REFERENCE ARCHITECTURE

ACCELERATING A.I. WORKLOADS AT LIGHTSPEED

UNLOCK NEW INSIGHTS WITH VAST DATA UNIVERSAL STORAGE AND NVIDIA DGX A100 SYSTEMS

JANUARY 2021

CONTENTS

EXECUTIVE SUMMARY

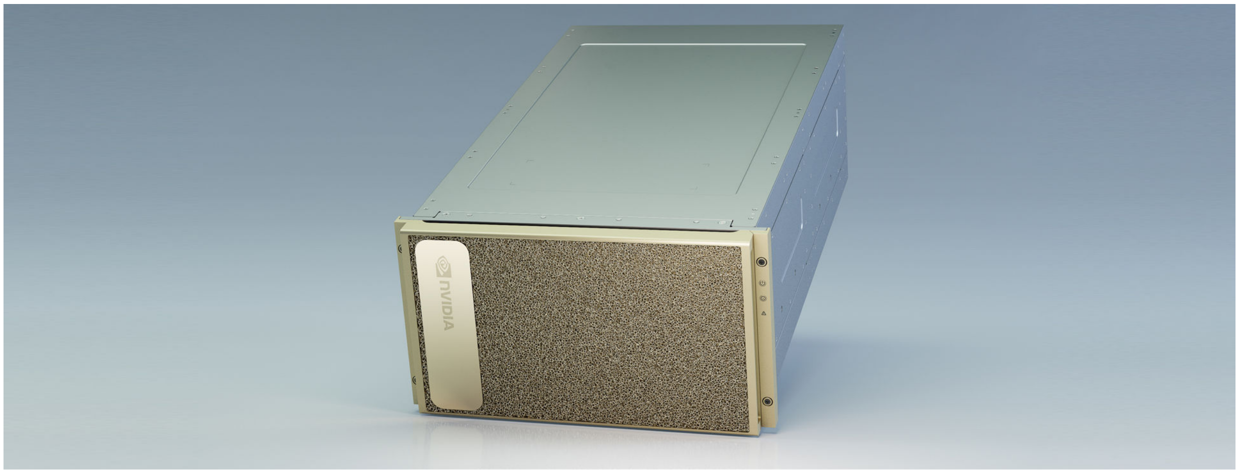
This document details the VAST Data Universal Storage reference architecture for machine learning (ML) and artificial intelligence (AI) workloads and contains benchmarking results obtained in partnership with NVIDIA. This reference design is implemented using VAST Data's LightSpeed all-flash storage system, four NVIDIA DGX™ A100 systems, and NVIDIA® Mellanox® Quantum™ InfiniBand and Spectrum Ethernet switches.

The operation and performance of this system was validated using industry-standard benchmark tools. Based on the testing results, VAST Data Universal Storage and NVIDIA DGX A100 systems demonstrate excellent performance for training and inferencing workloads. Customers looking to build infrastructure for new AI initiatives can deliver high – and predictable – performance to meet any machine-learning and deep-learning workload requirements by easily and linearly scaling compute and storage resources.

INTRODUCTION

About the NVIDIA DGX A100 System

The DGX A100 system is the universal system for all AI workloads, offering unprecedented compute density, performance, and flexibility in the world's first 5 petaflop AI system. Built on the revolutionary NVIDIA A100 Tensor Core GPU, the DGX A100 system enables enterprises to consolidate training, inference, and analytics workloads into a single, unified data center AI infrastructure. More than a server, the DGX A100 system is the foundational building block of AI infrastructure and part of the NVIDIA end-to-end data center solution created from over a decade of AI leadership by NVIDIA. The DGX A100 system integrates exclusive access to a global team of AI-fluent experts that offer prescriptive planning, deployment, and optimization expertise to help fast-track AI transformation.



About VAST Data LightSpeed

VAST Data, the pioneer of the Universal Storage concept, delivers all-flash storage systems for a wide range of workloads and use-cases. While built for high-performance, VAST has also eliminated the tradeoff between storage capacity and performance by delivering high-capacity and high-performance, at a price-point significantly lower than traditional all-flash storage systems.

