



Mellanox ConnectX[®]-5 Firmware Release Notes

Rev 16.26.1040

NOTE:

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



Mellanox Technologies
350 Oakmead Parkway Suite 100
Sunnyvale, CA 94085
U.S.A.
www.mellanox.com
Tel: (408) 970-3400
Fax: (408) 970-3403

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

Mellanox®, Mellanox logo, Connect-IB®, ConnectX®, CORE-Direct®, GPUDirect®, LinkX®, Mellanox Multi-Host®, Mellanox Socket Direct®, UFM®, and Virtual Protocol Interconnect® are registered trademarks of Mellanox Technologies, Ltd.

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

All other trademarks are property of their respective owners.

Table of Contents

Release Update History	6
Chapter 1 Overview	7
1.1 Supported Devices	7
1.2 Supported Cables and Modules	10
1.2.1 Validated and Supported 1GbE Cables	10
1.2.2 Validated and Supported 10GbE Cables	10
1.2.3 Validated and Supported 25GbE Cables	12
1.2.4 Validated and Supported 40GbE Cables	13
1.2.5 Validated and Supported 50GbE Cables	13
1.2.6 Validated and Supported 100GbE Cables	14
1.2.7 Validated and Supported 200GbE Cables	15
1.2.8 Validated and Supported QDR Cables	15
1.2.9 Validated and Supported FDR10 Cables	15
1.2.10 Validated and Supported FDR Cables	16
1.2.11 Validated and Supported EDR / 100Gb/s Cables	16
1.2.12 Validated and Supported HDR / 200Gb/s Cables	17
1.3 Tested Switches	18
1.3.1 Tested 10GbE Switches	18
1.3.2 Tested 40GbE Switches	18
1.3.3 Tested 100GbE Switches	18
1.3.4 Tested QDR Switches	19
1.3.5 Tested FDR Switches	19
1.3.6 Tested EDR / 100Gb/s Switches	19
1.4 Tools, Switch Firmware and Driver Software	20
1.5 Supported FlexBoot, UEFI	21
1.6 Revision Compatibility	21
Chapter 2 Changes and New Features in Rev 16.26.1040	22
Chapter 3 Known Issues	23
Chapter 4 Bug Fixes History	30
Chapter 5 Firmware Changes and New Feature History	35
Chapter 6 PreBoot Drivers (FlexBoot/UEFI)	49
6.1 FlexBoot Changes and New Features	49
6.2 UEFI Changes and Major New Features	49
Chapter 7 Unsupported Features and Commands	50
7.1 Unsupported Features	50
7.2 Unsupported Commands	50

Chapter 8 Supported Non-Volatile Configurations 51

List of Tables

Table 1:	Release Update History	6
Table 2:	Supported Devices	7
Table 3:	Validated and Supported 1GbE Cables	10
Table 4:	Validated and Supported 10GbE Cables	10
Table 5:	Validated and Supported 25GbE Cables	12
Table 6:	Validated and Supported 40GbE Cables	13
Table 7:	Validated and Supported 50GbE Cables	13
Table 8:	Validated and Supported 100GbE Cables	14
Table 9:	Validated and Supported 200GbE Cables	15
Table 10:	Validated and Supported QDR Cables	15
Table 11:	Validated and Supported FDR10 Cables	15
Table 12:	Validated and Supported FDR Cables	16
Table 13:	Validated and Supported EDR / 100Gb/s Cables	16
Table 14:	Validated and Supported HDR Cables	17
Table 15:	Tested 10GbE Switches	18
Table 16:	Tested 40GbE Switches	18
Table 17:	Tested 100GbE Switches	18
Table 18:	Tested QDR Switches	19
Table 19:	Tested FDR Switches	19
Table 20:	Tested EDR Switches	19
Table 21:	Tools, Switch Firmware and Driver Software	20
Table 22:	Supported FlexBoot, UEFI	21
Table 23:	Changes and New Features in Rev 16.26.1040	22
Table 24:	Ethernet Rate Limit per VF in RoCE Mode Limitations	23
Table 25:	Ethernet Rate Limit per VF in InfiniBand Mode Limitations	23
Table 26:	Known Issues	23
Table 27:	Bug Fixes History	30
Table 28:	Firmware Changes and New Feature History	35
Table 29:	Unsupported Features	50
Table 30:	Supported Non-Volatile Configurations	51

Release Update History

Table 1 - Release Update History

Release	Date	Description
Rev 16.26.1040	September 29, 2019	Initial version of this firmware release. This version introduces New Features Section 2, “Changes and New Features in Rev 16.26.1040” , on page 22 and Bug Fixes (see Section 4, “Bug Fixes History” , on page 30).

1 Overview

These are the release notes for the ConnectX®-5 adapters firmware Rev 16.26.1040.

This firmware supports the following protocols:

- InfiniBand - SDR, QDR, FDR10, FDR, EDR
- Ethernet - 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 56GbE¹, 100GbE
- PCI Express 4.0, supporting backwards compatibility for v3.0, 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 4)

Device Part Number	PSID	Device Name	FlexBoot	UEFI x86	UEFI ARM	Enable/disable exprom Feature
MCX512A-ACUT	MT_0000000425	ConnectX®-5 EN network interface card, 10/25GbE dual-port SFP28, PCIe3.0 x8, UEFI Enabled (x86/ARM), tall bracket	Present (Enabled)	Present (Enabled)	Present (Enabled)	Exists
MCX512A-ADAT	MT_0000000361	ConnectX®-5 Ex EN network interface card, 25GbE dual-port SFP28, PCIe3.0/4.0 x8, tall bracket	Present (Enabled)	Present (Disabled)	Not Present	Exists
MCX562A-ACAB	MT_0000000241	ConnectX®-5 EN network interface card for OCP 3.0, with host management, 25GbE Dual-port SFP28, PCIe3.0 x16, Thumbscrew (Pull Tab) bracket	Present (Enabled)	Present (Enabled)	Not Present	Exists
MCX4121A-ACUT	MT_0000000266	ConnectX®-4 Lx EN network interface card, 25GbE dual-port SFP28, PCIe3.0 x8, UEFI Enabled, tall bracket	Present (Enabled)	Present (Enabled)	Present (Disabled)	Exists
MCX566A-CDA	MT_0000000242	ConnectX®-5 Ex EN network interface card for OCP 3.0, with host management, 100GbE Dual-port QSFP28, PCIe4.0 x16, Internal Lock bracket	Present (Enabled)	Present (Disabled)	Not Present	Exists
MCX545B-CCUN	MT_0000000419	ConnectX®-5 EN network interface card for OCP2.0, Type 1, with host management, 100GbE, single-port QSFP28, PCIe3.0 x16, UEFI Enabled, no bracket	Present (Enabled)	Present (Enabled)	Not Present	Exists
MCX542B-ACUN	MT_0000000427	ConnectX®-5 EN network interface card for OCP2.0, Type 1, with host management, 25GbE dual-port SFP28, PCIe3.0 x8, UEFI Enabled (x86/ARM), no bracket Halogen free	Present (Enabled)	Present (Enabled)	Present (Enabled)	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 4)

Device Part Number	PSID	Device Name	FlexBoot	UEFI x86	UEFI ARM	Enable/disable exprom Feature
MCX512F-ACHT	MT_0000000416	ConnectX®-5 EN network interface card, with host management, 25GbE Dual-port SFP28, PCIe3.0 x16, UEFI Enabled, tall bracket	Present (Enabled)	Present (Enabled)	Not Present	Exists
MCX545A-CCUN	MT_0000000418	ConnectX®-5 EN network interface card for OCP2.0, Type 2, with host management, 100GbE, single-port QSFP28, PCIe3.0 x16, UEFI Enabled, no bracket	Present (Enabled)	Present (Enabled)	Not Present	Exists
MCX516A-CCHT	MT_0000000417	ConnectX®-5 EN network interface card, with host management 100GbE dual-port QSFP28, PCIe3.0 x16, UEFI Enabled, tall bracket	Present (Enabled)	Present (Enabled)	Not Present	Exists
MCX566A-CCAI	MT_0000000348	ConnectX®-5 EN network interface card for OCP 3.0, with host management, 100GbE Dual-port QSFP28, PCIe3.0 x16, Internal Lock bracket	Present (Enabled)	Present (Disabled)	Not Present	Exists
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
MCX546A-BCAN	MT_0000000069	ConnectX®-5 EN network interface card for OCP, 40GbE dual-port QSFP28, PCIe3.0 x16, no bracket, ROHS R6	Present (Enabled)	Present (Disabled)	Not Present	Exists
MCX546A-CDAN	MT_0000000058	ConnectX-5 Ex network interface card for OCP; 100GbE dual-port QSFP28; PCIe4.0 x16; no bracket; ROHS R6	Present (Enabled)	Present (Disabled)	Not Present	Exists
MCX545A-CCAN	MT_0000000157	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

