Mellanox ConnectX®-5 Firmware Release Notes

Rev 16.24.4020

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

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

Table of Contents

List of Tables

Release Update History

Table 1 - Release Update History

Release	Date	Description
Rev 16.24.4020	December 02, 2018	Initial version of this firmware release. This version introduces new Changes and Features (see Section 2, "Changes and New Features in Rev 16.24.4020", on page 16 and Section 4, "Bug Fixes History", on page 25).

1 Overview

These are the release notes for the ConnectX®-5 adapters firmware Rev 16.24.4020. This firmware supports the following protocols:

- InfiniBand - SDR, QDR, FDR10, FDR, EDR
- Ethernet - 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 56GbE[1], 100GbE
- PCI Express 4.0, supporting backwards compatibility for v3.0, v2.0 and v1.1

1.1 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.

1.1.1 Validated and Supported 1GbE Cables

Table 2 - 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

1.1.2 Validated and Supported 10GbE Cables

Table 3 - Validated and Supported 10GbE Cables

Speed	Cable OPN #	Description
10GB/S	CAB-SFP-SFP-1M	Arista 10GBASE-CR SFP+ Cable 1 Meter
10GB/S	CAB-SFP-SFP-3M	Arista 10GBASE-CR SFP+ Cable 3 Meter
10GB/S	CAB-SFP-SFP-5M	Arista 10GBASE-CR SFP+ Cable 5 Meter
10GB/S	MC2309124-004	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 4M
10GB/S	MC2309124-005	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 5M
10GB/S	MC2309130-001	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 1M
10GB/S	MC2309130-002	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 2M
10GB/S	MC2309130-003	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 3M
10GB/S	MC2309130-00A	Mellanox Passive Copper Cable ETH 10GBE 10GB/S QSFP TO SFP+ 0.5M
10GB/S	MC2609125-004	Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 4M
10GB/S	MC2609125-005	Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 5M
10GB/S	MC2609130-001	Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 1M

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 3 - Validated and Supported 10GbE Cables

Speed	Cable OPN #	Description
10GB/S	MC2609130-002	Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 2M
10GB/S	MC2609130-003	Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 3M
10GB/S	MC2609130-0A1	Mellanox Passive Copper Hybrid Cable ETH 40GBE TO 4X10GBE QSFP TO 4X SFP+ 1.5M
10GB/S	MC3309124-004	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 4M
10GB/S	MC3309124-005	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 5M
10GB/S	MC3309124-006	Mellanox® Passive Copper Cable, ETH 10GbE, 10Gb/s, SFP+, 6m
10GB/S	MC3309124-007	Mellanox® Passive Copper Cable, ETH 10GbE, 10Gb/s, SFP+, 7m
10GB/S	MC3309130-001	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 1M
10GB/S	MC3309130-002	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 2M
10GB/S	MC3309130-003	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 3M
10GB/S	MC3309130-00A	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 0.5M
10GB/S	MC3309130-0A1	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 1.5M
10GB/S	MC3309130-0A2	Mellanox Passive Copper Cable ETH 10GBE 10GB/S SFP+ 2.5M
10GB/S	MFM1T02A-LR-F	Mellanox Optical Module ETH 10GBE 10GB/S SFP+ LC-LC 1310NM LR up to 10KM
10GB/S	MFM1T02A-SR-F	Mellanox Optical Module ETH 10GBE 10GB/S SFP+ LC-LC 850NM SR up to 300M
10GB/S	SFP-10G-SR	Cisco 10GBASE-SR SFP+ transceiver module for MMF, 850-nm wavelength, LC duplex connector
10GB/S	SFP-H10GB-CU1M	Cisco 1-m 10G SFP+ Twinax cable assembly, passive
10GB/S	SFP-H10GB-CU3M	Cisco 3-m 10G SFP+ Twinax cable assembly, passive
10GB/S	SFP-H10GB-CU5M	Cisco 5-m 10G SFP+ Twinax cable assembly, passive
10GB/S	FTLX1471D3BCL-ME	10GBASE-LR SFP+ 1310nm 10km DOM Transceiver Module
10GE	BN-QS-SP-CBL-5M	40G QSFP+ to 4xSFP+ DAC Breakout Direct Attach Cable 5m
10GE	BN-QS-SP-CBL-5M	40G QSFP+ to 4xSFP+ DAC Breakout Direct Attach Cable 5m

1.1.3 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 4 - Validated and Supported 25GbE Cables

Speed	Cable OPN #	Description
25GbE	FTLF8536P4BCL	Finisar SFP+ transceivers 25Gb/s
25GbE	LTF8507-PC07	Hisense active fiber cable, 25GbE

Table 4 - 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-A01A	Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 1.5m
25GbE	MCP2M00-A02A	Mellanox® Passive Copper cable, ETH, up to 25Gb/s, SFP28, 2.5m
25GbE	MCP7F00-A001	Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 TO 4XSFP28 1M
25GbE	MCP7F00-A002	Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 TO 4XSFP28 2M
25GbE	MCP7F00-A003	Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 TO 4XSFP28 3M
25GbE	MCP7F00-A003-AM	Mellanox® passive copper hybrid cable, ETH 100GbE to 4x25GbE, QSFP28 to 4xSFP28, 3M 30AWG
25GbE	MCP7F00-A005AM	Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 to 4XSFP28 5M
25GbE	MCP7F00-A01A	Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 to 4XSFP28 1.5M
25GbE	MCP7F00-A02A	Mellanox Passive Copper Hybrid Cable ETH 100GBE TO 4X25GBS QSFP28 to 4XSFP28 2.5M
25GbE	MFA2P10-Axxx	Mellanox® active optical cable 25GbE, SFP28, up to 100m
25GbE	MMA2P00-AS	Mellanox® transceiver, 25GbE, SFP28, LC-LC, 850nm, SR, up to 100m
25GbE	SFP-H25G-CU1M	25GBASE-CR1 Copper Cable 1-meter
25GbE	SFP-H25G-CU2M	25GBASE-CR1 Copper Cable 2-meter
25GbE	SFP-H25G-CU3M	25GBASE-CR1 Copper Cable 3-meter
25GbE	MMA2P00-AS	Mellanox® transceiver, 25GbE, SFP28, LC-LC, 850nm, SR, up to 100m
25GbE	MFA2P10-A100	Mellanox® active optical cable 25GbE, SFP28, 100m
25GbE	MFA7A50-C030	Mellanox Active Fiber Hybrid Solution ETH 100GBE TO 4X25GBE QSFP28 TO 4XSFP28 30M
25GbE	MCP2M00-A005E26L	Mellanox Passive Copper Cable, ETH, UP TO 25GB/S, SFP28, 5M, BLACK, 26AWG, CA-L

1.1.4 Validated and Supported 40GbE Cables

Table 5 - Validated and Supported 40GbE Cables

Speed	Cable OPN #	Description
NA	MAM1Q00A-QSA	Mellanox® cable module, ETH 10GbE, 40Gb/s to 10Gb/s, QSFP to SFP+
NA	MAM1Q00A-QSA28	Mellanox® cable module, ETH 25GbE, 100Gb/s to 25Gb/s, QSFP28 to SFP28
40GbE	MC2210126-004	Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 4m
40GbE	MC2210126-005	Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 5m
40GbE	MC2210128-003	Mellanox Passive Copper Cable ETH 40GBE 40GbE QSFP 3M
40GbE	MC2210130-001	Mellanox Passive Copper Cable ETH 40GBE 40GbE QSFP 1M
40GbE	MC2210130-002	Mellanox Passive Copper Cable ETH 40GBE 40GbE QSFP 2M
40GbE	MC2210130-00A	Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 0.5m
40GbE	MC2210130-00B	Mellanox® Passive Copper Cable, ETH 40GbE, 40GbE, QSFP, 0.75m
40GbE	MC2210310-XXX	Mellanox Active Fiber Cable ETH 40GBE 40GbE QSFP from 3M up to 100M
40GbE	MC2210411-SR4L	Mellanox Optical Module 40GbE QSFP MPO 850NM UP TO 30M
40GbE	MC2210411-SR4	Mellanox Optical Module 40GbE QSFP MPO 850NM UP TO 100M
40GbE	MC2210411-SR4E	Mellanox Optical Module 40GbE QSFP MPO 850NM UP TO 300M
40GbE	QSFP-40G-SR-BD	Cisco 40GBASE-SR-BiDi, duplex MMF
40GbE	QSFP-40G-SR4	Cisco 40GBASE-SR4, 4 lanes, 850 nm MMF
40GbE	QSFP-H40G-ACU10M	Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 10-meter, active
40GbE	QSFP-H40G-AOC10M	Cisco 40GBase-AOC QSFP direct-attach Active Optical Cable, 10-meter
40GbE	QSFP-H40G-CU1M	Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 1-meter, passive
40GbE	QSFP-H40G-CU3M	Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 3-meter, passive
40GbE	QSFP-H40G-CU5M	Cisco 40GBASE-CR4 QSFP direct-attach copper cable, 5-meter, passive

