N94-33826

Distributed Active Archive Center

Lee Bodden
Hughes STX
Code 902.2, GSFC, NASA
Greenbelt, MD 20771
Telephone: (301) 286-9413
bodden@eosdata.gsfc.nasa.gov

Phil Pease
NASA
Code 902.2, GSFC
Greenbelt, MD 20771
Telephone: (301) 286-4418
pease@eosdata.gsfc.nasa.gov

Jean-Jacques Bedet
Hughes STX
Code 902.2, GSFC, NASA
Greenbelt, MD 20771
Telephone: (301) 513-1646
bedet@eosdata.gsfc.nasa.gov

Wayne Rosen
Hughes STX
Code 902.2, GSFC, NASA
Greenbelt, MD 20771
Telephone: (301) 286-3526
rosen@eosdata.gsfc.nasa.gov

Abstract

The Goddard Space Flight Center Version 0 Distributed Active Archive Center (GSFC V0 DAAC) is being developed to enhance and improve scientific research and productivity by consolidating access to remote sensor earth science data in the pre-EOS time frame. In cooperation with scientists from the science labs at GSFC, other NASA facilities, universities, and other government agencies, the DAAC will support data acquisition, validation, archive and distribution. The DAAC is being developed in response to EOSDIS Project Functional Requirements as well as from requirements originating from individual science projects such as SeaWiFS, Meteor3/TOMS2, AVHRR Pathfinder, TOVS Pathfinder, and UARS. The GSFC V0 DAAC has begun operational support for the AVHRR Pathfinder (as of April, 1993), TOVS Pathfinder (as of July, 1993) and the UARS (September, 1993) Projects, and is preparing to provide operational support for SeaWiFS (August, 1994) data. The GSFC V0 DAAC has also incorporated the existing data, services, and functionality of the DAAC/Climate, DAAC/Land, and the Coastal Zone Color Scanner (CZCS) Systems.

Introduction

This paper presents the architecture of the DAAC which includes two SGI 4D/440 mini-supercomputers and numerous smaller computers including: an HP 730, MicroVAX II, VAX 3900, SGI 4D/35 and three SUNs all configured in a distributed environment. The DAAC contains two different mass data storage systems, a Cygnet 1803 12" WORM Optical Jukebox and a Metrum RSS 600 VHS Automatic Tape Cartridge System. Both systems are being configured under the UniTree File Management System. The DAAC also supports a host of peripheral devices including two 9-track tape drives, three 8 mm tape drives, two 3480 tape drives, two 4 mm, two CD ROM drives, over 40 GB of magnetic disk storage, ten X-terminals and over 25 Macintoshes and personal computers. The DAAC's distributed environment includes two ethernet Local Area Networks, an FDDI network interface, two appletalk networks, and a T1/T3 link. This paper presents the advantages and disadvantages of the chosen architectural approach of the DAAC including a discussion of the cost trade-off analyses justifying the decisions made by the DAAC. This paper also discusses the system performance characteristics in terms of throughput rates and volumes for the data ingested into the DAAC's archive and for data distribution conducted by the DAAC. The percentages of data distributed on different media, and the medias popularity is also discussed.

GSFC V0 DAAC Mission

The Distributed Active Archive Center (DAAC) is a component of NASA's Earth Observing System (EOS) Data and Information System (EOSDIS). The EOSDIS acquires Earth science data, derives scientifically useful data products, archives the data products and makes them available to the Earth science researchers. The EOSDIS currently includes eight DAAC sites. These DAAC sites are generally oriented around scientific disciplines and are multi-agency.

A DAAC consists of three components, a Product Generation System (PGS) that generates derived data products, a Data Archive and Distribution System (DADS) that stores the data products and distributes requested products to a researcher, and a Information Management System (IMS) that are used by researchers as a catalog of all the DAAC products from which he/she can select specific data files of interest. The IMS allows the user to select data based on time, spatial location, geophysical parameter and/or instrument. The IMS will also provide a capability to browse interesting data products as an aid to ordering the data. The IMS at all the DAACs are interoperable so the user sees the holding of all the DAACs and can order them from any DAAC he/she logs into.

