NVIDIA DGX GB200 User Guide

NVIDIA Corporation

Aug 29, 2025

Overview

The *NVIDIA DGX GB200 System User Guide* is also available as a PDF.

This document contains detailed information about the hardware and software stack of an NVIDIA DGX GB200 system. In the context of this guide, the system is defined to be a rack scale solution for GPUs that are connected by NVIDIA NVLink®.

This rack scale architecture enables the following:

- Support for higher GPU and module power.
- Improved serviceability with modular building blocks.
- Support for a large L1 NVLink domain size.
- Scalable and non-scalable solutions in a common infrastructure.

Chapter 1. Hardware

The DGX GB200 is a rack scale solution for graphics processing units (GPUs) connected by an NVLink through the NVLink passive copper cable cartridge backplane. The complete DGX GB200 rack scale solution comprises compute trays with one or two compute boards, NVLink switch trays, an NVLink passive copper cable backplane, power shelves, a bus bar, and liquid cooling manifolds.

1.1. Rack Configuration

The DGX GB200 rack scale architecture is an NVIDIA rack with 72-GPU NVL domains. Each 72-GPU rack contains:

- 18x 1RU compute trays, each with 2 Grace CPUs and 4 Blackwell GPUs.
- 9x 1RU NVLink switch trays.
- 2x TOR Switches for management.
- Power shelves for supplying power to all trays and switches.

The rear side of the system provides access to the cable management system, the inlets and outlets to the liquid cooling manifolds and the manifolds themselves, the cable cartridges, and the power bus bar.

1.2. Compute Trays

The DGX GB200 compute tray is an enclosure that holds the compute boards and peripheral accessory boards and runs an OS image. The compute trays are cooled by liquid that runs up and down the rack through manifolds, then through the cold plates that are attached to the CPUs and the GPUs in the tray. The rest of the components like networking and storage devices are air cooled, which is pushed through the system by the fans. The compute trays in a rack are interconnected via NVLink through the connectors at the back of the tray, which enables communication with the other compute trays via the NVSwitch trays.

The following table provides information about the DGX GB200 compute tray configuration.

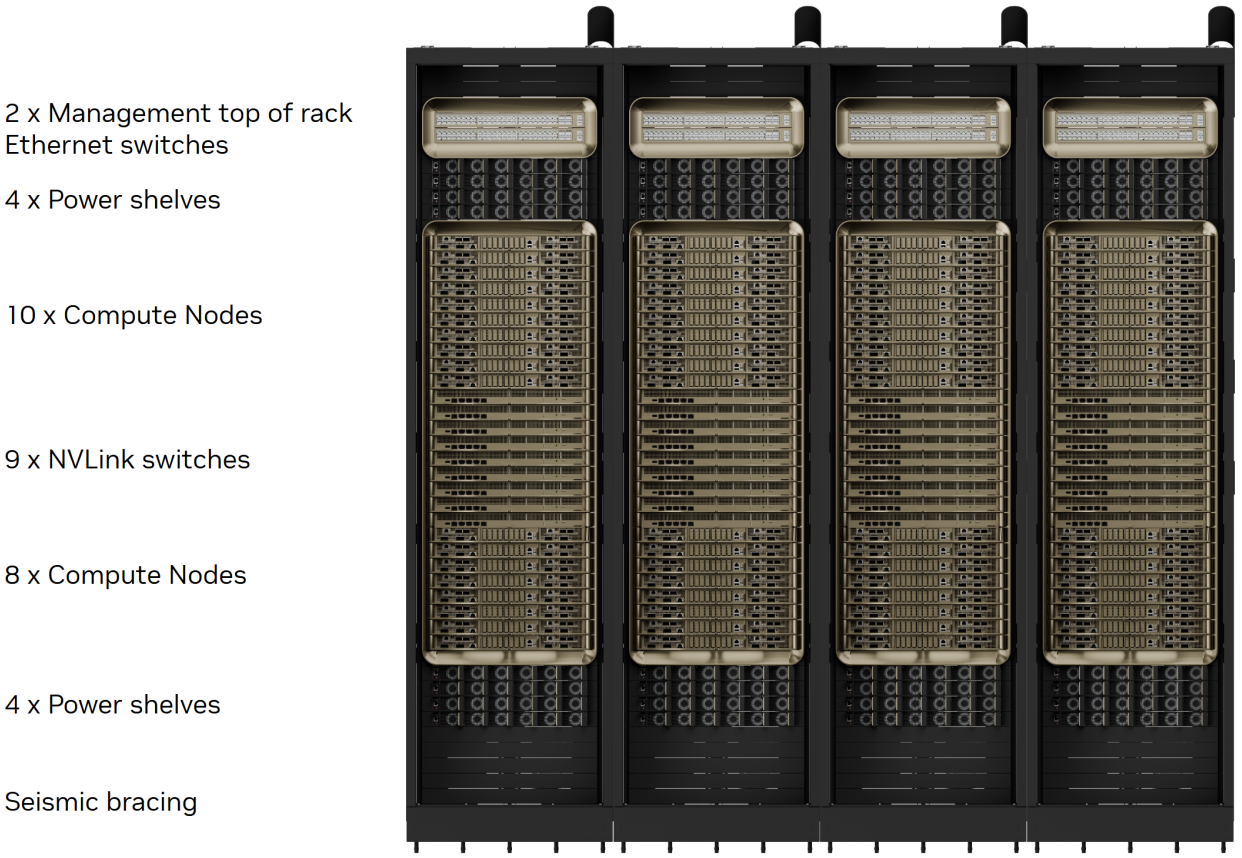


Fig. 1: DGX GB200 Rack Configuration - Front view of four 72-GPU racks

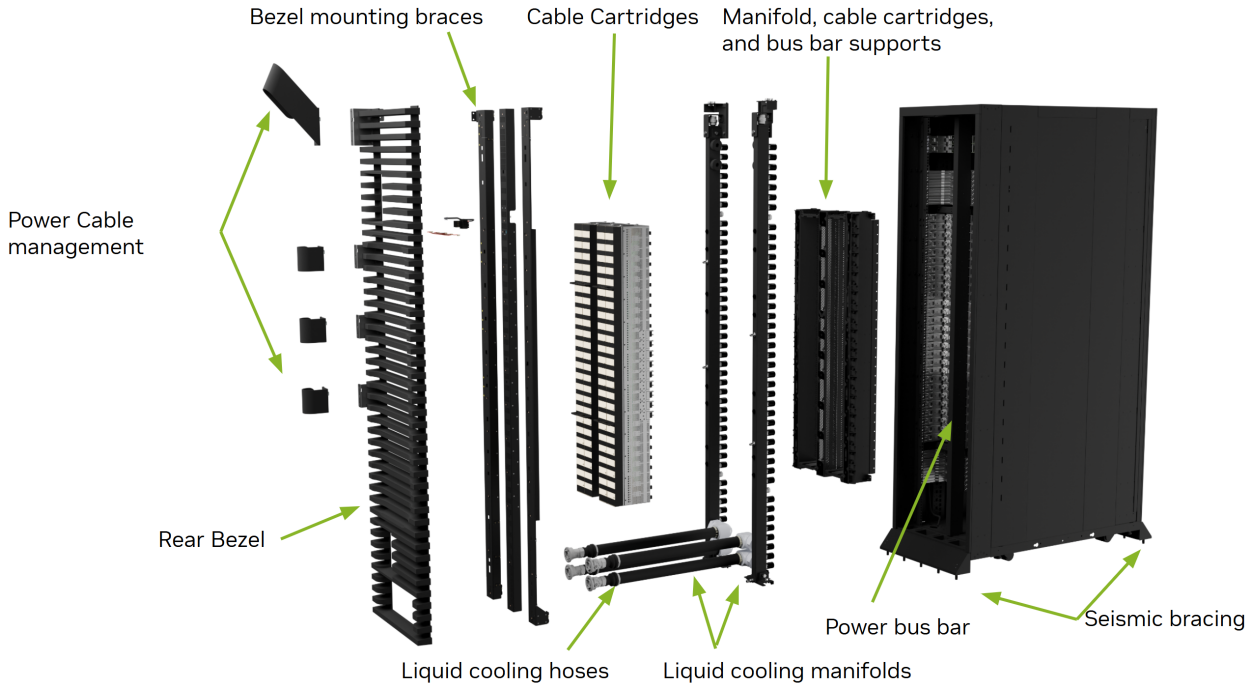


Fig. 2: DGX GB200 Rack Configuration - Rear view

Fig. 3: DGX GB200 Compute Tray

Component	Function	
Compute	Compute node	2x Grace™ CPUs and 4x Blackwell™ GPUs
Networking	Cluster network	4x NVIDIA ConnectX-7 single port 400G OSFP NIC
	Storage/management network	2x NVIDIA BlueField-3 DPU, dual port 400G Infiniband or Ethernet
	Out-of-band management network	1x 1GbE x RJ45 from the compute tray BMC module
		2x 1GbE x RJ45 from the BlueField-3 BMC interface
Storage	Data cache	4x 3.84TB E1.S NVMe per compute tray with software RAID 0
	Boot drive	1x 1.92TB M.2 NVMe

Out-of-Band (OOB) management of the compute tray hardware resources is provided by a combination of Board Management Controller (BMC) and Host Management Controller (HMC) microcontrollers. External access is available through standard BMC and console interfaces.

The figure below shows a top view of the DGX GB200 compute tray and identifies the main components.

The figure below shows a front view of the DGX GB200 compute tray and identifies the main components.

1.3. NVLink Switch Trays

The NVLink switch system delivers an unprecedented 57.6 terabits per second (Tbps) of full duplex bandwidth for the fifth-generation NVLink in a 1U design.

