DDN A³I® SOLUTIONS WITH NVIDIA DGX™ A100 SYSTEMS

Fully-integrated and optimized data platforms for accelerated at-scale AI, Analytics and HPC

EXECUTIVE SUMMARY

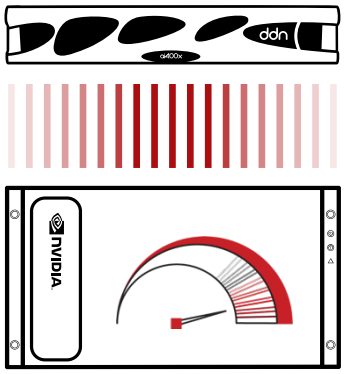
DDN A³I Solutions are proven at-scale to deliver highest data performance for AI and HPC applications running on GPUs in a DGX A100 system. DDN AI400X appliances provides up to 60X more throughput and 50X more IOPS than NFS-based data platforms, and scale predictably to ensure optimal application performance as AI requirements grow. DDN fully integrates GPUDirect Storage and demonstrates full GPU saturation, up to 162 GiB/s per DGX A100 system. The AI400X appliance enables GPU systems at all scale globally, including NVIDIA Selene, the largest SuperPOD with DGX A100 currently in operation, ranked #7 on the latest IO500 list.

DDN A³I Solutions with NVIDIA DGX A100 Systems

DDN A³I solutions are architected to achieve the most from at-scale AI, Analytics and HPC applications running on DGX systems. They are designed to provide extreme amounts of performance, capacity and capability through a tight integration between DDN and NVIDIA systems. Every layer of hardware and software engaged in delivering and storing data is optimized for fast, responsive, and reliable access.

DDN A³I solutions are designed, developed, and optimized in close collaboration with NVIDIA. The deep integration of DDN AI appliances with DGX systems ensures a predictable and reliable experience. DDN A³I solutions are highly configurable for flexible deployment in a wide range of environments and scale seamlessly in capacity and capability to match evolving workload needs. DDN A³I solutions are deployed globally and at all scale, from a single DGX system all the way to the largest NVIDIA DGX SuperPOD™ with DGX A100 in operation today.

DDN brings the same advanced technologies used to power the world's largest supercomputers in a fully-integrated package for DGX systems that's easy to deploy and manage. DDN A³I solutions are proven to provide maximum benefits for at-scale AI, Analytics and HPC workloads on DGX systems.



The DDN A³I Shared Parallel Architecture

The DDN A³I shared parallel architecture and client protocol provides superior performance, scalability, security, and reliability for DGX systems. Multiple parallel data paths extend from the drives all the way to containerized applications running on the GPUs in the DGX system. With DDN's true end-to-end parallelism, data is delivered with high-throughput, low-latency, and massive concurrency in transactions. This ensures applications achieve the most from DGX systems with all GPU cycles put to productive use. Optimized parallel data-delivery directly translates to increased application performance and faster completion times. The DDN A³I shared parallel architecture also contains redundancy and automatic failover capability to ensure high reliability, resiliency, and data availability in case a network connection or server becomes unavailable.

The DDN A³I client's NUMA-aware capabilities enable strong optimization for DGX systems. It automatically pins threads to ensure I/O activity across the DGX system is optimally localized, reducing latencies and increasing the utilization efficiency of the whole environment. Further enhancements reduce overhead when reclaiming memory pages from page cache to accelerate buffered operations to storage. The A³I DDN shared parallel architecture provides proven enablement and acceleration for AI infrastructure and workloads on DGX systems.

ai400x

Get Proven Performance with the DDN AI400X Appliance

The DDN AI400X is a turnkey appliance, fully-integrated and optimized for the most intensive AI and HPC workloads on DGX systems. The appliance is proven and well-recognized to deliver highest performance, optimal efficiency, and flexible growth for DGX deployments at all scale. A single appliance can deliver up to 48GB/s of throughput and well over 3 million IOPS to clients via a HDR100 or 100 GbE network, and can scale predictably in performance, capacity and capability. The AI400X appliance is available in all-nvme and hybrid NVME/HDD configurations for maximum efficiency and best economics. The unified namespace simplifies end-to-end deep learning workflows with integrated secure data ingest, management, and retention capabilities.

The AI400X achieves the most GPU performance, streamlines workflows, eliminates data management overhead. It enables customers to scale seamlessly, limitlessly and with full-confidence as workflow requirements increase. The appliance software is feature rich and includes extensive data management capabilities, robust data protection and security frameworks, intelligent analytics and analysis engines, and integrates a modern hybrid S3 object interface. The software also includes several advanced features ideal for deployments with multiple DGX systems, notably full support for container applications and secure multi-tenancy. It iinterfaces easily with file, object and cloud-based data repositories for ingest and archive.

