Release Notes

MT25408 ConnectX™ Firmware fw-25408

Supporting: InfiniBand, Ethernet, FCoE, VPI

Rev 2.5.000

MT25408 ConnectX Firmware fw-25408 Release Notes

Document Number:

Mellanox Technologies, Inc.
2900 Stender Way
Santa Clara, CA 95054
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

1 Overview

These are the release notes for the ConnectX™ and ConnectX™ EN adapters firmware, fw-25408 Rev 2.5.000. This firmware supports the following protocols:

- InfiniBand
- Ethernet
- Fibre Channel over Ethernet (FCoE)
- 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 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	Comments
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	Firmware support for InfiniBand only or Ethernet only configurations is at GA level; Firmware support for FCoE and VPI is at Beta level;
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	Ethernet firmware is at GA level;
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:

2 Revision Compatibility

- Firmware fw-25408 Rev 2.5.000 complies with the following programmer's reference manuals:
  - *ConnectX Programmer's Reference Manual (PRM), Rev 0.38 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 Major New Features

- Support for the Ethernet, FCoE, and VPI protocols on the ConnectX devices
- Support for the ConnectX EN device (Ethernet or FCoE only)
- Added QoS support according to Section 7.6, "Virtual Lanes Mechanism," in InfiniBand Architecture Specification, Vol. 1, Release 1.2.1
- Added Automatic Path Migration (APM) support according to Section 17.2.8, "Automatic Path Migration," in InfiniBand Architecture Specification, Vol. 1, Release 1.2.1
- Support for Local/Remote Invalidate (Beta level)
- Support for Mellanox Technologies' QDR Autonegotiation protocol
- Support for Data Integrity Field (DIF) Offload
- Added a mechanism to check and prevent the device from running illegal port speed configurations (Note: the RUN_FW command will fail)
- The *log_page_size* parameter is now configurable in the ModStatConfig command
- Support for the Ethernet *Per Priority Pause* feature

4 Bug Fixes

The following table describes known issues from previous releases of ConnectX™ IB firmware which were fixed in this firmware release.

Table 2 - Bug Fixes

	Issue	Description	Discovered in	Fixed in
1.	PCI Express compliancy issues	• 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
2.	INTA may be lost under stress	Fixed. (Ref. ID: 44473)	2.3.000	2.5.000
3.	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
4.	QP may get stuck upon Responder Gather Error	Fixed.	2.3.000	2.5.000
5.	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
6.	UC QP CQE with Error causes corruption	Fixed.	2.3.000	2.5.000
7.	Query_CQ/Query_EQ commands may return the old consumer_index	Fixed.	2.3.000	2.5.000
8.	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
9.	Possible FW internal error on a very noisy link	Fixed. (Ref. ID:41526)	2.3.000	2.5.000
10.	QueryDebugMSG command returns wrong status	Fixed. (Ref. ID: 44744)	2.3.000	2.5.000
11.	Dropping a ReadResponse packet may lead to ‘retry exceeded’	Fixed.	2.3.000	2.5.000
12.	CQ moderation parameters are wrongly configured	Fixed. (Ref. ID: 45570)	2.3.000	2.5.000
13.	False generation of CQE with error (vendor code 0x6f) upon large stress	Fixed. (Ref. ID: 45317)	2.3.000	2.5.000
14.	Bandwidth degradation if SetPort command is not called	Fixed.	2.3.000	2.5.000
15.	SQERR2RTS command followed by an error causes QP to be unfunctional	Fixed. (Ref. ID: 45828 45848)	2.3.000	2.5.000

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	NA	Change the "log2_uar_bar_megabytes" .ini parameter under the [HCA] section as follows: log2_uar_bar_megabytes = 5	NA
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)	NA	NA	NA
3.	BAR resizing on an enabled system	Changing bar sizes when a system is enabled may cause the device to hang (ID: 24208)	NA	NA	NA
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 *ConnectX EN Programmer's Reference Manual*	NA	Query all capabilities upon boot	2.6.000

6 Firmware-Enabled InfiniBand Features

The following table lists the enabled InfiniBand features by this firmware release. Enabled features are marked with 'v'.

Table 4 - Enabled ConnectX IB Features (Sheet 1 of 5)

Feature	Item/Command	Enabled ?
Transport Services		
	RC	v
	UD	v
	MLX	v
	UC	v
SRQ		v
SMA		
	internal	v
	external	v
	guid info	v
	pkey table	v
	node info	v
	port info	v
	node description	v
	SL2VL	v
	VL arbitration	v
	GSA	v
	traps	v
Special QP		v
Recoverable Errors (RNR, retransmission)		v
Unrecoverable Error (access violation)		v
Events		
	general event handling	v
	completion events	v

Table 4 - Enabled ConnectX IB Features (Sheet 2 of 5)

Feature	Item/Command	Enabled ?
Affiliated Asynchronous Events		
	srq limit	v
	path migration events	v
	srq last wqe reached	v
	send queue drained	
	communication established	v
Affiliated Asynchronous Errors		
	cq error	v
	local wq catastrophic error	v
	path migration error	v
	invalid request local WQ error	v
	Local access violation WQ error	v
	Local SRQ catastrophic error	v
Unaffiliated Events		
	port state changed	v
	client reregister	v
	gpio event	
	command interface	v
IB Link Speed		
	SDR	v
	DDR (1.1)	v
	DDR (1.2)	
	QDR	v
BlueFlame		v
QoS		v
PCIe		
	basic config cycle	v
	extended config cycle	v
	expansion ROM	v
	interrupts	v
	MSIX	v
	PCIe gen2	v

Table 4 - Enabled ConnectX IB Features (Sheet 3 of 5)

