UNIT II
E- COMMERCE and WWW

The need for E-commerce stems from the demand within business and government must make better use of computing i.e. to better apply computer technology to improve business process and information exchange both within the an organization and across the organization. E-commerce is used to devote proper exchange of business information using EDI, E-mail, Electronic bulletin boards, EFT(electronic fund transfer) and other similar technologies.

E-Commerce is used to describe a new online approach to perform traditional function such as payment and fund transfer, order entry and processing inventory management involving cargo tracking, electronic catalogue etc.

Advertising, marketing and customer support functions are also a part of E-commerce application.

No single technology can provide the full potential of E-commerce. Therefore we require an integrated architecture which is revolving in the form of WWW as E-commerce is becoming more matured. Thus we need to develop sophisticated applications on WWW.

Architectural framework of E-commerce:

A Frame Work is intended to define and create tools that integrate the information found in today's closed system and allow the development of E-commerce applications.

Architectural framework should focus on synthesizing the diverse resources already in place incorporation to facilitate the integration of data and software for better use and application.

The E-commerce applications architecture consists of 6 layers of functionality or services. They are

1. Application Services
2. Brokerage Services
3. Interface support layer
4. secure messaging & EDI
5. Middleware, structured document interchange.
6. Network infrastructure and providing communication services.

1. Application services:

It will be composed of existing and future applications based on innate architecture. The three distinct classes of E-commerce applications can be distinguished as

(a) Consumer to Business
(b) Business to Business
(c) Intra organization.

(a) Consumer to Business:

We call this enterprise market place transaction. In market place transaction customer learn about product differently through Electronic publishing by them differently using Electronic cash and secure payment and have them developed differently.

(b) Business to Business:

This is called as market link transaction. Here business, govt and other organizations depend on computer to computer communication as a fast, economical dependable way to conduct business transactions. They include the use of EDI and E-mail for Purchasing goods and services, buying information and consulting services, submitting requests for proposals and receiving proposals.

(c) Intra Organizational transactions:

This is called as market driven transaction. A company becomes market driven by dispersing throughout the firm information about his customers and competitors by spreading strategic and tactical decision making so that all units can participate and by continuously monitoring their customer commitment. To maintain relationships that are critical, to deliver superior customer value management, most pay close attention to both before and after sales.

A market driven business develops a comprehensive understanding of its customer business and how customers in the immediate and downstream markets perceive value. Three major components of market driven transactions are

(i) Customer orientation through product and service customization
(ii) Cross functional coordination through enterprise integration, marketing and advertising.
(iii) Customer service.

2. Information Brokerage and management:

This layer provides service integration through the notion of information brokerages. Information brokerage is used to represent an intermediary which provides service integration between customer and information providers, given some constraints such as low price, fast service, profit maximization for a client.
Information brokerage addresses the issue of adding value to the information that is retrieved.
Brokerage function can support data management and traditional transaction services. Brokerage may provide tools to accomplish more sophisticated tasks such as time delay updates or feature comparative transaction.
At the heart of this layer lies the work flow scripting environment that built on software agent model that coordinate work and data flow among support services. Software agents are mobile programmers that have been called as “healthy viruses” , “digital butlers” , and “intelligent agents”. Agents are encapsulations of users instructions that perform all kinds of tasks in electronic market places spread across the network.

3. Interface support service:

The third layer interface and support services will provide interface for e-commerce applications such as interactive catalogues and will support directory services etc., functions necessary for information search and access. Interactive catalogues are customized interface to consumer applications such as home shopping. An interactive catalogue is an extension of paper based catalogues and incorporates additional features such as sophisticated graphics and video to make advertising more attractive.

Directories on the other hand operate behind the scenes and attempt to organize the huge amounts of information and transactions generated to facilitate electronic commerce. Directory services databases make data from any server appear as a local file. Thus directories play an important role in information management functions.

4. **Secure messaging and structure document interchange service:**
   The importance of fourth layer is secured messaging. Messaging is a software that sits between the network infrastructure and the clients or e-commerce applications.

   Messaging services offer solutions for communicating non formatted data such as letters, memo, reports etc as well as formatted data such as purchase order, shipping notices and invoice etc. messaging support both for synchronous (immediate) and asynchronous (delay) messaging. When a message is sent work continuous (software does not wait for response). This allows the transfer of messages through store and forward methods.
   With messaging tools people can communicate and work together more effectively, no matter where they are located.
   The main disadvantages of messaging are the new types of applications it enables , which appear to be more complex especially to traditional programmers.

