Release Notes

MT25208 InfiniHost III Ex Firmware

(MT23108 InfiniHost-Compatible Mode)

FW-25208 Rev 4.7.400

MT25208 InfiniHost III Ex Firmware FW-25208 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 MT25208 InfiniHost III Ex firmware, FW-25208 Rev 4.7.400. It is appropriate for the MT25208 device functioning as an MT23108 InfiniHost-Compatible device. This firmware supports the Mellanox HCA Adapter Cards listed in Table 1.

Note: After burning new firmware to an HCA board, reboot the machine so that the new firmware can take effect.

Table 1 - Supported HCA Adapter Cards

HCA Card OPN	Code Name	Description
MHEL-CFXXX-T[1] (previously MTLP25208)	Lion Cub SDR	InfiniHost III Ex PCI Express HCA Adapter Card. Note: IB ports support operation at single data rate (SDR) only.
MHXL-CFXXX-T[1] (previously MTLP23108)	Lion Cub 128 DDR / Lion Cub 256 DDR	InfiniHost III Ex PCI Express HCA Adapter Card with 128 / 256MB of local memory. Note: IB ports support operation at double data rate (DDR).

1. XXX reflects the size of on-board memory (in MB): 128, 256, or 512.

The document consists of the following sections:

- "Major New Features" (page 3)
- "Known Issues" (page 4)
- "Bug Fixes" (page 6)
- "Invariant Sector (IS) Changes / Fixes" (page 7)
- "History" (page 8)
- "Creating a Device Configuration (.ini) File" (page 10)

2 Major New Features

- A new firmware versioning method is now applied
- Configuration of SDRAM interface weak output drivers per bus (via .ini file)
- Added support for the SRQ Limit event

3 Known Issues

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

Table 2 - Known Issues

Index	Issue	Description	Current Implemented Workaround in FW	Possible Workaround	Patch Release (fix)	Scheduled Release (fix)
1.	MSIx vectors	Writing to MSIx vectors (Address/Data/Mask) does not take immediate effect. There may be MSIx messages that leave the device according to the old vector.	NA	Commit a PCI configuration cycle after the MSIx modification	NA	NA
2.	QPC.Flight_LIM	QPC field – no HW limit, infinite WQEs on send.	NA	NA	NA	NA
3.	QUERY_DDR	Query does not return JEDEC vendor ID yet. Scope of status is limited to active / not active.	NA	NA	NA	NA
4.	RTR2RTS_QPEE; SQD2RTS_QPEE: changing optional fields rra_max and ra_buf_index is not supported.	The optional fields rra_max and ra_buf_index are not supported in the RTR2RTS_QPEE and SQD2TRS_QPEE commands.	Change requests for these fields will not take effect, and no error indication is provided.	Mask these optional fields	NA	NA
5.	PCI 2.3 control and status - for interrupts	InfiniHost III Ex does not support PCI2.3 control and status bits for interrupts.	NA	NA	NA	NA
6.	Change of memory bars on a disabled system	Changing memory bars size / addresses between SYS_DIS and SYS_EN may cause the InfiniHost III Ex to hang (ID: 24206)	NA	NA	NA	NA
7.	BAR resizing on an enabled system	Changing bar sizes when a system is enabled may cause the InfiniHost III Ex to hang (ID: 24208).	NA	NA	NA	NA
8.	SW reset via configuration cycles	SW reset via config cycles may create double PCI- Express completions for the configuration transaction.	NA	If InfiniHost III Ex boots in memory controller mode, perform power cycle / hot reset after restoring the flash.	NA	NA
9.	SW reset is performed during a configuration transaction	If SW reset is performed while a configuration transaction is outstanding, it may create double PCI- Express completions for the configuration transaction.	NA	Do not perform SW reset during configuration cycles.	NA	NA

3.1 Unsupported *InfiniHost Programmer's Reference Manual* Changes

The following features of *InfiniHost Programmer's Reference Manual* are not scheduled to be supported:

1. Flight lim value in QPC may show a value other than 0'1111 even when set for unlimited usage.

4 Bug Fixes

The following table describes known issues from previous releases of InfiniHost III Ex firmware which were fixed in this firmware release.