This paper focuses on the DADS component. Data are ingested into the DADS primarily over the EOSDIS dedicated computer network either from instrument data capture facilities, other DAACs, or from the DAAC's own PGS. Metadata information is extracted or created from each data file and loaded into the IMS database. The data are archived to on-line (magnetic disk), near-line (robotics storage system), or off-line (on the shelf) storage. When an order for data is received via the IMS the data are copied from the archive to either magnetic disk for network (FTP) distribution or to magnetic tape (8mm, 4 mm, and 9 track are standard media supported).

The EOSDIS and the DAAC elements are being developed in an evolutionary manner with Version 0 being the initial system. Version 0 is intended to demonstrate the concept of an interoperable set of distributed archive centers and to prototype various aspects of the system. The version 0 will operate with pre-EOS satellite data sets, either currently existing or missions between now and the EOS flights.

Requirements

The GSFC V0 DAAC archive will contain about 20 Terabytes by FY97. The amounts of data expected from the projects and sources interfacing with the DAAC is shown in Figure 1. Rates for data delivery into the DADS are expected to reach 17 GB/day via a computer network. FTP data distribution and other networking activity is expected to double this figure for a total network load of 30 to 40 GB/day. Estimates are that distribution volumes may reach 50 to 60 GB/day. It is estimated that for tape distribution, 50% will be on 8mm cartridges, 33% on 9-track 6250 bpi round tapes and 17% on 4 mm cartridges. Distribution on prepublished CD-ROMs will also be supported.

The researcher will be able to order and receive small amounts (TBD) of data via network transmission during an interactive session while logged on to the DAAC's computers. Larger amounts of data will be available for distribution on the various media supported by the DAAC. The guideline is that all orders will be filled within 30 days, with 3 days response time being a desirable goal. Specific data sets have been identified by a Science User Working Group as being high priority (expect a lot of scientific interest) and with this prioritization, the DAAC have organized their on-line, near-line, and off-line archive storage to have the higher priority data more readily available.

One group of data that will also be stored on-line and accessed interactively by a user are the browse products. Browse products are reduced resolution images used as an aid for selecting and ordering data. The user will need to have the analytical tools required to display these browse images. Other data products such as scientific documentation describing the data sets will also be available for ordering.

Data compression is planned prior to archiving in order to reduce storage needs. The DAAC will encourage users to accept data in compressed form but will decompress the data prior to distribution to the user if desired. Data compression is also being recommended for the data being transmitted into the DAAC from the various supported science projects.

Strategy and Approach

The approach being used to meet the above requirements begins with an analysis of the system capabilities. This analysis initially was done using crude spreadsheet calculations of overall bandwidth for networks and published write rates for various peripheral devices. These simple calculations were used to arrive at a hardware configuration that was "in the ballpark" and the computer and two each of each peripheral were ordered. This initial configuration provided a platform for software development and for making performance measurements to gain better throughput figures.

We then initiated the development of a computer simulation of the workload, configuration, and operation of the DADS. Performance measurements were made to determine parameter values to be used in the computer model (e.g. actual device transfer rates achieved using operational software) and overall system throughput was calculated and compared with the simulation results (e.g., simultaneous ingest and multiple distribution activity).

With the validation of the computer model, using the benchmark measurements, the model can then be selectively modified to assess changes to the system configuration (e.g., faster processors, more tape drives, more disk space, use of data compression, number of operators) or workload (e.g., different proportion of distribution media types requested, greater number of requests for data).

Finally, after all hardware and software development and integration is completed the DAAC will perform a formal test of the systems ability to meet the performance requirements. These tests will also include stress testing to determine the upper limits of the processing capabilities of the DAAC.

Trade-off Analyses

The trade-off analyses for the DAAC began with the evaluation of different computers and operating systems. The types of media that would be supported and the drives were also analyzed. Most importantly, the DAAC evaluated the mass data storage hardware currently available and the file management systems that will support the hardware.

