QM9700/QM9790 1U NDR 400Gb/s InfiniBand Switch Systems User Manual

Exported on Oct/24/2022 09:50 AM

Table of Contents

Relevant for Models: QM9700 and QM9790

This manual describes the installation and basic use of the NVIDIA 1U NDR InfiniBand switch systems based on the NVIDIA Quantum™-2 switch ASIC. This manual is intended for IT managers and system administrators.

Ordering Information

System Model	Part Number	Legacy Part Number	Description
QM9700	920-9B210-00FN-0M0	MQM9700-NS2F	64-ports NDR, 32 OSFP ports, managed, power-to-connector (P2C) airflow (forward)
	920-9B210-00RN-0M2	MQM9700-NS2R	64-ports NDR, 32 OSFP ports, managed, connector-to-power (C2P) airflow (reverse)
QM9790	920-9B210-00FN-0D0	MQM9790-NS2F	64-ports NDR, 32 OSFP ports, unmanaged, P2C airflow (forward)
	920-9B210-00RN-0D0	MQM9790-NS2R	64-ports NDR, 32 OSFP ports, unmanaged, C2P airflow (reverse)

Related Documentation

Document	Description
InfiniBand Architecture Specification Volume 1 Release 1.5	The InfiniBand Trade Association (IBTA) InfiniBand® Specification at https://www.infinibandta.org.
MLNX-OS® User Manual	This document contains information regarding the configuration and management of the MLNX-OS® software. https://www.nvidia.com/en-us/networking/ethernet/switch-software/.
Hands-on workshops	Visit https://academy.nvidia.com/en/infiniband-customized-training/.
On-site/remote services	For any tailor-made service, contact: nbu-services-sales@nvidia.com.

Revision History

A list of the changes made to this document are provided in Document Revision History.

Introduction

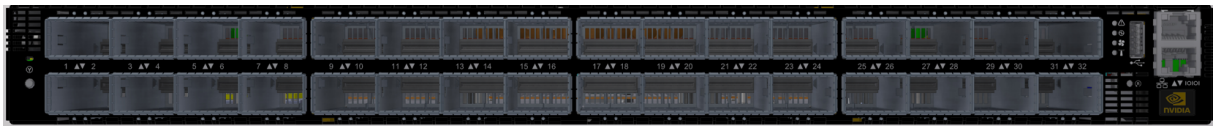
The NVIDIA Quantum-2-based QM9700 and QM9790 switch systems deliver an unprecedented 64 ports of NDR 400Gb/s InfiniBand per port in a 1U standard chassis design. A single switch carries an aggregated bidirectional throughput of 51.2 terabits per second (Tb/s), with a landmark of more than 66.5 billion packets per second (BPPS) capacity. Supporting the latest NDR technology, NVIDIA Quantum-2 brings a high-speed, extremely low-latency and scalable solution that incorporates state-of-the-art technologies such as Remote Direct Memory Access (RDMA), adaptive routing, and NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™.Unlike any other networking solution, NVIDIA InfiniBand provides self-healing network capabilities, as well as quality of service (QoS), enhanced virtual lane (VL) mapping, and congestion control to provide the highest overall application throughput. As an ideal rack-mounted InfiniBand solution, the QM9700 and QM9790 NDR InfiniBand fixed-configuration switches allow maximum flexibility, as they enable a variety of topologies, including Fat Tree, SlimFly, DragonFly+, multi-dimensional Torus, and more. They're also backwards compatible to previous generations and include expansive software ecosystem support.

Today's complex research demands ultra-fast processing of high-resolution simulations, extreme-size datasets, and complex, highly parallelized algorithms that need to exchange information in real time. The QM9700 NDR InfiniBand switches extend NVIDIA In-Network Computing technologies and introduce the third generation of NVIDIA SHARP technology, SHARPv3. Creating virtually unlimited scalability for large data aggregation through the data center network, participating in the application's runtime and reducing the amount of data needed to traverse the network.

By implementing NVIDIA port-split technology, the QM9700 and QM9790 switches provide a double-density radix for 200Gb/s (NDR200) data speeds, reducing the cost of network design and network topologies. Supporting up to 128 ports of 200Gb/s, NVIDIA delivers the densest top-of-rack (TOR) switch available on the market. The QM9700 family of switches enables small to medium-sized deployments to scale with a two-level Fat Tree topology while reducing power, latency, and space requirements.

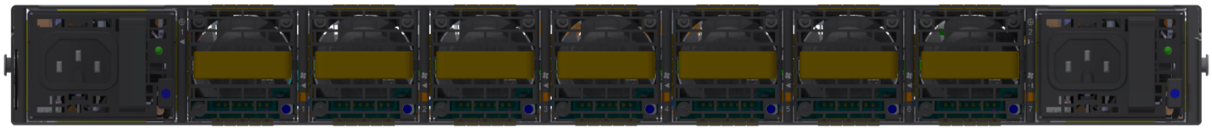
The internally managed QM9700 switch features an on-board subnet manager that enables simple, out-of-the-box bringup for up to 2,000 nodes. Running the NVIDIA MLNX-OS® software package, the subnet manager delivers full chassis management through command-line interface (CLI), web-based user (WebUI), Simple Network Management Protocol (SNMP), or JavaScript Object Notation (JSON) interfaces.The externally managed QM9790 switch can utilize the advanced NVIDIA Unified Fabric Manager (UFM®) feature sets to empower data center operators to efficiently provision, monitor, manage, preventatively troubleshoot, and maintain the modern data center fabric, to realize higher utilization and reduce overall opex.

QM9700 Front View



QM9790 Front View

QM9700 and QM9790 Rear View



For additional airflow options, see Airflow.

Speed and Switching Capabilities

The table below describes maximum throughput and interface speed per system model.

64 NDR non-blocking ports with aggregate data throughput up to 51.2Tb/s

System Model	NDR 400Gb/s OSFP Interfaces	Max Throughput
QM9700	32	51.2Tb/s
QM9790	32	51.2Tb/s

Management Interfaces, PSUs and Fans

The table below lists the various management interfaces and available replacement parts per system model.

System Model	USB	MGT	I²C	Console	Replaceable PSU	Replaceable Fan
QM9700	Front (USB3.0 type A)	Front (1 port)	NA	Front	Yes	Yes
QM9790	NA	NA	Front (USB3.0 type A)	NA	Yes	Yes

Features

For a full feature list, please refer to the system's product brief. Go to https://www.nvidia.com/en-us/networking/. In the main menu, click on Products > InfiniBand > Switch Systems, and select the desired product page.

Certifications

The list of certifications (such as EMC, Safety and others) per system for different regions of the world is located on the Mellanox website at http://www.mellanox.com/page/environmental_compliance.

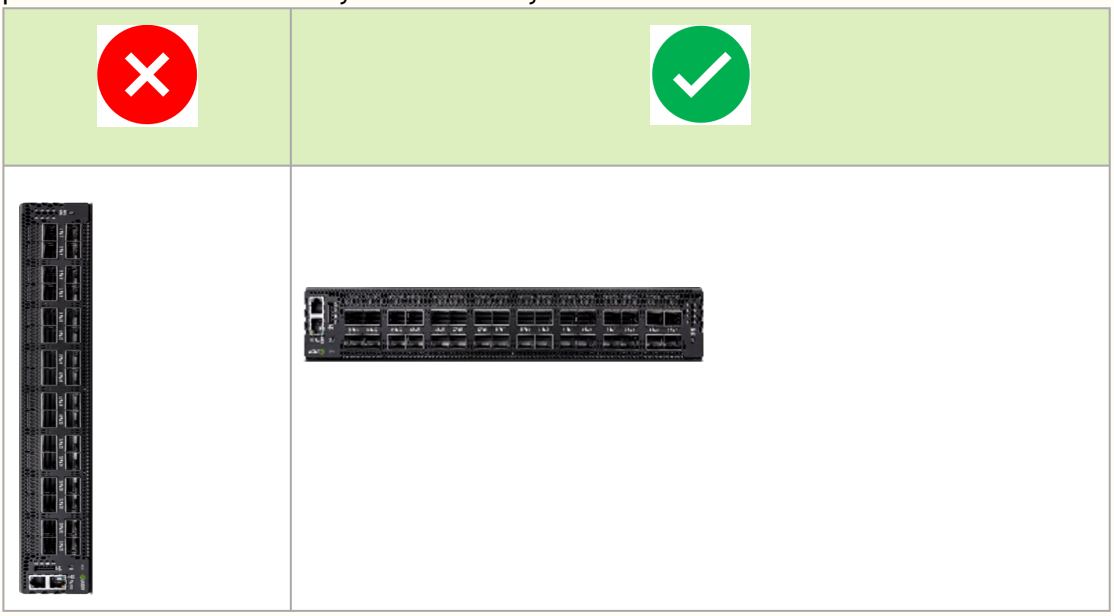
Installation

System Installation and Initialization

Installation and initialization of the system require attention to the normal mechanical, power, and thermal precautions for rack-mounted equipment.

The rack mounting holes conform to the EIA-310 standard for 19-inch racks. Take precautions to guarantee proper ventilation in order to maintain good airflow at ambient temperature.

Due to thermal considerations, the switch systems must be installed in a horizontal position. do not install the systems vertically.



- Unless otherwise specified, NVIDIA products are designed to work in an environmentally controlled data center with low levels of gaseous and dust (particulate) contamination.
- The operation environment should meet severity level G1 as per ISA 71.04 for gaseous contamination and ISO 14644-1 class 8 for cleanliness level.

The installation procedure for the system involves the following phases:

Step	Procedure	See
1	Follow the safety warnings	Safety Warnings

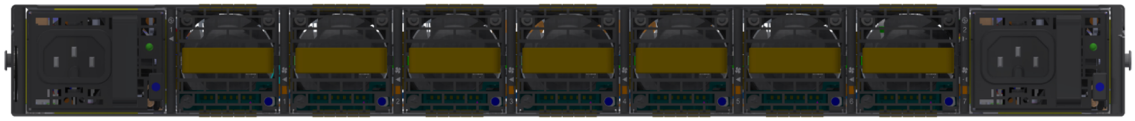
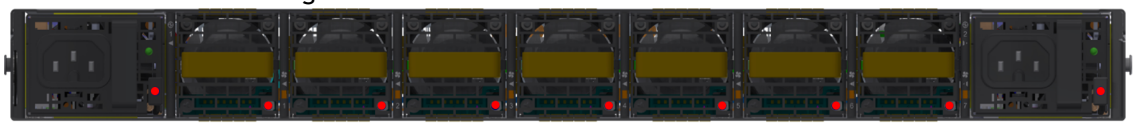
Step	Procedure	See
2	Pay attention to the air flow consideration within the system and rack	Air Flow
3	Make sure that none of the package contents is missing or damaged	Package Contents
4	Mount the system into a rack enclosure	19" System Mounting Options
5	Power on the system	Initial Power On
6	Perform system bring-up	System Bring-Up of Managed Systems
7	[Optional] FRU replacements	FRU Replacements

Safety Warnings

Prior to the installation, please review the Safety Warnings. Note that some warnings may not apply to all models.

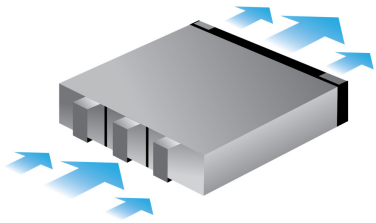
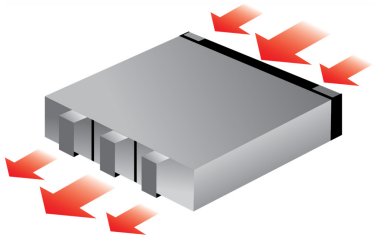
Air Flow

NVIDIA systems are offered with two air flow patterns:

- Power (rear) side inlet to connector side outlet - marked with blue dots that are placed on the power inlet side.
  Air Flow Direction Marking - Power Side Inlet to Connector Side Outlet
  
- Connector (front) side inlet to power side outlet - marked with red dots that are placed on the power inlet side.
  Air Flow Direction Marking - Connector Side Inlet to Power Side Outlet
  

- All servers and systems in the same rack should be planned with the same airflow direction.
- All FRU components need to have the same air flow direction. A mismatch in the air flow will affect the heat dissipation.

The table below provides an air flow color legend and respective OPN designation.

Direction	Description and OPN Designation
	Power side inlet to connector side outlet. Blue indicators are placed on the power inlet side. OPN designation is “-F”.
	Connector side inlet to power side outlet. Red indicators are placed on the power inlet side. OPN designation is “-R”.

