

NVIDIA ConnectX-7 Cards for OCP Spec 3.0 User Manual

Table of Contents

1	Abo	out This Manual	5
	1.1	Ordering Part Numbers	5
	1.2	Intended Audience	6
	1.3	Technical Support	6
	1.4	Related Documentation	6
	1.5	Document Conventions	7
2	Inti	roduction	8
	2.1	Product Overview	8
	2.2	System Requirements	9
	2.3	Package Contents	9
	2.4	Features and Benefits	10
3	Sup	oported Interfaces	15
	3.1	ConnectX-7 Layout and Interface Information	15
	3.2	Interfaces Detailed Description	17
	3.2.	1 ConnectX-7 IC	17
	3.	.2.1.1 NVIDIA Multi-HostTM Support	
	3.2.2	PCI Express Interface	18
	3.2.3		
	3.2.4	· · · · · · · · · · · · · · · · · · ·	
	3.2.5		
	3.2.6		
	3.2.	7 Voltage Regulators	21
4	Har	rdware Installation	22
	4.1	Safety Warnings	22
	4.2	Installation Procedure Overview	22
	4.2.	1 System Requirements	23

	4.2.1.1	Hardware Requirements	23
	4.2.1.2	Airflow Requirements	23
	4.2.1.3	Software Requirements	24
	4.2.2	Safety Precautions	24
	4.2.3	Pre-Installation Checklist	25
	4.2.4	Cables and Modules	26
	4.2.5	Identifying the Card in Your System	27
	4.2.6	Extraction Instructions	28
5	Driver I	nstallation	29
	5.1 Wind	dows Driver Installation	29
	5.1.1	Software Requirements	29
	5.1.2	Downloading WinOF-2 Driver	30
	5.1.3	Installing WinOF-2 Driver	31
	5.1.3.1	Attended Installation	31
	5.1.3.2	Unattended Installation	41
	5.1.4	Firmware Upgrade	43
	5.2 Linux	x Driver Installation	43
	5.2.1	Prerequisites	43
	5.2.2	Downloading MLNX_OFED	44
	5.3 VMw	are Driver Installation	69
	5.3.1	Hardware and Software Requirements	69
	5.3.2	Installing NATIVE ESXi Driver for VMware vSphere	70
	5.3.3	Removing Earlier NVIDIA Drivers	70
	5.3.4	Firmware Programming	71
6	Updatin	g Adapter Firmware	72
7	Settina	Link Type of High-Speed Ports	73
•	_	config	
0			
8	i rouble	shooting	14

	8.1	General Troubleshooting	74
	8.2	General Troubleshooting Linux Troubleshooting	75
	8.3	Windows Troubleshooting	76
9	Spe	ecifications	78
	9.1	MCX75343AMC-NEAC / MCX75343AMS-NEAC Specifications	
	9.2	MCX753436MS-HEAB / MCX753436MC-HEAB Specifications	
	9.3	Mechanical Drawings and Dimensions	83
	9.3.		
	9.3.2	2 Bracket Drawings and Dimensions	84
10) N	Nonitoring	85
	10.1	Thermal Sensors	85
	10.2	Heatsink	85
11	l Fi	inding the MAC and Serial Number on the Adapter Card	86
12		·	
1 4	ע :	ocument Revision History	00

1 About This Manual

This is the User Guide for adapter cards based on the NVIDIA® ConnectX®-7 integrated circuit device for Open Compute Project Spec 3.0. These adapters' connectivity provides the highest-performing low latency and most flexible interconnect solution for servers supporting OCP spec 3.0 adapter cards used in Enterprise Data Centers and High-Performance Computing environments.

1.1 Ordering Part Numbers

The table below provides the ordering part numbers (OPN) for the available ConnectX-7 cards for OCP Spec 3.0.

NVIDIA SKU	Legacy OPN	Port Type	Supported Speed	OCP3.0 Form Factor	PCIe Express	Multi-Host and Socket Direct	Crypto	Secure Boot	Bracket Type	Lifecycle
900-9X760- 0018-MB2	MCX753436MC- HEAB	Dual-port QSFP112	InfiniBand: NDR200/HDR Ethernet: 200GbE (Default Mode)	Small Form Factor (SFF)	PCle 5.0 x16	√	✓	√	Thumbscrew (Pull Tab) Bracket	Mass Production
900-9X760- 0078-MB0	MCX753436MS- HEAB	Dual-port QSFP112	InfiniBand: NDR200/HDR Ethernet: 200GbE (Default Mode)	Small Form Factor (SFF)	PCIe 5.0 x16	√	-	√	Thumbscrew (Pull Tab) Bracket	Mass Production
900-9X7AX- 004NMC0	MCX75343AMC- NEAC	Single-port OSFP	InfiniBand: NDR (Default Mode) Ethernet: 400GbE	Tall Small Form Factor (TSFF)	PCle 5.0 x16	✓	√	✓	Thumbscrew (Pull Tab) TSFF Bracket	Mass Production
900-9X7AX- 003NMC0	MCX75343AMS- NEAC	Single-port OSFP	InfiniBand: NDR (Default Mode) Ethernet: 400GbE	Tall Small Form Factor (TSFF)	PCIe 5.0 x16	1	-	✓	Thumbscrew (Pull Tab) TSFF Bracket	Mass Production

1.2 Intended Audience

This manual is intended for the installer and user of these cards.

The manual assumes basic familiarity with InfiniBand and Ethernet network and architecture specifications.

1.3 Technical Support

Customers who purchased NVIDIA products directly from NVIDIA are invited to contact us through the following methods:

- URL: <u>www.nvidia.com</u> → Support
- E-mail: enterprisesupport@nvidia.com

Customers who purchased NVIDIA M-1 Global Support Services, please see your contract for details regarding Technical Support.

Customers who purchased NVIDIA products through an NVIDIA-approved reseller should first seek assistance through their reseller.

1.4 Related Documentation

MLNX_OFED for Linux User Manual and Release Notes	User Manual describing OFED features, performance, band diagnostic, tools content, and configuration. See MLNX_OFED for Linux Documentation .
WinOF-2 for Windows User Manual and Release Notes	User Manual describing WinOF-2 features, performance, Ethernet diagnostic, tools content, and configuration. See <u>WinOF-2 for Windows Documentation</u> .
NVIDIA VMware for Ethernet User Manual	User Manual and release notes describing the various components of the NVIDIA ConnectX® NATIVE ESXi stack. See

InfiniBand Architecture Specification Release 1.2.1, Vol 2 - Release 1.4, and Vol 2 - Release 1.5	InfiniBand Specifications
IEEE Std 802.3 Specification	IEEE Ethernet Specifications
Open Compute Project 3.0 Specifications	https://www.opencompute.org/
PCI Express Specifications	Industry Standard PCI Express Base and Card Electromechanical Specifications. Refer to PCI-SIG Specifications .
LinkX Interconnect Solutions	LinkX InfiniBand cables and transceivers are designed to maximize the performance of High-Performance Computing networks, requiring high-bandwidth, low-latency connections between compute nodes and switch nodes. NVIDIA offers one of the industry's broadest portfolio of QDR/FDR10 (40Gb/s), FDR (56Gb/s), EDR/HDR100 (100Gb/s), HDR (200Gb/s) and NDR (400Gb/s) cables, including Direct Attach Copper cables (DACs), copper splitter cables, Active Optical Cables (AOCs) and transceivers in a wide range of lengths from 0.5m to 10km. In addition to meeting IBTA standards, NVIDIA tests every product in an end-to-end environment ensuring a Bit Error Rate of less than 1E-15. Read more at LinkX Cables and Transceivers.

1.5 Document Conventions

When discussing memory sizes, MB and MBytes are used in this document to mean size in mega Bytes. The use of Mb or Mbits (small b) indicates the size in megabits. In this document, PCIe is used to mean PCI Express.

2 Introduction

2.1 Product Overview

This is the *User Guide* for InfiniBand/VPI adapter cards based on the ConnectX-7 integrated circuit device for OCP Spec 3.0. These adapters' connectivity provides the highest-performing and most flexible interconnect solution for servers supporting OCP 3.0 used in Enterprise Data Centers and High-Performance Computing environments. ConnectX-7 network adapters are offered in two form factors and various flavors: stand-up PCle and Open Compute Project (OCP) Spec 3.0 cards. This user manual covers the OCP 3.0 cards, for the low-profile PCle stand-up cards hardware user manual, please refer to <u>ConnectX-7 PCle Stand-up Cards User Manual</u>.

Providing up to two ports of NDR200 and 200GbE or a single port of NDR and 400GbE connectivity, and PCIe Gen 5.0 x16 host connectivity, ConnectX-7 is a member of NVIDIA's world-class, award-winning, ConnectX family of network adapters. Continuing NVIDIA's consistent innovation in networking, ConnectX-7 provides agility and efficiency at every scale. ConnectX-7 delivers cutting-edge performance and security for uncompromising data centers.

In addition to the Small Form Factor (SFF) form factor, ConnectX-7 for OCP 3.0 cards are available in the newly added Tall-SFF (TSFF) spec form factor, taking into account the added height of the card to allow better thermal performance. Contact NVIDIA for more details.

2.2 System Requirements

Item	Description
PCI Express slot	In PCIe x16 Configuration PCIe Gen 5.0/4.0 (16 GT/s / 32GT/s) through x16 edge connector.
System Power Supply	Refer to Specifications
Operating System	 In-box drivers for major operating systems: Linux: RHEL, Ubuntu Windows Virtualization and containers VMware ESXi (SR-IOV) Kubernetes OpenFabrics Enterprise Distribution (OFED) OpenFabrics Windows Distribution (WinOF-2)
Connectivity	 Interoperable with 1/10/25/40/50/100/200/400 Gb/s Ethernet switches and SDR/DDR/EDR/HDR100/HDR/NDR200/NDR InfiniBand switches Passive copper cable with ESD protection Powered connectors for optical and active cable support

2.3 Package Contents

Category	Qty	Item
Cards	1	ConnectX-7 adapter card for OCP 3.0
Accessories	1	Thumbscrew (Pull Tab) Bracket

For more detailed information, see **Specifications**.

2.4 Features and Benefits

This section describes hardware features and capabilities. Please refer to the relevant driver and/or firmware release notes for feature availability.

Feature	Description						
InfiniBand Architecture Specification v1.5 compliant	ConnectX-7 delivers low latency, high bandwidth, and computing efficiency for high-performance computing (HPC), artificial intelligence (AI), and hyperscale cloud data center applications. ConnectX-7 is InfiniBand Architecture Specification v1.5 compliant. InfiniBand Network Protocols and Rates:						
	Protocol	Standard	Rate	(Gb/s)	Comments		
			4x Port (4 Lanes)	2x Ports (2 Lanes)			
	NDR/NDR200	IBTA Vol2 1.5	425	212.5	PAM4 256b/257b encoding and RS-FEC		
	HDR/HDR100	IBTA Vol2 1.4	212.5	106.25	PAM4 256b/257b encoding and RS-FEC		
	EDR	IBTA Vol2 1.3.1	103.125	51.5625	NRZ 64b/66b encoding		
	FDR	IBTA Vol2 1.2	56.25	N/A	NRZ 64b/66b encoding		
Up to 400Gb/s Ethernet	ConnectX-7 adapter cards comply with the following IEEE 802.3 standards: 400GbE / 200GbE / 100GbE / 50GbE / 40GbE / 25GbE / 10GbE						
	Protocol			MAC Rate	MAC Rate		
	IEEE802.3ck			100/200/400Gb/s Gigabit Ethernet (Include ETC enhancement)			
	IEEE802.3cd			50/100/200/400Gb/s Gigabit Ethernet			
	IEEE802.3bs			(Include ETC enhancement)			
		IEEE802.3cm					
	IEEE802.3cn						
	IEEE802.3cu						