All major computers available on the market today were evaluated during this analysis. Each computer was evaluated against the following criteria:

- MIPS
- Internal BUS throughput
- Individual magnetic disk storage capacity
- Magnetic disk transfer rates
- Power requirements
- Drives supported including interface mode
- Total number of drives supported
- Long-range maintenance
- Network connectivity
- Product reliability
- Applications s/w supported (DBMS,tools)
- Cost
- MFLOPS
- SCSI & IPI channel throughputs
- Total magnetic disk storage capacity
- Operating system/planned upgrades
- Space requirements
- Availability of device drivers
- Upgrade path
- Total memory capacity
- Product quality
- Procurement vehicle
- File management systems supported
- Standards supported(x-window/motif)

The information for the criteria listed above was collected and compared, and the SGI 4D/440 S and 4D/440 VGX computers were selected. This computer provided a very cost-effective MIPS/$ with an eight CPU expansion capacity per computer. The SCSI channel and internal BUS throughput rates were fast enough to meet the requirements of the DAAC and the disk storage capacity could also be expanded to meet the DAAC's needs. The space efficient SGI had high marks for quality and reliability with a low maintenance record. The SGIs also provided fddi and ethernet network connectivity.

The peripheral drives and corresponding media were also investigated as part of the overall computer system using the following additional criteria:

- Drive transfer rates
- Media capacity
- Available device drivers
- Popularity of media and drive
- File search time
- Media longevity
- Compatibility with host computer
- Cost of media and drives

Using this criteria, the DAAC selected SGI 8 mm, 4 mm, QIC, 9-track, and CD-ROM drives. Many of these drives were third party hardware sold by SGI. The risk of having interface problems with the computers was greatly reduced by selecting drives that have been thoroughly integrated. Two Fujitsu 3480 drives (one with a stacker) were also procured. This wide range of peripherals allows the DAAC to provide support to a broad base of users, an important consideration for the DAAC.

The DAAC analyzed mass storage hardware systems and the corresponding file management systems. The criteria used in this analysis were:

- Drive transfer rates
- Media capacity
- Mass storage system capacity
- Available device drivers
- Power requirements
- Reliability in the field
- Procurement vehicle
- Cost of file management system/licenses
- Maturity of file management system
- Data format and standards
- Adherence to IEEE Mass Storage Reference Model
- Multiple mass data storage systems supported
- File search times
- Media longevity
- Expandability and upgrade paths
- Compatibility with host computer
- Quality
- Maintenance costs
- Cost of hardware, media and drives
- Space requirements
- Functionality provided
- Supports hierarchical file migration
- Vendor support
- Integration support

The results of some of these analyses are shown in Figures 2 and 3. The Cygnet 1803 12" WORM Optical Jukebox and the Metrum RSS 600B VHS Automated Tape Library were selected along with UniTree as the file management system. The optical media inside the Cygnet jukebox provides the DAAC with a long-life substrate for its most important data. The Cygnet jukebox also provides rapid access for files that require it such as browse data. The media cost does prohibit all of the data from being place on the Cygnet jukebox. The Metrum provides a slightly higher throughput than the Cygnet jukebox with a very cost-effective $/TB ratio. The low cost of the media makes the Metrum the DAAC's selection for where most of the data will be stored. The UniTree file management system is the only system that can support both mass storage systems, although support for the Cygnet jukebox and the dual support capability were introduced into UniTree at the request of the DAAC. The selection of a file management system was essential in avoiding expensive development and maintenance costs associated with providing this functionality as part of the software development effort.

Hardware and Software Selected

The Silicon Graphics Inc. 4D/440 VGX computer is a four CPU machine that was selected for the IMS. It can be upgraded to an eight CPU version (4D/480) by simply plugging in additional boards. The ease of expansion and the relatively inexpensive cost was a factor in the selection of this system. Other factors are the commercial software packages available for this platform.

The database manager product used in the IMS is Oracle. Oracle was chosen primarily because in had been successfully used previously on other data systems that the DAAC organization continues to operate. Another factor is that the Oracle product on the SGI computer can use any and all of the processors available and thus as the need requires additional CPU boards can be added. A feature used with oracle is configuring for separate tables and interface from remote machines for software development activities, system testing activities, and for operational activities. The large number of platforms for which Oracle is available provides flexibility in future system configuration changes.

