

Red Hat Enterprise Linux (RHEL) 8.1 Driver Release Notes

RHEL 8.1

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 2020. Mellanox Technologies Ltd. All Rights Reserved.

Mellanox®, Mellanox logo, ASAP2 - Accelerated Switch and Packet Processing®, BlueField®, BlueOS®, CloudX logo, Connect-IB®, ConnectX®, CORE-Direct®, GPUDirect®, HPC-X®, LinkX®, Mellanox CloudX®, Mellanox HostDirect®, Mellanox Multi-Host®, Mellanox NEO®, Mellanox NVMeDirect®, Mellanox OpenCloud®, Mellanox OpenHPC®, Mellanox PeerDirect®, Mellanox ScalableHPC®, Mellanox Socket Direct®, PeerDirect ASYNC®, SocketXtreme®, StoreX®, UCX®, UCX Unified Communication X®, UFM®, Unbreakable-Link®, and Virtual Protocol Interconnect® are registered trademarks of Mellanox Technologies, Ltd.

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

All other trademarks are property of their respective owners.

Table of Contents

Table of Contents	3
List Of Tables	4
Chapter 1 Overview	5
1.1 Supported HCAs Firmware Versions	6
1.2 SR-IOV Support	6
1.3 RoCE Support	6
1.4 VXLAN Support	7
1.5 DPDK Support	7
1.6 Open vSwitch Hardware Offloads Support	7
Chapter 2 Changes and New Features	8
Chapter 3 Certifications	10
3.1 RHEL NIC Qualification	10
Chapter 4 Known Issues	11

List Of Tables

Table 1:	Supported Uplinks to Servers	5
Table 2:	Supported HCAs Firmware Versions	6
Table 3:	SR-IOV Support.	6
Table 4:	RoCE Support	6
Table 5:	VXLAN Support.	7
Table 6:	DPDK Support.	7
Table 7:	Open vSwitch Hardware Offloads Support	7
Table 8:	Changes and New Features.	8
Table 9:	Known Issues	11

1 Overview

These are the release notes of Red Hat Enterprise Linux (RHEL) 8.1 Driver Release Notes. This document provides instructions on drivers for Mellanox Technologies ConnectX® based adapter cards with Red Hat Enterprise Linux (RHEL) 8.1 Inbox Driver environment.

This version supports the uplinks to servers described in the table below.

Table 1 - Supported Uplinks to Servers

Uplink/HCAs	Uplink Speed	Supported Driver
ConnectX®-6	<ul style="list-style-type: none"> InfiniBand: SDR, EDR, HDR Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 100GigE 	mlx5_core (includes the ETH functionality as well), mlx5_ib
BlueField® ^a	<ul style="list-style-type: none"> Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, and 100GigE 	mlx5_core (includes the ETH functionality as well)
Innova™ IPsec EN	<ul style="list-style-type: none"> Ethernet: 10GigE, 40GigE 	mlx5_core (includes the ETH functionality as well)
ConnectX®-5	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR, FDR10, EDR Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 56GigE^b, and 100GigE 	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX®-4	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR, FDR10, EDR Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 56GigE^b, and 100GigE 	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX®-4 Lx	<ul style="list-style-type: none"> Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, and 50GigE 	mlx5_core (includes the ETH functionality as well)
ConnectX®-3/ ConnectX®-3 Pro	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR10, FDR Ethernet: 10GigE, 40GigE and 56GigE^b 	mlx4_core, mlx4_en, mlx4_ib
Connect-IB®	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR10, FDR 	mlx5_core, mlx5_ib

a. BlueField is supported as a standard ConnectX-5 Ethernet NIC only.

b. 56GbE is a Mellanox proprietary 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) 8.1 driver supports the following Mellanox network adapter cards firmware versions:

Table 2 - Supported HCAs Firmware Versions

HCA	Recommended Firmware Rev.
ConnectX®-6	20.25.2006
BlueField®	18.25.1600
ConnectX®-5	16.25.1020
ConnectX®-4 Lx	14.25.1020
ConnectX®-4	12.25.1020
ConnectX®-3 Pro	2.42.5000
ConnectX®-3	2.42.5000
Connect-IB®	10.16.1200

1.2 SR-IOV Support

Table 3 - SR-IOV Support

Driver	Support	Notes
mlx4_core, mlx4_en, mlx4_ib	ETH Infiniband - Technical Preview	Running InfiniBand (IB) SR-IOV requires IB Virtualization support on the OpenSM (Session Manager).
mlx5_core (includes ETH functionality), mlx5_ib	ETH Infiniband - Technical Preview	This capability is supported only on OpenSM provided by Mellanox, that is not available in Inbox. This support can be achieved by running the highest-priority OpenSM on a Mellanox switch in an IB fabric. The switch SM can support this feature by enabling the virt flag (<code># ib sm virt enable</code>). Note: This capability is not tested over Inbox environment and considered Tech Preview.