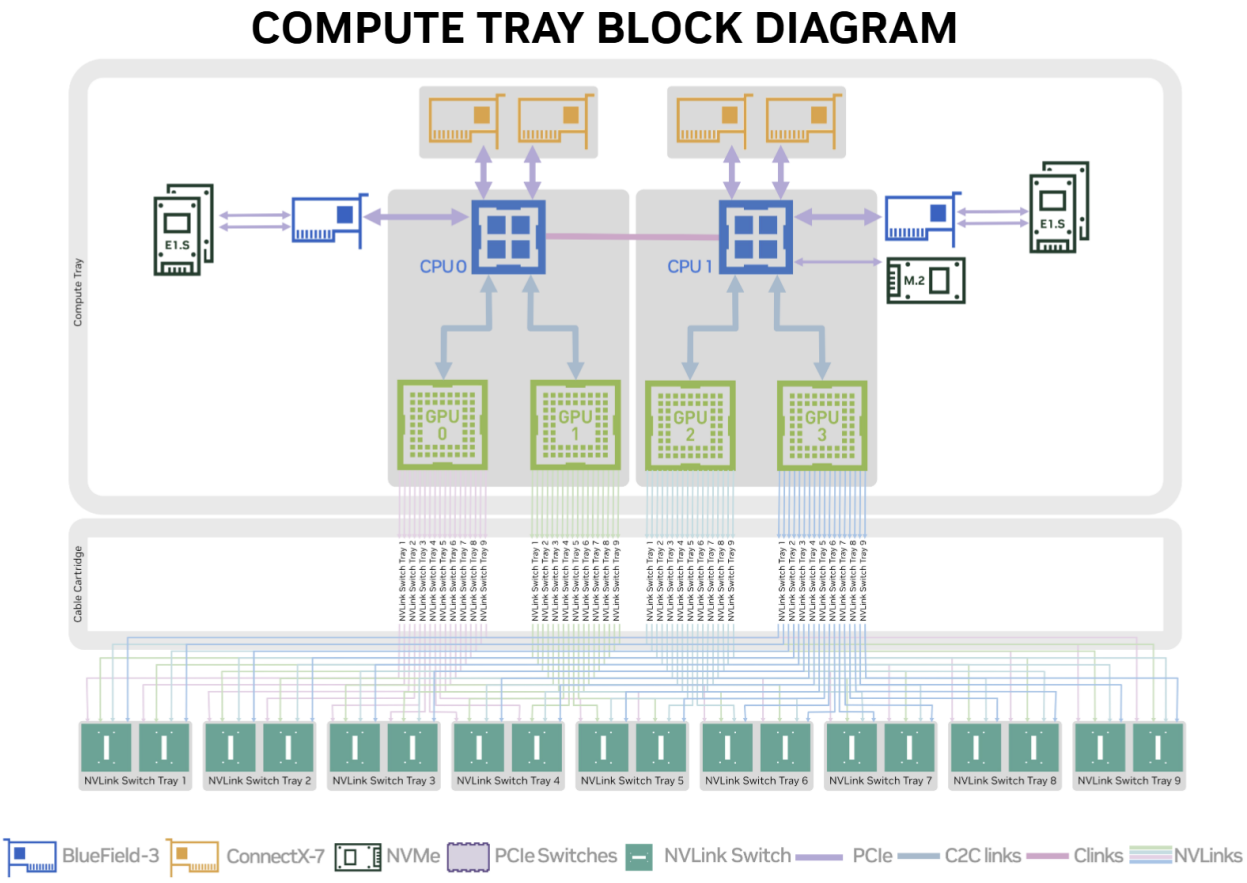


Fig. 4: DGX GB200 Compute Tray - Block diagram

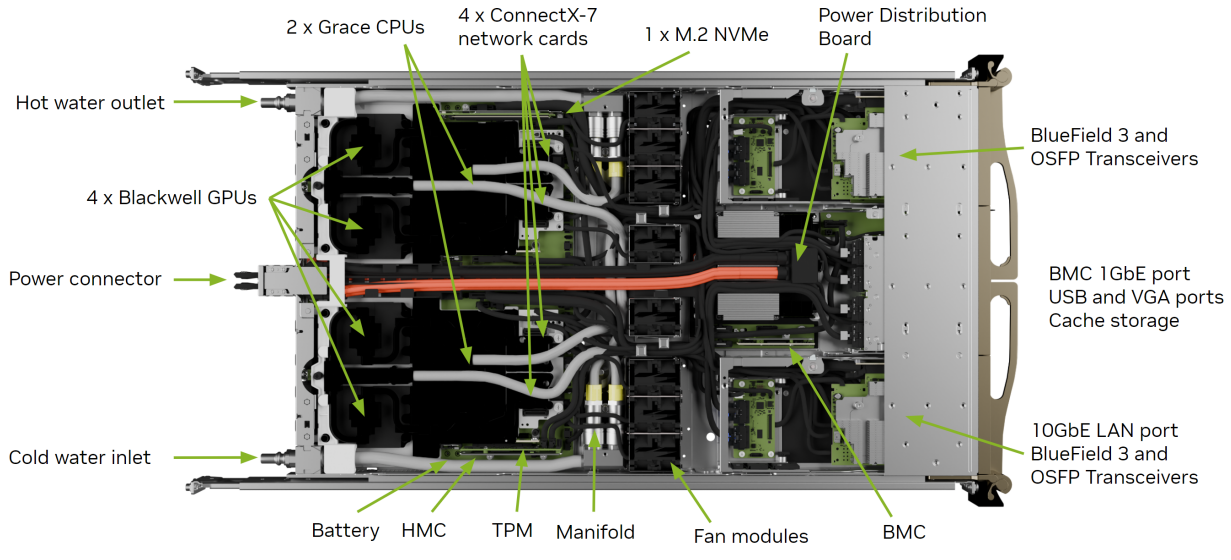


Fig. 5: DGX GB200 Compute Tray - Top view

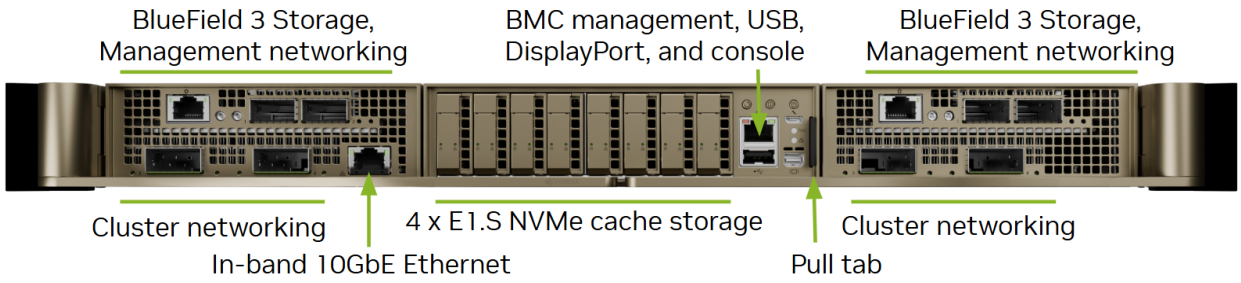


Fig. 6: DGX GB200 Compute Tray - Front view

Each switch tray has 2x NVLink switches, and each switch tray has the following components:

- 2x NVLink NVSwitches, each of which has 72 NVLink ports
- NVOS 2.0 or later for NVLink switch life cycle provisioning and telemetry
- BMC for attestation and recovery
- FM and SM as the NVLink control planes that run on the switch tray
- CPU ERoT for security

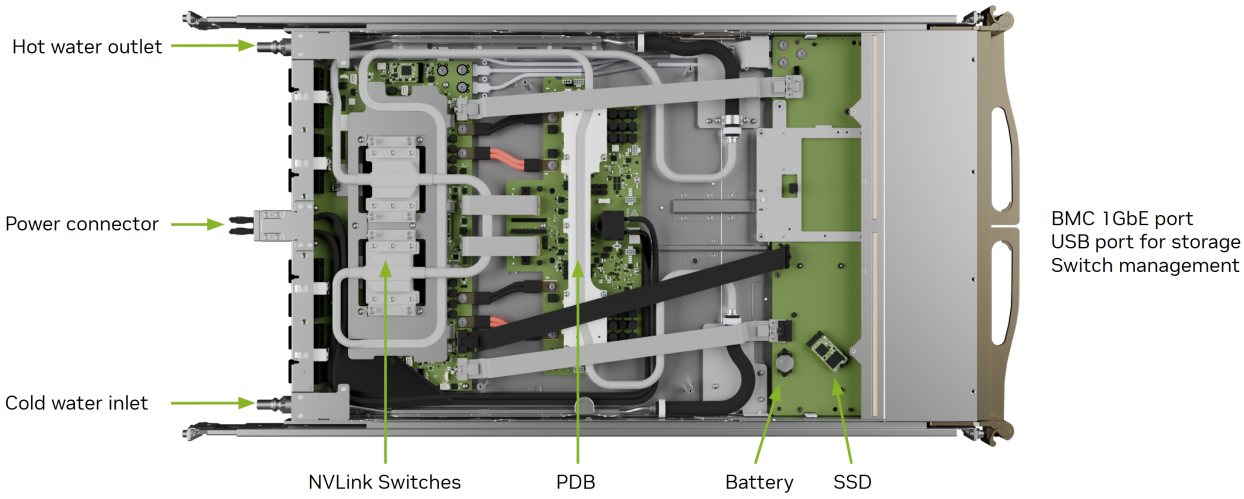


Fig. 7: NVLink Switch Tray - Top view

1.4. Power Shelves

The DGX GB200 rack uses a bus bar structure to distribute power in the rack. Power whips energize power shelves from a remote power panel. The power shelves convert AC power into nominal 50V-51V DC output and distribute it through the bus bar to the rack components. Multiple power shelves are used in the rack to supply redundant power.

The rack power consumption is approximately 120kW. The power shelf uses six air-cooled 5.5kW PSUs in eight power shelves that provide N+N redundancy and the required input power of 33kW per power shelf.

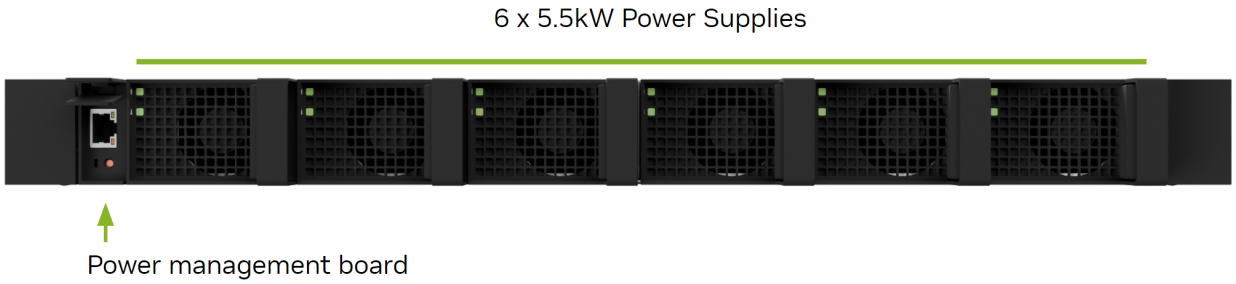


Fig. 8: Power Shelf - Top view

1.5. Top-of-Rack Ethernet Switches

Top-of-rack (TOR) switches are used to connect all BMCs in the compute trays (including BlueField-3 BMCs) and the switch trays to the management network. The following system and switch ports are connected to this switch:

- Out-of-band management BMC from the compute trays
- BlueField-3 BMCs from the compute trays

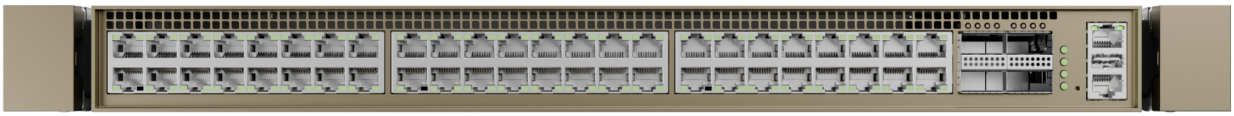


Fig. 9: DGX GB200 TOR Ethernet Switch

1.6. Console Access

Each NVLink switch provides an RJ45 interface for RS232 serial access. The default serial baud rate is 115200 baud.

1.7. Leak Detection

Leak detection in the DGX GB200 rack's liquid cooling system prevents damage to equipment, ensures system reliability and uptime, protects data integrity, and maintains safety. Early detection helps avoid costly repairs, downtime, and regulatory compliance issues, optimizes the cooling efficiency, and prevents secondary damage to infrastructure.

Chapter 2. Networking

The NVIDIA DGX GB200 system leverages a sophisticated hybrid networking approach to enable seamless, high-performance communication both within the rack and across multiple racks. This model combines the unparalleled speed and coherence of NVLink with the traditional, highly scalable InfiniBand and Ethernet technologies.

- NVLink forms the ultra-fast, memory-coherent scale-up fabric within each rack, making 72 GPUs act as one.
- InfiniBand provides the high-bandwidth, low-latency scale-out compute fabric between racks, enabling massive multi-node AI training clusters.
- Ethernet handles storage, management (in-band and out-of-band), and external connectivity, ensuring all components are integrated and manageable.

2.1. NVLink Networking

NVIDIA NVLink is a high-speed, low-latency interconnect technology designed to overcome PCIe bandwidth limitations in multi-GPU and GPU-to-CPU communication, crucial for HPC and AI platforms. Its key features include ultra-high bandwidth (e.g., 1.8 TB/s bidirectional per GPU with NVLink 5.0) and extremely low latency for direct GPU-to-GPU data transfer. NVLink enables unified memory access across GPUs, simplifying programming, and scales efficiently through NVSwitch chips, which act as non-blocking switches between servers, creating high-bandwidth fabrics.

The NVLink fabric interconnects are provided by the connectors at the back of the trays, and connect to the cable cartridges that communicate all compute and NVLink switch trays up and down the back of the system. This fabric is managed by NVIDIA Fabric Manager (or NVLSM), integrated with GPU drivers and the CUDA software stack (like NCCL), which configures the NVSwitch topology and monitors performance. Additionally, NVSwitch incorporates in-network computing (SHARP) to accelerate collective operations, further optimizing distributed AI training by performing computations within the network.

For further and detailed information on NVIDIA NMX, please refer to NVIDIA NMX Documentation

2.2. Network Overview

The DGX GB200 rack installation uses several discrete networks:

1. `externalnet` - This network provides external communication for the headnode to the enterprise network or to the Internet if allowed.

2. `internalnet` - This network provides provisioning capabilities as well as system management to the control plane nodes.
3. `dgxnet` - This network is used to provision and manage the compute trays.
4. `ipminet` - This network provides Base Command Manager and the headnode access to the out-of-band interfaces of all the components in the rack, including the compute trays, NVLink switch trays, the power shelves, the top-of-rack switches, and the control plane nodes.
5. `computenet` - This network is used for node communication across racks (east/west traffic).
6. `storagenet` - This network provides each node access to storage.
7. `failovernet` - This network is only used by the headnodes when configured in redundant failover configuration.

