HAMMERSPACE

Hammerspace Data Platform

A High-Performance Solution to Unify, Automate, and Accelerate Access & Orchestration of Unstructured Data

By Floyd Christofferson

Version 2 - January 2026

The Problem of Utilizing Global Data In a Siloed World

In This Paper

- Why traditional, storage-centric architectures fragment unstructured data across silos, sites, and clouds, and why this now limits modern data usage.
- How file and object metadata, when bound to storage platforms, drives copy-based data movement, duplication, and operational complexity.
- How the Hammerspace Data Platform uses a data-centric approach to bring all data together into a single parallel global file system to unify access and automate orchestration without migrating data.
- Why AI and other high-performance workloads expose the limits of siloed storage and serve as a stress test for data infrastructure at scale.
- How enterprise and research organizations apply this architecture in practice to support large scale computing, AI, and other data-intensive workloads.

Introduction

Enterprise IT environments have always been inherently heterogeneous. Unstructured data is rarely stored in a single file or object storage system or location, but is instead distributed across multiple vendor platforms, data centers, edge locations, and public clouds. This fragmentation has long increased operational complexity, risk, and cost, forcing organizations to manage and access increasingly distributed data across incompatible storage silos.

The problem is that what was once an operational inconvenience has become a serious business limiter as AI-driven use cases place new demands on that same data. What has changed is not the existence of silos, but the way data is being used in an AI era.

Unlike in the past, when data would age down to lower-performing tiers or into archives, modern workflows increasingly require data to remain accessible and reusable across its entire lifecycle. Analytics, distributed applications, collaboration, and automation all depend on the ability to access unstructured data wherever it resides, often by different use cases, applications, and users at the same time. Data that was once considered cold or archival must now be routinely revisited and combined with other datasets, which has elevated the silo problem to a core business issue rather than merely a costly nuisance.

AI workloads represent the most extreme expression of this shift, which in practical terms flips the traditional data lifecycle on its head. Training, inferencing, and agentic workflows demand high-performance parallel access to large volumes of unstructured data that may be distributed across silos, sites, and clouds. These requirements expose the limits of copy-based data management and vendor-locked storage architectures that IT teams have been working around for decades. AI does not create a new problem, but rather reveals the full cost of existing ones.

In AI initiatives, this directly impacts time to results and ROI. But even in traditional data environments it drives unnecessary duplication, infrastructure sprawl, and operational overhead.

When data remains locked inside proprietary silos, organizations are forced into costly tradeoffs: accept fragmented access and inefficiency, or build new repositories to make data usable across teams and workflows.

What organizations need is a way to activate any existing data to unify access without migrating it, duplicating it, or disrupting ongoing operations.

The Hammerspace Data Platform addresses this need with a solution that provides unified file and object access combined with automated data orchestration spanning storage from any vendor across sites and clouds, even at the extreme performance scales needed for AI workloads.

Without requiring customers to migrate data off their legacy storage, Hammerspace aggregates file and object metadata to create a parallel global file system with automated data orchestration that bridges them all.

No proprietary client software is required, nor any alterations to existing underlying storage. Users and applications access data on existing or new storage as a cross-platform global namespace via standard SMB, NFS, S3, and pNFS v4.2 protocols.

This enables data orchestration to become a background operation that is non-disruptive to users or applications, transparently staging data as needed to feed high-performance workloads on-premises or in the cloud. This also enables cross-platform control of other data services for data tiering, migration, protection, deduplication, and more.

Whether supporting traditional enterprise workflows or the most demanding AI pipelines, Hammerspace transforms long-standing data silos from an operational burden into a flexible, shared resource.

Users and applications gain seamless global access to their data, while IT teams gain policy-driven control to manage data placement and access non-disruptively across heterogeneous environments. Whether supporting traditional enterprise workflows or the most demanding AI pipelines, Hammerspace transforms long-standing data silos from an operational burden into a flexible, shared resource.

In this paper, we examine the architectural roots of unstructured data fragmentation and why traditional approaches to managing storage silos have become increasingly problematic. We then provide a more detailed overview of the Hammerspace Data Platform, including its key capabilities and example use cases.

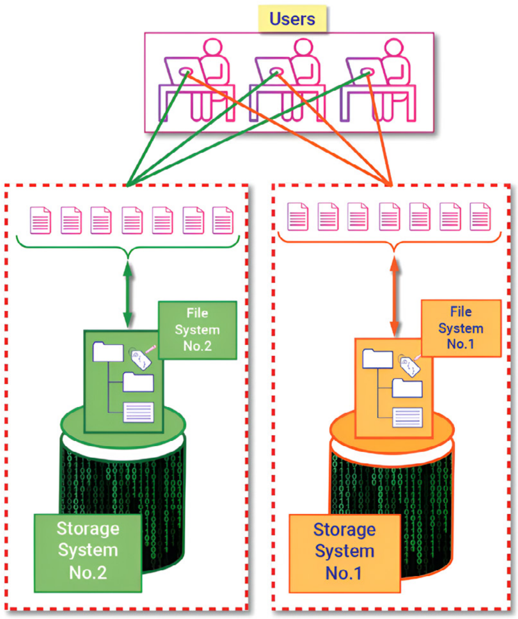


Figure 1: Although file systems were exported as shares on local area networks decades ago, file system metadata is still embedded in each vendor's storage system in proprietary formats. This means moving data to a different storage platform creates a forked copy of both the file essence and its metadata. This fragmentation is the root of the silo problem that pervades the industry to this day.

The Root of the Problem

The root of the silo problem lies not in storage hardware itself, but in how file systems have evolved and where they remain embedded today. Despite dramatic advances in storage performance, scale, and cost efficiency over the past several decades, the fundamental model for accessing unstructured data has changed very little. Users and applications still rely on file systems to translate raw blocks of data into files and directory structures that humans can understand.

Historically, file systems were part of a computer's operating system and were bound to locally-attached storage, making each system a standalone data island. Network-attached storage introduced shared access by exporting file systems over the network, effectively elevating the file system out of a single computer. This further expanded when scale-out NAS architectures were introduced, extending that architecture across multiple nodes for larger data volumes.

As data volumes inevitably grew, additional storage systems became necessary, fragmenting access to and management of the data.

As a result, moving data between platforms, locations, or tiers still requires copying both the data itself and the metadata that defines it. This creates redundant versions and fragments user and application access.