The AI400X appliance is designed for rapid deployment, easy management and support. It's fully-validated and deployed with hundreds of DGX client nodes. The AI400X is provides best performance for all workloads and data types. It is the most-proven data platform with maximum operational flexibility at all-scale for DGX systems.

DDN A³I Solutions Guide with NVIDIA DGX A100 Systems

ddn.com/a3i

Deploy Rapidly with Fully-Validated Reference Architectures

DDN proposes reference architectures for single and multi-node configurations including DGX POD and SuperPOD. They are documented in the DDN A³I Solutions Guide with NVIDIA DGX A100 Systems available from the DDN website.

The DDN AI400X is a turnkey appliance for at-scale DGX deployments. DDN recommends the AI400X as the optimal data platform for DGX system deployments. The AI400X delivers maximum GPU performance for every workload and data type in a dense, power efficient 2RU chassis. The AI400X simplifies the design, deployment and management of DGX systems and provides predictable performance, capacity and scaling. The AI400X arrives fully configured, ready to deploy and installs in minutes. The appliance is designed for seamless integration with DGX systems, and enables customers to move rapidly from test to production. DDN provides complete expert design, deployment, and support services globally and ensures best customer experience. The DDN field engineering organization has already deployed hundreds of solutions for customers based on the A³I reference architectures.

As general guidance, DDN recommends an AI400X for every four DGX systems in a DGX POD (Figure 1). These configurations can be adjusted and scaled easily to match specific workload requirements. For the storage network, DDN recommends HDR200 InfiniBand technology in a non-blocking topology, with redundancy to ensure data availability. DDN recommends use of at least two HDR200 connections per DGX system to the storage network.

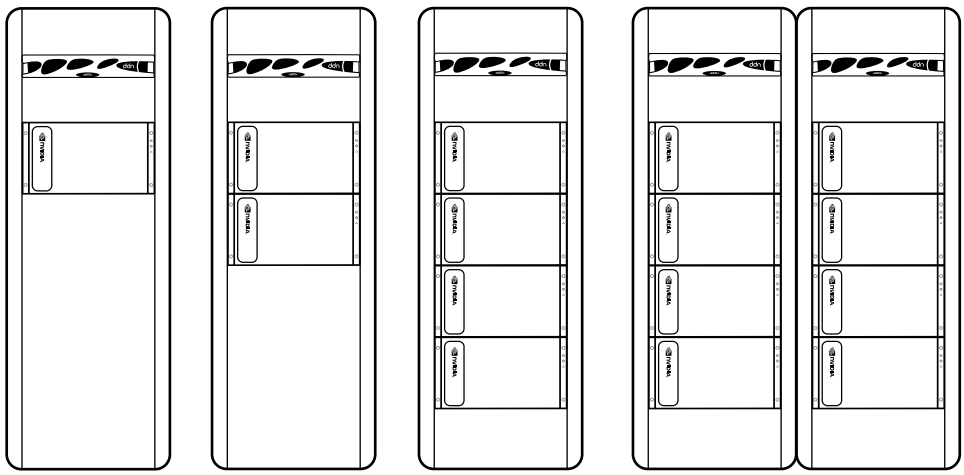
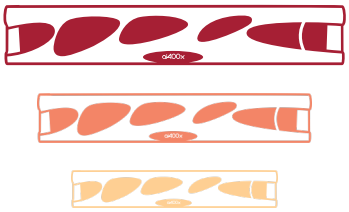


Figure 1. *Rack illustrations for DDN A³I reference configurations*



Scale Predictably with Multiple DDN AI400X Appliances

DDN A^{3}I solutions are widely recognized to deliver proven, seamless and predictable scaling in multiple dimensions. Every AI400X appliance provides well-defined performance, capacity and capability. This makes it simple to design a data platform that can scale reliably to meet evolving AI workflow needs. Testing demonstrates linear read and write throughput scaling with AI400X appliances (Figure 2). A single AI400X appliance can provide over 48 GB/s of read throughput and 35 GB/s of write throughput to a single DGX system, as illustrated on the left of the graph. An expanded system with two AI400X appliances is demonstrated to deliver over 99 GB/s read throughput and 72GB/s write throughput to a single DGX system, through a single mount point. This demonstrates that the full performance can be delivered to a single client node, and can be distributed with multi-node deployments.