Table 2 - Supported Devices (Sheet 3 of 4)

Device Part Number	PSID	Device Name	FlexBoot	UEFI x86	UEFI ARM	Enable/disable exprom Feature
MCX512F-ACAT	MT_0000000183	ConnectX®-5 EN network interface card, 25GbE Dual-port SFP28, PCIe3.0 x16, tall bracket, ROHS R6	Present (Enabled)	Present (Disabled)	Not Present	Exists
MCX511F-ACA	MT_0000000182	ConnectX-5 EN network interface card; 25GbE single-port SFP28; PCIe4.0 x16; ROHS R6	Present (Enabled)	Present (Disabled)	Not Present	Exists
MCX542B-ACA	MT_0000000248	ConnectX-5 EN network interface card for OCP; with host management; 25GbE dual-port SFP28; PCIe3.0 x8; no bracket; ROHS R6 Halogen free	Present (Enabled)	Present (Enabled)	Not Present	Exists
MCX542A-ACAN	MT_0000000167	ConnectX®-5 EN network interface card for OCP, with host management, 25GbE dual-port SFP28, PCIe3.0 x16, no bracket, ROHS R6 Halogen free	Present (Enabled)	Present (Disabled)	Not Present	Not Present
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
MCX553Q-ECAS	MT_0000000309	ConnectX®-5 VPI adapter card with Multi-Host, EDR IB (100Gb/s) and 100GbE, Single-port QSFP28, PCIe3.0 x4 on board, external connectors to 3x auxiliary cards?, Short bracket	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
MCX545A-ECA_BX	MT_0000000092	ConnectX-5 VPI network interface card for OCP EDR IB (100Gbs) and 100GbE; single-port QSFP28; PCIe3.0 x16; no bracket; ROHS R6; Halogen free	Present (Enabled)	Present (Disabled)	Not Present	Exists
MCX545B-ECAN	MT_0000000207	ConnectX-5 VPI network interface card for OCP; with host management; EDR IB (100Gb/s) and 100GbE; single-port QSFP28; PCIe3.0 x16; no bracket; 8mm Heat Sink; 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

Table 2 - Supported Devices (Sheet 4 of 4)

Device Part Number	PSID	Device Name	FlexBoot	UEFI x86	UEFI ARM	Enable/disable exprom Feature
MCX556A-EDAT	MT_0000000009	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_0000000023	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
MCX546A-EDAN	MT_0000000135	ConnectX-5 VPI network interface card for OCP; EDR IB (100Gb/s) and 100GbE dual-port QSFP28; PCIe4.0 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
10GbE	BN-QS-SP-CBL-5M	40G QSFP+ to 4xSFP+ DAC Breakout Direct Attach Cable 5m
10GbE	BN-QS-SP-CBL-5M	40G QSFP+ to 4xSFP+ DAC Breakout Direct Attach Cable 5m
10GbE	CAB-SFP-SFP-1M	Arista 10GBASE-CR SFP+ Cable 1 Meter
10GbE	CAB-SFP-SFP-3M	Arista 10GBASE-CR SFP+ Cable 3 Meter
10GbE	CAB-SFP-SFP-5M	Arista 10GBASE-CR SFP+ Cable 5 Meter
10GbE	FTLX1471D3BCL-ME	10GBASE-LR SFP+ 1310nm 10km DOM Transceiver Module
10GbE	L45593-D178-B50	QSFP-4SFP10G-CU5M
10GbE	MC2309124-004	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 4M
10GbE	MC2309124-005	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 5M

Table 4 - Validated and Supported 10GbE Cables

Speed	Cable OPN #	Description
10GbE	MC2309130-001	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 1M
10GbE	MC2309130-002	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 2M
10GbE	MC2309130-003	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 3M
10GbE	MC2309130-00A	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 0.5M
10GbE	MC2609125-004	Mellanox Passive Copper Hybrid Cable ETH 40GbE TO 4X10GBE QSFP TO 4X SFP+ 4M
10GbE	MC2609130-001	Mellanox Passive Copper Hybrid Cable ETH 40GbE TO 4X10GBE QSFP TO 4X SFP+ 1M
10GbE	MC2609130-002	Mellanox Passive Copper Hybrid Cable ETH 40GbE TO 4X10GBE QSFP TO 4X SFP+ 2M
10GbE	MC2609130-003	Mellanox Passive Copper Hybrid Cable ETH 40GbE TO 4X10GBE QSFP TO 4X SFP+ 3M
10GbE	MC2609130-0A1	Mellanox Passive Copper Hybrid Cable ETH 40GbE TO 4X10GBE QSFP TO 4X SFP+ 1.5M
10GbE	MC3309124-004	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 4M
10GbE	MC3309124-005	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 5M
10GbE	MC3309124-006	Mellanox® Passive Copper Cable, ETH 10GbE, 10Gb/s, SFP+, 6m
10GbE	MC3309124-007	Mellanox® Passive Copper Cable, ETH 10GbE, 10Gb/s, SFP+, 7m
10GbE	MC3309130-001	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 1M
10GbE	MC3309130-002	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 2M
10GbE	MC3309130-003	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 3M
10GbE	MC3309130-004	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 4M
10GbE	MC3309130-005	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 5M
10GbE	MC3309130-006	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 6M
10GbE	MC3309130-007	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 7M
10GbE	MC3309130-00A	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 0.5M
10GbE	MC3309130-0A1	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 1.5M
10GbE	MC3309130-0A2	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 2.5M
10GbE	MFMT102A-LR-F	Mellanox Optical Module ETH 10GBE 10GB/S SFP+ LC-LC 1310NM LR up to 10KM
10GbE	SFP-10G-SR	Cisco 10GBASE-SR SFP+ transceiver module for MMF, 850-nm wavelength, LC duplex connector
10GbE	SFP-H10GB-CU1M	Cisco 1-m 10G SFP+ Twinax cable assembly, passive
10GbE	SFP-H10GB-CU3M	Cisco 3-m 10G SFP+ Twinax cable assembly, passive
10GbE	SFP-H10GB-CU5M	Cisco 5-m 10G SFP+ Twinax cable assembly, passive
10GbE	FTLX8571D3BCL-ME	10gb SFP 850nm Optic Transceiver

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	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-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	MFA7A50-CXXX	Mellanox Active Fiber Hybrid Solution ETH 100GbE to 4X25GBE QSFP28 to 4XSFP28 up to 30M
25GbE	MMA2L20-AR	Mellanox® optical transceiver, 25GbE, 25Gb/s, SFP28, LC-LC, 1310nm, LR up to 10km
25GbE	MMA2P00-AS	Mellanox® transceiver, 25GbE, SFP28, LC-LC, 850nm, SR, up to 100m
25GbE	MFM1T02A-SR-P	Mellanox® Optical Module ETH 10GbE 10GB/S SFP+ LC-LC 850NM SR up to 300M
25GbE	SFP-H25G-CU1M	25GBASE-CR1 Copper Cable 1-meter
25GbE	SFP-H25G-CU2M	25GBASE-CR1 Copper Cable 2-meter

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-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-AOC10M	Cisco 40GBase-AOC QSFP direct-attach Active Optical Cable, 10-meter

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	CAB-Q-Q-100GbE-3M	Passive 3 meter , QSFP+ to QSFP+ QSFP100 TWINAX 103.125Gbps-CR4
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
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
100GbE	MMS1C10-CM	Mellanox® active optical module, 100Gb/s, QSFP, MPO, 1310nm, PSM4
100GbE	MMA1L30-CM	Mellanox® optical module, 100GbE, 100Gb/s, QSFP28, LC-LC, 1310nm, CWDM4, up to 2km
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	QSFP-40/100-SRBD	100GbE and 40GBASE SR-BiDi QSFP Transceiver, LC, 100m OM4 MMF

Table 8 - Validated and Supported 100GbE Cables

Speed	Cable OPN #	Description
100GbE	TR-FC13L-N00	100G QSFP28 Optical Transceivers, QSFP28 LR4 (10km)
100GbE	SO-QSFP28-LR4	QSFP28, 100GBase, 1310nm, SM, DDM, 10km, LC
100GbE	10137499-4050LF	HP 5m splitter 100G to 4X25G
100GbE	FTLF8519P3BTL-N1	1000BASE-SX and 2G Fibre Channel (2GFC) 500m Industrial Temperature SFP Optical Transceiver

