

publishing  
on

cd

chapter

fifteen

For conveying large amounts of information, CD can be much less expensive than on-line distribution, and it can sometimes even be quicker!

For example, CDs can be distributed by overnight mail to 1,000 locations for about \$12,000, including production of the CDs and automated shipping and handling.

**t i p**

**A 650-MB CD can store about 13,000 TIFF images, along with a simple index and a viewer to handle the documents. Compared to the simple cost of copying and shipping 13,000 paper pages, CD distribution offers tremendous financial and environmental advantages.**

If it costs two cents per page to duplicate a 13,000-page rich text or image document, that means production cost for each document is \$260. Even if it cost one cent/page on a web press to print the copies, that's still \$130 per 13,000-page document. Getting paper and toner costs that low is virtually impossible, not to mention the cost of hardware and labor.

And printed on 20-pound bond paper, those 13,000 pages weigh 130 pounds, so shipping will not be cheap. And if the cheapest methods are used, it will not be fast. Obviously, it's financially impractical to ship 130-pound documents to 1,000 different locations very often.

**At the famously advertised price of 10 cents per minute, it would cost \$378 to transmit over a 28.8 modem the contents of one CD-ROM full of information.**

On a straight cost basis, mass-produced CD-ROM titles, in quantities as small as a few hundred copies, are very inexpensive. If the CD printing costs 50 cents per copy in quantity, that's a rate of 260 nominal image pages for one cent.

The cost comparison is overwhelming, and the communications time comparison also favors CD-ROM distribution for large quantities of information. To move 650 MB to each of these 1,000 locations through a 28.8 modem running at 3K/second would take 63 hours. So it would require a minimum of three modems at each location just to match the speed of CD distribution via overnight shipping. Considering the connect charges, whether they be simple long distance or some sort of leased line, it is easy to conclude that communications charges will far exceed the cost of \$12 per location offered by CD.

To move 650 MB to each location through a 56-KB line running at 56 KB/second would take 3.3 hours. Obviously, telecommunications options are available that can and do easily move this volume on a constant basis. In fact, 56-KB lines are among the slowest of the advanced communications available, compared to ISDN, T1 and T3 lines. But the cost of communications follows the simple market rule of “you get what you pay for,” and high speed comes at high cost. If a dedicated 56-KB line is \$650 per month, at each of the 1,000 locations in this example, the communications cost can rise astronomically. No one can happily justify a \$650,000/month phone bill to move this amount of information.

tip

**A different problem arises when very large volumes of digital files must be accessed by many people in dispersed locations. At a point that can be easily measured in dollars and number of users, the original cheap data dissemination technology still makes sense: a dial-up bulletin board system, or BBS.**

**To distribute hundreds of thousands of images daily among many hundreds of data entry operators working at home, a dial-up BBS and a 28.8 modem is the proven winner. It's cheaper and faster than all the alternatives.**

Electronic distribution of this magnitude will be more practical when ISDN is universally available. After all, at 128 KB, 650 MB can be moved in just about 87 minutes. At two cents a minute, transmission costs would be only \$1.74. Even at a rate of 10 cents per minute, a dual-channel ISDN could move the CD's worth of data for less than \$12. And the information would get there at least 20 hours quicker than overnight shipping.

And technically, 87,000 minutes of total transmission time would be required to move these 13,000 pages to 1,000 locations. But that happens in a “lights-out” mode, so it does not figure into the equation. However, that's a lot of bandwidth, and bandwidth is like water, a shared resource.

And, of course, on the receiving end a spare 650 MB of storage would be necessary. A pittance on a network, perhaps, but a fairly demanding load for a laptop. Information shipped on a CD comes with its own portable storage media, a considerable advantage.