Package Contents

Before installing your new system, unpack it and check against the parts list below that all the parts have been sent. Check the parts for visible damage that may have occurred during shipping.

The QM9700 and QM9790 package content is as follows:

- 1 – System
- 1 – Rail kit
- 4 – Power cables Type C14 to C15
- 1 - Harness: HAR000631 – Harness RS232 2M cable – DB9 to RJ-45 (only in QM9700)
- 2 – Cable retainers
- 32 - OSFP thermal caps

If anything is damaged or missing, contact your sales representative at Networking-support@nvidia.com.

19” System Mounting Options

By default, the systems are shipped with the rail kit described in Tool-Less Rail Kit.

Tool-Less Rail Kit

Kit Part Number	Legacy Kit Part Number	Rack Size and Rack Depth Range
930-9BRKT-00JM-000	MTEF-KIT-I-TL	600-800 mm

Prior to the installation procedure, inspect all rail-kit components and make sure none of them is missing or damaged. If anything is missing or damaged, contact your NVIDIA representative at Networking-support@nvidia.com.

The following parts are included in the tool-less rail kit (see figure below):

- 2x System Rails (A)
- 2x Rack Rails (B)

Rail Kit Parts

A	B
	

Prerequisites:

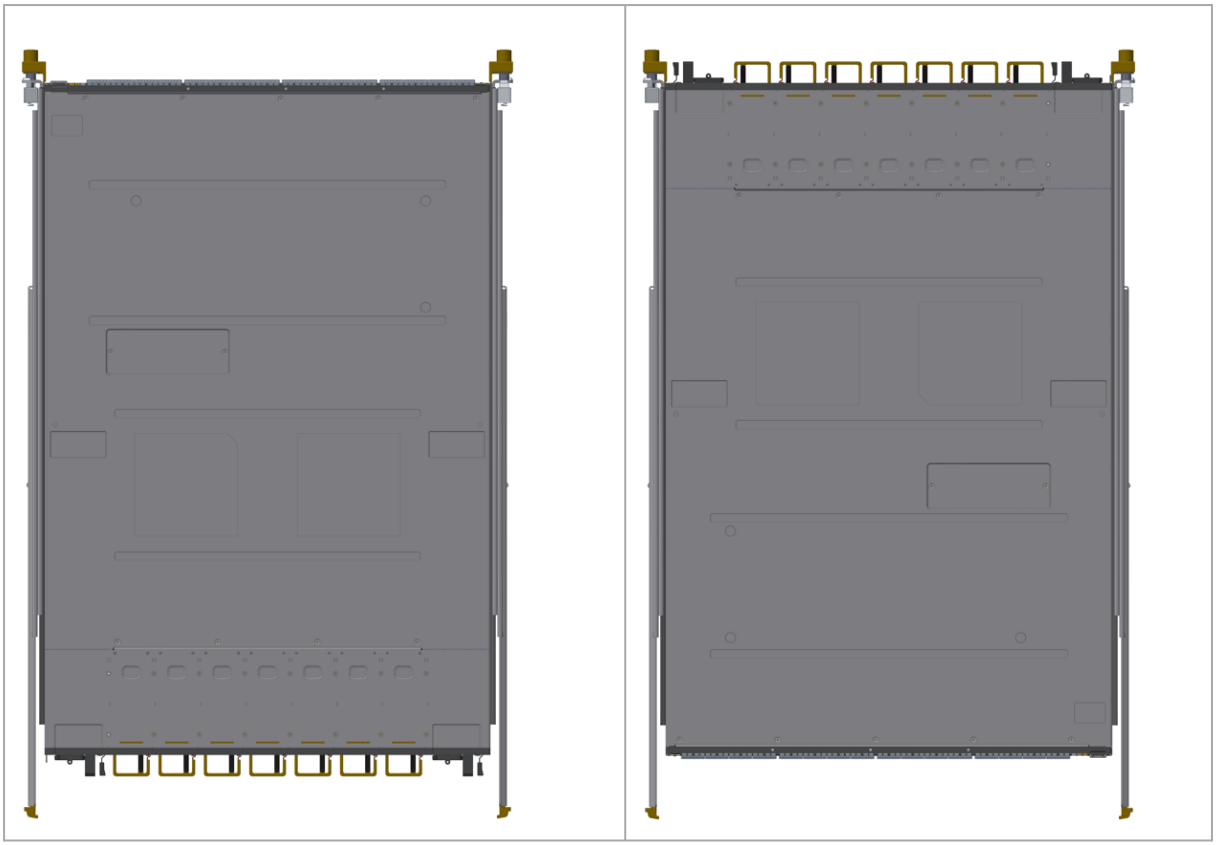
Before mounting the system to the rack, select the way you wish to place the system. Pay attention to the airflow within the rack cooling, connector and cabling options.

While planning how to place the system, review the two installation options shown in the table below, and consider the following points:

- Make sure the system air flow is compatible with your installation selection. It is important to keep the airflow within the rack in the same direction.
- Note that the part of the system to which you choose to attach the rails (the front panel direction, as demonstrated in Option 1 or the FRUs direction, as demonstrated in Option 2) will determine the system's adjustable side. The system's part to which the brackets are attached will be adjacent to the cabinet.
- The FRUs, as well as high-speed and MNG cables, must be extracted for replacement as part of the switch service. Consider this when planning the switch installation.

Switch Rails Installation - Top View

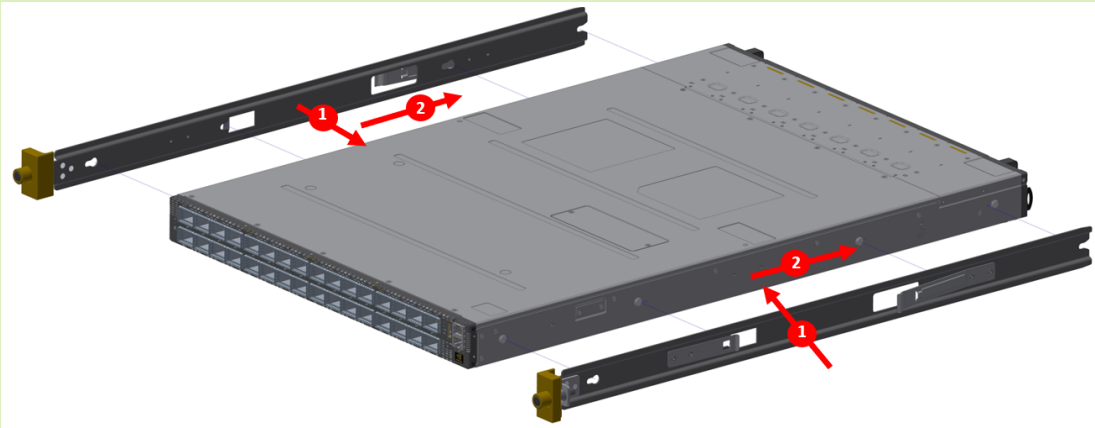
Front Side (Ports)	Rear Side (FRUs)



⚠ The following steps include illustrations that show front side (ports) installation, yet all instructions apply to all installation options.

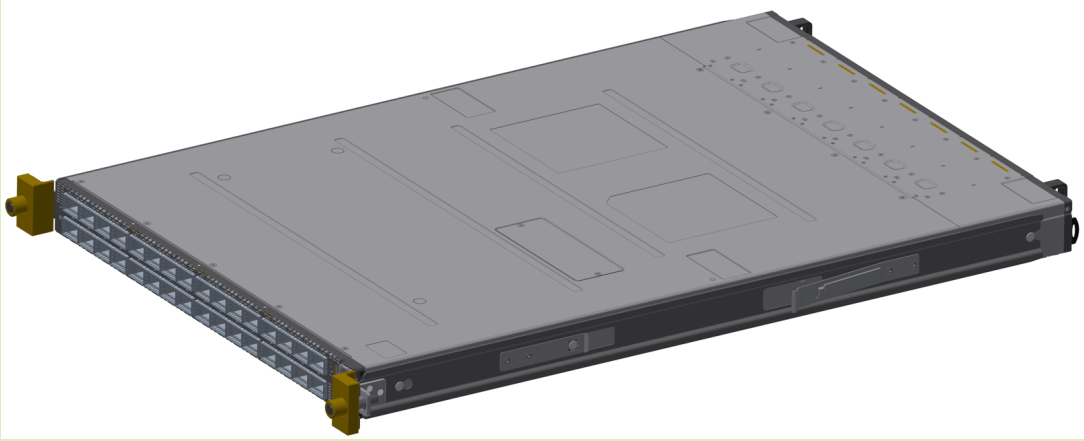
1. Attach the left and right system rails (A) to the switch.

Attaching the System Rails (A) to the Switch



2. Secure the assembly by gently pushing the system chassis' pins through the slider key holes, until locking occurs

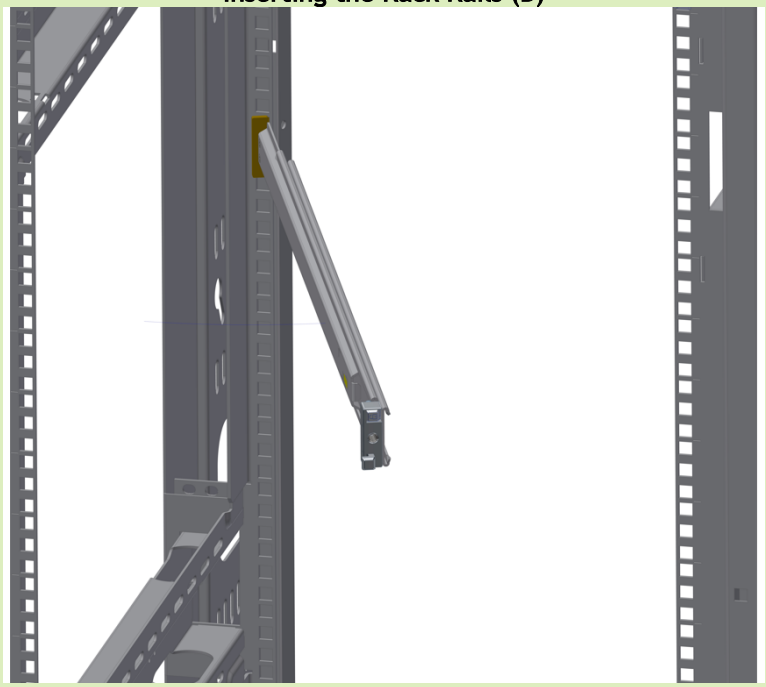
Securing the System in the Switch Rails (A)



Chassis' Pins in the Rails' Slots Locking them in a Fixed Position

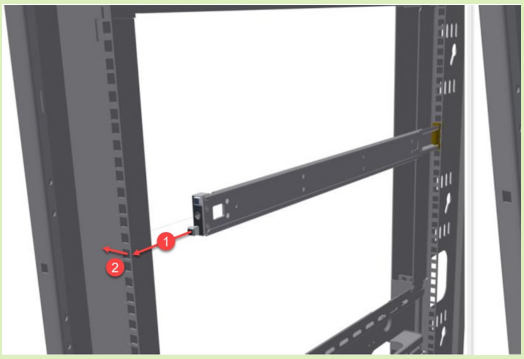
3. Mount both of the rack rails (B) into the rack by angularly inserting the brakes located at the rails edges into the designated slots in the rack unit, as shown in the following figure:

Inserting the Rack Rails (B)

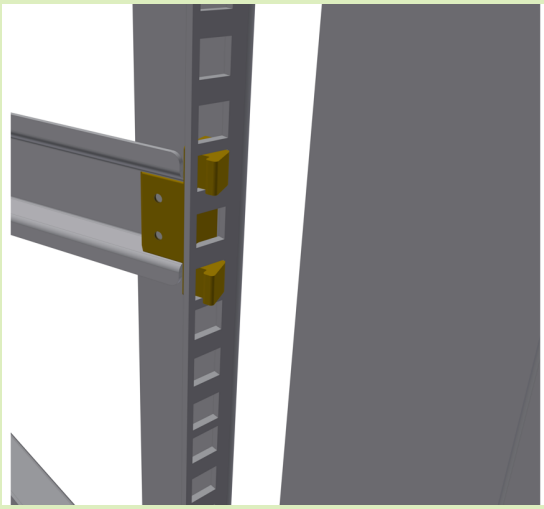


4. Align both rack rails (B) to sit horizontally in parallel to the rack assembly. By straightening the rails' angular position, their breaks will be caught and locked in the rack's slots.

