ORNL/TM-2001/25

PROBE PROJECT STATUS AND ACCOMPLISHMENTS

February 1, 2001

Prepared by
Randall D. Burris
Group Leader

ORNL/TM-2001/25

PROBE PROJECT STATUS AND ACCOMPLISHMENTS

R. D. Burris
M. K. Gleicher
H. H. Holmes
N. L. Meyer
D. L. Million

February 1, 2001

Prepared by
OAK RIDGE NATIONAL LABORATORY
P.O. Box 2008
Oak Ridge, Tennessee, 37831-6285
Managed by
UT-Battelle, LLC
for the
U.S. DEPARTMENT OF ENERGY
Under contract DE-AC05-00OR22725

CONTENTS

INTRODUCTION

The Probe project has completed its first full year of operation. In this document we will describe the status of the project as of December 31, 2000. We will describe the equipment configuration, then give brief descriptions of the various projects undertaken to date. We will mention first those projects performed for outside entities and then those performed for the benefit of one of the Probe sites. We will then describe projects that are under consideration, including some for which initial actions have been taken and others which are somewhat longer-term.

1. CURRENT CONFIGURATION

Appendices A and B detail the configurations of the Probe installations at the Oak Ridge National Laboratory (ORNL) and the National Energy Research Scientific Computing Center (NERSC).

2. PROJECTS AND ACTIVITIES

We have divided the projects into those performed at the request of external entities and those whose primary thrust was internal (specific to either or both sites or otherwise internally motivated). Each category is divided into completed and active projects. The third section describes projects that are under consideration or have been begun but are currently inactive for some reason.

2.1 WORK FOR EXTERNAL ENTITIES

2.1.1 Completed Projects

2.1.1.1 Lawrence Livermore National Laboratory (LLNL) S80 Test

The first project completed in ORNL's Probe installation was a study of the CPU performance of the IBM RS/6000 Model S80 Enterprise Server, performed at the urgent request of LLNL. LLNL had a tight deadline that required considerable effort to achieve, but the testing was concluded on time. Two letters of appreciation resulted. Summarizing the results, the S80 demonstrated excellent CPU performance and the ability to sustain very high single-channel and aggregate I/O throughput. See http://www.csm.ornl.gov/PROBE/S80.html for details.

2.1.1.2 Testing at the Request of the High Performance Storage System (HPSS) Collaboration

Several activities were performed to assist in the testing and support of HPSS. In one series, ORNL's StorageTek Redwood tape drives were used to test and validate HPSS version 4.2 (no Redwood drives are available in the Houston/HPSS testbed). In another activity, ORNL was designated a beta-test site for HPSS version 4.2 for the purposes of testing HPSS Solaris core servers and movers in a mixed IBM/Solaris environment. IBM staff performed the tests using Probe equipment. In tests of this nature, no earth-shaking conclusions are produced. In each case the testing was successful and the product is now in production. See http://www.csm.ornl.gov/PROBE/Pprojects.html.

2.1.1.3 TX/Database Management System (DBMS) Testing

One of the infrastructure elements of HPSS is Encina's Structured File System (SFS); it is the file system in which HPSS metadata are stored. "TX" refers to one of its approaches to transactional semantics. There has been interest for some time in evaluating database products as a possible replacement for SFS. ORNL did some preliminary work to investigate DB2 and Oracle database management systems and their ability to support the transaction processing required by HPSS. As the work proceeded, however, the HPSS project decided to eliminate the TX approach. The licenses for DB2 and Oracle will be retained and maintained to serve as a resource for prototyping in subsequent HPSS investigations (see "Bake-off" below).

2.1.1.4 University of Vermont

A researcher at the University of Vermont had a need for a data set of significant size for use in a data-mining project. Probe supplied a 5 GB ORNL global-climate dataset. See http://www.csm.ornl.gov/PROBE/bigdata.html.

2.1.1.5 Linear Tape Open (LTO) Beta Test

The PROBE testbed at NERSC has a beta test agreement for the new IBM 3584 tape library with LTO tape technology. The goal of this beta test was to assess the operation of the library and drives with AIX version 4.3.3, including performance and load tests for the LTO tape drives and the library. This project is described in a technical paper on the Web (in progress).

2.1.1.6 LTO-Associated HPSS Development

As a developer site for HPSS, NERSC also integrated the LTO system into HPSS, including developing a new LTO Physical Volume Repository and modifications to storage system management and the mover. That capability is scheduled for release in HPSS version 4.3, due out later in 2001.

2.1.1.7 Create a FibreChannel Disk Array Test Environment

Three FibreChannel disk array systems were evaluated in the NERSC Probe environment, including a DataDirect Networks EV-5000 Redundant Array of Independent Disk (RAID) array with a Brocade FibreChannel switch, a StorageTek 9176 RAID disk array, and a Sun T3 RAID disk array. NERSC's experience with installation, configuration, testing, problems encountered (including HBA/switch compatibility issues), and performance achieved are described in a technical memorandum. See http://hpcf.nersc.gov/storage/hpss/probe/fcdisk.html.

