



Mellanox DPDK

Release Notes

Rev 16.11_4.0

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

© Copyright 2016. Mellanox Technologies LTD. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, CloudX logo, Connect-IB®, ConnectX®, CoolBox®, CORE-Direct®, EZchip®, EZchip logo, EZappliance®, EZdesign®, EZdriver®, EZsystem®, GPUDirect®, InfiniHost®, InfiniScale®, Kotura®, Kotura logo, Mellanox Federal Systems®, Mellanox Open Ethernet®, Mellanox ScalableHPC®, Mellanox Connect Accelerate Outperform logo, Mellanox Virtual Modular Switch®, MetroDX®, MetroX®, MLNX-OS®, NP-1c®, NP-2®, NP-3®, Open Ethernet logo, PhyX®, SwitchX®, Tiler®, Tiler logo, TestX®, The Generation of Open Ethernet logo, UFM®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo are registered trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

For the most updated list of Mellanox trademarks, visit <http://www.mellanox.com/page/trademarks>

Contents

Document Revision History	5
1 Overview	6
1.1 System Requirements	6
2 Changes and Major New Features in Rev 16.11_4.0	7
3 mlx4 and mlx5 PMD Drivers Features	8
4 Known Issues	9
5 Bug Fixes	13
6 Changes and New Features History	17

List of Tables

Table 1: Document Revision History	5
Table 2: System Requirements.....	6
Table 3: Changes and Major New Features.....	7
Table 4: Known Issues	9
Table 5: Bug Fixes	13
Table 6: Changes and New Features History.....	17

Document Revision History

Table 1: Document Revision History

Release	Date	Description
Rev 16.11_4.0	November 14, 2017	Initial release of this DPDK version.

1 Overview

These are the release notes for mlx5 DPDK Poll-Mode Driver (PMD), based on dpdk.org 16.11 release for Mellanox ConnectX®-4 onwards Ethernet adapters.



NOTE: This MLNX_DPDK release is limited to mlx5 driver and intended for pre-releasing new and experimental features that are to be integrated into the upstream dpdk.org.

For Mellanox official dpdk PMD releases, refer to the latest official GA releases in dpdk.org.

1.1 System Requirements

Table 2: System Requirements

Specification	Value
Network Adapter Cards	ConnectX®-4 / ConnectX®-4 Lx / ConnectX®-5 / ConnectX®-5 Ex network adapter card. (These must be configured to work in ETH mode.)
Firmware	<ul style="list-style-type: none"> • ConnectX-5/ConnectX-5 Ex: v16.21.1000 • ConnectX-4: v12.21.1000 • ConnectX-4 Lx: v14.21.1000
Linux Driver Stack	MLNX_OFED_LINUX-4.2-1.0.0.0
Tested Operating Systems and Kernels	<ul style="list-style-type: none"> • RHEL/CentOS7.2 • RHEL/CentOS7.3 • Ubuntu 16.10 • Ubuntu 14.04 <p>Note: For additional OSes, please contact Mellanox Support (support@mellanox.com).</p>
Minimum memory requirements	16 GB RAM
Transport	Ethernet
CPU Arch	x86

2 Changes and Major New Features in Rev 16.11_4.0

Table 3: Changes and Major New Features

Driver	Changes
ConnectX-4/5 PMD, mlx5	<ul style="list-style-type: none">• Added support for RSS on inner standard and non-standard VXLAN header• Added support for standard and non-standard VXLAN inner RX checksum offload• Added support for VF representors• Added single core Performance Optimizations• Bug fixes (see Bug Fixes)

3 mlx5 PMD Drivers Features

Feature	mlx5 PMD
Supported NICs	ConnectX®-4, ConnectX®-4 Lx ConnectX®-5, ConnectX®-5 Ex
PCI mapping	Function per port
KVM SR-IOV	Yes
Scattering/gathering RX/TX packets	Yes
Multiple RX (with RSS/RCA) and TX queues	Yes
IPv4/v6, TCP IPv4/v6, UDP IPv4/v6 RSS	Yes
Standard and non-standard VXLAN RSS on inner header	Yes
Number of RSS queues	Any
Get and Set RSS key per flow type (rss_hf)	Yes
Multiple MAC addresses	Yes
VLAN filtering	Yes
Link state information	Yes
Software counters/statistics	Yes
Start/stop/close operations	Yes
Multiple physical ports host adapter	Yes
Promiscuous mode	Yes
Multicast Promiscuous	Yes
TX and RX Checksums hardware offloading	Yes
Standard and non-standard VXLAN TX and RX Checksum hardware offloading	Yes
Flow Director	Yes
RX VLAN stripping	Yes
TX VLAN insertion	Yes
Port extended statistics	Yes
TSO	Yes
VXLAN TSO	Yes
VXLAN TSO with packets without Ethernet header	Yes
VF Representors	Yes

4 Known Issues

The following are DPDK Known Issues.

