MHEA28-XT & MHGA28-XT Memfree InfiniHost III Ex PCI Express x8 HCA Card

Board User's Manual
Rev 1.10

MHEA28-XT & MHGA28-XT InfiniHost III Ex PCI Express MemFree HCA Adapter Card Board User's Manual

Document Number: 2151BM

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

Table of Contents

About This Manual

This *Board User's Manual* describes the following InfiniHost III Ex PCI Express x8 MemFree HCA Cards:

- MHEA28-XT (Code Name: Lion Mini) supporting operation of IB ports at SDR (Single Data Rate)
- MHGA28-XT (Code Name: Lion Mini DDR) supporting operation of IB ports at DDR (Double Data Rate)

The manual provides details as to the interfaces of the boards, specifications, required software and firmware for operating the boards, and relevant documentation.

Intended Audience

This manual is intended for the installer and user of the MHEA28-XT or MHGA28-XT HCA cards.

The manual assumes basic familiarity with the InfiniBand™ architecture specification.

Related Documentation

Table 1 - Documents List

MHEA28-XT & MHGA28-XT InfiniHost III Ex PCI Express MemFree HCA Adapter Card Board User's Manual Document no. 2151BM	This manual.
MHEA28-XT & MHGA28-XT InfiniHost III Ex PCI Express MemFree HCA Card Quick Start Guide Document no. 2143UG	A quick start guide for helping in quick hardware installation, software and firmware installation, and usage of the adapter cards.
InfiniHost III Programmer's Reference Manual Document no. 2248PM	A reference describing the interface used by developers to write a driver for MemFree Mellanox InfiniHost III devices.
InfiniHost III Ex MT25208 Hardware Reference Manual Document no. 2246HM	Reference for hardware engineers responsible for designing systems and boards incorporating InfiniHost III Ex components.
InfiniBand Administration (IBADM) Package User's Manual Document no. 2130UM	User's Manual describing the utilities included in the IBADM tools package for system administration of an InfiniBand cluster.
Mellanox MST User's Manual Document no. 2125SM	This manual describes various tools and utilities, included in the Mellanox Software Tools (MST) package, for accessing, burning firmware, and tracing Mellanox silicon devices.
VAPI HCA Device Driver Release Notes Version 4.0.3 or later	InfiniHost III device driver release notes

1 Overview

The MHEA28-XT and MHGA28-XT are dual-port 4X InfiniBand PCI Express x8 Adapter Cards based on the InfiniHost III Ex MT25208 third generation InfiniBand Host Channel Adapter (HCA). Each 4X InfiniBand port on the MHEA28-XT operates at SDR (10Gb/s), whereas each 4X InfiniBand port on the MHGA28-XT operates at SDR (10Gb/s) or DDR (20Gb/s).

InfiniHost III Ex features include a PCI Express x8 interface, hardware transport, advanced per queue pair (QP) QoS services and support for millions of QPs and completion queues (CQs). The MHEA28-XT / MHGA28-XT host channel adapter is a PCI Express Revision 1.0a compatible card, with dual 4X InfiniBand compliant connectors for copper cables.

> Note: The MHEA28-XT HCA card (only) utilizes media detect circuitry enabling external fiber media adapter modules (such as the Emcore QTR3400), allowing for fiber channel communication across up to 300m (at SDR only).

The cards feature MemFree technology which eliminates the requirement for local memory on the PCI Express adapter cards. The cards can be inserted into PCI Express x8 or higher slots of standard server, blade server, storage, and communications platforms to enable InfiniBand system area networks. The InfiniHost III Ex core features a full hardware implementation of the InfiniBand architecture memory protection and translation tables, as well as hardware transport. This drastically reduces CPU overhead to enable the host processor to spend its cycles on applications and not on communications. This advanced third generation design achieves industry leading bandwidth performance coupled with low latency for technical compute clusters, data centers and storage applications.