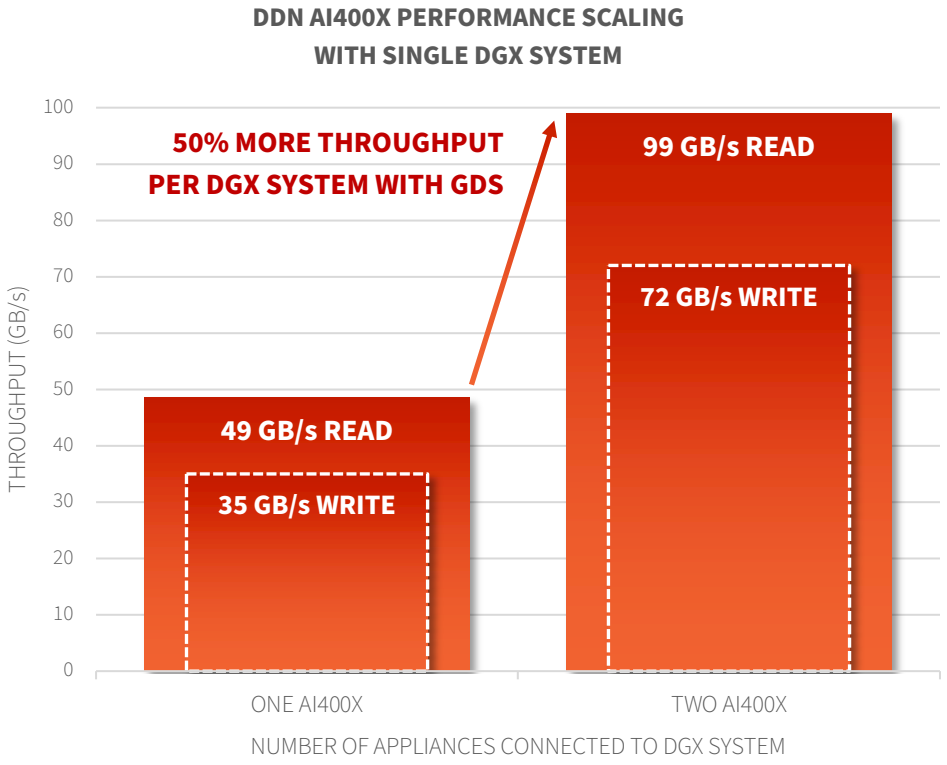


Figure 2. Scaling throughput performance on DGX system with AI400X appliances

Additional testing demonstrates that the DDN A3I shared parallel architecture enables a single DGX system to achieve scaled throughput and IOPS peak performance (Figure 3). The left graph illustrates peak read throughput performance of 99 GB/s to a single mount, single DGX system from dual AI400X appliances. This is 33X more read throughput than NFS, and nearly 10X more than NFS with ROCE. The left graph demonstrates peak IOPS read performance up to 4.7 million IOPS to a single DGX system, single mount with the same configuration. This is 46X more IOPS than NFS. This testing also clearly demonstrates that the DDN AI400X delivers uncompromising performance for a wide variety of data intensive workload, using a wide variety of data types.

