Hammerspace MLPerf® Storage v2.0 Benchmark Results

Delivering the best price/performance storage for GPU computing

TECHNICAL BRIEF

Summary

The MLCommons MLPerf Storage benchmark is designed to demonstrate the performance of data storage platforms supporting AI/ML workloads. The aim is to provide AI system architects and decision makers with useful data to evaluate storage performance for machine learning, deep learning, and other forms of GPU computing.

Hammerspace has been involved with MLPerf Storage since the beginning, submitting results for v1.0 of the benchmark in 2024, and v2.0 in July 2025. This technical brief includes all these results and context around them, including:

- Background on MLCommons and the MLPerf Storage Benchmark
- A summary of Hammerspace's results, and the hardware and software configurations used to achieve them
- Discussion of Hammerspace results relative to other participating vendors

The results highlight the performance, simplicity, efficiency and flexibility that make Hammerspace the data platform for AI anywhere.

The Hammerspace Advantage

Results achieved with both versions of the MLPerf Storage benchmark demonstrate that Hammerspace achieves HPC-level performance using the standard networking and interfaces of enterprise storage, both on bare metal and in the cloud, and does so with incredible efficiency.

Why is this important? Parallel file systems are essential for the demands of high-performance computing. But these file systems originated in the academic and scientific research communities, which operate under unique priorities and constraints. As a result, traditional parallel file systems are not enterprise IT friendly.

- They require specialized client software for every server accessing the storage
- Applications need to be designed to operate with those proprietary interfaces
- Exotic networking such as InfiniBand or Slingshot is needed for performance
- Such systems can be fragile, requiring high levels of specialized care to avoid downtime
- They are complex to install and tune, taking weeks or more to ready for production

These complexities have led many organizations to select scale-out NAS platforms as an easier-to-use and more familiar alternative. Unfortunately, such systems require twice the number of servers and twice the number of network ports compared to parallel file systems that create a direct data path between clients and storage, and they cannot deliver the performance required for HPC-class workloads at scale.

Hammerspace brings the best of these technologies together—while overcoming the negative challenges of each—to deliver the best price/performance storage solution for GPU computing. Hammerspace does this by using:

- **At least 50% fewer servers and 50% fewer network ports** than scale-out NAS architectures such as Dell PowerScale, Qumulo, and VAST. The corresponding power and cooling savings frees up wattage for the compute environment
- **Standard Ethernet connectivity**, eliminating the need for a specialized second network such as InfiniBand, which is used by other MLPerf Storage participants including DDN, HPE, and WEKA
- **Existing Linux client servers** to connect to the file system without proprietary client software or specialized hardware
- **Standard NFS and S3 interfaces** which work natively with existing applications

For the most recent runs, Hammerspace demonstrated another leap in simplicity and efficiency, by utilizing a Tier 0 implementation that activates local NVMe drives across a cluster of client machines instead of relying on networked external storage servers. Unifying local drives within a cluster as the shared high-performance storage tier, and mobilizing files with parallel file system efficiency and locality-aware data placement, Hammerspace reduces the required hardware footprint and conserves even more wattage for GPUs.

For more information on the architecture and features of the Hammerspace Data Platform, visit https://hammerspace.com/.

About MLCommons® and the MLPerf Storage Benchmark

> *"MLCommons is an Artificial Intelligence engineering consortium, built on a philosophy of open collaboration to improve AI systems. Through our collective engineering efforts with industry and academia we continually measure and improve the accuracy, safety, speed, and efficiency of AI technologies–helping companies and universities around the world build better AI systems that will benefit society."* (from https://mlcommons.org/about-us/)

The MLPerf Storage Benchmark Suite consists of several simulated AI workloads for model training and checkpointing. Hammerspace has run two of the training workloads:

- **3D U-Net:** A visual ML workload segmenting 3D medical imagery. This is a bandwidth-intensive test that opens large files in small batches and reads them sequentially.
- **ResNet50:** A deep learning convolutional neural network that excels at image classification – detecting objects within images and classifying them accordingly. This test involves concurrently reading many smaller (~100KB) samples from within a large number (>1000) of larger (>100MB) files. Compared to 3D U-Net this workload consists of smaller, more random I/O.

For each test, the goal is to demonstrate the maximum number of simulated GPUs ("Accelerators" in MLPerf terminology) that the storage system can simultaneously supply with data, keeping the utilization of every simulated GPU at 90% or higher. Total throughput is also reported.

Keep in mind that the MLPerf Storage benchmark is about the storage, not the clients. Clients exist only to provide the load on the storage system, and the benchmarks are constructed to enable this to be done in a very cost-effective way. This is why GPUs are simulated vs. requiring the use of real GPU hardware, and why you will see clients that simulate far more GPUs than can fit in a real server. MLPerf clients are not standardized and do not correspond to real-world application servers.

There are two divisions defined by MLCommons for testing, Open and Closed. Closed division rules establish a level playing field by requiring the use of a defined set of benchmark tuning parameters and options. This is intended to enable apples-to-apples comparisons between storage systems. Open division submissions have more flexibility for tuning both the benchmark and storage system configuration, and results are not directly comparable. The Open division is designed to allow vendors to showcase unique approaches or features that provide benefits when running AI/ML workloads.

MLPerf benchmarks have several unique aspects. One is that they are time limited. Results submitted during a defined time window that successfully complete the review process are considered "verified" results. Results submitted outside this window are not subject to review and are considered "unverified." During the pre-publication review, every participant may examine and question the results of every other participant. This environment of "coopetition" ensures the results are trustworthy.

Hammerspace has completed multiple rounds of testing, summarized in the table below. The most recent results are listed and discussed first. As new tests are run, they will be added to this document.

Benchmark Version	Workload	Simulated GPU	Division	Verification	Location	Implementation Type
2.0	3D U-Net	H100	Closed	Verified	On-Prem	pNFS Tier 0
1.0	3D U-Net	H100	Open	Unverified	On-Prem	pNFS Tier 1
1.0	3D U-Net	H100	Open	Unverified	On-Prem	pNFS Tier 0
1.0	3D U-Net	A100	Closed	Verified	Cloud	pNFS Tier 1
1.0	3D U-Net	H100	Closed	Verified	Cloud	pNFS Tier 1
1.0	ResNet50	A100	Closed	Verified	Cloud	pNFS Tier 1
1.0	ResNet50	H100	Closed	Verified	Cloud	pNFS Tier 1

MLPerf Storage v2.0

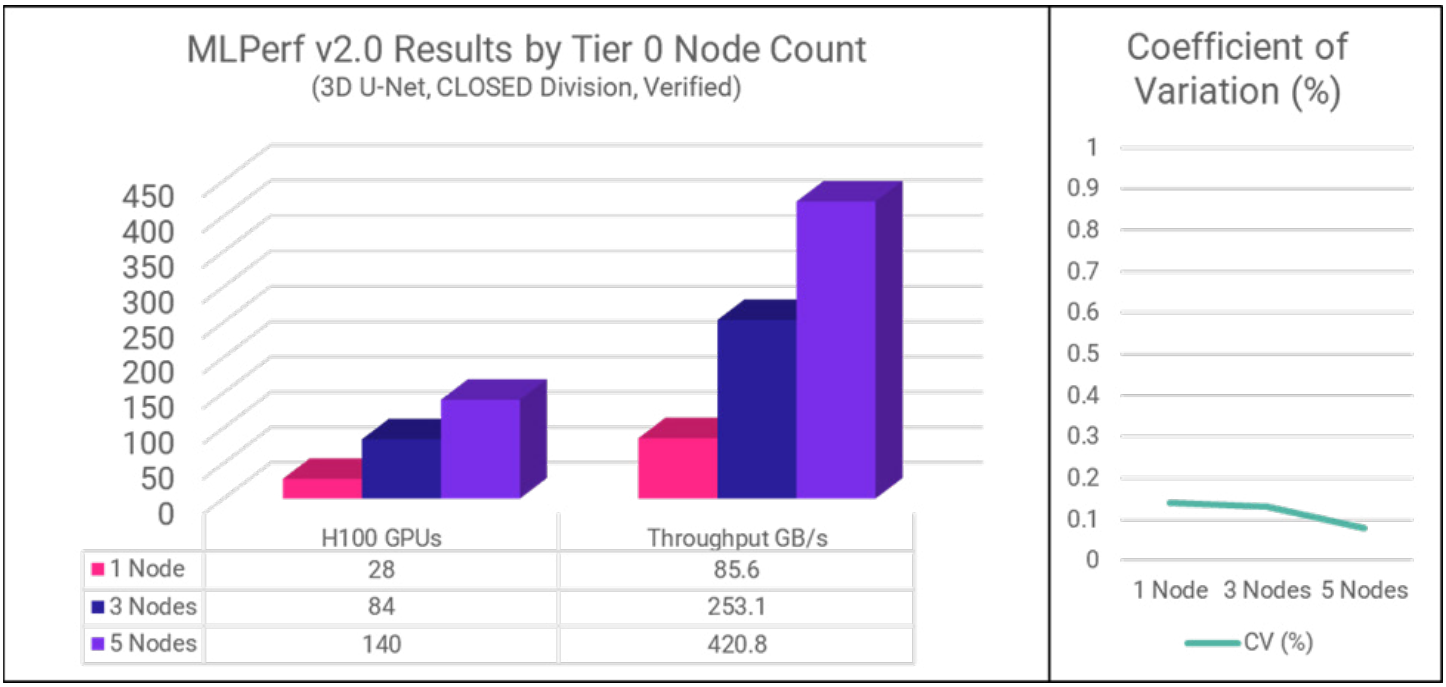
For the most recent version of the MLPerf Storage benchmark, Hammerspace focused on demonstrating the high performance and unmatched efficiency of Tier 0.

Hammerspace Tier 0 aggregates the local NVMe storage already present across a cluster of compute servers into the Hammerspace shared file system. This parallel global file system unifies this local NVMe across multiple GPU servers with external file and object storage, and cloud instances to create a unified global namespace. It uses pNFSv4.2 with FlexFiles and locality-aware orchestration to intelligently place and protect data. The result is globally shared data accessible at local PCIe bus speeds. Tier 0 enables organizations to activate the AI-ready infrastructure they already own, significantly reducing the cost and time required to get AI projects started.

Test Results

For this round, the 3D U-Net benchmark was chosen, using simulated H100 GPUs. It's the most bandwidth-intensive of the MLPerf Storage benchmarks, highlighting parallel I/O throughput as well as memory and CPU efficiency. Three configurations were tested, with one, three, and five Tier 0 nodes, respectively. The table and graph below summarize the results.

Tier 0 Node Quantity	H100 GPUs Supported	Total Throughput	Mean GPU Utilization	Coefficient of Variation
1	28	85.6 GB/s	94.7%	0.14%
3	84	253.1 GB/s	95.0%	0.13%
5	140	420.8 GB/s	96.4%	0.08%



Notice that both the number of GPUs supported and throughput scale linearly as the number of Tier 0 nodes increases. This demonstrates the full capabilities of the best case scenario, where the primary dataset can reside entirely on the host. As the cluster scales, peak performance will be dependent on the system configuration and the proportion of locally resident data, while aggregate performance will continue to grow. Because Tier 0 nodes are the compute nodes, and compute nodes typically outnumber storage nodes by a wide margin, the Tier 0 approach is more scalable than any traditional storage silo.