Feature	Item/Command	Enabled ?
Command Interface		
	hcr	v
Initialization Commands		
	map_fa	v
	unmap_fa	v
	run_fw	v
	set_icm_size	v
	map_icm_aux	v
	unmap_icm_aux	v
	map_icm	v
	unmap_icm	v
	query_dev_lim	v
	query_fw	v
	query_adapter	v
	mod_stat_comfig	v
	init_hca	v
	close_hca	v
	init_post	v
	close_ib	v
	query hca	v
	init_vm	
	set ib	v

Table 4 - Enabled ConnectX IB Features (Sheet 4 of 5)

Feature	Item/Command	Enabled ?
QP commands		
	RST2INIT	v
	INIT2INIT	v
	INIT2RTR	v
	RTR2RTS	v
	RTS2RTS	v
	RTS2SQD	
	SQERR2RTS	v
	SQD2SQD	
	2ERR	v
	2RST	v
	QueryQP	v
	SUSPEND_QP	v
	unsuspend_qp	v
	special QP	v
	conf special qp	v
	mad_ifc	v
TPT commands		
	SW2HW MPT	v
	QueryMPT	v
	HW2SW MPT	v
	WRITE MTT	v
	READ MTT	v
	SyncTPT	v
	MODIFY MPT	v
SRQ Commands		
	SW2HW SRQ	v
	HW2SW SRQ	v
	MODIFY SRQ - ARMING	v
	RESIZE SRQ	

Table 4 - Enabled ConnectX IB Features (Sheet 5 of 5)

Feature	Item/Command	Enabled ?
Multicast Commands		
	WRITE_MGM	v
	READ_MGM	v
	MGID_HASH	v
EQ Commands		
	map_eq	v
	sw2hw_eq	v
	hw2sw_eq	v
	qurey_eq	v
CQ Commands		
	sw2hw_cq	v
	hw2sw_cq	v
	query_cq	v
	modify_cq	v
Debug Commands		
	QueryDebugMSG	v
	SetDebugMSG	v
	DiagRprt	v
Bind Memory Window		
FMR		v
APM		v
Multicast		v
LED blinking		v
GPIO		v
IWARP Verbs		
RMC		
DIF		v
Extended Atomic		v
IPoIB stateless offload		v
Inline scatter		
Virtualization		
XRC		v

7 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 7.1 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.

7.1 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-25218-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
```

8 History of Fixed Issues

Table 5 - History of Bug Fixes (Sheet 1 of 2)

	Issue	Description	Discovered in	Fixed in
1.	QUERY_FW fails after RUN_FW	The command QUERY_FW fails after running the RUN_FW command	2.2.000	2.3.000
2.	HCA stall	The HCA might stall in any of the following scenarios: • 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
3.	QUERY_QP errors	Wrong QUERY_QP command in the following cases: • Returns wrong values (ID: 42078, 40707) • Enters the error state erroneously (ID: 43110)	2.2.000	2.3.000
4.	IB & PCI Express links quality	General improvements	2.2.000	2.3.000
5.	Incomplete support for PCI Express 2.0 configuration header	Fixed	2.2.000	2.3.000
6.	Wrong trap generation rate	The HCA might exceed the maximum trap generation rate upon processing different trap types	2.2.000	2.3.000
7.	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
8.	Possible ICM corruption	Possible ICM (Interconnect Context Memory) corruption upon large stress (ID: 42529)	2.2.000	2.3.000
9.	Performance	HCA performance improvements for the following cases: • Upon receiving multiple ACK packets • Upon multiple QPs in error state (ID:43377) • Upon multiple RNR NACKs	2.2.000	2.3.000
10.	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
11.	Possible multicast corruption	Fixed (ID: 43301)	2.2.000	2.3.000
12.	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
13.	SchedQueue corruption	Fixed (ID: 43289)	2.2.000	2.3.000
14.	Wrong SL2VL mapping upon set_sl2vl	Fixed	2.2.000	2.3.000
15.	False MAD packet drops	The HCA might drop MAD packet erroneously under large stress	2.2.000	2.3.000
16.	PCI Express 2.0 x1 link fails to rise	Fixed	2.2.000	2.3.000

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

	Issue	Description	Discovered in	Fixed in
17.	Command timeouts	The HCA times out commands while closing multiple QPs	2.2.000	2.3.000
18.	False internal error generation	Fixed	2.2.000	2.3.000
19.	Transport timeouts	Multiple RNR NACKs may lead to transport timeouts (ID: 44160)	2.2.000	2.3.000
20.	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
21.	Wrong ***sl*** and/or port number returned	The QUERY_QP command may return a wrong ***sl*** value and/or a wrong port number (ID: 40707)	2.1.000	2.2.000
22.	HCA stall	The HCA might stall upon stress involving RNR Nacks and RDMA reads (ID: 41918)	2.1.000	2.2.000
23.	QP corruption	QP corruption may occur following a CQ_overrun	2.1.000	2.2.000
24.	Sched Queue corruption	Sched Queue corruption may occur upon multiple re-transmissions	2.1.000	2.2.000
25.	False SRQ WQE limit event	A false SRQ WQE limit event is generated due to a race condition	2.0.164	2.1.000
26.	Wrong Dt value returned	The QUERY_FW command may return a wrong Dt value	2.0.164	2.1.000
27.	HCA hangs	The device hangs in one of the following cases: • 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
28.	High ACK latency	Delays in ACK may cause multiple local ACK timeouts	2.0.164	2.1.000
29.	HCA performance	HCA performance may be impacted in the following conditions: • QPs in error state • Slow QP context handling	2.0.164	2.1.000
30.	IB link stability issues		2.0.164	2.1.000
31.	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