Data Storage and Mobility for Modern High Performance Computing Workloads

SOLUTIONS BRIEF

Overview

The use cases, applications, and workflow patterns of high-performance computing (HPC) are rapidly evolving. The use of HPC has broadened from government and academia to industry. HPC systems are now tasked with heterogeneous workloads, including real-time and event-triggered use cases. Users, data, and computing power are geographically dispersed. Data must be mobile, searchable, and sharable through all stages, from simulation and modeling to visualization, analysis, and archive. The integration of Artificial Intelligence and Machine Learning (AI/ML) into workflows is enabling new capabilities and insights.[1] Containerization and automation make workloads portable and burstable.

These changes in the HPC landscape are driving challenging requirements for data storage and data mobility. Hammerspace solves these challenges by combining a high-performance global parallel file system with data orchestration capabilities to form a unique global data environment.

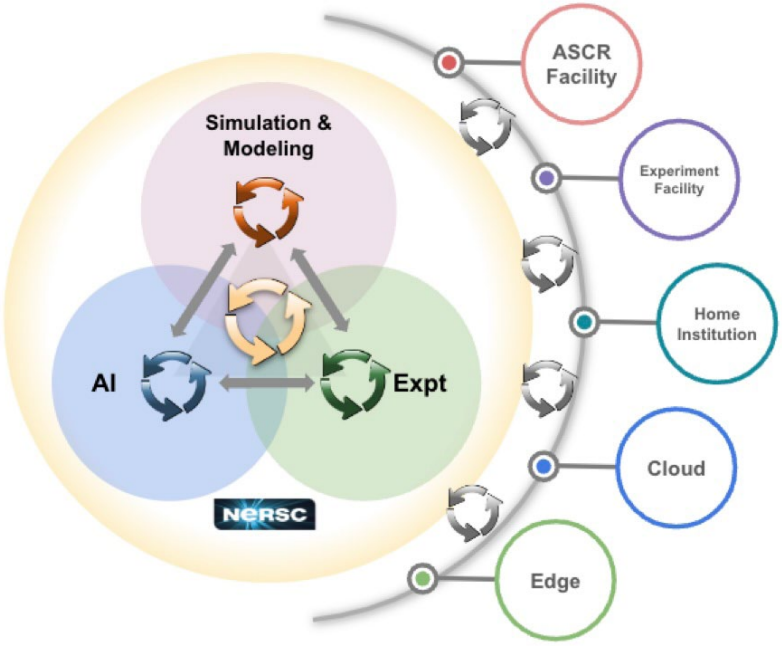


Figure 1 - Integrated Science workflow example, NERSC

Solution Summary

Traditional HPC environments have multiple silos of storage from different vendors. Boutique parallel file systems serve high-performance scratch storage, scale-out NAS handles pre- and post- job workflow stages, cloud storage is used for cloud bursting and collaboration, and long-term archives live on object storage or tape. Managing data access and protection requires separate tools for each silo, and orchestrating data flow between silos involves yet another layer of software. Data is frequently copied from file system to file system, wasting time and storage capacity.

