PURESTORAGE®

NVIDIA

CONFIGURATION GUIDE

AIRI Architecture Overview and Configuration

Contents

Abstract

As deep learning (DL) training and other computationally-intensive workloads become increasingly prevalent in data centers, the need to scale out GPU-accelerated computing infrastructure becomes a significant challenge that IT administrators must address. Deep learning is a fundamentally different workload than traditional enterprise applications running on CPU-based servers. DL necessitates consideration of specific networking, storage, and infrastructure management approaches proven to enable improved scalability, performance, and cost-effective manageability.

This guide provides a prescriptive solution for IT administrators to deploy a validated solution designed for DL, inference, data exploration, and other computationally intensive algorithms that enable DL today. An AIRI A100 POD is a platform that includes NVIDIA® DGX™ A100 systems, Pure Storage® FlashBlade® storage, networking, NVIDIA NGC™ DL containers, and DGX POD management systems and software. This guide is designed to support organizations looking to satisfy the needs of multiple workloads, from providing as-a-service access for small interactive jobs to supporting cluster-wide jobs that make full use of multi-GPU and multinode resources.

Introduction

AIRI is an optimized platform containing DGX A100 systems, Pure Storage FlashBlade storage, utility servers, and networking switches to support various job types. AIRI's design allows for multiple concurrent small interactive jobs, multinode AI model training, inference, and data exploration work, all using NVIDIA AI software.

Because of its integrated nature, AIRI is exceptionally scalable, allowing for flexible deployments based on workload, environmental, and budget constraints. In this Configuration Guide, we demonstrate four-node systems as an example.

> **Note:** We used two- and four-node DGX A100 configurations as example configurations for validating this guide. Contact an NVIDIA Partner for other possible configurations.

Network Architecture

The DGX A100 system has four networks:

- **Compute fabric:** Connects the eight 200GbE NVIDIA Mellanox ConnectX®-6 HCAs (Host Channel Adapters) from each DGX A100 through separate network planes for inter-node communication. You can configure them in IB or Ethernet mode.
- **Storage fabric:** Uses one port from each dual-port ConnectX-6 HCA to connect with FlashBlade, cabled for 100GbE.
- **In-band management:** Uses the remaining two QSFP56 ports on the DGX A100 system to connect to a dedicated Ethernet network for provisioning and job scheduling.
- **Out-of-band management:** Connects the 1GbE RJ45 BMC port of each DGX A100 system to an additional Ethernet switch for node management and monitoring.

