<u>UNIT V</u>

Internet and Website Establishment: Introduction – Technologies for web servers – Internet tools relevant to Commerce – Internet Applications for Commerce – Internet charges – Internet Access and Architecture – Searching the Internet- Case study.

2 MARKS

1. List the web browser features.

- Forms completion.
- E-mail support.
- Date inquiry.
- Customize by user.
- Rapid integration with other applications.
- Hypertext.

2. List the internet application for commerce.

- Direct selling or marketing of a company's existing products and services.
- Selling advertising space.
- Charging fees for the actual content accessible on a web site.
- Charging fees for online transaction or links.

3. Mention the 3 ways to provide information as home pages on the web.

Roll on your own- It is a low cost entry level strategy to create your own WWW pages and post them on the web.

Outsource- for many business working with a commercial web services provider is a good choice. You can create pages, get the assistance of a consultant or an market agent, then lease commercial web pages.

In-house development- Business with the resources to hire or train the necessary personnel can set up web servers within their business. It provides the best control over access, design, and content.

4. What is a routing arbiter?

A Routing Arbiter (RA) is an element that is introduced into the NAP architecture. RA organization implements the concept of policy-based network routing that enables routing of traffic between different network operators.

5. List the effects of Routing Arbiter.

The effects of the RA as follows:

- Route servers
- Network management systems
- Routing arbiter database
- Routing engineering

6. Mention the processing components of gathering information.

The process consists of 3 main components:

- Sourcing information
- Index, catalog, or database creation
- The search engine

7. What is a search engine spider?

A spider also known as a robot or a crawler is actually just a program that follows or crawls links throughput the internet, grabbing content from sites and adding it to search engine indexes.

8. List the uses spiders. (Nov 2012)

A spider is a program that autonomously explores the Web and newsgroups and takes some action upon the information it finds. This action may be as counting the number of web links found or as complex as indexing the entire text of a Web page or newsgroup.

The primary uses of spiders are as follows.

- Link validation
- HTML validation
- "What's new" monitoring
- Indexing

9. Define Multithreading?

Multithreading is a sophisticated and refers to the ability to support paths of execution within a single address space. Older operating systems achieve multitasking by creating multiple processes, which creates a great deal of overhead.

10. Describe the benefits of Frame relay?

Frame relay has two benefits:

- **Speed:** It is no longer necessary to carryout error controls and corrections between each node due to the improvements in transmission media.
- Sharing costly bandwidth: Frame relay allows users to share costly, high throughput channels over a single access line, and it uses a "hubbing" approach to distribute traffic over a wide area.

11. Define Multiprocessing?

Multiprocessing is defined as the ability to support the concurrent execution of several tasks on multiple processors. This implies that ability to use more than one CPU for executing programs. The processors can be tightly or loosely coupled.

12. Define multitasking?

Multitasking means that the server operating systems can run multiple programs and give the illusion that they are running simultaneously by switching control between them.

Two types of multitasking are used:

- Preemptive
- No preemptive

13. List some services of the internet?

Some services of the internet are:

- Individual to group communications
- Information Transfer and delivery services
- Information Databases
- Information processing services
- Resource-sharing services

14. What are all the key search features in search engine?

The following lists key technology features that should be included in search engines.

- Free-text search
- Automatic morphology
- Word indexing
- Lexical affinity search
- Ranking and relevance scoring

15. What is InfoSeek?

The search engine accepts natural language questions, such as, "What is a good restaurant in New York City that serves fresh morels?"But also provides keyword search capability. InfoSeek returns the most relevant matches.

16. What is WebCrawler?

WebCrawler is faster than InfoSeek and lets the user view more matches on the screen simultaneously. It does not screen its matches for relevance as much as InfoSeek does; hence users may find matches that InfoSeek omitted.

17. What is Lycos?

Lycos is the slowest of the "big three," but it is also the most detailed. The Lycos robot indexes not only the sites that it visits, but also every link on those sites.

18. What is Savvy Search?

One can search the entire previously mentioned database simultaneously using a tool called Savvy Search. It is not a search engine itself rather; it is a client that executes searches on several engines in parallel.

19. What is USENET?

If you are searching for something current, then Usenet is an excellent resource. Usenet is not a network and Usenet does not even need the Internet. Rather, what drives Usenet is akin to an agreement set up between those who want to distribute and those who want to read newsgroups.

20. What is Veronica?

Veronica is a self-updating database of Gopher documents. Veronica allows users to search all of the Gopher sites in the world. By entering a word or words, directories, programs, and articles with those words in them will show up in a menu for users to browse. For purposes of a Veronica search, the title is the name of the resource as listed on its home Gopher server.

21.What is meant by NAP access?(Apr 2014)

NAP access:

Sprint supports access from an NSP's router located at Sprint's physical NAP facility. The NSP can supply and maintain the router which interfaces directly with Sprint's NAP, or it can choose a sprint-supplied router.

22.Write a note on internet browser?(Apr 2014)

Although **browsers** are primarily intended to use the World Wide Web, they can also be used to access information provided by web servers in private networks or files in file systems. The major web **browsers** are Firefox, **Internet**Explorer, Google Chrome, Opera, and Safari.

23. Define the term Internet. (Apr 2012)

A global computer network providing a variety of information and communication facilities, consisting of interconnected networks using standardized communication protocols.

The Internet is the global system of interconnected computer networks that use the Internet protocol suite (TCP/IP) to link billions of devices worldwide.

