OPTICAL STORAGE DEVICES AND THEIR IMPACT ON LIBRARIES AND INFORMATION RETRIEVAL SYSTEMS

H Y MAHAKUTESHWAR
NICFOS Computer Centre
CFTRI, Mysore-570013

INTRODUCTION

Optical storage devices represent some of the most viable innovations for the improvement of the cost-performance relationships in computer and information management fields. About ten years ago it was becoming evident that optical storage of information would have revolutionary impact on the office automation as well as the publishing industry. Commercial exploitation of optical storage has come more slowly than expected in the industry. But everything is now set for rapid growth in at least some segments of this new technology. Just ten years ago it was not possible to even dream that the most promising element of the optical storage technology would be based on a new form of consumer phonograph. One of the most important optical storage technologies viz., CD-ROM (Compact Disk-Read Only Memory) - is based on this new form of consumer phonograph and it offers tremendous application potential specially in the field of information storage and retrieval systems involving large bibliographic databases.

One of the main reasons for this rapid growth is the continuing demand placed by database users for cheaper, more reliable and high density data storage devices to keep up with the capacities of their mainframe systems[3]. As the volume of data that a machine is capable of processing becomes greater, the capacity of the accompanying storage devices must also increase to handle this large volume. For the last 35 years, the density of magnetic recording media has doubled every two years. Increase in recording density in rigid devices has reduced the read/write head to medium spacing from 25 microns to 0.3 micron during the period. Though current laboratory experiments are investigating a spacing of just 0.1 micron, it is thought that for many purposes, magnetic media have reached their optimum. Though removable winchester disks of 300 MB capacity are expected in the near future, many are looking to optical erasable disks with a capacity measured in gigabytes as the future solution to their storage problems.

BILLION DOLLARS INDUSTRY

Already up to 11,000 CD ROM players have been sold and the majority of them have gone to system integrators. By 1990 the optical storage devices are expected to outperform the magnetic storage devices such as floppy disks in terms of sales. Though estimates vary, Mr. William Gates, Chairman, Microsoft, however expects only 10% penetration of office desktop systems.

Whatever may be the exact rate of growth, the market is expected to be measured in hundreds of millions of dollars. The San Jose market research firm “Dataquest”, gives a conservative estimate that sales will have reached a value of $700 million by 1990. Ed Rothchild, consultant and publisher of “Optical Memory News” goes further to predict a market of $13 billion by the same date. Wherever the truth may lie, optical disks seem to stay, displacing magnetic disks from their more than 3 decades of utilities [3].

ORIGIN OF THE TECHNOLOGY

Progress achieved in home audio entertainment, consumer electronics or high fidelity (Hi-fi) systems is remarkable. With the entry into the age of computerised electronics, a revolution is being witnessed in the machines for playback music. The digital audio disk (DAD) has been launched carrying the commercial name of compact disk (CD) and is most likely to supplement if not completely replace the audio disk or LP.

DAD OR CD AUDIO TECHNOLOGY

Now a new type of ‘record’ and ‘record player’ are in the western market. The CD is only 120
mm (about 5") in diameter, yet it has a playing time of 60 minutes with the sound information on one side.

The CD rotates like a record, it is not touched by a stylus but scanned by a small low powered laser built into the CD player. The disk provides information about the sound in the form of binary pulses or Os and Is, in digital terms. The rate at which the bit stream is obtained by detecting microscopic pits on the disk is constant and is equivalent to 4 million bits per second. The Disk rotates at 500 rpm when the laser starts at the center and slows down to 200 rpm as it follows the spiral of pits to the outside.

Having been detected by the laser and converted to electrical pulses by a photo-diode, the digital codes are separated into two channels for stereophonic reproduction, and eventually converted to analogue sound using digital to analogue converters (DACs).

