# Robotic Tape Library System Level Testing at NSA, Present and Planned

Michael F. Shields
Department of Defense
9800 Savage Road
Fort Meade, Maryland 20755
Tel: +1-301-688-9509
Fax: +1-301-688-9454

In the present era of declining Defense budgets, increased pressure has been placed on our Agency to utilize Commercial Off The Shelf (COTS) solutions to incrementally solve a wide variety of our computer processing requirements. With the rapid growth in processing power, significant expansion of high performance networking, and the increased complexity of applications data sets, the requirement for high performance, large capacity, reliable and secure, and most of all affordable robotic tape storage libraries has greatly increased. Additionally, the migration to a heterogeneous, distributed computing environment has further complicated the problem. With today's open system compute servers approaching yesterday's supercomputer capabilities, the need for affordable, reliable secure Mass Storage Systems (MSS) has taken on an ever increasing importance to our processing centers' ability to satisfy operational mission requirements. To that end, NSA has established an in-house capability to acquire, test, and evaluate COTS products. Its goal is to qualify a set of COTS MSS libraries, thereby achieving a modicum of standardization for robotic tape libraries which can satisfy our low, medium, and high performance file and volume serving requirements. In addition, NSA has established relations with other Government Agencies to complement this in-house effort and to maximize our research, testing, and evaluation work. While the preponderance of the effort is focused at the high end of the storage ladder, considerable effort will be extended this year and next at the server class or mid range storage systems.