24. What are the two most common publishing systems? (Nov 2012)

The electronic publishing process follows a traditional publishing process but differs . Ink-jet or laser printer or via a print on demand system. Three attributes of digital technology: XML tags to define content, style .The most common file format is .epub, used in many e-book .A desktop publishing system allows you to use different typefaces, specify various margins.

25.Write about electronic payment system. (Apr 2013)

An e-commerce payment system facilitates the acceptance of electronic payment for online transactions. Also known as a sample of Electronic Data Interchange (EDI), e-commerce payment systems have become increasingly popular due to the widespread use of the internet-based shopping and banking.

26.What is SOA GOVERNANCE? (Apr 2013)

SOA governance is a concept used for activities related to exercising control over services in serviceoriented architecture (SOA) solutions. One viewpoint, from IBM and others, is that SOA governance is an extension (subset) of IT governance which itself is an extension of corporate governance.

27. Write short notes on Internet & Intranets.(Nov 2014)

Internet

A global computer network providing a variety of information and communication facilities, consisting of interconnected networks using standardized communication protocols.

Intranet

A local or restricted communications network, especially a private network created using World Wide Web software.

28.Mention Some Internet Tools.(Nov 2014)(Apr 2015)

- Telnet
- E-Mail
- FTP
- Gopher
- World Wide Web

29. Comparison between Website & Webpage.(Apr 2015)

Website

A site (location) on the World Wide Web. Each Web site contains a home page, which is the first document users see when they enter the site. The site might also contain additional documents and files. Each site is owned and managed by an individual, company or organization.

Webpage

A web page *or* webpage is a document commonly written in HyperText Markup Language (HTML) that is accessible through the Internet or other network using a browser. A web page is accessed by entering a URL address and may contain text, graphics, and hyperlinks to other web pages and files.

11 MARKS

1. Describe about Browsing and providing Information. (or) Internet Charges.

INTERNET CHARGES:

Browsing for information:

A user needs to get the computer server and software choose and ISP, and obtain a suitable connection. A machine is required for the user to access the internet the machine could be an internet terminal a PC or a high speed server depending on what type of info. A person needs, and how much the person is willing to pay.

With a PC a user the user can be connected and configured as a client on the internet by a low or high speed connection. Depending on power and capacity the user wants the cost of the PC ranges from a couple hundred dollars to a couple thousand dollars. There are many options to choose for internet connections.

Browsing and providing information:

There are three ways to provide information as home pages on the web. They are

Roll on your own- It is a low cost entry level strategy to create your own WWW pages and post them on the web.

Outsource- for many business working with a commercial web services provider is a good choice. You can create pages, get the assistance of a consultant or an market agent, then lease commercial web pages.

In-house development- Business with the resources to hire or train the necessary personnel can set up web servers within their business. It provides the best control over access, design, and content.

Settlements:

The financial settlement between internet providers is an important issue. This is clear when you understand that the internet is no more than a set of local and long distance networks. The internet is WWW and many we servers and internet hosts are outside of the united states. The settlements will require complex accounting mechanisms whose costs might be passed on the end user.

2. What is the internet architecture? Or Explain the Internet Access and Architecture with neat Diagram. Or Explain about E- Commerce and internet access in detail. (April 2013)

INTERNET ACCCESS AND ARCHITECTURES:

It is define a Meta network a constantly changing collection of thousands of individual networks intercommunicating with a common protocol.

The internet architecture is described in it name, a short from of the compound word internetworking. This architecture is based in the very specification of the standard TCP/IP protocol designed to connect any two networks which may be very different in internal hardware,

software and technical design. Once two networks are interconnected communication with TCP/IP is enabled end-to-end, so that any node on the internet has the near magical ability to communicate with

any other no matter where they are. This openness of design has enabled the internet architecture to grow to a global scale.

Routing arbiters:

The routing arbiters to provide stable coherent routing in the internet with the internet doubling every 13 months this is not as easy to it might be. The problem is compounded by the withdrawal of the NFSNET service and the proliferation of internet service providers and exchange points. A brief view from the RA perspective is given with some attention to tools and techniques that will facilitate the contained growth of the internet in size, features and function.

The effects of the RA as follows:

- Route servers
- Network management systems
- Routing arbiter database
- Routing engineering

Example of NAP architecture

As an illustrate example, sprint chose FDDI as an initial architecture for its NAP; sprint has planned to upgrade to an ATM switch in the near future. In the interim, switched FDDI hubs may be deployed when the NAP loads are projected to exceed what the FDDI ring can support. An Ethernet interface is also available for use by any NSP that requires that type of interface.



NAP access:

Sprint supports access from an NSP's router located at Sprint's physical NAP facility. The NSP can supply and maintain the router which interfaces directly with Sprint's NAP, or it can choose a sprint-supplied router. This approach allows NSPs to connect to the sprint NAP via a serial link or a switched service from dispersed locations using any type of WAN technology the NSP choose to use. WAN technologies supported include the following:

- Dedicated circuit access
- Switched service access

Dedicated circuit access:

For a dedicated circuit connection from an NSP to the Sprint NAP, the NSP can choose to provide the link via Sprint, a local exchange carrier (LEC), a competitive LEC (CLEC), or another interexchange carrier in conjunction with either a LEC or a CLEC.



Switched service access:

The Sprint NAP may also be accessed by any one of three switched services: ATM, frame relay, or SMDS.

