

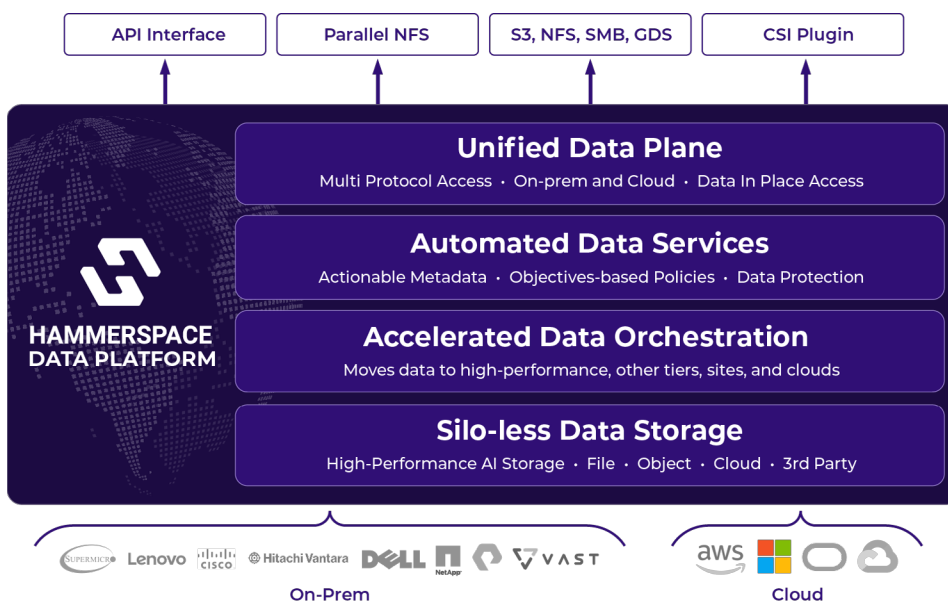
Hammerspace and Microsoft Azure

SOLUTION GUIDE

Unify, Orchestrate and Accelerate Your Data in the Cloud

Hammerspace helps Azure Cloud customers unlock the full value of their data by breaking down silos across on-premises, cloud providers, and cloud regions. Together with Azure, we make it easy to move data where it's needed, speed up AI and HPC projects, and simplify hybrid-cloud operations. The result is faster time to value, lower infrastructure complexity, and greater consumption of Azure compute and storage resources.

Hammerspace is available in the Microsoft Azure Marketplace and runs on standard Azure VMs.



Solution Benefits:

- Unify Distributed, Siloed Data with a Global Namespace
- Migrate Data from On-Prem and Other Clouds to Azure
- Run Demanding AI/HPC Workloads in the Cloud
- Simplify Hybrid-Cloud Infrastructure
- Streamline Cloud Data Pipelines and Workflows
- Drive Azure Compute and Storage Consumption



Deploying Hammerspace on Azure

Hammerspace software is deployed on standard Azure infrastructure. There are two components of the Hammerspace software architecture:

- **Metadata Servers**, called Anvils, host the metadata database, act as the control plane, and provide routing information to the clients. There are two Anvil metadata servers per cluster, in other words per site, for high-availability.
- **Data Services nodes**, or DSX nodes, are used to scale out client access across NFSv3, SMB, and S3 protocols, can act as a data store in the cloud and on-prem, and can also act as a data mover for NAS-to-NAS and NAS-to-Object data mobility. There are 2 to 60 DSX nodes per cluster depending on the performance and scale requirements.

Hammerspace can also add storage volumes from any Linux storage server and make those volumes part of the Hammerspace file system. For example, you can add high-performance storage by using Azure VMs with NVMe storage.