Vendor lock-in, copy-based data movement, and the proliferation of storage silos persist as a result.

The fact that file system metadata remains trapped within proprietary storage systems is the underlying cause of data fragmentation.

The Silo Problem Emerges at Scale

As data volumes grew beyond the capacity of individual NAS systems, organizations inevitably had to purchase additional storage platforms to meet demand. These systems were often optimized for different performance, cost, or workload characteristics, were frequently sourced from different vendors, and were a mix of file and object storage in the data center. Each new platform introduced its own file system and metadata domain, further fragmenting access to data and increasing complexity.

At small scales, this fragmentation could be managed. IT teams relied on manual processes, scripts, or point solutions to move or replicate data between systems and locations, as needed.

However, as environments grew these approaches became increasingly brittle. Each copy of data represented not just additional storage consumption, but a fork in visibility, access control, and lifecycle management. Determining which copy was authoritative, current, or properly protected became an ongoing operational burden for IT teams, but a critical compliance and governance risk for data owners.

Importantly, these challenges were not the result of poor tooling or insufficient process. They were a direct consequence of an architecture in which data access and mobility required copying, and in which metadata could not be shared across storage boundaries.

Silos Go Global

The emergence of cloud computing amplified these problems. While cloud platforms promised elastic capacity and global reach, they did not eliminate silos. Instead, they increased them. Public cloud storage systems introduced new geographic regions as well as new file systems, object stores, and metadata models with the goal of reducing capex and infrastructure by outsourcing data storage through a consumption-based model. In practice, cloud storage often became yet another silo alongside existing on-premises infrastructure.

The result is an environment in which an organization’s data exists everywhere, but access remains localized, inconsistent, and tightly bound to the storage systems on which the data was originally created.

Most organizations now operate hybrid environments that may span on-premises data centers, multiple cloud vendors or regions, and edge locations. In these environments, unstructured data is increasingly distributed by necessity rather than design. Traditional approaches to data movement, which rely on replication, gateways, or manual orchestration, struggle to scale across such decentralized architectures.

What was once an inconvenience in a single data center (when there was more or less a 1:1 ratio of application to data set) becomes a structural problem when extended globally and to multiple use cases for the data. Copy-based workflows multiply costs, increase operational risk, and introduce delays that are incompatible with modern data usage patterns. At the same time, the growing need by AI and other modern use cases to reuse and recombine data across locations makes fragmentation increasingly visible and problematic to users and applications.

The Architectural Impasse

Taken together, these trends expose a fundamental architectural impasse. Storage-centric data architectures were never designed to support unified, cross-platform data access at scale. Yet modern data usage increasingly demands exactly that.

Solving this problem requires more than incremental improvements to existing tools. It requires decoupling the metadata from the storage platforms that hold the data, and rethinking how unstructured data is accessed, shared, and placed in motion across heterogeneous environments.

This shift from a storage-centric model to a data-centric one sets the stage for a new approach to managing unstructured data at scale.

The Hammerspace Data Platform

The Hammerspace Data Platform was designed to address the architectural limitations of storage-centric data environments by unifying control and access across silos through a data-centric model. Rather than introducing yet another proprietary silo, Hammerspace creates a standards-based data environment that unifies heterogeneous storage systems and automates data management across them, while remaining invisible to users and applications.

At its core, the platform is built around a high-performance, parallel global file system that provides unified access to unstructured data across sites and clouds. This unified data plane is paired with automated data services and accelerated orchestration, allowing data to be placed, moved, and managed dynamically across silos, sites, and clouds without disrupting data access to users or applications. In this way, Hammerspace enables organizations to treat their entire data estate as a shared global resource, with the agility to adapt to changing data uses and application requirements, including AI-driven workloads.

Multi-protocol, Standards-Based Access with No Proprietary Client Layer

At the top of the architecture, Hammerspace presents data through standard, non-proprietary interfaces. Users and applications can now access their entire data estate using familiar protocols such as SMB, NFS, S3, pNFS v4.2 with Flex Files, CSI for containerized environments. APIs are available for integration with higher-level workflows and services, and an MCP Server interface will be introduced soon to enable direct, AI-native access to data by model-driven pipelines and agents.