Aligning the Rack Rails (B) Angular Position



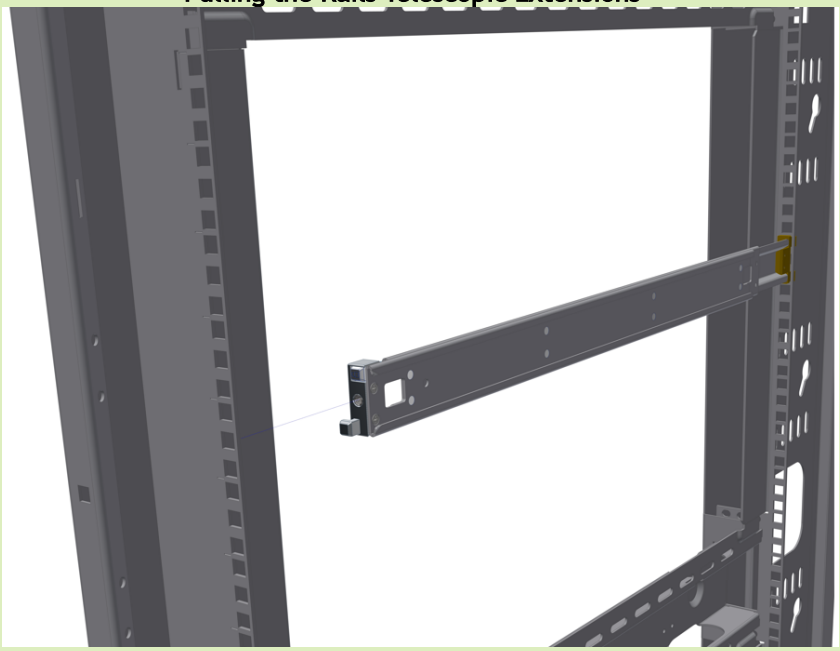
The Breaks are Caught and Locked in the Rack's Slots



Rack Rails Fully Inserted and Locked in the Rack Assembly

5. Pull the rack rails' telescopic extensions all the way to the rack's opposite side, and insert the latches at the rails' free edges to the rack's slots. A click should be heard as the spring latches are fully inserted and locking occurs.

Pulling the Rails Telescopic Extensions



Inserting the Spring Latches to the Rack's Slots

To mount the system into the rack:

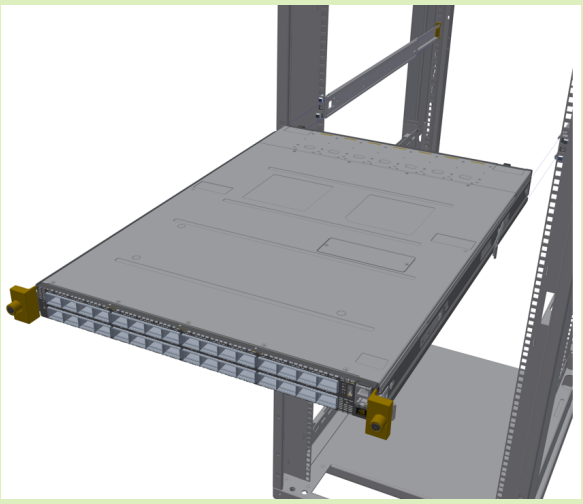
> At least two people are required to safely mount the system in the rack.

While your installation partner is supporting the system's weight, perform the following steps:

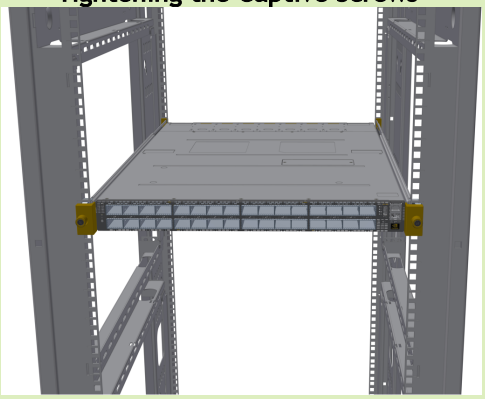
6. Slide the rails installed on the system into the channels in the rack rails. Push them forward until the locking mechanism is activated on both sides, and a click is heard.
7. Tighten the captive screws on both sides to further secure the system to the rack's posts.

Sliding the System's Rails (A) into the Rack Rails (B)

Sliding the System Rails (A) into the Rack Rails (B)



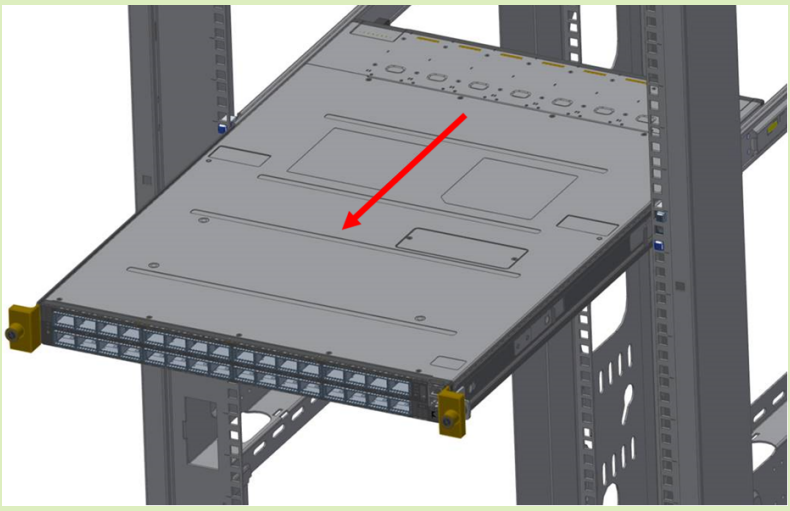
Tightening the Captive Screws



To remove the system from the rack:

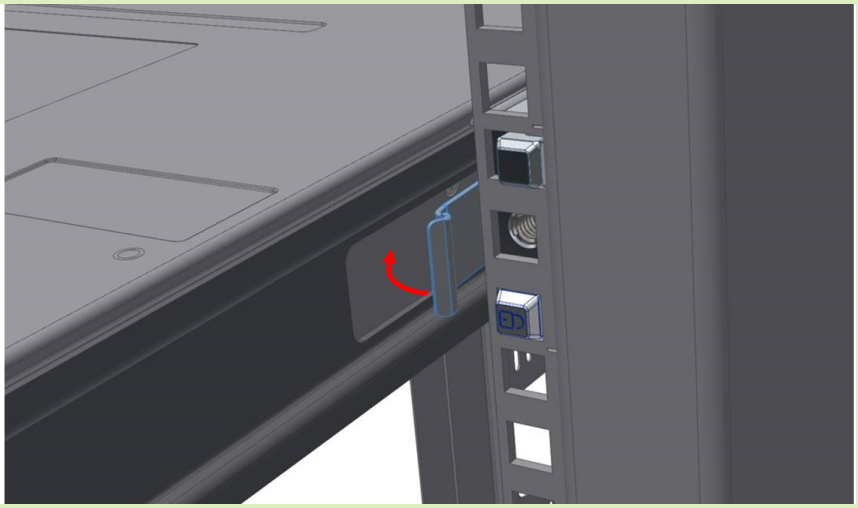
1. Turn off the system and disconnect it from peripherals and from the electrical outlet. While your installation partner is supporting the system’s weight:
2. Loosen the captive screws attaching the system's rails to the rack's posts.
3. Use two hands to pull the system out until the rails are stopped.

Pulling the System Out



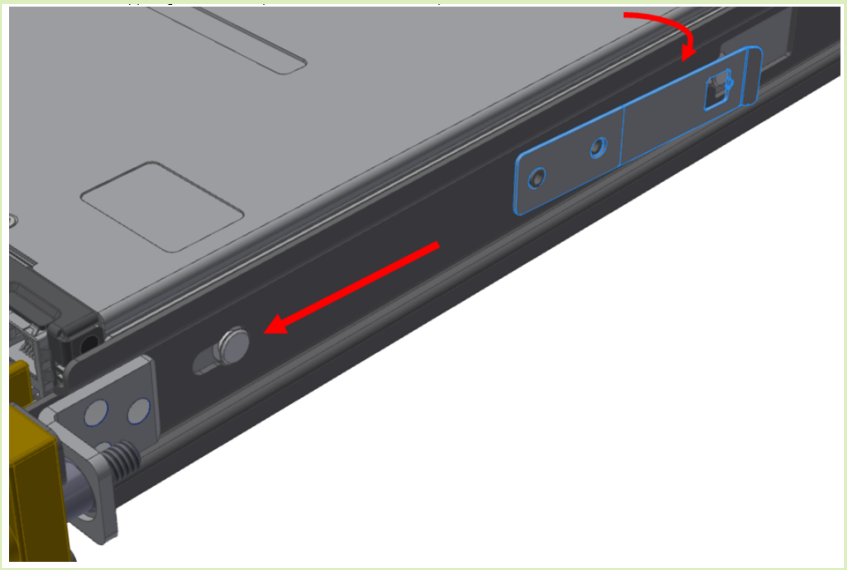
4. Press the spring latches on both sides of the rack, and continue to pull the system out until the rack rails are clear of the system's rails.

Pressing the Spring Latches on Both Sides



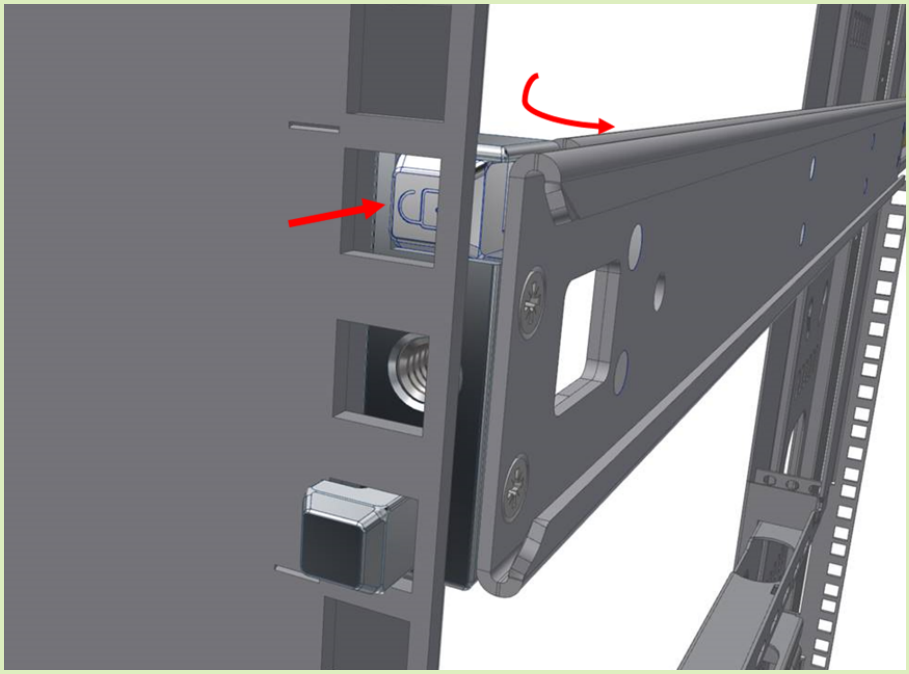
5. Remove the rails from the system. Release the metal latches and pull out the rails, so the system's pins will be removed out of the oval slots.

Removing the Rails from the System



6. Remove the rails from the rack by pressing the lock button, and pull the rails outside of the rack assembly.

Pressing the Lock Button to Remove the Rails from the Rack



Cable Installation

Power Cable and Cable Retainer

In some switch models, the product's package includes cable retainers. It is highly recommended to use them in order to secure the power cables in place.

When installing retainers for the PSUs of the QM97x0 switch systems, please adhere to the following instructions:

1. Verify the integrity of the retainer assembly, as demonstrated in the below table:- The snaps' push-pins must have visible edges with no broken or torn parts.
   - The shoulders' pins should be in-tact and must not be bent inwards.

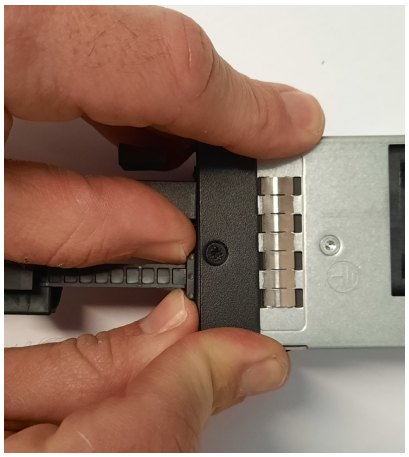
Proper Condition	Improper Condition
	

2. It is advised to place the PSU on a flat, stable surface. While you secure the PSU in place, use two thumbs to insert the retainer's two snaps into the designated holes located near the AC inlet. Make sure that the retainer's plastic loop is facing upwards, as demonstrated in the below table.

