

# PNY DATA CENTER SOLUTIONS



**PNY**  
GRAPHICS | HPC | AI

**PRO**



**nvidia**

# BENEFITS OF PNY DATA CENTER SOLUTIONS

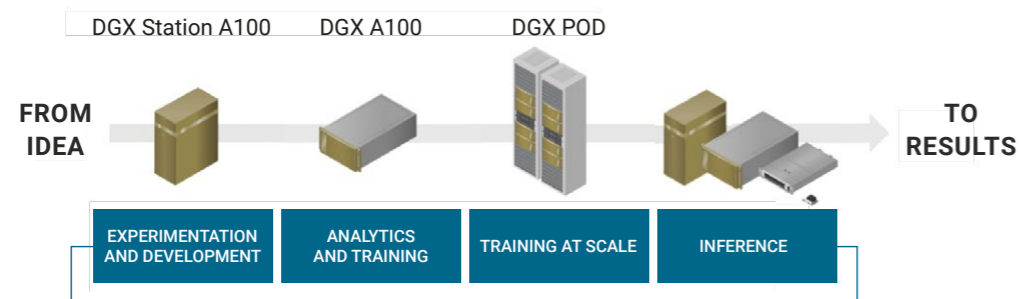
Today's enterprise needs an end-to-end strategy for AI innovation to accelerate time to insights and reveal new business frontiers. To stay ahead of the competition, they also need to construct a streamlined AI development workflow that supports fast prototyping, frequent iteration, and continuous feedback, as well as robust infrastructure that can scale in an enterprise production setting.

## ONE SYSTEM FOR ALL AI INFRASTRUCTURE

AI INFRASTRUCTURE RE-IMAGINED, OPTIMIZED, AND READY FOR ENTERPRISE AI-AT-SCALE

Flexible AI Infrastructure that adapts to the pace of enterprise:

- One universal building block for the AI data center
- Uniform, consistent performance across the data center
- Any workload on any node : any time
- Limitless capacity planning with predictably great performance with scale



## AI DRIVES REAL BUSINESS SUCCESS

Customer services, like Pinterest™, use deep learning for image recognition, curating rich custom experiences for its users. Financial businesses like PayPal™ are using NVIDIA GPU-accelerated deep learning for fraud detection.

Consulting firm Accenture™'s R&D arm and other businesses are using deep learning to detect Internet security threats. Healthcare benefits the most of the application of AI and deep learning.

This sweeping change in the landscape of data-driven business is taking place across all industries. It's changing how farmers herd their animals and leading us to rethink transportation. The first self-driving public transit system is already up and running, giving us a glimpse into futuristic cities that will exist within our lifetimes.

## PNY KEY SELLING POINT

PNY provides the most comprehensive computing, storage and networking portfolio for Datacenters. Its vast expertise and experience enable to support the challenges of the Datacenter coming from a broad range of industries such as Healthcare, Retail, Higher Education and more.

With its constant focus on innovation, quality, reliability, and support of excellence, the company helps its customers to empower and manage their infrastructure in the most technical and financial efficient ways.



**ARGONNE NATIONAL LABORATORY**  
World's First DGX A100  
Supercomputer Fighting COVID-19

24-node Cluster of DGX A100 Systems, 192 A100 GPUs,  
Mellanox High-Speed Low-Latency Network Fabric,  
120 PetaFLOPS of AI Computing Power for Scientific Research