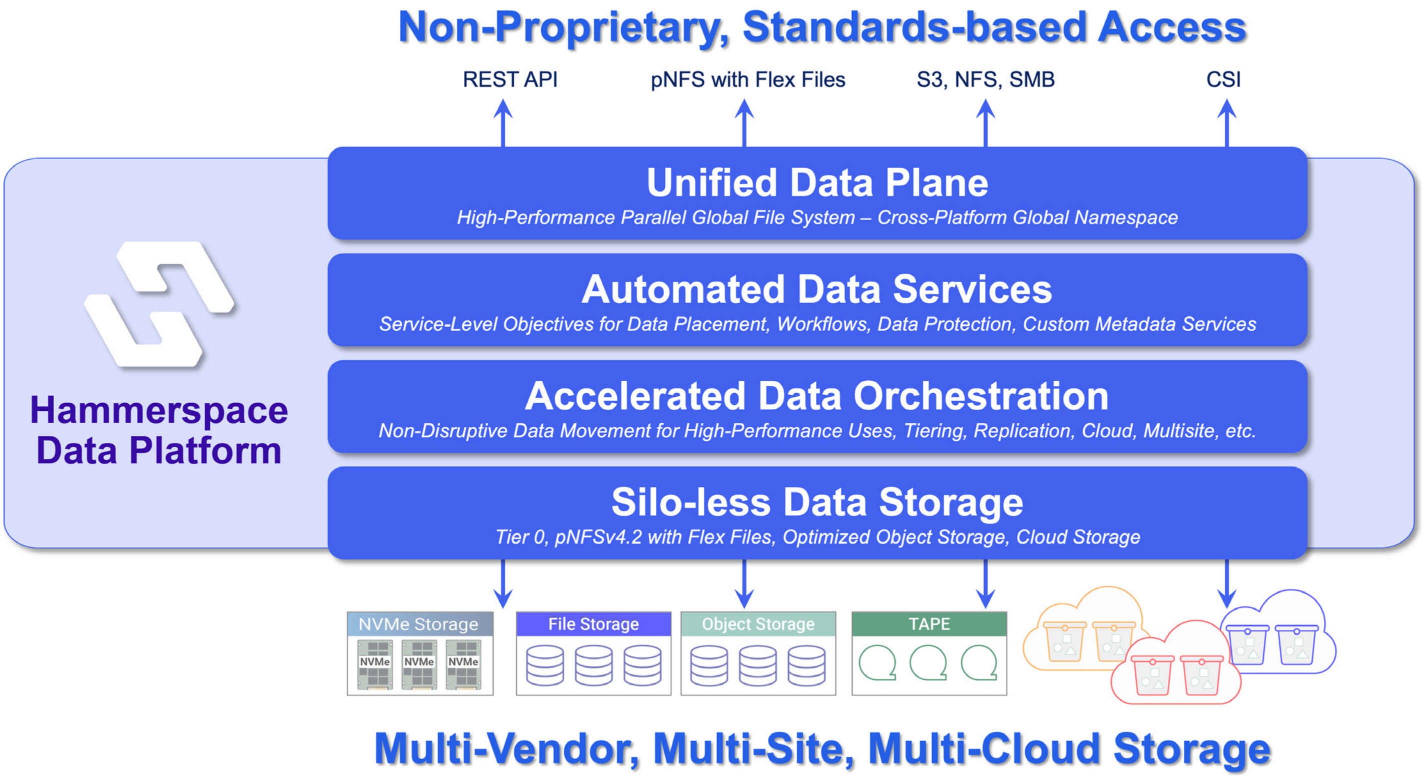
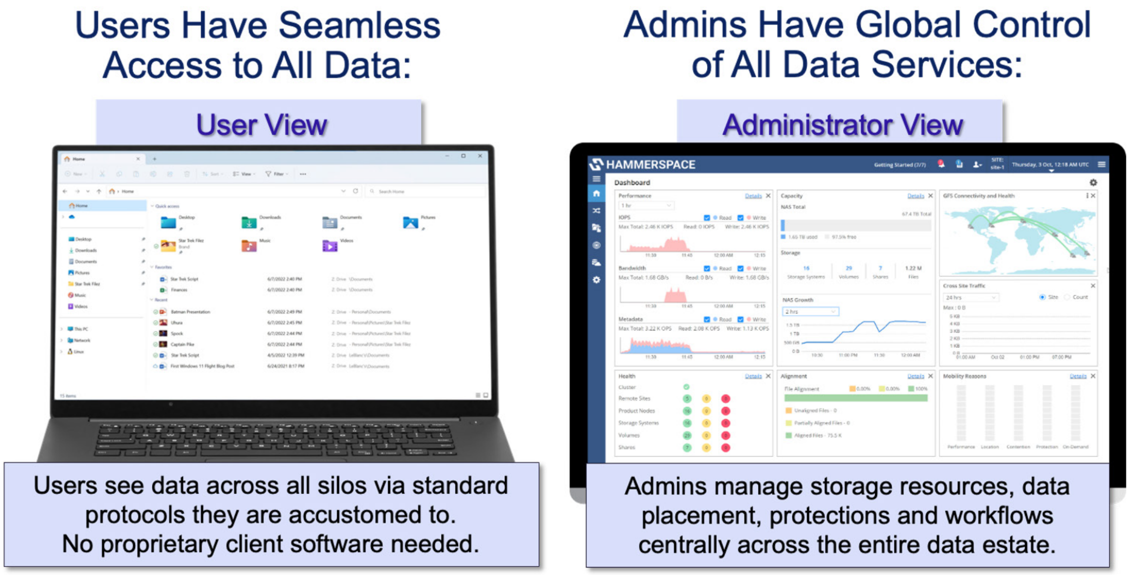


Figure 2: This is a logical view of the Hammerspace Data Platform. No proprietary client software needed for access or alterations to storage of any type, from any vendor, enabling non-disruptive integration into existing data environments.

Because Hammerspace focuses on industry-standard access methods, no proprietary client software is required on application servers, workstations, or compute clusters. From the perspective of users and applications, data appears exactly where it is expected to be, in the same file and directory structures they already understand. In many cases, users are unaware that Hammerspace is present at all. Data movement between storage systems or locations is invisible. The data continues to appear in the same share or bucket, regardless of where the underlying bits have been moved. With appropriate permissions, users can access their data across all silos and locations through the same standard interfaces they used before.



This design is particularly important in AI environments, where introducing custom agents or non-standard clients into GPU clusters can add operational risk and complexity. By remaining transparent and standards-based, Hammerspace integrates cleanly into existing compute environments without disrupting established workflows.

Unified Data Plane for Cross-Platform, High-Performance Access

The key to making this effective, however, is elevating the metadata layer out of the storage systems to create a unified data plane and single global namespace. This is done by assimilating metadata from existing storage systems while leaving data in place. Even though the data itself remains on existing storage silos, Hammerspace presents it as a single, coherent cross-platform global namespace that bridges any vendor, location, or storage type.