Over the past year, we have performed extensive testing of several high performance, high capacity Mass Storage Systems. In the open systems arena, we have evaluated the Convex based EMASS FileServ hierarchal storage management (HSM) product, see figure 1. Initially, the system was tested for use in one of our processing areas as the deep

Figure 1 - CONVEX Based EMASS FileServ HSM
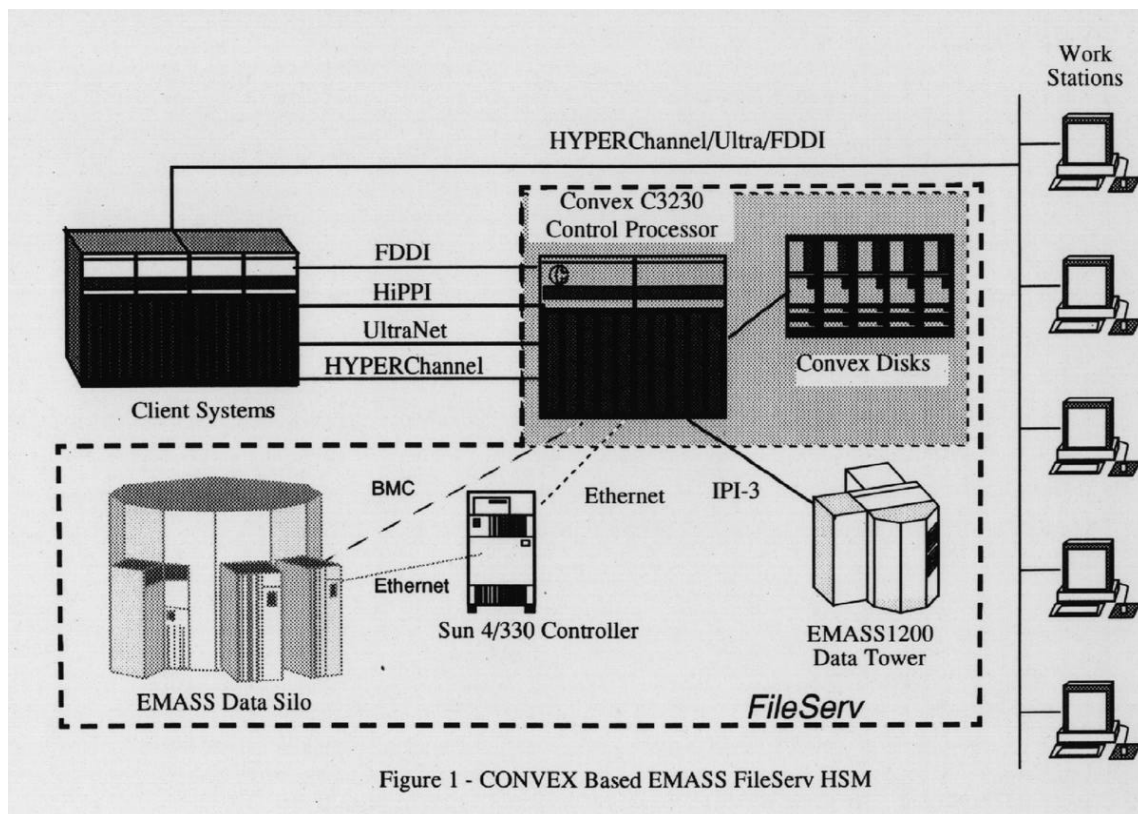
storage/archive for multiple server class Unix based systems. These client systems were networked using FDDI to the HSM which managed the multiple clients' stored files. Classes were created, disk and tape capacity were dedicated to each client, and policies were established to tune the system for each client's storage and retrieval needs. A dedicated client system under the control of the test team was also included in the configuration under test, so as to baseline the load and to control feeds and flows as the test progressed. This element (a dedicated client system) is a recommended must for any system level test. To establish a consistent approach to testing this and other Mass Storage Systems, a standard test approach was developed. The first phase of this standard test was to qualify all of the vendor's commands and extensions and to verify that they operated as advertised. Once this was completed, we used the dedicated Test Client System to generate files of varying sizes and frequencies. This was essential to establish the baseline load. We then would vary the feeds and flows and measure the change as multiples of our baseline load, (e.g. 2X, ... 10X, etc.). Since almost every HSM's performance is highly dependent upon file size, we established three sizes of files (small: 1.5 MB, medium: 10-15 MB, and large 150 MB) in order to adequately categorize the system's end-end performance. Over the duration of our testing we carefully controlled the file size parameter by phase so as to measure the optimal disk and tape allocations for each client system. Our goal was not to break the system, but to establish the optimal range in which to have it operate most efficiently.

