CD-ROM networks: from platform-dependent local networks to seamless global access

KI-TAT LAM
Hong Kong University of Science & Technology Library, Hong Kong

Abstract: This paper discusses CD-ROM networking technology and the issues of managing a CD-ROM network in an academic library environment. These issues include search software installation, workstation management, network licensing and usage monitoring. Despite problems with the above issues, the prospect of CD-ROM networking technology looks promising, since there is a growing interest in and a greater acceptance of the technology. Future trends of CD-ROM networking are also discussed as more and more CD-ROM database publishers and online service vendors are providing their databases on the Internet. This move would undoubtedly enlarge the user base and change libraries' collection development strategies.

Keywords: CD-ROM networking

1. Introduction

CD-ROM was introduced to the information world in early 1985 (Ref 1). At that time, very few people could foresee the impact of this shiny silver disc in changing the way of accessing information. Even fewer could envision that by equipping CD-ROM with networking capabilities, the mechanism of providing electronic information in a library environment could be greatly transformed. Libraries do not have to depend solely on expensive online services for electronic information.

BiblioFile was the first CD-ROM title introduced to Hong Kong libraries in 1987. The availability of this handy cataloguing resource indirectly stimulated the discontinuation of a joint cataloguing effort among Hong Kong academic libraries. Before that time, MARC tapes received from the British Library and the Library of Congress were loaded to a central database located at the University of Hong Kong which provided multi-user access for member libraries. Access to the BiblioFile CD-ROM workstation, however, was restricted to one person at a time. Feeling frustrated waiting in the queue, one day a colleague asked the author whether it was possible to install the CD-ROM drives on the network. The idea of putting CD-ROM discs in a networking environment was not new at that time. In fact, the networking solution began to appear in the western market in 1989.

In 1991, the Hong Kong University of Science & Technology (HKUST) installed the first CD-ROM network system in Hong Kong. Since then, the two Hong Kong public library systems, other academic libraries and some special libraries have followed the trend; some even provide remote dial-in access to the CD-ROM databases. Today, HKUST Library has over 150 networked CD-ROM drives and seven networked CD-ROM jukeboxes (each holding up to 240 discs), serving the campus with around 1000 gigabytes of online electronic information.

2. CD-ROM networking technology

After nearly a decade of development, CD-ROM networking technology has become more reliable and sophisticated. In summary, the different types of library CD-ROM networks are identified below.

2.1. Attaching CD-ROM drives to a workstation

This architecture simply attaches CD-ROM drives to a workstation as a sharable resource. It is a straightforward structure that links CD-ROM drives on a peer-to-peer network (e.g. Microsoft Windows for
Workgroups, LANtastic, etc.). Access to the drives is then granted to other peer workstations; just like granting an attached printer for use by the peer. Peer-to-peer networking is mostly suitable for small organisations. Its performance deteriorates sharply when too many workstations are added. Thus, such a topology will only be useful in small library settings — or in a 'large' library aiming at providing a small dedicated network of CD-ROMs.

2.2. Attaching CD-ROM drives to a fileserver

CD-ROM drives can also be accessed on the network by attaching them to a fileserver. Third party software is required to make the drives appear as a mounted network volume on the fileserver. In a Novell network environment, such products based on NLM (Netware Loadable Module) technology are available; for example, SCSI Express, CORELSCSI and others. (Ref 2)

2.3. CD-ROM server

The most popular CD-ROM networking topology employed in Hong Kong libraries is based on CD-ROM servers. A CD-ROM server is actually a computer (usually a DOS PC) with a number of SCSI CD-ROM drives. Proprietary CD-ROM networking software (e.g. Meridian Data's CD-Net, Logicraft's LanCD, etc.) is required.

