

# MINES ParisTech Analyzes Rich Data Assets to Understand and Improve Materials

Leading French Engineering School Performs Cutting-Edge Industry Research Using High-Performance Panasas ActiveStor®

## SUMMARY

### CUSTOMER

MINES ParisTech

### SIZE

2,000 employees across the organization, 200 at the Materials Research Center

### LOCATION

Evry, France

### INDUSTRY

Higher education and academic research

### CHALLENGE

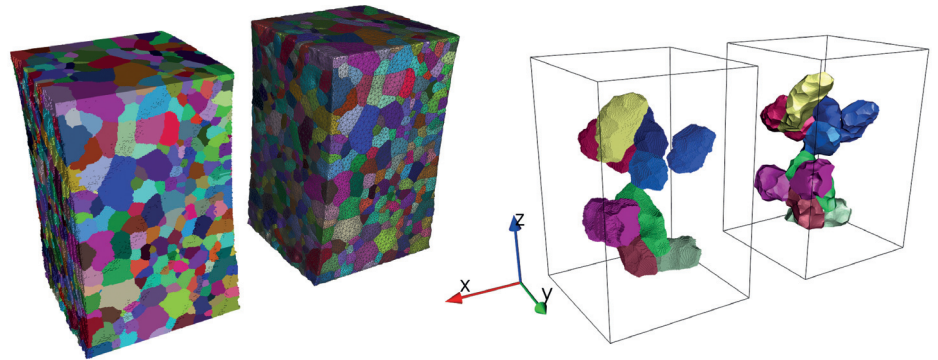
- Collect, analyze, and store huge volumes of data from experimental observation techniques
- Support data-intensive analysis and simulation with a unified storage solution
- Improve researcher productivity by eliminating data access bottlenecks

### SOLUTION

- Offers consistent performance and rapid data access, even for heavy concurrent and mixed workloads
- Supports data-intensive research techniques, such as tomography, 3D image correlation, and 4D statistical studies
- Scales to meet the needs of a growing research community with expanding data volumes
- Eliminates the need to move or copy data files, simplifying data access

### RESULTS

- Increases data access speed by 4X to 5X, improving researcher productivity
- Boosts system performance through a more balanced distribution of data
- Eliminates system downtime while enhancing availability and resiliency
- Reduces system administration and maintenance effort
- Meets immediate storage needs while providing room to grow



4X to 5X increase in data access speed for high-performance computing

“With Panasas ActiveStor, researchers save time since they no longer have to move or copy data files before they can process them. They gain productivity since there are no more HPC bottlenecks to access the storage system. Bandwidth is four to five times higher than before. And file-sharing is more convenient, thanks to shorter response times and higher throughput.”

#### Gregory Sainte-Luce

IT Manager at the Materials Research Center and member of the IT board, MINES ParisTech

MINES ParisTech, France’s most prestigious engineering school, works with companies and universities from around the globe. The public institution holds more than €30 million in annual research contracts.

Increasing the number of partnerships is critical for MINES ParisTech – as these contracts not only expand the school’s research efforts but also enable the creative discoveries that improve lives worldwide.

Researchers depend on experimental observation techniques that employ high-throughput 2D imaging and 3D

nondestructive imaging. These technologies create huge volumes of data, which researchers collect and analyze daily using sophisticated tools such as image analysis and mechanical simulations

To process the data, researchers previously had to move it between several storage systems, which compromised productivity and resulted in data duplication. Management and maintenance of multiple storage systems was also time-consuming for IT.

To accommodate new tools such as tomography and 3D image correlation, the Center needed a unified storage solution capable of providing excellent performance while supporting continued data growth. Another priority was providing a positive user experience for everyone – including 70 researchers who conduct high-performance computing (HPC) and 130 users of the system’s file-sharing applications.

MINES ParisTech evaluated various parallel file and object-based storage solutions before selecting Panasas’ ActiveStor®, a plug-and-play HPC appliance that runs the PanFS® parallel file system. “We chose ActiveStor for its unique combination of high performance, easy scalability, and simple administration,” says Gregory Sainte-Luce, IT manager at the Materials Research Center and member of the IT board of MINES ParisTech.