In the early phases of our testing, we spent significant time comparing the file sent to the HSM with the data stored. By performing check sums on each file stored for about a two week period, we discovered a flaw in the Convex D2 tape driver microcode, which was quickly fixed by the vendor. After two weeks of verifying that all of the data in each file was successfully written to tape and could be retrieved, this testing was suspended. At our Agency, data integrity is paramount and must always perform at 100%. Suffice it to say that, although we were about the 20th customer for this commercial product, no other customer had experienced the data integrity problem in their facility. We believe that this was due to their testing approach. Although we do NOT normally perform this degree of data integrity testing for most commercial products, it is strongly recommended that it be done for any new tape drive that is introduced. Since we were using the EMASS ER90 Helical Scan D2 drives, we felt it necessary to verify data integrity at a high confidence level; as our tests indicated, this was a wise step. We will also do this for IBM's NTP and STK's REDWOOD drives before they are placed into production.

The next phase of our testing was aimed at sustainability and reliability. Since our storage paradigm is to have all Mass Storage Systems located in unmanned spaces and to
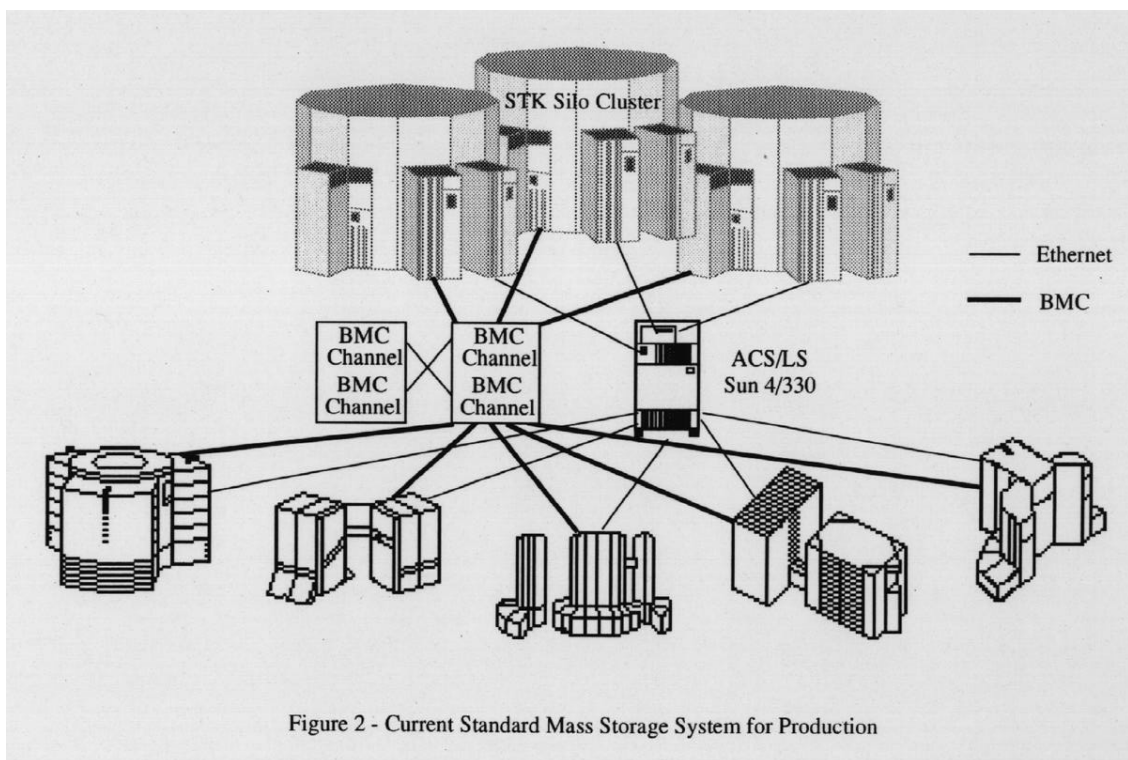

