Innova-2™ Flex Open Adapter Card Documentation

NOTE:
THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT (“PRODUCT(S)”) 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) 970-3400
Fax: (408) 970-3403

Table of Contents

1 25Gb/s Ethernet and 100Gb/s VPI Application Acceleration Platforms 6
2 Innova-2™ Flex Adapter Card - Software and Firmware Bundle Release Notes 8
2.1 Overview 8
2.1.1 Innova-2 Flex Open Bundle Contents 8
2.1.2 Supported Platforms/Operating Systems 8
2.1.3 List of Supported Cables 9
2.2 Changes and New Features 10
2.2.1 Changes and New Features History 10
2.2.2 Upgrading to Bundle 18.12 11
2.2.3 Bundle Interoperability 11
2.3 Known Issues 11
2.4 Bug Fixes 12
2.5 Introcuction 13
2.5.1 Product Overview 13
2.5.1.1 Block Diagram 14
2.5.1.2 Features and Benefits 15
2.5.1.3 System Requirements 18
2.6 Interfaces 18
2.6.1 InfiniBand Interface 18
2.6.2 Ethernet Interface 18
2.6.3 PCI Express Interface 18
2.6.4 LED Indications 19
2.6.5 JTAG Interface 19
2.6.6 FPGA DDR4 Interface 19
2.6.6.1 Importing CSV Inputs from Vivado 20
2.6.7 FPGA PCIe ConnectX-5 Interface 24
2.6.8 FPGA OpenCAPI Interface 24
2.6.9 FPGA QSPI Flash Interface 24
2.6.10 On-Board Clock 25
2.7 Hardware Installation 26
2.7.1 System Requirements 26
2.7.1.1 Hardware 26
2.7.1.2 Operating Systems/Distributions 26
2.7.2 Safety Precautions 26
2.7.3 Pre-installation Checklist 26
2.7.4 Bracket Installation Instructions 26

2.7.4.1 Removing the Existing Bracket 27
2.7.4.2 Installing the New Bracket 27
2.7.5 Card Installation Instructions 27
2.7.6 Cables and Modules 27
2.7.6.1 Cable Installation 27
2.7.6.2 Xilinx Programming Cable 28
2.7.7 Identify the Card in Your System 28
2.8 Specifications 29
2.8.1 MNV303212A-ADLT Specifications 29
2.8.2 MNV303611A-EDLT Specifications 31
2.8.3 Innova-2 LEDs 33
2.8.3.1 Network LEDs Operation 33
2.8.3.2 FPGA LEDs 34
2.8.3.3 General LEDs 34
2.8.4 Board Mechanical Drawing and Dimensions 34
2.9 Innova-2 Flex Open Card Driver 36
2.9.1 Linux Driver Installation 36
2.9.1.1 Hardware and Software Requirements 37
2.9.2 Downloading Mellanox OFED 37
2.9.3 Installing Mellanox OFED 38
2.9.3.1 Installation Script 38
2.9.3.2 Installation Results 42
2.9.3.3 Installation Logging 42
2.9.3.4 Driver Load upon System Boot 43
2.9.3.5 mlnxofedinstall Return Codes 43
2.9.3.6 Uninstalling MLNX_OFED 43
2.9.3.7 UEFI Secure Boot 44
2.9.3.8 Enrolling Mellanox's x.509 Public Key on Your Systems 44
2.9.3.9 Removing the Signature from Kernel Modules 45
2.10 Using the Innova-2 Flex Open Bundle 45
2.10.1 Installing the Innova-2 Flex Open Application 46
2.10.2 Running the Innova-2 Flex Open Application 46
2.10.2.1 Identifying the System State 47
2.10.3 Diagnostic Capabilities 49
2.10.3.1 Running the Diagnostics 49
2.10.3.2 JTAG Access to the FPGA 52
2.10.4 Burning Capabilities 54
2.10.4.1 Running the Burning Flows 54

2.10.5 Switching between Images 55
2.10.5.1 Flash Format 55
2.10.6 Burning the FPGA Image through the JTAG 56
2.10.7 Reloading the Innova-2 User Image 58
2.10.7.1 Procedure Scenario 58
2.10.7.2 Finding the PCI Switch 59
2.10.7.3 Disabling/Enabling the PCI Link between the FPGA and the PCI Switch 59
2.10.7.4 Reloading the User Image 59
2.10.8 Innova-2 Flex Open Bundle Content 60
2.10.8.1 FPGA Images on Card: Factory/Flex/User 61
2.10.8.2 Innova-2 Flex 25G vs. Innova-2 Flex VPI 61
2.11 Using the FPGA on the Adapter Card 61
2.12 Appendixes 62
2.12.1 Finding the MAC and Serial Number on the Adapter Card 62
2.12.2 Safety Warnings 62
2.13 User Manual Revision History 63
2.14 Release Update History 64

1 25Gb/s Ethernet and 100Gb/s VPI Application Acceleration Platforms

About This Manual

Mellanox Innova-2 Flex Open is a family of innovative adapters that combine the advanced ConnectX®-5 VPI network controller ASIC with a state-of-the-art FPGA. Maximizing network efficiency and scalability, Innova-2 Flex offers customers an open platform for developing custom- made offloads for a range of applications, including Storage, High Performance Computing (HPC), Machine Learning, Security, Networking, and more.

Innova-2 Flex Open dual-port 25Gb/s Ethernet and 100Gb/s VPI network adapters combine ConnectX-5 with a fully open programmable Xilinx® FPGA, utilizing the Xilinx Vivado Design Suite development environment and Mellanox tools suite. The Vivado license is to be obtained from Xilinx.

With Innova-2 Flex Open, FPGA resources are fully dedicated to the customer's application logic. Innova-2 Flex Open reduces TCO by combining the FPGA acceleration with the network card on a single PCIe slot. The FPGA is connected to the host via an embedded PCIe switch supporting x8 Gen4, and is therefore visible to the host as a PCIe device.

This User Manual includes the following attachments:

1. Verilog_VHDL_and_Xilinx_Design_Constrains.zip:
   - Verilog (.v) User pinout top level
   - VHDL (.vhd) User pinout top level
   - Xilinx Design Constrains (.xdc): Pinout location and type, clock rates, and device configuration
2. DDR Configuration Excel
3. Innova-2 Flex Open Interface Pinouts
4. Mellanox Innova-2™ Flex Xilinx PCI Express DMA with DDR4 Example Design

Ordering Part Numbers

The table below provides the ordering part numbers (OPN) for the available ConnectX-5 Ethernet adapter cards.

IC in Use	OPN	Marketing Description
ConnectX®-5	MNV303212A-ADLT	Innova-2 Flex Open for Application Acceleration, dual-port SFP28, 25GbE, KU15P, 8GB, No Crypto, PCI4.0 x8, HHHL, active heat sink, tall bracket
	MNV303611A-EDLT	Innova-2 Flex Open VPI, dual-port QSFP28, EDR / 100GbE, KU15P, No memory, No Crypto, PCI4.0 x8, HHHL, passive heat sink, 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.

Recommended Documents and Tools

In order to write your own FPGA logic, the table below lists the recommended Xilinx Vivado tools and documents. Xilinx tools are not provided by Mellanox, please contact your Xilinx representative for details and purchasing information.

Document Name/Tools	Location
Xilinx Vivado Tools and Documents	
VIVADO Tool Datasheet	https://www.xilinx.com/content/dam/xilinx/support/documentation/sw_manuals/xilinx2018_3/ug910-vivado-getting-started.pdf
VIVADO Release Notes	https://www.xilinx.com/support/documentation/sw_manuals/xilinx2018_3/ug973-vivado-release-notes-install-license.pdf
VIVADO Programming and Debugging	https://www.xilinx.com/support/documentation/ sw_manuals/xilinx2018_3/ug908-vivado-programming-debugging.pdf
Xilinx Ultrascale Plus Documentation	https://www.xilinx.com/products/silicon-devices/fpga/kintex-ultrascale-plus.html#documentation
Xilinx Ultrascale Overview	https://www.xilinx.com/support/documentation/data_sheets/ds890-ultrascale-overview.pdf
Ultrascale Packaging and Pinout	https://www.xilinx.com/support/documentation/user_guides/ug575-ultrascale-pkg-pinout.pdf
Mellanox Related Documents	
Mellanox Firmware Tools (MFT) User Manual Document no. 2204UG	User Manual describing the set of MFT firmware management tools for a single node. See http://www.mellanox.com/page/management_tools
Product Release Notes	Release Notes for Innova-2 Adapter Card.
Performance Tuning Guidelines for Mellanox Network Adapters	Community post describing important tuning parameters and settings that can improve performance for Mellanox drivers. See https://community.mellanox.com/s/article/performance-tuning-for-mellanox-adapters
IEEE Std 802.3 Specification	This is the IEEE Ethernet specification: http://standards.ieee.org/getieee802
PCI Express® Base Specification and Errata for the PCI Express Base Specification	Describes the PCI Express® architecture, interconnect attributes, fabric management, and the programming interface required to design and build systems and peripherals that are compliant with the PCI Express Specification. http://pcisig.com/specifications

Revision History

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

2 Innova-2™ Flex Adapter Card - Software and Firmware Bundle Release Notes

2.1 Overview

These are the release notes for the GA release of Innova_Flex_Open_18_12. Please refer to the following table for the list of supported software, firmware and firmware tools versions. For a list of firmware changes and new features in ConnectX-5, please ask your Mellanox representative for the latest ConnectX5-FW-16_24_4020 Release Notes.

2.1.1 Innova-2 Flex Open Bundle Contents

Innova-2 Flex Open Bundle Contents Versions

Content	Version
FPGA Factory Image Version	192
FPGA Flex Image Version	193
Vivado Version	Vivado 2017.3
PCI Burn Application Version	18.12.00

Mellanox OFED and MFT are not included in the bundle and can be obtained here:
Mellanox Firmware, OFED and MFT

Version	Reference
FW 16.24.4000	with MLNX_OFED 4.5-1.0.1.0
FW 16.25.1000	with MLNX_OFED 4.6
MFT Version: mft-4.11.0	http://www.mellanox.com/page/management_tools
Mellanox OFED Version: Rev 4.5-1.0.1.0.3	http://www.mellanox.com/page/products_dyn?product_family=26&mtag=linux_sw_drivers

2.1.2 Supported Platforms/Operating Systems

Table 4 - Supported Platforms/Operating Systems

Operating Systems	X86	PPC - LE	PPC - BE
RHEL 7.5	X	X	X
RHEL 7.6 ALT		X	

Operating Systems	X86	PPC - LE	PPC - BE
Ubuntu 18.04.01	X	X	
SLES 15 SP0	X		

2.1.3 List of Supported Cables

Validated and Supported 10/40GbE Cables

Speed	Cable OPN #	Length Tested (m)	Description
25GbE	MCP2M00-A005AM	5	Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 5m, 26AWG
25GbE	MCP2M00-A005E26L	5	Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 5m, Black, 26AWG, CA-L
25GbE	FTLF8536P4BCL	-	Finisar 25GE SR SFP28 Optical Transceiver
25GbEx4	MCP7F00-A002	2	Mellanox® passive copper hybrid cable, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 2m, 30AWG
25GbEx4	MCP7F00-A002R30N	2	Mellanox® passive copper hybrid cable, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 2m, Colored, 30AWG, CA-N
25GbEx4	MCP7F00-A005R26L	5	Mellanox® passive copper hybrid cable, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 5m, Colored, 26AWG, CA-L
25GbEx4	MCP7F00-A005AM	5	Mellanox® passive copper hybrid cable, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 5m, 26AWG
25GbEx4	MFA7A50-C030	30	Mellanox® active fiber hybrid solution, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 30m
10GbEx4	MC2609125-005	5	Mellanox® passive copper hybrid cable, ETH 40GbE to 4x10GbE, QSFP to 4xSFP+, 5m
10GbE	MC3309124-005	5	Mellanox® passive copper cable, ETH 10GbE, 10Gb/ s, SFP+, 5m
10GbE	MC3309124-007	7	Mellanox® passive copper cable, ETH 10GbE, 10Gb/ s, SFP+, 7m

Speed	Cable OPN #	Length Tested (m)	Description
10GbE	MFM1T02A-SR	-	Mellanox® SFP+ optical module for 10GBASE-SR
25GbE	MMA2P00-AS	-	Mellanox® transceiver, 25GbE, SFP28, LC-LC, 850nm, SR, up to 100m
1GbE	MC3208411-T	-	Mellanox® module, ETH 1GbE, 1Gb/s, SFP, Base-T, up to 100m

2.2 Changes and New Features

2.2.1 Changes and New Features History

Changes and New Features History

Feature/Change	Description
Version 18.12.00	
Reload User Image	Added the option to reload user image. See description in "Reload User Image" in the User Manual.
Version 18.11.00	
JTAG Access	Added the option to enable JTAG Access. See description in “JTAG Access to the FPGA” in the User Manual.
Factory Image	Added a Factory Image. See description in "FPGA images on card: Factory/ Flex/User” in the User Manual.
FPGA Power Control	Added the ability to control FPGA power while the Flex Image is running.
FPGA Temperature	Added the ability to obtain FPGA temperature while User Image is active.
Image Burn	Added the ability to burn User Image without using the Innova-2 on- board DDR device.
Innova-2 Flex Application Log	Added a log for Innova-2 Flex application. Resides in /var/log/ morse_install.log".
Innova-2 Flex 25G/Innova-2 Flex VPI Support	Added support for Innova-2 Flex VPI

2.2.2 Upgrading to Bundle 18.12

To perform bundle update from all previous versions to Rev. 18.11 or 18.12, please refer to the 18.11 Release Notes. Note that the instructions in the Release Notes refer to the 18.11 bundle, but upgrading to 18.12 is done in the same manner.

⚠ OpenSM must be shut-off before performing the bundle upgrade.

2.2.3 Bundle Interoperability

Please note that no bundle interoperability is guaranteed.
Components of an Innova Flex Open bundle (FPGA images, ConnectX-5 firmware, BOPE driver and Flex Open Application) are compatible and have been tested only within a certain bundle and there is no assurance that they will work with components from other bundles.

2.3 Known Issues

The following table describes known issues in this bundle release and possible workarounds. For list of known issues for ConnectX-5 firmware, please refer to http://www.mellanox.com/page/ firmware_table_ConnectX5EN.

Known Issues

Internal Ref.	Issue
-	**Description:** The "Reload User Image" feature is not supported.
	Workaround: N/A
	Keywords: Reload User Image
1557678	**Description:** The lspci command shows the wrong speed and width. The device has a physical PCIe interface of x8, yet in the system the device reports a maximum PCIe width of x16.
	Workaround: N/A
	Keywords: lspci, speed, width
1608025	**Description:** For Innova-2 Flex VPI, in Dual-port Bidirectional mode, the performance is lower than specified in the spec.
	Workaround: N/A
	Keywords: Dual-port Bidirectional

Internal Ref.	Issue
1481202	**Description:** FPGA images on flash should be burned using the Innova-2 Flex application. If a JTAG cable is used, it is highly recommended to perform the following actions: 1. Burn the image to the flash 2. Perform power-cycle Note: It is not recommended to load the image to the FPGA and then perform system reboot, as this may leave the FPGA PCI core in an unstable state.
	Workaround: N/A
	Keywords: JTAG image burn, PCI instability

2.4 Bug Fixes

The following table describes bug fixes in this bundle release.

Bug Fixes History

Internal Ref.	Issue
1580893	**Description:** When the User image is set to load, after server reboot there is a small chance that it will not be visible over PCI. This occurs due to a suboptimal loading process of the FPGA image that may cause the OS PCI enumeration to complete before the FPGA is visible over PCI. Please contact Mellanox support in case this behavior is consistent. This will be fixed in future releases.
	Keywords: User image, server reboot, PCI
	Discovered in release: 18.11
	Fixed in release: 18.12
1480139	**Description:** When two Innova_2_Flex_Open cards are installed on a server, the Innova-2 Flex Open application does not function.
	Keywords: burning tool
	Discovered in release: 18.07
	Fixed in release: 18.11
1381592	**Description:** When a User Image is burned through PCI burn tool, the tool deletes another 4KB from the flash after last image byte.

Internal Ref.	Issue
	Keywords: burn tool, flash
	Discovered in release: 18.05
	Fixed in release: 18.07
1382476	**Description:** Burning progress bar will wrap around to zero for images larger than 40 MB. This has no affect on the burn process.
	Keywords: burn progress
	Discovered in release: 18.05
	Fixed in release: 18.07
1383037	**Description:** If the user burns a corrupted User Image and power cycles the machine (which loads the FPGA from flash), the FPGA will not be programmed.
	Keywords: User Image, FPGA
	Discovered in release: 18.05
	Fixed in release: 18.07

2.5 Introcuction

2.5.1 Product Overview

Provides the ordering part number, port speed, number of ports, and PCI Express speed.