ATM- The NSP connects directly via a UNI to Sprint's ATM network. When broadband intercarrier interface (B-ICI) is available, an NSP may also use it to connect another carrier's ATM network to a sprint's network in order to reach the NAP.

Frame relay. The NSP connects directly via a frame relay link or a frame relay network to a Sprint-provided router. A NSP may also use B-ICI to connect another carrier's frame relay network to Sprint's ATM network.

SMDS (Switched Multi-Megabit Data Service). The NSP connects directly via an SMDS link or network to a sprint-provided router. An NSP may use B-ICI to connect another carrier's frame relay network to Sprint's ATM network.



3.Describe about searching operations in Internet and its Limitations.

SEARCHING THE INTERNET:

As the amount of content on-line increases, it becomes more and more difficult for users to find what they are looking for on the Internet.

Gathering information:

The process consists of 3 main components:

- Sourcing information
- Index, catalog, or database creation
- The search engine

SPIDERS AND SEARCH ENGINE:

A spider also known as a robot or a crawler is actually just a program that follows or crawls links throughput the internet, grabbing content from sites and adding it to search engine indexes.

Spiders only can follow links from one page to another and from one site to another. That is the primary reason why links to your site are so important links to your website from other website will give the search engine spiders more food to chew on. The more time they find links to your site the more times they will stop by and visit. Google especially relies on its spiders to create their vast index of listings.

Spider find web pages by following links from other web pages, but you can also submit your web pages directly to a search engine or directory and request a visit by their spider. In fact it a good idea to manually submit your site a human edited directory such as yahoo and usually spiders from other search

engines will find it and add it to the database. It can be useful to submit your URL straight to the various search engines as well, but search cheating.

Uses of spiders:

A spider is a program that autonomously explores the Web and newsgroups and takes some action upon the information it finds. This action may be as counting the number of web links found or as complex as indexing the entire text of a Web page or newsgroup.

The primary uses of spiders are as follows.

- Link validation
- HTML validation
- "What's new" monitoring
- Indexing

An advantage of using spiders is that databases can be updated automatically, allowing users to be reasonably sure they are receiving the latest information.

Limitations of spiders:

The information databases complied by spiders are useful; however, there are some limitations to spiders concerning resource discovery. The main limitation is that there is just too much information available through the Internet and the amount is growing all the time. Although spider-generated databases are generally more current in totality than manually complied databases, the fact that new information is being added every day precludes the spider-generated database from being completely up-to-date. Also, adding to the problem is the fact that a large portion of the data on the Internet is dynamic in nature.

To address redundant and irrelevant database listing developed by spiders, a standard for spider exclusion has been developed. This standard describes the use of simple structured text files at an easily found place on a server to specify which parts of a URL space should be avoided by spiders.

This process can be used to give specific instructions to individual spiders, given that some behave more sensibly than others or are known to specialize in particular areas. Currently, this standard is voluntary, but there is pressure from various Internet agencies for users of spiders to comply.

Search engines

A search engine is a program that searches through a database. In the context of the Internet, search engine is most often equated to search forms that request programs to look through databases of HTML documents gathered by spiders or through manual data gathering processes.

Accessing the information databases- An HTTP server provides a method of running database searches from within an HTML document. This method is called the Common Gateway Interface (CGI). The server can pass information to the CGI program and get back any results returned by the database search.



High-level details of using CGI and gateway software- Processing software is called directly from the HTML search document by the HTTP client. The processing software reads the query input passes to it via the CGI call.

CGI use:



Multithread queries:



Multithreading queries are performed by using a gateway program which conforms to CGI. The program works with an HTML form describing the various search engines and the type of queries that can be performed. When users wish to perform a query, they retrieve the HTML form, select the desired search engines, fill in the query parameter, and submit the form. The query is then submitted to an HTTPD (HTTP daemon) process which starts the multi-threaded query gateway giving the query information.

When a multithreaded gateway receives the query, it creates a thread for each search engine. Each thread then sends the query to the gateways using HTTP and waits for the response. When responses have been received from all search engines, the multithreaded gateway creates an index bringing all responses into one document which is sent to the user.



What if a search engine was programmed to boost the ratings of sites with a particular political

inclination? People would for ex, see a lot of sites at the top of the list advocating a certain political stance which contrary pages could be ranked way down where most people would not see them.

The search terms entered by users also provide an immediate snap shot of people thoughts. Google has a whimsical feature that tracks how often different topics are searched for such information would be invaluable to adventures, politicians, newspaper reporters, commercial organizations etc because they would get a clue about what hot people use indexes of searches terms as indicators of frame is Madonna being searched for less often? Is Britney spears on the way up or down in her career? You can often get clues from the trends in search terms submitted to search engines.

Different search engines handle moral issues differently. Some reward sponsors by inserting their link into search results regardless of words used in the search terms. Often these sponsored links are not identified as such leaving people to wonder why bobs hardware appeared in the results of their search for Britney spears.

Other search engines pride themselves on more reputable behaviour. They only include sponsored links if they are relevant to a search Google goes further and keeps sponsored links separate links separate from search results and clearly identifies them as sponsored.

Many people try to work how different search engine calculate the positioning in search results. Most engines keep their formula secret to prevent people exploiting it to their sites popularity artificial inflated.

