



Release Notes

MT25208 InfiniHost™ III Ex Firmware

Mem-Free Mode: PCI DevID 25218 (Decimal)

fw-25218 Rev 5.3.000

© Copyright 2007. Mellanox Technologies, Inc. All Rights Reserved.

MT25208 InfiniHost(TM) III Ex Mem-Free Firmware fw-25218 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

Mellanox Technologies

1 Overview

These are the release notes for the MT25208 InfiniHost™ III Ex firmware, fw-25218 Rev 5.3.000. It is appropriate for the MT25208 device functioning in Mem-Free mode (that is, without local memory). For the most updated list of HCA Adapter Cards supported by this firmware, visit the firmware download pages via <http://www.mellanox.com>.

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

The document consists of the following sections:

- “Revision Compatibility” (page 3)
- “Major New Features” (page 3)
- “Bug Fixes” (page 4)
- “Invariant Sector (IS) Changes / Fixes” (page 5)
- “Known Issues” (page 6)
- “Creating a Device Configuration (.ini) File” (page 7)
- “History of Fixed Issues” (page 8)

2 Revision Compatibility

- Firmware fw-25218 Rev 5.3.000 complies with the *InfiniHost™ III Programmer’s Reference Manual (PRM), Rev. 0.89, Document No. 2248PM*.

3 Major New Features

- Active cable support
- Added the following debug features:
 - The new .ini parameter `absorb_ecc_unrecoverable` can be used to configure the HCA to treat DIMM unrecoverable errors as recoverable
 - The new .ini parameter `gpio_num_indicates_fsr_en` can be used to set a specific GPIO to indicate whether Fast Self-Refresh mode is enabled or disabled. See the command `ENABLE_LAM` in the PRM.

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 1 - Bug Fixes

	Issue	Description	Discovered in	Fixed in
1.	Doorbell loss upon UAR page size >= 64KB	Fixed (ID: 39840)	5.2.000	5.3.000
2.	False Fatal Error	A false Fatal Error is reported upon a WQE fetch translation error (ID: 39843)	5.2.000	5.3.000
3.	Schedule Queue corruption	May cause an internal error upon closing QPs in a condition of RNR stress (ID: 40576)	5.2.000	5.3.000
4.	Multicast performance	Performance may be degraded upon receiving a multicast packet that is not a member of any multicast group (ID:40668)	5.2.000	5.3.000
5.	MTT corruption	Fixed (ID: 42024)	5.2.000	5.3.000
6.	Duplicate receive CQE	Under very stressful conditions, multiple receive CQEs may be generated (ID: 40730)	5.2.000	5.3.000
7.	Incomplete implementation of the ModStatConfig command	It is possible to re-enable a port using a MAD after disabling it with the command ModStatConfig (ID: 39018)	5.2.000	5.3.000
8.	Wrong Revision return value by NodeInfo	NodeInfo returns the wrong Revision (device revision) value for a DDR device (ID: 39357)	5.2.000	5.3.000
9.	Send CQE in error compliance	A Send CQE with error is generated only if the QP was in RTS prior to moving to the Error state (ID: 41203)	5.2.000	5.3.000
10.	QUERY_FW fails after RUN_FW	The command QUERY_FW fails after running the RUN_FW command	5.2.000	5.3.000
11.	Wrong trap generation rate	The HCA might exceed the maximum trap generation rate upon processing different trap types	5.2.000	5.3.000
12.	Suspend_QP command timeout	Fixed (ID: 42473)	5.2.000	5.3.000

5 Invariant Sector (IS) Changes / Fixes

None in this release.

Mellanox Technologies

6 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.	DoorBell loss	DoorBells may be lost on systems with a 64-KByte page size upon heavy stress conditions	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 Infini-Host™ 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

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 3 - History of Fixed Issues (Sheet 1 of 4)