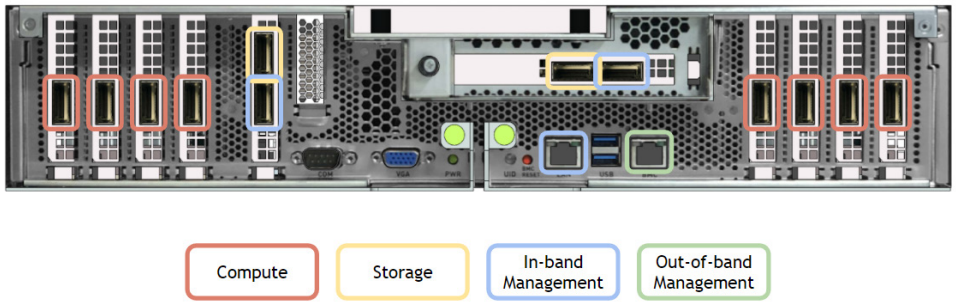

Figure 1. Network connections for DGX A100 system

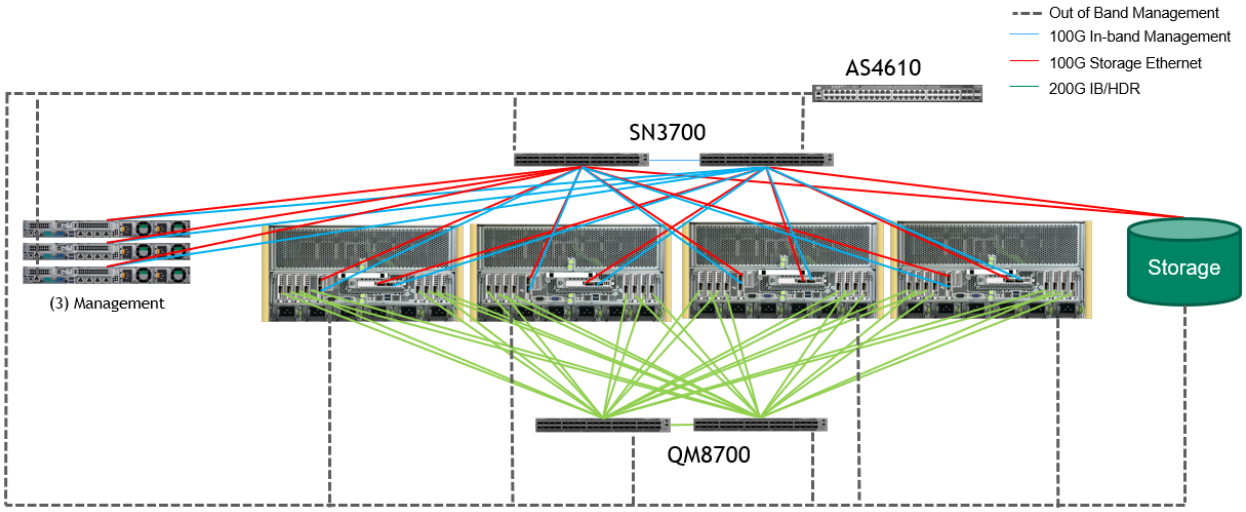

Figure 2. AIRI network connectivity overview

Compute Fabric

For a DGX A100 system, the inter-GPU compute fabric may be InfiniBand (IB) or RDMA over Converged Ethernet (RoCE). This configuration is entirely independent of the DGX-to-FlashBlade fabric configuration. Each DGX A100 system has eight connections for the compute fabric (Figure 3). The fabric design maximizes performance for typical communications traffic of AI workloads and provides some redundancy in the event of hardware failures and minimizing cost.

InfiniBand is the preferred interconnect between DGX A100 systems because it provides performance and scalability with little management and switch equipment overhead. InfiniBand is the default interconnect between DGX nodes regardless of whether you've selected IB or RoCE for the storage fabric. NVIDIA officially supports roCE as an alternative configuration of the compute fabric.

This guide details both network configurations.

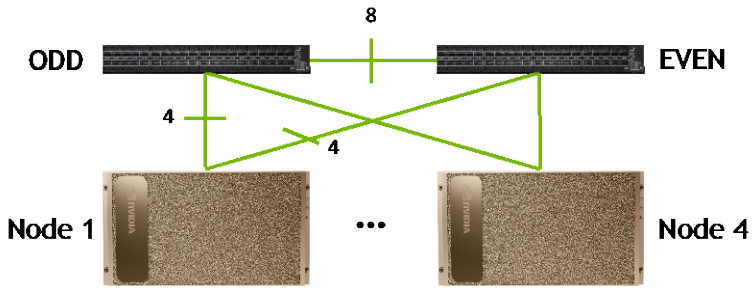

Figure 2. Compute fabric topology for a four-node system

Note: Eight links are provisioned between the ODD-EVEN switch to support some multinode traffic patterns, which must span switches. In cases where the workload is known not to need this, it may be reduced to two or four links.

InfiniBand

If using IB mode, the compute fabric may utilize QM8700 Series Mellanox Quantum HDR 200Gb/s InfiniBand Smart Switches. Each switch includes 40 QSFP56 ports used for communication to each DGX A100 system and between switches in the compute fabric. All connections are HDR, maximizing the bandwidth between network elements. No InfiniBand partitioning or other segmentation is used, with the QM8700 switch providing the subnet manager for the compute fabric. Connection to the out-of-band management ports on the switch to the out-of-band management fabric may be made if needed, but it is not critical to the AIRI system's operation.

