

NVIDIA ConnectX-5 Ethernet Adapter Cards User Manual

Table of contents

Introduction	9
Interfaces	15
Hardware Installation	17
Driver Installation	32
Linux Driver Installation	32
Windows Driver Installation	57
VMware Driver Installation	67
Updating Adapter Firmware	71
Troubleshooting	73
Specifications	76
Monitoring	89
Finding the MAC on the Adapter Card	90
Document Revision History	94

List of Figures

Figure 0. Image2021 8 11 8 44 35
Figure 1. Image2021 8 11 8 56 28
Figure 2. Image2021 8 11 8 44 35
Figure 3. Image2021 8 11 8 56 28
Figure 4. Image2021 8 11 8 44 35
Figure 5. Image2021 8 11 8 56 28
Figure 6. Image2021 8 11 8 44 35
Figure 7. Image2021 8 11 8 56 28
Figure 8. Image2021 8 11 8 44 3
Figure 9. Image2018 11 8 16 16 16
Figure 10. Image2018 11 15 11 34 36
Figure 11. Image2018 11 19 15 49 54
Figure 12. Image2018 11 8 16 25 32
Figure 13. Image2018 11 8 16 32 19
Figure 14. Image2018 11 8 16 32 40
Figure 15. Image2018 10 2 16 16 3
Figure 16. Image2018 11 14 15 55 33
Figure 17. Image2018 11 8 16 41 39
Figure 18. Image2018 11 8 16 42 15

Figure 19. Procedure Heading Icon
Figure 20. Procedure Heading Icon
Figure 21. Welcome Page
Figure 22. EULA
Figure 23. Destination Folder
Figure 24. Firmware Upgrade
Figure 25. Setup Type
Figure 26. Installer Features Screen
Figure 27. Ready 2 Install
Figure 28. FW Upgrade Icon
Figure 29. Installation Progress
Figure 30. Installation Completed
Figure 31. Image2018 11 11 16 7 44
Figure 32. Image2018 11 11 16 8 15
Figure 33. Image2018 11 11 16 9 2
Figure 34. Image2018 11 11 16 9 22
Figure 35. Image2023 12 18 10 54 55
Figure 36. Image2018 11 8 19 13 49
Figure 37. Image2018 11 8 19 14 6
Figure 38. Image2018 11 8 19 14 49
Figure 39. Image2018 11 8 19 15 15

Figure 40. lmage2018 11 8 19 15 28	
Figure 41. Image2018 11 8 19 15 43	
Figure 42. Image2018 11 8 19 15 56	
Figure 43. Image2020 11 10 10 20 48	
Figure 44. Image2018 11 8 19 16 8	



(i) Note

The ConnectX-5 product line has moved to end-of-life.

About This Manual

This User Manual describes NVIDIA® ConnectX®-5 and ConnectX®-5 Ex Ethernet Single and Dual SFP28 and QSFP28 port PCI Express x8/x16 adapter cards. It provides details as to the interfaces of the board, specifications, required software and firmware for operating the board, and relevant documentation.

EOL'ed (End of Life) Ordering Part Numbers

IC in Use	NVIDIA SKU	Legacy OPN	Marketing Description
900- 9X5AZ- 0053-ST		MCX51 2A- ADAT	ConnectX®-5 Ex EN network interface card, 25GbE dual -port SFP28, PCIe Gen 3.0/4.0 x8 , tall bracket
Connec tX®-5 Ex	9X5AD- 6A- ConnectX®-5 Ex EN network interface ca		ConnectX®-5 Ex EN network interface card, 100GbE dual -port QSFP28, PCle Gen 4.0 x16 , tall bracket
	900- 9X5AD- 0054-ST0	MCX51 6A- BDAT	ConnectX®-5 Ex EN network interface card, 40GbE dual -port QSFP28, PCle Gen 4.0 x16 , tall bracket
Connec tX®-5	900- 9X5AZ- 0053-0T3	MCX51 2A- ACAT	ConnectX®-5 EN network interface card, 25GbE dual -port SFP28, PCIe Gen 3.0 x8 , tall bracket
	900- 9X5AZ- 0053-ST4	MCX51 2A- ACUT	ConnectX®-5 EN network interface card, 10/25GbE dual -port SFP28, PCle Gen 3.0 x8 , UEFI Enabled (x86/Arm) , tall bracket
	900- 9X5AZ-	MCX51 2F-	ConnectX®-5 EN network interface card, 25GbE dual -port SFP28, PCIe Gen 3.0 x16 , tall bracket

IC in Use	NVIDIA SKU	Legacy OPN	Marketing Description
	0053-ST6	ACAT	
	900- 9X5AZ- 0053-ST0	MCX51 2F- ACHT	ConnectX®-5 EN network interface card, with host management, 25GbE Dual-port SFP28, PCle Gen 3.0 x16, UEFI Enabled, tall bracket
	900- 9X5AD- 0015-ST0	MCX51 5A- GCAT	ConnectX®-5 EN network interface card, 50GbE single -port QSFP28, PCIe Gen 3.0 x16 , tall bracket
	900- 9X5AD- 0055-ST0	MCX51 6A- GCAT	ConnectX®-5 EN network interface card, 50GbE dual -port QSFP28, PCIe Gen 3.0 x16 , tall bracket
	900- 9X5AD- 0016-ST1	MCX51 5A- CCAT	ConnectX®-5 EN network interface card, 100GbE single -port QSFP28, PCIe Gen 3.0 x16 , tall bracket
9X5AD- 5A- port QSFI		5A-	ConnectX®-5 EN network interface card, 100GbE single -port QSFP28, PCle Gen 3.0 x16 , UEFI Enabled (Arm, x86) , tall bracket
	900- 9X5AD- 0056-ST6	MCX51 6A- CCHT	ConnectX®-5 EN network interface card, with host management 100GbE dual-port QSFP28, PCIe Gen 3.0 x16, UEFI Enabled, tall bracket
	900- 9X5AD- 0056-ST1	MCX51 6A- CCAT	ConnectX®-5 EN network interface card, 100GbE dual -port QSFP28, PCIe Gen 3.0 x16 , tall bracket

Intended Audience

This manual is intended for the installer and user of these cards. The manual assumes basic familiarity with Ethernet network and architecture specifications.

Technical Support

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

URL: https://www.nvidia.com > SupportE-mail: enterprisesupport@nvidia.com

Customers who purchased NVIDIA 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.

Related Documentation

MLNX_O FED for Linux User Manual and Release Notes	User Manual describing OFED features, performance, band diagnostic, tools content, and configuration. See <u>MLNX_OFED for Linux Documentation</u> .
WinOF-2 for Window s 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 VMware® ESXi Drivers Documentation.
NVIDIA Firmwar e Utility (mlxup) User Manual and Release Notes	NVIDIA firmware update and query utility used to update the firmware. Refer to Firmware Utility (mlxup) Documentation.

NVIDIA Firmwar e Tools (MFT) User Manual	User Manual describing the set of MFT firmware management tools for a single node. See MFT User Manual.
IEEE Std 802.3 Specifica tion	IEEE Ethernet Specifications
PCI Express Specifica tions	Industry Standard PCI Express Base and Card Electromechanical Specifications. Refer to PCI-SIG Specifications.
LinkX Intercon nect Solution s	NVIDIA LinkX Ethernet 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 portfolios of 40Gb/s, 56Gb/s and 100Gb/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 Ethernet 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® Ethernet Cables and Transceivers.

Document Conventions

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

Revision History

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

Introduction

Product Overview

This is the User Guide for Ethernet adapter cards based on the ConnectX®-5 and ConnectX®-5 Ex integrated circuit device. These adapters connectivity provide the highest performing low latency and most flexible interconnect solution for PCI Express Gen 3.0/4.0 servers used in Enterprise Data Centers and High-Performance Computing environments.

The following provides the ordering part number, port speed, number of ports, and PCI Express speed. Each adapter comes with two bracket heights - short and tall.

ConnectX-5 Ex Ethernet Adapter Cards

Model	ConnectX-5 Ex Ethernet Adapter Cards				
Part Number	MCX512A-ADAT MCX516A-BDAT MCX516A-CDAT				
Ethernet Data Rate	1/10/25 Gb/s 1/10/25/40 Gb/s 10/25/40/50/100 Gb/s				
Network Connector Type	Dual-port SFP28 Dual-port QSFP28 Dual-port QSFP28				
PCI Express Connectors	PCIe Gen 3.0/4.0 x16 ^{(a), (b)} SERDES @ 8.0GT/s / 16.0GT/s				
Dimensions	2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile				
RoHS	RoHS Compliant				
IC Part Number	MT28808A0-FCCF-EV				

a. PCIe 4.0×16 bus can supply a maximum bandwidth of 256Gb/s (= 16×16 GT/s, including overhead), and can support 200Gb/s when both network ports of the card run at 100Gb/s.

ConnectX-5 Ethernet 25Gb/s Adapter Cards

b. The MCX516A-CDAT card has been tested and certified with PCIe 3.0 servers. PCIe 4.0 interface will be tested when servers with Gen 4.0 support become available.

