



Mellanox ConnectX[®]-5 Firmware Release Notes

Rev 16.22.1002

NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT ("PRODUCT(S)") AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES "ASIS" 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 NON-INFRINGEMENT 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

© Copyright 2018. Mellanox Technologies Ltd. All Rights Reserved.

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

All other trademarks are property of their respective owners.

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

Table of Contents

| | |
|---|-----------|
| Release Update History | 6 |
| Chapter 1 Overview | 7 |
| 1.1 Supported Devices | 7 |
| 1.2 Supported Cables and Modules | 8 |
| 1.2.1 Validated and Supported 1GbE Cables | 8 |
| 1.2.2 Validated and Supported 10GbE Cables | 8 |
| 1.2.3 Validated and Supported 25GbE Cables | 10 |
| 1.2.4 Validated and Supported 40GbE Cables | 11 |
| 1.2.5 Validated and Supported 50GbE Cables | 12 |
| 1.2.6 Validated and Supported 100GbE Cables | 12 |
| 1.2.7 Validated and Supported QDR Cables | 14 |
| 1.2.8 Validated and Supported FDR10 Cables | 14 |
| 1.2.9 Validated and Supported FDR Cables | 15 |
| 1.2.10 Validated and Supported EDR Cables | 15 |
| 1.3 Tested Switches | 16 |
| 1.3.1 Tested 10GbE Switches | 16 |
| 1.3.2 Tested 40GbE Switches | 17 |
| 1.3.3 Tested 100GbE Switches | 17 |
| 1.3.4 Tested QDR Switches | 18 |
| 1.3.5 Tested FDR Switches | 18 |
| 1.3.6 Tested EDR Switches | 18 |
| 1.4 Tools, Switch Firmware and Driver Software | 18 |
| 1.5 Supported FlexBoot, UEFI | 19 |
| 1.6 Revision Compatibility | 20 |
| Chapter 2 Changes and New Features in Rev 16.22.1002 | 21 |
| Chapter 3 Known Issues | 24 |
| Chapter 4 Bug Fixes History | 31 |
| Chapter 5 Firmware Changes and New Feature History | 38 |
| Chapter 6 FlexBoot Changes and New Features | 49 |
| 6.1 FlexBoot Known Issues | 51 |
| 6.2 FlexBoot Bug Fixes History | 57 |
| Chapter 7 UEFI Changes and Major New Features | 58 |
| 7.1 UEFI Known Issues | 58 |
| Chapter 8 Unsupported Features and Commands | 59 |
| 8.1 Unsupported Features | 59 |
| 8.2 Unsupported Commands | 59 |

Chapter 9 Supported Non-Volatile Configurations 60

List of Tables

| | | |
|-----------|--|----|
| Table 1: | Release Update History | 6 |
| Table 2: | Supported Devices | 7 |
| Table 3: | Validated and Supported 1GbE Cables | 8 |
| Table 4: | Validated and Supported 10GbE Cables | 8 |
| Table 5: | Validated and Supported 25GbE Cables | 10 |
| Table 6: | Validated and Supported 40GbE Cables | 11 |
| Table 7: | Validated and Supported 50GbE Cables | 12 |
| Table 8: | Validated and Supported 100GbE Cables | 12 |
| Table 9: | Validated and Supported QDR Cables | 14 |
| Table 10: | Validated and Supported FDR10 Cables | 14 |
| Table 11: | Validated and Supported FDR Cables | 15 |
| Table 12: | Validated and Supported EDR Cables | 15 |
| Table 13: | Tested 10GbE Switches | 16 |
| Table 14: | Tested 40GbE Switches | 17 |
| Table 15: | Tested 100GbE Switches | 17 |
| Table 16: | Tested QDR Switches | 18 |
| Table 17: | Tested FDR Switches | 18 |
| Table 18: | Tested EDR Switches | 18 |
| Table 19: | Tools, Switch Firmware and Driver Software | 18 |
| Table 20: | Supported FlexBoot, UEFI | 19 |
| Table 21: | Changes and New Features in Rev 16.22.1002 | 21 |
| Table 22: | Known Issues | 24 |
| Table 23: | Bug Fixes History | 31 |
| Table 24: | Firmware Changes and New Feature History | 38 |
| Table 25: | FlexBoot Changes and New Features | 49 |
| Table 26: | FlexBoot Known Issues | 51 |
| Table 27: | FlexBoot Bug Fixes History | 57 |
| Table 28: | UEFI Changes and New Features | 58 |
| Table 29: | UEFI Known Issues | 58 |
| Table 30: | Per-physical Port Settings | 60 |
| Table 31: | Global Settings | 60 |
| Table 32: | Per host/function Settings | 60 |
| Table 33: | Per host Settings | 60 |

Release Update History

Table 1 - Release Update History

| Release | Date | Description |
|----------------|-------------------|--|
| Rev 16.22.1002 | March 12, 2018 | Added Bug Fix 1231791. See Section 4, “Bug Fixes History” , on page 31 |
| | February 28, 2018 | Initial version of this firmware release. This version introduces new Changes and Features (see Section 2, “Changes and New Features in Rev 16.22.1002” , on page 21) and Bug Fixes (see Section 4, “Bug Fixes History” , on page 31). |

1 Overview

These are the release notes for the ConnectX®-5 adapters firmware Rev 16.22.1002. This firmware supports the following protocols:

- InfiniBand - SDR, QDR, FDR10, FDR, EDR
- Ethernet - 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 56GigE¹ and 100GigE
- PCI Express 3.0, supporting backwards compatibility for v2.0 and v1.1

1.1 Supported Devices

This firmware supports the devices and protocols listed in [Table 2](#)

Table 2 - Supported Devices (Sheet 1 of 2)

| Device Part Number | PSID | Device Name | FlexBoot | UEFI x86 | UEFI ARM | Enable/disable exprom feature |
|--------------------|---------------|--|-------------------|--------------------|-------------|-------------------------------|
| MCX512A-ACAT | MT_0000000080 | ConnectX®-5 EN network interface card, 10/25GbE dual-port SFP28, PCIe3.0 x8, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX515A-CCAT | MT_0000000011 | ConnectX-5 EN network interface card, 100GbE single-port QSFP28, PCIe3.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX515A-GCAT | MT_0000000087 | ConnectX®-5 EN network interface card, 50GbE single-port QSFP28, PCIe3.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX516A-BDAT | MT_0000000123 | ConnectX®-5 Ex EN network interface card, 40GbE dual-port QSFP28, PCIe 4.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX516A-CCAT | MT_0000000012 | ConnectX-5 EN network interface card, 100GbE dual-port QSFP28, PCIe3.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX516A-CDAT | MT_0000000013 | ConnectX-5 Ex EN network interface card, 100GbE dual-port QSFP28, PCIe4.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX516A-GCAT | MT_0000000090 | ConnectX®-5 EN network interface card, 50GbE dual-port QSFP28, PCIe3.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX545A-ECAN | MT_0000000077 | ConnectX®-5 VPI network interface card for OCP EDR IB (100Gb/s) and 100GbE, single-port QSFP28, PCIe3.0 x16, no bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX555A-ECAT | MT_0000000010 | ConnectX-5 VPI adapter card, EDR IB (100Gb/s) and 100GbE, single-port QSFP28, PCIe3.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX556A-ECAT | MT_0000000008 | ConnectX-5 VPI adapter card, EDR IB (100Gb/s) and 100GbE, dual-port QSFP28, PCIe3.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |

1. 56GbE is a Mellanox propriety link speed and can be achieved while connecting a Mellanox adapter cards to Mellanox SX10XX switch series or connecting a Mellanox adapter card to another Mellanox adapter card.

Table 2 - Supported Devices (Sheet 2 of 2)

| Device Part Number | PSID | Device Name | FlexBoot | UEFI x86 | UEFI ARM | Enable/disable exprom feature |
|--------------------|--------------|---|-------------------|--------------------|-------------|-------------------------------|
| MCX556A-EDAT | MT_000000009 | ConnectX-5 Ex VPI adapter card, EDR IB (100Gb/s) and 100GbE, dual-port QSFP28, PCIe4.0 x16, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX556M-ECAT-S25 | MT_000000023 | ConnectX®-5 VPI adapter card with Multi-Host Socket Direct supporting dual-socket server, EDR IB (100Gb/s) and 100GbE, dual-port QSFP28, 2x PCIe3.0 x8, 25cm harness, tall bracket, ROHS R6 | Present (Enabled) | Present (Disabled) | Not Present | Exists |
| MCX545A-CCAN | MT_000000157 | ConnectX-5 EN network interface card for OCP 100GbE; single-port QSFP28; PCIe3.0 x16; no bracket; ROHS R6; | Present (Enabled) | Present (Disabled) | Not Present | Exists |

1.2 Supported Cables and Modules

Please refer to the LinkX™ Cables and Transceivers web page

(<http://www.mellanox.com/products/interconnect/cables-configurator.php>) for the list of supported cables.

1.2.1 Validated and Supported 1GbE Cables

Table 3 - Validated and Supported 1GbE Cables

| Speed | Cable OPN # | Description |
|-------|--------------|--|
| 1GB/S | MC3208011-SX | Mellanox Optical module, SX, 850nm |
| 1GB/S | MC3208411-T | Mellanox® module, ETH 1GbE, 1Gb/s, SFP, Base-T, up to 100m |

1.2.2 Validated and Supported 10GbE Cables

Table 4 - Validated and Supported 10GbE Cables

| Speed | Cable OPN # | Description |
|--------|----------------|--|
| 10GB/S | CAB-SFP-SFP-1M | Arista 10GBASE-CR SFP+ Cable 1 Meter |
| 10GB/S | CAB-SFP-SFP-3M | Arista 10GBASE-CR SFP+ Cable 3 Meter |
| 10GB/S | CAB-SFP-SFP-5M | Arista 10GBASE-CR SFP+ Cable 5 Meter |
| 10GB/S | MC2309124-004 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 4M |
| 10GB/S | MC2309124-005 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 5M |
| 10GB/S | MC2309130-001 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 1M |
| 10GB/S | MC2309130-002 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 2M |

Table 4 - Validated and Supported 10GbE Cables

| Speed | Cable OPN # | Description |
|--------|----------------|--|
| 10GB/S | MC2309130-003 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 3M |
| 10GB/S | MC2309130-00A | Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 0.5M |
| 10GB/S | MC2609125-004 | Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 4M |
| 10GB/S | MC2609125-005 | Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 5M |
| 10GB/S | MC2609130-001 | Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 1M |
| 10GB/S | MC2609130-002 | Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 2M |
| 10GB/S | MC2609130-003 | Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 3M |
| 10GB/S | MC2609130-0A1 | Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 1.5M |
| 10GB/S | MC3309124-004 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 4M |
| 10GB/S | MC3309124-005 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 5M |
| 10GB/S | MC3309124-006 | Mellanox® Passive Copper Cable, ETH 10GbE, 10Gb/s, SFP+, 6m |
| 10GB/S | MC3309124-007 | Mellanox® Passive Copper Cable, ETH 10GbE, 10Gb/s, SFP+, 7m |
| 10GB/S | MC3309130-001 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 1M |
| 10GB/S | MC3309130-002 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 2M |
| 10GB/S | MC3309130-003 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 3M |
| 10GB/S | MC3309130-00A | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 0.5M |
| 10GB/S | MC3309130-0A1 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 1.5M |
| 10GB/S | MC3309130-0A2 | Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 2.5M |
| 10GB/S | MFM1T02A-LR-F | Mellanox Optical Module ETH 10GBE 10GB/S SFP+ LC-LC 1310NM LR UP TO 10KM |
| 10GB/S | MFM1T02A-SR-F | Mellanox Optical Module ETH 10GBE 10GB/S SFP+ LC-LC 850NM SR UP TO 300M |
| 10GB/S | SFP-10G-SR | Cisco 10GBASE-SR SFP+ transceiver module for MMF, 850-nm wavelength, LC duplex connector |
| 10GB/S | SFP-H10GB-CU1M | Cisco 1-m 10G SFP+ Twinax cable assembly, passive |
| 10GB/S | SFP-H10GB-CU3M | Cisco 3-m 10G SFP+ Twinax cable assembly, passive |
| 10GB/S | SFP-H10GB-CU5M | Cisco 5-m 10G SFP+ Twinax cable assembly, passive |

1.2.3 Validated and Supported 25GbE Cables



The 25GbE cables can be supported in ConnectX-5 adapter cards only when connected to the MAM1Q00A-QSA28 module.

