



ABES

Engineering College

College Code 032

Estd. 2000

Internet & Web Technology

PRESENTED BY: Sanjeev Soni

Assistant Professor

ABES- IT Dept



INTERNET & WWW (WORLD WIDE WEB)



What is the Internet?

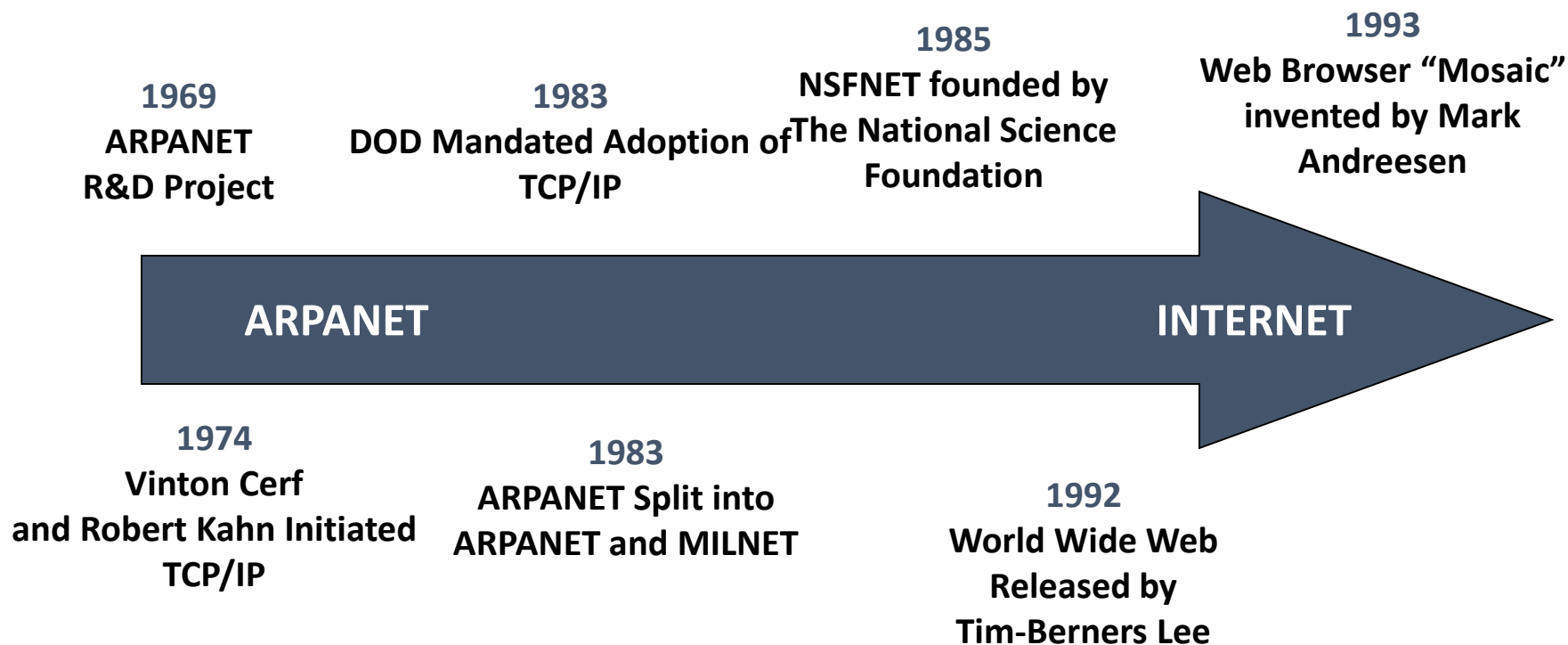
- Internet is a global network of interconnected networks and devices that exchange information.
- The internet is a publicly available i.e. global network of computers.
- It interconnects billions of Users. For example:
 - Businesses
 - Citizens
 - Governments
 - Academic Institutions
 - Research Centers
 - Libraries
 - Etc.
- Still Rapidly Expanding
- Involves Numerous Technologies (Not a Single Technology)

Web Technology

- Web Technology refers to the various tools and techniques that is utilized in the process of communication between different types of devices over the Internet.



Internet History Milestones



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.

Applications of the Internet

- The internet is used for a many of applications. For example:
 - Sending and receiving email.
 - Searching and browsing information archives.
 - Copying files between computers.
 - Conducting financial transactions.
 - Navigating (in your car, smart scooter, smart bike, or other)
 - Playing interactive games.
 - Video and music streaming.
 - And so more...



The World Wide Web (WWW)

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

- Birth of the World Wide Web
 - 1992 Tim Berners-Lee, a physicist in Geneva Switzerland came up with the term World Wide Web - WWW - W3 who is called father of WWW.
 - June 1993 Marc Andreessen & others at NCSA - University of Illinois developed Mosaic - 1st Web Browser.
- The World Wide Web is a system of interlinked hypertext documents accessed via the internet.
- The world wide web is a advanced and complicated system for capturing and delivering universal information .
- The world wide web consortium (W3C) definition of the web: "*The World Wide Web is the universe of network-accessible information, an embodiment of human knowledge*" .