Tricks like including a lot of irrelevant keywords in the META tags have been popular with web authors. A site on a band for ex, would include irrelevant but popular keywords like Pamela Anderson, nude porn free. Many search engines now pay little attention to keywords for their reason. Google them completely.

Another trick is to repeat keywords many times on a page, often with the text made invisible by setting the text color to the page color. Again some search engines know this trick and will ignore blatant mass-listings.

Oddly Google will occasionally list page that don't have your search terms but other pages linked to that page with those words. This led to the Google bomb originally a fun trick but potentially a powerful technique to manipulate and distort Google results.

4. Briefly explain about search engine and its relevant ranking methods. (Nov 2014)

Social implications of search engines:

Sounds odd, doesn't? Social implications of search engines? You might as well talk about the social implications of screw drivers. They are just tools aren't they? In fact search engines have heavy moral and social responsibilities and can widely considerable influence. They believe that a person searching for info. On the church would be getting a satisfied and biased list of links. A bit of a controversy corrupted with people arguing that Google was censoring the internet.

Search engine relevance ranking methods:

Keyword matching: simply checks to see how many of the search terms appear on a page ex: if searching

for King Henry, a page with both the search terms would be judged more relevant than a page with only one of the terms.

Frequency: If a page mentioned king Henry frequently, it's a good sign the page is more valuable than another page a that just mentions him once.

Positioning:

Pages with king Henry n the page title or in a heading or formatted as hold would probably be more valuable than another page that did not emphasize the words or positioning them in a prominent place. A page that has the search terms near the beginning of the text would be more highly ranked than a page that had the search terms near the end of the text.

HTML web pages also have an invisible section, the head where META tags can be inserted by the pages author. If search terms appear in these tags it is a clue that the words are important to the page and it will tend to be ranked more highly.

There is a continuing battle to be the best search engine. While their services are free to users search engines make a great deal of money through advertising and sponsorship. Although they all do the same job, they use a variety if techniques to do it.

Some search engines also take into account whether as site is listed in prestigious listings. One ex is DMOZ a list of services selected by human editors. A sites inclusion in this list is an indicator of the quality of the site and search engine like Google take this us to account with TCP/IP is enabled end to end so that any node on the internet magical ability to communicate with any other no matter where they are.

The companies running the internet backbone operate very high bandwidth networks relied on by governments, corporations and large organizations and other internet service providers. As always a larger scale introduces new phenomena the no. of packets flowing through the switches on the backbones is so large that it exhibits the kind of complex non linear patterns found in natural analog systems like the flow of water or development of the rings of Saturn.

Key search features (in most search engines):

The following lists key technology features that should be included in search engines.

- Free-text search: The ability of the search engine to accept single words, phrases, or sentences. Common words are ignored, and the documents are elected based on the words in the query.
- Automatic morphology: Nouns and verbs in English occur in many tenses; the search process automatically generates the various forms of the words and uses in the query.
- Word indexing: Processing of collections of documents by building indexes which contain the location of every instance of every word. It is the indexes, not the documents that are searched.
- Lexical affinity search: Search on words that occur close to the text form of the query.
- Ranking and relevance scoring: Using a ranking algorithm, the search engine assigns every document that contains at least one of the queried words a numerical score.

These items do not by themselves guarantee a good search engine but assist users in better defining their informational requirements.

4.List the Search tools available in internet? (or) Internet Tools (Apr 2014).

Actual search tools:

Yahoo: This is currently considered the most complete directory resource on the WWW. It is a categorized, menu-driven directory of web resources.

InfoSeek: The search engine accepts natural language questions, such as, "What is a good restaurant in New York City that serves fresh morels?"but also provides keyword search capability. InfoSeek returns

the most relevant matches.

AltaVista: This is a search engine that is very likable. Often it may be all a user needs.

Open Text: This is a technically superior search engine on the Internet and produces fast, relevant responses.

WebCrawler: WebCrawler is faster than InfoSeek and lets the user view more matches on the screen simultaneously. It does not screen its matches for relevance as much as InfoSeek does; hence users may find matches that InfoSeek omitted.

Lycos: Lycos is the slowest of the "big three," but it is also the most detailed. The Lycos robot indexes not only the sites that it visits, but also every link on those sites.

Savvy Search: One can search the entire previously mentioned database simultaneously using a tool called Savvy Search. It is not a search engine itself rather; it is a client that executes searches on several engines in parallel.

Galaxy: Galaxy searches for resources on the Web, including Gopher and Telnet.

Usenet: If you are searching for something current, then Usenet is an excellent resource. Usenet is not a network and Usenet does not even need the Internet. Rather, what drives Usenet is akin to an agreement set up between those who want to distribute and those who want to read newsgroups.

Archie: If you are searching for computer software, then Archie is an excellent resource for navigating through the content on anonymous FTP archives throughout the world.

Veronica: Veronica is a self-updating database of Gopher documents. Veronica allows users to search all of the Gopher sites in the world. By entering a word or words, directories, programs, and articles with those words in them will show up in a menu for users to browse. For purposes of a Veronica search, the title is the name of the resource as listed on its home Gopher server.

Gopher: Gopher is a service that allows users to view and obtain files, programs, and software. When accessing Gopher information with a graphic browser, users are presented with a series of menu choices, much like the directory structure one sees in the Microsoft Windows File Manager.

Ftp: Anonymous FTP permits users to access remote systems without actually having user accounts on the systems. Users can FTP to any anonymous site in the world using hypertext interface.