Table 5 - Validated and Supported 25GbE Cables

| Speed | Cable OPN # | Description |
|-------|------------------|---|
| 25GbE | FTLF8536P4BCL | Finisar SFP+ transceivers 25Gb/s |
| 25GbE | LTF8507-PC07 | Hisense active fiber cable, 25GbE |
| 25GbE | MCP2M00-A001 | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 1m |
| 25GbE | MCP2M00-A002 | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 2m |
| 25GbE | MCP2M00-A003 | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 3m |
| 25GbE | MCP2M00-A003AP | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 3m, 26AWG |
| 25GbE | MCP2M00-A005E26L | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 5m, Black, 26AWG, CA-L |
| 25GbE | MCP2M00-A00A | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 0.5m |
| 25GbE | MCP2M00-A01A | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 1.5m |
| 25GbE | MCP2M00-A01A | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 1.5m |
| 25GbE | MCP2M00-A02A | Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 2.5m |
| 25GbE | MCP7F00-A001 | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 TO 4XSFP28 1M |
| 25GbE | MCP7F00-A002 | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 TO 4XSFP28 2M |
| 25GbE | MCP7F00-A003 | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 TO 4XSFP28 3M |
| 25GbE | MCP7F00-A003-AM | Mellanox® passive copper hybrid cable, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 3M 30AWG |
| 25GbE | MCP7F00-A005AM | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 to 4XSFP28 5M |
| 25GbE | MCP7F00-A01A | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 to 4XSFP28 1.5M |
| 25GbE | MCP7F00-A02A | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 to 4XSFP28 2.5M |
| 25GbE | MFA2P10-Axxx | Mellanox® active optical cable 25GbE, SFP28, up to 100m |
| 25GbE | MMA2P00-AS | Mellanox® transceiver, 25GbE, SFP28, LC-LC, 850nm, SR, up to 100m |
| 25GbE | SFP-H25G-CU1M | 25GBASE-CR1 Copper Cable 1-meter |
| 25GbE | SFP-H25G-CU2M | 25GBASE-CR1 Copper Cable 2-meter |
| 25GbE | SFP-H25G-CU3M | 25GBASE-CR1 Copper Cable 3-meter |

Table 5 - Validated and Supported 25GbE Cables

| Speed | Cable OPN # | Description |
|-------|------------------|---|
| 25GbE | MMA2P00-AS | Mellanox® transceiver, 25GbE, SFP28, LC-LC, 850nm, SR, up to 100m |
| 25GbE | MFA2P10-A100 | Mellanox® active optical cable 25GbE, SFP28, 100m |
| 25GbE | MFA7A50-C030 | Mellanox Active Fiber Hybrid Solution ETH 100GBE TO 4X25GBE QSFP28 TO 4XSFP28 30M |
| 25GbE | MCP2M00-A005E26L | Mellanox Passive Copper Cable, ETH, UP TO 25GB/S, SFP28, 5M, BLACK, 26AWG, CA-L |

1.2.4 Validated and Supported 40GbE Cables

Table 6 - Validated and Supported 40GbE Cables

| Speed | Cable OPN # | Description |
|-------|-------------------|---|
| NA | MAM1Q00A-QSA | Mellanox® cable module, ETH 10GbE, 40Gb/s to 10Gb/s, QSFP to SFP+ |
| NA | MAM1Q00A-QSA28 | Mellanox® cable module, ETH 25GbE, 100Gb/s to 25Gb/s, QSFP28 to SFP28 |
| 40GbE | MC2210126-004 | Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 4m |
| 40GbE | MC2210126-005 | Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 5m |
| 40GbE | MC2210128-003 | Mellanox Passive Copper Cable ETH 40GBE 40GbE QSFP 3M |
| 40GbE | MC2210130-001 | Mellanox Passive Copper Cable ETH 40GBE 40GbE QSFP 1M |
| 40GbE | MC2210130-002 | Mellanox Passive Copper Cable ETH 40GBE 40GbE QSFP 2M |
| 40GbE | MC2210130-00A | Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 0.5m |
| 40GbE | MC2210130-00B | Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 0.75m |
| 40GbE | MC2210310-XXX | Mellanox Active Fiber Cable ETH 40GBE 40GbE QSFP from 3M up to 100M |
| 40GbE | MC2210411-SR4L | Mellanox Optical Module 40GbE QSFP MPO 850NM UP TO 30M |
| 40GbE | MC2210411-SR4 | Mellanox Optical Module 40GbE QSFP MPO 850NM UP TO 100M |
| 40GbE | MC2210411-SR4E | Mellanox Optical Module 40GbE QSFP MPO 850NM UP TO 300M |
| 40GbE | QSFP-40G-SR-BD | Cisco 40GBASE-SR-BiDi, duplex MMF |
| 40GbE | QSFP-40G-SR4 | Cisco 40GBASE-SR4, 4 lanes, 850 nm MMF |
| 40GbE | QSFP-H40G-ACU10M | Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 10-meter, active |
| 40GbE | QSFP-H40G-AOC-10M | Cisco 40GBase-AOC QSFP direct-attach Active Optical Cable, 10-meter |
| 40GbE | QSFP-H40G-CU1M | Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 1-meter, passive |
| 40GbE | QSFP-H40G-CU3M | Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 3-meter, passive |

Table 6 - Validated and Supported 40GbE Cables

| Speed | Cable OPN # | Description |
|-------|----------------|---|
| 40GbE | QSFP-H40G-CU5M | Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 5-meter, passive |

1.2.5 Validated and Supported 50GbE Cables

Table 7 - Validated and Supported 50GbE Cables

| Speed | Cable OPN # | Description |
|-------|--------------|--|
| 50GbE | MCP7H00-G001 | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 2X50GBS QSFP28 TO 2XQSFP28 1M |
| 50GbE | MCP7H00-G002 | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 2X50GBS QSFP28 TO 2XQSFP28 2M |
| 50GbE | MCP7H00-G003 | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 2X50GBS QSFP28 TO 2XQSFP28 3M |
| 50GbE | MCP7H00-G01A | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 2X50GBS QSFP28 TO 2XQSFP28 1.5M |
| 50GbE | MCP7H00-G02A | Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 2X50GBS QSFP28 TO 2XQSFP28 2.5M |
| 50GbE | MFA7A20-C020 | Mellanox® active fiber hybrid solution, ETH 100GbE to 2x50GbE, QSFP28 to 2xQSFP28, 20m |

1.2.6 Validated and Supported 100GbE Cables

Table 8 - Validated and Supported 100GbE Cables

| Speed | Cable OPN # | Description |
|--------|------------------|--|
| 100GbE | MCP1600-C001 | Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 1M |
| 100GbE | MCP1600-C002 | Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 2M |
| 100GbE | MCP1600-C003 | Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 3M |
| 100GbE | MCP1600-C005AM | Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP, 5m, 26AWG |
| 100GbE | MCP1600-C005E26L | Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP28, 5m, Black, 26AWG, CA-L |
| 100GbE | MCP1600-C00A | Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 0.5M |
| 100GbE | MCP1600-C01A | Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP, LSZH, 1.5m |
| 100GbE | MCP1600-C02A | Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP, LSZH, 2.5m |

Table 8 - Validated and Supported 100GbE Cables

| Speed | Cable OPN # | Description |
|--------|------------------|--|
| 100GbE | MCP1600-C03A | Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP, PVC, 3.5m 26AWG |
| 100GbE | MCP7F00-A005R26L | Mellanox® passive copper hybrid cable, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 5m, Colored, 26AWG, CA-L |
| 100GbE | MCP7H00-G005R26L | Mellanox® passive copper hybrid cable, ETH 100GbE to 2x50GbE, QSFP28 to 2xQSFP28, 5m, Colored, 26AWG, CA-L |
| 100GbE | MFA1A00-C003 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 3m |
| 100GbE | MFA1A00-C005 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 5m |
| 100GbE | MFA1A00-C010 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 10m |
| 100GbE | MFA1A00-C015 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 15m |
| 100GbE | MFA1A00-C020 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 20m |
| 100GbE | MFA1A00-C030 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 30m |
| 100GbE | MFA1A00-C050 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 50m |
| 100GbE | MFA1A00-C100 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 100m |
| 100GbE | MFS1200-C005 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 5m |
| 100GbE | MFS1200-C010 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 10m |
| 100GbE | MFS1200-C015 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 15m |
| 100GbE | MFS1200-C020 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 20m |
| 100GbE | MFS1200-C030 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 30m |
| 100GbE | MFS1200-C050 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 50m |
| 100GbE | MFS1200-C100 | Mellanox® Active Fiber Cable, ETH 100GbE, 100GbE, QSFP, LSZH, 100m |
| 100GbE | MMA1B00-C100_B | Mellanox® transceiver, up to 100GbE, QSFP28, MPO, 850nm, up to 100m OM3 |
| 100GbE | MMA1B00-C100D | Mellanox® Transceiver, 100GbE, QSFP28, MPO, 850nm, up to 100m |

Table 8 - Validated and Supported 100GbE Cables

| Speed | Cable OPN # | Description |
|--------|--------------|--|
| 100GbE | MMA1L10-CR | Mellanox® optical transceiver, 100GbE, 100GbE, QSFP28, LC-LC, 1310nm, LR4 up to 10km Note: Only revision A2 and above. |
| 100GbE | MMS1C00-C500 | Mellanox® transceiver, 100GbE, QSFP28, MPO, 1550nm PSM4, up to 2km |
| 100GbE | MMS1C00-C500 | Mellanox® transceiver, 100GbE, QSFP28, MPO, 1550nm PSM4, up to 2km |
| 100GbE | MMS1C00-CM | Mellanox® transceiver, 100GbE, QSFP28, MPO, 1550nm PSM4, up to 2km for internal use only |
| 100GbE | MMS1C10-CM | Mellanox® active optical module, 100Gb/s, QSFP, MPO, 1310nm, PSM4 |
| 100GbE | MMS1C00-CM | Mellanox® transceiver, 100GbE, QSFP28, MPO, 1550nm PSM4, up to 2km |

1.2.7 Validated and Supported QDR Cables

Table 9 - Validated and Supported QDR Cables

| Speed | Cable OPN # | Description |
|-------|---------------|---|
| QDR | MC2206125-007 | Mellanox Passive Copper Cable IB QDR 40GB/S QSFP 7M |
| QDR | MC2206126-006 | Mellanox Passive Copper Cable IB QDR 40GB/S QSFP 6M |

1.2.8 Validated and Supported FDR10 Cables

Table 10 - Validated and Supported FDR10 Cables

| Speed | Cable OPN # | Description |
|-------|---------------|---|
| FDR10 | MC2206128-004 | Mellanox Passive Copper Cable VPI UP TO 40GB/S QSFP 4M |
| FDR10 | MC2206128-005 | Mellanox Passive Copper Cable VPI UP TO 40GB/S QSFP 5M |
| FDR10 | MC2206130-001 | Mellanox Passive Copper Cable VPI UP TO 40GB/S QSFP 1M |
| FDR10 | MC2206130-002 | Mellanox Passive Copper Cable VPI UP TO 40GB/S QSFP 2M |
| FDR10 | MC2206130-003 | Mellanox Passive Copper Cable VPI UP TO 40GB/S QSFP 3M |
| FDR10 | MC2206130-00A | Mellanox Passive Copper Cable VPI UP TO 40GB/S QSFP 0.5M |
| FDR10 | MC2206310-XXX | Mellanox Active Fiber Cable IB QDR/FDR10 40GB/S QSFP from 3M up to 100M |
| FDR10 | MFS4R12CB-XXX | Mellanox Active Fiber Cable VPI UP TO 40GB/S QSFP from 3M up to 100M |

1.2.9 Validated and Supported FDR Cables

Table 11 - Validated and Supported FDR Cables

| Speed | Cable OPN # | Description |
|-------|----------------|--|
| FDR | MC2207126-004 | Mellanox Passive Copper Cable VPI UP TO 56GB/S QSFP 4M |
| FDR | MC2207128-003 | Mellanox Passive Copper Cable VPI UP TO 56GB/S QSFP 3M |
| FDR | MC2207128-0A2 | Mellanox Passive Copper Cable VPI UP TO 56GB/S QSFP 2.5M |
| FDR | MC2207130-001 | Mellanox Passive Copper Cable VPI UP TO 56GB/S QSFP 1M |
| FDR | MC2207130-002 | Mellanox Passive Copper Cable VPI UP TO 56GB/S QSFP 2M |
| FDR | MC2207130-00A | Mellanox Passive Copper Cable VPI UP TO 56GB/S QSFP 0.5M |
| FDR | MC2207130-0A1 | Mellanox Passive Copper Cable VPI UP TO 56GB/S QSFP 1.5M |
| FDR | MC2207310-100 | Mellanox Active Fiber Cable VPI UP TO 56GB/S QSFP from 3M up to 100M |
| FDR | MC2207310-XXX | Mellanox Active Fiber Cable VPI UP TO 56GB/S QSFP from 3M up to 100M |
| FDR | MC2207312-XXX | Mellanox Active Fiber Cable VPI UP TO 56GB/S QSFP from 3M up to 300M |
| FDR | MC220731V-XXX | Mellanox® Active Fiber Cable, VPI, up to 56Gb/s, QSFP, up to 100m |
| FDR | MC2207411-SR4L | Mellanox Optical Module IB FDR 56GB/S QSFP MPO 850NM up to 30M |
| FDR | MCP170L-F001 | Mellanox® Passive Copper Cable, VPI, up to 56Gb/s, QSFP, LSZH, 1m |
| FDR | MCP170L-F002 | Mellanox® Passive Copper Cable, VPI, up to 56Gb/s, QSFP, LSZH, 2m |
| FDR | MCP170L-F003 | Mellanox® Passive Copper Cable, VPI, up to 56Gb/s, QSFP, LSZH, 3m |

1.2.10 Validated and Supported EDR Cables

Table 12 - Validated and Supported EDR Cables

| Speed | Cable OPN # | Description |
|-------|---------------------------|--|
| EDR | MCP1600-E001 ^a | Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 1M |
| EDR | MCP1600-E002 ^a | Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 2M |
| EDR | MCP1600-E003 | Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 3M |
| EDR | MCP1600-E004A26 | Mellanox® Passive Copper cable, IB EDR, up to 100Gb/s, QSFP28, 4m, Blue, 26AWG |
| EDR | MCP1600-E00A ^a | Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 0.5M |
| EDR | MCP1600-E01A ^a | Mellanox® Passive Copper cable, VPI, up to 100Gb/s, QSFP, LSZH, 1.5m |
| EDR | MCP1600-E02A | Mellanox® Passive Copper cable, VPI, up to 100Gb/s, QSFP, LSZH, 2.5m |
| EDR | MFA1A00-E005 ^a | MELLANOX Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 5m |

Table 12 - Validated and Supported EDR Cables

| Speed | Cable OPN # | Description |
|-------|---------------------------|--|
| EDR | MFA1A00-E010 ^a | MELLANOX Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 10m |
| EDR | MFA1A00-E015 ^a | MELLANOX Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 15m |
| EDR | MFA1A00-E020 | MELLANOX Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 20m |
| EDR | MFA1A00-E030 | MELLANOX Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 30m |
| EDR | MFA1A00-E050 | MELLANOX Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 50m |
| EDR | MFA1A00-E100 | MELLANOX Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 100m |
| EDR | MFS1200-E005 | Mellanox® Active Fiber Cable, IB EDR, up to 100Gb/s, QSFP, LSZH, 5m |
| EDR | MFS1200-E010 | Mellanox® Active Fiber Cable, IB EDR, up to 100Gb/s, QSFP, LSZH, 10m |
| EDR | MFS1200-E015 | Mellanox® Active Fiber Cable, IB EDR, up to 100Gb/s, QSFP, LSZH, 15m |
| EDR | MFS1200-E020 | Mellanox® Active Fiber Cable, IB EDR, up to 100Gb/s, QSFP, LSZH, 20m |
| EDR | MFS1200-E030 | Mellanox® Active Fiber Cable, IB EDR, up to 100Gb/s, QSFP, LSZH, 30m |
| EDR | MFS1200-E050 | Mellanox® Active Fiber Cable, IB EDR, up to 100Gb/s, QSFP, LSZH, 50m |
| EDR | MFS1200-E100 | Mellanox® Active Fiber Cable, IB EDR, up to 100Gb/s, QSFP, LSZH, 100m |
| EDR | MMA1B00-E100 | Mellanox® Transceiver, IB EDR, up to 100Gb/s, QSFP28, MPO, 850nm, up to 100m |

a. Forward Error Correction (FEC) is deactivated on this cable.

