



Connect. Accelerate. Outperform.™

Mellanox OFED for FreeBSD Release Notes

Rev 2.1.6

NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT (“PRODUCT(S)”) AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES “AS-IS” WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER’S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCT(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



Mellanox Technologies
 350 Oakmead Parkway Suite 100
 Sunnyvale, CA 94085
 U.S.A.
www.mellanox.com
 Tel: (408) 970-3400
 Fax: (408) 970-3403

© Copyright 2015. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, Connect-IB®, CoolBox®, CORE-Direct®, GPUDirect®, InfiniBridge®, InfiniHost®, InfiniScale®, Kotura®, Kotura logo, MetroX®, MLNX-OS®, PhyX®, ScalableHPC®, SwitchX®, TestX®, UFM®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo are registered trademarks of Mellanox Technologies, Ltd.

CyPU™, ExtendX™, FabricIT™, FPGADirect™, HPC-X™, Mellanox Care™, Mellanox CloudX™, Mellanox Open Ethernet™, Mellanox PeerDirect™, Mellanox Virtual Modular Switch™, MetroDX™, NVMeDirect™, StPU™, Switch-IB™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Table of Contents

Table of Contents	3
List of Tables	4
Chapter 1 Introduction	5
1.1 Main Features in This Release	5
1.2 Supported Platforms and Operating Systems	5
1.3 Unsupported Functionality/Features	5
1.4 Supported HCAs	5
Chapter 2 Changes and Fixes in Rev 2.1.6	6
Chapter 3 Known Issues	7
Chapter 4 Bug Fixes History	8

List of Tables

Table 1:	Supported Platforms and Operating Systems	5
Table 2:	Changes and Fixes in Rev 2.1.6	6
Table 3:	Known Issues.....	7
Table 4:	Fixed Bugs List	8
Table 5:	Change Log History.....	9

1 Introduction

These are the release notes for Mellanox Technologies' driver for FreeBSD Rev 2.1.6 driver kit for Mellanox adapter cards supporting the following uplinks to servers:

- InfiniBand: QDR, FDR10, FDR
- Ethernet: 10GigE, 40GigE

1.1 Main Features in This Release

- Packet Pacing, also known as “rate limit” support
Note: Packet Pacing requires firmware version 2.34.8110. The firmware is part of Packet Pacing: Supporting Package.
- IP based addressing for RoCE
- EEPROM cable information reader

1.2 Supported Platforms and Operating Systems

The following are the supported OSs in Mellanox OFED for FreeBSD Rev 2.1.6:

Table 1 - Supported Platforms and Operating Systems

Operating System	Platform
FreeBSD 10.1	AMD64



It is strongly recommended to use the 10-stable version for user space.

1.3 Unsupported Functionality/Features

The following are the unsupported functionalities/features in FreeBSD Rev 2.1.6:

- ConnectX®-2 Adapter cards

1.4 Supported HCAs

Mellanox OFED for FreeBSD Rev 2.1.6 supports the following Mellanox network adapter cards:

- ConnectX®-3/ConnectX®-3 Pro Rev 2.34.5000

For official firmware versions, please see:

http://www.mellanox.com/content/pages.php?pg=firmware_download

2 Changes and Fixes in Rev 2.1.6

Table 2 - Changes and Fixes in Rev 2.1.6

Category	Description
RoCE	Added support for IP based addressing
Ethernet	Removed “Oversized header or SG list” warning, and added a counter which counts the amount of this events.
	Fixed the shown stats in sysctl: tx bytes and rx_packets are now showing the correct values.
Cables	EEPROM cable information reader.
Packet Pacing	Packet Pacing at GA level.

3 Known Issues

The following is a list of general limitations and known issues of the various components of this Mellanox OFED for FreeBSD release.

Table 3 - Known Issues

Index	Category	Description	Workaround
1.	VLAN	On rare occasions kernel panic might occur when removing VLAN interface while running multiple traffic threads over it.	Stop traffic over the relevant VLAN interface before removing it.
2.	Arch	Occasionally fails to load mlx4_core module on servers that have six cards or more.	-
3.	RoCE	No support for UD mode.	-
4.		Since the number of GIDs per port is limited, there cannot be more than the allowed IP addresses configured to Ethernet devices that are associated with the port. The allowed number is "124" for a single function machine.	-
5.		No support for multicast IPv6 based addressing	
6.	Packet Pacing	Closing multiple Packet Pacing sockets at once takes a relatively long time.	-
7.		When the driver is compiled with Packet Pacing, IB modules will not be supported.	-
8.	Performance	Kernel panic might occur when setting rx_size value during heavy traffic.	-
9.		Interface numbering increments after changing link type.	-
10.	Interface Presentation	Small disturbance in traffic rate of port one while the other port on the same card is deliberately going up and down repeatedly. When the second port stops changing, the traffic stabilizes.	
11.		Incorrect supported link speed information when using "ifconfig -m <interface>" for 10GbE adapter cards. The cards show 40GbE instead of 10GbE.	

