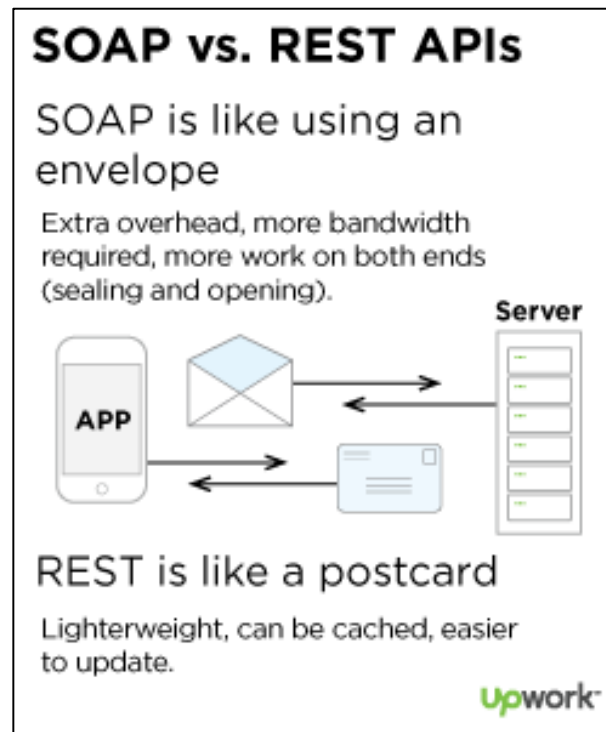
SOAP – Simple Object Access Protocol

SOAP is an XML based messaging protocol which allows programs running on different operating systems such as Windows and Linux to communicate using HTTP and XML.



REST – Representational State Transfer

- REST was defined by Roy Fielding in 2000.
- REST is an architectural style for developing web services. It uses HTTP for accessing resources, but unlike SOAP the web service developed using REST are light weight and use less bandwidth.



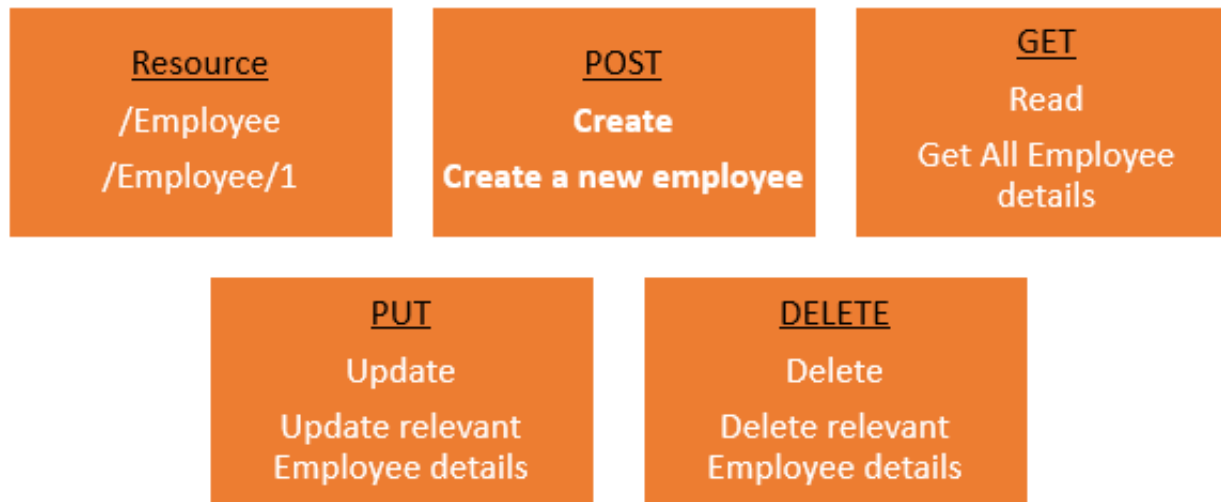
REST – Representational State Transfer

- REST defines a way of accessing resources such as documents, pictures or videos which reside on a different environment.
- The key elements of RESTful implementation are:
 1. **Resources** – the resource itself.
 2. **Request Verbs** – to describe what to do with the resource. The verbs include **GET, POST, DELETE, PUT**.
 3. **Request Header** – additional information such as authorization details.
 4. **Request Body** – contains Data, usually sent when POST is used.
 5. **Response Body** – an XML document with the data.
 6. **Response Status Codes** – returned with the response indicating whether there is any error or response has been sent.

REST - Example

Let us assume a RESTful web service is defined at a particular location. The client can use any HTTP verbs to make the request. For Example,

- **POST** – This would be used to create a new employee using the RESTful web service
- **GET** - This would be used to get a list of all employee using the RESTful web service
- **PUT** - This would be used to update all employee using the RESTful web service
- **DELETE** - This would be used to delete all employee using the RESTful web service



Request Verbs on a single employee record