Figure 2 - Current Standard Mass Storage System for Production

be remotely monitored by a geographically separated command center, production storage systems must be highly reliable and be capable of degraded mode operations. They must operate for long periods of time without operator/maintenance intervention to justify their existence. Our current standard for reliability is the STK silo which is our main line Mass Storage System for today's production, see figure 2. Over the past year, all of the drives/controllers have been upgraded to 36 track, and we are about 35% completed with the infusion of 800 MB tapes. Over the past five years, we have had only

a small level of problems with these systems as they only require preventative maintenance at 6 month intervals. With self contained cleaning cartridges, they have proven to be highly reliable and satisfy our personnel staffing limitations.

Our Convex/EMASS HSM was initially tested with 4 ER90 D2 Helical Scan drives which were housed in an Odetics Data Tower. Its capacity was about 5.7 TBs. During the reliability/sustainability testing phase, we experienced significant difficulty with the robotics. Problems encountered included several instances of "stuck tapes", several dropped tapes, excessive mechanical wear on the cassettes themselves, and repeated failure of the robotic hub itself. Over a nine month period, five hub failures were experienced. The lack of reliability of the hub in large measure caused the Government to fail the system acceptance test. While the contractor went to yeoman efforts to attempt to correct these deficiencies, the problem persisted. A side effect of the robotics failures made endurance testing of the drives impossible; even still, we experienced a fair level of problems which made the drive questionable for "lights out" use. The principal problems encountered with the drives were head related. We determined that in order to have a margin of safety, we needed to have operators clean the heads shortly after 50 head/tape contact hours of use. In order to accurately monitor remotely when this event occurred, software had to be written which accessed firmware counters in the drives/controllers. In addition, operators had to monitor the error correction code counters in the drives. Again, software had to be written to enable this activity. We discovered that as soon as a drive had to employ the second level ECC, it was prudent to vary the drive off line to preclude a permanent write error. Employing this technique, we never encountered a hard write error. However, the degree of monitoring by our test team and operations personnel was deemed excessive. Another significant problem encountered was in the quality of the replacement heads. While some heads greatly exceeded their warranty hour limit., others failed prematurely. We concluded that this resulted from poor quality control in the manufacturing process. However, we noted that once a set of heads got past the 70-100 hour mark, they tended to be reliable for their design life, and often exceeded it. But once again, the degree of monitoring and maintenance intervention made this library unsuitable for our lights out processing scenario.
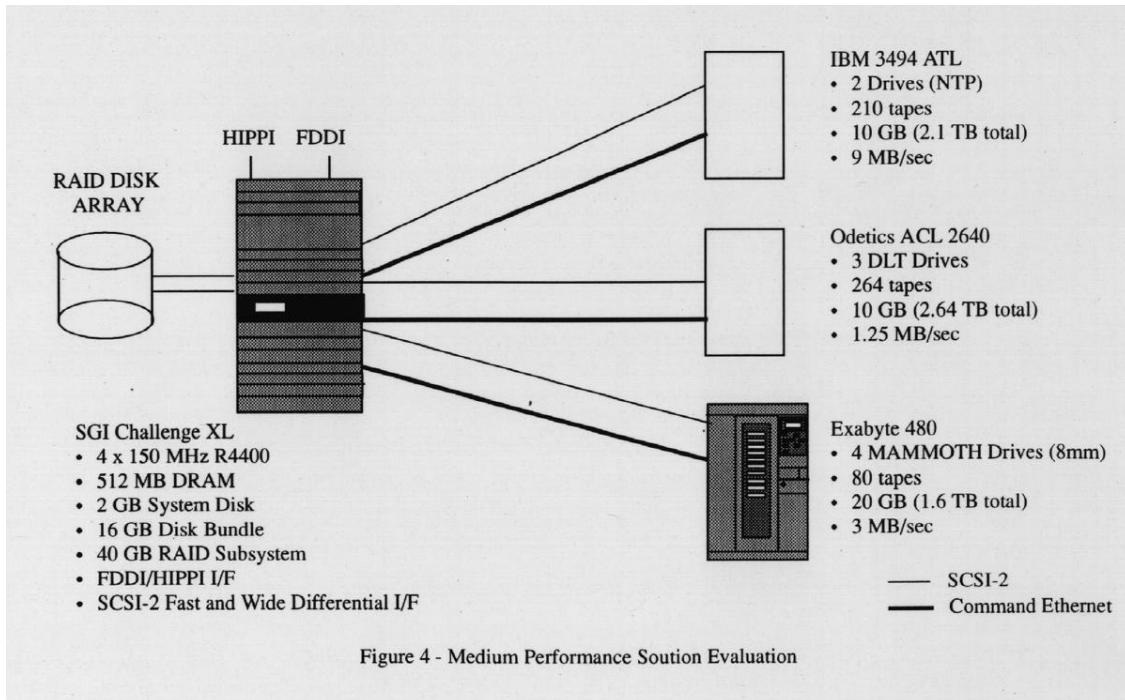
As a result of the aforementioned problems, the Government acquired an STK POWDERHORN 36 track/800 MB per tape Silo system with eight 4490 drives. This system was connected to our Convex/EMASS HSM as the second archive. It underwent its acceptance testing without a single problem. With the long tape, it provided a capacity of 4.4 TBs. The Government was highly interested in verifying that the EMASS FileServ software could effectively control two archives with different drive types. Since the client systems had a preponderance of small files, on the order of 6-10 MBs, the STK robotics and drives outperformed the Odetics/ER90 configuration. However, when the file sizes were changed to 100-150 MBs, the ER90s were more efficient. Testing of the mixed mode archive continued until the Government reached a level of confidence that the system could perform as advertised. At that time, the Odetics Tower and ER90 drives were deinstalled and returned to the integration contractor. Noting the deficiencies