In practice, a CD player is an extremely complex piece of equipment. Being compatible with most existing systems, conventional disks can still be used alongside the CD player without abandoning the treasured collection of LPs[2]

WHAT IS CD-ROM

CD-ROM player and disk technology is an extension of CD-audio technology using exactly the same recording and reading techniques as the most successful DAD players. Thus optical disks are going to be practically of very low cost. Giga disks, as they are called sometimes, are expected to cost between $10 and $20 per gigabyte in the next few years. In addition they provide mass storage (currently between 500 MB and 4 GB) enabling a lot of data to be stored on very little physical space. Being small and lightweight, optical disks are transportable and therefore offer an alternative means of database distribution to on-line services. Already it is possible to build up a complete database system including all the hardware for just $30,000 to $40,000. Optical disk drives now cost $1000 to $1500 with a disk capacity equivalent to 10,000 pages of typed script. Within few years this capacity is expected to increase to over 250,000 pages for the same price. Linked via a suitable interface to a personal computer (PC) a CD-ROM player and disk allow an end user access to database without the cost constraint of variable charges (such as telecommunication tariffs and host's time related access prices).

With these cost, storage and distribution characteristics CD-ROM is widely considered to have important application potentials, notably in publishing, library, documentation, archival and information retrieval systems[1].

Presently there are three main types of optical storage devices viz.:

(i) the read only optical disk (CD-ROM)
(ii) the write once read many (WORM) optical disk and
(iii) the erasable/re-recordable optical disk.

There are also other kinds of compact disk called compact disk interactive (CD-I), compact disc video (CD-V), compact disc interactive video (CD-IV), laser vision-read only memory (LV-ROM) etc. whose description is beyond the scope of this article.

Currently only the CD-ROM and WORM disks are in manufacture. In the future write once optical disks are expected to dominate the market. Erasable disks are also in the horizon. Already at the prototype stage, they may be seen in use very shortly [3].

STRUCTURE OF CD ROM

The basis of the composition of CD ROMs is a plastic substrate. Usually the substrate is a polycarbonate, but there are experiments using polymethyl metacrylate (PMMA) instead. The plastic substrate supports a reflective aluminium layer just 500 Å thick. Over this there is another very thin protective layer of plastic. The interface between the layers has a series of more or less equally spaced pits or bumps. Such pits would be about 0.6 mm across and 1.6 mm apart on a typical disk. The pits follow a long circular pattern- usually a spiral as on an LP recording.

The information embossed on the disk at manufacture in this way is read by a 780 nano-
OPTICAL STORAGE DEVICES

A meter laser beam focused on the aluminium layer. The laser beam passes through the thin plastic coating without difficulty, and also through small amounts of debris and dust on the disk. The disks are recorded in a constant-linear-velocity mode. This gives the disks a capacity of at least 540 MB; a figure which is rapidly increasing. The storage face is divided into blocks of 2048 bytes with a typical field of 270,000 blocks on one disk. The linear density of optical disks is about 35,000 bits per inch. With a track spacing of 1.6 microns, an optical disk can store at least 20 times the capacity of a magnetic disk—some 560 million bits per square inch of surface.

STRUCTURE OF CD PLAYER

A series of prisms, beam splitters, mirror lenses, wavelength filters and polarising filters directs the laser beam and detects the patterns of scattered light from the disk. The major components are mounted at the side of the drive while others are transported on a moving carriage which carries the optical-head reader across the disk surface. It is this moving carriage which gives the optical disks their high random access capability.

Apart from the tracking mechanism, there are two types of movement in aiming the laser beam at the disk. The first of these is a fine adjustment galvo-mirror which deflects the laser beam to modify its angle of incidence in relation to the CD ROM. This gives the tracking head the fine degree of accuracy that is necessary. The second is the in and out movement of lens mounted voice-coil type actuator, which maintains the focus of the aluminium surface in relation to the tracking head.

Many optical disk systems contain embossed tracking grooves that absorb or scatter light. The tracking photodetector regulates the running of the drive according to the intensity of light falling on the walls of the tracking grooves and on the bottom of them. The variation in intensity of light is picked up by the photodetector which adjusts the lead orientation angle accordingly.