2.1.2 Active Projects

2.1.2.1 "Bake-off"

The HPSS collaboration is considering replacing the metadata engine, the Encina/SFS product, with another commercial database management system. In a new activity spawned as a result of the TX/DBMS project mentioned above, the performance of Oracle, DB2, and SFS in HPSS-relevant operations will be tested in Probe. ORNL staff and subcontracted staff have joined with HPSS developers

to perform the testing. Preliminary results indicate Oracle performance is comparable to DB2 performance and both improve upon SFS.

2.1.2.2 Beta Test of Gigabyte System Network (GSN) Hardware and Drivers for IBM RS/6000

Genroco has built network interface cards for several platforms (IBM, Compaq, and Sun) and operating-system software ("drivers") for each. With ORNL, they have been beta-testing the IBM hardware and drivers on the RS/6000 Model S80. At this writing a penultimate version has demonstrated a record transfer rate exceeding 150 megabytes/second between the S80 and an SGI Origin 2000. Testing will continue as Genroco revises the driver to address some minor problems.

2.1.2.3 Test of Gigabit Ethernet/FibreChannel Bridge

Genroco beta-tested a device that bridges FibreChannel disks to Gigabit Ethernet interface cards and uses the Scheduled Transfer (ST) protocol. It was tested at ORNL using Sun T300 disks and the Compaq DS20 server. In the testing, the server performed I/O to the FibreChannel disks through the server's Gigabit Ethernet interface (rather than through its FibreChannel interface) exhibiting transfer rates up to 54 megabytes/second. As products of this type mature, high-performance FibreChannel equipment will be available through less-expensive and more easily-managed Ethernet switches and interface cards. ORNL is also testing such transfers using another Genroco device, a GSN bridge with Gigabit Ethernet and FibreChannel blades. See http://www.csm.ornl.gov/PROBE/iSR.html for details.

2.1.2.4 GRID FTP Daemon

NERSC has installed the GRID FTP server on the Probe system so that security issues and GRID connectivity can be tested.

2.2 SITE-INITIATED PROJECTS

2.2.1 Completed Projects

2.2.1.1 Hierarchical Storage Interface (HSI)

HSI is an application providing a very friendly and powerful interface to HPSS (see http://www.csm.ornl.gov/PROBE/hsi.html and http://www.sdsc.edu/Storage/hsi). The author of HSI, Mike Gleicher, under contracts with ORNL and NERSC, has made extensive improvements to HSI.

- The HSI non-DCE HPSS client API library has been extended to provide the ability to communicate with multiple HPSS systems in a single session, and to freely switch between these sessions. HSI makes use of this capability, but it is not restricted to just HSI. The ability within HSI to treat multiple HPSS systems as logical "drives" is a simple but powerful concept that will make it easier for researchers to make use of resources at several sites without requiring cross-cell authentication.

- I/O performance has been improved. The IPI-3 project at NERSC funded the initial I/O rewrite in HSI, and PROBE-funded work has resulted in even greater performance improvements. The new "buffer pool" code results in fully double-buffered I/O (for both reads and writes) irrespective of the number of transfer threads, and decouples the HSI buffer size from the mover buffer size and the VV block size. PROBE also funded the work to make use of multiple network interfaces if they are available, and to make use of restricted TCP ports at sites with firewalls. These developments are now in production use at NERSC.

- Long-haul network performance has been improved. In addition to the continuing investigation of bottlenecks, PROBE funded the changes in HSI and the non-DCE server to use multiple concurrent sockets for inter-HPSS copies.

- The ability for HSI to communicate with different-release HPSS systems has been added. (In HPSS release 4.1.1.4 a network parameter was changed in a way that was incompatible with earlier releases.) HSI can determine at run time which level of HPSS the server system is running, and can determine other site-dependent features such as the available Classes of Service. If possible, this ability will be preserved for the HPSS 4.2 conversion, so that the ability to communicate with multiple HPSS systems will not require sites to move forward in lockstep to new versions. See http://www.csm.ornl.gov/PROBE/hsi.html.

- The new HSI has been put into production at ORNL, CalTech, University of Maryland, Indiana University, Maui High Performance Computing Center, LLNL, and the San Diego Supercomputer Center, and is being put into production at NERSC. Roughly 15 other sites use HSI, either for a primary user interface or for administrative functions. The improved HSI is much more powerful; it will have considerable impact on the community.

2.2.1.2 Improve ORNL-NERSC Bandwidth

The average bandwidth between ORNL and NERSC was seen to be approximately 250 KB/second, far below the peak of roughly 11 MB/second the hardware should allow. An initial project to find and remedy the cause has been completed. Increasing the buffer sizes at both ends has resulted in typical rates of roughly 4 MB/second with higher rates achieved until congestion limits are reached. For more information see http://www.csm.ornl.gov/PROBE/nerscband.html and http://hpcf.nersc.gov/storage/hpss/probe/bw.html.

2.2.1.3 Gigabit Ethernet, Jumbo Frames

When using Gigabit Ethernet interfaces, the standard maximum data packet size is approximately 1500 bytes. Each such packet requires an interrupt of the server's operating system for processing. At gigabit rates, these interrupts use a terrific amount of CPU time.

