# PANASAS®

# PanFS

## The Preeminent High-Performance Parallel File System

Panasas® PanFS® is an integrated portable storage platform that combines a high-performance distributed and clustered parallel file system, volume manager, and software-based erasure-coding data protection. Running on networked commodity compute servers, PanFS uses parallel and redundant access to data storage nodes to deliver the highest performance with unlimited scalability, exceptional performance, enterprise reliability, and ease of management.

PanFS is the reliable, autonomic data engine that orchestrates networked servers into a single file system that serves data to clients at hundreds of gigabytes per second with little to no manual intervention. It automatically recovers from any failures, provides network-distributed erasure coding, and continuously balances the data load across the computers. PanFS also ensures data integrity via background scrubbing and encrypts stored data for the highest levels of protection from unwanted exposure.

Built on a foundation laid over 20 years ago and now in its ninth generation of commercial deployment, PanFS is ideal for scientific computing, modeling and simulations, high-performance data analytics, and AI/ML workloads in multiple industries including manufacturing, life sciences, semiconductor design, financial services, energy, universities, government, and media & entertainment.

### PanFS Architecture

There are three fundamental components that comprise the PanFS parallel file system: data storage nodes, metadata director nodes, and the DirectFlow® driver on client systems. As shown in Figure 1, the three elements work together to power the PanFS parallel file system. Director and storage nodes are commodity compute servers with high-speed networking connections dedicated to running PanFS software. File system workloads are divided between director nodes that process metadata and storage nodes that process data.

### DirectFlow Client Driver

The Panasas DirectFlow client driver is a loadable file system implementation that can be installed on Linux client systems and used by application programs like any other file system.
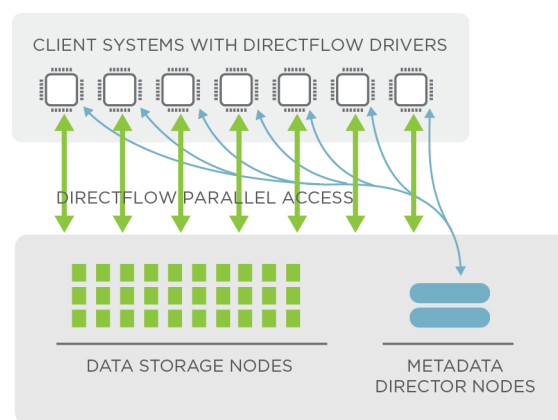


**Figure 1.** The PanFS parallel file system.

DirectFlow clients work with director nodes and storage nodes to deliver fully POSIX-compliant and cache-coherent file system behavior, from a single namespace, across all the servers in the compute cluster. For each file an application accesses, the DirectFlow client reads and writes directly in parallel to all storage nodes that hold that file's data. The DirectFlow client is the key to eliminating bottlenecks, application scalability issues, and erratic performance common to scale-out NAS systems.

### NFS and SMB/CIFS Gateway

PanFS provides high-performance NFSv3 and SMB/CIFS v3 access through director nodes that act as gateways translating NFS and SMB/CIFS operations into DirectFlow operations. This ability allows clients such as laptops and workstations to access the same namespace and data. But, as it is parallel and direct, the DirectFlow protocol will always be the highest performance path to PanFS storage.

### Unlimited Scalability

For more capacity or storage performance, just add more storage nodes. For more metadata processing performance, add more director nodes. In PanFS scale-out storage systems, there simply is no maximum performance or maximum capacity.

## Reliability that Increases with Scale

Any system can experience failure and, as systems grow larger, those failures can become more frequent and severe. However, PanFS reliability actually increases with scale due to its linear scale-out reconstruction time improvement in the event of a storage node failure.



**Figure 2.** PanFS Dynamic Data Acceleration illustrated on an ASU-100.
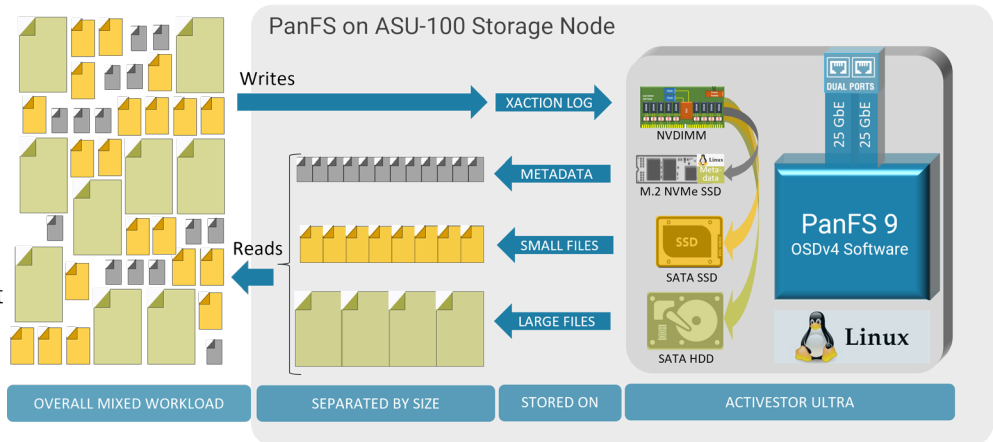
## Dynamic Data Acceleration and Mixed Workloads

PanFS Dynamic Data Acceleration (DDA) technology takes the complexity out of tiered high-performance storage systems by maximizing the efficiency of all storage media in a seamless, all-hot system that matches I/O patterns to storage device capabilities. DDA automatically adapts to changing file sizes and mixed workloads without the need for tuning or manual intervention, as illustrated with an ASU-100 storage node in Figure 2.

PanFS architecture delivers exceptionally high and consistent performance for workloads that include a wide range of file sizes, access patterns, and changes in volume over time. The result is a dynamic performance environment compared to other parallel file systems that require time-consuming and laborious tuning and retuning as workloads change.

## Enterprise-Grade Data Protection and Availability

PanFS uses software-based, network-distributed erasure codes to separately protect individual files, rather than traditional RAID groups that only offer drive-level protection. This eliminates the multiple risks, costs, and performance penalties of traditional RAID architectures.

## Multiple Levels of Security

To prevent unauthorized data access while PanFS is online, PanFS supports file system ACLs and SELinux security labels. PanFS also offers hardware-based encryption-at-rest with zero performance impact to prevent unauthorized data access while PanFS is offline by using industry-standard self-encrypting drives (SEDs). PanFS is KMIP compliant, allowing customers to use proven KMIP security key management servers, including a tested and validated integrated solution with industry leading CipherTrust Manager from Thales.

## Surprising Simplicity, Streamlined Management

As scale-up storage architectures grow in complexity, manageability often suffers. PanFS file system architecture simplifies management by offering a single global namespace that easily scales. PanFS automates key workflows such as new storage discovery and load balancing to streamline performance, and offers enterprise data services such as reporting, snapshots, and user quota enforcement. In addition, automated capacity balancing and centralized management features streamline management as storage volumes grow.

## More Information

For more information about PanFS, contact your local Panasas representataive or visit **panasas.com/products/panfs**.

### 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 | Panasas Headquarters | Panasas EMEA | Panasas APAC | Panasas China |
|---|---|---|---|---|
| 1-888-PANASAS | San Jose, CA, USA | Oxford, United Kingdom | Sydney, Australia | Shanghai, China |
| info@panasas.com | **Panasas Research & Development** | emeainfo@panasas.com | apacinfo@panasas.com | chinainfo@panasas.com |
| | Pittsburgh, PA, USA | | | |