



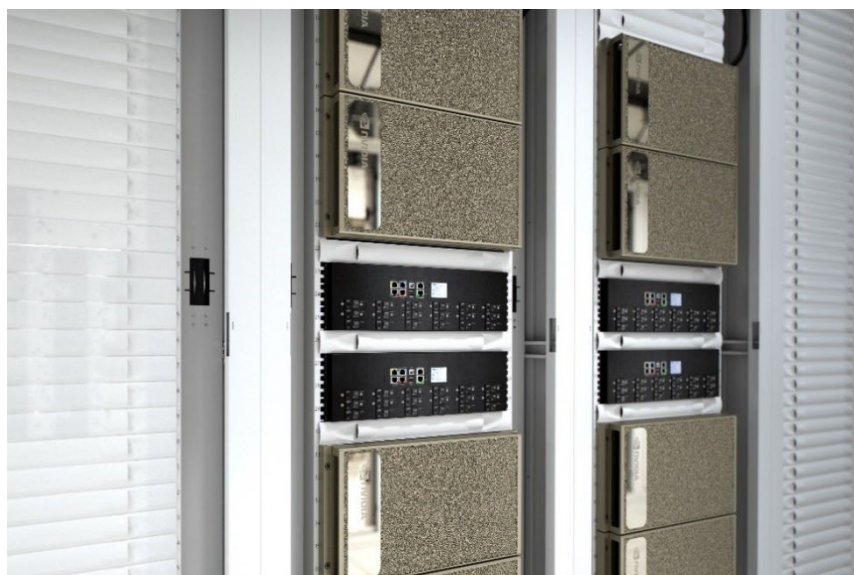
# NVIDIA DGX BasePOD: The Infrastructure Foundation for Enterprise AI

## Reference Architecture

Featuring NVIDIA DGX B200, H200 and H100 Systems

# Abstract

The number of use cases for AI within an enterprise, including examples such as language modeling, cybersecurity, autonomous systems, and healthcare, continues to expand quickly. Not only have the number of use cases grown, but model complexity and data sources also are growing. The system required to process, train, and serve these next generation models must also grow. Training models commonly use dozens of GPUs for evaluating and optimizing different model configurations and parameters. Training data must be readily accessible to all the GPUs for these new workloads. In addition, organizations have many AI researchers that must train numerous models simultaneously. Enterprises need the flexibility for multiple developers and researchers to share these resources as they refine and bring their AI stack to production.



[NVIDIA DGX BasePOD™](#) provides the underlying infrastructure and software to accelerate deployment and execution of these AI workloads. By building upon the success of NVIDIA DGX systems, DGX BasePOD is a prescriptive AI infrastructure for enterprises, eliminating the design challenges, lengthy deployment cycle, and management complexity traditionally associated with scaling AI infrastructure. Powered by [NVIDIA Base Command™](#), DGX BasePOD provides the essential foundation for AI development optimized for enterprise.

This reference architecture discusses the key components of DGX BasePOD and provides a prescriptive design for DGX BasePOD solutions.

# Contents

Chapter 1.	DGX BasePOD Overview.....	1
	NVIDIA Networking.....	2
1.1	Partner Storage Appliance.....	2
1.2	NVIDIA Software.....	2
1.2.1	NVIDIA Base Command.....	2
1.2.2	NVIDIA NGC.....	4
1.2.3	NVIDIA AI Enterprise.....	4
Chapter 2.	Core Components.....	5
2.1	NVIDIA DGX Systems.....	5
2.1.1	NVIDIA DGX B200 System.....	5
2.1.2	NVIDIA DGX H200 and H100 Systems.....	6
2.2	NVIDIA Networking Adapters.....	8
2.2.1	NVIDIA Networking Adapters.....	8
2.3	NVIDIA Networking Switches.....	9
2.3.1	NVIDIA QM9700 Switch.....	9
2.3.2	NVIDIA SN4600C switch.....	9
2.3.3	NVIDIA SN2201 Switch.....	10
2.4	Control Plane.....	10
Chapter 3.	Reference Architectures.....	11
3.1	DGX BasePOD with NDR400 Compute Fabric.....	11
3.1.1	System Architecture.....	12
3.1.2	Switches and Cables.....	13
Chapter 4.	Summary.....	14