Telnet: Telnet allows users to remotely log in to other computers and can give them access to databases and many information services.

WAIS: WAIS lets the user search indexed information to find articles containing groups of chosen words. WAIS is like Gopher in that it shields users from having to know on which computer the information resides; but unlike Gopher, WAIS does the searching for the user.

Internet Directory Services: There are a number of ways to create resources that provide access to email addresses of individuals and institutions and contact information for Internet services. Directories to Internet addresses for people and services indexed by name are often referred to as white pages, while those directories that allow access to addresses of services by category are usually called yellow pages.

6.Explain in detail about internet applications for commerce?(Apr 2014)

A growing segment of the Internet is Electronic Commerceor E-Commerce. Consumers are looking for suppliers selling products and services on the Internet. Meanwhile, suppliers are looking for buyers to increase their market share. The vast amount of available information causes a great deal of problems for both ends. Searching, a task executed online, isnot only time consuming but boring as well. Intelligent Agents suited for this type of task.

A typical campus Internet Infrastructure is a collection of communication devices and networks of varying types and configurations connected together to form one network: the

"Internet". The tower and the high speed disk are used as a data warehouse for suppliers' products and consumers' needs and services. The remaining devices on the network provide user interface to the network clients.

Agents executing on such network will surf through the information on the Internet, resulting in the selection of specific information of interest to the user. Intelligent Agents provide a mechanism for conducting ECommerce. Using Intelligent Agents, the user composes the task offline, logs to the Internet to issue the task and logs off. At a later time the user can log to the Internet to collect the results of the task execution. Typical Internet Infrastructure

SIMULATION TEST

The NET-Computer was simulated in a network of three hosts: ainur, kira and yavanna connected. The hosts connected to networks A, B and C are located on different floors of the same building and run UNIX operating systems. Simulation test network The objective of the simulation test is to test the following Agents characteristics:

- Mobility: Agents navigate the network and reach destination address (host).
- Agency or task execution: Agents carry out useful work on behalf of the user.
- Intelligence: Agents follow a set of rules predefined by user.
 - 1. Direct Selling
 - 2. Selling and space
 - 3. Charging for content
 - 4. Charging for services

Direct selling

A gamut of companies has begun selling their products directly on the Web. The Web's reach can transform a small company into a global distributor. Large corporations that already have their distribution networks in place often find the Web to be a niche channel and many still think it is too early to be profitable. Some believe the reason is that "Internet selling is too new," others believe that people are not certain, and changing behavior takes some time.³¹ Smaller companies and socalled cyberpreneurs are moving faster and getting more substantive results. For industries overrun by corporate chains, the Web may help smaller companies level the playing field. An information-rich Web site can help specialist retailers provide the same services as a fancy store in a big city.³⁰

While the "no location" aspect of Web-based commerce can be advantageous, companies are also finding out that customers have a hard time discovering a particular site among the hundreds of thousands out there. As a result, more Web merchants are paying a sales commission as well as an advertising fee in exchange for prominent placement on high-traffic web sites, such as search engines and home pages of on-line service providers. For example, Amazon.com and 1-800-Flowers have entered into long-term exclusive agreements with America Online to gain access to the service's 8.5 million customers. These two 1997 agreements were valued at \$44 million in revenue. In exchange, AOL will provide premiere placement on AOL's orginal service or on its heavily trafficked Internet site. On-line users will be able to click on the ads and be connected to the Amazon.com or 1-800-Flowers sites. Stand-alone Web sites face challanges in attracting potential customers.

8.4.2 Selling ad space

Many companies have started advertising their products and services on the Web. It was estimated that companies spent \$10 million to advertise on the Web in 1995 and the figure has been larger in recent years.³² Companies selling ads on the Web say that this business model is a natural extension of their other lines of business. For instance, Career Mosaic functions as an on-line classified-ad and job database. Internet users going to the site can view ads organized by type of work and corporate profiles. The site lists several thousand jobs.¹⁹ There is no standard rate structure yet emerged for Internet advertisements. DealerNet charges a flat \$995 fee to put a car dealership on the Internet, plus a \$500/month maintenance fee.³²

4.3 Charging for content

Content providers charge for subscriptions and also count on advertising revenue from their sites. Observers believe that advertising drives business. Forrester expected companies would spend \$2.2 billion advertising on the Web by year 2000.³² A few daily newspapers have started charging for content. For instance, San Jose Mercury News began charging fees, ranging from about \$1 to \$5 per month, for its Mercury Center news site in 1995.³²

4 Charging for services

Another model involves charging for some type of services such ^{as} searching databases, providing space, linking, and other services ^{on a} Web site. *Industry.Net* offers business a place to shop for goods.^{II} Charges manufacturers and suppliers \$3000 to \$8000 a year to ^{main} tain an electronic storefront on its site. *Industry.Net* generated \$30^{min} tain in 1995. The business not only allows users to access the data^{bit} also allows them to place orders by e-mail.³⁰

7. What are the technologies used for web servers? Explain. (April 2012)(Nov 2014)(April 2015)

Designing a web presentation for a company or an organization requires or to sort out the content to present, set goals, decide on topics, then organize the layout and navigation of the web pages. As previously discussed, to publish documents on the web, a server that makes available documents and media to the browser that requests them in needed. Whenever the browser is pointing to a web documents, the browser communicates with the server to get at that document.