The IMS user interface was implemented using the JYACC Applications Manager (JAM). This product allowed us to create both the interface for alphanumeric users and for graphical users without needing to develop separate programs. JAM also supported interface with the Oracle database product and allowed running the interface from remote systems without additional license

costs. The wide variety of platforms for which JAM is available provides flexibility in future configuration changes.

The SGI 4D/440 S is a four CPU machine in a server configuration that is the computer system selected for the DADS activities. This server configuration substitutes the graphics hardware [in the VGX model] with additional I/O capacity. Like the IMS machine this system is expandable to eight CPUs. Cost and availability of software was a factor in selecting this computer system. Also, having the same operating system for both the IMS and DADS makes system support easier.

The DAAC selected two mass data storage systems for its archive. The first is a Cygnet model 1803 12" WORM Jukebox with two ATG Gigadisc model GD9001 WORM drives. The ATG WORM platters hold 4.5 GB per side. With the two drive configuration, the Cygnet holds 131 platters providing a total storage capacity of 1179 GB. The second mass storage system is a Metrum model RSS-600B Automated Tape Library system with four model RSP-2150 VHS Cartridge Tape Drive Subsystems. The DAAC is currently using ST-120 VHS cartridges, that hold 14.5 GB per cartridge. The RSS-600B holds 600 cartridges providing a total storage capacity of 8700 GB. The Metrum system can also be used with ST-160 VHS cartridges that hold 18 GB/cartridge yielding a storage capacity of 10800 GB. The DAAC will be storing its low level (L1) data on WORM because of it's reported long life characteristic and the higher level (L2, L3, and L4) on VHS tape because this data is more likely to be reprocessed and replaced as better scientific processing algorithms are developed.

UniTree Central File Manager (UCFM) from Titan Client/Server Technologies and Open Vision was selected to manage the archive. Agreements were reached to introduce into this version (1.6.1) of UniTree support for mixed mass storage media. It also has been enhanced to support asynchronous I/O and thus can take advantage of the multiple CPUs of the SGI 4D/440 machine to give improved performance for simultaneous archive and multiple distribution activities.

Hardware and Network Architecture

The hardware architecture of the DAAC is shown in Figure 4. The functionality of the DAAC was distributed over two computer systems in the operational configuration; the Information Management System (IMS) and the Data Archive and Distribution System (DADS).

For distribution of the large number of data orders, anticipated for 8 mm cartridges and 4 mm DAT media, several of the distribution tape drives are configured in a tape stacker configuration. This will reduce the workload on the operations staff for mounting and dismounting of media.

Functional Capabilities

Information Management System (IMS)

Users will connect to the IMS computer through the GSFC V0 EOSDIS Ethernet LAN [network]. The user interacts with the IMS system through an interface program that support either an alphanumeric or graphics terminals. There are actually two IMS interface programs; one is an EOSDIS IMS interface that interacts with all of the DAAC sites (there are currently eight) and the other interacts only with the local (i.e., GSFC) DAAC. With the EOSDIS IMS the user sees the holding at all the DAACs while the GSFC local IMS only sees the GSFC DAAC holdings.

Either of these IMS user interfaces is used to search a database containing metadata information for the DAAC data holdings in order to identify and then request desired data. Users may also order browse data, for data sets of interest, that may be viewed on his/her local workstation, and to directly order the corresponding data file from the browse viewer program. Orders for data are stored in an order database. Ordered data may be retrieved over the network or copied to media and mailed to the user.

Data Archive and Distribution System (DADS)

The DADS provides two main functions; the ingest and archiving of data and copying of data from the archive to a disk for network distribution or to media for distribution by that mechanism. Most

of the data to be ingested into the DAAC are transmitted over the GSFC V0 EOSDIS FDDI LAN using a client/server program to transfer this data in a fully automated manner. When the data arrives on the ingest staging disk the ingest program extracts metadata information from the data that is then loaded into the IMS metadata database. If the ingest process is successful the data are then moved to the UniTree staging disk for archiving.