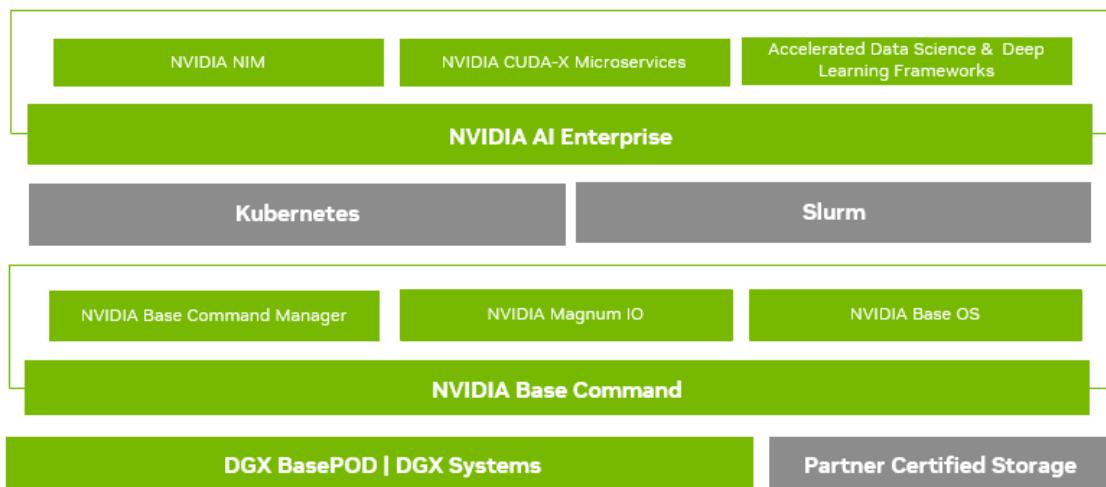
---

# Chapter 1. DGX BasePOD Overview

DGX BasePOD is an integrated solution consisting of NVIDIA hardware and software components, MLOps solutions, and third-party storage. Leveraging best practices of scale-out system design with NVIDIA products and validated partner solutions, customers can implement an efficient and manageable platform for AI development. The designs in this DGX BasePOD reference architecture (RA) support developer needs, simplify IT manageability, and infrastructure scaling from two nodes to dozens with certified storage platforms from an industry-leading ecosystem. Optional MLOps solutions can be integrated with DGX BasePOD to enable a full stack solution to shorten AI model development cycles and speed the ROI of AI initiatives.

Figure 1 highlights the various components of NVIDIA DGX BasePOD. Each of these layers is an integration point that users typically would have to build and tune before an application could be deployed. The designs in the RA simplify system deployment and optimization using a validated prescriptive architecture.

**Figure 1. Layers of integration for DGX BasePOD**



# NVIDIA Networking

InfiniBand and Ethernet technologies enable networking functionality in DGX BasePOD. Proper networking is critical to ensuring that DGX BasePOD does not have any bottlenecks or suffer performance degradation for AI workloads. For more information on the products and technologies that enable this, refer to [NVIDIA Networking](#).

## 1.1 Partner Storage Appliance

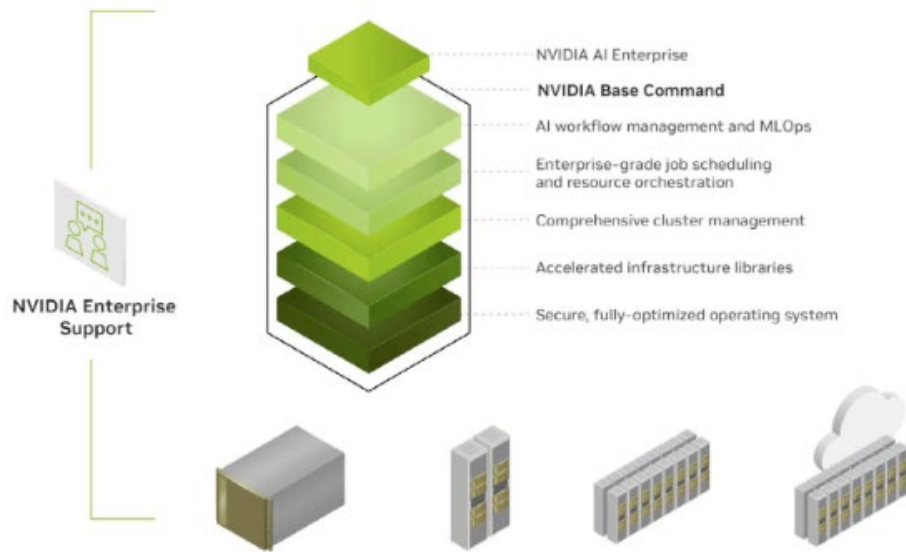
DGX BasePOD is built on a proven storage technology ecosystem. As NVIDIA validated storage partners introduce new storage technologies into the marketplace, they will qualify these new offerings with DGX BasePOD to ensure design compatibility and expected performance for known workloads. Every storage partner has performed rigorous testing to ensure that applications receive the highest performance and throughput when deployed with DGX BasePOD.

