CSE 256
INTERNET TECHNOLOGIES

TUTORIAL SHEET-1

Presidency University, Bengaluru

Welcome to the exciting and rapidly evolving world of Internet and web programming!

Introduction…………..course overview

You'll begin by learning the client-side programming technologies used to build web pages and applications that are run on the client (i.e., in the browser on the user's device).

You'll use Hyper Text Markup Language 5 (HTML5) and Cascading Style Sheets 3 (CSS3)

The recent releases of HTML and CSS technologies—to add powerful, dynamic and fun features and effects to web pages and web applications, such as audio, video, animation, drawing, image manipulation, designing pages for multiple screen sizes

You'll learn JavaScript—the language of choice for implementing the client side of Internet-based applications

you'll learn server-side programming—the applications that respond to requests from client-side web browsers, such as searching the Internet, checking your bank-account balance, ordering a book from Amazon, bidding on an eBay auction and
ordering concert tickets.

there's also an emphasis on Ajax development, which helps
you create better-performing, more usable applications.

A brief history of internet

The Internet—a global network of computers—was made possible by the convergence of computing and communications technologies.

In the late 1960s, ARPA (the Advanced Research Projects Agency) rolled out blueprints for networking the main computer systems of about a dozen ARPA-funded universities and research institutions.

The history of telecommunication and data transport is a long one.

The Internet is not alone in providing instantaneous digital communication. Earlier technologies like radio, telegraph, and the telephone provided the same speed of communication in an analog form.

What Is the Internet?

The Internet is a huge collection of computers connected in a communications network.

These computers are of every imaginable size, configuration, and manufacturer.

In fact, some of the devices connected to the Internet—such as plotters and printers—are not computers at all.

The innovation that allows all these diverse devices to communicate with each other is a single, low-level protocol named Transmission Control Protocol/Internet Protocol (TCP/IP).

Circuit switching   or   Packet Switching

Earlier days of network connections are referred to as circuit switching.

Low bandwidth was one of the major issues with circuit switching

Bandwidth is a measurement of how much data can (maximally) be transmitted along an Internet connection.

Normally measured in bits per second (bps), this measurement differs according to the type of Internet access technology you are using.
A dial-up 56-Kbps modem has far less bandwidth than a 10-Gbps fiber optic connection.

A packet-switched network does not require a continuous connection.

Instead it splits the messages into smaller chunks called packets and routes them to the appropriate place based on the destination address.

The packets contained address, error-control and sequencing information.

The packets can take different routes to the destination

Internet Protocol Addresses

To promote the growth and unification of the disparate networks, a suite of protocols was invented to unify the networks.

A protocol is the name given to a formal set of publicly available rules that manage data exchange between two points.

Communications protocols allow any two computers to talk to one another, so long as they implement the protocol.

The protocol for communicating over the ARPANET became known as TCP—the Transmission Control Protocol.

TCP ensured that messages were properly routed from sender to receiver and that they arrived intact.

As the Internet evolved, organizations worldwide were implementing their own networks for both intraorganization (i.e., within the organization) and interorganization (i.e., between organizations) communications.

A wide variety of networking hardware and software appeared.

One challenge was to get these different networks to communicate.

ARPA accomplished this with the development of IP—the Internet Protocol, truly creating a network of networks, the current architecture of the Internet.

The combined set of protocols is now commonly called TCP/IP.

Each computer on the Internet has a unique IP address.

The Internet Protocol (IP) address of a machine connected to the Internet is a unique 32-bit number.

IP addresses usually are written (and thought of) as four 8-bit numbers, separated by periods.

The four parts are separately used by Internet-routing computers to decide where a message must go next to get to its destination.

Domain Names

Because people have difficulty dealing with and remembering numbers, machines on the Internet also have textual names.

These names begin with the name of the host machine, followed by progressively larger enclosing collections of machines, called domains.

There may be two, three, or more domain names.

Because IP addresses are the addresses used internally by the Internet, the fully qualified domain name of the destination for a message, which is what is given by a browser user, must be converted to an IP address before the message can be transmitted over the Internet to the destination.

These conversions are done by software systems called name servers, which implement the Domain Name System (DNS).

The World Wide Web

In 1989, a small group of people led by Tim Berners-Lee at Conseil Européen pour la Recherche Nucléaire (CERN) or European Organization for Particle Physics proposed a new protocol for the Internet, as well as a system of document access to use it.

The intent of this new system, which the group named the World Wide Web, was to allow scientists around the world to use the Internet to exchange documents describing their work.

Main postulates of WWW are as follows

1. A technology for sharing information via hyperlinked text documents. Berners-Lee called his invention the HyperTextMarkup Language (HTML).

2. He also wrote communication protocols to form the backbone of his new information system, In particular, he wrote the Hypertext Transfer Protocol (HTTP)—a communications protocol used to send information over the web.

3. The units of information on the Web have been referred to by several different names; among them, the most common are pages, documents, and resources.

4. A Uniform Resource Locator (URL) to uniquely identify a resource on the WWW. Each web page on the Internet is associated with a unique URL.