Figure 3. NVIDIA Mellanox QM8700 HDR 200 Gb/s InfiniBand Switch

RoCE

If using RoCE, the compute fabric may utilize NVIDIA Mellanox Spectrum SN3700 Open Ethernet Switches. The 32 QSFP56 ports on compute fabric switches are configured for 200Gb/s operation and used for connecting to the DGX A100 systems. The switch is configured for RoCE traffic and uses VLANs for each discrete subnet configured on each DGX A100 NIC. Support is not available for All-All communication between GPUs in the cluster as it requires Layer 3 routing. This may be supported in a future update.

Figure 4. NVIDIA Mellanox Spectrum SN3700 Open Ethernet Switch

Storage Fabric

The storage fabric employs an Ethernet network fabric (Figure 6). In most configurations, the connections on the DGX A100 system will be bond Ethernet using the Linux bond driver.

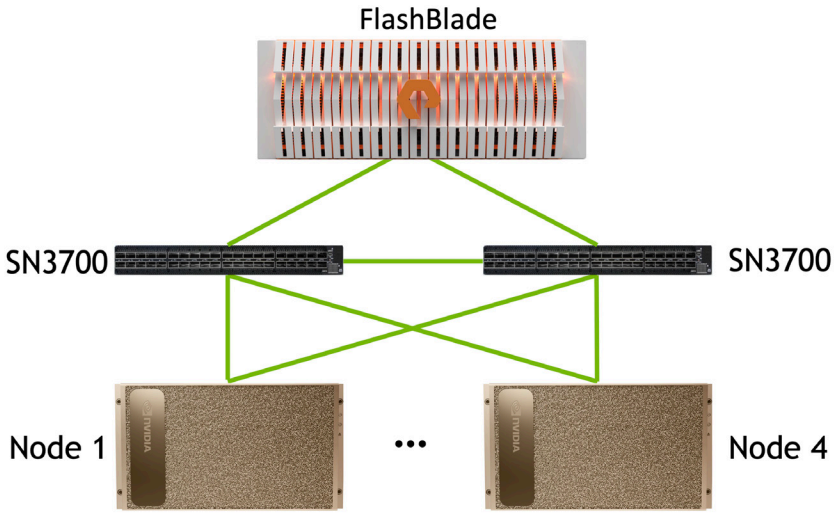

Figure 5. Storage fabric topology for a four-node AIRI deployment

Each node has two connections originating on one port on each of the dual-port NVIDIA Mellanox ConnectX®-6 HCAs. For DGX POD, the recommended number of Inter-Peer Links is 4x200G, although 2x200G is also acceptable.

Note: The AIRI configuration guide supports using Top-of-Rack switches from NVIDIA Mellanox, Cisco, or Arista.

In-Band Management Network

The in-band Ethernet network provides connectivity for cluster services such as Slurm and Kubernetes and other services outside of the cluster, such as the NGC registry, code repositories, and data sources.

The in-band network may use the SN3700 switches (Figure 5) from the storage fabric. For most configurations, two SN3700 switches will be sufficient to support all the in-band and storage network connections. In-band management and storage networks are VLAN-separated in typical deployments. Many storage controllers currently support only 100GbE, so that additional switch ports may be required depending on the storage controller used. These switches may be connected directly to the data center core switch, with traffic routed to the storage or management network as needed to integrate with existing infrastructure and cluster access mechanisms.

Out-of-Band Management Network

The out-of-band Ethernet network is used for system management via the Baseboard Management Controller (BMC) and provides connectivity to manage all networking equipment. Out-of-band management is critical to the cluster's operation by providing low usage paths that ensure management traffic does not conflict with other cluster services.

The out-of-band management network is based on 1GbE NVIDIA Mellanox AS4610 switches (Figure 7), running the DevOps-friendly Cumulus Linux network operating system. These switches are usually also connected to the data center management network. It is recommended that the serial console of the out-of-band management network switches be connected to a console server in the data center to facilitate reconfiguration and connectivity in the event of a network failure.