Feature	Description	
	IEEE 802.3bj IEEE 802.3bm	100 Gigabit Ethernet
	IEEE 802.3by Ethernet Consortium25	25/50 Gigabit Ethernet
	IEEE 802.3ba	40 Gigabit Ethernet
	IEEE 802.3ae	10 Gigabit Ethernet
	IEEE 802.3cb	2.5/5 Gigabit Ethernet (For 2.5: support only 2.5 x1000BASE-X)
	IEEE 802.3ap	Based on auto-negotiation and KR startup
	IEEE 802.3ad IEEE 802.1AX	Link Aggregation
	IEEE 802.1Q IEEE 802.1P VLAN tags and priority	
	IEEE 802.1Qau (QCN) Congestion Notification IEEE 802.1Qaz (ETS) EEE 802.1Qbb (PFC) IEEE 802.1Qbg IEEE 1588v2 IEEE 802.1AE (MACSec) Jumbo frame support (9.6KB)	
 SPI - includes 256Mbit SPI Quad Flash device. FRU EEPROM - Stores the parameters and personality of the card. The EEPROM capacity is 128I FRU I2C address is (0x50) and is accessible through the PCIe SMBus. (Note: Address 0x58 is reset of the card of the card. The EEPROM capacity is 128I FRU I2C address is (0x50) and is accessible through the PCIe SMBus. (Note: Address 0x58 is reset of the card of the card. The EEPROM capacity is 128I FRU I2C address is (0x50) and is accessible through the PCIe SMBus. (Note: Address 0x58 is reset of the card of the card. The EEPROM capacity is 128I FRU I2C address over logical through the PCIe SMBus. (Note: Address 0x58 is reset of the card of the card of the card. The EEPROM capacity is 128I FRU I2C address over logical through the PCIe SMBus. (Note: Address 0x58 is reset of the card of		
		apsulated formats such as NVGRE and VXLAN. While this solves

Feature	Description
	CPU. ConnectX-7 effectively addresses this by providing advanced NVGRE and VXLAN hardware offloading engines that encapsulate and de-capsulate the overlay protocol.
Quality of Service (QoS)	Support for port-based Quality of Service enabling various application requirements for latency and SLA.
Hardware-based I/O Virtualization	ConnectX-7 provides dedicated adapter resources and guaranteed isolation and protection for virtual machines within the server.
Storage Acceleration	A consolidated compute and storage network achieves significant cost-performance advantages over multi-fabric networks. Standard block and file access protocols can leverage: RDMA for high-performance storage access NVMe over Fabric offloads for the target machine NVMe over TCP acceleration
SR-IOV	ConnectX-7 SR-IOV technology provides dedicated adapter resources and guaranteed isolation and protection for virtual machines (VM) within the server.
High-Performance Accelerations	 Collective operations offloads Vector collective operations offloads MPI tag matching MPI_Alltoall offloads Rendezvous protocol offload
RDMA Message Rate	330-370 million messages per second.
Secure Boot	The secure boot process assures booting of authentic firmware/software that is intended to run on ConnectX-7. This is achieved using cryptographic primitives using asymmetric cryptography. ConnectX-7 supports several cryptographic functions in its HW Root-of-Trust (RoT) that has its key stored in on-chip FUSES.
Secure Firmware Update	The Secure firmware update feature enables a device to verify digital signatures of new firmware binaries to ensure that only officially approved versions can be installed from the host, the network, or a Board Management Controller (BMC). The firmware of devices with "secure firmware update" functionality (secure FW), restricts access to specific commands and registers that can be used to modify the firmware binary image on the flash, as well as commands that can jeopardize security in general. For further information, refer to the MFT User Manual.

Feature	Description
Advanced storage capabilities	Block-level encryption and checksum offloads.
Host Management	ConnectX-7 technology maintains support for host manageability through a BMC. ConnectX-7 PCle stand-up adapter can be connected to a BMC using MCTP over SMBus or MCTP over PCle protocols as if it is a standard NVIDIA PCle stand-up adapter card. For configuring the adapter for the specific manageability solution in use by the server, please contact NVIDIA Support. • Protocols: PLDM, NCSI • Transport layer – RBT, MCTP over SMBus and MCTP over PCle • Physical layer: SMBus 2.0 / I2C interface for device control and configuration, PCle • PLDM for Monitor and Control DSP0248 • PLDM for Firmware Update DSP026 • IEEE 1149.6 • Secured FW update • FW Recovery • NIC reset • Monitoring and control • Network port settings • Boot setting
RDMA and RDMA over Converged Ethernet (RoCE)	ConnectX-7, utilizing IBTA RDMA (Remote Data Memory Access) and RoCE (RDMA over Converged Ethernet) technology, delivers low-latency and high-performance over InfiniBand and Ethernet networks. Leveraging datacenter bridging (DCB) capabilities as well as ConnectX-7 advanced congestion control hardware mechanisms, RoCE provides efficient low-latency RDMA services over Layer 2 and Layer 3 networks.
NVIDIA PeerDirect™	PeerDirect™ communication provides high-efficiency RDMA access by eliminating unnecessary internal data copies between components on the PCIe bus (for example, from GPU to CPU), and therefore significantly reduces application run time. ConnectX-7 advanced acceleration technology enables higher cluster efficiency and scalability to tens of thousands of nodes.
CPU Offload	 Adapter functionality enables reduced CPU overhead allowing more available CPU for computation tasks. Flexible match-action flow tables Open VSwitch (OVS) offload using ASAP^{2®}

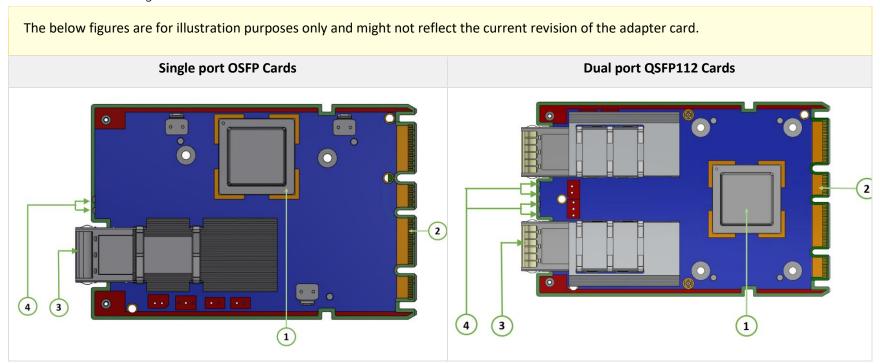
Feature	Description	
	Tunneling encapsulation/decapsulation	
NVIDIA Multi-Host	NVIDIA® Multi-Host technology enables next-generation Cloud, Web 2.0 and high-performance data centers to design and build new scale-out heterogeneous compute and storage racks with direct connectivity between multiple hosts and the centralized network controller. This enables direct data access with the lowest latency to significantly improve densities and maximizes data transfer rates. For more information, please visit NVIDIA Multi-Host Solutions .	

3 Supported Interfaces

This section describes the ConnectX-7 supported interfaces. Each numbered interface that is referenced in the figures is described in the following table with a link to detailed information.

3.1 ConnectX-7 Layout and Interface Information

The below figures show the component side of the NVIDIA ConnectX-7 adapter card. Each numbered interface that is referenced in the figures is described in the following table with a link to detailed information.



Item	Interface	Description	
1	ConnectX-7 IC	ConnectX-7 Integrated Circuit	
2	PCI Express Interface	PCIe Gen 5.0 through x16 edge connector	
3	Networking Interfaces	Network traffic is transmitted through the adapter card QSFP112/OSFP connectors. The networking connectors allow for the use of modules, optical and passive cable interconnect solutions.	
4	Networking Ports LEDs	Two I/O LEDs per port to indicate speed and link status	
	FRU EEPROM	FRU EEPROM capacity 16KB	
	SMBus Interface	Allows BMC connectivity using MCTP over SMBus or MCTP over PCIe protocols.	
	<u>Voltage</u> <u>Regulators</u>	Voltage supply pins that feed onboard regulators.	
	CPLD Interface	Controls the networking port logic LED (LED0) and implements the OCP 3.0 host scan chain.	
	<u>Heatsink</u>	Dissipates the heat.	

3.2 Interfaces Detailed Description

3.2.1 ConnectX-7 IC

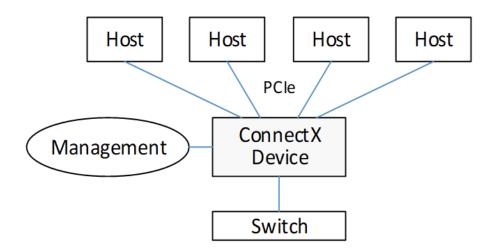
The ConnectX-7 family of adapter IC devices delivers two ports of NDR200/200GbE) or a single-port of NDR/400GbE connectivity paired with best-inclass hardware capabilities that accelerate and secure cloud and data-center workloads.

3.2.1.1 NVIDIA Multi-HostTM Support

In addition to building exceptionally high bandwidth to the data center, the ConnectX-7 device enables leveraging this speed across the entire data center utilizing its NVIDIA Multi-Host feature.

Using its 16-lane PCI Express interface, a single ConnectX-7 device can provide 400GbE interconnect for up to four independent hosts without any performance degradation.

The figure below shows a ConnectX-7 device with NVIDIA Multi-Host connected to four separate hosts, each with a PCIe x4 interface, on one side to a switch on the other side.



The below bifurcation are optional for the adapter's x16 PCIe interface:

• x1 PCle x16, x1 PCle x8, x1 PCle x4

- x2 PCle x8, x2 PCle x4, x2 PCle x2, x2 PCle x1
- x4 PCle x4, x4 PCle x2, x4 PCle x1

Multi-host capable cards also support Socket-Direct applications and work as regular Single-Host cards, depending on the type of server they are plugged into, assuming the server complies with the OCP 3.0 spec.

According to the OCP 3.0 spec, the adapter card advertises its capability through the PRSNTB[3:0]# pins. The server determines the configuration through the BIF[2:0]# pins, which it drives to the adapter card.

The **NVIDIA OCP3.0** card has an internal logic that uses the BIF[2:0]# data and determines the correct operating mode to boot at. The combination of the PRSNTB[3:0]# and BIF[2:0]# pins deterministically sets the PCIe lane width for a given combination of OCP 3.0 cards and baseboard. The logic and the decoding table can be found in the OCP 3.0 spec (Chapter **3.5 PCIe Bifurcation Mechanism)**.

For example:

the NVIDIA OCP 3.0 Multi-host adapter drives 0100 on PRSNTB[3:0]# to the server.

If Server Drivers	Adapter PCIe Mode
000	Single-Host Mode: x1 PCle x16
001	Socket Direct Mode: x2 PCIe x8
010	Socket Direct Mode: x4 PCle x4
101	Multi-Host Mode: x2 PCIe x8
110	Multi-Host Mode: x4 PCIe x4

3.2.2 PCI Express Interface

The table below describes the supported PCIe interface in ConnectX-7 OCP 3.0 adapter cards.

- PCle Gen 5.0 compliant, 4.0, 3.0, 2.0 and 1.1 compatible
- 2.5, 5.0, 8.0, 16.0 and 32GT/s link rate x16
- Support for PCIe bifurcation: Auto-negotiates to x16, x8, x4, x2, or x1
- NVIDIA Multi-Host™ supports connection of up to 4x hosts
- Transaction layer packet (TLP) processing hints (TPH)

- PCle switch Downstream Port Containment (DPC)
- Advanced error reporting (AER)
- Access Control Service (ACS) for peer-to-peer secure communication
- Process Address Space ID (PASID)
- Address translation services (ATS)
- Support for MSI/MSI-X mechanisms
- Support for SR-IOV

3.2.3 Networking Interfaces

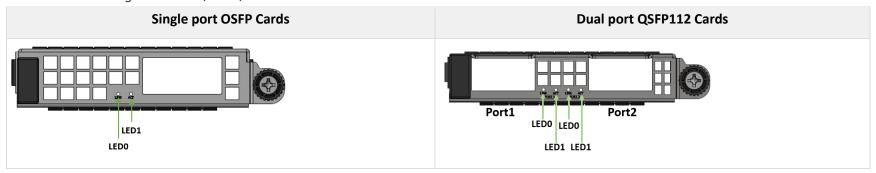
The adapter card includes special circuits to protect from ESD shocks to the card/server when plugging copper cables.