In 2020, VAST Data introduced LightSpeed, both as a product and storage philosophy for AI-based workloads that optimizes for simplicity and speed. New VAST Data Universal Storage clusters built using LightSpeed enclosures deliver 2X the I/O throughput compared to previous enclosures. In addition, VAST added support for a number of NFS client-side enhancements to achieve best-in-class throughput for AI applications and GPU servers without requiring complex parallel file systems.

NFS-over-RDMA

The Network File System (NFS) is a popular protocol used for accessing files remotely over networks and is widely used in enterprise environments for numerous applications because of its simplicity. In specialized environments, however, NFS has historically been deemed too slow for high-performance requirements.

But by adding support for Remote Direct Memory Access (RDMA), NFS can now deliver greater efficiency of data transfers by bypassing client-side CPU and memory when leveraging InfiniBand or RDMA-over-Converged Ethernet (RoCE) networks. This results in higher throughput and lower latency while preserving the operational simplicity and familiarity with this industry standard storage protocol.

NFS Multipath

While NFS-over-RDMA optimizes throughput over a single client-side network interface, some applications, such as those running on GPUs, need more throughput than a single network connection can provide. By default, NFS uses one IP address and TCP connection for I/O limiting the total throughput of most NAS systems to 2 GB/s. NFS multipath facilitates using multiple IP addresses and network connections for I/O, in effect aggregating different network links to provide access to a single file system while increasing total throughout.

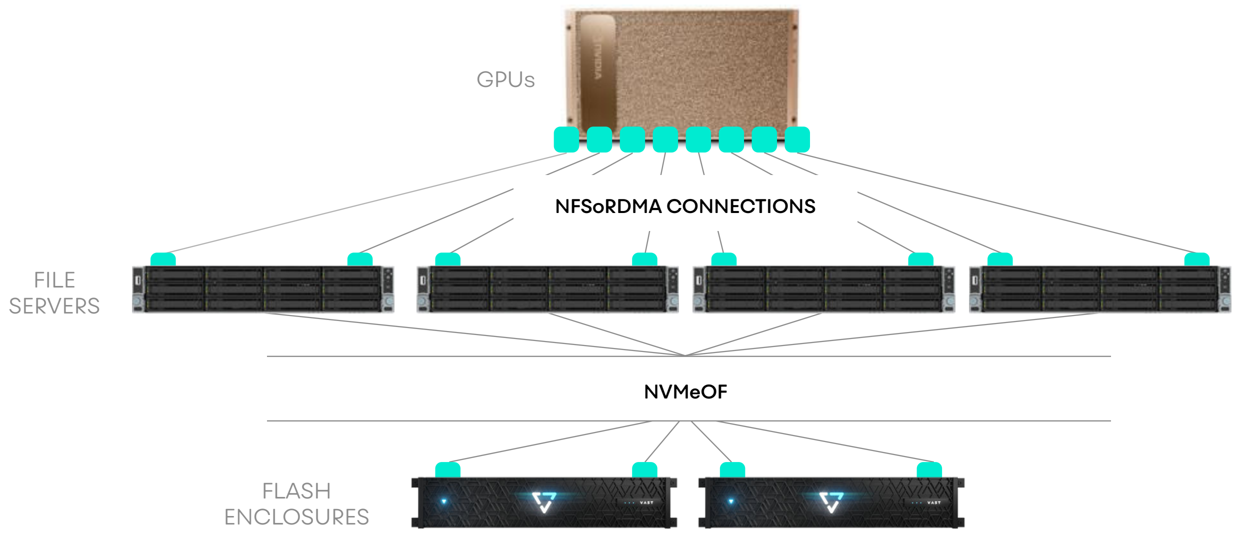


Figure 1 – NFS Multipath aggregates multiple DGX A100 network interfaces to provide better remote storage throughput with built-in high-availability.

VAST Data has developed a new multipath solution, based on nconnect (found in NFS 4.1), and has integrated this functionality back into NFS 3.x. These open-source enhancements have been submitted upstream for inclusion in both previous and future Linux distributions.