1.3 Tested Switches

1.3.1 Tested 10GbE Switches

Table 13 - Tested 10GbE Switches

| Speed | Switch Silicon | OPN # / Name | Description | Vendor |
|----------|----------------|--------------|----------------------------|---------|
| 10/40GbE | N/A | 3064 | 48-port 10Gb/40Gb Switch | Cisco |
| 10/40GbE | N/A | 7050Q | 16-port 40Gb Switch | Arista |
| 10/40GbE | N/A | 7050S | 48-port 10Gb/40Gb Switch | Arista |
| 10GbE | N/A | 5548 | Cisco 10GB ETH switch | Cisco |
| 10GbE | N/A | G8264 | BNT 10/40GB ETH switch | BNT |
| 10GbE | N/A | QFX3500 | Juniper 10/40GB ETH switch | Juniper |

Table 13 - Tested 10GbE Switches

| Speed | Switch Silicon | OPN # / Name | Description | Vendor |
|-------|----------------|--------------|-----------------------------|----------|
| 10GbE | N/A | S4810P-AC | 48-port 10Gb/40Gb Switch | Force10 |
| 10GbE | SwitchX® | SX1016X-1BFR | 64-Port 10GbE Switch System | Mellanox |

1.3.2 Tested 40GbE Switches

Table 14 - Tested 40GbE Switches

| Speed | Switch Silicon | OPN # / Name | Description | Vendor |
|----------|----------------|--------------|--------------------------------|----------|
| 10/40GbE | N/A | 3064 | 48-port 10Gb/40Gb Switch | Cisco |
| 10/40GbE | N/A | 7050Q | 16-port 40Gb Switch | Arista |
| 10/40GbE | N/A | 7050S | 48-port 10Gb/40Gb Switch | Arista |
| 40GbE | N/A | 3132Q | Cisco 40GB ETH switch | Cisco |
| 40GbE | N/A | 7050QX | 32-port 40Gb Switch | Arista |
| 40GbE | N/A | G8316 | BNT 40GB RackSwitch G8316 | BNT |
| 40GbE | N/A | S6000 | 32-port 40Gb Switch | Dell |
| 40GbE | SwitchX® | SX1036B-1BFR | 36-Port 40/56GbE Switch System | Mellanox |

1.3.3 Tested 100GbE Switches

Table 15 - Tested 100GbE Switches

| Speed | Switch Silicon | OPN # / Name | Description | Vendor |
|--------|----------------|-------------------|--|----------|
| 100GbE | N/A | 7060CX | 32-port 100Gb Switch | Arista |
| 100GbE | N/A | 93180YC-EX | 48 x 10/25-Gbps fiber ports and 6 x 40/100-Gbps Quad Small Form-Factor Pluggable 28 (QSFP28) ports | Cisco |
| 100GbE | N/A | C3232C | High-Density, 100 Gigabit Ethernet Switch | Cisco |
| 100GbE | N/A | CE8860-4C-EI | 24x10GE (SFP+) or 25GE (SFP28) and 2x100GE switch | Huawei |
| 100GbE | Spectrum | SN2410-CB2F | 48-port 25GbE + 8-port 100GbE Open Ethernet ToR Switch System | Mellanox |
| 100GbE | Spectrum | SN2700-CS2R | 32-port Non-blocking 100GbE Open Ethernet Spine Switch System | Mellanox |
| 100GbE | Spectrum | SN2740-CB2F1 | 32-port Non-blocking 100GbE Open Ethernet Spine Switch System | Mellanox |
| 100GbE | N/A | Wedge 100-32X R04 | 32-port 100G QSFP28 - Leaf/Spine Switch, power-to-port airflow, DC Power | Edgecore |

1.3.4 Tested QDR Switches

Table 16 - Tested QDR Switches

| Speed | Switch Silicon | OPN # / Name | Description | Vendor |
|-------|-----------------|--------------|---|----------|
| QDR | N/A | 12300 | 36-Port 40Gb QDR Infiniband Switch, Management Module, Dual Power | QLogic |
| QDR | InfiniScale® IV | IS5025Q-1SFC | 36-port 40Gb/s InfiniBand Switch Systems | Mellanox |
| QDR | InfiniScale® IV | Switch 4036 | Grid Director™ 4036E | Mellanox |

1.3.5 Tested FDR Switches

Table 17 - Tested FDR Switches

| Speed | Switch Silicon | OPN # / Name | Description | Vendor |
|-------|----------------|--------------|--|----------|
| FDR | SwitchX® | SX6018F-1SFR | 18-port 56Gb/s InfiniBand/VPI Switch Systems | Mellanox |
| FDR | SwitchX® | SX6036F-1BFR | 36-port 56Gb/s InfiniBand/VPI Switch Systems | Mellanox |
| FDR | SwitchX® | SX6506 | 108-Port 56Gb/s InfiniBand Director Switch | Mellanox |
| FDR | SwitchX®-2 | SX6710-FB2F2 | 36-port 56Gb/s InfiniBand/VPI Switch Systems | Mellanox |

1.3.6 Tested EDR Switches

Table 18 - Tested EDR Switches

| Speed | Switch Silicon | OPN # / Name | Description | Vendor |
|-------|----------------|--------------|--|----------|
| EDR | Switch-IB | SB7790-EB2F | 36-port EDR 100Gb/s InfiniBand Switch Systems | Mellanox |
| EDR | Switch-IB 2 | SB7800-ES2R | 36-port Non-blocking Managed EDR 100Gb/s InfiniBand Smart Switch | Mellanox |

1.4 Tools, Switch Firmware and Driver Software

Firmware Rev 16.22.1002 is tested with the following tools, Switch firmware, and driver software:

Table 19 - Tools, Switch Firmware and Driver Software

| | Supported Version |
|--------------------------------|-------------------------|
| MLNX_OFED | 4.3-1.0.1.0/4.2-1.2.0.0 |
| MLNX_EN (MLNX_OFED based code) | 4.3-1.0.1.0/4.2-1.2.0.0 |
| WinOF-2 | 1.90/1.80 |

Table 19 - Tools, Switch Firmware and Driver Software

| | Supported Version |
|------------------------------|--|
| MFT | 4.9.0/4.8.0 |
| VMware | <ul style="list-style-type: none"> ESXi 6.5 v4.16.12.12 ESXi 6.0 v4.15.12.12 |
| MLNX-OS | <ul style="list-style-type: none"> SwitchX: 3.6.4930 Switch-IB: 3.6.4930 Switch-IB 2: 3.6.4930 Spectrum: 3.6.4930 |
| SwitchX®/SwitchX®-2 Firmware | 9.4.4040 |
| Spectrum™ Firmware | 13.1530.0136 |
| SwitchX-IB™ Firmware | 11.1530.0136 |
| SwitchX-IB 2 Firmware | 15.1530.0136 |
| InfiniScale® V Firmware | 7.4.3000/v7.4.2200 |
| Linux Inbox Drivers | <ul style="list-style-type: none"> Ubuntu 14.04.3 Ubuntu 14.04.4 Ubuntu 15.04 Ubuntu 15.10 Ubuntu 16.04 Ubuntu 16.04.1 Ubuntu 16.10 SLES12 SLES12.1 SLES12.2 RHEL6.6 RHEL6.7 RHEL6.8 RHEL7.1 RHEL7.2 RHEL7.3 |
| Windows Inbox Drivers | Windows Server 2016 |

1.5 Supported FlexBoot, UEFI



Please be aware that not all firmware binaries contain FlexBoot or UEFI, support may vary between cards (see [Section 1.1, “Supported Devices”, on page 7.](#))

Firmware Rev 16.22.1002 supports the following FlexBoot:

Table 20 - Supported FlexBoot, UEFI

| Expansion ROM | Supported Version |
|---------------|-------------------|
| FlexBoot | 3.5.403 |
| UEFI | 14.15.19 |

1.6 Revision Compatibility

Firmware Rev 16.22.1002 complies with the following programmer's reference manual:

- *Mellanox Adapters Programmer's Reference Manual (PRM), Rev 0.45 or later*, which has Command Interface Revision 0x5. The command interface revision can be retrieved by means of the `QUERY_FW` command and is indicated by the field `cmd_interface_rev`.

2 Changes and New Features in Rev 16.22.1002

Table 21 - Changes and New Features in Rev 16.22.1002

| Feature/Change | Description |
|--|---|
| Disable SL/diff Flow | Added support for disable SL/diff flow to avoid performance degradation for single queue using multiple priorities. This functionality should not be used when DCB (PFC, ETS) is enabled. |
| Software Reset Flow | <p>Software Reset Flow enables the device to recover from fatal errors. The flow includes software detection of a fatal error, automatic creations of an mstdump file for future debug by the software, and resetting of the device. The feature is enabled using an mlxconfig command.</p> <p>Note: The flow is currently not supported on Multi host devices, Socket Direct devices and devices running management traffic (NCSI, MCTP).</p> |
| Steering Discard Packet Counters | <p>Any received packet which is dropped by the device is accounted for. To enable this functionality, the following counters were added to count the discard packets (per vport):</p> <ul style="list-style-type: none"> • <code>nic_receive_steering_discard</code>: Number of packets that completed the NIC Receive Flow Table steering, and were discarded because they did not match any flow in the final Flow Table. • <code>receive_discard_vport_down</code>: Number of packets that were steered to a VPort, and discarded because the VPort was not in a state to receive packets. • <code>transmit_discard_vport_down</code>: Number of packets that were transmitted by a vNIC, and discarded because the VPort was not in a state to transmit packets. |
| Pause Frame Duration and XOFF Resend Time | Increased the Pause Frame Duration and the XOFF Resend Time to the maximum value defined by the specification. |
| PCI Relax Ordering | <p>mlxconfig configuration can now enable or disable forced PCI relaxed ordering in mkey_context.</p> <p>If this feature is enabled, the software per mkey configuration is ignored.</p> |
| Push/Pop VLAN | <p>Added support for Push/Pop VLAN, new FLOW TABLE ENTRY actions. These new actions are used by the driver to implement Q-in-Q functionality.</p> <p>For further information, please refer to the PRM section <i>Flow Table</i></p> |
| Packet Pacing | <p>Added support for Packet Pacing in ConnectX-5 adapter cards. Packet Pacing (traffic shaping) is a rate-limited flow per Send QPs. A rate-limited flow is allowed to transmit a few packets before its transmission rate is evaluated, and the next packet is scheduled for transmission accordingly. Setting and changing the rate is done by modifying the QP.</p> <p>Note: Packet Pacing is not functional in ConnectX-5 Multi host adapter cards.</p> |

Table 21 - Changes and New Features in Rev 16.22.1002

| Feature/Change | Description |
|--|--|
| vport Mirroring | <p>Packets are mirrored based on certain mirroring policy. The policy is set using the “set FTE command” that supports forward action in the ACL tables (ingress/egress).</p> <p>The firmware support the following destination list format:</p> <ol style="list-style-type: none"> 1. new destination vport (analyzer) 2. another Flow Table <p>this way, the driver can forward the SX/RX packet related to the vport once it reaches the ACL table (forward it to the analyzer vport).</p> |
| Resiliency: Special Error Event | <p>Firmware uses error events to monitor the health of core transport engines, both Rx and Tx, and to detect if a system hang occurred and was not cured by other error mechanisms. Upon such detection, events are sent to the driver to perform any required action (e.g., software reset).</p> |
| QP’s Creation Time | <p>Accelerated QP’s creation time.</p> |
| SR-IOV LID based Routing Mode | <p>SR-IOV default routing mode is now LID based. The configuration change is available via mlxconfig tool. Note that in such mode, the VF will get its own LID, hence the GRH is not required.</p> <p>Note: LID based routing support for vports is supported using SM v4.8.1</p> |
| Expansion ROM | <p>Added PXE and UEFI to additional ConnectX- 5adapter cards. ConnectX-5 now holds PXE and x86-UEFI</p> |
| Host Chaining | <p>Host Chaining allows the user to connect ("chain") one server to another without going through a switch, thus saving switch ports.</p> <p>Host Changing algorithm is as follow:</p> <ul style="list-style-type: none"> • Received packets from the wire with DMAC equal to the host MAC are forwarded to the local host • Received traffic from the physical port with DMAC different than the current MAC are forwarded to the other port: <ul style="list-style-type: none"> • Traffic can be transmitted by the other physical port • Traffic can reach functions on the port's Physical Function • Device allows hosts to transmit traffic only with its permanent MAC • To prevent loops, the received traffic from the wire with SMAC equal to the port permanent MAC is dropped (the packet cannot start a new loop) <p>For Host Chaining limitation, see Known Issue #1178792 in Section 3, “Known Issues”, on page 24</p> |
| Fast path VLs | <p>Enabled fast path VLs which have lower latency (less than 2.55us) than slow path VLs. Fast path mapping can be configured using OpenSM configuration file.</p> |