Protocol	Specifications	
	The network ports are compliant with the IEEE 802.3 Ethernet standards listed in <u>Features and Benefits</u> . Ethernet traffic is transmitted through the networking connectors on the adapter card.	
	The network ports are compliant with the <i>InfiniBand Architecture Specification, Release 1.5.</i> InfiniBand traffic is transmitted through the cards' networking connectors.	

3.2.4 Networking Ports LEDs Specifications

There are two I/O LEDs per port to indicate port speed and link status.

- LEDO is a bi-color LED (Yellow and Green)
- LED2 is a single-color LED (Green)



State	· · · · · ·		Single Color LED (Green)	
Beacon command for locating the adapter card	1Hz blinking Yellow			OFF
Error	4Hz blinking Yellow Indicates an error with the link. The error can be one of the following:			ON
	Error Type	Description	LED Behavior	
	I ² C	I ² C access to the networking ports fails	Blinks until error is fixed	
	Over- current	Over-current condition of the networking ports	Blinks until error is fixed	
Link Up	In full port speed: the Green LED is solid In less than full port speed: the Yellow LED is solid		ON	
Physical Activity	The Green LED will blink.		Blinking	

3.2.5 FRU EEPROM

FRU EEPROM allows the baseboard to identify different types of OCP 3.0 cards. FRU EEPROM is accessible through SMCLK and SMDATA. FRU EEPROM address is defined according to SLOT_ID0 and SLOT_ID1 and its capacity is 16KB.

3.2.6 SMBus Interface

ConnectX-7 technology maintains support for manageability through a BMC. ConnectX-7 OCP 3.0 adapter can be connected to a BMC using MCTP over SMBus or MCTP over PCIe protocols as if it is a standard NVIDIA OCP 3.0 adapter. For configuring the adapter for the specific manageability solution in use by the server, please contact NVIDIA Support.

3.2.7 Voltage Regulators

The adapter card incorporates a CPLD device that implements the OCP 3.0 host scan chain and controls the networking port logic LED (LEDO). It draws its power supply from the 3.3V_EDGE and 12V_EDGE rails.

4 Hardware Installation

Installation and initialization of ConnectX-7 adapter cards for OCP Spec 3.0 require attention to the mechanical attributes, power specifications, and precautions for electronic equipment.

4.1 Safety Warnings

Safety warnings are provided here in the English language. For safety warnings in other languages, refer to <u>Adapter Installation Safety Instructions</u>.

4.2 Installation Procedure Overview

The installation procedure involves the following steps:

Step	Procedure	Direct Link
1	Check the system's hardware and software requirements.	Refer to <u>System Requirements</u>
2	Pay attention to the airflow consideration within the host system	Refer to <u>Airflow Requirements</u>
3	Follow the safety precautions	Refer to <u>Safety Precautions</u>
4	Follow the pre-installation checklist	Refer to Pre-Installation Checklist
5	(Optional) Replace the assembled OCP 3.0 bracket with the desired form factor bracket	Refer to OCP 3.0 Bracket Replacement Instructions
6	Install the adapter card in the system	Refer to <u>Installation Instructions</u>
7	Connect cables or modules to the card	Refer to <u>Cables and Modules</u>
8	Identify the adapter card in the system	Refer to <u>Identify the Card in Your System</u>

4.2.1 System Requirements

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 operating environment should meet severity level G1 as per ISA 71.04 for gaseous contamination and ISO 14644-1 class 8 for cleanliness level.

4.2.1.1 Hardware Requirements

For proper operation and performance, please make sure to use a PCle slot with a corresponding bus width and that can supply sufficient power to your card. Refer to the Specifications section of the manual for more power requirements.

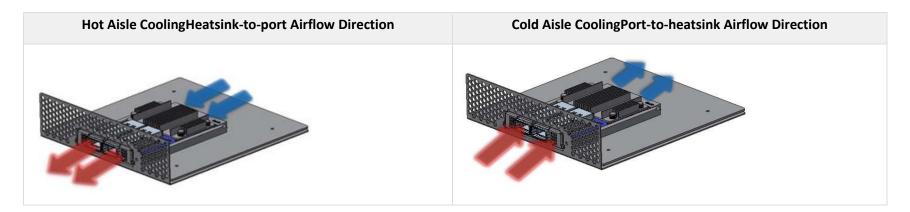
A system with a PCI Express x16 slot for OCP spec 3.0 is required for installing the card.

4.2.1.2 Airflow Requirements

ConnectX-7 adapter cards are offered with two airflow patterns: from the heatsink to the network ports, and vice versa, as shown below.

Please refer to the "Specifications" chapter for airflow numbers for each specific card model.

All cards in the system should be planned with the same airflow direction.



4.2.1.3 Software Requirements

- See Operating Systems/Distributions section under the Introduction section.
- Software Stacks NVIDIA OpenFabric software package MLNX_OFED for Linux, WinOF-2 for Windows, and VMware. See the <u>Driver Installation</u> section.

4.2.2 Safety Precautions

The adapter is being installed in a system that operates with voltages that can be lethal. Before opening the case of the system, observe the following precautions to avoid injury and prevent damage to system components.

- 1. Remove any metallic objects from your hands and wrists.
- 2. Make sure to use only insulated tools.
- 3. Verify that the system is powered off and is unplugged.
- 4. It is strongly recommended to use an ESD strap or other antistatic devices.

4.2.3 Pre-Installation Checklist

1. **Unpack the ConnectX-7 adapter card.** Unpack and remove the ConnectX-7 card. Check the parts for visible damage that may have occurred during shipping. Please note that the cards must be placed on an antistatic surface.

Please note that if the card is removed hastily from the antistatic bag, the plastic ziplock may harm the EMI fingers on the networking connector. Carefully remove the card from the antistatic bag to avoid damaging the EMI fingers.

2. Shut down your system if active:

Turn off the power to the system, and disconnect the power cord. Refer to the system documentation for instructions. Before you install the ConnectX-7 card, make sure that the system is disconnected from power.

4.2.4 Cables and Modules

To obtain the list of supported NVIDIA cables for your adapter, please refer to the Cables Reference Table at Networking Configuration Tools.

Cable Installation

- 1. All cables can be inserted or removed with the unit powered on.
- 2. To insert a cable, press the connector into the port receptacle until the connector is firmly seated.
 - a. Support the weight of the cable before connecting the cable to the adapter card. Do this by using a cable holder or tying the cable to the rack.
 - b. Determine the correct orientation of the connector to the card before inserting the connector. Do not try and insert the connector upside down. This may damage the adapter card.
 - c. Insert the connector into the adapter card. Be careful to insert the connector straight into the cage. Do not apply any torque, up or down, to the connector cage in the adapter card.
 - d. Make sure that the connector locks in place.

When installing cables make sure that the latches engage.

Always install and remove cables by pushing or pulling the cable and connector in a straight line with the card.

- 3. After inserting a cable into a port, the Yellow or Green LEDO indicator will light 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). See Adapter Card LED Operations.
- 4. After plugging in a cable, lock the connector using the latching mechanism particular to the cable vendor. When a logical connection is made, Green LED1 will light. When data is being transferred, Green LED1 will blink.
- 5. Care should be taken as not to impede the air exhaust flow through the ventilation holes. Use cable lengths which allow for routing horizontally around to the side of the chassis before bending upward or downward in the rack.
- 6. To remove a cable, disengage the locks and slowly pull the connector away from the port receptacle. LED indicator will turn off when the cable is unseated.

To uninstall the adapter card, see Uninstalling the Card.

4.2.5 Identifying the Card in Your System

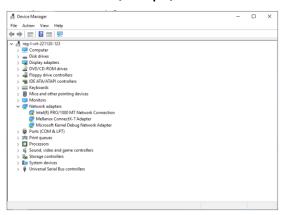
On Linux

Get the device location on the PCI bus by running Ispci and locating lines with the string "Mellanox Technologies":

lspci |grep -i Mellanox Network controller: Mellanox Technologies MT2910 Family [ConnectX-7]

On Windows

- Open Device Manager on the server. Click Start => Run, and then enter devmgmt.msc.
- 2. Expand **System Devices** and locate your NVIDIA ConnectX-7 adapter card.
- 3. Right click the mouse on your adapter's row and select **Properties** to display the adapter card properties window.
- 4. Click the **Details** tab and select **Hardware Ids** (Windows 2022/2019/2016/2012 R2) from the **Property** pull-down menu. **PCI Device (Example)**



1. In the **Value** display box, check the fields VEN and DEV (fields are separated by '&'). In the display example above, notice the sub-string "PCI\VEN_15B3&DEV_1021": VEN is equal to 0x15B3 – this is the Vendor ID of Mellanox Technologies; and DEV is equal to 1021 (for ConnectX-7) – this is a valid NVIDIA PCI Device ID.

If the PCI device does not have an NVIDIA adapter ID, return to Step 2 to check another device.

The list of NVIDIA PCI Device IDs can be found at the PCI ID repository.

4.2.6 Extraction Instructions

Unable to render include or excerpt-include. Could not retrieve page.

5 Driver Installation

Please use the relevant driver installation section.

- Windows Driver Installation
- Linux Driver Installation
- VMware Driver Installation

5.1 Windows Driver Installation

For Windows, download and install the latest WinOF-2 for Windows software package available via the NVIDIA website at: <u>WinOF-2 webpage</u>. Follow the installation instructions included in the download package (also available from the download page).

The snapshots in the following sections are presented for illustration purposes only. The installation interface may slightly vary, depending on the operating system in use.

5.1.1 Software Requirements

Description	Package
Windows Server 2022	MLNX_WinOF2- <version>_All_x64.exe</version>
Windows Server 2019	
Windows Server 2016	
Windows Server 2012 R2	
Windows 11 Client (64 bit only)	
Windows 10 Client (64 bit only)	

Description	Package
Windows 8.1 Client (64 bit only)	

Note: The Operating System listed above must run with administrator privileges.

5.1.2 Downloading WinOF-2 Driver

To download the .exe file according to your Operating System, please follow the steps below:

- 1. Obtain the machine architecture.
 - a. To go to the Start menu, position your mouse in the bottom-right corner of the Remote Desktop of your screen.
 - b. Open a CMD console (Click Task Manager-->File --> Run new task and enter CMD).
 - c. Enter the following command.

cho %PROCESSOR_ARCHITECTURE%

On an x64 (64-bit) machine, the output will be "AMD64".

- 2. Go to the WinOF-2 web page at: https://www.nvidia.com/en-us/networking/ > Products > Software > InfiniBand Drivers (Learn More) > Nvidia WinOF-2.
- 3. Download the .exe image according to the architecture of your machine (see <u>Step 1</u>). The name of the .exe is in the following format: MLNX_WinOF2-<version>_<arch>.exe.

Installing the incorrect .exe file is prohibited. If you do so, an error message will be displayed.

For example, if you install a 64-bit .exe on a 32-bit machine, the wizard will display the following (or a similar) error message: "The installation package is not supported by this processor type. Contact your vendor"

5.1.3 Installing WinOF-2 Driver

The snapshots in the following sections are for illustration purposes only. The installation interface may slightly vary, depending on the used operating system.

This section provides instructions for two types of installation procedures, and both require administrator privileges:

- Attended Installation
 An installation procedure that requires frequent user intervention.
- <u>Unattended Installation</u>
 An automated installation procedure that requires no user intervention.

5.1.3.1 Attended Installation

The following is an example of an installation session.