Model	ConnectX-5 Ethernet Adapter Cards			
Part Number	MCX512A- ACAT	MCX512A- ACUT	MCX512F- ACAT	MCX512F- ACHT
Ethernet Data Rate	1/10/25 Gb/s			
Network Connector Type	Dual-port SFP28			
PCI Express Connectors	PCIe Gen 3.0 x8 SerDes @ 8.0GT/s		PCle Gen 3.0 x16 SerDes @ 8.0GT/s	
Dimensions	2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile			
Host Management	-	-	-	√
UEFI Enabled (x86/Arm)	-	√	-	√
RoHS	RoHS Compliant			
IC Part Number	MT27808A0-FCCF-EV			

ConnectX-5 Ethernet 50Gb/s Adapter Cards

Model	ConnectX-5 Ethernet Adapter Cards		
Part Number	MCX515A-GCAT MCX516A-GCAT		
Ethernet Data Rate	10/25/40/50 Gb/s		
Network Connector Type	Single-port QSFP28 Dual-port QSFP28		
PCI Express Connectors	PCIe Gen 3.0 x16 SerDes @ 8.0GT/s		
Dimensions	2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile		
Host Management	-	-	
UEFI Enabled (x86/Arm)	-	-	
RoHS	RoHS Compliant		
IC Part Number	MT27808A0-FCCF-EV		

ConnectX-5 Ethernet 100Gb/s Adapter Cards

Model	ConnectX-5 Ethernet Adapter Cards					
Part Number	MCX515A-CCAT	MCX515A-CCUT	MCX516A- CCAT	MCX516A- CCHT		
Ethernet Data Rate	10/25/40/50/100	Gb/s				
Network Connector Type	Single-port QSFP28					
PCI Express Connectors	PCIe Gen 3.0 x16 SerDes @ 8.0GT/s					
Dimensions	2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile					
Host Management		-	-	√		
UEFI Enabled (x86/Arm)	-	✓	-	√		
RoHS	RoHS Compliant					
IC Part Number	MT27808A0-FCCF-EV					

For more detailed information see **Specifications**.

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
PCI Express (PCIe)	Uses PCIe Gen 3.0 (8GT/s) and Gen 4.0 (16GT/s) through an x8 or x16 edge connector. Gen 1.1 and 2.0 compatible.

Feature	Description	
Up to 25 Gigabit Etherne t	NVIDIA adapters comply with the following IEEE 802.3 standards: • 100GbE/ 50GbE / 40GbE / 25GbE / 10GbE / 1GbE • IEEE 802.3bj, 802.3bm 100 Gigabit Ethernet • IEEE 802.3by, Ethernet Consortium25, 50 Gigabit Ethernet, supporting all FEC modes • IEEE 802.3ba 40 Gigabit Ethernet • IEEE 802.3by 25 Gigabit Ethernet • IEEE 802.3ae 10 Gigabit Ethernet • IEEE 802.3ap based auto-negotiation and KR startup • Proprietary Ethernet protocols (20/40GBASE-R2, 50GBASE-R4) • IEEE 802.3ad, 802.1AX Link Aggregation • IEEE 802.1Q, 802.1P VLAN tags and priority • IEEE 802.1Qau (QCN) • Congestion Notification • IEEE 802.1Qaz (ETS) • IEEE 802.1Qbb (PFC) • IEEE 802.1Qbg • IEEE 1588v2 • Jumbo frame support (9.6KB)	
Memory	 SPI Quad - includes 128Mbit SPI Quad Flash device (W25Q128FVSIG device by ST Microelectronics). Available in QSFP28 OPNs: FRU EEPROM - Stores the parameters and personality of the card. The EEPROM capacity is 128Kbit. FRU I2C address is (0x50) and is accessible through the PCIe SMBus. (Note: A ddress 0x58 is reserved.) 	
Overlay Network s	In order to better scale their networks, data center operators often create overlay networks that carry traffic from individual virtual machines over logical tunnels in encapsulated formats such as NVGRE and VXLAN. While this solves network scalability issues, it hides the TCP packet from the hardware offloading engines, placing higher loads on the host CPU. ConnectX-5 effectively addresses this by providing advanced NVGRE and VXLAN hardware offloading engines that encapsulate and de-capsulate the overlay protocol.	
RDMA and RDMA over Converg	ConnectX-5, utilizing IBTA RDMA (Remote Data Memory Access) and RoCE (RDMA over Converged Ethernet) technology, delivers low-latency and high-performance over Band and Ethernet networks. Leveraging data center bridging (DCB) capabilities as well as ConnectX-5 advanced congestion control	

Feature	Description
ed Etherne t (RoCE)	hardware mechanisms, RoCE provides efficient low-latency RDMA services over Layer 2 and Layer 3 networks.
NVIDIA PeerDir ect™	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-5 advanced acceleration technology enables higher cluster efficiency and scalability to tens of thousands of nodes.
CPU Offload	Adapter functionality enabling reduced CPU overhead allowing more available CPU for computation tasks. Open vSwitch (OVS) offload using ASAP ^{2(TM)} • Flexible match-action flow tables • Tunneling encapsulation/decapsulation
Quality of Service (QoS)	Support for port-based Quality of Service enabling various application requirements for latency and SLA.
Hardwa re- based I/O Virtualiz ation	ConnectX-5 provides dedicated adapter resources and guaranteed isolation and protection for virtual machines within the server.
Storage Accelera tion	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 • Erasure Coding • T10-DIF Signature Handover
SR-IOV	ConnectX-5 SR-IOV technology provides dedicated adapter resources and guaranteed isolation and protection for virtual machines (VM) within the server.
NC-SI	The adapter supports a Network Controller Sideband Interface (NC-SI), MCTP over SMBus and MCTP over PCIe - Baseboard Management Controller interface.

Feature	Description
High- Perform ance Accelera tions	 Tag Matching and Rendezvous Offloads Adaptive Routing on Reliable Transport Burst Buffer Offloads for Background Checkpointing
UEFI Enabled	Enabled in MCX516A-CCHT and MCX512F-ACHT. UEFI Enabled (x86/Arm) is enabled in MCX515A-CCUT and MCX512A-ACUT.

Operating Systems/Distributions

- RHEL/CentOS
- Windows
- FreeBSD
- VMware
- OpenFabrics Enterprise Distribution (OFED)
- OpenFabrics Windows Distribution (WinOF-2)

Connectivity

- Interoperable with 1/10/25/40/50/100 Gb/s Ethernet switches
- Passive copper cable with ESD protection
- Powered connectors for optical and active cable support

Interfaces

Ethernet SFP28 and QSFP28 Interfaces

The network ports of the ConnectX®-5 adapter card are compliant with the IEEE 802.3 Ethernet standards listed in Features and Benefits. Ethernet traffic is transmitted through the SFP28/QSFP28 connectors on the adapter card.



↑ Warning

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

PCI Express Interface

The table below describes the supported PCIe interface in ConnectX-5 and ConnectX-5 Ex adapter cards.

IC Model	Supported PCIe Interface	Features
ConnectX -5	PCIe Gen 3.0 (1.1 and 2.0 compatible) through x8 or x16 edge connectors	Link Rates: 2.5. 5.0 or 8.0GT/s. Auto Negotiation to: x16, x8, x4, x2 or x1. Support for MSI/MSI-X mechanisms.
ConnectX -5 Ex	PCIe Gen 3.0/4.0 (1.1 and 2.0 compatible) through x16 edge connectors	Link Rates: 2.5. 5.0, 8.0 or 16.0GT/s. Auto Negotiation to: x16, x8, x4, x2 or x1. Support for MSI/MSI-X mechanisms.

LED Interface

There is one bi-color (yellow/green) I/O LED per port to indicate link status located on the adapter card.

State	LED Behavior
Beacon command for locating the adapter card	Blinks yellow at 1Hz
Error	Blinks yellow at 4Hz until error is resolved. The reason for the error may be: • Failure of I ² C access to the port • Over-current conditions of the port
Link Activity	Blinks green reflecting the link speed
Link Up	Constant green light

SMBus Interface

ConnectX-5 technology maintains support for manageability through a BMC. ConnectX-5 PCIe stand-up adapter can be connected to a BMC using MCTP over SMBus or MCTP over PCIe protocols as if it is a standard NVIDIA PCIe stand-up adapter. For configuring the adapter for the specific manageability solution in use by the server, please contact NVIDIA Support.

Voltage Regulators

The voltage regulator power is derived from the PCI Express edge connector 12V supply pins. These voltage supply pins feed onboard regulators that provide the necessary power to the various components on the card.

Hardware Installation

Installation and initialization of ConnectX-5 adapter cards require attention to the mechanical attributes, power specification, and precautions for electronic equipment.

Safety Warnings



Note

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

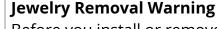
Please observe all safety warnings to avoid injury and prevent damage to system components. Note that not all warnings are relevant to all models.

Note that not all warnings are relevant to all models.



General Installation Instructions

Read all installation instructions before connecting the equipment to the power source.





Before you install or remove equipment that is connected to power lines, remove jewelry such as bracelets, necklaces, rings, watches, and so on. Metal objects heat up when connected to power and ground and can meltdown, causing serious burns and/or welding the metal object to the terminals.



