1.0 Background of the IRIS Project

In the fall of 1989, the University of Houston Libraries were awarded a $99,852 Research and Demonstration Grant from the U. S. Department of Education's College Library Technology and Cooperation Grants Program to develop an Intelligent Reference Information System (IRIS) over a two-year period. 

It is estimated that Federal funds will pay for approximately 51% of the IRIS Project's costs.

The primary goals of IRIS Project are to: (1) network diverse CD ROM database resources in the Information Services Department of the University of Houston Libraries; and (2) create an expert system that will identify appropriate reference materials to meet users' needs, including networked CD-ROM's, stand-alone CD-ROM's, and print reference works. Users of the CD-ROM network will be able to consult the expert system, then connect to CD-ROM databases that are recommended by the expert system.

The IRIS Project hopes to provide some useful data about networking CD-ROM databases, test the success of such an endeavor from the user's perspective, study the effectiveness of using an expert system in such an environment, identify intellectual and technical issues surrounding the use of expert systems, and enhance the understanding of the profession in these areas. The work of the IRIS Project is conducted by the Project Director and eleven librarians, serving on four committees: the Electronic Publications Instruction Group, the Knowledge Engineering Group, the Project Management Group, and the Research and Evaluation Group.

Prior to the project, the University of Houston Libraries had accomplished pioneering work in the area of automated reference assistance by developing the Information Machine and the Index Expert programs. The Information Machine provides general directional and library use information through a menu-driven interface. Index Expert is an expert system designed to assist users in selecting appropriate indexes and abstracts based on a subject approach. Both systems have been successfully implemented in the University of Houston Libraries. A complete discussion of the IRIS Project is beyond the scope of this paper, which will primarily focus on technical aspects of the CD-ROM network. The reader is encouraged to consult the cited references for further information about the IRIS Project.

2.0 History of CD-ROM Use

Prior to the IRIS Project, the University of Houston Libraries had significant expertise in implementing CD-ROM technology. The Information Services Department began exploring CD-ROM databases in 1986 and served as a beta-test site for Compact Disclosure. During the summer of 1987, the University of Houston Libraries implemented three CD-ROM workstations consisting of one IBM-PC/XT and two OCLC M300's. These machines were equipped with 640 KB RAM (Random Access Memory), 20 MB hard disk drives, and CGA color monitors. Each workstation was dedicated to a particular product (one for Compact Disclosure and two for Ondisc ERIC). These two databases were immensely popular with the University of Houston Libraries' clientele. Shortly after implementation, however, it became evident that a more diverse set of databases would be necessary to meet client demand.
The following year the University of Houston Libraries expanded the collection of databases to include: ABI/INFORM Ondisc, Compact Disclosure, CD-Plus Medline, MLA International Bibliography, GPO/CAT PAC (Monthly Catalog), Ondisc ERIC, PAIS on CD-ROM, PsycLIT, Science Citation Index Compact Disk Edition, sociofile, and SUPERMAP. The number of workstations was also expanded from three to nine. Since there were more databases than workstations, the configuration had to allow more than one database to run on each workstation. The three dedicated workstations were retained because of heavy use, but the six remaining workstations were configured to run any of the other CD-ROM databases. The University of Houston Libraries named this service point the "Electronic Publications Center." Quickly the use of this service doubled, then tripled. Just prior to the installation of the network, the annual cumulative use total was eight times as much as the first year. Lines and queues were common occurrences in the Electronic Publications Center. It was evident that somehow the CD-ROM resources needed to be made more widely available.

3.0 Network Requirements

In preparing to network CD-ROM resources, the University of Houston Libraries explored a variety of local area network options. It would be necessary to create an environment that could support a great diversity of data resources. Part of the project was to test access to a wide-array of databases, including citation, full-text, graphic, and numeric databases. To accomplish this end, the network environment needed to be robust. The University of Houston Libraries had installed an IBM Token Ring network running the IBM PC LAN software for shared administrative and staff functions. The Libraries were very pleased with the reliability and performance of the IBM Token Ring network; however, a more sophisticated networking operating system would be required to support IRIS.