Mean GPU utilization indicates the percentage of time the GPUs are being kept busy vs. waiting. To 'pass' the MLPerf Storage benchmark, all GPUs must be kept at 90% or higher utilization. Higher is better, since the goal is to minimize GPU idle time.

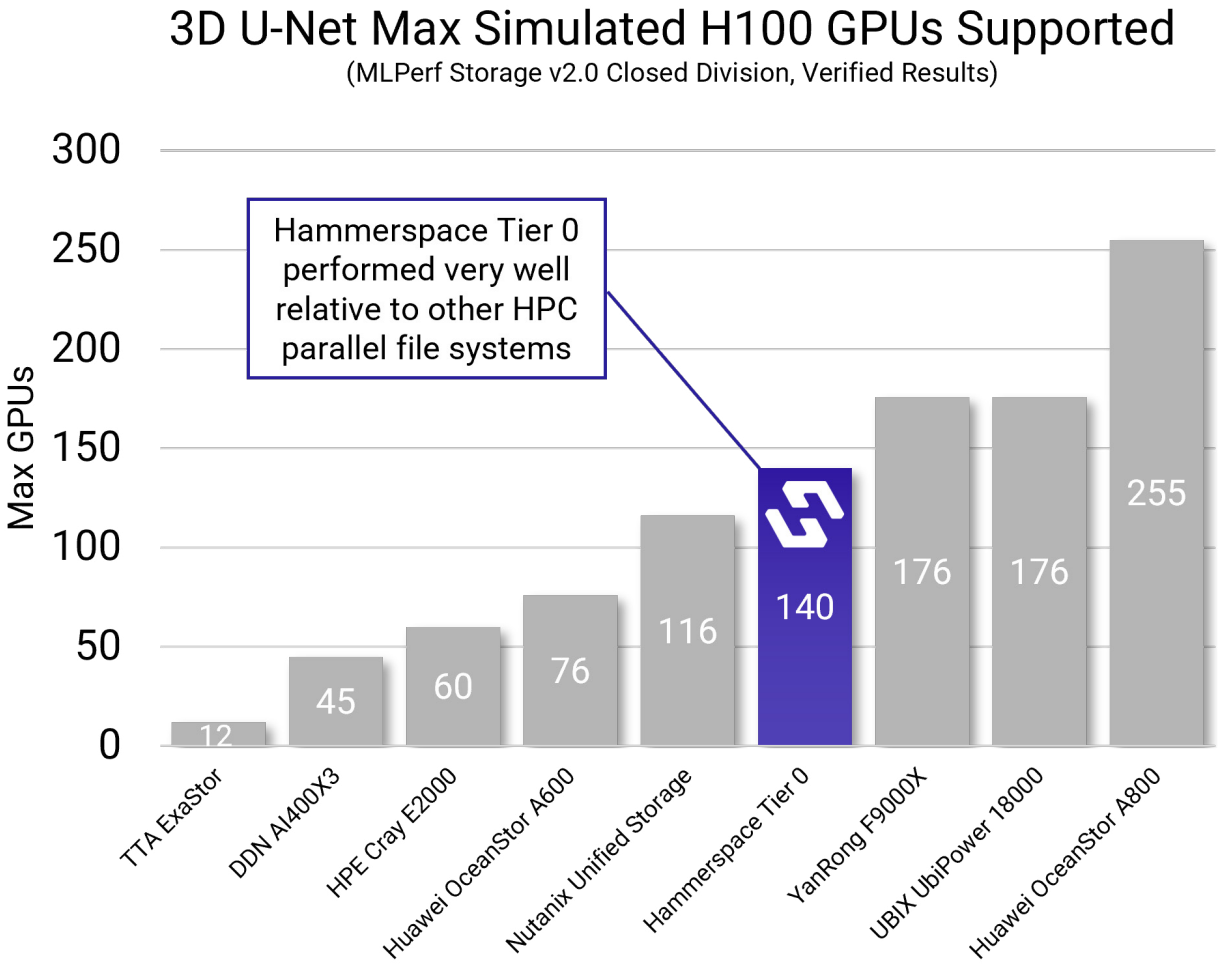
Coefficient of variation (CV) is a measure of the difference in the results between multiple runs of the same test. The MLPerf Storage benchmark requires that each test be run five times, with results falling within 5% of each other. This ensures that results are truly reproducible. The very low CV achieved by Hammerspace indicates that system performance was both stable and predictable.

One factor positively impacting these results was an optimization in the NFS data path called LOCALIO. This optimization engineered by Hammerspace is present beginning in Linux kernel v6.12 and was further enhanced in the v6.14 kernel. LOCALIO enables the NFS client and server to detect when they are running on the same host. When this is the case, the client essentially gets direct access to the local file system, bypassing the network stack and NFS server. This lowers the latency for every local storage transaction.

It's also worth noting that in v2.0 of the MLPerf Storage benchmark, the 3D U-Net workload was enhanced to allow direct IO, another optimization that Hammerspace took full advantage of.

Results Competitive Comparison

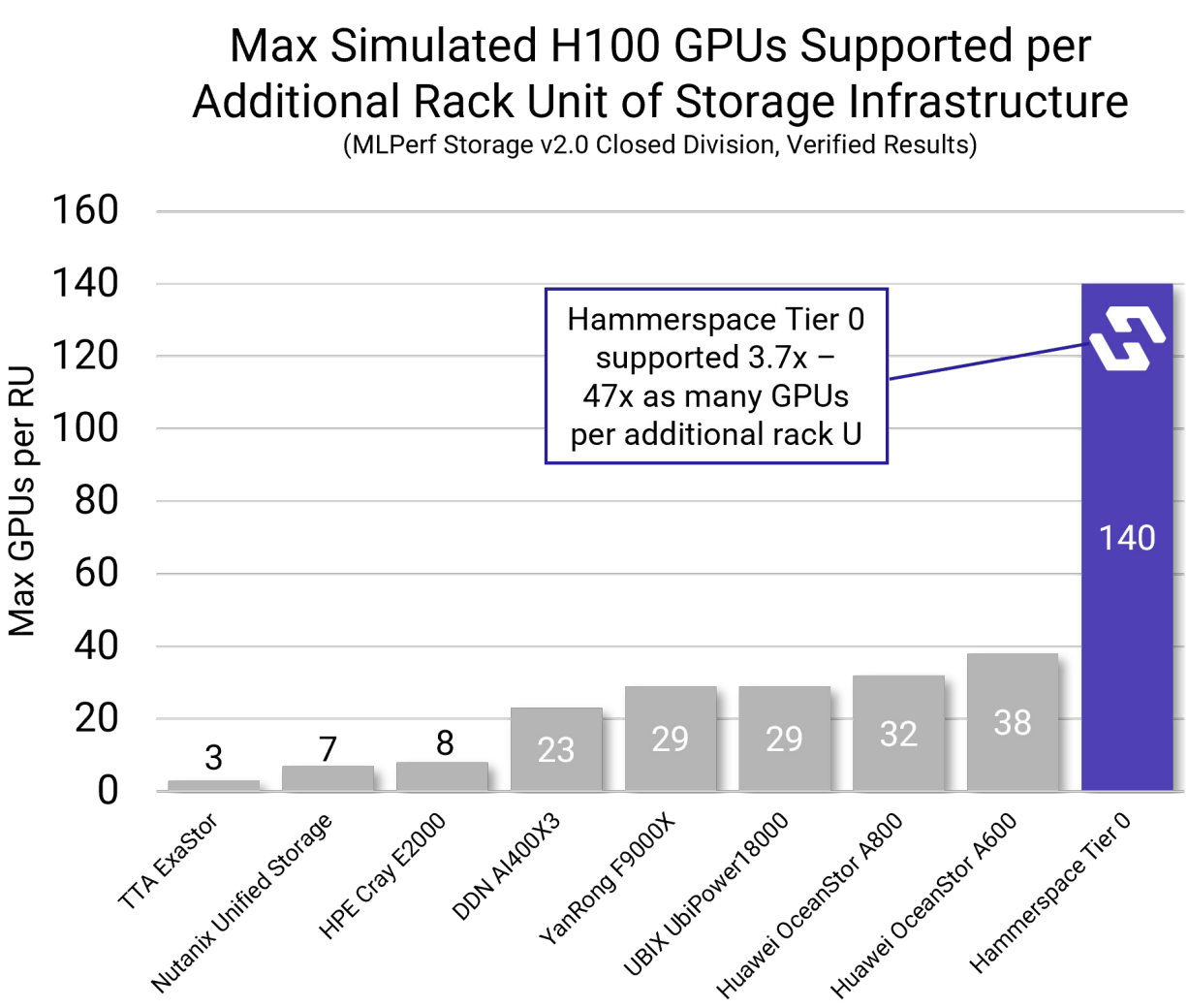
To ensure meaningful and fair comparisons, the following discussion includes only vendors who performed the 3D U-Net H100 test using on-premises shared file configurations. This graph shows the best result submitted by each vendor in terms of the number of GPUs supported:



As you can see, Hammerspace Tier 0 delivered an excellent result—better than many other parallel file system participants. However, there is another way to look at this data that is just as revealing and relevant – through the lens of efficiency.

Datacenters everywhere are constrained by power, cooling, and often rack space, and AI workloads—requiring power-hungry GPU servers—magnify the problem. Every watt dedicated to storage infrastructure is one less that's available for GPUs. In short, efficiency matters.