HTML:

General features: HTML is the markup language used to create web documents. It is loosely a subset of the SGML (standardized General Markup Language).SGML is used to describe the general structure of various kinds of documents. The primary focus of SGML, and therefore HTML, is the content of document, not its appearance. Browsers decode the HTML instructions and display the document on the requester's screen. The theory behind this is that most documents have common elements, for example, titles, paragraphs, or lists, and if these elements are defined, they can be labeled as the appropriate parts of the documents. The elements of the Web documents are labeled through the use of HTML tags. The following is a list of basic tags for HTML.

Tag	Use		
<html><\HTML></html>	The entire HTML document		
<head></head>	The head, or prologue, of the HTML document		
<center></center>	Center text		
<title></title>	The title of the document		
< <u>A</u> > <u A>	Illustrate reference and anchor portions		
	Make word boldface		
<i></i>	Italicize word		
<h4></h4>	Fourth-level heading		
<p></p>	Paragraph		
	Comment		

HTML does little to describe the exact placement or appearance of any element on the page, since there is no way of knowing what platform in the document going to be viewed on, the size of the screen, the font that are installed on the platform, or if there are any fonts at all. In addition to the creating tabs, the user may also create links between documents, create list set page breaks Identify addresses and quotations and inserts any special characters and colors.

Editors and converters are the programs that can help publishers to write HTML pages. These Programs tend to fall into two categories: Editors, in which HTML is directly written, and The converters which convert the output of some other word processing program into HTML.

Most of the programmers are essentially text editors with extra menu items of buttons that insert Appropriate HTML tags into the text.HTML based text editors are useful and easy to use because publishers do not have to remember tags and do not have to type them all. Some examples of editors are as follows:

Editor

Description

• HTML editor It is an application that allows the insertion of tags into a file and the result is in WYSIWYG fashion. The tags are shown in a lighter

color and the surrounding text. There are options to hide the tags in the documents to depict the full affect.

- Microsoft Internet It is a plug in for words for windows 6.0 that allows user to create Assistant HTML files in the word and then save them as HTML.
- Template packages For Microsoft word version 2.0 or 6.0 that allows user to assign for MS Word styles in a word document and select HTML features from a toolbar.
- WordPerfect It is plug in for word profile for windows. It include template for editing files for the web and a converter program that supports the ability to add links and convert the document to HTML.

These converters takes files from many popular word processing programs and convert them to HTML. Using these converters, a user may create documents in a program and then convert the results. Converters help put existing documents on the web quickly.

Tables are critical for summarizing information in a way that can be quickly and easily grasped. Tables are ranked high among some of the viewer's best information designs. They help structures the data so that viewers of the document can get in and out of the pages with ease.

Initially, support for tables was somewhat limited.

There are two ways to create home pages: do it yourself or hire a consultant who knows HTML and has the knowledge of HTML. As described, there are tools available to help the creation of home pages. These specialists can combine the text, image, audio, and video.

Some people argue that if an organization's pages are not attractive, well-designed, and user-Friendly, people will not be interested and the organization will not get much out of it.

Tool	Туре	platform	platform	
BBEdit HTML Extensions ^a	Template	BBEdit or for Mac		
Fm2html ^b	MIF converter	UNIX		
Hot Metal ^c	Editor	Mac, pc, UNIX		
HTML Assistant ^d	Editor	PC		
Web Maker ⁱ	MIF converter	UNIX		
Web Weaver ^j	Editor	Mac		

HTML tools (partial list)

Given a pointer to a piece of information on the Internet, the browser has to able to access that Information or operate in some way based on the contents of that pointer. For hypertext web Documents, this means that the browser must be able to interconnect with the server using the HTTP protocol. Since the web can also manage information contained on FTP and gopher servers, in new postings, in mail, and so on, the browser has to support these tools as well. Each page that is loaded from the web to the single document, written in HTML that includes the text of the document, its structure, and any links to the other documents, images, and other media. The browser interprets the HTML markup code contained in that documents and formats and displays the document.

Hypermedia:

For graphical output, there are two kinds of images that a web browser can handle: inline image and external images. An inline image is a graphic on web page that does not contain a link to another place; it is there for illustration only. External images are not directly loaded on a web page; they are only downloaded at the request of the reader.

Inline images are specified using the tag in the HTML file. An example of an image tag is which makes references to a GIF (Graphics Interchange Format) files containing scanned pictures of Dan's cat, Micio-Micio. However inline images and External images provide a design and presentation for a document that can be alluring to the Customer. Therefore, publishing on a web would be more attractive and beneficial if some image Formats were included to help capture the viewer's attention. Once an image is placed on the web, the publisher may then perform any text and image alignment, place any links on the image needed, place transparent backgrounds on the image, create borders, and so on. Some of the more newly introduced technologies include the interlacing effects. Interlacing a GIF image does not change the appearance of the image but rather changes the effect of how the image is saved and loaded. The image, as it is loading, has the appearance of fading-in line by line. To create these effects, publisher needs tools for creating interlaced GIF images.

Gif, is the most widely used graphics formats on the web today. GIF was developed by CompuServe to fill the need for a cross-platform image formats.GIF files are predominately used for logos, icons, line, art, and other simple images. The problems with GIF, at the moment, is that the form of compression it uses, LZW, is patented. Unisys, the owner of the patent, has requested that developers who use the GIF formats after 1994 pay a per-copy royalty for the use of LZW. Because of the problems with the patent on LZW and the possibility of it costing publishers to use the format, the GIF formats may fade away in the future to be replaced with some other, more freely available format.