Security was another concern. Some clients in the Electronic Publications Center had proved that, given the opportunity, they would use the workstations for whatever purpose they chose, despite reasonably restrictive security measures. Thus, to avoid an increase of these problems in a network setting, the network operating system needed to have good security features. Given that the exact nature of the behavior of CD-ROM databases in a network was an unknown variable, the network needed to be flexible and expandable. It was also a distinct possibility that the network would be used for other types of data resources other than CD-ROM, and that over time the network would be greatly enlarged throughout the building, to the branches, and across the campus.

To the degree possible, centralizing the searching software on a file server was also desirable. This approach would make updating and maintaining the software for each database much easier. In addition, other network management software, such as menu programs, statistical packages, and utilities could be more easily used and maintained if they were available from a central file server.

4.0 Description of the IRIS Network

The University of Houston Libraries utilize an IBM Token-Ring network for the IRIS Project. The IBM Token-Ring network was chosen in part because of previous experience with it and in part because of a belief that it would perform well under moderate and heavy loads. The University of Houston Libraries' network uses the IBM Cabling System, with major cables being Type 1 cable. From a physical perspective, the IBM Token-Ring network is a combination of ring and star topologies. Cable runs from Multistation Access Unit (MAU) to Multistation Access Unit to form a ring. From each MAU, cable emanates in a star pattern to individual workstations. This arrangement permits malfunctioning workstations to be removed from the network without bringing the entire network down, and it enhances troubleshooting. Logically, the IBM Token-Ring functions as a ring. For example, on a very simple network with two workstations, workstation A receives a "token," changes the token to a "frame," adds control information and data to it, and sends the frame to workstation B; workstation B processes the frame and sends the frame back to workstation A; workstation A changes the frame back to a token and sends it to workstation B. The token continuously
moves in a circular fashion around the network until one of the workstations wants to send data.

The IRIS Project is running Novell NetWare 2.15 revision B as a network operating system. NetWare provides high levels of reliability, stability, security, and expandability. It is considered a major player in the computing industry and is widely supported. NETBIOS is used on the network servers and workstations. NETBIOS provides additional network communication services that permit multiple CD-ROM servers to operate on the network.

The file server is a Club American Model 320, 20 MHz, 80386 microcomputer with an ESDI 150 MB hard disk drive. The file server and each workstation has a Western Digital TokenCard network interface card.

Two Meridian Data CD Net Model 314 CD-ROM servers are used to provide access to 19 CD-ROM databases. Each 20 MHz 80386 server houses 10 CD-ROM drives (up to 14 CD-ROM drives can be housed in each unit). The Meridian CD Net software (3.01) as well as the NetWare DOS Client software (3.01) are running on the Meridian servers.

Originally, the workstations were a mix of 80286 and 80386SX machines, but due to a need for large-body machines elsewhere in the University of Houston Libraries, the 80286 machines were replaced. All workstations are now Club American Model 316SX, 16 MHz, 80386SX microcomputers with 40 MB hard disk drives and EGA color monitors. The performance of these workstations has been remarkably good.

Eight public network workstations are located in the Electronic Publications Center, one staff workstation is located at the Information Desk, and the servers as well as another staff workstation are located in the University of Houston Libraries' central computer room. The computer room provides a climate controlled, secure location for the computing equipment. The following CD-ROM databases are currently available on the network: ABI/INFORM Ondisc, Art Index, Biological and Agricultural Index, Business Dateline Ondisc, Compact Disclosure, Compendex Plus, Computer Library, Social Science Index, The New Grolier Electronic Encyclopedia, Humanities Index, Microsoft Bookshelf, Ondisc ERIC, Periodical Abstracts Ondisc, PsycLIT, Social Sciences Index, sociofile, The Software Toolworks World Atlas, Statistical Masterfile, and SUPERMAP. Future plans include additional electronic publications, both on CD-ROM and on magnetic media.

5.0 Procurement Process

Acquiring network licenses for the 19 databases needed by the IRIS project was a complicated process that required about two months of negotiation with database vendors. At the time, many database vendors were uncertain about what contractual terms and license fees were appropriate for network use, and the contracts that resulted from these negotiations were not very uniform. A variety of restrictions on use were included in these contracts, including number of simultaneous users, number of network workstations, location of network workstations, and end user affiliation.