About NVIDIA GPUDirect Storage

Eliminating potential system bottlenecks increases the performance and efficiency of AI workload environments. NVIDIA GPUDirect™ Storage (GDS), one of a family of technologies that are part of NVIDIA Magnum IO™, enables a direct path for direct memory access (DMA) transfers between storage and GPU memory without burdening CPUs. This results in dramatically lower CPU utilization within the GPU server, but also increased storage bandwidth and decreased latency.

As an early NVIDIA GDS partner, VAST Data now supports GDS via LightSpeed v3.4+. Using NFS-over-RDMA as the storage protocol, GDS-enabled applications can now benefit from the increased performance between NVIDIA DGX systems and VAST Data's all-flash storage.

Storage Requirements for AI Applications

When designing infrastructure for AI applications, storage architects typically must make tradeoffs (e.g. persistent versus non-persistent, flash drives versus spinning disks) to balance the needs of performance, capacity, and cost. The wrong balance can often lead to subpar training performance, limited by the rate at which data can be read and re-read from storage.

Caching and Tiering Considerations

Traditionally, a storage and caching hierarchy is employed to counter the tradeoffs typically found across performance, capacity, and cost. The table below shows a typical hierarchy illustrating the capacity and performance tradeoffs common to a DGX A100 system.

Storage Hirearchy Level	Technology	Total Capacity	Performance
RAM	DDR4	1TB per node (upgradeable to 2TB)	> 200 GB/s per node
Local Cache	NVMe	15TB per node (upgradeable to 30TB)	> 55 GB/s per node
External Storage	Varies	Varies depending on specific needs	Varies

Caching data in local RAM provides the best performance for reads. This caching is transparent once the data are read from a storage filesystem. However, the size of RAM is limited and less cost effective than other storage and memory technologies. Local NVMe storage is a more cost-effective way to provide caching close to the GPUs. However, manually replicating datasets to the local disk can be tedious and time consuming and is not a viable long-term storage location for datasets.

While external storage solutions vary wildly in terms of performance, choosing an external storage system capable of delivering performance that meets or exceeds the local cache eliminates the need for complicated caching strategies.

VAST Data's innovative approach to all-flash storage will blur the performance lines between local cache and external storage. The benchmarking below will show how VAST's use of NFS can close the performance gap that exists with traditional enterprise NAS appliances.

Workloads and Dataset Size Considerations

An ideal DGX POD architecture should also strike a good balance between cost and performance across a range of common deep learning model types, including Image Processing, Recommendation, and Natural Language Processing. The table below captures three different performance levels required for common AI workloads and dataset sizes.

Storage Performance Level	Work Description	Dataset Size
Good	Natural Language Processing	Most all datasets fit in cache
Better	Image processing with Compressed images, ImageNet/ ResNet-50	Many to most datasets can Fit within the local node's cache
Best	Training with 1080p, 4K, or uncompressed images, offline inference, ETL	Datasets are too large to fit into cache, massive first epoch I/O requirements, workflows that only read the dataset once
Note: The actual performance requirements of the external storage is not provided because the answer greatly depends on the types of models that are being trained and the size of the dataset.		

Understanding specific workload requirements can greatly instruct how infrastructure should be designed for AI applications. However, if an external storage system is capable of scaling both performance and capacity from day one, accommodating changing workload requirements will be much more feasible.

VAST Data's Disaggregated and Shared Everything (DASE) architecture truly scales performance and capacity to meet the requirements of any AI workload today and tomorrow.

SOLUTION OVERVIEW

While VAST Data's DASE architecture linearly scales to both larger capacity and higher performance configurations, this reference architecture serves as guidance for the ideal storage deployment for a pod of four DGX A100 systems. This combination of VAST Data's Universal Storage, AI computing hardware and software from NVIDIA, and Infiniband and Ethernet networking from NVIDIA provides insight into the performance expected across a representative set of AI workloads.

