NVIDIA BlueField-2 BF2500 InfiniBand/Ethernet DPU Controller User Manual

Table of Contents

About This Manual

This User Manual describes NVIDIA® BlueField®-2 BF2500 DPU Controller card. It provides details as to the interfaces of the board, specifications, required software and firmware for operating the card, hardware installation, driver installation and bring-up instructions.

EOL'ed (End of Life) Ordering Part Numbers

NVIDIA SKU	Legacy OPN	New Description
900-9D205-0056-ST0	MBF2H516B-EENOT	NVIDIA® BlueField®-2 P-Series BF2500 DPU Controller, **100GbE**/**EDR**/**HDR100** InfiniBand/Ethernet Dual-Port QSFP56, PCIe Gen4 **x16**, Crypto Disabled, **16GB** on-board DDR, 1GbE OOB management, Tall Bracket, **FHHL**
900-9D205-0066-ST0	MBF2H516B-EEEOT	NVIDIA® BlueField®-2 P-Series BF2500 DPU Controller, **100GbE**/**EDR**/**HDR100** InfiniBand/Ethernet Dual-Port QSFP56, PCIe Gen4 **x16**, **Crypto Enabled**, **16GB** on-board DDR, 1GbE OOB management, Tall Bracket, **FHHL**
900-9D205-0058-ST0	MBF2H515B-HENOT	NVIDIA® BlueField®-2 P-Series BF2500 DPU Controller, **200GbE**/**HDR** InfiniBand/Ethernet Single-Port QSFP56, PCIe Gen4 **x16**, Crypto Disabled, **16GB** on-board DDR, 1GbE OOB Management, Tall Bracket, **FHHL**
900-9D205-0068-ST0	MBF2H515B-HEEOT	NVIDIA® BlueField®-2 P-Series BF2500 DPU Controller, **200GbE**/**HDR** InfiniBand/Ethernet Single-Port QSFP56, PCIe Gen4 **x16**, **Crypto Enabled**, **16GB** on-board DDR, 1GbE OOB Management, Tall Bracket, **FHHL**

Overview of Document Content

Section	Section Description
Introduction	Provides a general overview of the BlueField-2 BF2500 DPU Controller Cards and discusses its benefits and features.
Supported Interfaces	Provides descriptions of all supported interfaces.
Pin Description	Provides interfaces' Pinouts description.
Cables and Cabling Configurations	Describes the required cables and cabling configurations.
Thermal Sensors	Describes the available thermal sensors on the BlueField-2 IPU
Hardware Installation	Describes the procedures for installing and uninstalling the Controller Card.
Bring-Up and Driver Installation	Describes driver installation and bring-up instructions

Troubleshooting	Describes potential system problems - you can use this troubleshooting information to identify and resolve the problem.
Specifications	Lists the physical, electrical, operational and regulatory specifications of the BlueField-2

Intended Audience

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

Technical Support

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

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

Related Documentation

Related Documentation

InfiniBand Architecture Specification	InfiniBand Trade Association (IBTA) InfiniBand® specification Release 1.3.1, November 2, 2016 and Vol. 2, Release 1.4 , and Vol 2 - Release 1.5.
IEEE Std 802.3 Specification	IEEE Ethernet specification.
PCI Express Specifications	Industry Standard PCI Express Base and Card Electromechanical Specifications.
NVIDIA LinkX Interconnect Solutions	The NVIDIA® LinkX® product family of cables and transceivers provide the industry's broadest portfolio of QDR/FDR10 (40Gb/s), FDR (56Gb/s), EDR/HDR100 (100Gb/s), HDR (200Gb/s) and NDR (400Gb/s) cables, including Direct Attach Copper cables (DACs), copper splitter cables, Active Optical Cables (AOCs) and transceivers in a wide range of lengths from 0.5m to 10km. In addition to meeting IBTA standards, NVIDIA tests every product in an end-to-end environment ensuring a Bit Error Rate of less than 1E-15.
BlueField DPU Platform BSP Documentation	This guide provides product release notes as well as information on the BSP and how to develop and/or customize applications, system software, and file system images for the BlueField platform.

InfiniBand Architecture Specification	InfiniBand Trade Association (IBTA) InfiniBand® specification Release 1.3.1, November 2, 2016 and Vol. 2, Release 1.4 , and Vol 2 - Release 1.5.
DOCA SDK Software Documentation	NVIDIA DOCA SDK software.

Document Conventions

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

Revision History

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

Introduction

This is the User Manual for Technologies BlueField®-2 BF2500 DPU Controller. This document provides details as to the interfaces of the card, specifications, required software and firmware for operating the board, and a step-by-step plan of how to bring up the BlueField-2 BF2500 DPU Controller.

System Requirements Overview

Motherboard	PCI Express 3.0/4.0-compliant motherboard with **x16** interface
System Power Supply	• Minimum 75W or greater system power supply through the PCIe x16 interface • P-Series DPU controllers with PCIe Gen 4.0 x16 require additional 75W through a supplementary 6-pin ATX power supply connector. NOTE: The power cable is not included in the package. It should be part of system wiring or it can be ordered separately as a system accessory.
Operating System	BlueField-2 DPU Controller can be shipped with Ubuntu – a Linux commercial operating system – which includes the OFED stack, and is capable of running all customer-based Linux applications seamlessly. The cards also support CentOS and have an out-of-band 1GbE management interface. For more information, please refer to the BlueField-2 Software User Manual.
Connectivity	• Interoperable with up to 200Gb/s InfiniBand and Ethernet switches • Passive copper cable with ESD protection • Powered connectors for optical and active cable support

Package Contents

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

> ⚠ If anything is damaged or missing, contact your sales representative at support@.com.

Card Package

	FHHL DPUs
Cards	1x BlueField-2 DPU
Accessories	1x tall bracket (shipped assembled on the card)

Accessories Kit

The accessories kit should be ordered separately. OPN: MBF20-DKIT.

	FHHL DPUs
Cables	1x USB 2.0 Type A to Mini USB Type B cable
	1x USB 2.0 Type A to 30pin Flat Socket

> ⚠ For FHHL 100Gb/s P-Series DPUs, you need a 6-pin PCIe external power cable to activate the card. The cable is not included in the package. For further details, please refer to External PCIe Power Supply Connector.

Product Overview

BlueField-2 BF2500 DPU Controller features the second generation BlueField-2 I/O Processing Unit (IPU) – an innovative and high-performance programmable networking engine. The IPU integrates an array of eight powerful 64-bit Armv8 A72 cores interconnected by a coherent mesh with a DDR4 memory controller and a dual-port Ethernet network controller. Providing unmatched scalability and efficiency, the BF2500 DPU Controller is the ideal adapter to accelerate the most demanding workloads in datacenter, cloud, service provider and storage environments.

> ⚠ The BlueField BF2500 DPU Controller should be installed only in a JBOF and JBOD Systems as it functions as a PCIe root-complex (RC) initiating PCIe bus operations. Installing it in a regular host system may damage the card.

Ideal Solution for JBOF and JBOD Systems

BlueField-2 I/O Processing Unit (IPU) is a highly integrated and efficient controller, optimized for NVMe storage systems, Network Functions Virtualization (NFV), Cloud, and Machine Learning workloads. BlueField-2 integrates all the discrete components of a storage system appliance into a single chip, including Arm core CPUs, PCIe switch, and a network controller, making it the premier solution for building Just-a-Bunch-Of-Flash (JBOF) systems, All-Flash-Array, and storage appliances for NVMe over Fabrics. With an integrated NVMe-oF offload accelerator, the BF2500 DPU Controller has a superior performance advantage over existing JBOF systems, significantly reducing storage transaction latency, while increasing IOPs (I/O operations per second).

BlueField-2 DPU Controllers

Part Number	MBF2H516B-EEEOT	MBF2H516B-EENOT	MBF2H515B-HEEOT	MBF2H515B-HENOT
Data Rate	**InfiniBand**: SDR/DDR/QDR/FDR/EDR/HDR100 **Ethernet**:1/10/25/40/50/100 Gb/s		**InfiniBand**: SDR/DDR/QDR/FDR/EDR/HDR100/HDR **Ethernet**:1/10/25/40/50/100/200 Gb/s	
Network Connector Type	Dual-port QSFP56		Single-port QSFP56	
Form Factor	PCIe Full Height, Half Length **Dimensions**: 168mm x 111mm (6.61in x 4.37in)			
Heatsink Height	Single Slot **Height**: 14.91mm (from PCB surface)			
PCIe Support[a]	PCIe Gen **3.0 / 4.0** SERDES @ 8.0GT/s / 16.0GT/s		PCIe Gen **4.0** SERDES @ 16.0GT/s	
On-board DDR4 Memory	Single-channel with 8 DDR4 8 bit + ECC (64bit + 8bit ECC) **16GB** @ 3200MT/s		Single-channel with 8 DDR4 8 bit + ECC (64bit + 8bit ECC) **16GB** @ 3200MT/s	
1GbE OOB Management	∫	∫	∫	∫
Crypto	Enabled	Disabled	Enabled	Disabled

Part Number	MBF2H516B-EEEOT	MBF2H516B-EENOT	MBF2H515B-HEEOT	MBF2H515B-HENOT
RoHS	RoHS Compliant			
BlueField SoC	BlueField-2 P-Series 8 cores (high-bin)			

a. **Note**: Refer to BlueField-2 Software and Firmware release notes for the availability of PCIe Gen 4.0 capabilities.

For more detailed information see Specifications.

Features and Benefits

This section describes hardware features and capabilities. Please refer to the software release notes for feature availability.

> ⚠ It is recommended to upgrade your BlueField product to the latest software and firmware versions available in order to enjoy the latest features and bug fixes.