The candidate likely to replace GIF is jpeg, named after the group that developed it the Progressive JPFG file format loads images up to three times more quickly than the previous GIF formats and provides faster intermediate image recognition, so users on slow connections can view color-rich images quickly. In facts, with Native Progressive JPEG decompression, less than 10 percent of the image needs to be loaded to be recognizable.

JPEG was designed for the storage of photographic images. Unlike GIF, JPEG images can have any number of colors, and the compression formula it uses works well for photographic patterns, so the files sizes it creates from photographs are considerably smaller than those that GIF produces. In addition, JPEG files have just begun to widely supported by browsers.

Including sound files on a web page can provide customers with important information. In addition, it may provide welcome messages for readers or sound clips of an organization. To include a link to a sound on a web page, the sound sample must first be in the correct format. Currently, the only cross-platform sound file format for the web is Sun Microsystems 'AU format. The following is an illustration of linking an AU format file: (AIFF format, 357K)

The current standard video format for the web is Motion Picture Expert Group (MPEG). Apple has also introduced the QuickTime format that has increasingly gained popularity. To include video files on the web pages, one must first have the correct file

extension.MPEG files have the extension .mpg or.mpeg while QuickTime has .move extensions. The following is an illustration of linking a video to an image:

<IMG SRC="film.gif"(12Meg)

- For graphical output, there are two kinds of images that a web browser can handle; inline images and external images.
- An inline image is a graphic on the Web page that does not contain a link to another place.
- External images are not directly loaded on a Web page; they are only downloaded at the request of the reader.
- Inline images are specified using the tag in the HTML file.
- Eg. which makes reference to a GIF(Graphics Interchange Format)
- Some of the more newly introduced technologies include the interlacing effect.
- Interlacing a GIF image does not change the appearance of the image but rather changes the effect of how the image is saved and loaded.
- GIF is the most widely used graphics format on the Web today.
- GIF files are predominately used for logos, icons, line art, and other simple images.
- The problem with GIF, at the moment, is that the form of compression it uses, LZW, is patented.
- The progressive JPEG file format loads images up to three times more quickly than the previous GIF format and provides faster intermediate image recognition, so users on slow connections can view color-rich images quickly.
- JPEG was designed for the storage of photographic images.
- Unlike GIF, JPEG images can have any number of color, and the compression formula it uses works well for photographic patterns.
- To include a link to a sound on a Web page, the sound sample must first be in the correct format.
- The only cross-platform sound file format for the Web is Sun Microsystem' AU format.
- The following is an illustration of linking an AU format file:

(AIFF format, 357K)

- The current standard video format for the Web is Motion Picture Expert Group (MPEG).
- MPEG files have the extension .mpg and .mpeg while QuickTime has .mov extensions.
- To link the video files to a web page is similar to the functionality of linking a sound.
- The following is an illustration of linking a video to an image:

<IMG SRC="film.gif"(12 Meg)

Data collection:

- As many sites owners know by now, most web servers automatically create an access log of usage data, recording the domain name or IP address of sites accessing the file, dates and times of access, which files were viewed, and the sizes of those files.
- A three-digit return code indicates whether file requests were fulfilled or rejected, helping the web masters check the validity of links.
- The agent file records the brand and version of browser being used; the referrer file identifies the page from which a visitor is linking and where the visitor travels within the site.
- These log files can be linked to one another, run separately, or may not be used at all.

- Information on the visitor's travels through a site can identify underused pages and possibly suggest better navigational patterns.
- And attention to the error log can prevent visitors from being stopped by a broken link.

Publishing systems:

- There are many browsers/publishing systems that allow users to create Web documents.
- Companies are now beginning to understand the benefits of publishing their information on the WWW.
- Information about a company has always been available for anyone with the time, resources, and energy to collect it manually.
- The Web is shrinking the time it takes to gather this information and publish it: once it is published, the web increases the speed at which the information can be collected.
- Web/intranet publishing enables internal corporate departments to provide timely information on new product announcements or marketing videos and to help increase employee productivity, improve competitiveness, enhance customer service, and reduce internal publishing costs.
- Publishing on the web goes a long way in helping to acquire customers and maintain their loyalty.
- Two of the most common publishing systems are Netscape Publisher and Adobe PageMaker.
 - The following subsections highlight the features of each (systems from other vendors are also entering the market).

Netscape:

- Netscape Publishing System has allowed enterprises to distribute their own publications and services on the Internet by offering content providers flexible and integrated software for organizing, customizing, and delivering text, graphics, audio, or video documents to users around the world.
- Netscape Publishing System is an integrated solution that is easily customized and can deploy quickly.
- It can be used as a standalone solution or be integrated with current software systems.
- It has the ability to enhance existing systems to become Internet-enabled and to provide increased functionality to customers.
- In addition to storing and providing access to information, the Netscape Publishing System allows publishers to add value to information by tailoring it to the requirements of individuals.
- Netscape Publishing system automatically searches new documents for predefined keywords and create automatic hyperlinks.
- Netscape Public system features an easy-to-use graphical interface and the ability to work with content derived from multiple sources.
- It exploits the security capabilities of Netscape Commerce Server and the ease of use of Netscape Navigator to provide a system that is industrial strength yet easy to manage and develop standards, Netscape in the past has worked with the Internet Engineering Task Force and the World Wide Web consortium. Table 8.2 describes key features.