## 1.2 NVIDIA Software

### 1.2.1 NVIDIA Base Command

[NVIDIA Base Command](#) (Figure 2) powers every DGX BasePOD, enabling organizations to leverage the best of NVIDIA software innovation. Enterprises can unleash the full potential of their investment with a proven platform that includes enterprise-grade orchestration and cluster management, libraries that accelerate compute, storage and network infrastructure, and an operating system (OS) optimized for AI workloads.

Figure 2. NVIDIA Base Command features and capabilities with DGX BasePOD



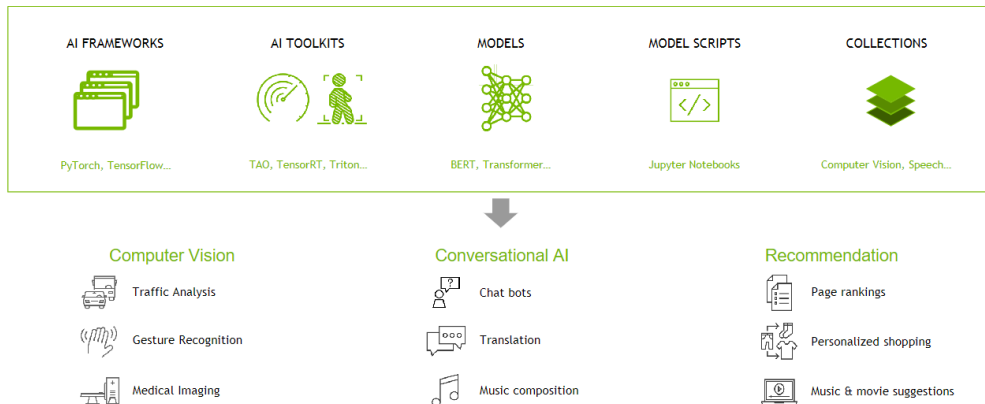
DGX BasePOD hardware is further optimized with acceleration libraries that know how to maximize the performance of AI workload across a GPU, the DGX system and an entire DGX cluster, speeding data access, movement, and management from system I/O to storage to network fabric.

Base Command provides integrated cluster management from installation and provisioning to ongoing monitoring of systems—from one to hundreds of DGX systems. Base Command also supports multiple methods for workflow management. Either Slurm or Kubernetes can be used to allow for optimal scheduling and management of system resources within a multi-user environment.

## 1.2.2 NVIDIA NGC

NVIDIA NGC™ (Figure 3) provides software to meet the needs of data scientists, developers, and researchers with various levels of AI expertise.

Figure 3. NGC catalog overview



Software hosted on NGC undergoes scans against an aggregated set of common vulnerabilities and exposures (CVEs), crypto, and private keys. It is tested and designed to scale to multiple GPUs and in many cases, to multi-node, ensuring users maximize their investment in DGX systems.

## 1.2.3 NVIDIA AI Enterprise

NVIDIA AI Enterprise is the end-to-end software platform that brings generative AI into reach for every enterprise, providing the fastest and most efficient runtime for generative AI foundation models developed with the NVIDIA DGX platform. With production-grade security, stability, and manageability, it streamlines the development of generative AI solutions. NVIDIA AI Enterprise is included with DGX SuperPOD for enterprise developers to access pretrained models, optimized frameworks, microservices, accelerated libraries, and enterprise support.