Table 4: Known Issues

Internal Ref.	Issues
-	<p>Description: VLAN insertion and TSO are not supported with MPW.</p> <p>Workaround: Disable MPW.</p> <p>Keywords: VLAN, TSO, MPW</p>
-	<p>Description: Sending data between vports using representors in ConnectX-4 Lx is currently not functional.</p> <p>Workaround: Disable MPW to enable the representors.</p> <p>Keywords: VF representors, ConnectX-4 Lx, MPW</p>
-	<p>Description: PMD does not recognize the packet as tunneled in case of a bad inner checksum.</p> <p>Workaround: N/A</p> <p>Keywords: Tunneled, inner checksum,</p>
-	<p>Description: Mellanox PMDs does not support InfiniBand.</p> <p>Workaround: N/A</p> <p>Keywords: InfiniBand</p>
-	<p>Description: Mellanox PMDs does not support DPDK integrated shared library.</p> <p>Workaround: Use the PMD compiled as a shared library.</p> <p>Keywords: DPDK integrated shared library</p>
-	<p>Description: Hardware queue counters are not implemented.</p> <p>Workaround: N/A</p> <p>Keywords: Hardware queue counters</p>
-	<p>Description: The Primary and Secondary multi process model is currently supported in TX only.</p> <p>Workaround: multi process model is supported in DPDK 17.11</p> <p>Keywords: The Primary and Secondary multi process model</p>
-	<p>Description: Bond PMD does not configure the needed MAC address to Bond slaves in case of failover. Active-Passive Bonding mode is available in case of KVM SR-IOV when 2 VMs are configured with the same MAC.</p> <p>Workaround: N/A</p> <p>Keywords: Active - Passive Bonding mode</p>
-	<p>Description: When the packet does not include TCP/UDP header and the hardware checksum offloading is enabled, L4 checksum is reported as bad.</p> <p>Workaround: N/A</p> <p>Keywords: L4 checksum report with hardware checksum offloading</p>
-	<p>Description: Promiscuous mode does not work when SR-IOV is enabled.</p>

Internal Ref.	Issues
	<p>Workaround: N/A</p> <p>Keywords: Promiscuous mode and SR-IOV</p>
-	<p>Description: In testpmd: Although promiscuous mode fails in SR-IOV, it is still shown as enabled This is DPDK implementation bug.</p> <p>Workaround: N/A</p> <p>Keywords: Promiscuous mode and SR-IOV</p>
-	<p>Description: Using OVS 2.5.0 requires several additional patches for it to function properly</p> <p>Workaround: Use OVS 2.6.0 branch instead. For further information, refer to the Mellanox Community.</p> <p>Keywords: OVS support</p>
-	<p>Description: Working with multiple containers where separate network namespaces are created is not supported</p> <p>Workaround: N/A</p> <p>Keywords: Containers</p>
-	<p>Description: Enabling Auto-inline feature might cause host to crush,</p> <p>Workaround: N/A</p> <p>Keywords: Auto-inline, inline</p>
-	<p>Description: When a Multicast or Broadcast packet is sent to the SR-IOV VF, it is returned back to the port.</p> <p>Workaround: N/A</p> <p>Keywords: Multicast / Broadcast Self-loopback on SR-IOV VF</p>
-	<p>Description: System hang during heavy traffic and stopping DPDK process</p> <p>Workaround: Disable ethernet driver steering rules with ethtool. example: ethtool -U eth2 flow-type ether dst E4:1D:2D:EA:CC:30 action -1 loc 0</p> <p>Keywords: System hang during heavy traffic</p>
-	<p>Description: Broadcast packets are not being received on the KVM VM.</p> <p>Workaround: N/A</p> <p>Keywords: Broadcast packets on KVM VM</p>
-	<p>Description: When using default firmware configuration, LCS events are not received when <code>rte_eth_dev_set_link_down/up</code> are called.</p> <p>Workaround:</p> <ol style="list-style-type: none"> Run: <pre>mlxconfig -d /dev/mst/mt4115_pciconf0 set KEEP_ETH_LINK_UP_P1=0 KEEP_ETH_LINK_UP_P2=0</pre> Reboot the server. <p>Keywords: LCS events</p>

