PANASAS®

CUSTOMER SUCCESS

Summary

Challenge CACR's limited staff and budget called for a reliable, cost effective, and easy to manage high performance storage solution that would keep the university at the forefront of computational science and engineering.

Solution

• Panasas ActiveStor scale-out NAS appliances with the integrated PanFS parallel file system

• Comprehensive Panasas technical support

Results

• Greater performance, extreme scalability, and enhanced reliability

• Simple and care-free storage system management, allowing CACR staff to focus on supporting research instead of IT infrastructure

• Easy upgrade path
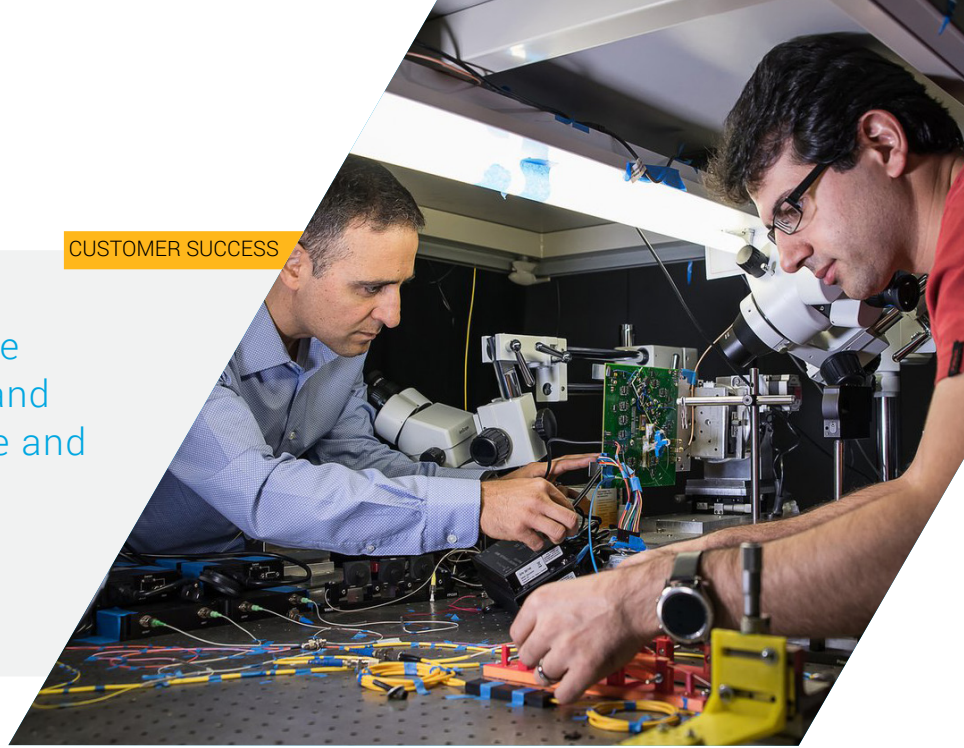
• Superior support and service

California Institute of Technology
Uses Panasas® ActiveStor™ to Accelerate Scientific Discovery

The Center for Advanced Computing Research (CACR) at the California Institute of Technology (Caltech) operates large-scale computing facilities for numerous campus research groups who have big data design and discovery requirements. CACR has a full-time, 25-person staff with expertise in data-intensive scientific discovery, physics- based simulation, scientific software engineering, visualization, and novel computer architectures. It provides technical assistance such as porting code, designing and specifying resources, and advanced IT integration.

Caltech

![PANASAS]

"We have tremendous confidence in the manageability, reliability, and performance of Panasas storage and in the integrity of the company."

Sharon Brunett
Senior Scientist
California Institute of Technology

The Center for Predictive Modeling and Simulation In her role as a researcher, Brunett is a PSAAP software integration team member working on operational, performance, and scalability aspects of the Center's simulations. An overarching objective of the PSAAP program is to predict hypervelocity impact phenomena with quantified uncertainties. Variables include geometry (thickness and obliquity), layout (target plates), impact velocities (2 – 10 Km/s), impactor materials (steel and nylon), and target materials (Fe, Ta). Simulations include multi-scale parallel OTM runs, producing large numbers of datasets. One of the more challenging aspects of the simulations from an I/O perspective, is archiving and curating the results of the simulations. A wide variety of file access and updating methods are used, requiring that the underlying parallel file system be robust.