Over-temperature

This equipment should not be operated in an area with an ambient temperature exceeding the maximum recommended: 55°C (131°F). An airflow of 200LFM at this maximum ambient temperature is required for HCA cards

	and NICs. To guarantee proper airflow, allow at least 8cm (3 inches) of clearance around the ventilation openings.
\(\frac{1}{2^2}\)	During Lightning - Electrical Hazard During periods of lightning activity, do not work on the equipment or connect or disconnect cables.
	Copper Cable Connecting/Disconnecting Some copper cables are heavy and not flexible, as such, they should be carefully attached to or detached from the connectors. Refer to the cable manufacturer for special warnings and instructions.
	Equipment Installation This equipment should be installed, replaced, or serviced only by trained and qualified personnel.
	Equipment Disposal The disposal of this equipment should be in accordance to all national laws and regulations.
	Local and National Electrical Codes This equipment should be installed in compliance with local and national electrical codes.
	Hazardous Radiation Exposure
	 Caution – Use of controls or adjustment or performance of procedures other than those specified herein may result in hazardous radiation exposure. For products with optical ports. CLASS 1 LASER PRODUCT and reference to the most recent laser standards: IEC 60 825-1:1993 + A1:1997 + A2:2001 and EN 60825-1:1994+A1:1996+

Installation Procedure Overview

A2:20

The installation procedure of ConnectX-5 adapter cards involves the following steps:

Step	Procedure
1	Check the system's hardware and software requirements.
2	Pay attention to the airflow consideration within the host system.

Step	Procedure
3	Unpack the product.
4	(Optional) Replace the full-height mounting bracket with the supplied short bracket.
5	Install the adapter card in the system.
6	Connect cables or modules to the card.
7	Identify ConnectX-5 adapter card in the system.

System Requirements

Hardware Requirements

(i) Important

Unless otherwise specified, NVIDIA products are designed to work in an environmentally controlled data center with low levels of gaseous and dust (particulate) contamination.

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

A system with a PCI Express slot of corresponding bus width is required for installing the card.

(i) Important

For proper operation and performance, please make sure to use a PCIe slot with a corresponding bus width and that can supply

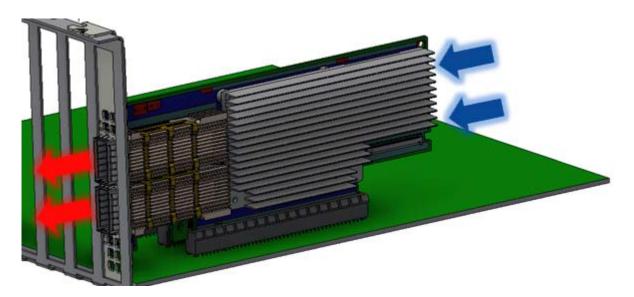
sufficient power to your card. Refer to the <u>Specifications</u> section of the manual for more power requirements.

Airflow Requirements

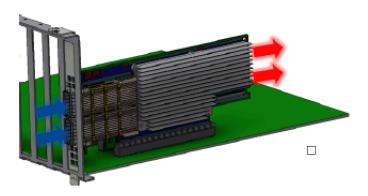
ConnectX-5 adapter cards are offered with two airflow patterns: from the heatsink to the network ports, and vice versa, as shown below.

Please refer to the <u>Specifications</u> section for airflow numbers for each specific card model.

Airflow *from the heatsink* to the network ports:



Airflow *from the network ports* to the heatsink:



(i) Important

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

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.

Unpacking the Product

(i) Important

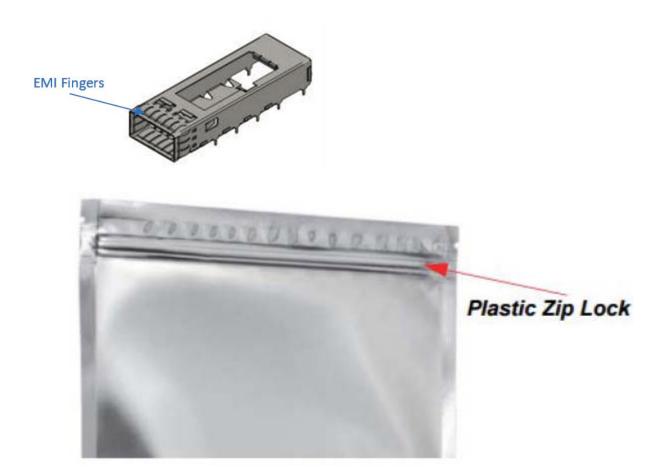
The adapter is being installed in a system that operates with voltages that can be lethal. Before opening the case of the system, observe the following precautions to avoid injury and prevent damage to system components.

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

- 1. Unpack the adapter package and place them on an antistatic surface, and verify you have received the following items:
 - 1. ConnectX-5 Adapter Card, with a tall bracket assembled on the card.
 - 2. Short bracket (accessory)
- 2. Check the parts for visible damage that may have occurred during shipping.

(i) Important

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.



3. 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-5 card, make sure that the system is disconnected from power.

Bracket Replacement Instructions

The card is usually shipped with an assembled high-profile bracket. If this form factor is suitable for your requirements, you can skip the remainder of this section and move to <u>Installation Instructions</u>. If you need to replace the high-profile bracket with the short bracket that is included in the shipping box, please follow the instructions in this section.



Important

Due to risk of damaging the EMI gasket, it is not recommended to replace the bracket more than three times.

To replace the bracket you will need the following parts:

- The new brackets of the proper height
- The 2 screws saved from the removal of the bracket

Removing the Existing Bracket

- 1. Using a torque driver, remove the two screws holding the bracket in place.
- 2. Separate the bracket from the ConnectX-5 card.



(i) Important

Be careful not to put stress on the LEDs on the adapter card.

3. Save the two screws.

Installing the New Bracket

1. Place the bracket onto the card until the screw holes line up.



(i) Important

Do not force the bracket onto the adapter card.

2. Screw on the bracket using the screws saved from the bracket removal procedure above.



(i) Important

Use a torque driver to apply up to 2 lbs-in torque on the screws.

Installation Instructions

This section provides detailed instructions on how to install your adapter card in a system.

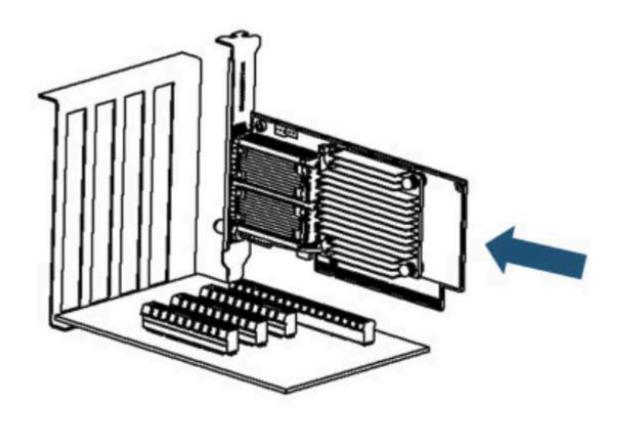


/ Warning

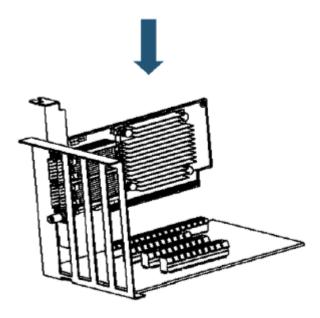
Please note that the following figures are for illustration purposes only.

1. Before installing the card, make sure that the system is off and the power cord is not connected to the server. Please follow proper electrical grounding procedures.

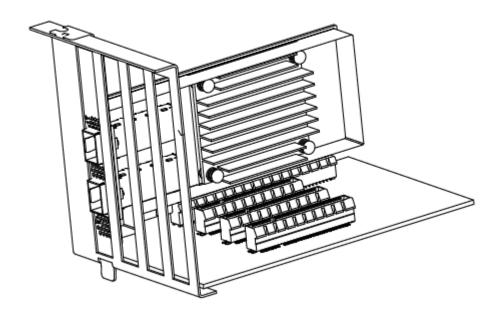
- 2. Open the system case.
- 3. Place the adapter in an available PCI Express slot.



4. Applying even pressure at both corners of the card, insert the adapter card into the PCI Express slot until firmly seated.



5. When the adapter is properly seated, the port connectors are aligned with the slot opening, and the adapter faceplate is visible against the system chassis.



6. Secure the adapter with the adapter clip or screw.



To uninstall the adapter card, see <u>Uninstalling the Card</u>.

Cables and Modules

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.
 - 1. 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.
 - 2. 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.
 - 3. 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.
 - 4. Make sure that the connector locks in place.



↑ Warning

When installing cables make sure that the latches engage.



Important

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 Green LED 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 <u>LED Interface</u> under the Interfaces section.
- 4. After plugging in a cable, lock the connector using the latching mechanism particular to the cable vendor. When data is being transferred the Green LED will blink. See LED Interface under the Interfaces section.
- 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. The LED indicator will turn off when the cable is unseated.

Identifying the Card in Your System

On Linux

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

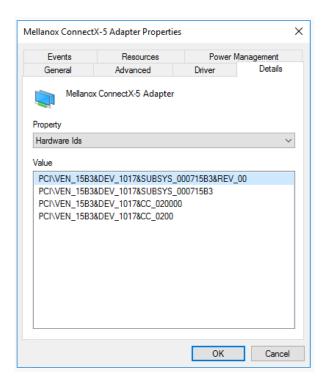
Ispci | grep -i Mellanox

Network controller: Mellanox Technologies MT28800 Family [ConnectX-5]

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-5 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 2012/R2/2016) from the **Property** pull-down menu.

