



Mellanox ConnectX[®]-2 Firmware (fw-ConnectX2) Release Notes

Rev 2.6.648

www.mellanox.com

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 NON-INFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



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

Mellanox Technologies, Ltd.
PO Box 586 Hermon Building
Yokneam 20692
Israel
Tel: +972-4-909-7200
Fax: +972-4-959-3245

© Copyright 2011. Mellanox Technologies. All rights reserved.

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

FabricIT, MLNX-OS and SwitchX 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
Chapter 2 Revision Compatibility	5
Chapter 3 Changes and Major New Features	6
3.1 Changes in Rev 2.6.648 From Rev 2.6.000-----	6
3.2 Changes in Rev 2.6.000 From Rev 2.5.000-----	6
Chapter 4 Bug Fixes	7
Chapter 5 Known Issues	8
Chapter 6 Creating a Device Configuration (.ini) File	9
6.1 Configuration (.ini) File Format -----	9
Chapter 7 History of Fixed Issues	10

1 Overview

These are the release notes for the ConnectX®-2 and ConnectX®-2 EN adapters firmware, fw-ConnectX2 Rev 2.6.648. This firmware supports the following protocols:

- InfiniBand
- Ethernet
- Fibre Channel over Ethernet (FCoE)
- Virtual Protocol Interconnect (VPI) – this capability enables ConnectX-2 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 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>.

Note: 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.

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	
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	

The document consists of the following sections:

- “Revision Compatibility” on page 5
- “Changes and Major New Features” on page 6
- “Bug Fixes” on page 7
- “Known Issues” on page 8
- “Creating a Device Configuration (.ini) File” on page 9

2 Revision Compatibility

Firmware fw-25408 Rev 2.6.648 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*.
- *ConnectX EN Programmer's Reference Manual (PRM), Rev 0.18 or later*, which has Command Interface Revision 0x2. The command interface revision can be retrieved by means of the QUERY_FW command and is indicated by the field *ifc_rev*.

3 Changes and Major New Features

3.1 Changes in Rev 2.6.648 From Rev 2.6.000

- Support for IBTA specification v1.2.1
- Support for InfiniBand over Ethernet

3.2 Changes in Rev 2.6.000 From Rev 2.5.000

- VPI support is now at GA level
- Support for QDR interoperability with InfiniScale IV switch platforms
- Support for Reserved LKey
- Support for Fast Registration Work Request (FRWR)
- Support for congestion control
- Support for Raw QP
- Support for QUERY_ADAPTER.vsd_vendor_id parameter
- Disabled the PCI Express Extended Tag feature
- Changes to the SET_PORT command (as described in PRM rev 0.39):
 - Support for the Reset Capability Mask (rcm) field
 - Added the ability to configure VLCap and MTUCap using the command SET_PORT
 - Support for the multicast loopback disable bit field
- Support for the sniffer feature
- Support for PortXmitWait Port Counter
- Added a re-read WQE bit to the WQE Control segment

The new bit is located at bit offset 6 of the first Dword in the Control Segment. See Table 19, “General - Ctrl Segment Format” in ConnectX Programmer's Reference Manual (PRM), Rev 0.38. Firmware that supports the new bit sets the device capability bit 15 at offset 0x44 to 1. See Table 145, “Query Device Capabilities Layout” in ConnectX Programmer's Reference Manual (PRM), Rev 0.38

- Added support for twinax cables used for connecting Single and Dual SR ConnectX® EN 10 Gigabit Ethernet Adapters
- Ethernet performance improvement
- Firmware configuration (.ini) file change: To reflect a new assignment scheme for the PCI Express SubSystem ID parameter of each Mellanox adapter card, .ini files of standard Mellanox adapter cards now list the parameter hca_header_subsystem_id

4 Bug Fixes

Table 2 - Bug Fixes

	Issue	Description
1.	Congestion Control packet can cause the HCA to get stuck	Fixed
2.	Wrong handling of SW2HW/QUERY_MPT with status=FREE	Fixed

5 Known Issues

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

Table 3 - Known Issues

Index	Issue	Description	Current Implemented Workaround in FW	Possible Workaround	Scheduled Release (fix)
1.	UAR Bar is too small for 64k-page machines	The small BAR causes driver loading to fail	N/A	Change the "log2_uar_bar_megabytes" .ini parameter under the [HCA] section as follows: log2_uar_bar_megabytes = 5	N/A
2.	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	N/A
3.	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	N/A
4.	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 EN Programmer's Reference Manual</i>	N/A	Query all capabilities upon boot	N/A
5.	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	N/A	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

6 Creating a Device Configuration (.ini) File

Mellanox firmware burning tools enable setting and/or changing configuration variables by the use of an optional configuration (.ini) file. This is needed in case the default values of some variables do not suit a user's specific system requirements. This section describes how to create this configuration file.