The following images identify the network ports used to communicate to these networks on the compute and switch trays. Note that the number corresponds to the network they are connected to in reference to the numbered list above.

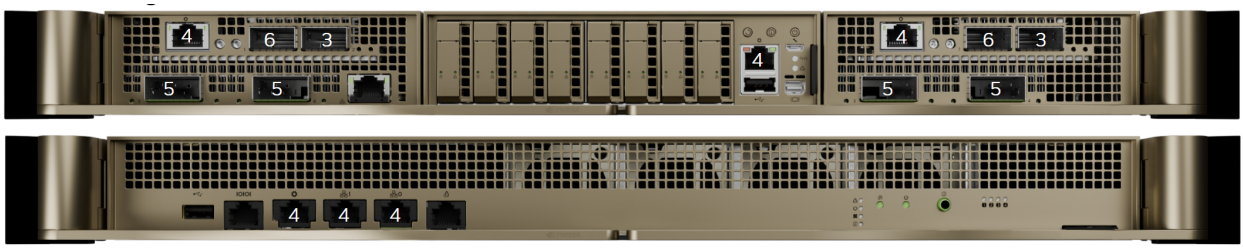


Fig. 1: DGX GB200 Compute and Switch Trays - Identification of network port locations

The following image describes the port types available in the compute tray, and how DGX OS identifies each of the ports. Note that the port names in black are the default configurations. Since the BlueField-3 card ports can be switched to InfiniBand mode, those network port names are called out in light gray.

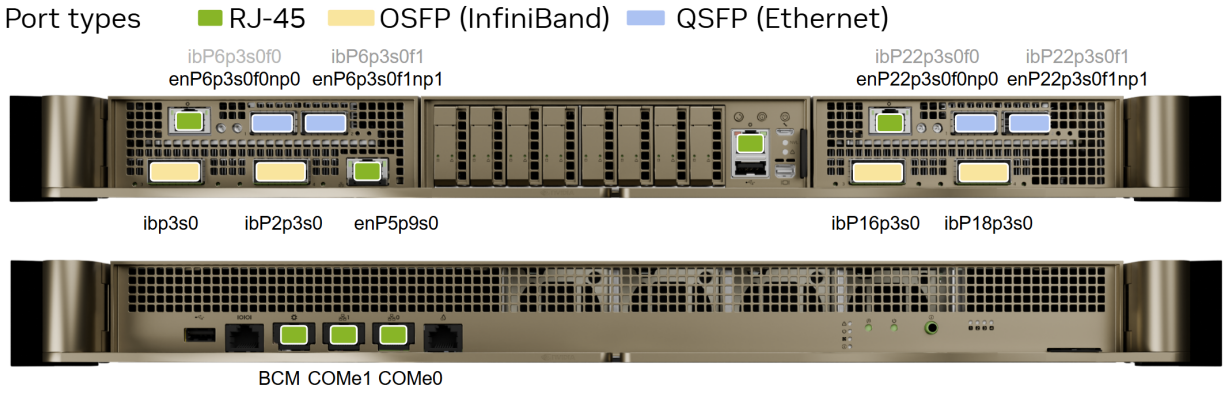


Fig. 2: DGX GB200 Compute and Switch Trays - Identification of network port types

2.3. Compute Tray Network Ports

The following table describes the function and type of each of the compute tray network ports.

Table 1: Compute Tray Network Ports

Port	Port Type	Network / switch
ConnectX-7 port	OSFP	computenet (InfiniBand)
BlueField-3 left port	QSFP	storagenet (Ethernet)
BlueField-3 right port	QSFP	dgxnet (Ethernet)
Bluefield-3 RJ-45 port	RJ45	ipminet (Ethernet)
1 GbEthernet port LAN	RJ45	Not Connected
1 GbEthernet Port BMC	RJ45	ipminet (Ethernet)

2.4. Network Interface Names

The following table describes the network interface names for each of the ports on the compute tray.

Table 2: Network Interface Names

Port	PCIe Bus	Interface name	RDMA
Left Bay OSFP P3	00:03:00.0	ibp3s0	mlx5_0
Left Bay OSFP P4	02:03:00.0	ibp2p3s0	mlx5_1
Right Bay OSFP P3	10:03:00.0	ibp16p3s0	mlx5_4
Right Bay OSFP P4	12:03:00.0	ibp11p3s0	mlx5_5
Left Bay BF3 P1	06:03:00.0	enP6p3s0f0np0	mlx5_2
Left Bay BF3 P2	06:03:00.1	enP6p3s0f1np1	mlx5_3
Right Bay BF3 P1	16:03:00.0	enP22p3s0f0np0	mlx5_6
Right Bay BF3 P2	16:03:00.1	enP22p3s0f1np1	mlx5_7
LAN		enP5p9s0	

2.5. NVLink Switch Network Ports

The following table describes the function and type of each of the NVLink switch network ports.

Table 3: NVLink Switch Network Ports

Port (from left to right)	Port Type	Network / switch
USB port	USB	Not applicable
RS232 to CPU UART or for BMC	RJ45	Not connected
1GbEthernet port BMC	RJ45	ipminet (Ethernet)
1GbEthernet for switch management	RJ45	ipminet (Ethernet)
1GbEthernet for switch management	RJ45	ipminet (Ethernet)
Leakage connector	RJ45	Not connected

Chapter 3. Storage

NVIDIA DGX GB200 systems are part of SuperPODs designed for massive-scale AI and HPC workloads. Their storage architecture is a critical component for feeding data to the GPUs efficiently. The system uses a multi-tiered storage hierarchy that includes internal storage as well as network-connected storage over a high-speed network.

3.1. Internal Storage

Each DGX compute tray within the DGX GB200 is equipped with local NVMe storage for caching, and a single M.2 NVMe drive is used for booting the operating system.

For local caching, the E1.s NVMe drives are configured in RAID 0 as a single volume. During the first read of a dataset from shared storage, the DGX system's software can automatically cache a copy of the data to the local NVMe devices using `cachefilesd`, if configured. For subsequent reads of the same data, applications can then access it from the local NVMe cache, which provides significantly faster access than retrieving it again across the network. This caching process is transparent to users and applications.

Local storage can also be used for checkpointing acceleration. While primary checkpoints typically go to high-performance storage for reliability, local NVMe can be used for temporary, faster checkpointing before being flushed to shared storage, or for smaller, very frequent checkpoints within an epoch.

3.2. External Storage

The DGX GB200's primary data source is a high-performance, shared parallel file system available from partners as described in the SuperPOD reference architecture. This storage is accessed over converged Ethernet by default, although InfiniBand can also be used. The shared storage holds the massive datasets required for AI training and serves as the original data source.

For complete information on sizing and performance requirements, please refer to the reference architecture available at Storage Architecture — NVIDIA DGX SuperPOD.

Chapter 4. Software

The DGX GB200 software stack is composed of software and firmware that runs on the various compute nodes, switch nodes, and the power shelves.

4.1. Mission Control Software

NVIDIA Mission Control combines the workload aware intelligence with full-stack observability (this includes datacenter facilities integration) to ensure that developer jobs are executed on the right resources, intelligently orchestrated to maximize throughput and uptime so that the infrastructure as a whole is working in concert with the datacenter center facilities supporting it.

AI practitioners can seamlessly manage jobs and select power profiles based on workload type and power vs. performance goals. Topology-aware scheduling aligns the job at multiple layers ensuring the entire rack functions like one big GPU. And thanks to fast checkpointing and tiered restart, no more finding out halfway into your run that your job crashed and you just lost a week of work based on the last checkpoint.

Jobs that use Mission Control are rapidly checkpointed and when errors or faults occur the software automatically isolates those problems, and then quickly restarts the job without manual intervention. With unified, integrated telemetry the software is able to isolate the specific component that is faulty and cordons off affected infrastructure, moving workloads to healthy nodes while taking into consideration the system topology.

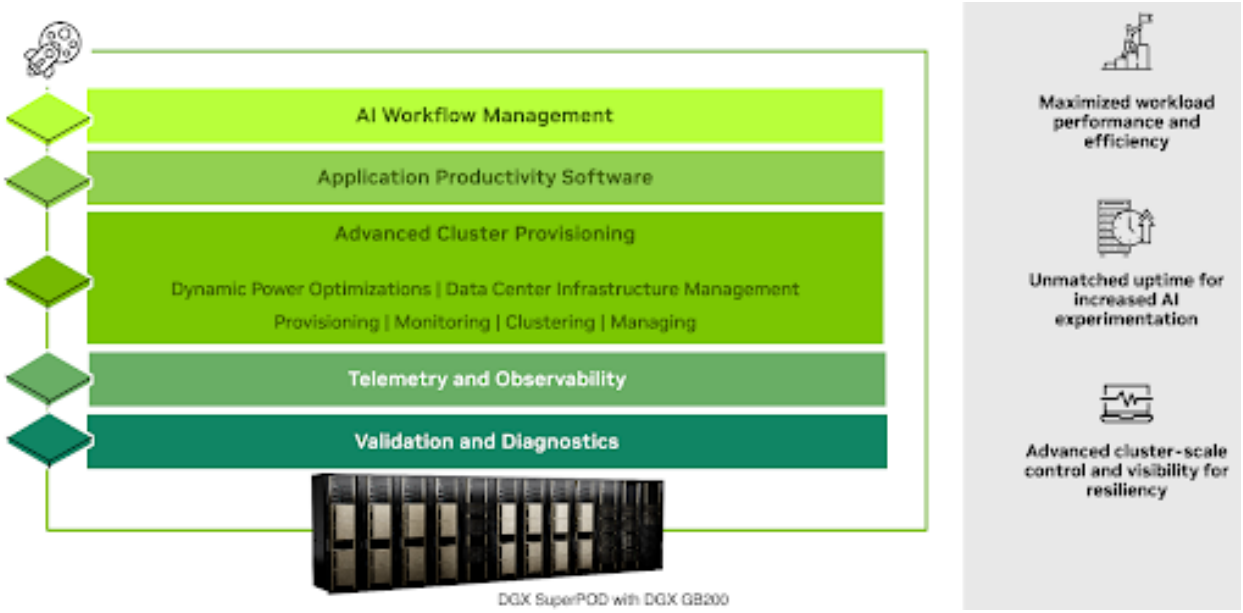
It's important to remember that the infrastructure relies on the facilities that support it. This software monitors power and cooling, and can even detect leaks, to ensure the proper functioning of the facilities' power and cooling systems, such as those supplying power to a rack or providing liquid cooling.

In conclusion, Mission Control provides end-to-end observability for the AI factory, addressing the challenge of not only deploying but also maintaining complex systems. This software offers cluster-scale control and visibility, improving infrastructure resiliency. It allows for effortless scaling and management of your cluster, simplifies the balancing of performance and power efficiency, and enables regular health checks to validate hardware and performance throughout the infrastructure lifecycle.

NVIDIA Mission Control also provides the following software benefits:

- *Simplified cluster setup and provisioning*: New automation and standardized application programming interfaces to accelerate time to deployment with integrated inventory management, and visualizations.
- *Seamless workload orchestration*: Simplifies Slurm and Kubernetes workflows.
- *Energy-optimized power profiles*: Balances power requirements and tunes GPU performance for various workload types with developer-selectable controls.

- *Autonomous job recovery*: Identifies, isolates, and recovers from inefficiencies without manual intervention to maximize developer productivity and infrastructure resiliency.
- *Customizable dashboards*: Tracks key performance indicators with access to critical telemetry data about clusters.
- *On-demand health checks*: Validates hardware and cluster performance throughout the infrastructure lifecycle.
- *Building management integration*: Provides enhanced coordination with building management systems to provide more control for power and cooling events, including rapid leakage detection.



4.2. CUDA Software Platform

NVIDIA CUDA® allows applications to use GPUs for accelerated computing and provides more than 150 libraries to create high-performance applications across programming languages and use cases. Here are some additional benefits:

- CUDA provides the tools to optimize and debug your GPU accelerated application.
- CUDA-enabled applications run on the GB200 Superchips and leverage the NVLink network to provide distributed processing across every GPU in the DGX GB200 system.

Refer to the NVIDIA CUDA documentation for more information about CUDA.

4.3. Internode Memory Exchange Service

The NVIDIA Internode Memory Exchange/Management (IMEX) is a secure service that facilitates the mapping of GPU memory over NVLink between the GPUs in an NVLink domain across the OS/node boundary using memory export and import operations. The service is started on all nodes in the NVLink domain during system startup or during the launch of the CUDA job.

