

Electronic Commerce

2. E-Commerce System Architecture

V Rajaraman



V Rajaraman is with the Jawaharlal Nehru Centre for Advanced Scientific Research and the Indian Institute of Science, Bangalore. Several generations of scientists and engineers in India have learnt computer science using his lucidly written textbooks on programming and computer fundamentals.

Part 1. What is E-Commerce?, *Resonance*, Vol.5, No.10, 13-23, 2000.

In this part we will depict the infrastructure required for e-commerce as a six layered architecture and describe the three bottom layers.

The system architecture for e-commerce consists of several layers, each layer providing some services. At the bottom is the physical network which connects the computers participating in e-commerce and at the top level are the e-commerce applications. We will now examine each layer and see what services they provide. The layered architecture is shown in *Table 1*.

Physical Network

If computers are to communicate with one another they should be physically connected. Most businesses have a Local Area Network (LAN) connecting all their computers. The LAN usually connects machines with an unshielded twisted pair (UTP) of copper wires. Computers connected to a LAN using UTP can communicate at the rate of 1 Gigabits/sec even though 100 Mega bits/sec is more common. The number of computers that can be connected to a segment of a LAN is limited to around 16. Larger LANs are made by connecting smaller LAN segments by what are known as *bridges*. Besides UTP, one may use fibre optic cables to interconnect computers if higher speed is needed. Mobile computers may also be connected to a LAN using wireless communication.. Those interested in detailed information on physical networks may refer to the book by Stallings [1].

When two businesses want to communicate with one another their LANs are connected using what is known as a *router* and a *modem* to the telephone network provided by the Department of



Table 1. Layered architecture for e-commerce.

Logical Layers	Services in each layer								
Application layer	Customer to Business Business to Business Customer to Customer								
Middleman Services	Mail order houses, Hosting services Payment schemes, Electronic cash Mailing services Value Added Networks Digital Signature Certification Authority								
Secure Messaging	Encryption EDI Firewalls Digital Signature								
World Wide Web Services	<table border="0"> <tr> <td data-bbox="490 835 597 895">Structured documents</td> <td data-bbox="681 800 1054 895"> HTTP Hypertext transfer protocol HTML Hypertext Markup Language XML Extensible Markup Language </td> </tr> <tr> <td data-bbox="490 930 597 991">Compound documents</td> <td data-bbox="681 930 836 956">OLE, opendoc</td> </tr> <tr> <td data-bbox="490 999 642 1025">Search Engines</td> <td></td> </tr> <tr> <td data-bbox="490 1034 655 1060">Software Agents</td> <td></td> </tr> </table>	Structured documents	HTTP Hypertext transfer protocol HTML Hypertext Markup Language XML Extensible Markup Language	Compound documents	OLE, opendoc	Search Engines		Software Agents	
Structured documents	HTTP Hypertext transfer protocol HTML Hypertext Markup Language XML Extensible Markup Language								
Compound documents	OLE, opendoc								
Search Engines									
Software Agents									
Logical Network	<table border="0"> <tr> <td data-bbox="584 1107 742 1203"> INTERNET INTRANET EXTRANET </td> <td data-bbox="715 1116 742 1194">} </td> <td data-bbox="790 1142 952 1168">TCP/IP Protocol</td> </tr> </table>	INTERNET INTRANET EXTRANET	}	TCP/IP Protocol					
INTERNET INTRANET EXTRANET	}	TCP/IP Protocol							
Physical Network	PSTN (Public Switched Telephone Network) Wireless Optical fibre UTP copper wires LAN Bridges Routers								



Box 1. Modems

Modem is a contraction for *Modulator Demodulator*. Telephone lines were designed to carry human conversations which lie in the frequency range 30 Hz to 3000 Hz and were not meant to transmit digital data which is a train of pulses. Thus modems were designed to convert 1 and 0 into signals at two different frequencies in the range of 1 to 2 KHz. Early modems could carry digital data at a rate of only 1200 bits/sec. Improvement in digital signal processing, error correcting codes and data compression has now resulted in modems which can work at very high speeds of 56 Kbps on the same old telephone networks. The copper wires by themselves have a bandwidth of around 2 Mhz (for 2 Km length) but the speed is limited by the switching equipment at the telephone exchanges. New technologies which bypass telephone exchange and provide higher bandwidth are on the horizon. Besides analog telephone lines, lately DOT is providing Integrated Services Digital Network (ISDN) which provide speed in excess of 64 Kbps for high bandwidth digital traffic.