Jumbo frames use a 9000 byte data packet, with a consequent reduction in interrupt processing. In testing at ORNL using the S80 and H70 processors, transfer rates consistent with the full bandwidth (over 90 megabytes/second) were obtained with CPU loads of less than one processor on each server. The transfer rate could be maintained indefinitely. In one test, ten terabytes were transferred over a period of 32 hours for a sustained 93 megabytes per second. See http://www.csm.ornl.gov/PROBE/S80.html#giga for more information.

As a consequence of these observations, jumbo frames have been implemented in the production HPSS system and the IBM SP, AlphaServer SC, and Origin 2000 supercomputers at ORNL.

2.2.1.4 Two-Stripe T300-Gigabit Ethernet Jumbo-Frame Performance

ORNL has two Sun/MaxStrat T300 FibreChannel RAID arrays, two FibreChannel interfaces in the RS/6000 Model S80, and two Gigabit Ethernet interfaces in the S80 and in the RS/6000 Model H70. Thus there is sufficient hardware to study the throughput of a two-stripe HPSS file transfer between the two nodes. Using the HSI application, two stripes could be read at 115 megabytes/second and written at 88 megabytes/second. Corresponding rates with parallel FTP were 92 and 58 megabytes/second, respectively. See http://www.csm.ornl.gov/PROBE/2stripe.html for details.

2.2.1.5 Configuration of StorageTek 9840 FibreChannel Tape Drives.

There are several elements involved in supporting a FibreChannel tape drive, including hardware interfaces, the operating system, software "drivers" for FibreChannel itself, and software drivers for tape drives. There are multiple sources for the interfaces and the drivers. After extensive research and many trials, complicated by a dearth of documentation, ORNL successfully configured StorageTek 9840 drives and made them operational. An HPSS Operational Service Bulletin documenting the configuration process was provided at IBM/Houston's request. One request for assistance in configuring the drives has already been received and satisfied. The drives have been added to the ORNL production system. See http://www.csm.ornl.gov/PROBE/Pprojects.html for more information.

2.2.1.6 StorageTek SCSI 9840 Tape Drive Testing

ORNL tests of StorageTek SCSI 9840 tape drives were also performed. The drives were used with HPSS 4.1.1.1 over AIX 4.3.3 on the RS/6000 Model S80. Prior to these tests the drives had not been used on an S80 or with HPSS running over AIX 4.3.3. The goal of the test was to verify correct operation, and that was observed. These drives have also been added to the ORNL production system. See http://www.csm.ornl.gov/PROBE/9840.html.

2.2.1.7 Establishing an HPSS System on RS/6000 S80

At the time of procurement, the RS/6000 Model S80 computer required the AIX 4.3.3 version of the operating system; HPSS had not been tested on that AIX release. At ORNL HPSS was successfully compiled, installed, configured, and tested on the S80. The S80 has since been used to compile the latest releases of HPSS (HPSS 4.1.1.4 and then 4.2), to test FibreChannel disks and tapes, and to test HPSS movers linked with Gigabit Ethernet jumbo frames. ORNL's intention is to upgrade the S80's HPSS system immediately after each patch or major release. See also http://www.csm.ornl.gov/PROBE/aix433.html.

2.2.1.8 Establishing an HPSS system on AIX 4.3.3 on an IBM H70

At NERSC, a first installation of HPSS on AIX 4.3.3 on an IBM H70 was performed, including compilation, installation, configuration, and testing. This installation has subsequently been used to evaluate FibreChannel disk arrays and the IBM LTO library and tape system.

2.2.1.9 SFS Create/Delete Tests

Los Alamos National Laboratory (LANL) tested two RS/6000 servers with different memory architectures (switched vs. bus). They found that Encina/SFS, the repository for HPSS metadata, exhibited exceptionally good performance on the bus machine. The Model S80 at ORNL has the switched memory architecture, so we investigated the S80's performance on the same tasks. For metadata-only tasks the performance of the S80 compared well to the bus architecture results at LANL. The performance of the S80 for write-data operations was significantly better than was the case for the bus machine. See http://www.csm.ornl.gov/PROBE/sfs.html for more information.

2.2.1.10 StorageTek FibreChannel Equipment Installation

In preparation for testing of HPSS movers, ORNL purchased FibreChannel disk equipment (including disks, storage processors, FibreChannel switch, and host interfaces for IBM, Compaq, Sun, and SGI computers) from StorageTek. The installation and configuration processes were so challenging that the work became a project in itself. A description of the process, notes taken during the work and the final

documentation have been posted on the Web. See http://www.csm.ornl.gov/PROBE/fiber.html for lessons learned and configuration summaries.

2.2.2 Active Projects

2.2.2.1 High-Performance Visualization

One of the primary motivations for the creation of Probe is the investigation of high-bandwidth transfers from storage to visualization systems. The primary purpose of the GSN and ST activities is to develop and test a mechanism for such transfers. At this writing the GSN switch has been installed and connected to the Origin 2000 Reality Monster. A project to visualize the results of a simulation of a supernova explosion - using Probe servers, storage resources, and the GSN equipment - has begun. See http://www.csm.ornl.gov/PROBE/Pprojects.html for details.