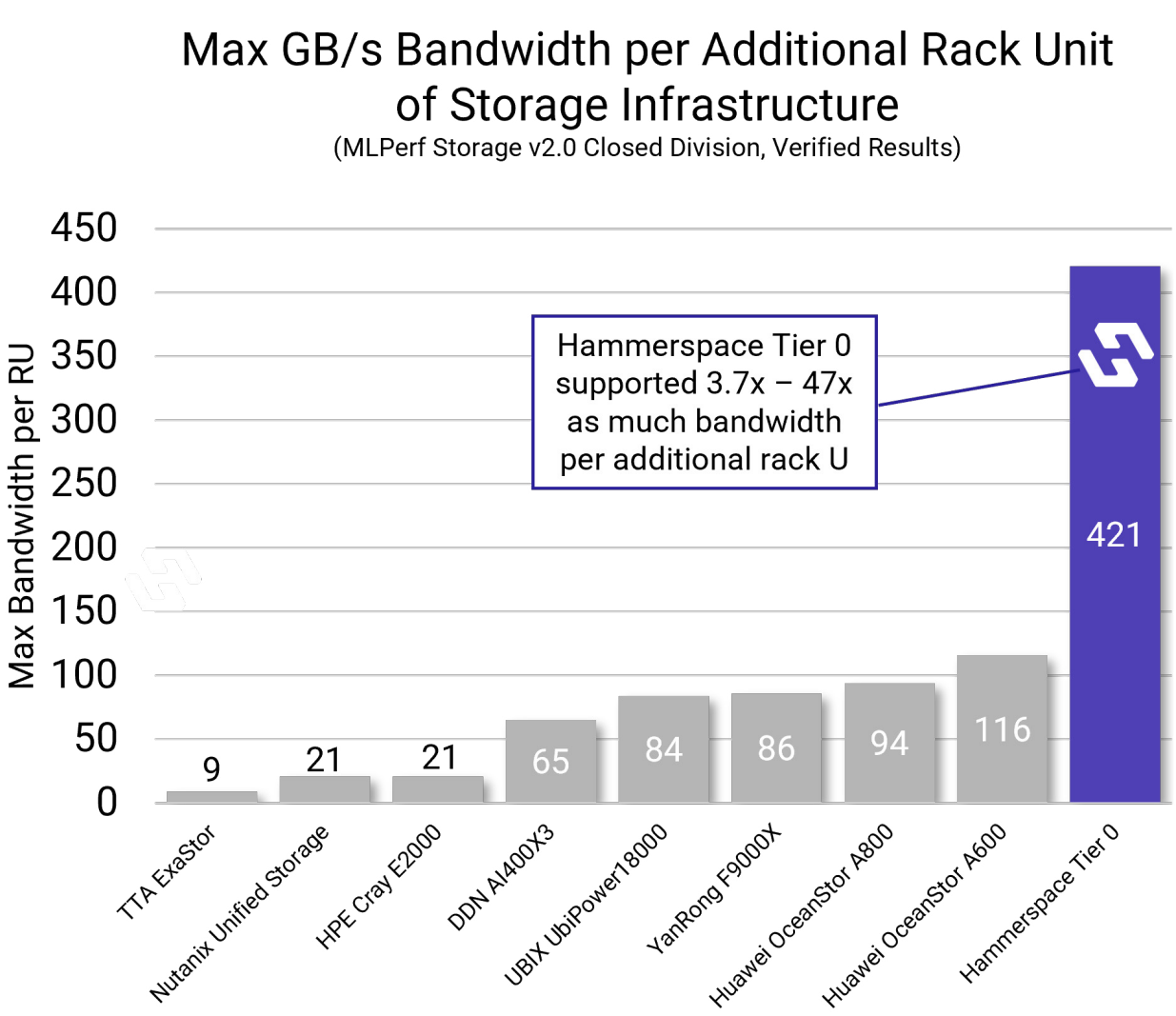
While actual power dissipation figures are not available for the MLPerf Storage submissions, rack units can serve as a useful proxy: the more rack units a solution requires, the more power it is likely to use.



When you look at the number of GPUs supported per additional rack unit of storage infrastructure, Hammerspace Tier 0 stands above the rest, with a result 3.7x that of the next most efficient system, and a massive 47x that of the least efficient participant.

In a real production deployment, GPU servers (for which benchmark clients are a proxy) run AI workloads. “Additional rack unit of storage infrastructure” refers to the additional space taken by the storage solution, over and above the compute servers/benchmark clients.

Because the storage for Tier 0 resides inside the GPU servers, the only added hardware needed for this benchmark run was a single 1U Hammerspace Anvil metadata server. In production it’s typical to run two Anvils for high availability, but even then Hammerspace would be 85% more efficient than the next best entry, supporting 70 H100 GPUs per U vs. 38. Looking at max GB/s bandwidth reveals a similar story: Hammerspace Tier 0 is 3.7x as efficient as the next best entry.



Why Tier 0 Matters

The benefits of Tier 0, including the resource efficiency described above, are not just academic. They can have a dramatic positive impact on AI projects. Tier 0 benefits include:

Simplicity:

- Get started with the storage and network infrastructure that's already in place
- No client or agent software to install
- No special networking, just Ethernet

Performance:

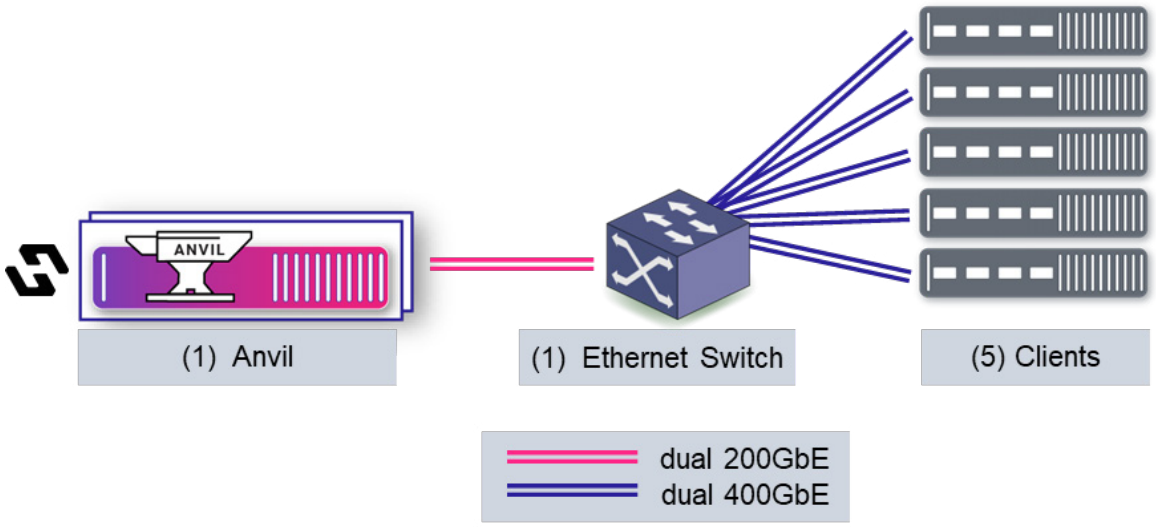
- Tier 0 storage is up to 10x faster than networked storage
- Tier 0 increases performance both on premises and in the cloud
- Tier 0 performance scales with compute
- Increased GPU utilization, faster checkpoints, reduced inferencing times

Efficiency:

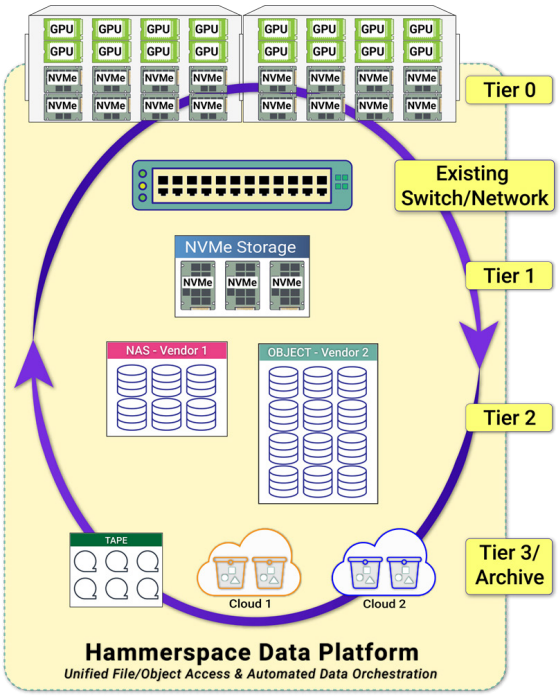
- Less external shared storage needed
- Less power, rack space, and networking vs. external shared storage
- Faster time to value – activate Tier 0 in hours, not days or weeks

Test Configuration

The diagram below illustrates the test configuration. Clients run the benchmark code. With Tier 0, they also house the NVMe drives and thus act as storage servers. The Anvil is responsible for metadata operations and cluster coordination tasks – no data flows through it. Clients mount the shared file system via parallel NFS (pNFS) v4.2, accessing the storage directly after receiving a layout from the Anvil.



The benchmark configuration is a bit artificial in its limited scope. Typically, Tier 0 is just one tier of shared, persistent storage in a more comprehensive Hammerspace infrastructure that may include network-attached Tier 1 NVMe, object storage, and more across multiple sites and clouds.



MLPerf Storage v1.0 (Open Division)

The tests reviewed in this section are in the Open division and the results are not verified by MLCommons Association, since the tests were run after the v1.0 benchmark submission window had closed.

Testing was performed in the Hammerspace performance test lab on physical servers, using the MLPerf Storage 3D U-Net workload. To highlight the positive performance benefits of Tier 0, two configurations were used. The first (configuration A) was a typical pNFSv4.2 Tier 1 NAS configuration using dedicated Linux Storage Servers (LSS). Two runs were completed with this configuration, one using 200GbE client connectivity, and one using 400GbE client connectivity. The second configuration (configuration B) used a Tier 0 configuration, where the storage devices reside within the clients themselves.

With Tier 0, the NVMe drives in each client machine are promoted from islands of underutilized and isolated local storage to become unified across the cluster as a new tier of ultra-fast storage that is part of the Hammerspace parallel global file system. An NFS protocol bypass (LOCALIO), engineered by Hammerspace and contributed upstream to standard Linux provides performance acceleration. Because Hammerspace leverages the software already built into the Linux kernel, it consumes only a tiny amount of CPU utilization, leaving more server resources for running business workloads.

To learn more about the benefits and applications of this new tier of ultra-fast shared storage, see the Tier 0 page on the Hammerspace web site.

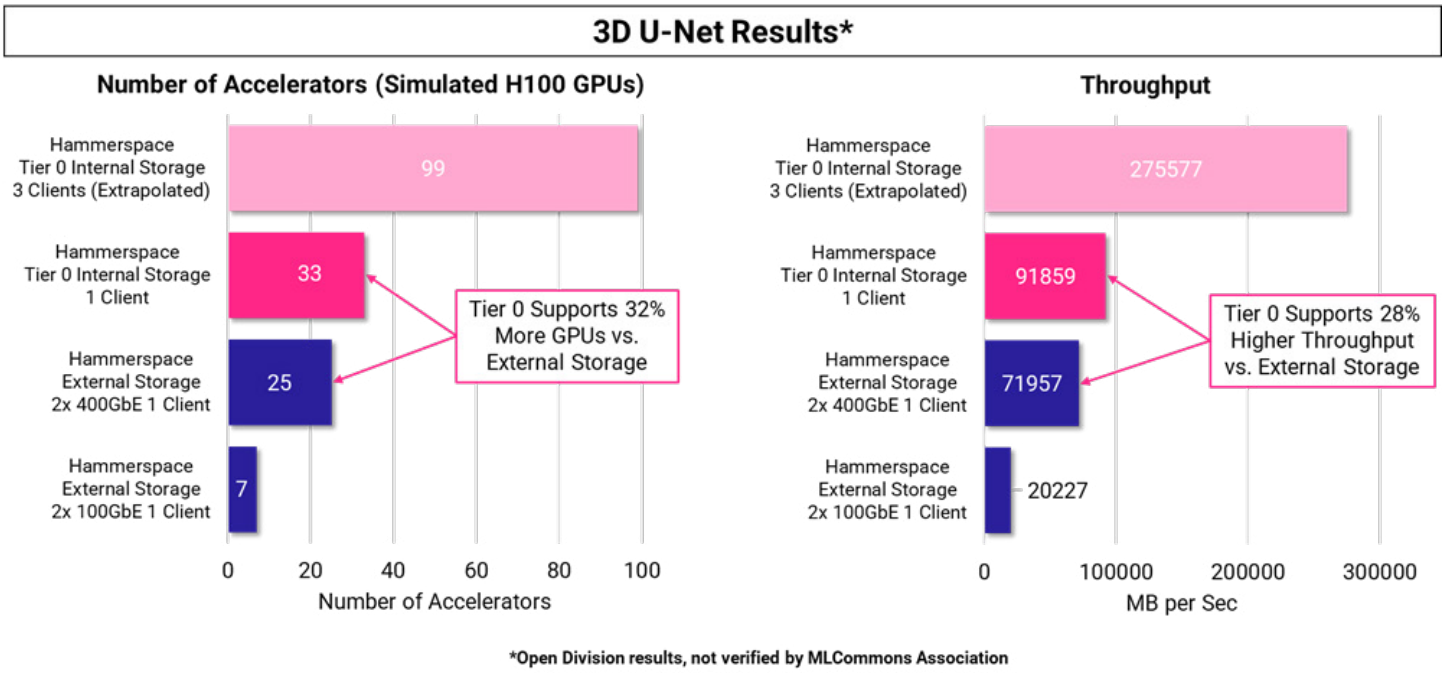
Test Results

The graph below summarizes the results of these tests.