1.1.5 Validated and Supported 50GbE Cables

Table 6 - 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

1.1.6 Validated and Supported 100GbE Cables

Table 7 - Validated and Supported 100GbE Cables

Speed	Cable OPN #	Description
100GbE	MCP1600-C001	Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 1M
100GbE	MCP1600-C002	Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 2M
100GbE	MCP1600-C003	Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 3M
100GbE	MCP1600-C005AM	Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP, 5m, 26AWG
100GbE	MCP1600-C005E26L	Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP28, 5m, Black, 26AWG, CA-L
100GbE	MCP1600-C00A	Mellanox Passive Copper Cable ETH 100GBE 100GBS QSFP LSZH 0.5M
100GbE	MCP1600-C01A	Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP, LSZH, 1.5m
100GbE	MCP1600-C02A	Mellanox® Passive Copper cable, ETH 100GbE, 100GbE, QSFP, LSZH, 2.5m
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	MMA1L10-CR	Mellanox® optical transceiver, 100GbE, 100GbE, QSFP28, LC-LC, 1310nm, LR4 up to 10km **Note:** Only revision A2 and above.
100GbE	MMA1L30-CM	Mellanox® optical module, 100GbE, 100GbE, QSFP28, LC-LC, 1310nm, CWDM4, up to 2km
100GbE	MMS1C00-C500	Mellanox® transceiver, 100GbE, QSFP28, MPO, 1550nm PSM4, up to 2km

Table 7 - Validated and Supported 100GbE Cables

Speed	Cable OPN #	Description
100GbE	MMS1C00-C500	Mellanox® transceiver, 100GbE, QSFP28, MPO, 1550nm PSM4, up to 2km
100GbE	MMS1C00-CM	Mellanox® transceiver, 100GbE, QSFP28, MPO, 1550nm PSM4, up to 2km for internal use only
100GE	CAB-Q-Q-100G-3M	Passive 3 meter , QSFP+ to QSFP+ QSFP100 TWINAX 103.125Gbps-CR4
100GE	QSFP-40/100-SRBD	100G and 40GBASE SR-BiDi QSFP Transceiver, LC, 100m OM4 MMF
100GE	MCP1650-H001E30	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP28, PVC, 1m, white pultab, 30AWG
100GE	MCP1650-H01AE30	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP28, PVC, 1.5m, white pultab, 30AWG

1.1.7 Validated and Supported QDR Cables

Table 8 - 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

1.1.8 Validated and Supported FDR10 Cables

Table 9 - Validated and Supported FDR10 Cables

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

1.1.9 Validated and Supported EDR Cables

Table 10 - Validated and Supported EDR Cables

Speed	Cable OPN #	Description
EDR	MCP1600-E001[a]	Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 1M
EDR	MCP1600-E002[a]	Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 2M

Table 10 - Validated and Supported EDR Cables

Speed	Cable OPN #	Description
EDR	MCP1600-E003	Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 3M
EDR	MCP1600-E004A26	Mellanox® Passive Copper cable, IB EDR, up to 100Gb/s, QSFP28, 4m, Blue, 26AWG
EDR	MCP1600-E00A[a]	Mellanox Passive Copper Cable VPI 100GB/S QSFP LSZH 0.5M
EDR	MCP1600-E01A[a]	Mellanox® Passive Copper cable, VPI, up to 100Gb/s, QSFP, LSZH, 1.5m
EDR	MCP1600-E02A	Mellanox® Passive Copper cable, VPI, up to 100Gb/s, QSFP, LSZH, 2.5m
EDR	MFA1A00-E005[a]	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 5m
EDR	MFA1A00-E010[a]	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 10m
EDR	MFA1A00-E015[a]	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 15m
EDR	MFA1A00-E020	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 20m
EDR	MFA1A00-E030	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 30m
EDR	MFA1A00-E050	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 50m
EDR	MFA1A00-E100	Mellanox Active Fiber Cable, VPI, up to 100Gb/s, QSFP, 100m
EDR	MMA1B00-E100	Mellanox® Transceiver, IB EDR, up to 100Gb/s, QSFP28, MPO, 850nm, up to 100m
EDR	MCP1650-H001E30	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP28, PVC, 1m, white pultab, 30AWG
EDR	MCP1650-H01AE30	Mellanox® Passive Copper cable, IB HDR, up to 200Gb/s, QSFP28, PVC, 1.5m, white pultab, 30AWG

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

1.2 Tested Switches

1.2.1 Tested 10GbE Switches

Table 11 - 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

1.2.2 Tested 40GbE Switches

Table 12 - 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

1.2.3 Tested 100GbE Switches

Table 13 - 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

Table 13 - Tested 100GbE Switches

Speed	Switch Silicon	OPN # / Name	Description	Vendor
100GbE	N/A	CE8860-4C-EI	24x10GE (SFP+) or 25GE (SFP28) and 2x100GE switch	Huawei
100GbE	Spectrum	SN2410-CB2F	48-port 25GbE + 8-port 100GbE Open Ethernet ToR Switch System	Mellanox
100GbE	Spectrum	SN2700-CS2R	32-port Non-blocking 100GbE Open Ethernet Spine Switch System	Mellanox
100GbE	Spectrum	SN2740-CB2F1	32-port Non-blocking 100GbE Open Ethernet Spine Switch System	Mellanox
100GbE	N/A	Wedge 100-32X R04	32-port 100G QSFP28 - Leaf/Spine Switch, power-to-port airflow, DC Power	Edgecore
100GbE	N/A	QFX5200-32C-32	32-port 100GbE Ethernet Switch System	Juniper
100GbE	N/A	S6820-56HF	48 SFP+ + 8 QSFP Ports 100GE Switch Ethernet	H3C

1.2.4 Tested QDR Switches

Table 14 - 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

1.2.5 Tested FDR Switches

Table 15 - 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

1.2.6 Tested EDR Switches

Table 16 - Tested EDR Switches

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

1.3 Tools, Switch Firmware and Driver Software

Firmware Rev 16.24.4020 is tested with the following tools, Switch firmware, and driver software:

Table 17 - Tools, Switch Firmware and Driver Software

	Supported Version
MLNX_OFED	4.5-1.0.1.0
MLNX_EN (MLNX_OFED based code)	4.5-1.0.1.0
WinOF-2	2.10
MFT	4.11.0
VMware	• ESXi 6.7 v4.17.13.8 • ESXi 6.5 v4.16.13.5 • ESXi 6.0 v4.15.13.2
MLNX-OS	• Switch-IB: 3.7.1000 • Switch-IB 2: 3.7.1000 • Spectrum: 3.7.1000
Quantum™ Firmware	27.1920.0620
Spectrum™ Firmware	13.1910.0618
SwitchX-IB™ Firmware	11.1910.0618
SwitchX-IB 2 Firmware	15.1910.0618

1.4 Supported FlexBoot, UEFI

Firmware Rev 16.24.4020 supports the following FlexBoot:

Table 18 - Supported FlexBoot, UEFI

Expansion ROM	Supported Version
FlexBoot	3.5.603
UEFI	14.17.13

1.5 Revision Compatibility

Firmware Rev 16.24.4020 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 16.24.4020