Here are the key features:

- Facilitates memory sharing across compute nodes.
- Manages the lifecycle of the shared memory.
- Registers for the memory import/unimport events with the GPU driver.
- Does not directly communicate with CUDA or user applications.
- Communicates across nodes using the compute node's network employing TCP/IP and gRPC connections.
- Runs exclusively on compute nodes.

IMEX domain vs. IMEX channel

- An IMEX domain is an OS instance or a group of securely connected OS instances that use the IMEX service daemon in a multi-node system. *On single-node systems, the IMEX daemon is not required.*
- **An IMEX channel is a communication path exposed by the GPU driver within an IMEX domain to allow sharing memory securely in a multi-user environment.**
  - An IMEX channel is a logical entity that is represented by a /dev node.
  - The IMEX channels are global resources within the IMEX domain.
  - When exporter and importer CUDA processes have been granted access to the same IMEX channel, they can securely share memory.

4.4. NVIDIA Switch Tray Software

With the fourth generation of NVSwitches, NVIDIA has implemented a unified architecture that spans across NVLink, InfiniBand, and Ethernet switches. The NVSwitch trays that implement these fourth-generation NVSwitches have an NVOS image installed on them.

4.4.1. NVIDIA Switch OS

The NVLink Switch Tray comes with the NVIDIA NVSmanagerwitch™ operating system (NVOS) that enables the management and configuration of NVIDIA's switch system platforms. NVOS provides a suite of management options, incorporates an industry-standard command line interface (CLI), and OpenAPI (Swagger) that allows system administrators to easily configure and manage the system.

NVOS includes the NVLink Subnet Manager (NVLSM), the Fabric Manager (FM), NMX services such as NMX-Controller and NMX-Telemetry, and the NVSwitch firmware.

4.4.2. NMX-Controller

The NMX-Controller is a cluster application for fabric Software Defined Network (SDN) services. In the DGX GB200 the SDN services include SM and FM.

The NVOS cluster infrastructure includes the cluster applications package file in the NVOS image. The packages are automatically installed with the NVOS image installation and upgrade process, which ensures a hassle-free setup.

4.4.3. Fabric Manager

FM configures the NVSwitch memory fabrics to form a large memory fabric among the participating GPUs and monitors the NVLinks that support the fabric. At a high level, FM has the following responsibilities:

- Configures routing among NVSwitch ports.
- Sets up the GPU side routing/port map if applicable.
- Coordinates with the GPU driver to initialize the GPUs.
- Monitors the fabric for NVLink and NVSwitch errors.

4.4.4. NVLink Subnet Manager

NVLink Subnet Manager (NVLSM) originated from the IB networking and added additional logic to manage the NVSwitch and NVLinks in the switch trays. At a high level, the NVLSM provides the following functionalities in NVSwitch-based systems:

- Discovers the NVLink network topology.
- Assigns a logical identifier (LID) to the GPU and NVSwitch NVLink ports.
- Calculates and programs the switch forwarding table.
- Programs the Partition Key (PKEY) for NVLink partitions.
- Monitors the changes in the NVLink fabric.

4.4.5. NMX-Telemetry

NMX-T is a subsystem that collects, aggregates, and transmits telemetry data from various devices, applications, and platforms. In the context of this guide, it collects and aggregates from all the NVLink switches in a NVLink domain.

4.5. DOCA Software Framework

The DOCA framework unlocks the potential of the NVIDIA® BlueField® networking platform by enabling the rapid creation of applications and services that offload, accelerate, and isolate data center workloads. It lets developers create software-defined, cloud-native, DPU- and SuperNIC-accelerated services with zero-trust protection, which addresses the performance and security demands of modern data centers. DOCA-Host includes the host drivers and tools for BlueField and ConnectX® devices in the rack.

4.6. NVIDIA Collective Communication Library

GPU communication libraries like NVIDIA Collective Communications Library (NCCL), NVIDIA NVSHMEM, and Unified Communication X (UCX) provide a framework for efficient and scalable inter-GPU communication. Communication performance is crucial for the overall Deep Learning (DL) performance and HPC applications that run on multi-node racks and large clusters. They support a variety

of interconnect technologies including PCIe, NVLink, NVLink Network (known as NVLink Multi-node), InfiniBand, RoCE, and so on.

NCCL provides inter-GPU communication primitives that are topology-aware, removing the need for developers to optimize their applications for specific systems. NCCL implements collective communication and point-to-point send/receive primitives for intra- and inter-node data transfers. Refer to the NCCL documentation for more information.

NVSHMEM provides a partitioned global address space (PGAS) for data that spans the memory of multiple GPUs. In addition to the CPU and CUDA stream communication APIs, NVSHMEM also provides device APIs for communication. These APIs enable low latency communication between GPUs and allows applications to do fine-grained and/or fused overlap of compute and communication from the CUDA kernel. Refer to the NVIDIA NVSHMEM documentation for more information.

UCX is a framework that provides a common set of CPU-initiated tag-matching, active-messaging, and remote memory access APIs over a variety of network protocols and hardware, which enhances scalability and performance in distributed computing systems. Refer to NVIDIA UCX documentation for more information.

4.7. Data Center GPU Manager

NVIDIA Data Center GPU Manager (DCGM) is a suite of tools that manage and monitor NVIDIA datacenter GPUs and NVSwitches in cluster environments. It includes active health monitoring, comprehensive diagnostics, system alerts, and governance policies including power and clock management. The suite can be used as a standalone option by infrastructure teams and easily integrates into cluster management tools, resource scheduling, and monitoring products from NVIDIA partners. Refer to the NVIDIA DCGM documentation for more information.

4.8. NVIDIA Firmware Tools (MFT)

The MFT package, which is a set of network ASIC firmware management tools, is used for the following reasons:

- Generating a standard or customized firmware image.
- Querying for firmware information.
- Flashing a firmware image.

The MFT package for DGX GB200 can be found in the accompanying software release.

Chapter 5. Safety

This section provides information about how to safely use the NVIDIA DGX™ GB200 system.

5.1. Safety Information

To reduce the risk of bodily injury, electrical shock, fire, and equipment damage, read this document and observe all warnings and precautions in this guide before installing or maintaining your server product.

In the event of a conflict between the information in this document and information provided with the product or on the website for a particular product, the product documentation takes precedence.

Your server should be integrated and serviced only by technically qualified persons.

You must adhere to the guidelines in this guide and the assembly instructions in your server manuals to ensure and maintain compliance with existing product certifications and approvals. Use only the described, regulated components specified in this guide. Use of other products I components will void the UL Listing and other regulatory approvals of the product and may result in noncompliance with product regulations in the region(s) in which the product is sold.

5.2. Safety Warnings and Cautions

To avoid personal injury or property damage, before you begin installing the product, read, observe, and adhere to all of the following safety instructions and information.

The following safety symbols may be used throughout the documentation and may be marked on the product and the product packaging.

- CAUTION: Indicates the presence of a hazard that may cause minor personal injury or property damage if the CAUTION is ignored.
- WARNING: Indicates the presence of a hazard that may result in serious personal injury if the WARNING is ignored.

Indicates potential hazard if indicated information is ignored.

Indicates shock hazards that result in serious injury or death if safety instructions are not followed.

Indicates hot components or surfaces

Indicates do not touch fan blades, may result in injury.

Shock hazard: The product might be equipped with multiple power cords. - To remove all hazardous voltages, disconnect all power cords. - High leakage current ground (earth) connection to the Power Supply is essential before connecting the supply.

Recycle the battery.

The rail racks are designed to carry only the weight of the server system. Do not use rail-mounted equipment as a workspace. Do not place additional load onto any rail-mounted equipment.

5.3. Intended Application Uses

This product was evaluated as Information Technology Equipment (ITE), which may be installed in offices, schools, computer rooms, and similar commercial type locations.

The suitability of this product for other product categories and environments (such as medical, industrial, residential, alarm systems, and test equipment), other than an ITE application, may require further evaluation.

5.4. Site Selection

Choose a site that is:

- Clean, dry, and free of airborne particles (other than normal room dust).
- Well-ventilated and away from sources of heat including direct sunlight and radiators.
- Away from sources of vibration or physical shock.
- In regions that are susceptible to electrical storms, we recommend you plug your system into a surge suppressor and disconnect telecommunication lines to your modem during an electrical storm.

- Provided with a properly grounded wall outlet.
- Provided with sufficient space to access the power supply cord(s), because they serve as the product's main power disconnect.

5.5. Equipment Handling Practices

To reduce the risk of personal injury or equipment damage, do the following:

- Conform to local occupational health and safety requirements when moving and lifting equipment.
- Use mechanical assistance or other suitable assistance when moving and lifting equipment.

5.6. Electrical Precautions

5.6.1. Power and Electrical Warnings

Caution

The power button, indicated by the stand-by power marking, DOES NOT completely turn off the system AC power; standby power is active whenever the system is plugged in. To remove power from system, you must unplug the AC power cord from the wall outlet. Make sure all AC power cords are unplugged before you open the chassis, or add or remove any non hot-plug components.

Do not attempt to modify or use an AC power cord if it is not the exact type required. A separate AC cord is required for each system power supply.

Some power supplies in servers use Neutral Pole Fusing. To avoid risk of shock use caution when working with power supplies that use Neutral Pole Fusing.

The power supply in this product contains no user-serviceable parts. Do not open the power supply. Hazardous voltage, current and energy levels are present inside the power supply. Return to manufacturer for servicing.

When replacing a hot-plug power supply, unplug the power cord to the power supply being replaced before removing it from the server.

To avoid risk of electric shock, tum off the server and disconnect the power cords, telecommunications systems, networks, and modems attached to the server before opening it.

5.6.2. Power Cord Warnings

Caution

To avoid electrical shock or fire, check the power cord(s) that will be used with the product as follows:

- Do not attempt to modify or use the AC power cord(s) if they are not the exact type required to fit into the grounded electrical outlets.

- The power cord(s) must meet the following criteria:
  - The power cord must have an electrical rating that is greater than that of the electrical current rating marked on the product.
  - The power cord must have safety ground pin or contact that is suitable for the electrical outlet.
  - The power supply cord(s) is/ are the main disconnect device to AC power. The socket outlet(s) must be near the equipment and readily accessible for disconnection.
  - The power supply cord(s) must be plugged into socket-outlet(s) that is /are provided with a suitable earth ground.

5.7. System Access Warnings

To avoid personal injury or property damage, the following safety instructions apply whenever accessing the inside of the product:

- Turn off all peripheral devices connected to this product.
- Turn off the system by pressing the power button to off.
- Disconnect the AC power by unplugging all AC power cords from the system or wall outlet.
- Disconnect all cables and telecommunication lines that are connected to the system.
- Retain all screws or other fasteners when removing access cover(s). Upon completion of accessing inside the product, refasten access cover with original screws or fasteners.
- Do not access the inside of the power supply. There are no serviceable parts in the power supply.
- Return to manufacturer for servicing.
- Power down the server and disconnect all power cords before adding or replacing any non hot-plug component.
- When replacing a hot-plug power supply, unplug the power cord to the power supply being replaced before removing the power supply from the server.

Caution

If the server has been running, any installed processor(s) and heat sink(s) may be hot. Unless you are adding or removing a hot-plug component, allow the system to cool before opening the covers. To avoid the possibility of coming into contact with hot component(s) during a hot-plug installation, be careful when removing or installing the hot-plug component(s).

Caution

To avoid injury do not contact moving fan blades. Your system is supplied with a guard over the fan, do not operate the system without the fan guard in place.

5.8. Rack Mount Warnings

The following installation guidelines are required by UL to maintain safety compliance when installing your system into a rack.

The equipment rack must be anchored to an unmovable support to prevent it from tipping when a server or piece of equipment is extended from it. The equipment rack must be installed according to the rack manufacturer's instructions.

Install equipment in the rack from the bottom up with the heaviest equipment at the bottom of the rack.

Extend only one piece of equipment from the rack at a time.

You are responsible for installing a main power disconnect for the entire rack unit. This main disconnect must be readily accessible, and it must be labeled as controlling power to the entire unit, not just to the server(s).

To avoid risk of potential electric shock, a proper safety ground must be implemented for the rack and each piece of equipment installed in it.

Elevated Operating Ambient- If installed in a closed or multi-unit rack assembly, the operating ambient temperature of the rack environment may be greater than room ambient. Therefore, consideration should be given to installing the equipment in an environment compatible with the maximum ambient temperature (Tma) specified by the manufacturer.