PCI Device (Example)



5. 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_1003": VEN is equal to 0x15B3 – this is the Vendor ID of NVIDIA; and DEV is equal to 1018 (for ConnectX-5) – this is a valid NVIDIA PCI Device ID.



If the PCI device does not have a 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.

Uninstalling the Card

Safety Precautions

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

- 1. Remove any metallic objects from your hands and wrists.
- 2. It is strongly recommended to use an ESD strap or other antistatic devices.
- 3. Turn off the system and disconnect the power cord from the server.

Card Removal

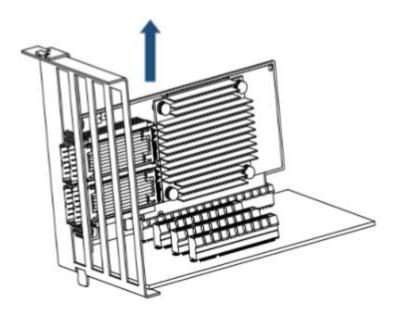
Please note that the following images are for illustration purposes only.



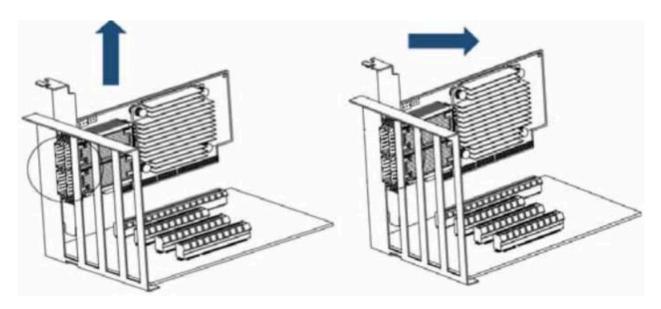
Note

Please note that the following images are for illustration purposes only.

- 1. Verify that the system is powered off and unplugged.
- 2. Wait 30 seconds.
- 3. To remove the card, disengage the retention mechanisms on the bracket (clips or screws).
- 4. Holding the adapter card from its center, gently pull the adapter card out of the PCI Express slot.



5. When the port connectors reach the top of the chassis window, gently pull the adapter card in parallel to the motherboard.



Driver Installation

Please use the relevant driver installation section.

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

Linux Driver Installation

This section describes how to install and test the NVIDIA OFED for Linux package on a single server with a ConnectX-5 adapter card installed.

Prerequisites

Requirements	Description
Platforms	 A server platform with one of the following adapters installed: ConnectX®-5 Ex (EN) (firmware: fw-ConnectX5) ConnectX®-5 (EN) (firmware: fw-ConnectX5)
Required Disk Space for Installation	1GB
Operating System	Linux operating system. For the list of supported operating system distributions and kernels, please refer to the MNLX_OFED Release Notes
Installer Privileges	The installation requires administrator (root) privileges on the target machine.

Downloading NVIDIA OFED

1. Verify that the system has an NVIDIA network adapter installed.

The following example shows a system with an installed NVIDIA adapter card:

lspci -v | grep Mellanox

86:00.0 Network controller [0207]: Mellanox Technologies MT27620 Family

Subsystem: Mellanox Technologies Device 0014

86:00.1 Network controller [0207]: Mellanox Technologies MT27620 Family

Subsystem: Mellanox Technologies Device 0014

2. Download the ISO image to your host.

The image's name has the format MLNX_OFED_LINUX-<ver>-<OS label><CPU arch>.iso.
You can download and install the latest OpenFabrics Enterprise Distribution (OFED) software package available via the NVIDIA web site at nvidia.com/en-us/networking
_ Products Software InfiniBand Drivers NVIDIA MLNX OFED

1.

- 1. Scroll down to the Download wizard, and click the Download tab.
- 2. Choose your relevant package depending on your host operating system.
- 3. Click the desired ISO/tgz package.
- 4. 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.

SHA256 MLNX_OFED_LINUX-<ver>--<OS label>.iso

Installing MLNX_OFED

Installation Script

The installation script, mlnxofedinstall, performs the following:

o Discovers the currently installed kernel

- 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)
- 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.



Warning

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

- 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).
- 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.

↑ Warning

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-kernelsupport' option below.

↑ Warning

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

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

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

Do you want to continue?[y/N]:y

See log file /tmp/mlnx_ofed_iso.21642.log

Building OFED RPMs. Please wait...

Removing OFED RPMs...

Created /tmp/MLNX_OFED_LINUX-x.x-x-rhel7.3-x86_64-ext.tgz

- 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.

- 1. Log in to the installation machine as root.
- 2. Mount the ISO image on your machine.

host1# mount -o ro,loop MLNX_OFED_LINUX-<ver>--<OS label>--<CPU arch>.iso /mnt

3. Run the installation script.

/mnt/mlnxofedinstall Logs dir: /tmp/MLNX_OFED_LINUX-x.x-x.logs This program will install the MLNX_OFED_LINUX package on your machine. Note that all other Mellanox, OEM, OFED, RDMA or Distribution IB packages will be removed.

Those packages are removed due to conflicts with MLNX_OFED_LINUX, do not reinstall them.

Starting MLNX_OFED_LINUX-x.x.x installation ...

Installation finished successfully.

Attempting to perform Firmware update...

Querying Mellanox devices firmware ...

↑ Warning

For unattended installation, use the --force installation option while running the MLNX_OFED installation Script:/mnt/mlnxofedinstall --force

↑ Warning

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 --forceFor example:./mlnxofedinstall --without-dkms --add-kernelsupport --kernel 3.13.0-85-generic --without-fw-update -forceNote that the path to kernel sources (--kernelsources) 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-XPart Number: MCXXXX-XXXPSID: MT_<version>PCI Device Name: 0b:00.0Base MAC: 0000e41d2d5cf810Versions: Current AvailableFW XX.XX.XXXXStatus: Up to date

/\ Warning

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-XPart Number: MCXXXX-XXXPSID: MT <version>PCI Device Name: 0b:00.0Base MAC: 0000e41d2d5cf810Versions: Current Available FW XX.XX.XXXXStatus: 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.

4. 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.

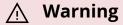
Action \ Adapter	Driver Restart	Standard Reboot (Soft Reset)	Cold Reboot (Hard Reset)
Standard ConnectX- 4/ConnectX-4 Lx or higher	-	+	-
Adapters with Multi-Host Support	-	-	+
Socket Direct Cards	-	-	+

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".

- 5. (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:
- HCA firmware version
- 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

Soft ware	 Most of MLNX_OFED packages are installed under the "/usr" directory except for the following packages which are installed under the "/opt" directory: 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
Firm ware	 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:

```
Logs 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:



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

```
blacklist mlx5_core
blacklist mlx5_ib
```

- 2. Set "ONBOOT=no" in the "/etc/infiniband/openib.conf" file.
- 3. 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

Return Code	Meaning
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
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

Soft ware	 Most of MLNX_OFED packages are installed under the "/usr" directory except for the following packages which are installed under the "/opt" directory: 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
Firm ware	 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:

Logs 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

- 1. Log into the installation machine as root.
- 2. 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

3. 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

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

4. Install the key.

```
# sudo rpm --import RPM-GPG-KEY-Mellanox
warning: rpmts_HdrFromFdno: Header V3 DSA/SHA1 Signature, key ID
6224c050: NOKEY
Retrieving 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
Is this ok [y/N]:
```

5. Check that the key was successfully imported.

```
# rpm -q gpg-pubkey --qf '%{NAME}-%{VERSION}-%{RELEASE}\t%{SUMMARY}\n' | grep Mellanox gpg-pubkey-a9e4b643-520791ba gpg(Mellanox Technologies <support@mellanox.com>)
```

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

```
[mlnx_ofed]
name=MLNX_OFED Repository
baseurl=file:///<path to extracted MLNX_OFED package>/RPMS
enabled=1
gpgkey=file:///<path to the downloaded key RPM-GPG-KEY-Mellanox>
gpgcheck=1
```

7. Check that the repository was successfully added.

```
# yum repolist
Loaded plugins: product-id, security, subscription-manager
This system is not registered to Red Hat Subscription Management. You
can use subscription-manager to register.
repo id repo name status
mlnx_ofed MLNX_OFED Repository 108
rpmforge RHEL 6Server - RPMforge.net - dag 4,597
repolist: 8,351
```

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

- 1. Log into the installation machine as root.
- 2. 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
```

3. 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/
Note: 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

Checking if all needed packages are installed...

Building MLNX_OFED_LINUX RPMS . Please wait...

Creating metadata-rpms for 3.10.0-957.21.3.el7.x86_64 ...

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

WARNING: that it contains unsigned rpms, therefore, you need to disable the gpgcheck

WARNING: by setting 'gpgcheck=0' in the repository conf file.