1.2.7 Validated and Supported 200GbE Cables

Table 9 - Validated and Supported 200GbE Cables

Speed	Cable OPN #	Description
200GbE	MCP1650-V001E30	Mellanox® Passive Copper cable, 200GbE, 200Gb/s, QSFP56, LSZH, 1m, black pultab, 30AWG
200GbE	MCP1650-V002E26	Mellanox® Passive Copper cable, 200GbE, 200Gb/s, QSFP56, LSZH, 2m, black pultab, 26AWG
200GbE	MCP1650-V003E26	Mellanox® Passive Copper cable, 200GbE, 200Gb/s, QSFP56, LSZH, 3m, black pultab, 26AWG
200GbE	MCP1650-V00AE30	Mellanox® Passive Copper cable, 200GbE, 200Gb/s, QSFP56, LSZH, 0.5m, black pultab, 30AWG
200GbE	MCP1650-V02AE26	Mellanox® Passive Copper cable, 200GbE, 200Gb/s, QSFP56, LSZH, 2.5m, black pultab, 26AWG
200GbE	MCP7H50-V003R26	Mellanox® Passive Copper hybrid cable, 200GbE 200Gb/s to 2x100Gb/s, QSFP56 to 2xQSFP56, colored, 3m, 26AWG

1.2.8 Validated and Supported QDR Cables

Table 10 - 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

1.2.9 Validated and Supported FDR10 Cables

Table 11 - 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

Table 11 - Validated and Supported FDR10 Cables

Speed	Cable OPN #	Description
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

1.2.10 Validated and Supported FDR Cables

Table 12 - 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	MC2207130-001	Mellanox® Passive Copper Cable, VPI, up to 56Gb/s, QSFP, 1m
FDR	MC2207130-00A	Mellanox® Passive Copper Cable, VPI, up to 56Gb/s, QSFP, 0.5m
FDR	MC2207310-003	Mellanox® Active Fiber Cable, VPI, up to 56Gb/s, QSFP, 3m
FDR	MC2207310-010	Mellanox® Active Fiber Cable, VPI, up to 56Gb/s, QSFP, 10m
FDR	MC2207310-015	Mellanox® Active Fiber Cable, VPI, up to 56Gb/s, QSFP, 15m
FDR	MC2207310-100	Mellanox® Active Fiber Cable, VPI, up to 56Gb/s, QSFP, 100m
FDR	MC2207312-XXX	Mellanox® Active Fiber Cable, VPI, up to 56Gb/s, QSFP, up to 100m
FDR	MC220731V-XXX	Mellanox® Active Fiber cable, VPI, up to 56Gb/s, QSFP, up to 100m
FDR	MC2207411-SR4L	Mellanox® Optical Module, VPI, up to 56Gb/s, QSFP, MPO, 850nm, up to 30m

1.2.11 Validated and Supported EDR / 100Gb/s Cables

Table 13 - Validated and Supported EDR / 100Gb/s Cables

Speed	Cable OPN #	Description
EDR	MCP1600-E001	Mellanox Passive Copper Cable VPI 100Gb/s QSFP LSZH 1M
EDR	MCP1600-E002	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-E005	Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 5M
EDR	MCP1600-E00A	Mellanox Passive Copper Cable VPI 100Gb/s QSFP LSZH 0.5M
EDR	MCP1600-E01A	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	MCPIOPT-E002	Mellanox® Passive Copper cable, VPI, up to 100Gb/s, QSFP, LSZH, 2m
EDR	MFA1A00-E005	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 5m
EDR	MFA1A00-E006	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 6m
EDR	MFA1A00-E010	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 10m
EDR	MFA1A00-E015	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

Table 13 - Validated and Supported EDR / 100Gb/s Cables

Speed	Cable OPN #	Description
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	MMA1B00-E100	Mellanox® Transceiver, IB EDR, up to 100Gb/s, QSFP28, MPO, 850nm, up to 100m
EDR	MMA1L30-CM	Mellanox® optical module, 100Gb/s, QSFP28, LC-LC, 1310nm, CWDM4, up to 2km

1.2.12 Validated and Supported HDR / 200Gb/s Cables

Table 14 - Validated and Supported HDR Cables

Speed	OPN # / Name	Description
HDR	MCP1650-H001E30	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP28, PVC, 1m, white pultab, 30AWG
HDR	MCP1650-H002E26	Mellanox Passive Copper Cable, IB HDR, up to 200GB/S, QSFP56, LSZH, 2M, black pultab, 26AWG
HDR	MCP1650-H003E26 / P06149-B26	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP56, LSZH, 3m, black pulltab, 26AWG
HDR	MCP1650-H00AE30	Mellanox Passive Copper Cable, IB HDR, up to 200Gb/s, QSFP56, LSZH, 0.5M, black pultab, 30AWG
HDR	MCP1650-H01AE30	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP28, PVC, 1.5m, white pultab, 30AWG
HDR	MCP1650-H01AE30	Mellanox Passive Copper Cable, IB HDR, up to 200Gb/s, QSFP56, LSZH, 1.5M, black pultab, 30AWG
HDR	MCP1650-H02AE26	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP28, PVC, 2.5m, white pultab, 26AWG
HDR	MCP7H50-H003R26	Mellanox® passive copper hybrid cable, IB HDR 200Gb/s to 2x100Gb/s, QSFP56 to 2xQSFP56, LSZH, colored, 3m, 26AWG
HDR	MCP7H50-H01AR30	Mellanox® passive copper hybrid cable, IB HDR 200Gb/s to 2x100Gb/s, QSFP56 to 2xQSFP56, LSZH, colored, 1.5m, 30AWG



HDR links raise with RS-FEC.

1.3 Tested Switches

1.3.1 Tested 10GbE Switches

Table 15 - 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
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 16 - 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 17 - 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

Table 17 - Tested 100GbE Switches

Speed	Switch Silicon	OPN # / Name	Description	Vendor
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 100GbE QSFP28 - Leaf/Spine Switch, power-to-port airflow, DC Power	Edgecore
100GbE	N/A	QFX5200-32C-32	32-port 100GbE Ethernet Switch System	Juniper
100GbE	N/A	S6820-56HF	48 SFP+ + 8 QSFP Ports 100GbE Switch Ethernet	H3C

1.3.4 Tested QDR Switches

Table 18 - 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 19 - 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 / 100Gb/s Switches

Table 20 - Tested EDR Switches

Speed	Switch Silicon	OPN # / Name	Description	Vendor
EDR	Switch-IB	MSB7790-EB2F	36-port EDR 100Gb/s InfiniBand Switch Systems	Mellanox
EDR	Switch-IB 2	MSB7800-ES2R	36-port Non-blocking Managed EDR 100Gb/s InfiniBand Smart Switch	Mellanox

1.4 Tools, Switch Firmware and Driver Software

The following are the drivers' software, tools, switch/HCA firmware versions tested that you can upgrade from or downgrade to when using firmware Rev 16.26.1040:

Table 21 - Tools, Switch Firmware and Driver Software

	Supported Version
MLNX_OFED	4.7-x.0.0.0 / 4.6-1.0.1.1
MLNX_EN (MLNX_OFED based code)	4.7-x.0.0.0 / 4.6-1.0.1.1
WinOF-2	2.30 / 2.20
MFT	4.13.0 / 4.12.0
MLNX-OS	3.8.2004
Onyx	3.8.2004
ConnectX-5 Firmware	16.25.1020 / 16.24.1000
SwitchX-IB™ Firmware	11.2000.2046
SwitchX-IB 2 Firmware	15.2000.2046
Linux Inbox Drivers	<ul style="list-style-type: none"> • RH7.6 • Ubuntu 16.04.05
Windows Inbox Drivers	<ul style="list-style-type: none"> • Windows 2012 • Windows 2012 R2 • Windows 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 16.26.1040 supports the following FlexBoot:

Table 22 - Supported FlexBoot, UEFI

Expansion ROM	Supported Version
FlexBoot	3.5.803
UEFI	14.19.14

1.6 Revision Compatibility

Firmware Rev 16.26.1040 complies with the following programmer’s reference manual:

- *Mellanox Adapters Programmer’s Reference Manual (PRM), Rev 0.47 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 16.26.1040

Table 23 - Changes and New Features in Rev 16.26.1040

Feature/Change	Description
Rev. 16.26.1040	
Address Translation Service (ATS)	Added Address Translation Service (ATS) support for MKEY and UMEM.
VPD	Added support for exposing the VPD on the VF.
ICMD and Diagnostic Counters	Enabled the firmware by using the ICMD commands to deal with diagnostic counters similar to cmdif. They can be called via the vsec space. The counters' values are returned only via the tracer. The ICMD Query Caps indicate support and expose the list of the supported counters.
Hairpin Drop Counter	Added support for Hairpin Drop Counter.
User Context Object (DEVX)	This is a containerized sandbox per user, to access PRM command securely by using General Object commands, UMEM and UCTX contexts. The allowed functionalities of this capability depend on the user permissions. The following functionalities are still managed by the Kernel: <ul style="list-style-type: none"> • Resource cleaning • UCTX stamping • Blocking the physical address and IRQ from these UCTX
DEVX Support for Asynchronous Events	Added support for reporting the supported affiliated and unaffiliated asynchronous events to DEVX users through the command interface.
Hairpin and TM RNDV QPs in DEVX	Added support for Hairpin and TM RNDV QPs to work with DevX.
Software Managed Steering Tables	Added support for creating software managed steering tables in eSwitch/FDB.
Zero-Touch-RoCE Counters	Zero-Touch-RoCE counters are now available to the user for debuggability purposes when using the Zero-Touch-RoCE feature.
Security Hardening Enhancements	This release contains important reliability improvements and security hardening enhancements. Mellanox recommends upgrading your device firmware to this release to improve the device firmware security and reliability.
Bug Fixes	See Section 4, “Bug Fixes History”, on page 30