Issue	Description	Discovered in	Fixed in
CQ context corruption	<ul style="list-style-type: none"> May occur in a multiple CQ contexts environment (ID: 37432-3) May occur under stress (ID:37213-4, 36897) May cause HCA stall (ID:36209) Upon SuspendQP in conjunction with QP Error (ID:38876) 	5.1.400	5.2.000
Missing or corrupted CQE with Error	CQE with error may be lost or corrupted upon QP moving to error while receiving a multi-packet message (ID:37049-50)	5.1.400	5.2.000
EQ_ARM doorbell	May be lost (ID: 37872)	5.1.400	5.2.000
ICM context corruption	May occur when executing MAP_ICM command (ID:37790)	5.1.400	5.2.000
relaxed_ordering_en bit is not PCI Express compliant	PCI Express bit relaxed_ordering_en is R/W instead of Read Only	5.1.400	5.2.000
RDMA_READ and ATOMIC Deadlock	Internal leak may occur when doing RDMA_READS and ATOMIC on Send Queues (ID: 27761)	5.1.400	5.2.000
E2E credits miscalculation	May occur on OOS/RNR Nack (ID:37308/28) or on suspend_qp (ID:37608)	5.1.400	5.2.000
LinkSpeedEnable	LinkSpeedEnable is not applied correctly (ID: 36592)	5.1.400	5.2.000
Fatal error in large-stress condition	FW may fail (FW_INTERNAL_ERROR) upon large stress (ID: 38845,39393)	5.1.400	5.2.000
Bringing IB port up on DDR device	DDR IB link does not go up after being in port DISABLE state (ID:36586)	5.1.400	5.2.000
Vendor Specific MADs	<ul style="list-style-type: none"> Wrong Vendor Key (V_KEY) authentication and mismatch behavior (ID:36127) Wrong Revision ID in General Info MAD (ID: 38025) 	5.1.400	5.2.000
UNMAP_FA command may get stuck if IB port runs at DDR	Fixed (ID: 36757,39325)	5.1.400	5.2.000
Potential FW deadlock	FW may deadlock upon heavy stress (ID: 39393)	5.1.400	5.2.000
HCA may hang upon SW reset	If the HCA has a pending In-Bound read and a SW reset occurs simultaneously, the HCA may hang (ID: 38053,39660,39661)	5.1.400	5.2.000
Corrupted Read/Atomic response	Possible corruption of Read or Atomic responses (ID:36353, 37014, 37026)	5.1.400	5.2.000
Serdes Parameters	Serdes Parameters are not applied correctly to port 2 (ID:37782-3)	5.1.400	5.2.000
QUERY_DEBUG_MSG	msg_hdr_size is wrong (ID:38428)	5.1.400	5.2.000
Out-of-Sequence packets following RNR Nack may cause back-pressure	Slow handling of Out-of-Sequence packets following RNR Nack may cause back-pressure (ID: 38958)	5.1.400	5.2.000
Incorrect Query_FW dt field	If the command QUERY_FW is run before the command RUN_FW, the returned QUERY_FW dt (Debug Trace) bit is returned with the wrong value (ID: 38428)	5.1.400	5.2.000
Failure upon posting first WQE at a random location	The first WQE can be posted to any random location in the WQE buffer. FW fails if this location is not the first. (ID: 39413-4)	5.1.400	5.2.000
SendWQ with size 1	SendWQ with size 1 causes a wrong WQE address in CQE (ID:33302)	5.1.0	5.1.400
ACK delay when sender is back-pressured	This causes a transport timeout to the remote peer	5.1.0	5.1.400

Table 3 - History of Fixed Issues (Sheet 2 of 4)

Issue	Description	Discovered in	Fixed in
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)	5.1.0	5.1.400
Unreliable-QP context corruption under stress	(ID:31927)	5.1.0	5.1.400
Exit Self-Refresh is not followed by INIT operation	Fixed	5.1.0	5.1.400
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)	5.1.0	5.1.400
Wrong Client ReRegister bit	Client ReRegister bit may be wrong (ID:33958,35356)	5.1.0	5.1.400
Wrong APM behavior	Packet from wrong port may be dropped even if MigReq is not expected. Also, a bad path packet may be allowed in MIGRATED state. (ID: 35179)	5.1.0	5.1.400
FW deadlock upon event queue overrun	An event queue overrun may cause a firmware deadlock	5.1.0	5.1.400
Wrong M_Key check	Given MKeyProtectBits<2, if a SubnGet(NodeInfo) with a wrong M_Key is received while the M_Key Lease Period counter is already active (due to a previous M_Key violation), the counter gets stopped instead of continuing with the original count (ID: 33388)	5.1.0	5.1.400
Inbox/outbox alignment	Inbox/outbox alignment to 4K is not enforced (ID: 29323)	5.1.0	5.1.400
APM with changed SL may result in corruption	A schedule queue corruption may occur upon an APM with a changed SL (ID: 32672)	5.1.0	5.1.400
Wrong behavior upon UC send with first message packet dropped	If the first packet of an Unreliable Connection gets discarded, consequent UC behavior may be wrong (ID: 34689)	5.1.0	5.1.400
QP Flush during an SQ Drain may result in corruption	Flushing the QP during an SQ Drain may cause a schedule queue corruption (ID: 34879)	5.1.0	5.1.400
Bind Memory Window checks	The Bind Memory Window command does not check that all index bits of the window to be bound are identical to those of an old window (ID: 35035)	5.1.0	5.1.400
wrong behavior on SuspendQP in UD	SuspendQP in a UD may cause a local QP error (ID 35524)	5.1.0	5.1.400
Bad params in QueryQP()	RecvQP/SRQ return a wrong DoorBellRecordIndex upon a Query QP (ID: 35001)	5.1.0	5.1.400
UAR context not in address 0	The HCA does not work if its UAR context table is not set as address 0x0 of ICM (ID: 35001)	5.1.0	5.1.400
SERR# (System Error)	The Signaled System Error Status bit is not set after an Error message is sent (ID: 35458, 35703). The enable_serr Control bit has no effect. (ID: 23409, 26881, 35693)	5.1.0	5.1.400
Wrong MAD status	MAD status may be "Invalid combination of Method and Attribute" instead of "Invalid Attribute" (ID: 35226)	5.1.0	5.1.400
Link up in 1X rather than 4X	Signal Detect must happen before RX power-up; otherwise, the port comes up in 1X mode instead of 4X (ID:32401)	5.1.0	5.1.400
Send Work queue corruption	HW may overrun Send WQ in case of multiple RDMA READs or in case of entering Limited State (ID: 35649)	5.1.0	5.1.400
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)	5.0.1	5.1.0