---

https://www.nersc.gov/assets/NERSC-10/Workflows-Archetypes-White-Paper-v1.0.pdf

© 2024 - Hammerspace

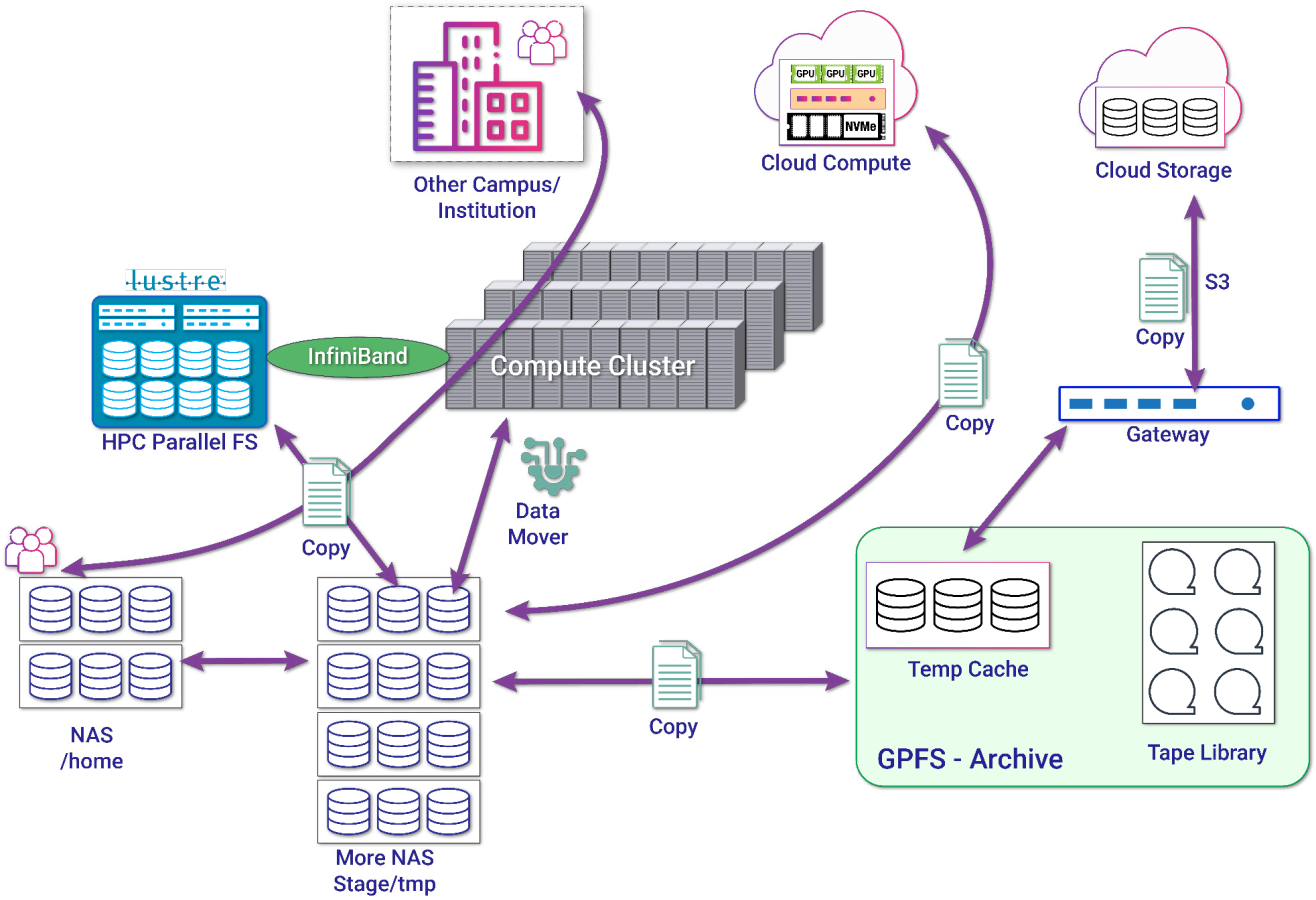


Figure 2 - Typical HPC Storage Environment Before Hammerspace

Hammerspace software creates a global data environment that simplifies and streamlines HPC data management in these crucial ways:

- Consolidating multiple silos, sites, and clouds
- Standardizing data services across heterogenous storage systems and sites
- Automating data orchestration with objective-based policies
- Creating high-performance parallel file systems with pNFS-based Hyperscale NAS

The integration of AI/ML into HPC workflows is enabling new capabilities and insights, but putting pressure on data storage infrastructures