5. A software program (later called web server software) that can respond to HTTP requests.

6. A program (later called a browser) that can make HTTP requests from URLs and that can display the HTML it receives.

Web Browsers

When two computers communicate over some network, in many cases one acts as a client and the other as a server.

The client initiates the communication, which is often a request for information stored on the server, which then sends that information back to the client.

The Web, as well as many other systems, operates in this client-server configuration.

Documents provided by servers on the Web are requested by browsers, which are programs running on client machines.

They are called browsers because they allow the user to browse the resources available on servers.

Web Servers

Web servers are programs that provide documents to requesting browsers.

Servers are slave programs: They act only when requests are made to them by browsers running on other computers on the Internet.

The most commonly used Web servers are Apache, which has been implemented for a variety of computer platforms,

and Microsoft's Internet Information Server (IIS), which runs under Windows operating systems

World Wide Web Consortium

In October 1994, Tim Berners-Lee founded an organization—the World Wide Web Consortium (W3C)—devoted to developing nonproprietary, interoperable technologies for the World Wide Web.

One of the W3C's primary goals is to make the web universally accessible—regardless of disability, language or culture.

The W3C is also a standards organization. Web technologies standardized by the W3C are called Recommendations.

Current and forthcoming W3C Recommendations include the Hyper Text Markup Language 5 (HTML5), Cascading Style Sheets 3 (CSS3) and the Extensible Markup Language (XML).

Governance

Internet Engineering Task Force (IETF)- The technical underpinning and standardization of the core protocols (IPv4 and IPv6)

Internet Corporation for Assigned Names and Numbers (ICANN)- coordinates the assignment of unique identifiers for use on the Internet, including domain names, Internet Protocol (IP) addresses, application port numbers in the transport protocols etc...

Intranet or Internet

One of the more common terms you might encounter in web development is the term "intranet" (with an "a"), which refers to an Internet network that is local to an organization or business.

Intranet resources are often private, meaning that only employees (or authorized external parties such as customers or suppliers) have access to those resources.

Internet is a broader term that encompasses both private (intranet) and public networked resources.

CSE 256
INTERNET TECHNOLOGIES

TUTORIAL SHEET-2

Presidency University, Bengaluru

Welcome to the exciting and rapidly evolving world of Internet and web programming!

The Client-Server Model

The web is sometimes referred to as a client-server model of communications.

In the client-server model, there are two types of actors: clients and servers.

The server is a computer agent that is normally active 24 hours a day, 7 days a week, listening for queries from any client who make a request.

A client is a computer agent that makes requests and receives responses from the server, in the form of response codes, images, text files, and other data.

Client machines are the desktops, laptops, smart phones, and tablets you see everywhere in daily life.

These machines have a broad range of specifications regarding operating system, processing speed, screen size, available memory, and storage.

In the most familiar scenario, client requests for web pages come through a web browser.

The essential characteristic of a client is that it can make requests to particular servers for particular resources using URLs and then wait for the response.

These requests are processed in some way by the server.

The server in this model is the central repository, the command center, and the central hub of the client-server model.

It hosts web applications, stores user and program data, and performs security authorization tasks.

The essential characteristic of a server is that it is listening for requests, and upon getting one, responds with a message.

The exchange of information between the client and server is often referred as the request-response loop.

Server Types

- Web servers. A web server is a computer servicing HTTP requests. This typically refers to a computer running web server software such as Apache or Microsoft IIS (Internet Information Services).

- Application servers. An application server is a computer that hosts and executes web applications, which may be created in PHP, ASP.NET, Ruby on Rails, or some other web development technology.

- Database servers. A database server is a computer that is devoted to running a Database Management System (DBMS), such as MySQL, Oracle, or SQL Server, that is being used by web applications.

- Mail servers. A mail server is a computer creating and satisfying mail requests, typically using the Simple Mail Transfer Protocol (SMTP).

- Media servers. A media server (also called a streaming server) is a special type of server dedicated to servicing requests for images and videos. It may run special software that allows video content to be streamed to clients.

Web Basics

a web page is nothing more than an HTML (HyperText Markup Language) document (with the extension .html or .htm) that describes to a web browser the document's content and structure.

Hyperlinks

HTML documents normally contain hyperlinks, which, when clicked, load a specified web document.

Both images and text may be hyperlinked.

When the mouse pointer hovers over a hyperlink, the default arrow pointer changes into a hand with the index finger pointing upward.

Often hyperlinked text appears underlined and in a different color from regular text in a web page.

URL Formats

All URLs have the same general format:

The first part of the URL is the protocol that we are using.

```
http://www.funwebdev.com/index.php?page=17#article
```
Protocol   Domain   Path   Query String   Fragment

The domain identifies the server from which we are requesting resources. Since the DNS system is case insensitive, this part of the URL is case insensitive. Alternatively, an IP address can be used for the domain.

The path is a familiar concept to anyone who has ever used a computer file system. The root of a web server corresponds to a folder somewhere on that server.

*Query strings is a way of passing information such as user form input from the client to the server. In URLs, they are encoded as key-value pairs delimited by "&" symbols and preceded by the "?" symbol.