Reduced Air Flow -Installation of the equipment in a rack should be such that the amount of air flow required for safe operation of the equipment is not compromised.

Mechanical Loading- Mounting of the equipment in the rack should be such that a hazardous condition is not achieved due to uneven mechanical loading.

Circuit Overloading- Consideration should be given to the connection of the equipment to the supply circuit and the effect that overloading of the circuits might have on overcurrent protection and supply wiring. Appropriate consideration of equipment nameplate ratings should be used when addressing this concern.

Reliable Earthing- Reliable earthing of rack-mounted equipment should be maintained.

Particular attention should be given to supply connections other than direct connections to the branch circuit (for example, the use of power strips).

5.9. Electrostatic Discharge

Caution

ESD can damage drives, boards, and other parts. We recommend that you perform all procedures at an ESD workstation. If one is not available, provide some ESD protection by wearing an antistatic wrist strap attached to chassis ground (any unpainted metal surface) on your server when handling parts.

Always handle boards carefully. They can be extremely sensitive to ESD. Hold boards only by their edges. After removing a board from its protective wrapper or from the server, place the board component side up on a grounded, static free surface. Use a conductive foam pad if available but not the board wrapper. Do not slide board over any surface.

5.10. Other Hazards

5.10.1. CALIFORNIA DEPARTMENT OF TOXIC SUBSTANCES CONTROL

Perchlorate Material - special handling may apply. See www.dtsc.ca.gov/perchlorate.

Perchlorate Material: Lithium battery (CR2032) contains perchlorate. Please follow instructions for disposal.

5.10.2. NICKEL

NVIDIA Bezel. The bezel's decorative metal foam contains some nickel. The metal foam is not intended for direct and prolonged skin contact. Please use the handles to remove, attach or carry the bezel. While nickel exposure is unlikely to be a problem, you should be aware of the possibility in case you are susceptible to nickel-related reactions.

5.10.3. Battery Replacement

Caution

There is the danger of explosion if the battery is incorrectly replaced. When replacing the battery, use only the battery recommended by the equipment manufacturer.

Dispose of batteries according to local ordinances and regulations. Do not attempt to recharge a battery.

Do not attempt to disassemble, puncture, or otherwise damage a battery.

□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□ □□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□

5.10.4. Cooling and Airflow

Caution

Carefully route cables as directed to minimize airflow blockage and cooling problems. For proper cooling and airflow, operate the system only with the chassis covers installed.

Operating the system without the covers in place can damage system parts. To install the covers:

- Check first to make sure you have not left loose tools or parts inside the system.
- Check that cables, add-in cards, and other components are properly installed.
- Attach the covers to the chassis according to the product instructions.

The equipment is intended for installation only in a Server Room/ Computer Room where both these conditions apply:

- Access can only be gained by SERVICE PERSONS or by USERS who have been instructed about the reasons for the restrictions applied to the location and about any precautions that shall be taken.
- Access is through the use of a TOOL or lock and key, or other means of security, and is controlled by the authority responsible for the location.

Chapter 6. System Control and Configuration

6.1. System Operations

To bring a DGX GB200 rack online, power on and configure the NVSwitch trays first to establish the NVLink fabric, then power on the compute trays. This sequence ensures high-bandwidth GPU-to-GPU communication is active from the start, preventing potential inter-node communication issues and avoiding the need for compute node restarts.

Before bringing the system online, ensure all infrastructure components are operational: power shelves energized, top-of-rack switches active, and NMX-M services available. You can verify these components are operational through Base Command Manager (BCM).

BCM manages all system components. Use BCM to power on the NVLink switch trays first, followed by the compute trays. Each NVLink switch tray boots with rack-specific configuration, enabling direct GPU communication across all compute trays via NVLink.

After the switch trays are operational, power on the compute trays. As they boot, BCM confirms that the latest updates and configurations are installed in the operating system. The compute trays establish communication with the NVLink switch trays and other resources such as storage and additional compute trays.

Power operations are orchestrated through NVIDIA Mission Control. Refer to the Administration Guide for a complete list of available system operations.

6.2. Operating System Updates

The operating system comes pre-installed on each tray but is managed by Base Command Manager. When a compute tray boots, it checks for updates over the network. If updates are available, they are installed before booting; otherwise, the tray boots quickly from local storage. This ensures each compute tray boots with the correct operating system and latest configuration.

You can install additional software and system configurations directly to the software images in the Base Command Manager head node. These changes become persistent once assigned to a specific compute tray, providing a single point of management for all operating systems needed in the cluster.

For details on customizing software images configured to boot on system trays, refer to the NVIDIA Mission Control Administration Guide.

6.3. Firmware Updates

Compute and switch trays require periodic firmware updates. Firmware update procedures are released with specific instructions for each release, as different components may require specific actions. Not all components need updates in every release, so always review the firmware release notes for the complete recipe (firmware packages and version numbers).

While firmware can be applied to individual components, it is typically released as a recipe to update all applicable components simultaneously. NVIDIA Mission Control provides features to perform firmware updates efficiently instead of manually updating components one at a time.

6.4. Health Checks

NVIDIA Mission Control continuously monitors system health, verifying the functionality of compute and switch trays. It also monitors critical factors such as potential system leaks that may require immediate operator attention.

Health check issues are highlighted on the NVIDIA Mission Control dashboards, allowing system administrators to investigate the source of problems. Issues may stem from software misconfiguration, hardware problems, or data center conditions such as air temperature or liquid cooling system issues.

6.4.1. Redfish Commands for Compute Tray

The following Redfish commands can be used to manage power operations for compute trays. You can use these commands in two ways:

1. **Using environment variables**: Set the BMC IP address and credentials as environment variables before running the command (replace the IP address with your BMC's IP address):

   ```
   export bmcip="192.168.1.100"
   ```

   Then run the curl commands as shown in the examples below.

2. **Manual substitution**: Alternatively, you can manually replace `${bmcip}` with your BMC's IP address in each command.

The examples below use the environment variable approach with default credentials. Be sure to replace with your actual BMC credentials.

For advanced troubleshooting and support, see the *Debug Log Collection* section at the end of this document.

6.4.1.1 Power Control Operations

This section provides commands for remotely controlling host power states and behaviors using Redfish and `curl`.

6.4.1.1.1 Virtual A/C Cycle

Simulates a power drop (like pulling and reinserting the power cord) without physically disconnecting the power.

```
curl -k -u "root:openBmc" \
     -H "Content-Type: application/json" \
     -X POST \
     -d '{"ResetType":"AuxPowerCycleForce"}' \
     https://${bmcip}/redfish/v1/Chassis/BMC_0/Actions/Oem/NvidiaChassis.
↪AuxPowerReset
```

6.4.1.1.2 Graceful Shutdown

Initiates a clean shutdown of the host system, allowing the operating system to properly terminate all services and processes.

```
curl -s -k -u "root:openBmc" \
     -X POST \
     -d '{"ResetType": "GracefulShutdown"}' \
     https://${bmcip}/redfish/v1/Systems/System_0/Actions/ComputerSystem.Reset
```

6.4.1.1.3 Force Power Off

Immediately cuts power to the host without waiting for the OS to shut down.

```
curl -s -k -u "root:openBmc" \
     -X POST \
     -d '{"ResetType": "ForceOff"}' \
     https://${bmcip}/redfish/v1/Systems/System_0/Actions/ComputerSystem.Reset
```

6.4.1.1.4 Power On Host

Turns the host on.

```
curl -s -k -u "root:openBmc" \
     -X POST \
     -d '{"ResetType": "On"}' \
     https://${bmcip}/redfish/v1/Systems/System_0/Actions/ComputerSystem.Reset
```

6.4.1.1.5 Power Cycle Host

Performs a full power cycle (off and then back on).

```
curl -s -k -u "root:openBmc" \
     -X POST \
     -d '{"ResetType": "PowerCycle"}' \
     https://${bmcip}/redfish/v1/Systems/System_0/Actions/ComputerSystem.Reset
```

6.4.1.1.6 Set Power Restore Policy: Always Off

After power loss, the system will remain off when power is restored.

```
curl -s -k -u "root:openBmc" \
     -X PATCH \
```

(continues on next page)

(continued from previous page)

```
     -d '{"PowerRestorePolicy": "AlwaysOff"}' \
     https://${bmcip}/redfish/v1/Systems/System_0
```

6.4.1.1.7 Set Power Restore Policy: Last State

After power loss, the system will return to its previous state (on or off) when power is restored.

```
curl -s -k -u "root:openBmc" \
     -X PATCH \
     -d '{"PowerRestorePolicy": "LastState"}' \
     https://${bmcip}/redfish/v1/Systems/System_0
```

6.4.1.1.8 Set Power Restore Policy: Always On

After power loss, the system will automatically power on when power is restored.

```
curl -s -k -u "root:openBmc" \
     -X PATCH \
     -d '{"PowerRestorePolicy": "AlwaysOn"}' \
     https://${bmcip}/redfish/v1/Systems/System_0
```

6.4.1.2 HMC/BMC Reset Operations

This section provides commands for remotely resetting the Host Management Controller (HMC) and BMC using Redfish and `curl`. These operations are useful for troubleshooting, applying firmware updates, or recovering from certain system faults. Use these commands with caution, as some reset types may impact system availability or trigger a host restart.

> **Note**
>
> Use these reset commands with care. For most scenarios, prefer a graceful restart. Forced resets should only be used if the controller is unresponsive or as directed by support documentation.

6.4.1.2.1 HMC Graceful Restart

Performs a controlled restart of the HMC, allowing it to shut down services cleanly before rebooting. This is the preferred method for most maintenance scenarios.

```
curl -k -u "root:openBmc" \
     -X POST \
     -d '{"ResetType": "GracefulRestart"}' \
     https://${bmcip}/redfish/v1/Managers/HGX_BMC_0/Actions/Manager.Reset
```

6.4.1.2.2 HMC Force Restart

Immediately restarts the HMC without waiting for services to shut down gracefully.

> **Warning**
>
> This is not recommended except in emergency situations, as it may lead to a host restart or data loss.

```
curl -k -u "root:openBmc" \
     -X POST \
     -d '{"ResetType": "ForceRestart"}' \
     https://${bmcip}/redfish/v1/Managers/HGX_BMC_0/Actions/Manager.Reset
```

6.4.1.2.3 BMC Reset

Performs a forced reset of the BMC. This operation is typically used after firmware updates or if the BMC becomes unresponsive. The following example includes a descriptive payload, which may be required for certain firmware update workflows.

```
curl -k -u "root:openBmc" \
     -X POST \
     -d '{"ResetType": "ForceRestart", "Description": "BMC bundle update curl
↪"}' \
     https://${bmcip}/redfish/v1/Managers/BMC_0/Actions/Manager.Reset
```

6.4.1.3 Firmware and Image Management

This section provides Redfish and `curl` commands for managing firmware updates and querying firmware image information on the compute tray. These operations are essential for keeping your system up to date and verifying the integrity and type of installed firmware.

6.4.1.3.1 Firmware Bundle Update

Uploads and applies a firmware bundle to the system. The following command uses a multipart form upload to send the update file to the BMC.

```
curl -k -u "root:openBmc" --header 'Expect:' --location --request POST \
     "https://${bmcip}/redfish/v1/UpdateService/update-multipart" \
     -F 'UpdateParameters={"Targets":["/redfish/v1/Chassis/HGX_Chassis_0"],
↪"ForceUpdate":true};type=application/json' \
     -F UpdateFile="@${FILENAME}"
```

6.4.1.3.2 Get Firmware Versions (with expand)

Retrieves detailed firmware version information for all components, using the `expand` query to include nested inventory data.

```
curl -k -u "root:openBmc" -X GET \
     "https://${bmcip}/redfish/v1/UpdateService/FirmwareInventory?expand=.
↪$levels=1" \
     | jq -r '.Members[] | [.Id, .Version] | @tsv' | column -t
```

6.4.1.3.3 Check BMC Firmware Signing Type

Checks the signing type of the BMC firmware image, which helps confirm the authenticity and security of the installed firmware.

> **Note**
>
> Replace `<active image slot>` with the appropriate image slot identifier for your system.

```
curl -k -u "root:openBmc" -X GET \
     "https://${bmcip}/redfish/v1/Chassis/ERoT_BMC_0/Oem/NvidiaRoT/
↪RoTProtectedComponents/BMC_0/Images/<active image slot>" \
     | grep SigningType
```

6.4.1.3.4 Check HMC Firmware Signing Type

Checks the signing type of the HMC firmware image.