Table 21 - Changes and New Features in Rev 16.22.1002

| Feature/Change | Description |
|---|--|
| Hairpin | Hairpin enables ingress traffic on the network port to egress on the same port or the 2nd port of the adapter. Hairpin enables hardware forwarding of packets from the receive queue to the transmit queue, thus fully offloading software gateways to the hardware. The queues can be allocated on different PCI functions, thus enabling packets' forwarding between different NIC ports. |
| Coherent Accelerator Processor Interface (CAPI v2) | The Coherent Accelerator Process Interface (CAPI) enables the user to attach a coherent accelerator to a Power and OpenPower based platforms. This solution delivers performance that exceeds today's I/O-attached acceleration engines. Note: This feature is available only with IBM Power 9 CPUs. |
| NVMe-oF Target Offload over DC transport | The NVMe-oF target offload provides the IO data path functionality of an NVMe over Fabrics Front-End subsystem transferring the IO operations to NVMe PCIe subsystems. |
| Bug Fixes | See Section 4, "Bug Fixes History" , on page 31 |

3 Known Issues

The following table describes known issues in this firmware release and possible workarounds. For a list of old firmware Know Issues, please see ConnectX-5 Firmware Archived Known Issues file

(http://www.mellanox.com/pdf/firmware/ConnectX5-Firmware_Archived_Known_Issues.pdf)

Table 22 - Known Issues (Sheet 1 of 7)

| Internal Ref. | Issue |
|---------------|---|
| 1316221 | Description: Health counter increases every 50ms instead of 10ms. |
| | Workaround: N/A |
| | Keywords: Health counter |
| | Discovered in Version: 16.22.1002 |
| 1316221 | Description: In very rare cases, triggering a function level reset while running NVMeF offload traffic might cause a response capsule that carries a bad command identifier of 0 to be sent. |
| | Workaround: N/A |
| | Keywords: NVMeF offload |
| | Discovered in Version: 16.22.1002 |
| 1298377 | Description: When a packet is sent on a non-native port, a LAG or a RoCE dual port, and it reaches the ingress mirroring entry, the packet sends the RX a meta data loopback syndrome, on the non-native port, resulting in the packet reaching the wrong meta_data table. |
| | Workaround: N/A |
| | Keywords: Steering, mirroring |
| | Discovered in Version: 16.22.1002 |
| 1306342 | Description: Signature-accessing WQEs sent locally to the NVMeF target QPs that encounter signature errors, will not send a SIGERR CQE. |
| | Workaround: N/A |
| | Keywords: Signature-accessing WQEs, NVMeF target |
| | Discovered in Version: 16.22.1002 |
| 1308236 | Description: Packet Pacing is not functional in ConnectX-5 Multi host adapter cards. |
| | Workaround: N/A |
| | Keywords: Packet Pacing, ConnectX-5 Multi host cards |
| | Discovered in Version: 16.22.1002 |
| 1309104 | Description: ParaVport is not supported in ConnectX-5. |
| | Workaround: N/A |
| | Keywords: ParaVport |
| | Discovered in Version: 16.22.1002 |

Table 22 - Known Issues (Sheet 2 of 7)

| Internal Ref. | Issue |
|-----------------|---|
| 1178792 | <p>Description: Host Chaining Limitations:</p> <ul style="list-style-type: none"> • MAC address must not be changed • Both ports should be configured to Ethernet when host chaining is enabled. • The following capabilities cannot function when host chaining is enabled: <ul style="list-style-type: none"> • SR-IOV • DSCP • NODNIC • Load balancing • LAG • Dual Port RoCE (multi port vHCA) <p>Workaround: N/A</p> <p>Keywords: Host Chaining</p> <p>Discovered in Version: 16.22.1002</p> |
| 1284452/1282926 | <p>Description: mlxconfig tool presents all possible expansion ROM images, instead of presenting only the existing images.</p> <p>Workaround: N/A</p> <p>Keywords: mlxconfig</p> <p>Discovered in Version: 16.22.1002</p> |
| 1277762 | <p>Description: An Ethernet multicast loopback packet is not counted (even if it is not a local loopback packet) when running the <code>nic_receive_steering_discard</code> command.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet multicast loopback packet</p> <p>Discovered in Version: 16.22.1002</p> |
| 1190753 | <p>Description: When a dual-port VHCA sends a RoCE packet on its non-native port, and the packet arrives to its affiliated vport FDB, a mismatch might happen on the rules that match the packet source vport.</p> <p>Workaround: N/A</p> <p>Keywords: Dual-port VHCA, RoCE packet, vport FDB</p> <p>Discovered in Version: 16.22.1002</p> |
| 1079027/1126921 | <p>Description: Occasionally, when adding module info page for Bell 1G BaseT module to the mlxlink data, the information is not updated correctly.</p> <p>Workaround: N/A</p> <p>Keywords: Bell 1G BaseT module, mlxlink</p> <p>Discovered in Version: 16.21.2010</p> |
| 1168594 | <p>Description: RoCE Dual Port Mode (a.k.a Multi-Port vHCA: MPV) is not supported in Multi-Host setups.</p> <p>Workaround: N/A</p> <p>Keywords: Multi-Port vHCA, Multi-Host</p> <p>Discovered in Version: 16.21.1000</p> |

Table 22 - Known Issues (Sheet 3 of 7)

| Internal Ref. | Issue |
|-----------------|---|
| 1133394 | Description: Running loopback and outbound traffics simultaneously in Multi-Host set-ups results in inbound/outbound traffic consuming more bandwidth than loopback traffic. |
| | Workaround: N/A |
| | Keywords: Loopback and outbound traffic, bandwidth |
| | Discovered in Version: 16.21.1000 |
| 1171013 | Description: Signature Handover Operations is not supported when FPP (Function-Per-Port) mode is disabled. |
| | Workaround: N/A |
| | Keywords: Signature Handover Operations, FPP |
| | Discovered in Version: 16.21.1000 |
| 1176407/1171665 | Description: Secure Firmware devices block any "not secure" access to the CR-Space, private ICMDs and Flash random access commands. Thus, tools/capabilities such as ibdump or wqedump or packet sniffing do not function properly. |
| | Workaround: Load a secure customer token to use any blocked tools. |
| | Keywords: Secure Firmware |
| | Discovered in Version: 16.21.1000 |
| 1119458 | Description: When RoCE Dual Port Mode is enable, if VHCA0 has a different VLAN stripping configuration on its E-SW vport context than the configuration on the affiliated vport (VHCA1), the NIC steering on the VLAN might be wrong for the single-port VHCA traffic. |
| | Workaround: N/A |
| | Keywords: VLAN, dual port affiliation |
| | Discovered in Version: 16.21.1000 |
| 1145910 | Description: Using E-SW VLAN insertion with a different priority than the QP's native priority might cause QoS issues. Issues might be observed while using PUSH VLAN action in Flow Steering. |
| | Workaround: N/A |
| | Keywords: E-SW VLAN Insertion |
| | Discovered in Version: 16.21.1000 |
| 1163425 | Description: Running mlxfwreset on ConnectX-5 Socket-Direct adapter cards on Windows OS is currently not functional. |
| | Workaround: Reboot the server |
| | Keywords: mlxfwreset, ConnectX-5 Socket-Direct |
| | Discovered in Version: 16.21.1000 |

Table 22 - Known Issues (Sheet 4 of 7)

| Internal Ref. | Issue |
|---------------|--|
| 1043521 | Description: Upgrading old ConnectX-5 firmware using MFT tools version 4.8.0 and above, requires a one time, non failsafe operation. If this is performed, a server reboot is needed (not mlxconfig). Note: Not rebooting the server after the upgrade will result in future updates done using mlxconfig to be lost (until server reboot). |
| | Workaround: N/A |
| | Keywords: MFT, Firmware Update |
| | Discovered in Version: 16.21.1000 |
| 912315 | Description: In standby (WoL) mode, although the total IC consumption is as set, the actual current consumption in 3.3V rail is higher in 20-40mA than the set values. Consumption will be decreased in the next firmware release. |
| | Workaround: N/A |
| | Keywords: WoL, IC consumption |
| | Discovered in Version: 16.21.1000 |
| 1031669 | Description: Using Flow Tag in Flow Table Entry together with Header Rewrite is currently not supported. |
| | Workaround: N/A |
| | Keywords: Header Rewrite, Flow Tag |
| | Discovered in Version: 16.21.1000 |
| 1122414 | Description: In a Multi-Host setup, in case of multicast traffic, if a host modifies the packet's format, other hosts might receive the modified packet instead of the original packet. |
| | Workaround: N/A |
| | Keywords: Multi-Host, multicast traffic |
| | Discovered in Version: 16.21.1000 |
| 1122452 | Description: After switching to CAPI mode the driver cannot be unloaded and a server reboot through the OS might hang the server. |
| | Workaround: To restart the server perform a power cycle through OpenBMC. |
| | Keywords: CAPI |
| | Discovered in Version: 16.21.1000 |
| 1110219 | Description: In some cases, CAPI page miss flow is not resolved. |
| | Workaround: N/A |
| | Keywords: CAPI |
| | Discovered in Version: 16.21.1000 |

Table 22 - Known Issues (Sheet 5 of 7)

| Internal Ref. | Issue |
|-----------------|--|
| 1068382 | Description: NVMeF target offload is not supported when the device is configured to support more than 32 functions (VFs + PFs). |
| | Workaround: N/A |
| | Keywords: NVMeF target offload, VFs, PFs |
| | Discovered in Version: 16.21.1000 |
| 1072337 | Description: If a packet is modified in e-sw flow steering, the SX sniffer Flow Table (of the VF) will see the sniffed packet after the modification. |
| | Workaround: N/A |
| | Keywords: SX sniffer Flow Table |
| | Discovered in Version: 16.21.1000 |
| 1070650 | Description: A multicast packet that arrives from e-sw loopback (VF2VF) and hits some FTE with MODIFY action in the FDB is sent to the source vport (for local loopback) with the modified headers. |
| | Workaround: N/A |
| | Keywords: Multicast packet |
| | Discovered in Version: 16.20.1010 |
| 1077244 | Description: In the following cases, "rx_buffer_passed_thres_phy" would not indicate fullness even when "rx_discards_phy" indicates a drop: <ul style="list-style-type: none"> • For single port devices: <ul style="list-style-type: none"> • Flow control is not enabled on the buffer • fullness_threshold is configured to 99% • MTU size smaller than 4K • For dual port devices: <ul style="list-style-type: none"> • Flow control is not enabled on the buffer • fullness_threshold is configured to 99% when MTU size is smaller than 6.5K • fullness_threshold is configured to 98% when MTU size is smaller than 4K |
| | Workaround: N/A |
| | Keywords: rx_buffer_passed_thres_phy |
| | Discovered in Version: 16.20.1010 |
| 1090492 | Description: FLR is not supported when a function receives NVMeF traffic on the target side. |
| | Workaround: N/A |
| | Keywords: FLR, NVMeF |
| | Discovered in Version: 16.21.1000 |
| 1095081/1093055 | Description: Latency sensitive is not supported in NVMeF and Tag Matching offload QP. Note: Enabling this feature can cause the machine to hang. |
| | Workaround: N/A |
| | Keywords: NVMeF, Tag Matching, Latency sensitive |
| | Discovered in Version: 16.21.1000 |

Table 22 - Known Issues (Sheet 6 of 7)

| Internal Ref. | Issue |
|-----------------|---|
| 1114798 | Description: Tag Matching DC transport does not support GRH. |
| | Workaround: N/A |
| | Keywords: Tag Matching DC, GRH |
| | Discovered in Version: 16.21.1000 |
| 1086254/1090475 | Description: Packet header rewrite (modification) and packet encapsulation are not supported when one ConnectX-5 [Ex] VPI port is configured as InfiniBand and the other as Ethernet. |
| | Workaround: N/A |
| | Keywords: ConnectX-5 VPI; ConnectX-5 Ex VPI; InfiniBand; Ethernet; Ports; Header Rewrite, Packet encapsulation |
| | Discovered in Version: 16.20.1010 |
| 1063904 | Description: Messages with mkey signature on offset > 4GB are not supported. |
| | Workaround: N/A |
| | Keywords: Signature retransmission |
| | Discovered in Version: 16.20.1010 |
| 1063148 | Description: Pause duration: Physical port counters count in 512bits quantas, instead of microseconds. |
| | Workaround: To normalize the counter, do not change the speed: $\text{counter_value_in_microsec} = \text{current_counter_value} * 512 / \text{port_speed}$ |
| | Keywords: Pause duration, Physical port counters |
| | Discovered in Version: 16.20.1010 |
| 1059975 | Description: NVMeF limitation: <ul style="list-style-type: none"> • Transaction size - up to 128KB per IO (non-inline) • Support up to 16K connections • Support single namespace per drive • Staging buffer size must be at least 16MB in order to allow SRQ size of 64 entries |
| | Workaround: N/A |
| | Keywords: NVMeF |
| | Discovered in Version: 16.20.1010 |
| 1059975 | Description: When using NVMeF, DESTORY XRQ command failure might occur. |
| | Workaround: N/A |
| | Keywords: NVMeF |
| | Discovered in Version: 16.20.1010 |