Key Observations

32% more GPUs were supported, and 28% higher throughput was obsetved when running the 3D U-Net test using Tier 0 storage internal to the GPU server vs. extemal shared storage. This was true even when the GPU server was connected to the external storage with two 400GbE intetfaces.

Tier 0 provides meaningfully higher performance compared to Tier 1 networked external storage, even when the GPU server has high-bandwidth connections to the extemal storage, and especially if network bandwidth is constrained, as was the case with the client connected via 2x I00GbE.



The results of these tests show clear benefits to using Tier 0. There are several observations worth highlighting.

Tier 0 eliminates network bandwidth constraints

It is common knowledge that high-performance storage is required to keep GPUs busy processing data. The test results show that high-speed networking is also critical. Based on the jump from 7 to 25 GPUs between the 2x 100GbE and 2x 400GbE clients (and a commensurate jump in throughput), it's obvious that the 100GbE interfaces were a serious bottleneck.

The only thing better than a fast network is no network. Eliminating the network and using storage local to the GPUs provides the best possible performance. As shown in the graph, using Tier 0 local storage enabled the client to support 32% more simulated H100 GPUs than it could when accessing storage over 2x 400GbE, with 28% higher aggregate throughput.

In practical terms, this means GPU server clusters with underutilized local NVMe storage and sub-optimal networking, Hammerspace software can provide the performance benefits of a major network infrastructure upgrade without the cost or headaches of replacing existing NICs and switches.

Tier 0 performance is linearly scalable

Tier 0 enables GPU servers to work on data that is stored locally as a shared resource across the cluster. Hammerspace data orchestration is used to place data onto Tier 0 storage, protect that data, and offload checkpoint files and computation results to other tiers of storage. Because processing is local, performance scales linearly as more GPU servers with Tier 0 storage are added to the cluster.

On the graphs below, the dashed outline shows this linear scaling potential as an extrapolated result. If three clients with Tier 0 storage were used for this test, the number of supported GPUs and aggregate throughput would triple vs. the single-client results. See the Explanation of Verified Results vs. Extrapolated Results section below for a thorough explanation of the use of extrapolated results.

Tier 0 provides Capex and Opex benefits

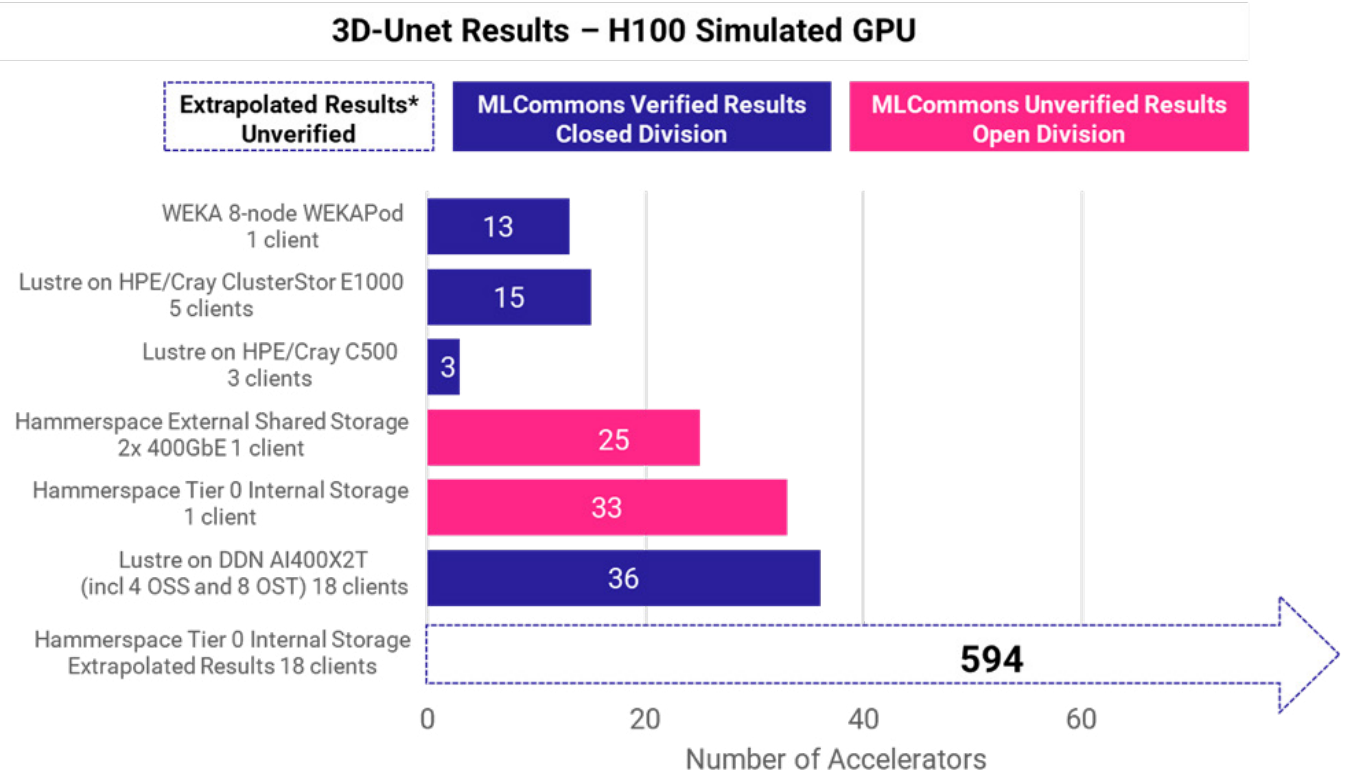
With Tier 0, Hammerspace unifies existing GPU-local NVMe storage across a server cluster into a global shared file system, eliminating the obstacles that otherwise make it impractical to use. This has numerous financial benefits.

- **External Storage Offset:** Using local NVMe storage reduces the amount of external network-attached Tier 1 high-performance storage needed, along with the associated network, power, and cooling expenses.
- **Operational Time Savings:** Hammerspace software enables existing storage to be used in minutes, freeing up time that would otherwise be spent installing and configuring external storage and networking hardware.
- **CPU Efficiency:** Unlike traditional parallel file systems that require resource-hungry proprietary client software, Tier 0 operates with virtually zero CPU overhead. This preserves more server resources for running business workloads.
- **Increased GPU Efficiency:** Tier 0 reduces checkpointing durations from minutes to seconds. This unlocks significant additional GPU compute capacity, enabling jobs to be completed more quickly without investing in additional hardware. A full analysis of the checkpointing use case is here: https://hammerspace.com/whitepaper-tier-0-checkpointing/

Results Competitive Comparison

The comparisons below show the maximum number of accelerators delivered by each vendor, independent of the number of clients used. This methodology was chosen because benchmark clients are not standardized nor do they correspond 1:1 to real-world application servers, as explained in the *About MLCommons and the MLPerf Benchmark* section of this document.

Based on that methodology and noting the differences in both client type and client count, the chart below shows Hammerspace results relative to other parallel file system vendor vendors.



Extrapolated results show expected performance at the same client count submitted by DDN and are based on Hammerspace internal analysis, not verified by MLCommons. Both single-client and multi-client results are shown.

The single Hammerspace Tier 0 client supported 33 simulated H100 GPUs, using only 1U of storage (within the 1U client) and a total of 3U of rack space (2x 1U for the Anvils, 1U for the client/storage server). This shows that Tier 0 configurations are very space- and power-efficient.

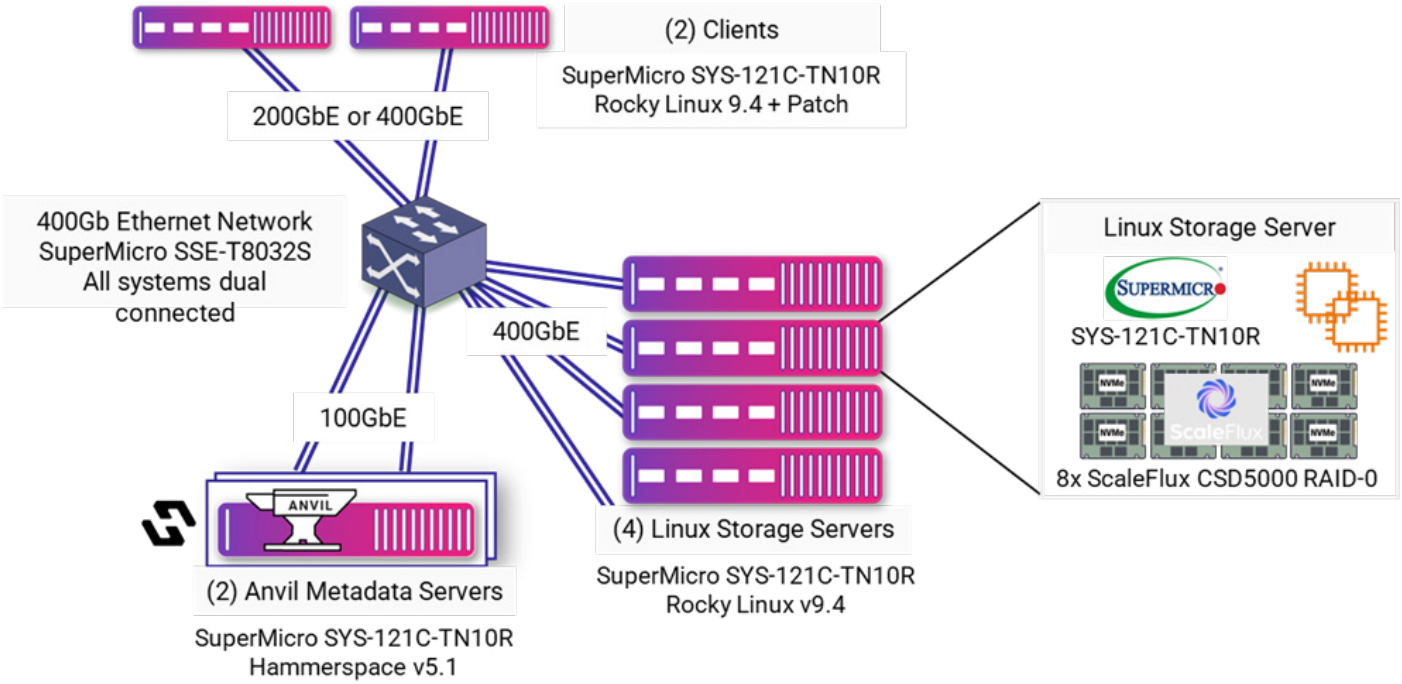
The extrapolation shows expected Hammerspace results based on the same number of clients used by DDN for their verified results. This extrapolation is valid because with Tier 0, every client works on data housed on its local NVMe drives, therefore there are no performance dependencies between clients.