Type of communication	Megs per minute	Estimated Total Cost	Comments
CD - Shipped Overnight - 18 Hrs.	0.6 Mb/Minute	\$25.00	Cost includes CD, duplication, shipping
Paper - Shipped Overnight - 18 Hrs. for 13,000 pages*	0.6 Mb/Minute	\$560.00 (\$300 ship, \$260 print & paper)	Cost includes duplication (and paper), shipping
28.8 modem - 63 Hrs.	0.17 Mb/Minute	\$378 @ .10/min.	Slow & expensive for large volume
ISDN - 15.75 Hrs. (112 Kbps)	0.69 Mb/Minute	\$18.90 @ .02/min.	Cost includes .02 per minute charges
56 Kb - 31.5 Hrs.	0.34 Mb/Minute	\$56.87 @ .015/min.	\$650 / Mo. Flat Rate
T1 - 1.544 Mbps - 70 Min.	9.26 Mb/Minute	\$1.96 @ 0.28/min.	\$1200 / Mo. Flat Rate

Estimated costs to ship 650 Mb coast to coast.

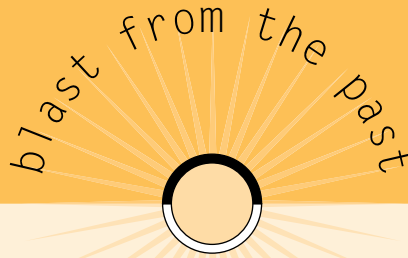
\*Note: assumes 13,000 pages of images at 50K/page. At 2K per page for text, 650 Mb of data is 343,455 pages.

### **Sony Explains The Technology In 8X CD Drives**

"Sony designed the CSD-880E CD-ROM drive to overcome problems inherent with high-speed performance. The new, more robust spindle motor minimizes vibration, spins the disks more smoothly and lasts longer. Sony also expanded the frequency bandwidth of the optical pickup, improved the RF (radio frequency) amplification circuitry, and developed a new digital signal processor for better control over functions such as error correction and speed. All of these improvements combine to lessen vibration and maximize readability for superior performance and faster data access.

"Besides including components for 8X performance, the CSD-880E CD-ROM drive is also designed to utilize very little CPU processing power to increase net system performance. In addition, a large 256-KB buffer enhances system throughput.

"Whether viewing a full-motion video clip, looking up information in a multimedia encyclopedia or playing the latest action-packed game, the CSD-880E CD-ROM drive retrieves information fast and opens files smoothly. Not only does the drive feature a quick average access time of 160 milliseconds, but its high data transfer rate of 1.2 MB per second enables support for MPEG-2 video."<sup>1</sup>



### From The Foreword To CD ROM: The New Papyrus

"Microsoft is extremely excited about the vast potential of this emerging CD ROM technology for a number of reasons. For one thing, CD ROM is quite different from any other medium in existence, whether it be television, movies, video, slides, audio, books or personal computers. In fact, one might look at CD ROM as the summation or combination of almost all of these. But what is it that makes CD ROM so special? On the technical side, because all the information is stored in digital form, the medium is similar to floppy disks, hard disks and other magnetic media. However, the transportable nature of the compact disk and the cost per bit put CD ROM in a class all by itself."

William H. Gates, Microsoft Press, © 1986

## Archiving to CD

CD offers an excellent means of relatively short-term archival storage, for example, for the next several years. Compared to physical or even microfilm archives, documents may need to be stored for 100 years or more. Some of the manufacturers have introduced CD media that is projected to have a 100-year shelf life, but no one expects CD media to last that long. So even though the media is perfectly fine, it may be difficult to maintain the drives themselves for such periods.

Therefore, if we embrace the rate of change, we see the digitization of the information as the ultimate archive technique. Digital documents can be easily migrated to each new generation of storage media as they are developed and attain widespread acceptance.

# Write It Once

If possible, the entire collection should be organized and written once to the CD to take optimal advantage of storage space and file structure.

The content manager, whether called a digital librarian, Webmaster or collection director, must always put the needs of the user foremost in determining the optimum file characteristics of a particular collection. There are also industry standards to be considered, for easier access for the reader:

“One thing described in these standards is the physical format of the information recorded on the compact disc. The physical format refers to not only the dimensions of the disk, but specifies how the information is laid out on it.

“Think of it like a book. Suppose that all books were made in one standard size, perhaps the size of a paper-back novel, and contained the same number of pages. These are the physical dimensions of the book. Now if a book standard describes the size of the type, spacing between lines and page margins in the book, it specifies how the information is formatted. This is the same way these standards document the format of information on a compact disc.”