Innova-2 Flex 25G MNV303212A-ADLT Open Adapter Card

Ordering Part Number (OPN)	MNV303212A-ADLT
Data Transmission Rate	Ethernet: 25Gb/s
Network Connector Types	Dual-port SFP28
PCI Express (PCIe) SerDes Speed	PCIe 3.0/4.0 x8 16GT/s
RoHS	R6

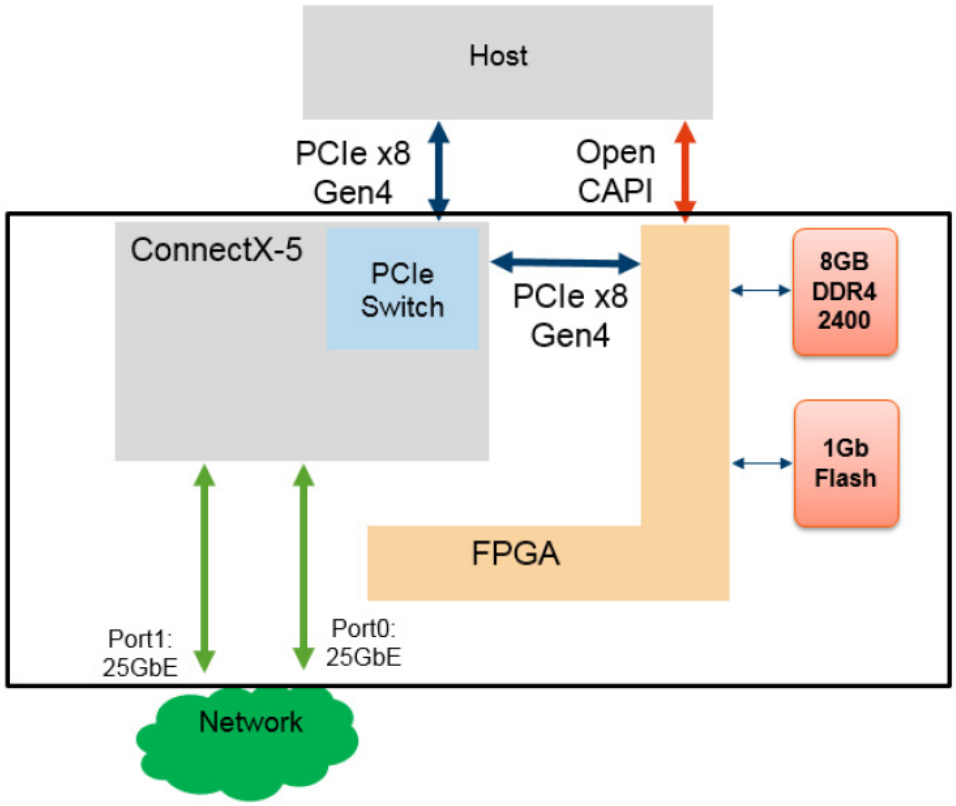
Adapter IC Part Number	ConnectX-5: MT27808A0-FCCF FPGA: XCKU15P-FFVE1517-2- i
Device ID (decimal)	4119

Innova-2 Flex VPI MNV303611A-EDLT Open Adapter Card

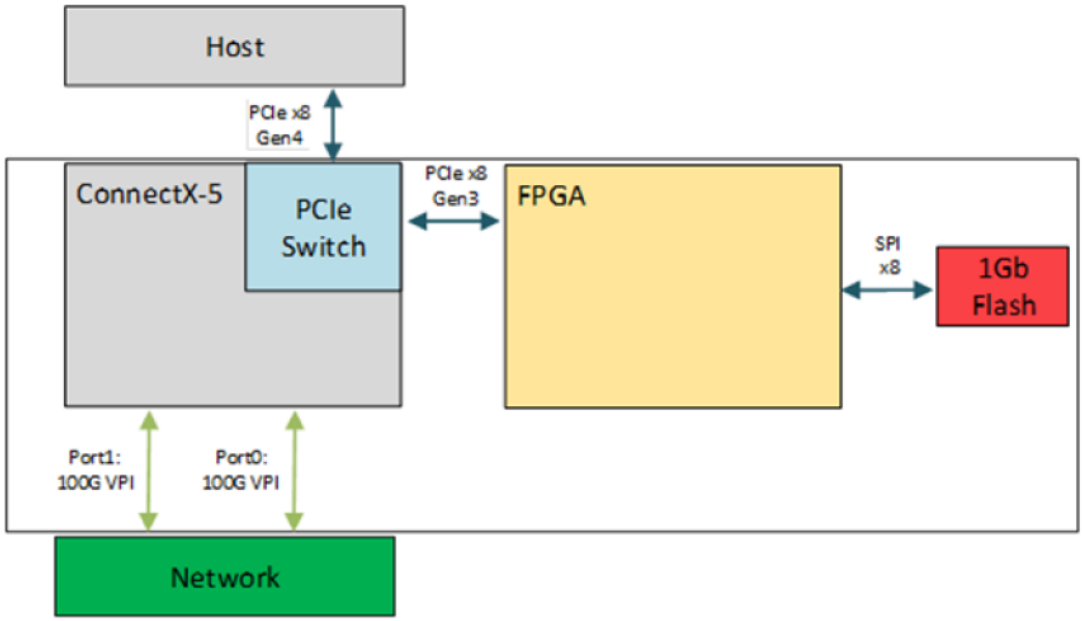
Ordering Part Number (OPN)	MNV303611A-EDLT
Data Transmission Rate	QDR/EDR and 40/100GbE
Network Connector Types	Dual-port QSFP InfiniBand and Ethernet
PCI Express (PCIe) SerDes Speed	PCIe 3.0/4.0 x8 16GT/s
RoHS	R6
Adapter IC Part Number	ConnectX-5: MT27808A0-FCCF FPGA: XCKU15P-FFVE1517-2-1
Device ID (decimal)	4119

2.5.1.1 Block Diagram

MNV303612A-ADLT Adapter Card Block Diagram



MNV303611A-EDLT Adapter Card Block Diagram



2.5.1.2 Features and Benefits

Innova-2 Flex Features [a]

Feature	Description
PCI Express (PCIe)	• PCIe Gen 4.0, 3.0, 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 PCIe Atomic • OpenCAPI support (Open Coherent Accelerator Processor Interface) - In MNV303212A-ADLT only • TLP (Transaction Layer Packet) Processing Hints (TPH) • Access Control Service (ACS) for peer-to-peer secure communication • Advance Error Reporting (AER) • Process Address Space ID (PASID) Address Translation Services (ATS) • Support for MSI/MSI-X mechanisms

Speed	Mellanox adapters comply with the following IEEE 802.3* standards: • 25GbE / 10GbE • IEEE 802.3bj, 802.3bm 25 Gigabit Ethernet • IEEE 802.3by, Ethernet Consortium 25 Gigabit • 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 • IEEE 1588v2 • Jumbo frame support (9600B) **Note:** • Up to 25 Gigabit Ethernet for MNV303212A-ADLT • Up to 100 Gigabit Ethernet for MNV303611A-EDLT
Memory	8GB DDR4-2400 (in MNV303212A-ADLT only)
RDMA and RDMA over Converged Ethernet (RoCE)	The Innova-2 Flex Open adapter supports RoCE specifications delivering low- latency and high- performance over Ethernet networks. Leveraging data center bridging (DCB) capabilities as well as Innova-2 Flex Open adapter advanced congestion control hardware mechanisms, RoCE provides efficient low-latency RDMA services over Layer 2 and Layer 3 networks.
Mellanox PeerDirect™	PeerDirect™ communication provides high efficiency RDMA access by eliminating unnecessary internal data copies between components on the PCIe bus (for example, from FPGA to network peer), and therefore significantly reduces application run time. Innova-2 Flex Open 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.
Hardware-based I/O Virtualization	Innova-2 Flex Open adapter SR-IOV technology provides dedicated adapter resources and guaranteed isolation and protection for virtual machines (VMs) within the server. I/O virtualization with Innova-2 Flex Open adapter gives data center administrators better server utilization while reducing cost, power, and cable complexity, allowing more Virtual Machines and more tenants on the same hardware.

a. This section describes hardware features and capabilities. Please refer to the driver release notes for feature availability. See Mellanox Related Documents.

Innova-2 Flex Benefits

Benefit	Description
Open platform for custom-made logic accelerations	Innova-2 Flex Open card holds a Xilinx KU15P FPGA with 520K LUTs, 70Mb of internal RAM and 1970 DSP blocks. It is a powerful tool that allows users to implement top-of-the-line acceleration engines. MNV303212A-ADLT is packed with on-board 8MB DDR4@2400MHz and connected with x8 PCIe gen4 to the host - delivering over 100Gb/s of throughput, the FPGA is primed to deliver top- notch performance and meet the most demanding offload tasks.
Single PCIe slot for advanced network adapter and custom acceleration	Innova-2 Flex Open packs Mellanox state of the art ConnectX-5 network adapter ASIC with Xilinx KU15P FPGA in one convenient HHHL adapter card, thus obviating the necessity for two separate cards for networking and offload. This frees up PCI slots for other applications, as well as saving on power.
Ease of Deployment	Innova-2 Flex Open is fully compatible with software running on other Mellanox ConnectX-5 adapters[a], and therefore allows a seamless transition from your current adapter to the Flex Open adapter.
Security	Customers can integrate their proprietary security application logic and utilize both the Innova-2 Flex's FPGA offload capabilities and advanced network features to achieve highly optimized and unique solutions. Custom hardware encryption accelerators and Innova-2 Flex modular board architecture can be combined to serve a variety of security use cases, including network-distributed Denial-of Service (DDoS) protection, software encryption and more.
Storage	Innova-2 Flex Open offers the ability to accelerate and scale different types of storage applications. For example, it can help deliver transparent data compression, decompression or de-duplication capabilities, improving overall storage utilization while not adding any additional load on the CPU.
HPC and Machine Learning	Innova-2 Flex Open helps to address the main challenges of HPC and Machine Learning by enabling higher speed and lower latency interconnect, and by performing customer-specific in-network compute. Innova-2 Flex Open delivers the ability to program customer-specific in-network compute logic on top of Innova-2 Flex Open accelerations, which include Tag Matching, RDMA and GPUDirect® with rCUDA. Communication latency can be reduced by an order of magnitude, and Machine Learning inference and training stages can be greatly accelerated by offloading specific computational elements to Innova-2 Flex Open.
Innova-2 Flex Open For Media & Entertainment	Innova-2 Flex Open adapters deliver a unique, compliant and differentiated solution for streaming applications through custom accelerations. Using Innova-2 Flex Open, multimedia application gain both scalability, by being able to handle multiple 4K/8K streams in a single host, and efficiency, with lower CPU and PCIe bandwidth utilization. Video applications can leverage Innova-2 Flex further with: • Packet pacing, natively offloaded by the ConnectX-5 within Innova 2 Flex, which can be further customized through FPGA logic • Video compression for multi-viewers applications • Seamless Protection switch to implement video redundancy

a. For specific driver availability, please refer to the Software Release Notes.

2.5.1.3 System Requirements

Operating Systems/Distributions	• RHEL • Ubuntu • SLES • FreeBSD For more information, please refer to the Innova-2™ Flex Open for Application Acceleration IPsec EN Adapter Card Release Notes. Please refer to Vivado release notes for supported Operating Systems.
Connectivity (Cables/Switches)	• Interoperable with Ethernet switches (up to 25GbE for MNV303212A-ADLT and 100GbE for MNV303611A-EDLT) • Passive copper cable with ESD protection • Powered connectors for optical and active cable support • SlimSAS OpenCAPI (Open Interface Architecture) connector 25Gb/s x8 (in MNV303212A-ADLT only)
Hardware	The Innova-2 Flex Open adapter connects to a system through a standard x8 PCIe slot. The FPGA relies on the ConnectX-5 embedded PCIe switch to connect through the PCIe fabric.

2.6 Interfaces

Each adapter card may include some or all of the interfaces described in this chapter, as well as special circuits to protect from ESD shocks to the card/server when plugging copper cables.

2.6.1 InfiniBand Interface

The network ports of the ConnectX®-5 adapter cards are compliant with the I*nfiniBand Architecture Specification, Release 1.3*. InfiniBand traffic is transmitted through the cards' SFP28 connectors in MNV303212A-ADLT, or QSFP28 connectors in MNV303611A-EDLT.

2.6.2 Ethernet Interface

The network port of the Innova-2 Flex Open adapter cards is compliant with the IEEE 802.3 Ethernet standards listed in "Features". The MNV303212A-ADLT interface provides two SFP28 ports of 10/25Gb/s, and the MNV303611A-EDLT provides two QSFP28 ports of 40/100Gb/s through the ConnectX-5 adapter. The FPGA does not have direct access to this interface.

2.6.3 PCI Express Interface

The Innova-2 Flex Open adapter cards support PCI Express 3.0/4.0 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. For pinout of the PCIe interface, please refer to the attached *Innova-2 Flex Open Interface Pinouts* excel or to the attached Verilog and XDC files.

2.6.4 LED Indications

For Innova-2 Flex Open adapter card LED specifications, please refer to “Innova-2 LEDs”.

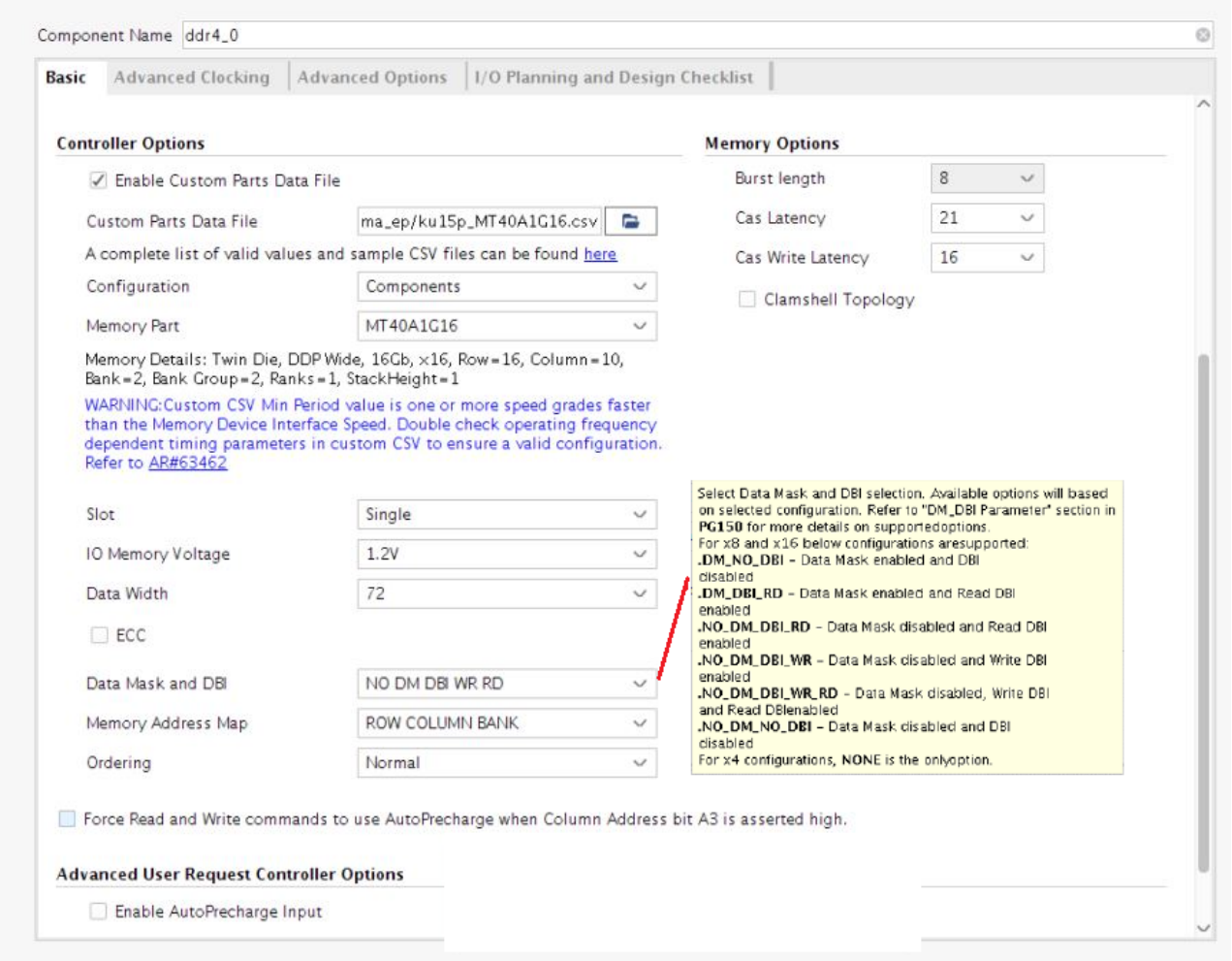
2.6.5 JTAG Interface

JTAG interface is used for direct connection to Xilinx development tools through a JTAG cable, for FPGA configuration. This interface is only intended as a hardware backup when there is no other way to configure the FPGA. See “Xilinx Programming Cable”.