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 24 - Ethernet Rate Limit per VF in RoCE Mode Limitations

Adapter Card	Dual Port Device				Single Port Device	
	w/o LAG (TOTAL_VFS>32)		With LAG (TOTAL_VFS<32)		w/o LAG	
	w/o QoS	Full QoS	w/o QoS	Full QoS	w/o QoS	Full QoS
ConnectX-5	127	127	64	64	127	127

Table 25 - Ethernet Rate Limit per VF in InfiniBand Mode Limitations

Adapter Card	Dual Port Device		Single Port Device	
	w/o LAG		w/o LAG	
	w/o QoS	Full QoS	w/o QoS	Full QoS
ConnectX-5	127	127	127	127

Table 26 - Known Issues (Sheet 1 of 7)

Internal Ref.	Issue
1888306	Description: Occasionally Bluescreen might occur when using mlxfwreset for Socket Direct devices on Windows.
	Workaround: N/A
	Keywords: mlxfwreset, Socket Direct,
	Discovered in Version: 16.26.1040
1895191	Description: SX_RDMA is not supported when Dual Port RoCE feature is enabled. Due to this behavior, packets sent on other port are be processed by the SX_RDMA table.
	Workaround: Use SX_RDMA with Dual Port GVMI instead.
	Keywords: SX_RDMA, Dual Port RoCE, Dual Port GVMI
	Discovered in Version: 16.26.1040
1919403	Description: Hardware arbitration is currently disabled in OCP3.0 cards. It will be supported on future releases for the same hardware.
	Workaround: N/A
	Keywords: Hardware arbitration, OCP3.0
	Discovered in Version: 16.26.1040
1912117	Description: The sw_reset option is not supported when ATS is enabled.
	Workaround: N/A
	Keywords: ATS, sw_reset
	Discovered in Version: 16.26.1040

Table 26 - Known Issues (Sheet 2 of 7)

Internal Ref.	Issue
1836465	<p>Description: When using the hairpin feature, and using VLAN strip or using the “modify esw vport context” command, the packets can have an incorrect VLAN header. Meaning, using VLAN push/pop may not work properly when using vport context VLAN.</p> <p>The features that may be affected by this and not work properly are:</p> <ul style="list-style-type: none"> • Host chaining • Mirroring in FDB • TTL modify in FDB • VGT+
	Workaround: N/A
	Keywords: E-switch vport context, VLAN
	Discovered in Version: 16.26.1040
1842278	<p>Description: DC LAG can function only in case there is a single PF per port without any active VFs.</p>
	Workaround: N/A
	Keywords: DC LAG
1796628	<p>Description: Due to performance considerations, unicast loopback traffic will go through the NIC SX tables, and multicast loopback traffic will skip the NIC SX tables.</p>
	Workaround: N/A
	Keywords: Performance, unicast loopback traffic, multicast loopback traffic
	Discovered in Version: 16.26.1040
1797493	<p>Description: Firmware asserts may occur when setting the PF_BAR2_SIZE value higher than the maximum supported size (maximum PF_BAR2_SIZE is 4 for .</p>
	Workaround: Configure within limits (NIC PF_BAR_SIZE <= 4).
	Keywords: Multi-GVMI, Sub-Function, SFs, BAR2
	Discovered in Version: 16.26.1040
1761271	<p>Description: CWD4 AOM cable is currently not supported.</p>
	Workaround: N/A
	Keywords: Modules/Cables
	Discovered in Version: 16.26.1040
1752009	<p>Description: When working with Multi-GVMI and SR-IOV, and with a high number of Virtual Functions and sub-functions, the driver start may fail for the VFs/sub-functions.</p>
	Workaround: Decrease the number of configured VFs or sub-functions using mlxconfig PF_BAR2_SIZE or NUM_OF_VFS.
	Keywords: Multi-GVMI, SR-IOV
	Discovered in Version: 16.25.1020
1762142	<p>Description: PF / ECPF FLR does not clear all its dependent sub-functions. QUERY_ESW_FUNCTIONS and ALLOC/DEALLOC_SF commands might fail / show allocated SFs after PF FLR.</p>
	Workaround: Perform a graceful shutdown, and not an FLR.
	Keywords: Multi-GVMI, SF, Sub-Functions, FLR
	Discovered in Version: 16.25.1020

Table 26 - Known Issues (Sheet 3 of 7)

Internal Ref.	Issue
1768814/1772474	Description: Due to hardware limitation, REG_C cannot be passed over loopback when the FDB action is forwarded to multiple destinations.
	Workaround: N/A
	Keywords: Connection-Tracking
	Discovered in Version: 16.25.1020
1770736	Description: When a PF or ECPF with many VFs (SR-IOV), and/or SFs (Multi-GVMI) triggers an FLR, PCIe completion timeout might occur.
	Workaround: Increase the PCIe completion timeout.
	Keywords: Multi-GVMI, SR-IOV, Sub-Function, Virtual Function, PF FLR
	Discovered in Version: 16.25.1020
1716334	Description: When mlxconfig.PF_BAR2_EN is enabled, configuring more than 255 PCI functions will raise an assert.
	Workaround: When working with BAR2, configure SR-IOV to align to the 255 PCI functions limitation. mlxconfig.NUM_OF_VFS controls the number of configured SR-IOV VFs. e.g.: <ul style="list-style-type: none"> Smart NICs: 2 External Host PFs, 2 ARM ECPFs, 125 VFs per PF. Non-smart NICs: 2 External Host PFs, 126 VFs per PF
	Keywords: Multi-GVMI, PF_BAR2_EN, Sub-Functions, SR-IOV, VFs
	Discovered in Version: 16.25.1020
1699214	Description: NODNIC VF is partially tested. It is fully tested only in ConnectX-5 adapter cards.
	Workaround: N/A
	Keywords: NODNIC VF
	Discovered in Version: 16.25.1020
1699214	Description: NODNIC VF is partially tested. It is fully tested only in ConnectX-5 adapter cards.
	Workaround: N/A
	Keywords: NODNIC VF
	Discovered in Version: 16.25.1020
1749691	Description: On rare occasions, when using Socket-Direct devices, inband burning through the external port might fail.
	Workaround: N/A
	Keywords: Socket-Direct, inband burning
	Discovered in Version: 16.25.1020
1689186	Description: Changing priority to TC map during traffic might cause packet drops.
	Workaround: N/A
	Keywords: QoS
	Discovered in Version: 16.25.1020
1604699	Description: Ethernet RFC 2819 counter ether_stats_oversize_pkts and Ethernet IEEE 802.3 counter a_frame_too_long_errors share the same resource. Clearing each of them will affect the other.
	Workaround: N/A
	Keywords: Counters
	Discovered in Version: 16.25.1020

Table 26 - Known Issues (Sheet 4 of 7)

Internal Ref.	Issue
1558250	Description: eSwitch owner may receive NIC_VPORT_CONTEXT events from vPorts that are not necessarily armed using the nic_vport_context_arm_change_even_tbit.
	Workaround: N/A
	Keywords: Port event, NODNIC
	Discovered in Version: 16.25.1020
-	Description: In Ethernet mode, at 10/40GbE speeds, only NO-FEC in Force mode is supported. Other user configurations are overridden.
	Workaround: N/A
	Keywords: Ethernet, 10GbE, 40GbE, RS-FEC
	Discovered in Version: 16.25.1020
1574876	Description: DC RoCE LAG is functional only if the router posts VRRP address as the source MAC.
	Workaround: N/A
	Keywords: DC RoCE LAG
	Discovered in Version: 16.25.1020
1498399	Description: If the XRC switches between SRQ/RMPs while there is an outstanding ODP on the responder XRC QP, a CQE with an error might be generated (that is not a PFAULT abort).
	Workaround: N/A
	Keywords: XRC SRQ/RMP ODP
	Discovered in Version: 16.25.1020
1546401	Description: vport_tc and para_vport_tc are not supported in this version.
	Workaround: N/A
	Keywords: SR-IOV vport_tc and para_vport_tc
	Discovered in Version: 16.24.1000
1546492	Description: Executing the update_lid command while the IB port sniffer utility is active can stop the utility.
	Workaround: N/A
	Keywords: IB Sniffer
	Discovered in Version: 16.24.1000
1537898	Description: Initializing a function while the IB port sniffer utility is active can stop the utility.
	Workaround: N/A
	Keywords: IB Sniffer
	Discovered in Version: 16.24.1000
1523577	Description: When modifying the TTL in the NIC RX, the CQE checksum is not recalculated automatically. The limitation is indicated by the ttl_checksum_correction bit. If the ttl_checksum_correction=0, the capability is not functioning properly.
	Workaround: N/A
	Keywords: multi_prio_sq, VF
	Discovered in Version: 16.24.1000

