Hammerspace Data Platform IO500 ISC '25 Benchmark Results

Twice the Performance. One-Eighth the Infrastructure.
Built for AI and HPC Without the Complexity.

TECHNICAL BRIEF

Summary

The IO500 Foundation created the IO500 storage benchmark to measure the performance of storage systems under High-Performance Computing (HPC) workloads, typically those used for Modeling and Simulation (Mod/Sim) in government, academic, and industry research environments. This benchmark has long been well known in the HPC community but virtually unheard of elsewhere. That is changing with the rise of enterprise AI, which is driving HPC-class performance requirements into traditional enterprise IT architectures. However, AI projects bring different constraints and priorities than specialized HPC Mod/Sim use cases, so the IO500 results must be interpreted carefully in the context of enterprise AI.

Hammerspace submitted results for the IO500 benchmark for the first time at the 2025 International Supercomputing show (ISC25). This technical brief summarizes the results of that testing, including:

- Background on the IO500 benchmark and the individual tests within it.
- A summary of Hammerspace's results relative to other vendors, including the test setup used for the benchmark.
- A discussion of the advantages of the Hammerspace standards-based parallel file system architecture for enterprise AI as compared to scale-out NAS and HPC parallel file systems.

The results prove that simplicity doesn't have to come at the cost of storage inefficiency. Hammerspace demonstrated impressive performance with a compact hardware footprint and feature set compatible with the expectations and experience of enterprise IT teams.

Key Results:

- **Hammerspace delivered 2x the IO500 10-Node challenge score and 3x the bandwidth of VAST – using just nine nodes compared to VAST's 128.**
    - This results in higher performance with just a fraction of the hardware, power, and infrastructure complexity to deploy and manage.
- **Hammerspace landed among the Top 10 highest performers on the IOEasy Write and IOEasy Read tests** in the 10-node challenge.

Why Enterprises Should Care About the IO500

Enterprise AI infrastructure leaders may assume that IO500 results are only relevant to elite HPC labs and traditional modeling and simulation use cases in industry. But today's AI workloads—training, inferencing, RAG, fine-tuning—place similar demands on infrastructure: they're IOPS-heavy, metadata-sensitive, and require sustained, high-bandwidth access to massive data sets.

What matters isn't just raw performance—it's how that performance is achieved:

- Can it start small and scale without complexity?
- Can it integrate with existing enterprise storage?
- Can it deliver results without requiring a forklift upgrade?
- Can it avoid creating yet another storage silo to handle the performance requirements of AI?

This is why Hammerspace's IO500 results are relevant to enterprises, confirming that enterprise AI infrastructure can achieve HPC-class performance without specialized hardware, networking, or proprietary clients – using a standards-based parallel global file system that works with the infrastructure and data that's already in place.

> "Too often, achieving performance means copying all your data, moving your applications, and rearchitecting your environment—just to make your data AI-ready. It shouldn't have to be that way."

With Hammerspace, organizations can assimilate existing siloed data into a unified global namespace – without migrating data or creating yet another data storage silo. This makes current environments instantly AI-ready, eliminating the need to re-architect systems or copy massive data sets. It's high performance made easy.

Hammerspace IO500 Results and Discussion

The IO500 tests measure systems based on both bandwidth and metadata performance to provide a holistic view. The detailed test results are boiled down to three numbers: a composite bandwidth figure, a composite IOPS figure, and a calculated overall numeric score.

Hammerspace Results, 10-Node Research Cluster ISC2		
Overall Score	Bandwidth GiB/s	IOPS kIO/s
63.77	37.25	109.16