— “A Brief History of CD Development,” which notes the various Color Book conventions developed primarily by Sony and Philips

This site offers a concise, in-depth history of the development of CD technology:

<http://www.octave.com/ricoh/why.html#stand>

## tip

**In a mass production of hundreds or thousands of CDs, the collection must be tested in the folder and directory structure on hard disk before being mastered for CD publication. It's the only way to ensure that all links will remain intact and that indexing will function.**

## **T** CD Speeds: 1X, 2X, 4X, 8X ...

In the ISO 9660 standard, upon which the original music CDs were designed, the data-transfer rate was 150 KB per second, and this became the basis for "1X." By this measure, a 4X drive has a speed of 600 KB per second, an 8X drive has a speed of 1.2 MB per second (1 MB per second is 10 times faster than 1 KB per second) and so on. The original music CDs, at 1X speed, had a maximum playing time of 74 minutes, so it would have taken 74 minutes to read the entire contents of one CD. At the task of pure data transfer, an 8X CD drive could read all of the data from a full CD in a little over 9 minutes. Of course, these peak speeds are not useful for all applications, such as audio and video, where the presentation speed of the CD contents is of primary importance.

The content manager must also decide on PDF creation options in Distiller and PDF Writer and indexing options in Catalog. Such features as Word Stemming and Thesaurus add to the overhead of the database, increasing the size of the files required for the index.

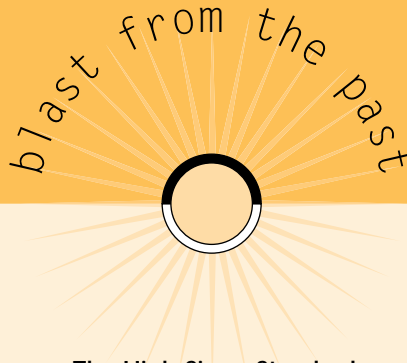
If the collection of documents is relatively small, and let's face it, the 650-MB capacity of a single CD-ROM makes most requirements relatively small, every single Search Expansion option should probably be included.

But in voluminous applications, every variable should be explored and the value of each should be weighed. Of course, it is nice to include Bookmarks, Thumbnails and every possible Search option in the database, but the value of such luxuries must be judged for each application.

You should publish a few sample collections to judge the effectiveness of search tools and options on a specific set of documents. You can experiment by doing Web searches on

In the PDF format files created by Distiller or PDF Writer, there are many options that affect the size of the files. And in HTML, many options affect the files' size of a page, including graphics, backgrounds, Java applets, animated GIF, frames and so on. The author and publisher and Webmaster must look carefully at both the benefits and costs of rich content.

In many cases, CD-ROM publications can go all out with the relatively luxurious capacity of 650 MB and the extremely high-speed access of 8X CD reader drives.



### **The High Sierra Standard**

Philips and Sony did not define the file structure or the logical file format for CD-ROMs. They left that up to other manufacturers concerned with digital computer data. This could have resulted in utter chaos because each operating system had a different format.

Two people working at Silver Platter, Bela Havatny and Parke Lightbown, called a meeting in 1985 of all companies working with CD-ROMs. Individuals from such companies as Digital Equipment Corp. (DEC), Microsoft, Hitachi, Sony, Apple Computers, Philips and 3M attended the meeting.

The meeting took place at the High Sierra Casino and Hotel in Lake Tahoe, Nevada, so when they formed a committee they called themselves the High Sierra Group. They defined a standard that was based on the Yellow Book, a standard owned by Philips and Sony. The standard addressed the requirements of reading CDs on various platforms: Macintosh, MS-DOS, UNIX and VMS. It also was set up as an international standard so foreign characters such as Kanji could be used. It was even designed to later accept WORM (Write Once, Read-Many) drives and rewritable media.