These software are usually DOS-based and consist of two parts: the server module and the client module. The server module runs on the CD-ROM server. It listens to the network for requests to access data on the CD-ROM drives. When asked, the server will retrieve the data from the CD-ROM drives and send it through the network to the client workstation (also a DOS PC). The client module, which runs on the client workstation, will then pick up the data from the network and send it to the CD-ROM application. Apart from providing the connectivity services, the CD-ROM networking software will also provide other functions including data caching, metering, and load balancing. This approach is applicable on a wide range of network operating system (e.g. Novell, Lan Manager, DEC Pathworks) and on top of a number of protocols (e.g. TCP/IP, DECNet).

2.4. Multi-platforms access and remote dial-in

Since most of the CD-ROM networking applications used in libraries are PC-based, solutions are required for non-PC users on the network (e.g. Macintoshes, Unix workstations, dumb terminals, etc.), as well as for remote dial-in users that are not on the network (Ref 3). This is usually accomplished by adding a blackbox to the network. Users on the network will open a terminal session and login to the blackbox via a terminal server. The blackbox will act as a client workstation and redirect the keyboard and screen input/output back to the users (again, via the terminal server). What is in the blackbox? The major component is an array of PC-processor boards, acting as client workstations. These boards are connected through serial lines to the terminal server.

The set-up for remote dial-in PC-users is very similar to the one mentioned above. Instead of connecting the processor boards to a terminal server, they are connected to a pool of dial-in modems (a communication gateway). Remote communication software (e.g. PCAnywhere, Carbon Copy, Close-Up, etc.) is run on both the remote PC and on the processor board for the redirection of keyboard and screen input/output. Although this approach is workable, a lot of problems exist in the following areas: correct terminal emulation and keyboard mapping; downloading/printing of data directly from CD-ROM applications to the user; speed performance; type of display mode supported (e.g. Monochrome, VGA or SVGA); handling of graphics and multimedia; and support of a mouse device.

3. Management issues

CD-ROM networks provide database access for less money, and with fewer complications, than was available previously. Asian libraries can save significant amounts of money by using discs, rather than paying long-distance charges and higher access fees to the major international dial-up vendors. (It is only in the recent years that these online services vendors started to market their products on the Internet — a move that may make them economically competitive. This issue is discussed further in Section 4).

However, CD-ROM networks suffer from a number of limitations (Ref 4). In a campus environment, there are hundreds of networked CD-ROM titles and thousands of workstations. One cannot expect the
user to modify their workstation’s boot-up settings, to mount the CD-ROM drive manually, or to type commands at the operating system prompt for loading the search software each time they access a database. In short, a standard workstation set-up and an automatic program loading procedure are a must. To achieve this, the following solutions are recommended.

3.1. CD-ROM search software installation

It is important that CD-ROM search software is installed on a fileserver rather than kept in each workstation. This strategy will greatly simplify the task of updating the software, and at the same time it eliminates the trouble of reinstalling the software on a workstation whenever it is accidentally modified or deleted by users.

Avoid buying CD-ROM titles that are not network-capable. Software is not network-capable if it creates temporary/configuration files with unique names and in a unique directory. In the worst case, you may still install them on to the network fileserver. Create a batch file that will copy the search software to the workstation’s local hard disk, and then run the software from the local hard disk. This strategy has a copying overhead, but it works.

CD-ROM network administrators may prefer to give read-write privileges to public workstations for accessing the search software on the fileserver. Although this can save their time on handling those programs that require read-write access privileges to the residing directory, it raised a serious security issue. Files on the server are subject to infection by computer viruses if they are not properly protected. Users may crack the menu system and gain access to the operating system; if network files are not flagged as read-only, the cracker will have the chance to delete files. It is therefore highly recommended that only read privilege be granted to CD-ROM users. To deal with those programs that require read-write privileges, you may adopt the batch copying strategy discussed above.

Some software requires a key-disk to be inserted in the floppy disk drive before running. It is impossible to put this type of CD-ROM title on a network environment. Other search software may require a login password. Obviously, they are not designed with public access in mind. To handle this, you may use a batch file to display the password on screen before loading the program.