- Provides information access in ways not previously possible
 - Hyperlinked (Hypertext)
 - Graphical user interface
 - Pictorial and non-text information
 - Information that changes rapidly
 - Immediate access
 - Anyone can author a web site
 - Multi-user access to the same information (try that with a book)
 - Easily searchable information

Evolution of Web

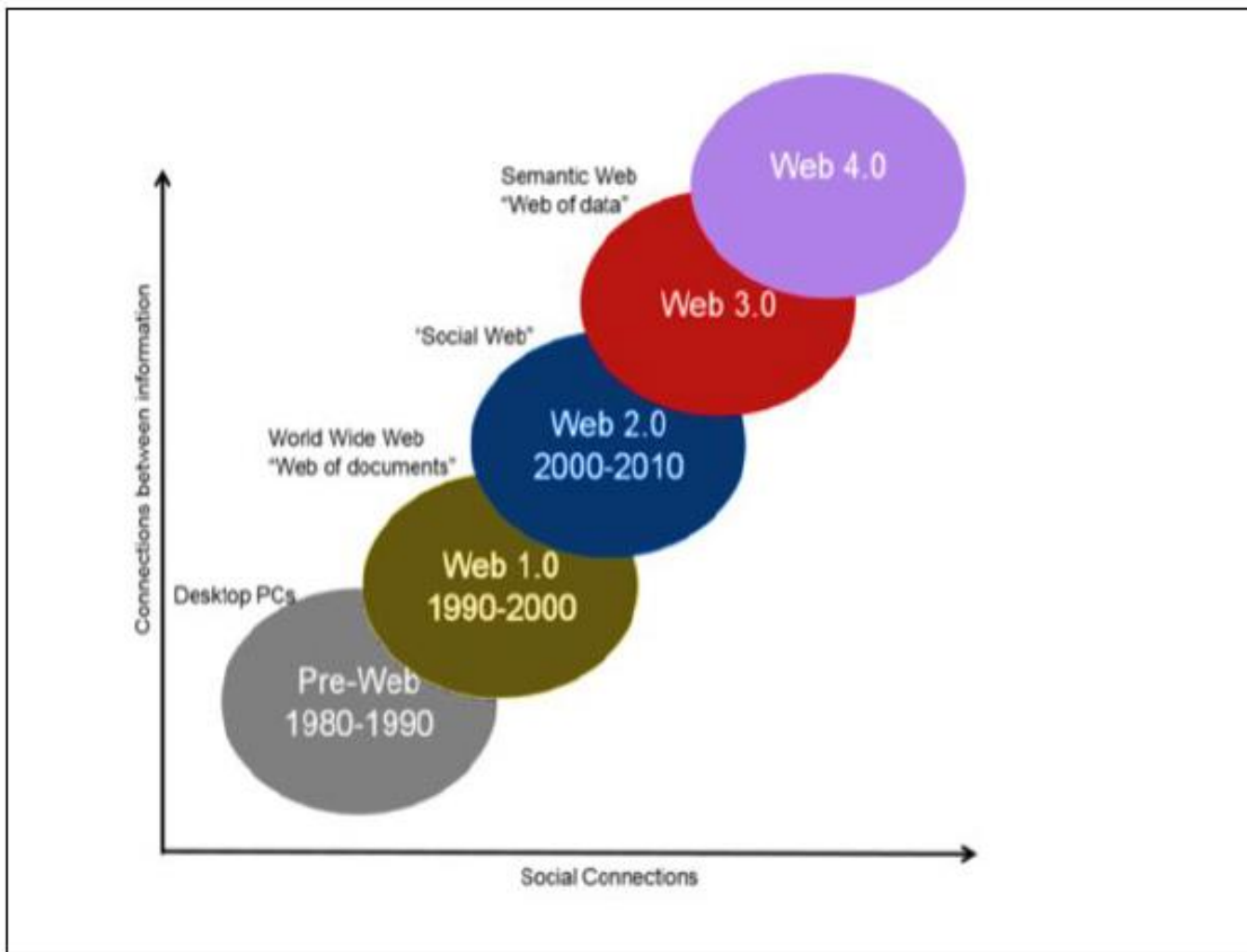


Diagram: World Wide Web Timeline

Evolution of Web

Web 1.0 / 2.0 / 3.0 Summary

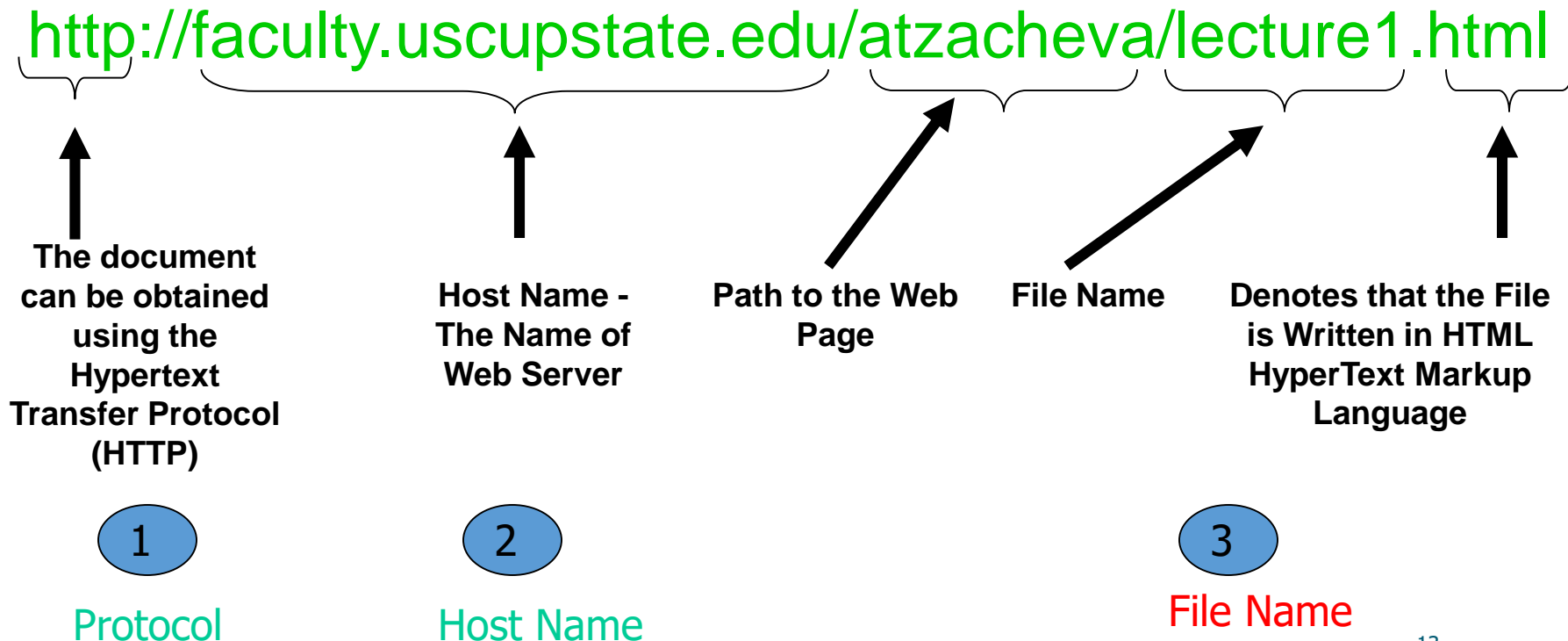
Crawl	Walk	Run
Web 1.0	Web 2.0	Web 3.0
Mostly Read-Only	Wildly Read-Write	Portable & Personal
Company Focus	Community Focus	Individual Focus
Home Pages	Blogs / Wikis	Lifestreams / Waves
Owning Content	Sharing Content	Consolidating Content
Web Forms	Web Applications	Smart Applications
Directories	Tagging	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

Functionality of the WWW

- The functionality of the WWW is based on 3 main standards:
 - **URL (Universal Resource Locator)**
 - **HTML (Hypertext Markup Language)**
 - **HTTP (Hypertext transfer Protocol)**

The Universal Resource Locator (URL)

Each page of information on the web has a unique address called the URL at which it can be found





Hyper Text Markup Language

- HTML is a universal, simple language for formatting text. It also allows embedding of graphics, and linking of documents via 'hyperlinks'.
- 'Hypertext' documents allow readers to freely move around the document, following links to subjects of interest.
- HTML is a language that is used to specify the structure of documents for retrieval across the Internet using browsers.



A portion of the html document from the course web site:

```
<a  
href="../../../SIMS201/Fall2006/ComputerNetworks_Chapter19.ppt">Co  
mputer Networks</a>
```

HTTP



- The standard Web transfer protocol is HTTP (HyperText Transfer Protocols). It is used for document exchange between servers and clients (typically browsers) in the WWW.
- To retrieve a document, the client first sends a request to the web server and waits for a reply
- An HTTP daemon/service (a program that waits for http requests) on the server then handles the request and the document is sent to the client (over a connection established using TCP/IP – Transmission Control Protocol / Internet Protocol)
- We recall from previous lectures, that a Protocol is “Agreed upon sets of rules that provide order to a system or situation.” In this case, a set of rules/standards for communicating on the Internet.