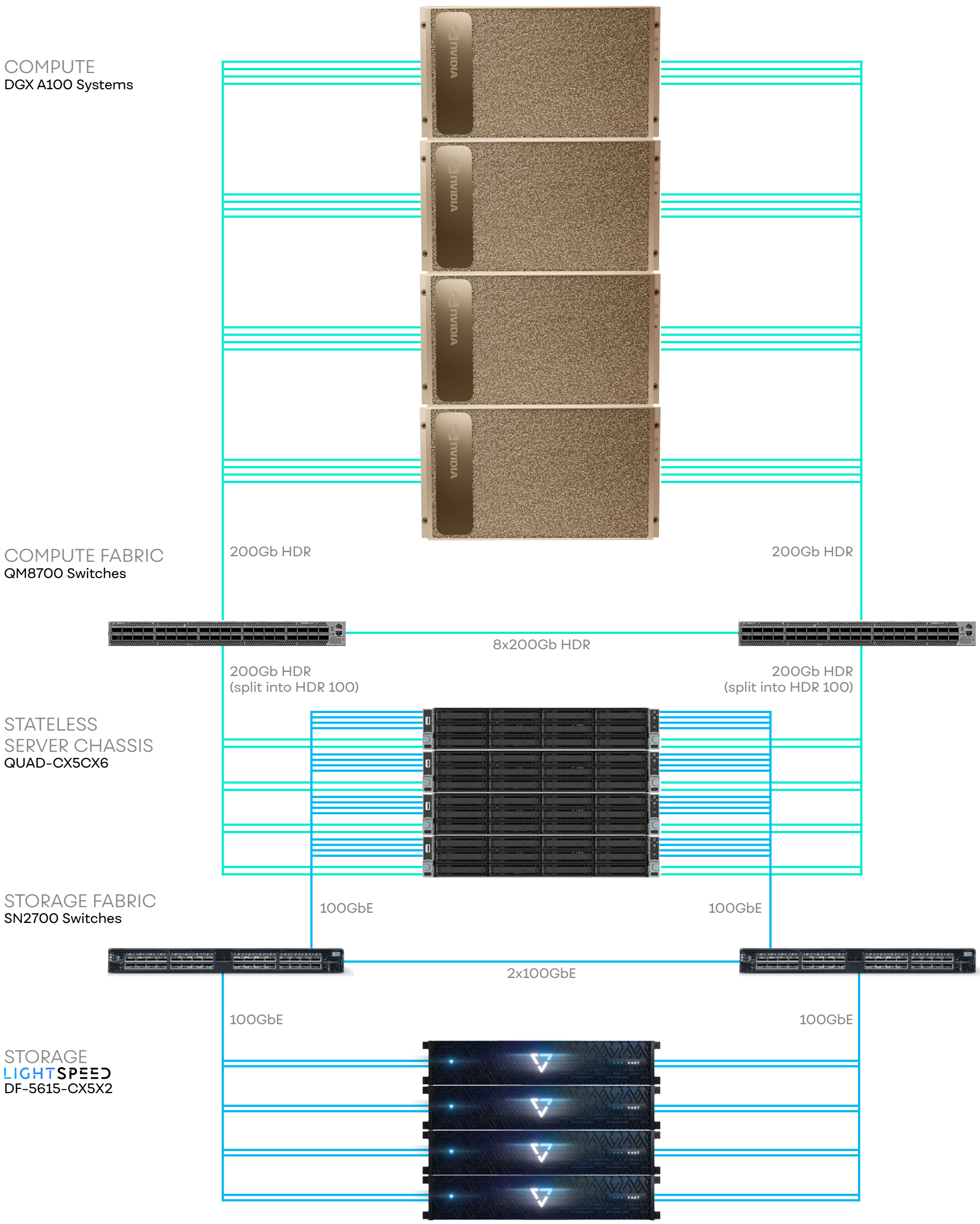


Figure 2 – VAST Data validated architecture with DGX A100 systems

The remaining sections outline the specific hardware, software, and network configurations used for all the testing described throughout this paper.

Hardware Requirements

The following table lists the hardware components that were used for this reference design.