Figure 6. 1GbE NVIDIA Mellanox AS4610 Data Center Open Ethernet Switch

NVIDIA AI Software

The value of the AIRI architecture extends beyond its hardware. AIRI is a complete system providing the major components for system management, job management, and optimizing workloads to ensure quick deployment, ease of use, and high availability.

- The software stack begins with the DGX Operating System (DGX OS), which is tuned and qualified for DGX A100 systems.
- Nvidia GPU Cloud (NGC) is a key component of AIRI, providing the latest DL frameworks. NGC provides packaged, tested, and optimized containers for quick deployment, ease of use, and the best performance on NVIDIA GPUs.
- Lastly, essential tools such as CUDA-X, Magnum IO, and RAPIDS provide developers the tools they need to maximize DL, HPC, and data science performance in multinode environments.

DGX OS and POD Management

NVIDIA AI software running on AIRI provides a high-performance DL training environment for large-scale multi-user AI software development teams. In addition to the DGX OS, it contains cluster management, orchestration tools and workload schedulers (DGX POD management software), NVIDIA libraries and frameworks, and optimized containers from the NGC container registry. For additional functionality, NVIDIA's DGX POD management software includes third-party opensource tools recommended by NVIDIA and tested to work on AIRI racks with the NVIDIA AI software stack. Support for these tools is available directly from third-party support structures.

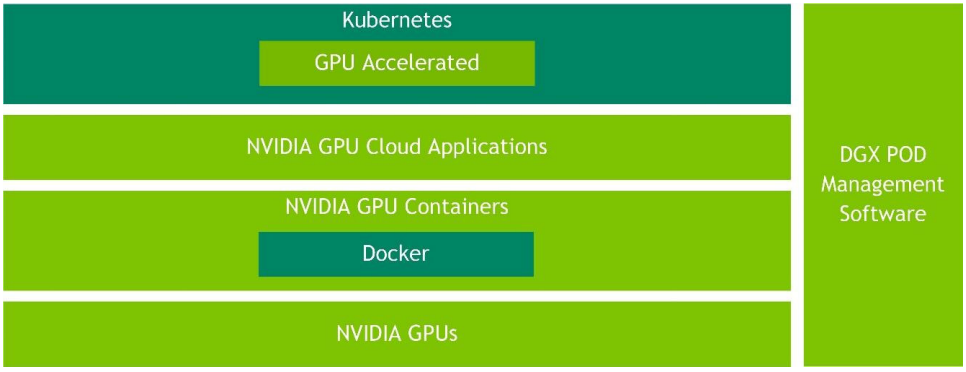

Figure 7. AI software stack

The foundation of the NVIDIA AI software stack is the DGX OS. Whether used with DGX OS with Ubuntu or Red Hat, the DGX software includes certified GPU drivers, a network software stack, pre-configured local caching, NVIDIA data center GPU management (DCGM) diagnostic tools, and GPU-enabled container runtime, all certified to work with NVIDIA NGC containers.

The DGX POD management software is composed of various services running on the Kubernetes container orchestration framework for fault tolerance and high availability. Services are provided for network configuration (DHCP) and fully automated DGX OS software provisioning over the network (PXE). The DGX OS software can be automatically re-installed on demand by the DGX POD management software.