Table 19 - Changes and New Features in Rev 16.24.4020

Feature/Change	Description
TTL RX	Enables the device to modify incoming packets' TTL from the uplink to a vport using eswitch rules when sw_fdb_ipv4_ttl_modify_enable 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/ 10G Optical Modules	This new capability accelerates the network to auto-sense the port speed and use it when using a 25/10G optical module. Meaning, if the used module is 25G but the port is a 10G port, the speed used for that network will be 10G.
Package ID	Enabled Package ID configuration using server strap according OCP 3.0.
Management Query Information Strings (MQIS)	Added MQIS support, “Part Number” and “Description" information for secured adapter cards.
Bug Fixes	See Section 4, “Bug Fixes History”, on page 25

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 20 - 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	-	64	-	127

Table 21 - 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

Table 22 - Known Issues (Sheet 1 of 8)

Internal Ref.	Issue
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
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.
	Workaround: N/A
	Keywords: Remote mirroring
	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

Table 22 - Known Issues (Sheet 2 of 8)

Internal Ref.	Issue
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
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
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
1446378	**Description:** RoCE DC is not supported in LAG mode.
	Workaround: N/A
	Keywords: RoCE DC, LAG
	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

Table 22 - Known Issues (Sheet 3 of 8)

Internal Ref.	Issue
1350794	**Description:** Encapsulation / Decapsulation support in steering has the following limitations: • 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: 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
1316221	**Description:** Health counter increases every 50ms instead of 10ms.
	Workaround: N/A
	Keywords: Health counter
	Discovered in Version: 16.22.1002
1316221	**Description:** In very rare cases, triggering a function level reset while running NVMf offload traffic might cause a response capsule that carries a bad command identifier of 0 to be sent.
	Workaround: N/A
	Keywords: NVMF offload
	Discovered in Version: 16.22.1002
1298377	**Description:** When a packet is sent on a non-native port, a LAG or a RoCE dual port, and it reaches the ingress mirroring entry, the packet sends the RX a meta data loopback syndrome, on the non-native port, resulting in the packet reaching the wrong meta_data table.
	Workaround: N/A
	Keywords: Steering, mirroring
	Discovered in Version: 16.22.1002
1306342	**Description:** Signature-accessing WQEs sent locally to the NVMeF target QPs that encounter signature errors, will not send a SIGERR CQE.
	Workaround: N/A
	Keywords: Signature-accessing WQEs, NVMeF target
	Discovered in Version: 16.22.1002
1308236	**Description:** Packet Pacing is not functional in ConnectX-5 Multi host adapter cards.
	Workaround: N/A
	Keywords: Packet Pacing, ConnectX-5 Multi host cards
	Discovered in Version: 16.22.1002

Table 22 - Known Issues (Sheet 4 of 8)

Internal Ref.	Issue
1309104	**Description:** ParaVport is not supported in ConnectX-5.
	Workaround: N/A
	Keywords: ParaVport
	Discovered in Version: 16.22.1002

1178792	**Description:** Host Chaining Limitations:
	• MAC address must not be changed
	• Both ports should be configured to Ethernet when host chaining is enabled.
	• The following capabilities cannot function when host chaining is enabled:
	  • SR-IOV
	  • DSCP
	  • NODNIC
	  • Load balancing
	  • LAG
	  • Dual Port RoCE (multi port vHCA)

	Workaround: N/A
	Keywords: Host Chaining
	Discovered in Version: 16.22.1002
1277762	**Description:** An Ethernet multicast loopback packet is not counted (even if it is not a local loopback packet) when running the `nic_receive_steering_discard` 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: Dual-port VHCA, RoCE packet, vport FDB
	Discovered in Version: 16.22.1002
1079027/1126921	**Description:** Occasionally, when adding module info page for Bell 1G BaseT module to the mlxlink data, the information is not updated correctly.
	Workaround: N/A
	Keywords: Bell 1G BaseT module, mlxlink
	Discovered in Version: 16.21.2010
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

Table 22 - Known Issues (Sheet 5 of 8)

Internal Ref.	Issue
1133394	**Description:** Running loopback and outbound traffics simultaneously in Multi-Host setups results in inbound/outbound traffic consuming more bandwidth than loopback traffic.
	Workaround: N/A
	Keywords: Loopback and outbound traffic, bandwidth
	Discovered in Version: 16.21.1000
1171013	**Description:** Signature Handover Operations is not supported when FPP (Function-Per-Port) mode is disabled.
	Workaround: N/A
	Keywords: Signature Handover Operations, FPP
	Discovered in Version: 16.21.1000
1176407/1171665	**Description:** Secure Firmware devices block any "not secure" access to the CR-Space, private ICMDs and Flash random access commands. Thus, tools/capabilities such as ibdump or wqedump or packet sniffing do not function properly.
	Workaround: Load a secure customer token to use any blocked tools.
	Keywords: Secure Firmware
	Discovered in Version: 16.21.1000
1119458	**Description:** When RoCE Dual Port Mode is enable, if VHCA0 has a different VLAN stripping configuration on its E-SW vport context than the configuration on the affiliated vport (VHCA1), the NIC steering on the VLAN might be wrong for the single-port VHCA traffic.
	Workaround: N/A
	Keywords: VLAN, dual port affiliation
	Discovered in Version: 16.21.1000
1145910	**Description:** Using E-SW VLAN insertion with a different priority than the QP's native priority might cause QoS issues. Issues might be observed while using PUSH VLAN action in Flow Steering.
	Workaround: N/A
	Keywords: E-SW VLAN Insertion
	Discovered in Version: 16.21.1000
1163425	**Description:** Running mlxfwreset on ConnectX-5 Socket-Direct adapter cards on Windows OS is currently not functional.
	Workaround: Reboot the server
	Keywords: mlxfwreset, ConnectX-5 Socket-Direct
	Discovered in Version: 16.21.1000
1043521	**Description:** Upgrading old ConnectX-5 firmware using MFT tools version 4.8.0 and above, requires a one time, non failsafe operation. If this is performed, a server reboot is needed (not mlxconfig). **Note:** Not rebooting the server after the upgrade will result in future updates done using mlxconfig to be lost (until server reboot).
	Workaround: N/A
	Keywords: MFT, Firmware Update
	Discovered in Version: 16.21.1000

Table 22 - Known Issues (Sheet 6 of 8)

Internal Ref.	Issue
912315	**Description:** In standby (WoL) mode, although the total IC consumption is as set, the actual current consumption in 3.3V rail is higher in 20-40mA than the set values. Consumption will be decreased in the next firmware release.
	Workaround: N/A
	Keywords: WoL, IC consumption
	Discovered in Version: 16.21.1000
1031669	**Description:** Using Flow Tag in Flow Table Entry together with Header Rewrite is currently not supported.
	Workaround: N/A
	Keywords: Header Rewrite, Flow Tag
	Discovered in Version: 16.21.1000
1122414	**Description:** In a Multi-Host setup, in case of multicast traffic, if a host modifies the packet's format, other hosts might receive the modified packet instead of the original packet.
	Workaround: N/A
	Keywords: Multi-Host, multicast traffic
	Discovered in Version: 16.21.1000
1122452	**Description:** After switching to CAPI mode the driver cannot be unloaded and a server reboot through the OS might hang the server.
	Workaround: To restart the server perform a power cycle through OpenBMC.
	Keywords: CAPI
	Discovered in Version: 16.21.1000
1110219	**Description:** In some cases, CAPI page miss flow is not resolved.
	Workaround: N/A
	Keywords: CAPI
	Discovered in Version: 16.21.1000
1068382	**Description:** NVMeF target offload is not supported when the device is configured to support more than 32 functions (VFs + PFs).
	Workaround: N/A
	Keywords: NVMeF target offload, VFs, PFs
	Discovered in Version: 16.21.1000
1072337	**Description:** If a packet is modified in e-sw flow steering, the SX sniffer Flow Table (of the VF) will see the sniffed packet after the modification.
	Workaround: N/A
	Keywords: SX sniffer Flow Table
	Discovered in Version: 16.21.1000
1070650	**Description:** A multicast packet that arrives from e-sw loopback (VF2VF) and hits some FTE with MODIFY action in the FDB is sent to the source vport (for local loopback) with the modified headers.
	Workaround: N/A
	Keywords: Multicast packet
	Discovered in Version: 16.20.1010