Quantity	Component
4	NVIDIA DGX A100 System
2	NVIDIA Mellanox QM8700 InfiniBand Switch
4	VAST Data File Server (Model: QUAD-CX5CX6)
4	VAST Data Flash Enclosure (Model: DF-5615-CX5X2)
2	NVIDIA Mellanox SN2700 Ethernet Switch

Software Requirements

The following table lists specific software versions used for this reference design and benchmarking.

Software	Version
DGX OS	4.99.9
Docker Container Platform	19.03.8
Container runtime	nvcr.io/nvidia/mxnet:20.06-py3 – Mlperf test tensorflow:20.05-tf2-py3 – other tests
NCCL version	2.7.5
FIO version	3.1
mdtest version	3.3.0
MOFED version	5.0-2.1.8.0
VAST OS	2.3.0-sp14

Network Configuration

VAST Data's unique disaggregated architecture enables the support of both Ethernet and InfiniBand networks, for both client-facing networks and back-end storage fabrics. In this specific reference architecture, the DGX A100 system's InfiniBand compute fabric is leveraged for VAST Data's storage systems, resulting in high-throughput, low-latency I/O performance using a simple NFS client protocol:

- Each DGX A100 system is connected via eight (8) 200Gb HDR InfiniBand ports.
- Each VAST Data File Server is dual-homed, connected via one (1) 100Gb HDR InfiniBand port (using a 2-way splitter from the 200Gb switch port) to service storage requests from the DGX A100 systems, as well as connecting into a backend storage (Ethernet) fabric via two (2) 100GbE ports.
- Each VAST Data Flash Enclosure is connected via four (4) 100GbE ports into the backend storage (Ethernet) fabric.
- Both InfiniBand and Ethernet switches are deployed in a pair for high-availability.

TESTING METHODOLOGY AND RESULTS

This solution's configuration was tested by measuring each benchmark while scaling from one to four DGX A100 systems against VAST Data's storage appliances.

NCCL Tests: all_reduce_perf

The NVIDIA Collective Communications Library (NCCL) implements multi-GPU and multi-node collective communication primitives that are performance optimized for NVIDIA GPUs. This NCCL test is intended to verify scalability across multiple DGX A100 systems to ensure there are no inherent bottlenecks for GPU-to-GPU communication within subsequent benchmarks.

For a single DGX A100 system, the maximum throughput should be limited to the internal NVIDIA NVLink® interconnect bandwidth. For multiple DGX A100 systems, the maximum throughput should approach the total aggregate bandwidth of all network adapters used for GPU-to-GPU communication within a single system.