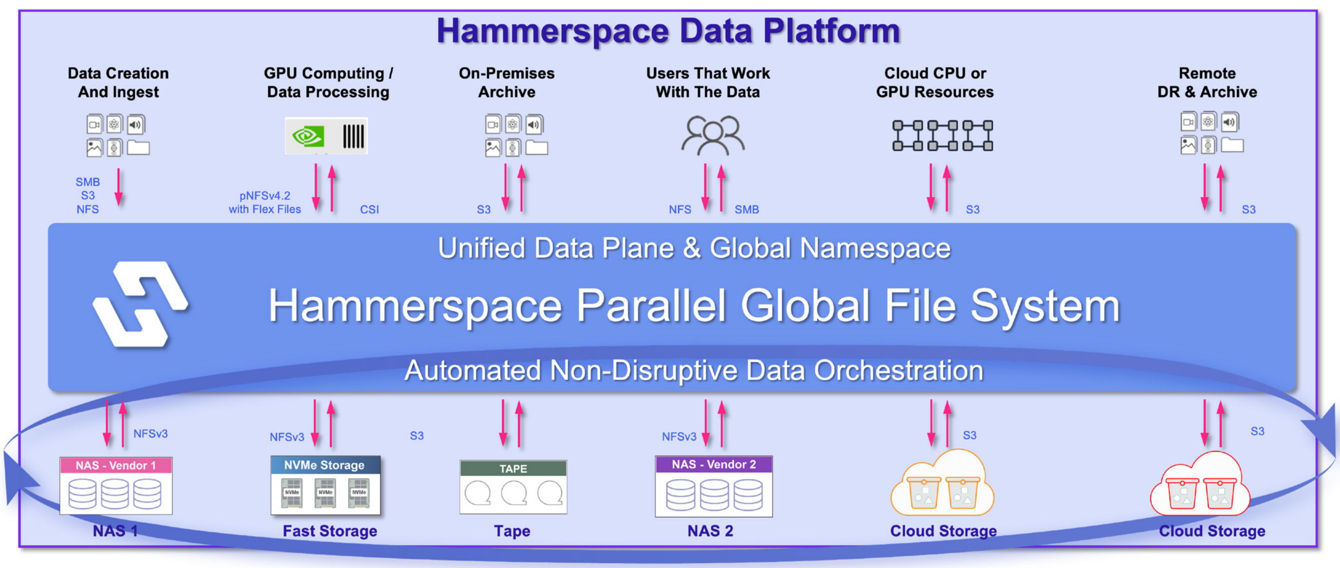


Figure 3: Hammerspace rapidly assimilates metadata from data in place on existing storage, without the need to copy or migrate it anywhere. No agents or alterations to underlying storage are required, nor is any proprietary client software needed on user systems. This enables persistent unified access all data globally via standard protocols, regardless of where the data is stored today or moves to tomorrow.

This unified data plane is designed to support the high levels of concurrency and throughput required by modern AI and analytics workloads. Multiple applications, users, and compute clusters can access the same data simultaneously, even when those datasets span multiple sites or clouds. By decoupling data access from storage boundaries, Hammerspace eliminates the need to create duplicate copies of data simply to make it accessible.

Automated Data Services Driven by Policy

By elevating the metadata layer into a unified, parallel global file system above the storage layer and exposing it through standards-based interfaces, Hammerspace enables a wide range of data services to run as background operations without disrupting users or applications, even on live data that is actively in use. These services are governed by service-level objectives within Hammerspace, defined by data owners or administrators on a global basis across all storage types and locations.

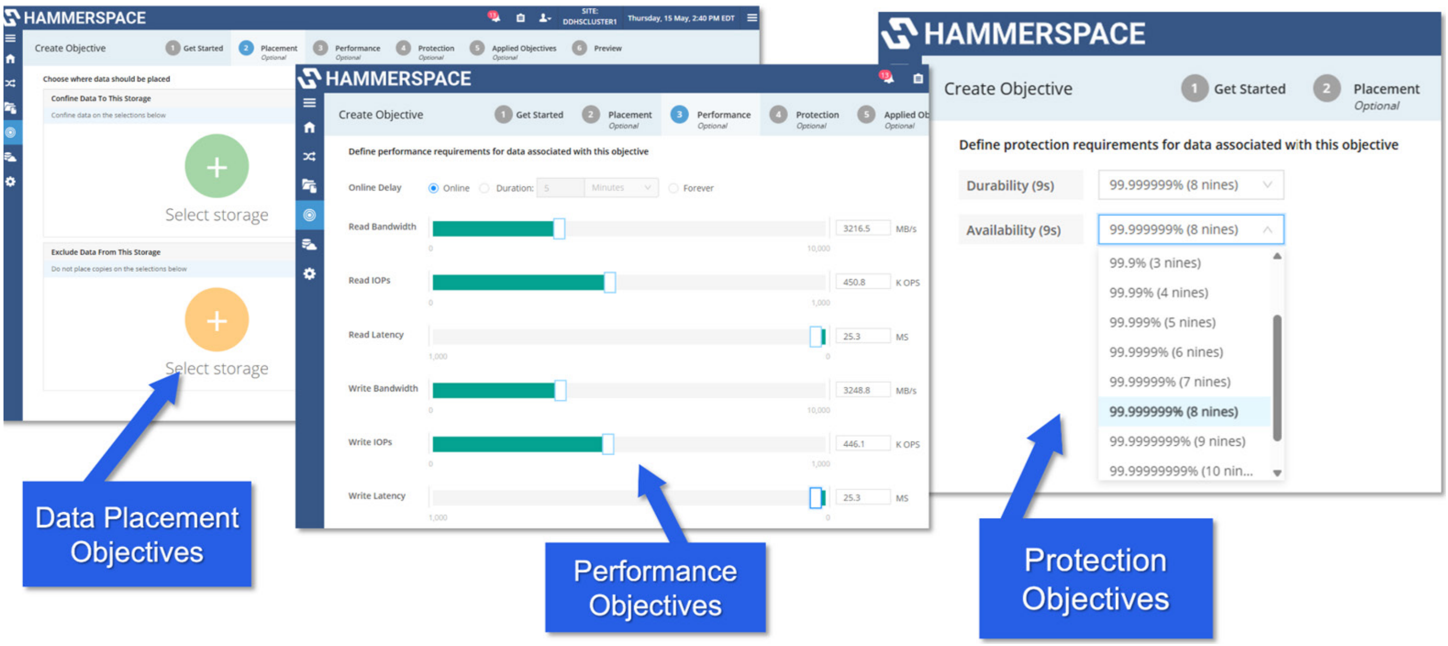


Figure 4: Hammerspace automates a wide range of data services via service-level objectives, which can be defined at a file-granular level to datasets. These apply globally, regardless of which storage the data lives on today, or moves to tomorrow.

Policies can control data placement, workflow behavior, data protection, lifecycle management, and the application of custom metadata. Because these services operate at the metadata level, they can be applied consistently across the entire data estate. This allows organizations to manage their data centrally, even as underlying storage platforms and use cases change over time.

In AI-driven environments, this policy-based control is critical. Data on existing storage can now be staged or aligned with GPU resources non-disruptively as needed, while remaining part of the same global namespace before, during, and after processing. This eliminates the need to create net-new repositories for AI pipelines simply to make data accessible.