Figure 3 - CONVEX Based EMASS FileServ HSM with GRAU ABBA/2 and NTP Drives
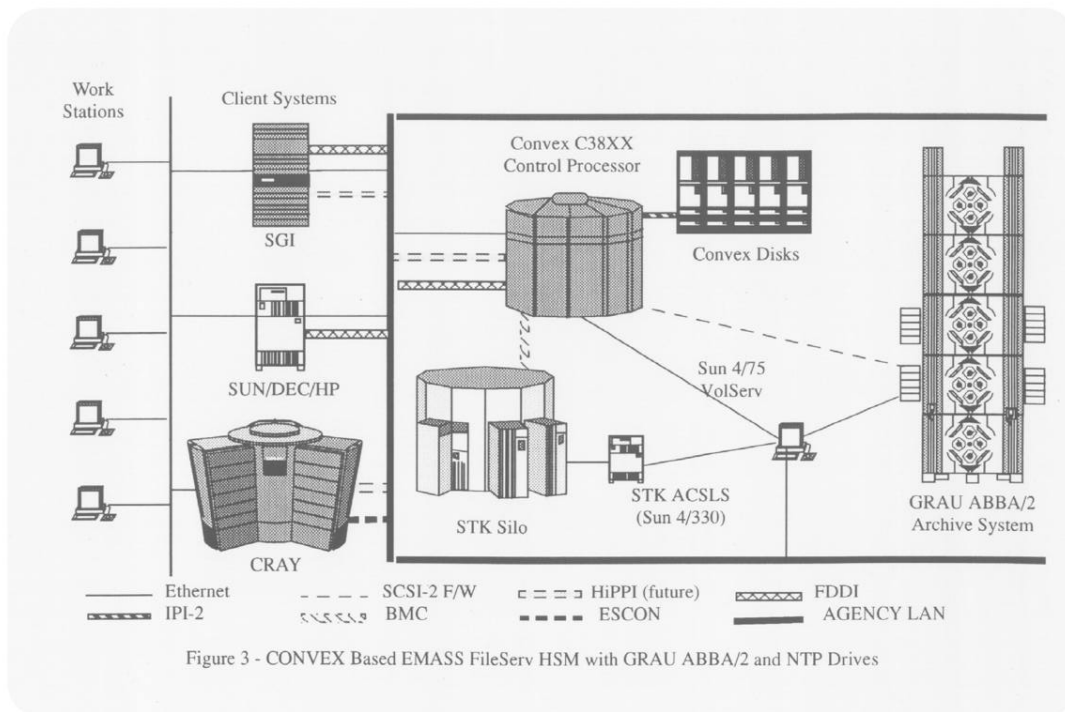
encountered during the testing, the contractor offered to deliver a Grau ABBA/2 robotic library with IBM NTP drives as a replacement, see figure 3. This system will be integrated and tested with the Government's loading scenarios at the contractor's facility prior to shipping the system to NSA in January 1996.
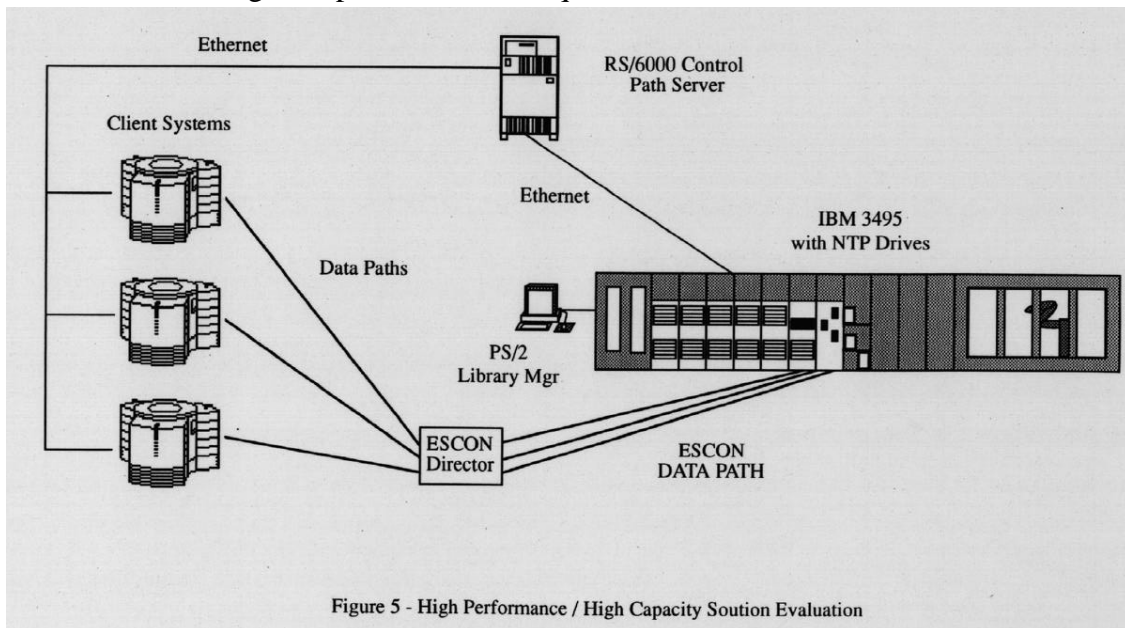
The Convex/EMASS FileServ System with STK Silo and drives is now in production at NSA. While the steps cited above are somewhat skewed to our specific clients/networks, we believe that our test approach is sound and generally applicable to any robotic tape HSM. In early 1994, we applied the same testing approach to a Sun/AMASS/Metrum RS48 robotic tape system. Once again, the theme was to verify the command set and functionality, verify the data integrity, evaluate the reliability and sustainability of the drives/robotics, and to categorize the sustained throughput of the system. We found this to be a stable product for low performance Mass Storage requirements.