Table 22 - Known Issues (Sheet 7 of 8)

Internal Ref.	Issue

1077244	Description: In the following cases, "rx_buffer_passed_thres_phy" would not indicate fullness even when "rx_discards_phy" indicates a drop:
	• For single port devices:
	  • Flow control is not enabled on the buffer
	  • fullness_threshold is configured to 99%
	  • MTU size smaller then 4K
	• For dual port devices:
	  • Flow control is not enabled on the buffer
	  • fullness_threshold is configured to 99% when MTU size is smaller than 6.5K
	  • fullness_threshold is configured to 98% when MTU size is smaller than 4K

	Workaround: N/A
	Keywords: rx_buffer_passed_thres_phy
	Discovered in Version: 16.20.1010
1090492	**Description:** FLR is not supported when a function receives NVMeF traffic on the target side.
	Workaround: N/A
	Keywords: FLR, NVMeF
	Discovered in Version: 16.21.1000
1095081/1093055	**Description:** Latency sensitive is not supported in NVMeF and Tag Matching offload QP. **Note:** Enabling this feature can cause the machine to hang.
	Workaround: N/A
	Keywords: NVMeF, Tag Matching, Latency sensitive
	Discovered in Version: 16.21.1000
1114798	**Description:** Tag Matching DC transport does not support GRH.
	Workaround: N/A
	Keywords: Tag Matching DC, GRH
	Discovered in Version: 16.21.1000
1086254/1090475	**Description:** Packet header rewrite (modification) and packet encapsulation are not supported when one ConnectX-5 [Ex] VPI port is configured as InfiniBand and the other as Ethernet.
	Workaround: N/A
	Keywords: ConnectX-5 VPI; ConnectX-5 Ex VPI; InfiniBand; Ethernet; Ports; Header Rewrite, Packet encapsulation
	Discovered in Version: 16.20.1010
1063904	**Description:** Messages with mkey signature on offset > 4GB are not supported.
	Workaround: N/A
	Keywords: Signature retransmission
	Discovered in Version: 16.20.1010

Table 22 - Known Issues (Sheet 8 of 8)

Internal Ref.	Issue
1059975	**Description:** NVMeF limitation: • Transaction size - up to 128KB per IO (non-inline) • Support up to 16K connections • Support single namespace per drive • Staging buffer size must be at least 16MB in order to allow SRQ size of 64 entries
	Workaround: N/A
	Keywords: NVMeF
	Discovered in Version: 16.20.1010
1059975	**Description:** When using NVMeF, DESTORY XRQ command failure might occur.
	Workaround: N/A
	Keywords: NVMeF
	Discovered in Version: 16.20.1010
1031744	**Description:** Same flow counter cannot be used on different table types.
	Workaround: N/A
	Keywords: Flow counter
	Discovered in Version: 16.20.1010
1009067	**Description:** In case of an ip_protocol match (on UDP/TCP) related to fragmented packet, the l4_type match might be missed when the hardware steering does not see the L4 headers.
	Workaround: Add to the driver ip_frag match for all steering rules that use ip_protocol match.
	Keywords: ip_protocol match, L4 headers
	Discovered in Version: 16.20.1010

4 Bug Fixes History

Table 23 lists the bugs fixed in this release. For a list of old firmware Bug Fixes, please see ConnectX-5 Firmware Archived Bug Fixes file http://www.mellanox.com/pdf/firmware/ConnectX5-Firmware_Archived_Bug_Fixes.pdf)

Table 23 - Bug Fixes History (Sheet 1 of 9)

Internal Ref.	Issue
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: 16.24.4020
1434510	**Description:** Fixed a PCIe bifurcation issue.
	Keywords: PCIe
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.24.1000
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

Table 23 - Bug Fixes History (Sheet 2 of 9)

Internal Ref.	Issue
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
1288757	**Description:** Closed the vport as part of the fast teardown flow, to prevent Ack to be sent without been scatter to memory.
	Keywords: Fast Teardown Flow
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.23.1020
1356954	**Description:** Fixed an issue that occasionally caused the link to drop after running the 'reboot' command due to I2C being stuck.
	Keywords: I2C stuck, link in polling
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.23.1020
1390343	**Description:** Fixed an issue that caused the tx_pause_storm_warning_events and the tx_pause_storm_error_events counters to increase even when there was TX pause.
	Keywords: TX pause storm events
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.23.1020
1391584	**Description:** Fixed an issue related to the QoSConfig MADs that occasionally caused the firmware to issue an assert under certain device configurations.
	Keywords: QoSConfig
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.23.1020
1333823	**Description:** Fixed an issue where on high loaded Power9 systems with On-Demand-Paging (ODP) enabled, QP might have entered an error state in case the retry counter was exceeded during page fault recovery. As a result, if the page was fetched after QP has already been closed, page_fault_resume error message would appear.
	Keywords: ODP, PPC, Power9
	Discovered in Version: 16.21.2010
	Fixed in Release: 16.23.1020

Table 23 - Bug Fixes History (Sheet 3 of 9)

Internal Ref.	Issue
1333168	**Description:** Fixed an issue that resulted in a wrong sensor state report when running the PLDM sensor reading command and when a negative temperature was set with the hook.
	Keywords: Temperature
	Discovered in Version: 16.22.1002
	Fixed in Release: 16.23.1020
1342760	**Description:** Fixed a speed issue for link mode Speed25GBaseSR.
	Keywords: Link speed
	Discovered in Version: 16.21.2010
	Fixed in Release: 16.23.1020
1063148	**Description:** Pause duration: Physical port counters count in 512bits quantas, instead of microseconds.
	Keywords: Pause duration, Physical port counters
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.23.1020
1231791	**Description:** Fixed an issue that caused the driver to return a wrong logical OR of the 2 physical ports, when querying the vport state when the LAG was enabled.
	Keywords: LAG, vport
	Discovered in Version: 16.21.2010
	Fixed in Release: 16.22.1002
1252833	**Description:** Increased the Full Wire Speed (FWS) threshold value to improve EDR link results.
	Keywords: Full Wire Speed (FWS) threshold, EDR
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.22.1002
1260985	**Description:** Added the option to avoid reconfiguration of QoS tables upon link toggling to reduce packet loss and improve performance.
	Keywords: ECN, QoS
	Discovered in Version: 16.21.2010
	Fixed in Release: 16.22.1002
1262477	**Description:** Fixed an issue that caused traffic to hang when Responder Not Ready (RNR) flow was used.
	Keywords: RoCE Lossy, ECN
	Discovered in Version: 16.21.2010
	Fixed in Release: 16.22.1002
1080868/ 1109484	**Description:** Tag Matching supports up to 16K connections.
	Keywords: Tag Matching
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.22.1002

Table 23 - Bug Fixes History (Sheet 4 of 9)

Internal Ref.	Issue
1084581	**Description:** Target NVMEoF offload for 4 SSDs are 950K IOPS in ConnectX-5 Ex.
	Keywords: Performance
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.22.1002
1096454	**Description:** The HCA does not always identify correctly the presets at the 8G EQ TS2 during speed change to Gen4. As a result, the initial Gen4 Tx configuration might be wrong which might cause speed degrade to Gen1.
	Keywords: Gen4 TX configuration
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.22.1002
1281622	**Description:** Fixed an issue that resulted in “Destroy LAG” command failure if a VFs received an FLR while its affinity QPs were open.
	Keywords: ECMP / SR-IOV LAG
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.22.1002
1172293	**Description:** When RoCE Dual Port mode is enabled, tcpdump is not functional on the 2nd port.
	Keywords: Dual Port vHCA, Multi-port, RoCE Dual Port
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.22.1002
1184961/ 1076206	**Description:** Fixed an issue that occasionally cased the keepalive packet to fail and the FIO connection to disconnect (error =5).
	Keywords: FIO, NVMeF
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.22.1002
1124226	**Description:** Fixed an issue that caused QP connection timeout due to firmware not being able to handle duplicate packets with AckReq bit set. The fix stopped ignoring duplicate AckReq packets to avoid timeout on the sender side.
	Keywords: QP connection timeout
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.21.2010
1179155	**Description:** MPFS load balance (`DUP_MAC_ACTION==LOAD_BALANCE(1)`) is not working as long as `IB_ROUTING_MODE/SRIOV_IB_ROUTING_MODE` is configured to LID.
	Keywords: MPFS load balance, LID, GID
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.21.2010

