

PRODUCT BRIEF

The data engine for all your applications: The **PanFS** architecture makes more possible

Focus on innovation with an intelligent storage solution designed for HPC and AI/ML

The Challenge

If you're a product developer, academic researcher, materials scientist, or any other kind of knowledge worker, you want to focus on innovation without being held back by your data. Despite the problem-solving potential of high performance computing (HPC), artificial intelligence (AI), and machine learning (ML), processing data at the speed and volume that innovation requires is often difficult. Teams are grappling with frequent storage downtime, constant tuning, and disparate data silos, as well as the need to hire expensive HPC storage experts. These challenges result in:

- Reduced productivity as IT and developer teams struggle to access disjointed data and are stuck waiting for systems to either input and output data or restore to full operability after a failure
- Unstructured data spread across complex storage systems making team collaboration difficult and requiring specialized management processes
- An ever-growing total cost of ownership (TCO) as expert staff, time, and money are needed to maintain complicated and outdated storage systems

The Panasas solution for powering HPC and AI/ML

To support their HPC and AI/ML needs, large organizations such as academic institutions and global enterprises need the right storage infrastructure with tools built in to make managing massive volumes of data simple. PanFS®, the Panasas operating system that powers Panasas' storage platforms, liberates teams from data storage headaches and gives them space to focus on their research by meeting all industry-wide needs.

HPC performance and scalability:

PanFS supports HPC environments, allowing researchers to process large amounts of data quickly. Performance increases as data volumes scale.

Enterprise-grade reliability: With PanFS, you avoid frequent, long, unplanned downtime and benefit from optimized data integrity. High availability is built in with automated online failure recovery and rapid parallel reconstruction.

Simple manageability: HPC storage specialists are expensive. PanFS offers a system that automatically manages data placement without user intervention, eliminating significant manual tuning that can destabilize HPC storage systems and allowing your experts to focus on the matters they specialize in without being held back by storage.

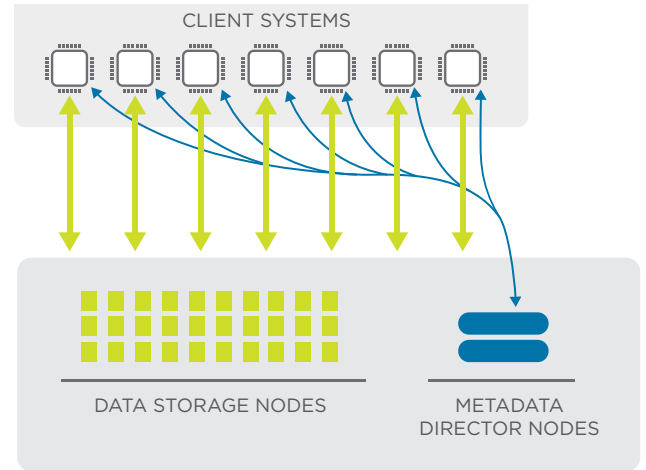
Data workload flexibility: PanFS provides teams with rapid access to heavy concurrent and mixed workloads from a single storage space, providing a consolidated storage environment for all your HPC and AI/ML applications – all while maintaining consistent performance across your data environment. The PanFS data engine allows you to think big when it comes to your data. With a user-friendly interface and powerful data management tools, it removes the manual work, unreliability, and inefficiencies associated with HPC and AI/ML storage.

As the pioneer in scale-out parallel file systems, as well as a leader in supporting both industry and academic research, Panasas has engineered PanFS to eliminate the headaches associated with HPC storage and the limitations associated with enterprise storage. Manageability and reliability are hardwired into the design’s DNA, enabling your teams to innovate faster and more easily than ever before.

Key Benefits of PanFS

Scale limitlessly with enterprise-grade reliability and lower TCO

- Scalable data capacity and performance, so your organization’s growth is a blessing, never a curse
- A rock-solid reliability architecture means fewer



PanFS Parallel File System.

unplanned outages and downtime that lead to wasted resources

- Expanded application access to data without needing a team of HPC storage experts

Eliminate data silos to foster collaboration

- Storage consolidation for managing mixed HPC and AI/ML workloads all from the same platform
- User-friendly interface that facilitates consolidation rather than siloing, for all staff, not just storage experts
- Streamlined workflows that save time and complexity due to a central repository where data stays throughout its analysis

Manage HPC storage operations without all the specialized staff

- Simple set-it-and-forget-it management
- Single-vendor, first-class customer support
- A platform that allows organizations to host home directories and project files under one global namespace

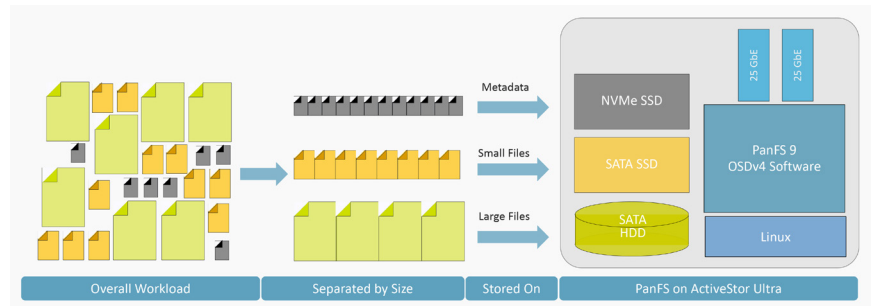
Achieve higher productivity

- Rapid data access for heavy concurrent and mixed workloads
- Consistent performance
- Ability to work with data in a single place without the need to move or copy data files to separate environments.