Accelerated, Non-Disruptive Data Orchestration

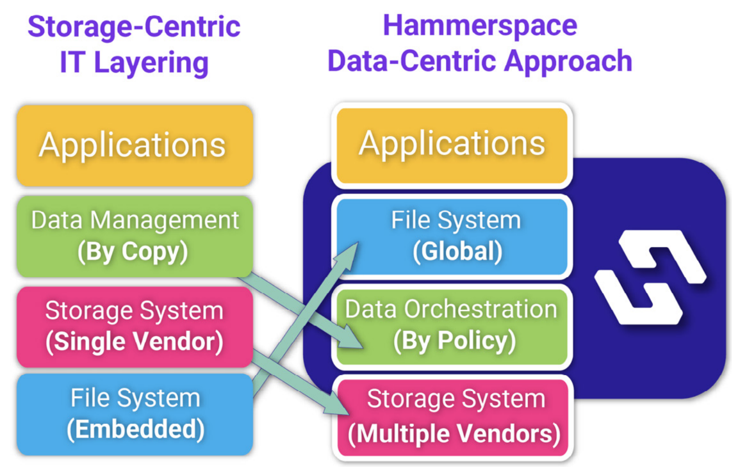


Figure 5: Hammerspace elevates the file system out of the infrastructure layer, creating a persistent unified data plane. This enables data orchestration and other data services to be automated as background operations, non-disruptive to users and applications even on live data that is actively in use.

At the core of the platform is automated data orchestration, which enables data to be moved or staged non-disruptively across storage systems, sites, and clouds. Unlike traditional copy-based approaches, orchestration in Hammerspace is metadata-driven and occurs in the background, without breaking access or creating forked copies of data.

Data can be orchestrated to high-performance NVMe-based Tier 0 storage to feed GPU clusters, replicated across sites for resiliency, or placed into one or more cloud storage platforms as needed. These operations are transparent to users and applications, which continue to see a consistent unified view of the data throughout its lifecycle.

This powerful orchestration capability allows organizations to support extreme performance requirements for AI training and inferencing, while still managing data holistically across a diverse set of storage resources.

Silo-less Support for Any Storage Backend

As a core capability of its standards-based data platform, Hammerspace supports virtually any storage type, including NVMe-based systems, traditional file storage, object storage, tape, and cloud-based storage services. Existing storage investments remain in place and continue to be used as part of a unified data environment.

By assimilating file and object metadata from data in place on existing storage, Hammerspace eliminates the need to migrate data into a new repository to gain unified access. This allows data to be accessed, managed, and orchestrated independently of the storage platforms on which it resides, enabling a truly data-centric architecture that is no longer constrained by the proprietary characteristics or limitations of individual storage systems.

Operational Flexibility for a Changing Data Landscape

By decoupling data access and control from underlying storage platforms, the Hammerspace Data Platform gives organizations a level of operational flexibility that is virtually impossible to achieve with traditional, storage-centric architectures. Data placement, performance, cost, and location can all become variables that are managed dynamically, rather than fixed decisions made at the time data is created.

Storage platforms can be refreshed, expanded, or retired as non-disruptive background operations. Data can be aligned with compute resources as needs change, whether for analytics, collaboration, or AI-driven processing. New workflows can be introduced without reorganizing data or creating additional storage silos. Platform refreshes and infrastructure changes can be executed as background operations, without requiring disruptive, brute-force data migrations.

This flexibility allows organizations to adapt existing environments to new requirements without retooling infrastructure or disrupting users.

As an organization's data usage continues to evolve, particularly with the growing demands of AI and large-scale analytics, this ability to adapt becomes critical. Rather than forcing organizations to choose between performance, cost, or accessibility, a data-centric approach allows all three to be balanced continuously across the entire data estate.

In this way, the Hammerspace Data Platform enables organizations to modernize how data is accessed, managed, and put to use, while preserving existing investments and minimizing operational disruption.

Why AI Is the Defining Workload for a Data-Centric Platform

The Hammerspace Data Platform architecture described in the previous section is broadly applicable across modern enterprise workloads. However, no workload exposes the limitations of siloed storage-centric architectures more clearly than AI.

AI-driven pipelines place sustained, simultaneous demands on data access, performance, and reuse that traditional enterprise data center environments were never designed to handle. As a result, AI becomes the most effective way to evaluate whether a data platform can truly operate as a shared, global resource rather than a collection of loosely-connected storage systems.

AI Pipelines Consume Data Differently

The same unstructured data may be used repeatedly across training, inferencing, analytics, and downstream processing, often by multiple pipelines and teams at the same time. Data that was created months or years earlier can suddenly become relevant again without warning for high-performance AI use cases.

In these environments, access patterns shift from one-to-one relationships to many-to-many relationships. Multiple GPU clusters, often distributed across sites or cloud regions, require concurrent access to shared datasets at high throughput and low latency. This makes AI uniquely sensitive to fragmentation, duplication, and siloed or inconsistent data access.

A platform that cannot present distributed data as a coherent, globally accessible resource with automated data orchestration will struggle to support AI at scale, regardless of raw storage performance.

Aligning Data with Compute, Not Infrastructure

Unlike traditional applications, AI workflows do not operate on isolated datasets with linear lifecycles.

AI also changes the economics of data placement. GPU resources are expensive, and increasingly distributed in remote cloud-based compute environments. To support new AI pipelines, compute is often provisioned where GPU capacity is available rather than where data happens to reside. In this model, forcing data to be copied, rehydrated, or reorganized each time different compute resources are required becomes untenable very fast.

The data-centric architecture of the Hammerspace Data Platform directly addresses this challenge. By decoupling data access and control from underlying storage systems, Hammerspace allows data to be aligned dynamically with compute resources wherever they are without breaking access, duplicating datasets, or disrupting other workflows. Data remains part of the same global namespace before, during, and after processing, even as it is staged to different performance tiers or locations.

This capability is not specific to AI, but the demands of AI pipelines make it essential.

Performance at Scale Without Fragmentation

AI pipelines are also unforgiving of inefficiency. If data access stalls, GPU pipelines stall. If access semantics change between locations or tiers, workflows break. Meeting these requirements with copy-based approaches typically leads to data sprawl, governance gaps, and rising operational costs. Because Hammerspace operates at the metadata level, performance and access consistency can be maintained without introducing forked views of data.

Hammerspace also uniquely includes the ability to aggregate otherwise underutilized local NVMe storage on GPU or CPU-based compute nodes as a seamless part of the global namespace. This Tier 0 storage activates existing capacity within the servers to deliver performance levels exceeding any networked storage available on the market. The data itself remains part of a unified data estate that continues to be governed, protected, and reused globally.