5. **Middleware services:**
   Middleware is a relatively new concept that emerged only recently. Middleware is a mediator between diverse software programs that enable them to talk with one another. It solves all the interface, translation, transformation and interpretation problems that were driving application programmers crazy.

   Another reason for Middleware is the computing shift from application centric to data centric. i.e., remote data controls all of the applications in the network instead of applications controlling data. To achieve data centric computing middleware services focus on three elements.

   (1) Transparency
   (2) Translation security management
   (3) Distributed object management and services

(1) Transparency:
Transparency implies that users should be unaware that they are accessing multiple systems. Transparency is essential for dealing with higher level issues than physical media interconnections that the underlying network infrastructure is in charge of. Transparency is accomplished using middleware that facilitates a distributed computing environment. This gives users and applications transparent access to data, computation and other resources across collection of multi vendor heterogeneous systems.

(2) Transaction security management:
The two broad categories of security (management) services for transaction processing are

(a) Authentication
(b) Authorization.

Transaction integrity must be given for business that cannot afford any loss or inconsistency in data. For E-commerce , middleware provides qualities expected in a standard transaction processing (T.P) system i,e. the so called ACID (Atomocity, consistency, isolation, Durability).

(3) Distributed Object Management:

Object orientation is proving fundamental to the proliferation of network based application for the following reasons.

It is hard to write a network based application without either extensive developer retaining or technology that adopts the difficulties of the network. objects are defined as combination of data and instructions acting on the data. objects are an evolution of more traditional programming concept of functions and procedures.

A natural instance of an object in E-commerce is a document. A document carries data and often carries instructions about the action to be performed on the data.

Middleware acts as an integrator for various standard protocols such as TCP(transmission control protocol) IP (Internet protocol), OLL

Hypertext Publishing

Web provides a functionality necessary for e-commerce. The web has become an umbrella for wide range of concepts and technology that differ markedly in purpose and scope which include hypertext publishing concept, the universal reader concept and the client server concept.

Hypertext publishing promotes the idea of seamless information world in which all online information can be accessed and retrieved. In a constant and simple way hypertext publishing is a primary application of web interest in hypermedia. On the internet (called distributed or global hypermedia). As accelerated shortly following the success of web media and browser. This success has been aided by more powerful work station high resolution graphic display faster network communication and decreased cost for large online service.

Hypertext Vs hypermedia:

Hypertext

Hypertext is an approach information management in which data are shared in the network of document connect by links (this link represents relationship between nodes.

Hypermedia

A hypermedia system is made up of nodes (documents) and links (pointers). A node generally represents a simple concept and idea. Nodes can contain texts, graphics, audio, video images etc. nodes are connected to other nodes by links. The movement between nodes is made by activating links which connect related concept or nodes links can be bidirectional.

Hypertext is a simple context based on the association of nodes through links. A node from which a link is originated is called the reference or the anchor link and a node at which a link ends is called referent. The movement between the links is made possible by activating links. The promise of hypertext lies in the ability to produce large complex richly connected and crossed reference bodies of information.

Benefits of Hypermedia:

1. hypermedia documents are much more flexible than conventional documents.
2. hypermedia documents offer video sequences animation and even compute programs.
3. its power and appeal increases when it is implemented in computing environments that include network , micro computers , work stations, high resolution displays and large online storage.
4. it provides dynamic organization.
5. hypermedia systems provides non-linear innovative way of accessing and restricting network documents.

Technology behind the web:

Information providers (publishers) run programs called servers from which the browsers can obtain information. These programs can either be web servers that understand the hypertext transfer protocol (HTTP) , "gateway" programs that convert an existing information format to hypertext, or a non-HTTP server that web browsers can access i.e FTP or Gopher servers.

Web servers are composed of two major parts.

1. the hypertext transfer protocol (HTTP) for transmitting documents between servers and clients .
2. HTML format for documents.

The link between HTML files & HTTP server is provided by Uniform Resource Locator (URL).