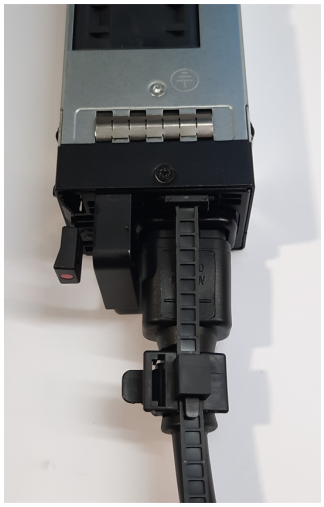
> ⚠ For demonstration purposes, the images in this document show C2P (Connector-to-Power) airflow PSUs with red latches, yet the instructions apply to P2C (Power-to-Connector) PSUs with blue latches as well.

Correct Insertion	Incorrect Insertion
	

3. Push the retainer until the shoulders' pins (in blue circles below) are open and aligned with the PSU front panel, as shown in the following table:

Fully Mated Retainer	
	

4. Make sure that the retainer is fully locked in place by gently attempting to pull it outwards.
5. Open the plastic loop and route the AC cord through it. Locate the loop over the AC cord, as shown in the following table, and fasten it tightly.

Proper Loop Placement	Improper Loop Placement
	

Each cable retainer can be used once only. Once the retainer has been fully inserted and the shoulders' pins have been adjusted, the retainer cannot be used again, and should be discarded if pulled out.

Port Cables

All cables can be inserted or removed with the unit powered on.

To insert a cable, press the connector into the port receptacle until the connector is firmly seated. The LED indicator, corresponding to each data port, will light when the physical connection is established. When a logical connection is made, the relevant port LED will turn on.

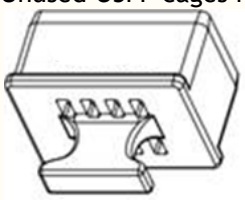
To remove a cable, disengage the locks and slowly pull the connector away from the port receptacle. The LED indicator for that port will turn off when the cable is unseated.

For full cabling guidelines, ask your NVIDIA representative for a copy of *NVIDIA Cable Management Guidelines and FAQs Application Note*.

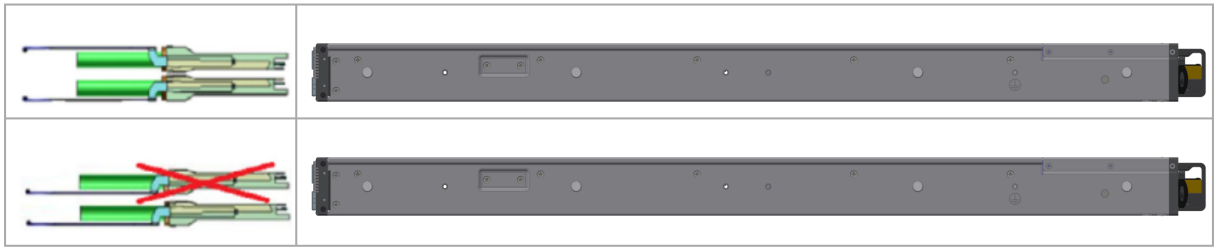
For more information about port LEDs, refer to Port LEDs.

Do not force the cable into the cage with more than 40 newtons/9.0 pounds/4kg force. Greater insertion force may cause damage to the cable or to the cage.

Unused OSFP cages must be closed with the thermal caps supplied with the system.



Cable Orientation



Splitter (Breakout) Cables and Adapters

The breakout option is intended for users planning to run HDR100 using ConnectX-6 only.

The breakout cable is a unique capability, where a single physical quad-lane OSFP port is divided into 2 dual-lane ports. It maximizes flexibility by enabling end users to use a combination of dual-lane and quad-lane interfaces according to the specific requirements of their network. In the QM97x0 systems, each connector contains two ports, and all system ports may be split into 2-lane ports. For the systems splitting options, see QM9700/QM9790 Splitting Options below.

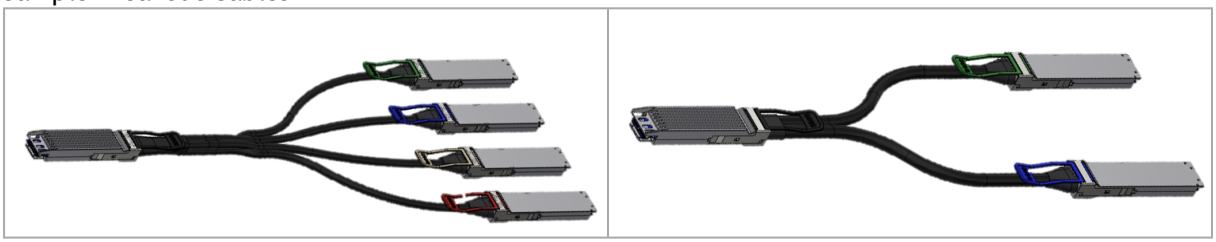
Splitting a port changes the notation of that port from x/y to x/y/z with “x/y” indicating the previous notation of the port prior to the split and “z” indicating the number of the resulting single-lane port (1,2). Each sub-physical port is then handled as an individual port. For example, splitting

port 5 into 2 lanes results in ports 1/5/1 and 1/5/2, and splitting it into 4 lanes results in ports 1/5/1/1, 1/5/1/2, 1/5/2/1 and 1/5/2/2. For full notation schematics, see Port Notation Schematics.

The following behavior should be expected when disconnecting a 1:2 splitter cable (from cages in both the upper and lower rows):

- When you disconnect a cable marked as “1”, the CLI <cage number>/1 will always go down, and the left LED of the cage will be turned off.
- When you disconnect the cable marked as “2”, the CLI <cage number>/2 will always go down, and the right LED of the cage will be turned off.

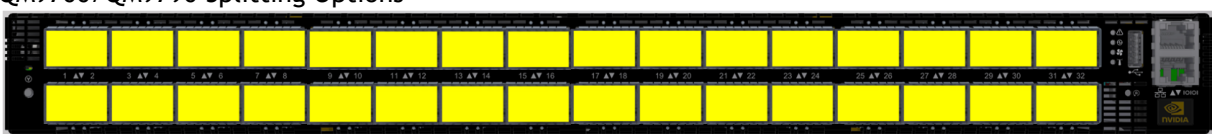
Sample Breakout Cables



- Splitting the interface deletes all configuration on that interface.
- This feature is available only for Quantum ASIC systems.
- In order to be able to use this feature, the system profile command must be activated with split-ready configuration (cross-reference to system profile command).

For more information on how to change the system’s profile to allow Split-Ready configuration, how to change the module type to a split mode, and how to unsplit a split port, please refer to the "InfiniBand Switching" chapter in the latest *MLNX-OS® User Manual*.

QM9700/QM9790 Splitting Options



Split x2 All OSFP ports are splittable. Each OSFP cage contains two ports of 400G, and each port can be split to two. There are no blocking requirements.

Port Notation Schematics

Two port notation profiles can be selected for the QM97x0 NDR switch systems. In both cases, each cage in the system's front panel holds two ports from the same ASIC, and the cage numbers are global:

1. ASIC/Cage/Port:

1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
1/1/1, 1/1/2	1/3/1, 1/3/2	1/5/1, 1/5/2	1/7/1, 1/7/2	1/9/1, 1/9/2	1/11/1, 1/11/2	1/13/1, 1/13/2	1/15/1, 1/15/2	1/17/1, 1/17/2	1/19/1, 1/19/2	1/21/1, 1/21/2	1/23/1, 1/23/2	1/25/1, 1/25/2	1/27/1, 1/27/2	1/29/1, 1/29/2	1/31/1, 1/31/2
1/2/1, 1/2/2	1/4/1, 1/4/2	1/6/1, 1/6/2	1/8/1, 1/8/2	1/10/1, 1/10/2	1/12/1, 1/12/2	1/14/1, 1/14/2	1/16/1, 1/16/2	1/18/1, 1/18/2	1/20/1, 1/20/2	1/22/1, 1/22/2	1/24/1, 1/24/2	1/26/1, 1/26/2	1/28/1, 1/28/2	1/30/1, 1/30/2	1/32/1, 1/32/2
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32

2. ASIC/Cage/Port/Split:

1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
1/1/1/1, 1/1/1/2 1/1/2/1, 1/1/2/2	1/3/1/1, 1/3/1/2 1/3/2/1, 1/3/2/2	1/5/1/1, 1/5/1/2 1/5/2/1, 1/5/2/2	1/7/1/1, 1/7/1/2 1/7/2/1, 1/7/2/2	1/9/1/1, 1/9/1/2 1/9/2/1, 1/9/2/2	1/11/1/1, 1/11/1/2 1/11/2/1, 1/11/2/2	1/13/1/1, 1/13/1/2 1/13/2/1, 1/13/2/2	1/15/1/1, 1/15/1/2 1/15/2/1, 1/15/2/2	1/17/1/1, 1/17/1/2 1/17/2/1, 1/17/2/2	1/19/1/1, 1/19/1/2 1/19/2/1, 1/19/2/2	1/21/1/1, 1/21/1/2 1/21/2/1, 1/21/2/2	1/23/1/1, 1/23/1/2 1/23/2/1, 1/23/2/2	1/25/1/1, 1/25/1/2 1/25/2/1, 1/25/2/2	1/27/1/1, 1/27/1/2 1/27/2/1, 1/27/2/2	1/29/1/1, 1/29/1/2 1/29/2/1, 1/29/2/2	1/31/1/1, 1/31/1/2 1/31/2/1, 1/31/2/2
1/2/1/1, 1/2/1/2 1/2/2/1, 1/2/2/2	1/4/1/1, 1/4/1/2 1/4/2/1, 1/4/2/2	1/6/1/1, 1/6/1/2 1/6/2/1, 1/6/2/2	1/8/1/1, 1/8/1/2 1/8/2/1, 1/8/2/2	1/10/1/1, 1/10/1/2 1/10/2/1, 1/10/2/2	1/12/1/1, 1/12/1/2 1/12/2/1, 1/12/2/2	1/14/1/1, 1/14/1/2 1/14/2/1, 1/14/2/2	1/16/1/1, 1/16/1/2 1/16/2/1, 1/16/2/2	1/18/1/1, 1/18/1/2 1/18/2/1, 1/18/2/2	1/20/1/1, 1/20/1/2 1/20/2/1, 1/20/2/2	1/22/1/1, 1/22/1/2 1/22/2/1, 1/22/2/2	1/24/1/1, 1/24/1/2 1/24/2/1, 1/24/2/2	1/26/1/1, 1/26/1/2 1/26/2/1, 1/26/2/2	1/28/1/1, 1/28/1/2 1/28/2/1, 1/28/2/2	1/30/1/1, 1/30/1/2 1/30/2/1, 1/30/2/2	1/32/1/1, 1/32/1/2 1/32/2/1, 1/32/2/2
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32

Initial Power On

Each system's input voltage is specified in the Specifications chapter.

The power cords should be standard 3-wire AC power cords including a safety ground and rated for 15A or higher.

> The system platform will automatically power on when AC power is applied. There is no power system. Check all boards, power supplies, and fan tray modules for proper insertion before plugging in a power cable.

1. Plug in the first power cable.
2. Plug in the second power cable.
3. Wait for the system upload process.

   > It may take up to five minutes to turn on the system. If the System Status LED shows amber after five minutes, unplug the system and call your NVIDIA representative for assistance.

4. Check the frontal System Status LEDs and confirm that all of the LEDs show status lights consistent with normal operation (initially flashing, and then moving to a steady color) as shown below. For more information, refer to LED Notifications.

   System Status LEDs 5 Minutes After Power On

After inserting a power cable and confirming the green System Status LED light is on, make sure that the Fan Status LED shows green. If the Fan Status LED is not green, unplug the power connection and check that all fan modules are inserted properly and that the mating connector of the fan unit is free of any dirt and/or obstacles. If no obstacles were found and the problem persists, call your NVIDIA representative for assistance.