Table 3 - History of Fixed Issues (Sheet 3 of 4)

Issue	Description	Discovered in	Fixed in
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.	5.0.1	5.1.0
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)	5.0.1	5.1.0
IB Tx phase detector should be opened when link state is Config Debounce	(ID: 24332)	5.0.1	5.1.0
MLX QP may be broken under heavy load of DoorBells	(ID: 30544)	5.0.1	5.1.0
FW deadlock in an environment of BIND and HW2SW_MPT	(ID: 29814)	5.0.1	5.1.0
Internal error when receiving RDMA_READ or ATOMIC requests	RDB access fails when MTT base != 0x0 for an RDMA_READ or ATOMIC request, and internal error occurs. (ID: 29701)	5.0.1	5.1.0
QP deadlocks when doing a RETRY	(ID: 29676)	5.0.1	5.1.0
Big UAR pages	Support for Big UAR pages is not complete. (ID: 29496)	5.0.1	5.1.0
QP.send_doorbell_record_index not returned in queryQP	(ID: 29488)	5.0.1	5.1.0
CQ error may cause FW to deadlock when doing HW2SW_CQ	(ID: 29439)	5.0.1	5.1.0
Multicast Index miscalculation	Multicast Index miscalculation may cause dropping of multicast packets instead of inserting them. (ID: 29469)	5.0.1	5.1.0
CQ error or QP error together with 2ERR_QPEE may cause CommandIF to hang	(ID: 29431,29737)	5.0.1	5.1.0
Schedule Queue corruption may cause QPs to deadlock	(ID: 29292)	5.0.1	5.1.0
FW deadlock when flushing a QP	(ID: 29277)	5.0.1	5.1.0
SRQ deadlocks when QP goes to error	(ID: 29174)	5.0.1	5.1.0
After a Catastrophic Error, HCA start may fail	(ID: 29066)	5.0.1	5.1.0
Port state ACTIVE_DIFFER should be reported as ACTIVE	(ID: 28811)	5.0.1	5.1.0
DIMM Unrecoverable Error not detected	(ID: 28902)	5.0.1	5.1.0
EQC.intr for the Catastrophic Error EQ is hard wired to 0x0	It now can be any legal value (including MSIx) (ID: 28815,28377)	5.0.1	5.1.0
e2e credits may not be seen for a short period. Thus UC QP may drop packets, and RC QP may RNR nack	(ID: 28825)	5.0.1	5.1.0
SRQ performance is too low	(ID: 28702)	5.0.1	5.1.0
MSIx vector race when updating MSIx Table	(ID: 26599)	5.0.1	5.1.0
SQ DoorBells may be lost in rare cases	(ID: 28646)	5.0.1	5.1.0

Table 3 - History of Fixed Issues (Sheet 4 of 4)

Issue	Description	Discovered in	Fixed in
UD starvation because RC ACKs are not arriving	UD messages are not sent because RC ACKs are not arriving (ID:28374,28427)	5.0.1	5.1.0
SerDes electrical stress may occur if VDDIO > 1.2V	(ID: 28385)	5.0.1	5.1.0
Request Notification for multiple CQEs may take the wrong "multiple" number, and wait for more CQEs in order to notify	(ID: 28096)	5.0.1	5.1.0
MPT window count is corrupted when BIND is used excessively	(ID: 27759)	5.0.1	5.1.0
Fast Self Refresh feature is not functional	(ID: 30014)	5.0.1	5.1.0
DIMM timing parameters	DIMM timing parameters are not configured correctly	5.0.1	5.1.0

The following table describes Invariant Sector changes and/or fixes with respect to the last firmware release.

Table 4 - IS Changes / Fixes

	Issue	Description	Discovered in	Fixed in
1.	PCI Express x2	A PCI Express x2 link width is not allowed anymore	5.1.400	5.2.000

Mellanox Technologies