Feature	Description
PCI Express (PCIe)[a]	PCIe Gen 3.0 (8GT/s) and Gen 4.0 (16GT/s) through an x16 edge connector. Gen 1.1 and 2.0 compatible
200Gb/s DPU Controller	BlueField-2 offers the highest throughput DPU Controller, supporting HDR 200b/s InfiniBand and 200Gb/s Ethernet and enabling any standard networking, clustering, or storage to operate seamlessly over any converged network leveraging a consolidated software stack.
InfiniBand Architecture Specification v1.3 compliant	BlueField-2 DPU Controller delivers low latency, high bandwidth, and computing efficiency for performance-driven server and storage clustering applications. BlueField-2 DPU Controller is InfiniBand Architecture Specification v1.3 compliant.
Up to 200 Gigabit Ethernet	DPU Controllers comply with the following IEEE 802.3 standards: 200GbE / 100GbE / 50GbE - IEEE 802.3bj, 802.3bm 100 Gigabit Ethernet - IEEE 802.3by, Ethernet Consortium25, 50 Gigabit Ethernet, supporting all FEC modes

Feature	Description
	- IEEE 802.3ba 40 Gigabit Ethernet - IEEE 802.3by 25 Gigabit Ethernet - IEEE 802.3ae 10 Gigabit Ethernet - IEEE 802.3ap based auto-negotiation and KR startup - IEEE 802.3ad, 802.1AX Link Aggregation - IEEE 802.1Q, 802.1P VLAN tags and priority - IEEE 802.1Qau (QCN) - Congestion Notification - IEEE 802.1Qaz (ETS) - IEEE 802.1Qbb (PFC) - IEEE 802.1Qbg - IEEE 1588v2 - Jumbo frame support (9.6KB)
On-board Memory	• DDR4 16GB on-board, 3200Mt/s • 16GB eMMC NAND flash memory for software • SPI Flash for NIC firmware
BlueField-2 IPU	The BlueField-2 IPU integrates eight 64-bit Armv8 A72 cores interconnected by a coherent mesh network, one DRAM controller, an RDMA intelligent network adapter supporting up to 200Gb/s, an embedded PCIe switch with endpoint and root complex functionality, and up to 16 lanes of PCIe Gen 3.0/4.0.
Overlay Networks	In order to better scale their networks, data center operators often create overlay networks that carry traffic from individual virtual machines over logical tunnels in encapsulated formats such as NVGRE and VXLAN. While this solves network scalability issues, it hides the TCP packet from the hardware offloading engines, placing higher loads on the host CPU. BlueField BF2500 DPU Controller effectively addresses this by providing advanced NVGRE and VXLAN hardware offloading engines that encapsulate and de-capsulate the overlay protocol.
RDMA and RDMA over Converged Ethernet (RoCE)	BF2500 card, utilizing RDMA (Remote Data Memory Access) and RoCE (RDMA over Converged Ethernet) technology, delivers low-latency and high-performance over Ethernet networks. Leveraging data center bridging (DCB) capabilities as well as advanced congestion control hardware mechanisms, RoCE provides efficient low-latency RDMA services over Layer 2 and Layer 3 networks.
PeerDirect®	PeerDirect communication provides high-efficiency RDMA access by eliminating unnecessary internal data copies between components on the PCIe bus (for example, from GPU to CPU), and therefore significantly reduces application run time. BlueField BF2500 DPU Controller advanced acceleration technology enables higher cluster efficiency and scalability to tens of thousands of nodes.
Quality of Service (QoS)	Support for port-based Quality of Service enabling various application requirements for latency and SLA.

Feature	Description
Storage Acceleration	A consolidated compute and storage network achieves significant cost-performance advantages over multi-fabric networks. Standard block and file access protocols can leverage RDMA for high-performance storage access. • NVMe over Fabric offloads for the target machine • T10-DIF Signature Handover BlueField-2 SmartNIC may operate as a co-processor offloading specific storage tasks from the host, isolating part of the storage media from the host, or enabling abstraction of software-defined storage logic using the BlueField-2 Arm cores. On the storage initiator side, BlueField-2 SmartNIC can prove an efficient solution for hyper-converged systems to enable the host CPU to focus on computing while all the storage interface is handled through the Arm cores.
NVMe-oF	Nonvolatile Memory Express (NVMe) over Fabrics is a protocol for communicating block storage IO requests over RDMA to transfer data between a host computer and a target solid-state storage device or system over a network. BlueField-2 DPU Controller may operate as a co-processor offloading specific storage tasks from the host using its powerful NVMe over Fabrics Offload accelerator.
SR-IOV	DPU Controller SR-IOV technology provides dedicated adapter resources and guaranteed isolation and protection for virtual machines (VM) within the server.
High-Performance Accelerations	• Tag Matching and Rendezvous Offloads • Adaptive Routing on Reliable Transport • Burst Buffer Offloads for Background Checkpointing
GPU Direct	The latest advancement in GPU-GPU communications is GPUDirect RDMA. This new technology provides a direct P2P (Peer-to-Peer) data path between the GPU Memory directly to/from the HCA devices. This provides a significant decrease in GPU-GPU communication latency and completely offloads the CPU, removing it from all GPU-GPU communications across the network. BlueField BF2500 DPU Controller uses high-speed DMA transfers to copy data between P2P devices resulting in more efficient system applications.
Crypto	The BlueField-2 DPU Controller crypto enabled versions include a BlueField-2 IC which supports accelerated cryptographic operations. In addition to specialized instructions for bulk cryptographic processing in the Arm cores, an offload hardware engine accelerates public-key cryptography and random number generation are enabled.
Security Accelerators	A consolidated compute and network solution based on BlueField DPU Controller achieves significant advantages over a centralized security server solution. Standard encryption protocols and security applications can leverage BlueField-2 compute capabilities and network offloads for security application solutions such as: • Connection Tracking • Statefull firewall
Out-of-Band Management	The BlueField-2 DPU Controller incorporates a 1GbE RJ45 out-of-band port that allows the network operator to establish trust boundaries in accessing the management function to apply it to network resources. It can also be used to ensure management connectivity (including the ability to determine the status of any network component) independent of the status of other in-band network components.

a. Refer to BlueField Software and Firmware release notes for the availability of PCIe Gen 4.0 capabilities

Supported Interfaces

Layout and Connector Information

The BF2500 DPU Controller is a **FHHL** card. The component height on the top and bottom of the card complies with the PCIe specification. The below figures show the connector and LED designations.

The below figures are for illustration purposes only and might not reflect the current revision of the BF2500 card.

BF2500 DPU Controller - Component Side
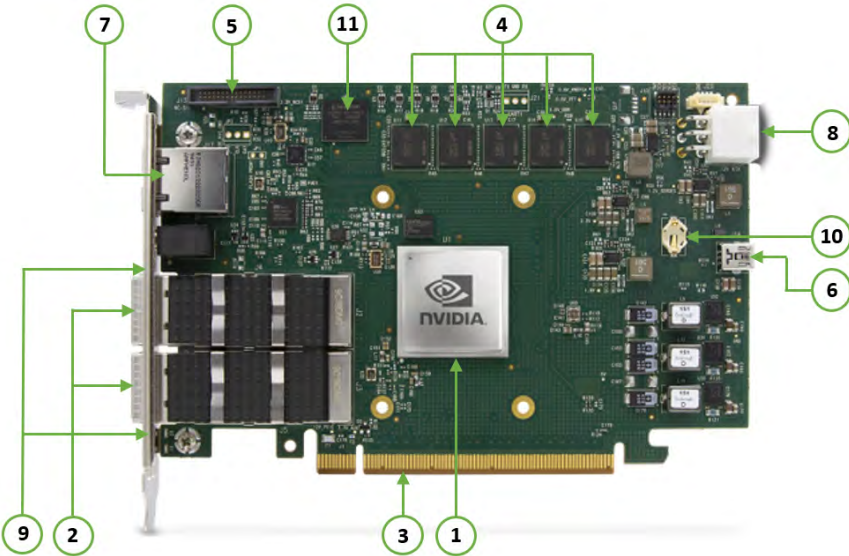
BF2500 DPU Controller - Print Side
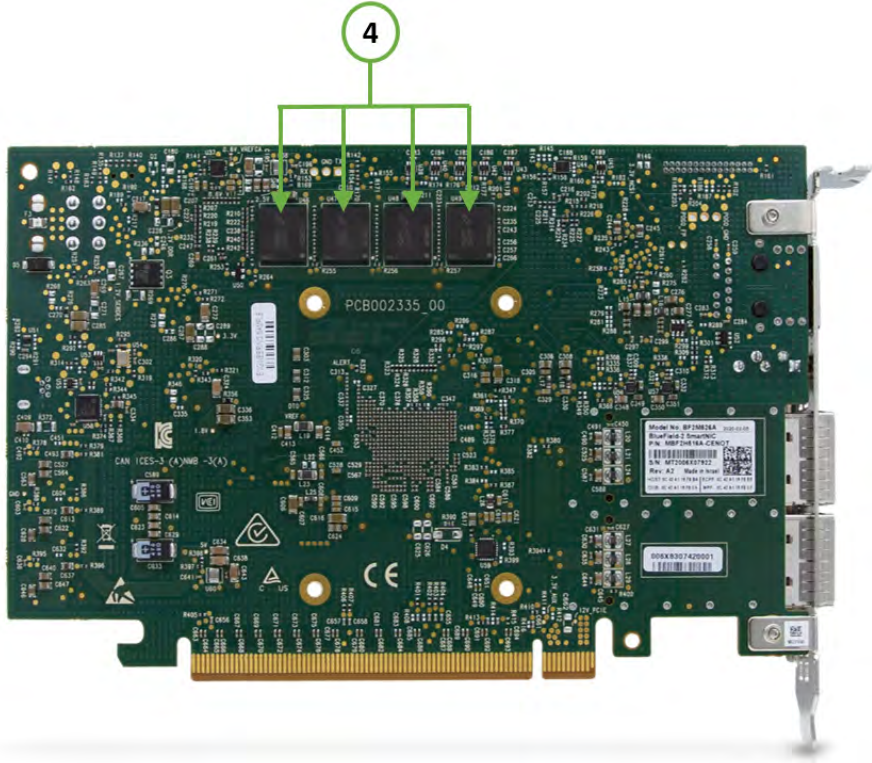
Callout	Interface	Description
1	BlueField-2 Data Processing Unit (DPU)	BlueField-2 DPU with heatsink
2	PCIe x16 Gen 3.0/4.0 Edge Connector	The interface between the BlueField-2 Controller and the JBOF system.
3	InfiniBand QSFP56 Interfaces	The network ports of the ConnectX®-6 adapter cards are compliant with the InfiniBand Architecture Specification, Release 1.3. InfiniBand traffic is transmitted through the cards' QSFP56 connectors.
	Ethernet QSFP56 Interfaces	Ethernet traffic is transmitted through the controller's QSFP56 connectors. The QSFP56 connectors allow for the use of Optical modules and cable interconnect solutions.
4	DDR4 SODIMM Memory	Single-channel with 8 DDR4 8 bit + ECC (64bit + 8bit ECC) 16GB @ 3200MT/s
5	NC-SI Management Interface	Connectivity channel between the BMC CPU and the BlueField-2 Arm Array
6	USB Connector	USB Type-B Bracket mounted on card
7	1GbE OOB Management Interface	1GbE BASE-T OOB management interface.
8	External PCIe Power Supply Interfaces	Along with the PCIe interface, this connector provides power to the board.
9	Networking LED Interfaces	One bi-color LED per port.
10	RTC Battery	Battery for RTC.
11	eMMC Interface	16GbE x8 NAND flash10

Interfaces Detailed Description

BlueField-2 Data Processing Unit (DPU)

BlueField-2 is a family of advanced DPU solutions that integrate a coherent mesh of 64-bit Arm v8 A72 cores, a ConnectX network adapter front-end and a PCI Express switch into a single chip. The powerful DPU architecture includes an Armv8 multicore processor array and enables customers to develop sophisticated applications and highly differentiated feature sets. BlueField-2 leverages the rich Arm software ecosystem and introduces the ability to offload the host server software stack.