Distribution on media is 8mm tape, 4 mm DAT, 9-track 6250 bpi round tape, and on CD-ROM (if available). These media were specified by the EOSDIS project as the standard distribution media that all DAACs must support. Distribution processing is an automated process where the DADS software communicates with the order database, on the IMS system through the GSFC V0 EOSDIS FDDI LAN, to retrieve information needed to fill a users order for data. Scheduling and resource management software is used to control the data distribution (and ingest) activities.

Performance Characteristics

The following performance throughput characteristics have been measured for the DAAC:

Metrum (via UniTree)	1.05 MB/sec	(RSP 2150 drive rate	1.92 MB/sec)
Cygnet (via UniTree)	.5 MB/sec	(GD9001 drive rate	.8 MB/sec)
Magnetic Disks	2.3 MB/sec		
Drives:			
8 mm (tar, 8500 mode)	.42 MB/sec		
4 mm (tar)	.17 MB/sec		
9 track (tar)	.17 MB/sec		

Distribution

The DAAC collects statistics on the types of media and data sets requested by the user. These statistic for the past year are presented in Figure 5. The following data sets are currently available through the DAAC's on-line system:

Data Sets	Transfer Mechanism
AVHRR Pathfinder	Network (Data Transfer Program)
TOVS Pathfinder	Network (FTP)
CZCS	Network (NFS)
UARS	Network (Data Transfer Program)

The DAAC also still supports the DAAC/Land and the DAAC/Climate heritage data sets. Data is available from these data sets by contacting the DAAC's User Support Office.

The current ingest and distribution rates are:

Current Ingest Volume	60 GB/month
Current Distribution Volume	125 GB/month

Future Growth

The following data sets will be available in FY94 through the DAAC's on-line system:

Data Sets	Transfer Mechanism
SeaWiFS	Network (Data Transfer Program)
4D Assimilated	TBD
TOGA-COARE	Network (FTP) and undetermined media
Meteor3/TOMS2	TBD

The future ingest and distribution rates for FY94 are predicted to be:

Future Ingest Volume	510 GB/month
Future Distribution Volume	1800 GB/month

The system hardware will also be upgraded in FY94 as shown in Figure 6. An SGI Challenger L computer will be procured for the DADS and will be used to support distribution (not shown in Figure 7). Six additional disk drives at 2.3 GB each and three 8 mm and one 4 mm drive will also be procured and installed. These disks will be used to expand distribution staging for media and ftp orders. The disks will also be used to expand storage for ingest and UniTree staging areas. The additional drives are required to support an ever increasing data media distribution load.

Conclusion

The GSFC V0 DAAC has been very successful in meeting its goals to date. It has investigated much of the technology that may be needed for the ECS Version 1 System. The careful selection of the hardware and software components of the system has produced a high-quality product that is meeting requirements and current workload needs. The DAAC system will continue a planned expansion to meet anticipated future needs.