- 1. Double click the .exe and follow the GUI instructions to install MLNX_WinOF2.
- 2. [Optional] Manually configure your setup to contain the logs option (replace "LogFile" with the relevant directory).

```
ILNX_WinOF2_<revision_version>_All_Arch.exe /v"/l*vx [LogFile]"
```

3. [Optional] If you do not want to upgrade your firmware version (i.e., MT_SKIPFWUPGRD default value is False).

```
ILNX_WinOF2_<revision_version>_All_Arch.exe /v" MT_SKIPFWUPGRD=1"
```

4. [Optional] If you do not want to install the Rshim driver, run.

```
ILNX_WinOF2_<revision_version>_All_Arch.exe /v" MT_DISABLE_RSHIM_INSTALL=1"
```

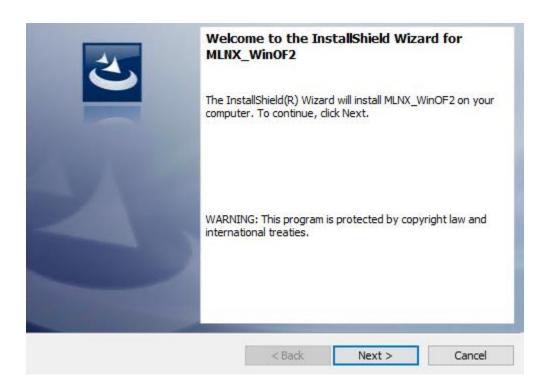
The Rshim driver installanion will fail if a prior Rshim driver is already installed. The following fail message will be displayed in the log:

"ERROR!!! Installation failed due to following errors: MlxRshim drivers installation disabled and MlxRshim drivers Installed, Please remove the following oem inf files from driver store: <oem inf list>"

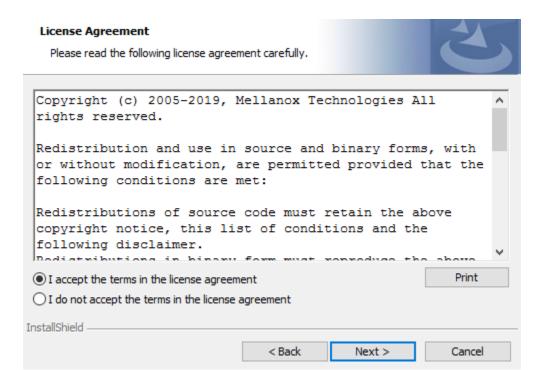
5. **[Optional]** If you want to skip the check for unsupported devices, run.

LNX_WinOF2_<revision_version>_All_Arch.exe /v" SKIPUNSUPPORTEDDEVCHECK=1"

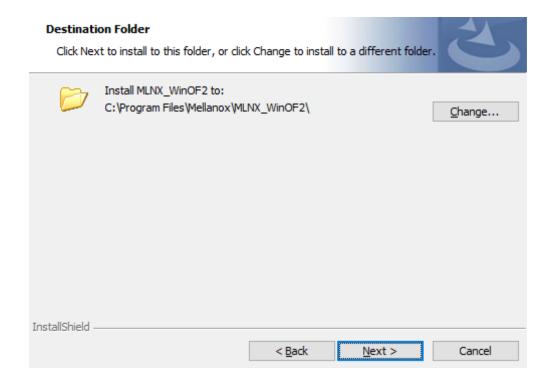
6. Click Next in the Welcome screen.



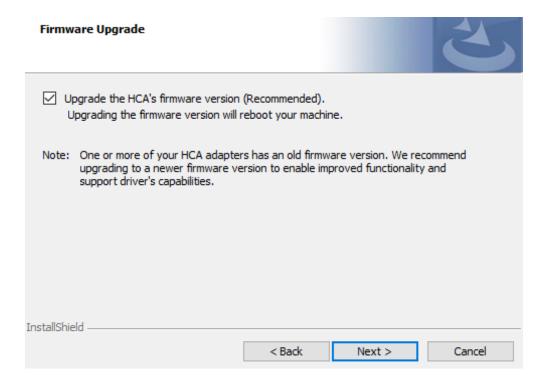
7. Read and accept the license agreement and click Next.



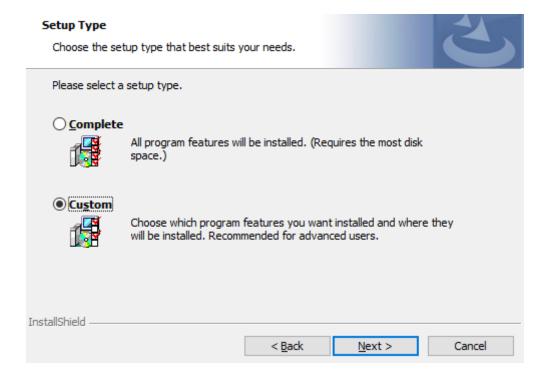
8. Select the target folder for the installation.



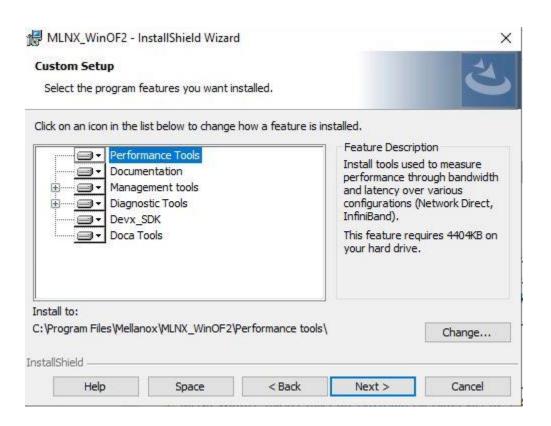
- 9. The firmware upgrade screen will be displayed in the following cases:
 - o If the user has an OEM card. In this case, the firmware will not be displayed.
 - o If the user has a standard NVIDIA® card with an older firmware version, the firmware will be updated accordingly. However, if the user has both an OEM card and a NVIDIA® card, only the NVIDIA® card will be updated.



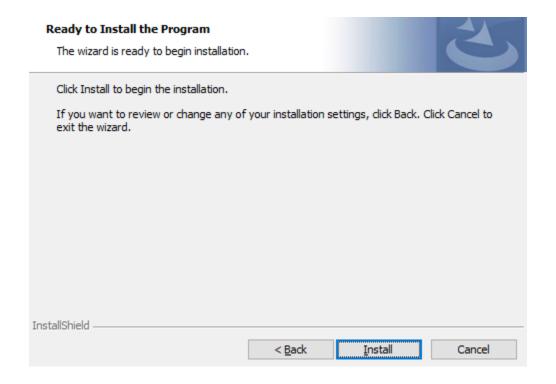
10. Select a Complete or Custom installation, follow <u>Step a</u> onward.



- a. Select the desired feature to install:
 - Performances tools install the performance tools that are used to measure performance in user environment
 - Documentation contains the User Manual and Release Notes
 - Management tools installation tools used for management, such as mlxstat
 - Diagnostic Tools installation tools used for diagnostics, such as mlx5cmd
- b. Click Next to install the desired tools.



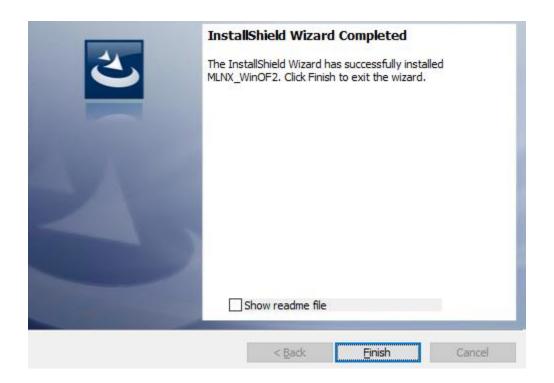
11. Click Install to start the installation.



12. In case firmware upgrade option was checked in <u>Step 7</u>, you will be notified if a firmware upgrade is required (see **1**).



13. Click Finish to complete the installation.



5.1.3.2 Unattended Installation

If no reboot options are specified, the installer restarts the computer whenever necessary without displaying any prompt or warning to the user.

To control the reboots, use the */norestart* or */forcerestart* standard command-line options.

The following is an example of an unattended installation session.

- 1. Open a CMD console-> Click Start-> Task Manager File-> Run new task-> and enter CMD.
- 2. Install the driver. Run:

```
ILNX_WinOF2-[Driver/Version]_<revision_version>_All_-Arch.exe /S /v/qn
```

3. **[Optional]** Manually configure your setup to contain the logs option:

```
ILNX_WinOF2-[Driver/Version]_<revision_version>_All_-Arch.exe /S /v/qn /v"/l*vx
[LogFile]"
```

4. [Optional] if you wish to control whether to install ND provider or not (i.e., MT_NDPROPERTY default value is True).

```
LNX_WinOF2-[Driver/Version]_<revision_version>_All_Arch.exe /vMT_NDPROPERTY=1
```

5. [Optional] If you do not wish to upgrade your firmware version (i.e., MT_SKIPFWUPGRD default value is False).

```
ILNX_WinOF2-[Driver/Version]_<revision_version>_All_Arch.exe /vMT_SKIPFWUPGRD=1
```

6. [Optional] If you do not want to install the Rshim driver, run.

```
ILNX_WinOF2_<revision_version>_All_Arch.exe /v" MT_DISABLE_RSHIM_INSTALL=1"
```

The Rshim driver installanion will fail if a prior Rshim driver is already installed. The following fail message will be displayed in the log:

"ERROR!!! Installation failed due to following errors: MlxRshim drivers installation disabled and MlxRshim drivers Installed, Please remove the following oem inf files from driver store: <oem inf list>"

7. **[Optional]** If you want to enable the default configuration for Rivermax, run.

```
MLNX_WinOF2_<revision_version>_All_Arch.exe /v"MT_RIVERMAX=1 /l*vx
C:\Users\<user>\log.txt "
```

8. [Optional] If you want to skip the check for unsupported devices, run/

```
MLNX_WinOF2_<revision_version>_All_Arch.exe /v" SKIPUNSUPPORTEDDEVCHECK=1"
```

5.1.4 Firmware Upgrade

If the machine has a standard NVIDIA® card with an older firmware version, the firmware will be automatically updated as part of the NVIDIA® WinOF-2 package installation. For information on how to upgrade firmware manually, please refer to MFT User Manual.

If the machine has a DDA (pass through) facility, firmware update is supported only in the Host. Therefore, to update the firmware, the following must be performed:

- 1. Return the network adapters to the Host.
- Update the firmware according to the steps in the MFT User Manual.
- 3. Attach the adapters back to VM with the DDA tools.

5.2 Linux Driver Installation

This section describes how to install and test the MLNX_OFED for Linux package on a single server with a ConnectX-7 adapter card installed.

5.2.1 Prerequisites

Requirements	Description
Platforms	A server platform with a ConnectX-7 InfiniBand/Ethernet adapter card installed.
Required Disk Space for Installation	1GB

Requirements	Description
Operating System	Linux operating system. For the list of supported operating system distributions and kernels, please refer to the MLNX_OFED Release Notes.
Installer Privileges	The installation requires administrator (root) privileges on the target machine.

5.2.2 Downloading MLNX_OFED

1. Verify that the system has a network adapter installed by running lspci command. The below table provides output examples per ConnectX-7 card configuration.

ConnectX-7 Card Configuration	Output Examples	
Single-port PCIe x16 Card	[root@mftqa-009 ~]# lspci grep mellanox -ia 3:00.0 Infiniband controller: Mellanox Technologies MT2910 Family [ConnectX-7]	
Dual-port PCIe x16 Card	[root@mftqa-009 ~]# lspci grep mellanox -ia 86:00.0 Network controller: Mellanox Technologies MT2910 Family [ConnectX-7] 86:00.1 Network controller: Mellanox Technologies MT2910 Family [ConnectX-7]	

2. Download the ISO image to your host.

The image's name has the format MLNX_OFED_LINUX-<ver>-<0S label><CPU arch>.iso.

You can download and install the latest OpenFabrics Enterprise Distribution (OFED) software package available via the NVIDIA web site at $\underline{\text{nvidia.com/en-us/networking}} \rightarrow \text{Products} \rightarrow \text{Software} \rightarrow \text{InfiniBand Drivers} \rightarrow \underline{\text{NVIDIA MLNX_OFED}}$

a.

- i. Scroll down to the Download wizard, and click the Download tab.
- ii. Choose your relevant package depending on your host operating system.
- iii. Click the desired ISO/tgz package.

- iv. To obtain the download link, accept the End User License Agreement (EULA).
- 3. Use the Hash utility to confirm the file integrity of your ISO image. Run the following command and compare the result to the value provided on the download page.

```
HA256 MLNX_OFED_LINUX-<ver>-<OS label>.iso
```

Installing MLNX_OFED

Installation Script

The installation script, mlnxofedinstall, performs the following:

- Discovers the currently installed kernel
- o Uninstalls any software stacks that are part of the standard operating system distribution or another vendor's commercial stack
- o Installs the MLNX_OFED_LINUX binary RPMs (if they are available for the current kernel)
- o Identifies the currently installed InfiniBand and Ethernet network adapters and automatically upgrades the firmware **Note**: To perform a firmware upgrade using customized firmware binaries, a path can be provided to the folder that contains the firmware binary files, by running --fw-image-dir. Using this option, the firmware version embedded in the MLNX_OFED package will be ignored.