---

# Chapter 2. Core Components

The compute nodes with HCAs and switch resources form the foundation of the DGX BasePOD. The specific components used in the DGX BasePOD Reference Architectures are described in this section.

## 2.1 NVIDIA DGX Systems

NVIDIA DGX BasePOD configurations use DGX B200, DGX H200, and H100 systems. The systems are described in the following sections.

### 2.1.1 NVIDIA DGX B200 System

The [NVIDIA DGX B200 system](#) (Figure 4) offers unprecedented compute density, performance, and flexibility.

**Figure 4. DGX B200 system**



Key specifications of the DGX B200 system are:

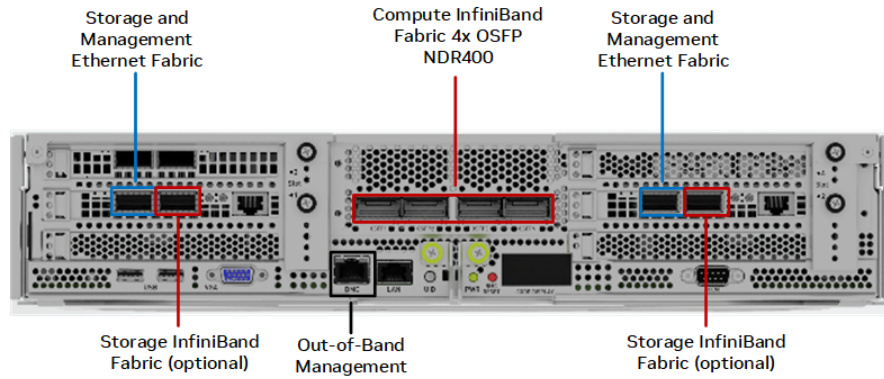
- > Built with eight NVIDIA B200
- > 1.4TB of GPU memory space
- > 4x OSFP ports serving 8x single-port NVIDIA ConnectX-7 VPI



- > Up to 400Gb/s InfiniBand/Ethernet 2x dual-port QSFP112 NVIDIA BlueField-3 DPU
- > Dual 5th generation Intel® Xeon® Scalable Processors

Rear ports of the DGX B200 CPU tray are shown in .

**Figure 5. DGX B200 CPU tray rear ports**



Four of the ConnectX-7 OSFP are used for the compute fabric. Each pair of dual-port BlueField-3 HCAs (NIC mode) provide parallel pathways to the storage and management fabrics. The out-of-band (OOB) port is used for BMC access.

## 2.1.2 NVIDIA DGX H200 and H100 Systems

The [DGX H200 system](#) (Figure 6) is the latest DGX system and the AI powerhouse that is accelerated by the groundbreaking performance of the [NVIDIA Hopper GPU](#).

**Figure 6. DGX H200 and H100 system**



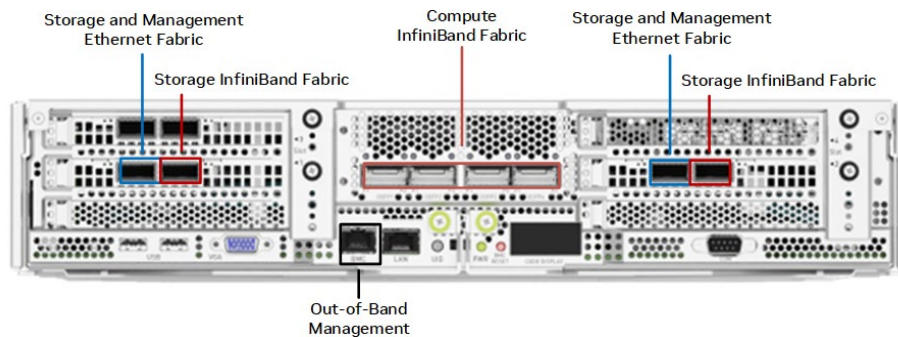
Key specifications of the DGX H200 and H100 system are:

- > Eight NVIDIA Hopper GPUs.
- > 1,128 GB total GPU memory for H200.
- > 640 GB total GPU memory for H100.
- > Four NVIDIA NVSwitch™ chips.
- > Dual Intel® Xeon® Platinum 8480C processors, 112 cores total, 2.00 GHz (Base), 3.80 GHz (Max Boost) with PCIe 5.0 support.