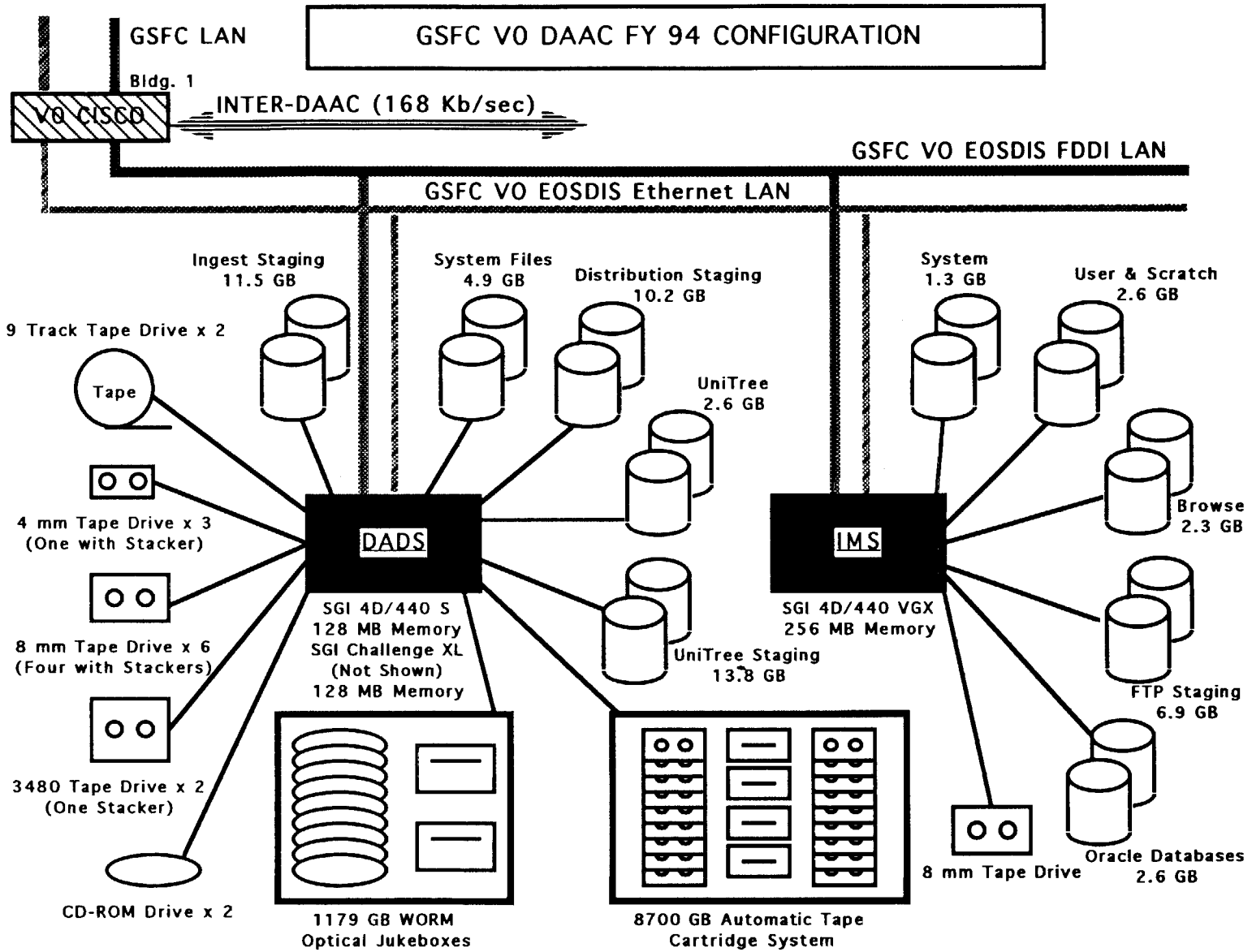
Comparison Of File Management Systems

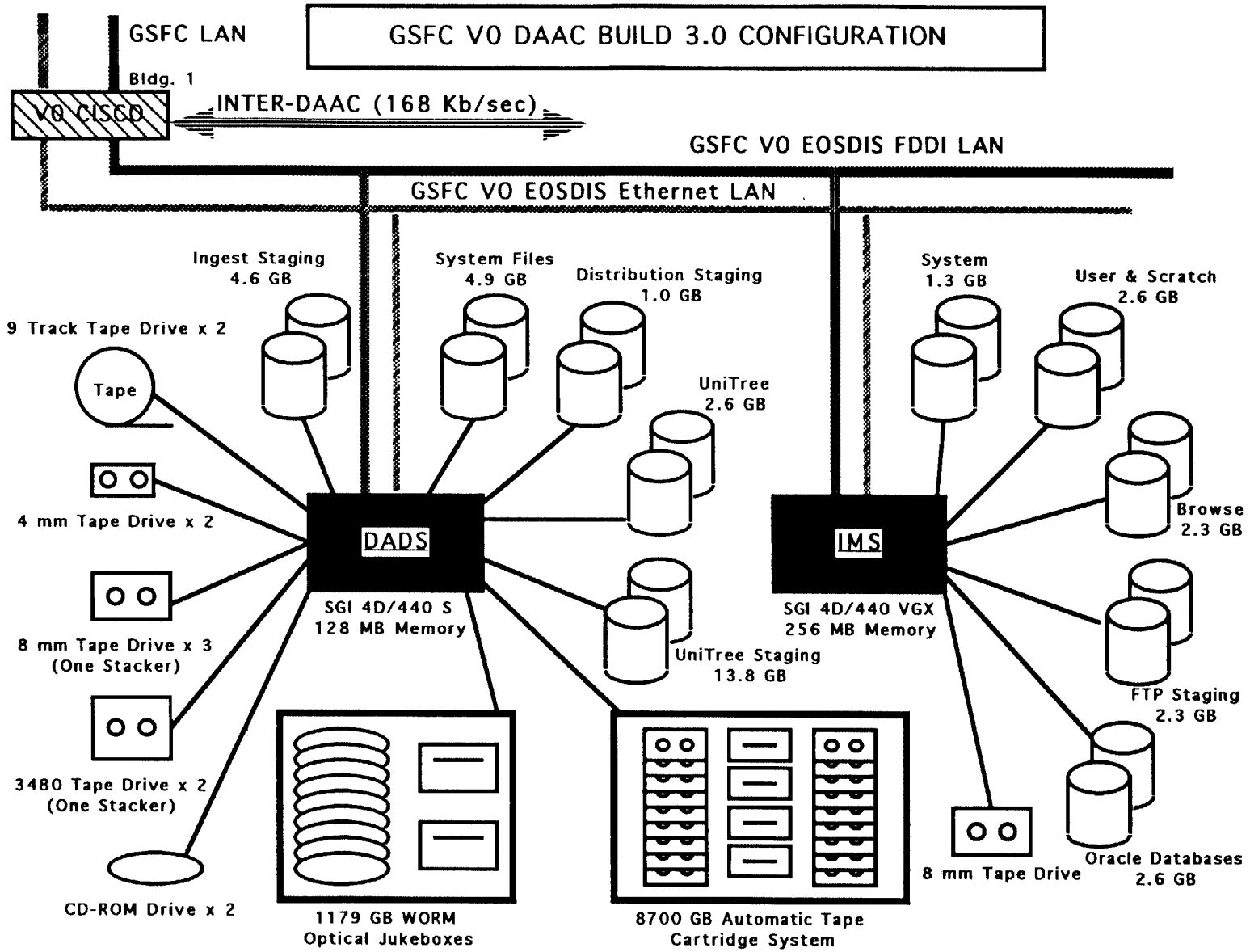
File Management System	FileServ	Optical Archiving System (OAS)	Storage Server (UniTree-Based)	UniTree
Vendor	Sun Coast Softworks, Inc. Cygnet Systems, Inc.	Aquidneck, Inc. Cygnet Systems, Inc.	Loral	Titan (See Note)
Characteristics	S/W Packages	Programmed H/W Device	Programmed H/W Device Transputers for I/O	S/W Package
Supports IEEE Mass Storage System Reference Model	Yes	Yes	Partially	Yes
Supports Cygnet WORM 12" Optical Jukebox	Yes	Yes	Yes	Yes (With Modifications)
Supports Metrum VHS Auto. Tape Cartridge System	No (Optical Systems Only)	No (Optical Systems Only)	Yes (With Modifications)	Yes
Additional Modifications Needed by Vendor	Yes	Yes	Yes	TBD
Device Drivers Included	Yes	Yes	Yes	Yes
File Format	ANSI Tape Label Format	Sequential 8 mm Format (Must Be Read by OAS)	UniTree Unique Format (Must be Read by UniTree)	UniTree Unique Format (Must Be Read by UniTree)
Supports Hierarchical File System With Auto. Migration	No	No	Yes	Yes
Allows Designation of Data Storage Location	Yes	Yes	No	Yes
Supports File Name Database	Partially (Volume Locations Only)	Yes	Yes	Yes
Source Code Available	Yes	No	No	Yes
Amount of Integration Required	Applications Interface to S/W Package	Minimal Emulates an 8 mm Drive	Minimal Used as a Server	Applications Interface to S/W Package
Hosted to SGI Computer Platforms	No	No	Not Required	Yes
Rehost to SGI or Other Modification Costs	$4 K 1 Month Effort	$34 K 2-4 Month Effort	Estimate in Progress (Metrum Mods Extensive)	$TBS