1.1 Product Features

- PCI Express x8 version 1.0a compatible card
- Two Port InfiniBand Version 1.1 compatible Host Channel Adapter
- InfiniBand Compatible Verbs API interface for both Linux and Windows operating systems
- Dual 4X InfiniBand ports with standard copper connectors (each port runs at 10Gb/s for MHEA28-XT, and at 10Gb/s or 20Gb/s for MHGA28-XT)
- Hardware support for up to 16 million QPs, EEs and CQs
- Memory Protection and Translation Tables fully implemented in hardware
- IB Native layer 4 DMA hardware acceleration
- Multicast support
- Programmable MTU size from 256 to 2K bytes
- Eight Virtual Lanes supported plus Management Lane
- Support for InfiniBand transport mechanisms (UC, UD, RC, RAW)
- EEPROM used for storing Vital Product Data (VPD)

1.2 Key Mellanox Features

- Embedded InfiniRISC Processors for Management & Subnet Management Agent (SMA)
- Integrated Physical Layer SerDes

- Integrated GSA (General Service Agents)
- Low-Latency Communication Technology
- Flexible Completion Mechanism Support (Completion Queue, Event, or Polled operation)

1.3 Operating Systems Support

MHEA28-XT and MHGA28-XT can operate in Linux and Windows operating system environments. See “Firmware and Software” to obtain device drivers.

1.4 Firmware and Software

The MHEA28-XT and MHGA28-XT cards come with the latest version of the appropriate firmware available at the time of manufacturing. Firmware updates are periodically provided, and the most recent firmware can be obtained from the ‘Firmware Downloads’ page reached from www.mellanox.com. The firmware can be updated using the **ibf-wmg** tool of the IBADM package (see “Documents List” on page 5).

Both low profile HCA cards include verbs interface and basic device drivers for both Windows and Linux operating systems. In addition, the card includes an internal Subnet Management Agent (SMA) and General Service Agents, eliminating the requirement for an external management agent CPU. The HCA is fully compatible with the open source OpenIB software suite.

The OpenIB software is available through the InfiniBand Gold Distribution which includes software for database clustering, high performance computing, communications, and storage applications. This collection consists of drivers, protocols, and management applications in a simple to install package. To download the package go to www.mellanox.com. The board is also supported by open source applications and management package. Complete initialization, diagnostic and management utilities are also provided to facilitate quick system bring up.

Alternatively, it is possible to download the HCA driver and other software from *Products / HCA Board Products / MemFree InfiniHost III Ex PCI Express Adapter Card* as well as from *Code Releases / InfiniHost III Ex /* on the DDS at www.mellanox.com.

1.5 MHEA28-XT / MHGA28-XT Contents

- MHEA28-XT / MHGA28-XT PCI Express HCA Adapter Card

2 MHEA28-XT / MHGA28-XT Board

2.1 I/O Interfaces

The MHEA28-XT / MHGA28-XT board includes the following interfaces:

- Two 4X InfiniBand Copper Connectors
- PCI Express x8 Edge Connector
- I/O Panel LEDs
- I^2C compatible connector (for debug)

2.1.1 InfiniBand Interface

The MHEA28-XT / MHGA28-XT board provides two 4X InfiniBand v1.1 4X connectors for external InfiniBand cables. The MT25208 device is compliant with the IBTA specification 1.1. The MT25208 silicon device has two 4X ports A and B, which connect to 4X connectors 1 and 2 of the MHEA28-XT / MHGA28-XT, respectively. The standard InfiniBand connectors include a "media detect circuit" that supports external InfiniBand fiber solutions such as the Emcore Model QTR3400 Smart Module.

2.1.2 PCI Express Interface

The PCI Express interface is version 1.0a compatible. The MT25208 can be either a root complex initiating the PCI Express bus operations or a downstream device responding to PCI bus operations. The PCI Express bus can connect to either a host CPU in an HCA application or to an I/O device (such as Gigabit Ethernet) when used as a Target Channel Adapter.

2.1.2.1 PCI Express Board Features

- Low profile, short PCI Express 8x expansion board (approximately 2.5 x 4.6 inches) with auto-negotiation features.

2.1.3 LED Assignment

The board has four LEDs located on the I/O panel - 2 LEDs per 4X port. The physical link (green) illuminates once VAPI (InfiniBand Verbs API) is started and a physical connection is made between two nodes. The data activity link (yellow) illuminates once the InfiniBand network is discovered over the physical link. The activity link is a steady yellow when it is discovered but no data is being passed. The activity link blinks when data is being passed. If the LEDs are not active, either the physical or the logical (or both) connections have not been established. See Figure 2, "I/O Panel with Dual Ports and LEDs," on page 13.

