PANASAS®

DATASHEET

All Storage Architectures Are Not Equal

Superior Architecture Leads to Superior Performance and Reliability

Performance and Reliability. Those are the touchstones against which any storage system is measured. Several factors combine to determine how well two storage solutions compare, but the most important is their architectures. Performance, reliability, and management complexity are direct outcomes of architecture.

Comparing architectures is the act of comparing how each element of the system is designed and built, and how well they each integrate with the others to achieve the overall goals.

There are two major classes of storage architectures: legacy ones developed for Enterprise systems, and ones such as Panasas' PanFS® that have been developed for the scale and performance needs of HPC environments. This data sheet will compare the two and point out the significant differences that clearly place one ahead of the other.

Legacy Architectures

Enterprise file storage has evolved over the years from a single computer (a "NAS head") serving several clients to multiple NAS heads with attached arrays of drives and hundreds of clients.

The NAS heads are connected via a private high-speed "back-end network" whose sole purpose is to allow those NAS heads to talk to each other. The NAS heads do both metadata processing such as directories and file timestamps, plus data movement as the files are read and written.

Indirect Back-End Network – When a client sends a data request to one of the NAS heads, the NAS head first checks its local drives for that data. If the data is found, it's returned to the client by that same NAS head.

However, if the desired data is not local to that NAS head, the request is forwarded through the backend network to the NAS head that does own that data. That second NAS head checks its local drives for the data, retrieves it from them, and sends it back through the backend network to the original NAS head, which then sends it on to the requesting client.

Increased Cost – Purchasing the redundant network switches and additional dual-port network interface cards inside each NAS head to build the back-end network is a significant cost.

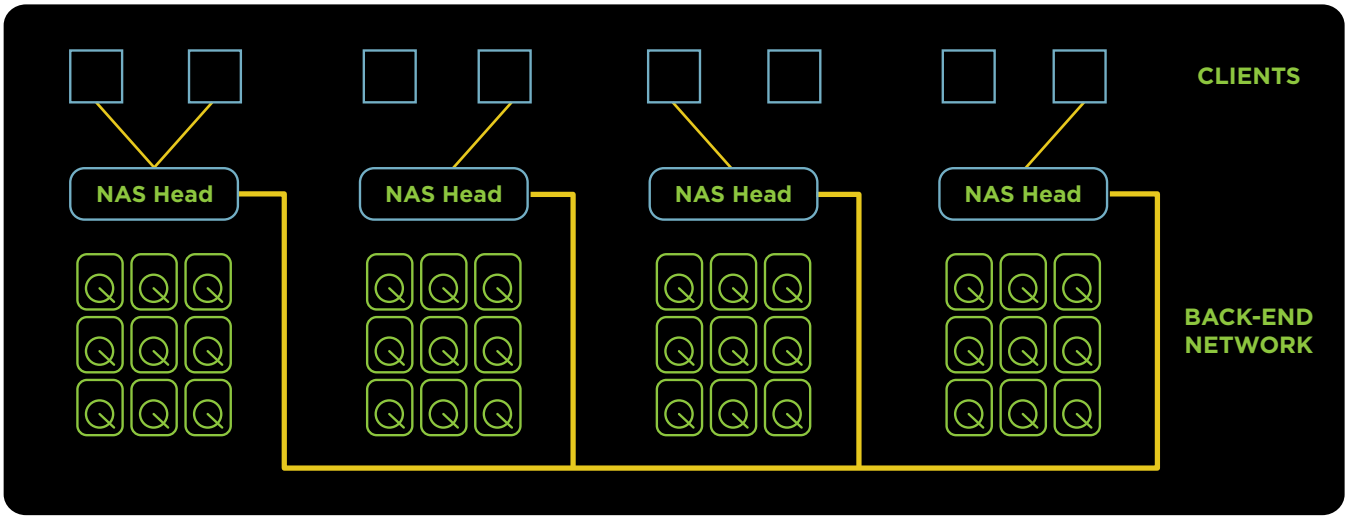

Figure 1. Legacy Architectures

Decreased Performance – A file read or write operation will take longer to complete (increased latency) because of having to hop through that backend network. Those reads and writes may also have lower bandwidth if the system is under heavy load and the backend network becomes congested.

Hotspots and Bottlenecks – The files held by the storage system may not be spread evenly across the NAS heads, so you may experience hotspots, bottlenecks, and erratic performance when one dataset is being heavily accessed. One NAS head can be overloaded while the others may be lightly loaded, or idle.

RAID roups – File data is stored in traditional RAID groups, small groups of drives that include redundancy information so that a failed drive can be rebuilt. Rebuilding a drive requires a spare drive and can only leverage the limited number of other drives in that RAID group, dragging out the rebuild operation.

There is a better way. One that removes the NAS heads entirely and significantly increases performance and reliability.

Direct Parallel File System

Unlike the NAS heads from a legacy architecture system, a parallel filesystem separates its metadata processing from its data movement, running them on computer systems dedicated to one purpose or the other. In PanFS, they're called Director Nodes and Storage Nodes, respectively.