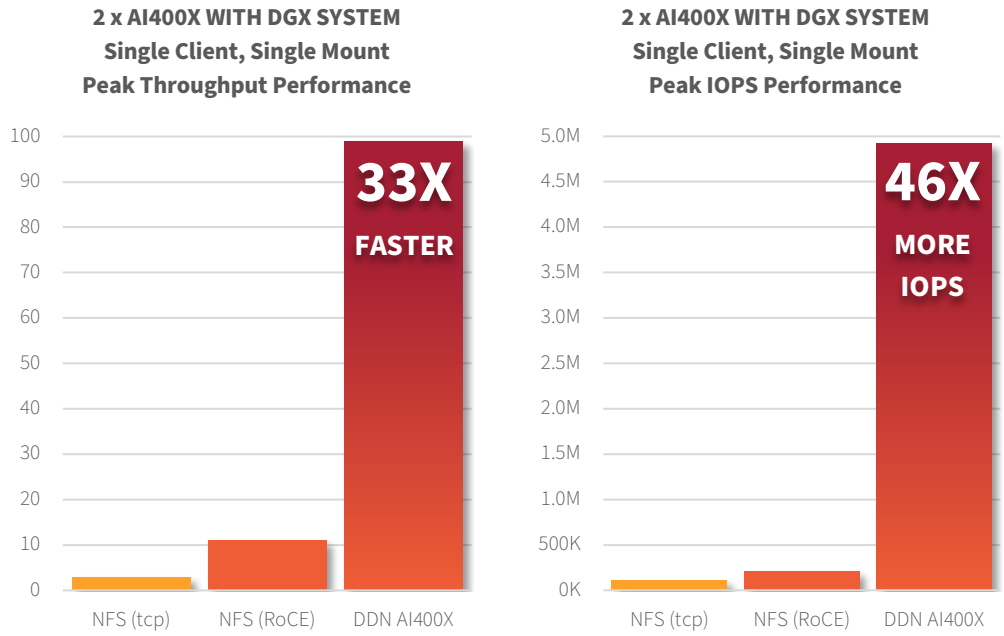


Figure 3. Peak performance for single mount on singe DGX system

DDN A^{3}I solutions are currently deployed at extreme scale and power the largest NVIDIA SuperPOD with DGX A100 in operation globally. Over 280 DGX systems access the shared DDN data platform simultaneously and engage a wide variety of HPC, AI and Analytics workloads using mixed data types. The deployment for this project started with ten AI400X appliances and three additional expansions of ten AI400X, for a total of forty AI400X appliances. At every phase, the ten DDN appliances delivered additional capacity and nearly 500 GB/s of throughput. Fully deployed, the forty AI400X appliances deliver 2 TB/s of throughput to all DGX systems in the SuperPOD, from a single unified namespace. This greatly simplifies data management, and eliminates the need to copy, move or tier data between different storage locations.

The DDN A³I shared parallel architecture is fully-optimized to deliver peak performance while ensuring most efficient use of system resources. Testing demonstrates that DDN can deliver data with high-throughput and minimal overhead to a DGX system (Figure 4). On the left of the graph are performance claims published from an NFS-based solution provider. A single client can read up to 33 GB/s throughput of data. However, this requires that nearly all CPU resources on the DGX system be consumed for IO and leaves nothing for applications running on the node. On the right of the graph, DDN demonstrates that an application can read data from two AI400X appliances with up to 99 GB/s of throughput, over 3X competitor claims and 33X more than with regular NFS. The DDN client software on the DGX system is fully-optimized and requires only 23% of CPU utilization to handle the full 99 GB/s throughput and leaves the vast majority of CPU resources available for other applications. This is especially important with deep learning, as certain steps like image decoding on ingest are run on CPU. The AI400X appliances ensures that DGX systems can achieve maximum throughput with highest efficiency.

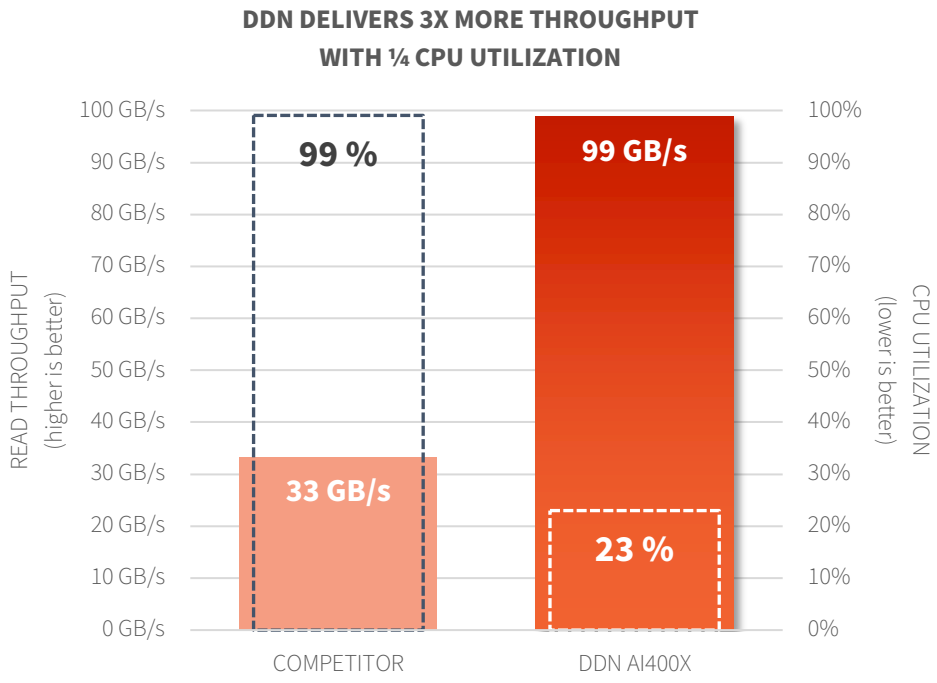


Figure 4. DGX system CPU utilization with DDN AI400X and NFS-based solution claims

Accelerate your AI Applications with DDN Shared Parallel Architecture

The DDN A^{3}I shared parallel architecture delivers data to GPUs with high-throughput, low-latency and massive concurrency. This ensures that all GPU cycles are put to productive use and achieve maximum AI and HPC application performance on DGX systems with any data type. For distributed workloads, performance scales linearly and maintains full GPU saturation as multiple GPUs are engaged. This contrasts heavily with legacy network protocols like NFS which are designed for modest workloads, small volumes of data and repeatedly proven inadequate to meet the demands of modern workloads running on GPUs.