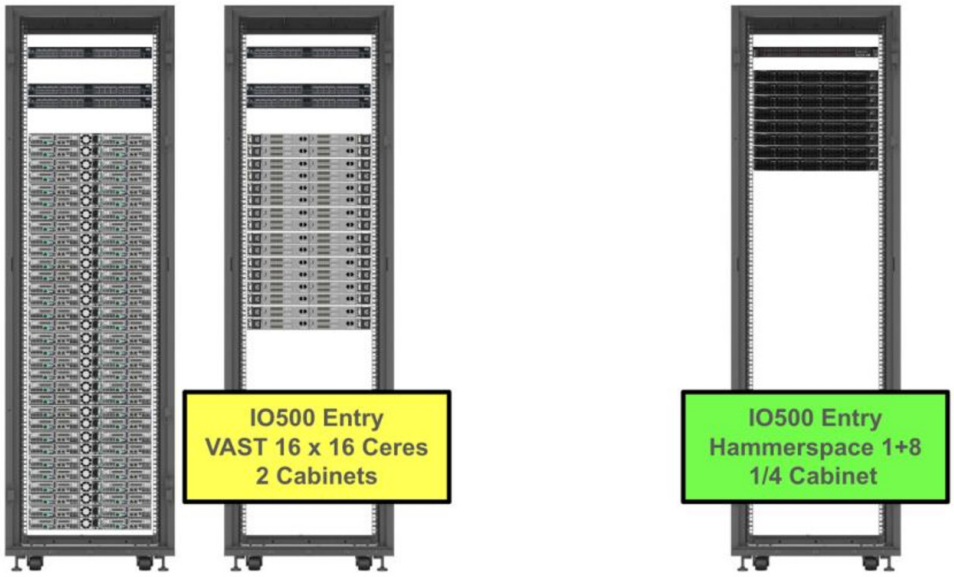
Looking at the “10-Node Research” list, Hammerspace ranks alongside traditional HPC file systems like Lustre, Spectrum Scale, WekaFS, and BeeGFS—systems often tuned for peak “hero” numbers in highly customized environments. But our results tell a different story.

68	SC19	CSIRO	bracewell	Dell/ThinkParQ	beegfs
69	SC19	DDN	AI400	DDN	Lustre
70	ISC25	Hammerspace	Hammerspace Performance Lab	Hammerspace	pNFS v4.2 via Hammerspace Anvil Metadata Service
71	SC21	Raytheon	MCK-HPC	Penguin Computing	WekaFS
72	SC23	Center for Research Informatics at University of Chicago	Randi	IBM	Spectrum Scale

Hammerspace delivered leading performance using standard Linux, standard Ethernet networking, standard NVMe storage servers, and a single metadata controller. No custom hardware. No exotic fabrics. No proprietary clients. Even our metadata search used the default Linux “pfind” command as compiled with the IO500 source—not a benchmark-optimized variant.

This benchmark validates a critical shift: Enterprise AI doesn't need complexity to get performance. The future of enterprise AI infrastructure is standards-based and enterprise-ready—built on the hardware and networks organizations already have, not on proprietary silos.

While vendors like VAST and Pure Storage are aggressively pursuing the AI market, both require disruptive infrastructure overhauls: forklift upgrades, massive hardware footprints, and migrating applications to their new silos. Hammerspace takes a fundamentally different approach. We stand alone in delivering superior performance—2X the IO500 score, 3X the bandwidth, and 31% higher metadata path IOPS than VAST, for example —while using just a fraction of the infrastructure. **No new silos. No data copying. No re-platforming. Just performance and efficiency, on enterprise terms.**



	Overall Score	Bandwidth GiB/s	IOPS kIO/s	Server Nodes	Capacity PiB
VAST Data	31.08	11.61	83.22	128	3.8
Hammerspace	63.77 **(2X Higher)**	37.25 **(3X Higher)**	109.16 **(31% Higher)**	9 **(119 Fewer)**	0.492 **(~1/8th)**

Hammerspace turned in solid results overall, and our data platform did particularly well on the test known as “IOEasy.” This test focuses on streaming, large-block IO, as used when loading training data, model saving, and similar operations. Hammerspace placed in the top ten results for IOEasy Write and Read among entrants on the Ten-Node Research list.