# NVIDIA-CERTIFIED SYSTEMS™ FOR ENTERPRISES

SYSTEMS CERTIFIED BY NVIDIA FOR ACCELERATED COMPUTING FROM LEADING PARTNERS

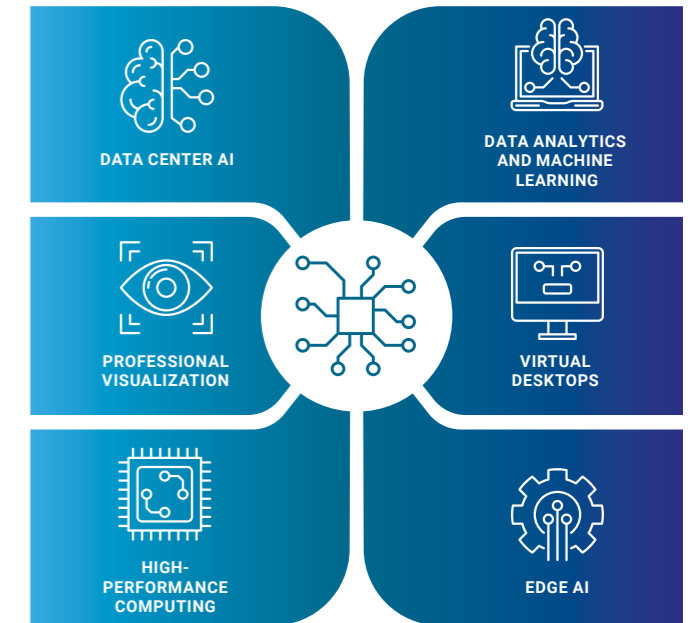
## NVIDIA-CERTIFIED SERVERS

NVIDIA-Certified servers create the essential platform for the evolution of enterprise data centers, delivering infrastructure that can handle a diverse range of accelerated workloads. The certification test suite exercises the performance and functionality of a configured server by running a set of software that represents a wide range of real-world applications.

### Benefits



### For all kind of workloads



## ENTERPRISE SOLUTIONS



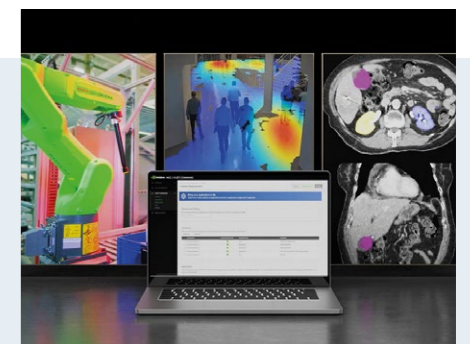
### NVIDIA AI Enterprise

NVIDIA AI Enterprise is an end-to-end, cloud-native suite of AI and data analytics software, optimized and certified by NVIDIA to run on VMware vSphere with NVIDIA-Certified Systems.



### NVIDIA Omniverse™ Enterprise

NVIDIA Omniverse™ Enterprise revolutionizes design collaboration and simulation. Globally dispersed teams can accelerate their workflows with one-click interoperability between leading software tools, and seamlessly collaborate in a shared virtual world running from the data center.



### NVIDIA Fleet Command™

NVIDIA Fleet Command™ is a cloud service that securely deploys, manages, and scales AI applications across distributed edge infrastructure. It's a turnkey solution for AI lifecycle management, offering streamlined deployments, layered security, and detailed monitoring capabilities.

# NVIDIA® DGX™ SYSTEMS

SETTING THE BAR FOR ENTREPRISE AI INFRASTRUCTURE

NVIDIA DGX™ systems are purpose-built to meet the demands of enterprise AI and data science, delivering the fastest start in AI development, effortless productivity, and revolutionary performance for insights in hours instead of months.



## NVIDIA DGX™ A100

AI Training, Inference, and Analytics

The third generation of the world's most advanced AI system, unifying all AI workloads.



## NVIDIA DGX Station™ A100

AI Workgroup Appliance

Server-class workstation ideal for experimentation and development by teams. No data center required.

