



Mellanox ConnectX® Firmware (fw-ConnectX) Release Notes

Rev 2.9.1000

Last Updated: October 10, 2013

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

Mellanox Technologies, Ltd.
 Beit Mellanox
 PO Box 586 Yokneam 20692
 Israel
www.mellanox.com
 Tel: +972 (0)74 723 7200
 Fax: +972 (0)4 959 3245

© Copyright 2013. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, CORE-Direct®, InfiniBridge®, InfiniHost®, InfiniScale®, MLNX-OS®, PhyX®, SwitchX®, UFM®, Virtual Protocol Interconnect® and Voltaire® are registered trademarks of Mellanox Technologies, Ltd.

Connect-IB™, ExtendX™, FabricIT™, Mellanox Open Ethernet™, Mellanox Virtual Modular Switch™, MetroX™, MetroDX™, ScalableHPC™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Table of Contents

Table of Contents	3
Chapter 1 Overview	4
1.1 Revision Compatibility	5
Chapter 2 Changes and Major New Features	6
2.1 Changes in Rev 2.9.1000 From Rev 2.8.0600	6
2.2 Changes in Rev 2.8.0600 From Rev 2.8.0000	6
2.3 Changes in Rev 2.8.0000 From Rev 2.7.9112	6
2.4 Changes in Rev 2.7.9112 From Rev 2.7.8100	6
2.5 Changes in Rev 2.7.8100 From Rev 2.7.200	6
2.6 Changes in Rev 2.7.700 From Rev 2.7.200	6
2.7 Changes in Rev 2.7.200 From Rev 2.7.000	6
Chapter 3 Known Issues	7
Chapter 4 Bug Fixes	9

1 Overview

These are the release notes for the ConnectX® and ConnectX® EN adapters firmware, fw-ConnectX Rev 2.9.1000. This firmware supports the following protocols:

- InfiniBand
- Ethernet
- Fibre Channel over Ethernet (FCoE) – FCoE is at a beta level
- Virtual Protocol Interconnect (VPI) – this capability enables ConnectX devices to support the InfiniBand, Ethernet and DCE network standards, including auto-sensing of the network protocol to which each device port is connected. This feature is not available with ConnectX EN.

This firmware supports the devices and protocols listed in Table 1. For the most updated list of adapter cards supported, visit the firmware download pages via <http://www.mellanox.com>.

Notes

1. Firmware Rev 2.9.1000 is compatible with MFT tools v2.6.0 and above. Using older versions of MFT tools will cause failure during image creation.
2. After burning new firmware to an adapter card, reboot the machine so that the new firmware can take effect. If you do not reboot, you will get an error in the RUN_FW command.
3. The following Mellanox adapters are not supported in MLNX_OFED 1.5.2-2.1.0:
 - MHGH29B-XSR - Dual 4X IB DDR Port, PCIe Gen2 x8 2.0 (5GT/s), Short Bracket, RoHS-R6 HCA Card, CX4 Connectors. (PSID: MT_0D10110008)
 - MHGH29B-XTR - Dual 4X IB DDR Port, PCIe Gen2 x8 2.0 (5GT/s), Tall Bracket, RoHS-R6 HCA Card, CX4 Connectors. (PSID: MT_0D20110008)

Table 1 - PCI Device ID

PCI Device ID (Decimal)	Device Part Number	Device Name	Supported Protocols
25408	MT25408A0-FCC-SI	ConnectX, Dual Port 10Gb/s InfiniBand / 10GigE Adapter IC with PCIe 2.0 x8 2.5GT/s Interface	InfiniBand, Ethernet, FCoE, VPI
25418	MT25408A0-FCC-DI	ConnectX, Dual Port 20Gb/s InfiniBand / 10GigE Adapter IC with PCIe 2.0 x8 2.5GT/s Interface	
26418	MT25408A0-FCC-GI	ConnectX, Dual Port 20Gb/s InfiniBand / 10GigE Adapter IC with PCIe 2.0 x8 5.0GT/s Interface	
26428	MT25408A0-FCC-QI	ConnectX, Dual Port 40Gb/s InfiniBand / 10GigE Adapter IC with PCIe 2.0 x8 5.0GT/s Interface	