To begin with, the .ini file is a text file is composed of one or several configuration sections (see Section 6.1 for the format and/or an example). It is recommended to include, under the appropriate sections, only those variables that need to be changed.

A firmware release includes a reference file called fw-25218-defaults.ref. This file contains the list of all variables which can be configured by a configuration (.ini) file. For each variable the reference file includes a short explanation, the [<section>] it should be under, the range of possible values, and a line with the default setting of the variable which is assumed by the firmware release.

To create the .ini file, simply copy the lines with the variables you wish to set, paste them under their appropriate [<section>] headings, and change the setting values as desired.

6.1 Configuration (.ini) File Format

The .ini file is composed of one or more sections with variable settings. Each section in the file starts with its name between square brackets, e.g. [ADAPTER], [HCA], [IB], etc. The section name is followed by one or more lines of configuration settings and comments, as in the .ini file example shown below. Note that comment lines start with a semicolon.

Excerpt from fw-25408-defaults.ref:

```

;;;; VPD support can be Disabled/Enabled
;;;; Under [ADAPTER] section
;;;; Boolean parameter. Possible values: true, false .
vpd_enable = true

```

Example of a .ini file:

```

;Begin of .ini file
[ADAPTER]
vpd_enable = false
;This is a comment line
;End of .ini file

```

7 History of Fixed Issues

Table 4 - History of Bug Fixes (Sheet 1 of 4)

	Issue	Description	Discovered in	Fixed in
1.	Link may come up as SDR instead of DDR on some systems	Fixed	2.5.000	2.6.000
2.	Possible live lock in QP upon retransmission stress	Fixed – (ID: 49870, 52066)	2.5.000	2.6.000
3.	Wrong link state reported during link speed negotiation	Fixed – (ID: 49951)	2.5.000	2.6.000
4.	CQs may be generated after CQ overrun error	Fixed – (ID: 49982)	2.5.000	2.6.000
5.	Slow handling of configuration cycles	Fixed – (ID: 49807)	2.5.000	2.6.000
6.	Wrong fields in CQE-w-Error on XRC QP	Fixed – (ID: 49742)	2.5.000	2.6.000
7.	Wrong handling of Remote Invalidate Error	Fixed	2.5.000	2.6.000
8.	Multiple RNR Nack may cause slowdown	Fixed – (ID: 49559)	2.5.000	2.6.000
9.	QUERY_DEV_CAP.apm bit was fixed at 0 even though APM was active	Fixed – (ID: 49548)	2.5.000	2.6.000
10.	PCIe physical errors upon entering L1 state	Fixed – (ID: 52025)	2.5.000	2.6.000
11.	PCI_CFG.interrupt_disable has no impact	Fixed – (ID: 53350)	2.5.000	2.6.000
12.	Non-default setting of VLCap via .ini does not take effect	Fixed	2.5.000	2.6.000
13.	SET_PORT may lead to non-optimal RX buffer reallocation if opvl was less than vlcap	Fixed	2.5.000	2.6.000
14.	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
15.	Adapter may generate PCIe transactions with wrong function ID	Fixed	2.5.000	2.6.000
16.	Adapter may generate PCIe ERR_NON_FATAL in case of an unsupported request	Fixed	2.5.000	2.6.000
17.	Ethernet only: Different VLAN priorities for WQE and QP may cause wrong SchedQ allocation	Fixed	2.5.000	2.6.000
18.	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 4 - History of Bug Fixes (Sheet 2 of 4)