# NVIDIA DGX POD™ & SUPERPOD™

BUILDING ENTERPRISE AI REQUIRES A SOLUTION THAT'S ENTERPRISE-READY

A solution that eliminates the complexity of designing AI infrastructure, integrated into your existing environment without upheaval, and that enables a simplified, fast deployment experience. NVIDIA DGX PODS, SuperPODS and PNY AI Cluster solution delivers the world's best platform for enterprise AI innovation, from AI infrastructure building block to Data Center as a product. Start Small, Scale Predictably in Response to Business Demand



## NVIDIA DGX POD™

A reference architecture that incorporates best practices for compute, networking, storage, power, cooling, and more, in an integrated AI infrastructure design built on NVIDIA DGX, from 2 to 10 nodes. Reference architectures available in association with storage partners such as DDN and NetAPP.

## NVIDIA DGX SUPERPOD™

The Fastest Path to AI-Innovation at Scale.

DGX SuperPOD now offers a complete, cloud native, AI data center for enterprise.

Become an AI Scaler, from 20 to 140 DGXA100, and more.



## System Specifications

	NVIDIA DGX A100 640 GB	NVIDIA DGX A100 320 GB	NVIDIA DGX Station A100 320 GB	NVIDIA DGX Station A100 160 GB
GPUS	8x NVIDIA A100 80 GB GPUs	8x NVIDIA A100 40 GB GPUs	4x NVIDIA A100 80 GB GPUs	4x NVIDIA A100 40 GB GPUs
GPU Memory	640 GB total	320 GB total	320 GB total	160 GB total
Performance	5 petaFLOPS AI 10 petaOPS INT8		2,5 petaFLOPS AI 5 petaOPS INT8	
NVIDIA NVSwitches	6		-	
System Power Usage	6,5 kW max		1,5 kW at 100-120 Vac	
CPU	Dual AMD Rome 7742, 128 cores total, 2.25 GHz (base) 3.4 GHz (max boost)		Single AMD Rome 7742, 64 cores total, 2.25 GHz (base) 3.4 GHz (max boost)	
System Memory	2 TB	1 TB	512 GB DDR4	
Networking	8x Single-Port Mellanox ConnectX-6 VPI 200 Gb/s HDR InfinBand 2x Dual-Port Mellanox Connect X- 6 VPI 10/25/50/100/200 Gb/s Ethernet	8x Single-Port Mellanox ConnectX-6 VPI 200 Gb/s HDR InfinBand 1x Dual-Port Mellanox Connect X- 6 VPI 10/25/50/100/200 Gb/s Ethernet	Dual-port 10Gbase-T Ethernet LAN Single-port 1Gbase-T Ethernet BMC management port	
Storage	OS: 2x 1.92TB M.2 NVME drives Internal Storage: 30 TB (8x 3.84 TB) U.2 NVMe drives	OS: 2x 1.92TB M.2 NVME drives Internal Storage: 15 TB (4x 3.84 TB) U.2 NVMe drives	OS: 1x 1.92 TB NVME drive Internal storage: 7,68 TB U.2 NVME drive	

## Example for a 140-node DGX SuperPOD configuration

Technology	Component	Description
NVIDIA DGX A100 System with eight 40 GB / 80 GB GPUs	Compute Nodes	1120 DGX A100 SXM4 GPUs • 89.6 TB of HBM2 memory • 336 AI PFLOPS via Tensor Cores • 280 TB System RAM • 4.4 PB local NVMe • 600 GBps NVLink bandwidth per GPU • 4.8 TBps total NVSwitch bandwidth per node
NVIDIA Mellanox Quantum™ QM8790 HDR 200 Gb/s IB Smart Switch	Compute Fabric	Full fat-tree network built with eight connections per DGX A100 system
NVIDIA Mellanox Spectrum® SN3700C switch	Storage Fabric	Fat-tree network with two connections per DGX A100 system
NVIDIA Mellanox Spectrum® SN3700C switch	In-band Management Network	One connection per DGX A100 system
NVIDIA Mellanox AS4610 Switch	Out-of-band Management Network	One connection per DGX A100 system
DeepOps	Management Framework	Software tools for deployment and management of SuperPOD nodes and resources
NVIDIA Magnum IO™ Technology	DGX SuperPOD Software Stack	Suite of library technologies that optimize GPU communication performance
NVIDIA CUDA-X™ Technology		A collection of libraries, tools, and technologies that maximize application performance on NVIDIA GPUs
NVIDIA NGC	User Runtime Environment	Containerized DL and HPC applications, optimized for performance
Slurm		Orchestration and scheduling of multi-GPU and multi-node jobs