### **The Birth Of ISO 9660**

"The High Sierra standard was accepted by the majority of CD-ROM manufacturers at the time. Within a year it was adopted by the International Standards Organization (ISO) with minor changes. They gave it the cryptic name ISO 9660. It is, however, an important name to remember."<sup>2</sup>

Now that most CD drives come bundled with the computer, standardization is virtually complete. However, when deciding upon CD-recording hardware and software, it is very important to be sure that the equipment meets the predominant industry standards to assure the future usefulness of your published CD.



services such as InfoSeek or Open Text. The more you practice, the more you'll understand what tools and options will deliver the results your audience will appreciate. See Chapter 7 for details on Acrobat Search and Chapter 12 for details on advanced searching techniques.

For the latest information on CD-ROM developments and standards, go to:

<http://www.uscchi.com/cdrom/>

[http://www.nta.no/teiktronikk/4.93.dir/Maartmann\\_Moe\\_E.html](http://www.nta.no/teiktronikk/4.93.dir/Maartmann_Moe_E.html)

**"CD-ROM as a medium is physically identical to the audio CD that was introduced to the world market in 1984.**

**The audio CD has been an unprecedented success in the consumer market. In fact, as far as I know, only two products have wildly surpassed the (often very optimistic) prognoses made by American market research companies: the CD and the fax machine."**

**By Erling Maartmann-Moe,  
12/23/93**

## Acrobat Catalog And Search on CD

On CD-ROM, user software is usually included on the CD. The fact that ISO 9660-compliant CDs can run on many platforms, from DOS to UNIX, makes CD an ideal medium for physical distribution of digital documents.

An example of this is the Adobe Acrobat search engine, which can be used to publish a CD-ROM. A content manager can use Acrobat Catalog to index the collection, to create a database of documents that will be fully searchable on the CD-ROM.

By including the Acrobat Reader for all platforms, users on DOS, Mac, UNIX and Windows can enjoy the benefits of this CD-ROM. The storage requirement for all four readers is about 6 MB.

# Production Via CD-Recordable

CD-R drives now cost less than \$1,000 and come complete with software for writing files to so-called CD-ROM disks. Since the original acronym of CD-ROM meant Compact Disk-Read Only Memory, it's a bit of an oxymoron to say "CD-recordable." Of course, the technology itself is a great combination of widely readable media, a modern-day floppy disk, with the new CD-recordable drives that allow writing files on CDs.

These CD-Rs can be expanded beyond individual usage by simple "ganging" on a network to produce multiple copies of a collection simultaneously. Ganging refers to the practice of connecting tens or hundreds of low-cost disk drives to the network. The alternative for offering network access to large numbers of CD-ROM disks is to connect a jukebox to the net. However, even if the jukebox has multiple picker arms and multiple drives, the ganged array of many CD drives containing many CDs is inherently faster. Since CD drives cost less than \$100, they can be arrayed in stacks. The advantage to the user and the network is that there is no latent or waiting time while a jukebox robots seeks, grabs and loads a CD. All CDs are online, as opposed to the near-line storage of jukeboxes on networks.

Of course, absent intense usage and faced with large volumes of CDs, jukeboxes offer excellent performance and great value.

CD-R is also very useful in producing copies of voluminous collections on an ad hoc or demand basis. In either case, it is essential to have a single-source database complete with all files and indexes to write in one shot to the CD.

## Mass Production

Mass production is the next step for volume more than you realistically produce one at a time on an individual CD-recordable.

When considering the potential of mass CD-ROM publication, you must weigh costs vs. potential enjoyment by thousands of users. The same 40 or 50 cents per page will produce a high-performance CD-ROM collection or a massive collection of boring accounting data. Let's face it; even the most interesting CD encyclopedia is boring compared to best-selling game of *Myst*.

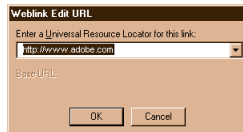
The reason for the excitement of *Myst* is that it embodied all of the smartest technologies that could possibly be put on CD. When you installed *Myst*, you got a Windows version of Apple's Quick Time movie viewer. And you got a database that

tracked your progress through an extremely complex series of situations and choices. The CD offered movies, sound, interactive programming ... Myst first verified the richness of CD publication.

