Mellanox
TECHNOLOGIES

MT25204 InfiniHost III Lx Firmware

fw-25204 Rev 1.2.000

MT25204 InfiniHost III Lx Firmware 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 MT25204 InfiniHost III Lx firmware, fw-25204 Rev 1.2.000. 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
MHES14-XT / MHES14-XS MHES14-XTC / MHES14-XSC	Tiger SDR family	InfiniHost III Lx PCI Express x4 MemFree HCA Adapter Card.
MHES18-XT / MHES18-XS MHES18-XTC / MHES18-XSC	Cheetah SDR	InfiniHost III Lx PCI Express x8 MemFree HCA Adapter Card.
MHGS18-XT / MHGS18-XS MHGS18-XTC / MHGS18-XSC	Cheetah DDR	InfiniHost III Lx PCI Express x8 MemFree HCA Adapter Card.

The document consists of the following sections:

2 Major New Features

- Improved multi-core latency
- Added the ability to configure the PCI Express Ack Latency Timer via the firmware INI file. See the parameters **auto_ack_latency_timer_limit_8x** and **ack_latency_timer_limit_8x** in the `fw-25204-defaults.ref` file.
- Added the ability to configure the PCI Express SerDes via the firmware INI file. See the parameters **PCIe_OBPreAmp, PCIe_OBVoltage, PCIe_OBPreEmpPreAmp, PCIe_OBPreEmpOut,** and **PCIe_Equal** in the `fw-25204-defaults.ref` file.

3 Bug Fixes

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

Table 2 - Bug Fixes

	Issue	Description	Discovered in	Fixed in
1.	QUERY_DEBUG_MSG	msg_hdr_size is wrong (ID:38428)	1.1.000	1.2.000
2.	Suspend QP problem	Suspend QP may cause corruption of CmdIf (ID: 38325) and/or QP context corruption (ID: 38876)	1.1.000	1.2.000
3.	Wrong link state returned to GetPort-Info query on DDR Auto-negotiation	While a link is in the DDR Auto-negotiation process, it may report a POLLING state for the physical link and an INITIALIZE state for the logical link while it should report POLLING and Link-Down, respectively	1.1.000	1.2.000
4.	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)	1.1.000	1.2.000
5.	relaxed_ordering_en bit is not PCI Express compliant	PCI Express bit relaxed_ordering_en is R/W instead of Read Only	1.1.000	1.2.000
6.	Potential deadlock upon sending MSIX	Sending MSIX interrupt messages may cause a deadlock in a heavy-stress condition	1.1.000	1.2.000
7.	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)	1.1.000	1.2.000
8.	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)	1.1.000	1.2.000
9.	Fatal error in large-stress condition	FW may fail (FW_INTERNAL_ERROR) upon large stress (ID: 38845,39393)	1.1.000	1.2.000
10.	Low local timeout threshold impacts large clusters performance	The minimum acceptable Local_ACK_Timeout parameter was changed from 0xC to 0xD to reduce unnecessary timeouts	1.1.000	1.2.000

4 Invariant Sector (IS) Changes / Fixes

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

Table 3 - IS Changes / Fixes

Issue	Description	Discovered in	Fixed in
Failure upon access to Flash addresses higher or equal to 1MByte on single Flash device HCA cards	Newer versions of Tiger/Cheetah HCA cards have a single 2MByte Flash device. In case Debug FW is burnt onto this Flash device, access to addresses equal to or higher than 1MByte fail	1.1.000	1.2.000

5 Known Issues

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

Table 4 - Known Issues

Index	Issue	Description	Current Implemented Workaround in FW	Possible Workaround	Patch Release (fix)	Scheduled Release (fix)
1.	Possible boot failure off Debug FW on some MHES18-XTC Rev A2 cards	InfiniHost III Lx device may fail to boot from the Secondary image of Debug FW on some MHES18-XTC Rev A2 cards (PSID: MT_03B0120001)		Burn Debug FW again so the device will boot from the Primary image	NA	NA
2.	Possible WRITE_MTT failure	WRITE_MTT may fail if first MTT_INDEX is not even		WRITE_MTT to an even address	NA	NA
3.	IB performance counters	IB performance counters are not visible through GSI (ID: 28304)	NA	NA	NA	NA
4.	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 modifi-cation	NA	NA
5.	RTR2RTS_QPEE; SQD2RTS_QPEE: chang-ing 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 pro-vided.	Mask these optional fields	NA	NA
6.	PCI 2.3 control and status for interrupts	InfiniHost III Lx does not sup-port PCI2.3 control and status bits for interrupts.	NA	NA	NA	NA
7.	Change of memory bars on a disabled system	Changing size / addresses of memory bars between SYS_DIS and SYS_EN may cause the InfiniHost III Lx to hang (ID: 24206)	NA	NA	NA	NA
8.	BAR resizing on an enabled system	Changing bar sizes when a sys-tem is enabled may cause the InfiniHost III Lx to hang (ID: 24208).	NA	NA	NA	NA
9.	SW reset via configuration cycles	SW reset via configuration cycles may create double PCI Express completions for the configuration transaction.	NA	If InfiniHost III Lx boots in memory con-troller mode, perform power cycle / hot reset after restoring the flash	NA	NA

