Mellanox DPDK

Release Notes

Rev 2.1_1.1

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.
Hakidma 26
Ofer Industrial Park
Yokneam 2069200
Israel
www.mellanox.com
Tel: +972 (0)74 723 7200
Fax: +972 (0)4 959 3245

Contents

List of Tables

1 Overview

These are the release notes for mlx4 and mlx5 DPDK Poll-Mode Driver (PMD) for Mellanox ConnectX®-3/ConnectX®-3 Pro and ConnectX®-4/ConnectX®-4 Lx Ethernet adapters.

1.1 System Requirements

Table 1: System Requirements

Specification	Value
Network Adapter Cards	Mellanox ConnectX®-3, ConnectX®-3 Pro, ConnectX®-4, ConnectX®-4 Lx network adapter card. (This must be configured to work in ETH mode.)
Firmware	• ConnectX4: 12.12.0780 • ConnectX-4 Lx: 14.12.0780 • ConnectX3/Pro: 2.35.5100
Driver Stack	MLNX_OFED-3.1-x.x.x
Operating Systems and Kernels	All OSs supported by MLNX_OFED
Minimum memory requirements	16 GB RAM
Transport	Ethernet
CPU Arch	x86

2   Changes and Major New Features

Table 2: Changes and Major New Features

Release	Driver	Changes
Rev 2.1_1.1 from DPDK 2.1 inbox	ConnectX-4 PMD, mlx5	- Added support for ConnectX®-4 and ConnectX®-4 LX NICs - Added support for MLNX_OFED 3.1-1.0.0 - DPDK fixes (see section Table 5: DPDK v2.1 Bug Fixes)
Rev 2.1_1.1 from MLNX DPDK 2.0_2.8.4	ConnectX-3 PMD, mlx4	- Removed support for VMware ESX 5.5 - Added support for Accelerated verbs – PMD does not include libibverbs and libmlx4

3   mlx4 and mlx5 PMD Drivers Features

Feature	mlx4 PMD	mlx5 PMD
Supported NICs	ConnectX®-3 ConnectX®-3 Pro	ConnectX®-4 ConnectX®-4 Lx
PCI mapping	address per adapter	address per port
KVM SR-IOV	Yes	No
Scattering/gathering RX/TX packets	Yes	Yes
Multiple RX (with RSS/RCA) and TX queues	Yes	Yes
IPv4, TCP IPv4, UDP IPv4 RSS	Yes	Yes
IPv6 RSS	Yes	No
VXLAN RSS	Yes	According to the outer header
Number of RSS queues	Power of 2	Any
Get and Set RSS key per flow type (rss_hf)	No	Yes
Multiple MAC addresses	Yes	Yes
VLAN filtering	Yes	Yes
Link state information	Yes	Yes
Software counters/statistics	Yes	Yes
Start/stop/close operations	Yes	Yes
Multiple physical ports host adapter	Yes	Yes
Promiscuous mode	Yes	Yes
Multicast Promiscuous	Yes	No
Checksums hardware offloading	Yes	No
Checksum VXLAN hardware offloading	Yes	No

4 Known Issues

4.1 General Known Issues

Table 3: General Known Issues

Subject	Description	Workaround
InfiniBand	Mellanox PMDs does not support InfiniBand.	N/A
Support for DPDK integrated shared library	Mellanox PMDs does not support DPDK integrated shared library	PMD compiled as shared library can be used
Hardware counters	Hardware counters are not implemented.	N/A
The Primary and Secondary multi process model	The Primary and Secondary multi process model is currently not supported	Use multithreaded model instead of multi process
Performance degradation when SGE_NUM =1	There is performance degradation for small packets when PMD is compiled with SGE_NUM = 4 compared to the performance when SGE_NUM = 1	If scattered packets are not used compile PMD with SGE_NUM = 1

4.2 mlx4 PMD Known Issues

Table 4: mlx4 PMD Known Issues

Subject	Description	Workaround
Multicast / Broadcast Self-loopback on SR-IOV VF	When a Multicast or Broadcast packet is sent to the SR-IOV VF, it is returned back to the port.	N/A
VXLAN TX checksum hardware offloading	The bad checksum of sent packets is not fixed by the hardware.	Use software checksum calculation
Hardware checksum offloading on ConnectX-3	Hardware checksum offloading is not supported in ConnectX-3	Use ConnectX3-Pro if hardware checksum offloading is needed
Sending abnormal packets	Sending abnormal packets, smaller then 16B and bigger than the configured MTU causes PMD queues to enter an error state.	run dev_stop and dev_start