Created /tmp/MLNX_OFED_LINUX-5.2-0.5.5.0-rhel7.6-x86_64-ext.tgz

4. Open the tarball.

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

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

```
[mlnx_ofed]
name=MLNX_OFED Repository
baseurl=file:///<path to extracted MLNX_OFED package>/RPMS
enabled=1
gpgcheck=0
```

6. Check that the repository was successfully added.

yum repolist

Loaded plugins: product-id, security, subscription-manager This system is not registered to Red Hat Subscription Management. You can use subscription-manager to register.

repo id repo name status

mlnx_ofed MLNX_OFED Repository 108

rpmforge RHEL 6Server - RPMforge.net - dag 4,597

repolist: 8,351

Installing MLNX_OFED Using the YUM Tool

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

1. View the available package groups by invoking:

yum search mlnx-ofedmlnx-ofed-all.noarch: MLNX_OFED all installer package (with KMP support) mlnx-ofed-all-user-only.noarch: MLNX_OFED all-user-only installer package (User Space packages only) mlnx-ofed-basic.noarch: MLNX_OFED basic installer package (with KMP) support) mlnx-ofed-basic-user-only.noarch: MLNX_OFED basic-user-only installer package (User Space packages only) mlnx-ofed-bluefield.noarch: MLNX_OFED bluefield installer package (with KMP support) mlnx-ofed-bluefield-user-only.noarch : MLNX_OFED bluefield-user-only installer package (User Space packages only) mlnx-ofed-dpdk.noarch: MLNX_OFED dpdk installer package (with KMP support) mlnx-ofed-dpdk-upstream-libs.noarch: MLNX_OFED dpdk-upstream-libs installer package (with KMP support) mlnx-ofed-dpdk-upstream-libs-user-only.noarch: MLNX_OFED dpdkupstream-libs-user-only installer package (User Space packages only) mlnx-ofed-dpdk-user-only.noarch: MLNX_OFED dpdk-user-only installer package (User Space packages only) mlnx-ofed-eth-only-user-only.noarch: MLNX_OFED eth-only-user-only installer package (User Space packages only) mlnx-ofed-guest.noarch: MLNX_OFED guest installer package (with KMP) support) mlnx-ofed-guest-user-only.noarch: MLNX_OFED guest-user-only installer package (User Space packages only)

mlnx-ofed-hpc.noarch : MLNX_OFED hpc installer package (with KMP support)

mlnx-ofed-hpc-user-only.noarch : MLNX_OFED hpc-user-only installer package (User Space packages only)

mlnx-ofed-hypervisor.noarch : MLNX_OFED hypervisor installer package (with KMP support)

mlnx-ofed-hypervisor-user-only.noarch: MLNX_OFED hypervisor-user-only installer package (User Space packages only)

mlnx-ofed-kernel-only.noarch : MLNX_OFED kernel-only installer package (with KMP support)

mlnx-ofed-vma.noarch: MLNX_OFED vma installer package (with KMP support)

mlnx-ofed-vma-eth.noarch : MLNX_OFED vma-eth installer package (with KMP support)

mlnx-ofed-vma-eth-user-only.noarch : MLNX_OFED vma-eth-user-only installer package (User Space packages only)

mlnx-ofed-vma-user-only.noarch : MLNX_OFED vma-user-only installer package (User Space packages only)

mlnx-ofed-vma-vpi.noarch : MLNX_OFED vma-vpi installer package (with KMP support)

mlnx-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:

mlnx-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) mlnx-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) mlnx-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) mlnx-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) mlnx-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) mlnx-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) mlnx-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) mlnx-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) mlnx-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)
mlnx-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)
mlnx-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)
mlnx-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 repository configuration file.

2. Install the desired group.



/ Warning

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

- 1. Log into the installation machine as root.
- 2. 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.
- 3. Create an apt-get repository configuration file called "/etc/apt/sources.list.d/mlnx_ofed.list" with the following content:

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

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

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

5. Verify that the key was successfully imported.

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

6. Update the apt-get cache.

sudo apt-get update

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

- 1. Log into the installation machine as root.
- 2. 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

3. 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/

Note: 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

Checking if all needed packages are installed...

Building MLNX_OFED_LINUX RPMS . Please wait...

Creating metadata-rpms for 3.10.0-957.21.3.el7.x86_64 ...

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

WARNING: that it contains unsigned rpms, therefore, you need to disable the gpgcheck

WARNING: by setting 'gpgcheck=0' in the repository conf file.

Created /tmp/MLNX_OFED_LINUX-5.2-0.5.5.0-rhel7.6-x86_64-ext.tgz

4. Open the tarball.

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

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

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

6. 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:

1. View the available package groups by invoking:

```
# apt-cache search mlnx-ofed-
apt-cache search mlnx-ofed ........
knem-dkms - DKMS support for mlnx-ofed kernel modules
mlnx-ofed-kernel-dkms - DKMS support for mlnx-ofed kernel modules
mlnx-ofed-kernel-utils - Userspace tools to restart and tune mlnx-ofed
kernel modules
mlnx-ofed-vma-vpi - MLNX_OFED vma-vpi installer package (with DKMS support)
mlnx-ofed-kernel-only - MLNX_OFED kernel-only installer package (with DKMS support)
mlnx-ofed-bluefield - MLNX_OFED bluefield installer package (with DKMS support)
mlnx-ofed-hpc-user-only - MLNX_OFED hpc-user-only installer package (User Space packages only)
```

mlnx-ofed-dpdk-user-only - MLNX_OFED dpdk-user-only installer package (User Space packages only)

mlnx-ofed-all-exact - MLNX_OFED all installer package (with DKMS support) (exact)

mlnx-ofed-all - MLNX_OFED all installer package (with DKMS support) mlnx-ofed-vma-vpi-user-only - MLNX_OFED vma-vpi-user-only installer package (User Space packages only)

mlnx-ofed-eth-only-user-only - MLNX_OFED eth-only-user-only installer package (User Space packages only)

mlnx-ofed-vma-user-only - MLNX_OFED vma-user-only installer package (User Space packages only)

mlnx-ofed-hpc - MLNX_OFED hpc installer package (with DKMS support) mlnx-ofed-bluefield-user-only - MLNX_OFED bluefield-user-only installer package (User Space packages only)

mlnx-ofed-dpdk - MLNX_OFED dpdk installer package (with DKMS support) mlnx-ofed-vma-eth-user-only - MLNX_OFED vma-eth-user-only installer package (User Space packages only)

mlnx-ofed-all-user-only - MLNX_OFED all-user-only installer package (User Space packages only)

mlnx-ofed-vma-eth - MLNX_OFED vma-eth installer package (with DKMS support)

mlnx-ofed-vma - MLNX_OFED vma installer package (with DKMS support) mlnx-ofed-dpdk-upstream-libs-user-only - MLNX_OFED dpdk-upstream-libs-user-only installer package (User Space packages only)

mlnx-ofed-basic-user-only - MLNX_OFED basic-user-only installer package (User Space packages only)

mlnx-ofed-basic-exact - MLNX_OFED basic installer package (with DKMS support) (exact)

mlnx-ofed-basic - MLNX_OFED basic installer package (with DKMS support) mlnx-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

2. Install the desired group.

apt-get install '<group name>'

Example:

apt-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:# aptget 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 Performance Tuning Guide for NVIDIA Network Adapters.

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.

Software Requirements

Description	Package
Windows Server 2022	
Windows Server 2019	
Windows Server 2016	
Windows Server 2012 R2	MLNX_WinOF2- <version>_All_x64.exe</version>
Windows 11 Client (64 bit only)	
Windows 10 Client (64 bit only)	
Windows 8.1 Client (64 bit only)	

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

Downloading WinOF-2 Driver

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

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

echo %PROCESSOR_ARCHITECTURE%



↑ Warning

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 Step

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 64bit .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"

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.

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).

MLNX_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).

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

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

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

↑ Warning

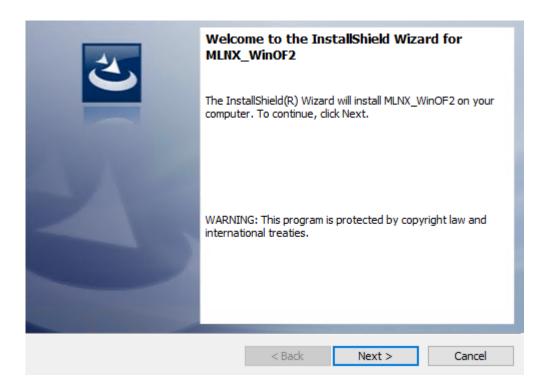
The Rshim driver installanion will fail if a prior Rshim driver is already installed. The following fail message will be displayed in the log:

"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>"

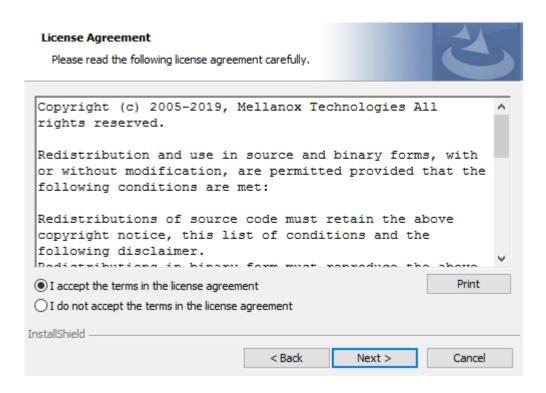
5. **[Optional]** If you want to skip the check for unsupported devices, run.

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

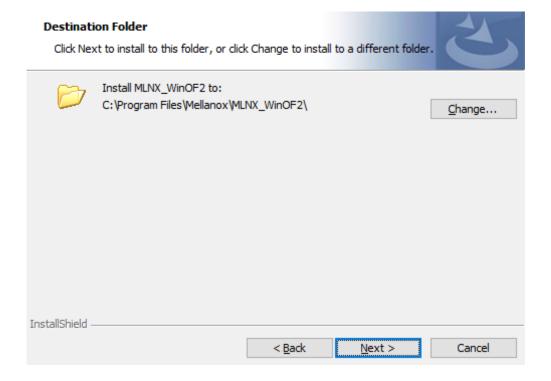
6. Click Next in the Welcome screen.