	Issue	Description	Discovered in	Fixed in
19.	Wrong aliasing in address steering mode	This issues applies only to the VMware® Infrastructure 3 v3.5 operating system in netq mode	2.5.000	2.5.900
20.	Possible packet dropping though the pause policy is set	Fixed	2.5.000	2.5.900
21.	Wrong PCI Class Code for Ethernet Network Controller	This issue applies to Ethernet devices only. Fixed.	2.5.000	2.5.900
22.	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
23.	PPP does not work for an MTU of 9600	Fixed	2.5.000	2.5.900
24.	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
25.	INTA may be lost under stress	Fixed. (Ref. ID: 44473)	2.3.000	2.5.000
26.	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
27.	QP may get stuck upon Responder Gather Error	Fixed.	2.3.000	2.5.000
28.	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
29.	UC QP CQE with Error causes corruption	Fixed.	2.3.000	2.5.000
30.	Query_CQ/Query_EQ commands may return the old consumer_index	Fixed.	2.3.000	2.5.000
31.	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
32.	Possible FW internal error on a very noisy link	Fixed. (Ref. ID:41526)	2.3.000	2.5.000
33.	QueryDebugMSG command returns wrong status	Fixed. (Ref. ID: 44744)	2.3.000	2.5.000
34.	Dropping a ReadResponse packet may lead to 'retry exceeded'	Fixed.	2.3.000	2.5.000
35.	CQ moderation parameters are wrongly configured	Fixed. (Ref. ID: 45570)	2.3.000	2.5.000
36.	False generation of CQE with error (vendor code 0x6f) upon large stress	Fixed. (Ref. ID: 45317)	2.3.000	2.5.000
37.	Bandwidth degradation if SetPort command is not called	Fixed.	2.3.000	2.5.000

Table 4 - History of Bug Fixes (Sheet 3 of 4)

	Issue	Description	Discovered in	Fixed in
38.	SQERR2RTS command followed by an error causes QP to be unfunctional	Fixed. (Ref. ID: 45828 45848)	2.3.000	2.5.000
39.	QUERY_FW fails after RUN_FW	The command QUERY_FW fails after running the RUN_FW command	2.2.000	2.3.000
40.	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
41.	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
42.	IB & PCI Express links quality	General improvements	2.2.000	2.3.000
43.	Incomplete support for PCI Express 2.0 configuration header	Fixed	2.2.000	2.3.000
44.	Wrong trap generation rate	The HCA might exceed the maximum trap generation rate upon processing different trap types	2.2.000	2.3.000
45.	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
46.	Possible ICM corruption	Possible ICM (Interconnect Context Memory) corruption upon large stress (ID: 42529)	2.2.000	2.3.000
47.	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
48.	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
49.	Possible multicast corruption	Fixed (ID: 43301)	2.2.000	2.3.000
50.	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
51.	SchedQueue corruption	Fixed (ID: 43289)	2.2.000	2.3.000
52.	Wrong SL2VL mapping upon set_sl2vl	Fixed	2.2.000	2.3.000
53.	False MAD packet drops	The HCA might drop MAD packet erroneously under large stress	2.2.000	2.3.000
54.	PCI Express 2.0 x1 link fails to rise	Fixed	2.2.000	2.3.000
55.	Command timeouts	The HCA times out commands while closing multiple QPs	2.2.000	2.3.000

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

	Issue	Description	Discovered in	Fixed in
56.	False internal error generation	Fixed	2.2.000	2.3.000
57.	Transport timeouts	Multiple RNR NACKs may lead to transport timeouts (ID: 44160)	2.2.000	2.3.000
58.	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
59.	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
60.	HCA stall	The HCA might stall upon stress involving RNR Nacks and RDMA reads (ID: 41918)	2.1.000	2.2.000
61.	QP corruption	QP corruption may occur following a CQ_ overrun	2.1.000	2.2.000
62.	Sched Queue corruption	Sched Queue corruption may occur upon multiple re-transmissions	2.1.000	2.2.000
63.	False SRQ WQE limit event	A false SRQ WQE limit event is generated due to a race condition	2.0.164	2.1.000
64.	Wrong Dt value returned	The QUERY_FW command may return a wrong Dt value	2.0.164	2.1.000
65.	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
66.	High ACK latency	Delays in ACK may cause multiple local ACK timeouts	2.0.164	2.1.000
67.	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
68.	IB link stability issues		2.0.164	2.1.000
69.	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