Explanation of Verified Results vs. Extrapolated Results

To provide buyers with a somewhat normalized result, some charts in this section (MLPerf Storage v1.0 Open Division) and the next section (MLPerf Storage v1.0 Closed Division) show a combination of measured results (both MLCommons verified and unverified) and extrapolated results. The intent of the extrapolated results is to communicate Hammerspace's linear performance scalability, based on these facts:

- For the MLPerf Storage v1.0 Closed division verified results, Hammerspace submitted both single client results and results based on five clients, which showed linearly scalable performance in that range with the MLPerf Storage 3D U-Net workload.
- Hammerspace has demonstrated linear performance scalability in real-world production LLM training environments up to 1,000 storage nodes and 3,000 GPU clients with 24,000 total GPUs.

Test Configuration A



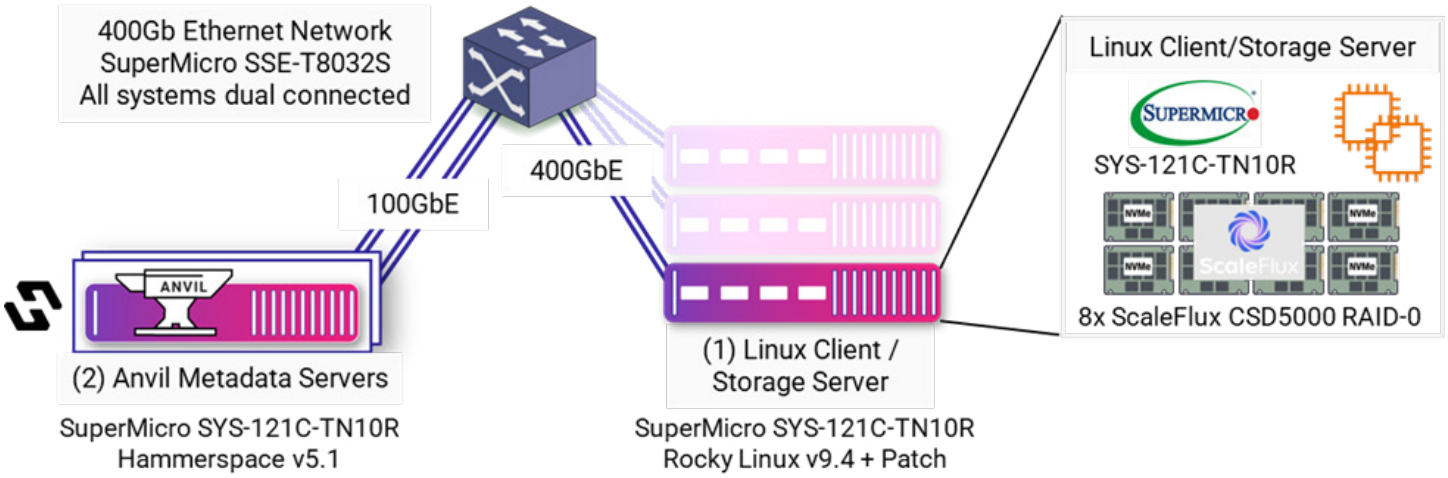
Configuration A is the pNFSv4.2 Tier 1 NAS configuration using external shared storage. This test is used to show the same Hammerspace software and storage server hardware running the benchmark as a baseline for comparison with the Tier 0 results that follow. This configuration consists of two redundant Anvil metadata servers in an active/passive configuration, four LSSs, and two clients. The Anvil servers are responsible for metadata operations and cluster coordination tasks, while the LSSs serve test data using internal ScaleFlux CSD5000 Computational Storage NVMe drives. No computational storage features of the ScaleFlux drives were used in this test.

It's important to emphasize that the LSSs are just standard Linux servers exporting NFSv3, with no added software.

All client systems mounted a Hammerspace share using standard pNFSv4.2 with the FlexFile layout type. Like the LSSs, the clients are standard Linux servers. Unlike other parallel file systems, Hammerspace does not require installation of any proprietary software on clients to achieve peak performance.

Clients and storage servers were connected to the network using 2x 200GbE or 2x 400GbE interfaces each. Anvil nodes were each connected via 2x 100GbE. Since Anvils are only involved in metadata communication (no data flows through them), 100GbE was sufficient.

Test Configuration B



Configuration B was used to demonstrate Tier 0 performance. Two redundant active/passive Anvil metadata servers are responsible for metadata operations and cluster coordination tasks. The client in this test has two roles. It runs the benchmark code as usual, but is also the storage server, serving the test data from internal ScaleFlux CSD5000 Computational Storage NVMe drives. No computational storage features of the ScaleFlux drives were used in this test.

It's important to emphasize that the combination client/storage server machine is just a standard Linux server with no added software. The internal storage is exported via NFSv3 and mounted using pNFSv4.2. Though the file system metadata path traverses the network to the Anvils, the data path remains entirely within the client/storage server host, providing the client with direct access to the local file system using the Tier 0 NFS protocol bypass. This direct data path boosts throughput and reduces latency.

The client/storage server was connected to the network using 2x 400GbE interfaces. Anvil nodes were each connected via 2x 100GbE links. Since Anvils are only involved in metadata communication (no data flows through them), 100GbE was sufficient.

The two additional faded client/storage server machines shown in the diagram reflect the configuration that would generate the extrapolated three-client results.

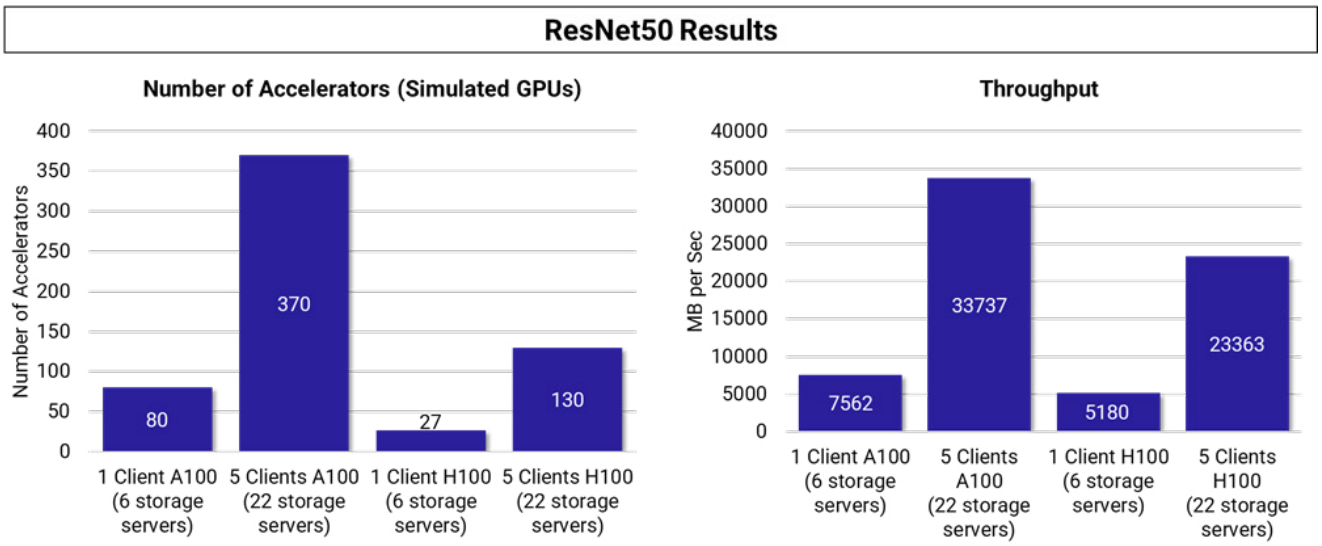
MLPerf Storage v1.0 (Closed Division)

The results of this test in the Closed division are verified by MLCommons Association as shown on the MLPerf Storage Results page.

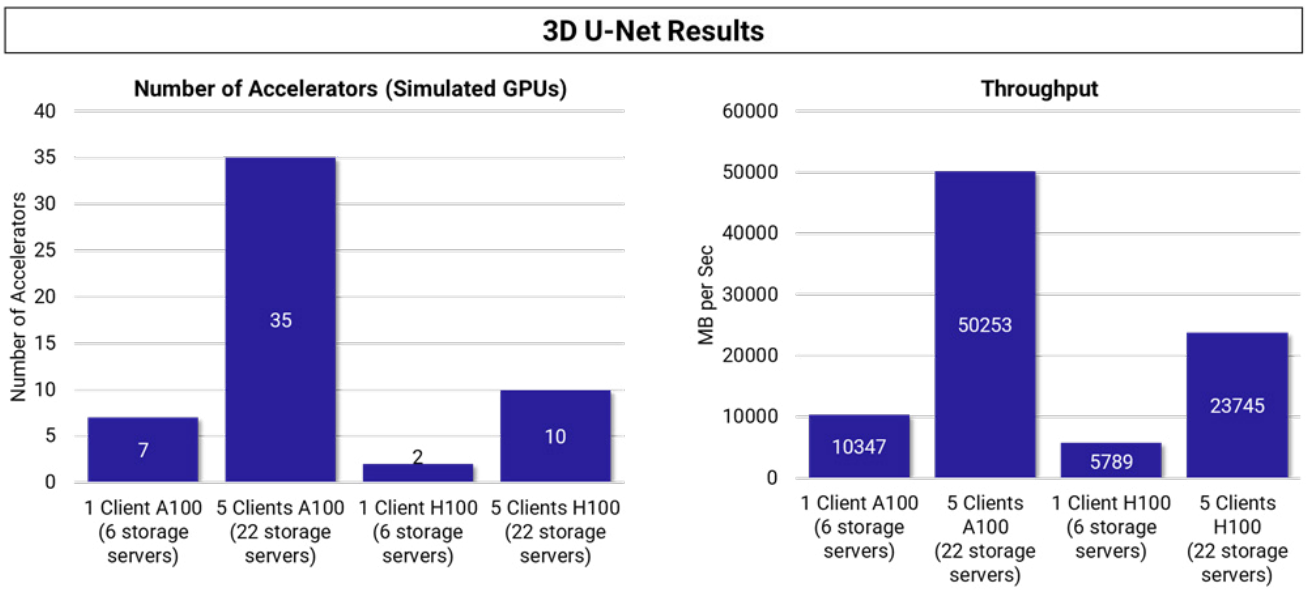
For this initial release of the MLPerf Storage benchmark, Hammerspace participated in the official submission period with a configuration based on cloud infrastructure. Cloud was selected to show the performance that can be achieved in standard cloud instances without specialized hardware designs.