Ten-Node Research List IOEasy Write				
IOEasy Write Rank	IO500 Score	System Name	Filesystem Type	IOEasy Write GiB/s
1	1	Cheeloo-1	OceanFS2	2480.69
2	6	Aurora	DAOS	1473.27
3	3	ParaStor	ParaStor	905.49
4	10	Endeavour	DAOS	538.83
5	4	StarStor	StarStor	502.88
6	14	SuperMUC-NG Phase2	DAOS	455.22
7	11	SuperMUC-NG-Phase2-10	DAOS	434.74
8	33	Cumulus	Lustre	429.28
9	70	Hammerspace	pNFSv4.2	420.35
10	17	Cumulus	DAOS	360.44

Ten-Node Research List IOEasy Read				
IOEasy Write Rank	IO500 Score	System Name	Filesystem Type	IOEasy Write GiB/s
1	1	Cheeloo-1	OceanFS2	2438.99
2	6	Aurora	DAOS	1301.33
3	3	ParaStor	ParaStor	776.42
4	4	StarStor	StarStor	770.07
5	10	Endeavour	DAOS	586.05
6	35	MEGWARE XiRAID Classic	Lustre	580.14
7	70	Hammerspace	pNFSv4.2	446.79
8	15	Meadowgate	DAOS	394.94
9	33	Cumulus	Lustre	381.45
10	9	OceanStor Pacific	OceanFS	357.78

For organizations building AI infrastructure, Hammerspace software proves you can achieve performance at scale without creating a new silo or undergoing a forklift upgrade.

Hammerspace Efficiency/Simplicity Advantage

Enterprise AI teams are prioritizing simplicity and open standards as they build out their data infrastructures. Traditional HPC parallel file systems such as Lustre, IBM Storage Scale, and BeeGFS are anything but simple, requiring proprietary clients, special networks, and complex administration. Modern Network Attached Storage (NAS) systems that use standard protocols present a simpler option, but this simplicity comes with a hidden cost: Disaggregated and scale-out NAS systems require significantly more infrastructure to reach performance goals – and hit scalability limits as storage node counts grow. These systems are easier to use, but trade efficiency and scalability for that ease. What's worse, to achieve maximum performance some of these systems require customers to install proprietary client software on their application servers, adding complexity and vendor lock-in.

Hammerspace proves that you don't have to choose between simplicity and efficiency. Our architecture delivers performance that rivals HPC parallel file systems – without the need for any new client software, exotic storage, custom networking, or large-scale server sprawl. The result is high performance, low friction, and a clear path to AI readiness, on the familiar infrastructure that's already in place in standard enterprise IT data centers.

Parallel File Systems, AI, and the Enterprise

Parallel file systems are needed to achieve the extreme scalability and performance necessary to support HPC workloads, such as AI/DL and Mod/Sim. Historically parallel file systems such as Lustre, BeeGFS, and others have been deployed for these use cases, which have predominantly been in specialized HPC environments. But the problem is these file systems are famously complex, very fragile, subject to downtime, and require expensive networking such as InfiniBand to perform at the level required. Most importantly, these solutions require specialized client software for every server accessing the storage, which requires that applications be designed to operate with the proprietary interfaces of the system.

> "Enterprise AI needs simplicity—but not at the cost of infrastructure sprawl. Hammerspace proves there's a better way. 2-3x more performance with 93% less hardware."

While classic HPC research environments have adapted to this reality, the expansion of AI has driven HPC-class workloads into the broader enterprise data center, where such limitations are not acceptable. The added cost, operational complexity, and specialized talent required to keep classic HPC file systems operational is not practical in enterprise data environments, especially since the deployment and optimization of these HPC file systems can take weeks or even months to fine tune.

These complexities have caused many organizations to assess scale-out NAS as an alternative to HPC parallel file systems. Scale-out NAS solutions are easier to use, include enterprise reliability features,

support standard Ethernet networking, and are well understood by conventional IT staff in enterprise data centers. Unfortunately, scale-out NAS cannot deliver the performance needed for high-performance use cases at scale.

- Scale-out NAS systems require 2x the number of servers and 2x the number of network ports when compared with parallel file systems, which create a direct data path between clients and storage. Scale-out NAS storage solutions also struggle to deliver performance at scale.
- Traditional HPC parallel file systems like Lustre require proprietary client software that adds complexity and ongoing administrative burden and costs. They require custom hardware, and exotic and expensive networking technology like InfiniBand and Slingshot.

The Hammerspace Data Platform: The Best of Both Worlds

With its Data Platform architecture, Hammerspace brings together the best of these technologies while overcoming the negative challenges of each, to deliver the best price/performance storage for GPU computing in AI, ML, HPC, and deep learning.