> **Note**
>
> Replace `<active image slot>` with the appropriate image slot identifier for your system.

```
curl -k -u "root:openBmc" -X GET \
     "https://${bmcip}/redfish/v1/Chassis/HGX_ERoT_BMC_0/Oem/NvidiaRoT/
↪RoTProtectedComponents/HGX_BMC_0/Images/<active image slot>" \
     | grep SigningType
```

6.4.1.4 Write Protection

This section provides commands for managing hardware write protection on the system using Redfish and `curl`. Write protection helps prevent unauthorized changes to critical system components.

6.4.1.4.1 Enabling Global Write Protect

Enables hardware write protection for the system.

```
curl -s -k -u 'root:openBmc' -X PATCH \
     -d '{ "Oem":{"Nvidia":{"HardwareWriteProtectEnable":true}}}' \
     https://${bmcip}/redfish/v1/Chassis/Chassis_0
```

6.4.1.4.2 Disabling Global Write Protect

Disables hardware write protection for the system.

```
curl -s -k -u 'root:openBmc' -X PATCH \
     -d '{ "Oem":{"Nvidia":{"HardwareWriteProtectEnable":false}}}' \
     https://${bmcip}/redfish/v1/Chassis/Chassis_0
```

6.4.1.4.3 Reading Global Write Protect

Reads the current hardware write protection status.

```
curl -s -k -u 'root:openBmc' -X GET \
     https://${bmcip}/redfish/v1/Chassis/Chassis_0 | jq '.Oem.Nvidia.
↪HardwareWriteProtectEnable'
```

Note

Enabling write protection is recommended after firmware updates or configuration changes to prevent accidental or unauthorized modifications.

6.4.1.5 Account Management

This section provides commands for managing user accounts on the BMC using Redfish and `curl`. You can create new accounts or update passwords for existing accounts.

6.4.1.5.1 Creating a New Account

Creates a new user account with specified username, password, and role.

```
curl -s -k -u "root:openBmc" \
     -X POST \
     -H "Content-Type: application/json" \
     -d '{"UserName": "newuser", "Password": "newPassword", "RoleId":
↪"Operator"}' \
     https://${bmcip}/redfish/v1/AccountService/Accounts
```

Note

You can set the `RoleId` to “Administrator”, “Operator”, or “ReadOnly” based on the required access level for the user.

6.4.1.5.2 Updating Password for an Existing Account

Updates the password for an existing account (e.g., the `root` user).

```
curl -k -X PATCH -H "Content-Type: application/json" \
     -u root:CURRENT_PASSWORD \
     https://${bmcip}/redfish/v1/AccountService/Accounts/root \
     --data '{ "Attributes": { "Password": "NEW_PASSWORD" } }'
```

Note

Replace `CURRENT_PASSWORD` with the current password for the account, and `NEW_PASSWORD` with the desired new password.

6.4.1.6 Event and Log Management

This section provides commands for clearing event logs and managing system logs using Redfish and `curl`. These operations help maintain log clarity and support troubleshooting.

6.4.1.6.1 Clear HMC Event Log

Clears all entries from the HMC event log. Use this command after reviewing or archiving log data to maintain log clarity and free up space.

```
curl -k -u "root:openBmc" -X POST \
     https://${bmcip}/redfish/v1/Systems/HGX_Baseboard_0/LogServices/EventLog/
↪Actions/LogService.ClearLog
```

6.4.1.6.2 Clear BMC Event Log

Removes all entries from the BMC event log. This is useful for resetting the log after resolving issues or before starting new troubleshooting steps.

```
curl -k -u "root:openBmc" -X POST \
     https://${bmcip}/redfish/v1/Systems/System_0/LogServices/EventLog/
↪Actions/LogService.ClearLog
```

6.4.1.7 Reset and Factory Defaults

This section provides commands for restoring system components to their factory default state. These operations are useful for recovery, troubleshooting, or preparing a system for redeployment.

6.4.1.7.1 Reset BMC to Factory Defaults

Restores the BMC to its original factory settings. All user configurations and custom settings will be lost. Use this if the BMC is unresponsive or you need to clear all settings.

```
curl -k -u "root:openBmc" -X POST \
     https://${bmcip}/redfish/v1/Managers/BMC_0/Actions/Manager.
↪ResetToDefaults \
     -d '{"ResetToDefaultsType": "ResetAll"}'
```

6.4.1.7.2 Reset HMC to Factory Defaults

Resets the Host Management Controller (HMC) to its factory default configuration. This will remove all custom settings and user data from the HMC.

```
curl -k -u "root:openBmc" -X POST \
     https://${bmcip}/redfish/v1/Managers/HGX_BMC_0/Actions/Manager.
↪ResetToDefaults \
     -d '{"ResetToDefaultsType": "ResetAll"}'
```

6.4.1.7.3 Clear the eMMC (This Will Clear the Journal)

Securely erases the eMMC storage, including the system journal. Use this operation with caution, as it will permanently delete all data stored on the eMMC.

```
curl -k -u "root:openBmc" -X POST \
     https://${bmcip}/redfish/v1/Managers/HGX_BMC_0/Actions/Oem/eMMC.
↪SecureErase
```

6.4.1.7.4 Reset SBIOS to Defaults

Restores the system BIOS (SBIOS) settings to their factory defaults. This can help resolve configuration issues or prepare the system for a fresh deployment.

```
curl -k -u "root:openBmc" -X POST \
     https://${bmcip}/redfish/v1/Systems/System_0/Bios/Actions/Bios.ResetBios
```

6.4.1.8 Serial Over LAN and Host Console Access

This section provides commands for accessing the host system console remotely using Serial Over LAN (SOL) or SSH. These methods are useful for troubleshooting, remote management, and system recovery.

6.4.1.8.1 IPMI - Host SOL Access

Establishes a Serial Over LAN (SOL) session to the host using IPMI. This allows you to interact with the system console remotely, even if the OS is not running.

```
ipmitool -C 17 -H ${bmcip} -U root -P OpenBmc -I lanplus sol activate
```

To exit the SOL session, use ~. (tilde followed by a period).

6.4.1.8.2 SSH - Host SOL Access

Provides SSH access to the host's SOL interface. This is an alternative to IPMI SOL for remote console access.

```
ssh root@${bmcip} -p 2200
```

6.4.1.9 Network Controls

This section provides commands for managing network protocol access on the BMC, including enabling or disabling IPMI and understanding credential behavior after a factory reset.

6.4.1.9.1 Disable IPMI

Disables the IPMI protocol on the BMC. This is recommended for enhanced security if IPMI remote management is not required.

```
curl -s -k -u 'root:openBmc' -X PATCH \
     https://${bmcip}/redfish/v1/Managers/BMC_0/NetworkProtocol \
     -d '{ "IPMI": {"ProtocolEnabled": false} }'
```

6.4.1.10 BMC Credentials After Factory Reset

After a BMC factory reset, the default username and password are `admin`/`admin`. You will be prompted to change the username and password at the first login to the system. Any BMC factory reset will restart this sequence.

> **Note**
>
> For security, always change the default credentials immediately after a factory reset.

6.4.1.11 Debug Log Collection

When troubleshooting complex issues or working with NVIDIA support, it is often necessary to collect detailed debug logs from the compute tray. The preferred method for capturing and packaging these logs is to use the NVDEBUG tool, which gathers all relevant system, firmware, and event logs into a single archive for analysis.

6.4.1.11.1 Capturing Debug Logs with NVDEBUG

1. **Access the BMC or host system** where NVDEBUG is available. Ensure you have the necessary permissions to run diagnostic tools.
2. **Run the NVDEBUG tool** according to your system's documentation. The typical command is:

   ```
   nvdebug --collect --output /path/to/save/debug-logs.tar.gz
   ```

   Replace `/path/to/save/debug-logs.tar.gz` with your desired output location.
3. **Transfer the log archive** to your support contact or upload it as instructed by NVIDIA support.

> **Note**
>
> NVDEBUG automatically collects a comprehensive set of logs, including BMC event logs, firmware versions, hardware inventory, and system health data. This is the recommended approach for efficient and complete log collection.

Chapter 7. Rack Reboot Sequence

To ensure that software is in a known-correct state, running a cold or warm reboot sequence may be required. This topic describes the general reboot and verification process. However, to safeguard system integrity, it is recommended to use the Mission Control to perform the reboot sequence. Refer to the Mission Control documentation for more information.

7.1. Cold Reboot Sequence

The cold rack reboot sequence should be executed for the following use cases:

- **After** the initial software installation and setup and **before** running a workload for the first time.
- **After** firmware upgrades and **before** re-running workloads.
- **Before** running the full rack diags.
- Maintenance or service operations, including but not limited to compute tray, switch tray, trunk links or a cable cartridge replacement.

Use the method that is most appropriate to complete the procedure below. You need all IP addresses for BMC, host, and power shelves.

> **Note**
>
> The sequence diverges at step 8 for NMX-C provisioning depending on whether you plan to run a multi-node workload or an L11 diag. For the diags use-case, the NMX-C services are started by the diags as part of its deployment but for regular software workloads these services have to be started manually after a cold reboot,

1. AC power off all the power shelves in the rack that will turn off all PSUs.

   The shelves can be powered off in parallel.
2. Wait three to five minutes for the capacitors to discharge from all the PSUs.
3. Ping all compute/switch tray BMCs and OS IPs to ensure that they are not reachable on the network.

   > **Note**
   >
   > Steps 1-3 are not applicable when you power up the rack for the first time.

4. AC power on all the power shelves in the rack, which will turn on all PSUs.

   The shelves can be powered on in parallel.
5. Wait for two minutes for the compute tray and switch tray BMCs to power up.
6. Ping all compute/switch tray BMC IPs to ensure they are reachable on the network.
7. The switch tray continues to boot to NVOS.
8. Wait for two more minutes for the switch trays to boot to NVOS.
9. To provision the switch tray, complete the following tasks:
   - To run the L11 diagnostics, log in to each switch tray and verify the system health check (refer to Switch Tray Health Check).
   - To run a multi-node workload:
     1. Manually repeat the entire cluster provisioning (refer to Cluster Provisioning flow).
     2. Set up and start the NMX-C services followed by the health checks verification (refer to Switch Tray Verification).
10. Power on all compute nodes from the compute tray BMC.
11. Wait five minutes.
12. Ping the compute nodes to ensure that they are reachable.
13. Run the verification checks for the compute tray (refer to *Post-reboot Verification* for more information).

7.2. Warm Reboot Sequence

The cold rack reboot sequence takes a long time. For workarounds to certain issues that are usually documented in the associated release notes, the following warm rack reboot procedure ensures that the GPUs, NVLinks, and the Switch ASICs come up in the correct state.

1. Power off all compute trays through the BMC using IPMI, Redfish APIs or other supported APIs. Here is an example using IPMI commands:

```
$ ipmitool chassis power off
$ ipmitool power off
```

2. Stop the NMX-C service on the switch node that is running the service.

```
$ nv action stop cluster app nmx-controller
```

3. Start the NMX-C on the selected switch node.

```
$ nv action start cluster app nmx-controller
```

   The cluster must already be in enabled state before you restart the NMX-C.
4. Verify NMX-C is running. Status should return *ok*.

```
$ nv show cluster apps running
Name            Status  Reason  Additional Information
--------------  ------  ------  ------------------------------
```

(continues on next page)

(continued from previous page)

```
nmx-controller  ok              CONTROL_PLANE_STATE_CONFIGURED
nmx-telemetry   ok
```

5. Power up all the compute trays through BMC using IPMI, redfish APIs or other supported APIs. Here is an example using IPMI commands:

```
ipmitool chassis power on
ipmitool power on
```

6. Run all the verification checks for the compute tray (refer to *Post-reboot Verification* for more information).

7.3. Post-reboot Verification

Complete the following verification checks when a GPU is reset, the compute tray is rebooted, or the entire rack is cold or warm rebooted. For detailed instructions on performing each verification step, refer to the MNNVL User Guide Verification section.

1. Verify that `nvidia-persistenced` is running and is in a good state..

   Refer to Checking the nvidia-persistenced Service for more information.

2. Verify that nvidia-imex service is active and running.

   Refer to Checking the nvidia-imex Service for more information.