At the heart of BlueField-2 is the ConnectX-6 Dx network offload controller with RDMA and RDMA over Converged Ethernet (RoCE) technology, delivering cutting-edge performance for networking and storage applications such as NVMe over Fabrics. Advanced features include an embedded virtual switch with programmable access lists (ACLs), transport offloads and stateless encaps/decaps of NVGRE, VXLAN, and MPLS overlay.

Encryption

> ⚠ Applies to Crypto enabled OPNs.

BlueField-2 BF2500 DPU Controller addresses the concerns of modern data centers by combining hardware encryption accelerators with embedded software and fully integrated advanced network capabilities, making it an ideal platform for developing proprietary security applications. It enables a distributed security architecture by isolating and protecting each individual workload and providing flexible control and visibility at the server and workload level, controlling risk at the server access layer. BlueField-2 builds security into the DNA of the data center and enables prevention, detection, and response to potential threats in real-time. BlueField-2 DPU Controller can deliver powerful functionality, including encryption of data-in-motion, and data-at-rest (AES-XTS) bare-metal provisioning, stateful L4 firewall and more.

InfiniBand QSFP56 Interface

The network ports of the BlueField-2 DPU Controllers are compliant with the *InfiniBand Architecture Specification, Release 1.3*. InfiniBand traffic is transmitted through the cards' QSFP56 connectors.

Ethernet QSFP56 Interfaces

The network ports of the BlueField-2 BF2500 DPU Controller are compliant with the IEEE 802.3 Ethernet standards listed in Ethernet traffic is transmitted through the cards' QSFP56 connectors.

QSFP56 Port Over-current Protection

The power consumption for each QSFP56 module is controlled by a 5W current limiting power distribution switch. Every power limiting switch is controlled by a signal, which is enabled by default. This signal is originated from the thermal shutdown mechanism. When the output load exceeds the current limit threshold, the power switch limits the current to a safe level.

PCI Express Interfaces

The PCI Express bus is a high-speed interface used to connect the DPU Controller to other peripheral devices (the BMC, PCIe switch or directly to the NVMe SSDs backplane), and operates as a Root-Complex (RC) initiating PCIe bus operations. The BlueField-2 BF2500 DPU Controller supports PCI Express Gen 3.0/4.0 (1.1 and 2.0 compatible) through x16 edge connector which serves as the main interface of the card. This connector enables the connectivity of the system to the BlueField-2 PCIe interface.

The following lists PCIe interface features:

- PCIe Gen 4.0 and 3.0 compliant, 2.0 and 1.1 compatible
- 2.5, 5.0, or 8.0, or 16.0 GT/s link rate x16
- Auto-negotiates to x8, x4, x2, or x1
- Support for MSI/MSI-X mechanisms

DDR4 Memory

The BlueField-2 BF2500 DPU Controller incorporates one on-board DDR4 memory sized 16GB, 64-bit data and 8-bit Error Correcting Code (ECC) memory.

The following lists the supported DDR4 data rates:

- 1600 Mb/s
- 1866 Mb/s
- 2133 Mb/s
- 2400 Mb/s
- 2666 Mb/s
- 2933 Mb/s
- 3200 Mb/s

GbE OOB Management

The BlueField-2 BF2500 DPU Controller incorporates a 1GbE RJ45 out-of-band port that allows the network operator to establish trust boundaries in accessing the management function to apply it to network resources. It can also be used to ensure management connectivity (including the ability to determine the status of any network component) independent of the status of other in-band network components.

1GbE OOB Management LEDs Interface

There are two I/O LEDs, LED1 is green and LED2 is Amber, to indicate link activity as described in the below table.

LED Pin	LED Definition	Link Activity
LED2 (Amber)	OFF	Link OFF
	ON	Link ON (Any speed)
LED1 (Green)	OFF	No Activity
	Blinking	Activity (RX, Tx)

NC-SI Management Interface

The Network Controller Sideband Interface (NC-SI) enables the connection of a Baseboard Management Controller (BMC) to a card for the purpose of enabling out-of-band remote manageability. NC-SI over the RMII interface is routed through the RMII connector with a flat cable.

The UART0 debug interface is connected to the NC-SI connector through the following pins:

- Pin # 30 - UART0 RX
- Pin # 28 - UART0 TX
- Pin # 26 - GND

The UART interface is complaint with TTL 3.3V voltage level.

UART Interface Connectivity

The USB to UART cable that supports TTL voltage levels can be used to connect the UART Interface for Arm console access. See an example of TTL 3.3V voltage levels cable in the below figure.
Please note that the USB to UART cable is not used for NC-SI management purposes.

It is prohibited to directly connect any RS-232 cable! Only TTL 3.3V voltage level cables are supported!

USB to UART Cable - Example

Mini USB Type B Connector

The BlueField-2 BF2500 DPU Controller uses a Mini USB 2.0 Type B device. The purpose of the USB connector is debugging and loading new versions of the operating systems and firmware.

External PCIe Power Supply Interfaces

The BlueField-2 BF2500 DPU Controller consumes power from two power sources; an external +12V through an 6-pin ATX connector (reference #9) and an additional +12V through the PCIe x16 edge connector. The BF2500 DPU Controller's special circuitry balances the power consumption between these two power sources. Each power source must be capable of supplying up to 75W of power. The 3.3V AUX from the card's PCIe edge connector is used for logics on the BF2500 DPU Controller. This logic includes power for buffering logic for signals to the card's edge connector, power for the FRU EEPROM, and

miscellaneous items. The 3.3V from PCIe edge connector is floating on the board. For the external PCIe power supply pins, please refer to External Power Supply Interfaces.

> ⚠ The use of an external power supply interface is restricted to the ATX 6-pin PCIe power connector only.

eMMC Interface

The BlueField-2 BF2500 DPU Controller incorporates an eMMC interface on the card's print side. The eMMC is 16GB x8 NAND flash and is used for Arm boot, operating system storage and disk space.

Networking LED Interfaces

There is a one bi-color I/O LED per port to indicate link status. See the below figure for the networking LEDs locations on the BF2500 DPU Controller bracket.

LED1 and LED2 Link Status Indications - Ethernet Protocol:

LED Color and State	Description
Off	A link has not been established
Beacon command for locating the adapter card	1Hz blinking Yellow
Error	4Hz blinking Yellow Indicates an error with the link. The error can be one of the following: | Error Type	Description	LED Behavior
---	---	---
I^2C	I^2C access to the networking ports fails	Blinks until error is fixed
Over-current	Over-current condition of the networking ports	Blinks until error is fixed	

Solid green	Indicates a valid link with no active traffic
Blinking green	Indicates a valid link with active traffic

LED1 and LED2 Link Status Indications - InfiniBand Protocol:

LED Color and State	Description
Off	A link has not been established
Beacon command for locating the adapter card	1Hz blinking Yellow
Error	4Hz blinking Yellow Indicates an error with the link. The error can be one of the following: <table><tr><th>Error Type</th><th>Description</th><th>LED Behavior</th></tr><tr><td>I^{2}C</td><td>I^{2}C access to the networking ports fails</td><td>Blinks until error is fixed</td></tr><tr><td>Over-current</td><td>Over-current condition of the networking ports</td><td>Blinks until error is fixed</td></tr></table>
Solid amber	Indicates an active link
Solid green	Indicates a valid (data activity) link with no active traffic
Blinking green	Indicates a valid link with active traffic

RTC Battery

The BF2500 DPU Controller incorporates a COIN TYPE LITHIUM BATTERY CR621 for RTC (Real Time Clock).

eMMC Interface

The BlueField-2 DPU incorporates an eMMC interface on the card's print side. The eMMC is 16GbE x8 NAND flash and is used for Arm boot, operating system storage and disk space.

Pin Description

PCI Express Interface

PCIe x16 Gen 4.0 Edge Connector

The BF2500 Controller Card uses a PCI Express x16 connector (component and print sides) according to the PCI Express 4.0 specifications. The following two tables provide component side pins description and print side pins description. For further details, please refer to PCIe x16 Gen 3.0/4.0 Edge Interface.

> ⚠ HSOp(x) and HSOn(x) stand for High Speed Output and HSIp(x) and HSIn(x) stand for High Speed Input. The direction is relative to the motherboard or system board.

Side A

Pin#	Pin Name	Pin Definition	Description
A1	PRSNT1#	INPUT	Card presence detect, connected to PRSNT2#
A2	12V	POWER	12V power
A3	12V	POWER	12V power
A4	GND	GND	Ground
A5	TCK	NC	Not connected
A6	TDI	INPUT	on board loop

Pin#	Pin Name	Pin Definition	Description
A7	TDO	OUTPUT	
A8	TMS	NC	Not connected
A9	3.3V	POWER	3.3V power
A10	3.3V	POWER	3.3V power
A11	PERST#	OUTPUT	PCIe Interface Reset#, drived by SoC
Key			
A12	GND	GND	Ground
A13	REFCLK+	OUTPUT	PCIe Reference Differential Clock. Not connected (default)
A14	REFCLK-	OUTPUT	PCIe Reference Differential Clock. Not connected (default)
A15	GND	GND	Ground
A16	PERP0	OUTPUT	SoC TX
A17	PERN0	OUTPUT	SoC TX
A18	GND	GND	Ground
A19	RSVD	NC	Not connected, option for SER_DO
A20	GND	GND	Ground
A21	PERP1	OUTPUT	SoC TX

Pin#	Pin Name	Pin Definition	Description
A22	PERN1	OUTPUT	SoC TX
A23	GND	GND	Ground
A24	GND	GND	Ground
A25	PERP2	OUTPUT	SoC TX
A26	PERN2	OUTPUT	SoC TX
A27	GND	GND	Ground
A28	GND	GND	Ground
A29	PERP3	OUTPUT	SoC TX
A30	PERN3	OUTPUT	SoC TX
A31	GND	GND	Ground
A32	RSVD	NC	Not connected, option for SER_CLK
A33	RSVD	NC	Not connected, option for SER_CAP
A34	GND	GND	Ground
A35	PERP4	OUTPUT	SoC TX
A36	PERN4	OUTPUT	SoC TX
A37	GND	GND	Ground

Pin#	Pin Name	Pin Definition	Description
A38	GND	GND	Ground
A39	PERP5	OUTPUT	SoC TX
A40	PERN5	OUTPUT	SoC TX
A41	GND	GND	Ground
A42	GND	GND	Ground
A43	PERP6	OUTPUT	SoC TX
A44	PERN6	OUTPUT	SoC TX
A45	GND	GND	Ground
A46	GND	GND	Ground
A47	PERP7	OUTPUT	SoC TX
A48	PERN7	OUTPUT	SoC TX
A49	GND	GND	Ground
A50	RSVD	NC	Not connected
A51	GND	GND	Ground
A52	PERP8	OUTPUT	SoC TX
A53	PERN8	OUTPUT	SoC TX