Test Results

Hammerspace test results are shown in the figures below.



In the ResNet50 image classification workload simulation, a Hammerspace system with 22 flash-based Linux storage servers (LSSs) drove 370 simulated A100 GPUs and 130 simulated H100 GPUs to > 90% utilization, delivering 33.7GB/s and 23.3GB/s aggregate read performance, respectively.



With the 3D U-Net simulated image segmentation workload, this system drove 35 simulated A100s and 10 simulated H100s, delivering 50.3GB/s and 23.7GB/s, respectively.

A smaller configuration with six LSSs performed admirably as well, demonstrating the scalability of the system. It supported 80 simulated A100s at 7.6GB/s and 27 simulated H100s at 5.2GB/s in the ResNet50 test, and 7 simulated A100s and 2 simulated H100s at 10.3GB/s and 5.8GB/s, respectively, in the 3D U-Net test.

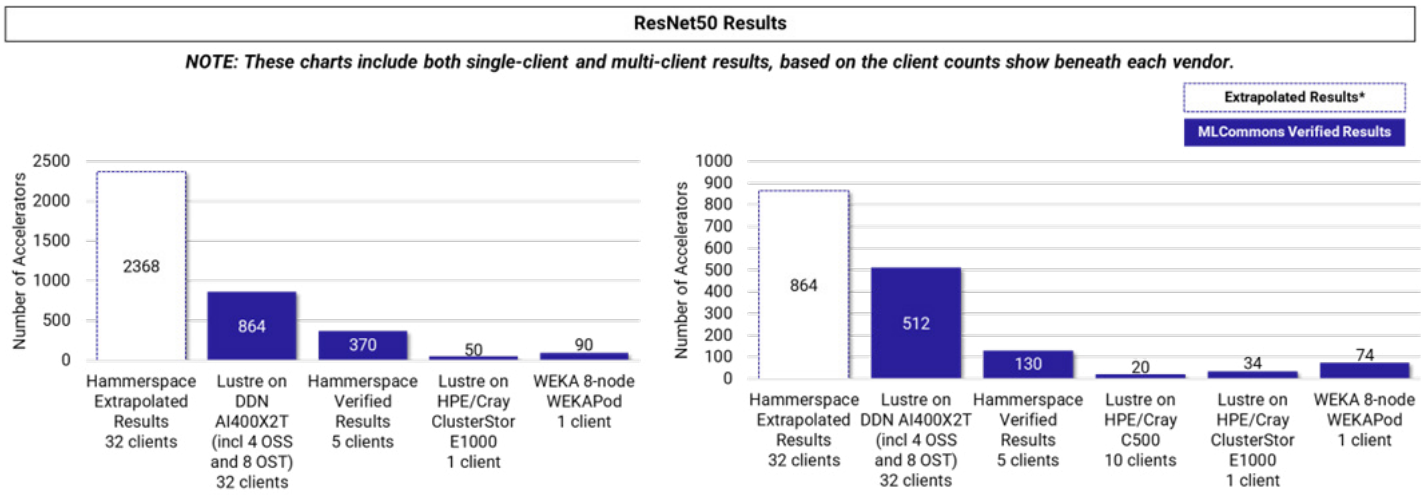
Both Hammerspace configurations used standard Ethernet networking and standard pNFSv4.2, requiring no special client-side software or agents. Neither was tested to its limits.

Results Competitive Comparison

The comparisons below show the maximum number of accelerators delivered by each vendor, independent of the number of clients used. This methodology was chosen because benchmark clients are not standardized nor do they correspond 1:1 to real-world application servers, as explained in the *About MLCommons and the MLPerf Benchmark* section of this document.

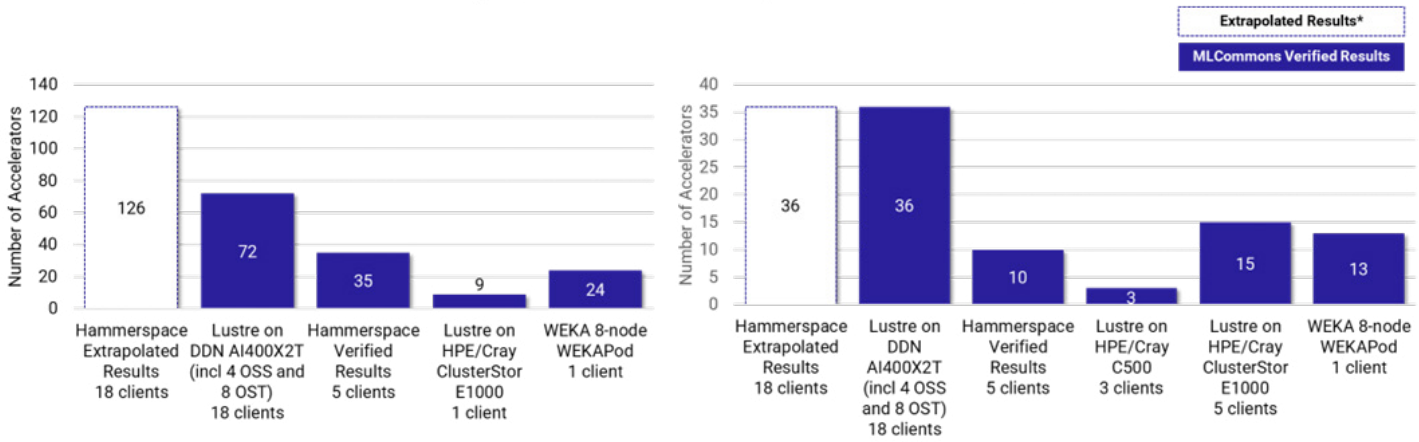
Based on that methodology and noting the differences in both client type and client count, the graphs below show Hammerspace results relative to other parallel file system vendor vendors.

Note that the graphs below contain extrapolated results to highlight the ability of Hammerspace to scale. Extrapolated results are not verified by MLCommons Association. The reasons for including extrapolated results are outlined in more detail in the *Explanation of Verified Results vs. Extrapolated Results* section of the MLPerf Storage v1.0 (Open Division) portion of this document

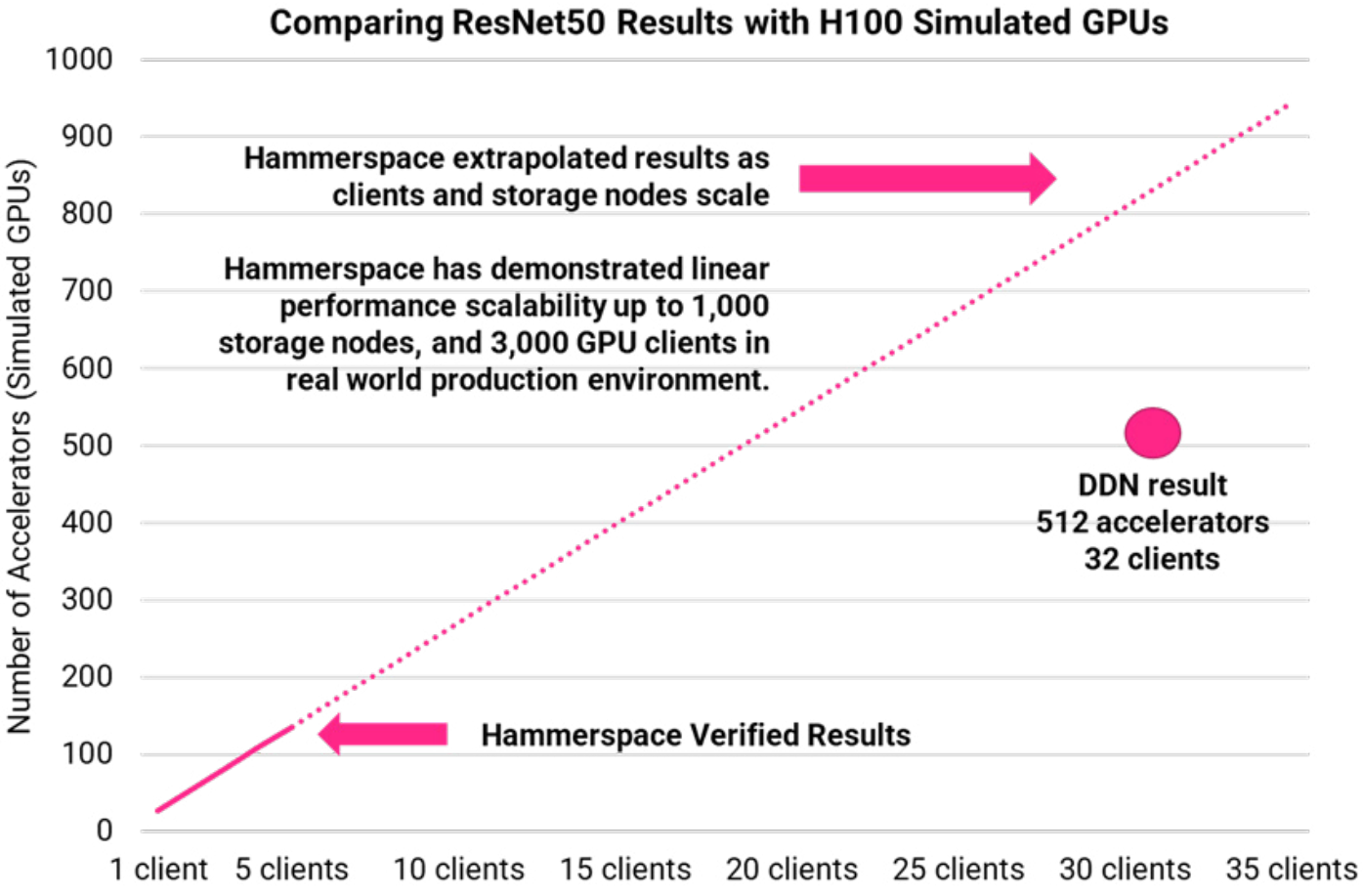


U-Net3D Results

NOTE: These charts include both single-client and multi-client results, based on the client counts shown beneath each vendor.