Table 23 - Bug Fixes History (Sheet 5 of 9)

Internal Ref.	Issue
1155392	**Description:** Fixed an issue that caused a SX engine deadlock (the SX engine handles software port/ priority changes for a specific Send Queue) when more than a single SX engine handled the "prio diff" flow simultaneously, thus caused the hardware to get stuck. The issue happened as the firmware releases the SX engine and waits for various operation to complete. However, due to a race that allows a different transaction to get into the SX engine and cause the lock to be taken by it, the SX engine release is prevented.
	Keywords: SX engine deadlock, system hang
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.21.2010
1168271	**Description:** Fixed an issue that caused the system to hang while changing QPTS/QPDPM/QPDP parameters during traffic.
	Keywords: QPTS/QPDPM/QPDP, system hang
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.21.2010
1190215	**Description:** Fixed unfairness between senders in RoCE LAG while ECN is configured.
	Keywords: RoCE LAG, ECN
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.21.2010
1121688	**Description:** Fixed an issue which displayed diagnostic counters only on the adapter that was initialized first, which is the counters' owner. The owner received correct values, while the other adapter only received zeros.
	Keywords: Diagnostic counters
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.21.2010
1175146	**Description:** Fixed an issue that caused the rdma_cm traffic to fail on the 2nd port when more than 32 VFs were configured, when the RoCE Dual Port vHCA (a.k.a Multi-Port vHCA: MPV) feature was enabled.
	Keywords: MPV
	Discovered in Version: 16.21.1000
	Fixed in Release: 16.21.2010
1167218/ 1168567	**Description:** Fixed an issue related to RDMA_CM driver that might have caused the QP Rate Limit to be activate unexpectedly and reduce the bandwidth significantly on this QP.
	Keywords: Performance
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.21.1000
1122718	**Description:** Fixed an issue that caused low throughput when ECN was enabled in a many-to-one scenario.
	Keywords: ECN
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.21.1000

Table 23 - Bug Fixes History (Sheet 6 of 9)

Internal Ref.	Issue
1149487	**Description:** Fixed an issue that caused ConnectX-5 Ex Virtual Function to be recognized as a ConnectX-5 Virtual Function device.
	Keywords: Virtualization
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.21.1000
1090723	**Description:** Fixed an issue that wrongly reported the maximum temperature in a setup as the current temperature regardless of the actual temperature.
	Keywords: PCI Gen4 receiver
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.21.1000
1134407	**Description:** PCI Gen4 receiver stability enhancements.
	Keywords: PCI Gen4 receiver
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.21.1000
1099880	**Description:** Disabled the option to write to the protected modules to avoid receiving NACK upon module initialization.
	Keywords: Thermal temperature
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.21.1000
1014078	**Description:** Enabled connecting 5m 40GbE cables to SwitchX, SwitchX-2 based switches.
	Keywords: Cables, SwitchX, SwitchX-2
	Discovered in Version: 16.19.1200
	Fixed in Release: 16.21.1000
1054335/ 1054671	**Description:** Fixed the issue where when using UD RoCE multicast traffic over SR-IOV, packets were scattered to all the attached QPs in the e-sw (PF and its VFs) and not only on the vport that was specified in the e-se FDB.
	Keywords: UD RoCE multicast traffic, SR-IOV
	Discovered in Version: 16.20.1010
	Fixed in Release: 16.21.1000
1060650	**Description:** Fixed a link issue on Intel 10GbE Optical module PN: R8H2F, Y3KJN.
	Keywords: Intel 10GbE Optical module
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
1047533	**Description:** Fixed an issue that caused the TX traffic not to send packets when using VF index (ARI) bigger than 127.
	Keywords: VFs
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010

Table 23 - Bug Fixes History (Sheet 7 of 9)

Internal Ref.	Issue
1009614	**Description:** Fixed a scaling issue with more than 1k QPs for ECN by moving from per QP caching to per IP to allow better scale with number of host in the fabric.
	Keywords: Performance
	Discovered in Release: 16.19.1200
	Fixed in Release: 16.20.1010
1041108	**Description:** Enabled firmware resync of the internal clocks after getting out of the standby mode to prevent PTP time sync from getting out of sync after system warm-rebooted due to system getting into a low-power (standby) mode.
	Keywords: PTP time sync, standby mode
	Discovered in Release: 16.19.1200
	Fixed in Release: 16.20.1010
1047693	**Description:** When running RoCE over VRRP, enabled the device to receive RoCE packet with different source MAC than the original RoCE packet's destination MAC, to allow routing between different subnets.
	Keywords: RoCE over VRRP, Destination MAC
	Discovered in Release: 16.19.1200
	Fixed in Release: 16.20.1010
1050234	**Description:** Fixed an issued that caused LLDP not to enable PFC configuration currently when DCBX transitioning flow control configurations was set from Global Pause to PFC.
	Keywords: RoCE Lossy & ECN
	Discovered in Release: 16.19.1200
	Fixed in Release: 16.20.1010
1047533	**Description:** Rephrased and improved external troubleshoot messages in PDDR register.
	Keywords: PDDR register
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
999261	**Description:** Improved SR-IOV performance.
	Keywords: SR-IOV
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
954822	**Description:**: The `ipoib_enhanced_offloads` indication in the HCA capabilities reports `0` while `SRIOV_EN=1`.
	Keywords: SR-IOV, IPoIB Offloads
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010

Table 23 - Bug Fixes History (Sheet 8 of 9)

Internal Ref.	Issue
981598	**Description:** Fixed an issue on an ETH port with SR-IOV enabled that prevented packets from reaching the BMC (failure in steering loopback resolution) if the BMC addresses were configured after VF initialization, and the VF was trying to send traffic to the BMC (that located on the same phy port).
	Keywords: BMC, SR-IOV, packets
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
906144	**Description:** Fixed an issue which caused the rate limiter not to function when setting a rate to tc 7.
	Keywords: QOS - ETH - rate limit per TC
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
893261	**Description:** Fixed the PCIe TX glitch during Recovery.Speed state of the link training to PCIe Gen3.
	Keywords: PCIe TX glitch
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
1002190	Description: Fixed an issue related to the PortRcvDataVLExtended/PortXmitDataVLExtended parameter that caused the counters' value to be reported in octets instead of dwrods.
	Keywords: Counters
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
1025741/ 781339/ 1050373	Description: QP ULP modes 0 and 1 cannot be assigned to the same Multicast group.
	Keywords: Multicast Group (MCG), QPs
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
913451	Description: Fixed an issue in standby (WoL) modes only that caused the actual current consumption in 1.2V rail to be higher by<33mA than the advertised values although the total IC consumption is as advertised.
	Keywords: Standby (WoL) modes, current consumption
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
852744	**Description:** Mapping an SL to VL 15 is currently not supported. Trying to do so, will cause a health buffer fatal internal error report.
	Keywords: SL to VL mapping
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
937318	Description: Setting more than 8K QPs with hca_sq_owner == 1 connected to an RNDV XRQ is currently not supported.
	Keywords: QPs, RNDV XRQ
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010

Table 23 - Bug Fixes History (Sheet 9 of 9)

Internal Ref.	Issue
860574/ 860716	Description: Fixed performance issues to improve Packet Pacing performance.
	Keywords: Performance, Packet Pacing
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
1019003/ 1019039/ 995878	Description: Fixed an issue causing physical errors observed on Ixia 100GbE receiver.
	Keywords: Ixia 100GbE receiver
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010
954826	Description: Fixed an issue that caused the ipoib_enhanced_offloads indication (in the HCA capabilities) to report 0 while SRIOV_EN=1.
	Keywords: ipoib_enhanced_offloads indication
	Discovered in Release: 16.18.1000
	Fixed in Release: 16.20.1010

5 Firmware Changes and New Feature History

Table 24 - Firmware Changes and New Feature History (Sheet 1 of 13)

Feature/Change	Description
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.
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. • 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.

