ACTIVE *Archive*

CHALLENGE:

A research community, comprised of about 3,000 users accessing and storing large volumes of content, needed a better way to manage its data and accommodate future growth.

SOLUTION:

- *Active archive providing a file-based system to manage tape and disk (2 PB capacity)*
- *Spectra Logic T950 library*
- *SGI Data Migration Facility (DMF)*
- *IBM Tivoli for host and user backups*

RESULTS/BENEFITS:

- *Significant reduction in CapEx and OpEx*
- *Highly scalable to handle future growth*
- *Reliable data accessibility through proactive integrity monitoring*

Why Active Archive?

- *Superior expandability and scalability*
- *High data reliability*
- *Fast, easy data accessibility and portability*
- *High data storage density*

CASE STUDY

University of Minnesota's Supercomputing Institute for Advanced Computational Research

Minnesota Supercomputing Institute Manages High Volume Data Growth With Active Archive

Research Institute Selects Active Archive To Support High Volume Supercomputing and Scientific Computing Environment

The Supercomputing Institute for Advanced Computational Research is an interdisciplinary research program spanning all colleges of the University of Minnesota. The institute provides supercomputing resources and user support to faculty and their research groups. It is a key program in the University's broad-based digital technology effort, supporting collaborative research within the University and the state of Minnesota as well as education related to supercomputing and scientific computing for undergraduates and graduates.

The Minnesota Supercomputing Institute (MSI) was awarded a National Institutes of Health (NIH) grant for research that required a large data archive to support its research.

"Receiving the NIH grant was a huge win for us. We have over 3,000 users accessing and storing a variety of data – anything from computational fluid dynamics to genome specimens to weather modeling," said Jeff McDonald, Assistant Director for HPC Operations at Minnesota Supercomputing Institute. "The NIH funding helped support our growth and allowed us to find a solution that best fit the needs of our user community."

Replacing an Antiquated System
MSI's existing storage system was more than a decade old and needed to be replaced. The MSI IT operations group reviewed its current environment and recognized that keeping all of its data on disk would be cost prohibitive. At the same time, it was
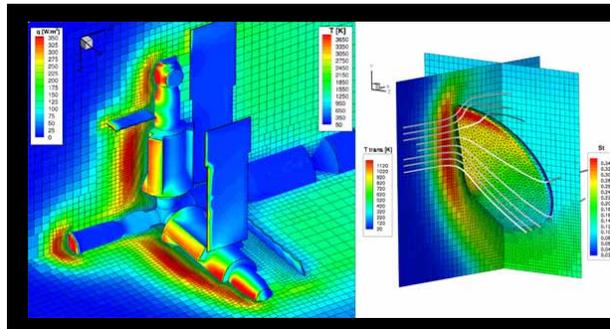
Shown at left: Imagery from the MSI website featuring research on molecular simulation of hypersonic flow.

> **"We manage and store large volumes of data so it was important to select a system that is scalable and can grow as our capacity needs increase," said McDonald. "Equally important, was that the solution bring a high level of reliability."**

University of Minnesota Research Institute

incurring significant service maintenance fees and the proprietary nature of its tape technology limited its ability to integrate new solutions.

"We needed an enterprise-class storage solution that could not only backup and archive the data for the NIH project, but be partitioned to manage the backup data for nearly 3,000 MSI accounts and provide fast and easy data access," said McDonald.

Efficient Storage and Access to Data Through Active Archive

Due to expected high volume data growth, MSI needed a scalable research storage and archive system. MSI researched the capabilities of various storage solutions and realized an active archive to tape and disk solution would best suit its needs. It selected the Spectra Logic T950 enterprise tape library, integrated it with MSI's existing disk system and added SGI DMF to manage the archive storage system. By applying a file system to expand over disk and tape library technologies, MSI can now enable a persistent view of the data in its archives. In addition, the new system guarantees that the researchers will have reliable, online and efficient access to their archived data.

In addition to accessibility, the active archive is easily expandable and provides the high scalability MSI requires to store large quantities of critical data in support of both the NIH project and the University's data.

Repository for Research Data

MSI's active archive system is partitioned into two sections. One side can hold up to 700 TB of data and is comprised of a data migration facility with additional disk drives and a file system for accessing data. It is used for migrating aged data to tape as well as backing up the NIH project systems. The other partitioned side can hold up to 1 PB of data and is comprised of an IBM Tivoli suite that sends files to the tape library for archiving and manages the MSI's backup data.

Researchers are allocated a specific amount of storage space on the system and data in the active archive is backed up for a 30-day period and a full system backup is performed and retained for an approximate 10-year period.

Explosive Data Growth

MSI is experiencing exponential growth in the amount of data it stores each year. At the initial implementation of its new active archive solution, the tape library contained approximately 200 TB of data. Now, two years later, it has grown to 400 TB and is expected to increase capacity to more than 1 PB within the next six months as additional online storage is added.

Unlike MSI's previous tape library, which was limited and not expandable, the Spectra Logic T950 enterprise tape library is capable of scaling to much larger capacities, and it provides industry leading density that minimizes floor space requirements. In addition, the increased capacity capabilities of LTO tapes have allowed MSI to store more data on fewer tapes.

Scalable and Reliable

"We manage and store large volumes of data so it was important to select a system that is scalable and can grow as our capacity needs increase," said McDonald. "Equally important, was that the solution bring a high level of reliability."

Since the initial deployment of the active archive solution, MSI has purchased additional LTO-5 tape media, added an additional T950 Media Expansion Frame and a new disk storage system – all of which have been seamlessly integrated within the expanded solution.

Moving Forward

MSI plans to purchase an additional 1 PB of storage that will support LTO-5 tape drives and can be upgraded to LTO-6 in the future. In addition, MSI is planning to incorporate and leverage LTFS tapes to allow users to move data between heterogeneous systems. This will enable researchers to have local access to data for those who opt to take their tapes with them.
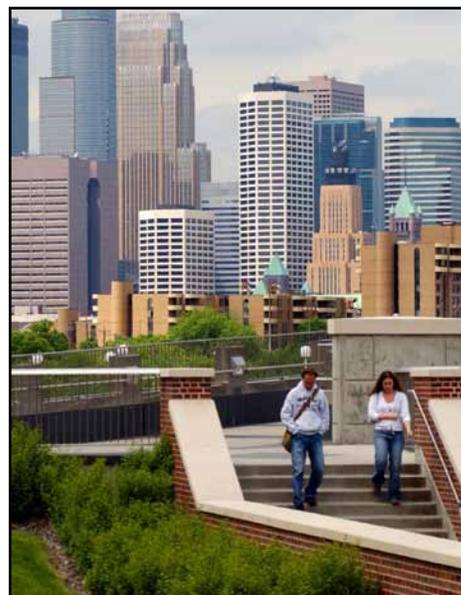
U of M students on campus with skyline of Minneapolis behind.