Comparison Of File Management Systems

PERFORMANCE (Jukebox Only)	Writes Equal to OAS Faster Reads/Direct Access	Writes Equal to FileServ Slower Reads/Sequential Access	UniTree Overhead; Slower Than FileServ and OAS Faster Than UniTree/Titan	UniTree Overhead; Slower Than FileServ and OAS UniTree = 150K SLOC
COST	$22.5 K	$48.5 K	$TBS (Estimated at 100 - $120 K)	$TBS (Estimated at $100 - $250 K)
RELIABILITY	Extensively Implemented on Jukebox Systems	Extensively Implemented on Jukebox Systems	Limited Implementation	One Operable System with SGI & Metrum

Storage Systems	Storage Capacity GB	Media Capacity GB	Transfer Rate MB/sec	Access Time secs	Floor Space sq. ft.	Cost/TB $	Total Cost $
Cygnet 1803 WORM 12" Optical Jukebox (SONY Drives)	858	6.55	0.48	7	83.44	$256,410	$220,000
Cygnet 1803 WORM 12" Optical Jukebox (ATG Drives)	1179	9	1.5	7	83.44	$174,724	$206,000
Kodak WORM 14" Optical Jukebox (Kodak Drives)	1020	10.2	1	14	117.9	$294,118	$300,000
Metrum RSS-600 VHS Automatic Tape Cartridge System	8700	14.5	2	98	111	$37,931	$330,000
E-MASS 200 (D2) 19mm Data Tower	6450	25	15	128	201.33	$102,326	$660,000
Storage Tek 4400 Silos	1200	0.2	4	53	476	$541,667	$650,000
SONY DMS-600M (D1) 19mm Digital Mass Storage System	27720	41.25	32	188	130.85	$32,468	$900,000


GSFC LAN
GSFC V0 DAAC FY 94 CONFIGURATION
Bldg. 1
V0 CISCO
INTER-DAAC (168 Kb/sec)
GSFC V0 EOSDIS FDDI LAN
GSFC V0 EOSDIS Ethernet LAN
Ingest Staging
11.5 GB
System Files
4.9 GB
Distribution Staging
10.2 GB
System
1.3 GB
User & Scratch
2.6 GB
9 Track Tape Drive x 2
Tape
UniTree
2.6 GB
4 mm Tape Drive x 3
(One with Stacker)
DADS
IMS
Browse
2.3 GB
8 mm Tape Drive x 6
(Four with Stackers)
SGI 4D/440 S
128 MB Memory
SGI Challenge XL
(Not Shown)
128 MB Memory
SGI 4D/440 VGX
256 MB Memory
UniTree Staging
13.8 GB
FTP Staging
6.9 GB
3480 Tape Drive x 2
(One Stacker)
CD-ROM Drive x 2
1179 GB WORM
Optical Jukeboxes
8700 GB Automatic Tape
Cartridge System
8 mm Tape Drive
Oracle Databases
2.6 GB


GSFC V0 DAAC BUILD 3.0 CONFIGURATION
GSFC LAN
Bldg. 1
V0 CISCO
INTER-DAAC (168 Kb/sec)
GSFC V0 EOSDIS FDDI LAN
GSFC V0 EOSDIS Ethernet LAN
Ingest Staging
4.6 GB
System Files
4.9 GB
Distribution Staging
1.0 GB
System
1.3 GB
User & Scratch
2.6 GB
9 Track Tape Drive x 2
Tape
UniTree
2.6 GB
4 mm Tape Drive x 2
DADS
IMS
Browse
2.3 GB
8 mm Tape Drive x 3
(One Stacker)
SGI 4D/440 S
128 MB Memory
SGI 4D/440 VGX
256 MB Memory
UniTree Staging
13.8 GB
FTP Staging
2.3 GB
3480 Tape Drive x 2
(One Stacker)
CD-ROM Drive x 2
1179 GB WORM
Optical Jukeboxes
8700 GB Automatic Tape
Cartridge System
8 mm Tape Drive
Oracle Databases
2.6 GB