Web Client/Server Architecture

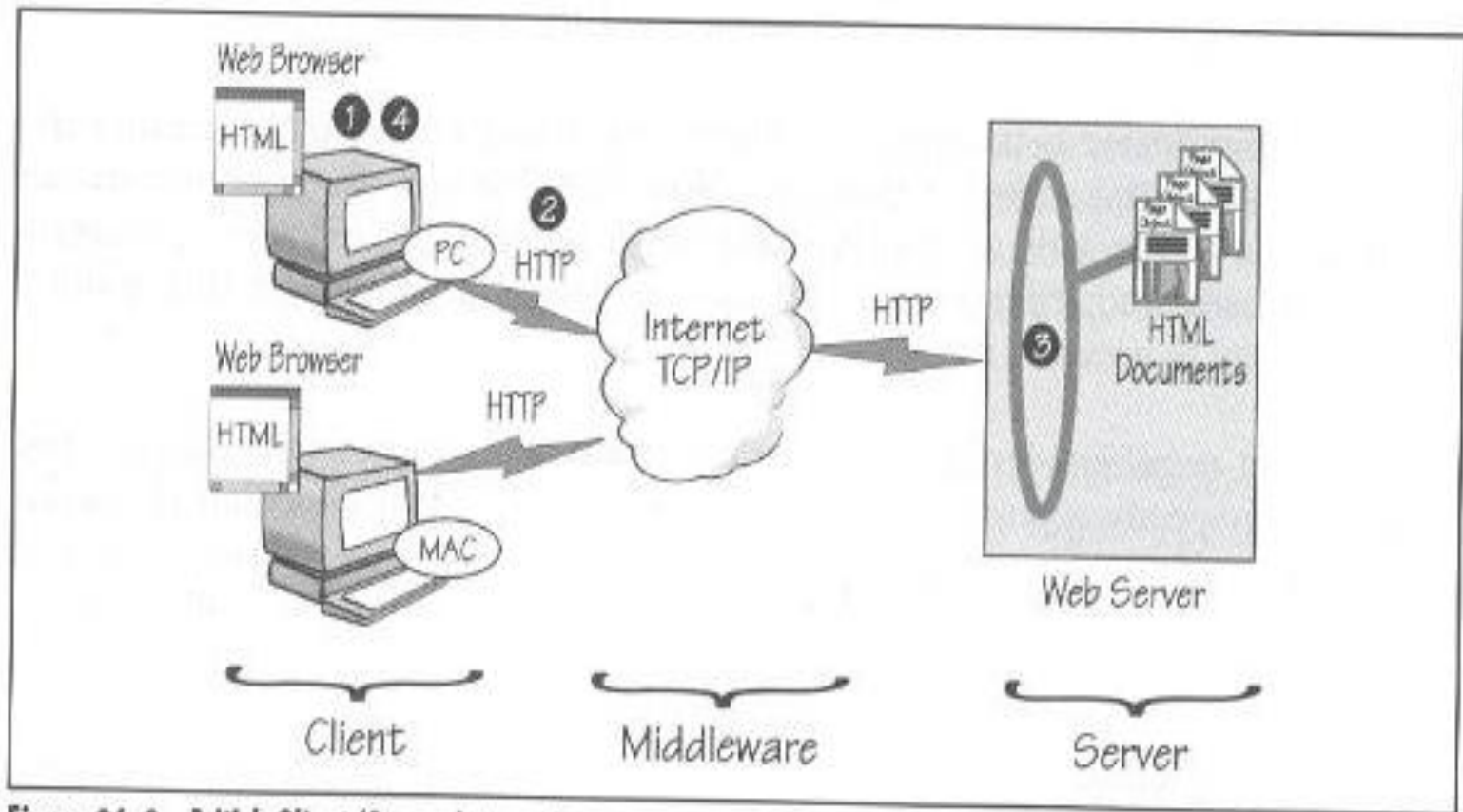


Figure 26-2. A Web Client/Server Interaction.



Web Technology

Web Technology refers to the various tools and techniques that are utilized in the process of communication between different types of devices over the Internet. A web browser is used to access web pages. Web browsers can be defined as programs that display text, data, pictures, animation, and video on the Internet. Hyperlinked resources on the World Wide Web can be accessed using software interfaces provided by Web browsers.

Web Technology can be classified into the following sections:

- **World Wide Web (WWW):** The World Wide Web is based on several different technologies: Web browsers, Hypertext Markup Language (HTML), and Hypertext Transfer Protocol (HTTP).
- **Web Browser:** The web browser is an application software to explore www (World Wide Web). It provides an interface between the server and the client and requests to the server for web documents and services.

Web Technology

- **Web Server:** Web server is a program which processes the network requests of the users and serves them with files that create web pages. This exchange takes place using Hypertext Transfer Protocol (HTTP).
- **Web Pages:** A webpage is a digital document that is linked to the World Wide Web and viewable by anyone connected to the internet has a web browser.
- **Website:** A website is a collection of web pages and related content that is identified by a common domain name and published on at least one web server. Websites are typically dedicated to a particular topic or purpose, such as news, education, commerce, entertainment or social networking.

Types of Website

- Static Website
- Dynamic Website
- **Web Development:** Web development refers to the building, creating, and maintaining of websites. It includes aspects such as web design, web publishing, web programming, and database management. It is the creation of an application that works over the internet i.e. websites.



Types of Web Development

Frontend Development: The part of a website that the user interacts directly is termed as front end. It is also referred to as the 'client side' of the application.

Backend Development: Backend is the server side of a website. It is the part of the website that users cannot see and interact. It is the portion of software that does not come in direct contact with the users. It is used to store and arrange data.