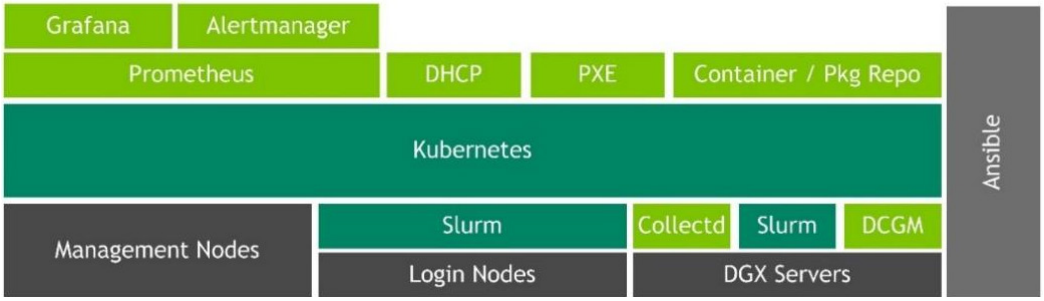

Figure 8. DGX POD management software

DeepOps, NVIDIA's DGX POD management software, allows users to provision, deploy, manage, and monitor DGX servers and NVIDIA Mellanox switches in a DGX POD. Services are hosted in Kubernetes containers for fault tolerance and high availability. The DGX POD management software leverages the Ansible configuration tool to install and configure all the tools and packages needed to run the DGX systems. System data collected by Prometheus is reported through Grafana. Alertmanager can use the collected data to send automated alerts.

For sites operating in an air-gapped environment or with additional on-premises services, a local container registry can be run on the Kubernetes management layer to provide services to the cluster. This local registry can mirror NGC containers as well as OS and Python packages.

DeepOps can deploy Slurm or Kubernetes as the orchestration and workload manager.

NVIDIA NGC

The NGC (Figure 10) provides a range of container images that meet the needs of data scientists, developers, and researchers with various levels of AI expertise. These users can quickly deploy AI frameworks with containers, get a head start with pre-trained models or model training scripts, and use domain-specific workflows and Helm charts for the fastest AI implementations, giving them faster time-to-solution.

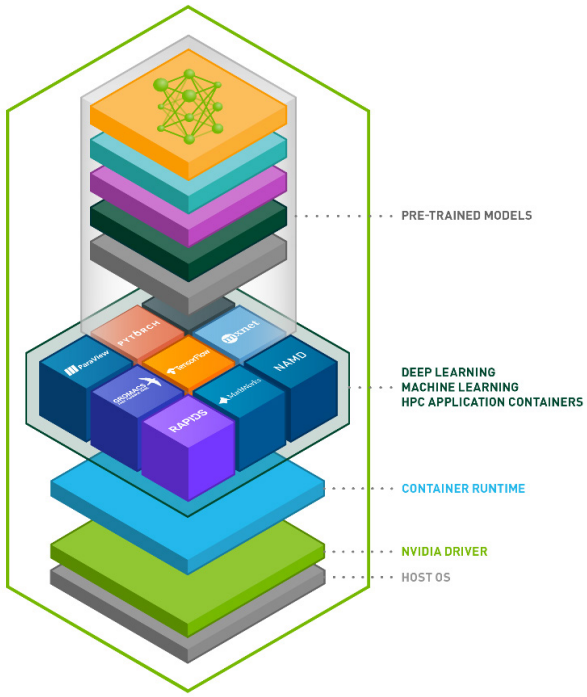

Figure 10. NGC components

Spanning AI, data science, and high-performance computing (HPC), the container registry on NGC features an extensive range of GPU-accelerated software for NVIDIA GPUs. The NGC hosts containers for the top AI and data science software. Containers are tuned, tested, and optimized by NVIDIA. Other containers for additional HPC applications and data analytics are fully tested and made available by NVIDIA. NGC containers provide powerful and easy-to-deploy software proven to deliver the fastest results, allowing users to build solutions from a tested framework with complete control.

Developers should stay abreast of significant changes in the software stack. Taking advantage of library-based performance improvements means faster time-to-production with new models. The NGC makes it easy for data teams to consume the latest end-to-end application stack with the best performance. It's a one-stop-shop where data scientists obtain a single Docker image with the latest libraries and tools pre-integrated.

Major Components

When planning an AIRI deployment, you need to consider several factors to determine the total floorspace and resources needed. The design explicitly described in this configuration guide uses at minimum two data center racks: one as a utility rack for login, management, and storage servers, plus additional racks for the DGX A100 systems that comprise the minimum compute muscle. When building out a four-DGX AIRI system, typical configurations may consume between two and five racks, depending on data center power and cooling.

InfiniBand Compute Fabric

This section outlines the major components of a four-node DGX POD with InfiniBand compute fabric.