## Responsive, Scalable Storage Technology

Located at the Materials Research Center, the ActiveStor system operates 24x7, providing users with responsive parallel access to storage and consistent high performance, even under heavy concurrent and mixed workloads.

The Center's HPC workloads range from a few hours to several days, running around the clock. The most resource-intensive HPC task is tomography data analysis, a process that can last several days. A single tomography job execution can require 10GB to 100GB of storage, and the Center runs hundreds of these jobs in its experiments. Largescale finite element simulations of the material behavior require up to 1 terabyte (TB) of storage per simulation, the Center must ensure that researchers have rapid data access.

The Panasas solution has met and exceeded user and IT expectations. "Researchers save time since they no longer have to move or copy data files before they can process them," according to Sainte-Luce. "They gain productivity since there are no more HPC bottlenecks to access the storage system. Bandwidth is four to five times higher than before. And file-sharing is more convenient, thanks to shorter response times and higher throughput."

## Consistently High Performance

With ActiveStor, users now have rapid, concurrent access to storage, even when workloads are mixed or usage is extremely high.

The Center conducts 4D statistical studies, which assess mechanical stress that occurs over time. The 4D testing process automatically creates a digital twin of the material sample, allowing analysts to simulate performance and validate the behavior models used by industrial partners.

"Analysis of extremely rich 4D data sets can take several months or even a year, and we are in the process of automating

this analysis with our BIGMECA initiative," says Henry Proudhon, a researcher at the Center. "We typically process 1 TB of data for each day of the experiment, with 15 to 20 days of experiments per year. An average study could require 100 HPC runs for six hours per day."

The Center's BIGMECA initiative uses techniques based on artificial intelligence that analyze massive data sets to generate forecasts of a part's lifespan, according to its geometry and characteristics, with a methodology that combines experimentation, simulation and statistical learning.

As a user, Proudhon appreciates how ActiveStor increases the availability of HPC data, speeding time to results. "Panasas handles large volumes that can be expanded or reduced quickly, as needed," says Sainte-Luce. "We currently run at 80% of capacity and our load level is 17% to 20%, according to Panasas monitoring. Data access is five times faster for HPC storage, and we've even seen better performance and convenience for users who are not using HPC. We have virtually no more duplicated data. ActiveStor handles the entire workload of the research lab without any performance issues."

What's more, centralizing its storage servers and arrays into a single converged solution helped the Center reduce users' need for data transfer between separate storage devices.

## Simple Manageability

ActiveStor reduced the Center's storage management requirements. The solution's management simplicity eliminated engineering effort, while reducing overall administrative time. Automatic aggregation of storage resources frees administrators from the need to configure complex hierarchical storage management rules.

The intuitive Web-based graphical user interface allows one administrator to handle quota and volume management and maintenance in just 40 hours per year. ActiveStor automatically sends quota

warnings to administrators via e-mail, letting them quickly and dynamically adjust disk space as needed, without any interruption of service.

"Adding a new shelf is a risk-free and seamless process that takes less than one hour," says Sainte-Luce. "ActiveStor automatically rebalances our existing stored data between all the storage servers in the newly expanded system - without any performance degradation. The more balanced distribution of data significantly improved our system performance."

## More Reliable, Less Costly Storage

Since deploying ActiveStor, users have experienced no system downtime. Multiple metadata and file system management servers provide redundancy, increasing availability and resilience.

"Administrators spend less time managing storage, which allows them to focus on other activities," says Sainte-Luce. "Across the board, ActiveStor's seamless operation has helped us reduce the cost of storage management."

## Flexibility to Support Growth

ActiveStor offers MINES ParisTech the scalability needed to meet its immediate and future storage requirements. Panasas handles the current workload of the entire research group, and it can easily accommodate changing needs, such as expanding or reducing hard and soft volume quotas. "Scalability is one of the key advantages of our Panasas solution," says Sainte-Luce. "It's virtually effortless for us to scale the technology to meet demand."

ActiveStor also helps researchers accomplish more. "When researchers can access their scientific data faster, achieve higher job throughput, and perform more large computations, they can complete more jobs," says Sainte-Luce. "Their efficiency helps us sell more contracts to our partners and grow our business. We're very pleased to have Panasas as our partner."