Table 3 - Bug Fixes

Issue	Description	Discovered in	Fixed in
ACK delay when sender is back-pressured	This causes a transport timeout to the remote peer	4.7.0	4.7.400
RNR timer is always set to 0x1F	RNR timer for a Sender retry is always configured to 0x1F regardless of the QP set value (ID:31959)	4.7.0	4.7.400
Unreliable-QP context corruption	(ID:31927)	4.7.0	4.7.400
SDRAM INIT operation when exiting Self Refresh not implemented	Now it is implemented	4.7.0	4.7.400
ECC errors handling	ECC errors may not be reported at all, may cause a wrong SERR# assertion, or may be reported with the wrong ColumnAddress. (ID:31614)	4.7.0	4.7.400
Link up in 1X rather than 4X	Signal Detect must happen before RX powerup; otherwise, the port comes up in 1X mode instead of 4X (ID:32401)	4.7.0	4.7.400
MADs:PortInfo Get()	When querying for information about an InfiniHost III Ex IB port via its other IB port, the wrong Local port number is returned. Instead of the number of the second port, the one which received the MAD packets, the number of the first port is being returned. (ID: 24177)	4.6.2	4.7.0
A Concurrent Bind and Deallocate for the same Memory Window will prevent closing the Memory Region of this Window	Bind and Deallocate modify the same 'unprotected' variable of the Memory Region. If both operations are attempted simultaneously for the same Memory Window, the variable does not get updated correctly. This prevents closing the Memory Region as the corrupted variable value may indicate that a Memory Window is still bound to it.	4.6.2	4.7.0
Requester ScatterList corruption upon CQ error	A CQ error can cause corruption in the Requester ScatterList Database. As a result QPs may move to error, and the device may stop sending packets (ID: 30670)	4.6.2	4.7.0
IB Tx phase detector should be opened when link state is Config Debounce	(ID: 24332)	4.6.2	4.7.0
FW deadlock in an environment of BIND and HW2SW_MPT	(ID: 29814)	4.6.2	4.7.0
QP deadlocks when doing a RETRY	(ID: 29676)	4.6.2	4.7.0
Big UAR pages	Support for Big UAR pages is not complete (ID: 29496)	4.6.2	4.7.0
Multicast Index miscalculation	Multicast Index miscalculation may cause dropping of multicast packets instead of inserting them. (ID: 29469)	4.6.2	4.7.0
CQ error or QP error together with 2ERR_QPEE may cause CommandIF to hang	(ID: 29431,29737)	4.6.2	4.7.0
FW deadlock when flushing a QP	(ID: 29277)	4.6.2	4.7.0
SRQ deadlocks when QP goes to error	(ID: 29174)	4.6.2	4.7.0
After a Catastrophic Error, HCA start may fail	(ID: 29066)	4.6.2	4.7.0

Table 3 - Bug Fixes (Continued)

Issue	Description	Discovered in	Fixed in
Port state ACTIVE_DIFFER should be reported as ACTIVE	(ID: 28811)	4.6.2	4.7.0
DIMM Unrecoverable Error not detected	(ID: 28902)	4.6.2	4.7.0
EQC.intr for the Catastrophic Error EQ is hard wired to 0x0	It now can be any legal value (including MSIx) (ID: 28815,28377)	4.6.2	4.7.0
SRQ performance is too low	(ID: 28702)	4.6.2	4.7.0
MSIx vector race when updating MSIx Table	(ID: 26599)	4.6.2	4.7.0
UD starvation	UD messages are not sent because RC ACKs are not arriving (ID: 28374,28427)	4.6.2	4.7.0
SerDes electrical stress may occur if VDDIO > 1.2V	(ID: 28385)	4.6.2	4.7.0
A SendQ connected to an SRQ may get a wrong RETRY_EXCEEDED	MPT window count is corrupted when BIND is used excessively (ID: 29953)	4.6.2	4.7.0
Fast Self Refresh feature	Fast Self Refresh feature is not functional. (ID: 30014)	4.6.2	4.7.0
DIMM timing parameters	DIMM timing parameters are not configured correctly	4.6.2	4.7.0

5 Invariant Sector (IS) Changes / Fixes

None in this release.