2.6.6 FPGA DDR4 Interface

⚠ This section applies to MNV303212A-ADLT only.

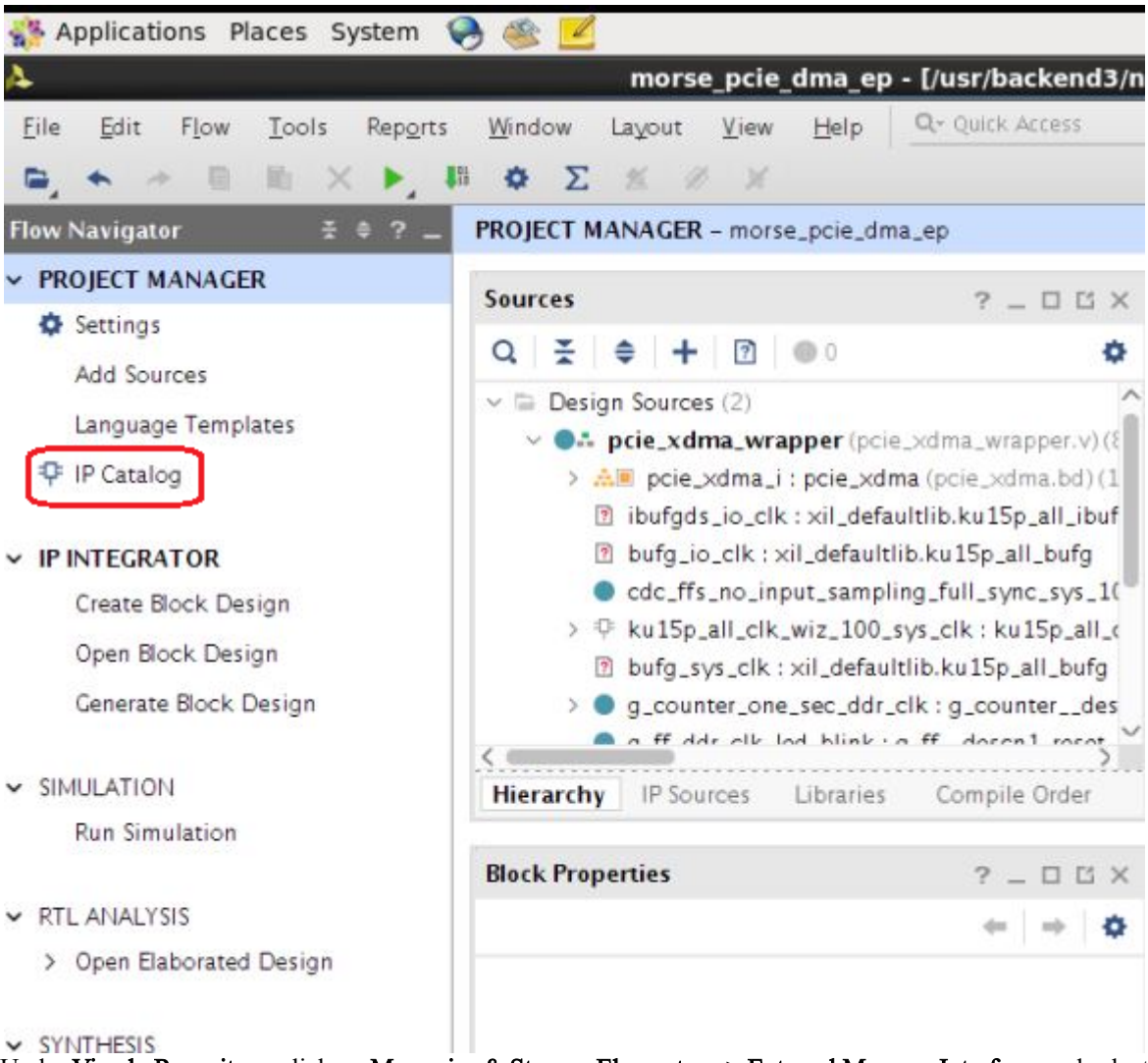
The Innova-2 Flex Open adapter supports up to 8GB DDR4 FPGA dedicated memory with a single channel 64b + 8b ECC interface. The memory is ECC protected. When working with Xilinx DDR4 controller, it is recommended to enable the DBI option by default by selecting “NO DM DBI WR RD” as shown in the following figure. For further details, please refer to DDR4 controller documentation listed in "Recommended Documents and Tools". For pinout of the DDR4 interface, please refer to the attached Innova-2 Flex Open Interface Pinouts excel or to the attached Verilog and XDC files.



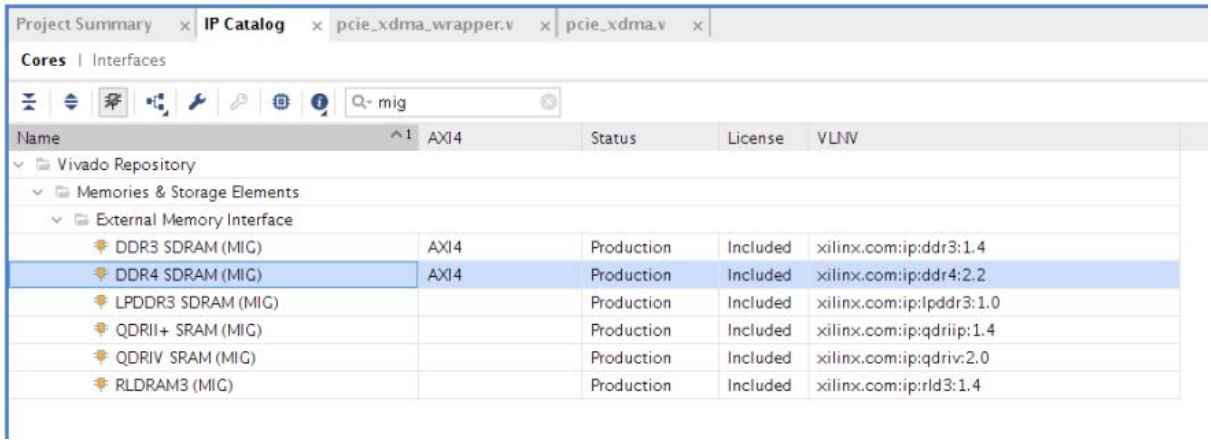
2.6.6.1 Importing CSV Inputs from Vivado

To import CSV inputs from Vivado when generating the DDR4 IP Core, please perform the following steps:

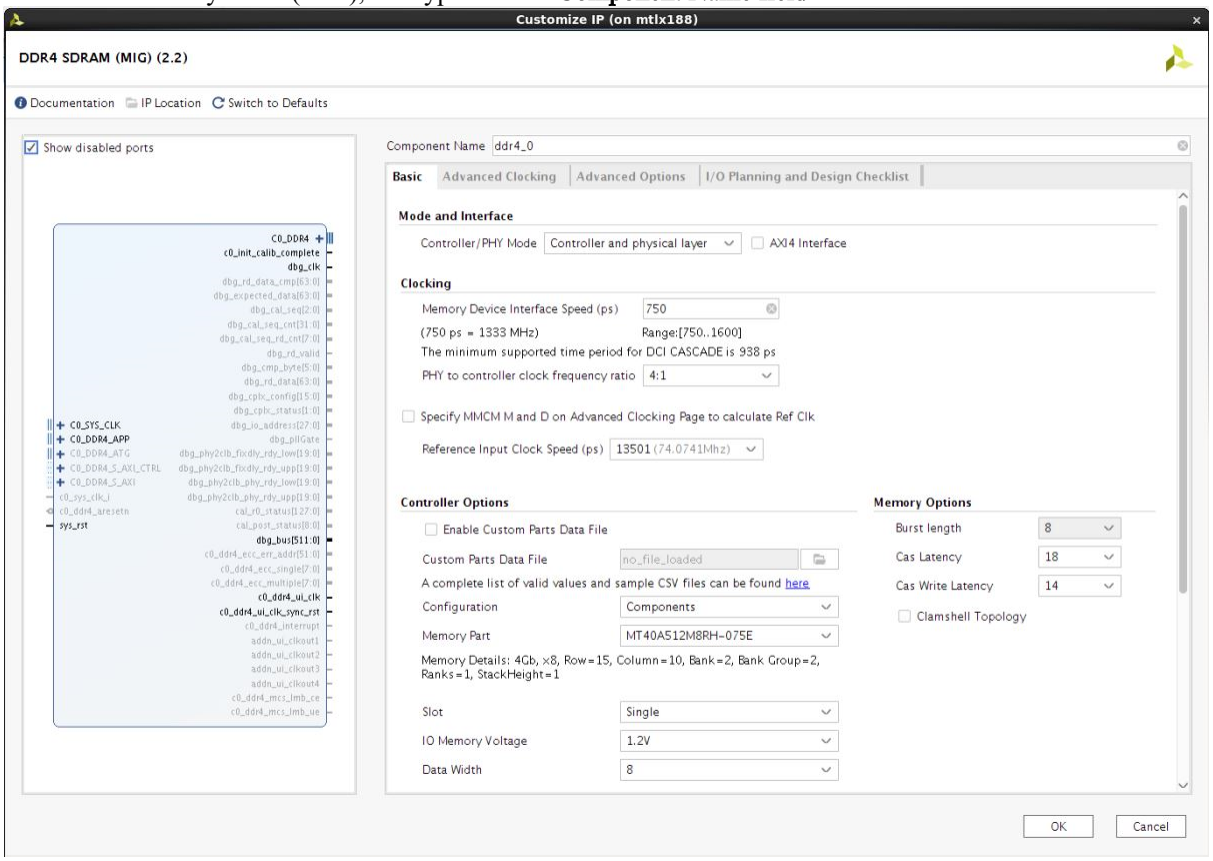
1. In the **Project Manager** navigation panel, click on **IP Catalog**.

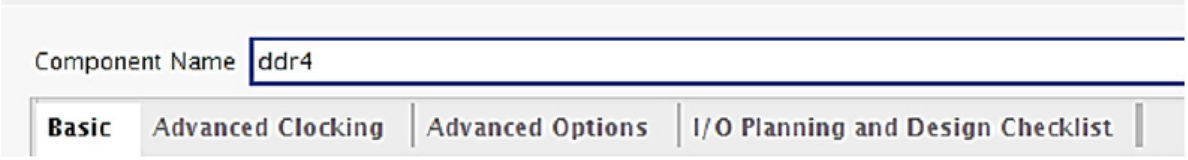


2. Under **Vivado Repository,** click on **Memories & Storage Elements** ---> **External Memory Interface**, and select DDR4 SDRAM (MIG).

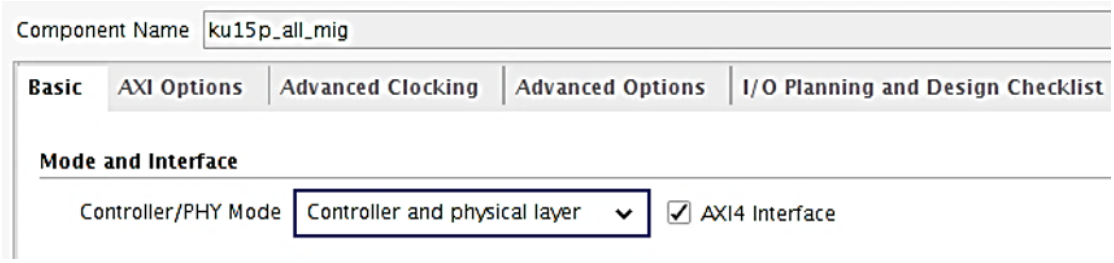


3. Choose a name for your IP (ddr4), and type it in the **Component Name field**.

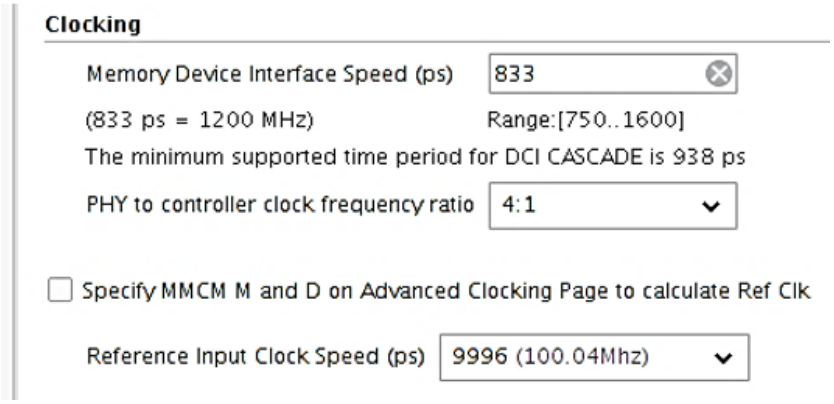




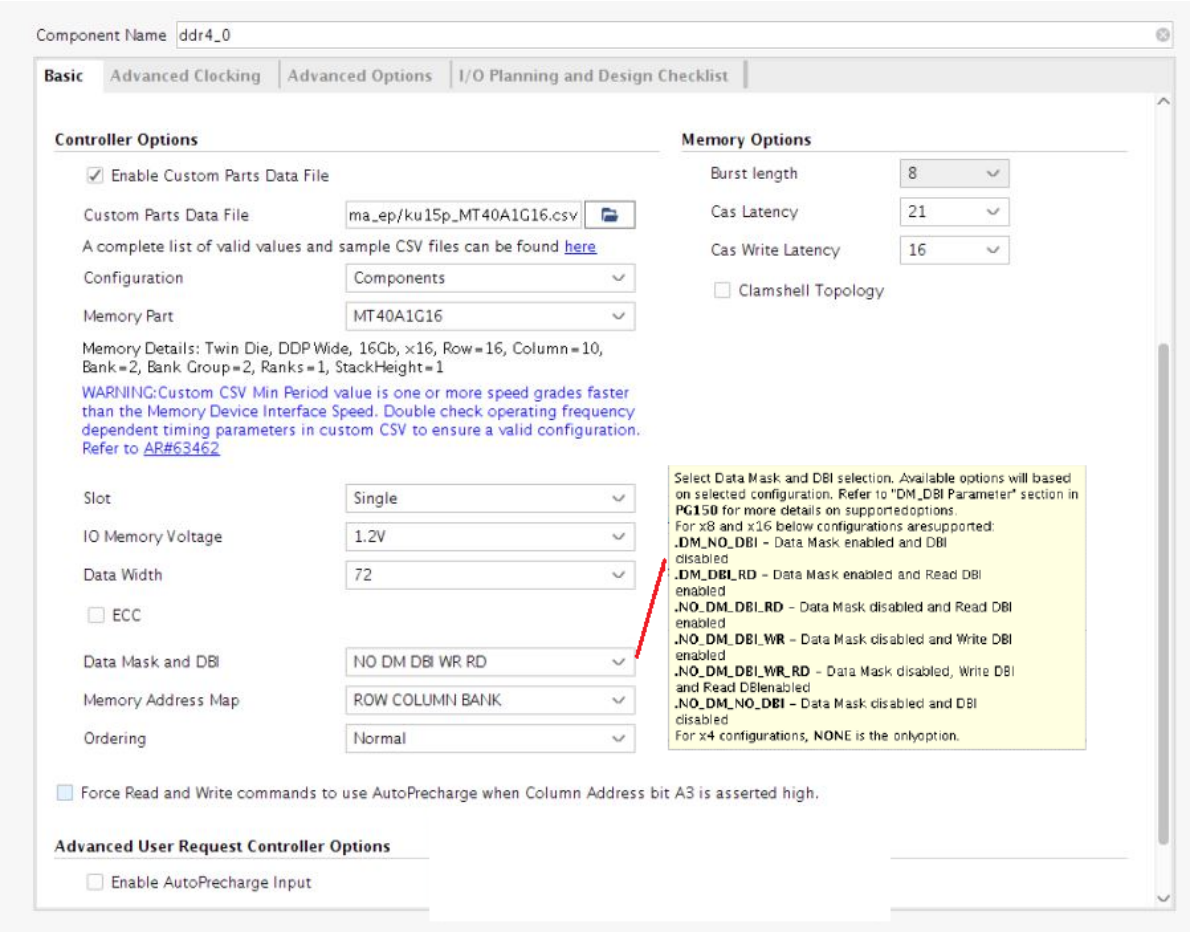
4. In the **Basic** tab under **Mode and Interface**, select **Controller and physical layer**. The AXI4 Interface can be added by checking the adjacent checkbox.



5. Under **Clocking**:
   a. In the **Memory Device Interface Speed (ps)** field, select 833 (1200MHz, to run the DDR at 2400MHz).
   b. In the **Reference Input Clock Speed (ps)** field, select 9996 (100.04MHz), which is the DDR clock on the board.