In this way, AI workloads serve as a practical validation of a data-centric approach. Such workloads no longer require a separate data infrastructure, but they do demand that the underlying data platform operate globally, at scale, and at high performance without resorting to unneeded data copies.

Setting the Stage for Real-World Proof

The ability to support AI pipelines without re-platforming data or adding more data silos is a direct outcome of Hammerspace's architectural choices described earlier. The following example illustrates how this approach has been applied in practice to support AI at extreme scale, using existing infrastructure and standard access methods.

The Most Common Customer Use Cases

Taking Advantage of the Hammerspace Data Platform

The most common customer use cases for the Hammerspace Data Platform stem directly from its ability to unify access to unstructured data and orchestrate it dynamically across storage systems, sites, and clouds. By creating a unified data environment, Hammerspace allows organizations to treat their existing data estate as a shared resource that can be accessed and aligned with compute resources wherever it is needed.

Rather than relying on copy-based workflows or dedicated repositories, customers rely on Hammerspace to make data on existing silos globally visible and immediately usable, while automating placement and performance through policy-driven objectives. This enables data to remain persistent and accessible throughout its full lifecycle on whichever storage tier or location is required. As requirements change, this happens without disrupting users or applications.

These capabilities are especially important for modern workloads that depend on high-performance, parallel access to large, distributed datasets. AI, analytics, and other data-intensive workflows frequently require access to data that spans locations, vendors, and tiers, including data that would otherwise be isolated in archives or remote systems. By bridging these silos into a unified data platform, Hammerspace enables such use cases without re-architecting infrastructure or introducing additional layers of complexity.

Below are some of the most common use cases for the Hammerspace Data Platform:

Artificial Intelligence and Machine Learning (Training, Inference, and Agentic AI)

Hammerspace is commonly used to accelerate enterprise AI pipelines by eliminating data bottlenecks that starve GPUs. Organizations use Hammerspace to unify distributed datasets and dynamically place data on high-performance tiers—such as GPU server-local NVMe (Tier 0 storage)—while maintaining a single global namespace across silos and locations. This ensures that training and inference workloads have low-latency, high-throughput access to data without rewriting applications or managing specialized storage silos. The result is higher GPU utilization, faster model iteration, and more responsive AI inference at scale.

High-Performance Computing (HPC) and Research

In HPC environments, Hammerspace provides parallel file system performance without the operational complexity of proprietary clients or rigid storage architectures. Research institutions and enterprises use Hammerspace to support I/O-intensive workloads such as simulations, modeling, and checkpoint/restart operations. By combining standards-based file access with policy-driven data orchestration and Tier 0 performance, Hammerspace reduces job completion times, accelerates time-to-insight, and simplifies data management across clusters and sites.

Hybrid-Cloud and Multi-Site Operations

Hammerspace is frequently deployed to enable seamless data access and mobility across on-premises, cloud, and geographically-distributed environments. Organizations use the platform to unify data across sites under a single global namespace while automating data placement, replication where needed, and lifecycle policies. This allows teams to move workloads to the most appropriate location—edge, core, or cloud—without manual data movement or workflow disruption. The value is greater agility, improved collaboration, and consistent data access regardless of where applications run.

Storage Consolidation and Infrastructure Modernization

Many organizations adopt Hammerspace as part of a broader storage modernization initiative to reduce sprawl and operational overhead. Hammerspace consolidates disparate NAS, object, and direct-attached storage into a unified global namespace, while allowing existing infrastructure to remain in place. By abstracting storage resources within a global namespace with its intelligent data orchestration layer, organizations simplify management, improve resource utilization, and extend the value of prior investments—without forcing disruptive migrations or lock-in to proprietary hardware.

Customer Examples:

The following examples illustrate how Hammerspace capabilities are applied in practice at extreme scale.

Meta

High-Performance Data Platform for Training Frontier Large Language Models

- More than 40 PB of active data
- Sustained performance to feed tens of thousands of GPUs in parallel
- Linear scalability across 1,000 storage nodes
- Standards-based architecture with no proprietary client or agent software needed
- Hybrid-cloud agility with extreme performance

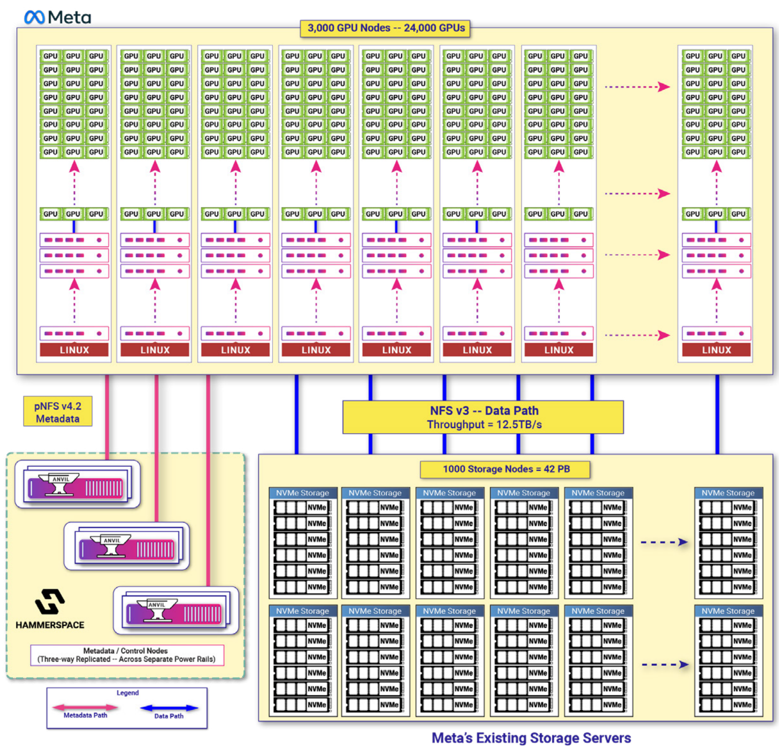
When Meta had a mandate to accelerate their AI strategy with the development of the Llama large language model, they needed a data platform with the performance to feed their massive GPU cluster from a 1,000-node storage cluster containing 42PB of data. Performance and scalability were both critical, but it also needed to be standards-based. Installing and maintaining proprietary client software across their GPU cluster was a non-starter.