Table 1 - LEDs

Port 1	LED Name
	Physical Link - Green
	Logical Link - Yellow
Port 2	**LED Name**
	Physical Link - Green

Table 1 - LEDs (Continued)

Port 1	LED Name
	Logical Link - Yellow

2.1.4 I2C Interface

A three pin header is provided as the I2C interface. Please refer to the *Mellanox MST Tools User's Manual* (see "Documents List" on page 5) for usage.

2.2 Power

The MHEA28-XT / MHGA28-XT board receives 3.3v and 12v power from the PCI Express Edge connector. All other required power is generated by on-board switch mode regulators.

2.3 Memory

The MHEA28-XT /MHGA28-XT supports multiple memory devices through the PCI Express, Flash, and I2C interfaces.

2.3.1 System Memory

The MHEA28-XT / MHGA28-XT utilizes the PCI Express interface to store and access IB fabric connection information on the system memory.

2.3.2 Flash ROM

The MHEA28-XT / MHGA28-XT board supports 4MB of Flash ROM space via the CPU interface of the MT25208 InfiniHost III Ex device. The details of the Flash supported are:

- 8 bit data
- 22 bit address
- Programming via the CPU bus

Intel's TE28F320J3C-110 Flash device used on the MHEA28-XT / MHGA28-XT board has the following characteristics:

- 32 Mb Uniform (4MB)
- 110 ns access time

2.3.3 EEPROM

The board incorporates an EEPROM that is accessible through an I2C interface. The EEPROM is used for storing the Vital Product Data (VPD). VPD format adheres to the *PCI Local Bus specification rev 2.3* VPD definition. EEPROM capacity is 512 bytes.

2.4 Jumper Configuration

Table 2 - Jumper Configuration

Ref. #	Name	Description	Option	Default Configuration	Comments
J4	Flash	Flash present/ not present.	No Shorting Block (Default): Flash Present Shorting Block present: Flash not present 	No Shorting Block	Header 1x2

Note: See Figure 1, "PCI Express Expansion Board," on page 12 for the Jumper location.

3 Mechanical and Other Information

3.1 MHEA28-XT Board Diagram

Note: The MHGA28-XT has exactly the same footprint of MHEA28-XT, therefore its diagram is not included.