- > 2 TB of DDR5 system memory.
- > Four OSFP ports serving eight single-port NVIDIA ConnectX-7 VPI, 2x dual-port QSFP112 NVIDIA ConnectX-7 VPI, up to 400 Gb/s InfiniBand/Ethernet.
- > 10Gb/s onboard NIC with RJ45, 100 Gb/s Ethernet NIC, BMC with RJ45.
- > Two 1.92 TB M.2 NVMe drives for DGX OS, eight 3.84 TB U.2 NVMe drives for storage/cache.

The rear ports of the DGX H200 and H100 CPU tray are shown in .

**Figure 7. DGX H200 and H100 CPU tray rear ports**



Four of the OSFP ports serve eight ConnectX-7 HCAs for the compute fabric. Each pair of dual-port ConnectX-7 HCAs provide parallel pathways to the storage and management fabrics. The OOB port is used for BMC access.

## 2.2 NVIDIA Networking Adapters

NVIDIA DGX B200 and DGX H200 and H100 systems are equipped with NVIDIA® ConnectX®-7 network adapters. The DGX B200 has both ConnectX-7 and NVIDIA BlueField-3 network adapters. The network adapters are described in this section.



**Note:** Going forward, HCA will refer to network adapter cards configured for InfiniBand and NIC for those configured for Ethernet.

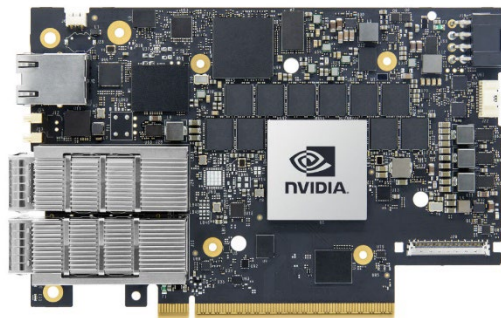
### 2.2.1 NVIDIA Networking Adapters

The ConnectX-7 VPI Adapter (Figure ) is the latest ConnectX Adapter line. It can provide 25/50/100/200/400G of throughput. NVIDIA DGX system use ConnectX-7 and BlueField-3 (NIC Mode) HCAs to provide flexibility in DGX BasePOD deployments with NDR400 and RoCE. Specifications are available [here](#).

Figure 8. NVIDIA ConnectX-7 HCA



Figure 9. NVIDIA BlueField-3 HCA



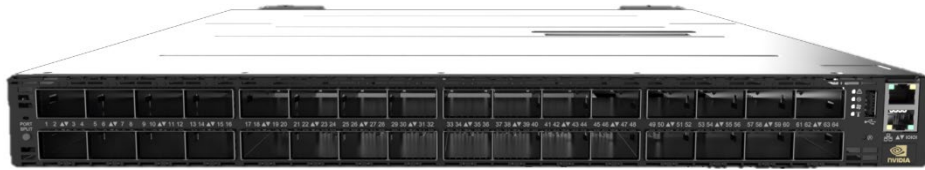
## 2.3 NVIDIA Networking Switches

DGX BasePOD configurations can be equipped with four types of NVIDIA networking switches. The switches are described in this section, with how the switches are being deployed in the Reference Architectures section.

### 2.3.1 NVIDIA QM9700 Switch

NVIDIA QM9700 switches (Figure 10) with NDR InfiniBand connectivity power the compute fabric in NDR BasePOD configurations. ConnectX-7 single-port adapters are used for the InfiniBand compute fabric. Each NVIDIA DGX system has dual connections to each QM9700 switch, providing multiple high-bandwidth, low-latency paths between the systems.

Figure 10. NVIDIA QM9700 switch



### 2.3.2 NVIDIA SN4600C switch

NVIDIA SN4600C Switches (Figure 11) offer 128 total ports (64 per switch) to provide redundant connectivity for in-band management of the DGX BasePOD. The NVIDIA SN4600C switch can provide for speeds between 1 GbE and 100 GbE.