ADVANTAGES

The main attractive feature of optical disk storage is the flexibility of the media. The optical disk can store data in four forms viz.:

(i) audio
(ii) compressed audio
(iii) video and
(iv) compressed video

It can also store in usual data formats. With the introduction of optical disk technology, the complex formatting procedures used by computers for storage of data generated from varied inputs such as sound and video are simplified.

In future, optical disks are expected to have extremely high access speeds reaching up to 150 K bytes/sec. For small systems their speed will show a great advantage over the on-line baud (bits/sec) rates of 1200 bits/sec.

The compact nature of optical disk systems gives them a great advantage where space is limited. Disk densities are 40 times more than those of magnetic disks. Storage modules of the order of tens of gigabytes in a few cubic feet can be built up in a single system.

There are two main reasons which hold back optical disk storage at the moment. The first is the fact that few write-once disks and no erasable disks are currently available in the market. Until these are available, optical disks are only useful for databases which seldom need updating. A good example of such databases is the bibliographic databases of the secondary or the abstracting periodicals. As soon as erasable disks are available, the situation will change dramatically.

Another attractive feature of CD-ROM is that it uses the same technology as CD audio. Hence manufacturers' investment in production facilities are minimised. Given the capital costs involved, this is an important economic advantage.

LIBRARY AND INFORMATION SCIENCE APPLICATIONS

Hutnik [3] opines that library applications feature high on the list of products available on optical disk. Grolier Electronic Publishing have produced an optical disk version of the "American Academic Encyclopaedia" for $199. A
library of public-domain-PC-software (software available freely for copying and use as popularly known in USA) is offered by Library Corp. A firm by name Reference Technology has produced a CD ROM disk which offers over 9000 different software packages. Most of these products are also sold along with complete system hardware because of the low cost of optical disk readers.

In general it is library applications which will first see the influence of CD ROMs. Libraries, archives, hospitals and insurance records are all ideal sites for optical disk storage technology. With this in mind Oxford University Press is putting the Oxford English Dictionary on optical disk. The product should be available by 1990.

SOME OF THE DATABASES AVAILABLE ON CD ROM

Chemical Abstracts (Health and Safety in Chemistry)

A portion of Chemical Abstracts service is being made available on CD ROM. Some 50,000 abstracts in the area of environmental chemistry and toxicity are readable on compact optical disks as "Chemical Abstracts: Health and Safety in Chemistry". The disks use search and retrieval software for use with Digital's Micro VAX I and II and VAX station II computers and the IBM PC XT. The CD ROM file is offered on an annual subscription basis with quarterly disk updates for $1195.

UKMARC file, the Whitaker file and the Conference Proceedings Index

The British Library demonstrated the advantages of using the compact disk for storage at the London Online Meeting and earlier at the October 85 Frankfurt Book Fair. The demonstration disk which has the potential for accommodating 150,000 pages, contained samples of three of the library's Blaise online bibliographic files, i.e. the UKMARC, the Whitaker file and the Conference Proceedings Index[6].

CAB Abstracts

The compact disk version of Commonwealth Agricultural Bureaux is now being tested in Mexico, Italy and Malawi. The CD ROM which represents an alternative form for the online version, contains nearly a quarter of a million abstracts record for 1984 and part of 1985, a 40,000 word thesaurus, and the Micro Basis retrieval software in two versions for IBM PC and Micro VAX I/II.

The disks will eventually be tested in 30 countries many of them in the Third World. The CAB system group are presently assessing the viability of CD ROM application for further distribution of CAB abstracts database, expert systems, books and unique databases formed from subsets of various source databases[4].

H.W. Wilson Company Database

H.W. Wilson Company is bringing out 12 separate CD ROM databases such as Applied Science Technology Index, Art Index, Biography Index, Social Science Index etc., with four different access modes.

Educational Resources Information Centre (ERIC) databases

OCLC has become the third organisation after Dialog and Silver Platter to market ERIC databases on CD-ROM.