Telecommunications (DOT) (see Box 1). This network is known as Public Switched Telephone Network (PSTN) or Plain Old Telephone System (POTS). PSTNs are not very secure. Thus if two businesses want to closely collaborate they have their own private leased communication lines.

Logical Network

The single most important technology which has enabled the growth of e-commerce is the *internet*. The internet connects tens of millions of computers spread all over the world enabling them to exchange information and share resources [2]. The major applications of internet are exchange of electronic mail, exchange of files (text as well as multimedia), storing information in a form which allows other computers connected to the internet to access it and remote logging into a computer and using it to run programs. For two computers in different locations to communicate it is necessary that:

The single most important technology which has enabled the growth of e-commerce is the *internet*.

- i) Each one must have a unique address
- ii) Messages originating from one computer must be routed to the destination via a Public Switched Telephone Network. It may be local, national or international PSTNs.
- iii) There must be software in each computer to format the

messages using commonly agreed rules so that the messages are properly routed and interpreted. This commonly agreed set of rules is called the *Internet Protocol* (IP).

The unique address required by a computer in order to access the internet is called its IP address. The IP address is a 4 byte address and is expressed in what is known as the dotted decimal format, eg. 202.42.128.3. The IP address for a business or an individual is provided by the Internet Service Provider (ISP). IP addresses are an important and scarce resource particularly because the number of computers connected to the internet is rapidly growing. The IP address is converted into a string of characters for ease of remembering and grouped into *domains*. For example in the address:iisc.ernet.in, the top domain is the country name abbreviated *in*, the ISP is *ernet* who is the host of *iisc*. IP addresses are controlled by an international authority known as Internet Corporation for Assigned Names and Numbers (ICANN) and a hierarchical organization of addresses allows this authority to decentralise assigning of addresses. For example, ernet is given a range of IP addresses by ICANN and it allocates a subset of addresses to iisc which in turn allocates addresses to various departmental servers. The clients connected to the departmental servers are then given their unique address by the department. Domain names are very important in e-commerce as they provide immediate brand recognition. Thus there has been a problem known as cyber squatting which has led to legal wrangles (see *Box 2*).

Domain names are very important in e-commerce as they provide immediate brand recognition.

The internet protocol breaks up a message sent from a source to a destination into a number of *packets*. A packet consists of two parts, the part containing the information which is called the *payload* and a part called the *header* (see *Figure 1*). The header consists of the source and destination addresses, the serial number of the packet, error detection bits and other control bits and is used to route the packet to destination address. Messages are broken into packets as it reduces the cost of transmission and improves the fault tolerance of transmission. The cost is reduced as a number of packets (may be belonging to different