Now every user has cheap access to the tools that were used to create this ground-breaking CD adventure. All of the tools that were used to enhance this fantasy can be used to make all forms of information accessible:

- A tech manual can be enhanced through Quick Time movies that demonstrate a procedure in a way that a series of drawings could never convey.
- A drawing or graph can be mapped at high resolution to provide mouse-click access to information specifically related to that spot in the image.
- A verbal explanation can be heard through PC speakers with a click on a Hints box.

## Web Links



Entering a Web link into an Acrobat file is simple.

Acrobat Web links connect the rich and sturdy format of CD-ROM diskettes to the ever-changing and updated World Wide Web.

The function of Web links depends upon a smooth integration of the user environment. For example, ideally a user could be navigating a CD-ROM publication and have access to some form of network connection to the Internet. This connection can provide access to constantly updated information referenced on the central source of the CD.

With so many Internet service providers offering “unlimited Internet access” for less than \$20 per month, more and more people are going to have their remote PCs connected to the Web full time, 24 hours a day, seven days a week.

Home PCs this year are designed to cost about \$2,500 and they offer an absolute treasure chest of power. Former luxuries are now included for the small office/home office market, such as 17-inch monitors, stereo sound and 3-D graphics. A PC with 32 MB of RAM, a couple of gigabytes of fast hard drive storage, running a 32-bit multi-tasking system like Win95 or NT can easily stay online constantly, performing complex tasks.

The Internet is just another network, so why not stay online?

Of course, you may need another phone line, or perhaps in the near future the phone company will offer ISDN, or the cable TV networks will offer two-way service, or perhaps even cel phone or satellite connection to the Internet will be affordably available.

Your configuration will determine the success of the linkage, depending on your Internet access and software setup. For example, if Acrobat is properly installed as the Helper App with Netscape Navigator, this linkage will be virtually seamless. If you're on a network with a direct Internet connection, you'll barely notice the leap from the CD to the Web.

### **T** Naming PDF Files

The safest bet is to name files according to the least common denominator, which is the DOS convention for names. This means that eight characters can form the filename and three characters can form the extension. For example, FILENAME.EXT, where FILENAME equals the eight-character file name, and EXT equals the three-character extension.

To be fully compliant with the ISO 9660 Interchange Level 1 conventions, use only the letters A to Z, the numbers 0 to 9, and only the underscore and period characters in an eight-by-three file name.

# Summary

Virtually every new PC now includes a CD-ROM reader drive, usually at 4X or 8X speed. A lot of new software is now distributed on CD, along with all user documentation. As a user, you'll print your own documents as needed—at your cost for toner and paper.

Distribution on CD-ROM can be very timely and very cost-effective compared to all other alternatives when a large volume of information must be conveyed. Documents with Acrobat Web links and other hyperlinks can be distributed on CD-ROM and have access to dynamic updates via the World Wide Web.

Looking into the future, with both the past of discarded technology and the present of new technology constantly being announced, it is logical to question the future viability of the currently ubiquitous CD-ROM media.

But even as we ask the question, the basis of the question is changing. One hundred years ago it was vitally important to print books in inks that would not destroy the paper pages. At the time, there was no new media on the foreseeable horizon that could replace paper and books.

Today, when it is obvious, given the history of 9-track tape and 8-track cartridges, that physical media comes and goes, the requirement to preserve information has a new solution. In the past, to move information from an outdated media to a new media, such as from paper to word processing or database format, required physical effort and labor.

Now that information is in digital form, it will be almost instantaneous to transfer the information onto whatever new media may evolve. The information itself, once embodied in digital form, is liberated from the physical media upon which it temporarily resides.

Given all the above, the current dominant, robust physical medium is compact disk. Of course, like all things, their time too shall pass. But because the information is digital, it will naturally migrate to the next hardware and software world.

1 <http://www.sel.sony.com/SEL/ccpg/text/ccpg/flash.html#pr3>, Press Release, Multimedia Products Div. Of Sony Information Technologies of America, makers of Trinitron monitors, 6/3/96.

2 <http://www.octave.com/ricoh/why.html#stand>