Frontend Languages

Frontend Languages: The front-end portion is built by using some languages which are discussed below:

- **HTML:** HTML stands for Hypertext Markup Language. It is used to design the front-end portion of web pages using a markup language. HTML is the combination of Hypertext and Markup language. Hypertext defines the link between the web pages. The markup language is used to define the text documentation within the tag which defines the structure of web pages.
- **CSS:** Cascading Style Sheets fondly referred to as CSS is a simply designed language intended to simplify the process of making web pages presentable. CSS allows you to apply styles to web pages. More importantly, CSS enables you to do this independent of the HTML that makes up each web page.
- **JavaScript:** JavaScript is a famous scripting language used to create magic on the sites to make the site interactive for the user. It is used to enhancing the functionality of a website to running cool games and web-based software.
- **AJAX:** Ajax is an acronym for Asynchronous Javascript and XML. It is used to communicate with the server without refreshing the web page and thus increasing the user experience and better performance.

Backend Languages

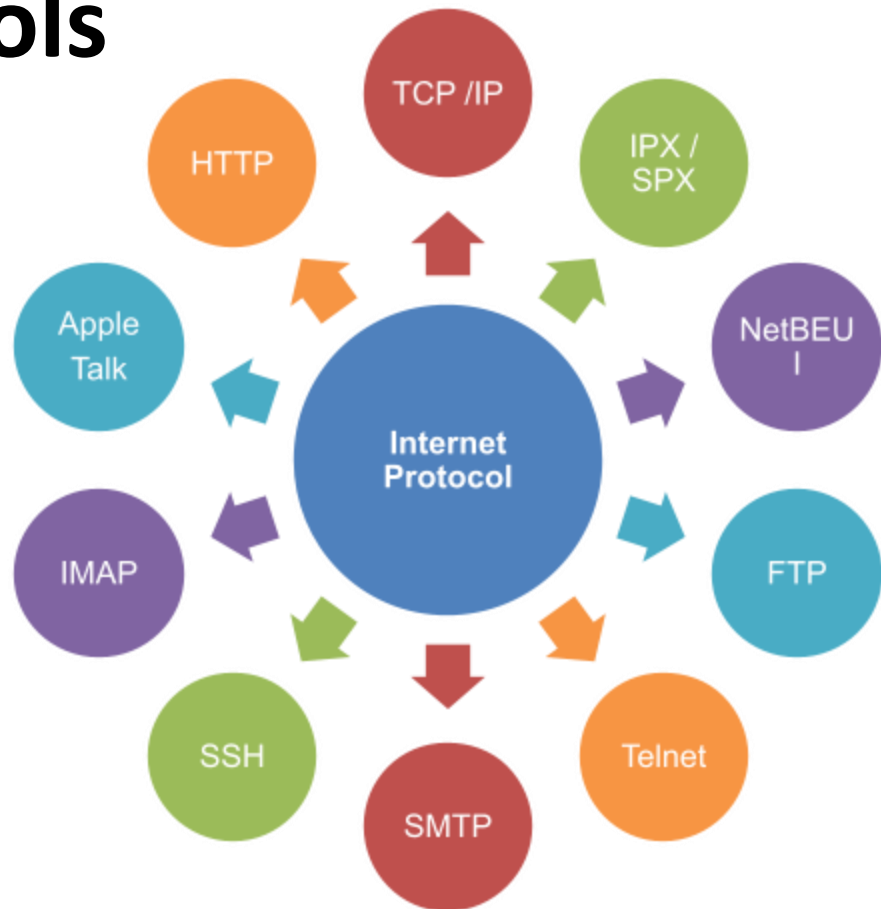
Backend Languages: The back end portion is built by using some languages which are discussed below:

- **PHP:** PHP is a server-side scripting language designed specifically for web development. Since PHP code executed on the server-side, so it is called a server-side scripting language.
- **Node.js:** Node.js is an open-source and cross-platform runtime environment for executing JavaScript code outside a browser. You need to remember that NodeJS is not a framework, and it's not a programming language. Most people are confused and understand it's a framework or a programming language. We often use Node.js for building back-end services like APIs like Web App or Mobile App. It's used in production by large companies such as Paypal, Uber, Netflix, Wallmart, and so on.
- **Python:** Python is a programming language that lets you work quickly and integrate systems more efficiently.

Backend Languages

- **Ruby:** Ruby is a dynamic, reflective, object-oriented, general-purpose programming language. Ruby is a pure Object-Oriented language developed by Yukihiro Matsumoto. Everything in Ruby is an object except the blocks but there are replacements too for it i.e procs and lambda. The objective of Ruby's development was to make it act as a sensible buffer between human programmers and the underlying computing machinery.
- **Java:** Java is one of the most popular and widely used programming languages and platforms. It is highly scalable. Java components are easily available.
- **JavaScript:** JavaScript can be used as both (front end and back end) programming.
- **Golang:** Golang is a procedural and statically typed programming language having the syntax similar to C programming language. Sometimes it is termed as Go Programming Language.
- **C#:** C# is a general-purpose, modern and object-oriented programming language pronounced as "C sharp".
- **DBMS:** The software which is used to manage database is called Database Management System (DBMS)

Internet Protocols





Protocols

- 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 are here
 - TCP
 - FTP/IP
 - SMTP/POP3
 - Telnet



Protocols - 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

Protocols – FTP/IP

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.