After evaluating a broad selection of vendors–including using their own in-house technology–Meta selected Hammerspace to power their AI training workload for Llama 2 & 3. Hammerspace enabled Meta to achieve the extreme performance and scalability needed for this use case with commodity networking and server hardware.

The next stage of Meta's AI journey created another challenge: To maintain their pace of innovation and train Llama 4, Meta needed access to more GPU nodes than they could deploy in their own data centers. And their use of AI had expanded to include inferencing and analytics for their internal business needs, as well as continuing to train and release new versions of Llama. To solve these problems, Meta secured GPU resources in both the Oracle (OCI) and AWS clouds, and now needed a way to dynamically route workloads between on-prem and cloud-based GPUs.

Using Hammerspace, Meta was able to extend their on-premises workloads to Oracle cloud seamlessly and efficiently, without the need for data migration. Of key importance in this architecture was Hammerspace's ability to dynamically extend on-premises workloads to compute resources in one or more clouds without altering user/application workflows.

One component of this success was the capabilities of Hammerspace's native S3 connector, which is significantly different from S3 gateways or bolt-on S3 interfaces in other platforms. Rather than translating or proxying data between protocols, Hammerspace provides a unified data plane to the same data that natively supports both NFS and S3 access. Leveraging this standards-based approach, NFS-based applications can access data in S3 buckets as if they were files, with no workflow changes. Conversely, S3-native applications can access data stored on NFS directly as though they were buckets, without copying or caching.



This matters because workflows can be orchestrated seamlessly between on-premises and cloud compute resources without any change to the user experience.

As in Meta's on-premises environment, Hammerspace was able to deliver high-throughput, low-latency performance and linear scalability to feed cloud-based GPU clusters, and is now used to automate data movement and workload execution across on-premises, OCI, and AWS infrastructure.

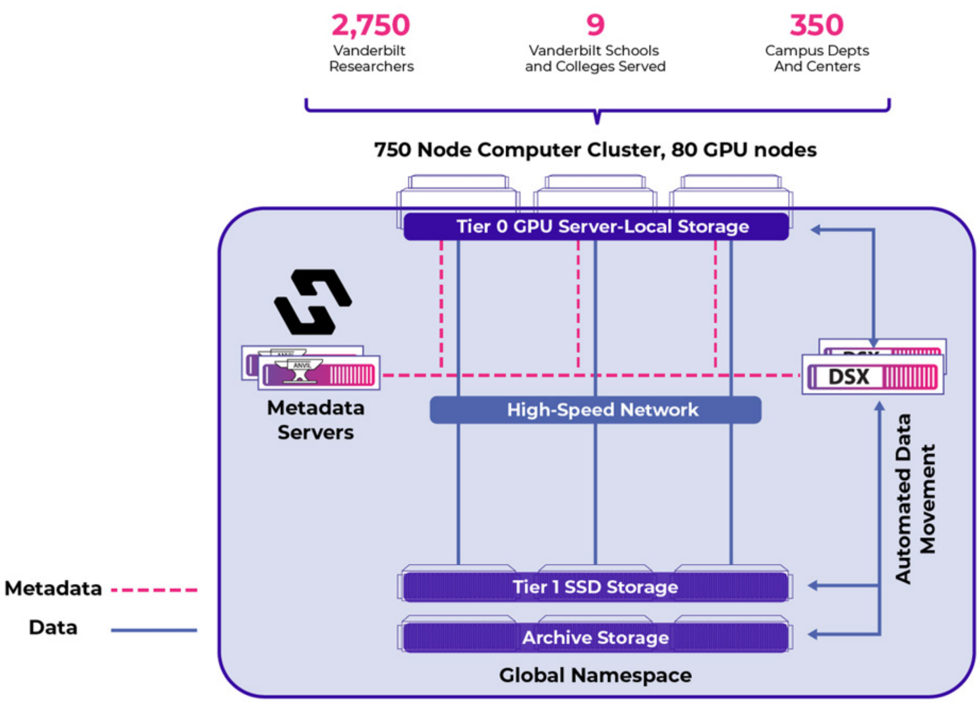
Vanderbilt ACCRE

Modern Data Platform for a High-Performance Computing (HPC) Research Center

- Unified 10 PB of storage capacity across Tier 0, Tier 1, and archival storage
- Sustained performance to feed a 750-node compute cluster, including 80 GPU nodes
- 48% reduction in storage costs
- Flexible, composable architecture for research workloads

As data volumes surged across hundreds of research projects and demand for AI and GPU-accelerated workloads expanded, the Vanderbilt Advanced Computing Center for Research and Education (ACCRE) needed a modern data platform capable of supporting next-generation scientific computing. ACCRE operates a large-scale, campus-wide HPC environment serving thousands of researchers across disciplines, but relied on fragmented storage systems for performance, capacity, and archival tiers. This drove high costs, operational complexity, and limited agility.

After evaluating multiple approaches, ACCRE selected Hammerspace to unify its research data infrastructure, reduce storage costs, and deliver storage services more dynamically to researchers. Using Hammerspace, ACCRE deployed a 10-petabyte, next-generation research data environment built on a unified global namespace spanning GPU server-local NVMe (Tier 0), new commodity storage servers, and its multi-petabyte LStore archive. Hammerspace's open, composable architecture enabled ACCRE to activate Tier 0 in GPU and CPU server-local storage for high-performance AI and simulation workloads. This approach eliminated reliance on proprietary storage appliances, and enabled them to better use resources they already had.



Storage resources can now be dynamically composed and reconfigured on a per-project basis, giving researchers high-throughput access to data regardless of where it physically resides.

Beyond performance and flexibility, economics were a critical driver. By consolidating siloed storage systems under Hammerspace and pairing commodity hardware with existing archival capacity, ACCRE is on track to reduce storage costs by 48% while establishing a scalable foundation for future consolidation.

Today, Hammerspace enables ACCRE to support diverse HPC, AI, and data-driven research workloads with greater agility, lower cost, and a unified data platform designed to scale with future generations of scientific discovery.

