1995

NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

AN ANALYSES OF CD-ROM TECHNOLOGY FOR SPACELINK

Prepared By:	Willard A. Smith, Ph.D.
Academic Rank:	Associate Professor
Institution and Department	Tennessee State University Department of Physics, Mathematics and Computer Science
NASA/MSFC:	
Office:	Office of Education Programs
MSFC Colleague:	Jeff Ehmen

AN ANALYSIS OF CD-ROM TECHNOLOGY FOR SPACELINK

by

Willard A. Smith Ph.D.

This study first examined the possibility of placing NASA CD's in the Spacelink Public Electronic Library (SPEL). The goal of this task was to place these materials on the Internet. The second part of the study addressed the possibility of using CD's as storage for portions of SPEL. These files would also be on the Internet and available to the users of Spacelink.

CD technology was not designed with the Internet in mind. The original concept for CD Technology was to provide high volume, read only storage, for stand alone, high powered computer systems. This provides the user with data, processed images, sound, user interaction features or any combination of these features. Many CD-ROM's perform very well in this environment. For CD's to work, in this manner, on the Internet, a high speed connection (T1 or faster) is required. Many of the teachers served by Spacelink are on much slower connections. Many schools are slow to place phone lines in the classrooms. T1 connections for a classroom are an even larger issue and much less likely to be installed.

Current CD-ROM's present several issues related to their potential use, by the nature of the original design criteria of CD-ROM's. CD's also vary considerably by when and where they were produced. These issues fall into the following three broad categories.

Multi-Media CD-ROM's are the source of the current excitement in the public market. Multi-Media CD's feature the full array of data, graphics, sound, and interactive capabilities. These CD's come with descriptions of the hardware and software needed to run the disc. In most cases the CD-ROM is customized to a specified hardware and software platform. Different versions of each CD are produced for Macs, PCs, and UNIX based machines. These CD's are copyrighted and therefore can not be placed on Spacelink. The only known NASA CD of this type, "Welcome to the Planets," is already on the Internet.

Text and/or image CD-ROM's are more like the "free world atlas, dictionary and encyclopedia" provided with new computer systems. This is a greatly downgraded version of the multi-media CD-ROM's described above. The visual images are often of less quality. The interactive capabilities and sound, where they exist, are seriously limited. This group of CD's carries all the limitations related to the operating platform discussed above.

Some few NASA CD's are text and/or image CD's and come with text, images, and the "necessary executable software" to use the CD's. The visual quality of these CD's is excellent. The difficulty with these CD's is that the software provided is usually not executable. To solve this problem, source code of the software must be found and loaded onto the computer to be used in the process. This software must be recompiled, debugged and/or modified to correct the limiting conditions. In most cases this means doing this for each CD-ROM because no standard systems were used to develop the CD's or the software. This must be done for each platform to

be supported. If different versions of the operating system, or different hardware, from the "playing" system, were used to produce the CD, the discs may be unusable.

Compressed Raw Data CD's are raw data as captured from the acquisition platform. NASA has produces hundreds of CD's in this format. NASA CD's of this type are the data as received from Voyager, Mariner or other space crafts. Most are unprocessed data and are stored in compressed data format. On the CD is source code for the software needed to decompress and store the images on a hard drive in JPEG, gif, Raster, or other formats. This software, usually written in C or FORTRAN, must be compiled, link edited, and stored in executable form before beginning the processing of the images on the CD. The software in these CD's is not standardized. The success of the compilation, link editing, and development of an executable module depends, on the compatibility of the platforms of the original host and the user system. Software is also different for each of the NASA missions. Imaging viewing software is usually not included on the CD but must reside on the platform where the image is to be viewed.

The NASA CD's in this group are stored in compressed data format. Each CD contains approximately 2400 images (files). When decompressed, a file will often be 3+ times larger than in compressed mode. Many of the expanded files average 800k and range up to 9 meg of needed hard disc storage per image to allow for viewing of an image. That is a very large file to be transmitted over the Internet, especially if a slow link is in the path of the transmission. The receiving machine must have real memory large enough to capture the image.

