Red Hat Enterprise Linux (RHEL) 7.5 Driver Release Notes

RHEL 7.5

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

Table of Contents

List Of Tables

1　Overview

These are the release notes of Red Hat Enterprise Linux (RHEL) 7.5 Driver Release Notes. This document provides instructions on drivers for Mellanox Technologies ConnectX® based adapter cards with Red Hat Enterprise Linux (RHEL) 7.5 Inbox Driver environment.

This version supports the following uplinks to servers:

Table 1 - Supported Uplinks to Servers

Uplink/HCAs	Uplink Speed	Supported Driver
ConnectX®-5	• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 56GigE[a], and 100GigE	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX®-4	• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 56GigE[a], and 100GigE	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX®-4 Lx	• Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, and 50GigE	mlx5_core (includes the ETH functionality as well)
Connect-IB®	• InfiniBand: SDR, QDR, FDR10, FDR	mlx5_core, mlx5_ib
ConnectX®-3/ ConnectX®-3 Pro	• InfiniBand: SDR, QDR, FDR10, FDR • Ethernet: 10GigE, 40GigE and 56GigE[a]	mlx4_core, mlx4_en, mlx4_ib
ConnectX®-2	• InfiniBand: SDR, DDR • Ethernet: 10GigE, 20GigE	mlx4_core, mlx4_en, mlx4_ib

a. 56GbE is a Mellanox propriety link speed and can be achieved while connecting a Mellanox adapter cards to Mellanox SX10XX switch series or connecting a Mellanox adapter card to another Mellanox adapter card.

1.1　Supported HCAs Firmware Versions

Red Hat Enterprise Linux (RHEL) 7.5 Driver supports the following Mellanox network adapter cards firmware versions:

Table 2 - Supported HCAs Firmware Versions

HCA	Recommended Firmware Rev.	Additional Firmware Rev. Supported
Connect-IB®	10.16.1200	N/A
ConnectX®-5	16.22.1002	N/A
ConnectX®-4 Lx	14.22.1002	N/A
ConnectX®-4	12.22.1002	N/A
ConnectX®-3 Pro	2.42.5000	N/A
ConnectX®-3	2.42.5000	N/A
ConnectX®-2	2.9.1000	N/A

1.2 SR-IOV Support

Table 3 - SR-IOV Support

Driver	Support
mlx4_core, mlx4_en, mlx4_ib	ETH Infiniband - Technical Preview[a]
mlx5_core (includes ETH functionality), mlx5_ib	ETH Infiniband - Technical Preview[a]

a. Technical Preview is not fully supported production feature.

1.3 RoCE Support

Table 4 - RoCE Support

Driver	Support
mlx4 - RoCE v1/v2	Yes
mlx5 - RoCE v1/v2	Yes

1.4 VXLAN Support

Table 5 - VXLAN Support

Driver	Support
mlx4 - VXLAN offload	Yes
mlx5 - VXLAN offload	Yes (without RSS)

1.5 DPDK Support

Table 6 - DPDK Support

Driver	Support
mlx4	Yes
mlx5	Yes

1.6 Open vSwitch Hardware Offloads Support

Table 7 - Open vSwitch Hardware Offloads Support

Driver	Support
mlx4	No
mlx5	Yes

2   Changes and New Features

Table 8 - Changes and New Features

Driver	Feature/Change	Description
mlx5	On Demand Paging	Added On-Demand-Paging (ODP), a technique to alleviate much of the shortcomings of memory registration.
	Support 4K User Access Regions (UAR)	Allow a more efficient use of device memory mapped I/O as UAR (User Access Region). Each process requires an area from the device memory in order to trigger operations on the device such as sending a message to the wire. This area is taken from the memory covered by BAR0-1 of the device. This kind of memory is called UAR. This feature will make more efficient use of this area allowing more processes to run concurrently. The effect is especially significant on PPC machines since they use 64KB for a system page.
	SR-IOV Bandwidth Share for Ethernet/RoCE	Added the ability to guarantee the minimum rate of a certain VF in SR-IOV mode.
	RoCE Diagnostics and ECN Counters	Added support for additional RoCE diagnostics and ECN congestion counters under /sys/class/infiniband/mlx5_0/ports/1/hw_counters/ directory.
	Local Loopback	Improved performance by rendering Local loopback (unicast and multicast) disabled by mlx5 driver by default while local loopback is not in use. The mlx5 driver keeps track of the number of transport domains that are opened by user-space applications. If there is more than one user-space transport domain open, local loopback will automatically be enabled.
	Enhanced IPoIB	Added support for Enhanced IPoIB feature, which enables better utilization of features supported in ConnectX-4 adapter cards, by optimizing IPoIB data path and thus, reaching peak performance in both bandwidth and latency. Enhanced IPoIB is enabled by default

