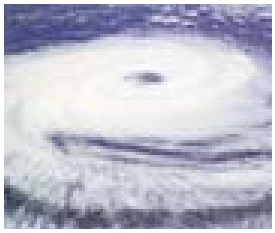
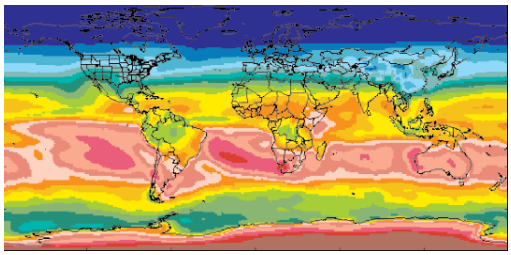
WINCHESTERSYSTEMS®
Purpose-Built Storage

NASA's Earth Science Community Gets Fast, Reliable Access to Growing Petabyte of Data

Using FlashDisk® with RAID 6



The Atmospheric Science Data Center (ASDC) was established in 1991 as a division of the NASA Langley Research Center to support NASA's Earth Science Enterprise and the U.S. Global Change Research Program. ASDC is responsible for the processing, archival, and distribution of NASA Earth Science data in the areas of radiation budget, clouds, aerosols, and tropospheric chemistry. After capturing the data through NASA's satellite system, ASDC makes the data available to the public (with its primary audience consisting of scientists, researchers, and students) through its website, ***http://EOSweb.larc.nasa.gov***. Last year, ASDC distributed over 275 terabytes of free data to more than 34,000 customers worldwide.



ASDC's current data archive volume has reached 1.5 petabytes, and the volume increases on a daily basis. The center's overarching goal is to make the data available in a fast, reliable format to its end users on a moment's notice. Systems Engineer David Cordner explains how ASDC's technological and data needs have evolved over the past ten years. The first big challenge was figuring out how to cache, store, and process all of the satellite data that was streaming in on a regular basis. The center accomplished this by dedicating a big machine room where large tape silos of archival data could be stored. Tape drives were chosen for economical reasons as disk archiving large volumes of data would be cost prohibitive. The data is now captured, processed, and distributed using Linux, open source, and high-end server and data storage systems from Big Iron vendors such as EMC and SGI.

Over the years, the next major challenge was learning how to get the large and growing files out to the scientific community, which Cordner describes as the "next generation systems distribution and access challenge." This situation reached a critical turning point two years ago as access demands increased dramatically. This accelerated access rate compounded the data distribution issue to a point where the systems engineers (a team of five) knew the time had come to invest in disk storage space that had two critical features: (1) reliable storage and (2) rapid access. The team was ready to consider more cost-effective, off-the-shelf commercial equipment from a landscape of data storage vendors ranging from Big Iron companies to independent manufacturers.

Winchester Systems Responds to an RFP

The center went out to bid, issuing an RFP with these specifications: 30 ter-

abytes of SATA storage with a Fibre Channel (FC) connection and RAID 6 data protection. One of the vendors to respond was Winchester Systems. The engineering team knew they had a winner when Winchester Systems recommended its FlashDisk SATA 4500 series. Cordner comments: "Winchester Systems price performance tipped the scale in its favor. We were looking for Fibre Channel disk arrays that offered reliable, dense, and high-speed performance at an economical price, and that's what Winchester Systems delivered."

Leveraging FlashDisk for Caching and Rapid Access

It took only a few hours for Winchester Systems to install the SA4500 series for a total of 36 terabytes. While each unit contains 12 terabytes, only 10 terabytes are used, reserving the remaining 2 terabytes for providing Raid 6 reliability.

Cordner describes the dual-purpose configuration, emphasizing that disk caching has become an increasingly critical function to meet the demands of the growing archive system. "We use a variety of disk systems from many vendors. The data from our partners who provide the satellite information to us is initially staged onto a set of disks that are not Winchester Systems. The archive system then offloads the data from that drop box area and stages it on the Winchester Systems disks, which are configured in a StorNext SAN architecture for integrity checking, manipulation, and cataloging. When the archive is finished verifying that the data is good, StorNext places the data on a tape system in a silo that has robotic control over the tape drives."

The Winchester Systems/StorNext configuration allows multiple archiving processors to access and manipulate the data in a high performance environment. The processes that occur while the data is on the Winchester Systems disks are critical because they verify that the data files are accurate and complete. The RAID 6 environment prevents data from being lost or corrupted. Should a disk fail, as it surely does, Winchester Systems disk arrays continue to perform without losing or corrupting the data.

> "Winchester Systems price performance tipped the scale in its favor."
>
> "We've got a good value for the purchase, and that's how we measure our ROI."
>
> David Cordner, Systems Engineer

ASDC also sets aside a few terabytes of Winchester disk space outside of the SAN architecture for databases that keep track of data location, content, and availability. The robust storage capability of the Winchester Systems disk arrays is critical for maintaining the integrity of all of those databases. Again, the center relies on the Winchester Systems' ability to continue working through disk failures to keep the archive databases intact.

Winchester Systems Keeps the Center Running Smoothly

The Winchester Systems solution has had a positive impact on ASDC in that the day-to-day operations run without a hiccup, which ultimately means the center provides its users with outstanding customer service. Since FlashDisk SATA had been installed nearly two years ago, ASDC has had no downtime due to data failures. Cordner remarks: "The greatest compliment we can give the folks at Winchester Systems is that we don't have to deal with them much. That's just the way we like it."

The Best Value for the Money

Cordner is quick to note that the center's return on investment is measured by how much money the center doesn't have to pay keeping the system running smoothly. Cordner concludes: "We don't have to spend money care feeding the disk cache equipment. It sits there and runs without any manpower interference. We've got a good value for the purchase, and that's how we measure our ROI."

NASA images courtesy of:

http://EOSweb.larc.nasa.gov

http://visibleearth.nasa.gov

101 Billerica Ave., Bldg. 5, Billerica, MA 01862 • 800-325-3700 • 781-265-0200 • fax: 781-265-0201 • www.winsys.com