For storage appliances connected over Ethernet, the NVIDIA SN4600C switches are also used. The ports on the NVIDIA DGX dual-port network adapters are used for both in-band management and storage connectivity.

Figure 11. NVIDIA SN4600C switch



## 2.3.3 NVIDIA SN2201 Switch

NVIDIA SN2201 switches (Figure 12) offer 48 ports to provide connectivity for OOB management. OOB management provides consolidated management connectivity for all components in BasePOD.

Figure 12. NVIDIA SN2201 switch



## 2.4 Control Plane

The minimum requirements for each server in the control plane are:

- > 2 × Intel x86 Xeon Gold or better
- > 512 GB memory
- > 1 × 6.4 TB NVMe for storage
- > 2 × 480 GB M.2 RAID for OS
- > 4 × 200 Gbps network
- > 2 × 100 GbE network

---

# Chapter 3. Reference Architectures

DGX BasePOD is a flexible solution that offers multiple prescriptive architectures. These architectures are adaptable to support the evolving demands of AI workloads.

## 3.1 DGX BasePOD with NDR400 Compute Fabric

DGX BasePOD is a flexible solution that offers multiple prescriptive architectures. These architectures are adaptable to support the evolving demands of AI workloads.

The components of the DGX BasePOD are described in Table 1.