"We tried other storage solutions with parallel file systems in the past. Some were just not worth the continued investment. Users complained about slow response times and the file systems were a hassle to manage. The Panasas file system just works. The hardware and the software are well tested before being integrated into our environment. We wanted a solution that was big enough and fast enough for our big data workloads, and one that wouldn't get in our way. We wanted, as much as possible, to be able to take file system support and usability for granted. Panasas does a darn good job of letting us do that"

Sharon Brunett
Senior Scientist
California Institute of Technology

Starting with CACR as a systems support team member some twenty years ago, Sharon Brunett now manages the CACR computing facility and systems staff. In addition, she is a researcher for Caltech's Predictive Science Academic Alliance Program (PSAAP) which supports computing projects for the U.S. Department of Energy. As CACR facilities and operations manager, Brunett oversees a broad collection of computer equipment, including multiple computing clusters from various manufacturers, high performance and commodity networks, and storage systems.
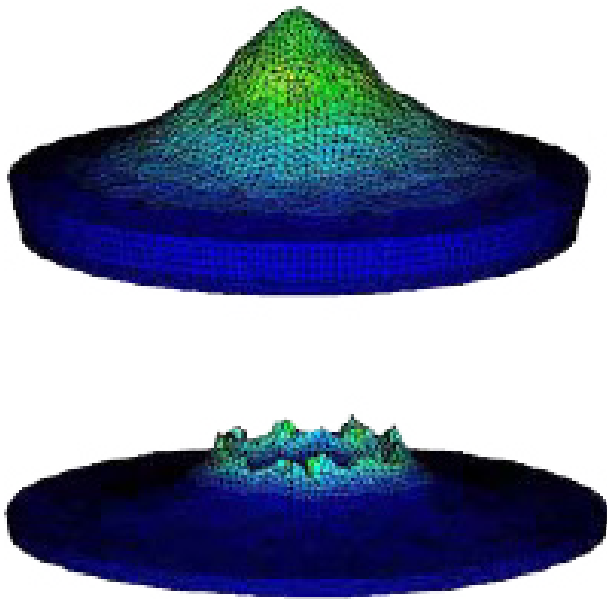
When a new HPC project for the Department of Energy was initiated at CACR that required a parallel file system, Brunett reviewed the storage and file system options that were available. Two of those options, Lustre and PVFS, were free, open source products. "Free isn't always quite so free when you factor in the impact on users, integration, and support challenges in your particular environment," noted Brunett. CACR experienced a loss of non-critical data and endured problematic system administration tools with limited vendor support with some of its prior storage systems. Brunett was understandably concerned with manageability, reliability, and stability of both the parallel file system and its underlying storage hardware when weighing new storage solution options.

The Panasas Solution
Panasas offered a high performance storage solution with a fully integrated parallel file system. Panasas ActiveStor proved to be extremely reliable and fully supported. Brunett noted that glowing Panasas customer references, including those from Los Alamos National Laboratory and others,

![PANASAS]

"The Panasas file system just works. The hardware and the software are well tested before being integrated into our environment."

Sharon Brunett
Senior Scientist
California Institute of Technology

werea critical factor in CACR's decision to purchase Panasas storage. "Dataintegrity, system reliability, and support were the motivating factors to deploy a production quality parallel file system from a company focusing solely on high performance storage, rather than a roll-your-own solution," said Brunett. "We've been pleased ever since.

We have limited funds and a limited staff, so we try to select vendors and products that meet the needs of our users and that therefore make our lives easier. Managing Panasas storage is straightforward. When the primary storage administrator is unavailable, even I can step in and help support the Panasas system, should a drive fail."

Summary

Caltech chose Panasas for ease of use, data integrity, and service – and ActiveStor has consistently delivered for over four years. Panasas is now the largest capacity, high performance parallel storage system at CACR. The current deployment includes five Panasas ActiveStor shelves totaling 180TB of capacity. The university is currently upgrading two legacy ActiveStor 7 shelves to faster, higher capacity ActiveStor 11.

ActiveStor has been an extremely reliable parallel storage platform for CACR. It has greatly reduced tedious storage support efforts, eliminated many of the administration hassles of file system management, and removed user complaints about lackluster performance and application response times.