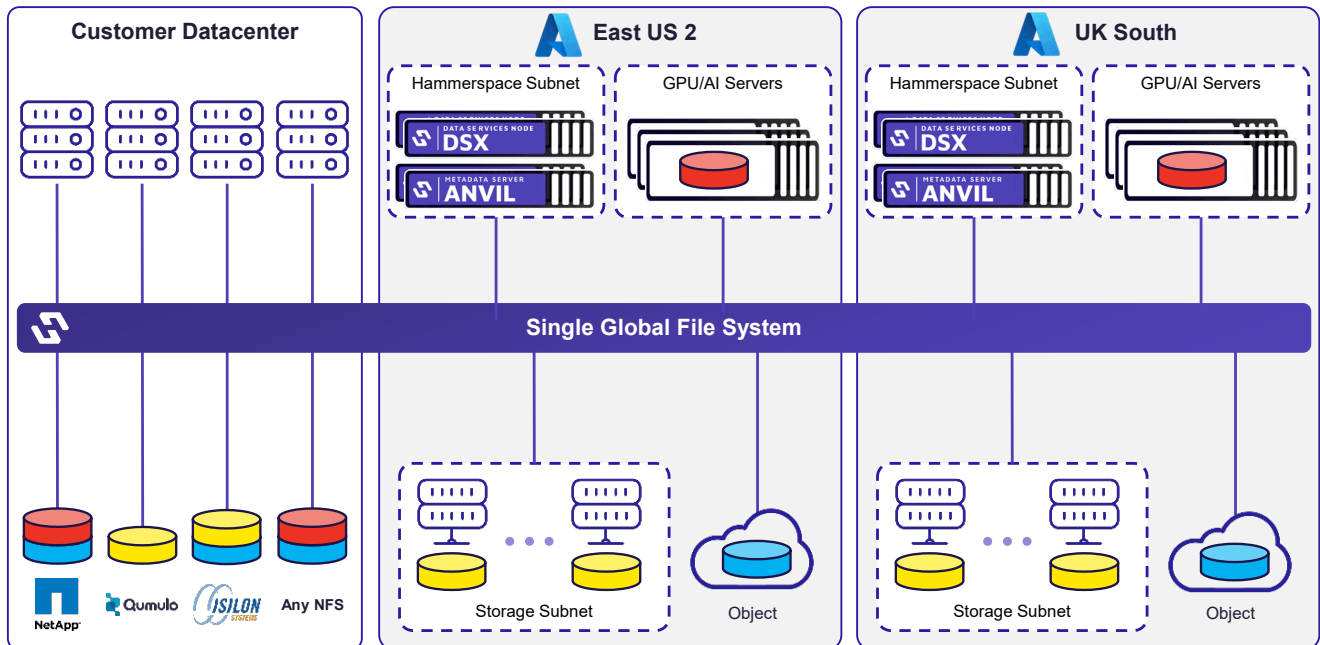
Hammerspace allows you to add the NVMe storage volumes that are local to the Azure GPU VMs, and aggregate these volumes together to create a tier of ultra-fast, low-latency NVMe storage that is called "Tier 0."

With Tier 0, files stored on these local NVMe volumes can be protected by mirroring data on different nodes, and data can also be tiered off to external storage - for example Azure blob storage - for long term storage. No matter where the files are stored physically within the Hammerspace global namespace, the files and objects remain visible and accessible via NFS, SMB, and S3. For additional information, read the [Getting Started Guide](#), and the [Hammerspace Azure Marketplace Deployment Guide](#).



Example Hybrid-Cloud Deployment

An example hybrid-cloud deployment is shown in the figure below, which shows one customer data center and two Azure cloud regions connected with a Hammerspace global file system.



In this example, the Hammerspace metadata servers are deployed on (2) Azure VMs, and the DSX nodes are also deployed on (2) Azure VMs. Each of the Azure cloud regions have GPU VMs that are connected to the Hammerspace Data Platform.

The local drives in these VMs are used as a Tier 0 layer for the applications running on those VM's. A Tier 1 storage layer is also being used. These are Linux VM's with internal NVMe drives in them. And finally, Azure blob storage is used as a tiering bucket as well as the shared object store used to transport data between sites.