Example:

```
/mlnxofedinstall --fw-image-dir /tmp/my_fw_bin_files
```

If the driver detects unsupported cards on the system, it will abort the installation procedure. To avoid this, make sure to add --skip-unsupported-devices-check flag during installation.

Usage

/mnt/mlnxofedinstall [OPTIONS]

The installation script removes all previously installed OFED packages and re-installs from scratch. You will be prompted to acknowledge the deletion of the old packages.

Pre-existing configuration files will be saved with the extension ".conf.rpmsave".

- o If you need to install OFED on an entire (homogeneous) cluster, a common strategy is to mount the ISO image on one of the cluster nodes and then copy it to a shared file system such as NFS. To install on all the cluster nodes, use cluster-aware tools (suchaspdsh).
- o If your kernel version does not match with any of the offered pre-built RPMs, you can add your kernel version by using the "mlnx_add_kernel_support.sh" script located inside the MLNX_OFED package.

On Redhat and SLES distributions with errata kernel installed there is no need to use the mlnx_add_kernel_support.sh script. The regular installation can be performed and weak-updates mechanism will create symbolic links to the MLNX OFED kernel modules.

If you regenerate kernel modules for a custom kernel (using --add-kernel-support), the packages installation will not involve automatic regeneration of the initramfs. In some cases, such as a system with a root filesystem mounted over a ConnectX card, not regenerating the initramfs may even cause the system to fail to reboot.

In such cases, the installer will recommend running the following command to update the initramfs:

dracut -f

On some OSs, dracut -f might result in the following error message which can be safely ignore.

libkmod: kmod_module_new_from_path: kmod_module 'mdev' already
exists with different path

The "mlnx_add_kernel_support.sh" script can be executed directly from the mlnxofedinstall script. For further information, please see '-- add-kernel-support' option below.

On Ubuntu and Debian distributions drivers installation use Dynamic Kernel Module Support (DKMS) framework. Thus, the drivers' compilation will take place on the host during MLNX_OFED installation.

Therefore, using "mlnx_add_kernel_support.sh" is irrelevant on Ubuntu and Debian distributions.

Example: The following command will create a MLNX_OFED_LINUX ISO image for RedHat 7.3 under the /tmp directory.

```
./MLNX_OFED_LINUX-x.x-x-rhel7.3-x86_64/mlnx_add_kernel_support.sh -m
/tmp/MLNX_OFED_LINUX-x.x-x-rhel7.3-x86_64/ --make-tgz

Iote: This program will create MLNX_OFED_LINUX TGZ for rhel7.3 under /tmp
directory.

Il Mellanox, OEM, OFED, or Distribution IB packages will be removed.

No you want to continue?[y/N]:y
see log file /tmp/mlnx_ofed_iso.21642.log

Building OFED RPMs. Please wait...

Removing OFED RPMs...

Treated /tmp/MLNX_OFED_LINUX-x.x-x-rhel7.3-x86_64-ext.tgz
```

- o The script adds the following lines to /etc/security/limits.conf for the userspace components such as MPI:
 - * soft memlock unlimited
 - * hard memlock unlimited

• These settings set the amount of memory that can be pinned by a userspace application to unlimited. If desired, tune the value unlimited to a specific amount of RAM.

For your machine to be part of the InfiniBand/VPI fabric, a Subnet Manager must be running on one of the fabric nodes. At this point, OFED for Linux has already installed the OpenSM Subnet Manager on your machine.

For the list of installation options, run:

```
/mlnxofedinstall --h
```

Installation Procedure

This section describes the installation procedure of MLNX_OFED on NVIDIA adapter cards.

- a. Log in to the installation machine as root.
- b. Mount the ISO image on your machine.

```
ost1# mount -o ro,loop MLNX_OFED_LINUX-<ver>-<OS label>-<CPU arch>.iso /mnt
```

c. Run the installation script.

```
mnt/mlnxofedinstall
ogs dir: /tmp/MLNX_OFED_LINUX-x.x-x.logs
his program will install the MLNX_OFED_LINUX package on your machine.
lote that all other Mellanox, OEM, OFED, RDMA or Distribution IB packages will be removed.
hose packages are removed due to conflicts with MLNX_OFED_LINUX, do not reinstall them.
starting MLNX_OFED_LINUX-x.x.x installation ...
......
nstallation finished successfully.
httempting to perform Firmware update...
huerying Mellanox devices firmware ...
```

For unattended installation, use the --force installation option while running the MLNX_OFED installation script:

/mnt/mlnxofedinstall --force

MLNX_OFED for Ubuntu should be installed with the following flags in chroot environment:

./mlnxofedinstall --without-dkms --add-kernel-support --kernel <kernel version in chroot> --without-fw-update --force

For example:

./mlnxofedinstall --without-dkms --add-kernel-support --kernel 3.13.0-85-generic --without-fw-update --force

Note that the path to kernel sources (--kernel-sources) should be added if the sources are not in their default location.

In case your machine has the latest firmware, no firmware update will occur and the installation script will print at the end of installation a message similar to the following:

Device #1:

Device Type: ConnectX-X
Part Number: MCXXXX-XXX

PSID: MT_<version>

PCI Device Name: 0b:00.0
Base MAC: 0000e41d2d5cf810
Versions: Current Available

FW XX.XX.XXXX Status: Up to date

In case your machine has an unsupported network adapter device, no firmware update will occur and one of the error messages below will be printed. Please contact your hardware vendor for help with firmware updates. Error message #1:

Device #1:

Device Type: ConnectX-X
Part Number: MCXXXX-XXX

PSID: MT_<version>

PCI Device Name: 0b:00.0 Base MAC: 0000e41d2d5cf810 Versions: Current Available FW XX.XX.XXXX

Status: No matching image found

Error message #2:

The firmware for this device is not distributed inside NVIDIA driver:

0000:01:00.0 (PSID: IBM2150110033)

To obtain firmware for this device, please contact your HW vendor.

d. **Case A**: If the installation script has performed a firmware update on your network adapter, you need to either restart the driver or reboot your system before the firmware update can take effect. Refer to the table below to find the appropriate action for your specific card.

Action \ Adapter	Driver Restart	Standard Reboot (Soft Reset)	Cold Reboot (Hard Reset)
Standard ConnectX- 4/ConnectX-4 Lx or higher	-	+	-
Adapters with Multi- Host Support	-	-	+
Socket Direct Cards	-	-	+

Case B: If the installations script has not performed a firmware upgrade on your network adapter, restart the driver by running: "/etc/init.d/openibd restart".

- e. (InfiniBand only) Run the hca_self_test.ofed utility to verify whether or not the InfiniBand link is up. The utility also checks for and displays additional information such as:
- o HCA firmware version
- o Kernel architecture
- Driver version
- Number of active HCA ports along with their states

Node GUID
 For more details on hca_self_test.ofed, see the file docs/readme_and_user_manual/hca_self_test.readme.

After installation completion, information about the OFED installation, such as prefix, kernel version, and installation parameters can be retrieved by running the command /etc/infiniband/info. Most of the OFED components can be configured or reconfigured after the installation, by modifying the relevant configuration files. See the relevant chapters in this manual for details.

The list of the modules that will be loaded automatically upon boot can be found in the /etc/infiniband/openib.conf file.

Installing OFED will replace the RDMA stack and remove existing 3rd party RDMA connectors.

Installation Results

Software	 Most of MLNX_OFED packages are installed under the "/usr" directory except for the following packages which are installed under the "/opt" directory: 	
	o fca and ibutils	
	iproute2 (rdma tool) - installed under /opt/Mellanox/iproute2/sbin/rdma	
	The kernel modules are installed under	
	 /lib/modules/`uname -r`/updates on SLES and Fedora Distributions 	
	 /lib/modules/`uname -r`/extra/mlnx-ofa_kernel on RHEL and other RedHat like Distributions 	
	 /lib/modules/`uname -r`/updates/dkms/ on Ubuntu 	

Firmware

- The firmware of existing network adapter devices will be updated if the following two conditions are fulfilled:
 - The installation script is run in default mode; that is, without the option '--without- fw-update'
 - The firmware version of the adapter device is older than the firmware version included with the OFED ISO image

Note: If an adapter's Flash was originally programmed with an Expansion ROM image, the automatic firmware update will also burn an Expansion ROM image.

 In case your machine has an unsupported network adapter device, no firmware update will occur and the error message below will be printed.

"The firmware for this device is not distributed inside NVIDIA driver: 0000:01:00.0 (PSID: IBM2150110033) To obtain firmware for this device, please contact your HW vendor."

Installation Logging

While installing MLNX_OFED, the install log for each selected package will be saved in a separate log file.

The path to the directory containing the log files will be displayed after running the installation script in the following format:

Example:

```
ogs dir: /tmp/MLNX_OFED_LINUX-4.4-1.0.0.0.IBMM2150110033.logs
```

Driver Load Upon System Boot

Upon system boot, the NVIDIA drivers will be loaded automatically.



To prevent the automatic load of the NVIDIA drivers upon system boot:

a. Add the following lines to the "/etc/modprobe.d/mlnx.conf" file.

```
lacklist mlx5 core
lacklist mlx5 ib
```

- b. Set "ONBOOT=no" in the "/etc/infiniband/openib.conf" file.
- c. If the modules exist in the initramfs file, they can automatically be loaded by the kernel. To prevent this behavior, update the initramfs using the operating systems' standard tools. Note: The process of updating the initramfs will add the blacklists from step 1, and will prevent the kernel from loading the modules automatically.

mInxofedinstall Return Codes

The table below lists the mlnxofedinstall script return codes and their meanings.

Return Code	Meaning
0	The Installation ended successfully
1	The installation failed
2	No firmware was found for the adapter device
22	Invalid parameter
28	Not enough free space
171	Not applicable to this system configuration. This can occur when the required hardware is not present on the system

Return Code	Meaning
172	Prerequisites are not met. For example, missing the required software installed or the hardware is not configured correctly
173	Failed to start the mst driver

Software	 Most of MLNX_OFED packages are installed under the "/usr" directory except for the following packages which are installed under the "/opt" directory:
	o fca and ibutils
	 iproute2 (rdma tool) - installed under /opt/Mellanox/iproute2/sbin/rdma
	The kernel modules are installed under
	 /lib/modules/`uname -r`/updates on SLES and Fedora Distributions
	 /lib/modules/`uname -r`/extra/mlnx-ofa_kernel on RHEL and other RedHat like Distributions
	/lib/modules/`uname -r`/updates/dkms/ on Ubuntu
o Firmware	 The firmware of existing network adapter devices will be updated if the following two conditions are fulfilled:
	 The installation script is run in default mode; that is, without the option 'without- fw-update'

The firmware version of the adapter device is older than the firmware version included with the OFED ISO image

Note: If an adapter's Flash was originally programmed with an Expansion ROM image, the automatic firmware update will also burn an Expansion ROM image.

 In case your machine has an unsupported network adapter device, no firmware update will occur and the error message below will be printed.

"The firmware for this device is not distributed inside NVIDIA driver: 0000:01:00.0 (PSID: IBM2150110033) To obtain firmware for this device, please contact your HW vendor."

Installation Logging

While installing MLNX_OFED, the install log for each selected package will be saved in a separate log file.

The path to the directory containing the log files will be displayed after running the installation script in the following format:

Example:

```
ogs dir: /tmp/MLNX_OFED_LINUX-4.4-1.0.0.0.IBMM2150110033.logs
```

Uninstalling MLNX_OFED

Use the script /usr/sbin/ofed_uninstall.sh to uninstall the MLNX_OFED package. The script is part of the ofed-scripts RPM.

Additional Installation Procedures

Installing MLNX_OFED Using YUM

This type of installation is applicable to RedHat/OL and Fedora operating systems.