The last part of a URL is the optional fragment. This is used as a way of requesting a portion of a page.

*[will be explained in later sessions]

The Hypertext Transfer Protocol

All Web communications transactions use the same protocol: the Hypertext Transfer Protocol (HTTP).

HTTP consists of two phases: the request and the response.

Each HTTP communication (request or response) between a browser and a Web server consists of two parts: a header and a body.

The header contains information about the communication

The body contains the data of the communication if there is any.

The HTTP establishes a TCP connection on port 80 (by default).

The HTTP protocol defines several different types of requests, each with a different intent and characteristics.

The most common requests are the GET and POST request.

GET request
In this request one is asking for a resource located at a specified URL to be retrieved. Whenever you click on a link, type in a URL in your browser, or click on a book mark, you are usually making a GET request.

POST request

This method is normally used to transmit data to the server using an HTML form. A post request sends form data as part of the HTTP message, not as part of the URL. since the data is not transmitted in the URL, it is seen to be a safer way of transmitting data

Response codes are integer values returned by the server as part of the response header. These codes describe the state of the request, including whether it was successful, had errors, requires permission, and more.

200: OK :-The 200 response code means that the request was successful.

301: Moved Permanently:- Tells the client that the requested resource has permanently moved. Codes like this allow search engines to update their databases to reflect the new
location of the resource. Normally the new location for that resource is returned in the response.

304: Not Modified:- If the client so requested a resource with appropriate Cache-Control
headers, the response might say that the resource on the server is no newer than the one in the client cache. A response like this is just a header, since we expect the client to use a cached copy of the resource.

401: Unauthorized:- Some web resources are protected and require the user to provide credentials to access the resource. If the client gets a 401 code, the request will have
to be resent, and the user will need to provide those credentials.

404: Not found:- 404 codes are one of the only ones known to web users. Many browsers will display an HTML page with the 404 code to them when the requested resource was not found.

414: Request URI too long:- URLs have a length limitation, which varies depending on the server software in place. A 414 response code likely means too much data is likely trying to be submitted via the URL.

307: Temporary redirect :-This code is similar to 301, except the redirection should be considered temporary.

400: Bad Request:- If something about the headers or HTTP request in general is not correctly adhering to HTTP protocol, the 400 response code will inform the client.

500: Internal server error:- This error provides almost no information to the client except to say the server has encountered an error.

advantages of web applications include:

- Accessible from any Internet-enabled computer.

- Usable with different operating systems and browser applications.

- Easier to roll out program updates since only software on the server needs to be updated and not on every desktop in the organization.

- Centralized storage on the server means fewer security concerns about local.

Some of these disadvantages include:

- Requirement to have an active Internet connection (the Internet is not always available everywhere at all times).

- Security concerns about sensitive private data being transmitted over the Internet.

- Concerns over the storage, licensing, and use of uploaded data.

- Problems with certain websites on certain browsers not looking quite right.

Static Websites versus Dynamic Websites

Static website consists only of HTML pages that look identical for all users at all times.

dynamic website page content is being created at run time by a program created by a programmer; this page content can vary from user to user.

Client-Side Scripting versus Server-Side Scripting

Client-side scripting:- method of interacting with web browser

Client-side scripting with JavaScript can be used to validate user input, to interact with the browser, to enhance web pages, and to add client/server communication between a browser and a web server.

server-side scripts:- allow user to interact with server

CSE 256
INTERNET TECHNOLOGIES

TUTORIAL SHEET-3

Presidency University, Bengaluru

Where Is the Internet?