Table 26 - Known Issues (Sheet 5 of 7)

Internal Ref.	Issue																																
1414290	Description: When getting an inline scatter CQE on IB striding RQ, the stride index in the CQE will be zero.																																
	Workaround: N/A																																
	Keywords: Scatter CQE																																
	Discovered in Version: 16.24.1000																																
1475490	Description: Reboot is not supported on any host during the PLDM firmware burning process.																																
	Workaround: N/A																																
	Keywords: PLDM																																
	Discovered in Version: 16.23.1020																																
1332714	Description: The maximum “read” size of MTRC_STDB is limited to 272 Bytes.																																
	Workaround: Set the MTRC_STDB.read_size to the maximum value of 0x110=272 Bytes																																
	Keywords: Access register, MTRC_STDB, tracer to dmesg, fwtrace to dmesg																																
	Discovered in Version: 16.23.1020																																
1408994	Description: FTE with both forward (FWD) and encapsulation (ENCAP) actions is not supported in the SX NIC Flow Table.																																
	Workaround: N/A																																
	Keywords: SX NIC Flow Table																																
	Discovered in Version: 16.23.1020																																
1350794	<p>Description: Encapsulation / Decapsulation support in steering has the following limitations:</p> <ul style="list-style-type: none"> Encapsulation / Decapsulation can be open on the FDB only if all VFs are non active. Encapsulation / Decapsulation supports single mode only: FDB / NIC. Opening tables of both types is not supported. Encapsulation / Decapsulation per device support: <table border="1"> <thead> <tr> <th></th> <th></th> <th>NIC</th> <th>FDB</th> <th></th> </tr> </thead> <tbody> <tr> <td rowspan="2">ConnectX-4</td> <td>encap</td> <td>NO</td> <td>YES</td> <td>non MH</td> </tr> <tr> <td>decap</td> <td>NO</td> <td>NO</td> <td></td> </tr> <tr> <td rowspan="2">ConnectX-4 Lx</td> <td>encap</td> <td>NO</td> <td>YES</td> <td>non MH</td> </tr> <tr> <td>decap</td> <td>NO</td> <td>YES</td> <td></td> </tr> <tr> <td rowspan="2">ConnectX-5</td> <td>encap</td> <td>YES</td> <td>YES</td> <td></td> </tr> <tr> <td>decap</td> <td>YES</td> <td>YES</td> <td></td> </tr> </tbody> </table>			NIC	FDB		ConnectX-4	encap	NO	YES	non MH	decap	NO	NO		ConnectX-4 Lx	encap	NO	YES	non MH	decap	NO	YES		ConnectX-5	encap	YES	YES		decap	YES	YES	
			NIC	FDB																													
	ConnectX-4	encap	NO	YES	non MH																												
		decap	NO	NO																													
ConnectX-4 Lx	encap	NO	YES	non MH																													
	decap	NO	YES																														
ConnectX-5	encap	YES	YES																														
	decap	YES	YES																														
Workaround: N/A																																	
Keywords: Steering Encapsulation / Decapsulation																																	
Discovered in Version: 16.23.1020																																	
1027553	Description: While using e-switch vport sVLAN stripping, the RX steering values on the sVLAN might not be accurate.																																
	Workaround: N/A																																
	Keywords: e-sw vport sVLAN stripping, RX steering																																
	Discovered in Version: 16.24.1000																																
1799917	Description: Untagged CVLAN packets in the Steering Flow Tables do not match the SVLAN tagged packets.																																
	Workaround: N/A																																
	Keywords: Steering Flow Tables, CVLAN/SVLAN packets																																
	Discovered in Version: 16.23.1020																																

Table 26 - Known Issues (Sheet 6 of 7)

Internal Ref.	Issue
1504073	Description: When using ConnectX-5 with LRO over PPC systems there might be backpressure to the NIC due to delayed PCI writes operations. In this case bandwidth might drop from line-rate to ~35Gb/s. Packet loss or pause frames might also be observed.
	Workaround: Look for an indication of PCI back pressure (“outbound_pci_stalled_wr” counter in ethtools advancing). Disabling LRO helps reduce the back pressure and its effects.
	Keywords: Flow Control, LRO
	Discovered in Version: 16.23.1020
1277762	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.
	Workaround: N/A
	Keywords: Ethernet multicast loopback packet
	Discovered in Version: 16.22.1002
1190753	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.
	Workaround: N/A
	Keywords: RoCE, vport FDB
	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
1168594	Description: RoCE Dual Port Mode (a.k.a Multi-Port vHCA: MPV) is not supported in Multi-Host setups.
	Workaround: N/A
	Keywords: Multi-Port vHCA, Multi-Host
	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
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

Table 26 - Known Issues (Sheet 7 of 7)

Internal Ref.	Issue
1059975	<p>Description: NVMeF limitation:</p> <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 <p>Workaround: N/A</p> <p>Keywords: NVMeF</p> <p>Discovered in Version: 16.22.1010</p>

4 Bug Fixes History

Table 27 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 27 - Bug Fixes History (Sheet 1 of 5)

Internal Ref.	Issue
1803791	Description: On rare occasions, when firmware coalesce Host stuck events occur, a async event might be delayed to be reported, and not be triggered until the next time the PCIe hangs on one of the hosts.
	Keywords: PCIe Error Notification
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1870690	Description: Fixed an issue that resulted in unexpected queue pairs transitioned to error in lossy tests.
	Keywords: RoCE Lossy
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1723460	Description: Limited the maximum amount of dumps created on a PF.
	Keywords: Dump files
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1824111	Description: Renamed the GMP Mellanox Vendor Specific External Capability mask enum from <code>IsDiagnosticCountersSupported</code> to <code>IsDiagnosticDataSupported</code> .
	Keywords: GMP Mellanox Vendor Specific External Capability mask DiagnosticData
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1784290	Description: Fixed a stability issue in RoCE retransmissions under stress affecting Zero-Touch-RoCE.
	Keywords: Zero-Touch-RoCE
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1822787	Description: Fixed an issue that caused a function to misbehave when a PCIe TLP was set with a poisoned indication.
	Keywords: PCIe TLP
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1788895	Description: Fixed an issue that caused large number of packet to drop when running Jumbo frames with TTL rewrite.
	Keywords: Jumbo frames, TTL
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1806874	Description: Limited the number of the elements in the QoS tree 2K. Note: Creating more than 250 Vport_tc for every TC is not allowed.
	Keywords: VQoS tree , 255Vfs
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040

Table 27 - Bug Fixes History (Sheet 2 of 5)

Internal Ref.	Issue
1769862	Description: Fixed the query QP flow. Instead of always taking port 1, FW will now reply the proper port, 1 or 2, for the dual port RoCE net device.
	Keywords: Counters
	Discovered in Version: 16.24.1020
	Fixed in Release: 16.26.1040
1771326	Description: Fixed an issue that caused the nack counters to constantly be reported as "0".
	Keywords: Counters
	Discovered in Version: 16.24.1020
	Fixed in Release: 16.26.1040
1775228	Description: In a rare scenario when the driver is executing the "2err" command and the QP is in SQ drain state, the firmware might post event of broken WQ instead of sending error CQEs on all the WQEs.
	Keywords: QP, WQE, CQE
	Discovered in Version: 16.25.1020
	Fixed in Release: 16.26.1040
1750224	Description: Fixed an issue that resulted in packets sometimes being dropped while setting the software owner Flow Table as root.
	Keywords: Flow Table
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1771921	Description: Fixed an issue that prevented users with non-port owner privilege from using the "read DCBX access registry key" REGID_DCBX_APP/REGID_DCBX_PARAM.
	Keywords: DCBX
	Discovered in Version: 16.24.1000
	Fixed in Release: 16..25.1020
1726413	Description: Fixed an issue that resulted in inaccurate counters' value (negative value) when querying the vPort counters during traffic.
	Keywords: Counters
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1747677	Description: Fixed an issue that prevented the firmware from working with multiple priorities when in Host Chaining mode.
	Keywords: Host Chaining
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1682830	Description: Fixed a rare issue that caused the RX to hang when the server went into the Standby mode (aux power).
	Keywords: RX
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020

Table 27 - Bug Fixes History (Sheet 3 of 5)