The AI400X appliance delivers faster, scalable AI application performance with DGX system. Testing with PyTorch, a very commonly used deep learning framework demonstrates 3X higher application throughput and maintains linear performance scaling with multiple GPUs (Figure 5). This contrasts heavily when using NFS, which fails to fully-engage a single GPU and cripples the performance of the DGX system to less than a third of its capabilities. This test clearly demonstrates that efficient data delivery to GPUs with the DDN shared parallel architecture directly translates to increased AI application performance, and that this is maintained at-scale with additional GPUs and client nodes engaged.

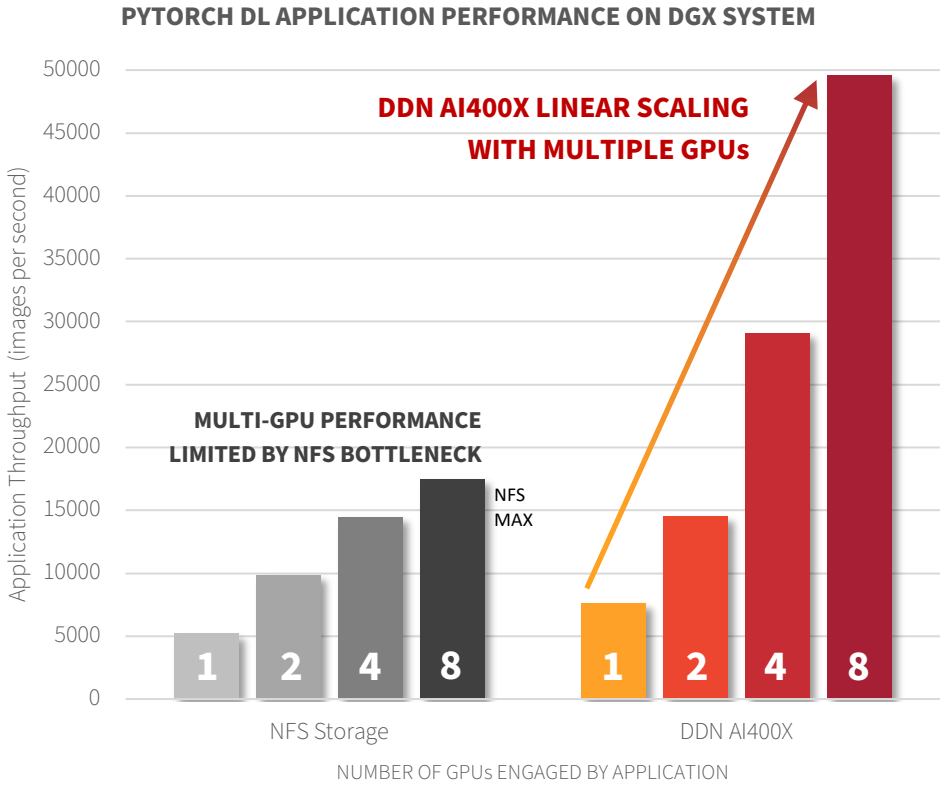


Figure 5. Scaling PyTorch application performance on DGX system

The benefits of the DDN shared parallel architecture for AI applications extend throughout the entire lifecycle of AI data including ingest, labeling, processing and archiving for long-term retention and reuse. Deep learning applications like PyTorch can take advantage of optimized data formats to achieve faster and more efficient runtime results.

TFRecords is a highly optimized file format compatible with PyTorch that enables the conversion of discrete data and metadata asset collections into series of streamlined binary files. This process significantly reduces the amount of dataset preparation time required before running the deep learning application. To be utilized, discrete assets must be split into training, testing, and validation sets that are stored in a specific folder structure and shuffled to avoid biased data distribution. This requires tedious data handling and attention to maintain proper shuffling. TFRecords provide a consolidated dataset that is easy to maintain and distribute and that eliminates the need for file manipulation.

The DDN shared parallel architecture furthers these benefits by allowing concurrent delivery of discrete data and metadata assets from source datasets to the conversion application, and rapid write of the binary file to persistent storage. In this demonstration, a dataset with 1.9 million data and metadata files spread across thousands of folders is being condensed to 1150 TFRecords binary files in a single directory, over 3X faster with DDN compared to NFS (Figure 6).