Table 1 - PCI Device ID

PCI Device ID (Decimal)	Device Part Number	Device Name	Supported Protocols
25448	MT25448A0-FCC-SE	ConnectX EN, Dual Port 10GigE Adapter IC with PCIe 2.0 x8 2.5GT/s Interface	Ethernet
26448	MT26448A0-FCC-TE	ConnectX EN, Dual Port 10GigE Adapter IC with PCIe 2.0 x8 2.5GT/s Interface	
25458	MT25458A0-FCC-SE	ConnectX EN, Dual Port 10BASE-T Adapter IC with PCIe 2.0 x8 2.5GT/s Interface	
26458	MT25458A0-FCC-TE	ConnectX EN, Dual Port 10BASE-T Adapter IC with PCIe 2.0 x8 5.0GT/s Interface	

1.1 Revision Compatibility

Firmware fw-ConnectX Rev 2.9.1000 complies with the following programmer's reference manual:

- *ConnectX Programmer's Reference Manual (PRM), Rev 0.39 or later*, which has Command Interface Revision 0x3. 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 Major New Features

2.1 Changes in Rev 2.9.1000 From Rev 2.8.0600

- MPI performance enhancement: #97601 enigma bit

2.2 Changes in Rev 2.8.0600 From Rev 2.8.0000

- Bug Fixes - see Section 3, “Known Issues,” on page 7

2.3 Changes in Rev 2.8.0000 From Rev 2.7.9112

Added support for the following features:

- Signal integrity improvements in IBTA 1.2.1 mode (Attenuation Based Algorithm)
- Eye opener MAD
- Cable info MAD
- CQ.OI to the MODIFY_CQ command to enable CORE-Direct
- SET_NODE command
- PCIe device Serial Number capability
- LoosyVI MAD

2.4 Changes in Rev 2.7.9112 From Rev 2.7.8100

- Support for link speed/width changing via SET_PORT

2.5 Changes in Rev 2.7.8100 From Rev 2.7.200

- Support HCA class code configuration
- Support Link and Width change using the SET_PORT command
- Local invalidation enabled on Fibre Channel QPs

2.6 Changes in Rev 2.7.700 From Rev 2.7.200

- Added RDMA over Converged Ethernet (RoCE) support

2.7 Changes in Rev 2.7.200 From Rev 2.7.000

- IBTA spec 1.2.1 speed autonegotiation compliance for all standard Mellanox adapter cards

3 Known Issues

The following table describes known issues in this firmware release and possible workarounds.

Table 2 - Known Issues

Index	Issue	Description	Workaround	Scheduled Release (fix)
1.	RoCE performance in Lossy fabric	Massive packet drops in RoCE may cause low performance in lossy fabrics		N/A
2.	UAR Bar is too small for 64k-page machines	The small BAR causes driver loading to fail	Change the "log2_uar_bar_megabytes" .ini parameter under the [HCA] section as follows: log2_uar_bar_megabytes = 5	N/A
3.	GUID of ConnectX-2 Ethernet adapter cards	On ConnectX-2 Ethernet adapter cards, there is a mismatch between the GUID value returned by firmware management tools and that returned by fabric/driver utilities that read the GUID via device firmware (e.g., using ibstat). Mlxburn/flint return 0xffff as GUID while the utilities return a value derived from the MAC address. For all driver/firmware/software purposes, the latter value should be used.	N/A. Please use the GUID value returned by the fabric/driver utilities (not 0xffff).	
4.	Change of memory BARs on a disabled system	Changing memory bars size / addresses between SYS_DIS and SYS_EN may cause the device to hang (ID: 24206)	N/A	N/A
5.	BAR resizing on an enabled system	Changing BAR sizes when a system is enabled may cause the device to hang (ID: 24208)	N/A	N/A
6.	Ethernet only: Must query all capabilities upon boot	If not all capabilities are queried upon boot, then the query command may fail. See the QUERY_CAP command in <i>ConnectX Programmer's Reference Manual</i>	Query all capabilities upon boot	N/A

Table 2 - Known Issues (Continued)

Index	Issue	Description	Workaround	Scheduled Release (fix)
7.	Disrupting QDR negotiation may lead to port rising as SDR	Disconnecting an IB cable (or closing the port) during QDR negotiation and then reconnecting (or reopening) may cause the adapter to bring up the port at SDR	Disconnect the cable (or close the port) again and then reconnect (reopen). To avoid this scenario, wait for QDR negotiation to finish prior to disconnecting the cable (or closing the port) and reconnecting (or reopening). The following are two possible methods to verify QDR negotiation is complete: a. The physical (green) LED is on. b. A query of LinkPhyState using a GetPortInfo MAD indicates LinkUp.	N/A
8.	MTNIC driver is not supported			
9.	InfiniBand Static rate is not supported			
10.	Wrong reporting of transceiver type			
11.	Autosensing after cable reconnection to different port protocol	If you disconnect a cable from an IB/Eth port and reconnect it to an Eth/IB port (different protocol), the link may not rise	Restart the driver (openibd restart)	N/A
12.	Some boot applications may fail to boot on VPI cards due to the infiniband PCIe.class_ID issue		Change class_ID by applying the following setting to the .ini file: Under the [HCA] section, enter: hca_header_class_code = 0x028000 eth_class_en_ib = true	N/A