As a public institution in the State of Texas, the University of Houston Libraries were required to follow strict purchasing guidelines—a practice common to all publicly-funded agencies. The nature of these requirements determined the level of flexibility in selecting network hardware and software, not to mention workstation hardware and software. For most items, selection was limited to suppliers on the state contract; however, some of the items needed for networking could be purchased off contract.

One crucial element was the certified compatibility of the file server with the NetWare software. The most critical aspect of NetWare compatibility was the disk drive. Fortunately, the Club American vendor on state contract also handled NetWare and could guarantee the compatibility of the equipment.

Another complexity of dealing with state contract purchase was the length of time between the submission of an order and the arrival of the items. Delivery times for some needed
items were sixty days or more.

The process was also slowed by the level of review necessary for a grant project. Additional paperwork was required, and the University's Office of Sponsored Programs had to review purchases to ensure that the University of Houston Libraries were following Federal government guidelines.

The University of Houston Libraries had been installing Copicard units on all photocopiers and CD-ROM workstations to assist in defraying the cost of printing supplies. Such a commitment, however, demanded that each new equipment acquisition be compatible with the Copicard units. Arrangements were made for a trial of the Club American hardware with the Copicard hardware and software. Both companies were willing to communicate with each other and were helpful in satisfying our request.

Table 1 shows estimated network costs.

Table 1. Estimated Network Costs

<table>
<thead>
<tr>
<th>Item</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD-ROM Servers</td>
<td>$34,890</td>
</tr>
<tr>
<td>NetWare Server</td>
<td>$7,700</td>
</tr>
<tr>
<td>Network Workstations*</td>
<td>$20,866</td>
</tr>
<tr>
<td>MAU's</td>
<td>$2,340</td>
</tr>
<tr>
<td>Cable and Installation</td>
<td>$3,118</td>
</tr>
<tr>
<td>CD-ROM Subscriptions</td>
<td>$48,000</td>
</tr>
<tr>
<td>Total</td>
<td>$116,914</td>
</tr>
</tbody>
</table>

* The cost of purchasing five new workstations and upgrading five existing workstations is shown here. The existing five workstations were later replaced.

6.0 Network Implementation

Implementation of the IRIS network proceeded in two broad phases: (1) testing a small-scale prototype network, and (2) testing the full-scale network.

6.1 Small-Scale Network Test

Given the size of the IRIS Project, it seemed wise to begin by installing a test network in a staff area. This arrangement permitted ample working room, a quiet work space, and physical security. The Information Services department had available space at the time. Having the test network in this location was also convenient for demonstration purposes.

The test network consisted of the NetWare server, two Meridian Data CD Net 314 CD-ROM servers, two workstations, and one MAU. Cables were conveniently placed behind the equipment to allow for easy connecting and disconnecting, which occurred quite frequently in the test mode.

6.1.1 NetWare

Installing NetWare was a somewhat daunting task. Portions of the process were lengthy and complex. Fortunately, Novell provided extensive documentation; sometimes it seemed like too much documentation. Several NetWare books proved to be useful during the installation process.

One of the first challenges was the preparation of the disk drive for NetWare. A test, called COMPSURF, needed to be run to check the surface of the platters for any irregularities and to prepare the drive to receive data. This test could take 12 to 18 hours or more to complete and could be run unattended. The better part of wisdom suggested that the server be set up for the test prior to closing the department for the day to allow the
computer to do what it does best on its own. In the morning, messages indicated that the process had been aborted. The next night this was done again with the same result.

Investigation revealed that the University's custodial staff were turning off the power to the Information Services department over night. This problem was resolved by taping official-looking messages to the door of the circuit box controlling this area of the building. Fortunately, the custodial staff obliged our request.

NetWare used a program called NETGEN to place the appropriate files on the file server and to configure it according to the implementer's wishes. There were many options. Fortunately, there were also defaults, but these were often not well documented. Nonetheless, we often relied on default settings as a starting point.

NetWare used another program called SHGEN to create the "shell" for each workstation on the network. This process created the IPX program and it was considerably less tedious than NETGEN, as one would expect.

In addition, installing more than one Meridian server on a network running NetWare required the use of the NETBIOS program or an emulator, which provided another level of network services. Novell provided a NETBIOS emulator with their software. This program needed to be copied to each workstation.