3. Verify that all the links are active.

   Refer to Checking the NVLink status for more information.

4. Verify fabric state to make sure all GPUs have a Completed state and a Success status.

   Refer to Checking the fabric health for more information.

5. Verify peer-to-peer topology to ensure all GPUs show OK.

6. If a GPU does not show OK, reset the GPU and return to step 1.

   Refer to Checking the p2p topology for details.

Chapter 8. System Health Check

NVIDIA provides customers a diagnostics and management tool called NVIDIA System Management, or NVSM. The `nvsm` command can be used to determine the system's health, identify component issues and alerts, or run a stress test to ensure all components are in working order while under load.

The following instructions show how to perform a health check on the DGX GB200 system.

1. Establish an SSH connection to the DGX B200 system.
2. Run a basic system check.

   ```
   sudo nvsm show health
   ```

3. Verify that the output summary shows that all checks are Healthy and that the overall system status is Healthy.

For more information about the `nvsm` command, refer to the NVIDIA System Management User Guide.

Chapter 9. Compliance

The NVIDIA DGX™ H100/H200 Server is compliant with the regulations listed in this section.

9.1. United States

Federal Communications Commission (FCC) FCC Marking (Class A)

This device complies with part 15 of the FCC Rules. Operation is subject to the following two conditions: (1) this device may not cause harmful interference, and (2) this device must accept any interference received, including any interference that may cause undesired operation of the device.

NOTE: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference in which case the user will be required to correct the interference at his own expense.

California Department of Toxic Substances Control: Perchlorate Material - special handling may apply. See www.dtsc.ca.gov/perchlorate.

9.2. United States/Canada

TÜV Rheinland of North America is accredited as a Nationally Recognized Testing Laboratory (NRTL), by OSHA (The Occupational Safety and Health Administration) in the United States, and as a Product Certification Body by SCC (Standards Council of Canada) in Canada. Refer to https://www.tuv.com/usa/en/ctuvus-certification.html

cTUVus Mark

9.3. Canada

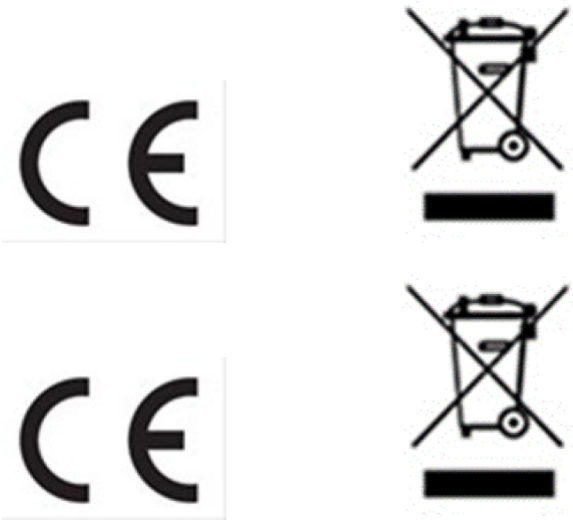
Innovation, Science and Economic Development Canada (ISED) CAN ICES-3(A)/NMB-3(A)

The Class A digital apparatus meets all requirements of the Canadian Interference-Causing Equipment Regulation.

Cet appareil numerique de la class A respecte toutes les exigences du Reglement sur le materiel brouilleur du Canada.

9.4. CE

European Conformity; Conformité Européenne (CE)



This is a Class A product. In a domestic environment this product may cause radio frequency interference in which case the user may be required to take adequate measures.

This device bears the CE mark in accordance with Directive 2014/53/EU. This device complies with the following Directives:

- EMC Directive A, I.T.E Equipment.
- Low Voltage Directive for electrical safety.
- RoHS Directive for hazardous substances.
- Energy-related Products Directive (ErP).

The full text of EU declaration of conformity is available at the following URL: http://www.nvidia.com/support

A copy of the Declaration of Conformity to the essential requirements may be obtained directly from NVIDIA GmbH (Bavaria Towers – Blue Tower, Einsteinstrasse 172, D-81677 Munich, Germany).

9.5. Australia and New Zealand

Australian Communications and Media Authority

This product meets the applicable EMC requirements for Class A, I.T.E equipment.

9.6. Brazil

INMETRO

9.7. Japan

Voluntary Control Council for Interference (VCCI)

この装置は、クラスA機器です。この装置を住宅環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。
VCCI－A

この装置は、クラスA機器です。この装置を住宅環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。
VCCI－A

This is a Class A product.

In a domestic environment this product may cause radio interference, in which case the user may be required to take corrective actions. VCCI-A.

2008年、日本における製品含有表示方法、JISC0950が公示さ れました。製造事業者は、2006年7月１日
以降に販売さ れる電気・・・電子機器の特定化学物質の含有に付きまし て情報提供を義務付けら れまし た。
製品の部材表示に付きまし ては、以下をご覧く ださい。¶

A Japanese regulatory requirement, defined by specification JIS C 0950, 2008, mandates that manufacturers provide Material Content Declarations for certain categories of electronic products offered for sale after July 1, 2006.¶

To view the JIS C 0950 material declaration for this product, visit¶

¶

Japan RoHS Material Content Declaration

日本工業規格JIS C¶
0950:2008により、2006年 7 月 1 日以降に販売される特定分野の電気および電子機器について、製造者による含有物質の表示が義務付けられます。¶

機器名称：リ バ

¶ 主な分類	特定化学物質記号					
	Pb	Hg	Cd	Cr(VI)	PBB	PBDE
筐体	除外項目	0	0	0	0	0
プリント基板	除外項目	0	0	0	0	0
プロセッサー	除外項目	0	0	0	0	0
マザーボード	除外項目	0	0	0	0	0
電源	除外項目	0	0	0	0	0
システムメモリ	除外項目	0	0	0	0	0
ハードディスクドライブ	除外項目	0	0	0	0	0
機械部品（ファン、ヒートシンク、ベゼル¶ ..)	除外項目	0	0	0	0	0
ケーブル/コネクター	除外項目	0	0	0	0	0
はんだ付け材料	0	0	0	0	0	0
フラックス、クリームはんだ、ラベル、その他消耗品	0	0	0	0	0	0

注：¶

1.「0」は、特定化学物質の含有率が日本工業規格 JIS C 0950:2008 に記載されている含有率基準値より低いことを示します。¶

2.「除外項目」は、特定化学物質が含有マークの除外項目に該当するため、特定化学物質について、日本工業規格 JIS C¶
0950:2008 に基づく含有マークの表示が不要であることを示します。¶

¶

3.「0.1wt%超」または「0.01wt%超」は、特定化学物質の含有率が日本工業規格JIS C 0950:2008 に記載されている含有率基準値を超えていることを示します。

A Japanese regulatory requirement, defined by specification JIS C 0950: 2008, mandates that manufacturers provide Material Content Declarations for certain categories of electronic products offered for sale after July 1, 2006.

Product Model Number: P3687 Server

Major Classification	Symbols of Specified Chemical Substance					
	Pb	Hg	Cd	Cr(VI)	PBB	PBDE
Chassis	Exempt	0	0	0	0	0
PCA	Exempt	0	0	0	0	0
Processor	Exempt	0	0	0	0	0
Motherboard	Exempt	0	0	0	0	0
Power supply	Exempt	0	0	0	0	0

System memory	Exempt	0	0	0	0	0
Hard drive	Exempt	0	0	0	0	0
Mechanical parts (fan, heat sink, bezel...)	Exempt	0	0	0	0	0
Cables/Connectors	Exempt	0	0	0	0	0
Soldering material	0	0	0	0	0	0
Flux, Solder Paste, label and other consumable materials	0	0	0	0	0	0

Notes:

1. "0" indicates that the level of the specified chemical substance is less than the threshold level specified in the standard, JIS C 0950:2008.
2. "Exempt" indicates that the specified chemical substance is exempt from marking and it is not required to display the marking for that specified chemical substance per the standard, JIS C 0950: 2008.
3. "Exceeding 0.1wt%" or "Exceeding 0.01wt%" is entered in the table if the level of the specified chemical substance exceeds the threshold level specified in the standard, JIS C 0950: 2008.

9.8. South Korea

Korean Agency for Technology and Standards (KATS)

R-R-WT1-P3687

A급 기기 (업무용 방송통신기자재)	이 기기는 업무용(A급) 전자파적합기기로서 판매자 또는 사용자는 이 점을 주의하시기 바라며, 가정외의 지역에서 사용하는 것을 목적으로 합니다.

Class A Equipment (Industrial Broadcasting & Communication Equipment). This equipment Industrial (Class A) electromagnetic wave suitability equipment and seller or user should take notice of it, and this equipment is to be used in the places except for home.

Korea RoHS Material Content Declaration

확인 및 평가 양식은 제품에 포함 된 유해 물질의 허용 기준의 준수에 관한

문 준비	상호 :	엔비디아홍콩홀딩즈 리미티드(영업소)	법인등록번호	110181-0036373
	대표자성명	카렌테레사번즈	사업자등록번호:	120-84-06711
	주소	서울특별시 강남구 영동대로 511, 2101호 (삼성동,		

제품 내용

제품의 종류	해당없음	제품명(규격)	해당없음
세부모델명(번호):	해당없음	제품출시일	해당없음
제품의 중량	해당없음	제조, 수입업자	엔비디아

엔비디아의 그래픽 카드제품은 전기 전자제품 및 자동차의 자원순환에 관한 법률 시행령 제 11조 제 1항에 의거한 법 시행행규칙 제 3조에에따른 유해물질함유 기준을 확인 및 평가한 결과, 이를 준수하였음을 공표합니다.

구비서류 : 없음

작성방법

① 제품의 종류는 "전기.전자제품 및 자동차의 자원순환에관한 법률 시행령" 제 8조 제 1항 및 제 2항에 따른 품목별로 구분하여 기재합니다.

② 전기 전자 제품의 경우 모델명 (번호), 자동차의 경우, 제원관리번호를 기재합니다.

③ 해당제품의 제조업자 또는 수입업자를 기재합니다.

Confirmation and Evaluation Form Concerning the Adherence to Acceptable Standards of Hazardous Materials Contained in Products

Statement Prepared by	Company Name:	Nvidia HongKong Holding Ltd.Korea branch	Corporate Identification Number:	110181-0036373
	Name of Company Representative:	Karen Theresa Burns	Business Registration Number:	120-84-06711
	Address	2788 San Tomas Expressway, Santa Clara, CA 95051		

Product Information

Product Category:	N/A	Name of Product:	N/A
Detailed Product Model Name (Number):	N/A	Date of first market release:	N/A
Weight of Product:	N/A	Manufacturer and/or Importer:	NVIDIA Corporation

This for is publicly certify That NVIDIA Company has undergone the confirmation and evaluation procedures for the acceptable amounts of hazardous materials contained in graphic card according to the regulations stipulated in Article 3 of the 'Status on the Recycling of Electrical and Electronic Products, and Automobiles' and that company has graphic card adhered to the Enforcement Regulations of Article 11, Item 1 of the statute.

Attachment: None

* Preparing the Form

① Please indicate the product category according to the categories listed in Article 8, Items 1and 2 of the ' Enforcement Ordinance of the Statute on the Recycling of Electrical, Electronic and Automobile Materials'

② For electrical and electronic products, please indicate the Model Name (and number). For automobiles, please indicate the Vehicle Identification Number.

③ Please indicate the name of manufacturer and/or importer of the product.

9.9. China

China Compulsory Certificate

No certification is needed for China. The NVIDIA DGX H100/H200 system is a server with power consumption greater than 1.3 kW.

China RoHS Material Content Declaration

产品中有害物质的名称及含量

The Table of Hazardous Substances and their Content

根据中国《电器电子产品有害物质限制使用管理办法》

as required by China's Management Methods for Restricted of Hazardous Substances Used in Electrical and Electronic Products

部件名称 Parts	有害物质 Hazardous Substances					
	铅 (Pb)	汞 (Hg)	镉 (Cd)	六价铬 (Cr(VI))	多溴联苯 (PBB)	多溴联苯醚 (PBDE)
机箱 Chassis	X	O	O	O	O	O
印刷电路部件 PCA	X	O	O	O	O	O
处理器 Processor	X	O	O	O	O	O
主板 Motherboard	X	O	O	O	O	O
电源设备 Power supply	X	O	O	O	O	O
存储设备 System memory	X	O	O	O	O	O
硬盘驱动器 Hard drive	X	O	O	O	O	O
机械部件（风扇、散热器、面板等) Mechanical parts (fan, heat sink, bezel...)	X	O	O	O	O	O
线材/连接器 Cables/Connectors	X	O	O	O	O	O

焊接金属 Soldering material	O	O	O	O	O	O
助焊剂，锡膏，标签及其他耗材 Flux, Solder Paste, label and other consumable materials	O	O	O	O	O	O

本表格依据SJ/T 11364-2014 的规定编制

The table according to SJ/T 11364-2014

O：表示该有害物质在该部件所有均质材料中的含量均在GB/T 26572-2011 标准规定的限量要求以下。

O: Indicates that this hazardous substance contained in all of the homogeneous materials for this part is below the limit requirement in GB/T 26572-2011.