Library and Information Science Abstracts (LISA)

Silver Platter has released the Library and Information Science Abstracts on CD-ROM. The disc includes the database from 1969 to 1986 and the database and the software would be updated annually. This is even operational for retrieval with Informatics India - a Bangalore based private firm.

OTHER DATABASES

Some of the other databases available on CD-ROM are:

1. Dissertation Abstracts
2. Aquatic Science and Fisheries
3. MEDLINE CD
OPTICAL STORAGE DEVICES

4. Poisindex: A toxicology database
5. Dialog on disc
7. INVESTEXT
8. DISCLOSURE II
9. PROMT
10. ABI/INFORM
11. Media general's Market File
12. Marquis' Who's who in Finance and Industry
13. Academic American Encyclopaedia (Electronic Encyclopaedia etc. [7]).

ACTIVITIES CONCERNING CD ROM

The world's first international conference on CD ROM technology was held in March 86 and brought together over 900 software developers and database publishers. Organised by Microsoft, the conference was an official recognition of the increasing importance of optical storage devices to the computer industry.

Another body, the Optical Disk Forum, met under the auspices of the European Commission (DG XIII) in Luxembourg on 22nd January 1986, to discuss the early implementation of standards for data retrieval from optical disks. Standards already exist at the level of hardware for CD ROM, and agreements on standards for the rapidly evolving database market are urgently needed to maximise the potential of the new technology.

The Optical Disk Forum assembled information providers, CD ROM manufacturers and vendors with the aim of achieving a joint working standard for both Europe and US. It was proposed to set up a technical committee to analyse the emerging American recommendations, to react to them, and propose modifications wherever necessary to suit European requirements.

A conference to discuss the European requirements was held on 16th May 1986 in Luxembourg[6].

Harwel, the producer of STATUS software, made an announcement to hear from publishers who wish to exploit CD ROM technology so that a CD ROM club could be formed. About 50 enquiries were received by July 86 regarding the proposed CD ROM club. The club would enable the new technology to be aimed at specialised professional groups.

CD ROM clubs could offer new cross roads for information providers to exploit the full potential of the European market[5].

CONCLUSIONS

CD-ROM technology works and is most likely to penetrate the conventional storage technology and stay long. However, if optical disks are to succeed, some kind of standard must be agreed upon on both hardware and software. Efforts are also on in this direction among the European countries and the US. Since CD ROMs are interchangeable and retrieval is intrinsic, they bring in considerable amount of flexibility in their use. Closer assessment of the impact of CD-ROM on library and information science would indicate that it might act as an initial substitute for host on-line services such as DIALOG, ORBIT, ESA-IRS etc. The future tendency might be to use host on-line services for current search and also for general subjects not falling directly in line with the special subject fields of interest. For retrospective search for 5 yearly periods, CD-ROMs are likely to be used. Ultimately the end users will be the professionals from institutions and occasionally individuals for whom on-line searches would become an independent day-to-day routine.

Optical disks will be a boon for India providing a cheap and satisfactory solution to database problems. The working conditions here are radically different when compared with the West. Telecommunication networks are often unreliable, and so in-house CD-ROM based systems may become more prominent. Moreover, high humidity during monsoon, and the summer dust will not affect these new storage devices which are made of very stable material (polycarbonate covered with thin aluminium layer followed by another very thin protective layer of plastic) and are not easily damaged. They can even be read through a layer of dust since laser beam is used for reading.
Large databanks can be stored securely on CD-ROMs. Handbooks, encyclopaedias and other reference materials, besides bibliographic databases in the fields of agriculture, medicine, engineering etc., will be available on CD-ROMs for the first time at the touch of a key on the keyboard of PCs. Thus, information transfer - which is as important as technology transfer in modern times- will be enormously speeded up and India's access to technical information and data will be at par with the West.

REFERENCES


4. Compact CAB. Information Market 1986, 44, 7