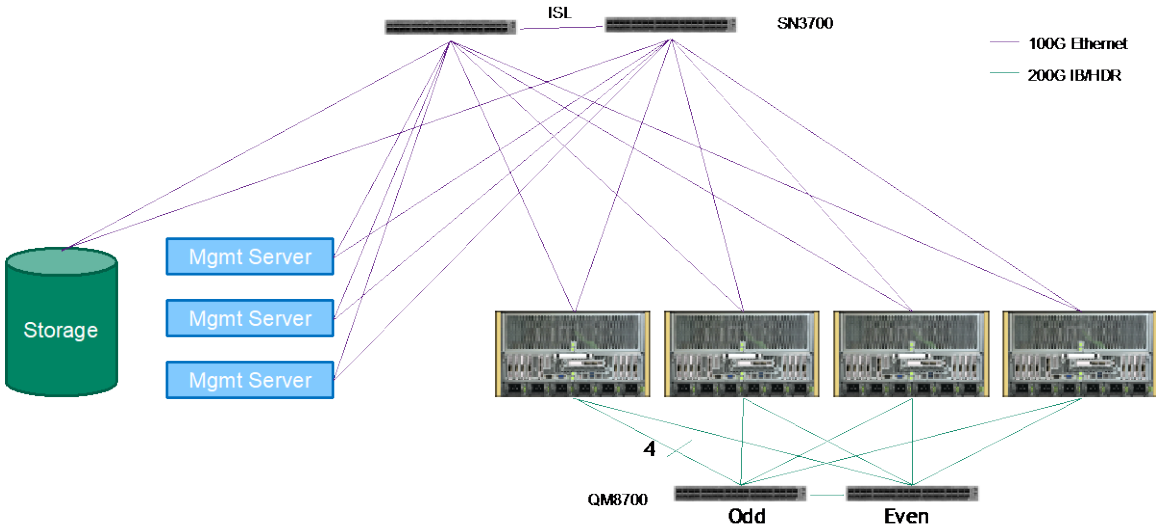

Figure 11. Logical topology for a four-node DGX-POD with InfiniBand compute fabric

Quantity	SKU	Description
4	DGXA-2530A+P2CMI00	DGX A100 P3687 System 8X 40GB GPUs
4	965-23687-0010-000	NIC dual-port 200Gb IB or Ethernet NVIDIA Mellanox CX-6 VPI
3	MGMT-SERVER	Up to customer

Table 1. Compute

Quantity	SKU	Description
2	MQM8700-HS2F	NVIDIA Mellanox Quantum HDR InfiniBand switch, 40 QSFP56 ports, two power supplies¬†(AC), x86 dual-core, standard depth, P2C airflow, rail kit
32	MFS1S00-H010E	NVIDIA Mellanox active fiber cable, IB HDR, up to 200Gb/s, QSFP56, LSZH, black pulltab, 10m
8	MCP1650-H00AE30	NVIDIA Mellanox passive copper cable, IB HDR, up to 200Gb/s, QSFP56, LSZH, 0.5m, black pulltab, 30AWG

Table 2. Compute networking

Quantity	SKU	Description
1	STOR-ENDPOINT	Pure Storage FlashBlade

Table 3. Storage

Quantity	SKU	Description
2	MSN3700-CS2F	NVIDIA Mellanox Spectrum-2 based 100GbE 1U open Ethernet switch with Onyx, 32 QSFP28 ports, two power supplies (AC), x86 CPU, standard depth, P2C airflow, Rail Kit
1	4610-54T-O-AC-B	1 GE RJ45 1u open Ethernet switch with ONIE, 48-port GE RJ45 port + 4x10G SFP+, two power supplies (AC), integrated ARM A9 CPU, short depth, P2C airflow. Rail kit must be purchased separately
8	MFA1A00-C010	NVIDIA Mellanox active fiber cable, ETH 100GbE, 100 Gb/s, QSFP, LSZH, 10m
2	MCP1600-C00AE30N	NVIDIA Mellanox passive copper cable, ETH 100GbE, 100Gb/s, QSFP28, 0.5m, black, 30AWG, CA-N
6	MCP1600-C002E30N	NVIDIA Mellanox passive copper cable, ETH 100GbE, 100Gb/s, QSFP28, 2m, black, 30AWG, CA-N
16	CAT5-CABLE	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 10m (from utility rack to GPU racks)
3	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)
2	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)
2	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)
1	RKIT-85-4610	Rackmount kit for AS4610-54T
1	UPGR-CUM-1G	UPGR-CUM-1G - Cumulus Linux software license for Edgecore AS4610 1G management Ethernet switch. Requires purchase SUP-4610-xxx
8	MCP1600-C002E30N	NVIDIA Mellanox passive copper cable, ETH 100GbE, 100Gb/s, QSFP28, 2m, black, 30AWG, CA-N
1	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)