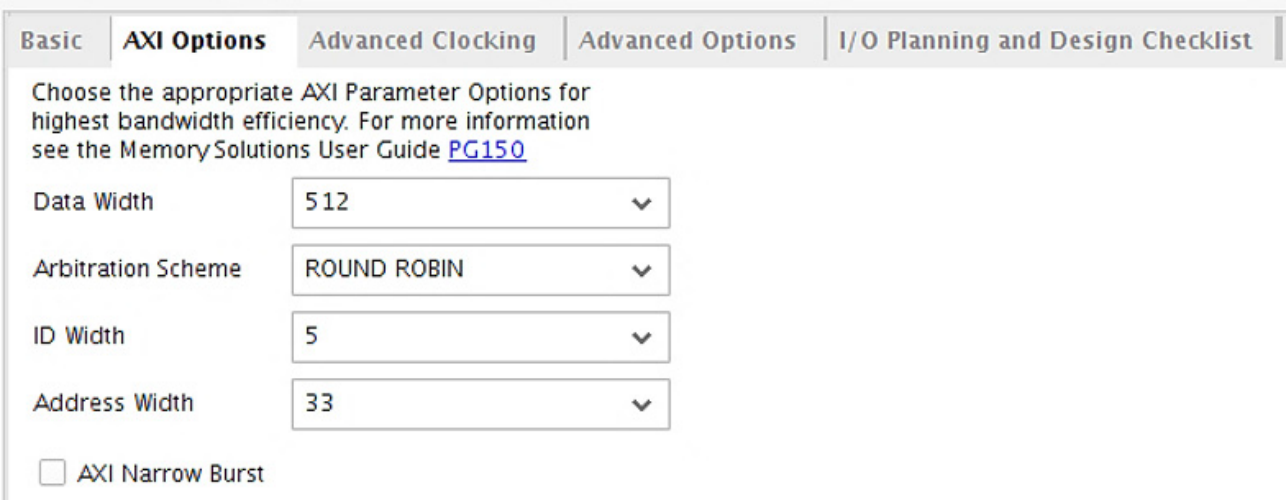
6. Under **Controller Options**, check the **Enable Custom Parts Data File** checkbox, and upload the CSV file that is attached to this document. In the **Memory Part** drop-down menu, select MT40A1G16, and in **Data Mask and DBI**, it is recommended to select "NO DM DBI WR RD" = Data Mask disabled, Write and Read DBI enabled. The other parameters are set by the CSV file.



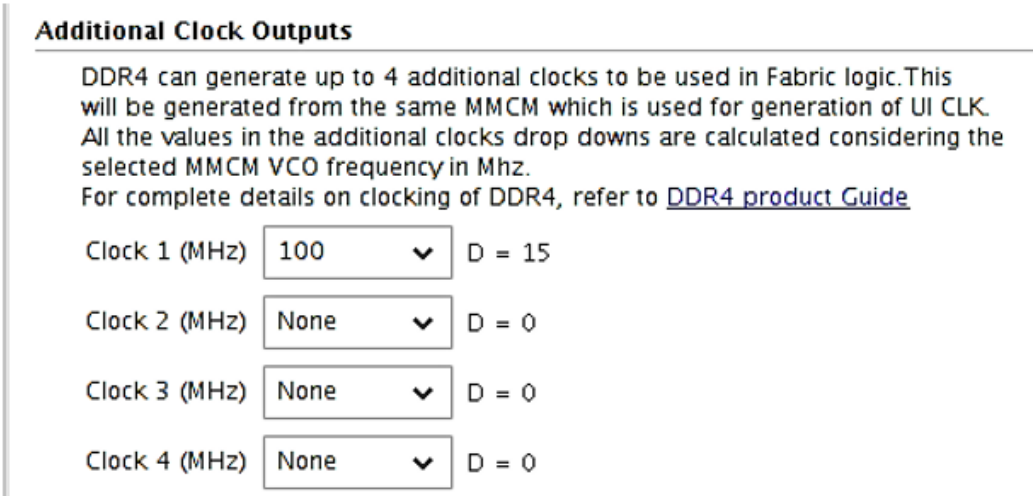
Note that the ECC checkbox is checked following the AXI4 I/F selection (the data should include the ECC).

2.6.6.1.1 Additional Recommendations

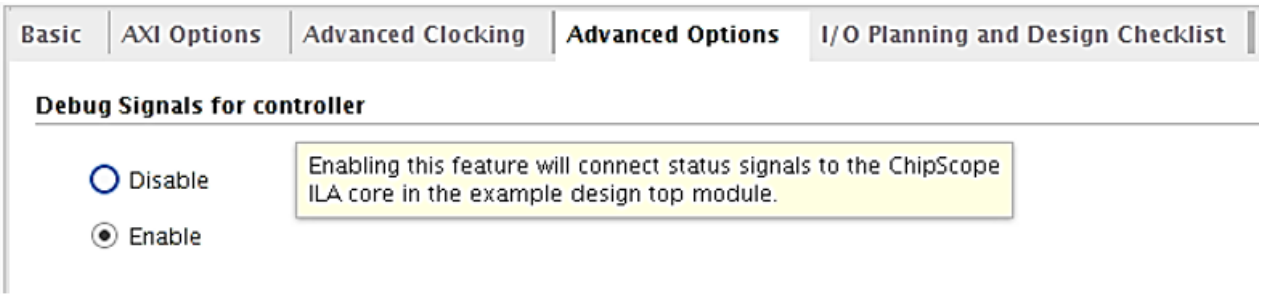
In the **AXI Options** tab, customize the fields as follows:



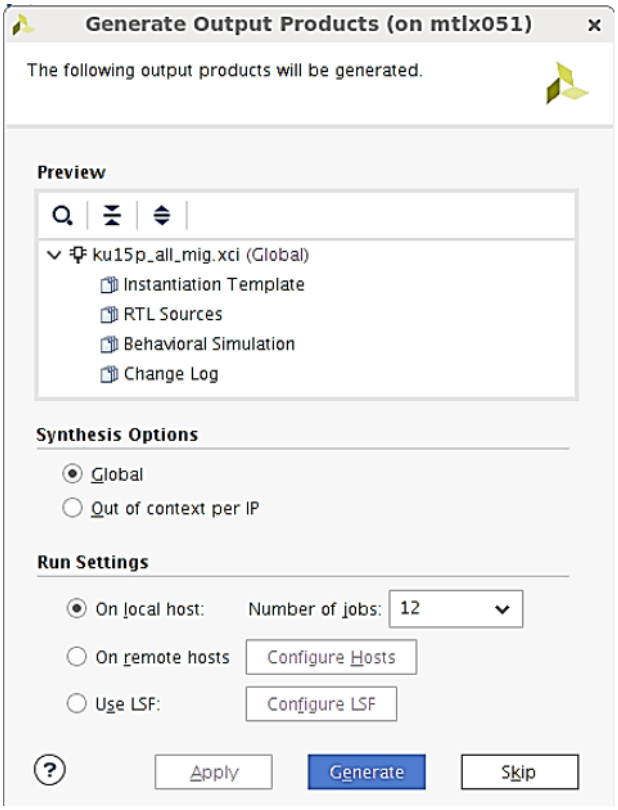
Under **Additional Clock Outputs**, customize the fields as follows:



In the **Advanced Options** tab, under **Debug Signals for Controller** select **Enable**.



Click OK and select **Global for Synthesis Options:**



2.6.7 FPGA PCIe ConnectX-5 Interface

The FPGA is connected to the ConnectX-5 embedded PCIe Gen3 x8 switch and is accessible both from the ConnectX-5 and the host over the PCIe interface. PCI reset to the FPGA is driven by the ConnectX-5 adapter.

2.6.8 FPGA OpenCAPI Interface

⚠ This section applies to MNV303212A-ADLT only.

The FPGA is connected to a SlimSAS OpenCAPI 25Gb/s x8 connector. For pinout of the OpenCAPI interface, please refer to the attached *Innova-2 Flex Open Interface Pinouts* excel or to the attached Verilog and XDC files.

2.6.9 FPGA QSPI Flash Interface

The FPGA is connected to two x4 on-board MICRON TECHNOLOGY (MT25QU512ABB8E12-0SIT) QSPI flashes.

FPGA QSPI Flash Interface

Signal	Total Delay
D0	0.35135232
D1	0.355097342

Signal	Total Delay
D2	0.35009532
D3	0.353372756
CS0#	0.350112638
CCLK	0.355390305
D4	0.268570741
D5	0.263691389
D6	0.266098594
D7	0.265581939
CS1#	0.266061071
CCLK	0.268427867

2.6.10 On-Board Clock

On-Board Clock

Clock	Pin Number
DDR CLK 100.04MHz LVDS (in MNV303212A-ADLT only)	AP24 (P), AP25 (N)
EMCCLK 150MHz LVCMOS	AM14
SYSCLK 100MHz LVDS	AR14 (P), AT14 (N)
PCIe internal link 100MHz LVDS	AB27 (P), AB28 (N) [GTY128 CLK0]
OpenCAPI 156.25MHz LVDS (in MNV303212A-ADLT only)	T27 (P), T28 (N) [GTY131 CLK0] P27 (P), P28 (N) [GTY132 CLK0]

2.7 Hardware Installation

2.7.1 System Requirements

2.7.1.1 Hardware

A system with a standard x8 PCIe slot.

2.7.1.2 Operating Systems/Distributions

Please refer to “Operating Systems/Distributions”.

2.7.2 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.

2.7.3 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.

2.7.4 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 “Card Installation Instructions".

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

2.7.4.1 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.

2.7.4.2 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.

2.7.5 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.

2.7.6 Cables and Modules

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

2.7.6.1 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 "Network LEDs Operation".
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 "Network LEDs Operation".
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.
    a. 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.

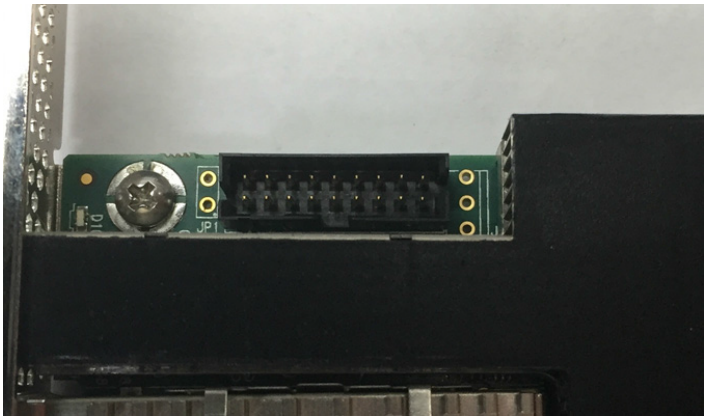
2.7.6.2 Xilinx Programming Cable

> ⚠ Xilinx programming cable is not provided by Mellanox - please contact your Xilinx representative for details.

The Xilinx programming cable is a USB to JTAG probe required to directly access the FPGA HW. Please refer to http://www.xilinx.com/products/boards-and-kits/hw-usb-ii-g.html.

It is possible to connect the flat cable that comes with the Xilinx platform cable USB II to the on- board JTAG connector.

The on-board JTAG connector is 9 pins, dual row, 2mm pitch. The flat cable has a female connector 7 pins, dual row, 2mm pitch. The cable should be connected to the on-board connector according to the key in the shroud, (indicated with the yellow rectangle) leaving 2 unconnected pins on each side.



The other end of the Xilinx Platform Cable USB II should be connected to the user's computer.