These three types of CD's left several options related to placing CD's on Spacelink. One option was to place one CD on each of the ten drives now located on the SunSparc 1000 and leave then for some specified period. At the end of this period, rotate all or part of the ten CD's and replace them with new selections. The cost of additional equipment was minimal. Some disc storage space would be required. A major cost would be in programming time to compile, debug and modify the software to make the images available to Spacelink users. The ten drives would make only a fraction of the 500 plus CD's already produced by NASA available to users. Each time the CD's were changed the SPEL menus would have to be modified to reflect the changes. The cost of staff time and management problems of this option made it undesirable.

The Jet Propulsion Laboratory has placed some NASA CD's in a Jukebox for access over the Internet. Another option was to develop and place information about NASA CD's on the menus of Spacelink and point to the location providing the CD services. This eliminates the need to place the physical CD's of Spacelink at MSFC. Depending on the detail provided on the Spacelink menus the cost of this option was small. This option has several disadvantages. The indexes and menus on the host systems at other NASA sites are totally foreign to our teacher-student clientele. The JPL host system, where most of these CD's reside, consists of 7 nodes at this time. Access to complete services at JPL is restricted to NASA and NASA funded scientists. The help desk at JPL is staffed only part time. The service at the other nodes seems

process to be completed. This would tie up lines on Spacelink. This option was eliminated because it seemed to do little or nothing for our primary audience.

A third option was to purchase a large CD Jukebox or a series of smaller Jukeboxes, and integrate the NASA and other Educational CD's into SPEL. This would provide control of our services and support for the teachers and students. Changes and additions to the SPEL menus would have to be made only one time. This had several disadvantages. This option would require additional hardware and possibly some Hierarchical File Management (HFM) software. Integration of the files into the Spacelink menus would require up to a full person year. Backup procedures would have to be developed to prevent backing up the CD's on the readers. As the CD collection grows additional Jukeboxes would be needed to support this option. The approximate $50,000 front end costs of hardware and software and continued staff cost made this option undesirable.

Another option was to combine the ten readers we now have with a 500 CD jukebox. Place the most popular CD's in the ten readers we now have and the remainder of the collection on the Jukebox. This makes full use of the current system hardware. This makes access to the most popular CD's virtually as quick as the files on the hard drives. This option provides a logical growth path as the collection of CD's grows. The cost of this option would be considerable in the new hardware needed. Some method of identifying the most popular CD's would have to be developed and monitored. This option brought with it the costs and limitations of the two options it combined. This option was also undesirable.

With the above options being undesirable or unrealistic, the issue of loading parts of SPEL onto CD's produced here at Marshall was examined. The technology of the CD-ROM has progressed to the point that it is possible to archive data files on to CD-ROM's, from a working database system, and have these files on the CD's continue to be an active part of the information system. This process involves the hardware of a Young Minds, Inc., CD Studio Intelligent Controller, and a Kodak CD recorder. The process is driven by MakeDisc Premastering Software using the ISO 9660/Rock Ridge processing format. (1)

The process discussed here was developed by Young Minds, Inc. of Redlands, California. This application does not require a separate workstation and can be directly linked to the SPARCServer 1000 now running Spacelink. The complete package of the MakeDisc Premastering Software; Young Minds, CD Studio Intelligent Controller; and a Kodak PCD Writer 225; has a federal price of $14,329.80. This is also available on other hardware platforms.

ISO 9660 is the file system standard that has made CD-ROM technology so widely accepted and used. The ISO 9660/Rock Ridge is the most open of the file creation and access systems related to the ISO standards. UNIX, as the most open of the current operating systems, is best suited to take advantage of the ISO 9660/Rock Ridge standard for open file structures. The ISO 9660/Rock Ridge standard allows the implementation of the UNIX/POSIX (the file structure of SPEL) file structure as the file index and structure of the CD. This is designed to be directly accessible as a physical and logical extension of the UNIX/POSIX file structure of the on-line hard drives. (2)

The software to drive this process is MakeDisc Premastering Software from Young Minds, Inc. To develop a CD someone collects and identifies those files that are to be placed on the CD. This includes the necessary information for identification and technical definition of the needed details to identify the CD, what is stored on the CD, and how the files are stored on the platter.

After the files to be placed on the CD and necessary identification and processing information are collected, the MakeDisc software becomes a "fill in the blanks," and a "point and click" process. The GUI and instructions are designed to allow for unsophisticated file managers to develop the necessary information and expertise to complete this process with relative ease. The software stores of the files, in CD format, on the disc, in the Young Minds Intelligent Controller. This allows for a preview of the CD before the disc is burned. This disc is fully editable and can be fully tested on the Young Minds CD Studio before the gold CD is burned.