Pin#	Pin Name	Pin Definition	Description
A54	GND	GND	Ground
A55	GND	GND	Ground
A56	PERP9	OUTPUT	SoC TX
A57	PERN9	OUTPUT	SoC TX
A58	GND	GND	Ground
A59	GND	GND	Ground
A60	PERP10	OUTPUT	SoC TX
A61	PERN10	OUTPUT	SoC TX
A62	GND	GND	Ground
A63	GND	GND	Ground
A64	PERP11	OUTPUT	SoC TX
A65	PERN11	OUTPUT	SoC TX
A66	GND	GND	Ground
A67	GND	GND	Ground
A68	PERP12	OUTPUT	SoC TX
A69	PERN12	OUTPUT	SoC TX

Pin#	Pin Name	Pin Definition	Description
A70	GND	GND	Ground
A71	GND	GND	Ground
A72	PERP13	OUTPUT	SoC TX
A73	PERN13	OUTPUT	SoC TX
A74	GND	GND	Ground
A75	GND	GND	Ground
A76	PERP14	OUTPUT	SoC TX
A77	PERN14	OUTPUT	SoC TX
A78	GND	GND	Ground
A79	GND	GND	Ground
A80	PERP15	OUTPUT	SoC TX
A81	PERN15	OUTPUT	SoC TX
A82	GND	GND	Ground

Side B

Pin#	Pin Name	Pin Definition	Description
B1	12V	POWER	12V power
B2	12V	POWER	12V power
B3	12V	POWER	12V power
B4	GND	GND	Ground
B5	SMCLK	I/O	SMBUS_SCL - connected to BlueField IPU
B6	SMDAT	I/O	SMBUS_SDA - connected to BlueField IPU
B7	GND	GND	Ground
B8	3.3V	POWER	3.3V power
B9	TRST#	NC	Not connected
B10	3.3V_AUX	POWER	3.3V auxilary power
B11	WAKE	NC	Not connected, option connect to NIC_GPIO[13]
Key			
B12	RSVD	NC	Not connected, option for SER_DI
B13	GND	GND	Ground

Pin#	Pin Name	Pin Definition	Description
B14	PETP0	IN	SoC RX
B15	PETN0	IN	SoC RX
B16	GND	GND	Ground
B17	PRESNT2#	NC	Not connected
B18	GND	GND	Ground
B19	PETP1	IN	SoC RX
B20	PETN1	IN	SoC RX
B21	GND	GND	Ground
B22	GND	GND	Ground
B23	PETP2	IN	SoC RX
B24	PETN2	IN	SoC RX
B25	GND	GND	Ground
B26	GND	GND	Ground
B27	PETP3	IN	SoC RX
B28	PETN3	IN	SoC RX
B29	GND	GND	Ground

Pin#	Pin Name	Pin Definition	Description
B30	RSVD	NC	Not connected
B31	PRESNT2#	NC	Not connected
B32	GND	GND	Ground
B33	PETP4	IN	SoC RX
B34	PETN4	IN	SoC RX
B35	GND	GND	Ground
B36	GND	GND	Ground
B37	PETP5	IN	SoC RX
B38	PETN5	IN	SoC RX
B39	GND	GND	Ground
B40	GND	GND	Ground
B41	PETP6	IN	SoC RX
B42	PETN6	IN	SoC RX
B43	GND	GND	Ground
B44	GND	GND	Ground
B45	PETP7	IN	SoC RX

Pin#	Pin Name	Pin Definition	Description
B46	PETN7	IN	SoC RX
B47	GND	GND	Ground
B48	PRESNT2#	NC	Not connected
B49	GND	GND	Ground
B50	PETP8	IN	SoC RX
B51	PETN8	IN	SoC RX
B52	GND	GND	Ground
B53	GND	GND	Ground
B54	PETP9	IN	SoC RX
B55	PETN9	IN	SoC RX
B56	GND	GND	Ground
B57	GND	GND	Ground
B58	PETP10	IN	SoC RX
B59	PETN10	IN	SoC RX
B60	GND	GND	Ground
B61	GND	GND	Ground

Pin#	Pin Name	Pin Definition	Description
B62	PETP11	IN	SoC RX
B63	PETN11	IN	SoC RX
B64	GND	GND	Ground
B65	GND	GND	Ground
B66	PETP12	IN	SoC RX
B67	PETN12	IN	SoC RX
B68	GND	GND	Ground
B69	GND	GND	Ground
B70	PETP13	IN	SoC RX
B71	PETN13	IN	SoC RX
B72	GND	GND	Ground
B73	GND	GND	Ground
B74	PETP14	IN	SoC RX
B75	PETN14	IN	SoC RX
B76	GND	GND	Ground
B77	GND	GND	Ground

Pin#	Pin Name	Pin Definition	Description
B78	PETP15	IN	SoC RX
B79	PETN15	IN	SoC RX
B80	GND	GND	Ground
B81	PRESNT2#	OUT	X16 PCIe card detection, connected to PRESNT1#
B82	RSVD	NC	Not connected

Power Sequencing

The BF2500 Controller Card requires the 12V ATX PCIe power to come up prior or at the same time to the +12V at the PCIe edge connector.

NC-SI Management Interface

The below table lists the NCSI Management Interface pins description. For further details, please refer to NC-SI Management Interface.

Pin #	Signal Name	I/O	Description	
1	REF_CLK	Input	50M REF CLK for NCSI BUS	
2	GND	GND	Ground	
3	ARB_IN	Input	NCSI hardware arbitration input	

Pin #	Signal Name	I/O	Description	
4	GND	GND	Ground	
5	ARB_OUT	Output	NCSI hardware arbitration output	
6	GND	GND	Ground	
7	RX_D0	Output	Received data	Output fo

Pin #	Signal Name	I/O	Description	
				r S o C
8	GND	GND	Ground	
9	RX_D1	Output	Receive data	Output for SoC
10	GND	GND	Ground	
11	CRS_DV	Output	Carrier se	

Pin #	Signal Name	I/O	Description	
			nse/ Receive Data Valid	
12	GND	GND	Ground	
13	TX_D0	Input	Transmit data	Input for SoC
14	GND	GND	Ground	

Pin #	Signal Name	I/O	Description	
15	TX_D1	Input	Transmit data	Input for SoC
16	GND	GND	Ground	
17	TX_EN	Input	Transmit enable	
18	GND	GND	Ground	
19	NC			
20	NC			

Pin #	Signal Name	I/O	Description	
21	I2C_SDA	Bidirectional	I2C Serial Data	GW_ARM1
22	GND	GND	Ground	
23	I2C_SCL	Bidirectional	I2C Serial Clock	GW_ARM1
24	GND	GND	Ground	
25	GND	GND	Ground	

Pin #	Signal Name	I/O	Description	
26	GND	GND	Ground	
27	NC			
28	UART_TX	Output	Transmit data	Output for SoC
29	NC			
30	UART_RX	Input	Receive data	

External Power Supply Interfaces

The below table lists the External Power Supply pins of the external power supply interface on the BF2500 Controller card. For further details, please refer to External PCIe Power Supply Interfaces.

> ⚠ The mechanical pinout of the 6-pin external +12V power connector is shown below. This +12V connector is sometimes referred to as EPS-12V. Care should be taken to ensure the power is applied to the correct pins as some 6-pin ATX type connector can have different pinouts.

Pin #	Signal Name	Description
1	12V	ATX supplied 12V
2	12V	ATX supplied 12V
3	12V	ATX supplied 12V
4	GND	Power Return
5	Sense	Indicates proper connectivity
6	GND	Power Return

External +12V Mechanical Pinout Diagram
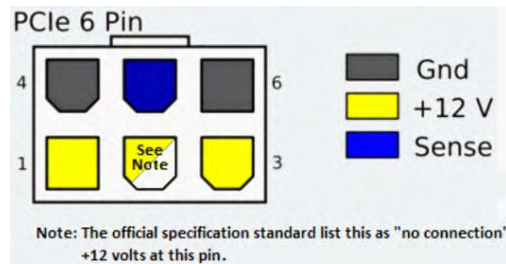
UART Interface

The below table lists the UART0 interface pins from the BlueField Controller Card perspective. For further details, please refer to UART0 Interface.
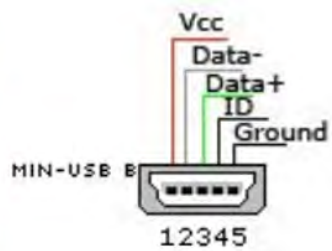
> ⚠ The electrical levels of the UART0 interface is TTL (3.3V).

3.5mm Audio Jack

Mini USB Interface

The below table lists the Mini USB interface pins. For further details, please refer to Mini USB Interface.

Pin	Name	Description
1	Vcc	+5v
2	D-	Data -
3	D+	Data +
4	ID	N/A
5	GND	Ground

Cables and Cabling Configurations

RJ45 Cable for the 1GbE OOB Management Interface

USB Interface Cabling

Debugging and loading new versions of the operating systems and firmware requires the use of Type A to Type B USB 2.0 Cable. The following figure shows an example of Type A to Type V USB Cable.

Thermal Sensors

The BlueField-2 BF2500 DPU Controller incorporates the BlueField-2 IPU which operates in the range of temperatures between 0C and 105C.

There are three thermal threshold definitions for the BlueField-2 IPU which impact the overall system operation state:

- **Warning** – 105°C: On managed systems only: When the device crosses the 100°C threshold, a Warning Threshold message will be issued by the management SW, indicating to system administration that the DPU Controller has crossed the Warning threshold. Note that this temperature threshold does not require nor lead to any action by hardware (such as BlueField-2 DPU Controller shutdown).
- **Critical** – 120°C: When the device crosses this temperature, the firmware will automatically shut down the device.
- **Emergency** – 130°C: In case the firmware fails to shut down the device upon crossing the Critical threshold, the device will auto-shutdown upon crossing the Emergency (130°C) threshold.

The BlueField-2 thermal sensors can be read through the system's SMBus. The user can read these thermal sensors and adapt the system airflow in accordance with the readouts and the needs of the above-mentioned IPU thermal requirements.

Hardware Installation

Installation and initialization of the BlueField-2 BF2500 DPU Controller require attention to the mechanical, power, and precautions for rack-mounted equipment.

Safety Warnings

> ⓘ Safety warnings are provided here in the English language. For safety warnings in other languages, refer to the DPU Controller Installation Safety Instructions.

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

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

Installation Procedure Overview

The installation procedure of BlueField-2 DPU Controller involves the following steps:

Step	Procedure	Direct Link
1	Unpack the package and confirm that you have received all the required components	Refer to Package Contents
2	Check the system's hardware and software requirements.	Refer to System Requirements
3	Pay attention to the airflow consideration within the JBOF system	Refer to Airflow Requirements
4	Follow the pre-installation checklist	Refer to Pre-Installation Checklist
5	Install the BF2500 DPU Controller in the JBOF system	Refer to Installation Instructions

7	Connect the external power supply to the BF2500 DPU Controller
8	Power on the system
9	Perform initial system bring-up

System Requirements

Hardware Requirements

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

A system that supports PCI Express high power cards is required for installing the card.