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

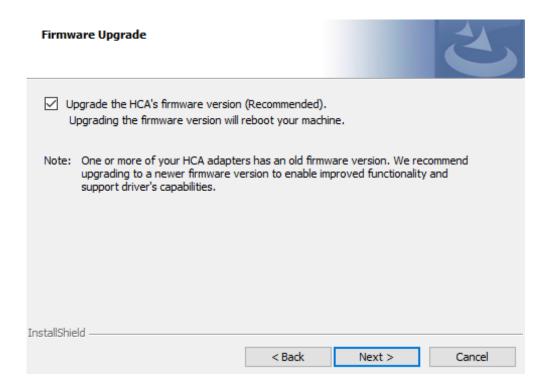


8. Select the target folder for the installation.

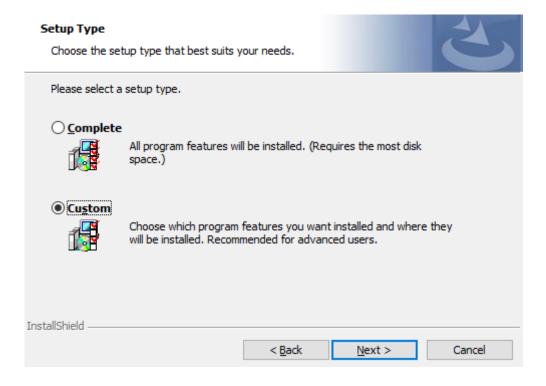


- 9. The firmware upgrade screen will be displayed in the following cases:
 - o If the user has an OEM card. In this case, the firmware will not be displayed.

 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.

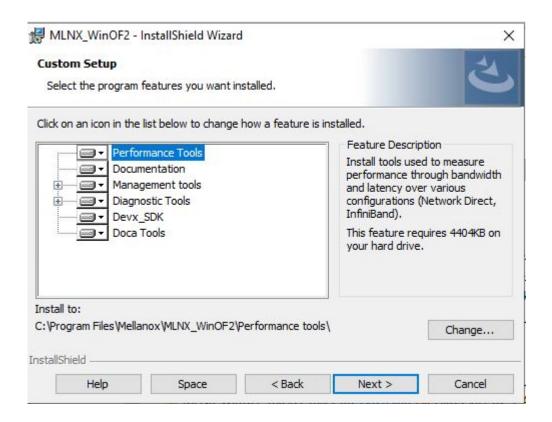


10. Select a Complete or Custom installation, follow <u>Step a</u> onward.

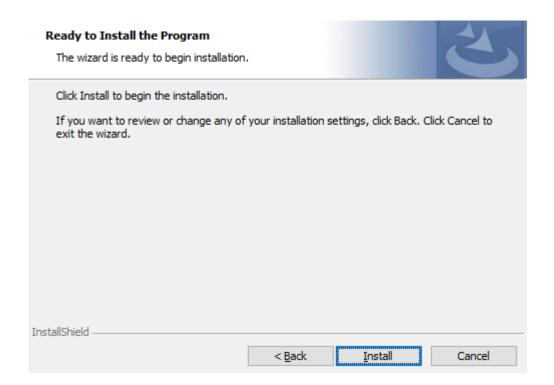


1. Select the desired feature to install:

- 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
- 2. Click Next to install the desired tools.



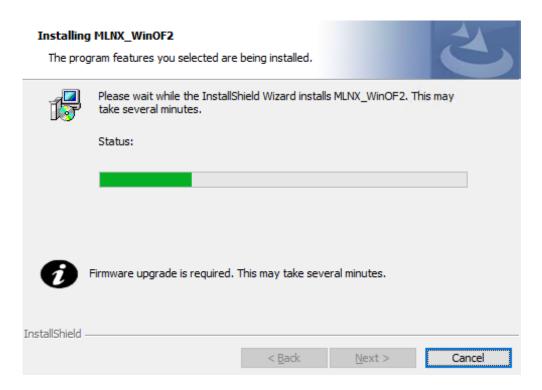
11. Click Install to start the installation.



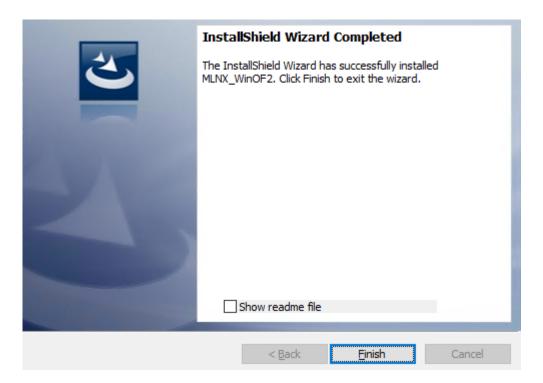
12. In case firmware upgrade option was checked in <u>Step 7</u>, you will be notified if a firmware upgrade is required (see



).



13. Click Finish to complete the installation.



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:

MLNX_WinOF2-[Driver/Version]_<revision_version>_All_-Arch.exe /S /v/qn

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

MLNX_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).

MLNX_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).

MLNX_WinOF2-[Driver/Version] < revision_version > All_Arch.exe /vMT_SKIPFWUPGRD=1

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

MLNX_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 /I*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"

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

VMware Driver Installation

This section describes VMware Driver Installation.

Hardware and Software Requirements

Requirement	Description
Platforms	A server platform with an adapter card based on one of the following devices: • ConnectX®-5 (EN) (firmware: fw-ConnectX5) • ConnectX®-5 Ex (EN) (firmware: fw-ConnectX5)
Operating System	ESXi 6.5
Installer Privileges	The installation requires administrator privileges on the target machine.

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.

esxcli software vib list | grep nmlx nmlx5-core 4.16.8.8-10EM.650.0.0.4240417 MEL PartnerSupported 2017-01-31 nmlx5-rdma 4.16.8.8-10EM.650.0.0.4240417 MEL PartnerSupported 2017-01-31

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

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.

Firmware Programming

- 1. Download the VMware bootable binary images v4.6.0 from the Firmware Tools (MFT) site.
 - 1. ESXi 6.5 File: mft-4.6.0.48-10EM-650.0.0.4598673.x86_64.vib
 - 2. 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.

Updating Adapter Firmware

Each adapter card is shipped with the latest version of qualified firmware at the time of manufacturing. 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 NVIDIA 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.

[server1]# ./mlxup

Querying Mellanox devices firmware ...

Device Type: ConnectX-5

Part Number: MCX516A-CDAT

Description: ConnectX®-5 Ex EN network interface card, 100GbE dual-port QSFP28,

PCIe4.0 x16, tall bracket PSID: MT_2190110032

PCI Device Name: 0000:06:00.0 Base GUID: e41d2d0300fd8b8a

Versions: Current Available FW 16.23.1020 16.24.1000

Status: Update required

Device Type: ConnectX-5

Part Number: MCX516A-CDAT

Description: ConnectX®-5 Ex EN network interface card, 100GbE dual-port QSFP28,

PCIe4.0 x16, tall bracket PSID: MT_2170110021

PCI Device Name: 0000:07:00.0

Base MAC: 0000e41d2da206d4 Versions: Current Available FW 16.24.1000 16.24.1000

Status: Up to date

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

Troubleshooting

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 communicati on 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 PCle 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 PCle 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.

Linux Troubleshooting

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

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> 	

Specifications

MCX512A-ADAT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Dual SFP28 (copper and optical)			
Protocol Support	Ethernet: 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X Data Rate: 1/10/25 Gb/s Ethernet PCI Express Gen3/Gen4: SERDES @ 8.0GT/s / 16.0GT/s (respectively), 8 lanes (2.0 and 1.1 compatible)			
	Voltage: 12V			
	Power	Cable Type		
	Typical Power ^(a)	Passive Cables	12.5W	
Power and Airflow	Maximum Power	Passive Cables	17.5W	
		1.5W Active Cables	21W	
	Maximum power available through SFP28 port: 1.5W			
	Airflow	Airflow Specifications.		
	Temperature	Operational	0°C to 55°C	
		Non-operational	-40°C to 70°C ^(b)	
Environment	Humidity	Operational	10% to 85% relative humidity	
al		Non-operational	10% to 90% relative humidity	
	Altitude (Operational)	3050m		
Regulatory	Safety	CB / cTUVus / CE		

EMC	CE / FCC / VCC	CI / ICES / RCM
RoHS	RoHS complia	nt

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

MCX516A-BDAT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Dual QSFP28 (copper and optical)			
Protocol Support	Ethernet: 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X Data Rate: 1/10/25/40 Gb/s Ethernet PCI Express Gen3/Gen4: SERDES @ 8.0GT/s / 16.0GT/s (respectively), 16 lanes (2.0 and 1.1 compatible)			
	Voltage: 12V			
	Power	Cable Type		
	Typical Power ^(a)	Passive Cables	18.4W	
Power and	Maximum Power	Passive Cables	21.6W	
Airflow Airflow		1.5W Active Cables	24.9W	
		2.5W Active Cables	26.1W	
	Maximum power available through QSFP28 port: 3.5W			
	Airflow Specifications.			
Environm	Temperature	Operational	0°C to 55°C	
ental		Non-operational	-40°C to 70°C ^(b)	
	Lumidity	Operational	10% to 85% relative humidity	
	Humidity	Non-operational	10% to 90% relative humidity	