Get HPC and AI/ML capabilities with enterprise-grade features

- Consolidated storage that makes it possible for big enterprises to reliably deploy HPC and AI/ML

- High availability and data integrity
- Full data management tool suite for data visibility, movement, and protection wherever your data is located – across platforms and in the cloud.



Dynamic Data Acceleration.

Unlimited scalability

HPC, AI, and ML are powerful technologies that represent realistic approaches to breakthrough innovations. But it can be difficult to access such capabilities at scale or in a manner that's feasible for large organizations with massive amounts of data and complex requirements. The PanFS data engine consists of three components: Director Nodes to manage file system metadata, Storage Nodes to store data and metadata, and the DirectFlow client driver that reads and writes files stored by the PanFS operating system. Both the Storage Nodes and the Director Nodes can scale out. If a team needs more storage capacity or performance, they can simply add more Storage Nodes. If they need greater metadata processing, they can add more Director Nodes. There is never the fear of "maxing out" performance with PanFS. The entire system is designed to offer linear scale out. If a team adds 50% more storage nodes, they'll obtain 50% more performance and 50% more capacity.

No more data silos

There is no longer any time to have storage islands for individual groups within an organization. PanFS brings your different workloads with multiple data types together, so you can reach insights and innovations faster. One of the challenges of HPC is dealing with a range of file types and access patterns, as well as dynamic and fast-changing workloads. When teams work with other solutions, they have to manually re-tune to accommodate this variety. This time-consuming labor isn't necessary with Panasas. Our solutions offer exceptional mixed workload performance through Dynamic Data Acceleration (DDA). DDA makes the most of the underlying hardware. By using a careful balance of storage media technologies, such as SSDs (NVMe and SATA) and HDDs, our solution provides exceptional performance at lower cost per TB.

Manage HPC storage operations without specialized staff

Your people don't have to worry about understanding the intricacies of this system. All of their work happens through a graphic user interface (GUI) or command line interface (CLI) that runs on a Director Node. The Director Node automatically processes file system metadata, such as directories and file attributes, and coordinates the work of the Storage Nodes and DirectFlow Client drivers. This work includes file access management, failure recovery, and data reliability operations. Teams never have to interact with the Storage Nodes or the DirectFlow Client driver. PanFS offers a single file system namespace called a "realm."

Realize higher productivity

PanFS is a parallel file system. DirectFlow enables it to outperform the standard NFS or SMB/CIFS protocols. When PanFS stores a file, it stripes it across multiple Storage Nodes, so different segments of the files are located on different physical devices, allowing for parallel reading and writing of files and increased performance when accessing them. Enterprise storage solutions typically run file access through a lengthy process that lowers the ability to scale performance. First, the file access request goes to a "head node" that runs NFS or SMB/CIFS. Then, it goes through a backend network to get to other nodes that hold the HDDs and SSDs that hold the files. This system leads to data traffic building up at the head node. As a direct file system, PanFS takes a different approach that eliminates bottlenecks. The DirectFlow client speaks directly to the Storage Nodes that are holding the files, offering limitless and consistent performance scaling.



HPC and AI/ML with enterprise-grade features

Enterprises no longer have to make a tradeoff between cutting-edge technologies such as HPC, AI, and ML and the enterprise-grade features they need. PanFS offers erasure coding (EC) that stripes data across many storage nodes for advanced data protection. Once PanFS stripes the data, it creates a map to each file in each Storage Node, so every striped component can be easily found. The DirectFlow Client then uses that striping map to know which Storage Node to look at when trying to access a file. PanFS is POSIX-compliant and offers the ability to take per-volume snapshots for easy, user-directed recovery, as well as encryption at rest and access control lists. Plus, PanFS offers tools to manage your data across multiple environments with PanView analytics tools and the PanMove mobility suite. You can sync, backup, or archive data between Panasas systems, across third-party storage, and in the cloud.

Optimize your storage infrastructure with PanFS

The Panasas team specifically engineered PanFS to optimize HPC and AI/ML storage performance, reliability, and manageability. Your organization can choose the underlying storage system that's right for your data environment —whether that's all-flash, hybrid, massive capacity, or a combination — and then use the PanFS suite of software tools to get the maximum performance with the most reliability at the best price. It enables versatility and data consolidation, so teams can manage mixed workloads while drastically reducing outages and minimizing manual tuning. Panasas' PanFS storage software provides the innovation engine needed to power the next chapter of research and development.

About Panasas

Panasas builds a portfolio of data solutions that deliver exceptional performance, unlimited scalability, and unparalleled reliability—all at the best total cost of ownership and lowest administrative overhead. The Panasas data engine accelerates AI and high-performance applications in manufacturing, life sciences, energy, media, financial services, and government. The company's flagship PanFS® data engine and ActiveStor® storage solutions uniquely combine extreme performance, scalability, and security with the reliability and simplicity of a self-managed, self-healing architecture. The Panasas data engine solves the world's most challenging problems: curing diseases, designing the next jetliner, creating mind-blowing visual effects, and using AI to predict new possibilities.



Worldwide Office
1-888-PANASAS
info@panasas.com

Panasas Headquarters
San Jose, CA, USA
Panasas Research & Development
Pittsburgh, PA, USA

Panasas EMEA
Oxford, United Kingdom
+44 20 7751 2276
emeainfo@panasas.com

Panasas APAC
Selangor, Malaysia
+60 12 385 3028
aong@panasas.com

Panasas China
Shanghai, China
+86 181 1613 1226
twang@panasas.com