2.2.2.2 High-Bandwidth Wide-Area-Network Bulk Data Transfers

When a network is busy, packets may be dropped. TCP/IP recovers from lost packets by retransmitting them, and it attempts to avoid making congestion worse by a "slow restart" algorithm. The result is reduced effective bandwidth, observed in the ORNL-NERSC case to be roughly 4 megabytes/second on a link that should be capable of over 10 megabytes/second. When the network is to be used to transmit extremely large files, as is required for ORNL's Global Climate and Human Genome projects, this level of performance is unsatisfactory.

ORNL/Probe has undertaken a research project to develop a restart mechanism that will recover much more rapidly (but without taking an unfair portion of the bandwidth) from dropped packets. The effect should be to raise average throughput nearly to the peak bandwidth assigned to the circuit. Subsequent to satisfactory testing, FTP and HSI will be modified to use the protocol.

2.2.2.3 HPSS Movers Using Gigabit Ethernet Network Connectivity and FibreChannel Disks

ORNL has purchased servers from IBM, Compaq, SGI, and Sun with the goal of testing/tuning HPSS mover operation using FibreChannel disks and Gigabit Ethernet network interfaces. The equipment is installed and operational.

The first round of testing used the "dd" utility and is complete on all four platforms with the disks configured as "RAID 5" (optimal for small-transfer workloads). Tests of HPSS mover software then were undertaken with the disks still in the RAID 5 configuration. We then discovered that mover software for the Compaq platform was not part of the normal HPSS distribution; we are now working with Compaq to obtain that software. Mover tests in the RAID 5 configuration on the remaining three platforms have been completed.

The next step of testing was to repeat the tests with the disks in the RAID 3 configuration (preferred for very large transfers). Tests using "dd" were performed on all four platforms and mover tests on three. See http://www.csm.ornl.gov/PROBE/Pprojects.html for current status and details.

2.2.2.4 Modeling Storage

The acquisition, storage and use of terabytes of data requires hundreds of pieces of equipment and very complex applications. Intuition is of limited value in establishing optimal and cost-effective configurations and procedures. ORNL has established a project to develop a model of the entire storage scenario, from acquisition through analysis, first modeling HPSS. Various data sources and analyses (high-energy physics experiments, for instance) could be added as additional projects.

At this time a network modeling tool, OPNET, has been purchased and installed. Discussions with various possible sources of data within IBM and StorageTek have been held. An ORNL HPSS developer is investigating acquiring performance data from the HPSS 4.2 Gatekeeper server and the Real Time Monitor. Another ORNL staff member will be acquiring file transfer data from the Atmospheric Radiation Measurement system. IBM/Houston is cooperating by providing performance data and an IBM HPSS developer will be participating as time permits. A student at the University of North Dakota, Aric Broeking, and his advisor, Thomas Wiggin, have begun developing the model as Aric's Master's Thesis.

2.2.2.5 GSN Bridging

Genroco offers a GSN Bridge, a device that can aggregate up to eight channels of 100 megabyte/second links into one GSN uplink. The eight channels can be a mixture of Gigabit Ethernet and FibreChannel links. The bridge with two Gigabit Ethernet interfaces and two FibreChannel interfaces has been installed and configured at ORNL for testing the bridging of FibreChannel to GSN, of Gigabit Ethernet to GSN, and of Gigabit Ethernet to FibreChannel.

GSN is also known as HiPPI 6400, reflecting the 6400 megabit/second transfer rate of the specification. ORNL has purchased a GSN switch, populated at this time with six ports, and network interface cards for the SGI Origin 2000 supercomputer and for IBM RS/6000 and Compaq Alpha platforms.

Testing to date has successfully transferred files between each pair of computers with rates as high as 150 megabytes/second. Subsequent testing will involve Genroco's driver for the Compaq Tru64 version 5 operating system (the current driver is for Tru64 version 4). Version 5 is the version in use on ORNL's Compaq AlphaServerSC supercomputers.

2.2.2.6 Remote Mover

NERSC and ORNL have begun to establish an HPSS installation that includes a mover node at NERSC which is part of the ORNL Probe installation and DCE cell. We intend to define a hierarchy with two levels of disk (local and remote) so that HPSS will be handling the migration of files from one site to the other. NERSC will investigate configurations which put the data into a Class of Service that does not stage the data. This will allow retrieval at the remote site without having the data flow through disks on the originating site. This arrangement is targeted at activities involving users at two HPSS sites by having HPSS manage the transfer between sites.

2.2.2.7 HSI Server-to-Mover Transfers

HSI is becoming a widely used and preferred interface to HPSS. ORNL and NERSC are supporting the development of direct HPSS to HPSS copies without moving the data through the controlling client. This ability is especially relevant to national laboratory collaborations, where investigators often need to move data between laboratories. High energy nuclear physics, climate modeling, and human genome projects all have this need.

