



Connect. Accelerate. Outperform.™

Mellanox ConnectX®-3 Poll-Mode Driver

Release Notes

Rev 2.7.4

www.mellanox.com

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 NONINFRINGEMENT 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

Mellanox Technologies, Ltd.
Beit Mellanox
PO Box 586 Yokneam 20692
Israel
www.mellanox.com
Tel: +972 (0)74 723 7200
Fax: +972 (0)4 959 3245

© Copyright 2014. Mellanox Technologies. All Rights Reserved.

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

ExtendX™, FabricIT™, HPC-X™, Mellanox Open Ethernet™, Mellanox PeerDirect™, Mellanox Virtual Modular Switch™, MetroDX™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Contents

- 1 DPDK Poll-Mode Driver (PMD) Overview 5**
 - 1.1 System Requirements 5
- 2 Changes and Major New Features..... 6**
- 3 Known Issues..... 7**
- 4 Bug Fixes 9**

List of Tables

Table 1: System Requirements	5
Table 2: Changes and Major New Features	6
Table 3: PMD Known Issues	7
Table 4: PMD Bug Fixes	9

1 DPDK Poll-Mode Driver (PMD) Overview

librte_pmd_mlx4 is the DPDK Poll-Mode Driver (PMD) for Mellanox ConnectX®-3 Ethernet adapters. This driver is based on libibverbs and currently supports:

- Hardware TX and RX checksum offloading
- Scattering/gathering RX/TX packets
- Multiple RX (with RSS/RCA) and TX queues
- Multiple MAC addresses
- VLAN filtering
- Link state information
- Software counters/statistics
- Start/stop/close operations
- Multiple physical ports host adapter
- DPDK 1.6.0 or above from dpdk.org <<http://dpdk.org/>>

1.1 System Requirements

Table 1: System Requirements

Specification	Value
Network Adapter Cards	Mellanox ConnectX®-3 and ConnectX®-3 Pro network adapter card. (This must be configured to work in ETH mode.)
Firmware	2.33.5000
Driver Stack	MLNX_OFED-2.4-1.0.0
Operating Systems and Kernels	All OSs supported by MLNX_OFED
# CPU	Minimum 4 CPUs for VM
Minimum memory requirements	16 GB RAM
Transport	Ethernet
CPU Arch	x86 Power8

2 Changes and Major New Features

Table 2: Changes and Major New Features

Release	Changes
Rev 2.7.4	<ul style="list-style-type: none"> • Added support for: <ul style="list-style-type: none"> • DPDK 1.8 • MLNX_OFED 2.4-1.0.0 • Performance improvements for: <ul style="list-style-type: none"> • RX and TX • SR-IOV with RECV INLINE and optimized steering mode • Added support for TX and RX checksum offloading • VXLAN traffic is received and steered between PMD's RSS queues • Removed support for DPDK 1.6 • Bug fixes (see section Bug Fixes)
Rev 2.6.9	<ul style="list-style-type: none"> • Added support for MLNX_OFED 2.3-2.0.0 • Changed compilation instructions. For further information, please refer to the Quick Start Guide. • Bug fixes (see section Bug Fixes)
Rev 2.6.8	<ul style="list-style-type: none"> • Added support for: <ul style="list-style-type: none"> • DPDK 1.7 • MLNX_OFED 2.3-X-X-X • Power8 (with DPDK 1.6)Performance improvements for: <ul style="list-style-type: none"> • Promiscuous mode • RSS • TX • Added support for sending large packets (8192< packet size < 16384) • Added auto detection of Virtual Function (VF) devices – removed the need to use MLX4_NO_SRIOV=1 for bare metal environment • Added 4 new functions to device capabilities : <ul style="list-style-type: none"> • mtu_get & mtu_set • flow_ctrl_get, & flow_ctrl_set • Added bring up kernel interface during initialization stage • Added support for Multicast Promiscuous mode on SR-IOV environment • Added new option to compilation script: -t <RTE_TARGET> • Added new options to configuration script: -p <PAGE SIZE>, -n <HUGE PAGES NUM> • Removed support for: <ul style="list-style-type: none"> • VMware • DPDK version < 1.6 • Bug Fixes (see section Bug Fixes)
Rev 2.5	<ul style="list-style-type: none"> • Added support for VMware SR-IOV with PMD for DPDK running in guest VM • Added support for DPDK 1.5 and 1.6 from dpdk.org

Release	Changes
	<ul style="list-style-type: none"> Performance improvements Bug fixes (see section Bug Fixes)
Rev 2.0	<ul style="list-style-type: none"> Added support for KVM SR-IOV with PMD for DPDK running in guest VM Added support for Multiple NICs on the same host Bug fixes (see section Bug Fixes)
Rev 1.5	<ul style="list-style-type: none"> Added support for Multi cores Added support for RSS Added support for ConnectX®-3 Pro
Rev 1.0	<ul style="list-style-type: none"> Added support for Single core Added support for ConnectX®-3

3 Known Issues

The following table describes known issues in Mellanox PMD, and existing workarounds.