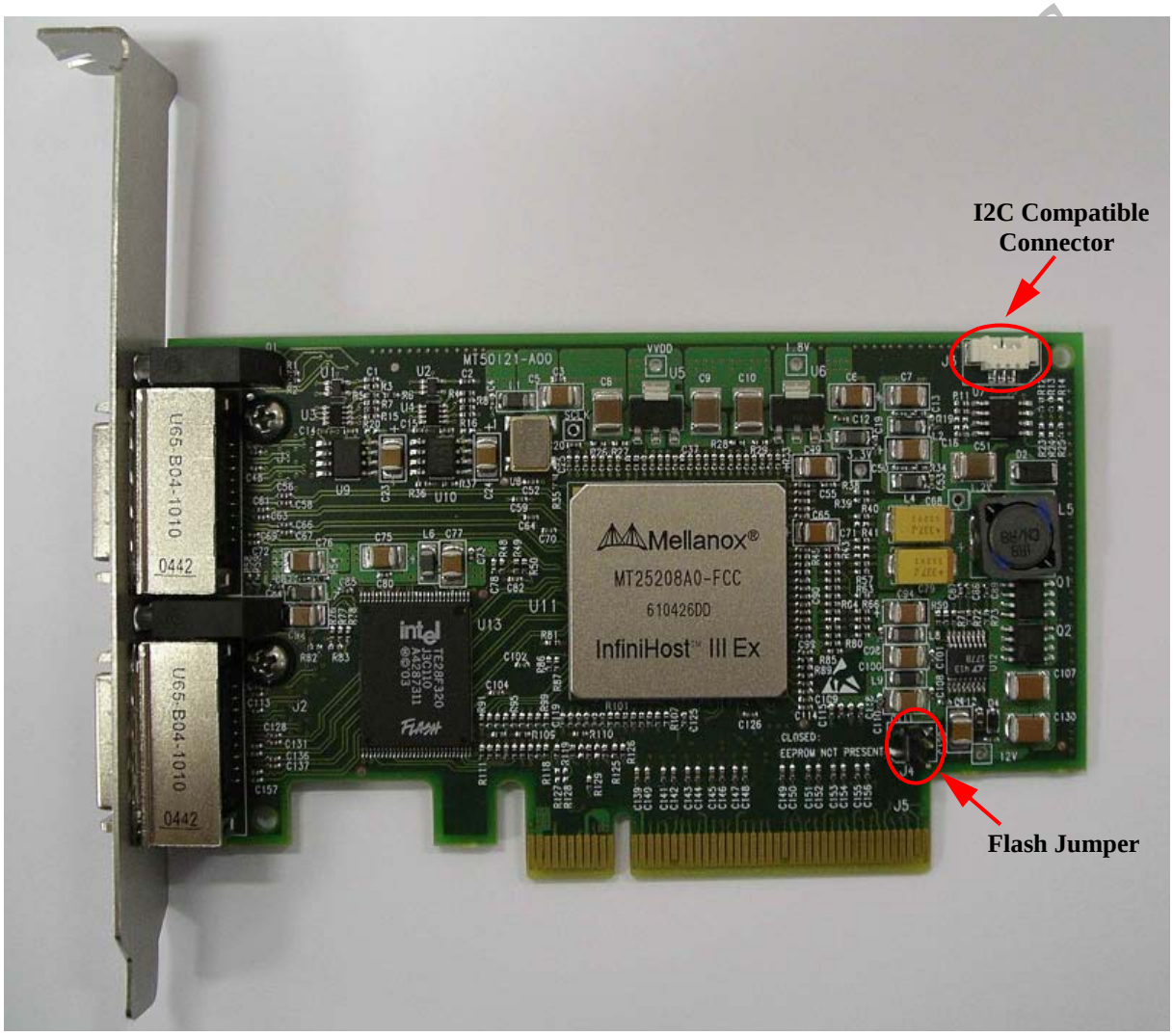

Figure 1: PCI Express Expansion Board

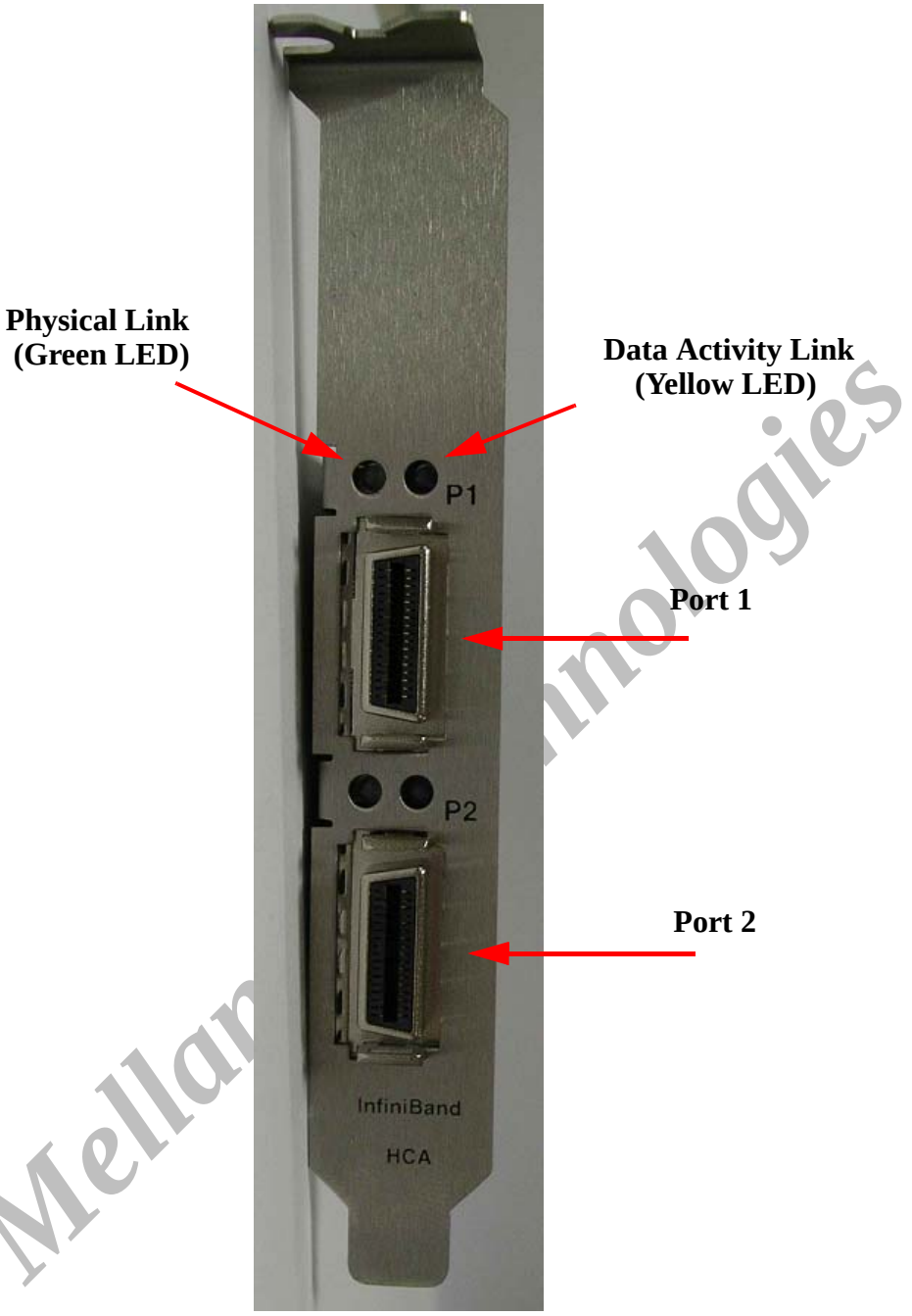

Figure 2: I/O Panel with Dual Ports and LEDs

3.2 Air Flow Requirement

Per the PCI specification, the MHEA28-XT / MHGA28-XT board needs to be provided with 200 LFM at 55C ambient temperature.

4 Vital Product Data (VPD) Format

This chapter includes the following sections:

4.1 VPD for MHEA28-XT (Lion Mini)

Table 3, "VPD Format for MHEA28-XT" shows Mellanox PCI VPD (Vital Product Data) format details for the MHEA28-XT HCA card. This definition complies with the format defined in the *PCI 2.3 Specification, Appendix I*.

Table 3 - VPD Format for MHEA28-XT

Offset (Decimal)	Item	Value	Format	Description
0	Large Resource Type ID String Tag (0x02)	0x82		
1	Length	0x9		
3	Data	'Lion Mini'	Alphanumeric	
12	Large Resource Type VPD-R Tag (0x10)	0x90		
13	Length	0x51		
15	VPD Keyword	'PN'	Numbers	Add-in Card Part Number
17	Length	0x4		
18	Data	'MHEL-C-T'		
39	VPD Keyword	'EC'	Alphanumeric	Engineering change level of the card (Rev)
41	Length	0x4		
42	Data	'A-02'		PCB Revision
46	VPD Keyword	'SN'	Alphanumeric	Serial Number
48	Length	0x18		
49	Data	'MTYYWWFXXXXX'		
73	VPD Keyword	'V0'		Misc. Information
75	Length	0x10		
76	Data	'PCI Ex x8'		
92	VPD Keyword	'RV'		
94	Length	0x1		
95	Data	Checksum		
96	Large Resource Type VPD-W Tag (0x11)	0x91		
97	Length	0x9C		

Table 3 - VPD Format for MHEA28-XT (Continued)