2.2.2.8 ORNL-NERSC High-Bandwidth Transfers

Transfers between the sites still do not exhibit expected bandwidth. ORNL, NERSC, Lawrence Berkeley Laboratory, and ESnet staff are investigating the possibility that packets are being dropped at a router somewhere in the path.

2.2.2.9 HPSS Compatibility With New Infrastructure Products

HPSS is tested on a specific set of infrastructure products (including DCE, DFS, Encina, Encina's SFS, and Sammi) and on two platforms, IBM/AIX and Sun/Solaris. The various products have different release schedules, so it is usually the case that soon after HPSS is released, some infrastructure product comes out with a new release. The HPSS test team has its hands full testing the functionality of new HPSS patches and releases. They cannot test/certify all combinations of HPSS and infrastructure product releases.

ORNL/Probe has offered to instantiate the latest release of HPSS over the latest releases of infrastructure products. This effort would compile, build and run HPSS in the new environment. Success will provide HPSS customers with some confidence that HPSS operates correctly with the later infrastructure.

2.2.2.10 SCSI-FibreChannel Bridge Testing

ORNL has eight IBM 3590E SCSI tape drives and a need to connect them to a FibreChannel interface for transfer-rate and packaging reasons. To that end a SCSI-FibreChannel Bridge has been acquired and used to connect two 3590E drives to a Probe HPSS node. Testing in Probe is complete. The next step, shifting the production 3590 tape drives to the bridge, is pending and expected within a week or so of this writing. When complete, ORNL will retire several obsolete IBM RS/6000 MicroChannel nodes which until recently were the primary HPSS movers.

2.3 PROJECTS UNDER CONSIDERATION

A variety of attractive and valuable projects are in the pipeline but not being actively performed at this writing. Some are blocked by the availability of vendor software or network equipment, others by scheduling priorities.

2.3.1 ESnet III

Probe will be an early tester of the ESnet III network. At this writing we are awaiting ESnet hardware installation.

2.3.2 Scheduled Transfer (ST)

ST is a software technology that bypasses much of the operating-system work ordinarily performed in high-bandwidth transfers. ORNL has acquired three ST licenses from Genroco for installation on the two Compaq AlphaServer SC supercomputers and the Probe Compaq DS20 server being used in HPSS mover testing. The first study will investigate ST over Gigabit Ethernet links, but it cannot proceed until Genroco completes driver software for Tru64 version 5.

2.3.3 Redundant Array of Independent Tapes (RAIT)/Redundant Array of Independent Library (RAIL)

StorageTek is developing the RAIT system as a PathForward project. The equipment would afford considerable operational advantages to the ORNL production environment, including ease of administration and operation, increased data protection, and higher tape bandwidth. A RAIT system might also be important to future data-sharing products from several vendors. Procurement of a RAIT system is underway at ORNL.

2.3.4 Storage Area Networks (SAN)/IP

The work to investigate bridging FiberChannel to Gigabit Ethernet is a preliminary to studying SANs using IP. Early products are just entering the market; some are tuned for local area networks and others for wide-area networks. They hold considerable promise for wide-area storage transfers and for less-expensive SANs.

2.3.5 Texas Memory Systems (TMS) Solid State Disk Systems

TMS is interested in placing systems at HPSS sites for testing; success would enhance their possible market. We are considering two possible uses of TMS solid-state disks. NERSC would test the use of the solid-state disk to store metadata, evaluate transaction performance and look at how to achieve adequate protection of data. A mirror using conventional disks is being considered as one solution to the data protection issue. We also believe it would be valuable to retrieve statistical information from the TMS box to analyze SFS usage patterns. ORNL is interested in the possibility of using the TMS equipment as a communications buffer between a supercomputer and a visualization engine.

2.3.6 Fibre to Fibre Transfers Within HPSS

Internal transfers within HPSS can benefit significantly from FibreChannel to FibreChannel data transfers. NERSC is reviewing the upcoming SN6000 unit from StorageTek as an entry point into this area. The SN6000 allows the attachment of both FibreChannel disks and FibreChannel tapes; software to support transfers between them is scheduled for availability in the fourth quarter of 2001.

2.3.7 Cross Cell

Users often have projects involving data in two or more HPSS installations. Such users could transfer data between the two sites by FTP, or much more easily using the HSI application described earlier in this report. Such a user might also need to work conveniently on both sites. To do so in a secure manner requires authentication and authorization at both sites, which in turn requires administrative actions to establish cross-cell trust. Once such trust is established, the user should be able to work in both HPSS cells. Tests of that capability, initially between installations at ORNL and subsequently between ORNL and NERSC, are being established.

2.3.8 IBM/Compaq Gigabit Ethernet

Jumbo frames are supported in the same manner on the IBM and Compaq supercomputers as on the Alteon switches that link the two. ORNL intends to investigate transfers between the systems, then try at least two such parallel links between the systems. ORNL will also test ST between the machines over the parallel links when the ST drivers are available.