Internal Ref.	Issues
-	<p>Description: Packets that have UDP destination port 4791 will be silently dropped by the NIC which may cause to low performance of 0, packet loss tests. This port is a reserved port and should not be used.</p> <p>Workaround: Use a different port as the UDP destination port</p> <p>Keywords: UDP port 4791</p>
-	<p>Description: Auto inline feature is supported only on a single port ConnectX-4 adapter card (is not supported on ConnectX-4 Lx) and cannot be used for 0 packet loss performance</p> <p>Workaround: Use auto-inline only with supported mode</p> <p>Keywords: Auto inline</p>
-	<p>Description: When destination MAC of a sent packet is the same as the ports MAC, the packet is returned to the port and is not sent to the network. Note: On ConnectX-4 Lx, the packet is sent when txq_mpw_en is set to 1 (the default configuration).</p> <p>Workaround: N/A</p> <p>Keywords: The dest MAC of the sent packet</p>
-	<p>Description: The maximum supported number of descriptors in TX/RX queues is equal to 32K. When txq_inline parameter is used the number is smaller and is equal to $32K / (txq_inline / 32 + 1)$. When auto-inline is used, the maximum supported number of descriptor is 6K If the number of descriptors of TX or RX queues is higher than the above maximum, segmentation fault will occur</p> <p>Workaround: Use maximum number of TX and RX descriptors as explained.</p> <p>Keywords: Maximum supported size of TX/RX queue</p>
-	<p>Description: Performance degradation of small messages might occur when MTU is higher than the mbuf size</p> <p>Workaround: N/A</p> <p>Keywords: Performance of small messages when MTU > mbuf size</p>
-	<p>Description: Adding MAC address on KVM VM is not supported</p> <p>Workaround: N/A</p> <p>Keywords: Adding MAC address on KVM VM</p>
-	<p>Description: DPDK has a flag per port to indicate the VLAN stripping state (On/Off), that updated when using rte_eth_dev_set_vlan_offload() API. This flag is not updated when calling the rte_eth_dev_set_vlan_strip_on_queue() API (VLAN stripping per queue), and can cause some confusion when mixing between those two APIs.</p> <p>Workaround: Use only one API to configure RX stripping</p> <p>Keywords: RX VLAN Stripping has unexpected behavior</p>
-	<p>Description: TX VLAN offloading insertion is supported only when txq_mpw_en is set to 0. By default is not supported since MPW is enabled</p> <p>Workaround: Set txq_mpw_en=0. Please check the Quick Start Guide for more information regarding command line options.</p>

Internal Ref.	Issues
	Keywords: TX VLAN insertion offloading with ConnectX-4 LX
-	Description: Changing MTU during traffic results in segmentation fault.
	Workaround: Stop the ports before changing the MTU
	Keywords: Changing MTU during traffic
-	Description: Flow Control settings are not allowed on VM
	Workaround: Set Flow Control settings on the Hypervisor
	Keywords: Flow Control settings on KVM VM
-	Description: Allmulticast on VM is not supported
	Workaround: N/A
	Keywords: Allmulticast on KVM VM
-	Description: When configuring the MTU on the VM make sure the MTU on the hypervisor is the same as on the VM
	Workaround: N/A
	Keywords: MTU configuration on KVM VM

5 Bug Fixes

Table 5: Bug Fixes

Internal Ref.	Issue
DPDK v16.11 Bug Fixes	
	Description: Fixed Tx doorbell memory barrier
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed a deadlock due to buffered slots in Rx SW ring
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed a TSO segment size verification
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed a Tx stats error counter logic
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed a Tx stats error counter definition
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed the number of segmentation assumption in SSE Tx
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed memory barrier definition
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed a trimming software ring for vectored Rx
	Keywords: mlx5
	Discovered in Release: DPDK 16.11
	Fixed in Release: MLNX_DPDK 16.11_4.0