Direct – In a parallel filesystem, data doesn't flow through a NAS head or even a Director Node, the data moves directly over the network between the clients and the Storage Nodes. There are no extra networking hops or NAS heads to bottleneck things.

Parallel – A parallel filesystem stripes each file across many Storage Nodes so it can be read and written in parallel. That multiplies the bandwidth that a single file can be read/written by a single client and is the core reason parallel filesystems are used in HPC, they provide extremely fast access to your data.

File System – Each file is separately striped across a random subset of the Storage Nodes. This evenly spreads the incoming load from the client systems across all the Storage Nodes, preventing hotspots or bottlenecks.

ASD Director Nodes – Are dedicated to processing metadata such as directories and timestamps, cross-client cache coherency for POSIX conformance, and storage cluster health monitoring. Data does not pass through Director Nodes except when a client is using the legacy NFS or CIFS protocols.

ASU Storage Nodes – Are dedicated to storing and retrieving user data. PanFS Storage Nodes are a carefully balanced combination of NVDIMMs, NVMe SSDs, SSDs, and HDDs that provide optimal performance, at excellent price/performance, for mixed-workload environments.

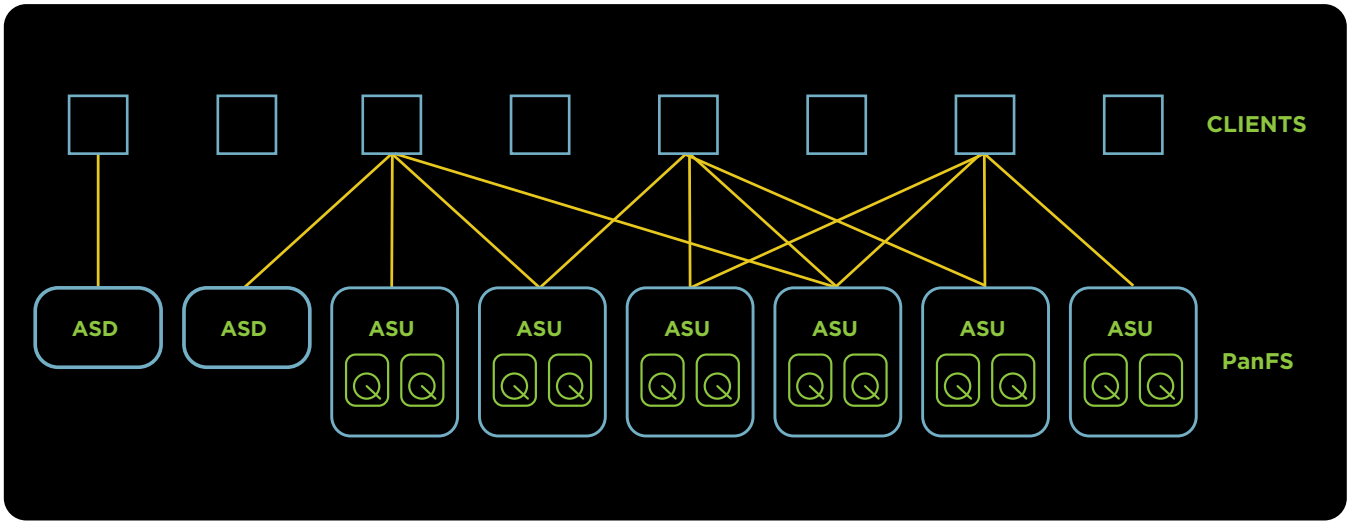

Figure 2. Direct Parallel File System

DirectFlow Protocol Client – To gain the advantages of a parallel filesystem, a kernel driver needs to be installed on the clients. The kernel drivers for the legacy NFS and CIFS protocols are in the Linux kernel already, while the driver for the PanFS DirectFlow protocol needs to be installed.

Erasure Coded Reliability – When a file is striped across Storage Nodes, an erasure code is used so that the failure of a configurable number of drives or whole Storage Nodes can be transparently recovered from. Continuous background scrubbing of those erasure codes prevents buildup of latent drive errors.

Reliability Increases with Scale – After the failure of a drive or even a whole Storage Node, all the Storage Nodes join forces to return the PanFS system to full levels of protection for your data files. As a result, as the number of Storage Nodes increases, the time to recover decreases, making overall data reliability go up.

Conclusion

Compared to a legacy architecture NAS product, a Direct Parallel File system consistently delivers an order of magnitude more bandwidth than a Legacy Architecture NAS product, is much more efficient in terms of performance delivered from a given storage capacity, and is much less prone to the hotspots, bottlenecks, and erratic performance common to Legacy systems.

Since there is no maximum number of Storage Nodes in a parallel filesystem such as PanFS, and growth of performance is just as linear as growth of capacity, there simply is no maximum performance of a parallel filesystem.

All those advantages derive from the architecture of a parallel file system. The result is an HPC environment with significantly improved performance and reliability, and a much lower cost to own and maintain.

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
emeainfo@panasas.com

Panasas APAC
Sydney, Australia
apacinfo@panasas.com

Panasas China
Shanghai, China
chinainfo@panasas.com