Table 4. Storage networking

Quantity	SKU	Description
4	718-A10000+P2CMI36	DGX A100 system 8X 40GB GPUs support, three years
3	SUP-MGMT-3S-SERVER	3-year support for management nodes - ESTIMATE
1	SUP-4610-54T-3SNBD	NVIDIA Mellanox Technical Support and Warranty - Silver 3-year with NBD on-site support for 4610-54T series switch

2	SUP-QM8700-3SNBD	NVIDIA Mellanox Technical Support and Warranty - Silver 3-year with NBD on-site support for QM8700 series switch
2	SUP-SN3700-3SNBD	NVIDIA Mellanox Technical Support and Warranty - Silver 3-year with NBD on-site support for SN3700 series switch
1	SUP-STOR-3S-ENDPOINT	Support for FlashBlade

Table 5. Support

RoCE Compute Fabric

This section lists the major components of a four-node DGX POD with dedicated Ethernet (RoCE) compute fabric (Figure 12).

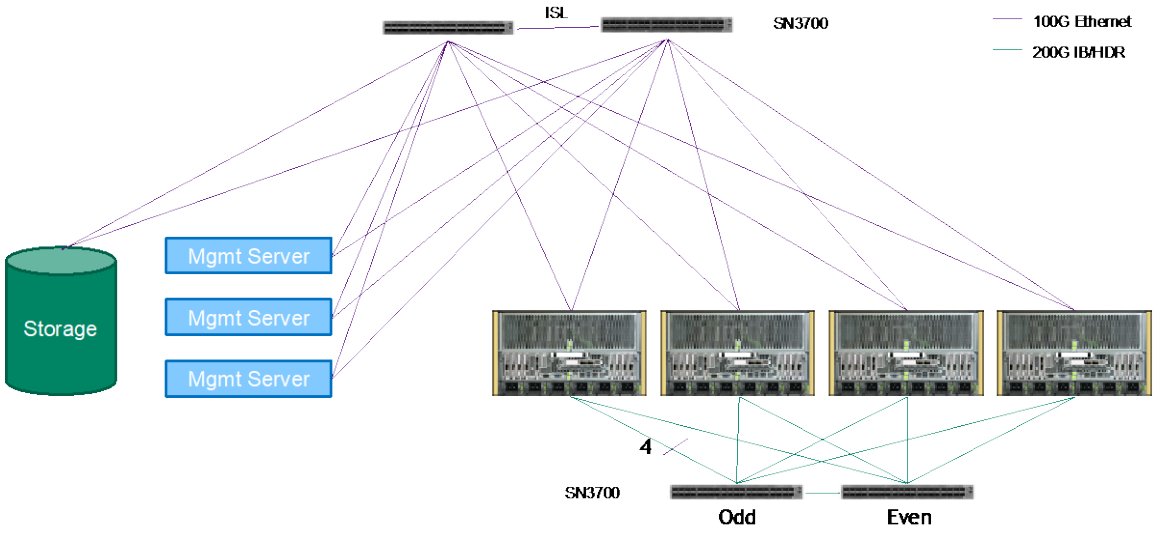

Figure 12. Logical topology for a four-node DGX-POD with dedicated RoCE compute fabric

