Turnkey NVIDIA DGX POD/SuperPOD
With DDN A³I Solutions

Overview

The next generation artificial intelligence (AI) supercomputing infrastructure, providing the computational power necessary to train today’s state-of-the-art deep learning (DL) models and to fuel innovation well into the future. The AMAX Turnkey NVIDIA DGX SuperPOD delivers groundbreaking performance and is designed to solve the world’s most challenging computational problems.

Simplified AI
Predictable performance, capacity and scaling

Seamless Integration
Move rapidly to production at scale

Tested and Validated
Reference architecture for the enterprise

Expert Services
Leverage our experience for your peace of mind

Fully Configured
Turnkey solution ready to deploy and install in minutes

Highlights

Networking
• 4x NVIDIA Mellanox QM8700
  » 2 Storage Network Switches
  » 2 Compute Network Switches

Compute
• 2x NVIDIA’s DGX A100 Systems

Cluster Management
• Bright Cluster Manager

Support & Services
• AMAX Professional Services

Storage and Data Management
• 1x DDN AI400X

DGX SUPERPOD SOLUTION FOR ENTERPRISE

• 100-700 PFLOPS AI system
• 20-140 NVIDIA DGX A100 systems with NVIDIA BlueField DPUs
• 1-10PB high-performance storage
• 200Gbps NVIDIA networking fabric
Solving the Challenge of Large-Scale, Multi-Node AI Infrastructure

NVIDIA DGX SuperPOD is designed to tackle the most important challenges of AI at scale, delivering unmatched levels of multi-system training. Traditional large compute clusters are constrained by the complexity of scaling inter-GPU communications as configurations become larger and computation is parallelized over more and more nodes. This results in diminishing performance returns. DGX SuperPOD solves this scaling problem by optimizing every component in the system for the unique demands of multi-node AI infrastructure.

High-Performance Infrastructure in a Single Solution—Optimized for AI

NVIDIA DGX SuperPOD brings together a design-optimized combination of AI computing, network fabric, storage, and software. Its compute foundation is built on NVIDIA DGX™ A100, the universal system for all AI workloads, which provides unprecedented compute density, performance, and flexibility. NVIDIA DGX A100 systems, available with up to 640 gigabytes (GB) of total GPU memory each, feature the world’s most advanced accelerator, the NVIDIA A100 Tensor Core GPU, enabling enterprises to consolidate training, inference, and analytics in a unified, easy-to-deploy AI infrastructure.

Solving the Challenge of Large-Scale, Multi-Node AI Infrastructure

NVIDIA DGX SuperPOD is designed to tackle the most important challenges of AI at scale, delivering unmatched levels of multi-system training. Traditional large compute clusters are constrained by the complexity of scaling inter-GPU communications as configurations become larger and computation is parallelized over more and more nodes. This results in diminishing performance returns. DGX SuperPOD solves this scaling problem by optimizing every component in the system for the unique demands of multi-node AI infrastructure.

High-Performance Infrastructure in a Single Solution—Optimized for AI

NVIDIA DGX SuperPOD brings together a design-optimized combination of AI computing, network fabric, storage, and software. Its compute foundation is built on NVIDIA DGX™ A100, the universal system for all AI workloads, which provides unprecedented compute density, performance, and flexibility. NVIDIA DGX A100 systems, available with up to 640 gigabytes (GB) of total GPU memory each, feature the world’s most advanced accelerator, the NVIDIA A100 Tensor Core GPU, enabling enterprises to consolidate training, inference, and analytics in a unified, easy-to-deploy AI infrastructure.

Solving the Challenge of Large-Scale, Multi-Node AI Infrastructure

NVIDIA DGX SuperPOD is designed to tackle the most important challenges of AI at scale, delivering unmatched levels of multi-system training. Traditional large compute clusters are constrained by the complexity of scaling inter-GPU communications as configurations become larger and computation is parallelized over more and more nodes. This results in diminishing performance returns. DGX SuperPOD solves this scaling problem by optimizing every component in the system for the unique demands of multi-node AI infrastructure.