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.

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)

IP Address - 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**.

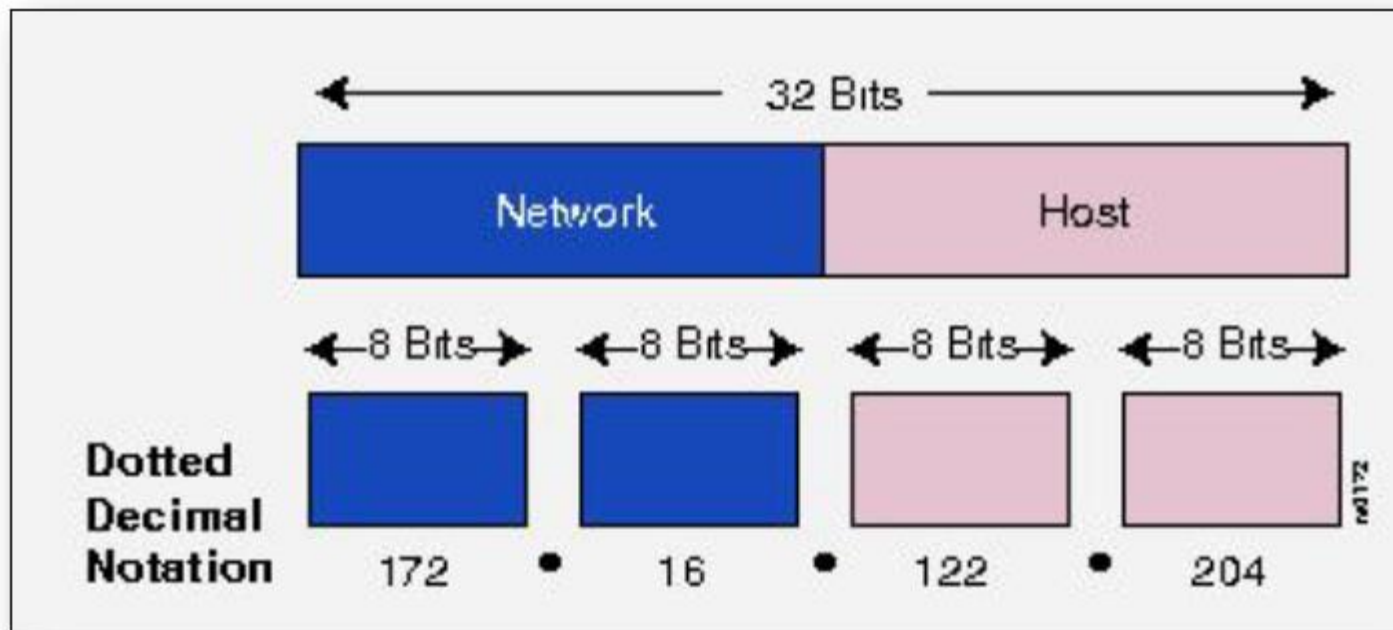


Diagram: IPV4 Address Format

IP Address – 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.

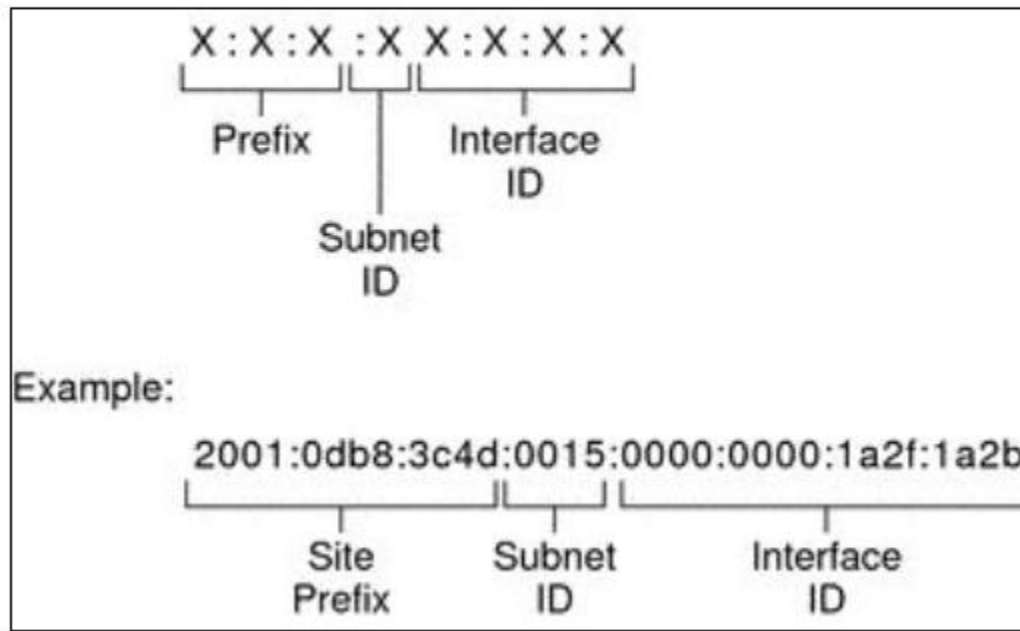


Diagram: IPV6 Address Format

Protocols – UDP/TELNET

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.