Setting up MLNX_OFED YUM Repository

- a. Log into the installation machine as root.
- b. Mount the ISO image on your machine and copy its content to a shared location in your network.

```
mount -o ro,loop MLNX_OFED_LINUX-<ver>-<OS label>-<CPU arch>.iso /mnt
```

c. Download and install NVIDIA's GPG-KEY:

The key can be downloaded via the following link:

http://www.mellanox.com/downloads/ofed/RPM-GPG-KEY-Mellanox

d. Install the key.

```
sudo rpm --import RPM-GPG-KEY-Mellanox
varning: rpmts_HdrFromFdno: Header V3 DSA/SHA1 Signature, key ID 6224c050:
    NOKEY
tetrieving key from file:///repos/MLNX_OFED/<MLNX_OFED file>/RPM-GPG-KEY-Mellanox
importing GPG key 0x6224c050:
Userid: "Mellanox Technologies (Mellanox Technologies - Signing Key v2)
    <support@mellanox.com>"
From : /repos/MLNX_OFED/<MLNX_OFED file>/RPM-GPG-KEY-Mellanox
s this ok [y/N]:
```

e. Check that the key was successfully imported.

f. Create a yum repository configuration file called "/etc/yum.repos.d/mlnx_ofed.repo" with the following content:

```
mlnx_ofed]
name=MLNX_OFED Repository
naseurl=file:///<path to extracted MLNX_OFED package>/RPMS
nabled=1
npgkey=file:///<path to the downloaded key RPM-GPG-KEY-Mellanox>
npgcheck=1
```

g. Check that the repository was successfully added.

```
yum repolist
oaded plugins: product-id, security, subscription-manager
his system is not registered to Red Hat Subscription Management. You can use
subscription-manager to register.
epo id repo name
status
lnx_ofed MLNX_OFED Repository
108
pmforge RHEL 6Server - RPMforge.net - dag
4,597
epolist: 8,351
```

Setting up MLNX_OFED YUM Repository Using --add-kernel-support

- a. Log into the installation machine as root.
- b. Mount the ISO image on your machine and copy its content to a shared location in your network.

```
mount -o ro,loop MLNX_OFED_LINUX-<ver>-<OS label>-<CPU arch>.iso /mnt
```

c. Build the packages with kernel support and create the tarball.

```
/mnt/mlnx_add_kernel_support.sh --make-tgz <optional --kmp> -k $(uname -r) - m /mnt/
lote: This program will create MLNX_OFED_LINUX TGZ for rhel7.6 under /tmp directory.

lote: This program will create MLNX_OFED_LINUX TGZ for rhel7.6 under /tmp directory.

lote: This program will create MLNX_OFED_LINUX TGZ for rhel7.6 under /tmp directory.

lote: This program will create MLNX_OFED_LINUX TGZ for rhel7.6 under /tmp directory.

lote: This program will create MLNX_OFED_LINUX TGZ for rhel7.6 under /tmp directory.

lote: This program will create MLNX_OFED_LINUX TGZ for rhel7.6 under /tmp /tmp directory.

lote: This program will create MLNX_OFED_LINUX-5.2-0.5.5.0-rhel7.6-x86_64-ext.tgz
```

d. Open the tarball.

```
cd /tmp/
tar -xvf /tmp/MLNX_OFED_LINUX-5.2-0.5.5.0-rhel7.6-x86_64-ext.tgz
```

e. Create a YUM repository configuration file called "/etc/yum.repos.d/mlnx_ofed.repo" with the following content:

```
mlnx_ofed]
wame=MLNX_OFED Repository
waseurl=file:///<path to extracted MLNX_OFED package>/RPMS
what is nabled=1
pgcheck=0
```

f. Check that the repository was successfully added.

```
yum repolist
oaded plugins: product-id, security, subscription-manager
his system is not registered to Red Hat Subscription Management. You can use subscription-manager to register.
epo id repo name status
lnx_ofed MLNX_OFED Repository 108
pmforge RHEL 6Server - RPMforge.net - dag 4,597
epolist: 8,351
```

Installing MLNX_OFED Using the YUM Tool

After setting up the YUM repository for MLNX_OFED package, perform the following:

a. View the available package groups by invoking:

```
yum search mlnx-ofed-
lnx-ofed-all.noarch : MLNX OFED all installer package (with KMP support)
lnx-ofed-all-user-only.noarch : MLNX OFED all-user-only installer package
   (User Space packages only)
lnx-ofed-basic.noarch : MLNX OFED basic installer package (with KMP support)
lnx-ofed-basic-user-only.noarch : MLNX OFED basic-user-only installer package
   (User Space packages only)
lnx-ofed-bluefield.noarch : MLNX OFED bluefield installer package (with KMP
lnx-ofed-bluefield-user-only.noarch : MLNX OFED bluefield-user-only installer
  package (User Space packages only)
lnx-ofed-dpdk.noarch : MLNX OFED dpdk installer package (with KMP support)
lnx-ofed-dpdk-upstream-libs.noarch : MLNX OFED dpdk-upstream-libs installer
   package (with KMP support)
lnx-ofed-dpdk-upstream-libs-user-only.noarch : MLNX OFED dpdk-upstream-libs-
  user-only installer package (User Space packages only)
lnx-ofed-dpdk-user-only.noarch : MLNX OFED dpdk-user-only installer package
  (User Space packages only)
lnx-ofed-eth-only-user-only.noarch : MLNX OFED eth-only-user-only installer
  package (User Space packages only)
lnx-ofed-guest.noarch : MLNX OFED guest installer package (with KMP support)
lnx-ofed-quest-user-only.noarch : MLNX OFED quest-user-only installer package
  (User Space packages only)
hlnx-ofed-hpc.noarch : MLNX OFED hpc installer package (with KMP support)
lnx-ofed-hpc-user-only.noarch : MLNX OFED hpc-user-only installer package
   (User Space packages only)
lnx-ofed-hypervisor.noarch : MLNX OFED hypervisor installer package (with
  KMP support)
lnx-ofed-hypervisor-user-only.noarch : MLNX OFED hypervisor-user-only
  installer package (User Space packages only)
lnx-ofed-kernel-only.noarch : MLNX OFED kernel-only installer package (with
  KMP support)
hlnx-ofed-vma.noarch : MLNX OFED vma installer package (with KMP support)
lnx-ofed-vma-eth.noarch : MLNX OFED vma-eth installer package (with KMP
  support)
lnx-ofed-vma-eth-user-only.noarch : MLNX OFED vma-eth-user-only installer
  package (User Space packages only)
lnx-ofed-vma-user-only.noarch : MLNX OFED vma-user-only installer package
   (User Space packages only)
lnx-ofed-vma-vpi.noarch : MLNX OFED vma-vpi installer package (with KMP
lnx-ofed-vma-vpi-user-only.noarch : MLNX OFED vma-vpi-user-only installer
  package (User Space packages only
```

where:

mlnx-ofed-all	Installs all available packages in MLNX_OFED
mlnx-ofed-basic	Installs basic packages required for running NVIDIA cards
mlnx-ofed-guest	Installs packages required by guest OS
mlnx-ofed-hpc	Installs packages required for HPC
mlnx-ofed-hypervisor	Installs packages required by hypervisor OS
mlnx-ofed-vma	Installs packages required by VMA
mlnx-ofed-vma-eth	Installs packages required by VMA to work over Ethernet
mlnx-ofed-vma-vpi	Installs packages required by VMA to support VPI
bluefield	Installs packages required for BlueField
dpdk	Installs packages required for DPDK
dpdk-upstream-libs	Installs packages required for DPDK using RDMA-Core
kernel-only	Installs packages required for a non-default kernel

Note: MLNX_OFED provides kernel module RPM packages with KMP support for RHEL and SLES. For other operating systems, kernel module RPM packages are provided only for the operating system's default kernel. In this case, the group RPM packages have the supported kernel version in their package's name.

Example:

```
lnx-ofed-all-3.17.4-301.fc21.x86 64.noarch : MLNX OFED all installer package
   for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-basic-3.17.4-301.fc21.x86 64.noarch : MLNX OFED basic installer
  package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-guest-3.17.4-301.fc21.x86 64.noarch : MLNX OFED guest installer
  package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-hpc-3.17.4-301.fc21.x86 64.noarch : MLNX OFED hpc installer package
  for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-hypervisor-3.17.4-301.fc21.x86 64.noarch : MLNX OFED hypervisor
   installer package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-vma-3.17.4-301.fc21.x86 64.noarch : MLNX OFED vma installer package
   for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-vma-eth-3.17.4-301.fc21.x86 64.noarch : MLNX OFED vma-eth installer
   package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-vma-vpi-3.17.4-301.fc21.x86 64.noarch : MLNX OFED vma-vpi installer
   package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-hypervisor-3.17.4-301.fc21.x86 64.noarch : MLNX OFED hypervisor
  installer package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-vma-3.17.4-301.fc21.x86 64.noarch : MLNX OFED vma installer package
   for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
lnx-ofed-vma-eth-3.17.4-301.fc21.x86 64.noarch : MLNX OFED vma-eth installer
  package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
hlnx-ofed-vma-vpi-3.17.4-301.fc21.x86 64.noarch : MLNX OFED vma-vpi installer
   package for kernel 3.17.4-301.fc21.x86 64 (without KMP support)
```

When using an operating system different than RHEL or SLES, or you have installed a kernel that is not supported by default in MLNX_OFED, you can use the mlnx_add_kernel_support.sh script to build MLNX_OFED for your kernel.

The script will automatically build the matching group RPM packages for your kernel so that you can still install MLNX_OFED via yum. Please note that the resulting MLNX_OFED repository will contain unsigned RPMs, therefore, you should set 'gpgcheck=0' in the

b. Install the desired group.

repository configuration file.

Installing MLNX_OFED using the "YUM" tool does not automatically update the firmware.

To update the firmware to the version included in MLNX_OFED package, run: # yum install mlnx-fw-updater

Installing MLNX_OFED Using apt-get

This type of installation is applicable to Debian and Ubuntu operating systems.

Setting up MLNX_OFED apt-get Repository

a. Log into the installation machine as root.

- Extract the MLNX_OFED package on a shared location in your network.
 It can be downloaded from https://www.nvidia.com/en-us/networking/ → Products → Software → InfiniBand Drivers.
- c. Create an apt-get repository configuration file called "/etc/apt/sources.list.d/mlnx_ofed.list" with the following content:

```
leb file:/<path to extracted MLNX_OFED package>/DEBS ./
```

d. Download and install NVIDIA's Technologies GPG-KEY.

```
wget -qO - http://www.mellanox.com/downloads/ofed/RPM-GPG-KEY-Mellanox | sudo apt-key add -
```

e. Verify that the key was successfully imported.

```
apt-key list
pub 1024D/A9E4B643 2013-08-11
iid Mellanox Technologies <support@mellanox.com>
sub 1024g/09FCC269 2013-08-11
```

f. Update the apt-get cache.

```
sudo apt-get update
```

Setting up MLNX_OFED apt-get Repository Using --add-kernel-support

- a. Log into the installation machine as root.
- b. Mount the ISO image on your machine and copy its content to a shared location in your network.

```
mount -o ro,loop MLNX_OFED_LINUX-<ver>-<OS label>-<CPU arch>.iso /mnt
```

c. Build the packages with kernel support and create the tarball.

```
/mnt/mlnx_add_kernel_support.sh --make-tgz <optional --kmp> -k $(uname -r) -
    m /mnt/
Jote: This program will create MLNX_OFED_LINUX TGZ for rhel7.6 under /tmp
    directory.
Do you want to continue?[y/N]:y
See log file /tmp/mlnx_iso.4120_logs/mlnx_ofed_iso.4120.log

Thecking if all needed packages are installed...
Suilding MLNX_OFED_LINUX RPMS . Please wait...
Freating metadata-rpms for 3.10.0-957.21.3.el7.x86_64 ...

JARNING: If you are going to configure this package as a repository, then
    please note

JARNING: that it contains unsigned rpms, therefore, you need to disable the
    gpgcheck
JARNING: by setting 'gpgcheck=0' in the repository conf file.
Freated /tmp/MLNX_OFED_LINUX-5.2-0.5.5.0-rhel7.6-x86_64-ext.tgz
```

d. Open the tarball.

```
cd /tmp/
tar -xvf /tmp/MLNX_OFED_LINUX-5.2-0.5.5.0-rhel7.6-x86_64-ext.tgz
```