2.7.7 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 MT28800 Family [ConnectX-5]
```

2.8 Specifications

2.8.1 MNV303212A-ADLT Specifications

MNV303212A-ADLT Air Flow Specifications

Parameter	Description			
Marketing Description	Innova-2 Flex Open for Application Acceleration, dual-port SFP28, 25GbE, KU15P, 8GB, No Crypto, PCI4.0 x8, HHHL, active heat sink, tall bracket			
Physical	**Size:** 167.65mm X 68.90mm			
	Connector: Dual SFP28 (Copper and optical)			
Protocol Support	**Ethernet:** 25GBASE-R, 1000BASE-CX, 1000BASEKX, 10GBASE-SR, 10GBASE-LR,10GBASE-ER,10GBASE-CR, 10GBASE-KR			
	Data Rate: 10/25 Gb/s - Ethernet			
	PCI Express Gen3/4: SERDES @ 16GT/s, 8 lanes (2.0 and 1.1 compatible)			
Power and Environmental	**Voltage:** 12V			
	Power Consumption	**Use Case**	**Cable**	**PCIe Card Power**
	Ambient Temperature: 25°C			
	Typical Power	No FPGA logic	Passive cables	18W
		FPGA consumes 20W	Passive cables	41W
	Maximum Power	FPGA consumes 35W	Passive cables	59W
		FPGA consumes 35W	1.5W Active cables	63W
	Maximum power the card supports	75W[a]		

Parameter	Description
	Temperature: Operational 0°C to 55°Cb Non-operational -40°C to 70°C
	Humidity: 90% relative humidity[c]
	Air Flow: See "MNV303212A-ADLT Air Flow Specifications".
Regulatory	**Safety:** IEC/EN 60950-1:2006 ETSI EN 300 019-2-2 IEC 60068-2- 64, 29, 32
	RoHS: RoHS-R6
Cable Support	To obtain the list of supported Mellanox cables for your adapter, please refer to the Cables Reference Table.

a. In order to attain the specified consumption by the board, it is required to place it inside a x16 PCI slot.

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

c. For both operational and non-operational states.

MNV303212A-ADLT Air Flow Specifications

Ta = 35°C			Ta = 45°C			Ta = 55°C		
FPGA Power (W)	**Cable Type**	**Airflow Requirement (LFM)**	**FPGA Power (W)**	**Cable Type**	**Airflow Requirement (LFM)**	**FPGA Power (W)**	**Cable Type**	**Airflow Requirement (LFM)**
42	Active	300	42	Active	650	42	Active	Not supported
	Passive			Passive			Passive	
40	Active	300	40	Active	550	40	Active	Not supported
	Passive			Passive			Passive	
35	Active	300	35	Active	400	35	Active	800

Ta = 35°C			Ta = 45°C			Ta = 55°C		
	Passive			Passive			Passive	
30	Active	300	30	Active	300	30	Active	700
	Passive			Passive			Passive	600
25	Active	300	25	Active	300	25	Active	600
	Passive			Passive			Passive	400

2.8.2 MNV303611A-EDLT Specifications

MNV303611A-EDLT Specifications

Parameter	Description
Marketing Description	Innova-2 Flex Open VPI, dual-port QSFP28, EDR / 100GbE, KU15P, No memory, No Crypto, PCI4.0 x8, HHHL, passive heat sink, tall bracket
Physical	**Size:** 161.3mm X 68.90mm
	Connector: Dual QSFP28 (Copper and optical)
Protocol Support	**VPI:** 25GBASE-R, 1000BASE-CX, 1000BASEKX, 10GBASE-SR, 10GBASE-LR,10GBASE-ER,10GBASE-CR, 10GBASE-KR
	Data Rate: EDR / 100GbE
	PCI Express Gen3/4: SERDES @ 16GT/s, 8 lanes (2.0 and 1.1 compatible)

Parameter	Description			
Power and Environmental	**Voltage:** 12V			
	Power Consumption	**Use Case**	**Cable**	**PCIe Card Power**
	Ambient Temperature: 25°C			
	Typical Power	No FPGA logic	Passive cables	22W
		FPGA consumes 20W	Passive cables	45W
	Maximum Power	FPGA consumes 35W	Passive cables	63W
		FPGA consumes 35W	3.5W Active cables	71W
	Maximum power the card supports	75W[a]		
	Temperature: Operational 0°C to 55°C[b] Non-operational -40°C to 70°C			
	Humidity: 90% relative humidity[c]			
	Air Flow: See "MNV303611A-EDLT Air Flow Specifications".			
Regulatory	**Safety:** IEC/EN 60950-1:2006 ETSI EN 300 019-2-2 IEC 60068-2- 64, 29, 32			
	RoHS: RoHS-R6			
Cable Support	To obtain the list of supported Mellanox cables for your adapter, please refer to the Cables Reference Table.			

a. In order to attain the specified consumption by the board, it is required to place it inside a x16 PCI slot.

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

c. For both operational and non-operational states.

MNV303611A-EDLT Air Flow Specifications

Ta = 35°C			Ta = 45°C			Ta = 55°C		
FPGA Power (W)	Cable Type	Airflow Requirement (LFM)	FPGA Power (W)	Cable Type	Airflow Requirement (LFM)	FPGA Power (W)	Cable Type	Airflow Requirement (LFM)
35	Active	600	35	Active	800	35	Active	Not supported
	Passive			Passive			Passive	
30	Active	500	30	Active	700	30	Active	Not supported
	Passive			Passive			Passive	
25	Active	550	25	Active	700	25	Active	Not supported
	Passive	500		Passive	650		Passive	850
20	Active	400	20	Active	600	20	Active	Not supported
	Passive			Passive			Passive	750

2.8.3 Innova-2 LEDs

2.8.3.1 Network LEDs Operation

LEDs D1 and D2 provide information on ConnectX-5 Port 1 and Port 2 accordingly. The following table defines the LEDs behavior.

Network LEDs Operation

LED	Function
Off	A link has not been established
Blinking Amber[a]	6 Hz blinking Amber indicates a problem with the link

LED	Function
Solid Green	Indicates a valid link with no active traffic
Blinking Green	Indicates a valid logical link with active traffic

a. 1 Hz Blinking Amber occurs due to running a beacon command for locating the adapter card.

2.8.3.2 FPGA LEDs

FPGA LEDs Indications

LED Name	FPGA Pinout	Function
D18	A6	For User Image: Programmable For Innova-2 Flex Image: DDR clock blink
D19	B6	For User Image: Programmable For Innova-2 Flex Image: DDR HW BIST status (on - success, off - failure/in progress)

2.8.3.3 General LEDs

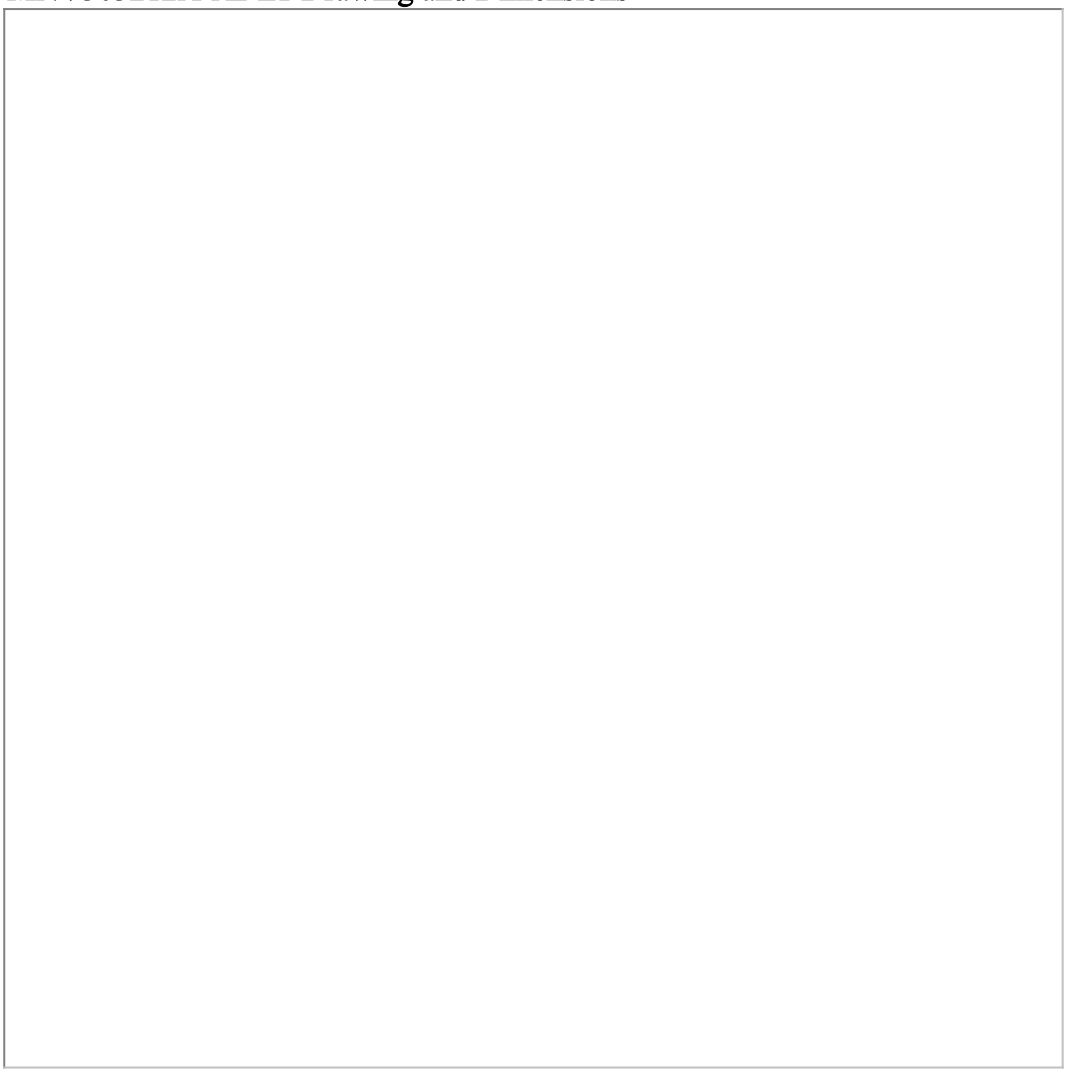
General LEDs Indications

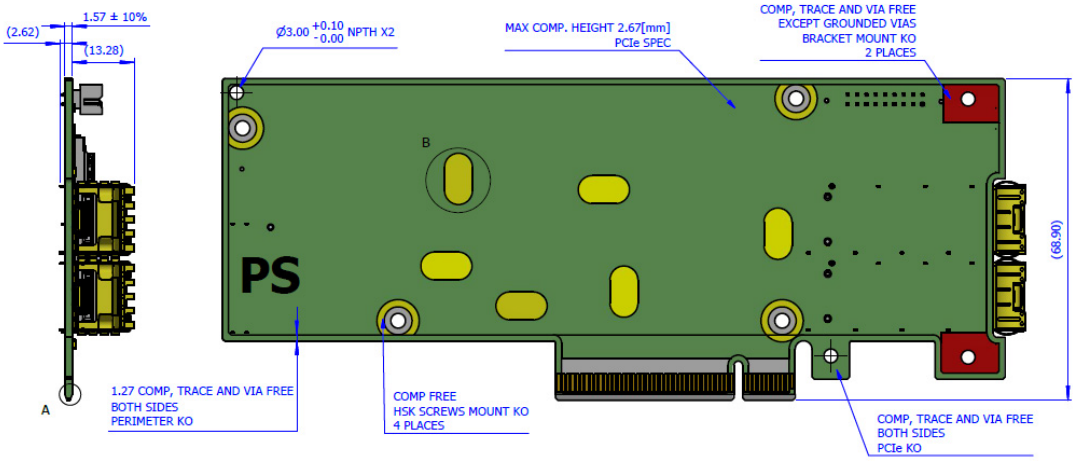
LED Name	Function
D10	On when ConnectX power is stable
D15	Red when PROG_B or DONE are low. Green when Done is high

2.8.4 Board Mechanical Drawing and Dimensions

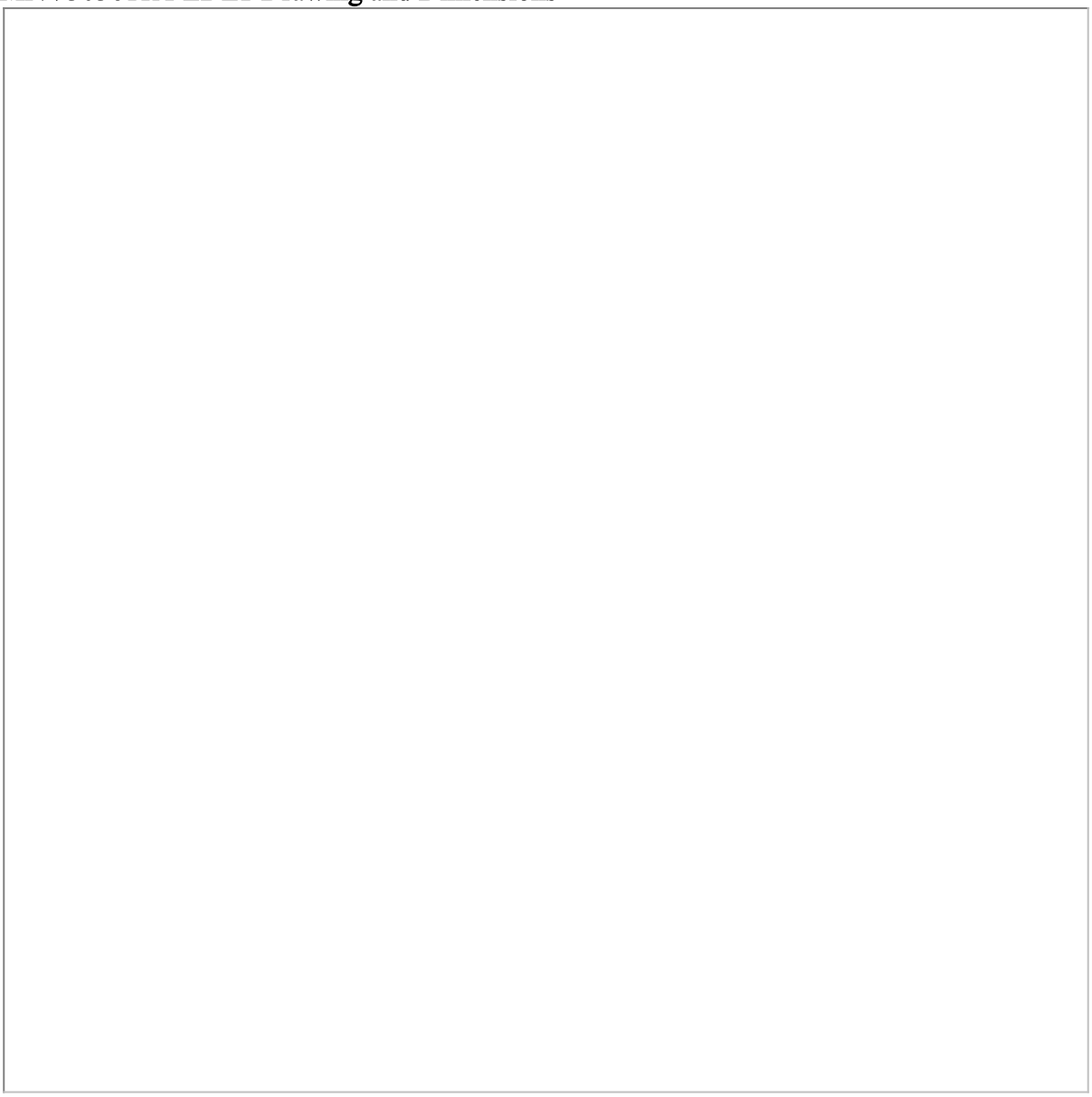
All dimensions are in millimeters.
All the mechanical tolerances are +/- 0.1mm.

MNV303212A-ADLT Drawing and Dimensions





MNV303611A-EDLT Drawing and Dimensions



2.9 Innova-2 Flex Open Card Driver

2.9.1 Linux Driver Installation

A MLNX_OFED package that supports the Innova-2 Flex adapter can be obtained through http://www.mellanox.com. This chapter describes how to install and test the Mellanox OFED for Linux package on a single host machine with Innova-2 Flex Open adapter hardware installed.

MLNX_OFED should be installed before installing the following content from the bundle.

2.9.1.1 Hardware and Software Requirements

Software and Hardware Requirements

Requirements	Description
Platforms	A server platform with an adapter card based on one of the following Mellanox Technologies' adapter devices: • MT4119 ConnectX®-5 (VPI, IB, EN) (firmware: fw-ConnectX5) For the list of supported architecture platforms, please refer to the Mellanox OFED Release Notes file in http://www.mellanox.com/page/products_dyn?product_family=26.
Required Disk Space for Installation	1GB
Device ID	The list of Mellanox Technologies PCI Device IDs can be found in the PCI ID repository at http://pci-ids.ucw.cz/read/PC/15b3.
Operating System	Linux operating system. For the list of supported operating system distributions and kernels, please refer to the Mellanox OFED Release Notes file in http://www.mellanox.com/page/products_dyn?product_family=26.
Installer Privileges	The installation requires administrator privileges on the target machine.

2.9.2 Downloading Mellanox OFED

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 MT28800 Family [ConnectX-5]
   Subsystem: Mellanox Technologies Device 0024
   ```

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 Innova-2 Flex adapter can be obtained through Mellanox support.
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
   ```

2.9.3 Installing Mellanox OFED

2.9.3.1 Installation Script

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 InfiniBand and Ethernet network adapters and automatically upgrades the firmware.
  Note: The firmware will not be updated if you run the install script with the '--without-fw-update' option..

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 `"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]
    ```

    where:

`[--make-iso]`	Create MLNX_OFED ISO image
`[--make-tgz]`	Create MLNX_OFED tarball (default)

[-t|--tmpdir <local work dir>]	Local work directory
[--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]	Enable verbose mode
[-n|--name]	Name of the package to be created
[-y|--yes]	Answer "yes" to all questions
[--force]	Force removing packages that depend on MLNX_OFED

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.

For further information, see the help file.

where:

-c|--config <packages config_file>	Example of the configuration file can be found under docs
-n|--net <network config_file>	Example of the network configuration file can be found under docs

-k|--kernel-version <kernel version>	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/Dis- able of SRI-OV in a non-volatile configuration through uEFI and/or tool will override this flag.
--add-kernel-support	Add kernel support (Run mlnx_add_kernel_sup- port.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>	Do not install package
--with-fabric-collector	Install fabric-collector package

Installation Procedure:

1. Login 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, 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:*
> ---
> *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*

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.

2.9.3.2 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: ...
	  Please contact your hardware vendor for help on firmware updates.

2.9.3.3 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
```

2.9.3.4 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:*

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

```
blacklist mlx4_core
blacklist mlx4_en
blacklist mlx5_core
```

2.9.3.5 mlnxofedinstall Return Codes

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

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

2.9.3.6 Uninstalling MLNX_OFED

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

2.9.3.6.1 Updating the Device Manually

In case you ran the `mlnxofedinstall` script with the `'--without-fw-update'` option or you are using an OEM card and now you wish to (manually) update firmware on your adapter card(s), you need to perform the steps below.

The following steps are also appropriate in case you wish to burn newer firmware that you have obtained from Mellanox Support.

1. Get the device's PSID.

```
mst start
mst status
flint -d <mst device> q | grep PSID
PSID
```

2. Get the firmware BIN file provided by Mellanox for the adapter card.
3. Burn the firmware, using mlxup, Mellanox Update and Query Utility - http://www.mellanox.com/page/mlxup_firmware_tool.

```
mlxup -i <fw_file.bin>
```

4. Reboot your machine after the firmware burning is completed.

2.9.3.7 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.

2.9.3.8 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 that:
- The 'mokutil' package is installed on your system, and
- The system is booted in UEFI mode.

1. Download the x.509 public key.

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

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
```

3. Reboot the system.

The pending MOK key enrollment request will be noticed by shim.efi and it will launch Mok- Manager.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`

2.9.3.9 Removing the 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:

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 -
   taint- ing 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. Thus, it is much easier to miss this mes- sage. You will not see it on repeated tests where you unload and reload a kernel module until you reboot. There is no way to eliminate this message.
2. Update the initramfs on RHEL systems with the stripped modules.

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

2.10 Using the Innova-2 Flex Open Bundle

Mellanox provides the Innova-2 Flex Open Bundle, which includes the Innova-2 Flex Open FPGA Image and Innova-2 Flex Open application (for detailed content of the bundle see "Innova-2 Flex Open Bundle Content". The bundle allows the user to:

- Perform FPGA and board diagnostics - see "Diagnostic Capabilities".
- Burn a user image - see "Burning Capabilities".
- Determine whether the Innova-2 Flex Open or the User Image will run on the FPGA - See "Identifying the System State".

Important Notes:

- If the Innova-2 Flex image is not burned onto the card upon shipment, the user must burn the Innova-2 Flex Image provided by Mellanox via JTAG. See "Burning the FPGA Image Through the JTAG".
- The user can use Innova-2 Flex Open application to burn the user image and for FPGA interface diagnostics only if the FPGA Innova-2 Flex Image is loaded. Otherwise, the application will display a reduced menu which only allows the user to load the Innova-2 Flex Image.
- When using the Innova-2 Flex image, the user must work with the Innova-2 Flex Open application only.
- The user cannot overwrite the Innova-2 Flex Image.

2.10.1 Installing the Innova-2 Flex Open Application

1. Copy the provided Innova-2 Flex Open application package to a local temporary directory (i.e: /tmp).
2. Enter the temporary directory:

```
cd /tmp
```

3. Extract the Innova-2 Flex Open application installation package:

```
tar zxvf Innova_2_Flex_Open_xx_xx.tar.gz
```

4. Navigate to the application folder (/app):

```
cd Innova_2_Flex_Open_xx_xx/app
```

5. Navigate to the driver folder (/driver):

```
cd Innova_2_Flex_Open_xx_xx/driver
```

6. Compile:

```
make clean; make
```

2.10.2 Running the Innova-2 Flex Open Application

1. Run the application:

```
sudo ./innova2_flex_app -v
```

2. Detect the system's state as instructed in "Identifying the System State". If the system state is as shown in Example A, proceed to Step 3, if the system state is as shown in Example B or Example C, proceed to Step 4.
3. Perform the following:
   a. Exit the application using the Exit option.
   b. Navigate to the driver folder

```
cd ../driver/
```

   c. Install the driver.