Two Power Inlets - Electric Caution Notifications:

- Risk of electric shock and energy hazard. The two power supply units are independent. Disconnect all power supplies to ensure a powered down state inside of the switch platform.
- ACHTUNG Gafahr des elektrischen Schocks. Entferrnen des Netzsteckers elnes Netzteils spannungsfrei. Um alle Einhieten spannungsfrei zu machen sind die Netzstecker aller Netzteile zu entfernen.
- ATTENTION Risque de choc et de danger e'lectriques. Le de'branchment d'une seule alimentation stabilise'e ne de'branch uniquement qu'un module "Alimentation Stabilise'e". Pour isoler completement le module en cause, Il faut de'brancher toutes les alimentations stabilise'es.
- 電擊與能源危害的危險。所有 PSU 均各自獨立。將所有電源供應器斷電，確保交換器平台內部在電源關閉狀態。

System Bring-Up of Managed Systems

The bring-up procedures described in this section do not apply to unmanaged/externally managed systems. Such systems are ready for operation after power-on.

In order to query the system, perform firmware upgrade or other firmware operation. Refer to the latest Mellanox Firmware tools (MFT) located on https://network.nvidia.com/products/adapter-software/firmware-tools/.

In order to obtain the firmware version of the externally managed system:

1. Run the following command from a host:

```
# flint -d <device> q
```

2. Compare the results of this command with the latest version for your system posted on https://network.nvidia.com/products/adapter-software/firmware-tools/.
3. If the current version is not the latest version, follow the directions in the *MFT User Manual* to burn the new firmware.

Configuring Network Attributes

The procedures described in this chapter assume that you have already installed and powered-on the system according to the instructions in this document. Since the system comes with a pre-

configured DHCP, you may find the explanation in Disable Dynamic Host Configuration Protocol (DHCP) sufficient. In case manual configuration is required, please refer to the instructions in Manual Host Configuration.

Disable Dynamic Host Configuration Protocol (DHCP)

DHCP is used for automatic retrieval of management IP addresses.

If a user connects through SSH, runs the wizard and turns off DHCP, the connection is immediately terminated, as the management interface loses its IP address. In such a case, the serial connection should be used.

<localhost># ssh admin@<ip-address>

```
Mellanox MLNX-OS Switch Management
Password:
Mellanox Switch
Mellanox configuration wizard
Do you want to use the wizard for initial configuration? yes
Step 1: Hostname? [my-switch]
Step 2: Use DHCP on mgmt0 interface? [yes] no
<localhost>#
```

Manual Host Configuration

To perform initial configuration of the system:

1. Connect a host PC to the Console RJ45 (IOIOI) port of the system, using the supplied harness cable (DB9 to RJ45). Make sure to connect to the Console RJ45 port and not to the (Ethernet) MGT () port.
2. Configure a serial terminal program (for example, HyperTerminal, minicom, or Tera Term) on your host PC with the settings described in the table below. Once you perform that, you should get the CLI prompt of the system.

   Serial Terminal Program Configuration

Parameter	Setting
Baud Rate	115200
Data bits	8
Stop bits	1
Parity	None
Flow Control	None

3. The boot menu is prompted.

```
...                                                                    .
This terminal is not active for input or output while booting.


    Boot Menu                                                          .

------------------------------------------------------------------------
0: <image #1>
1: <image #2>
------------------------------------------------------------------------

      Use the ^ and v keys to select which entry is highlighted.
      Press enter to boot the selected image or 'p' to enter a
      password to unlock the next set of features.

    Highlighted entry is 0:
```

⚠ Select “0” to boot with software version installed on partition #1.
Select “1” to boot with software version installed on partition #2.

The boot menu features a countdown timer. It is recommended to allow the timer to run out by not selecting any of the options.

4. Login as admin and use admin as password. If the machine is still initializing, you might not be able to access the CLI until initialization completes. As an indication that initialization is ongoing, a countdown of the number of remaining modules to be configured is displayed in the following format: “<no. of modules> Modules are being configured”.
5. Go through the Switch Management configuration wizard.
   IP Configuration by DHCP

Wizard Session Display (Example)	Comments
`Do you want to use the wizard for initial configuration? yes`	You must perform this configuration the first time you operate the switch or after resetting the switch to the factory defaults. Type “y” and then press <Enter>.
`Step 1: Hostname? [switch-1]`	If you wish to accept the default hostname, then press <Enter>. Otherwise, type a different hostname and press <Enter>.

Wizard Session Display (Example)	Comments
`Step 2: Use DHCP on mgmt0 interface? [yes]`	Perform this step to obtain an IP address for the switch. (mgmt0 is the management port of the switch.) - If you wish the DHCP server to assign the IP address, type “yes” and press <Enter>. If you type “no” (no DHCP), then you will be asked whether you wish to use the “zeroconf” configuration or not. If you enter “yes” (yes Zeroconf), the session will continue as shown in the "IP zeroconf configuration" table. If you enter “no” (no Zeroconf), then you need to enter a static IP, and the session will continue as shown in the "Static IP configuration" table.
`Step 3: Enable IPv6 [yes]`	Perform this step to enable IPv6 on management ports. If you wish to enable IPv6, type “yes” and press <Enter>. If you enter “no” (no IPv6), then you will automatically be referred to Step 5.
`Step 4: Enable IPv6 autoconfig (SLAAC) on mgmt0 interface`	Perform this step to enable StateLess address autoconfig on external management port. If you wish to enable it, type “yes” and press <Enter>. If you wish to disable it, enter “no”.
`Step 5: Use DHCPv6 on mgmt0 interface? [yes]`	Perform this step to enable DHCPv6 on the MGMT0 interface.
`Step 6: Update time?`	Perform this step to change the time configured. Press enter to leave the current time.
`Step 7: Enable password hardening?`	Perform this step to enable/ disable password hardening on your machine. If enabled, new passwords will be checked upon configured restrictions. If you wish to enable it, type “yes” and press <Enter>. If you wish to disable it, enter “no”.

Wizard Session Display (Example)	Comments
`Step 8: Admin password (Must be typed)? <new_password>`	To avoid illegal access to the machine, please type a password and then press <Enter>. Starting from the 3.8.2000 release, the user must type in the admin password upon initial configuration. Due to Senate Bill No. 327, this stage is required and cannot be skipped.
`Step 9: Confirm admin password? <new_password>`	Confirm the password by re-entering it. Note that password characters are not printed.
`Step 10: Monitor password (Must be typed)? <new_password>`	To avoid illegal access to the machine, please type a password and then press <Enter>. Starting from the 3.8.2000 release, the user must type in the admin password upon initial configuration. Due to Senate Bill No. 327, this stage is required and cannot be skipped.
`Step 11: Confirm monitor password? <new_password>`	Confirm the password by re-entering it. Note that password characters are not printed.

`You have entered the following information:`
`Hostname: <switch name>`
`Use DHCP on mgmt0 interface: yes`
`Enable IPv6: yes`
`Enable IPv6 autoconfig (SLAAC) on mgmt0 interface: yes`
`Enable DHCPv6 on mgmt0 interface: no`
`Update time: <current time>`
`Enable password hardening: yes`
`Admin password (Enter to leave unchanged): (CHANGED)`
`To change an answer, enter the step number to return to.`
`Otherwise hit <enter> to save changes and exit.`
`Choice: <Enter>`
`Configuration changes saved.`
`To return to the wizard from the CLI, enter the "configuration jump-start" command`
`from configuration mode. Launching CLI...`
`<switch name> [standalone: master] >`
	The wizard displays a summary of your choices and then asks you to confirm the choices or to re-edit them.
	Either press <Enter> to save changes and exit, or enter the configuration step number that you wish to return to.
	To run the command "configuration jump-start" you must be in Config mode.

IP Configuration by DHCP for Modular Switch Systems

Wizard Session Display (Example)	Comments
`Do you want to use the wizard for initial configuration? yes`	You must perform this configuration the first time you operate the switch or after resetting the switch to the factory defaults. Type "y" and then press <Enter>.

Wizard Session Display (Example)	Comments
`Step 1: Hostname? [switch-1]`	If you wish to accept the default hostname, then press <Enter>. Otherwise, type a different hostname and press <Enter>.
`Step 2: Use DHCP on mgmt0 interface? [yes]`	Perform this step to obtain an IP address for the switch. (mgmt0 is the management port of the switch.) If you wish the DHCP server to assign the IP address, type “yes” and press <Enter>. If you type “no” (no DHCP), then you will be asked whether you wish to use the “zeroconf” configuration or not. If you enter “yes” (yes Zeroconf), the session will continue as shown in the IP zeroconf configuration" table. If you enter “no” (no Zeroconf), then you need to enter a static IP, and the session will continue as shown in the "Static IP configuration" table.
`Step 3: Enable IPv6 [yes]`	Perform this step to enable IPv6 on management ports. If you wish to enable IPv6, type “yes” and press <Enter>. If you enter “no” (no IPv6), then you will automatically be referred to Step 5.
`Step 4: Enable IPv6 autoconfig (SLAAC) on mgmt0 interface`	Perform this step to enable StateLess address autoconfig on external management port. If you wish to enable it, type “yes” and press <Enter>. If you wish to disable it, enter “no”.
`Step 5: Use DHCPv6 on mgmt0 interface? [yes]`	Perform this step to enable DHCPv6 on the MGMT0 interface.
`Step 6: Admin password (Press <Enter> to leave unchanged)? <new_password>`	To avoid illegal access to the machine, please type a password and then press <Enter>.
`Step 7: Confirm admin password? <new_password>` `(this step only happens if you change the password)`	Confirm the password by re-entering it. Note that password characters are not printed.
`Step 9: HA Chassis Management IP netmask? (Example: [255.255.255.0])`	Perform this step to configure the box IPv4 netmask. If you wish to accept the default value, type “yes” and press <Enter>. Otherwise, enter the desired box IPv4 netmask
`Step 10: HA Chassis IPv6 address? (Example: [fdfd:fdfd:7:145::1000:4814])`	Perform this step to configure the box IPv6. If you wish to accept the default value, type “yes” and press <Enter>. Otherwise, enter the desired box IPv6
`Step 11: HA Chassis Management IPv6 masklen? (Example: [33])`	Perform this step to configure the box IPv6 masklen. If you wish to accept the default value, type “yes” and press <Enter>. Otherwise, enter the desired box IPv6 masklen.

Wizard Session Display (Example)	Comments
You have entered the following information: Hostname: <switch name> Use DHCP on mgmt0 interface: yes Enable IPv6: yes Enable IPv6 autoconfig (SLAAC) on mgmt0 interface: yes Enable DHCPv6 on mgmt0 interface: yes Admin password (Enter to leave unchanged): (CHANGED) HA Chassis IP address: 10.6.166.200 HA Chassis Management IP netmask: 255.255.255.0 HA Chassis IPv6 address: fdfd:fdfd:7:145::1000:4814 HA Chassis Management IPv6 masklen: 33 To change an answer, enter the step number to return to. Otherwise hit <enter> to save changes and exit. Choice: <Enter> Configuration changes saved. To return to the wizard from the CLI, enter the "configuration jump-start" command from configuration mode. Launching CLI... <switch name> [standalone: master] >	The wizard displays a summary of your choices and then asks you to confirm the choices or to re-edit them. Either press <Enter> to save changes and exit, or enter the configuration step number that you wish to return to. To run the command "configuration jump-start" you must be in Config mode.