Table 24 - Firmware Changes and New Feature History (Sheet 2 of 13)

Feature/Change	Description
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 25
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.
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 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 25
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.

Table 24 - Firmware Changes and New Feature History (Sheet 3 of 13)

Feature/Change	Description
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): • `nic_receive_steering_discard`: 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. • `receive_discard_vport_down`: Number of packets that were steered to a VPort, and discarded because the VPort was not in a state to receive packets. • `transmit_discard_vport_down`: 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 mkey_context. If this feature is enabled, the software per mkey configuration is ignored.
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 *Flow Table*
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.

Table 24 - Firmware Changes and New Feature History (Sheet 4 of 13)

Feature/Change	Description
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- 5adapter 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:
	• 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:
	  • 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 17

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.
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 25
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: • current_q_under_processor_handle • total_q_under_processor_handle • qp_priority_update_flow

Table 24 - Firmware Changes and New Feature History (Sheet 5 of 13)

Feature/Change	Description
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 25
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 ia 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 *"Tag Matching and Rendezvous Offload"*
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.
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 *Flow Table*
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.

Table 24 - Firmware Changes and New Feature History (Sheet 6 of 13)

Feature/Change	Description
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: • 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: • 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 25
Rev. 16.20.1010	
NVME-oF Target Offload	**[Beta]** Performance and stability improvements.
Tag Matching	**[Beta]** Performance and stability improvements. For further information, refer to the PRM section "Tag Matching and Rendezvous Offload"
Tag Matching DC	**[Alpha]** Added support for Tag Matching DC (RNDV/EAGER traffic)
Adaptive Routing (Out-Of-Order)	Added support for send/receive Out-of-Order RDMA packets required by the Adaptive Routing. Adaptive Routing (AR) enables the switch to select the output port based on the port's load.
DSCP	Added trust level for QoS prioritization according to the DSCP or PCP.
	[Beta] Added ingress buffer management for: • ingress traffic mapping to a buffer according to priority • buffers sizes and lossless parameters

Table 24 - Firmware Changes and New Feature History (Sheet 7 of 13)

Feature/Change	Description
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 31, "Supported Non-Volatile Configurations," on page 57)
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 25
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 25
Rev. 16.18.2000 (Beta)	
Virtual Extensible LAN (VXLAN) encapsulation/decapsulation	**[Beta Level]** Virtual Extensible LAN (VXLAN) is a network virtualization technology that improves scalability problems associated with large cloud computing deployments. It tunnels Ethernet frames within Ethernet + IP + UDP frames. Mellanox implements VXLAN encapsulation and decapsulation in the hardware.
NVME-oF Target Offload	**[Beta Level]** Added support for NVMe over Fabrics (NVME-oF) offload, an implementation of the new NVME-oF standard target (server) side in hardware.
Tag Matching	**[Beta Level]** Added support for offloading MPI tag matching to HCA. It matches the source send operations to the destination receivers.
VLAN Switch Tagging (VST)	Enables the virtual machine interface to have no VLAN tag over it, thus allowing VLAN tagging to be handled by the Hypervisor.
On Demand Paging (ODP)	Added supported for Demand Paging (ODP).
Rev. 16.18.1000 (Beta)	
NVM Express over Fabrics (NVMf)	NVMf is a protocol for communicating block storage IO requests over RDMA. For further information, please refer to the PRM section *"NVMe over Fabric Target Application Offload (NVMf)"*.

Table 24 - Firmware Changes and New Feature History (Sheet 8 of 13)

Feature/Change	Description
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 *"Tag Matching and Rendezvous Offload"*
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: • 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 24 - Firmware Changes and New Feature History (Sheet 9 of 13)

Feature/Change	Description
GENEVE & IP-in-IP Stateless Offload	Added support for IP-in-IP and GENEVE network protocols encapsulated into IP frame (L2 tunneling). Encapsulation is suggested as a means to alter the normal IP routing for datagrams, by delivering them to an intermediate destination that would otherwise not be selected based on the (network part of the) IP Destination Address field in the original IP header. **Note:** For driver support, please see the Release Notes/User Manual of the relevant OS driver.
Resilient RoCE	Resilient RoCE is the ability to send RoCE traffic over a lossy network (a network without flow control enabled), without the need to enable flow control on the network. The ability is accomplished by enabling ECN on both the Switch and the Host.
Power MGMT	Added support for PCIe Express standard "Slot capability register" message (PCIe base rev 3.1, section 6.9 – "Slot Power Limit Control") 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. 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). Custom and OEM branded card based on ConnectX-5 Ex may be configured by INI to support/not-support the Power management feature. 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: • `echo "MLNX_RAW_TLV_FILE" > /tmp/power_conf_tlv.cfg; echo "0x00000004 0x00000088 0x00000000 0xc0000000" >> /tmp/power_conf_tlv.cfg` • `mlxconfig -d <device> -f /tmp/power_conf_tlv.cfg set_raw` • `mlxfwreset -d <device> reset` For details on ConnectX-5 Ex power, please refer to ConnectX-5 Ex Datasheet

Virtual Functions (VF) per Port	The maximum Virtual Functions (VF) per port is 64.
	Note: When increasing the number of VFs, the following limitations must be taken into consideration:
	`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)`
	Note: For the maximum number of VFs supported by your driver, please refer to your drivers' Release Notes or User Manual.

QoS per VFs	**[InfiniBand Only]** Added support for multiple VLs in SR-IOV/mutlihost environments. **Note:** The number of VLs can be configured by the NVCONFIG. The default VL number is 4 VLs.
HCA Port Flap Counter	Added support for Port Flap Counter.
Fixed Buffer Size (KSM)	Limits the buffer size for all entries to improve performance. KSM is used when associating Key Length My Virtual Address (KLMs) with fixed memory size.

Table 24 - Firmware Changes and New Feature History (Sheet 10 of 13)

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: • 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 tear-down or FLR)
No Driver NIC (NODNIC) Performance Improvement	Added support for: • 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: • 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.

Table 24 - Firmware Changes and New Feature History (Sheet 11 of 13)

Feature/Change	Description
FCS no scatter / FCS check	Enables the user to control whether or not to scatter Frame Check Sequence (FCS) or to check FCS functionality.
PRBS Patterns Generation and Tuning	A new PHY test mode in which the device can generate different PRBS patterns for SerDes tuning purpose. For further information, please refer to PRM registers: PPAOS, PPTT, PPRT.
Management Controller Transport Protocol (MCTP) over PCI	Added support for MCTP host management over PCI
OCBB / OCSD support after mlxfwreset	Added support for OCBB/OCSD memory pointers restoration after mlxfwreset
MCTP media migration	Added support for MCTP media migration between SMBUS and PCI
Cables	Removed the RX amplitude configuration on some cable types
IPoIB checksum and LSO offload	Added IPoIB checksum and LSO offload support
Scatter FCS in RQ	Enables software to scatter or strip FCS in RQ.
CQE Time Stamping	Keeps track of the creation of a packet. A time-stamping service supports assertions of proof that a datum existed before a particular time.
Priority Flow Control (PFC)	Applies pause functionality to specific classes of traffic on the Ethernet link.
RDMA retransmission counters	Custom port counters provide the user a clear indication about RDMA send/ receive statistics and errors.
Link Layer Discovery Protocol (LLDP)	The Link Layer Discovery Protocol (LLDP) is a vendor-neutral Link Layer protocol in the Internet Protocol Suite used by network devices for advertising their identity, capabilities, and neighbors on a IEEE 802 LAN. The protocol is formally defined in IEEE 802.1AB.
Flow Steering Counters	Provides a clear indication of Flow Steering statistics and errors.
WQE Inline Header	The minimal amount of packet headers inlined in the WQE's Eth Segment.
table-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: 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 24 - Firmware Changes and New Feature History (Sheet 12 of 13)

Feature/Change	Description
Ethernet Network	• 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	• PCIe Function Level Reset (FLR) • Power Management L2/L3 flow support

PRM	• Strided SRQ
	• Self Loopback support
	• Transport Domain support
	• CQ2EQ remapping
	• Added support for the following commands:
	  • `MODIFY/QUERY_ESW_VPORT_CONTEXT`
	  • `QUERY/MODIFY_CONG_STATUS`
	  • `QUERY/MODIFY_CONG_PARAMS`
	  • `QUERY_CONG_STATISTICS`
	  • `ADD/DELETE_VXLAN_UDP_DPORT`

Virtualization	• VXLAN/NVGRE Stateless offload In this release, NVGRE is supported through Windows ONLY • SR-IOV EN
Performance	• CQE zipping
Misc	• Wake-on-Lane/Standby • FlexBoot/UEFI support
Non-Volatile Configuration	• Non-Volatile Configuration (NVConfig). For the complete list, lease refer to Section 9, on page 57.
Port management	• Enabled port management. Now one port can be set as Ethernet and one as InfiniBand.

