



## AI OPTIMISED STORAGE FOR DEEP LEARNING ACCELERATION

Maximizing GPU usage of your DGX systems throughout your AI journey

### 30% FASTER TRAINING

Real life deep learning projects show a massive 30% improvement in training times when compared to other solutions.

Excellent performance with the standard storage synthetic benchmarks, with bandwidth, latency and IOPS, leaving others behind.

### 50% LOWER COST

Cost and affordability are a key design focus. By removing the need for expensive storage controllers, costs are dramatically reduced, and more of your investment is spent on GPU and NVMe resource providing greater productivity and ROI.

### 100% SCALABILITY

With Scale-UP and Scale-OUT, both capacity and performance are not limited. Regardless of where you start, scaling is simple, fast and on demand.

Providing a future proof solution and benefiting from the lowering trend of flash costs.

### EXTENDING NVIDIA'S DGX RESOURCE

NVIDIA's DGX range has helped shape the AI landscape and changed future possibilities. However, the DGX range has limited internal space for NVMe flash storage, an essential element for performance and overall capability.

PNY 3S-2400 AI Optimised Storage Server creates a **central pool of ultra-low latency NVMe** which can be shared amongst one or multiple DGX servers. Providing each DGX with the ideal level of resource without the need for upfront over investment.

Simply connected via NVIDIA compatible EDR InfiniBand / 100Gbe, the unique RDMA protocol ensures the NVMe resource is seen and performs as if it were internal to the DGX.

### BLISTERING PERFORMANCE FOR AI WORKFLOWS

Today's AI servers consume and analyse data at much higher rates than traditional storage solutions can deliver. Resulting in low GPU utilisation and dramatically extending training times as well as decreasing productivity.

PNY 3S-2400 has been developed from the ground up for AI workloads and NVIDIA DGX optimisation. Delivering ultra-low latency and tremendous bandwidth at a price which allows more investment to be made on GPU resource and less on expensive, slower storage.

*Ensuring your project's funds are better spent, and your team are more productive by taking full advantage of the DGX capability.*

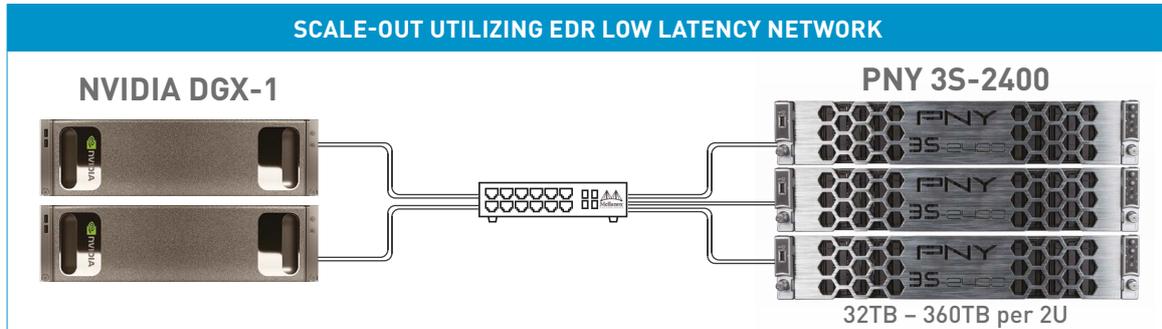
### START SMALL. SCALE ONLY WHEN NEEDED

Unlike most storage solutions where the initial investment dictates future growth and performance, often resulting in the need to overspend for potential future growth.

PNY's NVmesh design can scale in stages to suit your project without any limitation. Just purchase the capacity and performance you need today and feel secure that as you scale, so can your capacity and performance.



OR



## TAILORED RESILIENCE

Regardless of your resilience needs, PNY 3S-2400 storage server will maintain ultra-low latency and high performance for AI / DL workflows.

Resilience options range from RAID 0, through RAID 5 and 6 to full mirrored solutions with No Single Point of Failure.

## TECHNICAL SUPPORT

With full support and a range of on-site options, you can choose the package most effective for your organisation or project

## CLEAR AND SIMPLE USER INTERFACE



## FLEXIBLE DESIGN FLEXIBLE SOLUTIONS

The below shows three examples of configurations, starting with the entry level bundle, a 2U 32TB solution.



Common Solutions	2U Starter Bundle based on 8TB drives	120TB configuration based on 15TB drives	360TB on 2 nodes configuration based on 15TB drives
Form factor	2U	2U	4U
Drives	4 x 8TB	8 x 15TB	24 x 15TB (12 per node)
Capacity	32TB	120TB or 90TB in RAID6	360TB or 300TB in RAID 6
Max Capacity	192TB (24 drives)	360TB (24 drives)	720TB (48 drives)
Latency	< 100 µS	< 100 µS	< 100 µS
Bandwidth max	12GBsec (4 drives)	23GBsec	46GBsec
IOPS	2 million	4 million	8 million
Resnet50 Benchmarks (256)	6 500 img/sec* (based on a single DGX-1)	13 000 img/sec* (based on two DGX-1)	26 000 img/sec* (based on four DGX-1)
Connectivity	2 x Independent EDR InfiniBand / 100Gbe	2 x Independent EDR InfiniBand / 100Gbe	4 x Independent EDR InfiniBand / 100Gbe
NVMesh HW RAID	RAID 0, 1	RAID 0, 1, 10, 5**, 6**	RAID 0, 1, 10, 5**, 6**

\* 6 500 img/sec is the maximum efficiency for a NVIDIA DGX-1. Additional DGX-1's will increase the results  
 \*\*RAID5 and RAID6 based upon 15TB configurations, available in single of multiple nodes

CONTACT YOUR PNY REPRESENTATIVE AT

**PNYPRO@PNY.EU**

FOR FURTHER CONFIGURATION OPTIONS