When running the IPX network shell only, the CD-ROM drives in the Meridian servers were mapped directly to drive letters on the workstations. These drive assignments could not be changed without editing files and rebooting the workstations. Under the NETBIOS emulation, by using Meridian's MOUNT or CMOUNT mapping utilities, the workstation's logical drives could be dynamically allocated to any of the CD-ROM drives.

The NetWare implementation was a time-consuming process because there was a substantial learning curve involved. NetWare is a complex network environment that should not be attempted by the technically faint-at-heart.

Once the initial hurdles were crossed, success was evident at the network level. The small test network appeared to be communicating appropriately.

6.1.2 Meridian Data CD Net 314 Servers

Once the NetWare file server and the workstations were configured, the Meridian CD-ROM servers had to be installed. For the most part, these servers came ready to plug and play. Some software adjustments were necessary, such as naming each server and generating the appropriate shell for the network. The Meridian servers connected to the network just as any other workstations would. They were not identified by NetWare as servers. It was the Meridian software running on the CD-ROM servers and on each workstation that permitted mapping to specific CD-ROM drives and multiuser access to these resources. The Meridian software and CD-ROM server hardware were optimized to provide rapid multiuser access to CD-ROM databases on the network. Workstations communicated with the Meridian servers using NetWare and NETBIOS network protocols.

In addition to software adjustments on the Meridian servers, each workstation required the presence of certain programs and environment variables. Each workstation needed to run the Microsoft CD-ROM Extensions to identify what type of CD-ROM device would be used, how many of them could be mapped at once, and what size memory buffer was required. Each of the Microsoft CD-ROM Extensions settings had memory (RAM) implications. There were other variables to set in the DOS environment, such as the name and number of Meridian servers. These adjustments were fairly well-documented in the Meridian manual, and little trouble was encountered in this arena.

6.1.3 CD-ROM Databases
Since the University of Houston Libraries had purchased a number of CD-ROM databases from H. W. Wilson and UMI, these vendors' CD ROM searching software packages were installed first on the network. Presumably, once the software was installed, all of that vendors' products would run. This approach had psychological as well as pragmatic advantages.

Another guiding principle was to install as much of the software as possible on the file server's hard disk. This arrangement would simplify maintenance and updating of the network. Several vendor products, however, did not run reliably--or at all--in this configuration and had to be installed on individual workstations instead.

The initial installation of CD-ROM products went remarkably well. Two people being able to search the same database at the same time or to search multiple databases from one workstation was an exciting prospect. Many of the staff who had followed the installation process closely were quite excited as the system unfolded. However, as was discovered later, two workstations on a small test network are quite different from ten workstations on a fully implemented network.

The CD-ROM installation process was not without its challenges. Incomplete or inaccurate documentation for network installations caused many frustrating moments. For nearly all of the products, calls had to be made to the vendors to request additional information. Since many phone calls to a number of vendors were necessary, an accurate impression of the technical support capabilities for each vendor was attained. The quality, helpfulness, and attitudes of the technical support staff varied greatly. In several cases, vendor representatives gave prompt, accurate, and courteous assistance. Others were argumentative, ill-prepared, and non-committal. Clearly, the issue of providing adequate network support will need to be addressed by vendors in the future.

Another challenge faced in this process was caused by misbehaving or poorly designed software. Although many software sins can be hidden in stand-alone mode where only one product is running on a workstation, the network environment tended to accentuate program weaknesses. Theoretically, software that was written for a single-user environment should also run in a multiuser environment (unless the application entails transaction processing, which CD-ROM does not); however, in reality, problems arose. Some products performed poorly, required the constant attention of the network manager, or even ground the network to a halt. The nature of these problem ran the gamut from hard-coded sub-directory assignments to circumventing the Microsoft CD-ROM Extensions or other standard conventions.

Despite the challenges faced, the IRIS network had sixteen databases up and running after two weeks of intensive effort. At least, these products ran most of the time on the small test network. Minor glitches remained, such as not being able to dependably view the results of a search on one vendor's products and an occasional freeze up with another vendor's products.