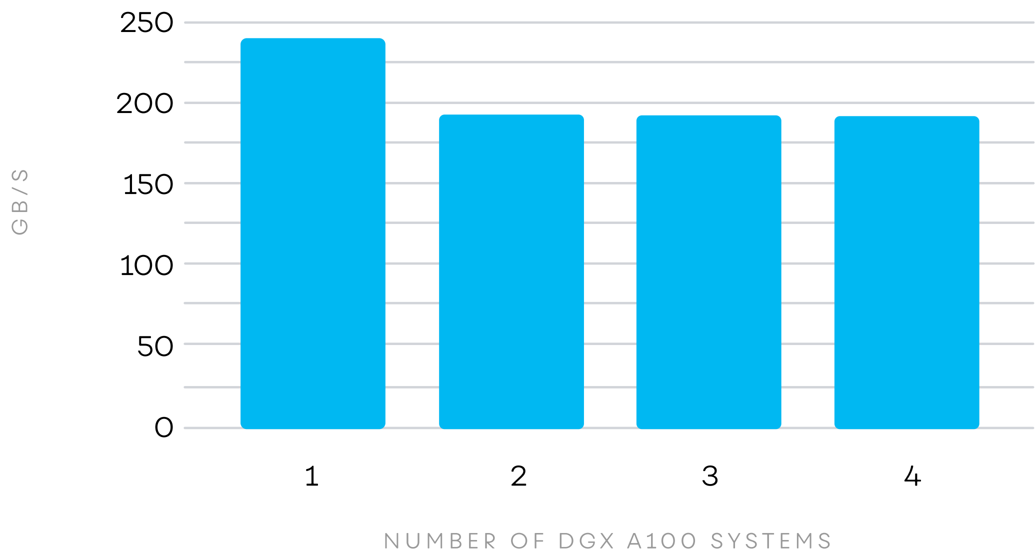


Figure 3 – NCCL bandwidth test results

The results of the NCCL test show a single DGX A100 system's inter-GPU bandwidth reaching the maximum throughput of the internal NVLink switching capabilities. When scaling to two, three, and four DGX A100 systems, the inter-GPU bandwidth across multiple systems approaches the aggregate bandwidth of the eight InfiniBand network adapters allocated for compute per DGX A100.

FIO Bandwidth & IOPS

Flexible IO tester (FIO) is an open-source synthetic benchmark tool useful for measuring I/O of local or external/remote storage targets. FIO can generate various IO workloads and supports the ability to scale testing using multiple clients simultaneously.

Separate tests were run to measure VAST Data's storage bandwidth in order to simulate both 100% read and 100% write workload scenarios.

Below are the specific FIO configuration parameters used for these tests:

- ioengine = posixaio
- direct = 1
- blocksize = 1024k for bandwidth test, 4k for IOPS test
- numjobs = 144 for bandwidth test, 288 for IOPS test
- iodepth = 16
- size = 4194304k

