A Peephole in to The World of File Sharing

By

Firstborn Roy Sumer
Information Scientist
Central Library
North Eastern Hill University
Mayurbhanj Complex: Nonthymmai
Shillong: 793014
E-mail : fbornsumer@yahoo.com

ABSTRACT

Paper discusses the latest developments in file sharing on internet. The functioning of peer to peer network as file sharing method is discussed in detail. Concludes that peer to peer network method will be most suited for library consortia

KEYWORDS : File sharing, Information retrieval, Peer to Peer Network, Library consortia

0. INTRODUCTION

Leaping into the 21st Century saw the need for the conventional library to introduce computers or high-tech electronic gadgets and database management systems for its day-to-day functions. This became evident, due to the rapid advancement of Information Technology especially in the form of various applications over the Internet leading to Information Explosion. Thus the need of the hour to handle this is the integration of all library functions into a network so that information retrieval, storage and dissemination can be managed in an organized, efficient and faster way. To balance this scenario, library professionals, both in academics or administration environment have to keep pace with in order to familiarize themselves with this environment. They too have to gather, equip, share this knowledge and most importantly, adopt it for handling their day-to-day profession. Therefore, for development what is essentially required is for this type of network to expand. This could be achieved by investing in technological infrastructure especially in the form of distributed information systems. In turn, it would definitely encourage library and information science professionals to communicate quite frequently. By adopting this, would mean the transition into a new area of distributed information storage, information retrieval and network architecture commonly known as file sharing and Peer-to-Peer network respectively designed according to some standards and protocols.

1. PEER TO PEER NETWORK

The sharing of files or documents, which is stored in your hard disk with two or more computers through a network, is known as File Sharing. This concept is not new at all, as presently the trend is that information can be shared by hosting it on a website, send as an attachment via e-mail or by physically passing on to your friends or colleagues through floppies or CD ROM or through a LAN. Retrieving information through the former would mean that you have to type in the keyword/title/author and wait for the search engine to return you with a link(s) from a central server. There would be no guarantee that you will retrieve specific information as desired. The latter requires you to make a personal request to the author(s) with a hope that he or she would oblige. So, in both cases we have to adopt a search and wait policy in which uncertainty prevails. Alternatively, the absence of a centralized web server where all the information is stored represents the basic standard structure in the architecture of a Peer-to-Peer network. Here, each and every individual computer distributed throughout the network runs a piece of software that functions simultaneously as a server and a client since the transfer of information in this case is bi-directional. Information, which is meant to be shared, resides on any of the computers in their respective shared folders. Retrieving it does not depend on a single central web server but can be obtained from any peer machine or computer that is 'active' on the network. In this case, the chance of retrieving, specifically, a desired piece of information is high due to the presence of many sources. This is a classic example, which differentiates it from a client-server network.

2. ISSUES FOR FILE SHARING IN LIBRARIES

The adoption of file-sharing process among library and information professionals raises various issues that need to be addressed, enlightened, and debated so as to allow the users to come to a common platform. Complex Intellectual Property Rights issues, which include Copyright Infringement, be clarified to avoid violations. To start with let us decide that files or documents especially technical essays which have already been published be allowed to be shared by their respective author(s). A file has to be save with a standard format or extension, preferably .pdf. Naming the file to be shared is crucial and needs to be carefully done according to some standard and format. The reason is that usually in this type of design, the search is done only on the name of the file not the content. This can be decided by inclusion of the author(s) name and one or two keywords along with the title in the filename. Searching of files too has to be decided either by title/author/keywords or by the combination of the above. Its relevance with the available search options should be adopted.

As a justification for the proposal, an investigation was carried out using a PentiumIII, 650 MHz, 64MB RAM running on WindowsMe and a 56Kbps modem (dial-up access) on a renowned Peer-to-Peer network known as Gneutella that is fully functional on the Internet and which we can avail of for free. A few client softwares, available as freeware can be downloaded from their respective websites. In our case, we chose XoloX1.57 build 768 because of obvious reasons. Its simplicity in its interface, ease to install and guarantee to work. There are three main functions which we can

perform using the client software apart from hooking the computer to the network. We can search, transfer and view our downloaded files. Switching between these three activities can be done using the three separate available buttons. For searching a document, the name of the file we are looking for is entered in the search area under the Peer-to-Peer file sharing option. Select 'documents' as the type of file we are interested in. A request for the particular document was sent out instantly after clicking on the 'go' button. In less than a minute, the search window is displayed with results from other peer machines that match the query. A requested filename e.g. 'the.pdf' returns randomly with 109 results within a time frame of 15 minutes which keeps on incrementing as more peer machines acknowledge. All files listed contain the word 'the' in their names and are in Adobe Acrobat format. Thus knowing its exact name is vital in order to get the desired file instantly. This software has a unique feature called a Score. The score represents the number of peer machines that is currently online which has the same file. Double-clicking on a filename with the highest score increases the chance of getting the file early. The filename then can be viewed on the transfer window as XoloX connects to the peer machine to start downloading it. A file can be downloaded very quickly if connection is attain simultaneously with multiple peer machines that contain the file. Once downloading is completed, the file resides in 'downloads' – a sub-folder of XoloX under program files. Similarly, sharing of downloaded files can be done by clicking on the 'sharing' tab of the preference dialog box under the main menu Options. Location of completed and incomplete downloads can be assigned manually under the download tab. In case Internet is disconnected, the incompleted files that are being downloaded resumes their downloading process once its connected again. A few other options such as controlling of uploading speed, filtering and activating antivirus protection during downloading too are present. XoloX even supports web search, web address and ring tones search.

The complete process describe above, revolves around a concept in which machine 1 sends a request to machine 2. Machine 2 records the request and forwards it to machine 3. Machine 3 returns this matching request if available locally to machine 1 through machine2. Otherwise, machine 3 forwards it to machine 4 and so on so forth till the request, which has a 'time-to-live' dies off. Each peer machine recognizes another on the network by their IP Addresses.The success of the above network is its simplicity since it uses HTTP (HyperText Transfer Protocol). The advantage here is that two sites can still communicate, even if one is behind a firewall, assuming that this firewall communicates to the web server through port 80.

3. CONCLUSION

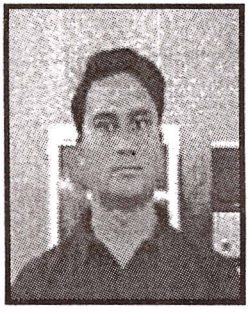
Thus the real challenge lies in the ability to develop a standard Peer-to-Peer network running similar applications meant strictly for the library and information science community and following protocol that is supported by the Internet, especially the World Wide Web. Moving in this direction, would be a positive step since the future in promoting and evolving library and information science into newer dimensions, will depend heavily on how we are able to exploit the Internet. This would immensely benefit the professionals

sharing their resources in their pursuit for academic excellence. The latest consortia and search engines should adopt this concept, as it has immense potential for libraries of the future.

REFERENCES

1. How File-Sharing works – Marshall Brain

BRIEF BIOGRAPHY OF AUTHOR



Mr. Firstborn Roy is Information Scientist from Central Library, North Eastern Hill University, Shillong. He has done B.E. (Electronics & Communication), P.G.D.C.A. and BLISc. He is actively involve in the current automation process in the Central Library NEHU.