Static IP Configuration

Wizard Session Display (Example)
Do you want to use the wizard for initial configuration? y Step 1: Hostname? [switch-112126] Step 2: Use DHCP on mgmt0 interface? [yes] n Step 3: Use zeroconf on mgmt0 interface? [no] Step 4: Primary IP address? 192.168.10.4 Mask length may not be zero if address is not zero (interface mgmt0) Step 5: Netmask? [0.0.0.0] 255.255.255.0 Step 6: Default gateway? 192.168.10.1 Step 7: Primary DNS server? Step 8: Domain name? Step 9: Enable IPv6? [yes] yes Step 10: Enable IPv6 autoconfig (SLAAC) on mgmt0 interface? [no] no Step 11: Update time? [yyyy/mm/dd hh:mm:ss] Step 12: Enable password hardening? [yes] yes Step 13: Admin password (Enter to leave unchanged)? You have entered the following information: Hostname: switch-112126 Use DHCP on mgmt0 interface: no Use zeroconf on mgmt0 interface: no Primary IP address: 192.168.10.4 Netmask: 255.255.255.0 Default gateway: 192.168.10.1 Primary DNS server: Domain name: Enable IPv6: yes Enable IPv6 autoconfig (SLAAC) on mgmt0 interface: no Update time: yyyy/mm/dd hh:mm:ss Enable password hardening: yes Admin password (Enter to leave unchanged): (unchanged) To change an answer, enter the step number to return to. Otherwise hit <enter> to save changes and exit. Choice: Configuration changes saved. To return to the wizard from the CLI, enter the "configuration jump-start" command from configure mode. Launching CLI... <hostname>[standalone: master] >

IP Zeroconf Configuration

Wizard Session Display (Example)

```
Configuration wizard

Do you want to use the wizard for initial configuration? y

Step 1: Hostname? [switch-112126]
Step 2: Use DHCP on mgmt0 interface? [no]
Step 3: Use zeroconf on mgmt0 interface? [no] yes
Step 4: Default gateway? [192.168.10.1]
Step 5: Primary DNS server?
Step 6: Domain name?
Step 7: Enable IPv6? [yes] yes
Step 8: Enable IPv6 autoconfig (SLAAC) on mgmt0 interface? [no] no
Step 9: Update time? [yyyy/mm/dd hh:mm:ss]
Step 10: Admin password (Enter to leave unchanged)?

You have entered the following information:

Hostname: switch-112126
Use DHCP on mgmt0 interface: no
Use zeroconf on mgmt0 interface: yes
Default gateway: 192.168.10.1
Primary DNS server:
Domain name:
Enable IPv6: yes
Enable IPv6 autoconfig (SLAAC) on mgmt0 interface: yes
Update time: yyyy/mm/dd hh:mm:ss
Enable password hardening: yes
Admin password (Enter to leave unchanged): (unchanged)

To change an answer, enter the step number to return to.
Otherwise hit <enter> to save changes and exit.

Choice:

Configuration changes saved.

To return to the wizard from the CLI, enter the “configuration jump-start”
command from configure mode. Launching CLI...
<hostname> [standalone: master] >
```

IP Zeroconf Configuration **for Modular Switch Systems**

Wizard Session Display (Example)

```
Configuration wizard

Do you want to use the wizard for initial configuration? y

Step 1: Hostname? [switch-mgmt1]
Step 2: Use DHCP on mgmt0 interface? [yes]
Step 3: Enable IPv6? [yes]
Step 4: Enable IPv6 autoconfig (SLAAC) on mgmt0 interface? [no]
Step 5: Enable DHCPv6 on mgmt0 interface? [yes]
Step 6: Admin password (Enter to leave unchanged)?
Step 7: HA Chassis IP address: [10.6.166.200]
Step 8: HA Chassis Management IP netmask: [255.255.255.0]
Step 9: HA Chassis IPv6 address: [fdfd:fdfd:7:145::1000:4814]
Step 10: HA Chassis Management IPv6 masklen: [33]


You have entered the following information:

1. Hostname: sw-mantaray-201-mgmt1
2. Use DHCP on mgmt0 interface: yes
3. Enable IPv6: yes
4. Enable IPv6 autoconfig (SLAAC) on mgmt0 interface: no
5. Enable DHCPv6 on mgmt0 interface: yes
6. Admin password (Enter to leave unchanged): (unchanged)
7. HA Chassis IP address: 10.6.166.200
8. HA Chassis Management IP netmask: 255.255.255.0
9. HA Chassis IPv6 address: fdfd:fdfd:7:145::1000:4814
10. HA Chassis Management IPv6 masklen: 33

To change an answer, enter the step number to return to.
Otherwise hit <enter> to save changes and exit.

Choice:
Configuration changes saved.

To return to the wizard from the CLI, enter the “configuration jump-start”
command from configure mode. Launching CLI...
<hostname> [standalone: master] >
```

6. Check the mgmt0 interface configuration before attempting a remote (for example, SSH) connection to the switch. Specifically, verify the existence of an IP address.

```
switch # show interfaces mgmt0

Interface mgmt0 status:
  Comment          :
  Admin up         : yes
  Link up          : yes
  DHCP running     : yes
  IP address       : 10.12.67.34
  Netmask          : 255.255.0.0
  IPv6 enabled     : yes
  Autoconf enabled: no
  Autoconf route   : yes
  Autoconf privacy: no
  DHCPv6 running   : no
  IPv6 addresses   : 1

  IPv6 address:
    fe80::268a:7ff:fe53:3d8e/64
```

```
Speed           : 1000Mb/s (auto)
Duplex          : full (auto)
Interface type  : ethernet
Interface source: physical
MTU             : 1500
HW address      : 00:02:c9:11:a1:b2

Rx:
 11700449  bytes
    55753  packets
        0  mcast packets
        0  discards
        0  errors
        0  overruns
        0  frame

Tx:
 5139846  bytes
   28452  packets
       0  discards
       0  errors
       0  overruns
       0  carrier
       0  collisions
    1000  queue len
```

Configuring the Switch with ZTP

Zero-touch Provisioning (ZTP) automates initial configuration of switch systems at boot time. It helps minimize manual operation and reduce customer initial deployment cost.

For more information, please refer to section "Zero-touch Provisioning".

Rerunning the Wizard

To rerun the wizard:

1. Enter Config mode. Run:

```
switch > enable
switch # config terminal
```

2. Rerun the wizard. Run:

```
switch (config) # configuration jump-start
```

Starting the Command Line (CLI)

1. Set up an Ethernet connection between the switch and a local network machine using a standard RJ-45 connector.
2. Start a remote secured shell (SSH) to the switch using the command "ssh -l <username> <switch ip address>".

```
rem_mach1 >  ssh -l <username> <ip address>
```

3. Log into the switch (default username is admin, password admin).
4. Read and accept the EULA when prompted.
5. Once the following prompt appears, the system is ready to use.

```
NVIDIA MLNX-OS Switch Management
```

```
Password:
Last login: <time> from <ip-address>


NVIDIA Switch
Please read and accept the End User License Agreement located at:
https://www.mellanox.com/related-docs/prod_management_software/MLNX-OS_EULA.pdf
switch >
```

Remote Connection

Once the network attributes are set, you can access the CLI via SSH or the WebUI via HTTP/ HTTPs.

To access the CLI, perform the following steps:

1. Set up an Ethernet connection between the system and a local network machine using a standard RJ45 cable.
2. Start a remote secured shell (SSH) using the command: ssh -l <username> <IP_address>

```
# ssh -l <username> <ip_address>
Mellanox MLNX-OS Switch Management

Password:
```

3. Login as admin (default username is *admin*, password is *admin*).
4. Once you get the CLI prompt, you are ready to use the system.

For additional information about MLNX-OS, refer to the *MLNX-OS User Manual* located on https://docs.nvidia.com/networking/category/mlnxos.

FRU Replacements

Power Supply

NVIDIA systems are equipped with two replaceable power supply units work in a redundant configuration. Either unit may be extracted without bringing down the system.

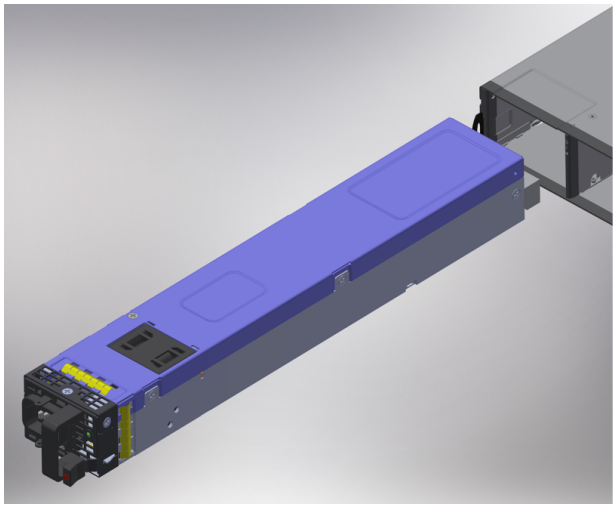
> Make sure that the power supply unit that you are NOT replacing is showing all green, for both the power supply unit and the rear System Status LEDs.

> Power supply units have directional air flows similar to the fan module. The fan module airflow must coincide with the airflow of all of the power supply units. If the power supply unit airflow direction is different from the fan module airflow direction, the system's internal temperature will be affected. For power supply unit air flow direction, refer to Air Flow.

To extract a power supply unit:

1. Remove the power cord from the power supply unit.
2. Grasping the handle with your hand, push the latch release with your thumb while pulling the handle outward. As the power supply unit unseats, the power supply unit status LEDs will turn off.

3. Remove the power supply unit.
   PS Unit Pulled Out



To insert a power supply unit:

1. Make sure the mating connector of the new unit is free of any dirt and/or obstacles.

   > Do not attempt to insert a power supply unit with a power cord connected to it.

2. Insert the power supply unit by sliding it into the opening, until a slight resistance is felt.
3. Continue pressing the power supply unit until it seats completely. The latch will snap into place, confirming the proper installation.
4. Insert the power cord into the power supply connector.
5. Insert the other end of the power cord into an outlet of the correct voltage.

   > The green power supply unit indicator should light. If it does not, repeat the whole procedure to extract the power supply unit and re-insert it.

Fans

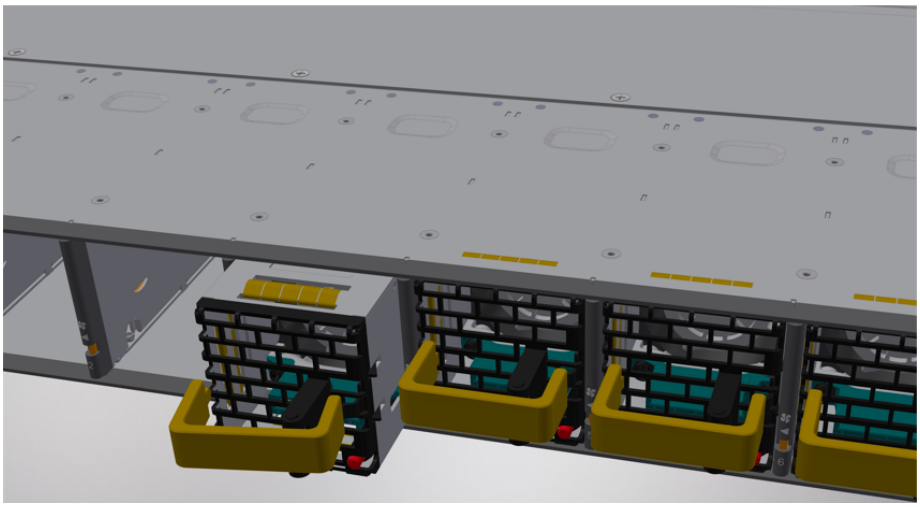
The system can fully operate if one fan FRU is dysfunctional. Failure of more than one fan is not supported.

> Make sure that the fans have the air flow that matches the model number. An air flow opposite to the system design will cause the system to operate at a higher (less than optimal) temperature. For power supply unit air flow direction, refer to Air Flow.

To extract a fan unit:

1. Extract the fan by pulling the gold handle outwards. As the fan unit unseats, its status LEDs will turn off.
2. Remove the fan unit.

Fan Module Latches



To remove or replace a fan unit, gently pull out its handle while pushing the latch release with your index finger.

To insert a fan unit:

1. Make sure the mating connector of the new unit is free of any dirt and/or obstacles.
2. Insert the fan unit by sliding it into the opening until slight resistance is felt. Continue pressing the fan unit until it seats completely.

The green Fan Status LED should light. If not, extract the fan unit and reinsert it. After two unsuccessful attempts to install the fan unit, power off the system before attempting any system debug.

Software Management

Managed systems come with an embedded management CPU card that runs MLNX-OS® management software. The MLNX-OS systems management package and related documentation can be downloaded at https://docs.nvidia.com/networking/category/mlnxos.

InfiniBand Subnet Manager

The InfiniBand Subnet Manager (SM) is a centralized entity running in the system. The SM applies network traffic related configurations such as QoS, routing, partitioning to the fabric devices. You can view and configure the Subnet Manager parameters via the CLI/WebUI. Each subnet needs one subnet manager to discover, activate and manage the subnet.

Each network requires a Subnet Manager to be running in either the system itself (system based) or on one of the nodes which is connected to the fabric (host based).

> ⚠ No more than two subnet managers are recommended for any single fabric.

The InfiniBand Subnet Manager running on the system supports up to 2048 nodes. If the fabric includes more than 2048 nodes, you may need to purchase Mellanox's Unified Fabric Manager (UFM®) software package.

Each subnet needs one subnet manager to discover, activate and manage the subnet.

Each network requires a Subnet Manager to be running in either the system itself (system based) or on one of the nodes which is connected to the fabric (host based).

The subnet manager (OpenSM) assigns Local IDentifiers (LIDs) to each port connected to the fabric, and develops a routing table based on the assigned LIDs.