Hammerspace delivers the best price/performance by using:

- At least 50% fewer servers and 50% fewer network ports than scale-out NAS architectures such as Dell PowerScale, Qumulo, and VAST
- Standard Ethernet connectivity, eliminating the need for a specialized second network, such as InfiniBand, which is used in the IO500 results published by others using parallel file systems from DDN, HPE and WEKA
- Existing Linux client servers to connect to the file system without the need for any additional client software beyond what they already have in the standard Linux kernel
- Existing applications natively designed to interface via NFS, SMB, or S3 can do so, without the need to redesign for a parallel file system interface

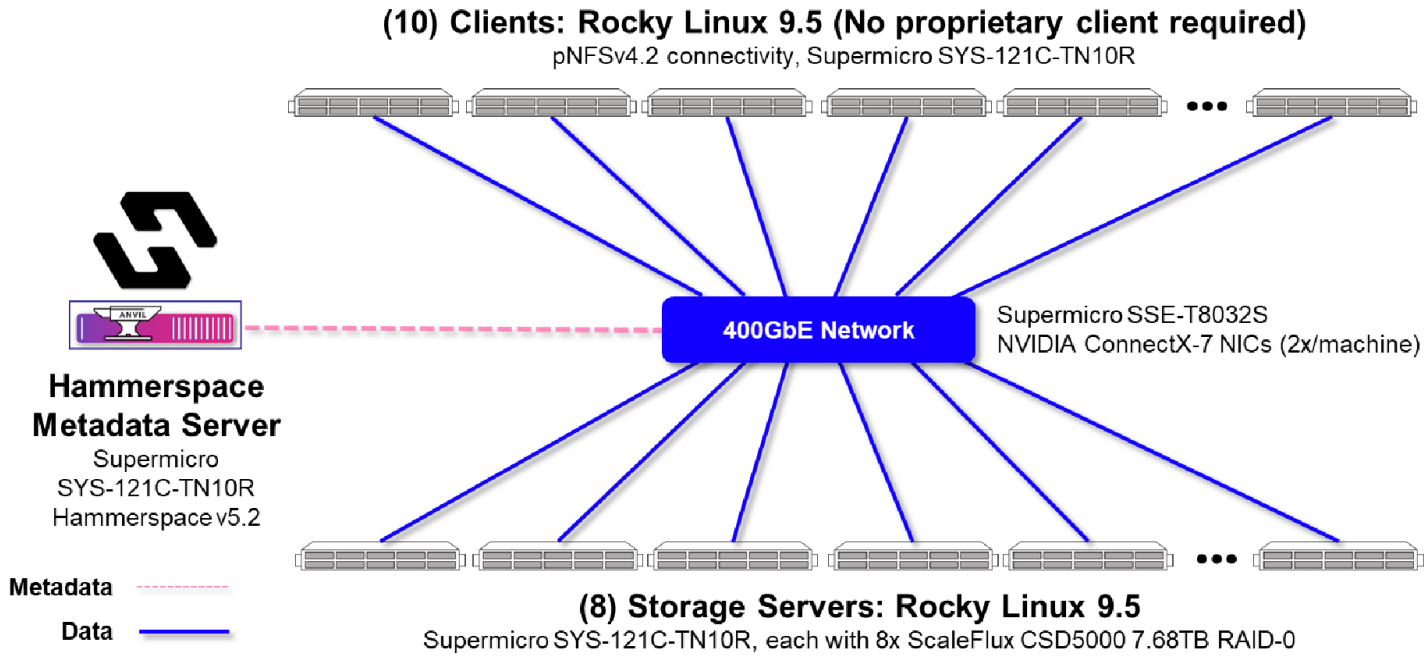
By reducing the number of servers and switches, this also reduces the amount of power and cooling required. While this can save costs, most importantly it also frees up wattage for the compute environment.