NSA will evaluate the following server class systems during CY95. For the medium

RAID DISK
ARRAY

HIPPI FDDI

IBM 3494 ATL
• 2 Drives (NTP)
• 210 tapes
• 10 GB (2.1 TB total)
• 9 MB/sec

Odetics ACL 2640
• 3 DLT Drives
• 264 tapes
• 10 GB (2.64 TB total)
• 1.25 MB/sec

SGI Challenge XL
• 4 x 150 MHz R4400
• 512 MB DRAM
• 2 GB System Disk
• 16 GB Disk Bundle
• 40 GB RAID Subsystem
• FDDI/HIPPI I/F
• SCSI-2 Fast and Wide Differential I/F

Exabyte 480
• 4 MAMMOTH Drives (8mm)
• 80 tapes
• 20 GB (1.6 TB total)
• 3 MB/sec

——— SCSI-2
—— Command Ethernet

Figure 4 - Medium Performance Soution Evaluation

performance solution we have acquired an SGI Challenge series computer running the AMASS software. Three different robotic/drive configurations will be tested, see figure 4. They are IBM 3494/NTP for high performance/high capacity, Quantum DLT/Odetics 2640 for medium performance/capacity, and Exabyte 480/Mammoth for low/medium performance/capacity. Once again we will use the same approach as outlined above to evaluate/categorize these configurations.

For the high end high performance/high capacity robotic tape requirements of the Scientific Processing Complex, we have acquired two different volume servers. The first

Ethernet

RS/6000 Control
Path Server

Client Systems

Ethernet

IBM 3495
with NTP Drives

Data Paths

PS/2
Library Mgr

ESCON
Director

ESCON
DATA PATH

Figure 5 - High Performance / High Capacity Soution Evaluation

is an IBM 3495 L20 with 8 NTP drives which is being qualified by Cray Research Inc (CRI), see figure 5. Once this qualification is completed, the system will be fielded at NSA and will undergo in-house testing in late CY95. The second high performance/high capacity system to be tested is a Grau ABBA/2 robotics with NTP drives. It also will be qualified by a cooperative effort by E-Systems and CRI. Once the system is qualified, it will be shipped to NSA and will undergo in-house testing in early CY96. Both of these systems will have ESCON connectivity to the drives, which will facilitate sharing of the system by any of our Crays.

We feel that the Mass Storage community should establish performance benchmarks for products to aid the customer community in selecting the right Mass Storage products for their operational requirements. Our experience is that none of the vendors can provide the right size system configuration for any customer's needs. Today the entire burden for system sizing and delivered performance rests on the customer. Vendors need to perform and disseminate more evaluation information. They need to cover as broad a range as possible, either with their own testing or teamed with another, larger vendor who has the resources needed to perform the tests. Given a broad enough range of tests, customers should be able to take the results and extrapolate the expected performance characteristics for their environment. Solutions should be predictable and must include control processor network bandwidth, memory and disk needs, channels and I/O bandwidth, numbers of drives, and controllers, and robotics speeds. All that really matters for a Mass Storage System is the end-to-end sustainable bandwidth for stores and retrieves between the clients and the HSM. The industry must address some form of performance benchmark standards which will be the first step in aiding the customer in selecting the right system configuration for their unique problem. We have SpecMarks for processors, TP benchmarks for Data Base machines, but have nothing for storage systems. This area must be addressed soonest by the storage vendors. NSA has invested significant in-house resources in evaluating just a few of the systems available on the commercial marketplace. Scalability of the number of files that the HSM can manage is another key area of uncertainty. NSA has discovered the EMASS FileServ HSM has scalability limits largely caused by their use of Ingres RDBMS software in their commercial product. While this is a temporal limit, scalability testing of the product by vendors in-house, prior to first customer ship is a must for companies to survive. Competition dictates that this must be done and done quickly.

In summary, our approach should be clear as we have standardized on three different robotic tape systems for our high end processors across our various computer complexes. These are IBM 3494/5, Grau ABBA/2, and STK 4400 silos. The drives used with these robotics include: IBM/STK 36 track, NTP, and D3. For our server class systems we have a similar approach envisioned, as outlined above. The principal common entity for the server class problem is the AMASS HSM product. The specific drive/robotics and platform will be selected based on the required performance and capacity. Regarding high performance file servers, we will evaluate the scalability of the EMASS FileServ product during early 1996 and make our decision regarding its suitability for 150+ TB

libraries. Our ultimate goal for the future would be to have one logical shared robotic tape library, accessible by any of our computer complexes.