The completed Studio work is transmitted to a Kodak PCD Writer 225 for processing. This is a dual speed "burner" and produces CD's at the speed needed for the CD drives now attached to the Sun SPARCServer 1000. This CD burning device does not require certified media and non-certified media are some less costly.

This newly produced CD is then ready to be placed on the CD drivers on the Spacelink system. At this time, with the deletion of the files now on the CD, form the Hard Drives, this makes the materials on the CD directly accessible to the users of Spacelink.

Updates to the SPEL that come after this CD is produced are added to this data set at the time of the production of the next CD. New Materials are integrated into the "new" CD, at the proper location, with the materials, on the current CD. The materials on the Current CD are saved on disc, or on tape, at the time of production of the "original CD" for future updates. The process described above of Premastering, testing, and burning are repeated and the new CD replaces the previous volume. This allows for dynamic growth of the files and directories and the addition of items that were not available when the previous CD volume was produced.

Several technical issues remain to be examined before transferring Spacelink files to CD's. Backup procedures remain to be tested. Documentation indicates that the backup tapes on the Spacelink System can be set to backup only the hard disc. This is the desirable option as the CD's would soon hold a very large volume of data. The nature of the CD is such that little or no need would exist for a backup. A simple inexpensive backup would be to burn a duplicate disc at the time of each production. This process would take only 15 minutes and cost approximately ten dollars. This backup CD could then be stored, off site with the backup tapes, as part of a good disaster recovery plan.

The directory interface and links with SPEL remain to be tested. Young Minds will supply references in the field using the system. Questions need to be formulated to take advantage of these contacts. The technical staff at Young Minds indicated that they supply systems to be tested and evaluated before purchase.

Speed of access and transfer of data are related technical issues yet to be fully tested. The directory search seems to be a bit slower. In testing this issue with CD's on line at JPL, the difference, between CD access delays, and the normal Internet delays for host contact and response, seems to be little if any. The transfer rate of a duel speed CD is approximately 300k bytes per second. Any one who has any network delays will see this as a good transfer rate. This would allow for the transmission of a good size gif much faster than many of our uses can receive, and store the data

The frequencies of production of "new" CD's and use of the equipment needs to be addressed. This is directly related to what data is be to placed on this media. A plan for what will trigger the production of the "next" CD needs to be developed. Is the production of the next CD time driven, volume driven, or ad hoc? A careful plan needs to be formulated before the testing of software and hardware and the questionings of references.

Some cost issues also remain. The price of this proposed hardware and software package is $14,329.80. There would also be some additional costs for cabling and racks. The human cost of the initial set up of the system and learning to use the software and equipment must also be considered.

Outside production of the Archived CD's is an option. If the decision is made to use CD's on Spacelink, and to have them produced at a private company, like Disc Manufacturing here in Huntsville, each gold CD with a ten to fifteen day turnaround time would cost $350.00+. The collection of the data to be placed on the CD and learning the technology well enough to do this would remain the responsibility of the Spacelink Staff. At this price, it would take little time and production of only a few CDs, to recover the cost of the proposed system.

An additional option to offset cost is to develop some slick CD's (The Best of Spacelink) for sale at the Space Museum or souvenir stores at the NASA Centers. With the gold master, the cost of a thousand CD order would be less than $2.50 per CD. If a winner was produced, and more CD's were needed, an order of 5,000 CD copies would cost approximately $1.50 per CD. CD's sell for about $20 each. Profits of 12 to 15 dollars each would result in income of 60 to 75 thousand dollars profit on the sale on an order of 5000 CD's.

An important additional consideration is the relative cost of Hard Disc to CD's Currently an 8 gigabyte Hard Drive costs approximately $5000. With the $14,000 plus cost of the identified system, 24 gigabytes of Hard Disc could be added to the Spacelink System.

This analysis did not identify an optimal solution to the use of CD's on Spacelink. Several options were analyzed that do not seem to be desirable at this time. Further study is needed to determine if CD's are a desirable addition to the Spacelink System.

---

1. Young Minds, Inc. "Documents and Sales Materials," Redlands, California, July 1995
2. Clayton Summers "Introduction to ISO 9660," Wilmington , Delaware; Disc Manufacturing. Inc., 1993.