Table 8 - Changes and New Features

Driver	Feature/Change	Description
mlx5	Precision Time Protocol (PTP)	Added support for PTP feature in IPoIB offloaded devices. This feature allows for accurate synchronization between the distributed entities over the network. The synchronization is based on symmetric Round Trip Time (RTT) between the master and slave devices. The feature is enabled by default.
	Tag Matching Offload	Added support for hardware Tag Matching offload with Dynamically Connected Transport (DCT).
	GRE Tunnel Offloads	Added support for the following GRE tunnel offloads: • TSO over GRE tunnels • Checksum offloads over GRE tunnels • RSS spread for GRE packets
	Dropless Receive Queue (RQ)	Added support for the driver to notify the FW when SW receive queues are overloaded.
	OVS enhancements	• Ipv6 support for VXLAN offloads (encap/decap) • Encapsulation flow enhancement • Expose min-inline mode to libmlx5 • mlx5 offloading of TC pedit (header re-write) action • mlx5 neigh update • Add support for drop action for steering rules • Header re-write • Flow counters
mlx4	RSS Support	RSS (Receive Side Scaling) technology allows to spread incoming traffic between different receive descriptor queues. Assigning each queue to different CPU cores allows to better load balance the incoming traffic and improve performance.
DPDK	Mellanox PMDs enabled by default	Mellanox PMDs (Poll Mode Driver) are open source upstream drivers, embedded within DPDK releases which are integrated with major Linux distribution releases. Mellanox PMDs enable advanced ConnectX NIC hardware offload capabilities freeing up the server CPU cores for application needs.
General	RHEL NIC Qualification	Successfully passed RHEL NIC Qualification as described in https://github.com/ctrautma/RHEL_NIC_QUALIFICATION/ Covering: • ConnextX-3, ConnextX-4, ConnectX-5 adapter cards • Kernel Data Path, OVS, OVS-DPDK • PVP (Physical-Virtual-Physical), Performance (VSPerf), OVS functionality.

3   Known Issues

The following table describes known issues in this release and possible workarounds..

Table 9 - Known Issues (Sheet 1 of 2)

Internal Ref.	RedHat Ref.	Description
1646807	1615865	**Description:** Due to a missing kernel support, creating 127 VFs on a ConnectX-5 adapter card fails. The failure may occur even with lower number of VFs (higher then 60).
		Workaround: N/A
		Keywords: ConnectX-5, VFs
1263525	1533918	**Description:** mlx5 module fails to load while using kernel debug.
		Workaround: N/A
		Keywords: Debug
1284047	-	**Description:** Bandwidth degradations might occur due to PTI (Page Table Isolation) Intel's CPU security fix.
		Workaround: Perform one of the following: • Disable PTI during run time by writing 0 to `/sys/kernel/debug/x86/pti_enabled`. • Another option is adding "nopti" or "pti=off" to grub.conf.
		Keywords: Performance
1336618	-	**Description:** On rare occasions, under heavy traffic and loading/unloading the mlx4_en, mlx4_ib and ml4_core drivers may cause VPD access failure.
		Workaround: N/A
		Keywords: mlx4, Firmware
-	1189428	**Description:** kdump over mlx5 driver is not supported. However, it is supported on Ethernet interfaces but it requires preserving enough memory to support it.
		Workaround: As kdump is supported on Ethernet interfaces, make sure to preserve enough memory by adding `crashkernel=512M` to the kernel command line.
		Keywords: kdump, mlx5
-	-	**Description:** Atomic Operations in Connect-IB® are limited on little-endian machines (e.g. x86).
		Workaround: N/A
		Keywords: Atomic Operations in Connect-IB®

Table 9 - Known Issues (Sheet 2 of 2)

Internal Ref.	RedHat Ref.	Description
-	1450329	**Description:** Occasionally, under special circumstances, Virtual Function can be removed.
		Workaround: N/A
		Keywords: mlx5, SR-IOV
-	1460579	**Description:** Bond creation on mlx4 Virtual Function interfaces might fail with error -524.
		Workaround: N/A
		Keywords: mlx4, SR-IOV
-	1462591	**Description:** InfiniBand mlx5 SR-IOV is not supported using OpenSM.
		Workaround: Use the SM on the Mellanox Switch/UFM/mlnx_opensm
		Keywords: mlx5, SR-IOV, OpenSM