InfiniBand Network	• Dynamically Connected (DC) transport
	Note: There is no interoperability between ConnectX-5 and ConnectX-4 adapter cards when using DC.
	• Unreliable Datagram Connection transport
	• Atomic Operation
	• CORE-Direct®
	  • Provides Collective Off-loading in HCA
	  • Frees CPU to perform computation in parallel with collective operations
	• T10 DIF pipeline Data Integrity Signature off-loading (at beta level)
	• User Memory Registration (UMR)
	• Automatic Path Migration
	• On Demand Paging (ODP) - Memory can now be used without pinning memory beforehand.
	• Congestion Control
	• Shrink Address Vectors for RC and UD
	• Programmable Port/Node GUID

Table 24 - Firmware Changes and New Feature History (Sheet 13 of 13)

Feature/Change	Description
General	• 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	• NC-SI over RMII support
MAD	• Config space address in MAD management class 0x09

6 FlexBoot Changes and New Features

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

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

Version	Description
Rev. 3.5.603	
Enable/Disable UEFI X86, UEFI ARM Exprom using Ctrl+B	Added support to Enable/Disable UEFI X86, UEFI Arm Exprom in FlexBoot using the Ctrl+B menu.
Rev. 3.5.504	
FlexBoot UI	Added *"PXE boot without fail-over to iSCSI"* and *"iSCSI boot without fail-over to PXE boot"* options for legacy_boot_prtocol configuration. For further information, refer to Mellanox PreBoot Drivers User Manual.
	Enabled PXE to expose the current link speed of the system when in ETH mode.
Rev. 3.5.403	
Enable/Disable FlexBoot in EXPROM via mlxconfig	Added PXE support to additional ConnectX- 5 adapter cards. Enabling/Disabling FlexBoot in ConnectX- 5 in EXPROM is done via mlxconfig. The default value is: • FLEXBOOT enable
VLAN Priority	Set the default VLAN priority to 0.
Link Aggregation Control Protocol (LACP)	LACP support is disabled by default. It can be enabled via mlxconfig.
Rev. 3.5.305	
PXE Boot	Added ESC option as an abort key during PXE boot process.
FlexBoot Link Aggregation Control Protocol (LACP)	Enabled/disabled FlexBoot LACP support by editing the INI configuration.
Upstream sync	Synced the source with iPXE (upstream sync)
Rev. 3.5.210	
Promiscuous VLAN mode	Added support for promiscuous VLAN mode.
MTU	**[InfiniBand]** Added support for configurable MTU.
Expansion ROM version	Enabled expansion ROM (`exp_rom`) version exposition according to the new specification (e.g. expose ARCH in flint tool).
FlexBoot UI	Added a FlexBoot menu support for `NV_POWER_CONF`. Now power consumption configuration is supported from the FlexBoot menu.
	Enhanced FlexBoot/firmware debug capability using Flexboot UI. Added the `reg_dump` option to the `panic_behavior` configuration in the FlexBoot menu
Rev. 3.5.110	
Networking	Ethernet only: The MTU value is set to 1500 upon driver's bring up.
Rev. 3.5.109	
Performance	Performance enhancements in Ethernet mode

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

Version	Description
FlexBoot UI	Added support for "Undi network wait timeout"
	Enhanced FlexBoot/firmware debug capability using Flexboot UI
Upstream sync	Synced the source with iPXE (upstream sync)

6.1 FlexBoot Known Issues

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

Internal Ref.	Description
1422604	**Description:** The FlexBoot TFTP enters an infinite loop when it tries to resolve the server name by a DNS request, and the DNS server dose not response to the request.
	Workaround: Terminate the TFTP process by pressing Ctrl + C
	Keywords: Boot, TFTP, DNS
	Discovered in Version: 3.5.504
1295727	**Description:** In Secure Host mode, the Ctrl + B option will be “read only” and changes will not be applied and may cause unknown behavior.
	WA: N/A
	Keywords: Secure Host mode, Ctrl + B option
	Discovered in Version: 3.5.403
1066544	**Description:** Chain-loading boot-loaders that works with interrupts fails to boot on multi-host adapter cards
	Workaround: N/A
	Keywords: Chain-loading, boot
	Discovered in Version: 3.5.403
1157875	**Description:** Pressing any of the arrow keys during boot might cause the boot process to be aborted.
	Workaround: N/A
	Keywords: Abort boot, arrows, FlexBoot
	Discovered in Release: 3.5.305
1149467	**Description:** Chain-loading "ipxe.pxe" and "undionly.kpxe" over InfiniBand is currently not supported when using DHCP client identification based on InfiniBand 32-bit Prefix+GUID (as with FlexBoot).
	Workaround: N/A
	Keywords: FlexBoot, chainload, InfiniBand, undionly.kpxe, ipxe.pxe
	Discovered in Version: 3.5.305
841198	**Description:** FlexBoot fails to boot when the following occurs: • Boot priority is set to iSCSI • The iSCSI TCP/IP parameters via DHCP is disabled • iSCSI boot fails or iSCSI boot to target configuration is set to disable
	Workaround: N/A
	Keywords: PXE boot, iSCSI

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

Internal Ref.	Description
843377/849223	**Description:** The physical MAC assigned via the boot menu is displayed as zeroes instead of the set MAC when ConnectX-4 VPI adapter card is configured as InfiniBand.
	Workaround: N/A
	Keywords: Physical MAC, Boot menu
656001	**Description:** Booting from WDS and Windows DHCP server when only Option 66 is enabled (without Option 67), is not supported.
	Workaround: N/A
	Keywords: DHCP
776057	**Description:** Citrix PVS boot is not supported.
	Workaround: N/A
	Keywords: Citrix PVS boot
689460	**Description:** FlexBoot uses system UUID to generate the client DUID-UUID as per RFC 6355, the data conveyed with DHCPv6 Code 1 (Option ID).
	Workaround: N/A
	Keywords: DUID-UUID
928217	**Description:** Installing ESXi 6.5/6.0 on iSCSI target is currently not supported.
	Workaround: N/A
	Keywords: ESXi 6.5/6.0, iSCSI target
-	**Description:** To use the DHCP server to identify ipxe requests when using `undionly.kpxe` or `ipxe.pxe` when booting over IB requires special configuration. (see the Workaround below).

	Workaround: Add to the DHCP host declaration the MAC identification alongside the option 61 DUID.
	For example:
	`host ib-client1 {`
	`option dhcp-client-identifier =`
	`ff:00:00:00:00:00:02:00:00:02:c9:00:<Port-GUID> ;`
	`hardware ethernet <Port-MAC> ;`
	`fixed-address <IPoIB Address> ;`
	`filename "ipxe.pxe" ;`
	`if exists user-class and option user-class = "iPXE" { filename “pxelinux.0” ; }`
	`}`

	Keywords: undionly.kpxe or ipxe.pxe

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

Internal Ref.	Description
-	**Description:** Due to interoperability issue between the ESXi installer and the lpxelinux bootloader, when trying to install ESXi 6.5 on iSCSI target using lpxelinux.0 as a bootloader, a PSOD occurs.