***The extrapolated results show the expected result at the same number of clients as submitted by DDN. These extrapolated results are based on Hammerspace internal analysis and have not been verified by MLCommons.**



Test Configuration Comparison

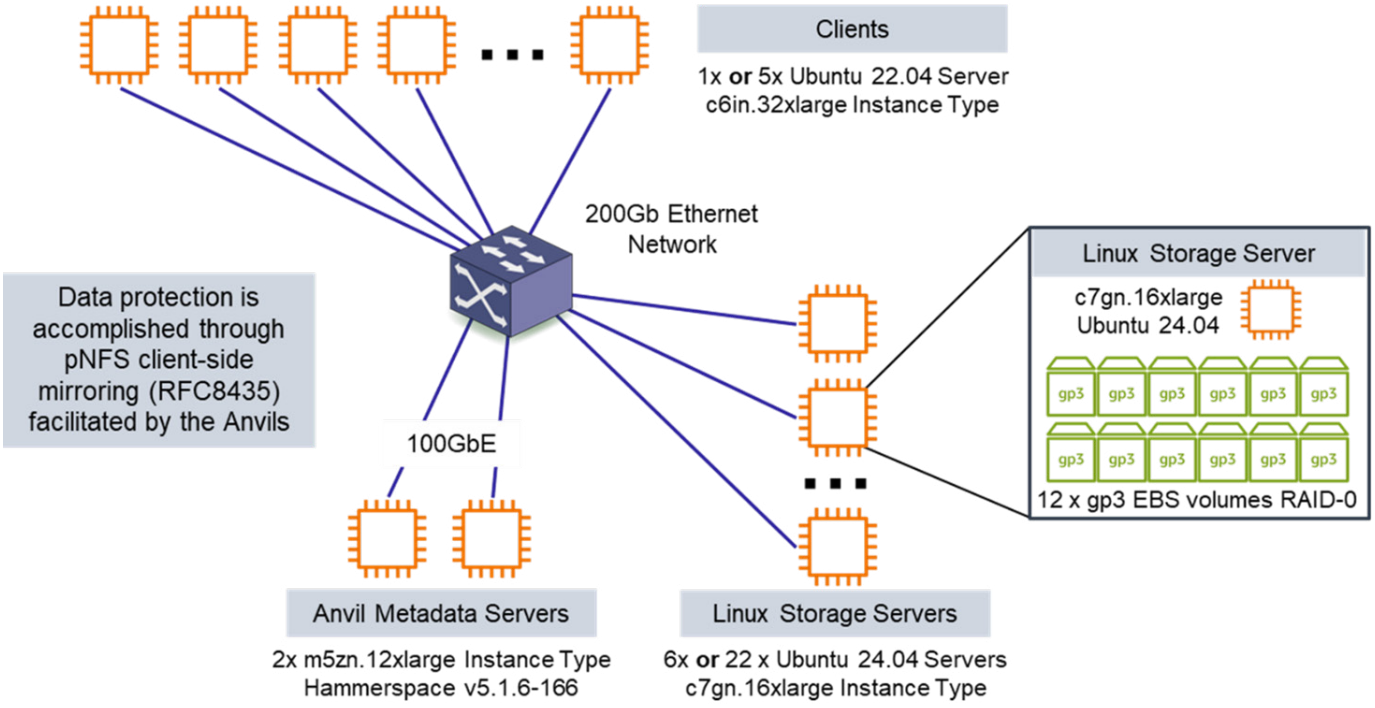
Vendor and Systems	DDN AI400X2 Turbo	Hammerspace	HPE Cray C500	HPE Cray ClusterStor E1000	Weka 8-node WEKApod
Client Description	Bare metal nodes 1x CPU, 92GB RAM	AWS C6in.32xlarge 128x vCPUs, 256GB RAM	Single CPU Socket, 131GB RAM	Single CPU Socket, 131GB RAM	NVIDIA HGX H100 Dual Socket, 2048GB RAM
Network Type	InfiniBand HDR100 links (one per client)	200GbE Virtual Network	InfiniBand	HPE Slingshot	400Gb/s NDR InfiniBand
Storage Server Description	DDN AI400X2 Turbo w/24x 14TB Phison NVMe Drives	C7gn.16xlarge, AWS Graviton3 64x vCPUs, 128GB RAM, 12x 500GB GP3 EBS Volume, RAID0	Cray C500	Cray ClusterStor E1000 w/2x AFAs, 24x 7.68TB Drives	1U WEKApod Servers AMD 48-core, 384GB RAM, 14x 3.84TB TLC NVMe Drives

The notable differences are summarized in the chart below, which highlights that Hammerspace was the only parallel file system vendor that used cloud infrastructure for this testing, used standard Ethernet, and standard NFS connectivity for this submission.

	Hammerspace	DDN	HPE	Weka
Infrastructure	**Cloud**	Physical Hardware	Physical Hardware	Physical Hardware
Networking	**Ethernet**	Infiniband	Infiniband / Slingshot	Infiniband
Client Connectivity	**Standard NFS**	Lustre Client	Lustre Client	Weka Client

Test Configuration

Testing was performed using Amazon Web Services (AWS) public cloud infrastructure. Cloud was selected for this MLPerf submission to show the performance that can be achieved in standard cloud instances without specialized hardware designs.



The system consisted of two redundant Anvil metadata servers in an active/passive configuration and up to 22 Linux storge server (LSS) nodes. The Anvil servers are responsible for metadata operations and cluster coordination tasks, while the LSSs serve test data using associated solid-state EBS volumes. Single-client tests used six LSSs, and multiple-client tests used all 22.

It's important to emphasize that the LSS nodes are just standard Linux servers exporting NFSv3, with no added software.

All client systems mounted a Hammerspace share using standard pNFSv4.2 with Flexible Files layouts.

Clients and LSSs were connected to the network using 200GbE interfaces. Anvil nodes were connected via 100GbE. Since Anvils are only involved in metadata communication (no data flows through them), they did not require 200GbE.

Conclusion

The benchmarks that are useful to AI and infrastructure architects are those that simulate real-world workloads. This is why Hammerspace likes the MLCommons® MLPerf Storage benchmark suite and is actively involved in efforts to expand and improve it. MLPerf Storage simulates a variety of realistic AI/ML workloads, providing relevant data points for organizations evaluating storage architectures for AI.

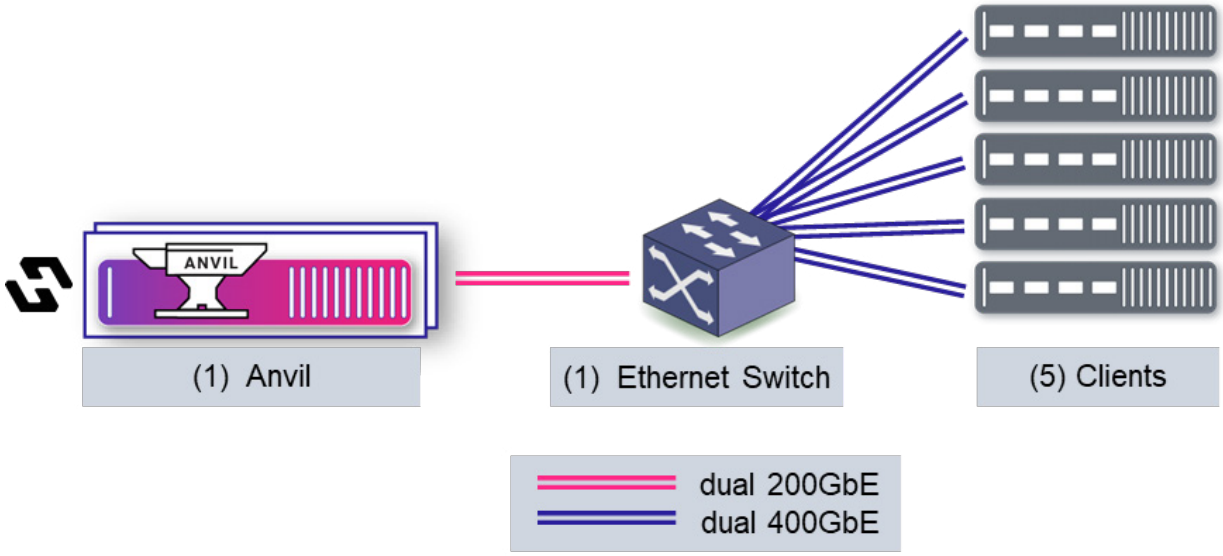
Hammerspace's results on the MLPerf Storage benchmarks with both on-premises and cloud infrastructure show that the Hammerspace Data Platform is ideally suited to be the data platform to provide the best utilization of data center resources to accelerate AI workloads and enable AI anywhere.

Appendix: Hardware, Software, Settings Details

Below are the details of the hardware and software configurations used for each test discussed in this document, along with any settings used to achieve the reported results.

MLPerf Storage v2.0

Test Configuration



Hardware Details

The hardware configuration used for these runs is detailed in the tables below.

Anvil

Resource	Qty	Specification
System	1	Lenovo SR630 v3
CPU	2	Intel Xeon Gold 6542Y 60MB Cache, 2.90 GHz, 48-Core
Memory	16	Samsung M321R4GA3PB0-CWMXJ 32GB DDR5-5600 ECC RDIMM Total 512GB RAM
Boot Drive	2	Samsung MZQL21T9HCJR-00A07 1.92TB PCIe4 x4 U.2 SSD (Mirrored)
Network Adapters	1	NVIDIA CX755106AS-HEAT ConnectX-7 Dual 200GbE QSFP112 PCIe5 x16
Storage Drives	2	ScaleFlux CSD5000 7.68TB NVMe PCIe5 x4 U.2 (Metadata - Mirrored)

MLPerf Client

Resource	Qty	Specification
System	1	Supermicro SYS-121C-TN10R
CPU	2	Intel Xeon Gold 6542Y 60MB Cache, 2.90 GHz, 48-Core
Memory	16	Micron MEM-DR564L-CL02-ER56 64GB DDR5-5600 ECC RDIMM, Total 1TB RAM
Boot Drive	2	Micron 7450 Pro MTFDKBG1T9TFR 1.92 TB NVMe PCIe4 M.2 TLC
Network Adapters	2	NVIDIA MCX75310AAS-NEAT ConnectX-7 Single 400GbE OSFP PCIe5 x16
Storage Drives	10	ScaleFlux CSD5000 7.68TB NVMe PCIe5 x4 U.2 (RAID-0)

Network

Resource	Qty	Specification
Ethernet Switch	1	Supermicro SSE-T8032S Twin-port 2x 400GbE Open Networking Switch

Storage Rack Units

Resource	Qty	Rack U each	Total Rack Units
Anvil	1	1	1
		Configuration Total	1

Software Details

The software running on each component type in this configuration is detailed below.

Anvil

Hammerspace is packaged as a single installation ISO that includes a Linux operating system, application code, and all dependencies.

- Hammerspace: v5.2.2-39
- Kernel: 6.12.24-8.hs.91.el8.x86_64
- Pd-protod RPM: pd-protod-5.2.2-86124.el8.x86_64.rpm

MLPerf Client

- Standard, unmodified MLPerf benchmark v2.0 code
- Rocky Linux v9.5 kernel 6.12.24-8.hs.91.e19.x86_64

Network Switch

- SONiC OS v4.5.0 Enterprise Advanced.

Settings

Specific settings used to achieve the results shown are listed below.

Anvil

- Increased the number of initial instances created using this setting in the /pd/dme/configuration.ini file:

  [layoutManager]
  maxInitialNASInstances = 5

- Hammerspace enables “Objectives” to be set by administrators. Objectives are declarative policies describing desired data-handling behavior, including aspects such as data durability. For this configuration one objective was set, specifying data durability of nine nines.