6.2 Full-Scale Network Test

After the successful implementation of the test network, the time came to move the equipment to its final resting place and to connect the actual public workstations. The three servers (NetWare server and two Meridian servers) were moved to the computer room that also houses the integrated library system equipment. Eight public workstations in the Electronic Publications Center were connected to the network, the cabling and MAU's having been previously installed. One workstation remained in Information Services, intended for network administration. The process of installing network boards in the workstations and connecting them to the network went smoothly. Once all of the hardware was connected, it was time to test the network. The network booted the first time; however, this success was fleeting. At the observable network level, connectivity had been achieved; however, once the applications (i.e., the CD-ROM databases) were tested, new problems arose. The CD-ROM databases would work if one or two users were searching in the same product at the same time. But higher numbers of users froze the workstations and the
network. Even when products did work, the performance was highly variable. Some products were more problematic than others.

**6.2.1 Diagnosing the Problem**

Diagnosing network troubles is akin to determining the desires of a new-born infant. Even with previous experience, the most one can hope for is a reasonably accurate guess. Starting from the simplest component and moving to increasingly more complex ones was the approach taken.

**6.2.1.1 The Network and NetWare**

Through the process of elimination, components were tested. One at a time cables, connectors, and MAU's were removed from the ring and the network was rebooted. If the same problem persisted, the component was reconnected and another removed. Several potential weak points were identified via this method, but none that yielded significant improvement in network performance.

The cable run to the Information Services department was from a previous LAN connection installed several years before. The caliber and condition of the cable were suspect, but removing it from the network did not resolve the entire problem. Nonetheless, a significant improvement in network performance occurred when the cable was disconnected. The network management workstation was then moved from Information Services to the computer room.

A cable tester was acquired. The suspect cable proved to be problematic, but other cables were fine. More sophisticated network diagnostic equipment was available, but it was expensive ($1,500 to $25,000) and topology-, NIC-, and operating system specific. Consequently, the acquisition of this type of costly diagnostic equipment was seen as a last resort. The network level of communication could also be tested by utilities provided with some network operating systems. NetWare came with COMCHECK, a program that ran on each workstation to verify that each is sending and receiving messages on the network. Once the cabling difficulties had been resolved, this program indicated that the network itself was not the cause of continuing problems.

**6.2.1.2 The Meridian CD-ROM Servers**

Throughout the implementation process, representatives from Meridian were in regular communication with the University of Houston Libraries. It became evident after all of the testing that there must be a problem related to the Meridian servers. The hardware appeared to be functional, but something was not working appropriately. It seemed likely that the Meridian software was not functioning properly in the NETBIOS-oriented Token-Ring network environment.

One problem was that the University of Houston Libraries had purchased "plug-compatible" Token-Ring boards on state contract, and, at that time, Meridian did not support these boards. Consequently, Meridian staff needed to become familiar with these boards before potential software compatibility problems could be addressed.

Meridian representatives were extremely helpful in resolving the server problems. The University of Houston Libraries had access to Meridian's technical support, software development, and engineering staff. Meridian sent faster, more reliable NIC's for the servers, and it provided the University of Houston Libraries with two new versions of its CD-ROM server software as well as the NetWare DOS Client software. Once these changes were made, performance improved dramatically and the network was highly reliable.

**6.3 Additional Implementation Challenges**
While the primary network problems were being addressed, other technical issues needed to be resolved as well. During the course of testing, a number of product errors had been linked to the lack of sufficient memory on each workstation, a common problem in network environments.

Each workstation had device drivers, Microsoft CD-ROM Extensions, network drivers, and DOS environment variables loaded in conventional memory below the 640 KB limit, leaving a fairly small amount of useable RAM. Many CD-ROM products required additional RAM in order to function properly. Fortunately, each workstation was equipped with one megabyte of RAM. With appropriate memory management software, certain portions of the these memory resident programs could be moved into high memory (the memory between 640 KB and 1 MB).

The AboveLAN program, which was chosen because it is a software only solution, performed these functions easily. This program ran on each workstation and freed up approximately 37 KB of RAM below 640 KB.

Memory management continues to be a challenge for the IRIS Project. Additional solutions are being investigated to yield even greater RAM savings.