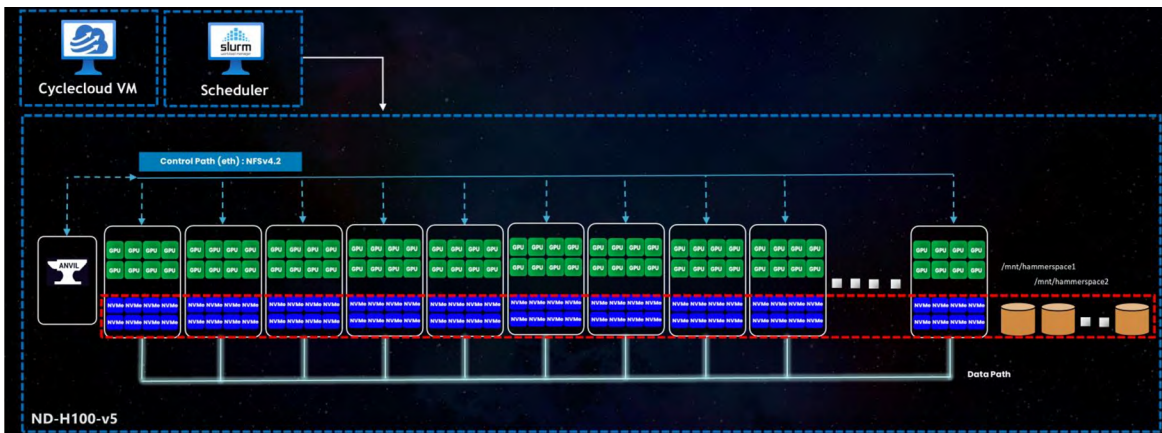
Performance Benchmarking on Microsoft Azure

To validate Hammerspace Tier 0 performance at scale, Hammerspace partnered with Microsoft Azure engineering to conduct joint benchmarking on a large Azure GPU cluster earlier this year. The goal was to measure both raw I/O throughput for traditional HPC workloads and real-world checkpointing performance for generative AI training jobs.



Test Environment

- **Cluster:** 128-node GPU cluster built on **ND H100 v5** instances, Azure's most advanced AI/HPC-optimized VM family.
- **Data Platform:** Hammerspace Data Platform, deployed natively on Azure infrastructure to unify and accelerate data access.
- **Orchestration:** **Azure CycleCloud** for cluster management and **Slurm** as the job scheduler.

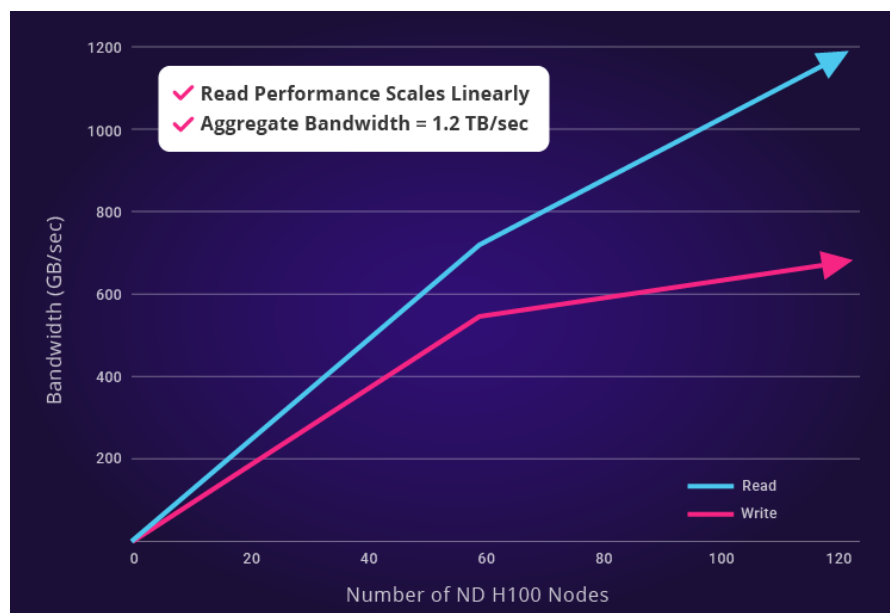


Benchmarks Conducted

IOR Performance Benchmark: The industry-standard IOR benchmark was used to measure read/write throughput across the parallel file system.