The system should be able to support 75W through the PCIe x16 interface and an additional 75W through the ATX PCIe power connector

Airflow Requirements

BlueField-2 BF2500 DPU Controller is offered with one airflow pattern: from the BlueField-2 SoC to the QSFP28 ports.

> ⚠ All systems in the same rack should be planned with the same airflow direction. All components need to have the same airflow direction.

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

Software Requirements

- Software Stacks - The BF2500 DPU Controller is shipped with Linux based Operating System burned on it which includes all needed drivers. For more information, please refer to the BlueField-2 Software User Manual.
- Check System Requirements Overview under Introduction for more details.

Safety Precautions

The BF2500 card being installed in a system that operates with voltages that can be lethal. Before opening the case of the system, observe the following precautions to avoid injury and prevent damage to system components.

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

Pre-Installation Checklist

1. **Unpack the BF2500 DPU Controller**

Unpack and remove the BF2500 DPU Controller. Check against the package contents list that all the parts have been sent. Check the parts for visible damage that may have occurred during shipping. Please note that the cards must be placed on an antistatic surface. For package contents please refer to Package Contents.

> ⚠ Please note that if the card is removed hastily from the antistatic bag, the plastic ziplock may harm the EMI fingers on the QSFP56 connector. Carefully remove the card from the antistatic bag to avoid damaging the EMI fingers.

2. **Turn off the power to the JBOF system.**

Turn off the power to the JBOF system, and disconnect the power cord and remove the cover. Refer to the JBOF system documentation for instructions. Before you install the BF2500 DPU Controller, make sure that the system is disconnected from power and any networks.

Installation Instructions

This section provides detailed instructions on how to install your BlueField-2 2500 DPU Controller in a system.

> ❗ The BlueField-2 BF2500 DPU Controller should be installed only in a JBOF System as it functions as a PCIe root-complex (RC) initiating PCIe bus operations. Installing it in a regular host system may damage the card.

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

Step 1. Open the system case.

Step 2. Locate an available PCI Express x16 slot.

Step 3. Applying even pressure at both corners of the card, insert the BF2500 DPU Controller in a PCI Express slot until firmly seated.

Step 4. Secure the bracket to the system with the bracket screw.

Step 5. Close the system case

Cables and Modules

Networking Cables

1. All networking 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. Verify 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).

4. After plugging in a cable, lock the connector using the latching mechanism particular to the cable vendor. When data is being transferred the Green LED will blink. See Networking LED Interfaces.

5. Care should be taken as not to impede the air exhaust flow through the ventilation holes. Use cable lengths that allow for routing horizontally around to the side of the chassis before bending upward or downward in the rack.

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

Bring-Up and Driver Installation

⚠ It is recommended to upgrade your BlueField product to the latest software and firmware versions in order to enjoy the latest features and bug fixes.

BlueField Software

NVIDIA provides software that enables users to fully utilize the BlueField® DPU and enjoy the rich feature-set it provides. Using BlueField software packages, users are able to:

- Quickly and easily boot an initial Linux image on your development board
- Port existing applications to and develop new applications for BlueField
- Patch, configure, rebuild, update or otherwise customize your image
    - Debug, profile, and tune their development system using open source development tools taking advantage of the diverse and vibrant Arm ecosystem.

The BlueField family of DPU devices combines an array of 64-bit Arm v8 A72 cores coupled with the ConnectX® interconnect. Standard Linux distributions run on the Arm cores allowing common open source development tools to be used. Developers should find the programming environment familiar and intuitive which in turn allows them to quickly and efficiently design, implement and verify their control-plane and data-plane applications.

BlueField SW ships with the NVIDIA BlueField Controller Cards. BlueField SW is a reference Linux distribution based on the Yocto Poky distribution and extended to include the NVIDIA OFED stack for Arm and a Linux kernel which supports NVMe-oF. This SW distribution is capable of running all customer-based Linux applications seamlessly. Yocto also provides an SDK that contains an extremely flexible cross-build environment allowing software targeted for the BlueField DPU to build on virtually any x86 server running any Linux distribution.

The following are other software elements delivered with BlueField DPU:

- Arm Trusted Firmware (ATF) for BlueField
- UEFI for BlueField
- Hardware Diagnostics
- NVIDIA OFED stack
- NVIDIA MFT

For the BlueField software and software user manual, please contact your local support.

> ⚠ Before installing the preferred OS on the BlueField Controller Card, make sure you install the card only in a JBOF system. Installing it in a host system may damage the card.

Software On eMMC

The BlueField Controller Card boots off eMMC upon power-up. The image flashed on the eMMC from the factory is the Yocto Linux.

Run the following command to discover the BlueField Software version:

```
cat /etc/bluefield_version
```

Yocto Distribution Installation

The BlueField tarball comes with pre-built Yocto images that can be installed. Please refer to The BlueField Software User Manual for the basic Yocto Installation.The core-image-full image is a full root filesystem image that is appropriate for imaging on the rootfs partition of the eMMC. Refer to the BlueField Software User Manual for instructions on how to build it yourself, or contact your local support to get the image.

1. Prepare the host environment. For more information, refer to Preparing the Host-Side Environment.
2. Boot the BlueField Controller Card over USB using the samples/install.bfb image.
3. Refer to samples/README.install for instructions.
4. To boot the BlueField Controller Card over USB from the server host, run:

```
cat install.bfb > /dev/rshim0/boot
```

5. To prepare the eMMC for Yocto installation from the Arm, run:

```
/opt/mlnx/scripts/bfinst --fullfs /tmp/core-image-full-bluefield.tar.xz
```

6. After the installation is done, execute lowercase reboot on the Arm.

```
shutdown -r now
```

7. Verify the version via:

```
cat /etc/bluefield_version
```

PXE Server Configuration on Host Side

Before installing CentOS 7 on the BlueField Controller Card, you need to configure the PXE server on the host side (x86) to allow the deployment of the CentOS image over the BlueField Controller Card.

Download the CentOS installation iso file from the following link:

```
# Download the centos installation iso file from http://mirror.centos.org/altarch/7/isos/aarch64/CentOS-7-aarch64-
Everything.iso
# cd <BF_INST_DIR>/distro/rhel/pxeboot
# ./setup.sh -d <BF_INST_DIR> -i <centos-installation.iso> [-c <ttyAMA0 >]
```

⚠️ UART0 (ttyAMA0) is used by default, or you can use "-c ttyAMA0" to manually specify UART0.

Installing Linux on BlueField Controller Card

This section demonstrates CentOS 7.4 installation on the BlueField Controller Card. Other OSs work similarly with the PXE boot installation process.

❗ Before installing the preferred OS on the BlueField Controller Card, make sure you install the BlueField Controller Card in a JBOF System. Installing it in a host system may damage the card.

Software Requirements

- CentOS 7.4 Linux OS. To get CentOS 7.4 image, run:

```
wget http://archive.kernel.org/centos-vault/altarch/7.4.1708/isos/aarch64/CentOS-7- aarch64-Everything.iso
```

> ⚠ Some required drivers do not compile and load if running CentOS 5.x or earlier.

> ⚠ Please note that CentOS 7.5 is not supported.

- Access to the latest BlueField Controller Card SW bundle:
  NVIDIA uses box.com to distribute BlueField software. Contact your sales/support representative for a custom link to download BlueField software releases.

- In this document, we assume the tarball BlueField-1.0.alphaX.XXXXX.tar.gz is extracted at / root, to do this, run the following command:

```
tar -xvf BlueField-1.0.alphaX.XXXXX.tar.xz -C /root
```

Preparing Host-Side Environment

Before installing the preferred OS on the BlueField Controller Card, the host must be set up for it to be capable of provisioning the BlueField Controller Card. The RShim USB driver is installed on the host to communicate with the RShim device on the BlueField DPU. The RShim USB driver must be installed so that it can push the initial bootloader and supply the OS image for PXE boot through the USB connection.

> ⚠ This process only needs to be done on the host machine which is provisioning the BlueField Controller Card, it is not required on the end machine.

Setup Procedure With Installation Script

If the host is running CentOS 7 (or equivalent) on the host, you may run a script to complete all the steps detailed in Preparing the Host-Side Environment.

```
/root/BlueField-1.0.alphaX.XXXXX/distro/rhel/pxeboot/setup.sh    \
-d /root/BlueField-1.0.alphaX.XXXXX/ \
-i /root/CentOS-7-aarch64-Everything.iso \
-o /root/dd-rhel7.4-mlnx-ofed-4.2-1.4.10.0-aarch64.iso \
-c ttyAMA0\
-k
```

Note that there should be no firewall blocking the IP communication between the BlueField Controller Card and the server host machine. If a firewall exists, disable it with the following commands:

```
iptables -F
iptables -t
nat -F
```

- The "-d" flag points to where the tar file has been extracted from, the script uses this directory to find all the source code it needs.
- The "-i" flag points to the OS installation disk. This is the image that is accessed via PXE boot to install the OS on the BlueField Controller Card.
- The "-o" flag points to the NVIDIA OFED driver disk for Arm. Download and extract it from http://www.mellanox.com/page/products_dyn?product_family=34.
- The "-c" flag specifies the default UART port for the OS to use since the BlueField DPU has two Arm UARTs. For the BlueField Controller Card, "ttyAMA0" is used, which is UART0.
- The "-t" flag is optional and needed for nonpxe boot. When specified and given the argument of what Controller card is set (BlueField Controller Card in this case), it generates a "nonpxe.bfb" file which contains the install kernel and rootfs. If this file is pushed to the RShim boot device, it automatically runs the installation process and skips the initial UEFI PXE boot operations. (the -t flag). Please refer to distro/rhel/pxe/README.
- The optional "-k" flag kickstarts auto-installation based on a default kickstart file which is installed as /var/pxe/ks/ks.cfg (optional).

Setup Procedure Without Installation Script

If the host is running CentOS 7 or equivalent, please refer to Preparing the Host-Side Environment for a simpler way to perform the installation using an installation script.

The following sections demonstrate CentOS 7 installation, however, installation in other environments should be relatively similar.

Step 1: Set up RShim Interface

The RShim driver communicates with the RShim device on the BlueField DPU. The RShim is in charge of many miscellaneous functions of the DPU, including resetting the Arm cores, providing the initial bootstream, and using the TMFIFO and the RShim network, to exchange network and console data with the host.

The RShim device can be reached by the host via the USB connector.

Step 2: Install RShim Drivers

To install the kernel modules, please follow the instruction in section RShim Host Driver.

To install the kernel modules, please follow the instruction in NVIDIA BlueField DPU Family Software Documentation under BlueField Software Overview > Installation and Initialization > RShim Host Driver.

Step 3: Configure TFTP Server

The host should be configured to act as a TFTP server to the BlueField Controller Card via the USB RShim network. This server provides the required files by the BlueField Controller Card to perform the PXE boot for installing the preferred OS.