e. Create an apt-get repository configuration file called "/etc/apt/sources.list.d/mlnx_ofed.list" with the following content:

```
leb [trusted=yes] file:/<path to extracted MLNX_OFED package>/DEBS ./
```

f. Update the apt-get cache.

```
sudo apt-get update
```

Installing MLNX_OFED Using the apt-get Tool

After setting up the apt-get repository for MLNX_OFED package, perform the following:

a. View the available package groups by invoking:

```
apt-cache search mlnx-ofed-
pt-cache search mlnx-ofed ......
nem-dkms - DKMS support for mlnx-ofed kernel modules
hlnx-ofed-kernel-dkms - DKMS support for mlnx-ofed kernel modules
lnx-ofed-kernel-utils - Userspace tools to restart and tune mlnx-ofed kernel
  modules
lnx-ofed-vma-vpi - MLNX OFED vma-vpi installer package (with DKMS support)
lnx-ofed-kernel-only - MLNX OFED kernel-only installer package (with DKMS
lnx-ofed-bluefield - MLNX OFED bluefield installer package (with DKMS
  support)
lnx-ofed-hpc-user-only - MLNX OFED hpc-user-only installer package (User
   Space packages only)
lnx-ofed-dpdk-user-only - MLNX OFED dpdk-user-only installer package (User
  Space packages only)
lnx-ofed-all-exact - MLNX OFED all installer package (with DKMS support)
   (exact)
hlnx-ofed-all - MLNX OFED all installer package (with DKMS support)
lnx-ofed-vma-vpi-user-only - MLNX OFED vma-vpi-user-only installer package
   (User Space packages only)
lnx-ofed-eth-only-user-only - MLNX OFED eth-only-user-only installer package
   (User Space packages only)
lnx-ofed-vma-user-only - MLNX OFED vma-user-only installer package (User
   Space packages only)
hlnx-ofed-hpc - MLNX OFED hpc installer package (with DKMS support)
lnx-ofed-bluefield-user-only - MLNX OFED bluefield-user-only installer
  package (User Space packages only)
lnx-ofed-dpdk - MLNX OFED dpdk installer package (with DKMS support)
lnx-ofed-vma-eth-user-only - MLNX OFED vma-eth-user-only installer package
   (User Space packages only)
lnx-ofed-all-user-only - MLNX OFED all-user-only installer package (User
   Space packages only)
lnx-ofed-vma-eth - MLNX OFED vma-eth installer package (with DKMS support)
lnx-ofed-vma - MLNX OFED vma installer package (with DKMS support)
lnx-ofed-dpdk-upstream-libs-user-only - MLNX OFED dpdk-upstream-libs-user-
  only installer package (User Space packages only)
lnx-ofed-basic-user-only - MLNX OFED basic-user-only installer package (User
   Space packages only)
lnx-ofed-basic-exact - MLNX OFED basic installer package (with DKMS support)
lnx-ofed-basic - MLNX OFED basic installer package (with DKMS support)
lnx-ofed-dpdk-upstream-libs - MLNX OFED dpdk-upstream-libs installer package
   (with DKMS support)
```

where:

mlnx-ofed-all	MLNX_OFED all installer package
mlnx-ofed-basic	MLNX_OFED basic installer package
mlnx-ofed-vma	MLNX_OFED vma installer package
mlnx-ofed-hpc	MLNX_OFED HPC installer package
mlnx-ofed-vma-eth	MLNX_OFED vma-eth installer package
mlnx-ofed-vma-vpi	MLNX_OFED vma-vpi installer package
knem-dkms	MLNX_OFED DKMS support for mlnx-ofed kernel modules
kernel-dkms	MLNX_OFED kernel-dkms installer package
kernel-only	MLNX_OFED kernel-only installer package
bluefield	MLNX_OFED bluefield installer package
mlnx-ofed-all-exact	MLNX_OFED mlnx-ofed-all-exact installer package
dpdk	MLNX_OFED dpdk installer package
mlnx-ofed-basic-exact	MLNX_OFED mlnx-ofed-basic-exact installer package
dpdk-upstream-libs	MLNX_OFED dpdk-upstream-libs installer package

b. Install the desired group.

```
pt-get install '<group name>'
```

Example:

pt-get install mlnx-ofed-all

Installing MLNX_OFED using the "apt-get" tool does not automatically update the firmware.

To update the firmware to the version included in MLNX_OFED package, run: # apt-get install mlnx-fw-updater

Performance Tuning

Depending on the application of the user's system, it may be necessary to modify the default configuration of network adapters based on the ConnectX® adapters. In case that tuning is required, please refer to the <u>Performance Tuning Guide for NVIDIA Network Adapters</u>.

5.3 VMware Driver Installation

This section describes VMware Driver Installation.

5.3.1 Hardware and Software Requirements

Requirement	Description
Platforms	A server platform with an adapter card based on NVIDIA devices:
	ConnectX®-7 (InfiniBand/Ethernet) (firmware: fw-ConnectX7)
Operating System	ESXi 8.x
Installer Privileges	The installation requires administrator privileges on the target machine.

5.3.2 Installing NATIVE ESXi Driver for VMware vSphere

Please uninstall all previous driver packages prior to installing the new version.

To install the driver:

- 1. Log into the ESXi server with root permissions.
- 2. Install the driver.

```
> esxcli software vib install -d <path>/<bundle_file>
```

Example:

```
> esxcli software vib install -d /tmp/MLNX-NATIVE-ESX-ConnectX-4-5_4.16.8.8-10EM-650.0.0.4240417.zipesxcli
```

- 3. Reboot the machine.
- 4. Verify the driver was installed successfully.

After the installation process, all kernel modules are loaded automatically upon boot.

5.3.3 Removing Earlier NVIDIA Drivers

Please unload the previously installed drivers before removing them.

To remove all the drivers:

- 1. Log into the ESXi server with root permissions.
- 2. List all the existing NATIVE ESXi driver modules. (See Step 4 in Installing NATIVE ESXi Driver for VMware vSphere.)
- 3. Remove each module:

```
> esxcli software vib remove -n nmlx5-rdma
> esxcli software vib remove -n nmlx5-core
```

To remove the modules, you must run the command in the same order as shown in the example above.

4. Reboot the server.

5.3.4 Firmware Programming

- 1. Download the VMware bootable binary images v4.6.0 from the Firmware Tools (MFT) site.
 - a. ESXi 6.5 File: mft-4.6.0.48-10EM-650.0.0.4598673.x86_64.vib
 - b. MD5SUM: 0804cffe30913a7b4017445a0f0adbe1
- 2. Install the image according to the steps described in the MFT User Manual.

The following procedure requires custom boot image downloading, mounting and booting from a USB device.

6 Updating Adapter Firmware

Each adapter card is shipped with the latest version of qualified firmware at the manufacturing time. However, NVIDIA issues firmware updates occasionally that provide new features and bug fixes. To check that your card is programmed with the latest available firmware version, download the mlxup firmware update and query utility. The utility can query for available Mellanox adapters and indicate which adapters require a firmware update. If the user confirms, mlxup upgrades the firmware using embedded images. The latest mlxup executable and documentation are available in mlxup - Update and Query Utility.

Firmware Update Example

```
[server1]# ./mlxup
Querying Mellanox devices firmware ...
Device Type: ConnectX-7
Part Number:
               MCX753436MC-HEAB
Description: NVIDIA ConnectX-7 OCP3.0 SFF Adapter Card, 200GbE (default mode) / NDR200 IB, Dual-port OSFP112, Multi-Host and
Socket Direct capable, PCIe 5.0 x16, Crypto Enabled, Secure Boot Enabled, Thumbscrew (Pull Tab) Bracket
PSID:
               MT 2190110032
PCI Device Name: 0000:06:00.0
Base GUID: e41d2d0300fd8b8a
Versions: Current Available
              FW 16.00.0000 16.00.0000
Status: Up to date
Device Type: ConnectX-7 Part Number:
                                         MCX753436MC-HEAB
Description: NVIDIA ConnectX-7 OCP3.0 SFF Adapter Card, 200GbE (default mode) / NDR200 IB, Dual-port QSFP112, Multi-Host and
Socket Direct capable, PCIe 5.0 x16, Crypto Enabled, Secure Boot Enabled, Thumbscrew (Pull Tab) Bracket
PCI Device Name: 0000:07:00.0
Base MAC:
             0000e41d2da206d4
Versions:
               Current Available
               FW 16.20.1000 16.24.1000
Status:
               Update required
Perform FW update? [y/N]: y
Device #1: Up to date
Device #2: Updating FW ... Done
Restart needed for updates to take effect.
Log File: /var/log/mlxup/mlxup-yyyymmdd.log
```

7 Setting Link Type of High-Speed Ports

The default networking port configuration of ConnectX-7 InfiniBand/Ethernet adapter cards is listed in the below table.

OPN	Data Transmission Rate	Default Mode
MCX753436MC-HEAB	NDR200/HDR and 200GbE	200GbE
MCX753436MS-HEAB	NDR200/HDR and 200GbE	200GbE
MCX75343AMC-NEAC	NDR and 400GbE	400GbE
MCX75343AMS-NEAC	NDR and 400GbE	400GbE

To change the networking high-speed ports mode, you can either use the mlnxconfig or UEFI tools.

UEFI can be used to configure the adapter card device before the operating system is up while mlnxconfig can be used once the operating system is up. According to your preference, use one of the below tools:

7.1 mlnxconfig

The mixconfig tool allows the user to change some of the device configurations without reburning the firmware. The configuration is also kept after reset. By default, mixconfig shows the configurations that will be loaded in the next boot. For more information and instructions, refer to minxconfig.

7.2 UEFI

PreBoot drivers first initialize the adapter device, senses the port protocol – Ethernet or InfiniBand, and brings up the port. Then it connects to a DHCP server to obtain its assigned IP address and network parameters, and also to obtain the source location of the kernel/OS to boot from. The DHCP server instructs the PreBoot drivers to access the kernel/OS through a TFTP server, an iSCSI target, or some other service. For more information and instructions, refer to <u>UEFI</u>.

8 Troubleshooting

8.1 General Troubleshooting

Server unable to find the adapter	 Ensure that the adapter is placed correctly Make sure the adapter slot and the adapter are compatible Install the adapter in a different PCI Express slot Use the drivers that came with the adapter or download the latest Make sure your motherboard has the latest BIOS Try to reboot the server
The adapter no longer works	 Reseat the adapter in its slot or a different slot, if necessary Try using another cable Reinstall the drivers for the network driver files may be damaged or deleted Reboot the server
Adapters stopped working after installing another adapter	 Try removing and re-installing all adapters Check that cables are connected properly Make sure your motherboard has the latest BIOS
Link indicator light is off	 Try another port on the switch Make sure the cable is securely attached Check you are using the proper cables that do not exceed the recommended lengths Verify that your switch and adapter port are compatible

Link light is on, but with no communication established	 Check that the latest driver is loaded Check that both the adapter and its link are set to the same speed and duplex settings
Event message received of insufficient power	 When [adapter's current power consumption] > [PCle slot advertised power limit] – a warning message appears in the server's system even logs (Eg. dmesg: "Detected insufficient power on the PCle slow")
	 It's recommended to use a PCIe slot that can supply enough power.
	• If a message of the following format appears — "mlx5_core 0003:01:00.0: port_module:254:(pid 0): Port module event[error]: module 0, Cable error, One or more network ports have been powered down due to insufficient/unadvertised power on the PCIe slot" please upgrade your Adapter's firmware.
	 If the message remains – please consider switching from Active Optical Cable (AOC) or transceiver to Direct Attached Copper (DAC) connectivity.