FOR MORE INFORMATION:

Visit: [WWW.PNY.EU](http://WWW.PNY.EU)

The PNY logo is a registered trademark of PNY Technologies, Inc. All other trademarks are the property of their respective owners. All rights reserved. - © 2021 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, NVIDIA-Certified Systems, NVIDIA Omniverse, NVIDIA Fleet Command, NVIDIA DGX POD, NVIDIA DGX SuperPOD, NVIDIA DGX A100 and NVIDIA DGX Station are trademarks and/or registered trademarks of NVIDIA Corporation. All company and product names are trademarks or registered trademarks of the respective owners with which they are associated. Features, pricing, availability, and specifications are all subject to change without notice. - © 2021 Argonne National Laboratory. All rights reserved. ACMK0727

# PNY 3S STORAGE SERVERS FOR DATACENTERS

AI OPTIMISED STORAGE FOR DEEP LEARNING ACCELERATION AND INFERENCE



PNY 3S Storage Servers create a central pool of ultra low latency NVMe which can be shared amongst one or multiple NVIDIA DGX Servers, providing each NVIDIA DGX with the ideal level of resources without the need for upfront over investment.

## System Specifications

	PNY 3S-1050	PNY 3S-2450
Drives	Up to 10 NVMe	Up to 24 NVMe
Storage Capacity (raw)	Up to 150TB	Up to 360TB
Storage Capacity (protected R5)	Up to 135TB	Up to 345TB
Expansion units	Up to 5	Up to 5
Total expansion capacity	Up to 750TB	Up to 2PB
RAID Levels	0, 1, 10, 5, 6 (50, 60)	0, 1, 10, 5, 6 (50, 60)
Bandwidth	23GB/s	23GB/s
IOPS	3M IOPS	3M IOPS
Latency	35µs	35µs
Connectivity	2 x QSFP56 HDR InfiniBand /200Gb Ethernet	2 x QSFP28 EDR InfiniBand /100Gb Ethernet



### 50% Faster Training

Real life deep learning projects show a massive 50% 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.



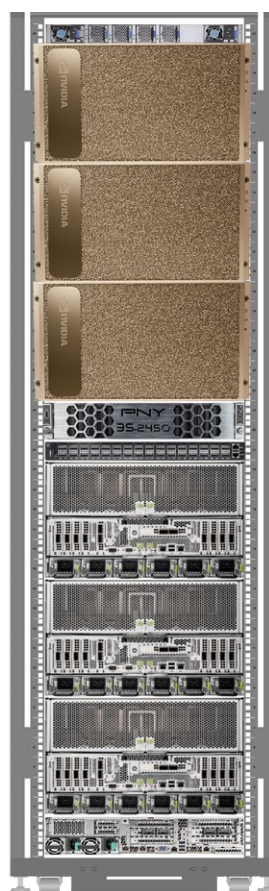
### 50% Lower Cost

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 up to 360TB within 2U and a massive 150TB within the 1U, even solutions starting at 30TBs have can scale in stages that suit your project.



## PNY AI CLUSTER

Pre-configured AI Infrastructure stacks, fully integrated and installed by PNY. Combining the NVIDIA DGX A100 compute solution, NVIDIA Networking and PNY 3S-STORAGE solution, from 2 to 10 nodes. There is no need for multiple storage nodes or controllers, everything needed is contained and automated within the single appliance.