> ⚠ Configuring the TFTP server requires a TFTP package. If it is not installed, install it via "yum install tftp" or "apt-get tftp", depending on your Linux distribution.
> Note: On some versions, the TFTP package cannot be found. In such cases, install "xinetd".

1. Extract the OS image and copy the required PXE boot components:

```
mount -t iso9660 -o loop CentOS-7-aarch64-Everything.iso /mnt
mkdir -p /var/lib/tftpboot/centos/7.4
cp /mnt/EFI/BOOT/BOOTAA64.EFI /var/lib/tftpboot/ cp /mnt/EFI/BOOT/grubaa64.efi /var/lib/tftpboot/
cp /mnt/images/pxeboot/vmlinuz /var/lib/tftpboot/centos/7.4
cp /mnt/images/pxeboot/initrd.img /var/lib/tftpboot/centos/7.4/initrd-orig.img
```

2. Patch the initrd with the eMMC driver and TMFIFO (RShim network) driver:

```
mkdir -p /tmp/.bfcentos mkdir -p $/tmp/.bfinstdd cd /tmp/.bfcentos
xzcat /var/lib/tftpboot/centos/7.4/initrd-orig.img | cpio -idm mount
/root/BlueField-1.0.alpha3.10409/distro/rhel/bluefield_dd/bluefield_dd-4.11.0- 22.el7a.aarch64.iso
/tmp/.bfinstddmkdir -p usr/lib/modules/4.11.0-22.el7a.aarch64/updates/cp
/tmp/.bfinstdd/lib/modules/4.11.0-22.el7a.aarch64/updates/dw_mmc*.ko usr/lib/mod- ules/
4.11.0-22.el7a.aarch64/updates/cp
/tmp/.bfinstdd/lib/modules/4.11.0-22.el7a.aarch64/updates/tmfifo.ko usr/lib/mod- ules/4.11.0-22.el7a.aarch64/
updates/cp
/root/BlueField-1.0.alpha3.10409/distro/rhel/bluefield_dd/bluefield_dd-4.11.0- 22.el7a.aarch64.iso ./
bluefield_dd.iso
umount /tmp/.bfinstdd; rmdir /tmp/.bfinstdd chown root:root
* -R
depmod -b
/tmp/.bfcentos 4.11.0-22.el7a.aarch64
find . | cpio
-oc | xz --check=crc32 --lzma2=dict=32MiB >
/var/lib/tftpboot/centos/7.4/ initrd.img
```

> ⚠ These commands assume that you are using kernel version "4.11.0-22.el7a.aarch64". If you are using a different version, utilize the corresponding bluefield_dd.iso. If none is found, compile one by running the following: source /path/to/SDK/environment-setup-aarch64-poky-linux; unset LDFLAGS; ./build- dd.sh   /path/to/kernel-devel-4.11.0-44.el7a.aarch64.rpm

3. Change the grub configuration to PXE boot over the right location:

```
cat >/var/lib/tftpboot/grub.cfg <<EOF
menuentry 'Install centos/7.4 AArch64 - BlueField'
--class red --class
gnu-linux
--class gnu --class os {
linux (tftp)/centos/7.4/vmlinuz ro ip=dhcp method=http://192.168.100.1/centos7 inst.dd=/bluefield_dd.iso
console=ttyAMA0
initrd (tftp)/centos/7.4/initrd.img
} EOF
```

4. Start the TFTP server:

```
systemctl restart tftp
```

> ⚠ Based on the system, the user may need to use "system TFTP restart" instead. Also, if required, the user might need to switch use "xinetd" instead of "TFTP".

Step 4: Set Up the DHCP Server

DHCP server set up on the host is required for BlueField Controller Card to get a private IP from the host for PXE boot process completion. Configure the correct server names and domain names so that the BlueField Controller Card can connect to the network via the host later on.

1. Get the server/domain names on the host:

```
bash-4.2$ cat /etc/resolv.conf
# Generated by NetworkManager search internal.mlnx.com labs.mlnx
nameserver 10.15.2.29
nameserver 10.15.2.16
```

This example shows that the domains are internal.mlnx.com and labs.mlnx, and the names of the servers are 10.15.2.29 and 10.15.2.16.

2. Set up the DHCP config file accordingly:

```
cat >/etc/dhcp/dhcpd.conf <<EOF
allow booting;
allow bootp;
subnet 192.168.100.0 netmask 255.255.255.0 {
range 192.168.100.10 192.168.100.20;
option broadcast-address 192.168.100.255;
option routers 192.168.100.1;
option domain-name-servers 10.15.2.29 10.15.2.16;
# Set the domain search according to the network configuration option domain-search "internal.tilera.com"
"mtbu.labs.mlnx"; next-server 192.168.100.1;
filename "/BOOTAA64.EFI";
}
```

```
# Specify
the IP address for
this client. host pxe_client {
hardware ethernet 00:1a:ca:ff:ff:01; fixed-address 192.168.100.2;
} EOF
```

> ⚠ It is recommended to back up the previous dhcpd.conf file before overwriting it.

Step 5: Set Up the HTTP Server

The TFTP server allows the PXE boot to load the initrd and kernel. The BlueField Controller Card obtains all the other required sources through the network, thus, making it necessary to set up an HTTP.

> ⚠ Setting up the HTTP server requires the HTTP package. If it is not installed, please install it via "yum install httpd" or "apt-get httpd", depending on your Linux distribution.

To configure the http server to serve the contents of the installation disk, run the following command:

```
cat >/etc/httpd/conf.d/pxeboot.conf <<EOF Alias /centos7
/mnt
<Directory /mnt>
    Options Indexes FollowSymLinks Require ip 127.0.0.1
    192.168.100.0/24
</Directory>
EOF
systemctl enable httpd systemctl restart
httpd
```

Flashing BlueField Controller Card Bootloader Code

Before installing an OS, flash the bootloader code first. The BlueField Controller Card is shipped with an initial bootloader code, and should be updated with the following instructions.

Opening Terminal Connection to BlueField Controller Card

To open a console window to the BlueField Controller Card, a terminal application is required. The application "minicom" is used for the flow, however, any standard terminal application can work, e.g. "screen".

> ⚠ Install minicom by running "yum install minicom" or "apt-get install minicom".

1. On the host, type "minicom" to open minicom on the current terminal, use "minicom -s" to set it up.
2. Go to the settings menu by pressing "Ctrl-a + o" (the setting menu opens by default when launching with the "-s" option). Navigate to the "Serial port setup" submenu and set the "Serial Device" to the one connected (should be one of the /dev/ttyUSBx if using the serial- UART cable).
3. Change the baud rate to 115200 8N1, and ensure that the hardware and software flow control are set to "No".
   Minicom Settings - Example

```
+----------------------------------------------------+
| A -    Serial Device      : /dev/ttyUSB0           |
|                                                    |
| C -    Callin Program     :                        |
| D -    Callout Program    :                        |
| E -    Bps/Par/Bits       : 115200 8N1             |
| F - Hardware Flow Control : No                     |
| G - Software Flow Control : No                     |
|                                                    |
|      Change which setting? █                       |
+----------------------------------------------------+
        | Screen and keyboard  |
        | Save setup as dfl    |
        | Save setup as..      |
        | Exit                 |
        +----------------------+
```

4. Select "Save setup as dfl" in order not to have to set it again in the future.

Using Initial Install Bootstream

1. On the host side, ensure that the RShim driver is running:

```
$ systemctl status rshim
```

An RShim device is located under the /dev directory, if you only have one, it should be "rshim0":

```
[root@bu-lab02 ~]# ls /dev/rshim0/
boot       console net       rshim
```

The boot device is used to push the bootstream to the BlueField Controller Card. Upon writing to it, it automatically resets the Arm cores so that it is booted using the pushed bootstream.

> ⚠ The console device can be used as a console instead of the serial-USB console. The primary bootloader does not support this device, however, UEFI and Linux support it. In cases where the special UART adapter board is unavailable, this can be used instead.

2. Push the initial install bootstream to the BlueField Controller Card:

```
cat /root/BlueField-1.0.alphaX.XXXXX/sample/install.bfb > \
/dev/rshim0/boot
```

On the terminal, various boot messages appear until Linux is loaded. This is the Yocto embedded Linux running off the kernel initramfs pushed in the bootstream.

3. When prompted, type in "root" to get to the command prompt without any password.
   Yocto Log

```
done.
Starting OpenBSD Secure Shell server: sshd
done.
Starting rpcbind daemon...done.
starting statd: done
exportfs: can't open /etc/exports for reading
NFS daemon support not enabled in kernel
Starting syslogd/klogd: done

Poky (Yocto Project Reference Distro) 2.3.1 bluefield /dev/ttyAMA0

bluefield login: root
root@bluefield:~# ▮
CTRL-A Z for help |115200 8N1 | NOR | Minicom 2.6.2 | VT102 |      Offline
```

4. After Linux is loaded, in the terminal, run the /opt/mlnx/scripts/bfrec script to update the bootloader.

Installing CentOS 7.4 on BlueField Controller Card

> ⚠ If the error "no root is found" appears in the installation process, check or disable the firewall as needed on the server host machine.

Full PXE Boot Installation

1. Get to the UEFI boot menu.
   a. Reboot the BlueField Controller Card by typing "reboot" on the console. A "UEFI firmware…" message should appear and the screen clears.
   b. Press ESC several times until you enter the UEFI boot menu.

   UEFI Boot Menu

   ```
   Continue                                This selection will
   Select Language        <Standard English>   direct the system to
   Boot Manager                            continue to booting
   Device Manager                          process
   Boot Maintenance Manager




   ▮



   ↕Move Highlight        <Enter>=Select Entry
   ```

   c. On the host, restart the DHCP and TFTP service:

   ```
   systemctl restart dhcpd
   systemctl restart tftp #might be xinetd
   ```

d. Navigate to the Boot Manager.
UEFI Boot Manager

```
**************************************************************************
6                              Boot Manager                              6
**************************************************************************

                                                Device Path :
    Boot Option Menu                            VenHw(F019E406-8C9C-11
                                                E5-8797-001ACA00BFC4)/
    Linux from rshim                            Image
    Linux from mmc0
    CentOS 7.4
    EFI Misc Device
    EFI Network
    EFI Internal Shell

      and  to change option, ENTER to select an
    option, ESC to exit




                                                      ▄
**************************************************************************
6                                                                        6
6 =Move Highlight        <Enter>=Select Entry      Esc=Exit              6
**************************************************************************
```

e. Select EFI Network, it will then use the TFTP service on the host to discover all available PXE boot options. Shortly after, a "..Fetching Netboot Image" message will appear enabling CentOS installation.
   Option to Install CentOS

```
I
  Install centos/7.4 AArch64 - BlueField                       ⬚







    Use the  and  keys to change the selection.
    Press 'e' to edit the selected item, or 'c' for a command prompt.
```

f. Select CentOS download.

