Success Story

DATA LIFECYCLE MANAGEMENT

SGI DATA LIFECYCLE MANAGEMENT AT NASA AMES:

Efficient, Cost-effective Storage Solutions for Massive Data

"We wanted something that gave us a growth path and was easy to integrate and maintain. DMF was very successful for us."

– Harper Pryor, Senior Systems Engineer, GMAO

A countrywide user base of up to 800 scientists accesses the immense compute resources at NASA Advanced Supercomputing (NAS) facility at NASA Ames Research Center at Moffett Field, CA, which had the only 1024 SGI® Origin® 3800 system and a 512-processor SGI® Altix® compute server. Project Columbia, over a 4.5 month period ending in October, has twenty 512 SGI Altix nodes installed and is projected to be the world's fastest supercomputer. NAS users write terabytes of data every day, and with Project Columbia are expected to write on average 11TB a day. As stored data has grown from 100 terabytes to over a petabyte, the NAS team has minimized the cost and complexity of data growth using SGI® InfiniteStorage Data Lifecycle Management solution DataMigration Facility (DMF).

Grand Challenge Science and Mushrooming Data

"One of the reasons we have this facility is to do grand challenge science," says Alan Powers, High End Computing Lead at NAS. On a typical day, 200-300 users are logged on to NAS resources, and their work is multidisciplinary — planet modeling, ocean modeling, nanotechnology, astrobiology. Not all are involved in earth and space sciences, however. NASA scientists working with Dr. Michael DeBakey of Houston's Baylor College of Medicine ran computational fluid dynamics models to develop a miniaturized ventricular-assist device (VAD) called the NASA/DeBakey heart pump. The device, designed to aid heart patients, received NASA's Commercial Invention of the Year award.

Roughly 80% of the users of the NAS facility access its resources from other NASA research centers — Glenn, Goddard, Jet Propulsion Laboratory, Johnson, Langley, and Marshall. Industrial users also have access under government agreements. They work on Chapman, NAS's one-of-a-kind, single-system-image, 1024-processor SGI Origin 3800 system with 256GB of memory and 20TB of RAID storage. These users also use another one-of-a-kind system — a 512-processor SGI Altix compute server. "It's the largest single system image Linux box in the world," says Powers. "It's doing super science for us."

As of May 2004, the data access for most users' day-to-day work is handled by a 20-processor SGI® Origin® 3000 server with 10GB of memory and 20TB of RAID storage, backed up by a 10-processor Origin® 2800 system. The Origin 3000 server writes a terabyte of data on a typical day, and may write 6TB on exceptional days. Because of Project Columbia archive requirements, the main DMF system is planned to be upgraded to a 96 processor Origin 3000 with 48 GB of memory and over 200 TB of RAID storage. The combined output of all the facility's systems is mind-bending.

"In June 1999 we were writing about 80GB a day total using the previous hierarchical storage management (HSM) software running on two Convex systems," says Powers. "In January 2004 we had a single machine writing about a terabyte a day with a peak of four terabytes, and another machine writing 800GB. We keep 700-800TB of data nearline." NAS uses SGI storage technology to manage this massive output.

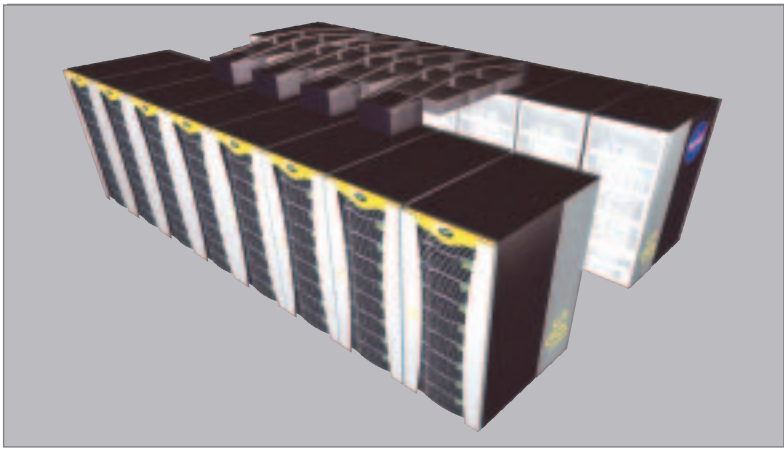
DMF: Keeping Ahead of a Data Avalanche

In 1999 NASA's Data Assimilation Office (DAO), a major user of NAS resources, was creating too much data to manage cost-effectively with traditional approaches. They needed a solution that would allow for phenomenal growth and fast data access, while minimizing the cost and complexity of capacity and management. With 100TB already under management, DAO considered the previous HSM, as well as a number of vendor offerings to address these issues, and asked Powers to install DMF to manage the DAO data.

"The homegrown product was difficult to maintain and to keep current with new storage technologies," says Harper Pryor, Senior Systems Engineer, GMAO (Goddard Modeling and Assimilation Office, formerly DAO). "We wanted something that gave us a growth path and was easy to integrate and maintain. DMF was very successful for us."

"One major consideration was the ease with which DMF could be implemented," says Powers. "We had to move quickly. We had this Y2K problem, and we needed a solution NOW. Whatever we picked, we had to get it in quickly and get it up easily, and it had to mesh with DAO's back-end storage resources.

"We also wanted minimal staffing requirements. We had the equivalent of four fulltime developers and four fulltime support people caring for the homegrown system, which was running with 100TB of storage. Today, with DMF and about eight times the storage requirements, we only need the equivalent of one fulltime person, and we have faster access."