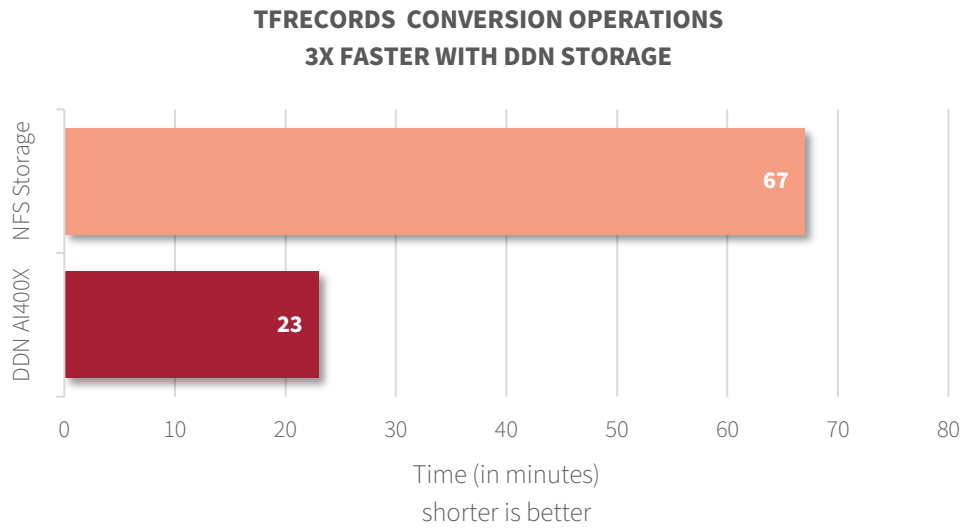


Figure 6. Comparing TFRecords conversion operation duration

TFRecords also streamlines PyTorch at runtime. Discrete assets must be opened individually, generating tremendous overhead for the data delivery and storage systems. A consolidated TFRecord binary file is more efficient as it only requires a single file open operation and allows the entire dataset to be held into a block of memory. This also enables applications to shuffle data at random places throughout the workflow and dynamically split training, testing and validation sets. This provides tremendous agility, efficiency and acceleration to Pytorch applications.

Testing demonstrates that Pytorch application performs at significantly higher throughput with both optimized and discreet data sets using DDN AI400X compared to NFS (Figure 7). The graph on the left illustrates application performance using discreet data set comprised of individual JPEG files. The DDN shared parallel architecture enables 4X faster image ingest than NFS. On the right, the same application ingests the same data that has been converted to TFRecords. The application performs at much higher throughput than using a discreet dataset and the DDN shared parallel architecture further compounds the application performance benefits, with 3X faster ingest than using a legacy NFS data platform.

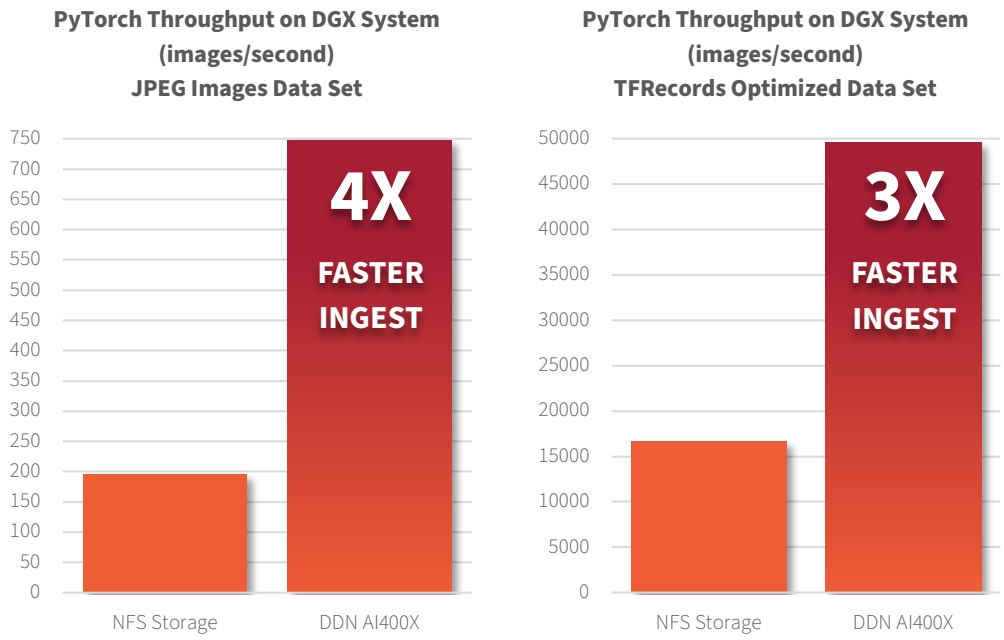
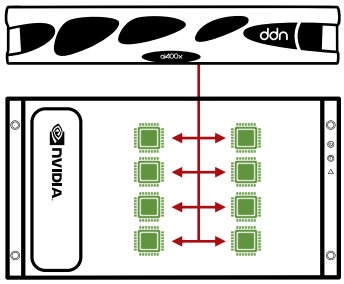


