



DDN A³I[®] Storage Platforms

DDN's A³I (Accelerated, Any-Scale AI) solutions break new ground for Artificial Intelligence (AI) and Deep Learning (DL), providing unmatched flexibility for your organization's AI needs.

Engineered from the ground-up for the AI-enabled data center, A³I solutions are optimized for ingest, training, data transformations, replication, metadata and small data transfers.

DDN offers flexibility in platform choice with the all-flash NVMe AI400X2, a hybrid flash and hard drive storage platform which leverages parallel access to flash and deeply expandable HDD storage. The AI400X2 supports a scale-out model with solutions starting at a few TBs yet scalable to 10s of PBs.

FULLY INTEGRATED, GPU-OPTIMIZED DATA PLATFORMS

Easy to deploy, A³I solutions are turn-key, pre-configured, and provide the most capable scale-out platform for capacity and performance.

FULL GPU SATURATION

Fully optimized for all types of I/O patterns and data layouts, A³I solutions deliver data to applications, ensuring full GPU resource utilization even with distributed applications running on multiple computing servers. Performance testing on the DDN A³I architecture has been conducted with all widely-used DL frameworks

(TensorFlow, Horovod, Torch, PyTorch, NVIDIA[®] TensorRT™, Caffe, Caffe2, CNTK, MXNET and Theano). Using the A³I intelligent client, containerized applications can engage the full capabilities of the data infrastructure, and that the AI servers achieves full GPU saturation consistently for DL workloads.

CAPACITY-EFFICIENT AI STORAGE

A³I solutions provide flexible capacity expansion options, with up to 500TB of scale-out NVMe capacity per AI400X2 appliance.

HIGHEST RESILIENCY, RELIABILITY & SECURITY AT SCALE

Engineered to provide the highest data availability and maximum system uptime, all A³I hardware and software components are integrated as a fully redundant system.

UNIFIED NAMESPACE

A³I solutions allow for consolidation of hot training data and warm expanding data libraries into a single platform, providing easy data access from a unified interface.

MULTI-TENANCY & QUOTA SUPPORT

A³I solutions can be secured on a per-tenant basis that ensure users and applications can only access the data that they're entitled to. Advanced quota controls provide easy management of file system consumption at the user, group, and project level.

PROVIDING UNMATCHED STORAGE PERFORMANCE & FLEXIBILITY FOR ARTIFICIAL INTELLIGENCE & DEEP LEARNING

10X SCALING EFFICIENCY

- > True parallel architecture performance, efficiency, GPU utilization, and storage capacity at any scale

10X DL ACCELERATION

- > Delivers full, real-time acceleration for all workloads concurrently and continuously.

ROCK SOLID & PROVEN AT SCALE

- > Data protection, integrity, declustering, redundancy ideal for mission critical applications.

EFFORTLESS DEPLOYMENT

- > Fully-integrated and optimized for AI workloads and GPU enabled solutions.

>_ AI400X2-QLC & AI400X2-TURBO



	AI400X2 QLC	AI400X2	AI400X2 TURBO
	MORE USEABLE CAPACITY	OPTIMAL PERFORMANCE/CAPACITY	EXTRA PERFORMANCE BOOST
Parallel File System Performance	90 GB/s (r), 70 GB/s (w), 3.5M IOPs	90 GB/s (r), 65 GB/s (w), 3M IOPs	115 GB/s (r), 75 GB/s (w)*
Appliance Capacity	1 or 2 PB useable (QLC)	120, 250, 500 TB useable (TLC)	120, 250, 500 TB useable (TLC)
Appliance Requirements	6 RU • 4.4 KW • 15K BTU/h	2 RU • 2.2 KW • 7.5K BTU/hr	2 RU • 2.2 KW • 7.5K BTU/hr
Controller Host Posts	HDR200/100GbE/200GbE QSFP 56 (8) or NDR200/200GbE QSFP 112 (8)	HDR200/100GbE/200GbE QSFP 56 (8) or NDR200/200GbE QSFP 112 (8)	NDR200/200GbE QSFP 112 (16)

* AI400X2-Turbo GA in Q1 2024. Current numbers are projected performance.