Uniform Resource Locator:
The documents that the browsers display are hypertext that contains pointers to other documents. The browser allows us to deal with the pointer in a transparent way that is select the pointer we are presented with a text to which it points. This pointer is implemented by using a concept which is central to web browser known as URL.

URL's are streams used as address of objects (documents, images etc) on the web. URL marks the unique location on the internet so that a file or a service can be found.

URL's follow a consistent pattern that the first part describes the type of the resources, second part gives the name of the server posting the resources and the third part gives the full name of resources.
e.g : FTP://server.address / complete file.name

URL are central to web architecture. That fact is that it is easy to address an object anywhere on the internet is essential for the system to scale & for the information space to be independent os network and server topology.

Hypertext Transfer Protocol (HTTP):

It is the simple request response protocol that is currently run over TCP and is the basis of WWW. HTTP is a protocol for transferring information efficiently between the requesting client and server. The data transferred may be plain text , hypertext images or anything else. When a user browses the web objects are retrieved in rapid succession from often widely dispersed servers.

HTTP is used for retrieving documents in an unbounded & extensible set of formats. It is an internet protocol. It is similar in its readable, text based style to the file transfer (FTP) & the network news (NNTP) protocols that have been used to transfer files and news on the internet for many years.

When objects are transferred over network, information about them is transferred in HTTP Header. The set of headers is an extension of the multi purpose internet mail extension (MIME) set. This design decision was taken to open the door to integration of hypermedia mail , news and information access.

HTTPD Servers (Hypertext transfer protocol domain)

The server that are used to publish information via WWW servers are called HTTPD servers. While choosing a web server flexibility, ease of administrator, security features, familiarity and performance are considered.

It is important to evaluate the tasks for which the web server is used. A server used for internet based marketing & technical support task will need more powerful server than the web server used internally within a firewall for distributing memos and bulletins. HTTPD servers are ideal for companies that want tp provide multitude of services ranging from product information to technical support.

HTML (Hypertext markup language)

At the heart of the web is a simple page description language called HTMl. It is a common basic language of interchange for hypertext that forms the fabric of the web. It is based on an international electronic document standard called Standard generalized markup language (SGML)

HTML enables document orientation for the web by embedding control codes in ASCII (American standard code for information interchange) text to designate titles, headings, graphics and the hypertext links, making links of SGML's powerful linking capabilities. HTML was meant to be a language of communication which actually flows over the network HTML was designed to be sufficiently simply as to be produced easily by the people and automatically generated by the programs.

HTML Forms

Forms support is an important element for doing online business. Forms are necessary for gathering user information conducting surveys and also providing interactive services.

Forms make web browsing an interactive process for the user and the provider. They provide the means to collect and act upon the data entered by end users. Forms also open up a number of possibilities for online transactions such as restricting specific news articles, specifying such as request , soliciting customer feedback or ordering products. The number of features are available for building forms including text boxes, radio buttons, check boxes.

Common Gateway Interface Services (CGI)

An important aspect of web server development is application gateways. More specifically it is CGI. CGI is a specification for communicating data between web server and other application server. CGI is used whenever web server needs to send or receive data from another application.

A CGI script is a program that negotiate the movement between web server and an outside application. CGI scripts may be written virtually any high level language such as C, Perl (Practical extraction and reporting language), Java scripts etc.

Security on the web:

Security and confidentiality are essential before business conduct, financial transactions over the internet has become a big problem due to the increasing number of application oriented towards commerce. Therefore commercial application requires that the client and server be able to authenticate each other and exchange data confidentiality. This exchange has three basic properties.

1. clients are confident about servers they are communicating with server authentication.
2. client conversation with server is private using encryption.
3. client conversation cannot be tampered or inter separated with data integrity.

Categories of Internet data & Transaction:

Several categories of data must be encrypted making internet data security an interesting challenge.

Public Data :

Public data have no security distinctions and can be read by anyone. Such data should be protected from unauthorized tampering or modification because a reader may perform damaging actions on its contents.

Copyright Data:

Copyright data have content that is copyrighted but not secret. The Owner of the data is willing to provide it but wishes to ensure that the user has paid for it. The objective is to maximize the revenue and security.

Confidential Data:

Confidential data contains material that is secret but whose existence is not secret such data include bank account systems, personal files etc. such material may be referenced by public or copyright data.

Secret Data:

Secret data existence is a secret such data might include algorithms which is necessary to monitor, log all access to secret data.