The global network of networks is implemented via millions of miles of copper wires and fiber optic cables, as well as via hundreds of thousands or even millions of server computers and probably an equal number of routers, switches, and other networked devices, along with many thousands of air conditioning units and specially constructed server rooms and buildings.
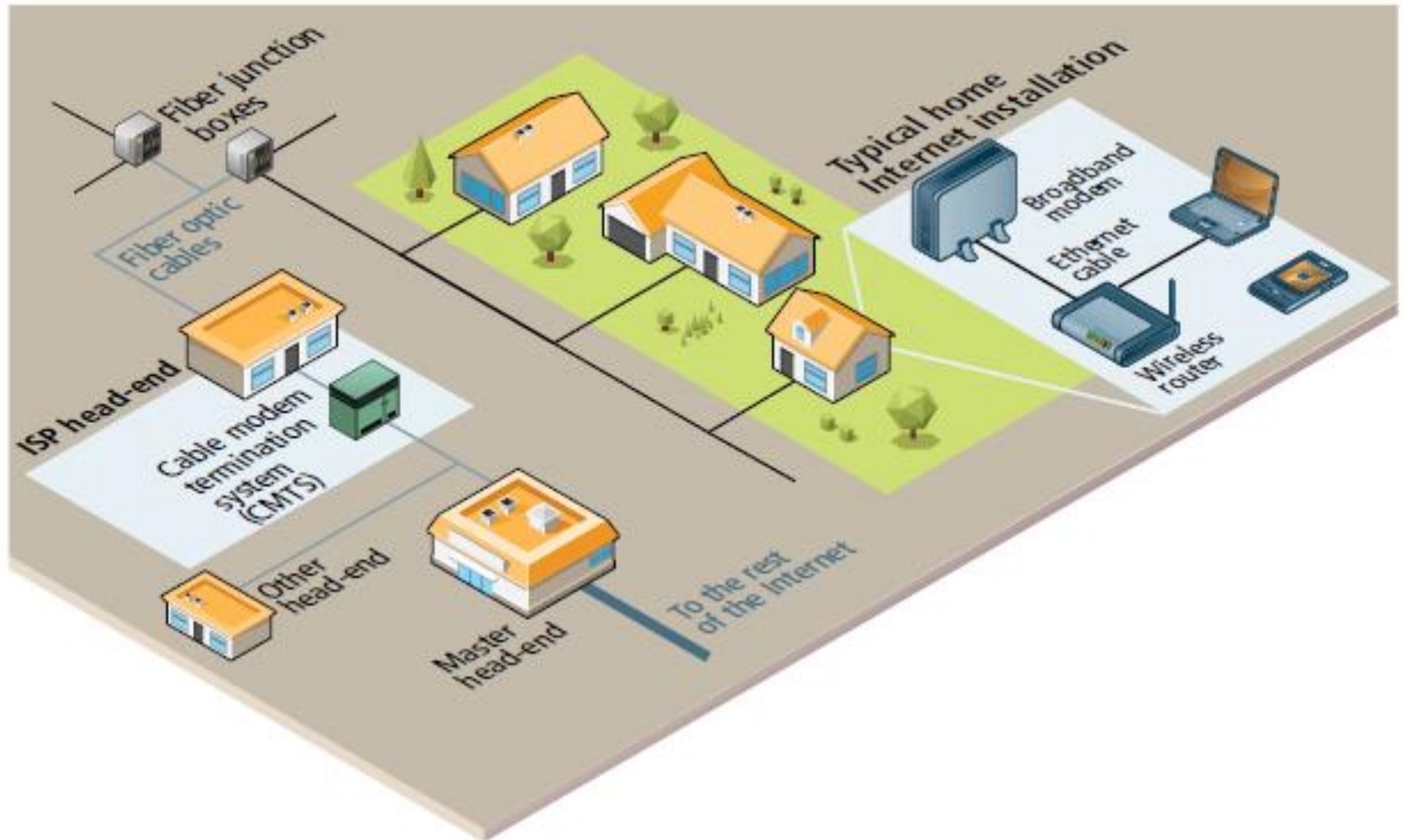
The **broadband modem** (also called a cable modem or DSL modem) is a **bridge** between the network hardware outside the house (typically controlled by a phone or cable company) and the network hardware inside the house.

These devices are often supplied by the ISP.

2. router

The wireless router is perhaps the most visible manifestation of the Internet in one's home, in that it is a device we typically need to purchase and install.

Routers are in fact one of the most important and ubiquitous hardware devices that make the Internet work.

At its simplest, a router is a hardware device that forwards data packets from one network to another network.

When the router receives a data packet, it examines the packet's destination address and then forwards it to another destination by deciding the best path to send the packets.

3. Fiber optic cable

Fiber optic cable (or simply optical fiber) is a glass-based wire that transmits light and has significantly greater bandwidth and speed in comparison to metal wires.

In some cities (or large buildings), you may have fiber optic cable going directly into individual buildings; in such a case the fiber junction box will reside in the building.

These fiber optic cables eventually make their way to an ISP's head-end, which is a facility that may contain a cable modem termination system (CMTS) or a digital subscriber line access multiplexer (DSLAM) in a DSL-based system.

Internet Protocols- An OS abstraction

The Internet exists today because of a suite of interrelated communications protocols.

A protocol is a set of rules that partners in communication use when they communicate.

We have already mentioned one of these essential Internet protocols, namely TCP/IP.