# NVIDIA NETWORKING SOLUTIONS

UNLOCKING SYSTEM PERFORMANCE

Networking technology is at the heart of a new "agile computing" model. NVIDIA® Networking Solutions unlock the full potential of your Data Center.



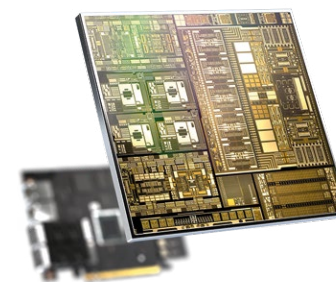
## ACCELERATING COMPUTE AND CLOUD DATA

NVIDIA Networking Solutions technologies is a leading supplier of end-to-end Ethernet and InfiniBand intelligent interconnect solutions and services for servers, storage and hyper-converged infrastructure. With NVIDIA Mellanox solutions, businesses can take advantage of their Data Center full potential while reducing total cost of ownership.

## SMART HIGH PERFORMANCE SOLUTIONS

There is a variety of smart high performance solutions, multi-core processors, network adapters, switches, cables, software and silicon.

### Complete Infiniband and Ethernet portfolio



### DPU / SmartNICs

The NVIDIA® BlueField-2® data processing unit (DPU) is modern data center infrastructure on a chip, delivering accelerated software-defined networking, storage, security, and management services.

Leveraging faster speeds and innovative In-Network Computing, NVIDIA InfiniBand smart adapters achieve extreme performance and scale.



### Switches

Best-in-class Ethernet switch portfolio, in Edge, Leaf, and Spine form factors, ranging in speeds from 10 to 400Gbps. adapters, switches, cables, software and silicon.

NVIDIA InfiniBand switch systems deliver the highest performance and port density for HPC AI, Web 2.0, big data, cloud infrastructure and enterprise data centers.



### Interconnect

LinkX InfiniBand DACs, AOCs and transceivers meet or exceed all the IBTA industry standards for QDR, FRD, EDR, and HDR products.

LinkX Ethernet DACs, AOCs and Optical Transceivers meet or exceed all of the IEEE 802.3xx industry standards for 1G, 10G, 25G, 50G and 100G products.



FOR DISCOVERING ALL OUR RECOMMENDATIONS:  
Please visit our [PNY Networking Solutions](https://products.pny.eu/)  
[HTTPS://PRODUCTS.PNY.EU/](https://products.pny.eu/)

# PNY ADVANTAGE

## “PNY PROVIDES UNSURPASSED SERVICE AND COMMITMENT TO ITS DATA CENTER CUSTOMERS”

- 20 Years expertise selling NVIDIA GPU Solutions
- Strong alliances with technological suppliers
- Dedicated head count for Sales, Marketing and Support
- Local Pre and post sales support
- Direct tech support hotlines
- Pre-sales tools, support and configuration assistance
- Dedicated Field Application Engineers, added-value on site installation
- Published product support and training materials
- Advanced replacement options for mission-critical deployments
- Long product life cycles and availability
- Loyalty partner channel programs
- Dedicated support programs
- Strong logistic and operation abilities
- Equipment loan for strategic opportunities
- PNY LAB: technological centre to support development of AI, HPC and VDI Solutions

## MARKETS WE SERVE



HPC



Data Center



AI/Machine Learning



Clouds



Storage



Security



Telecom



Media & Entertainment

## CONTACT

### PNY Technologies Europe

Zac du Phare  
9 rue Joseph Cugnot - BP 40181  
33708 Mérignac Cedex, France

Tel: +33 (0)5 56 13 75 75

### PNY Technologies GmbH

Schumanstraße 18a  
52146 Würselen  
Germany

Tel: +49 (0)2405/40848-0

### PNY Technologies Middle East Fze

Jafza View 19 308, Jebel Ali Free Zone,  
PO Box 263897,  
Dubai

Tel: +97148814966

e-mail: [SALES@PNY.EU](mailto:SALES@PNY.EU)