Quantity	SKU	Description
4	DGXA-2530A+P2CMI00	DGX A100 P3687 system 8X 40GB GPUs
4	965-23687-0010-000	NIC dual-port 200Gb IB or Ethernet NVIDIA Mellanox CX-6 VPI
3	MGMT-SERVER	Up to customer

Table 6. Compute

Quantity	SKU	Description
2	MSN3700-VS2F	NVIDIA Mellanox Spectrum-2 based 200GbE 1U open Ethernet switch with Onyx, 32 QSFP56 ports, two power supplies¬†(AC), x86 CPU, standard depth, P2C airflow, rail kit
32	MFS1S00-V010E	NVIDIA Mellanox active fiber cable, 200GbE, 200Gb/s, QSFP56, LSZH, black pulltab, 10m

Table 7. Compute networking

Quantity	SKU	Description
1	STOR-ENDPOINT	Pure Storage FlashBlade
32	MFS1S00-V010E	NVIDIA Mellanox active fiber cable, 200GbE, 200Gb/s, QSFP56, LSZH, black pulltab, 10m

Table 8. Storage

Quantity	SKU	Description
2	MSN3700-CS2F	NVIDIA Mellanox Spectrum-2 based 100GbE 1U open Ethernet switch with Onyx, 32 QSFP28 ports, two power supplies (AC), x86 CPU, standard depth, P2C airflow, Rail Kit
8	MFA1A00-C010	NVIDIA Mellanox active fiber cable, ETH 100GbE, 100Gb/s, QSFP, LSZH, 10m
2	MCP1600-C00AE30N	NVIDIA Mellanox passive copper cable, ETH 100GbE, 100Gb/s, QSFP28, 0.5m, black, 30AWG, CA-N
8	MCP1600-C002E30N	NVIDIA Mellanox passive copper cable, ETH 100GbE, 100Gb/s, QSFP28, 2m, black, 30AWG, CA-N
1	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)
1	4610-54T-O-AC-B	1GE RJ45 1U open Ethernet switch with ONIE, 48-port GE RJ45 port + 4x10G SFP+, two power supplies (AC), integrated ARM A9 CPU, short depth, P2C airflow. Rail kit must be purchased separately
1	UPGR-CUM-1G	UPGR-CUM-1G - Cumulus Linux software license for Edgecore AS4610 1G management Ethernet switch. Requires purchase SUP-4610-xxx
1	RKIT-85-4610	Rackmount kit for AS4610-54T
16	CAT5-CABLE	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 10m (from utility rack to GPU racks)
3	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)
2	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)
2	CAT5-CABLE-SHORT	Cat5e snagless unshielded (UTP) network patch Ethernet cable-blue, 3m (within utility rack)
6	MCP1600-C002E30N	NVIDIA Mellanox passive copper cable, ETH 100GbE, 100Gb/s, QSFP28, 2m, black, 30AWG, CA-N

Table 9. Storage and Management Network

Quantity	SKU	Description
8	718-A10000+P2CMI36	DGX A100 system 8X 40GB GPUs support, 3 Years
3	SUP-MGMT-3S-SERVER	3-year support for management nodes
1	SUP-4610-54T-3SNBD	NVIDIA Mellanox technical support and warranty - Silver 3-year with NBD on-Site support for 4610-54T series switch
1	SUP-STOR-3S-ENDPOINT	Support for FlashBlade
2	SUP-SN3700-3SNBD	NVIDIA Mellanox technical support and warranty - Silver 3-year with NBD on-site support for SN3700 series switch
2	SUP-SN3700-3SNBD	NVIDIA Mellanox technical support and warranty - Silver 3-year with NBD on-site support for SN3700 series switch

Table 10. Support

Component Configurations

For specific questions around configuring DGXs, please reach out to NVIDIA Support.

Additional Resources

- Learn more with the Pure Storage AIRI Reference Architecture with DGX A100.
- Read about NVIDIA AIRI.
- Discover Pure Storage AIRI.

Pure Storage, Inc.
650 Castro Street, #400
Mountain View, CA 94041