These protocols have been implemented in every operating system

Layered architecture- A quick revision
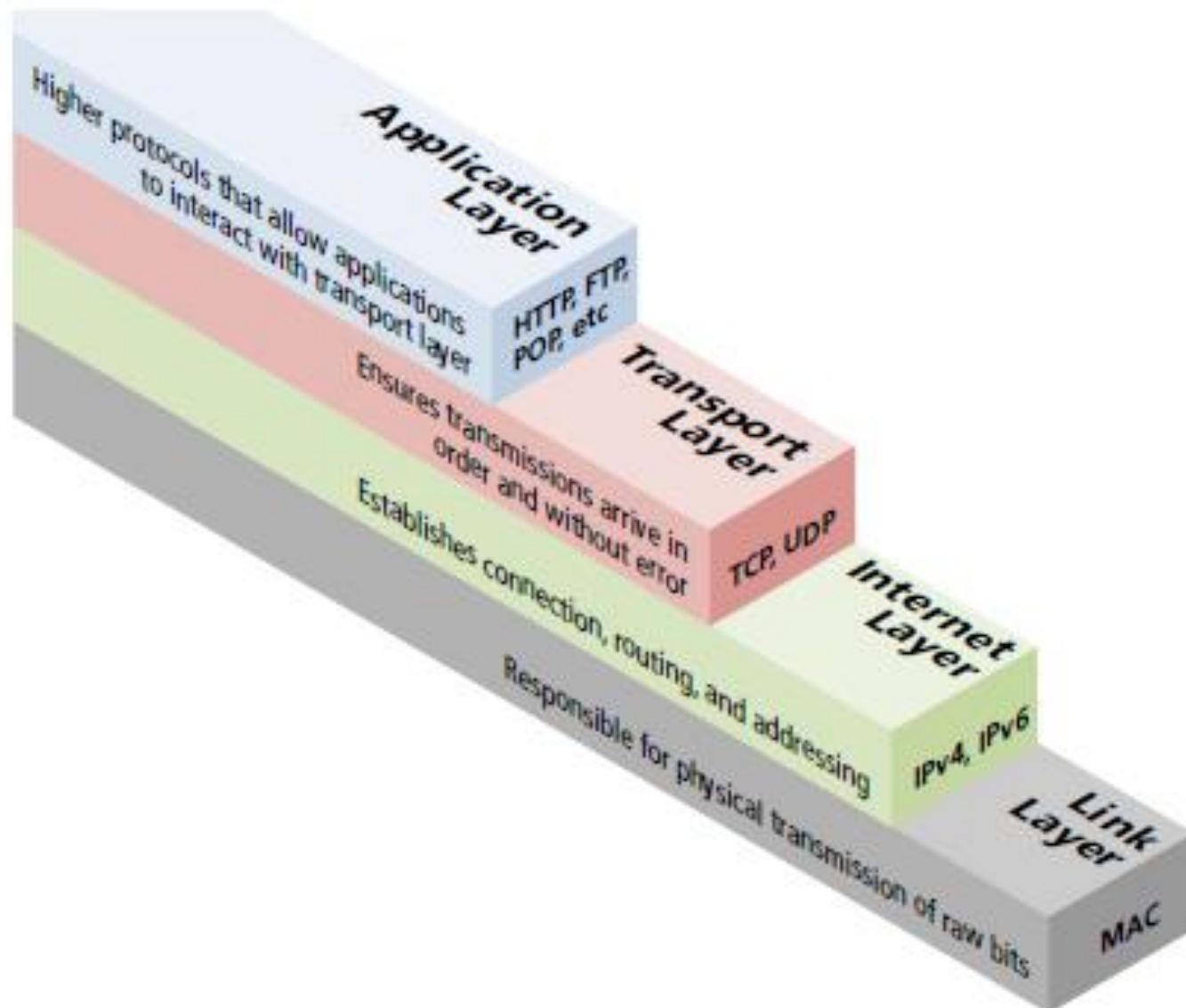
link layer

The link layer is the lowest layer, responsible for both the physical transmission across media (wires, wireless) and establishing logical links.

It handles issues like packet creation, transmission, reception, error detection, collisions, line sharing

The MAC (media access control) addresses:- unique 48- or 64-bit identifiers assigned to network hardware and which are used at the physical networking level

The Internet layer

The Internet layer (sometimes also called the IP Layer) routes packets between communication partners across networks.

The Internet layer provides "best effort" communication.

It sends out the message to the destination, but expects no reply, and provides no guarantee the message will arrive intact, or at all.

The Internet uses the Internet Protocol (IP) addresses to identify destinations on the Internet.

Every device connected to the Internet has an IP address, which is a numeric code that is meant to uniquely identify it.

The transport layer

The transport layer ensures transmissions arrive in order and without error.

The data is broken into packets formatted according to the Transmission Control Protocol (TCP).

Each data packet has a header that includes a sequence number, so the receiver can put the original message back in order, no matter when they arrive.

Each packet is acknowledged back to the sender so in the event of a lost packet, the transmitter will realize a packet has been lost since no ACK arrived for that packet.

That packet is retransmitted, and although out of order, is reordered at the destination

Application layer

Application layer protocols implement process-to-process communication and are at a higher level of abstraction in comparison to the low-level packet and IP address protocols in the layers below it.

There are many application layer protocols. A few that are useful to web developers include:

■ HTTP. The Hypertext Transfer Protocol is used for web communication.
■ SSH. The Secure Shell Protocol allows remote command-line connections to servers.
■ FTP. The File Transfer Protocol is used for transferring files between computers.
■ POP/IMAP/SMTP. Email-related protocols for transferring and storing email.
■ DNS. The Domain Name System protocol used for resolving domain names to IP addresses.