A typical installation using the OFED package will run the OpenSM subnet manager at system start up after the drivers are loaded. This automatic OpenSM is resident in memory, and sweeps the fabric approximately every 5 seconds for new adapters to add to the subnet routing tables.

Upgrading Software (on Managed Systems)

Software and firmware updates are available from the NVIDIA Support website. Check that your current revision is the same one that is on the NVIDIA website. If not upgrade your software. Copy the update to a known location on a remote server within the user's LAN.

Use the CLI or the GUI in order to perform software upgrades. For further information please refer to the Upgrading MLNX-OS® Software section in the MLNX-OS Software User Manual.

Be sure to read and follow all of the instructions regarding the updating of the software on your system.

Managed systems do not require Firmware updating. Firmware updating is done through the MLNX-OS management software. The system comes standard with a management software module for

system management called Mellanox Operating System (MLNX-OS). MLNX-OS® is installed on all NVIDIA Mellanox Quantum™ based managed systems. MLNX-OS® includes a CLI, WebUI, SNMP, system management software and IB management software (OpenSM).

The Ethernet ports for remote management connect to Ethernet systems. These systems must be configured to 100Mb/1Gb/s auto-negotiation.

Updating Firmware on Externally Managed Systems

There are two methods to update system firmware:

- (Typical) In-band via a switch network port across a cable connecting the server to the switch port.
- (Non-typical) Via the I²C port of the switch using an NVIDIA MTUSB-1 device connecting to a server's USB port on the one end and to the I²C port of the switch on the other.

Firmware updates should normally be conducted in-band. The use of the MTUSB-1 device is intended for cases of debug or firmware corruption and should be conducted by NVIDIA FAEs or Support engineers, or by trained users at the customer's site.

Both types of updates require the installation of NVIDIA Mellanox Firmware Tools (MFT) package. The MFT package and user manual are available for download under https://network.nvidia.com/products/adapter-software/firmware-tools/. Please select the package that suits your operating system.

In order to obtain information regarding the externally managed system, you must download the NVIDIA Mellanox MFT tools from https://network.nvidia.com/products/adapter-software/firmware-tools/.

Select and download the release that matches your system. Follow the instructions in the User Manual https://docs.nvidia.com/networking/category/mft to get the tools.

Updating Firmware In-band (Typical)

Check the currently programmed firmware on the system and compare it to the latest firmware available under https://network.nvidia.com/support/firmware/firmware-downloads/ (check under Quantum™ Switch Systems).

In order to obtain the firmware version of the externally managed system:

1. Obtain the LID of the target system. The following instructions use one of the utilities provided by the installed MFT package. (Other methods are described in the *MFT User Manual*) by performing the following:
    a. Mark the GUID printed on the inventory pull-out tab of the system.
    b. Run the command ibnetdiscover and search for the row starting with the word "Switch" and indicates the GUID of the system.
    c. Mark the displayed LID on that row (a decimal number).
2. Run the following command from a host:

```
# flint -d <device> q#
```

3. Compare the results of this command with the latest version for your system posted on https://network.nvidia.com/support/firmware/firmware-downloads/ (select the Quantum™ System page).
4. If the current version is not the latest version, follow the directions in the MFT User manual to burn the new firmware inband.

For further information, please refer to *MFT User Manual* at https://docs.nvidia.com/networking/category/mft.

Interfaces

The systems support the following interfaces:

- Data interfaces - InfiniBand
- 10/100/1000Mb Ethernet management interface (RJ45)*
- USB port (USB Type A)*
- RS232 Console port (RJ45)**
- I²C interface*
- Reset button
- Status and Port LEDs

*This interface is not found in managed systems.

**This interface is not found in externally managed systems.

In order to review the full configuration options matrix, refer to Management Interfaces, PSUs and Fans.

Data Interfaces

The data interfaces use OSFP connectors. The full list of interfaces per system is provided in Speed and Switching Capabilities.

Each OSFP port consists of 2 logical InfiniBand ports, and can be connected with OSFP cable or connector for 40/56/100/200/400 Gb/s. The system offers Class 8 (17W) OSFP112 transceivers support.

Speed

InfiniBand speed is auto-adjusted by the InfiniBand protocol. NVIDIA systems support QDR/FDR/EDR/HDR/NDR InfiniBand.

- FDR is an InfiniBand data rate, where each lane of a 4X port runs a bit rate of 14.0625Gb/s with 64b/66b encoding, resulting in an effective bandwidth of 56.25Gb/s.
- EDR is an InfiniBand data rate, where each lane of a 4X port runs a bit rate of 25Gb/s with 64b/66b encoding, resulting in an effective bandwidth of 100Gb/s.
- HDR is an InfiniBand data rate, where each lane of a 4X port runs a bit rate of 50Gb/s with 64b/66b encoding, resulting in an effective bandwidth of 200Gb/s.
- NDR is an InfiniBand data rate, where each lane of a 4X port runs a bit rate of 100Gb/s with 64b/66b encoding, resulting in an effective bandwidth of 400Gb/s.

RS232 (Console)

The Console port labeled " IOIOI " is an RS232 serial port on the front side of the chassis that is used for initial configuration and debugging. Upon first installation of the system, you need to

connect a PC to this interface and configure network parameters for remote connections. Refer to Configuring Network Attributes to view the full procedure.

> ⚠ This interface is not found in externally managed systems.

Management

The Management RJ45 Ethernet ports labeled “  ” provide access for remote management. The management ports are configured with auto-negotiation capabilities by default (100MbE to 1000GbE). The management ports’ network attributes (such as IP Address) need to be pre-configured via the RS232 serial console port or by DHCP before use. Refer to Configuring Network Attributes to view the full procedure.

> ⚠ This interface is not found in externally managed systems.

> ⚠ Make sure you use only FCC compliant Ethernet cables.

USB

The USB interface is USB3.0 type A compliant and can be used by MLNX-OS software to connect to an external disk for software upgrade or file management. The connector comes in a standard micro USB shape. To view the full matrix of micro USB configuration options, refer to Management Interfaces, PSUs and Fans.

> ⚠
> - USB 1.0 is not supported.
> - Do not use excessive force when inserting or extracting the USB disk to and from the connector.
> - This interface is not found in externally managed systems.

I²C

The I²C connector is combined with the USB connector, and is located on the front side of the system. It can be used with the I²C DB9 to micro USB splitting harness.

> ⚠
> - This interface is not found in managed systems. It is available in QM9790 systems only.
> - Apart from the initial configuration, I²C interface is made exclusively for debugging and troubleshooting. Only FAEs are authorized to connect through it.

Only original NVIDIA cables supplied with the switch package can be used to connect a switch system to the server.

Connecting any cable other than the NVIDIA supplied console cable may cause an I²C hang. Using uncertified cables may damage the I²C interface.

Refer to the .Replacement Parts Ordering Numbers v2.4 appendix for harness details.

Reset Button

The reset button is located on the front side of the system under the USB port. This reset button requires a tool to be pressed.

Do not use a sharp pointed object such as a needle or a push pin for pressing the reset button. Use a flat object to push the reset button.

- To reset the system, push the reset button for less than 15 seconds.
- When using an Onyx (MLNX-OS) based system, keeping the reset button pressed for more than 15 seconds will reset the system and the “admin” password, this should allow you to enter without a password and set a new password for the user “admin”.

LEDs

See LED Notifications.

LED Notifications

The system’s LEDs are an important tool for hardware event notification and troubleshooting.

LEDs Symbols

Symbol	Name	Description	Normal Conditions
	System Status LED	Shows the health of the system.	Green/Flashing green when booting
	Power Supply Units LEDs	Shows the health of the power supply units.	Green

Symbol	Name	Description	Normal Conditions
	Fan Status LED	Shows the health of the fans.	Green
 *	Unit Identifier LED	Lights up on command through the CLI.	Off or blue when identifying a port

System Status LED

System Status LED - Front Side

Front Panel	Description
	The LED in the red rectangle shows the system's status.

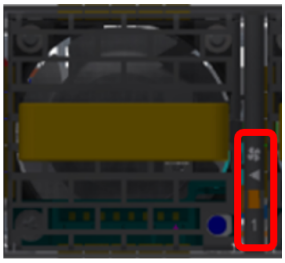
It may take up to five minutes to turn on the system. If the System Status LED shows amber after five minutes, unplug the system and call your NVIDIA representative for assistance.

System Status LED Assignments

LED Behavior	Description	Action Required
Solid Green	The system is up and running normally.	N/A
Flashing Green	The system is booting up. This assignment is valid on managed systems only.	Wait up to five minutes for the end of the booting process.
Solid Amber	Major error has occurred. For example, corrupted firmware, system is overheated, etc.	If the System Status LED shows amber five minutes after starting the system, unplug the system and call your NVIDIA representative for assistance.

Fan Status LED

Fan Status LED - Front and Rear Sides

Front Panel	Description	Rear Panel
	Both of these LEDs in the red rectangles show the fans' status.	

Fan Status Front LED Assignments

LED Behavior	Description	Action Required
Solid Green	All fans are up and running.	N/A
Solid Amber	Error, one or more fans are not operating properly.	The faulty FRUs should be replaced.

Fan Status Rear LED Assignments (One LED per Fan)

LED Behavior	Description	Action Required
Solid Green	A specific fan unit is operating.	N/A
Solid Amber	A specific fan unit is missing or not operating properly.	The fan unit should be replaced.

Risk of Electric Shock! With the fan module removed, power pins are accessible within the module cavity. Do not insert tools or body parts into the fan module cavity.

Power Supply Status LEDs

There are two power supply inlets in the system (for redundancy). The system can operate with only one power supply connected. Each power supply unit has a single 2 color LED that indicates the status of the unit.

Power Status LED

Rear Side Panel

Power Supply Unit Status Front LED Assignments

LED Behavior	Description	Action Required
Solid Green	All plugged (one or two) power supplies are running normally.	N/A
Solid Amber	One or both of the power supplies are not operational or not powered up/ the power cord is disconnected.	Make sure the power cord is plugged in and active. If the problem resumes, the FRUs might be faulty, and should then be replaced.

The power supply status LEDs on the rear side of the system are located on the PSUs themselves. Each PSU has a single 2 color LED.

Power Supply Unit Status Rear LED Assignments

LED Behavior	Description	Action Required
Solid Green	All PS units are connected and running normally.	N/A
Flashing Green 1Hz	AC present / Only 12VSB on (PSU off) or PSU in Smart-on state.	Call your NVIDIA representative for assistance.
Amber	AC cord unplugged or AC power lost while the second power supply still has AC input power.	Plug in the AC cord of the faulty PSU.
	PS failure (including voltage out of range and power cord disconnected).	Check voltage. If OK, call your NVIDIA representative for assistance.

LED Behavior	Description	Action Required
Flashing Amber	Power supply warning events where the power supply continues to operate; high temp, high power, high current, slow fan.	Call your NVIDIA representative for assistance.
Off	No AC power to all power supplies.	Call your NVIDIA representative for assistance.

Unit Identification LED

The UID LED is a debug feature, that the user can use to find a particular system within a cluster by turning on the UID blue LED.

To activate the UID LED on a switch system, run:

```
switch (config) # led MGMT uid on
```

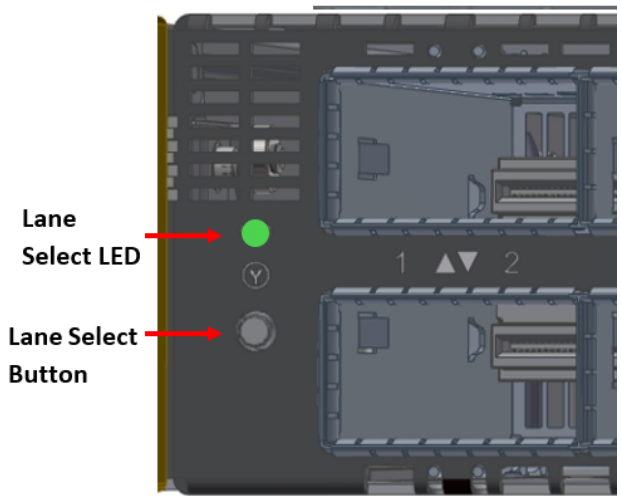
To verify the LED status, run:

```
switch (config) # show leds
Module LED Status
------------------------------------------------------------------------------
MGMT UID Blues
```

To deactivate the UID LED on a switch system, run:

```
switch (config) # led MGMT uid off
```

Port LEDs



Each time you press on the Lane Select Button, the Port LEDs display will switch to a different state, as follows:

Lane Select Button States

State	LED Status	Ports LED Indication
0 (Default)	LED is off	4x || 2xA
1	LED is on	4x || 2xB

The port LEDs behavior indicates the ports' state, as follows:

Port LEDs in InfiniBand System Mode

LED Behavior	Description	Action Required
Off	Link is down.	Check the cable.
Solid Green	Link is up with no traffic.	N/A
Flashing Green	Link is up with traffic.	N/A
Solid Amber	Link is up.	Wait for the Logical link to raise. Check that the SM is up.
Flashing Amber	A problem with the link.	Check that the SM is up.