IOR simulates the access patterns of HPC applications, stressing the file system with both sequential and parallel I/O.

Results showed **linearly scalable read performance to 128 nodes**, achieving an aggregate throughput of **1.2 TB/sec**.

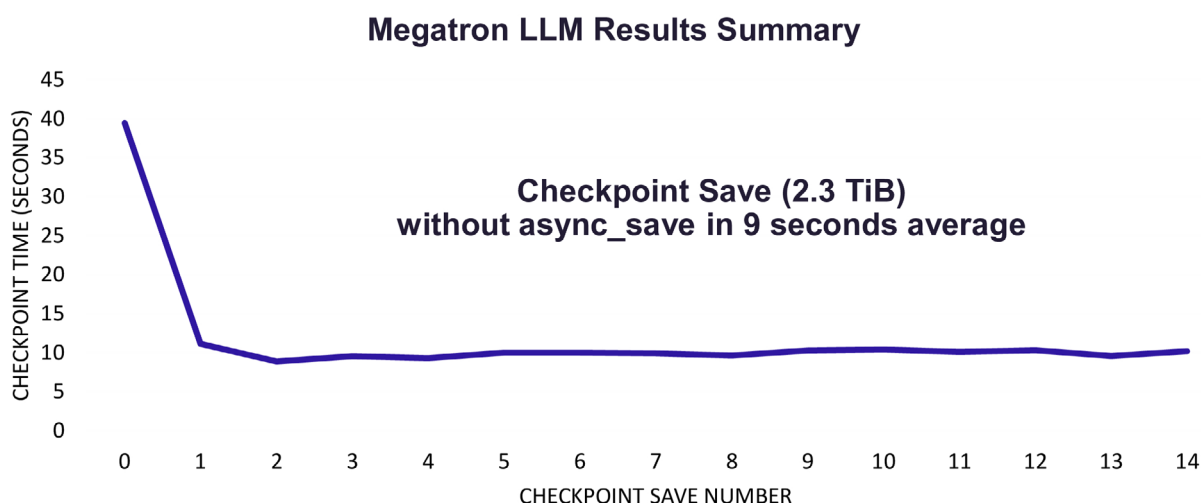


Megatron LLM Training: To capture performance in an AI-native use case, we trained the open-source **Megatron Large Language Model** on the same ND H100 v5 cluster.

Testing focused on **checkpointing**, a critical factor in training efficiency where GPUs often idle while model state is written to storage.

Hammerspace Tier 0 delivered rapid checkpoint writes – **an average of 9 seconds for a 2.3 TiB checkpoint write** – reducing checkpoint times significantly compared to traditional storage architectures and minimizing GPU idle cycles.

Based on these tests, checkpointing introduced only ~9 seconds of overhead every 1000 iterations (1.94 hours of compute), resulting in **>99.9% time efficiency** attributable to storage performance. Actual utilization will vary based on checkpoint frequency, model size, and training configuration.



Key Outcomes

- **Proven at Scale on Azure:** Validated on 128 Azure ND H100 v5 instances, ensuring confidence for customers deploying large AI/HPC clusters.
- **HPC & AI Ready:** Delivered top-end throughput on IOR while also optimizing real AI workflows with reduced checkpoint overhead.
- **Cloud-Native Simplicity:** Deployed directly on Azure using CycleCloud and Slurm, providing a standards-based environment that mirrors on-prem HPC infrastructure.
- **Optimized GPU ROI:** Faster checkpointing keeps GPUs productive, enabling customers to maximize utilization of their Azure ND H100 v5 investments.

By combining HPC-proven throughput with AI-specific workflow acceleration, these benchmarks confirm that Hammerspace on Microsoft Azure delivers the performance and efficiency required for the most demanding workloads.