4 Bug Fixes History

Table 4 lists the bugs fixed in this release.

Table 4 - Fixed Bugs List

#	Description	Discovered in Release	Fixed in Release
1.	Fixed memory leaks upon InfiniBand modules unload.	FreeBSD-10.0 inbox driver	2.1.5
2.	Removed debug prints from a not debug kernel	2.1	2.1.5
3.	Fixed an issue of very low TCP/UDP traffic performance when using minimal MTU.	2.1	2.1.5
4.	Fixed an issue of ping without fragmentation which does not pass when MTU=6552	2.1	2.1.5
5.	Fixed kernel panic occurring when auto loading mlnx modules on a setup with multiple HCAs	2.1	2.1.5
6.	Fixed memory leaks upon driver unload.	FreeBSD-10.0 inbox driver	2.1
7.	Adjusted driver boot up sequence from being loaded last to a specific boot up order for OFED modules.	FreeBSD-10.0 inbox driver	2.1
8.	Fixed an issue preventing a static LAG configuration over Mellanox interfaces.	FreeBSD-10.0 inbox driver	2.1
9.	Fixed traffic balancer behavior when using LAG LACP mode.	FreeBSD-10.0 inbox driver	2.1
10.	Fixed kernel panic occurring when using 'iftop' utility while running multiple traffic threads.	FreeBSD-10.0 inbox driver	2.1

Change Log History

Table 5 - Change Log History

Release	Category	Description
2.1.5	Performance	Enlarged the number of RX rings for better spread of incoming traffic
		Added a tuning process for dual NUMA architecture
	Memory	Resolved memory leaks in InfiniBand modules
	Multicast	Added support for Multicast with RoCE
	VLAN	Added support for VLAN with RoCE
	Packet Pacing (Beta level)	Packet pacing, also known as “rate limit”, defines a maximum bandwidth allowed for a TCP connection. Limitation is done by Hardware, where each QP (transmit queue) has a rate limit value from which it calculates the delay between each packet sent.
RoCE (RDMA over Converged Ethernet)	RoCE allows InfiniBand (IB) transport applications to work over an Ethernet network. RoCE is enabled only for drivers that support VPI.	
2.1	Ethernet	<ul style="list-style-type: none"> • VLAN TX/RX offloads insertion/stripping Received VLAN traffic will be stripped from the VLAN tag by the hardware. • Interrupt moderation By default, the driver uses adaptive interrupt moderation for the receive path, which adjusts the moderation time to the traffic pattern. • Statistics/Extended counters. Counters are used to provide information about how well an operation system, an application, a service or a driver is performing. • RX offloads: <ul style="list-style-type: none"> • Large Receive Offload (LRO) for IPv4 and IPv6 - Increasing inbound throughput of high bandwidth network connections by reducing CPU overhead. • Hardware checksum - Support the Receive Checksum Offload mode. • RSS (Receive Side Scaling) - The RSS hash function distributes received traffic throughout RX rings. • TX offloads: <ul style="list-style-type: none"> • TCP Segmentation offload (TSO) for IPv4 and IPv6 - increase outbound throughput by reducing CPU overhead. • Hardware checksum - Support the Outgoing Checksum Offload mode. • Inline - Using inline data. • Blue Flame - Added Blue Flame support. • Promiscuous mode - multicast included. • PPP - Point to Point protocol • Configuration modification - Support modification of MAC address, MTU size and RX/TX queues size.

Table 5 - Change Log History

Release	Category	Description
2.1 (cont.)	Performance	Added adaptive moderation improvements
		Added 40GigE Out-Of-Box improvements
		Added IPv6 support for LRO, TSO
		Added general performance improvements
		Incremented the number of used TX and RX queues
	Multicast Filtering	Added exact match multicast filtering
	Driver Load	Enabled link bring up upon driver load
		Enabled VPI ports boot as Ethernet by default