Another issue requiring attention was the creation of menus to guide users to available databases, to provide easy access, and to minimize security risks. NetWare came with a built-in menu program, but it was limited to fewer menu entries than the IRIS application needed and users could easily exit from the menus into NetWare. Some CD-ROM searching software would not run on the file server unless very liberal directory rights were granted to network users, creating a security risk. Thus, a different menu program option was required.

To meet this need, the Saber LAN Administration Pack was acquired. This package also included a statistics gathering program, a report generator, a metering program to ensure licensing compliance for network products, a program that prevents users from exiting the menu system, and other network utilities. The menu software permitted relatively easy modification of menu entries. Menu entries could be password protected, assigned to particular workstations, and nested. The Saber software ran on the file server, making maintenance and updating manageable.

6.4 Total Network Implementation Time

Implementation began May 21, 1990, and the network was made publicly available August 20, 1990. During that time, the equivalent of about four weeks of intensive effort was invested in getting the network up and running. The remaining time can best be described as "hurry up and wait." Obviously, other library-related tasks had to continue in the interim; no new staff were hired as a result of this project. Furthermore, technical challenges were frequently resolved at odd hours after removing oneself from the firing range. Thus, the entire period was not spent directly resolving network problems.

7.0 Network Evaluation

Based on the previous success of CD-ROM products, it was anticipated that the IRIS LAN would be a popular service; however, actual use quickly exceeded our expectations. Table 2 shows use statistics for the first full six months of operation. Each use represents one CD-ROM session; multiple searches may have been conducted during that session.

Table 2. IRIS LAN Use Statistics 9/1/90-2/28/91

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ABI/Inform Ondisc</td>
<td>1,249</td>
<td>1,561</td>
<td>1,369</td>
<td>380</td>
<td>542</td>
<td>1,205</td>
</tr>
<tr>
<td>Art Index</td>
<td>59</td>
<td>90</td>
<td>85</td>
<td>31</td>
<td>22</td>
<td>63</td>
</tr>
<tr>
<td>Biological and Agricultural Index</td>
<td>84</td>
<td>150</td>
<td>123</td>
<td>87</td>
<td>69</td>
<td>144</td>
</tr>
<tr>
<td>Business Dateline Ondisc</td>
<td>602</td>
<td>737</td>
<td>737</td>
<td>201</td>
<td>269</td>
<td>619</td>
</tr>
</tbody>
</table>
The IRIS Project is currently analyzing performance benchmarks for the Meridian CD-ROM network. In general, performance is quite good, even when the network is heavily loaded. The network has been rock solid since the service became public, and there has been very little network downtime. Routine network maintenance (e.g., putting CD-ROM update disks on the network) takes a maximum of two hours a week, and little other work is required to keep the network operational. Overall, the IRIS Project has been very satisfied with the network, and it would not hesitate to purchase additional hardware and software from our current vendors to support expanded CD-ROM network efforts.

8.0 Conclusion

CD-ROM networking is a relatively new technology. Unless libraries purchase true turn-key CD-ROM networks that include network installation and maintenance services, they should expect that a fairly high level of technical expertise will be required to support these networks. As the size and complexity of these networks increases, technical support needs will also increase. This paper has provided a candid "behind the scenes" look at some of the technical issues that the University of Houston Libraries faced in its CD-ROM network implementation. Some of the challenges described here were the result of being an early user of CD-ROM networking technology, local hardware compatibility and cabling problems, and the overall scope of the CD-ROM networking effort. Some challenges were caused by the fact that, for the most part, CD-ROM vendors have not modified their single-user products to run smoothly in networked environments. Some were simply the result of the inherent complexity of CD-ROM networks. CD-ROM networking offers libraries a powerful and effective tool for providing increased access to electronic information. For libraries with adequate fiscal resources, a desirable information provision strategy may be to put high-use databases on mainframe computers, medium-use databases on CD-ROM networks, and low-use databases on stand-alone CD-ROM workstations. Of course, some databases will only be available in CD-ROM format, and the demand for high- or medium-use databases in this category is likely to be adequately met only through networking.

References


6. Thomas C. Wilson, "Zen and the Art of CD-ROM Network License Negotiation," *The Public-Access Computer Systems Review* 1, no. 2 (1990): 4-14. (To obtain this article, BITNET users can send the command GET WILSON PRV1N2 to LISTSERV@UHUPVM1.)

**Citation**