Table 4 - Known Issues (Continued)

Index	Issue	Description	Current Implemented Workaround in FW	Possible Workaround	Patch Release (fix)	Scheduled Release (fix)
10.	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
11.	Flash CRC error	InfiniHost III Lx fails to report a Flash CRC error.	NA	NA	NA	NA

5.1 Unsupported InfiniHost III Programmer's Reference Manual Features

• Power Management per the PCI Express Base Specification, Rev. 1.0a is not implemented.

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-25204-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-25204-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

Table 5 - History of Bug Fixes

	Issue	Description	Discovered in	Fixed in
1.	Wrong PSID in INI Files of MHES18-XS and MHGS18-XS	See "Correcting Wrong PSID in MHES18-XS.ini & MHGS18-XS.ini Files of fw-25204 Rev 1.0.700/1.0.800" in Rev 1.1.0 Release Notes	1.0.800/ 1.0.700	1.1.0/ 1.0.800/ 1.0.700
2.	CQ context corruption	May occur in a multiple CQ contexts environment. (ID: 37432-3)	1.0.800	1.1.0
3.	QP context corruption	May occur under stress. (ID:37213-4,36897,36209)	1.0.800	1.1.0
4.	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)	1.0.800	1.1.0
5.	CQ_ARM doorbell	May be lost under stress (ID:37133)	1.0.800	1.1.0
6.	EQ_ARM doorbell	May be lost. (ID: 37872)	1.0.800	1.1.0
7.	ICM context corruption	May occur when executing MAP_ICM command. (ID:37790)	1.0.800	1.1.0
8.	Send WQ corruption	May happen for POST_LIST with RDMA_READs. (ID:35649)	1.0.800	1.1.0
9.	RDMA_READ and ATOMIC Dead-Lock	Internal leak may occur when doing RDMA_READs and ATOMIC on Send Queues. (ID: 27761)	1.0.800	1.1.0
10.	E2E credits miscalculation	May occur upon OOS/RNR Nack (ID:37308/28)	1.0.800	1.1.0
11.	Multicast Deadlock	May occur in Receive-Side-errors environment. (ID:35186-9)	1.0.800	1.1.0
12.	**only_ddr_ib_support** (.ini parameter) is not supported	Fixed (ID: 35151)	1.0.800	1.1.0
13.	LinkSpeedEnable	LinkSpeedEnable is not applied correctly. (ID: 36592)	1.0.800	1.1.0
14.	Bringing IB port up on DDR device	DDR IB link is not going up after being put in port DISABLE state. (ID:36586)	1.0.800	1.1.0
15.	Wrong MAD status	MAD status can be "Invalid combination of Method and Attribute" instead of "Invalid Attribute". (ID: 35226)	1.0.800	1.1.0
16.	Wrong Client ReRegister bit	Fixed (ID:33958,35356)	1.0.800	1.1.0
17.	Vendor Specific MADs	Wrong Vendor Key (V_KEY) authentication and mismatch behavior. (ID:36127) Wrong Revision ID in General Info MAD. (ID: 38025)	1.0.800	1.1.0
18.	SUSPEND_QP	May cause: Command TimeOut and QP context corruption. (ID: 35109) Wrong e2e_credits calculation. (ID: 37608) In UD QP: may cause local QP error. (ID: 35524)	1.0.800	1.1.0
19.	Wrong APM behavior	Packet from wrong port are dropped even if Migration Request is set. Also, a bad path packet may be allowed in MIGRATED state. (ID: 35179) APM can cause a FW deadlock. (ID:37092)	1.0.800	1.1.0
20.	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)	1.0.800	1.1.0