Internal Ref.	Issue
-	Description: Fixed a wrong meter rate for small packets by excluding packet CRC size.
	Keywords: Meter rate
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1682727	Description: Fixed an issue that slowed the driver's unload process. The dmesg showed the following message: <i>teardown_hca, resource is still in use, gvmi=XX, type=0x2b</i>
	Keywords: Driver unload
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1717993	Description: Fixed an issue that required pages' calculations on SR-IOV, thus causing the adapter to return 'not enough resources' to the driver's VF.
	Keywords: SR-IOV
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1657685	Description: Fixed an issue that cause an unexpected pause counter to be transmitted when the HCA was in tear down mode.
	Keywords: Pause counter, Ethernet
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1679530	Description: Fixed an issue that cause performance degradation in the system when configured priority to TC mapping where some priorities were mapped to TC7. This issue occurred in scenarios such as hairpin QP in system, e.g. host chaining / ttl WA
	Keywords: Performance
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1615586	Description: Fixed a rare issue that caused the QP to falsely transition into the error state as a result of handling duplicate read/atomic request followed by memory key invalidation.
	Keywords: CQE
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1689159	Description: Fixed the gateway configurations when sending a MAD IB packet after a ETH packet, to prevent MAD transactions timeout when one port was set as IB, and one port as ETH.
	Keywords: Dual port with IB and ETH, MAD Timeout
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1697465	Description: Fixed an issue related to RoCE Lossy Access Register that caused the logic of the condition not to be returned according to the systems' requirements.
	Keywords: RoCE Lossy Access Register
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020

Table 27 - Bug Fixes History (Sheet 4 of 5)

Internal Ref.	Issue
1706349	Description: Fixed an ECN compatibility issue with Broadcom RNICs.
	Keywords: ECN
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1582423	Description: In the case of multi-destinations transmission where the last destination is set to encapsulation & wire, the packet for the last destination will not be send.
	Keywords: Remote mirroring
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1678824	Description: Fixed an issue that prevented the user to enable the port after disabling it in the VF NODNIC.
	Keywords: VF NODNIC
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1606289	Description: Enlarged the number of modify fields to 16 to avoid IPv6 header rewrite failure.
	Keywords: IPv6 header rewrite
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1627973	Description: Fixed an issue that prevented IB QP counters for Acks/Responses from working as a results the NACK/OOS counters showed as zero.
	Keywords: IB QP counters for Acks/Responses
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1554104	Description: Set the stateless offloads cap to be permanently '1'.
	Keywords: Stateless offloads cap
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1541193	Description: Fixed an issue that cause high connection setup latency on guest VMs.
	Keywords: Connection latency
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.25.1020
1554104	Description: Fixed a rare case that cause the transmission side to hang and an assertion in the RDMA ODP.
	Keywords: RDMA ODP
	Discovered in Version: 16.24.1000
	Fixed in Release: 16.24.1000
1434510	Description: Fixed a PCIe bifurcation issue.
	Keywords: PCIe
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.24.1000

Table 27 - Bug Fixes History (Sheet 5 of 5)

Internal Ref.	Issue
1424906	Description: Fixed a deadlock in RX related to the “send-invalidate” flow, that resulted in RX getting stuck.
	Keywords: RX, deadlock
	Discovered in Version: 16.23.1020
	Fixed in Release: 16.24.1000
929504	Description: Fixed rare errors in RX that resulted in double completion.
	Keywords: RX
	Discovered in Version: 16.23.1020
	Fixed in Release: 16.24.1000
1284452/ 1282926	Description: Fixed an issue that caused the mlxconfig tool to present all possible expansion ROM images, instead of presenting only the existing images.
	Keywords: mlxconfig
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.24.1000
1475993	Description: Aligned the default tuning type in PHY TEST MODE to the device protocol.
	Keywords: PHY
	Discovered in Version: 16.23.1020
	Fixed in Release: 16.24.1000
1403211	Description: When a device is operating in Safe Mode state, and the user issues the mlxfwreset command, the device might fail to come-up correctly after the reset. Note: Do not run mlxfwreset when operating in a Safe Mode state.
	Keywords: mlxfwreset
	Discovered in Version: 16.23.1020
	Fixed in Release: 16.24.1000
1295606	Description: Fixed an issue related to PCIe "Surprise link down" event reporting capability.
	Keywords: PCIe
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.24.1000
1434863	Description: Fixed an issue that resulted in the link partner experiencing false active linkup when plugging in a base-T cable to a closed port.
	Keywords: Interfaces
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.24.1000
1424873	Description: Modifying VMQoS rate limiter parameters during traffic might cause transmission failure.
	Keywords: VMQoS, rate limiter
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.24.1000

5 Firmware Changes and New Feature History

Table 28 - Firmware Changes and New Feature History (Sheet 1 of 14)

Feature/Change	Description
Rev. 16.25.1020	
OOB Support AR in IB Networks	Added SL based support for enabling adaptive routing only for traffic running on specific set of SLs. This capability is applicable to application and is enforced by the OpenSM (MADs).
ConnectX-5 OCP 3.0	Added the ability to read from the server whether or not the Socket-Direct is the requested mode on ConnectX-5 OCP 3.0 adapter cards.
ASAP² Offloading VXLAN Decapsulation with HW LRO	Added support for performing hardware Large Receive Offload (HW LRO) on VFs with HW-decapsulated VXLAN. For further information on the VXLAN decapsulation feature, please refer to ASAP ² User Manual under www.mellanox.com -> Products -> Software -> ASAP ² .
Multiple Output Ports - Enhance ASAP²	Added the ability to send to multiple destinations encapsulated data and each destination can have its own encapsulation data.
VSC Security	VSC security includes the mechanisms which will prevent a reasonable host from affecting other hosts from using VSC.
Extended Responder Not Ready (RNR)	Enabled Responder Not Ready (RNR) configuration. Now the number of RNRs can be set higher than 6 and lower than indefinite.
Zero Touch RoCE	Zero touch RoCE enables RoCE to operate on fabrics where no PFC nor ECN are configured. This makes RoCE configuration a breeze while still maintaining its superior high performance.
ODP support for SRQ & XRC	Added support for send opcode operations targeting a SRQ/RMP with the receive WQEs using ODP memory. In case the receive WQE receives an ODP, the device will generate ODP notifications (EQE) and PFAULT will abort CQEs. Note: It is recommended to prefetch the memory used by the receive WQEs to reduce ODP occurrence as these have significant latencies and will cause a performance degradation.
TTL RX	Enables the device to modify incoming packets' TTL from the uplink to a vport using eswitch rules when <code>sw_fdb_ipv4_ttl_modify_enable</code> is set to true.
QP Counters and Firmware Errors per PID	QP counters (RDMA errors) and flow counters (traffic) are now set per PID or UID and available through the driver and the user space tool.
Firmware Burning using DMA Pages	This new capability accelerates the firmware burning process by using Direct Memory Access (DMA) pages.
Auto-Sensing when using 25/10GbE Optical Modules	This new capability accelerates the network to auto-sense the port speed and use it when using a 25/10GbE optical module. Meaning, if the used module is 25GbE but the port is a 10GbE port, the speed used for that network will be 10GbE.
Package ID	Enabled Package ID configuration using server strap according OCP 3.0.
DPDK UIO	This capability provides a solution for improving user space drivers development, generic user space IO device services.
Management Query Information Strings (MQIS)	Added MQIS support, "Part Number" and "Description" information for secured adapter cards.
SR-IOV Offloading	Added the ability to probe one representor each time for a specific representor ID when bonding is enabled.

Table 28 - Firmware Changes and New Feature History (Sheet 2 of 14)