4 Bug Fixes

Table 3 - Firmware v2.9.1000 Bug Fixes

	Issue	Description	Discovered in	Fixed in
1.	Wrong UAR mapping in pages larger than 4K: • Blueflame did not work on UAR pages larger than 4K	Fixed	2.8.0600	2.9.1000
2.	Congestion Notification Packet may cause HCA to stall under extreme conditions	Fixed	2.8.0600	2.9.1000
3.	When in unsolicited mode a completion message is not sent on every completion	Fixed	2.8.0600	2.9.1000
4.	PCIe compatibility issue with a specific PCIe root complex	Fixed	2.8.0000	2.8.0600
5.	HCA may hang upon CLOSE_PORT command during data stress (ID 91657)	Fixed	2.7.000	2.8.0000
6.	Ethernet: Potential packet drops in Pause mode	Fixed	2.7.000	2.8.0000
7.	QUERY_DEV_CAP.log_max_bf_regs_per_page fixed from 8 to 3	Fixed	2.7.000	2.8.0000
8.	PCIe slow handling of configuration cycles may cause NMI	Fixed	2.7.000	2.7.9112
9.	Potential delay in responding to configuration cycles upon data stress	Fixed	2.7.000	2.7.8100
10.	High Bit Error Rate on active cables	Fixed	2.7.000	2.7.8100
11.	PCI Express compliancy: The ConnectX device may transmit before the PCI Express receiver detect is enabled	Fixed	2.7.000	2.7.8100
12.	Slow VPD performance might cause NMI	Fixed	2.7.000	2.7.8100
13.	Possible wrong link state during link bring up	Fixed	2.7.000	2.7.700
14.	High Bit Error Rate on active cables	Fixed	2.7.000	2.7.700
15.	Potential delay in responding to configuration cycles upon data stress	Fixed	2.7.000	2.7.700
16.	RoCE issue: Wrong GID in VPI	Fixed	2.7.700	2.8.0000
17.	RoCE issue: Missing GUIDs in Ethernet only cards	Fixed	2.7.700	2.8.0000
18.	InfiniBand auto-negotiation issues	Fixed	2.6.900	2.7.000
19.	MOD_STAT_CFG.port_en command does not apply to Ethernet	Fixed	2.6.000	2.6.900
20.	Mellanox auto-negotiation is not optimized for environments with significant crosstalk	Fixed	2.6.000	2.6.900
21.	CQ/EQ doorbell is lost in Configuration cycles	Fixed	2.6.000	2.6.900
22.	Congestion issues	Fixed	2.6.000	2.6.900

Table 3 - Firmware v2.9.1000 Bug Fixes

23.	SW2HW_MPT command may fail when MPT is in FREE state (used for FRWR)	Fixed	2.6.000	2.6.900
24.	Adapter may get stuck upon multicast traffic stress	Fixed	2.6.000	2.6.900
25.	IB Spec Release 1.2 link data-rate issues	Fixed	2.6.000	2.6.900
26.	Adapter may get stuck if QPC extended auxiliary table base address is not aligned to 1MB	Fixed	2.6.000	2.6.900
27.	Port mod_stat_cfg.disable does not take effect if the device is configured as an Ethernet device	Fixed	2.6.000	2.6.300
28.	HCA might hang when entering PCIe L1 mode	Fixed	2.6.000	2.6.200
29.	Link may come up as SDR instead of DDR on some systems	Fixed	2.5.000	2.6.000
30.	Possible live lock in QP upon retransmission stress	Fixed – (ID: 49870, 52066)	2.5.000	2.6.000
31.	Wrong link state reported during link speed negotiation	Fixed – (ID: 49951)	2.5.000	2.6.000
32.	CQs may be generated after CQ overrun error	Fixed – (ID: 49982)	2.5.000	2.6.000
33.	Slow handling of configuration cycles	Fixed – (ID: 49807)	2.5.000	2.6.000
34.	Wrong fields in CQE-w-Error on XRC QP	Fixed – (ID: 49742)	2.5.000	2.6.000
35.	Wrong handling of Remote Invalidate Error	Fixed	2.5.000	2.6.000
36.	Multiple RNR Nack may cause slowdown	Fixed – (ID: 49559)	2.5.000	2.6.000
37.	QUERY_DEV_CAP.apm bit was fixed at 0 even though APM was active	Fixed – (ID: 49548)	2.5.000	2.6.000
38.	PCIe physical errors upon entering L1 state	Fixed – (ID: 52025)	2.5.000	2.6.000
39.	PCI_CFG.interrupt_disable has no impact	Fixed – (ID: 53350)	2.5.000	2.6.000
40.	Non-default setting of VLCap via .ini does not take effect	Fixed	2.5.000	2.6.000
41.	SET_PORT may lead to non-optimal RX buffer reallocation if opvl was less than vlcap	Fixed	2.5.000	2.6.000
42.	Modified PLL parameter settings in .ini	Some PLL parameter settings were changed to allow a longer period for PLL stabilization	2.5.000	2.6.000
43.	Adapter may generate PCIe transactions with wrong function ID	Fixed	2.5.000	2.6.000
44.	Adapter may generate PCIe ERR_NON_FATAL in case of an unsupported request	Fixed	2.5.000	2.6.000
45.	<u>Ethernet only</u> : Different VLAN priorities for WQE and QP may cause wrong SchedQ allocation	Fixed	2.5.000	2.6.000
46.	Wrong VLAN Priority in PPP mode	The ConnectX device may send pause frames for the wrong priority in PPP mode (ID: 49533)	2.5.000	2.5.900