Table 22 - Known Issues (Sheet 7 of 7)

| Internal Ref. | Issue |
|---------------|--|
| 1031744 | Description: Same flow counter cannot be used on different table types. |
| | Workaround: N/A |
| | Keywords: Flow counter |
| | Discovered in Version: 16.20.1010 |

4 Bug Fixes History

Table 23 lists the bugs fixed in this release. For a list of old firmware Bug Fixes, please see ConnectX-5 Firmware Archived Bug Fixes file

http://www.mellanox.com/pdf/firmware/ConnectX5-Firmware_Archived_Bug_Fixes.pdf

Table 23 - Bug Fixes History (Sheet 1 of 7)

| Internal Ref. | Issue |
|---------------------|---|
| 1231791 | Description: Fixed an issue that caused the driver to return a wrong logical OR of the 2 physical ports, when querying the vport state when the LAG was enabled. |
| | Keywords: LAG, vport |
| | Discovered in Version: 16.21.2010 |
| | Fixed in Release: 16.22.1002 |
| 1252833 | Description: Increased the Full Wire Speed (FWS) threshold value to improve EDR link results. |
| | Keywords: Full Wire Speed (FWS) threshold, EDR |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.22.1002 |
| 1260985 | Description: Added the option to avoid reconfiguration of QoS tables upon link toggling to reduce packet loss and improve performance. |
| | Keywords: ECN, QoS |
| | Discovered in Version: 16.21.2010 |
| | Fixed in Release: 16.22.1002 |
| 1262477 | Description: Fixed an issue that caused traffic to hang when Responder Not Ready (RNR) flow was used. |
| | Keywords: RoCE Lossy, ECN |
| | Discovered in Version: 16.21.2010 |
| | Fixed in Release: 16.22.1002 |
| 1080868/ 1109484 | Description: Tag Matching supports up to 16K connections. |
| | Keywords: Tag Matching |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.22.1002 |
| 1084581 | Description: Target NVMeoF offload for 4 SSDs are 950K IOPS in ConnectX-5 Ex. |
| | Keywords: Performance |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.22.1002 |

Table 23 - Bug Fixes History (Sheet 2 of 7)

| Internal Ref. | Issue |
|---------------------|---|
| 1096454 | Description: The HCA does not always identify correctly the presets at the 8G EQ TS2 during speed change to Gen4. As a result, the initial Gen4 Tx configuration might be wrong which might cause speed degrade to Gen1. |
| | Keywords: Gen4 TX configuration |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.22.1002 |
| 1281622 | Description: Fixed an issue that resulted in “Destroy LAG” command failure if a VFs received an FLR while its affinity QPs were open. |
| | Keywords: ECMP / SR-IOV LAG |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.22.1002 |
| 1172293 | Description: When RoCE Dual Port mode is enabled, tcpdump is not functional on the 2nd port. |
| | Keywords: Dual Port vHCA, Multi-port, RoCE Dual Port |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.22.1002 |
| 1184961/ 1076206 | Description: Fixed an issue that occasionally cased the keepalive packet to fail and the FIO connection to disconnect (error =5). |
| | Keywords: FIO, NVMeF |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.22.1002 |
| 1124226 | Description: Fixed an issue that caused QP connection timeout due to firmware not being able to handle duplicate packets with AckReq bit set. The fix stopped ignoring duplicate AckReq packets to avoid timeout on the sender side. |
| | Keywords: QP connection timeout |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.21.2010 |
| 1179155 | Description: MPFS load balance (DUP_MAC_ACTION==LOAD_BALANCE(1)) is not working as long as IB_ROUTING_MODE/SRIOV_IB_ROUTING_MODE is configured to LID. |
| | Keywords: MPFS load balance, LID, GID |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.21.2010 |

Table 23 - Bug Fixes History (Sheet 3 of 7)

| Internal Ref. | Issue |
|---------------------|---|
| 1155392 | Description: Fixed an issue that caused a SX engine deadlock (the SX engine handles software port/priority changes for a specific Send Queue) when more than a single SX engine handled the "prio diff" flow simultaneously, thus caused the hardware to get stuck. The issue happened as the firmware releases the SX engine and waits for various operation to complete. However, due to a race that allows a different transaction to get into the SX engine and cause the lock to be taken by it, the SX engine release is prevented. |
| | Keywords: SX engine deadlock, system hang |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.21.2010 |
| 1168271 | Description: Fixed an issue that caused the system to hang while changing QPTS/QPDPM/QPDP parameters during traffic. |
| | Keywords: QPTS/QPDPM/QPDP, system hang |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.21.2010 |
| 1190215 | Description: Fixed unfairness between senders in RoCE LAG while ECN is configured. |
| | Keywords: RoCE LAG, ECN |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.21.2010 |
| 1121688 | Description: Fixed an issue which displayed diagnostic counters only on the adapter that was initialized first, which is the counters' owner. The owner received correct values, while the other adapter only received zeros. |
| | Keywords: Diagnostic counters |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.21.2010 |
| 1175146 | Description: Fixed an issue that caused the rdma_cm traffic to fail on the 2nd port when more than 32 VFs were configured, when the RoCE Dual Port vHCA (a.k.a Multi-Port vHCA: MPV) feature was enabled. |
| | Keywords: MPV |
| | Discovered in Version: 16.21.1000 |
| | Fixed in Release: 16.21.2010 |
| 1167218/ 1168567 | Description: Fixed an issue related to RDMA_CM driver that might have caused the QP Rate Limit to be activate unexpectedly and reduce the bandwidth significantly on this QP. |
| | Keywords: Performance |
| | Discovered in Version: 16.20.1010 |
| | Fixed in Release: 16.21.1000 |

Table 23 - Bug Fixes History (Sheet 4 of 7)

| Internal Ref. | Issue |
|---------------------|---|
| 1122718 | Description: Fixed an issue that caused low throughput when ECN was enabled in a many-to-one scenario. |
| | Keywords: ECN |
| | Discovered in Version: 16.20.1010 |
| | Fixed in Release: 16.21.1000 |
| 1149487 | Description: Fixed an issue that caused ConnectX-5 Ex Virtual Function to be recognized as a ConnectX-5 Virtual Function device. |
| | Keywords: Virtualization |
| | Discovered in Version: 16.20.1010 |
| | Fixed in Release: 16.21.1000 |
| 1090723 | Description: Fixed an issue that wrongly reported the maximum temperature in a setup as the current temperature regardless of the actual temperature. |
| | Keywords: PCI Gen4 receiver |
| | Discovered in Version: 16.20.1010 |
| | Fixed in Release: 16.21.1000 |
| 1134407 | Description: PCI Gen4 receiver stability enhancements. |
| | Keywords: PCI Gen4 receiver |
| | Discovered in Version: 16.20.1010 |
| | Fixed in Release: 16.21.1000 |
| 1099880 | Description: Disabled the option to write to the protected modules to avoid receiving NACK upon module initialization. |
| | Keywords: Thermal temperature |
| | Discovered in Version: 16.20.1010 |
| | Fixed in Release: 16.21.1000 |
| 1014078 | Description: Enabled connecting 5m 40GbE cables to SwitchX, SwitchX-2 based switches. |
| | Keywords: Cables, SwitchX, SwitchX-2 |
| | Discovered in Version: 16.19.1200 |
| | Fixed in Release: 16.21.1000 |
| 1054335/ 1054671 | Description: Fixed the issue where when using UD RoCE multicast traffic over SR-IOV, packets were scattered to all the attached QPs in the e-sw (PF and its VFs) and not only on the vport that was specified in the e-se FDB. |
| | Keywords: UD RoCE multicast traffic, SR-IOV |
| | Discovered in Version: 16.20.1010 |
| | Fixed in Release: 16.21.1000 |

Table 23 - Bug Fixes History (Sheet 5 of 7)

| Internal Ref. | Issue |
|---------------|--|
| 1060650 | Description: Fixed a link issue on Intel 10GbE Optical module PN: R8H2F, Y3KJN. |
| | Keywords: Intel 10GbE Optical module |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 1047533 | Description: Fixed an issue that caused the TX traffic not to send packets when using VF index (ARI) bigger than 127. |
| | Keywords: VFs |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 1009614 | Description: Fixed a scaling issue with more than 1k QPs for ECN by moving from per QP caching to per IP to allow better scale with number of host in the fabric. |
| | Keywords: Performance |
| | Discovered in Release: 16.19.1200 |
| | Fixed in Release: 16.20.1010 |
| 1041108 | Description: Enabled firmware resync of the internal clocks after getting out of the standby mode to prevent PTP time sync from getting out of sync after system warm-rebooted due to system getting into a low-power (standby) mode. |
| | Keywords: PTP time sync, standby mode |
| | Discovered in Release: 16.19.1200 |
| | Fixed in Release: 16.20.1010 |
| 1047693 | Description: When running RoCE over VRRP, enabled the device to receive RoCE packet with different source MAC than the original RoCE packet's destination MAC, to allow routing between different subnets. |
| | Keywords: RoCE over VRRP, Destination MAC |
| | Discovered in Release: 16.19.1200 |
| | Fixed in Release: 16.20.1010 |
| 1050234 | Description: Fixed an issued that caused LLDP not to enable PFC configuration currently when DCBX transitioning flow control configurations was set from Global Pause to PFC. |
| | Keywords: RoCE Lossy & ECN |
| | Discovered in Release: 16.19.1200 |
| | Fixed in Release: 16.20.1010 |
| 1047533 | Description: Rephrased and improved external troubleshoot messages in PDDR register. |
| | Keywords: PDDR register |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |

Table 23 - Bug Fixes History (Sheet 6 of 7)

| Internal Ref. | Issue |
|--------------------------------|--|
| 999261 | Description: Improved SR-IOV performance. |
| | Keywords: SR-IOV |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 954822 | Description: The <code>ipoib_enhanced_offloads</code> indication in the HCA capabilities reports 0 while <code>SRIOV_EN=1</code> . |
| | Keywords: SR-IOV, IPoIB Offloads |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 981598 | Description: Fixed an issue on an ETH port with SR-IOV enabled that prevented packets from reaching the BMC (failure in steering loopback resolution) if the BMC addresses were configured after VF initialization, and the VF was trying to send traffic to the BMC (that located on the same phy port). |
| | Keywords: BMC, SR-IOV, packets |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 906144 | Description: Fixed an issue which caused the rate limiter not to function when setting a rate to <code>tc 7</code> . |
| | Keywords: QOS - ETH - rate limit per TC |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 893261 | Description: Fixed the PCIe TX glitch during Recovery.Speed state of the link training to PCIe Gen3. |
| | Keywords: PCIe TX glitch |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 1002190 | Description: Fixed an issue related to the <code>PortRcvDataVLExtended/PortXmitDataVLExtended</code> parameter that caused the counters' value to be reported in octets instead of dwrods. |
| | Keywords: Counters |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 1025741/ 781339/ 1050373 | Description: QP ULP modes 0 and 1 cannot be assigned to the same Multicast group. |
| | Keywords: Multicast Group (MCG), QPs |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |

Table 23 - Bug Fixes History (Sheet 7 of 7)

| Internal Ref. | Issue |
|--------------------------------|--|
| 913451 | Description: Fixed an issue in standby (WoL) modes only that caused the actual current consumption in 1.2V rail to be higher by <33mA than the advertised values although the total IC consumption is as advertised. |
| | Keywords: Standby (WoL) modes, current consumption |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 852744 | Description: Mapping an SL to VL 15 is currently not supported. Trying to do so, will cause a health buffer fatal internal error report. |
| | Keywords: SL to VL mapping |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 937318 | Description: Setting more than 8K QPs with hca_sq_owner == 1 connected to an RNDV XRQ is currently not supported. |
| | Keywords: QPs, RNDV XRQ |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 860574/ 860716 | Description: Fixed performance issues to improve Packet Pacing performance. |
| | Keywords: Performance, Packet Pacing |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 1019003/ 1019039/ 995878 | Description: Fixed an issue causing physical errors observed on Ixia 100GbE receiver. |
| | Keywords: Ixia 100GbE receiver |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |
| 954826 | Description: Fixed an issue that caused the ipoib_enhanced_offloads indication (in the HCA capabilities) to report 0 while SRIOV_EN=1. |
| | Keywords: ipoib_enhanced_offloads indication |
| | Discovered in Release: 16.18.1000 |
| | Fixed in Release: 16.20.1010 |