Feature/Change	Description
Virtio Supported Hardware Offloads (DPDK)	<p>Added the following capabilities related to Virtio Supported Hardware Offloads (DPDK):</p> <ul style="list-style-type: none"> • VXLAN encapsulation/decapsulation • VXLAN GPE encapsulation/decapsulation • Packet Modification: <ul style="list-style-type: none"> • NAT (IPv4, IPv6, TCP, UDP) • Routing (L2) • TTL encapsulation/decapsulation • TCP SEQ and ACK • ICMP offloading: match on ICMP type, code and identifier fields • HA and LB: LAG • Flow Aging • Flow Metering • Firmware fast update
SR-IOV Supported Hardware Offloads	<p>Added the following capabilities related to SR-IOV Supported Hardware Offloads:</p> <ul style="list-style-type: none"> • VXLAN encapsulation/decapsulation • VXLAN GPE IPv4 & IPv6 encapsulation/decapsulation • Packet Modification: <ul style="list-style-type: none"> • NAT (IPv4, IPv6, TCP, UDP) • Routing (L2) • TTL encapsulation/decapsulation • TCP SEQ and ACK • VF group rate limit support • VF and PF LAG
Connection Tracking	Added metadata registry C “header modify” and “match rule” support to be used by the connection tracking functionality.
NODNIC Connectivity	Enabled NOIDNIC connectivity for VFs.
VDP	Enabled the option to read adapter’s VPD data (PN, SN etc) from the VM using lspci.
mlxconfig	Renamed the BOOT_RETRY_CNT1 parameter to BOOT_RETRY_CNT.
vPort: Multiple Guest Virtual Machine Identifier (GVMI) per Function	<p>Added multiple e-switch vPort PCI function by GVMI, per vPort for the use cases below:</p> <ul style="list-style-type: none"> • VM Live Migration SW fallback path • Container offloads • Scalable IOV
Reduced Firmware Upgrade Time	<p>Reduced firmware upgrade time using mlxfwreset tool to ~3 seconds. Using this capability requires enabling PARTIAL_RESET_EN in mlxconfig and using MFT version 4.12.0 and up. The “PARTIAL” refers to not resetting the port modules (which is not mandatory for firmware upgrades).</p> <p>Note: Currently this capability only supports firmware upgrade and downgrades to firmware versions newer than XX.25.1020.</p>
Bug Fixes	See Section 4, “Bug Fixes History”, on page 30
Rev. 16.24.1000	
Layer 3 Encapsulation	Added support for an additional layer (Layer 3) of packet processing at the hypervisor level that enables adding and removing protocol headers (e.g., the MAC address is removed during encapsulation, and added during decapsulation) for the encapsulated traffic.
QoS Enhanced Transmission Selection (ETS)	Enabled QoS ETS for systems with 64 VFs to better allocate bandwidth in the NIC.
Time to Live (TTL) Modification	Enabled TTL modification for received packets.

Table 28 - Firmware Changes and New Feature History (Sheet 3 of 14)

Feature/Change	Description
VF Mirroring	Mirrors the traffic from/to one VF to a dedicate admin VF for monitoring and traffic analysis. Note that in this process packets are duplicated and different packet modifications apply to different duplications.
Transmission Histogram Counters	Added support for the transmission histogram counter set as part of the Ethernet extended group counters.
Events Generation by the Hardware upon Counter Incrementation	Enabled the hardware to generate an event upon counter incrementation, in order to reduce an overhead from the software from reading rarely updated counters such as error counters.
NODNIC Connectivity	Enables NOIDNIC connectivity to the network through the e-switch and not directly to the physical port.
QP and Mkey Values	Enabled setting the QP and the Mkey values by the software upon these resources creation.
PCIe Atomic	Enabled advanced PCIe atomic operations. The HCA will perform PCIe atomic operations as a requestor towards the host memory when receiving compatible atomic messages from the network, and according to the configuration of NV_SW_OFFLOAD_CONFIG pci_atomic_mode field and the PCI AtomicOp Requester Enable bit in the Device Control 2 register.
TTL Modification for Rx NIC (Steering)	Enabled TTL modification in the Rx NIC steering. When modifying the TTL in the Rx NIC, the CQE checksum will not recalculated automatically. Note: TTL modification in the FDB for traffic from the network is currently not supported.
TIR Destination from the FDB	Enabled a single TIR destination from the FDB.
WRED	Changed the WRED default mode to OFF for Multi-Host adapter cards.
TX Steering Rule on in WQE Ethernet Segment	Added support for TX steering rule on flow_table_metadata in WQE Ethernet segment.
L3 Encapsulation/Decapsulation in the Reformat Context Allocation	Added L3 encapsulation/decapsulation support in the reformat context allocation. <ul style="list-style-type: none"> L3 encapsulation removes L2 headers and adds generic L3 tunnel encapsulation. L3 decapsulation removes the generic L3 tunnel decapsulation and L2 header.
Flow Steering Header Modification	Added support for flow steering header modification (header rewrite) for IPv4 TTL header for loopback traffic (VF-VF/VF-PF). Note: TTL modification for traffic from the network is currently not supported.
Teardown: Fast Mode	[Developers only] Moved the fast teardown HCA cap bit to offset 0x1c.4:1.
Virtual Functions/QoS	Enabled Virtual Functions to read QPDPM/QPDP/QPTS.
Message vs. Payload based flow control QP Configuration	Added support for requester QP packet based on E2E credits mode. The new flow control supports HCA-to-switch RDMA traffic packet-based End-2-End.
Multi PCI RDMA IB	This capability enables the user to expose two PCI/IB devices per network port.
Bug Fixes	See Section 4, “Bug Fixes History”, on page 30
Steering	Enabled TTL modification in the RX NIC steering. Note: TTL modification in the FDB for traffic from the network is currently not supported.
Rev. 16.23.1020	
RoCE DC	Added support for Dynamically Connected Transport (DC) in RoCE in ConnectX-5 adapter cards.
Mini-CQE Compression	Enabled Rx mini-CQE compressed format for striding RQ.

Table 28 - Firmware Changes and New Feature History (Sheet 4 of 14)

Feature/Change	Description
TX Steering Rules	Enabled TX steering rule on the application meta-data from the WQE. This functionality implements meta-data Reg A steering and meta-data Reg A and Reg B rewriting.
Tunneling Offload/steering Match Criteria	Added MPLS over UDP and MPLS over GRE protocols for tunneling offload/steering match criteria.
InfiniBand	Added support for IPoIB non-default Partition Keys (PKeys). Now the PKey values can be modified in the PKey table without the need of recreating the IPoIB (underlay) QPs.
SR-IOV in Multi-Host/Socket-Direct	[Beta] Added support for SR-IOV (up to 63 VFs) in Multi-Host/Socket-Direct.
Virtualization	Reduced firmware's memory consumption to increase the supported number of VFs per PF to up to 100.
Tools/Driver Version	Added support for QUERY_DRIVER_VERSION command. This command allows the PF driver to query its VFs driver version which was set by the SET_DRIVER_VERSION command.
Resiliency	Shutting Down RDMA QPs with Excessive Retransmissions is a mechanism used to detect excessive retransmissions for an RC connection, and to close the connection in response to it. If the number of retransmissions due to a Local Ack Timeout, NAK-Sequence Error, or Implied NAK, during a specified period, exceeds the specified threshold, the QP will be handled as if the IB spec defined Retry Count was exceeded.
Diagnostic Counters	Added new diagnostic counters to evaluate the number of ICMC hits and misses for particular resources.
Bug Fixes	See Section 4, "Bug Fixes History" , on page 30
Rev. 16.22.1002	
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	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. Note: The flow is currently not supported on Multi host devices, Socket Direct devices and devices running management traffic (NCSI, MCTP).
Steering Discard Packet Counters	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): <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	mlxconfig configuration can now enable or disable forced PCI relaxed ordering in <code>mkey_context</code> . If this feature is enabled, the software per mkey configuration is ignored.

Table 28 - Firmware Changes and New Feature History (Sheet 5 of 14)

Feature/Change	Description
Push/Pop VLAN	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. For further information, please refer to the PRM section <i>Flow Table</i>
QOS "Rate Limit"	Added support to limit the transmission rate of individual InfiniBand port Service Levels. This capability is configurable through a new vendor-specific MAD (QosConfigSL). For more details see Mellanox Vendor Specific MAD Specification Rev 1.4 or above.
Packet Pacing	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. Note: Packet Pacing is not functional in ConnectX-5 Multi host adapter cards.
vport Mirroring	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). The firmware support the following destination list format: 1. new destination vport (analyzer) 2. another Flow Table 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).
Resiliency: Special Error Event	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).
QP's Creation Time	Accelerated QP's creation time.
SR-IOV LID based Routing Mode	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. Note: LID based routing support for vports is supported using SM v4.8.1
Expansion ROM	Added PXE and UEFI to additional ConnectX- 5 adapter cards. ConnectX-5 now holds PXE and x86-UEFI
Host Chaining	Host Chaining allows the user to connect ("chain") one server to another without going through a switch, thus saving switch ports. Host Chaining algorithm is as follow: <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) For Host Chaining limitation, see Known Issue #1178792 in Section 3, "Known Issues" , on page 23
Fast path VLs	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.

Table 28 - Firmware Changes and New Feature History (Sheet 6 of 14)

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 30
Rev. 16.21.2010	
Query vPort Environments (Debug Counters)	Debug counters are a group of counters that handle traffic performance issue related to firmware overhead in transport flow. The following are the additional counters added to this firmware version: <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 30
Rev. 16.21.1000	
Coherent Accelerator Processor Interface (CAPI v2)	[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. Note: This feature is available only with IBM Power 9 CPUs.
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	Added support for Tag Matching Offload with RC/DC transport. 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> "
Hairpin	[Beta] 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.

Table 28 - Firmware Changes and New Feature History (Sheet 7 of 14)

Feature/Change	Description
Virtual Extensible LAN (VXLAN) Encapsulation Offloads over RDMA in SR-IOV	Added support for VXLAN encapsulation offloads over RDMA in SR-IOV. 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.
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.
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 30
Rev. 16.20.1010	

Table 28 - Firmware Changes and New Feature History (Sheet 8 of 14)

Feature/Change	Description
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 30, “Supported Non-Volatile Configurations,” on page 51)
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 30
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 30
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.