2.3.9 DBMS/HPSS R&D Associated With HPSS 5.0

ORNL has offered the use of its DB2 and Oracle licenses in testing and development associated with the re-engineering of HPSS. This work would build upon and expand the "Bake-off" project described earlier. The re-engineering effort is still in the early design stage.

2.3.10 GUI Interfaces to Storage

A graphical interface to storage would greatly increase convenience for casual users. NERSC is considering prototypes of web interfaces based on (1) a Web/GRID interface currently being developed and (2) a GUI/XHSI interface.

2.4 COLLABORATIONS

2.4.1 GENROCO

ORNL has established a collaborative relationship with Genroco, maker and vendor of GSN hardware and ST software. Genroco is also prominent in SAN/IP activities, bridging various high-bandwidth network technologies to GSN, use of ST over ATM, and soon 10-gigabit Ethernet. They have held demonstrations at CERN and in Japan.

2.4.2 National Science Foundation (NSF) Memorandum Of Understanding

Two NSF Project Managers, Chuck Koelbel and Maria Zemankova, have expressed interest in Probe and in making access to Probe available to their Principal Investigators. Chuck Koelbel is writing the Memorandum Of Understanding; Maria Zemankova has recently reiterated her strong support of such involvement.

2.4.3 LLNL

ORNL and LLNL are pursuing a collaboration linking our storage testbeds. One project would investigate high-bandwidth wide-area bulk transfers. At this writing negotiations are proceeding between ESnet, Qwest, and several national laboratories with the goal of establishing a very high bandwidth test network.

2.4.4 IBM and StorageTek

ORNL and NERSC have collaborative relationships with IBM in the hardware and software (HPSS) arena. ORNL also has a collaborative relationship with StorageTek. In each case the collaboration allows beta testing of new equipment and capabilities.

2.4.5 Brookhaven National Laboratory (BNL)

NERSC and BNL have collaborative relationships in the area of High Energy and Nuclear Physics data bases and data processing. One area of investigation is query optimization, which incorporates knowledge of file locations on the physical tapes in the HPSS storage systems. Data retrieval for multiple queries can be combined, and optimum ordering of file retrievals can speed up tape processing. This research is also looking at optimizing the original placement of data on tape, given knowledge about typical queries.

2.4.6 NERSC and GRID Collaboration

A GRID interface to storage is available in the NERSC Probe environment.

3. PRESENTATIONS AND PAPERS

- Presentation to the HPSS User's Forum, September 1999, R. D. Burris.
- Presentation to the HEPiX in October 1999, R. D. Burris.
- Presentation to the Eighth Goddard Conference on Mass Storage Systems and Technologies, March 29, 2000, R. D. Burris.
- Presentation to Dan Hitchcock in Oak Ridge in April 2000, R. D. Burris.
- Presentation to the ESCC Meeting, April 25-27, 2000, co-authored by R. D. Burris and H. H. Holmes and presented by H. H. Holmes.
- Presentation to Fred Johnson in Oak Ridge, June 20, 2000, R. D. Burris.
- Presentation to the HPSS User's Forum, July 26, 2000, R. D. Burris.
- Poster Paper at the High Performance Distributed Computing Conference in Pittsburgh, PA, August 1, 2000, co-authored by R. D. Burris, D. L. Million, S. R. White, M. L. Gleicher, and H. H. Holmes.
- Presentation to the ESnet Steering Committee, September 13, 2000, by R. D. Burris.
- Probe was described in the What's New section of the IBM HPSS web site (since superceded).
- ORNL is referenced as collaborating in the improvement of HSI in the on-line publication NPACI and SDSC Online dated September 20, 2000.