Buggy CD-ROM software may load a program into the workstation’s memory and fail to release it when finished. One way of cleaning these residues is to make use of mark-and-release memory utilities available in the public domain. The mark TSR (Terminate and Stay Resident) program will mark the current free memory address. You then run the search software. When finished, a release program will be called to clean up the memory between the last mark and the current free memory address. These utilities only work with programs that stay in the DOS conventional memory.

3.2. Workstation management

It is important that a standard boot-up configuration be maintained on the workstations to access the CD-ROM networks. This will ensure the workstations have a transparent environment. The CD-ROM network administrator can then install and test a new CD-ROM search software based on the known workstation environment before rolling it out to the public. This requirement, however, is not easy to accomplish; especially in a distributed network administration environment where different departments of the same organisation may have their own autonomy of configuring their workstations and file-servers.

One of the big headaches in the DOS environment is dealing with the conventional memory limitation. There are still a lot of DOS-based CD-ROM search software that require huge amounts of free conventional memory. After loading DOS, the network operating system and the CD-ROM client module, the free conventional memory available for the CD-ROM applications becomes very limited. The situation will get even worse when the CD-ROM applications require loading of extra TSR. In the case of CJK (Chinese, Japanese, Korean) applications, you need to load the CJK processing system before being able to display the vernacular characters, and this accounts for another 100–200 KB of valuable conventional memory. It is therefore very important to make sure that the standard workstation configuration will take up as little conventional memory as possible, and will not conflict with other system priorities. This, again, requires careful planning by network administrators.

It is important to reserve a standard set of DOS drive letters on the workstation for mapping to the CD-ROM drives on the network. Users should not be allowed to use this range of drive letters for other purposes. During installation, many CD-ROM search software requires hard-coding a drive letter to their configuration file. If this drive letter is not reserved and is being used for other purposes, the application
will be unable to run on that workstation.

In a campus environment that has a lot of public workstations, it is paramount to provide a menu interface for users. Most menuing software has a feature that requires a password before exiting to the operating system prompt. This will greatly help in preventing files on the workstation from being modified or deleted by users. A menu interface can also present the CD-ROM titles on the screen in an organised manner. When the user selects an application from the menu, a script or a batch program can be launched. This batch program will set up the required environment for the application (e.g. setting environment variables, mapping a drive letter to the CD-ROM drive, copying files or programs to the local hard disk, etc.) and automatically load the search software.

More and more CD-ROM search software now runs on Microsoft Windows. It is not a simple task to support mixed DOS and Windows applications on the same menu interface. There are two approaches to handle this. If you have more DOS applications than Windows, you may want to boot the workstation to a DOS menu interface. When a Windows application is selected, a DOS batch program will load Windows and run the application right away without the PROGMAN shell. Although the Windows loading overhead is quite large, it works quite well. Another approach is to run DOS applications from within Windows. When the DOS application is selected from the Windows desktop, a PIF (Program Information File) will be executed to load the application in a DOS window. The PIF is very similar to a batch program but it is not well constructed. It is very difficult to write a PIF that can function at the same level as a batch program.

Microsoft has now announced the release of Windows 95, but it is still unknown what impact it will have on handling DOS CD-ROM applications.

3.3. Network licensing and usage monitoring

Many CD-ROM vendors require the purchase of a network license. It is important to ask the vendor whether the number of licensed users is the total number of workstations on the network or the number of concurrent users accessing the title. Although there are thousands of workstations on the campus network, not all of them would be used for accessing the CD-ROM networks. In such cases, the vendor is actually asking for the size of the campus community to determine the pricing for a site license.

CD-ROM vendors will usually require the libraries to restrict the access to the number of concurrent users purchased. Some CD-ROM networking vendors will build this metering function into their products. If the number of concurrent users exceeds the number specified, the access session will be aborted. CD-ROM access metering is very difficult to achieve at the network operating system level. The network operating system may be able to restrict the number of users accessing a CD-ROM server, but it cannot restrict or monitor the access to a particular drive on the server. One way to handle this problem is to develop a custom-made program on your own.