Table 3 - Firmware v2.9.1000 Bug Fixes

47.	Wrong aliasing in address steering mode	This issues applies only to the VMware® Infra-structure 3 v3.5 operating system in netq mode	2.5.000	2.5.900
48.	Possible packet dropping though the pause policy is set	Fixed	2.5.000	2.5.900
49.	Wrong PCI Class Code for Ethernet Network Controller	This issue applies to Ethernet devices only. Fixed.	2.5.000	2.5.900
50.	Bringing up/down of an adapter port may lead to PHY errors on the second adapter port	Applies only to dual-port adapter cards. Fixed (ID: 51356)	2.5.000	2.5.900
51.	PPP does not work for an MTU of 9600	Fixed	2.5.000	2.5.900
52.	PCI Express compliancy issues	<ul style="list-style-type: none"> • Fixed L1 and L0s power states compliancy issues • Fixed PCIE-CV test completion_timeout failure • Fixed interoperability issue with all available PCIe Gen. 2.0 servers (Ref. ID: 43852) 	2.3.000	2.5.000
53.	INTA may be lost under stress	Fixed. (Ref. ID: 44473)	2.3.000	2.5.000
54.	Modifying SRQ number using RTS2RTS	Modifying SRQ number using RTS2RTS does not guarantee that no new CQEs will be generated using the old SRQ number. Fixed.	2.3.000	2.5.000
55.	QP may get stuck upon Responder Gather Error	Fixed.	2.3.000	2.5.000
56.	Wrong handling of SL mismatch between WQE and MLX QP	An SL mismatch between WQE and MLX QP may cause the QP to get stuck. Fixed.	2.3.000	2.5.000
57.	UC QP CQE with Error causes corruption	Fixed.	2.3.000	2.5.000
58.	Query_CQ/Query_EQ commands may return the old consumer_index	Fixed.	2.3.000	2.5.000

Table 3 - Firmware v2.9.1000 Bug Fixes

59.	CQ error may cause corruption	A CQ error such as an overrun may cause CQ corruption, leading to a wrong CQ number in the CQ error event or to an internal FW error. Fixed.	2.3.000	2.5.000
60.	Possible FW internal error on a very noisy link	Fixed. (Ref. ID:41526)	2.3.000	2.5.000
61.	QueryDebugMSG command returns wrong status	Fixed. (Ref. ID: 44744)	2.3.000	2.5.000
62.	Dropping a ReadResponse packet may lead to 'retry exceeded'	Fixed.	2.3.000	2.5.000
63.	CQ moderation parameters are wrongly configured	Fixed. (Ref. ID: 45570)	2.3.000	2.5.000
64.	False generation of CQE with error (vendor code 0x6f) upon large stress	Fixed. (Ref. ID: 45317)	2.3.000	2.5.000
65.	Bandwidth degradation if SetPort command is not called	Fixed.	2.3.000	2.5.000
66.	SQERR2RTS command followed by an error causes QP to be unfunctional	Fixed. (Ref. ID: 45828 45848)	2.3.000	2.5.000
67.	QUERY_FW fails after RUN_FW	The command QUERY_FW fails after running the RUN_FW command	2.2.000	2.3.000
68.	HCA stall	The HCA might stall in any of the following scenarios: <ul style="list-style-type: none"> • If running the command SET_DEBUG_MESSAGE (ID:42128) • Under large stress (ID: 43385, 43378) • Upon closing a large number of QPs (ID: 43697) • If the WQE SL is different than the QP Context SL in a UD QP (ID: 41423) • Upon multiple retransmissions 	2.2.000	2.3.000