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.

Protocols – SMTP/POP3

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.

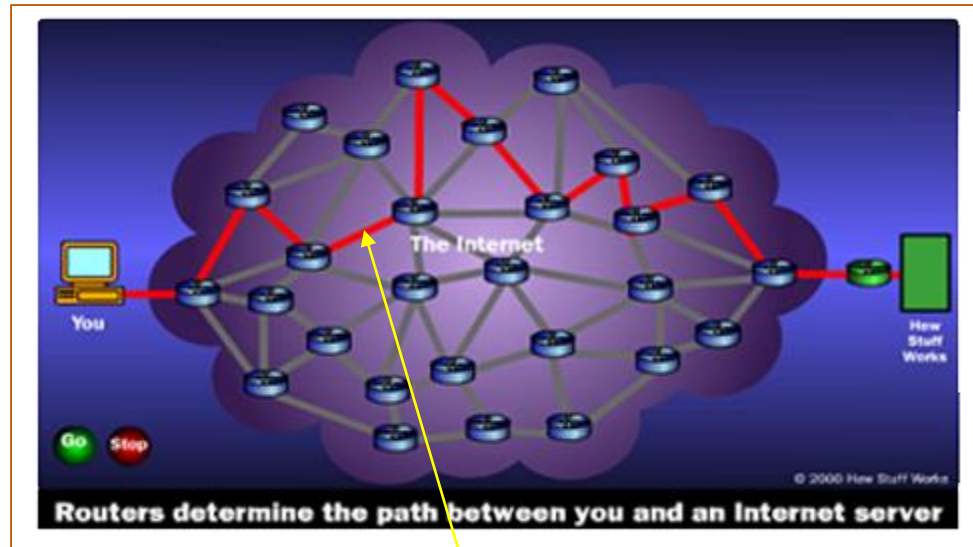
How the Internet works



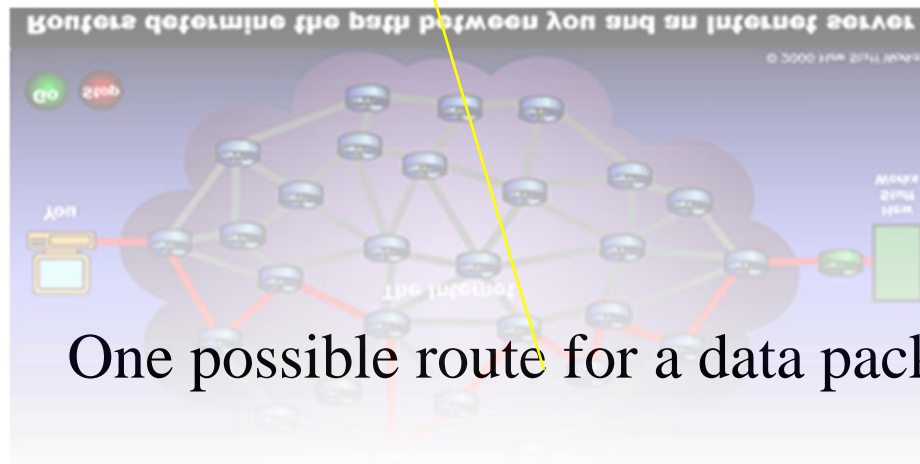
Underlying principle of the Internet: *Packet Switching & Circuit Switching*

- The internet was first conceived in 1969 as ARPANET (Advanced Research Projects Agency Network) which was used to connect computers in geographically dispersed research centers
- ARPANET was the first major implementation of a *packet switched* network
- Previously, data communications were based on *circuit switching* as in traditional telephone networks, where a dedicated link is set up for the duration of the call
- The invention of *packet switching* in the early 1960's enabled the transmission of data over links that are not established beforehand. In packet switching, data is assembled into packets which are then independently routed (using routers) to their destination through various links over the network. This leads to the distribution of resources, an increase in resource utilization, and increased fault tolerance since a packet may take alternative routes to reach its destination

Source



Destination



One possible route for a data packet

Addressing Schemes

We know that packets are sent over the network, but how does a packet reach its destination?

Three addressing schemes are used in sending information across the Internet.

- **Organizationally-Unique Identifier (OUI)**
 - This is the 24-bit number that uniquely identifies vendor, manufacturer or other organization globally or worldwide.
 - These are purchased from incorporated (IEEE) registration authority.
- **IP Address**
 - All devices connected to the internet are identified by a unique 32-bit IP address
 - It is via IP addresses that computers in the internet can identify each other
- **Domain Name System (DNS)**
 - Hierarchical, alphanumeric addressing scheme that is a “synonym” of an IP address



The IP Address

- Unique 32-bit (4 byte) logical address
 - 129.174.1.8 (This one belongs to GMU)
- Made up of two parts
 - Network Number
 - Identifies a network
 - Must be assigned by the Internet Network Information Center (InterNIC)
 - Host Number
 - Identifies a host on a network
 - Assigned by the local network administrator