Most of the menuing software will keep a log of the use of its menu options. These transaction logs can be compiled into usage reports. Since the license fee for networked CD-ROM titles is not cheap, it is important for the library to monitor the performance of each individual CD-ROM title.

4. CD-ROM network and the Internet

The development of larger, faster, and cheaper hard disks and the rapid growth of Internet access in the information world will certainly have a significant impact on the progress of CD-ROM networks as a medium for providing online information services.

4.1. Hard disk vs CD-ROM

CD-ROM has been well-known for its capability of holding large amounts of data (0.6 GB). However, it suffers from the limitation of access speed. To resolve the speed problem, many CD-ROM network software provide powerful data caching capability. Other CD-ROM drive manufacturers have been focusing on developing faster CD-ROM drives and even larger storage capability.

However, the appearance of faster, larger and cheaper hard disk drives in the market has caused much attention. With the currently available technology, the access speed of hard disks is significantly faster than a CD-ROM. And, to our surprise, we found that the new generation of hard disks in the market are sold at a price comparable to CD-ROM drives. The price here means the dollar per megabyte of storage. It is expected that the trend of loading CD-ROM data to hard disk will be growing in the immediate
future. In fact, commercial products such as SilverPlatter’s ERL have set an example. CD-ROM data is loaded to a hard disk on a Unix platform to form an ERL server available on the Internet.

4.2. Standard retrieval protocol and client/server technology

There has been a strong demand to standardise the communication between the search software and the database engine in a client/server environment. An emerging standard that is designed for this purpose, Z39.50, has finally received much popularity. Z39.50, an information retrieval protocol, specifies how the client sends search queries to the database server; and how the database server responds and returns the search results. More and more database producers and vendors are adopting this standard.

CD-ROM database vendors will start to develop client/server products based on this retrieval protocol. To achieve this, they separate their search engine from the search software. The search engine will be installed on a server machine that has access to CD-ROM drives and/or a big hard disk, holding the database. The search software will be available for a range of client platforms, including Windows, Macintosh, X Windows and VT terminal sessions. The Z39.50 retrieval protocol will be used for communication between the search software and the search engine. Both the server and the client workstation will be connected to the Internet.

It is very unlikely that the server machine will be running under DOS because of many inhibiting factors of this operating system in a network environment. In other words, CD-ROM networking vendors will have to shift their emphasis on DOS-based CD-ROM servers to products that can run on other operating systems, such as Unix and Windows NT.

4.3. Internet access

Recently many large CD-ROM database vendors and online service vendors started to put their databases on the Internet. They want to take advantage of the popularity of the Internet and to extend their markets to this ever-increasing community in the information world. Since World Wide Web (WWW) browsers (e.g. Netscape) are popular tools for Internet access, some of these vendors have built a WWW gateway on their database server to allow Internet users to access their databases without specific hardware/software installation on their computers.

It is very likely that CD-ROM database vendors will continue to distribute their products in CD-ROM format for some time. But they may start to consolidate their CD-ROM products so that a unique search software can be used to access both the locally mounted CD-ROM database and the remote database they put up on the Internet.

5. Conclusion

CD-ROM was originally designed for use as a data carrying media for personal computers. It was then quickly adopted by database publishers as a medium for distributing their databases. As the need to have simultaneous access of these databases by multiple users arose, CD-ROM networking vendors started to emerge. However, these CD-ROM network systems have frequently suffered from a number of limitations on their implementation and use. Current development of this technology is to focus on allowing faster data access, user location independence, multi-platform access, and a customisable and standard user interface.

With the ever-increasing popularity of the Internet, the adoption of a client/server approach and the development of the standardised information retrieval protocols (Z39.50), the CD-ROM networking business is moving in a new direction. CD-ROM database sets will be available on the Internet, searchable by user-friendly World Wide Web client programs — independent of size, location, platform or access.

Ki Tat Lam  
Library  
Hong Kong University of Science & Technology  
Clear Water Bay  
Kowloon  
Hong Kong  
Tel: +852 2358 6702  
Fax: +852 2358 1043
References