Table 3 - Firmware v2.9.1000 Bug Fixes

69.	QUERY_QP errors	Wrong QUERY_QP command in the following cases: <ul style="list-style-type: none"> • Returns wrong values (ID: 42078, 40707) • Enters the error state erroneously (ID: 43110) 	2.2.000	2.3.000
70.	IB & PCI Express links quality	General improvements	2.2.000	2.3.000
71.	Incomplete support for PCI Express 2.0 configuration header	Fixed	2.2.000	2.3.000
72.	Wrong trap generation rate	The HCA might exceed the maximum trap generation rate upon processing different trap types	2.2.000	2.3.000
73.	Client Reregister event not generated	The HCA might fail to generate a Client Reregister event under large stress. (ID: 42232)	2.2.000	2.3.000
74.	Possible ICM corruption	Possible ICM (Interconnect Context Memory) corruption upon large stress (ID: 42529)	2.2.000	2.3.000
75.	Performance	HCA performance improvements for the following cases: <ul style="list-style-type: none"> • Upon receiving multiple ACK packets • Upon multiple QPs in error state (ID:43377) • Upon multiple RNR NACKs 	2.2.000	2.3.000
76.	Wrong wqe_index in Receive CQE with Error	This can occur when running stress IPoIB CM tests. (ID: 43076)	2.2.000	2.3.000
77.	Possible multicast corruption	Fixed (ID: 43301)	2.2.000	2.3.000
78.	Wrong limit on number of supported EQ UARs	The HCA now supports the requested number of EQ UARs specified in INIT_HCA	2.2.000	2.3.000
79.	SchedQueue corruption	Fixed (ID: 43289)	2.2.000	2.3.000

Table 3 - Firmware v2.9.1000 Bug Fixes

80.	Wrong SL2VL mapping upon set_sl2vl	Fixed	2.2.000	2.3.000
81.	False MAD packet drops	The HCA might drop MAD packet erroneously under large stress	2.2.000	2.3.000
82.	PCI Express 2.0 x1 link fails to rise	Fixed	2.2.000	2.3.000
83.	Command timeouts	The HCA times out commands while closing multiple QPs	2.2.000	2.3.000
84.	False internal error generation	Fixed	2.2.000	2.3.000
85.	Transport timeouts	Multiple RNR NACKs may lead to transport timeouts (ID: 44160)	2.2.000	2.3.000
86.	Opcode/Input Modifier verification	Command Opcode/ Input Modifier values are now checked for correctness. If a wrong value is provided, the command status indicates the error.	2.2.000	2.3.000
87.	Wrong <i>sl</i> and/or port number returned	The QUERY_QP command may return a wrong <i>sl</i> value and/ or a wrong port number (ID: 40707)	2.1.000	2.2.000
88.	HCA stall	The HCA might stall upon stress involving RNR Nacks and RDMA reads (ID: 41918)	2.1.000	2.2.000
89.	QP corruption	QP corruption may occur following a CQ_ overrun	2.1.000	2.2.000
90.	Sched Queue corruption	Sched Queue corruption may occur upon multiple re-transmissions	2.1.000	2.2.000
91.	False SRQ WQE limit event	A false SRQ WQE limit event is generated due to a race condition	2.0.164	2.1.000
92.	Wrong Dt value returned	The QUERY_FW command may return a wrong Dt value	2.0.164	2.1.000

Table 3 - Firmware v2.9.1000 Bug Fixes

93.	HCA hangs	The device hangs in one of the following cases: <ul style="list-style-type: none"> • upon retry – due to local_ack_timeout • upon retry – due to RNR Nack • upon ringing a CQ doorbell for an invalid QP • upon stress conditions (IDs: 41543,732/6,755,778) 	2.0.164	2.1.000
94.	High ACK latency	Delays in ACK may cause multiple local ACK timeouts	2.0.164	2.1.000
95.	HCA performance	HCA performance may be impacted in the following conditions: <ul style="list-style-type: none"> • QPs in error state • Slow QP context handling 	2.0.164	2.1.000
96.	IB link stability issues		2.0.164	2.1.000
97.	High QP closing duration	Closing QPs with outstanding posted WQs may take a long time due to slow CQE with error generation	2.0.164	2.1.000