Figure 7. PyTorch application performance on DGX system with different data set formats

The AI400X appliance delivers faster, scalable AI application performance with DGX system, and the DDN shared parallel architecture provide clear acceleration and benefits at every stage of the end-to-end AI workflow, for every data type.



Maximize Throughput and Efficiency with NVIDIA GPUDirect Storage

DDN A^{3}I solutions interface directly with GPU memory for fastest and most efficient I/O operations possible. DDN is the first to fully integrate GDS which enables a direct DMA data path between GPU memory and storage, thus avoiding a bounce buffer through the CPU. This direct path increases system bandwidth while decreasing latency and utilization load on the CPU and GPU. The DDN shared parallel architecture combined with GDS enables customers to maximize DGX system I/O capabilities. With GDS, DDN can deliver over 162 GiB/s of throughput directly to GPU memory on a single DGX system, fully-saturating the network interfaces on the server, and delivering 57% more throughput than available over standard data paths. This significantly improves AI, Analytics and HPC application performance on DGX systems. GDS is fully implemented in current generation DDN AI storage appliances and validated with all GDS-supported DGX systems and with multi-node deployments.

Testing demonstrates significant read and write throughput performance benefits of DDN A^{3}I with GDS for DGX systems (Figure 8). The graph compares peak read and write throughput from a single client node. With GDS, the client performance increases 1.6X.

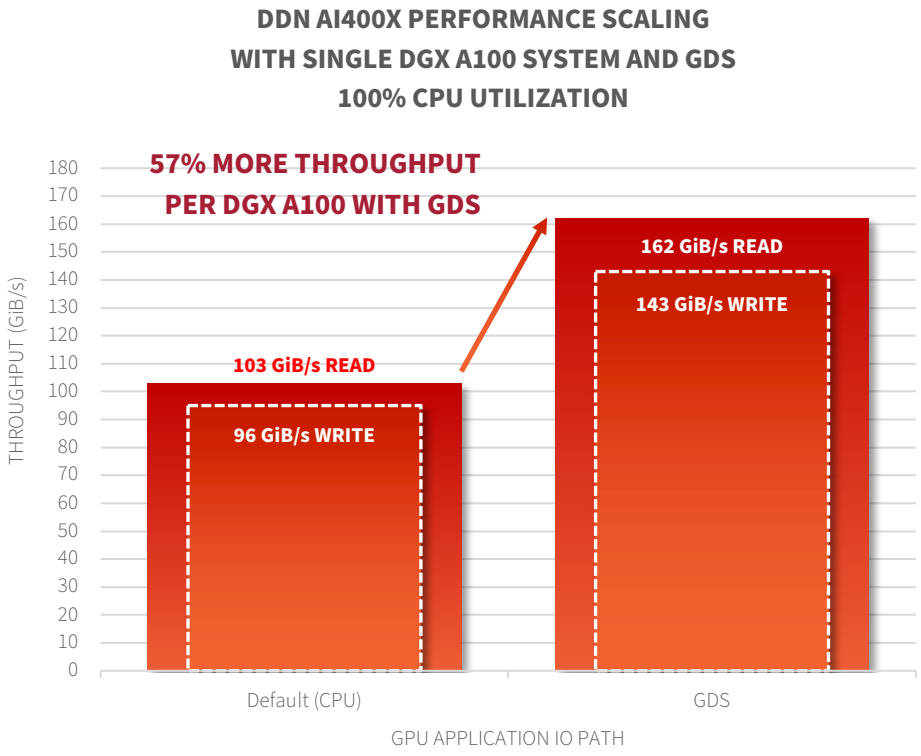


Figure 8. DDN A^{3}I with GDS increased throughput performance on DGX system

The architecture of the DGX system enables individual GPUs to consume up to 22 GiB/s of data from the network interface cards located on the same PCIe switch. Testing demonstrates that the DDN shared parallel architecture can fully saturate read throughput for all eight GPUs in the DGX system simultaneously, delivering over 162 GiB/s with linear performance scaling as more GPUs are engaged. The results also demonstrate almost full saturation of the eight HDR200 network interface cards simultaneously from a single shared data platform (Figure 9).

