

Data Management Software

Managing Access to Petabytes of Data

Every day, five of NASA's initial Earth Observing System satellites beam down terabytes of scientific data about the earth, oceans, and atmosphere. This data is processed and made available to climate researchers around the world using NASA's Earth Observing System Data and Information System (EOSDIS) Core System—one of the most ambitious data gathering and distribution systems ever conceived. The storage of this irreplaceable information, which spans multiple media types and locations, is entrusted to Quantum's data management software.

ESCALATING DATA GROWTH

Since going online in 1999, NASA's EOSDIS has become one of the world's largest and most active repositories of data. In 2001, EOSDIS filled more than 3.4 million requests for data from 5,000 different users, ranging from international climatologists to high school students. Every day, the EOSDIS system receives 3 terabytes of new data from its five satellites, and distributes 2 terabytes of data from its four linked data centers to researchers around the world. Currently, EOSDIS provides hundreds of thousands of global users access to more than 20 million files which together contain over a petabyte of information. This data bank is the richest set of information ever collected on the Earth's climate. By the end of the decade, the system is expected to hold more than 10 petabytes.

“...we needed a robust storage system that could scale to handle massive data sets stored on multiple types of media. We also needed a stable solution that could support our 24/7 operation.”

EOSDIS Project Technical Director, Raytheon

The core components of EOSDIS were developed by Raytheon Imagery and Geospatial Systems. “This satellite data is a national resource that will provide tremendous insight into how the earth is changing,” says Raytheon Project Technical Director, who has been with the project since it began almost 10

years ago. “Storing it, organizing it, and making it available to researchers are critical tasks and a huge challenge.”

Setting up a storage system of this magnitude presented major challenges. The sheer scale of the task meant that automated tape libraries would be the primary storage. The EOSDIS architects needed to structure the system so that applications would be able to write data and read it from the libraries. These applications included both the software that processes the data and the database that lists the data sets available to researchers. “When we designed EOSDIS, we knew we needed a robust storage system that could scale to handle massive data sets stored on multiple types of media,” Raytheon Project Technical Director explains. “We also needed a stable solution that could support our 24/7 operation.”

HIGH-PERFORMANCE DATA ACCESS WITH DISK AND TAPE

Key to the EOSDIS process was developing a system for managing high-speed disk cache and for keeping file metadata persistently available on disk for rapid online searches. Above all, the system needed to be able to manage millions of files, petabytes of data, and millions of transactions a year in a resilient 24x7 operational environment.

The storage solution which NASA selected combined the strengths of disk, tape, and data management software. The data is managed through a combination of disk storage, used primarily as caching, and several high-capacity automated tape



“...we needed a robust storage system that could scale to handle massive data sets stored on multiple types of media. We also needed a stable solution that could support our 24/7 operation.”

EOSDIS Project Technical Director, Raytheon

SOLUTION OVERVIEW

- Application: Automated storing and accessing of NASA satellite data
- Quantum's Data Management Software
- Tape Storage: Automated tape libraries
- Server and Storage Environment: SGI IRIX servers

KEY BENEFITS

- Vendor support
- High reliability and scalability
- Support of multiple media types
- High speed search and retrieval capabilities

systems in four separate data centers. Quantum's data management software ties the whole system together and enables both EOSDIS staff and users to access it effectively.

Quantum's data management software transforms libraries—typically considered "offline" storage—into truly online direct-access mass storage, enabling users, administrators, and applications to effortlessly archive terabytes or petabytes of data. Quantum's data management software presents automated tape libraries as one device and one mount point, through a standard UNIX file system interface consisting of directories and files. Users or applications can access files residing on tape libraries in the same way they would with disk devices. Quantum's data management software manages disk cache, maintains metadata online in highly available server clusters, and scales to manage an effectively unlimited number of files.

Overall commitment to this project from the primary vendors has been one of the keys to its success. A project like this only runs smoothly when all of the lead vendors work together to ensure full compatibility across multiple product generations.

EOSDIS Project Technical Director, Raytheon

"The completed system lets us get the most efficient use possible out of our storage resources. Today, we make 1.8 petabytes of data and about 22 million files available to the public," Raytheon's Project Technical Director continued. "It gives us a high-performance, very resilient storage environment, easy access to the data from all our applications, and it gives us massive resource leverage. All that is seen by the end-users who request data in the database of available data sets or images. All of the file retrieval,

staging to disk cache, and delivery as a file or duplicate set of media happens in the background." To the users who access the data through the EOSDIS website, data appears as though it were stored in a single database.

Support from Quantum that helped with system architecture and provided ongoing service was also critical to the project's success. Raytheon's Project Technical Director concludes, "Overall commitment to this project from the primary vendors has been one of the keys to its success. A project like this only runs smoothly when all of the lead vendors work together to ensure full compatibility across multiple product generations."

"The completed system lets us get the most efficient use possible out of our storage resources...It gives us a high-performance, very resilient storage environment, easy access to the data from all our applications, and it gives us massive resource leverage."

EOSDIS Project Technical Director,
Raytheon

ABOUT NASA'S EOS

The Earth Observing System (EOS) is the centerpiece of NASA's Earth Science Enterprise. It consists of a science component and a data system, supporting a coordinated series of polar-orbiting and low-inclination satellites for long-term global observations of the land surface, biosphere, solid Earth, atmosphere, and oceans. By enabling improved understanding of the Earth as an integrated system, the EOS program has benefits for us all. For more details, please see <http://eosps0.gsfc.nasa.gov>

To contact your local sales office, please visit www.quantum.com

Quantum
BE CERTAIN

©2014 Quantum Corporation (NYSE:QTM). All rights reserved. Quantum and the Quantum logo are registered trademarks of Quantum Corporation and its affiliates in the United States and/or other countries. All other trademarks are the property of their respective owners.

CS00087A-v03