



Release Notes

IPoIB Driver

Rev 1.7.0

Mellanox Technologies

© Copyright 2005. Mellanox Technologies, Inc. All Rights Reserved.

IPoIB Driver 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 IPoIB (IP over IB) driver Rev 1.7.0. The IPoIB driver enables the transmission of IP packets over an InfiniBand fabric.

Note: For supported platforms and operating systems, please refer to the *Mellanox IB Gold Distribution Release Notes*.

This document includes the following sections:

- “Software Dependencies” (page 3)
- “Major New Features” (page 3)
- “Known Issues And Unsupported Features” (page 4)
- “Fixed Bugs” (page 4)
- “Main Verification Flows” (page 5)

2 SOFTWARE DEPENDENCIES

The IPoIB driver depends on the installation of the IB Gold Distribution stack with OpenSM running.

3 MAJOR NEW FEATURES

- This IPoIB driver is based on the following IETF drafts:
 - draft-ietf-ipoib-architecture-04.txt
 - draft-ietf-ipoib-ip-over-infiniband-07.txt

Please refer to these documents (www.ietf.org) for features of IPoIB.

3.1 New Features

1. `ipoib_max_address_vector`: Defines the maximum number of addresses the HCA device will be able to create. This number reflect the maximum number of destination ports the IPoIB driver can possibly connect to. The default value is 2048.
2. `ipoib_tx_ring_size`: Defines the maximum number of DoorBells allowed by the IPoIB driver in queue. `txqueuelen` will be twice this number. The default value is 128.
3. Performance enhancement: Running the NetPerf benchmark once with the previous (1.6.0) release of IPoIB and once with this new 1.7.0 release showed a 20% improvement in bandwidth with the new release. Following are the details of the system used: CPU: Dual Xeon @3.4GHz (Intel), OS: SuSE Linux 9.1, Kernel: 2.6; and an MHEL-CF128-T InfiniHost III Ex Adapter card.

4 KNOWN ISSUES AND UNSUPPORTED FEATURES

The following table lists the compliancy features not supported in this release.

Table 1 - Known Issues And Unsupported Compliancy Features

Issue/Limitation	Compliance	Description
If ib0 or ib1 is brought down and then brought up again, the IPoIB 'Mac' address will not be the same		Note: This issue does not apply in case of the command: <code>'/etc/init.d/openibd restart'</code> If you execute the following sequence (for ib0 or ib1): <pre>> ifconfig ib0 down > ifconfig ib0 <ip addr></pre> then the IPoIB 'Mac' address after the new bring-up will differ from the one established on the first bring-up. Consequently, the previously established connectivity to ports on other device nodes will be lost. There are two ways for a device 'on the other side' to re-establish the connectivity: 1. Automatic: A kernel timeout will clear its cache entry where the old connectivity was defined. 2. Manual: By running the command: <code>arp -d <ip_addr></code>
No IPv6 support		The IPoIB driver does not support IPv6
The IPoIB driver does not support the flag "-b" in the tcp benchmark		
Ifconfig does not give the right MAC address on kernel 2.6		On kernel 2.6, the IPoIB driver uses 20-byte MAC addresses. This is not supported by ifconfig. In order to obtain the MAC address, it is possible to use the "ip addr" command
DHCP client does not work on kernel 2.6		
For kernel 2.4 only: IP Multicast addresses that differ only in bits [27:23] are not supported		In case of IP multicast aliasing, the Linux kernel does not notify the IPoIB driver of any update to the MAC multicast filter. Therefore, if such multicast addresses are used, there is a chance that the IPoIB driver will not add all of them to the IPoIB multicast groups.

5 FIXED BUGS

The following table lists the fixed bugs in this release.

Table 2 - Fixed Bugs

Issue	Description	Found in version
The IPoIB driver does not delete AHs on kernel 2.6	The Driver now deletes the AH whenever the entry is deleted from the ARP cache	1.6.0
Kernel oops occurring after reboot	If the kernel was compiled with the spinlock debug feature enabled, a kernel oops could occur after reboot	1.6.0
IP Multicast addresses that differ only in bits [27:23] are not supported (except for kernel 2.4)	See description in Section 4 above. This problem has been fixed for all supported kernels except 2.4.	1.6.0

6 MAIN VERIFICATION FLOWS

In order to verify the correctness of the IPoIB initiator the following tests were run.

Table 3 - Verification Tests

Test	Description/Comments	Flags
TTCP	<i>Client-server 1 to 1</i> test. Data transfers in different buffer sizes.	-l,-n
Netperf	<i>Client-server 1 to many</i> test. Data transfer in different buffer sizes. A server may serve many clients.	-n,-m
Netpipe	<i>Client-server 1 to 1</i> test. Data transfers in many different buffer sizes. Connect/Disconnect tests.	-I, -2
Iperf	<i>Client-server 1 to many</i> test. Also a bi-directional test. A test can open many clients.	-P, -l,-d
Apache WebServer	http server	
lftp	http client	
Wget	http client	
Apache Benchmark	http client	
ftp	ftp protocol	
inetd services	telnet, ssh, vsftpd	
ping	With various packet sizes	
Linux ltp tests http://prdownloads.sourceforge.net/ltp/ltp-full-20040603.tgz	All Linux network orders	
IPoIB Infiniband Plugfest test http://www.infinibandta.org/	IPoIB tests from Plugfest	
Tests written by Mellanox: - Big cluster emulating test - Stress tests	These tests exercise sending packets at their maximum send rate and sending huge messages	
Bad flow verification tests	Inject changes during run-time: Change of OpenSM, change of IB Link status, CTRL-C, CTRL-Z, removal of a module	

Mellanox Technologies