```
make clean make
sudo insmod mlx_fpga_bope.ko
```

d. Return to "Installing the Innova-2 Flex Open Application".

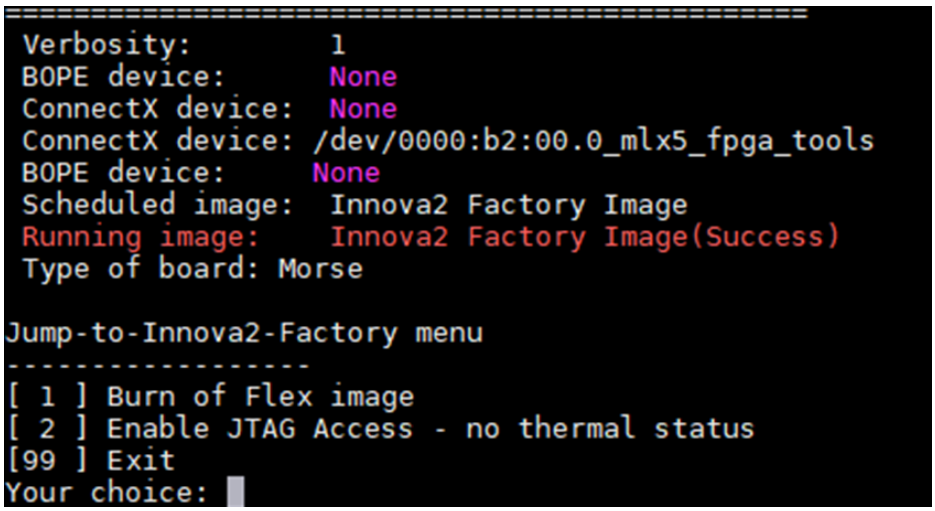
4. You will now see the "Jump to Innova-2 Flex Image" menu with one single item: "Set Innova-2 Flex Image active".
   a. See "Innova-2 Flex Open Bundle Content" in order to change to Innova-2 Flex image.

Return to "Installing the Innova-2 Flex Open Application".

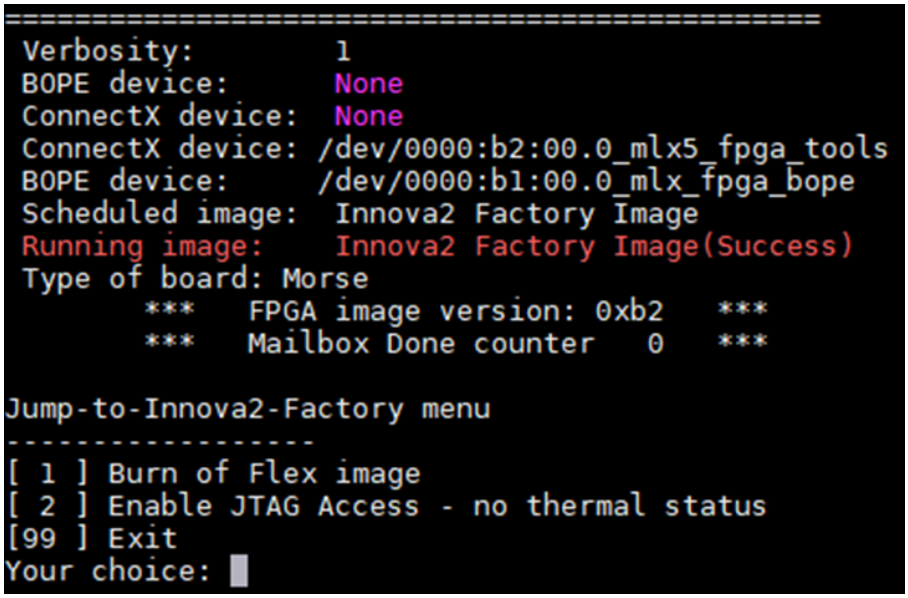
2.10.2.1 Identifying the System State

Examples:

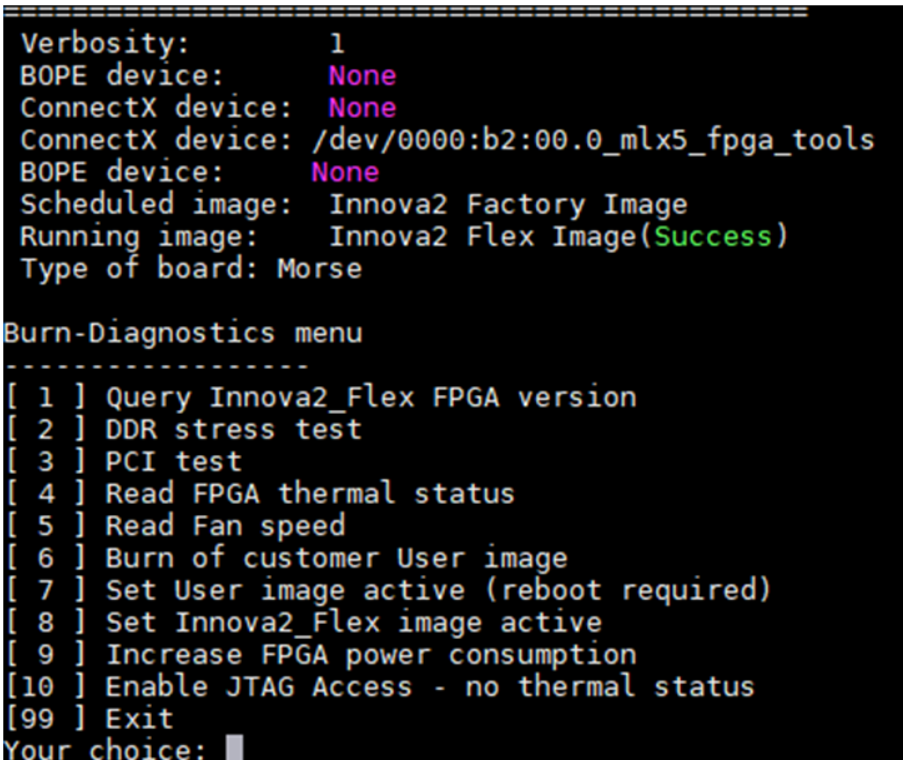
1. In the figure below, the Innova-2 Factory Image is running, and there is no Innova-2 FPGA PCI driver loaded.



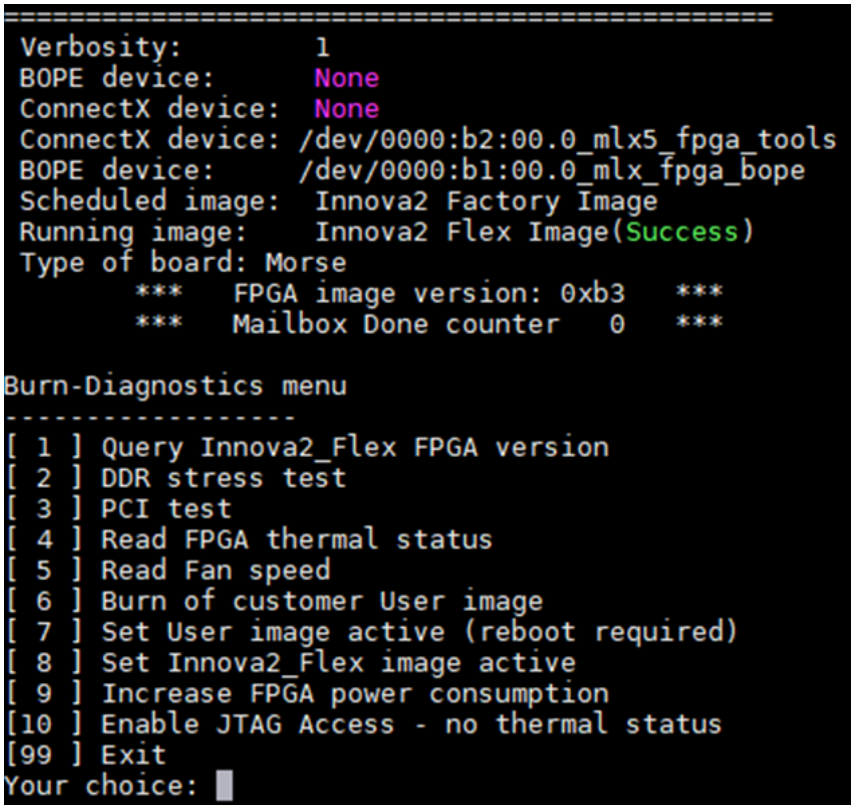
2. In the figure below, the Innova-2 Factory Image is running, and the Innova-2 FPGA PCI driver is loaded. In the BOPE device info, the FPGA image version is displayed.



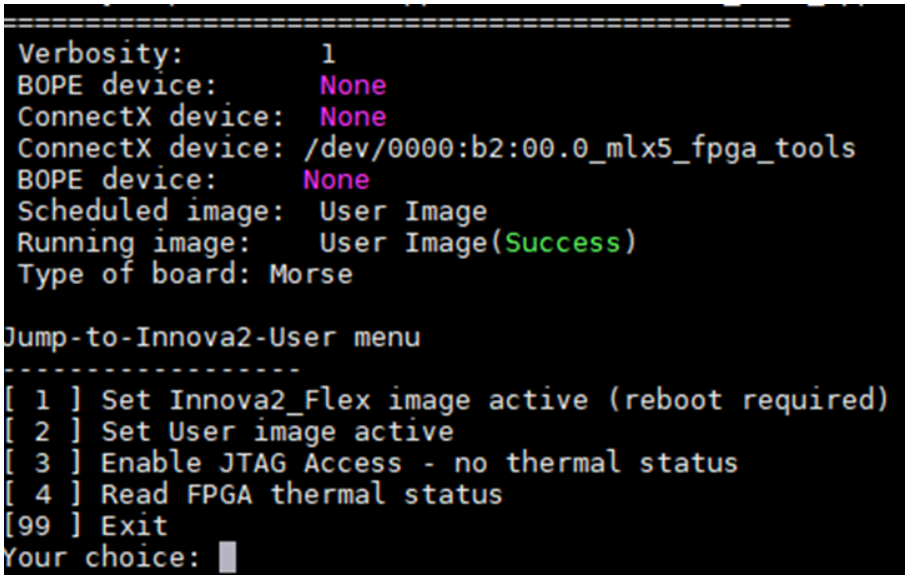
3. In the figure below, the Innova-2 Flex Image is running, and the Innova-2 FPGA driver is not loaded.



4. In the figure below, the Innova-2 Flex Image is running, and the Innova-2 FPGA PCI driver is loaded.



5. In the figure below, the User Image is running (the Innova-2 FPGA PCI driver is not required).



2.10.3 Diagnostic Capabilities

- Query FPGA version - reads the Innova-2 Flex Image FPGA version and presents it to the user.
- DDR Stress BIST (Built-in Self Test):
  - Cyclic test LFSR (Linear Feedback Shift Register) address - data, which can be either 1s, 0s or pseudo-random, is written to pseudo-random address until every DDR address is written to. The test then reads back the sequence and compares to the expected sequence. A new seed is used for the pseudo-random address sequence in every new cycle. The test continues until terminated by the user.
  - Cyclic test incremental address - data, which can be either 1s, 0s or pseudo-random, is written to incremental address until every DDR address is written to. The test then reads back the sequence and compares to the expected sequence. The test continues until terminated by the user.
- Single test - writes data all over the DDR space and validates that data is written properly.
- PCI test - tests the PCIe interface between the host and the FPGA.
- FPGA Thermal status - reads the FPGA temperature and presents it to the user (in Celsius).
- Fan speed - reads the Fan speed and presents it to the user (in RPM).
- Increase FPGA power consumption - reads the FPGA power level, and presents it to the user. The user can set one of the FPGA power levels: 1,2...10.

2.10.3.1 Running the Diagnostics

⚠ Note: The Innova-2 Flex Open bundle must be installed prior to the diagnostics

1. To view the Innova-2 Flex Open application options, run:

```
./innova2_flex_app -help
```

2. Run the application:

```
./innova2_flex_app -v
```

3. Check which image is active by reading the Running Image field. If the running image is Innova-2 Flex, continue to Step 4. Otherwise, switch to Innova-2 Flex Image, as instructed in "Switching between Images".

```
==============================================
 Verbosity:         1
 BOPE device:       None
 ConnectX device:   None
 ConnectX device: /dev/0000:b2:00.0_mlx5_fpga_tools
 BOPE device:     /dev/0000:b1:00.0_mlx_fpga_bope
 Scheduled image:   Innova2 Factory Image
 Running image:     Innova2 Flex Image(Success)
 Type of board: Morse
        ***   FPGA image version: 0xb3   ***
        ***   Mailbox Done counter   0   ***

Burn-Diagnostics menu
-----------------
[ 1 ] Query Innova2_Flex FPGA version
[ 2 ] DDR stress test
[ 3 ] PCI test
[ 4 ] Read FPGA thermal status
[ 5 ] Read Fan speed
[ 6 ] Burn of customer User image
[ 7 ] Set User image active (reboot required)
[ 8 ] Set Innova2_Flex image active
[ 9 ] Increase FPGA power consumption
[10 ] Enable JTAG Access - no thermal status
[99 ] Exit
Your choice:
```

4. Run the required tests or query, by choosing one of the menu options. For DDR Stress Test -
   - After selecting DDR Stress Test, the following screen should be shown:

```
Your choice: 2

DDR stress test
-----------------
[ 1 ] Cyclic test LFSR address
[ 2 ] Cyclic test incremental address
[ 3 ] Single test
[99 ] Back
Your choice: 1

Test is running. Press Enter to interrupt

Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Waiting for test to finish...
Test finished with status "success"
DDR stress test
```

   - Choose the type of the required test. Description of the different tests can be found in "Diagnostic Capabilities".
   - In case the cyclic test was selected:

- Press Enter to stop the test.
- Wait for the test to finalize
- Once finished, the test will print its result (success/fail)

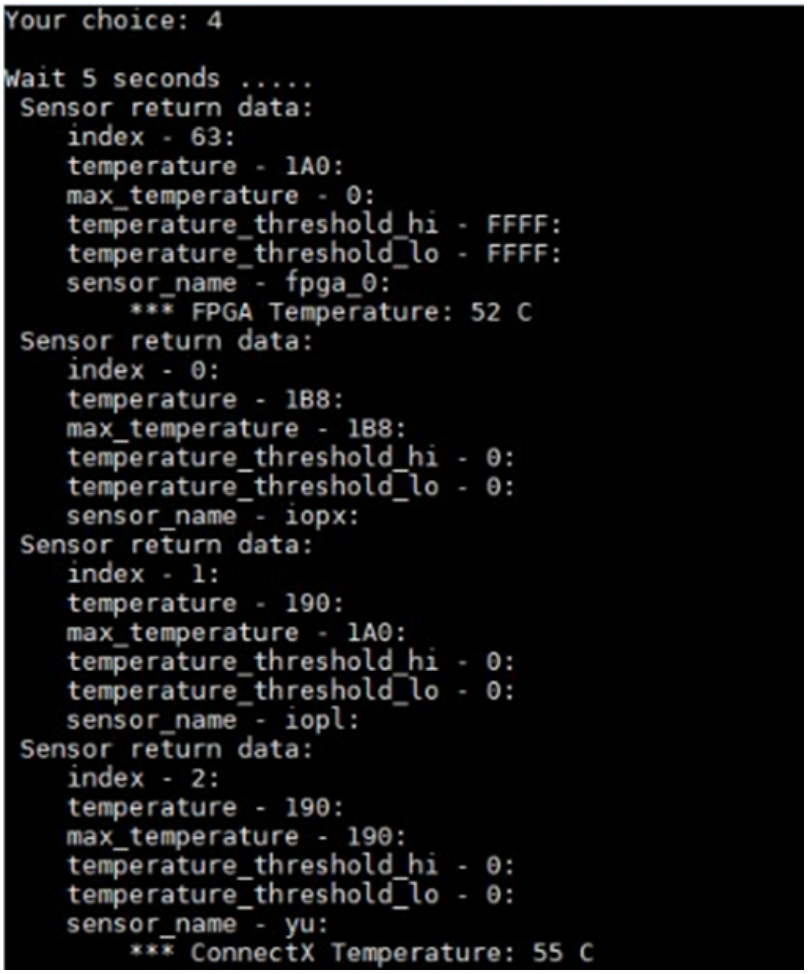
5. Set FPGA power diagnostic.
   - In the Burn-Diagnostics menu, select Option 9 - “Increase FPGA power consumption”.

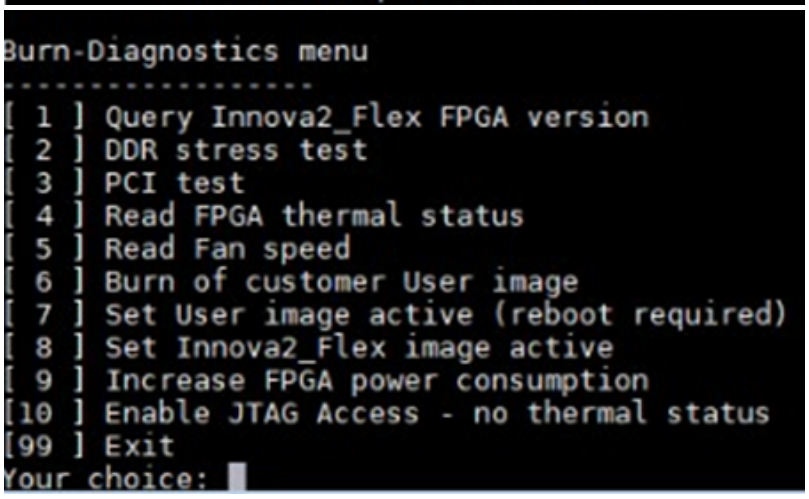
```
==============================================
 Verbosity:         3
 BOPE device:       None
 ConnectX device:   None
 ConnectX device: /dev/0000:b2:00.0_mlx5_fpga_tools
 BOPE device:     /dev/0000:b1:00.0_mlx_fpga_bope
 Scheduled image:  Innova2 Factory Image
 Running image:    Innova2 Flex Image(Success)
 Type of board: Morse
        ***   FPGA image version: 0xb3   ***
        ***   Mailbox Done counter   0   ***