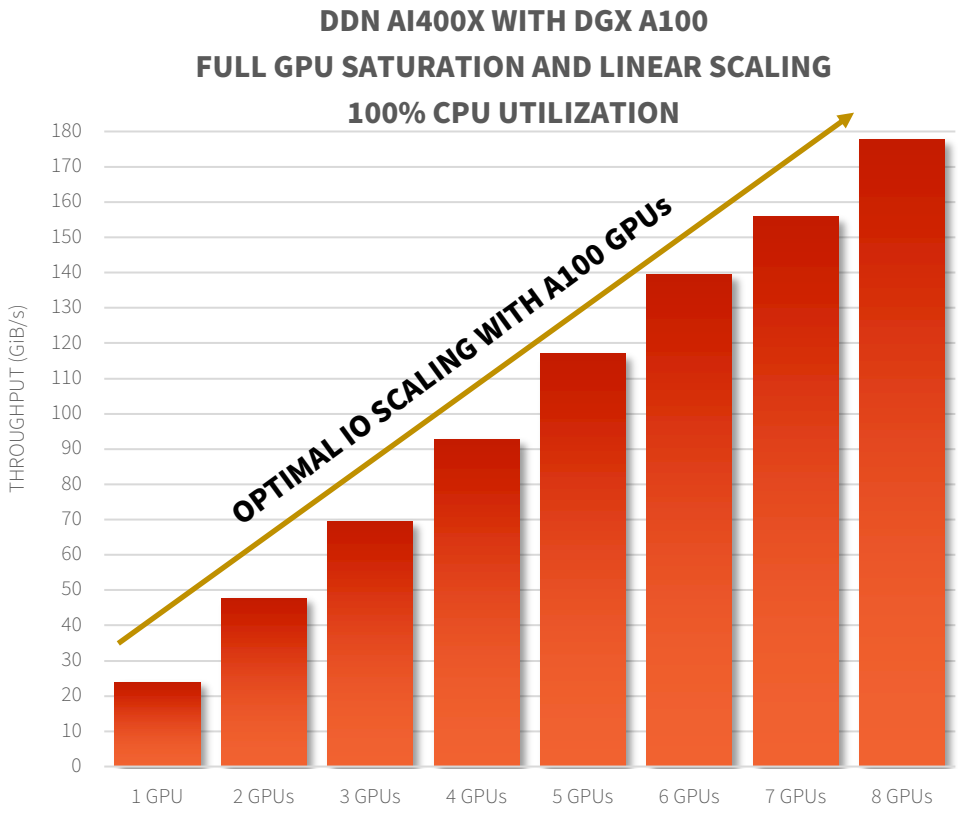


Figure 9. DDN A³I with GDS GPU throughput scaling on DGX system

Clearly, GDS provides significant advantage for applications running on DGX systems, and the DDN AI400X is proven to provide the performance required to maximize the value of GDS, especially with at-scale workloads and multi-node deployments. DDN appliances fully-integrate GDS and the capability can be enabled seamlessly for customers looking to engage in the ongoing early access program with NVIDIA. Contact DDN for more information.

Contact DDN to Unleash the Power of Your AI

DDN has long been a partner of choice for organizations pursuing at-scale data-driven projects. Beyond technology platforms with proven capability, DDN provides significant technical expertise through its global research and development and field technical organizations.

A worldwide team with hundreds of engineers and technical experts can be called upon to optimize every phase of a customer project: initial inception, solution architecture, systems deployment, customer support and future scaling needs.

Strong customer focus coupled with technical excellence and deep field experience ensures that DDN delivers the best possible solution to any challenge. Taking a consultative approach, DDN experts will perform an in-depth evaluation of requirements and provide application-level optimization of data workflows for a project. They will then design and propose an optimized, highly reliable and easy to use solution that best enables and accelerates the customer effort.

Drawing from the company's rich history in successfully deploying large scale projects, DDN experts will create a structured program to define and execute a testing protocol that reflects the customer environment and meet and exceed project objectives. DDN has equipped its laboratories with leading GPU compute platforms to provide unique benchmarking and testing capabilities for AI and DL applications.

Contact DDN today and engage our team of experts to unleash the power of your AI projects.

About DDN

DataDirect Networks (DDN) is the world's leading big data storage supplier to data-intensive, global organizations. DDN has designed, developed, deployed, and optimized systems, software, and solutions that enable enterprises, service providers, research facilities, and government agencies to generate more value and to accelerate time to insight from their data and information, on premise and in the cloud.

V2 11/20