X：表示该有害物质至少在该部件的某一均质材料中的含量超出GB/T 26572-2011 标准规定的限量要求。

X: Indicates that this hazardous substance contained in at least one of the homogeneous materials used for this part is above the limit requirement in GB/T 26572-2011.

此表中所有名称中含 "X" 的部件均符合欧盟 RoHS 立法。

All parts named in this table with an "X" are in compliance with the European Union's RoHS Legislation.

Note: The referenced Environmental Protection Use Period Marking was determined according to normal operating use conditions of the product such as temperature and humidity.

9.10. Taiwan

Bureau of Standards, Metrology & Inspection (BSMI)

R33088
RoHS

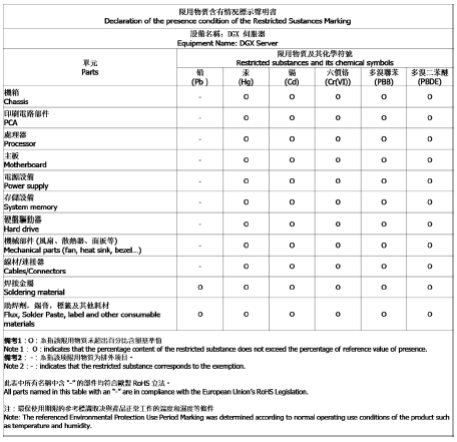
警告使用者:
此為甲類資訊技術設備，於居住環境中使用時，可能會造成射頻擾動，在此種情況下，使用者會被要求採取某些適當的對策

報驗義務人：

香港商輝達香港控股有限公司台灣分公司·・統一編號：80022300

臺北市內湖區基湖路8號.

Taiwan RoHS Material Content Declaration



限用物質含有情況標示聲明書 Declaration of the presence condition of the Restricted Sustances Marking						
設備名稱：DGX 伺服器 Equipment Name: DGX Server						
單元 Parts	限用物質及其化學符號 Restricted substances and its chemical symbols					
	鉛 (Pb)	汞 (Hg)	鎘 (Cd)	六價鉻 (Cr(VI))	多溴聯苯 (PBB)	多溴二苯醚 (PBDE)
機箱 Chassis	-	o	o	o	o	o
印刷電路部件 PCA	-	o	o	o	o	o
處理器 Processor	-	o	o	o	o	o
主板 Motherboard	-	o	o	o	o	o
電源設備 Power supply	-	o	o	o	o	o
存儲設備 System memory	-	o	o	o	o	o
硬盤驅動器 Hard drive	-	o	o	o	o	o
機械部件 (風扇、散熱器、面板等) Mechanical parts (fan, heat sink, bezel...)	-	o	o	o	o	o
線材/連接器 Cables/Connectors	-	o	o	o	o	o
焊接金屬 Soldering material	o	o	o	o	o	o
助焊劑、錫膏、標籤及其他耗材 Flux, Solder Paste, label and other consumable materials	o	o	o	o	o	o

備考1：O：表指該用物質未超出百分比含量基準值
Note 1：O：indicates that the percentage content of the restricted substance does not exceed the percentage of reference value of presence.
備考2：-：表指該限用物質為排外項目。
Note 2：-：indicates that the restricted substance corresponds to the exemption.

此表中所有名稱中含"-"的部件均符合歐盟 RoHS 立法。
All parts named in this table with an "-" are in compliance with the European Union's RoHS Legislation.

注：環保使用期限的參考標識取決於產品正常工作的溫度和濕度等條件
Note: The referenced Environmental Protection Use Period Marking was determined according to normal operating use conditions of the product such as temperature and humidity.

9.11. Russia/Kazakhstan/Belarus

Customs Union Technical Regulations (CU TR)

This device complies with the technical regulations of the Customs Union (CU TR)

ТЕХНИЧЕСКИЙ РЕГЛАМЕНТ ТАМОЖЕННОГО СОЮЗА О безопасности низковольтного оборудования (ТР ТС 004/2011)

ТЕХНИЧЕСКИЙ РЕГЛАМЕНТ ТАМОЖЕННОГО СОЮЗА Электромагнитная совместимость технических средств (ТР ТС 020/2011)

Технический регламент Евразийского экономического союза “Об ограничении применения опасных веществ в изделиях электротехники и радиоэлектроники” (ТР ЕАЭС 037/2016)

Federal Agency of communication (FAC)

This device complies with the rules set forth by Federal Agency of Communications and the Ministry of Communications and Mass Media.

Federal Security Service notification has been filed.

9.12. Israel

SII

ודא שלמות ותקינות כבל החשמל והתקע אין להכניס או להוציא את התקע מרשת החשמל בידיים רטובות . אין לפתוח את המכשיר , במקרה של בעיה כלשהו יש לפנות למעבדת השירות הקרובה. יש להרחיק את המכשיר מנוזלים . במקרה של ריח מוזר, רעשים שמקורם במכשיר , יש לנתקו מיידית מרשת החשמל ולפנות למעבדת שירות המכשיר מיועד לשימוש בתוך המבנה , ולא לשימוש חיצוני ולא לשימוש בסביבה לחה. אין לחתוך, לשבור, ולעקם את הכבל החשמל. אין להניח חפצים על הכבל החשמל או להניח לו להתחמם יתר על המידה , שכן עלול לגרום לנזק ,דליקה או התחשמלות . יש להקפיד לחזק את התקן הניתוק במצב תפעולי מוכן לשימוש. אזהרה: אין להחליף את כבל הזינה בתחליפים לא מקוריים, חיבור לקוי עלול לגרום להתחשמלות המשתמש. בשימוש על כבל מאריך יש לוודא תקינות מוליך הארקה שבכבל .

9.13. India

Bureau of India Standards (BIS)

Authenticity may be verified by visiting the Bureau of Indian Standards website at http://www.bis.gov.in.

India RoHS Compliance Statement

This product, as well as its related consumables and spares, complies with the reduction in hazardous substances provisions of the “India E-waste (Management and Handling) Rule 2016”. It does not contain lead, mercury, hexavalent chromium, polybrominated biphenyls or polybrominated diphenyl ethers in concentrations exceeding 0.1 weight % and 0.01 weight % for cadmium, except for where allowed pursuant to the exemptions set in Schedule 2 of the Rule.

9.14. South Africa

South African Bureau of Standards (SABS)

This device complies with the following SABS Standards:

SANS 2332: 2017/CISPR 32:2015 SANS 2335:2018/ CISPR 35:2016

National Regulator of Compulsory Specification (NRCS)

This device complies with following standard under VC 8055:

SANS IEC 60950-1

9.15. Great Britain (England, Wales, and Scotland)

UK Conformity Assessed

This device complies with the following Regulations:

- SI 2016/1091: Electromagnetic Compatibility (EMC)
- SI 2016/1101: The Low Voltage Electrical Equipment (Safety)
- SI 2012/3032: The Restriction of the Use of Certain Hazardous Substances in Electrical and Electronic Equipment (As Amended)

A copy of the Declaration of Conformity to the essential requirements may be obtained directly from NVIDIA Ltd. (100 Brook Drive, 3rd Floor Green Park, Reading RG2 6UJ, United Kingdom)

Chapter 10. Third-Party License Notices

This NVIDIA product contains third party software that is being made available to you under their respective open source software licenses. Some of those licenses also require specific legal information to be included in the product. This section provides such information.

10.1. Micron msecli

The msecli utility is provided under the following terms:

Micron Technology, Inc. Software License Agreement PLEASE READ THIS LICENSE AGREEMENT ("AGREEMENT") FROM MICRON TECHNOLOGY, INC. ("MTI") CAREFULLY: BY INSTALLING, COPYING OR OTHERWISE USING THIS SOFTWARE AND ANY RELATED PRINTED MATERIALS ("SOFTWARE"), YOU ARE ACCEPTING AND AGREEING TO THE TERMS OF THIS AGREEMENT. IF YOU DO NOT AGREE WITH THE TERMS OF THIS AGREEMENT, DO NOT INSTALL THE SOFTWARE. LICENSE: MTI hereby grants to you the following rights: You may use and make one (1) backup copy the Software subject to the terms of this Agreement. You must maintain all copyright notices on all copies of the Software. You agree not to modify, adapt, decompile, reverse engineer, disassemble, or otherwise translate the Software. MTI may make changes to the Software at any time without notice to you. In addition MTI is under no obligation whatsoever to update, maintain, or provide new versions or other support for the Software. OWNERSHIP OF MATERIALS: You acknowledge and agree that the Software is proprietary property of MTI (and/or its licensors) and is protected by United States copyright law and international treaty provisions. Except as expressly provided herein, MTI does not grant any express or implied right to you under any patents, copyrights, trademarks, or trade secret information. You further acknowledge and agree that all right, title, and interest in and to the Software, including associated proprietary rights, are and shall remain with MTI (and/or its licensors). This Agreement does not convey to you an interest in or to the Software, but only a limited right to use and copy the Software in accordance with the terms of this Agreement. The Software is licensed to you and not sold.

DISCLAIMER OF WARRANTY:

THE SOFTWARE IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. MTI EXPRESSLY DISCLAIMS ALL WARRANTIES EXPRESS ORmIMPLIED, INCLUDING BUT NOT LIMITED TO, NONINFRINGEMENT OF THIRD PARTY RIGHTS, AND ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. MTI DOES NOT WARRANT THAT THE SOFTWARE WILL MEET YOUR REQUIREMENTS, OR THAT THE OPERATION OF THE SOFTWARE WILL BE UNINTERRUPTED OR ERROR-FREE. FURTHERMORE, MTI DOES NOT MAKE ANY REPRESENTATIONS REGARDING THE USE OR THE RESULTS OF THE USE OF THE SOFTWARE IN TERMS OF ITS CORRECTNESS, ACCURACY, RELIABILITY, OR OTHERWISE. THE ENTIRE RISK ARISING OUT OF USE OR PERFORMANCE OF THE SOFTWARE REMAINS WITH YOU. IN NO EVENT SHALL MTI, ITS AFFILIATED COMPANIES OR THEIR SUPPLIERS BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL, INCIDENTAL, OR SPECIAL DAMAGES (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF PROFITS, BUSINESS INTERRUPTION, OR LOSS OF INFORMATION) ARISING OUT OF YOUR USE OF OR INABILITY TO USE THE SOFTWARE, EVEN IF MTI

HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Because some jurisdictions prohibit the exclusion or limitation of liability for consequential or incidental damages, the above limitation may not apply to you.

TERMINATION OF THIS LICENSE: MTI may terminate this license at any time if you are in breach of any of the terms of this Agreement. Upon termination, you will immediately destroy all copies the Software.

GENERAL: This Agreement constitutes the entire agreement between MTI and you regarding the subject matter hereof and supersedes all previous oral or written communications between the parties. This Agreement shall be governed by the laws of the State of Idaho without regard to its conflict of laws rules.

CONTACT: If you have any questions about the terms of this Agreement, please contact MTI's legal department at (208) 368-4500. By proceeding with the installation of the Software, you agree to the terms of this Agreement. You must agree to the terms in order to install and use the Software.

10.2. Mellanox (OFED)

MLNX_OFED <http://www.mellanox.com/> is provided under the following terms:

Copyright (c) 2006 Mellanox Technologies. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Chapter 11. Notices

11.1. Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation (“NVIDIA”) makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), 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.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer (“Terms of Sale”). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer’s own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer’s product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. 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 under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, “MATERIALS”) ARE BEING PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA’s aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

11.2. Trademarks

NVIDIA, the NVIDIA logo, DGX, DGX-1, DGX-2, DGX A100, DGX H100, DGX H200, DGX Station, and DGX Station A100 are trademarks and/or registered trademarks of NVIDIA Corporation in the Unites States and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright