



Mellanox Innova[™] IPsec Ethernet Adapter Card User Manual

Rev 1.8



NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT □PRODUCT(S)^N AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES AS-IS- WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER'S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCT(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



Mellanox Technologies 350 Oakmead Parkway Suite 100 Sunnyvale, CA 94085 U.S.A. www.mellanox.com Tel: (408) 9703400 Fax: (408) 9703403

© Copyright 2019. Mellanox Technologies Ltd. All Rights Reserved.

Mellanox®, Mellanox logo, Accelio®, BridgeX®, CloudX logo, CompustorX®, ConnectIB®, ConnectX®, CoolBox®, CORE-Direct®, EZchip®, EZchip logo, EZappliance®, EZdesign®, EZdriver®, EZsystem®, GPUDirect®, InfiniHost®, InfiniBridge®, InfiniScale®, Kotura®, Kotura logo, Mellanox CloudRack®, Mellano: CloudXMellanox®, Mellanox Federal Systems®, Mellanox HostDirect®, Mellanox Multiost®, Mellanox Open Ethernet®, Mellanox OpenCloud®, Mellanox OpenCloud Logo®, Mellanox PeerDirect®, Mellanox ScalableHPC® Mellanox StorageX®, Mellanox TuneX®, Mellanox Connect Accelerate Outperform logMellanox Virtual Modular Switch®, MetroX®, MLNX-OS®, NP-1c®, NP-2®, NP-3®, NPS®, Open Ethernet logo, PhyX®, PlatformX®, PSIPHY®, SiPhy®, StoreX®, SwitchX®, Tilera®, Tilera logo, TestX®, TuneX®, The Generation of Open Ethernet logo, UFM®, Unbreakable Link®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo art registered trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

For the most updated list of Mellanox trademarks, visit http://www.mellanox.com/page/trademarks



Table of Contents

Table of Contents				
List of Tables				
List of Figu	List of Figures			
Revision Hi	isto	ry	. 8	
Chapter 1	Intr	oduction	10	
:	1.1	Product Overview	11	
:	1.2	Features and Benefits	12	
:	1.3	Block Diagram	14	
:	1.4	Operating Systems/Distributions	14	
:	1.5	Connectivity	15	
:	1.6	Related Documents	15	
Chapter 2	Inte	erfaces	16	
	2.1	Ethernet QSFP Interface	16	
:	2.2	PCI Express Interface	16	
:	2.3	LED Interface	16	
Chapter 3	Har	dware Installation	17	
3	3.1	System Requirements	17	
		3.1.1 Hardware	17	
		3.1.2 Operating Systems/Distributions	17	
ŝ	3.2	Safety Precautions	17	
:	3.3	Pre-installation Checklist	17	
:	3.4	Bracket Installation Instructions	17	
		3.4.1 Removing the Existing Bracket	18	
		3.4.2 Installing the New Bracket	18	
	3.5	Card Installation Instructions.	18	
	3.6	Cables and Modules	19	
	~ -	3.6.1 Cable Installation	19	
	3./	Identify the Card In Your System.	20	
Chapter 4	Ne	llanox Innova IPsec Offload Overview	21	
4	4.1	Security Engines and IPsec Protocols	21	
		4.1.1 Ottloaded IPsec Protocols and Internet Protocols	22	
		4.1.1.1 IPSEC Protocols	22	
	4.2	IPsec Offload Kernel and Driver.	22	



		4.2.1 Mellanox Innova IPsec Ethernet Driver Module	22
		4.2.2 mlx5_fpga_tools Module	23
		4.2.3 Key Generation and Exchange	23
	4.3	IPsec Offload for DPDK Applications	24
Chapter 5	i Ps	ec Offload Software Installation and Operation	25
	5.1	Installation via MLNX_OFED	25
		5.1.1 Downloading Mellanox OFED	25
		5.1.2 Installing Mellanox OFED	25
		5.1.2.1 Installation Procedure	. 28
		5.1.2.2 Installation Results	. 30
		5.1.2.3 Installation Logging	. 30
		5.1.2.4 Driver Load Upon System Boot.	. 30
		5.1.2.5 mlnxofedinstall Return Codes	. 31
		5.1.3 Uninstalling Mellanox OFED	31
		5.1.4 Updating Firmware After Installation	31
		5.1.5 UEFI Secure Boot	32
		5.1.5.1 Enrolling Mellanox's x.509 Public Key On your Systems	. 32
		5.1.5.2 Removing Signature from Kernel Modules	. 33
	5.2	Installation of Kernel Module with IPsec Offload	34
		5.2.1 Obtaining the Kernel Modules	34
		5.2.2 Installing the Kernel and Driver	34
		5.2.3 Installing the Customized iproute2 Utility	35
	5.3	Operating the IPsec Offload.	37
		5.3.1 Loading/Unloading the Module	37
		5.3.1.1 Automatic Load	. 37
		5.3.1.2 Manual Load/Unload.	. 37
		5.3.2 Setting up an Offloaded IPsec Connection	37
		5.3.3 Destroying IPsec Tunnels	39
		5.3.4 IPsec Offload Statistics	39
Chapter 6	mb	د_fpga Tool	41
	6.1	Tool Requirements	41
	6.2	mlx_fpga Synopsis	41
	6.3	Examples of mlx_fpga Usage	42
		6.3.1 Adding FPGA mst Device Interface	42
		6.3.1.1 Burning the FPGA's Flash Device Using the mlx_fpga Burning Tool	. 42
		6.3.1.2 Loading Tool	. 42
		6.3.1.3 Debugging Tool	. 43
		6.3.1.4 Update FPGA Image.	. 43
Chapter 7	Up	dating Mellanox Innova IPsec Adapter Card Firmware	44
Chapter 8	C Tro	oubleshooting	46
	81	General	46



8.	.2 Li	nux	47
Chapter 9 S	peci	fications	48
9.	.1 N	INV101512A-BCIT Specifications.	48
9.	.2 N	INV101511A-BCIT Specifications.	49
9.	.3 N	Iellanox Innova IPsec EN LEDs Iellanox Innova IPsec EN LEDs	50
	9.	3.1 Network LEDs Operation	51
	9.	3.2 FPGA Debug LEDs	51
	9.	3.3 FPGA Load-Flow Debug LEDs	52
9.	.4 Bo	oard Mechanical Drawing and Dimensions	52
9.	.5 Bi	racket Mechanical Drawing	54
Appendix A	Fast	t Installation and Update	56
	A.1	Hardware Installation	56
	A.2	Content of Mellanox Innova IPsec Bundle	56
	A.3	Software, Firmware and Tools Installation	56
	A.4	Software, Firmware and Tools Update	58
	A.5	OFED Installation with Script	59
Appendix B	Finc	ding the MAC and Serial Number on the Adapter Card \ldots .	51
Appendix C	Safe	ety Warnings	6 2
Appendix D	Ave	ertissements de sécurité d'installation (Warnings in French)	64
Appendix E	Sich	erheitshinweise (Warnings in German)	6 6
Appendix F	Adv	ertencias de seguridad para la instalación (Warnings in Spa	nish) 68



List of Tables

Table 1:	Revision History Table
Table 2:	MNV101511A-BCIT Mellanox Innova IPsec Passive Cooling Adapter Card 11
Table 3:	MNV101512A-BCIT Mellanox Innova IPsec Active Cooling Adapter Card 11
Table 4:	Features
Table 5:	Documents List
Table 6:	mInxofedinstall Return Codes31
Table 7:	ethtool IPsec Offload Counters
Table 8:	MNV101512A-BCIT Specifications Table48
Table 9:	MNV101511A-BCIT Specifications Table49
Table 10:	Physical and Logical Link Indications51
Table 11:	FPGA Debug LEDs51
Table 12:	FPGA Load-Flow Debug LEDs52



List of Figures

Figure 1:	Mellanox Innova IPsec EN Adapter Card Block Diagram	14
Figure 2:	IPsec Solution Layers and Components	23
Figure 3:	MNV101511A-BCIT/MNV101512A-BCIT LEDs Placement (Example)	50
Figure 4:	Mechanical Drawing of MNV101511A-BCIT	52
Figure 5:	Mechanical Drawing of MNV101512A-BCIT	53
Figure 6:	Single-Port Tall Bracket	54
Figure 7:	Single-Port Short Bracket	55
Figure 8:	MNV101511A-BCIT Board Label	61
Figure 9:	MNV101512A-BCIT Board Label	61



Revision History

This document was printed on August 29, 2019.

Date	Rev	Comments/Changes
August 2019	1.8	 Updated Table 4, "Features," on page 12: removed Hardware-based I/O Virtualization and authentication algorithms. Updated Section 1.4, "Operating Systems/Distributions," on page 14. Updated Figure 3, "MNV101511A-BCIT/MNV101512A-BCIT LEDs Placement (Example)," on page 50. Mellanox Innova is a trademark of Mellanox Technologies Ltd.
February 2018	1.7	 Updated Section 5.2.2, "Installing the Kernel and Driver," on page 34 Updated Section 5.1.4, "Updating Firmware After Installation," on page 31
December 2017	1.6	 Updated Table 8, "MNV101512A-BCIT Specifications Table," on page 48 Updated Table 9, "MNV101511A-BCIT Specifications Table," on page 49
November 2017	1.5	 Added Table 3, "MNV101512A-BCIT Mellanox Innova IPsec Active Cooling Adapter Card," on page 11 Updated Section 5.1, "Installation via MLNX_OFED," on page 25 Updated Section 5.1, "Installation via MLNX_OFED," on page 25 Added Table 9, "MNV101511A-BCIT Specifications Table," on page 49 Added Figure 5, "Mechanical Drawing of MNV101512A-BCIT," on page 53 Updated Appendix A.3, "Software, Firmware and Tools Installation," on page 56 Updated Appendix A.4, "Software, Firmware and Tools Update," on page 58 Added Appendix A.5, "OFED Installation with Script," on page 59
June 2017	1.4	 Removed MCX4732A-BCIT from document Changed mlx_ipsec to mlx5_core Updated "memory" in Section 1.2, "Features and Benefits," on page 12 Updated Section 4.2, "IPsec Offload Kernel and Driver," on page 22 Updated Section 4.3, "IPsec Offload for DPDK Applications," on page 24 Updated Section 5.1, "Installation via MLNX_OFED," on page 25 Updated Section 5.3.1, "Loading/Unloading the Module," on page 37 Updated Section 6.2, "mlx_fpga Synopsis," on page 41 Updated Table 8, "MNV101512A-BCIT Specifications Table," on page 48 Added Appendix A.3, "Software, Firmware and Tools Installation," on page 56 Updated Appendix A.4, "Software, Firmware and Tools Update," on page 58



Table 1 - Revision History Table

Date	Rev	Comments/Changes
January 2017	1.3	 Updated Section 5.2.1, "Obtaining the Kernel Modules," on page 34 Updated Section 5.3.2, "Setting up an Offloaded IPsec Connection," on page 37 Updated Section 5.3.3, "Destroying IPsec Tunnels," on page 39 Updated Section 6.1, "Tool Requirements," on page 41 Updated Section 6.2, "mlx_fpga Synopsis," on page 41 Added Section 6.3.1.1, "Burning the FPGA's Flash Device Using the mlx_fpga Burning Tool," on page 42 Updated Section 6.3.1.2, "Loading Tool," on page 42 Updated Chapter 7, "Updating Mellanox Innova IPsec Adapter Card Firmware" on page 44 Added Figure 7, "Single-Port Short Bracket," on page 55 Added Appendix A, "Fast Installation and Update," on page 56
September 2016	1.2	 Added MNV101511A-BCIT across document: Section 1.1, "Product Overview," on page 11 Section 9.1, "MNV101512A-BCIT Specifications," on page 48 Figure 3, "MNV101511A-BCIT/MNV101512A-BCIT LEDs Placement (Example)," on page 50 Figure 4, "Mechanical Drawing of MNV101511A-BCIT," on page 52 Added Chapter 5, "IPsec Offload Software Installation and Operation" on page 25 Updated Section 5.2.2, "Installing the Kernel and Driver," on page 34 Updated Section 5.3.1, "Loading/Unloading the Module," on page 37 Updated Section 5.3.2, "Setting up an Offloaded IPsec Connection," on page 37 Added Section 5.3.3, "Destroying IPsec Tunnels," on page 39 Removed Innova IPsec 4 Lx EN Card Drivers Updated Section 8.1, "General," on page 41 Updated Section 9.1, "MNV101512A-BCIT Specifications," on page 48 Updated Appendix B, "Finding the MAC and Serial Number on the Adapter Card," on page 64
July 2016	1.1	 Changed mlx_accel_ipsec to mlx_ipsec. Added Section 4.2.2, "mlx5_fpga_tools Module," on page 23 Updated Section 5.2.1, "Obtaining the Kernel Modules," on page 34 Updated Section 5.2.2, "Installing the Kernel and Driver," on page 34 Updated Section 5.2.3, "Installing the Customized iproute2 Utility," on page 35 Added Section 5.3.4, "IPsec Offload Statistics," on page 39 Added "Update FPGA Image" on page 43 Updated Chapter 5.1,"Installation via MLNX_OFED" on page 25
April 2016	1.0	First Release



1 Introduction

This is the User Guide for Mellanox Innova[™] IPsec adapter card based on the ConnectX®-4 Lx Ethernet (EN) integrated circuit device with an on-board FPGA device.

The Mellanox Innova IPsec EN adapter card provides security acceleration for IPsec-enabled networks while taking advantage of the ConnectX-4 Lx EN network controller's best-in-class performance, unmatched scalability, and efficiency.

The constantly growing demand for security and privacy in modern data centers, private and public clouds, Web 2.0 infrastructure, and telecommunication systems, requires the use of security protocols. IPsec is a protocol suite for secure Internet Protocol (IP) communications by authenticating and encrypting each IP packet of a communication session. However, the high computing power required by the IPsec algorithms consumes expensive CPU cycles and limits network connection performance.

The Mellanox Innova IPsec EN adapter offloads the processing of the IPsec algorithms, frees up the CPU, and eases network bottlenecks.

The adapter integrates advanced network capabilities and encryption offloading in one card, utilizing only a single PCIe slot for both networking and crypto functions.

The Mellanox Innova IPsec EN adapter also brings ConnectX-4 Lx's industry leading technologies: hardware support for RDMA over Converged Ethernet, Ethernet stateless offload engines, overlay networks, GPUDirect® technologies, and more.

This chapter covers the following topics:

- Section 1.1, "Product Overview," on page 11
- Section 1.2, "Features and Benefits," on page 12
- Section 1.3, "Block Diagram," on page 14
- Section 1.4, "Operating Systems/Distributions," on page 14
- Section 1.5, "Connectivity," on page 15
- Section 1.6, "Related Documents," on page 15



1.1 Product Overview

This section provides the ordering part number, port speed, number of ports, and PCI Express speed for the various models.

Ordering Part Number (OPN)	MNV101511A-BCIT - HHHL card with Xilinx Kintex® UltraScale™ XCKU060
Data Transmission Rate	Ethernet: 10/40Gb/s
Network Connector Types	Single-port QSFP
PCI Express (PCIe) SerDes Speed	PCIe 3.0 x8 8GT/s
RoHS	R6
Adapter IC Part Number	MT27711A0-FDCF-BE
Device ID (decimal)	610
Cooling Method	Passive Cooling

Table 3 - MNV101512A-BCIT Mellanox Innova IPsec Active Cooling Adapter Card

Ordering Part Number (OPN)	MNV101512A-BCIT - HHHL card with Xilinx Kintex® UltraScale™ XCKU060
Data Transmission Rate	Ethernet: 10/40Gb/s
Network Connector Types	Single-port QSFP
PCI Express (PCIe) SerDes Speed	PCIe 3.0 x8 8GT/s
RoHS	R6
Adapter IC Part Number	MT27711A0-FDCF-BE
Device ID (decimal)	610
Cooling Method	Active Cooling



1.2 Features and Benefits

Table 4 - Features^a

PCI Express (PCIe)	Uses PCIe Gen 3.0 (1.1 and 2.0 compatible) through an x8 edge connector up to 8GT/s
10/40 Gigabit Ethernet	Mellanox adapters comply with the following IEEE 802.3* standards: – IEEE Std 802.3ba 40 Gigabit Ethernet – IEEE Std 802.3ae 10 Gigabit Ethernet – IEEE Std 802.3ad, Link Aggregation – IEEE Std 802.1Q, 1P VLAN tags and priority – IEEE Std 802.1Qau Congestion Notification – IEEE Std 802.1Qbg – IEEE P802.1Qaz D0.2 ETS – IEEE P802.1Qbb D1.0 Priority-based Flow Control
Memory	 PCI Express - stores and accesses Ethernet fabric connection information and packet data SPI - includes two SPI Flash devices: one 16MB SPI Flash device (W25Q128FVSIG by WINBOND-NUVOTON) for ConnectX-4 Lx device one 512Mb SPI flash (MT25QL512ABB8E12 by MICRON TECHNOL-OGY) for the FPGA device EEPROM - accessible through the I2C-compatible interface. The EEPROM capacity is 128Kb. 2GByte DDR4 - PC-1600MT/sec Soldered on board
IPsec Offload	The Mellanox Innova IPsec adapter provides offloading of compute intensive encryption/decryption, which are used by the IPsec protocol. Support for Linux and Windows IPsec software interfaces ensures native integration with existing IPsec applications, with no required changes to the user's software. IPsec offloading is handled by the combination of the ConnectX-4 Lx network controller and an on-board FPGA, providing high performance and flexibility for future enhancements and customizations. The FPGA is connected to the ConnectX-4 Lx through a 'bump-in-the-wire' topology, hence encryption and decryption are performed inline with the network flow. This results in lower latency and additional savings of CPU resources compared to other IPsec protocol solutions, be it through software or alternative accelerators.
Overlay Networks	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. The Mellanox Innova IPsec adapter effectively addresses this by providing advanced NVGRE, VXLAN and GENEVE hardware offloading engines that encapsulate and de-capsulate the overlay protocol headers, enabling the traditional offloads to be performed on the encapsulated traffic for these and other tunneling protocols (GENEVE, MPLS, QinQ, and so on). With the Mellanox Innova IPsec adapter, data center operators can achieve native performance in the new network architecture.



Tabl	e 4	- F	eat	ure	sa
------	-----	-----	-----	-----	----

RDMA and RDMA over Converged Ethernet (RoCE)	Mellanox Innova IPsec adapter supports RoCE specifications delivering low- latency and high- performance over Ethernet networks. Leveraging data center bridging (DCB) capabilities as well as Mellanox Innova IPsec adapter advanced congestion control hardware mechanisms, RoCE provides efficient low-latency RDMA services over Layer 2 and Layer 3 networks.
Mellanox PeerDirect®	Mellanox 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. Mellanox Innova IPsec adapter 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.
Quality of Service (QoS)	Support for port-based Quality of Service enabling various application requirements for latency and SLA
Storage Acceleration	A consolidated compute and storage network achieves significant cost- performance advantages over multi-fabric networks. Standard block and file access protocols can leverage RDMA for high-performance storage access.
Distributed RAID	Mellanox Innova IPsec adapter delivers advanced Erasure Coding offloading capability, enabling distributed Redundant Array of Inexpensive Disks (RAID), a data storage technology that combines multiple disk drive components into a logical unit for the purposes of data redundancy and performance improvement. Mellanox Innova IPsec adapter's Reed-Solomon capability introduces redundant block calculations, which, together with RDMA, achieves high performance and reliable storage access.

a. This section describes hardware features and capabilities. Please refer to the driver release notes for feature availability. See <u>Section 1.6, "Related Documents," on page 15</u>.



1.3 Block Diagram



Figure 1: Mellanox Innova IPsec EN Adapter Card Block Diagram

1.4 Operating Systems/Distributions¹

• RHEL/CentOS

^{1.} Please refer to the driver release notes for feature availability.

Introduction



1.5 Connectivity

- Interoperable with 10Gb and 40Gb Ethernet switches
- Passive copper cable with ESD protection
- Powered connectors for optical and active cable support

1.6 Related Documents

Table 5 - Documents List

Document's Name	Location
<i>Mellanox Firmware Tools (MFT) User Manual</i> Document no. 2204UG	User Manual describing the set of MFT firmware management tools for a single node. See <u>http://www.mellanox.com/page/management_tools</u>
<i>Mellanox OFED for Linux User Manual</i> Document no. 2877	User Manual describing OFED features, performance, tools content and configuration. See <u>http://www.mellanox.com</u> => Products => Software => InfiniBand/VPI Drivers => Mellanox OpenFabrics Enterprise Distribution for Linux (MLNX_OFED)
Performance Tuning Guidelines for Mellanox Network Adapters Document no. 3368	User Manual describes important tuning parameters and settings that can improve performance for Mellanox drivers. See <u>http://www.mellanox.com/related-docs/prod_software/</u> <u>Performance_Tuning_Guide_for_Mellanox_Network_Adapters</u> .pdf
MellanoxO FED for Linux Driver Release Notes	Release notes for Mellanox Technologies' MLNX_OFED for Linux driver kit for Mellanox adapter cards: <u>http://www.mellanox.com</u> => Products => Software => InfiniBand/VPI Drivers => Mellanox OpenFabrics Enterprise Distribution for Linux (MLNX_OFED) => Release Notes
IEEE Std 802.3 Specification	This is the IEEE Ethernet specification http://standards.ieee.org/getieee802
PCI Express 3.0 Specifications	Industry Standard PCI Express 3.0 Base and Card Electromechanical Specifications. https://pcisig.com/specifications
IETF IPsec Specifications	https://tools.ietf.org/html/rfc4301



2 Interfaces

Each adapter card includes the following interfaces:

- "Ethernet QSFP Interface"
- "PCI Express Interface"
- "LED Interface"

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

2.1 Ethernet QSFP Interface

The network port of the Mellanox Innova IPsec adapter is compliant with the IEEE 802.3 Ethernet standards listed in Table 4, "Features," on page 12. For connecting to an SFP+ interface, you can use a Mellanox QSA (QSFP to SFP+) adapter module.

2.2 PCI Express Interface

The Mellanox Innova IPsec adapter card supports PCI Express 3.0 (1.1 and 2.0 compatible) through an x8 edge connector. The device can be either a master initiating the PCI Express bus operations or a slave responding to PCI bus operations. The following lists PCIe interface features of the Mellanox Innova IPsec adapter card:

- PCIe Gen 3.0 compliant, 1.1 and 2.0 compatible
- 2.5, 5.0, or 8.0GT/s link rate x8
- Auto-negotiates to x8, x4, x2, or x1
- Support for MSI/MSI-X mechanisms

2.3 LED Interface

For Mellanox Innova IPsec adapter card LED specifications, please refer to Section 9.3, "Mellanox Innova IPsec EN LEDs," on page 50.



3 Hardware Installation

3.1 System Requirements

3.1.1 Hardware

A system with a PCI Express x8 slot is required for installing the card.

3.1.2 Operating Systems/Distributions

Please refer to Section 1.4, "Operating Systems/Distributions," on page 14.

3.2 Safety Precautions



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

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

3.3 Pre-installation Checklist

- 1. Verify that your system meets the hardware and software requirements stated above.
- 2. Shut down your system if active.
- 3. After shutting down the system, turn off power and unplug the cord.
- 4. Remove the card from its package. Please note that the card must be placed on an antistatic surface.
- 5. Check the card for visible signs of damage. Do not attempt to install the card if damaged.

3.4 Bracket Installation Instructions

The card is usually shipped with a tall bracket installed. If this form factor is suitable for your requirements, you can skip the remainder of this section and move to Section 3.5, "Card Installation Instructions," on page 18.

If you need to replace it with the short bracket that is included in the shipping box, please follow the instructions in this section.





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 bracket of the proper height
- The 2 screws saved from the removal of the bracket
- The 2 fiber washers saved from the removal of the bracket

3.4.1 Removing the Existing Bracket

1. Remove the two screws holding the bracket in place. The bracket comes loose from the card.



Be careful not to put stress on the LED.

2. Save the two screws and the two fiber washers.

3.4.2 Installing the New Bracket

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



Do not force the bracket onto the card. You may have to gently push the LEDs using a small screwdriver to align the LEDs with the holes in the bracket.

- 2. Screw on the bracket using the screws and washers saved from the bracket removal procedure above.
- 3. Make sure that the LEDs are aligned onto the bracket holes.
- 4. Use a torque driver to apply up to 2.9 lbs-in torque on the screws.

3.5 Card Installation Instructions

- 1. Open the system case.
- 2. Place the adapter in a standard PCI Express slot.
- 3. Applying even pressure at both corners of the card, insert the adapter card into the slot until it is firmly seated. When the adapter is properly seated, the adapter port connectors are aligned with the slot opening, and the adapter faceplate is visible against the system chassis.





Do not use excessive force when seating the card, as this may damage the system or the adapter.

3.6 Cables and Modules

To obtain the list of supported cables for your adapter, please refer to <u>http://www.mellanox.com/</u> products/interconnect/cables-configurator.php.

3.6.1 Cable Installation

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



When installing cables make sure that the latches engage.



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

- 3. After inserting a cable into a port, the Amber 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 Section 9.3, "Mellanox Innova IPsec EN LEDs," on page 50.
- 4. After plugging in a cable, lock the connector using the latching mechanism particular to the cable vendor. When a logical connection is made, the Green LED will light. When data is being transferred the Green LED will blink. See Section 9.3, "Mellanox Innova IPsec EN LEDs," on page 50.



- 5. Care should be taken as not to impede the air exhaust flow through the ventilation holes. Use cable lengths which allow for routing horizontally around to the side of the chassis before bending upward or downward in the rack.
- 6. To remove a cable, disengage the locks and slowly pull the connector away from the port receptacle. LED indicator will turn off when the cable is unseated.

3.7 Identify the Card in Your System

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

lspci |grep -i Mellanox Network controller: Mellanox Technologies MT27710 Family [ConnectX-4 Lx]



4 Mellanox Innova IPsec Offload Overview

The Mellanox Innova IPsec EN adapter is pre-programmed with a Mellanox IPsec offload FPGA logic, offering encryption, decryption and authentication for IPsec security protocol suite.

The IPsec offload solution offers three major benefits:

- 1. Offloads compute intensive crypto algorithms from the host CPU, thus freeing up the CPU and easing network bottlenecks.
- 2. Since the crypto process occurs on the FPGA, which acts as a 'bump-in-the-wire', the traffic reaches the ConnectX-4 Lx plain so that the various ConnectX-4 Lx networking and stateless offloads can be applied to that traffic.
- 3. The existing IPsec implementation in Linux kernel requires the network stack to process the packet before and after the crypto processing of the packet. As 'bump-in-the-wire', Mellanox Innova IPsec prevents traffic from undergoing the kernel network stacks process more than once.

With these benefits, IPsec offload allows the adapter to reach full wire speed with IPsec secured traffic on the wire while reducing CPU utilization.

IPsec offload is supported in two modes - kernel mode (Section 4.2, "IPsec Offload Kernel and Driver," on page 22) and DPDK mode (Section 4.3, "IPsec Offload for DPDK Applications," on page 24).

4.1 Security Engines and IPsec Protocols



For list of supported crypto algorithms please refer to *Mellanox Innova IPsec EN Release Notes*. Additional crypto algorithms can be added based on business needs.

The crypto algorithms in the Mellanox Innova IPsec adapter is a symmetric encryption and authentication using either the AES-GCM mechanism (described in GCM-Spec), the encryption of AES-CBC (described in CBC-Spec) and/or the authentication by:

- HMAC-SHA-1
- HMAC-SHA2 (256, 384, 512)

Please refer to HMAC-Spec and SHA-Spec for further details.

The crypto engines are designed to deliver full wire speed operation in a wire rate of 40G. These crypto engines are integrated with IPsec-ESP protocol mechanism which is elaborated in rfc4106, or with IPsec-AH, as described in rfc4302.



4.1.1 Offloaded IPsec Protocols and Internet Protocols

This section lists IPsec protocols and Internet Protocols that can be offloaded to the Mellanox Innova IPsec adapter.



For list of supported protocols, please refer to Mellanox Innova IPsec EN Release Notes.

4.1.1.1 IPsec Protocols

• ESP modes - Tunnel mode, Transport mode

4.1.1.2 Internet Protocols

- IPv4
- IPv6

4.2 IPsec Offload Kernel and Driver

In order to install the kernel and driver, please refer to Chapter 5,"IPsec Offload Software Installation and Operation" on page 25.

The Mellanox Innova IPsec offload solution is integrated into the latest IPsec framework in the Linux kernel, IP-XFRM framework, using the IP-XFRM offload API provided by the kernel.

The IP-XFRM framework is exposed to the user through various software implementations for IPsec connection creation and management (such as iproute2, libreswan, strongswan and others).

Upon setting up an IPsec connection, the user can choose whether to enable the Mellanox Innova IPsec offload on the specific IPsec security association (SA) that is created once the connection is generated. See Section 5.3.2, "Setting up an Offloaded IPsec Connection," on page 37. Security associations that are not set to be offloaded will still undergo encryption/decryption operations by the Linux kernel.

4.2.1 Mellanox Innova IPsec Ethernet Driver Module

The Mellanox Innova IPsec adapter has a dedicated driver in the form of a kernel module, mlx5_core.ko. The driver performs the following:

- Configures the offload settings and modes in HW.
- Manages the offloaded security associations database in HW and ensures its validity.
- Ensures and maintains the flow of packets from kernel network stack to the Mellanox Innova IPsec adapter for offloading of encryption and from the Mellanox Innova IPsec adapter to kernel network stack after decryption offloading.

Figure 2 illustrates the IPsec solution layers and components.





Figure 2: IPsec Solution Layers and Components

4.2.2 mlx5_fpga_tools Module

mlx5_fpga_tools module is included in the new kernel installation.

The module allows opening and configuring the character device to be used by the dedicated mlx_fpga tool for various purposes. Please refer to Chapter 6, "mlx_fpga Tool" on page 41.

The module is not loaded by default and not required for IPsec offload.

To load it run:

modprobe mlx5 fpga tools

The module depends on the mlx5_core module.

4.2.3 Key Generation and Exchange

The Mellanox Innova IPsec adapter currently supports offloading of the encryption, decryption and authentication of IPsec traffic. The key generation and exchange protocol, whether done manually or through IKE protocol, remains within complete ownership of the userspace software that is used for IPsec connection creation and management (such as iproute2, libreswan, strongswan and others) and is not affected by the HW or the supplied IPsec kernel module.

The Mellanox IPsec kernel module will only be invoked by the kernel offload API once the key and SPI values are determined (whether manually or by IKE) and crypto offload is enabled. The module will update the security association database on the FPGA/DDR so that crypto offload can occur while traffic is running.



4.3 IPsec Offload for DPDK Applications

mlx5_core module offers offload support for raw Ethernet and kernel bypass drivers by exposing a user interface to control the offloaded security associations in the FPGA.

Mellanox provides a DPDK Poll Mode Driver (PMD) which makes use of this interface. PMD provides a new API for DPDK applications to open/close offloaded security associations (control path) while transmitting/receiving traffic through them (data path). The data path is still done with kernel network stack bypass, providing the application with the benefits of both DPDK acceleration and security offload (encryption/decryption).

Please refer to Mellanox Innova IPsec EN Release Notes for supported versions.





5 IPsec Offload Software Installation and Operation

5.1 Installation via MLNX_OFED



OFED installation on top of the kernel module that is provided in this bundle is not supported.

As of version 4.2, MLNX_OFED supports Mellanox Innova IPsec EN adapter card. This type of installation is applicable to RedHat 7.1, 7.2, 7.3 and 7.4 operating systems and Kernel 4.13.

Prerequisites

In order to obtain Mellanox Innova IPsec offload capabilities once MLNX_OFED is installed, make sure kernel v4.13 or newer is installed with the following configuration flags enabled:

- CONFIG_XFRM_OFFLOAD
- CONFIG_INET_ESP_OFFLOAD
- CONFIG_INET6_ESP_OFFLOAD

5.1.1 Downloading Mellanox OFED

Step 1. Verify that the system has a Mellanox network adapter (HCA/NIC) installed.

The following example shows a system with an installed Mellanox HCA:

lspci -v | grep Mellanox 06:00.0 Network controller: Mellanox Technologies MT27710 Family [ConnectX-4] Subsystem: Mellanox Technologies Device 0024

Step 2. Download the ISO image to your host.

The image's name has the format MLNX_OFED_LINUX-<ver>-<OS label><CPU arch>.iso. An ISO image for the Mellanox Innova Flex adapter can be obtained through Mellanox support.

Step 3. Use the md5sum 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.

host1\$ md5sum MLNX OFED LINUX-<ver>-<OS label>.iso

5.1.2 Installing Mellanox OFED

The installation script, mlnxofedinstall, performs the following:

- Discovers the currently installed kernel
- Uninstalls any software stacks that are part of the standard operating system distribution or another vendor's commercial stack
- Installs the MLNX_OFED_LINUX binary RPMs (if they are available for the current kernel)



• Identifies the currently installed Ethernet network adapters and automatically¹ upgrades the firmware

Usage

./mnt/mlnxofedinstall [OPTIONS]

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



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

- If you need to install Mellanox 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 (such as pdsh).
- 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 under the docs/ directory.



On Redhat 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.

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



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

 $Therefore, using \verb"mlnx_add_kernel_support.sh" is irrelevant on Ubuntu distributions.$

Usage:

```
mlnx_add_kernel_support.sh -m|--mlnx_ofed <path to MLNX_OFED directory> [--make-iso|--
make-tgz]
[--make-iso] Create MLNX_OFED ISO image.
[--make-tgz] Create MLNX_OFED tarball. (Default)
[-t|--tmpdir <local work dir>]
[--kmp] Enable KMP format if supported.
[-k | --kernel] <kernel version> Kernel version to use.
[-s | --kernel-sources] <path to the kernel sources> Path to kernel headers.
[-v|--verbose]
[-n|--name] Name of the package to be created.
[-y|--yes] Answer "yes" to all questions
[--force] Force removing packages that depends on MLNX_OFED
```

1. The firmware will not be updated if you run the install script with the '--without-fw-update' option.



Example

The following command will create a MLNX_OFED_LINUX ISO image for RedHat 7.2 under the /tmp directory.

```
# ./MLNX_OFED_LINUX-x.x-x-rhel7.2-x86_64/mlnx_add_kernel_support.sh -m /tmp/
MLNX_OFED_LINUX-x.x-x-rhel7.1-x86_64/ --make-tgz
Note: This program will create MLNX_OFED_LINUX TGZ for rhel7.2 under /tmp directory.
All Mellanox, OEM, OFED, or Distribution 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-rhel7.1-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 user space application to unlimited. If desired, tune the value unlimited to a specific amount of RAM.

-c config <packages config_file=""></packages>	Example of the configuration file can be found under docs
-n net <network config_file=""></network>	Example of the network configuration file can be found under docs
-k kernel-version <kernel version></kernel 	Use provided kernel version instead of 'uname -r'
-p print-available	Print available packages for current platform and create corresponding ofed.conf file
with-32bit	Install 32-bit libraries
without-32bit	Skip 32-bit libraries installation (Default)
without-depcheck	Skip Distro's libraries check
without-fw-update	Skip firmware update
fw-update-only	Update firmware. Skip driver installation
force-fw-update	Force firmware update
force	Force installation
all hpc basic msm	Install all, hpc, basic or Mellanox Subnet manager packages correspondingly
vma vma-vpi	Install packages required by VMA to support VPI
vma-eth	Install packages required by VMA to work over Ethernet
with-vma	Set configuration for VMA use (to be used with any installation parameter).
guest	Install packages required by guest os



hypervisor	Install packages required by hypervisor os
-v -vv -vvv	Set verbosity level
umad-dev-rw	Grant non root users read/write permission for umad devices instead of default
umad-dev-na	Prevent from non root users read/write access for umad devices. Overrides 'umad-dev-rw'
enable-affinity	Run mlnx_affinity script upon boot
disable-affinity	Disable mlnx_affinity script (Default)
enable-sriov	Burn SR-IOV enabled firmware - Note: Enable/Disable of SRIOV in a non-volatile configuration through uEFI and/ or tool will override this flag.
add-kernel-support	Add kernel support (Run mlnx_add_kernel_support.sh)
skip-distro-check	Do not check MLNX_OFED vs Distro matching
hugepages-overcommit	Setting 80% of MAX_MEMORY as overcommit for huge page allocation
-q	Set quiet - no messages will be printed
without- <package></package>	Do not install package
with-fabric-collector	Install fabric-collector package.

5.1.2.1 Installation Procedure

- Step 1. Login to the installation machine as root.
- Step 2. Mount the ISO image on your machine.

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

Step 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, or Distribution IB packages will be removed.
Uninstalling the previous version of MLNX_OFED_LINUX
Starting MLNX_OFED_LINUX-x.x.x installation ...
......
Installation finished successfully.
Attempting to perform Firmware update...
Querying Mellanox devices firmware ...
```





In case your machine has an unsupported network adapter device, no firmware update will occur and the error message below will be printed. Please contact your hardware vendor for help on firmware updates.

```
Error message:
Device #1:
```

vice #1:	
Device:	0000:05:00.0
Part Number:	
Description:	
PSID:	MT_2410110034MT_2490110032
Versions:	Current Available
FW	14.12.1000 N/A
Status:	No matching image found

Step 4. Reboot the machine if the installation script performed firmware updates to your network adapter hardware. Otherwise, restart the driver by running: "/etc/init.d/ openibd restart"

After the installer completes, information about the Mellanox OFED installation such as prefix, kernel version, and installation parameters can be retrieved by running the command /etc/ infiniband/info.

Most of the Mellanox 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.



5.1.2.2 Installation Results

Software	 Most of MLNX_OFED packages are installed under the "/usr" directory except for the following packages which are installed under the "/opt" directory: openshmem, bupc, fca and ibutils The kernel modules are installed under /lib/modules/`uname -r`/extra/mlnx-ofa_kernel on RHEL and other RedHat like Distributions /lib/modules/`uname -r`/updates/dkms/ on Ubuntu
Firmware	 The firmware of existing network adapter devices will be updated if the following two conditions are fulfilled: The installation script is run in default mode; that is, without the option 'without-fw-update' The firmware version of the adapter device is older than the firmware version included with the Mellanox 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. I Querying device E Can't auto detect fw configuration file:

5.1.2.3 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: "Logs dir: /tmp/MLNX_OFED_LINUX-<version>.<PID>.logs".

Example:

Logs dir: /tmp/MLNX_OFED_LINUX-x.x-x.logs

5.1.2.4 Driver Load Upon System Boot

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

> To prevent automatic load of the Mellanox drivers upon system boot:

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

blacklist	mlx4	core
blacklist	mlx4	en
blacklist	mlx5	core



5.1.2.5 mlnxofedinstall Return Codes

Table 6 lists the mlnxofedinstall script return codes and their meanings.

Table 6 - mlnxofedinstall Return Codes

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

5.1.3 Uninstalling Mellanox OFED

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

5.1.4 Updating Firmware After Installation

The firmware and FPGA update package (mlnx-fw-updater) is installed under "/opt/mellanox/ mlnx-fw-updater" folder. The latest FW and FPGA update package can be downloaded from <u>www.mellanox.com</u> => Products => Ethernet => SmartNICs => Innova IPsec SmartNIC Ethernet => Download tab.

You can run the following update script using one of the modes below:

/opt/mellanox/mlnx-fw-updater/mlnx_fpga_updater.sh

1. With -u flag to provide URL to the software package (tarball). Example:

./mlnx_fpga_updater.sh -u http://www.mellanox.com/downloads/fpga/ipsec/ Innova_IPsec_<version>.tgz

2. With -t flag to provide the path to the downloaded tarball. Example:

./mlnx_fpga_updater.sh -t <Innova_IPsec_bundle_file.tgz>

3. With -p flag to provide the path to the downloaded and extracted tarball. Example:

./mlnx_fpga_updater.sh -p <Innova_IPsec_extracted_bundle_directory>



For more information on the script usage, you can run mlnx_fpga_updater.sh -h.



Its is recommended to perform firmware and FPGA upgrade on Mellanox Innova IPsec cards using this script only.

5.1.5 UEFI Secure Boot

All kernel modules included in MLNX_OFED for RHEL7 are signed with x.509 key to support loading the modules when Secure Boot is enabled.

5.1.5.1 Enrolling Mellanox's x.509 Public Key On your Systems

In order to support loading MLNX_OFED drivers when an OS supporting Secure Boot boots on a UEFI-based system with Secure Boot enabled, the Mellanox x.509 public key should be added to the UEFI Secure Boot key database and loaded onto the system key ring by the kernel.

Follow these steps below to add the Mellanox's x.509 public key to your system:



- Prior to adding the Mellanox's x.509 public key to your system, please make sure:
- the 'mokutil' package is installed on your system
- the system is booted in UEFI mode

Step 1. Download the x.509 public key.

wget http://www.mellanox.com/downloads/ofed/mlnx_signing_key_pub.der

Step 2. Add the public key to the MOK list using the mokutil utility.

You will be asked to enter and confirm a password for this MOK enrollment request.

mokutil --import mlnx signing key pub.der

Step 3. Reboot the system.

The pending MOK key enrollment request will be noticed by shim.efi and it will launch MokManager.efi to allow you to complete the enrollment from the UEFI console. You will need to enter the password you previously associated with this request and confirm the enrollment. Once done, the public key is added to the MOK list, which is persistent. Once a key is in the MOK list, it will be automatically propagated to the system key ring and subsequent will be booted when the UEFI Secure Boot is enabled.



To see what keys have been added to the system key ring on the current boot, install the 'keyutils' package and run: #keyctl list %:.system_keyring



5.1.5.2 Removing Signature from Kernel Modules

The signature can be removed from a signed kernel module using the 'strip' utility which is provided by the 'binutils' package.

strip -g my module.ko

The strip utility will change the given file without saving a backup. The operation can be undo only by resigning the kernel module. Hence, we recommend backing up a copy prior to removing the signature.

> To remove the signature from the MLNX_OFED kernel modules:

Step 1. Remove the signature.

rpm -qa | grep -E "kernel-ib|mlnx-ofa_kernel|iser|srp|knem" | xargs rpm -ql | grep
"\.ko\$" | xargs strip -g

After the signature has been removed, a massage as the below will no longer be presented upon module loading:

"Request for unknown module key 'Mellanox Technologies signing key: 61feb074fc7292f958419386ffdd9d5ca999e403' err -11"

However, please note that a similar message as the following will still be presented:

"my_module: module verification failed: signature and/or required key missing - tainting kernel"

This message is presented once, only for each boot for the first module that either has no signature or whose key is not in the kernel key ring. So it's much easier to miss this message. You won't see it on repeated tests where you unload and reload a kernel module until you reboot. There is no way to eliminate this message.

Step 2. Update the initramfs on RHEL systems with the stripped modules.

mkinitrd /boot/initramfs-\$(uname -r).img \$(uname -r) --force



5.2 Installation of Kernel Module with IPsec Offload



OFED installation on top of the kernel module that is provided in this bundle is not supported.

5.2.1 Obtaining the Kernel Modules

The kernel modules described in Section 4.2, "IPsec Offload Kernel and Driver," on page 22 are a part of a special Linux kernel installation bundle provided by Mellanox. The bundle includes the latest kernel installation files and other related components:

- FPGA image bin file
- Kernel RPM files
- MFT tarball file
- Firmware bin files
- Example IPsec offload scripts

To download the bundle, please refer to <u>www.mellanox.com</u> => Ethernet => SmartNICs => Innova IPsec SmartNIC Ethernet => Download tab.

5.2.2 Installing the Kernel and Driver



Please make sure that the latest FW, FPGA image and MFT versions are installed. Please refer to the *Mellanox Innova IPsec EN Adapter Card Release Notes* for the latest versions.

Once you have obtained the kernel RPM file, the file can be installed by performing the following steps:

- 1. Run: rpm -i kernel-<kernel version>.rpm / rpm -i kernel-devel-<kernel version>.rpm
- 2. Verify that the initial RAM disk image has been created:
 - a. Run ls /boot/ and look for the relevant initramfs and vmlinuz files that match the kernel version you just installed (names should match the RPM name).
- 3. Please verify that the new kernel is added to the bootloader configuration file. If the new kernel was not added automatically, you can add it manually:
 - a. For example, open the /boot/grub/grub.conf file for editing (the boot menu configuration file) and add a new menu entry for the new installed kernel. Example of menu entry to be added (replace the vmlinuz and init-ramfs names with the new kernel file names and modify the entry title as desired):

```
title upstream-4.7 rc5 for FPGA
root (hd0,0)
kernel /vmlinuz-4.7.0-rc5+ root=/dev/sda2 console=tty0 console=ttyS0,115200n8 rhgb
initrd /initramfs-4.7.0-rc5+.img
```



b. If using grub2, open the /boot/grub2/grub.cfg for editing and add a new menu entry for the new installed kernel. Example of menu entry to be added (replace the vmlinuz and initramfs names with the new kernel file names and modify the entry title as desired):

```
menuentry 'Upstream 4.12.0-rc4+' --class rhel fedora --class qnu-linux --class qnu --
class os --unrestricted $menuentry id option 'gnulinux-4.12.0-rc4+-advanced-2d912b91-
d2e5-44fd-8040' {
       load video
        set gfxpayload=keep
       insmod gzio
        insmod part msdos
        insmod xfs
        set root='hd0,msdos1'
        if [ x$feature platform search hint = xy ]; then
         search --no-floppy --fs-uuid --set=root --hint='hd0,msdos1' f663c5e3-1cbb-
4f88-b65e-f9848f1458c9
       else
         search --no-floppy --fs-uuid --set=root f663c5e3-1cbb-4f88-b65e-f9848f1458c9
        fi
        linux16 /vmlinuz-4.12.0-rc4+ root=/dev/sda2 console=tty0 control=ttyS0,115200n8
rhqb
        initrd16 /initramfs-4.12.0-rc4+.img
```

- 4. Once the kernel installation is complete, reboot your system and select the relevant kernel to load from the grub menu.
- 5. Optional It is also possible to change the default entry value to the index of the new entry (the indexes are zero-based) so that the new kernel will be loaded by default. This is done by changing the index value next to the word "default" at the beginning of the grub.conf file. For grub2, you can use the command: *grub2-set-default*.
- **Note:** To confirm that the required kernel version is loaded, use the "uname -r" command. The output indicates the kernel version and name.
- **Note:** Installing the kernel modules will also install the following Mellanox device driver modules mlx5_core and mlx5_ib.

5.2.3 Installing the Customized iproute2 Utility

iproute2 is a user space utilities package that controls TCP/IP networking configuration in the kernel. It includes commands such as:

• *ip*: for management of network tables and network interfaces. It is also used to configure packet transformation policies and the security associations (SAs) attached to those policies. *ip* utility is used to set up IPsec policies on security associations.

Mellanox provides a customized iproute2 utility set which exposes new flags in the ip xfrm utility to allow the user control of the IPsec tunnel offload state. These flags provide the option to enable offload for IPsec SAs.

1. Obtain the customized iproute2 RPM file by contacting Mellanox support (File Name: iproute2-<version>.x86_64.rpm)



2. Install the utility using the following command: rpm -i --force iproute2-<version>.x86_64.rpm

Once the installation is complete, you will have the modified iproute2 utility that supports the IPsec offload flags installed in your system.

Note: There are several additional user space applications that provide an interface to configure IPsec policies and SAs: Strongswan (which has IPsec offload support as of version 5.5.3), Libreswan (which has IPsec offload support as of version 3.21) and more. Please refer to the release notes of the above mentioned user space applications for more information on IPsec offload support.



5.3 Operating the IPsec Offload

5.3.1 Loading/Unloading the Module

5.3.1.1 Automatic Load

The Mellanox Innova IPsec Ethernet driver, mlx5_core, is loaded automatically by the kernel when a Mellanox Innova IPsec card is installed.

5.3.1.2 Manual Load/Unload

1. Load/unload mlx5_core using one of the following commands:

insmod mlx5_core
modprobe mlx5_core
rmmod mlx5 core(unload command)

Note: Unloading the IPsec offload module while there are active IPsec offloaded connections is not supported and the result is undefined. For proper and stable operation of the HW and SW, the offloaded IPsec connection must be terminated via the proper utility before module unload. It is recommended to flush the existing IPsec XFRM states before restarting the mlx5_core module in case there are offloaded security associations. It can be done by running the following command: ip xfrm state flush; ip xfrm pol flush.

5.3.2 Setting up an Offloaded IPsec Connection

IPsec secured connection can be opened through the iproute2 utility. For offload support, please use the iproute2 version that is modified and supplied by Mellanox (see Section 5.2.3, "Installing the Customized iproute2 Utility," on page 35).

In order to configure an IPsec secured connection between hosts, it is necessary to:

- 1. Configure the security association (SA) intended for use, with its relevant parameters (such as: crypto algorithm, key length, ESP mode, the SA ID, traffic direction of th SA and more).
- 2. Configure the xfrm policy which defines the type of traffic that will undergo encryption or decryption. It also sets the tunnel IP addresses which encapsulate the packet when working in ESP tunnel mode.

The following example shows how to configure a host (one side of an IPsec connection) with an offloaded IPsec tunnel using the iproute2 utility. In this example, the tunnel is set in IPv4 mode with AES-GCM128 crypto algorithm. The keys are added manually.

1. Set the egress traffic security parameters: *ip xfrm state add src*¹ 192.168.7.2 dst² 192.168.7.9 *proto esp spi*³ 0x4c250336 reqid⁴ 0x4c250336 mode tunnel aead 'rfc4106(gcm(aes))' 0x44e6625f4d2fb01b03cc9baefe9b5c8de9d7b9c1⁵ 128 offload dev ens8 ⁶ dir out⁷

^{1.} The IP addresses of the src host of the egress traffic. Modify it with your own relevant addresses.

^{2.} The IP addresses of the destination host of the egress traffic. Modify it with your own relevant addresses.



Set the ingress traffic security parameters: *ip xfrm state add src*¹ 192.168.7.9 dst² 192.168.7.2 proto esp spi³ 0x0f2e596c reqid 0x0f2e596c mode tunnel aead 'rfc4106(gcm(aes))' 0x44e6625f4d2fb01b03cc9baefe9b5c8de9d7b9c1⁴ 128 offload dev ens8 dir in⁵

Note: *offload dev ens8 dir out* and *offload dev ens8 dir in* are the new flags which instruct the iproute2 utility to enable HW offload for the specified security policy.

- 3. Apply the new egress traffic security policy: *ip xfrm policy add src 192.168.7.2 dst*⁶ *192.168.7.9 dir out tmpl*⁷ *src 192.168.7.2 dst*⁸ *192.168.7.9 proto esp reqid 0x4c250336 mode tunnel*
- 4. Apply the new ingress traffic security policy: *ip xfrm policy add src 192.168.7.9 dst 192.168.7.2 dir in tmpl src 192.168.7.9 dst 192.168.7.2 proto esp reqid 0x0f2e596c mode tunnel*
- **Note:** The above example shows how to configure a host on one side of the IPsec secured connection. The peer host must undergo the same flow listed above only with the traffic directions inverted. That is, the settings of the egress traffic in this example are the settings of the ingress traffic for the peer host.

Once configured, the existing xfrm states (SAs) and policies can be seen using the following commands:

- 1. ip xfrm state to view all the xfrm states in the kernel.
- 2. ip xfrm pol to view all the xfrm policies in the kernel.

When viewing the xfrm states in the system, the flag dir in/dir out (depending on the traffic direction of the state), under the "crypto offload parameters" section, will indicate that this state is offloaded by an offload device. If these flags are not present, it indicates that encryption/ decryption is not offloaded for this xfrm state and remains within the kernel scope.

- 6. The relevant network interface name replace with your own.
- 7. out/in traffic direction of this IPsec tunnel setting.

3. SPI value for ingress traffic - add your own desired value.

5. out/in - traffic direction of this IPsec tunnel setting.

^{3.} SPI value for egress traffic - add your own desired value.

^{4.} SA request id - this ID is used as a reference to the new SA (for modification, destruction, attaching to a policy). Any number can be chosen here.

^{5.} The 128 bit key concatenated with the constant initialization vector (IV) that are used for the encryption of the egress traffic.

^{1.} The IP addresses of the src host of the ingress traffic. Modify it with your own relevant addresses.

^{2.} The IP addresses of the destination host of the ingress traffic. Modify it with your own relevant addresses.

^{4.} The 128 bit key concatenated with the constant initialization vector (IV) that are used for the decryption of the ingress traffic. This traffic key does not have to be similar to the egress traffic key.

^{6.} The IP addresses of the inner (original) packet to undergo transformation and tunnel encapsulation.

^{7.} Indicates that we are about to define the template of the outer IP header of our tunnel.

^{8.} The tunnel source and destination IP addresses - can be different than the inner packet IP address.



5.3.3 Destroying IPsec Tunnels

The process of destroying an offloaded IPsec tunnel remains unchanged and is similar to the destruction of a non-offloaded tunnel using the iproute2 "ip" utility.

It is required to close both the xfrm policies and the xfrm states to completely terminate a session.

A complete flush of all the policies and states can be done with the following commands:

- 1. ip xfrm state flush
- 2. ip xfrm pol flush

To delete a specific xfrm policy use the "ip xfrm policy delete <policy parameters>" command. To delete a specific xfrm state use the "ip xfrm state delete <state parameters>" command.

When flushing all xfrm states, the flow cache is flushed automatically, so this additional operation is not required.

5.3.4 IPsec Offload Statistics

The FPGA contains several counters which provide information and statistics on the offload operation.

These counters are a part of the network interface counters and can be viewed using the ethtool

-S <interface_name> command.

Note: The mlx5 core module must be loaded for the counters to appear in ethtool.

Table 7 - ethtool IPsec Offload Counters

Name	Description
ipsec_dec_in_packets	Total packets received for decryption by FPGA.
ipsec_dec_out_packets	Number of packets that were received for decryption, decrypted and successfully authenticated by FPGA.
ipsec_dec_bypass_packets	Number of packets that were bypassed by FPGA in decryption direction.
ipsec_enc_in_packets	Total packets received for encryption by FPGA.
ipsec_enc_out_packets	Number of packets that were received for encryption, encrypted and successfully authenticated by FPGA.
ipsec_enc_bypass_packets	Number of packets that were bypassed by FPGA in encryption direction.
ipsec_dec_drop_packets	Number of packets dropped by decryption engine. This can be as a result of having inband metadata in packet or corrupted decryption.
ipsec_dec_auth_fail_packets	Number of packets dropped by decryption engine due to authentication issue.



Table 7 - ethtool IPsec Offload Counters

Name	Description
ipsec_enc_drop_packets	Number of packets dropped by encryption engine. This can be as a result of more VLAN tags than the number supported by FPGA, having inband metadata or miss in SADB.
ipsec_add_sa_success	Total amount of SAs successfully added by FPGA.
ipsec_add_sa_fail	Total amount of failed SA add commands by FPGA. This can be a result of adding an already valid SA.
ipsec_del_sa_success	Total amount of SAs successfully removed by FPGA.
ipsec_del_sa_fail	Total amount of failed SA remove commands by FPGA. This can be a result of remove command on invalid SA.
ipsec_cmd_drop	Total amount of failed commands. This can be a result of failure to parse command.



6 mlx_fpga Tool

mlx_fpga tool allows the user to burn and update a new FPGA image on Mellanox Innova IPsec adapter card. The tool also enables the user to read/write individual registers in the FPGA configuration space.

6.1 Tool Requirements

- Mellanox Innova IPsec EN adapter card with an FPGA device
- Download MFT via: <u>www.mellanox.com</u> => Products => Software => InfinBand/VPI Drivers => MFT - Firmware Tools. For supported MFT version, please refer to *Mellanox Innova IPsec EN Release Notes*.
- Extract the TGZ and run install.sh
- Load mlx5_fpga_tools module. See Section 4.2.2, "mlx5_fpga_tools Module," on page 23.
- Start mst service with the fpga lookup flag (mst start --with_fpga)

6.2 mlx_fpga Synopsis

mlx_fpga [-d <device>] < read <addr> | write <addr> <value> | b <image path> | clear_semaphore | reset | load | query

where:

-d device <device></device>	FPGA mst device interface
-v version	Display version info
-h help	Display help message
r read <addr></addr>	Read debug register in address
w write <addr> <data></data></addr>	Write data to debug register in address
b burn <bin></bin>	Burn image on flash
l load	Load image from flash (factory - load image from factory flash)
clear_semaphore	Unlock flash controller semaphore
reset	Reset FPGA (fpga)
q query	Query general FPGA information



6.3 Examples of mlx_fpga Usage

6.3.1 Adding FPGA mst Device Interface

```
# modprobe mlx5_fpga_tools
# mst start --with_fpga
# mst status
MST modules:
------
MST PCI module is not loaded
MST PCI configuration module is not loaded
MST devices:
------
No MST devices were found nor MST modules were loaded.
You may need to run 'mst start' to load MST modules.
FPGA devices:
-------//dev/mst/mt4117_pciconf0_fpga_i2c
/dev/mst/mt4117_pciconf1_fpga_rdma<sup>a</sup>
```

a. It is recommended to use the RDMA device as it uses the "fast path" to the FPGA. I2C is used for recovery purposes when RDMA is not functional.

6.3.1.1 Burning the FPGA's Flash Device Using the mlx_fpga Burning Tool

mlx_fpga tool burns a .bin file onto the FPGA flash device.



It is recommended to burn the FPGA device using an RDMA device as it uses the "fast path" to the FPGA thus minimizing the burning time.

Step 1. Burn the image.

mlx fpga -d <device> burn image.bin

Step 2. Load the FPGA image from flash according to Section 6.3.1.2, "Loading Tool," on page 42 or power cycle the machine for change to take effect.

6.3.1.2 Loading Tool

• Load an FPGA image from user configurable flash:

mlx_fpga -d <device> 1/load <optional: load options>

where <optional: load options> is:

factory	Load FPGA image from factory flash
user	Load FPGA image from user flash [default option]





6.3.1.3 Debugging Tool

- Reading One Debug Register:
 # mlx fpga -d <device> read 0x0
- Writing One Debug Register:

mlx_fpga -d <device> write 0x0 0x0

6.3.1.4 Update FPGA Image

In order to verify the new image burned to the FPGA, the user can use mlx_fpga tool to read the following registers:

Name	Address	Range	Default	RW	Description
image_version	0x900000	31:00:00	0x0	RO	Version of the image
image_date	0x900004	31:00:00	0x0	RO	Image date of creation. The hex number is actually the decimal value, i.e. 0x12011995 means 12/ 01/1995 in DD/MM/YY: bits [31:24] = day of creation bits [23:16] = month of creation bits [15:0] = year of creation
image_time	0x900008	31:00:00	0x0	RO	Image time of creation. The hex number is actually the decimal value, i.e. 0x00015324 means 01:53:24 in HH:MM:SS: bits [23:16] = hour (0023) bits [15:8] = minutes (0059) bits [7:0] = seconds (0059)



7 Updating Mellanox Innova IPsec Adapter Card Firmware



This section applies only when updating the ConnectX-4 Lx firmware. In order to burn and update the FPGA image, please refer to Chapter 6, "mlx_fpga Tool" on page 41.

Each card is shipped with the latest version of qualified ConnectX-4 Lx firmware at the time of manufacturing. However, Mellanox issues firmware updates occasionally. Please contact Mellanox for the correct Firmware version.

Firmware can be updated on the stand-alone single card using the **flint** tool of the *Mellanox Firmware Tools (MFT)* package. Please contact Mellanox for the correct MFT package.

The following steps describe how to retrieve the PSID (firmware identification) and programmed firmware version of your adapter card. They also describe how to update the card with the latest firmware version available.

- 1. Retrieve the PSID and firmware version:
 - a. Install the MFT package.
 - b. Enter: mst start.
 - c. Get the Mellanox *mst device name* using the command "mst status". The mst device name will be of the form: /dev/mst/mt4117_pciconf0.
 - d. Get the PSID (firmware identification) and programmed firmware version using the command ' flint -d <mst device> q', where <mst device> is the device retrieved in step c.



The shown versions and/or parameter values in the example below may not reflect the latest or actual values for this product, and are included here for illustration purposes only.

flint -d /dev/m	st/mt4117_pci_cr0 q		
Image type:	ConnectX-4 Lx		
FW Version:	14.12.1000		
Device ID:	4117		
Chip Revision:	0		
Description:	Node	Port1	Port2
Sys image			
GUIDs:	000002c90000020	0 000002c900000201	000002c900000202
000002c90000020	3		
MACs:			
000002c90200	000002c90201		
Board ID:	(MT 2410110034MT	2490110032)	
VSD:		-	
PSID:	MT 2410110034MT	2490110032	

- 2. To burn the new FW image to ConnectX-4 Lx flash:
 - a. Install mst package and start mst as in section 1 above



b. To burn the firmware, run:

mlxburn -d /dev/mst/mt4117_pciconf0 -i <fw.bin>

c. To load the firmware, run:

mlxfwreset -d /dev/mst/mt4117_pciconf0 reset -y



8 Troubleshooting

- See "General" on page 46.
- See "Linux" on page 47.

8.1 General

	Ensure that the adapter is placed correctly
	Make sure the adapter slot and the adapter are compatible
Server unable to find the adapter	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
	Reseat the adapter in its slot or a different slot, if necessary
The adapter no longer	• Try using another cable
works	• Reinstall the drivers for the network driver files may be damaged or deleted
	Reboot the server
Adapters stopped	Try removing and re-installing all adapters
working after installing	Check that cables are connected properly
another adapter	Make sure your motherboard has the latest BIOS
	Engure that adopter driver/a is loaded
	• Ensure that adapter driver/s is loaded
	• Iry another port on the switch
Link indicator light is off	• Make sure the cable is securely attached
	• Check you are using the proper cables that do not exceed the recommended
	Ichiguis
	• Verify that your switch and adapter port are compatible
Link light is on, but with	Check that the latest driver is loaded
no communication	• Check that both the adapter and its link are set to the same speed and duplex set-
established	tings
	Verify the Mellanox Innova IPsec kernel is loaded
FPGA not found on mst status	Load mlx5_fpga_tools module
	Start mlx_fpga tool
	uname -r
	modprobe mlx5 fpga tools
	mst startwith fpga
	mst status
	Reinstall MFT





8.2 Linux

Environment Information	cat/etc/issue uname –a cat/proc/cupinfo grep 'model name' uniq ofed_info head -1 ifconfig –a 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: <u>http://www.mellanox.com/content/</u> <u>pages.php?pg=management_tools&menu_section=34</u> Refer to the User Manual for installation instructions. Once installed, run: mst start mst start flint -d <mst_device> q</mst_device>	
Ports Information	ibstat lbv_devinfo	
Firmware Version Upgrade	To download the latest firmware version refer to http://www.mellanox.com/supportdownloader	
Collect Log File	/var/log/messages dmesg > system.logF	



9 Specifications

9.1 MNV101512A-BCIT Specifications

Table 8 - MNV101512A-BCIT Specifications Table

Physical	Size: 2.7 in. x 6.6 in. (68.9 mm x167.65 mm)
r nysicai	Connector: Single QSFP (Copper and optical)
Protocol Support	Ethernet: 40GBASE-CR4, 40GBASE-KR4, 40GBASE-SR4, 40GBASE-ER4, 40GBASE-R2, 10GBASE-SR, 10GBASE-LR,10GBASE-ER, 10GBASE-CR, 10GBASE-KR
i rotocoi Support	Data Rate: 10/40Gb/s – Ethernet
	PCI Express Gen3: SERDES @ 8.0GT/s, 8 lanes (2.0 and 1.1 compatible)
	Voltage: 3.3Vaux, 12V
	10GbE Typ Power: Passive Cables: 18.7W
	10GbE Max Power: Passive Cables: 20.2W 1.5W Active Cables: 21.7W
	40GbE Typ Power: Passive Cables: 30W
Power and	40GbE Max Power: Passive Cables: 31.5W 1.5W Active Cables: 33W
Environmentar	Max power available through QSFP port: 1.5W
	Temperature: Operational 0°C to 55°C ^a Non-operational -40°C to 70°C
	Humidity: 90% relative humidity ^b
	Air Flow ^{cd} : 100LFM at 35°C/45°C 200 LFM at 55°C
Regulatory	Safety: IEC/EN 60950-1:2006 ETSI EN 300 019-2-2 IEC 60068-2- 64, 29, 32
	RoHS: RoHS-R6 compliant
Cable Support	Please refer to http://www.mellanox.com/products/interconnect/cables-configurator.php

a. Ambient temperature may vary. Please contact Mellanox technical support if further assistance is needed.

b. For both operational and non-operational states.

c. Air flow is measured ~1" from the heat sink between the heat sink and the cooling air inlet.

d. Airflow requirements may vary according to ambient temperature and other parameters. Please contact Mellanox technical support if further assistance is needed.



9.2 MNV101511A-BCIT Specifications

Table 9 - MNV101511A-BCIT Specifications Table

Dhysiaal	Size: 2.7 in. x 6.6 in. (68.9 mm x167.65 mm)
rnysicai	Connector: Single QSFP (Copper and optical)
Protocol Support	Ethernet: 40GBASE-CR4, 40GBASE-KR4, 40GBASE-SR4, 40GBASE-ER4, 40GBASE-R2, 10GBASE-SR, 10GBASE-LR,10GBASE-ER, 10GBASE-CR, 10GBASE-KR
	Data Rate: 10/40Gb/s – Ethernet
	PCI Express Gen3: SERDES @ 8.0GT/s, 8 lanes (2.0 and 1.1 compatible)
	Voltage: 3.3Vaux, 12V
	10GbE Typ Power: Passive Cables: 16.7W
	10GbE Max Power: Passive Cables: 18.2W 1.5W Active Cables: 19.7W
	40GbE Typ Power: Passive Cables: 30W
Power and	40GbE Max Power: Passive Cables: 29.5W 1.5W Active Cables: 31W
Environmental	Max power available through QSFP port: 1.5W
	Temperature: Operational 0°C to 55°C ^a Non-operational -40°C to 70°C
	Humidity: 90% relative humidity ^b
	10GbE Air Flow ^{cd} : 250LFM at 35°C 300LFM at 45°C 350LFM at 55°C
	40GbE Air Flow: 300LFM at 35°C/ 350LFM at 45°C 500 LFM at 55°C
Regulatory	Safety: IEC/EN 60950-1:2006 ETSI EN 300 019-2-2 IEC 60068-2- 64, 29, 32
	RoHS: RoHS-R6 compliant
Cable Support	Please refer to http://www.mellanox.com/products/interconnect/cables-configurator.php

a. Ambient temperature may vary. Please contact Mellanox technical support if further assistance is needed.

b. For both operational and non-operational states.

c. Air flow is measured ~ 1 " from the heat sink between the heat sink and the cooling air inlet.

d. Airflow requirements may vary according to ambient temperature and other parameters. Please contact Mellanox technical support if further assistance is needed.



9.3 Mellanox Innova IPsec EN LEDs

Figure 3: MNV101511A-BCIT/MNV101512A-BCIT LEDs Placement¹ (Example)



Group A LEDs: Network LEDs - these LEDs indicate the network link status. See Section 9.3.1, "Network LEDs Operation," on page 51 for details.

Group **B** LEDs: Debug LEDs - indicate memory calibration done, memory BIST done, ConnectX-4 Lx link up is with traffic, Heartbeat and power good. See Section 9.3.2, "FPGA Debug LEDs," on page 51 for details.

Group C LEDs: FPGA load-flow Debug LEDs - see Section 9.3.3, "FPGA Load-Flow Debug LEDs," on page 52.



The image shown is for informational purposes and does not necessarily represent the latest revision of the card.

^{1.} The adapter card is shipped with the heat sink assembled.





9.3.1 Network LEDs Operation

Table 10 - Physical and Logical Link Indications

LED	Function	LED Symbol
Amber - physical link	 Constant on indicates a good physical link Blinking indicates a problem with the physical link If neither LED is lit, then the physical link has not been established 	D1
Green - logical (data activity) link	 Constant on indicates a valid logical (data activity) link without data transfer. Blinking indicates a valid logical link with data transfer If only the green LED is lit and the Amber LED is off, then the logical link has not been established 	

9.3.2 FPGA Debug LEDs

Note: D2-D9 are the "symbols" of these LEDs as printed on the board.

LED Symbols	LED Function
D2	Power Good - Or on all POWER-GOOD inputs. Expected LED ON.
D3	Clock - the LED starts blinking once out of reset and the clock is running. Expected blinking LED 1Hz.
D4	DDR Calibration DONE - the LED will be ON after power-up, if DDR calibration is successful.
D5	DDR BIST Passed - DDR Built In Test runs once after power-up. LED will turn on if test passes successfully.
D6	ConnectX Port Ready - the LED will be ON when FPGA-ConnectX link is up.
D7	ConnectX Port Traffic - the LED will blink when there is FPGA-ConnectX traffic (TX/SX).
D8	Network Port Ready - the LED will be ON when FPGA-Network link is up.
D9	Network Port Traffic - the LED will blink when there is FPGA-Network traffic (TX/SX).

Table 11 - FPGA Debug LEDs



9.3.3 FPGA Load-Flow Debug LEDs

Table 12 - FPGA Load-Flow Debug LEDs

LED	LED Symbol and Function
Green - power good Off - power issue	D10 - Power Good
Red - during configuration Green - when complete	D11- Configuration Done Indication
Red - factory default Green - user image	D12 - Configuration Image Selection

9.4 Board Mechanical Drawing and Dimensions



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









Figure 5: Mechanical Drawing of MNV101512A-BCIT



9.5 Bracket Mechanical Drawing



Figure 6: Single-Port Tall Bracket



Specifications



Figure 7: Single-Port Short Bracket



Appendix A: Fast Installation and Update

A.1 Hardware Installation

- 1. Shut down your system if active.
- 2. After shutting down the system, turn off power and unplug the cord.
- 3. Place the adapter in a standard PCI Express slot.

For further details, please refer to Chapter 3, "Hardware Installation" on page 17.

A.2 Content of Mellanox Innova IPsec Bundle

Mellanox provides an IPsec bundle which includes the following:

- FPGA image bin file
- Kernel RPM files
- MFT tarball file
- Firmware bin files
- Offload scripts (xfrm, iproute)

A.3 Software, Firmware and Tools Installation

The following instructions apply to installation only. If the bundle is already installed, please refer to Appendix A.4, "Software, Firmware and Tools Update," on page 58.

Please make sure to install in the following order:

Step 1. Download the bundle from <u>www.mellanox.com</u> => Products => SmartNICs => Innova IPsec => FW & SW. Each card is shipped with the latest version of the qualified FPGA image and firmware at the time of manufacturing. Please download the Mellanox Innova IPsec bundle that matches the FPGA image burned on your card.

> To install the kernel:

- Step 1. Locate the RPM files in the Kernel folder:
 - rpm -i kernel-<kernel_version>.rpm
 - rpm -i kernel-devel-<kernel_version>.rpm
- Step 2. Reboot your system and select the relevant kernel to load from the grub menu.

> To install MFT:

- Step 1. Untar the MFT tar file.
- **Step 2.** Install MFT by running:

install.sh

Step 3. Start MFT:

Step a. modprobe mlx5_fpga_tools



Step b. mst start --with fpga

Step c. mst status

a. It is recommended to use the RDMA device as it uses the fast path to the FPGA. I2C is used for recovery purposes when RDMA is not functional.

> To install the FPGA image:

- **Step 4.** In the bundle folder/directory look for the installation script: mlnx_fpga_updater.sh
 - Step a. The script is located in <downloaded bundle folder >/Scripts
 - Step b. Once you have the script you can run it in the following ways:
 - 1. Basic:

./Scripts/mlnx_fpga_updater.sh

2. 40G image is the default image. It is possible to choose between 10G and 40G:

```
./Scripts/mlnx_fpga_updater.sh -f 10g
./Scripts/mlnx fpga updater.sh -f 40g
```

3. Device - If there is only one RDMA device it will be auto installed. If there is more than 1 device, it is possible to choose a device:

```
./Scripts/mlnx_fpga_updater.sh -d /dev/mst/mt4117_pciconf0
______All
/dev/mst/mt4117_pciconf0
/dev/mst/mt4117_pciconf1
```

Step 5. From this point on the script will install the FPGA image, the FW and will also ask if to install the MFT and do a reset at the end.



A.4 Software, Firmware and Tools Update

Step 1. To download the bundle, please refer to <u>www.mellanox.com</u> => Products => Smart-NICs => Innova IPsec => FW & SW

> To install the most updated kernel:

Step 1. Locate the RPM files in the Kernel folder:

- rpm -i kernel-<kernel_version>.rpm
- rpm -i kernel-devel-<kernel_version>.rpm
- Step 2. Reboot your system and select the relevant kernel to load from the grub menu.

For further details, please refer to Chapter 5, "IPsec Offload Software Installation and Operation" on page 25.

> To update MFT:

Step 1. Untar the MFT tar file.

Step 2. Install MFT by running:

install.sh

Step 3. Start MFT:

Step a. modprobe mlx5_fpga_tools

Step b. mst start --with_fpga

```
Step c. mst status
```

```
# modprobe mlx5_fpga_tools
# mst start --with_fpga
# mst status
MST modules:
------
MST PCI module is not loaded
MST PCI configuration module is not loaded
MST devices:
------
No MST devices were found nor MST modules were loaded.
You may need to run 'mst start' to load MST modules.
FPGA devices:
-------
/dev/mst/mt4117_pciconf0_fpga_i2c
/dev/mst/mt4117_pciconf1_fpga_rdma<sup>a</sup>
```

a. It is recommended to use the RDMA device as it uses the fast path to the FPGA. I2C is used for recovery purposes when RDMA is not functional.

> To update the FPGA image:

Step 4. In the bundle folder/directory look for the installation script: mlnx_fpga_updater.sh

Step a. The script is located in <downloaded bundle folder >/Scripts

Step b. Once you have the script you can run it in the following ways:



1. Basic:

./Scripts/mlnx_fpga_updater.sh

2. 40G image is the default image. It is possible to choose between 10G and 40G:

```
./Scripts/mlnx_fpga_updater.sh -f 10g
./Scripts/mlnx fpga updater.sh -f 40g
```

3. Device - If there is only one RDMA device it will be auto installed. If there is more than 1 device, it is possible to choose a device:

```
./Scripts/mlnx_fpga_updater.sh -d /dev/mst/mt4117_pciconf0
------
All
/dev/mst/mt4117_pciconf0
/dev/mst/mt4117 pciconf1
```

Step 5. From this point on the script will install the FPGA image, the FW and will also ask if to install the MFT and do a reset at the end.

A.5 OFED Installation with Script

- **Step 1.** Download OFED from <u>www.mellanox.com</u> => Products => Software => InfiniBand/ VPI drivers => Linux SW/Drivers. In this case after OFED installation, the kernel and MFT are already installed.
- Step 2. Get URL of IPsec bundle from <u>www.mellanox.com</u> => Products => SmartNICs => Innova IPsec => Download
- Step 3. You can run the following update script using one of the modes below:

/opt/mellanox/mlnx-fw-updater/mlnx fpga updater.sh

1. With -u flag to provide URL to the software package (tarball). Example:

./mlnx_fpga_updater.sh -u http://www.mellanox.com/downloads/fpga/ipsec/ Innova IPsec <version>.tgz

2. With -t flag to provide the path to the downloaded tarball. Example:

./mlnx fpga updater.sh -t <Innova IPsec bundle file.tgz>

3. With -p flag to provide the path to the downloaded and extracted tarball. Example:

./mlnx_fpga_updater.sh -p <Innova_IPsec_extracted_bundle_directory>

For more information on the script usage, you can run mlnx_fpga_updater.sh -h.



It is recommended to perform a firmware and FPGA upgrade on Mellanox Innova IPsec cards using this script only.



Appendix B: Finding the MAC and Serial Number on the Adapter Card

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



The revision indicated on the labels in the following figures do not necessarily represent the latest revision of the card.



Figure 9: MNV101512A-BCIT Board Label





Appendix C: Safety Warnings

1. Installation Instructions



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

2. Over-temperature



This equipment should not be operated in an area with an ambient temperature exceeding the maximum recommended: 55° C (131°F).

To guarantee proper air flow, allow at least 8cm (3 inches) of clearance around the ventilation openings.

3. During Lightning - Electrical Hazard



During periods of lightning activity, do not work on the equipment or connect or disconnect cables.

4. 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.

5. Equipment Installation



This equipment should be installed, replaced, or serviced only by trained and qualified personnel.

6. Equipment Disposal



Disposal of this equipment should be in accordance to all national laws and regulations.

7. Local and National Electrical Codes



This equipment should be installed in compliance with local and national electrical codes.



8. Hazardous Radiation Exposure



Caution – Use of controls or adjustment or performance of procedures other than those specified herein may result in hazardous radiation exposure.



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+ A2:20.



Appendix D: Avertissements de sécurité d'installation (Warnings in French)

1. Instructions d'installation



Lisez toutes les instructions d'installation avant de brancher le matériel à la source d'alimentation électrique.

2. Température excessive



Ce matériel ne doit pas fonctionner dans une zone avec une température ambiante dépassant le maximum recommandé de 55°C (131°F). Un flux d'air de 200LFM à cette température ambiante maximale est nécessaire. En outre, pour garantir un bon écoulement de l'air, laissez au moins 8 cm (3 pouces) d'espace libre autour des ouvertures de ventilation.

3. Orages – dangers électriques



Pendant un orage, il ne faut pas utiliser le matériel et il ne faut pas brancher ou débrancher les câbles.

4. Branchement/débranchement des câbles en cuivre



Les câbles en cuivre sont lourds et ne sont pas flexibles, il faut donc faire très attention en les branchant et en les débranchant des connecteurs. Consultez le fabricant des câbles pour connaître les mises en garde et les instructions spéciales.

5. Installation du matériel



Ce matériel ne doit être installé, remplacé ou entretenu que par du personnel formé et qualifié.

6. Elimination du matériel



L'élimination de ce matériel doit s'effectuer dans le respect de toutes les législations et réglementations nationales en vigueur.

7. Codes électriques locaux et nationaux



Ce matériel doit être installé dans le respect des codes électriques locaux et nationaux.



8. Exposition au rayonnement grave



Mise en garde – l'utilisation de commandes ou de réglages ou l'exécution de procédures autres que ce qui est spécifié dans les présentes peut engendrer une exposition au rayonnement grave.



PRODUIT LASER DE CLASSE 1 » et références aux normes laser les plus récentes CEI 60 825-1



Appendix E: Sicherheitshinweise (Warnings in German)

1. Installationsanleitungen



Lesen Sie alle Installationsanleitungen, bevor Sie das Gerät an die Stromversorgung anschließen.

2. Übertemperatur



Dieses Gerät sollte nicht in einem Bereich mit einer Umgebungstemperatur über der maximal empfohlenen Temperatur von 55°C (131°F) betrieben werden. Es ist ein Luftstrom von 200 LFM bei maximaler Umgebungstemperatur erforderlich. Außerdem sollten mindestens 8 cm (3 in.) Freiraum um die Belüftungsöffnungen sein, um einen einwandfreien Luftstrom zu gewährleisten.

3. Bei Gewitter - Elektrische Gefahr



Arbeiten Sie während eines Gewitters und Blitzschlag nicht am Gerät, schließen Sie keine Kabel an oder ab.

4. Anschließen/Trennen von -Kupferkabel



Kupferkabel sind schwer und nicht flexible. Deshalb müssen sie vorsichtig an die Anschlüsse angebracht bzw. davon getrennt werden. Lesen Sie die speziellen Warnungen und Anleitungen des Kabelherstellers.

5. Geräteinstallation



Diese Gerät sollte nur von geschultem und qualifiziertem Personal installiert, ausgetauscht oder gewartet werden.

6. Geräteentsorgung



Die Entsorgung dieses Geräts sollte unter Beachtung aller nationalen Gesetze Bestimmungen erfolgen.

7. Regionale und nationale elektrische Bestimmungen t



Dieses Gerät sollte unter Beachtung der regionalen und nationalen elektrischen Bestimmungen installiert werden.



8. Strahlenkontak



Achtung – Nutzung von Steuerungen oder Einstellungen oder Ausführung von Prozeduren, die hier nicht spezifiziert sind, kann zu gefährlichem Strahlenkontakt führen.



Klasse 1 Laserprodukt und Referenzen zu den aktuellsten Lasterstandards : ICE 60 825-1



Appendix F: Advertencias de seguridad para la instalación (Warnings in Spanish)

1. Instrucciones de instalación



Antes de conectar el equipo a la fuente de alimentación, leer todas las instrucciones de instalación.

2. Sobrecalentamiento



No se debe utilizar el equipo en un área con una temperatura ambiente superior a la máxima recomendada: 55°C(131°F). Además, para garantizar una circulación de aire adecuada, se debe dejar como mínimo un espacio de 8 cm (3 pulgadas) alrededor de las aberturas de ventilación.

3. Cuando hay rayos: peligro de descarga eléctrica



No utilizar el equipo ni conectar o desconectar cables durante períodos de actividad de rayos.

4. Conexión y desconexión del cable Copper



Dado que los cables de cobre son pesados y no son flexibles, su conexión a los conectores y su desconexión se deben efectuar con mucho cuidado. Para ver advertencias o instrucciones especiales, consultar al fabricante del cable.

5. Instalación de equipos



La instalación, el reemplazo y el mantenimiento de este equipo estarán a cargo únicamente de personal capacitado y competente.

6. Eliminación de equipos



La eliminación definitiva de este equipo se debe efectuar conforme a todas las leyes y reglamentaciones nacionales.

7. Códigos eléctricos locales y nacionales



Este equipo se debe instalar conforme a los códigos eléctricos locales y nacionales.



8. Exposición a niveles de radiación peligrosos



Precaución: el uso de controles o ajustes o la realización de procedimientos distintos de los que aquí se especifican podrían causar exposición a niveles de radiación peligrosos.



PRODUCTO LÁSER DE CLASE 1 y referencia a las normas de láser más recientes: IEC 60825-1