Table 28 - Firmware Changes and New Feature History (Sheet 9 of 14)

Feature/Change	Description
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)	NVMeF 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
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.
Dropless 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.

Table 28 - Firmware Changes and New Feature History (Sheet 10 of 14)

Feature/Change	Description
GENEVE & IP-in-IP Stateless Offload	<p>Added support for IP-in-IP and GENEVE network protocols encapsulated into IP frame (L2 tunneling).</p> <p>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.</p> <p>Note: For driver support, please see the Release Notes/User Manual of the relevant OS driver.</p>
Resilient RoCE	<p>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.</p> <p>The ability is accomplished by enabling ECN on both the Switch and the Host.</p>
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 optic 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" >></code> <code>/tmp/power_conf_tlv.cfg</code> • <code>mlxconfig -d <device> -f /tmp/power_conf_tlv.cfg</code> <code>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/mutlihost 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.</p> <p>KSM is used when associating Key Length My Virtual Address (KLMs) with fixed memory size.</p>

Table 28 - Firmware Changes and New Feature History (Sheet 11 of 14)

Feature/Change	Description
NULL Mkey	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.
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.
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-plain in Mellanox NIC hardware (Mellanox Embedded Switch or eSwitch) while maintaining OVS control-plain 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.

Table 28 - Firmware Changes and New Feature History (Sheet 12 of 14)

Feature/Change	Description
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-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)

Table 28 - Firmware Changes and New Feature History (Sheet 13 of 14)

Feature/Change	Description
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
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, please refer to Section 8, on page 51.
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 <ul style="list-style-type: none"> 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 28 - Firmware Changes and New Feature History (Sheet 14 of 14)

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 RMIII support
MAD	<ul style="list-style-type: none"> • Config space address in MAD management class 0x09

6 PreBoot Drivers (FlexBoot/UEFI)

6.1 FlexBoot Changes and New Features

For further information, please refer to FlexBoot Release Notes

<https://docs.mellanox.com> --> Software --> Firmware Management --> PreBoot Drivers (FlexBoot/UEFI)

6.2 UEFI Changes and Major New Features

For further information, please refer to UEFI Release Notes

<https://docs.mellanox.com> --> Software > Firmware Management --> PreBoot Drivers (FlexBoot/UEFI)

7 Unsupported Features and Commands

7.1 Unsupported Features

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

Table 29 - Unsupported Features

Feature	ConnectX-5
The following service types: <ul style="list-style-type: none"> • SyncUMR • Mellanox transport • RAW IPv6 	No
INT-A not supported for EQs only MSI-X	No
PCI VPD write flow (RO flow supported)	No
Streaming Receive Queue (STRQ) and collapsed CQ	No
Subnet Manager (SM) on VFs	No
RoCE LAG in Multi-Host/Socket-Direct	No

7.2 Unsupported Commands

- QUERY_MAD_DEMUX
- SET_MAD_DEMUX
- CREATE_RQ - MEMORY_RQ_RMP
- MODIFY_LAG_ASYNC_EVENT

8 Supported Non-Volatile Configurations

Table 30 - Supported Non-Volatile Configurations

Configuration	mlxconfig Parameter Name	Class	TLV ID
NV_MEMIC_CONF	MEMIC_BAR_SIZE	GLOBAL (0)	0x6
	MEMIC_SIZE_LIMIT		
NV_HOST_CHAINING_CONF	HOST_CHAINING_MODE		0x8
	HOST_CHAINING_DESCRIPTOR		
	HOST_CHAINING_TOTAL_BUFFER_SIZE		
NV_FLEX_PARS_CONF	FLEX_PARSER_PROFILE_ENABLE		0xe
	FLEX_IPV4_OVER_VXLAN_PORT		
NV_ROCE_1_5_CONF	ROCE_NEXT_PROTOCOL		0x10
NV_INTERNAL_RESOURCE_CONF	ESWITCH_HAIRPIN_DESCRIPTOR		0x13
	ESWITCH_HAIRPIN_TOT_BUFFER_SIZE		
NV_GLOBAL_PCI_CONF	NON_PREFETCHABLE_PF_BAR		0x80
	NUM_OF_VFS		
	SRIOV_EN		
	PF_LOG_BAR_SIZE		
	VF_LOG_BAR_SIZE		
	NUM_PF_MSIX		
	NUM_VF_MSIX		
NV_TPT_CONF	INT_LOG_MAX_PAYLOAD_SIZE		0x82
NV_POWER_CONF	SW_RECOVERY_ON_ERRORS	0x88	
	RESET_WITH_HOST_ON_ERRORS		
	ADVANCED_POWER_SETTINGS		
NV_SW_OFFLOAD_CONFIG	CQE_COMPRESSION	0x10a	
	IP_OVER_VXLAN_EN		
	PCI_ATOMIC_MODE		
	LRO_LOG_TIMEOUT0		
	LRO_LOG_TIMEOUT1		
	LRO_LOG_TIMEOUT2		
	LRO_LOG_TIMEOUT3		
NV_IB_DC_CONF	LOG_DCR_HASH_TABLE_SIZE	0x190	
	DCR_LIFO_SIZE		
NV_VPI_LINK_TYPE	LINK_TYPE	PHYSICAL_PORT (2)	0x12

Table 30 - Supported Non-Volatile Configurations

Configuration	mlxconfig Parameter Name	Class	TLV ID
NV_ROCE_CC	ROCE_CC_PRIO_MASK	PHYSICAL_PORT (2)	0x107
	ROCE_CC_ALGORITHM		
NV_ROCE_CC_ECN	CLAMP_TGT_RATE_AFTER_TIME_INC	PHYSICAL_PORT (2)	0x108
	CLAMP_TGT_RATE		
	RPG_TIME_RESET		
	RPG_BYTE_RESET		
	RPG_THRESHOLD		
	RPG_MAX_RATE		
	RPG_AI_RATE		
	RPG_HAI_RATE		
	RPG_GD		
	RPG_MIN_DEC_FAC		
	RPG_MIN_RATE		
	RATE_TO_SET_ON_FIRST_CNP		
	DCE_TCP_G		
	DCE_TCP_RTT		
	RATE_REDUCE_MONITOR_PERIOD		
	INITIAL_ALPHA_VALUE		
	MIN_TIME_BETWEEN_CNPS		
	CNP_802P_PRIO		
CNP_DSCP			
NV_LLDP_NB_CONF	LLDP_NB_DCBX	PHYSICAL_PORT (2)	0x10a
	LLDP_NB_RX_MODE		
	LLDP_NB_TX_MODE		
NV_LLDP_NB_DCBX	DCBX_IEEE	PHYSICAL_PORT (2)	0x18e
	DCBX_CEE		
	DCBX_WILLING		
NV_KEEP_LINK_UP	KEEP_ETH_LINK_UP	PHYSICAL_PORT (2)	0x190
	KEEP_IB_LINK_UP		
	KEEP_LINK_UP_ON_BOOT		
	KEEP_LINK_UP_ON_STANDBY		
NV_QOS_CONF	NUM_OF_VL	PHYSICAL_PORT (2)	0x192
	NUM_OF_TC		
	NUM_OF_PFC		
NV_MPFS_CONF	DUP_MAC_ACTION	PHYSICAL_PORT (2)	0x196
	SRIOV_IB_ROUTING_MODE		
	IB_ROUTING_MODE		

Table 30 - Supported Non-Volatile Configurations

Configuration	mlxconfig Parameter Name	Class	TLV ID
NV_HCA_CONF	PCI_WR_ORDERING	HOST-FUNCTION (3)	0x112
	MULTI_PORT_VHCA_EN		
NV_EXTERNAL_PORT_CTRL	PORT_OWNER		0x192
	ALLOW_RD_COUNTERS		
	RENEG_ON_CHANGE		
	TRACER_ENABLE		
NV_ROM_BOOT_CONF2	IP_VER		0x195
	BOOT_UNDI_NETWORK_WAIT		
NV_ROM_UEFI_CONF	UEFI_HII_EN		0x196
NV_ROM_UEFI_DEBUG_LEVEL	BOOT_DBG_LOG		0x206
	UEFI_LOGS		
NV_ROM_BOOT_CONF1	BOOT_VLAN		0x221
	LEGACY_BOOT_PROTOCOL		
	BOOT_RETRY_CNT		
	BOOT_LACP_DIS		
	BOOT_VLAN_EN		
NV_ROM_IB_BOOT_CONF	BOOT_PKEY	0x222	
NV_PCI_CONF	ADVANCED_PCI_SETTINGS	HOST (7)	0x80
SAFE_MODE_CONF	SAFE_MODE_THRESHOLD		0x82
	SAFE_MODE_ENABLE		