> ⚠ This process may take few minutes as it fetches data over the USB network. Running "ifconfig" on the host and monitoring the RX/TX packets on the "tmfifo_net0" network indicates that the fetching data process is not complete.

g. Follow the installation instructions in the configuration menu. Recommended settings are included.

> ⚠ These configuration inputs are not needed when the kickstart option "-k" is specified when running the setup.sh script.

```
============================================================================
============================================================ VNC

Text mode provides a limited set of installation options. It does not offer custom partitioning for full
control over the disk layout. Would you like to use VNC mode instead?

1)  Start VNC

2)  Use text mode

Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: 2
============================================================================
============================================================================
Installation:main* 2:shell 3:log 4:storage-lo> Switch tab: Alt+Tab | Help: F1
1) [x] Language settings    2) [!] Time settings
(English (United States))   (Timezone is not set.)
3) [!] Installation source  4) [!] Software selection (Processing...)      (Processing...)
5) [!] Installation Destination 6) [x] Kdump
(No disks selected) (Kdump is enabled)
7)  [x] Network configuration    8) [!] Root password
(Wired (eth0) connected)    (Password is not set.)
9) [!] User creation
(No user will be created)
Please make your choice from above ['q' to quit | 'b' to begin installation | 'r' to refresh]: 2
============================================================================
============================================================================
Time settings Timezone: not set
NTP servers:not configured

1)  Set timezone
2)  Configure NTP servers
Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: 1
============================================================================
============================================================================
Timezone settings Available regions
Available regions
```

```
1)  Europe  6)  Pacific     10) Arctic
2)  Asia    7)  Australia   11) US
3)  America 8)  Atlantic    12) Etc
4)  Africa  9)  Indian
5)  Antarctica


Please select the timezone.
Use numbers or type names directly [b to region list, q to quit]: 11
=====================================================================================
=====================================================================================
Timezone settings


Available timezones in region US
1)  Alaska  4) Eastern  6) Mountain
2)  Arizona 5) Hawaii   7) Pacific
3)  Central
Please select the timezone.
Use numbers or type names directly [b to region list, q to quit]: 4
=====================================================================================
=====================================================================================
Installation

1) [x] Language settings    2) [x] Time settings (English (United States))      (US/Eastern timezone)
3) [x] Installation source  4) [x] Software selection (http://192.168.100.1/centos7)      (Minimal
Install)
5) [!] Installation Destination 6) [x] Kdump
(No disks selected) (Kdump is enabled)
7) [x] Network configuration    8) [!] Root password
(Wired (eth0) connected)    (Password is not set.)
9) [!] User creation
(No user will be created)
Please make your choice from above ['q' to quit | 'b' to begin installation | 'r' to refresh]: 4
=====================================================================================
=====================================================================================
Base environment Software selection

Base environment

1)  [x] Minimal Install 6) [ ] Server with GUI
2)  [ ] Compute Node    7) [ ] GNOME Desktop
3)  [ ] Infrastructure Server  8) [ ] KDE Plasma Workspaces
```

```
4)  [ ] File and Print Server   9) [ ] Development and Creative
5)  [ ] Basic Web Server    Workstation
Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: 9
================================================================================
================================================================================
Base environment Software selection

1)  [ ] Minimal Install 6)  [ ] Server with GUI
2)  [ ] Compute Node    7)  [ ] GNOME Desktop
3)  [ ] Infrastructure Server  8)  [ ] KDE Plasma Workspaces
4)  [ ] File and Print Server  9)  [x] Development and Creative
5) [ ] Basic Web Server Workstation
Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: c
================================================================================
================================================================================
Installation

1) [x] Language settings    2) [x] Time settings (English (United States))    (US/Eastern timezone)
3) [!] Installation source  4) [!] Software selection (Processing...)      (Processing...)
5) [!] Installation Destination 6) [x] Kdump
(No disks selected) (Kdump is enabled)
7) [x] Network configuration    8) [!] Root password
(Wired (eth0) connected)    (Password is not set.)
9) [!] User creation
(No user will be created)
Please make your choice from above ['q' to quit | 'b' to begin installation | 'r' to refresh]: 5
================================================================================
================================================================================
Probing storage...
Installation Destination

[x] 1) : 13.75 GiB (mmcblk0)

1 disk selected; 13.75 GiB capacity; 1007.5 KiB free ...

Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: c
================================================================================
================================================================================
Autopartitioning Options

[ ] 1) Replace Existing Linux system(s) [x] 2) Use All Space
```

```
[ ] 3) Use Free Space

Installation requires partitioning of your hard drive. Select what space to use for the install target.

Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: c
================================================================================
================================================================================
Partition Scheme Options [ ] 1) Standard Partition [ ] 2) Btrfs
[x] 3) LVM
[ ] 4) LVM Thin Provisioning

Select a partition scheme configuration.

Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: 1
================================================================================
================================================================================
Partition Scheme Options [x] 1) Standard Partition [ ] 2) Btrfs
[ ] 3) LVM

[ ] 4) LVM Thin Provisioning

Select a partition scheme configuration.

Please make your choice from above ['q' to quit | 'c' to continue | 'r' to refresh]: c Generating
updated storage configuration
Checking storage configuration...

================================================================================
================================================================================
Installation

1)  [x] Language settings   2) [x] Time settings (English (United States))     (US/Eastern timezone)
3) [x] Installation source  4) [x] Software selection (http://192.168.100.1/centos7)       (Development
and Creative
5) [x] Installation Destination     Workstation) (Automatic partitioning    6) [x] Kdump
selected)    (Kdump is enabled)
7) [x] Network configuration    8) [!] Root password
(Wired (eth0) connected)     (Password is not set.)
9) [!] User creation
(No user will be created)
Please make your choice from above ['q' to quit | 'b' to begin installation | 'r' to refresh]: 8
```

```
================================================================================
================================================================================
Please select new root password. You will have to type it twice.

Password:
Password (confirm):
================================================================================
================================================================================
Question

The password you have provided is weak: The password fails the dictionary check
-   it is based on a dictionary word. Would you like to use it anyway?

Please respond 'yes' or 'no': yes
================================================================================
================================================================================
Installation

1)  [x] Language settings   2) [x] Time settings (English (United States))     (US/Eastern timezone)
3) [x] Installation source  4) [x] Software selection (http://192.168.100.1/centos7)       (Development
and Creative
5) [x] Installation Destination     Workstation) (Automatic partitioning    6) [x] Kdump
selected)   (Kdump is enabled)
7) [x] Network configuration   8) [x] Root password (Wired (eth0) connected)       (Password is set.)
9) [ ] User creation
(No user will be created)
Please make your choice from above ['q' to quit | 'b' to begin installation | 'r' to refresh]: b
```

h. Enter "b" and press "enter" to initiate the installation process.
i. Press "Enter" to reboot into CentOS.
   CentOS Installation Completion Screen

```
Installing iwl6000-firmware (1347/1348)
Installing words (1348/1348)
Performing post-installation setup tasks
Installing boot loader
.
Performing post-installation setup tasks
.
Configuring installed system
.
Writing network configuration
.
Creating users
.
Configuring addons
.
Generating initramfs
.
Running post-installation scripts
.
        Use of this product is subject to the license agreement found at /usr/sh

        Installation complete.  Press return to quit
[anaconda] 1:main* 2:shell  3:log  4:storage-lo> Switch tab: Alt+Tab | Help: F1
```

Non-PXE Boot Installation

When the setup script is run with the "-t" option, it generates a nonpxe.bfb file at the directory where the script is run. The directory contains the install kernel and rootfs which are usually loaded by UEFI during the initial PXE boot stage. Thus, if pushing this file, the host TFTP server no longer needs to be used and UEFI would automatically load the install kernel and rootfs from the boot FIFO. Together with the "-k" kickstart option, the host can be configured to initiate non-PXE boot and automatic CentOS installation, as long as the host HTTP and DHCP servers are working. To kick off the installation process, run the following command on the host:

```
cat nonpxe.bfb > /dev/rshim0/boot; sleep 2; systemctl restart dhcpd
```

MLNX_OFED Installation

> ⚠ This section is relevant to non-Yocto Operating Systems only.

Installing MLNX_OFED on Arm Cores

Prerequisite Packages for Installing MLNX_OFED

- MLNX_OFED installation requires some prerequisite packages to be installed on the system.
  Currently, CentOS installed on the BlueField Controller Card has a private network to the host via the USB connection, and it can be used to Secure Copy Protocol (SCP) all the required packages. However, it is recommended for the BlueField Controller Card to have a direct access to the network to use "yum install" to install all the required packages. For direct access to the network, set up the routing on the host via:

```
iptables -t nat -o eth0 -A POSTROUTING -j MASQUERADE
echo 1 > /proc/sys/net/ipv4/ip_forward
systemctl restart dhcpd
```

> ⚠ "eth0" is the outgoing network interface on the host. Change this according to your system requirements.

> ⚠ These commands are not saved in Linux startup script, and might be needed to be applied again after host machine reboots.

- Reset the BlueField Controller Card network for Internet connection (access to the web) as long as the host is connected:

```
[root@localhost
~]# ifdown eth0;
ifup eth0 [root@localhost
~]# ping google.com
PING google.com (172.217.10.142) 56(84)
bytes of data.
64 bytes from lga34s16-in-f14.1e100.net (172.217.10.142): icmp_seq=1 ttl=53 time=19.2 ms
64 bytes from lga34s16-in-f14.1e100.net (172.217.10.142): icmp_seq=2 ttl=53 time=17.7 ms
64 bytes from
lga34s16-in-f14.1e100.net (172.217.10.142): icmp_seq=3 ttl=53 time=15.8 ms
```

- Run "yum install" to install all the required MLNX_OFED packages:

```
yum install rpm-build
yum group install "Development Tools" yum install kernel-devel-`uname -r`
yum install valgrind-devel libnl3-devel python-devel yum install tcl tk
```

Note that this is not needed if you installed CentOS 7 with the kickstart ("-k") option.
/auto/sw_mc_soc_project/distro/rhel/kernel-devel-4.11.0-22.el7a.aarch64.rpm

Removing Pre-installed Kernel Module

There are cases where the kernel is shipped with an earlier version of the mlx5_core driver taken from the upstream Linux code. This version does not support the BlueField Arm, but is loaded before the MLNX_OFED driver, and therefore, needs to be removed.