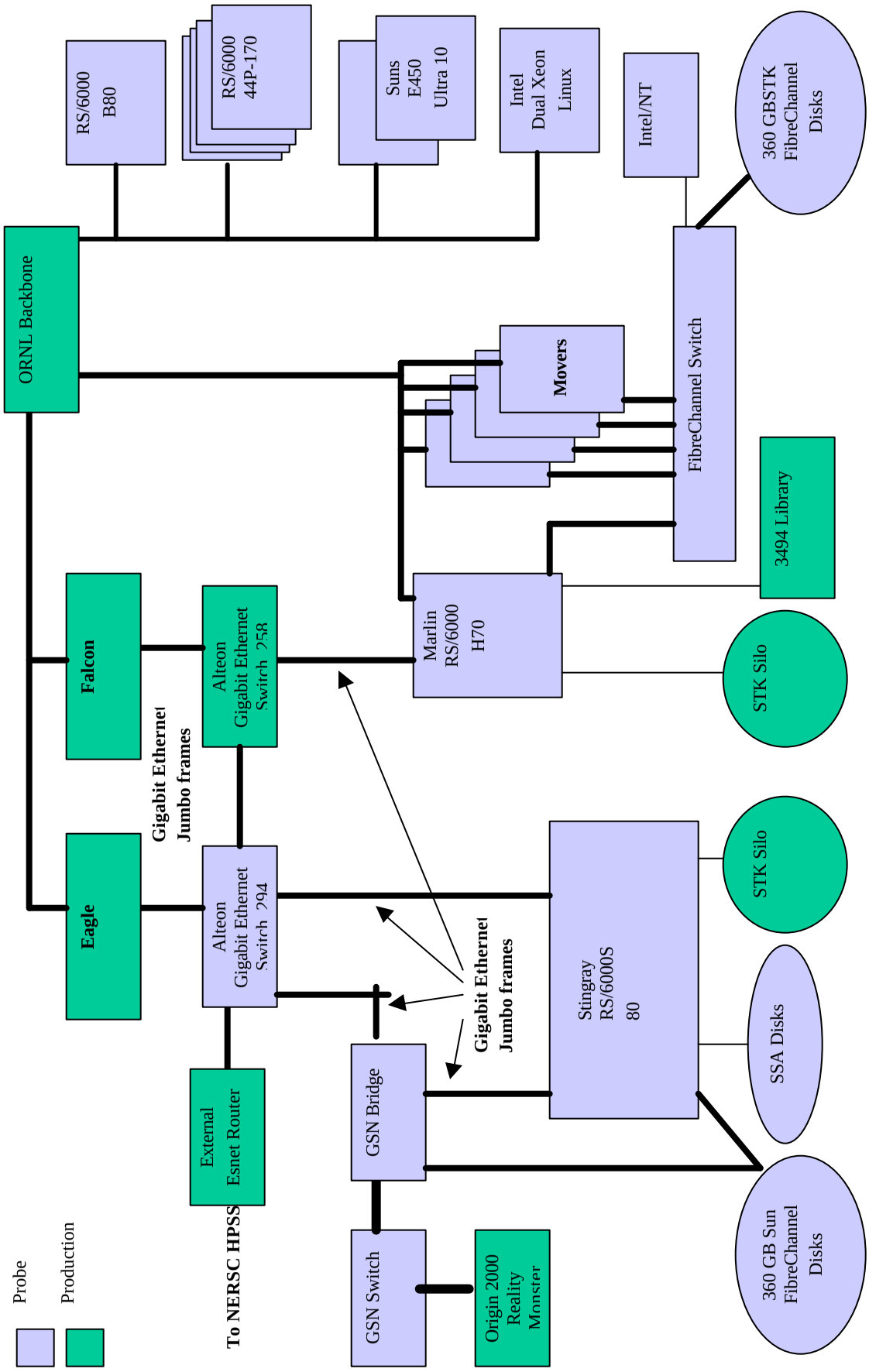
SUMMARY

The year 2000 has seen Probe evolve from a simple testbed to an institution with several collaborative relationships. It has moved from an initial focus on its own sites to involvement in testing on behalf of, and in cooperation with, several external entities. While continuing to perform valuable testing on storage and network hardware and software, it now supports network research and the development of important applications. We anticipate further expansions in scope and collaborations during 2001, as well as participation in challenging new SciDAC projects (if funded), while continuing to support storage and network research and to initiate studies of new applications, drivers and equipment.