Despite the variety of data, security and verification are necessary for all
Types because of the sensitivity of information being transferred and to protect the consumer form various forms of fraud and misconduct.

WWW based security schemes:
Several methods can provide security in the web framework. This includes the following.

1. **SHTTP:- (Secured Hypertext Transfer Protocol)**

   SHTTP will enable the incorporation of various cryptographic messages, formats such as digital signature Algrithms (DSA) & RSA standards into the both their client & servers.

2. **SSl :- (Security Socket Layer)**

   SSL uses RSA security to wrap security information around TCP/IP based protocols. The benefits of security socket layer over secured HTTP is that SSl is not restricted to HTTP. But can also be used for security for FTP & TELNET.

3. **SHEN :**

   It is the security scheme for the web sponsored by www. It is not non-commercial or more research oriented security & is similar to SHTTP.

   SHEN Security scheme for the web:
   SHEN provides for three separate security – related mechanisms.

   1. *Weak authentication with low maintenance overhead and without patent or export restrictions.* A user identity must be established as genuine. Unauthorized access must be improbable but need not be secure from all possible forms of attack.
   2. *Strong authentication via public key exchange.* A user identity must be established as genuine. Unauthorized access must be impossible except by random chance or by access to unknown technology.
   3. Strong encryption of message content. The data must not be transmitted in a form comprehensible to a third party, an identified party acts as guarantor in this respect.

Messaging Security Issues:

In order to conduct electronic commerce on the internet, including the WWW, messages must be electronically transmitted in some manner. In addition to the general concern of data security, a primary concern is the non-refutable linking of message contents to individuals and businesses.

Several important security services are required to ensure reliable, trustworthy electronic transmission of business messages. The primary security services are interrelated. The five security services are:

SECURITY ISSUE	SECURITY OBJECTIVE	SECURITY TECHNIQUES
Confidentiality	Privacy of messages	Encryption
Message Interit	Detecting message tamerin	Hashing
Authentication	Origin verification	Digital signatures challenge- Response passwords Biometric devices
Non-Repudiation	Proof of Origin, receipt, and contents(sender cannot falsely deny sending or receiving the message)	Bi-directional hashing Digital signatures Transaction certificates Time Stamps Confirmation services.
Access controls	Limiting entry to authorized users	Firewalls Passwords Biometric devices.

1. Confidentiality: when a message is sent electronically, the sender and receiver may desire that the message remain confidential, and thus not read by any other parties. Analogies can be drawn to traditional mail and phone systems. In regular mail systems, the sender uses an envelope to conceal the inside contents rather than writing the information on a post card.

For E-commerce, keeping order details and credit information confidential during the transmission is a major security concern. Further, trading partners sharing design specifications also want to ensure the confidentiality of their messages so that proprietary design specifications can be viewed only by the sender and the intended receiver of the information. The most effective technique for masking a message is encryption.

2. Integrity : when a message is sent electronically, both the sender and receiver want to ensure that the message received is exactly the same as the message transmitted by the sender. A message that has not been altered in any way, either intentionally or unintentionally, is said to have maintained its **integrity**. For electronic commerce verifying that the order details sent by purchaser have not been altered is one major security concern. An effective cryptographic means of ensuring message integrity is through the use of **hashing** , where a "hash" of the message is computed using an algorithm and the message contents. The hash value is sent along with the message; then, upon receipt, a hash is calculated by the recipient using the same hashing algorithm. The two hash values (received and calculated) are compared, and a match can indicate that the message is the same as that sent.

3.Authentication: when an electronic message is received by a user or a system, the identity of the sender needs to be verified(i.e.authenticated) in order to determine if the sender is who he claims to be. To identify a user at least one of the following types of information is generally required

- Something you have(e.g., a token)
- Something you know(e.g., a PIN) or
- Something you are (e.g., fingerprints or signatures)

4.Non-Repudiation: the term **repudiate** means to accept as having rightful authority or obligation as in refusing to pay a debt because one refuses to acknowledge that the debt exists. For business transactions, unilateral repudiation of a transaction by either party un acceptable and can result in legal action. Well designed electronic commerce system provide for **non-repudiation**, which is the provision for irrefutable proof of the origin receipt, and contents of an electronic message.