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 ^a

a. Technical Preview is not a fully supported production feature.

2 Changes and New Features

Table 8 - Changes and New Features

Driver/ Component	Feature/Change	Description
General	ConnectX-5 Socket Direct Adapter Support	Added support for ConnectX-5 Socket Direct and hardware certification.
mlx5	DevX Interface	Added support for DevX interface which enables direct access to the mlx5 driver from the user space using KABI mechanism. This minimizes the dependency between user-space and kernel, such that user-space functionality can be added with minimal to no kernel changes.
	RoCE Invariant Cyclic Redundancy Code (ICRC) Counter Exposure	Added the ability to query for the counter that counts RoCE packets with corrupted ICRC.
	Tunnel Protocol (MPLSoGRE, MPLSoUDP, VxLan) Rx/Tx Decap/Encap Offload	Added the capability to create flow actions that can change packet headers.
	IB Device Name	Added the ability to rename the user's visible IB device name from a vendor-specific name (e.g. mlx5_0) to any other name.
	ODP Pre-fetch	Added support for pre-fetching a range of on-demand-paging (ODP) memory region (MR). This reduces latency by making pages with RO/RW permissions available before the actual IO is conducted.
	HDR Link Speed Exposure	Added support for HDR link speed over ConnectX-6 and above adapter cards.
	Double VLAN Pop/ Push (QinQ) Action	Added support for offloading match on both outer and inner VLAN tags for QinQ.
	VF LAG	Added support for High Availability and load balancing for Virtual Functions of different physical ports in SwitchDev SR-IOV mode.
	Equal Cost Multi-Path (ECMP)	Added support for offloading ECMP rules by tracking software multipath route and related next-hops, and reflecting this as port affinity to the hardware.
	Remote Mirroring	Added support for additional mirroring output in SwitchDev mode. The mirroring port may either be a local or a remote VF, using VxLAN or GRE encapsulations.
	GRE HW Offloading	Added the ability to offload TC filters set on GRE interfaces to NICs that support ASAP ² .
VxLAN/GRE Tunneling over VLAN	Added support for VxLAN and GRE tunnel encap/decap offload over Ethernet-tagged packets.	

Table 8 - Changes and New Features

Driver/ Component	Feature/Change	Description
	Uplink Representors	Exposed PF (uplink) representors in SwitchDev mode, similarly to VF representors, as an infrastructure improvement for SmartNICs.
	General mlx5 Driver Update	Updated the driver's base Upstream kernel version to v5.0.
mlx4	Configuration Parameters Setting through Devlink	Added the ability to set the configuration parameters through Devlink. Each device registers supported configuration parameters table.
	General mlx4 Driver Update	Updated the driver's base Upstream kernel version to v5.0.
rdma-core	Version Update	Updated rdma-core version to v22.3-1.el8.
mstflint		Updated mstflint version to v4.12.0-1.el8.
VMA		Updated VMA version to v8.7.7-1.el8.

3 Certifications

3.1 RHEL NIC Qualification

RHEL 8.0, Successfully passed RHEL NIC qualification has passed successfully as described in:
https://github.com/ctrautma/RHEL_NIC_QUALIFICATION/tree/8.0-Beta

Covering:

- ConnectX-4 Lx and ConnectX-5 adapter cards
- OVS functional, OVS non-offload, OVS-offload, OVS-DPDK

4 Known Issues

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

Table 9 - Known Issues

Internal Ref.	Description
1284047	Description: BW degradations due to PTI (Page Table Isolation) in Intel's CPU security fix
	Workaround: PTI can be disabled in run time by writing 0 to <code>/sys/kernel/debug/x86/pti_enabled</code> . Another option is adding "nopti" or "pti=off" to <code>grub.conf</code> .
	Keywords: Performance
1610281	Description: Setting speed to 56Gb/s on ConnectX-4 causes FW syndrome (0x1a303e)
	Workaround: N/A
	Keywords: ConnectX-4, syndrome
1609804	Description: Kernel panic during MTU change under stress traffic
	Workaround: N/A
	Keywords: panic, MTU
1578022	Description: OVS offload: fragmented traffic is not offload. When sending traffic with packets bigger than MTU, traffic runs but is not offloaded.
	Workaround: N/A
	Keywords: OVS offload, fragmentation