	Workaround: Use FlexBoot (or iPXE) to load mboot.c32 directly instead of pxelinux.0 using the script below:
	`#!ipxe`
	`set base /nfs/Esxi-6.5_INBOX`
	`chain ${base}/mboot.c32 -c ${base}/boot.cfg BOOTIF=01-${mac:hexhyp}`
	
	where the "`set base ...`" specifies a suitable absolute path.
	Note: iPXE does not need an absolute path, however, mboot.c32 requires it.

	Keywords: mboot.c32, PSOD,
976878	**Description:** When using bootloader grub2 to boot WDS, if the WDS boot fails, an RSOD might appear.
	Workaround: N/A
	Keywords: Bootloader grub2, WDS, RSOD
-	**Description:** Several BIOS vendors have limited boot-vector space and may not display FlexBoot in their boot menu.
	Workaround: Disable the embedded NIC boot agent in BIOS
	Keywords: BIOS
-	**Description:** In several BIOS, the server might hang during FlexBoot booting due to wrong configuration of the PMM.
	Workaround: N/A
	Keywords: BIOS
-	**Description:** Only EBX, ESI, DS, ES registers can be saved in Boot Entry.
	Workaround: N/A
	Keywords: BIOS
-	**Description:** If a client returned control to the BIOS after a successful connection to an iSCSI target (but did not boot from it), then, unexpected behavior may occur.
	Workaround: Follow the instructions described in the FlexBoot UM for the proper iSCSI boot/install
	Keywords: BIOS
673114/821899	**Description:** FlexBoot banner might not be shown in some BIOSes.
	Workaround: N/A
	Keywords: BIOS

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

Internal Ref.	Description
-	**Description:** In some cases, PXE boot will not work if the client was given only the filename without next-server (siaddr).
	Workaround: N/A
	Keywords: PXE Boot
-	**Description:** PXE boot after iSCSI boot with static configuration is currently not supported.
	Workaround: N/A
	Keywords: PXE Boot
-	**Description:** Boot over VLAN with IB port is currently not supported.
	Workaround: N/A
	Keywords: PXE Boot
-	**Description:** Some faulty boot loaders do not close the underlying UNDI device which may result in unexpected behavior and possible system crash after the OS starts to load.
	Workaround: N/A
	Keywords: PXE Boot
-	**Description:** Chain-loading gPXE stack is not supported.
	Workaround: N/A
	Keywords: PXE Boot
647143	**Description:** Executing a partial boot loop while only downloading the NBP and selecting local-boot is unsupported and may cause undefined behavior.
	Workaround: N/A
	Keywords: PXE Boot
-	**Description:** iSCSI over IB is not tested.
	Workaround: N/A
	Keywords: iSCSI
-	**Description:** iSCSI over DCB is not supported.
	Workaround: N/A
	Keywords: iSCSI
-	**Description:** FlexBoot supports only a single active iSCSI connection. Thus, when iSCSI-boot via Port 1 succeeds to connect but fails to boot, it will fail to connect via Port 2.
	Workaround: N/A
	Keywords: iSCSI

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

Internal Ref.	Description
-	**Description:** Boot retries is currently not functional when booting from iSCSI.
	Workaround: N/A
	Keywords: iSCSI
655800	**Description:** iSCSI over IPv6 is not supported.
	Workaround: N/A
	Keywords: iSCSI
-	**Description:** Boot menu is displayed as READ ONLY if the HCA card does not support flash configuration.
	Workaround: N/A
	Keywords: User Interface
-	**Description:** FlexBoot Boot Menu will not be visible in serial output.
	Workaround: N/A
	Keywords: User Interface
-	**Description:** Large Receive Offload (LRO) and iSCSI may not interoperate due to a bug in current Linux kernel distributions.
	Workaround: Disable LRO in the IPoIB module when using iSCSI. See the Mellanox FlexBoot user's manual for details under the Diskless Machines chapter (InfiniBand Ports).
	Keywords: Networking
-	**Description:** Setting the number of Virtual Functions higher than the machine's memory capability may cause memory issues and system instability.
	Workaround: N/A
	Keywords: Virtualization
-	**Description:** SLAM, FTP, HTTPS and SRP are currently not supported.
	Workaround: N/A
	Keywords: Protocols
-	**Description:** Occasionally, using the Spanning Tree Protocol (STP) in the switches may cause packet drops and boot failure in the system.
	Workaround: Enable the "edgemode" if disabled on the switch, or use either portfast or edgemode functionality on the switch ports connected to the NICs.
	Keywords: Protocols
655800	**Description:** IPv6 can only run if a RADVD service is running in the network.
	Workaround: N/A
	Keywords: Protocols

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

Internal Ref.	Description
-	**Description:** IPv6 over IB is not supported.
	Workaround: N/A
	Keywords: Protocols
655800	**Description:** Enabling IPv6 first and then IPv4 is currently not supported.
	Workaround: N/A
	Keywords: Protocols

6.2 FlexBoot Bug Fixes History

Table 27 - FlexBoot Bug Fixes History

Version	Issue
1072419	**Description:** The FlexBoot DHCP loops indefinitely when it continuously gets NACK on the DHCP requests On some setups, it might also cause an RSOD after a a continues looping.
	Keywords: Bootloader grub2, WDS, RSOD
	Discovered in Release: 3.5.305
	Fixed in Release: 3.5.504
1157875	**Description:** Pressing any of the arrow keys during boot might cause the boot process to be aborted.
	Keywords: Abort boot, arrows, FlexBoot
	Discovered in Release: 3.5.305
	Fixed in Release: 3.5.403
1113560	**Description:** Fixed an issue that prevented the first iSCSI target parameters to be reset to their default values.
	Keywords: iSCSI target
	Discovered in Release: 3.5.110
	Fixed in Release: 3.5.305

7 UEFI Changes and Major New Features

Table 28 - UEFI Changes and New Features

Category	Description
Rev. 14.17.13	
Secure Boot	Added support for secure boot on Arm servers for Mellanox adapter cards.
Rev. 14.16.15	
Hii Menu	Added *"PXE boot without fail-over to iSCSI"* and *"iSCSI boot without fail-over to PXE boot"* options for legacy_boot_prtocol configuration. For further information, refer to Mellanox PreBoot Drivers User Manual.
	Enabled UEFI to expose the current link speed of the system in the Hii menu.
Rev. 14.15.19	
Enable/Disable UEFI in EXPROM via mlxconfig	Added UEFI support to additional ConnectX- 5 adapter cards. • ConnectX-5 adapter cards are compiled with x86-UEFI Enabling/Disabling UEFI in ConnectX- 5 in EXPROM is done via mlxconfig. The default values are: • UEFI_X86 disabled • UEFI_AARCH64 disabled

7.1 UEFI Known Issues

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

Table 29 - UEFI Known Issues

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

8 Unsupported Features and Commands

8.1 Unsupported Features

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

Table 30 - Unsupported Features

Feature	CX-5
The following service types: • 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

8.2 Unsupported Commands

- QUERY_MAD_DEMUX
- SET_MAD_DEMUX
- CREATE_RQ - MEMORY_RQ_RMP
- MODIFY_LAG_ASYNC_EVENT

9 Supported Non-Volatile Configurations

Table 31 - 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_DESCRIPTORS		
	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_DESCRIPTORS		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 31 - 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		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		0x10a
	LLDP_NB_RX_MODE		
	LLDP_NB_TX_MODE		
NV_LLDP_NB_DCBX	DCBX_IEEE		0x18e
	DCBX_CEE		
	DCBX_WILLING		
NV_KEEP_LINK_UP	KEEP_ETH_LINK_UP		0x190
	KEEP_IB_LINK_UP		
	KEEP_LINK_UP_ON_BOOT		
	KEEP_LINK_UP_ON_STANDBY		
NV_QOS_CONF	NUM_OF_VL		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 31 - 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_CNT1		
	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		