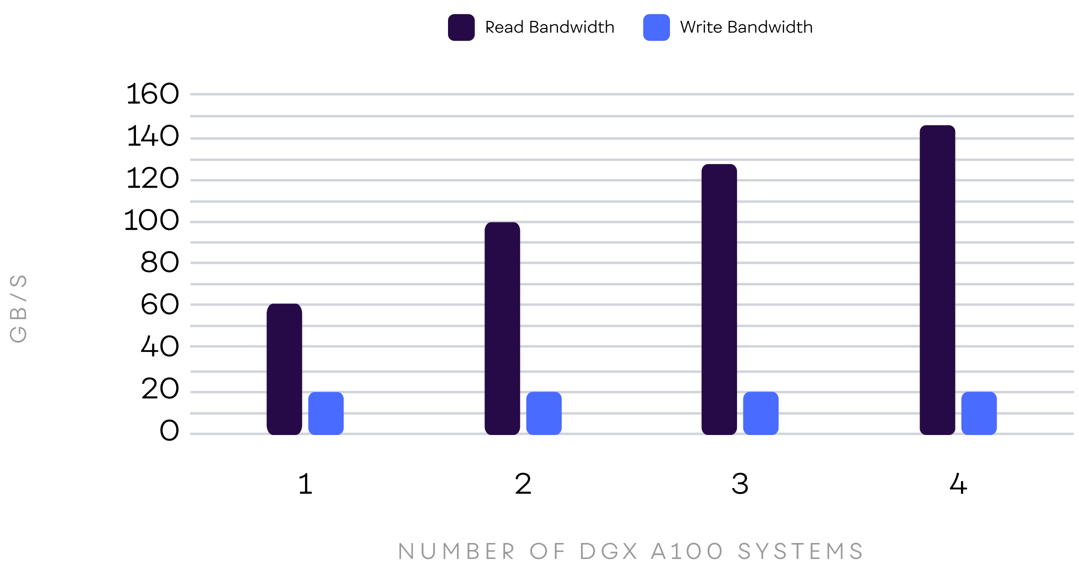


Figure 4 – FIO read and write bandwidth results

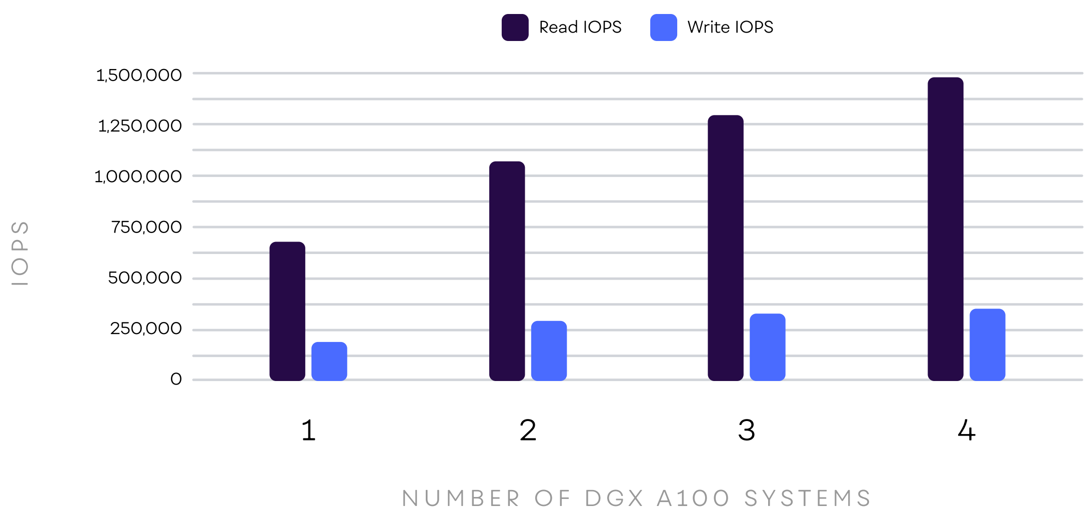


Figure 5 – FIO read and write IOPS results

The FIO test results demonstrate more than 140GB/s of read bandwidth and nearly 1.5M read IOPS from four DGX A100 systems. This joint NVIDIA and VAST Data reference architecture is intelligently designed to leverage additional 200Gb NICs within each DGX A100 for storage traffic when needed, while still allowing for maximum multi-system inter-GPU bandwidth. The tests also show 20GB/s of write performance (the maximum bandwidth rated for this specific storage configuration), plenty for the read-oriented nature of AI workloads.

MLPerf Training v0.7 – ResNet-50

MLPerf Training is an industry-standard benchmark suite for measuring how fast systems can train models to a target quality metric. This particular test uses the ResNet-50 neural network, a well-known image classification network, that can be used with the ImageNet dataset. With a quality target of 75.90% classification, this test workload is both sufficiently computationally-intensive and I/O-intensive.

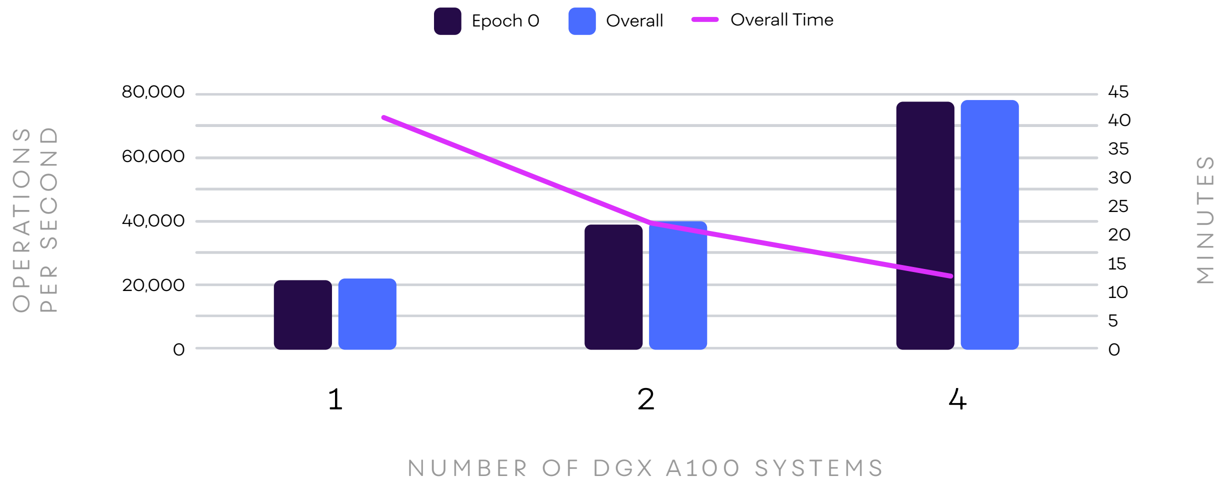


Figure 6 – MLPerf ResNet-50 test results

The MLPerf test results show an impressive near-zero delta between performance during epoch zero compared to the average overall performance of subsequent epochs. Since the first epoch reads data directly from the file system into local cache (where it is read from on all subsequent epochs), the filesystem is most heavily stressed during the first epoch and has a smaller impact on the later passes. The fact that there is so low of a difference between the first epoch and the overall average demonstrates that VAST Data's filesystem has no issues keeping up with the demand for this test.