Decipher IP address

Converting a 32-bit Internet Address to Dotted Decimal Format

Recall binary to decimal conversion

- An Internet address, known as an IP address for “Internet Protocol” is comprised of four binary octets, making it a 32-bit address.
- IP addresses, difficult for humans to read in binary format, are often converted to “dotted decimal format”
- To convert the 32-bit binary address to dotted decimal format, divide the address into four 8-bit octets and then convert each octet to a decimal number.
- Each octet will have one of 256 values (0 through 255)

192.48.29.253

(Example of an IP address in dotted decimal form)

IP address conversion

Convert the following 32-bit Internet address into dotted decimal format:

01011110000101001100001111011100

1) Divide the IP address into four octets

01011110

00010100

11000011

11011100

2) Convert each binary octet into a decimal number

$01011110 = 64 + 16 + 8 + 4 + 2 = 94$

$00010100 = 16 + 4 = 20$

$11000011 = 128 + 64 + 2 + 1 = 195$

$11011100 = 128 + 64 + 16 + 8 + 4 = 220$

3) Write out the decimal values separated by periods

94.20.195.220

The Domain Name System

- We would go crazy if we would have to remember the IP addresses of all the web sites that we wanted to visit



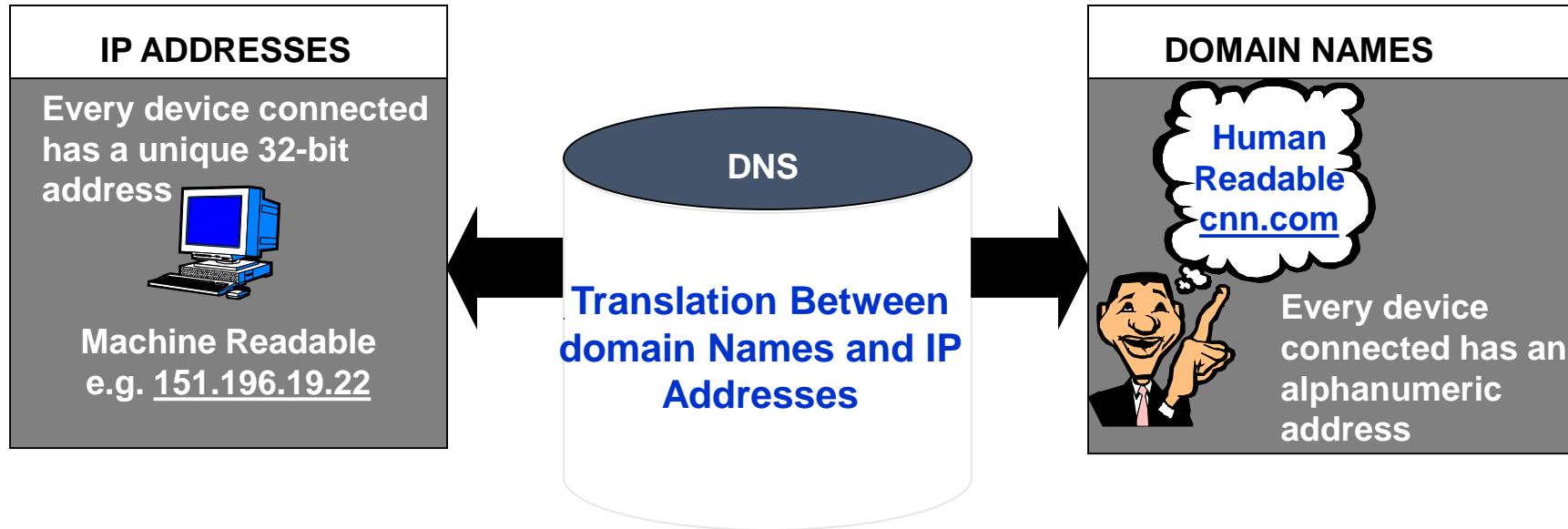
- The *Domain Name System* translates between domain names and IP addresses of devices connected to the Internet
 - A *domain name* (a part of the URL) is a unique alphanumeric name such as gmu.edu
 - The top level domain name is edu and the secondary level domain name is gmu in the above example (there could be up to 127 levels, but more than 4 is rare)



Examples of top level domains

- **Generic top level domains**
 - .com
 - .biz
 - .info
 - .edu
 - .mil
 - .net, etc.
- **Country codes (2 character codes)**
 - .jp, .sw, .us, etc.

DNS



- ❑ IP address and domain name allocation requires central administration to avoid duplication
- ❑ Previously administered by U.S. government contract (NSI)
- ❑ In 1998, technical coordination assigned to ICANN (Internet Corporation for Assigned Names and Numbers).



How DNS works

- When you type a URL into the address portion of your browser, the browsers' first task is to convert the domain name into an IP address so that it can send a request to the appropriate web server at that IP address for the web page that you are requesting
- To accomplish this task, the browser interacts with a *name server*, a machine that maintains lists to translate the domain names into IP addresses
- Once the browser receives the IP address from the name server, it sends a request for the web page that you are requesting
- Let us retrieve the IP address of the name server that we are using...

Thank You !!!