5 Firmware Changes and New Feature History

Table 24 - Firmware Changes and New Feature History (Sheet 1 of 11)

| Feature/Change | Description |
|--|--|
| Rev. 16.21.2010 | |
| Query vPort Environments (Debug Counters) | <p>Debug counters are a group of counters that handle traffic performance issue related to firmware overhead in transport flow.</p> <p>The following are the additional counters added to this firmware version:</p> <ul style="list-style-type: none"> • current_q_under_processor_handle • total_q_under_processor_handle • qp_priority_update_flow |
| Address Ordering | Enables address ordering on ConnectX-5 Multi-Host adapter cards. It allows PCIe READ transaction to bypass the PCIe WRITE transaction, when both transactions are not on the same page. |
| Bug Fixes | See Section 4, “Bug Fixes History”, on page 31 |
| Rev. 16.21.1000 | |
| Coherent Accelerator Processor Interface (CAPI v2) | <p>[Beta] The Coherent Accelerator Process Interface (CAPI) enables the user to attach a coherent accelerator to a Power and OpenPower based platforms. This solution delivers performance that exceeds today’s I/O-attached acceleration engines.</p> <p>Note: This feature is available only with IBM Power 9 CPUs.</p> |
| NVME-oF Target Offload | Added support for NVMe over Fabrics (NVME-oF) offload, an implementation of the new NVME-oF standard target (server) side in the hardware. |
| Tag Matching RC/DC transport | <p>Added support for Tag Matching Offload with RC/DC transport.</p> <p>In Tag Matching, the software holds a list of matching entries called matching list. Each matching entry contains a tag and a pointer to an application buffer. The matching list is used to steer arriving messages to a specific buffer according to the message tag. The action of traversing the matching list and finding the matching entry is called tag matching.</p> <p>For further information, refer to the PRM section “<i>Tag Matching and Rendezvous Offload</i>”</p> |
| Hairpin | <p>[Beta] Hairpin enables ingress traffic on the network port to egress on the same port or the 2nd port of the adapter.</p> <p>Hairpin enables hardware forwarding of packets from the receive queue to the transmit queue, thus fully offloading software gateways to the hardware. The queues can be allocated on different PCI functions, thus enabling packets’ forwarding between different NIC ports.</p> |
| Virtual Extensible LAN (VXLAN) Encapsulation Offloads over RDMA in SR-IOV | <p>Added support for VXLAN encapsulation offloads over RDMA in SR-IOV.</p> <p>Virtual Extensible LAN (VXLAN) is a network virtualization technology that improves scalability problems associated with large cloud computing deployments. It tunnels Ethernet frames within Ethernet + IP + UDP frames.</p> |

Table 24 - Firmware Changes and New Feature History (Sheet 2 of 11)

| Feature/Change | Description |
|---|---|
| Flow Table Entries Enhancements | Enabled adapter support for up to 1 million Flow Table Entries. For further information, please refer to the PRM section <i>Flow Table</i> |
| Hardware Accelerated 802.1ad VLAN (Q-in-Q Tunneling) | Q-in-Q tunneling allows the user to create a Layer 2 Ethernet connection between two servers. The user can segregate a different VLAN traffic on a link or bundle different VLANs into a single VLAN. |
| Memory Mapped to InterConnect (MEMIC) | Added support for locked ICMC data buffer in order to improve latency by saving the PCI 'round trip'. For further information, refer to the PRM. |
| QoS: Tx Rate Limiter | Added support for VF rate limiter and bandwidth share in ConnectX-5. |
| Single PF for InfiniBand Dual Port Device | Added support for InfiniBand native (No SR-IOV) dual port device (Function per port is disabled). In this mode virtualization is not supported and ISSI = 0. |
| Explicit Congestion Notification (ECN) | Enabled ECN by default. |
| RoCE Dual Port Mode (a.k.a Multi-Port vHCA: MPV) | Enables the usage of a dual port Virtual HCA (vHCA) to share RDMA resources (e.g., MR, CQ, SRQ, PDs) across the two Ethernet (RoCE) NIC network ports and display the NIC as a dual port device. For this feature to function properly, the following requirements must be met: <ul style="list-style-type: none"> • Either the LAG or the Dual Port mode is enabled by the driver • Dual port device: both ports must be set as ETH • In ConnectX-4/ConnectX-4 Lx adapter cards, the maximum allowed number of VFs per PF is 32. • Function per port is enabled Note: This feature is only supported in single host device |
| DSCP | Added QPDPM register to support dynamic mapping between DSCP and priority. |
| | Added trust level for QoS prioritization according to the DSCP or PCP. |
| | Added ingress buffer management for: <ul style="list-style-type: none"> • ingress traffic mapping to a buffer according to priority • buffers sizes and lossless parameters |
| Steering Rules Rate Improvement | Improved steering rules update rate to up to 50K rules per sec. |
| Windows SR-IOV Enhanced eIPoIB | Enabled Windows SR-IOV Enhanced eIPoIB (without Secure Connection) for Windows-over-Windows setups. |
| Driver CR Dump | crdump operation takes a snapshot of the device's crspace dword-by-dword. It enables the driver to collect debug information upon firmware failure. |

Table 24 - Firmware Changes and New Feature History (Sheet 3 of 11)

| Feature/Change | Description |
|---|---|
| Secured Firmware Update | Secure Firmware Updates provides devices with the ability to verify digital signatures of new firmware binaries, in order to ensure that only officially approved versions are installed on the devices. Note: This feature is only available in adapter cards that support this feature. |
| Cables | Changed the default FEC mode for cables with attenuation 16 and below from RS to FC. |
| ECN | Enabled ECN (CongestionControl) by default for all priorities on Ethernet ports. |
| Bug Fixes | See Section 4, “Bug Fixes History” , on page 31 |
| Rev. 16.20.1010 | |
| NVME-oF Target Offload | [Beta] Performance and stability improvements. |
| Tag Matching | [Beta] Performance and stability improvements. For further information, refer to the PRM section “Tag Matching and Rendezvous Offload” |
| Tag Matching DC | [Alpha] Added support for Tag Matching DC (RNDV/EAGER traffic) |
| Adaptive Routing (Out-Of-Order) | Added support for send/receive Out-of-Order RDMA packets required by the Adaptive Routing. Adaptive Routing (AR) enables the switch to select the output port based on the port's load. |
| DSCP | Added trust level for QoS prioritization according to the DSCP or PCP. [Beta] Added ingress buffer management for: <ul style="list-style-type: none"> • ingress traffic mapping to a buffer according to priority • buffers sizes and lossless parameters |
| Secured Firmware Updates | [Beta] Secure Firmware Updates provides devices with the ability to verify digital signatures of new firmware binaries, in order to ensure that only officially approved versions are installed on the devices. Note: This feature is only available in adapter cards that support this feature. |
| Multi-Host/Socket Direct Routing to be LID based | [InfiniBand only] Changed the Multi-Host/Socket Direct routing to be LID based instead of GID based. Thus, GRH/GID index is not required. Note: This feature requires SM 4.8.1 and above. |
| Relaxed Ordering | [Beta] Added support for relaxed ordering write in memory keys. |
| RDMA Counters | Enhanced RDMA counter |
| TLV for PCI class code | Added 2 new per Host TLVs (see Table 33, “Per host Settings,” on page 60) |

Table 24 - Firmware Changes and New Feature History (Sheet 4 of 11)

| Feature/Change | Description |
|---|---|
| Header Rewrite | Modifies the packets header. |
| Fast Teardown | Enables fast unloading driver by using Teardown HCA with op_mode=1 (force_close). For further information, refer to the PRM. |
| IPoIB Virtualization | Added support for enhanced IPoIB (QP.ulp == 2) in virtualized system (SR-IOV / Multi-Host / Socket Direct) |
| SFP Power Flow Improvement (level 2,1) | Added support for SFP power class. |
| Bug Fixes | See Section 4, “Bug Fixes History” , on page 31 |
| Rev. 16.19.1200 | |
| General | This is the first GA version of the ConnectX-5/Ex adapter cards. |
| Bug Fixes | See section Section 4, “Bug Fixes History” , on page 31 |
| Rev. 16.18.2000 (Beta) | |
| Virtual Extensible LAN (VXLAN) encapsulation/decapsulation | [Beta Level] Virtual Extensible LAN (VXLAN) is a network virtualization technology that improves scalability problems associated with large cloud computing deployments. It tunnels Ethernet frames within Ethernet + IP + UDP frames. Mellanox implements VXLAN encapsulation and decapsulation in the hardware. |
| NVME-oF Target Offload | [Beta Level] Added support for NVMe over Fabrics (NVME-oF) offload, an implementation of the new NVME-oF standard target (server) side in hardware. |
| Tag Matching | [Beta Level] Added support for offloading MPI tag matching to HCA. It matches the source send operations to the destination receivers. |
| VLAN Switch Tagging (VST) | Enables the virtual machine interface to have no VLAN tag over it, thus allowing VLAN tagging to be handled by the Hypervisor. |
| On Demand Paging (ODP) | Added supported for Demand Paging (ODP). |
| Rev. 16.18.1000 (Beta) | |
| NVM Express over Fabrics (NVMeF) | NVMf is a protocol for communicating block storage IO requests over RDMA. For further information, please refer to the PRM section <i>“NVMe over Fabric Target Application Offload (NVMeF)”</i> . |
| Tag Matching | In Tag Matching, the software holds a list of matching entries called matching list. Each matching entry contains a tag and a pointer to an application buffer. The matching list is used to steer arriving messages to a specific buffer according to the message tag. The action of traversing the matching list and finding the matching entry is called tag matching. For further information, refer to the PRM section <i>“Tag Matching and Rendezvous Offload”</i> |
| RX Loss (BaseT link down indication) | Added logical link indication in SFP to BaseT modules and disabled logical link when peer port is down. |
| SFP Rate | Added support for 10GbE in 25GbE SFP optical modules |

Table 24 - Firmware Changes and New Feature History (Sheet 5 of 11)

| Feature/Change | Description |
|--|--|
| PDDR | Enables mlxlink tool to collect data on the PHY link status and provides link down reasons and additional link related information. |
| KR Tx Response | Enabled TX configuration response and movement during Link Training in Ethernet. |
| Phy Test mode | Added support at lane rate of 12.89Gb. |
| Performance | Improved performance for Send Queues (SQs) transmitting multiple priorities in a single Traffic Class (TC) configuration. |
| Droptless TCP | Added the ability to avoid packet drops due to temporary lack of posted Receive buffers (WQEs), for trusted Receive Queues (RQs). |
| Head of Queue (HoQ) per TC | Limits the amount of time a packet may head a Traffic Class (TC) transmission queue, without being transmitted. Stale packets are discarded. Active by default for TCs adhering to link level flow control |
| User Access Region (UAR) 4KB Granularity Allocation | UAR page size currently is set to 4KB and not according to what the system page size determines. |
| No Driver NIC (NODNIC) Performance Improvement | Improved performance of: <ul style="list-style-type: none"> • Doorbell from User Access Region (UAR) • Clear interrupt from User Access Region (UAR) |
| Counters | Added support for additional transport counters. |
| Scatter to CQE on Sender for DC | Enabled scatter-to-CQE for sent packets for DC. |
| CQ modify | Enabled moderation period modification in CQ modify command. |
| Network traffic between UEFI-Shell and OS | Enabled network traffic between UEFI-Shell and OS. |
| non-RDMA capable VFs | Enabled the PF to force disable RoCE for its VFs. |
| Loopback Enabled/Disabled | Enabled VNIC the control to enable/disable its local loopback traffic. |
| RDMA RX Flow Table | Added the option to open a receive RDMA Flow Table and to forward RoCE traffic to some destination QP. |
| GENEVE & IP-in-IP Stateless Offload | Added support for IP-in-IP and GENEVE network protocols encapsulated into IP frame (L2 tunneling). Encapsulation is suggested as a means to alter the normal IP routing for datagrams, by delivering them to an intermediate destination that would otherwise not be selected based on the (network part of the) IP Destination Address field in the original IP header. Note: For driver support, please see the Release Notes/User Manual of the relevant OS driver. |
| Resilient RoCE | Resilient RoCE is the ability to send RoCE traffic over a lossy network (a network without flow control enabled), without the need to enable flow control on the network. The ability is accomplished by enabling ECN on both the Switch and the Host. |

Table 24 - Firmware Changes and New Feature History (Sheet 6 of 11)

| Feature/Change | Description |
|--|--|
| Power MGMT | <p>Added support for PCIe Express standard "Slot capability register" message (PCIe base rev 3.1, section 6.9 – "Slot Power Limit Control")</p> <p>When ConnectX-5 Ex based adapter is inserted to a PCIe slot that supports the reporting of the slot power limit control, the ConnectX-5 Ex may disable the 2nd port if PCIe slot message, showing that the power in this slot is insufficient. If not, both ports will stay enabled.</p> <p>In cases where ConnectX-5 Ex based adapter is inserted to a PCI slot that DOES NOT support the "Slot capability register" message, and the adapters' configuration is 2 active optical cables/ transceivers, only one port will be enabled (the first inserted optic).</p> <p>Custom and OEM branded card based on ConnectX-5 Ex may be configured by INI to support/not-support the Power management feature.</p> <p>In hosts which do not support the "Slot capability register" Message and have enough power to support 2 active optical cable, the user will have the option to override the configuration resulted from "Slot capability register" by running the following NVconfig command:</p> <ul style="list-style-type: none"> • <code>echo "MLNX_RAW_TLV_FILE" > /tmp/power_conf_tlv.cfg;</code> • <code>echo "0x00000004 0x00000088 0x00000000 0xc0000000" >> /tmp/power_conf_tlv.cfg</code> • <code>mlxconfig -d <device> -f /tmp/power_conf_tlv.cfg set_raw</code> • <code>mlxfwreset -d <device> reset</code> <p>For details on ConnectX-5 Ex power, please refer to ConnectX-5 Ex Datasheet</p> |
| Virtual Functions (VF) per Port | <p>The maximum Virtual Functions (VF) per port is 64.</p> <p>Note: When increasing the number of VFs, the following limitations must be taken into consideration:</p> <pre>server_total_bar_size >= (num_pfs)*(2log_pf_uar_bar_size + 2log_vf_uar_bar_size*total_vfs) server_total_msix >= (num_pfs)*(num_pf_msix + num_vfs_msix *total_vfs)</pre> <p>Note: For the maximum number of VFs supported by your driver, please refer to your drivers' Release Notes or User Manual.</p> |
| QoS per VFs | <p>[InfiniBand Only] Added support for multiple VLs in SR-IOV/multi-host environments.</p> <p>Note: The number of VLs can be configured by the NVCONFIG. The default VL number is 4 VLs.</p> |
| HCA Port Flap Counter | <p>Added support for Port Flap Counter.</p> |
| Fixed Buffer Size (KSM) | <p>Limits the buffer size for all entries to improve performance. KSM is used when associating Key Length My Virtual Address (KLMs) with fixed memory size.</p> |
| NULL Mkey | <p>This entry (null_mkey) is use to indicate non-present KLM/KSM entries. When accessing is, it causes the device to generate page fault event.</p> |