APPENDIX A

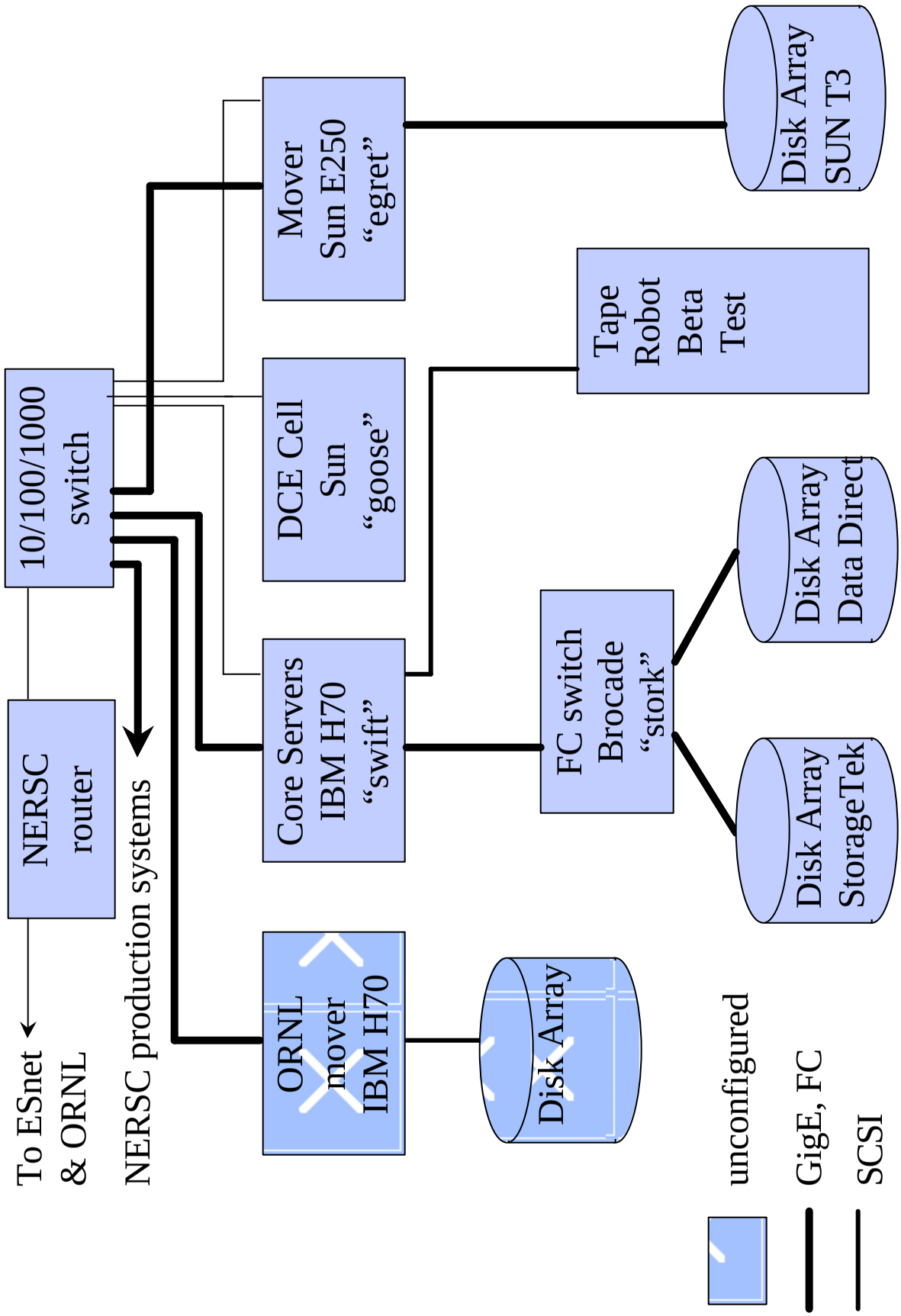
CONFIGURATION OF ORNL EQUIPMENT AS OF DECEMBER, 2000

APPENDIX A CONFIGURATION OF ORNL EQUIPMENT AS OF DECEMBER, 2000



APPENDIX B

CONFIGURATION OF NERSC EQUIPMENT AS OF JANUARY 2001


Disk Array
SUN T3
Mover
Sun E250
"egret"
Tape
Robot
Beta
Test
10/100/1000
switch
DCE Cell
Sun
"goose"
Disk Array
Data Direct
FC switch
Brocade
"stork"
NERSC
router
Core Servers
IBM H70
"swift"
Disk Array
StorageTek
To ESnet
& ORNL
NERSC production systems
ORNL
mover
IBM H70
Disk Array
unconfigured
GigE, FC
SCSI

INTERNAL DISTRIBUTION

1. R. A. Alexander
2. L. F. Arrowood
3. A. S. Bland
4. R. D. Burris
5. T. Dunigan
6. G. A. Geist
7 C. A. Giles
8 R. A. McCord
9. D. L. Million
10. G. Ostrouchov
11. N. Samatova
12. D. A. Steinert
13. R. J. Toedte
14. J. B. White, III
15. V. L. White
16. W. R. Wing
17. S. D. Woods
18. B. A. Worley
19. T. Zacharia
20. Central Research Library
21. ORNL Laboratory Records (RC)
22-23. ORNL Laboratory Records—OSTI

EXTERNAL DISTRIBUTION

24. John Blaylock, LANL
25. Jeff Bongianino, GDE Systems
26. Aric Broeking, University of North Dakota
27. Shreyas Cholia, LBNL
28. Danny Cook, LANL
29. Bob Coyne, IBM
30. Mike Devaney, PNL
31. Keith Fitzgerald, LNL
32. Jim Fox, University of Washington
33. Mark Gary, LNL
34. Krzysztof Genser, FML
35. Mike Gleicher, Gleicher Enterprises
36. Otis Graf, IBM
37. Phil Greene, StorageTek
38. Annette Hamala, IBM
39. Dan Hitchcock, Department of Energy
40. Harvard Holmes, LBNL
41. Fred Johnson, Department of Energy
42. Nancy Johnston, LBNL
43. William E. Johnston, LBNL
44. Hilary Jones, SNL
45. Charles H. Koelbel, NSF
46. Jae Kerr, IBM
47. Bill Kramer, LBNL
48. Greg Lefelar, IBM
49. Joe Lopez, San Diego Supercomputer Center
50. C. William McCurdy, LBNL
51. Nancy Meyer, LBNL
52. Thomas Ndousse, Department of Energy
53. Bill Nickles, ANL
54. John Noe, SNL
55. Ramin Nosrat, IBM
56. Bernard O'Lear, NCAR
57. Juliet Pao, NASA
58. James Patton, CalTech
59. Don Petravick, FNL
60. Walter Polanski, Department of Energy
61. Bill Rahe, SNL
62. Arie Shoshani, LBNL
63. Horst Simon, LBNL
64. Ivan Sipos, Compaq
65. John Sobolewski, University of New Mexico
66. Danny Teaff, IBM
67. Dick Watson, LNL
68. Thomas Wiggen, University of North Dakota
69. Dave Wiltzius, LNL
70. Don Woelz, GENROCO
71. Maria Zemankova, NSF