	SCALE-OUT NAS	LEGACY PARALLEL FILE SYSTEMS	HAMMERSPACE DATA PLATFORM
SAFE Reliability, Availability, Serviceability	✓	✗	✓
EASY Standards-Based, Plug-N-Play	✓	✗	✓
FAST HPC-Class Performance	✗	✓	✓
AFFORDABLE Cost Effective at Scale	✗	✓	✓

Because Hammerspace Data Platform software supports any server hardware and storage solutions from any vendor, including in the cloud, buyers are free to purchase hardware from any source, including OCP servers such as those being used by Meta in their AI Research Supercluster. Meta chose Hammerspace as their high-performance solution for provisioning their Llama 2 and Llama 3 LLM training pipelines, because only Hammerspace demonstrated the linear scalability to achieve over 12TB/sec over standard networking, feeding data between Meta's existing 1,000-node NVMe storage cluster and a 3,000-node GPU cluster with 24,000 GPUs in total. No other vendor came close. For Llama 4, Hammerspace's ability to extend on-premises environments into the cloud enabled Meta to seamlessly expand that project into Oracle cloud, with plans to further expand to other cloud providers.

Hammerspace IO500 Test Configuration

The Hammerspace test setup was straightforward: 10 clients, eight Linux storage servers with NVMe drives, and a single metadata server, all connected over 400GbE.



The hardware is powerful, but not remarkably so – it's off-the shelf stuff. There's nothing special about the network, either, it's just Ethernet using RDMA (RoCE). Nothing you couldn't find in any enterprise data center. File access was via pNFSv4.2 with FlexFiles which is standard and built into every modern Linux distribution today.

This bears emphasizing – the client machines are just running stock Linux. There is no special client software, special kernel, or agent of any kind installed or required. The storage servers are also just running Linux, in this case a build with additional performance patches that we are contributing upstream into the open-source kernel. This is what a standards-based system looks like.

Note that “standards-based” doesn’t mean “standards-only.” The NFS standards define a set of basic tools. With any craft, quality depends more on the skill and experience of the craftsman than the tools themselves. Nobody has more experience with pNFSv4.2 and FlexFiles than Hammerspace, which contributed the standard into Linux in 2018, and has since added over 2,500 performance-related enhancements to the standard kernel. It is our pNFS metadata server implementation that enables the highest performance. But when that raw performance is combined with a comprehensive set of enterprise-class data services, the result is the Hammerspace Data Platform, a uniquely powerful option for organizations wrestling with how to support AI workloads.

About the Virtual Institute for I/O IO500 Benchmark

Goals of the Virtual Institute for I/O (VI4IO) are:

- Provide a platform for I/O researchers and enthusiasts for exchanging information.
- Foster training and international collaboration in the field of high-performance I/O.
- Track/encourage the deployment of large storage systems by hosting information about high-performance storage systems.

The philosophical cornerstones of VI4IO are:

- Treat contributors/participants equally.
- Allow free participation without any fee inclusive to all.
- Independent of vendors/research facilities”.

From https://hps.vi4io.org/_media/research/publications/2018/dltvifiatikl18-the_virtual_institute_for_i_o_and_the_io_500.pdf

IO500 participants range from small organizations to huge government labs, with entries in four categories:

Category	Who Can Submit	Client Count	Client Size	Storage System Capacity/Nodes
Production	Production Customer	Any	Unlimited	Unlimited
10-Node Production	Production Customer	10	Unlimited	Unlimited
Research	Anyone	Any	Unlimited	Unlimited
10-Node Research	Anyone	10	Unlimited	Unlimited

As you might expect, the scores cover a wide range, about five orders of magnitude. This is due to the range of different systems participating, and the fact that clients and storage systems tested can be of unlimited size, (and cost). If you want the top spot you'll need to spend millions of dollars to get there.

The "10-Node" categories introduce a limit of 10 on the client count (clients are what run the benchmark code) as a way to compare setups that are a little more constrained. Hammerspace submitted results in the "10-Node Research" category.

The benchmark consists of a range of tests focusing on different workload patterns summarized in the table below. Tests involve read, write, delete, and stat operations. Results are output for each test and then combined and summarized to calculate an overall score.

Workload	**Details**	**AI Workload Relevance**
IOEasy	File per process, large block sequential IO	Similar to training data loading and model saving
IOHard	Shared single file for all processes, 47KB+ IO	Small, random IO typical in optimizer state and KV cache updates
MDEasy	Directory per process, zero-length files	Pure metadata speed test
MDHard	Small files (3901 bytes) Single shared directory for all processes	Mimics checkpoint/restore logic, where file create/delete dominates
Find	Finding relevant objects based on patterns	patterns Metadata search operation – custom Find algorithms are allowed