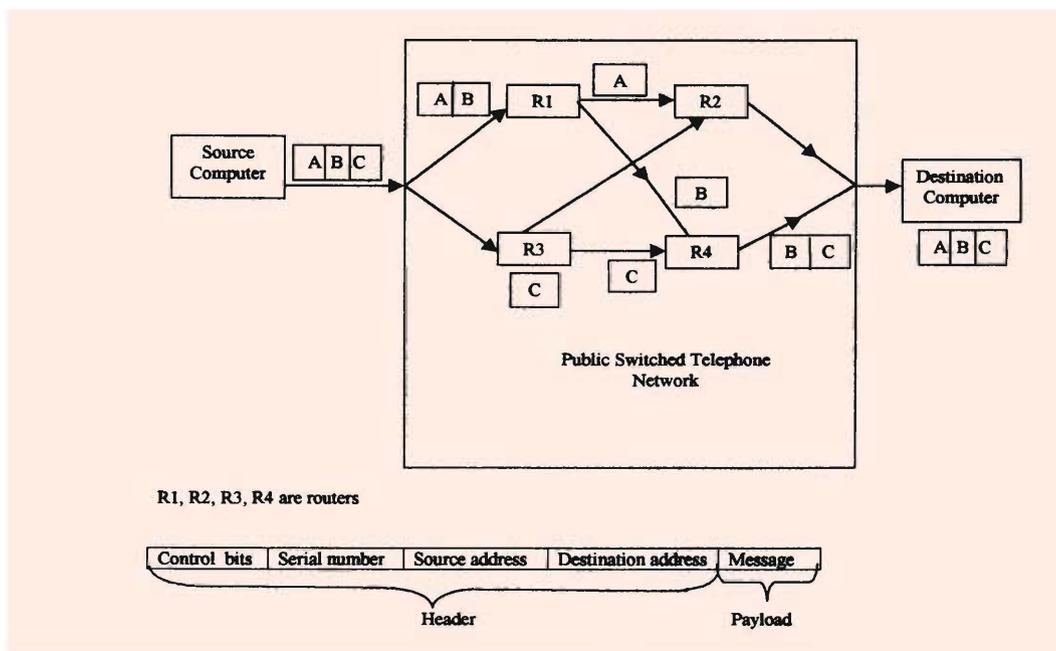


Box 2. Cyber Squatting.

With Net based shopping becoming popular the value of internet addresses has been realised by both old companies and new entrants. There have been disputes between organizations with similar addresses. An interesting dispute has been between etoys.com a toy store in U.S.A and etoy.com an artist's group in Europe on their rights for their respective addresses. The Internet Corporation for Assigned Names and Numbers (ICANN) has adopted a uniform domain name dispute resolution policy in October 1999. This policy states that an owner of a domain name being disputed by a complainant must obey the decision of an international cyberspace tribunal if accused of 'bad faith' registration of domain name also known as 'cybersquatting'. If the bad faith registration is proven ICANN will remove the domain name of the squatter and assign it to the legitimate party. In spite of this international policy, some countries (notably U.S.A), are passing their own legislation which may lead to lengthy litigation and in some cases severe penalties. There is thus a move now to refer international disputes to World Intellectual Property Organization which resolves disputes on copyrights, trade marks, etc.

messages) can be assembled and sent along any free communication channel. The packets are stored in routers along the path and forwarded to another router when the communication link is free. This is called *packet switching*. It is also fault tolerant because if a line is not working the stored packet can be sent along another line which may be working. Observe that differ-

Figure 1. Packet communication used in internet.



ent packets belonging to a message may travel along different paths. They are finally assembled at the destination using the serial number of each packet. The major disadvantage of packet switching is that the time taken for a message to reach a destination cannot be predicted. This is not a disadvantage for applications such as email, file transfer etc., but is a disadvantage for real-time messages such as voice conversations and video transmissions. Currently work is going on to improve the internet protocol to allow real-time data transmission also. The routing of data as packets is illustrated in *Figure 1*. One of the major advantages of internet is its ability to connect computers from any manufacturer and LANs using different technologies together into a uniform access system by enforcing that the computers use a software layer conforming to the internet protocol known as TCP/IP (Transmission Control Protocol/Internet Protocol). It was realised by many business/organizations with a large number of computers and departmental LANs that if they also use the internet protocol the entire network can easily share resources very efficiently. Such a network used by an organization to interconnect their computers is called an *intranet*. Thus a corporate intranet will enable internal email, web pages, information transfer, remote login etc., provided by an internet with high security as no PSTN is used. When two or more organizations closely collaborate and trust one another they can connect their corporate intranets using a leased line or leased circuits in a PSTN. Such a network will use internet protocols but is a 'private internet' among the collaborating organizations. Such a network is called an *extranet*. Extranet is normally more secure compared to internet as it is a closed group of users.

Currently work is going on to improve the internet protocol to allow real-time data transmission.

World Wide Web Services

The world wide web is a global multimedia information service available on the internet. It consists of linked web pages (or documents). Each web page is prepared using a language known as HTML (Hyper Text Markup Language). HTML has features to embed links within web pages pointing to other web pages, multimedia files and data bases. Web pages are stored on what

The world wide web is a global multimedia information service available on the internet.