Offset (Decimal)	Item	Value	Format	Description
99	VPD Keyword	'V1'		EFI Driver Version
101	Length	0x6		
102	Data	'N/A'	Number	
108	VPD Keyword	'YA'		Asset Tag
110	Length	0x20		
111	Data	'N/A'	Alphanumeric	'N/A'
143	VPD Keyword	'RW'		Remaining read/write area
145	Length	0x6d		
146	Data	Reserved (0x00)		
255	Small Resource Type END Tag (0x11)	0x78		
256	Mellanox Read Only Mask	0x0...0	Numbers	
355	Mellanox Read/Write Mask	0x1..1	Numbers	
511	Mellanox Read Only Mask	0x0	Numbers	

4.2 VPD for MHGA28-XT (Lion Mini DDR)

Table 4, “VPD Format for MHGA28-XT” shows Mellanox PCI VPD (Vital Product Data) format details for the MHGA28-XT HCA card. This definition complies with the format defined in the *PCI 2.3 Specification, Appendix I.*

Table 4 - VPD Format for MHGA28-XT

Offset (Decimal)	Item	Value	Format	Description
0	Large Resource Type ID String Tag (0x02)	0x82		
1	Length	0xD		
3	Data	“Lion mini DDR”	Alphanumeric	
16	Large Resource Type VPD-R Tag (0x10)	0x90		
17	Length	0x4F		
19	VPD Keyword	“PN”	Numbers	Add in Card Part Number
21	Length	0x15		
22	Data	“MHGA28-XT”		
43	VPD Keyword	“EC”	Alphanumeric	Engineering Change Level of the card (rev)
45	Length	0x2		
46	Data	"A1		"PCB revision
48	VPD Keyword	“SN”	Alphanumeric	Serial Number
50	Length	0x18		
51	Data	“00..00XXXX..XX”		
75	VPD Keyword	“V0”		Misc Information
77	Length	0x10		
78	Data	“PCI EX x8”		
94	VPD Keyword	“RV”		
96	Length	0x1		
97	Data	Checksum		
98	Large Resource Type VPD-W Tag (0x11)	0x91		
99	Length	0x9A		
101	VPD Keyword	“V1”		EFI Driver version
103	Length	0x6		
104	Data	“N/A”	Number	
110	VPD Keyword	“YA”		Asset Tag
112	Length	0x20		
113	Data	“N/A”	Alphanumeric	“N/A”
145	VPD Keyword	“RW”		Remaining read/write area
147	Length	0x6b		

Table 4 - VPD Format for MHGA28-XT (Continued)

Offset (Decimal)	Item	Value	Format	Description
148	Data	Reserved (0x00)		
255	Small Resource Type END Tag (0x11)	0x78		
256	Mellanox Read Only Mask	0x0....0	Numbers	
357	Mellanox Read/Write Mask	0x1...1	Numbers	
511	Mellanox Read Only Mask	0x0	Numbers	

Appendix A: Specifications

A.1 MHEA28-XT Specifications

Physical		**Power and Environmental**	
Size: Air Flow: 10Gb/s Connector:	117mm x 69mm 200LFM @55C InfiniBand MicroGigaCN (Copper)/ Emcore (Fiber Channel)	Voltage: Maximum Power: Temperature:	12V, 3.3V 8.2W 0 to 55 Celsius
Protocol Support		**Regulatory**	
InfiniBand: QoS: RDMA Support:	Auto-Negotiation 10Gb/s, 2.5Gb/s 8 InfiniBand Virtual Lanes for all ports Yes, All Ports		

A.2 MHGA28-XT Specifications

Physical		**Power and Environmental**	
Size: Air Flow: 20Gb/s Connector:	117mm x 69mm 200LFM @55C InfiniBand MicroGigaCN (Copper)/ Emcore (Fiber Channel)	Voltage: Maximum Power: Temperature:	12V, 3.3V 11.4W 0 to 55 Celsius
Protocol Support		**Regulatory**	
InfiniBand: QoS: RDMA Support:	Auto-Negotiation (10Gb/s, 2.5Gb/s) or (20Gb/s, 5Gb/s) 8 InfiniBand Virtual Lanes for all ports Yes, All Ports		

Revision History

Table 1 - Revision History

Date	Rev	Comments/Changes
July 2005	1.10	Added the MHGA28-XT HCA board part number supporting IB DDR operation, added its specifications in Appendix A: "Specifications," on page 19 and updated the power number for MHEA28-XT.
Dec. 2004	1.00	Created the document