Burn-Diagnostics menu
-----------------
[ 1 ] Query Innova2_Flex FPGA version
[ 2 ] DDR stress test
[ 3 ] PCI test
[ 4 ] Read FPGA thermal status
[ 5 ] Read Fan speed
[ 6 ] Burn of customer User image
[ 7 ] Set User image active (reboot required)
[ 8 ] Set Innova2_Flex image active
[ 9 ] Increase FPGA power consumption
[10 ] Enable JTAG Access - no thermal status
[99 ] Exit
Your choice: 9
```

   - Upon selecting Option 9, the following menu should appear:

```
Your choice: 9


Increase FPGA Power (current power level 4):
-----------------
[ 0 ] Set FPGA Power level 0 - only base power without incre
[ 1 ] Set FPGA Power level 1
[ 2 ] Set FPGA Power level 2
[ 3 ] Set FPGA Power level 3
[ 4 ] Set FPGA Power level 4
[ 5 ] Set FPGA Power level 5
[ 6 ] Set FPGA Power level 6
[ 7 ] Set FPGA Power level 7
[ 8 ] Set FPGA Power level 8
[ 9 ] Set FPGA Power level 9
[10 ] Set FPGA Power level 10
[99 ] Exit
Your choice: 4
```

   - Choose the one of FPGA power levels. For example, Option 4, “Set FPGA Power level 4”. The following screen should appear:





2.10.3.2 JTAG Access to the FPGA

The Innova-2 Flex Open Application access to FPGA may interfere with the on-board JTAG access to FPGA. If external JTAG cable access is required (for example, see "Burning the FPGA Image Through the JTAG") the application must first enable this.

1. Select the "Enable JTAG Access - no thermal status" in the "Jump-to Innova2-Use" menu, or in the "Burn-Diagnostics" menu.

```
Your choice: 10


Disable JTAG Access menu
------------------
[ 1 ] Disable JTAG Access - enable thermal status
[99 ] Exit
Your choice:
```

The FPGA is now disconnected, and you can burn the FPGA image through the JTAG.

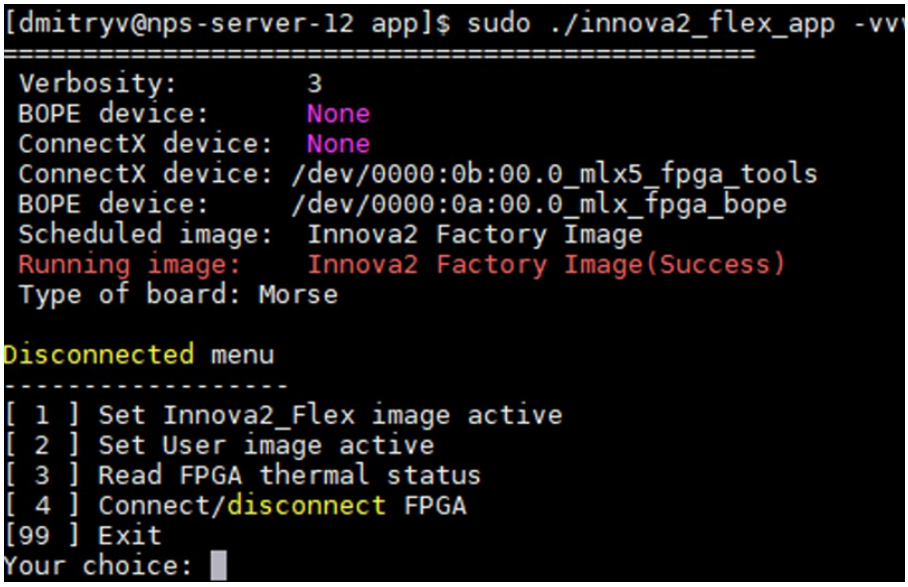
2. After burning, you can connect the FPGA by selecting "Disable JTAG Access - enable Thermal status", and continue working with the Innova-2 Flex Open application.

```
Your choice: 1

        ***   FPGA image version: 0xb3   ***
        ***   Mailbox Done counter   0   ***

Burn-Diagnostics menu
-------------------
[ 1 ] Query Innova2_Flex FPGA version
[ 2 ] DDR stress test
[ 3 ] PCI test
[ 4 ] Read FPGA thermal status
[ 5 ] Read Fan speed
[ 6 ] Burn of customer User image
[ 7 ] Set User image active (reboot required)
[ 8 ] Set Innova2_Flex image active
[ 9 ] Increase FPGA power consumption
[10 ] Enable JTAG Access - no thermal status
[99 ] Exit
Your choice:
```

3. Disconnect the FPGA, and check the status.

```
Connect/disconnect FPGA
-------------------
[ 1 ] Query current status
[ 2 ] Connect FPGA
[ 3 ] Disconnect FPGA
[99 ] Back
Your choice: 1

         *** FPGA is disconnected ***

Connect/disconnect FPGA
-------------------
[ 1 ] Query current status
[ 2 ] Connect FPGA
[ 3 ] Disconnect FPGA
[99 ] Back
Your choice:
```

The “Disconnect” state is saved when the Innova-2 Flex Open application is closed. After running the Innova-2 Flex Open application, the system is in “Disconnected” state.



4. Close the application and start burning Flex through the JTAG.
5. After burning and following the power cycle, open the application and select the “Connect FPGA” option in the "Disconnected" menu.

2.10.4 Burning Capabilities

Note: The Innova-2 Flex Open bundle must be installed prior to the image burning.

The Innova-2 Flex Open application and image allow the user to burn an FPGA User Image to the flash through the PCI, and to switch back and forth between the images that run on the FPGA - the Innova-2 Flex Image and the User Image. After burning and activating the User Image, the user can return to the Innova-2 Flex Image (for diagnostics or PCI-burn).

2.10.4.1 Running the Burning Flows

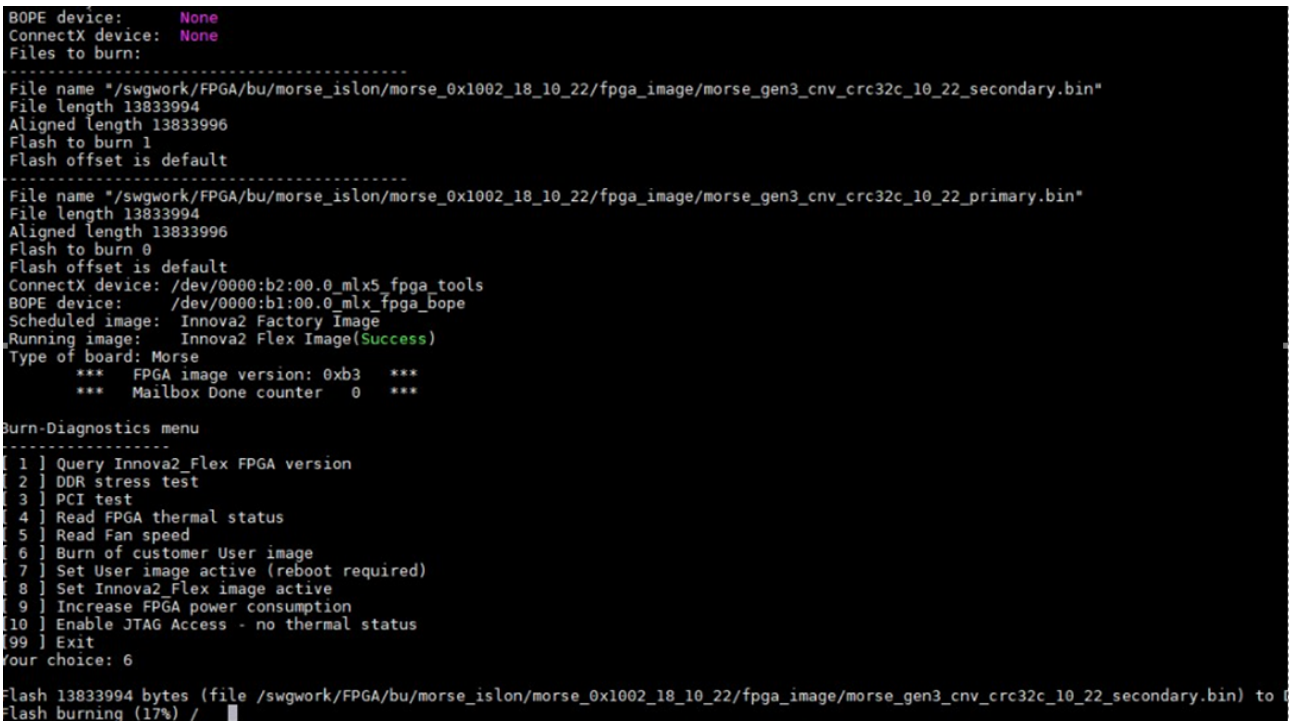
1. To view the Innova-2 Flex Open application options:

```
innova2_flex_app -help
```

2. To run the application, see the below example:
   - To program two files at flash 0 and flash_1:

```
/innova2_flex_app -b first_file_name.bin,0 second_file_name.bin,1 -v
```

3. When the application opens, choose “Burn of customer User image” to burn the image.
4. During the burning process the application will show progress and a message will indicate once burn is done.
5. For activating the burned image, see “Switching between Images”.



2.10.5 Switching between Images

The user can determine which image from the Flash will run on the FPGA.

> ⚠ In case the FPGA User Image is corrupted and/or does not exist but the user chose to set it as active, the booting action will fail and the FPGA will appear un-programmed after reboot.

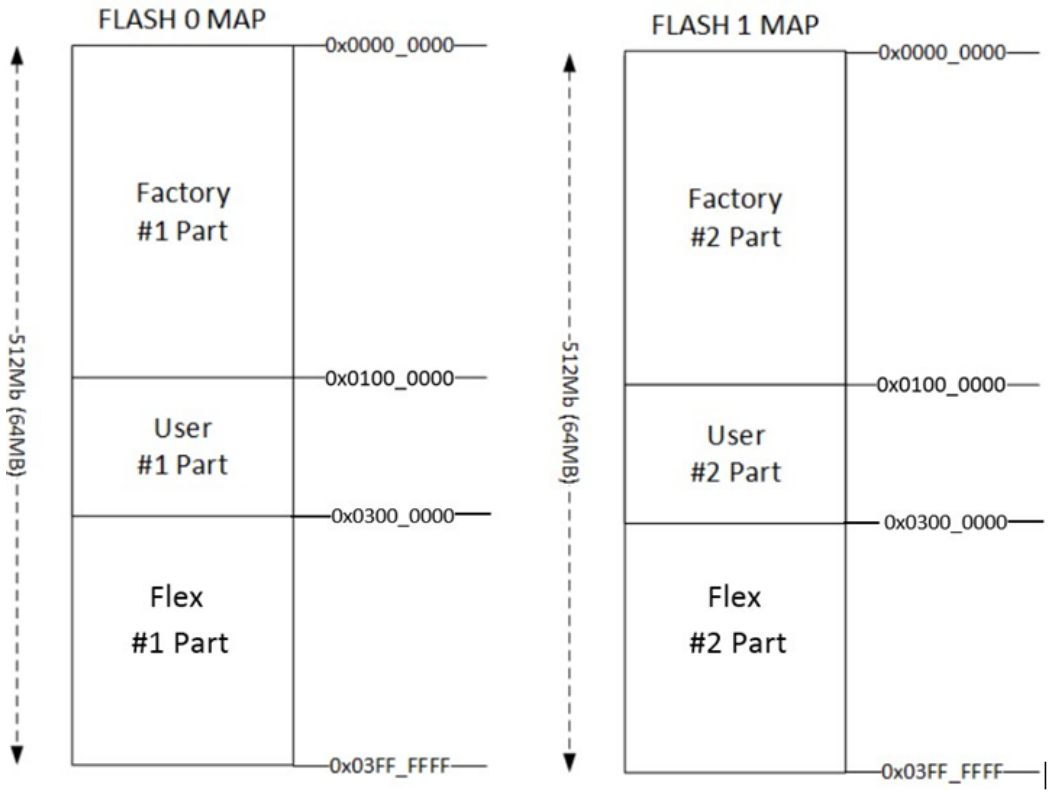
1. Open the application:

```
./innova2_flex_app -v
```

2. In case the Innova-2 Flex Image is running, select the “Switch to Innova-2 Flex Image” option in order to move to the Innova-2 Flex Image. In case User Image or no image is running, choose the “Switch into Innova-2 Flex Image” option in order to move to Innova-2 Flex Image.

2.10.5.1 Flash Format

The below figure shows the flash format when both the Innova-2 Flex Image and User Image are burned on the flash (i.e. when the Innova-2 Flex Image is not overwritten).



2.10.6 Burning the FPGA Image through the JTAG

⚠ Burning an FPGA image via JTAG should be done from an external server connection. Dropping a PCI link during FPGA burning can cause the server to stall and self-reboot, thereby causing the burning process to fail. When JTAG is connected to an external server, even if the target server stalls, the burning process will complete successfully.

To burn the Flash via Vivado Lab Edition 2017.3:

⚠ If the Innova-2 Flex Open Application is running, it may interfere with the JTAG connectivity. See "JTAG Access to the FPGA" before you connect a JTAG cable.

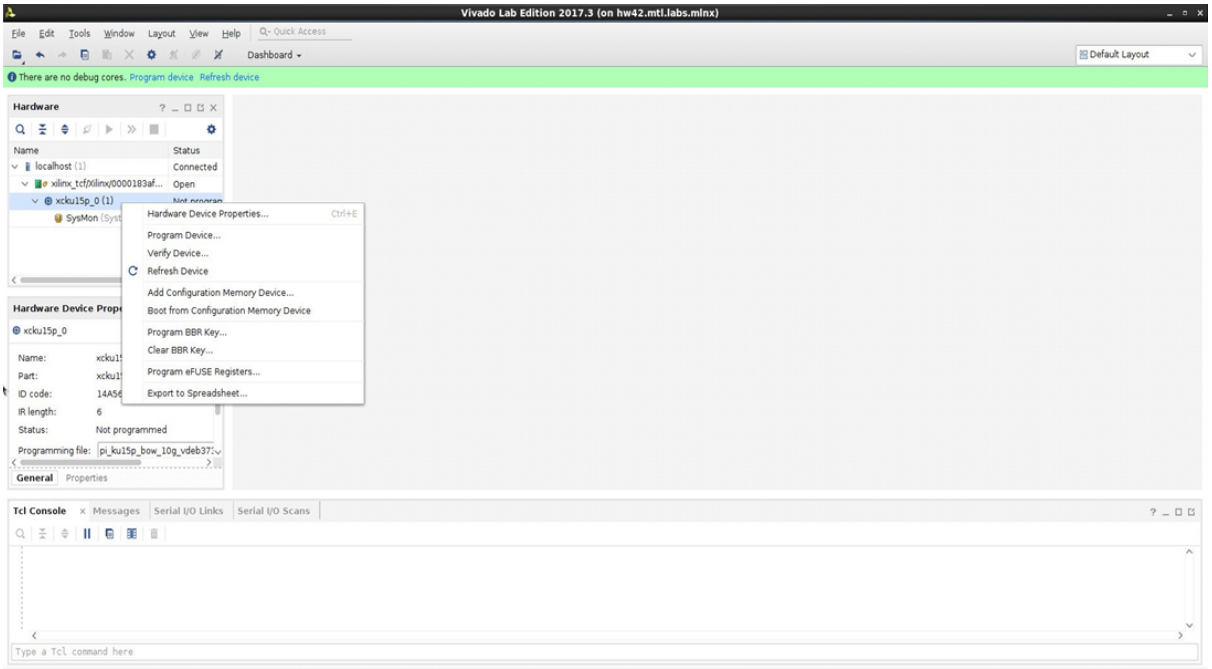
1. Go to the following link, and install Vivado Lab Edition 2017.3: - https://www.xilinx.com/support/download/index.html/content/xilinx/en/downloadNav/vivado-design-tools/2017-3.html.
2. Connect the JTAG cable. For information on JTAG connection, see Xilinx Programming Cable.

⚠ Xilinx programming cable is not provided by Mellanox. Please contact your Xilinx representative for details.

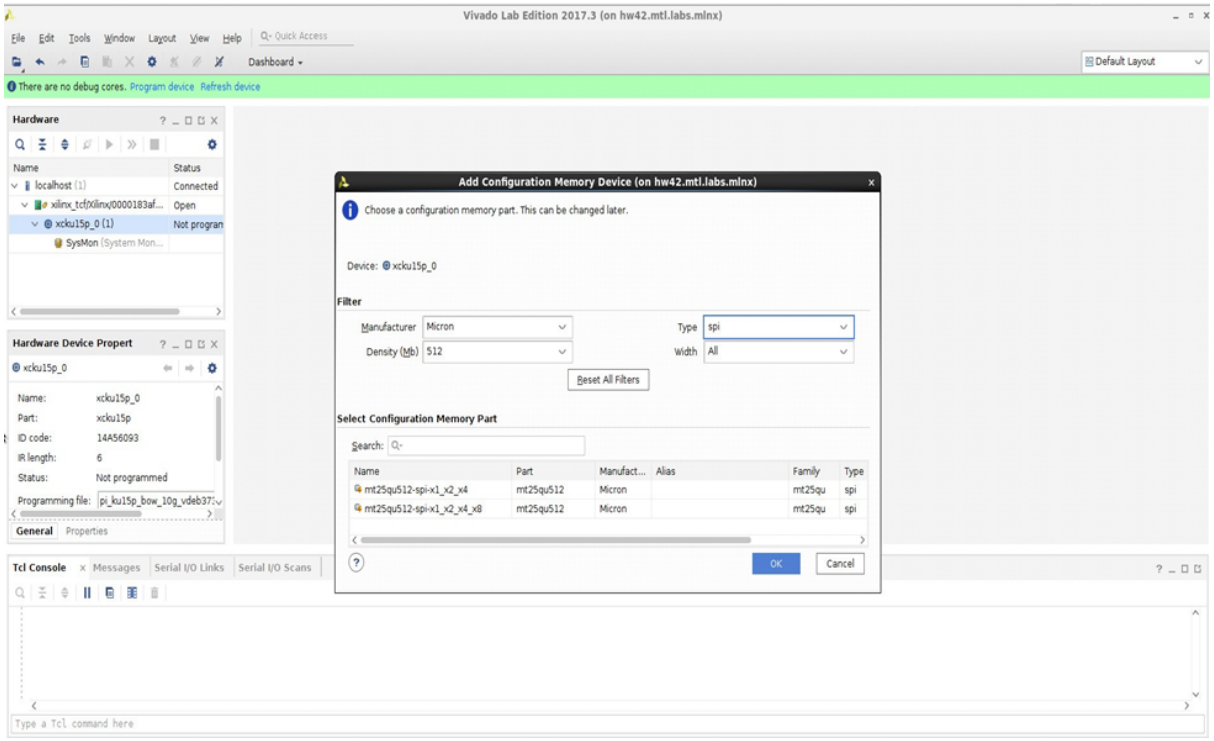
To add a configuration memory device:

1. Open Vivado Lab Edition 2017.3.
2. Click "Open Target" and then "Auto Connect".
3. Click on the "Open Hardware Manager."
4. On the left corner (Hardware) - make sure you can see 'xcku15p' device.

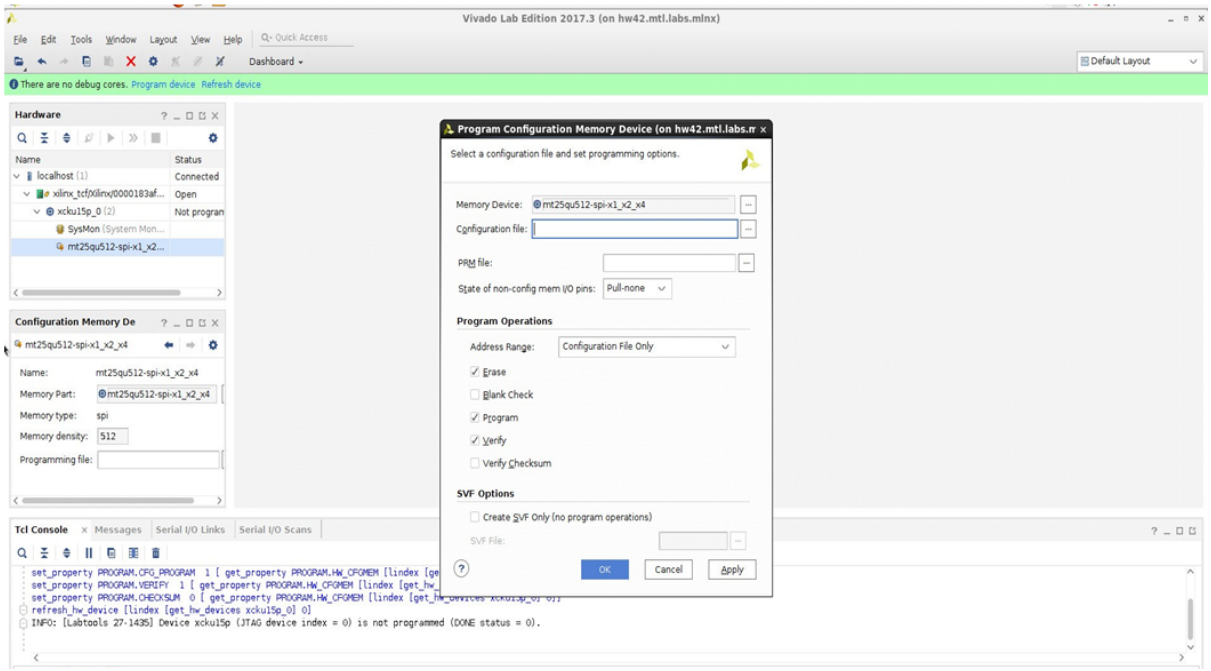
5. Right click on the 'xcku15p' and 'Add Configuration Memory Device'.



6. Select the appropriate device (x1_x2_x4 or x1_x2_x4_x8).



7. Choose the BIN file. See “Innova-2 Flex Open Bundle Content”.



8. Burn the flash.
9. Run server full power cycle.

2.10.7 Reloading the Innova-2 User Image

> ⚠ The reloading procedure is under the customer’s responsibility. Errors can lead to a failure of the Innova-2 network card, and even to a failure of the server. The link between the FPGA and the PCI switch should be disabled before reloading the FPGA User Image, and enabled after reloading. Two consoles are used: the console where the Innova2_Flex_app is running, and the console for PCI commands.

2.10.7.1 Procedure Scenario

In order to reload the Innova-2 User Image, start the first console, and launch the Innova2_flex-_app, as instructed in “Running the Innova-2 Flex Open Application”. The second terminal should be open, as it is used for disabling/enabling the PCI link between the FPGA and the PCI switch.

In the second console:

Step 1. Find the PCI switch in the list of PCI devices

Step 2. Save the FPGA device PCI configuration registers

Step 3. Disable the PCI link between the FPGA and the PCI switch

In the first console:

Step 4. Reload the FPGA User Image

Move back to the second console:

Step 5. Enable PCI link between FPGA and PCI switch (second console)

Step 6. Copy back the PCI configuration registers (second console)

2.10.7.2 Finding the PCI Switch

Get the device location on the PCI bus by running the below lspci command, and locating lines with the "Mellanox" string, as instructed in “Identify the Card in Your System”.

> ⚠ lspci -v | grep Mellanox

Example:

- ConnectX5 Device has 3e:00.0 PCI bdf ([bus:][device.][func]) address
- Innova-2 FPGA PCI driver (BOPE-driver) has 3d:00.0 PCI bdf address
- The PCI switch has 3c:08.0 bdf address

```
3b:00.0 PCI bridge: Mellanox Technologies MT28800 Family [ConnectX-5 PCIe Bridge]
3c:08.0 PCI bridge: Mellanox Technologies MT28800 Family [ConnectX-5 PCIe Bridge]
3c:10.0 PCI bridge: Mellanox Technologies MT28800 Family [ConnectX-5 PCIe Bridge]
3d:00.0 Class 2000: Mellanox Technologies Innova-2 Flex Burn image
3e:00.0 Ethernet controller: Mellanox Technologies MT27800 Family [ConnectX-5]
3e:00.1 Ethernet controller: Mellanox Technologies MT27800 Family [ConnectX-5]
```

2.10.7.3 Disabling/Enabling the PCI Link between the FPGA and the PCI Switch

Disabling example:

```
sudo setpci -s 3c:08.0 0x70.w=0x50
```

Enabling example:

```
sudo setpci -s 3c:08.0 0x70.w=0x40
```

Examples legend:

- 3c:08.0 is the bdf address of the PCI switch
- 0x70 is the capability register base address + link control register (0x60+0x10)
- 0x40 is the last value in this register
- 0x50 is the last value in this register + disable link bit (0x40 + 0x10)

2.10.7.4 Reloading the User Image

In the Burn-Diagnostic Menu, select Option 5 - "Reload User image”. The following message will appear:

```
"Reload feature may hang up your station if PCI is not disabled!!! Do you want to run
this feature? [y/]"
```

Make sure that the PCI link is disabled, as instructed in "Disabling/Enabling the PCI Link between the FPGA and the PCI Switch".

```
[dmitryv@nps-server-37 app]$ sudo ./innova2_flex_app -vvv
Device /dev/*_mlx_fpga_bope not found
================================================
 Verbosity:         3
 BOPE device:       None
 ConnectX device:   None
 ConnectX device: /dev/0000:3e:00.0_mlx5_fpga_tools
 BOPE device:      None
 Scheduled image:   User Image
 Running image:     User Image(Success)

Jump-to-Innova2-User menu
-----------------
[ 1 ] Set Innova2_Flex image active (reboot required)
[ 2 ] Set User image active
[ 3 ] Enable JTAG Access - no thermal status
[ 4 ] Read FPGA thermal status
[ 5 ] Reload User image
[99 ] Exit
Your choice: 5

 Reload feature may hang up your station if PCI is not disabled!!!
  Do you want to run this feature?  [y/n] y

Jump-to-Innova2-User menu
-----------------
[ 1 ] Set Innova2_Flex image active (reboot required)
[ 2 ] Set User image active
[ 3 ] Enable JTAG Access - no thermal status
[ 4 ] Read FPGA thermal status
[ 5 ] Reload User image
[99 ] Exit
Your choice:
```

Enable the PCI link on the second console, as instructed in "Disabling/Enabling the PCI Link between the FPGA and the PCI Switch".

2.10.8 Innova-2 Flex Open Bundle Content

⚠ MLNX_OFED should be installed before installing the following content from the bundle. See "Innova-2 Flex Open Card Driver".

Mellanox provides an Innova-2 Flex Open bundle - Innova_Flex_Open, which includes the following directories:

- FPGA_image/ - This folder contains two subfolders: Factory_image/ and Flex_image/, and additional auxiliary files. See "Using the Innova-2 Flex Open Bundle".
- Factory_image/ - This folder contains the Innova-2 Factory Image binary files:
  - innova2_factory_open_fpga_primary_file.bin - for flash 0
  - innova2_factory_open_fpga_secondary_file.bin - for flash 1
- Flex_image/
  - innova2_flex_open_fpga_primary_file.bin - for flash 0
  - innova2_flex_open_fpga_secondary_file.bin - for flash 1
- FW/ - This folder contains two subfolders: Morse_FW/ for Innova-2 Flex 25G cards, and MorseQ_FW/ for Innova-2 Flex VPI cards.
  - Morse_FW/
    - fw-ConnectX5-rel-(version number).bin - This folder contains Innova-2 Flex 25G firmware bin files for ConnectX-5.
  - MorseQ_FW/

  - fw-ConnectX5-rel-(version number).bin - This folder contains Innova-2 Flex VPI firmware bin files for ConnectX-5.
- Driver/ - Contains a direct PCI device driver for Innova-2 Flex
- App/ - Contains an Innova-2 Flex user application
- Update_script -contains the "install.sh" script for automatic update of previously installed 18_7 , 18_10 and 18_11 bundles
- main.conf, 18_7.conf, 18_10.conf and 18_11.conf versions - configuration files for the update script

2.10.8.1 FPGA Images on Card: Factory/Flex/User

In the Innova-Flex Open 18_07 bundle, the FPGA was released with a single image on board - the Flex Image. This image was burned at offset 0x0, and contained the Diagnostics and Burn- over-PCI features. At the time - this image was referred to as "Factory Image". It had the ability, along with the Flex-Open application, to burn a User Image on the flash.

In this bundle (18.12) and all future bundles, the FPGA will be shipped with two Mellanox images - a **Factory Image** located at offset 0x0, and a **Flex Image** located at offset 0x3000000. When burned by the customer using the application, the **User Image** will be located at offset 0x1000000 in the flash.

The Flex Image harbors the same functionality as before - it provides the Diagnostics and Burn- over-PCI features. This image may be updated in future releases with new features. The Factory Image is there to provide a fail-safe method to update the Flex Image. In case of a flex update failure, the system will fall back to Factory Image to allow re-burning of the Flex Image. The Factory Image is only there for this scenario, and cannot be used for diagnostics or for User Image burning.

2.10.8.2 Innova-2 Flex 25G vs. Innova-2 Flex VPI

The FPGA images in this bundle (and all future bundles) are compatible with both Innova-2 Flex 25G and Innova-2 Flex VPI boards. The only difference lies in the boards' ConnectX firmware. This is due to the different network ports on the cards (2x25Gb/s for Innova-2 Flex 25G and 2x40/100 Gb/s for Innova-2 Flex VPI). The bundle includes ConnectX firmware for Innova-2 Flex 25G and for Innova-2 Flex VPI in two separate folders - Morse_FW and MorseQ_FW.

2.11 Using the FPGA on the Adapter Card

Mellanox Innova™- 2 Flex Open dual-port network adapter combines ConnectX-5 with a fully open programmable FPGA. FPGA applications can be easily developed and deployed, utilizing Mellanox tools suite and the Xilinx standard development environment.

- For more information on the Xilinx Vivado tools and documents, see "Xilinx Vivado Tools and Documents".
- For instructions how to burn the User Image using the Innova™- 2 Flex Open bundle, see "Running the Burning Flows".
- For instructions how to burn the User Image using JTAG, see "Burning the FPGA Image through the JTAG".
- For instructions how to run diagnostics to make sure that the FPGA is working properly, see "Diagnostic Capabilities".

⚠ The user is responsible for temperature monitoring and avoiding FPGA overheating.

⚠ The ConnectX network traffic and FPGA offload traffic are competing on the same PCI link resources. It is the user's responsibility to avoid oversubscription of the PCI link resources.

2.12 Appendixes

The document contains the following appendixes:

- Finding the MAC and Serial Number on the Adapter Card
- Safety Warnings

2.12.1 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.

2.12.2 Safety Warnings

Safety warnings are provided here in the English language. For safety warnings in other languages, refer to the Adapter Installation Safety Instructions document available on mellanox.com.

Please observe all safety warnings to avoid injury and prevent damage to system components. 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.

Jewelry Removal Warning

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 melt down, 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 air flow of 200LFM at this maximum ambient temperature is required.To guarantee proper airflow, allow at least 8cm (3 inches) of clearance around the ventilation openings.The line" An air flow of 200LFM at this maximum ambient temperature is required." is for HCA cards and NICs only!

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

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+ A2:20

2.13 User Manual Revision History

Date	Revision	Description of Changes
March 2019	2.0	Updated: • "Burning Capabilities"
February 2019	1.9	Migrated to online format; minor reorganization.

Date	Revision	Description of Changes

January 2018	1.8	• Updated:
		  • "Features" Table
		  • "Connectivity"
		  • "InfiniBand Interface"
		  • "Ethernet Interface"
		  • "PCI Express Interface"
		  • "JTAG Interface"
		  • "FPGA DDR4 Interface"
		  • "FPGA PCIe ConnectX-5 Interface"
		  • "FPGA OpenCAPI Interface"
		  • "On-Board Clock" Table
		• Added:
		  • "Innova-2 Flex VPI MNV303611A-EDLT Open Adapter Card" Table
		  • "MNV303611A-EDLT Adapter Card Block Diagram"
		  • "MNV303212A-ADLT Air Flow Specifications" Table
		  • "MNV303611A-EDLT Air Flow Specifications" Table
		  • "MNV303611A-EDLT Drawing and Dimensions"
		  • "Reloading the Innova-2 User Image" (and its sub-sections)
		• Added the following files as a zip attachment:
		  • Verilog (.v) User pinout top level
		  • VHDL (.vhd) User pinout top level
		  • Xilinx Design Constrains (.xdc): Pinout location and type, clock rates, and device configuration

2.14 Release Update History

Release Update History

Release	Date	Description
19.04	April 30, 2019	First release of FW version 16.24.4000 with MLNX_OFED version 4.5-1.0.1.0 and FW version 16.25.1000 with MLNX_OFED version 4.6