The world wide web is not internet.

Internet provides the infrastructure on which the world wide web is built.

are known as web servers. A web server can host one or more web pages. Observe that the world wide web is not internet. Internet provides the infrastructure on which the world wide web is built.

To locate a web page stored in the world wide web a scheme known as Uniform Resource Locator (URL) is used. An example of a URL is given below:

<http://www.freesoft.org/connected/index.html>

In this example http specifies the protocol to be used. In this case it is hypertext transfer protocol. This is the protocol used for web search. www.freesoft.org preceded by:// is the address (called domain name) of a computer (called a server) which is permanently connected to the internet. The computer may be located anywhere in the world. The part of the URL, namely, /connected/index.html is a path to the required file which stores the information. In this case the document index.html is stored in a folder named 'connected'.

There are other protocols used in the internet for other services. For example ftp:// is used for transferring files from one computer to another connected to the internet; ftp stands for file transfer protocol. For example

<ftp://freesoft.org/<filename>>

will transfer the contents of the specified file to your computer if you have access permission from the server.

The information on a web page can be

retrieved by a customer (or user)

using a web browser program

which runs on his/her desktop

computer

connected to the internet.

The information on a web page can be retrieved by a customer (or user) using a web browser program which runs on his/her desktop computer connected to the internet. There are many web browsers, the most popular of which are Netscape and Internet Explorer. The URL is entered in the location field of the browser screen. The browser program connects to the specified web server, interprets the HTML markups and displays the document on the browser screen. Web browsers have excellent Graphical User Interface (GUI) which simplifies ac-



cess to web pages to anyone. Most organizations now maintain a web page on a server in their organization or on a server which is rented by a service agency. Hosting of organization's web pages has now become an important business. These businesses keep a large number of powerful servers on their network with reliable connection to the internet. They create the web pages of the organization based on specifications given by them and continuously update them on request from the contracting organizations. Web presence is now essential for any business as it publicises their activity. Besides organizations individuals also create and maintain web page to 'sell themselves'.

Hosting of organization's web pages has now become an important business.

In order to create a webpage a language is needed which formats the page with nice background colours, graphics, links to other parts of the same document and links to other web pages either in the same server or other servers. This language is called Hypertext Markup Language (HTML). Hypertext markup language adds tags to text which can be interpreted by any program. A simple example is given below:

```
<HTML>
<HEAD>
<TITLE>    </TITLE>
</HEAD>
<BODY>
<H1> Analysis and Design of Information Systems. </H1>
<P> This is <B> second edition </B> of <I> Rajaraman's
</I> book. </P>
</BODY>
</HTML>
```

will display the following:

Analysis and Design of Information Systems.
This is **second edition** of *Rajaraman's* book.

Observe the various commands introduced with the text. Some of them are:

- Document delimiters such as <HTML>, <HEAD>, <TITLE> and <BODY>



As the use of inter and intranets is increasing most documents are now created using HTML format.

- Section heading <H1>. More levels are available
- Paragraph and other spacing commands such as <P> above
- Character attributes such as for bold face, <I> for italics
- Graphic images to be displayed with the document.
- Listing using bullets or sequence numbers
- ‘Anchor’ commands which specify text or images that can be clicked on to another HTML document either in the same server or another server.

Observe that when you design a web page you can pick words and tag them with anchor commands. When these words are clicked the tag activates a link to the specified page, graphics file, audio or video file. As the use of inter and intranets is increasing most documents are now created using HTML format. Standard word processor outputs can be converted to HTML format using tools. There are also specialised tools available to create web pages.

HTML is based on a much larger standard language known as Standard Generalized Markup Language (SGML). A dialect of SGML called XML (Extended Markup Language) is now becoming more popular as it allows designing documents tailored to a select audience [4].

We summarise below the desirable properties of a website. A website is a computer which has a dedicated internet/intranet connectivity and provides access to information stored in it through web browsers. The information is stored using HTML. The following are the desirable properties of a website.

A website is a computer which has a dedicated internet/intranet connectivity and provides access to information stored in it through web browsers.