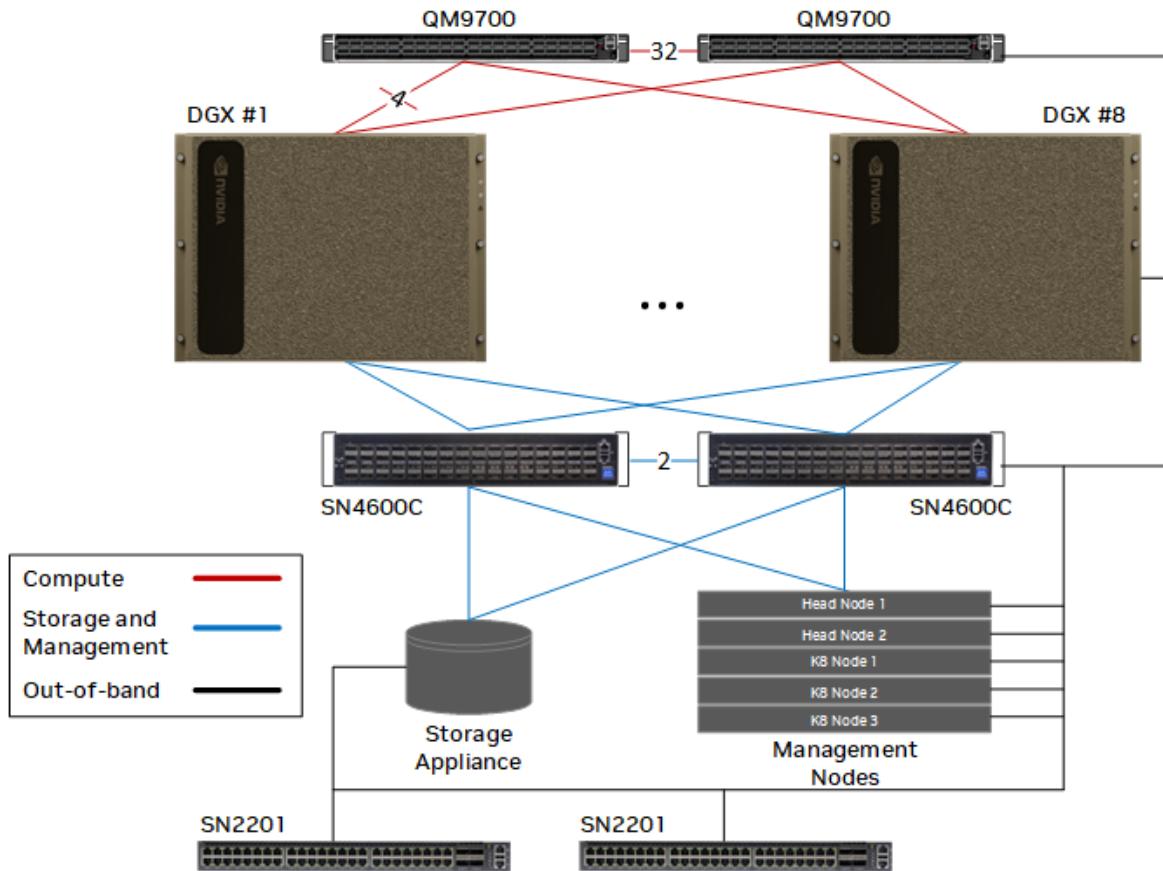
**Table 1. DGX BasePOD components**

Component	Technology
Compute nodes (2-8)	NVIDIA DGX B200 system with eight 180 GB B200 GPUs, NDR400 InfiniBand networking, and two NVIDIA BlueField-3 DPUs Or NVIDIA DGX H100 system with eight 80 GB H100 GPUs, NDR400 InfiniBand networking, and two NVIDIA ConnectX-7 NICs Or NVIDIA DGX H200 system with eight 141 GB H100 GPUs, NDR400 InfiniBand networking, and two NVIDIA ConnectX-7 NICs
Compute fabric	NVIDIA Quantum QM9700 NDR400 Gbps InfiniBand switch
Management and storage fabric	NVIDIA SN4600C switches
OOB management fabric	NVIDIA SN2201 switches
Control plane	See Section 2.4

### 3.1.1 System Architecture

Figure 12 depicts the architecture for the DGX BasePOD for up to eight DGX nodes with NDR InfiniBand. BasePOD with DGX B200 and H200 and H100 systems use eight compute connections from each node running at NDR400. The complete architecture has three networks, an InfiniBand-based compute network, an Ethernet fabric for system management and storage, and an OOB management network.

Figure 13. DGX BasePOD with up to eight systems with NDR400



Included in the reference architecture are five dual-socket x86 servers for system management. Two nodes are used as the head nodes for Base Command Manager. The three additional nodes provide the platform to house specific services for the deployment. This could be login nodes for a Slurm-based deployment, or Kubernetes for MLOps-based partner solutions. Any OEM server that meets the minimum requirements for each node described in Table 1 can be used. All management servers are configured in a high-availability (HA) pair (or triple), a failure of a single node won't lead to the outage of the BasePOD service.

## 3.1.2 Switches and Cables

Table 2 shows the number of cables and switches required for various deployments of DGX BasePOD. These designs are built with active optical cables or direct attached copper. Alternatively, DGX BasePOD may be deployed with transceivers and fiber cables.

**Table 2. Switches and cables**

Components	Part Number <sup>1</sup>	# of DGX	
		4	8
NVIDIA Quantum QM9700 switch	920-9B210-00FN-0M0	2	2
NDR Fiber Cables, 400 Gbps, DGX to IB Switches	980-9I570-00N030	16	32
System 2x400G OSFP Flat-top Multimode Transceivers on DGX Systems	980-9I51A-00NS00	16	32
Switch 2x400G OSFP Finned- top Multimode Transceivers	980-9I510-00NS00	8	16
NDR InfiniBand DAC for Switch ISL	980-9IA0J-00N002	16	32
NVIDIA SN2201 switch with Cumulus Linux, 48 RJ45 ports, P2C	920-9N110-00F1-0C0	1	2
NVIDIA SN4600C switch with Cumulus Linux, 64 QSFP28 ports, P2C	920-9N302-00F7-0C2	2	2
1 GbE Cat 6 cables	N/A	29	45
NVIDIA active fiber cable, ETH 100GbE, 100Gb/s, QSFP, LSZH, 30m, DGX to Inband	980-9I13N-00C030	8	16
100 Gbps QSFP Passive Cable for Inband Switch ISL	980-9I54C-00V001	2	2
NVIDIA active fiber cable, ETH 100GbE, 100Gb/s, QSFP, LSZH, 10m, OOB to Inband	980-9I13N-00C010	2	4
BCM management servers	Varies	5	5
NVIDIA active fiber cable, ETH 100GbE, 100Gb/s, QSFP, LSZH, 10m, Management Servers to Inband	980-9I13N-00C010	10	10

<sup>1</sup> Part number will depend on exact cable lengths needed based on data center requirements.