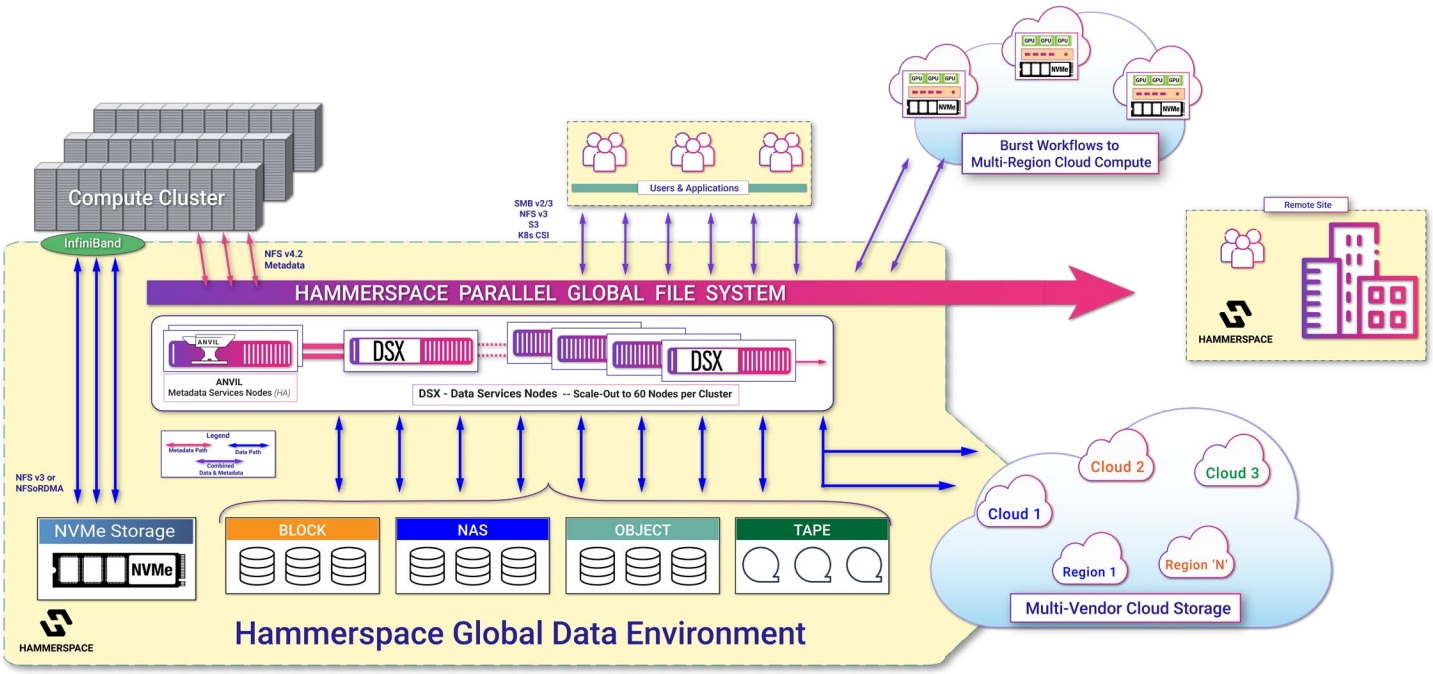


Figure 3 - HPC Storage Environment with Hammerspace

Solution Benefits

Parallel file system performance without vendor lock-in	Global file system spans storage technologies, sites, and clouds	Control data placement to meet workflow requirements
• Create high performance parallel file systems for scratch and pre/post processing with storage servers from any vendor or mixed vendors • Replace proprietary file systems with standard Linux pNFS • RDMA and GPUDirect Storage enabled for maximum performance	• Provide access to all data, wherever it resides • Use cloud storage for collaboration or burst workloads to cloud compute • Reduced management burden leaves more time for research	• Assimilate existing data in place – no lengthy storage migrations • Automatically place and relocate data based on policy objectives • Assign custom metadata for organization and automation

Storage Silo Consolidation

HPC environments traditionally incorporate file systems from multiple vendors, each with their own special properties. For example, Figure 4 portrays a "simplified" view of the file systems in use at Lawrence Berkeley NERSC. There are five different file systems shown, from four different vendors, to accommodate different use case profiles.