5.Access Controls: Electronic commerce systems, particularly those using the internet and the WWW, require a certain amount of data sharing. Limiting access to data and systems only to authorized users is the objectives of **access controls**. Some form of authentication procedure is typically employed in access controls in order to gain entry into the desired part of the system. The emerging attribute certificate or “privilege management” technology promises to be a highly effective form of access control provided it is implemented correctly. Firewalls can also be used to implement additional screening mechanisms.

Encryption Techniques:

Confidentiality of electronic messages is a necessity of electronic commerce application. The primary method of achieving confidentiality is **encryption .** messages are initially created in a form that is readable and understandable by the sender, and by any other individuals as well if they have access to the message. The message, when it is in this form is commonly reffered to as **clear text** or **plaintext** .

Encryption is defined as the transformation of data, via a cryptographic mathematical process into a form that is unreadable by anyone who does not posses the appropriate secret key. That data in this unreadable form is commonly referred to as **cipher text.** If a message is intercepted and read, it will be useless since the cipher text message is unintelligible to any party not possessing the secret key. In order to be able to read and understand the message, the encrypted message must be transformed back to its original state- the clear text. The process os restoring cipher text to clear text is called **decryption.**

The key contains the binary code used to mathematically transform a message, two types of cryptographic mechanisms can be used to provide an encryption capability: **Symmetric** cryptography where entities share a common secret key; and a public key cryptography (also known as **Asymmetric** cryptography) where each communicating entity has a unique pair (a public key and a private key).

For symmetric and asymmetric encryption, the relative strength of the cryptography is most commonly measured by length of the key, in bits. However it should be noted that the true strength of the confidentiality service may depend on a number of variables associated with the encryption function :

- The security protocol used to invoke the encryption function.
- The trust in the platform executing the protocol or application.
- The cryptographic algorithm.
- The length of the key(s) used for encryption/decryption.
- The protocol used to manage/generate those keys.
- The storage of secret keys(key management keys and encryption keys).

The strength of a system usually increases as the key length increases. This is because a longer key length implies a larger number of possible keys, which makes searching for the correct key a more time consuming process. Any key length less than 64-bits is no longer considered to be secure.

Symmetric Encryption Keys:
In symmetric key systems, both the sender and the receiver of the message must have access to the same key. This shared secret key is used to both encrypt and decrypt the message.

Asymmetric Cryptography:
In 1976, a concept referred to as public key cryptography was introduced by Whitefield Diffie and martin Hellman, called the **Diffie-hellman** technique. The public-key method allows a sender and a receiver to generate a shared, secret key over an insecure telecommunications line. This process uses an algorithm based on the sender's and receiver's public and private information. The following steps are used

1. The sender determines a secret value a.
2. A related value , A, is derived from a. A is made public.
3. The receiver determines a secret value b.
4. A related value, B is derived from b. B is made public.
5. the Diffie-Hellman algorithm is used to calculate a secret key corresponding the key pairs (a, B) and (b, A).

the sender knows his private value, a and the receiver's public value, B. the receiver knows her private value, b , and the sender's public value, A. the secret key is generated from (a, B) and (b, A) by an algorithm that makes it computationally infeasible to calculate the secret key from solely knowing the two public values, A and B. In order to generate the secret key, one of the secret values must be known. The secret key is shared avoiding the problem of transmitting it over a insecure telecommunications line.

Good encryption practices:

The following are the few good encryption practices that foster stronger security.

1. **Password maintenance:** never share your secret password. A password can be used to protect your private key, and therefore your digital signature.
2. **keylength:** use an appropriate key length whenever possible. The longer the key length, the greater the security. For domestic use a key length of at least 64-bits should be used .
3. **compressed files:** in order to reduce transmission time, data compression is frequently used to reduce the size of a file. Most loss less data compression techniques are based on removing redundancy from the file.

Questions:

1. Briefly explain about Architectural frame work of E-Commerce.
2. Explain different types of E-commerce Applications.
3. Explain about Hypertext publishing and advantages of Hyper media documents.
4. Briefly discuss the technology behind the Web.
5. Write short notes on HTML forms.
6. Write short notes on
    a) HTTP
    b) HTTPD servers
    c) URL
    d) CGI
    e) HTML
7. Briefly Discuss the security on the web
8. List and explain various WWW based security schemes.
9. Write about various Encryption techniques.