---

## Chapter 4. Summary

Every enterprise wants to leverage AI to improve their products, services, and processes. But many struggle with how to operationalize AI at scale. Production-ready AI infrastructure requires blending a complex set of leading-edge hardware and software into a complete solution. This takes time and expertise to design, is difficult to deploy, and expensive to support across a multilayered technology stack from a variety of vendors.

With DGX BasePOD, NVIDIA has done the work for you—solving the complexity of designing AI infrastructure, systemizing it to power AI development and deployment, and simplifying its management. NVIDIA DGX BasePOD incorporates tested and proven design principles into an integrated AI infrastructure solution that incorporates best-of-breed NVIDIA DGX systems, NVIDIA software, NVIDIA networking, and an ecosystem of high-performance storage to enable AI innovation for the modern enterprise.

## NVIDIA Reference Architecture Notice and Disclaimers

NVIDIA Reference Architecture documentation is NVIDIA confidential information and embodies substantial intellectual property and engineering know-how. NVIDIA grants you limited permissions to use NVIDIA Reference Architecture to create NVIDIA platform-enabled data center clusters with NVIDIA technologies. You may not use the NVIDIA Reference Architecture to develop competing products or technologies or assisting a third party in such activities. Except as expressly stated in this Notice & Disclaimer statement, no other license or right is granted to you by implication, estoppel or otherwise. NVIDIA objects to and rejects any additional or different terms you may propose.

NVIDIA Reference Architecture is subject to change without notice. NVIDIA technologies may require enabled hardware, software, or service activation. Your costs and results may vary. This document is not a commitment to develop, release, or deliver any technology, code, or functionality. NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice. It is your sole responsibility to evaluate and determine the applicability of any information contained in this document and ensure that the product is suitable and fit for the application planned by you. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party or a license from NVIDIA.

NVIDIA products or services are subject to the applicable NVIDIA product or sales terms and conditions that are provided in conjunction with the relevant NVIDIA product or service. NVIDIA's provision of these resources does not expand or otherwise alter NVIDIA's applicable warranties or warranty disclaimers for the NVIDIA products or services.

You will not use the NVIDIA Reference Architecture or NVIDIA's confidential information to: (i) identify or support an assertion or potential assertion of any intellectual property rights (including patent, copyright, or trade secret); or (ii) prepare, file, amend, or prosecute any patent applications. You agree to grant NVIDIA a limited, worldwide, nonexclusive, perpetual, irrevocable, non-sublicensable, nontransferable, royalty-free, fully-paid license to any patent claim that you may file after accessing NVIDIA Reference Architecture that covers the subject matter disclosed herein.

**Warranty Disclaimer.** NVIDIA Reference Architecture materials disclosed are provided "AS IS". To the maximum extent permitted by applicable law, NVIDIA disclaims all warranties and representations of any kind, whether express, implied, or statutory, relating to or arising under this agreement, including, without limitation, the warranties of title, noninfringement, merchantability, fitness for a particular purpose, usage of trade, and course of dealing.

**Limitation of Liability.** To the maximum extent permitted by applicable law, in no event will NVIDIA be liable for any (i) indirect, punitive, special, incidental, or consequential damages, or (ii) damages for the (a) cost of procuring substitute goods or (b) loss of profits, revenues, use, data, or goodwill arising out of or related to this document, whether based on breach of contract, tort (including negligence), strict liability, or otherwise, and even if NVIDIA has been advised of the possibility of such damages and even if a your remedies fail their essential purpose. Additionally, to the maximum extent permitted by applicable law, NVIDIA's total cumulative aggregate liability for any and all liabilities, obligations, or claims arising out of or related to this document will not exceed one-hundred U.S. dollars (US\$100).

NVIDIA, NVIDIA HGX, NVIDIA DGX SuperPOD, NVIDIA DGX Cloud, NVIDIA ConnectX, NVIDIA BlueField, NVIDIA Spectrum, NVIDIA NVLink, NVIDIA Base Command, NVIDIA CUDA, NVIDIA Magnum IO, NVIDIA NetQ, and the NVIDIA logo are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

© 2024 NVIDIA Corporation & Affiliates. All rights reserved.