Table 24 - Firmware Changes and New Feature History (Sheet 7 of 11)

| Feature/Change | Description |
|--|--|
| Out-of-Band Online Firmware Update: Firmware Update over PLDM | PLDM firmware burning is based on the DMTF spec DSP0267 (draft 9). The feature enables upgrading firmware and expansion ROM images using the PLDM protocol over MCTP (over PCIe). By doing so, a supporting BMC can query and upgrade the firmware without using OS based tools. |
| New Group in Ports Performance Counters (PPCNT) | Added a new physical layer statistics counters group. The new group includes BER counters, FEC error correction, clear time, and additional physical layer counters. For further information, please refer to the Ethernet Adapters Programming Manual (PRM) . |
| Permanent Link Up Mode | Enables the user to set a certain link up state for an unlimited period of time. This mode has 3 states: <ul style="list-style-type: none"> • Aux power (standby) • Reboot/boot/driver unloaded - the server is active and no driver is up • Driver is up - at least one driver is up (the time between init HCA and teardown or FLR) |
| No Driver NIC (NODNIC) Performance Improvement | Added support for: <ul style="list-style-type: none"> • Doorbell from User Access Region (UAR) • Clear interrupt from User Access Region (UAR) |
| Firmware Resiliency: Suppress Pauses | Allows the user to configure the adapter card to stop sending pauses after x when the receive port is unavailable (in a hang state). |
| Performance Back-pressure Counters | Added support for new performance counters. |
| Data Center Bridging Exchange (DCBX) | DCBX is used by DCB devices to exchange configuration information with directly connected peers. DCBX uses Link Layer Discovery Protocol (LLDP) to exchange parameters between two link peers. For further information, please refer to the PRM. |
| Access Register: Default Values Revert | Allows network port registers to revert to their default values when the driver is restarted or the host is rebooted. |
| Link up Modes | Added additional network link up modes. The new modes decide when to keep the network link up. The new modes are: <ul style="list-style-type: none"> • keep_eth_link_up • keep_ib_link_up • keep_link_up_on_boot • keep_link_up_on_standby |
| Explicit Congestion Notification (ECN) | Explicit Congestion Notification (ECN) is an extension to the Internet Protocol and to the Transmission Control Protocol. ECN allows end-to-end notification of network congestion without dropping packets. |

Table 24 - Firmware Changes and New Feature History (Sheet 8 of 11)

| Feature/Change | Description |
|---|--|
| RoCE Link Aggregation (RoCE LAG) | RoCE Link Aggregation provides failover and link aggregation capabilities. In this mode, only one IB port, that represents the two physical ports, is exposed to the application layer. For further information, please refer to the PRM. |
| OVS Offload | Mellanox Accelerated Switching And Packet Processing (ASAP ²) Direct technology allows to offload OVS by handling OVS data-plane in Mellanox NIC hardware (Mellanox Embedded Switch or eSwitch) while maintaining OVS control-plane unmodified. |
| FCS no scatter / FCS check | Enables the user to control whether or not to scatter Frame Check Sequence (FCS) or to check FCS functionality. |
| PRBS Patterns Generation and Tuning | A new PHY test mode in which the device can generate different PRBS patterns for SerDes tuning purpose. For further information, please refer to PRM registers: PPAOS, PPTT, PPRT. |
| Management Controller Transport Protocol (MCTP) over PCI | Added support for MCTP host management over PCI |
| OCBB / OCSD support after mlxfwreset | Added support for OCBB/OCSD memory pointers restoration after mlxfwreset |
| MCTP media migration | Added support for MCTP media migration between SMBUS and PCI |
| Cables | Removed the RX amplitude configuration on some cable types |
| IPoIB checksum and LSO offload | Added IPoIB checksum and LSO offload support |
| Scatter FCS in RQ | Enables software to scatter or strip FCS in RQ. |
| CQE Time Stamping | Keeps track of the creation of a packet. A time-stamping service supports assertions of proof that a datum existed before a particular time. |
| Priority Flow Control (PFC) | Applies pause functionality to specific classes of traffic on the Ethernet link. |
| RDMA retransmission counters | Custom port counters provide the user a clear indication about RDMA send/receive statistics and errors. |
| Link Layer Discovery Protocol (LLDP) | The Link Layer Discovery Protocol (LLDP) is a vendor-neutral Link Layer protocol in the Internet Protocol Suite used by network devices for advertising their identity, capabilities, and neighbors on a IEEE 802 LAN. The protocol is formally defined in IEEE 802.1AB. |
| Flow Steering Counters | Provides a clear indication of Flow Steering statistics and errors. |
| WQE Inline Header | The minimal amount of packet headers inlined in the WQE's Eth Segment. |

Table 24 - Firmware Changes and New Feature History (Sheet 9 of 11)

| Feature/Change | Description |
|--|---|
| table-miss Flow | A flow table may include a table-miss flow entry, which renders all Match Fields wildcards. If a packet does not match a flow entry in a flow table, this is a table miss. The behavior on a table miss depends on the table configuration. A table-miss flow entry in the flow table may specify how to process unmatched packets. |
| Strided WQE User Space | Striding RQ is a receive queue comprised by work queue elements (i.e. WQEs), where multiple packets of LRO segments (i.e. message) are written to the same WQE. |
| SR-IOV (EN eSwitch & RoCE) | Single Root IO Virtualization (SR-IOV) is a technology that allows a physical PCIe device to present itself multiple times through the PCIe bus. |
| Vector Calculation/ Erasure Coding Offload | Uses the HCA for offloading erasure coding calculations. |
| Link params modification via access registers | The change includes the following: <ol style="list-style-type: none"> 1. Changed port configuration which required link re-training (such as speed) 2. PAOS down 3. PAOS up This change, will cause the link to toggle and new configurations to take effect. |
| Checksum Calculation on Image/Device | Flint utility allows performing an MD5 checksum on the non-persistent sections of the firmware image. For further information, please refer to MFT User Manual . |
| Port Link | Reduced the port link-up time when negotiating according to Clause 73 (DME) |
| Ethernet Network | <ul style="list-style-type: none"> • Large Receive Offload (LRO) • Large Send Offload (LSO) • Receive Side Scaling (RSS) • Global Pause • RoCEv1.0/RoCEv2.0 • Flow Steering • Sniffer Ethernet • Multi packet WQE • Minimal Bandwidth Guarantee (ETS) • Explicit Congestion Notification (ECN) • Priority Flow Control (PFC) |
| PCI | <ul style="list-style-type: none"> • PCIe Function Level Reset (FLR) • Power Management L2/L3 flow support |

Table 24 - Firmware Changes and New Feature History (Sheet 10 of 11)

| Feature/Change | Description |
|-----------------------------------|--|
| PRM | <ul style="list-style-type: none"> • Strided SRQ • Self Loopback support • Transport Domain support • CQ2EQ remapping • Added support for the following commands: <ul style="list-style-type: none"> • MODIFY/QUERY_ESW_VPORT_CONTEXT • QUERY/MODIFY_CONG_STATUS • QUERY/MODIFY_CONG_PARAMS • QUERY_CONG_STATISTICS • ADD/DELETE_VXLAN_UDP_DPORT |
| Virtualization | <ul style="list-style-type: none"> • VXLAN/NVGRE Stateless offload In this release, NVGRE is supported through Windows ONLY • SR-IOV EN |
| Performance | <ul style="list-style-type: none"> • CQE zipping |
| Misc | <ul style="list-style-type: none"> • Wake-on-Lane/Standby • FlexBoot/UEFI support |
| Non-Volatile Configuration | <ul style="list-style-type: none"> • Non-Volatile Configuration (NVConfig). For the complete list, lease refer to Section 9, on page 60. |
| Port management | <ul style="list-style-type: none"> • Enabled port management. Now one port can be set as Ethernet and one as InfiniBand. |
| InfiniBand Network | <ul style="list-style-type: none"> • Dynamically Connected (DC) transport Note: There is no interoperability between ConnectX-5 and ConnectX-4 adapter cards when using DC. • Unreliable Datagram Connection transport • Atomic Operation • CORE-Direct® <ul style="list-style-type: none"> • Provides Collective Off-loading in HCA • Frees CPU to perform computation in parallel with collective operations • T10 DIF pipeline Data Integrity Signature off-loading (at beta level) • User Memory Registration (UMR) • Automatic Path Migration • On Demand Paging (ODP) - Memory can now be used without pinning memory beforehand. • Congestion Control • Shrink Address Vectors for RC and UD • Programmable Port/Node GUID |

Table 24 - Firmware Changes and New Feature History (Sheet 11 of 11)

| Feature/Change | Description |
|------------------------|--|
| General | <ul style="list-style-type: none"> • Thermal monitoring and protection • Port LEDs indications • NVConfig Tool • Suspend to RAM (S3) support • Diagnostic counters vendor-specific MAD support, as defined by VS-MAD spec version 1.2 • Physical Port Counter - Beta level • Q Counter - Beta level • Firmware burning (using mstflint) when the driver is down • CPLD field upgrade • V Port commands |
| Host management | <ul style="list-style-type: none"> • NC-SI over RMIi support |
| MAD | <ul style="list-style-type: none"> • Config space address in MAD management class 0x09 |

6 FlexBoot Changes and New Features

For further information, please refer to FlexBoot Release Notes (www.mellanox.com > Software > InfiniBand/VPI Drivers > FlexBoot).

Table 25 - FlexBoot Changes and New Features (Sheet 1 of 2)

| Version | Description |
|--|--|
| Rev. 3.5.403 | |
| Enable/Disable FlexBoot in EXPROM via mlxconfig | <p>Added PXE support to additional ConnectX- 5 adapter cards.</p> <p>Enabling/Disabling FlexBoot in ConnectX- 5 in EXPROM is done via mlxconfig. The default value is:</p> <ul style="list-style-type: none"> • FLEXBOOT enable <p>Please note, not all cards are compiled with FlexBoot. For the full list of the OPNs compiled with FlexBoot, please refer to Section 1.1, “Supported Devices”, on page 7</p> |
| VLAN Priority | Set the default VLAN priority to 0. |
| Link Aggregation Control Protocol (LACP) | LACP support is disabled by default. It can be enabled via mlxconfig. |
| Rev. 3.5.305 | |
| PXE Boot | Added ESC option as an abort key during PXE boot process. |
| FlexBoot Link Aggregation Control Protocol (LACP) | Enabled/disabled FlexBoot LACP support by editing the INI configuration. |
| Upstream sync | Synced the source with iPXE (upstream sync) |
| Rev. 3.5.210 | |
| Promiscuous VLAN mode | Added support for promiscuous VLAN mode. |
| MTU | [InfiniBand] Added support for configurable MTU. |
| Expansion ROM version | Enabled expansion ROM (<code>exp_rom</code>) version exposition according to the new specification (e.g. expose ARCH in flint tool). |
| FlexBoot UI | Added a FlexBoot menu support for <code>NV_POWER_CONF</code> . Now power consumption configuration is supported from the FlexBoot menu. |
| | Enhanced FlexBoot/firmware debug capability using Flexboot UI. Added the <code>reg_dump</code> option to the <code>panic_behavior</code> configuration in the Flex-Boot menu |
| Rev. 3.5.110 | |
| Networking | Ethernet only: The MTU value is set to 1500 upon driver’s bring up. |
| Rev. 3.5.109 | |
| Performance | Performance enhancements in Ethernet mode |
| FlexBoot UI | Added support for "Undi network wait timeout" |
| | Enhanced FlexBoot/firmware debug capability using Flexboot UI |

Table 25 - FlexBoot Changes and New Features (Sheet 2 of 2)

| Version | Description |
|---------------|---|
| Upstream sync | Synced the source with iPXE (upstream sync) |

6.1 FlexBoot Known Issues

Table 26 - FlexBoot Known Issues (Sheet 1 of 6)

| Internal Ref. | Description |
|---------------|---|
| 1295727 | Description: In Secure Host mode, the Ctrl + B option will be “read only” and changes will not be applied and may cause unknown behavior. |
| | WA: N/A |
| | Keywords: Secure Host mode, Ctrl + B option |
| | Discovered in Version: 3.5.403 |
| 1066544 | Description: Chain-loading boot-loaders that works with interrupts fails to boot on multi-host adapter cards |
| | Workaround: N/A |
| | Keywords: Chain-loading, boot |
| | Discovered in Version: 3.5.403 |
| 1157875 | Description: Pressing any of the arrow keys during boot might cause the boot process to be aborted. |
| | Workaround: N/A |
| | Keywords: Abort boot, arrows, FlexBoot |
| | Discovered in Release: 3.5.305 |
| 1149467 | Description: Chain-loading "ipxe.pxe" and "undionly.kpxe" over InfiniBand is currently not supported when using DHCP client identification based on InfiniBand 32-bit Prefix+GUID (as with FlexBoot). |
| | Workaround: N/A |
| | Keywords: FlexBoot, chainload, InfiniBand, undionly.kpxe, ipxe.pxe |
| | Discovered in Version: 3.5.305 |
| 841198 | Description: FlexBoot fails to boot when the following occurs: <ul style="list-style-type: none"> • Boot priority is set to iSCSI • The iSCSI TCP/IP parameters via DHCP is disabled • iSCSI boot fails or iSCSI boot to target configuration is set to disable |
| | Workaround: N/A |
| | Keywords: PXE boot, iSCSI |
| 843377/849223 | Description: The physical MAC assigned via the boot menu is displayed as zeroes instead of the set MAC when ConnectX-4 VPI adapter card is configured as InfiniBand. |
| | Workaround: N/A |
| | Keywords: Physical MAC, Boot menu |