6 History

Table 4 - History of Fixed Bugs

Issue	Description	Discovered in	Fixed in
Consumer Index corruption in a Completion Queue	When using Increment_CI doorbells, to increment the CI in more than 1, CI may advance wrongly, casing a false CQ overrun, or not detecting a real overrun (ID: 27893)	4.6.1	4.6.2
Memory Region WindowCount corruption	When deallocating a window and trying to bind it simultaneously, the Region entry WindowsCount may be corrupted. (ID: 26829)	4.6.1	4.6.2
Access to VPD with partial ByteEnables	A configuration access to VPD with partial ByteEnables may cause VPD corruption (ID: 27690)	4.5.3	4.6.1
MAD with bad methods	a MAD with an unsupported method should be dropped (ID: 27472)	4.5.3	4.6.1
QP in retry may halt sending	A QP that is executing a message retry may hang and stop sending it (ID: 27252)	4.5.3	4.6.1
Q_Key source changes from QPC /WQE	If the Q_Key source changes from QPC /WQE, InfiniHost generates a packet with a wrong Q_key (ID: 21987)	4.5.3	4.6.1
Successful TimeOut-Driven-APM may cause QP context corruption	When an APM occurs as a result of TimeOut, QP context may be corrupted (ID: 26632)	4.5.3	4.6.1
Closing-QP commands get stuck	The commands 2RST_QPEE and RST2ERR_QPEE get stuck if a bad NACK is sent simultaneously (ID: 26243)	4.5.3	4.6.1
Validation of duplicate RDMA_READ/Atomic	Duplicate RDMA_READs/Atomics are not validated against the original requests (ID: 26247)	4.5.3	4.6.1
APM EQE due to TimeOut	An APM that resulted from a TimeOut does not generate an “APM succeeded” EQE (ID: 25948)	4.5.3	4.6.1
Binding Memory Windows across a 4GB boundary	There is an error in binding a Memory Window across a 4GB boundary (ID: 25958)	4.5.3	4.6.1
PCI express DeviceControl.unsupported_error_report_enable is R/W (Invariant Sector)	When flash is corrupted, bit 3 in device control (unsupported error reporting enable) is Read Only. It must be read/write (ID: 25570)	4.5.0	4.5.3
MAD with wrong BaseVersion or ClassVersion	MAD with wrong BaseVersion or ClassVersion should return status 0x1 (ID: 25888)	4.5.0	4.5.3
PortInfo.ResponseTime	The returned PortInfo.ResponseTime is too short (ID: 10597)	4.5.0	4.5.3
MAD with wrong AttributeModifier	MAD with wrong AttributeModifier should cause the response-MAD to have status INVALID_ATTR (ID: 25875)	4.5.0	4.5.3
MTTs in addresses > 4GB	If MTTs are in address > 4GB (e.g. in HIDE DDR), the device may get stuck (ID: 25877)	4.5.0	4.5.3
FW Debug version Data Section corruption	In Debug version of FW, FW Linker can cause a corruption in the Data Section of iRISC (ID: 25774)	4.5.0	4.5.3
PCI express DeviceControl.unsupported_error_report_enable is R/W	When flash is corrupted, bit 3 in device control (unsupported error reporting enable) was Read Only. It must be read/write (ID: 25570)	4.5.0	4.5.3

Issue	Description	Discovered in	Fixed in
Link phy error threshold	link phy error threshold was set to a value larger by 1 than the one accepted in portInfo SET() (ID: 25854)	4.5.0	4.5.3
QP state mis-calculation	QPSTATE is shifted in a bit when QPC is non-cacheable. This may cause an unsolicited ACK to be dropped (ID: 25814)	4.5.0	4.5.3
local_ca_ack_delay	QUERY_DEV_LIM.local_ca_ack_delay was 16, which implies 266ms. Changed to 15. Driver can choose to return a lower number for in a non-stress case (ID: 25584)	4.5.0	4.5.3
CQE with error counters	The error counters sq_num_wrfe and rq_num_wrfe miss some increments (ID: 24463)	4.5.0	4.5.3
2err_qp/2rst_qpee starvation	If a CQE with Error is scheduled, and a heavy back-pressure is applied by UpLink, QP may be put to end of queue, and wait a long time till next chance to generate the CQE w/ Error (ID: 25044)	4.5.0	4.5.3
EQE loss	In stress, an ARM doorbell that should discover a non-empty CQ may be bypassed by a SET_CI doorbell. Thus CQ will seem empty, and an EQE will not be generated (ID: 24911)	4.5.0	4.5.3
MSIX memory mapped tables endianess	MSIX vectors and pending bits have wrong endianess (ID: 22794)	4.5.0	4.5.3
PCI compliancy and init changes (Invariant Sector)	Revision number changed to 21	4.0.1	4.5.0

Table 5 - Invariant Sector Fixed Issues History

Issue	Description	Discovered in	Fixed in
PCI memory space must be disabled by default	PCI memory space must be disabled by default (ID: 27341)	4.5.3	4.6.1
SW reset may block PCIe message generation	In some cases, a SW reset may block the generation of future PCIe messages by the device (ID: 21721,27015)	4.5.3	4.6.1

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-25208-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-25208-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
```