Simplified NERSC File Systems

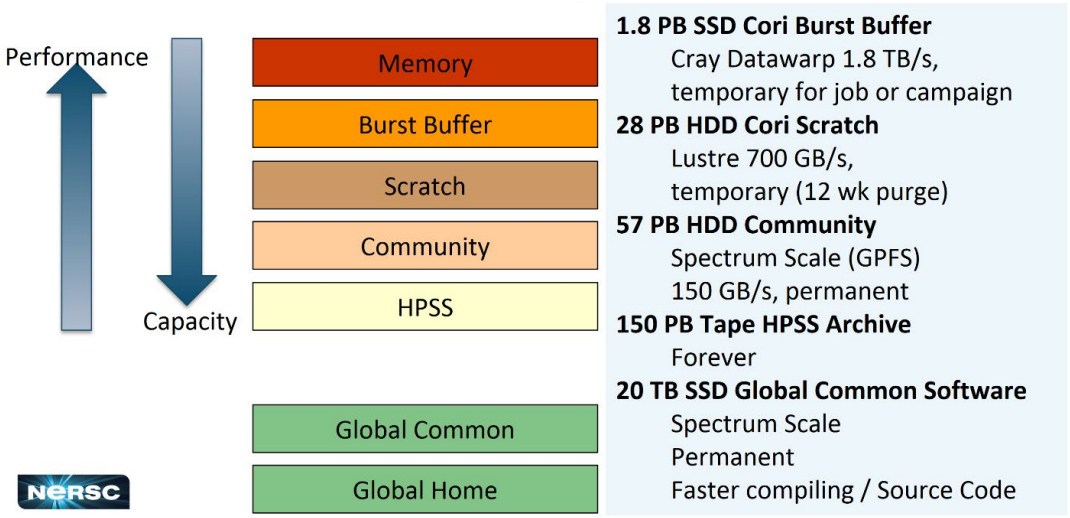


Figure 4 - Lawrence Berkley NERSC HPC File Systems

Most or all these use case profiles can be served by a single Hammerspace instance. There will still be multiple exports and multiple storage systems, but all within a Hammerspace global data environment. This results in a system with fewer points of management and valuable added capabilities.

The typical vendor approach to storage consolidation is to offer a new, better, bigger silo to replace multiple existing silos. This forces lengthy and risky data migration efforts, as existing files must be copied into the new storage every few years. In HPC environments this type of physical storage consolidation doesn't work. Never mind the pain of migrating petabytes of data — there is no single storage system that can meet the performance, scale, and cost requirements of every HPC storage tier simultaneously.

Instead of physical consolidation, Hammerspace logically consolidates storage silos. All are brought under a single umbrella through a process known as assimilation. Assimilation imports all file system metadata into Hammerspace, while data remains in place on existing storage.

Compared to typical PB-scale data migrations, which can take months to years, assimilation takes just hours to days. As an example, a recent 40PB assimilation was accomplished in two days. Even more impressive, assimilation uses a combination of background and on-demand methods to minimize disruption. As soon as an assimilation job is started, all data may be accessed from Hammerspace as if the assimilation has already finished.

Logical consolidation with Hammerspace acknowledges the necessity of multiple types of storage while providing many of the benefits of physical storage consolidation. Additional capabilities such as the multi-site global namespace and support for cloud storage provide added value by eliminating the need for single-purpose gateway products.

Multiple exports or shares may be created in Hammerspace, but they are abstracted and can span multiple storage systems. Additional capacity may be added at any time, with data migrated to newer storage as a background orchestration task. Storage systems will come and go, but what users and applications see — exports and shares — remains undisturbed, maintaining continuity and user productivity.

Standardized Data Services

Because Hammerspace disaggregates and elevates the file system above the storage, data services such as snapshots and clones, auditing, versioning, and data placement are implemented in a uniform way across all storage, at all sites.

Having a uniform set of tools to manage data protection and security across heterogeneous storage systems reduces the management burden. This is of particular benefit in academic and scientific environments, leaving more time for research activities. It also makes data loss less likely by reducing the chance that errors and misconfigurations will go unnoticed.

The diagram below enumerates the global data services provided by Hammerspace.