To remove the kernel module from the initramfs, run the following command:

```
mkdir /boot/tmp
cd /boot/tmp
gunzip < ../initramfs-4*64.img | cpio -i
rm -f lib/modules/4*/updates/mlx5_core.ko
rm -f lib/modules/4*/updates/tmfifo*.ko
cp ../initramfs-4*64.img ../initramfs-4.11.0-22.el7a.aarch64.img-bak
find | cpio -H newc -o | gzip -9 > ../initramfs-4*64.img
rpm -e mlx5_core
depmod -a
```

Installing MLNX_OFED on the BlueField Controller Card

1. Copy the MLNX_OFED image to the BlueField Controller Card via the USB network. The MLNX_OFED images should be provided in the software drop:

```
scp MLNX_OFED_LINUX-4.2-1.4.8.0-rhel7.4alternate-aarch64.iso \
root@192.168.100.2:/root
```

2. Mount the image on the BlueField Controller Card:

```
mount /root/MLNX_OFED_LINUX-4.2-1.4.8.0-rhel7.4alternate-aarch64.iso /mnt
```

3. Install MLNX_OFED.
   If the kernel on the BlueField is 4.11.0-22.el7a.aarch64, run:

```
cd /mnt
# ./mlnxofedinstall --bluefield
```

If the kernel is different than 4.11.0-22.el7a.aarch64, run:

```
cd /mnt
# ./mlnxofedinstall --add-kernel-support --skip-repo
```

> ⚠ For OFED to support DPDK, use the arguments "--upstream-libs" and "--dpdk".

```
./mlnxofedinstall --distro rhel7.4alternate --add-kernel-support --upstream-libs
--dpdk
```

This step might take longer than expected to be completed. If you are using a different pack- age than the required one, run "yum install".

> ⚠ If the date is not set correctly while installing MLNX_OFED, first, set the date (e.g date -s 'Mon Feb 5 15:02:10 EST 2018'), then run the installation.

4. Restart openibd:

```
/etc/init.d/openibd restart
```

Updating BlueField Controller Card Firmware

> ⚠ The below commands apply to MBF1M616A-CSNAT. The commands vary per OPN.

The below steps demonstrate how to manually update the firmware if the automatic process fails. The firmware image can be found in the BlueField Software package.

1. Copy the firmware image to the BlueField Arm:

```
scp fw-BlueField-rel--XX_XX_XXXX-MBF1M6X6A-CSNA_Ax.ini.bin \
root@192.168.100.2:/root
[root@localhost ~]# mst start
Starting MST (Mellanox Software Tools) driver set Loading MST PCI module - Success
Loading MST PCI configuration module - Success Create devices
Unloading MST PCI module (unused) - Success [root@localhost ~]# mst status
MST modules:
------------
MST PCI module is not loaded
MST PCI configuration module loaded
MST devices:
------------
/dev/mst/mt41682_pciconf0
- PCI configuration cycles access.
domain:bus:dev.fn=0000:04:00.0 addr.reg=88 data.reg=92
Chip revision is: 00
```

2. The output indicates that the device is "/dev/mst/mt41682_pciconf0". To update the firmware:

```
flint -d /dev/mst/mt41682_pciconf0 b \
-i /root/ fw-BlueField-rel-XX_XX_XXXX-MBF1M6X6A-CSNA_Ax.ini.bin
```

When using the mlx and ini files, use the following command instead:

```
mlxburn -d /dev/mst/mt41682_pciconf0 -fw fw-BlueField.mlx -c bf.ini
```

To burn the firmware which comes with OFED after OFED is installed, run:

```
/opt/mellanox/mlnx-fw-updater/firmware/mlxfwmanager_sriov_dis    -force
```

3. Power cycle the BlueField Controller Card for the new firmware to take effect.

```
root@bluefield:~# flint -i /opt/fw-Bluefield-red-18_99_4608_MBF1M6X6A-CSNA_Ax-
Flexboot-3.5.404_UEFI-14.15.20.bin -d /dev/mst/mt41682_pciconf0 b

    Current FW version flash:     18.24.0013
    New FW version:               18.99.4608

Burning FW image without signatures - 55%
Burning FW image without signatures - OK
Restoring signature
-I- To load new FW run mlxfwreset or reboot machine.
root@bluefield:~#
```

⚠ After MLNX_OFED is installed on the Arm cores, use the mlx5_core driver to use the two Ethernet ports on the BlueField Controller Card. If the Ethernet ports on the BlueField Controller Card are connected to the network, there is no need to bridge the host via RShim net to access the network.

To install the kernel modules, please follow the instruction in section RShim Host Driver.

Troubleshooting

General Troubleshooting

The BF2500 Controller Card no longer works	• Reseat the BF2500 Controller Card in its slot or a different slot, if necessary • Try using another cable • Reinstall the drivers for the network driver files may be damaged or deleted • Reboot the server
BF2500 Controller Card stopped working after installing another BF2500 card	• Try removing and re-installing all BF2500 Controller Cards • Check that cables are connected properly • Make sure your motherboard has the latest BIOS
Link indicator light is off	• Try another port on the switch • Make sure the cable is securely attached • Check you are using the proper cables that do not exceed the recommended lengths • Verify that your switch and BF2500 Controller Card port are compatible
Link light is on, but with no communication established	• Check that the latest driver is loaded • Check that both the BF2500 Controller Card and its link are set to the same speed and duplex settings

Specifications

> ⚠ Ensure your system can support 75W or greater system power supply through the PCIe x16 interface and an additional 75W through the supplementary 6-pin ATX power supply connector

MBF2H516B-EE[N/E]OT Specifications

Physical	Form Factor: FHHL Dimensions: 167.65mm x 11.15mm (6.6in. x 4.37in.)		
	Connector: Dual Ethernet QSFP56 (copper and optical)		
Protocol Support	**InfiniBand**: IBTA v1.3[a] **Auto-Negotiation:** 1X/2X/4X SDR (2.5Gb/s per lane), DDR (5Gb/s per lane), QDR (10Gb/s per lane), FDR10 (10.3125Gb/s per lane), FDR (14.0625Gb/s per lane), EDR (25Gb/s per lane) port, HDR100 (2 lane x 50Gb/s per lane)		
	Ethernet: 100GBASE-CR4, 100GBASE-KR4, 100GBASE-SR4, 56GBASE-R4, 550G Ethernet Consortium, 40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 25GBASE-CR/CR-S, 25GBASE-SR, 25GBASE-LR, 25G Ethernet Consortium, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-CX4, SGMII, 1000BASE-X		
	Data Rate	InfiniBand	SDR/DDR/QDR/FDR/EDR/HDR100/HDR
		Ethernet	1/10/25/40/50/100 Gb/s
	PCI Express Gen 3.0 / 4.0: SERDES @ 8.0GT/s / 16GT/s, 16 lanes (2.0 and 1.1 compatible)		
Power	Voltage: 12V		
	Maximum power available through QSFP56 port: 5W		
Environmental	Temperature	Operational	0°C to 55°C
		Non-operational	-40°C to 70°C[a]

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

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

MBF2H515B-HE[N/E]OT Specifications

Physical	Form Factor: FHHL Dimensions: 167.65mm x 11.15mm (6.6in. x 4.37in.)		
	Connector: Dual Ethernet QSFP56 (copper and optical)		
Protocol Support	**InfiniBand**: IBTA v1.3[a] **Auto-Negotiation**: 1X/2X/4X SDR (2.5Gb/s per lane), DDR (5Gb/s per lane), QDR (10Gb/s per lane), FDR10 (10.3125Gb/s per lane), FDR (14.0625Gb/s per lane), EDR (25Gb/s per lane) port, HDR100 (2 lane x 50Gb/s per lane), HDR (50Gb/s per lane) port		
	Ethernet: 200GBASE-CR4, 200GBASE-KR4, 200GBASE-SR4, 100GBASE-CR4, 100GBASE-KR4, 100GBASE-SR4, 50GBASE-R2, 50GBASE-R4, 40GBASE-CR4, 40GBASE-KR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 40GBASE-R2, 25GBASE-R, 20GBASE-KR2, 10GBASE-LR,10GBASE-ER, 10GBASE-CX4, 10GBASE-KX4, 10GBASE-CR, 10GBASE-KR, SGMII, 1000BASE-CX, 1000BASE-KX, 10GBASE-SR		
	Data Rate	InfiniBand	SDR/DDR/QDR/FDR/EDR/HDR100/HDR
		Ethernet	1/10/25/40/50/100/200 Gb/s

	PCI Express Gen 3.0 / 4.0: SERDES @ 8.0GT/s / 16GT/s, 16 lanes (2.0 and 1.1 compatible)		
Power	Voltage: 12V		
	Maximum power available through QSFP56 port: 5W		
Environme ntal	**Temperature**	**Operational**	0˚C to 55˚C
		Non- operational	-40˚C to 70˚C[a]
	Humidity	**Operational**	10% to 85% relative humidity
		Non- operational	10% to 90% relative humidity
Regulatory	Safety: CB / cTUVus / CE		
	EMC: CE / FCC / VCCI / ICES / RCM		
	RoHS: RoHS Compliant		

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

BlueField-2 BF2500 Mechanical Drawing and Dimensions

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

167.65

111.15
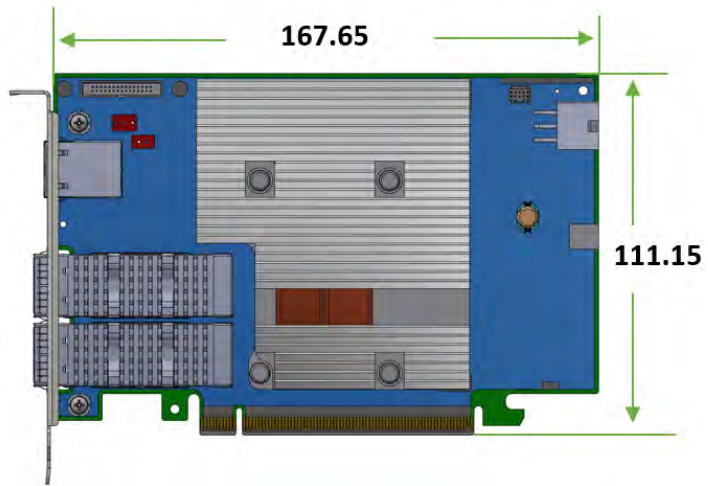
Bracket Mechanical Drawing and Dimensions

121.0

18.4
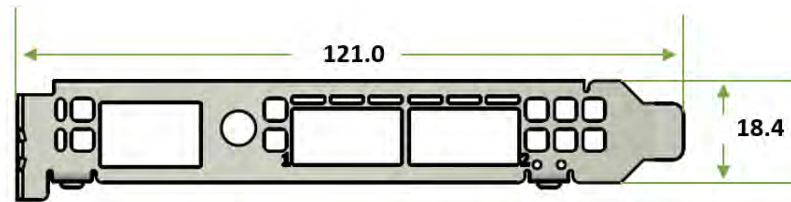
Finding the GUID/MAC and Serial Number on the Card

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

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

Board Label Example

Document Revision History

Date	Comments/Changes
May. 2023	Updated Specifications - added non-operational storage temperature specifications
Sep. 2021	Updated "System Requirements Overview".
Dec. 2020	- Updated the NCSI interface pinouts. - Removed the UART interface from "Supported Interfaces". - Updated "Package Contents".
Dec. 2020	Updated the NCSI interface pinouts.
Mar. 2020	First release