Table 3: PMD Known Issues

Subject	Description	Workaround
InfiniBand	Mellanox PMD does not support InfiniBand.	N/A
RSS	RSS hash key and options cannot be modified.	N/A
Hardware counters	Hardware counters are not implemented.	N/A
Promiscuous and SR-IOV	Promiscuous mode does not work when SR-IOV is enabled.	N/A
Promiscuous and SR-IOV	In testpmd: Promiscuous mode failed on SR-IOV but still showed as enabled	N/A
Multicast Promiscuous mode and SR-IOV	Virtual Machines (VM) with enabled Multicast Promiscuous mode do not receive the multicast traffic that other VMs on the same hypervisor are registered to.	N/A
Optimized steering mode	When optimized steering mode is enabled, the packet with DST MAC as port's MAC will not be received. In addition, when promiscuous mode is enabled, the packet with DST MAC as port's MAC will not be received	<ul style="list-style-type: none"> Add "DPDK" MAC using the <code>rte_eth_dev_mac_addr_add DPDK API</code> In case of Bare Metal, use the promiscuous mode to receive packets with DST MAC that is not the port's MAC.
	When optimized steering mode is	<ul style="list-style-type: none"> Add "DPDK" MAC using

Subject	Description	Workaround
	<p>enabled, the error messages below will be reported in dmesg:</p> <pre>mlx4_core 0000:21:00.0: command QP_FLOW_STEERING_ATTACH (0x65) failed: in_param=0xbfd1ba000, in_mod=0x10, op_mod=0x0, fw status = 0x4 <mlx4_ib> __mlx4_ib_create_flow: Device managed flow steering is disabled. Fail to register network rule..</pre> <p>The following message will be received from PMD:</p> <pre>"Cannot configure normal flow. If optimized steering is enabled (options mlx4_core log_num_mgm_entry_size=-7) , please check RN and QSG for more information"</pre>	<p>the <code>rte_eth_dev_mac_addr_add</code> DPDK API</p> <ul style="list-style-type: none"> In case of Bare Metal, use the promiscuous mode to receive packets with DST MAC that is not the port's MAC.
Number of configured RSS queues must be power of 2	Number of configured RSS queues must be power of 2, if configured differently, the number of queues that are used is the smaller number that is power of 2	Use only power of 2 RSS queues
libibverbs warning	<p>The following warning appears when running DPDK application with ConnectX-3 PMD</p> <pre>libibverbs: Warning: couldn't open config directory '/usr/local/etc/libibverbs .d'</pre>	Ignore this warning
Multiple RX VLAN filters	Although multiple RX VLAN filters can be configured, only the first one functions properly.	Future Release
RX VLAN filter and optimized steering mode	RX VLAN filter does not function if the optimized steering mode is enabled (<code>log_num_mgm_entry_size=-7</code>)	Configure default steering mode (<code>log_num_mgm_entry_size=-1</code>)
1Gb/s port's link reported as 10Gb/s	When a 10Gb/s port is set as 1Gb/s, it reports a wrong speed, (10Gb/s instead of 1Gb/s)	Use <code>ethtool</code> to retrieve link speed

4 Bug Fixes

The following table describes the bug fixes of Mellanox PMD.

Table 4: PMD Bug Fixes

Subject	Description	Found in	Fixed in
rte_eth_dev_flow_ctrl_set cannot be used to disable flow ctrl	When trying to disable flow ctrl using <code>rte_eth_dev_flow_ctrl_set</code> it is always set to on	2.6.8	2.7.4
RX VLAN filters	RX VLAN filters do not function with PMD 2.6.8	2.5	2.7.4
Kernel crush with RSS application	Kernel crush might occur when a DPDK application which uses multiple queues (RSS) is shut down, (kill -9 or ^C)	2.6.9	2.7.4
ibv_register failures and local protection errors	When using function <code>rte_mempool_xmem_create(.., void *vaddr, ...)</code> to allocate mempool, if <code>vaddr != NULL</code> , mempool struct and data blocks were separated causing <code>ibv_reg_mr</code> failure and bad completion status on send and receive messages.	2.6.8	2.6.9
Memory corruptions	Fixed memory corruptions errors that happened after queues reconfiguration	2.5	2.6.8
Ubuntu	Fixed compilation on Ubuntu	2.5	2.6.8
VST mode on SR-IOV	Fixed VST mode with DPDK on SR-IOV	2.5	2.6.8
Performance	Occasionally, performance decreasing might occur in native systems.	2.0	2.5
Mismatch between PMD ConnectX® ports' MAC and the actual MAC.	Occasionally, the ConnectX® ports' MACs used by the PMD are not the actual ConnectX® MACs.	1.5	2.5