BUCK

Global Creative Agency Enables Collaboration at Scale Across Four Sites, Three Continents

- Unified global namespace spanning four sites across three continents
- Improved productivity of globally distributed creative teams
- Enabled BUCK to take on more projects and complete them faster
- Repurposed existing NAS systems as lower-cost capacity tiers

BUCK is a global creative company managing dozens of concurrent projects across studios in North America, Europe, and Australia, with large volumes of unstructured data flowing through performance-intensive creative workflows. Prior to Hammerspace, each location operated independently on different NAS platforms, and project data was copied manually between sites. This inhibited collaboration and limited BUCK's ability to seamlessly scale teams across geographies.

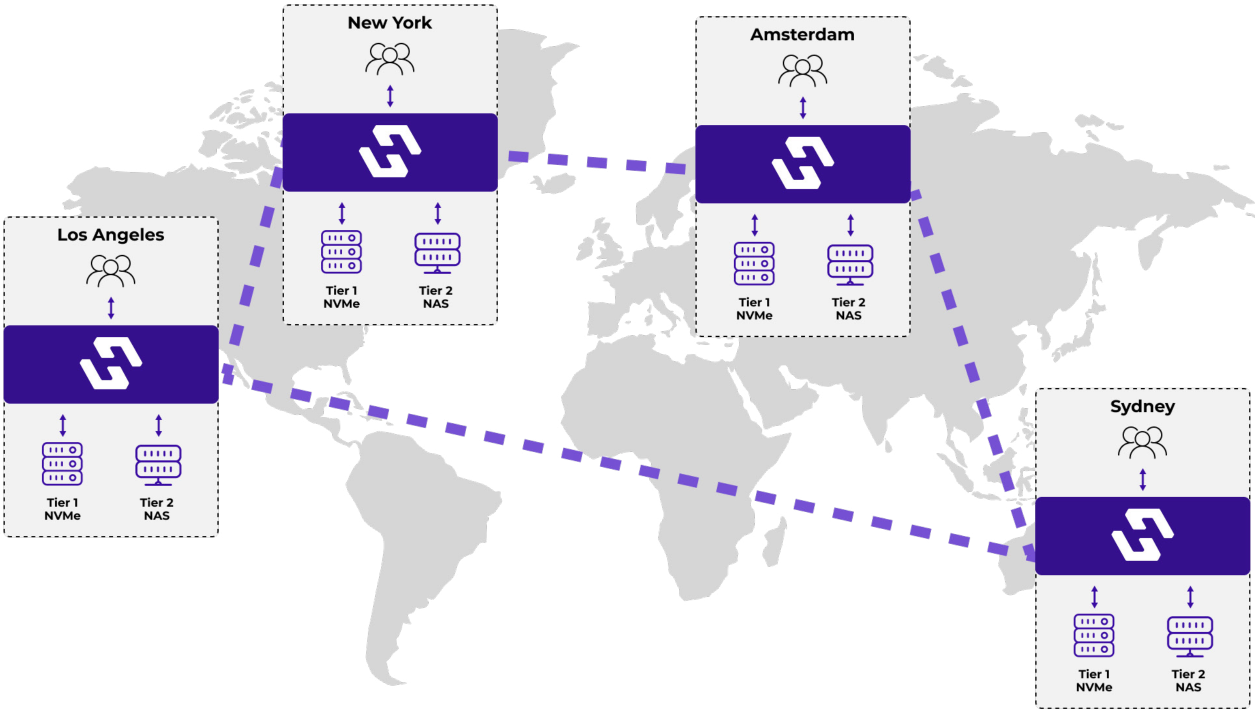
With a vision for a single, high-performance global namespace that was independent of storage appliances, BUCK selected Hammerspace to unify its data infrastructure and enable true global collaboration.

Using Hammerspace, BUCK rapidly deployed a multi-site environment spanning Los Angeles, New York, Amsterdam, and Sydney, creating a single live global system accessible by artists in all locations.

Hammerspace delivers high-performance shared storage for editing, rendering, and other performance-intensive creative workloads using standard NFS, SMB, and S3 protocols—without requiring proprietary clients or workflow changes. This allows geographically-distributed teams to collaborate on the same files in real time, enabling projects to be completed faster and allowing BUCK to staff projects globally rather than regionally.

Hammerspace also enabled BUCK to modernize its storage architecture by combining Supermicro NVMe servers for high-performance workloads with existing NAS systems repurposed as lower-cost capacity tiers, all managed as part of a unified global file system.

This storage-agnostic, data-centric approach gives BUCK the flexibility to scale efficiently, adapt to evolving creative and AI-driven workflows, and operate as a truly global organization without being constrained by the limitations of traditional, appliance-bound storage architectures.



Summary

As described throughout this paper, the challenges organizations face in managing unstructured data are not new, but the impact on their business is changing dramatically, especially with the rise of AI. Fragmentation across storage platforms, locations, and clouds has long introduced added costs and operational complexity. What has changed is how data is now used. Modern workflows increasingly require the same data to remain accessible, reusable, and performant across its entire lifecycle, often by multiple applications and distributed teams at the same time.

This shift is most visible in AI-driven workloads, but it applies broadly across analytics, collaboration, distributed processing, and other use cases. The traditional one-to-one relationship between applications and data has given way to many-to-many access patterns, longer-lived datasets, and far higher performance expectations. Architectures that depend on copy-based data movement or storage-bound file systems struggle to adapt to these demands without introducing duplication, further fragmentation, and inefficiency.

By elevating file and object metadata into a unified global data plane, organizations can create a shared data environment with their existing resources in which data can be accessed, orchestrated, and governed consistently across silos, sites, and clouds without migrating it into new repositories or disrupting users and applications.

The Hammerspace Data Platform was designed to enable this shift. By combining a standards-based parallel global file system with automated, policy-driven data orchestration, Hammerspace allows organizations to unify their existing data estate into a global resource that can be dynamically aligned with compute resources and evolving use cases. This architecture supports both traditional enterprise workloads and the extreme performance requirements of HPC and modern AI pipelines, without forcing organizations to re-architect infrastructure or abandon existing storage investments.

Addressing this challenge requires more than incremental improvements to existing storage tools. It requires a data-centric approach that decouples data access and control from the underlying storage platforms on which data resides today, or may move to tomorrow.

As data usage continues to evolve, the ability to access data globally, align it dynamically with compute, and deliver high performance without duplication will become increasingly critical. A data-centric architecture provides the foundation needed to meet these demands, enabling organizations to activate their data wherever it lives, today and in the future.

This is the Hammerspace innovation.