8.2 Linux Troubleshooting

Environment Information	cat /etc/issue uname -a cat /proc/cupinfo grep 'model name' uniq ofed_info -s ifconfig -a ip link show
Card Detection	ethtool <interface> ethtool -i <interface_of_mellanox_port_num> ibdev2netdev</interface_of_mellanox_port_num></interface>
Card Detection	Ispci grep -i Mellanox

Mellanox Firmware Tool (MFT)	Download and install MFT: MFT Documentation Refer to the User Manual for installation instructions. Once installed, run: mst start mst status flint -d <mst_device> q</mst_device>
Ports Information	ibstat ibv_devinfo
Firmware Version Upgrade	To download the latest firmware version, refer to the <u>NVIDIA Update and Query Utility</u> .
Collect Log File	cat /var/log/messages dmesg >> system.log journalctl (Applicable on new operating systems) cat /var/log/syslog

8.3 Windows Troubleshooting

Environment Information	From the Windows desktop choose the Start menu and run: msinfo32 To export system information to a text file, choose the Export option from the File menu. Assign a file name and save.
Mellanox Firmware Tool (MFT)	Download and install MFT: MFT Documentation Refer to the User Manual for installation instructions. Once installed, open a CMD window and run: WinMFT mst start mst status flint -d <mst_device> q</mst_device>

Ports Information	vstat	
Firmware Version Upgrade	Download the latest firmware version using the PSID/board ID from here . flint -d <mst_device> -i <firmware_bin_file> b</firmware_bin_file></mst_device>	
Collect Log File	 Event log viewer MST device logs: mst start mst status flint –d <mst_device> dc > dump_configuration.log</mst_device> mstdump <mst_device> dc > mstdump.log</mst_device> 	

9 Specifications

Please make sure to install the ConnectX-7 card in a PCIe slot that is capable of supplying the required power and airflow, as stated in the below table.

In Standby mode only port0 is available.

9.1 MCX75343AMC-NEAC / MCX75343AMS-NEAC Specifications

Physical	Adapter Card Size: Tall Small Form Factor (TSFF) OCP 3.0, 4.52 in. x 2.99 in (115.00mm x 76.00mm) Adapter Card Height: 0.7 in. (17.8mm)				
	Bracket Type: Thumbscrew (Pull Tab) Bracket				
Interfaces	See Supported Interfaces				
	PCI Express Gen 5.0/4.0	SERDES @ 16/32GT/s, x1	6 lanes (Gen 3.0 compatible)		
	Networking Port: Single-port OSFP Infir	iBand and Ethernet			
Protocol Support	Data Rate	InfiniBand	NDR/NDR200/HDR/HDR100/EDR/FDR/S	SDR	
		Ethernet	400/200/100/50/25/10 Gb/s		
	_		e), NDR200 (2 lanes x 100Gb/s per lane), HDR (4 5Gb/s per lane), 1X/2X/4X SDR (2.5Gb/s per lan		
	Ethernet: 400GAUI-4 C2M, 400GBASE-CR4, 200GAUI-4 C2M, 200GBASE-CR4, 100GAUI-2 C2M, 100GBASE-CR4, 100GBASE-CR2, 50GAUI-2 C2M, 50GAUI-1 C2M, 50GBASE-CR, 40GBASE-CR4, 25GBASE-R, 10GBASE-R, 10GBASE-CX4, 1000BASE-CX, CAUI-4 C2M, 25GAUI C2M, XLAUI C2M, XLPPI, SFI				
Capabilities	MCX75343AMC-NEAC	Secure Boot Enabled, Cry	ypto Enabled, Support for Multi-Host and Socket Direct		
	MCX75343AMS-NEAC	Secure Boot Enabled, Cry	Secure Boot Enabled, Crypto Disabled, Support for Multi-Host and Socket Direct		
Electrical and Thermal	Voltage: 12V_EDGE, +3.3V_EDGE				
Specifications	Typical power with passive cables in PCIe Gen 5.0 x16		MCX75343AMC-NEAC	25.9W	
			MCX75343AMS-NEAC	24.4W	
	Maximum power available through OS	FP port: 17W per port (Not ther	mally supported)		
	The complete electrical and thermal specifications are provided in "NVIDIA ConnectX-7 Electrical and Thermal Specifications" document, which is available at NVOnline following login.				
Environmental	Temperature	Operational	0°C to 55°C		
		Non-operational	-40°C to 70°C ^b		
	Humidity	Operational	10% to 85% relative humidity		
		Non-operational	10% to 90% relative humidity		
	Altitude (Operational)	3050m			
Regulatory	Safety: CB / cTUVus / CE				
	EMC: CE / FCC / VCCI / ICES / RCM / KC				
	RoHS: RoHS Compliant				

Notes: a. The ConnectX-7 adapters supplement the IBTA auto-negotiation specification to get better bit error rates and longer cable reaches. This supplemental feature only initiates when connected to another NVIDIA InfiniBand product.

b. The non-operational storage temperature specifications apply to the product without its package.

9.2 MCX753436MS-HEAB / MCX753436MC-HEAB Specifications

Physical	Adapter Card Size: Small Form Factor (SFF), 4.52 in. x 2.99 in (115.00mm x 76.00mm) Adapter Card Height: 0.59 in. (15.1mm)			
	Bracket Type: Thumbscrew (Pull Ta	b) Bracket		
Interfaces	See Supported Interfaces			
	PCI Express Gen 5.0/4.0	SERDES @ 16/32GT/s,	x16 lanes (4.0 and 3.0 compatible)	
	Networking Port: Dual-Port QSFP12	12 InfiniBand and Ethern	et	
Protocol Support	Data Rate	InfiniBand	NDR200/HDR/HDR100/EDR/FDR	/SDR
		Ethernet	200/100/50/25/10 Gb/s	
	Auto-Negotiation: NDR200 (2 lanes x 100Gb/s per lane), HDR (4 lanes x 50Gb/s per lane), HDR100 (2 lane x 50Gb/s per lane), EDR (25Gb/s per lane), FDR (14.0625Gb/s per lane), 1X/2X/4X SDR (2.5Gb/s per lane) Ethernet: 200GAUI-4 C2M, 200GAUI-4 C2C, 200GBASE-CR4, 200GBASE-KR4, 200GBASE-CR4, 200GBASE-KR4, 100GBASE-			
	CR4, 100GBASE-KR4, 100GBASE-SR4, 100GBASE-CR2, 100GBASE-KR2, 100GBASE-SR2, 50GBASE-R2, 50GBASE-R4, 40GBASE-CR4, 40GBASE-KR4, 40GBASE-KR4, 40GBASE-KR4, 40GBASE-KR4, 40GBASE-KR4, 40GBASE-KR4, 40GBASE-R2, 25GBASE-R, 20GBASE-KR2, 10GBASE-LR, 10GBASE-CR, 10GBASE-			
Capabilities	MCX753436MC-HEAB	Secure Boot Enabled, (Crypto Enabled, Support for Multi-Hos	t and Socket Direct
	MCX753436MS-HEAB	Secure Boot Enabled, (Crypto Disabled, Support for Multi-Ho	st and Socket Direct
Electrical and	Voltage: 12V_EDGE, +3.3V_EDGE			
Thermal	Maximum power available through QSFP112 port: 17W per port (Not thermally supported)			
Specifications	Typical power with passive cables in PCIe Gen 5.0 x16		MCX753436MC-HEAB	25.9W
			MCX753436MS-HEAB	24.5W
	The complete electrical and thermal specifications are provided in "NVIDIA ConnectX-7 Electrical and Thermal Specifications" document, which is available at NVOnline following login.			
Environmental	Temperature	Operational	0°C to 55°C	
		Non-operational	-40°C to 70°C ^b	

	Humidity	Operational	10% to 85% relative humidity	
		Non-operational	10% to 90% relative humidity	
	Altitude (Operational)	3050m		
Regulatory	Safety: CB / cTUVus / CE			
	EMC: CE / FCC / VCCI / ICES /	RCM / KC		
	RoHS: RoHS Compliant			

Notes: a. The ConnectX-7 adapters supplement the IBTA auto-negotiation specification to get better bit error rates and longer cable reaches. This supplemental feature only initiates when connected to another NVIDIA InfiniBand product.

b. The non-operational storage temperature specifications apply to the product without its package.

9.3 Mechanical Drawings and Dimensions

For the 3D Model of the card, please refer to the PID.

9.3.1 Board Drawings and Dimensions

All dimensions are in millimeters. The PCB mechanical tolerance is +/- 0.13mm.

ConnectX-7 Single-Port OSFP

ConnectX-7 Dual-Port QSFP112

9.3.2 Bracket Drawings and Dimensions

All dimensions are in millimeters. The bracket's mechanical tolerance is +/- 0.25mm.

Single-Port OSFP Thumbscrew Bracket

Dual-Port QSFP112 Thumbscrew Bracket

82.50

17.80

10 Monitoring

10.1 Thermal Sensors

The adapter card incorporates the ConnectX IC, which operates in the range of temperatures between 0°C and 105°C.

Three thermal threshold definitions impact the overall system operation state:

- **Warning** 105°C: On managed systems only: When the device crosses the 105°C threshold, a Warning Threshold message is issued by the management SW, indicating to system administration that the card has crossed the warning threshold. Note that this temperature threshold does not require nor lead to any action by hardware (such as adapter card shutdown).
- Critical 115°C: When the device crosses this temperature, the firmware automatically shuts down the device.
- **Emergency** 130°C: If the firmware fails to shut down the device upon crossing the critical threshold, the device automatically shuts down upon crossing the emergency (130°C) threshold.

The card's thermal sensors can be read through the system's SMBus. The user can read these thermal sensors and adapt the system airflow following the readouts and the needs of the above-mentioned IC thermal

10.2 Heatsink

A heatsink is attached to the ConnectX-7 IC to dissipate the heat. ConnectX-7 IC has a thermal shutdown safety mechanism that automatically shuts down the ConnectX-7 card in case of a high-temperature event, improper thermal coupling, or heatsink removal.

For the required airflow (LFM) per OPN, please refer to Specifications.

The below table lists the heatsink proprieties per OPN.

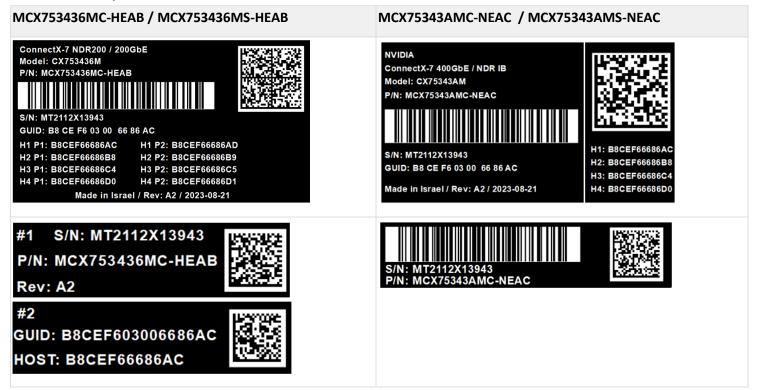
OPN	Heatsink
MCX753436AS-HEAB and MCX753436MS-HEAB	Attached by using two spring-loaded push pins.
MCX75343AMC-NEAC and MCX75343AMS-NEAC	Attached by using two spring-loaded push pins and three L-shaped mechanical holders, as can be seen in the above mechanical drawing. These L-shaped mechanical holders connect the PCB to the heatsink to create a stiffener.

11 Finding the MAC and Serial Number on the Adapter Card

Each adapter card has a different identifier printed on the label: serial number and the card MAC for the Ethernet protocol and the card GUID for the InfiniBand protocol. VPI cards have both a GUID and a MAC (derived from the GUID). IB only cards have GUID for the InfiniBand protocol.

The product revisions indicated on the labels in the following figures do not necessarily represent the latest revisions of the cards.

Board Label Example





12 Document Revision History

Date	Description of Changes
Dec. 2023	Updated regulatory label in Finding the MAC and Serial Number on the Adapter Card
Oct. 2023	 Updated <u>Finding the MAC and Serial Number on the Adapter Card</u> Updated the bracket dimensions in <u>Specifications</u>
Aug. 2023	Updated Finding the MAC and Serial Number on the Adapter Card
Jun. 2023	Updated Specifications with typical power numbers
May. 2023	Updated Specifications to include non-operational storage temperature specifications
Jan. 2023	Updated 400Gb/s supported Ethernet protocols in Specifications
Dec. 2022	First release of this consolidated user manual for all ConnectX-7 adapter cards for OCP 3.0

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.

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 other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2023 NVIDIA Corporation & affiliates. All Rights Reserved.