	Altitude (Operational)	3050m
	Safety	CB / cTUVus / CE
Regulator v	EMC	CE / FCC / VCCI / ICES / RCM
,	RoHS	RoHS compliant

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

MCX516A-CDAT Specifications

Physical		Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Dual QSFP28 (copper and optical)		
Protocol Support	R4, 50G Ethernet Co LR4, 40GBASE-ER4, Ethernet Consortium 10GBASE-CX4, SGM Data Rate: 1/10/25/ PCI Express Gen3/G 16 lanes (2.0 and 1.1 Note: PCIe 4.0 x16 but * 16GT/s, including of	Ethernet: 100GBASE-CR4, 100GBASE-KR4, 100GBASE-SR4, 56GBASE-R4, 50G Ethernet Consortium, 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X Data Rate: 1/10/25/40/50/100 Gb/s Ethernet PCI Express Gen3/Gen4: SERDES @ 8.0GT/s / 16.0GT/s (respectively), 16 lanes (2.0 and 1.1 compatible) Note: PCIe 4.0 x16 bus can supply a maximum bandwidth of 256Gb/s (=16 * 16GT/s, including overhead), and can support 200Gb/s when both network ports of MCX516A-CDAT run at 100Gb/s.		
Power and	Voltage: 12V	Voltage: 12V		
Airflow	Power	Cable Type		
	Typical Power ^(a)	Passive Cables	19.3W	
		Passive Cables	23.4W	
	Maximum Power	1.5W Active Cable	26.8W	
		2.5W Active Cable	29.0W	
	Maximum power available through QSFP28 port: 5W			
	Airflow See <u>Airflow</u>			

		<u>Specifications</u> .	
	Temperature	Operational	0°C to 55°C
		Non-operational	-40°C to 70°C ^(b)
Environmental	Humidity	Operational	10% to 85% relative humidity
		Non-operational	10% to 90% relative humidity
	Altitude (Operational)	3050m	
	Safety	CB / cTUVus / CE	
Regulatory	EMC	CE / FCC / VCCI / ICES / RCM	
	RoHS	RoHS-compliant	

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

MCX512A-ACAT and MCX512A-ACUT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Dual SFP28 Ethernet (copper and optical)		
Protocol Support	Ethernet: 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, SGMII / 1000BASE-X Data Rate: 1/10/25 Gb/s Ethernet PCI Express Gen3: SERDES @ 8.0GT/s, 8 lanes (2.0 and 1.1 compatible)		
Power and	Voltage: 12V		
Airflow	Power	Cable Type	
	Typical Power ^(a)	Passive Cable	9.5W
	Maximum Power	Passive Cable	11.8W
		1.5W Active Cable	15.1W
	Maximum power available through SFP28 port: 1.5W		

	Airflow	See <u>Airflow Specifications</u> .	
	Temperature	Operational	0°C to 55°C
		Non-operational	-40°C to 70°C ^(b)
Environm ental	Humidity	Operational	10% to 85% relative humidity
		Non-operational	10% to 90% relative humidity
	Altitude (Operational)	3050m	
_	Safety	CB / cTUVus / CE	
Regulator y	EMC	CE / FCC / VCCI / ICES / RCM	
	RoHS	RoHS-compliant	

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

MCX512F-ACAT and MCX512F-ACHT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Dual SFP28 Ethernet (copper and optical)			
Protocol Support	Ethernet: 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, SGMII / 1000BASE-X Data Rate: 1/10/25 Gb/s Ethernet PCI Express Gen3: SERDES @ 8.0GT/s, 16 lanes (2.0 and 1.1 compatible)			
	Voltage: 12V			
	Power	Cable Type		
Power and	Typical Power ^(a)	Passive Cable	12.7W	
Airflow		Passive Cable	15.1W	
	Maximum Power	1.5W Active Cable	18.5W	
	Maximum power available through SFP28 port: 2.5W			

Environm ental	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		
Airflow	See <u>Airflow Specifications</u> .			
	Safety	CB / cTUVus / CE		
Regulator y	EMC	CE / FCC / VCCI / ICES / RCM		
	RoHS	RoHS-compliant		

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

MCX515A-GCAT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Single QFP28 (copper and optical)		
Protocol Support	Ethernet: 50G Ethernet Consortium, 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X Data Rate: 1/10/25/40/50 Ethernet PCI Express Gen3: SERDES 8.0GT/s, 16 lanes (2.0 and 1.1 compatible)		
Power and	Voltage: 12V		
Airflow	Power	Cable Type	
	Typical Power ^(a)	Passive Cable	12.7W
		Passive Cable	15.1W
	Maximum Power	1.5W Active Cable	18.5W
		2.5W Active Cable	20.7W

	Maximum power available through QSFP28 port: 3.5W			
	Airflow	See <u>Airflow Specifications</u> .		
	_	Operational	0°C to 55°C	
	Temperature	Non-operational	-40°C to 70°C ^(b)	
Environmen tal	Humidity	Operational	10% to 85% relative humidity	
		Non-operational	10% to 90% relative humidity	
	Altitude (Operational)	3050m		
	Safety	CB / cTUVus / CE		
Regulatory	EMC	CE / FCC / VCCI / ICES / RCM		
	RoHS	RoHS-compliant		

MCX516A-GCAT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Dual QSFP28 (copper and optical)			
Protocol Support	Ethernet: 50G Ethernet Consortium, 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X Data Rate: 1/10/25/40/50 Gb/s Ethernet PCI Express Gen3: SERDES @ 8.0GT/s, 16 lanes (2.0 and 1.1 compatible)			
Power and	Voltage: 12V			
Airflow	Power	Cable Type		
	Typical Power ^(a)	Passive Cable	14.7W	
	Maximum Power	Passive Cable	17.2W	
		1.5W Active Cable	20.6W	

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

		2.5W Active Cable	22.8W	
	Maximum power available through QSFP28 port: 3.5W			
	Airflow	See <u>Airflow Specifications</u> .		
	Tamanayahuya	Operational	0°C to 55°C	
	Temperature	Non-operational	-40°C to 70°C ^(b)	
Environmen tal	Humidity	Operational	10% to 85% relative humidity	
		Non-operational	10% to 90% relative humidity	
	Altitude (Operational)	3050m		
	Safety	CB / cTUVus / CE		
Regulatory	EMC	CE / FCC / VCCI / ICES / RCM		
	RoHS	RoHS-compliant		

MCX515A-CCAT and MCX515A-CCUT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Single QSFP28 (copper and optical)			
Protocol Support	Ethernet: 100GBASE-CR4, 100GBASE-KR4, 100GBASE-SR4, 56GBASE-R4, 50G Ethernet Consortium, 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X Data Rate: 1/10/25/40/50/100 Gb/s Ethernet PCI Express Gen3: SERDES @ 8.0GT/s, 16 lanes (2.0 and 1.1 compatible)			
Power and	Power and Voltage: 12V			
Airflow	Power	Cable Type		
	Typical Power ^(a) :	Passive Cable	14.2W	

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

	Maximum Power:	Passive Cable	17.0W
		1.5W Active Cable	18.6W
		2.5W Active Cable	19.7W
	Maximum power availabl	e through QSFP28 port: 5W	
	Airflow	See <u>Airflow Specifications</u> .	
	Temperature	Operational	0°C to 55°C
		Non-operational	-40°C to 70°C ^(b)
Environmen tal	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	
	RoHS:	RoHS-compliant	

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

MCX516A-CCAT and MCX516A-CCHT Specifications

Physical	Size: 2.71 in. x 5.6 in. (68.90mm x 142.24 mm) – low profile Connector: Dual QSFP28 (copper and optical)
Protocol Support	Ethernet: 100GBASE-CR4, 100GBASE-KR4, 100GBASE-SR4, 56GBASE-R4, 50G Ethernet Consortium, 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X Data Rate: 1/10/25/40/50/100 Gb/s Ethernet PCI Express Gen3: SERDES @ 8.0GT/s, 16 lanes (2.0 and 1.1 compatible) Note: PCIe 3.0 x16 bus can supply a maximum bandwidth of 128Gb/s only (=16 * 8GT/s, including overhead), and therefore cannot support 200Gb/s

	when both network ports of MCX516A-CCAT run at 100Gb/s.			
	Voltage: 12V			
	Power	Cable Type		
	Typical Power ^(a) :	Passive Cable	16.2W	
Power and		Passive Cable	19.0W	
Airflow	Maximum Power:	1.5W Active Cable	22.3W	
		2.5W Active Cable	24.6W	
	Maximum power available through QSFP28 port: 5W			
	Airflow:	See <u>Airflow Specifications</u> .		
	Temperature	Operational	0°C to 55°C	
		Non-operational	-40°C to 70°C ^(b)	
Environmen tal	Humidity	Operational	10% to 85% relative humidity	
		Non-operational	10% to 90% relative humidity	
	Altitude (Operational)	3050m		
	Safety:	CB / cTUVus / CE		
Regulatory	EMC:	CE / FCC / VCCI / ICES / RCM		
	RoHS:	RoHS-compliant		

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

Airflow Specifications

Airflow (LFM) with airflow direction from heatsink to port.