The following is a listing of some additional standards that Netscape Publisher complies with .These standards help existing software packages such as plug-ins, Internet servers, and also many platforms be compatible with one another.

- *Native support.* Available for HTML, HTTP, FTP, NNTP, SMTP, MIME, S/MIME, S/MIME, and POP3 standards.
- *Cross –platform support.* Provide a common interface and common behavior among the different platforms.
- Open environments. Works in environments that support HTTP compliant network clients, including Winsock.

Minimum platform requirements are shown in Table 8.3.

The following is a listing of support platform for the Publisher.

 Apple Macintosh Macintosh system 7 or later Mac OS Power PC
Intel (x86)-based Window 3.1 and 3.11 Window for workgroups 3.11 Window 95 Windows NT (3.5 or higher)
UNIX Digital Equipment Corporation Alpha (OSF/12.0) Hewlett-Packard 700-series (HP-UK 9.03)
IBM RS/600 AIX 3.2 Silicon Graphics (IRIX 5.2)

Silicon Graphics (IRIX 5.2) Sun SPARC (Solaris 2.4, SunOS 4.1.3) 386/486/Pentium (BSDI)

Adobe PageMaker. AdobePageMaker is a Professional, cross-platform, desktop publishing program that allow the ability to design and pro-duce sophisticated publications. PageMaker combines text and graphics from a wide range of software application for printed or electronic delivery .PageMaker offers tools for producing professional-quality publications. PageMaker helps produce accurate color.

Registration	At time of registering, subscribers specify their interests. This information is used to filter through incoming documents for each subscriber. Subscribers scan their list and can then review the documents of interest.
Access Control	Access to documents can be controlled at the level of the entire publication or by the particular issue of a publication. Access controls can be applied to documents at various levels.
Document Managements	Electronic content such as text, graphics, audio and video can be created with a number of software tools or be imported to tag , convert the formats of, and load news and satellite feeds.
Simplified content Administration	A range of tools simplify the task of managing large document collections, including the ability to define searchable attributes, presentation formats, and profile.
Content Retrieval	A Natural Language Interface with support for dictionaries, synonyms and relevance ranking is providing to help users easily locate information.
An Open solution	The Publishing system is built on widely used Industry-standard technologies such as relational Database, text search engines , HTML,HTTP,SQL

Architecture	Modular system design to allow different modules to run on multiple systems, if necessary, to support high availability, scalability, security.		
Document handling	Multiple media types supported, including text, graphics, audio and video incoming feeds can include satellite and electronic news feeds and several other formats.		
Personalized Information access	Automatic filtering of information to generate personalized list of documents matching each subscriber's interests. Flexible number of Preferences can be stored per subscriber.		
Text search	Full Boolean search Capability including AND, where document OR and NOT. Relevance ranking, retrieved as a result of a query are ranked in order of their relevance.		
Elawihla Dilling	Symmet for both symptoms and real times and dit and myschages		

Flexible Billing Support for both subscriptions and real time credit card purchases. Option

Platform	Processor (minimum)	Disk space (MB)	Memory (minimum, MB)	Memory (recommended, MB)
Window	386SX	2	4	8
Macintosh	68020	2	4	8
UNIX	N/A	3	16	16

Table 8.3: Platform requirement

Adobe PageMaker includes a build in word with a spell Checker and search and replace features. In addition PageMaker words with text, graphics, image, spreadsheets and data from all software application.

PONDICHERRY UNIVERSITY QUESTIONS

2 <u>MARKS</u>

- 1. Define the term Internet. (Apr 2012) (Ref.Qn.No.23)
- 2. What is the usage of a browser? (Apr 2012) (Ref.Qn.No.22)
- 3. What are the two most common publishing systems? (Nov 2012) (Ref.Qn.No.24)
- 4. Give the primary uses of Spider program. (Nov 2012) (Ref.Qn.No.8)
- 5. Write about electronic payment system. (Apr 2013) (Ref.Qn.No.25)
- 6. What is SOA GOVERNANCE? (Apr 2013) (Ref.Qn.No.26)
- 7. What is meant by Nap access?(Apr-2014) (Ref.Qn.No.21)
- 8. Write a note on internet browser?(Apr-2014) (Ref.Qn.No.22)
- 9. Write short notes on Internet & Intranets.(Nov 2014) (Ref.Qn.No.27)
- 10. Mention Some Internet Tools.(Nov 2014(Apr 2015) (Ref.Qn.No.28)
- 11. Comparison between Website & Webpage.(Apr 2015) (Ref.Qn.No.29)

11 <u>MARKS</u>

- 1. What are the technologies used for web servers? Explain. (April 2012)(Nov 2014)(April 2015) (Ref.Qn.No7)
- 2. Discuss the tools of WWW. (April 2012) (Ref.Qn.No.5)
- 3. Describe the various internet tools relevant to commerce. (Nov 2012)(April 2014) (April 2015) (Ref.Qn.No.5)
- 4. Discuss the key technology features that should be included in search engines.(Nov 2012) (Ref.Qn.No.4).
- 5. Explain about various e-commerce considerations for internet applications. (April 2013) (April 2014) (Ref.Qn.No.6)
- 6. Explain about E- Commerce and internet access in detail. (April 2013) (Ref.Qn.No.2, Pg.no.7)
- 7. Explain the function of Search Engine.(Nov 2014) (Ref.Qn.No.4)