MLPerf Client

- The following changes were made to settings for the NFS client:
- /sys/module/nfs/parameters/localio_async_probe (Default Y) Setting is N
- /sys/module/nfs/parameters/localio_enabled (Default is N) Setting is Y
- /sys/module/nfs/parameters/localio_O_DIRECT_semantics (Default N) Setting is Y
- /sys/module/nfsv4/parameters/nfs4_files_uncacheable (Default N) Setting is Y

The following changes were made to settings for the NFS server:

- /sys/kernel/debug/nfsd/io_cache_read (Default 0) Setting is 2
- /sys/kernel/debug/nfsd/io_cache_write (Default 0) Setting is 2

Clients mounted the shared file system using this command:

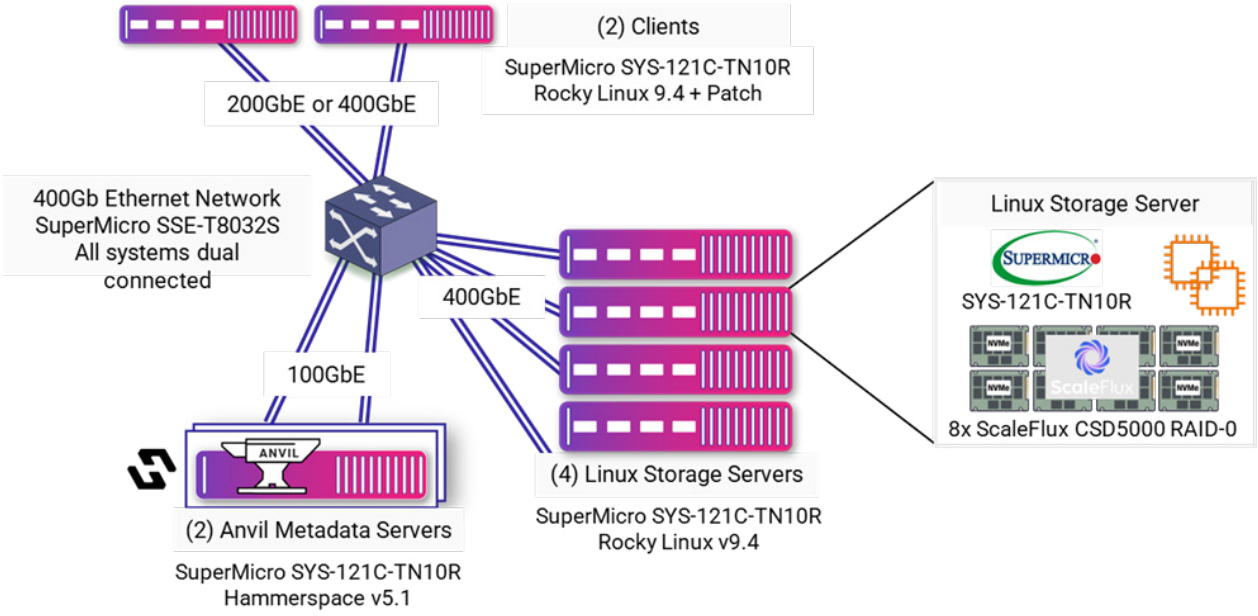
```
mount -t nfs -o nconnect=16,port=20491 192.168.0.161:/hs_test /mnt/hs_test
```

Network Switch

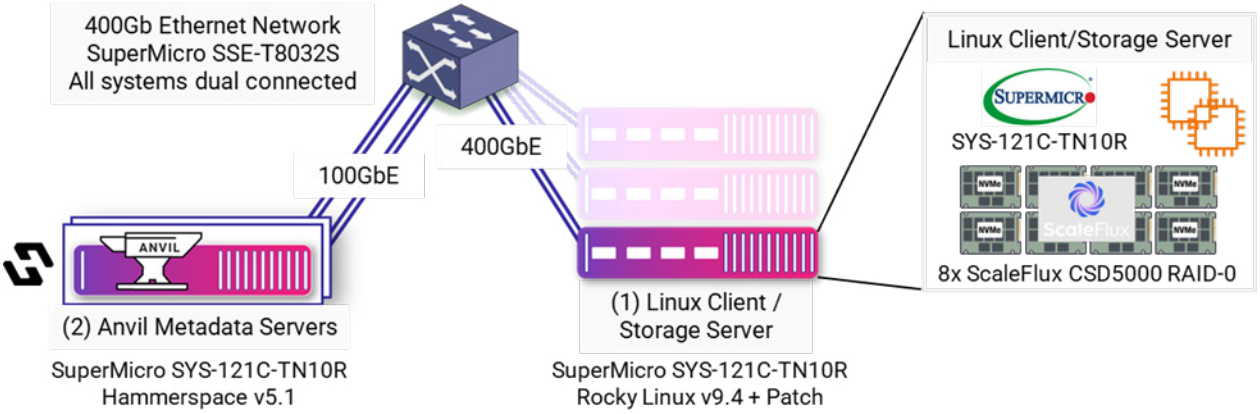
No tunings were applied to the network switch.

MLPerf Storage v1.0 (Open Division)

Test Configuration A



Test Configuration B



Hardware Details

The same hardware was used for both configurations described above.

Hammerspace Anvil Metadata Servers (2)

Component	Manufacturer	Model	Details
Chassis	SuperMicro	SYS-121C-TN10R	1U Dual Socket
CPU (x2)	Intel	Xeon Gold 6542Y	60MB Cache, 2.90 GHz, 24-Core
Memory (x16)	Micron	MEM-DR564L-CL02-ER56	64GB DDR5-5600 ECC RDIMM 1TB Total RAM per server
Boot Drive (x2)	Micron	7450 MTFDKBA960TFR	960GB NVMe PCIe4 M.2 3D TLC
NIC (x2)	NVIDIA	ConnectX-7 Dx	Dual 400GbE QSFP56 PCIe4 x16
Storage (x4)	ScaleFlux	CSD5000	7.68TB NVMe PCIe4 U.2

Linux Storage Servers (4) and Clients (2)

Component	Manufacturer	Model	Details
Chassis	SuperMicro	SYS-121C-TN10R	1U Dual Socket
CPU (x2)	Intel	Xeon Gold 6542Y	60MB Cache, 2.90 GHz, 24-Core
Memory (x16)	Micron	MEM-DR564L-CL02-ER56	64GB DDR5-5600 ECC RDIMM 1TB Total RAM per server
Boot Drive (x2)	Micron	7450 MTFDKBA960TFR	960GB NVMe PCIe4 M.2 3D TLC
NIC (x2)	NVIDIA	ConnectX-7 Dx	Dual 400GbE QSFP56 PCIe4 x16
Storage (x8)	ScaleFlux	CSD5000	7.68TB NVMe PCIe4 U.2

Network Switch

Manufacturer	Model	Details
SuperMicro	SSE-T8032S	Twin-port 2x 400GbE Open Networking Switch, SONiC OS

Software Details

The same software was used for both configurations described above.

Anvil

Hammerspace is packaged as a single installation that includes the Linux operating system, application code, and all dependencies.

Anvil nodes ran Hammerspace v5.1.

Linux Storage Servers

Linux Storage Servers ran Rocky Linux v9.4 with no additional patches.

Clients and Tier 0 Combined Client / Storage Server

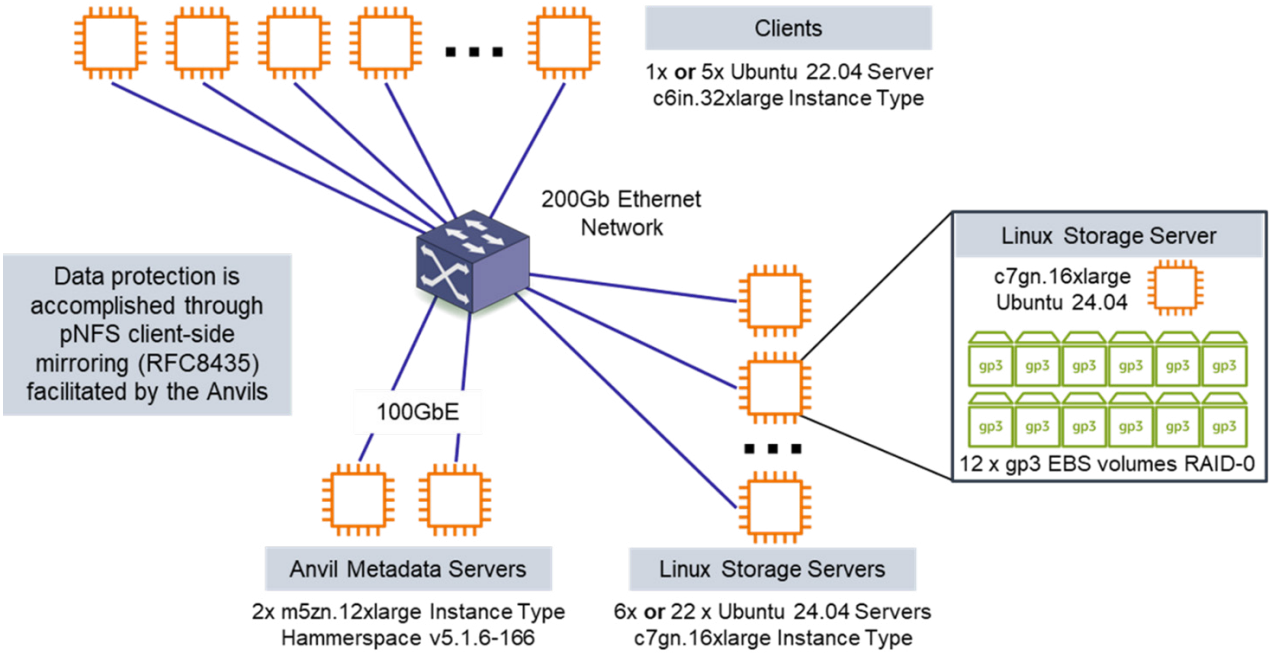
These machines ran Rocky Linux v9.4 with the additional of one upstream kernel patch to enable RDMA to honor the NFS nconnect mount option.

MLPerf Benchmark Code

Modified MLPerf benchmark code was used that enabled bypassing the client-side page cache. This is not required for Tier 0, it represents additional tuning to improve performance.

MLPerf Storage v1.0 (Closed Division)

Test Configuration



Hardware Details

AWS cloud infrastructure was used for this test as described above

Software Details

Anvil

Hammerspace is packaged as a single installation ISO that includes a Linux operating system, application code, and all dependencies.

The Anvil nodes ran Hammerspace v5.1.6-166.

Storage Servers

The storage servers ran Ubuntu Server Linux v24.04 LTS. Additionally, the NFS server was tuned to allow for more NFS daemons:

Set `RPCNFSDCOUNT=64` in `/etc/default/nfs-kernel-server`

Clients

Clients ran Ubuntu Server Linux v22.04.4 LTS. Clients communicated with the Anvil and Storage Servers through NFS. The mount command used was:

```
mount -t nfs -o vers=4.2,port=20492,nconnect=16 IP_OF_ANVIL:/mlperf /mnt/mlperf
```