Table 26 - FlexBoot Known Issues (Sheet 2 of 6)

| Internal Ref. | Description |
|---------------|---|
| 656001 | Description: Booting from WDS and Windows DHCP server when only Option 66 is enabled (without Option 67), is not supported. |
| | Workaround: N/A |
| | Keywords: DHCP |
| 776057 | Description: Citrix PVS boot is not supported. |
| | Workaround: N/A |
| | Keywords: Citrix PVS boot |
| 689460 | Description: FlexBoot uses system UUID to generate the client DUID-UUID as per RFC 6355, the data conveyed with DHCPv6 Code 1 (Option ID). |
| | Workaround: N/A |
| | Keywords: DUID-UUID |
| 928217 | Description: Installing ESXi 6.5/6.0 on iSCSI target is currently not supported. |
| | Workaround: N/A |
| | Keywords: ESXi 6.5/6.0, iSCSI target |
| 689460 | Description: To use the DHCP server to identify ipxe requests when using <code>undionly.kpxe</code> or <code>ipxe.pxe</code> when booting over IB requires special configuration. (see the Workaround below). |
| | Workaround: Add to the DHCP host declaration the MAC identification alongside the option 61 DUID. For example: <pre>host ib-client1 { option dhcp-client-identifier = ff:00:00:00:00:00:02:00:00:02:c9:00:<Port-GUID> ; hardware ethernet <Port-MAC> ; fixed-address <IPoIB Address> ; filename "ipxe.pxe" ; if exists user-class and option user-class = "iPXE" { filename "pxelinux.0" ; } }</pre> |
| | Keywords: <code>undionly.kpxe</code> or <code>ipxe.pxe</code> |

Table 26 - FlexBoot Known Issues (Sheet 3 of 6)

| Internal Ref. | Description |
|---------------|---|
| 928217 | <p>Description: Due to interoperability issue between the ESXi installer and the lpxelinux bootloader, when trying to install ESXi 6.5 on iSCSI target using lpxelinux.0 as a boot-loader, a PSOD occurs.</p> <p>Workaround: Use FlexBoot (or iPXE) to load mboot.c32 directly instead of pxelinux.0 using the script below:</p> <pre data-bbox="472 516 1289 632">#!/ipxe set base /nfs/Esxi-6.5_INBOX chain \${base}/mboot.c32 -c \${base}/boot.cfg BOOTIF=01- \${mac:hexhyp}</pre> <p>where the "set base ..." specifies a suitable absolute path.</p> <p>Note: iPXE does not need an absolute path, however, mboot.c32 requires it.</p> <p>Keywords: mboot.c32, PSOD,</p> |
| 976878 | <p>Description: When using bootloader grub2 to boot WDS, if the WDS boot fails, an RSOD might appear.</p> <p>Workaround: N/A</p> <p>Keywords: Bootloader grub2, WDS, RSOD</p> |
| 1072419 | <p>Description: The FlexBoot DHCP loops indefinitely when it continuously gets NACK on the DHCP requests. On some setups, it might also cause an RSOD after a continuous looping.</p> <p>Workaround: N/A</p> <p>Keywords: Bootloader grub2, WDS, RSOD</p> |
| - | <p>Description: Several BIOS vendors have limited boot-vector space and may not display FlexBoot in their boot menu.</p> <p>Workaround: Disable the embedded NIC boot agent in BIOS</p> <p>Keywords: BIOS</p> |
| - | <p>Description: In several BIOS, the server might hang during FlexBoot booting due to wrong configuration of the PMM.</p> <p>Workaround: N/A</p> <p>Keywords: BIOS</p> |
| - | <p>Description: Only EBX, ESI, DS, ES registers can be saved in Boot Entry.</p> <p>Workaround: N/A</p> <p>Keywords: BIOS</p> |

Table 26 - FlexBoot Known Issues (Sheet 4 of 6)

| Internal Ref. | Description |
|-------------------|---|
| - | Description: If a client returned control to the BIOS after a successful connection to an iSCSI target (but did not boot from it), then, unexpected behavior may occur. |
| | Workaround: Follow the instructions described in the FlexBoot UM for the proper iSCSI boot/install |
| | Keywords: BIOS |
| 673114/ 821899 | Description: FlexBoot banner might not be shown in some BIOSes. |
| | Workaround: N/A |
| | Keywords: BIOS |
| - | Description: In some cases, PXE boot will not work if the client was given only the filename without next-server (siaddr). |
| | Workaround: N/A |
| | Keywords: PXE Boot |
| - | Description: PXE boot after iSCSI boot with static configuration is currently not supported. |
| | Workaround: N/A |
| | Keywords: PXE Boot |
| - | Description: Boot over VLAN with IB port is currently not supported. |
| | Workaround: N/A |
| | Keywords: PXE Boot |
| - | Description: Some faulty boot loaders do not close the underlying UNDI device which may result in unexpected behavior and possible system crash after the OS starts to load. |
| | Workaround: N/A |
| | Keywords: PXE Boot |
| - | Description: Chain-loading gPXE stack is not supported. |
| | Workaround: N/A |
| | Keywords: PXE Boot |
| 647143 | Description: Executing a partial boot loop while only downloading the NBP and selecting localboot is unsupported and may cause undefined behavior. |
| | Workaround: N/A |
| | Keywords: PXE Boot |

Table 26 - FlexBoot Known Issues (Sheet 5 of 6)

| Internal Ref. | Description |
|---------------|--|
| - | Description: iSCSI over IB is not tested. |
| | Workaround: N/A |
| | Keywords: iSCSI |
| - | Description: iSCSI over DCB is not supported. |
| | Workaround: N/A |
| | Keywords: iSCSI |
| - | Description: FlexBoot supports only a single active iSCSI connection. Thus, when iSCSI-boot via Port 1 succeeds to connect but fails to boot, it will fail to connect via Port 2. |
| | Workaround: N/A |
| | Keywords: iSCSI |
| - | Description: Boot retries is currently not functional when booting from iSCSI. |
| | Workaround: N/A |
| | Keywords: iSCSI |
| 655800 | Description: iSCSI over IPv6 is not supported. |
| | Workaround: N/A |
| | Keywords: iSCSI |
| - | Description: Boot menu is displayed as READ ONLY if the HCA card does not support flash configuration. |
| | Workaround: N/A |
| | Keywords: User Interface |
| - | Description: FlexBoot Boot Menu will not be visible in serial output. |
| | Workaround: N/A |
| | Keywords: User Interface |
| - | Description: Large Receive Offload (LRO) and iSCSI may not interoperate due to a bug in current Linux kernel distributions. |
| | Workaround: Disable LRO in the IPoIB module when using iSCSI. See the Mellanox FlexBoot user's manual for details under the Diskless Machines chapter (InfiniBand Ports). |
| | Keywords: Networking |

Table 26 - FlexBoot Known Issues (Sheet 6 of 6)

| Internal Ref. | Description |
|---------------|---|
| - | Description: 56Gb/s is currently not supported. |
| | Workaround: N/A |
| | Keywords: Link Speed |
| - | Description: Setting the number of Virtual Functions higher than the machine's memory capability may cause memory issues and system instability. |
| | Workaround: N/A |
| | Keywords: Virtualization |
| - | Description: SLAM, FTP, HTTPS and SRP are currently not supported. |
| | Workaround: N/A |
| | Keywords: Protocols |
| - | Description: Occasionally, using the Spanning Tree Protocol (STP) in the switches may cause packet drops and boot failure in the system. |
| | Workaround: Enable the "edgemode" if disabled on the switch, or use either portfast or edgemode functionality on the switch ports connected to the NICs. |
| | Keywords: Protocols |
| - | Description: FCoE, BCV are not supported. |
| | Workaround: N/A |
| | Keywords: Protocols |
| 655800 | Description: IPv6 can only run if a RADVD service is running in the network. |
| | Workaround: N/A |
| | Keywords: Protocols |
| - | Description: IPv6 over IB is not supported. |
| | Workaround: N/A |
| | Keywords: Protocols |
| 655800 | Description: Enabling IPv6 first and then IPv4 is currently not supported. |
| | Workaround: N/A |
| | Keywords: Protocols |

6.2 FlexBoot Bug Fixes History

Table 27 - FlexBoot Bug Fixes History

| Version | Issue |
|---------|--|
| 1157875 | Description: Pressing any of the arrow keys during boot might cause the boot process to be aborted. |
| | Keywords: Abort boot, arrows, FlexBoot |
| | Discovered in Release: 3.5.305 |
| | Fixed in Release: 3.5.403 |
| 1113560 | Description: Fixed an issue that prevented the first iSCSI target parameters to be reset to their default values. |
| | Keywords: iSCSI target |
| | Discovered in Release: 3.5.110 |
| | Fixed in Release: 3.5.305 |

7 UEFI Changes and Major New Features

Table 28 - UEFI Changes and New Features

| Category | Description |
|--|--|
| Rev. 14.15.19 | |
| Enable/Disable UEFI in EXPROM via mlxconfig | <p>Added UEFI support to additional ConnectX- 5 adapter cards.</p> <ul style="list-style-type: none"> ConnectX-5 adapter cards are compiled with x86-UEFI <p>Enabling/Disabling UEFI in ConnectX- 5 in EXPROM is done via mlxconfig. The default values are:</p> <ul style="list-style-type: none"> UEFI_X86 disabled UEFI_AARCH64 disabled <p>For the full list of the OPNs, please refer to Section 1.1, “Supported Devices”, on page 7</p> |

7.1 UEFI Known Issues

The following is a list of general limitations and known issues of the various components of this UEFI release.

Table 29 - UEFI Known Issues

| Internal Ref. | Description |
|---------------|---|
| 1295727 | Description: In Secure Host mode, the Hii protocol will be “read only”, changes will not be applied and it may cause unknown behavior. |
| | WA: N/A |
| | Keywords: Secure Host mode |
| | Discovered in Version: 14.15.19 |
| 798073 | Description: UEFI driver is not supported on Supermicro X9DEW (BIOS version 3.0c). |
| | WA: N/A |
| | Keywords: BIOS, Supermicro X9DEW |
| - | Description: Burning the UEFI driver will remove the Flexboot driver (Legacy BIOS driver) from the firmware. |
| | WA: N/A |
| | Keywords: UEFI burning, Flexboot |

8 Unsupported Features and Commands

8.1 Unsupported Features

The following advanced feature are unsupported in the current firmware version:

- Service types not supported:
 - SyncUMR
 - Mellanox transport
 - PTP
 - RAW IPv6
 - PTP (IEEE 1588)
- INT-A not supported for EQs only MSI-X
- PCI VPD write flow (RO flow supported)
- Streaming Receive Queue (STRQ) and collapsed CQ
- Precise clock synchronization over the network (IEEE 1588)
- SM is not supported on VFs
- Socket-Direct currently does not support SR-IOV
- DC is not supported in: SR-IOV, and Ethernet (RoCE)
- Multi-Host/Socket-Direct are not supported in RoCE LAG

8.2 Unsupported Commands

- QUERY_MAD_DEMUX
- SET_MAD_DEMUX
- PAGE_FAULT_RESUME
- ACTIVATE_TRACER
- DEACTIVATE_TRACER
- ACCESS_REG_SPACE
- ACCESS_REG_SPACE_DWORD
- ACTIVATE/DEACTIVATE_TRACER
- QUERY/MODIFY_SCHED_QUEUE
- CREATE_RQ - MEMORY_RQ_RMP
- MODIFY_LAG_ASYNC_EVENT

9 Supported Non-Volatile Configurations

Table 30 - Per-physical Port Settings

| Name | Parameter Index |
|-----------------|-----------------|
| VPI settings | 0x12 |
| RoCE CC | 0x107 |
| RoCE CC ECN | 0x108 |
| LLDP_NB_DCBX | 0x18E |
| NV_QOS_CONF | 0x192 |
| NV_QOS_CAP | 0x193 |
| NV_KEEP_LINK_UP | 0x190 |

Table 31 - Global Settings

| Name | Parameter Index |
|--------------------------|-----------------|
| PCI settings | 0x80 |
| PCI setting capabilities | 0x81 |
| TPT settings | 0x82 |
| TPT capabilities | 0x83 |
| Option ROM ini | 0x100 |
| Option ROM capabilities | 0x101 |
| NV_SW_OFFLOAD_CONF | 0x10A |
| NV_PACKET_PACING | 0x10C |

Table 32 - Per host/function Settings

| Name | Parameter Index |
|---------------|-----------------|
| Wake-on-LAN | 0x10 |
| External Port | 0x192 |

Table 33 - Per host Settings

| Name | Parameter Index |
|-------------|-----------------|
| NV_PCI_CONF | 0x80 |
| NV_PCI_CAP | 0x81 |