In InfiniBand system mode, the LED indicator, corresponding to each data port, will light orange when the physical connection is established (that is, when the unit is powered on and a cable is plugged into the port with the other end of the connector plugged into a functioning port). When a logical connection is made the LED will change to green. When data is being transferred the light will blink green.

Inventory Pull-out Tab

The system's inventory parameters (such as serial number, part number and GUID address) can be extracted from the inventory pull-out tab on the lower left side of the rear panel.

Pull-out Tab

The images provided here are for illustration purposes only. The may not reflect the latest version of the product nor all available models.

Troubleshooting

Problem Indicator	Symptoms	Cause and Solution
LEDs	System Status LED is blinking for more than 5 minutes	Cause: MLNX-OS software did not boot properly and only firmware is running. Solution: Connect to the system via the console port, and check the software status. You might need to contact an FAE if the MLNX-OS software did not load properly.
	System Status LED is amber	Cause: • Critical system fault (CPU error, bad firmware) • Over temperature Solution: • Check environmental conditions (room temperature)
	Fan Status LED is amber	Cause: Possible fan issue Solution: • Check that the fan is fully inserted and nothing blocks the airflow • Replace the fan FRU if needed
	Front PSU Status LED is amber	Cause: Possible PSU issue Solution: • Check/replace the power cable • Replace the PSU if needed
	The activity LED does not light up (InfiniBand)	Make sure that there is an SM running in the fabric.

System boot failure	The last software upgrade failed on x86 based systems	Solution:
		• Connect the RS232 connector (CONSOLE) to a laptop.
		• Push the system's reset button.
		• Press the ArrowUp or ArrowDown key during the system boot. GRUB menu will appear. For example:
		
		`Default image: 'SX_X86_64 SX_3.4.0008 2014-11-10 20:07:51 x86_64'`
		`Press enter to boot this image, or any other key for boot menu`
		`Booting default image in 3 seconds.`
		`Boot Menu`
		`----------`
		`-`
		`0: SX_X86_64 SX_3.4.0008 2014-11-10 20:07:51 x86_64`
		`1: SX_X86_64 SX_3.4.0007 2014-10-23 17:27:34 x86_64`
		`----------`
		`-`
		`Use the ArrowUp and ArrowDown keys to select which entry is highlighted.`
		`Press enter to boot the selected image or 'p' to enter a password to unlock the next set of features.`
		`Highlighted entry is 0:`
		`"`
		
		• Select previous image to boot by pressing an arrow key and choosing the appropriate image.

Specifications

QM9700 and QM9790 Technical Specifications

Feature		Value
Mechanical	Size:	1.7” (H) x 17.2” (W) x26” (D), 43.6mm (H) x 438mm (W) x 660mm (D)
	Mounting:	19” rack mount
	Weight:	1 PSU: 13.6 kg 2 PSUs: 14.8 kg
	Speed:	40, 56, 100, 200, 400 Gb/s per port
	Connector cage:	32 OSFP
Environmental	Temperature:	Operational: Forward air flow: 0° to 35°C Reverse air flow: 0° to 35°C
		Non-Operational: -40° to 70°C
	Humidity:	Operational: 10%-85% non-condensing Non-Operational: 10%-90% non-condensing
	Altitude:	3050m
	Noise level:	78.4dBA at room temperature
Regulatory	Safety:	CB, cTUVus, CE, CU
	EMC:	EMC: CE, FCC, VCCI, ICES, RCM
	RoHS:	RoHS compliant
Power	Input Voltage:	1x/2x, 200-240Vac, 10A, 50/60Hz
	Global Power Consumption:	**QM9700:** Typical power with passive cables (ATIS): 747W Max power with active cables: 1,703W **QM9790** Typical power with passive cables (ATIS): 640W Max power with active cables: 1,593W
Main Devices	CPU (in QM9700 only):	Intel® Core™ i3 Coffee Lake
	Switch:	NVIDIA Quantum™-2 IC
Throughput	Switching:	25.6Tbps

Appendixes

The document contains the following appendixes:

Accessory and Replacement Parts

Ordering Part Numbers for Replacement Parts

Part Number	Legacy Part Number	Part Description
930-9BRKT-00JM-000	MTEF-KIT-I-TL	NVIDIA 19" racks ,Tool-less rail-kit for QM97xx system, Rack size 600-800mm
930-9BFAN-00IW-000	MTEF-FANF-L	400G 1U systems FAN MODULE W/ P2C air flow
930-9BFAN-00JA-000	MTEF-FANR-L	400G 1U systems FAN MODULE W/ C2P air flow
930-9NPSU-00JN-000	MTEF-PSR-AC-K	NVIDIA Power-Supply Unit, 2000W AC, C2P Airflow, For QM97xx switches, Power cord included
930-9NPSU-00J6-000	MTEF-PSF-AC-K	NVIDIA Power-Supply Unit, 2000W AC, P2C Airflow, For QM97xx switches, Power cord included
HAR000631	-	Harness RS232 2M cable - DB9 to RJ-45 (for managed switches only)
ACC001897	-	Power cord black 250V 15A 1830MM C14 TO C15 UL
ACC001899	-	Power cord black 250V 10A 1830MM C14 TO C15 EUR + CCC
ACC001850	-	OSFP thermal cap with openings for airflow

Thermal Threshold Definitions

Three thermal threshold definitions are measured by the Quantum™ ASICs, and impact the overall switch system operation state as follows:

- Warning - 105°C: On managed systems only: When the ASIC device crosses the 100°C threshold, a Warning Threshold message will be issued by the management software, indicating to system administration that the ASIC has crossed the Warning threshold. Note that this temperature threshold does not require nor lead to any action by hardware (such as switch shutdown).
- Critical - 120°C: When the ASIC device crosses this temperature, the switch firmware will automatically shut down the device.

- Emergency – 130°C: In case the firmware fails to shut down the ASIC device upon crossing its Critical threshold, the device will auto-shutdown upon crossing the Emergency (130°C) threshold.

Interface Specifications

OSFP Pin Description

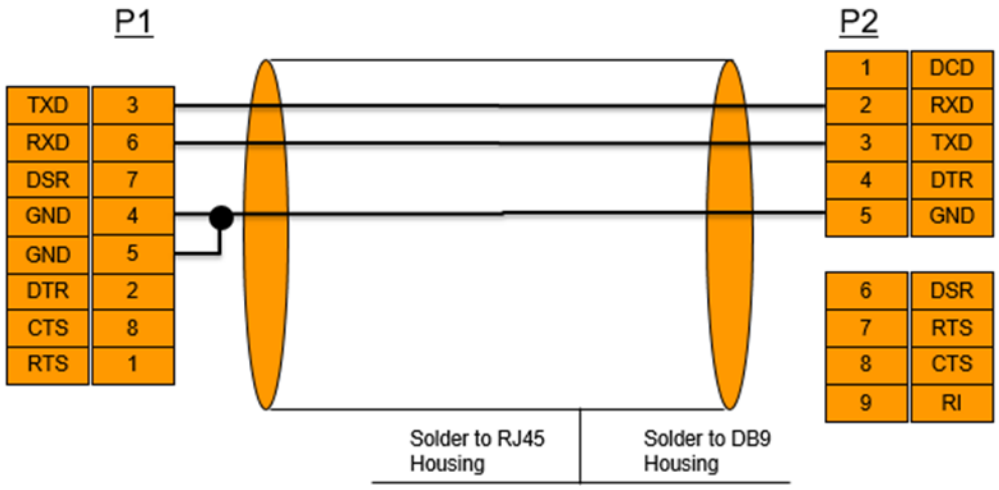
Net Name	PinNum	Signal Description
GND	1	Ground
TX2P	2	Transmitter Data Non-Inverted
TX2N	3	Transmitter Data Inverted
GND	4	Ground
TX4P	5	Transmitter Data Non-Inverted
TX4N	6	Transmitter Data Inverted
GND	7	Ground
TX6P	8	Transmitter Data Non-Inverted
TX6N	9	Transmitter Data Inverted
GND	10	Ground
TX8P	11	Transmitter Data Non-Inverted
TX8N	12	Transmitter Data Inverted
GND	13	Ground
SCL	14	2-wire Serial interface clock
VCC1	15	+3.3V Power
VCC1	16	+3.3V Power
LPWn_PRSn	17	PRSn Low-Power Mode / Module Present
GND	18	Ground
RX7N	19	Receiver Data Inverted
RX7P	20	Receiver Data Non-Inverted
GND	21	Ground
RX5N	22	Receiver Data Inverted
RX5P	23	Receiver Data Non-Inverted
GND	24	Ground
RX3N	25	Receiver Data Inverted
RX3P	26	Receiver Data Non-Inverted
GND	27	Ground
RX1N	28	Receiver Data Inverted
RX1P	29	Receiver Data Non-Inverted

Net Name	PinNum	Signal Description
GND	30	Ground
GND	31	Ground
RX2P	32	Receiver Data Non-Inverted
RX2N	33	Receiver Data Inverted
GND	34	Ground
RX4P	35	Receiver Data Non-Inverted
RX4N	36	Receiver Data Inverted
GND	37	Ground
RX6P	38	Receiver Data Non-Inverted
RX6N	39	Receiver Data Inverted
GND	40	Ground
RX8P	41	Receiver Data Non-Inverted
RX8N	42	Receiver Data Inverted
GND	43	Ground
INT_RSTn	44	INT/RSTn Module Interrupt / Module Reset
VCC2	45	+3.3V Power
VCC2	46	+3.3V Power
SDA	47	2-wire Serial interface data
GND	48	Ground
TX7N	49	Transmitter Data Inverted
TX7P	50	Transmitter Data Non-Inverted
GND	51	Ground
TX5N	52	Transmitter Data Inverted
TX5P	53	Transmitter Data Non-Inverted
GND	54	Ground
TX3N	55	Transmitter Data Inverted
TX3P	56	Transmitter Data Non-Inverted
GND	57	Ground
TX1N	58	Transmitter Data Inverted
TX1P	59	Transmitter Data Non-Inverted
GND	60	Ground

RJ45 to DB9 Harness Pinout

In order to connect a host PC to the Console RJ45 port of the system, a RS232 harness cable (DB9 to RJ45) is supplied.

RJ45 to DB9 Harness Pinout



RJ-45 Console and I²C interfaces are integrated in the same connector. Due to that, connecting any cable other than the NVIDIA supplied console cable may cause an I²C hang.

Using uncertified cables may damage the I²C interface. Refer to the Replacement Parts Ordering Numbers appendix for harness details.

Disassembly and Disposal

Disassembly Procedure

To disassemble the system from the rack:

1. Unplug and remove all connectors.
2. Unplug all power cords.
3. Remove the ground wire.
4. Unscrew the center bolts from the side of the system with the bracket.

   Support the weight of the system when you remove the screws so that the system does not fall.
5. Slide the system from the rack.
6. Remove the rail slides from the rack.
7. Remove the caged nuts.

Disposal

According to the WEEE Directive 2002/96/EC, all waste electrical and electronic equipment (EEE) should be collected separately and not disposed of with regular household waste. Dispose of this product and all of its parts in a responsible and environmentally friendly way.

Follow the instructions found at http://www.mellanox.com/page/dismantling_procedures for proper disassembly and disposal of the switch, according to the WEEE directive.

Document Revision History

Date	Revision	Description
July 2022	1.2	Updated OPNs in: • Ordering Information • Installation • Accessory and Replacement Parts Updated Cable Installation.
February 2022	1.1	Updated Cable Installation.
November 2021	1.0	Initial release

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. Neither NVIDIA Corporation nor any of its direct or indirect subsidiaries and affiliates (collectively: “NVIDIA”) make any 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.

Trademarks

NVIDIA, the NVIDIA logo, and Mellanox are trademarks and/or registered trademarks of NVIDIA Corporation and/or Mellanox Technologies Ltd. in the U.S. and in other countries. Other company and product names may be trademarks

of the respective companies with which they are associated.