CSE 256
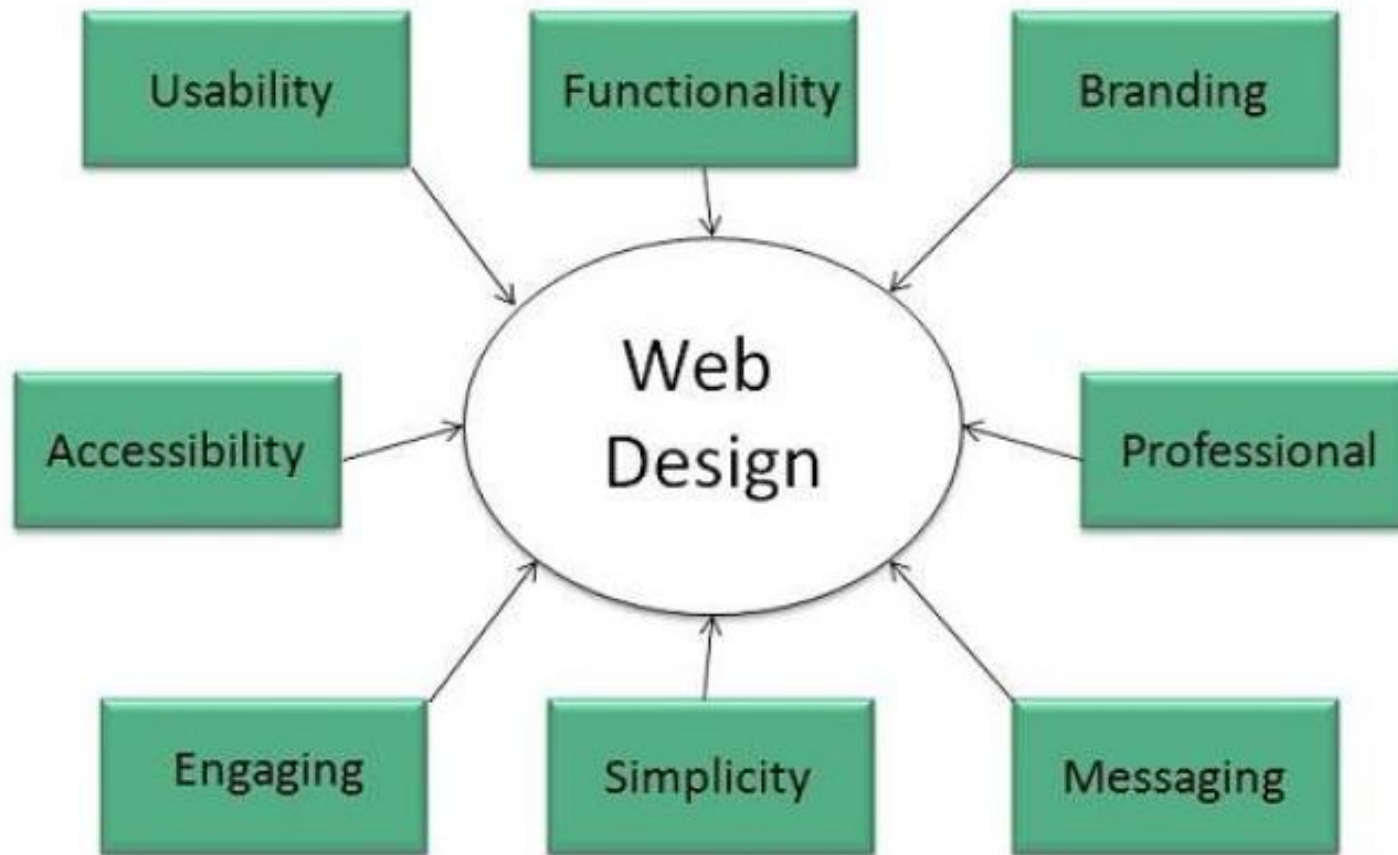# INTERNET TECHNOLOGIES

TUTORIAL SHEET-4

Presidency University, Bengaluru

Website Designing

Website Designing

Web designing has direct link to visual aspect of a web site. Effective web design is necessary to communicate ideas effectively.

Web Page Anatomy

Containing Block

Container can be in the form of page's body tag, an all containing div tag. Without container there would be no place to put the contents of a web page.

Logo

Logo refers to the identity of a website and is used across a company's various forms of marketing such as business cards, letterhead, broachers and so on.

Naviagation

The site's navigation system should be easy to find and use. Often the navigation is placed right at the top of the page.

# Web Page Anatomy

Content

The content on a web site should be relevant to the purpose of the web site.

Footer

Footer is located at the bottom of the page. It usually contains copyright, contract and legal information as well as few links to the main sections of the site.

Whitespace

It is also called as negative space and refers to any area of page that is not covered by type or illustrations.

LOGO

Navigation

Content

White Space

Footer
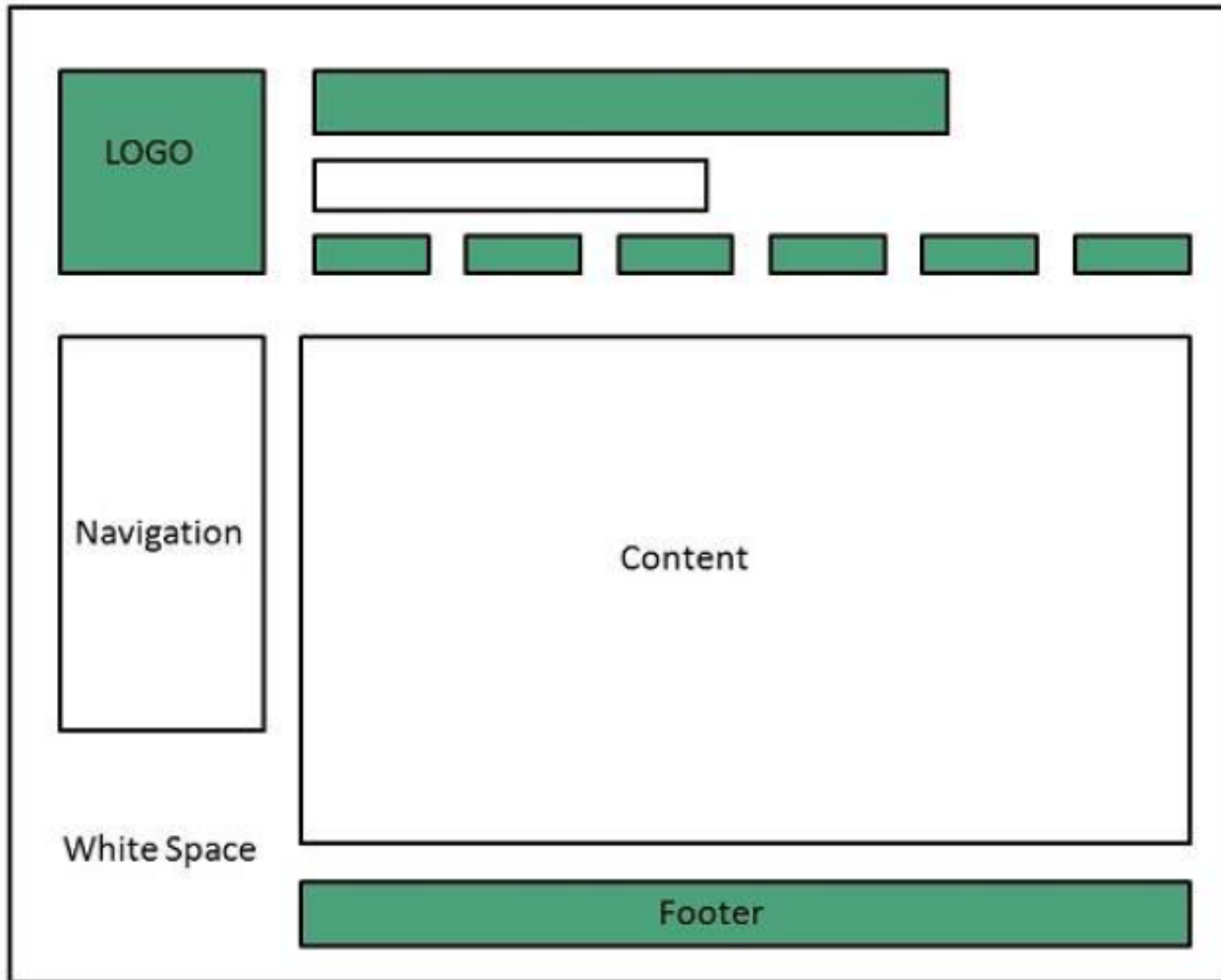
Web development

Web development refers to building website and deploying on the web. Web development requires use of scripting languages both at the server end as well as at client end.

Before developing a web site once should keep several aspects in mind like:
- What to put on the web site?
- Who will host it?
- How to make it interactive?
- How to code it?
- How to create search engine friendly web site?
- How to secure the source code frequently?
- Will the web site design display well in different browsers?

- Will the navigation menus be easy to use?
- Will the web site loads quickly?
- How easily will the site pages print?
- How easily will visitors find important details specific to the web site?
- How effectively the style sheets be used on your web sites?

Web Development Process

Web development process includes all the steps that are good to take to build an attractive, effective and responsive website.
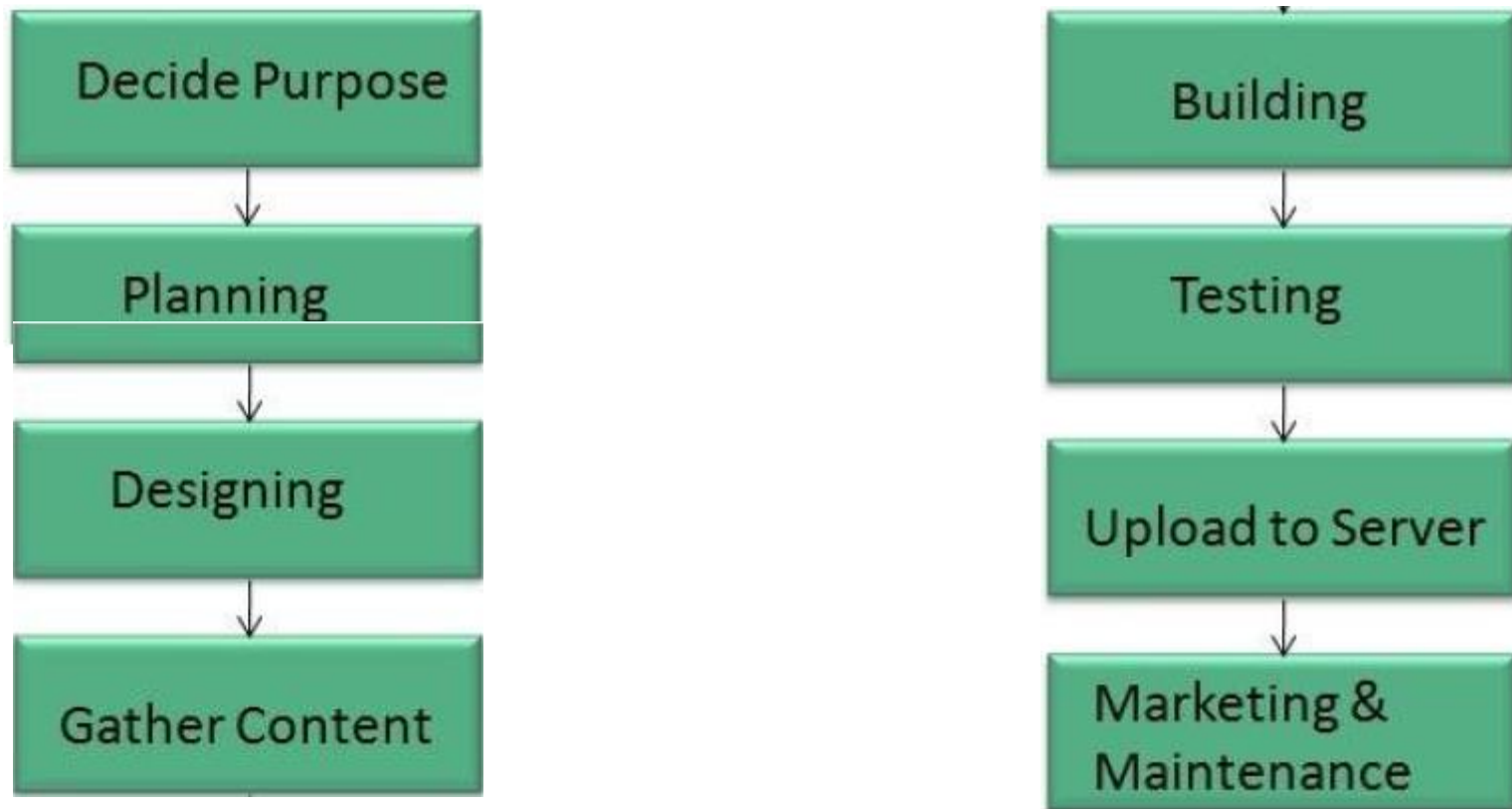
Websites Hosting

Web hosting is a service of providing online space for storage of web pages. These web pages are made available via World Wide Web.

The companies which offer website hosting are known as Web hosts.

The servers on which web site is hosted remain switched on 24 x7.

These servers are run by web hosting companies.
Each server has its own IP address.

Since IP addresses are difficult to remember therefore, webmaster points their domain name to the IP address of the server their website is stored on.

Types of Hosting

Shared Hosting

In shared hosting, the hosting company puts thousand of website on the same physical server.

Each customer has their own allocation of physical web space and a set of bandwidth limit.

As all websites share same physical memory, MYSQL server and Apache server, one website on the server experiencing high traffic load will affect performance of all websites on the server.

Virtual Private Server VPS

It is also known as Virtual Dedicated Server. It is a server which is partitioned into smaller servers.

In this customer is given their own partition, which is installed with its own operating system.

Unlike shared hosting, VPS doesn't share memory or processor time rather it allocates certain amount of memory and CPU to use which means that any problem on a VPS partition on the same drive will not affect other VPS customers.

Dedicated Server

In this kind of hosting, single dedicated server is setup for just one customer.

It is commonly used by the businesses that need the power, control and security that a dedicated server offers.

Reseller Hosting

A reseller acts as a middle man and sells hosting space of someone else's server.

Grid Hosting

Instead of utilizing one server, Grid Hosting spreads resources over a large number of servers. It is quite stable and flexible. The servers can be added or taken away from the grid without crashing the system.

Website Security Considerations

Websites are always to prone to security risks.

Cyber crime impacts your business by hacking your website.

Your website is then used for hacking assaults that install malicious software or malware on your visitor's computer.

Updated Software

It is mandatory to keep you software updated. It plays vital role in keeping your website secure.

Security Considerations

SQL Injection

It is an attempt by the hackers to manipulate your database.

It is easy to insert rogue code into your query that can be used to manipulate your database such as change tables, get information or delete data.

Cross Site Scripting XSS

It allows the attackers to inject client side script into web pages. Therefore, while creating a form It is good to endure that you check the data being submitted and encode or strip out any HTML.

Security Considerations

Error Messages

You need to be careful about how much information to be given in the error messages.

For example, if the user fails to log in the error message should not let the user know which field is incorrect: username or password.

Validation of Data

The validation should be performed on both server side and client side.

Security Considerations

Passwords

It is good to enforce password requirements such as of minimum of eight characters, including upper case, lower case and special character. It will help to protect user's information in long run.

Upload files

The file uploaded by the user may contain a script that when executed on the server opens up your website.

SSL

It is good practice to use SSL protocol while passing personal information between website and web server or database.