Internal Ref.	Issue
	Description: Fixed an MTU update Keywords: mlx5 Discovered in Release: DPDK 16.11 Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed L4 packet type support Keywords: mlx5 Discovered in Release: DPDK 16.11 Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed a missing packet type calculation Keywords: mlx5 Discovered in Release: DPDK 16.11 Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed a counting consumed Tx descriptors Keywords: mlx5 Discovered in Release: DPDK 16.11 Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed tunnel offload detection Keywords: mlx5 Discovered in Release: DPDK 16.11 Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed the TSO inline size Keywords: mlx5 Discovered in Release: DPDK 16.11 Fixed in Release: MLNX_DPDK 16.11_4.0
	Description: Fixed the start pointer of a compressed completion Keywords: mlx5 Discovered in Release: DPDK 16.11 Fixed in Release: MLNX_DPDK 16.11_4.0
-	Description: Fixed a redundant free of Tx buffer Keywords: mlx5 Discovered in Release: DPDK 16.07 Fixed in Release: MLNX_DPDK 16.11_3.0
-	Description: Fixed exception handling Keywords: mlx5 Discovered in Release: DPDK 16.07 Fixed in Release: MLNX_DPDK 16.11_3.0

Internal Ref.	Issue
-	<p>Description: Fixed index handling for Tx ring</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11_2.3</p> <p>Fixed in Release: MLNX_DPDK 16.11_3.0</p>
-	<p>Description: Fixed Tx when first segment size is too short</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_3.0</p>
-	<p>Description: Fixed Completion buffer size</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.07</p> <p>Fixed in Release: MLNX_DPDK 16.11_3.0</p>
-	<p>Description: Fixed PMD specific parameters defaults.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11_2.3</p> <p>Fixed in Release: MLNX_DPDK 16.11_3.0</p>
-	<p>Description: Fixed updating RETA.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_2.3</p>
-	<p>Description: Fixed extended statistics counters identification.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_2.3</p>
-	<p>Description: Fixed VLAN stripping indication.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_2.3</p>
-	<p>Description: Fixed extended statistics wrong number.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_2.3</p>
-	<p>Description: Fixed inline WQE consumption.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_2.3</p>

Internal Ref.	Issue
-	<p>Description: Fixed Tx WQE corruption caused by starvation.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_2.3</p>
-	<p>Description: Fixed Ethernet header re-writing.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_2.3</p>
-	<p>Description: Fixed Rx packet validation and type.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_1.5</p>
-	<p>Description: Changed the default RSS algorithm of the NIC to TOPLITZ.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_1.5/ MLNX_OFED 4.0-1.0.0.1</p>
-	<p>Description: Fixed unexpected TX drops.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_1.5</p>
-	<p>Description: Fixed data corruption with MPW on ConnectX-4 Lx.</p> <p>Keywords: mlx5</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_1.5</p>
-	<p>Description: Fixed IPv4 and IPv6 packet type report.</p> <p>Keywords: mlx5/mlx4</p> <p>Discovered in Release: DPDK 16.11</p> <p>Fixed in Release: MLNX_DPDK 16.11_1.5</p>

6 Changes and New Features History

Table 6: Changes and New Features History

Driver	Changes
Rev. 16.11_3.0	
ConnectX-4/5 PMD, mlx5	<ul style="list-style-type: none"> • Added Delay Drop Rx queue support • Added HW TSO support for VXLAN packets without the inner Ethernet header. Supported only for PFs. • Added HW checksum offload for VXLAN packets without the inner Ethernet header. Supported only for PFs. • Performance Optimizations. • Bug fixes
Rev. 16.11_2.3	
ConnectX-4/5 PMD, mlx5	<ul style="list-style-type: none"> • Added HW TSO support for non-tunneled and tunneled packets • [Beta] Added HW TSO support for VXLAN packets without the inner Ethernet header. Such capability is supported only for PFs. • [Beta] Added HW checksum offload for VXLAN packets without the inner Ethernet header. Such capability is supported only for PFs. • Optimized Enhanced MPW for ConnectX@-5 • Added support for RETA table update • Added support for MPLS RSS (MLNX_OFED feature) • Added 'out-of-buffer' counter to extended statistics • Bug fixes
Rev. 16.11_1.5	
ConnectX-4/5 PMD, mlx5	<ul style="list-style-type: none"> • Added support for ConnectX@-5 NICs [at Beta level] • Added support for optimized MPW for ConnectX@-5 • Added support for MLNX_OFED 4.0-x.x.x.x • Vectorized logic on TX path to improve CPU utilization (backport from DPDK 17.02) • Improved RSS spreading in case of non-power of two number of queues • Added support for port extended statistics (backported from DPDK 17.02) • Added Auto-inline support for ConnectX-4 single port NIC (backported from MLNX_DPDK 2.2_4.8)