Subject	Description	Workaround
Segmentation fault when killing application	On some operating systems, when killing DPDK application with several RSS queues it can generate segmentation fault	Close application without killing it
VXLAN hardware checksum offloading	VXLAN hardware checksum offloading is supported only if the used kernel supports it	N/A
L4 checksum report with hardware checksum offloading	When the packet does not include TCP/UDP header and the hardware checksum offloading is enabled, L4 checksum will be reported as bad	N/A
RSS	RSS hash key and options cannot be modified.	N/A
Promiscuous and SR-IOV	Promiscuous mode does not work when SR-IOV is enabled.	N/A
Promiscuous and SR-IOV	In testpmd: Although promiscuous mode fails in SR-IOV, it still shows as enabled This is DPDK implementation bug.	N/A
VLAN filtering	VLAN filtering is supported only with non-optimized steering mode.	See QSG, RX VLAN Filter section
Number of configured RSS queues must be power of 2	Number of configured RSS queues must be power of 2.	Use only power of 2 RSS queues
MLNX OFED 2.4	MLNX_OFED 2.4 and below is not supported	It is recommended to use MLNX_OFED 3.1

4.3 mlx5 PMD Known Issue

Table 4: mlx5 PMD Known Issues

Subject	Description	Workaround
Port hang	Occasionally, when ConnectX-4/ ConnectX-4 Lx ports are connected to a non-Mellanox NIC and CONFIG_RTE_LIBRTE_MLX5_SGE_WR_N >1, port stops receiving traffic from remote port	Configure CONFIG_RTE_LIBRTE_MLX5_SGE_WR_N = 1
KVM and VMware SR-IOV	SR-IOV is not supported	N/A
Checksums HW offloading	Checksums HW offloading is not supported	N/A
IPv6 RSS	IPv6 traffic is not spread between RSS queues	N/A
VXLAN RSS	VXLAN traffic is spread between RSS queues according to the outer header	

Subject	Description	Workaround
Multicast Promiscuous	Multicast Promiscuous is not supported	Use Promiscuous instead
Checksums hardware offloading	Checksum hardware offloading is not supported	N/A
Checksum VXLAN hardware offloading	Checksum VXLAN hardware offloading is not supported	N/A
VLAN filter and promiscuous	When promiscuous mode is enabled, the VLAN filter does not stop traffic from every VLAN to be received	When working with VLAN filter, disable promiscuous mode

Mellanox Technologies

5   Bug Fixes

Table 5: ConnectX-3 PMD Bug Fixes

Subject	Description	Found in	Fixed in
Multicast Promiscuous mode and SR-IOV	Fixed Virtual Machines (VM) with enabled Multicast Promiscuous mode do not receive the multicast traffic that other VMs on the same hypervisor are registered to.	MLNX_OFED 2.4	MLNX_OFED 3.0 and above
QP setup failure	Fixed QP setup failure on Debian 3.16.7-ckt7-1	DPDK 2.1 inbox	2.1_1.1
Memory corruption when fork() is used in the application	Fixed an unpredicted behavior due to "cow" mechanism when an application used fork() or system call that were not related to the networking.	DPDK 2.0 inbox	2.0_2.8.4
1Gb/s port's link reported as 10Gb/s	Fixed a wrong speed reporting when a 10Gb/s port was set as 1Gb/s (10Gb/s instead of 1Gb/s)	DPDK 2.0 inbox	2.0_2.8.4
Multiple RX VLAN filters	Fixed a multiple RX VLAN filters issue preventing multiple RX VLAN filters to function properly although all were configured. In this case, only the first one functioned properly.	DPDK 2.0 inbox	2.0_2.8.4

Table 6: DPDK v2.1 Bug Fixes

Subject	Description	Found in	Fixed in
testpmd: Added a missing type to RSS hash commands	DPDK uses a structure to get or set a new hash key (see eth_rte_rss_hash_conf). rss_hf field is used in rss_hash_get_conf to retrieve the hash key and in rss_hash_update to verify the key existence before trying to update it.	DPDK 2.1 inbox	2.1_1.1
testpmd: Fixed a missing initialization in the RSS hash show command	The "show port X rss-hash" command sometimes displays junk instead of the expected RSS hash key because the maximum key length is undefined. When the requested key is too large to fit in the buffer, rte_eth_dev_rss_hash_conf_get() does not update it.	DPDK 2.1 inbox	2.1_1.1
eal: Added support for 100 Gb/s link rate (and above)	uint16_t is not large enough to store a 100 Gb/s link rate. The API was updated to use uint32_t instead.	DPDK 2.1 inbox	2.1_1.1