- The address should be meaningful.
- It should be aesthetically pleasing.
- It should contain meaningful, well structured and current information.
- The scope and coverage of information in the site should be clearly specified.



- It must have good help and search features.
- It must have links to related sites.
- It must be easy to navigate using the links.
- It should provide facilities for users to give their comments.
- It should have capability to count and track the number of references (called hits) made to it.

The following web sites give good guidelines for authoring web sites.

- www.w3.org/provider/style
- www.ed.gov/internal/wwwstds.html
- www.sun.com/styleguide

The number of web pages in the world wide web run into tens of millions and is continuously growing. Information in the web is poorly structured and contains very useful information along with poor quality unauthenticated information. Finding relevant documents is not easy. There are many tools known as search engines [3] which aid users in their search. These engines (which are actually search programs) receive a user query, systematically explore the web to locate documents, evaluate their relevance and return a rank ordered list of documents to the user. Among well known search engines are Altavista (www.altavista.digital.com), Excite (www.excite.com), Hotbot (www.hotbot.com), Lycos (www.lycos.com), webcrawler (www.webcrawler.com) and Google (www.google.com). There are so many search engines being developed and used that there is a website which gives the latest news on search engines (www.searchenginewatch.com).

One more important technology in the third layer is software agents. Software agents are programs which are personalized to a user's special requirements or profile, continuously run and are semi-autonomous. For example, a simple software agent would monitor all incoming email messages, classify them, give routine replies to certain messages on behalf of the recipient, discard unimportant mail and bring to the attention of its client

Information in the web is poorly structured and contains very useful information along with poor quality unauthenticated information.

Software agents are programs which are personalized to a user's special requirements or profile, continuously run and are semi-autonomous.



messages which need personal attention. More complex software agents would be mobile. For example a software agent of a book publisher may be programmed to continuously monitor stock of paper, determine usage pattern of paper, search the web for suppliers when the stock is low, find suppliers whose cost is low and who deliver in time and even place an order on behalf of the publisher. The development of innovative software agents is a challenging current area of work which will make e-commerce much more powerful than what it is today.

Suggested Reading

Address for Correspondence

V Rajaraman
IBM Professor of Information
Technology
JNCASR
Bangalore 560 064, India.
Email:rajaram@serc.iisc.ernet.in

- [1] W Stallings, *Data and Computer Communications*, 5th Edition, Prentice-Hall of India, New Delhi 1998.
- [2] N Shrikhande, *Internet and the World Wide Web*, *Resonance*, Vol.2, No.2, pp.64-74, 1997.
- [3] T B Rajasekhar, *Web Search Engines*, *Resonance*, Vol.3, No.11, pp.40-53, 1998.
- [4] W J Pardi, *XML in action*, Microsoft Press, 1999.

DST Winter School on Organometallic Chemistry Chemistry Department, IIT, Kharagpur January 8-13, 2001

An intensive course for college and University teachers, aiming to up-grade teaching and research in Organometallic Chemistry in the country, will be held. Major topics to be covered are: *Structure and bonding in organometallic chemistry including PC-based demonstration of molecular modeling applications, Synthesis and reactivity of mono- and polynuclear organometallic compounds, Application of NMR in organometallic chemistry and Recent advances in the application of organometallic chemistry in organic synthesis and catalysis.* The course format includes lectures by leading national and international experts, experimental demonstrations, research paper presentation by participants, and panel discussion.

Selected participants will be provided local hospitality and travel support to the extent of Sleeper Class fare. Interested persons may please write to the Convener specifying date of birth, qualifications, subjects of teaching, areas of research, complete address including e-mail. Senior research scholars wishing to participate in this program may also apply with a letter of recommendation from their supervisors.

Dr. Sujit Roy, Convener,

2nd Winter School on Organometallic Chemistry, Chemistry Department, IIT Kharagpur 721 302

Fax: 03222-82252; E-mail: omschool@chem.iitkgp.ernet.in;

<http://www.iitkgp.ernet.in/acads/depts/chem/announce.html>

Last Date for receipt of applications: November 30, 2000