"In June 1999 we were writing about 80GB a day total using the previous Hierarchical Storage Management (HSM) software running on two Convex systems. In January 2004 we had a single machine writing about a terabyte a day with a peak of four terabytes, and another machine writing 800GB. We keep 700-800TB of data nearline."

– Alan Powers, High End Computing Lead, NAS

"The biggest thing about any DLM system is reliability, reliability, reliability. You don't want to lose any data. That's really what drove us to implement DMF."

– Alan Powers, High End Computing Lead, NAS

DATA LIFECYCLE MANAGEMENT

A Flexible, Efficient Storage Solution

DMF is currently running on two SGI servers at NAS — the 20-processor Origin end-user system and the 10-processor Origin backup system. The 20-processor system is connected to 20TB of RAID storage on TP9400 and TP9500 disk arrays. Ten STK 9940B and 10 STK 9840A tape drives provide massive nearline storage resources. The 10-processor system has 4TB of attached storage and 10 STK 9940 tape drives.

"DMF allows us to keep the most recent copy, or the most relevant data, on disk, so users can pull that information off quickly and do their runs on the high performance compute engines," says Powers. "If DMF didn't exist, we'd have to buy RAID. And RAID compared to tape drives is very expensive.

"We have a 20TB cache, which is enough for four or five days' activity. If the data fills the cache sooner than that, the older data blocks will be released. If the users need to pull it off tape, the data typically starts flowing in about a minute. So it's very transparent to them that it's on tape. And it's a huge saving in cost and response time."

DMF's flexibility enabled NAS to tailor storage type by file size, sending smaller files (1-100MB) to quick-seek dual-reel tape drives, where they can be read very efficiently, and sending larger files to larger capacity tape drives that can transfer data more quickly. Files less than a megabyte — usually source code or scripts — stay on the local RAID cache so they can be pulled up quickly.

NAS chose DMF because of its flexibility, low maintenance requirements, and ease of integration. But there was another basic factor. "The biggest thing about any data lifecycle management (DLM) system is reliability, reliability, reliability," says Powers. "You don't want to lose any data. That's really what drove us to implement DMF."

HOW DO YOU MOVE 500 Terabytes of Live Data Across the Country?

In 2003 the decision was made to relocate DAO's entire computational and storage workload from the NAS facility at Moffett Field to the NASA Center for Computational Sciences (NCCS) at the Goddard Space Flight Center in Greenbelt, Maryland.

The requirements for this process were deceptively simple: make the move completely transparent to users on both coasts. No loss of data; no interruption of service; no change in ownership, account privileges or restrictions. Of the 1.3 petabytes of NAS data stored online and under DMF, this project would be moving 500TB from coast to coast.

The first step was the creation of a NAS-like environment at Goddard, implementing the technologies and protocols that were successful at Ames. Goddard already had a 512-processor SGI Origin 3000 server that was split operationally into a 384-processor system and a 128-processor system. They upgraded an Origin 2000 system to a 96-processor Origin 3000 server. All three systems were used as compute servers. Goddard also upgraded a second Origin 2000 system to make it a 32-processor Origin 3000 DMF server with six TP9840 and 10 TP9940 tape drives.

NAS and DAO worked out a plan for a staged network transfer of the data. Then, at the 11th hour, they were told that there wasn't time for a network transfer. Since the network could not be used for the transferring of the data, an extra copy of data was written to tape media and shipped cross-country. NASA enlisted SGI's help to manage the transfer. NAS made incremental snapshots of the 500TB DAO data and wrote them to tape. SGI sent five shipments of tapes to Goddard.

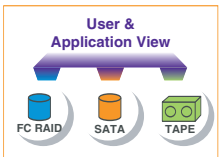
"It was enormous," says Pryor. "How do you make snapshots of the DAO database, copy all the data (while we were still making data at NAS), ship it to Goddard, get the data on line, then come back and take the next snapshot? To further complicate the situation, as we were writing less in California and more in Maryland, we shipped tape drives from NAS to NCCS. It was very interesting to see SGI carefully decommissioning the drives so they could send them to Goddard."

At the Goddard end, SGI engineers used a DMF tool called dmaudit to validate data as it was received and written onto the storage systems at NCCS. They used DMF to keep the databases in sync during the transfer.

"We sent the metadata along with the data," says Powers, "so they had enough information to validate the process. DMF had the smarts to let us downscale the use here as we upscaled the use of the one at Goddard." Having DMF on both coasts enabled NCCS to organize the data at Goddard the way it was organized at NAS. Researchers use the same calls to get their files, and have the same directory structures, allowing them to focus on science, not computer science.

"If we had moved it to another kind of DLM product, it would have been a nightmare," says Pryor. "It would have taken much longer. Some data would probably have been lost. This was not something that we could have done without SGI's help."

"When you're moving machines and data around, things can go wrong," says Powers. "The great thing about this whole process was that these guys from SGI knew that if something went wrong, they would make it right. That's an extremely important thing to me — knowing that your vendor has your best interests at heart. Their attitude is, 'Let's work together and solve this problem.'"



DMF Server virtualizes storage assets, automatically managing data by user policies to deliver near-infinite, easy to manage, lower cost capacity

Corporate Office
1500 Crittenden Lane
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America +1 800.800.7441
Latin America +55 11.5509.1455
Europe +44 118.925.7500
Japan +81 3.5488.1811
Asia Pacific +1 650.933.3000

3705 [11.2004]

J14712