	Cable Type				
OPN	Passive	Active 1.5W	Active 2.5W	Active 3.5W	Active 5W

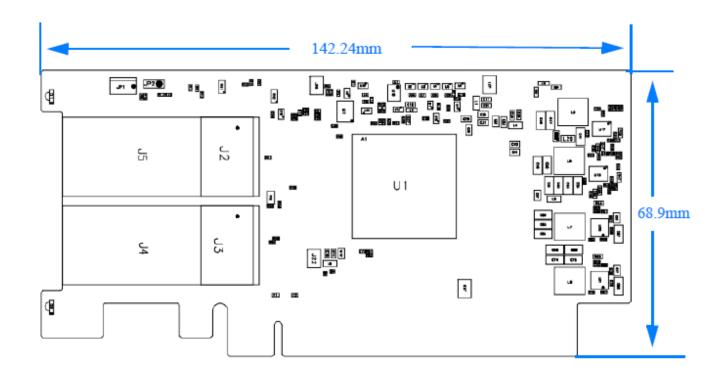
	Cable Type				
MCX512A-ADAT	200LF M	350LFM	N/A	N/A	N/A
MCX516A-BDAT	300LF M	350LFM	400LFM	400LFM	N/A
MCX516A-CDAT	400LF M	600LFM	850LFM	1000LFM	1400LFM
MCX512A-ACAT / MCX512A-ACUT	150LF M	300LFM	N/A	N/A	N/A
MCX512F-ACAT / MCX512F-ACHT	200LF M	300LFM	N/A	N/A	N/A
MCX515A-GCAT	300LF M	350LFM	400LFM	400LFM	N/A
MCX516A-GCAT	350LF M	500LFM	600LFM	650LFM	N/A
MCX515A-CCAT/ MCX515A-CCUT	300LF M	350LFM	400LFM	400LFM	950LFM
MCX516A-CCAT / MCX516A-CCHT	350LF M	500LFM	600LFM	650LFM	1300LFM

Board Mechanical Drawing and Dimensions

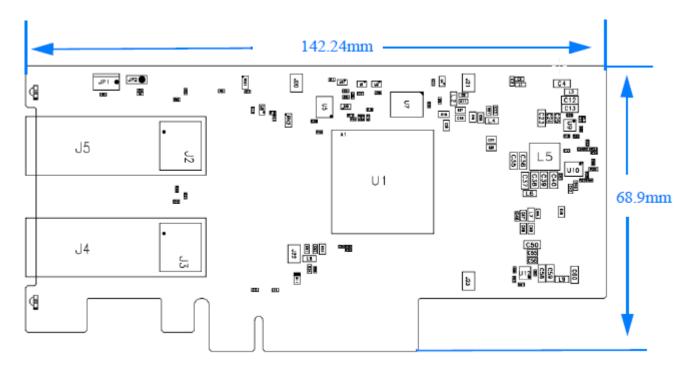


All dimensions are in millimeters. All the mechanical tolerances are +/- 0.1mm.

Mechanical Drawing of a Dual-Port x16 Card

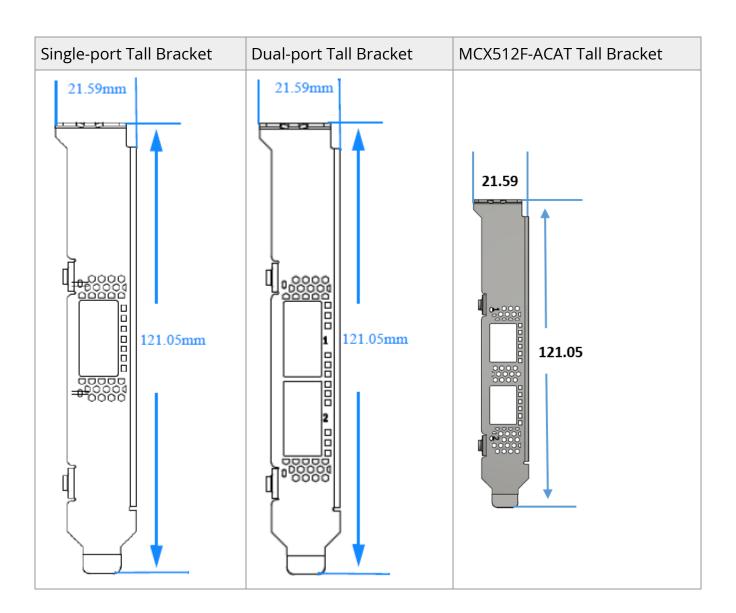


Mechanical Drawing of a Dual-Port x8 Card



Bracket Mechanical Drawing

Dimensions in millimeters.



Monitoring

Thermal Sensors

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

There are three thermal threshold definitions that 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 will be 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 will automatically shut down the device.
- **Emergency** 130°C: If the firmware fails to shut down the device upon crossing the Critical threshold, the device will auto-shutdown 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 in accordance with the readouts and the needs of the above-mentioned IC thermal requirements.

Adapter Heatsink

A heatsink is attached to the ConnectX-5 IC in order to dissipate the heat from the ConnectX- 5 IC. It is attached either by using four spring-loaded push pins that insert into four mounting holes or by screws.

ConnectX-5 IC has a thermal shutdown safety mechanism that automatically shuts down the ConnectX-5 card in cases of a high-temperature event, improper thermal coupling, or heatsink removal.

Finding the MAC on the **Adapter Card**

Each NVIDIA adapter card has a different identifier printed on the label: serial number and the card MAC for the Ethernet protocol.



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

MCX512A-ADAT Board Label (Example)

MCX512A-ACAT Board Label (Example)



MCX512F-ACAT Board Label (Example)



MCX512F-ACHT Board Label (Example)

TBD

MCX516A-BDAT Board Label (Example)



MCX515A-GCAT Board Label (Example)



MCX516A-GCAT Board Label (Example)

Model No: CX516A

ConnectX-5 50GbE
P/N: MCX516A-GCAT

S/N: MT0806X01504

MAC: 00 02 C9 27 05 1C

Made in Israel

MCX515A-CCAT Board Label (Example)



MCX516A-CCAT Board Label (Example)



MCX516A-CCHT Board Label (Example)

Model: CX516A

ConnectX-5 100GbE

P/N: MCX516A-CCHT

S/N: MT0806X01504

MAC: 00 02 C9 27 05 1C

Made in Israel



2020-10-22

MCX516A-CDAT Board Label (Example)

Model No: CX516A

ConnectX-5 100GbE

P/N: MCX516A-CDAT

S/N: MT0806X01504

MAC: 00 02 C9 27 05 1C

Made in Israel



2016-11-28 Rev: A1

Document Revision History

Date	Description of Changes
Jan. 2024	Added a note that the ConnectX-5 product line has moved to end-of-life.
Dec. 2023	Updated <u>Specifications</u> with MCX512F-ACAT bracket.
May. 2023	Added non-operational storage temperature specifications.
Nov. 2022	Updated Specifications tables.
Aug. 2022	Added a note on FRU EEPROM memory component under the Features and Benefits table.
Jul. 2021	Updated power numbers for MCX512A-ADAT.
Mar. 2021	Updated "Troubleshooting".
Jan. 2021	Updated LED Interface.
Oct. 2020	Updated board labels.
Jun. 2020	Added MCX512A-ADAT to relevant sections.
May. 2020	Added MCX515A-CCUT to relevant sections.
Feb. 2020	Added altitude criteria to "Specifications".
Oct. 2019	Updated the Ethernet supported protocols in "Specifications".

Date	Description of Changes
	Added the following OPNs to relevant sections:
Sep.	MCX516A-CCHT
2019	MCX512F-ACHT
	MCX512A-ACUT
Mar. 2019	Updated MCX512F-ACAT Specifications.
Dec. 2018	Migrated to online format; minor reorganization.
Nov. 2018	Added MCX512F-ACAT and MCX512A-ACAT support across the document
Jan. 2018	Added a note to System Requirements.
Nov. 2017	Updated Adapter Card LED Operations.
	Added MCX516A-BDAT support across the document and added the
	following:
	MCX516A-BDAT Specifications MCX513A-ACAT Board Label (Evaporate)
Sept. 2017	 MCX512A-ACAT Board Label (Example) Updated the following sections:
	Optiated the following sections: Oniver Installation: Linux
	MCX515A-GCAT Specification Table
	 MCX516A-GCAT Specification Table
	MCX515A-CCAT Specifications
	 MCX516A-CCAT Specifications
	 Airflow Specifications
Aug.	Added MCX515A-GCAT and MCX516A-GCAT support across the
2017	document and added the following:
	MCX516A-BDAT Specifications TableMCX516A-GCAT Specification Table
	 MCX516A-GCAT Specification Table Mechanical Drawing of the Single-port Adapter Cards
	 Single-port Tall Bracket
	MCX512A-ACAT Board Label (Example)
	MCX516A-GCAT Board Label (Example)

Date	Description of Changes			
	 Updated the following specifications tables: MCX515A-CCAT Specifications MCX516A-CCAT Specifications MCX516A-CDAT Specifications 			
July 2017	 Updated the following specifications tables: MCX515A-CCAT Specifications MCX516A-CCAT Specifications MCX516A-CDAT Specifications Updated Airflow Specifications 			
Apr. 2017	Updated Product OverviewUpdated Installation Instructions			
Mar. 2017	First release			

© Copyright 2023, NVIDIA. PDF Generated on 06/07/2024