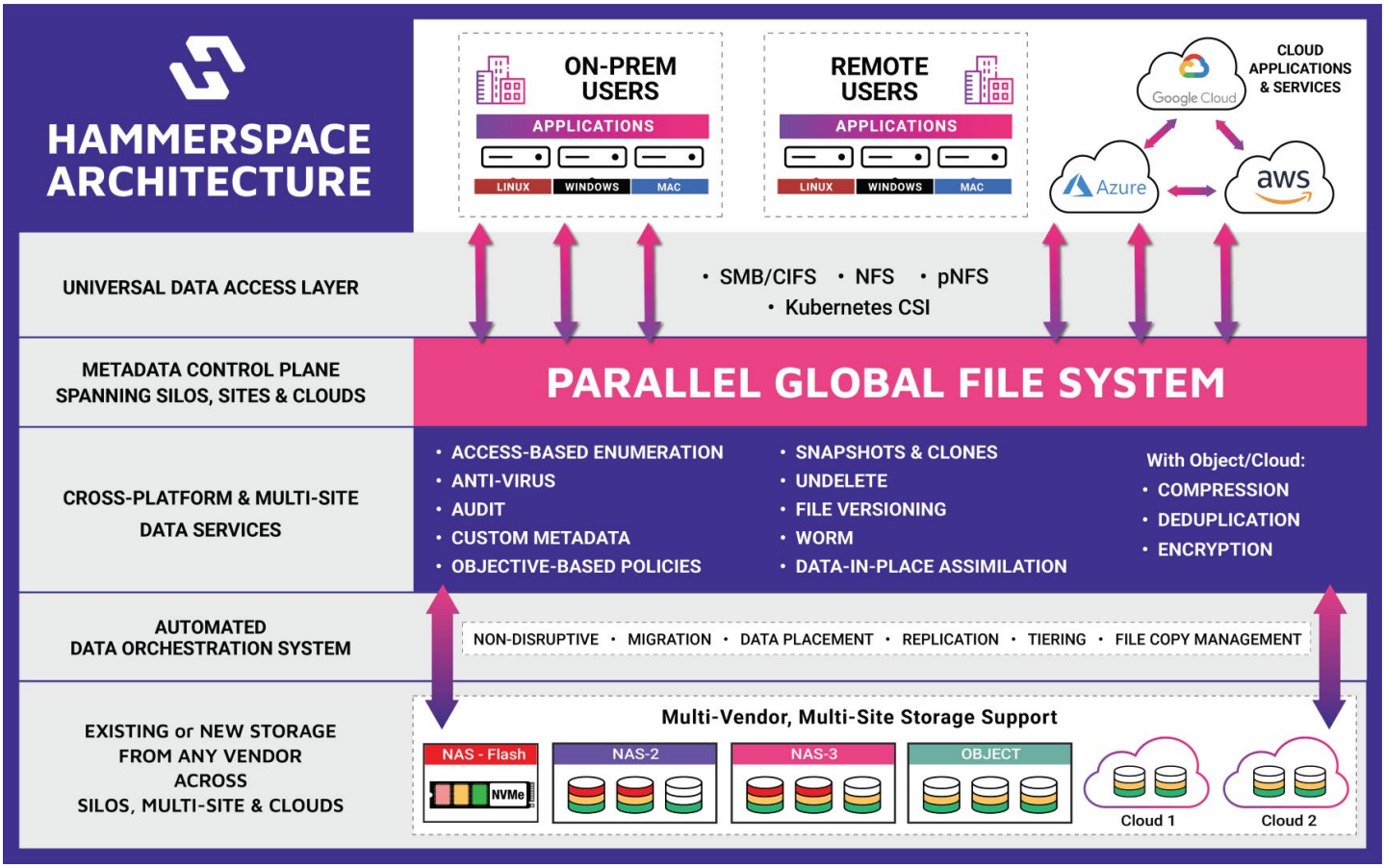


Figure 5 - Hammerspace Architecture

Automated Data Orchestration

The Hammerspace parallel global file system can span storage technologies from flash to tape, NAS to object and cloud. All tiers from high performance scratch to archive are part of the same data environment. This enables researchers to focus on the needs of the data, not the care and feeding of storage systems.

Once data is under Hammerspace control, it may be migrated, replicated, tiered, protected, and intelligently placed using automated data orchestration. Configurable policy objectives accomplish these tasks non-disruptively. Users and applications always have access to files, even when they are in motion, making migration and data placement activities transparent.

As soon as an assimilation job is started, all data may be accessed from Hammerspace as if the assimilation has already finished

Files and directories may be tagged with custom metadata, which follows the data across the global data environment, regardless of where it resides. Both standard and custom metadata may be used to drive objectives and workflow automation, track files and directories based on project or researcher, or other uses.

Hammerspace's built-in data orchestration eliminates the need for separate tools such as data mover software, archive systems, and cloud storage gateways.

Hyperscale NAS

Specialized HPC parallel file systems such as Lustre® and IBM Storage Scale™ (formerly GPFS™) provide high performance and scalability by separating metadata from data. This empowers client machines to interact directly with storage systems but requires custom client software to do so.

Hammerspace uses parallel NFS (pNFS) with the Flex Files layout type to realize the same architecture and achieve the same performance and scalability benefits using only the standard NFS client built into Linux. This Hyperscale NAS architecture achieves parallel throughput that scales linearly, limited only by the storage and network hardware available.

The open nature of Hyperscale NAS extends to the NFS server as well, enabling the creation of parallel file systems using any storage system that supports NFSv3, including white box or OCP servers running standard Linux.

While RDMA is not required to achieve high performance with the Hyperscale NAS architecture, it is supported, both using InfiniBand and RoCEv2. Hammerspace's implementation of Hyperscale NAS is also certified with NVIDIA GPUDirect Storage, for the fastest performance in workflows requiring GPUs.