SCALING VAST DATA UNIVERSAL STORAGE

This reference architecture, consisting of four (4) VAST Data Flash Enclosures, serves as a proven building block for designing and deploying modern AI infrastructure with DGX A100 systems. The performance demonstrated should serve as guidance for the minimum performance expected from this specific configuration. As organization's AI initiatives expand and datasets grow, VAST Data's DASE architecture helps seamlessly scale storage capacity (to exabytes) and storage performance linearly (to TB/s+) to meet the changing needs of every AI workload today and beyond.

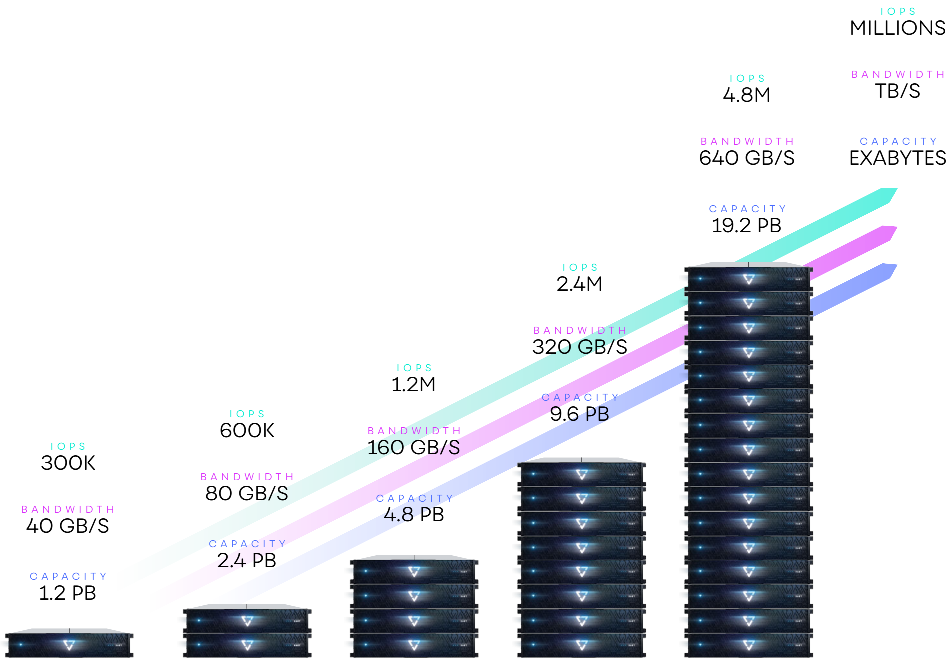


Figure 7 – VAST Data scales performance linearly with affordable 100% NVMe-flash systems

CONCLUSION

As AI becomes increasingly pervasive in our world - across all industries and verticals - organizations need a simpler way of providing modern IT infrastructure. Just as traditional compute (CPU-based) infrastructure is ill-suited for AI workloads, legacy storage architectures are no longer adequate in meeting the I/O demands of GPU-equipped systems.

This reference architecture provides evidence and a proven recipe for designing and delivering high-throughput storage infrastructure – without the need for complex parallel file systems. The results illustrate VAST Data's ability to linearly scale performance from one to four DGX A100 systems, perfect for AI workloads both large and small. With NVIDIA DGX A100 systems and VAST Data's Universal Storage, IT and data science teams alike no longer need to worry about infrastructure architecture and instead focus on unlocking new and innovative insights through AI.

To learn more about how **Universal Storage** can help accelerate your A.I. initiatives, reach out to us at hello@vastdata.com.